# Extensive Chemometric Investigations of the Multidrug Resistance in Strains of the Phytopathogenic Fungus *Penicillium Digitatum*

**Rudolf Kiralj\* and Márcia M. C. Ferreira**

Laboratório de Quimiometria Teórica e Aplicada, Instituto de Química, Universidade Estadual de Campinas, Campinas, SP, 13083-970, Brazil, E-mail: rudolf@iqm.unicamp.br

## Abstract

Multidrug resistance activities $pEC_{50}$ of sensitive, resistant, and moderately resistant strains of the pathogenic fungus *Penicillium digitatum* against triflumizole, fenarimol, bitertanol, pyrifenox, cycloheximide, acriflavine, and 4-nitroquinoline-*N*-oxide were studied by principal component and hierarchical cluster analyses. Genome descriptors for fungal cytochrome P450 sterol 14α-demethylase and multidrug efflux pump PMR1 were generated and correlated with $pEC_{50}$ by partial least squares. Toxicants were modeled at the PM3 level. Novel Activity – Structure Relationships (ASRs) were established to predict toxicant structural features from biological activities and to identify and classify the strains. New types of relationships to model and predict biological activities are Quantitative Genome – Activity Relationship (QGAR) and Quantitative Genome/Structure – Activity Relationship (QGSAR). QGAR for demethylation inhibitors had a reasonable regression [$Q^2 = 0.79$, $R^2 = 0.81$, Standard Error of Validation (SEV) = 0.34] and was extended into QGSAR with improved statistics ($Q^2 = 0.85$, $R^2 = 0.87$, SEV = 0.29). Conformers of toxicants with common hydrogen bonding and aromatic ring geometry indicate possible interactions with receptors such as cytochromes, efflux pumps, and regulatory proteins which activate fungal multidrug resistance.

## 1 Introduction

*CYP51* or *ERG11* gene [1 – 5] encodes one of the most frequent 14α-methyl sterol demethylases (ergosterol biosynthesis enzyme), CYP51 or $P450_{14DM}$, in different biological kingdoms (animals, plants, fungi and yeast, lower eukaryotes, bacteria). It catalyzes a three-step reaction which requires oxygen and NADPH for the oxidation of 14-methyl in some sterols into formic acid.

Demethylation inhibitors (DMIs) are aromatic nitrogen-based compounds (azoles, triazoles, imidazoles, pyrimidines, pyridines, *etc.*), used to interfere sterol biosynthesis in pathogenic fungi as pesticides and antimycotics [2 – 4], and as chemotherapeutics in postmenopausal breast tumor cells [2]. They also act against mycobacteria and streptomycetes [5]. The inhibitors used as antifungals prevent 14α-demethylation of lanosterol into 4,14-dimethylzymosterol, resulting in the depletion of ergosterol and accumulation of its sterol precursors [3]. Such altered sterol composition, after being integrated into fungal plasma membrane, causes the membrane disruption and cell death.

Structural studies of cytochrome CYP51 and its interactions with substrates and inhibitors [1, 6 – 11] have demonstrated the importance of hydrogen bonds, hydrophobic – hydrophobic, and aromatic – aromatic intermolecular interactions. Aromatic nitrogen from a DMI heterocycle binds chemically to the iron of the heme from CYP51, thus preventing the activation of the oxygen molecule.

The emergence of multidrug resistance by pathogenic fungi represents serious problems in agriculture [12, 13] and medicine for immunocompromised patients (HIV, cancer, surgery, *etc.*) [14, 15]. Fungi possess numerous resistance mechanisms [3, 4, 16], among which are *CYP51* alterations (point mutations, overexpression, gene amplification, *etc.*) and drug efflux (efflux pumps expression, overexpression, point mutations). Point mutations of fungal CYP51 [7, 8, 17, 18] occur at its entrance channel, by which substrates are differentiated from inhibitors in DMI-resistant strains. New resistance mechanisms have been reported recently: mutations of expression regulation factors [19] and toxicant-induced expression of efflux pumps [20 – 26].

Japanese researchers have made systematic characterization and DMI sensitivity bioassays of diverse strains of *Pen-*
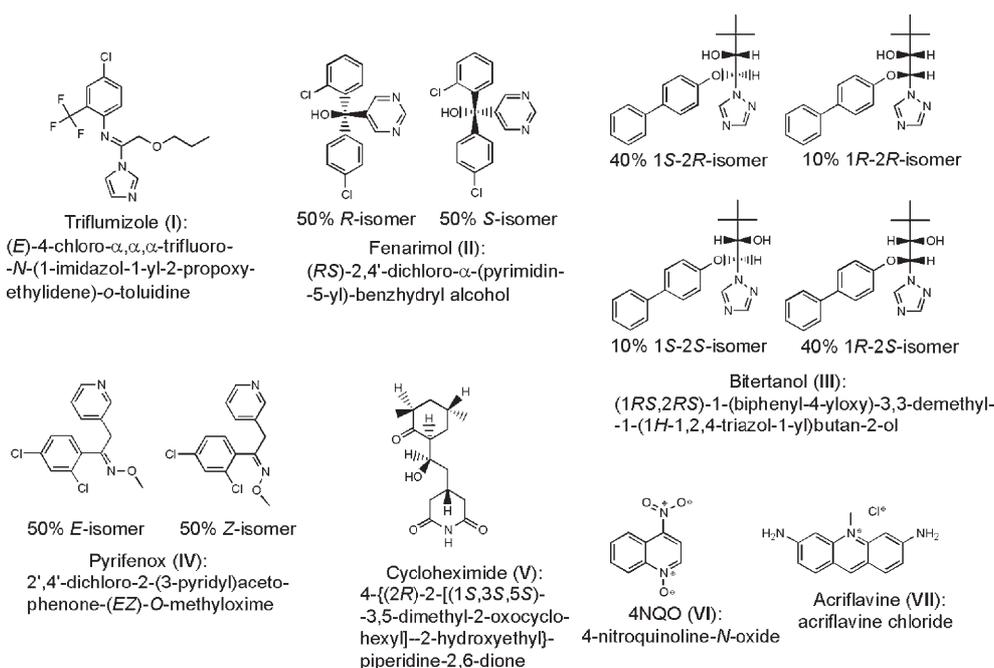
**Figure 1.** Structures, isomeric compositions, trivial, and IUPAC names of toxicants **I – VII**.

*icillium digitatum* [20 – 25], a phytopathogenic fungus (the green mold) that causes one of the most important post-harvest diseases of citrus fruits [27, 28]. They have reported two main resistance mechanisms: the CYP51 mechanism and the mechanism mediated by efflux pumps from the ATP-binding cassette (ABC) family of transporters [29 – 31]. In this work, resistance activities, and genetic structure of *P. digitatum* strains [20 – 24] related to seven toxicants **I – VII** (Figure 1) were explored by means of chemometric methods. **I–IV** are DMIs to which *P. digitatum* have already developed resistance [13], and **V – VII** are non-DMIs. Hierarchical Cluster Analysis (HCA), Principal Component Analysis (PCA), and Partial Least Squares (PLS) regression [32 – 35] were used to establish novel types of relationships: Activity–Structure Relationships (ASRs), Quantitative Genome–Activity Relationships (QGARs), and Quantitative Genome/Structure–Activity Relationships (QGSAR). This work gives more insight into resistance mechanisms in *P. digitatum*. The reported findings can be useful in the design of more potent DMIs and better exploration and design of DMIs sensitivity tests.

## 2 Methodology

### 2.1 Exploratory Analysis of the Datasets A and B

Biological (resistance) activities of *P. digitatum* strains against **I – VII** are defined in Table 1. The original data in standardized experiments were effective mass concentrations $EC_{50}$ and Minimal Inhibitory Concentration (MIC)

[20 – 24]. The former was averaged for seven strains whenever possible and transformed into $pEC_{50}$, forming the dataset A [matrix $(7 \times 7)$]. Effective Concentration for 50% ($EC_{50}$) and MIC from the same experiments [20 – 24] with a particular toxicant were used in the regression relationship between pMIC and $pEC_{50}$ (Table 1). New descriptors $a$, $b$, $c$, $|a|$, and $|c|$ for **I – VII** formed the dataset B [matrix $(5 \times 7)$]. PCA and HCA were performed for A and B. The data were autoscaled prior to all analyses in this work, and the complete linkage method were applied in all HCA. The analyses were carried out using Pirouette 3.11 [36] and Matlab 6.1 [37].

