Enzymatic Inactivation of Trichothecene Mycotoxins Associated with *Fusarium* Head Blight

By

Karl M. Wetterhorn

A dissertation submitted in partial fulfillment of

the requirements for the degree of

Doctor of Philosophy

(Biochemistry)

at the

UNIVERSITY OF WISCONSIN-MADISON

2018

Date of final oral examination: 12/3/2018

The dissertation is approved by the following members of the Final Oral Committee:

Ivan Rayment. Professor, Biochemistry
Aaron Hoskins. Associate Professor, Biochemistry
Brian G. Fox. Professor, Biochemistry
Nancy Keller. Professor, Bacteriology and Medical Microbiology and Immunology
Srivatsan Raman. Assistant Professor, Biochemistry

Abstract

Fusarium head blight (FHB) and *Fusarium* ear rot are devastating plant diseases that affect small grain cereals and maize on a global scale. Trichothecene mycotoxins produced by *Fusarium* fungi act as virulence factors and accumulate in the grain of infected crops. These mycotoxins are potent inhibitors of eukaryotic protein synthesis, posing a significant threat to both animal and human consumers. To date, one of the most promising approaches to combating FHB has been the creation of transgenic crops that constitutively overexpress trichothecene UDP-glucosyltransferases (UGTs), which inactivate trichothecenes and inhibit the spread of infection. These transgenic crops have demonstrated resistance to specific *Fusarium* species such as *Fusarium graminearum*, but the UGTs responsible for this resistance do not demonstrate broad specificity to the diverse group of over 200 trichothecene produced by *Fusarium* species. Consequently, these enzymes likely cannot provide resistance against other *Fusarium* species such as *Fusarium sporotrichioides*, which relies on T-2 toxin, among others, as a virulence factor. This report aims to establish a structural framework that will aid in the identification and engineering of new UGTs capable of glycosylating a broad range of trichothecenes including T-2 toxin. We accomplish this by presenting seven crystal structures of the rice UGT Os79 in complex with various trichothecene substrates and products, along with extensive kinetic analysis with eight chemically diverse trichothecene substrates. These data provide detailed structural and biochemical information that establishes the basis for trichothecene UGT specificity and substrate binding. Using this information, we designed an Os79 variant that exhibits specificity towards a wide range of type A and B trichothecene, including T-2 toxin. This, together development of a powerful selection protocol capable of identifying trichothecene resistance genes from large gene libraries, represent significant advances in the effort to combat FHB.

Acknowledgments

First and foremost, I would like to thank my adviser, Dr. Ivan Rayment, for his advice, guidance, and instruction throughout my graduate school career. I learned a great deal about science, research, and mentoring from Ivan, and I am grateful to have had him as an advisor. I would also like to thank all the members of the Rayment and Holden Labs for all the help and support that they have provided, especially Dr. Keenan Taylor, Dr. Becky Phillips, Mike Andreas, Dr. Haley Brown and Dr. Ari Salinger. Additionally, I would like to thank Dr. Norma Duke at the Structural Biology Consortium beamlines at the Advanced Photon Source for help with data collection and processing, and Dr. Susan McCormick and Dr. Mark Busman for producing and analyzing trichothecenes for my studies. Finally, I would like to thank my committee members, Dr. Aaron Hoskins, Dr. Brian Fox, Dr. Nancy Keller, and Dr. Vatsan Raman for all of the instrumental feedback and guidance throughout my time in graduate school. I would especially like to thank Dr. Aaron Hoskins and members of his lab for their help with my yeast and radiation research.

On a more personal note, I would like to thank all of the amazing friends that I have made in my time at the University of Wisconsin-Madison. Their support and encouragement made graduate school an incredible experience. I would especially like to thank Markus Nevil, Alex Justen, Dr. Tyler Stanage, Dr. Haley Brown, Megan Dowdle, Stephanie Maciuba, and Ana Lindahl for making my time in Madison unforgettable. Lastly, I would like to thank Christina and my family, Anna, Lynn, and Markus, for all of their support. I love you all!

Table of Contents

Chapter 1:

Fusarium head blight and trichothecene mycotoxins

This work was performed in collaboration with Dr. Ivan Rayment.

***Fusarium* head blight is a devastating plant disease**. *Fusarium* head blight (FHB) and *Fusarium* ear rot are devastating plant diseases that affect wheat (*Triticum* spp.), barley (*Hordeum vulgare*), maize (*Zea mays*), and other small grain cereals on a global scale (*1*). Infection is caused by ascomycete fungi of the genus *Fusarium* and can lead to crop losses of 10-60% (*1*). While there are dozens of *Fusarium* species that cause FHB around the world, members of the *Fusarium graminearum* species complex are the most prevalent agents of FHB, especially in North America, China, and the southern hemisphere (*2*). *Fusarium culmorum* and *Fusarium sporotrichioides* are more frequently found in the cooler regions of the world such as the United Kingdom, Northern Europe, and Canada (*3*).

Perhaps even more harmful than the devastating losses in crop yields, *Fusarium* species contaminate infected crops with highly toxin trichothecene mycotoxins which act as virulence factors during infection (*4*). The yield reduction and contamination of crops by FHB has led to enormous economic losses in North America and around the world. In fact, in the late 1990s the Midwestern United States alone lost an estimated $2.7 billion to FHB outbreaks (*5*). Such devastating occurrences have led to increased public and private funding for FHB research. Researchers have attempted to solve the problem with a verity of approaches, including genetic based crop resistance, the use of fungicides, and cultural practices such as tillage and crop rotation (*1, 6*).

Despite the funding and effort dedicated to combating FHB, efforts to prevent the disease have been largely unsuccessful. Indeed, a review of the state of wheat diseases published in 2018 found that very little progress has been made in combating FHB (*6*). Unfortunately, FHB remains a worldwide threat to cereal crop production and human and animal health. One of the most promising approaches to combating FHB is inactivation of the highly toxic trichothecene virulence

factors that *Fusarium* species rely in during infection. The research presented in this thesis focuses on this strategy, and several exciting developments in the area of enzymatic inactivation of trichothecenes will be presented. In order to understand the context of this research and the importance of trichothecenes to *Fusarium* infection, an understanding of *Fusarium* fungi and the mechanisms by which they infect cereal crops is necessary.

***Fusarium* infection of cereal crops**. *Fusarium graminearum* overwinters on infested crop residue resting on the soil surface (*7*). In the spring, flask-like fruiting bodies (perithecium) develop, and wet spring weather leads to the discharge of ascospores (*8*) (Figure 1). Specifically, rainfall and increasing relative humidity induce changes in osmotic pressure within the perithecium which lead to the build-up of turgor pressure. This pressure is released as the ascospores are ejected through a small slit (ostiole) at the tip of the perithecium (*8*) (Figure 1). Interestingly, *Fusarium* ascospores can be launched at speeds of 34 m s^{-1} with an acceleration of 870000 *g*, one of the fastest recorded accelerations in a biological organism (*9, 10*).

Ascospores (sexual reproduction) are considered the primary inoculum of *Fusarium graminearum* (*11*), but it has been suggested that splash dispersal of conidia (asexual reproduction) may also play a role in infection for *Fusarium graminearum* (*12*) and *Fusarium culmorum* (*13*). The weather-dependence of ascospore release is likely the reason that abundant rainfall during flowering and early grain development has been shown to correlate with severe FHB outbreaks (*14*). This weather-dependence is also probably why it has been found that severe FHB outbreaks typically only occur every four to five years in the United States, China, Europe, Africa, and Brazil (*6*), as wet, rainy spring weather may not occur every year.

Fusarium fungi can enter cereal florets either passively through natural openings, such as stomata, or actively by direct penetration (*15*). After entering the florets, *Fusarium* rapidly activate

a host of genes including various *Tri* genes, which are responsible for trichothecene biosynthesis (*16, 17*). The production of trichothecenes allow *Fusarium* fungi to spread from the site of initial infection. Specifically, it has been shown that when infected with *Fusarium* that are unable to produce trichothecenes, plants are able to stop the infection from spreading from the floret by developing a strong cell wall around the site of infection (*18*). This defense response is inhibited by trichothecenes, allowing trichothecene-producing *Fusarium* species to spread rapidly form the site of infection (*18*). This would suggest that inactivating trichothecenes is a promising approach for preventing the spread of fungal infection.

Trichothecene mycotoxins. Trichothecene mycotoxins are potent inhibitors of eukaryotic protein synthesis, which is thought to be their primary mechanism of toxicity in humans and animals (*19, 20*). Structural studies have demonstrated that trichothecenes bind in the A-site of the 60S peptidyl transferase center of the *Saccharomyces cerevisiae* ribosome (*21*). However, due to low resolution and poor data quality, the specific interactions between the rRNA of the ribosome and the trichothecenes could not be determined (*21*). These studies confirmed that the trichothecenes T-2 toxin and deoxynivalenol (DON) bind the eukaryotic ribosome in the same location, an interesting discovery given that these two trichothecenes represent two of the more chemically different trichothecenes from the more than 200 identified (*3, 22-25*) (Figure 2).

Trichothecenes form a highly diverse group of tricyclic sesquiterpenoid epoxides, all of which are characterized by a 12,13-epoxytrichothec-9-ene skeleton (Figure 2) (*23*). Variations in the substitution pattern of positions R1-R5 of the trichothecene skeleton account for the enormous diversity of these mycotoxins. Substitutions on the R1-R5 positions can vary in size and complexity from relatively small hydroxyls to larger moieties like acetyl and isovaleryl groups (Figure 2).

A single *Fusarium* species can produce a diverse set of trichothecenes, and different *Fusarium* species produce different combination of trichothecenes. Twenty-three *Fusarium* species and some of the trichothecenes that they produce are presented in Table 1. Table 1 includes only a subset of known *Fusarium* species, and it is likely that some of the species listed produce additional trichothecenes that are not reported here. *Fusarium culmorum*, *Fusarium graminearum*, and *Fusarium sporotrichioides*, three of the most prevalent FHB causing species, are bolded in Table 1.

As seen in Table 1, *Fusarium graminearum* synthesizes DON, nivalenol (NIV), and its acetylated derivative 4-ANIV, while T-2 toxin, HT-2, 4-ANIV, and DAS are produced by *Fusarium sporotrichioides* (*23, 25*). In addition to the trichothecene diversity of *Fusarium graminearum* and *Fusarium sporotrichioides*, Table 1 illustrates that the many different *Fusarium* species rely on diverse sets of trichothecenes as virulence factors (*24*). Not only can individual *Fusarium* species rely on different sets of trichothecenes, but a given *Fusarium* species can change which trichothecenes it produces as virulence factors. Indeed, between the years of 1998 and 2004 in western Canada the DON and 3ADON chemotype of *Fusarium graminearum* increased more than 14-fold over the DON and 15-ADON variety (*26*).

The trichothecene profile of a given *Fusarium* species is an significant factor when considering the impact on human and animal health, as varying substitutions on the trichothecene skeleton can have a dramatic effect on the toxicity to both plants and animals (*27*). In addition to the identities of trichothecenes present, the abundance of trichothecenes in harvested crops is also an important factor. Many studies have investigated the toxicity and abundance of trichothecenes in harvested crops (*28-31*).

Grain from wheat infected in the laboratory has shown DON levels as high as 500 mg/kg (mg/kg = ppm) 12 days after infection (*30*), while DON levels have been reported as high as 20 mg/kg in grain from severely affected fields (*1, 28*). These levels are well above the limit of 0.5 ppm for DON in products for final human consumptions set by the Food and Agriculture Organization of the United Nations, and they highlight the need for *Fusarium* resistant wheat. At 20 mg/kg DON, even grain from severely affected fields does not exceed the oral LD_{50} in mice of 78 mg/kg (*29*). Although wheat is the primary concern given its importance as a staple grain, barley is also a concern (*1*).

With the increase in popularity of craft beer in recent years, several studies have analyzed beer for the presence of trichothecenes. In fact, a recent study that tested 374 beers from around the world found that 348 and 289 beers (93 and 77%, respectively) contained deoxynivalenol 3-*O*-glucoside (D3G) and DON at levels above the limit of detection (*31*). The average concentration of D3G and DON across all beers tested was 0.007 ppm and 0.008 ppm respectively, with the highest contamination being 0.08 ppm D3G and 0.09 ppm DON (*31*). While none of these levels are above the accepted limit for DON in food consumption, the fact that most beers tested contained DON is concerning. Consuming even low amounts of DON can lead to immunological problems, vomiting, skin dermatitis, and hemorrhagic lesions (*32, 33*). Given the significant health risks that trichothecenes pose for both humans and animals, and the important role they play as virulence factors for *Fusarium* infection, inactivating trichothecenes is highly desirable and a key step in combating FHB and the health risks that it poses.

Inactivation of trichothecene mycotoxins. Trichothecenes are extremely stable small molecules, making their inactivation a challenge. Because they are heat stable, they are not degraded in normal food processing or autoclaving (*34, 35*), and their stability in highly acidic

conditions makes them stable in the stomach and after ingestion (*34, 36*). The toxicity of trichothecenes can be reduced in several ways, including 3-*O*-acetylation (*37*), 3-*O*-glucosylation (*38*), elimination of the characteristic 12,13-epoxide bond (*39*), and others (*40, 41*). Although the biochemical mechanism responsible for the reduced toxicity of glycosylated, acetylated, and deepoxidated trichothecenes has not been demonstrated, one explanation is that these modifications disrupt trichothecene binding to the A-site of the peptidyl transferase center of the eukaryotic ribosome.

Fusarium species have been shown to protect themselves from trichothecenes in part by the 3-*O*-acetylation of trichothecenes (*37, 42*). Although 3-*O*-acetylation has not been reported in plants, 3-*O*-glucosylation has been shown to play a significant role in trichothecene inactivation in plants. Conversion of DON to D3G was first observed in maize suspension cultures (*43*), and the first glucosyltransferase capable of glycosylating DON (DOGT1) was identified in *Arabidopsis thaliana* by screening an *Arabidopsis* cDNA library in a DON-sensitive yeast strain. DOGT1 conferred increased tolerance to DON when constitutively overexpressed in *Arabidopsis*, although it did not provide protection against NIV (*38*).

It has also been shown that an increased ability to form D3G in wheat is responsible for increased resistance to both the bleaching effects of DON and fungal spreading (*44*) and D3G has a significantly decreased ability to inhibit wheat ribosomes *in vitro* (*38*). Additionally, transcriptome analysis in a genotype of barley that has been shown to convert DON to D3G revealed the upregulation of several UDP-glucosyltransferases likely responsible for the glycosylation of DON (*45*).

Based on their ability to detoxify DON, plant UGTs have received considerable attention as potential transgenic resistance targets. Several UGTs from *Arabidopsis,* wheat, and barley with

the potential to form D3G have been identified based on homology to DOGT1 or based on transcriptome analysis following treatment with *Fusarium* or DON (*44-47*). Excitingly, transgenic wheat expressing *Hv*UGT13248, a UGT from barley, exhibited high levels of Type II resistance to *Fusarium graminearum* in greenhouse and field trials (*48*). Type II resistance is defined as the ability to resist the spread of disease, while Type I resistance is characterized by resistance to initial infection (*49*). The Type II resistance demonstrated by the transgenic wheat strain expressing *Hv*UGT13248 is likely the result of glycosylation of DON (*48*) and NIV (*50*) by the enzyme.

As previously discussed, trichothecenes inhibit plant defense mechanisms that are crucial to preventing the spread of *Fusarium* infection. It is consistent, then, that plants capable of inactivating DON by glycosylation display Type II resistance. These promising studies demonstrating resistance to the spread of *Fusarium* infection in transgenic crops expressing UGTs illustrate the importance of finding UGTs that can glycosylate not just DON and NIV, but also the many other diverse trichothecenes produced by *Fusarium* species.

Several UGT genes potentially associated with DON detoxification have been identified in *Arabidopsis thaliana* (*38*), wheat (*46*), barley (*45, 51*), *Brachypodium distachyon* (*52*), and rice (*53*). When tested for their ability to confer DON resistance in sensitized yeast, however, only one (*Hv*UGT13248) of four DON-induced barley UGT genes and two of six *Arabidopsis* UGT genes showed protection (*51*). This illustrates the difficulty in predicting the specificity of UGTs.

The difficulty in predicting which UGTs are capable of glycosylating trichothecenes highlights the need for structural information on trichothecene UGTs. Additionally, the success of DON UGTs at providing resistance to *Fusarium graminearum* illustrates the need for enzymes with broader or altered specificity capable of glycosylating the wide range of trichothecenes that are produced by various *Fusarium* species. To generate such enzymes, a more complete

understanding of the structural and enzymatic features that control specificity and activity in trichothecene UGTs is required. These factors prompted the work presented in Chapter 2.

Thesis content and scope. Chapter 2 of this thesis presents the first crystal structures of a trichothecene UGT (Os79) along with kinetic data for four trichothecene substrates. Three crystal structures of Os79 are presented, including Os79 in complex with UDP in an open conformation, in complex with UDP in a closed conformation, and in complex with UDP-2-fluoro-2-deoxy-D-glucose and trichothecene. Steady-state kinetic analysis is also presented, which demonstrates that Os79 glycosylates multiple trichothecene substrates such as DON, NIV, isotrichodermol, and HT-2 toxin, but not T-2 toxin. The crystal structures, along with the kinetic data, establish a foundation for understanding the catalytic mechanism and specificity of trichothecene UDP-glucosyltransferases. With a basic understanding of the structure of the UGT trichothecene binding pocket established, the next goal was to understand the structural determinants for specificity and leverage this knowledge to expand the number of trichothecenes Os79 is capable of glycosylating.

In Chapter 3, we present four additional structures of Os79, including one in complex with the glycosylated product D3G. We also present an extensive kinetic analysis of thirteen Os79 mutants with eight trichothecene substrates. This work establishes the basis of Os79 specificity and reveals that mutations to the acceptor binding pocket can broaden the specificity of the enzyme to include all trichothecenes tested. These data can be used to guide future transgenic crop efforts, which aim to increase resistance to the diverse group of trichothecenes utilized by *Fusarium* species responsible for infection of cereal crops around the world.

Although the development of a UGT with broad specificity towards trichothecene substrates is an enormous step forward in the effort to combat FHB, UGTs are not well suited for

inactivating trichothecenes in harvested grain due to the reversible nature of glycosylation. Removal of the glucose moiety restores the toxicity of the trichothecene (*38*). A logical alternative to consider is hydrolysis of the 12,13-epoxide bond. Unlike modifications at C3, hydrolysis of the epoxide bond is unlikely to be reversed in animal tissue. Unfortunately, at present, no enzymes have been identified that are capable of hydrolyzing the 12,13-epoxide. This was the motivation behind the development of a selection strategy aimed at identifying trichothecene resistance enzymes, which is presented in Chapter 4.

Chapter 4 outlines the development of a selection protocol which is capable of selecting for a trichothecene resistance gene in a library of over 1 million non-resistance genes. This selection protocol is a powerful method that can be used to identify enzymes capable of hydrolyzing the 12,13-epoxide bond of trichothecene from large genetic libraries.

Chapter 5 presents the structure and kinetic analysis of an aconitate isomerase from *Shewanella oneidensis* (PrpF). This work played a critical role in developing my skillset as a structural biologist.

