Structural and Biochemical Insights into Antifungal Drug Targets from Aspergillus fumigatus

A thesis submitted for the degree of Doctor of Philosophy

Stephanie Nguyen

Department of Molecular and Cellular Biology,

School of Biological Sciences, The University of Adelaide

July 2021

THE UNIVERSITY of ADELAIDE

Table of Contents

Publications Included in the Thesis:

Nguyen, S., Jovcevski, B., Pukala, T. L., Bruning, J. B. (2020). Nucleoside selectivity of *Aspergillus fumigatus* nucleoside-diphosphate kinase. *FEBS J*, **288**, doi: 10.1111/febs.15607

Nguyen, S., Truong, J. Q., Bruning, J. B (2021). Targeting unconventional pathways in pursuit of novel antifungals. *Front. Mol. Biosci.* **7**: 621366, doi: 10.3389/fmolb.2020.621366

Other Publications:

Truong, J. Q.*, **Nguyen, S.***, Bruning, J. B., Shearwin, K. E. (2021). Simplified heavy atom derivatization of protein structures via co-crystallization with the MAD tetragon B4C. *Acta Cryst. F.*, **F77**, doi: 10.1107/S2053230X21004052

Truong, J. Q.*, **Nguyen, S.***, Bruning, J. B., Shearwin, K. E (2021). Derivatization of Protein Crystals with I3C using Random Microseed Matrix Screening. *J. Vis. Exp.*, **167**, e61894, doi:10.3791/61894

Slynko, I.*, **Nguyen, S.***, Hamilton, E. M. C.*, Wisse, L. E., de Esch, I. J. P., de Graaf, C., Bruning, J. B., Proud, C. G., Abbink, T. E. M., van der Knaap, M. S. (2021). Vanishing white matter: Eukaryotic initiation factor 2B model and the impact of missense mutations. *Mol Genet Genom Med*. **e1593**, doi: 10.1002/mgg3.1593

Sternicki, L. M., **Nguyen, S.,** Pacholarz, K. J., Barran, P., Pendini, N. R., Booker, G. W., Huet, Y., Baltz, R., Wegener, K. L., Pukala T. L., Polyak, S. W. (2020) Biochemical characterisation of class III biotin protein ligases from *Botrytis cinerea* and *Zymoseptoria tritici*. *Arch. Biochem. Biophys.* **691**, 108509, doi: 10.1016/j.abb.2020.108509

Jia, Y., Shen, G., **Nguyen, S.,** Zhang, Y., Huang, K.-S., Ho, H.-Y., Hor, W.-S., Yang, C.-H., Bruning, J. B., Li, C., Wang, W.-L. (2020) Analysis of the mutation dynamics of SARS-CoV-2 reveals the spread history and emergence of RBD mutant with lower ACE2 binding affinity. *bioRxiv*. doi: 10.1101/2020.04.09.034942

* denotes that both authors contributed equally

Thesis Summary

Aspergillus fumigatus, the leading cause of invasive aspergillosis, has been recognised as a priority fungal pathogen by the World Health Organisation due to its impact on human health and the emergence of resistance against existing antifungal therapeutics. Ubiquitously found in the natural environment, *A. fumigatus* produces air-borne fungal spores that are routinely inhaled into the respiratory system where the initial infection is established. In severe cases, the fungus can subsequently invade multiple organs, causing invasive aspergillosis. Current approaches to treating invasive fungal infections require aggressive and expensive antifungal therapy with limited success due to misdiagnosis, poor drug pharmacokinetics and pharmacodynamics, high drug toxicity and low efficacy of existing drugs against resistant strains. As the mortality rate associated with invasive aspergillosis remains unacceptably high and primary therapy failure occurs in 40% - 70% for invasive aspergillosis patients, there is a need to identify new targets for the development of novel classes of antifungal therapeutics.

In order to effectively embark on a drug discovery project, intimate structural and biochemical knowledge of the target is required. Together, this information provides insights into protein function and furthers our understanding of their contribution to fungal survival and virulence. Furthermore, the structural data will be essential in guiding rational drug design efforts. This thesis focuses on four *A. fumigatus* enzymes from the purine biosynthesis, mannitol biosynthesis and glycolysis pathways that have been identified as promising antifungal drug targets. The structure and kinetic characteristics of each enzyme has been comprehensively studied using X-ray crystallography, analytical size-exclusion chromatography, native mass spectrometry and *in vitro* activity assays. Findings from this work has provided a necessary foundation for future antifungal drug discovery projects that target nucleoside diphosphate-kinase (NDK), guanosine monophosphate (GMP) synthase, mannitol-2-dehydrogenase (M2DH) and enolase.

(1) Seven crystal structures of NDK (1.6 Å – 2.3 Å) were solved, either in an unbound form, or bound to one of the six possible nucleoside triphosphate substrates used in its reaction. Analysis of the kinetic properties of the enzyme revealed the following order of preference: adenosine > guanosine > inosine > uridine > thymidine > cytidine substrates. By combining the structural and kinetic data obtained, the structural determinants that govern nucleoside selectivity in *A. fumigatus* NDK were determined.

(2) The first structure of a fungal GMP synthase enzyme from *A. fumigatus* was solved to a resolution of 2.3 Å. Analysis of this structure in comparison to the existing structure of the human homologue has revealed significant differences. Despite both being from eukaryotic species, the *A. fumigatus* GMP synthase lacks a D1 dimerisation domain that is observed in the human counterpart. Although denoted as the D1 dimerisation domain in the literature, this domain is known to prevent dimerisation. As a result, GMP synthase forms a dimeric complex, which is contrary to the monomeric human homologue. Analysis of binding-pocket residues in combination with rigorous analysis of the kinetic properties have revealed species-specific differences exist between fungal and human GMP synthases. This data will be imperative in the design of inhibitors that selectively target the fungal homologue.

(3) The crystal structure of *A. fumigatus* M2DH in an unbound state (1.8 Å) and bound to NADH (2.1 Å) were solved. Analysis of these structures have revealed a central binding cavity that is located between the N- and C-terminal domains of the enzyme. This binding pocket is predominantly positively charged that readily accommodates the negative charges derived from the phosphate groups of NADH. Kinetic analysis of the substrates

and co-factors used in the interconversion reaction have revealed that M2DH is likely to play a supporting role to the enzyme mannitol-1-phosphate-5-dehydrogenase in the biosynthesis of mannitol in fungi. As a proof of concept, we have shown that the activity of M2DH can be modulated by the small molecule 1,4-benzoquinone and have determined that the mechanism of inhibition is achieved by the covalent modification of all 5 cysteine residues.

(4) Although the canonical function of cytoplasmic enolase as a glycolysis enzyme is well understood, its 'moonlighting' functions as a virulence factor have not been explored in molecular detail. Enolase, when expressed on the surface of pathogenic bacteria and fungi, can function as a receptor for human proteins, including plasminogen. This interaction drives tissue invasion and facilitates nutrient acquisition, both of which function to accelerate infection. A novel approach to developing antifungal drugs aims to prevent the formation of the enolase-plasminogen complex. In order to better understand this interaction, the crystal structure of *A. fumigatus* enolase, unbound (2.0 Å) and bound to its endogenous substrates, 2-phosphoglycerate (1.9 Å) and phosphoenolpyruvate (2.3 Å) were determined. Enolase from *A. fumigatus* forms a dimeric complex in which both monomers are arranged in an anti-parallel orientation. Using the solved crystal structure of *A. fumigatus* enolase and the existing crystal structure of human plasminogen, we have developed a model of this complex using protein: protein docking software. These data provide insights to better understand this interaction and provides foundations to guide the design of peptidomimetics or small molecules that could potentially disrupt this interaction.

Statement of Declaration

I certify that this work contains no material which has been accepted for the award of any other degree or diploma in my name, in any university or other tertiary institution and, to the best of my knowledge and belief, contains no material previously published or written by another person, except where due reference has been made in the text. In addition, I certify that no part of this work will, in the future, be used in a submission in my name, for any other degree or diploma in any university or other tertiary institution without the prior approval of the University of Adelaide and where applicable, any partner institution responsible for the joint-award of this degree.

I acknowledge that copyright of published works contained within this thesis resides with the copyright holder(s) of those works.

I also give permission for the digital version of my thesis to be made available on the web, via the University's digital research repository, the Library Search and also through web search engines, unless permission has been granted by the University to restrict access for a period of time.

I acknowledge the support I have received for my research through the provision of an Australian Government Research Training Program Scholarship.

25th July, 2021

Acknowledgements

To my supervisor, Dr. John Bruning, thank you for introducing me to the world of protein crystallography! I have been so blessed to be able to complete my PhD in your lab. Over the last 3 years, you have taught me so many valuable skills – in and out of the lab. Thank you for your guidance, patience and cynical commentary. Without your mentorship, I could have never become the scientist I am today.

To my co-supervisor, Dr, Blagojce Jovcevski, thank you for being such a huge part of my PhD! Even when you weren't on my supervisory panel, you were an amazing friend and confidant. I always loved coming into your office to design experiments or plan papers or gossip or cry (mostly me) or eat. Thank you for being one of the biggest supporters during my PhD and I hope that we forever listed as 'co-authors' on each other's Google Scholar pages.

To my co-supervisor, Assoc. Prof. Keith Shearwin, thank you for being involved on the project! I am very grateful for the productive chats in your office about all things cloning-related. Without you (and your restriction enzymes), the Bruning Lab would have gone broke from ordering plasmids.

To Tara (I feel like you should be an honorary co-supervisor), thank you for your contributions to all of the manuscripts I have written during my PhD. Your expertise in mass spectrometry have been invaluable to the project.

To the Bruning Lab, especially Andrew, Bec, Jordan, Dan, Ben, James, Beth, thank you for being the most wonderful support network! I've had the best time eating snacks together, drinking in John's office, discussing last night's episode of Married at First Sight and complaining about the state of the lab.

To the people of the MLS, including Fiona, Tas and friends from the 3rd floor, thank you for the long chats in the hallways, meeting rooms and kitchens. I will always be grateful for the laughs and good times.

To my friends, Sandra, Minh, Tom, Nelz, thank you for looking out for me! Thank you for dragging me out to try new things and constantly calling the group chat. Without you all, I might never leave the house. You guys are the best.

To my partner, Jia, thank you for everything you do for me, especially for cooking dinner and cleaning the apartment while I'm writing this thesis. I truly appreciate your constant support and encouragement.

To my family, thank you for supporting me throughout this entire journey. Mum, Dad and Damian – your support means the world to me. I could have never done this without you.

List of Abbreviations

A. fumigatus	Aspergillus fumigatus
AMP	Adenosine monophosphate
ATP	Adenosine triphosphate
CTP	Cytidine triphosphate
Da	Dalton
DMSO	Dimethyl sulfoxide
DTT	Dithiothreitol
E. coli	Escherichia coli
EDTA	Ethylenediaminetetraacetic acid
FPLC	Fast protein liquid chromatography
GMP	Guanosine monophosphate
GTP	Guanosine triphosphate
h	Hour
HCl	Hydrochloric acid
HEPES	4-(2-hydroxyethyl)-1-piperazineethanesulfonic acid
IC_{50}	Half maximal inhibitory concentration
IMAC	Immobilised metal affinity chromatography
IM-MS	Ion mobility – mass spectrometry
IPTG	Isopropyl β-D-1-thiogalactopyranoside
ITP	Inosine triphosphate
k_{cat}	Turnover number
kDa	Kilodalton
K_M	Michaelis-Menten constant
LB	Luria Bertani
m/z	Mass to charge ratio
M2DH	Mannitol-2-dehydrogenase
MOPS	3-(N-morpholino)propanesulfonic acid
MS	Mass spectrometry
nanoESI-MS	Nanoelectrospray ionisation – mass spectrometry
NaOH	Sodium hydroxide
NDK	Nucleoside diphosphate-kinase
NH_4OAc	Ammonium acetate
PBS	Phosphate buffered saline
PEG	Polyethylene glycol
s	Second
SDS-PAGE	Sodium dodecyl-sulphate – polyacrylamide gel electrophoresis
SEC	Size-exclusion chromatography
TEV	Tobacco etch virus
Tris	2-amino-2-hydroxymethylpropane-1,3-diol
TTP	Thymidine triphosphate
Tween-20	Polyoxyethylene-sorbitan monolaurate
UTP	Uridine triphosphate

Chapter 1:

Insights into the pathogenic fungus, *Aspergillus fumigatus* and emerging targets for antifungal drug development

Chapter 1 Preface

Chapter 1 includes sections of a review article published in Frontiers of Molecular Biosciences that describes the features of fungal pathways that are of interest in this PhD project. The published article appears in Appendix A and the references are included in Chapter 7.

1.1 The growing burden of fungal infections demands the development of new classes of antifungals

Opportunistic fungal pathogens impose a significant societal and economic burden on the public health sector, particularly amongst immunocompromised patients. In the United States alone, it has been estimated that fungal infections incur hospitalisation and outpatient costs that exceed $7.2 billion USD annually (1). These costs are expected to rise in tandem with the expansion of the immunocompromised population and subsequent increase in the incidence of invasive mycoses. This predisposed population includes recipients of invasive surgery, chemotherapy and immunosuppressive therapy, as well as sufferers of human immunodeficiency virus (HIV)/acquired immunodeficiency syndrome (AIDS) and cystic fibrosis (2).

Global estimates suggest that over 150 million individuals are affected by serious fungal infections, which culminate in over 1.6 million deaths per annum (3). Common fungal pathogens associated with these opportunistic infections include *Aspergillus* spp. (most commonly *Aspergillus fumigatus*), *Candida albicans* (and other non-*albicans Candida* species), *Cryptococcus neoformans*, *Histoplasma capsulatum*, *Coccidioides imitis* and *Pneumocystis jirovecii* (4,5). Amongst these fungal pathogens, species of *Aspergillus*, *Candida, Pneumocystis* and *Cryptococcus* are responsible for over 90% of total global fungal mortalities and therefore, these fungal species pose the most prominent threat to human health (6). *A. fumigatus* preferentially colonises lung tissue, resulting in pulmonary aspergillosis (7). The severity of this fungal infection can escalate rapidly as the infection disseminates throughout the body and affects multiple organs. Infections caused by *A. fumigatus* can manifest as invasive aspergillosis which affects multiple organs, particularly the kidney and brain (8,9). This disease is associated with high rates of primary antifungal therapy failure which ultimately leads to high rates of mortality and significant burden on the public health system (10).

1.2 Overview of *Aspergillus fumigatus* and impact on human health

The project focuses on *A. fumigatus*, a saprophytic fungus that is ubiquitous in nature and commonly found in soil and decaying organic matter (11). Although *A. fumigatus* grows at an optimum temperature of 37°C and pH of 3.7 to 7.6, it shows high tolerance for environmental variability, which contributes to its adaptive response during infection (12). Through asexual reproduction, *A. fumigatus* produces hydrophobic conidia that are readily dispersed through the air (12). It has been predicted that adult humans will inhale between 100 and 1000 conidia from *A. fumigatus* per day where they enter the lower airways and deposit in the bronchioles and alveoli (13,14). In healthy humans, the conidia are effectively removed via mucociliary clearance or neutralised by alveolar macrophages and neutrophils (14). However, conidia can survive in immunocompromised hosts and within 4 to 6 hours, conidia shed their hydrophobic layers, swell and germinate to form hyphae (vegetative structures) (13).

Consequently, an invasive and systemic infection, collectively termed invasive aspergillosis can develop and persist in the lungs. The pathological features of the disease can vary between patients but generally include blood clotting and haemorrhage due to extensive hyphae growth (14). This disease is predicted to affect over 300,000 individuals globally and is associated with an exceedingly high mortality rate of up to 80% (3).

1.3 Current antifungal classes and treatments

Antifungal therapeutics are routinely administered to combat invasive mycoses. The four main classes of antifungals that currently exist function by targeting either the cell wall or cell membrane (Figure 1). They include azoles, allylamines, polyenes and echinocandins. The mechanism of action of each antifungal drug class and discussion of their biological targets have been extensively reviewed (15). A brief overview of currently available antifungals is given here.

1.4 Targeting the cell membrane

Ergosterol is a steroid alcohol that forms a major component of the fungal cell membrane. It is responsible for the maintenance of membrane fluidity and structural integrity, as well as the regulation of membrane permeability (16). Due to its implications in fungal survival, ergosterol biosynthesis enzymes have been the target of the allylamine and azole classes of antifungals, whereas the polyene class targets ergosterol function directly (15). Allylamines target squalene epoxidase, an enzyme that catalyses the conversion of squalene to lanosterol, whereas azoles function as non-competitive inhibitors of lanosterol 14a-demethylase by complexing with the haem prosthetic group (15,17,18). Both of these enzymes are involved in the later stages of ergosterol biosynthesis. Disruption of this pathway leads to an accumulation of upstream sterol precursors and ultimately, the depletion of ergosterol. This deficiency results in impaired cell membrane function, growth inhibition, and cell death (18,19). In comparison, polyenes induce fungal death by binding directly to ergosterol embedded in the cell membrane. Once incorporated into the membrane, this leads to the formation of pores, which increase membrane permeability to water, ions and non-electrolytes (20).

1.5 Targeting the cell wall

Due to the absence of a mammalian equivalent, several components of the fungal cell wall and the biosynthetic enzymes responsible for its construction have been the target for antifungal therapeutics. The cell wall, comprised of three main components including glucans, mannoproteins and chitin, plays an important role in maintaining fungal survival by maintaining cell rigidity and resisting osmotic stress (21). Furthermore, there are several receptors displayed on the surface of the cell wall that can interact with the host during infection and contribute to fungal virulence. Echinocandins disrupt these fundamental functions of the cell wall by non-competitively inhibiting (1,3)-β-glucan synthase, the enzyme responsible for maintaining and synthesising glucans. Insufficient biosynthesis of glucans leads to destabilisation of the cell wall and a compromised ability to resist osmotic pressure that results in cell lysis (22).

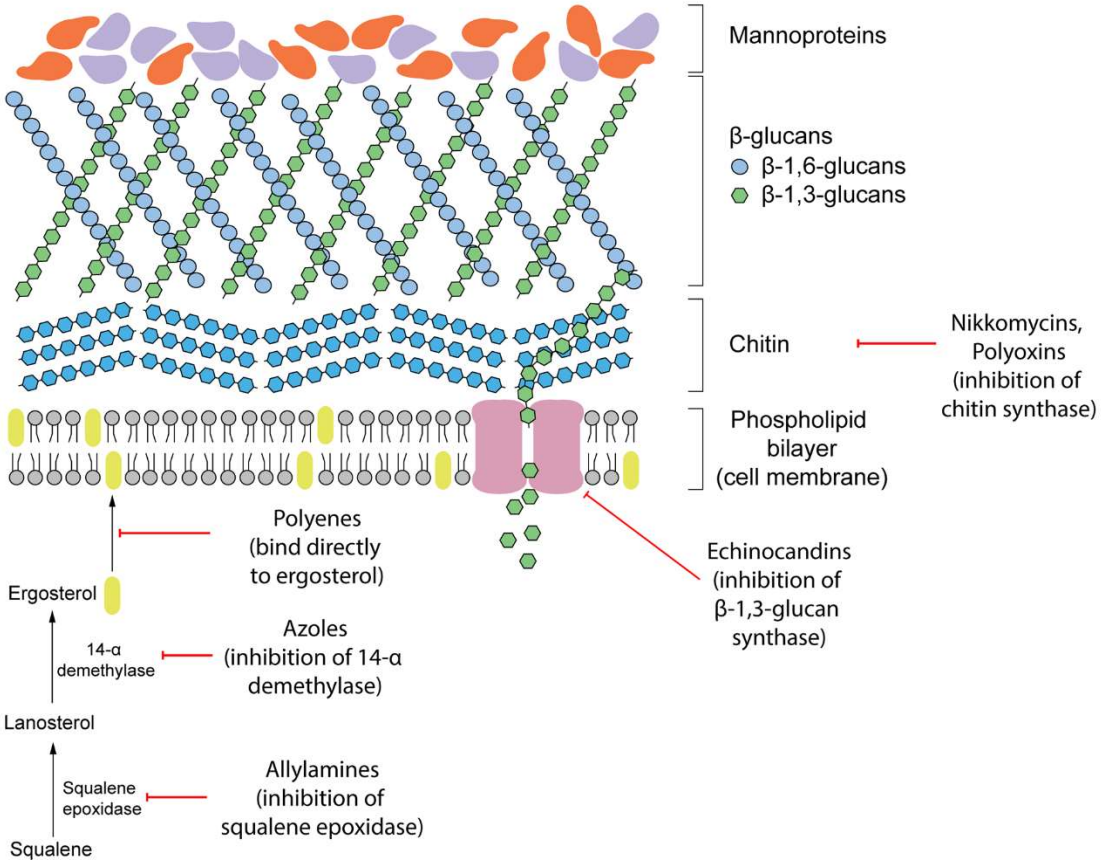


Figure 1: Existing antifungal drug classes that target the biosynthesis of the fungal cell wall components (nikkomycins, polyoxins, echinocandins), biosynthesis of the cell membrane (azoles, allylamines) or bind to ergosterol in the membrane directly (polyenes).

1.6 Emergence of antifungal resistance and the need for novel antifungal classes

The extensive use of antifungals in both agricultural and clinical settings has led to the rise of widespread antifungal resistance with significant consequences to food security and human health. To control fungal phytopathogens in agricultural crops, a range of antifungal drugs specifically for agricultural use are used that target mitochondrial respiration processes, including succinate dehydrogenase and Qo inhibitors (23). However, there is also significant overlap in the use of azoles to treat both human and plant fungal pathogens. In particular, they are frequently used in grain- and grass-growing environments, as well as in clinical settings, to control *Candida* spp., *Cryptococcus* spp. and *Aspergillus* spp. infections (24). As a result, bioactive azoles can persist in soil, water and fresh produce (25). Ultimately, it has been suggested that these environmental factors contribute to the rise in azole resistance currently being observed in clinical settings.

The prominent use of azoles has resulted in a shift in the prevalence of infections caused by *C. albicans* towards less susceptible species of *Candida*, including *C. glabrata*, *C. krusei* and *C. guilliermondii* (26-28). Fungal strains can also acquire drug resistance during therapy, which restricts the treatment options available to the patient and can ultimately cause antifungal therapy failure. Specifically, there are increasing concerns of echinocandin and azole resistance arising in species of *Candida* and *A. fumigatus*, respectively (28). In the United Kingdom, there has been a consistent increase in the proportion of *A. fumigatus* isolates with resistance to azoles growing from 5% in 2004 to 14% in 2008 and 20% in 2009 (29). Following a similar trend, the frequency of *A. fumigatus* clinical isolates with azole resistance in The Netherlands increased from 7.6% in 2013 to 14.7% in 2018 (30).

Several generations of antifungal drugs originating from these four classes have been developed to improve their pharmacokinetic profiles, increase potency and combat resistance. However, treatment of invasive mycoses amongst immunocompromised patients remains challenging. Primary antifungal therapy failure is a prevalent issue, affecting 20 – 60% of invasive candidiasis patients, 40 – 60% of invasive aspergillosis patients and 30 – 100% of invasive fusariosis patients (10). Although the role of underlying host factors and immune status must be considered in antifungal therapy failure, several issues arise from inherent limitations in currently available drugs. Low bioavailability in target tissues, drug toxicity, unfavourable drug-drug interactions and the emergence of drug resistance also contribute to the reduced efficacy of existing treatments (10).

As the number of antifungal classes available for clinical use is limited and the proportion of the population susceptible to invasive mycoses is expected to increase, there is a need to identify new targets for the design and development of novel classes of antifungals. Although cell membrane and cell wall biosynthesis enzymes have been traditional targets of antifungals, broadening our search to unconventional pathways can unlock an abundance of potential targets by exploiting the unique metabolic needs of fungi during pathogenesis.

In spite of the alarming rates of morbidity and mortality, the severity of invasive fungal infections remains underappreciated. Increases in disease incidence and prevalence of antifungal resistance highlights the need to identify novel targets and develop new classes of antifungals to manage mycoses amongst the immunocompromised population. There have been continual efforts to characterise enzymes involved in the biosynthesis of ergosterol or cell wall components, both of which are classic antifungal targets, to develop novel inhibitors (31-33). However, there has also been a notable shift in focus from these pathways exclusive to fungi to exploiting species-specific differences in shared pathways between fungi and humans (34-37).

To effectively establish infection, the fungus must adapt to a different niche within the human host, combat or circumvent the host immune response and obtain sufficient nutrients to reproduce and disseminate. Although these metabolic requirements may differ between fungal species, depending on their preferred infection site, disrupting shared metabolic pathways involved in these processes can impede fungal survival and pathogenesis. Targeting these pathways may present an elegant approach to develop novel classes of therapeutics with broad spectrum activity.

1.7 Emerging targets for antifungal development

There is a clear need to identify new targets to develop antifungal drug classes that are not subject to pre-existing resistance mechanisms. *A. fumigatus* is an opportunistic fungal pathogen that lacks specific virulence factors. Instead, its virulence is multifactorial and is not only dependent on the host immune system but also on its inherent biological traits that contribute to its pathogenicity (11). It has been suggested that *A. fumigatus* has acquired these traits to survive in an immunocompromised host from the hazards encountered in its natural environment (13). For example, the size of conidia (between 2 to 3 μm) not only facilitates wind dispersal but also allows for deep infiltration into the alveoli. Similarly, its thermotolerance and pH tolerance that enables *A. fumigatus* to thrive in its natural environment of compost heaps also aids its survival in harsh conditions within the human host (14). Furthermore, to complement its saprophytic lifestyle, *A. fumigatus* expresses numerous extracellular proteases and exhibits nutritional versatility that also contributes to its virulence (12,38). Thus, targeting genes that are involved in essential metabolic pathways or those that contribute to its opportunistic nature offer a promising approach to developing novel classes of antifungal therapeutics.

1.8 Glycolysis pathway

Glycolysis is a fundamental, multi-step metabolic process that occurs in most living organisms. The glycolysis pathway produces energy, in aerobic and anaerobic conditions, through the catabolism of sugars and provides useful intermediates for downstream biosynthetic pathways. The roles of these enzymes in glycolysis have been extensively studied but there has been growing interest in their additional 'moonlighting' functions that deviate from their canonical function. These 'moonlighting' functions may contribute to fungal survival and virulence mechanisms and have consequently been investigated as promising antifungal drug targets.

1.8.1 Enolase and host plasminogen

Enolase catalyses the penultimate step of glycolysis, interconverting 2-phosphoglycerate (2-PG) and phosphoenolpyruvate (PEP) (Figure 2A) (39). Consistent with its role in integral metabolic processes, a genetic knockout of enolase in *C. albicans* has been shown to reduce germination tube and hyphal formation, resulting in attenuated virulence and growth rate (40,41). Intriguingly, this enzyme fulfils its glycolytic roles in the cytoplasm but it is also expressed on the cell surface of many fungal and bacterial species. This has been observed in *A. fumigatus* and *C. albicans*, two of the most prominent etiological agents of invasive mycoses, as well as *Aspergillus flavus, Aspergillus terrus, Aspergillus nidulans, Candida glabrata, Saccharomyces cerevisiae*, and bacterial *Streptococci* and *Pneumococci* species (42-45). In *S.*

cerevisiae, it has been suggested that enolase secretion is mediated by an N-terminal 28 amino acid translocation sequence and operates through a SNARE-dependent pathway (46). Alignment of homologous enolase protein sequences from *S. cerevisiae* with *Aspergillus* spp. and *Candida* spp. has identified a highly conserved motif at the N-terminus. This suggests that this SNARE-dependent mechanism of enolase secretion may be conserved amongst many fungal species (45).

The function of surface-expressed enolase deviates from its canonical role in the cytoplasm and instead contributes to tissue invasion and nutrient acquisition during pathogenesis. It functions as a receptor for plasminogen, a zymogen circulating in the host bloodstream. Plasminogen is cleaved by plasminogen activating proteins (including tissue-plasminogen activator and urokinase plasminogen activator) to form active plasmin, which functions as a serine protease. The physiological function of the human plasminogen/plasmin system is to facilitate tissue remodelling, cell migration, haemostasis and wound healing, and induce inflammation and angiogenesis. It fulfils these roles by degrading proteins in the extracellular matrix, fibrin and fibrinogen in blood clots and targeting components of the complement system (47). Inactive plasminogen is proposed to dock onto a lysine-rich motif located on the surface-expressed enolase of fungal pathogens. While bound, host plasminogen activators are capable of recognising, binding and cleaving plasminogen to produce the activated plasmin (45). The serine protease activity of plasmin becomes concentrated at the site of fungal infection and this can accelerate tissue invasion and disease progression. An in-depth study that focused on surface-expressed enolase from *A. fumigatus* confirmed its ability to interact with human immune regulators including factor H, factor-H-like protein 1 (FHL-1), C4b-binding protein (C4BP) and plasminogen (48). Whilst bound to *A. fumigatus* enolase, factor H, FHL-1 and C4BP retained their normal cellular activity and plasminogen remained accessible to plasminogen activator proteins. When swollen *A. fumigatus* conidia coated with human plasminogen were exposed to human A549 epithelial cells or an epithelial monolayer, cellular metabolic activity was reduced by 41% and cell retraction was reduced (48). These observations are consistent with the 'moonlighting' role of surface-expressed enolase and their role in facilitating tissue invasion.

1.8.2 Immunogenic activity of enolase

The secretion and presentation of enolase on the cell surface is a common feature of many fungal species (43,45). As a surface-exposed receptor, it is capable of eliciting an immunogenic response that may be exploited for vaccine development. Preliminary studies have demonstrated the suitability of using enolase as an immunogenic agent to acquire modest protective effects against candidiasis in a murine model. When mice were challenged with *C. albicans* following subcutaneous vaccination with recombinant *C. albicans* enolase, an elevated antibody response was induced and fungal burden in major organs was reduced compared to non-immunised mice (49). In an independent study, *C. albicans* enolase was expressed and presented on *Lactobacillus casei* cells, then orally administered to mice prior to challenging with a normally lethal dose of *C. albicans*. These immunised mice produced higher IgG antibody titres and had improved survival rates (50).

1.8.3 Targeting surface-expressed enolase

Designing a peptidomimetic or small molecule inhibitor that specifically targets the plasminogen docking site of enolase using a structure-guided approach is a possible avenue for new therapeutics. This approach requires intimate knowledge of the interactions between both proteins. Although high resolution crystal structures of the full-length human Type II plasminogen (4DUR) and fungal enolase from *S. cerevisiae* (3ENL) are already available, the interaction interface has not yet been defined (51,52). There have been previous studies examining the molecular interaction between plasminogen and bacterial enolases. Studies conducted using enolase from *S. pneumoniae* have suggested two possible binding sites - the L3 loop on the alpha-beta (αβ) barrel, a structurally flexible region, or the lysine-rich C-terminal tail of the protein (42,53,54). Superimposition of enolase from prokaryotic and eukaryotic sources including *E. coli* (1E9I), human (3B97), *Enterococcus hirae* (1IYX), lobster (1PDZ), *Streptococcus pneumoniae* (1W6T) and *S. cerevisiae* (3ENL) reveal an overall similar fold and existence of a conserved, but structurally different, L3 loop (Figure 2B) (51,54-58). Therefore, it is likely that the plasminogen binding site is also conserved between homologous enolase enzymes but the mode of binding may differ between species. Building upon this work, efforts must be focused on elucidating the plasminogen binding interface of fungal enolase and structurally characterising this site to guide rational inhibitor design.

Evidence of the immunogenicity of fungal enolase, particularly the *C. albicans* homologue, raises the possibility of using antibody therapy as an adjunct to antifungal therapy. Therapeutic antibodies specific for enolase could potentially reduce tissue invasion and fungal dissemination by interfering with interactions between fungal enolase and host plasminogen. Antibodies raised in mice against recombinant *C. albicans* or *A. fumigatus* enolase successfully recognised plasminogen-bound enolase from *Candida* spp. and *Aspergillus* spp., respectively. However, this antibody-enolase-plasminogen complex still allowed plasminogen activating proteins to access, cleave and activate bound plasminogen (45). The viability of antibody therapy requires improved antibodies against enolase that recognise specific epitopes which interfere with this plasminogen activation process. Monoclonal antibodies that inhibit plasminogen activation have already been successfully raised against human α-enolase. During the process of fibrinolysis, plasminogen binds to surface-expressed α-enolase on leukocytic cells and this results in an accelerated rate of plasminogen activation. The plasminogen activation associated with binding to leukocytic cell lines was reduced by approximately 90% when monoclonal α-enolase antibodies (MAb 11G1) were introduced (59). These experiments provide a proof-of-principle that disruption of an enolase-plasminogen interaction can affect plasmin activity and therefore, this approach may be feasible when applied to an antifungal therapeutic strategy.

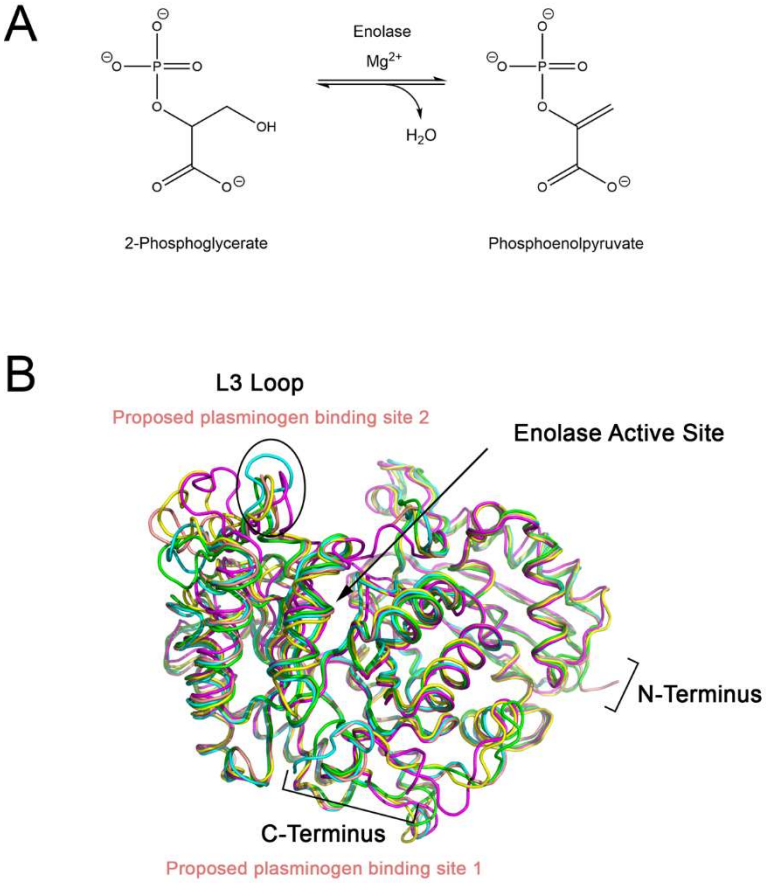


Figure 2: (A) Enolase catalyses the penultimate step of glycolysis, interconverting 2-phosphoglycerate (2-PG) and phosphoenolpyruvate (PEP) using magnesium as a cofactor. **(B)** X-ray crystal structures of enolase from different species have similar global folds. Enolase structures from *Escherichia coli* (1E9I; green), *Enterococcus hirae* (1IYX; cyan), lobster (1PDZ; yellow), *Streptococcus pneumoniae* (1W6T; salmon), and *Saccharomyces cerevisiae* (1ONE; magenta) were superimposed using *PyMOL* (51,54-58). The L3 loop, which shows the greatest structural deviation between the structures, is proposed to be the plasminogen binding loop for *S. pneumoniae*. Exposed lysine residues at the C-terminal end of enolase have also been proposed as a possible binding site.

1.9 Mannitol biosynthesis pathway

Mannitol biosynthesis is a crucial process that aids in fungal survivability and virulence. Mannitol-2-dehydrogenase (M2DH) and mannitol-1-phosphate 5-dehydrogenase (M1P5DH) are the primary biosynthesis enzymes that produce mannitol, an acyclic, six-carbon sterol alcohol (Figure 3). Present in all fungal structures, including mycelia, fruiting bodies and conidia, mannitol acts as a carbohydrate storage molecule, osmolyte and source of reducing power (60). During pathogenesis, mannitol production is increased to exploit its ability to quench reactive oxygen species (ROS) in order to resist host defences (60). The unique and diverse properties of mannitol, particularly in relation to human pathogenesis, have garnered considerable interest in the underlying mechanism of its biosynthesis and secretion during fungal infection.

1.9.1 Mannitol as a biomarker of fungal infection

The efficacy of antifungal therapy is dependent on the early detection and accurate identification of the invading fungal species. Traditional methods include culturing from patient samples, direct microscopical examination of fungal samples and histopathology testing of infected tissue (61). Although culturing is considered as the gold standard of diagnosis, the utility of this technique can suffer from the slow growth rates of filamentous fungi, environmental contamination and poor recovery of some fungal species from patient samples (61). Overall, these factors can delay adequate antifungal treatment, which can compromise the efficacy of these drugs. Therefore, non-culture-based methods of diagnosis, such as screening patient samples for biomarkers of fungal infection, offer a faster and less invasive alternative.
Several fungal pathogens, including *C. neoformans* and *Aspergillus* spp., have been shown to produce high levels of mannitol *in vitro* and in animal disease models of infection, highlighting the suitability for its use as a biomarker. In liquid cultures of twelve human isolates of *C. neoformans*, extracellular mannitol levels increased approximately 30-fold over a four day period in which rapid mannitol biosynthesis and secretion occurred predominantly in the stationary growth phase (62). Similarly, several *Aspergillus* species have also demonstrated an ability to produce mannitol from glucose *in vitro* (63). Closely mirroring these studies, mannitol levels in the cerebrospinal fluid (CSF) of a cortisone-treated rabbit disease model of cryptococcal meningitis also increased over time in response to infection with *C. neoformans* strain H99. Furthermore, there was a positive correlation identified between mannitol CSF levels and CSF colony forming units as well as CSF cryptococcal antigen titres, both of which are markers of infection severity (62). In an experimental model of aspergillosis using rats infected with *A. fumigatus*, high levels of mannitol were detected in liver tissue 12 hours post-infection and in serum 36 hours post-infection (64). In addition, a singularly surviving rat 48 hours post-infection also showed increased mannitol levels in lung, liver and kidney tissue (64).

From these animal models, the results indicate that infection of a human host with *C. neoformans* or *A. fumigatus* is also likely to induce upregulated mannitol production and secretion, localising in the CSF or serum, respectively. Analysis of the CSF from patients suffering from AIDS and cryptococcal meningitis confirm that mannitol content increases to sufficiently high levels to be easily detected and quantified using gas-liquid chromatography-linked mass spectroscopy (65). As humans lack an equivalent mannitol biosynthesis pathway, the detected mannitol is likely of fungal origin. However, mannitol content in the CSF could not be significantly correlated to cryptococcal antigen titres in humans, contrary to observations

made in the rabbit disease model. Taken together, these results provide valuable insights into the feasibility of using patient CSF and serum mannitol levels as a biomarker of fungal infection and presents possible methodologies that facilitate its use as a diagnostic tool.

1.9.2 Roles in virulence and protection against host defences

Fungal pathogens invading a human host experience multiple environmental stresses associated with occupying a new niche and chemical stresses invoked by the immune response. Sensing and adapting to abiotic stresses (temperature, pH), resisting phagocytosis from macrophages and neutralising ROS produced by neutrophils is critical for survival (66,67). Mannitol has been connected to many of these host-pathogen interactions associated with fungal survival and virulence and this provides a basis to understand the function of its rapid biosynthesis and secretion during infection.

Phenotypic characterisation of a *C. neoformans* mutant, which produces reduced levels of mannitol (Cn MLP) exemplifies the multitude of roles adopted by mannitol. The Cn MLP mutant showed similar growth rates, morphology and antifungal susceptibility as the wild-type strain, *C. neoformans* H99. However, Cn MLP had decreased tolerance to high temperatures and salinity (68). These findings are consistent with the proposed role of mannitol as an osmolyte in fungi (60). Although the loss of thermotolerance in Cn MLP cannot be directly linked to mannitol function, resistance to thermal stress in the fungus *Beauveria bassiana* has been partially attributed to trehalose, a metabolite produced from mannitol (Figure 3) (69). The Cn MLP isolate also demonstrated decreased resistance to oxidative stress, a common killing mechanism of the human immune system (70). Specifically, Cn MLP showed greater susceptibility to oxidative killing by polymorphonuclear neutrophils (PMNs) and ROS generated from an iron-hydrogen peroxide-iodide cytotoxic cell-free system, relative to the wild-type strain (70). Protection against ROS improved in a dose-dependent manner when PMNs were incubated with mannitol. Similarly, fungicidal activity was completely abolished when mannitol was added simultaneously to the ROS cell free system (70). In both scenarios, mannitol acted as an effective scavenger of ROS to confer protection against oxidative damage. In an experimental disease model, Cn MLP was 5,000-fold less virulent compared to *C. neoformans* H99 when intravenously injected into mice. Although both the mutant and wild-type strains proved to be lethal, 100% of mice succumbed to infection 51 days post-injection with *C. neoformans* H99, whereas 100% of mice injected with Cn MLP survived at the same time point (68). Together, the osmotolerance and thermotolerance conferred by intracellular mannitol and the protection against ROS by extracellular mannitol may aid fungal survival and virulence by resisting environmental and immune-induced stresses in the human body.

1.9.3 Roles in conidial stress response

Conidia are asexually-produced spores that are required for wide-spread fungal dispersal. Often dissipated through the air, conidia are inhaled into the respiratory system and at this developmental stage, fungal infection can be initiated (14). Conidia must be able to survive and adapt to this new environment to progress to subsequent germination stages and develop into a vegetative state. Mannitol has been implicated in the protection of conidia against these stress conditions in multiple species of fungi (68,71,72).

High concentrations of intracellular mannitol have been detected in conidia from *Aspergillus oryzae*, *Aspergillus clavatus* and *Aspergillus niger* (73-75). In order to assess mannitol function, a M1P5DH-null mutant of *A. niger* was generated. Disruption of M1P5DH resulted in a complete absence of mannitol in mycelia and a 30% reduction of intracellular mannitol content in conidia. Prolonged exposure to conditions that had no effect on the viability of wild-type conidia, such as high temperatures and oxidative stress induced by ROS, reduced the viability of M1P5DH-null conidia to 5% and 0.5%, respectively. To a lesser extent, these mutant conidia were more sensitive to freeze-drying and freeze-thawing processes. Supplementation of M1P5DH-null conidia with mannitol had a protective effect that reduced their sensitivity to heat and oxidative stress (71). These results indicate that mannitol is an important molecule that protects conidia against harsh environmental conditions.

In a separate study, the equivalent gene encoding M1P5DH in *Aspergillus fischeri* was deleted with unexpected effects. There was a severe reduction of mannitol in mycelia but an increase in mannitol and reduction of trehalose in its conidia (72). These conidia showed a modest increase in sensitivity to heat and oxidative stress, despite the higher mannitol content. Since trehalose and mannitol biosynthesis pathways are intrinsically linked and both sugars have been found in all stages of the fungal life cycle, it has been suggested that trehalose and mannitol may play overlapping roles in resistance to environmental stress (Figure 3) (72,76). There has also been growing interest in targeting components of the trehalose biosynthesis pathway for the development of novel antifungals (76). Thus, further investigation into the interplay between mannitol and trehalose in fungal survival is warranted to aid these efforts.

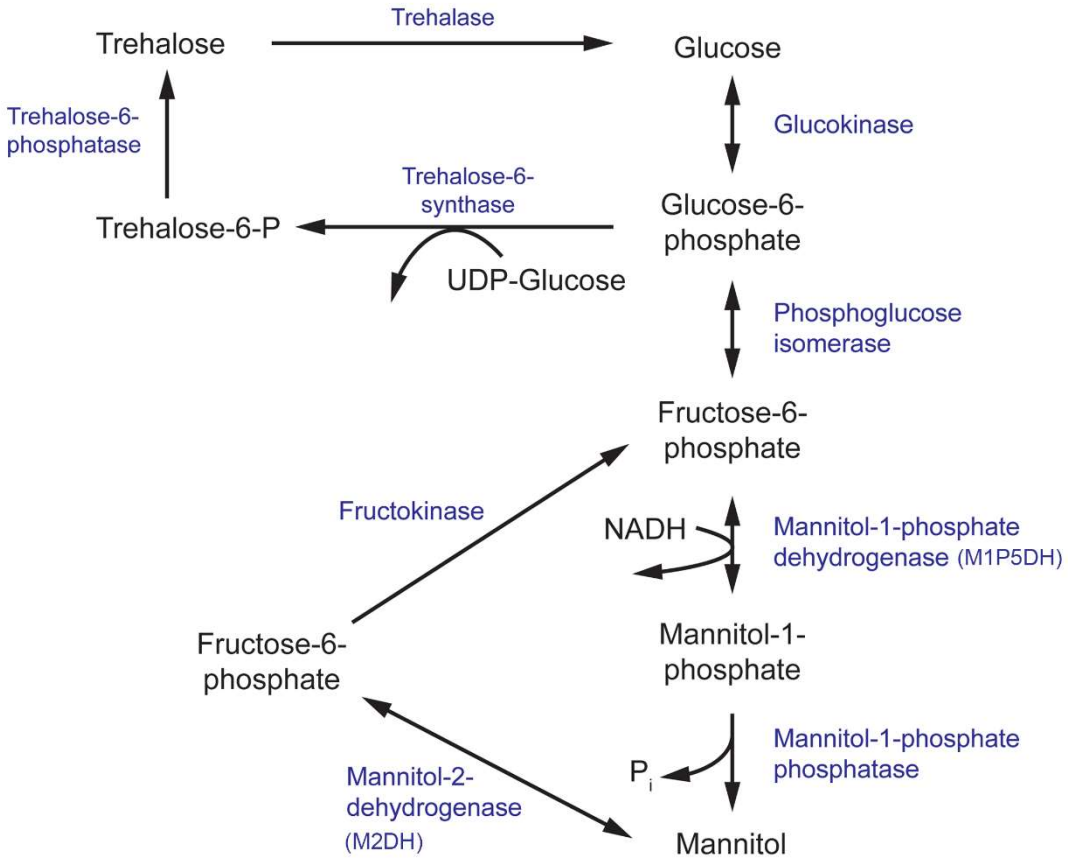


Figure 3: The mannitol and trehalose biosynthesis pathway in fungi. The primary biosynthesis enzymes of mannitol are mannitol-1-phosphate 5-dehydrogenase (M1P5DH) and mannitol-2-dehydrogenase (M2DH).

1.9.4 Targeting mannitol biosynthesis enzymes

Several lines of evidence that demonstrate the multifaceted roles of mannitol in the protection of fungi from stress conditions commonly experience during infection of a host exist (68,71,72).Accelerated mannitol production and secretion can confer thermotolerance and resistance to ROS generated from the human immune system needed to thrive in these harsh environments. Adding to the allure of targeting mannitol biosynthesis pathways in fungi is the fact that humans do not produce mannitol and therefore lack equivalent enzymes. The absence of a human homologue greatly simplifies the design and development of inhibitory compounds that specifically target these fungal enzymes. In turn, this will likely reduce the severity of off-target effects.

Efforts to target the mannitol biosynthesis enzyme should be focused on M1P5DH. In several species, knockout of the gene encoding M1P5DH (*mpdA*) and phenotypic characterisation of these mutants have shown significant decreases in intracellular mannitol levels (71,72). Although mannitol can be synthesised in an alternate pathway using M2DH, it appears that M1P5DH features in the predominant pathway. Hence, inhibition of M1P5DH is likely to have a greater effect on conidia viability and can be useful to halt initial stages of infection and dissemination.

Targeting the mannitol biosynthesis pathway can also be beneficial to delay the progression of existing infections. As demonstrated in the animal models infected with Cn MLP, decreased mannitol levels were associated with reduced virulence and ultimately, greater survival. By inhibiting the activity of these biosynthesis enzymes, mannitol production is reduced and this may increase fungal sensitivity to oxidative killing by components of the immune system.

1.10 Purine nucleotide biosynthesis pathway

Nucleobases, including purines and pyrimidines, are heterocyclic molecules that are found in all forms of life. They are obtained exogenously from the environment or can be synthesised endogenously. As they are involved in essential processes, including DNA and RNA synthesis, energy metabolism and signal transduction, the intracellular pool of nucleobases must be tightly maintained (77). Hence, disruption of the purine and pyrimidine biosynthesis pathway using small molecules has been explored as a potential approach to develop antimicrobials (37,78-80). *De novo* purine biosynthesis begins with phosphoribosylpyrophosphate (PRPP), which is ultimately converted to inosine monophosphate (IMP), a precursor to both adenine and guanine nucleobases, through a series of enzymatic reactions (Figure 4) (77). This section will focus on targeting enzymes responsible for the synthesis of guanine nucleobases.

1.10.1 Roles in virulence

Inosine monophosphate dehydrogenase (IMPDH) is the first committed and rate-limiting step in the guanosine triphosphate (GTP) biosynthesis pathway. Using NAD^+ as a cofactor, it catalyses the conversion of IMP to xanthosine monophosphate (XMP) via a dehydrogenase and hydrolysis reaction (81). IMPDH catalytic activity has been closely associated with microbial virulence, and has been the target of several drug discovery projects for antibiotics, antifungals,

antivirals, and anticancer therapeutics with great success (36,80,82,83). In terms of antifungals, there has been substantial progress in elucidating the mechanisms in which IMPDH contributes to *C. neoformans* survival and virulence. To first validate the rationality of IMPDH as an antifungal target, an IMPDH-deficient mutant of *C. neoformans* was produced and phenotypically characterised. This mutant exhibited reduced growth rates and impaired expression of virulence factors, resulting in smaller capsule sizes and reduced melanin expression (84). Importantly, these phenotypic defects could not be completely recovered with supplementation of exogenous guanine that may have been sourced from the human host during infection. Due to the direct involvement of IMPDH in virulence factor production, IMDPH-deficient *C. neoformans* was completely avirulent in a murine model of cryptoccoccus. As a result, fungal loads in the brain were cleared 3 days post-infection, total fungal clearance was achieved 2 weeks post-infection and ultimately, this resulted in a 100% survival rate 50 days post-infection (84). These experiments confirm that IMPDH is involved in essential processes associated with *C. neoformans* virulence and this presents a promising avenue for the design of antifungal compounds.

In the following step of the purine biosynthesis pathway, XMP produced by IMPDH is converted to guanosine monophosphate (GMP) by guanosine monophosphate synthase (GMP synthase). In a similar manner to IMPDH, GMP synthase has been extensively studied in *A. fumigatus*, *C. albicans* and *C. neoformans* and has been determined to be crucial in the survival and virulence of these fungal pathogens (37,85). A conditional knock-out of the gene encoding GMP synthase (*GUA1)* in *A. fumigatus* and *C. albicans*, and deletion in *C. neoformans*, inhibited growth on minimal media but was recovered when supplemented with exogenous guanine. These growth defects are consistent with the essential role of purines as components of DNA and RNA. Similar to the *in vivo* experiments conducted with IMPDH-deficient *C. neoformans*, disruption of *GUA1* in *A. fumigatus*, *C. albicans* and *C. neoformans* also rendered these strains avirulent in a murine model of infection (37,85). Phenotypic characterisation of Δ*GUA1 C. neoformans* revealed significant differences in the production of key virulence factors that may account for the abolished virulence. Despite supplementation with exogenous guanine, Δ*GUA1 C. neoformans* mutants produced less capsule, showed delayed melanin production and lacked detectable protease activity relative to their wild-type counterparts. Collectively, these virulence factors function to aid survival by resisting host defences and fungal dissemination. Specifically, the capsule protects against phagocytosis, melanin resists oxidative damage and secretion of proteases facilitates host tissue degradation for nutrient acquisition. Intriguingly, this effect was more pronounced at the optimum human body temperature of 37°C (85). Since *GUA1* is a determinant of virulence in several of these clinically-prominent fungal pathogens, inhibition of GMP synthase may be an ideal approach to develop antifungal drugs with broad spectrum activity.

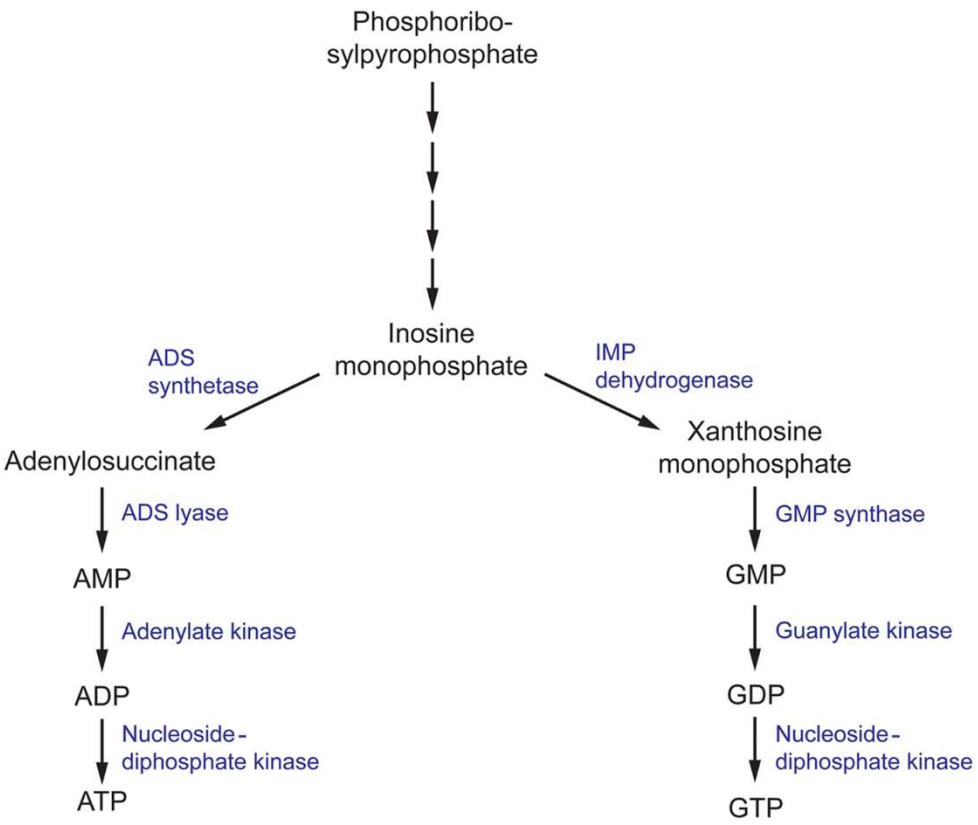


Figure 4: The purine metabolism pathway in fungi. ADS, (adenylosuccinate) synthetase; ADS, (adenylosuccinate) lyase; IMP, (inosine monophosphate) dehydrogenase; GMP, (guanosine monophosphate) synthase

1.10.2 Roles in survival

In the final steps of the purine biosynthesis pathway, GMP must be converted to an accessible form for its use in DNA and RNA synthesis and signal transduction. GMP is first converted to guanosine diphosphate and then nucleoside-diphosphate kinase (NDK) covalently attaches a phosphate moiety to form GTP (77). NDK is a promiscuous enzyme that can use all nucleoside diphosphates as substrates, with varying levels of selectivity, to generate their respective nucleoside triphosphates (86). Hence it is also involved in the adenine nucleoside synthesis pathway to produce adenosine triphosphate (ATP) as well as the synthesis of the remaining building blocks of DNA and RNA. Consistent with these fundamental processes, in *A. nidulans*, NDK is involved in various aspects of the life cycle, including hyphal growth and conidia production, which is necessary for its growth and dispersal (87). Similarly in *A. flavus*, disruption of one of the two copies of *ndk* impairs the development of conidia and sclerotia and affects plant virulence in a maize and peanut seed model (88). Identified as an essential gene for survival in *A. fumigatus* and *A. nidulans*, NDK has been identified as a promising target for antifungals against *Aspergillus* spp. (87,89).

1.10.3 Targeting purine biosynthesis pathways

Achieving species-selective targeting of *de novo* purine biosynthesis enzymes is an important consideration in the drug design process. This is especially important since bacteria, fungi and human share commonality in many of these enzymes (77,90,91). Thus, intimate structural and kinetic knowledge of the fungal and human homologues of these enzymes will be invaluable in the design of inhibitors with high potency and selectivity. Currently, structures of human IMPDH, isoforms I (92) and II (93), GMP synthase (94) and NDK (95) have been solved as well as structures of fungal IMPDH (*C. neoformans*) (84) and NDK (*A. fumigatus* and *A. flavus*) (86,96). Availability of these structures allows identification of key active site residues that differ between fungal and human enzymes and this may be integral in the design of highly selective inhibitors.

Knowledge of the enzyme structure and catalytic mechanism aids in identifying the most 'druggable' binding pocket. In the case of IMPDH, there are three possible sites: (i) the natural substrate IMP site, (ii) the co-factor NAD^+ site, or (iii) an allosteric site (79). On-going drug discovery projects targeting IMPDH from *Mycobacterium tuberculosis* and *Cryptosporidium parvum* have shown success in focusing on the co-factor binding site due to its divergence from the mammalian equivalent (80,97). In the case of *C. parvum* IMPDH, a high-throughput screening approach was employed to specifically target the NAD^+ site. Ten compounds, from a pool of 134 preliminary hits, were determined to be reversible inhibitors that were 9-fold to 400-fold more selective for the parasitic IMPDH over the human IMPDH isoform II (97). These experiments provide a proof of concept that IMPDH can be effectively and selectively targeted by small molecule inhibitors. When considering IMPDH from *C. neoformans*, a similar screening approach targeting the co-factor binding site may also be feasible due to its structural and functional deviation from human IMPDH (84). A structure of human IMPDH isoform II bound to selenazole-4-carboxyamide-adenine dinucleotide, an analogue of NAD^+, reveals differences in the residues that comprise the co-factor binding pocket between *C. neoformans* and human IMPDH. It can be inferred that the interaction between IMPDH and NAD^+ is stabilised by a π-π stacking interaction between the adenine ring of NAD^+ and R267 and Y296 of *C. neoformans* IMPDH, which corresponds to R253 and Y282 in human isoform I or H253 and F282 in human isoform II, respectively (Figure 5A) (84). Although many residues within

the binding pocket are invariant between *C. neoformans* and human isoforms I and II, the differences are likely to be sufficient to achieve species-selective inhibition. A high-throughput screen against IMPDH from *C. neoformans* has identified promising candidates (36). From a chemical library of 114,000 drug-like compounds, three 3-((5-substituted)-1,3,4-oxadiazol-2-yl)thio benzo[*b*]thiophene 1,1-dioxides were identified as having inhibitory activity against *C. neoformans* IMPDH. Through chemical modification, several analogues were then synthesised to assess the structure-activity relationships. Although many of these analogues inhibited *C. neoformans* IMPDH *in vitro* and had whole cell activity, they were also cytotoxic to human cell lines (36). These results illustrate the complexities associated with targeting pathways shared between fungi and humans. However, there is also great opportunity to build upon this work by trialling modifications that improve selectivity that retain the antifungal properties of the compound whilst reducing cytotoxic effects.

Similarly to IMPDH, inhibition of fungal GMP synthase may be achieved effectively by high-throughput screening of chemical libraries or fragment screening. Since GMP synthase is a bifunctional protein with two catalytic domains working in tandem to fulfil its function, it is imperative to select the most feasible binding pocket to target. The glutamine amidotransferase (GAT) domain releases ammonia from glutamine, which is shuttled to the ATP pyrophosphatase (ATPPase) domain where XMP is converted to GMP via an adenyl-XMP intermediate. There are pre-existing inhibitors that target the GAT domain such as acivicin (ACI) and 6-diazo-5-oxo-L-norleucine (DON) that inadvertently inhibit GMP synthase activity (98,99). However, targeting this catalytic module of GMP synthase would likely lead to non-specific and off-target effects due to the prevalence of GAT domains in other enzymes. Therefore, a more promising endeavour is to study the XMP- and ATP-binding pockets located within the ATPPase domain to achieve potent inhibition and high selectivity for the fungal enzyme over the mammalian equivalent. A novel compound, denoted ECC1385, identified from a synthetic compound library has been shown to inhibit *C. albicans* and *C. neoformans* GMP synthase activity *in vitro* and exhibits whole cell activity against a broad spectrum of fungi and bacteria, including multiple species of *Candida*, *A. fumigatus* (MF5668), *C. neoformans* (MY2062) and *Staphylococcus aureus* (MB2865) (37). Mechanism-of-action studies of ECC1385 have revealed that inhibition of GMP synthase is achieved through a different mode of action from ACI and DON. These data suggest that small chemical compounds that target the activity of the ATPPase domain is possible and presents a promising avenue to pursue.

As NDK has maintained its essential role in NTP biosynthesis throughout all domains of life, there is high structural and functional conservation between different species. This conservation is reflected in the largely invariant composition of binding pocket residues between fungal and human homologues (Figure 5B). However, key differences are revealed when the crystal structures of adenosine diphosphate (ADP)-bound NDK from human (2HVD), *A. fumigatus* (6XP7) and *A. flavus* (6K3H) are superimposed (Figure 5B) (86,88,100). Specifically, in the nucleoside binding pocket, Lys57 and C116 in both fungal NDK enzymes are replaced by Arg58 and I117 in the human structure, respectively. These discrepancies present an opportunity to selectively target the catalytic site and this may be beneficial in the development of a pan inhibitor with broad spectrum activity against multiple pathogenic fungi. Azidothymidine (AZT), also marketed as RETROVIR, is an antiretroviral medication currently used to treat HIV (101). It has been shown that AZT can inhibit NDK from *A. flavus in vitro* and exhibits antifungal activity *in vivo* (88). Docking of AZT to the crystal structure of *A. flavus* NDK predicts that it acts as a competitive inhibitor and interacts with residues R104, H117 and D120 in the catalytic site. Since AZT is an FDA-approved drug, there is potential to repurpose it as an antifungal drug or to use the AZT scaffold and improve upon its antimycotic properties.

An alternate approach may be explored to inactivate fungal NDK preferentially over the human homologue using non-competitive inhibitors. Allosteric sites are likely to show more variation between species, compared to the catalytic site, and therefore, development of a non-competitive inhibitor may be a more effective approach. Thorough kinetic analysis has revealed that ebselen, a cysteine-modifying compound, acts as a non-competitive inhibitor of human NDK secreted by airway epithelial cells (102). In this context, it is not used in an antifungal application. Instead, it is used to obtain stable and accurate measurements of nucleotides in airway surface liquid by disrupting nucleotide interconversion without the use of a nucleotide-derived inhibitor. Identifying allosteric sites in NDK from fungal pathogens, comparing them to the human homologue and exploiting their differences may be imperative to simultaneously achieve potent inhibition and high selectivity. Hence, it would be advantageous to explore non-competitive inhibitors of NDK that bind an allosteric site and influence catalytic activity distally.

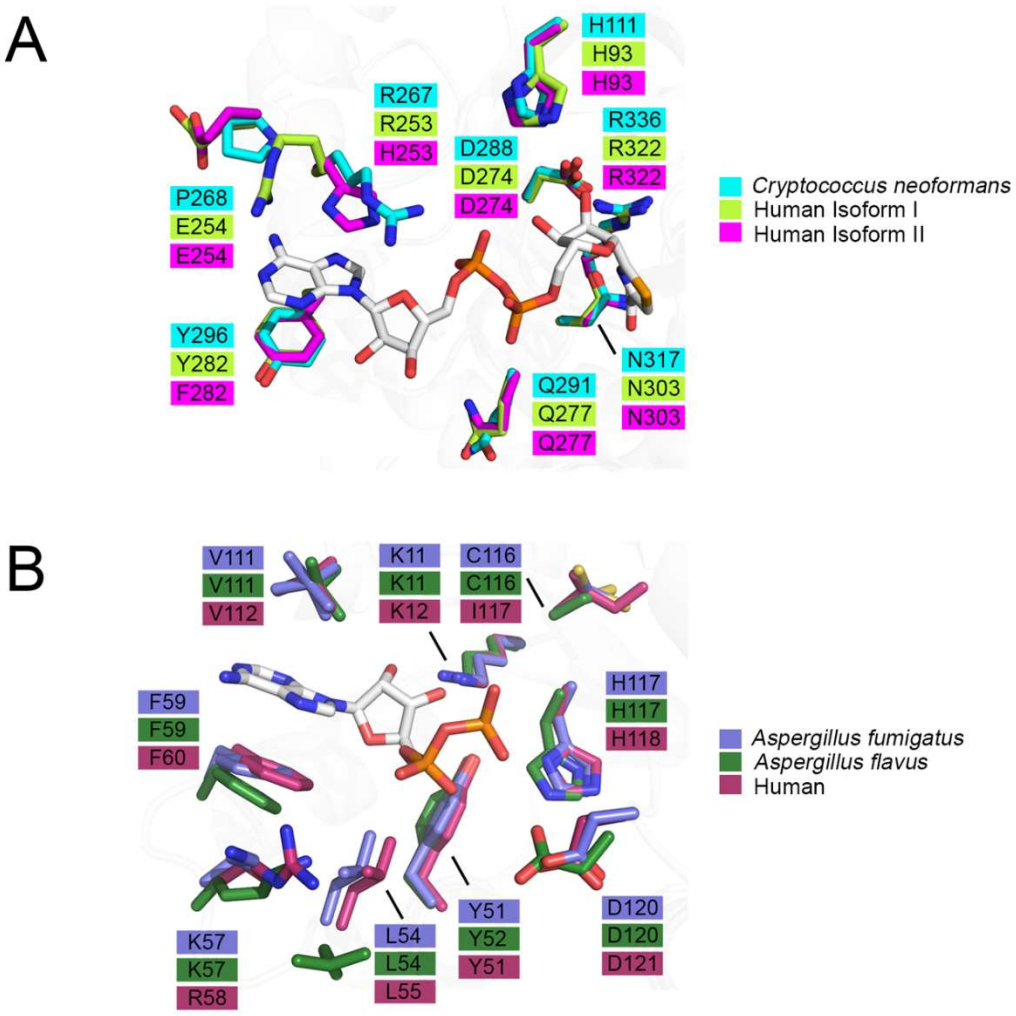


Figure 5: Binding pocket composition analysis reveals divergent residues between human and fungal homologues that may be exploited to achieve specificity. **(A)** Residues comprising the binding pocket of inosine monophosphate dehydrogenase from *Cryptococcus neoformans* (4AF0, cyan), human isoform I (1JCN, green), human isoform II (1B3O, magenta) and selenazole-4-carboxyamide-adenine dinucleotide, an NAD^+ analogue, are represented as sticks (84,92,93). **(B)** Residues comprising the binding pocket of nucleoside diphosphate-kinase from *Aspergillus fumigatus* (6XP7, purple), *Aspergillus flavus* (6K3H, dark green) and human (2HVD, maroon) and ADP, a natural substrate, are represented as sticks (86,88,100).

1.11 Aims and scope of the project

As the number of immunocompromised patients continues to grow, invasive fungal infections are becoming more pervasive. Due to inherent limitations in antifungal drug treatments and the emergence of resistance, the limited spectrum of currently available antifungals is a major barrier to effectively treat systemic fungal infections. Consequently, there is a need to identify and validate novel antifungal drug targets. Targeting eukaryotic proteins that are common to both fungi and humans is a difficult but necessary approach since the number of fungal enzymes that lack a human homolog and are essential for survival or virulence are scarce. Hence, there has been an increased interest in exploiting subtle structural and functional differences of metabolic enzymes shared by both fungi and humans.

This project focuses on the structural and biochemical characterisation of targets from the aforementioned pathways including nucleoside diphosphate-kinase, guanosine monophosphate synthase, mannitol-2-dehydrogenase and enolase from *A. fumigatus*. Data generated from X-ray crystallography, native mass spectrometry, analytical size-exclusion chromatography and *in vitro* activity assays were used to gain a better understanding of how these fundamental processes contribute to fungal survival and virulence. Furthermore, the structural data was used to probe targetable sites for rational drug design efforts and identify species-specific differences, if a human homologue existed, to achieve selectivity. Together, these experiments have laid the foundation for antifungal drug discovery projects that target unconventional, but promising, pathways.

Chapter 2 (Publication):

Nucleoside selectivity of *Aspergillus fumigatus* nucleoside-diphosphate kinase

Chapter 2 Preface

Chapter 2 is a manuscript published in the Federation of European Biochemical Societies (FEBS) Journal. In this paper, we have characterised the structure and function of nucleoside diphosphate kinase from *A. fumigatus*. Using X-ray crystallography and *in vitro* enzyme assays, we have also explored the structure-activity relationships between the six naturally occurring nucleoside triphosphate substrates used by nucleoside diphosphate kinase, I performed all X-ray crystallography experiments and *in vitro* enzyme assays. I performed the data analysis for these experiments, prepared the figures included in the manuscript, deposited structure factors and coordinate factors into the Protein Data Bank and prepared the manuscript.

Statement of Authorship

Title of Paper	Nucleoside selectivity of *Aspergillus fumigatus* nucleoside-diphosphate kinase
Publication Status	☑ Published ☐ Accepted for Publication ☐ Submitted for Publication ☐ Unpublished and Unsubmitted work written in manuscript style
Publication Details	This publication describes the structural determinants that govern nucleoside selectivity in nucleoside diphosphate-kinase from *Aspergillus fumigatus* using structural and kinetic data.

Principal Author

Name of Principal Author (Candidate)	Stephanie Nguyen		
Contribution to the Paper	Project conceptualisation, planned and performed experiments, analysed data, prepared, reviewed and edited the manuscript		
Overall percentage (%)	65		
Certification:	This paper reports on original research I conducted during the period of my Higher Degree by Research candidature and is not subject to any obligations or contractual agreements with a third party that would constrain its inclusion in this thesis. I am the primary author of this paper.		
Signature		Date	17/07/21

Co-Author Contributions

By signing the Statement of Authorship, each author certifies that:

i. the candidate's stated contribution to the publication is accurate (as detailed above);

ii. permission is granted for the candidate in include the publication in the thesis; and

iii. the sum of all co-author contributions is equal to 100% less the candidate's stated contribution.

Name of Co-Author	Dr. Blagojce Jovcevski		
Contribution to the Paper	Planned and performed experiments, analysed data, reviewed and edited the manuscript		
Overall Percentage (%)	15		
Signature		Date	20/07/21

Name of Co-Author	Assoc. Prof. Tara Pukala		
Contribution to the Paper	Planned experiments, reviewed and edited the manuscript		
Overall Percentage (%)	10		
Signature		Date	18/7/21

Name of Co-Author	Dr. John B. Bruning		
Contribution to the Paper	Project conceptualisation, planned experiments, reviewed and edited the manuscript		
Overall Percentage (%)	10		
Signature		Date	

The FEBS Journal

FEBS PRESS science publishing by scientists

Nucleoside selectivity of *Aspergillus fumigatus* nucleoside-diphosphate kinase

Stephanie Nguyen[1], Blagojce Jovcevski[2,3], Tara L. Pukala[2] and John B. Bruning[1]

1 Institute of Photonics and Advanced Sensing (IPAS), School of Biological Sciences, The University of Adelaide, Australia
2 Adelaide Proteomics Centre, School of Physical Sciences, The University of Adelaide, Australia
3 School of Agriculture, Food and Wine, The University of Adelaide, Australia

Keywords
Aspergillus fumigatus; nucleoside biosynthesis; nucleoside-diphosphate kinase; structural biology; substrate selectivity

Correspondence
J. B. Bruning, Institute of Photonics and Advanced Sensing (IPAS), School of Biological Sciences, The University of Adelaide, North Terrace, Adelaide, SA 5005, Australia
Tel: 8313 5218
E-mail: john.bruning@adelaide.edu.au

(Received 28 August 2020, revised 6 October 2020, accepted 19 October 2020)

doi:10.1111/febs.15607

Aspergillus fumigatus infections are rising at a disconcerting rate in tandem with antifungal resistance rates. Efforts to develop novel antifungals have been hindered by the limited knowledge of fundamental biological and structural mechanisms of *A. fumigatus* propagation. Biosynthesis of NTPs, the building blocks of DNA and RNA, is catalysed by NDK. An essential enzyme in *A. fumigatus*, NDK poses as an attractive target for novel antifungals. NDK exhibits broad substrate specificity across species, using both purines and pyrimidines, but the selectivity of such nucleosides in *A. fumigatus* NDK is unknown, impeding structure-guided inhibitor design. Structures of NDK in unbound- and NDP-bound states were solved, and NDK activity was assessed in the presence of various NTP substrates. We present the first instance of a unique substrate binding mode adopted by CDP and TDP specific to *A. fumigatus* NDK that illuminates the structural determinants of selectivity. Analysis of the oligomeric state reveals that *A. fumigatus* NDK adopts a hexameric assembly in both unbound- and NDP-bound states, contrary to previous reports suggesting it is tetrameric. Kinetic analysis revealed that ATP exhibited the greatest turnover rate ($321 \pm 33.0\ s^{-1}$), specificity constant ($626 \pm 110.0\ mM^{-1}{\cdot}s^{-1}$) and binding free energy change ($-37.0 \pm 3.5\ kcal{\cdot}mol^{-1}$). Comparatively, cytidine nucleosides displayed the slowest turnover rate ($53.1 \pm 3.7\ s^{-1}$) and lowest specificity constant ($40.2 \pm 4.4\ mM^{-1}{\cdot}s^{-1}$). We conclude that NDK exhibits nucleoside selectivity whereby adenine nucleosides are used preferentially compared to cytidine nucleosides, and these insights can be exploited to guide drug design.

Enzymes
Nucleoside-diphosphate kinase (EC 2.7.4.6).

Database
Structural data are available in the PDB database under the accession numbers: Unbound-NDK (6XP4), ADP-NDK (6XP7), GDP-NDK (6XPS), IDP-NDK (6XPU), UDP-NDK (6XPT), CDP-NDK (6XPW), TDP-NDK (6XPV).

Abbreviations
C_α, alpha carbon; IDP, ITP, inosine di- and tri-phosphate; IM-MS, ion-mobility mass spectrometry; NDK, nucleoside-diphosphate kinase; NH_4OAc, ammonium acetate; PB, phosphate buffer; RMSD, root mean square deviation; SEC, size-exclusion chromatography.

Introduction

Aspergillus fumigatus is a saprophytic fungus commonly found in the environment and is also the causative agent of devastating human infections such as aspergillosis. Within its ecosystem, it plays an essential role in nutrient recycling by contributing to the decomposition of organic matter. It primarily reproduces through asexual reproduction, forming small and hydrophobic conidia that are rapidly disseminated through the air [1]. Although harmless to humans with healthy immune systems, spores can colonise lung tissue in immunocompromised hosts and, in serious cases, develop into an invasive and systemic infection known as invasive aspergillosis [2]. This disease is estimated to affect over 300 000 individuals annually, and the associated mortality rate is approximated to be between 30% and 80%, depending on a number of patient factors [3,4]. Treatment requires aggressive antifungal therapy that is often problematic due to the inherent limitations of pre-existing antifungal drugs and the rise of antifungal resistance [5]. In an effort to better understand the nuances of fungal biology in the context of human disease and identify potential targets for antifungals, there has been growing interest in the study of essential metabolic pathways. One of which is the fungal de novo purine biosynthesis pathway with a specific interest in the purines.

Purine biosynthesis and salvage pathways have been extensively studied in clinically prominent pathogenic fungi including *Cryptococcus neoformans, Candida albicans* and *A. fumigatus* [6–8]. Disruption of the purine biosynthesis pathway by genetic knockout of key enzymes is associated with severe attenuation of virulence, or in some cases, complete avirulence [9]. Comprising of adenine, guanine and inosine nucleobases, purines are constructed from a base scaffold featuring a pyrimidine ring fused to an imidazole ring.