### 2.2 Exploratory Analysis of the Dataset C and Quantitative Relationships (QGAR and QGSAR)

The Japanese researchers have reported [20 – 25] two main mechanisms of the DMI-resistance of *P. digitatum*, both substantially determined or affected, directly or indirectly by fungal genome: (a) gene *CYP51* responsible for the expression of the $P450_{14DM}$ enzyme; (b) genes *PMR1*, *PMR3*, *PMR4*, and *PMR5* encoding multidrug efflux pumps PMR1, PMR3, PMR4, and PMR5, respectively, all belonging to ABC transporters. Six genome descriptors (Table 1) were generated from published data [20 – 24] and from the analysis of available pictorial results (photographs) [20 – 22] by applying an empirical graphics method with internal standards [38 – 41]. The descriptor PCR is the length of a promoter region of the *PdCYP51* gene: 0.25 kb for DMI-S and 0.75 kb for DMI-R (DMI-resistant) strains [23]. The promoter region corresponds to one or five units of a

**Table 1.** Definition of all biological activities and descriptors used in ASR, QGAR, and QGSAR analyses related to multidrug resistance of *P. digitatum*.

| Descriptor | Definition |
|---|---|
| Biological activities of fungal strains | |
| pMIC | $pMIC = -\log(MIC/mol/dm^3)$, where MIC is Minimal Inhibitory Concentration[a] |
| $pEC_{50}$ | $pEC_{50} = -\log(EC_{50}/mol/dm^3)$, where $EC_{50}$ is Effective Concentration for 50% inhibition[b] |
| $a$, $|a|$[c] | Free coefficient from regression equation $pMIC = a + b\,pEC_{50}$ |
| $b$[c] | Linear term coefficient from regression equation $pMIC = a + b\,pEC_{50}$ |
| $c$, $|c|$ | Ratio $c = a/b$ |
| Genome descriptors of fungal strains | |
| PMR1-g | Presence or absence of the native functional (nondisrupted) *PMR1* gene, or the presence of a *PMR1* gene from another plasmid |
| PMR1-e | Constitutive *PMR1* gene expression level (quantity of total RNA) in the absence of a toxicant, relative to DMI-S strains |
| PMR-t | *PMR1* expression level (quantity of total RNA) induced by a toxicant, triflumizole (0.5 μg/ml) |
| PCR | Length of the promoter fragment in the *PdCYP51* gene (the gene that encodes the target enzyme $P450_{14DM}$ of the DMIs in the strain PD5), corresponding to one or more copies of the *CYP51* transcriptional enhancer [23] |
| CYP51-e | Constitutive *CYP51* gene expression level (quantity of total RNA) in the absence of a toxicant, relative to DMI-S strains |
| CYP51-g | Number of the transcriptional enhancer copies in the *CYP51* gene |
| Molecular descriptors of DMIs **I – IV** | |
| Nπ[d] | Number of π-systems in a toxicant molecule |
| Lπ[e] | Maximum number of single bonds that separate these π-systems |
| Products of genome and molecular descriptors | |
| N*CYP51-g | Product of Nπ with CYP51-g |
| N*PCR | Product of Nπ with PCR |
| L*PMR1-e | Product of Lπ with PMR1-e |
| L*CYP51-e | Product of Lπ with CYP51-e |
| L*PCR | Product of Lπ with PCR |

[a] MIC, Minimal Inhibitory Concentration is the effective concentration at which no radial growth of a fungal culture was observed (100% radial growth inhibition).
[b] $EC_{50}$, Effective Concentration is the concentration inhibiting radial growth of a fungal culture by 50%.
[c] This regression equation was obtained for each of the seven toxicants **I – VII**.
[d] Nπ, number of π-systems in a DMI molecule, taking into account conjugation, separation, and steric hindrance between (hetero)aromatic rings and double bonds: **I**, 1; **II**, 3; **III**, 2; **Iv**, 3.
[e] Lπ, maximum number of single bonds that separate these π-systems: **I**, 1; **II**, – 3; **III** and **IV**, 3.

126 bp transcriptional enhancer in DMI-S and DMI-R strains, respectively. The dataset C1 (matrix $92 \times 6$ and $pEC_{50}$ vector $92 \times 1$) was formed from the six descriptors and activities obtained from 92 experiments for 24 diverse strains against **I – IV** [20–24]. The dataset C2 (matrix $29 \times 6$ and $pEC_{50}$ vector $29 \times 1$) was formed analogously for **V – VII** in interaction with *P. digitatum* strains. The dataset C, matrix $131 \times 6$ and $pEC_{50}$ vector $131 \times 1$, included C1 and C2. Products of molecular descriptors for **I – IV** with genome descriptors were added to C1 (dataset D). C, C1, C2, and D were used in PLS analysis to model activities $pEC_{50}$. C1 was analyzed by PCA and HCA to evaluate the contribution of CYP51 and PMR1 resistance mechanisms to the total fungal multidrug resistance.

### 2.3 Molecular Modeling of the Toxicants

**I – VI** and **VII** were modeled, molecular dynamics with manual conformational search in more complicated cases was performed, and the obtained conformers were opti-

mized at the PM3 semiempirical level using CHEM3D Ultra 6.0 [42]. **I – IV** have at least two hydrogen bonding (HB) acceptor sites (Figure 1) that can interact with receptors of the fungal resistance. Conformers of **I – IV** with two putative HB acceptors placed at the same side (*syn*-arrangement) or different sides (*anti*-arrangement) of the common fragment (interconnecting fragment with the five- or six-membered heterocycle) were modeled and optimized at the PM3 level within the CHEM3D package.

## 3 Results and Discussion

### 3.1 Exploratory Analysis of the Datasets A and B: ASRs

#### 3.1.1 Data Set A: Example I of ASRs

*Defining ASRs*: **I – VII** belong to three classes of compounds which interfere in different biochemical processes in *P. digitatum*, and thus undergo distinct fungal resistance
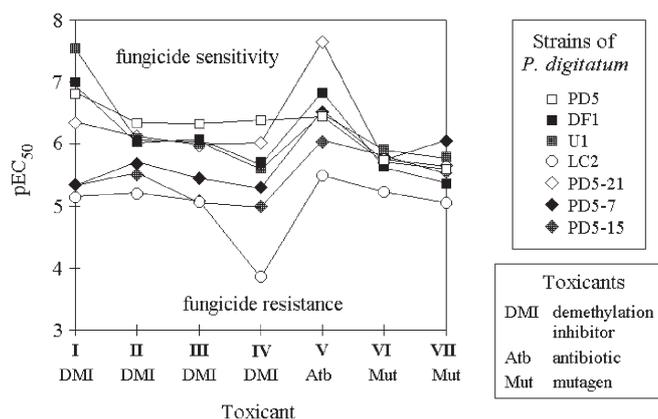
**Figure 2.** Comparison of the multidrug resistance activities $pEC_{50}$ of seven *P. digitatum* strains against toxicants **I – VII.**

mechanisms [20 – 22] (Table 2). **I – IV** are DMIs, used as fungicides against the cytochrome enzyme P450 sterol 14α-demethylase ($P450_{14DM}$) [12, 13]. **V** is an antibiotic (Atb) that interferes in the microbial protein synthesis [43], and **VI** and **VII** are mutagens (Mut) which act as effective nucleic acid intercalators and modifiers [44 – 46]. The strains have different baseline resistance (Figure 2): DMI-S strains PD5, DF1 and U1, field isolates; LC2 is a DMI-R strain, a field isolate; PD5-21 [abbreviated PD5(*PdCYP51*-P)-21], a mutant similar to PD5, also DMI-S; PD5-7 [abbreviated PD5(*PdCYP51*-L)-7] and PD5-15 [abbreviated PD5(*PdCYP51*-L)-15], mutants similar to LC2 with respect to DMIs and similar to PD5 with respect to non-DMIs, being DMI-moderately resistant (DMI-M). Correlations between the activities vary in large range (correlation coefficients are in the range of 0.16 – 0.98). The curves in Figure 2 are mutually nonparallel and vary from low (fungicide resistance, especially below 5) to high $pEC_{50}$ (fungicide sensitivity, especially above 7). The largest variations are for **I**, **IV**, and **V**, and smallest for Mut toxicants. DMI-S strains are distinguished from other strains in the case of DMI toxicants.

It has been shown in a previous work [47] that bacterial activities, when treated by PCA and HCA, could distinguish strains in accordance with their baseline resistance. Similarly, cluster analysis methods and PCA are used today extensively in taxonomy of living beings, taking into account functional, morphological, structural, phylogenetic, physiological, biochemical, and other characterizations of species and strains [48 – 53]. Such an exploratory analysis of biological activities that provide identification of structural patterns of toxicants or drugs is named in this work Activity-Structure Relationship (ASR). ASR is contrary to classical SAR which starts with a molecular structure (molecular descriptors) and ends with qualitative prediction of biological activities. Inverse QSAR and QSPR [52 – 57] require equations for calculation of activities from molecular descriptors, and by solving an inverse problem find out descriptors and candidate structures for

**Table 2.** The datasets A and B: multidrug resistance activities[a] of *P. digitatum* strains against toxicants **I – VII.**

Dataset A: Activities $pEC_{50}$

|  | PD5[b] | DF1[b] | U1[b] | LC2 | PD5-21 | PD5-7 | PD5-15 |
|---|---|---|---|---|---|---|---|
| **I** | 6.840 | 6.937 | 7.539 | 5.159 | 6.363 | 5.335 | 5.335 |
| **II** | 6.344 | 6.015 | 6.058 | 5.219 | 6.122 | 5.714 | 5.542 |
| **III** | 6.324 | 6.081 | 6.037 | 5.081 | 5.984 | 5.449 | 5.058 |
| **IV** | 6.391 | 5.664 | 5.595 | 3.847 | 6.023 | 5.294 | 4.979 |
| **V** | 6.449 | 6.847 | 6.495 | 5.500 | 7.657 | 6.546 | 6.064 |
| **VI** | 5.711 | 5.626 | 5.899 | 5.226 | 5.761 | 5.735 | 5.817 |
| **VII** | 5.595 | 5.373 | 5.771 | 5.053 | 5.636 | 6.068 | 5.528 |

Dataset B: LR descriptors

|  | $a$[c] | $b$[c] | $c$[c] | $|a|$[c] | $|c|$[c] |
|---|---|---|---|---|---|
| **I** | − 1.351 | 0.978 | − 1.383 | 1.351 | 1.383 |
| **II** | − 5.059 | 1.624 | − 3.116 | 5.059 | 3.116 |
| **III** | − 4.174 | 1.459 | − 2.861 | 4.174 | 2.861 |
| **IV** | 1.660 | 0.452 | 3.670 | 1.660 | 3.670 |
| **V** | − 0.002 | 0.822 | − 0.003 | 0.002 | 0.003 |
| **VI** | 2.048 | 0.438 | 4.666 | 2.048 | 4.666 |
| **VII** | − 6.224 | 1.937 | − 3.213 | 6.224 | 3.213 |

[a] The DMI resistance activity data referring to *P. digitatum* strains, based on experimental values of $EC_{50}$ [20 – 22]: DMI-S strains PD5, DF1, U1, and PD5-21; DMI-R strain LC2; DMI-M strains PD5-7 and PD5-15.
[b] Based on averaged experimental values of $EC_{50}$, taking into account the number of repetitions [20 – 22].
[c] LR descriptors for each toxicant, obtained from various measurements of $EC_{50}$ and MIC values with respect to diverse *P. digitatum* strains [20 – 22].

desired activity ranges. Therefore, these inverse approaches are different from the novel ASRs.

