Analysis of Structure-Activity Relationships of auxin – like molecules

Von der

Naturwissenschaftlichen Fakultät

der Gottfried Wilhelm Leibniz Universität Hannover

zur Erlangung des Grades

eines Doktors der Naturwissenschaften

Dr. rer. nat.

genehmigte Dissertation

von

Master en Biología Vegetal Noel Ferro Diaz

geboren am 28. 03.1971

in Havanna, Kuba

2006

Referent: Prof. Dr. H.-J. Jacobsen

Korreferent: PD Dr. Th. Bredow

Tag der Promotion: 07. 07. 2006

Analysis of Structure-Activity Relationships of auxin – like molecules

Von der

Naturwissenschaftlichen Fakultät

der Gottfried Wilhelm Leibniz Universität Hannover

zur Erlangung des Grades

eines Doktors der Naturwissenschaften

Dr. rer. nat.

genehmigte Dissertation

von

Master en Biología Vegetal Noel Ferro Diaz

geboren am 28. 03.1971

in Havanna, Kuba

2006

Abstract

Auxins have been the first plant growth regulators discovered several decades ago. They play a complex and important role in the dendritic network of the physiological processes in plants. Nevertheless, a correlation between their chemical structure and its biological implication could not be found. Many theories on correlations of structure and activity have been developed, which are based mainly on chemical intuition. This is the first approach to apply methods of computational chemistry combined with biostatistics. By the application of these methods, the distribution of the outer molecular orbitals on the molecule can be identified and also their impact on the biological activity.

The use of Molecular Quantum Similarity Measures (MQSM) on structures of auxin-like molecules resulted in a conceptual framework to classify auxin structures from a biostatistical point of view. Similarity matrices of both, Overlap and Coulomb, were used for semi-empirical optimizations. When using more than 200 auxin-like molecules, the Coulomb Matrix was able to discriminate best between differences in activity, which is caused by the decisive influence of the electrostatic interactions. First, a classification of auxins (quantum objects) by different clusters methods was performed, followed by the creation of a biological consensus variable, which – in the beginning - depended on biological data (bioassays) from literature. Then those structural groups were identified, which reveal a relation with the appropriate biological activity. This resulted in a classification of all molecules in a defined biological sense.

The classification of molecular quantum similarities solved some long-known confusing issues discussed in literature for years, like the inactivity of molecules like 8Cl-NAA or the activity of benzoic and phenolic compounds, since these molecules were sorted in the reasonable group by the approach used in this thesis.

Based on this conceptual model, highly standardized bioassays at a multi-dimensional scaling level and with parallel screening of different auxins were carried out for the very first time. The structure activity relationship approach was supported by *ab-initio* optimizations.

The whole concept proved to be valuable, since new active molecules (quantum objects) predicted via statistical grouping-analysis of MQSM were verified in

different bioactivity assays. The uncommon structure of the new active auxin-like molecule, 2, 6-dibromo-phenol, a non-carboxylated compound, fitted perfectly in the structure-activity concept developed here. The variable hardness (η) was found to play the major role in the correlation between structure and activity of auxins. Hardness (η), related with the biological activity of auxins, refers to a reaction of electronic arrangement. The chemical condition of the ring system determines the biological effects by the localization of the HOMO and HOMO-1 molecular orbitals.

Keywords: Auxin, structure-activity correlation, MQSM, bioassay, 2,6-dibromo-phenol

Zusammenfassung

Auxine wurden bereits vor mehreren Jahrzehnten als pflanzliche Wachstumsregulatoren beschrieben. Sie spielen eine komplizierte und zentrale Rolle im verzwiegten Netzwerk der physiologischen Wechselwirkungen in der Pflanze. Dennoch konnte bisher keine Korrelation zwischen ihrer chemischen Struktur und der biologischen Wirkweise gefunden werden. Viele Theorien zur Korrelationen von Struktur und Wirkung sind entwickelt worden, allerdings stützen sich diese hauptsächlich auf chemische Intuition. Im Rahmen der vorliegenden Doktorarbeit wurden erstmalig Methoden der theoretischen Chemie zur Ermittlung von Korrelationsparametern angewendet.

Im Rahmen dieser Arbeit wurde die Methode der "Molecular Quantum Similarity Measures" (MQSM) zur Strukturanalyse von auxinähnlichen Molekülen angewandt und lieferte ein erste konzeptionale Zusammenhänge, welche die Grundlage für die weiteren Arbeiten lieferte. Diese sollten eine Klassifizierung auxinartiger Strukturen vom biostatistischem Standpunkt aus ermöglichen. Ähnlichkeitsmatrizen von "Overlap-" und "Coulomb-"Eigenschaften wurden für semi-empirische Optimierungen verwendet. Für mehr als 200 auxinähnliche Moleküle lieferte die Coulomb-Matrix die beste Unterscheidung der Aktivitäten verschiedener Moleküle, woraus gefolgert werden kann, dass elektrostatischen Wechselwirkungen einen entscheidenden Einfluss auf die Auxin-wirkung haben.
Zunächst wurden die Auxine (Quantum Objekte) durch Einsatz verschiedener Cluster-Methoden klassifiziert. Anhand der Klassifizierung konnte eine biologische Einheitsvariable erstellt werden, deren Relevanz mit Daten von Bioaktivitätstests aus der Literatur gezeigt werden konnte. Anschließend wurden jene Strukturgruppen identifiziert, die eine enge Korrelation mit der jeweiligen biologischen Aktivität aufwiesen. Auf diese Weise konnten alle Molekülstrukturen in Bezug auf ihre biologische Aktivität hin klassifiziert werden.
Durch die Klassifizierung auf Basis von molekularen Quantum-Ähnlichkeiten konnten so einige Phänomene im Bereich der Auxinwirkung geklärt werden, die schon seit Jahrzehnten kontrovers diskutiert werden: beispielsweise die Inaktivität von Molekülen wie 8Cl-NAA oder die Auxinaktivität von Benzoesäure- und

Phenolgruppen, denn solche Moleküle tauchten in der Strukturklassifizierung in Bereichen auf, die ihrer Auinwirkung entsprechen.

Gestützt auf dieses Konzept wurde hochgradig standardisierte biologische Aktivitätstests parallel für mehere Auxine durchgeführt und multifaktoriell ausgewertet. Dieser Ansatz zur Klärung der Beziehungzwischen Auxinstruktur und Auxinwirkung wurde durch *ab initio* Optimierungen unterstützt.

Schließlich konnte das Konzept eindrucksvoll verifiziert werden, da neue, aktive Moleküle (Quantum-Objekte), die mittels statistischer Gruppenanalyse aus MQSM vorhergesagt wurden, in Biotests genau die erwarteten Eigenschaften aufwiesen. Die ungewöhnliche Struktur des neuen, aktiven auxinähnlichem Moleküls 2,6-diBromphenol, einem Molekül ohne Carboxylgruppe, passte perfekt in die entwickelten Zusammenhang zwischen Struktur und Aktivität. Als zentrale Variabel für diese Korrelaton stellte sich die "Hardness" (η), welche das Arrangement der Elektronen im Molekül widerspiegelt. Die chemischen Eigenschaften des Ringsystems bestimmen die biologischen Wirkungen durch die Lage der molekularen Orbitale HOMO und HOMO-1.

Stichworte: Auxin, Korrelation zwischen Struktur und Aktivität, MQSM, biologische Aktivitätstest, 2,6-di Bromphenol

Contents

ABSTRACT ..... **I**

ZUSAMMENFASSUNG ..... **III**

CONTENTS ..... **V**

ABBREVIATIONS ..... **VII**

CHAPTER 1:
GENERAL INTRODUCTION **1**

Plant growth regulators, overview 2
Evolution of auxin concept 3
Auxin perception, signal transduction and gene expression 6
Structure-activity 8
Chemical approach: 9
Physico-chemical approach 10
Binding site models 11
Literature 15

CHAPTER 2:
STRUCTURE–ACTIVITY ANALYSIS ON ECDYSTEROIDS: A STRUCTURAL AND QUANTUM CHEMICAL APPROACH BASED ON TWO BIOLOGICAL SYSTEMS **21**

CHAPTER 3:
COULOMB AND OVERLAP SELF - SIMILARITIES: A COMPARATIVE SELECTIVITY ANALYSIS OF STRUCTURE - FUNCTION RELATIONSHIPS FOR AUXIN - LIKE MOLECULES **34**

CHAPTER 4:
UNREVEALED STRUCTURAL REQUIREMENTS FOR AUXIN - LIKE MOLECULES BY THEORETICAL AND EXPERIMENTAL EVIDENCES **35**

CHAPTER 5:
SUPPLEMENTARY DISCUSSION **63**

Plant bioassays: a poor structural mirror *63*
Auxin, its molecular diversity and pleiotropic activity *64*
How to face the problems 65
Two sets of structural requirements for auxin-like molecules 68
General reflexions 69
Outlook and future prospect 70
Literature 71

APPENDIX **75**

PUBLICATIONS 75

Journals *75*

Book reports *77*

Congresses *77*

DECLARATION / ERKLÄRUNG 78

ACKNOWLEDGEMENTS 79

CURRICULUM VITAE *80*

Abbreviations

2,4 Br-PHAA	2,4-Dibromophenoxyacetic acid
2,4-Cl-PAA	2,4-Dichlorophenylacetic acid
2,4-D	2,4-Dichlorophenoxyacetic acid
2,6 Cl-PAA	2,6-Dichlorophenylacetic acid
2,6-Br-Phe	2,6-Dibromophenol
2,6-NO2-Phe	2,6-Dinitrophenol
2,4,5-T	2,4,5-Trichlorophenoxyacetic acid
2Cl-6NO2-Phe	2-Chloro-6-nitrophenol
2-F-BA	2-Fluorobenzoic acid
2-NAA	2-Naphthaleneacetic Acid
2-Naphtoic acid	2-Naphthoic Acid
2NO2-PHAA	2- Nitro Phenoxy acetic acid
3-F-PAA	3-Fluor Phenylacetic acid
3Me-PHAA	3 Methyl Phenoxy acetic acid
ABP1	auxin-binding protein 1
ASA	Atomic Shell Approximation
Dicamba	3,6-Dichloro-2-methoxybenzoic acid
DL-IndLacticAA	DL-Indole-3-lactic acid
ED50	Effective dosage at the 50% level
ER	Endoplasmic Reticulum
HOMO	Highest Occupied Molecular Orbital
I-3-Acetamide	Indole -3-acetamide
IAA	Indole-3-acetic acid
IBA	Indole-3-butyric acid
KDEL	KDEL Sequence
logP	Octanol-water partition coefficient
LUMO	Lowest Unoccupied Molecular Orbital
MQSM	Molecular Quantum Similarity Measures
NAA	Naphtalen acetic acid
Naphthoic a.	1-Naphthoic Acid
PCA	Principle Component Analysis,
PHAA	1-Phenoxy acetic acid
Picloran	4-Amino-3,5,6-Tricloro picolinic acid
QSAR	Quantitative structure-activity relationships
SAR	structure activity relationship
TIBA	2,3,5-Triiodo Benzoic acid
Trysben	2,3,6-Trichloro Benzoic Acid

Chapter 1

General Introduction

In plants, where the most rapid physiological response is thousand times slower than in animals, we have: fewer specialized cell types, passive circulatory system, sessility, most cells remain totipotent, a cell wall, no nervous system and a photosynthetic apparatus. Rather than the existence of a common biochemistry, plant and animals share points of contact between parallel biochemical systems (Verhey and Lomax 1993). It is not possible to expect the same cellular performance for chemical regulation.

The expressions “auxin” and “hormone” were used synonymously to coin "correlation carriers" at the beginning of the century in animals and plants, respectively. The hormone concept, as messenger substance, was accepted widely in biology. Besides auxin, several other plant messengers (phytohormones) are known today, but the differences expressed when compared to the animal system are gaining increased attention day by day. Contrary to animals, plants rarely have peptide hormones and brassinosteroids are the only steroid group with physiological significance in plants (Haubrick and Assmann 2006; Verhey and Lomax 1993). During the last few years it has turned out that the hormone concept, developed for animals, cannot easily be transferred to plants. Many significant gaps still exist in our knowledge about “hormone perception” and physiological changes in plant messengers. Currently, the hormonal regulation of plant life is one of the “hot spots” in biochemistry, physiology and plant molecular biology research (Kulaeva and Prokoptseva 2004).

However, practical applications of plant growth regulators have been highly recommended and are being exploited. Herbicides, tissue culture and rooting are the most recognized applications (Arteca 1995). Bioregulators represent the second most important issue in modern agriculture (Sasson 1993). Despite of the wide use of these growth regulators, the underlying cause-effect principles are only partially understood. This relationship is fully depending on many factors, which influence interaction regularizations between hormones and their pleiotropic effects.

Plant growth regulators, overview.

Indole-3acetic acid (IAA) was the first plant growth regulator to be isolated, and the most prominent auxin known so far. Auxins are generally characterized by a non-saturated ring (nucleus) with COOH in a side chain. Besides indole-3-acetic acid (IAA), 4-cloro-indol-3-acetic acid (4-Cl-IAA) and phenylacetic acid (PAA) (Sasse, 1991; Arteca, 1995) occur as natural auxin. IAA is found in both free and conjugated forms, the later forms are usually inactive. A huge number of synthetic auxins have been discovered within the last decades, which are mainly used in commercial applications (Davies, 1995).

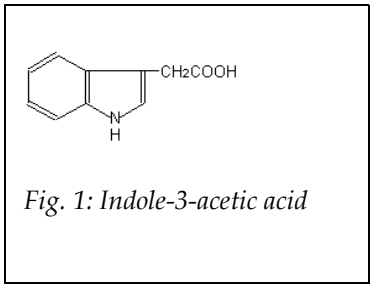


Fig. 1: Indole-3-acetic acid

Gibberellins were found to be the causal substances when abnormal rice growth occurred due to fungal infection. The fungus produced an ent-giberelano structure, of which main exponent is called gibberellic acid (first commercially available). Chemically these substances belong to the diterpens, of which around 90 compounds were known in the 1990s, actually there are 136 (Joo *et al.* 2005).

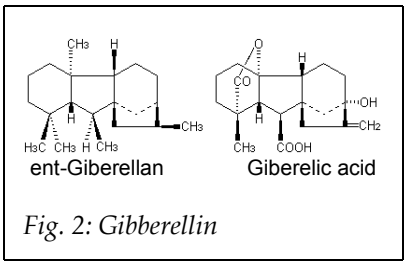


Fig. 2: Gibberellin

Adenine related substances provoke cell division in plant tissue culture. These substances were called

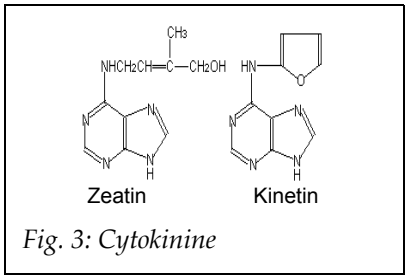


Fig. 3: Cytokinine

cytokinines. The first synthetic cytokinine from DNA was the kinetin (6-furfurilaminopurine), which produces cytokinesis in tobacco culture. Later, zeatin [6-(4-hidroxi-3-metil-trans-2-butenil-amino) purin] was isolated from immature corn. At present, there are approximately 20 purine derivatives known (Binns, 1994; Davies, 1995).

Abscisic acid (ABA), a sesquiterpene (15 C atoms), is another product broadly distributed in the plant kingdom . It can not only be found in higher plants, but also in algae, mushrooms, etc. and is associated generally with the abscission processes and dormancy (Davies, 1995; Arteca, 1995).

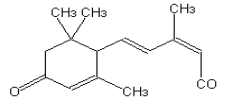


Fig. 4: Abscisic acid

The last compound belonging to the family of the "classical" phytohormones is ethylene, a simple non-saturated hydrocarbon that causes multiple answers in plants. In fact, its gas state confers the possibility to move through the intercellular spaces. It is biosynthesized from methionine and responds to stress (Davies, 1995).

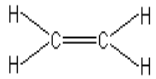


Fig. 5: Ethylen

Within the last twenty years, further plant growth substances have been described, as there are derivatives from brassinosteroide, polyamine, jasmonat (JA), salicylat (SA), and oligosaccharides. In general, these compounds are part of signal transduction cascades, which regulate the expression of essential genes for growth, development and plant defense (Aldington *et al.* 1991; Arteca 1995).

Evolution of auxin concept

A putative physiologicaly active substance, extracted from coleoptile tips was named auxin. Later, indole-3-acetic acid had become firmly established an auxin of higher plant tissues (Letham *et al.* 1978; Went and Thimann 1937). In the meantime, the number of putative auxins increased greatly up to hundreds of different chemical structures. Typically they represent small molecules with a common carboxyl group and usually a ring structure. Most are synthetic auxins, but only a few are naturally

occuring auxins, like indole-3-acetic acid, phenylacetic acid, and some of their derivatives.

The term "auxin" derived from the greek word "auxein" (= to increase) was introduced by Kölg and Haagen-Smit in 1931, when they isolated the compound Auxin-a (Kölg and Haagen-Smit 1931). That was the beginning of the first phase of evolution in the auxin concept characterized by the molecular structure definition described between 1930s and 1970s (Jönsson 1961; Katekar 1979; Koepfli *et al.* 1938; Went and Thimann 1937).

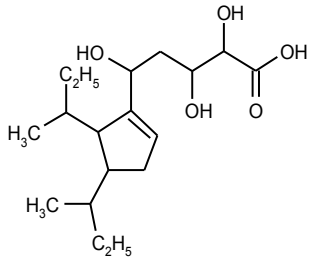


Fig. 6: Auxin-a

The main characteristic of this time period was the analysis of biological activities of many molecules, usually evaluated in many different types of "auxin tests". Went wrote in 1935: "the physiological name growth-substance and the chemical name auxin are interchangeable … of the different growth stages (initiation, differentiation, elongation, and maturation) elongation is the most spectacular and the one that can best be measured since it involves the greatest change in dimensions….", while root activity was not mentioned because of their exceptional behavior to auxin (Went 1935). Ten years later Went wrote in another review: "Chemical isolation and identification of indoleacetic acid from vascular plants has been accomplished. This makes it necessary to use the term auxin as a generic name for all substances, produced by plant as growth hormones or as correlation carriers, which gives response in the Avena test" (Went 1945). Additionally, Went proposed the common structural characteristics of the substances and indicated that the effect of auxin on the growing cell resemble a chemical reaction. Went adopted "auxin" as a chemical name without scientific evidences.

Following this idea, many dissimilar substances (molecules with different ring, without ring, with substituents in the rings or in the side chain, without side chain, etc.) were tested and the structural parameters for suitable biological activities were proposed, all under the name "auxin" (Fawcett *et al.* 1956; Harper and Wain 1969; Jönsson 1961; Koepfli *et al.* 1938; Porter and Thimann 1965).

Many definitions of "auxin" were suggested, such as:

- compounds that cause cell enlargement of plant cells (Nickell 1983);
- organic compounds, which promote growth (irreversible increase in volume) along the longitudinal axes, when applied in low concentrations to shoots of plants (Thimann 1948);
- plant growth regulator, natural or synthetic, are identified by certain operational terms, as auxin which by derivation cause an increase in size … (Thimann 1969),
- the generic name for a group of substances resembling the endogenous auxin molecule indole acetic acid (IAA) in action or in structure, and can be divided into several classes: the indol compounds, the phenoxi-acids compounds, the benzoic-acids compounds... (Leschem 1973);
- a compound that has a spectrum of biological activities similar to, but not necessarily identical with those of IAA. This include the ability to: 1) induce cell elongation in coleoptile or stem sections, 2) cell division in callus tissue join to cytokinin, 3) promote root formation to the cut surface... (Taiz and Zeiger 1998).

These varieties of definitions for auxins reflect the futility to establish a generally accepted definition. The definitions of auxins are usually based on physiological activities observed. Since the physiological activity of auxins is completly pleiotropic, it is impossible, to find a model on morphological level. This problem could not circumvent by a definition based on structural characteristics, since the structural distinctiveness or resemblance among auxins cannot be elucidate on their biological actions.

The first isolated causal molecules Auxin-a and Auxin-b are now predicted as a scientific fraud. The scientific scenario from 1930s did not allow for the estimation of such a molecular complexity and the experiments published by the Kögl lab were not well done (Wildman 1997). Not until the 1970s, IAA was conclusively identified in Picea, Pinus and at lest 18 angiosperms (Letham *et al.* 1978).

This was also the beginning of the second phase in auxin research, in which auxin was regarded as a signal transduction concept, similar to that of hormone action in animals. Ray (Ray *et al.* 1977) proposed the so-called first receptor candidate ABP1,

being the beginning of the molecular basis for auxin action. A great number of analyses were carried out on the following topics: auxin-perception, -transport machinery, -transport routes, -tunning genes and interactions with other hormones (overviews: (Berleth *et al.* 2004; Leyser 2002; Woodward and Bartel 2005)).

As a result of all these attempts, a modification of the perception of auxin from a generic concept as plant correlation carriers (Paál 1919) towards a chemical definition of a set of chemical structures, which is not consistent biologically, has occurred.

Auxin perception, signal transduction and gene expression

The biological activities of auxins are related to phenomenon such as biosynthesis and conjugation from evolutionary point of view. The evidences suggest that the apical regions of both charophytes and liverworts synthesize IAA via a tryptophan-independent pathway, with IAA levels being regulated, balanced by the rates of IAA biosynthesis and IAA degradation. Other terrestrial plants utilize the same class of biosynthetic pathway, but they have the additional potential to utilize IAA conjugation and conjugate hydrolysis reactions to achieve more precise spatial and temporal control of IAA levels (Cooke *et al.* 2002). That can even be considered being important for plant symbiont interaction (Grubb *et al.* 2004; Ludwig-Muller 2004). However, the option for conjugation or hydrolysis, respectively, is a general biological mechanism used to maintain internal equilibrium and adjusting its physiological processes.

The plant hormone's way of action shows some uniqueness. On one hand, the term sensitivity, introduces the availability of the receptor as a new factor (Trewavas 1982; Trewavas and Cleland 1983; Weyers *et al.* 1987) while on the other hand, the underlying concept of the structure-activity rule, stating that auxins act as a kind of co-enzyme or ergon at the growth center, which is a protein or enzyme surface of highly specific "shape", is not consistent anymore (Audus 1961).

Auxin perception is characterized by different auxin-binding sites and proteins described by different groups (Jacobsen 1984; Ray *et al.* 1977; Reinard *et al.* 1998). 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 (overwiew: (Napier *et al.* 2002)). Furthermore, the considerable speculation about specialized receptor functions for specific transporters like PIN (PIN-FORMED) proteins should be considered as well (Blakeslee *et al.* 2004; Friml *et al.* 2002; Geldner *et al.* 2001). Recently a new complex of three proteins SCF^{TIR1} out of which the transport inhibitor response 1 (TIR1) has been described as an auxin receptor (Dharmasiri *et al.* 2005; Kepinski and Leyser 2005).

The perception of the auxin molecules follows a signal transduction system to "inform" the different cell components. Furthermore auxin signal transduction is part of a multihormonal response network as well. Up to now, the only knockout plant for the one Gα identified in *Arabidopsis* showed a decreased cell division as a major trait in its phenotype which is part of the function of auxin – but auxin is not the only player in cell division. G protein subunits perhaps trigger a multi-signal of the cell cycle affected by auxin, and other hormones (Scherer 2002). The two-component systems in higher plants address several critical points with respect to cross talk, signal integration and specificity (Grefen and Harter 2004).

Considering auxin signal perception in plants, further discriminations have to be obeyed. The pH may exhibit a fundamental role, and also the response time. Usually, long term effects, like morphogensis or gene regulation use distinctive modes of action than fast auxin effects, which usually occur within minutes (are believed to occur on membrane bound receptors). Whereas slow auxin effects, like altered gene expression profiles occur (possibly via a signal transduction cascade) in the cell nuclei. Another group of auxin-interacting proteins represent membrane proteins have been convincingly implicated in auxin influx and efflux (Berleth *et al.* 2004; Okushima *et al.* 2005).

Plant hormone receptors have proven to be elusive research targets. The successes of describing receptors from animals and bacteria have not yet been matched for plants. ABP1 is still the most consequent candidate up to now. It could also be the first biological receptor with the major part localized in the ER (Jones 1994; Woo *et al.*

2002). However, the evaluation of a system in relation to the behaviour that might be expected as a *bona fide* receptor, does not say much in favor of an auxin receptor. Some criteria are not fulfilled by phytohormones (Venis 1985)and especially auxins, such as:-

- Binding specificity for different hormones analogues should be approximately in accordance with the relative biological activities of the compounds. Unfortunately, the inactive auxin 2-NAA shows the best binding affinity to ABP1 (Edgerton *et al.* 1994);
- Binding should lead to a hormone-specific, biological response. First of all, the pleiotropic responses for auxin-like molecules (Weyers and Paterson 2001). An ABP1-independent pathway was described recently, which is much more sensitive to IAA than the ABP1-dependent one (Yamagami *et al.* 2004);
- Binding may be limited to hormone-responsive tissue. But in plants this is a difficult term to define and still this is one of the most discussed issues in phytohormone research since the 1980s (Trewavas 1982; Weyers and Paterson 2001; Weyers *et al.* 1987).

Structure-activity

Auxin has become an indescribable biological phenomenon characterized by the parallel comprehension of the chemical and biological view-points. The assumption "structure generates properties" has been evaluated as a dynamic regularity of the hormone-receptor interaction exported from the animal model. Following the idea of "one receptor-one ligand" and based on bioassays (Steward and Krikorian 1971) have been developed some speculative concepts about the auxin molecular properties.

The first important attempt to formulate general rules for molecules exhibiting auxin activity was already formulated at the end of 1930s (Koepfli *et al.* 1938; Went and Thimann 1937). They stated requirements for molecules with a high auxin activity as follows:

- a ring system as a nucleus;
- at lest one double bond in the ring system;

- a side chain containing a carboxyl group with at least one atom removed from the ring;
- a particular space relationship between the carboxyl group and the ring.