The enzyme nucleoside-diphosphate kinase (NDK) catalyses the final step of the purine and pyrimidine biosynthesis pathway by producing nucleoside triphosphates (NTPs) from nucleoside diphosphates (NDPs) (Fig. 1A). It is a highly conserved enzyme that is ubiquitous in the living world and primarily functions in the maintenance of intracellular NTPs. These essential building blocks are required for DNA and RNA synthesis and act as mediators of signal transduction pathways. Regulated synthesis of NTPs is imperative to normal cellular function and survival of fungi in the natural environment which becomes increasingly important when considering the differences in nutrient availability within the human host. Air-dry soil composition studies have shown an abundance of guanine

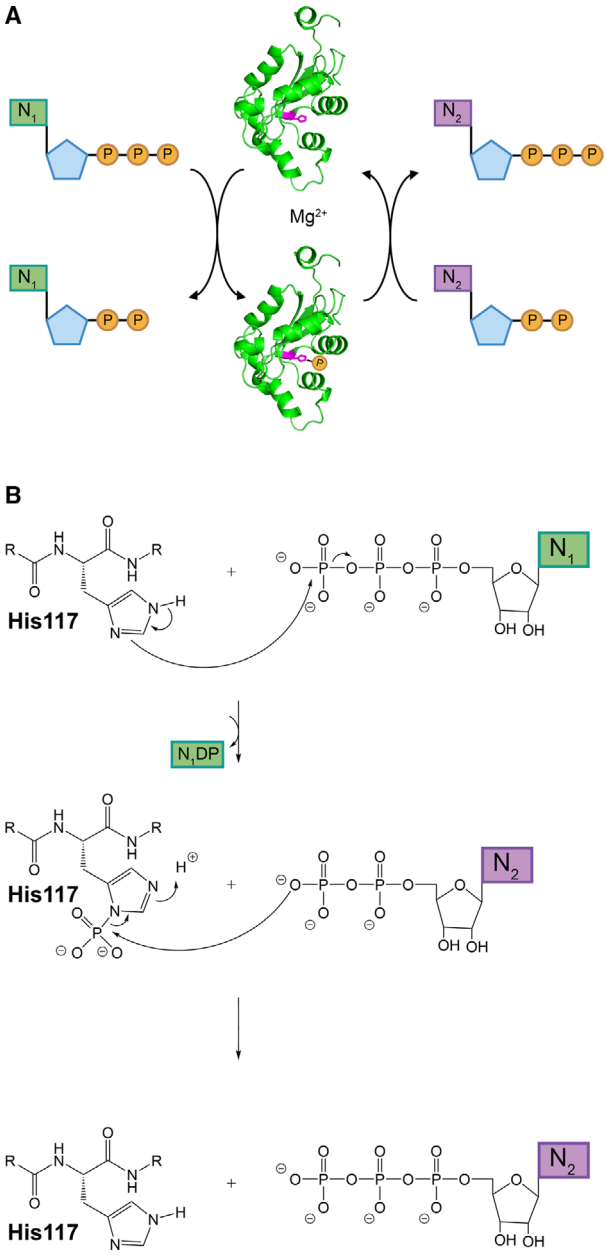


Fig. 1. Catalytic mechanism of nucleoside triphosphates by NDK. (A) Nucleoside-diphosphate kinase catalyses the synthesis of nucleoside triphosphates through a bi-bi ping-pong mechanism. (B) The active site His117 facilitates nucleophilic attack of the γ-phosphate of a nucleoside triphosphate (N_1) forming the phosphohistidine reaction intermediate. Once the corresponding nucleoside diphosphate (N_1DP) is released, the phosphate is then transferred to an acceptor NDP (N_2DP) to generate the corresponding NTP (N_2TP).

(41 mol %) and decreasing availabilities of cytosine (21 mol %), adenine (19 mol %), thymine (8 mol %) and uracil (5 mol %) [10]. However, once *A. fumigatus*

enters and colonises the human host, it experiences a greater abundance of cytosine (6.4 μM) and lower availability of adenine (4.4 μM) and guanine (0.4 μM) in human extracellular fluids [11]. In order to survive and thrive, *A. fumigatus* must adapt to a change in nucleoside availability.

In both *A. fumigatus* and *A. nidulans* (a closely related fungus to *A. fumigatus*), NDK has been identified as an essential gene, indicating that it may be an excellent target for novel antifungals [12,13]. NDK has been studied as a potential target to develop antiparasitic treatments for leishmaniasis and trypanosomiasis, both of which are caused by protozoan parasites [14]. In terms of antifungal therapies, *Aspergillus flavus* NDK has been similarly investigated [15]. Consistent with its integral role in DNA synthesis, NDK from *Aspergillus nidulans* has been implicated in essential processes associated with the cell cycle including hyphal growth whereas NDK from *A. flavus* has been shown to contribute to plant virulence [13,15]. In both *A. nidulans* and *A. flavus*, NDK is associated with conidiation [13,15]. The integral roles fulfilled by NDK in fungal survival and virulence strengthen its potential as a novel antifungal target. Since a human NDK homologue exists (64% amino acid sequence identity with *A. fumigatus* NDK), developing compounds with high selectivity for the fungal equivalent must be prioritised in drug design efforts. As the human NDK homologue has already been structurally and kinetically characterised, further investigation of the structure and function of NDK from individual fungal species will ascertain species-specific differences that will be imperative in guiding inhibitor design [16,17].

NDK operates through a bi-bi ping-pong mechanism with a Mg^{2+} ion cofactor. A donor NTP is the first substrate to bind. The active site His117 facilitates nucleophilic attack of the terminal phosphate, releasing the corresponding NDP and forming the phosphorylated NDK reaction intermediate (Fig. 1B) [18]. The phosphate is subsequently transferred to the acceptor NDP upon binding to produce the corresponding NTP [19]. Interestingly, homologues of NDK from different species (including prokaryotes, eukaryotes and viruses), exhibit promiscuity in terms of their ability to bind and utilise different NTP donors and NDP acceptors with differing degrees of selectivity [17,20,21]. Although the catalytic mechanism is well-established, the determinants of NDK substrate selectivity and nucleoside preference have not been explored in fungi. Since the purine and pyrimidine availability in both soil and within the human host is significantly different, understanding the basis of nucleoside selectivity may clarify how purine metabolism contributes to fungal virulence.

Here we present the first study to elucidate the determinants of nucleoside selectivity in fungal NDK by performing a comprehensive analysis of the structure and enzymology of *A. fumigatus* NDK. The structure of *A. fumigatus* unbound-NDK has been solved previously only in the unbound state, forming a crystallographic hexamer in the asymmetric unit but a tetramer in solution [22]. In our work, we have solved crystal structures of unbound and NDP-bound (ADP, GDP, IDP, UDP, TDP and CDP) *A. fumigatus* NDK and assessed the oligomeric state of the enzyme using analytical-size exclusion chromatography (SEC) and native mass spectrometry. We present compelling evidence that *A. fumigatus* NDK forms a hexameric biological assembly, a finding which contradicts that of an earlier study [22]. It is vital to discern the true oligomeric state of *A. fumigatus* NDK amongst conflicting arguments as the quaternary structure is intimately tied to enzyme function. At the time of writing, there are 51 structures of NDP-bound NDK deposited in the Protein Data Bank (PDB) in which all NDPs assume very similar binding modes. For the first time in any NDK of any species, we present evidence of a unique binding mode adopted by CDP and TDP when bound to *A. fumigatus* NDK and elucidate conformational changes associated with substrate binding. By combining high resolution structural data with kinetic characterisation studies, we provide a structural rationale that demonstrates how substrate binding modes influence binding, and ultimately, nucleoside selectivity in *A. fumigatus* NDK.

Results

Structural features and quaternary structure analysis of NDK

To probe the structural features of unbound-NDK versus NDK bound to nucleosides, we utilised X-ray crystallography. In total, seven structures were solved in with resolutions ranging from 1.6 to 2.3 Å (data processing and refinement statistics can be found in Table 1. The overall structure of *A. fumigatus* closely resembles a ferredoxin-like fold, characterised by an antiparallel β-sheet core (β1–β4) surrounded by eight α-helices (Fig. 2A). In addition, NDK features four 3_{10} helices (Fig. 2A). Each subunit has one active site capable of binding an NTP and a single Mg^{2+} ion, a cofactor required for catalysis. Once the first half reaction is completed and the phosphohistidine

Table 1. Data collection and model refinement statistics.

	Unbound-NDK	NDK-ADP	NDK-GDP	NDK-IDP	NDK-UDP	NDK-CDP	NDK-TDP
Wavelength (Å)	0.954	0.954	0.954	0.954	0.954	0.954	0.954
Resolution range (Å)	49.75–2.0 (2.071–2.0)	45.78–2.2 (2.279–2.2)	45.45–1.644 (1.703–1.644)	42.27–1.9 (1.968–1.9)	46.21–2.3 (2.382–2.3)	46.47–1.9 (1.968–1.9)	47.82–2.3 (2.382–2.3)
Space group	P 2_1 2_1 2_1	P 2_1 2 2_1	P 2_1 2 2_1	P 21 2 21	P 2_1 2_1 2_1	P 2_1 2_1 2_1	P 21 21 21
Unit cell dimensions *a, b, c* (Å)	93.4 101.6 114.1	53.1 71.8 118.9	53.0 70.3 119.2	52.9 70.307 119.133	66.6 126.8 134.9	90.7 101.8 113.9	67.223 126.073 136.062
α, β, γ (°)	90 90 90	90 90 90	90 90 90	90 90 90	90 90 90	90 90 90	90 90 90
Total reflections	768 855 (57 758)	280 780 (24 262)	785 733 (73 566)	499 739 (48 964)	722 513 (73 806)	1 098 985 (108 805)	735 282 (74 725)
Unique reflections	68 664 (6294)	23 361 (2204)	54 666 (5262)	35 789 (3498)	51 581 (5082)	83 485 (8233)	52 151 (5143)
Multiplicity	11.2 (9.2)	12.0 (11.0)	14.4 (14.0)	14.0 (14.0)	14.0 (14.5)	13.2 (13.2)	14.1 (14.5)
Completeness (%)	92.8 (86.1)	98.1 (94.8)	99.1 (97.0)	99.7 (99.8)	99.5 (99.9)	99.8 (99.4)	99.8 (99.9)
Mean *I*/sigma(*I*)	21.1 (6.9)	18.7 (9.4)	21.9 (3.8)	20.7 (5.3)	16.2 (5.0)	13.5 (4.9)	14.2 (3.3)
Wilson B-factor	12.5	13.7	18.2	21.0	30.0	18.7	35.6
R_{merge}[a]	0.090 (0.33)	0.10 (0.28)	0.077 (0.69)	0.091 (0.53)	0.12 (0.49)	0.15 (0.63)	0.14 (0.87)
R_{meas}[b]	0.094 (0.35)	0.11 (0.29)	0.080 (0.72)	0.095 (0.55)	0.12 (0.50)	0.15 (0.65)	0.14 (0.90)
R_{pim}[c]	0.027 (0.11)	0.030 (0.082)	0.021 (0.19)	0.025 (0.15)	0.033 (0.13)	0.042 (0.18)	0.038 (0.23)
CC1/2[d]	0.998 (0.956)	0.998 (0.965)	0.999 (0.942)	0.999 (0.959)	0.999 (0.974)	0.998 (0.938)	0.998 (0.939)
CC*	1 (0.989)	0.999 (0.991)	1 (0.985)	1 (0.99)	1 (0.993)	0.999 (0.984)	1 (0.984)
Reflections used in refinement	68 654 (6294)	23 325 (2204)	54 512 (5258)	35 668 (3491)	51 360 (5075)	83 477 (8233)	52 040 (5138)
Reflections used for R-free	3462 (351)	1115 (114)	1989 (192)	1797 (171)	2577 (230)	4276 (402)	2606 (233)
R_{work}[e]	0.13 (0.14)	0.18 (0.28)	0.16 (0.22)	0.18 (0.24)	0.18 (0.22)	0.13 (0.16)	0.19 (0.27)
R_{free}[f]	0.19 (0.22)	0.24 (0.37)	0.21 (0.26)	0.23 (0.31)	0.23 (0.33)	0.18 (0.22)	0.25 (0.32)
CC_{work}	0.976 (0.956)	0.958 (0.797)	0.974 (0.932)	0.972 (0.892)	0.968 (0.917)	0.977 (0.940)	0.964 (0.840)
CC_{free}	0.948 (0.914)	0.909 (0.715)	0.950 (0.904)	0.943 (0.775)	0.952 (0.791)	0.961 (0.876)	0.935 (0.693)
Number of non-hydrogen atoms	9262	4413	4666	4614	8628	9380	8480
Macromolecules	7281	3563	3716	3708	7158	7394	7138
Ligands	–	81	87	84	152	50	175
Solvent	1981	769	863	822	1318	1936	1167
Protein residues	929	456	459	459	923	932	922
RMS_{bonds} (Å)	0.006	0.002	0.006	0.002	0.002	0.006	0.002
RMS_{angles} (°)	0.78	0.62	0.97	0.54	0.56	0.85	0.48
Ramachandran favored (%)	97.82	97.33	98.23	98.01	97.80	97.72	97.69
Ramachandran allowed (%)	1.53	2.22	1.10	1.32	1.65	1.63	1.98
Ramachandran outliers (%)	0.65	0.44	0.66	0.66	0.55	0.65	0.33
Rotamer outliers (%)	0.51	1.31	1.23	1.99	1.71	1.01	1.45
Clashscore	2.20	3.20	1.99	2.27	2.15	3.17	2.42
Average B-factor	16.87	16.54	26.13	24.56	32.82	24.35	35.87
Macromolecules	13.22	15.25	23.06	22.56	31.03	20.37	34.51
Ligands	–	19.36	30.13	26.75	38.85	39.66	51.07
Solvent	30.27	22.22	38.95	33.35	41.84	39.15	41.89

Statistics for the highest-resolution shell are shown in parentheses.
[a] $R_{merge} = \sum_{hkl}\sum_i |I_i(hkl) - \langle I(hkl)\rangle| / \sum_{hkl}\sum_i I_i(hkl)$ [57].; [b] $R_{meas} = \sum_{hkl}\{N(hkl)/[N(hkl)-1]\}^{1/2} \times \sum_i |I_i(hkl) - \langle I(hkl)\rangle| / \sum_{hkl}\sum_i I_i(hkl)$ [58].; [c] $R_{pim} = \sum_{hkl}\{1/[N(hkl)-1]\}^{1/2} \times \sum_i |I_i(hkl) - \langle I(hkl)\rangle| / \sum_{hkl}\sum_i I_i(hkl)$ [59].; [d] $CC1/2 = \sum(x-\langle x\rangle)(y-\langle y\rangle) / [\sum(x-\langle x\rangle)^2 (y-\langle y\rangle)^2]^{1/2}$ [60].; [e] $R_{work} = \sum|F_o - F_c| / \sum|F_o|$ for all data with $F_o > 2\sigma(F_o)$, excluding data to calculate R_{free}.; [f] $R_{free} = \sum|F_o - F_c| / \sum|F_o|$ for all data with $F_o > 2\sigma(F_o)$, calculated from 5% of reflections, randomly chosen [61].

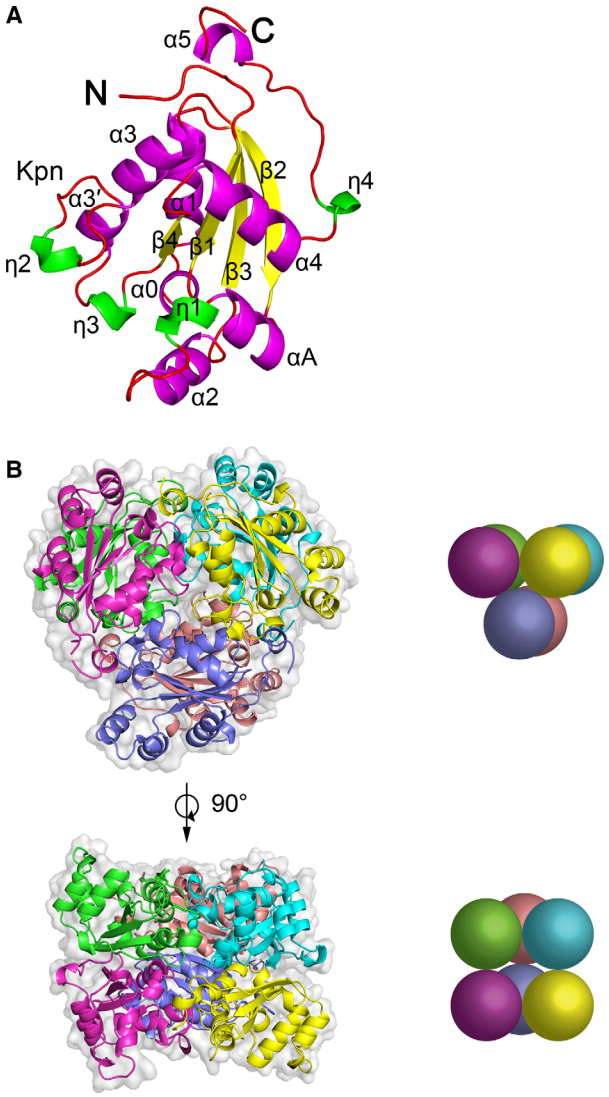


Fig. 2. Crystal structures of monomeric and hexameric NDK. (A) Ribbon diagram of monomeric NDK with secondary structure labelled using conventional nomenclature and colour-coded as follows: β-strand (yellow), α-helix (magenta), 3_{10} helix (green, annotated as η) and loops (red). The *Kpn* loop that forms the trimer interface is labelled (Kpn). Statistics of data processing and model refinement are shown in Table 1. (B) Ribbon diagram overlayed over a surface representation of *A. fumigatus* NDK shows the hexameric quaternary structure. The simplified diagram shows the arrangement as a stack of two trimers. The spheres are coloured in accordance with the individual monomers in the adjacent ribbon and surface representation.

intermediate is formed, the same site is occupied by the donor NDP.

Based on crystallographic symmetry, *A. fumigatus* NDK forms a hexamer which is arranged as a trimer of dimers (Fig. 2B). The *Killer of Prune* (*Kpn*) loop (historically named after mutations within this loop region of NDK from *Drosophila melanogaster*) is comprised of residues 97–103 [23]. In the *A. fumigatus* NDK structure, the *Kpn* loop in each monomer interacts to form a trimeric unit. This loop features a highly conserved proline (Pro100) that is known to play a role in stabilising the hexameric quaternary structure [24]. The trimer interface (average buried surface area of 818 Å^2) is maintained by an extensive hydrogen bond network and a salt bridge between Arg80 and Asp110 whereas the dimer interface (average buried surface area of 952 Å^2) predominantly features hydrogen bond interactions between the α1 helices and β2 sheets of each monomer.

Several lines of evidence have confirmed that the oligomeric state of *A. fumigatus* NDK is hexameric, consistent with eukaryotic homologues of this enzyme. To discern between a crystal contact and a true oligomeric interface, an interface surface area of 856 Å^2 has been used previously as a threshold to differentiate between monomeric and homodimeric proteins [25]. The interface surface area (defined as half the difference of the total surface area and interfacing surfaces) calculated between each subunit of *A. fumigatus* unbound-NDK ranges between 791.5 and 959.8 Å^2. Analysis of these macromolecular interfaces of NDK suggest that *A. fumigatus* NDK adopts a hexameric state.

To determine the oligomeric state of NDK in a solution state (and in turn confirm the crystal structures), analytical SEC and native ion-mobility mass spectrometry (IM-MS) were performed. The analytical SEC profiles of NDK in PBS, phosphate buffer (PB) and ammonium acetate (NH_4OAc, buffer used for MS experiments) were virtually identical with a single elution peak (~ 13.1 mL) corresponding to a molecular mass ~ 110 kDa, indicative of a hexamer (Fig. 3A). Native IM-MS is a high-resolution analytical technique capable of identifying the quaternary state and conformation of multimeric protein assemblies with high mass accuracy [26,27]. Here, native MS shows that both *A. fumigatus* unbound-NDK and ADP-NDK predominately forms a species with a molecular mass of 108.4 kDa (theoretical hexameric mass 108.1 kDa), corresponding to hexameric NDK (Fig. 3B–D, purple). In addition, IM-MS also shows that the ADP-NDK hexamer undergoes a conformational change (hexamer^{21+} 12.7 ms, hexamer^{22+} 12.6 ms) which is indicative of compaction compared to unbound-NDK (hexamer^{21+}: 15.1 ms, hexamer^{22+}: 13.8 ms) (Fig. 3B, inset). Furthermore, we also investigated the dissociation dynamics of NDK hexamers using both gas-phase induced (collision-induced, CID) and solution-induced dissociation approaches (Fig. 3C,

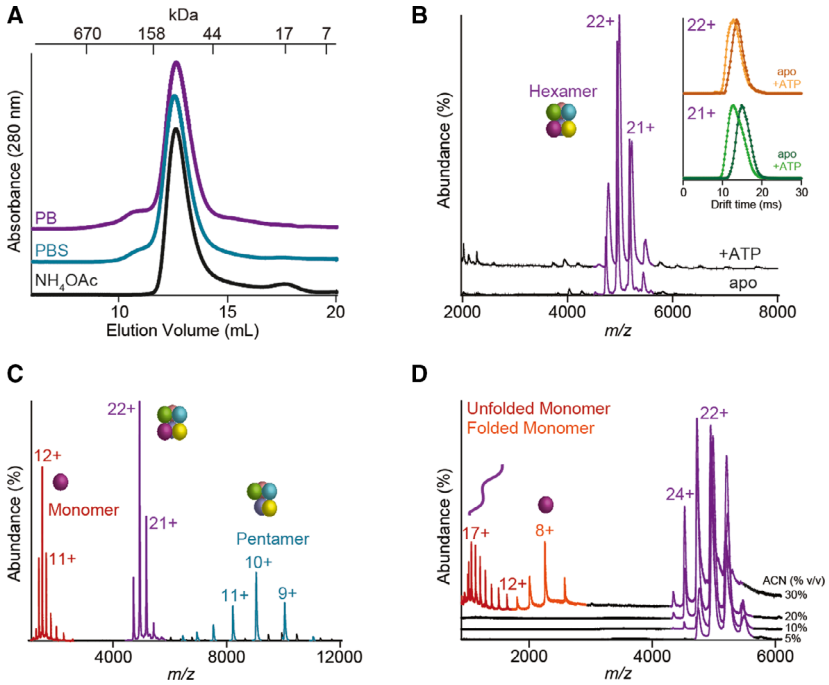


Fig. 3. Analytical-SEC and native IM-MS highlights changes in structure and dynamics of NDK hexamers. (A) Analytical-SEC of NDK (50 μM) in 50 mM phosphate buffer (PB, purple) (pH 7.4), PBS (green) (pH 7.0) and 200 mM ammonium acetate (NH_4OAc, black) (pH 6.8) indicates buffers do not affect the oligomeric state of NDK. Elution volumes of molecular mass standards are indicated above. (B) Native MS shows that both unbound-NDK and ADP-NDK (12.5 μM) are hexameric (purple). ATDs of unbound-NDK and ADP-NDK hexamers$^{21+/22+}$ show a conformational change in the hexamer upon ATP and Mg^+ addition. (C) CID-MS and solution-induced dissociation MS (D) of unbound-NDK hexamers (12.5 μM) shows that monomers dissociate from the hexamer. Dominant charge states that correspond to each oligomeric assembly are noted above.

D). Both CID-MS and solution-induced dissociation (via titration of acetonitrile, ACN) of NDK hexamers shows that monomers are the most labile unit in the hexameric assembly (Fig. 3C,D, red and orange). With this, the gas-phase properties mirror the solution-phase properties of NDK in terms of structure and dynamics. Overall, these data correspond with the crystallographic data demonstrating that unbound- and nucleoside-bound *A. fumigatus* NDK adopts a hexameric state and undergoes a conformational change upon nucleoside binding.

A. fumigatus NDK displays differences in binding pocket composition as compared to homologous enzymes

As NDK is involved in several essential cellular processes, it is unsurprising that the structure and function of these enzymes are conserved throughout all domains of life. Amino acid sequence alignment of NDK from *A. fumigatus* and several phylogenetically diverse species reveal high overall similarity. In comparison to *A. fumigatus*, eukaryotic NDKs share the highest identity (60% *Dictyostelium discoideum*; 64% *Homo sapiens* isoform A; 65% *Leishmania major*; 68% *Saccharomyces cerevisiae*; 82% *A. nidulans*) compared to prokaryotic NDKs (46% *Escherichia coli*; 49% *Halobacterium salinarum*; 51% *Staphylococcus aureus*) and viral NDKs (43% *Acanthamoeba polyphaga mimivirus* (APMV)). There is also high conservation of motifs that are common to both prokaryotes and eukaryotes. These include the Pro-Phe-Phe (PFF) and His-Gly-Ser-Asp (HGSD) motifs, both involved in substrate and cofactor binding, as well as the Ser-Gly-Pro (SGP) motif (Fig. 4A).

Despite differences in the polypeptide sequence, superimposition of the α-carbons (Cα) of eukaryotic, prokaryotic and viral NDK homologues reveal a similar global fold, reflected in their small root mean square deviation (RMSD) values (Fig. 5; Table 2). NDK from APMV shows the greatest variance in sequence identity to *A. fumigatus* and the highest structural perturbation based on RMSD whereas *L. major* shows the greatest similarity in global fold (Table 2). *S. cerevisiae* and *A. fumigatus* are closely related species, both members of the fungal kingdom, and their NDK sequences share a high degree of identity, yet the two structures show a moderate deviation in terms of their RMSD (0.771 Å) in comparison to more distantly related species (Table 2) (Fig. 4B).

The active site residues involved in the binding and stabilisation of ADP observed in *A. fumigatus* NDK show little deviation from homologous enzymes from *S. aureus* (PDB ID: 3Q8U), *H. sapiens* (PDB ID: 2HVD) and *L. major* (PDB ID: 3NGU) (Fig. 6A) [14,16,28]. Key to all structures is the π-π stacking interaction between the adenine ring and the equivalent Phe side chain. Furthermore, the interaction

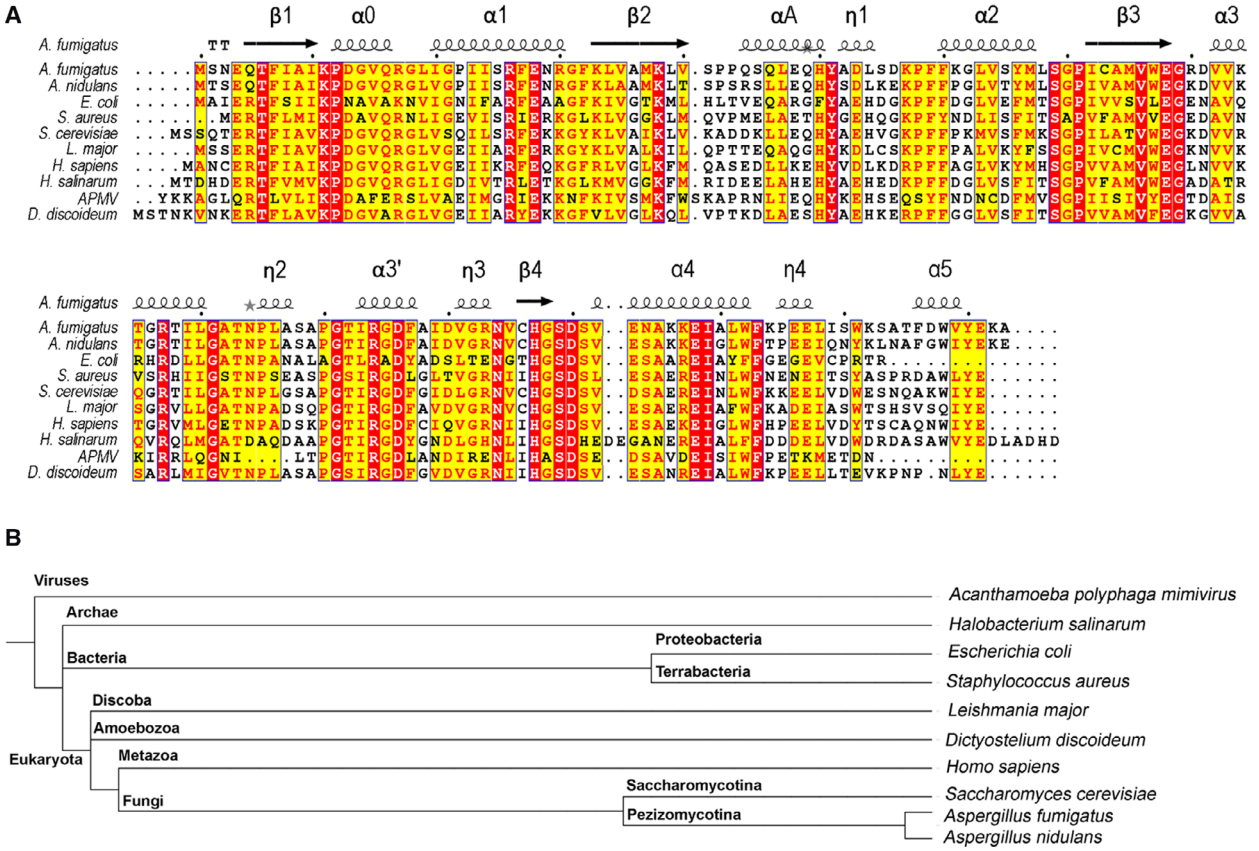


Fig. 4. Phylogenetic analysis and sequence alignment of NDK across prokaryotes and eukaryotes. (A) Multiple sequence alignment of NDK enzymes from *Aspergillus fumigatus* (UniProt Protein Database: Q7Z8P9), *Aspergillus nidulans* (UniProt Protein Database: Q8TFN0), *Escherichia coli* (UniProt Protein Database: P0A763), *Staphylococcus aureus* (UniProt Protein Database: P99068), *Saccharomyces cerevisiae* (UniProt Protein Database: P36010), *Leishmania major* (UniProt Protein Database: Q4Q7A6), *Homo sapiens* isoform A (UniProt Protein Database: P15531), *Halobacterium salinarum* (UniProt Protein Database: P61136), *Acanthamoeba polyphaga mimivirus* (UniProt Protein Database: Q5UQL3) and *Dictyostelium discoideum* (UniProt Protein Database: P22887) generated using T-COFFEE (http://tcoffee.crg.cat/) and visualised using ESPRIPT (http://espript.ibcp.fr) [52,53]. Strictly conserved residues are shown in white text on red background, well-conserved residues are shown in red text on yellow background, and non-conserved residues are shown in black text. Grey stars indicate residues that have alternate conformations. The secondary structure of *A. fumigatus* NDK is displayed above the sequence using conventional nomenclature where α denotes an α-helix, β denotes a β-strand, η denotes a 3_{10} helix, and TT denotes a strict β-turn. (B) Phylogenetic tree depicting the evolutionary relationships between representative prokaryotic and eukaryotic species selected for analysis of homologous NDK enzymes generated by INTERACTIVE TREE OF LIFE (https://itol.embl.de/) [54].

between the diphosphates and Thr and Arg side chains in equivalent positions is maintained for all but *S. aureus* NDK. In the *S. aureus* structure, the Thr interaction remains intact but Arg in a different position is involved. Common to all structures is the interaction between the ribose ring of ADP via the 2′ or 3′ hydroxyl groups, or both, and Asn. The main chain of Val also interacts with the 2′ hydroxyl group located on the ribose ring, a commonality to these structures, whereas Lys is involved in all species but *S. aureus* NDK.

Although the overall structure of NDK is similar between species, the residues comprising the binding pocket show species-specific differences. The binding pocket of *A. fumigatus* and *L. major* NDK exhibits the greatest sequence identity (28 out of 30 residues), followed by *H. sapiens* NDK (25 out of 30 residues) and *S. aureus* NDK (20 out of 30 residues) (Fig. 6 B). Predominantly, there are major differences to residues within the α2 helix (Gly62), α3/η2 loop (Ala92), α3′/η3 loop (Asp110) and β4 strand (Cys116). Residues that are not strictly conserved but maintain similar physiochemical properties reside in the η1 3_{10} helix (Asp53, Leu54), η1/α2 loop (Lys57), α2 helix (Tyr66, Met67), α3 helix (Leu90) and η3/β4 loop (Val115).

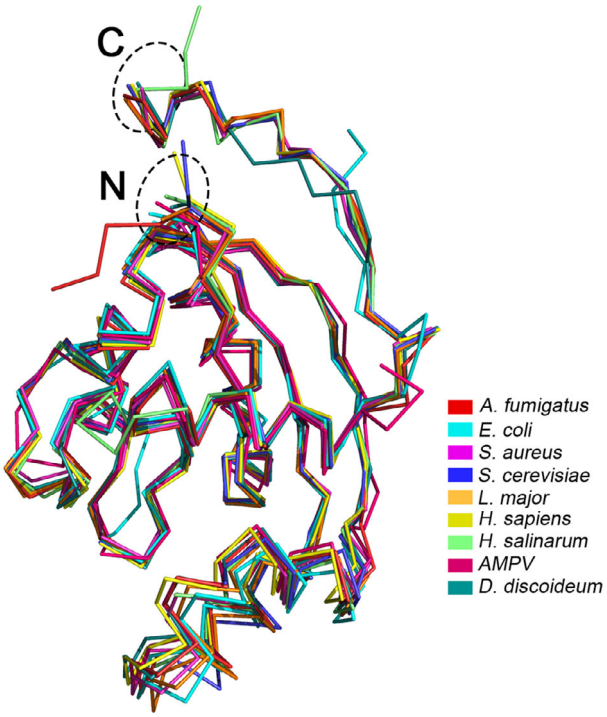


Fig. 5. Superimposed ribbon structures of NDK from various species. Overlapping structures show high conservation of structure despite deviations in amino acid sequence identity. Structures of NDK homologues are represented as follows: *Aspergillus fumigatus* (6XP4, red), *Escherichia coli* (2HUR, cyan), *Staphylococcus aureus* (3Q83, magenta), *Saccharomyces cerevisiae* (3B54, blue), *Leishmania major* (3NGS, orange), *Homo sapiens* isoform A (2HVD, yellow), *Halobacterium salinarum* (2AZ1, green), *Acanthamoeba polyphaga mimivirus* (2B8Q, burgundy), and *Dictyostelium discoideum* (1NDC, teal).

Table 2. Amino acid sequence alignment and root mean square deviation of alpha carbons from monomeric NDK across various species. Root mean squared deviation (RMSD) of alpha carbons (C_α) is reported in Å.

Comparison between *A. fumigatus* vs.	Sequence identity (%)	RMSD (C_α)
Acanthamoeba polyphaga mimivirus (2B8Q)	43	1.305 (156)
Escherichia coli (2HUR)	46	1.067 (138)
Halobacterium salinarum (2AZ1)	49	1.009 (148)
Staphylococcus aureus (3Q83)	51	0.780 (147)
Dictyostelium discoideum (1NDC)	60	1.158 (142)
Homo sapiens isoform A (2HVD)	64	0.762 (148)
Leishmania major (3NGS)	65	0.581 (149)
Saccharomyces cerevisiae (3B54)	68	0.771 (143)

CDP- and TDP-bound structures reveal a binding mode unique to *A. fumigatus* NDK

Although *A. fumigatus* NDK was co-crystallised with NTPs, all solved structures were bound to their respective NDP due to rapid substrate turnover (Fig. 7. The electrostatic potential surface map of NDP-bound NDK shows a surface-exposed, positively charged binding pocket that complements the negatively charged phosphate groups of NDPs (Fig. 8A). As observed in the GDP-, IDP- and UDP-bound structures, the catalytic Mg^{2+} binds in a neutral to positively charged region of the binding pocket. The Mg^{2+} forms a coordination sphere with five surrounding water molecules and an oxygen group from the α-phosphate of the NDP (coordination bond distances of 1.95–2.72 Å). Although magnesium chloride was added to the crystallisation medium, Mg^{2+} was not modelled in the ADP-, CDP- and TDP-bound structures of NDK due to ambiguous density.

Superimposition of the purines (ADP, GDP and IDP) shows a highly conserved binding mode (Fig. 8A). Although UDP shares the same binding pose as the purines, there are major differences in CDP and TDP that have not been observed previously in structures of NDK from any other species (Fig. 8B–D). Once bound to NDK, there is a contortion of the ribose ring of CDP that causes a 3.9 Å shift of the α-phosphate deeper into the pocket whereas the β-phosphate remains in a similar position to other NDPs (Figs 8C and 9A).

The binding mode of TDP is unique compared to the NDPs explored in this study due to its natural deoxyribose ring. The absence of a 2′ hydroxyl group on its deoxyribose ring limits the possible interactions between the deoxyribose ring and NDK active site residues and this has significant implications on its binding mode. As a result of this increased flexibility, the deoxyribose ring puckers when bound to NDK and the diphosphates are rotated outside of the binding pocket by 57° (Figs 8D and 9B).

Key π-π stacking interactions and hydrogen bonds stabilise NDP binding to the active site

The binding of an NDP to NDK is associated with a conformational change that stabilises substrate

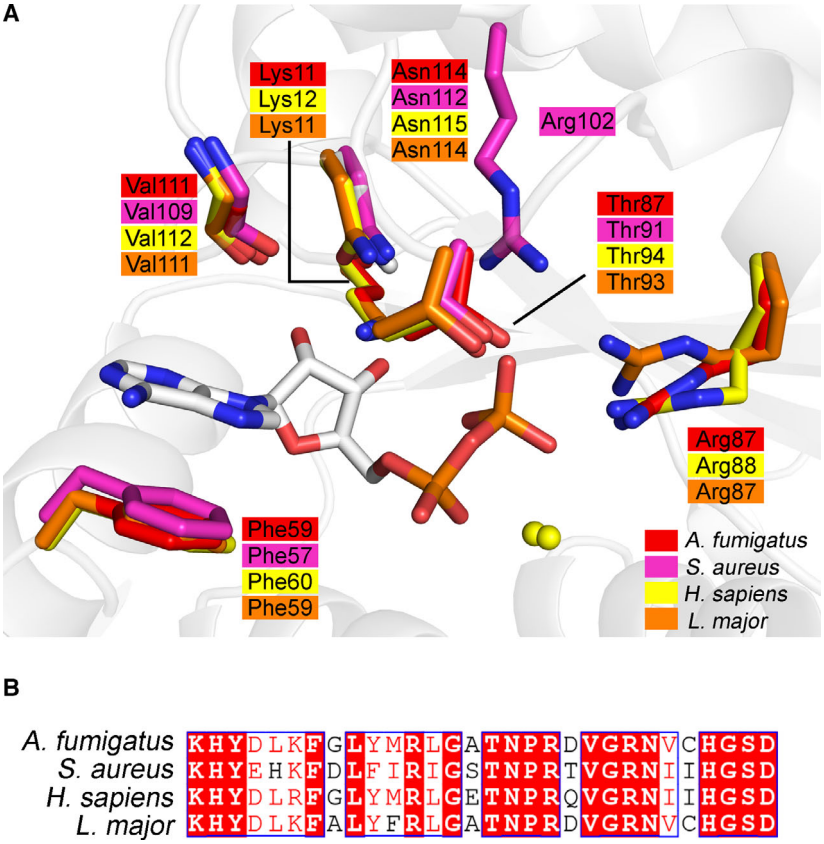


Fig. 6. Interactions between ADP and active site residues of homologous NDKs are highly conserved but residues that comprise the binding pocket are species-specific. (A) Superimposed active site of ADP-bound NDK and (B) sequence alignment of binding pocket residues from *Aspergillus fumigatus* (UniProt Protein Database: Q7Z8P9, PDB ID: 6XP7, red), *Staphylococcus aureus* (UniProt Protein Database: P99068, PDB ID: 3Q8U, magenta), *Homo sapiens* isoform A (UniProt Protein Database: P15531, PDB ID: 2HVD, yellow) and *Leishmania major* (UniProt Protein Database: Q4Q7A6, PDB ID: 3NGU, orange). ADP is represented as sticks and Mg^{2+} represented as yellow spheres. Sequence alignment was generated using CLUSTAL OMEGA (https://www.ebi.ac.uk/Tools/msa/clustalo/) and visualised using ESPRIPT (http://espript.ibcp.fr) [53,55]. Strictly conserved residues are shown in white text on red background, residues that are not conserved but have similar physico-chemical properties are shown in red text on white background, and non-conserved residues are shown in black text.

binding. The Phe59-containing loop moves to orientate the side chain into an optimum position to form a π-π stacking interaction with the nitrogenous base. This is common to all NDPs regardless of their base moiety (Fig. 10). This conformational change observed in the crystal structure is consistent with the observations from IM-MS experiments (Fig. 3B, inset).

Unique to the guanine ring of GDP is a 2′ amine group that forms a 3.0 Å hydrogen bond with Glu151 located on the C-terminal loop of an adjacent NDK monomer, acting to further stabilise binding of GDP in the active site (Fig. 10A). GDP is the only substrate observed in this study that interacts directly with other NDK subunits in the hexameric assembly. Focusing on the sugar moiety of NDPs, there are distinct differences in their interactions with active site residues that influence selectivity. Consistent amongst the purines and UDP, the Asn114 side chain forms a 2.6–3.1 Å hydrogen bond to the 3′ hydroxyl group and Lys11 forms a 2.4–3.0 Å hydrogen bond to both hydroxyl groups of the ribose ring (Fig. 10A–D). In the case of CDP and TDP, Lys11 interacts exclusively to the 3′ hydroxyl group (Fig. 10E,F). In all NDPs, except TDP, the oxygen group from the Val111 main chain forms a 3.0–3.5 Å hydrogen bond to the ribose ring. Interactions between the diphosphate group of NDPs and active site residues appear to be highly conserved. The side chains of Arg87 and Thr93 form hydrogen bonds with the β-phosphate of all NDPs (Fig. 10B).

As CDP and TDP adopt significantly different binding poses to the previously discussed NDPs, interactions with the NDK active site are unique. Despite the contortion of bound-CDP, the π-π stacking interaction between the cytidine ring and Phe59 remains intact. The puckered ribose ring maintains a 2.8 Å hydrogen bond with the Lys11 side chain but is no longer in close proximity to the Asn114 side chain. However, Asn114 can form a 3.1 Å hydrogen bond to the phosphoester bond between the ribose and phosphate moieties. Despite the alternate positioning of the diphosphate group of CDP, the β-phosphate maintains hydrogen bonds to Arg87 (2.7 Å) and Thr93 (2.2 Å) and the α-phosphate forms new interactions with Tyr51 (3.0 Å) and the catalytic histidine, His117 (2.6 Å) (Fig. 10E).

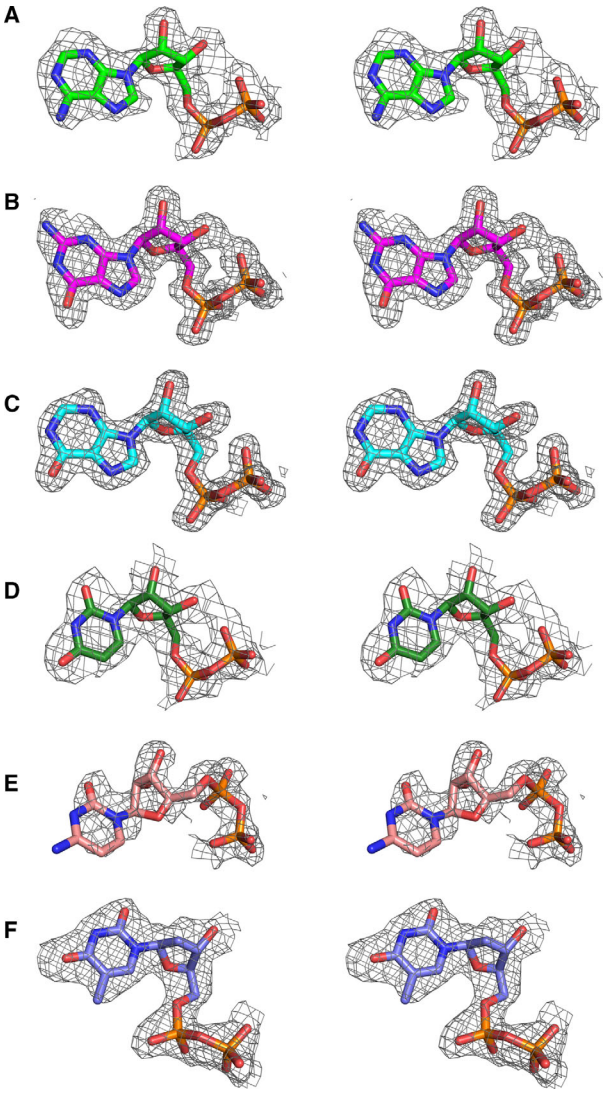


Fig. 7. Feature-enhanced electron density map (modified 2 m F_{obs}–DF_{model} σ_A-weighted map) of NDP ligands bound to *A. fumigatus* nucleoside diphosphate kinase. Maps are contoured at 1 σ for (A) GDP, (B) ADP, (C) IDP, (D) UDP, (E) CDP and (F) TDP in stereo walleye view [56].

TDP appears to make fewer interactions whilst bound to NDK. The lack of a 2′ hydroxyl group on the deoxyribose ring abolishes equivalent interactions observed in other NDPs but maintains a 3.0 Å hydrogen bond between the 3′ hydroxyl and the Lys11 side chain. Although the diphosphate group is rotated outside of the binding pocket, the β-phosphate maintains hydrogen bonds with Arg87 (3.0 Å) and Thr93 (3.3 Å) side chains, an interaction common to all other NDPs albeit with markedly longer bond distances (Fig. 10F).

Kinetic characterisation of *A. fumigatus* NDK and changes in binding free energy shows preference for adenine nucleosides and disfavours cytidine nucleosides

Although NDK can use all NTPs as substrates, selectivity is governed by structural differences in binding modes. Apparent Michaelis–Menten binding constants were measured for *A. fumigatus* NDK for all NTPs and showed clear preference for purines over pyrimidines (Table 3). Specifically, GTP has the lowest apparent K_M (0.402 ± 0.03 mM), followed closely by ATP (0.513 ± 0.08 mM) and ITP (0.617 ± 0.10 mM) whereas TTP has the highest apparent K_M (2.23 ± 0.30 mM). CTP and UTP had similar apparent K_M constants (1.32 ± 0.10 mM and 1.35 ± 0.20 mM, respectively). However, there are stark differences in the turnover rates (k_{cat}) and specificity constants (k_{cat}/K_M) calculated when different NTP substrates are used. Although GTP has the lowest apparent K_M, ATP is the preferred substrate, indicated by the high specificity constant measured (626 ± 110.0 $mM^{-1}{\cdot}s^{-1}$) but GTP closely follows (468 ± 41.0 $mM^{-1}{\cdot}s^{-1}$). Similarly, TTP has the highest apparent K_M yet CTP is the least favoured substrate with the lowest calculated specificity constant (40.2 ± 4.4 $mM^{-1}{\cdot}s^{-1}$). All kinetic parameters are summarised in Table 3 and representative binding curves are shown in Fig. 11.

The kinetic parameters are relatively consistent with the Gibbs free energy change of binding ($\Delta G_{binding}$) calculated for each NDP-bound NDK structure which considers the binding mode of the ligand and the interactions made with active site residues. Binding to ADP elicited the highest magnitude of change of negative $\Delta G_{binding}$, followed by GDP, UDP, CDP and IDP, and TDP induced the lowest $\Delta G_{binding}$ (Table 4). These data indicate that on a structural level, NDKs also show preference of binding to ADP but mostly disfavour binding to TDP. Considering both the structural and kinetic data, ATP showed the lowest $\Delta G_{binding}$ and the highest specificity constant, indicating that it is the preferred substrate of NDK. UTP has a two-fold higher apparent K_M than ITP yet both have comparable specificity constants and $\Delta G_{binding}$ values. This implies that there is little difference in preference between these two substrates. In terms of TTP, it has the highest apparent K_M, but a moderate turnover rate and specificity constant. The rotation of the TDP diphosphate group outside of the active site pocket has resulted in a binding free energy change penalty. Unexpectedly, the change in binding energy of CTP is most similar to UDP and GDP, despite having a higher apparent K_M and the lowest specificity constant

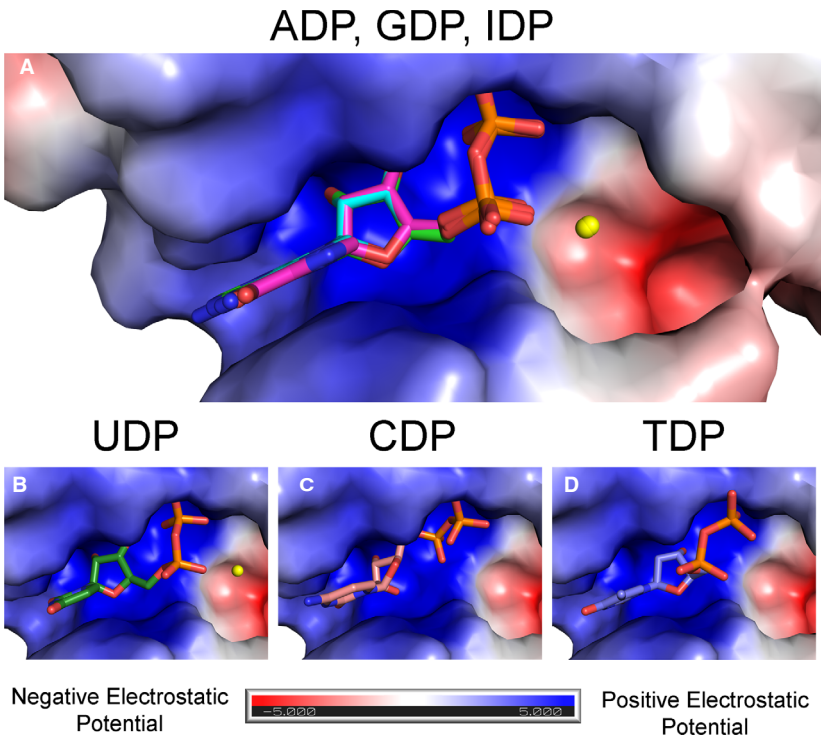


Fig. 8. Electrostatic surface maps of *A. fumigatus* NDK bound to purines and pyrimidines. (A) purines (ADP, GDP and IDP) and pyrimidines, (B) UDP, (C) CDP and (D) TDP. NDPs are colour coded as follows: ADP (green), GDP (magenta), IDP (cyan), UDP (dark green), CDP (salmon) and TDP (purple). Mg^{2+} is represented as a yellow sphere. Statistics of data processing and refinement are shown in Table 1.

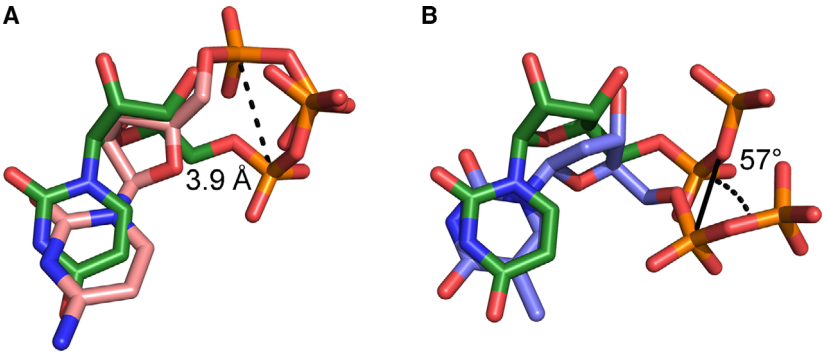


Fig. 9. TDP and CDP binds in a unique conformation to *A. fumigatus* NDK compared to the conventional pose adopted by UDP. (A) The diphosphate group of CDP (salmon) is shifted 5.9 Å deeper into the pocket relative to UDP (dark green) whereas (B) TDP (purple) rotates 57° outside of the binding pocket.

relative to all other tested substrates. Though it may bind tightly to the active site pocket, the turnover rate is significantly lower and is therefore a poor substrate in comparison to other nucleosides.

Discussion

In this study, we have analysed the structure, oligomeric state and kinetic characteristics of *A. fumigatus* NDK to understand the nuances of substrate selectivity. There are subtle differences in the substrate binding pocket of NDK from *A. fumigatus* that are species-specific, particularly between fungal and human homologues, which are important to identify in the context of drug design. Structural analysis of the binding pocket of ADP-NDK from *A. fumigatus* (PDB ID: 6XP7) and comparison to ADP-NDK *H. sapiens* (PDB ID: 2HVD) reveal distinct differences in their topography. The volume of the binding pockets are identical (~ 220 $Å^3$) yet the surface area of *A. fumigatus* (327 $Å^2$) is 15% smaller compared to that of *H. sapiens* (385 $Å^3$), indicating that the fungal binding pocket is comprised of fewer cavities and clefts [29]. Sequence alignment of the binding pocket residues identifies three distinct regions of high variability in the α3/η2 loop, α3′/η3 loop and β4 strand that may be exploited in the design of a competitive inhibitor (Fig. 6B). As an alternative, designing non-competitive inhibitors that target an allosteric site may be an elegant approach to simultaneously modulate catalytic

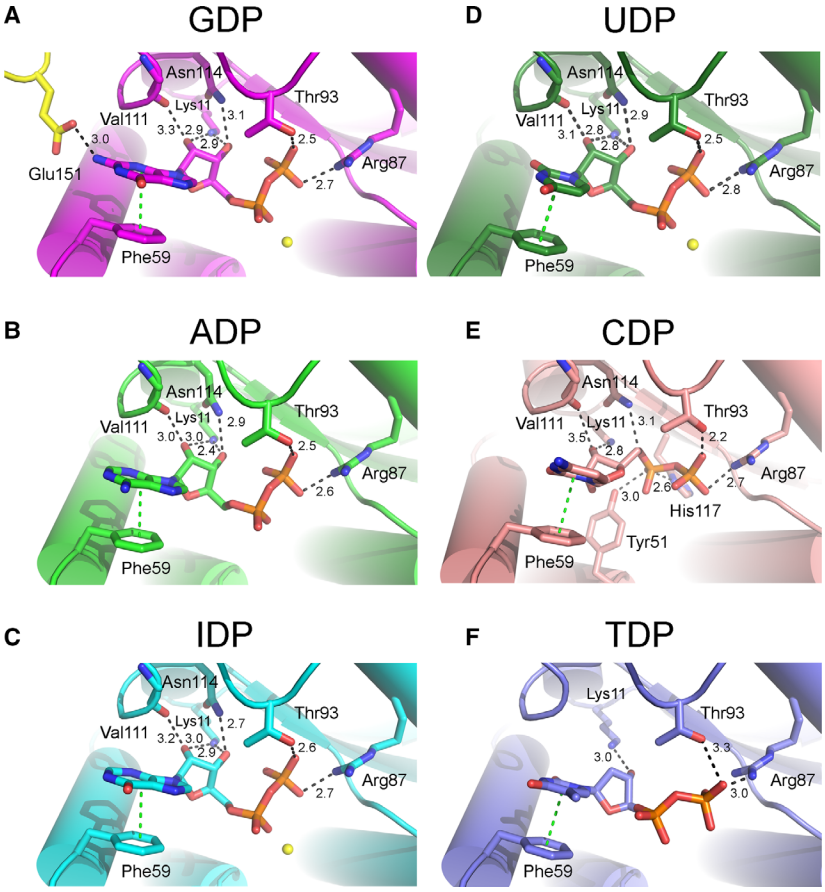


Fig. 10. Interactions between active site residues of *A. fumigatus* NDK with various purines and pyrimidines. (A) GDP, (B) ADP, (C) IDP, (D) UDP, (E) CDP and (F) TDP which naturally has a 2′ deoxy ribose ring. Key hydrogen bond interactions are shown as black dashed lines, π-π stacking interactions as green dashed lines, and all distances are reported in Å. Mg^{2+} is represented as a yellow sphere. All nucleoside diphosphates are represented as sticks.

activity whilst attaining selectivity between the fungal and human enzyme. At present, at least one non-competitive inhibitor of human NDK has been identified [30]. Although the binding surface has not yet been defined, it is expected that an allosteric site would vary considerably between different species compared to the active site and this should be considered in future drug design efforts.

Table 3. Apparent Michaelis–Menten kinetic constants of various NTPs used by *A. fumigatus* NDK. Kinetic parameters for all NTPs, excluding ATP, were determined using the hexokinase/glucose-6-phosphate dehydrogenase coupled enzyme assay, whereas for ATP, the pyruvate kinase/lactate dehydrogenase coupled enzyme assay was used. Apparent kinetic constants were measured from assays performed in duplicate from three replicates ($n = 3$).

Nucleotide	K_M (mM)	k_{cat} (s^{-1})	k_{cat}/K_M ($mM^{-1}·s^{-1}$)
GTP	0.402 ± 0.03	188 ± 5.0	468 ± 41.0
ATP	0.513 ± 0.08	321 ± 33.0	626 ± 110.0
ITP	0.617 ± 0.10	83.2 ± 13.0	135 ± 30.0
CTP	1.32 ± 0.10	53.1 ± 3.7	40.2 ± 4.4
UTP	1.35 ± 0.20	140 ± 22.0	104 ± 22.0
TTP	2.23 ± 0.30	178 ± 9.6	79.8 ± 12.0

It has been well documented that NDK from prokaryotes (i.e. *E. coli*) tend to form tetrameric complexes whereas NDK from eukaryotes (i.e. *S. cerevisiae*, *D. melanogaster*, *H. sapiens*) tend to form hexameric complexes [16,31,32]. Solution-based experiments such as small-angle X-ray scattering have shown good agreement between the molecular envelope and the hexameric crystal structure of *Leishmania braziliensis* NDK [33]. This has also been confirmed in NDK from *D. discoideum* using equilibrium ultracentrifugation and analytical SEC [34]. However, there is still an ongoing debate of the quaternary structure adopted by *A. fumigatus* NDK and how this relates to its biological function. In a previous paper, the crystal structure of *A. fumigatus* NDK was solved as a hexamer, arranged as a trimer of dimers [22]. However, analytical SEC indicated that the enzyme was tetrameric in solution [22]. Data obtained from the complementary techniques of crystallography, analytical SEC and native mass spectrometry unequivocally and consistently confirm that *A. fumigatus* NDK adopts a

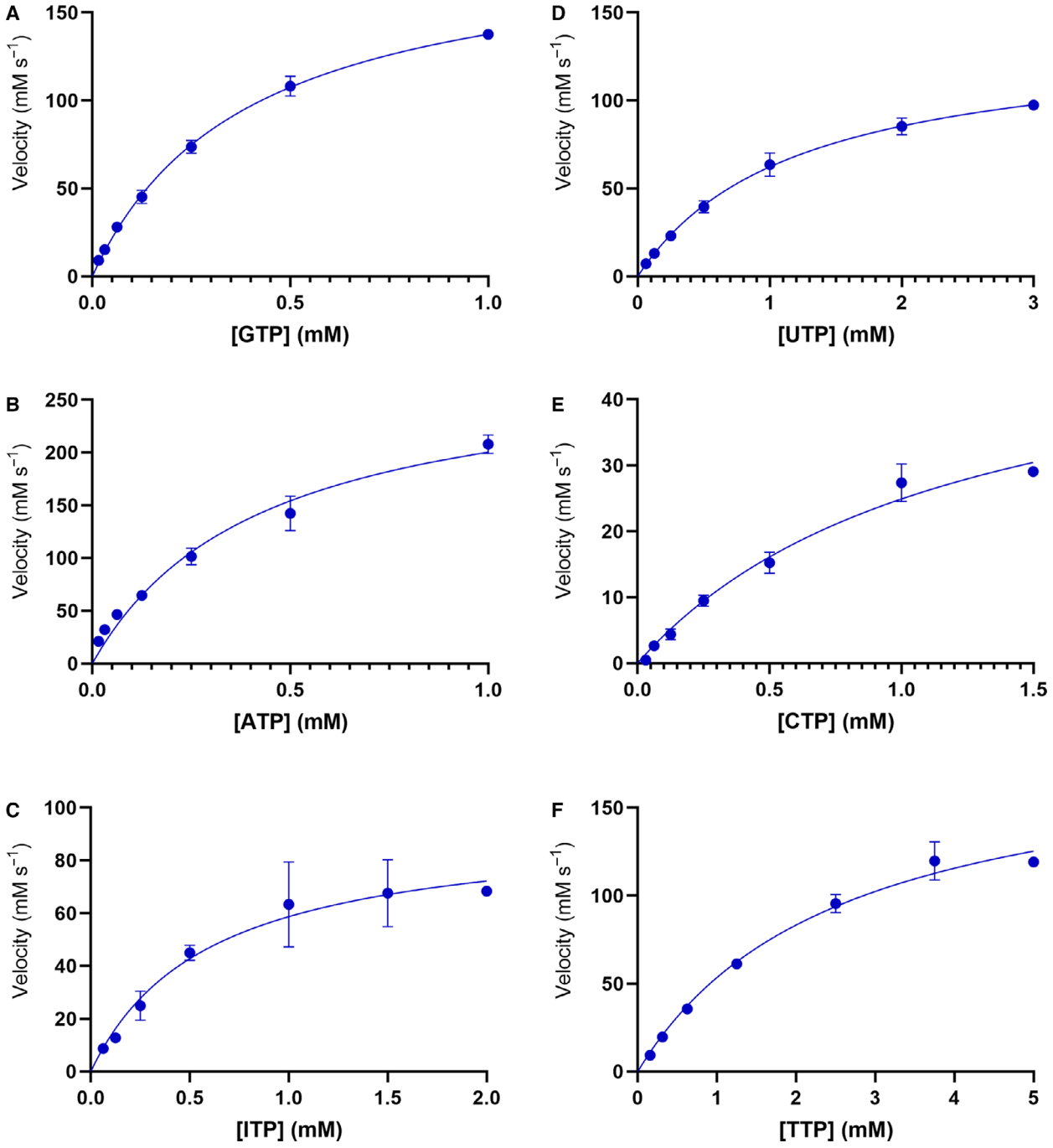


Fig. 11. Representative enzyme activity curves for NTPs. Binding curves for (A) GTP, (B) ATP, (C) ITP, (D) UTP, (E) CTP and (F) TTP represent single experiments. Each data point represents the mean of duplicate samples, and the error bars represent the standard deviation.

hexameric assembly, contradicting previous suggestions of a tetrameric species. Calculations of the buried surface area between dimer and trimer interfaces (952 and 818 Å^2, respectively) of *A. fumigatus* NDK are indicative of true oligomeric interfaces. The hexameric arrangement is also consistent with other eukaryotic homologues of NDK. As observed in NDK from *Mycobacterium tuberculosis*, the quaternary structure is intimately intertwined with enzymatic activity and function [35]. This is likely to be

Table 4. Gibbs free energy change upon binding of NDP to NDK. Free energies were calculated using ICM-Pro (MolSoft L. L. C.). The binding free energy change was calculated between ligand (NDPs) and protein (NDK) for each available monomer and averaged.

Nucleotide	$\Delta G_{binding}$ (kcal·mol^{-1})
GDP	−30.0 ± 4.3 (n = 3)
ADP	−37.0 ± 3.5 (n = 3)
IDP	−27.3 ± 0.9 (n = 3)
CDP	−26.8 (n = 1)
UDP	−29.3 ± 3.2 (n = 6)
TDP	−17.5 ± 3.7 (n = 6)

the case in *A. fumigatus* NDK since GDP-bound structures show that hydrogen bond interactions form within the monomer as well as with neighbouring subunits (Fig. 10A). As the oligomeric state has been shown to influence function, several lines of evidence have been presented in this study to clarify previous suggestions of a tetrameric species and instead confirm that NDK from *A. fumigatus* adopts a hexameric assembly.

Crystal structures of *A. fumgiatus* NDK bound to a full complement of naturally occurring NDPs have revealed key nuances in their binding poses that govern nucleoside selectivity. The conformation adopted by the purines and UDP were highly similar, featuring a conserved π-π stacking interaction between Phe59 and the nucleobase as well as an extensive hydrogen bond network between the ribose ring and diphosphate group with active side residues. However, CDP and TDP adopted unique binding modes that have not yet been observed in NDK from any other species and serve to explain why NDK disfavours thymidine and cytosine nucleosides as substrates. The CDP-bound structure shows puckering of the ribose ring upon binding to NDK which consequently changes the orientation of the diphosphate group within the active site (Fig. 9A). In this position, sub-optimal bond angles are adopted by CDP that destabilise the complex, exemplified in the low occupancy observed in the crystal structure. Intriguingly, the CDP-bound structure of NDK formed a hexamer in the asymmetric unit, yet only one monomer was bound to CDP whereas the remaining five were bound to phosphate. In comparison, all other NDP-bound structures of *A. fumigatus* NDK showed all monomers occupied by their respective NDP.

The binding pose adopted by TDP when bound to *A. fumigatus* NDK is unique in comparison to the conserved TDP binding modes observed in APMV (PDB ID: 2B8Q) and *D. discoideum* (PDB ID: 1NDC) NDK [21,36]. This is the first instance in which the diphosphate group rotates outside of the binding pocket whilst the thymidine ring and deoxyribose sugar maintains key interactions with active site residues, Phe59 and Lys11 (Fig. 9B). Without a 2′ hydroxyl group on the sugar moiety, TDP is poorly anchored into the binding site. This introduces flexibility to the diphosphate groups and allows TDP to adopt this highly unusual binding pose. In turn, this flexibility likely contributes to the high apparent K_M and low specificity constant measured. The rotation lengthens hydrogen bond distances observed in other NDP-bound structures and disfavours TDP binding. When comparing the binding pocket cavity of *A. mimivirus* NDK and *D. discoideum* NDK with *A. fumigatus* NDK, there are a number of anomalies that may account for the alternative TDP binding mode observed. The structure of *A. mimivirus* TDP-NDK (PDB ID: 2B8Q) shows a shallow binding pocket that has a volume of 115 Å^3 and surface area of 150 Å^2 relative to that of *D. discoideum* TDP-NDK (PDB ID: 1NDC, volume of 170 Å^3, surface area of 300 Å^2) and *A. fumigatus* TDP-NDK (PDB ID: 6XPV, volume of 200 Å^3, surface area 290 Å^2) [29]. In the *A. mimivirus* NDK structure, the loop between α3 and η2 (labelled in accordance with *A. fumigatus* NDK secondary structure in Fig. 4A) is positioned in an 'open' conformation that widens the binding pocket and easily accommodates the diphosphate group of TDP. The α3/η2 loop of *A. fumigatus* and *D. discoideum* NDK is positioned in a 'closed' conformation, narrowing the opening to the binding site and forming a deeper pocket, yet two different binding modes of TDP are observed. Intriguingly, residues within the α3/η2 loop of NDK have shown to deviate between species (Fig. 6B). In *A. fumigatus* NDK and *D. discoideum* NDK, this substitution from Ala to Val leads to subtle changes in the loop position (Fig. 4A). The *A. fumigatus* NDK α3/η2 loop is shifted 1.3 Å deeper into the binding pocket, relative to that of the *D. discoideum* NDK α3/η2 loop. This orientates the Thr93 side chain to a position that sterically clashes with the β-phosphate of TDP when bound in the conventional binding pose as observed in the *D. discoideum* NDK structure (Fig. 12).

The change in Gibbs free energy upon binding of an NDP to NDK calculated from solved crystal structures and the binding affinities and specific constants measured in enzyme assays can be used to determine the basis of NDK substrate selectivity. Compared to all other NTP substrates, GTP has the lowest apparent K_M. While all other NDPs can only form a π-π

stacking interaction between the Phe59 side chain and the respective nitrogenous base, the guanine ring contains a 2′ amine that can form an additional hydrogen bond to Glu151 from an adjacent NDK monomer (Fig. 10A). This unique interaction is the only instance of an NDP interacting with more than one NDK monomer observed in this study and is likely to contribute significantly to an increase in binding affinity. This effect is amplified when considering the hexameric arrangement of NDK. Although there is a modest increase in the apparent K_M of ATP relative to GTP, NDK rapidly catabolises ATP, resulting in more frequent use and greater preference of this substrate. The binding mode of CDP, particularly the shift of the α-phosphate deeper into the binding pocket, orients the coordinating oxygen group away from the Mg^{2+} binding site (Fig. 8C). Greater contortion and distance from Mg^{2+} likely contribute to the slower turnover rate of CTP and is therefore the least preferred substrate.

As observed in *A. fumigatus* NDK, homologues of NDK from different species have shown broad substrate specificity with an ability to utilise all naturally occurring NTPs as substrates [37,38]. Comparing the kinetic parameters of *A. fumigatus* NDK to eukaryotic homologues from the pathogenic fungus, *C. albicans*, and *H. sapiens* reveals similar trends but notable differences in their substrate selectivity. In all three species, GTP had the lowest apparent K_M, followed by ATP. However, the enzyme appeared to be clearly selective for GTP in the *C. albicans* and *H. sapiens* homologues whereas ATP was preferred by *A. fumigatus* NDK. Similarly, the least preferred substrate was CTP in both *C. albicans* and *A. fumigatus* NDK. The differences in the measured apparent K_M constants and specificity constants indicate that there are subtle structural differences that govern substrate selectivity in NDK from different species. The binding mode of CDP and TDP observed in this study is unique to *A. fumigatus* NDK and indicates that the position of the phosphate moiety as well as distance from the Mg^{2+} binding site appears to greatly influence catalytic turnover (Fig. 10E,F). In contrast, interactions made with adjacent NDK monomers, as observed in GDP-bound NDK, may increase binding affinity but is associated with a penalty to catalytic turnover (Fig. 10A) (Table 3). Finally, energetically favourable binding of ADP to NDK influenced by close hydrogen bond contacts to active site side-chains facilitates rapid turnover and is thus the preferred nucleoside. Considering the abundance of guanine in soils but greater availability of adenine in human extracellular fluid, the niche occupied by *A. fumigatus* may stimulate NDK activity and facilitate rapid cell growth and division to establish infection.

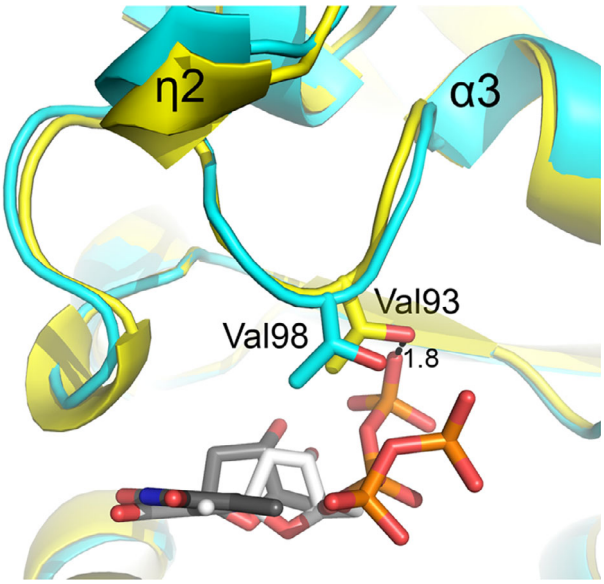


Fig. 12. The conventional binding mode of TDP sterically clashes with Val93 in the α3/η2 loop of *A. fumigatus* NDK. The crystal structure of *Dictyostelium discoideum* NDK (1NDC, cyan) bound to TDP (black, sticks) adopts the conventional binding mode compared to a novel binding mode identified in the crystal structure of *A. fumigatus* NDK (6XPV, yellow) bound to TDP (white, sticks). The α3/η2 loop in *A. fumigatus* NDK is shifted deeper into the binding pocket, re-orientating Val93 (Val98 in *D. discoideum* NDK) into a position that would sterically clash with the β-phosphate of TDP in the conventional binding mode.

The structural and functional characterisation of *A. fumigatus* NDK using X-ray crystallography and enzymatic assay data exemplifies the intricate interplay that exists between protein-ligand binding poses and substrate selectivity. The identification of two binding poses adopted by CDP and TDP that are specific to *A. fumigatus* NDK highlights the species-specific discrepancies that exist amongst homologous proteins. Together, these data further our understanding of the role NDK plays in NTP biosynthesis and the structural data provide opportunities to explore potential allosteric binding surfaces to target in future antifungal drug projects.

Materials and methods

Expression and purification of His_6-NDK

The open reading frame of NDK (Accession ID: Afu5g03490, *A. fumigatus* Af293) fused to an N-terminal

His_6 tag with a Gly-Ser dipeptide linker was commercially synthesised by GenScript and sub-cloned into pET11a for transformation into *Escherichia coli* BL21 (λDE3) cells. This construct was used to express His_6-tagged protein for crystallography experiments. Transformants were cultured in Luria broth at 37 °C until an optical density of 0.6 was reached and protein expression was induced with 0.5 mM IPTG at 16 °C for 16 h. Cells were harvested via centrifugation and the pellets were resuspended in Buffer A (20 mM Tris pH 8.0, 500 mM NaCl, 10 mM imidazole, 2 mM β-mercaptoethanol) and stored at −80 °C.

Cells were lysed using a M110L microfluidizer processor (Microfluidics, Westwood, MA, USA) and then clarified by centrifugation. The cell lysate was loaded onto a 5 mL ZetaSep Nickel NTA column (EMP Biotech, Berlin, Germany), pre-equilibrated with Buffer A, washed with 6 column volumes of 10% Buffer B (20 mM Tris pH 8.0, 500 mM NaCl, 250 mM imidazole, 2 mM β-mercaptoethanol) and eluted with an imidazole gradient from 10 mM to 250 mM. Fractions containing His_6-NDK were pooled and dialysed overnight against Storage Buffer (50 mM Tris pH 8.0, 0.5 mM EDTA, 5% v/v glycerol, 1 mM DTT). Protein purity was analysed via SDS/PAGE. His_6-NDK was concentrated to 15 $mg{\cdot}mL^{-1}$ using an Amicon Ultra-15 Centrifugal Filter Unit (10 kDa MWCO), then stored at −80 °C for crystallography experiments.

Expression and purification of His_6-TEV-NDK

Inverse PCR primers were used to insert a tobacco etch virus (TEV) proteolysis site between the His_6-tag and the NDK ORF for affinity tag removal. Methods used to express and purify His_6-TEV-NDK using nickel affinity chromatography are identical to that of His_6-NDK. Fractions containing His_6-TEV-NDK were pooled, supplemented with EDTA (final concentration 1 mM) and DTT (final concentration 0.5 mM) and tag removal was initiated with TEV (final concentration of 0.3 $mg{\cdot}mL^{-1}$) whilst dialysing overnight against Storage Buffer. The protein sample was loaded onto a 5 mL ZetaSep Nickel NTA column, pre-equilibrated with Buffer A and the unbound fraction containing untagged NDK was collected. Fractions were pooled and loaded onto a Hi Prep 26/10 desalting column pre-equilibrated in Storage Buffer. Fractions containing untagged NDK were collected, pooled and concentrated to 3.8 $mg{\cdot}mL^{-1}$ using an Amicon Ultra-15 Centrifugal Filter Unit (10 kDa MWCO), then stored at −80 °C for enzyme assay experiments.

Crystallisation

Unbound-NDK crystals were grown via hanging drop vapour diffusion by mixing a 1 : 1 ratio of NDK (15 $mg{\cdot}mL^{-1}$) to well solution containing 21% PEG 3350 and 0.45 M sodium malonate (pH 7.0) and incubated at 16 °C. Full-sized, diamond-shaped crystals of approximately 100 μm in length formed within 1 week.

Co-crystals of NDK with various nucleoside diphosphates were grown via hanging drop vapour diffusion. NDK (15 $mg{\cdot}mL^{-1}$) was incubated with $MgCl_2$ (10 mM) and each NTP for 10 min on ice before mixing in a 1 : 1 ratio with well solution and incubated at 16 °C (Table 5). To crystallise CDP-NDK, 0.1 μL of seed stock was added to the 1 : 1 protein to well solution mixture. This seed stock was generated by crushing NDK crystals co-crystallised with UTP (20 mM) and $MgCl_2$ (20 mM) which were formed in 21% PEG 3350, 0.15 M NaCl, 0.1 M HEPES pH 7.5 and diluted 1 : 100 in well solution. Full-sized, cubic (100 μm × 100 μm) and rod-like (50 μm × 100 μm) crystals formed within 1 week.

Data collection and processing

A single crystal was mounted and cryo-protected in *Paratone-N* (Hampton Research, Aliso Viejo, CA, USA), then cryo-cooled in liquid nitrogen. Diffraction data were collected at a wavelength of 0.954 Å at the Australian Synchrotron, part of ANSTO, using the Macromolecular Crystallography MX1 beamline for all NDK structures except for CDP-bound NDK which used the Macromolecular Crystallography MX2 beamline and made use of the Australian Cancer Research Foundation (ACRF) detector at 100 K [39,40]. Diffraction data were collected using an oscillation angle of 1° (yielding 360 frames) or 0.1° (yielding 3600 frames) per data set using the MX1 or MX2 beamline, respectively. Data from the MX1 beamline were integrated using IMOSFLM, followed by scaling and merging using AIMLESS from the CCP4 program suite [41,42]. Data from the MX2 beamline were integrated using XDS, converted to an mtz format using POINTLESS and then scaling

Table 5. Representative enzyme activity curves for (A) GTP, (B) ATP, (C) ITP, (D) UTP, (E) CTP and (F) TTP.

NTP	Final concentration of NTP (mM)	Well solution
GTP	4.8	21% PEG 3350, 0.25 M NaCl, 0.1 M HEPES (pH 7.5)
UTP		19% PEG 3350, 0.25 M NaCl, 0.1 M HEPES (pH 7.5)
ITP		19.5% PEG 3350, 0.20 M NaCl, 0.1 M HEPES (pH 7.5)
ATP	5	21% PEG 3350, 0.25 M NaCl, 0.1 M HEPES pH 7.5
TTP		18.5% PEG 3350, 0.15 M NaCl, 0.1 M HEPES (pH 7.5)
CTP	6.8	20% PEG 3350, 0.3 M Ammonium sulphate

and merging was performed by AIMLESS [42–44]. Statistics of data refinement for all NDK structures are summarised in Table 1. Crystals of unbound-NDK, CDP-NDK, TDP-NDK and UDP-NDK belong to space group $P2_12_12_1$ and were hexamers in the asymmetric unit whereas crystals of ADP-NDK, GDP-NDK and IDP-NDK belong to space group $P2_122_1$, forming trimers in the asymmetric unit and a hexamer with symmetry mates.

Structure refinement

The phase problem of NDK was solved by molecular replacement using PHASER MR from the CCP4 program suite [45]. The NDK from *Leishmania braziliensis* was used as a search model (PDB ID: 4KPC) [33]. Iterative cycles of manual rebuilding were completed using COOT, followed by model refinement using PHENIX.REFINE and model validation using MOLPROBITY [46–48]. Structures of unbound- and NDP-bound NDK were refined to a final R/R_{free} ratio of 13–19%/18–25%. The statistics of model refinement for all solved NDK structures are summarised in Table 1. Interactions between residues and bound ligands were analysed using POSEVIEW and residues between the interfaces were analysed using PDBEPISA [49,50].