Appendix I presents a patent application that was submitted to the United States Patent Office by the Wisconsin Alumni Research Foundation (WARF) on the enzyme that we designed using the data presented in chapters 2 and 3.

In Appendix II, we present the functional characterization of a soluble NADPH-cytochrome P450 reductase from *Fusarium graminearum* that could play a role in trichothecene biosynthesis.

References

1. McMullen, M., Bergstrom, G., De Wolf, E., Dill-Macky, R., Hershman, D., Shaner, G., and Van Sanford, D. (2012) A Unified Effort to Fight an Enemy of Wheat and Barley: *Fusarium* Head Blight, *Plant Dis 96*, 1712-1728.

2. Osborne, L. E., and Stein, J. M. (2007) Epidemiology of *Fusarium* head blight on small-grain cereals, *Int J Food Microbiol 119*, 103-108.

3. Bottalico, A., and Perrone, G. (2002) Toxigenic *Fusarium* species and mycotoxins associated with head blight in small-grain cereals in Europe, *Eur J Plant Pathol 108*, 611-624.

4. Proctor, R. H., Hohn, T. M., and McCormick, S. P. (1995) Reduced virulence of *Gibberella zeae* caused by disruption of a trichothecene toxin biosynthetic gene, *Mol Plant Microbe Interact 8*, 593-601.

5. Nganje, W. E., Bangsund, D. A., Leistritz, F. L., Wilson, W. W., and Tiapo, N. M. (2004) Regional economic impacts of *Fusarium* Head Blight in wheat and barley, *Rev Agr Econ 26*, 332-347.

6. Figueroa, M., Hammond-Kosack, K. E., and Solomon, P. S. (2018) A review of wheat diseases-a field perspective, *Mol Plant Pathol 19*, 1523-1536.

7. Trail, F., Gaffoor, I., Guenther, J. C., and Hallen, H. E. (2005) Using genomics to understand the disease cycle of the *fusarium* head blight fungus, *Gibberella zeae* (anamorph *Fusarium graminearum*), *Can J Plant Pathol 27*, 486-498.

8. Trail, F., and Common, R. (2000) Perithecial development by *Gibberella zeae*: a light microscopy study, *Mycologia 92*, 130-138.

9. Trail, F., Gaffoor, I., and Vogel, S. (2005) Ejection mechanics and trajectory of the ascospores of *Gibberella zeae* (anamorph *Fuarium graminearum*), *Fungal Genet Biol 42*, 528-533.

10. Trail, F., Xu, H. X., Loranger, R., and Gadoury, D. (2002) Physiological and environmental aspects of ascospore discharge in *Gibberella zeae* (anamorph *Fusarium graminearum*), *Mycologia 94*, 181-189.

11. Fernando, W. G. D., Miller, J. D., Seaman, W. L., Seifert, K., and Paulitz, T. C. (2000) Daily and seasonal dynamics of airborne spores of *Fusarium graminearum* and other *Fusarium* species sampled over wheat plots, *Can J Bot 78*, 497-505.

12. Jenkinson, P., and Parry, D. W. (1994) Splash Dispersal of Conidia of *Fusarium-Culmorum* and *Fusarium-Avenaceum*, *Mycol Res 98*, 506-510.

13. Wagacha, J. M., and Muthomi, J. W. (2007) *Fusarium culmorum*: Infection process, mechanisms of mycotoxin production and their role in pathogenesis in wheat, *Crop Prot 26*, 877-885.

14. Lacey, J., Bateman, G. L., and Mirocha, C. J. (1999) Effects of infection time and moisture on development of ear blight and deoxynivalenol production by *Fusarium* spp. in wheat, *Ann Appl Biol 134*, 277-283.

15. Walter, S., Nicholson, P., and Doohan, F. M. (2010) Action and reaction of host and pathogen during *Fusarium* head blight disease, *New Phytol 185*, 54-66.

16. Brown, N. A., Evans, J., Mead, A., and Hammond-Kosack, K. E. (2017) A spatial temporal analysis of the *Fusarium graminearum* transcriptome during symptomless and symptomatic wheat infection, *Molecular Plant Pathology 18*, 1295-1312.

17. Lysoe, E., Seong, K. Y., and Kistler, H. C. (2011) The Transcriptome of *Fusarium graminearum* During the Infection of Wheat, *Mol Plant Microbe In 24*, 995-1000.

18. Jansen, C., von Wettstein, D., Schafer, W., Kogel, K. H., Felk, A., and Maier, F. J. (2005) Infection patterns in barley and wheat spikes inoculated with wild-type and trichodiene synthase gene disrupted *Fusarium graminearum*, *P Natl Acad Sci USA 102*, 16892-16897.

19. Cundliffe, E., Cannon, M., and Davies, J. (1974) Mechanism of Inhibition of Eukaryotic Protein-Synthesis by Trichothecene Fungal Toxins, *P Natl Acad Sci USA 71*, 30-34.

20. Pestka, J. J. (2010) Deoxynivalenol: mechanisms of action, human exposure, and toxicological relevance, *Arch Toxicol 84*, 663-679.

21. de Loubresse, N. G., Prokhorova, I., Holtkamp, W., Rodnina, M. V., Yusupova, G., and Yusupov, M. (2014) Structural basis for the inhibition of the eukaryotic ribosome, *Nature 513*, 517-+.

22. Logrieco, A., Mule, G., Moretti, A., and Bottalico, A. (2002) Toxigenic *Fusarium* species and mycotoxins associated with maize ear rot in Europe, *Eur J Plant Pathol 108*, 597-609.

23. McCormick, S. P., Stanley, A. M., Stover, N. A., and Alexander, N. J. (2011) Trichothecenes: From Simple to Complex Mycotoxins, *Toxins 3*, 802-814.

24. Moretti, A., Susca, A., and Springer Science+Business Media (2016) *Mycotoxigenic fungi : methods and protocols*, Humana Press ; Springer, New York, NY.

25. Munkvold, G. P. (2017) *Fusarium* Species and Their Associated Mycotoxins, *Methods Mol Biol 1542*, 51-106.

26. Ward, T. J., Clear, R. M., Rooney, A. P., O'Donnell, K., Gaba, D., Patrick, S., Starkey, D. E., Gilbert, J., Geiser, D. M., and Nowicki, T. W. (2008) An adaptive evolutionary shift in *Fusarium* head blight pathogen populations is driving the rapid spread of more toxigenic *Fusarium graminearum* in North America, *Fungal Genet Biol 45*, 473-484.

27. Anderson, D. W., Black, R. M., Lee, C. G., Pottage, C., Rickard, R. L., Sandford, M. S., Webber, T. D., and Williams, N. E. (1989) Structure-activity studies of trichothecenes: cytotoxicity of analogues and reaction products derived from T-2 toxin and neosolaniol, *J Med Chem 32*, 555-562.

28. Berthiller, F., Dall'Asta, C., Schuhmacher, R., Lemmens, M., Adam, G., and Krska, R. (2005) Masked mycotoxins: determination of a deoxynivalenol glucoside in artificially and naturally contaminated wheat by liquid chromatography-tandem mass spectrometry, *J Agric Food Chem 53*, 3421-3425.

29. Forsell, J. H., Jensen, R., Tai, J. H., Witt, M., Lin, W. S., and Pestka, J. J. (1987) Comparison of acute toxicities of deoxynivalenol (vomitoxin) and 15-acetyldeoxynivalenol in the B6C3F1 mouse, *Food Chem Toxicol 25*, 155-162.

30. Savard, M. E., Sinha, R. C., Seaman, W. L., and Fedak, G. (2000) Sequential distribution of the mycotoxin deoxynivalenol in wheat spikes after inoculation with *Fusarium graminearum*, *Can J Plant Pathol 22*, 280-285.

31. Varga, E., Malachova, A., Schwartz, H., Krska, R., and Berthiller, F. (2013) Survey of deoxynivalenol and its conjugates deoxynivalenol-3-glucoside and 3-acetyl-deoxynivalenol in 374 beer samples, *Food Addit Contam A 30*, 137-146.

32. Pestka, J. J. (2008) Mechanisms of deoxynivalenol-induced gene expression and apoptosis, *Food Addit Contam A 25*, 1128-1140.

33. Ueno, Y. (1985) The Toxicology of Mycotoxins, *Crc Cr Rev Toxicol 14*, 99-132.

34. Lauren, D. R., and Smith, W. A. (2001) Stability of the *Fusarium* mycotoxins nivalenol, deoxynivalenol and zearalenone in ground maize under typical cooking environments, *Food Addit Contam 18*, 1011-1016.

35. Wolf-Hall, C. E., Hanna, M. A., and Bullerman, L. B. (1999) Stability of deoxynivalenol in heat-treated foods, *J Food Protect 62*, 962-964.

36. Duffy, M. J., and Reid, R. S. (1993) Measurement of the Stability of T-2 Toxin in Aqueous-Solution, *Chem Res Toxicol 6*, 524-529.

37. Kimura, M., Kaneko, I., Komiyama, M., Takatsuki, A., Koshino, H., Yoneyama, K., and Yamaguchi, I. (1998) Trichothecene 3-*O*-acetyltransferase protects both the producing organism and transformed yeast from related mycotoxins - Cloning and characterization of *Tri101*, *J Biol Chem 273*, 1654-1661.

38. Poppenberger, B., Berthiller, F., Lucyshyn, D., Sieberer, T., Schuhmacher, R., Krska, R., Kuchler, K., Glossl, J., Luschnig, C., and Adam, G. (2003) Detoxification of the *Fusarium* mycotoxin deoxynivalenol by a UDP-glucosyltransferase from *Arabidopsis thaliana*, *J Biol Chem 278*, 47905-47914.

39. Binder, J., Horvath, E. M., Heidegger, J., Ellend, N., Danner, H., Krska, R., and Braun, R. (1997) A bioassay for comparison of the toxicity of trichothecenes and their microbial metabolites, *Cereal Res Commun 25*, 489-491.

40. He, J. W., Hassan, Y. I., Perilla, N., Li, X. Z., Boland, G. J., and Zhou, T. (2016) Bacterial Epimerization as a Route for Deoxynivalenol Detoxification: the Influence of Growth and Environmental Conditions, *Front Microbiol 7*.

41. Carere, J., Hassan, Y. I., Lepp, D., and Zhou, T. (2018) The enzymatic detoxification of the mycotoxin deoxynivalenol: identification of DepA from the DON epimerization pathway, *Microb Biotechnol 11*, 1106-1111.

42. Kimura, M., Shingu, Y., Yoneyama, K., and Yamaguchi, I. (1998) Features of Tri101, the trichothecene 3-*O*-acetyltransferase gene, related to the self-defense mechanism in *Fusarium graminearum*, *Biosci Biotech Bioch 62*, 1033-1036.

43. Sewald, N., Vongleissenthall, J. L., Schuster, M., Muller, G., and Aplin, R. T. (1992) Structure Elucidation of a Plant Metabolite of 4-Desoxynivalenol, *Tetrahedron-Asymmetr 3*, 953-960.

44. Lemmens, M., Scholz, U., Berthiller, F., Dall'Asta, C., Koutnik, A., Schuhmacher, R., Adam, G., Buerstmayr, H., Mesterhazy, A., Krska, R., and Ruckenbauer, P. (2005) The ability to detoxify the mycotoxin deoxynivalenol colocalizes with a major quantitative trait locus for *fusarium* head blight resistance in wheat, *Mol Plant Microbe In 18*, 1318-1324.

45. Gardiner, S. A., Boddu, J., Berthiller, F., Hametner, C., Stupar, R. M., Adam, G., and Muehlbauer, G. J. (2010) Transcriptome Analysis of the Barley-Deoxynivalenol Interaction: Evidence for a Role of Glutathione in Deoxynivalenol Detoxification, *Mol Plant Microbe In 23*, 962-976.

46. Ma, L. L., Shang, Y., Cao, A. Z., Qi, Z. J., Xing, L. P., Chen, P. D., Liu, D. J., and Wang, X. E. (2010) Molecular cloning and characterization of an up-regulated UDP-glucosyltransferase gene induced by DON from *Triticum aestivum* L. cv. Wangshuibai, *Mol Biol Rep 37*, 785-795.

47. Shin, S., Torres-Acosta, J. A., Heinen, S. J., McCormick, S., Lemmens, M., Paris, M. P. K., Berthiller, F., Adam, G., and Muehlbauer, G. J. (2012) Transgenic *Arabidopsis thaliana* expressing a barley UDP-glucosyltransferase exhibit resistance to the mycotoxin deoxynivalenol, *J Exp Bot 63*, 4731-4740.

48. Li, X., Shin, S., Heinen, S., Dill-Macky, R., Berthiller, F., Nersesian, N., Clemente, T., McCormick, S., and Muehlbauer, G. J. (2015) Transgenic Wheat Expressing a Barley UDP-Glucosyltransferase Detoxifies Deoxynivalenol and Provides High Levels of Resistance to *Fusarium graminearum*, *Mol Plant Microbe In 28*, 1237-1246.

49. Mesterhazy, A. (1995) Types and components of resistance to *Fusarium* head blight of wheat, *Plant Breeding 114*, 377-386.

50. Li, X., Michlmayr, H., Schweiger, W., Malachova, A., Shin, S., Huang, Y. D., Dong, Y. H., Wiesenberger, G., McCormick, S., Lemmens, M., Fruhmann, P., Hametner, C., Berthiller, F., Adam, G., and Muehlbauer, G. J. (2017) A barley UDP-glucosyltransferase inactivates nivalenol and provides *Fusarium* Head Blight resistance in transgenic wheat, *J Exp Bot 68*, 2187-2197.

51. Schweiger, W., Boddu, J., Shin, S., Poppenberger, B., Berthiller, F., Lemmens, M., Muehlbauer, G. J., and Adam, G. (2010) Validation of a Candidate Deoxynivalenol-Inactivating UDP-Glucosyltransferase from Barley by Heterologous Expression in Yeast, *Mol Plant Microbe In 23*, 977-986.

52. Schweiger, W., Pasquet, J. C., Nussbaumer, T., Paris, M. P. K., Wiesenberger, G., Macadre, C., Ametz, C., Berthiller, F., Lemmens, M., Saindrenan, P., Mewes, H. W., Mayer, K. F. X., Dufresne, M., and Adam, G. (2013) Functional Characterization of Two Clusters of *Brachypodium distachyon* UDP-Glycosyltransferases Encoding Putative Deoxynivalenol Detoxification Genes, *Mol Plant Microbe In 26*, 781-792.

53. Michlmayr, H., Malachova, A., Varga, E., Kleinova, J., Lemmens, M., Newmister, S., Rayment, I., Berthiller, F., and Adam, G. (2015) Biochemical Characterization of a Recombinant UDP-glucosyltransferase from Rice and Enzymatic Production of Deoxynivalenol-3-*O*-beta-D-glucoside, *Toxins 7*, 2685-2700.

Tables

Table 1. *Fusarium* species and their associated trichothecene mycotoxins (*3, 22-25*). Bolded names indicate *Fusarium* species known to be the most prevalent causes of FHB.

	DAS	T-2 toxin	DON	HT-2	NIV	4,15diANIV	4-ANIV
Fusarium acuminatum	x	x		x			
Fusarium armeniacum		x		x			
Fusarium asiaticum			x		x		
Fusarium austroamericanum			x		x		
Fusarium boothii			x				
Fusarium camptoceras						x	
Fusarium cortaderiae					x		
Fusarium crookwellense					x	x	
Fusarium culmorum			x		x		
Fusarium episphaeria							x
Fusarium equiseti	x	x			x		x
Fusarium graminearum			x		x		x
Fusarium kyushuense		x			x		
Fusarium langsethiae	x	x		x			
Fusarium meridionale					x		
Fusarium mesoamericanum			x		x		
Fusarium nivale							x
Fusarium oxysporum		x		x			x
Fusarium poae	x						
Fusarium pseudograminearum			x		x		
Fusarium sambucinum	x	x		x		x	x
Fusarium semitectum	x						x
Fusarium sporotrichioides	x	x		x			x

Figures

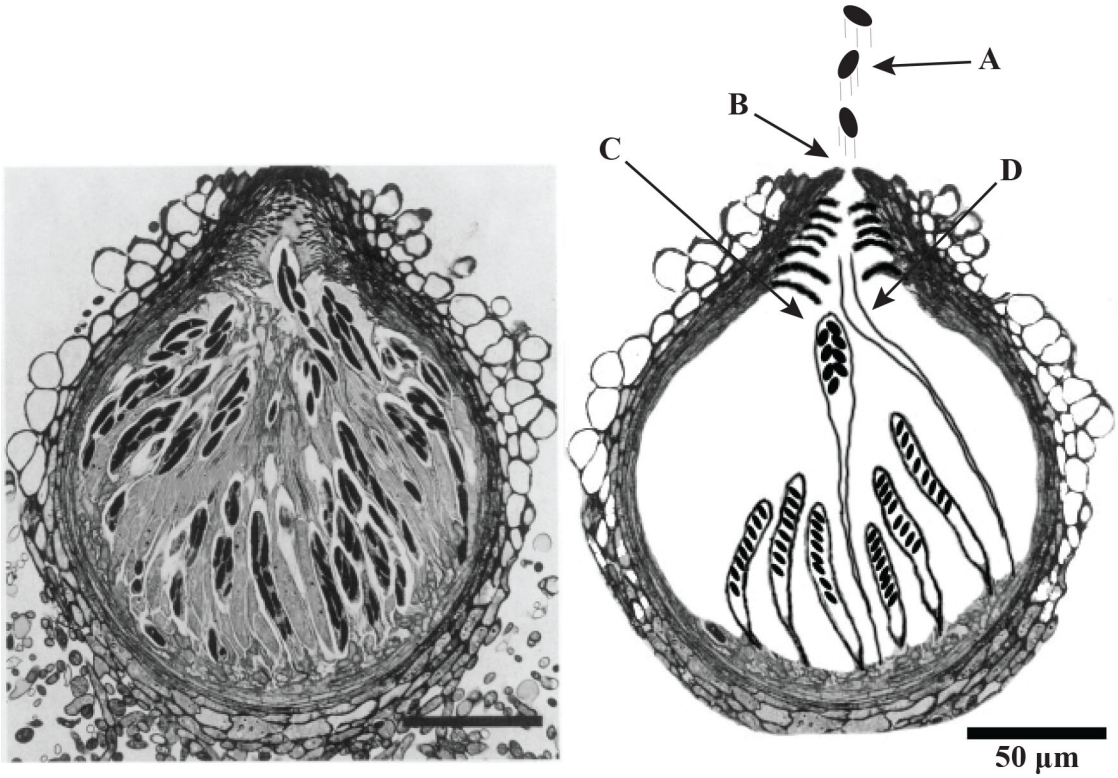


Figure 1. Longitudinal section of a mature perithecium from *Fusarium graminearum*. LEFT perithecium imaged by Trail *et al.* using light microscopy (*8*). RIGHT figure created to highlight the relevant parts using the microscopy image as a guide. (A) Ascospore after ejection from the perithecium. (B) ostiole (C) mature ascus, extending to the ostiolar canal to discharge. (D) An ascus retracting following discharge of ascospores.