They did not postulate anything on the physiological impact of these minimal requirements. Obviously, the rules are not compatible with the activities of certain naphthoic, benzoic acids and phenol derivatives, described as auxins later (Harper and Wain 1969; Jönsson 1961).

After the initial formulation of structural requirements for auxin molecules, two different proceedings can be distinguished, mainly issued in the 1950s: the chemical and the physico-chemical approach.

Chemical approach:

The theory of Hansch and Muir is related with the ortho-effect phenomenon (Hansch and Muir 1950), which deals with the Two Point Attachment theory. Here, a bond formation between the active site and an aromatic ring should occur. More detailed information on this type of chemical reaction between growth regulators and a plant substrate was published one year later: a reaction, in which the release of chloride ion is essentially connected with the physiological activity of the compound (Hansch *et al.* 1951). The analysis at this time was focused on benzoic derivatives. Muir *et al.* hypothized in 1967 that the position of attachment on the ring would depend on the particular combination of steric and electronic factors (Muir *et al.* 1967). Later, it was shown that the indole-3-acetic acid interaction with a receptor site may be non-covalent in nature (Katekar 1979).

The second purely chemical approach postulated in order to fit an auxin molecule into an attachment site was the Separation Charge Theory (Thimann and Leopold 1955). This has become one of the most known theories accepted in many text books even in the XXI century. Thimann himself found biological activities correlated to the N-H in the indole ring. These correlations were attributed to the charge of the nitrogen. But the low activity of 5,7-dichloroindole-3-acetic acid was a serious deviation, which could not be explained with this theory (Katekar 1979; Porter and

Thieman 1965). Jönnson, who analyzed the structure-activity relationship of more than 600 auxin molecules disregarded the Separation Charge Theory (Jönsson 1961). Further analysis using self-consistent field molecular orbital (SCF-MO) calculations did not support certain details of this theory. In fact for both natural auxin IAA and 2,4-D, the site regarded as carrying a positive charge was shown to exhibit a net negative charge (Farrimond *et al.* 1980; Farrimond *et al.* 1981). Recent calculations at ab-initio levels confirm that the position of the N in the pyrrole ring makes the indole more aromatic than its isomers and a substituent at position 3 does not change significantly the aromaticity properties of the indole system (Kiralj and Ferreira 2003).

Physico-chemical approach

Veldstra suggested that the action of an auxin consists in a "physico-chemical influencing of a boundary". The requirements were condensed into two points:

- a basal ring system with a high surface activity
- a carboxyl group in a very definite spatial position with respect to this ring system (Veldstra 1944).

Later, due to the increase of activity by chlorination of the phenoxyacetic acids, he had to postulate that a high surface activity in the ring system was not sufficient for the auxin action. A certain balance between the lipophile and the hydrophile part of the molecule was assumed to be essential (Jönsson 1961).

Veldstra made exhaustive analyses on the Two Point Attachment theory and he concluded that hydrophilic substituents (OH, NH_2) do not confer activity of the resulting derivates, but only lipophilic ones (Cl, Br, I, CH_3). A chemical attachment implies that a physiological response will occur, once the molecule is irreversibly fixed to the receptor (Veldstra 1953). Velstrad´s theory assumed that the auxin molecule is not bound by strong chemical bounds at the site of action but is loosely and reversibly attached by many weak bonds (hydrogen bridges, electrostatic attractions, van der Waals forces).

The Three Point Attachment theory attempt to explain certain phenomena, which were found to be inconsistent with Velstrad's theory. That theory bases on the findings that several compounds are essential for activity:

- a flat ring system,
- a hydrogen atom α to the carboxyl group,
- a special configuration of the side chain with respect to the ring, and
- the free rotation of the side chain at the bond joining to the ring seems to be structurally required for its activity (Fawcett *et al.* 1955; Fawcett *et al.* 1956; Wain and Wightman 1953).

It avoided the use of charge separation and the mechanism suggested two hydrophobic areas, either of which could complement aromatic ring systems, and a single positively charged site to accommodate the carboxylate group (Napier 2001). This theory did not consider the benzoic acids.

Binding site models

In the 1970s, biochemical based models were developed to elucidate the relationship between structure and activity. The first model using a binding site proposal was carried out by Kaethner (Kaethner 1977). The Conformational Change theory for auxin is far away from any rigid hypothesis like the Charge Separation theory above. Furthermore, the recognition conformation proposal coincides with the active auxin form suggested by Jönnson (Jönsson 1961) and the modulation conformation is equivalent to the form proposed by Velsdtra (Veldstra 1944). Additionally, Kaethner proposed a "floor" of the receptor site as responsible region for the hydrogen-bonding with the pyrrole nitrogen region of IAA, which is consistent with Thimann's theory (Porter and Thimann 1965). On one hand Thimann based his theory on the positive charge of N (analyzed above), which is not true and on the other hand, Kaethner´s theory was not proved experimentally.

The binding site model of Katekar (Katekar 1979), frequently considered as the first binding site model (Napier 2001), was a result of an intuitive analysis supported by a methodical examination of data accumulated by other authors. Katekar provided a

comprehensive biological focal point breaking the realistic view: "it is so far too early to predict how these findings (from the structure-activity analysis of more than 600 molecules) will influence the structure activity discussions" (Jönsson 1961). Katekar used just about 20 percent of the data from Jönson.

The definition of Katekar´s auxin receptor site is *ex hypothesi* complementary to the IAA molecule. Subsequently, the wide diversity of the remaining synthetic and natural auxin-like molecules were superimposed (Katekar 1979). This is inconsistent with the flexible proposal for all molecules since this proposal just relies on the deterministic conception of an IAA receptor. Katekar strengthened this inconsistence during his further analysis of IAA derivates for validating the theory (Katekar and Geissler 1982; Katekar and Geissler 1983; Katekar *et al.* 1987).

Kaethner and Katekar consequently introduced the pharmacophore concept in auxin related research. Napier mentioned that, if these pharmacophoric models were to be proven useful, they needed to be applied either to assist in discovery of novel ligands or for testing the structure of auxin-binding proteins (Napier 2001).

The general molecular requirements for auxins were confirmed by using results from binding assay with ABP1 (Edgerton *et al.* 1994). The modelling of the ABP1 suggests a conformational change of the ligand to achieve a binding site, which incorporates a metal ion (Warwicker 2001). This metal ion was confirmed experimentally, but the conformational change has not been confirmed up to now (Woo *et al.* 2002).

The growing capacity for computational chemistry permits novel variables (Tomic *et al.* 1998). Unfortunatelly, the model was reduced to a chemical point of view and the authors used a classical (animal-related) concept of hormone action. The use of Molecular Quantum Similarity Measures (MQSM) and the *LogP* and *LogD* indices were announced as a new method to predict biological activity within a set of about 100 compounds (Bertosa *et al.* 2003). Unfortunatelly, this issue is very contradictory to regular concepts in auxin research, mentioned above. The electronic and lipophilic effect of substituents on the ring cannot be assessed with some degree of reliability related with promotion of activity (Muir *et al.* 1967). Other concepts, mainly on the lipophilic character of certain substituents being a determinant factor for auxin activity is not substantiated, since auxin activity does not increase with the

increasing lipophilic character, and molecules of similar lipid solubility have very different auxin activities (Farrimond *et al.* 1981; Porter and Thimann 1965; Veldstra 1944). The existence of auxin carriers (Blakeslee *et al.* 2004; Friml *et al.* 2002) in the plasma membrane, even with directional implications, is another, very complex factor perturbing the influence of the lipophilic variables.

The pharmacophore identification problem is complicated substantially by the fact that ligands are very flexible molecules. Usually, ligands own many internal degrees of freedom. Each conformation may bind in the active site of the considered receptor (Dror *et al.* 2004). Most of the new approaches are focused on ABP1 as a supposed receptor (Bertosa *et al.* 2003; Kiralj and Ferreira 2005), but its high affinity for auxin molecules is only one additional variable, which has to be considered. The existence of more than this possible auxin receptor is already confirmed and widely accepted (Blakeslee *et al.* 2004; Dharmasiri *et al.* 2005; Kepinski and Leyser 2005; Napier *et al.* 2002; Ray *et al.* 1977). Therefore, the pharmacophore concept (the mapping of common structural features of active analogs that bind to the same receptor (Buehler 2003) can not overcome the structure-activity impasse of auxins.

To overcome these limitations of the auxin structure-activity concept, a new computational-biostatistical approach was developed, that focuses on the auxin chemical space in the biological context. The pleiotropic effects of plant hormones is a statistical regularity associated with the multi-receptor and signal transduction systems. Therefore the analysis of the structural consensus of the auxin-like molecules is treated as the invariant part, from which the phytohormone phenomena is statistically relative. That does not mean, that the phytohormone phenomenon depends exclusively on the ligand structure, but the ligand structure analysis is the point to define the degrees of freedom of the phenomenon.

The analyses presented in this work focus on the following main objectives:

1. to define a flexible methodology for analyzing dependences between structure and biological activities;
2. to search the region of Molecular Quantum Similarities Measures associate with the biological activities;

3. to classify auxin molecules based at the boundary Molecular Similarity - Biological Activity;
4. the development of parallel bioassay screenings of selected molecules to confirm the hypothesis of similarity;
5. the quantitative structure-activity relationships with fresh biological data.

Literature

Aldington S., Mcdougall G.J., Fry S.C. (1991) Structure-Activity-Relationships of Biologically-Active Oligosaccharides. Plant Cell and Environment 14: 625-636.

Arteca R.N. (1995) Plant Growth Substances: Principles and Applications. Chapman & Hall. New York, 311 p.

Audus L.J. (1961) Metabolism and mode of action. In: Ruhland W (ed) Encyclopediea of Plant Physiology. Springer, Berlin.

Berleth T., Krogan N.T., Scarpella E. (2004) Auxin signals - turning genes on and turning cells around. Current Opinion in Plant Biology 7: 553-563.

Bertosa B., Kojic-Prodic B., Wade R.C., Ramek M., Piperaki S., Tsantili-Kakoulidou A., Tomic S. (2003) A new approach to predict the biological activity of molecules based on similarity of their interaction fields and the logP and logD values: Application to auxins. Journal of Chemical Information and Computer Sciences 43: 1532-1541.

Blakeslee J.J., Bandyopadhyay A., Peer W.A., Makam S.N., Murphy A.S. (2004) Relocalization of the PIN1 auxin efflux facilitator plays a role in phototropic responses. Plant Physiology 134: 28-31.

Buehler L.K. (2003) What's in a Structure? PharmaGenomics 3: 20-21.

Cooke T.J., Poli D., Sztein A.E., Cohen J.D. (2002) Evolutionary patterns in auxin action. Plant Molecular Biology 49: 319-338.

Dharmasiri N., Dharmasiri S., Estelle M. (2005) The F-box protein TIR1 is an auxin receptor. Nature 435: 441-445.

Dror O., Shulman-Peleg A., Nussinov R., Wolfson H.J. (2004) Predicting molecular interactions in silico: I. A guide to pharmacophore identification and its applications to drug design. Current Medicinal Chemistry 11: 71-90.

Edgerton M.D., Tropsha A., Jones A.M. (1994) Modeling the Auxin-Binding Site of Auxin-Binding Protein-1 of Maize. Phytochemistry 35: 1111-1123.

Farrimond J.A., Elliott M.C., Clack D.W. (1980) Auxin Structure-Activity-Relationships - Benzoic-Acids and Phenols. Phytochemistry 19: 367-371.

Farrimond J.A., Elliott M.C., Clack D.W. (1981) Auxin Structure-Activity-Relationships - Aryloxyacetic Acids. Phytochemistry 20: 1185-1190.

Fawcett C.H., Wain R.L., Wightman F. (1955) Studies on plant growth-regulating substances. VIII. The growth promoting activity of certain aryloxy- and aryl-thyo-alkanecarboxilic acids. Ann. appl. Biol. 43: 342-354.

Fawcett C.H., Wain R.L., Wightman F. (1956) Plant Growth-Regulating Activity in Certain Carboxylic Acids Not Possessing a Ring Structure. Nature 178: 972-974.

Friml J., Wisniewska J., Benkova E., Mendgen K., Palme K. (2002) Lateral relocation of auxin efflux regulator PIN3 mediates tropism in Arabidopsis. Nature 415: 806-809.

Geldner N., Friml J., Stierhof Y.D., Jürgens G., Palme K. (2001) Auxin transport inhibitors block PIN1 cycling and vesicle traficking. Nature 413: 425-428.

Grefen C., Harter K. (2004) Plant two-component systems: principles, functions, complexity and cross talk. Planta 219: 733-742.

Grubb C.D., Zipp B.J., Ludwig-Müller J., Masuno M.N., Molinski T.F., Abel S. (2004) Arabidopsis glucosyltransferase UGT74B1 functions in glucosinolate biosynthesis and auxin homeostasis. Plant Journal 40: 893-908.

Hansch C., Muir R.M. (1950) The Ortho Effect in Plant Growth-Regulators. Plant Physiology 25: 389-393.

Hansch C., Muir R.M., Metzenberg R.L. (1951) Further Evidence for a Chemical Reaction between Plant Growth-Regulators and a Plant Substrate. Plant Physiology 26: 812-821.

Harper D.B., Wain R.L. (1969) Studies on plant growth - regulating substances. XXX. The plant growth - regulating activity of substituted phenols. Ann. appl. Biol. 64: 395-407.

Haubrick L.L., Assmann S.M. (2006) Brassinosteroids and plant function: some clues, more puzzles. Plant Cell and Environment 29: 446-457.

Jacobsen H.-J. (1984) 2 Different Soluble Cytoplasmic Auxin-Binding Sites in Etiolated Pea Epicotyls. Plant and Cell Physiology 25: 867-873.

Jones A.M. (1994) Auxin-binding proteins. Annual Review of Plant Physiology and Plant Molecular Biology 45: 393-420.

Jönsson A. (1961) Chemical structure and growth activity of auxins and antiauxins. In: Ruhland W (ed) Encyclopediea of Plant Physiology. Springer, Berlin, pp 959-1006.

Joo G.J., Kim Y.M., Kim J.T., Rhee I.K., Kim J.H., Lee I.J. (2005) Gibberellins-producing rhizobacteria increase endogenous gibberellins content and promote growth of red peppers. Journal of Microbiology 43: 510-515.

Kaethner T.M. (1977) Conformational Change Theory for Auxin Structure-Activity-Relationships. Nature 267: 19-23.

Katekar G.F. (1979) Auxins - Nature of the Receptor-Site and Molecular Requirements for Auxin Activity. Phytochemistry 18: 223-233.

Katekar G.F., Geissler A.E. (1982) Auxins II: the effect of chlorinated indolylacetic acids on pea stems. Phytochemistry 21: 257-260.

Katekar G.F., Geissler A.E. (1983) Structure-activity differences between indoleacetic acid auxins on pea and wheat. Phytochemistry 22: 27-31.

Katekar G.F., Winkler D.A., Geissler A.E. (1987) Hormone recognition in plants. In: Klämbt D (ed) Plant hormone receptor. Springer-Verlag, Berlin, pp 13-26.

Kepinski S., Leyser O. (2005) The Arabidopsis F-box protein TIR1 is an auxin receptor. Nature 435: 446-451.

Kiralj R., Ferreira M.M.C. (2003) Combined computational and chemometric study of ^{1}H-indole-3-acetic acid. International Journal of Quantum Chemistry 95: 237-251.

Kiralj R., Ferreira M.M.C. (2005) Chemometric and molecular modeling study of 1H-indole-3-acetic acid derivatives with auxin activity. Croatica Chemica Acta 78: 541-549.

Koepfli J.B., Thimann K.V., Went F.W. (1938) Phytohormones: structure and physiological activity. I. J. Biol. Chem. 122: 763-780.

Kölg F., Haagen-Smit A.J. (1931) Über die Chemie des Wuchsstoffs. Proc. Kon. Akad. Wetensch. Amsterdam 34: 1411-1416.

Kulaeva O.N., Prokoptseva O.S. (2004) Recent advances in the study of mechanisms of action of phytohormones. Biochemistry-Moscow 69: 233-247.

Leschem Y. (1973) The molecular and hormonal bases of plant-growth regulation. Pergamon Press, Oxford, 156 p.

Letham D.S., Goodwin P.B., Higgins T.J.V. (1978) The biochemistry of phytohormones and related compounds. Elsevier / North-Holland Biomedical Press, Amsterdam.

Leyser O. (2002) Molecular genetics of auxin signaling. Annual Review of Plant Biology 53: 377-398.

Ludwig-Müller J. (2004) From auxin homeostasis to understanding plant pathogen and plant symbiont interaction: Editor's research interests. Journal of Plant Growth Regulation 23: 1-8.

Muir R.M., Fujita T., Hansch C. (1967) Structure-Activity Relationship in Auxin Activity of Mono-Substituted Phenylacetic Acids. Plant Physiology 42: 1519-&.

Napier R.M. (2001) Models of auxin binding. Journal of Plant Growth Regulation 20: 244-254.

Napier R.M., David K.M., Perrot-Rechenmann C. (2002) A short history of auxin-binding proteins. Plant Molecular Biology 49: 339-348.

Nickell L.G. (1983) Plant Growth Regulators: Agricultural Uses. Springer Verlag, New York, 173 p.

Okushima Y., Overvoorde P.J., Arima K., Alonso J.M., Chan A., Chang C., Ecker J.R., Hughes B., Lui A., Nguyen D., Onodera C., Quach H., Smith A., Yu G.X., Theologis A. (2005) Functional genomic analysis of the AUXIN RESPONSE FACTOR gene family members in Arabidopsis thaliana: Unique and overlapping functions of ARF7 and ARF19. Plant Cell 17: 444-463.

Paál A. (1919) Über Phototropische Reizleitung. Jahrbuch für Wissenschaften Botanik: 406–458.

Porter W.L., Thimman K.V. (1965) Molecular requeriments for auxin action-I. Halogenated indoles and Indoleacetic acid. Phytochemistry 4: 229-243.

Ray P.M., Dohrmann U., Hertel R. (1977) Specificity of Auxin-Binding Sites on Maize Coleoptile Membranes as Possible Receptor-Sites for Auxin Action. Plant Physiology 60: 585-591.

Reinard T., Achmus H., Walther A., Rescher U., Klämbt D., Jacobsen H.J. (1998) Assignment of the auxin binding abilities of ABP(44) in gel. Plant and Cell Physiology 39: 874-878.

Sasson A. (1993) La alimentación del hombre del mañana. Unesco/Editorial Reverté, S.A., Barcelona, 807 p.

Scherer G.F. (2002) Secondary messengers and phospholipase A(2) in auxin signal transduction. Plant Molecular Biology 49: 357-372.

Steward F.C., Krikorian A.D. (1971) Plant, chemicals and growth. Academic Press, New York, 232 p.

Taiz L., Zeiger E. (1998) Plant physiology. Sinauer, 486 p.

Thimann K.V. (1948) Plant growth hormones. In: Thimann K.V. (ed) The hormones. Academic Press, New York, pp 5-119.

Thimann K.V. (1969) The auxins. In: Wilkins M.B. (ed) The physiology of plant growth and development. McGraw-Hill, New York, pp 3-45.

Thimann K.V., Leopold A.C. (1955) In: Pincus G., Thimann K.V. 'eds) The Hormones. Academic Press, New York, pp 1-56.

Tomic S., Gabdoulline R.R., Kojic-Prodic B., Wade R.C. (1998) Classification of auxin plant hormones by interaction property similarity indices. Journal of Computer-Aided Molecular Design 12: 63-79.

Trewavas A.J. (1982) Growth Substance Sensitivity - the Limiting Factor in Plant Development. Physiologia Plantarum 55: 60-72.

Trewavas A.J., Cleland R.E. (1983) Is Plant Development Regulated by Changes in the Concentration of Growth-Substances or by Changes in the Sensitivity to Growth-Substances. Trends in Biochemical Sciences 8: 354-357.

Veldstra H. (1944) Researches on Plant Growth Substances IV. Relation between chemical structure and physiological activity I. Enzymologia 11: 97-136.

Veldstra H. (1953) The Relation of Chemical Structure to Biological Activity in Growth Substances. Annual Review of Plant Physiology and Plant Molecular Biology 4: 151-178.

Venis M.A. (1985) Hormone binding sites in plants. Longman, New York London.

Verhey S.D., Lomax T.L. (1993) Signal-Transduction in Vascular Plants. Journal of Plant Growth Regulation 12: 179-195.

Wain R.L., Wightman F. (1953) Studies on Plant Growth-Regulating Substances.7. Growth-Promoting Activity in the Chlorophenoxyacetic Acids. Annals of Applied Biology 40: 244-&.

Warwicker J. (2001) Modelling of auxin-binding protein 1 suggests that its C-terminus and auxin could compete for a binding site that incorporates a metal ion and tryptophan residue 44. Planta 212: 343-347.

Went F.W. (1935) Auxin, the Plant-Growth Hormone. Botanical Review 1: 162-181.

Went F.W. (1945) Auxin, the Plant-Growth Hormone.2. Botanical Review 11: 487-496.

Went F.W., Thimann K.V. (1937) Phytohormones. Macmillan Company, New York.

Weyers J.D.B., Paterson N.W. (2001) Plant hormones and the control of physiological processes. New Phytologist 152: 375-407.

Weyers J.D.B., Paterson N.W., Abrook R. (1987) Towards a Quantitative Definition of Plant Hormone Sensitivity. Plant Cell and Environment 10: 1-10.

Wildman S.G. (1997) The auxin-A, B enigma: Scientific fraud or scientific ineptitude? Plant Growth Regulation 22: 37-68.

Woo E.J., Marshall J., Bauly J., Chen J.G., Venis M., Napier R.M., Pickersgill R.W. (2002) Crystal structure of auxin-binding protein 1 in complex with auxin. Embo Journal 21: 2877-2885.

Woodward A.W., Bartel B. (2005) A receptor for auxin. Plant Cell 17: 2425-2429.

Yamagami M., Haga K., Napier R.M., Iino M. (2004) Two distinct signaling pathways participate in auxin-induced swelling of pea epidermal protoplasts. Plant Physiol 134: 735-47.

Chapter 2

Structure–activity analysis on ecdysteroids: A structural and quantum chemical approach based on two biological systems

Noel Ferro, Juan E. Tacoronte, Thomas Reinard, Patrick Bultinck, Luis A. Montero

Journal of Molecular Structure: THEOCHEM **758** *(2006) 263–274*

ELSEVIER

Journal of Molecular Structure: THEOCHEM 758 (2006) 263–274

THEO CHEM

www.elsevier.com/locate/theochem

Structure–activity analysis on ecdysteroids: A structural and quantum chemical approach based on two biological systems

Noel Ferro [a,b,*], Juan E. Tacoronte [c], Thomas Reinard [b], Patrick Bultinck [d], Luis A. Montero [e]

[a] *Department of Agronomy, Forestry and Agronomy Faculty, University of Pinar del Río, Martí 270, Pinar del Río 20 100, Cuba*
[b] *LG Molekulargenetik, University of Hannover, Hannover, Germany*
[c] *Faculty of Chemistry, Center of Studies on Natural Products (C.E.P.N), University of Havana, Zapata e/Mazón and G. Vedado, Plaza, Ciudad Habana, Cuba*
[d] *Department of Inorganic and Physical Chemistry, Ghent University, Krijgslaan 281 (S3), 9000 Ghent, Belgium*
[e] *Laboratory of Computational and Theoretical Chemistry, (LQCT), Faculty of Chemistry, University of Havana, Zapata e/Mazón and G. Vedado, Plaza, Ciudad Habana, Cuba*

Received 16 September 2005; received in revised form 16 September 2005; accepted 5 October 2005
Available online 27 December 2005

Abstract

Besides their central role in the development of insects, ecdysteroids are widely found in other organisms as well. It is well established that ecdysteroids occur in various molecular forms as well as are valuable targets to identify structural requirements for the development of insecticides in favour of plant protection. Therefore, very advanced 3D- and 4D-QSAR have been applied to ecdysteroids. Our work, at present, is carried out by the arrangement use of quantum chemistry at semi-empirical level, molecular similarity measures and bio-statistical analysis. This strategy proficient gets into details of structure–activity of different ecdysteroid analogs. On the other hand, due to the analysis of a hormonal factor as multicellular scheme, in addition to the EC_{50} analysis in *Drosophila melanogaster*, the *Calliphora* test was taking into account. The influence of functional groups and different molecular properties (a total of 778) were cause-effect related to a set of 96 analogs for both tests. Additional theoretical analysis of hydrogen bonds and molecular orbitals were done as well. The work let to the achievement of a more realistic assessment in relation to the structure–activity, and it confirm that, geometrically, not all functional groups are important for bioactivity, only those whose contributions are involved in a quasi band of outer molecular orbitals. The discrimination of the most active molecules was done better by the use of the autosimilarity diagonal together with some other quantum and geometric variables.

Keywords: Ecdysteroids; Structure–activity; Plant protection; Biostatistical analysis; Molecular orbital

1. Introduction

Ecdysteroids by far are some of the most important steroid hormones in the biosphere in terms of quantity and diversity. Ecdysteroids are compounds related to ecdysone, a single nuclear hormone that can control differentiation, program cell death and proliferation in different tissues [1,2]. The so-called 'moulting and metamorphosis hormone' is produced by the prothoracic glands after brain activation during insect development. This exerts morphogenesis changes during gene activation, although another group of juvenile hormones controls the events of morphogenesis [3,4].

The ecdysteroids mediate their biological effects by either direct activation of gene transcription after binding to its receptor EcR-Usp or via hierarchical transcriptional regulations of several primary transcription factors [2,5]. The receptor of the ecdysteroids has a high affinity ($K_d = 30$ nM) and specificity with ecdysteroids. Following the classic steroids action mechanism of binding of the ecdysone receptor (EcR) to a ligand-nuclear inducible transcription factor, it must form a heterodimer with ultraspiracle (Usp), the homologue of retinoid-X receptor. The crystal structures of ligand-binding domains EcR-Usp heterodimer, in complex with ponasterone A, emphasizes the universality of heterodimerization as a general mechanism common to both vertebrates and invertebrates [4,6].

