Central University of Technology, Free State

THE VIABILITY OF IN-SILICO ANALYSIS OF CYP51s IN EUROTIOMYCETES SPECIES FOR IDENTIFYING NOVEL AZOLE RESISTANCE MECHANISM

By

Angela Mabona

Thesis Submitted in fulfillment of the requirements of the degree

Master of Health Sciences in Biomedical Technology

Department of Health Sciences

Central University of Technology, Free State

11 April 2022

Supervisor: Prof Samson Sitheni Mashele

© Central University of Technology, Free State

DECLARATION

I, Angela Mabona (SOUTH AFRICAN ID NUMBER: _________), hereby certify the thesis submitted by me for the degree MASTER OF HEALT SCIENCES IN BIOMEDICAL TECHNOLOGY, is my own independent work; and complies with the Code of Academic Integrity, as well as other relevant policies, procedures, policies, procedures, rules and regulations of the Central University of Technology (Free State). I hereby declare, that this research project has not been previously submitted before to any university or faculty for the attainment of any qualification. I further waive copyright of the dissertation in favour of the Central University of Technology (Free State).

© Central University of Technology, Free State

ABSTRACT

The classification and characterization of cytochrome P450 monooxygenases in Eurotiomycetes species of fungus have unfolded various structure-function correlates of the enzyme and its variants from the perspective of evolutionary history and drug resistance. The present thesis explored the attributes of azole resistance of Aspergillus strains in terms of point mutations and polymorphisms in *CYP51* and CYP51. The azole resistance in Eurotiomycetes was studied *Aspergillus fumigatus* as the case study. Studies reveal that Aspergillus houses CYP51A and CYP51B and both of them are required for functionality of the CYP51. The CYP51 is chemically 14-alpha demethylase that converts lanosterol to ergosterol. Ergosterol helps in the membrane fluidity and virulence of the fungal isolate. The enzyme is a putative target of azoles, especially the active heme molecule within the active site. The present study was based on in-silico analysis that involved BLAST, FASTA, phylogenetic analysis, and protein modeling to explore the novel single nucleotide polymorphisms and point mutations and its extrapolation to CYP51 structures as enlisted in the PDB database. The study showed that deletion mutations in CYP51 could lead to non-functional 14-alpha demethylase. Moreover, the deletion mutations could also potentiate the over activity of the ABC transporter proteins (efflux pump) because the alignment of the amino acid residues in mutated CYP51 nearly matched that for ABC transporter proteins in the same fungal isolate. It was further contended that mutations in *CYP51* promoter could have sensitized the binding of various transcription factors that reduced the expression of 14-alpha demethylase.

Keywords: CYP51, Aspergillus, Eurotiomycetes, azole, resistance,

© Central University of Technology, Free State

DEDICATION

This work is dedicated to my two beautiful children Tshepang Neo Sekabate and Neo-Entle Ofentse Sekabate. I would also like to extend the dedication to my mother Seane Suzan Mapheko Sehlare and my twin sister Angie Mabona Moeti.

© Central University of Technology, Free State

ACKNOWLEDGEMENTS

I would like to express my sincere and deepest gratitude to the following people:

1. Prof Khajamohiddin Syed, for introducing me to the p450 field. Thank you for your supervision and guidance within a short space of time that I have known you.
2. Prof Samson Sitheni Mashele, for your co-supervision for my study. Your understanding when I was desperate is broadly acknowledged.
3. Dr Hans Bamal for mentoring me and always encouraging me not to give up on my studies.
4. My colleagues at the South African National Control Laboratory for Biological Products, (SANCLBP), Mr Charles Tebogo Poonyane, Mrs Charity Mkhize and Dr Tondani Mboneni for your unending support and prayers while I was going through a difficult time in my life.
5. To my friends Mrs Anelisiwe Hyera, Shudufadzo Ndou and Linda Tshane, thank you for your friendship and help along the years.
6. The University of the Free State for giving me the opportunity to further my studies.

© Central University of Technology, Free State

Contents

DECLARATION 2
ABSTRACT 3
DEDICATION 4
ACKNOWLEDGEMENTS 5
RESEARCH OUTPUTS .... **Error! Bookmark not defined.**
LIST OF TABLES 8
LIST OF FIGURES 9
ABBREVIATIONS 10
CHAPTER 1: INTRODUCTION 11
1.1. Background 11
1.2. Problem Statement and rationale 16
1.3. Purpose of the Study 17
1.4. Aims and Objectives 18
1.5. Research Questions and Hypothesis 19
References 20
CHAPTER 2: LITERATURE REVIEW 22
2.1. CYP 450 monooxygenases 22
2.2. Nomenclature and Biochemistry of CYP 450 monooxygenases 23
2.2.1. Class I CYP 450s 26
2.2.2. Class II CYP 450s 27
2.3. Eurotiomycetes and their pathogenic potential 29
2.4. Antifungal mechanisms: Challenges and prospects 30
2.4.1. Efflux pumps 31
2.4.2. Target alterations 32
2.5. CYP53 Family: Role in Fungal Metabolism and as Novel Targets for Antifungal drugs 33
2.6. CYP 450 isozymes and Azole Resistance in Eurotiomycetes 36
2.7. Comparing azole resistance in Eurotiomycetes and Other fungal Species 39
2.8. Bioinformatics tools in Genome Data Mining and Homology Identification: Emergence of Statistical computing analysis 41
2.9. Genome wise studies designed to detect novel antifungal targets 44
References 47

© Central University of Technology, Free State

CHAPTER 3: METHODOLOGY 50
3.1. Study Design and Sampling 50
3.2. Procedure 50
3.2.1. BLAST 50
3.2.2.FASTA 50
3.2.3. Phylogenetic analysis 52
3.2.4.Statistical Computing Analysis 52
3.2.5 Protein Modelling 53
3.2.6. In-Silico and Statistical Software 53
3.3. Research Questions and Hypothesis Testing 54
References 57
CHAPTER 4: BLAST AND FASTA RESULTS 59
4.1. Query Sequence for Blast and relevant outputs 59
4.2. Azole resistance CYP51 Homology Matching with Different Aspergillum species 60
i. Query Sequences and FASTA with CYP51 of Different Aspergillum strains 60
4.3. Query sequences and FASTA with CYP450*Asperigillus fumigatus* 70
CHAPTER 5: HOMOLOGY MODELLING REPORT: AZOLE 83
RESISTANCE/MULTI-DRUG TRANSPORTER FOR ASPERGILLUS SPECIES 83
5.1. Model Building Report 83
5.2. Results 83
Models 83
Table 7: Template models 93
References 102
CHAPTER 6: PROTEIN MODELLING 100
References 124
CHAPTER 7: DISCUSSION, CONCLUSION, & FUTURE DIRECTIONS 126
7.1. Conclusion 135
7.2. Strength and Limitations 135
7.3. Future Directions 135
References 136

© Central University of Technology, Free State

LIST OF TABLES

Table 1: Protein Model…………………………………………………….82

Table 2: Template protein…………………………………………………83

Table 3: Ligand Tested Model………………..…………………………...83

Table 4: Protein Model………...………………………………………..84

Table 5: Template model…………………...……………………………...86

Table 6:Ligand tested Model………………………………………………88

Table 7: Template models……………………………………………..91

Table 8: Ramachandran Plot statistics…………………………………112

© Central University of Technology, Free State

LIST OF FIGURES

Fig 1: Phylogenetic Analysis……………………………………………….. 69

Fig 2: Phylogenetic analysis with Eurotiomycetes CYP51………………. 81

Fig 3: Protein Model……………………………………………………106

Fig 4: Local Quality Estimate…………..………………………………107

Fig 5: PDB Estimate…………………………………………..………..108

Fig 6: Protein Model……………………………………………………109

Fig 7: crystal structure of model protein…………………………………. 111

Fig 8: Ramachandran Plot………………………………………………. 112

Fig 9: Alignment of the models …………………………………….. 114

© Central University of Technology, Free State

ABBREVIATIONS

CYP450: Cytochrome P450 monooxygenase

CYP51: Cytochrome P450 monooxygenase isoform 53

CPR: Conditional promoter replacement

FAD: Flavine adenine nucleotide

GRACE: Gene replacement and conditioned expression

GWS: Genome wide sequencing

NRE: Negative regulatory element

SRS: Substrate recognition sequence

© Central University of Technology, Free State

CHAPTER 1: INTRODUCTION

1.1.Background

Efforts are now being aimed to decipher the molecular basis of complex diseases from the high throughput strategies that are available for identifying the dynamics of the bio-molecules that drive such diseases (Bencurova et al., 2018, p. 81). The successful completion of the human-genome project coupled with different technological advancements in the field of bioinformatics have provided a myriad of opportunities to scientists and investigators to undertake multidimensional analysis of different biological systems including fungi, bacteria, and humans (Sammut, Finn, Bateman, 2008, p. 210). Such attributes have been related to the enormous transitions in the field of transcriptomics. The affordable accessibility and ready availability of genomic data which support investigations have further expedited data generation in the field of bioinformatics. Since most the biological systems are studied at the genomic level, a large amount of relevant data is accessible through the public databases (Jensen et al., 2008, p.36). Hence, different computational-based methodologies have gained popularity for disseminating genomic data in attempt to ensure superior prognosis of a disease.

In this regard, In silico methods have generated wide interests across microbiologists, molecular biologists, and pharmacologists for identifying novel drug targets through computational analysis (Letunic & Bork, 2018, p.D493). In-silico analysis is derived from the phenomenon of the mass use of silicon that is inherent to computer chips. Since any form of computational data involves the mass use of computer chips they are often referred to in-silico analysis. In simplest terms, in-silico analysis is the analysis and prediction of future data from existing data through simulations conducted through the computer programs. In-silico analysis has become an integral part of systems biology because most of the structural and functional correlates of an organism are driven

© Central University of Technology, Free State

by bio molecules which are either genomes themselves or their expression products (Szklarczyk et al., 2014, p. D447). The aberrant expression of these genomes or their intrinsic features is related to the genesis of a disease from the perspective of the humans and the microbial species respectively. It is contended that the genomic expressions of microorganisms or humans regulate a wide array of cellular functions that either help to combat a disease or leads to the genesis of a disease. Therefore, biomedical scientists have always focussed on identifying the molecular basis of a disease to develop novel therapeutic strategies. However, historically such process involved the characterization of single molecules from a specific cell or a microorganism (Galperin et al., 2015, p. D261). On the contrary, such techniques often led to blind ends because most of the diseases caused by pathogenic fungi or bacteria are complex in nature and involved polygenic involvement rather than single molecule. Therefore, studying single molecules or single species at one time did not provide orchestrated information regarding the involvement of cellular pathways and its transition over time in the genesis of a disease. It is well-known that all the biological changes in a cell or a microorganism is governed by the expressional variations of different genes.

The ability of the cell to switch on of and switch off gene expression is the basis of different biological activities that are driven by the cell. It is contended that gene transcription is a facilitator of different pathogenomic events that governs the genesis of a disease. Hence, most of the interests have centred in identifying the gene expression profile of key genes or gene clusters that underpins the genesis of a disease. The comparison of gene expression profiles of various identical species of organisms or proteins help to predict the molecular basis of a disease as well as the mechanisms through by which they could switch-on or switch-off their expressions under various conditions. It is contended that the comparison of gene expression under various conditions it is possible to identify the role played by a single gene or its associated genes in regulating signalling cascades

© Central University of Technology, Free State

(Birney & Ensembl, 2003, p, 213). Expression profiling through in-silico analysis is also required for understanding the functions of a gene and whether they could act as potential therapeutic targets (Bateman et al., 2000, p. 263).

Different species of fungi are widely recognized for their pathogenic profile and economic importance. Fungi represent the largest biological kingdom comprising of diverse lower eukaryotic microorganisms. Different fungal species are pathogenic to both plants and animals. Moreover, such pathogens are continuously evolving new strategies for adapting to diverse ecological niches in spite of negative selection pressures. Hence, management of fungal diseases has raised concerns across scientists and health care professionals. Most of the researches in fungal pathogenesis have been limited to epidemiological studies. As a result, there was a limited scope in developing novel therapeutic targets against fungal microbes. Although various antifungal are effective and indicated against a wide array of diseases, such compounds are fast developing resistance against the target fungal species. Antimicrobial resistance also imposes significant concern across scientists and healthcare professional as it deteriorate the prognosis of a disease. Hence, scientists and researchers across the globe are in hunt for novel antifungal drugs that could overcome the challenge of antimicrobial resistance (Vandeputte, Ferrari, & Coste, 2012, n.p.).

Although different studies have implicated polymorphisms and functional correlates of CYP53 in ascomycete and basidiomycete species of fungi, no studies until date have explored the diversity of CYP53 family members in eurotiomycete species. These species of fungi represent a morphological and ecological disparate set of fungi that are strongly associated with the genesis of different life-threatening infections across animals and plants. Most of the members of the Eurotiomycetes serve significant contributions to human welfare. However, some of its members are also implicated in the genesis of life-threatening diseases. For example, mortality in humans

© Central University of Technology, Free State

due to Aspergillus fumigates infection is estimated to be 90%. The fungus can act as both primary and secondary pathogen. As per phylogenic classification, the class Eurotiomycetes (also known as Ascomycota, Pezizomycotina) comprises two major clades of ascomycetous fungi; subclass Eurotiomycetidae and subclass Chaetothyriomycetidae. The subclass Eurotiomycetidae was previously categorized as Plectomycetes due to the presence of enclosed ascomata and protunicate asci. The subclass Chaetothyriomycetidae is featured by enclosed ascomata with opening reminiscent those are common to Dothideomycetes or Soradariomycetes. Categorization and classification of CYP53 family members in eurotiomycete species would help to develop potent antifungal drugs against this class of fungi.

It is contended that potential drug targets can be developed if the target fungal genomes are appropriately sequenced and characterized. As a result, different organizations such as MIT and Harvard University are pursuing fungal genome sequencing projects. The center of interest in sequencing fungal genomes pivots around the characterization and identification of different isotypes of cytochrome P450 monooxygenase (CYP450). Genomic sequencing analysis confirmed the presence of various isotypes of the CYP450 enzyme across different fungal species. CYP450s are hemethiolate proteins those are ubiquitously present across different species of the biological kingdom (including humans). CYP450s play a major role in metabolizing toxicants and drugs in the respective organisms. Hence, such enzymes actively take part in primary and secondary metabolic reactionincluding xenobiotic (toxicant) inactivation. These enzymes are primarily considered as novel drug targets due to their importance in fungal pathophysiology.

Amongst different CYP450 isotypes, CYP51 (also referred as sterol-14a-demethylase) is the highest conserved isotype across different biological kingdoms. Hence, CYP51 has been the primary target of different antifungal drugs (such as azoles). CYP51 causes demethylation of

© Central University of Technology, Free State

lanosterol which converts the latter into ergosterol. Ergosterol is an important component of the fungal cell membranes and is responsible for fluidity of the membrane. Hence, CYP51 plays a major role in maintaining the integrity of the fungal cell membrane and in ensuring viability of the fungal cell. Such dynamics in metabolism help to ensure the titer of pathogenic fungi. Different studies suggest that fungal isolates have developed significant resistance against azole and azole derivatives. Moreover, the antifungal compounds which are currently used have major limitations because fungal metabolic pathways are almost similar host metabolic pathways. Such issues limit the use of novel antifungal drugs that are targeted towards CYP51.

On the contrary, studies conducted on CYP53 isotype of cytochrome p450 monooxygenase have shown promising results to act as a novel and alternate drug target to the currently available antifungal compounds. The hallmark of such isotypes is their selective presence in fungal species only. Hence, antifungal drugs against CYP53 isotypes would not inhibit metabolic pathways in the host (animals, mammals, or plants). The CYP53 isotypes functions as benzoate parahydroxylases as they hydrolyze benzoate to benzoate parahydroxylate. Benzoate is a naturally occurring antifungal compound and is highly toxic to fungal cells. The major actions of benzoate include disruption of the fungal cell membrane, inhibition of cellular and biochemical processes in fungi, changes in pH balance and induction of stress response in fungal species. Hence, parahydroxylation of benzoate is an important phenomenon for any fungal cell to overcome the challenges of negative selection pressure. CYP53 P450-mediated conversion of benzoate to benzoate parahydroxylate is the only known biochemical pathway in fungi which channels and detoxifies benzoate through the beta-ketoadipate pathway.

© Central University of Technology, Free State

1.2. Problem Statement and rationale

Vandeputte et al. (2012, n.p.) highlighted that antifungal resistance imposes critical healthcare challenges despite radical improvements in antifungal therapies. Over the past decade, the molecular mechanisms those underpins antifungal resistance have been extensively explored. However, managing fungal pathogens is a dilemma across clinicians and allied healthcare professionals. There are two facets of such dilemma; basic physiology of different fungal isolates is still obscure, and the growing prevalence of resistance to antimicrobial compounds. Although different authors have categorized and elucidated the functions of CYP53 in Ascomycetes and basidiomycete class of fungi, little is known regarding the structural-functional correlates of cytochrome p450 monooxygenase system in the fungal class “Eurotiomycetes.” Eurotiomycetes represents a morphologically and ecologically unique set of fungi which are capable of causing serious infections in both plants and animals. In fact, the pathogenic potential of Eurotiomycete class of fungi is speculated to be more than Ascomycetes and Basidiomycetes.

Till date, there have been no attempts to conduct in silico analysis of cytochrome P450 monooxygenases in studying the chronic infection caused Eurotiomycetes species. Although recent studies have highlighted the annotations of these CYP450 species, there is a gap in knowledge whether the annotations and phylogenetic analysis could help to predict antimicrobial resistance or sensitivity. In-silico analysis would provide the platform for supporting or challenging such speculations. Studies suggest that most fungal species related to Eurotiomycetes species are treated by azole drugs such as ketoconazole and itraconazole. However, these drugs are fast turning out to be resistant against the target fungal species. It is well known that the azole group of drugs acts by inhibiting CYP450 enzyme systems especially the CYP51 species which is involved in ergosterol synthesis. The CYP450s are a class of heme-thiolate proteins that are widely

© Central University of Technology, Free State

distributed across related and distant fungal species as well as in other organisms. Till date, there is no information how the CYP450 enzyme systems in Eurotiomycetes could act as putative drug targets for azoles and their derivatives. Recent annotation of this fungal species has provided an opportunity to study the distribution and sequencing of CYP450 in these organisms.

1.3.Purpose of the Study

Despite radical improvements in antifungal therapies over the past 30 years, antifungal resistance still remains a global concern across clinicians and researchers. Fungal isolates belonging to the class Eurotiomycetes is significantly associated with life-threatening infections in humans and plants. It is contended that novel antifungal compounds those target CYP553 would be effective against a wide range of pathogenic fungal isolates. The present study aimed to identify various CYP53 members in eurotiomycetes and their role in fungal physiology. The study was based on the analysis of comparative evolutionary and structural identification of different CYP53 members in the fungal class "Eurotiomycetes." The study deployed different bioinformatics tools to estimate the distribution (phylogenetic relationships) and structural conservation of such isotype of cytochrome p450 monooxygenase. The findings of the present study could elucidate the possibility of developing novel antifungal compounds against eurotiomycetes class of fungi. The antifungal action would pivot around the inhibition of the CYP53 family members in the target fungal population.

However, there has been no attempts to compare the isolated CYP450s of Eurotiomycete with related CYP450s from other fungal species that are either sensitive or have developed resistance against the azoles. Therefore, the present study aimed to identify whether CYP450s in Eurotiomycete species could act as putative target for the azole group of anti-fungals. The CYP450

© Central University of Technology, Free State

database of pathogenic Eurotiomycete would be compared with related fungal species that infest animal and plant species and against whom the azole group of drugs is active or resistant. The study would also reflect whether CYP53 could be a potent and selective target for antifungal compounds even in Eurotiomycetes. The extrapolations were based on the possible three-dimensional structural attributes of the CYP53 members that were explored in the present study. The study also investigated the possible causes for the existence of different paralogs of CYP53 in the genomes of eurotiomycetes. Once again, such possibilities were extrapolated from the sequence identity of CYP53 from different isolates of the fungal class 'Eurotiomycetes' and related (aligned) fungal species. This study was centered on Eurotiomycetes because till date no study explored the distribution or structural-functional correlates of CYP53 in this pathogenic class of fungi. The findings of the present study might be beneficial for clinicians and researchers in developing therapeutic strategies against Eurotiomycetes class of fungi. Such initiatives could reduce the prevalence of mortality due to life-threatening infections across at-risk individuals.

1.4. Aims and Objectives

The primary aim of this study was to undertake a comparative analysis of different cytochrome P450 monooxygenases that were reported in pathogenic and non-pathogenic strains of Eurotiomycetes. The objectives of the present study are as follows:

a. To identify the mutational hotspots in the CYP450 gene in Eurotiomycetes species
b. To identify phylogenetic tree of pathogenic Eurotiomycetes based on CYP450
c. To identify the mutational hotspots in the CYP51 gene in Aspergillus species
d. To identify phylogenetic tree of pathogenic Aspergillus based on CYP450

© Central University of Technology, Free State

e. To predict azole resistance from evolutionary deviation of pathogenic Eurotiomycetes from non-pathogenic Eurotiomycetes

f. To predict azole resistance from evolutionary deviation of pathogenic Aspergillus from non-pathogenic Aspergillus

1.5.Research Questions and Hypothesis

Whether in-silico analysis of CYP450s in Eurotiomycete species promises identification of putative targets that have turned resistant to azole group of drugs?

H0: In-silico analysis of CYP450s in Eurotiomycete species does not promise identification of putative targets that have turned resistant to azole group of drugs

H1: In-silico analysis of CYP450s in Eurotiomycete species promises identification of putative targets that have turned resistant to azole group of drugs

© Central University of Technology, Free State

References

Bencurova, E., Gupta, S. K., Sarukhanyan, E., & Dandekar, T. 2018. Identification of Antifungal Targets Based on Computer Modeling. *Journal of fungi (Basel, Switzerland)*, *4*(3), pp.81.

Birney E., Ensembl T 2003. Ensembl: A genome infrastructure. Cold Spring Harb. Symp. Quant. Biol. 68, pp. 213–215

Bateman A., Birney E., Durbin R., Eddy S.R., Howe K.L., Sonnhammer E.L. 2000 The Pfam protein families database. Nucleic Acids Res. 28, pp. 263–266.

Galperin M.Y., Makarova K.S., Wolf Y.I., Koonin E.V. 2015. Expanded microbial genome coverage and improved protein family annotation in the COG database. Nucleic Acids Res. 23, pp. D261–D269

Jensen L.J., Julien P., Kuhn M., von Mering C., Muller J., Doerks T., Bork P 2008. eggNOG: Automated construction and annotation of orthologous groups of genes. Nucleic Acids Res., p. 36

Letunic I., Bork P. 2018. 20 years of the SMART protein domain annotation resource. Nucleic Acids Res. 46, pp. D493–D496

Sammut S.J., Finn R.D., Bateman A. 2008. Pfam 10 years on: 10,000 families and still growing. Brief. Bioinform. 9, pp. 210–219

Szklarczyk D., Franceschini A., Wyder S., Forslund K., Heller D., Huerta-Cepas J., Simonovic M., Roth A., Santos A., Tsafou K.P.2014. STRING v10: protein–protein interaction networks, integrated over the tree of life. Nucleic Acids Res. 43, pp.D447–D452.

© Central University of Technology, Free State

Vandeputte, P, Ferrari, S, & Coste, A. 2012. Antifungal Resistance and New Strategies to Control Fungal Infections, International Journal of Microbiology, vol. 2012, Article ID 713687, 26 pages, 2012

© Central University of Technology, Free State

CHAPTER 2: LITERATURE REVIEW

2.1. CYP 450 monooxygenases

Cytochrome p450 monooxygenases are a group of heme-thiolate proteins which are widely distributed across the plant and animal kingdom. These proteins predominantly act as enzymes for major metabolic and biosynthetic processes within the biological Kingdom. They exhibit distinct and diverse catalytic activity on a wide range of substrates (Bernhardt, 2006). Hence, cytochrome p450 monooxygenases have been widely implicated for the production of different chemicals having value-added needs for humans, for the manufacture of pharmaceutical compounds such as antibiotics, fragrances, and degradation of foreign materials those are toxic to the animal or plant species (Chen et al., 2014). The study of cytochrome 450 monooxygenases has received by the attention due to their thermostable properties and for acting as potential target antibiotic for antifungal compounds. CYP 450 monooxygenases represent an extensive super family of diverse hemo-thiolate proteins. This super family of proteins exists in nature across all phylogenetic domains of life ranging from microscopic prokaryotes to multicellular eukaryotes (Hoffmeister & Keller, 2007). Archaea and bacteria are examples of microscopic prokaryotes those house CYP 450 monooxygenases, while lower eukaryotes containing such enzymes include protists, fungi, and insects. On the other hand, the plant and animal kingdom houses the most variant forms of CYP 50 monooxygenase enzyme systems those are witnessed in higher eukaryotes. The plant kingdom exhibits higher number and variety of CYP450 super family genes compared to animals, fungi, protists, Archaea, and bacteria (Deng, Carbone, Dean, 2007). Although fungal genomes have lower numbers of CYP 450 monooxygenase gene than plants, they exhibit the highest diversity of such genes across all phylogenetic domains. It is estimated that approximately 399 families of P450 gene are distributed across 2784 annotated fungal species. On the contrary, plant

© Central University of Technology, Free State

genomes contain only 129 CYP 450 families those are located across 4267 annotated plant species (Doddapaneni, Chakraborty, & Yadav, 2005).