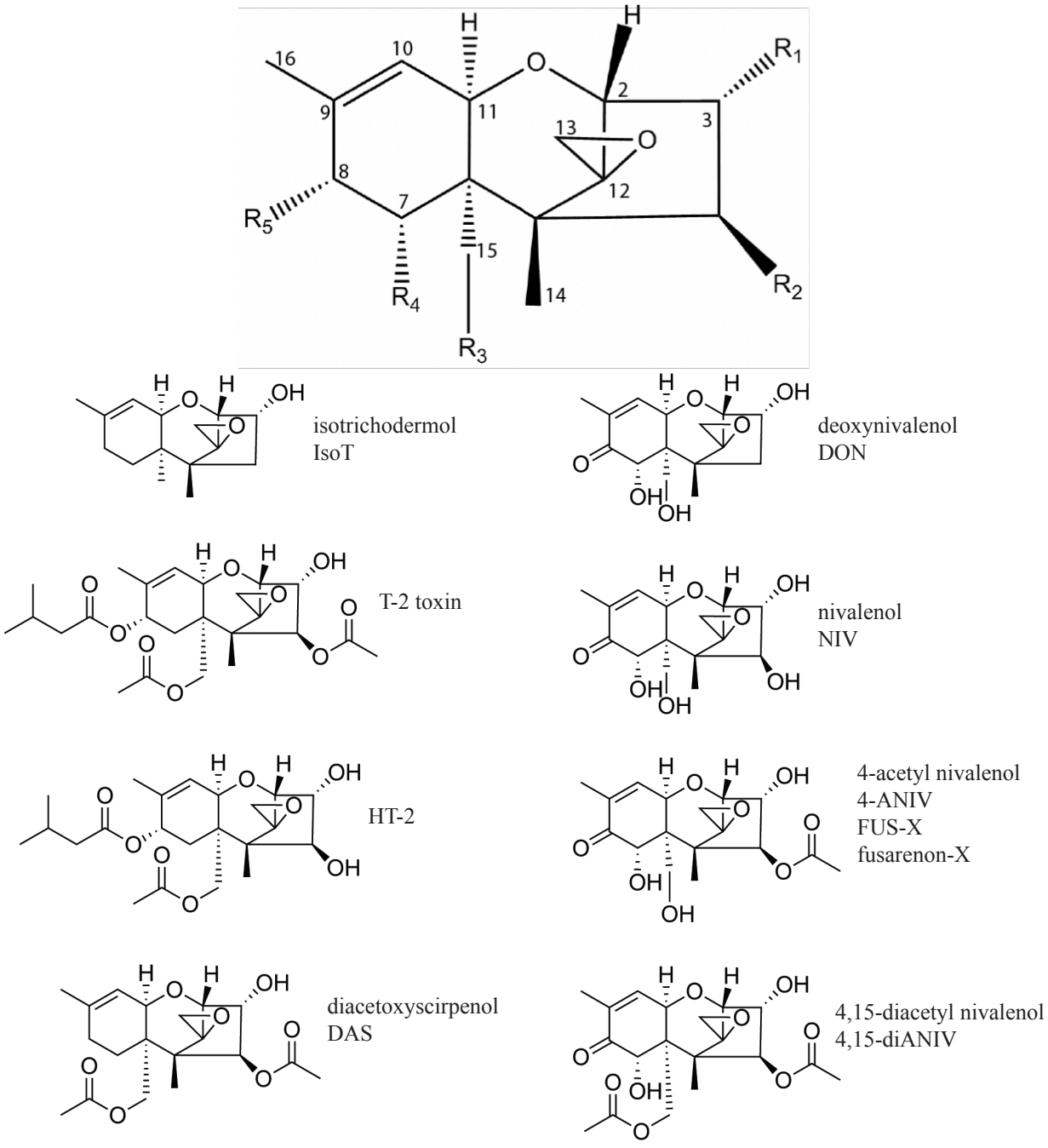


Figure 2. (Top) The conserved 12,13-epoxytrichothec-9-ene skeleton of trichothecenes with labeled carbon atoms and R-groups. (Bottom) Structures of eight trichothecenes relevant to this thesis with their full and abbreviated names directly to the right of each structure. Many trichothecenes are referred to by multiple names in the literate, when this is the case, all relevant names have been included. Over 200 trichothecenes have been identified (*3, 22-25*).

Chapter 2:

Crystal structure of Os79 (Os04g0206600) from *Oryza sativa*: A UDP-glucosyltransferase involved in the detoxification of deoxynivalenol

X-ray coordinates for the Os79·UDP complex in both open and closed conformations and Os79 complexed with UDP-2-deoxy- 2-fluoroglucose and trichothecene have been deposited with the Research Collaboratory for Structural Bioinformatics as Protein Data Bank entries 5TME, 5TMB, and 5TMD.

This work was performed in collaboration with Rachell Caniza, Dr. Sean Newmister, Dr. Mark Busman, Dr. Susan McCormick, Dr. Franz Berthiller, Dr. Gerhard Adam, and Dr. Ivan Rayment. Dr. Susan McCormick produced and purified all other trichothecenes. Dr. Mark Busman and Dr. Susan McCormick analyzed trichothecene products by LCMS. Sean Newmister performed all experiments from cloning to structure determination for the SeMet Os79 structure. All other cloning, protein purification, and X-ray crystallography were performed by Karl Wetterhorn with help from undergraduate research assistant Rachell Caniza. All structural determination, structural refinement, and kinetic analysis were performed by Karl Wetterhorn.

A version of this work was previously published as:

Wetterhorn KM, Newmister SA, Caniza RK, Busman M, McCormick SP, Berthiller F, Adam G, Rayment I. *Crystal structure of Os79 (Os04g0206600) from* Oryza sativa*: A UDP-glucosyltransferase involved in the detoxification of deoxynivalenol.* Biochemistry. 2016; 55(44): 6175-86.

Abbreviations

D3G, deoxynivalenol-3-*O*-glucoside

DON, deoxynivalenol

HEPES, 4-(2-hydroxyethyl)-1-piperazineethanesulfonic acid

MBP, maltose binding protein

MPD, 2-methyl-2,4-pentanediol

NiNTA, nickel-nitrilotriacetic acid

NIV, nivalenol

Os79*,* the product of the Os04g0206600 gene (original seq. NM_00105879)

SeMet, Seleno-Methionine

TCEP, tris(2-carboxyethyl)phosphine

TEV, Tobacco etch virus protease

U2F, UDP-2-fluoro-2-deoxy-D-glucose

TRI, trichothecene

Abstract

Fusarium head blight is a plant disease with significant agricultural and health impacts which affects cereal crops such as wheat, barley, and maize, and is characterized by reduced grain yield and the accumulation of trichothecene mycotoxins such as deoxynivalenol (DON). Studies have identified trichothecene production as a virulence factor in *Fusarium graminearum*, and have linked DON resistance to the ability to form DON-3-*O*-glucoside in wheat. Here, the structures of a deoxynivalenol:UDP-glucosyltransferase (Os79) from *Oryza sativa* are reported in complex with UDP in an open conformation, in complex with UDP in a closed conformation, and in complex with UDP-2-fluoro-2-deoxy-D-glucose and trichothecene at 1.8, 2.3, and 2.2 Å resolution respectively. The active site of Os79 lies in a groove between the N-terminal acceptor and the C-terminal donor-binding domains. Structural alignments reveal that Os79 likely utilizes a catalytic mechanism similar to other plant UGTs, with His 27 activating the trichothecene O3 hydroxyl for nucleophilic attack at C1′ of the UDP-glucose donor. Kinetic analysis of mutant Os79 revealed that Thr 291 plays a critical role in catalysis as either a catalytic acid or to position the UDP moiety during the nucleophilic attack. Steady-state kinetic analysis demonstrated that Os79 conjugates multiple trichothecene substrates such as DON, nivalenol, isotrichodermol, and HT-2 toxin, but not T-2 toxin. These data establish a foundation for understanding substrate specificity and activity in this enzyme, and can be used to guide future efforts to increase DON resistance in cereal crops.

Introduction

Fusarium head blight and *Fusarium* ear rot are devastating plant diseases that affect small grain cereals and maize on a global scale. Infection is caused by fungi of the genus *Fusarium*, where members of the *Fusarium graminearum* species complex are the most prevalent agents (*1, 2*). The trichothecene mycotoxins produced by these fungi are virulence factors and accumulate in the grain of infected plants. They are potent inhibitors of eukaryotic protein synthesis (*3*), posing a significant threat to both animal and human consumers (*4*).

Trichothecene mycotoxins are a highly diverse group of tricyclic sesquiterpenoid epoxides (*5*). Over 200 trichothecenes have been identified, all of which are characterized by a 12,13-epoxytrichothec-9-ene skeleton (Figure 1) (*6, 7*). Variations in the substitution pattern of the trichothecene backbone exist in different producing organisms. *F. graminearum* synthesizes either deoxynivalenol (DON) and its acetylated derivatives 3-acetyl-deoxynivalenol and 15-acetyl-deoxynivalenol, or nivalenol (NIV) and acetylated derivatives, while T-2 toxin and HT-2 toxin are produced by *F. sporotrichioides* (Figure 1) (*6*). Differences in substitution can have a dramatic effect on the toxicity to both plants and animals and have important implications in developing resistance strategies (*8*).

Previous research has been directed at understanding the role of the trichothecene mycotoxins in plant infection. Most efforts have focused on DON because it is the most prevalent mycotoxin associated with *Fusarium* head blight (*9, 10*). There is strong evidence to suggest that DON is a virulence factor for plant pathogenesis (*11*). Disruption of trichothecene biosynthesis by knocking out trichodiene synthase (*tri5*) led to *Fusarium* that were still capable of causing infection, however with decreased virulence in wheat and decreased ability to spread from the infection site (*11, 12*). The ability of DON to spread ahead of the fungus, causing bleaching and

necrosis, is likely due to the inhibition of protein synthesis caused by the toxin and thereby facilitates the spread of infection in host tissue. As a result DON resistance is considered an important component of *Fusarium* disease resistance (*13*).

One mechanism of DON resistance in plants is the ability to convert DON to deoxynivalenol-3-*O*-glucoside (D3G). It was shown that an increased ability to form D3G in hexaploid wheat was responsible for an increased resistance to both the bleaching effects of DON and fungal spreading (*13*). Additionally, transcriptome analysis of DON-treated barley revealed the up-regulation of several UDP-glucosyltransferases in a genotype that was shown to convert DON to D3G (*14*). D3G has a significantly decreased ability to inhibit wheat ribosomes *in vitro* (*15*). The first UGT capable of synthesizing D3G was cloned from *Arabidopsis thaliana* (*DOGT1*), and conferred increased tolerance to DON when constitutively overexpressed in *Arabidopsis* (*15*). Interestingly, DOGT1 overexpression did not confer protection against nivalenol, and was accompanied by a dwarfism phenotype associated with conversion of the brassinosteroid brassinolide to the inactive brassinolide-23-*O*-glucoside (*16*). Subsequently, other UGT genes potentially associated with DON detoxification have been identified in *Arabidopsis*, wheat (*17*) and barley (*14*). When tested for their ability to confer DON resistance in sensitized yeast, only one (*HvUGT13248*) out of four DON induced barley UGT genes and two out of six *Arabidopsis* UGT genes showed protection (*18*). This illustrates the difficulty in predicting which UGTs have the desired DON specificity among members of the very large gene family with 160-180 genes in diploid crop plants (*19-21*); however, this is a promising approach to controlling FHB. An additional problem is that the UGT genes are frequently located in gene clusters that seem to undergo rapid evolution, so that even highly similar genes in clusters have different substrate specificity (*22*).

In an effort to understand the determinants for DON specificity, the structures of an inverting deoxynivalenol:UDP-glucosyltransferase from rice, the product of the Os04g0206600 gene was determined. This enzyme was named Os79 in previous work based on the last digits of the original (removed) NCBI Reference Sequence: NM_001058779 (*22, 23*). It is highly similar to the HvUGT13249 from barley, but can be expressed efficiently in *E. coli* in contrast to this barley gene. In this work we determined the structures of Os79 at 1.8, 2.2, and 2.3 Å resolution in the open conformation in complex with UDP, open conformation in complex with UDP-2-fluoro-2-deoxy-D-glucose (U2F) and trichothecene (TRI), and closed conformation in complex with UDP, respectively.

Experimental Procedures

Trichothecenes. DON, T-2 toxin, HT-2 toxin, isotrichodermol, trichothecene, and nivalenol used in this study (Figure 1) were obtained from the USDA-ARS, Mycotoxin Prevention and Applied Microbiology Research Unit, Peoria, IL 61604, U.S.A. Deoxynivalenol-3-*O*-glucoside (D3G) was enzymatically produced and HPLC purified as previously described (*23*).

Cloning and Expression of Os79. The gene for Os79 was amplified by PCR using the forward primer 5'-ATGGGCTCTATGTCCACTCCTGC-3' and the reverse primer 5'-ATTGGAATACTTTGCTGCAAACTC-3' from plasmid pWS57 that contained a codon-optimized sequence for yeast (*23*). The resulting product was introduced into plasmid pKLD116, a pET31b derivative containing His_6-tagged maltose binding protein (MBP) followed by a TEV protease cleavage site (*24*), using an enzyme-free "Quikchange" method (*25, 26*). Os79 was overexpressed in *Escherichia col*i strain BL21 Codon Plus (DE3). Cultures from a single colony were used to inoculate 6 L Lysogeny broth (LB) supplemented with 100 μg/mL ampicillin and 30 μg/mL chloramphenicol. Expression of Os79 was induced with 1 mM isopropyl-β-D-

thiogalactopyranoside when cultures reached OD_{600} ~1.0. Induction was carried out at 16 °C for 20 h. Cells were harvested by centrifugation at 3,000 x g, washed with buffer containing 10 mM 4-(2-hydroxyethyl)-1-piperazineethanesulfonic acid (HEPES) pH 7.6 and 100 mM NaCl, and flash-frozen in liquid nitrogen. Cells were stored at -80 °C until use.

Site Directed Mutagenesis. Active site mutants of Os79 were generated using a single primer, PCR based method based on "Quikchange" mutagenesis (*25*). Mutagenic primers 5'-CCATTCCCGGCTGCTCAAGGTAACACCAATCCTATGTTACAGTTTGGAAG-3', 5'-GTAGACCAGCCAGAGTTTTAGTCTACGCCCCACATCTACCATGGGCTAGAAGAG-3', and 5'-GGTTTTGGTGTCATACGGAGTGGTTTCTACTTTTGATGTTGCTAAAC-3' were used to mutate His 27 to asparagine, Asp 120 to alanine, and Thr 291 to valine, respectively. All mutations were verified by DNA sequencing using BigDye® protocols (ABI PRISM). Reaction mixtures were resolved by the UW-Madison Biotechnology Center.

Expression of Os79 Selenomethionine Derivative. Seleno-Methionine (SeMet) Os79 was produced by metabolic inhibition (*27*). Briefly, transformed BL21 Codon Plus (DE3) cells harboring the gene encoding Os79 on pKLD116 were used to inoculate 4 L M9 minimal medium supplemented with 100 µg/mL ampicillin and 30 µg/mL chloramphenicol. An amino acid cocktail containing L-selenomethionine was added when the cells reached OD_{600} ~1.0. Cells were cooled to 16 °C and shaken for 30 min prior to induction with 1 mM isopropyl-β-D-thiogalactopyranoside. Cells were harvested as described above.

Purification of Native and Selenomethionine Os79. Native and SeMet Os79 were purified in the same manner. All purification steps were carried out on ice or at 4 °C. Twenty grams of *E. coli* cells expressing His_6-MBP-Os79 were resuspended in 120 mL buffer containing 20 mM HEPES pH 7.6, 50 mM NaCl, 0.2 mM tris(2-carboxyethyl)phosphine (TCEP), 1 mM

phenylmethanesulfonyl fluoride, and 60 mg lysozyme. Cells were lysed with 5 pulses (30 s) using a Qsonica Q700 sonicator, and the lysate was clarified by centrifugation at 40,000 rpm in a Ti 45 rotor (Beckman-Coulter) for 35 min. The concentration of NaCl and imidazole were raised to 300 mM and 20 mM, respectively, by the addition of 4 M stock solutions and loaded onto a 8 mL nickel-nitrilotriacetic acid (NiNTA; Qiagen) column equilibrated with NTA buffer: 50 mM potassium phosphate, pH 8.0, 300 mM NaCl, 20 mM imidazole, and 0.2 mM TCEP. After loading, the column was washed with 120 mL of NTA buffer. Os79 was eluted with a linear gradient of 20-300 mM imidazole in NTA buffer. Os79 containing fractions were identified by SDS-PAGE and Coomassie staining. His_6-tagged Tobacco etch virus (TEV) protease (*28*), was added at 1:40 molar ratio to cleave the His_6-MBP from Os79. The mixture was dialyzed overnight in a buffer containing 10 mM HEPES pH 7.6, 50 mM NaCl, 0.5 mM EDTA, 0.2 mM TCEP. The NaCl and imidazole concentrations were brought up to 300 mM and 20 mM, respectively, and the solution was passed over a 3 mL NiNTA column equilibrated with NTA buffer. The flow-through contained Os79, while the column retained TEV protease and undigested His_6-MBP Os79. Purified Os79 was concentrated using a stirred-cell pressure concentrator (Amicon) with a 30,000 molecular weight cutoff membrane (Millipore) to a final concentration of 12 mg/mL, estimated using calculated molar extinction coefficient of 57870 $M^{-1} \cdot cm^{-1}$ at 280 nm. The protein was dialyzed against a storage buffer containing 10 mM HEPES pH 7.6, 50 mM NaCl, 0.2 mM TCEP, drop frozen in 30 µL aliquots in liquid nitrogen, and stored at -80 °C.

Synthesis of UDP-2-fluoro-2-deoxy-D-glucose. UDP-2-fluoro-2-deoxy-D-glucose was synthesized in a reaction mixture of 100 mM triethanoamine pH 8.0, 4 mM $MgCl_2$, 10 mM β-mercaptoethanol, 1.8 mM 2-fluoro-2-deoxy-D-glucose, 1.9 mM UTP, 33 µM glucose-1,6-phosphate, 2 mM phospho(enol)pyruvic acid tri(cyclohexylammonium) salt, 60 U hexokinase, 80

U phosphoglucomutase from rabbit muscle, 200 U pyruvate kinase from rabbit muscle, 20 U Uridine-5′-diphosphoglucose pyrophosphorylase from baker's yeast, 40 U inorganic pyrophosphatase from baker's yeast (*29, 30*). The total reaction volume was 60 mL. Reactions were incubated at room temperature for 6 days after which they were filtered through a 10 kDa Amicon filter to remove the protein. Reaction mixtures were then loaded onto a 20 mL Mono Q (GE Healthcare) anion exchange column equilibrated with MilliQ water. Reaction components were eluted with a linear gradient from 30 to 35% ammonium acetate (2M, pH 4.0) over 13 CV with a flow rate of 3 mL/min. Fractions containing UDP-2-fluoro-2-deoxy-D-glucose were pooled and lyophilized until the ammonium acetate was completely removed. The identity of the product was confirmed by ESI TOF mass spectrometry (Mass Spectrometry/Proteomics Facility, University of Wisconsin) with the expected monoisotopic mass of the $[M+H]^+$ ion being 569.058 Da and the observed *m/z* being 569.040 Da.

