Coulomb and Overlap Self-Similarities: A Comparative Selectivity Analysis of Structure−Function Relationships for Auxin−like Molecules

Noel Ferro,*[,†,‡] Ana Gallegos,[§] Patrick Bultinck,[||] Hans-Jörg Jacobsen,[†] Ramón Carbó-Dorca,[§,||] and Thomas Reinard[†]

Institute for Plant Genetics, University of Hannover, Herrenhäuser Strasse 2, D-30419 Hannover, Germany, Department of Agronomy, Forestry and Agronomy Faculty, University of Pinar del Río, Cuba, Institute of Computational Chemistry, University of Girona, Catalonia, Spain, and Department of Inorganic and Physical Chemistry, Ghent University, Ghent, Belgium

Received November 11, 2005

Auxins are defined mainly by a set of physiological actions, but the structure-effect relationship still is based on chemical intuition. Currently a well-defined auxin molecular structure is not available. The existence of different auxin binding proteins and mechanisms of auxin action, the wide diversity of the auxin molecules, and the pleiotropic effects of auxin imply a completely different mechanism as described for the animal hormone concept. Here, we present a computational approach dealing with semiempirical optimizations of the auxin molecules themselves, which represent a number of about 250 different chemical structures. Our approach uses molecular quantum similarity measures and additional quantum variables for the analysis of auxin-like molecules. The finding of similarities in molecules by focusing basically on their electron structure results in new insights in the relationship of the different auxin groups. Additional statistical analysis allows the identification of relationships between similarity groups and their biological activity, respectively. It is postulated that the auxin-like molecular recognition depends more on specific molecular assembling states than on a specific ring system or side chain.

INTRODUCTION

A putative physiological active substance extracted from coleoptile tips was named auxin more than 50 years ago by Went. Later, indole-3-acetic acid had became firmly established as the auxin of higher plant tissues.[1,2] The number of putative auxins increased greatly up to hundreds of different chemical structures. In the 1970s, several groups tried to establish structure−activity relationships.[3,4] Nevertheless, no clear cause-effect relationship of auxins could be unraveled, even by applying modern molecular techniques.[5,6] An aggravating factor is that there are not one but numerous physiological effects caused by auxins, e.g. such important processes as cell division, cell growth by expansion, differentiation of root system, and gravitropic reaction of root, stem, and many more.[7] Unfortunately, most auxin-like substances are effective in some but not all auxin caused-effects. Still there is no defined molecular auxin structure; auxins are still characterized by phytophysiological definitions.

Certainly, the underlying concept of the structure−activity rules is that an auxin acts as a kind of coenzyme or ergon at the growth center, which (growth center) is a protein or enzyme surface of highly specific "shape".[8] However, different auxin-binding sites and proteins have been described by different groups.[9−12] The best characterized protein is the so-called Auxin-Binding-Protein 1 (ABP1). Definitively being an auxin-binding protein, its physiological role is debated, and it is not involved in all the different physiological auxin effects.[13] Furthermore, the considerable speculation about specialized receptor functions for specific transporters should be taking into account as well.[14] Recently a new complex of three proteins SCFTIR1 (where the transport inhibitor response 1 (TIR1) acts as receptor) has been described.[15−17]

Therefore it can be assumed that various auxin-interacting proteins with different binding specificities for the different auxin-like molecules occur in nature. To clarify these relationships, it is necessary to analyze the structures of the auxin-like molecules for common structural features, which can be used to group the different molecules into clusters of similar biological activities. To obtain reliable molecular models different computational methods have been combined. A conventional quantitative structure−activity relationship (QSAR) tries to explain and to predict activities by utilizing empirical descriptors. Unfortunately, the high structural diversity hampers the data structuring, which causes a large number of variables for the generation of outcome. This corresponds to the classical case of overparametrization. However, these large structural dissimilarities have common electronic sources as unsaturated rings and halogens. The ground-state electron density is the area wherein lie all molecular properties; therefore, this is the region to search for similarities and to associate with those activities.

* Corresponding author phone: +49-(0)511-762 3539; fax: +49-(0)-511-762 4088; e-mail: nferro71@yahoo.es.
[†] University of Hannover.
[‡] University of Pinar del Río.
[§] University of Girona.
[||] Ghent University.

10.1021/ci050491c CCC: $33.50
Published on Web 06/16/2006

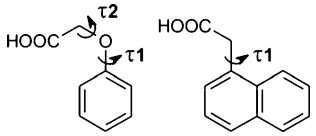


Figure 1. Angles of movement analyzed.

COMPUTATIONAL METHODS

Molecular Modeling. For the present study, 241 auxin molecules with different organic structures were chosen (Figure 1). In the first approach the molecular conformations were optimized using the MM+ force field in the Hyperchem program to improve the input geometries for the next optimization steps. No cutoffs for nonbonded interactions were used, and electrostatic interaction bond dipoles were used.[18] The final quantum chemical geometry optimizations were performed at the semiempirical level, using the software MOPAC v. 6. For indole-like molecule and other molecules with N or S in the rings, to avoid known problems of the structural geometry,[19] a first optimization with AM1 was performed. In the following step a second optimization was carried out by PM3, once the distances and angles related to the N and/or S atoms were frozen in the input matrix. In the case of the remaining molecules only PM3 was applied.

Analysis of side-chain movement for different auxin molecules was performed using a biostatistical approach defining the energy as variable (Figure 1). For the phenoxyacetic acid compounds an energy surface (ES) was plotted with two dihedral angles.

Molecular Quantum Similarity Measures. *(Quantitative) Structure−Activity Relationships ((Q)SAR)* are based on the well-known similarity principle: "the more similar two molecules are, the more similar their properties are". This statement requires a procedure to measure and quantify the degree of similarity between the compared systems.

Quantum similarity was performed by using the preceding semiempirical optimized structural conformations. This kind of similarity methods, as introduced by Carbó,[20] attempts to provide a measure of resemblance between two molecular structures, based on well-defined quantum-mechanical principles. In particular, a *Molecular Quantum Similarity Measure (MQSM)* can be expressed as the integral of the scalar product between the first-order molecular density functions associated with the molecules being compared and weighted by a positive definite operator (Ω)

$$Z_{AB}(\Omega) = \int\int \rho_A(\mathbf{r}_1)\Omega(\mathbf{r}_1,\mathbf{r}_2)\rho_B(\mathbf{r}_2)d\mathbf{r}_1 d\mathbf{r}_2 \qquad (1)$$

where A and B are the two molecules being compared, $\mathbf{r}_1$ and $\mathbf{r}_2$ are the electron coordinates, and ρ_A and ρ_B are the corresponding first-order density functions. Several types of MQSM can be defined according to the selection of the weighting operator, Ω.[21,22] In this study, *Overlap-like* and *Coulomb-like MQSM* have been considered.

Overlap MQSM gives a measure of the volume enclosed in the superposition of both molecules. In this case, the positive definite weight operator is the Dirac's delta distribution:

$$Z_{AB}(\Omega) = \int\int \rho_A(\mathbf{r}_1)\delta(\mathbf{r}_1,\mathbf{r}_2)\rho_B(\mathbf{r}_2)d\mathbf{r}_1 d\mathbf{r}_2 = \int \rho_A(\mathbf{r})\rho_B(\mathbf{r})d\mathbf{r} \qquad (2)$$

Coulomb MQSM. When the Coulomb operator is selected, a measure of the electronic repulsive Coulomb energy between the two charge densities is obtained:

$$Z_{AB}(\Omega) = \int\int \rho_A(\mathbf{r}_1)\delta(\mathbf{r}_1,\mathbf{r}_2)\rho_B(\mathbf{r}_2)d\mathbf{r}_1 d\mathbf{r}_2 = \int\int \rho_A(\mathbf{r}_1)\frac{1}{|\mathbf{r}_1 - \mathbf{r}_2|}\rho_B(\mathbf{r}_2)d\mathbf{r}_1 d\mathbf{r}_2 \qquad (3)$$

Computation of the Molecular Density Functions. To circumvent expensive computational calculations, the promolecular *Atomic Shell Approximation (ASA)*[23−25] has been used to compute density functions. This approximation considers the atoms in the molecule as spheres, approached by *1S* Gaussian functions and fitted to atomic ab initio density functions. The fitting ASA parameters, i.e., exponents and coefficients, can be downloaded from the Web site http://iqc.udg.es/cat/similarity/ASA/funcset.html. Then, the molecular electron density can be just computed as the sum of the discrete atomic density contributions

$$\rho_A^{\mathrm{ASA}} = \sum_{i\in A} w_i |S_i(\mathbf{r};\alpha_i)|^2 \qquad (4)$$

where i refers to the atomic shells, $\{w_i\}$ are the positive definite ASA coefficients, and $\{S_i\}$ is the set of normalized *1S* Gaussian type orbitals. It has been proved that the ASA densities differ less than 2% with the ab initio densities, so its use is clearly justified.[26] It has been shown by Bultinck *et al*. that the resulting MQSM values using ASA are nearly identical to the ones calculated using more involved quantum chemical methods.[27]

Determination of the Molecular Alignment. MQSM also depend on the relative orientation of the objects being compared. To align the molecular structures, the *maximum similarity superposition algorithm*[28,29] has been used. This field-based method considers that the optimal alignment is provided by the maximum value of the similarity measure for a given operator.