Analytical size-exclusion chromatography

The average oligomeric size of NDK was determined by analytical SEC. Samples (50 μM) were loaded onto a Superdex 200 10/300 GL analytical SEC (GE Healthcare, Chicago, IL, USA), which had been equilibrated in either 50 mM phosphate buffer (PB; pH 7.4), PBS (pH 7.4) or 200 mM ammonium acetate (NH_4OAc; pH 6.8) at a flow rate of 0.4 mL·min^{-1} at room temperature. The size-exclusion column was calibrated using standards (Sigma) containing bovine thyroglobulin (670 kDa), bovine γ-globulin (158 kDa), chicken ovalbumin (44 kDa), horse myoglobin (17 kDa) and bovine aprotinin (7 kDa) (Fig. 13).

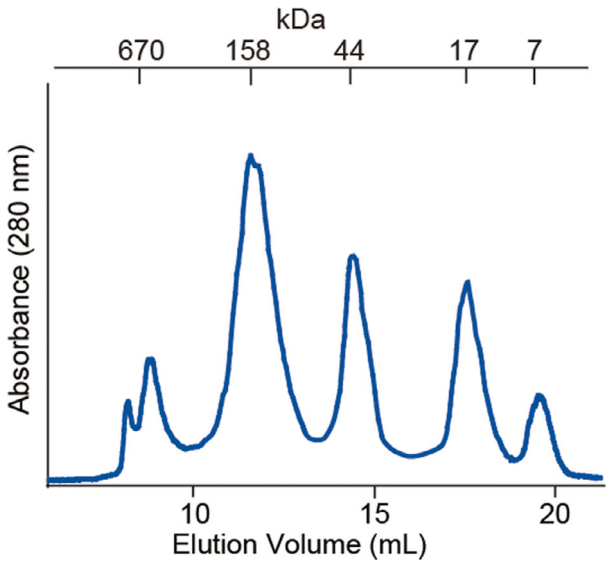


Fig. 13. Superdex 200 10/300 GL analytical SEC (GE Healthcare) profile calibrated using sizing standards (Sigma). Protein standards include bovine thyroglobulin (670 kDa), bovine γ-globulin (158 kDa), chicken ovalbumin (44 kDa), horse myoglobin (17 kDa) and bovine aprotinin (7 kDa).

Native ion-mobility mass spectrometry

The quaternary structure and dynamics of NDK was also examined using a Synapt G1 HDMS (Waters, Milford, MA, USA) using a nanoelectrospray ionisation source. NDK was dialysed into 200 mM NH_4OAc (pH 6.8) overnight at 4 °C. Unbound-NDK concentration was adjusted to 12.5 μM whilst ADP-NDK (12.5 μM) was formed by addition of ATP (5 mM) and magnesium acetate (5 mM) and desalted into 200 mM NH_4OAc using Biospin 6 columns (Bio-Rad, Hercules, CA, USA). Across all samples, 2 μL of protein was loaded onto gold-coated borosilicate glass capillaries prepared in-house. Instrument conditions were optimised to maintain non-covalent protein assemblies and key instrument parameters were as follows: capillary voltage: 1.60 kV; sampling cone: 50 V; extraction cone: 1.5 V; trap/transfer collision energy: 20/15 V; trap gas: 5.5 L·h^{-1}; backing gas: ~4.5 mbar. Ion-mobility (IM) MS parameters include IM cell wave height, 9 V; IM cell wave velocity, 350 m·s^{-1}; transfer t-wave height, 9 V; transfer t-wave velocity, 250 m·s^{-1}. Collison-induced dissociation (CID) MS was performed by increasing the trap-transfer collision energy up to 160 V (20 V increments). Titration of acetonitrile (5%, 10%, 20% and 30% v/v) was used for solution-induced dissociation of NDK hexamers. All mass spectra and arrival time distributions (ATDs) were analysed using MassLynx (v4.1) and DriftScope (v2.1) (both Waters).

NDK activity assay

NDK activity was monitored using a spectrophotometric assay adapted from methods described previously [37]. Analysis of binding kinetics of GTP, CTP, ITP, UTP and TTP utilised the hexokinase/glucose-6-phosphate dehydrogenase coupled enzyme system [37]. Reactions were prepared to a final concentration of 50 mM Tris (pH 8.0), 0.15 mM ADP, 5 mM $MgCl_2$, 10 mM KCl, 1 mM glucose, 0.2 mM $NADP^+$, 0.4 units hexokinase, 0.2 units glucose-6-phosphate dehydrogenase and 1.3 nM *A. fumigatus* NDK. Reactions were initiated with the addition of various concentrations of nucleoside triphosphates (GTP; 1–0.016 mM, CTP; 1.5–0.031 mM, ITP; 2–0.063 mM, UTP; 3–0.063 mM, TTP; 5–0.16 mM) at 30 °C and the production of NADPH was monitored at 340 nm using a PHERAstar FSX microplate

reader (BMG Lab Technologies, Offenburg, Germany). Analysis of the binding kinetics of ATP utilised the pyruvate kinase/lactate dehydrogenase couple enzyme system [37,51]. Reactions were prepared to a final concentration of 50 mM Tris (pH 8.0), 0.7 mM TDP, 5 mM $MgCl_2$, 10 mM KCl, 1 mM phosphoenolpyruvate, 0.2 mM NADH, 1.7 units pyruvate kinase, 2.6 units lactate dehydrogenase and 0.65 nM *A. fumigatus* NDK. Reactions were initiated with the addition of various concentrations of ATP (1–0.016 mM) at 30 °C and the consumption of NADH was monitored at 340 nm using a PHERAstar FSX microplate reader. Calculation of the Michaelis–Menten binding constants of all nucleoside triphosphates utilised three replicates plated in duplicate. All kinetic data were analysed using Prism 8.0 (GraphPad) and fitted to the Michaelis–Menten equation. Error propagation was calculated as follows: $\Delta\ (k_{cat}/K_M\ NDP) = (k_{cat}/K_M\ NDP)[(\Delta\ k_{cat}/k_{cat})^2 + (\Delta\ K_M\ NDP/K_M\ NDP)^2]^{1/2}$.

Calculation of binding energy change

The solved structures of NDK bound to ADP, GDP, IDP, UDP, CDP and TDP were imported into ICM-Pro Version 3.8-7b (MolSoft L. L. C., San Diego, CA, USA). The binding free energy change was calculated between ligand (NDPs) and protein (NDK) for each monomer and then averaged. The binding free energy change is calculated according to the following equation: $\Delta G_{bind} = (E^{comp}_{intra} - E^{parts}_{intra}) + (\Delta G^{comp}_{solv} - \Delta G^{parts}_{solv})$ where ΔG_{bind} is the binding free energy change, E^{comp} is the energy of the complex, E^{parts} is the interaction energy between the interacting parts, calculated from internal (*intra*) factors, ΔG^{comp}_{solv} represents the solvation free energy of the complex and ΔG^{parts}_{solv} represents the solvation free energy from the interacting parts.

Accession codes

The atomic coordinates and structure factors of unbound-NDK, ADP-NDK, GDP-NDK, IDP-NDK, UDP-NDK, CDP-NDK and TDP-NDK have been deposited to the Protein Data Bank with the following accession codes: 6XP4, 6XP7, 6XPS, 6XPU, 6XPT, 6XPW, 6XPV, respectively.

Acknowledgements

We acknowledge the support of the ANSTO in providing the facility used in this work. We thank Flinders Analytical (Flinders University, Australia) for access to IM-MS instrumentation. We thank Jun Aishima from the Australian Synchrotron for his expertise. SN is a recipient of an Australian Government Research Training Program stipend scholarship.

Conflict of interest

The authors declare no conflict of interest.

Author contributions

SN and JBB conceptualised the study; SN, BJ, TLP and JBB planned the experiments; SN and BJ performed the experiments; SN and BJ analysed the data; SN prepared the manuscript; and SN, BJ, TLP and JBB reviewed and edited the manuscript.

Peer Review

The peer review history for this article is available at https://publons.com/publon/10.1111/febs.15607.

References

1 Kwon-Chung KJ & Sugui JA (2013) *Aspergillus fumigatus* – what makes the species a ubiquitous human fungal pathogen? *PLoS Pathog* **9**, e1003743.
2 Dagenais TRT & Keller NP (2009) Pathogenesis of *Aspergillus fumigatus* in invasive aspergillosis. *Clin Microbiol Rev* **22**, 447–465.
3 Bongomin F, Gago S, Oladele RO & Denning DW (2017) Global and multi-national prevalence of fungal disease – estimate precision. *J Fungi* **3**, 57.
4 Baddley JW, Andes DR, Marr KA, Kontoyiannis DP, Alexander BD, Kauffman CA, Oster RA, Anaissie EJ, Walsh TJ, Schuster MG *et al.* (2010) Factors associated with mortality in transplant patients with invasive aspergillosis. *Clin Infect Dis* **50**, 1559–1567.
5 Nucci M & Perfect JR (2008) When primary antifungal therapy fails. *Clin Infect Dis* **46**, 1426–1433.
6 Morrow CA, Valkov E, Stamp A, Chow EWL, Lee IR, Wronski A, Williams SJ, Hill JM, Djordjevic JT, Kappler U *et al.* (2012) *De novo* GTP biosynthesis is citical for virulence of the fungal pathogen *Cryptococcus neoformans*. *PLoS Pathog* **8**, e1002957.
7 Chitty JL, Tatzenko TL, Williams SJ, Koh YQ, Corfield EC, Butrler MS, Roberston AA, Cooper MA, Kappler U, Kobe B *et al.* (2017) GMP synthase is required for virulence factor production and infection by *Cryptococcus neoformans*. *J Biol Chem* **292**, 3049–3059.
8 Rodriguez-Suarez R, Xu D, Veillette K, Davison J, Sillaots S, Kauffman S, Hu W, Bowman J, Martel N, Trosok S *et al.* (2007) Mechanism-of-action determination of GMP synthase inhibitors and target validation in *Candida albicans* and *Aspergillus fumigatus*. *Chem Biol* **14**, 1163–1175.
9 Chitty JL & Fraser JA (2017) Purine acquisition and synthesis by human fungal pathogens. *Microorganisms* **5**, E33.

10 Cortez J & Schnitzer M (1979) Purines and pyrimidines in soils and humic substances. *Soil Sci Soc Am J* **43**, 958–961.
11 Traut TW (1994) Physiological concentrations of purines and pyrimidines. *Mol Cell Biochem* **140**, 1–22.
12 Dinamarco TM, Brown NA, Couto de Almeida RS, Alves de Castro P, Savoldi M, de Souza Goldman MH & Goldman GH (2012) *Aspergillus fumigatus* calcineurin interacts with a nucleoside diphosphate kinase. *Microbes Infect* **14**, 922–929.
13 Lin X, Momany C & Momany M (2003) SwoHp, a nucleoside diphosphate kinase, is essential in *Aspergillus nidulans*. *Eukaryot Cell* **2**, 1169–1177.
14 Souza TA, Trindade DM, Tonoli CC, Santos CR, Ward RJ, Arni RK, Oliveira AH & Murakami MT (2011) Molecular adaptability of nucleoside diphosphate kinase b from trypanosomatid parasites: stability, oligomerization and structural determinants of nucleotide binding. *Mol Biosyst* **7**, 2189–2195.
15 Wang Y, Wang S, Nie X, Yang K, Xu P, Wang X, Liu M, Yang Y, Chen Z & Wang S (2019) Molecular and structural basis of nucleoside diphosphate kinase–mediated regulation of spore and sclerotia development in the fungus *Aspergillus flavus*. *J Biol Chem* **294**, 12415–12431.
16 Giraud MF, Georgescauld F, Lascu I & Dautant A (2006) Crystal structures of S120G mutant and wild type of human nucleoside diphosphate kinase A in complex with ADP. *J Bioenerg Biomembr* **38**, 261–264.
17 Schaertl S, Konrad M & Geeves MA (1998) Substrate specificity of human nucleoside-diphosphate kinase revealed by transient kinetic analysis. *J Biol Chem* **273**, 5662–5669.
18 Moréra S, Chiadmi M, LeBras G, Lascu I & Janin J (1995) Mechanism of phosphate transfer by nucleoside diphosphate kinase: X-ray structures of phosphohistidine intermediate of the enzymes from *Drosophila* and *Dictyostelium*. *Biochemistry* **34**, 11062–11070.
19 Lascu I & Gonin P (2000) The catalytic mechanism of nucleoside diphosphate kinase. *J Bioenerg Biomembr* **32**, 237–246.
20 Muñoz-Dorado J, Inouye S & Inouye M (1990) Nucleoside diphosphate kinase from *Myxococcus xanthus*. II. Biochemical characterization. *J Biol Chem* **265**, 2707–2712.
21 Jeudy S, Lartigue A, Claverie JM & Abergel C (2009) Dissecting the unique nucleotide specificity of mimivirus nucleoside diphosphate kinase. *J Virol* **83**, 7142–7150.
22 Hu Y, Jia X, Lu Z & Han L (2019) Characterization of crystal structure and key residues of *Aspergillus fumigatus* nucleoside diphosphate kinase. *Biochem Biophys Res Commun* **511**, 148–153.
23 Lascu I, Chaffotte A, Limbourg-Bouchon B & Veron M (1992) A pro/ser substitution in nucleoside diphosphate kinase of *Drosophila melanogaster* (mutation killer of prune) affects stability but not catalytic efficiency of the enzyme. *J Biol Chem* **267**, 12775–12781.
24 Lascu I, Deville-Bonne D, Glaser P & Véron M (1993) Equilibrium dissociation and unfolding of nucleoside diphosphate kinase from *Dictyostelium discoideum*. *J Biol Chem* **268**, 20268–20275.
25 Ponstingl H, Henrick K & Thornton JM (2000) Discriminating between homodimeric and monomeric proteins in the crystalline state. *Proteins* **41**, 47–57.
26 Benesch JL & Ruotolo BT (2011) Mass spectrometry: come of age for structural and dynamical biology. *Curr Opin Struct Biol* **21**, 641–649.
27 Jovcevski B & Pukala TL (2019) Mass Spectrometry and Its Applications, 1st edn. Hoboken, NJ: John Wiley & Sons Ltd..
28 Srivastava SK, Rajasree K & Gopal B (2011) Conformational basis for substrate recognition and regulation of catalytic activity in *Staphylococcus aureus* nucleoside di-phosphate kinase. *Biochim Biophys Acta* **1814**, 1349–1357.
29 Tian W, Chen C, Lei X, Zhao J & Liang J (2018) CASTp 3.0: computed atlas of surface topography of proteins. *Nucl Acids Res* **46**, W363–W367.
30 Semianrio-Vidal L, van Hesuden C, Mugesh G & Lazarowski ER (2010) Ebselen is a potent non-competitive inhibitor of extracellular nucleoside diphosphokinase. *Purinerg Signal* **6**, 383–391.
31 Wang H, Bao R, Jiang C, Yang Z, Zhou C-Z & Chen Y (2008) Structure of Ynk1 from the yeast *Saccharomyces cerevisiae*. *Acta Crystallogr F* **64**, 572–576.
32 Qian L & Liu X (2014) Purification, characterization and structure of nucleoside diphosphate kinase from *Drosophila melanogaster*. *Protein Expr Purif* **103**, 48–55.
33 Vieira PS, Giuseppe PO, Santos CR, Cunha EMF, de Oliveira AHC & Murakami MT (2015) Crystal structure and biophysical characterization of the nucleoside diphosphate kinase from *Leishmania braziliensis*. *BMC Struct Biol* **3**, 2.
34 Karlsson A, Mesnildrey S, Xu Y, Moréra S, Janin J & Véron M (1996) Nucleoside diphosphate kinase. Investigation of the intersubunit contacts by site-directed mutagenesis and crystallography. *J Biol Chem* **271**, 19928–19934.
35 Georgescauld F, Moynié L, Habersetzer J, Cervoni L, Mocan I, Borza T, Harris P, Dautant A & Lascu I (2013) Intersubunit ionic interactions stabilize the nucleoside diphosphate kinase of *Mycobacterium tuberculosis*. *PLoS One* **8**, e57867.
36 Cherfils J, Moréra S, Lascu I, Véron M & Janin J (1994) X-ray structure of nucleoside diphosphate kinase complexed with thymidine diphosphate and Mg^{2+} at 2-A resolution. *Biochemistry* **33**, 9062–9069.

37 Mourad N & Parks REJ (1966) Erythrocytic nucleoside disphosphokinase. II. Isolation and kinetics. *J Biol Chem* **241**, 271–278.
38 Biondi RC, Veron M, Walz K & Passeron S (1995) *Candida albicans* nucleoside-diphosphate kinase: purification and characterization. *Arch Biochem Biophys* **323**, 187–194.
39 Cowieson NP, Aragao D, Clift M, Ericsson DJ, Gee C, Harrop SJ, Mudie N, Panjikar S, Price JR, Riboldi-Tunnicliffe A *et al*. (2015) MX1: a bending-magnet crystallography beamline serving both chemical and macromolecular crystallography communities at the Australian Synchrotron. *J Synchrotron Radiat* **22**, 187–190.
40 Aragao D, Aishima J, Cherukuvada H, Clarken R, Clift M, Cowieson NP, Ericsson DJ, Gee C, Macedo S, Mudie N *et al*. (2018) MX2: a high-flux undulator microfocus beamline serving both the chemical and macromolecular crystallography communities at the Australian Synchrotron. *J Synchrotron Radiat* **25**, 885–891.
41 Battye TGG, Kontogiannis L, Johnson O, Powell HR & Leslie AGW (2011) iMOSFLM: a new graphical interface for diffraction-image processing with MOSFLM. *Acta Crystallogr D* **67**, 271–281.
42 Evans PR & Murshudov GN (2013) How good are my data and what is the resolution? *Acta Crystallogr D* **69**, 1204–1214.
43 Kabsch W (2010) XDS. *Acta Crystallogr D* **66**(Pt 2), 125–132.
44 Evans PR (2006) Scaling and assessment of data quality. *Acta Crystallogr D* **62**, 72–82.
45 McCoy AJ, Grosse-Kunstleve RW, Adams PD, Winn MD, Storoni LC & Read RJ (2007) Phaser crystallographic software. *J App Crystallogr* **40**, 658–674.
46 Emsley P & Cowtan K (2004) Coot: model-building tools for molecular graphics. *Acta Crystallogr D* **60**, 2126–2132.
47 Afonine PV, Grosse-Kunstleve RW, Echols N, Headd JJ, Moriarty NW, Mustyakimov M, Terwilliger TC, Urzhumtsev A, Zwart PH & Adams PD (2012) Towards automated crystallographic structure refinement with phenix.refine. *Acta Crystallogr D* **68**, 352–367.
48 Chen VB, Arendall WB, Headd JJ, Keedy DA, Immormino RM, Kapral GJ, Murray LW, Richardson JS & Richardson DC (2010) *MolProbity*: all-atom structure validation for macromolecular crystallography. *Acta Crystallogr D* **66**, 12–21.
49 Stierand K & Rarey M (2010) PoseView – molecular interaction patterns at a glance. *J Cheminform* **2**, P50.
50 Krissinel E & Henrick K (2007) Interference of macromolecular assemblies from crystalline state. *J Mol Biol* **372**, 774–797.
51 Bergmeyer HU (1974) Methods of Enzymatic Analysis, 2nd edn. New York, USA: Academic Press.
52 Notredame C, Higgins DG & Heringa J (2000) T-Coffee: a novel method for fast and accurate multiple sequence alignment. *J Mol Biol* **302**, 205–217.
53 Robert X & Gouet P (2014) Deciphering key features in protein structures with the new ENDscript server. *Nucl Acids Res* **42**, W320–W324.
54 Letunic I & Bork P (2019) Interactive Tree of Life (iTOL) v4: recent updates and new developments. *Nucl Acids Res* **47**, W256–W259.
55 Sievers F, Wilm A, Dineen D, Gibson TJ, Karplus K, Li W, Lopez R, McWilliam H, Remmert M, Söding J *et al*. (2011) Fast, scalable generation of high-quality protein multiple sequence alignments using Clustal Omega. *Mol Syst Biol* **7**, 539.
56 Afonine PV, Moriarty NW, Mustyakimov M, Sobolev OV, Terwilliger TC, Turk D, Urzhumtsev A & Adams PD (2015) FEM: feature-enhanced map. *Acta Crystallogr D* **71**, 646–666.
57 Arndt UW, Crowther RA & Mallet JFW (1968) A computer-linked cathode ray tube microdensitometer for X-ray crystallography. *J Phys E Sci Instrum* **1**, 510–516.
58 Diederichs K & Karplus PA (1997) Improved R-factors for diffraction data analysis in macromolecular crystallography. *Nat Struct Biol* **4**, 269–275.
59 Weiss MS & Hilgenfeld R (1997) On the use of merging *R*-factor as a quality indicator for X-ray data. *J Appl Crystallogr* **30**, 203–205.
60 Karplus PA & Diederichs K (2012) Linking crystallographic model and data quality. *Science* **336**, 1030–1033.
61 Brünger AT (1992) Free R value: a novel statistical quantity for assessing the accuracy of crystal structures. *Nature* **355**, 472–475.

Chapter 3 (Manuscript):

Structural insights into the antifungal drug target, guanosine monophosphate synthase from *Aspergillus fumigatus*

Chapter 3 Preface

Chapter 3 is a manuscript submitted for publication to Antimicrobial Agents and Chemotherapy that describes the characterisation of the structure and function of guanosine monophosphate synthase from *A. fumigatus* using X-ray crystallography, native mass spectrometry and *in vitro* enzyme assays. As a proof of concept, we have demonstrated the inhibitory activity of glutamine analogues, acivicin and 6-diazo-5-oxo-L-norleucine, against this enzyme. I performed all X-ray crystallography experiments and *in vitro* enzyme assays. I performed the data analysis for these experiments, prepared the figures included in the manuscript, deposited structure factors and coordinate factors into the Protein Data Bank and prepared the manuscript.

Statement of Authorship

Title of Paper	Structural insights into the antifungal drug target, guanosine monophosphate synthase from *Aspergillus fumigatus*
Publication Status	☐ Published ☐ Accepted for Publication ☑ Submitted for Publication ☐ Unpublished and Unsubmitted work written in manuscript style
Publication Details	This publication characterises the structural and kinetic properties of guanosine monophosphate synthase from *A. fumigatus* and demonstrates the inhibitory activity of two glutamine analogues against this enzyme.

Principal Author

Name of Principal Author (Candidate)	Stephanie Nguyen		
Contribution to the Paper	Project conceptualisation, planned and performed experiments, analysed data, prepared, reviewed and edited the manuscript		
Overall percentage (%)	65		
Certification:	This paper reports on original research I conducted during the period of my Higher Degree by Research candidature and is not subject to any obligations or contractual agreements with a third party that would constrain its inclusion in this thesis. I am the primary author of this paper.		
Signature		Date	17/07/21

Co-Author Contributions

By signing the Statement of Authorship, each author certifies that:

i. the candidate's stated contribution to the publication is accurate (as detailed above);

ii. permission is granted for the candidate in include the publication in the thesis; and

iii. the sum of all co-author contributions is equal to 100% less the candidate's stated contribution.

Name of Co-Author	Dr. Blagojce Jovcevski		
Contribution to the Paper	Planned and performed experiments, analysed data, reviewed and edited the manuscript		
Overall Percentage (%)	15		
Signature		Date	20/07/21

Name of Co-Author	Assoc. Prof. Tara Pukala			
Contribution to the Paper	Planned experiments, reviewed and edited the manuscript			
Overall Percentage (%)	10			
Signature			Date	18/7/21

Name of Co-Author	Dr. John B. Bruning			
Contribution to the Paper	Project conceptualisation, planned experiments, reviewed and edited the manuscript			
Overall Percentage (%)	10			
Signature			Date	21/7/21

Structural insights into the antifungal drug target, guanosine monophosphate synthase from *Aspergillus fumigatus*

Short title: GMP synthase from *Aspergillus fumigatus*

Stephanie Nguyen[1], Blagojce Jovcevski[2, 3], Tara L. Pukala[2], John B. Bruning[1*]

[1] Institute of Photonics and Advanced Sensing (IPAS), School of Biological Sciences, The University of Adelaide, South Australia, 5005, Australia

[2] School of Physical Sciences, Department of Chemistry, The University of Adelaide, South Australia, 5005, Australia

[3] School of Agriculture, Food and Wine, The University of Adelaide, South Australia, 5005, Australia

* Corresponding author:
Email: john.bruning@adelaide.edu.au

Keywords: Aspergillus fumigatus, guanosine monophosphate synthase, nucleotide biosynthesis, purine biosynthesis, antifungal target, structural biology

Abstract

Purine biosynthesis is a fundamental cellular process that sustains life by maintaining the intracellular pool of purines for DNA/RNA synthesis and signal transduction. As an integral determinant of fungal survival and virulence, enzymes in this metabolic pathway have been pursued as potential antifungal targets. Guanosine monophosphate (GMP) synthase has been identified as an attractive target as it is essential for virulence in the clinically prominent fungal pathogens, *Aspergillus fumigatus*, *Candida albicans* and *Cryptococcus neoformans*. However, the lack of structural information for GMP synthase has hindered drug design efforts. We present the first structure of *A. fumigatus* GMP synthase (2.3 Å) of fungal origin. Structural analyses of GMP synthase shows a distinct absence of the D1 dimerisation domain which is present in the human homologue. Interestingly, GMP synthase adopts a dimeric state, as determined by native mass spectrometry and gel filtration chromatography, which is contrary to the monomeric human homologue. Analysis of the substrate binding pockets of GMP synthase reveals key differences in the ATP and XMP binding sites that can be exploited for species-specific inhibitor drug design. Furthermore, we demonstrate the inhibitory activity of the glutamine analogues, acivicin (IC_{50} = 16.6 ± 2.4 µM) and 6-diazo-5-oxo-L-norleucine (IC_{50} = 29.6 ± 5.6 µM) against *A. fumigatus* GMP synthase. Together, these data provide crucial structural information required for specifically targeting *A. fumigatus* GMP synthase for future antifungal drug discovery endeavours.

Introduction

An immunocompromised state can arise from inherited diseases or acquired from malignancies, viral infections, such as HIV/AIDS, and treatment regimens such as corticoid steroid therapy, immunosuppressive therapy and chemotherapy (1). A weakened immune system increases the risk of contracting life-threatening infections, and consequently, infectious diseases have become the leading cause of morbidity and mortality amongst immunocompromised patients (2). The Transplant-Associated Infections Surveillance Network (TRANSNET) have revealed that invasive aspergillosis, primarily caused by the opportunistic and pathogenic fungus *Aspergillus fumigatus*, was the first and second most common fungal infection amongst haematopoietic stem cell transplant (HSCT) patients (43%) and recipients of solid-organ transplants (19%), respectively (3, 4). Although commonly associated with human disease, *A. fumigatus* is omnipresent in the natural environment. It occupies various niches, including soil and compost heaps which contributes to nutrient recycling, consistent with its saprophytic lifestyle (5). It reproduces asexually by producing and disseminating fungal spores that readily travel through the air. In an immunocompetent host, the defence mechanisms commonly used to neutralise these spores are ineffective and this can lead to the development of a broad spectrum of diseases, known as aspergillosis (6). In severe cases, the fungus invades the lung tissue and can then infiltrate the blood stream to establish a secondary infection in surrounding organs (7). This leads to the development of a systemic infection known as invasive aspergillosis. Despite the use of currently available antifungal therapies, the mortality rate associated with invasive aspergillosis remains unacceptably high in both HSCT and solid-organ transplant patients with a one-year survival rate of 25.4% and 59%, respectively (3, 4). Poor mortality rates and the increased prevalence of strains resistant to existing antifungals highlights the need to identify and characterise alternate targets for the development of a new class of antifungals.

There has been a growing interest in the biosynthetic enzymes involved in purine metabolism in terms of their role in sustaining fungal survival and contribution to virulence (8). Purines are essential for DNA and RNA synthesis, signal transduction and energy metabolism (9). Therefore, the intracellular purine pool must be tightly maintained by *de novo* biosynthesis and salvage pathways, the former being the focus of current antifungal drug discovery efforts (10-12). The validity of targeting key enzymes in this pathway in pathogenic fungi including *A. fumigatus, Candida albicans* and *Cryptococcus neoformans* have been affirmed in genetic knockout studies (10, 12). For example, inosine monophosphate dehydrogenase (IMPDH), an enzyme that catalyses the first committed step of guanosine triphosphate synthesis, has also been implicated in the production of several fungal virulence factors. IMPDH deficiencies in *C. neoformans* result in a reduction of capsule size and melanin expression, ultimately producing an avirulent strain when tested in a murine model of Cryptococcus (12). In the later stages of *de novo* purine biosynthesis, newly synthesised nucleoside diphosphates are converted to nucleoside triphosphates for use in downstream pathways. This reaction is catalysed by nucleoside diphosphate kinase, a highly conserved enzyme that has been shown to be essential for viability in several fungi including *A. fumigatus* and *A. nidulans* (13, 14). Both enzymes have been extensively characterised in terms of their structure and function and have been explored as potential antifungal drug targets (12, 15).

However, current work in the field and in this study has focused on guanosine monophosphate (GMP) synthase as a target to disrupt the purine biosynthesis pathway (8). This enzyme utilises xanthosine 5'-monophosphate (XMP) to produce GMP. This occurs via a two-step reaction where L-glutamine is first converted to L-glutamate via the activity of the N-terminal glutamine amidotransferase (GATase) domain. The reaction releases ammonia which is shuttled to the ATP pyrophosphatase (ATP-PPase) domain and is subsequently used to convert XMP to GMP, a precursor required to synthesise nucleoside triphosphates downstream in the pathway (16).

Reflective of the essential nature of purine nucleotides in fungal survival, disruption of GMP synthase by a conditional knockout in *A. fumigatus* and *C. albicans* or deletion of the encoding gene in *C. neoformans* resulted in growth defects on minimal media (10, 11). Furthermore, these GMP synthase-deficient mutants of *A. fumigatus*, *C. albicans* and *C. neoformans* were avirulent in murine models of infection (10, 11). Characterisation of the *C. neoformans* mutant revealed defects in the production of the capsule and melanin, as well as reduced activity of secreted proteases, all of which are known virulence factors (11). The indispensable role undertaken by GMP synthase in several aspects of fungal virulence reaffirms its validity as a broad-spectrum antifungal drug target. Thus, there is a need to develop a greater understanding of the biophysical characteristics of GMP synthase, including its structure, biological assembly and function, in order to effectively inhibit its activity. In particular, intimate knowledge of its three-dimensional structure will be imperative in the design of a drug that is both potent and selective for fungal GMP synthase over the human equivalent.

In this study, we have established the foundations for a drug discovery project that targets fungal GMP synthase. To guide rational drug design, we present the first crystal structure of a fungal homologue of GMP synthase from *A. fumigatus* (2.3 Å) and have rigorously characterised the biophysical and kinetic properties of this enzyme. We have also validated the inhibition of *A. fumigatus* GMP synthase activity *in vitro* using known inhibitors of the GATase domain, acivicin (ACI) and 6-diazo-5-oxo-L-norleucine (DON). Ultimately, these data provide essential structural and functional information required to guide rational drug design of an inhibitor that selectively targets fungal GMP synthase.

Results

Structure, domain architecture and biological assembly of A. fumigatus GMP synthase

The crystal structure of GMP synthase from *A. fumigatus* was solved to a resolution of 2.3 Å and had two molecules in the asymmetric unit. Each molecule of GMP synthase consists of 14 α-helices, 19 β-strands and 7 3_{10}-helices. There are loop regions within the structure that exhibit high mobility, reflective of their poor electron density. Hence, a 20-residue gap in the first monomer (residues 365 – 384) and a 5-residue gap in the second monomer (residues 366 – 373) were not modelled. GMP synthase from *A. fumigatus* possesses three distinct domains – an N-terminal GATase domain (residues 1 – 205), a C-terminal ATP-PPase domain (residues 213 – 431), both connected via a short loop (residues 206 – 212) and a dimerization domain, denoted as D2 according to the nomenclature used in the characterisation of human GMP synthase (residues 432 – 540) (Figure 1A). The GATase domain is classified as a Class I type, defined by its highly conserved catalytic triad, corresponding to residues Cys90, His184 and Glu186 in the *A. fumigatus* GMP synthase structure. This domain is responsible for the removal and transfer of ammonia from glutamine to the ATP-PPase domain where GMP is synthesised via an adenyl-XMP intermediate (17). The dimerization domain, D2, facilitates formation of the dimeric complex (Figure 1B).

There are two molecules of *A. fumigatus* GMP synthase in the asymmetric unit, forming an interface along the GATase domain of one monomer and the ATP-PPase domain of the other monomer. The calculated surface area of this interface is 323 $Å^2$ and so it is not likely to be reflective of the true oligomeric structure. However, a single monomer of GMP synthase can form a dimer with its symmetry mate via the D2 dimerization domain. This structural arrangement is representative of the biological assembly likely to be adopted by GMP synthase based on the structural function of dimerization domains and the large interface surface area of 1690 $Å^2$. The complementary techniques of analytical size-exclusion chromatography (SEC) and native mass spectrometry (MS) confirm the presence of a dimeric species (Figure 2). A predominant elution peak at 12.8 mL (~118 kDa) is observed in the size-exclusion chromatogram which corresponds to the dimeric form of GMP synthase (theoretical mass of 119.1 kDa) (Figure 2A). The oligomeric state was confirmed by native MS which also shows GMP synthase predominantly forms a dimer (measured mass 119.2 kDa) (Figure 2B, blue). GMP synthase in the presence of glutamine remains dimeric (Figure 2B) with an observed shift in charge state distribution of the dimer towards lower charge states (from 23+ to 20+ being the most dominant). This suggests a global conformational change (e.g. compaction or rearrangement) where basic residues on the dimeric surface are unable to be ionised. We further probed the conformational differences in GMP synthase dimers using IM-MS (Figure 2C-D) with the dimer^{21+} ions selected for comparative analysis (Figure 2B, grey box). The dimer^{21+} was selected for comparative analysis as it is easily observed in both the absence and presence of glutamine, as well as avoiding charge state overlap from other oligomeric species (monomer or tetramer). IM-MS spectra of GMP synthase dimers in both the absence (dimer^{21+} 10.5 ms, Figure 2C, blue) and presence (dimer^{21+} 10.8 ms, Figure 2C, green) of glutamine exhibit similar arrival time distributions (ATD) (Figure 2C). The ATD is a measure of the rotationally averaged collision cross section or ‘size’ of an ion, providing key structural information

149 including conformational changes and structural dynamics. Interestingly, the full width-half
150 maximum (FWHM) of the ATD of GMP synthase dimers in the absence of glutamine (4.53
151 ms) is significantly greater compared to GMP synthase in the presence of glutamine (2.21 ms)
152 (Figure 2D). The FWHM measurement provides valuable insights into structural dynamics as
153 it is a measure of structural heterogeneity, whereby individual oligomeric states sampling a
154 greater number of conformations (i.e. opening and closing) give rise to broader ATDs (15, 18-
155 20). This data, combined with the observed charge state distribution shifts, indicates that the
156 GMP synthase dimers adopt a conformational change with increased structural heterogeneity
157 upon glutamine addition. Furthermore, it is possible that this conformational change is needed
158 to enable ammonia tunnelling to the ATP-PPase catalytic site.

159 *Kinetic characterisation of GMP synthase enzyme activity*

160 The apparent Michaelis-Menten binding constants (K_M) of *A. fumigatus* GMP synthase were
161 determined for all substrates, apart from XMP, using a spectrophotometric activity assay. As
162 the method of detection used in this activity assay was not sufficiently sensitive at low
163 concentrations of XMP, the K_M value could not be accurately determined. The representative
164 binding curves used to determine the kinetic parameters are shown in Figure 3. The measured
165 binding constants were compared to homologous enzymes from prokaryotic sources, including
166 *Escherichia coli*, and eukaryotic sources, including the fungal pathogen *C. neoformans* and the
167 human equivalent of GMP synthase (Table 1). The Michaelis-Menten binding constants
168 measured for ATP and glutamine were consistently higher for *A. fumigatus* GMP synthase,
169 relative to that of *E. coli*, *C. neoformans* and *H. sapiens*. The binding kinetics of Mg^{2+} to *A.*
170 *fumigatus* GMP synthase appeared hyperbolic and from this data, a K_M value was calculated,
171 as opposed to a $K_{0.5}$ calculated for both *C. neoformans* and *H. sapiens* GMP synthase. The
172 hyperbolic fit indicates that Mg^{2+} is not subject to cooperative binding in *A. fumigatus* GMP
173 synthase.

Table 1: Apparent Michaelis-Menten binding constants (K_M) calculated for substrates of GMP synthase from *A. fumigatus* and compared to those of *Escherichia coli*, *Cryptococcus neoformans* and *Homo sapiens*. K_M values were calculated from three replicates of activity assays performed in triplicate.

Species	$k_{cat\ (ATP)}$ **(s^{-1})**	$K_{M\ (ATP)}$ **(µM)**	$k_{cat\ (ATP)}/K_M$ **($µM^{-1}\ s^{-1}$)**	$K_{M\ (XMP)}$ **(µM)**	$K_{M\ (Gln)}$ **(µM)**	$K_{M\ (Mg^{2+})}$ **(µM)**	$K_{0.5\ (Mg^{2+})}$ **(µM)**	**Reference**
A. fumigatus[a]	1657 ± 105.5	245 ± 15	6.76 ± 0.60	ND	2693 ± 119	1230 ± 140	NA	This study
E. coli[b]	ND	104 ± 44	ND	166 ± 43	ND	ND	ND	(21)
C. neoformans[c]	ND	77.5 ± 6	ND	65.9 ± 13	1130 ± 162	NA	1289 ± 66	(11)
H. sapiens[d]	ND	132 ± 7	ND	35.6 ± 1.8	406 ± 49	NA	1780 ± 70	(22)

* ND = not determined, NA = not applicable

[a] Enzyme assay conditions: 50 mM HEPES pH 7.5, 0.5 mM XMP, 2.5 mM ATP, 20 mM glutamine, 20 mM $MgCl_2$ at 40 °C

[b] Enzyme assay conditions: 60 mM HEPES pH 8.0, 5 mM ATP, 0.2 mM XMP, 20 mM $MgCl_2$, 200 mM NH_4Cl, 0.1 mM DTT, 0.8 mM EDTA at 40 °C

[c] Enzyme assay conditions: 50 mM HEPES pH 7.5, ATP, XMP and glutamine or $(NH_4)_2SO_4$ (at varying concentrations), 20 mM $MgCl_2$, 0.3 mM EDTA at 40 °C

[d] Enzyme assay conditions: 75 mM Tris-HCl pH 7.6, 2 mM ATP, 0.25 mM XMP, 2 mM glutamine, 10 mM $MgSO_4$ at 40 °C

Prokaryotic GMP synthases show high structural conservation to A. fumigatus GMP synthase

The architecture and spatial arrangement of domains in *A. fumigatus* GMP synthase closely resembles homologous enzymes from parasitic (*Plasmodium falciparum*, PDB: 4WIM) and prokaryotic (*Thermus thermophilus*, PDB: 2YWB; *Neisseria gonorrhoeae*, PDB: 5TW7; *Coxiella burnetii*, PDB: 3TQI; *Escherichia coli*, PDB: 1GPM) sources (23-27). Comparison of these GMP synthase structures reveal a conserved overall structural fold and preservation of the GATase and ATP-PPase functional domains as well as the D2 structural domain (Figure 4). Protein sequence alignments of homologous GMP synthases are also consistent with this observation and also reveal the presence of a large portion of the sequence that is unique to human GMP synthase (Figure 5). The greatest sequence deviation appears between *A. fumigatus* and *P. falciparum* GMP synthase (42% sequence identity) and consequently, the greatest structural perturbations are observed between these structures (root mean square deviation, RMSD of 4.1 Å, 535 C_{α}). GMP synthase from *A. fumigatus* bears the greatest resemblance to homologous enzymes from prokaryotic origins in terms of their sequence identity (47% - 49%) and structural similarity, as reflected by the small RMSD values measured (2.0 Å – 2.8 Å).

Human and A. fumigatus GMP synthase have different biophysical features

Comparison of the *A. fumigatus* GMP synthase structure, elucidated in this study, with the existing structure of human GMP synthase (PDB: 2VXO) reveals species-specific differences that may be exploited to design an inhibitor selective for the fungal equivalent. There are components of the domain architecture that are shared between *A. fumigatus* and human GMP synthase, including the GATase and ATP-PPase catalytic domains and the D2 dimerization domain (Figure 6A, B). However, human GMP synthase contains an insert between the ATP-PPase and D2 dimerization domain that forms an additional dimerization domain, denoted D1 (28). Superimposition of the *A. fumigatus* and human GMP synthase structures reveal significant perturbations in the overall arrangement of the individual domains (Figure 6B). However, alignment of the GATase domain (RMSD of 2.5 Å, 185 C_α), ATP-PPase domain (RMSD of 4.3 Å, 216 C_α) and D2 dimerization sub-domain (3.5 Å, 109 C_α) individually from *A. fumigatus* and human GMP synthase reveals high structural similarity reflective of the conservation of their function.

Despite the presence of two dimerization domains, D1 and D2, which show a high degree of structural similarity between them, human GMP synthase is predominantly monomeric in solution with some capability of forming a dimeric complex in the presence of substrates (28). This is in contrast to *A. fumigatus* GMP synthase that has been shown to form a dimeric interface via the D2 dimerization sub-domain, as demonstrated in the crystal structure, and confirmed using both analytical SEC and native mass spectrometry (Figures 1B and 2).

XMP- and ATP-binding site composition of GMP synthase from A. fumigatus and human reveals targetable differences for antimicrobial drug design

Structural alignment of the ATP-PPase domain from *A. fumigatus* and human GMP synthase reveals crucial deviations in the XMP-binding site composition that may be exploited to

achieve inhibitor selectivity. There is an invariant peptide sequence (Pro-Phe-Pro-Gly-Pro-Gly) found in homologues from both *A. fumigatus* and *H. sapiens* (Figure 6C). This motif, located on a flexible loop region, works in tandem with the more variable motif (Ser-Gly-Lys-Ala) to stabilise the xanthine base of XMP in the human GMP synthase structure. Due to its location, relative to XMP, both motifs are likely to be crucial for substrate binding (28). Of the 23 residues identified in the XMP-binding site, 5 residues are markedly different between human and fungal homologues. This includes the substitution of 3 residues within the variable xanthine-stabilising motif from Ser382, Lys384 and Ala385 in the human GMP synthase to Lys360, Pro362 and Ser633 in *A. fumigatus* GMP synthase, respectively (Figure 6C). There is also substitution of Pro612 to Tyr463 and Thr690 to Ile537 from human to *A. fumigatus* GMP synthase, respectively. Notably, Thr690 in human GMP synthase interacts with the phosphate moiety of XMP and therefore, substitution to isoleucine, a residue that is bulkier and more hydrophobic, may affect XMP binding kinetics. Sequence alignment of the ATP-binding site shows greater variability between *A. fumigatus* and *H. sapiens* GMP synthases (Figure 6D). Although some differences are biochemically conserved (Leu235 in *A. fumigatus* is equivalent to Val241 in *H. sapiens*), there are significant changes in the binding pocket (Leu263 and Asn265 in *A. fumigatus* is equivalent to His269 and Asp271 in *H. sapiens*) (Figure 6D).

In vitro inhibition of GMP synthase

As GMP synthase has two catalytic domains, inhibition of enzyme activity can be achieved by targeting either the GATase or ATP-PPase domains to effectively decouple the reaction. The glutamine analogues, ACI and DON, are known inhibitors of GMP synthase that function by targeting the GATase domain (10). Both inhibitors appeared to have inhibitory activity against *A. fumigatus* GMP synthase whereby ACI appeared to be more potent. A half maximal inhibitory concentration (IC_{50}) of 16.6 ± 2.4 µM was calculated for ACI, almost two-fold of

249 that of DON (IC_{50} = 29.6 ± 5.6 µM). Representative concentration-response curves for ACI

250 and DON are shown in Figure 7.

Discussion

The purine biosynthesis pathway is an integral metabolic process that produces purine nucleotide bases that, together with pyrimidine nucleotide bases, form the components of DNA and RNA. In addition to their essential contribution to life, the synthesis of guanosine nucleosides in particular has been linked to virulence in clinically-prominent pathogenic fungi including *A. fumigatus, C. albicans* and *C. neoformans* (10, 11). Although validated and endorsed as an attractive target for new antifungal drugs, the process of drug discovery has been hindered by a lack of structural data. In this study, we have determined the crystal structure of GMP synthase from *A. fumigatus*, the first homologue from a fungal species, and analysed its biophysical properties as part of the beginning of a novel antifungal drug discovery project.

Due to the importance of purines in DNA and RNA synthesis, the *de novo* biosynthesis pathway has been conserved throughout the eukaryotic domain of life. As a result, equivalent biosynthesis enzymes exist within *A. fumigatus* and humans. In order to develop an effective antimycotic agent that targets fungal GMP synthase, species specific differences must be elucidated to achieve selectivity. Analyses of the kinetic properties as well as comparisons between the existing crystal structure of human GMP synthase (2.5 Å) and the *A. fumigatus* GMP synthase structure (2.3 Å) reveal distinct differences in their domain architecture and active site composition which can be exploited to achieve selectivity.

The kinetic properties of *A. fumigatus* GMP synthase differ to those of *E. coli*, *C. neoformans* and *H. sapiens* GMP synthase, which may be reflective of the species-specific differences in substrate binding pocket shape and composition of both catalytic domains. The K_M value measured for ATP and glutamine from *A. fumigatus* GMP synthase is almost two-fold and seven-fold that of the human homologue, respectively. The most notable difference between

eukaryotic homologues of GMP synthase lies in the difference in binding of Mg^{2+}, an essential co-factor required for catalysis. Although the measured $K_M/K_{0.5}$ values of Mg^{2+} were comparable between species, *A. fumigatus* GMP synthase showed no indication of cooperative binding as observed in *C. neoformans* and human homologues.

The crystal structure of *A. fumigatus* GMP synthase more closely resembles that of bacterial and parasitic homologues, in terms of overall fold and domain architecture, despite being part of the eukaryotic domain of life (Figure 4). Conservation of the GATase and ATP-PPase domains allow the enzyme to retain their function in synthesising XMP and the existence of a singular dimerization domain facilitates the formation of a dimeric complex (Figure 1B). In addition to the GATase and ATP-PPase domain, which show high structural conservation to the equivalent domains in *A. fumigatus* GMP synthase, human GMP synthase features two dimerization domains, D1 and D2. The presence of an additional dimerization domain appears to prevent the formation of a dimer as human GMP synthase is predominantly monomeric with only a small dimeric population observed in the presence of substrates (28). As there are fundamental differences in the domain architecture and quaternary arrangement between *A. fumigatus* and *H. sapiens* GMP synthase, an alternative approach to targeting the fungal homologue that does not involve modulating enzyme activity may be possible. It has been suggested that the dimeric form of *P. falciparum* GMP synthase is essential for enzymatic activity (23). If this is also applicable to other dimeric homologues of GMP synthase, including *A. fumigatus*, it may be possible to target the dimer interface with small molecules or peptidomimetics.

The conformational changes undergone by GMP synthase during substrate binding and catalysis were observed using IM-MS which serves to further our understanding of its catalytic mechanism. It is known that the glutamine hydrolysis reaction occurs in the GATase domain, producing ammonia that is channelled to the ATP-PPase domain where it is used to produce

GMP. There is a substantial distance between the respective active sites of the GATase and ATP-PPase domains (approximate distance of 34 Å in the *A. fumigatus* structure) and so it has been suggested that the mechanism of ammonia tunnelling must involve large domain rearrangements. Evidence that supports this mechanism has been shown in studies of the *Plasmodium falciparum* GMP synthase homologue using X-ray crystallography, site-directed mutagenesis and molecular dynamics. These data have indicated that an 85° rotation of the GATase domain is needed to facilitate ammonia transfer (23). Using native IM-MS, we have also observed a dramatic conformational change coinciding with the addition of glutamine in *A. fumigatus* GMP synthase (Figure 2B-D) whereby dimers in the presence of glutamine exhibit an alteration in conformation, where a pronounced lower charge state distribution being observed denoting shielding of surface basic residues. In addition, glutamine induces GMP synthase dimers to adopt an increased number of conformations, as observed by the significant differences in ATD FWHM. Together, we provide further experimental evidence that the conversion of glutamine to glutamate, catalysed by the GATase domain, induces a large conformational change in GMP synthase that is necessary to tunnel ammonia to the ATP-PPase domain for use in converting XMP to GMP.

There are several approaches that can be undertaken to disrupt the catalytic activity of *A. fumigatus* GMP synthase. As discussed previously, targeting the protein-protein interface that forms during dimerization may be an effective method to exploit the differences in quaternary structure between the fungal and human homologues. Alternatively, functional binding pockets in the GATase and ATP-PPase catalytic domains can be explored to find small molecules that modulate enzyme activity. The class I GATase domain is commonly found as part of several biosynthetic enzymes required in purine and pyrimidine synthesis and features a characteristic and highly conserved Cys-His-Glu catalytic triad (29). Known inhibitors of the GATase domain include ACI and DON, both of which are glutamine analogues that have been shown

to have activity against *A. fumigatus* GMP synthase. As this domain family is found in several other enzymes, selectively targeting the GATase domain of *A. fumigatus* GMP synthase may prove to be difficult to achieve selectivity. Instead, targeting either the XMP- and ATP-binding sites, individually or simultaneously, of the ATP-PPase domain may be a more favourable approach. Closer inspection of the XMP-binding site reveals high conservation of motifs that are required for stabilisation of the xanthine component (Pro-Phe-Pro-Gly-Pro-Gly) and residues that interact with the phosphate moiety (Lys532, Ile537, Glu538) based on the XMP-bound *H. sapiens* GMP synthase structure (PDB ID: 2VXO) (Figure 6C). However, there also exists several divergent residues that can be exploited to selectively target the fungal active site over the human equivalent. As a result of the substitutions to the Ser-Gly-Lys-Ala motif, also involved in xanthine stabilisation, in *H. sapiens* GMP synthase to Lys-Gly-Pro-Ser in *A. fumigatus* GMP synthase, different chemical and steric environments are produced (Figure 6C).

In terms of the ATP-binding site, the residues that comprise this site can be inferred from the existing crystal structure of AMP-bound GMP synthase from *E. coli* (PDB ID: 1GPM) (27). Sequence alignments and structural superimpositions of the *E. coli* (PDB ID: 1GPM), *H. sapiens* (PDB ID: 2VXO) and *A. fumigatus* (PDB ID: 7MO6) GMP synthase structures reveal significantly more variability in the ATP-binding pocket in comparison to the XMP-binding pocket (Figure 6D). These divergences in sequence implies that the ATP-binding site may be the preferable site for the design of an inhibitor that modulates GMP synthase activity. Most notably, Leu263 and Asn265 in *A. fumigatus* GMP synthase are equivalent to His269 and Asp271 in the *H. sapiens* homologue, respectively (Figure 6D). Replacement of leucine, a small and hydrophobic residue, to histidine, an aromatic and charged residue, changes the chemical environment of the pocket. This is reflected in the almost two-fold increase in apparent $K_{M\ (ATP)}$ value measured for *A. fumigatus* GMP synthase compared to that of *H. sapiens* (Table 1). These subtle differences in binding pocket composition and measured kinetic

constants between *A. fumigatus* and *H. sapiens* GMP synthase shows high potential for the rational design of small molecule inhibitors that are selective for the fungal homologue.

Conclusions

In this study, we have performed rigorous structural and kinetic characterisation of *A. fumigatus* GMP synthase. These data are crucial to establish a strong foundation for future antifungal drug discovery projects that target GMP synthase. In particular, comparisons between the *A. fumigatus* and *H. sapiens* GMP synthases have revealed distinct divergences in domain architecture, quaternary structure and dynamics, active site composition and kinetic properties. These fundamental differences can be exploited in drug design to produce an inhibitor that is both potent and selective for the fungal homologue over the human equivalent. From this work, we therefore recommend that future drug design projects that target the fungal GMP synthase focus on the ATP-PPase domain, specifically the XMP- and ATP-binding sites.

Materials and Methods

Construction of GMPS-pDest17 Expression Vector and Protein Expression

Pelleted *A. fumigatus* Af293 was submerged in liquid nitrogen, ground into a fine powder and total RNA was extracted using TRIzole solution (Invitrogen) (30). First strand cDNA was synthesised using the SuperScript® III First-Strand Synthesis System (Invitrogen). The guanosine monophosphate synthase ORF (Afu3G01110, *A. fumigatus* Af293) was PCR amplified from *A. fumigatus* Af293 cDNA, inserted into pDONR221 (Invitrogen) and sub-cloned into pDEST17 (Invitrogen) using Gateway cloning (31). The GMP synthase ORF was attached to an N-terminal hexa-histidine tag, separated by an attR1 site and a TEV protease recognition site (ENLYFQG), derived from the pDEST17 vector. For protein expression, GMPS-pDEST17 was transformed into *E. coli* BL21 (λDE3) cells and cultured in Luria broth supplemented with ampicillin (0.2 mg/mL) at 37 °C until an optical density of 0.6 was reached. Protein expression was induced with 0.5 mM IPTG at 16 °C for 16 h. The cells were harvested via centrifugation and resuspended in Buffer A (20 mM Tris pH 8.0, 500 mM NaCl, 10 mM imidazole, 2 mM β-mercaptoethanol) and stored at -80 °C.

Protein Purification

Cells were thawed and then lysed using a M110L microfluidizer processer (Microfluidics). Lysate was clarified via centrifugation before loading onto a 5 mL ZetaSep Nickel NTA column (EMP Biotech), pre-equilibrated with Buffer A. The column was washed with 6 CV of 10% Buffer B (20 mM Tris pH 8.0, 500 mM NaCl, 250 mM imidazole, 2 mM β-mercaptoethanol) and bound GMP synthase was eluted using an imidazole gradient from 10 mM to 250 mM. Fractions containing GMP synthase were pooled and supplemented with EDTA (1 mM) and DTT (0.5 mM). To remove the hexa-histidine tag, TEV protease was added in a 1:1 mass ratio and incubated overnight at 4 °C whilst simultaneously dialysing against

Buffer A. The cleaved material was loaded onto a 5 mL ZetaSep Nickel NTA column (EMP Biotech) and the flow-through containing untagged GMP synthase was collected. The purity of GMP synthase was analysed via SDS-PAGE. Pure fractions were pooled and dialysed against Storage Buffer at 4 °C for 16 h (50 mM Tris pH 8.0, 0.5 mM EDTA, 5% glycerol, 1 mM DTT), concentrated using an Amicon Ultra-15 Centrifugal Filter Unit (30 kDa MWCO) to 24 mg/mL, flash frozen in liquid nitrogen and stored at -80 °C.

Crystallisation

GMP synthase crystals of poor morphology formed upon incubating GMP synthase (24 mg/mL) in the presence of ATP (8.2 mM) and $MgCl_2$ (4.2 mM), with well solution (25% PEG 3350, 0.15 M Bis-Tris pH 6.5) in a 1:1 ratio at 16 °C using hanging drop vapour diffusion. These crystals were used to make a seed stock for further optimisation using a random microseed matrix screening protocol adapted for grid screening (32). Briefly, GMP synthase crystals were crushed and transferred to 100 µL of well solution (25% PEG 3350, 0.15 M Bis-Tris pH 6.5), then diluted 1/100 with well solution for use as a working stock. To form high quality and diffracting crystals of apo-GMP synthase, 0.1 µL of 1/100 diluted seed stock was added to the 1:1 protein to well solution (21% PEG 3350, 0.05 M Bis-Tris pH 6.5) mixture, then incubated at 16 °C for 14 d.

Data Collection and Processing

A single crystal was mounted on a cryo-loop, cryo-protected in *Paratone-N* (Hampton Research) and flash frozen in liquid nitrogen. Diffraction data was collected at the Australian Synchrotron using the Macromolecular Crystallography MX2 beamline at a wavelength of 0.954 Å (33). Data was integrated using XDS, converted to an *mtz* format using *Pointless* (CCP4i), then scaled and merged using *Aimless* (CCP4i) (34-36). The GMP synthase crystals

had symmetry of the space group $P12_11$. Statistics of data collection and processing are summarised in Table 2.

Table 2: Statistics of data collection, data processing and structure refinement of the *Aspergillus fumigatus* GMP synthase structure from crystallographic data

Data Collection and Processing	***Aspergillus fumigatus* GMP Synthase**
Wavelength (Å)	0.954
Space Group	P 1 2_1 1
Unit Cell Dimensions	
a, b, c (Å)	47.3 159.8 76.2
α, β, γ (°)	90 108 90
Resolution (Å)	39.19 - 2.3 (2.382 - 2.3)
Unique Reflections	47516 (4762)
Multiplicity	6.6 (5.9)
Completeness (%)	98.9 (96.0)
R_{merge}[a]	0.21 (2.6)
R_{pim}[b]	0.09 (1.2)
Average I/σ (I)	6.1 (0.9)
CC1/2[c]	0.99 (0.34)
Structure Refinement	
No. Atoms	8209
Protein	7844
Water	365
R_{free}[d]	0.28 (0.46)
R_{work}[e]	0.23 (0.39)
RMS Deviations	
Bonds (Å)	0.002
Angles (°)	0.42
Ramachandran Analysis	
Outliers	0.3
Favoured	96.7

412 Statistics for the highest-resolution shell are shown in parentheses.

413 [a] $R_{merge} = \sum_{hkl} \sum_i | I_i(hkl) - \langle I(hkl) \rangle | / \sum_{hkl} \sum_i I_i(hkl)$ (37)

414 [b] $R_{pim} = \sum_{hkl} \{1/[N(hkl) - 1]\}^{1/2} \times \sum_i |I_i(hkl) - \langle I(hkl) \rangle| / \sum_{hkl} \sum_i I_i(hkl)$ (38)

415 [c] $CC1/2 = \sum (x - \langle x \rangle)(y - \langle y \rangle) / [\sum (x - \langle x \rangle)^2 (y - \langle y \rangle)^2]^{1/2}$ (39)

416 [d] $R_{free} = \sum |F_o - F_c| / \sum |F_o|$ for all data with $F_o > 2\sigma (F_o)$, calculated from 5% of reflections,
417 randomly chosen (40)

418 [e] $R_{work} = \sum |F_o - F_c| / \sum |F_o|$ for all data with $F_o > 2\sigma (F_o)$, excluding data to calculate R_{free}

419

Structure Refinement

The structure of GMP synthase was solved by molecular replacement using *Phaser MR* (CCP4i) with GMP synthase from *Thermus thermophilus* (2YWC) as a search model (41). Iterative cycles of manual rebuilding were performed with *Coot* followed by refinement using *Phenix.refine* (42, 43). Model validation was assessed using *MolProbity* (44). Protein structure visualisation and figure preparation were performed in *Pymol* version 2.3.4 (45). Analysis of protein interfaces was completed using *PDBe PISA* version 1.52 (46). Statistics of structure refinement are summarised in Table 2.

Analytical Size-Exclusion Chromatography (SEC)

The average oligomeric size of GMP synthase (100 µM) was determined by analytical SEC using a Superdex 200 10/300 GL analytical-SEC (GE Healthcare) equilibrated in either 50 mM phosphate-buffered saline (PBS, pH 7.4) or ammonium acetate (NH_4OAc, pH 6.8) at 0.4 mL min^{-1} at room temperature using methods described previously (15). The size-exclusion column was calibrated using molecular weight standards (Sigma-Aldrich).

Native & Ion-Mobility Mass Spectrometry

The quaternary structure of GMP synthase was examined using a Synapt G1 HDMS (Waters) using a nanoelectrospray ionisation source (15). GMP synthase was buffer exchanged into 200 mM ammonium acetate (NH_4OAc) (pH 6.8) using an Amicon Ultra-Centrifugal Filter (10,000 MWCO) at 4 °C. Protein concentration was adjusted to 15 µM and 2 µL of protein was loaded into gold-coated borosilicate glass capillaries prepared in-house. Key instrument parameters were as follows: capillary voltage (kV): 1.60; sampling cone (V): 50; extraction cone (V): 1.5; trap/transfer collision energy (V/V): 20/15; trap gas (L/hr): 5.5; backing gas (mbar): ~4.5. The conformation of GMP synthase (25 µM) in the presence of glutamine (0.5 mM) was examined by ion-mobility MS (IM-MS) (15). Arrival time distributions (ATD) of the GMP synthase

dimer^{21+} charge state were selected to determine changes in conformational states and minimize charge state overlap from other possible oligomeric species. IM-MS parameters included: IM cell wave height, 9 V; IM cell wave velocity, 350 m/s; transfer t-wave height, 9 V; transfer t-wave velocity, 250 m/s. All mass spectra and arrival time distributions (ATDs) were analysed using MassLynx (v4.1) and DriftScope (v2.1) (both Waters). Full width- half maximum (FWHM) of the dimer^{21+} ATD was analysed by unpaired t-test using Prism 8.0 (GraphPad).

GMP Synthase Spectrophotometric Activity Assay

Activity assays for *A. fumigatus* GMP synthase were performed in half area 96-well UV-transparent microplates using methods adapted from the cuvette-based assay described previously for *Cryptococcus neoformans* GMP synthase (11). Reaction mixtures for all assays contained 50 mM HEPES pH 7.5 and saturating conditions of all substrates (0.5 mM XMP, 20 mM glutamine, 2.5 mM ATP and 20 mM $MgCl_2$). Reactions were performed at 40 °C and were started by adding GMP synthase to a final concentration of 0.0125 mg/mL. The conversion from XMP (ε = 4.080 mM^{-1} cm^{-1}) to GMP (ε = 3.066 mM^{-1} cm^{-1}) was monitored using a PHERAstar FSX microplate reader (BMG LabTech) at 290 nm and the difference in molar absorptivity ($\Delta\varepsilon_{290}$ = 1.014 mM^{-1} cm^{-1}) was used to calculate the amount of GMP formed. Assays to determine steady-state kinetic parameters were conducted using saturated concentrations of three of the substrates whilst varying the concentration of the fourth substrate. Data from three biological replicates, plated in technical triplicate, were fitted using the Michaelis-Menten equation, defined as:

$$Y = (V_{max} \times x)/(\mathrm{K_M} + x)$$

Where V_{max} = maximum enzyme velocity, $\mathrm{K_M}$ = Michaelis-Menten constant, x = substrate concentration and Y = enzyme velocity

to calculate the kinetic parameters (GraphPad Prism 8.0).

GMP Synthase In vitro Inhibition Assays

In vitro inhibition assays were performed using the spectrophotometric activity assay as outlined above. Both DON and ACI (dissolved in water) were pre-incubated with GMP synthase (final concentration of 0.0125 mg/mL) in reaction mix containing 50 mM HEPES pH 7.5, 0.5 mM XMP, 2.5 mM ATP and 20 mM $MgCl_2$ for 5 min at 40 °C. Reactions were started by adding glutamine to a final concentration of 5 mM. Reaction rates were obtained for various concentrations of both inhibitors and plotted as a percentage of activity relative to the reaction containing no inhibitor. Data from three replicates, plated in triplicate, were fit to a dose response curve in GraphPad Prism 8.0 to calculate the inhibitory constant (IC_{50}) of the inhibitor. Concentrations of both inhibitors tested ranged from 0 mM – 125 mM.

Accession Codes

The atomic coordinates and structure factors of GMP synthase from *A. fumigatus* were deposited to the Protein Data Bank with accession code 7MO6.

Acknowledgements

We thank Flinders Analytical (Flinders University, Australia) for access to IM-MS instrumentation.

References

1. Chinen J, Shearer WT. 2010. Secondary immunodeficiencies, including HIV infection. J Allergy Clin Immunol 125:S195-S203.
2. Silva RF. 2010. Chapter 8: Fungal infections in immunocompromised patients. J Bras Pneumol 36:142-147.
3. Kontoyiannis DP, Marr KA, Park BJ, Alexander BD, Anaissie EJ, Walsh TJ, Ito J, Andes DR, Baddley JW, Brown JM, Brumble LM, Freifeld AG, Hadley S, Herwaldt LA, Kauffman CA, Knapp K, Lyon GM, Morrison VA, Papanicolaou G, Patterson TF, Perl TM, Schuster MG, Walker R, Wannemuehler KA, Wingard JR, Chiller TM, Pappas PG. 2010. Prospective surveillance for invasive fungal infections in hematopoietic stem cell transplant recipients, 2001-2006: Overview of the Transplant-Associated Infection Surveillance Network (TRANSNET) Database. Clin Infect Dis 50:1091-1100.
4. Pappas PG, Alexander BD, Andes DR, Hadley S, Kauffman CA, Freifeld AG, Anaissie EJ, Brumble LM, Herwaldt LA, Ito J, Kontoyiannis DP, Lyon GM, Marr KA, Morrison VA, Park BJ, Patterson TF, Perl TM, Oster RA, Schuster MG, Walker R, Walsh TJ, Wannemuehler KA, Chiller TM. 2010. Invasive fungal infections among organ transplant recipients: Results of the Transplant-Associated Infection Surveillance Network (TRANSNET). Clin Infect Dis 50:1101-1111.
5. Latgé JP. 1999. *Aspergillus fumigatus* and aspergillosis. Clin Microbiol Rev 12:310-350.
6. Abad A, Fernández-Molina JM, Bikandi J, Ramírez A, Margareto J, Sendino J, Hernando FL, Pontón J, Garaizar J, Rementeria A. 2010. What makes *Aspergillus fumigatus* a successful pathogen? Genes and molecules involved in invasive aspergillosis. Rev Iberoam Micol 27:155-182.

7. Brakhage AA. 2005. Systemic fungal infections caused by Aspergillus species: Epidemiology, infection process and virulence determinants. Curr Drug Targets 6:875-886.

8. Nguyen S, Truong JQ, Bruning JB. 2021. Targeting unconventional pathways in pursuit of novel antifungals. Front Mol Biosci 7:479.

9. Chitty JL, Fraser JA. 2017. Purine acquisition and synthesis by human fungal pathogens. Microorganisms 5:E33.

10. Rodriguez-Suarez R, Xu D, Veillette K, Davison J, Sillaots S, Kauffman S, Hu W, Bowman J, Martel N, Trosok S, Wang H, Zhang L, Huang L, Li Y, Rahkoodaee F, Ransom T, Gauvin D, Douglas C, Youngman P, Becker J, Jiang B, Roemer T. 2007. Mechanism-of-action determination of GMP synthase inhibitors and target validation in *Candida albicans* and *Aspergillus fumigatus*. Chem Biol 14:1163-1175.

11. Chitty JL, Tatzenko TL, Williams SJ, Koh YQ, Corfield EC, Butrler MS, Roberston AA, Cooper MA, Kappler U, Kobe B, Fraser JA. 2017. GMP synthase is required for virulence factor production and infection by *Cryptococcus neoformans*. J Biol Chem 292:3049-3059.

12. Morrow CA, Valkov E, Stamp A, Chow EWL, Lee IR, Wronski A, Williams SJ, Hill JM, Djordjevic JT, Kappler U, Kobe B, Fraser JA. 2012. *De novo* GTP biosynthesis is citical for virulence of the fungal pathogen *Cryptococcus neoformans*. PLoS Pathog 8:e1002957.

13. Dinamarco TM, Brown NA, Couto de Almeida RS, Alves de Castro P, Savoldi M, de Souza Goldman MH, Goldman GH. 2012. *Aspergillus fumigatus* calcineurin interacts with a nucleoside diphosphate kinase. Microbes Infect 14:922-929.

14. Lin X, Momany C, Momany M. 2003. SwoHp, a nucleoside diphosphate kinase, is essential in *Aspergillus nidulans*. Eukaryot Cell 2:1169-1177.

15. Nguyen S, Jovcevski B, Pukala TL, Bruning JB. 2020. Nucleoside selectivity of Aspergillus fumigatus nucleoside-diphosphate kinase. FEBS J 288:2398-2417.
16. Oliver JC, Linger RS, Chittur SV, Davisson VJ. 2013. Substrate activation and conformational dynamics of guanosine 5'-monophosphate synthetase. Biochemistry 52:5225-5235.
17. Fukuyama TT. 1996. Formation of an adenyl xanthosine monophosphate intermediate by xanthosine 5'-phosphate aminase and its inhibition by psicofuranine. J Biol Chem 241:4745-4749.
18. Jovcevski B, Kelly MA, Aquilina JA, Benesch JLP, Ecroyd H. 2017. Evaluating the effect of phosphorylation on the structure and dynamics of Hsp27 dimers by means of ion mobility mass spectrometry. Anal Chem 89:13275-13282.
19. Sanders HM, Jovcevski B, Carver JA, Pukala TL. 2020. The molecular chaperone β-casein prevents amorphous and fibrillar aggregation of α-lactalbumin by stabilisation of dynamic disorder. Biochem J 477:629-643.
20. Jovcevski B, Das S, Smid S, Pukala TL. 2020. Polyphenol honokiol and flavone 2′, 3′, 4′-trihydroxyflavone differentially interact with α-synuclein at distinct phases of aggregation. ACS Chem Neurosci 11:4469-4477.
21. Abbott JL, Newell JM, Lightcap CM, Olanich ME, Loughlin DT, Weller MA, Lam G, Pollack S, Patton WA. 2006. The effects of removing the GAT domain from *E. coli* GMP synthetase. Protein J 25:483-491.
22. Nakamura J, Lou L. 1995. Biochemical characterization of human GMP synthetase. J Biol Chem 270:7347-7353.
23. Ballut L, Violot S, Shivakumaraswamy S, Thota LP, Sathya M, Kunala J, Dijkstra BW, Terreux R, Haser R, Balaram H, Aghajari N. 2015. Active site coupling in

559 *Plasmodium falciparum* GMP synthetase is triggered by domain rotation. Nat
560 Commun 6.
561 24. Baba S, Kanagawa M, Yanai H, Ishii T, Kuramitsu S, Yokoyama S, Sampei G, Kawai
562 G. 2007. Crystal structure of GMP synthetase from *Thermus thermophilus*.
563 doi:10.2210/pdb2YWB/pdb.
564 25. Edwards TE, Abendroth J, Lorimer, D.,, Fairman JW, (SSGCID) SSGCfID. 2016.
565 Crystal structure of a GMP synthase (glutamine-hydrolyzing) from *Neisseria*
566 *gonorrhoeae*. doi:10.2210/pdb5TW7/pdb.
567 26. Franklin MC, Cheung J, Rudolph MJ, Burshteyn F, Cassidy M, Gary E, Hillerich B,
568 Yao ZK, Carlier PR, Totrov M, Love JD. 2015. Structural genomics for drug design
569 against the pathogen *Coxiella burnetii*. Proteins 83:2124-2136.
570 27. Tesmer JJ, Klem TJ, Deras ML, Davisson VJ, Smith JL. 1996. The crystal structure of
571 GMP synthetase reveals a novel catalytic triad and is a structural paradigm for two
572 enzyme families. Nat Struct Biol 3:74-86.
573 28. Welin M, Lehtio L, Johansson A, Flodin S, Nyman T, Tresaugues L, Hammarstrom
574 M, Graslund S, Nordlund P. 2013. Substrate specificty and oligomerization of human
575 GMP synthetase. J Mol Biol 425:4323.
576 29. van den Heuvel HH, Curti B, Vanoni MA, Mattevi A. 2004. Glutamate synthase: a
577 fascinating pathwya from L-glutamine to L-glutamate. Cell Mol Life Sci 61:669-681.
578 30. Bertuzzi M, van Rhijn N, Krappmann S, Bowyer P, Bromley MJ, Bignell EM. 2020.
579 On the lineage of *Aspergillus fumigatus* isolates in common laboratory use. Med
580 Mycol 59:7-13.
581 31. Walhout AJ, Temple GF, Brasch MA, Hartley JL, Lorson MA, van den Heuvel S,
582 Vidal M. 2000. GATEWAY recombinational cloning: application to the cloning of
583 large numbers of open reading frames or ORFeomes. Methods Enzymol 328:575-592.

584 32. Truong JQ, Nguyen S, Bruning JB, Shearwin KE. 2021. Derivatization of protein
585 crystals with I3C using random microseed matrix screening. JoVE:e61894.
586 33. Aragao D, Aishima J, Cherukuvada H, Clarken R, Clift M, Cowieson NP, Ericsson
587 DJ, Gee C, Macedo S, Mudie N, Panjikar S, Price JR, Riboldi-Tunnicliffe A, Rostan
588 R, Williamson R, Caradoc-Davies T. 2018. MX2: A high-flux undulator microfocus
589 beamline serving both the chemical and macromolecular crystallography communities
590 at the Australian Synchrotron. J Synchrotron Radiat 25:885-891.
591 34. Kabsch W. 2010. XDS. Acta Crystallogr D Biol Crystallogr 66 (Pt 2):125-132.
592 35. Evans PR. 2006. Scaling and assessment of data quality. Acta Crystallogr D Biol
593 Crystallogr 62:72-82.
594 36. Evans PR, Murshudov GN. 2013. How good are my data and what is the resolution?
595 Acta Crystallogr D Biol Crystallogr 69:1204-1214.
596 37. Arndt UW, Crowther RA, Mallet JFW. 1968. A computer-linked cathode ray tube
597 microdensitometer for X-ray crystallography. J Phys E: Sci Instr 1:510-516.
598 38. Weiss MS, Hilgenfeld R. 1997. On the use of merging *R*-factor as a quality indicator
599 for X-ray data. J Appl Crystallogr 30:203-205.
600 39. Karplus PA, Diederichs K. 2012. Linking crystallographic model and data quality.
601 Science 336:1030-1033.
602 40. Brünger AT. 1992. Free R value: a novel statistical quantity for assessing the
603 accuracy of crystal structures. Nature 355:472-475.
604 41. Winn MD, Ballard CC, Cowtan KD, Dodson EJ, Emsley P, Evans PR, Keegan RM,
605 Krissinel EB, Leslie AG, McCoy A, McNicholas SJ, Murshudov GN, Pannu NS,
606 Potterton EA, Powell HR, Read RJ, Vagin A, Wilson KS. 2011. Overview of the
607 CCP4 suite and current developments. Acta Crystallogr D Biol Crystallogr 67:235-
608 242.

42. Emsley P, Cowtan K. 2004. Coot: Model-building tools for molecular graphics. Acta Crystallogr D Biol Crystallogr 60:2126-2132.

43. Afonine PV, Grosse-Kunstleve RW, Echols N, Headd JJ, Moriarty NW, Mustyakimov M, Terwilliger TC, Urzhumtsev A, Zwart PH, Adams PD. 2012. Towards automated crystallographic structure refinement with phenix.refine. Acta Crystallogr D Biol Crystallogr 68:352-367.

44. Chen VB, Arendall WB, Headd JJ, Keedy DA, Immormino RM, Kapral GJ, Murray LW, Richardson JS, Richardson DC. 2010. *MolProbity*: All-atom structure validation for macromolecular crystallography. Acta Crystallogr D Biol Crystallogr 66:12-21.

45. Anonymous. The PyMOL Molecular Graphics System, Version 2.0 Schrödinger, LLC.

46. Krissinel E, Henrick K. 2007. Inference of macromolecular assemblies from crystalline state. J Mol Biol 372:774-797.

47. Notredame C, Higgins DG, Heringa J. 2000. T-Coffee: a novel method for fast and accurate multiple sequence alignment. J Mol Biol 302:205-217.

48. Robert X, Gouet P. 2014. Deciphering key features in protein structures with the new ENDscript server. Nucl Acids Res 42:W320-W324.