However, the fine structure of the molecular action mechanism of ecdysteroids, interaction ecdysteroid–protein, has not been explained until now. The hormone-receptor model of three interaction sites [4] is still unclear, essentially due to

* Corresponding author. Address: University of Hannover, Herrenhäuser Str. 2, D-30419 Hannover, Germany. Tel.: +49 511 762 4285.
E-mail address: nferro71@yahoo.es (N. Ferro).

0166-1280/$ - see front matter © 2005 Elsevier B.V. All rights reserved.
doi:10.1016/j.theochem.2005.10.027

the fact that one of these three sites (14α-OH) is lacking in some active molecules [7,8].

Here, we focus on the following objectives: (1) to determine quantitatively which key atoms or regions in the molecule trigger the action of the biological machinery, starting from the structure–activity relationship and the analysis of electronic properties; (2) to consider a cause-effect relationship between chemical features and biological activity of a series of ecdysteroid analogues by means of a mathematical–statistic approach by means of the uses of two biological approach.

2. Methods and techniques

2.1. Molecular modelling

First of all the three-dimensional molecular structure of the ecdysone molecule (Fig. 1) was built in Hyperchem (Hypercube). The geometry was optimized by using the MM+ [9] force field, including electrostatic bond dipoles and considering all non-bonded interactions. The quantum chemical calculations were made at the semi-empirical level using PM3 [10] with the MOPAC v. 6 program [11]. The ZINDO/S method (modified version of INDO) [12] was used to visualize the molecular orbitals using Hyperchem.

After the geometry optimisation of ecdysone, each analogue that was chosen, up to 96 molecules in total, was optimised once again in the same way. Different molecular variables were calculated, such as: (a) ionisation potential; (b) the energy of frontier orbitals (HOMO), and even that of contiguous orbitals from HOMO-1 to -5; (c) the atoms more involved in these molecular orbitals; (d) the oxygen atoms with contribution to the HOMO-3 and HOMO-4 orbitals; (e) the spatial distribution of HOMO, and HOMO-3 orbitals; (f) the distances between intra-molecular atoms, bond lengths, valence angles and dihedral angles (τ) among all the atoms of the ring in the molecules; (g) charge (defined by default in MOPAC, Cannolly surface) of each atom. Additional information such as position, quantity and kind of functional groups, were directly taken from each molecule. At the end, a total of 778 variables per molecule were calculated.

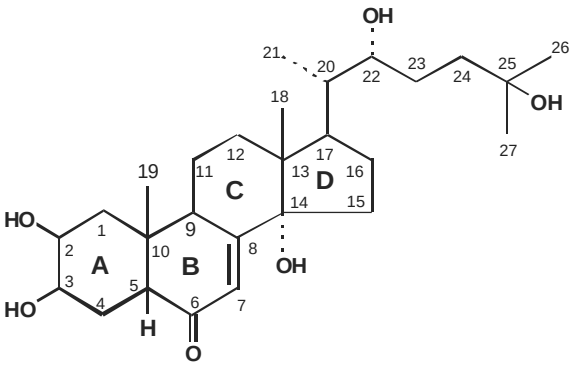


Fig. 1. Structural representation of the ecdysone rings.

2.2. Molecular similarity

Additional analysis of molecular quantum similarity measure (MQSM), applied to the molecules with *Callifora* activity and *Drosophila melanogaster* EC_{50} test, expressed as the integral of the scalar product between the first-order molecular density functions associated to the molecules being compared, and weighted by a positive definite two-electron operator (Ω) [13,14]. In this study, *Overlap-like* and *Coulomb-like* MQSM have been considered. In order to circumvent expensive computational calculations, the promolecular atomic shell approximation (ASA) [15–17] has been used to compute density functions. In order to align the molecular structures, the *maximum similarity superposition algorithm* [18] has been used.

Once computed, the overall set of pairwise MQSM can be stored in the so-called similarity matrix (SM): $Z=\{Z_{AB}\}$, where 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 compare the molecule with itself.

2.3. Statistical analysis

A statistical analysis of the ecdysone molecule and its analogues was performed using the calculated variables from both the geometry and quantum chemical approach. The energies data of ten outer molecular orbitals were analysed through cluster analysis for 100 different molecules. In case of confirmatory analyses like the influence of OH22 on the biological activity, the *t*-test was applied. PCA was also a valuable statistical method to make a mixture of the similarity Coulomb and Overlap matrix in principal components and reduce the repetitive information of the similarity matrix. The main task, then, was to discriminate the molecular variables involved in the biological activity by discrimination analysis with Wilks' lambda method. The computational software or programs used were R v. 2.1.1 and SPSS v. 12 (SPSS GmbH).

2.4. Hydrogen bond analysis

Hydrogen bonds can be formed between hydrogen atoms attached to electronegative atoms and lone pairs, especially on nitrogen and oxygen [19]. 'Non-bonding orbitals' (NBO) could be involved in the H-bonding between the ecdysteroid and receptor through OH2 and OH22. After the optimization, a lone pair is considered by more than 20% of contribution to the eigenvector of an outer molecular orbital. Subsequently, the interaction with a molecule of water near to the lone pair oxygen-orbital was analysed by means of a new optimization using geometric operations in Cartesian coordinates. This process allowed provoking a destabilization of the lone pair oxygen-orbital [20]. This was analysed by the semi-empirical methods PM3, AM1 and MNDO.

Geometrical analysis of probable H-bonds for OH2, OH3, O6, OH14 and OH22 were performed with PM3. Further analysis of the electrostatic potential (MEP) was performed

with Hyperchem. An ab initio single point calculation (DFT) with the 6-31G* basis set was performed. A map function range of $5.0 \times 10^{-2} e/a_0^3$ and the total charge density contour value $1.0 \times 10^{-3} e/a_0^3$ were used in plotting MEP.

3. Results

The biological machinery at both cellular and multicellular level mediates the influence of the molecular structure on bioactivity. The application of two independent tests resulted in models that were more realistic. The plot between structures and effect (*Calliphora* activity or EC_{50}) and its frequencies (Fig. 2) shows the irregular character of the relation structure function.

Analysis of the presence/absence of functional groups (OH), in *Calliphora*, was the opening step in this statistical analysis. Three categories were used: [Boil. activity $>, =, <1$] in which 1 is the activity of the naturally occurring ecdysone (Fig. 2) [21]. According to these analyses, the following functional groups are involved in discrimination: OHs in position C22, C5 and C14, as well as C5 and C14 together (Fig. 1). Looking at the discriminants cross validation (Tables 1 and 2), it is found that (a) the first group [Biol. Activity >1] is not statistically suitable due to the existence of solely one molecule. Although, the extracted variable OH groups in C5 and C14 together is important from the biochemical point of view. (b) The second [Biol. Activity $=1$] group has a prediction of 89.9% of membership; and (c) in the third group [Biol. Activity <1], not more than 54.5% are predicted. Furthermore, confirmatory statistical analysis (*t*-test) showed that OH in C22 position is significantly associated with biological activity (Fig. 3, Table 3). However, at this stage of our work the structure–activity relationship could not be totally explained by the presence/absence of functional groups itself. In the structure–activity relationship of ecdysteroids functional groups seem to be necessary but not sufficient condition.

Table 1
Discriminant functions

	Function	
	1	2
OH5_OH14	3.942	−2.171
OH_22	1.239	2.250
(Constant)	−6.442	−1.480

Then a functional analysis using geometric and quantum chemical variables was integrated through statistic–mathematical methods. As result, some variables were able to discriminate the biological differences.

3.1. Calliphora test

- Product of the contributions of atom O22 to the molecular orbital HOMO-3 and the atom O3 to molecular orbital HOMO-4 (O_{22},HOMO-3*O_3HOMO-4): these are orbitals with similar disposition on the ecdysteroids molecules from the ring A to the side chain, related to the negative electrostatic potential zones on the molecules and without statistical differences from energetic point of view. This shows the direct influence of a molecular orbital on the activity of a functional group. The contributions of atoms O3 and O22 to HOMO-3 and -4 are not continuous variable with values 0 (contribution) or 1 (no contribution).
- Mean energy of HOMO-3 and HOMO-4 ($\bar{x}_{(\mathrm{HOMO-3,HOMO-4})}$).
- Molecular self-similarity diagonal from the overlap-like MQSM matrix.

If: Biol. act: 1($\leq$10), 2($>$20 and $<$60), 3($\geq$100)

$$\mathrm{MO_{Positional}} = \mathrm{O_{22}, HOMO-4 * O_3, \ HOMO-3}$$

$$\mathrm{MO_{Energetic}} = \bar{x}_{(\mathrm{HOMO-3,HOMO-4})}$$

S, Overlap Self-similarity
Discriminant equations:

$$D_1 = -211.485 + 0.012S - 17.755\ \mathrm{MO_{Energy}} + 3.301\ \mathrm{MO_{Positional}}$$

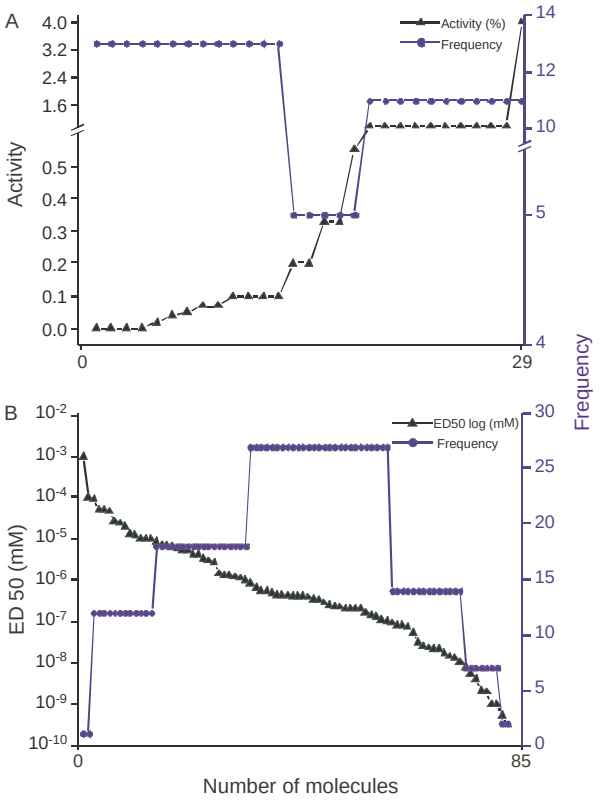


Fig. 2. Biological activity of ecdysteroid analogs in *Calliphora* test (A) and *Drosophila melanogaster* EC50 (B) [11]. In the graphic (starting from left to right) the increment of the biological activity depending on structural variations is shown (1 = 100% of activity of the natural occurrence of 20-hydroxyecdysone).

Table 2
Classification results of predicted group membership from discriminant function

Type		Group	Predicted membership			Total
			1	2	3	
Original	Count	1	1	0	0	1
		2	1	8	0	9
		3	0	5	6	11
	%	1	100	.0	.0	100
		2	11.1	88.9	.0	100
		3	.0	45.5	54.5	100
Cross-validated	Count	1	0	1	0	1
		2	0	8	1	9
		3	0	5	6	11
	%	1	0	100	.0	100
		2	0	88.9	11.1	100
		3	0	45.5	54.5	100

Cross validation is done only for those cases in the analysis. In cross validation, each case is classified by the functions derived from all cases other than that case. 71.4% of original grouped cases correctly classified. 66.7% of cross-validated grouped cases correctly classified. Biol. Activity >, =, < 1are respectively Group 1, 2 and 3.

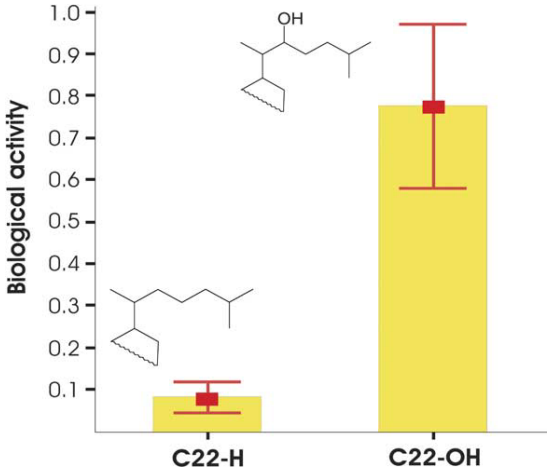


Fig. 3. Representation of biological activity of ecdysteroids as affected by OH in C22 position. Vertical lines represent standard errors. Significant *t*-value at the 0.01 level.

$$D_2 = 136.047 - 0.004S + 11.960\ \mathrm{MO}_{\mathrm{Energy}} + 1.150\ \mathrm{MO}_{\mathrm{Positional}}$$

$$\mathrm{LOO}_{\mathrm{Disc}} = 0.69$$

At this stage, the test of *D. melanogaster* with EC_{50} was considered as well. In case of *D. melanogaster*, discriminate the molecules with more or less than 1×10^{-6} in EC_{50} binding assay was a solution able to combine with the biological activity results of *Calliphora*. Discrimination for higher concentrations in a real hormonal activity will most likely depend on biological variables as a propagated effect. The main task was to identify common properties independently from the existence of different atoms arrangements and to use the existence of more molecules.

3.2. D. melanogaster

- Molecular self-similarity 'Principal Component 1' from the *Overlap-like* MQSM matrix.
- Energetic difference between the Molecular Orbitals HOMO-2 and HOMO-3. Show the relative depth of HOMO-3 in the molecular system.
- Angle among the atoms O3, O6 and the atom in the side chain to contribute to the outer molecular orbitals HOMO-3 or HOMO-4.

If: EC_{50}: $1(<10^{-7})$, $2(\geq 10^{-7})$ EC_{50}

S_{PCA_1} = OverlapSelf-Similarit (Principal component 1, Overlap matrix)

$$\mathrm{ED}_{2,3} = (\mathrm{HOMO} - 2_{\mathrm{energy}}) - (\mathrm{HOMO} - 3_{\mathrm{energy}})$$

$\alpha = \mathrm{Ang}_{3,6\mathrm{schH3\text{-}H4}}$

Discriminant equation:

$$D = 66.092 + 0.727S_{\mathrm{PCA\_1}} + 0.024\alpha + 3.118ED_{2,3}$$

Table 3
Biological activity of ecdysteroids with and without OH in C22 position

Calliphora	N	Mean	St. Dev.	St. Error	Confidence interval		Min	Max
					Min	Max		
Without (OH in C22)	6	0.103	0.133	0.054	−0.036	0.243	0	0.33
With (OH in C22)	15	0.936	0.934	0.241	0.419	1.453	0.1	4.00

The t-test substantiates the significant differences with $\alpha = 0.05$ ($t = 0.04517$**).

$$LOO_{Disc} = 0.718$$

By assuming that the constant equilibrium is a function of the structure of a molecule, two non-linear approximation of the free energy formalism ligand-receptor binding was done by means of discriminant analysis. The molecular variables are adequate in case of both biological systems and chemical. Biochemical and modelling viewpoint as well, because they are recognized as very reactive regions and atoms (Fig. 4). It is inferred that in order to make predictions of biological activity from the molecular structures (Fig. 4): (1) the functional groups are an indispensable part of the molecule but, (2) not all spatial dispositions are important to provoke a biological reaction and, (3) only some of them are spatially important, once these functional groups contribute to the reactivity properties of the molecule (e.g. the contribution of the atoms O3 and O22 and another atoms in the side chain to molecular orbital HOMO-3 and -4 or O6 to HOMO-2). As a result, they become strong causal candidates of the biological activity. Based on this, the equations described in Fig. 5 were statistically proven with good prediction fitness to activity and EC_{50}.

In biological systems, the Coulombic field is not the total representation of the electronic characteristics of the whole molecule. Behind each possible reaction, there is a theory able to support the molecule electronic constitution. HOMO is always the key external molecular orbital, but in this, QSAR HOMO-3 and HOMO-4 are the most important. Relating to this point, something irregular can be seen in these compounds. First, it is necessary to take into account the irregular distribution of the outer molecular orbitals on the large molecules like steroids. The energetic average of the difference between HOMO and HOMO-3 is not statistically related to the biological activity for the 21 studied molecules (results not shown). Considering the vicinity of the lowest degenerated molecular orbitals, HOMO is not sufficient to obtain a reliable biological activity. HOMO-3 has spatial resemblance in most of the molecules analysed, with the exception of some inactive molecules (Fig. 6, right panel). The curves show explicitly an energetic comparison among the different outer molecular orbitals for all the molecules (99) analysed (central panel). The statistical overview, by means of Cluster analysis of the energetic behaviour of outer MO, indicates the proximity of

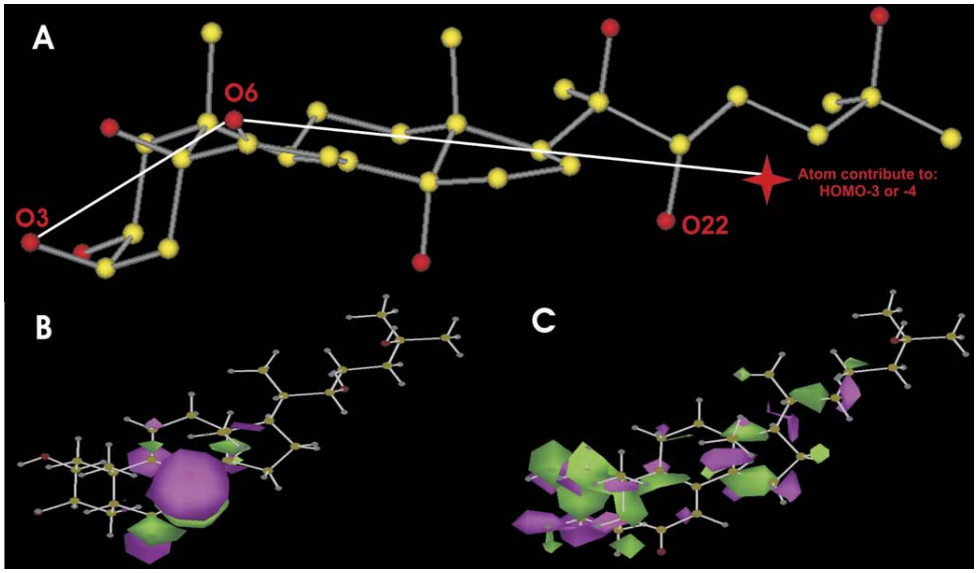


Fig. 4. Spatial representation on ecdysone of the main molecular variables related with the biological activity. (A) Geometric variables: interatomic distances (represented by brown lines), angles (represented by yellow lines), the atoms O3 y O22 are taxpayer to HOMO-3 orbital and the O6 is the mayor taxpayer to HOMO. (B) and (C) are three-dimensional representations of HOMO and HOMO-3 respectively (for interpretation of the reference to colour in this legend, the reader is referred to the web version of this article).

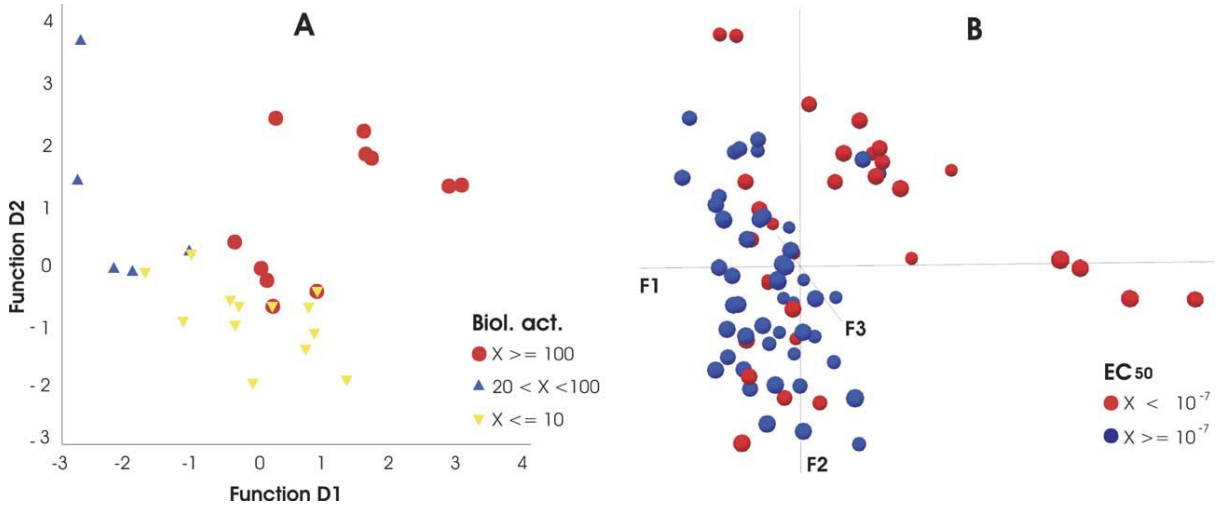


Fig. 5. Plots of Discriminat equation D1 versus Discriminat equation D2 for the studied molecules from Calliphora test (A). The discriminate analysis for the second case with *D. melanogaster* by means of other three variables and established by another equation is also show (B). Leave one out test (LOO) are 0.69 and 0.72, respectively.

the HOMO-2 and HOMO-3 and HOMO-4. This quasi band could be valid from biological point of view (Fig. 6, left panel).

The contributions of the concerning atoms O3 (ring A), O6 (ring B) and other (side chain) to this quasi band became important for biological activity. Therefore, the influence of some atoms should not be reduced to a simple geometrical spatial disposition, because the importance of the spatial disposition is mediated or conditioned by their electronic influence. OH22 functional group has forced energetic changes, with statistical significance on HOMO's -2, -3 and -4 in ecdysteroid analogs (in case of molecules analysed in *Calliphora*) by presence or absence (Fig. 7).

Two oxygen atoms on the ecdysone molecule are involved in '*n*' MO (Table 4). Consequently, H-bond between molecules of water and functional groups of ecdysone were predicted by mean of the semi-empirical MO methods (Table 4, Fig. 8). The change in contribution of the oxygen (OH in C22) to HOMO-2 was significant for PM3 and MNDO. However, for C3-OH in HOMO-3, the change in contribution of the oxygen was only significant when the PM3 method was used. The electrons of this oxygen (C22) are contributors to HOMO-2. This is a more external orbital than HOMO-3, the one for which the electrons of the oxygen (C3) are contributors. This result strengthens the hypothesis of H-bond formation by OH in C22 position.

In the case of the present model, comparing biological activities (Fig. 2) with structures represented in Fig. 9, two aspects remain unexplained. First, the analogues A and B have, respectively, 0.33 and 0.20 of biological activity without OH in

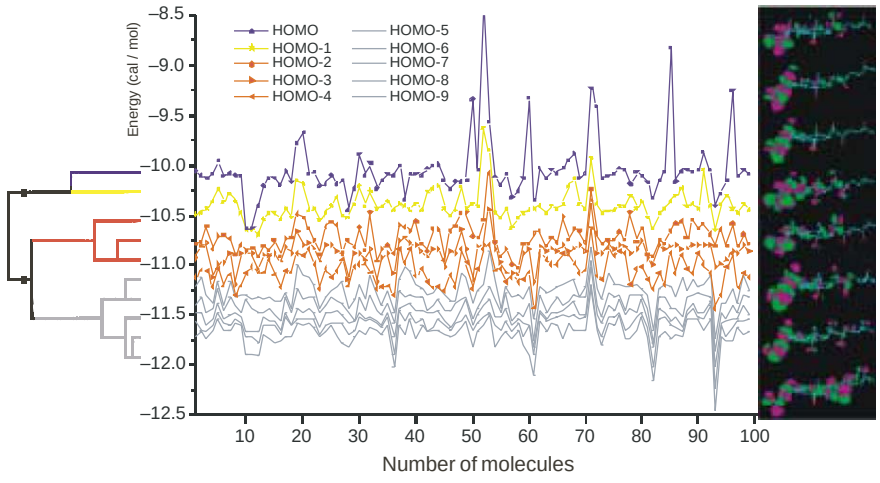


Fig. 6. Energetic behaviour of the outer molecular orbitals for all the molecules. The graph demonstrates degeneration on the most external molecular orbitals in the molecules with OH in 2, 3 and 5 positions. The convergence of energy for highest occupied molecular orbitals, from HOMO-3, HOMO-4 and HOMO-5 is shown without statistical difference. It is possible to observe the three-dimensional position of HOMO-3 for different molecules as well.