Presently, almost 22000 CYP 450s have been sequenced, annotated, and described. These enzyme systems have existed throughout the biological kingdom for billions of years and played a significant role in surviving environmental stress. For example, these enzymes have been instrumental in ensuring resistance to insecticides in mosquitoes (those act as vectors for malarial pathogens) (Chen et al., 2014). On the other hand, CYP 450s also play a pivotal role in the bioremediation of hexahydro-1,3,5-trinitro-,3,5,-triazine in contaminated soil. Different studies suggest that CYP 450s are not only implicated in metabolism of toxicants, but they became extensively modified in the course of evolution to perform different functions (Hargrove et al., 2012). However, most of these functions were related to ensure the adaptability and viability of the respective biological species under temperature extremes. For example, different CYP 450s became modified in some biological species to enable them to overcome high hydrostatic pressures or thermal stress of solfataric hot springs. On the other hand, some of its counterparts became modified in wood to utilize carbon that was fixed by photosynthesis. Hence, diversity and existence of CYP450s in lower and higher organisms across the biological kingdom is strongly mandated for ecological sustenance (Kelly & Kelly, 2013).

2.2. Nomenclature and Biochemistry of CYP 450 monooxygenases

The term “CYP 450 monooxygenase” was coined by Omura and Sato during 1962 to 1964. The protein was first named “pigment 450” because these proteins are hemoprotein in nature and exhibit unpredictable spectral properties as a reduced form of CO-bound complex. They exhibited

© Central University of Technology, Free State

maximum absorption at a wavelength of 450nm in the visual spectra. Hence, cytochrome P450 enzyme systems are classified as either monooxygenase or reductase depending on its function as an oxidant or reductant. The unique spectral properties of the CYP 450 super family is attributed to cyteine-thiolate group whereby it forms the 5th ligand of the of the heme-iron complex. However, the absorption spectrum of P450 super family was first elucidated in 1958 by Klingenberg. The nomenclature of CYP450s is based on the percentage homology between their amino acid sequences. The percept of CYP and 450 comes from cytochrome and their peak absorption spectra which are followed by the family number, subfamily annotation, and a unique E.C. (Enzyme Commission) number that denotes the functional property of the enzyme. A particular species of P450 is said to fall within a common family if it exhibits a 40% homology in their amino acid sequences with other members of the same family. Likewise, particular species of P450 is said to fall within a common subfamily if it exhibits a 60% homology in their amino acid sequences with other members of the same subfamily (Chen et al., 2014).

Although the CYP450s exhibit low sequence similarity in multiple sequence alignments across different species within a same family or subfamily, they exhibit high conservation in their topographical and structural domains across such species. Such findings suggest that the basic function of CYP 450s have remained strongly conserved over the period of evolutionary history (Kelly & Kelly 2013). In fact, the correlation between the structural and topographical orientation of different domains of the CYP 450 molecule is the new quest for defining and redefining their structural-functional attributes. The highly conserved and consensus sequences of the CYP450 domains hold the key of understanding the phylogenetic convergence and divergence of such proteins in the history of evolution. Such high conservation in their structural domains also reflect that the mechanism of extracting and transferring molecular oxygen to its respective substrates

© Central University of Technology, Free State

have remained similar during evolutionary history (Kodner et al., 2008). Cytochrome P450s are classified into respective clans based on a large group of families or groups compared to genes within a respective clan. This is because genes within the respective clans do not originate from a common ancestor. On the contrary, studies suggest that such genes have diverged from the common ancestor and are likely to share functional and structural traits those are common between each other. The spectral characteristics of CYP450s are based on spectral assay which works on the principle of the reaction between ferrous form of the molecule and carbon monoxide. The carbon monoxide-bound hemoprotein generates a spectrum which exhibits maximum absorption at a wavelength of 450nm in the visible spectra. Such spectral properties are attributed to the cysteine-thiolate axial ligand bound to the hemoprotein molecules of the CYP 450 families (Nebert, Wikvall, & Miller, 2013).

Monooxygenases (earlier referred as mixed function oxidases) are unique enzymes those catalyze molecular oxygen by extracting its first oxygen atom and inserting the same on the target substrate. On the other hand, the second atom in the oxygen molecule is reduced to water by the accepting two electrons from NADH or NADPH. However, an external reductase protein (such as NADPH reductase) is used to transfer the electrons from NADH or NADH to the second atom in the oxygen molecule. As a result, the ferrous state of the hemoprotein is converted to the ferric state and donates the first atom of the oxygen molecule to the respective and becomes reconverted into the ferrous form. Hence, if the second atom of the oxygen molecule was not reduced into water then the CYP 450 oxidation cycle would not have been viable. CYP 450s also act as potent biocatalysts to catalyze a wide variety of chemo-specific, region-specific, and stereo-specific reactions that involve activation of sp3 hybridized carbon atoms, dehalogenation, hydroxylation of aliphatic and aromatic hydrocarbons, sulfoxidation, oxidative deamination, desulfuration, epoxidation,

© Central University of Technology, Free State

peroxidation, and dealkylation of nitrogen, oxygen and sulfur atoms. However, the CYP 450s also carry out more unusual reactions such as carbon-carbon and carbon-oxygen phenol coupling, carbon-carbon bond cleavage, Baeyer-Villinger oxidation, and a plethora of rearrangement reactions such as ring formation and oxidative aryl migration and isomerization reactions. The diversity in the catalytic mechanisms of CYP 450s reflects their unique ability to bind with different substrates. The major substrates of CYP 450s include fatty acids, steroids, terpenes and terpenoids, prostaglandins, eicosanoids, fat soluble vitamins, bile acids, and a plethora of aromatic and aliphatic compounds those are primarily xenobiotics to the respective biological species. The major xenobiotics those are catalyzed by CYP 450s include pharmacological compounds, organic solvents, polycyclic aromatic hydrocarbons, antibiotics, pesticides, herbicides, anesthetics, and ethanol, carcinogens, and various alkyl and aryl hydrocarbon products. Syed et al. (2014) stated that CYP 450s are indeed versatile biocatalysts owing to the presence of flexible substrate recognition sequences and regions (SRS). The authors further stated that these proteins exhibit unique chemical reactivity and catalyze a diverse range of substrates.

2.2.1. Class I CYP 450s

The classes I CYP 450s are predominantly witnessed in bacteria and are located in their inner mitochondrial membrane. This class of CYP 450 enzymes is soluble in the cellular matrix and obtains the lone pair of electrons from the NADH-dependant FAD containing reductase. Such orientation in their electron abstraction mechanism is based on the formation of iron-sulfur cluster (2Fe-2S) and a FAD (flavine adenine nucleotide) containing reductase (adrenodoxin reductase). The first bacterial CYP 450 enzyme that was identified belonged to class of camphor hydrolase. The enzyme was designated as CYP101 and was isolated from the bacterial isolate *Pseudomonas putida*. Eventually, CYP101 became the first CYP 450 monooxygenase to be characterized both

© Central University of Technology, Free State

functionally and structurally. The electron transfer mechanism to CYP101 class of CYP 450 monooxygenase involved transfer of electrons from NADPH to the enzyme through FAD-containing reductase (putidaredoxin reductase) and an iron-sulfur cluster (putidaredoxin). The CYP101s are primarily responsible for metabolizing carbon compounds, xenobiotics, fatty acids and also for the synthesis of secondary metabolites such as antibiotics and antifungal compounds. On the contrary, the mammalian variations of CYP450s are involved in cleaving the side chain of cholesterol, mediating phase-1 reactions of drug or toxicant metabolism, mediating 11beta-hydroxylation of 11-deoxycortisol, and for the biosynthesis of vitamin D and mineralocorticoids (such as aldosterone).

2.2.2. Class II CYP 450s

The Class II CYP 450s is present as integral proteins in the endoplasmic reticulum. The protein is bound to the ER by a N-terminal anchor. The enzyme obtains the lone pair of electrons through the membrane-anchored NADPH-dependent reductase that contain both FAD and flavin mononucleotide (FMN) as cofactors. Due to their innate capability of catalyzing a diverse range of substrates, CYP450s have immense biological potential as drug target and production of human variables. Research on CYP450s has become an invaluable proposition in the field of biotechnology, pharmacokinetics, toxicological Sciences, and pharmaceutical chemistry. Class II CYP450sare used for the production of fine chemicals, fragrances, production of biofuels, and for their roles in biosensing and bioremediation techniques. The CYP450s have been asking you for the production of transgenic plants such as blue roses. Apart from its biotechnological potential introducing transgenic plants with superior aesthetic values, their role in producing disease resistant and insecticide resistant plants cannot be undermined. In fact, the demand for transgenic plants with aesthetic values has been estimated to be 27 billion U.S. dollars per annum. Hence,

© Central University of Technology, Free State

categorization and classification of novel CYP450s is strongly mandated from an economic point of view. The CYP450s are widely using the field of drug discovery and development. For example, CYP450s are traditionally used in the biotransformation of steroids active pharmaceutical ingredients. Such properties are mediated by their ability to transfer the side chain of steroids and cholesterol. They are also used to produce anticancer medications. For example, CYP450s in mycobacterium species cause hydroxylation of the 7th position of 1-limonene to perillyl alcohol which has potent anti-carcinogenic properties.

These enzymes are commonly used to detect clozapine, styrene, and cholesterol. The different isoforms of CYP450s are also used to detect pharmaceutical compounds by using nanoparticles carbon nanotubes. Such approaches exhibit greater device sensitivity and for achieving target therapeutic range in patients. On the contrary, CYP101A isoform of CYP450s are routinely used for the bioremediation of environmental toxins. For example, CYP101A of *P. putida* is used to hydroxylate camphor and for mediating site -directed mutagenesis. XPla is another compound that is widely used for bioremediation environmental toxicants. It is used in the microbial biodegradation of synthetic military explosives and pollutants such as RDX in contaminated soils. For the past few decades, scientists are in the search of novel renewable sources of energy. It is contented that cytochrome p 450 monooxygenases provide promising approaches in the genesis of first renewable sources of energy. These compounds could help to reduce exploitation of petroleum and other biofuels in the near future. Cytochrome p450 monooxygenases are also implicated in the production of biodegradable, sustainable, and renewable biofuels. The mutated variants of cytochrome p450 monooxygenases are capable of hydroxylating alkanes to alkalones. They are also used to decarboxylate and hydroxylate fatty acids.

© Central University of Technology, Free State

2.3. Eurotiomycetes and their pathogenic potential

Most of the members of the Eurotiomycetes serve significant contributions to human welfare (Kosmidis & Denning, 2015). However, some of its members are also implicated in the genesis of life-threatening diseases. For example, Aspergillus fumigates is life-threatening pathogen. Mortality in humans due to Aspergillus fumigates infection is estimated to be 90% (Patterson et al., 2000). The fungus can act as both primary and secondary pathogen. Hence, *Aspergillus fumigates* is an opportunistic pathogen and is primarily witnessed across immunocompromised individuals. As per phylogenic classification, the class Eurotiomycetes (also known as Ascomycota, Pezizomycotina) comprises two major clades of ascomycetous fungi; subclass Eurotiomycetidae and subclass Chaetothyriomycetidae. The subclass Eurotiomycetidae was previously categorized as Plectomycetes due to the presence of enclosed ascomata and protunicate asci.

The subclass Chaetothyriomycetidae is featured by enclosed ascomata with opening reminiscent those are common to Dothideomycetes or Soradariomycetes. Eurotiomycetes have been implicated across a wide range of pathologic conditions in both humans and plants. Apart from Asperigillus sp. different strains of Eurotiomycetes is responsible for the pathogenesis of various infections. *Histoplasma capsulatum* is a pathologic strain of Eurotiomycetes which is associated with respiratory infections. It is estimated that approximately 10% *Histoplasma capsulatum* infections are life-threatening in nature. The pathogen can cause infection in bones and other subcutaneous tissues of the body. Histoplasma infections are more prevalent in the geriatric community and immunocompromised individuals. On the other hand, *Blastomyces dermatitidis* is responsible for an infection called blastomycosis. It occurs in those individuals who are acutely immunocompromised. The pathogen primarily infects bone and connective tissues and can invade

© Central University of Technology, Free State

other organs. *Coccidioides immitis* is another type of pathologic strain of Eurotiomycetes that leads to coccidioidomycosis or valley fever in humans. The infection occurs in a disseminated and presents as acute infection in immunocompromised individuals. *Arthroderma Benhamiae* and *Trichophyton verrucosum* represent two other pathogenic strains of Eurotiomycetes which is responsible for various dermal infections in humans and animals. The disease is often featured by a loss of hair in affected individuals. Although most of the strains of Eurotiomycetes are pathogenic, some of the strains of this class of fungi exhibit economic viability. For example, *Asperigillus niger* is used to produce citric acid and proteins at a commercial level. On the other hand, *Penicillium sp* has been historically associated with the production of antibiotics.

2.4. Antifungal mechanisms: Challenges and prospects

Vandeputte, Ferrari, & Coste (2012) highlighted that antifungal resistance still remains a global concern, despite radical improvements in antifungal therapies over the past 30 years. Over the past decade, the molecular mechanisms underlying antifungal resistance have been extensively explored. The major mechanisms of antifungal resistance include deregulation of effector genes those confer microbial resistance against antifungal drugs. Such deregulations are speculated stem from point mutations or single nucleotide polymorphisms directly in the genes that codes for antifungal targets. The authors emphasized development of new and putative antifungal drugs those could over microbial resistance by altering the expression of different genes which leads to the genesis of such resistance. The major mechanisms that are deployed by fungi to overcome the challenge of antimicrobial therapy are to increase the expression and up-regulation of efflux pumps.

© Central University of Technology, Free State

2.4.1. Efflux pumps

Efflux pumps are integral proteins (belonging to the family ABC transporters) and play a major role in exhibiting resistance towards azoles. Efflux pumps such as CDR1 and CDR2 are commonly witnessed in *Candida albicans* and are responsible for Candida drug resistances 1 and 2 respectively. Studies suggest that such efflux pumps exhibit significant up-regulation on the cell membrane of some azole resistance fungi. Up-regulation in efflux pumps is speculated to stem from chemically unrelated inducers such as amorolfine, steroid hormones, and fluphenazine. The authors isolated different cis-acting factors that were responsible for the up-regulation of such genes in the given fungal species. Studies involving promoter deletions have indicated various regulatory elements in the CDR1 promoter region. These elements include a basal expression element, a drug response element, two steroid regulatory elements, and one NRE (negative regulatory element). Studies reflected that internal deletions in the basal expression element and the drug response element altered the basal and drug induced expression of the CDR1 gene. On the other hand, internal deletions in the negative regulatory element significantly increased the basal expression of cdr1. Such findings suggest that the altered expression and localization of different proteins associated with drug efflux underpinned the principle of antifungal resistance across a wide range of fungal isolates. However, studies also suggest that such changes in expression of the efflux pump proteins are not uniform across a diverse range of fungal isolates. Hence, development for novel antifungal drugs that aim to reduce efflux of antifungal drugs from the fungal cells might not be solely effective in overcoming the challenge of fungal invasion or resistance.

© Central University of Technology, Free State

2.4.2. Target alterations

Apart from the modulation of expression and turnover of efflux pumps, fungal pathogens are capable of developing antimicrobial resistance by decreasing the affinity of their antifungal targets. However, a reduction in the affinity of antifungal targets does not reduce the sensitivity in the target activity. The most exemplary example of target modulation is witnessed in resistance to azole drugs. Studies suggest that azole exhibits decreased affinity towards mutated lanosterol 14-alpha-demethylase. To recall, 14-alpha-demethylase plays a major role in converting azoles to their active antimicrobial entities. Studies showed that point mutations in the ERG11 gene (which codes for14 alpha-demethylase) completely inhibited the binding capacity of azole to its target. Most of these mutations in the ERG11 gene have been successfully demonstrated in *C. neoformans, C. albicans, and C. tropicalis.*

Likewise, in *Asperigillus fumigates* the 14 alpha-demethylase is encoded by two distinct forms of cytochrome p450 monooxygenases (CYP51a and CYP51b). Mutations in the CYP51a gene account for different types of azole resistance that are witnessed across clinical settings. The authors stated that "in these species of *Asperigillus*, the nature of nucleotide metabolism and the nature of amino acid substitution influenced resistance kinetics of the fungal species to different azole derivatives." It was also evident that "some clinical isolates shared common mutations in the CYP51a with environmental azole-resistant strains. Such findings suggest that clinical azole-resistant isolates might originate from the environment." (Vandeputte et al., 2012, p.13). Although there are different strategies that are aimed against azole, mutations in the ERG11 gene or CYP51 have been implicated in the development of resistance against echinocandin compounds. Such assumptions were confirmed in *S. cerevisiae and C. albicans*. Studies suggest that resistance to echinocandin compounds stem from point mutations in the FK1 or FKS2 genes. Identification of

© Central University of Technology, Free State

the locus of such mutations in the respective genes exposed two "hotspots" which were essential for the functional integrity of the FKS gene products that was associated with the enzymatic conversion of metabolites. However, unlike azoles and ERG11 mutations in the FKS1 genes did not jeopardize the affinity of beta-glucan synthase for its active target. However, there was some reduction in the processivity of the referred enzyme. Mutational hotspots have been identified across different genes that play a key role in fungal metabolic pathways. Such mutational hotspots have been identified in different fungal isolates such as *C. glabrata, C. krusei, and A.fumigatus*.

2.5. CYP53 Family: Role in Fungal Metabolism and as Novel Targets for Antifungal drugs

The members of the CYP53 family of CYP450s play a key role in fungal metabolic pathways. CYP53 family members play a significant role in both primary and secondary metabolism in fungi. The primary metabolic pathway includes the beta-ketoadipate pathway, while the secondary metabolic pathways are involved with the detoxification of phenolic compounds (Miceli & Kauffman, 2015). The beta-ketoadipate pathway is the convergent pathway for the metabolism of aromatic compounds in both bacteria and fungi.Being an integral part of the beta-ketoadipate pathway, CYP53 family members readily detoxifies benzoate and ensures survival of fungi. CYP53 family members hydroxylates benzoate into 4-hydroxybenzoate. The conversion of benzoate into 4-hydroxybenzoate is considered the prime reaction in the fungal metabolic pathway and benzoate metabolism. The 4-hydroxybenzoate subsequently leads to protocatechuate as the ring fission substrate. This reaction is critical as it helps fungal species to detoxify benzoate. Till date, this pathway is recognized as the sole pathway that is carried out by CYP53 family members to detoxify benzoate in fungal species. The role of CYP53 family members in such primary

© Central University of Technology, Free State

metabolic pathway have been proved and endorsed by different authors. In one study, deletion mutation in CYP53 significantly increased mortality in the fungal species. Such findings suggest that CYP53 could have a promising role as a novel and alternative drug target for different antifungal drugs.

Jawallapersand et al. (2014) highlighted that inhibitors directed towards CYP450 effectively inhibited CYP53 activity and promoted growth inhibition across different fungal species such as *A. niger, P. ostreatus, and C. lunatus*. On the other hand, novel antifungal drugs those targets CYP53 are more desirable than their counterparts which target CYP51. CYP53 does not have homologs in higher eukaryotes which differentiates it from CYP51. Since CYP51 exhibits a plethora of homologs in higher eukaryotes, antifungal that target CYP51 could be detrimental to the host itself. Hence, CYP53 could be considered a novel target for antifungal drugs based on its key attributes; role in primary and secondary metabolism, highly conserved primary and secondary structures, and lack of homologs in higher eukaryotes. Apart from being a novel target for antifungal drugs, the distribution of CYP53 family members across different fungal species has its own significance. Studies suggest that CYP53 can exist as multiple copies within different fungal species. These multiple copies are often mistaken as alleles. However, the variation in their loci with respect to different chromosomes makes them subtle as paralogs. In fact, some authors have concluded that the role of CYP53 extends beyond benzoate metabolism in basidiomycetes. Jawallapersand et al. (2014) reported that most of the ascomycetes have a single CYP53 in their genomes. On the contrary, the number of CYP53 in basidiomycete species far exceeds than their ascomycete counterparts. It is contended that the presence of multiple copies of CYP53 in basidiomycete species could be attributed to the duplication of CYP53 members post-speciation. Hence, the presence of multiple copies of CYP53 in the fungal genome underpins the probability

© Central University of Technology, Free State

of paralogous evolution. Jawallapersand et al. (2014) provided a conclusive explanation for the higher numbers of CYP53 that is witnessed in basidiomycete genome.

The presence of high numbers of CYP53 family members and generation of benzoic and para-hydroxy benzoate as an intermediate in the synthesis of veratryl alcohol implicate that basidiomycete play a significant role in the production of veratryl alcohol. Such functions of CYP53 help basidiomycete species to degraded and colonize wood. CYP53 also causes demethylation of stilbene. Stilbene is a class of molecule that is found in plants. Stilbene is preferably demethylated by CYP53D subfamily members found in *P.placenta*. Such findings suggest that CYP53 family members play a major role in detoxification and degradation of plant compounds to help basidiomycete for colonizing wood. Studies suggest that the members of the CYP53D subfamily are present in highest number in certain basidiomycete species and have evolved through paralogous evolution. To recall, paralogous genes are those genes which have different loci and functions in relation to the ancestral gene. Such findings suggest that paralogous evolution or presence of multiple copies of CYP53 confers different advantages to fungal species to overcome the environmental challenges of habitat displacement and threats of predation by toxic compounds. A critical appraisal on the differential distribution of CYP53 members in ascomycete and basidiomycete indicate that although their function is limited to detoxification of toxic compounds in ascomycete species, in basidiomycetes they enable the fungal species to overcome negative selection pressure.

© Central University of Technology, Free State

2.6. CYP 450 isozymes and Azole Resistance in Eurotiomycetes

Resistance of fungal isolates to azole group of anti-fungals is significantly on the rise. Resistance to azole group of anti-fungals is also witnessed across different species of Eurotiomycetes (such as *Aspergillus fumigatus*) (Becher et al., 2012). Although the prevalence of azole resistance is low in Aspergillus species, understanding the mechanism of azole resistance with respect to modulation of its cytochrome P450 monooxygenases provide conclusive insights regarding the development of azole resistance in other Eurotiomycetes species. Aspergillus is a thermo-tolerant and saprophytic mold that is responsible for various allergic and invasive diseases in both humans and animals. Since sporulation is an integral part of the reproductive cycle of Aspergillus, it can withstand adverse and unfavorable environmental conditions. As a result, individuals are often affected with invasive Aspergillosis. Nevertheless, the concerned stakeholders had superior prognosis over the years owing to the susceptibility of the fungal species to a wide array of azole anti-fungals such as voriconazole, itraconazole, and posaconazole. Hence, the azole group of anti-fungals has not only been beneficial for managing acute invasive aspergillosis but is also useful for managing patients with severe asthma and those who are hypersensitive to fungal pathogens. However, the emergence of resistance of Aspergillus to azole group of anti-fungals has raised significant public health concerns. Verweji et al. (2016) highlighted two routes of azole resistance in Aspergillus species; the environmental route and the patient route. Although the routes of development of azole resistance remain different, the mechanism of emergence of resistance in Aspergillus towards anti-fungal drugs remains almost the same. It is contended that any setting which increases the interaction between actively reproducing Aspergillus and anti-fungals (azoles) predisposes the risk of mutations that is responsible for the referred resistance. Such conditions

© Central University of Technology, Free State

could be present either in the patient or in the environment. Within a patient, Aspergillus often accesses the pulmonary cavity where asexual reproduction of the organism takes.

The spores produced from the asexual; reproduction is active hotspots for harboring azole resistant mutations. Genetic analysis studies suggest that Aspergillus undergoes various genetic modifications that confer resistance to the azole group of anti-fungals. Such speculations are substantiated by the diagnostic specimens of the microorganism which exhibit both azole-resistance and azole-susceptible forms. Although Aspergillus is a eukaryotic organism, the development of resistance against antimicrobials is similar and fast like bacterial isolates which are on the contrary prokaryotic organisms. However, resistance through horizontal gene transfer is not witnessed in fungal isolates which is one of the key features of resistance to antimicrobials in bacterial isolates. Hence, resistance to a specific class of anti-fungals is limited to one fungal species because resistance is transferred to the daughter isolates through vertical gene transfer. The resistance mechanisms to novel antimicrobials in aspergillum are attributed to the presence of point mutations in the CYP51A gene. To recall, the CYP51A is a isozyme of cytochrome P450 monooxygenases that is predominantly isolated in fungal species. In fact, the CYP51A gene is also incidentally the primary target of the azole group of drugs. Although the CYP51A gene is acknowledged as the major hot spot for the development of point mutations in Aspergillus, most resistance species of Aspergillus does not exhibit any mutations (either point or frame shift) within this gene. These findings suggest that azole resistance could be a function of point mutations in other genes apart from the CYP51A gene or it might be due to other mechanisms of resistance apart from point mutations. The resistance route through patient-based transmission involves multiple resistance mutations that could be isolated form a single sample, while those whose route of resistance through the environment involves only one azole resistance mechanism in most of

© Central University of Technology, Free State

the patients. The point mutations isolated in the patient-based resistance mechanisms in Aspergillus include base substitutions at G54, P216, F219, M220, G138, Y431, and G448 and also non-CYP51A resistance mechanisms (such as resistance in the ***HapE*** gene). On the contrary, the point mutations isolated in the environment-based route of resistance not only involve the specific domains of the CYP51 gene but also mutations in its transcription enhancer regions (the tandem repeats within the promoter region of the gene). The non-genomic regions which are mutational hotspots for conferring resistance in Aspergillus towards azole group of anti-fungals include the TR34/L98H region, T53 region, T46/Y121F/T289a regions. It is further observed that resistance transmitted though the environmental route exhibit low genetic diversity between azole-resistant isolates from unrelated individuals. On the contrary, the resistance transmitted though the patient route exhibit high genetic diversity between azole-resistant isolates from unrelated individuals. Although the fungal species is not a phtytopathogen, most azole-based fungicides are active against Aspergillus species those harbor different plant species. Studies suggest that most of the azole-group of fungicides used for preventing fungal infections in plants belong to the triazole class. As a result, the molecular structure of azole-based fungicides is significantly correlated with the triazole anti-fungals that is clinically recommended for the management of invasive Aspergillosis. It was hypothesized that the development of resistance in *A. fumigatus* could be attributed to the usage of azole group of fungicides for preventing infections in crops. Since the triazole anti-fungals bear molecular similarity with the triazole class of fungicides, development of cross-resistance is not surprising. Nevertheless, the resistance acquired through both routes of resistance exhibit the phenotypes to multiazole and panazole resistance.