$$Z_{AB}(\Omega;\Theta) = \max_{\Theta} \int\int \rho_A(\mathbf{r}_1)\Omega(\mathbf{r}_1,\mathbf{r}_2)\rho_B(\mathbf{r}_2;\Theta)d\mathbf{r}_1 d\mathbf{r}_2 \qquad (5)$$

The Θ operator represents the transformation of the coordinates of B in relation to the A coordinates. Once computed, the overall set of pairwise MQSM can be stored in the so-called *Similarity Matrix (SM)*: $\mathbf{Z} = \{Z_{AB}\}$, where $\mathbf{Z}$ is a squared matrix of dimension N, i.e., the number of compounds. The diagonal of the similarity matrix is composed by the so-called *Quantum Self-Similarity Measures (QS−SM)*, which compares the molecule with itself:

$$Z_{AA}(\Omega) = \int \rho_A(\mathbf{r})\Omega(\mathbf{r})\rho_A(\mathbf{r})d\mathbf{r} \qquad (6)$$

Statistical Treatment of Biological Activity and Quantum Similarity Matrices. The definition of the "target tissue" in plant hormones is not clearly established like in animal hormones. Auxin activity assays described in the literature (Table 1) are somehow heterogeneous regarding the type of auxin response or tissue analyzed and the scoring used by different authors. Therefore, we have developed a consensus variable, which concentrates on the maximal response of a substance in one of the used tests. We have

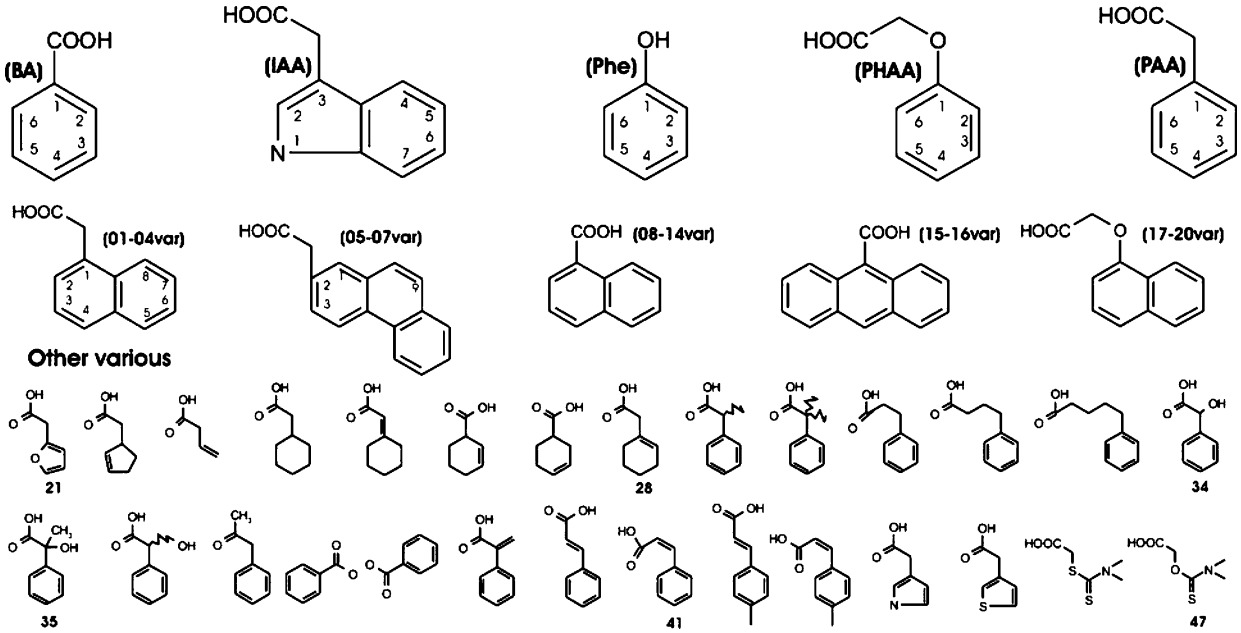


Figure 2. General structure of the analyzed auxin-like molecules.

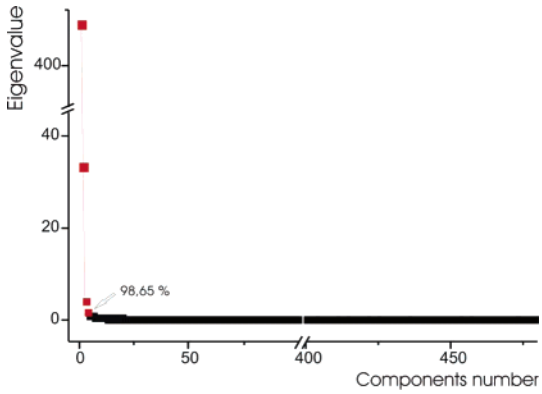


Figure 3. Scree plot of the Principal Component Analysis. The four selected principal components are over 1.

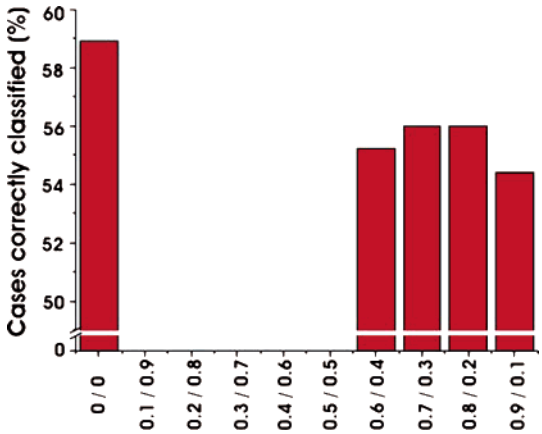


Figure 4. Discriminant analysis between active and inactive auxin molecules using different mixtures of Coulomb and Overlap quantum similarity matrix. The percent of cases correctly classified is shown in each mixture. X_{coul}/X_{over} represent probabilistic mixtures of the Coulomb/Overlap matrix where $p + q = 1$.

mainly focused on one type of assay: the coleoptile growth test. In one approach, we divided the auxin responses into active and nonactive ones (Figure 4). Later, we refined the consensus variable by introducing additional intermediate responses (Figure 5).

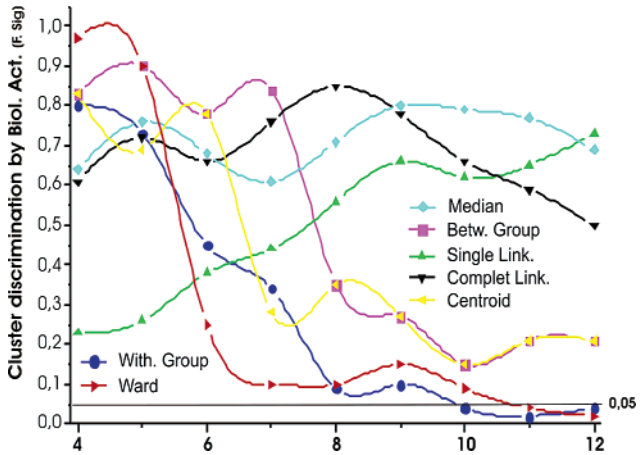


Figure 5. Analysis of the statistical significance of different grouping amount in different methods of cluster.

A statistical featuring of the auxin molecules was done using the calculated Overlap and Coulomb matrix. We were making probabilistic mixtures (7) of both matrices in order to develop a way to get all the possible information. One analysis without mixing matrices was performed as well (8).

$$[\text{mix } Z_{(i,j)}]_n = \{p[SM_{\text{over}}] + (1 - p)[SM_{\text{Coul}}]\}$$

$$p = 0.1...0.9;\ n = 9 \qquad (7)$$

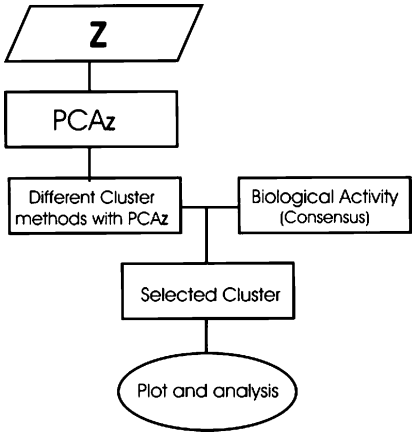
$$[\text{no} - \text{mix } Z_{(i,j)}] = \{[SM_{\text{over}}] \neq [SM_{\text{Coul}}]\} \qquad (8)$$

Next, a Principal Component Analysis (PCA) to the 10 resulting matrices was done in order to eliminate the repetitive information (*9*).

$$Z_{(i,j)} \approx Z_{(i,K)} K - \text{principal components} \qquad (9)$$

Each matrix was analyzed through cluster analysis (Euclidean distance) by means of the Principal Components. Thus the [no − mix] matrix was also mixed in the components. Therefore, it was possible to develop *Similarity Indices* without recurring information. Ten different com-

Chart 1: General Flow Scheme of Analysis



binations of the Overlap and Coulomb matrix were achieved. A Discriminant Analysis was applied to select the combinations related to the biological activities. The next step was to use different types of cluster analyses on the selected data to find causes of biological activities.