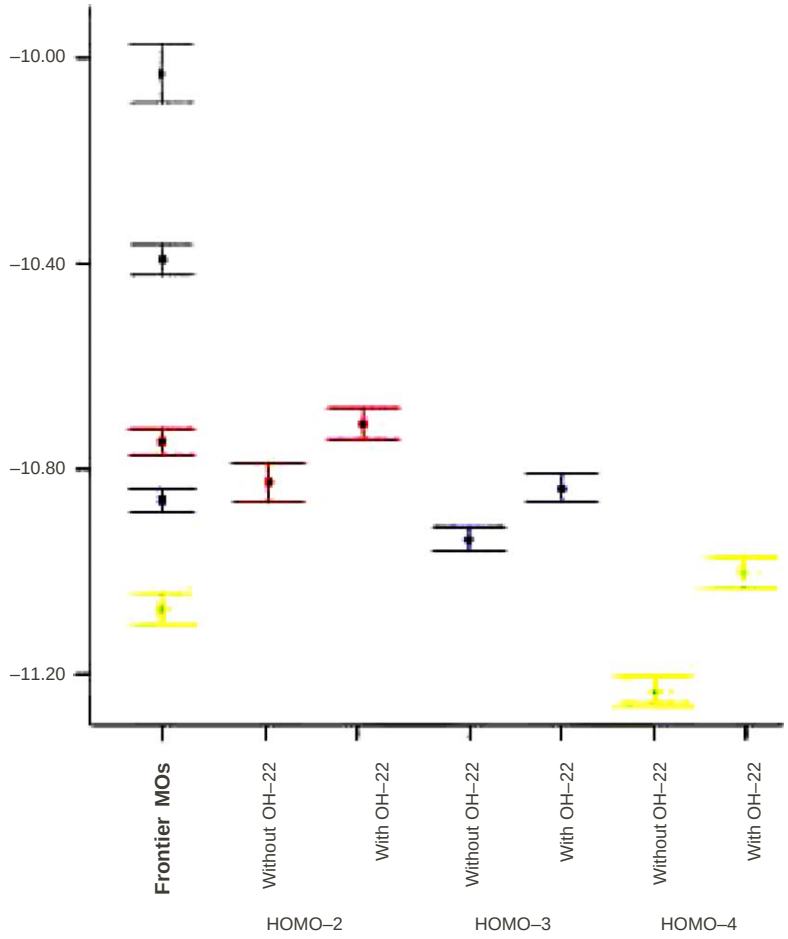


Fig. 7. Energetically, the statistical differences of highest occupied molecular orbitals for the 21 molecules studied, from HOMO to HOMO-4, are shown in the first column. Following, the significant influence of OH22 on energetic variation of HOMO-2, HOMO-3 and HOMO-4 was statistically confirmed (Mann–Whitney *U*-test, $\alpha = 0.05$). The standard errors are represented (vertical lines).

C22. Second, 5-β-Ecdysterone has four times the biological activity of the ecdysone.

The first assumption showed that OH in C22 is the decisive outcome, including the analysis of ligand-receptor interaction by H-bond. This disagreement could be solved, on one hand, with conjugations on other points of the side chain, where C26 is the most chemically likely biologically active point following the known reaction metabolism of ecdysone. Nevertheless, it needs a previous hydroxylation and the newly formed functional group will be on the outer side of the active molecular geometry (Figs. 4 and 10). On the other hand, it seems to be that no other reaction points are necessary; they only act as a catalyser of the biochemical activity, so that the analogous A and B form two new compounds metabolically by means of hydroxylation in position C22.

On the second aspect, it is interesting to mention that the OH in C5 position performs degeneration on the most external orbitals in the molecules where it exists (5-oxiecdysterone, Ponasterone C) (Fig. 6). Nevertheless, using the electrostatic potential (MEP) as a further analysis to clarify the action of these molecules (Fig. 11) we found that: both molecular regions influenced by OH at the C2, C3 and the side chain are the two most important areas to develop H-bonds. Additionally, the MEP is more negative on areas at the positions C2, C3 and O6 continuously, in the two molecules with 5-β-OH on the OHs. This is additional evidence for H-bond in this area and a reason to become 5-β-oxiecdysterone in the most active molecule, but in case of ponasterone C the OH22 is not in a clear position because of the OH24. An influence of 5-β-OH to the OH3 contribution on HOMO-3 is observed by means of the transposition of the ring A in ecdysteroids. Comparative calculus assumed even that 5-α-OH influence HOMO by way of O6. The O6 is a common important atom for all molecules and rest influence to the hydroxyl at ring A positions.

4. Discussion

Here we described the modelling of the interaction between the ecdysteroid receptor and different analogs of ecdysteroids via QSAR. The biological activity of ecdysteroids demands very specific chemical reactivities and physical properties. The existence of molecules, which show no correlation between biological activity and EC_{50}, could be attributed to biological variables in terms of mechanism. It depends normally on the

Table 4
The energy of highest occupied MOs in ecdysone

	HOMO-3	HOMO-2	HOMO-1	HOMO	LUMO
PM3					
Ecdysone	−10.98 (28% O2)	−10.75 (27.44% O22)	−10.48	−10.20	−0.25
Ecdysone–H2O C22	−11.05	−11.02 (0.8% O22)	−10.54	−10.28	−0.32
Ecdysone–H2O C2	−11.07 (0.2% O2)	−10.77	−10.51	−10.25	−0.28
AM1					
Ecdysone	−10.70 (7.54 % O2)	−10.69 (24.15 %O22)	−10.24	−10.09	−0.16
Ecdysone–H2O C22	−10.80	−10.75 (10.94 %O22)	−10.34	−10.21	−0.26
Ecdysone–H2O C2	−10.78 (31.62 % O2)	−10.61	−10.35	−10.19	−0.26
MNDO					
Ecdysone	−11.05 (41.67 % O2)	−10.93 (34.26% O22)	−10.56	−10.44	−0.40
Ecdysone–H2O C22	−11.09	−11.03 (2.22 % O22)	−10.55	−10.43	−0.38
Ecdysone–H2O C2	−11.10 (39.42 % O2)	−10.96	−10.58	−10.44	−0.41

Changes in the energy and oxygen contribution to each consequent eigenvector when H-bond in C2 and C22 position is carried out by theoretical methods.

properties of the ligand, changes of the balance between states of the receptor, channels, etc. [22]. In fact, EC_{50} as affinity binding assay is one more variable, which contributes to the biological activity as independent variable. Additionally, we will have also a dynamic binding where the affinity could be changed in the process. Finally, the trigger of a signal, as part of the efficacy of the cellular response, is mediated, as a key element, by the receptor structure [23]. Therefore, the use of both systems helps to determine the basic regions on the molecule being able to influence biological activity.

The OH groups in the molecule can act like alkyl or halogen. This depends on whether it modulates activity by virtue of effect on the physical properties (e.g. changes in the molecular electrostatic potential) or if they influence the chemical reactions significantly [24,25]. The OH groups in position C5, C14, C22 play a role in the binding behaviour. However, the sole existence of the functional group is not a sufficient condition for strong discrimination. As a consequence the molecular analysis has to be performed both partially and complete.

The spatial orientation of all oxygen atoms in steroids is not necessarily a restriction to carry out a reaction with the enzyme. In fact, there is a coincidence also reported by Brosa [26] that oxygen atoms are important, but not all of them. This approach [26] is not rational from the energetic point of view. The ligand, a steroid molecule with plane restriction will not need more than two main points to reach sufficient associated bond energy, once in contact with the receptor. Our results, from statistical and computational chemistry analysis show that the OH in C22 or another groups in the side chain are able to contribute to HOMO-3 or HOMO-4. Therefore, it becomes the most probable point of interaction in this region. Also it is already known, from the model of a putative receptor described before [4], that the side chain is the key region for interaction. On the other hand, the crystal complexes (2:1 indole-progesterone) of indole show two hydrogen bonds to the carbonyl oxygen at position O3 and O20 of the progesterone [27]. A model with the two equally important regions in ecdysteroids was predicted as well by means of CoMFA models [28]. Furthermore, the OH group in C14 is proposed to be a controversial third point, [4] although it could be a decisive piece of coordination with other groups, in some molecules like 5-β-oxiecdysterone. Harmatha [8] reported that this OH is not required for activity in ecdysone. The finding supports that two points bonding are adequate to support a reaction with the receptor. However, due to the common steroidal importance of O6 (with frequently contribute to HOMO-1) we found that an angle respect to the atom O3 and one atom in the side chain with contribution to HOMO-3 or HOMO-4 is a predicted variable. Parallel results were found by protein docking EcR-LBD analysis, but only related to OH22 [29].

Our analysis indicated that hydroxyls exert an influence on the electronic structure of the ecdysteroid molecules. In androstene the hydroxyl groups are also involved in changes on the electronic structure. Energies and enthalpies of mono

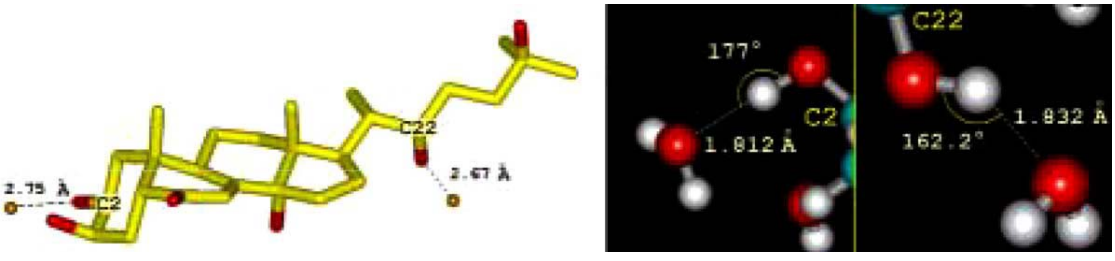


Fig. 8. H-Bond formed between the water molecule and ecdysone with the hydroxyls groups in position C2 and C22, as calculated by the PM3 Hamiltonian.

and di-hydroxy structures were the lowest [30]. Moreover, ab initio calculations of testosterone and epitestosterone showed changes on the enthalpy, dipole moment and energies of HOMO and LUMO [31]. Accordingly, it is convenient to remark that HOMO exerts its influence on the referred reactions by interactions with the hydrophobic pocket of the receptor [32]. Consequently, the study of three-dimensional characteristics in ligand-receptor interaction by means of molecular orbitals has already proved to be useful as descriptor [27]. In addition, the OHs have a qualitative influence on the energy, localization and degeneration of the HOMOs. Further intermolecular forces which may be important for biochemical reactions.

Particular attention should be paid to the region of the atoms C2 and C3 (Fig. 1). These are key points, since the direct involvement of ponasterone A has been reported in the

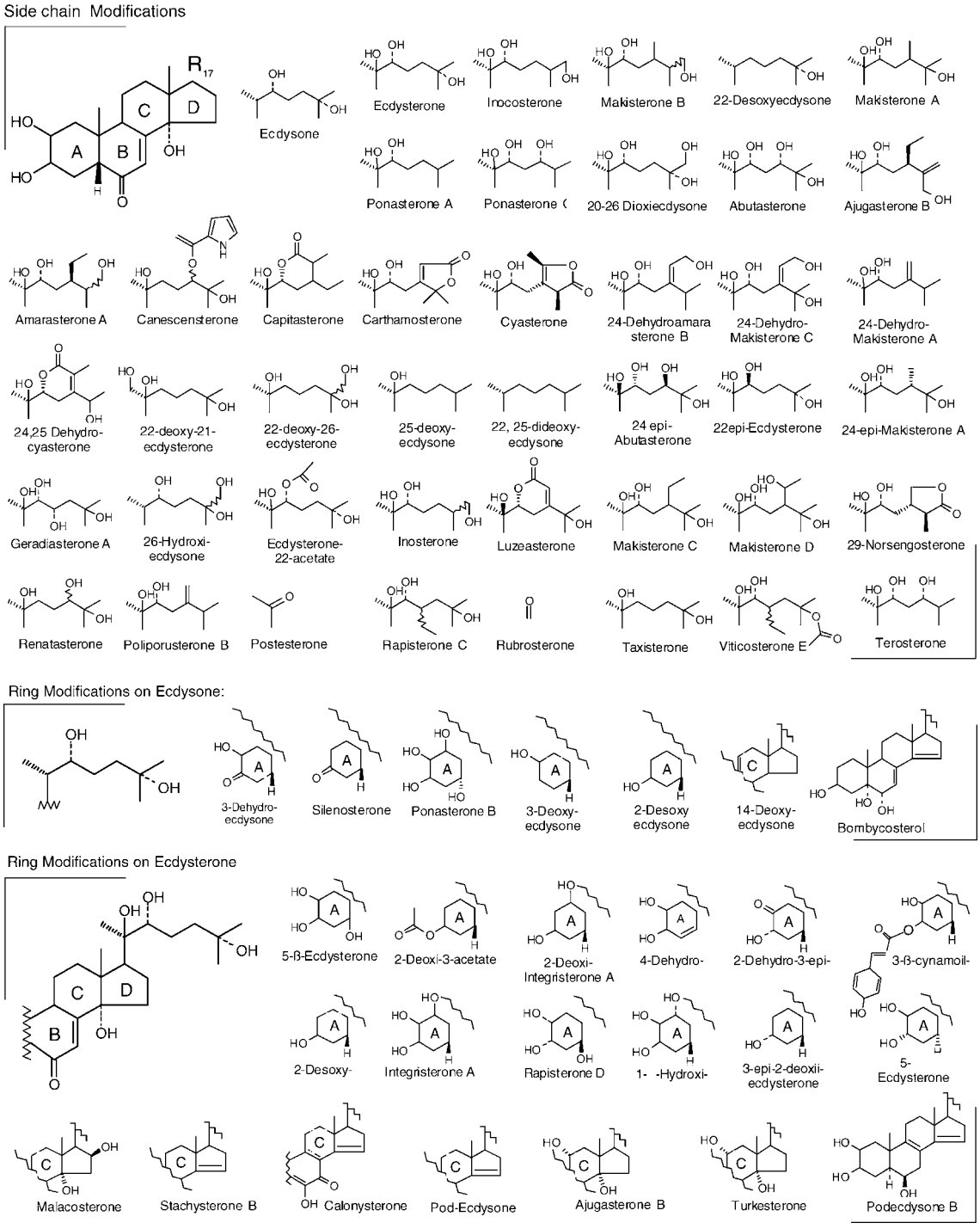


Fig. 9. Ecdysteroid and other molecules analysed in this work.

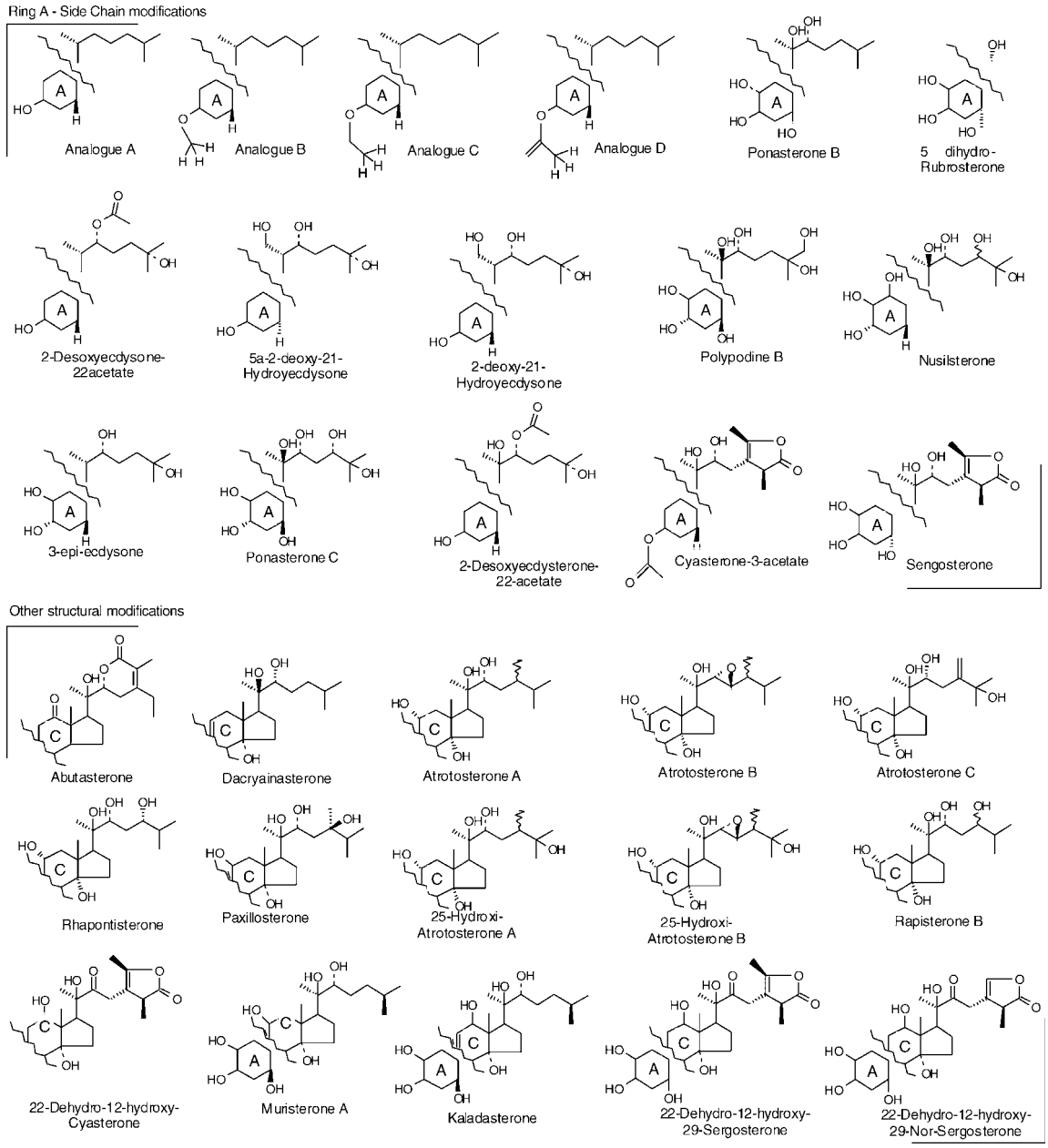


Fig. 9 (*continued*)

stabilization of the EcR-LBD structure [6]. In fact, the helical conformation of H2 is stabilized by interactions between the ponA C2- and C3-hydroxyl groups and residues of the H1-H2 loop, helix H5 and the ß-sheet [6]. An analysis of the electrostatic potential (MEP) show the importance of this zone due to a negative potential in the most active molecule (Fig. 11). This is closely related to the possibility to form H-bonds [33].

We decided to use water to develop this kind of bond with ecdysteroids, due to the aspects mentioned above and results published elsewhere: steroid-receptor interaction by means of H-bond (3–10 Kcal/mol) like 3D-QSAR and successful application on hydrophobic effect of substituents with respect to alterations in pharmacodynamic (pKi) as well as chemical equilibrium (pK_M) constants [24]. There are two main contributions to the H-bond: electrostatic attraction between oxygen and hydrogen and covalent contribution, which arises as a result the overlapping orbitals [34]. Usually, the MOs are delocalized on a large part of the molecule. Nevertheless, it can happen that they become very localized, with very specific lone-pair or π or σ character. In the former case, the lone-pair part of the MOs will be labelled n_x [35]. The stabilization of the zwitterionic structure occurs based on the destabilization of the MO localized essentially at the oxygen electron lone-pair

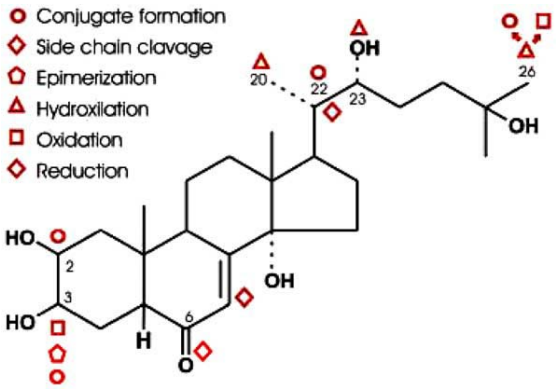


Fig. 10. Reactions of ecdysone metabolism.

(ecdysone), o-orbital, and concomitant with the stabilization of the MO essentially localized at the functional oxygen electron lone-pair, o-orbital (water) [20].

Three semi-empirical methods (MNDO, AM1 and PM3) were used to investigate H-bond on OH in positions C3 and C22. PM3 is more accurate than AM1 and both are more accurate than MNDO. However, semi-empirical NDDO calculations cannot predict the geometry of the hydrogen bond accurately [36]. Only PM3 can predict hydrogen bonds everywhere [26]. Therefore, PM3 has been used to optimize the geometry. The resulting geometry of water-OHs at positions C2, C3 and C22 are in agreement with possible hydrogen bonding [37,38] whereas O6 and OH14 are not likely able to form a hydrogen bond with significant probability.

Agonists like RH5849, tebufenozide and chromafenozide are used as insecticides (without equal potency EC_{50}), although they have only some common regions with the ecdysteroids. In the two better correspondences analyse, only three similar positions are important as a hydrogen-bond acceptor; C22 is one of them [39]. Consequently, not all the properties of the ecdysone molecule influence its biological effects; it has to be considered that the interaction of ligand and receptor is a dynamic one. Only some molecular properties of the frontier molecular orbitals and some atoms must be present. Therefore, we suggest that OH22 should be involved in a direct reaction to induce intermolecular forces while the other functional groups of the side chain are contributors to the molecular properties. As soon the molecule is attached to the receptor, rings A and B can interact whenever the distance between O3 and O22 of 11.3398 ± 0.3303 Å act as a geometric restriction for the action of the atom in position C3. The second most probable group is in the ring A, not clearly identified between OH2 or OH3 or both together. OH14 could facilitate intermolecular connection between the side chain and the ring A from electronic viewpoint.

5. Conclusions

Not all of the hormone molecule may be essential for the ultimate interaction of the hormone with the 'receptor substance' in the target cell. The overlap similarity matrix plays an important role in this analysis as an overall molecular descriptor. Therefore, not all functional groups are important to display biological activity from a purely geometrical point of view. The contribution of some atoms to three of the outer Molecular Orbitals localized along the molecules, are implicated in the structure-function as well. Although a functional group (OH) at position C22 is highly probable point of reactions for the biological activity in ecdysone analogs. It provokes variations of the energy of the HOMOs -2, -3 and -4 and is able to form H-bonds as well. OH22 is not a requisite key point in the side chain, in some molecular analogs other atoms are important, once they provided the fulfil condition of belonging at the orbital HOMO-3 or HOMO-4.

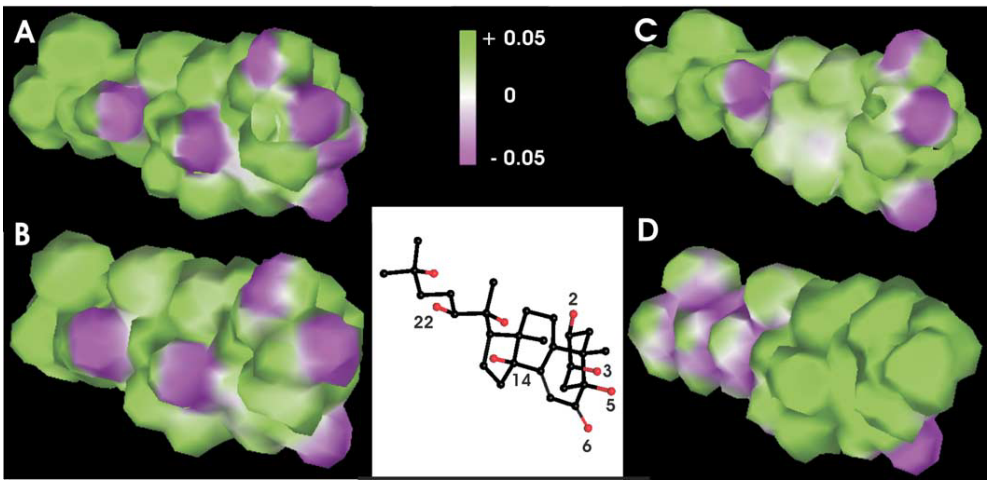


Fig. 11. Electrostatic potentials on the molecular surface computed at Hartree–Fock 6-31G*. Colour ranges, in e/a_0^3: green, represent the electron deficit regions; red, electron excess regions. (A) 5-(-Oxiexdysterone, (B) Ecdysone, (C) Pod-Ecdysone B, (D) Analog C (for interpretation of the reference to colour in this legend, the reader is referred to the web version of this article).

Elsewhere, in the ring B O6 definitely is the second essential point, common for all the steroid molecules, which in case of ecdysteroids make their contribution repeatedly to HOMO2. In addition, O3 (ring A) contribute to HOMO-3 or HOMO-4 as well. All these atoms are belonging to a quasi band amount HOMO-2, HOMO-3 and HOMO-4, energetically and spatially related to the biological activity. The angle between these three groups is crucial for the activity in the molecule. Both of the analysed systems are able to observe a common relationship respect to the predictors. Finally, the electrostatic potentials on the molecular surface reinforce the analysis above showing the same important points analysed before.

Acknowledgements

The authors are thankful to: DAAD for the financial support of the academic exchange. Italian Episcopal Conference and for the financial the ground support of the computer projects to the Laboratory of Computational Biology of Plant at Pinar del Río University.

