Inhibition of Protein Prenylation: a Novel Approach in Antifungal Therapy

by

Quincy Tseng

Department of Biochemistry
Duke University

Date:_______________________

Approved:

___________________________
Lorena S. Beese, Supervisor

___________________________
Paul L. Modrich

___________________________
Tao-shih Hsieh

___________________________
Leonard D. Spicer

___________________________
Matthew Redinbo

Thesis submitted in partial fulfillment of
the requirements for the degree of Master of Science in the Department of
Biochemistry in the Graduate School
of Duke University

2011
Inhibition of Protein Prenylation: a Novel Approach in Antifungal Therapy

by

Quincy Tseng

Department of Biochemistry
Duke University

Date:_______________________
Approved:

___________________________
Lorena S. Beese, Supervisor

___________________________
Paul L. Modrich

___________________________
Tao-shih Hsieh

___________________________
Leonard D. Spicer

___________________________
Matthew Redinbo

An abstract of a thesis submitted in partial fulfillment of the requirements for the degree of Master of Science in the Department of Biochemistry in the Graduate School of Duke University

2011
Abstract

Many proteins that are required for eukaryotic cell growth, differentiation, and cellular signaling pathways, including the Ras superfamily of GTP-binding proteins, go through prenylation, a post-translational modification of isoprenoid lipid attachment. Since prenylation of mutant Ras proteins is necessary for their oncogenic effects, protein prenyltransferases have been targets for cancer therapeutics. These enzymes have also been identified and shown to be essential in several fungal and protozoan organisms that exist as opportunistic human pathogens. Some of the mammalian protein prenyltransferase inhibitors, used as chemotherapeutic agents in clinical trials, exhibit antifungal properties and have the potential to be adapted for antifungal therapies. The prenylation pathway is therefore an attractive target for treating fungal infections.

The focus of this project is on the fungal organism *Aspergillus fumigatus*, a major opportunistic human pathogen of invasive aspergillosis that is becoming more common among transplant recipients and cancer and AIDS patients. To approach inhibitor design for *A. fumigatus* protein prenyltransferases, I have described an X-ray crystal structure of the *A. fumigatus* protein farnesyltransferase (FTase) in complex with farnesyl diphosphate and its active site features that include isoprenoid and peptide binding conformations, zinc coordination, and product exit groove. Additionally, an *in vitro* fluorescence-based activity assay was used to partially characterize the peptide substrate.
specificity of the enzyme. I have also obtained crystal structures of the *A. fumigatus* FTase in complex with two ethylenediamine-scaffold compounds and evaluated the binding mode and selectivity of the compounds relative to that in mammalian FTase. Given the limited effectiveness of current antifungal therapies, the work presented here may lead to the design of new fungal inhibitors with improved selectivity.
Table of Contents

Abstract ......................................................................................................................................... iv

List of Tables.................................................................................................................................. x

List of Figures ............................................................................................................................... xi

Chapter 1. Virulence, gene essentiality, and mechanisms of drug resistance in human pathogenic fungi .......................................................... 1

Introduction ............................................................................................................................. 1

Virulence factors of fungal pathogens .................................................................................... 3

Genes related to Aspergillus fumigatus growth and virulence ........................................... 6

Fungal cell wall ....................................................................................................................... 8

Antifungal agents: targets and mechanisms of action ......................................................... 12

Mechanisms of antifungal resistance .................................................................................. 16

Fungal efflux pumps: transcriptional control and drug resistance ................................. 18

Topology of fungal efflux pumps ...................................................................................... 20

Substrate selectivity of efflux transporters ....................................................................... 23

Summary ................................................................................................................................ 25

Chapter 2. Protein prenylation .............................................................................................. 27

Introduction ........................................................................................................................... 27

Roles of prenylation targets ............................................................................................... 29

Essentiality of protein prenylation ..................................................................................... 31

Mutant phenotypes of A. fumigatus prenylated gene products ........................................ 32

Overview of mammalian protein prenyltransferases ....................................................... 33
Chapter 4. Towards inhibition of *Aspergillus fumigatus* protein farnesyltransferase .......................... 83

Introduction........................................................................................................................... 83

Structure of *Aspergillus fumigatus* FTase inhibitor complex ........................................... 86

Inhibitor binding mode ..................................................................................................... 87

FTase inhibitor selectivity ................................................................................................. 90

Summary and future directions ......................................................................................... 92

Materials and methods........................................................................................................ 94

Crystallization..................................................................................................................... 94

Data collection and refinement......................................................................................... 95

FTase inhibitor specificity assay....................................................................................... 95

Chapter 5. Crystallization and preliminary X-ray diffraction analysis of human mismatch repair protein MutSβ ...................................................................................................... 96

Introduction........................................................................................................................... 96

Proteolytic fragment of human MSH3 ............................................................................ 100

DNA binding studies of hMutSβ ..................................................................................... 100

Crystallization of hMutSβ ................................................................................................. 102

Summary.............................................................................................................................. 103

Materials and methods...................................................................................................... 104

Cloning, expression, and purification of hMutSβ........................................................ 104

Limited proteolysis of hMutSβ....................................................................................... 105

DNA binding studies of hMutSβ ................................................................................... 106
List of Tables

Table 3-1: Data collection and refinement statistics for the binary complex ...................... 46
Table 3-2: X residue selectivity of rat, C. neoformans, and A. fumigatus FTases .................. 65
Table 3-3: Potential A. fumigatus Ca1a2X prenyltransferase substrates ................................ 69
Table 4-1: Data collection and refinement statistics ............................................................... 85
List of Figures

Figure 1-1: Essential genes in *A. fumigatus*, *S. cerevisiae*, and *C. albicans*. .............................. 7

Figure 1-2: Cell wall polysaccharides of *A. fumigatus* and *S. cerevisiae*. ................................. 9

Figure 1-3: Structures of antifungal agents in clinical use.......................................................... 13

Figure 1-4: Antifungal drug targets in the ergosterol biosynthetic pathway. ............................. 14

Figure 1-5: Timeline of antifungal development........................................................................ 15

Figure 1-6: Mechanisms of action of selected antifungals......................................................... 17

Figure 1-7: Fungal survival mechanism as a response to azole inhibition............................... 18

Figure 1-8: Predicted topological features of a fungal PDR transporter.................................. 22

Figure 2-1: Prenylation and modification of proteins................................................................. 28

Figure 2-2: Signaling pathways in mammalian cells involving Ras family members. .............. 29

Figure 2-3: Reaction pathway of protein farnesylation............................................................. 34

Figure 2-4: Topology diagram of rat FTase............................................................................... 35

Figure 2-5: Zinc coordination in human FTase ternary complex............................................. 36

Figure 2-6: Residues of rat FTase that interact with the CaaX motif...................................... 38

Figure 2-7: Exit groove residues of rat FTase.......................................................................... 40

Figure 2-8: Superposition of CaGGTase-I binary and rat GGTase-I ternary complexes. . 41

Figure 2-9: Superposition of CnFTase complexes. ................................................................. 42

Figure 3-1: Structure of *A. fumigatus* FTase........................................................................ 48

Figure 3-2: Superposition of *A. fumigatus* and rat FTases.................................................. 48

Figure 3-3: AfFTase isoprenoid binding site............................................................................ 50
Figure 3-4: Helix 4α-5α loop conformations in AfFTase, rFTase, and CnFTase. ...............52
Figure 3-5: AfFTase, rFTase, and CnFTase CaaX binding pockets. ..............................55
Figure 3-6: AfFTase, rFTase, and CnFTase product exit grooves.................................58
Figure 3-7: AfFTase and rFTase X residue selectivity assay ........................................62
Figure 3-8: X residue selectivity of AfFTase and rFTase ...............................................64
Figure 3-9: CVVX peptide modeled into AfFTase CaaX binding pocket .......................68
Figure 3-10: 11-mer peptides modeled into AfFTase peptide binding pocket ...............76
Figure 4-1: Chemical structures of ethylenediamine-scaffold compounds ..................85
Figure 4-2: Inhibitor binding modes of CCB2217 and SF122 ........................................88
Figure 4-3: SF102 and SF122 modeled in the AfFTase active site .................................89
Figure 4-4: Selectivity assay with titration of CCB2217 and SF122 ...............................91
Figure 5-1: Purification of MutSβ and deletion mutants ...............................................99
Figure 5-2: SDS-PAGE of human MutSβ limited proteolysis by subtilisin ......................99
Figure 5-3: Binding of MutSβ and MutSβΔ162 to a 41-bp heteroduplex .......................101
Figure 5-4: Human MutSβ crystals ..............................................................................102
Figure 5-5: Diffraction pattern of human MutSβ crystal form 2 ....................................109
Chapter 1. Virulence, gene essentiality, and mechanisms of drug resistance in human pathogenic fungi

Introduction

Immunological impairment has been a byproduct of medical treatments and trends in the last few decades, whether due to organ transplantation, cancer chemotherapy regimens, or even the advent of AIDS epidemic. Despite the relatively few numbers of human pathogenic fungi, aggressive use of immunosuppressive and antimicrobial agents has nonetheless led to an increase in opportunistic fungal infections that are most commonly attributed to the ubiquitous *Candida*, *Cryptococcus*, and *Aspergillus* species. Even with antifungal treatment, mortality rate among infected individuals is high, especially when the infection is of an invasive nature in which the fungus enters the bloodstream or the central nervous system. Given the high incidence of mycoses, it is clear that fungal pathogens are becoming a major threat to human health.

Non-*albicans* *Candida*, dematiaceous molds, and other *Aspergillus* species have been increasingly recognized as fungal pathogens, which exist mostly as yeasts and hyaline molds. *Candida* remains the most common cause of opportunistic mycoses with a mortality of more than 30% (Pfaller & Diekema, 2007; Gudlaugsson et al., 2003), and despite an increasing trend in the isolation of *C. glabrata*, *C. tropicalis*, *C. parapsilosis*, and *C. krusei*, *Candida albicans* still makes up the majority of *Candida* clinical isolates (Pfaller &
Diekema, 2007). Like *C. albicans*, *Cryptococcus neoformans* appears as spherical yeast cells in infected hosts and is the most common cause of central nervous system mycosis in AIDS patients (Walsh & Groll, 1999). In regions where highly active antiretroviral therapy is unavailable, mortality due to cryptococcosis can be up to 30% (Lin & Heitman, 2006). *Aspergillus fumigatus* is the most frequently isolated air-borne pathogenic mold that, like *C. neoformans*, may be inhaled by humans and trigger pulmonary infections. Most cases of aspergillosis are associated with leukemia, hematopoietic stem cell and solid organ transplantations (McCormick et al., 2010) and have an overall mortality ranging from 45 to 90% (Lin et al., 2001).

An estimated 2.6 million people became HIV-infected in 2009 and contributed to a worldwide total of about 33 million infected individuals that year (UNAIDS report, 2010). In 2008 alone, more than 27,000 solid organ transplant procedures were done in the United States (Wolfe et al., 2010). With an expanding population of immunocompromised hosts that is likely to remain high, the emergence of opportunistic fungal pathogens will continue to be affiliated with significant mortality and health care costs. This medical problem is further compounded by limited therapeutic options and developing antifungal resistance.

Challenges posed by increasing prevalence of invasive fungal infections therefore warrant the attention of scientists and clinicians in order to better deal with these microorganisms. The available whole-genomic sequence of pathogenic fungi has
complemented our understanding of their pathogenesis, virulence factors, and metabolism, as well as the identification of potential drug targets. A comprehensive study of fungal biochemistry will provide further insight into mechanisms of antifungal resistance and, ultimately, diagnostic and therapeutic advances.

Many current antifungals target essential gene products involved in pathways such as cell wall and ergosterol biosyntheses that are unique to fungi. An ideal first step in antifungal development is the identification of essential genes required for the growth of fungal pathogens, but given the rise in drug resistance, minimizing fungal virulence may be an alternative approach (Gauwerky et al., 2009). This chapter focuses on the *Candida*, *Cryptococcus*, and *Aspergillus* species and presents an overview of fungal biology and mechanisms in the context of drug therapy and development. In addition to fungal virulence, gene essentiality, and cell wall, antifungal agents and efflux-mediated drug resistance are also discussed here.

**Virulence factors of fungal pathogens**

Since transitioning from its ecological niche to a mammalian host can be extremely harsh for a fungus, environmental adaptation is essential to the survival of fungal pathogens. Having to adjust to fluctuating parameters in the forms of host immune system and abiotic stresses, pathogenic fungi possess virulence factors that contribute to their pathogenicity and adaptation to hostile environments. The ability to grow at physiological temperatures of 37°C and above is an important requirement for
virulence and is what separates the pathogenic species from their nonpathogenic counterparts (Cooney & Klein, 2008; van Burik & Magee, 2001). In *C. neoformans*, the calcium-dependent protein phosphatase calcineurin (encoded by *CNA1*) is found to be necessary for growth at 37°C (Odom et al., 1997).

The adherence to host cells and tissues for infection is important for many fungal pathogens. Adhesins, such as Hwp1p and the Als protein family, are present on the *C. albicans* cell wall surface and facilitates attachment to epithelial and endothelial cells (van Burik & Magee, 2001; Karkowska-Kuleta et al., 2009). While not as well studied, receptors on *C. neoformans* also allow the fungus to bind to human endothelial cells and cross the blood-brain barrier to cause meningitis (Karkowska-Kuleta et al., 2009).

Fungal pathogens often produce and secrete hydrolytic enzymes in order to degrade tissues for nutrients and dissemination with the host. *C. albicans* is known to use both secreted aspartyl proteinases and phospholipases to degrade host proteins and membranes. Similar phospholipases exist in *C. neoformans* and are responsible for membrane hydrolysis, cell lysis, and adhesion of the fungus to lung epithelium. *A. fumigatus* produces serine proteinase and metalloproteinase, both of which are active on elastin, a major component of lung tissue, during pulmonary infections (Karkowska-Kuleta et al., 2009).

Most fungal pathogens exhibit morphological dimorphism, which responds to environmental stimuli and plays a major role in pathogenicity. *C. albicans* has the ability
to switch from yeast cells to filamentous hyphal form, a transition triggered by changes in temperature, pH, or carbon dioxide concentration (Karkowska-Kuleta et al., 2009). Hyphae are more often associated with virulence, but the molecular basis for morphological change is not well understood. Similar to *C. albicans*, *A. fumigatus* has a virulent filamentous form but switches to conidial form for dissemination (van Burik & Magee, 2001). While *C. neoformans* can exist as chlamydospores and hyphae, it adopts a yeast form for infection (Lin & Heitman, 2006).

Other virulence factors pertain only to specific species, but perhaps this is just a reflection of a discrepancy in our level of understanding. The ability of *C. albicans* to tolerate different pH levels, ranging from acidic mucosal surfaces to slightly alkaline blood (Cooney & Klein, 2008), is regulated by *PHR1, PHR2*, and *RIM101*. The presence of melanin pigment in the cell wall of *C. neoformans* and *A. fumigatus* conidia provides protection against ultraviolet radiation, extreme temperatures, reactive oxygen species, and even antifungal agents. During infection, *C. neoformans* forms a thick polysaccharide capsule composed mostly of glucuronoxylomannan. Due to its negative charge, the capsule can prevent phagocytosis, as well as inhibit antigen presentation and leukocyte accumulation (van Burik & Magee, 2001; Karkowska-Kuleta et al., 2009). Many fungi employ siderophores to utilize host iron, an essential element for fungal growth and metabolism that can affect capsule synthesis in *C. neoformans* (Vartivarian et al., 1993). *A. fumigatus* is known to produce the secondary metabolite gliotoxin, which
can inhibit macrophage phagocytosis and T-cell activation and proliferation (Karkowska-Kuleta et al., 2009). Interestingly, this small, nonpolar compound is a dual inhibitor of mammalian protein farnesyltransferase and geranylgeranyltransferase I (Vigushin et al., 2004).

**Genes related to Aspergillus fumigatus growth and virulence**

The availability of whole-genomic sequence of pathogenic fungi (Jones et al., 2004; Loftus et al., 2005; Nierman et al., 2005) has accelerated the elucidation of genes essential for fungal growth, survival, and pathogenicity. The identification of these genes therefore constitutes an approach to uncovering and characterizing potential drug targets. Out of the 54 *A. fumigatus* genes which have orthologs that are essential for growth in *S. cerevisiae* and *C. albicans*, 35 of them have been identified to be essential for growth. These genes (Fig. 1-1) are involved in biological processes such as lipid, ergosterol, cell wall, amino acid, protein, and heme biosyntheses, as well as glycosylation, secretion, and RNA processing (Hu et al., 2007).

Genes related to *A. fumigatus* virulence are involved in processes such as thermotolerance, cell wall composition, resistance to immune response, toxins, nutrient uptake, signaling, and metabolic regulation. The *thtA, afpmt1, hsp1/asp f 12*, and *cgrA* genes enable growth at 37 °C and beyond, but the effect on gene expression in response to temperature fluctuation is not well understood. Genes responsible for cell wall composition have also been identified: α-1,3-glucan (*ags1-3*), β-1,3-glucan (*fks1* and *rho1-*...
Figure 1-1: Essential genes in *A. fumigatus*, *S. cerevisiae*, and *C. albicans*, as defined by a 3+ or greater growth phenotype. FKS1 and the ERG11 family, which will be discussed in more detail elsewhere in this chapter, are both essential for *A. fumigatus* growth. Reprinted from Hu et al., 2007.
4) and its elongation (gel1-7), chitin (chsA-G), and galactomannan (och1-4, mnn9, van1, anp1, and glfA). Surface components associated with resistance to immune response include hydrophobins (rodA/hyp1 and rodB), melanin (pksp/alb1, ayg1, arp1, arp2, abr1, and abr2), and efflux transporters (mdr1-4, atrF, abcA-E, and msfA-E) (Abad et al., 2010).

Toxins produced by *A. fumigatus*, such as gliotoxin (gliP and gliZ) and ribotoxins (asp f 1/mitF/res), can have immunosuppressive effects and alter protein synthesis. Besides the importance of zinc (zrfA-C, zafA, and pacC) and nitrogen (areA, rhbA, and cpcA) uptake to its virulence and growth, the fungus is also able to obtain nutrients by secreting phospholipases (plb1-3) and other proteases (alp/asp f 13, alp2, asp f 18, mep/asp f 5, mepB, and pep2) in order to break down host tissue. Additionally, regulatory mechanisms that serve as an adaptive response to environmental changes have been identified in *A. fumigatus*: mitogen-activated protein kinase pathways (sakA/hogA, mpkA-C, ste7, pbs2, mkk2, steC/ste11, bck1, and ssk2), G-protein-mediated signal transduction pathways (gpaA, gpaB, sfaD, gpgA, acyA, and pkaC1), and Ras proteins (rasA and rasB) (Abad et al., 2010).