*ASRs Application: Strains Classification and Characterization*: PCA of the dataset A (Figure 3) shows that the first three principal components (PCs) describe 96.02% of the original data. The strains form three groups with respect to baseline resistance, in 2-D arrangement along PC1 and PC2 (Figure 3a): DMI-S strains at PC2 < 0; DMI-R and DMI-M at PC2 > 0, separated at about PC1 = 0.35. Moreover, the average $pEC_{50}$ values show a systematic decrease (sensitivity decrease or resistance increase) along positive PC1 and also along positive PC2, with an exception of the strain PD5-21. The HCA dendogram (Figure 3e) confirms the strains clustering from PCA. Both the PCA and HCA plots exhibit differences among the DMI-S strains, because of which the DMI-S cluster is not compact.

*ASRs Application: Identification of Toxicant Structure Types and Strain – Toxicant Interactions*: Toxicants can be classified into three groups with respect to overall molecular topology (Figures 3b and 3d): condensed ring systems (**CR**: **VI** and **VII**), two-ring linear systems (**2RL**: **I** and **V**), and three-ring nonlinear systems (**3RN**: **II**, **III**, and **IV** with bent topology that mimics a three-ring structure). PC2 distinguishes DMIs and non-DMIs (dashed line in Figure 3b). The HCA dendogram (Figure 3f) discriminates **2RL** from the mixed (**CR**, **3RN**) cluster. Comparing the
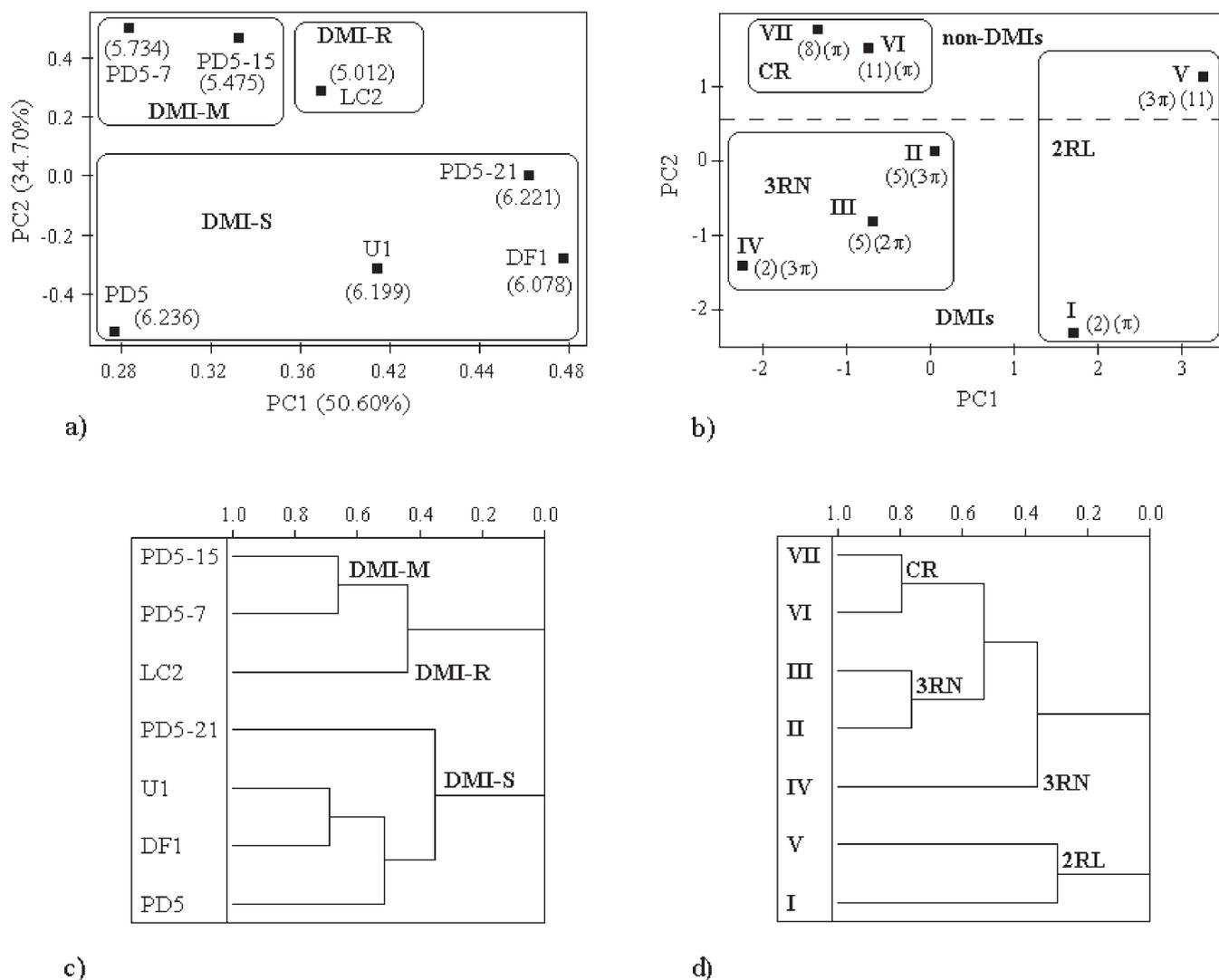
**Figure 3.** Exploratory analysis of the set of seven pEC$_{50}$ resistance activities: (a) PCA loadings plot with the first two PCs, showing the clustering of the *P. digitatum* strains and the average pEC$_{50}$ for each strain in brackets; (b) PCA scores plot with the first two PCs, showing the clustering of the toxicants with the number of hydrogen bonds and charge–charge interactions (integers in brackets) and the number of π-systems (π integers in brackets); (c) HCA dendogram for variables (the resistance activities) and the clustering of the *P. digitatum* strains; (d) HCA dendogram for the samples (toxicants) and their clustering.

loadings (Figure 3a) and scores (Figure 3b), it is noticeable that the baseline resistance (DMI-S, DMI-M, and DMI-S) arrangement from the loadings defines the samples clustering in the scores plot. That is why the DMIs **I – IV**, used in agricultural practice as pesticides, are placed in the central and lower parts of the scores plot, which corresponds to the DMI-S strains in the central and lower parts of the loadings plot. When toxicant – strain correlations are considered along PC1, the situation is the opposite due to different signs of the loadings (always positive) and scores (first negative and then positive) along this PC. Consequently, **I** is somewhat more effective against DMI-S strains (especially against PD5) than **II – IV** (**3RN** group). In the same sense, **V** is somewhat more effective than **VI** and **VII** (**CR** group) against PD5-7 and PD5-15 (DMI-M

strains). The position of the most resistant strain LC2 in the loadings does not correspond to any toxicant in the scores, meaning that among the studied toxicants, none showed a satisfactory growth inhibition of this strain.

PC1 distinguishes flexible from rigid toxicants: two-ring **V** lies at the positive end of PC1, whilst the most rigid **IV** (partially conjugated, delocalized, strained system) and **VII** (planar heteroaromatic system) are placed at the negative end of PC1. It can be said that PC1 increase is related to higher molecular flexibility and lower compactness (less rings, more pronounced linearity) in terms of scores. PC2 increase can be related to the elevated number of HB and charge-mediated intermolecular interactions (integer numbers in brackets in Figure 3b) due to increased hydrophilicity, polarity, and polarizability of the toxicants. In

terms of loadings (Figure 3a), it can be expected that DMI-M strains are more sensitive to hydrophilic toxicants and effective against lipophylic compounds, and the opposite is valid for DMI-S strains (although these strains express in average lower baseline resistance to all toxicants than DMI-R strains). The number of $\pi$-systems N$\pi$ (rings and double bonds enumerated in brackets, Figure 3b, see Table 1) in toxicant molecules can adopt different values with respect to the fungal strains. The $\pi$-system is continuous through conjugation in **I** and delocalization in **VI** and **VII**. Biphenyl in **III** is a unique $\pi$-system, whilst stereoelectronic effects in **IV** prevent the benzyl and imino groups to make a unique $\pi$-system. Rings in **II** are spatially close but separated by the central quaternary carbon. **V** contains three isolated double bonds as $\pi$-systems. N$\pi$ of **I** – **VII** indicates the importance of hydrophobic – hydrophobic and aromatic – aromatic interactions between toxicants and their receptors. N$\pi$ indicates the importance of molecular structure as the second factor (after the baseline fungal resistance) that defined the final fungal resistance/sensitivity with respect to the studied toxicants.

The presented ASR approach may be useful in predicting fungus – toxicant interactions based on known toxicant molecular structures and measured resistance activities of *P. digitatum* strains. The activities against an unknown toxicant, when treated by PCA and HCA, can provide useful information about the toxicant structure and properties. This is an item that should be further developed using results from various antifungal assays.

### 3.1.2 Data Set B: Exampled II of ASRs

*Defining (Unusual) ASRs*: Descriptors *a*, *b*, and *c* from linear regression (LR) relationship between pMIC and pEC$_{50}$ are illustrated for triflumizole (Figure F1 in Supplementary Material). The higher the pEC$_{50}$ value, the higher will be the pMIC, due to positive values of *b*. The EC$_{50}$ increase is correlated with a more pronounced MIC increase, as shown for strains LC2M, LC2, I2, and M1 (Figure F1). Since pMIC and pEC$_{50}$ are linearly related, some correlation of *a*, *b*, or *c* with molecular structure is expected.