Figures and Tables

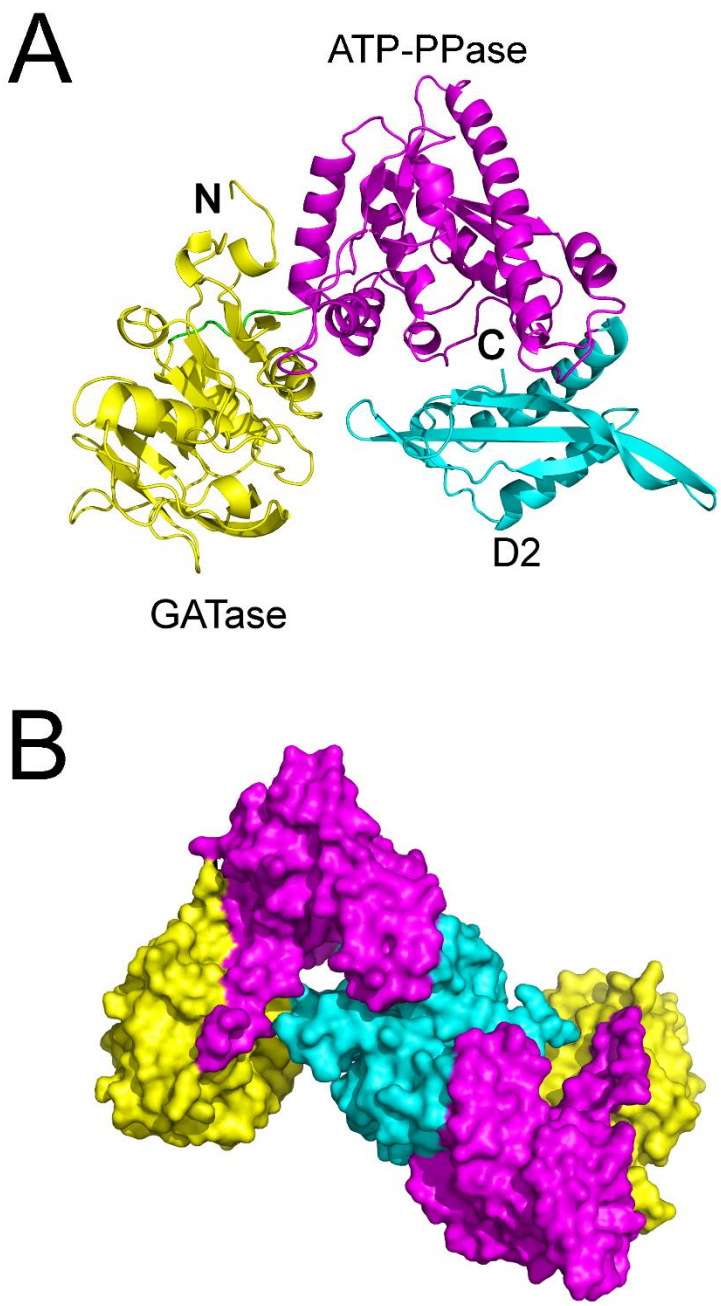
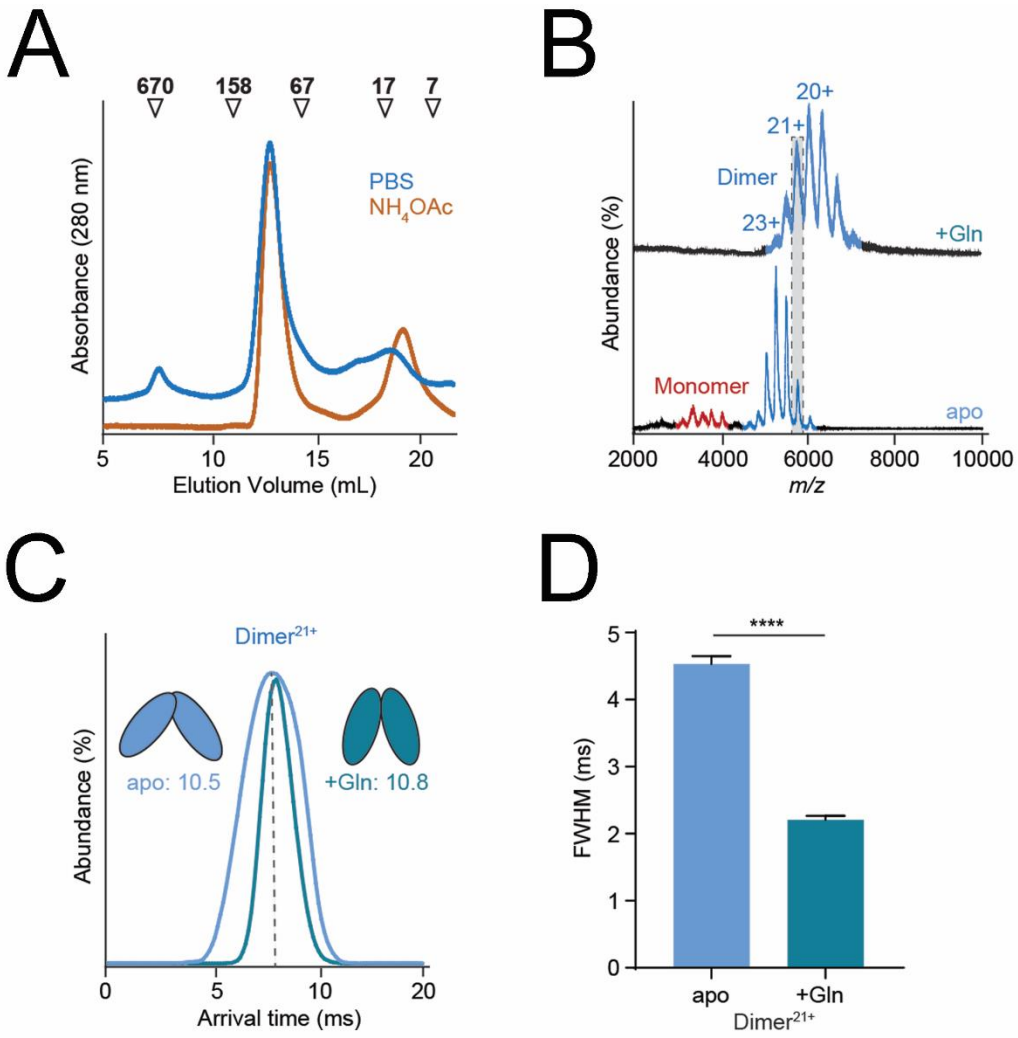
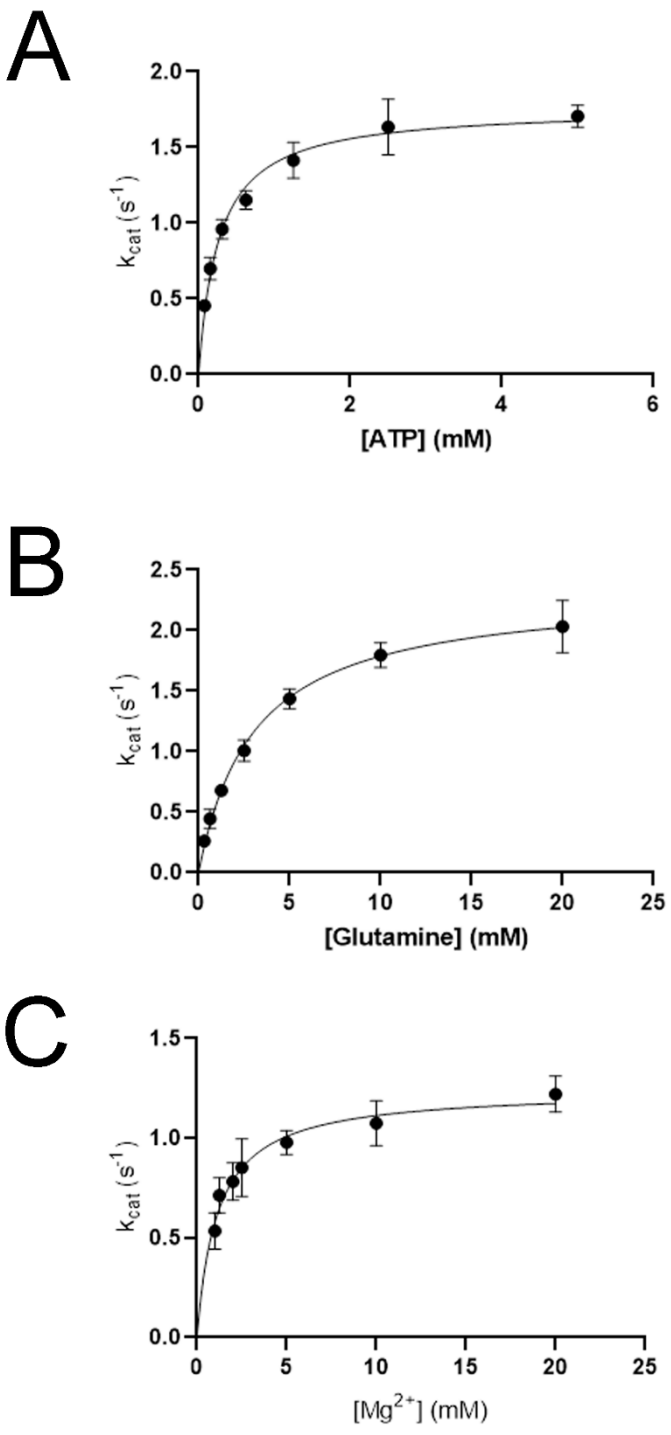


Figure 1



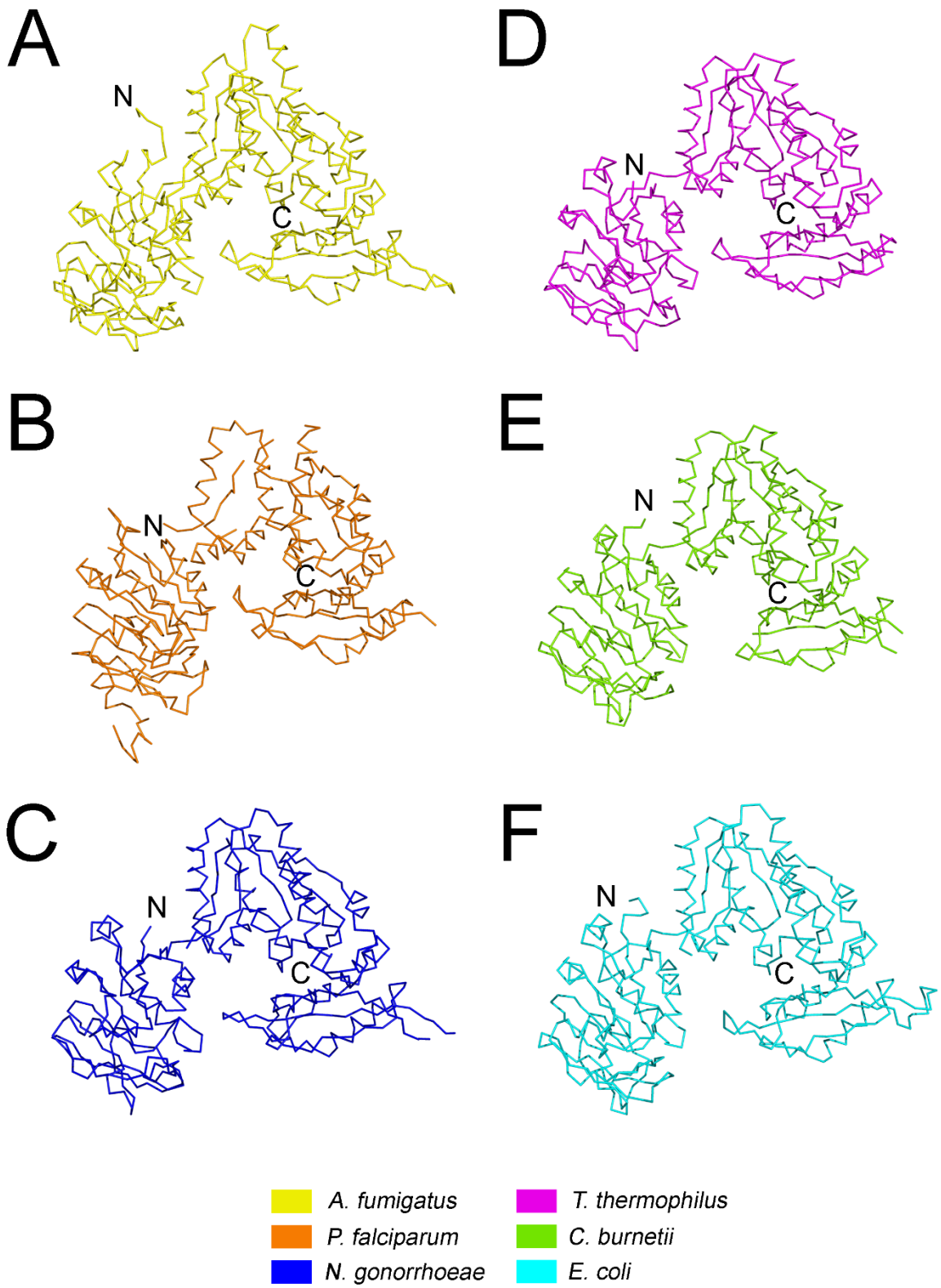
653 **Figure 2**

654



655 **Figure 3**

656


A
N
C
D
N
C
B
N
C
E
N
C
C
N
C
F
N
C
A. fumigatus
P. falciparum
N. gonorrhoeae
T. thermophilus
C. burnetii
E. coli

657 **Figure 4**

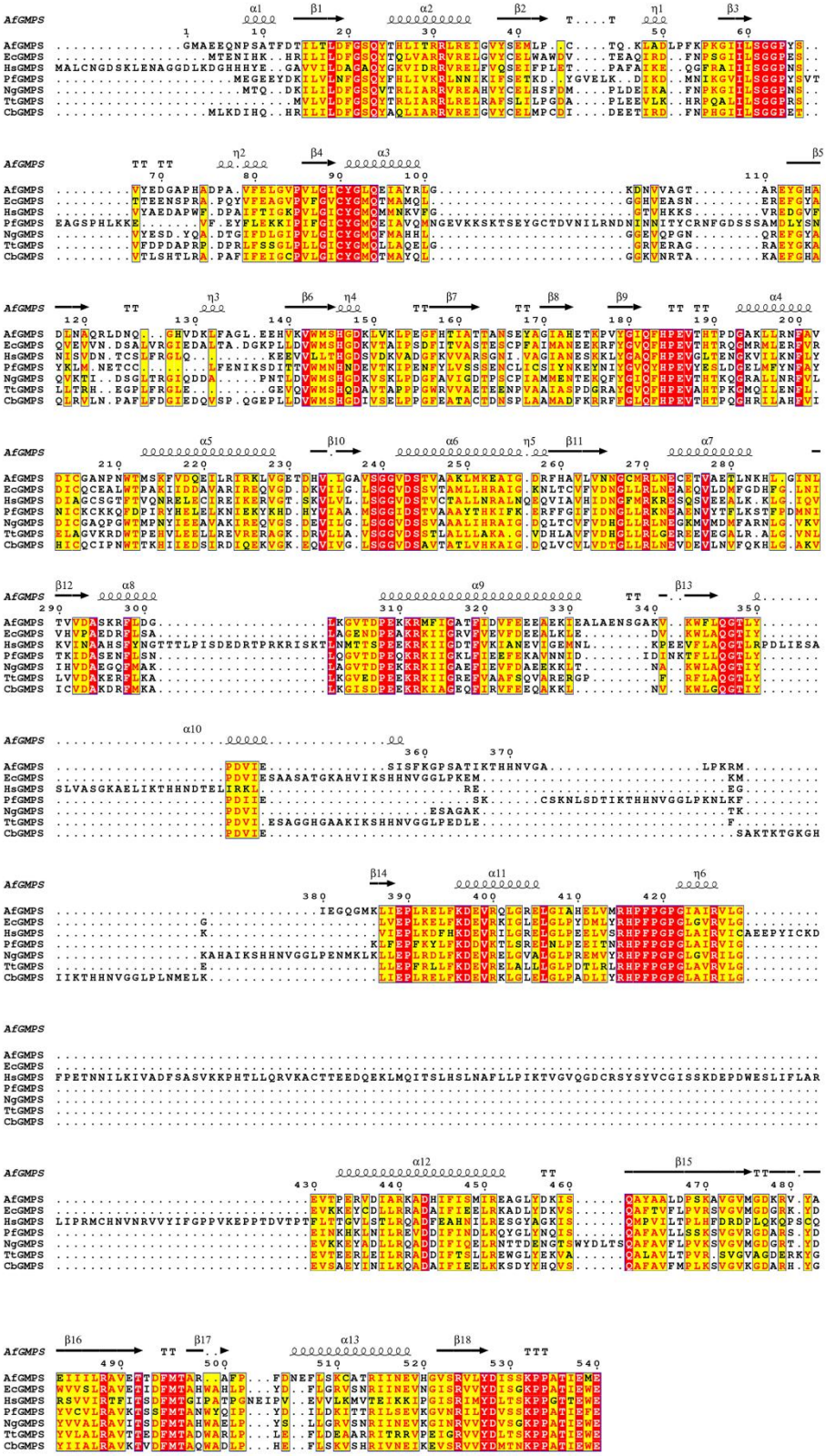
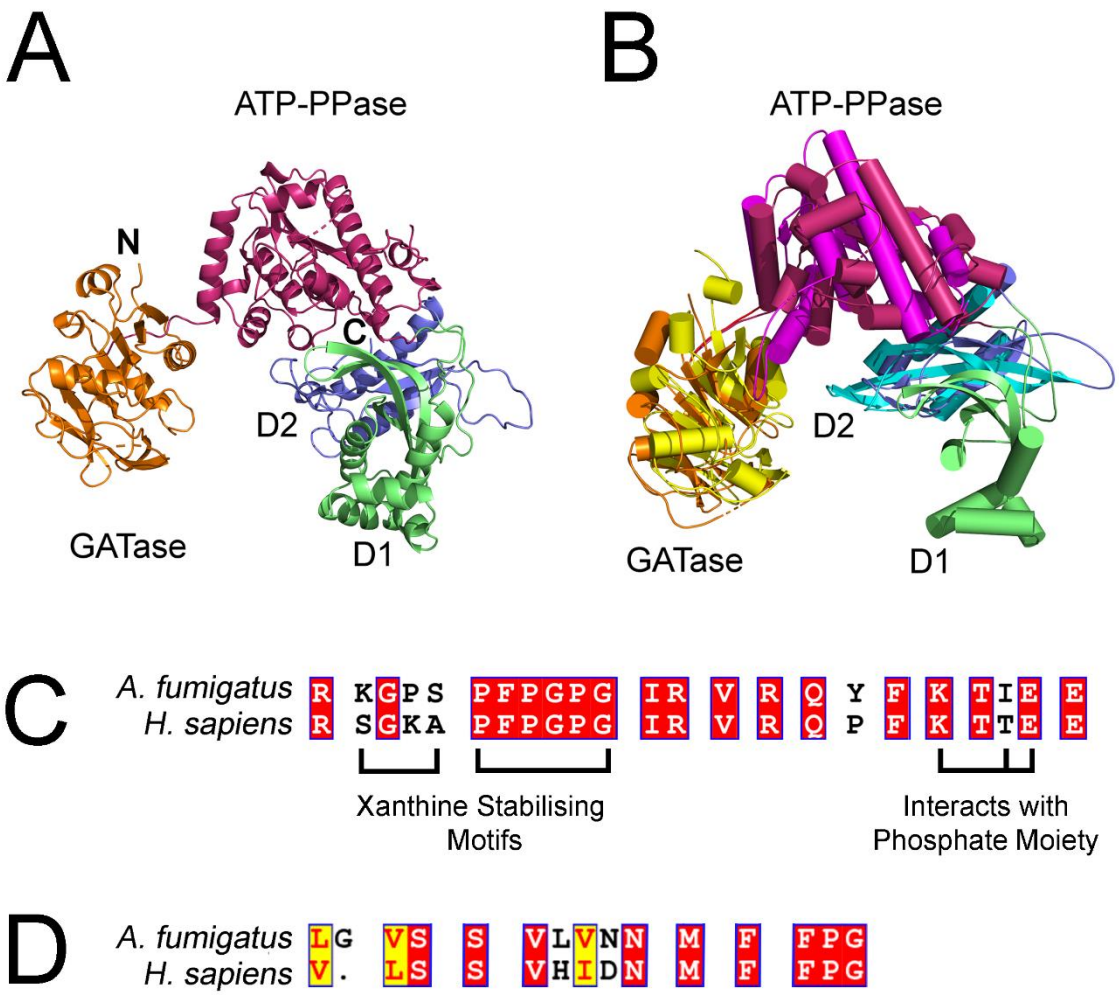


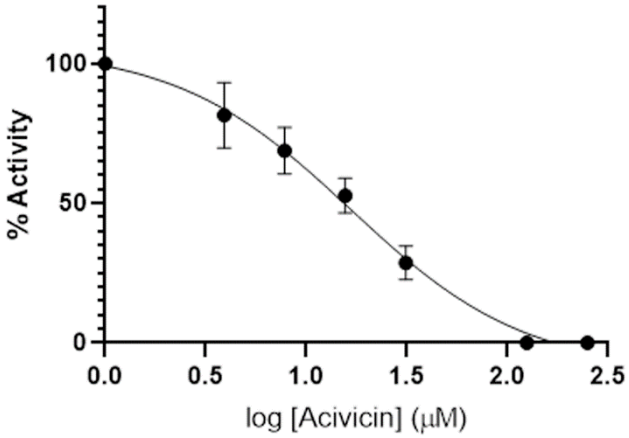
Figure 5



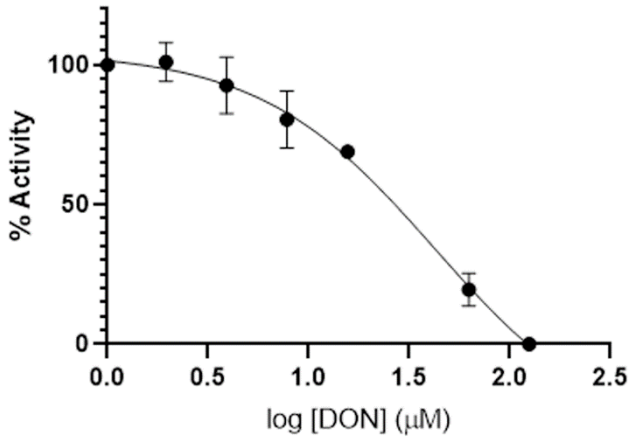
660 **Figure 6**

661

A



B



662

663 **Figure 7**

664

Figure Captions

Figure 1: Crystal structure of GMP synthase from *Aspergillus fumigatus* (2.3 Å). **(A)** The N-terminal glutamine amidotransferase (GATase) domain, coloured in yellow (residues 1 - 205) and the C-terminal ATP pyrophosphatase (ATP-PPase) domain, coloured in magenta (residues 213 - 431), are connected by a short loop region, coloured in green (residues 206 - 212). The C-terminal D2 dimerization domain is coloured in cyan (residues 432 – 540). **(B)** GMP synthase from *Aspergillus fumigatus* forms a dimeric complex via the D2 dimerization domain (cyan).

Figure 2: GMP synthase from *Aspergillus fumigatus* adopts a dimeric quaternary structure and adopts a 'closed' conformation upon glutamine (Gln) addition. **(A)** The analytical size-exclusion chromatograph shows the elution of GMP synthase as a predominant peak, corresponding to a dimeric species in phosphate-buffer saline (blue, PBS) and ammonium acetate (orange, NH_4OAc). **(B)** Native MS also shows GMP synthase (15 µM) is dimeric (blue) in both the absence (apo) and presence of Gln (0.5 mM) and the dimer^{21+} (~5,730 *m/z*) was selected for analysis (grey box). **(C)** ATD analysis of GMP synthase dimer^{21+} ions in the absence (apo, blue) and presence (+Gln, green) of Gln exhibiting similar ATD maxima (apo GMP synthase maxima indicated by dashed line). **(D)** FWHM for the ATD of GMP synthase dimer^{21+} ions in the absence and presence of Gln. Data presented in C and D are reported as mean ± SD (n = 3) (**** $p<0.001$).

Figure 3: Representative binding curves for **(A)** ATP, **(B)** glutamine and **(C)** Mg^{2+}. The data points represent the mean of triplicate samples and the error bars represent the standard deviation.

Figure 4: GMP synthase homologues show high conservation of the overall fold, domain architecture and domain arrangement. Homologous GMP synthase structures are displayed in

ribbon representation and colour coded as follows: **(A)** *Aspergillus fumigatus* (yellow), **(B)** *Plasmodium falciparum* (orange), **(C)** *Neisseria gonorrhoeae* (blue), **(D)** *Thermus thermophilus* (magenta), **(E)** *Coxiella burnetii* (green) and *Escherichia coli* (blue).

Figure 5: Protein sequence alignment of GMP synthase homologues from *A. fumigatus, H. sapiens, E. coli, P. falciparum, N. gonnorrhoeae, T. thermophilus* and *C. burnetti*. Strictly conserved residues are represented as white text on a red background, well-conserved residues are represented as red text on a yellow background and non-conserved residues are represented in black text. The secondary structure of *A. fumigatus* GMP synthase is represented above the protein sequence alignment where α, β, η and TT are used to represent an α-helix, β-strand, 3_{10} helix and a strict β-turn, respectively. The sequence alignment was carried out using T-Coffee and visualised using ESPript 3.0 (47, 48).

Figure 6: The domain architecture of human GMP synthase (PDB: 2VXO) is distinctly different from that of *Aspergillus fumigatus* GMP synthase. **(A)** The crystal structure of human GMP synthase features a GATase (orange), ATP-PPase (maroon) domain and D2 dimerization (purple) sub-domain which are conserved between species but the D1 dimerization (green) sub-domain is unique to human GMP synthase (28). **(B)** Superimposition of GMP synthase structures from *A. fumigatus* (GATase; yellow, ATP-PPase; magenta, D2: cyan) and human (colour coding retained) depicts the lack of a D1 dimerization sub-domain in the fungal equivalent. Alignment of the **(C)** XMP-binding site residues and **(D)** ATP-binding site residues of the GMP synthase ATP-PPase domain from *Aspergillus fumigatus* and *Homo sapiens* reveals key differences in composition.

Figure 7: Representative concentration-response curves for **(A)** acivicin and **(B)** 6-diazo-5-oxo-L-norleucine (DON). The data points represent the mean of triplicate samples and the error bars represent the standard deviation.

Chapter 4 (Manuscript):

Targeting the mannitol biosynthesis pathway in Aspergillus fumigatus: Characterisation and inhibition of mannitol-2-dehydrogenase

Chapter 4 Preface

Chapter 4 is an unpublished manuscript that describes the characterisation of the structure and function of mannitol-2-dehydrogenase from *A. fumigatus* using X-ray crystallography, native mass spectrometry and *in vitro* enzyme assays. As a proof of concept, we have demonstrated the inhibitory activity of the small molecule inhibitor, 1,4-benzoquinone, against this enzyme. I performed all X-ray crystallography experiments and *in vitro* enzyme assays. I performed the data analysis for these experiments, prepared the figures included in the manuscript, deposited structure factors and coordinate factors into the Protein Data Bank and prepared the manuscript.

Statement of Authorship

Title of Paper	Targeting the mannitol biosynthesis pathway in *Aspergillus fumigatus*: Characterisation and inhibition of mannitol-2-dehydrogenase
Publication Status	☐ Published ☐ Accepted for Publication ☐ Submitted for Publication ☑ Unpublished and Unsubmitted work written in manuscript style
Publication Details	This publication characterises the structural and kinetic properties of mannitol-2-dehydrogenase from *A. fumigatus* and demonstrates the inhibitory activity of the small molecule inhibitor, 1,4-benzoquinone, against this enzyme.

Principal Author

Name of Principal Author (Candidate)	Stephanie Nguyen		
Contribution to the Paper	Project conceptualisation, planned and performed experiments, analysed data, prepared, reviewed and edited the manuscript		
Overall percentage (%)	60		
Certification:	This paper reports on original research I conducted during the period of my Higher Degree by Research candidature and is not subject to any obligations or contractual agreements with a third party that would constrain its inclusion in this thesis. I am the primary author of this paper.		
Signature		Date	17/07/21

Co-Author Contributions

By signing the Statement of Authorship, each author certifies that:

i. the candidate's stated contribution to the publication is accurate (as detailed above);

ii. permission is granted for the candidate in include the publication in the thesis; and

iii. the sum of all co-author contributions is equal to 100% less the candidate's stated contribution.

Name of Co-Author	Dr. Blagojce Jovcevski		
Contribution to the Paper	Planned and performed experiments, analysed data, reviewed and edited the manuscript		
Overall Percentage (%)	10		
Signature		Date	20/07/21

Name of Co-Author	Chia-De Ruth Wang		
Contribution to the Paper	Planned and performed experiments, analysed data, reviewed and edited the manuscript		
Overall Percentage (%)	10		
Signature		Date	19/07/21

Name of Co-Author	Assoc. Prof. Tara Pukala		
Contribution to the Paper	Planned experiments, reviewed and edited the manuscript		
Overall Percentage (%)	10		
Signature		Date	18/7/21

Name of Co-Author	Dr. John B. Bruning		
Contribution to the Paper	Project conceptualisation, planned experiments, reviewed and edited the manuscript		
Overall Percentage (%)	10		
Signature		Date	21/7/21

Targeting the Mannitol Biosynthesis Pathway in *Aspergillus fumigatus*: Characterisation and Inhibition of Mannitol-2-Dehydrogenase

Stephanie Nguyen[1], Blagojce Jovcevski[2,3], C. Ruth Wang[2], Tara L. Pukala[2], John B. Bruning[1]

[1] Institute of Photonics and Advanced Sensing (IPAS), School of Biological Sciences, The University of Adelaide, South Australia, 5005, Australia

[2] School of Physical Sciences, Department of Chemistry, The University of Adelaide, South Australia, 5005, Australia

[3] School of Agriculture, Food and Wine, The University of Adelaide, South Australia, 5005, Australia

Keywords: Mannitol biosynthesis, *Aspergillus fumigatus*, aspergillosis, mannitol-2-dehydrogenase

Abstract

Infections caused by the opportunistic fungal pathogen *Aspergillus fumigatus* pose a serious burden on the public health system. The inherent limitations in existing antifungal drugs in conjunction with a rising emergence of antifungal resistance emphasizes an urgent need to identify and target alternative pathways that are crucial to survival and virulence. Targeting the fungal mannitol biosynthesis enzymes provides a promising avenue in the development of new antifungals due to the multifaceted roles mannitol fulfils in the fungal life cycle. However, a distinct lack of available structural information for these enzymes has hindered drug discovery efforts. In this study, the first crystal structure of mannitol-2-dehydrogenase from *A. fumigatus* in an unbound monomeric state (1.8 Å) and bound to its co-factor, NADH (2.1 Å) was solved. These structural data reveal a large, central cavity lined with positively charged residues that readily accommodates the negatively charged diphosphate moiety of NADH. This interaction is further stabilised by a network of hydrogen bond interactions and π-π stacking between Phe45 and the nicotinamide ring of NADH. Furthermore, we have performed rigorous kinetic characterization of *A. fumigatus* mannitol-2-dehydrogenase and demonstrated the dose-dependent inhibitory activity of 1,4-benzoquinone, a cysteine-modifying small molecule inhibitor, in an *in vitro* enzyme assay (IC_{50} = 1.2 ± 0.2 μM). Proteomics analysis also reveals 1,4-benzoquinone irreversibly modifies all five mannitol-2-dehydrogenase cysteine residues. From this work, we have established the foundations for a novel antifungal drug discovery project that targets the fungal mannitol biosynthesis pathway.

Introduction

Aspergillosis is an infection commonly contracted by immunocompromised patients and is caused primarily by the pathogenic fungus *Aspergillus fumigatus*. Spores produced by *A. fumigatus* are ubiquitously found in both indoor and outdoor environments and are constantly inhaled into the respiratory system[1]. However, they are neutralised in healthy human hosts by alveolar macrophages and neutrophils to avoid infection[2]. Patients who have an impaired immune system, caused by chemotherapy, corticoid therapy, viral infection or organ transplant patients who require immune-suppressing drugs, become highly susceptible to the

opportunistic nature of fungal pathogens[1]. Infections caused by *A. fumigatus* are localised to the lung tissue, as observed in patients with allergic bronchopulmonary aspergillosis (ABPA), but can progress to more severe forms such as invasive aspergillosis which can affect multiple organs[1]. Despite the availability of existing antifungal drugs used to treat aspergillosis, the mortality rate associated with these infections remains unacceptably high[3]. This is attributed to a myriad of inherent limitations of these drugs including high toxicity, poor bioavailability in target tissues, poor activity spectrum and unfavourable drug interactions[3]. These issues are further compounded by the growing rate of antifungal drug resistance. As the number of immunocompromised patients is expected to rise and the efficacy of existing antifungal drug treatment continues to decrease, there is a high demand for new classes of antifungals that target novel pathways to effectively treat human fungal infections. Targeting enzymes within the mannitol biosynthesis pathway has been proposed as a promising approach[4].

Mannitol is a six-carbon polyol that has been identified in various fungal structures including mycelia, fruiting bodies and conidia[4]. Biosynthesised primarily through the activity of mannitol-2-dehydrogenase (M2DH, UniProtKB Q4WQY4) and mannitol-1-phosphate 5-dehydrogenase (M1P5DH, UniProtKB Q4X1A4), mannitol acts as a carbohydrate reserve, store of reducing power, osmolyte, and quencher of reactive oxygen species (ROS), all of which are roles that contribute to fungal survival and pathogenicity. Mannitol has been previously implicated in fungal survivability, specifically in the high stress tolerance of conidia[5-7]. Disruption of mannitol production via knockout of the gene encoding M1P5DH in *Aspergillus niger*, a closely related species to *A. fumigatus*, conidia results in increased sensitivity to high temperatures, oxidative stress, freezing and lyophilisation[6]. Thus, mannitol has been suggested to provide a protective effect against stressful conditions.

Studies into the role of mannitol in establishing virulence have also been explored in two human fungal pathogens, *Cryptococcus neoformans* and *A. fumigatus*. Both fungal species have been shown to produce and secrete mannitol during infection in an experimental disease model[8, 9]. In rabbits with experimental meningitis inoculated with *C. neoformans*, higher levels of mannitol have been detected in cerebrospinal fluid[8]. Similarly, rats with experimental aspergillosis showed an increase in mannitol levels in liver tissue and serum[9]. To determine the effect of secreted fungal mannitol in these disease models, a low mannitol producing isolate of *C. neoformans* was injected intravenously into mice[5]. In this experimental disease model, 100% of mice infected with low mannitol producing isolates survived 60 days post challenge whereas 100% of mice infected with wild-type *C. neoformans* H99 succumbed to the infection by 51 days. The severely diminished production of mannitol rendered the isolate 5,000 times less virulent than wild-type[5]. Phenotypic characterisation of this mutant showed that it was also less resistant to environmental stresses such as temperature and salinity and was hypersensitive to ROS[10].

These studies suggest that mannitol plays a major role in fungal survival and pathogenicity in animal hosts. In addition, there are no human homologues of these mannitol biosynthesis enzymes and thus they have been suggested to be attractive antifungal drug targets[11]. In this study, we present the first crystal structure of M2DH from *A. fumigatus*, in an unbound state and bound to NADH. These structural data have been used to provide insights into the active

site residue composition, describe the characteristics of targetable substrate binding pockets and identify key interactions that stabilise the binding of NADH to the M2DH active site. The kinetic parameters of the interconversion reaction catalysed by M2DH are also described and we have demonstrated the inhibitory activity of 1,4-benzoquinone (1,4-BQ) against *A. fumigatus* M2DH. These data form the basis of a promising antifungal discovery project that specifically targets the mannitol biosynthesis pathway.

Materials and Methods

Construction of M2DH-pMCSG9 Expression Vector

Total RNA from *A. fumigatus* Af293 was extracted using TRIzole solution (Invitrogen) and first strand cDNA was synthesised using SuperScript® III First-Strand Synthesis System (Invitrogen). The open reading frame of M2DH (AFUA_4G14450) was PCR-amplified with a 5' NdeI site and a 3' HindIII site for restriction enzyme cloning, and a C-terminal 6x His tag for purification. The NdeI/HindIII digested PCR product was ligated to NdeI/HindIII digested pMCSG9 and the expression vector was transformed into *Escherichia coli* BL21 (λDE3) for protein expression.

Forward Primer: 5'- TTT *CAT ATG* GCA CCT CTC AAG CTC AAT AG – 3'
NdeI Homology to M2DH

Reverse Primer: 5'- TTA *AAG CTT* TTA **ATG GTG ATG GTG ATG GTG** AGA GCC GTT AAT GTA CTT GGG GAG CG – 3'
HindIII 6x His Tag Homology to M2DH

Expression and Purification of Mannitol-2-Dehydrogenase

M2DH was expressed in cells cultured in Luria broth supplemented with ampicillin (0.2 mg/mL) at 37 °C to an OD_{600} of 0.8. Protein expression was then induced for 16 h at 16 °C with the addition of 0.5 mM IPTG. Cells were harvested by centrifugation, resuspended in Buffer A (20 mM Tris-HCl pH 8.0, 500 mM NaCl, 10 mM imidazole, 2 mM β-mercaptoethanol), and stored at -80 °C until purification. Cells were thawed, then lysed in a M110 L microfluidizer processor (Microfluidics) and clarified by centrifugation before loading onto a 5 mL Zetasep Nickel NTA column equilibrated in Buffer A. The column was washed with 6 CV of a 90:10 ratio of Buffer A: Buffer B (20 mM Tris pH 8.0, 500 mM NaCl, 250 mM imidazole, 2 mM β-mercaptoethanol) to remove contaminants and M2DH was eluted using an imidazole gradient from 10 mM to 250 mM. The purity of M2DH was analysed using SDS-PAGE. Fractions containing M2DH were pooled, dialysed at 4 °C for 16 h into storage buffer (50 mM Tris-HCl pH 8.0, 0.5 mM EDTA, 5% glycerol, 1 mM DTT) and concentrated to 28.6 mg/mL in a 10,000 MWCO Amicon Ultra-15 Centrifugal Filter (Millipore). Aliquots of M2DH were flash frozen in liquid nitrogen and stored at -80 °C.

Crystallisation of Mannitol-2-Dehydrogenase and Data Collection

M2DH microcrystals that formed in 30% PEG 4000, 0.2 M $MgCl_2$, 0.1 M Tris pH 8.5 using sitting drop vapour diffusion were crushed and diluted 1:100 in reservoir solution to produce a seed stock. Crystals of unbound M2DH were grown using a modified random microseed matrix screening method[12]. Using hanging drop vapour diffusion, unbound M2DH crystals formed in a 0.1:1:1 ratio of seed stock, M2DH (28.6 mg/mL) and 25% PEG 3350, 0.15 M Tris pH 8.5, 0.2 M NaCl. Cubic crystals (100 µm x 100 µm) of unbound M2DH were fully formed after two weeks of incubation at 16 °C. NADH-bound M2DH crystals were obtained by adding 1 µL of NADH solution (30 mM) dissolved in 27% PEG 4000, 0.22 M lithium sulphate, 0.1 M Tris

pH 8.5 directly to unbound M2DH crystals, formed in 25% PEG 4000, 0.25 M lithium sulphate, 0.1 M Tris pH 8.5 and soaked for 17 h. Using a nylon CryoLoop, a single crystal was mounted, cryoprotected in *Paratone-N* (Hampton Research) and flash frozen in liquid nitrogen. Diffraction data of unbound M2DH crystals were collected at the Australian Synchrotron Macromolecular Crystallography Beam Line MX1 using an oscillation angle of 1° (yielding 360 frames) at a wavelength of 0.954 Å[13]. Diffraction data of NADH-bound M2DH crystals were collected at the Australian Synchrotron Macromolecular Crystallography Beam Line MX2 using the Australia Cancer Research Foundation (ACRF) detector[14]. Data was collected using an oscillation angle of 0.1° (yielding 3600 frames) at a wavelength of 0.954 Å.

Data Processing and Structure Refinement

Data integration was performed using *XDS*, converted to an mtz file using *Pointless*, then scaling and merging was performed using *Aimless* (CCP4 program suite)[15-17]. Molecular replacement was used to solve the phase problem using *Phaser MR* (CCP4 program suite) and M2DH from *Pseudomonas fluorescens* as an initial search model (1LJ8)[18]. The solved structure of *A. fumigatus* M2DH (7RK4) was then used as a search model for the NADH-bound *A. fumigatus* M2DH structure. Manual rebuilding was completed in *WinCoot* and structure refinement in *Phenix.refine*[19, 20]. Structure validation was assessed using *MolProbity*[21]. Statistics of data processing and structure refinement are summarised in Table 1. Protein structure visualisation and figure preparation was completed using *PyMOL* version 2.3.4[22]. Assessment of protein interfaces and domain architecture was completed using *PDBe PISA* version 1.52 and *InterPro*, respectively[23, 24].

Table 1. Statistics of data processing and structure refinement from diffraction data

Data Processing	**Apo M2DH**	**M2DH-NADH**
Wavelength (Å)	0.954	0.954
Space Group	C 1 2 1	C 1 2 1
Unit Cell Dimensions		
a, b, c (Å)	219.48 75.04 69.65	209.88 71.29 69.16
α, β, γ (°)	90 106.30 90	90 101.15 90
Resolution (Å)	37.52 – 1.80 (1.86 – 1.80)	42.91 – 2.10 (2.18 – 2.10)
Unique Reflections	99146 (9736)	58464 (5816)
Multiplicity	7.1 (7.3)	6.6 (6.6)
Completeness (%)	98.45 (97.44)	99.65 (99.85)
R_{merge}[a]	0.08 (1.05)	0.16 (1.95)
R_{meas}[b]	0.09 (1.13)	0.18 (2.12)
Average I/σ (I)	14.47 (2.08)	8.61 (1.94)
CC1/2[c]	0.99 (0.74)	0.99 (0.57)
Structure Refinement		
No. Atoms	9394	8340
Protein	7874	7656
Water	1520	596
Ligands	N/A	88
R_{work}[d]	0.16 (0.27)	0.19 (0.31)
R_{free}[e]	0.20 (0.31)	0.25 (0.37)
RMS Deviations		
Bonds (Å)	0.004	0.007
Angles (°)	0.64	0.84
Ramachandran Analysis		
Outliers	0	0
Favoured	97.20	96.74
Average *B*-factor	29.71	37.78
Macromolecules	27.60	37.46
Solvent	40.62	41.54
Ligands	N/A	39.53

Statistics for the highest-resolution shell are shown in parentheses.

[a] $R_{merge} = \sum_{hkl}\sum_i |\, I_i(hkl) - \langle I(hkl)\rangle| / \sum_{hkl}\sum_i I_i(hkl)$[25]

[b] $R_{meas} = \sum_{hkl}\{N(hkl)/[N(hkl)-1]\}^{1/2} \times \sum_i |I_i(hkl) - \langle I(hkl)\rangle| / \sum_{hkl}\sum_i I_i(hkl)$[26]

[c] $CC1/2 = \sum(x-\langle x\rangle)(y-\langle y\rangle)/[\sum(x-\langle x\rangle)^2\,(y-\langle y\rangle)^2]^{1/2}$[27]

[d] $R_{work} = \sum |F_o - F_c\,| / \sum |F_o|$ for all data with $F_o > 2\sigma\,(F_o)$, excluding data to calculate R_{free}

[e] $R_{free} = \sum |F_o - F_c\,| / \sum |F_o|$ for all data with $F_o > 2\sigma\,(F_o)$, calculated from 5% of reflections, randomly chosen[28]

Spectrophotometric Enzyme Assay for Kinetic Characterisation

Continuous activity assays of M2DH were performed as described previously with modifications[29]. Assay mixtures contained 100 mM Tris-HCl pH 8.5 and saturating concentrations of substrates (500 mM mannitol and 4 mM NAD^+ for the mannitol biosynthesis direction and 500 mM fructose and 1 mM NADH for the fructose biosynthesis direction). All assays were performed at 37 °C and the reaction was monitored using a PHERAstar FSX microplate reader (BMG LabTech) at 340 nm. The molar absorptivity of NADH (Δ_{340} = 6.22 mM^{-1} cm^{-1}) was used to calculate the formation of NADH in the mannitol biosynthesis direction or the consumption of NADH in the fructose biosynthesis direction. To determine the Michaelis-Menten binding constants (K_M), the concentration of mannitol (7.80 – 500 mM), NAD^+ (0.0625 – 4 mM), fructose (7.80 – 500 mM) or NADH (0.125 – 1 mM) was varied whilst maintaining remaining substrates at saturating conditions. The reaction was started by adding M2DH to a final concentration of 12.5 nM, except in the determination of the NADH K_M which required a lower final concentration of 6.22 nM. The kinetic parameters were determined using data from three biological replicates plated in technical triplicate. Data was fitted to the Michaelis-Menten equation using GraphPad Prism 9.

Spectrophotometric Enzyme Assay to Assess Inhibitory Activity of 1,4-Benzoquinone

Methods used to assay the inhibitory activity of 1,4-BQ against M2DH were based on methods used for kinetic characterisation, described above. Assay mixtures contained 100 mM Tris-HCl pH 8.5, 6.22 nM M2DH and 0.73 mM NAD^+ (equal to the experimentally determined K_M of NAD^+). Inhibitory assays were performed at 37 °C and the reaction was monitored using a PHERAstar FSX microplate reader (BMG LabTech) at 340 nm. The molar absorptivity of NADH (Δ_{340} = 6.22 mM^{-1} cm^{-1}) was used to calculate the formation of NADH in the mannitol biosynthesis direction. To determine the half maximal inhibitory concentration (IC_{50}), the concentration of 1,4-BQ (0.01 – 3.13 µM) was varied. The reaction was started by adding mannitol to a final concentration of 500 mM. The kinetic parameters were determined using data from three replicates plated in triplicate. Data was fitted to a dose-response curve using GraphPad Prism 9 according to the following equation:

$$Y = Bottom + \frac{Top - Bottom}{1 + 10^{(Log\ IC_{50} - X) \times Hill\ Slope}}$$

Where '*Top*' and '*Bottom*' are the upper and lower plateau points on the Y-axis, '*IC_{50}'* is the half maximal inhibitory concentration and '*Hill slope'* indicates the steepness of the family of curves.

Native mass spectrometry

The quaternary structure of M2DH was examined using a Synapt G1 HDMS (Waters) using a nanoelectrospray ionisation source[30]. M2DH was buffer exchanged into 200 mM ammonium acetate (NH_4OAc) (pH 6.8) using an Amicon Ultra-Centrifugal Filter (10,000 MWCO) at 4 °C. Protein concentration was adjusted to 15 µM and 2 µL of protein was loaded onto gold-coated borosilicate glass capillaries prepared in-house. Key instrument parameters were as follows:

capillary voltage (kV): 1.60; sampling cone (V): 50; extraction cone (V): 1.5; trap/transfer collision energy (V): 20/15; trap gas (L/hr): 5.5; backing gas (mbar): ~4.5.

In-solution digestion and LC-MS/MS analysis

Protein samples were prepared for LC-MS/MS analysis according to previously established methods with modificiations[31]. M2DH was diluted in protein dilution buffer (20 mM Tris-HCl pH 7.5, 100 mM NaCl) to a concentration of 4 µM, then incubated with a 20-fold molar excess of DTT at 24 °C for 2 h. Samples were washed three times with protein dilution buffer using an Amicon Ultra-Centrifugal Filter (10,000 MWCO) to remove excess DTT. Half of the DTT-treated M2DH sample was reserved (apo-treated). The remaining sample was then incubated with a 20-fold molar excess of 1,4-BQ at 24 °C for 2 h. Both apo- and 1,4-BQ-treated M2DH samples were lyophilised by freeze drying and stored at -80 °C.

Apo- and 1,4-BQ-treated M2DH samples (50 µg) were reconstituted in 7 M urea in 100 mM ammonium bicarbonate (100 µL). M2DH samples were digested using MS grade Trypsin/Lys-C protease mix (Promega) according to the manufacturer's two-step in-solution digestion protocol with modifications. Briefly, trypsin/Lys-C protease mix (15 µL) from a 100 ng/µL stock was added to each sample in a 1:30 protease:protein ratio (w/w), and the reaction was incubated for 3.5 hat 37 °C while shaking. Solutions were then diluted with 10 mM ammonium bicarbonate (700 µL) and incubated overnight at 37 °C while shaking. Digestions were quenched using trifluoroacetic acid (1% v/v final concentration). Samples were lyophilised prior to reconstitution in MilliQ water (100 µL) and purified using Pierce™ C18 Spin Columns (Thermo Fisher Scientific). Concentrations were verified using an extinction coefficient (ε205), of 31 mg/mL/cm[32].

Digested M2DH samples were analysed using an Ultimate 3000 nano-flow system (Thermo Fisher Scientific) coupled to a LTQ XL Orbitrap ETD mass spectrometer (Thermo Fisher Scientific). Approximately 1 µg of each peptide sample was pre-concentrated on a C18 trapping column (Acclaim PepMap 100 C18 75 µm x 20 mm, Thermo Fisher Scientific) at a flow rate of 5 µL/min using 2% acetonitrile 0.1% (v/v) trifluoracetic acid over 3 min. Peptides were separated using a 75 µm ID C18 column (Acclaim PepMap100 C18 75 µm x 50 cm, Thermo Fisher Scientific) at a flow rate of 0.3 µL/min where a linear gradient of 8% to 45% solvent B was applied over 32 min. This was followed by a 4 min wash with 90% solvent B prior to a 16 min equilibration process with 5% solvent B. Solvent A: 2% (v/v) acetonitrile, 0.1% (v/v) formic acid; Solvent B: 80% (v/v) acetonitrile, 0.1% (v/v) formic acid. LC-MS/MS acquisitions were controlled by Xcalibur version 2.1 (Thermo Fisher Scientific) and the mass spectrometer was operated in data-dependent mode. Spectra were acquired in positive ion mode over the mass range of 200 – 2000 *m/z* at a resolution of 60 000 in FT mode. The 5 most intense precursor ions were selected for further CID fragmentation using a dynamic exclusion duration of 180 s where the dynamic exclusion criteria included a minimum relative signal intensity of 1000 and ≥2 positive charge state. The isolation width used was 2.0 *m/z* and a normalised collision energy of 35 was applied.

Proteome Discoverer data analysis

Raw MS/MS data were searched against the target sequence of M2DH (Swiss-Prot accession number Q4WQY4) in Proteome Discoverer version 2.5 (Thermo Fisher Scientific). Parameters for the performed searches were as follows: tryptic peptides with a maximum of 2 missed cleavages were allowed, peptide mass tolerance of 20 ppm, fragment mass tolerance of 0.5 Da, methionine oxidation and customised cysteine modification by 1,4-BQ (mass shift of 156.0245 Da) set as variable modifications. The significance threshold was set as P-value < 0.05 and False Discovery Rate (FDR) set to 0.01.

Intact mass spectrometry analysis

Denatured protein mass spectra of both apo- and 1,4-BQ-treated M2DH were obtained using a Micromass Q-TOF 2 mass spectrometer (Waters). Sample (3 µL) was directly injected in 30 % (v/v) acetonitrile and 1% (v/v) formic acid. The instrument conditions were set as follows: polarity set to positive mode; a mass range of 850 – 3200 *m/z*; capillary voltage of 3.5 kV; Cone voltage: 20 V. Mass spectra acquisition and data analysis were performed using MassLynx 4.1 (Waters).

Analytical size-exclusion chromatography

The average oligomeric size of M2DH (50 µM) was determined by analytical SEC using a Superdex 200 10/300 GL analytical-SEC (GE Healthcare) using methods described previously[30].

Accession Codes

Coordinate files and structure factors of unbound M2DH and NADH-bound M2DH structures have been deposited in the Protein Data Bank under the following accession codes: 7RK4, 7RK5.

Results

Analysis of M2DH structure and oligomeric state

There has been a distinct lack of structural information for mannitol biosynthesis enzymes from fungal origins which has delayed the identification of targetable sites for antifungal drug discovery. To address this gap in knowledge, the X-ray crystal structure of M2DH from *A. fumigatus* in an unbound state (1.8 Å) was solved (Figure 1A). The structure is comprised of 22 α-helices, 16 β-strands, 10 3_{10}-helices and 2 π-helices, arranged into two domains (Figure 1A). The N-terminal domain (residues 1-201) features a Rossmann-type fold, characterised by the alternating βαβ motif, which is commonly observed in cofactor-binding proteins whereas the C-terminal domain (residues 232-475) is predominantly composed of α-helices. There are two molecules of M2DH in the asymmetric unit, forming a protein-protein interface encompassing a surface area of 879 $Å^2$. According to the Complex Formation Significance Score calculated by *PDBe PISA*, this dimeric interface is likely to be the result of crystal packing. In order to validate these results, analytical-SEC and native MS was used to determine the oligomeric state of *A. fumigatus* M2DH (Figure 1B-C). These data show the presence of a dominant species that corresponds to one monomeric unit of M2DH (measured molecular

weight: 57,465 Da, theoretical molecular weight: 57,442 Da) with a small proportion of dimeric species also observed (Figure 1B). Analytical-SEC of M2DH shows a single peak elution at 15.2 mL, indicative of monomeric M2DH (Figure 1C). These data indicate that M2DH from *A. fumigatus* appears to be functional in a monomeric state.

High structural similarity exists between homologous enzymes

Consistent with its conserved role in mannitol biosynthesis, the structure of M2DH from *A. fumigatus* shows a high degree of similarity in comparison to M2DH from *P. fluorescens* in terms of its overall fold and domain architecture (Figure 2). Although these homologous enzymes are derived from different domains of life and exhibit moderate sequence conservation (44% sequence identity), superimposition of the α-carbon atoms that comprise the backbone residues reveals a highly conserved overall fold, represented by their low RMSD of 3.1 Å (Figure 2). This superimposition also indicates that the characteristic N- and C-terminal domains adopted by M2DH are also conserved between homologous enzymes (Figure 2).

Characterisation of the NADH-binding site

The interconversion of mannitol and fructose, catalysed by M2DH, requires the use of an NAD^+/NADH co-factor. Visualisation of the electrostatic surface of M2DH reveals a large, central cavity that is formed at the boundary of the N- and C-terminal domains. This substrate-binding site is predominantly lined by basic amino acid residues. The strong electron density observed in the central cavity of holo-M2DH supports modelling of the NADH co-factor (occupancy of 0.79 and 0.92 for each monomer in the asymmetric unit) (Figure 4A). The abundance of positive charges within the active site readily accommodates the negatively charged phosphate groups that bridge the adenosine and nicotinamide riboside groups of NADH (Figure 3A). The nicotinamide riboside moiety is positioned deeper into this binding pocket whereas the adenosine component of NADH resides within a more surface exposed ridge in a region that is largely neutral in electrostatic potential (Figure 3A). The stabilisation of NADH within this binding pocket is facilitated by π-π stacking interactions, an extensive hydrogen bond network and hydrophobic interactions. The pyridine ring of nicotinamide forms a π-π stacking interaction with the Phe45 sidechain whereas the amide forms hydrogen bonds (2.9 Å and 3.1 Å) with the Thr242 backbone (Figure 3B). Additional hydrogen bonds are established between the bridging phosphate groups and the backbone of Gly44 (2.9 Å) and Phe45 (2.8 Å), as well as between both hydroxyl groups of the ribose in the adenosine moiety and the side chain of Asp77 (2.4 Å and 2.9 Å) (Figure 3B).

Kinetic characterisation and inhibition of M2DH

M2DH catalyses the interconversion of mannitol and fructose, using either NAD^+ or NADH as essential cofactors. The apparent Michaelis-Menten binding constants (K_M) measured for both carbohydrate substrates, mannitol and fructose, are consistent with those measured in kinetic characterisation experiments conducted with *A. fumigatus* M2DH performed previously (Table 2) (Figure 4A-D)[11]. However, the apparent K_M values measured for both cofactors, NAD^+ and NADH, recorded in this study were consistently higher by almost 7-fold and 11-fold, respectively. Calculations of the specificity constant ((k_{cat}/K_M) in s^{-1} mM^{-1}) indicate that M2DH

from *A. fumigatus* preferentially converts mannitol to fructose ($k_{cat}/K_{M\ (mannitol)}$ = 5.1 ± 0.55 s^{-1} mM^{-1}) as opposed to the opposite reaction ($k_{cat}/K_{M\ (fructose)}$ = 0.71 ± 0.12 s^{-1} mM^{-1}).

Mechanism of Inhibition of M2DH activity by 1,4-benzoquinone

The small molecule compound, 1,4-BQ, was previously identified as a potent inhibitor of M2DH from *Acetobacter xylinum* (K_i = 0.18 mM)[33]. Using an *in vitro* enzyme assay, 1,4-BQ was also shown to have dose-dependent inhibitory activity against *A. fumigatus* M2DH (Figure 4E). The half maximal inhibitory constant (IC_{50}) determined was 1.2 ± 0.2 nM. A potential mechanism in which 1,4-BQ inhibits M2DH activity is through covalent cysteine modification. Bottom-up proteomic and intact MS approaches were utilised to determine whether 1,4-BQ modifies M2DH cysteine residues (Figure 5). Bottom-up proteomic analysis of 1,4-BQ treated M2DH identified one of the five cysteine residues (Cys69) to contain this specific benzoquinone modification (+156 Da) (Figure 5A-B). Since no other cysteine-containing peptides were detected by LC-MS analysis intact MS analysis was also performed to determine whether other cysteine residues were modified by 1,4-BQ (Figure 5C). Intact MS shows that the majority of M2DH was modified with a small population of unmodified protein present (Figure 5C, pink). Furthermore, we also observed mass shifts corresponding to approximately five benzoquinone modifications (mass difference +785 Da) (Figure 5C, inset) indicating all cysteine residues were modified upon 1,4-BQ treatment. Together, this data shows that inhibition of M2DH activity by 1,4-BQ could be rationalised by irreversible cysteine modification. However, the mechanism appears to be non-specific to M2DH and is therefore, 1,4-BQ is likely to affect several other proteins containing cysteine residues.

Discussion

Mannitol biosynthesis is an integral metabolic process undergone by several pathogenic fungi in order to survive normal environmental conditions and during the infection of a host[8, 9]. When mannitol is produced in high abundance, it is able to efficiently fulfil its roles as a storage of carbohydrates, quencher of ROS, osmoregulator and regulator of pH[34]. These multifaceted roles ultimately contribute to the survival and persistence of pathogenic fungi within a human host. A novel approach to developing drugs that are effective against a broad spectrum of pathogenic fungi lies in the inhibition of the mannitol biosynthesis enzymes. In the existing literature, there have been some interest in targeting two key enzymes capable of producing mannitol from this pathway, M2DH and M1P5D[11]. However, there has been a distinct absence of structural data which has impeded drug discovery efforts. In this study, we present the first crystal structure of M2DH from *A. fumigatus* to establish the necessary foundations for a drug discovery project poised to accelerate the development of novel inhibitors that target the fungal mannitol biosynthesis pathway.

The crystal structures from *A. fumigatus* M2DH unbound and bound to NADH were solved to a resolution of 1.80 Å and 2.10 Å, respectively (Figure 1 and Figure 3). Analysis of the overall structure of *A. fumigatus* M2DH reveals a large degree of structural similarity, in terms of their overall fold and domain architecture, exists between homologous enzymes despite their low sequence similarity (Figure 2). Furthermore, the analysis of oligomeric interfaces in the crystal

structures, analytical-SEC and native mass spectrometry data confirm that M2DH from *A. fumigatus* is monomeric. Due to the highly conserved nature of these enzymes, the dissemination of a publicly available structure of M2DH from *A. fumigatus* may be of interest to other research groups as it can serve as a preferable template for the modelling of homologous enzymes from pathogenic fungi. A subsequent application of these homology models would be in *in silico* screening experiments of large compound or fragment libraries to identify novel inhibitors. This approach would be effective for homologous enzymes that may not be amenable to crystallisation.

In order to identify small molecule inhibitors that are effective against M2DH, it is imperative to define the most 'druggable' binding site using this information to define the target site for high throughput *in silico* screening experiments. There is a singular, central cavity that lies between the N- and C-terminal domains of M2DH, capable of binding the substrates and co-factors required in the interconversion reaction (Figure 3A). Based on structural superimposition between the NADH-bound *A. fumigatus* M2DH structure and the NAD^+- and D-mannitol-bound *P. fluorescens* M2DH structure, the co-factor and substrate bind closely to each other, each in a designated region of this pocket (Figure 6). Specifically, mannitol is positioned adjacent to the nicotinamide group of NAD^+/NADH and extends deeper into the positively charged pocket. From the co-factor-bound structure of *A. fumigatus* M2DH, NAD^+/NADH is stably held in the positively charged pocket by an extensive network of hydrogen bond and π-π-stacking interactions (Figure 3B). Closer inspection of this binding pocket shows a groove that extends above and below the adenine group of NADH into a solvent-exposed region (Figure 3A). Potential approaches to inhibitor design could explore analogues of NADH, particularly those that have chemical modifications on the adenine group to form additional interactions with surrounding residues in the solvent-exposed grooves. An alternate approach may involve joining an NADH and mannitol analogue using a short chemical, non-hydrolysable linker to exploit both regions of the binding pocket. This may be feasible due to the proximity of the two binding sites in the central cavity.

Kinetic characterisation studies were completed for all substrates used in the interconversion reaction (Table 2). Analysis of the specificity constant reveals a preference for M2DH to catalyse the turnover of mannitol to fructose, using NAD^+ as a co-factor, as opposed to the opposite reaction. Although M2DH can produce mannitol, these kinetic data indicate that M1P5DH may instead be primarily responsible for mannitol biosynthesis whereas M2DH only plays a supporting role. This has implications for the targeted disruption of the mannitol biosynthesis pathway in pathogenic fungi by small molecule inhibitors. An effective approach may require simultaneous inhibition of M2DH and M1P5DH enzymes to downregulate mannitol production. Phenotypic characterisation of a *C. neoformans* mutant that produced low levels of mannitol revealed a significant reduction in virulence compared to wild-type and an increased sensitivity to environmental stressors[5, 10]. Thus, dual targeting of the *A, fumigatus* mannitol biosynthesis enzymes may be sufficient to induce a comparable phenotype. As a result, this could increase fungal susceptibility to killing mechanisms employed by the immune system and resistance to changes in the host environment will be impaired.

As a proof of concept, we have shown that M2DH enzyme activity can be modulated using a small molecule inhibitor. We show that 1,4-BQ exhibited dose-dependent inhibition against M2DH from *A. fumigatus*. This compound has been shown to be effective against several enzymes, including monoamine oxidase, and its derivatives have been shown to have antimicrobial activity against *Staphylococcus aureus* and *Mycobacterium tuberculosis*[35, 36]. Mechanism of action studies that have focused on compounds that feature a 1,4-BQ scaffold indicate that it modifies the free thiol group on cysteine sidechains, resulting in irreversible inhibition by functioning as a covalent inhibitor[35, 37]. Our proteomics and MS data suggest that a similar mechanism elicits the inhibition of M2DH activity by 1,4-BQ through benzoquinone modifications of all M2DH cysteine residues. There are several structural analogues of 1,4-BQ in which chemical modifications have been added to the existing scaffold. Future work should explore the structure-activity relationships of these analogues as inhibitors of M2DH.

From this work, we have structurally and kinetically characterised the mannitol biosynthesis enzyme, M2DH from *A. fumigatus*. From these data, we have provided insights into targetable binding pockets to aid structure-guided drug design of novel inhibitors. As a proof of concept, we have shown that the small molecule 1,4-BQ can modulate the activity of M2DH. Together, these data provide the necessary foundations to establish an antifungal drug discovery project that specifically targets the mannitol biosynthesis pathway of pathogenic fungi.

Accession Codes

M2DH: Q4WQY4

M1P5DH: Q4X1A4

Acknowledgements

We thank the Adelaide Proteomics Centre (University of Adelaide) and Flinders Analytical (Flinders University, Australia) for access to MS instrumentation.

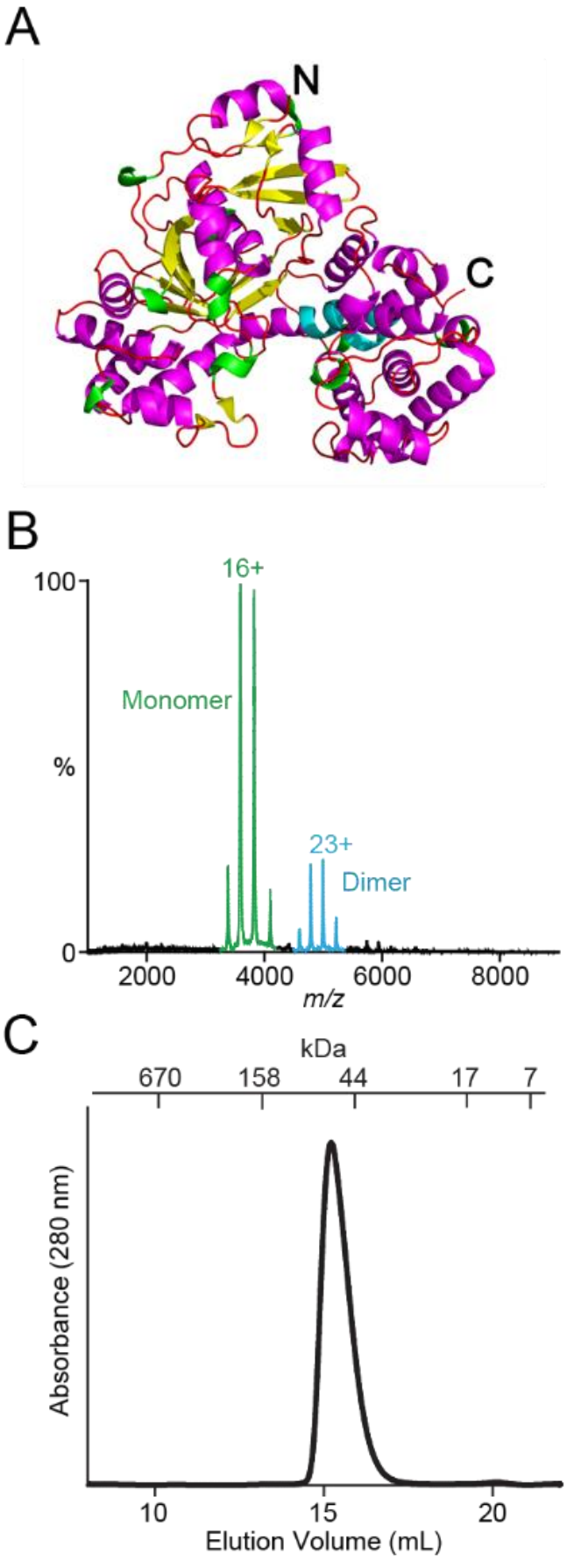

A
N
C
B
16+
100
Monomer
%
23+
Dimer
0
2000
4000
6000
8000
m/z
C
kDa
670
158
44
17
7
Absorbance (280 nm)
10
15
20
Elution Volume (mL)

Figure 1. Crystal structure of a monomeric M2DH from *Aspergillus fumigatus*. **(A)** The structure of M2DH in cartoon representation (1.8 Å, 7RK4). Secondary structure is colour-coded as follows: α-helix (magenta), β-strand (yellow), 3_{10}-helix (green), π-helix (cyan) and loops (red). **(B)** Native MS shows M2DH (15 μM in 200 mM ammonium acetate) is predominantly monomeric (green) with a small population of dimer (blue). **(C)** Analytical-SEC of M2DH (50 μM) in 50 mM phosphate buffer (pH 7.4). Elution volumes of molecular mass standards are indicated above.

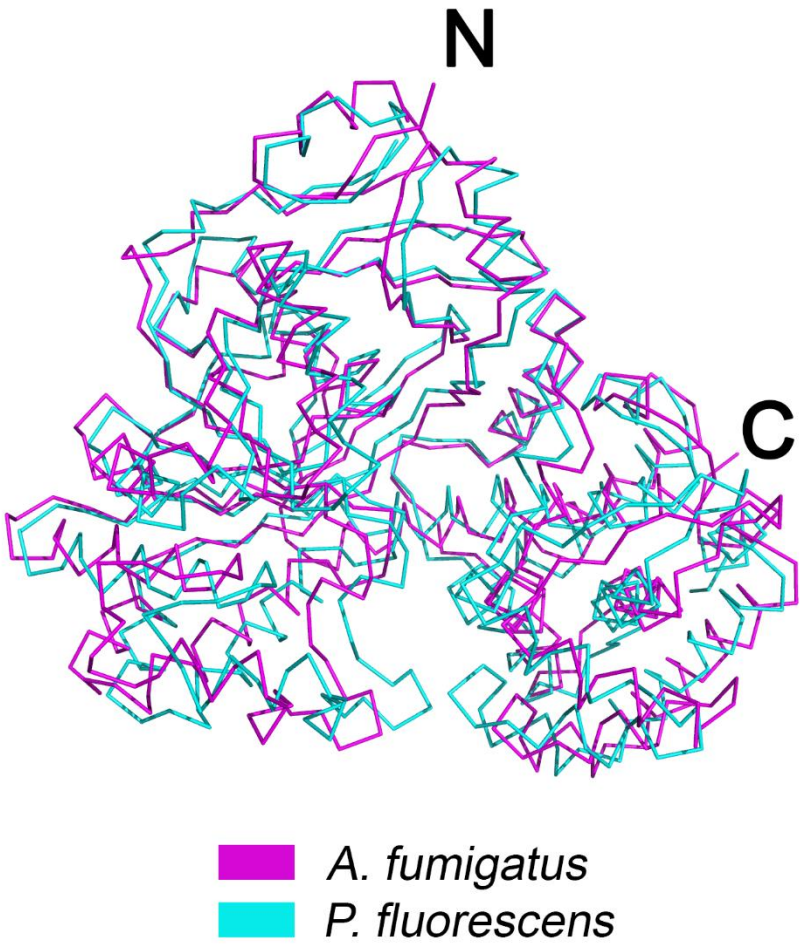
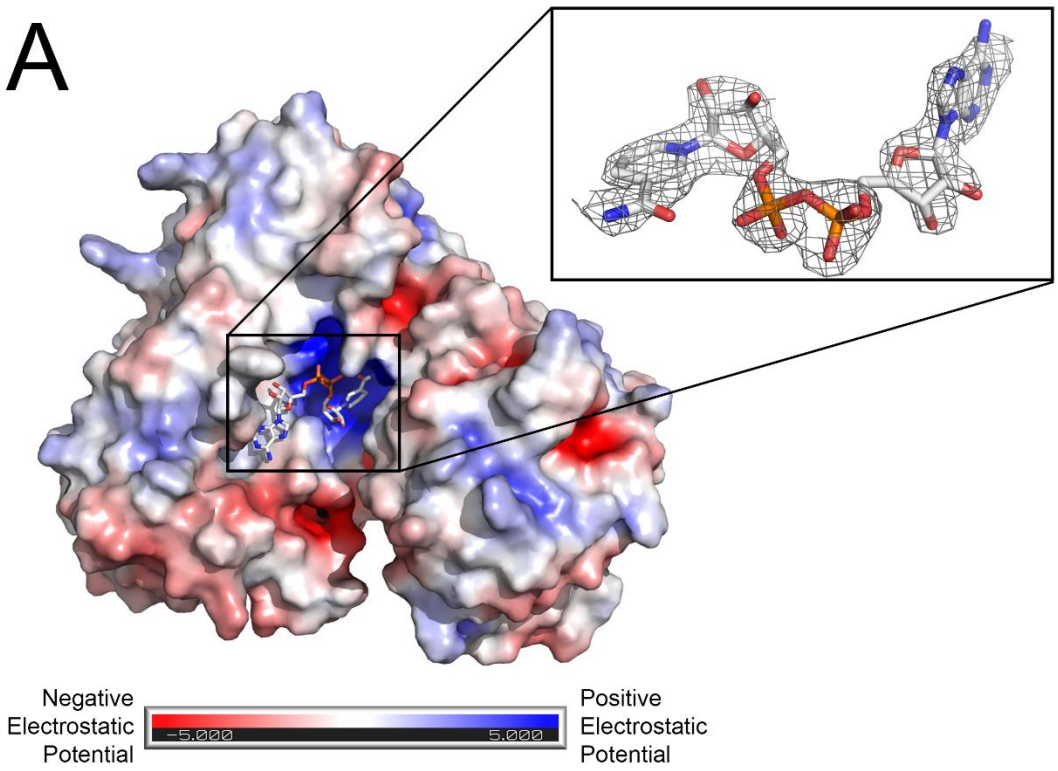


Figure 2. Homologues of M2DH show high structural similarity between species. The crystal structures of *A. fumigatus* (7RK4) and *P. fluorescens* (1LJ8) are coloured in magenta and cyan, respectively.



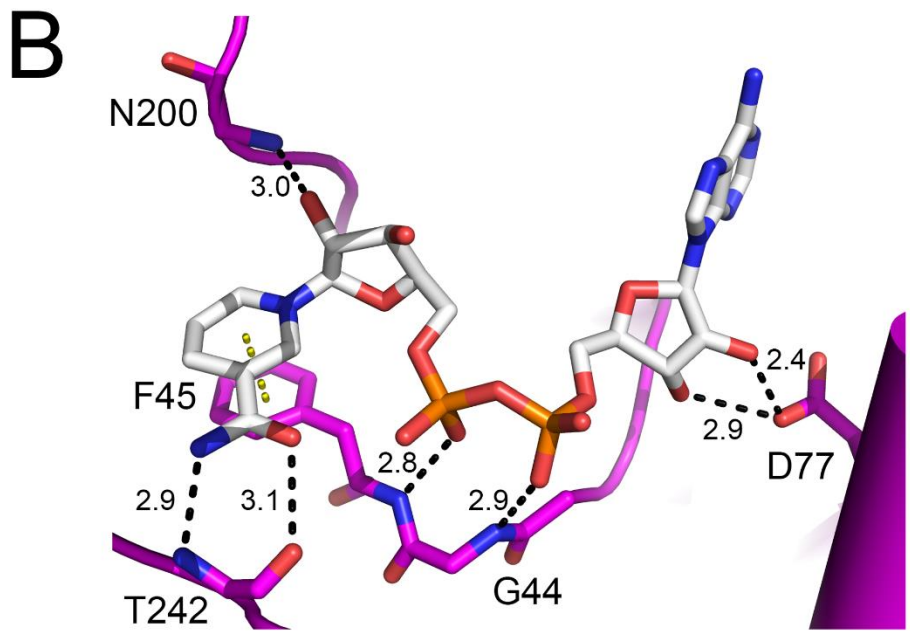


Figure 3. The reduced form of NADH bound to M2DH from *A. fumigatus*. **(A)** The electrostatic surface diagram shows NADH occupying a positively charged cavity. The simulated annealing composite omit map (2Fo-Fc) contoured at 1 σ supports the modelling of NADH (inset). **(B)** Extensive hydrogen bonding (black dashed lines) and π-π stacking (yellow dashed lines) between NADH (sticks) and binding site residues stabilise this interaction.

Table 2. Apparent Michaelis-Menten constants (K_M) and catalytic constant (k_{cat}) of substrates used in the mannitol biosynthesis and fructose biosynthesis reactions. Spectrophotometric activity assays were plated in triplicate and data from three replicates were used.

Substrate	k_{cat} **(s^{-1})**	K_M **(mM)**	k_{cat}/K_M **(s^{-1}/mM)**
Mannitol	81 ± 3.3	16 ± 1.6	5.1 ± 0.55
NAD^+	74 ± 5.1	0.73 ± 0.010	100 ± 7.1
Fructose	40 ± 3.9	56 ± 7.2	0.71 ± 0.12
NADH	63 ± 4.2	0.16 ± 0.014	390 ± 43

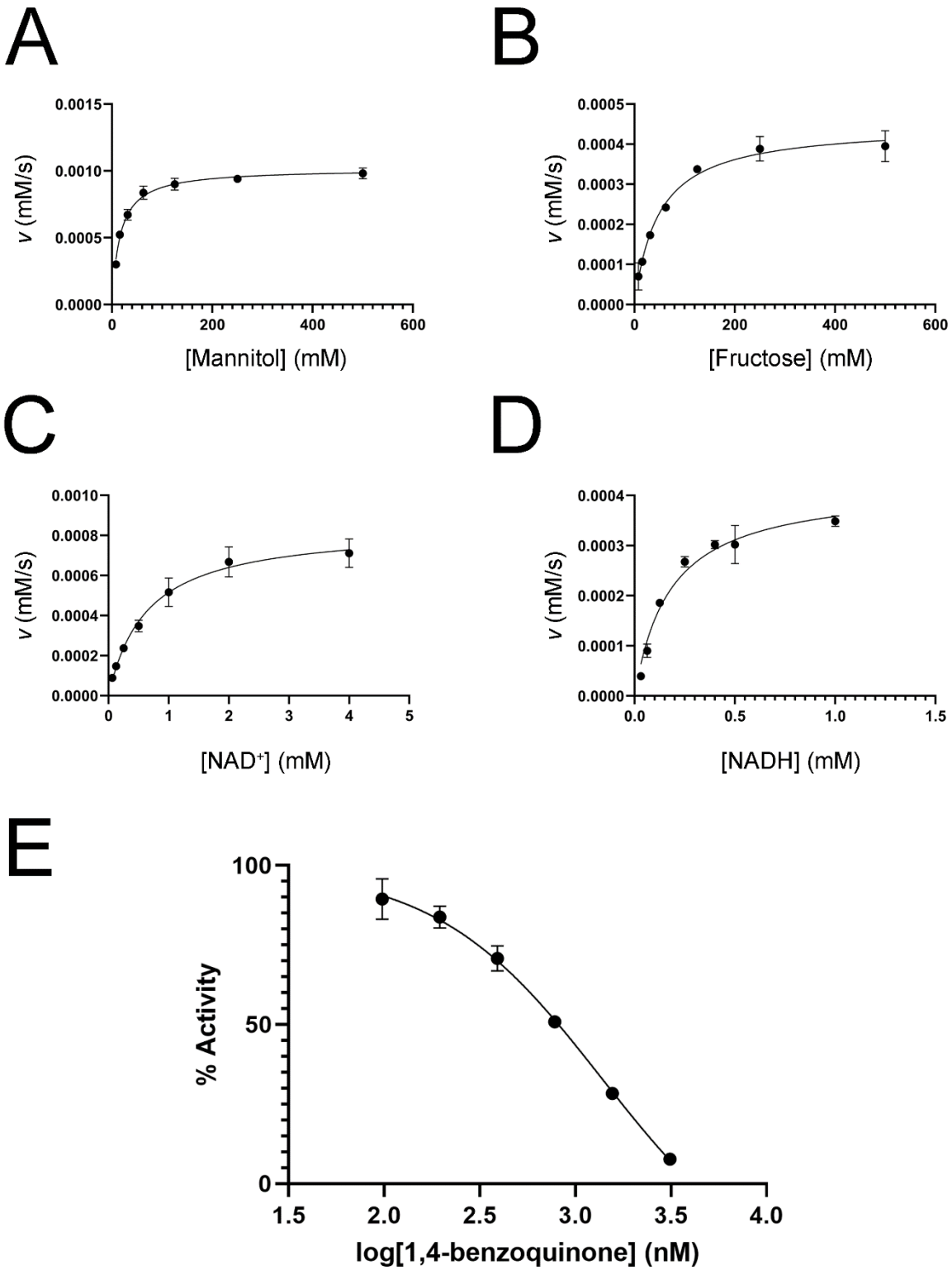

A
v (mM/s)
0.0015
0.0010
0.0005
0.0000
0
200
400
600
[Mannitol] (mM)
B
v (mM/s)
0.0005
0.0004
0.0003
0.0002
0.0001
0.0000
0
200
400
600
[Fructose] (mM)
C
v (mM/s)
0.0010
0.0008
0.0006
0.0004
0.0002
0.0000
0
1
2
3
4
5
[NAD+] (mM)
D
v (mM/s)
0.0004
0.0003
0.0002
0.0001
0.0000
0.0
0.5
1.0
1.5
[NADH] (mM)
E
% Activity
100
50
0
1.5
2.0
2.5
3.0
3.5
4.0
log[1,4-benzoquinone] (nM)

Figure 4. Representative steady-state M2DH activity assays plated in technical triplicate and fitted to the Michaelis-Menten equation to determine the Michaelis-Menten binding constant (K_M) and catalytic constants (k_{cat}) of the substrates, **(A)** mannitol, **(B)** fructose, **(C)** NAD^+ and **(D)** NADH. **(E)** 1,4-benzoquinone induces dose-dependent inhibition of *A. fumigatus* M2DH activity *in vitro*. The half maximal inhibitory constant (IC_{50}) measured was 1.2 ± 0.2 μM (n = 3).

