Journal of Computer-Aided Molecular Design, 8 (1994) 211–220
ESCOM

J-CAMD 224

A 3D QSAR approach to the search for geometrical similarity in a series of nonpeptide angiotensin II receptor antagonists

Laura Belvisi[a], Gianpaolo Bravi[a], Carlo Scolastico[a,*], Anna Vulpetti[a], Aldo Salimbeni[b] and Roberto Todeschini[c]

[a]*Dipartimento di Chimica Organica e Industriale Centro del C.N.R., via Venezian 21, I-20133 Milan, Italy*
[b]*Istituto LusoFarmaco d'Italia S.p.a., via Carnia 26, I-20132 Milan, Italy*
[c]*Dipartimento di Chimica Fisica e Elettrochimica, via Golgi 19, I-20133 Milan, Italy*

Received 15 June 1993
Accepted 23 July 1993

Key words: Binding affinity; Structure–activity relationship; Conformational analysis; Chemometric methods; Indirect drug design

SUMMARY

A 3D QSAR methodology based on the combined use of conformational analysis and chemometrics was applied to perform a comparative analysis of the 3D conformational features of 13 nonpeptide angiotensin II receptor antagonists showing different levels of binding affinity. Conformational analysis by using a molecular mechanics MM2 method was carried out for each of these structures to obtain conformational minima. These minima were described by ten interatomic distances which define the relative spatial disposition of five significant atoms belonging to relevant functional groups present in all the 13 molecules. The structure–activity relationship between the interatomic distances and the biological activity was then assessed by using chemometric methods (cluster analysis, principal component analysis, classification methods). With our indirect approach based on the search for geometrical similarity it was possible, even though structural information on the receptor active site was lacking, to identify the 3D geometrical requirements for the binding affinity of nonpeptide angiotensin II receptor inhibitors.

INTRODUCTION

Angiotensin II (AII), a linear octapeptide of sequence Asp-Arg-Val-Tyr-Ile-His-Pro-Phe produced by the renin-angiotensin system (RAS), is a powerful endogenous vasopressor. Drugs that inhibit the RAS have been shown [1] to be effective in treating human hypertension. A possible approach to interfering with the RAS is to inhibit the binding of AII to its receptor. In particular, the most recent pharmacological research is directed towards nonpeptide AII receptor antagonists [2–8] that, unlike peptide AII receptor inhibitors, can exhibit oral activity, have a long

*To whom correspondence should be addressed.