RESULTS

This work is based on the assumption that "similar chemicals behave similarly" and "structure generates properties". Therefore, we focused the analysis on molecular structure as the most reliable causal component in structure—activity functions. The results were projected onto the spectrum of biological activities described so far. In principle, we used the effect of different molecules on cell elongation as described before (Table 1). Therefore, we were able to consider the maximal spectrum of auxin molecules (Figure 2), which enabled us to obtain a highly representative statistical sample of the molecular population.

The great number of compounds with "auxin activity" is defined only by their physiological activities, since no general structural rules could be obtained up to now. At the end of 1970s, some new molecular methods began to be applied, techniques and ideas to explain this phenomenon molecularly.[4,30−32] An assessment of the relative activities among molecules has hampered the different assays, methods, and variables used by various authors. It has to be stressed out that our approach emphasizes a balance between activities and similarities rather than considering both on their own.

Once obtained, both similarity matrices for Coulomb and Overlap operators were probabilistically mixed in a sequential way. After each mixture a Principal Component Analysis was used to minimize the high repetitive information (Figure 3). In the case of the substances cyclohexylindeneacetic acid[33] and xanthates[34] we could detect that they are isolated from any group of similarity achieved by different matrices mixture methods. Since these substances are found inactive and usually not used as auxins, we decided to omit these from further analysis. When the Coulomb matrix was favored over Overlap matrix (Figure 4), we could discriminate between active and inactive compounds by these matrix mixtures. This led to the conclusion that the influence of the overlap matrix on biological activity is limited. The no mixture matrix delivered the highest biological information by means of discriminant analysis (Figure 4).

Consequently, the no mixture matrix was analyzed further by evaluation of seven cluster methods to find the relationship between similarities and biological activities (Figure 5). Ward[35,36] and Within Groups[36] are the subsequent methods able to discriminate compounds with different biological activities (Figure 5). The Ward method was the most stable in the biological discrimination along the grouping procedure, and it was more statistically significant in the analysis of variance by five times. The best discrimination in both methods was obtained for 11 cluster groups by analysis of variance (α=0.05).

A process of classification and analysis of the quantum objects was performed using the cluster mentioned before. The proximity matrix (239 $\times$ 239) from the cluster analysis, based on the Principal Components, was used as the similarity index. This enabled us to discriminate between three main nuclei of similarity along the diagonal, corresponding to the three biggest clusters (Figure 6). However, there is no possibility to detect differences in activity among them. It can be inferred that the influence of these molecules on cell expansion is not related with great dissimilarities but with close specific similarities.

The Ward method, based on the minimization of the variances, gave phenomenologically a consistent classification in compass with the sources of error from experimental viewpoint. According to the statistical consensus boundary between chemical similarities and biological properties, the compounds could be grouped in different classes. A confirmatory multiple analysis of means distinguished five groups (graphic at left down corner and dendrogram in the right panel of Figure 6). However, the high variance, as result of both molecular and biological systems, prevented their clear discrimination. A closer look at the graphic (Figure 6) uncovers the existence of two important statistical differences as well. Groups 1, 2, 8, and 11 reveal much less activity behaviors compared to the other groups.

The partial results required additional arguments to clarify the information by means of visualization and phenomenological analysis. Therefore, the distribution of the quantum objects within a three-dimensional similarity space was carried out using the three principal components.

The naturally occurring indole-3-acetic acid (IAA) and its synthetic stable substitutes 1-naphthaleneacetic acid (1-NAA) and 2,4-dichlorophenoxyacetic acid (2,4-D), three prominent compounds in the literature of plant growth regulators belonging to the large group of auxin-like molecules, share the same quantum spatial regions. IAA and 2,4-D are especially closely related.

It is noteworthy that the neighboring compounds share similar biological activities as well (Figure 7). Molecules with totally different side chain and rings system may result in a close relationship, visible by their graphical vicinity (Figure 7). While, molecules, which look related on first sight, can—structurally seen—be far away from those active compounds, like 8-Cl-NAA.

We also observed that the substitutions may easily affect the total shape of the electron molecular space. As an example, the substitution of Cl to 8-Cl-NAA results in a completely different location, now much closer to 2,3,6-dichlorobenzoic acid (Tryben) or 3-phenanthreneacetic acid than to NAA (Figure 7).

Additionally, information on the behavior of data analyzed was obtained by other statistical methods. The two correlation analyses, ring system—biological activity and ring system—

Table 1. Molecules Analyzed[a]

number	name	biocon-sensus	classes	ref	number	name	biocon-sensus	classes	ref
				Benzoic Acid (BA)					
01-BA	benzoic acid	0	1	4	18-BA	2-iodo-5-chlorobenzoic acid	1	8	4
02-BA	2-fluorobenzoic acid	0	2	4	19-BA	2-iodo-5-bromobenzoic acid	1	8	4
03-BA	2-chlorobenzoic acid	1	3	4	20-BA	2,5-diiodobenzoic acid	1	2	4
04-BA	2-bromobenzoic acid	1	4	4	21-BA	3-chloro-2-iodobenzoic acid	2	3	4
05-BA	2-nitrobenzoic acid	1	4	4	22-BA	2,3,4-trichlorobenzoic acid	0	3	4
06-BA	2-aminobenzoic acid	0	1	4	23-BA	2,3,5-trichlorobenzoic acid	1	5	4
07-BA	2-iodobenzoic acid	0	1	4	24-BA	2,3,6-trichlorobenzoic acid	3	9	4
08-BA	2-methylbenzoic acid	0	2	4	25-BA	2,4,5-trichlorobenzoic acid	0	4	4
09-BA	2,3-dichlorobenzoic acid	2	6	4	26-BA	2,4,6-trichlorobenzoic acid	0	8	4
10-BA	2,4-dichlorobenzoic acid	0	4	4	27-BA	3,4,5-trichlorobenzoic acid	0	3	4
11-BA	2,5-dichlorobenzoic acid	1	5	4	28-BA	2,3,5-triiodo-benzoic acid	3	4	4
12-BA	2,6-dichlorobenzoic acid	1	6	4	29-BA	2,3,6-tricloro-4-fluorobenzoic acid	2	8	4
13-BA	2-chloro-5-bromobenzoic acid	2	5	4	30-BA	2,6-dichloro-4-fluoro-3-nitro-BA	1	8	4
14-BA	2-chloro-5-iodobenzoic acid	3	7	4	31-BA	2-bromo-3-nitrobenzoic acid	2	1	4
15-BA	2-bromo-5-chlorobenzoic acid	2	6	4	32-BA	3-amino-2,5-dichlorobenzoic acid	1	7	4
16-BA	2,5-dibromobenzoic acid	2	4	4	33-BA	2,5-dichloro-3-nitrobenzoic acid	2	3	4
17-BA	2-bromo-5-iodobenzoic acid	3	5	4					
				Indole (IAA)					
01-IAA	1-indole acetic acid	0	8	33	37-IAA	7-aza-indoleacetic acid	3	5	4
02-IAA	2-indole acetic acid	0	5	4	38-IAA	7-Cl-indoleacetic acid	3	3	52
03-IAA	1-methyl indole acetic acid	1	8	4	39-IAA	7-hydroxyindoleaceic acid	1	5	53
04-IAA	2-benzofuraneacetic acid	1	10	33	40-IAA	7-methoxy-3-indolepropionic acid	0	8	33
05-IAA	2-bromoindol-3-acetic acid	3	7	52	41-IAA	7-methoxy indoleacetic acid	1	2	4
06-IAA	2-carboxy-3-indolebutyric acid	0	3	33	42-IAA	alpha-methyl-3-indoleacetic acid	3	4	33
07-IAA	2-carboxy-3-indolepropionic acid	0	10	33	43-IAA	4,7 Cl-indolacetic acid	1	3	53
08-IAA	2-cloroindol-3-acetic acid	3	8	4	44-IAA	6,7 Cl-indoleacetic acid	3	3	39
09-IAA	2-indolecarboxylic acid	0	5	33	45-IAA	benzothiophenoacetic acid	0	1	54
10-IAA	2-methylindol-3-acetic acid	1	6	52	46-IAA	ethyl-3-(2-bromoindole) acetate	3	10	52
11-IAA	2-methyl-indole	0	1	52	47-IAA	ethyl-3-(2-chloroindole) acetate	3	2	52
12-IAA	3-aza-indole	0	4	52	48-IAA	indoleacetic acid	3	4	33
13-IAA	3-CHO-indole	0	3	52	49-IAA	indene-3-acetic acid	2	3	53
14-IAA	3-COOH-indole	2	10	52	50-IAA	indene-3-propionic acid	1	3	53
15-IAA	oxindolone	0	8	53	51-IAA	indole	0	5	53
16-IAA	3-indolebutyric acid	3	5	33	52-IAA	indole-2-acetic acid	1	8	52
17-IAA	4,5 Cloro indoleacetic acid	2	2	39	53-IAA	indole-3-glycolic acid	1	2	53
18-IAA	3-indolepropionic acid	3	6	33	54-IAA	indole-3- glyoxylic ester	1	3	53
19-IAA	3-indolesuccinic acid	1	3	33	55-IAA	indole-3-isobutyric acid	1	3	53
20-IAA	3-indolevaleric acid	2	10	33	56-IAA	indole-4-acetic acid	1	3	53
21-IAA	skatole	0	5	33	57-IAA	DL-indole-3-lactic acid	1	1	53
22-IAA	4-Cl-indoleacetic acid	3	5	52	58-IAA	indolemethylene mevalonic	1	8	53
23-IAA	5,7-dichloro-2-methyl IAA	0	7	4	59-IAA	indoxylic acid	0	7	33
24-IAA	5,7-di-Cl-indoleacetic acid	1	6	52	60-IAA	indoxyl	0	6	33
25-IAA	5-Cl-indoleacetic acid	3	5	52	61-IAA	isatin	0	11	33
26-IAA	5-F-indoleacetic acid	3	3	52	62-IAA	methyl-3-(2-bromoindole) acetate	3	3	52
27-IAA	5-hydroxyindoleaceic acid	1	10	53	63-IAA	methyl-3-(2-chloroindole) acetate	3	10	52
28-IAA	5-methoxy-3-indolepropionic acid	0	5	33	64-IAA	methyl-3-indoleacetate	3	9	52
29-IAA	5-methyl-indoleacetic acid	2	2	52	65-IAA	N-adetyl-3-hydroxyindole	0	6	53
30-IAA	5-methylindole	0	8	53	66-IAA	*N*-methyl-indoleacetic acid	1	3	52
31-IAA	5-methoxy indoleacetic acid	1	5	4	67-IAA	5-Cl-7-Me-indoleacetic acid	1	3	39
32-IAA	4,6 Cl-indoleacetic acid	1	3	39	68-IAA	oxindole-3-aceonitrile	0	3	53
33-IAA	6-Cl-indoleacetic acid	3	5	52	69-IAA	indole-3-pyruvic acid	0	1	53
34-IAA	6-methoxy-3-indolepropionic acid	0	5	33	70-IAA	oxindolene-3-acetonitrile	0	10	53
35-IAA	6-methylindole	0	2	53	71-IAA	oxindolene-3-pyrubic	0	7	53
36-IAA	6-methoxyindoleacetic acid	1	10	4	72-IAA	Indole-3-acetamide	0	3	33
				Phenol (Phe)					
01-Phe	2,3-Cl-6-NO2-phenol	2	6	51	18-Phe	2,6-F-phenol	0	2	51
02-Phe	2,5-Cl-6-NO2-phenol	2	4	51	19-Phe	2,6-ICN-phenol	2	7	51
03-Phe	2,6-Br−CN-phenol	2	6	51	20-Phe	2,6-INO2-phenol	2	7	51
04-Phe	2,6-Br−I-phenol	2	1	51	21-Phe	2,6-I-phenol	2	2	51
05-Phe	2,6-Br−NO2-phenol	3	8	51	22-Phe	2,6-MeBr-phenol	1	2	51
06-Phe	2,6-Br-phenol	2	5	51	23-Phe	2,6-MeCl-phenol	0	2	51
07-Phe	2,6-Cl-3-NO2-phenol	2	6	51	24-Phe	2,6-MeCN-phenol	0	8	51
08-Phe	2,6-Cl−Br-phenol	2	5	51	25-Phe	2,6-MeI-phenol	0	6	51
09-Phe	2,6-Cl−CN-phenol	2	7	51	26-Phe	2,6-MeNO2-phenol	1	7	51
10-Phe	2 Cl, 6-I-phenol	2	5	51	27-Phe	2,6-Me-phenol	0	2	51
11-Phe	2,6-Cl−NO2-phenol	3	6	51	28-Phe	2,6-MoxiBr-phenol	0	2	51
12-Phe	2,6-Cl-phenol	2	6	51	29-Phe	2,6-NO2−CN-phenol	2	6	51
13-Phe	2,6-CN-phenol	0	2	51	30-Phe	2,6-NO2−OH-phenol	0	2	51
14-Phe	2,6-FCl-phenol	1	2	51	31-Phe	2,6-NO2-phenol	3	7	51
15-Phe	2,6-FCN-phenol	0	7	51	32-Phe	2,6-phenyl-Cl-phenol	1	7	51
16-Phe	2,6-FI-phenol	2	7	51	33-Phe	2,6-phenyl-NO-phenol	0	7	51
17-Phe	2,6-FNO2-phenol	1	7	51	34-Phe	2-CF-3,6-NO2-phenol	3	7	51