© Central University of Technology, Free State

2.7. Comparing azole resistance in Eurotiomycetes and Other fungal Species

Resistance to azoles as a function of mutation in the *ERG11* (CYP51) gene is also witnessed in Candida species. Studies suggest that amino acid substitutions in three mutational hotspots of ERG11p were responsible for the development of resistance against fluconazole in the target fungal species. Flowers et al. (2015) examined 63 isolates of Candida species that were resistant to fluconazole. The study showed that 55 out of the 63 isolates exhibited at least one single nucleotide polymorphism (SNP) within their ERG11 alleles. Amongst these, nine amino acid substitutions were identified as novel and were specific to the Candida species only. Molecular modeling and sequencing of the ERG11 isozymes reflected resistance against fluconazole stemmed from the mutations that were located within the active site of the enzyme. Likewise, resistance to fluconazole was also attributed to the presence of mutations in the fungal-specific external loop and on the proximal surface that potentially with the fungal loop or near the heme moiety of the enzyme. In another study, Xiang et al. (2013) performed site-directed mutagenesis isolated seventeen mutations in the ERG11 gene out of which seven mutations were classified as novel. Amino acid substitutions in the A114S, Y132H, Y132F, and K143R resulted into fourfold resistance of the referred fungal isolates to fluconazole and voriconazole. However, the resistance to itraconazole group of anti-fungals was not statistically significant in the referred fungal isolates. The study further showed that five of the mutations that contributed to the resistance against fluconazole and voriconazole in the target fungal species were located at or closer to the active site of ERG11. Various models for anti-fungal resistance have been elucidated by different authors, amongst which, the resistance involving the ERG11 is most studied. It is contended that mutational inactivation of Erg3 causes the utilization of alternative sterols in the cell membrane of yeasts. The uptake of exogenous sterols helps to alleviate the inhibition of endogenous sterol production by

© Central University of Technology, Free State

fluconazole. Apart from alteration of the sterol substrates to maintain the metabolism and viability of the cell membrane to confer resistance against fluconazole, certain fungal isolates also exhibit a high turnover of ATP-binding cassette and efflux pumps and through the major super family of the transporters that reduces the intracellular accumulation of azoles within the fungal isolates. Likewise, another study showed that inherent lower affinity of fluconazole to bind with the specific ERG11s could further reduce its potential to inhibit the ERG11s. On the other hand, certain studies suggest that increased expressions of the ERG11 protein could also help the fungal isolate to overcome the activity of azoles, thereby, conferring resistance to fluconazole or voriconazole. In another study, the authors showed that aneuploidy within the fungal genome could lead to genetic adaption post-exposure to the azoles. On the contrary, mutations in the ERG11 gene could result in the expression of certain novel proteins that could further reduce the sensitivity of the fungal isolate towards fluconazole or voriconazole.

In certain Candida species, resistance to fluconazole is also attributed to the mutations in the gene encoding the zinc-finger motif transcriptional regulator such as the Upc2p. The Upc2p in *Candida albicans* is a homolog of the ERG gene regulator pair (Upc2/Ecm22) in *S. cerevisiae*. The findings were supported from the genetic studies that reflected that over expression of Upc2/Ecm22 transcriptional regulators resulted in increased resistance to fluconazole while suppression of their expression alleviated the referred resistance in the given fungal isolates. Zarrin et al. (2008) showed that SNPs in one of the UCP22 alleles could confer fluconazole resistance across the fungal isolates even of such isolates did not over express the genes for efflux pumps. These findings suggest that transcriptional regulation of the ERG11 genes is responsible for the azole resistance that iswitnessed across various fungal isolates. It could be possible that the over expressed transcriptional regulators could act as co-repressors or co-activators for the ERG11

© Central University of Technology, Free State

gene. On the contrary, studies have shown that fungal isolates could still exhibit resistance to fluconazole even when there is no mutation in the referred transcriptional regulators such as UCP2 or ECM22. Nevertheless, Flowers et al. (2012) reported that fluconazole-resistant species of Candida over-expressed the UCP2 by two-fold compared to their counterparts who did not express fluconazole resistance. On the other hand, increased expression of the efflux pumps such as Mdr1p and Cdr1p also accounts for the fluconazole resistance that is witnessed across different Candida isolates. In this regard, the fungal isolates have common between the bacterial species that confer resistance through efflux-pumps mechanisms. For preventing resistance through efflux pump mechanisms, the dose of the antibiotic plays a crucial role in speculating the prognosis of an infection. However, such modulations have still not been identified for combating fungal isolates. Therefore, understanding the mechanisms for azole resistance would help to identify the putative targets for azoles in CYP51 gene in Eurotiomycetes. Likewise, it would also provide the opportunity to identify the single nucleotide polymorphisms in the transcriptional regulators in Aspergillus or related species of Eurotiomycetes in the transcriptional regulators for their *ERG11* gene.

2.8. Bioinformatics tools in Genome Data Mining and Homology Identification: Emergence of Statistical computing analysis

Genome sequencing provides an opportunity to explore different parts of the fungal genome that have remain conserved or underwent significant polymorphisms in the evolutionary history. Such analysis would help to identify novel drug targets that could be universally applicable across a diverse species of fungi. On the other hand, previous studies conducted with different fungal genes such as CYP51 have raised concerns in jeopardizing host metabolism. This is because CYP51 and their subfamily are present in higher eukaryotes. Studies suggest that the CYP53 family members

© Central University of Technology, Free State

of cytochrome P450 monooxygenase could act as novel and universal drug targets for antifungal compounds. Moreover, with the virtue of non-availability of their homologs in higher eukaryotes could overcome the challenges of destruction to the host cells. Such unique CYP53 targets have been already reported in ascomycete and basidiomycete species of fungi. Research reflects that almost 78% of the amino acid sequence in CYP53 family members is conserved in such species of fungi. Such findings suggest that inhibitors developed against any one of the CYP53 family members should be useful in overcoming the threats of different pathogenic fungi. Evidence also suggests that paralogs of CYP53 gene might confer different functions those are absolutely intrinsic to the fungal species from which they are isolated. Future studies should correlate evolutionary milestones which necessitated the duplication of the CYP53 gene in genomes of different fungi. Although different studies have implicated polymorphisms and functional correlates of CYP53 in ascomycete and basidiomycete species of fungi, no studies till date have explored the diversity of CYP53 family members in eurotiomycete species. These species of fungi represent a morphologically and ecologically disparate set of fungi that is strongly associated with the genesis of different life-threatening infections across animals and plants. Categorization and classification of CYP53 family members in eurotiomycete species would help to develop potent antifungal drugs against this class of fungi.

Different bioinformatics tools have paved the road for identifying complex evolutionary relationships between various proteins or organisms. One such bioinformatics tool is genome wide sequencing (GWS). GWS is a bioinformatics tool that helps to explore genomes of various organisms with respect to their evolutionary history. Such tools are also important in categorizing a family or subfamily of proteins those are intrinsic to different microbial species. GWS methods have found extensive use in identifying the set of genes in a genome or across multiple genomes

© Central University of Technology, Free State

that are aligned with different functional attributes in the respective species. Different authors have implemented standardized GWS tools such as NCBI-Batch web CD search tool to identify various families and subfamilies of proteins in lower and higher organisms. One of the important attributes of bioinformatics analysis is the BLAST program. BLAST helps to project a set of organisms or proteins based on their identity to each other. BLAST forms the basis of performing FASTA analysis. In FASTA, sequences of different target genomes or genes are selected and searched for alignment. The multiple sequence alignment helps to identify the gene clusters or nucleic acid sequences or amino acid sequences which were conserved over the period of evolutionary history. Likewise such analysis also helps to identify those sequences or amino acids that underwent mutation and single nucleotide polymorphisms over the period of evolution.

Both BLAST and FASTA analysis helps to compare nucleotide acid sequences or their expressions into amino acids. Such comparison serves to define their phylogenic convergence or diverge to and from each other. As a result of such divergence or convergence, the structural-functional correlates of different proteins and traits can be easily estimated or speculated. FASTA analysis thus help to construct a phylogenetic tree for a set of homologous proteins which might have originated from a common ancestral target but became mutated or modified to serve other functions. For example, different efflux pumps (which are one of the common mediators for antimicrobial) in bacteria and fungi have been shown to originate from a common ancestor. During their initial phase of evolutionary history these pumps accounted for the elimination of toxic compounds from the bacterial and fungal cells, which ensured the viability of concerned stakeholders against negative selective pressures. However, with the progression of evolution these efflux pumps became modified to eliminate antimicrobials from the bacterial and fungal cells. As a result, these pumps became associated with different antimicrobial resistance mechanisms with

© Central University of Technology, Free State

the progression of time. Such findings suggest the reliability and validity of bioinformatics tools in defining structural-functional modifications in the nucleic acid or amino acid sequences of different genomes.

2.9. Genome wise studies designed to detect novel antifungal targets

The improvements in existing antifungal medications and the limitations of antifungal resistance have provided major impetus in elucidating the physiology and biochemical pathways of different fungal isolates those were previously obscure. Such impetus is noted when a group of scientists first elucidated the structural-functional correlates of the genomes of 16 strains of *Candida albicans*. However, such simplification of other fungal isolates was not possible due to their significant similarities with metabolic pathways in higher organisms. Studies suggest that fungal genomes are associated with various functions those are common across their eukaryotic superiors. However, studies also implicate the homology in such sequences was also a key attribute that necessitated study of the genomes of different fungal isolates. More specifically, the reasons for such homology in the phylogeny tree implicate different facets of fungal physiology which are either unrevealed or under revealed. It is noted that genes that are essential for the survival of fungi are also putative targets for different antifungal compounds. Exploration of fungal genomes through genome-wide data mining provides robust and quick estimates for the identification of such targets. For example, some researchers have explored and identified such targets in *C. albicans* and *Asperigillus fumigates* through the implementation of genome wide data mining techniques only. The respective authors used gene replacement and conditioned expression (GRACE) and conditional promoter replacement (CPR) technologies to identify putative targets

© Central University of Technology, Free State

in the respective fungal species. In another study, the authors identified 54 genes of *A. fumigatus* based on the ortholog functions across *C. albicans* and *S. cerevisiae*. The authors showed that 35 such genes were essential in *A. fumigatus*. The differential distribution of various orthologs across closely related fungal species elucidated the functional correlates of such genes in fungal physiology. The authors further showed that although the erg11 gene family (comprising of CYP51 and ERG11B) were essential in *A. fumigatus*, the respective genes individually did not hadany functional significance on *A. fumigatus*.

Zarrin and Faramarzi (2018) evaluated azole-resistant based on molecular analysis of CYP51A gene in *Aspergillus fumigatus.* The authors isolated 58 strains of *Aspergillus fumigatus* that included environmental, clinical, and reference isolates. The azole susceptibility was tested for voricanozole and itraconazole followed by PCR of the CYP51A segments. The study showed that 15.5% of the isolates were resistant to voricanozole with MIC 4ug/ml. A 1500 base pair DNA fragment were amplified using the CYP51 gene for all the Aspergillus isolates. The sequence of fragments exhibited 99% identity with *A.fumigatus* CYP51A gene in GenBank. There was no point mutation in the CYP51A gene codons. Although Zarrin and Faramarzi (2018) could not find any point mutation or single nucleotide polymorphism in the CYP51A gene of Asperigillus, studies suggest that mutational hotspots in fungal genes could confer antifungal resistance. The major drawbacks of Zarrin and Faramarzi (2018) study are that they conducted study with 58 strains of Aspergillus only. Moreover, the number of isolates of Aspergillus that were susceptible to voricanozole were significantly higher compared to those that were resistant to it (45 versus 9, $p<0.05$). Hence, it is not surprising why the authors did not identify any point mutation in the respective gene. Likewise, the authors did not identify the consensus and conserved regions of the

© Central University of Technology, Free State

CYP51A in the respective isolates and their correlation with othe segments in determining evolutionary escapes with respect to treatment with azole drugs.

© Central University of Technology, Free State

References

Chen, W., Lee, M. K., Jefcoate, C., Kim, S. C., Chen, F., & Yu, J. H. (2014). Fungal cytochrome p450 monooxygenases: their distribution, structure, functions, family expansion, and evolutionary origin. *Genome biology and evolution*, *6*(7), 1620–1634. doi:10.1093/gbe/evu132

Bernhardt R (2006). Cytochromes P450 as versatile biocatalysts. J Biotechnol. 124, 128–145

Becher R, Wirsel SG. Fungal cytochrome P450 sterol 14 alpha-demethylase (CYP51) and azole resistance in plant and human pathogens. Appl Microbiol Biotechnol. 95, 825–840

Deng JX, Carbone I, Dean RA (2007). The evolutionary history of cytochrome P450 genes in four filamentous Ascomycetes. BMC Evol Biol. 7, 30

Doddapaneni H, Chakraborty R, Yadav JS 2005. Genome-wide structural and evolutionary analysis of the P450 monooxygenase genes (P450ome) in the white rot fungus *Phanerochaete chrysosporium*: evidence for gene duplications and extensive gene clustering. BMC Genomics. 6, 92.

Dunkel M, Günther S, Ahmed J, Wittig B, Preissner R, SuperPred: drug classification and target prediction. Nucleic Acids Research, volume 36, Issue suppl_2, 1July 2008, p W55-W59, https://doi.org/10.1093/nar/gkn307

Flowers, SA, Colón B, Whaley SG, Schuler MA, Rogers PD (2015). Contribution of Clinically Derived Mutations in ERG11 to Azole Resistance in *Candida albicans*. https://doi.org/10.1128/AAC.03470-14

© Central University of Technology, Free State

Hargrove TY, et al (2012). Structural complex of sterol 14α-demethylase (CYP51) with 14α-methylenecyclopropyl-Δ7-24, 25-dihydrolanosterol. J Lipid Res. 53,311–320.

Hoffmeister D, & Keller NP (2007). Natural products of filamentous fungi: enzymes, genes, and their regulation. Nat Prod Rep. 24, 393–416

Jawallapersand P, Mashele SS, Kovačič L, Stojan J, Komel R, Pakala SB, et al. (2014) Cytochrome P450 Monooxygenase CYP53 Family in Fungi: Comparative Structural and Evolutionary Analysis and Its Role as a Common Alternative Anti-Fungal Drug Target. PLoS ONE 9(9): e107209. https://doi.org/10.1371/journal.pone.0107209

Kelly SL, Kelly DE (2013). Microbial cytochromes P450: biodiversity and biotechnology. Where do cytochromes P450 come from, what do they do and what can they do for us? Philos Trans R Soc B. 368, 20120476.

Kosmidis C, & Denning DW (2015). The clinical spectrum of pulmonary aspergillosis. Thorax 70, 270–7

Kodner RB, Summons RE, Pearson A, King N, Knoll AH (2008). Sterols in a unicellular relative of the metazoans. Proc Natl Acad Sci 105, 9897–9902

Miceli MH, Kauffman CA (2015). Isavuconazole: a new broad-spectrum triazole antifungal agent. Clin Infect Dis 61, 1558–1565.

Nebert DW, Wikvall K, Miller WL (2013). Human cytochromes P450 in health and disease. Philos Trans R Soc B. 368, 20120431

© Central University of Technology, Free State

Patterson TF, Kirkpatrick WR, White M et al (2000). Invasive aspergillosis. Disease spectrum, treatment practices, and outcomes. Aspergillus Study Group. Medicine (Baltimore) 79, 250–260

Verweij, P. E., Chowdhary, A., Melchers, W. J., & Meis, J. F. (2016). Azole Resistance in *Aspergillus fumigatus*: Can We Retain the Clinical Use of Mold-Active Antifungal Azoles? *Clinical infectious diseases : an official publication of the Infectious Diseases Society of America*, *62*(3), 362–368. doi:10.1093/cid/civ885

Vandeputte, P, Ferrari, S, & Coste, A. 2012. Antifungal Resistance and New Strategies to Control Fungal Infections, International Journal of Microbiology, vol. 2012, Article ID 713687, 26 pages, 2012

Zarrin, M, Faramarzi, S, (2018). Study of Azole Resistant and CYP51a Gene in *Aspergillus Fumigatus* 6(5): 747 – 750. doi:10.3889/oamjms.2018.171

© Central University of Technology, Free State

CHAPTER 3: METHODOLOGY

3.1. Study Design and Sampling

The present study was carried out with secondary data and quantitative approach. The study was based on various in-silico methods that ranged from BLAST to modeling of the protein (CYP51) of interest.

3.2. Procedure

3.2.1. BLAST

BLAST analysis was initially carried out with CYP51 from Aspergillus species. The reason for carrying out the BLAST analysis was to identify the nucleic acid sequence and amino acid sequence of the CYP51 from different strains of Aspergillus. The Aspergillus strains were then explored based on their historical data on azole resistance which was obtained from evidence-based literature and recent publications (Song et al., 2016). Another aspect of the BLAST analysis was to select homologous proteins to CYP51 with the highest identity. However, the identity was selected at a cut off at minimum 98%. The structural configuration of the CYP51 genome was studied in terms of chain characteristics, sequence pattern and tandem repeats, and the chemical properties of the peptide chain such as hydrophobicity (Mueller et al.,2002). The peptide chain CYP51 was also studied from the perspective of evidence based literature that explained azole resistance in Aspergillus species. The azole resistant strains along with the homologous proteins were next subjected to FASTA.

© Central University of Technology, Free State

3.2.2.FASTA

The FASTA was carried out for multiple sequence analysis (MSA) with the homologous species of CYP51. The BLAST sequences were selected for the FASTA analysis. The MSA was done by comparing the first twenty species (with 98% to 100% fit-identity) with the subsequent nineteen species in decreasing order of identity that was obtained from the Uniprot –database (Pundir et al., 2016). The twenty species were ticked in the box and were subjected to the ALIGN button which took to the FASTA format. The FASTA was run with the twenty species to provide details on the similarity and identity on different parts of the genome. The FASTA alignment was further subjected to phylogenetic analysis that revealed the evolutionary trend and cladistics from where the species diverged over the period of time (Pundir et al., 2016). The FASTA analysis was done with appropriate template selection of the ABC transporter protein because it is one of the major proteins that are implicated in drug resistance of various microbes to antimicrobials including the azole group. The FASTA sequences were compared on the basis of chain lengths, DNA binding ability, and transmembrane helices orientation.

The amino acid properties of the respective chains were also searched for the similarity to explore the homology along with the hydrophobic properties, negative and positive charge residues orientation, the assembly of the aliphatic and aromatic groups along with polar residues and serine-threonine assembly (Huang et al, 2011). The similar positions were noted to study the conservativeness of the sequences based on consensus sequences and conserved sequences. The similar sequences would be searched for single nucleotide polymorphisms to extrapolate azole sensitiveness and azole resistance of the species of Aspergillus and the allied FASTA proteins.

© Central University of Technology, Free State

3.2.3. Phylogenetic analysis

Phylogenetic analysis was next carried out to evaluate the evolutionary divergence of the respective proteins based on their structure function correlates (Krimitzas et al., 2013). The structure function correlates were evaluated from evidence-based literature and was compared with the phylogenetic traits related to the parsimony principle (Mueller et al., 2002). The cladistics analysis with the divergence pattern helped to identify whether a specific protein of the CYP51 (and the organism that housed it) was becoming sensitive or resistance to the azole group of antibiotics. The resistance pattern was also correlated with the mutational hotspots in the similar sequences. The similarity and identity was carried out for Nucleotide and protein FASTA individually. The nucleotide FASTA helped to identify the single nucleotide polymorphisms along with the mis-sense or non-sense mutations that might have translated into the azole resistance. The phylogenetic tree was further used for prediction of evolutionary transitions on azole resistance and other antifungal (Samson et al., 2011).

3.2.4.Statistical Computing Analysis

SCA was carried out with position matrix of the conserve and consensus sequences to understand whether the mutational hotspots in one of the sequences is dependent on the mutation or orientation of the amino acid sequences in another sequence within the same peptide chain (Reese & Pearson, 2002). The prediction was based on the Pearson's correlation coefficient which include the length of the similar sequences, the number of single nucleotide polymorphisms observed in such positions, and the type of amino acids that differed on account of such polymorphisms (Mueller et al., 2002). The corretaion analysis would be also further supported by BLOSUM62 and BLOSUM 50 which would provide the deep scoring matrices (Pearson, 2013).

© Central University of Technology, Free State

3.2.5 Protein Modelling

Protein modelling was carried out with the PyMol software to understand the structure function correlates of the mutated versions of the CYP51 protein and non-mutated versions of the CYP51 protein to understand the putative targets for azoles and other anti-fungal. The protein modelling was also undertaken to evaluate the possible structure and class of antifungal that would be effective in mitigating antimicrobial resistance in Aspergillus species (Mueller et al., 2002).

3.2.6. In-Silico and Statistical Software

The in-silico software that was used for this study to conduct the BLAST, FASTA, and phylogenetic analysis was the UniProt and CLUSTAL-W software respectively while the PyMOl software was used for modelling the proteins of interest (Huang et al., 2011, Samson et al., 2011). The statistical analysis was carried out with SPSS software that includes correlation and regression analysis. The correlation analysis was conducted with different variables such as the length of the similar sequences, the mutational hotspots, and the mis-sense mutations in the conserved and consensus sequences. On the other hand, regression analysis was carried out with parsimony attributes and cladogram distances as the dependent variable with length of the similar sequences, the mutational hotspots, and the mis-sense mutations in the conserved and consensus sequences as the independent variables (Krimitzas et al., 2013). The regression analysis would be evaluated both at the goodness of fit level and the significance of the ANOVA as well as the p-value of the individual beta-coefficients of the independent variables with the dependent variable. All statistical tests of inference would be carried out at the 0.05 level of significance. The p-value of the intercept would be also explored because if the value is significant it would imply that there could be other variables length of the similar sequences, the mutational hotspots, and the mis-sense mutations in

© Central University of Technology, Free State

the conserved and consensus sequences that could influence parsimony and cladogram distances. The parsimony attributes and cladogram distances are an estimate of the predictable mutations or sensitivity of the CYP51 protein to azole and other antifungal (Samson et al., 2011).

3.3. Research Questions and Hypothesis Testing

The present study explored various main and sub-research questions research questions. The sub-research questions were conducted to answer the main research question in comprehensive and conclusive manner. The main research question explored in the study was:

MRQ: Whether in-silico analysis of CYP450s in Eurotiomycete species promises identification of putative targets that have turned resistant to azole group of drugs?

H0: In-silico analysis of CYP450s in Eurotiomycete species does not promise identification of putative targets that have turned resistant to azole group of drugs

H1: In-silico analysis of CYP450s in Eurotiomycete species promises identification of putative targets that have turned resistant to azole group of drugs

The sub-research questions and their respective hypothesis that were explored in this study were as follows:

SRQ1: Whether there is any significant correlation between length of the similar sequences, the mutational hotspots, and the mis-sense mutations in the conserved and consensus sequences, parsimony, and cladogram distance of the CYP51 protein and its homologues?

© Central University of Technology, Free State

HO: There is no significant correlation between length of the similar sequences, the mutational hotspots, and the mis-sense mutations in the conserved and consensus sequences, parsimony, and cladogram distance of the CYP51 protein and its homologues ($p>0.05$)

H1: There is significant correlation between length of the similar sequences, the mutational hotspots, and the mis-sense mutations in the conserved and consensus sequences, parsimony, and cladogram distance of the CYP51 protein and its homologues ($p<0.05$)

SRQ2: Whether parsimony of CYP51 in Aspergillus species and its homologues could be significantly predicted from length of the similar sequences, the mutational hotspots, and the mis-sense mutations in the conserved and consensus sequences?

H0: parsimony of CYP51 in Aspergillus species and its homologues could not be significantly predicted from length of the similar sequences, the mutational hotspots, and the mis-sense mutations in the conserved and consensus sequences ($p>0.05$)

H1: parsimony of CYP51 in Aspergillus species and its homologues could be significantly predicted from length of the similar sequences, the mutational hotspots, and the mis-sense mutations in the conserved and consensus sequences ($p< 0.05$)

SRQ3: Whether cladogram distance of CYP51 in Aspergillus species and its homologues could be significantly predicted from length of the similar sequences, the mutational hotspots, and the mis-sense mutations in the conserved and consensus sequences?

H0: cladogram distance of CYP51 in Aspergillus species and its homologues could not be significantly predicted from length of the similar sequences, the mutational hotspots, and the mis-sense mutations in the conserved and consensus sequences ($p>0.05$)

© Central University of Technology, Free State

H1: cladogram distance of CYP51 in Aspergillus species and its homologues could be significantly predicted from length of the similar sequences, the mutational hotspots, and the mis-sense mutations in the conserved and consensus sequences ($p < 0.05$).

© Central University of Technology, Free State

References

Huang Hongzhan, McGarvey Peter B, Suzek Baris E, Mazumder Raja, Zhang Jian, Chen Yongxing, Wu Cathy H. A comprehensive protein-centric ID mapping service for molecular data integration. Bioinformatics. 2011 Apr 15;27(8):1190–1191.