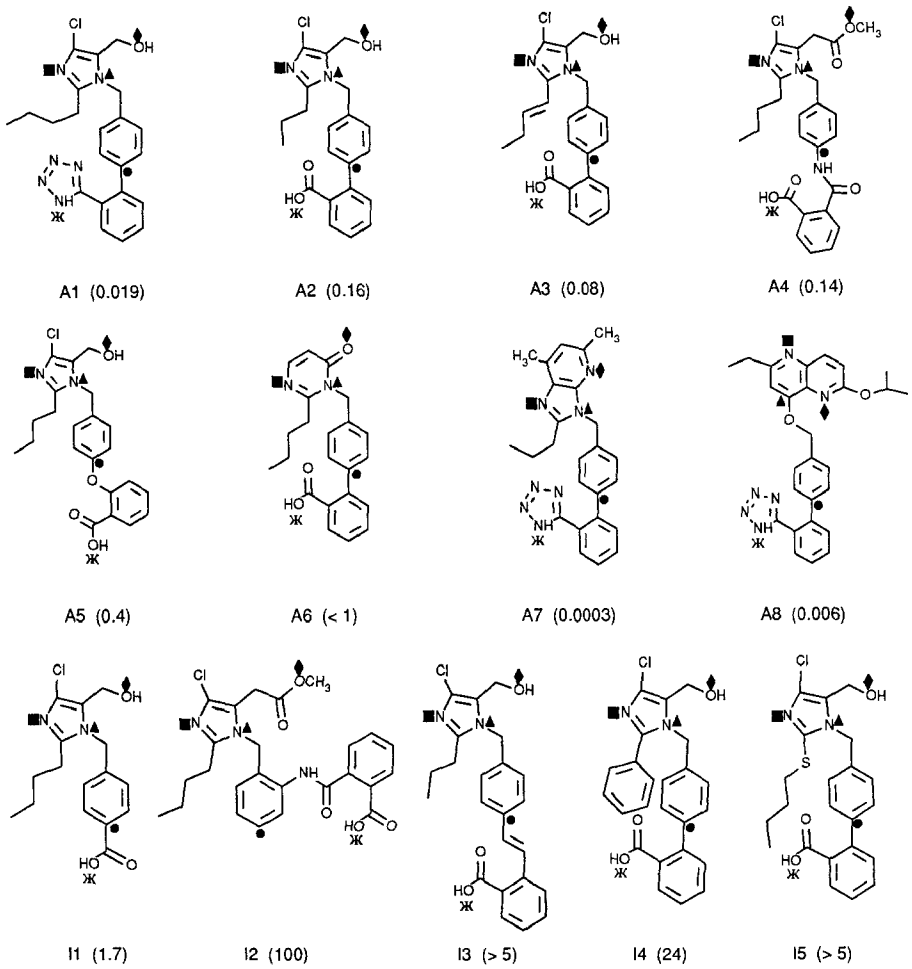

Fig. 1. Structures and binding affinities (IC_{50}, μM) of the 13 training set molecules. The labelled atoms (■, ▲, ♦, ●, ж) are considered in the definition of conformational descriptors.

plasma half-life, and do not exert partial agonism.

We have performed a 3D QSAR study and a computer graphics analysis on nonpeptide AII receptor inhibitors discovered by Du Pont [2–5]. As no direct structural information was available on the receptor active site during the research, an indirect approach was adopted, the aim being to identify the 3D geometrical requirements necessary for activity [9]. A comparative analysis of the molecular electrostatic potential (MEP) distributions of several Du Pont inhibitors was also performed to gain insight into the electrostatic requirements for AII antagonist activity [10]. On the basis of these previous studies, the extent to which Du Pont antagonists bind at the AII receptor seems to be related to both the relative spatial disposition of some relevant functionalities [9] and the electrostatic potential characteristics [10].

Continuing our theoretical studies on this promising pharmacological class, we have considered a more heterogeneous training set so as to take into consideration other structural requirements important for activity and inactivity. This training set also includes compounds which do not have a Du Pont-like structure. The training set used in the previous geometrical approach

(constituted by the Du Pont antagonists A1–A5 and I1–I4 of Fig. 1) was enlarged by three compounds (A6, A7 and A8) presenting a different heterocyclic ring with respect to the imidazole of the Du Pont antagonists, and by another Du Pont-like structure (I5), which presents a thiobutyl side chain at the imidazole 2-position. The structures and activity data (IC_{50} values) [2,6,7,11] of the molecules constituting the new extended training set are reported in Fig. 1.

The 13 selected AII antagonists are characterized by different structures and different levels of binding affinity for the same AII receptor: compounds showing IC_{50} values < 1 μM were considered active, whereas those possessing IC_{50} values > 1 μM were considered inactive.

A 3D QSAR methodology developed in our laboratories [9,12,13], based on the combined use of molecular modelling and chemometrics, was also applied in the present study in order to perform a comparative analysis of the 3D conformational features of these 13 AII antagonists, and to construct an improved, more generally valid, geometrical activity model for the nonpeptide antagonist–AII receptor interactions. The 3D features of each molecule were described by using the interatomic distances between five atoms belonging to functional groups that were present in all the training set molecules and which were, in a first analysis, considered responsible for the most significant interactions with the receptor. With respect to the previous study we have neglected one description point, the heteroatom electron withdrawal at the imidazole 4-position of the Du Pont inhibitors. In fact, this substituent is not present in the new molecules A6, A7 and A8; moreover, the presence of an electron-withdrawing substituent at the imidazole 4-position is not relevant for the biological activity of the Du Pont antagonists [3–5]. The five atoms chosen were the atom (B ▲), which directly binds the aromatic residue, and the −N= type nitrogen atom (N ■), both being present in the same heterocyclic ring; the heteroatom hydrogen-bond acceptor (A ◆), placed sideways to the heterocycle; the aromatic carbon atom (C ●), para to the group joining the heterocycle to phenyl; and the heteroatom of the acidic group (E ж) on the aromatic system (NH of the tetrazole for A1, A7 and A8, OH of the carboxyl group for all the other compounds) [14]. The relative spatial disposition of these five atoms was defined by all the ten possible interatomic distances (D1 = N–A, D2 = N–B, D3 = N–C, D4 = N–E, D5 = B–A, D6 = B–C, D7 = B–E, D8 = A–C, D9 = A–E, D10 = C–E). The five atoms considered in each compound are labelled in Fig. 1.