Crystallization of Os79·UDP. Os79 was screened for initial crystallization conditions in 144-condition sparse matrix screen developed in the Rayment laboratory. Single, diffraction quality crystals were grown at 4 °C by hanging drop vapor diffusion by mixing 2 µL of 9 mg/mL Os79 in 10 mM HEPES pH 7.6, 50 mM NaCl, 5 mM UDP-phenol with 2 µL well solution containing 50 mM HEPES pH 7.6, 10 mM glycine, 13% methyl ether polyethylene glycol 5000 (MEPEG 5K), 4% ethylene glycol. Hanging droplets were nucleated after 4 h from an earlier spontaneous crystallization event using a cat's whisker. Crystals grew to approximate dimensions of 100 x 100 x 300 µm within 10 days. The crystals were transferred stepwise to a cryoprotecting solution that contained 50 mM HEPES 7.6, 50 mM NaCl, 10 mM glycine, 13% MEPEG 5K, 20% ethylene glycol, 5 mM UDP-phenol and vitrified by rapid plunging into liquid nitrogen. Os79

crystallized in the space group $P2_12_12_1$ with unit cell dimensions of a = 59.2 Å, b = 83.2 Å, c = 98.8 Å and one chain in the asymmetric unit.

Crystallization of Os79 complexed with UDP 2-deoxy-2-fluoro-glucose and trichothecene. Os79 was screened for initial crystallization conditions in 144-condition sparse matrix screen developed in the Rayment laboratory. Single, diffraction quality crystals were grown at 4 °C by hanging drop vapor diffusion by mixing 2 µL of 17 mg/mL Os79 in 10 mM HEPES pH 7.6, 10 mM 2-deoxy-2-fluoro-glucose, 10 mM trichothecene with 2 µL well solution containing 100 mM MES pH 6.5, 200 mM glycine, 23% methyl ether polyethylene glycol 2000 (MEPEG 2K). Hanging droplets were nucleated after 24 h from an earlier spontaneous crystallization event using a cat's whisker. Crystals grew to approximate dimensions of 50 x 50 x 100 µm within 6 days. The crystals were transferred directly to a cryoprotecting solution that contained 10 mM 2-deoxy-2-fluoro-glucose, 10 mM trichothecene, 100 mM MES pH 6.5, 200 mM glycine, 23% MEPEG 2K and 15% glycerol and vitrified by rapid plunging into liquid nitrogen. Os79 crystallized in the space group $P2_12_12_1$ with unit cell dimensions of a = 59.1 Å, b = 82.9 Å, c = 98.6 Å and one chain in the asymmetric unit.

Crystallization of Os79 in the closed conformation. Os79 was likewise screened for suitable crystallization conditions. Single, diffraction quality crystals were grown at 4 °C by hanging drop vapor diffusion by mixing 2 µL of 17.4 mg/mL Os79 in 10 mM HEPES pH 7.6, with 2 µL well solution containing 100 mM piperazine-*N*,*N*′-bis(2-ethanesulfonic acid) pH 6.0, 80 mM NaCl, 20% 2-methyl-2,4-pentanediol (MPD), 15% taurine. Hanging droplets were nucleated after 24 h with 1 µL of crushed crystals from an earlier spontaneous crystallization event diluted by a factor of $1{:}10^5$ in 50% well solution in 10 mM HEPES pH 7.6. Crystals grew to approximate dimensions of 50 x 50 x 200 µm within 5 days. The crystals were soaked in 50% well solution pH

7.0 in 10 mM HEPES pH 7.6 with 10 mM D3G and 5 mM UDP for 4 h The crystals were transferred stepwise to a cryoprotecting solution that contained 100 mM piperazine-*N*,*N*′-bis(2-ethanesulfonic acid) pH 7.0, 80 mM NaCl, 30% MPD, 15% taurine and vitrified by rapid plunging into liquid nitrogen. Os79 crystallized in the space group $P3_221$ with unit cell dimensions of a = 105.5 Å, b = 105.5 Å, c = 100.2 Å and one chain in the asymmetric unit.

Data Collection and Refinement. X-ray data for the Os79·UDP·open and SeMet Os79 structures were collected at 100 K on the Structural Biology Center beamline 19BM at the Advanced Photon Source in Argonne, IL. X-ray data for the Os79·UDP·closed and Os79·U2F·Trichothecene complexes were collected on beamline 19ID. Diffraction data were integrated and scaled with HKL3000 (*31*). Data collection statistics are given in Table 1. The structure of Os79 was solved using multiple wavelength anomalous dispersion for SeMet Os79. Phasing and initial model building were conducted within the HKL3000 software suite: selenium sites were located using SHELXD (*32*), experimental phases were calculated with MLPHARE (*33*), density modification was performed with DM (*34*), and initial model building was performed with Buccaneer (*35*). This resulted in an initial model that could be extended by alternating cycles of manual building and least-squares refinement using COOT (*36*) and Refmac (*37*), respectively. The structures for Os79·UDP·open, Os79·UDP·closed and Os79·U2F·Trichothecene were determined using SeMet Os79 and Os79·UDP·open, respectively as molecular replacement search models in the program Phaser (*38*). Final models were generated by alternating cycles of manual building and least-squares refinement using Coot, Phenix and Refmac (*36, 37, 39*).

Glucosyltransferase Enzymatic Assays. Steady-state kinetic analyses of Os79 with trichothecene substrates were performed in a coupled continuous enzymatic assay at 23 °C in a 1 cm path length cuvette. Reactions were initiated by the addition of 30 µL of varied concentrations

of trichothecene and 0.5 mM UDP-glucose to 70 µL of Os79 (3.0 µg/mL final concentration), 3 units rabbit muscle lactic dehydrogenase and 2 units rabbit muscle pyruvate kinase (Sigma-Aldrich, buffered aqueous glycerol solution), 1.5 mM phosphoenolpyruvate, 100 µM β-NADH, 50 mM KCl, 10 mM $MnCl_2$, and 100 mM glycylglycine pH 8.0. Lactic dehydrogenase, pyruvate kinase, phosphoenolpyruvate, β-NADH, and Os79 were added to a master-mix containing the remaining reaction components prior to the initiation of each reaction. Reaction progress was followed by monitoring the decrease in A_{340} caused by the oxidation of β-NADH. The rates of reaction were determined at various trichothecene concentrations and fit by non-linear regression to the Michaelis-Menten equation using GraphPad Prism software. Fates of the trichothecenes were examined by LC-MS/MS analysis of the assay solutions, monitoring their conversion to the respective glucosides.

Results and Discussion

Structures of Os79 in Complex with UDP. Os79 was crystallized in an effort to understand the structural basis for deoxynivalenol glycosylation, and to identify the structural components responsible for substrate specificity in this enzyme. The structures of Os79 in complex with UDP were determined to 1.8 Å resolution for an open conformation and 2.3 Å resolution for a closed conformation. The overall fold of Os79 is shown in Figure 2. The enzyme crystallized with a monomer in the asymmetric unit. In the closed conformation this monomer contains 451 of 466 expected amino acids and extends from Gln 15-Ser 465. In the open conformation two additional regions in the protein were not modeled due to disorder: Ala 24-His 27 and Leu 255-Phe 266. The first break includes the proposed catalytic base (His 27), whereas the second break between Leu 255 and Phe 266 spans the linker that connects the N-terminal acceptor and C-terminal donor domains. The tertiary structure of Os79 is consistent with a GT-B

glycosyltransferase fold, consisting of an N-terminal acceptor-binding domain and a C-terminal donor-binding domain (Figure 2) (*40*). Each domain exhibits a similar α/β Rossmann-type fold and belongs to the UDP-glycosyltransferase/glycogen phosphorylase (UDPGT) superfamily. The N-terminal acceptor-binding domain contains a central seven-stranded parallel β-sheet surrounded on both sides by a total of eight α-helices. The C-terminal donor-binding domain contains a central six-stranded parallel beta sheet surrounded on both sides by nine α-helices. These domains are arranged adjacent to one another forming a globular protein in which the substrates bind in a deep cleft between the two domains (Figure 2A). In the closed conformation the 11 unresolved amino acids that connect the N-terminal acceptor and C-terminal donor domains lay down on top of this cleft, covering UDP (Figure 2B). The closed conformation was determined by soaking apo-crystals which were grown at pH 6.0 in a solution at pH 7.0 with 10 mM of the product D3G before freezing. Density for D3G was not observed in the active site. Crystals grown and treated in an identical manner, except soaked and frozen at pH 6.0, crystallized in the same $P3_221$ space group but in the open conformation. This suggests that the transition from a disordered to ordered conformation in this crystal form is a consequence of pH and not the presence of D3G in the lattice. At this point, it is unclear whether the shift in the conformation of the loop is an artifact of crystal packing or reflects a physiologically relevant property of Os79. Regardless of the physiological relevance of the position of the loop, the pH dependent conformational shift observed in the $P3_221$ crystal form highlights the low energy barrier associated with the repositioning of the loop.

Comparison with other Plant UGTs. The overall structure of Os79 is highly similar to that of the seven other structurally characterized plant UGTs (*41-47*) with root mean square deviation (RMSD) values varying between 1.9 Å and 2.3 Å as depicted in Table 3. As expected, the areas of greatest divergence surround the acceptor site. Interestingly, many plant UGTs,

including Os79, recognize a broad range of acceptor molecules, while only some utilize UDP-linked donor sugars other than UDP-glucose (*21*). These enzymes share a highly conserved PSPG motif, which is a signature motif for plant UGTs involved in glycosylation of secondary metabolites (*48, 49*).

Substrate Binding in Os79. Crystals of Os79 in complex with UDP were obtained by co-crystallization of Os79 with either UDP-phenol or UDP-glucose. Electron density was observed for only UDP regardless of which UDP-conjugate was used indicating that hydrolysis had occurred during crystallization. In an effort to trap the donor and acceptor in the active site, Os79 was co-crystallized with the non-reactive co-substrate analogs UDP-2-fluoro-2-deoxy-D-glucose (U2F) and the compound trichothecene, which lacks the C3-hydroxyl group to which the glucose moiety is attached. This led to the determination of a structure that mimics the Michaelis complex, with U2F and trichothecene bound in the active site (Figures 3 and 4). Strong electron density allowed for the unambiguous placement of U2F, however the electron density corresponding to trichothecene was weak, representing only half occupancy (Figure 3). This coupled with the small globular shape of trichothecene made placing the acceptor in the active site difficult. The orientation of trichothecene cannot be definitively determined from the current electron density, however the orientation that best fits the density makes chemical sense. If DON was present in this orientation this would position its C3 hydroxyl 2.5 Å from C1′ of U2F and 'in-line' for nucleophilic attack. This orientation also places the C3 hydroxyl 3.5 Å from the catalytic base, His 27 (Figure 4). Trichothecene binds predominantly within the N-terminal domain. It is bound in a large, mainly hydrophobic pocket composed of Phe 21, Ala 24, Met 79, Phe 189, Phe 199, Val 292, Trp 383, and Ala 384 (Figure 5A). Other trichothecene acceptor molecules for Os79 such as DON and HT-2 (Figure 1) have hydroxyls and other polar substituents at a combination of the C4,

C7, C8, and C15 positions on the trichothecene backbone. Polar residues His 122 and Gln 202 are ideally located in the acceptor-binding pocket to interact with these groups (Figure 5A). At this stage it is difficult to accurately predict the exact orientation that DON or HT-2 might adopt in the active site because of the ambiguity in the exact orientation of trichothecene. However, examination of Figure 5A clearly suggests that there is space associated with C7, C8, and C15 that could accommodate the hydroxyl and ketone groups on DON. Likewise, there appears to be enough room to fit in the hydroxyl moiety on HT-2 at C4, however conformational changes adjacent to C8 and C15 will most likely be required to accept the isovaleryl and acetyl moieties. Indeed, trichothecene substrates are distinguished from the planar, or near planar acceptors recognized by the other known plant UGTs by their noticeably contorted tricyclic ring structure, which likely requires an acceptor site that is more flexible or open to accommodate the comparatively bulky trichothecene acceptors. Some evidence of flexibility is seen in the active site loop.

The active site loop from Ala 24-His 27 was unresolved in the open conformation of Os79 in complex with UDP. The structural basis for the increased flexibility of this loop is not immediately obvious, and is unprecedented among known UGT structures, although large movements were seen between the active site residues of two independent MtUGT71G1 in the crystallographic asymmetric unit (*41*). It is possible that the poorly conserved insertion of Ala 24 and substitution of Ala 23 from a larger, hydrophobic side chain has given rise to the increased flexibility of this loop by disrupting the close hydrophobic contacts that are observed at this position in *Mt*UGT85H2, *Mt*UGT71G1, *Vv*GT1, and *At*UGTB1 (Figure 5B). An alanine or glycine residue at the position equivalent to Ala23 in Os79 is also observed in the *Arabidopsis* DOGT1, barley *Hv*UGT13248, and sorghum Sb06g002180 which possess DON:glucosyltransferase

activity (*22*). However, the specificity for DON must be controlled by more than this loop since the same change is seen in AtUGT73C6 which shows only limited activity and no activity in Os04g0206500 (Figure 5B). An inspection of the loop in the closed conformation reveals that there are no direct contacts between any of the residues in the loop and the UDP donor (Figure 6). Thus, although there is not an immediately apparent function for the loop in substrate binding or catalysis, it may play a subtler role. As noted earlier, the observed disorder-order transition in this region appears to be mediated in part by pH.

The U2F substrate lies in a long groove in the C-terminal domain of Os79. Numerous polar contacts with the glucose, uracil base, ribose sugar, and pyrophosphate moieties are observed as depicted in Figure 6. The C2′, C3′, C4′, and C6′ hydroxyls of the glucose moiety hydrogen bond with residues Gln 386, Asp 385, Trp 364, Gln 143, and Ser 142. Most of the interactions with UDP are highly conserved throughout the PSPG containing enzymes, which is not surprising given the fact that all these enzymes recognize UDP-linked donor molecules. The binding of UDP-glucose is conserved, as can be seen in superimpositions of *Vv*GT1, *At*UGT72B1, and *Mt*UGT71G1 onto Os79 (Figure 7). Additionally, Asp 385, which forms polar contacts with the C3′ and C4′ hydroxyl groups of glucose, occupies a structurally equivalent position in Os79 as Asp 374 in *Vv*GT1 and Asp 367 in *Ct*UGT78K6 (Figure 8). These residues have been shown to be involved in donor recognition and specificity (*41*), and are indicative of a preference to utilize UDP-glucose as opposed to other UDP-linked sugars. This likely represents the structural basis for the formation of DON-3-*O*-glucoside in DON resistant cereals (*13, 14*).

Catalytic Mechanism of Os79. Structural alignments with the known plant UGTs *Ct*UGT78K6 and *Vv*GT1 indicate that Os79 likely uses a similar catalytic mechanism. The H27-D120 proposed catalytic dyad of Os79 occupies a structurally equivalent position as H20-D119 in

*Vv*GT1 and H17-D117 in *Ct*UGT78K6 (Figure 8). The 3-hydroxyl group was modeled onto the trichothecene molecule bound in the active site (Figure 4). This model places the 3-hydroxyl 3.6 Å from the nitrogen of His 27 and 2.7 Å from the anomeric C1′ carbon of the glucose. As with other inverting plant UGTs, it is proposed that His 27 serves as a general base that deprotonates the 3-hydroxyl of the trichothecene acceptor (*50*). The resulting nucleophilic oxyanion then attacks the anomeric C1′ carbon of the glucose, resulting in the formation of the glycosylated toxin and UDP (Scheme 1). To test this hypothesis, the H27N/D120A double mutant was constructed, which indeed showed no activity within the sensitivity of the coupled enzymatic assay (a decrease of 1000 in activity or $k_{cat} < 1 \times 10^{-3}$ s^{-1}). It was originally hypothesized that the β-phosphate was protonated by His 361, which is 3.7 Å from one of the phosphate oxygens (Figure 6). Catalytic activity was not decreased in the H361A mutant however, suggesting that this histidine does not serve as a catalytic acid. Thr 291, which is 3.5 Å from two of the β-phosphate oxygens (Figure 4) was also considered as the catalytic acid. The T291V and T291A mutants are both catalytically inactive ($k_{cat} < 1 \times 10^{-3}$ s^{-1}) suggesting that Thr 291 plays a significant role in the catalytic mechanism. There are two possible functions for Thr 291. It could be playing a critical structural role in positioning the phosphate during the reaction, or it could function in proton transfer to the phosphate. Given the high pKa of threonine, it is likely that any protonation of the phosphate occurs through proton exchange with solvent. The location of the hydroxyl moiety for Thr 291 is consistent with a bifurcated hydrogen bond to the phosphoryl oxygens, though at this resolution it cannot be categorically proven.

Steady-State Kinetic Assays. Steady-state kinetic constants were determined for Os79 with four trichothecene substrates using a coupled-continuous enzymatic assay (Table 2). These data show K_M values in the low micromolar range and turnover numbers ranging between

1.07 – 0.38 s^{-1}. These data are similar to the previously published K_M and k_{cat} of 0.23 ± 0.06 mM and 0.57 s^{-1}, respectively for Os79 with DON as a substrate (*23*). The 4-fold decrease in K_M and 2-fold increase in k_{cat} observed in this study are likely the result of differences in experimental design. Michlmayr *et al.* (*23*) performed their kinetic analysis by measuring the product (after a reaction time leading to a maximum of 10% substrate consumption) directly by HPLC-MS using the entire MBP-containing fusion protein at 37 °C and pH 7.0 while in this study the reactions were performed at 23 °C and pH 8.0 with MBP removed with TEV protease. The exact K_M value could not be determined with isotrichodermol due to the sensitivity limit of the coupled enzymatic assay, which quantified product formation by following the loss of A_{340} due to the oxidation of NADH. In order to uphold initial velocity conditions less than 10% of the total substrate pool (0.5 μM) was converted to product in the reaction time course. This corresponds to a change in A_{340} of 0.003 AU, which approached the noise level in the enzyme coupled assay. Regardless, isotrichodermol was the best substrate tested in terms of catalytic efficiency.

It is unlikely that plant UGTs encounter isotrichodermol in nature because isotrichodermol is an early intermediate in trichothecene biosynthesis (*51*). However, it is conceivable that isotrichodermol was the first virulence factor in evolution and that plants confronted with it developed resistance to this compound. Extension of the trichothecene biosynthetic pathway within the fungi may have led to reduced affinity to the plant UGTs, which in turn requires evolution of better enzymes in plants. In this case, isotrichodermol establishes a baseline for the activity level of the Os79 enzyme because it is the least substituted trichothecene available that carries a 3-hydroxyl substitution (trichothecene used in the structural studies is not synthesized by *Fusarium*, but rather *Trichoderma* species). DON, NIV, and HT-2 toxin are more highly substituted and Os79 shows decreased catalytic efficiency towards them mainly owing to an

increase in K_M. Only a very low level of activity was detected with T-2 toxin as the substrate in the continuous assay. However when reactions were allowed to incubate at room temperature for 4 h and analyzed by LC-MS/MS, no T-2-glucoside was detected, but a small amount of HT-2 glucoside was detected in these reactions. Examination of the T-2 toxin starting material revealed a low level of contamination with HT-2 toxin indicating that the activity observed in the coupled assay was the result of HT-2 toxin glycosylation. These data indicate that T-2 toxin is not a substrate of Os79 as noted previously based on a lack of increased resistance in yeast transformants (*23*). This implies that acetylation of the C-4 hydroxyl perturbs substrate binding as this is the only difference between HT-2 and T-2 toxin.

