

Rothamsted Repository Download

A - Papers appearing in refereed journals

Martens, N. F., Bresso, E., Togawa, R. C., Urban, M., Antoniw, J. F., Maignet, B. and Hammond-Kosack, M. 2016. Searching for novel targets to control wheat head blight disease—I-Protein identification, 3D modeling and virtual screening. *Advances in Microbiology*. 6 (11), pp. 811-830.

The publisher's version can be accessed at:

- <https://dx.doi.org/10.4236/aim.2016.611079>

The output can be accessed at: <https://repository.rothamsted.ac.uk/item/8v3wv>.

© 14 September 2016. Licensed under the Creative Commons CC BY.

Searching for Novel Targets to Control Wheat Head Blight Disease—I-Protein Identification, 3D Modeling and Virtual Screening

Natália F. Martins¹, Emmanuel Bresso¹, Roberto C. Togawa¹, Martin Urban², John Antoniow², Bernard Maigret³, Kim Hammond-Kosack²

¹EMBRAPA Recursos Genéticos e Biotecnologia Parque Estação Biológica, Brasília, Brazil

²Department of Plant Biology and Crop Science, Rothamsted Research, Harpenden, UK

³CNRS, LORIA, UMR 7503, Lorraine University, Nancy, France

Email: natalia.martins@embrapa.br

How to cite this paper: Martins, N.F., Bresso, E., Togawa, R.C., Urban, M., Antoniow, J., Maigret, B. and Hammond-Kosack, K. (2016) Searching for Novel Targets to Control Wheat Head Blight Disease—I-Protein Identification, 3D Modeling and Virtual Screening. *Advances in Microbiology*, 6, 811-830.

<http://dx.doi.org/10.4236/aim.2016.611079>

Received: June 21, 2016

Accepted: September 11, 2016

Published: September 14, 2016

Copyright © 2016 by authors and Scientific Research Publishing Inc.

This work is licensed under the Creative Commons Attribution International License (CC BY 4.0).

<http://creativecommons.org/licenses/by/4.0/>



Open Access

Abstract

Fusarium head blight (FHB) is a destructive disease of wheat and other cereals. FHB occurs in Europe, North America and around the world causing significant losses in production and endangers human and animal health. In this article, we provide the strategic steps for the specific target selection for the phytopathogen system wheat-*Fusarium graminearum*. The economic impact of FHB leads to the need for innovation. Currently used fungicides have been shown to be effective over the years, but recently cereal infecting *Fusaria* have developed resistance. Our work presents a new perspective on target selection to allow the development of new fungicides. We developed an innovative approach combining both genomic analysis and molecular modeling to increase the discovery for new chemical compounds with both safety and low environmental impact. Our protein targets selection revealed 13 candidates with high specificity, essentiality and potentially assayable with a favorable accessibility to drug activity. Among them, three proteins: trichodiene synthase, endoglucanase-5 and ERG6 were selected for deeper structural analyses to identify new putative fungicides. Overall, the bioinformatics filtering for novel protein targets applied for agricultural purposes is a response to the demand for chemical crop protection. The availability of the genome, secretome and PHI-base allowed the enrichment of the search that combined experimental data *in planta*. The homology modeling and molecular dynamics simulations allowed the acquisition of three robust and stable conformers. From this step, approximately ten thousand compounds have been virtually screened against three candidates. Forty-five top-ranked compounds were selected from docking results as presenting better interactions and energy at the binding pockets and no toxicity. These compounds may act as inhibitors and lead to the development of new fungicides.

Keywords

Fusarium graminearum, Fungicide Development, 3D Protein Modeling, Target Selection

1. Introduction

Fusarium Head Blight disease (FHB), also known as scab, leads to significant losses in crop yield, low grain weights, low seed germination and contamination of grain with mycotoxins. Several species of *Fusarium* can cause FHB. Among these, *F. graminearum* (teleomorph *Gibberella zeae* (Schwein.) Petch), infects different parts of the plant. *Fusarium* penetrates through the stomata to the palea and lemma and destroys these tissues. The disease has a symptomless phase where the hyphae advance between the plant cells prior to host cell death [1]. *Fusarium* infection introduces several mycotoxins into the food chain, such as the trichothecene nivalenol (NIV), deoxynivalenol (DON) and the estrogenic mycotoxin zearalenone (ZEN), posing significant threat to humans and animals health. *Fusarium graminearum* has a very broad host range and can suppress the plant defense responses. Although agricultural practices such as crop rotation and an invasive tillage can limit the development of FHB, some critical factors for the control of *F. graminearum* are still unknown [2].

Synthetic fungicides are globally used to control plant diseases. According to Cools and Hammond-Kosack (2013) 95.5% of the wheat-growing area in UK is receiving fungicide sprays [3]. Nonetheless, fungicide resistance is a threat and leads to the search for new and more efficacious compounds [4]. The development of new fungicides in combination with genomic data can improve their specificity, lower costs and environmental impact. The *Fusarium graminearum* PH-1 genome was one of the first fungal pathogens to be sequenced [4] and recently an Australian isolate (CS3005) of *F. graminearum* genome *senso stricto* was reported by Gardiner *et al.* (2014) [5]. Ma *et al.* (2013) [6] reviewed the major discoveries performed by the genomics analyses of several Fusaria, focused on plant pathogenicity and secondary metabolites production. Furthermore, the compartmentalization of Fusaria genomes into core regions, and adaptive areas with a higher frequency of virulence genes was discussed. The *Fusarium graminearum* genome size is 36 Mb distributed in 4 chromosomes with a total of 13,718 annotated coding genes, 319 tRNAs, and genetic markers [4] [7].

Searching for novel targets to control wheat head blight disease remains a major challenge [1] [8]. The goal of selecting protein targets from the genome is to develop chemical compounds against isolated biological targets from a pathogenic organism or a disease using computer-aided design. The definition of a drug target is a naturally existing cellular or molecular structure involved in the pathology of interest where a chemical compound can act on [9]. To find a new compound for agricultural application, we firstly select protein targets from relevant databases and prepare those targets to the virtual screening (VS) step [10] [11]. This work is the first report of genomics

analysis to search for new targets and to develop fungicides using molecular structural based approach applied for agricultural purposes.

2. Results and Discussion

2.1. EST Sequences Analysis

At the early phase of infection, fungal growth is symptomless representing the majority of mycelium colonizing the host tissues. Secondly the hyphal colonization is observed where the plant cell collapses. During the infection process the fungus maintains living host cells, probably with the expression of various secreted proteins, that might increase during the progress of the infection [6] [12]. In order to explore the relationship between plant and pathogen induced *in planta*, we obtained ESTs from four infected wheat varieties at the early stage of infection [13] [14]. The total amount of 57,282 ESTs from four interactions between wheat and *Fusarium graminearum*, processed at Rothamsted Research were assembled to build 34,970 unigenes from which we trimmed those related to the *Triticum aestivum* plant to those related to *Fusarium*. For that purpose, BLASTx searches were performed on these 34,970 unigenes against the plant and the pathogen genomes. Moreover, as we were only interested in contigs and not in singlets, this procedure eliminated 10,242 contigs similar to wheat, 951 unknowns and retained 5,094 analogous to the *Fusarium*, which were considered to the database. These results are summarized in **Table 1**. The annotation of 10,242 wheat sequences

Table 1. Summarized results for targets candidates selection using the data set from *Fusarium graminearum* PH-1 genome, EST analysis, PHI-base and Secretome [59]-[61].

Data set	Target	FGSG locus number	Putative protein function	Phenotype in PHI-base	Cell localization	#copies in <i>Fusarium</i> genome	Protein size (aa)
EST analysis	1	FGSG_03537	trichodiene synthase	Mixed outcome	cytoplasm*	1	375
	2	FGSG_05906	triacylglycerol lipase precursor	Reduced virulence	cytoplasm	2*	349
PHI-base	3	FGSG_08731 FGSG_16980	casein kinase	Lethal	cytoplasm	2*	371
	4	FGSG_03795	endoglucanase 3	Not in PHI-base	unknown	1	382
	5	FGSG_00989	rhamnogalacturonase B	Not in PHI -base	unknown	1	540
Refined secretome	6	FGSG_08196	aspergillopepsin-2	Not in PHI-base	cell periphery!	1	260
	7	FGSG_11048	arabinogalactan	Not in PHI-base	extracellular	1	350
	8	FGSG_03813	arabinofuranosidase	Not in PHI-base	extracellular	1	499
	9	FGSG_03628	exoglucanase-6A	Not in PHI-base	extracellular	1	458
	10	FGSG_02658	endoglucanase-5	Not in PHI-base	extracellular!	1	380
	11	FGSG_11190	guanyl-specific ribonuclease	Not in PHI-base	extracellular!	1	132
Known targets	12	FGSG_02783 FGSG_16532	D-(24)-sterol C-methyltransferase	Not in PHI-base	cytoplasm	2*	381
	13	FGSG_09530 FGSG_06611	beta-tubulin	Not in PHI-base	cytoplasm	2	446 447

*Number of copies showed below, Invalidated by proteome experiments; (PHI-base reference number 1 PHI: 44; 2 PHI:432; 3 PHI:1235).

and of the 5,094 from *Fusarium graminearum* (named subgroup EST FG) are described in **Table 1** and **Table 2**, respectively (http://lbi.cenargen.embrapa.br/Fusarium_target_supp/). **Table 2** summarizes the ortholog search for the selected candidates.