The values of these 3D molecular descriptors were determined by carrying out a conformational analysis for each molecule of the training set. Chemometric techniques were then used to search for geometrical similarities and dissimilarities within the thousands of conformations derived from the conformational analysis, that is to say, to recognize the relative spatial dispositions of selected common functionalities assumed by only the AII antagonists showing the highest levels of receptor affinity. Such dispositions were considered to characterize the binding conformations, i.e. the conformations in which the inhibitors interact with the receptor site, and which therefore constitute the 3D geometrical requirements for receptor binding.

COMPUTATIONAL METHODS

Conformational analysis

Molecular mechanics calculations 'in vacuo' were performed with MacroModel Version 3.1 X and Batchmin Version 3.1 [15] on a Silicon Graphics Workstation. The usage-directed Monte

Carlo conformational search [16], the MM2 force field [17] and the TNCG energy minimization algorithm [18], implemented in these versions of this integrated software system, were used to find conformational minima of the training set molecules.

Chemometric elaboration

Cluster and principal component analyses were performed with the 6.06.01 Release SAS system software [19] running on an IBM 3090. We also used the recently developed classification tree method LDCT (Linear Discriminant Classification Tree) on a 386 PC [20].

RESULTS AND DISCUSSION

Conformational analysis

Conformational analysis was carried out performing molecular mechanics MM2 calculations, randomly varying the torsional space for each molecule of the training set. Minimum-energy conformations were determined, the hypothesis being that the conformation which interacts with the receptor is among them. For each molecule of the training set, all minimum-energy conformations within 8 kcal/mol of the global minimum were retained; the number of minima within this energy window and the total steric energy of the global minimum for each compound are reported in Table 1. Nevertheless, due to the high flexibility of the compounds, the total number of conformers (about 9000) remained too high to manage. Moreover, each conformer was described by ten interatomic distances. Therefore, the correlation between the conformational information and the biological activity or, in other words, the construction of a quantitative structure–activity relationship between interatomic distances and receptor-binding affinities, appeared to be a rather complex problem. For its solution, we adopted a reasoned sequence of chemometric techniques.

Statistic and chemometric methods

First of all, a cluster analysis [21] was performed for each molecule on the selected conformers within 8 kcal/mol of the global minimum, the aim being to assemble in clusters similar conformations on the basis of the ten interatomic distance values. An agglomerative hierarchical clustering procedure (unweighted linkage average method) which uses the Euclidean distance as similarity

TABLE 1
TOTAL STERIC ENERGY OF THE GLOBAL MINIMUM (SE, KCAL/MOL), TOTAL NUMBER OF MINIMUM-ENERGY CONFORMATIONS (N) WITHIN 8 KCAL/MOL OF THE GLOBAL MINIMUM AND TOTAL NUMBER OF CLUSTERS (n) OBTAINED BY CLUSTER ANALYSIS FOR COMPOUNDS A1–A8 AND I1–I5

Compd	SE	N	n	Compd	SE	N	n
A1	5.98	771	60	A8	21.55	763	61
A2	−2.11	732	61	I1	−5.97	281	30
A3	−1.23	511	48	I2	−8.58	1161	85
A4	−6.69	785	87	I3	2.46	870	59
A5	−6.50	821	58	I4	2.23	165	28
A6	−30.04	545	49	I5	−7.00	1625	83
A7	10.06	190	25				

measure was performed on autoscaled variables [21]. For each molecule we considered a similarity level so as to obtain a number of clusters comparable with about 10% (mean value) of the molecular conformations, verifying at the same time that the variance within each cluster was not too high. Seven hundred and thirty-four clusters were selected for a total of 13 training set molecules (the number of clusters obtained for each compound is reported in Table 1). The centroid of each cluster was then calculated; it corresponds to the ten-dimensional vector of the variables' means, i.e. it represents the average spatial disposition of the conformers forming the cluster. Therefore, in order to reduce the complexity of the data system, we substituted the (approximately) 9000 molecular conformers by the 734 centroids of the clusters, the hypothesis being that these average spatial dispositions could fully represent the complete conformational behaviour of the minima.