*ASRs Application: Identification of Toxicant Structure Types*: PCA and HCA of B (Figure 4) exhibit such correlations as common for all considered strains. Two PCs comprise 99.36% of the original data. The toxicants are distinguished according to the number of rings in the HCA dendogram (Figure 4c) and scores plot (Figure 4a): the cluster of molecules with three rings (**3R**), and the other of molecules with two rings (**2R**). In fact, **2R** is less compact than **3R**, consisting of two subclusters: **IV** and **VI** are more compact molecules (**VI**: condensed heteroaromatic system; **IV**: nonlinear strained system), whilst **I** and **V** are more linear molecules (**I**: linear conjugated and aromatic system; **V**: mainly saturated linear system). Condensed **VII** is distinguished from the subcluster of noncondensed **II** and **III**.

The toxicant structures arranged along the PC1 increase (**VII**→**II**→**III**→**I**→**V**→**IV**→**VI**) show a decrease in molecular complexity (size and compactness). On the other hand, the PC2 increase is followed by the increase in size-independent molecular compactness, from saturated **V** and conjugated **I** at PC2>0 to condensed **VI** and **VII** at PC2>0.

*ASRs Application: Identification of Toxicant – Strain Interactions*: Relationships among the descriptors in the loadings plot (Figure 4b) and HCA dendogram (Figure 4d) can be considered as three cases of the LR line in the pEC$_{50}$-pMIC space (Figure F1b with comments, Supplementary Material). Descriptors *a* and *c* are similar (axial intercepts) and form a subcluster. Positive *b* and |*a*| make another cluster, whilst |*c*| is practically isolated. First case: The PC1 increase is followed by the *b* decrease and positive signs of *a* and *c*. Small PC1 is related to small values of *b* and negative signs of *a* and *c*. The PC2 increase means an increase in absolute values of *a* and *c*. Therefore, the subcluster (**I**, **V**) is characterized by small values of all descriptors and negative *a* and *c*. Second case: The subcluster (**IV**, **VI**) lies at small values of *b* and positive *a* and *c*. Third case: The cluster (**II**, **III**, **VII**) is placed at large negative *a* and *c*, and large *b*. These cases determine the dose – response curve profiles for a toxicant with respect to *P. digitatum* strains. These observations can be useful in the design of sensitivity assays and more potent toxicants, taking into account that the equation pMIC = *a* + *b* pEC$_{50}$ can be written in another form EC$_{50}$ = *e*(MIC)$^{1/b}$ where *e* = 10$^{a/b}$. According to the latter form and Figure F1, effective toxicants are those for which MIC and EC$_{50}$ are small and consequently, the difference between MIC and EC$_{50}$ is also small. The last condition seems to be fulfilled for toxicants for which *b* tends to be *b* = 1 and *e* is a small number and *e* > 1.

### 3.2 Exploratory Analysis of the Data Set C and Quantitative Relationships (QGAR and QGSAR)

Genome descriptors and pEC$_{50}$ for DMIs are in Table 3 (dataset C1) and Table T1 in Supplementary Material (dataset C2). PMR1-g was set to 1 and 0 for the presence and absence of the native *PMR1* gene, respectively, and 0.5 for the presence of *PMR1* from another plasmid (supposing that the introduced *PMR1* could result in reduced resistance). Average relative expression level of the *PMR1* gene (PMR1-e) was set to 1 and 7 for DMI-S and DMI-R strains, respectively, and zero for strains without any functional *PMR1*. The number of the *CYP51* transcriptional enhancer copies (CYP51-g) was set to 1, 2, and 5 for DMI-S, DMI-M, and DMI-R strains, respectively [21, 23]. The analog values for PCR were 0.25, 0.37, and 0.75, respectively [23]. The *CYP51* gene expression level (CYP51-e) was set to 100 and 1 for DMI-R and DMI-S strains, respectively, and in some cases had other values (3 in four experiments, and 50 in another one). Triflumizole-induced
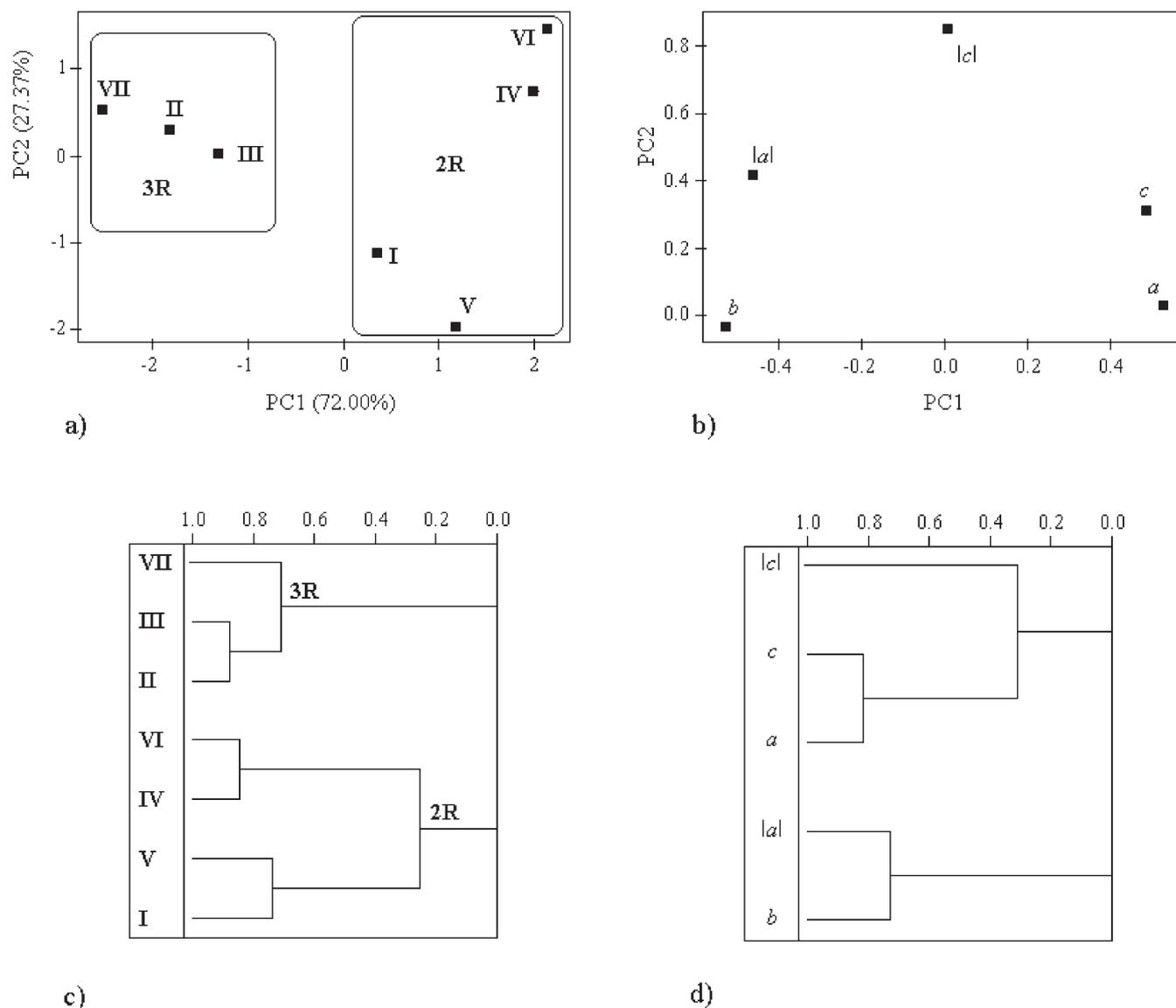
**Figure 4.** Exploratory analysis of the set of five LR descriptors ($a$, $b$, $c$, $|a|$, and $|c|$): (a) PCA scores plot with the first two PCs, showing the clustering of the toxicants; (b) PCA scores plot with the first two PCs, showing the clustering of the variables; (c) HCA dendogram for the samples (toxicants) and their clustering; (d) HCA dendogram for variables.

*PMR1* expression level (PMR1-t) was set to 100 (high), 10 (low), and 1 (none), independently of the baseline resistance. Two toxicant molecular descriptors, $N\pi$ and $L\pi$, were generated, and products of these descriptors with genome descriptors were manually selected in PLS variable selection (Table 1), forming the dataset D (descriptor matrix $92 \times 8$ and $pEC_{50}$ vector $92 \times 1$). It has been noticed in the first presented ASR plots (Figure 3b) that $N\pi$ is a factor that can affect fungal resistance to **I–VII**.