2.2. Target Selection

The selection strategy is summarized in **Figure 1**. To select the target a total of 20,466 protein sequences were retained: 13,321 sequences from the *Fusarium* genome, in which 31 proteins were identified as involved in the trichothecene pathway from GenRE-FGDB (*Fusarium graminearum* database). Additionally, 5094 EST sequences expressed in four wheat varieties infected with the pathogen and 1369 proteins from the wheat infected secretome prediction and 682 sequences from the Pathogen-Host Interactions database (PHI-base).

From 19,892 sequences only 2835 candidates were retained after considering protein annotation, phenotype description and expression *in planta*. Followed by the filtering for redundancy, cell localization and accessibility (cytoplasmic proteins), only 92 satisfied the criteria. Among them, 30 had a low number of copies in the genome and were considered small proteins. The final filter produced a set of 13 candidates, which had no orthologues within non host genomes. Among them, only three were finally chosen for the 3D modeling phase, namely trichodiene synthase, endoglucanase-5 and the ERG6 sterol C-methyltransferase. These proteins were considered as potential targets

Table 2. Ortholog identification for selected protein targets endoglucanase-5, trichodiene synthase and sterol C-methyltransferase in fungi database, plants, insects and human genome.

Locus number	Function	Genbank accession	Fungi	Ortholog			Reference
				Plants	Insects	human	
FGSG_02658	Endoglucanase-5	XP_382834.1	<i>F. pseudograminearum</i> <i>F. fujikuroi</i> <i>F. oxysporum</i> <i>F. verticillioides</i> <i>Magnaporthe oryzae</i> <i>Melanocarpus albomyces</i> <i>F. culmorum</i> <i>F. mesoamericanum</i> <i>F. austroamericanum</i> <i>F. boothii</i> <i>F. asiaticum</i>	None	None	None	[61]
FGSG_03537	Trichodiene synthase-TRI5	ESU09673.1	<i>F. pseudograminearum</i> <i>F. cerealis</i> <i>F. meridionale</i> <i>F. lunulosporum</i> <i>F. sambucinum</i> <i>F. sporotrichioides</i>	None	None	none	[22]
FGSG_02783	Sterol C-methyltransferase-ERG6	XP_382959.1	<i>F. pseudograminearum</i> <i>F. avenaceum</i> <i>F. fujikuroi</i> <i>F. oxysporum</i> <i>F. verticillioides</i>	None	None	None	[31]

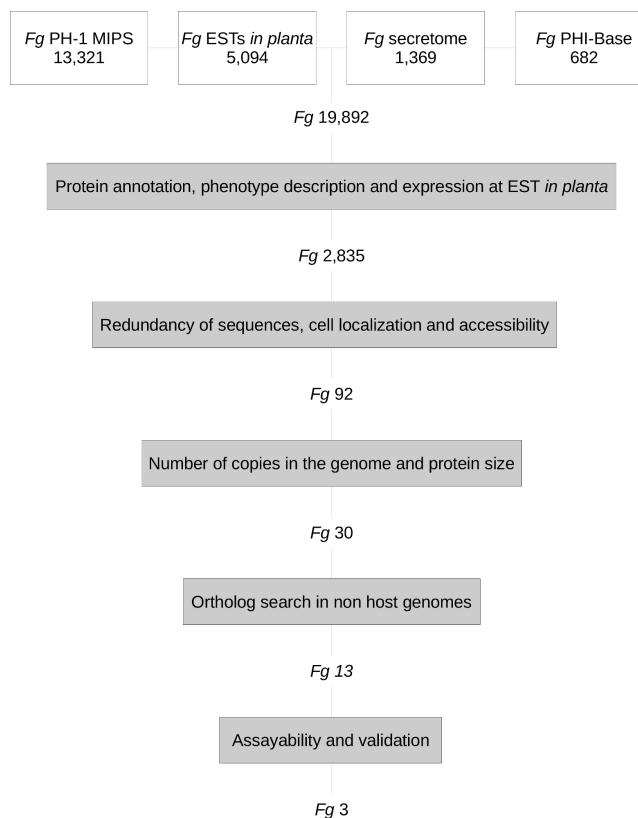


Figure 1. Schematic strategy for protein target selection for *Fusarium graminearum* databases.

representing a group of essential genes, for life or virulence on wheat, with a broad diversity of functions and cellular processes. The three potential targets are significant, unique and have a single function for the pathogen; all are expressed only by the pathogen and are assayable. Therefore, these candidates were selected as potentially and readily inhibited by low molecular weight compounds.

2.3. Endoglucanase-5

The gene FGSG 02658 is located on chromosome 1 and is predicted to encode an endoglucanase type K protein (E.C. number 3.2.1.4.), belonging to the cellulase class that catalyze the hydrolysis of the β -(1,4) glycosidic bonds of cellulose that primarily hydrolyses less well-ordered regions of cellulose by cutting at internal glycosidic bonds. Cellulases play an important role during infection, enabling the degradation of plant cell walls, pathogen penetration and growth through plant tissue [15] [16].

Protein sequence alignments between *Fusarium* endoglucanase FGSG 02658 and orthologs show high similarity with a well-defined conserved domain glycosyl hydrolase family 61 (GH61). To obtain a 3D homology model corresponding to the protein from *Fusarium* FGSG 02658 sequence (residue 1 to 380) was necessary to use complementary protocols for molecular modeling. Complete models were proposed by the HHPRED (one model), PHYRE2 (one model), TASSER (5 models) and ROBETTA (one model) server (Table 3). The comparison between these 3D models showed that the core se-

Table 3. Homology 3D predictions for the target endoglucanase-5 (FGSG_02658).

Method	Number of proposed models	Length of the model sequence	PDB templates used
MODWEB	4	22 - 229	1L8F
		21 - 227	2ENG
		344 - 378	1AZJ
		346 - 378	4BMF
HHPRED	1	1 - 380	1AZ6
M4T	1	20 - 230	10AZ 1HD5
PHYRES2	1	1 - 380	1L8F 2ENG 4BMF 1CBH
SwissProt	1	20 - 227	1L8F
TASSER	5	1 - 380	3ENG
		1 - 380	3CM9
		1 - 380	1W0S
		1 - 380	1W0R
		1 - 380	2OCN
ROBETTA	1	1 - 380	3ENG 3IOX 4BMF

quence (residues 20 - 220) is very similar in all predictions (and even with the predictions not retained for the incomplete sequence), an enormous diversity of proposals is given. These models were all submitted to a short 10 ns of molecular dynamics (MD) simulation to check their conformational changes compared to the starting homology structures. According to our selection procedure, the model given by the Robetta server appeared to be the most suitable and was the most stable from short 10 ns of MD simulations. Consequently, this last model (**Figure 2(C)**) was submitted to 100 ns of MD to validate the stability of the model. The evolution of this 3D structure was followed by the RMSD curve (**Figure 2(A)**) and mapped (**Figure 2(B)**), showing that the stability of the structure was achieved during the MD. Conformation obtained at frame 75 ns (**Figure 2(C)**) was retained as the representative conformer of this family and used later to perform the virtual screening campaign. In this conformation, the overall 3D core structure was conserved during the MD trajectory and consists of a six-stranded beta-barrel and a seventh strand that is not part of the barrel itself surrounded by short helices. The catalytic region is connected to a cellulose-binding domain through a flexible linker region. Two aspartic residues located on either side of the catalytic groove and positioned above the tyrosine residue in the N-terminus were maintained favorably to ensure their catalytic role similarly to experimentally described structures for endoglucanases [17].