Subsequently, an analysis of the descriptors of the system was performed on the reduced dataset (734 centroids and ten variables). A preliminary analysis of the correlation matrix revealed D4 = N–E, D7 = B–E ($\rho = 0.938$), D1 = N–A and D5 = B–A ($\rho = 0.804$) as the most correlated pairs of distances. Multivariate analysis of the descriptors was then carried out by performing a principal component analysis (PCA) [22] on the 734 centroids of the training set molecules to verify whether all ten molecular descriptors were required for a complete description of the system (autoscaled data were used). In fact, PCA is a powerful tool to detect the interdependence of variables and to obtain uncorrelated information. The results of the PCA showed that the first four principal components (PCs) explain about 83% of the total variance of the data system. It was therefore assumed that these four PCs would be sufficient to correctly describe the system and they were then used to analyse the relationships existing among the molecular descriptors. The cumulative per cent variance and the loading values of the first four PCs are reported in Table 2. Analysis of the coefficients of the original variables in the relevant four PCs allowed the evaluation of each variable's information and the recognition of the strongly correlated distances. D1 and D5 from one side and D4 and D7 from the other constitute the two detected groups of variables contributing the same information. The preliminary analysis of the correlation matrix had already pointed out the same pairs of distances as the most correlated ones. It therefore appeared possible to reduce a ten-dimensional problem to an eight-dimensional one by adopting only one distance within each group of correlated variables. Within the four possibilities (D1 and D4, D1 and D7, D4 and D5, D5 and D7) of selecting two uncorrelated distances, we decided to limit our choice to the two combinations involving four different description points (D1 and D7, D4 and D5). The final choice between these two remaining possibilities was made by analysing the behaviour of the two pairs of distances in the next phase of the chemometric elaboration. As the classification model (see below) constructed, using distances D1 and D7, exhibited a greater stability to the cross-validation procedure [23] than that constructed using D4 and D5, we chose the first pair of variables (D1 and D7) as uncorrelated distances to add to the other uncorrelated conformational descriptors (D2, D3, D6, D8, D9, D10). By means of PCA we completed the reduction of the complexity of the starting dataset without losing useful information.

An initial classification of the objects under study was then performed: centroids (a total of 449) of the active molecules were assigned to the active class and centroids of the inactive molecules (a total of 285) to the inactive class. A new classification method, linear discriminant classification tree (LDCT), was later employed to model the biological behaviour of the AII antagonists [20].

LDCT consists of the joint application of linear discriminant analysis (LDA) and classification

tree methods [24]; its aim is to integrate the classification tree methodology with the classical LDA procedure in order to use, at each step, a multivariate binary classifier instead of the common univariate criterion. Therefore, the result of a LDCT analysis is a multivariate decision tree, usually characterized by low complexity and easy interpretability.

In the first step of the analysis the discriminant function is calculated between two groups, obtained from a user-defined partition of the present classes (if only two classes are present, then groups and classes go together). The linear combination of all the variables which permits the maximum separation between the two groups is searched for [20]. If the separation is complete, then LDCT stops, otherwise the process continues by the same procedure applied on each successive subset (node) until the separation satisfies the user. The obtained tree structure is then analysed by cross-validation and the final choice of the optimal tree structure is made on this basis (see below). Each final node of the decision tree is usually called leaf.

The application of LDCT to our dataset of 734 centroids (objects), described by eight interatomic distances (variables) and distributed in two biological classes, has led to the decision tree represented in Fig. 2. It is characterized by six splits and each split is defined by a discriminant function as $R = a_1D_1 + a_2D_2 + a_3D_3 + a_6D_6 + a_7D_7 + a_8D_8 + a_9D_9 + a_{10}D_{10}$. At each step the significance of the variables in the classification can be determined from standardized coefficients of the linear discriminant functions. Analysis of the variable coefficients in all the discriminant functions of the tree model showed that variables D1, D2 and D10 are the most discriminant distances between the active and inactive classes, i.e. they had the highest weight in the construction of the decision tree; variables D3 and D6 possessed the lowest discriminating power.