Table 1. (Table 1 contd)

number	name	biocon-sensus	classes	ref	number	name	biocon-sensus	classes	ref
					Phenoxyacetic Acid (PHAA)				
01-PHAA	phenoxyacetic acid	0	11	4	20-PHAA	2,3-dichlorophenoxyacetic acid	2	7	4
02-PHAA	2-chlorophenoxyacetic acid	1	4	4	21-PHAA	2,4-dichlorophenoxyacetic acid	3	4	4
03-PHAA	3-chlorophenoxyacetic acid	2	6	4	22-PHAA	2,5-dichlorophenoxyacetic acid	2	3	4
04-PHAA	4-chlorophenoxyacetic acid	2	6	4	23-PHAA	2,6-dichlorophenoxyacetic acid	1	7	4
05-PHAA	2-bromophenoxyacetic acid	3	4	4	24-PHAA	3,4-dichlorophenoxyacetic acid	3	4	4
06-PHAA	3-bromophenoxyacetic acid	2	6	4	25-PHAA	3,5-dichlorophenoxyacetic acid	1	3	4
07-PHAA	4-bromophenoxyacetic acid	1	6	4	26-PHAA	2,3,4-trichlorophenoxyacetic acid	3	6	4
08-PHAA	2-iodophenoxyacetic acid	1	6	4	27-PHAA	2,3,5-trichlorophenoxyacetic acid	1	1	4
09-PHAA	4-iodophenoxyacetic acid	0	6	4	28-PHAA	2,3,6-trichlorophenoxyacetic acid	3	3	4
10-PHAA	2-methoxyphenoxyacetic acid	2	4	4	29-PHAA	2,4,5-trichlorophenoxyacetic acid	3	7	4
11-PHAA	3-methoxyphenoxyacetic acid	1	6	4	30-PHAA	2,4,6-trichlorophenoxyacetic acid	0	3	4
12-PHAA	4-methoxyphenoxyacetic acid	1	3	4	31-PHAA	3,4,5-trichlorophenoxyacetic acid	1	7	4
13-PHAA	2-methylphenoxyacetic acid	1	7	4	32-PHAA	2,4-dichloro-5-nitrophenoxyacetic acid	1	3	4
14-PHAA	3-methylphenoxyacetic acid	2	4	4	33-PHAA	2,4-dichloro-6-fluorophenoxyacetic acid	3	7	4
15-PHAA	4-methylphenoxyacetic acid	1	8	4	34-PHAA	2,4-dichloro-6-bromophenoxyacetic acid	0	1	4
16-PHAA	2-nitrophenoxyacetic acid	0	4	4	35-PHAA	2,4-dimethylphenoxyacetic acid	2	7	4
17-PHAA	3-nitrophenoxyacetic acid	1	2	4	36-PHAA	2,5-dimethylphenoxyacetic acid	1	4	4
18-PHAA	4-nitrophenoxyacetic acid	1	4	4	37-PHAA	2,6-dimethylphenoxyacetic acid	0	6	4
19-PHAA	2,4-dibromophenoxyacetic acid	2	3	4	38-PHAA	3,5-dimethylphenoxyacetic acid	0	4	4
					Phenylacetic Acid (PAA)				
01-PAA	phenylacetic acid	1	7	4	10-PAA	2,3-dichlorophenylacetic acid	3	11	4
02-PAA	2-nitrophenylacetic acid	0	1	4	11-PAA	2,4-dichlorophenylacetic acid	1	2	4
03-PAA	3-fluorophenylacetic acid	1	11	4	12-PAA	2,6-dichlorophenylacetic acid	3	4	4
04-PAA	3-nitrophenylacetic acid	0	8	4	13-PAA	2,4-dimethylphenylacetic acid	1	11	4
05-PAA	4-aminophenylacetic acid	1	11	4	14-PAA	3,5-dimethylphenylacetic acid	1	1	4
06-PAA	4-fluorophenylacetic acid	1	7	4	15-PAA	2,4-dinitrophenylacetic acid	0	8	4
07-PAA	4-iodophenylacetic acid	0	4	4	16-PAA	2,3,6-trichlorophenylacetic acid	3	7	4
08-PAA	4-nitrophenylacetic acid	0	11	4	17-PAA	2,4,6-trimethyl phenylacetic acid	0	4	4
09-PAA	4-phenylphenylacetic acid	0	4	4					
					Various				
01-var	1-alpha-naphthaleneacetic acid	3	3	33	25-var	cyclohexylindeneacetic acid	0	*b*	33
02-var	2-chloro naphthylacetic acid	3	1	4	26-var	*1-cycloheneacetic acid	0	11	33
03-var	8-Cl-naphthylacetic acid	0	9	4	27-var	*2-cyclohexeneacetic acid	1	5	33
04-var	2-naphthalenacetic acid	1	7	4	28-var	*1-cyclohexenephenylacetic acid	0	5	33
05-var	2-phenanthreneacetic acid	0	6	4	29-var	hydratropic acid (2-phenylpropionic acid)	1	6	33
06-var	3-phenanthreneacetic acid	0	8	4	30-var	alpha,alpha-dimethyl-alpha-toluid acid	0	6	33
07-var	9-phenanthreneacetic acid	1	10	4	31-var	hydrocinnamic acid (3-phenylpropionic acid)	1	1	33
08-var	1-naphthoic acid	1	9	4	32-var	gamma-phenylbutyric acid	0	11	33
09-var	2-cloronaphthoic acid	2	7	4	33-var	delta-phenylvaleric acid	0	6	33
10-var	8-cloronaphthoic acid	2	6	4	34-var	d and l-mandelic acid	0	4	33
11-var	8-bromonaphthoic acid	2	6	4	35-var	d-and-l-antrolactic acid	0	5	33
12-var	8-iodonaphthoic acid	2	10	4	36-var	dl-tropic acid	0		33
13-var	8-methylnaphthoic acid	2	10	4	37-var	benzyl methyl ketone	0	11	33
14-var	2-naphthoic acid	0	2	4	38-var	dibenzoyl peroxide	0	7	33
15-var	9-antracenecarboxilic acid	0	1	4	39-var	atropic acid (2-phenylpropenoic acid)	1	1	33
16-var	anthraceneacetic acid	2	6	33	40-var	trans-cinnamic acid	0	11	33
17-var	1-naphthoxyacetic acid	1	9	4	41-var	cis-cinnamic acid	1	4	33
18-var	2-naphthoxyacetic acid	2	11	4	42-var	trans-*p*-methylcinnamic acid	0	11	33
19-var	1-chloronaphthoxyacetic acid	0	6	4	43-var	cis-*p*-methylcinnamic acid	1	7	33
20-var	3-chloronaphthoxyacetic acid	0	1	4	44-var	pyrrole-3-acetic	0	1	53
21-var	furylacetic (isomer 1)	2	6	53	45-var	thiophene-3-acetic	0	11	53
22-var	cyclopenteneacetic acid	0	8	53	46-var	S-(carboximethyl)-dimethyldithiocarbamate	1	1	34
23-var	3-butenoic acid	0	11	33	47-var	O-alkyxantates	0	*b*	34
24-var	cyclohexaneacetic acid	0	6	33					