PLS models in QGARs and QGSARs are in Tables 3–5. The models I–III reconfirm the existence of two toxicant groups, **I–IV** and **V–VII**, which should be modeled separately. The model II for **IV–VII** has no acceptable statistics, probably because of the narrow range of $pEC_{50}$

values (0.8–0.9 in log units) and the fact that CYP51 descriptors do not have biochemical justification to describe the behavior of non-DMIs. The model IV for **I–IV** has satisfactory statistics, probably because of wide range of $pEC_{50}$ values (greater than 3.1 in log units), reasonable correlation coefficients (absolute values $\geq 0.5$, Table 5), and removed outliers (Table 4). Three outliers with too low resistance ($pEC_{50} > 6.8$) are related to **I** in interaction with DMI-S strains, whilst the other three outliers with too high resistance ($pEC_{50} < 4.3$) are related to **IV** in interaction with DMI-R strains. The model IV was externally validated (resulting in the model V) by dividing the dataset C1 into training and prediction sets, the latter being formed from arbitrarily selected samples (35% samples) in

**Table 3.** Genome descriptors,[a] experimental,[b] and predicted[c] resistance activities $pEC_{50}$ against DMIs (dataset C1).

| No. | Toxicant/strain[d] | PMR1-g | PMR1-e | CYP51-g | CYP51-e | PCR | PMR1-t | $pEC_{50exp}$ | $pEC_{50cal}$ |
|---|---|---|---|---|---|---|---|---|---|
| 1 | T/PD5 | 1 | 1 | 1 | 1 | 0.25 | 100 | 6.636 | 6.871 |
| 2 | T/DF1 | 1 | 1 | 1 | 1 | 0.25 | 100 | 6.937 | 6.781 |
| 3[e] | T/U1 | 1 | 1 | 1 | 1 | 0.25 | 100 | 7.539 | – |
| 4 | T1/LC2 | 1 | 7 | 5 | 100 | 0.75 | 100 | 5.225 | 5.184 |
| 5 | T/M1 | 1 | 7 | 5 | 100 | 0.75 | 100 | 5.432 | **5.184** |
| 6 | T/I1 | 1 | 7 | 5 | 100 | 0.75 | 100 | 5.093 | 5.184 |
| 7 | T/DIS03 | 0 | 0 | 5 | 100 | 0.75 | 1 | 5.895 | 6.093 |
| 8 | T/DIS33 | 0 | 0 | 5 | 100 | 0.75 | 1 | 5.971 | **6.093** |
| 9 | T/DIS96 | 0 | 0 | 5 | 100 | 0.75 | 1 | 5.858 | 6.093 |
| 10 | T/PD5 | 1 | 1 | 1 | 1 | 0.25 | 100 | 6.840 | 6.781 |
| 11 | T/PD5 – 21 | 1 | 1 | 1 | 3 | 0.25 | 100 | 6.363 | **6.871** |
| 12 | T/PD5 – 7 | 1 | 1 | 5 | 100 | 0.75 | 100 | 5.335 | 5.348 |
| 13 | T/PD5 – 15 | 1 | 1 | 5 | 100 | 0.75 | 100 | 5.335 | 5.348 |
| 14 | T/LC2 | 1 | 7 | 5 | 100 | 0.75 | 100 | 5.141 | **5.184** |
| 15 | T/PD5 | 1 | 1 | 1 | 1 | 0.25 | 100 | 6.948 | 6.781 |
| 16 | T/DISp12 | 0 | 0 | 1 | 1 | 0.25 | 1 | 7.694 | **7.526** |
| 17 | T/DISp21 | 0 | 0 | 1 | 1 | 0.25 | 1 | 7.694 | 7.526 |
| 18 | T/ECTp36 | 0 | 0 | 1 | 1 | 0.25 | 1 | 7.047 | **7.526** |
| 19 | T/LC2 | 1 | 7 | 5 | 100 | 0.75 | 100 | 5.146 | 5.184 |
| 20 | T/DIS33 | 0 | 0 | 5 | 100 | 0.75 | 1 | 6.177 | 6.093 |
| 21 | T/DIS33-Y4 | 0.5 | 7 | 5 | 100 | 0.75 | 100 | 5.641 | **5.184** |
| 22 | T/DIS33-Y8 | 0.5 | 7 | 5 | 100 | 0.75 | 100 | 5.179 | 5.184 |
| 23 | T/DIS33-B0 | 0.5 | 7 | 5 | 100 | 0.75 | 10 | 6.007 | 5.836 |
| 24 | T/DIS33-B13 | 0.5 | 7 | 5 | 100 | 0.75 | 10 | 6.124 | **5.836** |
| 25[e] | T/PD5-B1 | 0.5 | 7 | 5 | 100 | 0.75 | 10 | 6.937 | – |
| 26[e] | T/PD5-B2 | 0.5 | 7 | 5 | 100 | 0.75 | 10 | 6.840 | – |
| 27 | T/DIS33 | 0 | 0 | 5 | 100 | 0.75 | 1 | 6.283 | 6.093 |
| 28 | T/LC2 | 1 | 7 | 5 | 100 | 0.75 | 100 | 5.107 | **5.184** |
| 29 | T/PD5 | 1 | 1 | 1 | 1 | 0.25 | 100 | 6.937 | 6.781 |
| 30 | T/LC2 M | 1 | 7 | 2 | 50 | 0.37 | 100 | 5.840 | 6.270 |
| 31 | T/PD5 | 1 | 1 | 1 | 1 | 0.25 | 100 | 6.937 | **6.781** |
| 32 | T/DIS33 | 0 | 0 | 5 | 100 | 0.75 | 1 | 6.177 | 6.093 |
| 33 | T/LC2 | 1 | 7 | 5 | 100 | 0.75 | 100 | 5.189 | 5.184 |
| 34 | T/DISp21 | 0 | 0 | 1 | 1 | 0.25 | 1 | 7.539 | 7.526 |
| 35 | T/DIS5-L22 | 1 | 7 | 5 | 100 | 0.75 | 100 | 5.202 | **5.184** |
| 36 | T/DIS5-P26 | 1 | 1 | 1 | 1 | 0.25 | 100 | 6.937 | 6.781 |
| 37 | F/PD5 | 1 | 1 | 1 | 1 | 0.25 | 100 | 6.073 | 6.263 |
| 38 | F/DF1 | 1 | 1 | 1 | 1 | 0.25 | 100 | 6.015 | **6.263** |
| 39 | F/U1 | 1 | 1 | 1 | 1 | 0.25 | 100 | 6.058 | 6.263 |
| 40 | F/LC2 | 1 | 7 | 5 | 100 | 0.75 | 100 | 5.160 | 5.035 |
| 41 | F/M1 | 1 | 7 | 5 | 100 | 0.75 | 100 | 5.093 | **5.035** |
| 42 | F/I1 | 1 | 7 | 5 | 100 | 0.75 | 100 | 4.799 | 5.035 |
| 43 | F/PD5 | 1 | 1 | 1 | 1 | 0.25 | 100 | 6.073 | 6.263 |
| 44 | F/DIS07 | 0 | 0 | 5 | 100 | 0.75 | 1 | 6.002 | **6.136** |
| 45 | F/DIS33 | 0 | 0 | 5 | 100 | 0.75 | 1 | 6.073 | 6.136 |
| 46 | F/DIS96 | 0 | 0 | 5 | 100 | 0.75 | 1 | 5.952 | 6.136 |
| 47 | F/PD5 | 1 | 1 | 1 | 1 | 0.25 | 100 | 6.479 | **6.263** |
| 48 | F/PD5 – 21 | 1 | 1 | 1 | 3 | 0.25 | 100 | 6.122 | 6.295 |
| 49 | F/PD5 – 7 | 1 | 1 | 5 | 100 | 0.75 | 100 | 5.714 | 5.363 |
| 50 | F/PD5 – 15 | 1 | 1 | 5 | 100 | 0.75 | 100 | 5.542 | **5.363** |
| 51 | F/LC2 | 1 | 7 | 5 | 100 | 0.75 | 100 | 5.241 | 5.035 |
| 52 | F/PD5 | 1 | 1 | 1 | 1 | 0.25 | 100 | 6.475 | 6.263 |
| 53 | F/DISp12 | 0 | 0 | 1 | 1 | 0.25 | 1 | 7.219 | **7.036** |
| 54 | F/DISp21 | 0 | 0 | 1 | 1 | 0.25 | 1 | 6.520 | 7.036 |
| 55 | F/ECTp36 | 0 | 0 | 1 | 1 | 0.25 | 1 | 6.441 | **7.036** |
| 56 | F/LC2 | 1 | 7 | 5 | 100 | 0.75 | 100 | 5.230 | 5.035 |
| 57 | F/DIS33 | 0 | 0 | 5 | 100 | 0.75 | 1 | 6.241 | 6.136 |
| 58 | B/PD5 | 1 | 1 | 1 | 1 | 0.25 | 100 | 6.186 | **6.314** |
| 59 | B/DF1 | 1 | 1 | 1 | 1 | 0.25 | 100 | 6.081 | 6.314 |
| 60 | B/U1 | 1 | 1 | 1 | 1 | 0.25 | 100 | 6.037 | **6.314** |
| 61 | B/LC2 | 1 | 7 | 5 | 100 | 0.75 | 100 | 5.024 | 4.967 |