Krimitzas, A., Pyrri, I., Kouvelis, V. N., Kapsanaki-Gotsi, E., & Typas, M. A. (2013). A phylogenetic analysis of Greek isolates of Aspergillus species based on morphology and nuclear and mitochondrial gene sequences. *BioMed research international*, *2013*, 260395. https://doi.org/10.1155/2013/260395

Mueller T, Spang R, Vingron M (2002). Estimating amino acid substitution models: a comparison of Dayhoff's estimator, the resolvent approach and a maximum likelihood method. Mol. Biol. Evol. 19, 8–13

Pearson W. R. (2013). Selecting the Right Similarity-Scoring Matrix. *Current protocols in bioinformatics*, *43*, 3.5.1–3.5.9. https://doi.org/10.1002/0471250953.bi0305s43

Pundir, S., Martin, M. J., O'Donovan, C., & UniProt Consortium (2016). UniProt Tools. *Current protocols in bioinformatics*, *53*, 1.29.1–1.29.15. https://doi.org/10.1002/0471250953.bi0129s53

Reese JT, & Pearson WR (2002). Empirical determination of effective gap penalties for sequence comparison. Bioinformatics. 18:1500–1

Song, J., Zhai, P., Zhang, Y., Zhang, C., Sang, H., Han, G., Keller, N. P., & Lu, L. (2016). The *Aspergillus fumigatus* Damage Resistance Protein Family Coordinately Regulates Ergosterol Biosynthesis and Azole Susceptibility. *mBio*, *7*(1), e01919-15. https://doi.org/10.1128/mBio.01919-15

© Central University of Technology, Free State

Samson RA, Varga J, Meijer M, Frisvad JC (2011). New taxa in *Aspergillus* section *Usti* . *Studies in Mycology*. 69, 81–97

© Central University of Technology, Free State

CHAPTER 4: BLAST AND FASTA RESULTS

The present chapter reflected the multiple sequence alignment (FASTA) of CYP450s of the Aspergillum strains considered for this study as well as BLAST for finding out the sequence identity and evolutionary history of CYP450. The Blast analysis further reflected the mutational hotspots and the single nucleotide polymorphisms that could implicate the genesis of resistance towards the azole antibiotics. The FASTA was carried out on the outputs of BLAST for those proteins having at least 40% identity with Cytochrome P450 oxidoreductase (the primary protein with which BLAST was conducted).

4.1. Query Sequence for Blast and relevant outputs

A0A0E1S5J5	100.0%
A0A0J6YMN0	99.8%
E9CUH7	97.3%
A1DKZ2	42.8%
A0A0K8L8Q3	42.7%
A0A0S7DN39	42.3%
A0A2I1CEY9	42.6%
A0A397H4E7	41.8%
Q4WCN0	42.1%
A0A3R7JEG6	40.4%
A1CMZ8	40.7%
A0A2V1DR91	40.7%
B8NL59	40.1%

© Central University of Technology, Free State

A0A0F0IM13	39.9%
D4QC90	39.9%
A0A2G7FJL1	39.9%
A0A1F8AAA9	39.7%
A0A0L1IS94	40.7%
A0A0L1J786	40.3%
A0A1F7ZUA7	39.7%
A0A1V6Q3U3	38.6%
A0A2G7FJU7	40.5%
B8NI79	39.9%

BLAST with CYP450 indicated that azole resistance is primarily a function of the CYP51 protein because the first 20 sequences did not elucidate any ABC transporter protein homology with CYP450 species in general to indicate azole resistance.

4.2. Azole resistance CYP51 Homology Matching with Different Aspergillum species

i. Query Sequences and FASTA with CYP51 of Different Aspergillum strains

TR|Q9P8R0|Q9P8R0_ASPFM

MVPMLWLTAYMAVAVLTAILLNVVYQLFFRLWNRTEPPMVFHWVPFLGSTISYGIDPY

KF 60

TR|A0A0H4U5P5|A0A0H4U5P5_ASPFL -

MSWPRIGAYALLAFVAIMALNVTYQFLFRMLNKTRPPLVFHWIPFIGSTIHYGTDPYGF

59

© Central University of Technology, Free State

TR|G4WW88|G4WW88_ASPTU MAYLAVAGVYAFAALLVAIVLNVARQLLVR--

NEKEPPVVFHWIPFLGSTISYGMDPYAF 58

TR|G4WWA2|G4WWA2_ASPAW MALLAVAGVYAFAALLVAIVLNVTRQLLFR--

NEKEPPVVFHWIPFLGSTISYGMDPYAF 58

TR|G4WWA3|G4WWA3_ASPNG MALLAVAGVYAFAALLVAIVLNVTRQLLFR--

NEKEPPVVFHWIPFLGSTISYGMDPYTF 58

TR|G4WW93|G4WW93_9EURO MAYLAVAGAYAFAALLVAIVLNVARQLLVR--

NEKEPPVVFHWIPFLGSTISYGMDPYAF 58

TR|G4WW85|G4WW85_ASPNG MALLAVAGVYAFAALLVAIVLNVTRQLLFR--

NEKEPPVVFHWIPFLGSTISYGMDPYTF 58

TR|G4WW98|G4WW98_9EURO MAYLAVAGAYAFAALLVAIVLNVARQLLVR--

NEKEPPVVFHWIPFLGSTISYGMDPYAF 58

.* *.:. : ***. *::.* *...**:****:**:**** ** *** *

TR|Q9P8R0|Q9P8R0_ASPFM

FFACREKYGDIFTFILLGQKTTVYLGVQGNEFILNGKLKDVNAEEVYSPLTTPVFGSDVV

120

TR|A0A0H4U5P5|A0A0H4U5P5_ASPFL

FFSCREKYGDIFTFILLGRPTTVYLGTQGNEFILNGKLKDVNAEEVYSPLTTPVFGSDVV

119

TR|G4WW88|G4WW88_ASPTU

FFSCRKKYGDIFTFVLLGKKTTVYLGVQGNDFILNGKLKDVSAEEVYSPLTTPVFGSDV

V 118

© Central University of Technology, Free State

TR|G4WWA2|G4WWA2_ASPAW

FFSCRKKYGDIFTFVLLGQKTTVYLGVQGNDFILNGKLKDVSAEEVYSPLTTPVFGSDV

V 118

TR|G4WWA3|G4WWA3_ASPNG

FFSCRKKYGDIFTFVLLGQKTTVYLGVQGNDFILNGKLKDVSAEEVYSPLTTPVFGSDV

V 118

TR|G4WW93|G4WW93_9EURO

FFFCRKKYGDIFTFVLLGKKTTVYLGVQGNDFILNGKLKDVSAEEVYSPLTTPVFGSDV

V 118

TR|G4WW85|G4WW85_ASPNG

FFSCRKKYGDIFTFVLLGQKTTVYLGVQGNDFILNGKLKDVSAEEVYSPLTTPVFGSDV

V 118

TR|G4WW98|G4WW98_9EURO

FFSCRKKYGDIFTFVLLGKKTTVYLGVQGNDFILNGKLKDVSAEEVYSPLTTPVFGSDV

V 118

** **:********:***: ******.***:**********.******************

TR|Q9P8R0|Q9P8R0_ASPFM

YDCPNSKLMEQKKFIKYGLTQSALESHVPLIEKEVLDYLRDSPNFQGSSGRMDISAAMA

E 180

TR|A0A0H4U5P5|A0A0H4U5P5_ASPFL

YDCPNSKLIEQKKFIKFGLSQAALEAHVPLIEKEVEDYLAMSPNFHGTSGEVDIPAAMAE

179

© Central University of Technology, Free State

TR|G4WW88|G4WW88_ASPTU

YDCPNSKLMEQKKFIKFGLTQAALESHVQLIEKETLDYLRDSPRFNGASGVIDIPAAMAE

178

TR|G4WWA2|G4WWA2_ASPAW

YDCPNSKLMEQKKFIKFGLTQAALESHVQLIEKETLDYLRDSPRFNGASGVIDIPAAMAE

178

TR|G4WWA3|G4WWA3_ASPNG

YDCPNSKLMEQKKFIKFGLTQAALESHVQLIEKETLDYLRDSPRFNGASGVIDIPAAMAE

178

TR|G4WW93|G4WW93_9EURO

YDCPNSKLMEQKKFIKFGLTQAALESHVQLIEKETLDYLRDSPRFNGASGVIDIPAAMAE

178

TR|G4WW85|G4WW85_ASPNG

YDCPNSKLMEQKKFIKFGLTQAALESHVQLIEKETLDYLRDSPRFNGASGVIDIPAAMAE

178

TR|G4WW98|G4WW98_9EURO

YDCPNSKLMEQKKFIKFGLTQVALESHVQLIEKETLDYLRDSPRFNGASGVIDIPAAMAE

178

********:*******:**:* ***:** *****. *** **.*:*:** :** *****

TR|Q9P8R0|Q9P8R0_ASPFM

ITIFTAARALQGQEVRSKLTAEFADLYHDLDKGFTPINFMLPWAPLPHNKKRDAAHARM

R 240

© Central University of Technology, Free State

TR|A0A0H4U5P5|A0A0H4U5P5_ASPFL

ITIFTAGSALQGEEVRSKLTTEFAVLYHDLDKGFTPINFMLPWAPLPHNKKRDAAHARM

R 239

TR|G4WW88|G4WW88_ASPTU

ITIYTAARALQGEEVRKKLTAEFAELYHDLDKGFSPINFMLPWAPLPHNRKRDAAHARM

R 238

TR|G4WWA2|G4WWA2_ASPAW

ITIYTAARALQGEEVRKKLTAEFAELYHDLDKGFSPINFMLPWAPLPHNRKRDAAHARM

R 238

TR|G4WWA3|G4WWA3_ASPNG

ITIYTAARALQGEEVRKKLTAEFAELYHDLDKGFSPINFMLPWAPLPHNRKRDAAHARM

R 238

TR|G4WW93|G4WW93_9EURO

ITIYTAGRALQGEEVRKKLTAEFAELYHDLDKGFSPINFMLPWAPLPHNRKRDAAHARM

R 238

TR|G4WW85|G4WW85_ASPNG

ITIYTAARALQGEEVRKKLTAEFAELYHDLDKGFSPINFMLPWAPLPHNQKRDAAHARM

R 238

TR|G4WW98|G4WW98_9EURO

ITIYTAARALQGEEVRKKLTAEFAELYHDLDKGFSPINFMLPWAPLPHNRKRDAAHARM

R 238

:**. *:***.***:*** *********:**************:**********

© Central University of Technology, Free State

TR|Q9P8R0|Q9P8R0_ASPFM

SIYVDIINQRRLDGDKDSQKSDMIWNLMNCTYKNGQQVPDKEIAHMMITLLMAGQHSS

SS 300

TR|A0A0H4U5P5|A0A0H4U5P5_ASPFL

AIYIDIINKRRNAGNNVPEKLDMIGNLMQCTYKNGQPLPDKEIAHVMITLLMAGQHSSSS

299

TR|G4WW88|G4WW88_ASPTU EIYTDIINERRKNP--

DEEKSDMIWNLMHCTYKNGQPVPDKEIAHMMITLLMAGQHSSSS 296

TR|G4WWA2|G4WWA2_ASPAW EIYTDIINERRKNP--

DEEKSDMIWNLMHCTYKSGQPVPDKEIAHMMITLLMAGQHSSSS 296

TR|G4WWA3|G4WWA3_ASPNG EIYTDIINERRKNP--

DEEKSDMIWNLMHCTYKSGQPVPDKEIAHMMITLLMAGQHSSSS 296

TR|G4WW93|G4WW93_9EURO EIYTDIINERRKNP--

DEEKSDMIWNLMHCTYKNGQPVPDKEIAHMMITLLMAGQHSSSS 296

TR|G4WW85|G4WW85_ASPNG EIYTDIINERRKNP--

DEEKSDMIWNLMHCTYKSGQPVPDKEIAHMMITLLMAGQHSSSS 296

TR|G4WW98|G4WW98_9EURO EIYTDIINERRKNP--

DEEKSDMIWNLMHCTYKNGQPVPDKEIAHMMITLLMAGQHSSSS 296

** ****:**   :* *** ***:****.** :*******:**************

TR|Q9P8R0|Q9P8R0_ASPFM

ISAWIMLRLASQPKVLEELYQEQLANLGPAGPDGSLPPLQYKDLDKLPFHQHVIRETLRI

360

© Central University of Technology, Free State

TR|A0A0H4U5P5|A0A0H4U5P5_ASPFL

ISSWIMLRLASQPAVVEELYQEQLANLERTGPNGSLAPLQYKDFDNLPLHQNVIRETLRL

359

TR|G4WW88|G4WW88_ASPTU ISSWIMLRLASEPQVLEELYQEQLASLS--

NRNGVFEPLQYQDLDKLPLLQSVIKETLRI 354

TR|G4WWA2|G4WWA2_ASPAW ISSWIMLRLASEPQVLEELYQEQLASLS--

NRNGVFEPLQYQDLDKLPFLQSVIKETLRI 354

TR|G4WWA3|G4WWA3_ASPNG ISSWIMLRLASEPQVLEELYQEQLASLS--

NRNGVFEPLQYQDLDKLPFLQSVIKETLRI 354

TR|G4WW93|G4WW93_9EURO ISSWIMLRLASEPQVLEELYQEQLASLS--

NRSGVFEPLQYQDLDKLPLLQSVIKETLRI 354

TR|G4WW85|G4WW85_ASPNG ISSWIMLRLASEPQVLEELYQEQLASLS--

NRNGVFEPLQYQDLDKLPFLQSVIKETLRI 354

:*******:* *:**********.* ..* : ****:*:*:**: * **:****:

TR|Q9P8R0|Q9P8R0_ASPFM

HSSIHSIMRKVKSPLPVPGTPYMIPPGRVLLASPGVTALSDEHFPNAGCWDPHRWENQA

T 420

TR|A0A0H4U5P5|A0A0H4U5P5_ASPFL

HSSIHSLLRKVKNPLPVPGTPYVIPTSHVLLAAPGVTALSDEYFPNAMAWDPHRWETQA

P 419

© Central University of Technology, Free State

TR|G4WW88|G4WW88_ASPTU

HSSIHSIMRKVKNPLPVPGTSYVIPEDRVLLASPGVTALSDEYFPNATRWDPHRWENQL

D 414

TR|G4WWA2|G4WWA2_ASPAW

HSSIHSIMRKVKNPLPVPGTSYIIPEDHVLLASPGVTALSDEYFPNATRWDPHRWENQPD

414

TR|G4WWA3|G4WWA3_ASPNG

HSSIHSIMRKVKNPLPVPGTSYIIPEDHVLLASPGVTALSDEYFPNATRWDPHRWENQPD

414

TR|G4WW93|G4WW93_9EURO

HSSIHSIMRKVKNPLPVPGTSYVIPEDRVLLASPGVTALSDEYFPNATRWDPHRWENQPD

414

TR|G4WW85|G4WW85_ASPNG

HSSIHSIMRKVKNPLPVPGTSYIIPEDHVLLASPGVTALSDEYFPNATRWDPHRWENQPD

414

TR|G4WW98|G4WW98_9EURO

HSSIHSIMRKVKNPLPVPGTSYVIPEDRVLLASPGVTALSDEYFPNATRWDPHRWENQSD

414

******::****.******* *:** .:****:*********:**** *******.*

TR|Q9P8R0|Q9P8R0_ASPFM

KEQENDEVVDYGYGAVSKGTSSPYLPFGAGRHRCIGEKFAYVNLGVILATIVRHLRLFN

V 480

© Central University of Technology, Free State

TR|A0A0H4U5P5|A0A0H4U5P5_ASPFL

QENNKDDIVDYGYGAMSKGTSSPYLPFGAGRHRCIGEKFAYLNLAVIVATMVRHLRFS

NL 479

TR|G4WW88|G4WW88_ASPTU KE-

EEGEMVDYGYGSVSKGTASPYLPFGAGRHRCIGEKFAYVNLGVIIATIVRHLKLFNV 473

TR|G4WWA2|G4WWA2_ASPAW KE-

EDGEMVDYGYGSVSKGTASPYLPFGAGRHRCIGEKFAYVNLGVIIATIVRHLKLFNV 473

TR|G4WWA3|G4WWA3_ASPNG KE-

EDGEMVDYGYGSVSKGTASPYLPFGAGRHRCIGEKFAYVNLGVIIATIVRHLKLFNV 473

TR|G4WW93|G4WW93_9EURO KE-

EEGEMIDYGYGSVSKGTASPYLPFGAGRHRCIGEKFAYVNLGVIIATIVRHLKLFNV 473

TR|G4WW85|G4WW85_ASPNG KE-

EDGEMVDYGYGSVSKGTASPYLPFGAGRHRCIGEKFAYVNLGVIIATIVRHLKLFNV 473

TR|G4WW98|G4WW98_9EURO KE-

EEGEMVDYGYGSVSKGTASPYLPFGAGRHRCIGEKFAYVNLGVIIATIVRHLKLFNV 473

:* :..:::*****::****:**********************:**.**:**:****:: *:

TR|Q9P8R0|Q9P8R0_ASPFM DGKKGVPETDYSSLFSGPMKPSIIGWEKRSKNTSK----

515

TR|A0A0H4U5P5|A0A0H4U5P5_ASPFL

DGQTGVPDTDYSSLFSGPMKPARIRWERRAAKSG----- 513

TR|G4WW88|G4WW88_ASPTU

DGRKGVPGTDYSTLFSGPLKPAIVGWERRVQDNSKGSSN 512

© Central University of Technology, Free State

TR|G4WWA2|G4WWA2_ASPAW

DGRKGVPGTDYSTLFSGPMKPAIVGWERRFPDNSKGSLN 512

TR|G4WWA3|G4WWA3_ASPNG

DGRKGVPGTDYSTLFSGPMKPAIVGWERRFPDNIKGSMN 512

TR|G4WW93|G4WW93_9EURO

DGRKGVPGTDYSTLFSGPMKPAIVGWERRFPDNSKGSLN 512

TR|G4WW85|G4WW85_ASPNG

DGRKGVPGTDYSTLFSGPMKPAIVGWERRFPDHSKGSLN 512

TR|G4WW98|G4WW98_9EURO

DGRKGVPGTDYSTLFSGPMKPAIVGWERRFPDNSKGSLN 512

:.* ****:*****:**: : **:* .

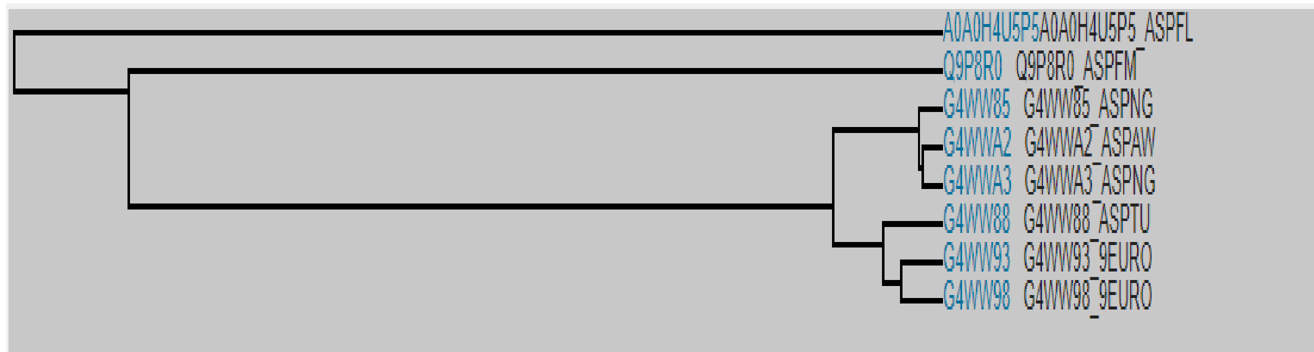
Identical positions: 358

Identity: 68.97%

Similar positions: 90

© Central University of Technology, Free State

Fig 1: Phylogenetic Analysis



The phylogenetic analysis reflected that the out-group was ASPFL, while the Aspergillus strain (ASPFM) evolved from the common ancestral origin of ASPFL. The phylogenetic analysis further revealed that the *Aspergillus* straints ASPNG, ASPAW, and Eurotiomycetes in general might have common ancestral origin to ASPFM (which is the most common Aspergillus strain) but they significantly formed distinct in-groups with evolutionary time period. As ASPNG and EURO had common ancestral origins, these strains might be the mutated versions of ASPFM and could house the genes for azole resistance.