[a] Bio-concensus: 0 − inactive; 1 − low active; 2 − middle active; 3 − high active. Classes: classification (11 groups) from the matrix without mix. [b] Molecules out of classification.

similarity grouping, were not significant statistically. The evidences suggest that the ring system itself does not confer any independent influence on the biological activity and hardly any influence on the similarity grouping of the quantum objects. Its effect is dissolved in the whole molecular system.

Counting a molecule as a whole system (Figure 8), further energetic analysis of the side chain was carried out. The movements of side chain of some active and inactive molecules were analyzed by behavior of their energy (Figure 8, surface energy for phenoxyacetic acid). It is vital to take into account the previous requirements of a molecule to be active. A short impression is shown by the dissimilar potential surface energy among them (Figure 8, PES). Analyzing the movement of the side chain is a further requisite. A molecule cannot be blocked on both planar sides with respect to the ring, that almost certainly makes it inactive, e.g. 2,4,6-Cl-phenoxyacetic acid with quantum similarity to NAA but inactive (Figure 7). However, it can be seen that the directional orientation of the side chain in

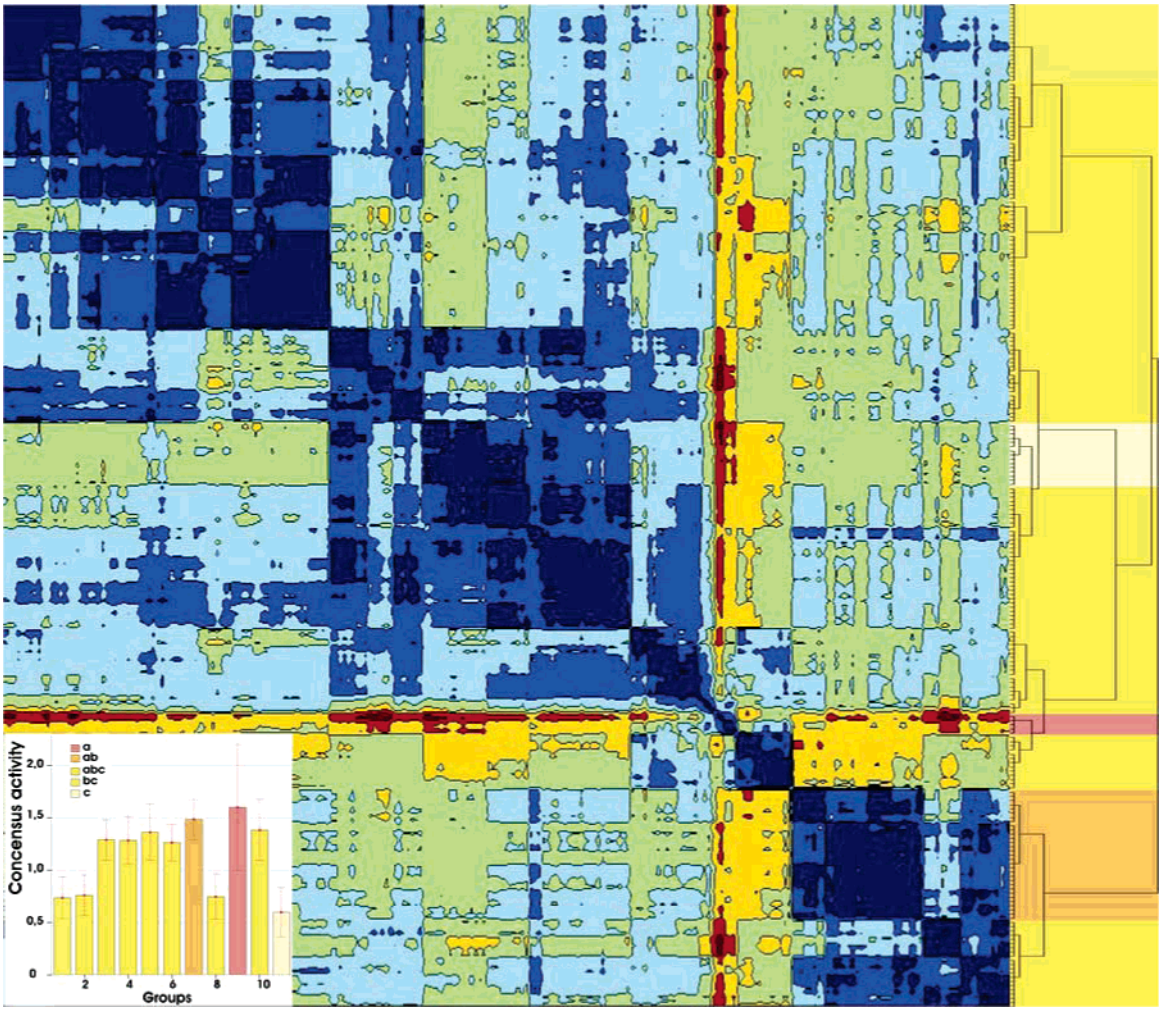


Figure 6. Graphical view of the similarity matrix of the cluster. Statistical differences of the biological activity are assessed through an ANOVA analysis for different groups of similarity. Duncan's multiple comparison analysis (α=0.05) was run to determine the grouping a, ab, abc, bc, and c as result (graphic left down).

one planar direction is a characteristic of active molecules (Figure 8) because of the decreasing gradient energy in one direction of the molecular plane. Even the probability of finding energies higher than the great mean in active molecules is analyzed as well (Figure 8). To prove this statement, active molecules "similar geometric angles with similar energies" and inactive molecules "alternated geometric angles with similar energies" were classified by means of cluster analysis. A simple structural overlay shows that the largest atom in one side of the ring of active molecules is fluorine (Figure 8, consensus below). It facilitated the side-chain orientation in this direction. It suggests that bigger atoms in this position, like Cl or Br, obstruct conformations in this area.

It suggests that, first, the side-chain movement of auxin is not an independent event from the combination of ring and its substitutions. Second, the huge auxin-like molecular diversity cannot be taken into account equivalently for each of the probable auxin carriers and/or receptors, even might not for one of them as well. The specific arrangements between rings and the kind of substitution near to the side chain are influencing energetically a directional conformational change of the side chain to a planar direction (Figure 8).

DISCUSSION

The chemical space, which encompasses the auxin definition, is a definitive important gap in plant biology. The molecular mechanism of auxin action is poorly understood, and the structures of most auxin-like molecules cannot be correlated with their biological activities. Applying the key-lock principle, common for enzyme−substrate-correlations, to the auxin biology, is definitively not coincidental with the observations made.