**Table 3.** (cont.)

| No. | Toxicant/strain[d] | PMR1-g | PMR1-e | CYP51-g | CYP51-e | PCR | PMR1-t | $pEC_{50exp}$ | $pEC_{50cal}$ |
|---|---|---|---|---|---|---|---|---|---|
| 62 | B/MI | 1 | 7 | 5 | 100 | 0.75 | 10 | 5.093 | **4.967** |
| 63 | B/I1 | 1 | 7 | 5 | 100 | 0.75 | 100 | 4.546 | 4.967 |
| 64 | B/DIS07 | 0 | 0 | 5 | 100 | 0.75 | 1 | 6.227 | **6.260** |
| 65 | B/DIS33 | 0 | 0 | 5 | 100 | 0.75 | 1 | 6.298 | 6.260 |
| 66 | B/DIS96 | 0 | 0 | 5 | 100 | 0.75 | 1 | 6.148 | **6.260** |
| 67 | B/LC2 | 1 | 7 | 5 | 100 | 0.75 | 100 | 5.024 | 4.967 |
| 68 | B/PD5 | 1 | 1 | 1 | 1 | 0.25 | 100 | 6.382 | **6.314** |
| 69 | B/PD5-21 | 1 | 1 | 1 | 3 | 0.25 | 100 | 5.984 | 6.344 |
| 70 | B/PD5-7 | 1 | 1 | 5 | 100 | 0.75 | 100 | 5.449 | 5.460 |
| 71 | B/PD5-15 | 1 | 1 | 5 | 100 | 0.75 | 100 | 5.324 | **5.460** |
| 72 | B/LC2 | 1 | 7 | 5 | 100 | 0.75 | 100 | 5.113 | 4.967 |
| 73 | B/PD5 | 1 | 1 | 1 | 1 | 0.25 | 100 | 6.424 | **6.314** |
| 74 | B/DISp12 | 0 | 0 | 1 | 1 | 0.25 | 1 | 7.324 | 7.114 |
| 75 | B/DISp21 | 0 | 0 | 1 | 1 | 0.25 | 1 | 7.324 | 7.114 |
| 76 | B/ECTp36 | 0 | 0 | 1 | 1 | 0.25 | 1 | 6.503 | **7.114** |
| 77 | B/LC2 | 1 | 7 | 5 | 100 | 0.75 | 100 | 5.112 | 4.967 |
| 78 | B/DIS33 | 0 | 0 | 5 | 100 | 0.75 | 1 | 6.382 | 6.260 |
| 79 | P/PD5 | 1 | 1 | 1 | 1 | 0.25 | 100 | 5.902 | 6.125 |
| 80 | P/DF1 | 1 | 1 | 1 | 1 | 0.25 | 100 | 5.664 | 6.125 |
| 81 | P/U1 | 1 | 1 | 1 | 1 | 0.25 | 100 | 5.595 | **6.125** |
| 82 | P/PD5 | 1 | 1 | 1 | 1 | 0.25 | 100 | 6.993 | 6.125 |
| 83 | P/PD5-21 | 1 | 1 | 1 | 3 | 0.25 | 100 | 6.023 | 6.165 |
| 84 | P/PD5-7 | 1 | 1 | 5 | 100 | 0.75 | 100 | 5.294 | **5.432** |
| 85 | P/PD5-15 | 1 | 1 | 5 | 100 | 0.75 | 100 | 4.979 | 5.432 |
| 86[e] | P/LC2 | 1 | 7 | 5 | 100 | 0.75 | 100 | 3.847 | – |
| 87 | P/PD5 | 1 | 1 | 1 | 1 | 0.25 | 100 | 6.993 | 6.125 |
| 88 | P/DISp12 | 0 | 0 | 1 | 1 | 0.25 | 1 | 7.023 | 6.925 |
| 89 | P/DISp21 | 0 | 0 | 1 | 1 | 0.25 | 1 | 7.039 | **6.925** |
| 90 | P/ECTp36 | 0 | 0 | 1 | 1 | 0.25 | 1 | 6.685 | 6.925 |
| 91[e] | P/LC2 | 1 | 7 | 5 | 100 | 0.75 | 100 | 3.842 | – |
| 92[e] | P/DIS33 | 0 | 0 | 5 | 100 | 0.75 | 1 | 4.253 | – |

[a] Six genome descriptors as reported in the literature [20–24], and estimated from these reports in cases of missing values.
[b] The $pEC_{50exp}$ activities from individual DMI sensitivity bioassays for 24 *P. digitatum* strains (wild types and mutants).
[c] The $pEC_{50cal}$ as predicted by PLS Model VII. Bold values are for the external validation set.
[d] Toxicants: T, triflumizole (**I**); F, fenarimol (**II**); B, bitertanol (**III**); and P, pyrifenox (**IV**).
[e] The six samples excluded from the final PLS modeling, because of absolute. Student residuals > 2.

such a way that representation of toxicants **I – IV** and strains with a different baseline resistance was proportional to that of the former. The model V has good statistics (Table 4). The regression vector of the model IV (Table 5) shows that the DMI resistance of *P. digitatum* is more consequence of the *CYP51* expression and less of the *PMR1* expression, which can agree with observations from Hamamoto *et al.* [21, 22] that the baseline resistance of *P. digitatum* strains is not determined by the *PMR1* expression or overexpression but by that of *CYP51*. New variable selection with genome and mixed descriptors resulted in an improved QGSAR model VI, which was successfully validated externally in the same way as the model IV (resulting in the model VII with good statistics, Tables 3 and 4). The mixed genome-toxicant descriptors in the regression vector indicate interactions between the fungal genome and DMIs, more probably with the expression of *CYP51* than of *PMR1* (Table 6). The increase in *P. digitatum* resistance, *i.e.*, decrease in $pIC_{50}$ is accompanied by negative contributions of all genome descriptors in the model IV,

meaning higher expression levels of the resistance mechanisms. All descriptors in the model VI positively contribute to the fungal resistance.

PCA and HCA on the dataset C1 (Figure 5) showed three PCs describing 99.51% of the original data and eight clusters of strains. PC1 and PC2 discriminate DMI-S, DMI-M, and DMI-R strains. CYP51 descriptors form a tight cluster and are major contributors to PC1 and minor to PC2. Two PMR1 descriptors (PMR1-g and PMR1-t) form another cluster with a minor contribution to PC1 and major to PC2. PMR1-e is isolated and has the major contribution to PC3 (not shown), and moderate to PC1 and PC2. It represents the *PMR1* gene overexpression not induced by toxicant and not depending on the number of the *PMR1* gene enhancer copies (there is only one copy in all *P. digitatum* strains), being an intrinsic property of *P. digitatum*. A new classification of *P. digitatum* strains (Figures 5a and 5c, Table 6) is proposed. S1 class contains the most DMI-susceptible strains with no functional *PMR1* gene and only one copy of the *CYP51* gene enhancer. The

**Table 4.** Comparison of PLS models with basic statistics.

| Parameters | Model I | Model II | Model III | Model IV | Model V | Model VI | Model VII |
|---|---|---|---|---|---|---|---|
| Relationship | QGAR | QGAR | QGAR | QGAR | QGAR | QGSAR | QGSAR |
| Dataset | C | C2 | C1 | C1 | C1 | D | D |
| Toxicants | **I – VII** | **V – VII** | **I – IV** | **I – IV** | **I – IV** | **I – IV** | **I – IV** |
| No. strains[a] | 24 | 14 | 24 | 22 | 22 | 22 | 22 |
| Descriptors[b] | 6 | 6 | 6 | 6 | 6 | 8 | 8 |
| Training set[c] | 131 | 39 | 92 | 86 | 56 | 86 | 56 |
| Ext. val. set[c] | 0 | 0 | 0 | 0 | 30 | 0 | 30 |
| Excluded[c] | 0 | 0 | 0 | 6 | 6 | 6 | 6 |
| PCs (%Var)[d] | 4(99.98%) | 2(84.24%) | 3(98.66%) | 3(99.45%) | 3(99.34%) | 5(96.78%) | 5(97.14%) |
| SEV[e] | 0.521 | 0.477 | 0.487 | 0.340 | 0.352 | 0.286 | 0.305 |
| SEP[e] | 0.509 | 0.462 | 0.475 | 0.333 | 0.341 | 0.271 | 0.279 |
| $Q^{2f}$ | 0.521 | 0.176 | 0.655 | 0.789 | 0.788 | 0.851 | 0.841 |
| $R^{2f}$ | 0.561 | 0.285 | 0.686 | 0.807 | 0.815 | 0.874 | 0.881 |
| $Q^2_{ext}$ [f] | – | – | – | – | 0.771 | – | 0.843 |
| R.e. $\geq$ 10.00%[g] | 25 | 6 | 11 | 4 | 3 (3) | 2 | 2 (0) |
| Max. R.e.[g] | 44.37% | 12.71% | 41.95% | 13.58% | 13.24% (14.64%) | 13.13% | 12.42% (8.83%) |
| Mean R.e.[g] | 6.41% | 6.60% | 5.60% | 3.99% | 4.25% (3.67%) | 3.31% | 3.23% (3.45%) |

[a] Number of *P. digitatum* strains related to the samples of particular datasets.
[b] QGAR models: six genome descriptors. QGSAR models: three genome descriptors and five products of molecular and genome descriptors.
[c] Number of samples in the training set, external validation set, and excluded samples (outliers).
[d] PCs, number of PCs (latent variables) used in the model, and %Var, percentage of the total variance contained in these PCs (latent variables).
[e] Standard deviations: SEV, Standard Error of Validation; SEP, Standard Error of Prediction.
[f] Correlation coefficients: $Q^2$, correlation coefficient of validation; $R^2$, correlation coefficient of prediction, $Q^2_{ext}$, correlation coefficient of external validation.
[g] Relative errors: R.e. $\geq$ 10.00%, number of samples with relative error $\geq$ 10.00%, Max. R.e., maximum relative error, mean R.e., mean relative error. These parameters are given in brackets for the external validation set.