4.3. Query sequences and FASTA with CYP450*Asperigillus fumigatus*

CLUSTAL O(1.2.4) multiple sequence alignment

TR|Q4WNT5|Q4WNT5_ASPFU

MVPMLWLTAYMAVAVLTAILLNVVYQLFFRLWNRTEPPMVFHWVPYLGSTISYGIDPY

KF 60

© Central University of Technology, Free State

TR|A1CXT9|A1CXT9_NEOFI

MVPMLLLTAYMAVAMLTAILLNVVYQLFFRLWNRTEPPMVFHWVPFLGSTISYGIDPY

KF 60

TR|I6LTQ6|I6LTQ6_9EURO

MVSMLLLTAYTAVAMLTVILLNVVYQLFFRLWNRTEPPMVFHWVPFLGNTISYGIDPY

KF 60

TR|A0A0K8L6G4|A0A0K8L6G4_9EURO

MVPMLLLTAYMAVAMLTVILLNVVYQLFFRLWNRTEPPMVFHWIPFLGSTISYGIDPYK

F 60

TR|A0A2I1C642|A0A2I1C642_9EURO

MVSMLLLTACMAVAMLTVILLNVVYQLFFRLWNRTEPPMVFHWVPFLGNTISYGIDPY

KF 60

TR|A0A3R7JL91|A0A3R7JL91_9EURO

MVPMLLLTAYMAVAMLTVILLNVVHQLFFRLWNRTEPPMVFHWVPFLGSTISYGIDPY

KF 60

TR|A0A397GX35|A0A397GX35_9EURO

MVSMLLLTAYTAVAMLAVILLNVVHQLFFRLWNRTEPPIVFHWVPFLGSTISYGIDPYK

F 60

TR|A1CGZ8|A1CGZ8_ASPCL

MLSLTLFGLYLVSATAVVILVNVVYQHLFRLRNRTEPPMVFHWIPFIGSTITYGIDPCKF

60

© Central University of Technology, Free State

TR|A0A2I2FZW3|A0A2I2FZW3_9EURO

MLSLASIGVYGLLAFVASIVLNLVYQWVFAKLNKTEPPLVFHWIPFLGNTISYGMDPYG

F 60

TR|A0A1E3BBY0|A0A1E3BBY0_9EURO

MLSISLLGVYMAVAMPIAILINVAYQLLFRRWNKTQPPLVFHWIPFLGSTISYGIDPYKF

60

```
*: : :    *   *::*..:* .*  *:*:**:****:*::*.**:**:**  *
```

TR|Q4WNT5|Q4WNT5_ASPFU FFACREK------------

YGDIFTFILLGQKTTVYLGVQGNEFILNGKLKDVNAEEVYS 108

TR|A1CXT9|A1CXT9_NEOFI FFACREK------------

YGDIFTFILLGQKTTVYLGVQGNEFILNGKLKDVNAEEVYS 108

TR|I6LTQ6|I6LTQ6_9EURO FFACRER------------

YGDIFTFILLGQKTTVYLGVQGNEFILNGKLKDVNAEEVYS 108

TR|A0A0K8L6G4|A0A0K8L6G4_9EURO FFACREK------------

YGDIFTFILLGQRTTVYLGVQGNEFILNGKLKDVNAEEVYS 108

TR|A0A2I1C642|A0A2I1C642_9EURO FFACREK------------

YGDIFTFILLGQKTTVYLGVQGNEFILNGKLKDVNAEEVYS 108

TR|A0A3R7JL91|A0A3R7JL91_9EURO LFACREN------------

YGDVFTFILLGQKTTVYLGVQGNEFILNGKLKDVNAEEVYS 108

TR|A0A397GX35|A0A397GX35_9EURO

FFACREKARSAEYCFVNWQYGDIFTFVLLGQKTTVYLGVQGNEFILNGKLKDVNAEEV

YS 120

TR|A1CGZ8|A1CGZ8_ASPCL FFACREK------------

YGNIFTFILLGQKVTVYLGVEGNEFILNGKLKDVNAEEVYS 108

TR|A0A2I2FZW3|A0A2I2FZW3_9EURO FFACREK------------

HGDIFTFILLGRKTTVYLGVQGNEFILNGKLKDVNAEEVYS 108

TR|A0A1E3BBY0|A0A1E3BBY0_9EURO FFDCQEK------------

YGDLFTFILLGQKTTVYLGVQGNEFILNGKLKDVNAEEVYS 108

:* *:*. :*::***:***::.******:********************

TR|Q4WNT5|Q4WNT5_ASPFU

PLTTPVFGSDVVYDCPNSKLMEQKKFIKYGLTQSALESHVPLIEKEVLDYLRDSPNFQGS

168

TR|A1CXT9|A1CXT9_NEOFI

PLTTPVFGSDVVYDCPNSKLMEQKKFIKYGLTQSALESHVPLIEKEVLDYLRNSPNFQGS

168

TR|I6LTQ6|I6LTQ6_9EURO

PLTTPVFGSDVVYDCPNSKLMEQKKFIKYGLTQSALESHVPLIEKEVLDYLRNSPNFQGS

168

TR|A0A0K8L6G4|A0A0K8L6G4_9EURO

PLTTPVFGSDVVYDCPNSKLMEQKKFIKYGLTQSALESHVPLIEKEVLDYLRNSPNFQGS

168

TR|A0A2I1C642|A0A2I1C642_9EURO

PLTTPVFGSDVVYDCPNSKLMEQKKFIKYGLTQSALEAHVPLIEKEVLDYLRKSPNFQGP

168

© Central University of Technology, Free State

TR|A0A3R7JL91|A0A3R7JL91_9EURO

PLTTPVFGSDVVYDCPNSKLMEQKKFIKYGLTQSALESHVPLIEKEVLDYLGNSPNFQGS

168

TR|A0A397GX35|A0A397GX35_9EURO

PLTTPVFGSDVVYDCPNSKLMEQKKFIKYGLTQSALESHVQLIEKEVLDYLRNSPNFQGS

180

TR|A1CGZ8|A1CGZ8_ASPCL

PLTTPVFGSDVVYDCPNSKFMEQKKFIKFGLTQSALEAHVPLIEKEVLDYLETSPRFQGT

168

TR|A0A2I2FZW3|A0A2I2FZW3_9EURO

PLTTPVFGSDVVYDCPNSKLMEQKKFIKFGLTQAALESHVPLIEKETLDYLATSPKFNGP

168

TR|A0A1E3BBY0|A0A1E3BBY0_9EURO

PLTTPVFGSDVVYDCPNAKLMEQKKFIKFGLTQSALESHVPLIEKEVLEYLDTSPYFQGS

168

****************:*:*********:****:***:** *****.*:** ** *:*

TR|Q4WNT5|Q4WNT5_ASPFU

SGRVDISAAMAEITIFTAARALQGQEVRSKLTAEFADLYHDLDKGFTPINFMLPWAPLPH

228

TR|A1CXT9|A1CXT9_NEOFI

SGQVDISAAMAEITIFTAARALQGQEVRSKLTAEFADLYHDLDKGFTPINFMLPWAPLPH

228

© Central University of Technology, Free State

TR|I6LTQ6|I6LTQ6_9EURO

SGQVDISAAMAEITIFTAARALQGQEVRSKLTAEFADLYHDLDKGFTPINFMLPWAPLPH

228

TR|A0A0K8L6G4|A0A0K8L6G4_9EURO

SGQVDISAAMAEITIFTAARALQGQEVRSKLTAEFADLYHDLDKGFTPINFMLPWAPLPH

228

TR|A0A2I1C642|A0A2I1C642_9EURO

SGQVDISAAMAEITIFTAARALQGEEVRSKLTAEFADLYHDLDKGFTPINFMLPWAPLPH

228

TR|A0A3R7JL91|A0A3R7JL91_9EURO

SGQVDISAAMAEITIFTAARALQGQEVRSKLTAEFADLYHDLDKGFTPINFMLPWAPLPH

228

TR|A0A397GX35|A0A397GX35_9EURO

SGQVDISAAMAEITIFTAARALQGQEVRSKLTAEFADLYHDLDKGFTPINFMLPWAPLPH

240

TR|A1CGZ8|A1CGZ8_ASPCL

SGLVDIAAAMAEITIFTAARALQGEEVRSKLTAEFADLYHDLDRCFTPVNFMFPWAPLP

R 228

TR|A0A2I2FZW3|A0A2I2FZW3_9EURO

SGRVDVGAAMAEITIFTAARALQGEEVRSKLTAEFADLYHDLDKGFTPINFMLPWAPLP

H 228

© Central University of Technology, Free State

TR|A0A1E3BBY0|A0A1E3BBY0_9EURO

SGRVNISAAMAEITIFTAARALQGEEVRSKLTAEFADLYHDLDKGFSPINFMLPWAPLPH

228

** *:::.*****************.******************: *:*:***:******:

TR|Q4WNT5|Q4WNT5_ASPFU NKKRDAAHARMRSIYVDIITQRRLD--

GEKDSQKSDMIWNLMNCTYKNGQQVPDKEIAHM 286

TR|A1CXT9|A1CXT9_NEOFI NKKRDAAHARMRSIYIDIINQRRLD--

GEKDSQKSDMIWNLMNSTYKNGQQVPDKEIAHM 286

TR|I6LTQ6|I6LTQ6_9EURO NKKRDAAHARMRSIYVDIINQRRLD--

GGKDSQKSDMIWNLMNSTYKNGQQVPDKEIAHM 286

TR|A0A0K8L6G4|A0A0K8L6G4_9EURO NKKRDAAHARMRSIYVDIINQRRLD--

GGKDSQKSDMIWNLMNSTYKNGQQVPDKEIAHM 286

TR|A0A2I1C642|A0A2I1C642_9EURO NKKRDAAHARMRSIYVDIINQRRLD--

GDKDSQKSDMIWNLMNSTYKNGQQVPDKEIAHM 286

TR|A0A3R7JL91|A0A3R7JL91_9EURO NKKRDAAHARMRSIYVDIINQRRLD--

GEKDSQKSDMIWNLMNSTYKNGQQVPDKEIAHM 286

TR|A0A397GX35|A0A397GX35_9EURO NKKRDAAHARMRSIYVDIINQRRLD--

GGMDSQKSDMIWNLMNCTYKNGQQVPDKEIAHM 298

TR|A1CGZ8|A1CGZ8_ASPCL NKKRDAAHVRMRAIYVDIINQRRRD--

GGEDTQKSDMIWNLMNCSYKNGQQVPDKEIAHM 286

TR|A0A2I2FZW3|A0A2I2FZW3_9EURO NKKRDAAHARMRAIYTDIINERRRA--

GESATHTSDMISNLMNCVYKNGTPVPDKEIAHM 286

© Central University of Technology, Free State

TR|A0A1E3BBY0|A0A1E3BBY0_9EURO

NNKRDAAHARMRAIYTDIINDRRHGEKGYNTTQTSDMISNLMNCTYKTGQPIPDKEIAH

M 288

*:******.***:** ***.:** * ::.**** ****. **.* :********

TR|Q4WNT5|Q4WNT5_ASPFU

MITLLMAGQHSSSSISAWIMLRLASQPKVLEELYQEQLANLGPAGPDGSLPPLQYKDLD

K 346

TR|A1CXT9|A1CXT9_NEOFI

MITLLMAGQHSSSSISAWIMLRLASQPKVLEELYQEQLANLGPVGPDGSLPPLQYKDLD

K 346

TR|I6LTQ6|I6LTQ6_9EURO

MITLLMAGQHSSSSISAWIMLRLASQPQVLEELYQEQLANLGPIGPDGNLPPLQYKDLDK

346

TR|A0A0K8L6G4|A0A0K8L6G4_9EURO

MITLLMAGQHSSSSIGAWIMLRLASQPEVLDELYQEQLANLGPIGPDGSLPPLQYEDLDK

346

TR|A0A2I1C642|A0A2I1C642_9EURO

MITLLMAGQHSSSSISAWIMLRLASQPRVLEELYQEQLANLGPIGPDGSLPPLQYRDLDK

346

TR|A0A3R7JL91|A0A3R7JL91_9EURO

MITLLMAGQHSSSSISAWIMLRLASQPEVLEELYQEQLANLGSIGPKGSLPSLQYKDLDK

346

© Central University of Technology, Free State

TR|A0A397GX35|A0A397GX35_9EURO

MITLLMAGQHSSSSISAWIMLRLASQPEVIEELYQEQLANLGRIGPGGSLPSLQYKDLDK

358

TR|A1CGZ8|A1CGZ8_ASPCL

MITLLMAGQHSSSSIGSWIMLRLASQPEVLEKLYQEQLDKLAQGGPGSNLRPLQYKDLE

L 346

TR|A0A2I2FZW3|A0A2I2FZW3_9EURO

MITLLMAGQHSSSSISSWIMLRLASQPEVIEELYQEQLAQLGPVGPDGTLPPLQYKDLDQ

346

TR|A0A1E3BBY0|A0A1E3BBY0_9EURO

MITLLMAGQHSSSSIGSWIMLRLAAHPEVIEELYQEQLANLDRDG--KKLPPLRYQDLDR

346

***************.:*******::*.*:::****** :*   *   .* *:*.**:

TR|Q4WNT5|Q4WNT5_ASPFU

LPFHQHVIRETLRIHSSIHSIMRKVKSPLPVPGTPYMIPPGRVLLASPGVTALSDEHFPN

406

TR|A1CXT9|A1CXT9_NEOFI

LPFHQHVVRETLRLHSSIHSLMRKVKSPLPVPGTPYMIPPGRVLLASPGVTALSDEHFPN

406

TR|I6LTQ6|I6LTQ6_9EURO

LPFHQHVVRETLRLHSSIHSLMRKVKSPLPVPGTPYVIPPGRVLLASPGVTALSDEHFPN

406

© Central University of Technology, Free State

TR|A0A0K8L6G4|A0A0K8L6G4_9EURO

LPFHQHVIRETLRLHSSIHSLMRKVKSPLPVPGTPYLIPPGRVLLASPGVTALSDEYFPN

406

TR|A0A2I1C642|A0A2I1C642_9EURO

LPFHQHVVRETLRLHSSIHSLMRKVKSPLPVPGSPYVIPPGRVLLASPGVTALSDEHFPN

406

TR|A0A3R7JL91|A0A3R7JL91_9EURO

LPLHQHVVRETLRLHSSIHSLMRKVKSPLPVPGTPYVIPPGRVLLASPGVTALSDEYFPN

406

TR|A0A397GX35|A0A397GX35_9EURO

LPFHQHVVRETLRLHSSIHSLMRKVKNPLSVPGTPYVIPPGRVLLASPGVTALSDEHFPG

418

TR|A1CGZ8|A1CGZ8_ASPCL

LPYHQHVIRETLRLHSSIHSILRKVKNTLSVPGTSYVIPPGRVLLASPGVTALSDEHFPN

406

TR|A0A2I2FZW3|A0A2I2FZW3_9EURO

LPLHQNVIRETLRVNNSIHSLMRKVKNPIPVPGTPYVIPTSHVLLASPGITALSDEYFPN

406

TR|A0A1E3BBY0|A0A1E3BBY0_9EURO

LPLHQNVIRETLRLHSSIHSLMRKVKNPLNVAGTPYIIPTSHVLLASPGVTALSDYYFPN

406

** **:*:*****:.:****::****.: * *: *:** .:*******:***** :**.

© Central University of Technology, Free State

TR|Q4WNT5|Q4WNT5_ASPFU AGCWDPHRWENQATKEQ---

ENDKVVDYGYGAVSKGTSSPYLPFGAGRHRCIGEKFAYVN 463

TR|A1CXT9|A1CXT9_NEOFI AGCWDPHRWENQAAKEQ---

ENDEVVDYGYGAVSKGTSSPYLPFGAGRHRCIGEKFAYVN 463

TR|I6LTQ6|I6LTQ6_9EURO AGCWDPHRWENQAAKEQ---

ENDEVVDYGYGAVSKGTSSPYLPFGAGRHRCIGEKFAYVN 463

TR|A0A0K8L6G4|A0A0K8L6G4_9EURO AGCWDPHRWENQASKEH---

ENDEVVDYGYGAVSKGTSSPYLPFGAGRHRCIGEKFAYLN 463

TR|A0A2I1C642|A0A2I1C642_9EURO AGCWDPHRWENQAAKEQ---

ENDEVVDYGYGAVSKGTSSPYLPFGAGRHRCIGEKFAYLN 463

TR|A0A3R7JL91|A0A3R7JL91_9EURO AGCWDPHRWENQAAKEQ---

ENDEVVDYGYGAVSKGTSSPYLPFGAGRHRCIGEKFAYLN 463

TR|A0A397GX35|A0A397GX35_9EURO AGRWDPHRWENQAAKEQ---

EDDEVVDYGYGAVSKGTSSPYLPFGAGRHRCIGEKFAYVN 475

TR|A1CGZ8|A1CGZ8_ASPCL ASRWDPQRWENQTTKED----

SGEMVDYGYGAVSKGTASPYLPFGAGRHRCIGEKFAYVN 462

TR|A0A2I2FZW3|A0A2I2FZW3_9EURO AMKWDPHRWETQAPTED---

KEEDVVDYGYGAVSKGTSSPYLPFGAGRHRCIGEKFAYVN 463

TR|A0A1E3BBY0|A0A1E3BBY0_9EURO

ASKWDPHRWDKGNQKDDNDDVNEEVVDYGYGAVSKGTSSPYLPFGAGRHRCIGEKFA

YVN 466

* ***:**:.   .:.   . ..:************:********************:*

© Central University of Technology, Free State

TR|Q4WNT5|Q4WNT5_ASPFU

LGVILATIVRHLRLFNVDGKKGVPETDYSSLFSGPMKPSIIGWEKRSKNTSK 515

TR|A1CXT9|A1CXT9_NEOFI

IGVILATIVRHLRLFNVDGKKGVPETDYSSLFSGPMKPSIIGWEKRSKDTSK 515

TR|I6LTQ6|I6LTQ6_9EURO

IGVILATIVRHLRLFNVDGKKGVPDTDYSSLFSGPMKPSIIRWEKRSKDSSK 515

TR|A0A0K8L6G4|A0A0K8L6G4_9EURO

IGVILATVVRHLRLFNVDEKKGVPETDYSSLFSGPMKPSIIGWEKRSKDVSK 515

TR|A0A2I1C642|A0A2I1C642_9EURO

IGVILATIVRHVRLFNVDGKKGVPETDYSSLFSGPMKPSIIGWEKRSKDTST 515

TR|A0A3R7JL91|A0A3R7JL91_9EURO

IGVILATIVRHLRLLNVDGKKGVPETDYSSLFSGPMKPSIIGWERRSKDTSK 515

TR|A0A397GX35|A0A397GX35_9EURO

IGVILATMVRHLRLFNVDGRKGVPETDYSSLFSGPMKPSIIGWEKRSDKSK- 526

TR|A1CGZ8|A1CGZ8_ASPCL

IGVILATLVRHLRLSNMDGKEGVPATDYSSLFSGPMKPSIIQWGKRSNDLSK 514

TR|A0A2I2FZW3|A0A2I2FZW3_9EURO

LTVIIASVVRTMRLHNVDGRKGVPGTDYSSLFSGPMKPAIIGWERRSAKSA- 514

TR|A0A1E3BBY0|A0A1E3BBY0_9EURO

LGVIVATIVRHVRFFNVDGKKGVPGTDYSSLFSGPVKPAVIGWERRDGQRKI 518

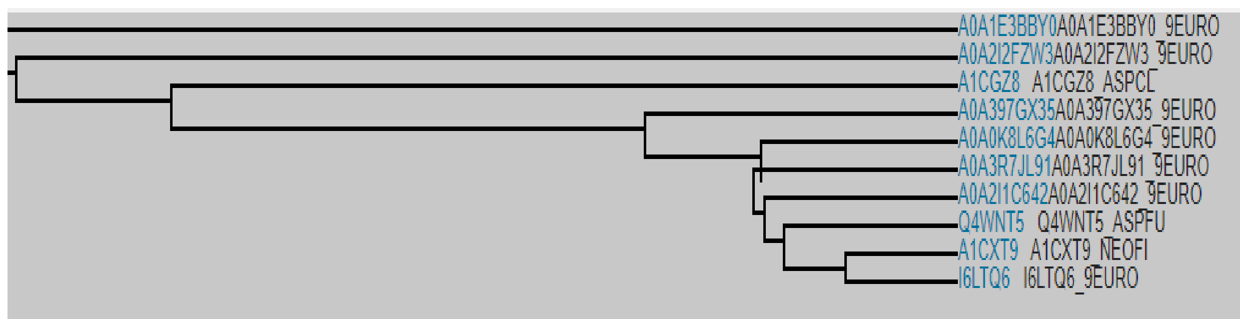
: **:*::** :*: *:* ::*** **********:**::* * :*. .

© Central University of Technology, Free State

Identity 331 positions 62.12%

Similarity 112 positions

Fig 2: Phylogenetic analysis with Eurotiomycetes CYP51



The FASTA with CYP450s revealed that the Aspergillus strains as well as those of the Eurotiomycetes family had common ancestry and the ancestry was evident as EURO strains being the out-group. The number of identical positions in CYP450 was 331 while that for CYP51 variant was 338. These findings suggest that the CYP51 had more conserved sequences than CYP450 over evolutionary time period. Such findings are not surprising because CYP51 in Aspergillus or other fungal strains is associated with specific function such as membrane fluidity and virulence while CYP450 accomplishes a plethora of functions from metabolism to drug transport. On the contrary, the number of similar positions in CYP51 was 90 compared to that of CYP450 which was 112. These findings suggested that CYP51 might have conserved sequences but they are more vulnerable to point mutations and single nucleotide substitutions. Therefore, it could be contended that azole resistance could stem from point mutations in CYP51 because it is the enzyme (14-alpha demethylase) that causes membrane fluidity and virulence of the fungal strain. Likewise, the referred enzyme (especially the active heme) is the major target of azoles.

© Central University of Technology, Free State

CHAPTER 5: HOMOLOGY MODELLING REPORT: AZOLE RESISTANCE/MULTI-DRUG TRANSPORTER FOR ASPERGILLUS SPECIES

5.1. Model Building Report

This document lists the results for the homology modelling project "Untitled Project" submitted to SWISS-MODEL workspace on Sept. 4, 2019, 8:39 a.m. The submitted primary amino acid sequence is given in Table T1.

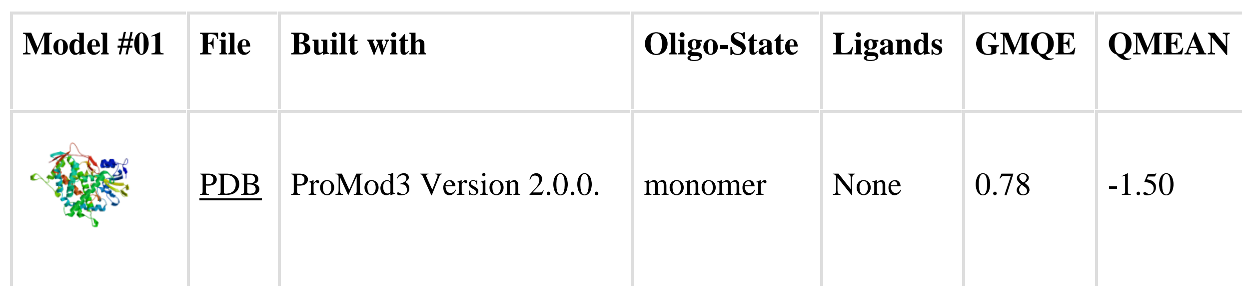
5.2. Results

The SWISS-MODEL template library (SMTL version 2019-08-28, PDB release 2019-08-23) was searched with BLAST (Camacho et al.) and HHBlits (Remmert et al.) for evolutionary related structures matching the target sequence in Table T1. For details on the template search, see Materials and Methods. Overall 1017 templates were found (Table T2).

Models

The following models were built (see Materials and Methods "Model Building"):

Table 1: Protein Model

Model #01	File	Built with	Oligo-State	Ligands	GMQE	QMEAN
	PDB	ProMod3 Version 2.0.0.	monomer	None	0.78	-1.50

© Central University of Technology, Free State

QMEAN -1.50

Cβ -1.46

All Atom -0.17

solvation 0.42

Torsion

Residue Number ★Protein Size (Residues)

-1.41

© Central University of Technology, Free State

Table 2: Template protein

Template	Seq Identity	Oligo-state	Found by	Method	Resolution	Seq Similarity	Range	Coverage	Description
6cr2.1.A	66.17	monomer	BLAST	X-ray	2.38Å	0.50	37 - 509	0.91	14-alpha sterol demethylase CYP51B

Table 3: Ligand Tested Model

Ligand	Added to Model	Description
HEM	✕ - Binding site not conserved.	PROTOPORPHYRIN IX CONTAINING FE
LFV	✕ - Binding site not conserved.	N-[(1R)-1-(2,4-dichlorophenyl)-2-(1H-imidazol-1-yl)ethyl]-4-{5-[2-fluoro-4-(2,2,2-trifluoroethoxy)phenyl]-1,3,4-oxadiazol-2-yl}benzamide

Target

MVPMLWLTAYMAVAVLTAILLNVVYQLFFRLWNRTEPPMVFHWVPFLGSTISYGIDPY

KFFFACREKYGDIFTFILLGQK

© Central University of Technology, Free State

6cr2.1.A ------------------------------------

PPVVFHWFPFIGSTISYGIDPYKFFFDCRAKYGDIFTFILLGKK

Target

TTVYLGVQGNEFILNGKLKDVNAEEVYSPLTTPVFGSDVVYDCPNSKLMEQKKFIKYGL

TQSALESHVPLIEKEVLDYLR

6cr2.1.A

TTVYLGTKGNDFILNGKLRDVCAEEVYSPLTTPVFGRHVVYDCPNAKLMEQKKFVKYG

LTSDALRSYVPLITDEVESFVK

Target

DSPNFQGSSGRMDISAAMAEITIFTAARALQGQEVRSKLTAEFADLYHDLDKGFTPINFM

LPWAPLPHNKKRDAAHARMR

6cr2.1.A

NSPAFQGHKGVFDVCKTIAEITIYTASRSLQGKEVRSKFDSTFAELYHNLDMGFAPINFM

LPWAPLPHNRKRDAAQRKLT

Target

SIYVDIINQRRLDGDKDSQKSDMIWNLMNCTYKNGQQVPDKEIAHMMITLLMAGQHSS

SSISAWIMLRLASQPKVLEELY

6cr2.1.A ETYMEIIKARRQAGSKKDSE-

DMVWNLMSCVYKNGTPVPDEEIAHMMIALLMAGQHSSSSTASWIVLRLATRPDIMEEL

Y

Target

QEQLANLGPAGPDGSLPPLQYKDLDKLPFHQHVIRETLRIHSSIHSIMRKVKSPLPVPGTP

YMIPPGRVLLASPGVTALS

6cr2.1.A QEQIRVLG-----

SDLPPLTYDNLQKLDLHAKVIKETLRLHAPIHSIIRAVKNPMAVDGTSYVIPTSHNVLSSP

GVTARS

Target

DEHFPNAGCWDPHRWENQATKEQENDEVVDYGYGAVSKGTSSPYLPFGAGRHRCIGE

KFAYVNLGVILATIVRHLRLFNV

6cr2.1.A

EEHFPNPLEWNPHRWDENIAASAEDDEKVDYGYGLVSKGTNSPYLPFGAGRHRCIGEQ

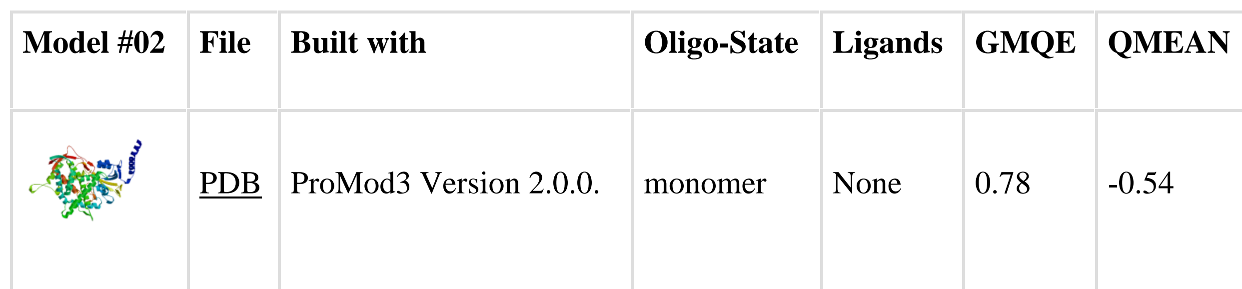
FAYLQLGTITAVLVRLFRFRNL

Target DGKKGVPETDYSSLFSGPMKPSIIGWEKRSKNTSK

6cr2.1.A PGVDGIPDTDYSSLFSKPLGRSFVEFEKR------

© Central University of Technology, Free State

Table 4: Protein Model



Model #02	File	Built with	Oligo-State	Ligands	GMQE	QMEAN
	PDB	ProMod3 Version 2.0.0.	monomer	None	0.78	-0.54

QMEAN	-0.54
Cβ	-1.28
All Atom	-0.38
solvation	0.59
torsion	-0.51

Residue Number

★Protein Size (Residues)

© Central University of Technology, Free State

Table 5: Template model

Template	**Seq Identity**	**Oligo-state**	**Found by**	**Method**	**Resolution**	**Seq Similarity**	**Range**	**Coverage**	**Description**
5jlc.1.A	48.48	monomer	BLAST	X-ray	2.40Å	0.44	7 - 509	0.96	Lanosterol 14-alpha demethylase

Table 6:Ligand tested Model

Ligand	**Added to Model**	**Description**
1PE	✕ - Not biologically relevant.	PENTAETHYLENE GLYCOL
1YN	✕ - Binding site not conserved.	2-[(2R)-butan-2-yl]-4-{4-[4-(4-{[(2R,4S)-2-(2,4-dichlorophenyl)-2-(1H-1,2,4-triazol-1-ylmethyl)-1,3-dioxolan-4-

© Central University of Technology, Free State

Ligand	Added to Model	Description
		yl]methoxy}phenyl)piperazin-1-yl]phenyl}-2,4-dihydro-3H-1,2,4-triazol-3-one
CL	✕ - Not biologically relevant.	CHLORIDE ION
CL	✕ - Not in contact with model.	CHLORIDE ION
HEM	✕ - Binding site not conserved.	PROTOPORPHYRIN IX CONTAINING FE
PGE	✕ - Not biologically relevant.	TRIETHYLENE GLYCOL

Target MVPMLWLTAYMAVAVLTAILLNVVYQLFFRLWNRTEPPMVFHWVPFLGSTISYGIDPY KFFFACREKYGDIFTFILLGQK

5jlc.1.A ------LAQRVSIMVALPFVYTITWQLLYSL-RKDRPPLVFYWIPWVGSAIPYGTKPYEFFEDCQKKYGDIFSFMLLGRI

© Central University of Technology, Free State

Target

TTVYLGVQGNEFILNGKLKDVNAEEVYSPLTTPVFGSDVVYDCPNSKLMEQKKFIKYGL

TQSALESHVPLIEKEVLDYLR

5jlc.1.A

MTVYLGPKGHEFIFNAKLADVSAEAAYSHLTTPVFGKGVIYDCPNHRLMEQKKFVKGA

LTKEAFVRYVPLIAEEIYKYFR

Target DSPNF---

QGSSGRMDISAAMAEITIFTAARALQGQEVRSKLTAEFADLYHDLDKGFTPINFMLPWA

PLPHNKKRDAAHA

5jlc.1.A

NSKNFKINENNSGIVDVMVSQPEMTIFTASRSLLGKEMRDKLDTDFAYLYSDLDKGFTPI

NFVFPNLPLEHYRKRDHAQQ

Target RMRSIYVDIINQRRLDGDKDSQKSDMIWNLM-

NCTYKNGQQVPDKEIAHMMITLLMAGQHSSSSISAWIMLRLASQPKVL

5jlc.1.A AISGTYMSLIKERR--

EKNDIQNRDLIDELMKNSTYKDGTKMTDQEIANLLIGVLMGGQHTSAATSAWCLLHLA

ERPDVQ

Target

EELYQEQLANLGPAGPDGSLPPLQYKDLDKLPFHQHVIRETLRIHSSIHSIMRKVKSPLPV

PGTPYMIPPGRVLLASPGV

© Central University of Technology, Free State

5jlc.1.A EELYQEQMRVL-----

NNDTKELTYDDLQNMPLLNQMIKETLRLHHPLHSLFRKVMRDVAIPNTSYVVPRDYHV

LVSPGY

Target TALSDEHFPNAGCWDPHRWE-

NQATKEQENDEVVDYGYGAVSKGTSSPYLPFGAGRHRCIGEKFAYVNLGVILATIVRHL

5jlc.1.A

THLQEEFFPKPNEFNIHRWDGDAASSSAAGGDEVDYGFGAISKGVSSPYLPFGGGRHRCI

GELFAYCQLGVLMSIFIRTM

Target RLFNVDGKKGVPETDYSSLFSGPMKPSIIGWEKRSKNTSK

5jlc.1.A KWRYPTEGETVPPSDFTSMVTLPTAPAKIYWEKR------

Primary amino acid sequence for which templates were searched and models were built.