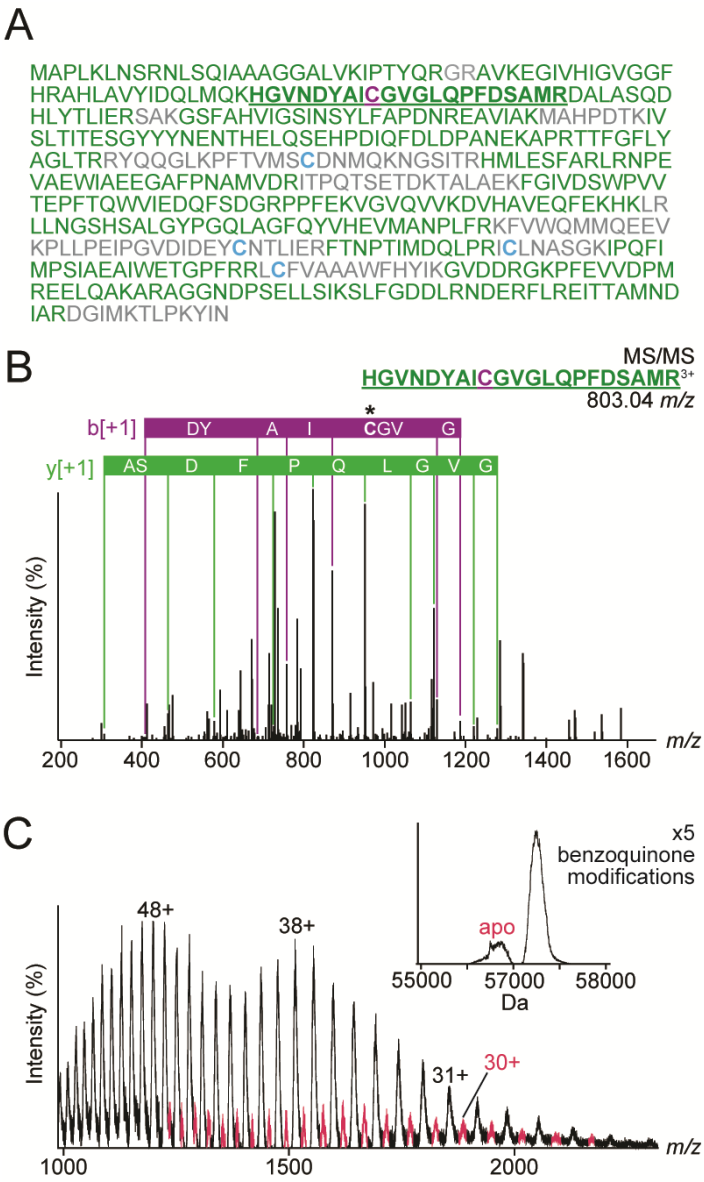


Figure 5. 1,4-BQ inhibits *A. fumigatus* M2DH activity by cysteine modification. **(A)** Amino acid sequence of *A. fumigatus* M2DH (UniProt ID: Q4WQY4) highlighting overall peptide sequence coverage (green) following trypsin/Lys-C digestion (76% sequence coverage). Cysteine residues that were covered (purple) and uncovered (blue) following digestion are highlighted. **(B)** MS/MS spectra of M2DH peptide (residues 61-81^{3+}, 803.04 *m/z*) identifies the modification of Cys69 (purple, *) induced by 1,4-BQ. **(C)** Intact MS of M2DH shows all cysteine residues were modified by 1,4-BQ with a small population of unmodified M2DH present (red, inset)

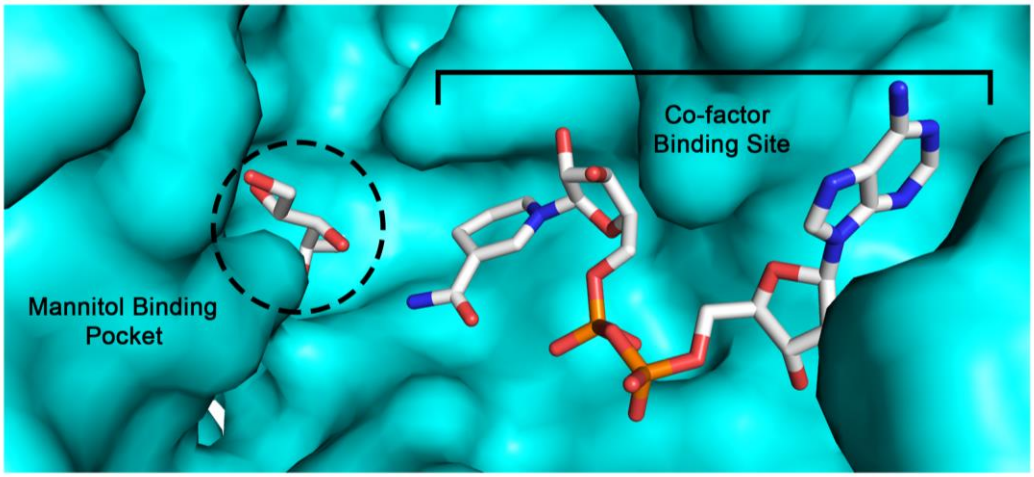


Figure 6: Structural superimposition of the NADH-bound *A. fumigatus* M2DH structure and the NAD^+- and D-mannitol-bound *P. fluorescens* M2DH structure reveals two distinct pockets located within the central cavity in which the substrates bind. The surface representation of *A. fumigatus* M2DH is coloured in cyan, the mannitol and NADH molecules are represented as sticks.

References

[1] Kousha, M., Tadi, R., and Soubani, A. O. (2011) Pulmonary aspergillosis: a clinical review, *Eur. Respir. J. 20*, 156-174.

[2] Dagenais, T. R. T., and Keller, N. P. (2009) Pathogenesis of *Aspergillus fumigatus* in invasive aspergillosis, *Clin. Microbiol. Rev. 22*, 447-465.

[3] Brown, G. D., Denning, D. W., Gow, N. A. R., Levitz, S. M., Netea, M. G., and White, T. C. (2012) Hidden killers: human fungal infections, *Med. Mycol. 4*, 165rv113.

[4] Nguyen, S., Truong, J. Q., and Bruning, J. B. (2021) Targeting unconventional pathways in pursuit of novel antifungals, *Front. Mol. Biosci. 7*, 479.

[5] Chaturvedi, V., Flynn, T., Neihaus, W. G., and Wong, B. (1996) Stress tolerance and pathogenic potential of a mannitol mutant of *Cryptococcus neoformans*, *Microbiology 142*.

[6] Ruijter, G. J. G., Bax, M., Patel, H., Flitter, S. J., van de Vondervoort, P. J. I., de Vries, R. P., van Kuyk, P. A., and Visser, J. (2003) Mannitol is required for stress tolerance in *Aspergillus niger* conidiospores, *Eukaryot. Cell 2*, 690-698.

[7] Wyatt, T. T., van Leeuwen, M. R., Wösten, H. A. B., and Dijksterhuis, J. (2014) Mannitol is essential for the development of stress-resistance ascospores in *Neosartorya fischeri* (*Aspergillus fischeri*), *Fungal Genet. Biol. 64*, 11-24.

[8] Wong, B., Perfect, J. R., Beggs, S., and Wright, K. A. (1990) Production of the hexitol D-mannitol by *Cryptococcus neoformans in vitro* and in rabbits with experimental meningitis, *Infect. Immun. 58*, 1664-1670.

[9] Wong, B., Brauer, K. L., Tsai, R. R., and Jayasimhulu, K. (1989) Increased amounts of the *Aspergillus* metabolite D-mannitol in serum of rats with experimental aspergillosis, *J. Infect. Dis. 160*, 95-103.

[10] Chaturvedi, V., Wong, B., and Newman, S. J. (1996) Oxidative killing of *Cryptococcus neoformans* by human neutrophils. Evidence that fungal mannitol protects by scavenging reactive oxygen intermediates., *J. Immunol. 156*, 3836-3840.

[11] Krahulec, S., Armao, G. C., Klimacek, M., and Nidetzky, B. (2011) Enzymes of mannitol metabolism in the human pathogenic fungus *Aspergillus fumigatus* – kinetic properties of mannitol-1-phosphate 5-dehydrogenase and mannitol 2-dehydrogenase, and their physiological implications, *FEBS J. 278*, 1264-1276.

[12] Truong, J. Q., Nguyen, S., Bruning, J. B., and Shearwin, K. E. (2021) Derivatization of protein crystals with I3C using random microseed matrix screening, *JoVE*, e61894.

[13] Cowieson, N. P., Aragao, D., Clift, M., Ericsson, D. J., Gee, C., Harrop, S. J., Mudie, N., Panjikar, S., Price, J. R., Riboldi-Tunnicliffe, A., Williamson, R., and Caradoc-Davies, T. (2015) MX1: A bending-magnet crystallography beamline serving both chemical and macromolecular

crystallography communities at the Australian Synchrotron, *J. Synchrotron. Radiat. 22*, 187-190.

[14] Aragao, D., Aishima, J., Cherukuvada, H., Clarken, R., Clift, M., Cowieson, N. P., Ericsson, D. J., Gee, C., Macedo, S., Mudie, N., Panjikar, S., Price, J. R., Riboldi-Tunnicliffe, A., Rostan, R., Williamson, R., and Caradoc-Davies, T. (2018) MX2: A high-flux undulator microfocus beamline serving both the chemical and macromolecular crystallography communities at the Australian Synchrotron, *J. Synchrotron Radiat. 25*, 885-891.

[15] Kabsch, W. (2010) XDS, *Acta Crystallogr. D Biol. Crystallogr. 66 (Pt 2)*, 125-132.

[16] Evans, P. (2006) Scaling and assessment of data quality, *Acta Crystallogr. D Biol. Crystallogr. 62*, 72-82.

[17] Evans, P. R., and Murshudov, G. N. (2013) How good are my data and what is the resolution?, *Acta Crystallogr. D Biol. Crystallogr. 69*, 1204-1214.

[18] McCoy, A. J., Grosse-Kunstleve, R. W., Adams, P. D., Winn, M. D., Storoni, L. C., and Read, R. J. (2007) Phaser crystallographic software, *J. App. Crystallogr. 40*, 658-674.

[19] Emsley, P., and Cowtan, K. (2004) Coot: Model-building tools for molecular graphics, *Acta Crystallogr. D Biol. Crystallogr. 60*, 2126-2132.

[20] Afonine, P. V., Grosse-Kunstleve, R. W., Echols, N., Headd, J. J., Moriarty, N. W., Mustyakimov, M., Terwilliger, T. C., Urzhumtsev, A., Zwart, P. H., and Adams, P. D. (2012) Towards automated crystallographic structure refinement with phenix.refine, *Acta Crystallogr. D Biol. Crystallogr. 68*, 352-367.

[21] Chen, V. B., Arendall, W. B., Headd, J. J., Keedy, D. A., Immormino, R. M., Kapral, G. J., Murray, L. W., Richardson, J. S., and Richardson, D. C. (2010) *MolProbity*: All-atom structure validation for macromolecular crystallography, *Acta Crystallogr. D Biol. Crystallogr. 66*, 12-21.

[22] The PyMOL Molecular Graphics System, Version 2.0 Schrödinger, LLC.

[23] Krissinel, E., and Henrick, K. (2007) Inference of macromolecular assemblies from crystalline state, *J. Mol. Biol. 372*, 774-797.

[24] Blum, M., Chang, H.-Y., Chuguransky, S., Grego, T., Kandasaamy, S., Mitchell, A., Nuka, G., Paysan-Lafosse, T., Qureshi, M., Raj, S., Richardson, L., Salazar, G. A., Williams, L., Bork, P., Bridge, A., Gough, J., Haft, D. H., Letunic, I., Marchler-Bauer, A., Mi, H., Natale, D. A., Necci, M., Orengo, C. A., Pandurangan, A. P., Rivoire, C., Sigrist, C. J. A., Sillitoe, I., Thanki, N., Thomas, P. D., Tosatto, S. C. E., Wu, C. H., Bateman, A., and Finn, R. D. . (2020) The InterPro protein families and domains database: 20 years on, *Nucleic Acids Res. 49*, D344-D354.

[25] Arndt, U. W., Crowther, R. A., and Mallet, J. F. W. (1968) A computer-linked cathode ray tube microdensitometer for X-ray crystallography, *J. Phys. E: Sci. Instr. 1*, 510-516.

[26] Diederichs, K., and Karplus, P. A. (1997) Improved *R*-factors for diffraction data analysis in macromolecular crystallography *Nat. Struct. Biol. 4*, 269-275.

[27] Karplus, P. A., and Diederichs, K. (2012) Linking crystallographic model and data quality, *Science 336*, 1030-1033.

[28] Brünger, A. T. (1992) Free R value: a novel statistical quantity for assessing the accuracy of crystal structures, *Nature 355*, 472-475.

[29] Groisillier, A., and Tonon, T. (2015) Determination of mannitol-2-dehydrogenase activity, *Bio Protoc. 5*, e1634.

[30] Nguyen, S., Jovcevski, B., Pukala, T. L., and Bruning, J. B. (2020) Nucleoside selectivity of *Aspergillus fumigatus* nucleoside-diphosphate kinase, *FEBS J. 288*, 2398-2417.

[31] Marshall, A. C., Kidd, S. E., Lamont-Friedrich, S. J., Arentz, G., Hoffman, P., Coad, B. R., and Bruning, J. B. (2019) Structure, mechanism, and inhibition of *Aspergillus fumigatus* thioredoxin reductase, *Antimicrob. Agents Chemother. 63*, e02281-02218.

[32] Anthis, N. J., and Clore, G. M. (2013) Sequence-specific determination of protein and peptide concentrations by absorbance at 205 nm, *Protein Sci. 22*, 851–858.

[33] Oikawa, T., Nakai, J., Tsukagawa, Y., and Soda, K. (1997) A novel type of D-mannitol dehydrogenase from Acetobacter xylinum: occurrence, purification, and basic properties, *Biosci. Biotech. Biochem. 61*, 1778-1782.

[34] Meena, M., Prasad, V., Zehra, A., Gupta, V. K., and Upadhyay, R. S. (2015) Mannitol metabolism during pathogenic fungal-host interactions under stressed conditions, *Front Microbiol 6*.

[35] Mostert, S., Petzer, A., and Petzer, J. P. (2017) The evaluation of 1,4-benzoquinones as inhibitors of human monoamine oxidase, *Eur. J. Med. Chem. 135*, 196-203.

[36] Carcamo-Noriega, E. N., Sathyamoorthi, S., Banerjee, S., Gnanamani, E., Mendoza-Trujillo, M., Mata-Espinosa, D., Hernández-Pando, R., Veytia-Bucheli, J. I., Possani, L. D., and Zare, R. N. (2019) 1,4-Benzoquinone antimicrobial agents against *Staphylococcus aureus* and *Mycobacterium tuberculosis* derived from scorpion venom, *Proc. Natl. Acad. Sci. U.S.A. 116*, 12642-12647.

[37] Diedrich, J. K., and Julian, R. R. (2010) Site selective fragmentation of peptides and proteins at quinone modified cysteine residues investigated by ESI-MS, *Anal. Chem. 82*, 4006-4014.

Chapter 5 (Manuscript):

Structural insights into the enolase-plasminogen complex

Chapter 5 Preface

Chapter 5 is a manuscript submitted for publication to Biochemistry that describes the characterisation of the structure and function of enolase from *A. fumigatus* using X-ray crystallography and *in vitro* enzyme assays. Using a combination of X-ray crystallography, mass spectrometry, protein sequence alignments and *in silico* docking, we have presented a structural model of the enolase-plasminogen complex. I performed all X-ray crystallography experiments, *in vitro* enzyme assays and *in silico* protein-protein docking. I performed the data analysis for these experiments, prepared the figures included in the manuscript, deposited structure factors and coordinate factors into the Protein Data Bank and prepared the manuscript

Statement of Authorship

Title of Paper	Structural insights into the enolase-plasminogen complex
Publication Status	☐ Published ☐ Accepted for Publication ☑ Submitted for Publication ☐ Unpublished and Unsubmitted work written in manuscript style
Publication Details	This publication characterises the structural and kinetic properties enolase from *A. fumigatus*. We have also used *in silico* protein docking methods to develop a structural model of the enolase-plasminogen complex.

Principal Author

Name of Principal Author (Candidate)	Stephanie Nguyen		
Contribution to the Paper	Project conceptualisation, planned and performed experiments, analysed data, prepared, reviewed and edited the manuscript		
Overall percentage (%)	65		
Certification:	This paper reports on original research I conducted during the period of my Higher Degree by Research candidature and is not subject to any obligations or contractual agreements with a third party that would constrain its inclusion in this thesis. I am the primary author of this paper.		
Signature		Date	17/07/21

Co-Author Contributions

By signing the Statement of Authorship, each author certifies that:

i. the candidate's stated contribution to the publication is accurate (as detailed above);

ii. permission is granted for the candidate in include the publication in the thesis; and

iii. the sum of all co-author contributions is equal to 100% less the candidate's stated contribution.

Name of Co-Author	Dr. Blagojce Jovcevski		
Contribution to the Paper	Planned and performed experiments, analysed data, reviewed and edited the manuscript		
Overall Percentage (%)	15		
Signature		Date	20/07/21

Name of Co-Author	Assoc. Prof. Tara Pukala			
Contribution to the Paper	Planned experiments, reviewed and edited the manuscript			
Overall Percentage (%)	10			
Signature			Date	18/7/21

Name of Co-Author	Dr. John B. Bruning			
Contribution to the Paper	Project conceptualisation, planned experiments, reviewed and edited the manuscript			
Overall Percentage (%)	10			
Signature			Date	21/7/21

A structural model of the human plasminogen and *Aspergillus fumigatus* enolase complex

Stephanie Nguyen[1]†, Blagojce Jovcevski[2,3], Tara L. Pukala[2], John B. Bruning[1]*

[1] Institute of Photonics and Advanced Sensing (IPAS), School of Biological Sciences, The University of Adelaide, North Terrace, South Australia, 5005, Adelaide, Australia

[2] Department of Chemistry, School of Physical Sciences, The University of Adelaide, North Terrace, South Australia, 5005, Adelaide, Australia

[3] School of Agriculture, Food and Wine, The University of Adelaide, North Terrace, South Australia, 5005, Adelaide, Australia

KEYWORDS: *Aspergillus fumigatus*, enolase, glycolysis, plasminogen, antifungal

ABSTRACT

The metabolic enzyme, enolase, is expressed in the cytoplasm where it plays a crucial role in maintaining cellular energy production through the process of glycolysis. However, enolase can fulfil 'moonlighting' roles that deviate from its known function. In pathogenic bacteria and fungi, enolase is also expressed on the cell surface where it functions as a virulence factor. Surface-expressed enolase is a receptor for human plasma proteins, including plasminogen, and this interaction facilitates nutrient acquisition and tissue invasion. A novel approach to developing antifungal drugs is to block the formation of this complex. In order to determine the interactions that govern complex formation, we have solved the first crystal structure of enolase from *A. fumigatus* (2.0 Å) and have shown that it preferentially adopts a dimeric quaternary structure using native mass spectrometry. We have also solved two additional structures of *A. fumigatus* enolase bound to their endogenous substrates, 2-phosphoglycerate and phosphoenolpyruvate, and performed rigorous kinetic characterization studies to better understand the details of its enzymatic function. From these data, we have produced a model of the *A. fumigatus* enolase and human plasminogen complex to provide key structural insights that will be paramount in the development of small molecules or peptidomimetics that mimic this interaction for future anti-fungal drug design.

Introduction

The prophylactic and therapeutic treatment of fungal infections is a necessary, but problematic, reality for immunocompromised patients. Impairments in the immune system, caused by immunosuppressive drugs required for organ transplants, chemotherapy, radiotherapy, oral steroids and infection by HIV/AIDS, lead to a higher susceptibility to infection by opportunistic fungal pathogens[1]. Members of the *Aspergillus* genus are amongst the most common fungal pathogens that cause serious mycoses[2]. Specifically, *Aspergillus fumigatus,* a saprophytic fungus that is ubiquitous to the natural environment, has been identified as the leading cause of many of these pulmonary diseases of different severities, including chronic pulmonary aspergillosis, allergic bronchopulmonary aspergillosis (ABPA) and invasive aspergillosis[2, 3]. Conidia produced during asexual reproduction are disseminated through the air and can enter the human pulmonary tract to establish infection[4]. ABPA can develop in response to *Aspergillus* allergens (denoted as 'Asp f' proteins). It has been estimated that ABPA affects 3.5 million individuals globally and is common in patients with asthma and cystic fibrosis[2]. This disease can progress from a mild to severe form, resulting in extensive damage to pulmonary tissue and respiratory failure[5]. At present, there are 23 known *A. fumigatus* allergens recorded in the Allergen Nomenclature database[6],one of which is enolase (Asp f 22), an enzyme that is well-characterized in its role in glycolysis and has also been implicated in fungal virulence. Since it is also a known allergen, it has been explored as an immunological agent for vaccines and studied extensively as a potential antifungal drug target, as reviewed previously[7].

Enolase, a member of the enolase superfamily, is a highly conserved enzyme that is responsible for catalyzing the interconversion of 2-phosphoglycerate (2-PG) and phosphoenolpyruvate (PEP)[8]. It is a metalloenzyme that requires ordered binding of two divalent metal cations, most commonly

Mg^{2+}, to catalyze this reaction[9]. When 2-PG binds to the enolase active site, its carboxyl group coordinates to the Mg^{2+} cofactors for neutralization and stabilization, followed by a conformational change of a neighboring loop to close the active site[10]. The active site lysine residue deprotonates C2 of 2-PG, facilitating the formation of an enolate intermediate (via. a dehydration reaction), there is a collapse of the oxyanion that results in a double bond rearrangement and removal of the hydroxyl group to form water by deprotonating the glutamate residue side chain[11]. Together, this reaction forms the penultimate step of glycolysis, and thus, the catalytic activity of enolase plays critical roles in fungal survival by facilitating glucose metabolism and energy homeostasis.

The metabolic pathways of glycolysis and gluconeogenesis are regulated by cytoplasmic enolase, yet there are indications that it undertakes 'moonlighting functions' that deviate from its canonical role when secreted and presented on bacterial and fungal cell surfaces[12]. Surface localization of enolase has been detected in prominent human pathogens including *A. fumigatus* and *Candida albicans*, as well as other fungi including *Aspergillus flavus*, *Aspergillus terrus*, *Aspergillus nidulans*, *Candida glabrata* and *Saccharomyces cerevisiae*[12-14]. In terms of bacterial species, this phenomenon has also been observed in both *Streptococci* and *Pneumococci*[15, 16]. Although enolase lacks a traditional N-terminal signal peptide, it is still able to translocate to the surface of fungi. It has been proposed that a 28 amino acid motif in the N-terminus of *S. cerevisiae* enolase mediates this translocation event[17]. This motif is highly conserved in other fungal enolases which suggests that translocation may occur via a similar mechanism observed in *S. cerevisiae*[12].

When enolase is presented on the fungal surface, it functions as a virulence factor that contributes to nutrient acquisition and tissue invasion during infection[18]. This is achieved by the recruitment of host plasminogen, a zymogen circulating in the bloodstream, to the fungal surface via binding to surface-expressed enolase[12]. The plasminogen binding site has been proposed to

occur through an internal motif located within the flexible L3 loop or the C-terminal lysine residues, yet the structural arrangement of the complex remains unclear[15, 19, 20]. Once the enolase-plasminogen complex is formed on the fungal surface, tissue- or urokinase-type plasminogen activators cleave plasminogen to form plasmin[12]. This results in activation of the serine protease activity of plasmin required to fulfil its physiological role of tissue remodeling and controlling blood coagulation by degrading extracellular matrix proteins and fibrin clots[21]. Pathogenic microbes can exploit this property by concentrating plasmin activity to the site of infection by interacting with enolase receptors, and this can accelerate tissue invasion and dissemination. A possible therapeutic approach may involve the design of inhibitors that block the interaction between fungal enolase and host plasminogen or preventing the activation of enolase-bound plasminogen. However, this requires a greater understanding of the structure and function of fungal enolases as well as elucidation of the structural arrangement of the enolase-plasminogen complex.

In this study, we present the first crystal structure of *A. fumigatus* enolase in an unbound form and bound to both substrates, 2-PG and PEP. We have also characterized the kinetic properties of this enzyme by determining Michaelis-Menten binding constants of 2-PG and PEP to *A.fumigatus* enolase. Furthermore, we have used the structure of *A. fumigatus* enolase to perform molecular docking and have developed a model of the enolase-plasminogen complex, providing invaluable insights to fuel future drug discovery projects.

Methods

Enolase-pET15b Expression Vector Construction, Protein Expression and Purification

The enolase ORF (AFUA_6G06770) from *A. fumigatus Af293* cDNA (kindly gifted from Professor James Fraser, University of Queensland) was amplified using PCR and fused to an N-terminal His-tag using Phusion polymerase [22]. An N-terminal *Nco*I and C-terminal *Bam*HI site was added for restriction enzyme cloning.

The PCR product and pET15b were treated with *Nco*I and *Bam*HI, then the digested products were ligated to form the Enolase-pET15b expression vector. The vector was transformed into *E. coli* BL21 (λDE3). Methods used for recombinant protein expression, lysis and purification using nickel-affinity chromatography followed methods described previously[23]. Recombinant enolase was purified to >90% homogeneity (assessed by SDS-PAGE) and buffer-exchanged into Storage Buffer (50 mM Tris-HCl pH 8.0, 0.5 mM EDTA, 5% v/v glycerol, 1 mM DTT) using size-exclusion chromatography (HiPrep™ 26/60 Sephacryl® S-200 HR, GE Healthcare). Protein was concentrated to 15.5 mg/mL using an Amicon Ultra-15 centrifugal filter unit (30 kDa MWCO) (Merck Millipore), flash frozen in liquid nitrogen and stored at -80 °C.

Crystallization of Enolase, Data Collection and Structure Refinement

Enolase crystals were grown using hanging drop vapor diffusion from a 1:1 ratio of enolase (15.5 mg/mL) and well solution. Unbound-enolase crystals formed in the presence of $MgCl_2$ (20 mM) in 23% v/v PEG 3350, 0.15 M Bis-Tris pH 5.5. Enolase co-crystallized with 2-PG (20 mM) and $MgCl_2$ (20 mM) in 23% PEG 3350, 0.1 M Bis-Tris pH 5.5, and enolase co-crystallized with PEP (10 mM) and $MgCl_2$ (20 mM) in 21% v/v PEG 3350, 0.05 M Bis-Tris pH 5.5.

A single crystal was mounted onto a cryo-loop, passed through *Paratone-N* (Hampton Research) and flash frozen in liquid nitrogen. Diffraction data was collected at 0.954 Å using the Macromolecular Crystallography MX1 beamline at the Australian Synchrotron, part of ANSTO[24]. An oscillation angle of 1° was used during data collection to yield 360 frames. Data integration was performed using *XDS*, then converted to an *mtz* format using Pointless and scaled and merged using *Aimless* [25-27]. Structures of unbound-, 2-PG- and PEP-enolase all contained two molecules of enolase in the asymmetric unit and the crystals had symmetry in space groups P121, $P12_11$ and $P12_11$, respectively.

Molecular replacement was used to solve the phase problem using *Phaser MR* from the *CCP4i* program suite[28]. A search model was prepared using *CHAINSAW* based on the crystal structure of enolase from *S. cerevisiae* (strain ATCC 204508/S288c) (PDB ID: 1ONE)[10, 29]. The unbound-enolase structure was used as a search model for molecular replacement thereafter for the 2-PG-enolase and PEP-enolase structures. Iterative cycles of manual rebuilding using *Coot* and structure refinement using *Phenix.refine* were completed and the model was validated using *MolProbity*[30-32]. Statistics of data collection and processing, and structure refinement are summarized in Table 1. Crystal structures were visualized in *PyMOL Version 2.3.4* and protein-ligand interactions were identified using *PoseView*[33, 34].

Native Mass Spectrometry

The oligomeric state of enolase, which was buffer-exchanged into 200 mM ammonium acetate (pH 6.8) using an Amicon Ultra-Centrifugal Filter (10,000 MWCO) at 4 °C, was examined using a Synapt G1 HDMS (Waters) using a nanoelectrospray ionisation source[23, 35]. Enolase (15 µM) was loaded into gold-coated borosilicate glass capillaries prepared in-house. Key instrument

parameters were as follows: capillary voltage (kV): 1.65; sampling cone (V): 50; extraction cone (V): 2.0; trap/transfer collision energy (V/V): 15/10; trap gas (L/hr): 5.5; backing gas (mbar): ~4.5.

Spectrophotometric Enolase Activity Assay

Enolase activity assays were performed in 96-well half-area UV-transparent plates (Corning) described previously[36], with modifications. The standard assay buffer used to determine steady state kinetic parameters included 50 mM Tris-HCl pH 8.5 and saturating concentrations of substrates (4 mM $MgCl_2$, 0.4 M KCl). To determine the apparent Michaelis-Menten binding constants, concentration ranges of 2 mM – 0.06 mM of 2-PG and 3 mM – 0.08 mM of PEP were used. The reactions were initiated with the addition of enolase (final concentration of 30.5 nM for 2-PG or 15.3 nM for PEP) and the reaction was monitored immediately at 290 nm at 37 °C using a PHERAstar FSX microplate reader (BMG Lab Technologies). The rate of PEP synthesis from 2-PG and the consumption of PEP to form 2-PG in the opposite reaction was measured using the extinction coefficient (ε_{240}= 1520 M^{-1} cm^{-1}). The data was fitted to the Michaelis-Menten equation using Prism 8.0 (GraphPad). Apparent Michaelis-Menten binding constants (K_M) and turnover number were calculated from three technical replicates performed in biological triplicate.

Multiple Sequence Alignments of Enolase Homologues

Multiple sequence alignments of enolase homologues were performed in *T-coffee* (http://tcoffee.crg.cat/) and visualized using *ESPript* (http://espript.ibcp.fr)[37, 38].

Molecular Docking of A. fumigatus Enolase and Human Plasminogen

An amino acid sequence alignment of *A. fumigatus* and *C. albicans* enolase was first performed. Subsequently, the motifs associated with human plasminogen binding in *C. albicans* enolase, previously identified by Satala, D. *et al.* using cross-linking mass spectrometry (MS), were located on the *A. fumigatus* sequence[39]. Protein-protein docking was performed using ICM-Pro Version 3.8-7b (MolSoft L. L. C.) by defining the epitopes associated with complex formation on the receptor (human plasminogen) and the ligand (*A. fumigatus* enolase). The subsequent model of the enolase-plasminogen complex was refined by minimizing free energy, refining clashing side-chains and optimizing geometric restraints.

Accession Codes

The atomic coordinates and structure factors for *A. fumigatus* enolase in an unbound state, bound to 2-PG only and bound to both 2-PG and PEP were deposited into the Protein Data Bank under the accession codes, 7RHV, 7RHW and 7RI0, respectively.

Results

Two molecules of *A. fumigatus* enolase were found in the asymmetric unit with each monomer arranged in an anti-parallel orientation (Figure 1). The monomer consists of 13 α-helices, 16 β-strands and 6 3_{10}-helices. There are two distinct domains, the N-terminal capping domain (α-helices 1-5 and β-strands 1-3) which confers substrate specificity, and the C-terminal β/α-barrel domain (α-helices 6-13 and β-strands 4-16) which catalyzes the interconversion reaction[40]. The dimer interface (buried surface area of 1767.5 $Å^2$) is stabilized by hydrogen bond interactions and

salt bridges between residues located on β1, β2, α6, α13, η1 and several flexible loops and turns. Activation of the enzyme is achieved by binding of divalent metal cations, primarily Mg^{2+}. The unbound-enolase structure shows one Mg^{2+} ion bound to each monomer. The Mg^{2+} ion forms a coordination sphere with the side chains of Asp246, Glu297 and Asp322, and three additional molecules of water. When the first Mg^{2+} binds to the enolase active site, a conformational change is induced to allow the substrate to bind, followed by binding of the second catalytic Mg^{2+}. As enolase was crystallized in the absence of substrates, the unbound-enolase structure shows a characteristic coordination sphere for only one of two possible Mg^{2+} ions.

As observed in the asymmetric unit, *A. fumigatus* enolase formed a dimeric complex. The interface area of this arrangement was calculated to be 1767.5 $Å^2$ which is sufficiently larger than the suggested surface area of 856 $Å^2$ to constitute a true oligomeric surface[41, 42]. The quaternary structure of *A. fumigatus* was then confirmed using native MS (Figure 2). These data indicate that enolase predominantly adopts a dimeric complex (measured molecular weight of 96.8 kDa, theoretical molecular weight of 96 kDa) and a small proportion of enolase is present as a monomer (measured molecular weight of 48.4 kDa, theoretical molecular weight of 48 kDa).

The structure of A. fumigatus enolase shows high conservation in comparison to homologous enzymes

Superimposition of the *A. fumigatus* enolase structures with homologous enzymes from different species shows high conservation of the overall fold, consistent with their high sequence identity (Figure 3A, Table 2). Closer inspection of the amino acid sequence of the L3 loop, the proposed internal plasminogen-binding motif, reveals high sequence identity between the eukaryotic homologues of enolase from fungal (*A. fumigatus*, *S. cerevisiae*, *C. albicans*, *Cryptococcus neoformans*) and human sources (Figure 3B). Although the L3 loops of two *Streptococcal* are

identical, there is significant variability in comparison to other prokaryotes, such as *Escherichia coli*.

The internal plasminogen-binding motif, originally identified in *Streptococcus pneumoniae*, has an amino acid sequence (FYDKERKVYD) that varies between species yet retains its plasminogen-binding function[19]. The starting Phe residue (F) and penultimate Tyr residue (Y) appear to be invariable in both eukaryotic and prokaryotic enolases (Figure 3B). The second residue of the motif tends to be aromatic in nature (Tyr or Phe) whereas the third residue encompasses a basic residue (Lys or Arg) in all species except *Streptococci* in which an acidic residue in that position is preferred (Asp). There is significant variability in residue positions 4 to 7 in terms of length and composition. This is followed by a basic residue (Lys) in position 8 or a hydrophobic residue (Val) in the case of *Streptococci*. In the final position, an acidic Asp residue is preferred, except in the case of *E. coli* in which a hydrophobic Val takes its place. The observed variability in the internal plasminogen-binding motif of enolase, particularly between eukaryotes and prokaryotes, indicates that the structural arrangement of the enolase-plasminogen complex or mechanism of formation may differ between species.

It has been reported previously that the L3 loop is located within one of four larger motifs identified as enolase binding sites for human plasminogen[39]. There is high sequence conservation evident between homologues of enolase in motifs 1 to 3, but significantly more variability in the fourth motif and the L3 loop, as discussed previously. Most notably, motifs 1 and 3 contain peptide sequences, PSGASTG and GANAIL, respectively, which are invariant between both eukaryotic and prokaryotic homologues of enolase (Figure 3B).

Active site interactions and activity between 2-PG- and PEP-bound enolase

Structures of enolase bound to either 2-PG or PEP were obtained using co-crystallization methods. The 2-PG-bound structure of enolase contained a 1:1 stoichiometric ratio of one 2-PG molecule per enolase monomer. Enolase co-crystallized with PEP yielded a structure that contained one molecule of 2-PG and one molecule of PEP bound to each enolase monomer. The simulated annealing composite omit maps (2Fo-Fc) generated for each substrate-bound structure support modelling of 2-PG and PEP (Figure 4). Both 2-PG and PEP adopt similar conformations whilst bound to the active site. Their binding is further stabilized by an extensive hydrogen bond network formed with surrounding residues (Figure 4). In the 2-PG-bound structure, the oxygen atoms from the phosphate moiety form several hydrogen bonds with the S40 (3.1 Å), R376 (2.8 Å) and S377 (2.6 Å) side chains, as well as the A39 (2.8 Å) and S377 (2.8 Å) backbone (Figure 4B). Furthermore, the carboxylic acid functional group of 2-PG forms an additional hydrogen bond with the K398 (2.5 Å) side chain. In the PEP-bound structure, these interactions observed previously in the 2-PG-bound structure are maintained (Figure 4C). In order to assess the kinetic parameters of *A. fumigatus* enolase, an *in vitro* enzyme assay was used to determine the apparent Michaelis-Menten kinetic constants and turnover number for substrates of both reactions in the interconversion of 2-PG and PEP are summarized in Table 3. Both 2-PG and PEP had similar K_M values (0.20 ± 0.01 mM and 0.31 ± 0.01 mM, respectively) and k_{cat} values (11.0 ± 2.0 s^{-1} and 19.0 ± 4.0 s^{-1}, respectively). Although the catalytic efficiency (k_{cat}/K_M) determined for both substrates are very similar, *A. fumigatus* enolase shows a slight preference for the binding and turnover of PEP in comparison to 2-PG.

Modelling of the A. fumigatus enolase and human plasminogen complex

The crystal structure of *A. fumigatus* enolase solved in this study and the existing crystal structure of human plasminogen (PDB ID: 4DUU) were used to model the enolase-plasminogen complex[43]. Previous modelling of the interaction between a homology model of *C. albicans* and human plasminogen, guided by cross-linking MS data that defined the binding interfaces on both proteins, was performed by Satala, D. *et al*[39]. We have built upon this work by defining the equivalent binding motifs on the surface of *A. fumigatus* enolase and human plasminogen to guide protein-protein docking. From this study we proposed a structural model of the interaction between *A. fumigatus* enolase and human plasminogen (Figure 5A). In this model, a segment of human plasminogen (TNPRAGLE) forms a loop structure that inserts into a groove on the enolase surface. This surface is formed by the residues that comprise motif 1 which appears to be stabilized by several hydrogen bond interactions (Figure 3B). The α-helical peptide (EECAAKCEE) sequence in human plasminogen lies in close proximity to the N-terminal α1 and α2 helices of enolase, forming hydrogen bond interactions with sidechains of neighboring residues (Thr72 on enolase surface and Glu28 and Glu29 on plasminogen surface).

Discussion

Although the role of enolase in the glycolysis pathway has been studied extensively, there is growing interest in its 'moonlighting roles' when expressed on the fungal cell surface. Its involvement in the recruitment of host plasminogen and subsequent concentration of plasmin activity to the site of infection assists nutrient acquisition and disease dissemination, both ultimately contributing to fungal virulence[18]. However, the structural arrangement of the fungal enolase-human plasminogen complex is largely unknown. Without this structural information, the design and development of small molecule or peptidomimetic inhibitors that would prevent complex formation would be an arduous process.

To gain greater insights into the structure and function of *A. fumigatus* enolase, structural and kinetic analyses were performed using data from X-ray crystallography and enzyme assays. Consistent with the highly conserved function of enolase in the universal pathway of glycolysis, *A. fumigatus* enolase shares an overall similar fold to homologous enzymes, as shown in a structural superimposition (Figure 3). The arrangement of the characteristic N-terminal capping domain and C-terminal β/α-barrel domain also appears to be highly conserved between species. The kinetic parameters determined for *A. fumigatus* enolase differ from *S. cerevisiae* enolase, despite both being of fungal origin. The Michaelis-Menten constant determined for 2-PG was higher for *A. fumigatus* enolase (0.20 ± 0.01 mM) than *S. cerevisiae* enolase (0.057 mM) whereas PEP were similar (0.31 ± 0.01 mM for *A. fumigatus* compared to 0.264 mM for *S. cerevisiae*)[44]. Although the overall folds of these enolase homologues are similar, the differences in kinetic parameters indicate that there may be subtle species-specific differences in the active site. However, it is likely to be difficult to exploit these active site differences to achieve selective targeting of enolase homologues from pathogenic bacteria and fungi in preference to the human

homologue due to the high conservation of structural features and catalytic mechanism. Thus, targeting the protein-protein interface generated from the enolase-plasminogen complex should be considered as an alternative approach in the development of antifungal drugs.

When comparing the structural motifs associated with plasminogen binding from enolase homologues, there are stark differences between species. There are clear structural perturbations in the L3 loop that contains the internal plasminogen-binding motif between residues 248 and 256 (FYDKERKVYD), originally identified in enolase from *S. pneumoniae* (Figure 3B)[19]. Similarly, there are species-specific differences in the amino acid sequence of this binding motif, particularly between eukaryotic and prokaryotic sources of enolase, indicating that species-specific binding mechanisms for human plasminogen exist (Figure 3B). From an antimicrobial drug design perspective, this is an important facet to consider. As there will be subtle differences in the complex formed between eukaryotic and prokaryotic homologues of enolase and human plasminogen, the protein-protein interfaces involved will be unique. Therefore, to effectively target the species-specific regions of the interface that are essential to complex formation, extensive knowledge of the structural arrangement of *A. fumigatus* enolase and plasminogen is required.

In pursuit of visualizing the interaction between fungal enolase and human plasminogen, we have combined data from X-ray crystallography, native MS, protein sequence alignments, in combination with published cross-linking MS data and *in silico* modelling to develop a structural model of the human plasminogen-*A. fumigatus* enolase complex[39]. The solved crystal structure combined with native MS confirms that *A. fumigatus* enolase forms a dimeric complex with each monomer arranged in an anti-parallel orientation (Figure 1 and 2). This finding is consistent with enolase homologues from both prokaryotic and eukaryotic species, apart from *S. pyogenes, Streptococcus suis* and *Streptoccocus pneumoniae* enolases which adopt an octameric structure[20,

[35, 45-47]. Exploring the oligomeric state adopted by fungal enolase is an important aspect of accurately modelling the enolase-plasminogen complex. In the structural model, a 1:1 stoichiometric ratio of *A. fumigatus* enolase and human plasminogen is assumed (Figure 5A). However, the oligomeric interface of enolase remains intact and accessible for a dimerization event to occur whilst bound to human plasminogen. Hence, with this currently proposed model, it is possible for two molecules of plasminogen to bind to a dimeric enolase receptor expressed on the fungal surface. Previously, it has been proposed that destabilization of the octameric structure of *S. pyogenes* enolase allows greater access of plasminogen to surface binding sites and subsequently, this leads to greater plasminogen activation[35]. When comparing the octameric *S. pyogenes* enolase structure, arranged as a tetramer of dimers, to the dimeric *A. fumigatus* structure, there are distinct differences in the arrangement of their monomeric units that may affect plasminogen binding. Superimposition of the dimeric unit of *S. pyogenes* onto *A. fumigatus* enolase from the proposed model of the enolase-plasminogen complex indicates that a monomer of *S. pyogenes* will block the available binding site for plasminogen. Hence, these observations reinforce the idea that species-specific mechanisms of enolase and plasminogen complex formation exist and destabilization of the quaternary structure may be necessary for plasminogen binding to occur.

From this *in silico* model, the structural features of the binding interface of the enolase-plasminogen complex can be analyzed and the feasibility of targeting this protein using small molecules can be assessed. These analyses are an important consideration since there are several challenges associated with the development of inhibitors that target protein-protein interfaces as opposed to the traditional targeting of enzyme active sites. Generally, protein-protein interfaces tend to be largely hydrophobic and lack well-defined grooves or pockets, leading to poor binding

by small molecules[48]. Analysis of the protein-protein interface reveals an abundance of Arg and Glu residues located on these surfaces which allows complex formation to be stabilized by an extensive network of hydrogen bonds and salt bridges[41]. In terms of the electrostatic potential and surface landscape of both proteins, it is encouraging to observe several regions of charged residues and large grooves and pockets that could potentially be targetable. Although the enolase-plasminogen interface encompasses a large buried surface area (~3006 $Å^2$), the residues that are involved in complex formation will not contribute equally to the binding free energy[49]. Instead, there tend to be clusters of functional residues that engage in critical interactions that disproportionally contribute to the binding affinity, known as 'hot spots'[49]. Therefore, to effectively disrupt complex formation, the hot spots in the enolase-plasminogen protein-protein interface must be targeted. The Knowledge-based FADE and Contacts server (KFC2) has identified two distinct regions on the enolase surface that contain clusters of hot spots between residues 38-106 and 301-369 (Figure 5B)[50, 51]. On the plasminogen surface, these hot spots are dispersed throughout the protein but generally cluster to regions between residues 3-76, 177-214, 460-472 and 495-517 (Figure 5B). These hot spot regions identified from the model of the enolase-plasminogen complex include motif 1 on the enolase surface and the N-terminal motif (EECAAKCEE) on the plasminogen surface, both of which were identified in cross-linking MS experiments (Figure 3B)[39]. The residues that comprise motif 1 appear to have regions of invariance that are consistent between prokaryotic and eukaryotic homologues of enolase that are flanked by regions of divergence. These species-specific differences may indicate that the mechanisms of enolase-plasminogen complex formation, and therefore, the protein-protein interfaces formed between these proteins, may differ between species.

Although the proposed structural model of *A. fumigatus* enolase and human plasminogen was developed using data from several complementary structural techniques, there are inherent limitations associated with the *in silico* methods used to perform protein-protein docking. Due to the rigid nature of docking, the model produced may not take into account protein flexibility and conformational changes that may occur upon complex formation. Hence, improvements to the docking strategy and validation of the proposed model must be considered. Molecular dynamics simulations of the association and dissociation of enolase and plasminogen may serve as a more robust approach to study the possible structure and mechanism of complex formation[52]. In terms of validation of the existing structural model, solving the structure of the protein complex will provide a wealth of information. However, the crystallization of a large protein complex is likely to prove difficult and will consequently become a bottleneck to the structure determination pipeline. Hence, solution-based structural techniques, such as small-angle X-ray scattering must be considered. Since the structures of human plasminogen (PDB: 4DUU) and *A. fumigatus* enolase (PDB: 7RHV) are now available, they can be docked into a molecular envelope of the complex determined using small angle X-ray scattering or using cryo-EM. These approaches could effectively validate proposed models derived from *in silico* docking or molecular dynamics simulations. Ultimately, validation of the existing structural model of the enolase-plasminogen complex will be essential to accurately guide antifungal drug design efforts that target this protein-protein interface.

FIGURES

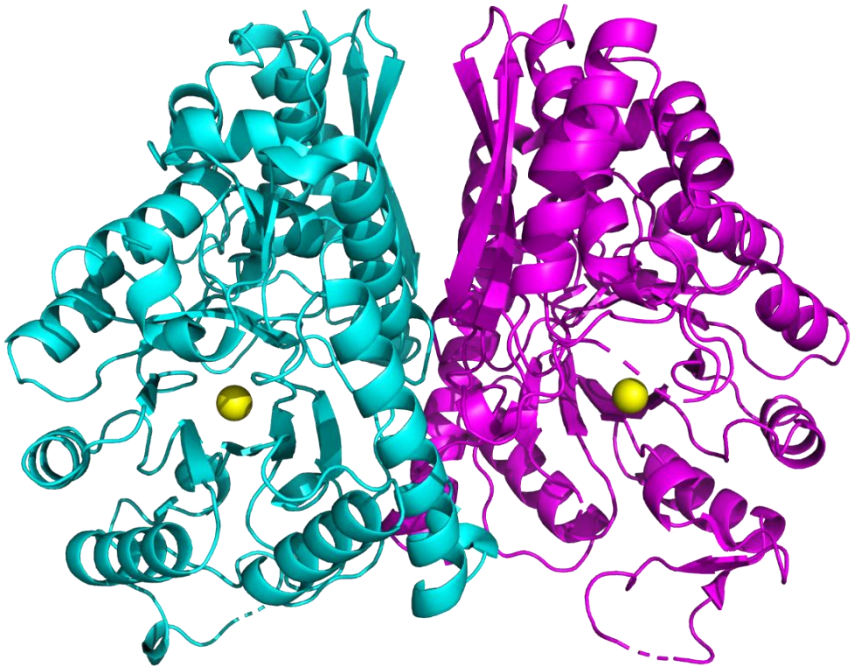


Figure 1. Crystal structure of *A. fumigatus* enolase (2.0 Å, PDB ID: 7RHV) contains two molecules in the asymmetric unit. Each molecule is depicted in cartoon representation (cyan and magenta) and adopts an anti-parallel arrangement. Each monomer of enolase is bound to a singular Mg^{2+} ion (yellow spheres).

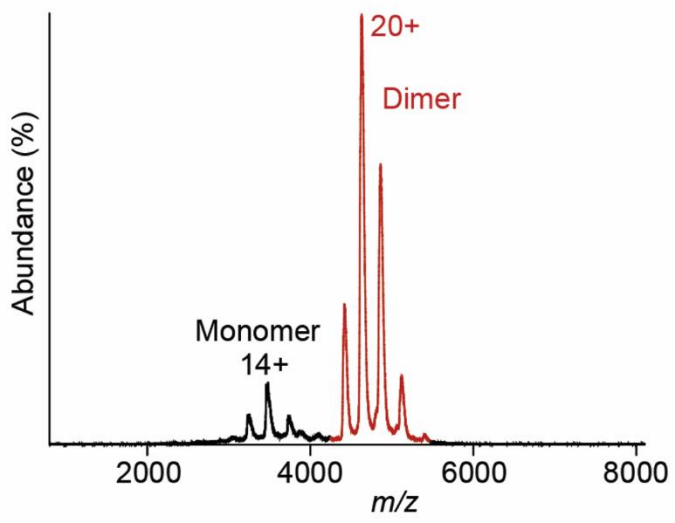
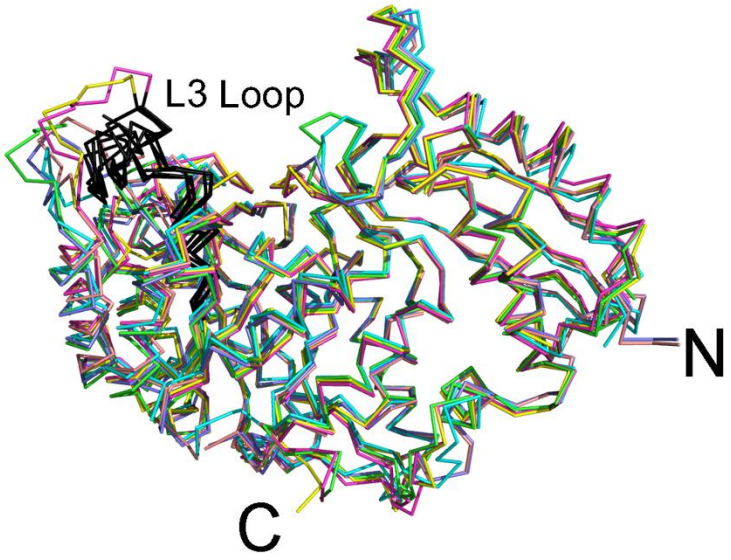


Figure 2. Native mass spectrum shows enolase from *A. fumigatus* preferably forms a dimeric complex. There is high abundance of dimeric enolase (red) and a small proportion of monomeric (black) enolase. Dominant charge states for each oligomer are indicated.

A


L3 Loop
N
C

B

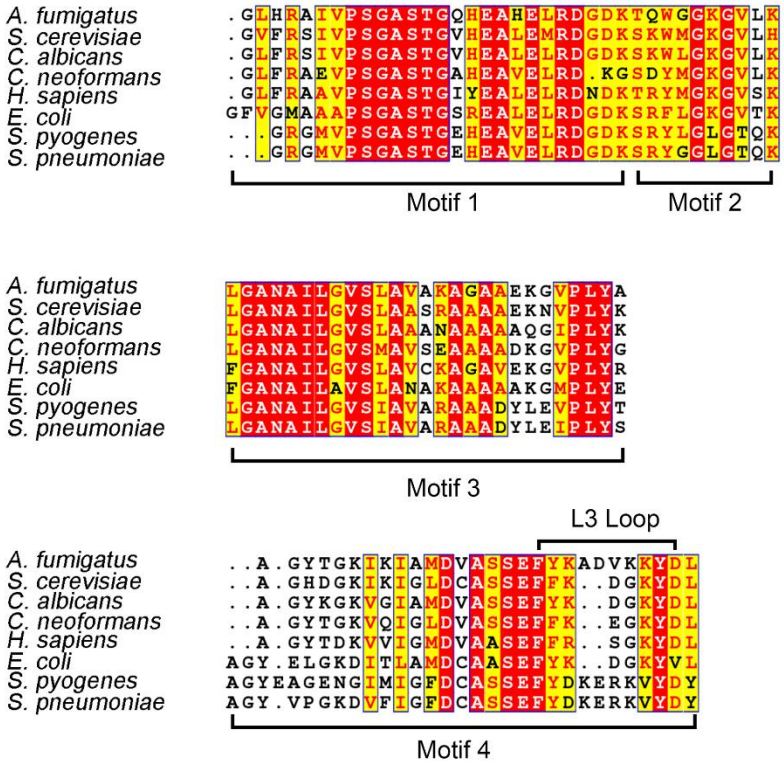

A. fumigatus
S. cerevisiae
C. albicans
C. neoformans
H. sapiens
E. coli
S. pyogenes
S. pneumoniae
Motif 1
Motif 2
Motif 3
L3 Loop
Motif 4

Figure 3. Superimposition of homologous enolase crystal structures from prokaryotic and eukaryotic sources reveal high structural conservation. **(A)** Crystal structures of enolase from *Aspergillus fumigatus* (cyan, PDB ID: 7RHV), *Saccharomyces cerevisiae* (pink, PDB ID: 1ONE), *Homo sapiens* (yellow, PDB ID: 3B97), *Escherichia coli* (purple, PDB ID: 1E9I), *Streptococcus pyogenes* (salmon, PDB ID: 3ZLH) and *Streptococcus pneumoniae* (green, PDB ID: 1W6T) are depicted in ribbon representation[10, 20, 35, 53, 54]. The L3 loop, a motif shown to be important in binding human plasminogen, is highlighted (black). **(B)** Sequence alignment of the four motifs associated with plasminogen binding, including the internal binding motif labelled as the 'L3 loop' from eukaryotic (*Aspergillus fumigatus*, *Saccharomyces cerevisiae*, *Candida albicans*, *Cryptococcus neoformans*, *Homo sapiens*) and prokaryotic (*Escherichia coli, Streptococcus pyogenes*, *Streptococcus pneumoniae*) homologues of enolase.

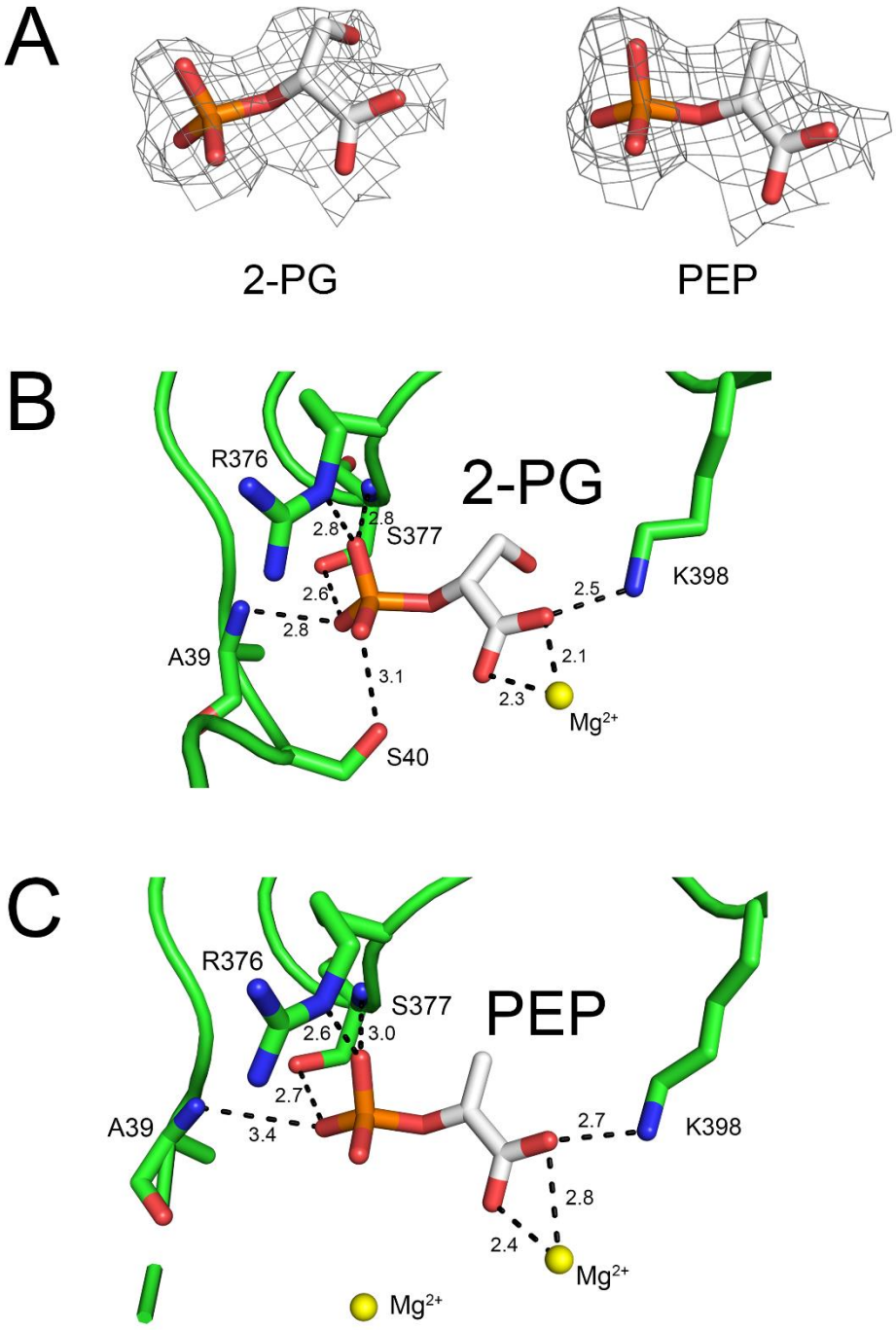
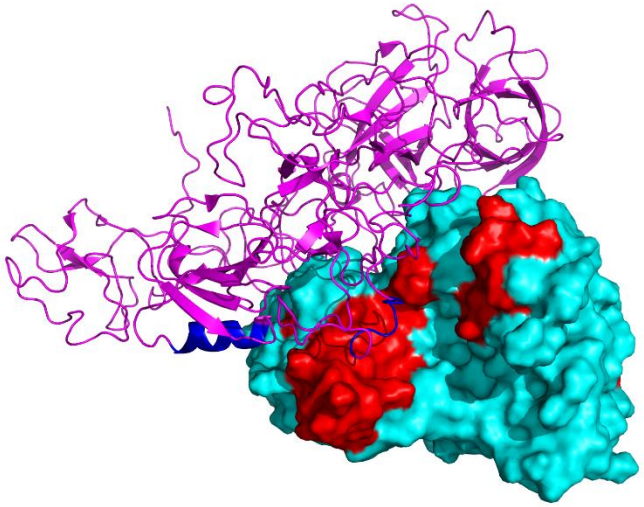

A
2-PG
PEP
B
R376
2-PG
S377
K398
A39
S40
Mg2+
C
R376
S377
PEP
A39
K398
Mg2+
Mg2+

Figure 4. Binding of the substrate, 2-phosphoglycerate (2-PG) and product, phosphoenolpyruvate (PEP) to the enolase active site is stabilized by an extensive hydrogen bond network. **(A)** Simulated annealing composite omit map (2Fo-Fc) contoured at 1 σ supports the modelling of 2-PG and PEP. Residues involved in hydrogen bonding to **(B)** 2-PG and **(C)** PEP are represented in sticks with distances measured in Å. Magnesium ions (Mg^{2+}) are represented as yellow spheres.

A



B

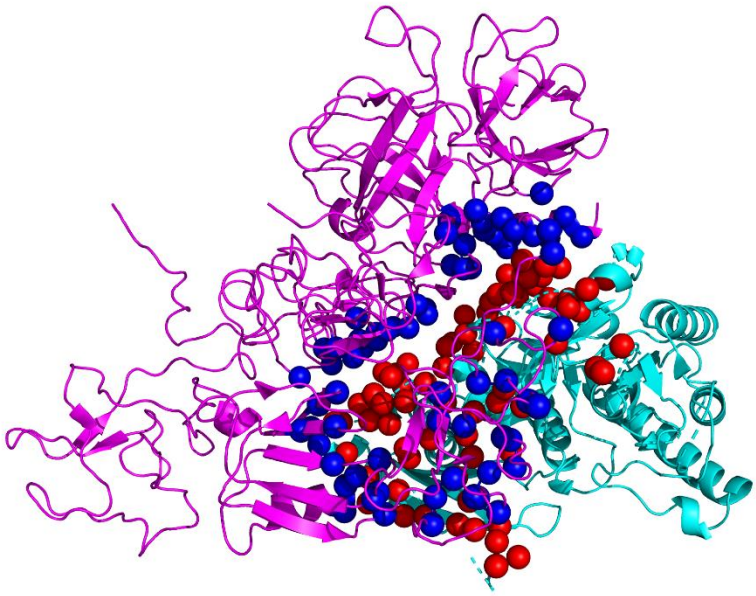


Figure 5. Structural model of the *A. fumigatus* enolase (PDB ID: 7RHV, cyan, surface representation) and human plasminogen complex (PDB ID: 4DUU, magenta,ribbon representation)[43]. **(A)** Regions shaded in red on the *A. fumigatus* enolase structure and in blue on the plasminogen complex correspond to motifs associated with plasminogen and enolase binding, respectively, as determined by cross-linking mass spectrometry experiments performed previously[39]. **(B)** Regions of the protein-protein interface that are predicted to substantially contribute to the binding free energy, known as 'hot spots' are mapped to the proposed structure of the enolase-plasminogen complex. The hot spot residues are represented as red and blue spheres on the ribbon structures of enolase and human plasminogen, respectively.

TABLES

Table 1. Statistics of data processing and structure refinement of diffraction data

Data Processing	**Unbound-Enolase**	**2-PG-Enolase**	**2-PG- and PEP-Enolase**
Wavelength (Å)	0.954	0.954	0.954
Space Group	P 1 2_1 1	P 1 2_1 1	P 1 2_1 1
Unit Cell Dimensions			
a, b, c (Å)	67.17 83.97 84.67	67.38 83.29 84.16	67.86 84.15 84.91
α, β, γ (°)	90.00 97.56 90.00	90.00 97.45 90.00	90.00 98.04 90.00
Resolution (Å)	42.39 – 2.0 (2.07 – 2.0)	37.65 – 1.90 (1.97 – 1.90)	42.50 – 2.30 (2.38 – 2.30)
Unique Reflections	63050 (6254)	72673 (7261)	42121 (4167)
Multiplicity	7.2 (7.3)	6.7 (7.0)	7.2 (7.3)
Completeness (%)	99.95 (99.97)	99.83 (99.96)	99.70 (98.68)
R_{merge} [a]	0.12 (0.41)	0.08 (0.94)	0.10 (0.32)
Average I/σ (I)	9.79 (3.55)	13.51 (1.96)	13.3 (4.7)
CC1/2 [b]	0.99 (0.96)	0.99 (0.77)	0.99 (0.97)
Structure Refinement			
No. Atoms	7640	7054	7225
Protein	6492	6454	6436
Water	1146	576	765
Ligands	2	24	24
R_{free} [d]	0.21 (0.26)	0.23 (0.47)	0.25 (0.39)
R_{work} [e]	0.16 (0.19)	0.19 (0.40)	0.21 (0.29)
RMS Deviations			
Bonds (Å)	0.004	0.003	0.002

Angles (°)	0.64	0.63	0.50
Ramachandran Analysis			
Outliers	0.24	0.35	0.35
Favoured	98.11	98.01	97.18
Average *B*-factor	26.10	39.39	24.11
Macromolecules	24.28	38.95	23.31
Solvent	36.44	44.34	30.81
Ligands	18.99	38.53	26.23

Statistics for the highest-resolution shell are shown in parentheses.

[a] $R_{merge} = \sum_{hkl} \sum_i | I_i(hkl) - \langle I(hkl) \rangle | / \sum_{hkl} \sum_i I_i(hkl)$ [55]

[c] $CC1/2 = \sum (x - \langle x \rangle)(y - \langle y \rangle) / [\sum (x - \langle x \rangle)^2 (y - \langle y \rangle)^2]^{1/2}$ [56]

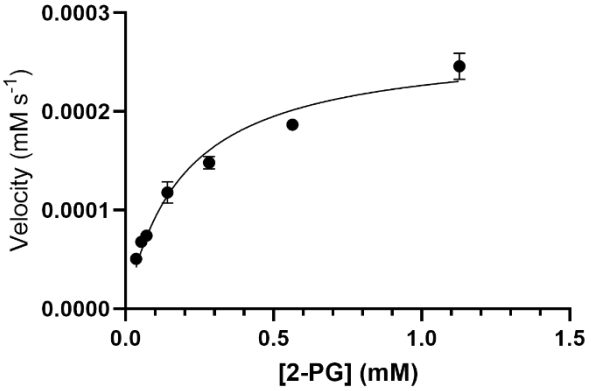
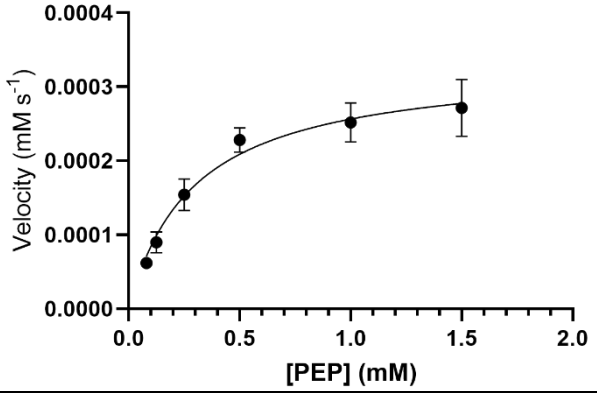
[d] $R_{free} = \sum |F_o - F_c| / \sum |F_o|$ for all data with $F_o > 2\sigma (F_o)$, calculated from 5% of reflections, randomly chosen [57]

[e] $R_{work} = \sum |F_o - F_c| / \sum |F_o|$ for all data with $F_o > 2\sigma (F_o)$, excluding data to calculate R_{free}

Table 2. Enolase from *A. fumigatus* shows high structural similarity and sequence identity to homologous enzymes from *Streptococcus pyogenes, Escherichia coli, Homo sapiens* and *Saccharomyces cerevisiae.*

Species	PDB ID	Sequence Identity Compared to *A. fumigatus* Enolase (%)	RMSD (Å, Cα)
Streptococcus pyogenes	3ZLH[35]	49	1.8, 435
Streptococcus pneumoniae	1W6T[20]	50	1.8, 435
Escherichia coli	1E9I[53]	51	1.8, 431
Homo sapiens	3B97[54]	64	1.6, 433
Saccharomyces cerevisiae	1ONE[10]	74	1.3, 435

Table 3. Apparent Michaelis-Menten kinetic constants (K_M) and turnover rates (k_{cat}) of *A. fumigatus* enolase measured for the substrates, 2-phosphoglycerate (2-PG) and phosphoenolpyruvate (PEP). Data reported is mean ± SD (n = 3).

Substrate	**Representative Binding Curves**	**K_M (mM)**	**k_{cat} (s^{-1})**	**k_{cat}/K_M (s^{-1} mM^{-1})**
2-PG		0.20 ± 0.01	11.0 ± 2.0	55 ± 10.4
PEP		0.31 ± 0.01	19.0 ± 4.0	61.3 ± 13.1

AUTHOR INFORMATION

Corresponding Author

John B. Bruning

Email: john.bruning@adelaide.edu.au

Present Addresses

† Present address: Chemical Biology Division, The Walter and Eliza Hall Institute of Medical Research, Parkville, Victoria, 3050, Australia

Author Contributions

The manuscript was written through contributions of all authors. All authors have given approval to the final version of the manuscript.

Funding Sources

SN is a recipient of an Australian Government Research Training Program stipend scholarship.

ACKNOWLEDGMENT

The authors would like to acknowledge ANSTO for providing the facility in which part of this work was completed. We thank Flinders Analytical (Flinders University, Australia) for access to MS instrumentation.

ABBREVIATIONS

2-PG, 2-phosphoglycerate; PEP, phosphoenolpyruvate; ABPA, allergic bronchopulmonary aspergillosis; MS, mass spectrometry

REFERENCES

[1] Kousha, M., Tadi, R., and Soubani, A. O. (2011) Pulmonary aspergillosis: a clinical review, *Eur. Respir. J. 20*, 156-174.

[2] Bongomin, F., Gago, S., Oladele, R. O., and Denning, D. W. (2017) Global and multi-national prevalence of fungal disease - estimate precision, *J. Fungi 3*, 57.

[3] Kosmidis, C., and Denning, D. W. (2015) The clinical spectrum of pulmonary aspergillosis, *Thorax 70*, 270-277.

[4] Kwon-Chung, K. J., and Sugui, J. A. (2013) *Aspergillus fumigatus* - what makes the species a ubiquitous human fungal pathogen?, *PLoS Pathog. 9*, e1003743.

[5] Shah, A., and Panjabi, C. (2014) Allergic aspergillosis of the respiratory tract, *Eur. Respir. Rev. 23*, 8-29.

[6] Pomés, A., Davies, J. M., Gadermaier, G., Hilger, C., Holzhaouser, T., Lidholm, J., Lopata, A. L., Mueller, G. A., Nandy, A., Radauer, C., Chan, S. K., Jappe, U., Kleine-Tebbe, J., Thomas, W. R., Chapman, M. D., van Hage, M., van Ree, R., Vieths, S., Raulf, M., and Goodman, R. E. (2018) WHO/IUIS allergen nomenclature: providing a common language, *Mol. Immunol. 100*, 3-13.

[7] Nguyen, S., Truong, J. Q., and Bruning, J. B. (2021) Targeting unconventional pathways in pursuit of novel antifungals, *Front. Mol. Biosci. 7*, 479.

[8] Ji, H., Wang, J., Guo, J., Li, Y., Lian, S., Guo, W., Yang, H., Kong, F., Zhen, L., Guo, L., and Liu, Y. (2016) Progress in the biological function of alpha-enolase, *Anim. Nutr. 2*, 12-17.

[9] Brewer, J. M. (1985) Specificity and mechanism of action of metal ions in yeast enolase, *FEBS Lett. 182*, 8-14.

[10] Larsen, T. M., Wedekind, J. E., Rayment, I., and Reed, G. H. (1996) A carboxylate oxygen of the substrate bridges the magnesium ions at the active site of enolase: Structure of the yeast enzyme complexed with the equilibrium mixture of 2-phosphoglycerate and phosphoenolpyruvate at 1.8 Å resolution, *Biochemistry 35*, 4349-4358.

[11] Reed, G. H., Poyner, R. R., Larsen, T. M., Wedekind, J. E., and Rayment, I. (1996) Structural and mechanistic studies of enolase, *Curr. Opin. Struct. Biol. 6*, 736-743.

[12] Funk, J., Schaarschmidt, B., Slesiona, S., Hallström, T., Horn, U., and Brock, M. (2016) The glycolytic enzyme enolase represents a plasminogen-binding protein on the surface of a wide variety of medically important fungal species, *Int. J. Med. Microbiol. 306*, 59-68.

[13] Edwards, S. R., Braley, R., and Chaffin, W. L. (1999) Enolase is present in the cell wall of *Saccharomyces cerevisiae*, *FEMS Microbiol. Lett. 177*, 211-216.

[14] Angiolella, L., Facchin, M., Stringaro, A., Maras, B., Simonetti, N., and Cassone, A. (1996) Identification of a glucan-associated enolase as a main cell wall protein of *Candida albicans* and an indirect target of lipopeptide antimycotics, *J. Infect. Dis. 173*, 684-690.

[15] Pancholi, V., and Fischetti, V. A. (1998) Alpha-enolase, a novel strong plasmin(ogen) binding protein on the surface of pathogenic *Streptococci*, *J. Biol. Chem. 273*, 14503-14515.

[16] Bergmann, S., Rohde, M., Chhatwal, G. S., and Hammerschmidt, S. (2001) α-Enolase of Streptococcus pneumoniae is a plasminogen-binding protein displayed on the bacterial cell surface, *Mol. Microbiol. 40*, 1273-1287.

[17] Miura, N., Kirino, A., Endo, S., Morisaka, H., Kuroda, K., Takagi, M., and Ueda, M. (2012) Tracing putative trafficking of the glycolytic enzyme enolase via SNARE-driven unconventional secretion, *Eukaryot. Cell 11*, 1075-1082.

[18] Dasari, P., Koleci, N., Shopova, I., A.,, Warternberg, D., Beyersdorf, N., Dietrich, S., Sahagún-Ruiz, A., Figge, M. T., Skerka, C., Brakhage, A. A., and Zipfel, P. F. (2019) Enolase From *Aspergillus fumigatus* is a moonlighting protein that binds the human plasma complement proteins factor H, FHL-1, C4BP, and plasminogen, *Front. Immunol. 10*, 2573.
[19] Bergmann, S., Wild, D., Diekmann, O., Frank, R., Bracht, D., Chhatwal, G. S., and Hammerschmidt, S. (2003) Identification of a novel plasmin(ogen)-binding motif in surface displayed α-enolase of *Streptococcus pneumoniae*, *Mol. Microbiol. 49*, 411-423.
[20] Ehinger, S., Schubert, W.-D., Bergmann, S., Hammerschmidt, S., and Heinz, D. W. (2004) Plasmin(ogen)-binding alpha-enolase from *Streptococcus pneumoniae*: crystal structure and evaluation of plasming(ogen)-binding sites, *J. Mol. Biol. 343*, 997-1005.
[21] Law, R. H. P., Abu-Ssaydeh, D., and Whisstock, J. C. (2013) New insights into the structure and function of the plasminogen/plasmin system *Curr. Opin. Struct. Biol. 23*, 836-841.
[22] Bertuzzi, M., van Rhijn, N., Krappmann, S., Bowyer, P., Bromley, M. J., and Bignell, E. M. (2020) On the lineage of *Aspergillus fumigatus* isolates in common laboratory use, *Med. Mycol. 59*, 7-13.
[23] Nguyen, S., Jovcevski, B., Pukala, T. L., and Bruning, J. B. (2020) Nucleoside selectivity of *Aspergillus fumigatus* nucleoside-diphosphate kinase, *FEBS J. 288*, 2398-2417.
[24] Cowieson, N. P., Aragao, D., Clift, M., Ericsson, D. J., Gee, C., Harrop, S. J., Mudie, N., Panjikar, S., Price, J. R., Riboldi-Tunnicliffe, A., Williamson, R., and Caradoc-Davies, T. (2015) MX1: A bending-magnet crystallography beamline serving both chemical and macromolecular crystallography communities at the Australian Synchrotron, *J. Synchrotron. Radiat. 22*, 187-190.
[25] Kabsch, W. (2010) XDS, *Acta Crystallogr. D Biol. Crystallogr. 66 (Pt 2)*, 125-132.
[26] Evans, P. (2006) Scaling and assessment of data quality, *Acta Crystallogr. D Biol. Crystallogr. 62*, 72-82.
[27] Evans, P. R., and Murshudov, G. N. (2013) How good are my data and what is the resolution?, *Acta Crystallogr. D Biol. Crystallogr. 69*, 1204-1214.
[28] McCoy, A. J., Grosse-Kunstleve, R. W., Adams, P. D., Winn, M. D., Storoni, L. C., and Read, R. J. (2007) Phaser crystallographic software, *J. App. Crystallogr. 40*, 658-674.
[29] Stein, N. (2008) CHAINSAW: a program formutating pdb files used as templates in molecular replacement, *J. Appl. Cryst. 41*, 641-643.
[30] Emsley, P., and Cowtan, K. (2004) Coot: Model-building tools for molecular graphics, *Acta Crystallogr. D Biol. Crystallogr. 60*, 2126-2132.
[31] Afonine, P. V., Grosse-Kunstleve, R. W., Echols, N., Headd, J. J., Moriarty, N. W., Mustyakimov, M., Terwilliger, T. C., Urzhumtsev, A., Zwart, P. H., and Adams, P. D. (2012) Towards automated crystallographic structure refinement with phenix.refine, *Acta Crystallogr. D Biol. Crystallogr. 68*, 352-367.
[32] Chen, V. B., Arendall, W. B., Headd, J. J., Keedy, D. A., Immormino, R. M., Kapral, G. J., Murray, L. W., Richardson, J. S., and Richardson, D. C. (2010) *MolProbity*: All-atom structure validation for macromolecular crystallography, *Acta Crystallogr. D Biol. Crystallogr. 66*, 12-21.
[33] The PyMOL Molecular Graphics System, Version 2.0 Schrödinger, LLC.
[34] Stierand, K., and Rarey, M. (2010) PoseView -- molecular interaction patterns at a glance, *J. Cheminform. 2*, P50.