Conclusions

In conclusion, the structures of Os79 reported here represent the first structures of a DON:UDP-glucosyltransferase and provide insight into the catalytic mechanism responsible for trichothecene glycosylation. It should be noted that it has recently been shown that increased detoxification of DON and increased *Fusarium* resistance can be achieved by overexpression of the barley HvUGT13248 in transgenic wheat or the Bradi5g03300 UGT in the model cereal species *Brachypodium distachyon* (*52, 53*). Our structures of the Os79 UGT provide insights into substrate specificity and donor binding. Although this study provides a significant first step, further structural studies are needed in order to determine the specific binding interactions of more substituted trichothecene substrates. Such studies may provide the structural information needed to engineer Os79 and related enzymes to accept other trichothecenes such as T-2 toxin as substrates.

Acknowledgements

We gratefully acknowledge Dr. Norma E. C. Duke and Dr. Stephan L. Ginell for assistance during the X-ray data collection at Argonne.

References

1. Starkey, D. E., Ward, T. J., Aoki, T., Gale, L. R., Kistler, H. C., Geiser, D. M., Suga, H., Toth, B., Varga, J., and O'Donnell, K. (2007) Global molecular surveillance reveals novel *Fusarium* head blight species and trichothecene toxin diversity, *Fungal Genet Biol 44*, 1191-1204.

2. van der Lee, T., Zhang, H., van Diepeningen, A., and Waalwijk, C. (2015) Biogeography of *Fusarium graminearum* species complex and chemotypes: a review, *Food Addit Contam Part A Chem Anal Control Expo Risk Assess 32*, 453-460.

3. Cundliffe, E., Cannon, M., and Davies, J. (1974) Mechanism of inhibition of eukaryotic protein synthesis by trichothecene fungal toxins, *Proc Natl Acad Sci U S A 71*, 30-34.

4. Pestka, J. J. (2010) Deoxynivalenol: mechanisms of action, human exposure, and toxicological relevance, *Arch Toxicol 84*, 663-679.

5. Evans, E., Holtom, A.M., and Hanson, J.R. (1973) *J. Chem. Soc. Chem. Commun. 1973*, 465.

6. McCormick, S. P., Stanley, A. M., Stover, N. A., and Alexander, N. J. (2011) Trichothecenes: from simple to complex mycotoxins, *Toxins 3*, 802-814.

7. Grove, J. F. (2007) The trichothecenes and their biosynthesis, *Fortschr Chem Org Naturst 88*, 63-130.

8. Anderson, D. W., Black, R. M., Lee, C. G., Pottage, C., Rickard, R. L., Sandford, M. S., Webber, T. D., and Williams, N. E. (1989) Structure-activity studies of trichothecenes: cytotoxicity of analogues and reaction products derived from T-2 toxin and neosolaniol, *J Med Chem 32*, 555-562.

9. Kim, H. S., Lee, T., Dawlatana, M., Yun, S. H., and Lee, Y. W. (2003) Polymorphism of trichothecene biosynthesis genes in deoxynivalenol- and nivalenol-producing *Fusarium graminearum* isolates, *Mycol Res 107*, 190-197.

10. Larsen, J. C., Hunt, J., Perrin, I., and Ruckenbauer, P. (2004) Workshop on trichothecenes with a focus on DON: summary report, *Toxicology letters 153*, 1-22.

11. Proctor, R. H., Hohn, T. M., and McCormick, S. P. (1995) Reduced virulence of *Gibberella zeae* caused by disruption of a trichothecene toxin biosynthetic gene, *Mol Plant Microbe Interact 8*, 593-601.

12. Jansen, C., von Wettstein, D., Schafer, W., Kogel, K. H., Felk, A., and Maier, F. J. (2005) Infection patterns in barley and wheat spikes inoculated with wild-type and trichodiene synthase gene disrupted *Fusarium graminearum*, *Proc Natl Acad Sci U S A 102*, 16892-16897.

13. Lemmens, M., Scholz, U., Berthiller, F., Dall'Asta, C., Koutnik, A., Schuhmacher, R., Adam, G., Buerstmayr, H., Mesterhazy, A., Krska, R., and Ruckenbauer, P. (2005) The ability to detoxify the mycotoxin deoxynivalenol colocalizes with a major quantitative trait locus for *Fusarium* head blight resistance in wheat, *Mol Plant Microbe Interact 18*, 1318-1324.

14. Gardiner, S. A., Boddu, J., Berthiller, F., Hametner, C., Stupar, R. M., Adam, G., and Muehlbauer, G. J. (2010) Transcriptome analysis of the barley-deoxynivalenol interaction: evidence for a role of glutathione in deoxynivalenol detoxification, *Mol Plant Microbe Interact 23*, 962-976.

15. Poppenberger, B., Berthiller, F., Lucyshyn, D., Sieberer, T., Schuhmacher, R., Krska, R., Kuchler, K., Glossl, J., Luschnig, C., and Adam, G. (2003) Detoxification of the *Fusarium* mycotoxin deoxynivalenol by a UDP-glucosyltransferase from *Arabidopsis thaliana*, *J Biol Chem 278*, 47905-47914.

16. Poppenberger, B., Berthiller, F., Bachmann, H., Lucyshyn, D., Peterbauer, C., Mitterbauer, R., Schuhmacher, R., Krska, R., Glossl, J., and Adam, G. (2006) Heterologous expression of *Arabidopsis* UDP-glucosyltransferases in *Saccharomyces cerevisiae* for production of zearalenone-4-*O*-glucoside, *Appl Environ Microbiol 72*, 4404-4410.

17. Lulin, M., Yi, S., Aizhong, C., Zengjun, Q., Liping, X., Peidu, C., Dajun, L., and Xiu, E. W. (2010) Molecular cloning and characterization of an up-regulated UDP-glucosyltransferase gene induced by DON from *Triticum aestivum L. cv. Wangshuibai*, *Molecular biology reports 37*, 785-795.

18. Schweiger, W., Boddu, J., Shin, S., Poppenberger, B., Berthiller, F., Lemmens, M., Muehlbauer, G. J., and Adam, G. (2010) Validation of a candidate deoxynivalenol-inactivating UDP-glucosyltransferase from barley by heterologous expression in yeast, *Mol Plant Microbe Interact 23*, 977-986.

19. Ross, J., Li, Y., Lim, E., and Bowles, D. J. (2001) Higher plant glycosyltransferases, *Genome biology 2*, REVIEWS3004.

20. Achnine, L., Huhman, D. V., Farag, M. A., Sumner, L. W., Blount, J. W., and Dixon, R. A. (2005) Genomics-based selection and functional characterization of triterpene glycosyltransferases from the model legume *Medicago truncatula*, *The Plant journal : for cell and molecular biology 41*, 875-887.

21. Caputi, L., Malnoy, M., Goremykin, V., Nikiforova, S., and Martens, S. (2012) A genome-wide phylogenetic reconstruction of family 1 UDP-glycosyltransferases revealed the expansion of the family during the adaptation of plants to life on land, *The Plant journal : for cell and molecular biology 69*, 1030-1042.

22. Schweiger, W., Pasquet, J. C., Nussbaumer, T., Paris, M. P., Wiesenberger, G., Macadre, C., Ametz, C., Berthiller, F., Lemmens, M., Saindrenan, P., Mewes, H. W., Mayer, K. F., Dufresne, M., and Adam, G. (2013) Functional characterization of two clusters of *Brachypodium distachyon* UDP-glycosyltransferases encoding putative deoxynivalenol detoxification genes, *Mol Plant Microbe Interact 26*, 781-792.

23. Michlmayr, H., Malachova, A., Varga, E., Kleinova, J., Lemmens, M., Newmister, S., Rayment, I., Berthiller, F., and Adam, G. (2015) Biochemical Characterization of a Recombinant UDP-glucosyltransferase from Rice and Enzymatic Production of Deoxynivalenol-3-*O*-beta-D-glucoside, *Toxins 7*, 2685-2700.

24. Rocco, C. J., Dennison, K. L., Klenchin, V. A., Rayment, I., and Escalante-Semerena, J. C. (2008) Construction and use of new cloning vectors for the rapid isolation of recombinant proteins from *Escherichia coli*, *Plasmid 59*, 231-237.

25. Chen, G. J., Qiu, N., Karrer, C., Caspers, P., and Page, M. G. (2000) Restriction site-free insertion of PCR products directionally into vectors, *BioTechniques 28*, 498-500, 504-495.

26. van den Ent, F., and Lowe, J. (2006) RF cloning: a restriction-free method for inserting target genes into plasmids, *Journal of biochemical and biophysical methods 67*, 67-74.

27. Van Duyne, G. D., Standaert, R. F., Karplus, P. A., Schreiber, S. L., and Clardy, J. (1993) Atomic structures of the human immunophilin FKBP-12 complexes with FK506 and rapamycin, *J Mol Biol 229*, 105-124.

28. Blommel, P. G., and Fox, B. G. (2007) A combined approach to improving large-scale production of tobacco etch virus protease, *Protein Expr Purif 55*, 53-68.

29. Barlow, J. N., and Blanchard, J. S. (2000) Enzymatic synthesis of UDP-(3-deoxy-3-fluoro)-D-galactose and UDP-(2-deoxy-2-fluoro)-D-galactose and substrate activity with UDP-galactopyranose mutase, *Carbohydr Res 328*, 473-480.

30. Hayashi, T., Murray, B. W., Wang, R., and Wong, C. H. (1997) A chemoenzymatic synthesis of UDP-(2-deoxy-2-fluoro)-galactose and evaluation of its interaction with galactosyltransferase, *Bioorganic & medicinal chemistry 5*, 497-500.

31. Minor, W., Cymborowski, M., Otwinowski, Z., and Chruszcz, M. (2006) HKL-3000: the integration of data reduction and structure solution--from diffraction images to an initial model in minutes, *Acta Crystallogr D Biol Crystallogr 62*, 859-866.

32. Schneider, T. R., and Sheldrick, G. M. (2002) Substructure solution with SHELXD, *Acta Crystallogr D Biol Crystallogr 58*, 1772-1779.

33. Otwinowski, Z. (1991) *Proceedings of the CCP4 Study Weekend. Isomorphous Replacement and Anomalous Scattering, eds. Wolf, W., Evans, P.R. & Leslie, A.G.W. (Science and Engineering Research Council, Daresbury Laboratroy, Warrington, U.K.)*, 80-86.

34. Cowtan, K. D., and Zhang, K. Y. (1999) Density modification for macromolecular phase improvement, *Progress in biophysics and molecular biology 72*, 245-270.

35. Cowtan, K. (2006) The Buccaneer software for automated model building. 1. Tracing protein chains, *Acta Crystallogr D Biol Crystallogr 62*, 1002-1011.

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

37. Murshudov, G. N., Vagin, A. A., and Dodson, E. J. (1997) Refinement of macromolecular structures by the maximum-likelihood method, *Acta Crystallogr D Biol Crystallogr 53*, 240-255.

38. 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 Appl Crystallogr 40*, 658-674.

39. Adams, P. D., Afonine, P. V., Bunkoczi, G., Chen, V. B., Davis, I. W., Echols, N., Headd, J. J., Hung, L. W., Kapral, G. J., Grosse-Kunstleve, R. W., McCoy, A. J., Moriarty, N. W., Oeffner, R., Read, R. J., Richardson, D. C., Richardson, J. S., Terwilliger, T. C., and Zwart, P. H. (2010) PHENIX: a comprehensive Python-based system for macromolecular structure solution, *Acta Crystallogr. D Biol. Crystallogr. 66*, 213-221.

40. Hu, Y., Chen, L., Ha, S., Gross, B., Falcone, B., Walker, D., Mokhtarzadeh, M., and Walker, S. (2003) Crystal structure of the MurG:UDP-GlcNAc complex reveals common structural principles of a superfamily of glycosyltransferases, *Proc Natl Acad Sci U S A 100*, 845-849.

41. Shao, H., He, X., Achnine, L., Blount, J. W., Dixon, R. A., and Wang, X. (2005) Crystal structures of a multifunctional triterpene/flavonoid glycosyltransferase from *Medicago truncatula*, *The Plant cell 17*, 3141-3154.

42. Offen, W., Martinez-Fleites, C., Yang, M., Kiat-Lim, E., Davis, B. G., Tarling, C. A., Ford, C. M., Bowles, D. J., and Davies, G. J. (2006) Structure of a flavonoid glucosyltransferase reveals the basis for plant natural product modification, *EMBO J 25*, 1396-1405.

43. Li, L., Modolo, L. V., Escamilla-Trevino, L. L., Achnine, L., Dixon, R. A., and Wang, X. (2007) Crystal structure of Medicago truncatula UGT85H2--insights into the structural basis of a multifunctional (iso)flavonoid glycosyltransferase, *J Mol Biol 370*, 951-963.

44. Brazier-Hicks, M., Offen, W. A., Gershater, M. C., Revett, T. J., Lim, E. K., Bowles, D. J., Davies, G. J., and Edwards, R. (2007) Characterization and engineering of the bifunctional N- and O-glucosyltransferase involved in xenobiotic metabolism in plants, *Proc Natl Acad Sci U S A 104*, 20238-20243.

45. Hiromoto, T., Honjo, E., Tamada, T., Noda, N., Kazuma, K., Suzuki, M., and Kuroki, R. (2013) Crystal structure of UDP-glucose:anthocyanidin 3-*O*-glucosyltransferase from *Clitoria ternatea*, *J Synchrotron Radiat 20*, 894-898.

46. Modolo, L. V., Li, L., Pan, H., Blount, J. W., Dixon, R. A., and Wang, X. (2009) Crystal structures of glycosyltransferase UGT78G1 reveal the molecular basis for glycosylation and deglycosylation of (iso)flavonoids, *J Mol Biol 392*, 1292-1302.

47. Hiromoto, T., Honjo, E., Noda, N., Tamada, T., Kazuma, K., Suzuki, M., Blaber, M., and Kuroki, R. (2015) Structural basis for acceptor-substrate recognition of UDP-glucose: anthocyanidin 3-*O*-glucosyltransferase from *Clitoria ternatea*, *Protein Sci 24*, 395-407.

48. Hughes, J., and Hughes, M. A. (1994) Multiple secondary plant product UDP-glucose glucosyltransferase genes expressed in cassava (*Manihot esculenta Crantz*) cotyledons, *DNA sequence : the journal of DNA sequencing and mapping 5*, 41-49.

49. Paquette, S., Moller, B. L., and Bak, S. (2003) On the origin of family 1 plant glycosyltransferases, *Phytochemistry 62*, 399-413.

50. Lairson, L. L., Henrissat, B., Davies, G. J., and Withers, S. G. (2008) Glycosyltransferases: structures, functions, and mechanisms, *Annu Rev Biochem 77*, 521-555.

51. McCormick, S. P., Alexander, N. J., and Proctor, R. H. (2006) *Fusarium* Tri4 encodes a multifunctional oxygenase required for trichothecene biosynthesis, *Can J Microbiol 52*, 636-642.

52. Li, X., Shin, S., Heinen, S., Dill-Macky, R., Berthiller, F., Nersesian, N., Clemente, T., McCormick, S., and Muehlbauer, G. J. (2015) Transgenic Wheat Expressing a Barley UDP-Glucosyltransferase Detoxifies Deoxynivalenol and Provides High Levels of Resistance to *Fusarium graminearum*, *Mol Plant Microbe Interact 28*, 1237-1246.

53. Pasquet, J. C., Changenet, V., Macadre, C., Boex-Fontvieille, E., Soulhat, C., Bouchabke-Coussa, O., Dalmais, M., Atanasova-Penichon, V., Bendahmane, A., Saindrenan, P., and Dufresne, M. (2016) A *Brachypodium* UDP-glycosytransferase confers root tolerance to deoxynivalenol and resistance to *Fusarium* infection, *Plant Physiol*.

54. Livingstone, C. D., and Barton, G. J. (1993) Protein sequence alignments: a strategy for the hierarchical analysis of residue conservation, *Comput Appl Biosci 9*, 745-756.

55. Waterhouse, A. M., Procter, J. B., Martin, D. M., Clamp, M., and Barton, G. J. (2009) Jalview Version 2--a multiple sequence alignment editor and analysis workbench, *Bioinformatics 25*, 1189-1191.

Tables

Table 1. Data Collection and Refinement Statistics

complex	Os79·UDP·open	Os79·UDP SeMet	Os79·UDP·closed	Os79·U2F·Trichothecene
PDB entry	5TME	Not deposited	5TMB	5TMD
space group	$P2_12_12_1$	$P2_12_12_1$	$P3_22_1$	$P2_12_12_1$
unit cell dimensions (Å)	a = 59.2, b = 83.2, c = 98.8	a = 59.1, b = 83.1, c = 99.0	a = 105.5, b = 105.5, c = 100.2	a = 59.4, b = 83.2, c = 99.4
data collection site	Argonne 19BM	Argonne 19BM	Argonne 19ID	Argonne 19ID
Wavelength (Å)	0.979	0.979	0.979	0.979
resolution range	50–1.78 (1.81-1.78)[a]	50-2.3 (2.34-2.3)	50-2.34 (2.38-2.34)	43.5–2.19 (2.27-2.19)
reflections: measured	696068	173871	488298	696068
reflections: unique	47323	22233	27447	25837
redundancy	7.9 (4.8)	7.8 (8.2)	10.3 (5.2)	7.8 (5.9)
completeness (%)	99.9 (100)	98.2	99.5 (95.3)	99.6 (95.9)
average I/σ	47.3 (4.1)	51.3 (16.2)	24.4 (3.0)	19.1 (5.0)
R_{merge} (%)[b]	5.3 (39.5)	8.4 (25.2)	16.3 (80.1)	5.3 (39.5)
R_{work}	18.6	16.1	19.2	16.4
R_{free}	22.8	21.4	24.4	20.4
$R_{overall}$	18.8	16.4	19.5	17.0
protein atoms	3421	3406	3495	3407
ligand atoms	25	25	25	53
water molecules	199	165	221	261
average B factors	27.1	31.3	37.2	28.4
Ramachandran (%)	--	--	--	--
most favored	98.2	95.7	96.0	97.0
allowed	1.6	4.1	3.8	3.0
disallowed	0.2	0.2	0.2	0
rms deviations	--	--	--	--
bond lengths (Å)	0.02	0.018	0.008	0.02
bond angles (deg)	2.1	1.87	1.17	1.96

[a]Values in parenthesis are for highest resolution shell

[b]R_{merge}= Σ| $I_{(hkl)}$ - I | x 100 /Σ| $I_{(hkl)}$ |, where the average intensity I is taken over all symmetry equivalent measurements and $I_{(hkl)}$ is the measured intensity for a given observation.