MVPMLWLTAYMAVAVLTAILLNVVYQLFFRLWNRTEPPMVFHWVPFLGSTISYGIDPY

KFFFACREKYGDIFTFILLGQKTTVYLGVQGNEFILNGKLKD

VNAEEVYSPLTTPVFGSDVVYDCPNSKLMEQKKFIKYGLTQSALESHVPLIEKEVLDYL

RDSPNFQGSSGRMDISAAMAEITIFTAARALQGQEVRSKLT

AEFADLYHDLDKGFTPINFMLPWAPLPHNKKRDAAHARMRSIYVDIINQRRLDGDKDS

QKSDMIWNLMNCTYKNGQQVPDKEIAHMMITLLMAGQHSSSS

ISAWIMLRLASQPKVLEELYQEQLANLGPAGPDGSLPPLQYKDLDKLPFHQHVIRETLRI

HSSIHSIMRKVKSPLPVPGTPYMIPPGRVLLASPGVTALS

DEHFPNAGCWDPHRWENQATKEQENDEVVDYGYGAVSKGTSSPYLPFGAGRHRCIGE

© Central University of Technology, Free State

KFAYVNLGVILATIVRHLRLFNVDGKKGVPETDYSSLFSGPMK

PSIIGWEKRSKNTSK

Table 7: Template models

Template	Seq Identity	Oligo-state	QSQE	Found by	Method	Resolution	Seq Similarity	Coverage	Description
6cr2.1.A	66.17	monomer		BLAST	X-ray	2.38Å	0.50	0.91	14-alpha sterol demethylase CYP51B
4uyl.1.A	66.17	monomer		BLAST	X-ray	2.81Å	0.50	0.91	14-ALPHA STEROL DEMETHYLASE
5frb.1.A	66.17	monomer		BLAST	X-ray	2.99Å	0.50	0.91	STEROL 14-ALPHA DEMETHYLASE, CYP51B
5eah.1.A	52.09	monomer		BLAST	X-ray	2.54Å	0.45	0.93	Lanosterol 14-alpha demethylase

© Central University of Technology, Free State

Template	Seq Identity	Oligo-state	QSQE	Found by	Method	Resolution	Seq Similarity	Coverage	Description
5v5z.1.A	48.80	monomer		BLAST	X-ray	2.90Å	0.44	0.97	Lanosterol 14-alpha demethylase
4ze2.1.A	51.88	monomer		BLAST	X-ray	2.30Å	0.45	0.93	Lanosterol 14-alpha demethylase
5esl.1.A	52.30	monomer		BLAST	X-ray	2.35Å	0.45	0.93	Lanosterol 14-alpha demethylase
5hs1.1.A	52.09	monomer		BLAST	X-ray	2.10Å	0.45	0.93	Lanosterol 14-alpha demethylase
5jlc.1.A	48.48	monomer		BLAST	X-ray	2.40Å	0.44	0.96	Lanosterol 14-alpha demethylase
4uyl.2.A	66.24	monomer		HHblits	X-ray	2.81Å	0.50	0.91	14-ALPHA STEROL DEMETHYLASE

© Central University of Technology, Free State

Template	Seq Identity	Oligo-state	QSQE	Found by	Method	Resolution	Seq Similarity	Coverage	Description
4uyl.1.A	66.24	monomer		HHblits	X-ray	2.81Å	0.50	0.91	14-ALPHA STEROL DEMETHYLASE
4uym.2.A	66.24	monomer		HHblits	X-ray	2.55Å	0.50	0.91	14-ALPHA STEROL DEMETHYLASE
4uym.1.A	66.24	monomer		HHblits	X-ray	2.55Å	0.50	0.91	14-ALPHA STEROL DEMETHYLASE
5frb.1.A	66.38	monomer		HHblits	X-ray	2.99Å	0.50	0.91	STEROL 14-ALPHA DEMETHYLASE, CYP51B
5eaf.1.A	52.09	monomer		BLAST	X-ray	2.65Å	0.45	0.93	Lanosterol 14-alpha demethylase

© Central University of Technology, Free State

Template	Seq Identity	Oligo-state	QSQE	Found by	Method	Resolution	Seq Similarity	Coverage	Description
4zdy.1.A	51.88	monomer		BLAST	X-ray	2.02Å	0.45	0.93	Lanosterol 14-alpha demethylase
5eqb.1.A	52.09	monomer		BLAST	X-ray	2.19Å	0.45	0.93	Lanosterol 14-alpha demethylase
5jlc.1.A	49.39	monomer		HHblits	X-ray	2.40Å	0.44	0.95	Lanosterol 14-alpha demethylase
5v5z.1.A	50.42	monomer		HHblits	X-ray	2.90Å	0.44	0.92	Lanosterol 14-alpha demethylase
4lxj.1.A	52.09	monomer		BLAST	X-ray	1.90Å	0.45	0.93	Lanosterol 14-alpha demethylase
5fsa.2.A	51.07	monomer		BLAST	X-ray	2.86Å	0.44	0.91	CYP51 VARIANT1
5fsa.1.A	51.07	monomer		BLAST	X-ray	2.86Å	0.44	0.91	CYP51 VARIANT1

© Central University of Technology, Free State

Template	Seq Identity	Oligo-state	QSQE	Found by	Method	Resolution	Seq Similarity	Coverage	Description
5tz1.1.A	51.17	monomer		HHblits	X-ray	2.00Å	0.44	0.91	Sterol 14-alpha demethylase
5fsa.2.A	51.17	monomer		HHblits	X-ray	2.86Å	0.44	0.91	CYP51 VARIANT1
5fsa.1.A	51.17	monomer		HHblits	X-ray	2.86Å	0.44	0.91	CYP51 VARIANT1
5tz1.2.A	51.17	monomer		HHblits	X-ray	2.00Å	0.44	0.91	Sterol 14-alpha demethylase
6fmo.1.A	32.75	monomer		BLAST	X-ray	3.18Å	0.38	0.89	Sterol 14alpha-demethylase
6fmo.4.A	32.75	monomer		BLAST	X-ray	3.18Å	0.38	0.89	Sterol 14alpha-demethylase
6fmo.1.A	29.68	monomer		HHblits	X-ray	3.18Å	0.36	0.85	Sterol 14alpha-demethylase
6fmo.4.A	29.68	monomer		HHblits	X-ray	3.18Å	0.36	0.85	Sterol 14alpha-demethylase
3zg3.1.A	32.27	monomer		BLAST	X-ray	2.90Å	0.38	0.85	STEROL 14-ALPHA

© Central University of Technology, Free State

Template	Seq Identity	Oligo-state	QSQE	Found by	Method	Resolution	Seq Similarity	Coverage	Description
									DEMETHYLASE
4coh.1.A	32.57	monomer		BLAST	X-ray	2.08Å	0.38	0.85	STEROL 14-ALPHA DEMETHYLASE
4h6o.1.A	32.27	monomer		BLAST	X-ray	2.80Å	0.38	0.85	Sterol 14-alpha demethylase
2wuz.1.A	32.57	homo-dimer	0.18	BLAST	X-ray	2.35Å	0.38	0.85	LANOSTEROL 14-ALPHA-DEMETHYLASE, PUTATIVE
4c28.1.A	32.57	monomer		BLAST	X-ray	2.03Å	0.38	0.85	STEROL 14-ALPHA DEMETHYLASE
2wx2.1.A	32.57	monomer		BLAST	X-ray	2.27Å	0.38	0.85	LANOSTEROL 14-ALPHA-

© Central University of Technology, Free State

Template	Seq Identity	Oligo-state	QSQE	Found by	Method	Resolution	Seq Similarity	Coverage	Description
									DEMETHYLASE
2wuz.1.A	29.75	homo-dimer	0.17	HHblits	X-ray	2.35Å	0.36	0.85	LANOSTEROL 14-ALPHA-DEMETHYLASE, PUTATIVE
4coh.1.A	29.82	monomer		HHblits	X-ray	2.08Å	0.36	0.85	STEROL 14-ALPHA DEMETHYLASE
2wx2.1.A	29.75	monomer		HHblits	X-ray	2.27Å	0.36	0.85	LANOSTEROL 14-ALPHA-DEMETHYLASE
2wx2.2.A	29.75	monomer		HHblits	X-ray	2.27Å	0.36	0.85	LANOSTEROL 14-ALPHA-DEMETHYLASE

© Central University of Technology, Free State

Template	Seq Identity	Oligo-state	QSQE	Found by	Method	Resolution	Seq Similarity	Coverage	Description
4c28.1.A	29.82	monomer		HHblits	X-ray	2.03Å	0.36	0.85	STEROL 14-ALPHA DEMETHYLASE
3khm.1.A	29.52	monomer		HHblits	X-ray	2.85Å	0.36	0.85	Sterol 14 alpha-demethylase
5te8.1.A	18.71	homo-trimer	0.17	HHblits	X-ray	2.70Å	0.29	0.87	Cytochrome P450 3A4
5vc0.2.B	18.71	homo-tetramer	0.03	HHblits	X-ray	2.70Å	0.29	0.87	Cytochrome P450 3A4
5a1p.1.A	18.71	monomer		HHblits	X-ray	2.50Å	0.29	0.87	CYTOCHROME P450 3A4
5a1r.1.A	18.71	monomer		HHblits	X-ray	2.45Å	0.29	0.87	CYTOCHROME P450 3A4
5vcc.1.B	18.71	homo-dimer	0.20	HHblits	X-ray	1.70Å	0.29	0.87	Cytochrome P450 3A4
4k9x.1.A	18.71	monomer		HHblits	X-ray	2.76Å	0.29	0.87	Cytochrome P450 3A4

© Central University of Technology, Free State

Template	Seq Identity	Oligo-state	QSQE	Found by	Method	Resolution	Seq Similarity	Coverage	Description
4i3q.1.A	18.71	monomer		HHblits	X-ray	2.60Å	0.29	0.87	Cytochrome P450 3A4
4i4h.1.A	18.71	monomer		HHblits	X-ray	2.90Å	0.29	0.87	Cytochrome P450 3A4

© Central University of Technology, Free State

References

- Waterhouse, A., Bertoni, M., Bienert, S., Studer, G., Tauriello, G., Gumienny, R., Heer, F.T., de Beer, T.A.P., Rempfer, C., Bordoli, L., Lepore, R., Schwede, T. SWISS-MODEL: homology modelling of protein structures and complexes. Nucleic Acids Res. 46(W1), W296-W303 (2018)
- Guex, N., Peitsch, M.C., Schwede, T. Automated comparative protein structure modeling with SWISS-MODEL and Swiss-PdbViewer: A historical perspective. Electrophoresis 30, S162-S173 (2009).
- Bienert, S., Waterhouse, A., de Beer, T.A.P., Tauriello, G., Studer, G., Bordoli, L., Schwede, T. The SWISS-MODEL Repository - new features and functionality. Nucleic Acids Res. 45, D313-D319 (2017).
- Benkert, P., Biasini, M., Schwede, T. Toward the estimation of the absolute quality of individual protein structure models. Bioinformatics 27, 343-350 (2011).
- Bertoni, M., Kiefer, F., Biasini, M., Bordoli, L., Schwede, T. Modeling protein quaternary structure of homo- and hetero-oligomers beyond binary interactions by homology. Scientific Reports 7 (2017).
  Camacho, C., Coulouris, G., Avagyan, V., Ma, N., Papadopoulos, J., Bealer, K., Madden, T.L. BLAST+: architecture and applications. BMC Bioinformatics 10, 421-430 (2009).

© Central University of Technology, Free State

CHAPTER 6: PROTEIN MODELLING

In this protein modelling chapter, the two closest members of azole resistance fungi CP51A_ASPFU and ATRR_ASPFU were compared. The selection was based on the SWISS-Model results that gave the close members structurally. However, the structural correlates were compared with the functions that were narrated in chapter 5. The former is the azole resistance fungus and the latter is the ABC drug transporter protein obtained from Candida. The FASTA for the same is shown in the alignment as follows:

SP|Q4WNT5|CP51A_ASPFU --

SP|Q4WI89|ATRR_ASPFU

MDGIGDGTESMGFDMPMLMNQQPHLFGSYGHDGSPVAPIFSNPTFQDEPSIGAADDNSD

A 60

SP|Q4WNT5|CP51A_ASPFU --

SP|Q4WI89|ATRR_ASPFU

KRRRIARACDMCRKKKIKCDGKMPKCSHCTNYKTDCVFTQVEKKRNPPKGAKYIEGLE

NR 120

SP|Q4WNT5|CP51A_ASPFU --

SP|Q4WI89|ATRR_ASPFU

LGRMESLLRLSGLLSEDDGGKTDLGTLEKRLADRYHASGSNTPHNPQKINIPSQSQIAMS

180

© Central University of Technology, Free State

SP|Q4WNT5|CP51A_ASPFU ---

SP|Q4WI89|ATRR_ASPFU

QQNSSSHYSTPRLESQSSPRTAATSPESQKESETEVEGLSDMMCSLVTNNCGETRYIGSS

240

SP|Q4WNT5|CP51A_ASPFU ----------------------------------MVPMLWLTAYMAVA----

VLTAIL-- 20

SP|Q4WI89|ATRR_ASPFU

SGFSIFSPKGIQWVNEKTGDNSFQEMISSAYVDDNKWMYWKPEIFSDIFARRVFKPLPPK

300

* * :: *:. :

SP|Q4WNT5|CP51A_ASPFU --LNVVYQLFFRLWNR----TEPPMVFHW-------

VPYLGSTISYGIDP---------- 57

SP|Q4WI89|ATRR_ASPFU

DEAMSLFKDFFDNFNCMFPLYHEPTFMHLVERQYSRDPYEGSGWWASINVVLAIAHRL

RV 360

::: ** :* . * .:* ** ** .*:

© Central University of Technology, Free State

SP|Q4WNT5|CP51A_ASPFU ----------

YKFFFACREKYGDIFTFILLGQKTTVYLGVQGNEFILNGKLKDVNA---- 103

SP|Q4WI89|ATRR_ASPFU

MSNLVPQEEDRKAWLYLKNAMGVLTELTMRNTDLLSVQALLGMSLFLQGTPNPQPSFF

LV 420

* :: :: * : : : . . .: * .::*:*. : :

SP|Q4WNT5|CP51A_ASPFU ---EEVYSPLTTPVFGSDVVYDCPNSKLMEQKKFIKYGLTQS-

ALESHV-PLIEKEVLDY 158

SP|Q4WI89|ATRR_ASPFU

AAAIRLSHSIGLHKRGSGFGLNPVEVEQRKRVFWIAYLLDKDICLRSGRPPVQDDDDMN

V 480

.: : **.. : : : :: :* * * :. .*.* *: :.: ::

SP|Q4WNT5|CP51A_ASPFU LRDSPNFQGSSGRVDISAAMAEITIFTAARALQG-----------

QEVRSKLTAEFADLY 207

SP|Q4WI89|ATRR_ASPFU

ELPSDDPPDNIGNVPLSDGRSKFNLFRSMCRFATIESKVYKRLYSAKASKQSDGELLNTI

540

* : .. *.* :* . :::.:* : : :. .: .*: :

SP|Q4WNT5|CP51A_ASPFU HDLDKGFTPINFMLPWAPLPHNKKRDAAHARM-------

RSIYVDII-TQRRLDGEKDSQ 259

© Central University of Technology, Free State

SP|Q4WI89|ATRR_ASPFU GELDKELEDWKDSIPLDFRPEHEIK-ASHTPLILHVVVLHFAYYNCLTTIHRMS------ 593

:*** : : :* *.:: : *:*: : : * : : * :*:.

SP|Q4WNT5|CP51A_ASPFU KSDMIWNLMNCTYKNGQ--QVPD--KE-----------IAHMMITLLMAGQHSSSSISA 303

SP|Q4WI89|ATRR_ASPFU VHHGYWTSRLSNYAIQGLNARPLNPRVFLSAVLCVTAARASINLIKYIPQ----GDFACV 649

. *. ..* * : : :*. : .. ..

SP|Q4WNT5|CP51A_ASPFU WIMLRLASQPKVLEELYQEQLANL---------------------GPAGPDGSLP---- 337

SP|Q4WI89|ATRR_ASPFU WLILYYPVS--ALVTLFANILQNPSDARARSDVKLMNVVVNFLSTLVSDESNGSIKRMLG 707

:: . .* *: : * * :**:

SP|Q4WNT5|CP51A_ASPFU ---------PLQYKDLDKLPFHQHVI---RETLRIHSSIHSIMRKVKSPLPVPGTPYMIP 385

SP|Q4WI89|ATRR_ASPFU LCGEFERIAKVVLDKAEKESYSKKKRKSPEEPVNLQQ---STPEEHPAPSPSTTQPTQAP 764

: ..:* : :: .* :.::. * .: :* * * *

© Central University of Technology, Free State

SP|Q4WNT5|CP51A_ASPFU PGRVLLASP---------

GVTALSDEHFPNAGCWDPHRWENQATKEQENDKVVDYGYGAV 436

SP|Q4WI89|ATRR_ASPFU SRNVPMSSPLFAENPGDPGGNTMAD----DAGGFASSREIP-----

---------GT--- 803

.* ::** * .::* :** : * *

SP|Q4WNT5|CP51A_ASPFU

SKGTSSPYLPFGAGRHRCIGEKFAYVNLGVILATIVRHLRLFNVDGKKGVP---ETDYSS

493

SP|Q4WI89|ATRR_ASPFU -TGVST-NIPPNIQAMPGIAQD--YQ------------

DMLSPDPLEGVSFADQPPYSA 846

.*.*: :* . *.:. * ::. * :** : **:

SP|Q4WNT5|CP51A_ASPFU -------LFSGPMK-------PSIIGWEKRSKNTSK------------- 515

SP|Q4WI89|ATRR_ASPFU

TANTPLSSFQQPFVPQDLWQMPMTIEWDWADMSTNFPVFDTNGPPHGGL 895

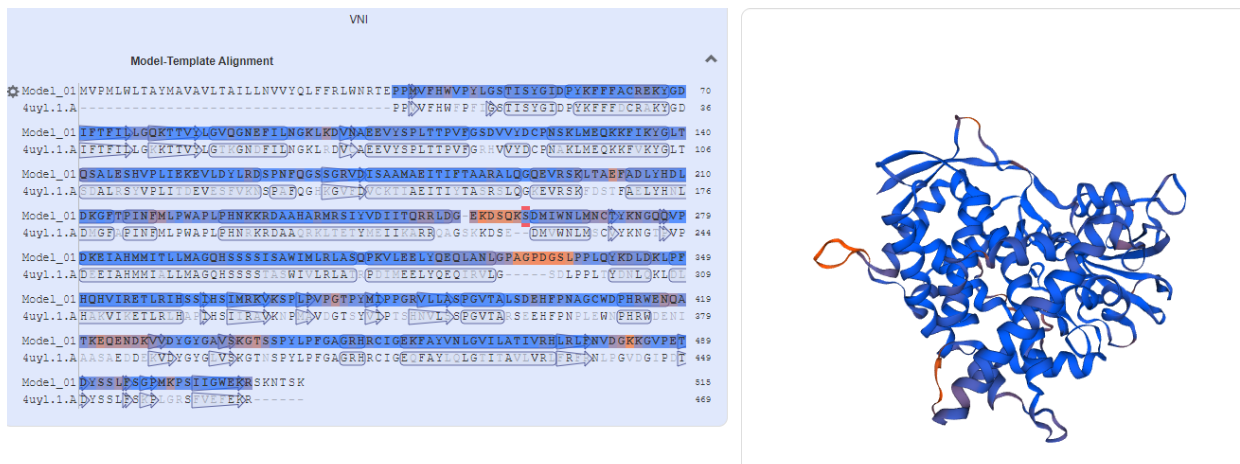
*. *: * * *: . .*.

The total identical positions and identity percentage that was revealed was 86 and 9% respectively. The number of similar positions was 156. The protein chain suggested that the azole resistance Aspergillus had 380 amino acid residues less compared to the ABC multidrug transporter protein. The loss of amino acids suggests that there could be minimum loss of the drug transporter (azole transporter) portions of the protein that led to the resistance of the pathogen against the antifungal

© Central University of Technology, Free State

drug. These findings are aligned with the previous studies that reported that deletion mutations in the CYP51 gene could attribute to azole resistance.

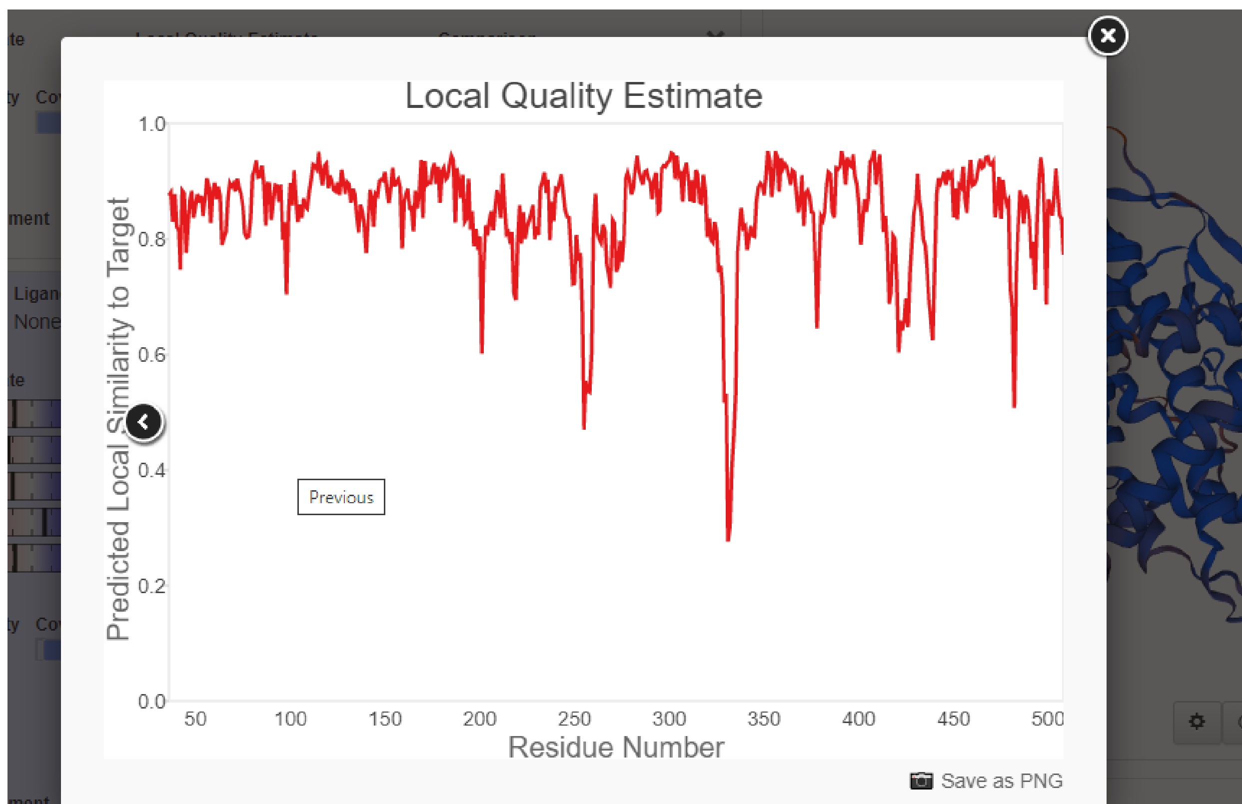
Fig 3: Protein Model



The protein modelling revealed that the mutational hotspots are present both within the folded chains that are mainly conserved regions (red) as well as in the N-terminal portions. Interestingly, this is the first study that showed that mutations in specific hotspots predisposed the chances of mutations in other hotspots for the same gene and protein. These assumptions were substantiated by Local quality estimate which showed predicted local similarity fluctuates in a rhythmic manner.