2.4. Trichodiene Synthase

The trichodiene synthase gene (*TR5*) is located on chromosome 2, locus FG03537.1

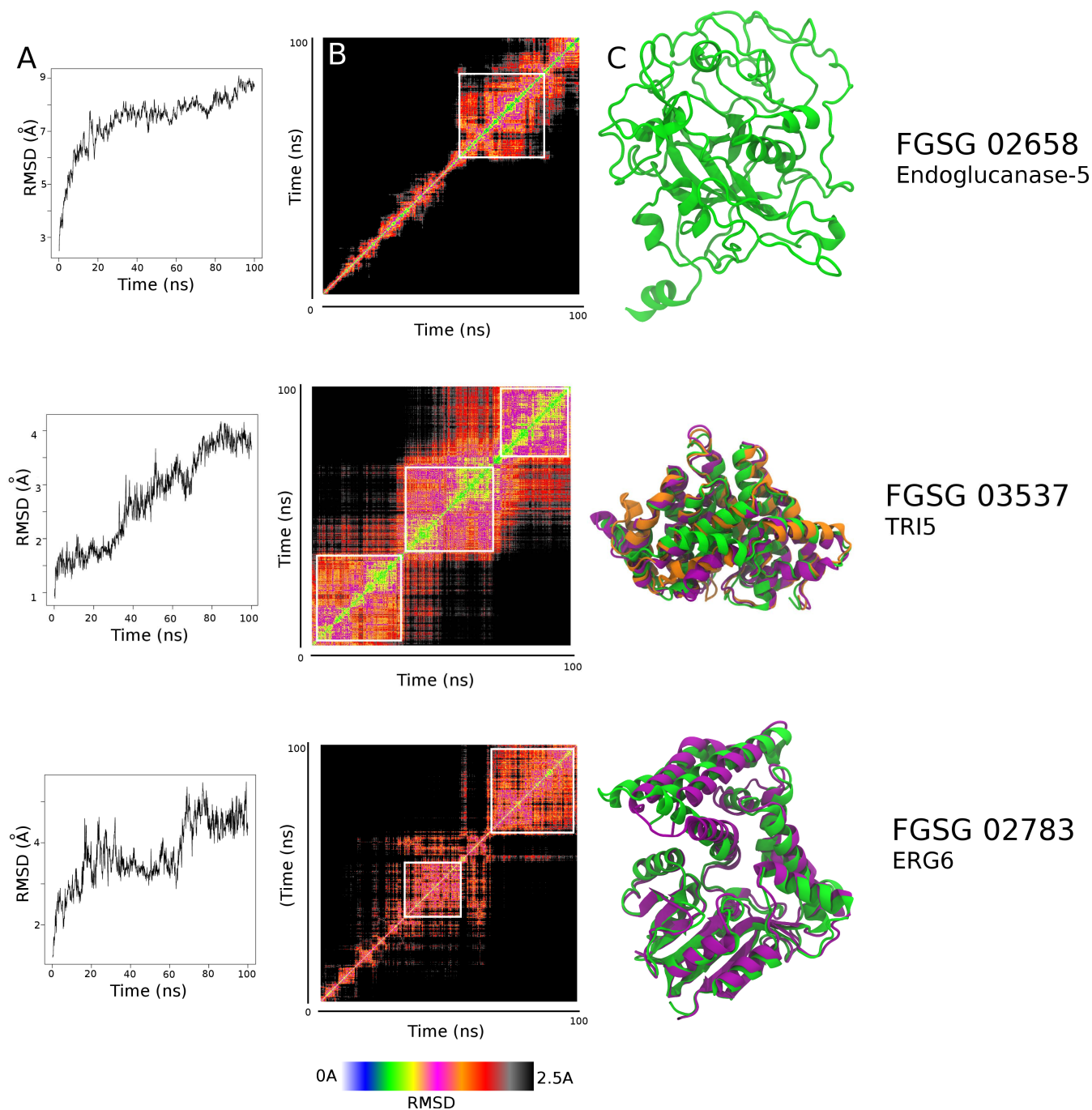


Figure 2. Molecular dynamics (MD) trajectory, RMSD map and three dimensional models for the three selected targets. FGSG 02658 Endoglucanase-5 (A) RMSD map obtained during 100 ns of MD; (B) RMSD map showing one representative conformer. Black color corresponds to high RMSD values between conformers; the white square represents the stable conformational family range; the arrow shows the representative conformer; (C) Proposed 3D model for *F. graminearum* endoglucanase-5. Colors represent the secondary structures. FGSG 03537 TRI5 (A) the RMSD map obtained during 100 ns of MD; (B) RMSD map showing three representative conformers. Black color corresponds to high RMSD values between conformers, the white square represents the stable conformational family range; (C) The proposed 3D structures, in purple the conformational state 1 (frame 200), in green 2 frame 500, in orange 3 frame 800. FGSG 02783 ERG6 (A) RMSD map obtained during 100 ns of MD. (B) RMSD plot for 100 ns MD simulation showing two representative conformers; (C) Proposed 3D structures, in purple conformer 1 (frame 450), in green conformer 2 (frame 850).

described as TRI5 GIBZE Trichodiene synthase (sesquiterpene cyclase) (TS). It contains two exons and encodes for a protein with 375 amino acids (XP 383713). Recently, the *TRI5* gene was studied in terms of phylogenetic relationships among *Fusarium* chemotypes and reported as having a highly conserved organization and a common expression pattern [18]-[20]. Trichodiene synthase catalyzes the production of trichodiene, which is the first committed step in the pathway [19] as part of the 15 biochemical steps required for DON biosynthesis. The enzyme is a sesquiterpene cyclase that catalyzes the formation of trichodiene in the biosynthesis of antibiotics and mycotoxins [21] [22]. It catalyzes the isomerization and cyclization of farnesyl pyro-phosphate to form trichodiene, the first cyclic intermediate in the biosynthetic pathway for trichothecenes. TRI5 serves to branch trichothecene biosynthesis from the isoprenoid pathway. Trichodiene synthase produces the intermediate trichodiene and is specifically induced *in planta*. This feature provides a substantial advantage over direct pathogen control by a chemical compound although it has been shown that the infection of *F. graminearum* in Arabidopsis can occur without the synthesis of DON toxins [23].

The homology modeling procedure used was similar to the one described above for FGSG 02858. Nevertheless, the homology modeling phase was straightforward as all homology servers gave similar results and due to the crystal structure PDB templates existing for an ortholog for trichodiene synthase (1 PDB structure for an apo and 5 for complexes forms), such as the 1JFA structure of trichodiene synthase from *Fusarium sporotrichioides* [22]. Compared to the X-ray crystal structure of this free recombinant, our 3D homology model of FGSG 03537 presented the overall conservation between the templates and the proposed model with an average RMSD 2.0Å for the peptide backbone. The trichodiene synthase structure is, therefore, formed by well conserved 17 α -helices, six of which (C, D, G, H, I and J) define a conical and hydrophobic active site cleft. The structural validation step through 100 ns MD simulation and the RMSD map analysis (Figure 2), shows three conformational families that appear as stable states for this protein. These conformers differed mainly by the position of the C-terminal helix and by some loops arrangements (Figure 2(C)). The position of the aspartate-rich motif DDSKD from residue 100 at the C-terminal up to the end of helix D was very well conserved during the MD trajectory, in good agreement with studies that investigated the mutagenesis of D100 and D101 which are important for catalytic activity [22] [24]. On the opposite wall of the active site cavity, at the C-terminal end of helix J, the “basic motif” DRRYR starting at residue 302 is also well positioned in all three of the models generated in this study. Mutagenesis studies of R304, Y305, and R306 in this motif similarly indicate the importance for catalysis. Consequently, these three models will be used in our structure-based virtual screening selection using an ensemble docking approach [10] [11] [25].

2.5. ERG6-D-(24)-Sterol C-Methyltransferase

The *ERG6* gene encodes for the enzyme D-(24)-sterol C-methyltransferase that catalyzes the attachment of a methyl group acting in a bifurcation point of the ergosterol bio-

synthesis pathway, locus FGSG 02783. The protein is located in the endoplasmic reticulum with a transmembrane portion and an active site positioned toward the cytoplasm [26]. The ERG6 protein is required for the formation of fungal plasma membranes, with an essential role in membrane stabilization and signaling. ERG6 does not occur in plants or insects and sterols can be recognized by a plant cell as a “non-self” signal in host defense [27]. In *Candida albicans*, mutants that do not produce D-(24)-sterol C-methyltransferase showed an increase in plasma membrane permeability. In *Saccharomyces cerevisiae*, ERG6 mutants have altered membrane fluidity and permeability [28]. The homology modeling of this protein was not trivial, and the Robetta server gave the more convincing model (mostly obtained from the 3BUS PDB structure) after the selection. Starting from this model and followed by 100 ns MD trajectory, it appeared that the molecular system was stable and that two conformational families depicted the protein conformational behavior (Figure 2(B)). Recently, Azam *et al.* (2014) [29] examined the structure of a 24-C-methyltransferase from *Leishmania infantum* and describe the prediction of the three-dimensional structure for the *Leishmania* orthologs based on theoretical data and structural details. The overall structure of ERG6 comprises ten alpha-helices and eight beta-strands and loops. The active site was previously described by Nes and co-workers (2002) [30] where the mutated residues 81–86 form a conserved motif of aromatic residues which is of considerable importance in inhibitor binding. Although ERG6 is a non-essential protein, as previously described in *C. albicans* and *A. fumigatus* [26] [31], this protein was added to the initial screening list since it was previously described as an already known drug target for the human pathogen *Paraccocidioides lutzi* and *P. brasiliensis* [11] [32].

2.6. Virtual Screening (VS) Results

All the 10,240 compounds selected from our virtual screening campaign were docked within the binding sites of the targets, as described above, for the major conformers' population. These molecules were next ranked according to their docking scores. After analysis of the whole scores distributions, the best 15 scored compounds obtained for each target were selected. These compounds strongly interact with several target amino acid residues through polar, aromatic and hydrophobic interactions to form the more stable protein/ligand complexes. These 45 ligands thus selected from their docking scores were consequently retained as valuable candidates for further experimental validation. Detailed information for each target is presented below:

VS on Endoglucanase-5 (FGSG 02658)

The top 15 candidate compounds retained for the endoglucanase-5 target are shown in Table 4. Main interacting residues were Cys347, Gly349 and Lys351. Furthermore, these three residues were interacting with the majority of the 15 compounds. Additionally, the protein residues Gly348, Ala362, Asp356 and Ser352 were associated to a few compounds among the 15 ranked. Examples of protein ligand interactions for the three best score and non-toxic compounds are given in Figure 3(A).