**Fungal cell wall**

The cell wall is a permeable barrier that provides the fungal cell rigidity, protection from environmental stress, and the ability to interact with the host. Its essentiality for fungal life and a lack of counterpart in humans make it an ideal drug target for antifungal development. The *A. fumigatus* mycelium cell wall consists of six
polysaccharides: β-1,3-glucan, chitin, galactomannan, α-1,3-glucan, β-1,3/1,4-glucan, and galactosaminogalactan (Bernard & Latge, 2001). As shown in Fig. 1-2, these polymers are divided according to their solubility in sodium hydroxide solution.

Figure 1-2: Cell wall polysaccharides of *Aspergillus fumigatus* and *Saccharomyces cerevisiae*. Reprinted from Latge, 2007.

β-1,3-glucan and chitin can be found in the alkali-insoluble fraction and are responsible for cell wall rigidity. β-1,3-glucan is highly branched with β-1,6 linkages,
constituting a core to which chitin, β-1,3/1,4-glucan, and galactomannan are covalently linked. Chitin is an $N$-acetylglucosamine polymer that is cross-linked to β-1,3-glucan via a β-1,4 linkage. β-1,3/1,4-glucan is a linear polymer of the repeating unit $3\beta$Glc1-4$\beta$Glc1 not seen in other fungi. α-1,3-glucan, being the major component of the alkali-soluble fraction, forms a pseudonigeran structure containing ~1% α-1,4 linkages in $A.\ fumigatus$. Galactomannan, present in both alkali-insoluble and alkali-soluble fractions, is a linear chain of α-mannose residues with short side chain of β-1,5-galactofuranose residues. Besides its attachment to β-1,3-glucan, galactomannan can also be membrane bound through a glycosylphosphatidylinositol (GPI) anchor and possibly be involved in plasma membrane-cell wall interactions. Galactosaminogalactan, a polymer of $N$-acetylgalactosamine and galactopyranose, is secreted by $Aspergillus$ species but can also be found in both alkali-insoluble and alkali-soluble fractions of mycelium but not conidial cell wall (Bernard & Latge, 2001).

Most of what is known about the conidial cell wall is limited to the outer layer, which is covered by rodlets that are composed of hydrophobins such as RodAp. These proteins confer hydrophobic properties to the conidia and allow for air buoyancy and dispersion. In addition to rodlets, the conidial surface also contains molecules that serve as receptors to various host proteins and facilitate adhesion (Bernard & Latge, 2001).

A comparison of cell wall structures among $S.\ cerevisiae$, $C.\ albicans$, and $A.\ fumigatus$ reveals some similarities in their polysaccharide cores. All three species
contain β-1,3-glucan branched with β-1,6 linkages, and like in _A. fumigatus_, yeast chitin is cross-linked to β-1,3-glucan via a β-1,4 linkage. β-1,6-glucan, which plays a role in interconnecting β-1,3-glucan, chitin, and proteins in _S. cerevisiae_, is absent in _A. fumigatus_. Similarly, the galactomannan, β-1,3/1,4-glucan, and galactosaminogalactan in _A. fumigatus_ have not been identified in yeasts. While glycoproteins have been found to be linked to yeast β-1,6-glucan, they are not bound to β-1,3-glucan in _A. fumigatus_ (Beauvais & Latge, 2001).

Several integral membrane proteins are responsible for the biosynthesis of cell wall polysaccharides. Chitin synthases (CHSs), which catalyze the polymerization of _N_-acetylglucosamine, have been divided into six classes according to their amino acid sequences, with Classes III, V, and VI being specific for filamentous fungi. Both the number of CHS genes and their essentiality vary in fungal species. Out of the four CHS genes in _C. albicans_, only Chs1p is essential for viability. None of the CHS genes identified in _S. cerevisiae_ and _A. fumigatus_ is essential, however. Each CHS consists of an N-terminal hydrophilic conserved region, a neutral conserved region, and a C-terminal hydrophobic region that is believed to anchor the enzyme to the membrane. The N-terminal and central regions are therefore predicted to be on the cytosolic side (Latge, 2007; Beauvais & Latge, 2001).

Responsible for β-1,3-glucan biosynthesis, the integral membrane protein complex glucan synthase contains both catalytic (Fksp) and regulatory (Rho1p) subunits.
The GTPase Rho1p is geranylgeranylated before being transported to the plasma membrane and activated by Rom2p. Fksp can have up to 16 transmembrane domains, and like chitin synthases, the number of FKS genes and their essentiality vary according to the fungal species. While none of the three FKS genes identified in yeasts is essential, the single FKS ortholog (FKS1) found in A. fumigatus is essential (Latge, 2007; Hu et al., 2007).

In A. fumigatus, β-1,3-glucan chains synthesized by the β-1,3-glucan synthase complex are eventually modified by β-1,3-glucanosyltransferases Bgt1p and Gel1p. Bgt1p eliminates laminaribiose from the reducing end of β-1,3-glucan and transfers the remaining portion to the nonreducing end of another β-1,3-glucan with a β-1,6 linkage. Gel1p elongates β-1,3-glucan chains and is attached to the membrane through a GPI anchor (Bernard & Latge, 2001; Beauvais & Latge, 2001).

**Antifungal agents: targets and mechanisms of action**

Antifungal agents (Fig. 1-3), despite their diversity, have traditionally targeted the fungal cell wall and plasma membrane, either directly or indirectly. One major target is the ergosterol biosynthetic pathway, which is responsible for the constitution of fungal plasma membrane. Azoles, the largest class of antifungal agents, are known to inhibit 14α-demethylation of lanosterol (Fig. 1-4) by binding to the protoporphyrin moiety of cytochrome P450-Erg11p, whose active site conformation differs among fungal species and several mammalian P450 monooxygenases. Substituents of azoles have thus
Figure 1-3: Structures of antifungal agents in clinical use. Adapted from Odds et al., 2003.
been modified to enhance their selectivity for fungal P450. Compared to fluconazole, new triazoles such as voriconazole, ravuconazole, posaconazole, and isavuconazole, have broader antifungal spectra (Odds et al., 2003). Isavuconazole, currently in phase III clinical trials, has several advantages over other triazoles: availability in an intravenous form, longer half-life, and \textit{in vitro} activity against zygomycetes (Thompson & Wiederhold, 2010).

![Diagram of antifungal drug targets in the ergosterol biosynthetic pathway](image)

**Figure 1-4: Antifungal drug targets in the ergosterol biosynthetic pathway. Reprinted from Lupetti \textit{et al.}, 2002.**

Two other classes of antifungal agents, allylamines and morpholines, also target the ergosterol biosynthetic pathway (Fig. 1-4). Allylamines, such as terbinafine, inhibit squalene epoxidase, which catalyzes the first oxygenation step in the pathway.
Amorolfine, a morpholine drug, inhibits Erg24p (Δ14 reductase) and Erg2p (Δ8-Δ7 isomerase) instead (Odds et al., 2003).

Unlike azoles, polyene antifungals such as amphotericin B bind to ergosterol directly and contribute to membrane leakage and eventual cell death (Ostrosky-Zeichner et al., 2010), but the exact mechanism remains unclear. Since ergosterol is a main component of fungal plasma membrane, amphotericin B has a broad antifungal spectrum. Despite its 10-fold higher binding affinity for ergosterol over cholesterol, amphotericin B is also toxic to mammalian cells (Moore et al., 2000). It has been available as an antifungal agent since the 1950s (Fig. 1-5), and liposomal reformulations have been utilized to reduce its nephrotoxicity (Odds et al., 2003).

![Figure 1-5: Timeline of antifungal development. ABCD, amphotericin B colloidal dispersion; ABLC, amphotericin B lipid complex; 5-FC, flucytosine; L-AmB, liposomal amphotericin B. Reprinted from Ostrosky-Zeichner et al., 2010.](image)
Instead of targeting cell membrane, echinocandins disrupt fungal cell wall by binding to the Fks1p subunit of β-1,3-glucan synthase complex and inhibiting synthesis of β-1,3-glucan polysaccharides in a non-competitive manner. Even though their binding mechanism to Fks1p is not well understood, echinocandins contain a long side chain known to play a role in their selectivity for different fungal species. Three echinocandin compounds, anidulafungin, caspofungin, and micafungin, have broad spectra of susceptible species and low toxicity but are limited to an intravenous form (Odds et al., 2003).

Other potential targets for antifungal drug development include chitin synthases and β-1,3-glucanosyltransferase, enzymes responsible for the biosynthesis of cell wall polysaccharides and their remodeling. Similarly, chitinases and β-1,3-glucanases, given their morphogenetic roles, have also been proposed as putative targets (Espinel-Ingroff, 2009), together with fatty acid synthase, phospholipases, mitochondrial ATPase activity, filamentation, and even fungal DNA. Chitotriosidase, on the other hand, is secreted by human macrophages and has demonstrated fungicidal activity (Gordon-Thomson et al., 2009). Fig. 1-6 summarizes the mechanisms of action of antifungal agents that are available or under development.

**Mechanisms of antifungal resistance**

Despite the selective targeting of various pathways by different classes of antifungal agents, fungal pathogens have developed antifungal resistance that counters
drug efficacy. Mutations in the *FKS* gene can result in resistance to echinocandins; similarly, *ERG11* mutations (Y132H, S405F, G464S, G465S, R467K, and I471T) or overexpression can contribute to azole resistance or even cross-resistance to drugs that target other steps of the ergosterol biosynthetic pathway (Lupetti et al., 2002). Some of the Erg11p mutations in *C. albicans* are known to lower the affinity of azoles for cytochrome P450 (Morschhauser, 2010).

Even with successful inhibition of 14α-demethylase by azoles, mutations in Δ5-Δ6 desaturase (encoded by *ERG3*) allow for continued fungal growth through an accumulation of 14α-methylfecoesterol instead of 14α-methyl-3,6-diol (Fig. 1-7). Low ergosterol levels, besides conferring amphotericin B cross-resistance, can subsequently reduce azole penetration and its intracellular concentration (Loffler et al., 2000). The vast majority of multidrug resistance in fungi, however, is attributed to the active
Figure 1-7: Fungal survival mechanism as a response to azole inhibition. ERG3* indicates an ERG3 mutation. Reprinted from Lupetti et al., 2002.

transport of structurally diverse drugs out of cells, which also causes intracellular drug level to be below the inhibitory concentration. This effect is analogous to the chemotherapeutic resistance observed in cancer cells due to overexpression of the efflux transporter P-glycoprotein (P-gp or ABCB1) (Morschhauser, 2010).

**Fungal efflux pumps: transcriptional control and drug resistance**

One major mechanism of azole resistance in pathogenic fungi is overexpression of plasma membrane efflux pumps. The transcriptional control of efflux pumps in *S. cerevisiae* appears to be conserved in many pathogenic fungi and serves as a response to antifungal or xenobiotic exposure. In yeast, gene expression is activated by the binding of dimeric zinc finger transcription factors ScPdr1p and ScPdr3p to promoters that
contain palindromic octanucleotide consensus binding sites. Certain molecules can bind to the xenobiotic binding domain of ScPdr1p and ScPdr3p and allow the transcription factors to interact with the ScGal11p subunit of the mediator complex, which can in turn recruit RNA polymerase II for the expression of efflux transporter genes such as ScPDR5, ScSNQ2, ScYOR1, ScTPO1, and ScFLR1 (Cannon et al., 2009).

*S. cerevisiae* and *C. albicans* each has ~30 ABC transporter genes. In contrast, 49 ABC proteins are found in *A. fumigatus* and 54 in *C. neoformans*. In *C. albicans*, zinc finger transcription factor CaTac1p regulates the expression of ABC transporters CaCdr1p and CaCdr2p. Like ScPdr1p and ScPdr3p, gain-of-function mutations in CaTac1p can elevate expression of efflux pumps and drug resistance. Together with CaCdr1p and CaCdr2p, MFS transporters CaMdr1p and CaFlu1p are involved in the transport of known antifungals. Out of these four transporters, CaCdr1p, CaCdr2p, and CaMdr1p are most responsible for azole resistance in *C. albicans*. Compared to CaMdr1p, however, CaCdr1p and CaCdr2p have broader substrate specificity, and CaCdr1p appears to play the major role in clinical *C. albicans* fluconazole resistance (Cannon et al., 2009).

In *C. neoformans*, azole resistance can be attributed to either mutations in the drug target CneErg11p or drug efflux by CneAfr1p and CneMdr1p. CneMdr1p belongs to the MDR subfamily of ABC proteins with a (TMD-NBD)$_2$ topology and is homologous to
human ABCB1. Overexpression of PDR transporter CneAfr1p results in increased azole resistance and virulence in mouse models (Cannon et al., 2009; Morschhauser, 2010).

In addition to 49 ABC transporters, *A. fumigatus* is predicted to have 278 MFS transporters, a total that is more than four times of efflux transporters found in yeasts. Surprisingly, these transporters are not strongly linked to clinical antifungal resistance in *A. fumigatus* despite their abundance. In *A. nidulans*, ABC transporters AtrAp, AtrBp, AtrCp, AtrC2p, and AtrDp have been associated with resistance to some fungicides and xenobiotics. The *A. fumigatus* MFS transporter Mdr3p and PDR transporter atrFp may also be involved in resistance to itraconazole (Cannon et al., 2009).

*A. fumigatus* is innately resistant to fluconazole, and genomic sequences of *A. fumigatus*, *A. nidulans*, and *A. oryzae* indicate that several enzymes in the ergosterol biosynthetic pathway are encoded by multiple copies of genes that belong to separate gene clusters. For example, *CYP51* (equivalent to *ERG11*) can be found in both *CYP51A* and *CYP51B* clusters, but only Cyp51Ap is involved in fluconazole resistance in *A. fumigatus*. Mutations in AfuCyp51Ap Gly54 or Met220 can also lead to itraconazole resistance (Cannon et al., 2009).

**Topology of fungal efflux pumps**

Two main classes, the ATP-binding cassette (ABC) proteins and major facilitator superfamily (MFS) transporters, actively move compounds across cell membrane using either ATP hydrolysis or an electrochemical proton gradient, respectively. Two
members of MFS transporters, which contain transmembrane domains (TMDs) made up of transmembrane spans (TMS), are associated with drug efflux: 12-TMS drug:\text{H}^+ antiporter 1 (DHA1) and 14-TMS DHA2 (Cannon et al., 2009). Residues in TMS5 of \textit{C. albicans} DHA1 Mdr1p have been shown to be critical for drug/H\text{+} transport (Pasrija et al., 2007); despite the involvement of Mdr1p, however, ABC transporters are believed to play the more dominant role in azole resistance.

In addition to two transmembrane domains, ABC transporters usually have two cytoplasmic nucleotide-binding domains (NBDs) for ATP binding and hydrolysis, but the arrangement of these domains varies depending on the type of ABC pump. Some exist as half-size transporters (one NBD and one TMD) that mostly function as either hetero- or homodimers, while others are full-size proteins in a single polypeptide. A (TMD-NBD)$_2$ orientation is seen in most multidrug resistance (MDR) and multidrug resistance-associated protein (MRP) subfamilies, but the pleiotropic drug resistance (PDR) subfamily has a reverse (NBD-TMD)$_2$ arrangement (Lamping et al., 2010). Predicted topological features of a full-size fungal PDR transporter are illustrated in Fig. 1-8. Of the three subfamilies, PDR transporters are most implicated in antifungal resistance (Cannon et al., 2009).

The NBDs contain several conserved motifs such as Walker A (P-loop), Q-loop, ABC signature (C-loop), pro-loop, Walker B, D-loop, and H-loop. Fungal PDR transporters typically consist of two asymmetrical NBDs: NBD1 has a conserved ABC
signature motif (ABC1; [V/I/L/C]SGGE) flanked by degenerate Walker A1 (GX2GXGC[S/T]) and Walker B1 (XWD) motifs, while NBD2 contains a degenerate ABC signature motif (ABC2; LNVEQ) and conserved Walker A2 and B2 motifs (Lamping et al., 2010). As seen in the crystal structures of *Staphylococcus aureus* half-size ABC transporter Sav1866 (Dawson & Locher, 2006) and the mouse homolog of human MDR transporter ABCB1 (Aller et al., 2009), these NBD components will likely form two composite nucleotide-binding pockets (CNBPs) in fungal transporters. Unlike CNBP2, which consists of Walker A2, Walker B2, and ABC1, it remains unclear whether CNBP1 (Walker A1, Walker B1, and ABC2) of PDR transporters can bind and hydrolyze ATP (Lamping et al., 2010).

The cytosolic side of fungal PDR proteins includes both NBDs and four intracellular loops (IL1, IL2, IL3, and IL4). Four small (EL1, EL2, EL4, and EL5) and two large (EL3 and EL6) extracellular loops populate the extracytosolic side instead.
Mutations that attenuate the inhibition of PDR transporters, such as *S. cerevisiae* Pdr5p and *C. albicans* Cdr1p, by small molecules are mostly located in EL3 and EL6. Since mutations in EL5 and EL6 also affect the transport of certain substrates, ELs are believed to be involved in substrate interaction and translocation. EL3 and EL6 both contain highly conserved cysteines that appear to be important for substrate specificity. Furthermore, the invariant glutamate of EL3 and aspartate of EL6 may be responsible for the binding and transport of positively charged substrates. The highly conserved glutamates in TMS3 and TMS9 may also play a similar role (Lamping et al., 2010).

**Substrate selectivity of efflux transporters**

Much of what is known about the transport mechanisms and substrate selectivity of efflux pumps has come from crystal structures of *S. aureus* multidrug binding transcription regulator QacR (Schumacher & Brennan, 2003), mouse homolog of human ABCB1 (Aller et al., 2009), and their respective drug and inhibitor complexes. Structural features of the mouse MDR transporter are consistent with the ATP switch model proposed by Higgins and Linton (Higgins & Linton, 2004). In this model, an inward-facing conformation represents a pretransport state and allows for the binding of substrate to a site in the TMD that is open to the inner leaflet of the membrane and cytoplasm. Binding of ATP likely causes a dimerization in the NBDs and produces a large structural change which results in an outward-facing conformation to facilitate substrate release. The transport cycle is then reset by ATP hydrolysis.
Cocrystal structures of mouse ABCB1 in complex with two stereoisomers of cyclic hexapeptide inhibitors, QZ59-RRR and QZ59-SSS, have revealed the molecular mechanism of substrate specificity, which involves different binding locations, orientation, and stoichiometry. One QZ59-RRR molecule binds to the “middle” site per ABCB1 and is surrounded by mostly hydrophobic residues on TMs 1, 5, 6, 7, 11, and 12; on the other hand, two QZ59-SSS molecules can bind to the transporter. The ligand in the “upper” site is mediated by hydrophobic aromatic residues on TMs 1, 2, 6, 7, 11, and 12, while the one in the “lower” site are close to polar and charged residues such as Gln721, Gln986, and Ser989 (Aller et al., 2009).