© Central University of Technology, Free State

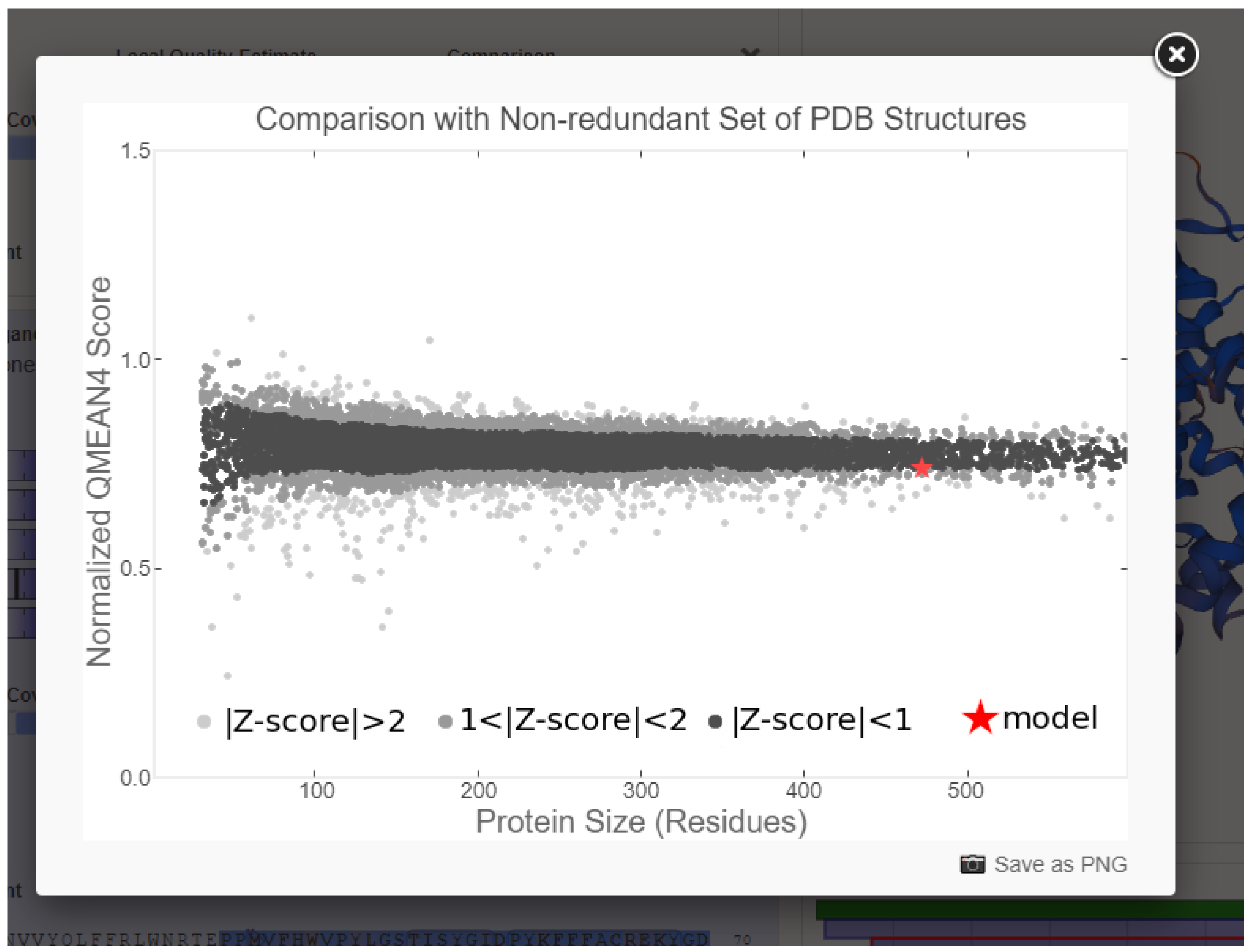
Fig 4: Local Quality Estimate



It could be possible that mutations (fluctuations) in one residue number or a set of residues could cause mutations (fluctuations) in other residues within the same molecule. These assumptions are substantiated by previous studies which had endorsed that the CYP51 is subjected to specific mutations and some of them are dependent on each other. However, there was no evidence till date that CYP51 of Eurotiomycetes also exhibit the same characteristics, especially in *Aspergillus fumigatus*. The conserved nature of CYP51 was further substantiated by the non-redundant PDB structures.

© Central University of Technology, Free State

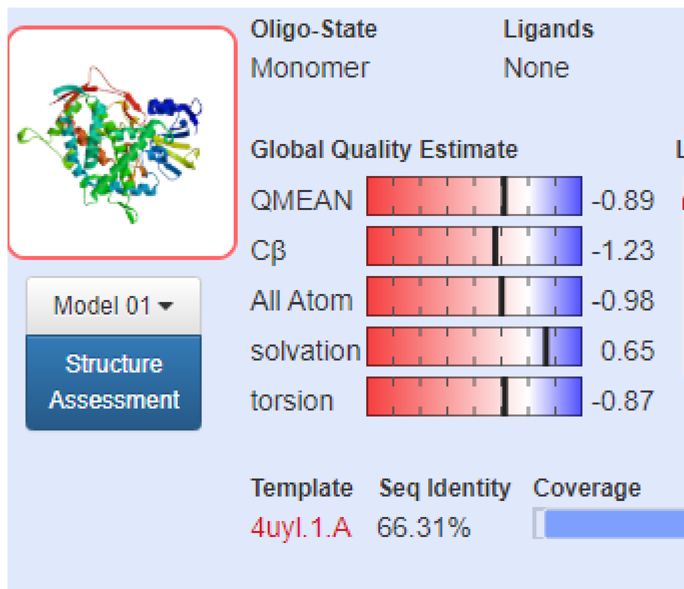
Fig 5: PDB Estimate



The non-redundant PDB structures showed that the 14-alpha demethylase (azole resistance) (model) marginally differed (yet significantly) in terms of QMEAN4 score from the normalized QMEAN4 score. This finding suggested that even minor fluctuations in PDB structure of CYP51 could attribute to azole resistance. The global quality estimates and comparison of azole resistance CYP51 with close variants are depicted as follows:

© Central University of Technology, Free State

Fig 6: Protein Model



14-ALPHA STEROL DEMETHYLASE

Seqres	KTPPVVFHWFPFIGSTISYGIDPYKFFFDCRAKYG	35
4uyl.1.A	KTPPVVFHWFPFIGSTISYGIDPYKFFFDCRAKYG	35
Seqres	DIFTFILLGKKTTVYLGTKGNDFILNGKLRDVCAE	70
4uyl.1.A	DIFTFILLGKKTTVYLGTKGNDFILNGKLRDVCAE	70
Seqres	EVYSPLTTPVFGRHVVYDCPNAKLMEQKKFVKYGL	105
4uyl.1.A	EVYSPLTTPVFGRHVVYDCPNAKLMEQKKFVKYGL	105
Seqres	TSDALRSYVPLITDEVESFVKNSPAFQGHKGVFDV	140
4uyl.1.A	TSDALRSYVPLITDEVESFVKNSPAFQGHKGVFDV	140

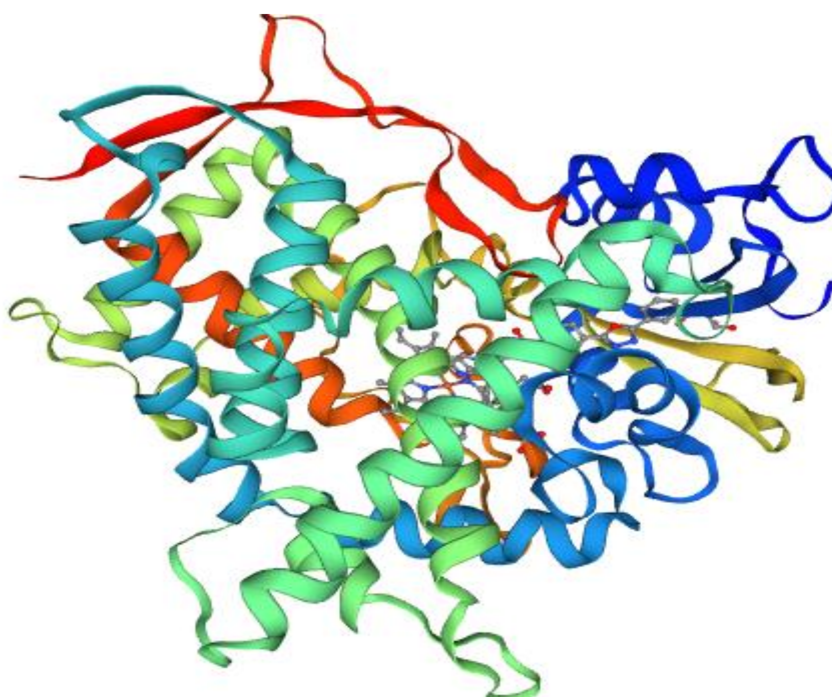
© Central University of Technology, Free State

Seqres	CKTIAEITIYTASRSLQGKEVRSKFDSTFAELYHN	175
4uyl.1.A	CKTIAEITIYTASRSLQGKEVRSKFDSTFAELYHN	175
Seqres	LDMGFAPINFMLPWAPLPHNRKRDAAQRKLTETYM	210
4uyl.1.A	LDMGFAPINFMLPWAPLPHNRKRDAAQRKLTETYM	210
Seqres	EIIKARRQAGSKKDSEDMVWNLMSCVYKNGTPVPD	245
4uyl.1.A	EIIKARRQAGSKKDSEDMVWNLMSCVYKNGTPVPD	245
Seqres	EEIAHMMIALLMAGQHSSSSTASWIVLRLATRPDI	280
4uyl.1.A	EEIAHMMIALLMAGQHSSSSTASWIVLRLATRPDI	280
Seqres	MEELYQEQIRVLGSDLPPLTYDNLQKLDLHAKVIK	315
4uyl.1.A	MEELYQEQIRVLGSDLPPLTYDNLQKLDLHAKVIK	315
Seqres	ETLRLHAPIHSIIRAVKNPMAVDGTSYVIPTSHNV	350
4uyl.1.A	ETLRLHAPIHSIIRAVKNPMAVDGTSYVIPTSHNV	350
Seqres	LSSPGVTARSEEHFPNPLEWNPHRWDENIAASAED	385
4uyl.1.A	LSSPGVTARSEEHFPNPLEWNPHRWDENIAASAED	385
Seqres	DEKVDYGYGLVSKGTNSPYLPFGAGRHRCIGEQFA	420
4uyl.1.A	DEKVDYGYGLVSKGTNSPYLPFGAGRHRCIGEQFA	420
Seqres	YLQLGTITAVLVRLFRFRNLPGVDGIPDTDYSSLF	455

© Central University of Technology, Free State

4uyl.1.A	YLQLGTITAVLVRLFRFRNLPGVDGIPDTDYSSLF	455
Seqres	SKPLGRSFVEFEKRH	470
4uyl.1.A	SKPLGRSFVEFEKRH	470

Fig 7: crystal structure of model protein



Crystal structure of sterol 14-alpha demethylase (CYP51B) from a pathogenic filamentous fungus *Aspergillus fumigatus* in complex with VNI showed that it was different from the azole resistance variant. The detailed structural correlates of azole resistance CYP51 in *Aspergillus fumigatus* is provided in the Ramachandran Plot as follows:

© Central University of Technology, Free State

Fig 8: Ramachandran Plot

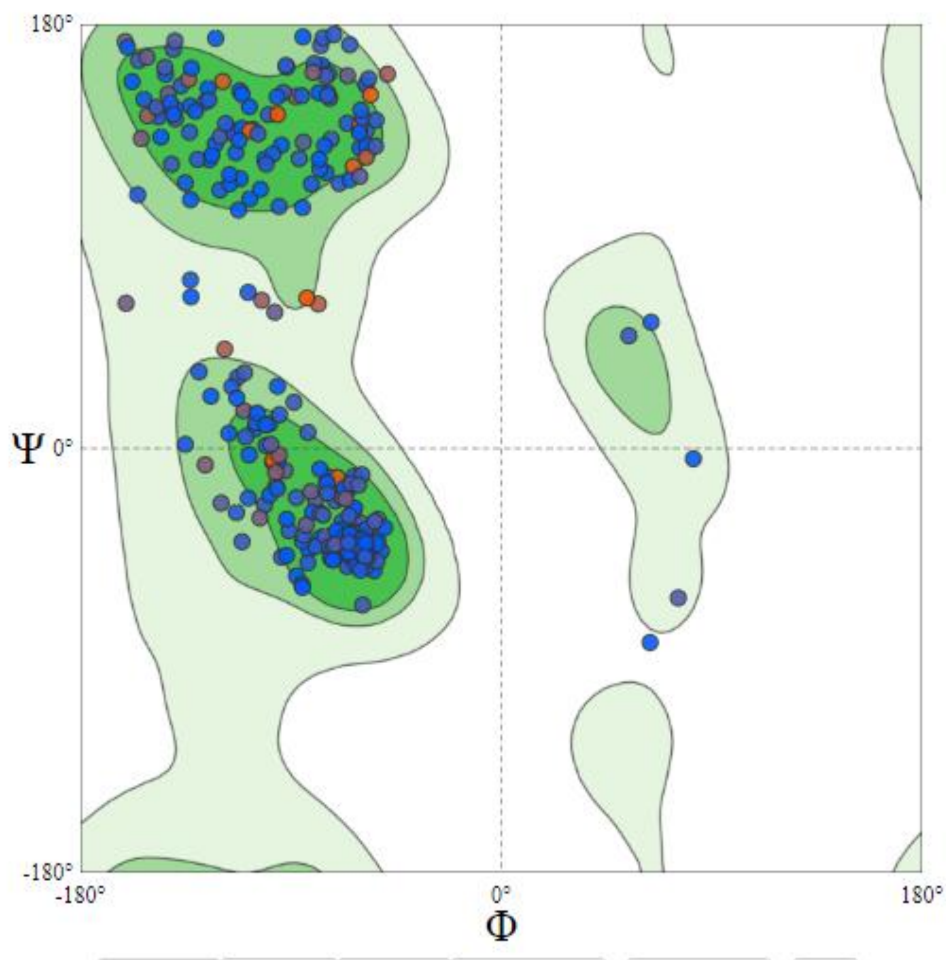


Table 8: Ramachandran Plot statistics

MolProbity Score	1.58
Clash Score	3.48

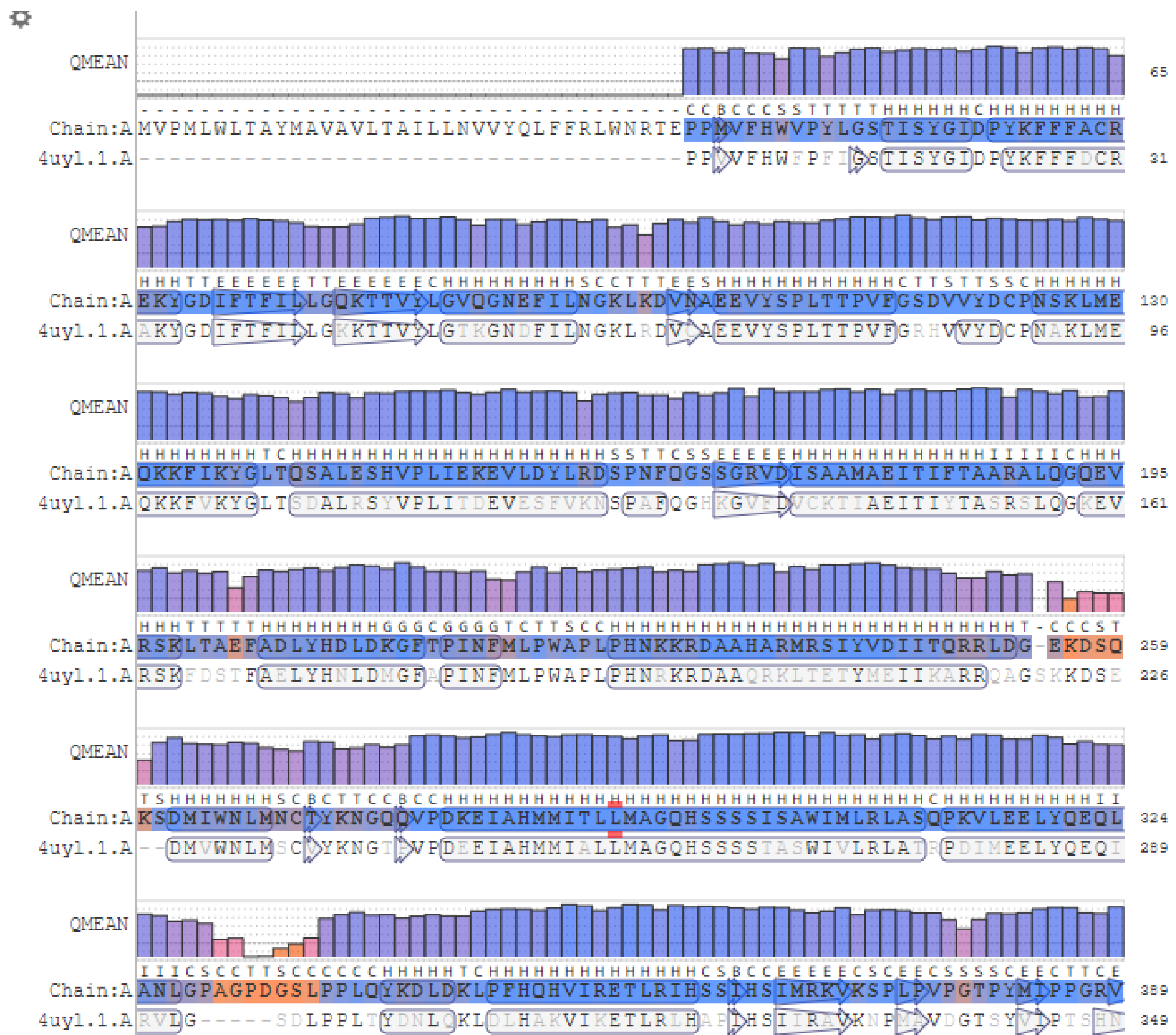
© Central University of Technology, Free State

Ramachandran Favoured	96.60%	
Ramachandran Outliers	0.85%	A221 LEU, A120 VAL, A225 PRO, A48 GLY
Rotamer Outliers	1.95%	A80 LYS, A146 SER, A377 VAL, A502 SER, A49 SER, A327 LEU, A494 LEU, A372 LYS
C-Beta Deviations	1	A265 TRP
Bad Bonds	0 / 3849	
Bad Angles	27 / 5221	(A221 LEU-A222 PRO), (A108 SER-A109 PRO), A404 PHE, A137 TYR, (A486 VAL-A487 PRO), (A148 VAL-A149 PRO), (A226 LEU-A227 PRO), A265 TRP, (A421 LYS-A422 GLU), (A337 PRO-A338 PRO), (A224 ALA-A225 PRO), A447 PHE, A264 ILE, A269 ASN, (A380 THR-A381 PRO), A228 HIS, (A112 THR-A113 PRO), A474 HIS, A365 HIS, A296 HIS, A413 HIS, (A336 LEU-A337 PRO), (A426 ASP-A427 LYS), A236 HIS, A438 LYS, A147 HIS
Cis Non-Proline	1 / 439	(A257 ASP-A258 SER)

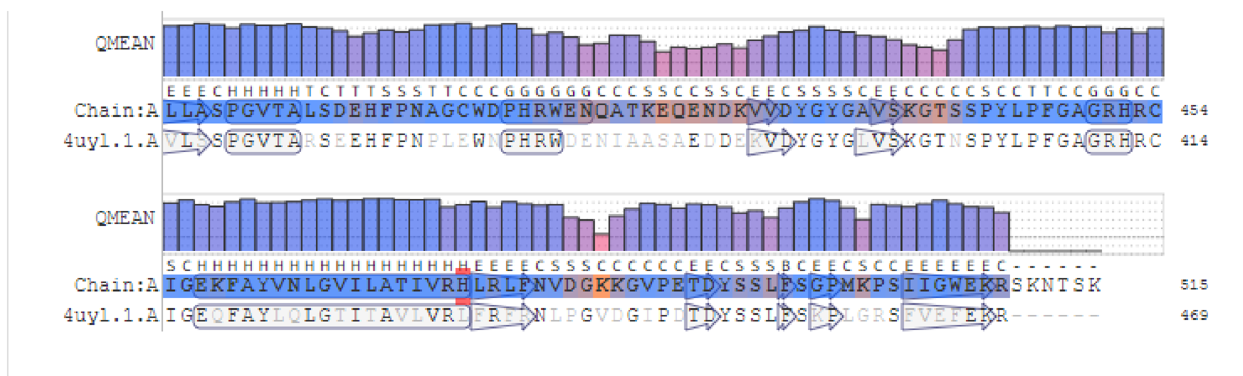
© Central University of Technology, Free State

Twisted Non-Proline	2 / 439	(A420 THR-A421 LYS), (A475 LEU-A476 ARG)
Twisted Prolines	1 / 33	(A500 LYS-A501 PRO)

Fig 9: Alignment of the models



© Central University of Technology, Free State



The Ramachandran plot endorsed that most of the CYP51 molecule is a B-pleated sheet and as a result there is a chance of finding folded chains and conserved regions. However, it could be possible that mutations in the molecule could change its conformation from Beta-pleated sheets to right handed ad left handed alpha helix. The change in conformation could disrupt the orientation of the active site in case of azole sensitive species while it could enter into a more beta pleated sheet to ensure that the active site s capable of converting lanosterol into ergosterol in case of azole–resistant isolates. CYP51 is a member of the cytochrome p450 monooxygenase super family that mediates the synthesis of ergosterol which is one of the important fungal-specific sterols. Moreover, CYP51 is a popular target of azole group of drugs. The protein mediates the permeability and fluidity of the fungal cell membrane by demethylation of 14 alpha position of lanosterol to form ergosterol (Bernhardt et al., 2018). Therefore formation of ergosterol is an important attribute of structural and functional correlates of a fungus. Cytochrome p450 monooxygenases are an abundant hemease super family which plays an important role in primary as well as secondary metabolism. These enzymes are primary responsible for catalyzing various oxidative processes and in playing a crucial role in heterogeneous metabolism within the fungal Kingdom.CYP51 belongs to the CYP super family and is the most conserved protein (Abastabar et al., 2019).

© Central University of Technology, Free State

These enzymes have strong specificity and bind to specific substrates such as lanosterol, obtusifoliol, and 24,25- dihydrolanosterol.CYP51 perform the catalytic reactions in three steps; in the first two reactions 14 alpha methyl is converted to methyl alcohol and methyl aldehyde. In the final step, the aldehyde group is transformed into formic acid with the formation of delta 14, 15-double bond. The 14-alpha demethylase (CYP51) of cytochrome p450s is the only invariant that is present in all steroid biosynthetic pathways. Although CYP51 is widely distributed across the fungal kingdom and with high conservation from species to species, differences still exist in types and subtypes in terms of phylogeny. For example, there could be only one CYP51 gene in Candida albicans while there could be two or more such genes in Ascomycota group. Interestingly, Aspergillus contains two CYP51 proteins, A and B (Alacazar-Fuoli et al., 2011). Certain species of Aspergillus could also contain a third variant of the CYP51 protein which is paralogous (extra copy) to either protein A or B. Studies on *Aspergillus fumigatus* suggest that CYP51B protein remains constitutively expressed while the A variant has to be induced for its functional activity (Abastabar et al., 2019).

The evidence also suggest that neither the A or the B variant is essential for in-vitro growth and virulence and only the simultaneous inactivation of both genes is lethal. A recent study has suggested that CYP51 could also perform indirect functions because deletion of the gene reduces mycelial elongation and invasive growth that is essential for the development of virulence (Camps et al., 2012). It is also observed that the deletion of CYP51 gene results into defects of reactive oxygen elimination which also reduces the virulence of the fungus. Moreover, the ERG-11 gene-deficient fungus is more susceptible to phagocytosis which suggests that the protein expression could help to escape the immune responses in the host as is with biofilms (Desai et al., 2014).

© Central University of Technology, Free State

Therefore, azoles could have both direct and indirect action on CYP51 (ERG-11) (Chen et al., 2018).

First of all, by inhibiting the enzyme, it prevents the fluidity of the plasma membrane of the fungus that is essential for mycelial elongation and invasive growth to cause virulence as well as make it susceptible to macrophages due to the non-production of ERG 11. The CYP51 plays an important role in virulence, which could be attributed to the multiple copies of the same gene or its complimentary invariants, which serve the function of mycelial growth and invasion. Therefore, azoles need to develop strategies inhibiting the enzyme ERG-11. On the contrary, resistance to azoles might be depicted by amplification of CYP51 genes within the fungus or other strategies (Flower et al., 2012). It could be possible that deletion mutations (that were observed in this study) in CYP51 could result into overproduction of ERG-11. However, future studies should substantiate such assumptions through experimental studies that would assay CYP51 mRNA expression through RT-qPCR or the amount of ERG-11 expressed in azole resistance and azole sensitive Aspergillus through western blotting (camps et al., 2012).

The present FASTA analysis reflected that there were deletions in the CYP51 in comparison to the ABC drug transporter protein. Since ABC drug transporter is an important mediator of efflux pump for antimicrobials that enter the microbes, deletion mutations might convert CYP51 to overactive ABC drug transporters that effluxes the azole out of the fungal cell. Such assumptions need to be substantiated experimentally by assaying the amount of azoles that came out of the fungal cell. The CYP51 is located in the inner face of the endoplasmic reticulum as a membrane mono-spanning enzyme, while its N-terminus is an amphipathic helix that links the catalytic unit of 14-alpha demythylase with the lipid bilayer. The multiple sequence alignment of CYP51 proteins indicates that they exhibit an identity between 36.5% and 93.9% across people

© Central University of Technology, Free State

and fungal isolates. The evidence suggests that the amino acid residues that are present within the folding of chains play a crucial role in enzymatic function. Therefore, it could be contended that changes in the amino acid residues within the folded chains might contribute to azole sensitivity. On the other hand, further folding of the protein chains that contain the conserved amino acid sequences might contribute to azole resistance. On the other hand, the residues that formed the surface of the CYP51 proteins active site remained highly conserved especially the Y118, F126, G127, V130, and T311 in the beta helix, beta helical turn and helix-1 signature regions.

Amino acid sequences that are conserved were classified into six substrate recognition sites and three motifs. Amongst them, the most conserved motif is a heme binding domain that contained the heme axial Cys ligand and the E-R-R triad which is forced by the EXXR and PER motifs. These motifs contributed to the locking of the heme pocket that ensured stabilization of the core structure. On the other hand, the six putative targets that house the substrate recognition are used landmarks for fungal CYP51s. Previous studies indicate that the CYP51B complex with tetrazole-based inhibitor VT-1598 has an optimized hydrogen bond between the phenoxymethyl oxygen of the inhibitor and imidazole ring nitrogen of HIs374 of CYP51. Studies further suggest that binding of azoles with CYP51 truncated its structure which led to the inhibition of fluidity of the cell membrane of the fungal isolate as well as its virulence (Zhang et al., 2019). The widespread use of CYP51 targeting molecules have resulted in the production of drug-resistant isolates. It is suggested that most of the resistance mechanisms associated with azoles are related to structural and genomic changes in CYP51. On the other hand, the binding of transcription factors such as Pdr1 and Stb5 and insertion elements Aft1 into the promoter region of CYP51 could reduce its sensitivity to the azoles (Gsaller et al., 2016). It could be possible that deletion mutations in CYP51 could promote the binding of the transcription factors to the CYP51 promoter which causes

© Central University of Technology, Free State

reduced expression of CYP51 or changes the functional conformation of the ERG-11 (Hagiwara et al., 2016).