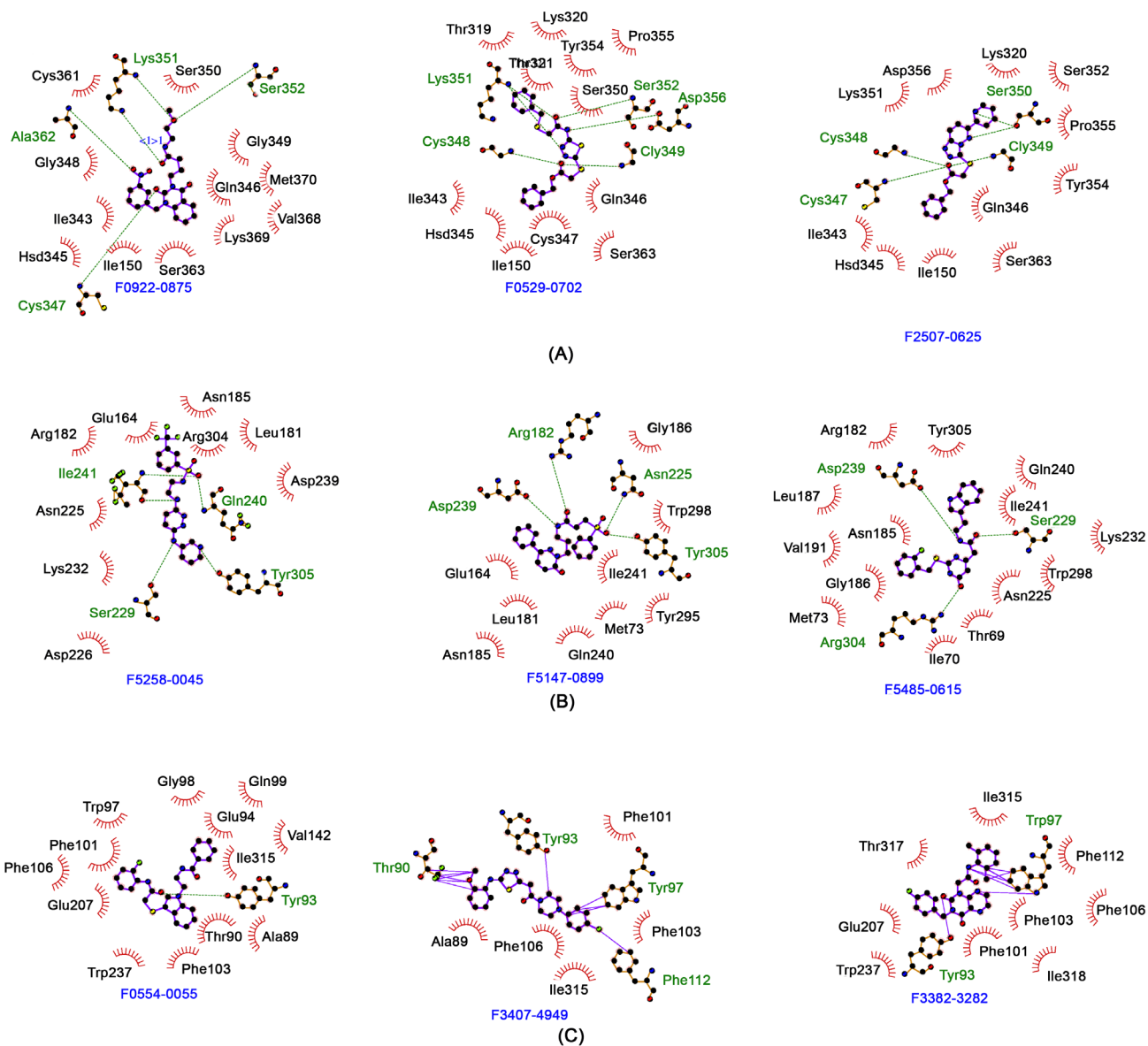


Figure 3. (A) Protein-ligand interactions between the top three compounds and endoglucanase-5; (B) Protein-ligand interactions between the top three compounds and trichodiene synthase; (C) Protein-ligand interactions between the top three compounds and ERG6. Hydrophobic interactions are in red with black residue name and H-bonds are in green.

VS on Trichodiene synthase (FGSG 03537)

The trichodiene synthase docking results are shown in **Table 5**. Several amino acid residues were found to interact with the majority of the top-ranked ligands, such as Asn225 (8 ligands among the 15 top-ranked), Gln240 (6 among 15 compounds), Ser229 (5 among 15), Tyr305 (8 among 15) and Ser229 (all compounds). The protein-ligand interactions found for the three best non-toxic compounds retained are given in **Figure 3(B)**.

VS on ERG6-D-(24)-sterol C-methyltransferase (FGSG 02783)

Table 4. Top-ranking synthetic compounds from Life Chemicals library after virtual screening against endoglucanase-5 unique conformer using GOLD docking software.

Rank	Name	IUPAC name	GOLD Score
1	F0922-0875	N-(2-methoxyethyl)-4-[1-[(3-nitrophenyl)methyl]-2,4-dioxoquinazolin-3-yl]butanamide	86.56
2	F0529-0702	benzyl 2-[[5-[(2-benzylsulfanylacetyl)amino]-1,3,4-thiadiazol-2-yl]sulfanyl]acetate	86.33
3	F2507-0625	benzyl 2-[(6-pyridin-2-yl-[1,2,4]triazolo[4,3-b]pyridazin-3-yl)sulfanyl]acetate	86.07
4	F0590-0218	1-(2,3-dihydroindol-1-yl)-2-[1-[(3-fluorophenyl)methyl]indol-3-yl]sulfonyl ethanone	84.07
5	F1126-0777	1-carbazol-9-yl-3-[2-(3,4-dimethoxyphenyl)ethylamino]propan-2-ol	83.24
6	F3382-0749	5-(1-benzyl-2,4-dioxoquinazolin-3-yl)-N-(thiophen-2-ylmethyl)pentanamide	82.98
7	F2563-0302	N-[(4-methoxyphenyl)methyl]-2-[(5-phenyl-4-pyrrol-1-yl-1,2,4-triazol-3-yl)sulfanyl]acetamide	82.98
8	F5037-0106	3-phenylsulfanyl-N-[2-[(3-phenyl-[1,2,4]triazolo[4,3-b]pyridazin-6-yl)oxy]ethyl]propanamide	82.8
9	F5037-1259	N-[2-[(3-phenyl-[1,2,4]triazolo[4,3-b]pyridazin-6-yl)oxy]ethyl]-1-thiophen-2-ylcyclopentane-1-carboxamide	82.51
10	F3398-1046	N-(2-fluorobenzyl)-3-(4-(piperidin-1-yl)-[1,2,4]triazolo[4,3-a]quinoxalin-1-yl)propanamide	82.5
11	F5016-2418	N1-((1-(benzofuran-2-carbonyl)piperidin-4-yl)methyl)-N2-(2-(cyclohex-1-en-1-yl)ethyl)oxalamide	82.5
12	F3010-0040	3-phenyl-N-(2-phenylethyl)-3-[(5-phenyl-1H-1,2,4-triazol-3-yl)sulfanyl]propanamide	82.45
13	F3406-4638	2-((5-(3-(benzo[d][1,3]dioxol-5-yl)-1,2,4-oxadiazol-5-yl)-2-oxopyridin-1(2H)-yl)methyl)-4H-pyrido[1,2-a]pyrimidin-4-one	82.4
14	F2650-0437	2-[[5-[(4-chlorophenyl)methyl]-4-pyrrol-1-yl-1,2,4-triazol-3-yl]sulfanyl]-N-(2-fluorophenyl)acetamide	82.15
15	F5037-1244	2-phenyl-N-(2-((3-phenyl-[1,2,4]triazolo[4,3-b]pyridazin-6-yl)oxy)ethyl)-2H-1,2,3-triazole-4-carboxamide	82.08

Table 5. Top-ranking synthetic compounds from Life Chemicals library after virtual screening against trichodiene synthase conformers' using GOLD docking software.