[35] Cork, A. J., Ericsson, D. J., Law, R. H. P., Casey, L. W., Valkov, E., Bertozzi, C., Stamp, A., Jovcevski, B., Aquilina, J. A., Whisstock, J. C., Walker, M. J., and Kobe, B. (2015) Stability of the octameric structure affects plasminogen-binding capacity of Streptococcal enolase, *PLoS One 10*, 21764.
[36] Boël, G., Pichereau, V., Mijakovic, I., Mazé, A., Poncet, S., Gillet, S., Giard, J.-C., Hartke, A., Auffray, Y., and Deutscher, J. (2004) Is 2-phosphoglycerate-dependent automodification of bacterial enolases implicated in their export?, *J. Mol. Biol. 337*, 485-496.
[37] Notredame, C., Higgins, D. G., and Heringa, J. (2000) T-Coffee: a novel method for fast and accurate multiple sequence alignment, *J. Mol. Biol. 302*, 205-217.
[38] Robert, X., and Gouet, P. (2014) Deciphering key features in protein structures with the new ENDscript server, *Nucl. Acids Res. 42*, W320-W324.
[39] Satala, D., Satala, G., Karkowska-Kuleta, J., Bukowski, M., Kluza, A., Rapala-Kozik, M., and Kozik, A. (2020) Structural insights into the interactions of candidal enolase with human vitronectin, fibronectin and plasminogen, *Int. J. Mol. Sci. 21*, 7843.
[40] Gerlt, J. A., Babbitt, P. C., Jacobson, M. P., and Almo, S. C. (2012) Divergent evolution in enolase superfamily: Strategies for assigning functions, *J. Biol. Chem. 287*, 29-34.
[41] Krissinel, E., and Henrick, K. (2007) Inference of macromolecular assemblies from crystalline state, *J. Mol. Biol. 372*, 774-797.
[42] Ponstingl, H., Henrick, K., and Thornton, J. M. (2000) Discriminating between homodimeric and monomeric proteins in the crystalline state, *Proteins: Struct., Funct., Bioinf. 41*, 47-57.
[43] Law, R. H. P., Caradoc-Davies, T., Cowieson, N. P., Horvath, A. J., Quek, A. J., Encanacao, J. A., Steer, D., Cowan, A., Zhang, Q., Lu, B. G., Pike, R. N., Smith, A. I., Coughlin, P. B., and Whisstock, J. C. (2012) The X-ray crystal structure of full-length human plasminogen, *Cell Rep. 1*, 185-190.
[44] Hannaert, V., Albert, M. A., Rigden, D. J., da Silva Giotto, M. T., Thiemann, O., Garratt, R. C., Van Roy, J., Opperdoes, F. R., and Michels, P. A. (2003) Kinetic characterization, structure modelling studies and crystallization of Trypanosoma brucei enolase, *Eur. J. Biochem. 270*, 3205-3213.
[45] Spring, T. G., and Wold, F. (1971) The purification and characterization of *Escherichia coli* enolase, *J. Biol. Chem. 248*, 6797-6802.
[46] Zhang, E., Brewer, J. M., Minor, W., Carreira, L. A., and Lebioda, L. (1997) Mechanism of enolase: The crystal structure of asymmetric dimer enolase - 2-phospho-D-glycerate/enolase - phosphoenolpyruvate at 2.0 Å resolution, *Biochemistry 36*, 12526-12534.
[47] Lu, Q., Lu, H., Qi, J., Lu, G., and Gao, G. F. (2012) An octamer of enolase from *Streptococcus suis*, *Protein Cell 3*, 769-780.
[48] Lu, H., Zhou, Q., He, J., Jiang, Z., Peng, C., Tong, R., and Shi, J. (2020) Recent advances in the development of protein–protein interactions modulators: mechanisms and clinical trials, *Signal Transduct. Tar. 5*, 213.
[49] Cukuroglu, E., Engin, H. B., Gursoy, A., and Keskin, O. (2014) Hot spots in protein–protein interfaces: Towards drug discovery, *Prog. Biophys. Mol. Bio 116*, 165-173.
[50] Darnell, S. J., Page, D., and Mitchell, J. C. (2007) An automated decision-tree approach to predicting protein interaction hot spots, *Proteins 68*, 813-823.

[51] Zhu, X., and Mitchell, J. C. (2011) KFC2: A knowledge-based hot spot prediction method based on interface solvation, atomic density, and plasticity features, *Proteins 79*, 2671-2683.
[52] Pan, A. C., Jacobson, D., Yatsenko, K., Sritharan, D., Weinreich, T. M., and Shaw, D. E. (2019) Atomic-level characterization of protein–protein association, *Proc. Natl. Acad. Sci. U.S.A. 116*, 4244-4249.
[53] Kuhnel, K., and Luisi, B. (2001) Crystal structure of the *Escherichia coli* RNA degradosome component enolase, *J. Mol. Biol. 313*, 583.
[54] Kang, H. J., Jung, S. K., Kim, S. J., and Chung, S. J. (2008) Structure of human alpha-enolase (hENO1), a multifunctional glycolytic enzyme, *Acta Crystallogr. D Biol. Crystallogr. 64*, 651-657.
[55] Arndt, U. W., Crowther, R. A., and Mallet, J. F. W. (1968) A computer-linked cathode ray tube microdensitometer for X-ray crystallography, *J. Phys. E: Sci. Instr. 1*, 510-516.
[56] Karplus, P. A., and Diederichs, K. (2012) Linking crystallographic model and data quality, *Science 336*, 1030-1033.
[57] Brünger, A. T. (1992) Free R value: a novel statistical quantity for assessing the accuracy of crystal structures, *Nature 355*, 472-475.

.

Chapter 6:

Discussion and future directions

6.1 Overall discussion

From this work, four crystal structures have been deposited to the Protein Data Bank of previously uncharacterised enzymes from the *A. fumigatus* proteome and established a solid foundation for four unique antifungal drug discovery projects that target unconventional pathways. From each of the pathways discussed, key enzymes were focused on that have high potential as antifungal drug targets due to their involvement in maintaining fungal survival and virulence. Extensive structural and kinetic characterisation has been completed and *in vitro* enzyme assays have been established for each target of interest. Future work should focus on the assessment of target feasibility before proceeding to the development of high-throughput *in silico* and screening methods to identify compounds that exhibit inhibitory activity from large compound and fragment libraries. Using the established crystallisation conditions identified in this work, the binding mode of these compounds to the target of interest can be determined. This structural information will be crucial in guiding chemical optimisation of the identified compounds to iteratively improve its potency and selectivity. Ultimately, these studies will use this structure-based drug discovery approach to develop novel antifungal drugs that are effective in the treatment of *A. fumigatus* infections.

6.2 Nucleoside selectivity of *Aspergillus fumigatus* nucleoside-diphosphate kinase

The metabolic enzyme, NDK, was chosen as a potential antifungal drug target as it had been previously shown to be essential for viability in several *Aspergillus* species, including *A. fumigatus* (87,89). In order to gain a better understanding of the structural aspects of this enzyme with the intention of targeting the active site, the crystal structure of unbound NDK from *A. fumigatus,* in addition to a full complement of nucleoside-diphosphate-bound structures were solved to investigate the structure-activity relationships of NDK and its 6 possible substrates. Analysis of the fungal NDK structure in comparison to the human homologue indicated that there was a high degree of similarity in terms of the overall fold of the enzyme, oligomeric state and active site residue composition (discussed in Chapter 2). The similarities between antifungal targets and their human homologues must be considered as there is high potential for toxic side-effects to arise as a result of off-target binding of small molecule inhibitors. This aspect was particularly important since NDK plays fundamental roles in survival by maintaining the nucleoside triphosphate pool required for DNA/RNA synthesis and signal transduction. Due to these structural similarities, it is anticipated that selective targeting

of the fungal homologue would be difficult to achieve and therefore, efforts should be focused on alternative antifungal drug targets, discussed later in this chapter (Chapter 6.3, 6.4 and 6.5).

6.3 Structural insights into the antifungal drug target, guanosine monophosphate synthase from *Aspergillus fumigatus*

GMP synthase is an enzyme in the purine biosynthesis pathway that has been extensively studied by multiple research groups and pharmaceutical companies (Merck & Co.) as an attractive antifungal drug target with high potential (37,85). This enzyme has been implicated as a determinant of virulence in several pathogenic fungi, including *A. fumigatus*, *C. albicans* and *C. neoformans* (37,85). Despite the growing interest in targeting GMP synthase to develop novel antifungal drugs, there has been a distinct lack of structural information available. There have been no crystal structures solved of GMP synthase of fungal origin, presenting a significant hindrance to drug discovery efforts. In order to address this gap in knowledge, the first crystal structure of GMP synthase from *A. fumigatus* was solved and rigorously characterised its structure and kinetics (discussed in Chapter 3). Although a human homologue of GMP synthase exists, we have discovered significant differences between the human and fungal structures in terms of the overall protein fold, domain architecture and active site composition. The *A. fumigatus* GMP synthase structure has a distinct absence of a second dimerization domain, which results in its ability to form a dimeric complex. This is in stark contrast to the human homologue, which is predominantly monomeric (94). Closer inspection of the XMP- and ATP- substrate binding sites reveal further divergences in the residue composition that can be exploited in inhibitor design to achieve selectivity. Furthermore, the kinetic properties determined for the substrates used in the GMP synthase reaction also differ to those determined for human GMP synthase, which indicates that there are subtle structural differences that govern substrate binding (103). The structural and kinetic differences identified from these data provide evidence to suggest that GMP synthase is a viable target for antifungal drug design and selective targeting of the fungal homologue is possible.

Future work must then focus on the discovery of chemical scaffolds that are able to modulate the enzymatic function of fungal GMP synthase. Although it was not discussed in Chapter 3, preliminary screening of large compound libraries using high-throughput *in silico* methods is currently being explored, as well as testing hit compounds using *in vitro* enzyme assays. In a collaboration formed with the biotechnology company AtomWise, deep learning AI technology has been used to performed compound screening using millions of compounds that target the ATP-binding pocket of *A. fumigatus* GMP synthesis. The top scoring 80 compounds from this

virtual screen have been acquired and screened using an established *in vitro* enzyme assay. Although these approaches are yet to yield a lead compound for further development, the intended approaches to build upon this work and improve the efficacy of compound screening will be discussed further in Chapter 6.6.

6.4 Targeting the mannitol biosynthesis pathway in *Aspergillus fumigatus*: Characterisation and inhibition of mannitol-2-dehydrogenase

The mannitol biosynthesis pathway has been extensively studied in *C. neoformans* and several *Aspergillus* species to define the role of mannitol in facilitating fungal viability and virulence during infection (64,68,70,73). From this work, it has been suggested that targeting the enzymes that facilitate mannitol production could have potential for a new class of antifungal drugs. Since there is no equivalent pathway in humans and therefore no human homologue of the mannitol biosynthesis enzymes, achieving selectivity is considerably simplified (104). In order to gain greater insights into the structure and function of M2DH from *A. fumigatus,* the X-ray crystal structure was solved in an unbound state and bound to the NADH co-factor. Analysis of these structures revealed a large central cavity that is predominantly lined with positively-charged residues that is capable of binding the carbohydrate substrate and NAD^+/NADH co-factor (discussed in Chapter 4). It was observed that the co-factor binding pocket was relatively large and expansive in comparison to that of the mannitol/fructose-binding pocket. This large region would be more accommodating to a greater range of chemical scaffolds and would provide more space and flexibility for chemical elaboration of lead compounds or fragment growing in a fragment-based drug discovery pipeline. For these reasons, *in silico* screening of a curated fragment library (5707 fragments purchased from Compounds Australia) was performed, targeting the co-factor binding pocket. The top 10 scoring hits from this screen were then tested in an *in vitro* enzyme assay to identify inhibitory activity. From the small-scale screening methods used, compounds have yet to be identified that exhibit M2DH inhibition. The intended approaches to build upon this work and improve the efficacy of compound screening will be discussed further in Chapter 6.6.

6.5 Structural insights into the enolase-plasminogen complex

Enolase is a highly conserved enzyme that plays important roles in glycolysis and gluconeogenesis (105). Although its purpose in the glycolysis pathway is well-characterised, it also has 'moonlighting' roles that deviate from its canonical function. In several pathogenic fungi and bacteria, enolase can be transported to the cell wall where it functions as a receptor for human proteins, such as plasminogen (45). The recruitment of plasminogen to the fungal surface via surface-expressed enolase and subsequent concentration of serine protease activity from activated plasmin at the site of infection facilitates tissue invasion (48). Since enolase is part of an ancient metabolic pathway, its structure and active site residue composition are highly conserved. Hence, selectively modulating enzyme activity by targeting the active site of fungal enolase in preference to the human homologue will likely be difficult. It has been proposed that disrupting enolase-plasminogen complex formation may be an innovative approach to delay disease progression. This approach will require intimate knowledge of the mechanism of complex formation and an accurate model of this protein-protein interaction. Although the structure of human plasminogen has been solved previously (PDB: 4DUU), there has been an absence of available structural information of enolase from a pathogenic fungus (52). In order to address this gap in knowledge, the X-ray crystal structure of *A. fumigatus* enolase was solved with the intention of developing a structural model of tshe enolase and human plasminogen complex (discussed in Chapter 5).

6.6 Future directions - next steps in the drug discovery pipeline

The availability of structural information of GMP synthase and M2DH, and the proposed structural model of the enolase-plasminogen complex, provides a strong foundation to pursue the next steps in the drug discovery pipeline. For GMP synthase and M2DH, there should be a focus on the identification of small molecule compounds that can modulate enzyme activity, whereas targeting enolase must focus on disruption of complex formation by targeting the protein-protein interface.

Both GMP synthase and M2DH projects are in the early stages of drug screening and so we have only explored a small subset of chemical compounds in the *in silico* screening experiments. It may beneficial to widen the search to include a larger sampling of different chemical scaffolds by using multiple compound libraries from the ZINC15 database (106). This database offers over 230 million compounds prepared for virtual screening, which can be

acquired and inputted into established high-throughput *in silico* screening methods. In addition, these compounds and their structural analogues are readily available for purchase to provide an efficient pathway from virtual screening to *in vitro* testing of hit compounds. Targeting alternative binding pockets in these enzymes may also be considered. This is particularly relevant for GMP synthase, which has two catalytic domains – the GATase and ATPPase domains. The AtomWise virtual screening focused on the ATP-binding site of the ATPPase domain but the XMP-binding site is also a viable option.

Fragment-based drug discovery has become an emerging approach to the generation of drug leads. These fragments tend to obey the 'rule of three', a collection of properties in which they comply to, including: (i) molecular weight of <300 Da, (ii) number of hydrogen bond donors <3, (iii) number of hydrogen bond acceptors <3 and (iv) the partition coefficient, a measure of lypophilicity, (logP) <3 (107). Due to their small size, hits identified in fragment screening can be further elaborated and chemically optimised. This is an advantage to chemical compounds that tend to be larger, bulkier and subject to steric limitations. However, these fragments tend to bind with low affinity and therefore, a powerful biophysical method of detection is required for screening.

Approaches to improve the efficiency of the screening process must be considered to validate hits from *in silico* screens of compound libraries as well as fragment screens. The *in vitro* enzyme assays used in the GMP synthase and M2DH projects were sufficient for screening small subsets of compounds (<100 compounds), but they will not suitable be for high-throughput screening. Instead, it would be highly advantageous to use a method to automate this screening process such as surface-plasmon resonance (SPR). Initially, a robust method for the immobilisation of GMP synthase and M2DH onto the sensor chip and conditions to facilitate effective screening must be established. Then, compound libraries can be prepared and flowed across the sensor chip where binding of those small molecules are detected. Due to the sensitivity of this technique, it is also applicable to fragment screening (108). The immobilisation of *A. fumigatus* M2DH, GMP synthase and a truncated version of GMP synthase that only contained the ATPPase domain (residues 211-557) onto a CM5 sensor chip was attempted using an amine coupling kit (Cytiva). When known substrates were flowed over the immobilised protein, there was no clear indication of binding, which can indicate that the binding sites were inaccessible after immobilisation. Hence, amine coupling does not appear to be a suitable method and alternative methods of protein capture should be considered such as using a Ni-NTA chip that captures His-tagged proteins or a GST capture kit (Cytiva) for the immobilisation of GST-tagged proteins (109). The GMP synthase and M2DH may be expressed with their His-tag intact or the coding sequences may be sub-cloned into suitable vectors to produce the GST-tagged proteins. This will increase the possible options available for protein

immobilisation and a suitable method that improves the accessibility of active sites can be determined.

Disruption of enolase-plasminogen complex formation is a more complicated process. Protein-protein interfaces are notoriously difficult to target since the interface tends to (i) expand across a large and hydrophobic surface area, (ii) contain few targetable grooves or pockets for small molecules to bind and (iii) exhibit high binding affinity between the two proteins, which can be difficult for small molecules to overcome (110). However, it has been shown that amino acid residues on the interface surface, known as 'hot-spots', can contribute disproportionately to the binding free energy and so these hot-spot regions should be the focus for targeting of the protein-protein interface (111). For these reasons, determining an accurate representation of the enolase-plasminogen complex is necessary to identify hot spots and define the target site for high-throughput *in silico* screening of compound and fragment libraries, as discussed previously.

There are several techniques that can be used in complement to each other to obtain structural information of the complex. X-ray crystallography would be the ideal approach to provide atomic-level resolution data but there are bottlenecks associated with obtaining diffracting crystals of large macromolecular complexes. Small-angle X-ray scattering (SAXS), a solution-based technique, or cryo-electron microscopy (cryo-EM), have been used to visualise large protein complexes (112,113). As the crystal structures of *A. fumigatus* enolase and human plasminogen are now available, they can be docked onto a molecular envelope obtained from SAXS or a cryo-EM map to visualise the relative orientations of each protein within the complex (114).

6.7 Conclusions

The foundational work from this thesis has provided the essential structural and functional characterisation of metabolic enzymes that are crucial to fungal viability and virulence. The pivotal next steps in the drug discovery pipeline will need to focus on the development of robust and high-throughput methods of compound and fragment screening that are suitable to these targets and the methods used to perform *in vitro* screening must be streamlined. Ultimately, these processes will be vital to progress towards the development of inhibitors against *A. fumigatus* GMP synthase, M2DH and enolase.

Chapter 7:

References

1. Benedict, K., Jackson, B. R., Chiller, T., and Beer, K. D. (2019) Estimation of direct healthcare costs of fungal diseases in the United States. *Clin Infect Dis.* **68**, 1791-1797
2. Nucci, M., and Marr, K. A. (2005) Emerging fungal diseases. *Clin Infect Dis.* **41**, 521-526
3. Bongomin, F., Gago, S., Oladele, R. O., and Denning, D. W. (2017) Global and multi-national prevalence of fungal disease - estimate precision. *J. Fungi* **3**, 57
4. Walsh, T. J., and Dixon, D. M. (1996) Spectrum of mycoses. in *Medical Microbiology* (Baron, S. ed.), 4th Ed., University of Texas Medical Branch at Galveston, Galveston, Texas. pp
5. Vandeputte, P., Ferrari, S., and Coste, A. T. (2012) Antifungal resistance and new strategies to control fungal infections. *Int. J. Microbiol.* **2012**
6. Brown, G. D., Denning, D. W., Gow, N. A. R., Levitz, S. M., Netea, M. G., and White, T. C. (2012) Hidden killers: human fungal infections. *Med. Mycol.* **4**, 165rv113
7. Latgé, J. P. (1999) *Aspergillus fumigatus* and aspergillosis. *Clin Microbiol Rev* **12**, 310-350
8. Schmiedel, Y., and Zimmerli, S. (2016) Common invasive fungal diseases: an overview of invasive candidiasis, aspergillosis, cryptococcosis, and *Pneumocystis* pneumonia. *Swiss Med Wkly* **14**, w14281
9. Shi, M., and Mody, C. H. (2016) Fungal infection in the brain: What we learned from intravital imaging. *Front Immunol.* **7**, 292
10. Nucci, M., and Perfect, J. R. (2008) When primary antifungal therapy fails. *Clin Infect Dis.* **46**, 1426-1433
11. Tekaia, F., and Latgé, J. P. (2005) *Aspergillus fumigatus*: saprophyte or pathogen? *Curr Opin Microbiol* **8**, 385-392
12. Kwon-Chung, K. J., and Sugui, J. A. (2013) *Aspergillus fumigatus* - what makes the species a ubiquitous human fungal pathogen? *PLoS Pathog.* **9**, e1003743
13. van de Veerdonk, F. L., Gresnight, M. S., Romani, L., Netea, M. G., and Latgé, J. P. (2017) *Aspergillus fumigatus* morphology and dynamic host interactions. *Nat Rev Microbiol.* **15**, 661-674
14. Dagenais, T. R. T., and Keller, N. P. (2009) Pathogenesis of *Aspergillus fumigatus* in invasive aspergillosis. *Clin. Microbiol. Rev.* **22**, 447-465
15. Mazu, T. K., Bricker, B. A., Flores-Rozas, H., and Ablordeppey, S. Y. (2016) The mechanistic targets of antifungal agents: An overview. *Mini Rev Med Chem.* **16**, 555-578
16. Gooday, G. W. (1995) Cell Membrane. in *The Growing Fungus* (Gow, N. A. R., and Gadd, G. M. eds.), Springer, Dordrecht. pp 63-74
17. Ryder, N. S. (1988) Mechanism of action and biochemical selectivity of allylamine antimycotic agents. *Ann N Y Acad Sci.* **544**, 208-220
18. Allen, D., Wilson, D., Drew, R., and Perfect, J. R. (2015) Azole antifungals: 35 years of invasive fungal infection management. *Expert Rev. Anti Infect. Ther.* **13**, 787-798
19. Ghannoum, M. A., and Rice, L. B. (1999) Antifungal agents: mode of action, mechanisms of resistance, and correlation of these mechanisms with bacterial resistance. *Clin Microbiol Rev* **12**, 510-517
20. Ermishkin, L. N., Kasumov, K. M., and Potzeluyev, V. M. (1976) Single ionic channels induced in lipid bilayers by polyene antibiotics amphotericin B and nystatine. *Nature* **262**, 698
21. Gow, N. A. R., Latgé, J. P., and Munro, C. A. (2017) The fungal cell wall: Structure, biosynthesis, and function. *Microbiol Spectr* **5**
22. Wiederhold, N. P., and Lewis, R. E. (2003) The echinocandin antifungals: an overview of the pharmacology, spectrum and clinical efficacy. *Expert Opin Investig Drugs* **12**, 1313-1333
23. Brauer, V. S., Rezende, C. P., Pessoni, A. M., de Paula, R. G., Rangappa, K. S., Nayaka, S. C., Gupta, V. K., and Almeida, F. (2019) Antifungal agents in agriculture: friends and foes of public health. *Biomolecules* **9**, 521

24. Azevedo, M.-M., Faria-Ramos, I., Cruz, L. C., Pina-Vaz, C., and Rodrigues, A. G. (2015) Genesis of azole antifungal resistance from agriculture to clinical settings. *J. Agric. Food Chem* **63**, 7463-7468
25. Hof, H. (2001) Critical annotations to the use of azole antifungals for plant protection. *Antimicrob Agents Chemother* **45**, 2987-2990
26. Hope, W., Morton, A., and Eisen, D. P. (2002) Increase in prevalence of nosocomial non-*Candida albicans* candidaemia and the association of *Candida krusei* with fluconazole use. *J Hosp Infect* **50**, 56-65
27. Pfaller, M. A., Diekema, D. J., Mendez, M., Kibbler, C., Erzsebet, P., Chang, S.-C., Gibbs, D. L., and Newell, V. A. (2006) *Candida guilliermondii*, an opportunist fungal pathogen with decrease susceptibility to fluconazole: geographic and temporal trends from the ARTEMIS DISK antifungal surveillance program. *J Clin Microbiol* **44**, 3551-3556
28. Perlin, D. S., Shor, E., and Zhao, Y. (2015) Update on antifungal drug resistance. *Curr Clin Microbiol Rep* **2**, 84-95
29. Bueid, A., Howard, S. J., Moore, C. B., Richardson, M. D., Harrison, E., Bowyer, P., and Denning, D. W. (2010) Azole antifungal resistance in *Aspergillus fumigatus*: 2008 and 2009. *J Antimicrob Chemother* **65**, 2116-2118
30. Lestrade P., Buil, J. B., van der Beek, M. T., Kuijper, E. J., van Dijk, K., Kampinga, G. A., Rjinders, B. J. A., Vonk, A. G., de Greeff, S. C., Schoffelen, A. F., van Dissel, J., Meis, J. F., Melchers, W. J. G., and Verweij, P. E. (2020) Paradoxal trends in azole-resistant *Aspergillus fumigatus* in a national multicenter surveillance program, the Netherlands, 2013–2018. *Emerg Infect Dis* **26**, 1447-1455
31. Marshall, A. C., Bond, C. S., and Bruning, J. B. (2018) Structure of *Aspergillus fumigatus* cytosolic thiolase: Trapped tetrahedral reaction intermediates and activation by monovalent cations. *ACS Catal.* **8**, 1973-1989
32. Hata, M., Ishii, Y., Watanabe, E., Uoto, K., Kobayashi, S., Yoshida, K. I., Otani, T., and Ando, A. (2010) Inhibition of ergosterol synthesis by novel antifungal compounds targeting C-14 reductase. *Med Mycol* **48**, 613-621
33. Urbina, J. M., Cortés, J. C. G., Palma, A., López, S. N., Zacchino, S. A., Enriz, R. D., Ribas, J. C., and Kouznetzov, V. V. (2000) Inhibitors of the fungal cell wall. Synthesis of 4-aryl-4-N-arylamine-1-butenes and related compounds with inhibitory activities on β(1–3) glucan and chitin synthases. *Bioorganic Med. Chem.* **8**, 691-698
34. Marshall, A. C., Kroker, A. J., Murray, L. A. M., Gronthos, K., Rajapaksha, H., Wegener, K. L., and Bruning, J. B. (2017) Structure of the sliding clamp from the fungal pathogen *Aspergillus fumigatus* (AfumPCNA) and interactions with Human p21. *FEBS J* **284**, 985-1002
35. Marshall, A. C., Kidd, S. E., Lamont-Friedrich, S. J., Arentz, G., Hoffman, P., Coad, B. R., and Bruning, J. B. (2019) Structure, mechanism, and inhibition of *Aspergillus fumigatus* thioredoxin reductase. *Antimicrob. Agents Chemother.* **63**, e02281-02218
36. Kummari, L. K., Butler, M. S., Furlong, E., Blundell, R., Nouwens, A., Silva, A. B., Kappler, U., Fraser, J. A., Kobe, B., Cooper, M. A., and Robertson, A. A. B. (2018) Antifungal benzo[b]thiophene 1,1-dioxide IMPDH inhibitors exhibit pan-assay interference (PAINS) profiles. *Bioorganic Med. Chem.* **26**, 5408-5419
37. Rodriguez-Suarez, R., Xu, D., Veillette, K., Davison, J., Sillaots, S., Kauffman, S., Hu, W., Bowman, J., Martel, N., Trosok, S., Wang, H., Zhang, L., Huang, L., Li, Y., Rahkoodaee, F., Ransom, T., Gauvin, D., Douglas, C., Youngman, P., Becker, J., Jiang, B., and Roemer, T. (2007) Mechanism-of-action determination of GMP synthase inhibitors and target validation in *Candida albicans* and *Aspergillus fumigatus*. *Chem Biol* **14**, 1163-1175
38. Abad, A., Fernández-Molina, J. M., Bikandi, J., Ramírez, A., Margareto, J., Sendino, J., Hernando, F. L., Pontón, J., Garaizar, J., and Rementeria, A. (2010) What makes *Aspergillus fumigatus* a successful pathogen? Genes and molecules involved in invasive aspergillosis. *Rev Iberoam Micol.* **27**, 155-182

39. Ji, H., Wang, J., Guo, J., Li, Y., Lian, S., Guo, W., Yang, H., Kong, F., Zhen, L., Guo, L., and Liu, Y. (2016) Progress in the biological function of alpha-enolase. *Anim. Nutr.* **2**, 12-17
40. de Backer, M. D., Nelissen, B., Logghe, M., Viaene, J., Loonen, I., Vandoninck, S., de Hoogt, R., Dewaele, S., Simons, F. A., Verhasselt, P., Vanhoof, G., Contreras, R., and Luyten, W. H. (2001) An antisense-based functional genomics approach for identification of genes critical for growth of *Candida albicans*. *Nat Biotechnol* **19**, 235-241
41. Ko, H.-C., Hsiao, T.-Y., Chen, C.-T., and Yang, Y.-L. (2013) *Candida albicans* ENO1 null mutants exhibited altered drug susceptibility, hyphal formation, and virulence. *J Microbiol* **51**, 345-351
42. Pancholi, V., and Fischetti, V. A. (1998) Alpha-enolase, a novel strong plasmin(ogen) binding protein on the surface of pathogenic *Streptococci*. *J. Biol. Chem.* **273**, 14503-14515
43. Edwards, S. R., Braley, R., and Chaffin, W. L. (1999) Enolase is present in the cell wall of *Saccharomyces cerevisiae*. *FEMS Microbiol. Lett.* **177**, 211-216
44. Bergmann, S., Rohde, M., Chhatwal, G. S., and Hammerschmidt, S. (2001) α-Enolase of Streptococcus pneumoniae is a plasminogen-binding protein displayed on the bacterial cell surface. *Mol. Microbiol.* **40**, 1273-1287
45. Funk, J., Schaarschmidt, B., Slesiona, S., Hallström, T., Horn, U., and Brock, M. (2016) The glycolytic enzyme enolase represents a plasminogen-binding protein on the surface of a wide variety of medically important fungal species. *Int. J. Med. Microbiol.* **306**, 59-68
46. Miura, N., Kirino, A., Endo, S., Morisaka, H., Kuroda, K., Takagi, M., and Ueda, M. (2012) Tracing putative trafficking of the glycolytic enzyme enolase via SNARE-driven unconventional secretion. *Eukaryot. Cell* **11**, 1075-1082
47. Law, R. H. P., Abu-Ssaydeh, D., and Whisstock, J. C. (2013) New insights into the structure and function of the plasminogen/plasmin system *Curr. Opin. Struct. Biol.* **23**, 836-841
48. Dasari, P., Koleci, N., Shopova, I., A.,, Warternberg, D., Beyersdorf, N., Dietrich, S., Sahagún-Ruiz, A., Figge, M. T., Skerka, C., Brakhage, A. A., and Zipfel, P. F. (2019) Enolase From *Aspergillus fumigatus* is a moonlighting protein that binds the human plasma complement proteins factor H, FHL-1, C4BP, and plasminogen. *Front. Immunol.* **10**, 2573
49. Li, W. Q., Hu, X. C., Zhang, X., Ge, Y., Zhao, S., Hu, Y., and Ashman, R. B. (2011) Immunisation with the glycolytic enzyme enolase confers effective protection against *Candida albicans* infection in mice. *Vaccine* **29**, 5526-5533
50. Shibasaki, S., Karasaki, M., Tafuku, S., Aoki, W., Sewaki, T., and Ueda, M. (2014) Oral immunization against candidiasis using *Lactobacillus casei* displaying enolase 1 from *Candida albicans*. *Sci Pharm* **82**, 697-708
51. Stec, B., and Lebioda, L. (1990) Refined structure of yeast apo-enolase at 2.25 Å resolution. *J Mol Biol* **211**, 235-248
52. Law, R. H. P., Caradoc-Davies, T., Cowieson, N. P., Horvath, A. J., Quek, A. J., Encanacao, J. A., Steer, D., Cowan, A., Zhang, Q., Lu, B. G., Pike, R. N., Smith, A. I., Coughlin, P. B., and Whisstock, J. C. (2012) The X-ray crystal structure of full-length human plasminogen. *Cell Rep.* **1**, 185-190
53. Bergmann, S., Wild, D., Diekmann, O., Frank, R., Bracht, D., Chhatwal, G. S., and Hammerschmidt, S. (2003) Identification of a novel plasmin(ogen)-binding motif in surface displayed α-enolase of *Streptococcus pneumoniae*. *Mol. Microbiol.* **49**, 411-423
54. Ehinger, S., Schubert, W.-D., Bergmann, S., Hammerschmidt, S., and Heinz, D. W. (2004) Plasmin(ogen)-binding alpha-enolase from *Streptococcus pneumoniae*: crystal structure and evaluation of plasming(ogen)-binding sites. *J. Mol. Biol.* **343**, 997-1005
55. Duquerroy, S., Camus, C., and Janin, J. (1995) X-ray structure and catalytic mechanism of lobster enolase. *Biochemistry* **34**, 12513-12523

56. Kuhnel, K., and Luisi, B. (2001) Crystal structure of the *Escherichia coli* RNA degradosome component enolase. *J. Mol. Biol.* **313**, 583
57. Hosaka, T., Meguro, T., Yamato, I., and Shirakihara, Y. (2003) Crystal Structure of *Enterococcus hirae* enolase at 2.8 Å resolution. *J Biochem* **133**, 817-823
58. Kang, H. J., Jung, S. K., Kim, S. J., and Chung, S. J. (2008) Structure of human alpha-enolase (hENO1), a multifunctional glycolytic enzyme. *Acta Crystallogr. D Biol. Crystallogr.* **64**, 651-657
59. López-Alemany, R., Longstaff, C., Hawley, S., Mirshahi, M., Fábregas, P., Jardí, M., Merton, E., Miles, L. A., and Félez, J. (2003) Inhibition of cell surface mediated plasminogen activation by a monoclonal antibody against alpha-Enolase. *Am J Hematol* **72**, 234-242
60. Meena, M., Prasad, V., Zehra, A., Gupta, V. K., and Upadhyay, R. S. (2015) Mannitol metabolism during pathogenic fungal-host interactions under stressed conditions. *Front. Microbiol.* **6**, 1019
61. Kozel, T. R., and Wickes, B. (2014) Fungal diagnostics. *Cold Spring Harb Perspect Med* **4**, a019299
62. Wong, B., Perfect, J. R., Beggs, S., and Wright, K. A. (1990) Production of the hexitol D-mannitol by *Cryptococcus neoformans in vitro* and in rabbits with experimental meningitis. *Infect. Immun.* **58**, 1664-1670
63. Birkinshaw, J. H., Charles, J. H. V., Hetherington, A. C., and Raistrick, H. (1931) On the production of mannitol from glucose by species of *Aspergillus*. *Philos Trans R Soc B* **22**, 153-171
64. Wong, B., Brauer, K. L., Tsai, R. R., and Jayasimhulu, K. (1989) Increased amounts of the *Aspergillus* metabolite D-mannitol in serum of rats with experimental aspergillosis. *J. Infect. Dis.* **160**, 95-103
65. Megson, G. M., Stevens, D. A., Hamilton, J. R., and Denning, D. W. (1996) D-Mannitol in cerebrospinal fluid of patients with AIDS and cryptococcal meningitis. *J Clin Microbiol* **34**, 218-221
66. Cooney, N. M., and Klein, B. S. (2008) Fungal adaptation to the mammalian host: it's a new world, after all. *Curr Opin Microbiol* **11**, 511-516
67. Khanna, N., Stuehler, C., Lünemann, A., Wójtowicz, A., Bochud, P. Y., and Leibundgut-Landmann, S. (2016) Host response to fungal infections - how immunology and host genetics could help to identify and treat patients at risk. *Swiss Med Wkly* **146**, w14350
68. Chaturvedi, V., Flynn, T., Neihaus, W. G., and Wong, B. (1996) Stress tolerance and pathogenic potential of a mannitol mutant of *Cryptococcus neoformans*. *Microbiology* **142**
69. Liu, Q., Ying, S. H., Feng, M. G., and Jiang, X. H. (2008) Physiological implication of intracellular trehalose and mannitol changes in response of entomopathogenic fungus *Beauveria bassiana* to thermal stress. *Antonie Van Leeuwenhoek* **95**, 65-75
70. Chaturvedi, V., Wong, B., and Newman, S. J. (1996) Oxidative killing of *Cryptococcus neoformans* by human neutrophils. Evidence that fungal mannitol protects by scavenging reactive oxygen intermediates. *J. Immunol.* **156**, 3836-3840
71. Ruijter, G. J. G., Bax, M., Patel, H., Flitter, S. J., van de Vondervoort, P. J. I., de Vries, R. P., van Kuyk, P. A., and Visser, J. (2003) Mannitol is required for stress tolerance in *Aspergillus niger* conidiospores. *Eukaryot. Cell* **2**, 690-698
72. Wyatt, T. T., van Leeuwen, M. R., Wösten, H. A. B., and Dijksterhuis, J. (2014) Mannitol is essential for the development of stress-resistance ascospores in *Neosartorya fischeri* (*Aspergillus fischeri*). *Fungal Genet. Biol.* **64**, 11-24
73. Horikoshi, K., Iida, S., and Ikeda, Y. (1965) Mannitol and mannitol dehydrogenases in conidia of *Aspergillus oryzae*. *J Bacteriol* **89**, 326-330
74. Corina, D. L., and Munday, K. A. (1971) Studies on polyol function in *Aspergillus clavatus*: a role for mannitol and ribitol. *J Gen Microbiol* **69**, 221-227

75. Witteveen, C. F. B., and Visser, J. (1995) Polyol pools in *Aspergillus niger*. *FEMS Microbiol Lett.* **134**, 57-62
76. Thammahong, A., Puttikamonkul, S., Perfect, J. R., Brennan, R. G., and Cramer, R. A. (2017) Central role of the trehalose biosynthesis pathway in the pathogenesis of human fungal infections: opportunities and challenges for therapeutic development. *Microbiol Mol Biol Rev* **81**, e00053-00016
77. Chitty, J. L., and Fraser, J. A. (2017) Purine acquisition and synthesis by human fungal pathogens. *Microorganisms* **5**, E33
78. Park, Y., Pacitto, A., Bayliss, T., Cleghorn, L. A., Wang, Z., Hartman, T., Arora, K., Ioerger, T. R., Sacchettini, J., Rizzi, M., Donini, S., Blundell, T. L., Ascher, D. B., Rhee, K., Breda, A., Zhou, N., Dartois, V., Jonnala, S. R., Via, L. E., Mizrahi, V., Epemolu, O., Stojanovski, L., Simeons, F., Osuna-Cabello, M., Ellis, L., MacKenzie, C. J., Smith, A. R., Davis, S. H., Murugesan, D., Buchanan, K. I., Turner, P. A., Huggett, M., Zuccotto, F., Rebollo-Lopez, M. J., Lafuente-Monasterio, M. J., Sanz, O., Diaz, G. S., Lelièvre, J., Ballell, L., Selenski, C., Axtman, M., Ghidelli-Disse, S., Pflaumer, H., Bösche, M., Drewes, G., Freiberg, G. M., Kurnick, M. D., Srikumaran, M., Kempf, D. J., Green, S. R., Ray, P. C., Read, K., Wyatt, P., Barry, C. E., Boshoff, H. I. (2018) Essential but not vulnerable: indazole sulfonamides targeting inosine monophosphate dehydrogenase as potential leads against *Mycobacterium tuberculosis*. *ACS Infect Dis.* **3**, 18-33
79. Shu, Q., and Nair, V. (2008) Inosine monophosphate dehydrogenase (IMPDH) as a target in drug discovery. *Med Res Rev* **28**, 219-232
80. Trapero, A., Pacitto, A., Singh, V., Sabbah, M., Coyne, A. G., Mizrahi, V., Blundell, T. L., Ascher, D. B., and Abell, C. (2018) Fragment-based approach to targeting inosine-5'-monophosphate dehydrogenase (IMPDH) from *Mycobacterium tuberculosis*. *J Med Chem* **61**, 2806-2822
81. Hedstrom, L. (2009) IMP dehydrogenase: Structure, mechanism and inhibition. *Chem Rev* **109**, 2903-2928
82. Markland, W., McQuaid, T. J., Jain, J., and Kwong, A. D. (2000) Broad-spectrum antiviral activity of the IMP dehydrogenase inhibitor VX-497: a comparison with Ribavirin and demonstration of antiviral additivity with alpha interferon. *Antimicrob Agents Chemother* **44**, 859-866
83. Floryk, D., and Thompson, T. C. (2010) Antiproliferative effects of AVN944, a novel inosine 5-monophosphate dehydrogenase inhibitor, in prostate cancer cells. *Int J Cancer* **123**, 2294-2302
84. Morrow, C. A., Valkov, E., Stamp, A., Chow, E. W. L., Lee, I. R., Wronski, A., Williams, S. J., Hill, J. M., Djordjevic, J. T., Kappler, U., Kobe, B., and Fraser, J. A. (2012) *De novo* GTP biosynthesis is citical for virulence of the fungal pathogen *Cryptococcus neoformans*. *PLoS Pathog.* **8**, e1002957
85. Chitty, J. L., Tatzenko, T. L., Williams, S. J., Koh, Y. Q., Corfield, E. C., Butrler, M. S., Roberston, A. A., Cooper, M. A., Kappler, U., Kobe, B., and Fraser, J. A. (2017) GMP synthase is required for virulence factor production and infection by *Cryptococcus neoformans*. *J Biol Chem* **292**, 3049-3059
86. Nguyen, S., Jovcevski, B., Pukala, T. L., and Bruning, J. B. (2020) Nucleoside selectivity of *Aspergillus fumigatus* nucleoside-diphosphate kinase. *FEBS J.* **288**, 2398-2417
87. Lin, X., Momany, C., and Momany, M. (2003) SwoHp, a nucleoside diphosphate kinase, is essential in *Aspergillus nidulans*. *Eukaryot Cell* **2**, 1169-1177
88. Wang, Y., Wang, S., Nie, X., Yang, K., Xu, P., Wang, X., Liu, M., Yang, Y., Chen, Z., and Wang, S. (2019) Molecular and structural basis of nucleoside diphosphate kinase–mediated regulation of spore and sclerotia development in the fungus *Aspergillus flavus*. *J Biol Chem* **294**, 12415-12431

89. Dinamarco, T. M., Brown, N. A., Couto de Almeida, R. S., Alves de Castro, P., Savoldi, M., de Souza Goldman, M. H., and Goldman, G. H. (2012) *Aspergillus fumigatus* calcineurin interacts with a nucleoside diphosphate kinase. *Microbes Infect* **14**, 922-929
90. Kilstrup, M., Hammer, K., Jensen, P. R., and Martinussen, J. (2005) Nucleotide metabolism and its control in lactic acid bacteria. *FEMS Microbiol Rev* **29**, 555-590
91. Sperling, O. (1988) Human Purine Metabolism. in *Myocardial Energy Metabolism* (De Jong, J. W. ed.), Springer, Dordrecht. pp 225-236
92. Risal, D., Strickler, M. D., and Goldstein, B. M. (2003) Crystal structure of the human type I inosine monophosphate dehydrogenase and implications for isoform specificity. Binary complex of human type-I inosine monophosphate dehydrogenase with 6-CL-IMP.
93. Colby, T. D., Vanderveen, K., Strickler, M. D., Markham, G. D., and Goldstein, B. M. (1999) Crystal structure of human type II inosine monophosphate dehydrogenase: Implications for ligand binding and drug design. *Proc. Natl. Acad. Sci. U.S.A.* **96**, 3531-3536
94. Welin, M., Lehtio, L., Johansson, A., Flodin, S., Nyman, T., Tresaugues, L., Hammarstrom, M., Graslund, S., and Nordlund, P. (2013) Substrate specificty and oligomerization of human GMP synthetase. *J Mol Biol* **425**, 4323
95. Chen, Y., Gallois-Montbrun, S., Schneider, B., Veron, M., Morera, S., Deville-Bonne, D., and Janin, J. (2003) Nucleotide binding to nucleoside diphosphate kinases: X-ray structure of human NDPK-A in complex with ADP and comparison to protein kinases. *J Mol Biol* **332**, 915-926
96. Wang, Y., and Wang, S. H. (2020) Molecular and structural basis of nucleoside diphosphate kinase regulating the spores and sclerotia development in *Aspergillus flavus*. The crystal of nucleoside diphosphate kinase from *Aspergillus flavus*.
97. Umejiego, N. N., Gollapalli, D., Sharling, L., Volftsun, A., Lu, J., Benjamin, N. N., Stroupe, A. H., Riera, T. V., Striepen, B., and Hedstrom, L. (2008) Targeting a prokaryotic protein in a eukaryotic pathogen: identification of lead compounds against *Cryptosporidiosis*. *Chem Biol* **15**, 70-77
98. Chittur, S. V., Klem, T. J., Shafer, C. M., and Davisson, V. J. (2001) Mechanism for acivicin inactivation of triad glutamine amidotransferases. *Biochemistry* **40**, 876-887
99. Ahluwalia, G. S., Grem, J. L., Hao, Z., and Cooney, D. A. (1990) Metabolism and action of amino acid analog anti-cancer agents. *Pharmacol Therapeut.* **46**, 243-271
100. Giraud, M. F., Georgescauld, F., Lascu, I., and Dautant, A. (2006) Crystal structures of S120G mutant and wild type of human nucleoside diphosphate kinase A in complex with ADP. *J Bioenerg. Biomembr.* **38**, 261-264
101. Furman, P. A., Fyfe, J. A., St Clair, M. H., Weinhold, K., Rideout, J. L., Greeman, G. A., Lehrman, S. N., Bolognesi, D. P., Broder, S., and Mitsuya, H. (1986) Phosphorylation of 3'-azido-3'-deoxythymidine and selective interaction of the 5'-triphosphate with human immunodeficiency virus reverse transcriptase. *Proc. Natl. Acad. Sci. U.S.A.* **83**, 8333-8337
102. Semianrio-Vidal, L., van Hesuden, C., Mugesh, G., and Lazarowski, E. R. (2010) Ebselen is a potent non-competitive inhibitor of extracellular nucleoside diphosphokinase. *Purinerg Signal* **6**, 383-391
103. Nakamura, J., and Lou, L. (1995) Biochemical characterization of human GMP synthetase. *J. Biol. Chem.* **270**, 7347-7353
104. Krahulec, S., Armao, G. C., Klimacek, M., and Nidetzky, B. (2011) Enzymes of mannitol metabolism in the human pathogenic fungus *Aspergillus fumigatus* – kinetic properties of mannitol-1-phosphate 5-dehydrogenase and mannitol 2-dehydrogenase, and their physiological implications. *FEBS J.* **278**, 1264-1276
105. Pancholi, V. (2001) Multifunctional α-enolase: its role in diseases. *Cell Mol. Life Sci.* **58**, 902-920
106. Irwin, J. J., Sterling, T., Mysinger, M. M., Bolstad, E. S., and Coleman, R. G. (2012) ZINC: A free tool to discover chemistry for biology. *J. Chem. Inf. Model* **52**, 1757-1768

107. Kirsch, P., Hartman, A. M., Hirsch, A. K. H., and Empting, M. (2019) Concepts and core priniciples of fragment-based drug design. *Molecules* **24**, 4309
108. Neumann, T., Junker, H.-D., Schmidt, K., and Sekul, R. (2007) SPR-based fragment screening: advantages and applications. *Curr Top Med Chem* **7**, 1630-1642
109. Khan, F., He M., and Taussig, M. J. (2006) Double-hexahistidine tag with high-affinity binding for protein immobilization, purification, and detection on ni-nitrilotriacetic acid surfaces. *Anal. Chem.* **78**, 3072-3079
110. Lu, H., Zhou, Q., He, J., Jiang, Z., Peng, C., Tong, R., and Shi, J. (2020) Recent advances in the development of protein–protein interactions modulators: mechanisms and clinical trials. *Signal Transduct. Tar.* **5**, 213
111. Cukuroglu, E., Engin, H. B., Gursoy, A., and Keskin, O. (2014) Hot spots in protein–protein interfaces: Towards drug discovery. *Prog. Biophys. Mol. Bio* **116**, 165-173
112. Nogales, E., and Scheres, S. H. W. (2015) Cryo-EM: A unique tool for the visualization of macromolecular complexity. *Mol. Cell.* **58**, 677-689
113. Lipfert, J., and Doniach, S. (2007) Small-angle X-ray scattering from RNA, proteins, and protein complexes. *Annu. Rev. Biophys. Biomol. Struct.* **36**, 307-327
114. Duhovny-Schneidman, D., and Hammel, M. (2018) Modeling structure and dynamics of protein complexes with SAXS profiles. *Methods Mol. Biol.* **1764**, 449-473

Appendix A (Publication):

Targeting unconventional pathways in pursuit of novel antifungals

Statement of Authorship

Title of Paper	Targeting unconventional pathways in pursuit of novel antifungals
Publication Status	☑ Published ☐ Accepted for Publication ☐ Submitted for Publication ☐ Unpublished and Unsubmitted work written in manuscript style
Publication Details	This publication discusses several unconventional fungal pathways that could be targeted for the development of novel antifungal therapies.

Principal Author

Name of Principal Author (Candidate)	Stephanie Nguyen		
Contribution to the Paper	Project conceptualisation, prepared, reviewed and edited the manuscript		
Overall percentage (%)	70		
Certification:	This paper reports on original research I conducted during the period of my Higher Degree by Research candidature and is not subject to any obligations or contractual agreements with a third party that would constrain its inclusion in this thesis. I am the primary author of this paper.		
Signature		Date	19/07/2021

Co-Author Contributions

By signing the Statement of Authorship, each author certifies that:

i. the candidate's stated contribution to the publication is accurate (as detailed above);

ii. permission is granted for the candidate in include the publication in the thesis; and

iii. the sum of all co-author contributions is equal to 100% less the candidate's stated contribution.

Name of Co-Author	Jia Quyen Truong		
Contribution to the Paper	Project conceptualisation, prepared, reviewed and edited the manuscript		
Overall Percentage (%)	20		
Signature		Date	19/07/2021

Name of Co-Author	Dr. John B. Bruning
Contribution to the Paper	Reviewed and edited the manuscript.
Overall Percentage (%)	10
Signature	

Date: 21/7/21

frontiers in Molecular Biosciences

REVIEW
published: 12 January 2021
doi: 10.3389/fmolb.2020.621366

Targeting Unconventional Pathways in Pursuit of Novel Antifungals

Stephanie Nguyen[1], *Jia Q. Truong*[2] *and John B. Bruning*[1]*

[1] Institute of Photonics and Advanced Sensing (IPAS), School of Biological Sciences, The University of Adelaide, Adelaide, SA, Australia, [2] School of Biological Sciences, The University of Adelaide, Adelaide, SA, Australia

The impact of invasive fungal infections on human health is a serious, but largely overlooked, public health issue. Commonly affecting the immunocompromised community, fungal infections are predominantly caused by species of *Candida*, *Cryptococcus*, and *Aspergillus*. Treatments are reliant on the aggressive use of pre-existing antifungal drug classes that target the fungal cell wall and membrane. Despite their frequent use, these drugs are subject to unfavorable drug-drug interactions, can cause undesirable side-effects and have compromised efficacy due to the emergence of antifungal resistance. Hence, there is a clear need to develop novel classes of antifungal drugs. A promising approach involves exploiting the metabolic needs of fungi by targeted interruption of essential metabolic pathways. This review highlights potential antifungal targets including enolase, a component of the enolase-plasminogen complex, and enzymes from the mannitol biosynthesis and purine nucleotide biosynthesis pathways. There has been increased interest in the enzymes that comprise these particular pathways and further investigation into their merits as antifungal targets and roles in fungal survival and virulence are warranted. Disruption of these vital processes by targeting unconventional pathways with small molecules or antibodies may serve as a promising approach to discovering novel classes of antifungals.

Keywords: ***Aspergillus fumigatus***, ***Candida albicans***, ***Cryptococcus neoformans***, **drug targets, drug discovery, antifungal**

OPEN ACCESS

Edited by:
Amit Prasad,
Indian Institute of Technology Mandi, India

Reviewed by:
Praveen Rao Juvvadi,
Duke University, United States
Vicente de Paulo Martins,
University of Brasilia, Brazil

***Correspondence:**
John B. Bruning
john.bruning@adelaide.edu.au

Specialty section:
This article was submitted to Molecular Diagnostics and Therapeutics, a section of the journal Frontiers in Molecular Biosciences

Received: *26 October 2020*
Accepted: *11 December 2020*
Published: *12 January 2021*

Citation:
Nguyen S, Truong JQ and Bruning JB (2021) Targeting Unconventional Pathways in Pursuit of Novel Antifungals. Front. Mol. Biosci. 7:621366. doi: 10.3389/fmolb.2020.621366

INTRODUCTION

Opportunistic fungal pathogens impose a significant societal and economic burden on the public health sector, particularly amongst immunocompromised patients. In the United States alone, it has been estimated that fungal infections incur hospitalization and outpatient costs that exceed $7.2 billion USD annually (Benedict et al., 2019). These costs are expected to rise in tandem with the expansion of the immunocompromised population and subsequent increase in the incidence of invasive mycoses. This predisposed population includes recipients of invasive surgery, chemotherapy, and immunosuppressive therapy, as well as sufferers of human immunodeficiency virus (HIV)/acquired immunodeficiency syndrome (AIDS) and cystic fibrosis (Nucci and Marr, 2005).

Global estimates suggest that over 150 million individuals are affected by serious fungal infections which culminate in over 1.6 million deaths per annum (Bongomin et al., 2017). Common fungal pathogens associated with these opportunistic infections include *Aspergillus* spp. (most commonly *Aspergillus fumigatus*), *Candida albicans* (and other non-*albicans Candida* species),

Cryptococcus neoformans, *Histoplasma capsulatum*, *Coccidioides imitis*, and *Pneumocystis jirovecii* (Walsh and Dixon, 1996; Vandeputte et al., 2012). Amongst these fungal pathogens, species of *Aspergillus*, *Candida, Pneumocystis*, and *Cryptococcus* are responsible for over 90% of total global fungal mortalities and therefore, these fungal species pose the most prominent threat to human health (Brown et al., 2012). Aspergillosis, candidiasis, and cryptococcosis often manifest as deep tissue mycoses that occupy different niches within the mammalian host. *A. fumigatus* preferentially colonizes lung tissue, resulting in pulmonary aspergillosis whereas *C. albicans* primarily infects the blood, resulting in systemic candidiasis, and *C. neoformans* establishes an initial infection in the lungs, causing pulmonary cryptococcosis (Latgé, 1999; Wenzel and Gennings, 2005; Sabiiti and May, 2012). The severity of these diseases can escalate rapidly as the infection disseminates throughout the body and affects multiple organs. Infections caused by *A. fumigatus* and *C. albicans* can manifest as invasive aspergillosis and invasive candidiasis, respectively, both of which target the kidney and brain, whereas *C. neoformans* can cause cryptococcal meningitis which affects the central nervous system (Bicanic and Harrison, 2005; Schmiedel and Zimmerli, 2016; Shi and Mody, 2016).

In this advanced stage, treatment of invasive mycoses requires aggressive and expensive antifungal therapy. However, the success of treatment is often impeded by (i) fundamental issues in the diagnostic stage, (ii) unfavorable characteristics that are inherent to pre-existing antifungal drugs or the (iii) emergence of antifungal resistance, all of which ultimately lead to primary antifungal therapy failure (Nucci and Perfect, 2008). In patients suffering from invasive candidiasis or invasive aspergillosis, the rate of failure can be as high as 60 and 70%, respectively (Nucci and Perfect, 2008). To address these issues, there have been advancements in diagnostic techniques and several iterations of pre-existing drugs have been developed to improve their pharmacological properties (Houšt et al., 2020; Kidd et al., 2020). However, new classes of antifungal drugs that bypass existing resistance mechanisms by targeting alternate pathways are yet to be discovered.

In spite of the alarming rates of morbidity and mortality, the severity of invasive fungal infections remains underappreciated. Increases in disease incidence and prevalence of antifungal resistance highlights the need to identify novel targets and develop new classes of antifungals to manage mycoses amongst the immunocompromised population. There have been continual efforts to characterize enzymes involved in the biosynthesis of ergosterol or cell wall components, both of which are classic antifungal targets, to develop novel inhibitors (Urbina et al., 2000; Hata et al., 2010; Marshall et al., 2018). However, there has also been a notable shift in focus from these pathways exclusive to fungi to exploiting species-specific differences in shared pathways between fungi and humans (Rodriguez-Suarez et al., 2007; Marshall et al., 2017, 2019; Kummari et al., 2018). To effectively establish infection, the fungus must adapt to a different niche within the human host, combat, or circumvent the host immune response and obtain sufficient nutrients to reproduce and disseminate. Although these metabolic requirements may differ between fungal species, depending on their preferred infection site, disrupting shared metabolic pathways involved in these processes can impede fungal survival and pathogenesis. Targeting these pathways may present an elegant approach to develop novel classes of therapeutics with broad-spectrum activity.

In this article, we have focused primarily on emerging targets for the development of novel antifungal classes. We have identified key enzymes involved in several targetable metabolic pathways and consolidated extensive research to define their roles in fungal survival and virulence. Furthermore, we have discussed their merits as potential targets and provided practical discussions to drive future drug design efforts from bench to bedside.

CURRENT ANTIFUNGAL CLASSES AND TREATMENTS

Antifungal therapeutics are routinely administered to combat invasive mycoses. The four main classes of antifungals that currently exist function by targeting either the cell wall or cell membrane (**Figure 1**). They include azoles, allylamines, polyenes and echinocandins. The mechanism of action of each antifungal drug class and discussion of their biological targets have been extensively reviewed (Mazu et al., 2016). In this review, we have provided a brief overview of the currently available antifungals.

Targeting the Cell Membrane

Ergosterol is a steroid alcohol that forms a major component of the fungal cell membrane. It is responsible for the maintenance of membrane fluidity and structural integrity, as well as the regulation of membrane permeability (Gooday, 1995). Due to its implications in fungal survival, ergosterol biosynthesis enzymes have been the target of the allylamine and azole classes of antifungals, whereas the polyene class targets ergosterol function directly (Mazu et al., 2016). Allylamines target squalene epoxidase, an enzyme that catalyzes the conversion of squalene to lanosterol whereas azoles function as non-competitive inhibitors of lanosterol 14α-demethylase by complexing with the haem prosthetic group (Ryder, 1988; Allen et al., 2015; Mazu et al., 2016). Both of these enzymes are involved in the later stages of ergosterol biosynthesis. Disruption of this pathway leads to an accumulation of upstream sterol precursors and ultimately, the depletion of ergosterol. This deficiency results in impaired cell membrane function, growth inhibition, and cell death (Ghannoum and Rice, 1999; Allen et al., 2015). In comparison, polyenes induce fungal death by binding directly to ergosterol embedded in the cell membrane. Once incorporated into the membrane, this leads to the formation of pores which increase membrane permeability to water, ions and non-electrolytes (Ermishkin et al., 1976).

Targeting the Cell Wall

Due to the absence of a mammalian equivalent, several components of the fungal cell wall and the biosynthetic enzymes responsible for its construction have been the target for antifungal therapeutics. The cell wall, comprised of three main components including glucans, mannoproteins, and chitin, plays

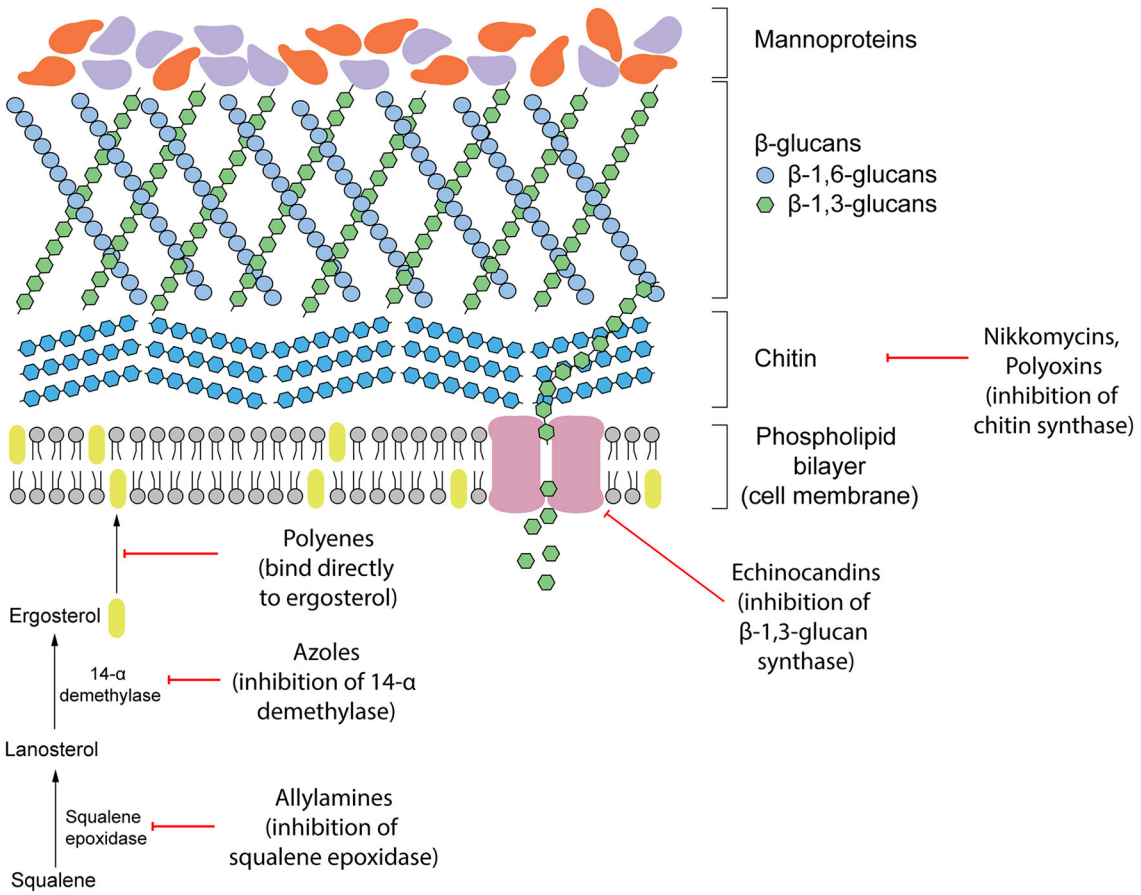


FIGURE 1 | Existing antifungal drug classes that target the biosynthesis of the fungal cell wall components (nikkomycins, polyoxins, echinocandins), biosynthesis of the cell membrane (azoles, allylamines) or bind to ergosterol in the membrane directly (polyenes).

an important role in maintaining fungal survival by maintaining cell rigidity and resisting osmotic stress (Gow et al., 2017). Furthermore, there are several receptors displayed on the surface of the cell wall that can interact with the host during infection and contribute to fungal virulence. Echinocandins disrupt these fundamental functions of the cell wall by non-competitively inhibiting (1,3)-β-glucan synthase, the enzyme responsible for maintaining and synthesizing glucans. Insufficient biosynthesis of glucans leads to destabilization of the cell wall and a compromised ability to resist osmotic pressure that results in cell lysis (Wiederhold and Lewis, 2003).

EMERGENCE OF ANTIFUNGAL RESISTANCE AND THE NEED FOR NOVEL ANTIFUNGAL CLASSES

The extensive use of antifungals in both agricultural and clinical settings has led to the rise of widespread antifungal resistance with significant consequences to food security and human health. To control fungal phytopathogens in agricultural crops, a range of antifungal drugs specifically for agricultural use are used that target mitochondrial respiration processes, including succinate dehydrogenase and Qo inhibitors (Brauer et al., 2019). However, there is also significant overlap in the use of azoles to treat both human and plant fungal pathogens. In particular, they are frequently used in grain- and grass-growing environments, as well as in clinical settings, to control *Candida* spp., *Cryptococcus* spp. and *Aspergillus* spp. infections (Azevedo et al., 2015). As a result, bioactive azoles can persist in soil, water, and fresh produce (Hof, 2001). Ultimately, it has been suggested that these environmental factors contribute to the rise in azole resistance currently being observed in clinical settings.

The prominent use of azoles has resulted in a shift in the prevalence of infections caused by *C. albicans* toward less susceptible species of *Candida*, including *C. glabrata*, *C. krusei*, and *C. guilliermondii* (Hope et al., 2002; Pfaller et al., 2006; Perlin et al., 2015). Fungal strains can also acquire drug resistance during therapy which restricts the treatment options available to the patient and can ultimately cause antifungal therapy failure. Specifically, there are increasing concerns of echinocandin and azole resistance arising in species of *Candida* and *A. fumigatus*, respectively (Perlin, 2015). In the UK, there has been a consistent increase in the proportion of *A. fumigatus* isolates with resistance

to azoles growing from 5% in 2004 to 14% in 2008 and 20% in 2009 (Bueid et al., 2010). Following a similar trend, the frequency of *A. fumigatus* clinical isolates with azole resistance in The Netherlands increased from 7.6% in 2013 to 14.7% in 2018 (Lestrade et al., 2020).

Several generations of antifungal drugs originating from these four classes have been developed to improve their pharmacokinetic profiles, increase potency, and combat resistance. However, treatment of invasive mycoses amongst immunocompromised patients remains challenging. Primary antifungal therapy failure is a prevalent issue, affecting 20–60% of invasive candidiasis patients, 40–60% of invasive aspergillosis patients and 30–100% of invasive fusariosis patients (Nucci and Perfect, 2008). Although the role of underlying host factors and immune status must be considered in antifungal therapy failure, several issues arise from inherent limitations in currently available drugs. Low bioavailability in target tissues, drug toxicity, unfavorable drug-drug interactions, and the emergence of drug resistance also contribute to the reduced efficacy of existing treatments (Nucci and Perfect, 2008).

As the number of antifungal classes available for clinical use is limited and the proportion of the population susceptible to invasive mycoses is expected to increase, there is a need to identify new targets for the design and development of novel classes of antifungals. Although cell membrane and cell wall biosynthesis enzymes have been traditional targets of antifungals, broadening our search to unconventional pathways can unlock an abundance of potential targets by exploiting the unique metabolic needs of fungi during pathogenesis. A summary of the proposed targets discussed in this review is outlined in **Table 1**.

EMERGING TARGETS FOR NEW ANTIFUNGAL DEVELOPMENT

Glycolysis Pathway

Glycolysis is a fundamental, multi-step metabolic process that occurs in most living organisms. The glycolysis pathway produces energy, in aerobic and anaerobic conditions, through the catabolism of sugars and provides useful intermediates for downstream biosynthetic pathways. The roles of these enzymes in glycolysis have been extensively studied but there has been growing interest in their additional "moonlighting" functions that deviate from their canonical function. These "moonlighting" functions may contribute to fungal survival and virulence mechanisms and have consequently been investigated as promising antifungal drug targets.

Enolase and Host Plasminogen

Enolase catalyzes the penultimate step of glycolysis, interconverting 2-phosphoglycerate (2-PG) and phosphoenolpyruvate (PEP) (**Figure 2**) (Ji et al., 2016). Consistent with its role in integral metabolic processes, a genetic knockout of enolase in *C. albicans* has shown to reduce germination tube and hyphal formation, resulting in attenuated virulence and growth rate (De Backer et al., 2001; Ko et al., 2013). Intriguingly, this enzyme fulfills its glycolytic roles in the cytoplasm but it is also expressed on the cell surface of many fungal and bacterial species. This has been observed in *A. fumigatus* and *C. albicans*, two of the most prominent etiological agents of invasive mycoses, as well as *Aspergillus flavus, Aspergillus terrus, Aspergillus nidulans, Candida glabrata, Saccharomyces cerevisiae* and bacterial *Streptococci* and *Pneumococci* species (Pancholi and Fischetti, 1998; Edwards et al., 1999; Bergmann et al., 2001; Funk et al., 2016). In *S. cerevisiae*, it has been suggested that enolase secretion is mediated by an N-terminal 28 amino acid translocation sequence and operates through a SNARE-dependent pathway (Miura et al., 2012). Alignment of homologous enolase protein sequences from *S. cerevisiae* with *Aspergillus* spp. and *Candida* spp. have identified a highly conserved motif in the N-terminus. This suggests that this SNARE-dependent mechanism of enolase secretion may be conserved amongst many fungal species (Funk et al., 2016).

The function of surface-expressed enolase deviates from its canonical role in the cytoplasm and instead contributes to tissue invasion and nutrient acquisition during pathogenesis. It functions as a receptor for plasminogen, a zymogen circulating in the host bloodstream. Plasminogen is cleaved by plasminogen activating proteins (including tissue-plasminogen activator and urokinase plasminogen activator) to form active plasmin which functions as a serine protease. The physiological function of the human plasminogen/plasmin system is to facilitate tissue remodeling, cell migration, hemostasis and wound healing, and induce inflammation and angiogenesis. It fulfills these roles by degrading proteins in the extracellular matrix, fibrin and fibrinogen in blood clots and targeting components of the complement system (Law et al., 2013). Inactive plasminogen is proposed to dock onto a lysine-rich motif located on the surface-expressed enolase of fungal pathogens. While bound, host plasminogen activators are capable of recognizing, binding, and cleaving plasminogen to produce the activated plasmin (Funk et al., 2016). The serine protease activity of plasmin becomes concentrated at the site of fungal infection and this can accelerate tissue invasion and disease progression. An in-depth study which focused on surface-expressed enolase from *A. fumigatus* confirmed its ability to interact with human immune regulators including factor H, factor-H-like protein 1 (FHL-1), C4b-binding protein (C4BP) and plasminogen (Dasari et al., 2019). Whilst bound to *A. fumigatus* enolase, factor H, FHL-1 and C4BP retained their normal cellular activity and plasminogen remained accessible to plasminogen activator proteins. When swollen *A. fumigatus* conidia coated with human plasminogen were exposed to human A549 epithelial cells or an epithelial monolayer, cellular metabolic activity was reduced by 41% and cell retraction was reduced (Dasari et al., 2019). These observations are consistent with the "moonlighting" role of surface-expressed enolase and their role in facilitating tissue invasion.

Immunogenic Activity of Enolase

The secretion and presentation of enolase on the cell surface is a common feature of many pathogenic fungal species (Edwards et al., 1999; Funk et al., 2016). As a surface-exposed receptor, it is capable of eliciting an immunogenic response that may be exploited for vaccine development. Preliminary studies have

TABLE 1 | Cellular function of the proposed antifungal protein targets and phenotypic characteristics of various fungi after genetic knockout/knockdown.

	Protein targets for antifungal drugs						
	Enolase	**Mannitol-2-Dehydrogenase**	**mannitol-1-phosphate 5-dehydrogenase**	**Inosine monophosphate dehydrogenase**	**Guanosine monophosphate synthase**	**Nucleoside diphosphate kinase**	**References**
Roles in fungi	• Catalyzes the penultimate step of glycolysis • Expressed on the surface and functions as a receptor for human plasminogen • Contributes to tissue invasion and nutrient acquisition during infection	• Biosynthesis of mannitol, a storage of carbohydrates, an osmolyte and source of reducing power • Quenches reactive oxygen species • Protects conidia against stressful conditions		• Biosynthesis of purine nucleobases required for signal transduction pathways, energy metabolism and DNA and RNA synthesis			Meena et al., 2015; Funk et al., 2016; Ji et al., 2016; Chitty and Fraser, 2017
Phenotype of genetic knockout/knockdown							
Aspergillus fumigatus					• Growth defects that could be recovered with exogenous guanine • Avirulent in a murine model of infection	• Essential for fungal survival	Rodriguez-Suarez et al., 2007; Dinamarco et al., 2012
Candida albicans	• Reduced germination tube and hyphal formation • Attenuated virulence and growth rate				• Growth defects that could be recovered with exogenous guanine • Avirulent in a murine model of infection		De Backer et al., 2001; Rodriguez-Suarez et al., 2007; Ko et al., 2013
Cryptococcus neoformans		• Decreased tolerance to high temperatures and salinity* • Decreased resistance to oxidative stress • 5,000-fold less virulent compared to wild-type *C. neoformans* H99 in a mouse model		• Reduced growth rates • Impaired expression of virulence factors ○ Smaller capsule size ○ Reduced melanin expression • Avirulent in a murine model of cryptococcus	• Growth defects that could be recovered with exogenous guanine • Avirulent in a murine model of infection • Impaired expression of virulence factors ○ Smaller capsule ○ Delayed melanin production ○ No detectable protease activity		Chaturvedi et al., 1996a,b; Morrow et al., 2012; Chitty et al., 2017
Aspergillus niger			• Decreased tolerance to high temperatures and oxidative stress • Greater sensitivity to freeze-drying and freeze-thawing				Ruijter et al., 2003
Aspergillus fischeri			• Modest increase in sensitivity to heat and oxidative stress in conidia				Wyatt et al., 2014

(Continued)

TABLE 1 | Continued

	Protein targets for antifungal drugs						
	Enolase	**Mannitol-2-Dehydrogenase**	**Mannitol-1-phosphate-5-dehydrogenase**	**Inosine monophosphate dehydrogenase**	**Guanosine monophosphate synthase**	**Nucleoside diphosphate kinase**	**References**
Aspergillus flavus						• Disruption of one of two copies impairs conidia and sclerotia development • Affects plant virulence in a maize and peanut seed model	Wang et al., 2019
Aspergillus nidulans						• Essential for fungal survival • Essential for hyphal growth and conidia production	Lin et al., 2003

*These observations are based on the phenotype of a mutant *Cryptococcus neoformans* strain that produces low levels of mannitol. However, the mutation that causes this phenotype has not been determined.

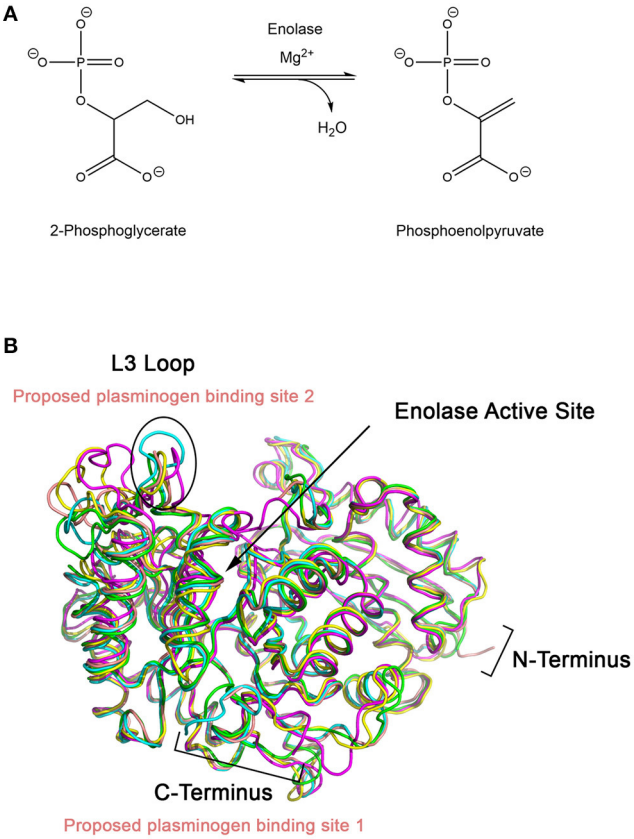


FIGURE 2 | (A) Enolase catalyzes the penultimate step of glycolysis, interconverting 2-phosphoglycerate (2-PG) and phosphoenolpyruvate (PEP) using magnesium as a cofactor. **(B)** X-ray crystal structures of enolase from different species have similar global folds. Enolase structures from *Escherichia coli* (1E9I; green), *Enterococcus hirae* (1IYX; cyan), lobster (1PDZ; yellow), *Streptococcus pneumoniae* (1W6T; salmon), and *Saccharomyces cerevisiae* (1ONE; magenta) were superimposed using *PyMOL* (Stec and Lebioda, 1990; Duquerroy et al., 1995; Kuhnel and Luisi, 2001; Hosaka et al., 2003; Ehinger et al., 2004; Kang et al., 2008). The L3 loop, which shows the greatest structural deviation between the structures, is proposed to be the plasminogen binding loop for *S. pneumoniae*. Exposed lysine residues at the C-terminal end of enolase have also been proposed as a possible binding site.

demonstrated the suitability of using enolase as an immunogenic agent to acquire modest protective effects against candidiasis in a murine model. When mice were challenged with *C. albicans* following subcutaneous vaccination with recombinant *C. albicans* enolase, an elevated antibody response was induced and fungal burden in major organs was reduced compared to non-immunized mice (Li et al., 2011). In an independent study, *C. albicans* enolase was expressed and presented on *Lactobacillus casei* cells, then orally administered to mice prior to challenging with a normally lethal dose of *C. albicans*. These immunized mice produced higher IgG antibody titers and had improved survival rates (Shibasaki et al., 2014).