LDCT has individuated four active, one inactive and two fuzzy leaves. A leaf was assigned to a class only if its specificity value [20] (i.e. the ratio between the number of objects of the more represented class and the total number of objects in the leaf) was greater than an imposed threshold value, otherwise it was not assigned but considered fuzzy. A specificity threshold value of 65% was adopted in the present study for the assignment of a class label to the final nodes of the decision tree. While specificity is synonymous with quality, the importance of a leaf, within the ensemble of leaves assigned to a class, is represented by the sensitivity value [20], i.e. the ratio

TABLE 2
CUMULATIVE PERCENT VARIANCE AND LOADINGS OF THE FIRST FOUR PRINCIPAL COMPONENTS OBTAINED BY PCA

Component identified	PC1	PC2	PC3	PC4
Cumulative % variance	28.24	54.09	71.27	83.04
D1:N–A	0.115	0.509	0.242	−0.007
D2:N–B	−0.092	−0.450	0.013	0.311
D3:N–C	0.102	−0.158	0.682	0.005
D4:N–E	0.498	−0.178	0.017	−0.373
D5:B–A	0.120	0.536	0.196	0.085
D6:B–C	0.216	−0.300	0.467	0.409
D7:B–E	0.527	−0.141	−0.103	−0.317
D8:A–C	0.244	0.140	−0.304	0.666
D9:A–E	0.498	−0.044	−0.294	0.187
D10:C–E	0.270	0.250	0.172	0.101

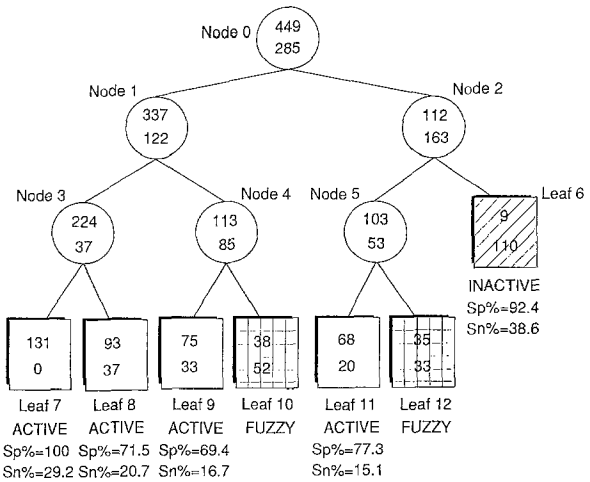

Fig. 2. LDCT model for nonpeptide AII antagonist dataset.

between the objects (in a leaf) of the class to which the leaf was assigned and the total number of objects in that class. Therefore leaf 7, which contains the greatest number of active centroids (131), is the most specific leaf, taking into account about 30% of the active centroids. Specificity (Sp%) and sensitivity (Sn%) values of each leaf are reported in Fig. 2.

Specificity and sensitivity can also be defined for the classes and the whole tree structure. For the rigorous definition of these parameters we refer to the original research paper on LDCT [20]. Another important classification parameter useful to characterize the tree model is the error rate percentage (ER%) [20], which is defined as the percentage ratio between the number of misclassified objects and the total number of objects. Misclassified objects are those centroids which, although belonging at the beginning to the active or inactive class, are given the opposite classification by the tree model.

The values of some relevant classification parameters characterizing our multivariate tree model are reported in the left column of Table 3. They express the capability of the tree model to describe, or to fit, the data of the training set but not the ability to make predictions on new data.

The tree model constructed by the LDCT method was then validated in order to estimate parameter values more suitable for prediction capability. The leave-one-out method of cross-validation [23] was adopted to check the structure of the tree model obtained. Each object is excluded one at a time and, on the basis of the discriminant functions calculated on the remaining objects, assigned to a leaf by following the tree structure. Cross-validated values of the classification parameters are reported in the right column of Table 3. It is worth noting that the cross-validated error rate, specificity and sensitivity only differ slightly from the corresponding estimates that have not been cross-validated.