[c]R_{factor}= Σ| $F_{(obs)}$ - $F_{(calc)}$ | x 100 /Σ| $F_{(obs)}$ |, where R_{work} refers to the R_{factor} for the data utilized in the refinement and R_{free} refers to the R_{factor} for 5% of the data that were excluded from the refinement.

Table 2. Steady state kinetic constants for various trichothecene substrates.

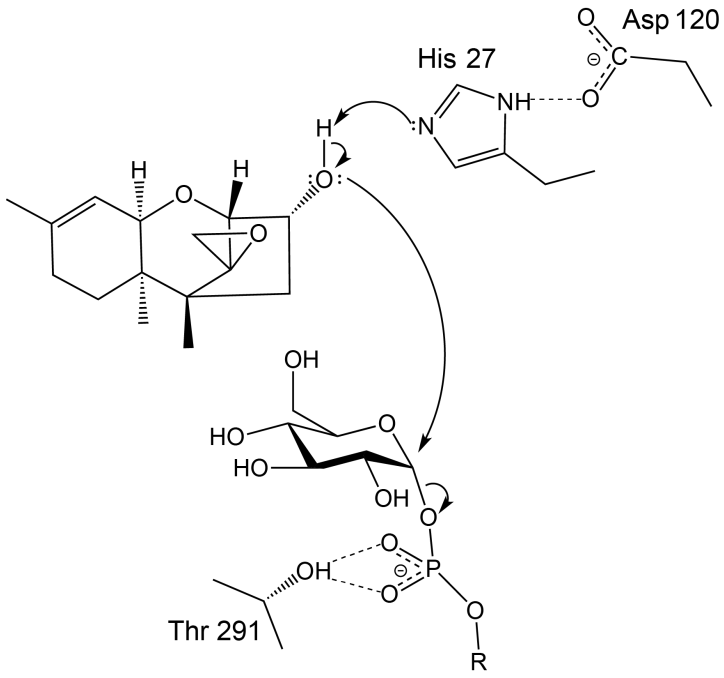
Substrate	K_M **(μM)**	k_{cat} **(s^{-1})**	k_{cat}/K_M **(s^{-1} M^{-1})**
Deoxynivalenol	61 ± 6	1.07 ± 0.04	1.7×10^4
HT-2	22 ± 2	0.85 ± 0.02	3.8×10^4
Isotrichodermol	< 1.5[a]	0.79 ± 0.02	$> 5.2 \times 10^5$
Nivalenol	35 ± 6	0.38 ± 0.02	1.1×10^4

[a]Not determined due to limited sensitivity of coupled continuous enzymatic assay.

Table 3. Comparison of Os79 with seven other plant UGT structures. Os79 is remarkably similar to these plant UGTs, with RMSD values ranging from 1.9-2.3.

PDB ID	**Species**	**Gene identification**	**Equivalent Cα's**	**RMSD (Å)**	**Reference**
2PQ6	*M. truncatula*	UGT85H2	375	1.9	(*43*)
2VCH	*A. thaliana*	UGT72B1	374	2.0	(*44*)
2C1Z	*V. vinerfera*	VvGT1	378	2.0	(*42*)
2ACW	*M. truncatula*	UGT71G1	362	2.1	(*41*)
3HBF	*M. truncatula*	UGT78G1	385	2.3	(*46*)
4REM	*C. ternatea*	UGT78K6	365	2.2	(*47*)
3WC4	*C. ternatea*	Ct3GT-A	365	2.2	(*45*)

Schemes



Scheme 1. Proposed catalytic mechanism for the glycosylation of isotrichodermol by Os79. The His 27-Asp 120 catalytic dyad serves as a general base that deprotonates the 3-hydroxyl of the trichothecene acceptor. The resulting nucleophilic oxyanion attacks the anomeric C1′ of glucose. Thr 291 plays a critical role in either positioning the phosphate or protonating it before the release of UDP. It is possible that Thr 291 serves both of these roles.

Figures

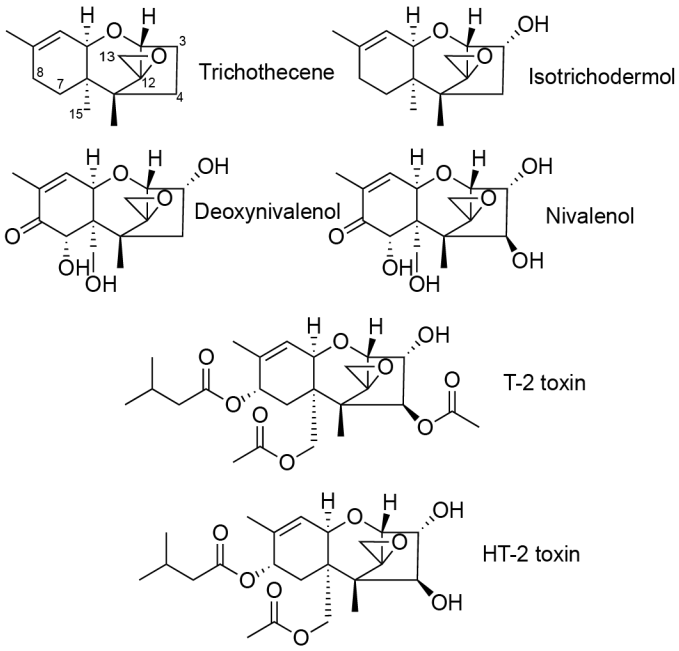


Figure 1. Trichothecene mycotoxins examined in this study. Glycosylation occurs at the 3-OH position.

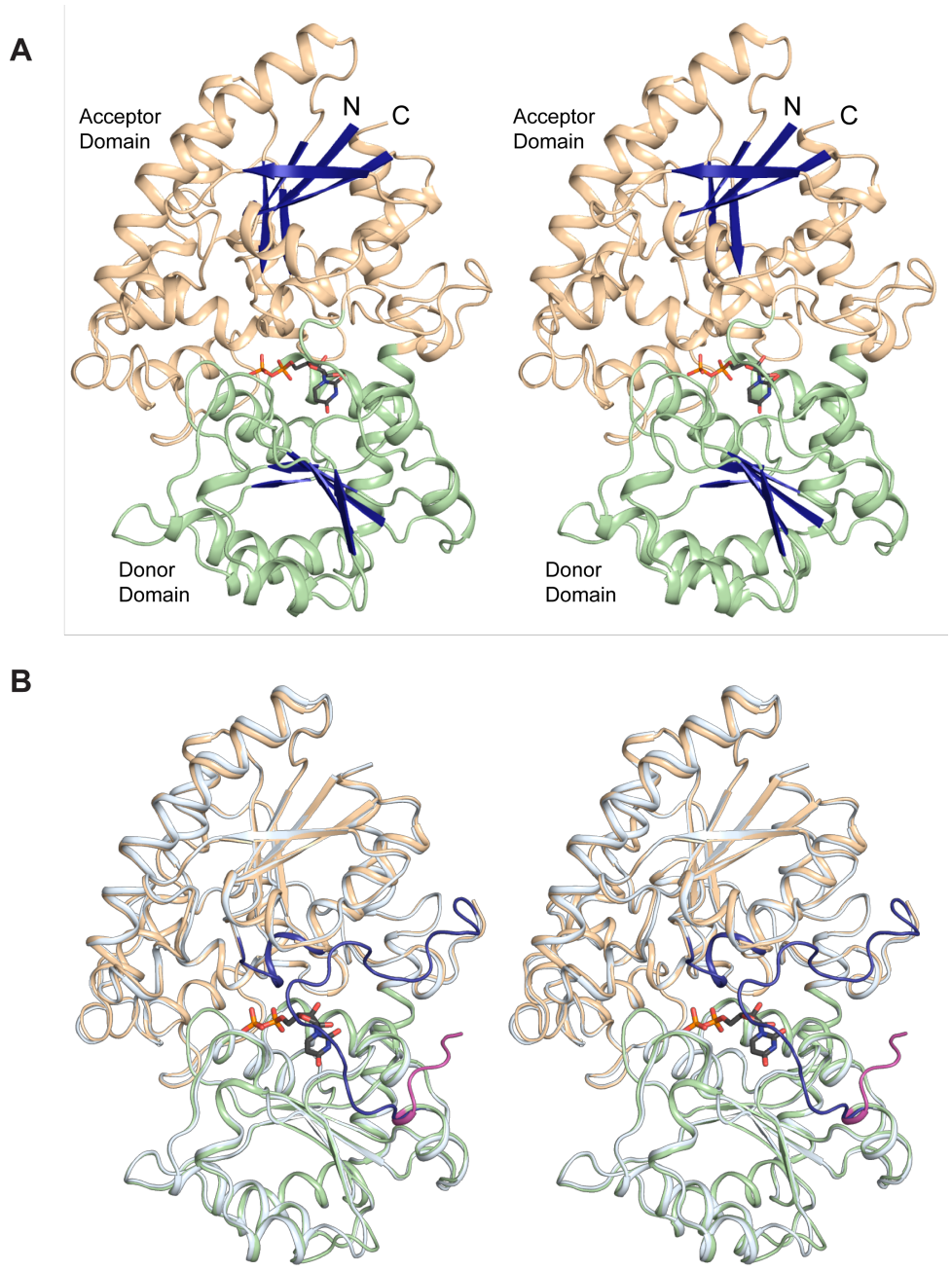


Figure 2. Stereo view of the Os79 monomer. (A) Ribbon representation of Os79 with the β-sheets in each of the two Rossmann-type α/β domains shown in blue. The substrates bind in the groove between the donor-binding domain (green) and the acceptor-binding domain (tan). UDP is shown in stick form and colored by element. (B) Comparison of the closed (white) and open (green and tan) conformations of Os79. In the closed conformation a loop (blue) covers UDP in the active site, in the open conformation this loop (magenta) is largely disordered.

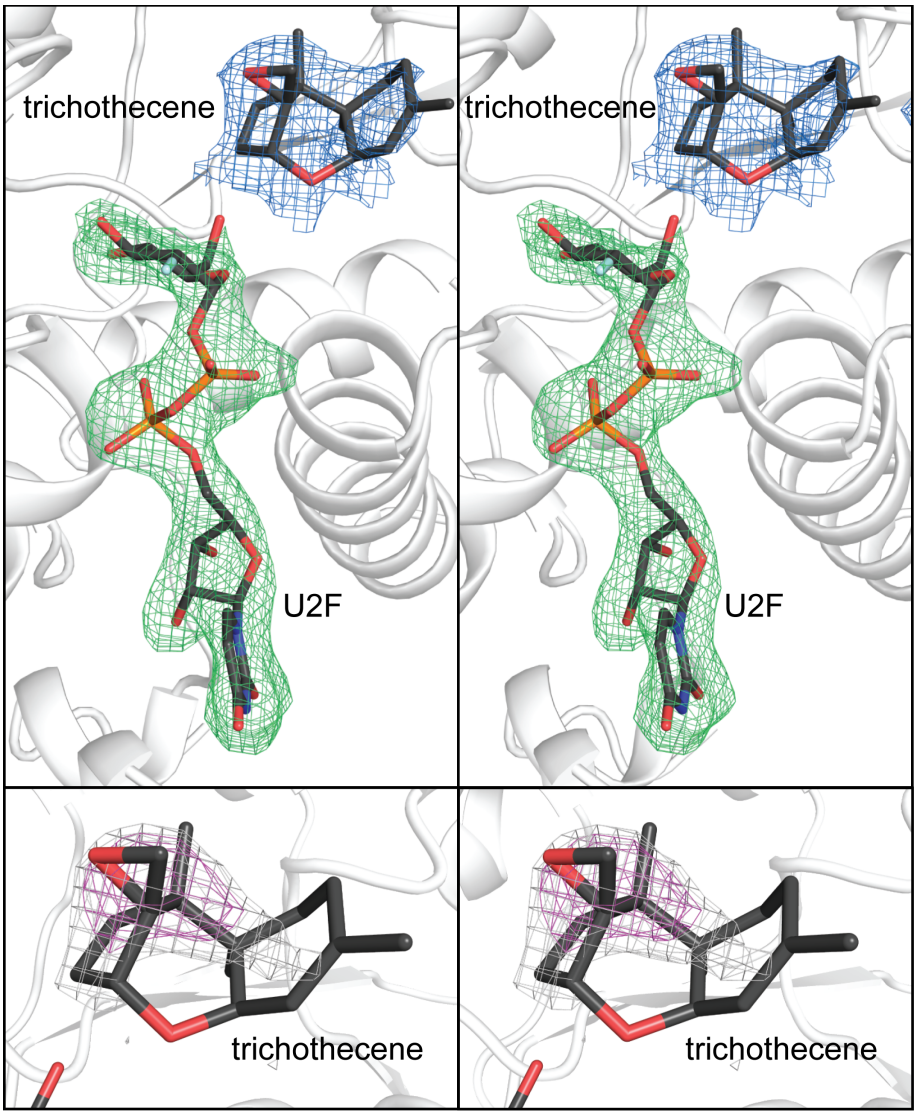


Figure 3. Electron density corresponding to the bound ligands U2F and trichothecene in complex with Os79. The maps were calculated with coefficients of the form F_o - F_c where the ligands were omitted from the final phase calculation and refinement. (Top) The U2F (green) and trichothecene (blue) maps were contoured to 3.0 σ and 1.3 σ, respectively. (Bottom) The electron density presented in the top panel corresponding to trichothecene was contoured to 2.0 σ (gray) and 2.5 σ (magenta).

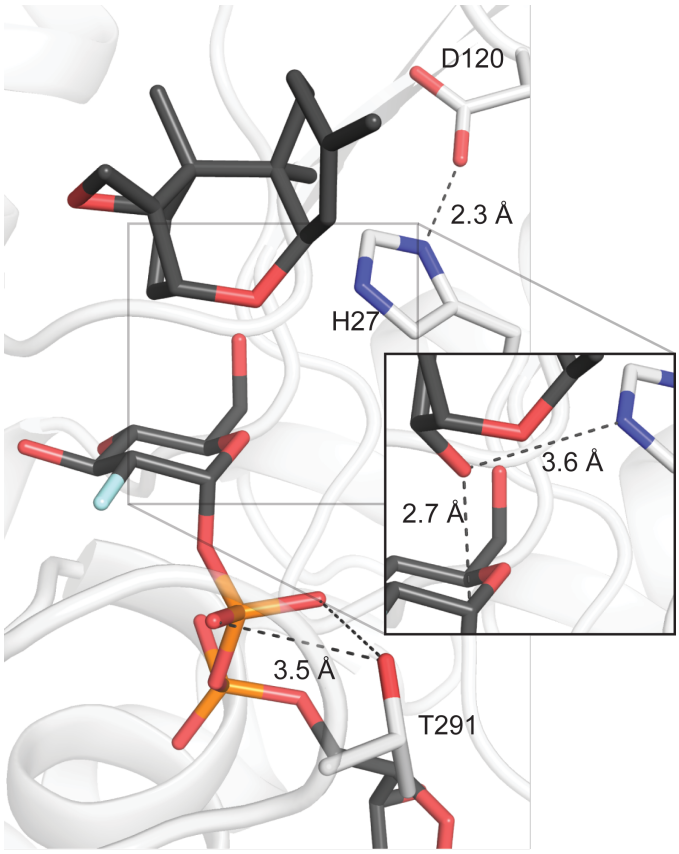


Figure 4. Active site of Os79 in complex with U2F and trichothecene. U2F and trichothecene are color by element. Inset is a model of the 3-hydroxyl of isotrichodermol in the active site.

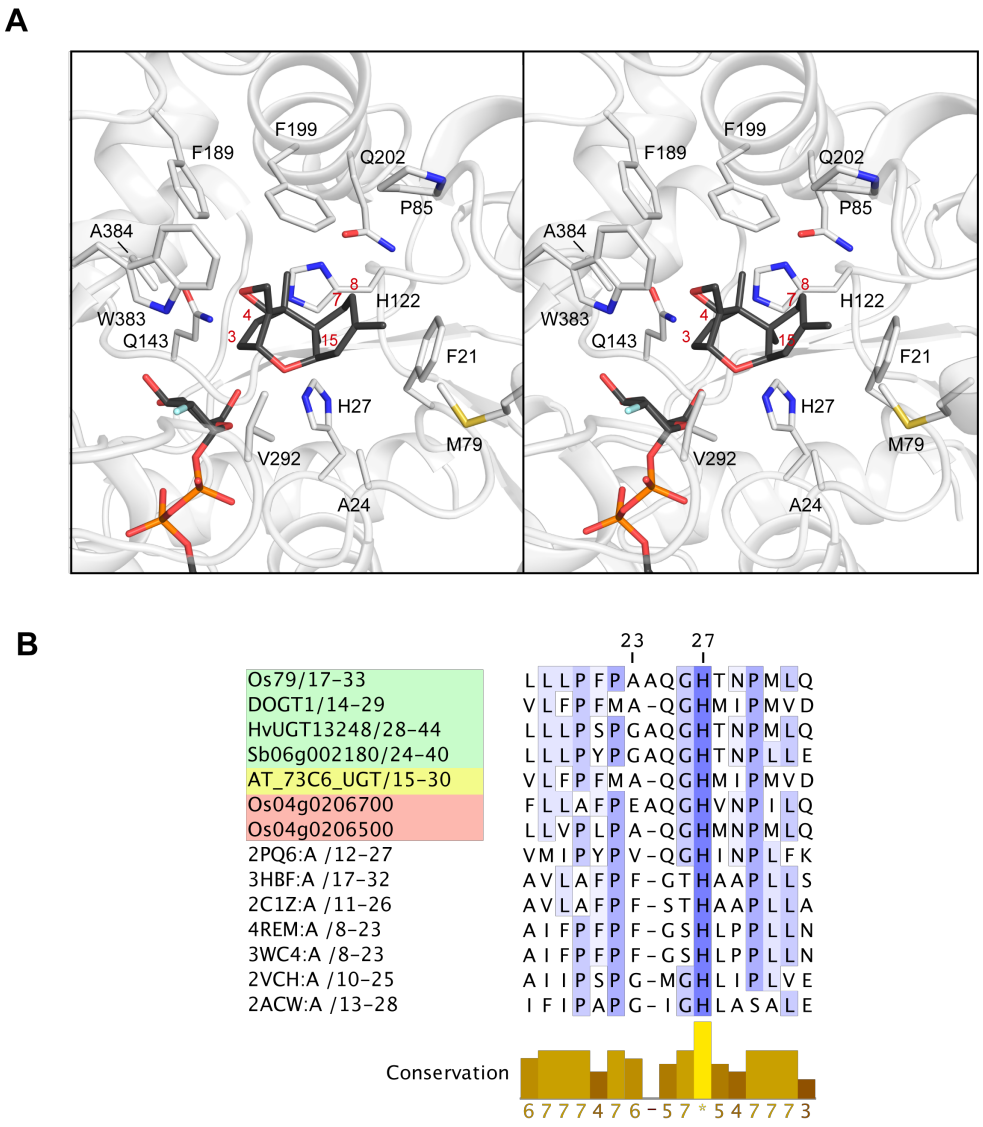


Figure 5. Stereo view of trichothecene bound in the active site of Os79 and sequence conservation surrounding the active site loop. (A) Trichothecene, U2F, and Os79 residues surrounding trichothecene in the acceptor-binding pocket are colored by element. The atoms substituted in DON, T2, and HT-2 are labeled. The only residues depicted that are not in the N-terminal domain are Phe 189, Ala 384, Trp 383 and Val 292. (B) Sequence alignment for the residues surrounding the active site loop that undergo a conformational transition in Os79 (Ala24-His27). The numbers in the conservation plot represent the degree of conservation, with one being the lowest, 9 being the highest, and a "*" indicating a completely conserved residue (*54, 55*). The sequences correspond to the structurally known plant UGTs, as identified by their RCSB accession codes, and closely related UGTs that either show DON glucosyltransferase activity, slight activity, or no such activity (highlighted in green, yellow, and red, respectively). **Os04g0206500 and** Os04g0206700 (formerly **NM_001058778.1 and** NM_001058780.1) are the most closely related UGTs to Os79 and do not show activity (*22*).