References

[1] T.W. Goodwin, D.H.S. Horn, P. Karlson, J. Koolman, K. Nakanishi, W.E. Robbins, J.B. Siddall, T. Takemoto, Nature 272 (1978) 122.
[2] D. Cakouros, T.J. Daish, Sh. Kumar, J. Cell Biol. 165 (2004) 631.
[3] P.C. Clegg, Introduction to Mechanism of Hormone Action. Heimann, London, 1969.
[4] J. Koolman, Zool. Sci. 7 (1990) 563.
[5] T. Bergman, V.C. Henrich, U. Schlattner, M. Lezzi, Biochem. J. 378 (2004) 779.
[6] I.M.L. Billas, T. Iwema, J.M. Garnier, A. Mitschler, N. Rochel, D. Moras, Nature 426 (2003) 91.
[7] R. Lafont, D.H.S. Horn, in: J. Koolman (Ed.), Phytoecdysteroids: Structures and Occurrence Ecdysone, Georg Thieme Verlag, Stuttgard, 1989.
[8] J. Harmatha, L. Dinan, R. Lafont, Insect Biochem. Mol. Biol. 32 (2002) 181.
[9] A. Hocquet, M. Langgard, J. Mol. Model. 4 (1998) 94.
[10] M.J.S. Steward, J. Comp. Chem. Soc. 10 (1989) 209.
[11] MOPAC v.6, Release 1.02, March 1997
[12] W.P. Anderson, H. Cundari, M.C. Zener, Int. J. Quantum Chem. 39 (1991) 31.
[13] R. Carbó, P.G. Mezey (Eds.), Advances in Molecular Similarity, vol. 2, JAI Press, Greenwich, CT, 1998.
[14] R. Carbó (Ed.), Molecular Similarity and Reactivity: From Quantum Chemical to Phenomenological Approaches, Kluwer, Amsterdam, 1995.
[15] L. Amat, R. Carbó-Dorca, Quantum similarity measures under atomic shell approximation: first order density fitting using elementary jacobi rotations, J. Comput. Chem. 18 (1997) 2023–2039.
[16] L. Amat, R. Carbó-Dorca, 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. 20 (1999) 911–920.
[17] L. Amat, R. Carbó-Dorca, Molecular electronic density fitting using elementary Jacobi rotations under atomic shell approximation, J. Chem. Inf. Comput. Sci. 40 (2000) 1188–1198.
[18] P. Constans, L. Amat, R. Carbó-Dorca, Toward a global maximization of the molecular similarity function: superposition of two molecules, J. Comput. Chem. 18 (1997) 826–846.
[19] F. Jensen, Introduction to Computational Chemistry, Wiley, New York, 1999.
[20] J.E. Parra-Mouchet, G. Zapata-Torres, W.H. Fink, C.P. Nash, J. Chil. Chem. Soc. 49 (2004) 2.
[21] Lafont, R. Harmatha, J. Marion-Poll, F. Dinan, L. Wilson, D. Ian, The Ecdysone Handbook, third ed., 2002.
[22] D. Colquhoun, Br. J. Pharmacol. 125 (1998) 923.
[23] M.D. Hollenberg, Fed. Am. Soc. Exp. Biol. J. 5 (1991) 178.
[24] W.A. Sexton, Chemical Constitutions and Biological Activity. E & F.N. Spon Ltd, London, 1963,
[25] C. Brosa, I. Zamora, E. Terricabras, L. Soca, R. Peracaula, C. Rodríguez-Santamarta, Lipids 32 (1997) 1341.
[26] C. Brosa, J.M. Capdevila, I. Zamora, Tetrahedrom 52 (1996) 2435.
[27] W.L. Duax, in: M. Bohl, D.L. Duax, (Eds.), Steroid Structure and Function from X-Ray Crystallographic Studies, Molecular Structure and Biological Activity of Steroids. New York, 1992.
[28] L. Dinan, R.E. Hormann, T. Fijimoto, J. Comput. Aided Mol. Des. 13 (1999) 185.
[29] I. Billas, D. Moras, Structural studies of insect nuclear receptors, in: G. Voss, G. Ramos (Eds.), Chemistry of Crop Protection: Progress and Prospects in Science and Regulation, Wiley, New York, 2003.
[30] C. Kubli-Garfias, J. Mendieta, R. Vazquez, THEOCHEM 388 (1996) 35.
[31] C. Kubli-Garfias, THEOCHEM 422 (1998) 167.
[32] G. Smeyers, L. Bouniam, N.J. Smeyers, A. Ezzamarty, A. Hernández-Laguna, C.I. Sainz-Diaz, Eur. J. Med. Chem. 33 (1998) 103.
[33] P. Politzer, Molecular electrostatic potentials, in: P. Bultinck, H. De Winter, W. Langenaeker, J.P. Tollenaere (Eds.), Computational Medicinal Chemistry for Drug Discovery, Marcel Dekker, New York, 2004.
[34] J.J.P. Stewart, Hydrogen bonds: semiempirical methods, Encyclopaedia of Computational Chemistry, 2, Wiley, Chichester, 1998.
[35] D. Dehareng, G. Dive, Int. J. Mol. Sci. 5 (2004) 301.
[36] M.C. Zerner, Semiempirical molecular orbital methods, in: K.B. Lipkowitz, D.B. Boyd (Eds.), Reviews in Computational Chemistry vol. 2, VCH Publishers, New York, 1991.
[37] J.W. Steed, J.L. Atwood, Supramolecular Chemistry, Wiley, New York, 2000.
[38] J-H. Lii, Hydrogen bonding: 2, Encyclopaedia of Computational Chemistry, vol. 2, Wiley, Chichester, 1998.
[39] A. Kasuya, Y. Sawada, Y. Tsukamoto, K. Tanaka, T. Toya, M. Yanagi, J. Mol. Model. 9 (2003) 58.

Chapter 3

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, Thomas Reinard.

J. Chem. Inf. Model. (JCICS), 46 (4), (2006) 1751-1762.

Chapter 4

Unreveled structural requirements for auxin - like molecules by theoretical and experimental evidences

Noel Ferro, Patrick Bultinck, Ana Gallegos, Ramon Carbó-Dorca
Hans-Jorg, Jacobsen, Thomas Reinard

Phytochemistry: Manuscript PHYTO 8690

Unrevealed structural requirements for auxin-like molecules by theoretical and experimental evidences.

Noel Ferro [1, 2], Hans-Jörg Jacobsen [1], Patrick Bultinck [4], Ana Gallegos [3], Ramon Carbo-Dorca [3,4], Thomas Reinard[1]

[1] Institute for Plant Genetics, University of Hannover, Herrenhäuser Str. 2, D-30419 Hannover, Germany, Fax: +49-(0)511-762 4088, Tel.: +49-(0)511-762 3539, E-mail: nferro71@yahoo.es or mail@reinard.de

[2] Dept. of Agronomy, Forestry and Agronomy Faculty, University of Pinar del Río, Cuba.

[3] Institute of Computational Chemistry, University of Girona, Catalonia, Spain.

[4] Dept. of Inorganic and Physical Chemistry, Ghent University, Ghent, Belgium.

Corresponding Author:
Dr. Thomas Reinard
Institute for Plant Genetics, University of Hannover
Herrenhäuser Str. 3, D-30419 Hannover
Tel. +49.511.762.4032
Fax. +49.511.762.4088
mail@reinard.de

Abstract

An computational-biostatistical approach, supported by *ab-initio* optimizations of auxin-like molecules, was used for to find biologically meaningful relationships between quantum variables and fresh biological data from different bioassays. It is proven that the auxin-like recognition requires different molecular assembling states. We suggest that the carboxyl group is not the determining factor in explaining the biological auxin-like conduct. The biological effects depends on the chemical condition of the ring system. The aim to find new active molecules (quantum objects) via statistical grouping-analysis of Molecular Quantum Similarity Measures was verified in bioactivity assays. Finally, this approach led to the discovery of a new active auxin-like molecule, 2, 6-dibromo-phenol, a non-carboxylated compound. This is the first publication on structure activity relationship of auxin-like molecules, which relays on highly standardized bioassays, multi-dimensional scaling, and parallel screnning of different auxins.

Keyword Index

Auxin, structure-activity relationship, molecular quantum similarity measures (MQSM), plant growth regulation (PGR)

Introduction

The analysis of chemical messengers is still one of the hot spots in plant physiology, biochemistry and molecular biology (Kulaeva and Prokoptseva, 2004), although phytohormones like auxin have been described decades ago (Went, 1935, 1945). Indole-3-acetic acid (IAA) and its analogs were found to be the typical auxin, mainly evaluated by cell elongation tests (Koepfli *et al.*, 1938, Porter and Thimann, 1965, Thimann, 1958, Thimann and Schneider, 1938, 1939, Went and Thimann, 1937). However, the chemical space, which encompasses the term „auxin“, is actually not easily achieved, since hundreds of substances were found to exhibit an auxin-like activity in several different bioassays. The classification of this huge group of auxins was attempted before by chemical intuition, but without convincing results (Jönsson, 1961, Katekar, 1979, Koepfli *et al.*, 1938, Porter and Thimann, 1965, Went and Thimann, 1937). Different reasons can be found for the inadequacy:

- most structure analysis was performed between 1930 and 1970, in a technological scenary unable to provide substances sufficiently purified;
- the term „auxin“ is a physiological definition, but a well-defined auxin molecular structure is still not available. The existence of different auxin binding proteins, the wide diversity of the auxin molecules and the pleiotropic effects of auxin prevented the establishment of a convincing correlation beween structure and biological activity up to now;
- all attempts to elucidate this relationship based on biological assays were performed in different labs and under different conditions and complicated the comparison of the different results obtained.

The biochemical mechanisms of auxin-like molecules are thought to be in charge of particular specificities on their structure – activity relationships. These represent the result of an evolutionary process in plant kingdom (Cooke *et al.*, 2002, Dibbfuller and Morris, 1992). IAA is still known to be an active molecule in all bioassays (Woodward and Bartel, 2005), and is reliant on its plasticity properties and metabolic interactions.

In a previous work, we were able to present a computational approach dealing with semi-empirical optimizations of the auxin molecules themselves. Our approach used Molecular Quantum Similarity Measures for the analysis of more than 240 auxin-like molecules (Ferro *et al.*, 2006b). The finding of similarities in these molecules by focusing basically on intermolecular interaction descriptor, enabled us, to cluster the auxins into different groups (Tab. 1). It was postulated that the auxin-like molecular recognition depends more on specific molecular assembling states than on a specific ring system or side chain.

Here, we present the relationships between similarity groups published before (Ferro *et al.*, 2006b) in the context of their biological activities. Furthermore, this is the first publication on structure activity relationship analysis of auxins, which relays on highly standardized bioassays, performed for all different auxins in parallel. By an integral approach we were able to detect both nominal and continuous cause – effect relationships from quantitative analysis, which are objects to explain biological attributes such as:

- pleiotropic activity,

- a high number of active molecules,
- differential specificity exposed by auxin-like molecules.

The data obtained from bioassays and biostatistical analysis enabled us to refine further the structure-activity relationships of auxin-like molecules.

Results and Discussion

Analysis and Clustering of Biological Activities

Auxins are defined mainly by a set of physiological actions, but the structure-effect relationship is still based on chemical intuition. Recently, we presented a computational approach dealing with semi-empirical optimizations of the auxin molecules themselves (Ferro *et al.*, 2006b) using Molecular Quantum Similarity Measures and additional quantum variables for the analysis of about 250 different auxin-like molecules. Additional statistical analysis identified relationships between eleven structural similarity groups. These groups could be assigned to five distinct groups according to their biological activity. However, the high variance especially of bioassay data prevented a clear discrimination between the five bioactivity groups. This originates in the limitations of bioactivity data, which were available from literature. Due to different background those datas might not be comparable at all. Nevertheless, the clustering of auxins according their molecule structures revealed convincing congruences with known biological functions. For instance, the naturally occurring Indole-3-acetic acid (IAA) and its synthetic analogs 1-Naphthalene-acetic acid (1-NAA) and 2,4-Dichlorophenoxyacetic acid (2,4-D) belonged to a group sharing the same quantum spatial regions. Furthermore, neighbouring compounds within a group share similar biological activities as well (Ferro *et al.*, 2006b).

To obtain a better correlation of structural properties with its biological activity, we decided to choose representatives out of each group for bioactivity analysis in defined and standardized bioassay systems. A list of the substances, which are the subject of this analysis, is shown in Tab. 1.

An ideal bioassay offers information from the primary reaction (Veldstra, 1944) but due to their pleiotropic effects, many factors in the plant may influence the biological activity of auxin, e.g. the interactions with carriers influences membrane-permeability (Benkova *et al.*, 2003, Paponov *et al.*, 2005), the effect of auxin on endocytosis (Paciorek *et al.*, 2005), or the rapid auxin conjugation (Cooke *et al.*, 2003). It is difficult to distinguish between the primary reactions or even between primary effects. Therefore, we decided to analyse the representatives of the different structural groups on the „classical“ auxin effects, callus, rooting and elongation growth (Fig. 1). Tab. 2 summarizes the means out of ten measured values from the different measurements performed. As expected, the biological activity depends on the structure and concentration of the substance, as it can be seen by the different responses of 1-NAA and 2, 6-Br-Phe depending on their concentration. However, IAA did not induce callus significantly in any concentration, but roots only. The low activity of the compounds ILA and IAM is likely linked to the metabolic pathways of IAA (Carreno-Lopez *et al.*, 2000) and not with their self activity. Some substances, like IAA and 1-NAA applied to maize seedlings at lower concentrations showed significant inhibitory effects on root growth (Fig. 1). This result is consistent with the

observation that the auxin-overexpressing mutants tend to inhibit root elongation and auxin-deficient mutants often show long primary root growth (Woodward and Bartel, 2005).

Due to the multi-receptor and / or signal system of auxin, it is assumed more than one way of action. Therefore various parameters were used for the evaluation of the biological activity (Niklas, 2003, Campanoni and Nick, 2005). These parameters derived from the standardized experiments (Fig. 1 and Tab. 2) and are defined as follows (Fig. 2b):

- Elongation of etiolated maize seedlings (hyp) in different concentration ranges: 10^{-9}–10^{-7}M (Hyp 9-7), 10^{-9}–10^{-4}M (Hyp 9-4), and 10^{-7}–10^{-4}M (Hyp 7-4), measured tobacco root length in same concentration range as above: 10^{-9}–10^{-7}M (Rl 9-7), 10^{-9}–10^{-4}M (Rl 9-4), and 10^{-7}–10^{-4}M (Rl 7-4). Furthermore callus induction (Call Ind) and Root induction (R ind) of tobacco explants were recorded as a yes-/no-response and calculated using the equation in Tab. 2 or just as a yes/no answer (Root ind and Call ind).
- The value ED50 represents the number of single data points, showing the effect on the application of the substance, e.g. An ED50 = 50 corresponds to five out of ten explants showed an response to the substance applied.
- Furthermore, we used the logP (Tab. 2), analysed by Veldstra (1944), as a variable based on lipophilicity evaluated with the QSAR method. The lipophilicity LogP correlates with membrane permeability and receptor binding of sample auxin molecules (Bertosa et al, 2003).
- Many putative auxin receptors have been described (overview: Napier *et al.*, 2002), but best characterized one 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. Nevertheless, several auxin-like substance have been analyzed for their binding behaviour to ABP1. Therefore, dissociation constants (Tab. 1) are valuable parameter, representing the fast auxin effects.

All resulting variables were submitted to a cluster analysis (Fig. 2b) as described in the Experimental Procedures/Statistical Analysis below. Interestingly, two separate clustered branches emerged from this analysis, representing the two prominent auxin responses: the effect of auxin on (elongation) growth in the upper branches and the effect of auxin on morphogenesis in the lower branches.

The reliability of the performed cluster analysis was confirmed, when using an ABP1-overexpressing tobacco mutant instead of wild type tobacco for the same experiments as presented in Fig. 1. It is noteworthy that the response of auxin on the auxin binding protein ABP1 is in the upper branches of Fig. 2b, indicating a physiological role of ABP1 on elongation growth only. ABP1 binds a series of auxins with affinities that, for the most part, correlate with the efficacy of the compound to stimulate cell elongation (Chen *et al.*, 2001, Jones, 1994, Jones *et al.*, 1998). Fig. 3 summarizes the vector construct used for tobacco transformation and the successful expression of ABP1 in the ER of transgenic tobacco was confirmed by purification of recombinant ABP1 using Strep tag based affinity purification. Using these transgenic plants, we performed bioassays on callus and root formation, as presented in Fig. 1. No differences to effects to wt-plants were observed, indicating that ABP1

participitates in elongation growth, but not in morphogenesis. This behaviour could be predicted from the cluster analysis of Fig. 2.

Principle Component Analysis

To reduce the dimensionality in our dataset without loosing the characteristics of our dataset that contributed most in its variance, we have chosen the Principle Component Analysis (PCA), a technique that can be used to simplify a dataset. PCA can be used for reducing dimensionality by keeping lower-order principal components and ignoring higher-order ones. The idea is that such low-order components often contain the "most important" aspects of the data. The PCA (Fig. 4) provides the chance to explore the individual contribution of the auxin like substances to distinguish biological effects. Three principal components were found to be informative with a percentage of variance of 49.88% (growth factor), 24.94 % (root induction factor), and 13.99 % (callus induction factor). An assessment of the relative biological activity (Fig. 4) shows a relation between chemical structure and biological activity by means of a multidimensional approach of the auxin effects.

Structure- function Relationships

Considering the relativity of the physiological effects and the formation of two groups of clustered variables (Fig. 2B, Fig. 4), it does not make sense to assume an unified structure - activity solution (Bertosa *et al.*, 2003). Auxin induction of root and callus are qualitative events, which depend on structure and concentration. Root induction is hardly linked to callus induction and it is dependent more on the properties of the molecule than on its concentration (Fig. 2B).

The molecular classification according to effects (by cluster analysis, Fig. 4) revealed one group, which is able to produce root and callus (●), a second group producing callus only (▲), and a third one, which is inactive (▼).

These groups could be discriminated in two ways by the use of chemical descriptors in a first approach (morphogenesis):

1.The variables molecular volume, HOMO energy, hardness [η=½(HOMO-LUMO], and one factor of Quantum Similarity measure Overlap not related to IAA (Fig.5A). The Overlap-self similarity matrix was processed by factorial analysis and shows similarity factors not related to the IAA molecule.

$$D_1 = 40.472 - 0.017\,Vol + 58.952\,\eta + 56.082\,\varepsilon\,HOMO + 0.379\,F3\_Sim - Ov \qquad \text{(eq. 1a)}$$

$$D_2 = -0.994 - 0.008\,Vol - 35.121\,\eta + 6.867\,\varepsilon\,HOMO + 1.136\,F3\_Sim - Ov \qquad \text{(eq. 1b)}$$

The influence of the indol-N-atom on the HOMO orbital, positions 10 and 11 on HOMO and HOMO-1 and once more the hardness (η) are considered (Fig. 5, B).

$$D_1 = 10.323 + 56.750\,\eta + 2.247\,N8\_HOMO - 1.855\,C10\_HOMO + 2.358\,P11 \quad \text{(eq. 2a)}$$

$$D_2 = 1.226 - 6.265\,\eta + 3.167\,N8\_HOMO + 1.660\,C10\_HOMO - 1.296\,P11 \quad \text{(eq. 2b)}$$

The 68.2% and 77.3 % grouped cases, respectively, were correctly classified by cross-validation. The joint use of descriptors of electronic molecular structures and intermolecular interaction descriptors facilitated the explanation of the biological behavior (Ferro *et al.*, 2006a). The indole-N-atom is an important contributor in both HOMOs. In all cases it is the most negative atom from the molecule as well, which has being commonly used as a molecular descriptor (Vaes *et al.*, 1996). The indole compounds were very specific in root induction, with a few or no callus being induced. Even in molecules without self activity (IAA metabolites) the indole ring system is an attractive molecular prerequisite for root induction. Additionally, correlation between biological activity and the existence of the N-indole was found by Porter and Thimann (1965), independent of the charge separation theory as explanation. Porter and Thimann (1965) as well as Kaethner (1977) created an special region for this N-indole in his binding site proposal. These remarks makes it
clear that the important role of the N-indole region for auxin action is documented long before. We proved statistically the importance of this region regarding to its significance to the outer molecular orbitals (HOMO and HOMO-1).
A statistical overview of the occupied outer molecular orbital elucidates the energetic closeness of HOMO and HOMO-1 orbitals, to which COOH, do not contribute significantly in indole compounds (Fig. 6). These two orbitals are energetically far away from the rest of the Molecular Orbitals. It means that auxin-like molecules, contrary to the bigger molecules such as ecdysteroids, the most probable orbitals to produce a reaction are HOMO and HOMO-1 (Ferro *et al.* 2006a). 1-NAA and 2,6-Br-Phe induced rooting or callus depending on the concentration, even if the second structure does not contain a COOH group. The contributions of atoms at positions 10 and 11 to molecular orbitals (HOMO, HOMO-1) infer a statistical resemblance between different ring systems depending on their substitutions. TIBA orbital localizations are quite different from the other molecules due to the influence of the two iodine atoms in ortho-position (Fig. 6).

The second approach (growth) is focused on the first group of the cluster (Fig. 2). The relative activity of the substances is characterized basically by the length of the primary root. In this case the hardness (η) results in a very significant variable. The linear regression equation without outliers is:

Biological Activity = 43,4098 + 159,8590 η (eq. 3)

Hardness results in a variable, responsible for the biological activity for the three time in the present work. The linear dependence is strongly related, but the outliers indicate that some points (molecules) do not line up with the rest of the analyzed molecules. Among the outliers, very familiar compounds can be found in both forms, active (IAA, IBA, and 1-NAA) and inactive (Trysben, 3-Me-PHAA, PHAA). In order to know the cause of these behaviours we performed a discriminant analysis between the

molecules adjusted linearly and outliers. Despite of the small statistical sample, the discriminant analysis considers the influence of the double significance at positions 8 and 9 to HOMO or HOMO-1 and the significance of C15 to HOMO (HOMO LUMO graphical distribution, Fig 7). These are two recognized auxin region: C15 is the position C4 for indole rings or C8 for naphthoxyacetic acid and substitutions in this regions are very important, like 4-Cl-IAA or 8-Cl-NAA. while the region 8 and 9 correspond to the position of the N-H of the indole system in non-indole compounds.

2,6-dibrom-phenol

Attractive and fully irregular results of the present work consist the positive biological activity of a non-carboxylated compound (2, 6-dibromo-phenol) as expected by a previous classification (Ferro *et al.*, 2006) (Fig. 1 and 4). It is the first time that, in practical terms, a new active compound is found via quantum similarity measures (Carbó-Dorca, pers. com.).

Statistically, the distances between the COOH group and the ring system were not significant. Biological activity and metabolism of phenol derivates have already been identified to be active disubstituted phenols at positions 2 and 6. But the issue was focused on the mimic of conformational geometries and charge separation of the COOH and NH2 groups in respect to the ring (Farrimond *et al.*, 1980, Harper and Wain, 1969, 1971). NO2 is a withdrawing substituent, while for unpolarized π-systems the dominant interaction is π-repulsion (Hunter *et al.*, 2001). Electron availability (Katekar, 1979) and the softness as measure of the chemical polarizability are influencing the degree of auxin-like activity. The ring system and its substitutions generate the decisive factors.

Of the nearly 3200 known naturally occurring organohalogen compounds, more than 1600 contain bromine (Gribble, 1999). 2,6-Br-Phenol is a versatile molecule known as pheromone and in marine algae (Whitfield *et al.*, 1999, Leonovich, 2004). As a biological remark is interesting to say that recent experiments with Gene Silencing Activity of siRNAs with a ribo-difluorotoluyl nucleotide has demonstrated the importance of stacking interactions rather than hydrogen bonding in the fidelity of DNA replication (Xia *et al.*, 2006).

Common analysis of both approaches

The chemical space, which encompasses the auxin definition, suggests a multi-dimensional molecular space characterized by the plasticity of its biological interactions. The hardness (η) gap between anti-bonding and bonding molecular orbitals and therefore a reflection of the molecular stability (Gilman, 1997) was commonly implied in every statistical result of this paper. Soft molecules are more active than hard molecules if electron transfer or rearrangement is necessary for the reaction (Pearson, 1986). This suggestion has been statistically confirmed for auxin-like molecules both by the analysis of self-similarity Coulomb matrix of almost 250 molecules (Ferro, *et al.* 2006b) and hardness with fresh biological data.

However, the explanation of the auxin behaviour could not be reduced to a chemical variable. The existence of a large number of auxin-like molecules and their pleiotropic effect implies the principle of “a separate key to a back door” to the enzyme-substrate correlations in auxins (Veldstra, 1944). The

physiological activity of a phytoregulator is a result of the interaction of its effector chemical fragments with the receptors of those systems which form the given character (quality) (Gafurov and Zefirov, 2004). Therefore, it can be seen that the abilities of the substituents to bind to the accessory binding areas (Katekar and Geissler, 1983) are the reasons, for statistically outlier of the structure-activity or for other variables reported for the discriminant analysis.

Other variables not commented before, because of the low percentages of the cross-validation of the multivariable analysis, were found statistically significant in specific situations. First, the number of Br and F is critical as well. Halogen substitutions in organic molecules will affect their metabolic degradations as well as their intrinsic activity and are very decisive in auxin activity (Katekar, 1979, Sexton, 1963). Second, the existence of position 4 (the methylene carbon sp2 hybrids) was determinant for molecules with high activity at lower concentrations (between 10^{-9} and 10^{-7}). It is reported before as buffer area to isolate the COOH from the ring system (Katekar, 1979, Katekar and Geissler, 1982, 1983).

The carboxyl group has been considered as the vital molecular site in auxins. Paradoxically, it is chemically and physically identical in all compounds, of which acidity is separated from the ring electronic effects by the buffering effect of the intervening methylene group (Katekar and Geissler, 1982, 1983). In contrast, tryptophan (indole) is the only heterocyclic aminoacid ring system, whose electronic structure has been preserved throughout all auxin analogs (Fig. 6). The COOH group does not influence any outer molecular orbital related to the activity (Fig 6). However, if a substance with fluoride substitutions in the carbon, which is more approximate to the side chain, is able to change the electronic structure and also the biological activity is totally different (Fig 6) (Zhang and Hasenstein, 2000). Therefore, differences in activity may be due to differences in ability to bind to the electron acceptor.

The structure-activity findings are consistent with both unspecific reactions like callus induction and very specific reactions like root induction. An analysis of the packing behaviour depending on size and chemical nature of the aromatic rings in the Protein Data Bank showed that the Tryptophan (indole) prefers edge-to-face interactions (Samanta *et al.*, 1999). Additional molecular interactions analysis confirms an obvious way to effect binding to a tryptophan by hydrogen bond to the indole NH proton. One way of edge-to-face interactions with indole ring is the formation of a NH…π bond (Taylor, 2002). Tryptophan could be the only heterocyclic amino acid, which may confer to the specificity of indolic auxins in root induction and particularly is responsible for the high activity of Indole-3-butiryc acid (IBA), whose distance between ring system and COOH is uncommon for active auxins (Jönsson, 1955).

Indolic compounds can induce roots at different concentrations without any influence on other tissues. Other molecules like 1-NAA or 2,6-Br-Phe provoked root inductions or a mass of undifferentiated plant cells (callus) at highest concentrations, which can be regarded as response to stress. Others are able to produce solely callus. This suggests unspecific non-bonding interactions of non-indolic rings. Successful biological auxin-like activity requires, essentially, both the preferred geometries of non-bonded contact and the likelihood of their occurrence.

Experimental

Bioassays

Maize seeds from KWS SAAT AG (Einbeck, Germany) and tobacco leaf explants (Nicotiana tabacum cv. samsun) were used to perform assays of growth and morphogenesis.