Many molecules have been evaluated in many different hypothetical "auxin tests". A relative behavior of compounds is strongly influenced by the type of assay performed, i.e., in *Avena* IAA is 1000 times more effective as 2,4-D; in split pea test 2,4-D is 12 times as effective as IAA, but in straight growth test IAA and 2,4-D have comparable activities.[37] Furthermore, some effects are the result of several underlying molecular mechanisms, as shown for the overall organ extension of the hypocotyl.[38] The endogenous auxin (indole-3-acetic acid) could influence the response to exogenous applied compounds. Finally, the purity of the compounds tested has not always been satisfactory, as in the case of the double chloride indoles.[39] The direction of the chemical−biological information cannot be elucidated. A given assay system detects physiologically active substances which have particular propensities, notably that they may be assayed by

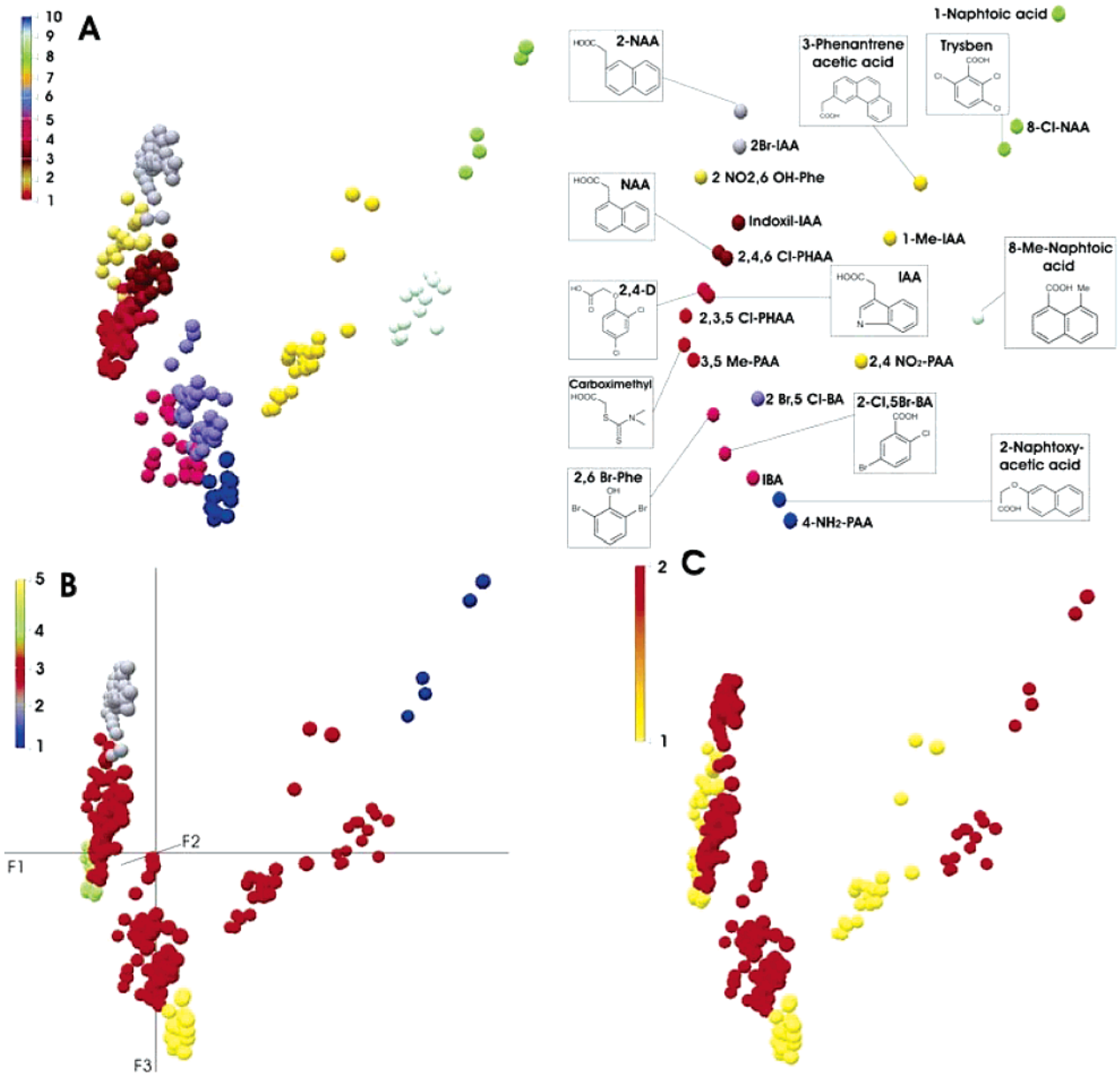


Figure 7. Relationship between similarity grouping and biological activity in auxin-like molecules. The use of different colors allows illustrating the similarity groups developed from the Coulomb and Overlap matrix (A); the statistical significant differences of biological activities detected between the groups of the similarity (B) and the two most important statistical differences in the biological activity of the analyzed molecules was emphasized (C). An overview of standing structures in the whole figure is given in superior right panel. The axes are the factors of a PCA.

a given test. The characteristics of a class of compounds as assayed may, therefore, be more a property of the assay system used than of the full range of biological activities of the substances in question.[37] Additionally, under Jacob's generality criterion, one would not perhaps wish to exclude a compound as a chemical controller solely because no other species than the one tested had been found to have the same system. This could be due to a new evolutionary development or (more likely perhaps) because of ignorance on our part.[40] For that reason, a consensus variable is an acceptable approach in order to compare the activities of all these molecules.

The influence of different mixtures of both matrices on the biological activity was the first step evaluated by discriminant analysis. The mixtures were not able to discriminate activity in any case, where the overlap operator was probabilistically benefited. Such an operator may be used when the volume of the components of a given molecular system is determinately viewed as steric field.[41,42] Katekar suggested the existence of steric obstruction (steric hindrance), which is referred to as size and shape limitations.[4,43,44] This is not in agreement with our results. The Coulomb descriptor selected statistically by discriminant analysis was able to distinguish compounds with biological activities in every favorable probabilistic mixture. This operator is known to better reflect existent electrostatic interactions.[42]

Once the Z_{AB} matrix was calculated, most of the information was found in the diagonal, and there is too much repetitive information in the remaining parts of this matrix. For our molecular approach we decided not to use the known Molecular Quantum Similarity Index, like off-diagonal element of the similarity measure.[41,42] To eliminate the

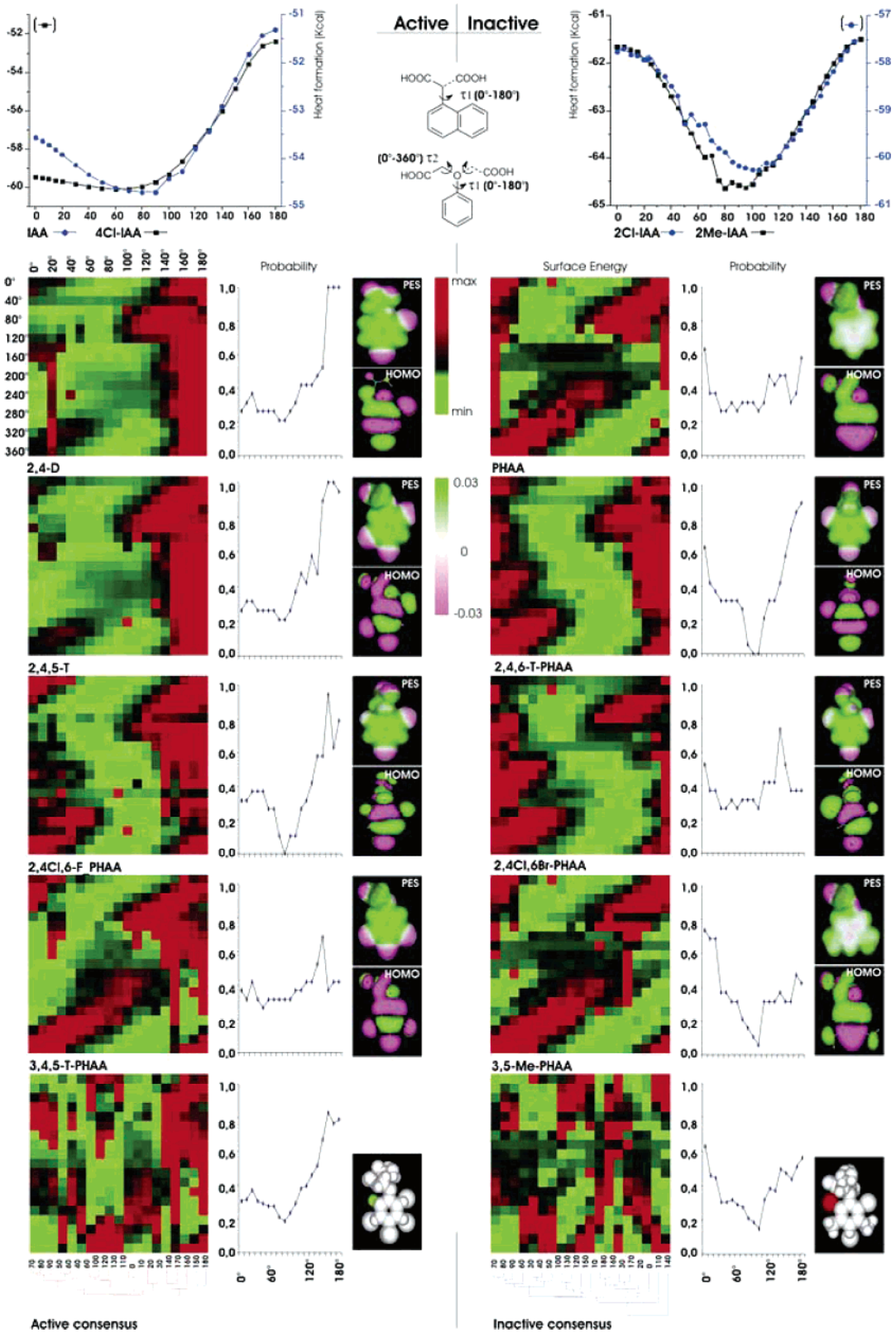


Figure 8. Analysis of the movement of the side chain of some active and nonactive auxin molecules. The energy surface of some phenoxyacetic compounds (PHAA) using second and third chain bond were modeled. It is possible to compare with other types of auxin side chains. The Potential Energy Surface (PES) was calculated, color ranges, in e/a^3_0: green, represents the electron deficit regions; red, electron excess regions. The first side-chain bond related with the movement of COOH group was not taken into account.

repetitive information referent to the quantum objects we converted the normal matrix ($Z_{AB} = Z_{(i,j)}$) to ($Z_{AB} \approx Z_{(i,K)}$), where K are Principal Components of this original Z_{AB} matrix. Then, a cluster analysis using the Principal Components ($Z_{(i,k)}$) was performed.