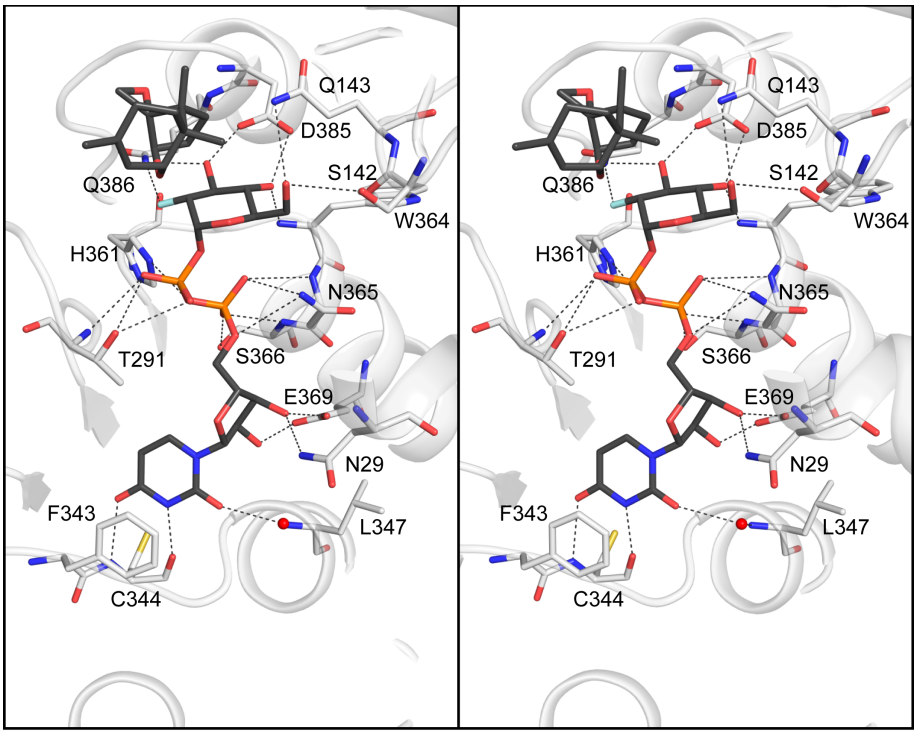


Figure 6. Stereo view of the Os79 polar interactions with the U2F donor molecule. Trichothecene, U2F, and Os79 residues surrounding U2F in the donor-binding pocket are colored by element.

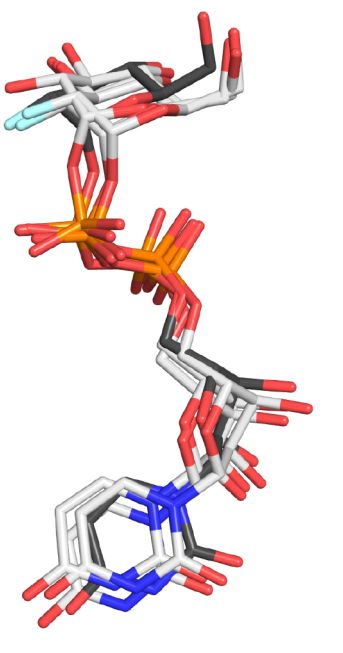


Figure 7. Structural alignments show that donor binding is conserved in plant UGT enzymes. U2F, UDP-Glc and U2F from the plant UGTs *At*UGT72B1, *Mt*UGT71G1, and *Vv*GT1, respectively are colored by element with white carbons. U2F from Os79 is colored by element with black carbons. All the donors are represented based on structural alignments of the respective protein molecules. This conservation of donor-binding implies a conserved catalytic mechanism among these plant UGTs.

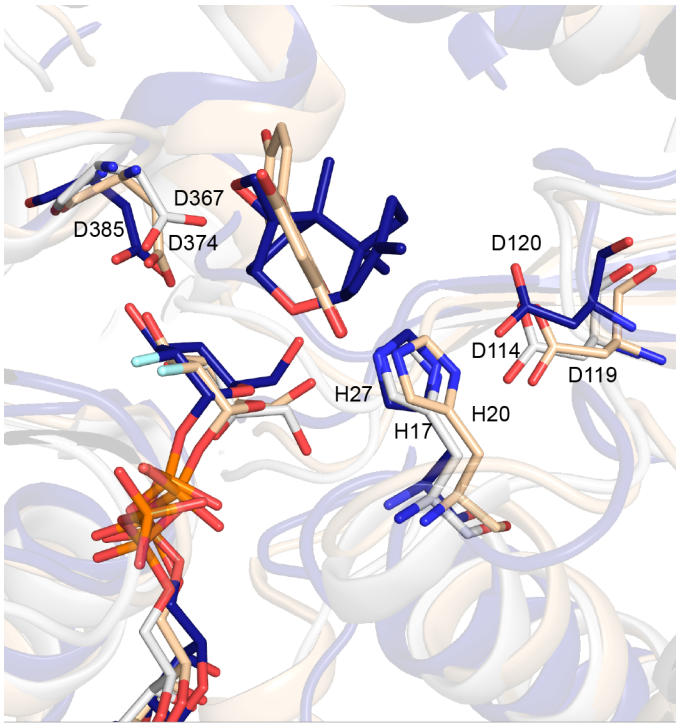


Figure 8. Alignment of the active sites of Os79, *Ct*UGT78K6, and *Vv*GT1. Structural alignments of Os79 (blue) with *Ct*UGT78K6 (white) and *Vv*GT1 (tan) show a similar arrangement of active site residues, suggesting a similar mechanism for Os79. This figure also shows that the trichothecene lies in a similar position to that observed for kaempherol where this is an acceptor substrate for the plant flavonoid glucosyltranferase *Vv*GT1 (*42*).

Chapter 3:

Determinants and Expansion of Specificity in a Trichothecene UDP-glucosyltransferase from Oryza sativa

X-ray coordinates for the Os79·UDP deoxynivalenol-3-*O*-glucoside, Q202A·UDP, H122A/L123A·UDP, and T291V·UDP complexes have been deposited with the Research Collaboratory for Structural Bioinformatics as Protein Data Bank entries 6BK3, 6BK0, 6BK2, and 6BK1.

This work was performed in collaboration with Kaitlyn Gabardi, Dr. Herbert Michlmayr, Dr. Alexandra Malachova, Dr. Mark Busman, Dr. Susan McCormick, Dr. Franz Berthiller, Dr. Gerhard Adam, and Dr. Ivan Rayment. Dr. Herbert Michlmayr and Dr. Alexandra Malachova produced and purified the D3G. Dr. Susan McCormick produced and purified all other trichothecenes. Dr. Mark Busman and Dr. Susan McCormick analyzed trichothecene products by LCMS. All cloning, protein purification, and X-ray crystallography were performed by Karl Wetterhorn with help from undergraduate research assistant Kaitlyn Girardi. All structural determination, structural refinement, and kinetic analysis were performed by Karl Wetterhorn.

A version of this work was previously published as:
Wetterhorn KM, Gabardi K, Michlmayr H, Malachova A, Busman M, McCormick SP, Berthiller F, Adam G, Rayment I. *Determinants and Expansion of Specificity in a Trichothecene UDP-glucosyltransferase from* Oryza sativa. Biochemistry. 2017 Nov 30;56(50):6585-96.

Abbreviations

4-ANIV, 4-acetyl nivalenol, FUS-X or fusarenon-X,

4,15-diANIV, 4,15-diacetyl nivalenol

D3G, deoxynivalenol-3-*O*-glucoside

DAS, diacetoxyscirpenol

DON, deoxynivalenol

FHB, *Fusarium* head blight

GT-1, family 1 of the glucosyltransferase superfamily

HEPES, 4-(2-hydroxyethyl)-1-piperazineethanesulfonic acid

HT-2, hydrolyzed T-2 toxin

MBP, maltose binding protein

MEPEG 5K, methyl ether polyethylene glycol 5000

MPD, 2-methyl-2,4-pentanediol

NiNTA, nickel-nitrilotriacetic acid

NIV, nivalenol

Os79*,* the product of the Os04g0206600 gene (original seq. NM_00105879)

rmsd, root mean square deviation

T-2, T-2 toxin

TCEP, tris(2-carboxyethyl)phosphine

TEV, Tobacco etch virus protease

TRI, trichothecene

U2F, UDP-2-fluoro-2-deoxy-D-glucose

UGTs, UDP-glycosyltransferases

Abstract

Family 1 UDP-glycosyltransferases in plants (UGTs) primarily form glucose conjugates of small molecules and, besides other functions, play a role in detoxification of xenobiotics. Indeed, overexpression of a barley UGT in wheat has been shown to control *Fusarium* head blight, where this is a plant disease of global significance that leads to reduced crop yields and contamination with trichothecene mycotoxins such as deoxynivalenol (DON), T-2 toxin and many other structural variants. The UGT Os79 from rice has emerged as a promising candidate for inactivation of mycotoxins on account of its ability to glycosylate DON, nivalenol and hydrolyzed T-2 toxin (HT-2). However, Os79 is unable to modify T-2 toxin (T-2), produced by pathogens such as *F. sporotrichioide*s and *F. langsethii*. Activity towards T-2 is desirable because it would allow a single UGT to inactivate co-occurring mycotoxins. Here, the structure of Os79 in complex with the products, UDP and deoxynivalenol-3-*O*-glucoside is reported together with a kinetic analysis of a broad range of trichothecene mycotoxins. Residues associated with the trichothecene binding pocket were examined by site-directed mutagenesis which revealed that trichothecenes substituted on the C4 position, which are not glycosylated by WT Os79, can be accommodated in the binding pocket by increasing its volume. The triple mutation H122A/L123A/Q202L, which increases the volume of the active site and attenuates polar contacts, led to strong and equivalent activity towards trichothecenes with C4 acetyl groups. This mutant enzyme provides the broad specificity required to control multiple toxins produced by different *Fusarium* species and chemotypes.

Introduction

Family 1 UDP-glycosyltransferases (UGTs) of plants are a large class of enzymes that modify their low molecular weight substrates through the addition of a sugar moiety from activated co-substrates such as UDP-glucose (*1, 2*). Typically, plant genomes contain a vast number of UGTs that can be recognized by the characteristic PSPG motif, a 44 amino acid sequence associated with the UDP-sugar binding site. For example, 107 UGT genes have been identified in the genome of *Arabidopsis* (*1*) and 180 (in rice) or more genes are present in diploid crop plants (*3*). The range of substrates to which the carbohydrates are transferred by plant UGTs is enormous and includes endogenous flavonoids, terpenoids, phenolics, saponins, sterols alkaloids and plant hormones, where the modifications serve a wide range of biological functions from defense mechanisms to regulation and changes in bioactivity (*4*).

Structural studies have revealed that family 1 plant UGTs adopt the GT-B protein fold (*5, 6*). This consists of an *N*-terminal and *C*-terminal domain that each exhibit a Rossmann fold. The *C*-terminal domain forms the binding site for the UDP-sugar (donor) whereas the *N*-terminal domain contributes mostly to the target substrate (acceptor) binding site. The active site lies in a groove between the *N*-terminal acceptor and the *C*-terminal donor-binding domains. Because of the structural constraints demanded by binding similar nucleotide sugar ligands, there is considerably more sequence similarity between family members in the *C*-terminal domain than the *N*-terminal domain.

Many plant UGTs demonstrate broad substrate specificity even as they maintain more stringent regioselectivity (*7-10*). This attribute is almost a prerequisite for plant glycosyltransferases, since there are many more glycosylated secondary metabolites observed than there are glycosyltransferases encoded in their genomes (*11, 12*). Broad specificity is also a

desirable feature in the application of plant UGTs to medicinal biochemistry when these are capable of generating glycosylated drug products or those which are able to glycosylate xenobiotics. Thus, understanding substrate specificity in plant UGTs in an effort to change or broaden specificity could prove very useful as a means of glycosylating a wider range of desired substrates. However, prediction of biological targets or specificity has proved difficult and in general can only be determined through experimentation (*12, 13*). Engineering of the acceptor binding pocket of plant UGTs has been shown to increase activity (*14*), and alter regioselectivity (*15, 16*). An exciting potential application for plant UGT engineering is the broadening of specificity of UGTs capable of glycosylating xenobiotics such as trichothecene toxins, which are virulence factors for *Fusarium* head blight (FHB) infection in cereal crops (*17, 18*).

Toxigenic *Fusarium* species cause a significant number of crop diseases worldwide (*19, 20*) and lead to crop loss and contamination with a variety of mycotoxins including the trichothecenes shown in Figure 1. Trichothecene mycotoxins are a highly diverse group of tricyclic sesquiterpenoid epoxides that serve as virulence factors of plant pathogenic *Fusarium* species, but are also toxic to both humans and livestock (*17, 18, 21, 22*). A primary cause of toxicity is the inhibition of protein biosynthesis by targeting the eukaryotic ribosome (*23, 24*). Glycosylation of deoxynivalenol (DON), one of the most common trichothecenes, reduces its bioavailability and toxicity (*25, 26*). As a consequence, UDP-glycosyltransferases have emerged as a promising strategy for controlling FHB in cereal crops (*27-29*). UGTs with specificity towards DON have been identified in *Arabidopsis thaliana*, barley, *Brachypodium distachyon*, rice, and sorghum (*25, 30-33*). Even so, there is a need for enzymes with broader or altered specificity because a wide range of trichothecenes are produced by the various *Fusarium* species that cause diseases on small grain cereals, maize or potato worldwide. In order to generate such

enzymes, a more complete understanding of the structural and enzymatic features that control specificity and activity in UGTs is required. This study focuses on the family 1 plant UGT Os79 (Os04g0206600) from *Oryza sativa* with the goal of understanding the structural determinants for specificity and expanding the number of substrates that it can accept.

The UGT Os79 already exhibits substantial transferase activity towards DON, nivalenol (NIV), isotrichodermol (IsoT), and HT-2 toxin (HT-2), but not to C4 acetylated compounds like T-2 toxin (T-2), DAS, 4-ANIV, or 4,15-diANIV (Figure 1). Recent structural studies of this enzyme with trichothecene bound in the active site provided insight into the general position of the substrate since it defines the approximate location of the acceptor binding site (*34*). However, it also left some uncertainty in how the other substrates bind in the active site since it represents a minimal inhibitor complex. The compound trichothecene is an inactive substrate because it lacks the 3-hydroxyl group which accepts the glucosyl moiety. It is also the most hydrophobic and least substituted member of the trichothecene family of mycotoxins. It is much smaller than other substrates such as HT-2 toxin. As such, it was unclear what structural features account for the wide specificity of Os79, and in particular why HT-2 is a substrate whereas T-2 is not, where the only difference is the absence of the C4 acetyl group in HT-2. To address these questions, we determined the structure of the product complex of Os79 with deoxynivalenol-3-*O*-glucoside (D3G) and UDP together with kinetic measurements on a series of closely related substrates that distinguish the effects of substitutions at the 4, 8, and 15 positions of the trichothecene skeleton (Figure 1). The structural determinants for specificity that became apparent from these observations were tested by site-directed mutagenesis which yielded an enzyme with considerably broader specificity than the wild-type enzyme as this list now includes T-2, 4-ANIV, and DAS. This was accomplished without greatly compromising the activity towards the

initial substrates and attests to the functional plasticity of this active site. In addition, the role of an absolutely conserved Thr/Ser in GT-1 family of glycosyltransferases that coordinates the *β*-phosphate oxygen of UDP was examined to understand its role in substrate orientation in the active site.

Experimental Procedures

Trichothecenes. Diacetoxyscirpenol (DAS), deoxynivalenol (DON), T-2 toxin (T-2), HT-2 toxin (HT-2), isotrichodermol (IsoT), trichothecene, 4-acetyl nivalenol (4-ANIV), 4,15-diacetyl nivalenol (4,15-diANIV), and nivalenol (NIV) used in this study (Figure 1) were obtained from the USDA-ARS, Mycotoxin Prevention and Applied Microbiology Research Unit, Peoria, IL 61604, U.S.A. Deoxynivalenol-3-*O*-glucoside (D3G) was enzymatically produced and HPLC purified as previously described (*33*).

Site-Directed Mutagenesis. The construction of the plasmid containing the Os79 wild-type gene used in this study was described previously (*34*). Briefly, the plasmid is a derivative of pET21 containing His_6-tagged maltose binding protein (MBP) followed by a TEV protease cleavage site (*35*), and the codon optimized Os79 WT gene. Mutants of *Os79* were generated using a single primer, PCR based method based on "QuikChange" mutagenesis (*36*). The following mutagenesis primers were used:

T291V: 5'-GGTTTTGGTGTCATACGGAGTGGTTTCTACTTTTGATGTTGCTAAAC-3',

T291S: 5'-GGTTTTGGTGTCATACGGAAGCGTTTCTACTTTTGATGTTGCTAAAC-3',

Q202A: 5'-TCCTGCCTTCTGTGAGGCGTCTATCGAGCAGTTTGCT-3',

Q202E: 5'-TCCTGCCTTCTGTGAGGAATCTATCGAGCAGTTTGCT-3',

Q202L: 5'-TCCTGCCTTCTGTGAGCTGTCTATCGAGCAGTTTGCT-3',

F199Q: 5'-CCTGAATTAACTCCTGCCCAATGTGAGCAATCTATCGAGC-3',

A384S: 5’-GCTATGCCTCACTGGTCTGATCAACCTACTATTAGCAAGTATG-3’,

Q143A: 5’-CAGCCGCATTTCTAAGTGCGCCATGTGCTGTGGAC-3’,

H122A/L123A:

5’-AGAGTTTTAGTCTACGATCCAGCGGCGCCATGGGCTAGAAGAGT-3’,

H122A/L123G:

5’-AGAGTTTTAGTCTACGATCCAGCGGGCCCATGGGCTAGAAGAGT-3’,

S203L: 5’-GCCTTCTGTGAGCAATTGATCGAGCAGTTTGCTG-3’,

S203A: 5’-GCCTTCTGTGAGCAAGCTATCGAGCAGTTTGCTG-3’,

All mutations were verified by DNA sequencing using BigDye® protocols (ABI PRISM). Reaction mixtures were resolved by the UW-Madison Biotechnology Center.