Therefore, changes in ERG-11 interaction with azoles could occur at the genomic level and at the protein level. Aspergillus has a unique ability to thrive inside mammalian hosts and external environment that is vital for treatment failure with azoles and for the emergence of azole-resistant species. Previous studies suggest that non-synonymous substitution of amino acids; transcription factors, tandem repeats, and Dap proteins are the major causes of azole resistance in Aspergillus. The substitutions in the amino acids due to single nucleotide polymorphisms or deletion mutations might result into truncated proteins which were unable to convert lanosterol into ergosterol. It might be also possible that the truncated proteins could behave as ABC transporter proteins that gained efflux function after such deletion mutations or substitutions. Indeed point mutations have been observed in lanosterol 14-alpha demethylase (CYP51) in G54, L98, G138, M220, and G448 (Albarrag et al., 2011).

These amino acid positions are considered as mutational hotspots. Interestingly, most of these amino acids are present in the folded chains of the protein which suggest that mutations could have led to changes in the three dimensional structure of the active site that is necessary for enzymatic conversion of lanosterol to ergosterol. Mutations in G54R/E/V and 138S led to cross-resistance with itraconazole and posaconazole. These findings further support that there is a conserved position within the enzyme for binding of ligands (azoles) which mediated their antifungal action. This is because mutations in the G448S were associated with voriconazole tolerance. When the substitution at G138s was reversed, the fungal isolate was susceptible to both itraconazole and posaconazole. The evidence further suggests that one mutation in the CYP51 could predispose the risk of different mutations and different azoles for the same fungal isolate. For example,

© Central University of Technology, Free State

heterologous expression of G54W significantly reduced susceptibility of Aspergillus isolates to itraconazole and posaconazole but had no effect o voriconazole. These findings suggest that the group of azoles should be changed if the fungal isolate are refractory to another group of azoles instead of considering treatment failure with azoles. In another study, the authors showed that mutations in the tandem repeat of the CYP51 promoter and enzyme leads to broad spectrum azole resistance. Hence, the broad spectrum azole resistance is a function of point mutations in the CYP51 gene while narrow spectrum azole resistance could be attributed to the CYP51 protein. The point mutations (deletions) in the tandem repeats of the promoters could have reduced the transcription efficiency of the CYP51 that reduces the titer of the functional enzyme (Zhang et al., 2019).

Studies have depicted that L98H could cause flexible changes in the BC loop and IH loop of *A. fumigatus* CYP51. As a result, the position of Tyr 107 and 121 side chains also changes. Such changes modify the ligand biding (azole binding) channels in CYP51A of Aspergillus which prevents the binding of azole to the active heme. The lack of binding of azole to the active heme ensures constitutive action of the CYP51 enzyme. The mutations are carried over in subsequent generations because they do not confer survival disadvantage to the referred fungus. On the contrary, Y121F substitution could disrupt the hydrogen bonding between tyrosine and heme center of CYP51A of *Aspergillus fumigatus* that results in the instability of the active center of the enzyme. The instability of the active center results in poor functionality of the CYP51A enzyme and reduces the conversion of lanosterol into ergosterol. Although most of the mechanisms associated with azole resistance in Aspergillus species is related to CYP51, there are other resistance mechanisms than CYP51 that confer azole resistance to the referred fungal isolates. For example, mutations in the efflux pump (ABC transporter protein) could also lead to azole

© Central University of Technology, Free State

resistance. These findings are aligned to the findings of the present study because it also showed that the CYP51 azole resistant variant has remarkable identity with the ABC transporter of the referred fungal isolates. Such findings suggest that the ABC transporter has gained more function by changes in point mutations or it might also be possible that the point mutations or deletions in the 14-alpha demethylase might have functionally converted into an efflux pump that is capable of ejecting out azoles from the fungal cell membrane(Alcazar-Fuoli, et al.,2011).

© Central University of Technology, Free State

References

Abastabar, M., Hosseini, T., Valadan, R., Lagzian, M., Haghani, I., and Aslani, N. (2019). Novel point mutations in CYP51a and CYP51b genes associated with itraconazole and posaconazole resistance in *Aspergillus clavatus* isolates. *Microb. Drug Resist.* doi: 10.1089/mdr.2018.0300

Albarrag, A. M., Anderson, M. J., Howard, S. J., Robson, G. D., Warn, P. A., and Sanglard, D. (2011). Interrogation of related clinical pan-azole-resistant *Aspergillus fumigatus* strains: G138C, Y431C, and G434C single nucleotide polymorphisms in CYP51A, upregulation of CYP51A, and integration and activation of transposon Atf1 in the CYP51A promoter. *Antimicrob. Agents Chemother.* 55, 5113–5121. doi: 10.1128/AAC.00517-11

Alcazar-Fuoli, L., Mellado, E., Cuenca-Estrella, M., and Sanglard, D. (2011). Probing the role of point mutations in the CYP51A gene from *Aspergillus fumigatus* in the model yeast *Saccharomyces cerevisiae*. *Med. Mycol.* 49, 276–284. doi: 10.3109/13693786.2010.512926

Bernhardt, A., Meyer, W., Rickerts, V., Aebischer, T., and Tintelnot, K. (2018). Identification of 14-alpha-lanosterol demethylase (CYP51) in *Scedosporium* species. *Antimicrob. Agents Chemother.* 62:e02599-17. doi: 10.1128/AAC.02599-17

Camps, S. M., Rijs, A. J., Klaassen, C. H., Meis, J. F., O'Gorman, C. M., and Dyer, P. S. (2012). Molecular epidemiology of *Aspergillus fumigatus* isolates harboring the TR34/L98H azole resistance mechanism. *J. Clin. Microbiol.* 50, 2674–2680

Chen, Y., Li, Z., Han, X., Tian, S., Zhao, J., and Chen, F. (2018). Elevated MIC values of imidazole drugs against *Aspergillus fumigatus* isolates with TR34/L98H/S297T/F495I mutation. *Antimicrob. Agents Chemother.* 62, e01549-17

© Central University of Technology, Free State

Desai, J. V., Mitchell, A. P., and Andes, D. R. (2014). Fungal biofilms, drug resistance, and recurrent infection. *Cold Spring Harb. Perspect. Med.* 4:a019729. doi: 10.1101/cshperspect.a019729

Flowers, S. A., Barker, K. S., Berkow, E. L., Toner, G., Chadwick, S. G., and Gygax, S. E. (2012). Gain-of-function mutations in UPC2 are a frequent cause of ERG11 upregulation in azole-resistant clinical isolates of *Candida albicans*. *Eukaryot. Cell* 11, 1289–1299.

Gsaller, F., Hortschansky, P., Furukawa, T., Carr, P. D., Rash, B., and Capilla, J. (2016). Sterol biosynthesis and azole tolerance is governed by the opposing actions of SrbA and the CCAAT binding complex. *PLoS Pathog.* 12:e1005775.

Hagiwara, D., Miura, D., Shimizu, K., Paul, S., Ohba, A., and Gonoi, T. (2017). A novel Zn2-Cys6 transcription factor AtrR plays a key role in an azole resistance mechanism of *Aspergillus fumigatus* by co-regulating CYP51A and cdr1B Expressions. *PLoS Pathog.* 13:e1006096. doi: 10.1371/journal.ppat.1006096

Zhang, J., Li, L., Lv, Q., Yan, L., Wang, Y., & Jiang, Y. (2019). The Fungal CYP51s: Their Functions, Structures, Related Drug Resistance, and Inhibitors. *Frontiers in microbiology*, *10*, 691. https://doi.org/10.3389/fmicb.2019.00691

© Central University of Technology, Free State

CHAPTER 7: DISCUSSION, CONCLUSION, & FUTURE DIRECTIONS

The growing prevalence of fungal diseases over the past few decades is a significant concern for the agriculturists as well as health care professionals. (Abastabar et al., 2019). Fungal diseases contribute to mortality rates across the globe per year, which is more than that caused by malaria. On the other hand, the agricultural yield and food security is significantly threatened due to the invasion of fungal species. The azole group of antifungals has remained the drug of choice for combating fungal diseases in humans as well as in the field of agriculture. The menace of fungal diseases has been expedited by the increased resistance of pathogenic fungus to the azoles(Zhang et al., 2019). Azole resistance is mediated by various mechanisms out of which point mutations in the target enzyme cytochrome p450 monooxygenase of the variant CYP51 is most common (Parker et al., 2014). The mutations in the referred enzyme are attributed to the selection pressure caused by the widespread use of triazole fungicides in agriculture as well as in health care settings.

The mutations similar to that observed in clinical isolates of the referred species during the 1990s in Candida albicans and *Aspergillus fumigatus* were also isolated in the field of agriculture (Lepesheva et al., 2007). However, the number and the location of the point mutations have significantly increased and changed respectively during the past decade. Interestingly, the present trend reflects that mutations to triazole antifungals emerge from the agricultural field and appear as pathogenic and resistant fungi in the clinical infections. It is speculated that the prevalence of azole resistance fungi and diseases significantly increase over the next two decades. This is because there would be a significant increase in the use of triazole antibiotics both in agriculture as well as in clinical settings. Although most fungi are saprophytic in nature, the number of fungal isolates that cause diseases in plants, animals, and humans is quite high. Some fungal pathogens are acknowledged as primary pathogens because they can cause disease in healthy and

© Central University of Technology, Free State

uncompromised hosts (Zhang et al., 2019),. The second group of fungal pathogens that cause diseases in immunocompromised hosts contribute to the opportunistic infections due to the health status of the patients is termed as secondary pathogens. As life-threatening bacterial infections have shown a steady increase in the turnover of patients in the Intensive Care units (ICUs), the use of antibacterial has parallel increased. The increased use of antibiotics and the simultaneous resistance of bacterial species against them predispose the risk of fungal superinfections. On the other hand, the decreased immune status of the patients coupled to the use of high dose steroids for combating clinical emergencies is another cause of fungal superinfections. However, the risk of fungal superinfections varies according to the type of antibiotic used and the immune status of the patient. For example, cefoperazone (a third generation cephalosporin) that is used for managing septicemia and life-threatening infections has one of the highest probabilities to increase the risk of a Candida super infection. As a result, the use of triazole antifungals would be on the rise to manage opportunistic fungal infections in immunocompromised patients.

The evidence suggests that 10% to 15% of nosocomial infections (hospital-acquired infections) are caused by opportunistic fungal pathogens. On the other hand, approximately, 300 million individuals across the globe contract at least one life-threatening fungal infection per year. The infection could range from minor episodes of vulvovaginal rash to chronic pulmonary infections such as aspergillosis (Sobel, 2007). Globally, the mortality rates due to fungal pathogens is estimated to be 1.35 million per annum compared to malaria which has a mortality rate of 0.63 million. The annual incidences of Candidaemia are estimated to be 300,000 globally and the mortality rates fall within 30% to 55%. Likewise, more than 10 million individuals are at the risk of developing invasive aspergillosis with the increased use of corticosteroid anticancer therapies. Studies suggest that almost 200,000 individuals across the globe develop invasive aspergillosis

© Central University of Technology, Free State

with the mortality rate of 50% even with triazole therapy. On the other hand, more than 4 million asthmatics across the globe develops bronchopulmonary aspergillosis, while another 3 to 13 million of the target population suffer from severe asthma due to fungal sensitization (Denning et al., 2013). Moreover, chronic fungal pulmonary aspergillosis could affect more than three million people globally per annum out of which 1.2 million contract the disease following tuberculosis. Different cryptococcal species account for meningitis in HIV patients that accounts for 60% of mortality in the target population. Cryptococcus species affect around 100,000 to 300,000 individuals per year in the United States. Filamentous fungi such as Aspergillus and Fusarium species also account for eye infections in visually impaired individuals. Likewise, the fungal diseases that are mainly observed in plants include brown rust in wheat, leaf scald in cereals, and apple scab. Most of the diseases are caused by Aspergillus and Fusarium species as they contain myotoxins (aflatoxins). The fungal diseases should be controlled in plants to ensure food security and crop field. For example, a study showed that there would be a fall in European wheat production by 7% to 12% if azole fungicides and not used. The decrease productivity in crop field would contribute to a financial loss of 2.4 billion Euros to 4.6 billion Euros by 2020 (Parker et al., 2014).

These findings suggested that most of the diseases caused by pathogenic fungus belong to Aspergillus, Candida, and Cryptococcus species. Hence, it is not surprising why there is a need for the increased use of triazole antifungals in healthcare settings as well as in agriculture. The azole antifungals work by inhibiting the synthesis of ergosterol in fungi. They could directly bind to the sterol 14-alpha demethylase which is a CYP51 protein. CYP51 is an important isomer of the super family of heme-containing enzymes; the cytochrome p450.The enzyme is located on the outer membrane of the endoplasmic reticulum and catalyzes the removal of methyl group at the C14

© Central University of Technology, Free State

position. The lone pair of electrons present in the nitrogen atom of the azole ring binds to the heme prosthetic group of CYP51. The prosthetic group is located within the active site of the enzyme, while the N1 substituent group interacts with the amino acids that are present in the lining of the active site pocket. The specificity of the azole compounds are determined by the interaction between the side groups of the azole compound and CYP51 (Isla et al., 2018). The binding of azole antifungals to the enzyme is non-competitive, which results in the depletion of the final fungal sterols followed by a concomitant accumulation of 14-methylated sterols. The accumulation of14-methylated sterols inhibit fungal growth by disrupting the structure of the cell membrane (Kushima et al., 2017)..

In other fungal species such as Candida albicans, other types of sterols accumulate that result into similar destruction of the cell membrane. The antifungal mechanism of action of azole compounds was first reported in 1944 in the form of imidazole (containing two N atoms within the azole ring) followed by the triazoles (having three N in the azole rings). In 1958, chlormidazole was the first antifungal compound to get an approval for topical use. From 1969 onwards, various imidazole-based antifungals have been recommended by the USFDA for topical applications. Ketoconazole was the first azole antifungal that was recommended for systemic use (Parker et al., 2014). However, the compound is contraindicated across various patient populations owing to its toxic profile. Fluconazole, the first systemic azole compound, obtained FDA approval for systemic use in the 1990s. In the late 1990s and early 2000s, the antifungal market witnesses the introduction of second generation triazoles such as voriconazole and posaconazole. The second generation triazol antifungals are primarily used to manage fluconazole-resistant strains. Moreover, these antifungals are also recommended for managing infections caused by *Aspergillus fumigatus*. However, there has been a significant increase in resistance towards triazole antibiotics over the

© Central University of Technology, Free State

last 10 years. It is contended that the progressive introduction of new generation of triazole antifungals which are intrinsically more active would help in the maintenance of ED50 values. Hence, there is a need to develop triazole antifungals that would have a low resistance potential or those that would be capable of overcoming the resistance caused by pathogenic fungi (Parker et al., 2014).

The emergence of azole resistance in fungal species could be attributed to the prophylactic use of azoles for preventing the risk of fungal super infections across healthcare settings. On the other hand, prolonged treatment regimes of antifungals in healthcare settings and agriculture have further increased the emergence of azole-resistance fungal species. Resistance to azole compounds could stem from phone mechanism. First of all, the affinity for the target enzyme CYP51 might be reduced by point mutations. Secondly, the amount of the target enzyme might be increased due to the up regulation of the gene followed by the efflux of the azole compounds due to the over expression of the efflux pump transporter proteins such as the ABC multidrug transporters (Flowers et al, 2015). On the other hand, secondary mutations in the target enzyme could further increase the probability of emergence of resistance to antifungals. For example, an ERG3 null mutant in Candida spp. was associated with resistance to azoles. The lack of ERG3 activity resulted in the production of 14-methyl fecosterol when treated is capable of preserving the functions of the plasma membrane in the referred fungal species. The complex resistance mechanisms witnessed in fungal species to the azole compounds is not fully understood. Although various studies have explored the resistance mechanisms to azoles in Candida species, only a few studies have explored resistance mechanism to azoles in Eurotiomycete species (for example, in Aspergillus).

© Central University of Technology, Free State

Flowers et al. (2015) showed that 55 out of the 63 isolates exhibited at least one single nucleotide polymorphism (SNP) within their ERG11 allele. However, nine amino acid substitutions were identified as novel and were specific to the Candida species only. These findings suggested that the resistance mechanism to azoles by pathogenic Aspergillus remain less understood. The resistant nature of ERG11 isozymes against fluconazole stemmed from the mutations that were clustered primarily in the active site of the enzyme. Likewise, resistance to fluconazole was also attributed to the presence of point mutations in the fungal-specific external loop as well as the proximal surface that potentially with the fungal loop or near the heme moiety of the enzyme. These findings suggested that mutational hotpsots in the fungal specific external loop and active site of the enzymes could attribute to the resistance mechanisms in azoles. The concept that point mutations are the major modes in conferring resistance to the azole compounds was based on the observation that amino acid substitutions in the A114S, Y132H, Y132F, and K143R resulted into fourfold resistance of the referred fungal isolates to fluconazole and voriconazole.

Hence, previous studies have reflected that base substituted mutations are the primary resistance mechanism in *Eurotiomycete* species for conferring resistance to antifungal species(Parker et al., 2014).Studies suggest that the number of novel mutations in CYP51 or ERG3 is a hallmark of *Eurotiomycete* species in conferring resistance to the azole antifungals. However, the frequency of novel mutations also makes it difficult to design novel antifungal compounds that would have the ability to manage azole-resistant *Eurotiomycete* species or those in which the probability of such resistance is high in the near future. Some resistant strains of fungal species tend to develop a plethora of adaptations to more than one resistance mechanisms(Parker et al., 2014). These findings are aligned with the resistance mechanisms and approaches observed in other microbes such as bacteria. Bacteria have various resistance mechanisms to combat the action of antibiotics.

© Central University of Technology, Free State

These mechanisms include production of beta-lactamses, an over activity of the efflux pumps, and mutation of the porin channels. Studies suggest that an antibiotic which was able to exhibit bactericidal action in resistant strains of bacteria becomes ineffective in the future in the same bacterial species. Hence, one cannot be assured that preventing the emergence of resistance through one mechanism in a fungal isolate would help to eradicate the referred species. This is because the referred species could overcome antifungal action by employing different resistance mechanisms. On the other hand, the mutations that affect the binding of azoles to the CYP51 enzymes are commonly found in azole-resistant fungal strains (Zhang et al., 2019).

Mutational hotspots in the referred enzyme have been extensively studied and exhibit a high correlation with mutational hotspots. Understanding the effects such mutations have on the interaction between CYP51 and azoles and the subsequent inhibition of CYP51 activity is necessary for designing novel antifungal drugs that would provide an answer to azole resistance for the present as well as for the future. On the contrary, the probability of adaptations and the frequency of emergence of new resistant mechanisms impose significant challenges in scientists for developing novel antifungal compounds based on synthetic modifications and in-vivo efficacy related to the antifungal compounds. Moreover, it is also time consuming to design and obtain approval for antifungal compounds that are likely to become resistant in the future. In this regard, in silico approaches based on computational chemistry and bioinformatics hold the promise for predicting resistance mechanisms in fungal species. The present study discussed the potential of in silico approaches in identifying and predicting novel resistance mechanisms in *Eurotiomycete* species of fungi, especially in Aspergillus (Zhang et al., 2019). Previously, there were difficulties in undertaking in silico approaches for identifying the resistance mechanisms to azoles in Eurotiomycete species due to the lack of molecular characterization of the CYP51 species in

© Central University of Technology, Free State

them. However, recent literature has shed adequate light on the molecular structure and mutational hotspots in CYP51 gene of the referred species. However, there is no study till date that has predicted the emergence of resistance to the azole antifungals based on in silico approaches. This is the first study that depicted the probable mutational hotpsots that could cause emergence in azole resistance in Eurotiomycete species including Aspergillus. The azole resistance group of CYP51 in Aspergillus was evaluated through BLAST and FASTA.

The BLAST format depicted the single nucleotide polymorphisms that have the highest probability through point mutations (including base substitutions). Likewise, the FASTA approach helped to identify the similarity of the azole resistance CYP51 with other proteins, apart from CYP51. The FASTA showed that the CYP51 molecules had highest sequence identity with the ABC transporter proteins which further suggested that emergence of resistance through one mechanism could predispose the resistance through other mechanisms. The CYP51 in Aspergillus species could also indicate that they mimic the natural ABC transporter proteins (that are not mutated). Such observations indicate that CYP51 could act as 14-alpha demethylase as well as ABC transporter protein. It is necessary to identify novel azole compounds that could exhibit effective antifungal action against Aspergillus species. Aspergillus infections are problem not only in immunocompromised patients but also in immunocompetent individuals. Aspergillus is primarily responsible for chronic bronchopulmonary Aspergillosis. Therefore, azole resistance imposes significant challenges for managing patients of cystic fibrosis.

The mortality rate amongst patients with multiazole-resistant invasive Aspergillus are estimated to be 88%.Such issues impose burden on the resources in healthcare settings because one has to rely more on amphotericin B and caspofungin for alleviating fungal infections (Zhang et al., 2019). There has been a significant increase in the number of Aspergillus isolates that showed resistance

© Central University of Technology, Free State

towards azole antifungals since 1998. The global ARTEMIS and SCARE surveillance program depicted that 5.8% and 3.6% of the clinical isolates of Aspergillus exhibited elevated MICs to more than one triazole compound in the United States and Europe, respectively. On the other hand, 10% of all clinical isolates of Aspergillus is resistant to itraconazole where 23% to 31% of the isolates of Aspergillus were resistant to azoles.

7.1. Conclusion

The present study showed that azole resistance in CYP51A and CYP51B of Aspergillus species occur at the gene and protein level. The gene changes include singe nucleotide polymorphisms and deletions while point substitutions are noted in amino acid sequences in the CYP51. The changes in the three dimensional conformation of the active site of the 14-alpha demethylase (CYP51A) protein impair the formation of ergosterol from lanosterol. As a result, there is marked decrease in membrane fluidity and mycelial growth that is essential for virulence of the fungal isolates. This study also showed that deletions in the tandem repeats of the promoter of CYP51 could have resulted into reduced expression of the14-alpha demethylase. Another possibility of azole resistance is that single nucleotide polymorphisms in the tandem repeats of the promoter might have increased its sensitivity to the transcription factors that causes down regulation of expression of CYP51A which again results into reduced titer of 14-alpha demethylase. The single nucleotide polymorphisms in the CYP51 could have resulted into folding of the heme active sites that is necessary to for binding with the ligands (azole). Finally, this is the first study that showed that mutations in CYP51 and azole resistance could be a alteration in the function of 14-alpha demethylase as an overactive ABC transporter. The sequence identity of the CYP51 protein and the ABC transporter coupled with previous evidence that azole resistance could occur through non-CYP51A mechanisms substantiated such assumptions.

© Central University of Technology, Free State

7.2. Strength and Limitations

The major strength of this study was that it implemented all the possible in silico methods to elucidate the structure-function correlates of azole resistance in CYP51 of Eurotiomycetes**.** The in silico analysis was not only limited to the polymorphisms in the CYP51 gene but was extrapolated to the structural anomalies in the CYP51 protein. However, the major limitation of this study was that it did not evaluate the in silico analysis on the light of protein expression or protein structure.

7.3. Future Directions

Future studies should explore altered mRNA expression through RT-qPCR to evaluate the point mutations in CYP51. The future studies should also assay the amount of ergosterol produced by azole-resistance and azole-sensitive Aspergillus. Such analysis would help to identify whether azole resistance primarily is a function of single nucleotide polymorphisms or point mutations in the amino acid sequence of CYP51. To substantiate the findings further, X-ray crystallographic structures of azole-resistance and azole-sensitive CYP51 of Aspergillus should be studied.

© Central University of Technology, Free State

References

Abastabar M., Hosseini T., Valadan R., Lagzian M., Haghani I., Aslani N. (2019). Novel point mutations in CYP51a and CYP51b genes associated with itraconazole and posaconazole resistance in *Aspergillus clavatus* isolates. *Microb. Drug Resist.* 10.1089/mdr.2018.0300

Denning DW, Pleuvry A, Cole DC (2013). Global burden of allergic bronchopulmonary aspergillosis with asthma and its complication chronic pulmonary aspergillosis in adults. Med Mycol. 51(4), 361–370

Isla, G., Leonardelli, F., Tiraboschi, I. N., Refojo, N., Hevia, A., and Vivot, W. (2018). First clinical isolation of an azole-resistant *Aspergillus fumigatus* isolate harboring a TR46 Y121F T289A mutation in south america. *Antimicrob. Agents Chemother.* 62, e00872-18. doi: 10.1128/AAC.00872-18

Kushima, H., Tokimatsu, I., Ishii, H., Kawano, R., Watanabe, K., and Kadota, J. I. (2017). A new amino acid substitution at G150S in Lanosterol 14-alpha demethylase (Erg11 protein) in Multi-azole-resistant Trichosporon asahii. *Med. Mycol. J.* 58, E23–E28. doi: 10.3314/mmj.16-00027

Lepesheva, G. I., and Waterman, M. R. (2007). Sterol 14alpha-demethylase cytochrome P450 (CYP51), a P450 in all biological kingdoms. *Biochim. Biophys. Acta* 1770, 467–477

Parker, J. E., Warrilow, A. G., Price, C. L., Mullins, J. G., Kelly, D. E., & Kelly, S. L. (2014). Resistance to antifungals that target CYP51. *Journal of chemical biology*, *7*(4), 143–161. https://doi.org/10.1007/s12154-014-0121-1

Sobel JD (2007). Vulvovaginal candidosis. Lancet. 369(9577),1961–1971.

© Central University of Technology, Free State

Zhang, J., Li, L., Lv, Q., Yan, L., Wang, Y., & Jiang, Y. (2019). The Fungal CYP51s: Their Functions, Structures, Related Drug Resistance, and Inhibitors. *Frontiers in microbiology*, *10*, 691. https://doi.org/10.3389/fmicb.2019.00691

© Central University of Technology, Free State