Rank	Name	IUPAC name	GOLD Score	Conformation
1	F5258-0045	N-(2-((6-(pyridin-3-ylamino)pyridazin-3-yl)amino)ethyl)-3-(trifluoromethyl)benzenesulfonamide	80.46	3
2	F5147-0899	N-(3-(6-oxo-3-phenylpyridazin-1(6H)-yl)propyl)-3-(phenylsulfonyl)propanamide	79.3	3
3	F5485-0615	N-(2-(1H-indol-3-yl)ethyl)-2-(2-((2-fluorobenzyl)thio)-6-oxo-1,6-dihydropyrimidin-4-yl)acetamide	78.86	2
4	F0554-0301	4-fluoro-N-[2-[3-(2-morpholin-4-yl-2-oxoethyl)sulfanylindol-1-yl]ethyl]benzamide	78.73	3
5	F1604-0352	2-[3-(benzenesulfonyl)-6-methyl-4-oxoquinolin-1-yl]-N-(2-methylphenyl)acetamide	78.16	3
6	F5864-0089	N-(2-((6-(1H-pyrazol-1-yl)pyrimidin-4-yl)amino)ethyl)-3-((6-methylpyridazin-3-yl)oxy)benzamide	77.89	3
7	F5913-0359	N-(2-(1H-indol-3-yl)ethyl)-2-(6-oxo-4-(thiophen-2-yl)pyrimidin-1(6H)-yl)acetamide	77.86	3
8	F2902-1688	2-(2-fluorophenoxy)-N-(2-(4-(phenethylamino)-1H-pyrazolo[3,4-d]pyrimidin-1-yl)ethyl)acetamide	77.71	3
9	F5630-0003	2-(benzo[d]isoxazol-3-yl)-N-(2-((4-methyl-6-(piperidin-1-yl)pyrimidin-2-yl)amino)ethyl)acetamide	76.46	2
10	F0554-0055	N-[2-[3-[2-(2-fluoroanilino)-2-oxoethyl]sulfanylindol-1-yl]ethyl]benzamide	76.28	3
11	F2563-0139	2-[(4-pyrrol-1-yl-1,2,4-triazol-3-yl)sulfanyl]-N-[4-(trifluoromethoxy)phenyl]acetamide	76.23	3
12	F5097-2588	N-[4-[2-oxo-2-(2,2,2-trifluoroethylamino)ethyl]phenyl]-3,3-diphenylpropanamide	76.15	3
13	F2902-1364	2-(1,2-benzoxazol-3-yl)-N-[2-[4-[(4-fluorophenyl)methylamino]pyrazolo[3,4-d]pyrimidin-1-yl]ethyl]acetamide	76.1	1
14	F3407-3740	4-[(2,4-dioxo-1H-quinazolin-3-yl)methyl]-N-[2-(1H-indol-3-yl)ethyl]benzamide	75.99	2
15	F0617-0211	N-[2-[3-[2-oxo-2-(oxolan-2-ylmethylamino)ethyl]sulfanylindol-1-yl]ethyl]thiophene-2-carboxamide	75.96	3

The docking results for this target are presented in **Table 6**. The amino acid residues found to interact with the selected ligands were Trp97 (7 ligands among the 15 top-ranked) and Tyr93 (8 compounds among 15), through pi-pi interactions. The protein-ligand interactions found for the three best non-toxic compounds are given in **Figure 3(C)**.

3. Conclusions

Overall, the searching approach for novel protein targets applied for agricultural purposes is a response to the demand for chemical crop protection. The search for new fungicides or antifungal compounds should be specific with no effect on the host or other organisms. The first part of this analysis was to select specific targets to start the virtual screening selection of chemical compounds. The availability of the genome, secretome, and PHI-base allowed and enriched the search, when used in combination with the experimental data provided from various *in planta* experiments (ESTs).

Several steps of selection enabled the choice of three major targets: an endoglucanase-5, a trichodiene synthase, and a 24-sterol C-methyltransferase. These three proteins are likely to represent a promising target group: the endoglucanase may be re-

Table 6. Top-ranking synthetic compounds from Life Chemicals library after virtual screening against ERG6 conformers' using GOLD docking software.

Rank	Name	IUPAC name	Score	Conformation
1	F0554-0055	N-[2-[3-[2-(2-fluoroanilino)-2-oxoethyl]sulfanylinol-1-yl]ethyl]benzamide	98.41	1
2	F3407-4949	1-(4-(4-chlorophenyl)piperazin-1-yl)-2-(5-((2-methoxyphenyl)amino)-1,3,4-thiadiazol-2-yl)ethanone	98.26	2
3	F3382-3282	N-(2,6-dimethylphenyl)-2-(3-(4-fluorobenzyl)-2,4-dioxo-3,4-dihydropteridin-1(2H)-yl)acetamide	97.95	2
4	F3382-6099	2-(5-amino-3-(ethylamino)-4-(3-(o-tolyl)-1,2,4-oxadiazol-5-yl)-1H-pyrazol-1-yl)-N-benzylacetamide	97.37	2
5	F3382-1309	N-cyclohexyl-2-(3-(4-methoxybenzyl)-4-oxo-3H-pyrimido[5,4-b]indol-5(4H)-yl)acetamide	95.69	2
6	F3382-0379	5-((2-(2-ethoxyphenyl)-5-methyloxazol-4-yl)methyl)-2-(p-tolyl)pyrazolo[1,5-a]pyrazin-4(5H)-one	95.59	1
7	F3382-7502	N-(2-(1H-indol-3-yl)ethyl)-3-(7-methyl-4-oxo-3,4,5,6,7,8-hexahydrobenzo[4,5]thieno[2,3-d]pyrimidin-2-yl)propanamide	95.58	2
8	F6089-1853	N-(2-(2-(3-fluorophenyl)thiazolo[3,2-b][1,2,4]triazol-6-yl)ethyl)-1-methyl-3-phenyl-1H-pyrazole-5-carboxamide	95.36	1
9	F5253-0280	N-(3-((4-(4-fluorophenyl)piperazin-1-yl)sulfonyl)propyl)-1H-indazole-3-carboxamide	95.03	2
10	F5485-0615	N-(2-(1H-indol-3-yl)ethyl)-2-(2-((2-fluorobenzyl)thio)-6-oxo-1,6-dihydropyrimidin-4-yl)acetamide	94.46	1
11	F2471-0720	N-(2-chlorobenzyl)-2-(5-((2-fluorobenzyl)oxy)-2-methyl-4-oxopyridin-1(4H)-yl)acetamide	94	2
12	F6089-8342	Phenyl 4-((3-benzhydrylureido)methyl)piperidine-1-carboxylate	93.63	1
13	F2605-0233	2-(4-ethoxyphenyl)-N-[2-[[5-(4-fluorophenyl)-1H-imidazol-2-yl]sulfanyl]ethyl]acetamide	93.62	2
14	F3406-0517	N-(2-(cyclohex-1-en-1-yl)ethyl)-2-(3-oxo-6-(p-tolylthio)-[1,2,4]triazolo[4,3-b]pyridazin-2(3H)-yl)acetamide	93.42	1
15	F2902-1688	2-(2-fluorophenoxy)-N-(2-(4-(phenethylamino)-1H-pyrazolo[3,4-d]pyrimidin-1-yl)ethyl)acetamide	93.41	1

quired for cell wall degradation and, therefore, the possibility to stop the disease before the development of symptoms; the trichodiene synthase, known to be required for mycotoxin synthesis, once inhibited could reduce the amount of toxins produced during infection; the 24-sterol C-methyltransferase is a well-known target for cellular growth inhibition, especially due to its specificity and absence in mammals.

The structure modeling was an important aspect when searching for new targets. Homology modeling and molecular dynamics proposed convincing 3D models at the atomic level, enhancing the future perspectives for the development of novel, efficient and specific fungicides to control plant diseases that do not impact on the environment. The experimental validation of the present virtual screening campaigns using the three targets models is under development.

4. Methods

4.1. Dataset Setting

To rank and select proteins from the *Fusarium graminearum* genome according to their potential as fungicides targets, we screened 13,321 proteins from the *F. graminearum* PH-1 genome [33], 5094 assembled sequences from ESTs *in planta*, the *F. graminearum* refined secretome with 1369 proteins [12] and the filtered PHI-base databank with 682 proteins from *F. graminearum* [34]. Thus, the identification of potential targets was based on the ensemble of sequences coming from the *Fusarium* genome, assembled ESTs from wheat infected with *Fusarium graminearum*, the secretome and PHI-base sequences. The expressed sequences *in planta* were produced by the construction of 18 cDNA libraries from *Fusarium graminearum*-infected wheat from four different varieties: Bobwhite, Piko, Sumai-3 and Gottingen-2 (supplementary material: http://lbi.cenargen.embrapa.br/Fusarium_target_supp/). The *in planta* libraries were designed to reveal both plant and fungal genes expressed at the early stage of infection. The BLASTx search was performed for the 5094 assembled ESTs to characterize their functions. Additionally, another BLASTx search was performed against three specific datasets: the predicted secretome, PHI-base version 3.4 and a group of selected *TRI* genes to avoid redundancy. All unigenes were submitted to BLAST2GO for further categorization of biological processes, cell localization, and molecular function. All ESTs were also submitted to BLASTx search against the high confidence reference proteins from the wheat genome [35] [36]. A total of 795 sequences from PHI-base which is a web-accessible database that catalogs experimentally verified pathogenicity, virulence and effectors genes, trimmed for *Fusarium graminearum* [34] and those from the total secretome [12].