Targeting Surface-Expressed Enolase

Designing a peptidomimetic or small molecule inhibitor that specifically targets the plasminogen docking site of enolase

using a structure-guided approach is a possible avenue for new therapeutics. This approach requires intimate knowledge of the interactions between both proteins. Although high resolution crystal structures of the full-length human Type II plasminogen (4DUR) and fungal enolase from *S. cerevisiae* (3ENL) are already available, the interaction interface has not yet been defined (Stec and Lebioda, 1990; Law et al., 2012). There have been previous studies examining the molecular interaction between plasminogen and bacterial enolases. Studies conducted using enolase from *S. pneumonia* have suggested two possible binding sites—the L3 loop on the alpha-beta (αβ) barrel, a structurally flexible region, or the lysine-rich C-terminal tail of the protein (Pancholi and Fischetti, 1998; Bergmann et al., 2003; Ehinger et al., 2004). Superimposition of enolase from prokaryotic and eukaryotic sources including *E. coli* (1E9I), human (3B97), *Enterococcus hirae* (1IYX), lobster (1PDZ), *Streptococcus pneumonia* (1W6T), and *S. cerevisiae* (3ENL) reveal an overall similar fold and existence of a conserved, but structurally different, L3 loop (**Figure 3**) (Stec and Lebioda, 1990; Duquerroy et al., 1995; Kuhnel and Luisi, 2001; Hosaka et al., 2003; Ehinger et al., 2004; Kang et al., 2008). Therefore, it is likely that the plasminogen binding site is also conserved between homologous enolase enzymes but the mode of binding may differ between species. Building upon this work, efforts must be focused on elucidating the plasminogen binding interface of fungal enolase and structurally characterizing this site to guide rational inhibitor design.

Evidence of the immunogenicity of fungal enolase, particularly the *C. albicans* homolog, raises the possibility of using antibody therapy as an adjunct to antifungal therapy. Therapeutic antibodies specific for enolase could potentially reduce tissue invasion and fungal dissemination by interfering with interactions between fungal enolase and host plasminogen. Antibodies raised in mice against recombinant *C. albicans* or *A. fumigatus* enolase successfully recognized plasminogen-bound enolase from *Candida* spp. and *Aspergillus* spp., respectively. However, this antibody-enolase-plasminogen complex still allowed plasminogen activating proteins to access, cleave, and activate bound plasminogen (Funk et al., 2016). The viability of antibody therapy requires improved antibodies against enolase that recognize specific epitopes that interfere with this plasminogen activation process. Monoclonal antibodies that inhibit plasminogen activation have already been successfully raised against human α-enolase. During the process of fibrinolysis, plasminogen binds to surface-expressed α-enolase on leukocytic cells and this results in an accelerated rate of plasminogen activation. The plasminogen activation associated with binding to leukocytic cell lines was reduced by ∼90% when monoclonal α-enolase antibodies (MAb 11G1) were introduced (López-Alemany et al., 2003). These experiments provide a proof-of-principle that disruption of an enolase-plasminogen interaction can affect plasmin activity and therefore, this approach may be feasible when applied to an antifungal therapeutic strategy.

MANNITOL BIOSYNTHESIS PATHWAY

Mannitol biosynthesis is a crucial process that aids in fungal survivability and virulence. Mannitol-2-dehydrogenase (M2DH) and mannitol-1-phosphate 5-dehydrogenase (M1P5DH) are the primary biosynthesis enzymes that produce mannitol, an acyclic, six-carbon sterol alcohol (**Figure 3**). Present in all fungal structures, including mycelia, fruiting bodies and conidia, mannitol acts as a carbohydrate storage molecule, osmolyte, and source of reducing power (Meena et al., 2015). During pathogenesis, mannitol production is increased to exploit its ability to quench reactive oxygen species (ROS) in order to resist host defenses (Meena et al., 2015). The unique and diverse properties of mannitol, particularly in relation to human pathogenesis, have garnered considerable interest in the underlying mechanism of its biosynthesis and secretion during fungal infection.

Mannitol as a Biomarker of Fungal Infection

The efficacy of antifungal therapy is dependent on the early detection and accurate identification of the invading fungal species. Traditional methods include culturing from patient samples, direct microscopical examination of fungal samples and histopathology testing of infected tissue (Kozel and Wickes, 2014). Although culturing is considered as the gold standard of diagnosis, the utility of this technique can suffer from the slow growth rates of filamentous fungi, environmental contamination, and poor recovery of some fungal species from patient samples (Kozel and Wickes, 2014). Overall, these factors can delay adequate antifungal treatment which can compromise the efficacy of these drugs. Therefore, non-culture-based methods of diagnosis, such as screening patient samples for biomarkers of fungal infection, offer a faster and less invasive alternative.

Several fungal pathogens, including *C. neoformans* and *Aspergillus* spp., have been shown to produce high levels of mannitol *in vitro* and in animal disease models of infection, highlighting the suitability for its use as a biomarker. In liquid cultures of 12 human isolates of *C. neoformans*, extracellular mannitol levels increased ∼30-fold over a 4 day period in which rapid mannitol biosynthesis and secretion occurred predominantly in the stationary growth phase (Wong et al., 1990). Similarly, several *Aspergillus* species have also demonstrated an ability to produce mannitol from glucose *in vitro* (Birkinshaw et al., 1931). Closely mirroring these studies, mannitol levels in the cerebrospinal fluid (CSF) of a cortisone-treated rabbit disease model of cryptococcal meningitis also increased over time in response to infection with *C. neoformans* strain H99. Furthermore, there was a positive correlation identified between mannitol CSF levels and CSF colony forming units as well as CSF cryptococcal antigen titers, both of which are markers of infection severity (Wong et al., 1990). In an experimental model of aspergillosis using rats infected with *A. fumigatus*, high levels of mannitol were detected in liver tissue 12 h post-infection and in serum 36 h post-infection (Wong

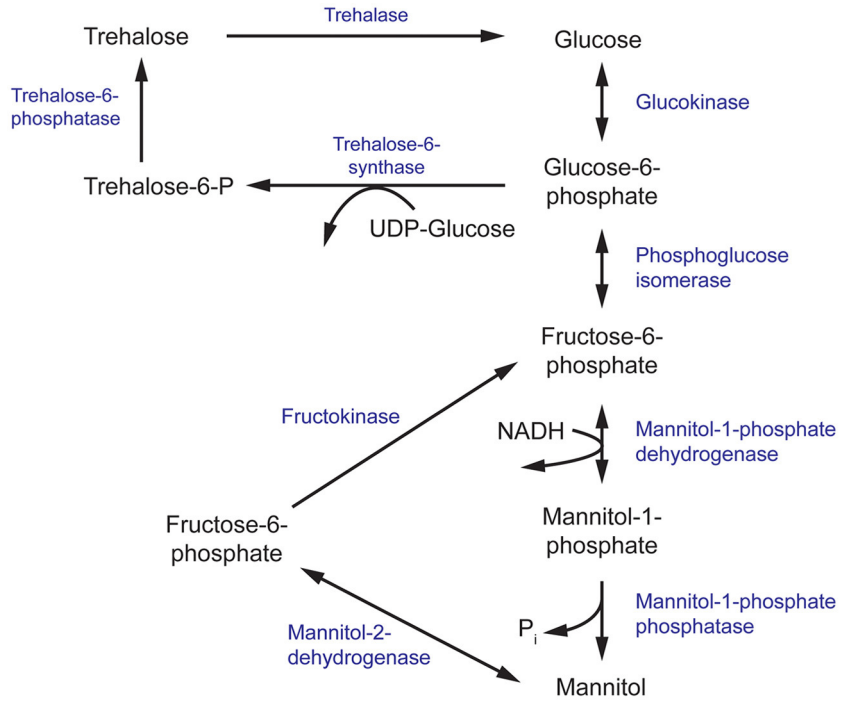


FIGURE 3 | The mannitol and trehalose biosynthesis pathway in fungi. The primary biosynthesis enzymes of mannitol, and the focus of this review, are mannitol-1-phosphate dehydrogenase (M1P5DH) and mannitol-2-dehydrogenase (M2D).

et al., 1989). In addition, a singularly surviving rat 48 h post-infection also showed increased mannitol levels in lung, liver, and kidney tissue (Wong et al., 1989). From these animal models, the results indicate that infection of a human host with *C. neoformans* or *A. fumigatus* is also likely to induce upregulated mannitol production and secretion, localizing in the CSF or serum, respectively. Analysis of the CSF from patients suffering from AIDS and cryptococcal meningitis confirm that mannitol content increases to sufficiently high levels to be easily detected and quantified using gas-liquid chromatography-linked mass spectroscopy (Megson et al., 1996). As humans lack an equivalent mannitol biosynthesis pathway, the detected mannitol is likely of fungal origin. However, mannitol content in the CSF could not be significantly correlated to cryptococcal antigen titers in humans, contrary to observations made in the rabbit disease model. Taken together, these results provide valuable insights into the feasibility of using patient CSF and serum mannitol levels as a biomarker of fungal infection and presents possible methodologies that facilitate its use as a diagnostic tool.

Roles in Virulence and Protection Against Host Defenses

Fungal pathogens invading a human host experience multiple environmental stresses associated with occupying a new niche and chemical stresses invoked by the immune response. Sensing and adapting to abiotic stresses (temperature, pH), resisting phagocytosis from macrophages and neutralizing ROS produced by neutrophils is critical for survival (Cooney and Klein, 2008; Khanna et al., 2016). Mannitol has been connected to many of these host-pathogen interactions associated with fungal survival and virulence and this provides a basis to understand the function of its rapid biosynthesis and secretion during infection.

Phenotypic characterization of a *C. neoformans* mutant which produces reduced levels of mannitol (Cn MLP) exemplifies the multitude of roles adopted by mannitol. The Cn MLP mutant showed similar growth rates, morphology, and antifungal susceptibility as the wild-type strain, *C. neoformans* H99. However, Cn MLP had decreased tolerance to high temperatures and salinity (Chaturvedi et al., 1996a). These findings are consistent with the proposed role of mannitol as an osmolyte in fungi (Meena et al., 2015). Although the loss of thermotolerance in Cn MLP cannot be directly linked to mannitol function, resistance to thermal stress in the fungus *Beauveria bassiana* has been partially attributed to trehalose, a metabolite produced from mannitol (**Figure 3**) (Liu et al., 2008). The Cn MLP isolate also demonstrated decreased resistance to oxidative stress, a common killing mechanism of the human immune system (Chaturvedi et al., 1996b). Specifically, Cn MLP showed greater susceptibility to oxidative killing by polymorphonuclear neutrophils (PMNs) and ROS generated from an iron-hydrogen peroxide-iodide cytotoxic cell-free system, relative to the wild-type strain (Chaturvedi et al., 1996b). Protection against ROS improved in a dose-dependent manner when PMNs were incubated with mannitol. Similarly, fungicidal activity was completely abolished when mannitol was added simultaneously to the ROS cell free system (Chaturvedi et al., 1996b). In both scenarios, mannitol

acted as an effective scavenger of ROS to confer protection against oxidative damage. In an experimental disease model, Cn MLP was 5,000-fold less virulent compared to *C. neoformans* H99 when intravenously injected into mice. Although both the mutant and wild-type strains proved to be lethal, 100% of mice succumbed to infection 51 days post-injection with *C. neoformans* H99 whereas 100% of mice injected with Cn MLP survived at the same time point (Chaturvedi et al., 1996a). Together, the osmotolerance and thermotolerance conferred by intracellular mannitol and the protection against ROS by extracellular mannitol may aid fungal survival and virulence by resisting environmental and immune-induced stresses in the human body.

Roles in Conidial Stress Response

Conidia are asexually-produced spores that are required for wide-spread fungal dispersal. Often dissipated through the air, conidia are inhaled into the respiratory system and at this developmental stage, fungal infection can be initiated (Dagenais and Keller, 2009). Conidia must be able to survive and adapt to this new environment to progress to subsequent germination stages and develop into a vegetative state. Mannitol has been implicated in the protection of conidia against these stress conditions in multiple species of fungi (Chaturvedi et al., 1996a; Ruijter et al., 2003; Wyatt et al., 2014).

High concentrations of intracellular mannitol have been detected in conidia from *Aspergillus oryzae*, *Aspergillus clavatus* and *Aspergillus niger* (Horikoshi et al., 1965; Corina and Munday, 1971; Witteveen and Visser, 1995). In order to assess mannitol function, a M1P5DH-null mutant of *A. niger* was generated. Disruption of M1P5DH resulted in a complete absence of mannitol in mycelia and a 30% reduction of intracellular mannitol content in conidia. Prolonged exposure to conditions that had no effect on the viability of wild-type conidia, such as high temperatures and oxidative stress induced by ROS, reduced the viability of M1P5DH-null conidia to 5 and 0.5%, respectively. To a lesser extent, these mutant conidia were more sensitive to freeze-drying and freeze-thawing processes. Supplementation of M1P5DH-null conidia with mannitol had a protective effect that reduced their sensitivity to heat and oxidative stress (Ruijter et al., 2003). These results indicate that mannitol is an important molecule that protects conidia against harsh environmental conditions.

In a separate study, the equivalent gene encoding M1P5DH in *Aspergillus fischeri* was deleted with unexpected effects. There was a severe reduction of mannitol in mycelia but an increase in mannitol and reduction of trehalose in its conidia (Wyatt et al., 2014). These conidia showed a modest increase in sensitivity to heat and oxidative stress, despite the higher mannitol content. Since trehalose and mannitol biosynthesis pathways are intrinsically linked and both sugars have been found in all stages of the fungal life cycle, it has been suggested that trehalose and mannitol may play overlapping roles in resistance to environmental stress (**Figure 3**) (Wyatt et al., 2014; Thammahong et al., 2017). There has also been growing interest in targeting components of the trehalose biosynthesis pathway for the development of novel antifungals (Thammahong et al., 2017). Thus, further investigation into the interplay between mannitol and trehalose in fungal survival is warranted to aid these efforts.

Targeting Mannitol Biosynthesis Enzymes

Several lines of evidence demonstrate the multifaceted roles of mannitol in the protection of fungi from stress conditions commonly experienced during infection of a host (Chaturvedi et al., 1996a; Ruijter et al., 2003; Wyatt et al., 2014). Accelerated mannitol production and secretion can confer thermotolerance and resistance to ROS generated from the human immune system needed to thrive in these harsh environments. Adding to the allure of targeting mannitol biosynthesis pathways in fungi is the fact that humans do not produce mannitol and therefore lack equivalent enzymes. This absence of a human homolog greatly simplifies the design and development of inhibitory compounds that specifically target these fungal enzymes. In turn, this will likely reduce the severity of off-target effects.

Efforts to target the mannitol biosynthesis enzyme should be focused on M1P5DH. In several species, knockout of the gene encoding M1P5DH (*mpdA*) and phenotypic characterization of these mutants have shown significant decreases in intracellular mannitol levels (Ruijter et al., 2003; Wyatt et al., 2014). Although mannitol can be synthesized in an alternate pathway using M2DH, it appears that M1P5DH features in the predominant pathway. Hence, inhibition of M1P5DH is likely to have a greater effect on conidia viability and can be useful to halt initial stages of infection and dissemination.

Targeting the mannitol biosynthesis pathway can also be beneficial to delay the progression of existing infections. As demonstrated in the animal models infected with Cn MLP, decreased mannitol levels were associated with reduced virulence and ultimately, greater survival. By inhibiting the activity of these biosynthesis enzymes, mannitol production is reduced and this may increase fungal sensitivity to oxidative killing by components of the immune system.

PURINE NUCLEOTIDE BIOSYNTHESIS PATHWAY

Nucleobases, including purines and pyrimidines, are heterocyclic molecules that are found in all forms of life. They are obtained exogenously from the environment or can be synthesized endogenously. As they are involved in essential processes, including DNA and RNA synthesis, energy metabolism and signal transduction, the intracellular pool of nucleobases must be tightly maintained (Chitty and Fraser, 2017). Hence, disruption of the purine and pyrimidine biosynthesis pathway using small molecules has been explored as a potential approach to develop antimicrobials (Rodriguez-Suarez et al., 2007; Shu and Nair, 2008; Du Pré et al., 2018; Park et al., 2018; Trapero et al., 2018). *De novo* purine biosynthesis begins with phosphoribosylpyrophosphate (PRPP) which is ultimately converted to inosine monophosphate (IMP), a precursor to both adenine and guanine nucleobases, through a series of enzymatic reactions (**Figure 4**) (Chitty and Fraser, 2017). This section will

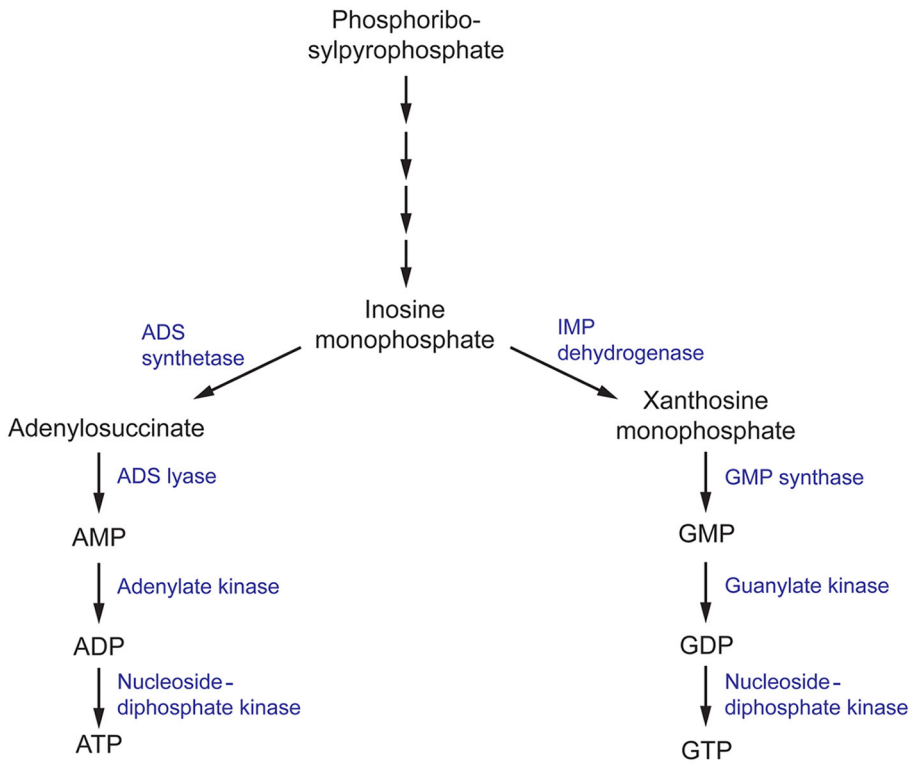


FIGURE 4 | The purine metabolism pathway in fungi. ADS, (adenylosuccinate) synthetase; ADS, (adenylosuccinate) lyase; IMP (inosine monophosphate) dehydrogenase; GMP, (guanosine monophosphate) synthase.

focus on targeting enzymes responsible for the synthesis of guanine nucleobases.

Roles in Virulence

Inosine monophosphate dehydrogenase (IMPDH) is the first committed and rate-limiting step in the guanosine triphosphate (GTP) biosynthesis pathway. Using NAD^+ as a cofactor, it catalyzes the conversion of IMP to xanthosine monophosphate (XMP) *via* a dehydrogenase and hydrolysis reaction (Hedstrom, 2009). IMPDH catalytic activity has been closely associated with microbial virulence, and has been the target of several drug discovery projects for antibiotics, antifungals, antivirals, and anticancer therapeutics with great success (Markland et al., 2000; Floryk and Thompson, 2010; Kummari et al., 2018; Trapero et al., 2018). In terms of antifungals, there has been substantial progress in elucidating the mechanisms in which IMPDH contributes to *C. neoformans* survival and virulence. To first validate the rationality of IMPDH as an antifungal target, an IMPDH-deficient mutant of *C. neoformans* was produced and phenotypically characterized. This mutant exhibited reduced growth rates and impaired expression of virulence factors, resulting in smaller capsule sizes and reduced melanin expression (Morrow et al., 2012). Importantly, these phenotypic defects could not be completely recovered with supplementation of exogenous guanine that may have been sourced from the human host during infection. Due to the direct involvement of IMPDH in virulence factor production, IMDPH-deficient *C. neoformans* was completely avirulent in a murine model of cryptoccoccus. As a result, fungal loads in the brain were cleared 3 days post-infection, total fungal clearance was achieved 2 weeks post-infection and ultimately, this resulted in a 100% survival rate 50 days post-infection (Morrow et al., 2012). These experiments confirm that IMPDH is involved in essential processes associated with *C. neoformans* virulence and this presents a promising avenue for the design of antifungal compounds.

In the following step of the purine biosynthesis pathway, XMP produced by IMPDH is converted to guanosine monophosphate (GMP) by guanosine monophosphate synthase (GMP synthase). In a similar manner to IMPDH, GMP synthase has been extensively studied in *A. fumigatus*, *C. albicans*, and *C. neoformans* and has been determined to be crucial in the survival and virulence of these fungal pathogens (Rodriguez-Suarez et al., 2007; Chitty et al., 2017). A conditional knock-out of the gene encoding GMP synthase (*GUA1*) in *A. fumigatus* and *C. albicans*, and deletion in *C. neoformans*, inhibited growth on minimal media but was recovered when supplemented with exogenous guanine. These growth defects are consistent with the essential role of purines as components of DNA and RNA. Similar to the *in vivo* experiments conducted with IMPDH-deficient *C. neoformans*, disruption of *GUA1* in *A. fumigatus*, *C. albicans*, and *C. neoformans* also rendered these strains avirulent in a murine model of infection (Rodriguez-Suarez et al., 2007;

Chitty et al., 2017). Phenotypic characterization of Δ*GUA1 C. neoformans* revealed significant differences in the production of key virulence factors that may account for the abolished virulence. Despite supplementation with exogenous guanine, Δ*GUA1 C. neoformans* mutants produced less capsule, showed delayed melanin production and lacked detectable protease activity relative to their wild-type counterparts. Collectively, these virulence factors function to aid survival by resisting host defenses and fungal dissemination. Specifically, the capsule protects against phagocytosis, melanin resists oxidative damage and secretion of proteases facilitates host tissue degradation for nutrient acquisition. Intriguingly, this effect was more pronounced at the optimum human body temperature of 37°C (Chitty et al., 2017). Since *GUA1* is a determinant of virulence in several of these clinically-prominent fungal pathogens, inhibition of GMP synthase may be an ideal approach to develop antifungal drugs with broad-spectrum activity.

Roles in Survival

In the final steps of the purine biosynthesis pathway, GMP must be converted to an accessible form for its use in DNA and RNA synthesis and signal transduction. GMP is first converted to guanosine diphosphate and then nucleoside-diphosphate kinase (NDK) covalently attaches a phosphate moiety to form GTP (Chitty and Fraser, 2017). NDK is a promiscuous enzyme that can use all nucleoside diphosphates as substrates, with varying levels of selectivity, to generate their respective nucleoside triphosphates (Nguyen et al., in press). Hence it is also involved in the adenine nucleoside synthesis pathway to produce adenosine triphosphate (ATP) as well as the synthesis of the remaining building blocks of DNA and RNA. Consistent with these fundamental processes, in *A. nidulans*, NDK is involved in various aspects of the life cycle, including hyphal growth and conidia production which is necessary for its growth and dispersal (Lin et al., 2003). Similarly in *A. flavus*, disruption of one of the two copies of *ndk* impairs the development of conidia and sclerotia and affects plant virulence in a maize and peanut seed model (Wang et al., 2019). Identified as an essential gene for survival in *A. fumigatus* and *A. nidulans*, NDK has been identified as a promising target for antifungals against *Aspergillus* spp. (Lin et al., 2003; Dinamarco et al., 2012).

Targeting Purine Biosynthesis Pathways

Achieving species-selective targeting of *de novo* purine biosynthesis enzymes is an important consideration in the drug design process. This is especially important since bacteria, fungi and humans share commonality in many of these enzymes (Sperling, 1988; Kilstrup et al., 2005; Chitty and Fraser, 2017). Thus, intimate structural and kinetic knowledge of the fungal and human homologs of these enzymes will be invaluable in the design of inhibitors with high potency and selectivity. Currently, structures of human IMPDH, isoforms I (Risal et al. unpublished) and II (Colby et al., 1999), GMP synthase (Welin et al., 2013), and NDK (Chen et al., 2003) have been solved as well as structures of fungal IMPDH (*C. neoformans*) (Morrow et al., 2012) and NDK (*A. fumigatus* and *A. flavus*) (Wang and Wang, 2019; Nguyen et al., in press). Availability of these structures allows identification of key active site residues that differ between fungal and human enzymes and this may be integral in the design of highly selective inhibitors.

Knowledge of the enzyme structure and catalytic mechanism aids in identifying the most "druggable" binding pocket. In the case of IMPDH, there are three possible sites: (i) the natural substrate site, IMP, (ii) the co-factor site, NAD^+, or (iii) an allosteric site (Shu and Nair, 2008). Ongoing drug discovery projects targeting IMPDH from *Mycobacterium tuberculosis* and *Cryptosporidium parvum* have shown success in focusing on the co-factor binding site due to its divergence from the mammalian equivalent (Umejiego et al., 2008; Trapero et al., 2018). In the case of *C. parvum* IMPDH, a high-throughput screening approach was employed to specifically target the NAD^+ site. Ten compounds, from a pool of 134 preliminary hits, were determined to be reversible inhibitors that were 9- to 400-fold more selective for the parasitic IMPDH over the human IMPDH isoform II (Umejiego et al., 2008). These experiments provide a proof of concept that IMPDH can be effectively and selectively targeted by small molecule inhibitors. When considering IMPDH from *C. neoformans*, a similar screening approach targeting the co-factor binding site may also be feasible due to its structural and functional deviation from human IMPDH (Morrow et al., 2012). A structure of human IMPDH isoform II bound to selenazole-4-carboxyamide-adenine dinucleotide, an analog of NAD^+, reveals differences in the residues that comprise the co-factor binding pocket between *C. neoformans* and human IMPDH. It can be inferred that the interaction between IMPDH and NAD^+ is stabilized by a π-π stacking interaction between the adenine ring of NAD^+ and R267 and Y296 of *C. neoformans* IMPDH which corresponds to R253 and Y282 in human isoform I or H253 and F282 in human isoform II, respectively (**Figure 5A**) (Morrow et al., 2012). Although many residues within the binding pocket are invariant between *C. neoformans* and human isoforms I and II, these differences are likely to be sufficient to achieve species-selective inhibition. A high-throughput screen against IMPDH from *C. neoformans* has identified promising candidates (Kummari et al., 2018). From a chemical library of 114,000 drug-like compounds, three 3-((5-substituted)-1,3,4-oxadiazol-2-yl)thio benzo[*b*]thiophene 1,1-dioxides were identified as having inhibitory activity against *C. neoformans* IMPDH. Through chemical modification, several analogs were then synthesized to assess the structure-activity relationships. Although many of these analogs inhibited *C. neoformans* IMPDH *in vitro* and had whole cell activity, they were also cytotoxic to human cell lines (Kummari et al., 2018). These results illustrate the complexities associated with targeting pathways shared between fungi and humans. However, there is also great opportunity to build upon this work by trialing modifications that improve selectivity that retain the antifungal properties of the compound whilst reducing cytotoxic effects.

Similarly to IMPDH, inhibition of fungal GMP synthase may be achieved effectively by high-throughput screening of chemical libraries or fragment screening. Since GMP synthase is a bifunctional protein with two catalytic domains working in tandem to fulfill its function, it is imperative to select the most feasible binding pocket to target. The glutamine amidotransferase

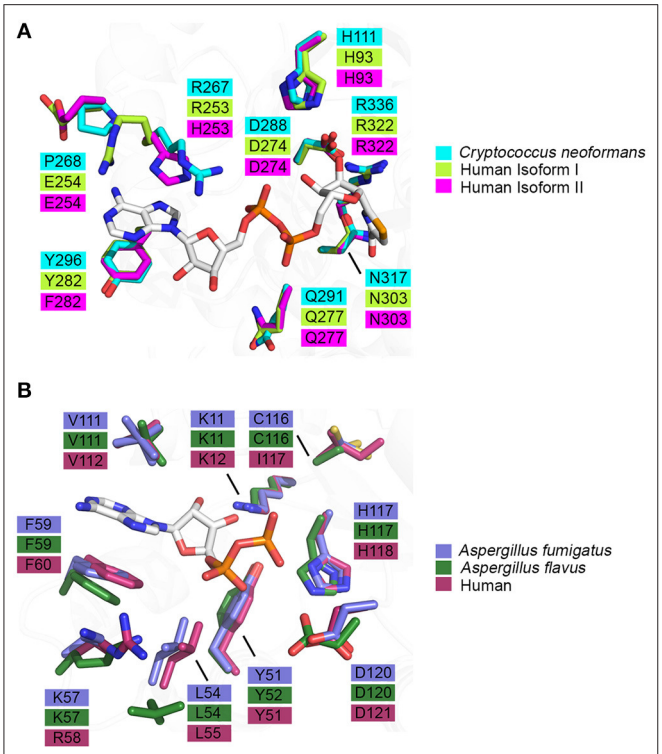


FIGURE 5 | Binding pocket composition analysis reveals divergent residues between human and fungal homologs that may be exploited to achieve specificity. **(A)** Residues comprising the binding pocket of inosine monophosphate dehydrogenase from *Cryptococcus neoformans* (4AF0, cyan), human isoform I (1JCN, green), human isoform II (1B3O, magenta) and selenazole-4-carboxyamide-adenine dinucleotide, an NAD^+ analog, are represented as sticks (Colby et al., 1999; Morrow et al., 2012) (Risal et al. unpublished). **(B)** Residues comprising the binding pocket of nucleoside di-phosphate kinase from *Aspergillus fumigatus* (6XP7, purple), *Aspergillus flavus* (6K3H, dark green) and human (2HVD, maroon) and adenosine diphosphate, a natural substrate, are represented as sticks (Giraud et al., 2006; Wang et al., 2019; Nguyen et al., in press).

(GAT) domain releases ammonia from glutamine which is shuttled to the ATP pyrophosphatase (ATPPase) domain where XMP is converted to GMP *via* an adenyl-XMP intermediate. There are pre-existing inhibitors that target the glutamine amidotransferase domain such as acivicin (ACI) and 6-diazo-5-oxo-L-norleucine (DON) that inadvertently inhibit GMP synthase activity (Ahluwalia et al., 1990; Chittur et al., 2001). However, targeting this catalytic module of GMP synthase would likely lead to non-specific and off-target effects due to the prevalence of GAT domains in other enzymes. Therefore, a more promising endeavor is to study the XMP- and ATP-binding pockets located within the ATPPase domain to achieve potent inhibition and high selectivity for the fungal enzyme over the mammalian equivalent. A novel compound, denoted ECC1385, identified from a synthetic compound library has been shown to inhibit *C. albicans* and *C. neoformans* GMP synthase activity *in vitro* and exhibits whole cell activity against a broad-spectrum of fungi and bacteria, including multiple species of *Candida*, *A. fumigatus* (MF5668), *C. neoformans* (MY2062), and *Staphylococcus aureus* (MB2865) (Rodriguez-Suarez et al., 2007). Mechanism-of-action studies of ECC1385 have revealed that inhibition of GMP synthase is achieved through a different mode of action from ACI and DON. These data suggest that small chemical compounds that target the activity of the ATPPase domain is possible and presents a promising avenue to pursue.

As NDK has maintained its essential role in NTP biosynthesis throughout all domains of life, there is high structural and functional conservation between different species. This conservation is reflected in the largely invariant composition of binding pocket residues between fungal and human homologs (**Figure 5B**). However, two key differences are revealed when the crystal structures of ADP-bound NDK from human (2HVD), *A. fumigatus* (6XP7), and *A. flavus* (6K3H) are superimposed (**Figure 5B**) (Giraud et al., 2006; Nguyen et al., in press). Specifically, in the nucleoside binding pocket, Lys57 and C116 in both fungal NDKs are replaced by Arg58 and I117 in the human structure, respectively. These discrepancies presents an opportunity to selectively target the catalytic site and this may be beneficial in the development of a pan inhibitor with broad-spectrum activity against multiple pathogenic fungi. Azidothymidine (AZT), also marketed as RETROVIR, is an antiretroviral medication currently used to treat HIV (Furman et al., 1986). It has been shown that AZT can inhibit NDK from *A. flavus in vitro* and exhibits antifungal activity *in vivo* (Wang et al., 2019). Docking of AZT to the crystal structure of *A. flavus* NDK predicts that it acts as a competitive inhibitor and interacts with residues R104, H117, and D120 in the catalytic site. Since AZT is an FDA-approved drug, there is potential to repurpose it as an antifungal drug or to use the AZT scaffold and improve upon its antimycotic properties. An alternate approach may be explored to inactivate fungal NDK preferentially over the human homolog using non-competitive inhibitors. Allosteric sites are likely to show more variation between species, compared to the catalytic site, and therefore, development of a non-competitive inhibitor may be a more effective approach. Thorough kinetic analysis has revealed that ebselen, a cysteine-modifying compound, acts as a non-competitive inhibitor of human NDK secreted by airway epithelial cells (Semianrio-Vidal et al., 2010). In this context, it is not used in an antifungal application. Instead, it is used to obtain stable and accurate measurements of nucleotides in airway surface liquid by disrupting nucleotide interconversion without the use of a nucleotide-derived inhibitor. Identifying allosteric sites in NDK from fungal pathogens, comparing them to the human homolog and exploiting their differences may be imperative to simultaneously achieve potent inhibition and high selectivity. Hence, it would be advantageous to explore non-competitive inhibitors of NDK that bind an allosteric site and influence catalytic activity distally.

Although not explicitly addressed in this review, there has been substantial progress made in targeting enzymes from the *de novo* pyrimidine biosynthesis pathway that is worth further discussion. Notably, a novel drug, known as F901318 or Olorofim), has shown antifungal activity against *A. fumigatus* and *A. flavus* by inhibiting the enzyme, dihydroorotate dehydrogenase (Du Pré et al., 2018). Furthermore, it has shown

efficacy against *A. fumigatus*, both azole-sensitive and azole-resistant strains, in a murine model of pulmonary aspergillosis (Oliver et al., 2016). F901318 is currently being assessed for its use in treating invasive fungal infections caused by *Lomentospora prolificans*, *Scedosporium* spp., *Aspergillus* spp., and other resistant fungi. The development of F901318 and progression into phase 2 clinical trials highlights the untapped potential of targeting shared metabolic pathways between fungi and humans for the discovery of novel antifungal drug classes.

CONCLUSIONS

As the number of immunocompromised patients continues to grow, invasive fungal infections are becoming more pervasive. Due to inherent limitations in antifungal drug treatments and the emergence of resistance, the limited spectrum of currently available antifungals is a major barrier to effectively treat systemic fungal infections. Consequently, there is a need to identify and validate novel antifungal drug targets. Targeting eukaryotic proteins that are common to both fungi and humans is a difficult but necessary approach since the number of fungal enzymes that lack a human homolog and are essential for survival or virulence are scarce. Hence, there has been an increased interest in exploiting subtle structural and functional differences of metabolic enzymes shared by both fungi and humans. In the field, significant progress has been made in the characterization of enolase, an enzyme implicated in tissue invasion, as well as enzymes from the mannitol biosynthesis and *de novo* purine biosynthesis pathways in many pathogenic fungi. In addition to the pathways explored in this review, *de novo* pyrimidine biosynthesis enzymes, intracellular redox environment regulators and key phosphatases, kinases and transcription factors that form cellular signaling networks have been explored as potential antifungal drug targets (Oliver et al., 2016; Marshall et al., 2019; Jin et al., 2020). Altogether, these experiments lay the foundation for antifungal drug discovery projects that target unconventional, but promising, pathways.

AUTHOR CONTRIBUTIONS

SN, JT, and JB: conceptualization, manuscript review, and editing. SN and JT: manuscript preparation. All authors contributed to the article and approved the submitted version.

FUNDING

SN and JT were recipients of an Australian Government Researching Training Program stipend scholarship.

ACKNOWLEDGMENTS

The authors would like to acknowledge the members of the Bruning Lab for their insightful discussions during the preparation of this manuscript.

REFERENCES

Ahluwalia, G. S., Grem, J. L., Hao, Z., and Cooney, D. A. (1990). Metabolism and action of amino acid analog anti-cancer agents. *Pharmacol. Therapeut.* 46, 243–271. doi: 10.1016/0163-7258(90)90094-I

Allen, D., Wilson, D., Drew, R., and Perfect, J. R. (2015). Azole antifungals: 35 years of invasive fungal infection management. *Expert Rev. Anti Infect. Ther.* 13, 787–798. doi: 10.1586/14787210.2015.1032939

Azevedo, M.-M., Faria-Ramos, I., Cruz, L. C., Pina-Vaz, C., and Rodrigues, A. G. (2015). Genesis of azole antifungal resistance from agriculture to clinical settings. *J. Agric. Food Chem.* 63, 7463–7468. doi: 10.1021/acs.jafc.5b02728

Benedict, K., Jackson, B. R., Chiller, T., and Beer, K. D. (2019). Estimation of direct healthcare costs of fungal diseases in the United States. *Clin. Infect. Dis.* 68, 1791–1797. doi: 10.1093/cid/ciy776

Bergmann, S., Rohde, M., Chhatwal, G. S., and Hammerschmidt, S. (2001). α-Enolase of *Streptococcus pneumoniae* is a plasminogen-binding protein displayed on the bacterial cell surface. *Mol. Microbiol.* 40, 1273–1287. doi: 10.1046/j.1365-2958.2001.02448.x

Bergmann, S., Wild, D., Diekmann, O., Frank, R., Bracht, D., Chhatwal, G. S., et al. (2003). Identification of a novel plasmin(ogen)-binding motif in surface displayed α-enolase of *Streptococcus pneumoniae*. *Mol. Microbiol.* 49, 411–423. doi: 10.1046/j.1365-2958.2003.03557.x

Bicanic, T., and Harrison, T. S. (2005). Cryptococcal meningitis. *Br. Med. Bull.* 72, 99–118. doi: 10.1093/bmb/ldh043

Birkinshaw, J. H., Charles, J. H. V., Hetherington, A. C., and Raistrick, H. (1931). On the production of mannitol from glucose by species of *Aspergillus*. *Philos. Trans. R. Soc. B* 22, 153–171.

Bongomin, F., Gago, S., Oladele, R. O., and Denning, D. W. (2017). Global and multi-national prevalence of fungal disease - estimate precision. *J. Fungi* 3:57. doi: 10.3390/jof3040057

Brauer, V. S., Rezende, C. P., Pessoni, A. M., de Paula, R. G., Rangappa, K. S., Nayaka, S. C., et al. (2019). Antifungal agents in agriculture: friends and foes of public health. *Biomolecules* 9:521. doi: 10.3390/biom9100521

Brown, G. D., Denning, D. W., Gow, N. A. R., Levitz, S. M., Netea, M. G., White, T. C., et al. (2012). Hidden killers: human fungal infections. *Med. Mycol.* 4:165rv113. doi: 10.1126/scitranslmed.3004404

Bueid, A., Howard, S. J., Moore, C. B., Richardson, M. D., Harrison, E., Bowyer, P., et al. (2010). Azole antifungal resistance in *Aspergillus fumigatus*: 2008 and 2009. *J. Antimicrob. Chemother.* 65, 2116–2118. doi: 10.1093/jac/dkq279

Chaturvedi, V., Flynn, T., Neihaus, W. G., and Wong, B. (1996a). Stress tolerance and pathogenic potential of a mannitol mutant of *Cryptococcus neoformans*. *Microbiology* 142(Pt 4), 937–943. doi: 10.1099/00221287-142-4-937

Chaturvedi, V., Wong, B., and Newman, S. J. (1996b). Oxidative killing of *Cryptococcus neoformans* by human neutrophils. evidence that fungal mannitol protects by scavenging reactive oxygen intermediates. *J. Immunol.* 156, 3836–3840.

Chen, Y., Gallois-Montbrun, S., Schneider, B., Veron, M., Morera, S., Deville-Bonne, D., et al. (2003). Nucleotide binding to nucleoside diphosphate kinases: X-ray structure of human NDPK-A in complex with ADP and comparison to protein kinases. *J. Mol. Biol.* 332, 915–926. doi: 10.1016/j.jmb.2003.07.004

Chittur, S. V., Klem, T. J., Shafer, C. M., and Davisson, V. J. (2001). Mechanism for acivicin inactivation of triad glutamine amidotransferases. *Biochemistry* 40, 876–887. doi: 10.1021/bi0014047

Chitty, J. L., and Fraser, J. A. (2017). Purine acquisition and synthesis by human fungal pathogens. *Microorganisms* 5:33. doi: 10.3390/microorganisms5020033

Chitty, J. L., Tatzenko, T. L., Williams, S. J., Koh, Y. Q., Corfield, E. C., Butrler, M. S., et al. (2017). GMP synthase is required for virulence factor production and infection by *Cryptococcus neoformans*. *J. Biol. Chem.* 292, 3049–3059. doi: 10.1074/jbc.M116.767533

Colby, T. D., Vanderveen, K., Strickler, M. D., Markham, G. D., and Goldstein, B. M. (1999). Crystal structure of human type II inosine monophosphate

dehydrogenase: implications for ligand binding and drug design. *Proc. Natl. Acad. Sci. U.S.A.* 96, 3531–3536. doi: 10.1073/pnas.96.7.3531

Cooney, N. M., and Klein, B. S. (2008). Fungal adaptation to the mammalian host: it's a new world, after all. *Curr. Opin. Microbiol.* 11, 511–516. doi: 10.1016/j.mib.2008.09.018

Corina, D. L., and Munday, K. A. (1971). Studies on polyol function in *Aspergillus clavatus*: a role for mannitol and ribitol. *J. Gen. Microbiol.* 69, 221–227. doi: 10.1099/00221287-69-2-221

Dagenais, T. R. T., and Keller, N. P. (2009). Pathogenesis of *Aspergillus fumigatus* in invasive aspergillosis. *Clin. Microbiol. Rev.* 22, 447–465. doi: 10.1128/CMR.00055-08

Dasari, P., Koleci, N., Shopova, I. A., Warternberg, D., Beyersdorf, N., Dietrich, S., et al. (2019). Enolase From *Aspergillus fumigatus* is a moonlighting protein that binds the human plasma complement proteins factor H, FHL-1, C4BP, and plasminogen. *Front. Immunol.* 10:2573. doi: 10.3389/fimmu.2019.02573

De Backer, M. D., Nelissen, B., Logghe, M., Viaene, J., Loonen, I., Vandoninck, S., et al. (2001). An antisense-based functional genomics approach for identification of genes critical for growth of *Candida albicans*. *Nat. Biotechnol.* 19, 235–241. doi: 10.1038/85677

Dinamarco, T. M., Brown, N. A., Couto De Almeida, R. S., Alves De Castro, P., Savoldi, M., De Souza Goldman, M. H., et al. (2012). *Aspergillus fumigatus* calcineurin interacts with a nucleoside diphosphate kinase. *Microbes Infect.* 14, 922–929. doi: 10.1016/j.micinf.2012.05.003

Du Pré, S., Beckmann, N., Almeida, M. C., Sibley, G. E. M., Law, D., Brand, A. C., et al. (2018). Effect of the novel antifungal drug F901318 (olorofim) on growth and viability of *Aspergillus fumigatus*. *Antimicrob. Agents Chemother.* 62:e00231-18. doi: 10.1128/AAC.00231-18

Duquerroy, S., Camus, C., and Janin, J. (1995). X-ray structure and catalytic mechanism of lobster enolase. *Biochemistry* 34, 12513–12523. doi: 10.1021/bi00039a005

Edwards, S. R., Braley, R., and Chaffin, W. L. (1999). Enolase is present in the cell wall of *Saccharomyces cerevisiae*. *FEMS Microbiol. Lett.* 177, 211–216. doi: 10.1111/j.1574-6968.1999.tb13734.x

Ehinger, S., Schubert, W.-D., Bergmann, S., Hammerschmidt, S., and Heinz, D. W. (2004). Plasmin(ogen)-binding alpha-enolase from *Streptococcus pneumoniae*: crystal structure and evaluation of plasming(ogen)-binding sites. *J. Mol. Biol.* 343, 997–1005. doi: 10.1016/j.jmb.2004.08.088

Ermishkin, L. N., Kasumov, K. M., and Potzeluyev, V. M. (1976). Single ionic channels induced in lipid bilayers by polyene antibiotics amphotericin B and nystatine. *Nature* 262, 698–699. doi: 10.1038/262698a0

Floryk, D., and Thompson, T. C. (2010). Antiproliferative effects of AVN944, a novel inosine 5-monophosphate dehydrogenase inhibitor, in prostate cancer cells. *Int. J. Cancer* 123, 2294–2302. doi: 10.1002/ijc.23788

Funk, J., Schaarschmidt, B., Slesiona, S., Hallström, T., Horn, U., and Brock, M. (2016). The glycolytic enzyme enolase represents a plasminogen-binding protein on the surface of a wide variety of medically important fungal species. *Int. J. Med. Microbiol.* 306, 59–68. doi: 10.1016/j.ijmm.2015.11.005

Furman, P. A., Fyfe, J. A., St Clair, M. H., Weinhold, K., Rideout, J. L., Greeman, G. A., et al. (1986). Phosphorylation of 3'-azido-3'-deoxythymidine and selective interaction of the 5'-triphosphate with human immunodeficiency virus reverse transcriptase. *Proc. Natl. Acad. Sci. U.S.A.* 83, 8333–8337. doi: 10.1073/pnas.83.21.8333

Ghannoum, M. A., and Rice, L. B. (1999). Antifungal agents: mode of action, mechanisms of resistance, and correlation of these mechanisms with bacterial resistance. *Clin. Microbiol. Rev.* 12, 510–517. doi: 10.1128/CMR.12.4.501

Giraud, M. F., Georgescauld, F., Lascu, I., and Dautant, A. (2006). Crystal structures of S120G mutant and wild type of human nucleoside diphosphate kinase A in complex with ADP. *J. Bioenerg. Biomembr.* 38, 261–264. doi: 10.1007/s10863-006-9043-0

Gooday, G. W. (1995). "Cell membrane," in *The Growing Fungus*, eds N. A. R. Gow and G. M. Gadd (Dordrecht: Springer), 63–74.

Gow, N. A. R., Latgé, J. P., and Munro, C. A. (2017). The fungal cell wall: structure, biosynthesis, and function. *Microbiol. Spectr.* 5, 1–25. doi: 10.1128/9781555819583.ch12

Hata, M., Ishii, Y., Watanabe, E., Uoto, K., Kobayashi, S., Yoshida, K. I., et al. (2010). Inhibition of ergosterol synthesis by novel antifungal compounds targeting C-14 reductase. *Med. Mycol.* 48, 613–621. doi: 10.3109/13693780903390208

Hedstrom, L. (2009). IMP dehydrogenase: Structure, mechanism and inhibition. *Chem. Rev.* 109, 2903–2928. doi: 10.1021/cr900021w

Hof, H. (2001). Critical annotations to the use of azole antifungals for plant protection. *Antimicrob. Agents Chemother.* 45, 2987–2990. doi: 10.1128/AAC.45.11.2987-2990.2001

Hope, W., Morton, A., and Eisen, D. P. (2002). Increase in prevalence of nosocomial non-*Candida albicans* candidaemia and the association of *Candida krusei* with fluconazole use. *J. Hosp. Infect.* 50, 56–65. doi: 10.1053/jhin.2001.1131

Horikoshi, K., Iida, S., and Ikeda, Y. (1965). Mannitol and mannitol dehydrogenases in conidia of *Aspergillus oryzae*. *J. Bacteriol.* 89, 326–330. doi: 10.1128/JB.89.2.326-330.1965

Hosaka, T., Meguro, T., Yamato, I., and Shirakihara, Y. (2003). Crystal structure of *Enterococcus hirae* enolase at 2.8 Å resolution. *J. Biochem.* 133, 817–823. doi: 10.1093/jb/mvg104

Houšt, J., SpíŽek, J., and Havlíček, V. (2020). Antifungal drugs. *Metabolites* 10:106. doi: 10.3390/metabo10030106

Ji, H., Wang, J., Guo, J., Li, Y., Lian, S., Guo, W., et al. (2016). Progress in the biological function of alpha-enolase. *Anim. Nutr.* 2, 12–17. doi: 10.1016/j.aninu.2016.02.005

Jin, J.-H., Lee, K.-T., Hong, J., Lee, D., Jang, E.-H., Kim, J.-Y., et al. (2020). Genome-wide functional analysis of phosphatases in the pathogenic fungus *Cryptococcus neoformans*. *Nat. Commun.* 11:4212. doi: 10.1038/s41467-020-18028-0

Kang, H. J., Jung, S. K., Kim, S. J., and Chung, S. J. (2008). Structure of human alpha-enolase (hENO1), a multifunctional glycolytic enzyme. *Acta Crystallogr. D Biol. Crystallogr.* 64, 651–657. doi: 10.1107/S0907444908008561

Khanna, N., Stuehler, C., Lünemann, A., Wójtowicz, A., Bochud, P. Y., and Leibundgut-Landmann, S. (2016). Host response to fungal infections - how immunology and host genetics could help to identify and treat patients at risk. *Swiss Med. Wkly.* 146:w14350. doi: 10.4414/smw.2016.14350

Kidd, S. E., Chen, S. C.-A., Meyer, W., and Halliday, C. L. (2020). A new age in molecular diagnostics for invasive fungal disease: are we ready? *Front. Microbiol.* 10:2903. doi: 10.3389/fmicb.2019.02903

Kilstrup, M., Hammer, K., Jensen, P. R., and Martinussen, J. (2005). Nucleotide metabolism and its control in lactic acid bacteria. *FEMS Microbiol. Rev.* 29, 555–590. doi: 10.1016/j.fmrre.2005.04.006

Ko, H.-C., Hsiao, T.-Y., Chen, C.-T., and Yang, Y.-L. (2013). *Candida albicans* ENO1 null mutants exhibited altered drug susceptibility, hyphal formation, and virulence. *J. Microbiol.* 51, 345–351. doi: 10.1007/s12275-013-2577-z

Kozel, T. R., and Wickes, B. (2014). Fungal diagnostics. *Cold Spring Harb. Perspect. Med.* 4:a019299. doi: 10.1101/cshperspect.a019299

Kuhnel, K., and Luisi, B. (2001). Crystal structure of the *Escherichia coli* RNA degradosome component enolase. *J. Mol. Biol.* 313, 583–592. doi: 10.1006/jmbi.2001.5065

Kummari, L. K., Butler, M. S., Furlong, E., Blundell, R., Nouwens, A., Silva, A. B., et al. (2018). Antifungal benzo[b]thiophene 1,1-dioxide IMPDH inhibitors exhibit pan-assay interference (PAINS) profiles. *Bioorganic Med. Chem.* 26, 5408–5419. doi: 10.1016/j.bmc.2018.09.004

Latgé, J. P. (1999). *Aspergillus fumigatus* and *Aspergillosis*. *Clin. Microbiol. Rev.* 12, 310–350. doi: 10.1128/CMR.12.2.310

Law, R. H. P., Abu-Ssaydeh, D., and Whisstock, J. C. (2013). New insights into the structure and function of the plasminogen/plasmin system. *Curr. Opin. Struct. Biol.* 23, 836–841. doi: 10.1016/j.sbi.2013.10.006

Law, R. H. P., Caradoc-Davies, T., Cowieson, N. P., Horvath, A. J., Quek, A. J., Encanacao, J. A., et al. (2012). The X-ray crystal structure of full-length human plasminogen. *Cell Rep.* 1, 185–190. doi: 10.1016/j.celrep.2012.02.012

Lestrade, P., Buil, J. B., Van Der Beek, M. T., Kuijper, E. J., Van Dijk, K., Kampinga, G. A., et al. (2020). Paradoxal trends in azole-resistant *Aspergillus fumigatus* in a national multicenter surveillance program, the Netherlands, 2013–2018. *Emerg. Infect. Dis.* 26, 1447–1455. doi: 10.3201/eid2607.200088

Li, W. Q., Hu, X. C., Zhang, X., Ge, Y., Zhao, S., Hu, Y., et al. (2011). Immunisation with the glycolytic enzyme enolase confers effective protection against *Candida albicans* infection in mice. *Vaccine* 29, 5526–5533. doi: 10.1016/j.vaccine.2011.05.030

Lin, X., Momany, C., and Momany, M. (2003). SwoHp, a nucleoside diphosphate kinase, is essential in *Aspergillus nidulans*. *Eukaryotic Cell* 2, 1169–1177. doi: 10.1128/EC.2.6.1169-1177.2003

Liu, Q., Ying, S. H., Feng, M. G., and Jiang, X. H. (2008). Physiological implication of intracellular trehalose and mannitol changes in response of entomopathogenic fungus *Beauveria bassiana* to thermal stress. *Antonie Van Leeuwenhoek* 95, 65–75. doi: 10.1007/s10482-008-9288-1

López-Alemany, R., Longstaff, C., Hawley, S., Mirshahi, M., Fábregas, P., Jardí, M., et al. (2003). Inhibition of cell surface mediated plasminogen activation by a monoclonal antibody against alpha-Enolase. *Am. J. Hematol.* 72, 234–242. doi: 10.1002/ajh.10299

Markland, W., Mcquaid, T. J., Jain, J., and Kwong, A. D. (2000). Broad-spectrum antiviral activity of the IMP dehydrogenase inhibitor VX-497: a comparison with Ribavirin and demonstration of antiviral additivity with alpha interferon. *Antimicrob. Agents Chemother.* 44, 859–866. doi: 10.1128/AAC.44.4.859-866.2000

Marshall, A. C., Bond, C. S., and Bruning, J. B. (2018). Structure of *Aspergillus fumigatus* cytosolic thiolase: trapped tetrahedral reaction intermediates and activation by monovalent cations. *ACS Catal.* 8, 1973–1989. doi: 10.1021/acscatal.7b02873

Marshall, A. C., Kidd, S. E., Lamont-Friedrich, S. J., Arentz, G., Hoffman, P., Coad, B. R., et al. (2019). Structure, mechanism, and inhibition of *Aspergillus fumigatus* thioredoxin reductase. *Antimicrob. Agents Chemother.* 63:e02281-18. doi: 10.1128/AAC.02281-18

Marshall, A. C., Kroker, A. J., Murray, L. A. M., Gronthos, K., Rajapaksha, H., Wegener, K. L., et al. (2017). Structure of the sliding clamp from the fungal pathogen *Aspergillus fumigatus* (AfumPCNA) and interactions with Human p21. *FEBS J.* 284, 985–1002. doi: 10.1111/febs.14035

Mazu, T. K., Bricker, B. A., Flores-Rozas, H., and Ablordeppey, S. Y. (2016). The mechanistic targets of antifungal agents: an overview. *Mini Rev. Med. Chem.* 16, 555–578. doi: 10.2174/1389557516666160118112103

Meena, M., Prasad, V., Zehra, A., Gupta, V. K., and Upadhyay, R. S. (2015). Mannitol metabolism during pathogenic fungal-host interactions under stressed conditions. *Front. Microbiol.* 6:1019. doi: 10.3389/fmicb.2015.01019

Megson, G. M., Stevens, D. A., Hamilton, J. R., and Denning, D. W. (1996). D-mannitol in cerebrospinal fluid of patients with AIDS and cryptococcal meningitis. *J. Clin. Microbiol.* 34, 218–221. doi: 10.1128/JCM.34.1.218-221.1996

Miura, N., Kirino, A., Endo, S., Morisaka, H., Kuroda, K., Takagi, M., et al. (2012). Tracing putative trafficking of the glycolytic enzyme enolase via SNARE-driven unconventional secretion. *Eukaryotic Cell* 11, 1075–1082. doi: 10.1128/EC.00075-12

Morrow, C. A., Valkov, E., Stamp, A., Chow, E. W. L., Lee, I. R., Wronski, A., et al. (2012). De novo GTP biosynthesis is critical for virulence of the fungal pathogen *Cryptococcus neoformans*. *PLoS Pathog.* 8:e1002957. doi: 10.1371/journal.ppat.1002957

Nguyen, S., Jovcevski, B., Pukala, T. L., and Bruning, J. B. (in press). Nucleoside selectivity of Aspergillus fumigatus nucleoside-diphosphate kinase. *FEBS J.* doi: 10.1111/febs.15607. [Epub ahead of print].

Nucci, M., and Marr, K. A. (2005). Emerging fungal diseases. *Clin. Infect. Dis.* 41, 521–526. doi: 10.1086/432060

Nucci, M., and Perfect, J. R. (2008). When primary antifungal therapy fails. *Clin. Infect. Dis.* 46, 1426–1433. doi: 10.1086/587101

Oliver, J. D., Sibley, G. E. M., Beckmann, N., Dobb, K. S., Slater, M. J., Mcentee, L., et al. (2016). F901318 represents a novel class of antifungal drug that inhibits dihydroorotate dehydrogenase. *Proc. Natl. Acad. Sci. U.S.A.* 113, 12809–12814. doi: 10.1073/pnas.1608304113

Pancholi, V., and Fischetti, V. A. (1998). Alpha-enolase, a novel strong plasmin(ogen) binding protein on the surface of pathogenic *Streptococci*. *J. Biol. Chem.* 273, 14503–14515. doi: 10.1074/jbc.273.23.14503

Park, Y., Pacitto, A., Bayliss, T., Cleghorn, L. A., Wang, Z., Hartman, T., et al. (2018). Essential but not vulnerable: indazole sulfonamides targeting inosine monophosphate dehydrogenase as potential leads against *Mycobacterium tuberculosis*. *ACS Infect Dis.* 3, 18–33. doi: 10.1021/acsinfecdis.6b00103

Perlin, D. S. (2015). Echinocandin resistance in Candida. *Clin. Infect. Dis.* 61, S612–S617. doi: 10.1093/cid/civ791

Perlin, D. S., Shor, E., and Zhao, Y. (2015). Update on antifungal drug resistance. *Curr. Clin. Microbiol. Rep.* 2, 84–95. doi: 10.1007/s40588-015-0015-1

Pfaller, M. A., Diekema, D. J., Mendez, M., Kibbler, C., Erzsebet, P., Chang, S.-C., et al. (2006). Candida guilliermondii, an opportunist fungal pathogen with decrease susceptibility to fluconazole: geographic and temporal trends from the ARTEMIS DISK antifungal surveillance program. *J. Clin. Microbiol.* 44, 3551–3556. doi: 10.1128/JCM.00865-06

Rodriguez-Suarez, R., Xu, D., Veillette, K., Davison, J., Sillaots, S., Kauffman, S., et al. (2007). Mechanism-of-action determination of GMP synthase inhibitors and target validation in *Candida albicans* and *Aspergillus fumigatus*. *Chem. Biol.* 14, 1163–1175. doi: 10.1016/j.chembiol.2007.09.009

Ruijter, G. J. G., Bax, M., Patel, H., Flitter, S. J., Van De Vondervoort, P. J. I., De Vries, R. P., et al. (2003). Mannitol is required for stress tolerance in *Aspergillus niger* conidiospores. *Eukaryotic Cell* 2, 690–698. doi: 10.1128/EC.2.4.690-698.2003

Ryder, N. S. (1988). Mechanism of action and biochemical selectivity of allylamine antimycotic agents. *Ann. N. Y. Acad. Sci.* 544, 208–220. doi: 10.1111/j.1749-6632.1988.tb40405.x

Sabiiti, W., and May, R. C. (2012). Mechanisms of infection by the human fungal pathogen *Cryptococcus neoformans*. *Future Microbiol.* 7, 1297–1313. doi: 10.2217/fmb.12.102

Schmiedel, Y., and Zimmerli, S. (2016). Common invasive fungal diseases: an overview of invasive candidiasis, aspergillosis, cryptococcosis, and Pneumocystis pneumonia. *Swiss Med. Wkly.* 14:w14281. doi: 10.4414/smw.2016.14281

Semianrio-Vidal, L., Van Hesuden, C., Mugesh, G., and Lazarowski, E. R. (2010). Ebselen is a potent non-competitive inhibitor of extracellular nucleoside diphosphokinase. *Purinerg. Signal* 6, 383–391. doi: 10.1007/s11302-010-9203-x

Shi, M., and Mody, C. H. (2016). Fungal infection in the brain: what we learned from intravital imaging. *Front. Immunol.* 7:292. doi: 10.3389/fimmu.2016.00292

Shibasaki, S., Karasaki, M., Tafuku, S., Aoki, W., Sewaki, T., and Ueda, M. (2014). Oral immunization against candidiasis using *Lactobacillus casei* displaying enolase 1 from *Candida albicans*. *Sci. Pharm.* 82, 697–708. doi: 10.3797/scipharm.1404-07

Shu, Q., and Nair, V. (2008). Inosine monophosphate dehydrogenase (IMPDH) as a target in drug discovery. *Med. Res. Rev.* 28, 219–232. doi: 10.1002/med.20104

Sperling, O. (1988). "Human purine metabolism," in *Myocardial Energy Metabolism*, ed J. W. De Jong (Dordrecht: Springer), 225–236.

Stec, B., and Lebioda, L. (1990). Refined structure of yeast apo-enolase at 2.25 Å resolution. *J. Mol. Biol.* 211, 235–248. doi: 10.1016/0022-2836(90)90023-F

Thammahong, A., Puttikamonkul, S., Perfect, J. R., Brennan, R. G., and Cramer, R. A. (2017). Central role of the trehalose biosynthesis pathway in the pathogenesis of human fungal infections: opportunities and challenges for therapeutic development. *Microbiol. Mol. Biol. Rev.* 81:e00053-16. doi: 10.1128/MMBR.00053-16

Trapero, A., Pacitto, A., Singh, V., Sabbah, M., Coyne, A. G., Mizrahi, V., et al. (2018). Fragment-based approach to targeting inosine-5'-monophosphate dehydrogenase (IMPDH) from *Mycobacterium tuberculosis*. *J. Med. Chem.* 61, 2806–2822. doi: 10.1021/acs.jmedchem.7b01622

Umejiego, N. N., Gollapalli, D., Sharling, L., Volftsun, A., Lu, J., Benjamin, N. N., et al. (2008). Targeting a prokaryotic protein in a eukaryotic pathogen: identification of lead compounds against *Cryptosporidiosis*. *Chem. Biol.* 15, 70–77. doi: 10.1016/j.chembiol.2007.12.010

Urbina, J. M., Cortés, J. C. G., Palma, A., López, S. N., Zacchino, S. A., Enriz, R. D., et al. (2000). Inhibitors of the fungal cell wall. Synthesis of 4-aryl-4-N-arylamine-1-butenes and related compounds with inhibitory activities on β(1–3) glucan and chitin synthases. *Bioorganic Med. Chem.* 8, 691–698. doi: 10.1016/S0968-0896(00)00003-1

Vandeputte, P., Ferrari, S., and Coste, A. T. (2012). Antifungal resistance and new strategies to control fungal infections. *Int. J. Microbiol.* 2012:713687. doi: 10.1155/2012/713687

Walsh, T. J., Dixon, D. M. (1996). "sSpectrum of mycoses," in *Medical Microbiology*, 4th Edn, eds S. Baron (Galveston, TX: University of Texas Medical Branch at Galveston).

Wang, Y., Wang, S., Nie, X., Yang, K., Xu, P., Wang, X., et al. (2019). Molecular and structural basis of nucleoside diphosphate kinase–mediated regulation of spore and sclerotia development in the fungus *Aspergillus flavus*. *J. Biol. Chem.* 294, 12415–12431. doi: 10.1074/jbc.RA119.007505

Wang, Y., Wang, S., Nie, X., Yang, K., Xu, P., Wang, X., et al. (2019). Molecular and structural basis of nucleoside diphosphate kinase-mediated regulation of spore

and sclerotia development in the fungus Aspergillus flavus. *J. Biol. Chem.* 294, 12415–12431.

Welin, M., Lehtio, L., Johansson, A., Flodin, S., Nyman, T., Tresaugues, L., et al. (2013). Substrate specificty and oligomerization of human GMP synthetase. *J. Mol. Biol.* 425, 4323–4333. doi: 10.1016/j.jmb.2013.06.032

Wenzel, R. P., and Gennings, C. (2005). Bloodstream infections due to Candida species in the intensive care unit: identifying especially high-risk patients to determine prevention strategies. *Clin. Infect. Dis.* 41, S389–S393. doi: 10.1086/430923

Wiederhold, N. P., and Lewis, R. E. (2003). The echinocandin antifungals: an overview of the pharmacology, spectrum and clinical efficacy. *Expert Opin. Investig. Drugs* 12, 1313–1333. doi: 10.1517/13543784.12.8.1313

Witteveen, C. F. B., and Visser, J. (1995). Polyol pools in *Aspergillus niger*. *FEMS Microbiol. Lett.* 134, 57–62. doi: 10.1111/j.1574-6968.1995.tb07914.x

Wong, B., Brauer, K. L., Tsai, R. R., and Jayasimhulu, K. (1989). Increased amounts of the Aspergillus metabolite D-mannitol in serum of rats with experimental aspergillosis. *J. Infect. Dis.* 160, 95–103. doi: 10.1093/infdis/160.1.95

Wong, B., Perfect, J. R., Beggs, S., and Wright, K. A. (1990). Production of the hexitol D-mannitol by *Cryptococcus neoformans in vitro* and in rabbits with experimental meningitis. *Infect. Immun.* 58, 1664–1670. doi: 10.1128/IAI.58.6.1664-1670.1990

Wyatt, T. T., Van Leeuwen, M. R., Wösten, H. A. B., and Dijksterhuis, J. (2014). Mannitol is essential for the development of stress-resistance ascospores in *Neosartorya fischeri* (*Aspergillus fischeri*). *Fungal Genet. Biol.* 64, 11–24. doi: 10.1016/j.fgb.2013.12.010

Conflict of Interest: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Appendix B (Publication):

Derivatization of protein crystals with I3C using random microseed matrix screening

Statement of Authorship

Title of Paper	Derivatization of protein crystals with I3C using random microseed matrix screening
Publication Status	☑ Published ☐ Accepted for Publication ☐ Submitted for Publication ☐ Unpublished and Unsubmitted work written in manuscript style
Publication Details	This publication describes a new technique to produce derivatised protein crystals using random microseed matrix screening whilst incorporating the phasing molecule, I3C.

Principal Author

Name of Principal Author (Candidate)	Stephanie Nguyen		
Contribution to the Paper	Contributed to data analysis, production of the video component, reviewed and edited the manuscript.		
Overall percentage (%)	40		
Certification:	This paper reports on original research I conducted during the period of my Higher Degree by Research candidature and is not subject to any obligations or contractual agreements with a third party that would constrain its inclusion in this thesis. I am the primary author of this paper.		
Signature		Date	19/07/2021

Co-Author Contributions

By signing the Statement of Authorship, each author certifies that:

i. the candidate's stated contribution to the publication is accurate (as detailed above);

ii. permission is granted for the candidate in include the publication in the thesis; and

iii. the sum of all co-author contributions is equal to 100% less the candidate's stated contribution.

Name of Co-Author	Dr. Jia Quyen Truong		
Contribution to the Paper	Project conceptualisation, planned and performed experiments, analysed data, assisted in the production of the video component, reviewed and edited the manuscript.		
Overall Percentage (%)	40		
Signature		Date	19/07/21

Name of Co-Author	Assoc. Prof. Keith Shearwin
Contribution to the Paper	Reviewed and edited the manuscript.
Overall Percentage (%)	10
Signature	Date: 19/07/2021

Name of Co-Author	Dr. John B. Bruning
Contribution to the Paper	Reviewed and edited the manuscript.
Overall Percentage (%)	10
Signature	Date: 21/7/21

jove

Derivatization of Protein Crystals with I3C using Random Microseed Matrix Screening

Jia Quyen Truong[*,1], **Stephanie Nguyen**[*,1,2], **John B. Bruning**[1,2], **Keith E. Shearwin**[1]

[1] School of Biological Sciences, The University of Adelaide [2] Institute of Photonics and Advanced Sensing (IPAS), School of Biological Sciences, The University of Adelaide

[*] These authors contributed equally

Corresponding Author

Keith E. Shearwin
keith.shearwin@adelaide.edu.au

Citation

Truong, J.Q., Nguyen, S., Bruning, J.B., Shearwin, K.E. Derivatization of Protein Crystals with I3C using Random Microseed Matrix Screening. *J. Vis. Exp.* (167), e61894, doi:10.3791/61894 (2021).

Date Published

January 16, 2021

DOI

10.3791/61894

URL

jove.com/video/61894

Abstract

Protein structure elucidation using X-ray crystallography requires both high quality diffracting crystals and computational solution of the diffraction phase problem. Novel structures that lack a suitable homology model are often derivatized with heavy atoms to provide experimental phase information. The presented protocol efficiently generates derivatized protein crystals by combining random microseeding matrix screening with derivatization with a heavy atom molecule I3C (5-amino-2,4,6-triiodoisophthalic acid). By incorporating I3C into the crystal lattice, the diffraction phase problem can be efficiently solved using single wavelength anomalous dispersion (SAD) phasing. The equilateral triangle arrangement of iodine atoms in I3C allows for rapid validation of a correct anomalous substructure. This protocol will be useful to structural biologists who solve macromolecular structures using crystallography-based techniques with interest in experimental phasing.

Introduction

In the field of structural biology, X-ray crystallography is regarded as the gold standard technique to determine the atomic-resolution structures of macromolecules. It has been utilized extensively to understand the molecular basis of diseases, guide rational drug design projects and elucidate the catalytic mechanism of enzymes[1,2]. Although structural data provides a wealth of knowledge, the process of protein expression and purification, crystallization and structure determination can be extremely laborious. Several bottlenecks are commonly encountered that hinder the progress of these projects and this must be addressed to efficiently streamline the crystal structure determination pipeline.

Following recombinant expression and purification, preliminary conditions that are conducive to crystallization must be identified which is often an arduous and time-consuming aspect of X-ray crystallography. Commercial

sparse matrix screens that consolidate known and published conditions have been developed to ease this bottleneck[3,4]. However, it is common to generate few hits from these initial screens despite using highly pure and concentrated protein samples. Observing clear drops indicates that the protein may not be reaching the levels of supersaturation required to nucleate a crystal. To encourage crystal nucleation and growth, seeds produced from pre-existing crystals can be added to the conditions and this allows for increased sampling of the crystallization space. Ireton and Stoddard first introduced the microseed matrix screening method[5]. Poor quality crystals were crushed to make a seed stock and then added systematically to crystallization conditions containing different salts to generate new diffraction-quality crystals that would not have otherwise formed. This technique was further improved by D'Arcy et al. who developed random microseed matrix screening (rMMS) in which seeds were introduced into a spare matrix crystallization screen[6,7]. This improved the quality of crystals and increased the number of crystallization hits on average by a factor of 7.

After crystals are successfully produced and an X-ray diffraction pattern is obtained, another bottleneck in the form of solving the 'phase problem' is encountered. During the data acquisition process, the intensity of diffraction (proportional to the square of the amplitude) is recorded but the phase information is lost, giving rise to the phase problem that halts immediate structure determination[8]. If the target protein shares high sequence identity to a protein with a previously determined structure, molecular replacement can be used to estimate the phase information[9,10,11,12]. Although this method is fast and inexpensive, model structures may not be available or suitable. The success of the homology model-based molecular replacement method drops significantly as sequence identity falls below 35%[13]. In the absence of a suitable homology model, *ab initio* methods, such as ARCIMBOLDO[14,15] and AMPLE[16], can be tested. These methods use computationally predicted models or fragments as starting points for molecular replacement. AMPLE, which uses predicted decoy models as starting points, struggles to solve structures of large (>100 residues) proteins and proteins containing predominately β-sheets. ARCIMBOLDO, which attempts to fit small fragments to extend into a larger structure, is limited to high resolution data (≤2 Å) and by the ability of algorithms to expand the fragments into a full structure.

If molecular replacement methods fails, direct methods such as isomorphous replacement[17,18] and anomalous scattering at a single wavelength (SAD[19]) or multiple wavelengths (MAD[20]) must be used. This is often the case for truly novel structures, where the crystal must be formed or derivatized with a heavy atom. This can be achieved by soaking or co-crystallizing with a heavy atom compound, chemical modification (such as 5-bromouracil incorporation in RNA) or labelled protein expression (such as incorporating selenomethionine or selenocysteine amino acids into the primary structure)[21,22]. This further complicates the crystallization process and requires additional screening and optimization.

A new class of phasing compounds, including I3C (5-amino-2,4,6-triiodoisophthalic acid) and B3C (5-amino-2,4,6-tribromoisophthalic acid), offer exciting advantages over pre-existing phasing compounds[23,24,25]. Both I3C and B3C feature an aromatic ring scaffold with an alternating arrangement of anomalous scatters required for direct phasing methods and amino or carboxylate functional groups that interact specifically with the protein and provide binding site specificity. The subsequent equilateral triangular

arrangement of heavy metal groups allows for simplified validation of the phasing substructure. At the time of writing, there are 26 I3C-bound structures in the Protein Data Bank (PDB), of which 20 were solved using SAD phasing[26].

This protocol improves the efficacy of the structure determination pipeline by combining the methods of heavy metal derivatization and rMMS screening to simultaneously increase the number of crystallization hits and simplify the crystal derivatization process. We demonstrated this method was extremely effective with hen egg white lysozyme and a domain of a novel lysin protein from bacteriophage P68[27]. Structure solution using the highly automated Auto-Rickshaw structure determination pipeline is described, specifically tailored for the I3C phasing compound. There exists other automated pipelines that can be used such as AutoSol[28], ELVES[29] and CRANK2[30]. Non-fully automated packages such as SHELXC/D/E can also be used[31,32,33]. This method is particularly beneficial to researchers who are studying proteins lacking homologous models in the PDB, by significantly reducing the number of screening and optimization steps. A prerequisite for this method is protein crystals or a crystalline precipitate of the target protein, obtained from previous crystallization trials.

Protocol

1. Experimental planning and considerations

1. Use pre-existing crystals of the protein of interest, preferably generated through vapor diffusion crystallization. For a generalized protocol of vapor diffusion crystallization, see Benvenuti and Mangani[34]. Other methods of crystallization such as microbatch under oil and free interface diffusion will require harvesting the crystals prior to crushing to generate microseeds.
2. In the preparation of a seed stock, use the highest quality crystals that can be sacrificed. The highest quality crystal can be judged visually based on morphology or the best diffracting crystal can be selected, if such data is available. It is very likely that even better quality crystals are obtained after optimization through seeding. In the case where no crystals are available, crystalline precipitate such as spherulites and needles can be used.
3. Identify salt crystals. Salt crystals can grow in crystallization screens and can look like protein crystals. Using salt crystals in rMMS will provide no benefit and will waste precious sample, so it is important to eliminate salt false positives.
    1. Salt crystals are loud when they are crushed. Crystals must be crushed to generate a seed stock, so this strategy is particularly relevant. If an audible crack sound is heard when crushing up the crystals, the crystal is likely to be salt.
    2. If the protein contains tryptophan and tyrosine residues, use ultraviolet fluorescence microscopy to identify protein crystals which fluoresce under these lighting conditions.
    3. Use Izit dye (methylene blue) to stain protein crystals to differentiate them to salt crystals which remain relatively unstained. However, this procedure is more destructive and is only recommended if one has crystals to spare from replicates of the same drop.

        NOTE: Although the aforementioned tests may give promising results, salt crystals may still be mistaken for protein crystals. In this case, diffraction

experiments can be used to definitively discern between a protein and salt crystal.

2. Preparation of lithium I3C stock

1. Measure out 120 mg of I3C (5-amino-2,4,6-triiodoisophthalic acid) into a 1.5 mL microcentrifuge tube.
2. Dissolve I3C in 200 µL of 2 M lithium hydroxide. The solution can be gently heated using a heat block at 40-60 °C to encourage dissolution. The resulting lithium I3C solution should be brown and has a concentration of 1 M. CAUTION: Lithium hydroxide is corrosive. Safety glasses, gloves and a lab coat should be worn.
3. Measure the pH of the solution. If necessary, add small amounts of 1 M hydrochloric acid or 2 M lithium hydroxide to adjust the pH to between 7 and 8. Add milliQ water to make the final solution volume to 400 µL. The concentration of the I3C stock solution is 0.5 M.
   NOTE: Step 2.3 is optional. The pH of the solution should be between pH 7-8 prior to any pH adjustment. This step should be performed if the protein of interest is strongly affected by pH. The protocol can be paused here. Lithium I3C can be kept in the dark at 4 °C for at least two weeks[35].

3. Addition of I3C to the protein stock

1. Method 1
   1. Add stock lithium I3C to a 150 µL aliquot of the target protein. The final concentration should be between 5-40 mM lithium I3C.
2. Method 2 (gentler method)
   1. Prepare a protein dilution buffer that matches the buffer of the target protein. To this dilution buffer, add stock lithium I3C to give a concentration of lithium I3C between 10-80 mM.
   2. Dilute the protein 1:1 with protein dilution buffer to give a final concentration of lithium I3C between 5-40 mM.
      NOTE: Some proteins will precipitate upon coming into contact with high concentrations of lithium I3C in method 1, while other proteins can tolerate it. Method 2 reduces the likelihood of precipitation. However, this method halves the protein concentration. For proteins that do not have an established crystallization protocol, a protein concentration of 10 mg/mL is generally recommended for initial crystallization screening. An initial molar ratio of I3C to protein of 8 is recommended. Protein concentration and molar ratio of I3C to protein can be optimized after the initial screen.