The maize seeds were soaked in sterile water to stimulate the germination. After 8 hours they were place on soaked cotton wool for 14 hours. After a negative selection of non-germinated and extreme seedling size, the maize seeds were rolled (10 seeds per roll) in filter paper. Each of them were placed vertically in a plastic flask, which contained 50 mL of the particular substance (auxin-type) solution (Tab. 1). The whole procedure was done under dark conditions at 20°C and the evaluations were accomplished four days later.

The sterile tobacco leaf explants were placed in a MS medium (Murashige and Skoog, 1962) supplemented with vitamins. The whole experiment was performed under a 12h photoperiod at 20°C. The different substances were always tested in parallel to achieve full comparability. The evaluation of data was performed after six weeks of growing.

Hypocotyl lengths of old seedlings and root lengths were measured, mean and standard deviation of ten measurements were calculated (Tab. 2). Callus and root formation at explants were either present (positive) or absent (negative).

Bioassays using an ABP-1 overexpressing mutant

We have established an ABP1 over-expression mutant in tobacco (*Nicotiana tabacum* cv. samsun) using the dual plasmid system (binary vector) pGreen II – pSoup for Agrobacterium tumefaciens (Hellens *et al.*, 2000). Transformation was performed as described by (Yao *et al.*, 2003). In short, ABP-1 (acc. no. AF389278), under the control of a constitutive promoter, was targeted into the ER lumen. A Strep-tag was integrated between ABP1 and the ER-retardation signal KDEL (Fig. 3a). After selection of using increasing concentrations of Phosphinothricin, ABP1 expressing clones were identified and verified. To verify the expression of recombinant ABP1, the protein was purified from transgenic tobacco plants (Fig.3b) using the StrepTag purification kit (IBA GmbH, Göttingen, Germany). SDS-PAGE and staining was performed as described by Lauer *et al.*, 2005. Additionally, these ABP1 overexpressing plants were used to analyze the activities of substances like IAA and 2,6-Br-Phe, and 2-NAA. Detailed data on transgenic plants overexpressing proteins dealing with auxin signal perception will be published elsewhere (Ferro *et al.*, in preparation).

Statistical assessment

First of all a molecular representation, based on Naphtoxyacetic acid, was carried out. This represents schematically the atom positions of any kind of auxin - like molecule, and it was able to be treated statistically (Fig. 7).

A statistical featuring of biological variables was done using a classification of the range standardized (-1 to 1) by different methods of cluster analysis. This yielded a consistent dendrogram with two

groups of related variables. Based on that, two different classifications of the molecules for each case were achieved.

The repetitive information of the similarity matrixes was eliminated by Principal Component Analysis (Ferro *et al.*, 2006b). These components can fully consider a discreet distribution of the quantum objects within a three-dimensional similarity space. Therefore, all molecules are not identically related from the quantum point of view and the components are used to find relations with effects.

Next, discriminant analysis were performed to find relationships between the biological classifications and molecular properties by means of both descriptors of intermolecular interaction and quantum-chemical descriptors related to intramolecular electronic properties (Ferro *et al.*, 2006a, Raevsky, 1999). Lineal Regression analysis was carried out in particular cases where it was consistent with the phenomenological facts.

Molecular Modelling and Quantum Molecular Similarity Measures

The first molecular conformations were optimized using the MM+ force field (in the Hyperchem program, no cutoffs for non-bonded interactions and electrostatic interaction bond dipoles (Allinger, 1977)) and additionally semi-empirical PM3 calculus (MOPAC v. 6, March 1997, Stewart, 1991). Subsequently, the final geometry was performed with quantum chemical optimizations at the ab-initio level, using Gaussian. The basis set defined for most molecules was 6-31G* (Petersson and Al-Laham, 1991). In case of the remaining molecules, which includes Iodine atoms was used the base CEP-31G (Stevens *et al.*, 1984).

Next, an analysis of Molecular Quantum Similarity Measure (MQSM) was applied to the molecules used for the biological tests. The Quantum similarity methods used in the present paper are essentially the same than those exposed in the previous work on auxins by the same authors (Ferro *et al.*, 2006b). A Molecular Quantum Similarity Measure (MQSM) (Carbó *et al.*, 1980) can be defined as the scalar product between the first–order molecular density functions (DF) of two compared molecules, weighted by a non–differential positive definite operator (Ω):

$$Z_{AB}(\Omega) = \int\int \rho_A(\mathbf{r}_1)\,\Omega\,(\mathbf{r}_1,\mathbf{r}_2)\rho_B(\mathbf{r}_2)d\mathbf{r}_1 d\mathbf{r}_2 \qquad \text{(eq. 4)}$$

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 the corresponding first–order density functions.

According to the form of the weighting operator, different types of MQSM can be defined. As previously described (Ferro *et al.*, 2006b), two kinds of MQSM have been used in the present study: the so–called Overlap QSM (Carbó *et al.*, 1980), and the Coulomb QSM (Carbó and Domingo, 1987). The molecular DF has been adjusted using the Promolecular Atomic Shell Approximation (ASA) (Amat and Carbó-Dorca, 2000, Gironés *et al.*, 1998). This electron density fitting algorithm adjusts the first–order molecular electronic density functions to linear combinations of spherically symmetric functions. In the present study, the presence of bromine and iodine atoms forced the election of the Huzinaga basis set, which provides fitted functions from H to Rn (Amat and Carbó-Dorca, 1999). The number of terms in the expansion of the atomic basis set for each atom can be found at

http://iqc.udg.es/cat/similarity/ASA/table432.html, whilst the ASA exponents and coefficients for each atom can be downloaded from the web site (http://iqc.udg.es/cat/similarity/ASA/Huzinaga432/).
Similarity measures also depend on the relative orientation of the molecules being compared. In this study the field–based maximum similarity superposition algorithm (Constans *et al.*, 1997) was used to superimpose molecular structures. Once calculated, the whole set of pairwise MQSM are stored in the Similarity Matrix (SM): $\mathbf{Z} = \{Z_{AB}\}$, where Z is a squared matrix of dimension N, i.e. the number of molecules.

Conclusions

Here we made an assessment of structure–property relationship of auxin–like molecules. Our strategy based on a multi-dimensional scale of the biological activity and a dynamic view of the structural requirements. It was demonstrated that the mixture of both electronic structure and intermolecular interaction descriptors was able to discriminate this multi-dimensional biological view. Our findings can open the spectrum of new structural relationships emerging from new molecules. For the first time, QMSM method has been useful to detect a new active molecule with unexpected characteristics showing empirical and experimental evidence. A compound without COOH (2, 6-Br-Phenol) in the side chain was able to induce root, callus and inhibit root elongation. Any influence of ABP1 over-expressed mutant on root induction was not detected with IAA or 2,6-Br-Phenol. Hardness (η = 1/2 HOMO-LUMO) represents a variable, statistically related to auxin activity. The molecular regions 8, 9 and 15 are statistically detected as significant depending on their influence on the outer Molecular Orbitals.

Acknowledgements

This work was supported by a grant from the Deutsche Akademischer Austauschdienst (DAAD) to N.F. and also by continuous support from the National Fund for Scientific Research in Flanders (FWO-Vlaanderen).

Abbreviations

Abreviations of auxin-like substances are summarized in Tab. 1. PCA: Principal Component Analysis, wt: wild type, MQSM: Molecular Quantum Similarity Measure, QSM: Quantum Similarity Measure, DF: density function.

Figures and Legends

Figure Captions

Figure 1:
Effects of auxin-like molecules on maize root inhibition (left panel). Concentrations of substances applied are indicated at the top of the panel. Control was done without any hormone added. One representative out of ten seedlings is shown. Measurement of 2Cl, 6NO2-Phe (A), and 2,6NO2-Phe (B) have been measured again two days later. These later analyses is shown at the right corner (A and B respectively).
Right panels show the influence of the different substances on tobacco leaves, either on root induction (upper right panel) or on callus induction (lower right panel). The concentrations applied are indicated at the left. Only those substance, which exhibted any effect, are shown but not those substances without any influence on the tobacco explants. The detailed values are summarized in Tab. 2.

Figure 2:
The matrix left represents the relationship among the different variables. The variables are plotted in rows and columns as indicated. Colors scheme proceeds from red → yellow → blue with increasing relationship of the variables. The corresponding dendrogram at the right is a multiscaling representation of the analyzed variables in all molecules tested by the different bioassays. It clearly can be seen that the dendrogram falls in two main branches. In the upper group auxin effects, related to elongation growth can be found, whereas in the lower group morphgenetic effects are combined. The two separated groups are also represented by the blue areas in the matrix.

Figure 3:
To proove the relevance of the clustering found in Fig. 2, we carried out all morphogenesis related bioassays using an ABP1 overexpressing tobacco mutant. The upper panel A represents a scheme of the cassette used for transformation of tobacco in the vector pGreen. As it is true for ABP1 in wt plants, the recombinant ABP1 is targeted into the ER lumen due to the KDEL sequence. Lower panel shows the purification of recombinant ABP1 from transgenic tobacco leaves. The purification was performed using the separation on a Streptavidin column, which interacts with the Streptavidin binding moeity tagged to the recombinant ABP1. Lane 1 contains a molecular weight marker, lane 2 the crude protein extract, lanes 2 and 3 different fractions eluted from streptavidin column containing purified ABP1.

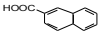
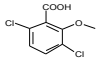
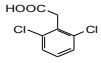
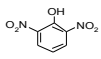
Figure 4:
Graphical representation of the PCA coefficients for the informative PCAs callus induction, root induction, and elongation growth. The coordinate system lists the coefficients against the various substances tested. This graph correlates the relative biological activity of each substance tested for with the physiological event, respectively.

Figure 5:
Overview of the structure-activity relationships of auxin-like molecules.
The scatter plots (panels A, B, C) correlate quantitative bioassay results (of callus (1A), root induction (1B), and root inhibition (2A)) with Hardness η. Bioassay data are also shown in Tab. 2. Each dot represents a molecule tested. Especially in panel C, a bordered group exhibits a linear dependence (triangles). Some outliers above and below are most interesting from the biological point of view (see text). Panel F shows details on the exceptions of outliers from the lineal dependence in panel C. Panel D is a representation of the discriminant functions from eq. 1A and 1b, and panel E shows the discriminant functions from eq. 2a and 2b. Different shapes represent the different molecular characterizations.

Figure 6:
Graphical view of spatial representation of the outer molecular Orbitals, HOMO and HOMO-1 in auxin-like molecules. Dendrogramm rechts

Figure 7:
Representation of the atom distributions to homogenize the analysis of auxin-like molecules by fix positions.

Table 1. *Substances for the biological and chemical analysis.*

Name	***Structure***	***CAS***	***pKd ABP1***	**LogP**	***Class*** [1]	***Che. Class. Induction*** [2]	***Out-lier***
Indole-3-acetic acid (IAA)		87-51-4	5,4	1,41	3 [abc]	root/call	x
Indole-3-butyric acid (IBA)		133-32-4	5	2,30	5 [abc]	root/call	x
DL-Indole-3-lactic acid (ILA)		832-97-3	-	1,22	1[bc]	root/call	
Indole-3-acetamide (IAM)		879-37-8	2,1	0,53	3 [abc]	root/call	
3-Methyl-1H-indole (Skatole)		83-34-1	-	2,29	5 [abc]	-	
1-Naphthalene acetic acid (1-NAA)		86-87-3	6,1	2,24	3 [abc]	root/call	x
2-Naphthaleneacetic Acid (2-NAA)		581-96-4	5,9	2,81	7 [ab]	inactive	
1-Naphthoic Acid		86-55-5	-	3,10	9 [a]	inactive	x
2-Naphthoic Acid		93-09-4	-	3,28	2 [abc]	-	
2,3,5-Triiodo Benzoic acid (TIBA)		88-82-4	5,1	5,03	4 [abc]	callus	
Picloram		1918021	-	0,30	-	callus	
Trysben		50317	-	2,71	9 [a]	callus	x
2-Fluorobenzoic acid (**2-F-BA)**		445-29-4	-	1,70	2 [abc]	-	x
Dicamba		1918-00-9	-	2,21	-	callus	
3-Fluor Phenylacetic acid		331-25-9	-	1,65	11[c]	callus	
2,4-Dichlorophenylacetic acid (**2,4-Cl-PAA)**		19719-28-9	-	1,75	2 [abc]	callus	
2,6-Dichlorophenylacetic acid (**2,6 Cl-PAA)**		6575-24-2	-	2,47	4 [abc]	callus	
Phenoxy acetic acid (PHAA)		122-59-8	3,8	1,34	11[c]	inactive	x
2,4,5-Trichlorphenoxyacetic acid (2,4,5-T)		93-76-5	-	3,31	7[ab]	callus	
2,4-Dichlorphenoxyacetic acid (2,4-D)		94-75-7	-	2,81	4 [abc]	-(callus)	
2- Nitro Phenoxyacetic acid (**2NO2-PHAA)**		1878-87-1	-	1,13	4 [abc]	inactive	
2,4-Dibromophenoxyacetic acid (**2,4 Br-PHAA)**		10129-78-9	-	1,86	3 [abc]	callus	
3 Methyl Phenoxy acetic acid (**3Me-PHAA)**		1643-15-8	-	1,78	4 [abc]	inactive	x
2,6-Dibromophenol (**2,6-Br-Phe**)		608-33-3	-	3,36	5 [abc]	root/call	
2-Chloro-6-nitrophenol (**2Cl-6NO2-Phe)**		603-86-1	-	2,55	6 [abc]	root/call	
2,6-Dinitrophenol (**2,6-NO2-Phe)**		573-56-8	-	1,37	7 [ab]	inactive	

Chemicals are from: Duchefa, Fluka and ABCR GmbH & Co KG, and Sigma-Aldrich. Picloram (4-Amino-3,5,6-Tricloro picolinic acid), Trysben (2,3,6-Trichloro Benzoic Acid); Dicamba (3,6-Dichloro-2-methoxybenzoic acid);
[1] Clasification (Ferro *et al.* , 2006b); [2]Quantitative information, Table 2.

Table 2:

Quantitative bioassay result. Each number was calculated by the equation: $Y_i = \frac{\left(\frac{\sum y_i}{c_j}\right)}{\sum c_j}$

Sustance	Root length [cm]			Hypo length [cm]			Root induction		Callus induction	
	-9,-7	*-9 , -4*	*-7, -4*	*-9,-7*	*-9 , -4*	*-7,-4*	*-9, -4*	*-7, -4*	*-9, -4*	*-7, -4*
IAA	6,096	4,772	2,788	5,362	5,045	4,567	0,38	0,47	0,17	0,18
IBA	8,924	7,054	4,523	5,155	4,944	4,830	0,24	0,42	0,19	0,34
1-NAA	6,558	5,092	3,003	5,175	4,976	4,806	0,06	0,11	0,56	1,00
TIBA	13,161	12,533	12,227	5,529	5,325	5,209	0,04	0,06	0,00	0,00
Picloran	11,595	9,103	6,338	5,403	5,109	4,899	0,00	0,00	0,46	0,82
Dicamba	11,642	8,894	5,730	5,350	5,173	5,065	0,13	0,23	0,44	0,79
DL-ILA	11,974	11,764	11,861	5,280	5,175	5,083	0,08	0,15	0,12	0,13
245-T	9,290	7,394	5,233	4,394	4,511	4,467	0,12	0,03	0,75	0,64
2-NAA	12,923	12,503	12,716	5,714	5,714	5,820	0,00	0,00	0,03	0,05
2,6 Cl-PAA	9,600	7,407	4,367	4,999	5,011	4,985	0,03	0,00	0,25	0,39
Trysben	13,927	12,697	11,769	5,534	5,509	5,493	0,00	0,00	0,16	0,29
2NO2-PHAA	10,348	9,622	9,107	5,332	5,159	4,965	0,00	0,00	0,00	0,00
3Me-PHAA	10,945	10,250	9,746	5,045	5,091	5,092	0,00	0,00	0,00	0,00
3-F-PAA	7,317	6,961	6,349	4,779	4,922	5,151	0,00	0,00	0,36	0,64
PHAA	8,904	8,896	8,792	5,015	4,932	4,720	0,00	0,00	0,00	0,00
Naphthoic a.	10,515	9,367	7,731	5,012	4,938	4,920	0,00	0,00	0,00	0,00
2,4 Br-PHAA	8,835	6,955	4,375	4,903	4,792	4,635	0,01	0,02	0,30	0,54
2,6-Br-Phe	9,127	7,820	3,906	5,152	5,185	3,572	0,12	0,21	0,19	0,34
2Cl-6NO2-Phe	12,905	12,258	12,297	5,573	5,419	5,308	0,00	0,00	0,13	0,23
2,6-NO2-Phe	12,698	11,575	10,751	5,703	5,648	5,549	0,00	0,00	0,03	0,05
I-3-Acetamide	12,132	11,662	11,470	5,379	5,282	5,083	0,19	0,34	0,04	0,00
2,4-Cl-PAA	9,230	7,911	6,441	4,918	4,901	4,794	0,00	0,00	0,09	0,16
skatol	11,954	11,481	11,014	5,650	5,554	5,344	-	-	-	-
2,4-D	10,839	8,035	3,946	5,517	5,127	4,650	-	-	-	-
2-F-BA	12,226	12,710	13,799	5,094	5,155	5,228	-	-	-	-
2-Naphtoic acid	14,107	12,232	10,848	5,902	5,638	5,497	-	-	-	-

Root and hypocotyl length were clustered in cm at different concentration ranges: 10^{-9}-10^{-7} M (=-9, -7), 10^{-9}-10^{-4} M (=-9, -4), and 10^{-7}-10^{-4} M (=-7, -4). Data for root and callus induction bases on a yes-no answer, calculated by the equation above

Figures

Figure 1

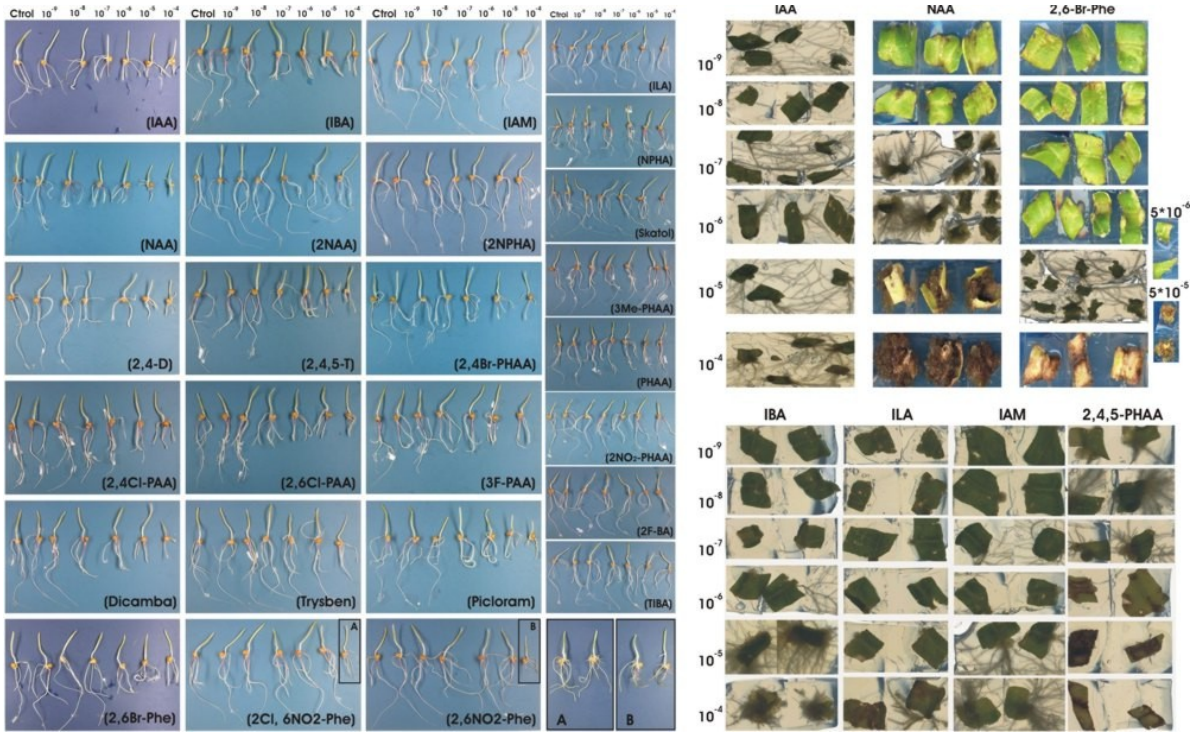


Figure 2

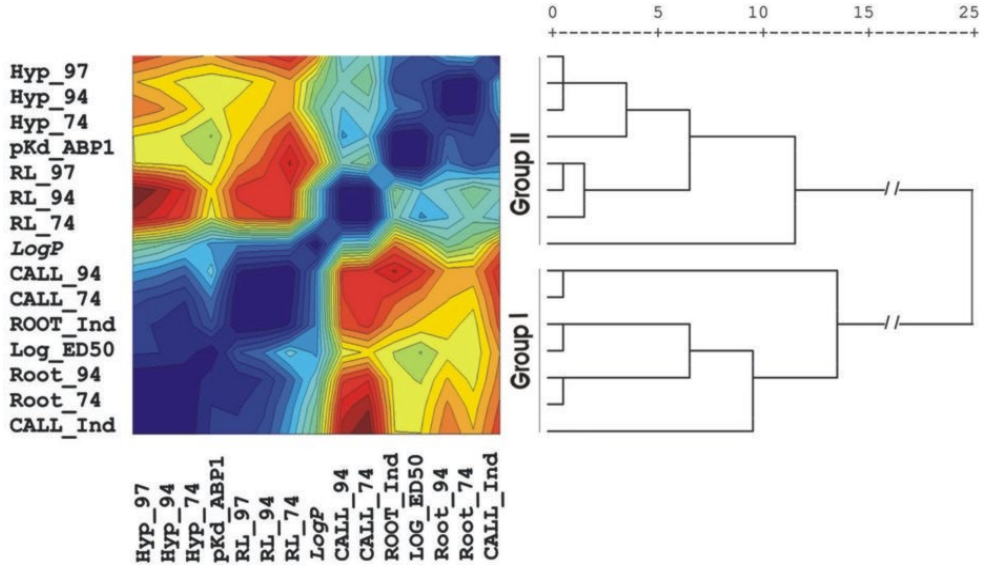

0
5
10
15
25
Hyp_97
Hyp_94
Hyp_74
pKd_ABP1
RL_97
RL_94
RL_74
LogP
CALL_94
CALL_74
ROOT_Ind
Log_ED50
Root_94
Root_74
CALL_Ind
Group II
Group I

Figure 3

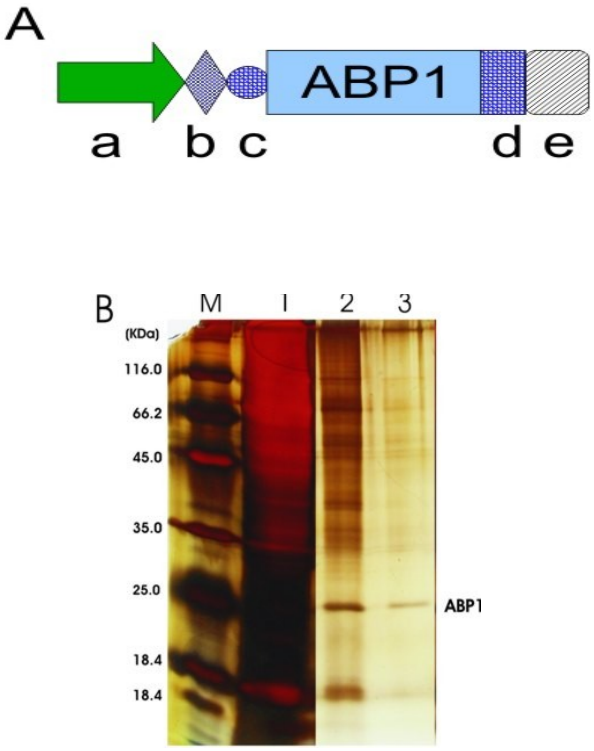


Figure 4

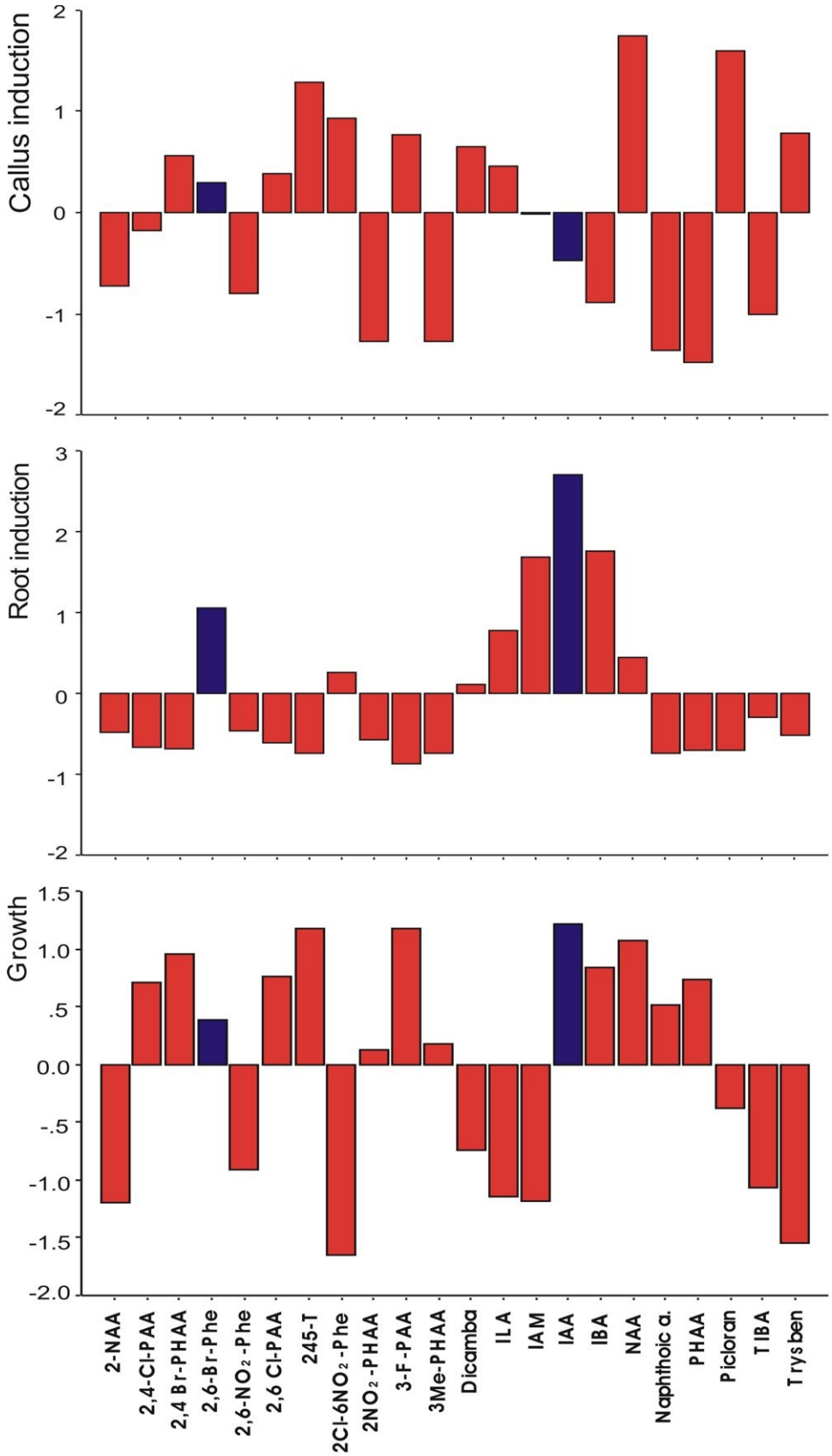

Callus induction
2
1
0
-1
-2
Root induction
3
2
1
0
-1
-2
Growth
1.5
1.0
.5
0.0
-.5
-1.0
-1.5
-2.0
2-NAA
2,4-Cl-PAA
2,4 Br-PHAA
2,6-Br-Phe
2,6-NO2-Phe
2,6 Cl-PAA
245-T
2Cl-6NO2-Phe
2NO2-PHAA
3-F-PAA
3Me-PHAA
Dicamba
ILA
IAM
IAA
IBA
NAA
Naphthoic a.
PHAA
Picloram
TIBA
Trysben