At this point, the cluster analysis was the key procedure. We developed a scheme to look for connections between dendrograms and biological activities. Bultinck *et al*. used clusters analysis for molecules to find groups of similarities, but he applied the index directly from the matrix and

biological activity was not considered.[45] The adequate strategy able to perform relationships between biological activities and similarities was based on the Ward method.[35] In this method, the error sum of square (ESS) is given by a functional relation; in the case of multivariable analysis, a minimization of the sum of the variance and covariance is the way to group the cases when minimal variation is allowed.

Both cluster methods (Ward and Within Groups), which can be applied to explain some relation between the quantum objects and their biological activities, are focused on the search for individual identities forming hierarchical cluster groups of mutually exclusive subsets.

Our results show that the major structural differences of the superior clusters are irrelevant for biological activity, in fact, small groups with minimal variance of the potential are crucial for activity up to now. The compounds analyzed are frequently active. Therefore, the molecular space analyzed covers almost the maximum molecular spectrum known of auxins but not the random spectrum of active and inactive molecules. Furthermore, the lack of correlation between the type of ring and both the biological activity and grouping by similarity index are further reasonable evidence to doubt the chemical intuition proposal,[46] which is also the basis directly or indirectly for further publications.

Katekar's work was mainly based on a work of Jönsson carried out almost 20 years before. Jönsson stated that it was far too early to predict how the finding at that time may influence the structure−activity discussions.[47] His approaches used the assumption that the chemical behavior in complex biological systems is context-dependent.

Another assessment to reveal relationships between structure and activity considered auxin as a pharmacophore and focused on the indole ring system.[4,44] Afterward, phytophore was another name concerned to the ABP1 binding site, but there was no relationship with a chemical definition.[48] This is the best known definition of pharmacophore in auxin, which describes critical substructures in a molecule but is also a less rigorous use of this concept. Although identification of pharmacophoric features is a potent approach, the concept of pharmacophore mapping strives to discover the common three-dimensional patterns present in diverse molecules that act at the same enzyme or receptor target site.[49] Both experimental evidences of multireceptor activity of auxins and data presented here strongly suggest the existence of more than one likely pharmacophoric center. Additionally, it is necessary to say that the biological information analyzed and the lack of a careful random conformational study of auxin-like molecules does not allow one to go deeper in the understanding of the phenomenon called auxin.

The possibility of conformational changes within the auxin molecules had been discussed at the end of the 1970s.[3] Different groups of molecules selected on their chemical or biological objective were analyzed.[4,6,31,46] [As they are checked for a recognition conformation, which suggests that 4-Cl-IAA would be less active because of the electron rich chloride atom,[4] or postulated, for the presence of the nitro and hydroxyl group in adjacent positions of the ring is important for the auxin activities,[31] or looked for a minimum energy conformation,[46] or proposed active conformers[6]]. Everything is focused on a rigid conformation idea, rather than the existence of random shapes, even analyzing molecules able to bind different carriers and/or receptors. It is assumed, basically, the approach of a single low-energy conformation whose rigidity is a major deficiency is provoked to find new approaches and methodologies to obviate this issue.[49] Our present studies suggest some energetically flexible conformations of the side chain[50] are adaptable for most of auxin molecules. Molecules with free energy movement in the side chain were not active. The complete block of the side chain was found to be harmful for biological activity as well, especially in the phenoxyacetic acid group. However, a reference point of guided energy flexibility is suggested as a positive issue for biological activity.

We propose that the angles at the end of the side chain are not important, but the one next to the ring is of crucial importance for the activity of the molecule. This point of view explains that the strong influence of ring substitutions in several cases, including the presence of two ortho-substituents per se, in phenoxyacetic acid, did not interfere with the activity but it was the size that did.[47] That is suggested by the energy of the side chain as well.

Several authors tried to explain the observation, that some molecules, of shapes resembling IAA or NAA, are inactive by conformational changes. However, the activity of compounds without a side chain like phenol of benzoic acid has to be explained. Our results with the inactive compound 8-chloronaphthalene acetic acid show no three-dimensional similarity with respect to typical auxins; therefore, at least for this molecule, the conformational problem referred to before[46] is the wrong approach. Furthermore, we found several phenolic and benzoic acidic substitutes with very high mimic abilities by our three-dimensional similarity approach. This could not be explained adequately before, by just chemical intuition based on classical quantum chemical calculus.[31]

CONCLUSIONS

Here we present an integrative method of different probabilistic mixtures of Molecular Quantum Similarity Measures matrix (Overlap and Coulomb) and the full range of biological activities of 241 auxin-like molecules. It was demonstrated that the Coulomb matrix, as an electrostatic potential descriptor, was the most important differentiating auxin activity by discriminant analysis. According to the boundary of relationship between molecular similarity and phenomenological information (tests) suitable clusterings were achieved, projecting the structural relationships emerging from different cluster methods onto the biological activity.

Empirically, the use of the conformational change has been used extensively; however, we consider other variables, which will result in more precise relation toward the biological activity. It is suggested that the conformational movement should be focused on its relationship to the electrostatic potential change. Data presented here will enable us to concentrate on a group of molecules with a high probability for possessing auxin activity, even if they could not account for when to apply the classical rules for active auxins. Several phenol and benzoic acid molecules were found to be spatially similar to the natural indole-3-acetic acid. On the other hand, molecules such as 8-chloro NAA are not related to the clusters of the known active auxin.

The existence of active molecules in a different, isolated group suggests that more than one enzyme acts as an auxin binding receptor, possibly depending on cellular compartment or physiological effect.

ACKNOWLEDGMENT

I would like to thank my family and especially my wife for their efforts to take care of my children without me (N.F.). This work was supported by a grant from the Deutsche Akademischer Austauschdienst (DAAD) to Noel Ferro and also by continuous support from the National Fund for Scientific Research in Flanders, Belgium (FWO-Vlaanderen).