**Table 5.** Regression vector components and correlation coefficients for the best models QGAR (model IV) and QGSAR (model VI).

| Descriptor | QGAR Model IV | | QGSAR Model VI | |
|---|---|---|---|---|
| | Regression vector | Correlation coefficient | Regression vector | Correlation coefficient |
| PMR1-g | 0.217 | 0.491 | – | – |
| PMR1-e | 0.127 | 0.694 | – | – |
| CYP51-g | 0.224 | 0.722 | − 0.540 | 0.722 |
| CYP51-e | 0.225 | 0.726 | − 0.594 | 0.726 |
| PCR | 0.224 | 0.722 | – | – |
| PMR1-t | 0.248 | 0.502 | − 0.447 | 0.502 |
| N*CYP51-g | – | – | 0.782 | 0.624 |
| N*PCR | – | – | − 0.620 | 0.556 |
| L*PMR1-e | – | – | − 0.178 | 0.677 |
| L*CYP51-e | – | – | 0.667 | 0.634 |
| L*PCR | – | – | − 0.485 | 0.564 |

**Table 6.** Classification of the strains into eight groups, based on six genome descriptors for the dataset C1.

| Group[a] | PMR1-g | PMR1-e | CYP51-g | CYP51-e | PCR | PMR1-t | DMI-R[b] |
|---|---|---|---|---|---|---|---|
| S1 | 0 | 0 | 1 | 1 | 0.25 | 1 | S |
| S2 | 1 | 1 | 1 | 1 | 0.25 | 100 | S |
| MR1 | 1 | 7 | 2 | 50 | 0.37 | 100 | M |
| MR2 | 0 | 0 | 5 | 100 | 0.75 | 1 | M |
| R1 | 1 | 1 | 5 | 100 | 0.75 | 100 | R |
| R2 | 0.5 | 7 | 5 | 100 | 0.75 | 10 | R |
| R3 | 0.5 | 7 | 5 | 100 | 0.75 | 100 | R |
| R4 | 1 | 7 | 5 | 100 | 0.75 | 100 | R |

[a] Groups with different DMI resistance: DMI-S strains (S1, S2), DMI-M strains (MR1, MR2), and DMI-R strains (R1, R2, R3, R4).
[b] DMI resistance of the strains: DMI-S (S), DMI-M (M), and DMI-R (R).
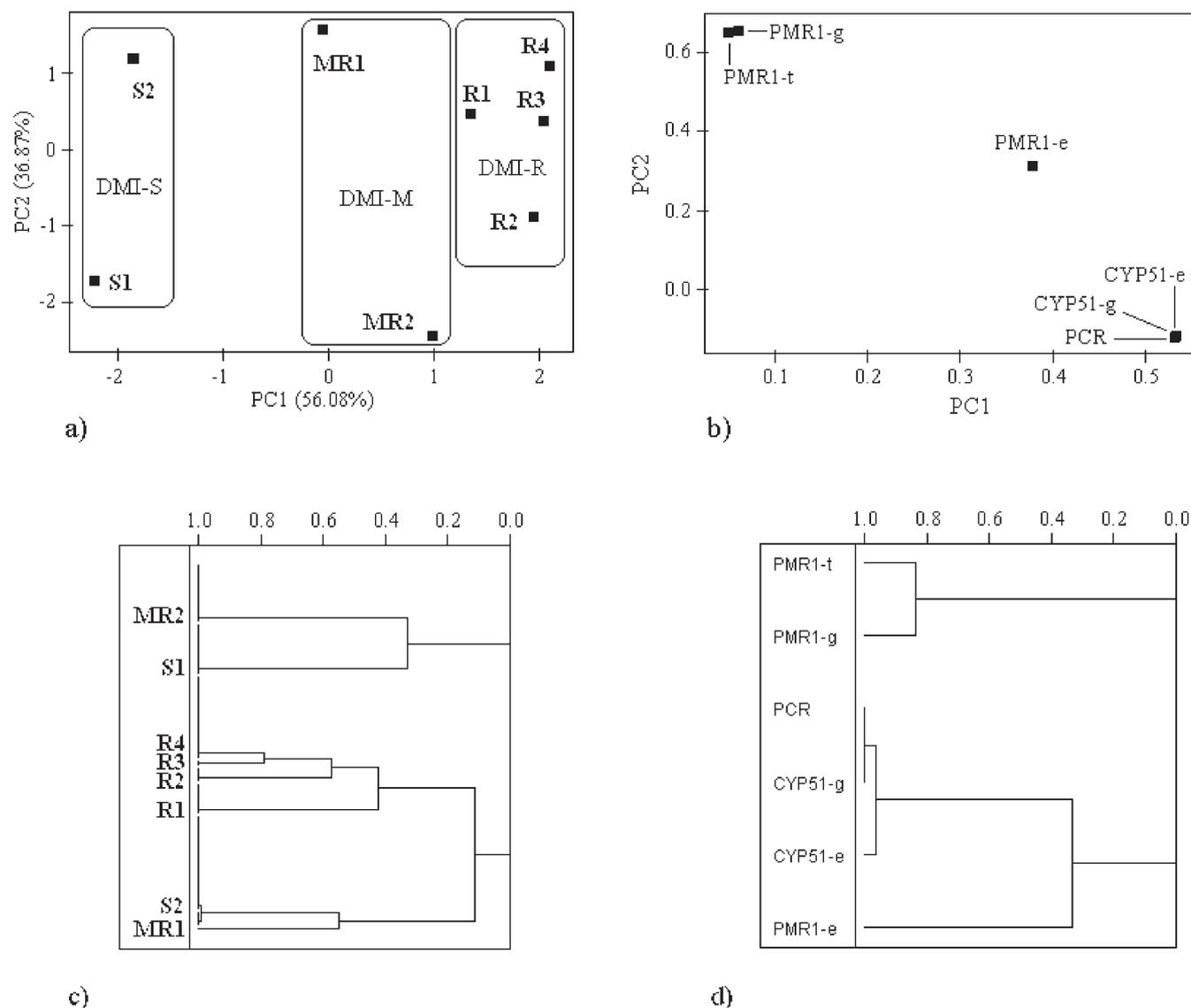
**Figure 5.** Exploratory analysis of the six genome variables (dataset C1): (a) PCA scores plot with the first two PCs, showing the clustering of the *P. digitatum* strains; (b) PCA loadings plot with the first two PCs, showing the clustering of the genome variables; (c) HCA dendogram for samples, showing eight cases of the variable combinations; (d) HCA dendogram for variables.

most DMI-resistant class R4 has completely functional *PMR1* and five copies of the *CYP51* enhancer. The total fungal resistance is mainly determined by the CYP51 and less by the PMR1 mechanism.

### 3.3 Toxicant Molecular Structures

The structures of **I–VII** (Figure 1) contain information on multiple bonds and stereochemistry which can provide sufficient three-dimensional structural information [58]. It is important to note that the structures of minimum energy conformers of all isomers of **I–VII** (Figures F2 and F3 with comments in Supplementary Material) agree with the structural patterns from ASRs (branching, HB potency).

Comparative (Figure 2) and exploratory analyses (Figure 3b) of biological activities also pointed out differences between DMIs and non-DMIs. The conformers with *syn-* and *anti*-arrangement of putative hydrogen bonds of all conformers of **I–IV** exhibit characteristic structural patterns when overlapped (Figure F3 in Supplementary Material with comments). The heterocyclic HB acceptor atom (a nitrogen atom) and other nitrogen (HB acceptor) or oxygen (HB donor/acceptor) atom in the branching region are approximately coplanar with the heterocycle. The *syn*-arrangement has more suitable HB and π-rings geometry for interaction with putative receptors such as efflux pumps or cytochromes. Two crystal structures of a DMI (fluconazole) complexed with CYP51 (PDB:1EA1 [7] pre-

sented in Figure F4 in Supplementary Material and PDB:2IJ7 [59]) confirm this observation.

Triflumizole (**I**) induces substantially the *PMR1* gene expression in all *P. digitatum* strains with this gene in a nondisrupted form [20, 22, 24], but poorly in the *CYP51* gene expression [21]. Other DMIs could provoke similar effects due to their interaction with fungal transcription factors which regulate cellular morphology [60, 61] and ABC efflux pumps production [62]. Some crystal structures of drug/toxicant – transcription factor complexes related to bacterial efflux pumps [63] suggest that future modeling of antifungals should be directed toward diverse targets such as CYP51 (DMIs), ABC efflux pumps (reversals, modifiers [64, 65]), and transcription factors [66].

## 4 Conclusions

ASRs and QGARs classified *P. digitatum* strains and toxicants with respect to structure and characteristics. The strains were discriminated according to their baseline resistance, depending mainly on the CYP51 resistance mechanism, and less on the PMR1 resistance mechanisms. Toxicants were distinguished according to function (demethylation inhibition) and structure (flexibility, branching, topology, compactness). The regression models showed satisfactory prediction power which makes them useful for practical purposes. Common three-dimensional distribution of HB and aromatic rings in DMIs can determine their interactions with receptors such as cytochromes, efflux pumps, and transcriptional factors which activate fungal multidrug resistance.

## Acknowledgements

**Supplementary Material Available.** Figures showing relationships between biological activities, structures of toxicant conformers and of a toxicant-receptor complex, a table with genome descriptors, and text discussing the figures. This material is free of charge via the Internet at http://www3.interscience.wiley.com/.

## References

[1] G. I. Lepesheva, M. R. Waterman, *Mol. Cell. Endocrinol.* **2004**, *215*, 165 – 170.

[2] J. A. Zarn, B. J. Brüschweiler, J. R. Schlatter, *Environ. Health Perspec.* **2003**, *111*, 255 – 261.

[3] M. A. Ghannoun, L. B. Rice, *Clin. Microbiol. Rev.* **1999**, *12*, 501 – 517.

[4] T. C. White, K. A. Marr, R. A. Bowden, *Clin. Microbiol. Rev.* **1998**, *11*, 382 – 402.

[5] K. J. McLeon, K. R. Marshall, A. Richmond, I. S. Hunter, K. Fowler, T. Kieser, S. S. Gurcha, G. S Besra, A. W. Munro, *Microbiology* **2002**, *148*, 2937 – 2949.

[6] S. Ekins, M. J. de Groot, J. P. Jones, *Drug Metab. Dispos.* **2001**, *29*, 936 – 944.

[7] L. M. Podust, T. L. Poulos, M. R. Waterman, *Proc. Natl. Acad. Sci. USA* **2001**, *98*, 3068 – 3073.

[8] L. M. Podust, L. V. Yermalitskaya, G. I. Lepesheva, V. N. Podust, E. A. Dalmasso, M. R. Waterman, *Structure* **2004**, *12*, 1937 – 1945.

[9] D. C. Lamb, M. Cannieux, A. G. S. Warrilow, S. Bak, R. A. Kahn, N. J. Manning, D. E. Kelly, S. L. Kelly, *Biochem. Biophys. Res. Commun.* **2001**, *284*, 845 – 849.