4. Making a seed stock

1. Make a rounded probe for crushing crystals.
   1. With a Bunsen burner on the blue flame, heat a Pasteur pipette towards its middle. Using a tweezer, pull the end of the Pasteur pipette to draw it out into a thin diameter of less than 0.3 mm.
   2. Once the midsection is thin enough, hold that segment in the flame to separate the pipette at this point and round the end of the pipette to finish the glass probe.
      NOTE: Rounded probe crystal crushers are sold by third party vendors. These are an alternative to making rounded probes.
2. Place five 1.5 mL microcentrifuge tubes on ice.

3. Under a light microscope, examine the crystallization tray for a suitable condition to generate microcrystals. Ideally, good morphology large crystals are selected. However, this technique also works with poor morphology crystals, needles, plates, microcrystals and spherulites.

4. Open up the crystallization tray well. For 96 well crystallization trays sealed with plastic, use a scalpel to cut the plastic sealing the well. For hanging drop trays sealed with grease, the coverslip can be removed using tweezers and inverted onto an even surface.

5. Transfer 70 µL of reservoir solution to a microcentrifuge tube and chill it on ice. To the other microcentrifuge tubes, add 90 µL of reservoir solution and return to ice to chill.
NOTE: If the reservoir does not have enough volume or does not exist (in the case of microbatch under oil), create crystallization reservoir by mixing the appropriate reagents.

6. Agitate the crystal in the drop using the crystal probe to thoroughly crush it up. The crystal needs to be completely crushed up which can be monitored under the microscope.

7. Remove all the liquid from the drop and transfer it to the microcentrifuge tube with the reservoir solution. Mix and subsequently take 2 µL of mixture from the microcentrifuge tube and add it back to the well. Rinse the well with the solution and transfer it to the microcentrifuge tube. Repeat this rinse step once more. From this point on, keep the microcentrifuge tube cold to avoid melting the microseeds in the mixture.

8. Vortex the tube at maximum speed at 4 °C for 3 min, stopping regularly to chill the tube on ice to prevent overheating.
NOTE: Some microseeding protocols add a polytetrafluoroethylene seed bead to the microcentrifuge tube to aid crystal crushing[7,36]. We have employed the technique without the use of a seed bead with success, but see no problems with utilizing a seed bead to crush up crystals.

9. Make a 1 in 10 serial dilution of the seed stock by sequentially transferring 10 µL between the chilled reservoir solutions.

10. Store seed stocks that will not be used immediately at -80 °C.

5. Setting up an rMMS screen

1. Setting up a 96 well screening plate using a liquid dispensing robot. In the absence of a robot, a multichannel pipette may also be used.
    1. Transfer 75 µL from a deep well block to a 96 well crystallization tray. Add 1 µL to the crystallization drop and 74 µL to the reservoir.
    2. Transfer 1 µL of protein supplemented with lithium I3C, made in step 2, to the crystallization drop.
    3. Transfer 0.1 µL of seed stock to the crystallization drop.
    4. Seal the plate with clear sealing tape and incubate the plate at a constant temperature to allow crystal growth.
2. Setting up a hanging drop screens
    1. Grease the edges of the hanging drop wells (hanging drop crystallization trays can be found in 24 and 48 well formats).
    2. Transfer 500 µL crystallization solution into reservoir.

3. Near the center of a glass cover slide, place a 1 µL drop of protein supplemented with lithium I3C, made in step 2.
4. Add 1 µL of the crystallization solution to the drop.
5. Transfer 0.1 µL of seed stock to the crystallization drop.
6. Invert the cover slide and seal the crystallization well by pushing the cover slide into the grease.
7. Incubate the plate at a constant temperature to allow crystal growth.
   NOTE: With new and untested seed stocks, it is recommended to use the most concentrated seed stock to maximize the chances of getting crystallization hits. Subsequent conditions can be set up with reduced seed concentration to optimize the number of crystals.

3. Inspect crystal trays under a microscope regularly for crystal growth. If crystals are of sufficient quality, they can be harvested for data collection. Crystals can also be used to generate new seed stocks and new rMMS screens to allow for iterative optimization.

6. Data collection

1. Harvest crystals using cryoloops, cryoprotect the crystals and flash cool them in liquid nitrogen. For additional information on flash cooling crystals, refer to Teng[37] and Garman and Mitchell[38].
   1. During the cryoprotection stage, if the crystal is passed through a new aqueous solution, I3C can be lost from the crystal due to it leeching into the cryoprotection solution. Use lithium I3C in the cryoprotection solution at a concentration that matches the crystallization condition to mitigate this.
   2. Crystals grown using this protocol have successfully been cryoprotected using Parabar 10312 oil based cryoprotectant (Hampton Research).
      NOTE: The protocol can be paused here while crystals are stored in liquid nitrogen.
      CAUTION: Liquid nitrogen can cause cold burns. Liquid nitrogen can also cause asphyxiation if used in enclosed spaces.
2. Mount the crystal on the X-ray source goniometer and collect diffraction data using the protocol specific for the X-ray source.
3. This technique relies on anomalous signal from iodine atoms in I3C. Thus, select the energy of the X-ray to maximize this signal.
   1. Set synchrotron X-ray sources with tunable energies as low as possible. For many macromolecular crystallography beamlines, the lowest configurable energy is 8000 to 8500 eV.
   2. Rotating anode X-ray sources cannot be tuned. Commonly used anode sources with copper have the Kα edge at 8046 eV, which provides a good anomalous signal for iodine (f" = 6.9 e). Anode sources with chromium have a Kα edge at 5415 eV, which provides a large anomalous signal for iodine (f" = 12.6 e).
4. Radiation damage is a significant problem during data collection as it will degrade the anomalous signal[39]. Select the exposure time and attenuation of the beam to achieve the best diffraction while minimizing radiation dose.

NOTE: In a similar phasing compound with the iodine atoms replaced with bromine atoms, radiation damage has been shown to cause the radiolysis of the carbon bromine bond and a reduction in the occupancy of the bromine atoms[24].

1. Use inverse beam SAD data collection as a collection strategy. The data is collected in wedges, with opposite wedges collected after each other. This allows Friedel pairs to be collected with an equivalent dose, resulting in an improved measurement of anomalous signal less affected by radiation damage. For example, an eight wedge strategy to collect 360° would involve collecting the data in the order of wedge 1 (0°-45°), wedge 2 (180°-225°), wedge 3 (46°-90°), wedge 4 (225°-270°), wedge 5 (90°-135°), wedge 6 (270°-315°), wedge 7 (135°-180°) and wedge 8 (315°-360°).

   NOTE: Continuous rotation is an alternative collection strategy to that of inverse beam data collection. For a recent comparison of the collection strategies, see Garcie-Bonte & Katona[40].

7. Data processing and structure solution

1. Perform data reduction on the diffraction data using XDS[41], with the aim of maximizing the anomalous signal. Data reduction input parameters are specific to the dataset and may require some trial and error. Here are some recommendations to start.
   1. Set FRIEDEL'S LAW=FALSE. Execute CORRECT twice, setting STRICT_ABSORPTION_CORRECT = TRUE and STRICT_ABSORPTION_CORRECT = FALSE. One run can have a higher anomalous signal than the other. Compare the anomalous signals between the runs using the 'Anomal Corr' and 'SigAno' disciplines in the output. This provides an indicator of data quality.
   2. Run SHELXC on the XDS_ASCII.HKL file for a more accurate indication of anomalous signal. The 'Ranom' discipline will give an indication of the anomalous signal at different resolutions.
2. Run POINTLESS[42] and AIMLESS[43] to scale the data. In AIMLESS, set the parameter ANOMALOUS ON. If the GUI is used, select the option **Separate anomalous pairs for outlier rejection and merging statistics**. Testing different resolution cutoffs may be required to maximize anomalous signal.
3. Solve the protein structure using Auto-Rickshaw automated crystal structure determination pipeline[44]. Auto-Rickshaw will attempt to solve the phase problem and build the crystal structure of the protein automatically with protein modelling and refinement software.
   1. For proteins without a homology model template, run the SAD protocol of Auto-Rickshaw in Advanced Mode. Enter the required parameters.
      1. Select PROTEIN as the molecule type.
      2. Enter the data collection wavelength in angstroms (Å).
      3. Select "I" as substructure element to indicate iodine atoms was used.
      4. Select "i3c" as substructure type to indicate I3C was the phasing molecule.
      5. Select "sub_direct" as the substructure determination method. This method employs SHELXD[32] to search for the substructure.

6. Select "3" as the number of expected substructure per monomer.
7. Enter "1" as the resolution cutoff of substructure search. This allows Auto-Rickshaw to automatically determine a suitable resolution cutoff.
8. Enter the number of residues in a single monomer, spacegroup of the dataset, and number of molecules in the asymmetric unit based on the Matthews coefficient.
9. Select the appropriate dissemination level of X-ray data that suits the needs. Selecting "AutoRickshaw developers" will allow Auto-Rickshaw developers to troubleshoot the run if problems arise.
10. Input the anomalous data as an mtz file.
11. Input the protein sequence as a seq, pir or txt file. A seq file can be generated in a text editor (such as Notepad++[9] on Windows or nano in Linux). Create a new file, enter the primary sequence of the protein as one long line or separated by line breaks. Save the file with the .seq file extension.
12. Enter an institutional email address.

4. Results are delivered via a web-link sent to the email address provided.

   NOTE: AutoRickshaw is an automated pipeline that invokes various crystallography software packages to solve an X-ray crystal structure[32,33,45,46,47,48,49,50,51,52,53,54,55,56,57,58].

   If the Auto-Rickshaw run fails to solve the structure, other Auto-Rickshaw settings can be tested. The structure determination method can be changed to "sub_phassade" to use Phaser[59] instead of SHELXD[32]. The number of expected substructure per monomer can be also increased or decreased.

5. During the experimental phasing of the crystal structure, Auto-Rickshaw will attempt to position heavy atoms in the unit cell, creating a substructure. The equilateral triangle arrangement of iodine atoms in I3C presents an efficient way of validating the substructure. If step 6.3 fails, validating the substructure could aid in troubleshooting structure solution.
   1. Download the list of heavy atom sites from the Auto-Rickshaw results page. It is a hyperlink called "heavy atom sites". This will download a text file with the heavy atom sites.
   2. Change the file extension of the file from .txt to .pdb.
   3. Open the PDB file in Coot[60]. Turn on symmetry to see other heavy atoms from neighboring asymmetric units.
   4. Measure the distances between the heavy atoms, including across asymmetric units. I3C will appear as an equilateral triangle with a side length of 6 angstroms. The presence of a triangle with these dimensions indicates the placements of those heavy atoms are correct.

Representative Results

Incorporating I3C into rMMS can generate new conditions supporting derivatized crystal growth

The efficacy of simultaneous rMMS screening and I3C derivatization was demonstrated in two proteins, hen egg white lysozyme (HEWL, obtained as a lyophilized powder) and the putative Orf11 lysin N-terminal domain (Orf11 NTD) from bacteriophage P68. Each protein was screened against PEG/ION HT under four different conditions including: unseeded, seeded, unseeded with I3C and seeded with I3C (**Figure 1**). For both proteins, the sole addition of I3C did not increase the number of conditions conducive to crystallization. In the case of Orf11 NTD, only one suitable condition was identified with and without I3C (**Figure 1B**). When I3C was added to the HEWL screens, the number of hits was reduced from 31 to 26, highlighting the added complexities of crystallisation when introducing phasing compounds (**Figure 1A**). Consistent with other studies, adding seed to commercial sparse matrix screens to generate an rMMS screen significantly increased the number of possible crystallization conditions for both proteins, resulting in a 2.1 and 6 fold increase for HEWL and Orf11 NTD, respectively[6,61] (**Figure 1**). Most importantly, simultaneous addition of I3C and seed increased the number of hits relative to an unseeded screen, demonstrating a 2.3 and 7 fold increase for HEWL and Orf11 NTD, respectively. Many of the crystals from rMMS in the presence of I3C show excellent crystal morphology (**Figure 2**).

Seeding allows careful control of crystal number in I3C rMMS screens

In microseeding experiments, the number of seeds introduced into a crystallization trial can be controlled by dilution of the seed stock and this allows for precise control of nucleation in the drop[7,36]. This often allows larger crystals to form since there is reduced competition of protein molecules at nucleation sites. This advantage also extends to the I3C-rMMS method and has been demonstrated successfully in both HEWL and Orf11 NTD. Recreation of a crystallization condition identified from the I3C-rMMS screen with a diluted seed stock yielded fewer but larger crystals (**Figure 3**).

SAD phasing can be used to solve the structures from crystals derived from rMMS I3C screen

Crystals grown using the diluted seed stock shown in **Figure 3** were used to solve the structure of the proteins using SAD phasing using diffraction data from a single crystal (**Figure 4**). Data was collected on the Australian Synchrotron MX1 beamline[62]. Detailed data collection and structure solution details are described elsewhere[27].

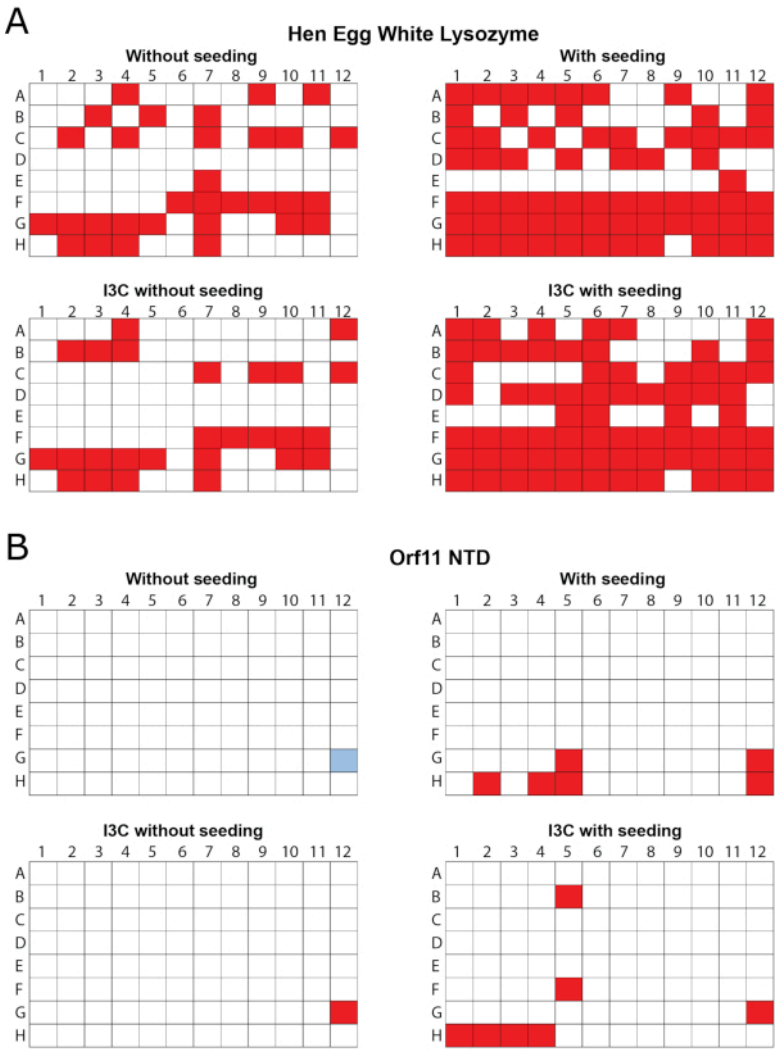


Figure 1 - rMMS was used to generate new conditions for crystal growth in the presence of I3C for two test proteins. 96 well vapor diffusion crystallization screens were carried out using commercial sparse matrix screens. (**A**) Hen egg white lysozyme was tested with the Index HT screen. Trays were seeded with HEWL crystals grown in 0.2 M ammonium tartrate dibasic pH 7.0, 20% (w/v) polyethylene glycol 3350. (**B**) Orf11 NTD from bacteriophage P68 was tested with the PEG/ION screen. Orf11 NTD trays were seeded from crystals from condition G12 from the unseeded screen, shown in blue. Conditions supporting crystal growth are shown in red. rMMS seeding in the presence and absence of I3C both gave significantly more crystal hits than unseeded trays. Figure adapted from Truong et al.[27]. Please click here to view a larger version of this figure.

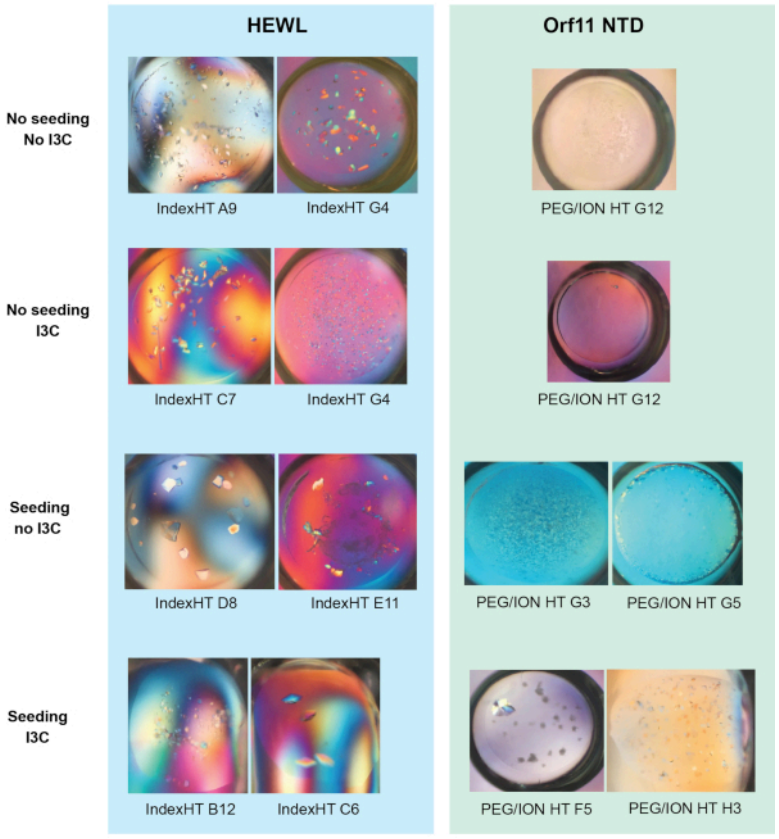


Figure 2 - Representative images of crystals grown from the vapor diffusion trials shown in Figure 1 (a) and (b).

Figure adapted from Truong et al.[27]. Please click here to view a larger version of this figure.

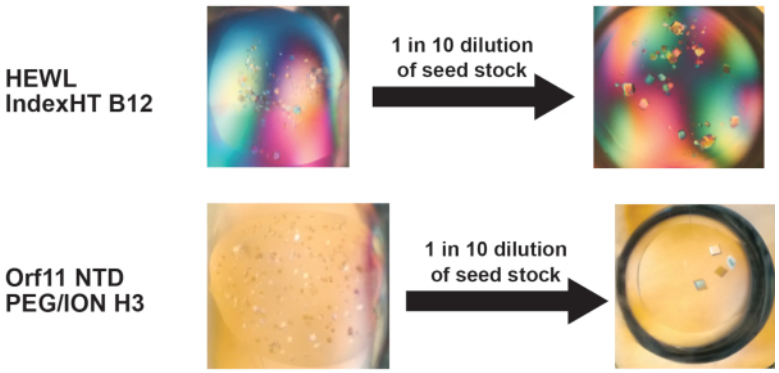


Figure 3 - Dilution of the seed stock is an effective way to reduce nucleation in a crystallization condition found using the I3C-rMMS method, to control the number of crystals that form. Reducing nucleation within a drop often results in crystals growing to larger dimensions. Figure adapted from Truong et al.[27]. Please click here to view a larger version of this figure.

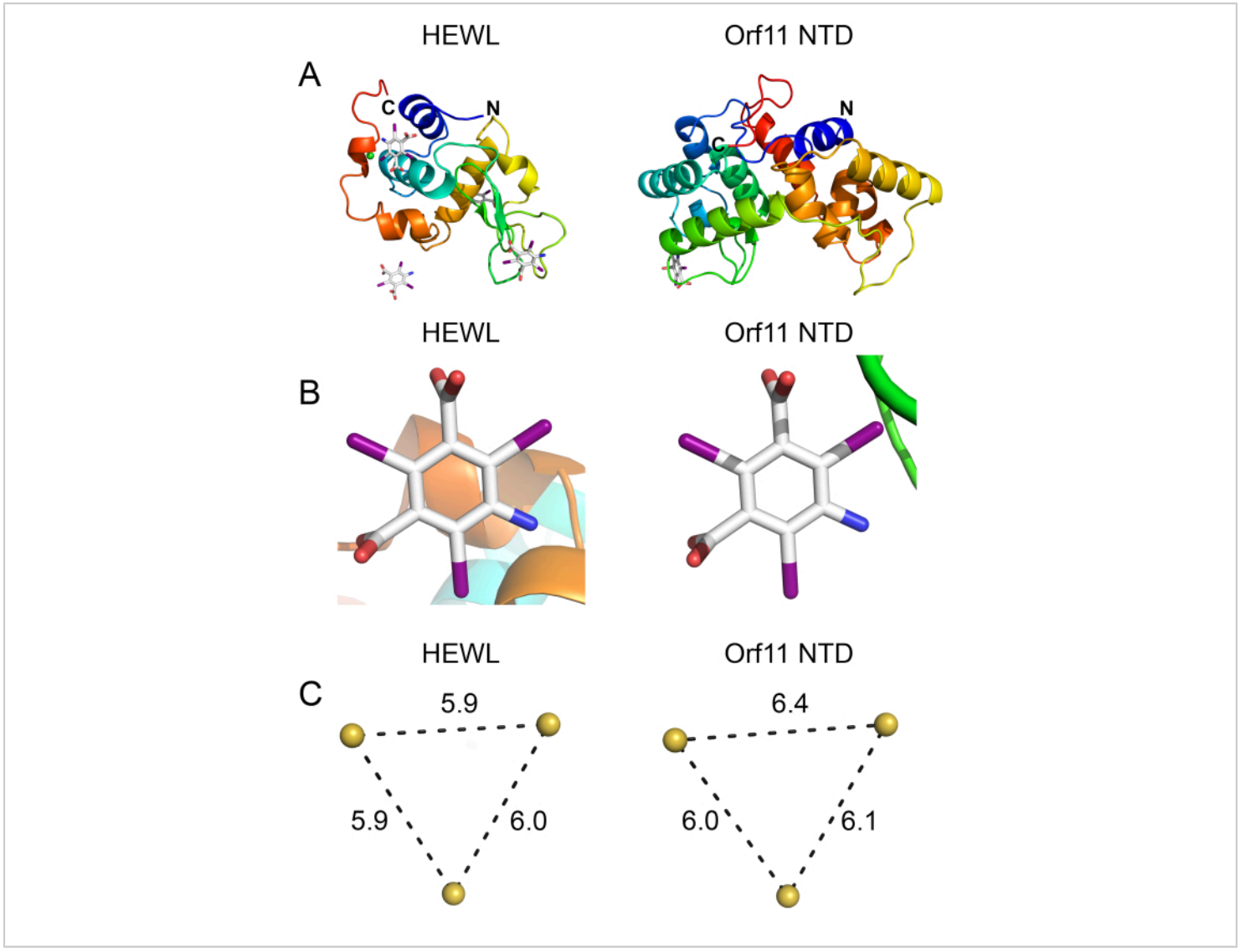


Figure 4 - Orf11 NTD (PDB ID 6O43) and HEWL (PDB ID 6PBB) were crystallized using the I3C-rMMS method and solved using Auto-Rickshaw SAD phasing. (**A**) Ribbon structures of HEWL and Orf11 NTD solved through experimental phasing. (**B**) I3C molecule bound to HEWL and Orf11 NTD. (C) Anomalous iodine atoms in I3C are arranged in an equilateral triangle of 6 Å. Thus the presence of this triangle in the phasing substructure indicates that there is an I3C molecule in that position. Please click here to view a larger version of this figure.

Discussion

Structure determination of a novel protein in the absence of a suitable homology model for molecular replacement requires experimental phasing. These methods require incorporation of heavy atoms into the protein crystal which adds a level of complexity to the structure determination pipeline and can introduce numerous obstacles that must be addressed. Heavy atoms can be incorporated directly into the protein through labelled expression using selenomethionine and selenocysteine. As this method is costly, laborious and can result in lower protein yields, labelled protein is often

expressed after crystallization conditions has been found and optimized with unlabeled protein. Alternatively, crystals can be derivatized by soaking in a solution containing heavy atoms[22,63,64]. This method often uses high quality crystals and is therefore performed after a robust crystallization method has already been developed. Successfully obtaining a derivatized crystal using this method requires further optimization of soaking procedures and screening of different phasing compounds, therefore adding more time to an already laborious process.

Co-crystallization of the protein with the heavy atom can be performed at the screening stage, thus efficiently streamlining the process and reducing crystal manipulation steps that can cause damage. However, there still exists the potential scenario of obtaining few initial crystallization hits and the problem of choosing a compatible heavy atom compound. Many currently available phasing compounds are incompatible with precipitants, buffers and additives commonly found in crystallization conditions. They may be insoluble in sulphate and phosphate buffers, chelate to citrate and acetate, react unfavorably with HEPES and Tris buffers or become sequestered by DTT and β-mercaptoethanol[21]. As the I3C phasing compound does not suffer from these incompatibilities, it is a robust phasing compound that could be amenable to many different conditions.

In this study, a streamlined method of producing derivatized crystals ready for SAD phasing through simultaneous co-crystallization of the I3C phasing compound and rMMS is presented. The combination of both techniques increases the number of crystallization hits, with many of the conditions having improved morphology and diffraction characteristics. In both Orf11 NTD and HEWL test cases, new conditions in the I3C-rMMS screen were identified that were absent when I3C was not present. Potentially, I3C may bind favorably to the protein, facilitating the formation and stabilization of crystal contacts[27]. In turn, this may induce crystallization and possibly improve diffraction characteristics. Besides being a compound compatible with sparse matrix screens, I3C is also an attractive phasing compound due to its intrinsic properties. The functional groups that alternate with iodine on the aromatic ring scaffold allow specific binding to proteins. This leads to greater occupancy and potentially reduces background signal[23]. Furthermore, the arrangement of anomalous scatterers in an equilateral triangle is obvious in the substructure and can be used to rapidly validate binding of I3C (**Figure 4B** and **4C**). Finally, it can produce an anomalous signal with tunable synchrotron radiation as well as chromium and copper rotating anode X-ray sources. Thus, it can be applied to many different workflows. As I3C is widely available and inexpensive to purchase, this approach is within reach for most structural biology laboratories.

There are several experimental considerations that must be addressed when using the I3C-rMMS method. This method cannot be applied if initial crystalline material of the protein cannot be obtained. In difficult cases, crystalline material from a homologous protein can also be used to generate seed stock. This cross-seeding approach to rMMS has shown some promising results[7]. Optimizing crystal number through dilution of the seed stock is a crucial step, which should not be overlooked, to maximize the chance of producing high quality large crystals and acquiring suitable diffraction data. If there are few I3C sites identified in the asymmetric unit, conditions conducive to crystallization should be further optimized with an increased concentration of I3C. This may increase the occupancy of I3C to maximize the anomalous signal and aid crystal derivatization.

There can be cases where this technique may not be the optimal method to derivatize protein crystals. As the size of a protein or protein-complex increases, the limited number of I3C sites on the protein surface may not provide sufficient phasing power to solve the structure. In these scenarios where protein size is suspected to be impeding phasing, selenomethionine labelling of the protein may be a more viable approach to phasing the protein. If the protein has adequate numbers of methionine residues in the protein (recommended having at least one methionine per 100 residues[65]) and high efficiency selenomethionine incorporation into a protein can be achieved (such as in bacterial expression systems[66]), multiple high occupancy selenium atoms will be present in the crystals to phase the structure.

In addition, some proteins may inherently be unsuited for derivatization with I3C. I3C binding sites on proteins are dependent on protein structure. There may exist proteins that naturally have few exposed patches compatible with I3C binding. Thus, it is not unforeseeable that there may be difficulties in co-crystallizing some target proteins with I3C.

Disclosures

The authors have nothing to disclose.

Acknowledgments

This research was undertaken on the MX1 beamline at the Australian Synchrotron, part of ANSTO. The authors would like to acknowledge members of the Shearwin and Bruning laboratories for discussions on this work. The authors would also like to acknowledge Dr. Santosh Panjikar and Dr. Linda Whyatt-Shearwin who contributed to the original work that pioneered this protocol.

The following funding is acknowledged: Australian Research Council (grant Nos. DP150103009 and DP160101450 to Keith E. Shearwin); University of Adelaide (Australian Government Research Training Program stipend scholarship to Jia Quyen Truong and Stephanie Nguyen).

References

1. Zheng, H., Hou, J., Zimmerman, M.D., Wlodawer, A., Minor, W. The future of crystallography in drug discovery. *Expert Opinion on Drug Discovery.* **9** (2), 125-137 (2014).
2. Oakley, A.J., Wilce, M.C.J. Macromolecular crystallography as a tool for investigating drug, enzyme and receptor interactions. *Clinical and Experimental Pharmacology and Physiology.* **27** (3), 145-151 (2000).
3. Jancarik, J., Kim, S.H. Sparse matrix sampling. A screening method for crystallization of proteins. *Journal of Applied Crystallography.* **24** (pt 4), 409-411 (1991).
4. Newman, J. et al. Towards rationalization of crystallization screening for small- To medium-sized academic laboratories: The PACT/JCSG+ strategy. *Acta Crystallographica Section D: Biological Crystallography.* **61** (10), 1426-1431 (2005).
5. Ireton, G.C., Stoddard, B.L. Microseed matrix screening to improve crystals of yeast cytosine deaminase. *Acta Crystallographica Section D: Biological Crystallography.* **60** (3), 601-605 (2004).
6. D'Arcy, A., Villard, F., Marsh, M. An automated microseed matrix-screening method for protein crystallization. *Acta Crystallographica Section D: Biological Crystallography.* **63** (4), 550-554 (2007).
7. D'Arcy, A., Bergfors, T., Cowan-Jacob, S.W., Marsh, M. Microseed matrix screening for optimization in protein crystallization: What have we learned?

Acta Crystallographica Section:F Structural Biology Communications. **70** (9), 1117-1126 (2014).

8. Taylor, G. The phase problem. *Acta Crystallographica - Section D Biological Crystallography.* **59** (11), 1881-1890 (2003).
9. Rossmann, M.G. The molecular replacement method. *Acta Crystallographica Section A.* **46** (2), 73-82 (1990).
10. McCoy, A.J., Grosse-Kunstleve, R.W., Adams, P.D., Winn, M.D., Storoni, L.C., Read, R.J. Phaser crystallographic software. *Journal of Applied Crystallography.* **40** (4), 658-674 (2007).
11. Millán, C., Jiménez, E., Schuster, A., Diederichs, K., Usón, I. ALIXE: a phase-combination tool for fragment-based molecular replacement. *Acta Crystallographica Section D.* **76** (3), 209-220 (2020).
12. Liebschner, D. et al. Macromolecular structure determination using X-rays, neutrons and electrons: Recent developments in Phenix. *Acta Crystallographica Section D: Structural Biology.* **75**, 861-877 (2019).
13. Abergel, C. Molecular replacement: Tricks and treats. *Acta Crystallographica Section D: Biological Crystallography.* **69** (11), 2167-2173 (2013).
14. Pröpper, K. et al. Structure solution of DNA-binding proteins and complexes with ARCIMBOLDO libraries. *Acta Crystallographica Section D: Biological Crystallography.* **70** (6), 1743-1757 (2014).
15. Rodríguez, D.D. et al. Crystallographic ab initio protein structure solution below atomic resolution. *Nature Methods.* **6** (9), 651-653 (2009).
16. Bibby, J., Keegan, R.M., Mayans, O., Winn, M.D., Rigden, D.J. AMPLE: A cluster-and-truncate approach to solve the crystal structures of small proteins using rapidly computed ab initio models. *Acta Crystallographica Section D: Biological Crystallography.* **68** (12), 1622-1631 (2012).
17. Green, D.W., Ingram, V.M., Perutz, M.F. The structure of haemoglobin - IV. Sign determination by the isomorphous replacement method. *Proceedings of the Royal Society of London. Series A. Mathematical and Physical Sciences.* **225** (1162), 287-307 (1954).
18. Blow, D.M., Rossmann, M.G. The single isomorphous replacement method. *Acta Crystallographica.* **14** (11), 1195-1202 (1961).
19. Wang, B.C. Resolution of phase ambiguity in macromolecular crystallography. *Methods in Enzymology.* **115** (C), 90-112 (1985).
20. Hendrickson, W.A. Determination of macromolecular structures from anomalous diffraction of synchrotron radiation. *Science.* **254** (5028), 51-58 (1991).
21. Pike, A.C.W., Garman, E.F., Krojer, T., Von Delft, F., Carpenter, E.P. An overview of heavy-atom derivatization of protein crystals. *Acta Crystallographica Section D: Structural Biology.* **72** (3), 303-318 (2016).
22. Dauter, Z., Dauter, M., Rajashankar, K.R. Novel approach to phasing proteins: Derivatization by short cryo-soaking with halides. *Acta Crystallographica Section D: Biological Crystallography.* **56** (2), 232-237 (2000).
23. Beck, T., Krasauskas, A., Gruene, T., Sheldrick, G.M. A magic triangle for experimental phasing of macromolecules. *Acta Crystallographica Section D: Biological Crystallography.* **64** (11), 1179-1182 (2008).
24. Beck, T., Gruene, T., Sheldrick, G.M. The magic triangle goes MAD: Experimental phasing with a bromine

derivative. *Acta Crystallographica Section D: Biological Crystallography.* **66** (4), 374-380 (2010).

25. Beck, T., Da Cunha, C.E., Sheldrick, G.M. How to get the magic triangle and the MAD triangle into your protein crystal. *Acta Crystallographica Section F: Structural Biology and Crystallization Communications.* **65** (10), 1068-1070 (2009).
26. Berman, H.M. et al. The Protein Data Bank. *Nucleic Acids Research.* (2000).
27. Truong, J.Q., Panjikar, S., Shearwin-Whyatt, L., Bruning, J.B., Shearwin, K.E. Combining random microseed matrix screening and the magic triangle for the efficient structure solution of a potential lysin from bacteriophage P68. *Acta Crystallographica Section D: Structural Biology.* **75** (7), 670-681 (2019).
28. Terwilliger, T.C. et al. Decision-making in structure solution using Bayesian estimates of map quality: The PHENIX AutoSol wizard. *Acta Crystallographica Section D: Biological Crystallography.* **65** (6), 582-601 (2009).
29. Holton, J., Alber, T. Automated protein crystal structure determination using ELVES. *Proceedings of the National Academy of Sciences of the United States of America.* **101** (6), 1537-1542 (2004).
30. Skubák, P., Pannu, N.S. Automatic protein structure solution from weak X-ray data. *Nature Communications.* **4** (2013).
31. Sheldrick, G.M. Crystal structure refinement with SHELXL. *Acta Crystallographica Section C: Structural Chemistry.* **71**, 3-8 (2015).
32. Schneider, T.R., Sheldrick, G.M. Substructure solution with SHELXD. *Acta Crystallographica Section D: Biological Crystallography.* **58** (10 I), 1772-1779 (2002).
33. Sheldrick, G.M. Macromolecular phasing with SHELXE. *Zeitschrift fur Kristallographie.* **217** (12), 644-650 (2002).
34. Benvenuti, M., Mangani, S. Crystallization of soluble proteins in vapor diffusion for x-ray crystallography. *Nature Protocols.* (2007).
35. Beck, T. *Sticky triangles : New tools for experimental phasing of biological macromolecules.* (2010).
36. Luft, J.R., DeTitta, G.T. A method to produce microseed stock for use in the crystallization of biological macromolecules. *Acta Crystallographica Section D: Biological Crystallography.* **55** (5), 988-993 (1999).
37. Teng, T.-Y. Mounting of crystals for macromolecular crystallography in a free-standing thin film. *Journal of Applied Crystallography.* **23** (5), 387-391 (1990).
38. Garman, E.F., Mitchell, E.P. Glycerol concentrations required for cryoprotection of 50 typical protein crystallization solutions. *Journal of Applied Crystallography.* **29**, 584-587 (1996).
39. Garman, E.F., Weik, M. X-ray radiation damage to biological samples: recent progress. *Journal of Synchrotron Radiation.* **26** (4), 907-911 (2019).
40. Garcia-Bonete, M.J., Katona, G. Bayesian machine learning improves single-wavelength anomalous diffraction phasing. *Acta Crystallographica Section A: Foundations and Advances.* **75**, 851-860 (2019).
41. Kabsch, W. XDS. *Acta Crystallographica Section D Biological Crystallography.* **66** (2), 125-132 (2010).
42. Evans, P. Scaling and assessment of data quality. *Acta Crystallographica Section D: Biological Crystallography.* **62** (1), 72-82 (2006).
43. Evans, P.R., Murshudov, G.N. How good are my data and what is the resolution? *Acta Crystallographica*

Section D: Biological Crystallography. **69** (7), 1204-1214 (2013).

44. Panjikar, S., Parthasarathy, V., Lamzin, V.S., Weiss, M.S., Tucker, P.A. Auto-Rickshaw: An automated crystal structure determination platform as an efficient tool for the validation of an X-ray diffraction experiment. *Acta Crystallographica Section D: Biological Crystallography.* **61** (4), 449-457 (2005).
45. Jones, T.A., Thirup, S. Using known substructures in protein model building and crystallography. *The EMBO journal.* **5** (4), 819-822 (1986).
46. Kleywegt, G.J., Jones, T.A. Template convolution to enhance or detect structural features in macromolecular electron-density maps. *Acta Crystallographica Section D: Biological Crystallography.* **53** (2), 179-185 (1997).
47. Perrakis, A., Morris, R., Lamzin, V.S. Automated protein model building combined with iterative structure refinement. *Nature Structural Biology.* **6** (5), 458-463 (1999).
48. Morris, R.J. et al. Breaking good resolutions with ARP/wARP. *Journal of Synchrotron Radiation.* **11** (1), 56-59 (2004).
49. Yao, D.Q. et al. SAD phasing by OASIS-2004: Case studies of dual-space fragment extension. *Acta Crystallographica Section D: Biological Crystallography.* **62** (8), 883-890 (2006).
50. Hao, Q. ABS: A program to determine absolute configuration and evaluate anomalous scatterer substructure. *Journal of Applied Crystallography.* **37** (3), 498-499 (2004).
51. Collaborative Computational Project Number 4 The CCP4 suite: Programs for protein crystallography. *Acta Crystallographica Section D: Biological Crystallography.* **50** (5), 760-763 (1994).
52. Sheldrick, G.M., Hauptman, H.A., Weeks, C.M., Miller, R., Usón, I. Ab initio phasing. *International Tables for Crystallography.* 333-345 (2006).
53. Smith, G.D. Matching selenium-atom peak positions with a different hand or origin. *Journal of Applied Crystallography.* **35** (3), 368-370 (2002).
54. Pannu, N.S., McCoy, A.J., Read, R.J. Application of the complex multivariate normal distribution to crystallographic methods with insights into multiple isomorphous replacement phasing. *Acta Crystallographica - Section D Biological Crystallography.* **59** (10), 1801-1808 (2003).
55. Pannu, N.S., Read, R.J. The application of multivariate statistical techniques improves single-wavelength anomalous diffraction phasing. *Acta Crystallographica Section D: Biological Crystallography.* **60** (1), 22-27 (2004).
56. De La Fortelle, E., Bricogne, G. Maximum-likelihood heavy-atom parameter refinement for multiple isomorphous replacement and multiwavelength anomalous diffraction methods. *Methods in Enzymology.* **276**, 472-494 (1997).
57. Cowtan, K. Joint CCP4 and ESF-EACBM Newsletter on Protein. *Crystallography.* **31**, 34-38, at <https://ci.nii.ac.jp/naid/10010645386/en/> (1994).
58. Terwilliger, T.C. Maximum-likelihood density modification. *Acta Crystallographica Section D: Biological Crystallography.* **56** (8), 965-972 (2000).

59. Read, R.J., McCoy, A.J. Maximum-likelihood determination of anomalous substructures. *Acta Crystallographica Section D: Structural Biology.* (2018).

60. Emsley, P., Lohkamp, B., Scott, W.G., Cowtan, K. Features and development of Coot. *Acta Crystallographica Section D: Biological Crystallography.* **66** (4), 486-501 (2010).

61. Till, M. et al. Improving the success rate of protein crystallization by random microseed matrix screening. *Journal of visualized experiments : JoVE.* (78), e50548 (2013).

62. McPhillips, T.M. et al. Blu-Ice and the distributed control system: Software for data acquisition and instrument control at macromolecular crystallography beamlines. *Journal of Synchrotron Radiation.* **9** (6), 401-406 (2002).

63. Nagem, R.A.P., Polikarpov, I., Dauter, Z. Phasing on Rapidly Soaked Ions. *Methods in Enzymology.* **374**, 120-137 (2003).

64. Taylor, G.L. Introduction to phasing. *Acta Crystallographica Section D: Biological Crystallography.* **66** (4), 325-338 (2010).

65. Hendrickson, W.A., Ogata, C.M. Phase determination from multiwavelength anomalous diffraction measurements. *Methods in Enzymology.* **276**, 494-523 (1997).

66. Doublié, S. Production of Selenomethionyl Proteins in Prokaryotic and Eukaryotic Expression Systems. *Macromolecular Crystallography Protocols. Methods in Molecular Biology.* 91-108 (2007).

Appendix C (Publication):

Simplified heavy atom derivatization of protein structures via co-crystallization with the MAD tetragon B4C

Statement of Authorship

Title of Paper	Simplified heavy atom derivatization of protein structures via co-crystallization with the MAD tetragon B4C
Publication Status	☑ Published ☐ Accepted for Publication ☐ Submitted for Publication ☐ Unpublished and Unsubmitted work written in manuscript style
Publication Details	This publication describes a new technique to produce derivatised protein crystals using random microseed matrix screening whilst incorporating the phasing molecule, I3C.

Principal Author

Name of Principal Author (Candidate)	Stephanie Nguyen		
Contribution to the Paper	Contributed to data analysis, prepared, reviewed and edited the manuscript.		
Overall percentage (%)	40		
Certification:	This paper reports on original research I conducted during the period of my Higher Degree by Research candidature and is not subject to any obligations or contractual agreements with a third party that would constrain its inclusion in this thesis. I am the primary author of this paper.		
Signature		Date	19/07/21

Co-Author Contributior.

By signing the Statement of Authorship, each author certifies that:

i. the candidate's stated contribution to the publication is accurate (as detailed above);

ii. permission is granted for the candidate in include the publication in the thesis; and

iii. the sum of all co-author contributions is equal to 100% less the candidate's stated contribution.

Name of Co-Author	Dr. Jia Quyen Truong		
Contribution to the Paper	Project conceptualisation, planned and performed experiments, analysed data, prepared, reviewed and edited the manuscript.		
Overall Percentage (%)	40		
Signature		Date	19/07/21

Name of Co-Author	Assoc. Prof. Keith Shearwin		
Contribution to the Paper	Reviewed and edited the manuscript.		
Overall Percentage (%)	10		
Signature		Date	19/07/2021

Name of Co-Author	Dr. John B. Bruning		
Contribution to the Paper	Reviewed and edited the manuscript.		
Overall Percentage (%)	10		
Signature		Date	21/7/21

STRUCTURAL BIOLOGY
COMMUNICATIONS

ISSN: 2053-230X
journals.iucr.org/f

Simplified heavy-atom derivatization of protein structures via co-crystallization with the MAD tetragon tetrabromoterephthalic acid

Jia Q. Truong, Stephanie Nguyen, John B. Bruning and Keith E. Shearwin

Acta Cryst. (2021). F**77**, 156–162

STRUCTURAL BIOLOGY COMMUNICATIONS

ISSN 2053-230X

Simplified heavy-atom derivatization of protein structures via co-crystallization with the MAD tetragon tetrabromoterephthalic acid

Jia Q. Truong,[a]‡ Stephanie Nguyen,[b]‡ John B. Bruning[b] and Keith E. Shearwin[a]*

[a]School of Biological Sciences, The University of Adelaide, Adelaide, South Australia 5005, Australia, and [b]Institute of Photonics and Advanced Sensing (IPAS), School of Biological Sciences, The University of Adelaide, Adelaide, South Australia 5005, Australia. *Correspondence e-mail: keith.shearwin@adelaide.edu.au

Received 10 March 2021
Accepted 15 April 2021

Edited by R. L. Stanfield, The Scripps Research Institute, USA

‡ These authors contributed equally.

Keywords: crystallography; experimental phasing; B4C; tetrabromoterephthalic acid; lysozyme; co-crystallization.

PDB reference: hen egg-white lysozyme, complex with tetrabromoterephthalic acid, 7kh5

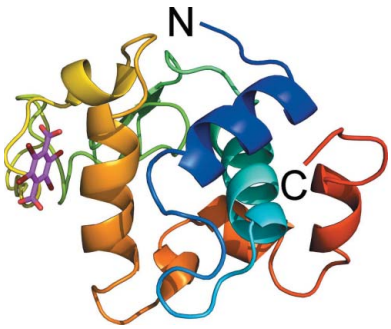
Supporting information: this article has supporting information at journals.iucr.org/f

The phase problem is a persistent bottleneck that impedes the structure-determination pipeline and must be solved to obtain atomic resolution crystal structures of macromolecules. Although molecular replacement has become the predominant method of solving the phase problem, many scenarios still exist in which experimental phasing is needed. Here, a proof-of-concept study is presented that shows the efficacy of using tetrabromoterephthalic acid (B4C) as an experimental phasing compound. Incorporating B4C into the crystal lattice using co-crystallization, the crystal structure of hen egg-white lysozyme was solved using MAD phasing. The strong anomalous signal generated by its four Br atoms coupled with its compatibility with commonly used crystallization reagents render B4C an effective experimental phasing compound that can be used to overcome the phase problem.

1. Introduction

X-ray crystallography is the gold-standard technique used to determine the three-dimensional structures of proteins and other macromolecules. At the time of writing, 88% of the total number of deposited structures in the Protein Data Bank (PDB) have been solved using X-ray crystallography (Burley *et al.*, 2019). Although the use of this technique is ubiquitous in the structural biology community, there are several challenges associated with its implementation that must be addressed in order to elucidate a protein structure from a diffraction pattern. One is that the conversion of an X-ray diffraction pattern requires the solution of the phase problem. Molecular replacement has quickly become the preferred method to overcome the phase problem due to its accessibility, speed and affordability. This technique approximates the phases of the target structure from structurally similar proteins that have previously been solved (Evans & McCoy, 2008; Rossmann, 1990). Despite the many benefits of using molecular replacement and the growing number of deposited structures in the PDB that are available for use as models, there are still limitations to this technique. Many proteins lack a suitable search model and so experimental phasing must be used to solve the phase problem.

To acquire initial phases experimentally, heavy atoms are incorporated into the crystal lattice at defined locations and phases are traditionally obtained using one of two methods: isomorphous replacement or anomalous dispersion (Cowtan, 2003). If isomorphous crystals (crystals with very similar unit-cell dimensions with protein molecules arranged similarly within the unit cell) can be obtained with and without the heavy atom, phases can be obtained using the isomorphous



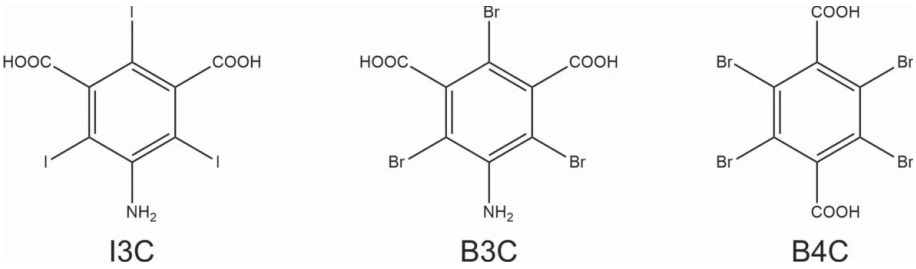


Figure 1
Chemical structure of the phasing compounds I3C (5-amino-2,4,6-triiodobenzene-1,3-dicarboxylic acid), B3C (5-amino-2,4,6-tribromobenzene-1,3-dicarboxylic acid) and B4C (tetrabromoterephthalic acid).

replacement method. In contrast, the anomalous dispersion method only requires a protein crystal to have a heavy atom incorporated. In this case, data collection includes measuring diffraction at an X-ray wavelength close to the absorption edge of the heavy atom. At these wavelengths, Friedel's law breaks down, resulting in anomalous scattering. The theories for phase determination using isomorphous replacement and anomalous scattering have been described previously (Taylor, 2010).

Methods of incorporating heavy atoms into crystals include labelling the protein (incorporating selenomethionine or selenocysteine amino acids; Strub *et al.*, 2003), chemical modification (such as 5-bromouracil labelling of RNA; Baugh *et al.*, 2000; Kieft *et al.*, 2002), and soaking and co-crystallization of the crystal with a heavy-atom compound. The last method, although a general method to derivatize crystals, often suffers from nonspecific binding, resulting in multiple binding sites with low occupancy in the unit cell and ultimately, low anomalous signal.

To address this problem, Tobias Beck, in his PhD dissertation *Sticky Triangles: New Tools for Experimental Phasing of Biological Molecules*, proposed a set of five new phasing compounds that could be used to derivatize protein crystals and incorporate heavy atoms into the lattice for phasing (Beck, 2010). These compounds have anomalous scattering halogen atoms, either bromine or iodine, attached to a benzene scaffold. In addition, the compounds were designed to have functional groups such as carboxyl, amine, methoxyl and hydroxyl groups surrounded by the heavy atoms on the benzene scaffold. These functional groups and the aromatic ring scaffold allow more specific interactions with the protein to improve occupancy and yield higher anomalous signals.

Two of these compounds, the magic triangle and the MAD triangle (Fig. 1), were later published in a series of journal articles (Beck *et al.*, 2008, 2009, 2010). The MAD triangle, 5-amino-2,4,6-tribromobenzene-1,3-dicarboxylic acid (B3C), has only been used to phase two structures in the PDB, whereas the magic triangle, 5-amino-2,4,6-triiodobenzene-1,3-dicarboxylic acid (I3C), has gained significant traction and has been incorporated into 21 structures in the PDB. Two suppliers of protein crystallography reagents, Hampton Research and Molecular Dimensions, also sell kits for I3C phasing, demonstrating the accessibility and feasibility of using I3C as a compound for experimental phasing.

There could be advantages to exploring phasing compounds outside of I3C and B3C. For larger proteins and protein complexes, the limited number of I3C or B3C sites on the protein surface may provide insufficient phasing power to successfully solve the phase problem. This limit could be overcome by employing phasing molecules with increased phasing power per molecule, such as tetrabromoterephthalic acid (B4C; Fig. 1). This compound, dubbed 'the MAD tetragon', has four anomalous scattering Br atoms and should provide increased anomalous signal over I3C and B3C. B4C has successfully been incorporated into crystals via soaking and has been used to phase the structures of thaumatin and thermolysin (Beck, 2010). This work has yet to be published in a peer-reviewed journal and we would like to explicitly credit Beck and the Sheldrick group as the first users of B4C as a phasing tool.

Recently, we have demonstrated that I3C can be combined with random microseed matrix screening (rMMS) to efficiently generate derivatized protein crystals (Truong *et al.*, 2019, 2021). To expand this screening technique to other phasing compounds, we have searched for other specific phasing ligands that could be effectively co-crystallized with different proteins. In this study, we have successfully co-crystallized B4C with hen egg-white lysozyme (HEWL) and confirm that B4C provides sufficient anomalous signal for structure solution using MAD phasing. This work expands the arsenal of existing phasing compounds and provides a proof of concept for the experimental phasing of other proteins.

2. Materials and methods

2.1. Crystallization

Tetrabromoterephthalic acid (B4C, catalogue No. 524441) and hen egg-white lysozyme (HEWL; catalogue No. L6876) were commercially acquired from Sigma–Aldrich. Lyophilized HEWL powder was dissolved in TBS (50 m*M* Tris–HCl pH 7.6, 150 m*M* NaCl) to a final concentration of 30 mg ml^{-1} as determined by UV absorbance at 280 nm. B4C was dissolved in 2 *M* LiOH solution to a final concentration of 0.5 *M*. To prepare HEWL samples for co-crystallization experiments, 0.5 *M* B4C was added directly to the HEWL solution to a final concentration of 40 m*M*. HEWL crystals were grown via sitting-drop vapour diffusion in 96-well Intelli-Plates (Art

Table 1
Statistics of data processing for HEWL co-crystallized with B4C.

Method	Sitting-drop vapour diffusion
Plate type	Art Robbins 96-well Intelli-Plate
Temperature (K)	289.15
Protein concentration (mg ml^{-1})	30
Buffer composition of protein solution	50 m*M* Tris–HCl pH 7.6, 150 m*M* NaCl
Composition of reservoir solution	20%(*w*/*v*) polyethylene glycol 3350, 0.2 *M* potassium sulfate
Volume and ratio of drop	1 µl:1 µl
Volume of reservoir (µl)	75

Robbins) by mixing 1 µl HEWL solution with 1 µl reservoir solution [20%(*w*/*v*) polyethylene glycol 3350, 0.2 *M* potassium sulfate] and equilibrating against 75 µl reservoir solution (Table 1). The crystals were harvested onto a cryoloop, passed through Paratone-N (Hampton Research) for cryoprotection and flash-cooled in liquid nitrogen for data collection (Teng, 1990).

2.2. Data collection and processing

A 1.3 Å resolution MAD data set was collected at a wavelength of 0.92 Å using an EIGER X 16M detector on the macromolecular beamline MX2 at the Australian Synchrotron, which is part of ANSTO (Aragão *et al.*, 2018). This wavelength produced the maximal anomalous signal close to the Br *K* edge, as determined by a fluorescence scan (Fig. 2). A total of 3600 diffraction images were collected with a 0.1° oscillation width and a crystal-to-detector distance of 170 mm. The diffraction data were processed using *XDS* (Kabsch, 2010) and were combined and scaled with *AIMLESS* (Evans & Murshudov, 2013) (Table 2).

2.3. Structure solution and refinement

The structure was solved using the two-wavelength multiple-wavelength anomalous diffraction (MAD) protocol of *Auto-Rickshaw*, the EMBL-Hamburg automated crystal structure-determination platform (Panjikar *et al.*, 2005). The diffraction data processed using *AIMLESS* (Evans & Murshudov, 2013) were used as input. Marker-atom structure-factor amplitude (F_A) values were calculated using *SHELXC* (Sheldrick *et al.*, 2001). Based on an initial analysis of the data,

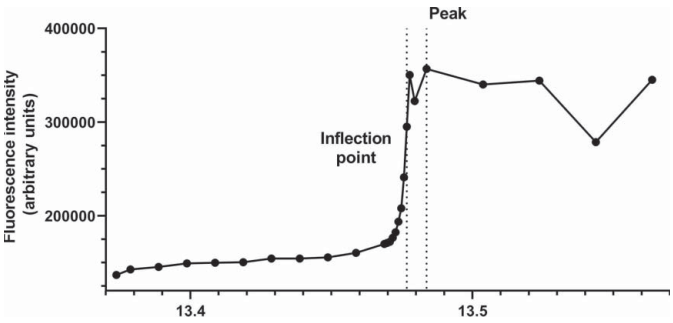


Figure 2
Fluorescence scan of HEWL co-crystallized with B4C. The peak wavelength (13.4837 keV) and inflection points (13.4767 keV) were assigned with *CHOOCH* (Evans & Pettifer, 2001)

Table 2
Statistics of data processing for HEWL co-crystallized with B4C.

Values in parentheses are for the highest resolution shell.

	Peak	Inflection
Diffraction source	MX2, Australian Synchrotron	MX2, Australian Synchrotron
Wavelength (Å)	0.9106	0.9200
Temperature (K)	100	100
Detector	EIGER X 16M	EIGER X 16M
Crystal-to-detector distance (mm)	170	170
Rotation range per image (°)	0.1	0.1
Total rotation range (°)	360	360
Space group	$P4_32_12$	$P4_32_12$
a, *b*, *c* (Å)	77.99, 77.99, 37.76	77.98, 77.98, 37.75
α, β, γ (°)	90, 90, 90	90, 90, 90
Mosaicity (°)	0.09	0.09
Resolution range (Å)	38.99–1.30	38.99–1.28
Total No. of reflections	734500 (32208)	753940 (31166)
No. of unique reflections	29530 (1477)	30377 (1428)
Completeness (%)	99.8 (97.1)	99.8 (96.2)
Multiplicity	24.9	24.8 (21.8)
$\langle I/\sigma(I)\rangle$	31.0 (2.4)	30.9 (2.5)
R_{meas}	0.052 (1.52)	0.050 (1.51)
Overall *B* factor from Wilson plot (Å^2)	16.54	16.97

the maximum resolution for substructure determination and initial phase calculation was set to 1.75 Å. *SHELXD* found 27 potential heavy-atom sites (Schneider & Sheldrick, 2002). The correct hand for the substructure was determined using *ABS* (Hao, 2004) and *SHELXE* (Sheldrick, 2002). Initial phases were calculated after density modification using *SHELXE* (Sheldrick, 2002). Using *ARP*/*wARP*, 94% (121 of 129 residues) of the model was built (Perrakis *et al.*, 1999; Morris *et al.*, 2004) and correctly docked into electron density. At this stage, the structure had an *R* factor of 27.4%. The structure was then iteratively rebuilt and refined using *Coot* (Emsley *et al.*, 2010) and *phenix.refine* (Afonine *et al.*, 2012) to an *R* and R_{free} of 19.38% and 22.35%, respectively. Structure-solution statistics are summarized in Table 3.

2.4. Accession numbers

The coordinates and structure factors for hen egg-white lysozyme co-crystallized with B4C have been deposited in the Protein Data Bank under accession number 7kh5.

3. Results

3.1. B4C is an effective compound for the experimental phasing of proteins

The anomalous signal from a single B4C molecule was sufficient to solve the crystal structure of HEWL to a resolution of 1.3 Å using MAD phasing. HEWL forms a compact and globular structure, consisting of four α-helices, three β-strands and three 3_{10}-helices (Fig. 3). Four disulfide bonds are formed between Cys6 and Cys127, between Cys30 and Cys115, between Cys64 and Cys80 and between Cys76 and Cys94. The overall fold and structure of HEWL solved by

Table 3
Statistics for structure solution of HEWL co-crystallized with B4C.

Values in parentheses are for the highest resolution shell.

Resolution range (Å)	38.99–1.30
Completeness (%)	99.8 (97.1)
No. of reflections, working set	29473 (2840)
No. of reflections, test set	978 (102)
Final R_{cryst}	0.19 (0.30)
Final R_{free}	0.22 (0.35)
No. of non-H atoms	
Total	1169
Protein	991
Ligand	16
Water	162
R.m.s. deviations	
Bonds (Å)	0.003
Angles (°)	0.60
Average *B* factors (Å^2)	
Overall	25.31
Protein	23.59
Ligand	31.14
Water	35.22
Ramachandran plot	
Most favoured (%)	98.43
Allowed (%)	1.57

experimental phasing with B4C is consistent with previously solved structures (data not shown).

3.2. B4C forms a characteristic arrangement of anomalous scatterers, and functional groups on the benzene scaffold of B4C facilitate protein binding

A single molecule of B4C was bound to one monomer of HEWL (with an occupancy of 27%), positioned between the α3 helix and the β3/η1 loop. When bound, B4C is positioned within an overall positively charged surface-exposed cleft with two *ortho*-arranged Br atoms protruding into a deeper pocket (Fig. 4*a*). Due to the arrangement of Br atoms on the benzene scaffold of B4C, the anomalous signal generated forms a characteristic 5.7 × 3.37 Å tetragon (Fig. 4*b*). Key hydrogen-bonding and hydrophobic interactions facilitate the binding of the compound to HEWL. When bound, the carboxylic acid functional group on the benzene scaffold of B4C forms a 2.9 Å hydrogen bond between the hydroxyl group and the side chain of Asn93 (Fig. 4*c*). Additional hydrophobic interactions formed between B4C and Cys76, Ile78, Ala90, Asn93 and Cys94 further stabilize binding. Analysis of the crystal contacts reveals two molecules of B4C, one from each symmetry mate, interlocking and likely aiding crystallization (Fig. 4*d*).

4. Discussion

Although the number of structures deposited in the PDB is steadily growing, a significant proportion of proteins remain that lack a suitable template for molecular replacement and therefore require experimental phasing. Commonly, phasing compounds containing heavy atoms are soaked into pre-existing crystals or co-crystallized with the protein to enable experimental phasing. However, many of these compounds have inherent limitations that hinder the structure-determination pipeline. Some phasing compounds, such as heavy-metal salts, show poor solubility, which can reduce their availability in both soaking and co-crystallization experiments. Halide-containing compounds may bind with low specificity, leading to low occupancy and poor anomalous scattering (Pike *et al.*, 2016). Furthermore, many of these phasing compounds are incompatible with common crystallization reagents, including phosphate, sulfate, citrate, acetate, tris(2-amino-2-hydroxymethyl-propane-1,3-diol) and HEPES [4-(2-hydroxyethyl)-1-piperazineethanesulfonic acid] buffers. Since there is no singular phasing compound that is universally compatible with all proteins and crystallization conditions, it is necessary to generate a library of available compounds that can be utilized in different situations.

In this study, we have demonstrated that B4C is an effective phasing compound by successfully solving the structure of the model protein HEWL. Due to the characteristic arrangement of Br atoms in B4C, the position of the B4C ligand was readily identified in the anomalous density map. In B4C, the Br atoms form a characteristic rectangle (5.7 × 3.37 Å) which can be used to easily validate that the substructure identified is correct (Fig. 4*b*).

The chemical composition of B4C offers several advantages that improve its qualities as a phasing compound. The unique arrangement of functional groups on the benzene scaffold affords superior binding specificity to the protein. In turn, this improves the occupancy, generates greater anomalous signal and thus reduces noise. As shown, a single molecule of B4C bound to HEWL provides sufficient anomalous signal for experimental phasing (Fig. 3).

B4C has a *para* arrangement of carboxyl groups that are surrounded by four Br atoms on the aromatic scaffold. This arrangement of functional groups differs from the arrangement of I atoms in I3C and Br atoms in B3C. Thus, B4C can recognize and bind different sites on the protein compared

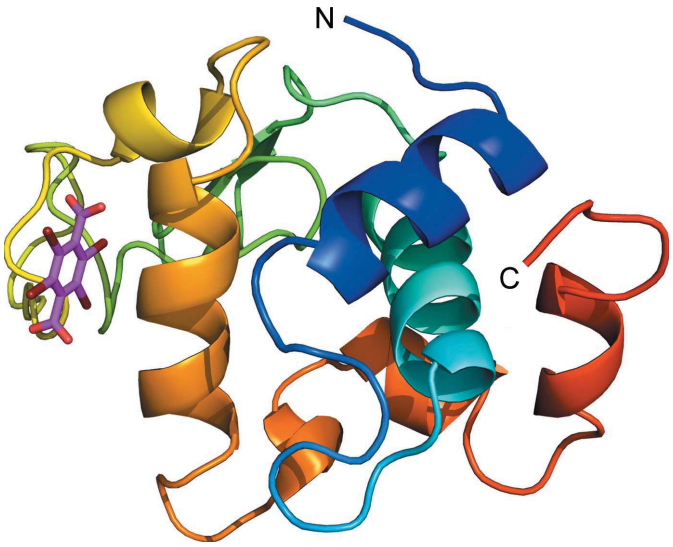


Figure 3
Hen egg-white lysozyme co-crystallized with B4C (PDB entry 7kh5). The crystal structure of HEWL is displayed in cartoon representation and B4C is represented in magenta as sticks.

with I3C and B3C. The two available crystal structures of I3C-bound HEWL (PDB entries 6pbb and 3e3d) show six positions at which I3C binds on the HEWL protein surface (Truong *et al.*, 2019; Beck *et al.*, 2008). The position where B4C binds on the protein surface differs from the previously occupied positions of I3C and does not clash spatially (Fig. 5). This result indicates that proteins that fail to be derivatized with I3C (and likely with B3C) could potentially be derivatized with B4C. B4C also has four anomalous scattering Br atoms, compared with three Br atoms in B3C, and thus in theory could provide greater anomalous signal. The presence of Br atoms also means that it can be phased by MAD phasing, instead of being restricted to single-wavelength anomalous dispersion (SAD) phasing as is the case for I3C. MAD can provide additional phasing information in comparison to the SAD method (Rice *et al.*, 2000).

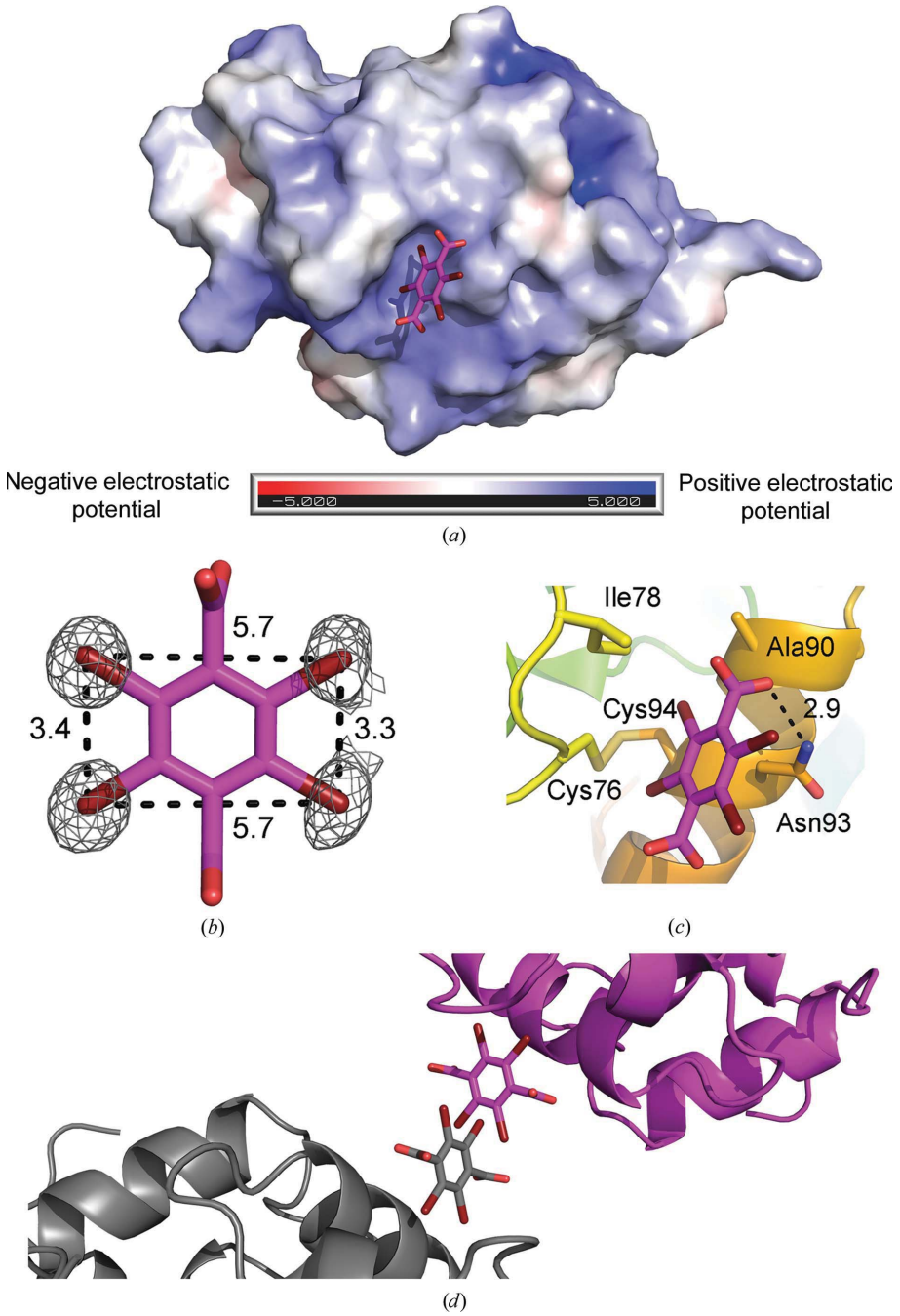


Figure 4
Characteristics of B4C binding to hen egg-white lysozyme. (*a*) Electrostatic surface representation of HEWL with B4C bound to a positively charged surface-exposed cleft. (*b*) The binding site of B4C in the lysozyme crystal overlaid with the substructure density of B4C (anomalous difference map contour of 5σ). This map was generated using *phenix.maps*. The four Br atoms arranged on the benzene scaffold form a characteristic tetragon with dimensions of 5.7 × 3.3–3.4 Å. (*c*) Binding of B4C to HEWL occurs through hydrogen-bond interactions (black dashed line) with the Asn93 side chain. All distances are reported in Å. (*d*) Symmetry mates of HEWL (depicted in grey and magenta) show B4C molecules interlocking at the interface of the crystal contact.

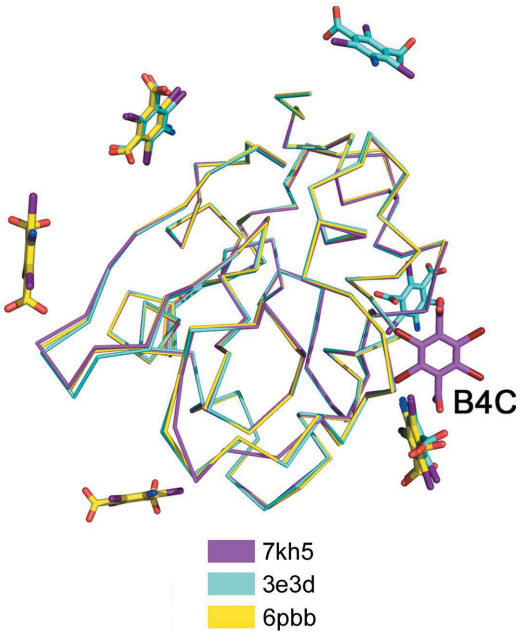


Figure 5
B4C can occupy different sites on a protein surface compared with I3C. Structural superimposition of HEWL bound to I3C (PDB entries 6pbb, yellow, and 3e3d, cyan) and HEWL bound to B4C (PDB entry 7kh5, magenta) reveals that B4C binds to a different surface patch on HEWL to I3C.

As B4C is a common starting reagent for the synthesis of polyesters, it is commercially available. It can be acquired easily in large quantities and is relatively inexpensive. It is also highly soluble in lithium hydroxide and other basic solutions, undergoing an acid–base reaction upon dissolving to produce a B4C solution that has an approximately neutral pH. This is beneficial for both co-crystallization and soaking experiments, since the process of crystallization and the crystals themselves are sensitive to changes in pH.

Additionally, B4C is compatible with co-crystallization experiments and thus derivatization using this compound is not entirely reliant on soaking. Co-crystallization reduces the number of potentially damaging crystal-handling steps. B4C also could be added to commercial sparse-matrix screens to identify suitable co-crystallization conditions whilst simultaneously growing derivatized protein crystals. A further application may involve combining B4C co-crystallization with rMMS to expand the number of potential conditions for derivatization of crystals. The efficiency of this technique has been demonstrated previously using the phasing compound I3C (Truong *et al.*, 2019).

In this study, we have co-crystallized HEWL with B4C and solved the structure using MAD phasing. The several advantageous properties of B4C, including its high phasing power and its compatibility with common crystallization reagents, highlight the effectiveness of B4C as an experimental phasing molecule. Due to its accessibility and affordability, B4C is an excellent addition to the arsenal of phasing compounds suitable for scenarios in which molecular replacement is not possible.

Acknowledgements

This research was undertaken in part using the MX2 beamline at the Australian Synchrotron, which is part of ANSTO, and made use of the Australian Cancer Research Foundation (ACRF) detector.

Funding information

The following funding is acknowledged: Australian Research Council (grant No. DP150103009 to Keith E. Shearwin; grant No. DP160101450 to Keith E. Shearwin); University of Adelaide (scholarship to Jia Q. Truong, Stephanie Nguyen).

References

Afonine, P. V., Grosse-Kunstleve, R. W., Echols, N., Headd, J. J., Moriarty, N. W., Mustyakimov, M., Terwilliger, T. C., Urzhumtsev, A., Zwart, P. H. & Adams, P. D. (2012). *Acta Cryst.* D**68**, 352–367.
Aragão, D., Aishima, J., Cherukuvada, H., Clarken, R., Clift, M., Cowieson, N. P., Ericsson, D. J., Gee, C. L., Macedo, S., Mudie, N., Panjikar, S., Price, J. R., Riboldi-Tunnicliffe, A., Rostan, R., Williamson, R. & Caradoc-Davies, T. T. (2018). *J. Synchrotron Rad.* **25**, 885–891.
Baugh, C., Grate, D. & Wilson, C. (2000). *J. Mol. Biol.* **301**, 117–128.
Beck, T. (2010). PhD thesis. Universität Göttingen, Göttingen, Germany.
Beck, T., da Cunha, C. E. & Sheldrick, G. M. (2009). *Acta Cryst.* F**65**, 1068–1070.
Beck, T., Gruene, T. & Sheldrick, G. M. (2010). *Acta Cryst.* D**66**, 374–380.
Beck, T., Krasauskas, A., Gruene, T. & Sheldrick, G. M. (2008). *Acta Cryst.* D**64**, 1179–1182.
Burley, S. K., Berman, H. M., Bhikadiya, C., Bi, C., Chen, L., Di Costanzo, L., Christie, C., Dalenberg, K., Duarte, J. M., Dutta, S., Feng, Z., Ghosh, S., Goodsell, D. S., Green, R. K., Guranović, V., Guzenko, D., Hudson, B. P., Kalro, T., Liang, Y., Lowe, R., Namkoong, H., Peisach, E., Periskova, I., Prlić, A., Randle, C., Rose, A., Rose, P., Sala, R., Sekharan, M., Shao, C., Tan, L., Tao, Y. P., Valasatava, Y., Voigt, M., Westbrook, J., Woo, J., Yang, H., Young, J., Zhuravleva, M. & Zardecki, C. (2019). *Nucleic Acids Res.* **47**, D464–D474.
Cowtan, K. (2003). *eLS*, https://doi.org/10.1038/npg.els.0002722.
Emsley, P., Lohkamp, B., Scott, W. G. & Cowtan, K. (2010). *Acta Cryst.* D**66**, 486–501.
Evans, G. & Pettifer, R. F. (2001). *J. Appl. Cryst.* **34**, 82–86.
Evans, P. & McCoy, A. (2008). *Acta Cryst.* D**64**, 1–10.
Evans, P. R. & Murshudov, G. N. (2013). *Acta Cryst.* D**69**, 1204–1214.
Hao, Q. (2004). *J. Appl. Cryst.* **37**, 498–499.
Kabsch, W. (2010). *Acta Cryst.* D**66**, 125–132.
Kieft, J. S., Zhou, K., Grech, A., Jubin, R. & Doudna, J. A. (2002). *Nat. Struct. Biol.* **9**, 370–374.
Morris, R. J., Zwart, P. H., Cohen, S., Fernandez, F. J., Kakaris, M., Kirillova, O., Vonrhein, C., Perrakis, A. & Lamzin, V. S. (2004). *J. Synchrotron Rad.* **11**, 56–59.
Panjikar, S., Parthasarathy, V., Lamzin, V. S., Weiss, M. S. & Tucker, P. A. (2005). *Acta Cryst.* D**61**, 449–457.
Perrakis, A., Morris, R. & Lamzin, V. S. (1999). *Nat. Struct. Biol.* **6**, 458–463.
Pike, A. C. W., Garman, E. F., Krojer, T., von Delft, F. & Carpenter, E. P. (2016). *Acta Cryst.* D**72**, 303–318.
Rice, L. M., Earnest, T. N. & Brunger, A. T. (2000). *Acta Cryst.* D**56**, 1413–1420.
Rossmann, M. G. (1990). *Acta Cryst.* A**46**, 73–82.
Schneider, T. R. & Sheldrick, G. M. (2002). *Acta Cryst.* D**58**, 1772–1779.
Sheldrick, G. M. (2002). *Z. Kristallogr.* **217**, 644–650.

Sheldrick, G. M., Hauptman, H. A., Weeks, C. M., Miller, R. & Usón, I. (2001). *International Tables for Crystallography*, Vol. F, edited by M. G. Rossmann & E. Arnold, pp. 333–345. Dordrecht: Kluwer Academic Publishers.
Strub, M. P., Hoh, F., Sanchez, J. F., Strub, J. M., Böck, A., Aumelas, A. & Dumas, C. (2003). *Structure*, **11**, 1359–1367.
Taylor, G. L. (2010). *Acta Cryst.* D**66**, 325–338.
Teng, T.-Y. (1990). *J. Appl. Cryst.* **23**, 387–391.
Truong, J. Q., Nguyen, S., Bruning, J. B. & Shearwin, K. E. (2021). *J. Vis. Exp.*, e61894.
Truong, J. Q., Panjikar, S., Shearwin-Whyatt, L., Bruning, J. B. & Shearwin, K. E. (2019). *Acta Cryst.* D**75**, 670–681.