A mechanism of substrate specificity which utilizes a central binding cavity that includes multiple drug-binding pockets has also been observed in *S. aureus* multidrug transporter Sav1866 (Dawson & Locher, 2006) and the multidrug binding transcription regulator QacR (Schumacher & Brennan, 2003), which uses negatively charged amino acids such as glutamate to neutralize positively charged substrates. Human ABCB1 is known to interact with itraconazole and ketoconazole (Wang et al., 2002), and as previously mentioned, fungal MDR orthologs are implicated inazole resistance. Given the use of multiple drug-binding pockets by ABCB1 in recognizing a wide variety of substrates such as anthracyclines, alkaloids, steroid hormones, local anesthetics, and dye molecules (Seelig, 1998), residues involved in the binding of QZ59-RRR, QZ59-SSS, and verapamil might not interact with triazoles at all. Without a crystal structure of efflux
transporter-triazole complex, the binding mechanism and selectivity of triazoles remain unclear.

Structural features of P-glycoprotein substrates that are required for interaction with the transporter have been proposed and involve either two or three electron donor groups with fixed spatial separation. Type I units contain two electron donor groups separated by $2.5 \pm 0.3 \text{ Å}$, while type II units consist of either two electron donor groups separated by $4.6 \pm 0.6 \text{ Å}$ or three electron donor groups with a $4.6 \pm 0.6 \text{ Å}$ separation of the outer two groups (Seelig, 1998). Itraconazole, ketoconazole, and fluconazole represent the type II unit of three electron donor groups (Wang et al., 2002), which serve as hydrogen bonding acceptors in substrate binding (Seelig, 1998).

**Summary**

As a result of medical treatments and trends in the last few decades, an increasing population of immunocompromised individuals has led to the emergence of opportunistic fungal pathogens as a threat to human health. Many current antifungals target essential gene products involved in pathways such as cell wall and ergosterol biosyntheses that are unique to fungi, but given limited therapeutic options and developing antifungal resistance, minimizing fungal virulence may be an alternative approach in antifungal development. The availability of whole-genomic sequence of pathogenic fungi has accelerated the elucidation of genes essential for fungal growth, survival, and virulence, further complementing our understanding of their pathogenesis.
and metabolism, as well as the identification of potential drug targets. Some of the recent efforts have been focused on the protein prenylation pathway, the topic of discussion in Chapter 2.
Chapter 2. Protein prenylation

Introduction

Many proteins required for eukaryotic cell growth, differentiation, and cellular signaling pathways undergo protein prenylation, a post-translational modification of isoprenoid lipid attachment by three protein prenyltransferases: protein farnesyltransferase (FTase), protein geranylgeranyltransferase type I (GGTase-I), and protein geranylgeranyltransferase type II (GGTase-II or RabGGTase). FTase and GGTase-I catalyze the covalent attachment of a 15-carbon (FPP) or 20-carbon (GGPP) isoprenoid lipid, respectively, to the thiol group of cysteine in a C-terminal CaaX recognition motif, where C represents the cysteine residue, followed by two aliphatic (aa) residues and a specificity-determining X residue. GGTase-II, on the other hand, catalyzes the transfer of two geranylgeranyl moieties to the C-terminal end of Rab proteins containing CC, CCXX, or CXC motifs (Eastman et al., 2006), but this attachment requires the interaction of Rab proteins with a carrier called Rab escort protein (Rep).

As illustrated in Fig. 2-1, most prenylated proteins are further modified in the endoplasmic reticulum after isoprenoid attachment in the cytoplasm. Following removal of the aaX tripeptide by the protease Rce1, isoprenylcysteine carboxymethyltransferase (Icmt) catalyzes the carboxymethylation of prenylcysteine using S-adenosylmethionine. With the exception of K-Ras4B, most Ras proteins are also
palmitoylated at the Golgi in a reaction catalyzed by mammalian DHHC9 and GCP16 where the C16 fatty-acid chain of palmitoyl CoA is transferred to one or more cysteines near the farnesylated cysteine (Gelb et al., 2006). Prenylation and subsequent modifications have the combined effect of increasing the hydrophobicity of proteins, allowing for proper membrane association and enhanced protein-protein interactions (Lane & Beese, 2006; Gelb et al., 2006; Rowinsky et al., 1999). K-Ras4B (Fig. 2-1) is believed to associate with the plasma membrane through electrostatic interactions between the acidic phospholipids and polylysine sequences upstream of the CaaX motif (Gelb et al., 2006).

Figure 2-1: Prenylation and modification of proteins. Farnesyl and palmitoyl moieties are shown in blue and red, respectively. Reprinted from Gelb et al., 2006.
Roles of prenylation targets

Proteins that undergo prenylation include ones from the Ras, Rab, Rho, and Rap families, the γ subunit of heterotrimeric G proteins, nuclear lamins, and centromeric proteins (Lane & Beese, 2006; Gelb et al., 2006). By switching between an active GTP-bound state and an inactive GDP-bound state, Ras proteins are small GTPases that play a major role in activation of signal transduction pathways mediating cellular processes such as growth, differentiation, and apoptosis (Rowinsky et al., 1999). After recruited and activated by Ras in the plasma membrane, Raf-1 phosphorylates mitogen-activated protein kinase kinases MEK1 and MEK2, which in turn phosphorylate MAPKs ERK1 and ERK2 (Fig. 2-2). In the nucleus, MAPKs can activate substrates such as the Elk-1 transcription factor of the E-twenty six (ETS) family (Rowinsky et al., 1999).

Figure 2-2: Signaling pathways in mammalian cells involving Ras family members. Reprinted from Berndt et al., 2011.
Mutational activation of the Ras subfamily, including H-, K-, and N-Ras, traps the Ras proteins in the active GTP-bound state and results in constitutive signal transduction (Basso et al., 2006). Accounting for ~20% of human cancers, these mutations include K-Ras in pancreatic and colorectal cancers, N-Ras in melanoma and hepatocellular cancer, and H-Ras in bladder and renal cell cancers. While K- and N-Ras are cross-prenylated by FTase and GGTase-I, H-Ras can only be farnesylated (Basso et al., 2006; Konstantinopoulos et al., 2007). Since prenylation of mutant Ras proteins is necessary for their oncogenic effects, protein prenyltransferases have been targets for cancer therapeutics.

Other members of the Ras superfamily include Rho and Rheb, which also function as small GTPases. Responsible for cytoskeletal reorganization and cellular motility, Rho proteins are predominantly geranylgeranylated but can also be farnesylated. They undergo modifications by Rce1 and Icmt and subsequently bind to Rho guanine nucleotide dissociation inhibitors (GDIs) for increased solubility and proper membrane delivery (Konstantinopoulos et al., 2007). Rheb, on the other hand, is activated by phosphorylation of tuberous sclerosis complex TSC1-TSC2 as part of the PI3K-AKT pathway and is exclusively farnesylated. It regulates mTOR signaling, which plays a role in cell growth, transcription, and cell survival through S6K and 4EBP1 activation (Berndt et al., 2011). Enhanced Rheb activity is often observed in human tumor cell lines and correlated with tumor predisposition (Basso et al., 2006).
Essentiality of protein prenylation

In mammals, protein farnesylation is essential for early embryonic proliferation but not for postnatal development and adult homeostasis (Mijimolle et al., 2005). FTase, while nonessential for Saccharomyces cerevisiae viability, is involved in growth and fitness (He et al., 1991). This implies that the loss of protein farnesylation is compensated by alternative mechanisms such as GGTase-I cross-prenylation (Trueblood et al., 1993; Whyte et al., 1997) and warrants the inhibition of FTase for cancer therapeutics (Sebti, 2005).

Compared to mammals and yeast, FTase has been shown to be essential in Cryptococcus neoformans (Vallim et al., 2004). The Ram2 gene, which encodes the common α subunit of FTase and GGTase-I, is required for Candida albicans viability (Song & White, 2003), but the GGTase-I β subunit (encoded by Cdc43) appears to be nonessential (Kelly et al., 2000). Even though the protein prenylation pathway is also present in other pathogenic protozoan organisms such as Plasmodium falciparum and Trypanosoma brucei, its essentiality has only been indirectly tested with mammalian FTase inhibitors (Chakrabarti et al., 1998; Ohkanda et al., 2001; Yokoyama et al., 1998). Nonetheless, these inhibitors have the potential to be adapted for antifungal therapies, making the prenylation pathway an attractive target for treating fungal infections.
Mutant phenotypes of \textit{A. fumigatus} prenylated gene products

Unlike in \textit{C. neoformans} and \textit{C. albicans}, essentiality of protein prenylation has not yet been established for \textit{Aspergillus fumigatus}, but inhibition of HMG-CoA reductase, a key enzyme in the biosynthesis of isoprenoid diphosphate and ergosterol, can lead to severe growth defects (Macreadie et al., 2006). Furthermore, mutant phenotypes of prenylated gene products in this fungal organism are known. The fungal Ras family of proteins is a group of monomeric GTPases that functions in a variety of cellular processes and plays a major role in growth and virulence. Unlike in \textit{Aspergillus nidulans}, \textit{rasA} is not essential in \textit{A. fumigatus}, but Δ\textit{rasA} mutants exhibit severe growth defects in hyphal and colony morphologies, delay in germination, and nuclear aberrancies. While Δ\textit{rasA} phenotypes can be enhanced at elevated temperatures and be rescued by exogenous cAMP in \textit{C. albicans} and \textit{C. neoformans}, these phenomenon are not observed in \textit{A. fumigatus}. Increased sensitivity to cell wall stress has also been identified in \textit{A. fumigatus} Δ\textit{rasA} mutants, but the mechanism by which RasA regulates cell wall integrity is unclear. This phenotype of cell wall instability, however, raises the possibility of targeting RasA for antifungal therapy (Fortwendel et al., 2008).

Similar to \textit{rasA}, \textit{rasB} is nonessential in \textit{A. fumigatus}, but Δ\textit{rasB} phenotypes include delayed germination, decreased growth rates on solid media, and abnormal hyphal morphology, leading to reduced virulence in a mouse model of invasive aspergillosis (Fortwendel et al., 2005). The \textit{rhbA} gene (a \textit{rheb} homolog), while not essential in \textit{A. fumigatus}...
is involved in nitrogen metabolism, which appears to be linked to pathogenicity (Hensel et al., 1998). Besides impaired growth on poor nitrogen sources, ΔrhbA mutants also show diminished virulence in a murine infection model (Panepinto et al., 2003). Farnesol, a quorum sensing molecule of *Candida albicans*, has been proposed to directly or indirectly target Rho1, the regulatory subunit of β-1,3-glucan synthase complex, and disturb cell wall integrity in *A. fumigatus*. Despite its added effect of altering growth phenotypes and interfering with localization of Rho1 and Rho3 to hyphal tips (Beauvais et al., 2001), farnesol has only a limited inhibition on the growth of *A. fumigatus*. This susceptibility is elevated in cell wall mutants Δmnt1 and ΔglfA, however, so farnesol has the potential to be included in a combination therapy for invasive aspergillosis (Dichtl et al., 2010).

**Overview of mammalian protein prenyltransferases**

Despite the importance of protein prenylation in both mammals and microorganisms, structural features of mammalian protein prenyltransferases, including isoprenoid and peptide binding conformations and specificities, zinc coordination, and product exit groove, are better characterized than that of fungal orthologs. As illustrated in Fig. 2-3, the major steps along the mammalian protein prenylation pathway have been delineated by X-ray crystal structures of FTase and GGTase-I in complex with various substrates (Long et al., 2002; Taylor et al., 2003). Kinetic studies of FTase reveal an ordered substrate binding mechanism in which FPP binds to the apo
enzyme first to form a binary substrate complex, followed by the binding of CaaX peptide to form a ternary substrate complex. A second FPP binds and forces the prenyl moiety of the farnesylated product to a new site known as the exit groove, and crystal structures of this displaced product complex are consistent with results from kinetic studies showing that product release is the rate-limiting step of the reaction (Furfine et al., 1995; Huang et al., 2000; Pickett et al., 2003).

Figure 2-3: Reaction pathway of protein farnesylation. (0) Crystal structure of unliganded rat FTase. The α subunit is colored in red and β subunit in blue. (1) Binary complex with FPP (blue) and Zn$^{2+}$ (magenta) bound. (2) Ternary complex with FPP analog (blue), Zn$^{2+}$, and CaaX peptide (yellow) bound. (3) Product complex with farnesylated peptide and Zn$^{2+}$ bound. (4) Displaced product complex with farnesylated peptide, Zn$^{2+}$, and a second FPP bound. Reprinted from Long et al., 2002.
Overall architecture and zinc binding

As obligate αβ heterodimers, FTase and GGTase-I share a common α subunit and a homologous β subunit. Even though the β subunits of the two enzymes are only about 25% identical in sequence, they are structurally very similar. The 15 anti-parallel α helices of α subunit are arranged in a crescent-shaped superhelix that wraps around the β subunit, which consists of 14 α helices in FTase and 13 in GGTase-I and adopts an α-α barrel with a funnel-shaped cavity at the center that serves as the active site (Park et al., 1997). A topology diagram of rat FTase is shown in Fig. 2-4.

Figure 2-4: Topology diagram of rat FTase. α helices are shown as open boxes, 3_10 helices as striped boxes, and β strands as arrows. Reprinted from Park et al., 1997.
As zinc metalloenzymes, both FTase and GGTase-I require Zn$^{2+}$ for catalytic activity. The zinc ion is coordinated by D297β, C299β, and H362β in FTase and D269β, C271β, and H321β in GGTase-I. In a ternary complex of rat FTase, five ligands (Fig. 2-5) are involved in zinc coordination: two ligands from D297β (2.0 and 2.6 Å), one from C299β (2.3 Å), one from H362β (2.1 Å), and one from CaaX cysteine (2.3 Å). Interaction of the CaaX cysteine with zinc is increased to 2.6 Å in a product complex (Lane & Beese, 2006).

![Zinc coordination in human FTase ternary complex. Reprinted from Lane & Beese, 2006.](image)

**Figure 2-5: Zinc coordination in human FTase ternary complex. Reprinted from Lane & Beese, 2006.**

**Binding and specificity of isoprenoid diphosphate**

FPP binds along one side of the hydrophobic cavity; the diphosphate moiety interacts with K164α, H248β, R291β, and Y300β in mammalian FTase and K164α, H219β, R263β, K266β, and Y272β in GGTase-I. The farnesyl portion interacts with several conserved aromatic residues and has essentially identical binding conformation as the
first three isoprene units of GGPP bound to GGTase-I (Lane & Beese, 2006). The fourth isoprene unit, however, is kinked by ~90° and accommodated by a space created from smaller threonine and phenylalanine residues as opposed to the bulky W102β and Y365β at the corresponding positions in FTase. A hypothesis of isoprenoid substrate specificity through a length-discriminating mechanism is supported by studies in which a W102T FTase mutant has enhanced preference for GGPP over FPP (Taylor et al., 2003). It is also consistent with the observation that prenylation efficiency of FTase and GGTase-I is highly dependent on the correct isoprenoid diphosphate substrate (Lane & Beese, 2006).

**Binding and specificity of CaaX peptide**

As seen in crystal structures of ternary complexes, binding of CaaX peptide does not change the conformation of either the enzyme or the bound isoprenoid, which provides a significant binding surface for the peptide substrate through van der Waals contact. The CaaX motif binds in an extended conformation, anchored at the cysteine residue by a direct zinc-coordination and at the C-terminal end by hydrogen bonds. The two points effectively allow the enzyme to discriminate against peptides that do not have the proper length or lack a cysteine at the correct position (Lane & Beese, 2006). In this binding mode, the a₁ position is solvent-exposed and can accept any amino acid. Specificity at the a₂ position is dominated by hydrophobic and steric interactions with W102β, W106β, Y361β, and the third isoprene of FPP (Fig. 2-6). The backbone carbonyl
oxygen of α₂ residue also forms a hydrogen bond with R202β of FTase (R173β in GGTase-I). The α₂ pocket accommodates valine, leucine, isoleucine, phenylalanine, tyrosine, proline, threonine, and methionine, but large, very polar, or charged residues are likely to be rejected (Reid et al., 2004).

Figure 2-6: Residues of rat FTase that interact with the CaaX motif. Reprinted from Long et al., 2000.

The X residue is in van der Waals contact with Y131α, A98β, S99β, W102β, H149β, A151β, and P152β in FTase (Long et al., 2000) but with T49β, H121β, A123β, and F174β in GGTase-I (Lane & Beese, 2006). Furthermore, its C-terminal carboxylate forms a hydrogen bond with Q167α (Fig. 2-6). Compared to that of GGTase-I, the X residue binding pocket of FTase is more polar and accepts methionine, serine, and glutamine. Kinetic and modeling studies reveal alanine, threonine, and cysteine as additional...
substrates but not larger residues such as tyrosine, tryptophan, and arginine (Reiss et al., 1991; Moores et al., 1991; Roskoski & Ritchie, 1998; Reid et al., 2004). A C-terminal phenylalanine fits in an alternative binding site, which can possibly also accommodate leucine, asparagine, or histidine (Reid et al., 2004).

An alternative model for CaaX selectivity involves two substrate classes that are farnesylated under either multiple-turnover (MTO) or single-turnover (STO) conditions. Selectivity of MTO peptides is consistent with the previously described CaaX model, but STO substrates appear to follow a different set of specificity requirement. They show a preference for Cys and discriminate against Leu at the a1 position. In addition, they favor small amino acids instead of aliphatic ones at the a2 position but have no preference at the X position (Hougland et al., 2010).

**Product formation and exit groove**

Structural changes after protein prenylation are mostly limited to the conformation of the isoprenoid moiety. With the release of pyrophosphate, the first and second isoprenes shift toward the CaaX peptide, which retains the same conformation after catalysis. Binding of a second isoprenoid diphosphate, however, pushes the isoprenoid moiety of the product to a new site called the exit groove. While the a2 and X residues occupy the same positions, the CaaX moiety switches from an extended conformation to a type I β-turn. As shown in Fig. 2-7, residues in the exit groove that
interact with the displaced farnesyl moiety include A92β, Y93β, C95β, L96β, W106β, S357β, and D359β (Long et al., 2002).

Figure 2-7: Exit groove residues (in blue) of rat FTase that interact with the displaced farnesyl moiety (brown). Reprinted from Long et al., 2002.

Overview of fungal protein prenyltransferases

Crystal structures of C. neoformans FTase (Hast et al., 2011) and C. albicans GGTase-I (Hast & Beese, 2008) indicate that although the two enzymes have similar overall architecture as their respective mammalian orthologs, their active sites exhibit significant divergence in comparison. In CaGGTase-I, F99β (T127β in rat GGTase-I) pushes the fourth isoprene away from the CaaX binding pocket, but C. albicans protein prenyltransferases appear to utilize a similar length-discriminating mechanism for isoprenoid substrate specificity as their mammalian counterparts (Hast & Beese, 2008).
Despite several non-conservative amino acid substitutions in the X residue binding pocket (Fig. 2-8), the X residue specificity of CaGGTase-I is comparable to that of the mammalian ortholog (Mazur et al., 1999). The exit groove of the fungal enzyme is also narrowed by ~4 Å, presenting a potential steric block to the displaced isoprenoid moiety (Hast & Beese, 2008).