Owing to the great stability of the multivariate tree model to the cross-validation procedure, we concluded that it could be reliably used to predict the biological behaviour of unedited nonpeptide AII inhibitors, on the basis of the values assumed by the eight selected interatomic distances.

LDCT has proved to be an excellent discriminant analysis method. In fact, due to the tree representation, it has been able to efficiently discriminate not only between the active and inactive classes, but also among different active situations. Therefore, in order to identify the geometrical

features of these situations, a statistical analysis was performed separately on the active centroids contained in each of the four discovered active leaves. The statistical parameters thus obtained (especially the mean values of the variables reported in Table 4) define four possible molecular geometries typical of only the pharmacologically more interesting compounds. These geometries were considered to be responsible for the binding affinity of the antagonists for the AII receptor.

Looking at the mean values of the distances, it can be observed that the geometries corresponding to leaves 8 and 11 seem to be very close to the two geometries found in the previous study on the nonpeptide AII antagonists with Du Pont skeleton only. In fact, leaves 8 and 11 are populated by centroids belonging only to Du Pont inhibitors. Leaf 9, which also contains Du Pont centroids, is characterized by a geometry representing an intermediate situation between geometries 8 and 11, but closer to the former. In particular, the geometries corresponding to leaves 8 and 11 significantly differ only in D7, D8 and D9 distance values, that is to say, in the relative spatial disposition of the acidic moiety with respect to the heterocycle ring and its substituents. Finally, leaf 7, which is populated exclusively by centroids belonging to the molecules A6, A7 and A8, introduces a new activity form, presenting a geometry differing from the previous ones, especially in the values of the distances D1 and D2 involving atoms N, A and B located on the heterocyclic ring.

CONCLUSIONS

The search for geometrical similarity by means of the above-described 3D QSAR approach, allowed the construction of an improved and more generally valid geometrical model for the activity of the nonpeptide AII receptor antagonists. Our methodology appears to be a promising and powerful tool to gain insight into the geometrical requirements for activity, especially when the receptor is unknown, the available biological activity data are not precise, and the ligands are conformationally very flexible.

Among the four different active model geometries proposed (Table 4), it is difficult to suggest the best set of distances for the 'best' AII inhibitor. In fact, the active leaves lie on the same tree level and none contain centroids derived from all the active molecules. However, on the basis of

TABLE 3
NOT CROSS-VALIDATED AND CROSS-VALIDATED FUZZY OBJECTS PERCENTAGE, ERROR RATE (ER%), SPECIFICITY (Sp%) AND SENSITIVITY (Sn%) FOR THE WHOLE TREE AND FOR THE ACTIVE AND THE INACTIVE CLASSES

Parameters	Not cross-validated	Cross-validated	Parameters	Not cross-validated	Cross-validated
Tree model			**Inactive class**		
Fuzzy obj.%	21.5	22.7	Fuzzy obj.%	29.8	30.2
ER%	13.5	14.2	ER%	31.6	33.0
Mean Sp%	82.1	81.0	Class Sp%	92.4	91.3
Mean Sn%	60.2	58.3	Class Sn%	38.6	36.8
Active class					
Fuzzy obj.%	16.3	18.0			
ER%	2.0	2.2			
Class Sp%	79.6	78.4			
Class Sn%	81.7	79.7			

TABLE 4
MEAN (Å), STANDARD DEVIATION, MINIMUM AND MAXIMUM (Å) OF THE SIGNIFICANT DISTANCES FOR THE ACTIVE CENTROIDS CONTAINED IN THE ACTIVE LEAVES OF THE LDCT MODEL