Figure 5

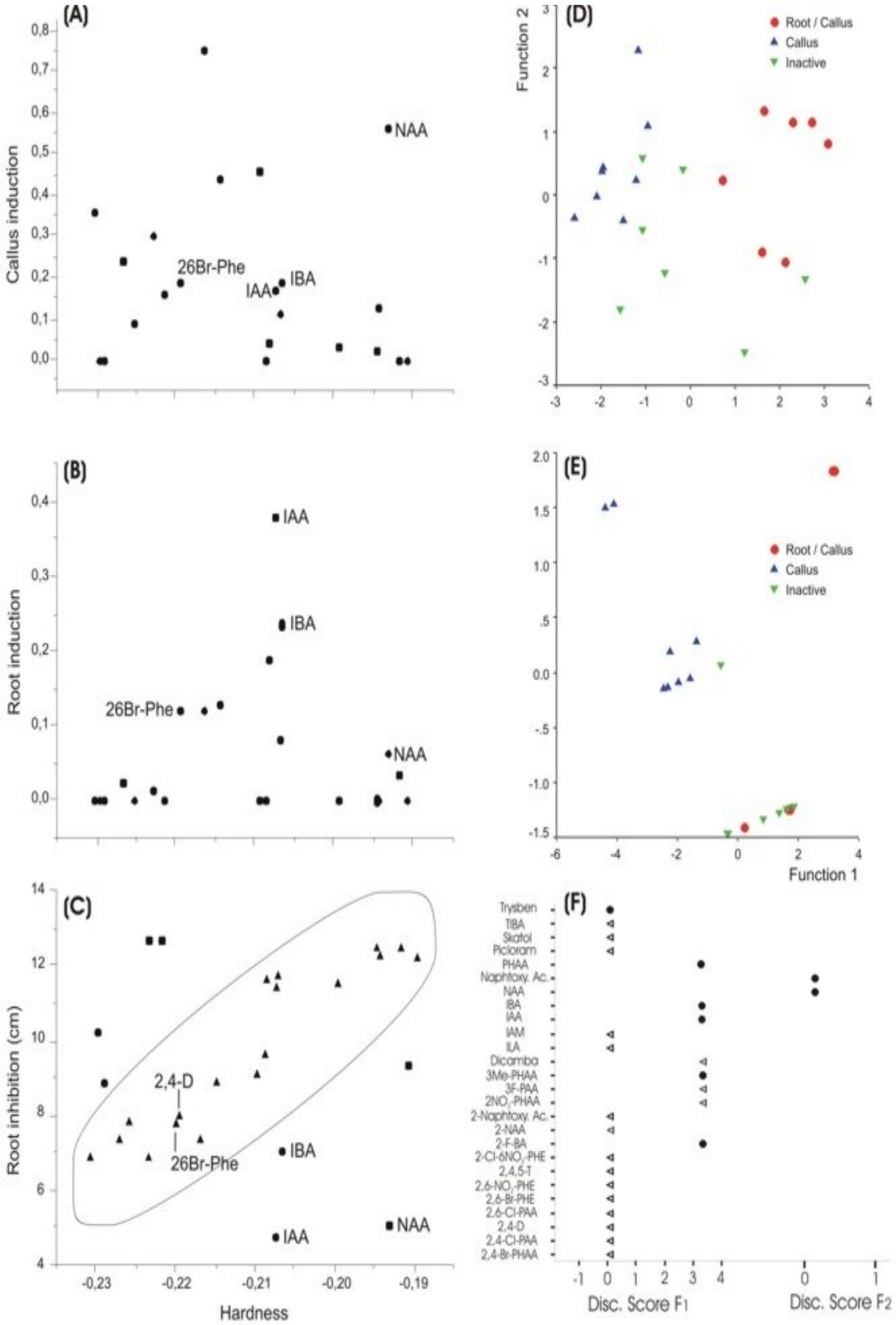

(A)
Callus induction
NAA
26Br-Phe
IBA
IAA
(B)
Root induction
IAA
IBA
26Br-Phe
NAA
(C)
Root inhibition (cm)
2,4-D
26Br-Phe
IBA
IAA
NAA
Hardness
(D)
Function 2
Root / Callus
Callus
Inactive
(E)
Root / Callus
Callus
Inactive
Function 1
(F)
Trysben
TIBA
Skatol
Picloram
PHAA
Naphtoxy. Ac.
NAA
IBA
IAA
IAM
ILA
Dicamba
3Me-PHAA
3F-PAA
2NO2-PHAA
2-Naphtoxy. Ac.
2-NAA
2-F-IBA
2-Cl-6NO2-PHE
2,4,5-T
2,6-NO2-PHE
2,6-Br-PHE
2,6-Cl-PAA
2,4-D
2,4-Cl-PAA
2,4-Br-PHAA
Disc. Score F1
Disc. Score F2

Figure 6

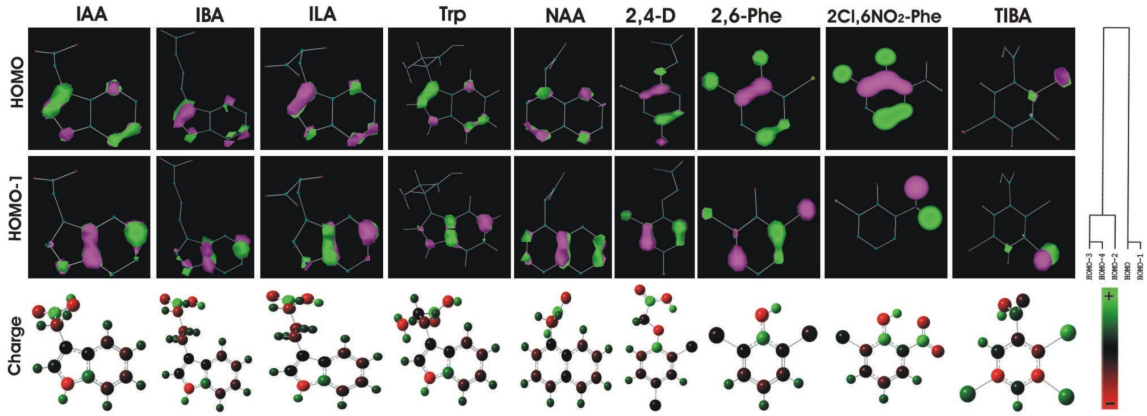
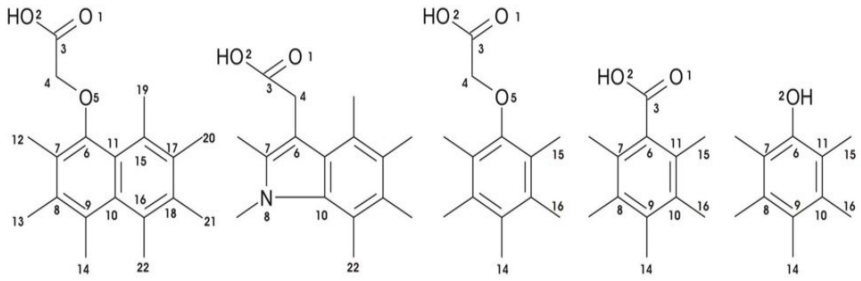

IAA
IBA
ILA
Trp
NAA
2,4-D
2,6-Phe
2Cl,6NO2-Phe
TIBA
HOMO
HOMO-1
Charge
+
-

Figure 7



References

Allinger, N.L., 1977. Conformational-Analysis .130. Mm2 - Hydrocarbon Force-Field Utilizing V1 and V2 Torsional Terms. J. Am. Chem. Soc. 99, 8127-8134.

Amat, L., Carbó-Dorca, R., 1999. Fitted electronic density functions from H to Rn for use in quantum similarity measures: cis-diammine-dichloroplatinum(II) complex as an application example. J. Comp. Chem. 20, 911-920.

Amat, L., Carbó-Dorca, R., 2000. Molecular electronic density fitting using elementary Jacobi rotations under atomic shell approximation. J. Chem. Inform. Mod. 40, 1188-1198.

Benkova, E., Michniewicz, M., Sauer, M., Teichmann, T., Seifertova, D., Jurgens, G., Friml, J., 2003. Local, efflux-dependent auxin gradients as a common module for plant organ formation. Cell 115, 591-602.

Bertosa, B., Kojic-Prodic, B., Wade, R.C., Ramek, M., Piperaki, S., Tsantili-Kakoulidou, A., Tomic, S., 2003. A new approach to predict the biological activity of molecules based on similarity of their interaction fields and the logP and logD values: Application to auxins. J. Chem. Inform. Mod. 43, 1532-1541.

Campanoni, P., Nick, P., 2005. Auxin-dependent cell division and cell elongation. 1-Naphthaleneacetic acid and 2,4-dichlorophenoxyacetic acid activate different pathways. Plant Phys. 137, 939-948.

Carbó, R., Domingo, L., 1987. LCAO-MO similarity measures and taxonomy. Int. J. Quantum Chem. 32, 517-545.

Carbó, R., Leyda, L., Arnau, M., 1980. How Similar Is a Molecule to Another? - An Electron-Density Measure of Similarity between Two Molecular-Structures. Int. J. Quantum Chem. 17, 1185-1189.

Carbó-Dorca, R., 2006. Personal comunication.

Carreno-Lopez, R., Campos-Reales, N., Elmerich, C., Baca, B.E., 2000. Physiological evidence for differently regulated tryptophan-dependent pathways for indole-3-acetic acid synthesis in Azospirillum brasilense. Mol. Gen. Genet. 264, 521-530.

Chen, J.G., Shimomura, S., Sitbon, F., Sandberg, G., Jones, A.M., 2001. The role of auxin-binding protein 1 in the expansion of tobacco leaf cells. Plant J. 28, 607-617.

Constans, P., Amat, L., Carbó-Dorca, R., 1997. Toward a global maximization of the molecular similarity function: Superposition of two molecules. J. Comp. Chem. 18, 826-846.

Cooke, T.J., Poli, D., Cohen, J.D., 2003. Did auxin play a crucial role in the evolution of novel body plans during the Late Silurian–Early Devonian radiation of land plants? In: Hemsley, A.R., Poole, I., (Eds.), The Evolution of Plant Physiology. Elsevier Ltd, California.

Cooke, T.J., Poli, D., Sztein, A.E., Cohen, J.D., 2002. Evolutionary patterns in auxin action. Plant Mol. Biol. 49, 319-338.

Dibbfuller, J.E., Morris, D.A., 1992. Studies on the Evolution of Auxin Carriers and Phytotropin Receptors - Transmembrane Auxin Transport in Unicellular and Multicellular Chlorophyta. Planta 186, 219-226.

Farrimond, J. A., Elliott, M. C., and Clack, D. W., 1980. Auxin Structure-Activity-Relationships - Benzoic-Acids and Phenols. Phytochem. 19, 367-371.

Ferro, N., Gallegos, A., Bultinck, P., Jacobsen, H.-J., Carbó-Dorca, R., Reinard, T., 2006. Coulomb and Overlap Self - Similarities: A comparative selectivity analysis of structure - function relationships for auxin - like molecules. J. Chem. Inform. Mod., In Press (manuscript ci050491c).

Ferro, N., Tacoronte, J.E., Reinard, T., Bultinck, P., Montero, L.A., 2005. Structure–activity analysis on ecdysteroids: A structural and quantum chemical approach based on two biological systems.THEOCHEM 758, 263-274.

Gafurov, R.G., Zefirov, N.S., 2004. Strategy of chemical design of phytoregulators and stress protectors with the given properties. Doklady Biological Sciences 399, 481-483.

Gilman, J.J., 1997. Chemical and physical "hardness". Materials Research Innovations 1, 71-76.

Gironés, X., Amat, L., Carbó-Dorca, R., 1998. A comparative study of isodensity surfaces using ab initio and ASA density functions. J. Mol. Graphics & Modelling 16, 190-196.

Gribble, G. W., 1999. The diversity of naturally occurring organobromine compounds. Chemical Society Reviews 28, 335-346.

Harper, D.B., Wain, R.L., 1969. Studies on plant growth - regulating substances. XXX. The plant growth - regulating activity of substituted phenols. Ann. appl. Biol. 64, 395-407.

Harper, D.B., Wain, R.L., 1971. Studies on plant growth - regulating substances. XXXI. The metabolism of certain 2,6 - disubstituted phenols within plant tissue. Ann. appl. Biol. 67, 395-408.

Hellens, R.P., Edwards, E.A., Leyland, N.R., Bean, S., Mullineaux, P.M., 2000. pGreen: a versatile and flexible binary Ti vector for Agrobacterium-mediated plant transformation. Plant Mol. Biol. 42, 819-832.

Hunter, C.A., Lawson, K.R., Perkins, J., Urch, C.J., 2001. Aromatic interactions. J.Chem. Soc. - Perkin Transactions 2, 651-669.

Jones, A.M., 1994. Auxin-Binding Proteins. Annual Review of Plant Physiology and Plant Mol. Biol. 45, 393-420.

Jones, A.M., Im, K.H., Savka, M.A., Wu, M.J., DeWitt, N.G., Shillito, R., Binns, A.N., 1998. Auxin-dependent cell expansion mediated by overexpressed auxin-binding protein 1. Science 282, 1114-1117.

Jönsson, A., 1955. Synthetic plnat hormones. VIII. Relationship between chemical structure and plant growth activity in the arylalkyl-, Aryloxyalkyl- and I*ndole-alkylcarboxylic acid series*. Svensk Kem, Tidskr 67, 166-187.

Jönsson, A., 1961. Chemical structure and growth activity of auxins and antiauxins. In: Ruhland, W., (Ed.), Encyclopediea of Plant Physiology, Vol. 14. Springer, Berlin,959-1006.

Kaethner, T. M., 1977. Conformational Change Theory for Auxin Structure-Activity-Relationships. Nature 267, 19-23.

Katekar, G.F., 1979. Auxins - Nature of the Receptor-Site and Molecular Requirements for Auxin Activity. Phytochem. 18, 223-233.

Katekar, G.F., Geissler, A.E., 1982. Auxins II: the effect of chlorinated indolylacetic acids on pea stems. Phytochem. 21, 257-260.

Katekar, G.F., Geissler, A.E., 1983. Structure-activity differences between indoleacetic acid auxins on pea and wheat. Phytochem. 22, 27-31.

Koepfli, J.B., Thimann, K.V.a., Went, F.W., 1938. Phytohormones: structure and physiological activity. J. Biol. Chem. 122, 763-780.

Kulaeva, O.N., Prokoptseva, O.S., 2004. Recent advances in the study of mechanisms of action of phytohormones. Biochemistry-Moscow 69, 233-247.

Lauer, B., Ottleben, I., Jacobsen, H.-J., Reinard, T., 2005. Production of a Single-Chain Variable Fragment Antibody against Fumonisin B1. J. Agric. Food Chem. 53, 899 – 904.

Leonovich, S. A., 2004. Phenol and lactone receptors in the distal sensilla of the Haller's organ in Ixodes ricinus ticks and their possible role in host perception. Experimental and Applied Acarology 32, 89-102.

Murashige, T., Skoog, F., 1962. A Revised Medium for Rapid Growth and Bio Assays with Tobacco Tissue Cultures. Physiol. Plant. 15, 473-&.

Napier, R.M., David, K.M., Perrot-Rechenmann, C., 2002. A short history of auxin-binding proteins. Plant Mol. Biol. 49, 339-348.

Niklas, K.J., 2003. The bio-logic and machinery of plant morphogenesis. American J. Bot. 90, 515-525.

Paciorek, T., Zazimalova, E., Ruthardt, N., Petrasek, J., Stierhof, Y.D., Kleine-Vehn, J., Morris, D.A., Emans, N., Jurgens, G., Geldner, N., Friml, J., 2005. Auxin inhibits endocytosis and promotes its own efflux from cells. Nature 435, 1251-1256.

Paponov, I.A., Teale, W.D., Trebar, M., Blilou, K., Palme, K., 2005. The PIN auxin efflux facilitators: evolutionary and functional perspectives. Trends Plant Sci. 10, 170-177.

Pearson, R.G., 1986. Absolute Electronegativity and Hardness Correlated with Molecular-Orbital Theory. Proc. Nat. Acad. Sci. U.S.A. 83, 8440-8441.

Petersson, G.A., Al-Laham, M.A., 1991. A complete basis set model chemistry. II. Open-shell systems and the total energies of the first-row atoms. J. Chem. Phys. 94, 6081-6090.

Porter, W.L.a., Thimann, K.V., 1965. Molecular requeriments for auxin action I. Halogenated indoles and indoleacetic acid. Phytochem. 4, 229-243.

Raevsky, O.A., 1999. Descriptors of molecular structure in computer-aided design of biologically active compounds. Uspekhi Khimii 68, 555-576.

Samanta, U., Pal, D., Chakrabarti, P., 1999. Packing of aromatic rings against tryptophan residues in proteins. Acta Crystallographica Section D-Biological Crystallography 55, 1421-1427.

Sexton, W.A., 1963. Chemical constitution and biological activity. E. & F. N. Spon LTD, London.

Stevens, W.J., Basch, H., Krauss, M., 1984. Compact effective potentials and efficient shared-exponent basis sets for the first- and second-row atoms. J. Chem. Phys. 81, 6026.

Stewart, J.J.P., 1991. Optimization of Parameters for Semiempirical Methods .3. Extension of Pm3 to Be, Mg, Zn, Ga, Ge, as, Se, Cd, in, Sn, Sb, Te, Hg, Tl, Pb, and Bi. J. Comp. Chem. 12, 320-341.

Taylor, R., 2002. Life-science applications of the Cambridge Structural Database. Acta Crystallographica Section D-Biological Crystallography 58, 879-888.

Thimann, K.V., 1958. Auxin activity of some indole derivatives. Plant Physiol. 33, 311-321.

Thimann, K.V.a., Schneider, L.C., 1938. Differential growth in plant tissues. Am. J. Bot. 25, 627-641.

Thimann, K.V.a., Schneider, L.C., 1939. Differential growth in plant tissues. II. a modified auxin test of high sensitivity. Am. J. Bot. 26, 792-797.

Vaes, W.H.J., Hamwijk, C., Ramos, E.U., Verhaar, H.J.M., Hermens, J.L.M., 1996. Partitioning of organic chemicals to polyacrylate-coated solid phase microextraction fibers: kinetic behavior and quantitative structure-property relationships. Anal. Chem. 68, 4458 – 4462.

Veldstra, H., 1944. Researches on Plant Growth Substances IV. Relation between chemical structure and physiological activity I. Enzymologia 11, 97-136.

Went, F.W., 1935. Auxin, the Plant-Growth Hormone. Botanical Review 1, 162-181.

Went, F.W., 1945. Auxin, the Plant-Growth Hormone .2. Botanical Review 11, 487-496.

Went, F.W.a., Thimann, K.V., 1937. Phytohormones. Macmillan Company.

Whitfield, F. B., Helidoniotis, F., Shaw, K. J., Svoronos, D., 1999. Distribution of bromophenols in species of marine algae from eastern Australia. J Agric. Food Chem. 47, 2367-2373.

Woodward, A.W., Bartel, B., 2005. Auxin: Regulation, action, and interaction. Annals of Botany 95, 707-735.

Xia, J., Noronha, A., Toudjarska, I., Li, F., Akinc, A., Braich, R., Frank-Kamenetsky, M., Rajeev, K. G., Egli, M., Manoharan, M., 2006. Gene Silencing Activity of siRNAs with a Ribo-difluorotoluyl Nucleotide. ACS Chem. Biol. 1, 176-183.

Yao J., Pang Y., Qi H., Wan B., Zhao X., Kong W., Sun X., Tang K., 2003. Transgenic tobacco expressing Pinellia ternata agglutinin confers enhanced resistance to aphids. Transgenic Res. 12, 715-722.

Zhang, N. G. and Hasenstein, K. H. (2000). Halogenated auxins affect microtubules and root elongation in Lactuca sativa. J. Plant Growth Reg. 19, 397-405.

Chapter 5

Supplementary Discussion.

Plant bioassays: a poor structural mirror

About one century ago the concept of hormone was coined by Starlin in the Royal Society. Until now, the hormone concept plays a crucial role in a wide range of research fields from chemistry to molecular biology. In our current concept of signal transduction mechanisms, hormone receptors occupy a central position- Our view on the structure function relationship of hormones has followed a progression of technology, that is needed to understand biological processes in details (Tata 2005).

The assignment of the specific region within a complex chemical space of an organic molecule that is responsible for the biological process, is a bold but critical endeavour. It is hampered by the fact that the biological action interacts only with a very small fraction (1x10-50%) of the relevant chemical space (Dobson 2004). In this direction, molecular biology has the potential to revolutionize pharmacology and medicine (Lipinski and Hopkins 2004).

In animals, a typical receptor is present and responds to hormone concentration in the range of a few μM. The relative activity is the only variation of the response and depends solely on the concentration. Even events like the propagation of the effect of receptor occupation may be important for the specific effect of the hormone. But not all hormone molecules in the target cell may be necessary for the ultimate interaction of the hormone with the "receptor substances" . Some parts of the molecule may act as transport vehicle, which protects the molecule during its biological journey. It should be emphasized that the molecular variables are able to cause biological effects. Certain flexibilization of the idea about how the molecular orbitals are

influencing the biological activity could be very interesting. From our results we could not rule out something conclusive, but the biological activity depends on the spatial distribution of the outer MO (Molecular Orbitals) on the molecule and the energetic differences among them. The use of intermolecular descriptors, which is based on the electronic structure, resulted in very good information on the interaction between ligand and receptor.

In plant science, especially true for hormones, like auxin, the pharmacological methods are not accurate. In fact, the cart was put before the horse, since molecular biology has changed the reflexions about the methods used for the work on plant hormones (Kepinski and Leyser 2005; McCourt 1999). All plant growth bioassays, which are based upon the responses of the preformed organs, the immediate stimulus merely "unblocks" some previous limitations, but in action, the molecules in question function as a part of a matrix of interacting and interlocking events. This philosophy is very different from the classical concept of hormonal action in animals (Steward and Krikorian 1971). Therefore, the auxin molecular diversity was not conceived under biological rules, but these biological rules were used to confirm chemical dependences in the biological systems.

Auxin, its molecular diversity and pleiotropic activity

Three main problems are dealing with auxins,

1. High amount of active molecules. That is a biological controversial issue because of the hormone concept. Affinity is described by the equilibrium constant for complex (AB) formation (K_{eq}= [AB]/[A_{free}] [B_{free}]), the free energy of the complex formation is ΔG_{AB} = -RTlnK_{eq}. Specificity is conveniently defined as the difference in affinity between ligands A and A´ ($\Delta\Delta G_{AA'} = \Delta G_{AB} - \Delta G_{A'B}$). Really in the hormone definition does not include the specificity, however it is one essential characteristic for this kind of molecular interaction. A very high specificity requires more stringent discrimination mechanisms, when competitors are similar and abundant, but more variable when there are few and distinct competitors present (Szwajkajzer and Carey 1997). Auxin

may depict the other paradigm in respect to hormone specificity (Clevenger 2003).The core problem is, auxins are dissimilar and abundant.