Crystal structures of *C. neoformans* FTase in complex with various substrates show that the helix 4α-5α loop of the enzyme can undergo substrate-dependent conformation change of ~7 Å (Fig. 2-9), a rearrangement not observed in mammalian protein prenyltransferases. Unlike rat FTase, the fungal ortholog shows preference for Leu and Tyr but discriminate against Cys, His, Thr, Ser, and Phe at the X position.
consistent with an alanine (hFTase A151β) to leucine (CnFTase L141β) substitution in the X residue binding pocket. Its exit groove, which contains multiple amino acid substitutions, is also widened by a helical movement of residues 80β-84β (Hast et al., 2011). Structural divergence between the two FTase orthologs can be exploited for design of CnFTase inhibitors with increased selectivity.

Figure 2-9: Superposition of CnFTase structures in apo/inhibitor-bound (gray) and CaaX-bound conformations (orange). FPP analog is shown in red and CVIM substrate in yellow. Reprinted from Hast et al., 2011.

Summary

More than 100 proteins required for eukaryotic cell growth, differentiation, and cellular signaling pathways undergo protein prenylation, a post-translational
modification of isoprenoid lipid attachment to the cysteine of a C-terminal CaaX recognition motif. This covalent attachment is catalyzed protein prenyltransferases that include FTase and GGTase-I, which are obligate αβ heterodimers that require Zn$^{2+}$ for activity. In addition to the protein prenylation pathway, structural features of mammalian protein prenyltransferases, including isoprenoid and peptide binding conformations and specificities, zinc coordination, and product exit groove, have been characterized by X-ray crystallography. Targets of protein prenylation include several Ras proteins, and since prenylation of Ras mutants is necessary for their oncogenic effects, protein prenyltransferases have been targets for cancer therapeutics. Some of the mammalian FTase inhibitors have the potential to be adapted for antifungal therapies by blocking protein prenylation in pathogenic fungi. Improving inhibitor selectivity, however, would benefit from structural details of fungal protein prenyltransferases such as C. neoformans FTase and C. albicans GGTase-I. In the following chapter, we present a high resolution crystal structure of the A. fumigatus FTase and a characterization of its substrate selectivity.
Chapter 3. Structure and substrate selectivity of *Aspergillus fumigatus* protein farnesyltransferase

Introduction

The focus of this project is on the fungal organism *Aspergillus fumigatus*. Species of the fungal genus *Aspergillus* play an important role in carbon and nitrogen metabolism and aerobic decomposition of organic materials. Their ability to produce airborne conidia by asexual reproduction allows them to be ubiquitous in the environment. The daily exposure to these large numbers of conidia is normally dealt with by the immune system. *Aspergillus fumigatus*, being the predominantly occurring species, is a major opportunistic human pathogen of invasive aspergillosis, which is becoming more common among transplant recipients and cancer and AIDS patients. Its high mortality rate (45 to 90%) and limited effectiveness of current antifungal therapies have emphasized the need for new drug development (Lin et al., 2001).

The first step toward inhibitor design for *A. fumigatus* protein prenyltransferases is obtaining high resolution structures of the enzymes, ideally at all steps of the reaction pathway. X-ray crystallography was chosen for structure determination due to the amount of foundational work already done on many protein prenyltransferases by the Beese lab and the size of these proteins (~80-100 kDa). The objectives of this structural work can be divided into two general categories: substrate and inhibitor complexes.
The first goal is to determine the structure of \textit{A. fumigatus} FTase and the basis of substrate specificity. Structures of binary and ternary substrate complexes are key components of understanding the molecular basis of substrate specificity, and a structural comparison to the mammalian enzyme will highlight potential differences and unique features in both the isoprenoid and CaaX peptide binding pockets, especially at the \( a_2 \) and X positions. Modeling studies can subsequently be done to expand on the substrate specificity hypothesis, which can be further tested by an \textit{in vitro} fluorescence-based activity assay. A detailed analysis of enzyme substrate specificity is a major step toward improving inhibitor selectivity.

The aim is to improve the resolution of \textit{A. fumigatus} FTase X-ray crystal structure and, through a structural comparison with the available mammalian FTase structures, identify and characterize any features of the active sites that may lead to the design of new fungal inhibitors with high selectivity. These features include isoprenoid and peptide binding conformations and specificities, zinc coordination, and product exit groove.

A 1.8 Å crystal structure of the \textit{A. fumigatus} FTase binary complex has already been solved by Michael Hast in the lab. In the present study, the resolution of this structure has been improved to 1.5 Å, matching the highest ever obtained for a published protein prenyltransferase crystal structure (PDB code 2H6F). Data
redundancy was also improved by merging two collected data sets. Data collection and refinement statistics for the binary complex are listed in Table 3-1.

<table>
<thead>
<tr>
<th>Table 3-1: Data collection and refinement statistics for the binary complex</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>AfFTase:FPP</strong></td>
</tr>
<tr>
<td>Resolution range Å</td>
</tr>
<tr>
<td>Data collection</td>
</tr>
<tr>
<td>Reflections, unique/total</td>
</tr>
<tr>
<td>I/σI</td>
</tr>
<tr>
<td>R_sym</td>
</tr>
<tr>
<td><strong>Unit cell</strong></td>
</tr>
<tr>
<td>Dimensions (Å)</td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td>Angles (°)</td>
</tr>
<tr>
<td>Space group</td>
</tr>
<tr>
<td>Completeness %</td>
</tr>
<tr>
<td>Multiplicity</td>
</tr>
<tr>
<td><strong>Refinement</strong></td>
</tr>
<tr>
<td>R_crys</td>
</tr>
<tr>
<td>R.m.s.d. for ideal bond length, Å</td>
</tr>
<tr>
<td>R.m.s.d. for ideal bond angles, °</td>
</tr>
<tr>
<td>Overall isotropic B factor, Å²</td>
</tr>
<tr>
<td><strong>Ramachandran plot</strong></td>
</tr>
<tr>
<td>Favorled, %</td>
</tr>
<tr>
<td>Allowed, %</td>
</tr>
</tbody>
</table>

**Structure of Aspergillus fumigatus FTase:FPP**

The *A. fumigatus* FTase (AfFTase) is a 97-kDa heterodimer composed of a 353-residue α subunit and a 519-residue β subunit, which share 36 and 44% sequence identity, respectively, with that of human FTase. With 1.65 Å r.m.s.d. calculated over all aligned Cα atoms, the two orthologs are similar in their overall structural architecture.
(Fig. 3-2a). Similar to what has been observed in human FTase, the α helices in AfFTase α subunit are arranged in a crescent-shaped domain that wraps around the α-α barrel core of β subunit (Fig. 3-1).

Compared to the 379-residue α subunit of human FTase, the AfFTase α subunit is smaller and contains an N-terminal region that is not rich in Pro and Glu. Like rFTase α subunit, the AfFTase α subunit consists of 15 anti-parallel α helices (Fig. 3-1). Helix 6α of rFTase is half a helical turn longer than that of AfFTase, but the loop connecting 5α and 6α is longer in AfFTase (11 residues) than in rFTase (5 residues). Similarly, the loop connecting helices 9α and 10α is much longer in AfFTase (28 residues) than in rFTase (6 residues).

For the AfFTase α subunit, the N-terminus (residues 1-28), C-terminus (residues 345-353), and two loops (residues 274-286 and 304-308) are missing in the electron density map. The missing residues 274-286 make up a loop connecting helices 12α and 13α that is potentially large enough to form an extra helix, such as 13α in C. albicans GGTase-I (Hast & Beese, 2008). The missing residues 304-308 are located between helices 13α and 14α, which are both one helical turn shorter than their corresponding helices in rFTase.

Compared to the 437-residue β subunit of human FTase, the AfFTase β subunit is larger but still mostly consists of 14 α helices forming an α-α barrel (Fig. 3-1). Loop connecting helices 7β and 8β is much longer in AfFTase (23 residues) than in rFTase (10
Figure 3-1: *A. fumigatus* protein farnesyltransferase (α subunit in red and β subunit in blue). Together with isoprenoid substrate FPP (green) and catalytic zinc ion (pink), the 15 α-helices of α subunit and termini of both subunits are labeled.

Figure 3-2: (a) AfFTase (α subunit in red and β subunit in blue) superimposed on rat FTase (PDB code 1FT2; gray). (b) Different view of the superposition. Only the β subunit of rat FTase (in gray) is illustrated in this panel. Yellow ribbons highlight the major structural differences in AfFTase β subunit in comparison to rat FTase.
residues) and forms a short helix. In addition, the N-terminus (residues 1-68) and two loops (residues 389-394 and 456-461) of the AfFTase β subunit are missing in the electron density map. Residues 389-394 are located between helices 11β and 12β, while residues 456-461 are situated between two anti-parallel β strands. Like in rFTase, these β strands in AfFTase β subunit are flanked by helices 13β and 14β, but despite similar topology, the β sheets in the two protein orthologs are oriented orthogonally to each other (Fig. 3-2b). The loop connecting the second β strand and helix 14β is longer in AfFTase (29 residues) than in rFTase (20 residues), and the 13 residues following the second β strand in AfFTase form a loop that occupies a completely different space than that occupied by the corresponding loop in rFTase. These structural differences, highlighted in yellow in Fig. 3-2b, are mostly located on the outer surface of β subunit and distant from the active site.

**Comparison of zinc, isoprenoid, and magnesium binding sites**

The *A. fumigatus* FTase:FPP coordinates were superimposed with that of *R. norvegicus* FTase binary (PDB code 1FT2) and ternary (PDB code 1D8D) complexes in order to structurally compare their active sites. However, given the 3.4 Å resolution of the 1FT2 structure, the rat FTase displaced product complex (PDB code 1KZO) was also used for comparison. Similar to that in rFTase, Zn\(^{2+}\) is coordinated by an aspartic acid (D365β), cysteine (C367β), and histidine (H433β) in AfFTase. Five ligands are involved in this coordination: two ligands from D365β (2.03 and 2.47 Å), one from C367β (2.30 Å),
one from H433β (2.15 Å), and a water molecule (2.02 Å). Residues lining the isoprenoid-binding cavity are entirely conserved between the two orthologs, but conformation of FPP in AfFTase binary complex is a bit different compared to that in rFTase binary complex; the curvature of kink between α-phosphate and the first isoprene is not as pronounced, leading to a more linearly extended FPP in AfFTase and a pyrophosphate shift toward the β subunit (Fig. 3-3).

It remains possible that the observed discrepancy in isoprenoid binding mode between the two orthologs is merely an artifact of significant difference in the resolution.
of the corresponding crystal structures. The overall FPP conformation in *A. fumigatus* FTase binary complex is more similar to that in rat FTase displaced product complex compared to the binary complex. The linear isoprenoid extension in AfFTase, especially conformations of the pyrophosphate and first isoprene, is also retained in the displaced product complex of the rat enzyme. The second and third isoprenes in AfFTase, however, are conformationally different from that in rat FTase binary and displaced product complexes. As illustrated in Fig. 3-3, AfFTase A313β (G200β in rat) pushes the second isoprene slightly toward the CaaX binding pocket, while an R252β (R202β in rat) side chain conformational change moves the third isoprene in an orthogonal direction. Shifts in second and third isoprenes are also present in the rat FTase ternary complex despite a slight length difference between FPP and FPT-II (a non-hydrolyzable FPP analog).

Side chain of AfFTase K107α (K164α in rat) is missing most of its electron density but is still within hydrogen bonding distance to the α-phosphate. As shown in Fig. 3-4, K164α adopts different side chain conformations in rFTase binary, ternary, and displaced product complexes, implying a certain degree of conformational flexibility in this particular residue and allowing it to hydrogen bond with the diphosphate moiety of FPP in a binary complex (Lane & Beese, 2006) but also stabilize peptide binding (Hast et al., 2011).
A slightly different conformational flexibility has also been reported for the conserved K107α of *C. neoformans* FTase (Hast et al., 2011). This residue is located in the middle of a loop that connects helices 4α and 5α and can undergo ligand-dependent conformational changes (Fig. 3-4). Upon peptide binding, the helix 4α-5α loop moves ~7 Å and adopts a conformation similar to that in rFTase, but unlike in CnFTase, the 4α-5α loop appears to be rigid and identical in various mammalian FTase complexes. There is only a minor loop movement of ~1.3 Å between AfFTase and either rFTase ternary or
displaced product complex, and without a structure of the AfFTase ternary complex, it is unclear whether the AfFTase 4α-5α loop undergoes substrate-dependent conformational changes. The missing side chain electron density of AfFTase K107α, however, suggests that multiple side chain conformations might be possible for this residue.

W153β and Y204β of AfFTase are located near the third isoprene of FPP and are conserved in rFTase as W102β and Y365β, which are known to play a role in excluding the fourth isoprene unit of GGPP (Taylor et al., 2003). This suggests that both FTase orthologs utilize length discrimination as the molecular basis of isoprenoid specificity. The hypothesis can be tested by mutating W153β to the corresponding residue in A. fumigatus GGTase-I and determining the mutant’s preference for GGPP over FPP. Even though magnesium dependence of AfFTase has yet to be established, the residues in mammalian FTase believed to coordinate Mg\(^{2+}\) and stabilize charge on the diphosphate moiety during the transition state are conserved in AfFTase (D423β, Y368β, and K107α) with minor positional adjustments to account for the pyrophosphate shift.

**Structural differences in the CaaX binding pocket**

Structures of AfFTase binary complex and rFTase ternary complex were superimposed in order to locate the CaaX binding pocket and evaluate residues that serve as determinants of substrate specificity. Residues around a1 are all conserved between the two orthologs, and there appears to be enough space to accommodate any
amino acid at the $a_1$ position, similar to what has been observed in rFTase (Reid et al., 2004).

$Y_{361}\beta$, $W_{106}\beta$, and $W_{102}\beta$, which interact with $a_2$ in rFTase, are conserved in AfFTase as $Y_{432}\beta$, $W_{157}\beta$, and $W_{153}\beta$, respectively (Fig. 3-5). Compared to $W_{102}\beta$ of the rFTase ternary complex, AfFTase $W_{153}\beta$ rotates ~30° clockwise along the $\beta$-$\gamma$ axis in order to avoid clashing with the third isoprene and $A_{201}\beta$, likely altering the shape of $a_2$ binding pocket. Since the isoprenoid forms a large portion of the CaaX binding surface in mammalian FTase via hydrophobic interaction, displacement of the second and third isoprenes toward the CaaX binding pocket effectively shrinks the $a_2$ pocket size, so $a_2$ selectivity might not be identical between the two orthologs. For example, a modeled $a_2$ isoleucine clashes with the second isoprene in AfFTase binary complex. The second isoprene has very little room to move because of $A_{313}\beta$, so either the $a_2$ binding pocket cannot accept an isoleucine, or the CaaX peptide has enough flexibility or space to adjust and allow isoleucine to fit in the $a_2$ pocket. The needed space may potentially come from AfFTase $Y_{109}\alpha$ ($Y_{166}\alpha$ in rat) side chain, which rotates ~45° counterclockwise toward the $\alpha$ subunit and provides additional room for $a_2$ and $X$ residue backbone (Fig. 3-5).

Unlike AfFTase $Y_{436}\beta$, CnFTase has a single non-conservative $N_{413}\beta$ substitution for rFTase $Y_{365}\beta$ in the $a_2$ binding pocket (Fig. 3-5). This substitution deepens the cavity and may contribute to $a_2$ specificity differences between the mammalian and $C. neoformans$ orthologs (Hast et al., 2011). Displacement of the third
isoprene is also less in CnFTase compared to AfFTase, so CnFTase W90β side chain conformation is more similar to that of rFTase W102β than of AfFTase W153β.

![Superposition of AfFTase binary complex (green), CnFTase ternary complex (orange), and rFTase ternary complex (pink) containing a CVIM substrate (yellow). Residues believed to interact with α and X residues (including those in the secondary binding pocket) are shown as sticks. Color labeling matches the corresponding protein ortholog.](image)

Fig. 3-5 also highlights the residues that can potentially contribute to any X residue selectivity differences between rFTase and AfFTase. With the exception of rFTase P152β (S202β in AfFTase), most residues lining the X residue binding pocket are conserved in the two orthologs, but many of them are positionally and conformationally
different. For example, AfFTase L147β is located on a $3_{10}$ helix which rotates $\sim 25^\circ$ clockwise compared to that in rFTase. This helix is followed by a loop that is partly constituted by A149β and S150β and, relative to the corresponding rFTase loop, shifts slightly toward the helical axis. These helical and loop movements cause L147β and S150β to move 2.96 Å and 1.73 Å, respectively, away from the X residue, increasing the size of binding pocket. The rFTase P152β to AfFTase S202β substitution further increases pocket size and alters polarity of binding surface, so the X residue specificity of the two protein orthologs is expected to be somewhat different.

Q167α, which anchors the C-terminus of CaaX peptide in mammalian FTase, is conserved in AfFTase (Q110α), but Q110α clashes with the C-terminus of a modeled X residue. Side chain of Q167α adopts different conformations in rFTase binary and ternary complexes, and the Q110α side chain conformation in AfFTase binary complex is more similar to that of Q167α in rFTase ternary complex than in rFTase binary complex. AfFTase Q110α might have enough conformational flexibility to accommodate the X residue, or alternatively, the CaaX peptide might be able to adjust and stay anchored to Q110α at its C-terminus.

Similar to AfFTase, most residues lining the X residue pocket of CnFTase are conserved with the exception of L141β in place of rFTase A151β (Hast et al., 2011), a substitution that alters the shape of the pocket. Side chain conformations of other CnFTase residues in the X residue binding pocket are overall more similar to that of
rFTase residues than of AfFTase residues (Fig. 3-5), especially CnFTase Y109α and R197β (Y109α and R252β in AfFTase, respectively). The 3_10 helix in which CnFTase L84β (L147β in AfFTase) is located also rotates clockwise relative to that in rFTase but to a lesser extent than the AfFTase helix (Fig. 3-6a). Similarly, the loop containing CnFTase A86β and S87β (A149β and S150β in AfFTase, respectively) no longer shifts as much toward the helical axis as the loop in AfFTase. These three CnFTase residues, as a result, do not move as far away from the X residue as the corresponding AfFTase residues (Fig. 3-5). Despite the fact that rFTase P152β (S202β in AfFTase) is conserved in CnFTase as P142β, the X residue specificity of all three orthologs is expected to be different.