4.2. Target Selection Strategy

Our systematic criteria for target filtering were defined in terms of priority, as our primary objective was to identify genes that could be used as targets for the development of new control options: 1) protein annotation, phenotype description and transcriptional evidence coming from EST analysis as unique genes expressed in early days of

infection; 2) we discarded redundant sequences, proteins with low structure similarity within the Protein Databank (PDB), cell localization prediction in nuclei, and candidates with low accessibility to chemical compounds; 3) the number of gene copies in the genome, and protein molecular size; 4) the possible target selection also considered the absence of orthologues in another organism such as insects, plants, humans and evolutionary conservation among other fungi. Therefore, our selection process consisted of several elimination rounds. The first step of this funnel consisted in using protein annotation, phenotype characterization at PHI-base and gene expression in the early infection phase from EST data collections. The next stage of our selection was: 1) the elimination of redundancy and choice of candidates expressed during the early days of infection; followed by 2) a B2GO categorization and cell localization within the cytoplasm and accessibility to chemicals; and finally 3) a BLAST search of the retained candidates against the PDB [37]. The third step consisted in a manual curation for some gene copies (one or two) in the genome and protein size (limited to 500 amino acids). Finally, searching for and then eliminating orthologs in other species with sequence identity criteria (above 60%) resulted in the final list. The end of these selections aimed to propose a first set of protein targets necessary to perform a structure-based design of putative new fungicides for agricultural purposes. We kept as representative the *F. graminearum* three proteins for further analysis.

4.3. Target Modeling

In the absence of experimentally solved 3D structures, computational methods were used to predict 3D protein models and provide information regarding protein functions and structures. Homology modeling has been shown to be efficient in methods to reach reasonable theoretical 3D models as soon as a suitable sequence alignment exists between the sequences of the template and the query [38] [39]. To build appropriate 3D models of our three selected targets, we used several homology modeling servers such as I-TASSER [40] [41], SwissProt [42], PHYRE2 [43], M4T [44], HHPRED [45], MODWEB [46] and ROBETTA [47]. Once the target protein was identified as the most suitable template for homology modeling, we used MUSCLE 3.8.31 for multiple sequence alignment with default parameters to check for sequence similarity and to verify the conservation of structural signatures (find diagonals option disabled, a maximum number of iterations: 16, no duration limitation and no more than 200 sequences). The software MODELLER was used to perform the homology modeling task with its default settings [48]. The crude model for each selected target was obtained and equilibrated. Next, their predictions were compared and analyzed, and the models corresponding to the largest part of the query protein sequence and presenting high similarities between several servers' proposals were kept according to structural index such as coverage and percentage of similarity. These 3D homology models were checked for their stability using short 10ns molecular dynamics and the most stable ones retained to be submitted to a long 100ns molecular dynamics simulation to ensure their structural behavior. For this purpose, each protein 3D model was first solvated with an 80 Å³ box of TIP3P ex-

PLICIT water molecules. Next, ions were added for ensuring the electrostatic neutrality of the whole protein and solvent systems. The NAMD program [49] was employed in conjunction with the CHARMM27 force field. The initial states of dynamics were generated from the homology models after 64,000 steps of conjugate gradients minimization protocol followed by an equilibration stage of 1ns. The simulations were carried out in the isobaric-isothermal ensemble, maintaining the pressure and the temperature at 1 atm and 300 K respectively by using Langevin dynamics and the Langevin piston approaches. The equations of motion were integrated with a 1fs time step. Long-range interactions were treated using the particle-mesh Ewald approach with an 11Å cut-off (switching distance 9Å) for the real space calculation. The calculation of forces and motion equations was repeated to generate the trajectories corresponding to the requested simulation times. A conformation of the whole molecular system was recorded every 1ps, generating for the long 100 ns runs conformational samples of 1000 frames each which was then analyzed using the VMD graphic software [50]. The analysis of the models was achieved by monitoring their Root-Mean-Square deviations (RMSD) during the simulations and by checking the conservation of their secondary structure elements along the molecular dynamic trajectories.

4.4. Virtual Screening

Virtual screening uses computer-based methods to discover new ligands on the basis of biological structures. This technique reduces the molecular database to a few hit compounds for a protein target based on structural features such as size and toxicity of the synthesizable chemical structures. The chemical library construction was performed according to the steps described at Beautrait *et al.* [10]; database creation and handling, substructure search, toxicity prediction and 3D structure generation. In order to limit the number of possible side effects, they were detected using three toxicity prediction web servers: Badapple [51], PAINS-Remover [52] and Protox [53]. The chemical library used was the one providing a high diversity of compounds and was accessed from Life Chemicals [54]. This databank contains 10,240 molecules and each one was docked within the binding site of the selected targets. For each stable conformer identified thanks to the MD simulation of each target, the binding pocket was detected and characterized with the LigSite program [55]. The docking was performed by GOLD [56] which has been recognized as one the best docking software [57]. As several stable conformers were identified for each target, we used the ensemble docking possibility available in GOLD. The use of such conformational ensembles was considered as an improved strategy in structure-based docking calculations [58]-[62]. For each docking, 100 starting ligand conformers were used in GOLD. All target conformers used were aligned in a common reference system and the center of the pocket cavity is an average of the individual centers found in each conformation. A sphere of 15Å was selected to define the binding region around this center.

Acknowledgements

The authors gratefully acknowledge CNPq (funding grant 400432/2012-9) EMBRAPA

Labex-Europe, University of Brasilia and CAPES for postdoc fellowship (#51/2013). MU and KHK receive support from BBSRC ISP Grant 20:20 Wheat (BB/J/00426X/1). The authors would also thank the funding resources of PHI-base, the UK Biotechnology and Biological Sciences Research Council (BBSRC) (BB/I/001077/1, BB/K020056/1) and receives additional support from the BBSRC as a National Capability (BB/J/004383/1).

References

- [1] Goswami, R.S. and Kistler, H.C. (2004) Heading for Disaster: *Fusarium graminearum* on Cereal Crops. *Molecular Plant Pathology*, **5**, 515-525. <http://dx.doi.org/10.1111/j.1364-3703.2004.00252.x>
- [2] Ravensdale, M., Rocheleau, H., Wang, L., Nasmith, C., Ouellet, T. and Subramaniam, R. (2014) Components of Priming-Induced Resistance to *Fusarium* Head Blight in Wheat Revealed by Two Distinct Mutants of *Fusarium graminearum*. *Molecular Plant Pathology*, **15**, 948-956.
- [3] Cools, H.J. and Hammond-Kosack, K.E. (2013) Exploitation of Genomics in Fungicide Research: Current Status and Future Perspectives. *Molecular Plant Pathology*, **14**, 197-210. <http://dx.doi.org/10.1111/mpp.12001>
- [4] Yonn M.Y., Cha, B. and Kim J.C. (2013) Recent Trends in Studies on Botanical Fungicides in Agriculture. *Plant Pathology Journal*, **29**, 1-9. <http://dx.doi.org/10.5423/PPJ.RW.05.2012.0072>
- [5] Gardiner, D.M., Stiller, J. and Kazan, K. (2014) Genome Sequence of *Fusarium graminearum* Isolate CS3005. *Genome Announcements*, **2**, e00108-14. <http://dx.doi.org/10.1128/genomeA.00227-14>
- [6] Ma, L.-J., van der Does, H.C., Borkovich, K.A., Coleman, J.J., Daboussi, M.-J., Di Pietro, A., *et al.* (2010) Comparative Genomics Reveals Mobile Pathogenicity Chromosomes in *Fusarium*. *Nature*, **464**, 367-373. <http://dx.doi.org/10.1038/nature08850>
- [7] Wong, P., Walter, M., Lee, W., Mannhaupt, G., Münsterkötter, M., Mewes, H.-W., *et al.* (2011) FGDB: Revisiting the Genome Annotation of the Plant Pathogen *Fusarium graminearum*. *Nucleic Acids Research*, **39**, D637-D639. <http://dx.doi.org/10.1093/nar/gkq1016>
- [8] Kazan, K., Gardiner, D.M. and Manners, J.M. (2012) On the Trail of a Cereal Killer: Recent Advances in *Fusarium graminearum* Pathogenomics and Host Resistance. *Molecular Plant Pathology*, **13**, 399-413. <http://dx.doi.org/10.1111/j.1364-3703.2011.00762.x>
- [9] Agüero, F., Al-Lazikani, B., Aslett, M., Berriman, M., Buckner, F.S., Campbell, R.K., *et al.* (2008) Genomic-Scale Prioritization of Drug Targets: The TDR Targets Database. *Nature Reviews Drug Discovery*, **7**, 900-907. <http://dx.doi.org/10.1038/nrd2684>
- [10] Beaudrait, A., Leroux, V., Chavent, M., Ghemtio, L., Devignes, M.-D.D., Smail-Tabbone, M., *et al.* (2008) Multiple-Step Virtual Screening Using VSM-G: Overview and Validation of Fast Geometrical Matching Enrichment. *Journal of Molecular Modeling*, **14**, 135-148. <http://dx.doi.org/10.1007/s00894-007-0257-9>
- [11] Abadio, A.K.R., Kioshima, E.S., Teixeira, M.M., Martins, N.F., Maigret, B. and Felipe, M.S.S. (2011) Comparative Genomics Allowed the Identification of Drug Targets against Human Fungal Pathogens. *BMC Genomics*, **12**, 75-85. <http://dx.doi.org/10.1186/1471-2164-12-75>
- [12] Brown, N.A., Antoniw, A., Hammond-Kosack, K.E. (2012) The Predicted Secretome of the Plant Pathogenic Fungus *Fusarium graminearum*: A Refined Comparative Analysis. *PLoS*