	Leaf 7				Leaf 8			
Distance	Mean	SD	Min.	Max.	Mean	SD	Min.	Max.
D1:N–A	3.716	0.137	3.536	4.142	4.421	0.102	4.231	4.622
D2:N–B	2.540	0.271	2.218	2.835	2.216	0.005	2.204	2.228
D3:N–C	6.584	0.998	4.145	9.150	6.587	0.297	5.708	7.115
D6:B–C	5.115	0.350	4.588	6.329	4.949	0.097	4.705	5.113
D7:B–E	6.882	1.404	3.952	9.954	6.900	1.569	4.257	10.62
D8:A–C	5.215	0.974	3.780	7.639	4.440	0.687	3.472	6.618
D9:A–E	6.809	2.012	2.503	10.48	5.116	2.091	2.596	10.53
D10:C–E	3.581	0.662	2.722	5.502	4.082	0.726	2.673	6.063
	Leaf 9				Leaf 11			
Distance	Mean	SD	Min.	Max.	Mean	SD	Min.	Max.
D1:N–A	4.563	0.483	3.869	5.812	4.788	0.551	4.319	5.876
D2:N–B	2.226	0.030	2.205	2.339	2.217	0.002	2.212	2.225
D3:N–C	6.578	0.370	4.873	7.060	6.390	0.231	5.881	7.070
D6:B–C	4.991	0.061	4.820	5.111	5.009	0.060	4.711	5.123
D7:B–E	6.575	1.150	5.015	10.15	8.120	1.424	5.444	10.36
D8:A–C	5.032	0.987	3.655	7.532	6.719	1.322	3.503	8.467
D9:A–E	5.724	2.218	2.534	10.72	9.781	2.026	4.277	13.29
D10:C–E	3.217	0.546	2.712	5.362	4.488	0.886	2.781	5.954

distance values, the number of the active model geometries can be reduced to three by keeping, from among the Du Pont leaves, only the most different 8 and 11 geometries.

The principal problems and limits of the geometrical activity model obtained are connected with the description of the biological behaviour of some Du Pont antagonists that differ in the side chain at the imidazole 2-position. In fact, the major part of the centroids misclassified by the geometrical activity tree model belong to molecules A3, I4 and I5, which feature different substitutions at C-2 of the imidazole on a common 2'-COOH-biphenyl skeleton. Improvement of the geometrical and structural model for activity, by extending the indirect theoretical approach to other molecular descriptors, seems necessary to better rationalize the activity of the nonpeptide AII receptor antagonists.

The comparative analysis of molecular electrostatic potential (MEP) distributions represents a first, and a particularly appropriate, step in this direction [10]. The actuality and the validity of the general scheme of this analysis and the importance of the results derived are fully confirmed by the above considerations concerning the limits of the geometrical activity model. Nonpeptide AII receptor inhibitors therefore present an interesting and complex SAR problem where both the geometrical and electronic properties of the compounds must be taken into account to fully rationalize their activity.

ACKNOWLEDGEMENTS

We wish to thank Professor W. Clark Still (Columbia University) for his permission to use

MacroModel for this research. Financial support from C.N.R. (Progetto Finalizzato Chimica Fine) and from a European Community grant (Human Capital and Mobility Program-A2, contract No. ERBCHRXCT 920027) is gratefully acknowledged. The calculations with the SAS system software were performed on the IBM 3090 at the Consorzio Interuniversitario Lombardo per l'Elaborazione Automatica (CILEA).