**Structural differences in the product exit groove**

Structures of AfFTase binary complex and rFTase displaced product complex were superimposed in order to locate the product exit groove, which serves as a major determinant of mammalian protein prenyltransferase turnover (Lane & Beese, 2006). Compared to CaaX binding pocket, the product exit groove exhibits more sequence divergence and structural variation between the two orthologs. Several residues that constitute one wall of the putative AfFTase exit groove (S143β, F144β, A146β, and L147β) are located on the same 3_10 helix which, as mentioned in the previous section, rotates ~25° clockwise compared to that in rFTase (Fig. 3-6a).
Figure 3-6: (a) Superposition of AfFTase binary complex (green) with CnFTase ternary complex (orange) and rFTase displaced product complex (yellow) containing a farnesylated peptide substrate. The isoprenoid moiety of farnesylated product is shown in red and the CaaX moiety in gray. Residues that are believed to interact with the displaced isoprenoid are shown as sticks. Color labeling matches the corresponding protein ortholog. (b) Different view of the superposition. Isoprenoid moiety is oriented toward the reader.

This helical rotation causes a Cα movement of up to ~5 Å away from the displaced isoprenoid moiety and appears to widen the exit groove, contrary to what has been observed in C. albicans GGTase-I (Hast & Beese, 2008); in actuality, the AfFTase exit groove is narrowed by two non-conserved substitutions: H428β (S357β in rat) and F144β (Y93β in rat). Due to the helical rotation, F144β no longer occupies the same space as rFTase Y93β but instead moves into the vicinity of rFTase A92β (Fig. 3-6b). Even though F144β moves ~3 Å away from the displaced isoprenoid, the size of F144β and H428β side chains counteracts that shift and constricts the groove to 6.83 Å at its narrowest.
point, potentially forming a minor steric clash with the third isoprene group of a
displaced isoprenoid moiety modeled in a position similar to that observed in rFTase.

On the other hand, a Cα shift of AfFTase A146β (C95β in rat) ~1.3 Å away from the
second displaced isoprene widens the exit groove (Fig. 3-6a) and might allow the
displaced isoprenoid moiety to adopt a different conformation in order to avoid the
potential steric clash with H428β and F144β. It remains unclear what effect the helical
rotation and H428β substitution may have on the kinetics of AfFTase product release.

Sequence divergence and structural variation observed in the CnFTase product
exit groove are slightly different compared to both mammalian FTase (Hast et al., 2011)
and AfFTase. As mentioned in the previous section, the α helix rotates clockwise to a
lesser degree in CnFTase than in AfFTase (Fig. 3-6a), so the exit groove in CnFTase
appears to be wider than in rFTase but narrower than in AfFTase. Upon closer
inspection, one end of the CnFTase exit groove is narrowed by two non-conserved
substitutions, K80β and R405β (A92β and S357β in rat, respectively), and their side
chains limit the groove to ~3.7 Å at its narrowest point.

The position of this bottleneck relative to the third displaced isoprene, however,
is different from that in both rFTase and AfFTase. Assuming an equivalent binding
mode, the third displaced isoprene is flanked by A92β and S357β in rFTase or F144β and
H428β in AfFTase. Side chains of K80β and R405β in CnFTase converge on one side of
the isoprene instead and do not directly sandwich it, allowing the exit groove to
accommodate a displaced isoprenoid moiety modeled in a position similar to that observed in rFTase (Fig. 3-6b). Despite potential minor steric clash between the third displaced isoprene and CnFTase R405β, the overall binding conformation of displaced isoprenoid moiety in rFTase and CnFTase is likely to be similar, but the kinetics of product release might be different between the two orthologs (Hast et al., 2011).

**Aspergillus fumigatus FTase X residue specificity**

The polar X residue binding pocket of mammalian FTase tends to exclude bulky residues such as tyrosine and tryptophan but accommodate methionine, glutamine, and serine. Alanine, threonine, and cysteine are also accepted to a lesser extent (Reid et al., 2004). From the structural analysis of *A. fumigatus* FTase CaaX binding pocket, a combination of helical rotation, loop movement, and single non-conserved amino acid substitution is likely to alter the X residue specificity of the fungal enzyme. The aim of this study is to qualitatively characterize the sequence preference of AfFTase at the X position. It is done to test the specificity hypothesis derived from comparison of residues around the X residue binding pockets in mammalian and fungal FTases. The study is also a response to the lack of success in both cocrystallizing and soaking in peptides and a step toward substrate optimization.

Substrate specificity was explored by determining the level of Ras farnesylation using a fluorescent FPP analog. Serving as peptide substrate, the Ras proteins used in the assay all contained a C-terminal CVVX motif and differed only in their terminal X
residue (C, E, F, G, H, I, K, L, M, N, P, Q, R, S, T, V, and Y). Ras A, D, and W were not included in the assay. Fluorescence on the farnesylated Ras was detected by a UV transillumination imaging system after the protein had been separated from both FTase subunits on a SDS-PAGE gel.

Based on undetectable levels of fluorescence signal (Fig. 3-7), Ras F and H appeared to be non-substrates for A. fumigatus FTase. I, N, Q, S, T, and V were farnesylated to a certain extent, but surprisingly, higher Ras concentration seemed to have an inhibitory effect on farnesylation in many cases. L and M turned out to be good substrates, and the assay was repeated for these two proteins at additional concentrations in order to better clarify the inhibition seen in other substrates. The same effect was observed for Ras L as it approached 20 μM, but up to 10 μM, farnesylation of M was directly proportional to its concentration.

One possible explanation for the inverse relationship between Ras concentration and farnesylation level is the dilution required to bring protein substrates to their different concentrations. Ras is stored in a buffer containing 20 mM HEPES pH 7.5, 500 mM NaCl, 5% glycerol, 5 μM GDP, and 5 mM DTT, while the reaction buffer used for diluting Ras contains 20 mM Tris pH 7.7, 150 mM NaCl, 10 mM MgCl₂, 5 μM ZnCl₂, 0.04% dodecylmaltoside, and 2 mM DTT. Serial dilution of Ras to various concentrations using reaction buffer means that the amount of reaction buffer components will be proportional to dilution volume and inversely related to final Ras
Figure 3-7: AfFTase and rFTase X residue selectivity assay
concentration. With the premise of maximizing dilution volume as a way to optimize farnesylation activity, low Ras stock concentration could potentially be detrimental to the reaction. To work around this issue, Ras L, M, N, Q, S, and V were overexpressed, purified, and concentrated to a high enough level so that the volume of Ras would be small relative to that of reaction buffer when diluted. In addition, a minor adjustment was made to the Ras buffer (20 mM Tris pH 7.7, 500 mM NaCl, 5% glycerol, 10 mM MgCl$_2$, 5 μM ZnCl$_2$, 5 μM GDP, and 5 mM DTT) so that it retained key components of the reaction buffer without compromising the quality of Ras in purification and storage. Results of the revised assay are shown in Fig. 3-8.

Based on the intensity of fluorescence signals, amino acid selectivity of AfFTase at the X position can be classified into three categories: M and Q as robust substrates, I, L, N, S, T, and V as weak substrates, and C, E, F, G, H, K, P, R, and Y as non-substrates. Changes made to selected Ras proteins, as mentioned in the previous paragraph, also eliminated the issue of farnesylation level being inversely proportional to substrate concentration. The assay, when repeated for rat FTase using Ras M, I, and R, showed that only M and I behaved as substrates (Fig. 3-8d), consistent with published studies (Reiss et al., 1991; Moores et al., 1991; Roskoski & Ritchie, 1998; Reid et al., 2004).

One observation worthy of note is the lack of farnesylation level-substrate concentration inverse relationship in rat and *C. neoformans* FTases even without Ras concentration and buffer revisions. In other words, unlike AfFTase, rat and *C. neoformans* FTases show a consistent behavior regardless of substrate concentration. This difference highlights the unique characteristics of these enzymes in terms of their substrate specificity and the response to changes in experimental conditions.
Figure 3-8: X residue selectivity of AfFTase (a)-(c) and rFTase (d).
*C. neoformans* FTases do not rely as heavily on the reaction buffer components for optimal activity. Magnesium and zinc ions are likely to be the necessary components for AfFTase activity, but the reason for its greater dependency on these divalent cations compared to the other two orthologs is unclear.

Table 3-2: X residue selectivity of rat, *C. neoformans*, and *A. fumigatus* FTases

<table>
<thead>
<tr>
<th>X residue</th>
<th>rFTase activity</th>
<th>CnFTase activity</th>
<th>AfFTase activity</th>
</tr>
</thead>
<tbody>
<tr>
<td>M</td>
<td>+</td>
<td>+</td>
<td>++</td>
</tr>
<tr>
<td>Q</td>
<td>+</td>
<td>+</td>
<td>++</td>
</tr>
<tr>
<td>I</td>
<td>+</td>
<td>+</td>
<td>+</td>
</tr>
<tr>
<td>V</td>
<td>+</td>
<td>+</td>
<td>+</td>
</tr>
<tr>
<td>N</td>
<td>+</td>
<td>+</td>
<td>+</td>
</tr>
<tr>
<td>L</td>
<td>-</td>
<td>+</td>
<td>+</td>
</tr>
<tr>
<td>H</td>
<td>+</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>F</td>
<td>+</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>C</td>
<td>+</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>S</td>
<td>+</td>
<td>-</td>
<td>+</td>
</tr>
<tr>
<td>T</td>
<td>+</td>
<td>-</td>
<td>+</td>
</tr>
<tr>
<td>Y</td>
<td>-</td>
<td>+</td>
<td>-</td>
</tr>
<tr>
<td>P</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>G</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>K</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>R</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>E</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
</tbody>
</table>

Table 3-2 summarizes the peptide substrate selectivity of rat (Hast et al., 2011), *C. neoformans* (Hast et al., 2011), and *A. fumigatus* FTases at the X position. M, Q, I, V, and N function as substrates across all three orthologs, while P, G, K, R, and E remain as
non-substrates. C, F, and H are farnesylated only by rat FTase and Y by CnFTase, but S and T appear to be substrates for both rat and \textit{A. fumigatus} FTases. L, a common mammalian GGTase-I substrate (Moores et al., 1991; Roskoski & Ritchie, 1998; Reid et al., 2004; Hartman et al., 2005), is prenylated by both fungal FTases.

As mentioned in the previous section, there are two amino acid substitutions in the X residue binding pockets of the three FTase orthologs: rat FTase A151β (A201β in \textit{Af}FTase) to CnFTase L141β and rat FTase P152β (P142β in CnFTase) to AfFTase S202β. These non-conserved substitutions, together with helical rotation and loop movement observed in fungal FTases, are likely to account for X residue specificity differences listed in Table 3-2. Like Phe, Leu and His are expected to bind in the same hydrophobic cavity adjacent to the X residue specificity pocket (Reid et al., 2004), implying that structural features of this cavity contribute to X residue selectivity in the three orthologs.

With the premise of a fixed CaaX backbone conformation, modeling studies were subsequently done to further examine the effect of sequence and structural variations on peptide substrate selectivity at the X position. By preserving the binding mode of TC21 KKSHTKCVIF peptide from rFTase ternary complex (PDB code 1TN7), Phe would come into steric clash with S150β, W153β, and R252β when modeled into the AfFTase hydrophobic cavity (Fig. 3-9a). This clash is due to moment of the loop in which S150β is located, clockwise rotation of W153β along its β-γ axis, and the side chain
conformation of R252β. Clashes with other residues in the binding pocket cannot be avoided despite different Phe rotamers.

As seen in Fig. 3-9a, AfFTase A313β pushes the second and third isoprenes of FPP toward the CaaX binding pocket and into potential steric conflict with the modeled α2 residue, even if that Ile is mutated to Val. The TKCVVM peptide from CnFTase ternary complex (PDB code 3Q75) was also modeled into the AfFTase peptide binding pocket and appeared to be a more appropriate choice than KKSKTKCVIF due to a slight bend in its CaaX backbone away from the third isoprene, minimizing steric clash between the α2 residue and FPP (Fig. 3-9b). When the TKCVVM Met was mutated to Phe (Fig. 3-9b) or His (Fig. 3-9c) with the same side chain conformation as that of Phe in KKSHTKCVIF, the X residue would avoid steric overlap with W153β and R252β but come into further clash with S150β, and clashes with other residues in the binding pocket would persist despite modeling different rotamers. The inability of AfFTase to accommodate a modeled Phe or His at the X position is hence consistent with selectivity results presented in Table 3-2.

Even though AfFTase is able to farnesylate Ras containing a C-terminal CVVL motif, Leu cannot be modeled into the secondary binding site without steric clash (Fig. 3-9e). When modeled into the X residue binding pocket, it comes into minor steric overlap with A201β (Fig. 3-9f), but this conflict can likely be avoided if the CaaX motif shifts further away from the third isoprene in order to completely eliminate clash.
Figure 3-9: CVVX peptide modeled into CaaX binding pocket of AfFTase binary complex (green), which is superimposed with CnFTase ternary complex (orange) and rFTase ternary complex (yellow). Steric clashes are represented by red spikes. (a) CaaX binding conformation (yellow) derived from KKSKTKCVIF of rFTase ternary complex. (b)-(f) CaaX binding conformation (orange) derived from TKCVVM of CnFTase ternary complex. (b) Arrow indicates a severe steric clash unable to be calculated by Probe.
between the a2 residue and FPP. As illustrated in Fig. 3-9d, the X residue binding pocket is able to accommodate a modeled Cys, so selectivity of the enzyme against Cys, unlike that for Leu, cannot be adequately explained by modeling.

In addition to modeling studies, occurrence frequency of each amino acid at the X position of CaaX sequences should somewhat reflect the X residue selectivity of protein prenyltransferases. Table 3-3 lists 112 potential prenylated proteins in the *Aspergillus fumigatus* genome that are identified through a PROSITE (Sigrist et al., 2010) search and the frequency of occurrence of each amino acid at the X position. While these proteins all contain a C-terminal CXXX motif (where X is any amino acid), there is no guarantee that they are all prenylated, and the frequencies of occurrence may therefore be skewed. Since 5% occurrence per amino acid is an even distribution, this percentage serves as a rough determinant of whether or not the occurrence frequency is high enough to imply selectivity for any particular amino acid.