REFERENCES AND NOTES

(1) Went, F. W. Auxin, the Plant-Growth Hormone. 2. *Bot. Rev.* **1945**, *11* (9), 487−496.
(2) Letham, D. S.; Goodwin, P. B.; Higgins, T. J. V. *The biochemistry of phytohormones and related compounds*; Elsevier/North-Holland Biomedical Press: Amsterdam, 1978; Vol. I.
(3) Kaethner, T. M. Conformational Change Theory for Auxin Structure-Activity-Relationships. *Nature* **1977**, *267* (5606), 19−23.
(4) Katekar, G. F. Auxins − Nature of the Receptor-Site and Molecular Requirements for Auxin Activity. *Phytochemistry* **1979**, *18* (2), 223−233.
(5) Tomic, S.; Gabdoulline, R. R.; Kojic-Prodic, B.; Wade, R. C. Classification of auxin related compounds based on similarity of their interaction fields: Extension to a new set of compounds. *Internet J. Chem.* **1998**, *1* (26), Cp1-U21.
(6) Tomic, S.; Gabdoulline, R. R.; Kojic-Prodic, B.; Wade, R. C. Classification of auxin plant hormones by interaction property similarity indices. *J. Comput.-Aided Mol. Des.* **1998**, *12* (1), 63−79.
(7) Kulaeva, O. N.; Prokoptseva, O. S. Recent advances in the study of mechanisms of action of phytohormones. *Biochemistry (Moscow)* **2004**, *69* (3), 233−247.
(8) Audus, L. J. Metabolism and mode of action. In *Encyclopediea of Plant Physiology*; Ruhland, W., Ed.; Springer: Berlin, 1961; Vol. 14.
(9) Batt, S.; Venis, M. A. Separation and Localization of 2 Classes of Auxin Binding-Sites in Corn Coleoptile Membranes. *Planta* **1976**, *130* (1), 15−21.
(10) Ray, P. M.; Dohrmann, U.; Hertel, R. Specificity of Auxin-Binding Sites on Maize Coleoptile Membranes as Possible Receptor-Sites for Auxin Action. *Plant Physiol.* **1977**, *60* (4), 585−591.
(11) Jacobsen, H. J. 2 Different Soluble Cytoplasmic Auxin-Binding Sites in Etiolated Pea Epicotyls. *Plant Cell Physiol.* **1984**, *25* (6), 867−873.
(12) Reinard, T.; Achmus, H.; Walther, A.; Rescher, U.; Klambt, D.; Jacobsen, H. J. Assignment of the auxin binding abilities of ABP(44) in gel. *Plant Cell Physiol.* **1998**, *39* (8), 874−878.
(13) Napier, R. M.; David, K. M.; Perrot-Rechenmann, C. A short history of auxin-binding proteins. *Plant Mol. Biol.* **2002**, *49* (3−4), 339−348.
(14) Leyser, O. Molecular genetics of auxin signaling. *Annu. Rev. Plant Biol.* **2002**, *53*, 377−398.
(15) Gray, W. M.; Kepinski, S.; Rouse, D.; Leyser, O.; Estelle, M. Auxin regulates SCFTIR1-dependent degradation of AUX/IAA proteins. *Nature* **2001**, *414* (6861), 271−276.
(16) Kepinski, S.; Leyser, O. Ubiquitination and auxin signaling: A degrading story. *Plant Cell* **2002**, *14*, S81−S95.
(17) Kepinski, S.; Leyser, O. The Arabidopsis F-box protein TIR1 is an auxin receptor. *Nature* **2005**, *435* (7041), 446−451.
(18) Allinger, N. L. Conformational-Analysis. 130. Mm2 - Hydrocarbon Force-Field Utilizing V1 and V2 Torsional Terms. *J. Am. Chem. Soc.* **1977**, *99* (25), 8127−8134.
(19) Edgerton, M. D.; Tropsha, A.; Jones, A. M. Modeling the Auxin-Binding Site of Auxin-Binding Protein-1 of Maize. *Phytochemistry* **1994**, *35* (5), 1111−1123.
(20) Carbó, R.; Leyda, L.; Arnau, M. How Similar Is a Molecule to Another? − An Electron-Density Measure of Similarity between Two Molecular-Structures. *Int. J. Quantum Chem.* **1980**, *17* (6), 1185−1189.
(21) Carbó, R. Molecular Similarity and Reactivity: From Quantum Chemical to Phenomenological Approaches. In *Molecular Similarity and Reactivity: From Quantum Chemical to Phenomenological Approaches*; Kluwer: Amsterdam, 1995; pp 3−30.
(22) Carbó-Dorca, R.; Mezey, P. G. JAI Press: Greenwich, CT,: 1998; Vol. 2, pp 1−42.
(23) Amat, L.; CarboDorca, R. Quantum similarity measures under atomic shell approximation: First order density fitting using elementary Jacobi rotations. *J. Comput. Chem.* **1997**, *18* (16), 2023−2039.
(24) Amat, L.; Carbó-Dorca, R. Fitted electronic density functions from H to Rn for use in quantum similarity measures: cis-diammine-dichloroplatinum(II) complex as an application example. *J. Comput. Chem.* **1999**, *20* (9), 911−920.
(25) Amat, L.; Carbó-Dorca, R. Molecular electronic density fitting using elementary Jacobi rotations under atomic shell approximation. *J. Chem. Inf. Comput. Sci.* **2000**, *40* (5), 1188−1198.
(26) Gironés, X.; Amat, L.; Carbó-Dorca, R. A comparative study of isodensity surfaces using ab initio and ASA density functions. *J. Mol. Graphics Modell.* **1998**, *16* (4−6), 190−196.
(27) Bultinck, P.; Carbó-Dorca, R.; Van Alsenoy, C. Quality of approximate electron densities and internal consistency of molecular alignment algorithms in molecular quantum similarity. *J. Chem. Inf. Comput. Sci.* **2003**, *43* (4), 1208−1217.
(28) Constans, P.; Amat, L.; Carbó-Dorca, R. Toward a global maximization of the molecular similarity function: Superposition of two molecules. *J. Comput. Chem.* **1997**, *18* (6), 826−846.
(29) Bultinck, P.; Kuppens, T.; Girone, X.; Carbó-Dorca, R. Quantum similarity superposition algorithm (QSSA): A consistent scheme for molecular alignment and molecular similarity based on quantum chemistry. *J. Chem. Inf. Comput. Sci.* **2003**, *43* (4), 1143−1150.
(30) Farrimond, J. A.; Elliott, M. C.; Clack, D. W. Charge Separation as a Component of Structural Requirements for Hormone-Activity. *Nature* **1978**, *274* (5669), 401−402.
(31) Farrimond, J. A.; Elliott, M. C.; Clack, D. W. Auxin Structure−Activity-Relationships - Benzoic-Acids and Phenols. *Phytochemistry* **1980**, *19* (3), 367−371.
(32) Farrimond, J. A.; Elliott, M. C.; Clack, D. W. Auxin Structure-Activity-Relationships - Aryloxyacetic Acids. *Phytochemistry* **1981**, *20* (6), 1185−1190.
(33) Koepfli, J. B.; Thimann, K. V.; Went, F. W. Phytohormones: structure and physiological activity. I. *J. Biol. Chem.* **1938**, *122*, 763−780.
(34) Fawcett, C. H.; Wain, R. L.; Wightman, F. Plant Growth-Regulating Activity in Certain Carboxylic Acids Not Possessing a Ring Structure. *Nature* **1956**, *178* (4540), 972−974.
(35) Ward, J. H. Hierarchical Grouping to Optimize an Objective Function. *J. Am. Stat. Assoc.* **1963**, *58* (301), 236-&.
(36) Anderberg, M. R. *Cluster analysis for applications*; Academic Press: New York, 1973; p 359.
(37) Steward, F. C.; Krikorian, A. D. *Plants, chemicals and growth*; Accademic Press, Inc.: New York, 1971; p 232.
(38) Barley, K. Why hypocotyl extension mutants need to be characterized at the cell level: a case study of axr3-1. *J. Exp. Bot.* **2004**, *55* (399), 1071−1078.
(39) Hatano, T.; Katayama, M.; Marumo, S. 5,6-Dichloroindole-3-Acetic Acid as a Potent Auxin − Its Synthesis and Biological-Activity. *Experientia* **1987**, *43* (11−12), 1237−1239.
(40) Weyers, J. D. B.; Paterson, N. W. Plant hormones and the control of physiological processes. *New Phytol.* **2001**, *152* (3), 375−407.
(41) Maggiora, G. M.; Petke, J. D.; Mestres, J. A general analysis of field-based molecular similarity indices. *J. Math. Chem.* **2002**, *31* (3), 251−270.
(42) Carbó-Dorca, R.; Gironés, X. Quantum Similarity and Quantum Structure−Activity Relationships. In *Computational medicinal Chemistry for drug Discovery*; Bultinck, P., De Winter, H., Langenaeker, W., Tollenaere, J. P., Eds.; Marcel Dekker: New York, 2004; pp 365−385.
(43) Katekar, G. F.; Geissler, A. E. Auxin II: the effect of chlorinated indolylacetic acids on pea stems. *Phytochemistry* **1982**, *21* (2), 257−260.
(44) Katekar, G. F.; Geissler, A. E. Structure-activity differences beteween indoleacetic acid auxins on pea and wheat. *Phytochemistry* **1983**, *22* (1), 27−31.
(45) Bultinck, P.; Carbó-Dorca, R. Molecular quantum similarity matrix based clustering of molecules using dendrograms. *J. Chem. Inf. Comput. Sci.* **2003**, *43* (1), 170−177.
(46) Katekar, G. F.; Winkler, D. A.; Geissler, A. E. Hormone recognition in plants. In *Plant hormone Receptor*; Klämbt, D., Ed.; Springer-Verlag: Berlin, 1987; pp 13−26.
(47) Jönsson, A. Chemical structure and growth activity of auxins and antiauxins. In *Encyclopediea of Plant Physiology*; Ruhland, W., Ed.; Springer: Berlin, 1961; Vol. 14, pp 959−1006.
(48) Jones, A. M. Auxin-binding proteins. *Annu. Rev. Plant Physiol. Plant Mol. Biol.* **1994**, *45*, 393−420.
(49) Mason, J. S.; Good, A. C.; Martin, E. J. 3-D Pharmacophores in drug discovery. *Curr. Pharm. Des.* **2001**, *7* (7), 567−597.
(50) Murrall, N. W.; Davies, E. K. Conformational Freedom in 3-D Databases. 1. Techniques. *J. Chem. Inf. Comput. Sci.* **1990**, *30* (3), 312−316.

(51) Harper, D. B.; Wain, R. L. Studies on plant growth - regulating substances. XXX. The plant growth - regulating activity of substituted phenols. *Ann. Appl. Biol.* **1969**, *64*, 395−407.
(52) Porter, W. L.; Thimman, K. V. Molecular requeriments for auxin action-I. Halogenated indoles and Indoleacetic acid. *Phytochemistry* **1965**, *4*, 229−243.
(53) Thimann, K. V. Auxin Activity of Some Indole Derivatives. *Plant Physiology* **1958**, *33* (5), 311−321.
(54) Went, F. W.; Thimann, K. V. *Phytohormones*; The Macmillan Company: New York, 1937.

CI050491C