REFERENCES

1 Reid, J.L. and Rubin, P.C., Physiol. Rev., 67 (1987) 725.
2 Carini, D.J. and Duncia, J.J.V., Eur. Pat. Appl. 253, 310 (1988) (Chem. Abstr. 109, 129008g, 1988).
3 Duncia, J.V., Chiu, A.T., Carini, D.J., Gregory, G.B., Johnson, A.L., Price, W.A., Wells, G.J., Wong, P.C., Calabrese, J.C. and Timmermans, P.B.M.W.M., J. Med. Chem., 33 (1990) 1312.
4 Carini, D.J., Duncia, J.V., Johnson, A.L., Chiu, A.T., Price, W.A., Wong, P.C. and Timmermans, P.B.M.W.M., J. Med. Chem., 33 (1990) 1330.
5 Carini, D.J., Duncia, J.V., Aldrich, P.E., Chiu, A.T., Johnson, A.L., Pierce, M.E., Price, W.A., Santella III, J.B., Wells, G.J., Wexler, R.R., Wong, P.C., Yoo, S.E. and Timmermans, P.B.M.W.M., J. Med. Chem., 34 (1991) 2525.
6 Mantlo, N.B., Chakravarty, P.K., Ondeyka, D.L., Siegl, P.K.S., Chang, R.S., Lotti, V.J., Faust, K.A., Chen, T.B., Schorn, T.W., Sweet, C.S., Emmert, S.E., Patchett, A.A. and Greenlee, W.J., J. Med. Chem., 34 (1991) 2919.
7 Bradbury, R.H., Edwards, M.P., Luke, R.W.A., Pearce, R.J., Roberts, D.A., Major, J.S. and Oldham, A.A., 202nd ACS Meeting, New York, NY, August 25–30, 1991, Abstr. no. 102.
8 Allen, E.E., Greenlee, W.J., Patchett, A.A., Walsh, T.F. and Chakravarty, P.K., Eur. Pat. Appl. 419, 048 (1991) (Chem. Abstr. 115, 208010d, 1991).
9 Belvisi, L., Salimbeni, A., Scolastico, C., Todeschini, R. and Vulpetti, A., Pharm. Pharmacol. Lett., 1 (1991) 57.
10 Belvisi, L., Bonati, L., Bravi, G., Pitea, D., Scolastico, C. and Vulpetti, A., J. Mol. Struct. (Theochem), 281 (1993) 237.
11 The IC_{50} value of compound A6 was determined by Istituto LusoFarmaco d'Italia.
12 Belvisi, L., Brossa, S., Salimbeni, A., Scolastico, C. and Todeschini, R., J. Comput.-Aided Mol. Design, 5 (1991) 571.
13 Cosentino, U., Moro, G., Pitea, D., Scolastico, S., Todeschini, R. and Scolastico, C., J. Comput.-Aided Mol. Design, 6 (1992) 47.
14 With regard to the selection of the 'reference atoms' in each compound, we followed the functionality correspondences suggested by the various pharmaceutical companies when reporting the discovery of novel nonpeptide AII receptor antagonists. Such a correspondence derives from the identification of the following key structural elements in all the compounds: a heterocycle ring, a lipophilic side chain, a hydrogen-bond acceptor group (both these features being present on the heterocycle), a suitably positioned acidic moiety, ionized at physiological pH, and an aromatic spacer group connecting the acidic group to the heterocycle. By separating the different antagonists according to these key features, we found ambiguities only in the selection of the heteroatom hydrogen-bond acceptor (A ♦) in molecules A4, A8 and I2. For compounds A4 and I2, we assumed that the two oxygen atoms of the ester group would introduce the same geometrical information. For compound A8, we chose the pyridine nitrogen atom rather than the oxygen atom of the alcoxy group because it appears, on the basis of SAR studies, to be more directly involved in the binding process [7,25].
15 MacroModel V3.1X Primer and Batchmin V3.1, Department of Chemistry, Columbia University, New York, NY, 1990.
16 Chang, G., Guida, W.C. and Still, W.C., J. Am. Chem. Soc., 111 (1989) 4379.
17 Burkert, U. and Allinger, N.L., Molecular Mechanics, ACS Monograph 177, Washington, DC, 1982.
18 Ponder, W.J. and Richards, F.M., J. Comput. Chem., 8 (1987) 116.
19 SAS V6.06.01, SAS Institute Inc., Cary, NC, 1989.
20 Todeschini, R. and Marengo, E., Chemom. Intell. Lab. Syst., 16 (1992) 25.
21 Massart, D.L. and Kaufman, L., The Interpretation of Analytical Chemical Data by the Use of Cluster Analysis, Wiley, New York, NY, 1983.
22 Mardia, K.V., Kent, J.T. and Bibby, J.M., Multivariate Analysis, Academic Press, London, 1988.
23 Efron, B., the Jack-knife, the Bootstrap and Other Resampling Plans, Society for Industrial and Applied Mathematics, Bristol, 1982.
24 Breiman, L., Friedman, J.H., Olshen, R.A. and Stone, C.J., Classification and Regression Trees, Wadsworth and Brooks, Monterey, 1984.
25 Bradbury, R.H., Allott, C.P., Dennis, M., Fisher, E., Major, J.S., Masek, B.B., Oldham, A.A., Pearce, R.J., Rankine, N., Revill, J.M. Roberts, D.A. and Russel, S.T., J. Med. Chem., 35 (1992) 4027.