**Table 3-3: Potential *A. fumigatus* CaaX prenyltransferase substrates**

<table>
<thead>
<tr>
<th>X residue</th>
<th>Frequency of occurrence (%)</th>
<th>Name</th>
<th>a1</th>
<th>a2</th>
</tr>
</thead>
<tbody>
<tr>
<td>M</td>
<td>13.4</td>
<td>Peptidyl-prolyl cis-trans isomerase H</td>
<td>G</td>
<td>E</td>
</tr>
<tr>
<td></td>
<td></td>
<td>WD repeat-containing protein</td>
<td>L</td>
<td>W</td>
</tr>
<tr>
<td></td>
<td></td>
<td>G-protein complex γ subunit Ste18/GpgA</td>
<td>T</td>
<td>V</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Smr domain protein</td>
<td>T</td>
<td>I</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Secretion pathway protein Sls2/Rcy1</td>
<td>N</td>
<td>I</td>
</tr>
<tr>
<td></td>
<td></td>
<td>F-box domain protein</td>
<td>T</td>
<td>V</td>
</tr>
<tr>
<td></td>
<td></td>
<td>LaeA-like methyltransferase</td>
<td>T</td>
<td>I</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Glycosyl hydrolase</td>
<td>S</td>
<td>V</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Ras small monomeric GTPase RasA</td>
<td>V</td>
<td>I</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Rheb small monomeric GTPase RhbA</td>
<td>I</td>
<td>L</td>
</tr>
<tr>
<td>Protein Name</td>
<td>Symbol</td>
<td>Score</td>
<td></td>
<td></td>
</tr>
<tr>
<td>--------------------------------------------------</td>
<td>--------</td>
<td>-------</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Ubiquitin C-terminal hydrolase family protein</td>
<td>I</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>SNARE protein Ykt6</td>
<td>V</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Chitin synthase activator (Chs3)</strong></td>
<td>V</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Ras homolog enriched in brain</td>
<td>L</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>DUF455 domain protein</td>
<td>V</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Protein mitochondrial targeting protein (Mas1)</td>
<td>A</td>
<td>S</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Peptide methionine sulfoxide reductase</td>
<td>P</td>
<td>A</td>
<td></td>
<td></td>
</tr>
<tr>
<td>DnaJ domain protein (Mas5)</td>
<td>A</td>
<td>P</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Peroxisomal membrane protein receptor Pex19</td>
<td>P</td>
<td>Q</td>
<td></td>
<td></td>
</tr>
<tr>
<td>SNF2 family helicase/ATPase</td>
<td>K</td>
<td>T</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Q 4.5</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Probable β-glucosidase F</td>
<td>R</td>
<td>V</td>
<td></td>
<td></td>
</tr>
<tr>
<td>G-protein complex α subunit GpaA/FadA</td>
<td>G</td>
<td>L</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Blue light-inducible protein Bli-3</td>
<td>K</td>
<td>S</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Golgi integral membrane protein (Cln3)</td>
<td>R</td>
<td>R</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Phosphoesterase</td>
<td>V</td>
<td>D</td>
<td></td>
<td></td>
</tr>
<tr>
<td>LysM domain protein</td>
<td>V</td>
<td>G</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Squalene-hopene-cyclase</td>
<td>L</td>
<td>A</td>
<td></td>
<td></td>
</tr>
<tr>
<td>V 5.4</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Ctr copper transporter family protein</td>
<td>V</td>
<td>L</td>
<td></td>
<td></td>
</tr>
<tr>
<td>ATP synthase subunit ATP9</td>
<td>K</td>
<td>Y</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Amino acid transporter</td>
<td>H</td>
<td>M</td>
<td></td>
<td></td>
</tr>
<tr>
<td>RPEL repeat protein</td>
<td>P</td>
<td>A</td>
<td></td>
<td></td>
</tr>
<tr>
<td>C6 transcription factor</td>
<td>I</td>
<td>S</td>
<td></td>
<td></td>
</tr>
<tr>
<td>LysM domain protein</td>
<td>V</td>
<td>G</td>
<td></td>
<td></td>
</tr>
<tr>
<td>N 6.3</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Tripeptidyl-peptidase sed5 (serine protease)</td>
<td>F</td>
<td>R</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Glutamine synthetase</td>
<td>G</td>
<td>G</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Pectin methylesterase</td>
<td>P</td>
<td>T</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Lipase</td>
<td>S</td>
<td>T</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Defensin domain protein</td>
<td>T</td>
<td>C</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Integral membrane protein (Pth11)</td>
<td>G</td>
<td>W</td>
<td></td>
<td></td>
</tr>
<tr>
<td>25D9-11</td>
<td>D</td>
<td>L</td>
<td></td>
<td></td>
</tr>
<tr>
<td>L 13.4</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Nuclear mRNA splicing factor</td>
<td>K</td>
<td>G</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Isocitrate dehydrogenase, NAD-dependent</td>
<td>K</td>
<td>R</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Primary amine oxidase</td>
<td>C</td>
<td>K</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Peptidyl-prolyl cis-trans isomerase</td>
<td>G</td>
<td>E</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Rho GTPase ModA</td>
<td>V</td>
<td>L</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Ras small monomeric GTPase RasB</td>
<td>V</td>
<td>I</td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Rho GTPase Rho1</strong></td>
<td>T</td>
<td>I</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Ras GTPase Rho2</td>
<td>I</td>
<td>V</td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Rho GTPase Rho3</strong></td>
<td>I</td>
<td>I</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Ras GTPase Rho4</td>
<td>K</td>
<td>I</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Ras GTPase Rac</td>
<td>T</td>
<td>I</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Ras small monomeric GTPase (Rsr1)</td>
<td>V</td>
<td>I</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Salicylate hydroxylase</td>
<td>S</td>
<td>M</td>
<td></td>
<td></td>
</tr>
<tr>
<td>3-oxo-5-alpha-steroid 4-dehydrogenase family protein</td>
<td>P</td>
<td>W</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Hydrophobin</td>
<td>S</td>
<td>K</td>
<td></td>
<td></td>
</tr>
<tr>
<td>------------</td>
<td>----</td>
<td>----</td>
<td></td>
<td></td>
</tr>
<tr>
<td>H</td>
<td>3.6</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>β-1,3-glucanosyltransferase Bgt1</td>
<td>P</td>
<td>K</td>
<td></td>
<td></td>
</tr>
<tr>
<td>DNA (apurinic or apyrimidinic site) lyase</td>
<td>Q</td>
<td>S</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Thermoresistant gluconokinase family protein</td>
<td>A</td>
<td>D</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Flavin-containing monoxygenase</td>
<td>D</td>
<td>G</td>
<td></td>
<td></td>
</tr>
<tr>
<td>F</td>
<td>1.8</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Isoamyl alcohol oxidase</td>
<td>I</td>
<td>A</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Inositol kinase kinase (UvsB)</td>
<td>A</td>
<td>F</td>
<td></td>
<td></td>
</tr>
<tr>
<td>C</td>
<td>6.3</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Splicing factor 3B subunit 10 (SF3b10)</td>
<td>S</td>
<td>G</td>
<td></td>
<td></td>
</tr>
<tr>
<td>6-hexanolactone hydrolase</td>
<td>L</td>
<td>G</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Yippee zinc-binding protein Moh1</td>
<td>V</td>
<td>V</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Rho-like small GTPase</td>
<td>V</td>
<td>V</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Class II aldolase/adducin domain protein</td>
<td>F</td>
<td>P</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Endo-1,4-β-xylanase</td>
<td>T</td>
<td>V</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Esterase/lipase/thioesterase family protein</td>
<td>L</td>
<td>G</td>
<td></td>
<td></td>
</tr>
<tr>
<td>MFS transporter (Accession# EDP54064)</td>
<td>G</td>
<td>K</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Quinone oxidoreductase</td>
<td>T</td>
<td>V</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Yip1 domain protein</td>
<td>F</td>
<td>I</td>
<td></td>
<td></td>
</tr>
<tr>
<td>MFS transporter (Accession# EDP54365)</td>
<td>S</td>
<td>D</td>
<td></td>
<td></td>
</tr>
<tr>
<td>NADH-ubiquinone oxidoreductase 64 kDa subunit</td>
<td>S</td>
<td>F</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Alpha/beta superfamily hydrolase</td>
<td>T</td>
<td>V</td>
<td></td>
<td></td>
</tr>
<tr>
<td>C6 transcription factor</td>
<td>K</td>
<td>I</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Cell wall protein</td>
<td>T</td>
<td>C</td>
<td></td>
<td></td>
</tr>
<tr>
<td>T</td>
<td>1.8</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>BZIP transcription factor</td>
<td>L</td>
<td>A</td>
<td></td>
<td></td>
</tr>
<tr>
<td>C6 finger domain protein</td>
<td>G</td>
<td>S</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Y</td>
<td>1.8</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Ankyrin repeat protein</td>
<td>A</td>
<td>L</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Vegetative incompatibility WD repeat protein</td>
<td>W</td>
<td>R</td>
<td></td>
<td></td>
</tr>
<tr>
<td>P</td>
<td>6.3</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>ATP dependent RNA helicase</td>
<td>L</td>
<td>P</td>
<td></td>
<td></td>
</tr>
<tr>
<td>GPI anchored protein</td>
<td>L</td>
<td>A</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Extracellular chitosanase CsnC</td>
<td>T</td>
<td>A</td>
<td></td>
<td></td>
</tr>
<tr>
<td>4-hydroxybenzoate polyprenyl transferase</td>
<td>V</td>
<td>K</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Aldehyde dehydrogenase AldA</td>
<td>P</td>
<td>F</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Protein tyrosine phosphatase (Pyp1)</td>
<td>E</td>
<td>N</td>
<td></td>
<td></td>
</tr>
<tr>
<td>DUF221 membrane protein</td>
<td>K</td>
<td>S</td>
<td></td>
<td></td>
</tr>
<tr>
<td>G</td>
<td>1.8</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Integral membrane protein</td>
<td>Y</td>
<td>Q</td>
<td></td>
<td></td>
</tr>
<tr>
<td>C6 transcription factor (Ctf1B)</td>
<td>L</td>
<td>S</td>
<td></td>
<td></td>
</tr>
<tr>
<td>K</td>
<td>3.6</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Kynureninase 1</td>
<td>G</td>
<td>R</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Sugar hydrolase</td>
<td>I</td>
<td>H</td>
<td></td>
<td></td>
</tr>
<tr>
<td>4'-phosphopantetheinyl transferase NpgA</td>
<td>T</td>
<td>K</td>
<td></td>
<td></td>
</tr>
<tr>
<td>DUF647 domain protein</td>
<td>I</td>
<td>V</td>
<td></td>
<td></td>
</tr>
<tr>
<td>R</td>
<td>3.6</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>EF hand domain protein</td>
<td>K</td>
<td>R</td>
<td></td>
<td></td>
</tr>
<tr>
<td>MFS monocarboxylate transporter</td>
<td>V</td>
<td>S</td>
<td></td>
<td></td>
</tr>
<tr>
<td>40S ribosomal protein S8</td>
<td>R</td>
<td>V</td>
<td></td>
<td></td>
</tr>
<tr>
<td>PKS-like enzyme</td>
<td>A</td>
<td>K</td>
<td></td>
<td></td>
</tr>
<tr>
<td>E</td>
<td>3.6</td>
<td>Glucokinase</td>
<td>Q</td>
<td>S</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Class II aldolase and adducin N-terminal domain protein</td>
<td>K</td>
<td>G</td>
</tr>
<tr>
<td></td>
<td></td>
<td>F-box domain and ankyrin repeat protein</td>
<td>F</td>
<td>Q</td>
</tr>
<tr>
<td></td>
<td></td>
<td>MFS transporter (Seo1)</td>
<td>V</td>
<td>D</td>
</tr>
<tr>
<td>A</td>
<td>4.5</td>
<td>Probable endo-β-1,4-glucanase D</td>
<td>V</td>
<td>S</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Integral membrane protein Pth11-like</td>
<td>T</td>
<td>K</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Ser/Thr protein phosphatase</td>
<td>E</td>
<td>D</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Oxidoreductase, short chain dehydrogenase/reductase family</td>
<td>A</td>
<td>S</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Ubiquitin conjugating enzyme</td>
<td>S</td>
<td>V</td>
</tr>
<tr>
<td>W</td>
<td>3.6</td>
<td>TOR Pathway phosphatidylinositol 3-kinase TorA</td>
<td>S</td>
<td>F</td>
</tr>
<tr>
<td></td>
<td></td>
<td>GAF domain nucleotide-binding protein</td>
<td>C</td>
<td>D</td>
</tr>
<tr>
<td></td>
<td></td>
<td>O-methyltransferase</td>
<td>S</td>
<td>V</td>
</tr>
<tr>
<td></td>
<td></td>
<td>C6 finger domain protein</td>
<td>L</td>
<td>S</td>
</tr>
<tr>
<td>D</td>
<td>1.8</td>
<td>Endoglucanase</td>
<td>Y</td>
<td>R</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Iron-sulfur cluster biosynthesis protein Isd11</td>
<td>R</td>
<td>W</td>
</tr>
</tbody>
</table>

A comparison of Table 3-2 and 3-3 shows that the X residue selectivity of AfFTase correlates well with occurrence frequencies of some amino acids but not others. Selectivity for M, Q, I, V, N, L, and S can be accounted for by their occurrence frequencies of approximately 5% or, in most cases, higher. Selectivity against H, F, Y, G, K, R, and E can similarly be rationalized by their < 5% occurrence. Even though T serves as a weak substrate for AfFTase, its 1.8% occurrence in all C-terminal CXXX motifs of the *A. fumigatus* genome is surprisingly low. On the contrary, occurrence frequencies of C and P, which are non-substrates, are higher than expected. While Ras A, D, and W are not included in the specificity assay, AfFTase is likely to select for A but not D and W based on the frequencies of these amino acids.
Even though L is farnesylated by AfFTase (Table 3-2) and has 13.4% occurrence, several of the proteins listed in Table 3-3 with a C-terminal CXXL motif are likely to be AfGGTase-I substrates instead, since many members of the Rho small GTPase family are known to be prenylated by mammalian GGTase-I (Reid et al., 2004). The ability of AfGGTase-I to prenylate Ras L has also been observed (data not shown) using a fluorescent GGPP analog in a qualitative assay similar to that described at the beginning of this section. Despite the report that RhoB (CKVL) serves as a substrate for both mammalian FTase and GGTase-I (Baron et al., 2000), AfFTase and GGTase-I appear to demonstrate a stronger overall cross-selectivity for a C-terminal Leu.

The list of potential *A. fumigatus* CaaX prenyltransferase substrates in Table 3-3 may also be useful for identifying pathways in antifungal development. As discussed in Chapters 1 and 2, some of these proteins (highlighted in bold in Table 3-3) are involved in fungal biosyntheses, pathogenicity, and efflux mechanisms. With the exception of β-1,3-glucanosyltransferases Bgt1p, the majority of them contain X residues that can be prenylated by AfFTase or GGTase-I, but it remains unclear whether they all function as prenylated substrates.

**Crystallization of *A. fumigatus* FTase ternary complex**

Without a crystal structure of the AfFTase ternary complex, the molecular basis of AfFTase X residue selectivity presented in this chapter is based on modeling studies that serve as a speculative extension of the binary complex structure. Similarly,
discussion of the exit groove relies on structures of the displaced product complexes of other FTase orthologs. A crystal structure of AfFTase ternary complex with Leu at the X position is of particular interest due to no available structures of FTase in complex with Leu-terminating peptides as well as the selectivity of AfFTase for a C-terminal Leu (Table 3-2).

In addition to peptides with Ala and Ser at the X position, several 11-mers previously used for the crystallization of mammalian FTase complexes were chosen for crystallization screening of AfFTase ternary and product complexes, including K-Ras4B (KKSKTKCVIM), Rap2a (DDPTASACNIQ), C. neoformans Ras1 (DKGCCRGCVVL), and its derivative (DKGCCRGCVVM). Most of the X residues of these peptides had been shown to be farnesylated by AfFTase (Table 3-2), and furthermore, sequences of DKGCCRGCVVL and DKGCCRGCVVM were identical to the last 11 residues of Ras L and M, respectively, from the selectivity assay. The peptide sequences used for crystallization screening of AfFTase complexes were therefore reasonable choices.

To screen for crystals of the ternary complex, peptides were either co-crystallized with AfFTase and FPT-II or soaked into crystals of the enzyme in complex with the FPP analog. Peptide concentration and length of soak were both varied in an attempt to capture a peptide substrate in the active site, and while FPT-II and Zn\(^{2+}\) were consistently observed in the active site, the peptide binding pocket remained empty in
Since the likelihood of peptide sequence as a contributor to weak or non-binding was low, peptide length was evaluated as another possible factor.

Structural comparison of ternary complexes indicates that although the CaaX motifs of peptides adopt similar extended conformations, upstream sequences can be conformationally diverse (Reid et al., 2004). Although a CaaX motif can be modeled into the AfFTase peptide binding pocket with minimal steric clash (Fig. 3-9f), the upstream portion of the 11-mers used for crystallization screening might come into steric conflict with other parts of the enzyme. In order to evaluate this hypothesis, two peptides used for crystallization screening, KKKSKTKCVIM and DDPTASACNIQ, were modeled into the peptide binding pocket of AfFTase binary complex (Fig. 3-10) based on their conformations in rat (PDB code 1D8D) and human FTases (PDB code 1TN6), respectively. In the case of KKKSKTKCVIM (Fig. 3-10a), the second Lys could potentially clash with side chain of AfFTase H105α (Q162α in rFTase), and because of the clockwise helical rotation discussed in previous sections, the first Lys could come into steric conflict with S143β (A92β in rFTase).

The AfFTase crystal packing also brings residues 270-282 (colored blue in Fig. 3-10) of a symmetry related molecule of the β subunit near the exit groove of the original molecule where the N-terminal portion of a bound peptide is likely to be located. Pressing against the peptide binding pocket, these residues form a small helix that is absent in mammalian and C. neoformans FTases and will likely encounter steric overlap
Figure 3-10: KKKSKTKCVIM (a) and DDPTASACNIQ (b) 11-mer peptides modeled into the peptide binding pocket of AfFTase binary complex (green), which is superimposed with rFTase (yellow) and human FTase (yellow) ternary complexes, respectively. Parts of the symmetry related molecules of AfFTase which superimpose well with that of the mammalian enzymes (red) are in gray, while structurally divergent parts are in blue.
with KKS of KKKSKTKCVIM (Fig. 3-10a) or the second Asp of DDPTASACNIQ (Fig. 3-10b). The same residue of the latter 11-mer can also potentially clash with H428β (S857β in rFTase).

The steric clashes described above have a collective effect of excluding peptides of certain lengths and offer a possible explanation for the unsuccessful crystallization of ternary complexes. Since clashes are limited to the N-terminal region of modeled peptides, further crystallization screening of AfFTase ternary and product complexes should be done using shorter peptides. Given the success of using Rap2a and CnRas1 sequences in AfFTase X residue selectivity assay and 6-mers in crystallizing various ternary complexes (Reid et al., 2004; Hast et al., 2011), peptides such as TKCVVM (K-Ras4B derivative), SACNIQ (Rap2a), and RGCVVL (CnRas1 derivative) should be heavily considered for future experiments. A cocrystal structure of AfFTase in complex with FPT-II and KCVVM peptide has been obtained recently.

**Optimization of Aspergillus fumigatus GGTase-I expression**

With its vector available, *A. fumigatus* GGTase-I was overexpressed in *E. coli* and purified for crystallization trials. About 0.5 mg of purified protein was obtained from a 10L growth. The aim of this study is to improve the overexpression of this protein for crystallization screening and optimization. An expression test was done on BL21 Star, C41, and C43 cell lines at 18 and 37°C. Protein was purified after cell lysis, and its expression level was determined by the intensity of bands on a Coomassie-stained SDS-
PAGE gel. Based on the results, the protein was overexpressed in BL21 Star at 18°C. The yield improved to ~16 mg from a 10L growth.

**Summary**

In addition to data redundancy, resolution of the *A. fumigatus* FTase binary complex crystal structure has been improved from 1.8 Å to 1.5 Å, matching the highest ever obtained for a published protein prenyltransferase crystal structure. This chapter describes the overall architecture of AfFTase complexed with the isoprenoid substrate FPP and structural features of the isoprenoid and peptide binding sites and product exit groove. Despite an overall architectural similarity between mammalian and *A. fumigatus* FTases, some parts of β subunit that are distant from the active site contain significant structural divergence.

Compared to that in the mammalian ortholog, binding of FPP in AfFTase adopts a more linearly extended conformation. A313β pushes the second and third isoprenes toward the CaaX binding pocket, which will likely affect positioning of the a₂ residue. Shape of the X residue binding pocket is altered by a helical rotation, loop movement, and perhaps even the side chain conformation of R252β; at the same time, polarity of binding surface is likely affected by a rFTase P152β to AfFTase S202β substitution. Taken together, the X residue specificity of the two protein orthologs is expected to be somewhat different.
The same helical rotation that changes the shape of X residue binding pocket also appears to widen the exit groove, but the channel is actually narrowed by two non-conserved substitutions. The displaced isoprenoid moiety might be able to adopt a different conformation in order to avoid potential steric clash with these two residues, so the overall effect on the kinetics of AfFTase product release is unclear.

X residue selectivity of AfFTase has been qualitatively characterized mainly as a response to the lack of success in crystallizing a ternary complex and a step toward substrate optimization. H, F, and C function as substrates for AfFTase but not rFTase, and unlike the mammalian orthologs, AfFTase and GGTase-I demonstrate cross-selectivity for a C-terminal Leu. Many of these selectivity differences can be accounted for either by modeling studies or the occurrence frequency of a particular X residue in the *A. fumigatus* genome.

With a high resolution structure of the AfFTase active site and knowledge of substrate selectivity, our attention now shifts to inhibitor design, the main topic to be discussed in Chapter 4.

**Materials and methods**

**Cloning, expression and purification of *A. fumigatus* FTase**

*A. fumigatus* FTase was overexpressed in C41(DE3) *E. coli* at 37°C using a pCDFDuet-1 vector that contained both the α and β subunits. Resuspended in 20 mM imidazole Ni-NTA buffer (1x PBS pH 7.4 and 200 mM NaCl) supplemented with
protease inhibitor and 10 mM MgCl₂, cells were lysed using a cell cracker and subsequently spun down. Washed Ni-NTA resin was added to the clarified lysate and stirred at 4°C before being used to pack a gravity column. The column was washed with 20 mM imidazole Ni-NTA buffer until no more unbound protein could be detected by mini-Bradford assay. 250 mM imidazole Ni-NTA buffer was then used to elute the protein.

Salt concentration of the eluate was adjusted slowly with three volumes of 1 M (NH₄)₂SO₄ Phenyl buffer (20 mM HEPES-NaOH pH 7.5, 5 μM ZnCl₂, and 5 mM DTT) before loaded onto a Phenyl Sepharose column (GE Healthcare). Protein was eluted from a gradient of 800 to 0 mM (NH₄)₂SO₄ Phenyl buffer. Pooled fractions were buffer exchanged into Phenyl buffer, concentrated to 10 mg/ml, and flash-frozen in liquid nitrogen.

**Crystallization**

10 mg/ml *A. fumigatus* FTase was incubated with three-fold molar excess FPP for crystallization of the binary complex or with three-fold molar excess FPT-II and KCVVM peptide for cocrystallization of the ternary complex. At 17°C, crystals with a rod-like morphology with dimensions up to 40 x 40 x 300 μm grew out of hanging drops containing a precipitating agent of 100 mM ammonium acetate pH 5.2, 500-700 mM LiCl, and 6-8% PEG 6000. Crystals were transferred to a stabilization solution composed of mother liquor with additional 50 mM LiCl and 2% PEG 6000 and then serially into
cryobuffer solutions supplemented with 6, 14, and 20% ethylene glycol for 5 minutes each before flash-cooling in liquid nitrogen. Due to the high binding constant of FPP, its presence was unnecessary during cryoprotection.