- ONE*, **7**, e33731. <http://dx.doi.org/10.1371/journal.pone.0033731>
- [13] Hamilton, J.P., Neeno-Eckwall, E.C., Adhikari, B.N., Perna, N.T., Tisserat, N., Leach, J.E., *et al.* (2011) The Comprehensive Phytopathogen Genomics Resource: A Web-Based Resource for Data-Mining Plant Pathogen Genomes. *Database*, **2011**, bar053. <http://dx.doi.org/10.1093/database/bar053>
- [14] Wise, R.P., Caldo, R.A., Hong, L., Shen, L., Cannon, E. and Dickerson, J.A. (2007) Barley-Base/PLEXdb. *Methods in Molecular Biology*, **406**, 347-363.
- [15] Kikot, G.E., Hours, R.A. and Alconada, T.M. (2009) Contribution of Cell Wall Degrading Enzymes to Pathogenesis of *Fusarium graminearum*: A Review. *Journal of Basic Microbiology*, **49**, 231-241. <http://dx.doi.org/10.1002/jobm.200800231>
- [16] Jenczmionka, N.J. and Schäfer, W. (2005) The Gpmk1 MAP Kinase of *Fusarium graminearum* Regulates the Induction of Specific Secreted Enzymes. *Current Genetics*, **47**, 29-36. <http://dx.doi.org/10.1007/s00294-004-0547-z>
- [17] King, B.C., Waxman, K.D., Nenni, N.V., Walker, L.P., Bergstrom, G.C. and Gibson, D.M. (2011) Arsenal of Plant Cell Wall Degrading Enzymes Reflects Host Preference among Plant Pathogenic Fungi. *Biotechnology for Biofuels*, **4**, 4. <http://dx.doi.org/10.1186/1754-6834-4-4>
- [18] Stępień, L. (2014) The Use of *Fusarium* Secondary Metabolite Biosynthetic Genes in Chemotypic and Phylogenetic Studies. *Critical Reviews in Microbiology*, **40**, 176-185. <http://dx.doi.org/10.3109/1040841X.2013.770387>
- [19] Foroud, N.A., Ouellet, T., Laroche, A., Oosterveen, B., Jordan, M.C., Ellis, B.E., *et al.* (2011) Differential Transcriptome Analyses of Three Wheat Genotypes Reveal Different Host Response Pathways Associated with *Fusarium* Head Blight and Trichothecene Resistance. *Plant Pathology*, **61**, 296-314. <http://dx.doi.org/10.1111/j.1365-3059.2011.02512.x>
- [20] Zook, M., Johnson, K., Hohn, T. and Hammerschmidt, R. (1996) Structural Characterization of 15-Hydroxytrichodiene, a Sesquiterpenoid Produced by Transformed Tobacco Cell Suspension Cultures Expressing a Trichodiene Synthase Gene from *Fusarium sporotrichioides*. *Phytochemistry*, **43**, 1235-1237. [http://dx.doi.org/10.1016/S0031-9422\(96\)00382-2](http://dx.doi.org/10.1016/S0031-9422(96)00382-2)
- [21] McCormick, S.P., Stanley, A.M., Stover, N. and Alexander, N.J. (2011) Trichothecenes: from Simple to Complex Mycotoxins. *Toxins*, **3**, 802-814. <http://dx.doi.org/10.3390/toxins3070802>
- [22] Rynkiewicz, M.J., Cane, D.E. and Christianson, D.W. (2001) Structure of Trichodiene Synthase from *Fusarium sporotrichioides* Provides Mechanistic Inferences on the Terpene Cyclization Cascade. *Proceedings of the National Academy of Sciences of the United States of America*, **98**, 13543-13548. <http://dx.doi.org/10.1073/pnas.231313098>
- [23] Cuzick, A., Urban, M. and Hammond-Kosack, K.E. (2008) *Fusarium graminearum* Gene Deletion Mutants *Map1* and *Tr5* Reveal Similarities and Differences in the Pathogenicity Requirements to Cause Disease on Arabidopsis and Wheat Floral Tissue. *New Phytologist*, **177**, 990-1000. <http://dx.doi.org/10.1111/j.1469-8137.2007.02333.x>
- [24] Cane, D.E., Shim, J.H., Xue, Q., Fitzsimons, B.C. and Hohn, T.M. (1995) Trichodiene Synthase. Identification of Active Site Residues by Site-Directed Mutagenesis. *Biochemistry*, **34**, 2480-2488. <http://dx.doi.org/10.1021/bi00008a011>
- [25] Abadio, A.K.R., Kioshima, E.S., Leroux, V., Martins, N.F., Maigret, B. and Felipe, M.S.S. (2015) Identification of New Antifungal Compounds Targeting Thioredoxin Reductase of *Paracoccidioides* Genus. *PLoS ONE*, **10**, e0142926. <http://dx.doi.org/10.1371/journal.pone.0142926>
- [26] Gaber, R.F., Copple, D.M., Kennedy, B.K., Vidal, M. and Bard, M. (1989) The Yeast Gene

- ERG6* Is Required for Normal Membrane Function but Is Not Essential for Biosynthesis of the Cell-Cycle-Sparking Sterol. *Molecular and Cell Biology*, **9**, 3447-3456. <http://dx.doi.org/10.1128/MCB.9.8.3447>
- [27] David, W., Jayasimha, P., Zhou, W., Kanagasabai, R., Jin, C., Jaradat, T.T., *et al.* (2004) Sterol Methyltransferase: Functional Analysis of Highly Conserved Residues by Site Directed Mutagenesis. *Biochemistry*, **43**, 569-576. <http://dx.doi.org/10.1021/bi035257z>
- [28] Vandeputte, P., Tronchin, G., Larcher, G., Ernoult, E., Bergès, T., Chabasse, D., *et al.* (2008) A Nonsense Mutation in the *ERG6* Gene Leads to Reduced Susceptibility to Polyenes in a Clinical Isolate of *Candida glabrata*. *Antimicrobial Agents and Chemotherapy*, **52**, 3701-3709. <http://dx.doi.org/10.1128/AAC.00423-08>
- [29] Azam, S.S., Abro, A., Raza, S. and Saroosh, A. (2014) Structure and Dynamics Studies of Sterol 24-C-Methyltransferase with Mechanism Based Inactivators for the Disruption of Ergosterol Biosynthesis. *Molecular Biology Reports*, **41**, 4279-4293. <http://dx.doi.org/10.1007/s11033-014-3299-y>
- [30] Nes, W.D., Marshall, J., Jia, Z., Jaradat, T.T., Song, Z. and Jayasimha, P. (2002) Active Site Mapping and Substrate Channeling in the Sterol Methyltransferase Pathway. *Journal of Biological Chemistry*, **277**, 42549-42556. <http://dx.doi.org/10.1074/jbc.M204223200>
- [31] Jensen-Pergakes, K. (1998) Sequencing, Disruption, and Characterization of the Sterol Methyltransferase (*ERG6*) Gene: Drug Susceptibility Studies in *Erg6* Mutants. *Antimicrobial Agents Chemother*, **42**, 1160-1167.
- [32] Felipe, M.S.S., Andrade, R.V., Arraes, F.B.M., Nicola, A.M., Maranhao, A.Q., Torres, F.A.G., *et al.* (2005) Transcriptional Profiles of the Human Pathogenic Fungus *Paracoccidioides brasiliensis* in Mycelium and Yeast Cells. *The Journal of Biological Chemistry*, **280**, 24706-24714. <http://dx.doi.org/10.1074/jbc.M500625200>
- [33] Cuomo, C.A., Güldener, U., Xu, J.R., Trail, F., Turgeon, B.G., Di Pietro, A., *et al.* (2007) The *Fusarium graminearum* Genome Reveals a Link between Localized Polymorphism and Pathogen Specialisation. *Science*, **317**, 1400-1402. <http://dx.doi.org/10.1126/science.1143708>
- [34] Winnenburg, R., Baldwin, T.K., Urban, M., Rawlings, C., Köhler, J. and Hammond-Kosack, K.E. (2006) PHI-Base: A New Database for Pathogen Host Interactions. *Nucleic Acids Research*, **34**, D459-D464. <http://dx.doi.org/10.1093/nar/gkj047>
- [35] Mayer, K.F.X., Rogers, J., Dole el, J., Pozniak, C., Eversole, K., Feuillet, C., *et al.* (2014) A Chromosome-Based Draft Sequence of the Hexaploid Bread Wheat (*Triticum aestivum*) Genome. *Science*, **345**, 1251788. <http://dx.doi.org/10.1126/science.1251788>
- [36] Ling, H.-Q., Zhao, S., Liu, D., Wang, J.J., Sun, H., Zhang, C., *et al.* (2013) Draft Genome of the Wheat A-Genome Progenitor *Triticum urartu*. *Nature*, **496**, 87-90. <http://dx.doi.org/10.1038/nature11997>
- [37] Berman, H.M., Westbrook, J., Feng, Z., Gilliland, G., Bhat, T.N., Weissig, H., *et al.* (2000) The Protein Data Bank. *Nucleic Acids Research*, **28**, 235-242. <http://dx.doi.org/10.1093/nar/28.1.235>
- [38] Cavasotto, C.N. and Phatak, S.S. (2009) Homology Modeling in Drug Discovery: Current Trends and Applications. *Drug Discovery Today*, **14**, 676-683. <http://dx.doi.org/10.1016/j.drudis.2009.04.006>
- [39] Berman, H.M., Bhat, T.N., Bourne, P.E., Feng, Z., Gilliland, G., Weissig, H., *et al.* (2000) The Protein Data Bank and the Challenge of Structural Genomics. *Nature Structural Biology*, **7**, 957-959. <http://dx.doi.org/10.1038/80734>
- [40] Zhang, Y. (2008) Protein Structure Prediction: When Is It Useful? *Current Opinion in*