2. Pleiotropic physiological effect. Another controversial issue is a relative behaviour of compounds strongly influenced on 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. The auxin specificity nature is depending on what kind of effect they are inducing (Steward and Krikorian 1971). It is very difficult to follow the reaction proper from phenomenological point of view by a single biological test.
3. The high molecular diversity. This is a very crucial point, since most structure activity theories on auxins have not been proven by strong and reproducible experiments. Only basing on experimental evidences with many exceptions or without sufficient and / or a representative statistical samples, reliable experimental data are highly needed (Farrimond et al. 1981; Hansch and Muir 1950; Kaethner 1977; Katekar 1979; Porter and Thimann 1965; Tomic et al. 1998a). In fact, in one hand it is very difficult to select chemical descriptors that are able to manage such different structures. On the other hand, in order to unify the molecular individualities, strategies for the selection of statistical variables is a difficult task.

How to face the problems

So far the broader range of auxin molecular diversity was analysed by Jönnson (Jönsson 1961). His work was also the key source for the comprehensive Katekar's publications (Katekar 1979; Katekar and Geissler 1982; Katekar and Geissler 1983; Katekar et al. 1987). While Jönnson´s work was forgotten, the intuitive principles of Katekar´s work is still the basis for most of the auxin research. Since the 1960s - 1970s clear and common dogma for an auxin molecule is well accepted: a rich electronic surface formed by different ring systems frequently combined with halogen substituents and a high interface activity. But the base of this view was already established by Veldstra about 20 years before (Veldstra 1944; Veldstra 1952; Veldstra

1953; Veldstra and Vandewesteringh 1951b; Veldstra and Vandewesteringh 1952). Furthermore Veldstra emphasized the importance of a balance between a defined hydrophilic part (carboxyl group) and a lipophilic part (ring system) (Veldstra and Vandewesteringh 1951a).

The way to face the structure activity-relationship must be, first of all, by using electronic descriptors of intermolecular interactions (Bertosa et al. 2003; Tomic et al. 1998a; Tomic et al. 1998b). That is a way to have much information about different molecules. In case of auxin we are exactly in the core problem of the possible interaction ligand-receptor.

First of all, it has to be figured out, which kind of similarity actually correlates with the biological activity. In case of auxin the difficulties aggravate due to the problematic historical background of plant bioassays (Steward and Krikorian 1971; Veldstra 1944; Veldstra 1953; Weyers and Paterson 2001).

This scenario led to our strategy:

- first, a consensus variable was developed, in which the general information on biological activities independent from test and tissue could be included. The reference point for this was just the maximal activity of the appropriate substance;
- second, different substances were biotested in parallel using different assays combined with a procedure similar to statistical multi-scaling analysis. These analyses are able to eliminate the redundant information raising from the biological context, thereby focussing on the proper reaction.

The statistical treatment of similarity matrices of both, Coulomb and Overlap operators, was preceded by a Principal Component Analysis, which allowed the minimisation of the highly repetitive information. Different methods of cluster analysis were applied and the relationship between the resulting clusters and biological activity was tested. 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 biological activity groups. In this way molecules with contradictory properties in biological acitvity and structural requirements, could be

enlightened, as it is true for the well-known inactivity issue of 8-Cl-NAA (Katekar et al. 1987).

The second approach necessarily followed the cyclic scientific method: hypothesis formulation, experimental design, and data analyses (Box et al. 1978). Thereby, some new insights were unraveled and the approach was convincingly confirmed by the (predicted by structure) activity of a compound without classical characteristics of auxins. The substance 2,6-bromophenol does not even contain a COOH group, but nevertheless, it belonged to the group of active substances in the structural classification made earlier by Molecular Quantum Similarity Methods (MQSM).

Moreover, variables were used, which are able to characterize the electron structure of molecules. This new approach is enabled by recent methods of calculus with high precision, like ab-initio prediction and which gave good results, thereby solving some old problems related with the inclusion of the atom N in the indole ring (Kiralj and Ferreira 2003). The statistical analysis of the energies of the outer molecular orbitals arrived at the conclusion that HOMO and HOMO-1 can form a quasi-band, while the other orbitals are energetically spoken far away from them. The additional analysis of participation of the different atoms on each of these outer molecular orbitals allowed the explanation of most of the biological information comprehended in auxin molecules.

The multi-scaling analysis based on the finding of similarities among the biological variables and allowed us to discriminate between two groups of variables. These variables are also linked in a very fine manner with the two central biological mechanisms related to auxin: morphogenesis and growth (cell elongation). Control experiments using an ABP1-over- expressing mutant confirmed that the position of the results (from binding assays) for these hormones are related with process of cell elongation. This has been found by different means and authors before.

Two sets of structural requirements for auxin-like molecules

The biological evidences suggested that the high diversity of auxin molecules can only be explained when more than one set of biological descriptors are used. Therefore, we splited the analysis. The results pointed out one region in the molecules with very high importance. This region is defined by the existence of the N-indole atom. The second region is somehow modifiable , depending on the HOMO - HOMO-1 localization between the atomic positions 10 - 11 (morphogenesis) and 15 (growth).

It should be emphasized that the distance between the carboxyl group and different parts of the ring had no significant influence from our sample (data sets), including the distances in respect of the more negative atoms in the ring. Therefore, it can be concluded that the balance between the hydrophilic and lipophilic regions (Veldstra and Vandewesteringh 1951a) is important, especially for cell growth, in which the C-sp2 carbon (at position 4 - Fig7 - Chapter 4) was found to be a determinant variable for biological activity. The buffer area, postulated by Katekar (Katekar and Geissler 1982), can be important for the occurrence of hight activities, but strict distance relationships are dispensable.

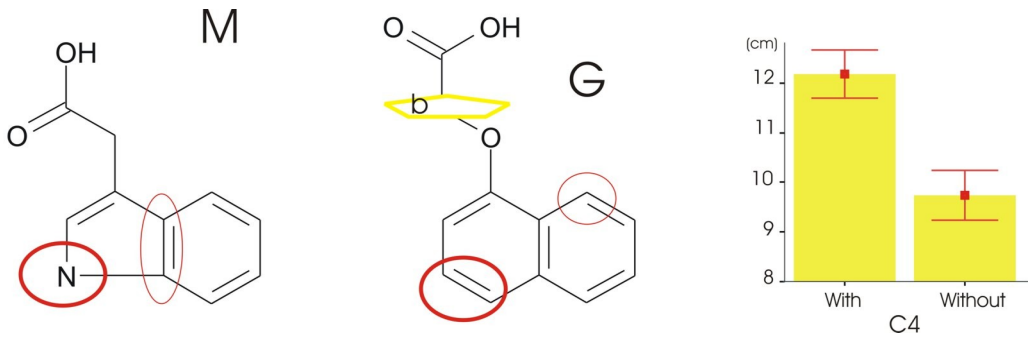


Fig. 7: Probably "accessory binding areas": Significant regions of the molecular orbitals related to events of morphogenesis (M) or growth (G). Region buffer important for growth regulation (b).

Morphogenesis (root induction) seems to be dominated by ring interactions and by the recognized N-indole region (Kiralj and Ferreira 2003). The significance is

supported by the observation that indole compound acts strongly on root induction but hardly on callus induction. It can be postulated that callus formation depends on the mimetic representation of the orbital structure of the N-indole ring in other kind of ring. Other rings might provoke non-specific interactions (callus induction).

General reflexions

Finally, a common variable was declared, which influences the biological activity of auxins at every level: the chemical hardness (η) is determined by the gap between the Highest Occupied Molecular Orbital (HOMO) and the Lowest Unoccupied Molecular Orbital (LUMO). The gap between anti-bonding and bonding molecular orbitals represents a reflection of the molecular stability (Gilman 1997). In short, soft molecules are more active than hard molecules if electron transfer or rearrangement is necessary for the reaction (Pearson 1986). This is exactly allocated by our results. The reaction of auxins involves electron arrangements independently from the biological signal system. This variable "hardness" (η) is the fundamental variable for all auxin molecules and their activities, thereby this variable elucidates the high variation of the auxin effects.

Actually, hydrogen bonding and stacking interactions are getting a more significant role being the prime reason for electrostatic interactions in biological systems. Experiments in gene silencing activity of siRNAs with a ribo-difluorotoluyl nucleotide revealed that the stacking interactions play a major role, rather than hydrogen bonding in the fidelity of DNA replication (Rebek et al. 1987; Xia et al. 2006).

Indole-3-acetic acid is a derivative from tryptophan, the only heterocyclic one of the usual amino acids. Therefore, the packing of the indole ring system is common in proteins. Here, it was proven by ab-initio Electronic Structure Calculation that the electronic structure of the indolic hetero-ring is conserved between both tryptophan and IAA, even when changes at the end of the side chain occurred. Therefore, IAA

could inherit the different packing characteristic of tryptophan (Samanta et al. 1999; Taylor 2002).

The strategies used for this thesis are the key of the chemical design, which will result in the generation of new synthetic auxins and/or comprehensive bases of the action mechanism. These approach have to fulfil the following requirements. This may be also extended to the the phytohormone context. It is already shown that the field of molecular interactions could be more complicated than expected (Gafurov and Zefirov 2004). Gafurov confirmed the existence of auxin and gibberellin mimetics by experimental evidences basing the molecular design on the following postulates:

1. even weak but simultaneous stimulating influences on some systems with a given character, can cause a strong responds;
2. to achieve this objective, these influences should be complementary and coordinated;
3. the physiological activity of a molecule is a result of the interaction of its effector chemical fragments with the receptors of those systems which form the given character (quality);
4. physical and chemical properties of the effector fragments and the whole molecule do not determine the regulation influence qualitatively.

Outlook and future prospect

Quantum Chemical Methods and biostatistical analysis will further assist in clarifing the statistical regularities of the biological matrix, representing interactions within the auxin scenery. A feed-back approach of computational chemistry and molecular biological methods will be a promising strategy to further unravel the phenomenon of the pleiotropic effects of phytoregulators (auxin).

The use of these techniques will open up new perspectives in auxin and plant hormone regulators research. The classification of the diversity of auxins-like molecules will unravel the connections at the biochemical and the molecular level. The new variables, developed in this thesis are strongly linked to fresh biological

evidences and they will enable us, to develop better classifications for the almost 1000 auxin-like molecules known from literature.

A consequent experimental design could assist to find new proteins involved in the metabolism of auxins. The analysis will focus on ligand structure and binding assays. The general mechanism bases on the concepts of ligand-specific receptor conformations and conditional efficacy and should consider the ligand-specific physiological response.

The introduction of point mutations can (at least theoretically) be used for further analysis of the interactions between ligand and different amino acids. The existence of more than one receptor in the auxin signal transduction chain demands to the discrimination of the structural-binding relationships for each receptor – ligand pair and the evaluation of the physiological relationships for each of them. Further non-additive influences from the molecular and biochemical context can be expected, which will result in a very complex system.

Literature

Bertosa B., Kojic-Prodic B., Wade R.C., Ramek M., Piperaki S., Tsantili-Kakoulidou A., Tomic S. (2003) A new approach to predict the biological activity of molecules based on similarity of their interaction fields and the logP and logD values: Application to auxins. Journal of Chemical Information and Computer Sciences 43: 1532-1541.

Box G.E.P., Hunter W.G., Hunter J.S. (1978) Statistics for Experimenters: An Introduction to Design, Data Analysis, and Model Building. John Wiley & Sons, 672 p.

Clevenger C.V. (2003) Nuclear localization and function of polypeptide ligands and their receptors: a new paradigm for hormone specificity within the mammary gland? Breast Cancer Research 5: 181-187.

Dobson C.M. (2004) Chemical space and biology. Nature 432: 824-828.

Farrimond J.A., Elliott M.C., Clack D.W. (1981) Auxin Structure-Activity-Relationships - Aryloxyacetic Acids. Phytochemistry 20: 1185-1190.

Gafurov R.G., Zefirov N.S. (2004) Strategy of chemical design of phytoregulators and stress protectors with the given properties. Doklady Biological Sciences 399: 481-483.

Gilman J.J. (1997) Chemical and physical "hardness". Materials Research Innovations 1: 71-76.

Hansch C., Muir R.M. (1950) The Ortho Effect in Plant Growth-Regulators. Plant Physiology 25: 389-393.

Jönsson A. (1961) Chemical structure and growth activity of auxins and antiauxins. In: Ruhland W (ed) Encyclopediea of Plant Physiology. Springer, Berlin, pp 959-1006.

Kaethner T.M. (1977) Conformational Change Theory for Auxin Structure-Activity-Relationships. Nature 267: 19-23.

Katekar G.F. (1979) Auxins - Nature of the Receptor-Site and Molecular Requirements for Auxin Activity. Phytochemistry 18: 223-233.

Katekar G.F., Geissler A.E. (1982) Auxins II: the effect of chlorinated indolylacetic acids on pea stems. Phytochemistry 21: 257-260.

Katekar G.F., Geissler A.E. (1983) Structure-activity differences between indoleacetic acid auxins on pea and wheat. Phytochemistry 22: 27-31.

Katekar G.F., Winkler D.A., Geissler A.E. (1987) Hormone recognition in plants. In: Klämbt D (ed) Plant hormone receptor. Springer-Verlag, Berlin, pp 13-26.

Kepinski S., Leyser O. (2005) The Arabidopsis F-box protein TIR1 is an auxin receptor. Nature 435: 446-451.

Kiralj R., Ferreira M.M.C. (2003) Combined computational and chemometric study of 1H-indole-3-acetic acid. International Journal of Quantum Chemistry 95: 237-251.

Lipinski C., Hopkins A. (2004) Navigating chemical space for biology and medicine. Nature 432: 855-861.

McCourt P. (1999) Genetic analysis of hormone signaling. Annual Review of Plant Physiology and Plant Molecular Biology 50: 219-243.

Pearson R.G. (1986) Absolute Electronegativity and Hardness Correlated with Molecular-Orbital Theory. Proceedings of the National Academy of Sciences of the United States of America 83: 8440-8441.

Porter W.L., Thimann K. V. (1965) Molecular requeriments for auxin action-I. Halogenated indoles and Indoleacetic acid. Phytochemistry 4: 229-243.

Rebek J., Askew B., Ballester P., Buhr C., Jones S., Nemeth D., Williams K. (1987) Molecular Recognition - Hydrogen-Bonding and Stacking Interactions Stabilize a Model for Nucleic-Acid Structure. Journal of the American Chemical Society 109: 5033-5035.

Samanta U., Pal D., Chakrabarti P. (1999) Packing of aromatic rings against tryptophan residues in proteins. Acta Crystallogr D Biol Crystallogr 55: 1421-7.

Steward F.C., Krikorian A.D. (1971) Plant, chemicals and growth. Academic Press, New York, 232 p.

Szwajkajzer D, Carey J (1997) Molecular and biological constraints on ligand-binding affinity and specificity. Biopolymers 44: 181-198.

Tata J.R. (2005) One hundred years of hormones - A new name sparked multidisciplinary research in endocrinology, which shed light on chemical communication in multicellular organisms. Embo Reports 6: 490-496.

Taylor R. (2002) Life-science applications of the Cambridge Structural Database. Acta Crystallogr D Biol Crystallogr 58: 879-88.

Tomic S., Gabdoulline RR, Kojic-Prodic B, Wade RC (1998a) Classification of auxin plant hormones by interaction property similarity indices. Journal of Computer-Aided Molecular Design 12: 63-79.

Tomic S., Gabdoulline R.R., Kojic-Prodic B., Wade R.C. (1998b) Classification of auxin related compounds based on similarity of their interaction fields: Extension to a new set of compounds. Internet Journal of Chemistry 1: Cp1-U21.

Veldstra H. (1944) Researches on Plant Growth Substances IV. Relation between chemical structure and physiological activity I. Enzymologia 11: 97-136.

Veldstra H. (1952) Halogenated Benzoic Acids and Related Compounds. Recueil Des Travaux Chimiques Des Pays-Bas-Journal of the Royal Netherlands Chemical Society 71: 15-32.

Veldstra H. (1953) The Relation of Chemical Structure to Biological Activity in Growth Substances. Annual Review of Plant Physiology and Plant Molecular Biology 4: 151-178.

Veldstra H., Vandewesteringh C. (1951a) Researches on Plant Growth Regulators.19. Structure Activity.4. Partially Hydrogenated Naphthoic Acids and Alpha-Alkyl-Phenylacetic Acids. Recueil Des Travaux Chimiques Des Pays-Bas-Journal of the Royal Netherlands Chemical Society 70: 1113-1126.

Veldstra H., Vandewesteringh C. (1951b) Researches on Plant Growth Regulators.20. Structure Activity.5. 2-Phenyl-Cyclopropane-1-Carboxylic Acids. Recueil Des Travaux Chimiques Des Pays-Bas-Journal of the Royal Netherlands Chemical Society 70: 1127-1135.

Veldstra H., Vandewesteringh C. (1952) On the Growth Substance Activity of Substituted Benzoic Acids - Preliminary Communication. Recueil Des Travaux Chimiques Des Pays-Bas-Journal of the Royal Netherlands Chemical Society 71: 318-320.

Weyers J.D.B., Paterson N.W. (2001) Plant hormones and the control of physiological processes. New Phytologist 152: 375-407.

Xia J., Noronha A., Toudjarska I., Li F., Akinc A., Braich R., Frank-Kamenetsky M., Rajeev K.G., Egli M., Manoharan M. (2006) Gene Silencing Activity of siRNAs with a Ribo-difluorotoluyl Nucleotide. ACS Chemical Biology 1: 176-183.

Appendix

Publications

Journals

- **Ferro, N.**, Jacobsen, H.-J., Bultinck, P., Gallegos, A., Carbó-Dorca, R., Reinard, T. Assessment of unrevealed structural requirements for auxin - like molecules by theoretical and experimental evidences. *Phytochemistry (eingereicht).*
- **Ferro, N.**, Gallegos, A., Bultinck, P., Jacobsen, H.-J., Carbó-Dorca, R., Reinard, T. Coulomb and Overlap Self - Similarities: A comparative selectivity analysis of structure - function relationships for auxin - like molecules. *J. Chem. Inf. Model.* In press. (manuscript ci050491c).
- **Ferro, N.**, Tacoronte, J. E., Reinard, T., Bultinck, P., Montero, L. A. Structure-Activity analysis on ecdysteroids: a structural and quantum chemical approach based on two biological systems. *J. of Molecular Structure: THEOCHEM* 758, (2006) 263-274.
- **Ferro, N.**, Tacoronte, J. E., Zaldo, A., Montero, L. A., González-Jonte, R. Analysis of the electronic and geometric properties of ecdisteroids. *CNIC (Chemical Sciences),* 2004.
- **Ferro, N.**, Gonzalez-jonte, R., Carnesoltas, D., Alvarez-Idaboy, J. R., Iglesias, M., Coll, F.
  Characterization of brassinosteroids and their analogs by quantum computational methods and biostatistics. *FCTL, Folia Chim. Theoret. Latina* -Vol. XXVII, p105-118, 1999.
- González, S., Diosdado, E., Rodríguez, J., Román, M.I., González, C., Garbey, P., Coll, D., Benítez, D., Abreu, C., Echemendía, D., Ramírez, I., **Ferro, N.**, Coll, F. Biological activities of new regulators on plants. *Biología*, 12: 28-34, 1998.

Book Reports

- **Ferro, N.** Getting into details about structure-activity relationships of auxins. British-Cuban Bioinformatic Workshop. 16-21 December, 2002. (http://karin.fq.uh.cu/biocubauk).
- **Ferro, N.**, Morales, L. O., Noda, A. L., González, S., Sistachs, V., Jomarron, I. Bioactivity studies of brassinosteroid analogue DI-31 on tobacco tissue. Summary Book of the X Scientific Seminar of the INCA. 1996.
- **Ferro, N.** Micropropagation of Nicotiana tabacum, var. H-2000. L., Book Short Reports, Biotecnologia'95, Havana, Cuba 1995.

Congresses

- Biotechnica 2005. Hannover, Germany. **Oct. 2005**
- DFG Schwerpunktprogramm: "Molecular Analysis of Phytohormone Action".– GBM. Würzburg Bundesrepublik Deutschland. **Apr. 2005**
- III International workshop on Tobacco Culture, Pinar del Río, Cuba. **Dez. 13-16, 2001**
- I Workshop on Bio-Informatic. Havana, Cuba. **Apr. 11-12, 2001**
- Britisch - Kuba Bioinformatik Workshop. **Dez. 2002**
- IV Seminar of Advanced Studies on Molecular Design and Bio-Informatics. University of Havana. **Feb. 2002**
- International Symposium on Natural Resources Management, IUFRO. Pinar del Río. Cuba. **Apr. 2000**
- Bioinformática 99. Introductory workshop of Bioinformatic. Havana. **Jun. 1999**
- Humboldt Symposium on Chemistry. University of Havana, Havana. **Mär. 1999**
- XXIV International Congress of Theoretical Chemistry. Puebla, Mexico. **Sep. 1998**
- IV provincial Shop of Guava of the Pinar (Psidium salutare). Pinar del Río. Cuba. **Jan. 1998**
- 50 Anniversary of the University of East. Santiago de Cuba. Cuba. **Oct. 1997.**
- BIOVEG'97. Center of Bioplantas. Ciego de Avila. Cuba. **1997**

- I Workshop about Plant Growth Regulators. Fac. Biology, U. Habana. Habana. Cuba. **Apr. 1997**
- X Scientific Seminar of the INCA. Havana. Cuba. **Nov. 1996.**
- XV Conference of Chemistry of the Universidad de Oriente. Santiago de Cuba. Cuba. **Nov. 27 – 30, 1996**
- International Scientific Pedagogic Workshop "The applied to Chemistry the Sustainable Agriculture.". Habana. Cuba. **Jul. 3 – 5, 1996**
- National FORUM of students of Agricultural Sciences. ISCAB. Bayamo. Cuba. Tutor. **Jun, 6 – 8, 1996**
- IV WORKSHOP OF BRASSINOSTEROIDS. Havana. Cuba. (Brassinosteroids). **Nov. 30 - Dic. 1 1995**
- I WORKSHOP OF BIOACTIVOS PRODUCTS. Havana Cuba. (Oligosacharides). **Nov. 30 - Dic. 1 1995**
- BIOTECCNOLOGIA-95. Havana. Cuba. **Nov. 1995**
- First International Symposia on Tropical and Subtropical Fruit-culture. National Fruit Station. Havana. Cuba. **Okt. 1995.**

Declaration / Erklärung

Hiermit erkläre ich an Eides statt, dass die vorliegende Dissertation nicht schon als Diplomarbeit oder ähnliche Prüfungsarbeit verwendet worden ist.

Hannover, den

Acknowledgements

First of all, I would like to thanks to **Annia M. Labrador** (my wife) who, por razones ajenas a mi voluntad, take care alone from my children from almost three years ago.
My father and my mother, because of the force they gave me to face the life.
I want to express my gratitude to the DAAD for the financial and personal support, which enabled this thesis.
Mi mas sincero y especial agradecimiento al **Dr. Thomas Reinard** for his support, as professional and Human Being, the interesting topic, an excellent supervision and the exchange of knowledge under the rules of the scientific passion.
Agradezco sinceramente al **Prof. Dr. Hans-Jörg Jacobsen** for his support, the hospitality in his institute, and for the high level of commitment to the professional promise and personal trust.
Personal thanks Prof. **Dr. Patrick Bultinck**, who support my ideas about the use of Theoretical Chemistry in Biology, offer to stay at his Department in Ghent as second home in Europe.
Al viejo amigo **Prof. Dr. Ramón Carbó-Dorca**. The application of his theory in auxins was my scientific dream, since he made an exposition in Havana more than 5 years ago. It's already come true.
Rehana, who kindly take care for my English in every article.
Herr und Frau Dirschel and Frau Müller for their support in Hannover
Many thanks will go to my colleagues, who supported me during the time of my PhD thesis, especially to André, Katha, Ilka, Jasmin, Isa and Maren.
I also like to express many thanks to **all my friends** who endured my moods in the final time of my thesis. Lina, Roberta, Ingo, Ingo´s Vater und Mutter, Jose, Marcela, Patricia, Cesar, Martin, Claudia... Gracias a la Familia Ulloa por acojerme entre ellos.

For my family

...

Curriculum Vitae

Noel Ferro Diaz

M.Sc. Planzenbiologie, Landwirtschaftsingenieur

Geboren am:	28. März 1971 in Havana.
Nationalität:	Kubanisch
Familienstand:	Verheiratet (durch politische Gründe von meiner Familie getrennt lebend)

Seit 4.2004	Doktorarbeit an der Universität Hannover, Institut f. Pflanzengenetik, Abt. II.
11.–12. 2004	DAAD Stipendium Universität Ghent, Belgien Spezielle Ausbildung auf dem Gebiet: „Quantum Molecular Similarity Measures“ (QMSM). bei Prof. Dr. Patrik Bultinck und Prof. Dr. Ramón Carbó-Dorca
10.2003 – 3.2004	Goethe - Institut, Göttingen Sprachkurs Deutsch – Intensiv
8.-12.2000	DSE Stipendium, Leipzig, Deutschland
4.– 6. 2000	Berufliche Aus – und Fortbildung auf dem Gebiet Biotechnologie Universität Havana. Fakultät für Chemie, Havana, Kuba Spezielle Ausbildung auf dem Gebiet: “Molekulares Design“ bei Prof. Dr. Luis Alberto Montero
2.1995 – 7.1999	Universität Havana. Fakultät für Biologie, Havana, Kuba Master: Pflanzenbiologie Fachrichtung: Pflanzenbiotechnologie (Hormonwirkung, Statistik and Datenanalyse) Masterarbeit: Analyse des Einflüsse von Brassinosteroiden-Analoge mit starrer Seitenkette in Tabak.

Studium	Universität Pinar del Río, Kuba	
1989 – 1994	Landwirtschaftsingenieur	
	Fachrichtung:	Pflanzenbiotechnologie
	Diplom Landwirtschafts-ingenieur	Analyse der exogenen Applikationen von Auxin und Cytokinin im pflanzlichen Organogeneseprozess.
Schule	Institut für Naturwissenschaften Friedrich Engels, Pinar del Rio, Kuba	
1984 – 1988	"Instituto Pre-Universitario Vocacional de Ciencias Exactas"	
	Vergleichbarer Abschluß: Abitur	