**Data collection and refinement**

Diffraction data were collected at 100 K on beamline 22-ID (SER-CAT) at the Advanced Photon Source. Crystals diffracted to ~ 1.8 Å and belonged to P2₁ space group with unit cell parameters of 63.3 x 90.5 x 83.3 Å³ (β = 110.6°). The data were reduced using *DENZO* and *SCALEPACK* (Otwinowski & Minor, 1997). Molecular replacement was carried out with the program *Phaser* (McCoy et al., 2007) using the structure of rat FTase as a search model. Crystals contained one molecule per asymmetric unit. Structures were built using *Coot* (Emsley et al., 2010), and model refinement was done in *REFMAC5* (Murshudov et al., 1997).

**FTase substrate specificity assay**

1 μM *A. fumigatus* FTase was incubated with 10 μM 3,7-dimethyl-8-(7-nitrobenzo[1,2,5]oxadiazol-4-ylamino)-octa-2,6-diene-1-pyrophosphate (NBD-GPP from Jena Bioscience) and *C. neoformans* Ras at various concentrations as indicated in Fig. 3-8. *R. norvegicus* FTase and Rap2a were included as control experiments. All reactions were carried out in 20 mM Tris pH 7.7, 150 mM NaCl, 10 mM MgCl₂, 5 μM ZnCl₂, 0.04% dodecylmaltoside, and 2 mM DTT at 37°C for 2 hours before quenched with SDS-PAGE
buffer. Samples were run on 4-20% gradient gels, and fluorescence on the farnesylated Ras was detected by a UV transillumination imaging system.

**A. fumigatus** GGTase-I expression test

Cell pellets were resuspended in B-PER Bacterial Protein Extraction Reagent (Thermo Scientific) supplemented with 1 mM CaCl$_2$, 0.1 mg/ml lysozyme (in 50 mM sodium acetate pH 4.5), and 13.3 units/ml DNase I. Protein was purified using the QIAGEN Ni-NTA Spin Kit. Samples were run on 4-20% SDS-PAGE gradient gels and Coomassie-stained.
Chapter 4. Towards inhibition of *Aspergillus fumigatus* protein farnesyltransferase

**Introduction**

Inhibitors of protein prenyltransferases are typically divided into three classes: isoprenoid-competitive, peptide-competitive, and bisubstrate analogs. Isoprenoid diphosphate analogs, such as the non-hydrolyzable FPT-II and α-hydroxyfarnesylphosphonic acid, retain a length-dependent hydrophobic group and a negatively charged moiety that mimics the diphosphate. Since FPP is used by other enzymes, not only do these compounds need to have higher FTase-binding affinity than FPP, they also need to possess high selectivity for FTase (Rowinsky et al., 1999).

Based on the finding that CaaX tetrapeptides with an aromatic α₂ residue serve as nonsubstrate FTase inhibitors, peptidomimetics have been developed as inhibitory compounds. As peptides, however, they tend to have low membrane permeability, poor cellular uptake, and are prone to proteolytic degradation. Nonpeptidomimetics, such as tipifarnib (R115777; Janssen) and L-778,123 (Merck Research Laboratories) have also been developed to bind in the CaaX binding pockets of FTase and even GGTase-I. Bisubstrate analogs that mimic both isoprenoid and CaaX substrates, despite being potent inhibitors of FTase, are not as well characterized in their mode of binding (Lane & Beese, 2006).
Besides characterizing structural features of the AfFTase active site, a second aim of the project is to explore and adapt some of the mammalian FTase inhibitors for our structural study of the \textit{A. fumigatus} enzyme. These compounds, most of which are peptidomimetics, have shown various degrees of inhibitory effects on both the mammalian and \textit{Plasmodium falciparum} FTases (Glenn et al., 2006). Structures of inhibitor complexes will provide further insight into features of the peptide binding site and a chance to reevaluate structure-activity relationship of the compounds, which has already been defined in the context of the mammalian enzyme (Hast et al., 2009). Additional characterization of the inhibitors for their IC$_{50}$ values will also reveal their selectivity between fungal and mammalian systems.

Compounds based on an ethylenediamine scaffold that inhibit both the mammalian and \textit{Plasmodium falciparum} FTases (Glenn et al., 2006; Hast et al., 2009) have been screened for their ability to bind to \textit{A. fumigatus} FTase in cocrystallization experiments. Structures of fungal FTase in complex with these compounds will provide an opportunity to reexamine structure-activity relationship of the scaffold substituents and, furthermore, a starting point for lead optimization. Modeling studies are likely to be insufficient or inaccurate due to conformational flexibility of the compounds and their multiple binding modes in the mammalian enzyme (Hast et al., 2009). Knowledge of experimental binding modes will also be instrumental in guiding molecular design.
In the present study, crystal structures of two *A. fumigatus* FTase inhibitor complexes have been determined. Chemical structures of the two ethylenediamine-scaffold inhibitors used for cocrystallization, CCB2217 and SF122, are illustrated in Fig. 4-1. Data collection and refinement statistics for the two complexes are listed in Table 4-1.

![Figure 4-1: Chemical structures of ethylenediamine-scaffold compounds](image)

Table 4-1: Data collection and refinement statistics

<table>
<thead>
<tr>
<th></th>
<th>AfFTase:FPP:CCB2217</th>
<th>AfFTase:FPP:SF122</th>
</tr>
</thead>
<tbody>
<tr>
<td>Resolution range Å</td>
<td>50.0-1.85</td>
<td>50.0-1.62</td>
</tr>
<tr>
<td>Data collection</td>
<td>ANL-APS SER-CAT</td>
<td>ANL-APS SER-CAT</td>
</tr>
<tr>
<td>Reflections, unique/total</td>
<td>70,861/285,037</td>
<td>103,825/771,465</td>
</tr>
<tr>
<td>I/σ</td>
<td>(1.9)</td>
<td>(2.5)</td>
</tr>
<tr>
<td>R&lt;sub&gt;sym&lt;/sub&gt;</td>
<td>10.1 (60.9)</td>
<td>8.4 (38.4)</td>
</tr>
</tbody>
</table>

**Unit cell**

<table>
<thead>
<tr>
<th></th>
<th>AfFTase:FPP:CCB2217</th>
<th>AfFTase:FPP:SF122</th>
</tr>
</thead>
<tbody>
<tr>
<td>Dimensions (Å)</td>
<td>a = 63.4</td>
<td>a = 63.4</td>
</tr>
<tr>
<td></td>
<td>b = 90.6</td>
<td>b = 90.4</td>
</tr>
<tr>
<td></td>
<td>c = 82.7</td>
<td>c = 83.0</td>
</tr>
<tr>
<td>Angles (°)</td>
<td>α = γ = 90; β = 110.8</td>
<td>α = γ = 90; β = 110.7</td>
</tr>
<tr>
<td>Space group</td>
<td>P2₁</td>
<td>P2₁</td>
</tr>
<tr>
<td>Completeness %</td>
<td>100 (100)</td>
<td>98.4 (86.9)</td>
</tr>
<tr>
<td>Multiplicity</td>
<td>3.8 (3.8)</td>
<td>7.1 (4.3)</td>
</tr>
</tbody>
</table>

**Refinement**

<table>
<thead>
<tr>
<th></th>
<th>AfFTase:FPP:CCB2217</th>
<th>AfFTase:FPP:SF122</th>
</tr>
</thead>
<tbody>
<tr>
<td>R&lt;sub&gt;cryst&lt;/sub&gt;</td>
<td>20.8</td>
<td>21.2</td>
</tr>
</tbody>
</table>
Structure of Aspergillus fumigatus FTase inhibitor complex

Similar to what has been observed in human FTase (Hast et al., 2009), binding of FPP and inhibitor does not induce significant conformational changes in the AfFTase active site. All of the structural features of the binary complex described in Chapter 3, such as the helix 4α-5α loop conformation (Fig. 3-4) and helical rotation (Fig. 3-6), are also present in the inhibitor complex. Binding modes of FPP in AfFTase binary and inhibitor complexes are essentially identical with a slight shift in the second and third isoprenes toward the CaaX binding pocket of the inhibitor complex. By forming a ternary complex with FPP, CCB2217 occupies both the CaaX binding pocket and part of the exit groove. Side chain conformations of residues in the active site are mostly similar with the exception of R252β, Y109α, and H148α. While H148α rotates ~30° counterclockwise along the β-γ axis, R252β moves ~0.9 Å away from the CaaX binding pocket, and Y109α rotates ~45° clockwise toward the pocket.
Inhibitor binding mode

Cocrystals of \textit{A. fumigatus} FTase with FPP and ethylenediamine scaffolds CCB2217 and SF122 were also obtained under similar conditions and had the same crystal form as the binary complex. Human FTase:SF102 complex (PDB code 3E37) was used for structural comparison due to its similarity to CCB2217 (Fig. 4-1). As illustrated in Fig. 4-2, superposition of \textit{A. fumigatus} FTase:FPP:CCB2217 and human FTase:SF102 complexes showed different inhibitor binding conformations in the two orthologs. Changes in the orientation of all four ethylenediamine scaffold substituents and a switch from bisubstrate to monosubstrate binding mode were both observed. The \textit{o}-methylphenyl sulfonamide no longer binds in the same place as the second isoprene of FPP, which allows an isoprenoid to occupy the FPP binding pocket. The \textit{N}-methylimidazole group, which coordinates the catalytic \textit{Zn}^{2+} in the human structure, rotates about 45° toward the FPP instead. The \textit{para}-benzonitrile substituent, directed at the exit groove in all the mammalian FTase inhibitor complexes, shifts 90° toward the third isoprene of FPP. The \textit{N}-Boc-piperidin-4-ylmethyl moiety that previously occupies the third isoprene is now oriented in the opposite direction at the exit groove (Hast et al., 2009).

CCB2217 contains an additional methyl group on the \textit{o}-methylphenyl sulfonamide of SF102, and residues around both the FPP and CaaX binding pockets are well conserved between \textit{A. fumigatus} and human FTases with a few subtle
Figure 4-2: Inhibitor binding modes of CCB2217 (left) and SF122 (right).

conformational differences. The different binding modes of ethylenediamine scaffold are probably due to its inherent conformational flexibility, which in turn allows for the binding of FPP. Factors contributing to the switch from monosubstrate to bisubstrate binding mode in the two orthologs, however, remain unclear.

To get a better sense of the switch in binding modes, SF102 and SF122 (retaining their conformations from PDB codes 3E37 and 3E34, respectively) were modeled in the AfFTase active site. R252β, a residue that is conserved in both rat and *C. neoformans* FTases, appears to be in steric clash with SF102 (Fig. 4-3a). A structural superposition of all three orthologs shows that the *A. fumigatus* Arg adopts a different conformation that results in its side chain being closer to the X residue binding pocket, which is now occupied by the N-Boc-piperidin-4-ylmethyl moiety of SF102. Clashes with SF122, on
the other hand, involve Q110α and W153β, two AfFTase residues with slightly different side chain conformations compared to those in the human enzyme (Fig. 4-3b).

Figure 4-3: SF102 (a) and SF122 (b) modeled in the AfFTase active site. *A. fumigatus* FTase structure in yellow, human FTase structure in green, and *C. neoformans* FTase structure in red.
**FTase inhibitor selectivity**

Cocrystallization experiments have identified two ethylenediamine compounds, CCB2217 and SF122, as potential inhibitors of *A. fumigatus* FTase. The aim of this study is to evaluate their selectivity for both mammalian and *A. fumigatus* FTases. Different inhibitor binding modes observed in the two orthologs might lead to changes in binding affinities. On the way to inhibitor design and optimization, these assays are done as a first step in characterizing the potency of two compounds in farnesylation inhibition and their cross-species reactivity.

Inhibitor specificity was determined in a fashion similar to that used for the substrate specificity assay and described in the previous section. Due to its strong fluorescence signal relative to that of Ras at equal concentrations, Rap2a was initially chosen as protein substrate. As a function of the amount of titrated ethylenediamine compound, inhibition would be inversely proportional to the level of Rap2a farnesylation. In spite of a clear dose-response inhibition, Rap2a was much better farnesylated by rat FTase than AfFTase in the absence of ethylenediamine compounds (data not shown), making it a poor substrate for inhibitor selectivity analysis. Ras CVVM was used as the protein substrate instead.

Based on the fluorescence (Fig. 4-4), both CCB2217 and SF122 inhibit Ras farnesylation by *A. fumigatus* and rat FTases. Regardless of a lack of electron density for one of the SF122 substituents in the *A. fumigatus* FTase inhibitor complex, the two
Figure 4-4: Selectivity assay with titration of CCB2217 (A) and SF122 (B).
compounds appear to be equally potent inhibitors of each enzyme. Furthermore, they also seem to be more selective for rat FTase compared to AfFTase. Ras farnesylation by rat FTase became undetectable starting at 5 μM CCB2217/SF122, but despite the dose-response inhibition, farnesylation by AfFTase was still going on at 20 μM CCB2217/SF122. This is also a more valid comparison of inhibitor selectivity because Ras CVVM looks to be more equally farnesylated by the two protein orthologs in the absence of CCB2217/SF122.

**Summary and future directions**

Crystal structures of the *A. fumigatus* FTase in complex with two ethylenediamine-scaffold compounds have been obtained. For the CCB2217 compound, changes in the orientation of all four ethylenediamine scaffold substituents and a switch from bisubstrate to monosubstrate binding mode were both observed, likely a result of its inherent conformational flexibility and also the side chain conformation of AfFTase R252β. The two ethylenediamine-scaffold compounds, while appearing to be equally potent inhibitors of Ras farnesylation by rat and *A. fumigatus* FTases, seem to be more selective for rat FTase compared to AfFTase.

By using the structure of the AfFTase inhibitor complex presented in this chapter, however, the ethylenediamine-scaffold compound CCB2217 can be modified in order to improve its potency and selectivity for the fungal ortholog. Compared to that of SF102 in human FTase, the binding mode of CCB2217 in AfFTase allows the N-Boc-
piperidin-4-ylmethyl moiety to extend much farther into the exit groove, which contains many of the structural and sequence divergences between the two FTase orthologs (Fig. 3-6). The alternative binding mode suggests the possibility of further building on this particular CCB2217 substituent to exploit the likely different steric constraints in the exit groove. If CCB2217 binds to mammalian FTase in a conformation similar to that of SF102, an extension of this substituent might also disrupt binding of CCB2217 to the mammalian enzyme and improve selectivity.

Even though no CCB2217 substituent probes the X residue binding pocket of AfFTase, a structural superposition of the AfFTase:FPP:CCB2217 and CnFTase ternary complexes indicates that the \textit{para}-benzonitrile moiety is in position for modification to exploit differences in the X residue binding pocket of mammalian and fungal orthologs. Additions to the 2-position of the \textit{para}-benzonitrile substituent will likely extend the compound into the X residue binding pocket. The conformation of this new modification, however, will in turn depend on the steric hindrance imposed by the nitrile group at the 1-position.

Given that both CCB2217 and SF122 appear to be more potent inhibitors of mammalian FTase and the observation that Zn$^{2+}$ is coordinated by N-methylimidazole group in all mammalian FTases, zinc coordination by the same substituent of CCB2217 may be another option in improving the binding affinity and potency of this compound for AfFTase. Evaluation of structure-activity relationship of ethylenediamine-scaffold
compounds in the context of this fungal enzyme is limited by the partial and complete lack of electron density of the N-Boc-piperidin-4-ylmethyl and biphenyl moieties of SF122, respectively. Ultimately, structures of AfFTase in complex with additional compounds, such as the nonpeptidomimetic mammalian FTase inhibitor tipifarnib, would be instrumental in defining the structural determinants of inhibition. A cocrystal structure of AfFTase in complex with FPT-II and tipifarnib has been obtained recently.

**Materials and methods**

**Crystallization**

10 mg/ml *A. fumigatus* FTase was incubated with three-fold molar excess FPP for crystallization of the binary complex. At 17°C, crystals with a rod-like morphology with dimensions up to 40 x 40 x 300 μm grew out of hanging drops containing a precipitating agent of 100 mM ammonium acetate pH 5.2, 500-700 mM LiCl, and 6-8% PEG 6000. Crystals were transferred to a stabilization solution composed of mother liquor with additional 50 mM LiCl and 2% PEG 6000 and then serially into cryobuffer solutions supplemented with 6, 14, and 20% ethylene glycol for 5 minutes each before flash-cooling in liquid nitrogen. Due to the high binding constant of FPP, its presence was unnecessary during cryoprotection.

Cocrystallization and cryoprotection of FTase with FPP and inhibitor were similar to that of the binary complex. Protein was first incubated with three-fold molar excess FPP followed by three-fold molar excess inhibitor. Crystals were obtained from
the same growth condition, albeit with a more cubic morphology for the CCB2217 complex. Inhibitor was included at a 50 μM concentration during cryoprotection.

**Data collection and refinement**

Diffraction data were collected at 100 K on beamline 22-ID (SER-CAT) at the Advanced Photon Source. Crystals diffracted to ~ 1.8 Å and belonged to P2₁ space group with unit cell parameters of 63.3 x 90.5 x 83.3 Å³ (β = 110.6°). The data were reduced using *DENZO* and *SCALEPACK* (Otwinowski & Minor, 1997). Molecular replacement was carried out with the program *Phaser* (McCoy et al., 2007) using the structure of rat FTase as a search model. Crystals contained one molecule per asymmetric unit. Structures were built using *Coot* (Emsley et al., 2010), and model refinement was done in *REFMAC5* (Murshudov et al., 1997).

**FTase inhibitor specificity assay**

1 μM *A. fumigatus* and *R. norvegicus* FTases were incubated with 6 μM Ras CVVM, 10 μM 3,7-dimethyl-8-(7-nitro-benzo[1,2,5]oxadiazol-4-ylamino)-octa-2,6-diene-1-pyrophosphate (NBD-GPP from Jena Bioscience) and CCB2217 or SF122 at various concentrations as indicated in Fig. 4-4. 6 μM Rap2a was included as a positive control. All reactions were carried out in 20 mM Tris pH 7.7, 150 mM NaCl, 10 mM MgCl₂, 5 μM ZnCl₂, 0.04% dodecylmaltoside, and 2 mM DTT at 37°C for 2 hours before quenched with SDS-PAGE buffer. Samples were run on 4-20% gradient gels, and fluorescence on the farnesylated Ras was detected by a UV transillumination imaging system.
Chapter 5. Crystallization and preliminary X-ray diffraction analysis of human mismatch repair protein MutSβ

Introduction

The DNA mismatch repair (MMR) pathway is responsible for correcting base substitution and extrahelical base lesions that arise from DNA replication and proofreading errors (Hsieh & Yamane, 2008; Iyer et al., 2006). Additionally, it is involved in the suppression of homeologous recombination (Iyer et al., 2006; Jiricny, 2006; Li, 2008) and activation of cell cycle checkpoints and signal apoptosis as a response to DNA damage produced by chemical agents and carcinogens (Hsieh & Yamane, 2008; Iyer et al., 2006; Kunkel & Erie, 2005). MMR defects in mammals can lead to serious biological consequences, including a ~1000-fold decrease in DNA replication fidelity (Hsieh & Yamane, 2008; Iyer et al., 2006). In humans, defects in MMR genes are responsible for predisposition to hereditary nonpolyposis colorectal cancer, development of a subset of sporadic tumors, and resistance to certain chemotherapeutic agents (Iyer et al., 2006; Kunkel & Erie, 2005; Li, 2008).