Structural Biology, **19**, 145-155. <http://dx.doi.org/10.1016/j.sbi.2009.02.005>

- [41] Zhang, Y. (2009) I-TASSER: Fully Automated Protein Structure Prediction in CASP8. *Proteins*, **77**, 100-113. <http://dx.doi.org/10.1002/prot.22588>
- [42] Kopp, J. and Schwede, T. (2006) The SWISS-MODEL Repository: New Features and Functionalities. *Nucleic Acids Research*, **34**, D315-D318. <http://dx.doi.org/10.1093/nar/gkj056>
- [43] Kelley, L.A. and Sternberg, M.J.E. (2009) Protein Structure Prediction on the Web: A Case Study Using the Phyre Server. *Nature Protocols*, **4**, 363-371. <http://dx.doi.org/10.1038/nprot.2009.2>
- [44] Fernandez-Fuentes, N., Madrid-Aliste, C.J., Rai, B.K., Fajardo, J.E. and Fiser, A. (2007) M4T: A Comparative Protein Structure Modeling Server. *Nucleic Acids Research*, **35**, W363-W368. <http://dx.doi.org/10.1093/nar/gkm341>
- [45] Söding, J., Biegert, A. and Lupas, A.N. (2005) The HHpred Interactive Server for Protein Homology Detection and Structure Prediction. *Nucleic Acids Research*, **33**, W244-W248. <http://dx.doi.org/10.1093/nar/gki408>
- [46] Pieper, U., Webb, B.M., Dong, G.Q., Schneidman-Duhovny, D., Fan, H., Kim, S.J., *et al.* (2014) ModBase, a Database of Annotated Comparative Protein Structure Models and Associated Resources. *Nucleic Acids Research*, **42**, D336-D346. <http://dx.doi.org/10.1093/nar/gkt1144>
- [47] Kim, D.E., Chivian, D. and Baker, D. (2004) Protein structure prediction and analysis using the Robetta Server. *Nucleic Acids Research*, **32**, W526-W523. <http://dx.doi.org/10.1093/nar/gkh468>
- [48] Bresso, E., Leroux, V., Urban, M., Hammond-Kosack, K.E., Maigret, B. and Martins N.F. (2016) Structure-Based Virtual Screening of Hypothetical Inhibitors of the Enzyme Longi-borneol Synthase—A Potential Target to Reduce *Fusarium* head blight disease. *Journal of Molecular Modeling*, **22**, 163-176. <http://dx.doi.org/10.1007/s00894-016-3021-1>
- [49] Phillips, J.C., Braun, R., Wang, W., Gumbart, J., Tajkhorshid, E., Villa, E., *et al.* (2005) Scalable molecular dynamics with NAMD. *Journal of Computational Chemistry*, **26**, 1781-1802. <http://dx.doi.org/10.1002/jcc.20289>
- [50] Humphrey, W., Dalke, A. and Schulten, K. (1996) VMD: Visual molecular dynamics. *Journal of Molecular Graphics*, **14**, 33-38. [http://dx.doi.org/10.1016/0263-7855\(96\)00018-5](http://dx.doi.org/10.1016/0263-7855(96)00018-5)
- [51] Yang, J., Ursu, O., Bologa, C., Waller, A., Sklar, L. and Oprea, T. (2013) The BADAPPLE promiscuity plugin for BARD Evidence-based promiscuity scores. *Journal of Cheminformatics*, **8**, 29-43. <http://dx.doi.org/10.1186/s13321-016-0137-3>
https://bard.nih.gov/BARD/static/documentation/BARD_ACS_3_Sep2013.pdf
- [52] Baell, J.B. and Holloway, G.A. (2010) New Substructure Filters for Removal of Pan Assay Interference Compounds (PAINS) from screening libraries and for their exclusion in bioassays. *Journal of Medicinal Chemistry*, **53**, 2719-2740. <http://dx.doi.org/10.1021/jm901137j>
- [53] Drwal, M.N., Banerjee, P., Dunkel, M., Wettig, M.R. and Preissner, R. (2014) ProTox: A Web Server for the in Silico Prediction of Rodent Oral Toxicity. *Nucleic Acids Research*, **42**, W53-W58. <http://dx.doi.org/10.1093/nar/gku401>
- [54] Akella L.B. and DeCaprio, D. (2010) Cheminformatics Approaches to Analyze Diversity in Compound Screening Libraries *Current Opinion in Chemical Biology*, **14**, 325-330.
- [55] Huang, B. and Schroeder, M. (2006) LIGSITEcsc: predicting ligand binding sites using the Connolly surface and degree of conservation. *BMC Structural Biology*, **6**, 19. <http://dx.doi.org/10.1186/1472-6807-6-19>
- [56] Verdonk, M.L., Cole, J.C., Hartshorn, M.J., Murray, C.W. and Taylor, R.D. (2003)

- MPROVED Protein-Ligand Docking Using GOLD. *Proteins: Structure, Function and Genetics*, **52**, 609-623. <http://dx.doi.org/10.1002/prot.10465>
- [57] Liebeschuetz, J.W., Cole, J.C. and Korb, O. (2012) Pose Prediction and Virtual Screening Performance of GOLD Scoring Functions in a Standardized Test. *Journal of Computer-Aided Molecular Design*, **26**, 737-748. <http://dx.doi.org/10.1007/s10822-012-9551-4>
- [58] Korb, O., Olsson, T.S.G., Bowden, S.J., Hall, R.J., Verdonk, M.L., Liebeschuetz, J.W., *et al.* (2012) Potential and Limitations of Ensemble Docking. *Journal of Chemical Information and Modeling*, **52**, 1262-1274. <http://dx.doi.org/10.1021/ci2005934>
- [59] Wang, Y., Yang, Xu, L., , H., , Q., Ma, Z. and Chu, C. (2005) Differential Proteomic Analysis of Proteins in Wheat Li Spikes Induced by *Fusarium graminearum*. *Proteomics*, **5**, 4496-4503. <http://dx.doi.org/10.1002/pmic.200401317>
- [60] Guo, L., Han, L., Yang, L., Zeng, H., Fan, D., Zhu, Y., *et al.* (2014) Genome and Transcriptome Analysis of the Fungal Pathogen *Fusarium oxysporum* f. sp. *cubense* Causing Banana Vascular Wilt Disease. *PLoS ONE*, **9**, e95543. <http://dx.doi.org/10.1371/journal.pone.0095543>
- [61] Yang, F., Jacobsen, S., Jørgensen, H.J.L., Collinge, D.B., Svensson, B. and Finnie, C. (2013) *Fusarium graminearum* and Its Interactions with Cereal Heads: Studies in the Proteomics era. *Frontiers in Plant Science*, **4**, 37. <http://dx.doi.org/10.3389/fpls.2013.00037>
- [62] Valente, M.T., Infantino, A. and Aragona, M. (2011) Molecular and Functional Characterization of an Endoglucanase in the Phytopathogenic Fungus *Pyrenochaeta lycopersici*. *Current Genetics*, **57**, 241-251. <http://dx.doi.org/10.1007/s00294-011-0343-5>



Scientific Research Publishing

Submit or recommend next manuscript to SCIRP and we will provide best service for you:

Accepting pre-submission inquiries through Email, Facebook, LinkedIn, Twitter, etc.

A wide selection of journals (inclusive of 9 subjects, more than 200 journals)

Providing 24-hour high-quality service

User-friendly online submission system

Fair and swift peer-review system

Efficient typesetting and proofreading procedure

Display of the result of downloads and visits, as well as the number of cited articles

Maximum dissemination of your research work

Submit your manuscript at: <http://papersubmission.scirp.org/>