[10] A. Tafi, R. Costi, M. Botta, R. Di Santo, F. Corelli, S. Massa, A. Ciacci, F. Manetti, M. Artico, *J. Med. Chem.* **2002**, *45*, 2720 – 2732.

[11] B. Rupp, S. Raub, C. Marian, H.-D. Höltje, *J. Comput. Aided Mol. Des.* **2005**, *19*, 149 – 153.

[12] W. Köller, W. F. Wilcox, *Plant Dis.* **1999**, *83*, 857 – 863.

[13] Resistance: Penicillium digitatum, *French Ministry of Agriculture and Fishery*, http://e-phy.agriculture.gouv.fr/ecoacs/3a1286htm, last access 24-11-2005.

[14] B. Wanke, M. S. Lazera, M. Nucci, *Mem. Inst. Oswaldo Cruz* **2000**, *95*, 153 – 158.

[15] Nguyen, Q. V., Hospital-Acquired Infections, e-Medicine.com, Inc. September 1, 2004 update, http://www.emedicine.com/ped/topic1619.htm, last access 24-11-2005.

[16] G. Schnabel, A. L. Jones, *Phytopathology* **2001**, *91*, 102 – 109.

[17] C. D. Stoud, *Structure* **2004**, *12*, 1921 – 1922.

[18] K. Asai, N. Tscuchimori, K. Okonogi, J. R. Perfect, O. Gotoh, Y. Yoshida, *Antimicrob. Agents Chemother.* **1999**, *43*, 1163 – 1169.

[19] K. Niimi, D. R. K. Harding, R. Parshot, A. King, D. J. Lun, A. Decoltignies, M. Niimi, S. Lin, R. D. Cannon, A. Goffeau, B. C. Monk, *Antimicrob. Agents Chemother.* **2004**, *48*, 1256 – 1271.

[20] R. Nakaune, K. Adachi, O. Nawata, M. Tomiyama, K. Akutsu, T. Hibi, *Appl. Environ. Microbiol.* **1998**, *64*, 3983 – 3988.

[21] H. Hamamoto, K. Hasegawa, R. Nakaune, Y. J. Kee, Y. Makizumi, K. Akutsu, T. Hibi, *Appl. Environ. Microbiol.* **2000**, *66*, 3421 – 3426.

[22] H. Hamamoto, O. Nawata, K. Hasegawa, R. Nakaune, Y. J. Lee, Y. Makizumi, K. Akutsu, T. Hibi, *Pestic. Biochem. Physiol.* **2001**, *70*, 19 – 26.

[23] H. Hamamoto, K. Hasegawa, R. Nakaune, Y. J. Lee, K. Akutsu, T. Hibi, *Pest. Manage. Sci.* **2001**, *57*, 839 – 843.

[24] R. Nakaune, H. Hamamoto, J. Imada, K. Akutsu, T. Hibi, *Mol. Genet. Genom.* **2002**, *267*, 179 – 185.

[25] Y. Matsuzaki, R. Nakaune, H. Hamamoto, K. Akutsu, T. Hibi, *PMR3* and *PMR4*, novel ABC transporter genes in the phytopathogenic fungus *Penicillium digitatum*, play different roles in multidrug susceptibility and pathogenicity, submitted (March 2004) to the EMBL/GenBank/DDJ databases.

[26] I. Stergiopoulos, J. G. M. von Nistelrooy, G. H. J. Kema, M. A. De Waard, *Pest. Manage. Sci.* **2003**, *59*, 1333 – 1343.

[27] S. Droby, V. Vinokur, B. Weiss, L. Cohen, A. Daus, E. E. Goldschmidt, R. Porat, *Phytopathology* **2002**, *92*, 393 – 399.

[28] G. J. Holmes, J. W. Eckert, *Phytopathology* **1999**, *89*, 716 – 721.

[29] M. Jasinski, E. Ducos, E. Martinoia, M. Boutry, *Plant Physiol.* **2003**, *131*, 1169 – 1177.

[30] M. H. Saier, Jr., *Microbiology* **2000**, *146*, 1775 – 1795.

[31] M. H. Saier, Jr., *Microbiol. Molec. Biol. Rev.* **2000**, *64*, 354 – 411.

[32] H. Martens, T. Naes, *Multivariate Calibration*, 2nd Edn., Wiley, New York **1989**.

[33] K. R. Beebe, R. Pell, M. B. Seasholtz, *Chemometrics: A Practical Guide*, Wiley, New York **1998**.

[34] M. M. C. Ferreira, *J. Braz. Chem. Soc.* **2002**, *13*, 742 – 753.

[35] M. M. C. Ferreira, A. M. Antunes, P. L. O. Volpe, *Quim. Nova* **1999**, *22*, 724 – 731.

[36] Pirouette 3.11, Infometrix, Woodinville, WA **2003**.

[37] Matlab 6.1, MathWorks, Inc., Natick, MA **2001**.

[38] R. Kiralj, M. M. C. Ferreira, *J. Mol. Graph. Mod.* **2003**, *21*, 499 – 515.

[39] R. Kiralj, M. M. C. Ferreira, *J. Braz. Chem. Soc.* **2003**, *14*, 20 – 26.

[40] R. Kiralj, Y. Takahata, M. M. C. Ferreira, *QSAR Comb. Sci.* **2003**, *22*, 430 – 448.

[41] R. Kiralj, M. M. C. Ferreira, *J. Mol. Graph. Mod.* **2006**, *25*, 126 – 145.

[42] Chem3D Ultra 6.0, CambridgeSoft.Com, Cambridge, MA **2000**.

[43] C Cristner, R. Wyrwa, S. Marsch, G. Külertz, R. Thiericke, S. Grabley, D. Schumann, G. Fischer, *J. Med. Chem.* **1999**, *42*, 3615 – 3622.

[44] Y. C. Fann, C. A. Metosh-Dickey, G. W. Winston, A. Sygula, D. N. R. Rao, M. B. Kadiiska, R. P. Mason, *Chem. Res. Toxicol.* **1999**, *12*, 450 – 458.

[45] B. G. Siim, K. O. Hicks, S. M. Pullen, P. L. van Zijl, W. A. Denny, W. R. Wilson, *Biochem. Pharm.* **2000**, *60*, 969 – 978.

[46] T. Ferenc, E. Janik-Spiechowicz, W. Bratkowska, D. Lopaczynska, H. Strozynski, A. Denys, A. Mordalska, *Mutation Res.* **1999**, *444*, 463 – 470.

[47] M. M. C. Ferreira, R. Kiralj, *J. Chemometr.* **2004**, *18*, 242 – 252.

[48] H. Peltroche-Llacsahuanga, S. Schmidt, M. Seibold, R. Lütticken, G. Haase, *J. Clin. Microbiol.* **2000**, *38*, 3696 – 3704.

[49] B. C. Viljoen, J. L. F. Kock, T. J. Britz, *J. Gen. Microbiol.* **1988**, *134*, 1893 – 1899.

[50] R. P. Lane, Morphometric Discrimination of Pest: Experimental Design, in: D. L. Hawkesworth (Ed.), *The Identification and Characterization of Pest Organisms*, CAB International, Wallingford, UK **1994**, pp. 215 – 224.

[51] A. Morgan, L. Boddy, J. E. M. Mordue, C. W. Morris, *Mycol. Res.* **1998**, *102*, 975 – 984.

[52] R. Russell, M. Paterson, A. Venancia, N. Lima, *Res. Microbiol.* **2004**, *155*, 507 – 513.

[53] P. Znidarsic, A. Pavko, *Food Technol. Biotechnol.* **2001**, *39*, 237 – 252.

[54] D. P., Jr. Visco, R. S. Pophale, M. D. Rintoul, J.-L. Faulon, *J. Mol. Graph. Mod.* **2002**, *20*, 429 – 438.

[55] C. Rücker, M. Meringer, A. Kerber, *J. Chem. Inf. Comput. Sci.* **2004**, *44*, 2070 – 2076.

[56] C. J. Churchwell, M. D. Rintoul, S. Martin, D. P. Visco, Jr., A. Kotu, R. S. Larson, L. D. Sillerud, D. C. Brown, J.-L. Faulon, *J. Mol. Graph. Mod.* **2004**, *22*, 263 – 273.

[57] S. J. Cho, W. Zheng, A. Tropsha, *J. Chem. Inf. Comput. Sci.* **1998**, *38*, 259 – 268.

[58] R. Kiralj, M. M. C. Ferreira, *J. Mol. Graph. Mod.* **2003**, *21*, 435 – 448.

[59] H. Seward, A. Roujeinikova, K. J. McLean, A. W. Munro, D. Leys, Crystal structure of the *Mycobacterium tuberculosis* P450 CYP121-fluconazole, The Protein Data Bank content, deposited on September 29, 2006 and release on November 7, 2006.

[60] B. R. Braun, A. D. Johnson, *Science* **1997**, *277*, 105 – 109.

[61] E. R. Sprague, M. J. Redd, A. D. Johnson, C. Wolberger, *EMBO J.* **2000**, *19*, 3016 – 3027.

[62] A. T. Coste, M. Karababa, F. Ischer, J. Bille, D. Sanglard, *Eukar. Cell* **2004**, *3*, 1639 – 1652.

[63] S. Grkovic, M. H. Brown, R. A. Skurray, *Microbiol. Molec. Biol. Rev.* **2002**, *66*, 671 – 701.

[64] Y.-H. Wang, Y. Li, S.-L. Yang, L. Yang, *J. Comput. Aided Mol. Des.* **2005**, *19*, 137 – 147.

[65] H. S. Garmony, R. W. Titball, *Infect. Immun.* **2004**, *72*, 6757 – 6763.

[66] M. V. Karamouzis, V. G. Gorgoulis, A. Papavassiliou, *Clin. Cancer Res.* **2002**, *8*, 949 – 961.