Unlike the prokaryotic MutS homodimer, mismatch recognition in eukaryotes utilizes at least two MutS homolog (MSH) heterodimers: MutSα (MSH2-MSH6) and MutSβ (MSH2-MSH3). The role of a third heterodimer, MSH4-MSH5, appears to be limited to meiotic recombination (Ross-Macdonald & Roeder, 1994; Snowden et al.,
While MutSα binds to base-base mispairs and insertion-deletion loops of 1-10 unpaired nucleotides, MutSβ preferentially targets two or more unpaired nucleotides (Acharya et al., 1996; Genschel et al., 1998; Palombo et al., 1996). Although the conserved Phe-X-Glu motif has been shown to play a major role in mismatch recognition in both prokaryotic MutS and eukaryotic MutSα (Lamers et al., 2000; Obmolova et al., 2000; Warren et al., 2007), the underlying molecular basis for substrate specificity of MutSβ remains poorly understood. Despite the presence of a conserved Phe-X-Glu motif in the MSH3 mismatch binding domain, MutSβ is only weakly active on base-base mispairs (Genschel et al., 1998; Harrington & Kolodner, 2007) and single unpaired nucleotides (Acharya et al., 1996; Genschel et al., 1998; Palombo et al., 1996). Additional studies in *S. cerevisiae* have implicated the mismatch binding domain of MSH2 in MutSβ-but not MutSα-mediated MMR (Lee et al., 2007), and the DNA-binding mode of MutSβ appears to vary depending on size of the loop (Dowen et al., 2010). Taken together, these observations suggest a unique mechanism of mismatch binding in MutSβ.

Besides MMR, MutSβ plays an unclear role in trinucleotide repeat expansion, a critical step in the pathogenesis of diseases such as Huntington’s disease, fragile X syndrome, and myotonic dystrophy (Mirkin, 2007; Pearson et al., 2005). Despite requirement of *msh2* and *msh3* in the expansion of trinucleotide repeat sequences (Manley et al., 1999; Savouret et al., 2004; Foiry et al., 2006; van den Broek et al., 2002), which have the propensity to adopt unusual structures (Mitas et al., 1995; Pearson et al., 2005).
2002; Smith et al., 1995; Yu et al., 1995), MutSβ exhibits identical biochemical properties when interacting with CAG repeats and insertion-deletion loops (Tian et al., 2009b). Microsatellite instability, changes in length of short, repetitive DNA sequences, has also been partially attributed to MSH3 deficiency (Risinger et al., 1996), and msh3 polymorphisms have been suggested to increase the risk of colorectal cancer (Berndt et al., 2007).

For MutS proteins, the central function of DNA binding is tightly regulated by their ATPase activity (Schofield & Hsieh, 2003). Nucleotide binding and ATPase activities of MutSβ, previously inferred from that of MutSα, appear to be distinct and lesion dependent (Owen et al., 2009; Tian et al., 2009a), implying an association between differential nucleotide-occupancy states and lesion specificity. This, together with a common binding mode of human MutSα for a variety of DNA lesions (Warren et al., 2007), necessitates a comparative structural analysis of both heterodimers. In the present study, recombinant human MutSβ and three deletion mutants have been overexpressed in a baculovirus system and purified to homogeneity. Here we report the crystallization and preliminary X-ray diffraction analysis of the protein bound to a duplex DNA containing an insertion-deletion loop. This chapter is a reprint of the work presented in Tseng et al., 2011. DNA binding studies were done by Dr. Ravi R. Iyer.
Figure 5-1: Purification of MutSβ and deletion mutants. (a) Representative chromatogram (shown for MutSβΔ162) of final purification step (Mono Q column) for MutSβ and deletion mutants. Absorbance at 280 nm vs. elution volume is shown in blue; percent elution buffer vs. elution volume is shown in red. (b) 12% SDS-PAGE of purified human MutSβ and its deletion mutants. Lane 1 shows molecular weight standards with weight (in kDa) indicated on the left. Lanes 2, 3, 4, and 5 show full-length MutSβ (FL) Δ162, Δ175, and Δ223. Location of both MSH2 and MSH3 subunits is indicated with labelled arrows. Reprinted from Tseng et al., 2011.

Figure 5-2: SDS-PAGE of human MutSβ limited proteolysis by subtilisin. Leftmost lane shows molecular weight standards with weight (in kDa) indicated on the left. Digestion time (in minutes) is shown at the top. MSH3 fragment was detected starting two minutes after digestion. Locations of full length (FL) MSH3, MSH3 proteolysis fragment, and MSH2 are indicated with labelled arrows. Reprinted from Tseng et al., 2011.
Proteolytic fragment of human MSH3

Full-length hMutSβ (Fig. 5-1) was subjected to partial proteolysis by subtilisin. A proteolytic fragment (Fig. 5-2) was identified by N-terminal sequencing to be an MSH3 fragment starting at residue 176. A similar proteolytic fragment of MSH6 was observed for human MutSα (Warren et al., 2007), suggesting like domain structures between MSH3 and MSH6.

DNA binding studies of hMutSβ

To determine whether truncation of the N-terminal 162 amino acids of MSH3 alters the heteroduplex binding of MutSβ, we compared the affinity and specificity of MutSβ and MutSβΔ162 for a 41-bp I/D heteroduplex and an otherwise identical homoduplex DNA by SPRS. As shown in Fig. 5-3a, the affinities of MutSβ and MutSβΔ162 for a CA dinucleotide insertion-deletion I/D heteroduplex are similar (4±1 and 5±2 nM, respectively), and both proteins prefer heteroduplex over homoduplex DNA. Furthermore, both MutSβ (Fig. 5-3b) and MutSβΔ162 (Fig. 5-3c) dissociate with similar t_{1/2} values (~2.5-3 seconds) from heteroduplex DNA upon challenge with 1 mM ATP, indicating that that functional interaction of DNA and nucleotide binding sites is retained in MutSβΔ162. Similar results were obtained with MutSβΔ223 (data not shown).
Figure 5-3: Binding of MutSβ and MutSβΔ162 to a 41-bp heteroduplex containing a CA/I/D insertion-deletion loop. (a) Binding of MutSβ (circles) and MutSβΔ162 (squares) to a 41-bp heteroduplex (closed symbols) or control homoduplex (open symbols) was assessed by SPRS (Section 2.3). Mass response units at saturation (~400 s) are plotted as a function of protein concentration. Data were fit by non-linear least squares regression to a rectangular hyperbola yielding $K_d$ values of $4\pm 1$ nM for the wild type protein and $5\pm 2$ nM for MutSβΔ162 for heteroduplex binding. Binding to homoduplex DNA was not saturable up to 100 nM protein. Dissociation of MutSβ (b) or MutSβΔ162 (c) from a 41-bp heteroduplex (solid lines) and homoduplex (dotted lines) DNA substrates upon challenge (arrow) with 1 mM ATP was assessed by SPRS.

Reprinted from Tseng et al., 2011.
Crystallization of hMutSβ

MutSβΔ162 was initially cocrystallized with ADP and a 14-bp duplex DNA containing a CA insert (5’-CCTAGCGCAGCGGTTC-3’ annealed with 5’-GAACCGCCGCTAGG-3’). By using 1 μl protein solution and 1 μl precipitant, crystals with a bipyramidal morphology (form 1; Fig. 5-4a) grew out of drops containing a
precipitating agent of 100 mM HEPES pH 7.5, 200 mM calcium acetate, 5 mM cadmium chloride, and 9-12% PEG 8000 at 290 K. Further screening by varying DNA substrate length and end base pair sequence led to a second crystal form using the MutSβΔ223 construct. Crystallization of MutSβΔ223 with a revised duplex (5’-TCCCTAGCGCAGCGGTTCCGA-3’ annealed with 5’-CGGAACCGCCGCTAGGGAC-3’) yielded three-dimensional rectangular crystals (form 2; Fig. 5-4b) from 100 mM trisodium citrate pH 5.6, 250 mM lithium sulfate, and 12-15% PEG 4000 at 290 K. Crystals reached full size in five days and grew up to ~200 x 80 x 40 μm³.

Form 1 crystals were transferred to a stabilization solution (100 mM HEPES pH 7.5, 200 mM calcium acetate, 17% PEG 8000, 150 mM KCl, and 0.5 mM EDTA) and serially to cryobuffer solutions supplemented with 7, 13.5, 20, and 27% ethylene glycol. Similar step transfer cryocooling done with glycerol, PEG 8000, calcium acetate, and PEG 200 compromised diffraction quality. Form 2 crystals were cryoprotected by stepwise transfer into mother liquor solutions supplemented with ethylene glycol, up to a final ethylene glycol concentration of 22% before flash-cooling in liquid nitrogen.

**Summary**

In contrast to the repair of both base-base mismatches and insertion-deletion loops by a single homodimer in prokaryotes, eukaryotic MMR employs multiple heterodimeric MutS homologs. This partition of substrate recognition is puzzling, given the relatively low cellular abundance of MutSβ (Drummond et al., 1997; Genschel et al.,
and its functional overlap with MutSα in repairing extrahelical nucleotides (Acharya et al., 1996; Genschel et al., 1998). Furthermore, despite the presence of a conserved Phe-X-Glu motif in both MSH3 and MSH6 mismatch binding domains, base-base mispairs are ideal substrates for MutSα but not MutSβ (Palombo et al., 1996). Structural characterization and comparison of both MutSβ and MutSα bound to heteroduplex DNA will likely clarify this ambiguity, provide insight into the molecular mechanism of eukaryotic mismatch recognition, and establish a basis for studying repeat instability and segregation fidelity.

**Materials and methods**

**Cloning, expression, and purification of hMutSβ**

Three hMutSβ deletion mutants were generated from a pFastBac Dual vector containing msh2 and msh3 by PCR mutagenesis. These constructs (MutSβΔ162, Δ175, and Δ223) lack the first 162, 175, and 223 residues of MSH3, respectively. The truncation in MutSβΔ162 is based on sequence alignment with hMutSαΔ341 (Warren et al., 2007), whereas the MutSβΔ223 deletion eliminates predicted disordered regions before the putative MSH3 mismatch binding domain (Notredame et al., 2000). MutSβΔ175 is a proteolytic fragment identified by subtilisin digestion. These mutants were expressed in Sf9 insect cells and purified as previously described for the full-length protein (Genschel et al., 1998) using buffers supplemented with 10 μg/ml leupeptin, 10 μg/ml aprotinin, 1 μg/ml pepstatin A, 0.1% (v/v) saturated phenylmethylsulfonyl fluoride,
and 1 mM dithiothreitol. Cell lysate was centrifuged (18,500 x g) and loaded onto a Q-Sepharose column equilibrated with 250 mM KCl buffer (25 mM HEPES-KOH pH 7.5, 0.5 mM EDTA, and 250 mM KCl) and directly connected to a single-stranded DNA-cellulose column. After the two columns were disconnected, the single-stranded DNA-cellulose column was washed with 300 mM KCl buffer and subsequently with 300 mM KCl buffer containing 2.5 mM MgCl₂. MutSβ was eluted with 300 mM KCl buffer containing 2.5 mM MgCl₂ and 1 mM ATP. Pooled fractions were diluted to 100 mM KCl and loaded onto a second Q-Sepharose column equilibrated with 100 mM KCl buffer. Fractions eluted with 250 mM KCl buffer were diluted to 100 mM KCl and loaded onto an 8-ml Mono Q 10/100 GL column (GE Healthcare). Protein was eluted with an 80-ml linear gradient of 100-370 mM KCl buffer at 0.5 ml/min and stored at 193 K in 150 mM KCl buffer. Purity of the protein was confirmed on a 12% SDS-PAGE gel (Fig. 5-1).

**Limited proteolysis of hMutSβ**

Full-length hMutSβ (3.6 μM) was incubated with 5 mM MgCl₂, 2-fold molar excess of ADP, and 1.2-fold molar excess of a 14-bp duplex DNA containing a CA insert (5’-CCTAGCGCAGCGGTTC-3’ annealed with 5’-GAACCGCCGCTAGG-3’). The mixture was subjected to partial proteolysis by subtilisin (1:167 protease:protein molar ratio) at 310 K in a buffer containing 25 mM HEPES pH 7.5 and 150 mM NaCl. Samples were quenched with 5 mM phenylmethylsulfonyl fluoride at the indicated time points, run on a 12% SDS-PAGE gel, and electroblotted onto PVDF membrane. After
Coomassie Blue staining, proteolytic fragment (Fig. 5-2) was excised and submitted for peptide sequencing (University of Texas Medical Branch Biomolecular Resource Facility).

**DNA binding studies of hMutSβ**

Heteroduplex DNA substrates (41-bp) used for surface plasmon resonance spectroscopy (SPRS) experiments contained a centrally positioned CA dinucleotide insertion-deletion (I/D) loop and were prepared as described previously (Blackwell et al., 2001) by annealing HPLC-purified synthetic oligonucleotides 5′-AGCCGAATTTTCTAGCTTGAGCTAGCAATTCGGCG-3′ (top strand) and 5'-CGCCGAATTGCTAGCAAGCTCAATCGAGTCTAAAAATTCGGCT-3′ (bottom strand, CA insertion underlined). Corresponding homoduplexes were prepared by annealing the top strand with the oligonucleotide 5′-AGCCGAATTTTCTAGCTTGAGCTAGCAATTCGGCG-3′. The top strand contained a 5′ biotin tag for immobilization on a streptavidin-coated surface.

MutSβ binding to DNA was analyzed by SPRS on a Biacore 2000 (GE Healthcare). Indicated concentrations of MutSβ or MutSβΔ162 (in 25 mM HEPES-KOH pH 7.5, 150 mM KCl, 1 mM MgCl₂, 1 mM dithiothreitol, 0.02% surfactant P20) were allowed to flow at 20 μl/min over a streptavidin chip derivatized with 200 response units of a biotinylated 41-bp homo- or heteroduplex DNA. Protein flow was carried out for 5 minutes followed by a 1 min wash with running buffer. MutSβ dissociation from
DNA upon ATP challenge was then monitored by injection of 1 mM ATP in running buffer. Dissociation kinetic data were fit by non-linear regression to a single exponential decay function.

**DNA substrates**

Oligonucleotides (Midland Certified Reagent Co.) were annealed in 10 mM sodium cacodylate pH 7.4, 50 mM NaCl, 10 mM MgSO$_4$, and 0.5 mM EDTA. For initial crystal screening, a 14-bp duplex DNA containing a CA insert (5’-CCTAGCGCAGCGGTTC-3’ annealed with 5’-GAACCGCCGCTAGG-3’) was derived from the G·T mispair heteroduplex used for the crystallization of hMutSα (Warren et al., 2007). DNA length (5’-CCTAGCGCAGCGGTTC-3’ annealed with 5’-GACCGCCGCTAGG-3’, 5’-CTAGCGCAGCGGTTC-3’ annealed with 5’-GAACCGCCGCTAGG-3’, 5’-CAGCGCAGCGGTTC-3’ annealed with 5’-GAACCGCCGCTAGG-3’, 5’-TCCCTAGCGCAGCGGTTCCTCGA-3’ annealed with 5’-CGGAACCCGCGCTAGGAC-3’) and end base pair sequences (5’-CCTAGCGCAGCGGTTC-3’ annealed with 5’-AAACCGCCGCTAGG-3’, 5’-GCTAGCGCAGCGGTTC-3’ annealed with 5’-GAACCGCCGCTAGG-3’, 5’-CCTAGCGCAGCGGTTC-3’ annealed with 5’-CAACCGCCGCTAGG-3’) were subsequently explored for screening and optimization.
Crystallization

Full-length hMutSβ, Δ162, Δ175, and Δ223 were used for crystallization trials. Protein samples at 29 μM were incubated with 2.5 mM MgCl₂, 10 mM dithiothreitol, twofold molar excess of ADP, and 1.2-fold molar excess of loop-containing duplex DNA of various lengths. Sitting-drop vapor-diffusion method and a set of sparse-matrix crystallization screens from QIAGEN were used for initial crystallization trials at 290 K. Screening was done using the Mosquito robot (TTP LabTech) in two-well MRC crystallization plates using a reservoir volume of 82 μl and drops consisting of 200 nl protein solution and 200 nl reservoir solution.

Data collection and processing

Diffraction data were collected at 100 K on beamline 22-ID (SER-CAT) at the Advanced Photon Source and beamline 12.3.1 (SIBYLS) at the Advanced Light Source. Crystal data is summarized in Table 1. Crystal form 1 diffracted to ~10 Å resolution and belonged to a tetragonal space group with unit cell parameters of 300 x 300 x 260 Å³. Crystal form 2 belonged to the C2 space group with unit cell dimensions 185 x 134 x 102 Å³. Form 2 crystals diffracted to a resolution of ~3.1 Å (Fig. 5-5), and two data sets from these crystals were merged to produce a more complete data set. Data were reduced using XDS (Kabsch, 2010) and scaled to a resolution of 3.1 Å with an \( R_{merge} \) of 7.5% and an overall completeness of 100.0%. Solvent content analysis suggested a Matthews coefficient (Matthews, 1968) of 2.88 Å³ Da⁻¹, corresponding to one molecule in the
asymmetric unit with a solvent content of 57%. Analysis of reduced data was carried out using SCALA (Collaborative Computational Project, Number 4, 1994). Molecular replacement (MR) was carried out with the program Phaser (McCoy et al., 2007) using the structure of human MutSα (PDB code 2O8B) (Warren et al., 2007) as a search model. Although Phaser provided a clear MR solution (both rotational and translational Z scores > 20) with the expected one molecule per asymmetric unit, the resulting maps were not of sufficient quality to build the 976 residues necessary to complete the MSH3 subunit. Expression of selenomethionine-substituted protein to provide experimental phasing information is under way.

Figure 5-5: A diffraction pattern of crystal form 2 collected on beamline 12.3.1 at the Advanced Light Source. Resolution ring corresponds to 3.05 Å. Reprinted from Tseng et al., 2011.
References