Expression of Os79 mutants. Os79 mutants were overexpressed in *E. coli* strain BL21 Codon Plus (DE3). Cultures from a single colony were used to inoculate 6 L Lysogeny broth (LB) supplemented with 100 µg/mL ampicillin and 30 µg/mL chloramphenicol. Expression of Os79 was induced with 1 mM isopropyl-β-D-thiogalactopyranoside when cultures reached OD_{600} ~0.8. Induction was carried out at 16 °C for 20 h. Cells were harvested by centrifugation at 3,000 x g, washed with buffer containing 10 mM HEPES pH 7.6 and 100 mM NaCl, and flash-frozen in liquid nitrogen. Cells were stored at -80 °C until use.

Purification of Os79. All Os79 mutants were purified in a similar manner. All purification steps were carried out on ice or at 4 °C. Around twenty grams of *Escherichia coli* cells expressing His_6-MBP-Os79 were resuspended in 100 mL buffer containing 20 mM HEPES pH 7.6, 50 mM NaCl, 0.2 mM tris(2-carboxyethyl)phosphine (TCEP), 1 mM PMSF, 50 nM leupeptin (Peptide International), 70 nM E-65 (Peptide International), 2 nM aprotinin (ProSpec), and 2 mM AEBSF (Gold BioTechnology) and 50 mg lysozyme. Cells were lysed with 5 pulses

(45 s) using a Qsonica Q700 sonicator, and the lysate was clarified by centrifugation at 40,000 rpm in a Type 45 Ti Rotor (Beckman-Coulter) for 30 min. The concentration of NaCl and imidazole were raised to 300 mM and 20 mM, respectively, by the addition of 4 M stock solutions and loaded onto a 5 mL nickel-nitrilotriacetic acid (NiNTA; Qiagen) column equilibrated with NTA buffer: 50 mM HEPES, pH 7.6, 300 mM NaCl, 20 mM imidazole, and 0.2 mM TCEP. After loading, the column was washed with 120 mL of NTA buffer. Os79 was eluted with 40 mL NTA buffer containing 300 mM imidazole. His_6-tagged Tobacco etch virus (TEV) protease (*37*), was added at 1:40 molar ratio to cleave the His_6-MBP from Os79. The mixture was dialyzed overnight in a buffer containing 10 mM HEPES pH 7.6, 50 mM NaCl, 0.5 mM EDTA, 0.2 mM TCEP. The NaCl and imidazole concentrations were brought up to 300 mM and 20 mM, respectively, and the solution was passed over a 5 mL NiNTA column equilibrated with NTA buffer. The flow-through contained Os79, while the column retained TEV protease, His_6-MBP, and undigested His_6-MBP Os79. Purified Os79 was concentrated using a centrifugal filter (Amicon) with a 30,000 nominal molecular weight limit to a final concentration of 10-20 mg/mL, estimated using calculated molar extinction coefficient of 57870 $M^{-1} \cdot cm^{-1}$ at 280 nm. Proteins were dialyzed against a storage buffer containing 10 mM HEPES pH 7.6, 0.2 mM TCEP, drop-frozen in 30 µL aliquots in liquid nitrogen, and stored at -80 °C.

Crystallization of Os79·D3G·UDP. Os79 was screened for initial crystallization conditions with a 144-condition sparse-matrix screen developed in the Rayment laboratory. Single, diffraction quality crystals were grown at 4 °C by hanging drop vapor diffusion by mixing 2 µL of Os79 at 12 mg/mL in 10 mM HEPES pH 7.6, 50 mM NaCl, 30 mM D3G, 5 mM UDP with 2 µL well solution containing 100 mM sodium acetate pH 5.0, 40% pentaerythritol propoxylate 426, 320 mM NaCl (*38*). Hanging droplets were nucleated after 24 h from an earlier

spontaneous crystallization event using a cat's whisker. Crystals grew to approximate dimensions of 100 x 100 x 300 μm within 15 days. The crystals were transferred to a cryoprotecting solution that contained 100 mM sodium acetate pH 5.0, 40% pentaerythritol propoxylate 426, 320 mM NaCl, 30 mM D3G, 5 mM UDP and vitrified by rapid-plunging into liquid nitrogen. Os79 crystallized in the space group $P3_221$ with unit cell dimensions of a = 104.5 Å, b = 104.5 Å, c = 98.3 Å and one chain in the asymmetric unit.

Crystallization of Os79 Q202A·UDP. Os79 Q202A was screened for initial crystallization conditions as described above. Single, diffraction quality crystals were grown at 4 °C by hanging drop vapor diffusion by mixing 2 μL of 11 mg/mL Os79 Q202A in 10 mM HEPES pH 7.6, 5 mM UDP-glucose with 2 μL well solution containing 100 mM 3-(*N*-Morpholino)propanesulfonic acid pH 7.5, 15% methyl ether polyethylene glycol 5000 (MEPEG 5K), 160 mM potassium glutamate. Hanging droplets were nucleated after 24 h from an earlier spontaneous crystallization event using a cat's whisker. Crystals grew to approximate dimensions of 100 x 100 x 300 μm within 8 days. The crystals were transferred stepwise to a cryoprotecting solution that contained 100 mM 3-(*N*-Morpholino)propanesulfonic acid pH 7.5, 15% MEPEG 5K, 15% glycerol, 5 mM UDP-glucose and vitrified by rapid plunging into liquid nitrogen. Os79 Q202A crystallized in the space group $P2_12_12_1$ with unit cell dimensions of a = 59.4 Å, b = 83.2 Å, c = 99.1 Å and one chain in the asymmetric unit.

Crystallization of Os79 T291V·UDP. The Os79 protein complex **T291V·UDP** was screened for initial crystallization conditions as described above. Single, diffraction quality crystals were grown at 23 °C by hanging drop vapor diffusion by mixing 2 μL of 10 mg/mL T291V Os79 in 10 mM HEPES pH 7.6, 50 mM NaCl, 5 mM UDP-glucose with 2 μL well solution containing 50 mM HEPES pH 7.5, 32% MEPEG 5K. Hanging droplets were nucleated

after 24 h from an earlier spontaneous crystallization event using a cat's whisker. Crystals grew to approximate dimensions of 75 x 75 x 300 μm within 11 days. The crystals were transferred to a cryoprotecting solution that contained 50 mM HEPES 7.5, 32% MEPEG 5K, 5 mM UDP-glucose and vitrified by rapid plunging into liquid nitrogen. Os79 crystallized in the space group $P2_12_12_1$ with unit cell dimensions of a = 59.4 Å, b = 83.2 Å, c = 98.7 Å and one chain in the asymmetric unit.

Crystallization of Os79 H122A/L123A·UDP. Os79 was screened for initial crystallization conditions as described above. Single, diffraction quality crystals were grown at 23 °C by hanging drop vapor diffusion by mixing 2 μL of 11 mg/mL Os79 in 10 mM HEPES pH 7.6, 50 mM NaCl, 5 mM UDP with 2 μL well solution containing 50 mM HEPES pH 7.0, 18% methyl ether polyethylene glycol 2000. Hanging droplets were nucleated after 24 h from an earlier spontaneous crystallization event using a cat's whisker. Crystals grew to approximate dimensions of 75 x 75 x 300 μm within 4 days. The crystals were transferred stepwise to a cryoprotecting solution that contained 50 mM HEPES 7.0, 20% methyl ether polyethylene glycol 2000, 15% glycerol, and 5 mM UDP and vitrified by rapid plunging into liquid nitrogen. Os79 crystallized in the space group $P2_12_12_1$ with unit cell dimensions of a = 59.3 Å, b = 82.9 Å, c = 99.0 Å and one chain in the asymmetric unit.

Data collection and Refinement. X-ray data for the Os79 structures were collected at 100 K on the Structural Biology Center beamline 19BM at the Advanced Photon Source in Argonne, IL. Diffraction data were integrated and scaled with HKL3000 (*39*). Data collection statistics are given in Table 1. The structures were determined by molecular replacement using coordinates from the RCSB (accession number 5TME) as the molecular replacement search

model in the program Phaser (*34, 40*). Final models were generated by alternating cycles of manual building and least-squares refinement using Coot, Phenix and Refmac (*41-43*).

Glucosyltransferase Enzymatic Assays. Steady-state kinetic analyses of Os79 mutants with trichothecene substrates were performed in a coupled continuous enzymatic assay at 23 °C in a 1 cm path length cuvette. Reactions were initiated by the addition of varying volumes of trichothecenes to a master mix containing Os79 (3.0 μg/mL final concentration), 3 units rabbit muscle lactic dehydrogenase and 2 units rabbit muscle pyruvate kinase (Sigma-Aldrich, buffered aqueous glycerol solution), 1.5 mM phosphoenolpyruvate, 100 μM β-NADH, 50 mM KCl, 10 mM $MnCl_2$, and 100 mM glycylglycine pH 8.0 to yield a final volume of 100 μL. Lactic dehydrogenase, pyruvate kinase, phosphoenolpyruvate, β-NADH, Os79, and UDP-glucose were added to a master-mix containing the remaining reaction components prior to the initiation of each reaction. Reaction progress was followed by monitoring the decrease in A_{340} caused by the oxidation of β-NADH. The rates of reaction were determined at various trichothecene concentrations and fit by non-linear regression to the Michaelis-Menten equation using GraphPad Prism software. In several cases, LC-HR-MS and LC-HR-MS/MS (Agilent 6550 iFunnel QTOF) was used to confirm the identity of the glycosylated product. In these cases, the assay that was used was identical to the coupled assay described above with the exception that none of the components for the coupling system were added. The reaction mixtures were filtered in centrifugal filters (Amicon) with a 30,000 nominal molecular weight limit, lyophilized, and analyzed by LC-MS/MS.

Results and Discussion

The goal of this study was to establish the structural determinants that define acceptor specificity for the wild-type (WT) trichothecene UDP-glucosyltransferase from *Oryza sativa*,

Os79, and test the resultant hypotheses through a combination of structural and kinetic studies on mutant proteins (Tables 1 and 2). As described, this study shows that the wild-type enzyme intrinsically exhibits broad specificity, but surprisingly this can be expanded to include a wider range of substrates without great loss of catalytic efficiency. As part of this study the role of a serine/threonine in the binding site, that is completely conserved in family 1 plant UGTs, was also examined to gain a greater understanding of those factors that influence enzyme activity in the glucosyltransferase superfamily.

Role of the conserved Thr 291 in Os79. The previously proposed catalytic mechanism for Os79 requires that His 27 deprotonates the 3-hydroxyl of the trichothecene acceptor as the oxygen attacks the anomeric C1′ atom of glucose resulting in the formation of the glycosylated trichothecene and UDP (Scheme 1) (*34*). Three residues were reported to be critical for Os79 activity, including the proposed catalytic dyad of His 27-Asp 120 and Thr 291 (*34*). As was shown previously, substitution of Thr 291 with a valine abolished all transferase activity (*34*). It was hypothesized that Thr 291 was either important for positioning the β-phosphate or acted as a catalytic acid to protonate the phosphate for UDP release, or perhaps a combination of both. The potential structural role of Thr 291 is an important factor when considering Os79 specificity and acceptor binding. If Thr 291 plays a structural role in positioning the β-phosphate, this would restrict any potential movement of the UDP-glucose and provide a stringent constraint on the position of the 3-hydroxyl. This constraint would help to narrow the potential positions of the trichothecene acceptor. To address the potential structural role of Thr 291, the structure of Os79 T291V in complex with UDP was determined to 1.61 Å resolution (PDB: 6BK1) (Table 1).

As seen in Figure 2A, when the hydroxyl of Thr 291 is replaced with a methyl group in Os79 T291V the phosphate of UDP adopts a substantially different position in the active site

compared to that seen in WT Os79. In the WT Os79 structure there is a hydrogen bond between the threonine hydroxyl and a phosphate oxygen atom of length about 2.5 Å. This is clearly lost when threonine is replaced by a valine. The change in the position of UDP is likely a direct result of the substitution since crystals of Os79 T291V and WT protein structures were grown at the same pH and under very similar crystallization conditions.

If this interaction is important, then it should be present in all plant GT-1 glycosyltransferases. Indeed, this appears to be true. Comparison of the structures of six plant UGT structures that have been determined with UDP bound in their active sites reveals that the presence of a hydroxyl moiety within hydrogen bonding distance of a β-phosphate oxygen is a conserved characteristic in all of these enzymes (Figure 2B). The proteins were aligned by superimposing an α-helix and two beta strands that surround the UDP binding site, because the overall architecture of these proteins varies substantially at the periphery of the protein, away from the UDP binding site. These three secondary structural elements that underlie the framework for the UDP binding site are very similar across the six proteins and align with an average rmsd (root mean square deviation) of 0.30 Å for the alpha carbons. In every case, there is a threonine or serine hydroxyl positioned within 3.0 Å of the phosphate oxygen. The hydroxyl containing residues in each of the six structures are as follows: 2VG8=S277, 2VCH=S277, 3HBF=S282, 3HBJ=S282, 2C1X=T280, 2C9Z=T280, Os79=T291 (*44-46*). As seen in Figure 2B, in every structure the phosphate adopts almost exactly the same orientation. This provides further evidence that the change in the orientation of the phosphate of UDP in the T291V structure is caused by the lack of a hydroxyl group and the inability to form a hydrogen bond. This serine or threonine residue has not been identified as a catalytically critical residue in any publication about plant UGTs to date. Here, we demonstrate that it is critical for orienting the

phosphate in the active site. While this does not preclude it from participating as a catalytic acid, it does confirm a structural role for this side chain. This is likely a broad finding that applies to many, if not all, plant UGTs. With the question regarding the role of Thr 291 in the catalytic mechanism of Os79 resolved, understanding the substrate specificity for Os79, and how it might be broadened, was the next goal of this study.

Steady-state kinetic parameters for WT Os79. Many plant UGTs demonstrate broad substrate specificity, a useful characteristic when considering enzymes that glycosylate xenobiotics. It was previously reported that Os79 exhibits broad specificity and can glycosylate DON, HT-2, IsoT, and NIV but not T-2 (*34*). T-2 is commonly produced by *Fusarium* species in Europe, Asia, Africa, and Australia (*47-51*). Given the importance of detoxifying T-2, expanding the specificity of Os79 to include this trichothecene and others would be of great benefit. Given that the only difference between T-2 and HT-2 toxin is the C4 acetyl group that is present on T-2 in place of the hydroxyl on HT-2 it was hypothesized that this acetyl group prevented T-2 from binding in the active site of Os79. It can then be predicted that 4-acetyl nivalenol (4-ANIV), commonly known as FUS-X, will not be a substrate for Os79. 4-ANIV is acetylated at the C4 position but lacks the C15 acetyl and C8 isovaleryl groups of T-2 (Figure 1). Indeed, glycosylation could not be detected with our coupled assay with 4-ANIV, 4,15,diANIV or DAS. This confirms that the 4-acetyl group is responsible for precluding T-2, 4-ANIV, 4,15,diANIV and DAS from the acceptor binding pocket. To gain insight into the structural aspects of the acceptor binding pocket that contribute to its inability to accommodate an acetyl group on the C4 position, the structure of Os79 with the glycosylated trichothecene product D3G in the active site was determined.

Structure of Os79 in complex with UDP and D3G. Os79 was crystallized in the presence of the product D3G in an effort to understand the structural components responsible for substrate specificity and to gain insights into the nature of trichothecene binding in the acceptor pocket. The structure of Os79 in complex with UDP and D3G was determined to 2.17 Å resolution (PDB: 6BK3) (Table 1). The overall structure is very similar to the previously solved structure of Os79 with trichothecene (TRI) and UDP-2-fluoro-2-deoxy-D-glucose (U2F) bound in the active site with an rmsd of 0.7 Å for structurally equivalent α-carbons. Electron density corresponding to DON and UDP was observed, however there was no clear electron density corresponding to the glucose moiety of D3G and the moiety was not modeled (Figure 3A). This is likely the result of the flexibility of the glucose moiety.

There are two major conformational changes in Os79 with D3G bound compared to the structure with trichothecene bound in the active site (Figure 3B). A loop from Ser 288 to Val 297 is shifted 9.5 Å away from the active site and the region that extends from Trp 316 to Lys 336, which is composed of a loop and two short α-helices, is shifted 5.9 Å away from the active site. These conformational changes highlight the flexibility of the acceptor binding region of the protein, and may provide insights into the structural basis for UDP release after the donor sugar has been transferred. Aligning the structures with TRI or D3G bound in the active site reveals that the trichothecene skeleton of DON is rotated about 45 degrees compared to the backbone of TRI. This change is likely the result of the presence of glucose on C3 and would suggest that the trichothecene backbone rotates in the acceptor binding pocket after the reaction is completed and before the glycosylated product is released. It is likely that the orientation of trichothecene is more representative than D3G of the positioning of DON in the active site prior to glycosylation. Using the orientation of trichothecene as a reference, His 122, which is only 4.2 Å away from

C4, was identified as a residue that could clash with the C4 acetyl of the substrates that WT Os79 is unable to glycosylate. With the structural information as a guide, the role of His 122 and other residues identified as potential contributors to the specificity of Os79 were investigated by kinetic analysis.

Steady-state kinetic assays on Os79 mutants. The remarkable substrate plasticity of Os79 prompted the question of whether there are critical residues in the trichothecene binding pocket that facilitate this ability. It also raised the question whether it might be possible to expand the substrate specificity to accommodate trichothecenes with large substituents such as acetyl groups at the C4 position. Examination of the structures with trichothecene and D3G suggested seven residues that might influence binding (Figure 4). Three of these seven side chains, Phe 199, Gln 202, and Ser 203, are located on an α-helix that forms the back of the acceptor pocket as it is presented in Figure 4. His 122 and Leu 123 are on a loop located in the bottom right of the binding pocket. Gln 143 and Ala 384 are on two separate loops in the back left side and the upper left side of the binding pocket respectively. Steady-state kinetic constants were determined for eight mutant proteins where these included single changes and combinations thereof (Table 2). The effect of each of these mutations is discussed below.

Role of Gln 202. Gln 202 is located on an α-helix in the back of the acceptor binding pocket as it is presented in Figure 4. In our proposed structural model of DON binding in the acceptor pocket the carboxamide oxygen of Gln 202 is within hydrogen bonding distance (2.3 Å) of the C7 hydroxyl of DON (Figure 4). To examine whether this residue contributes to specificity it was changed to a glutamate, alanine, and leucine. The K_M value of Os79 Q202E is 17.5-fold higher than the K_M of WT with DON as a substrate where this is accompanied with a small increase in k_{cat}. The Q202E substitution maintains a similar residue size at the 202 position

but adds a negative charge. It is difficult to predict the molecular origin for kinetic effect of this electrostatic change without knowing the on-rate or off-rate for the substrate as both of these enter into the definition of K_M, but the Q202E substitution may still contribute a hydrogen bond to the enzyme substrate complex. To test whether there is a hydrogen bond that plays a role in DON binding between this hydroxyl and Gln 202 the kinetic constants for Os79 Q202A were determined. There is no significant difference in the K_M value of Os79 Q202A compared to WT, (Table 2) indicating that either Gln 202 does not play a role in DON binding or that a water molecule can substitute the place of the side chain. Interestingly, the Q202L substitution decreases the K_M for DON by 4.7-fold without a major change in the value of k_{cat} (Table 2), which is in contrast to the increase generated by the Q202E mutation. Together these substitutions emphasize the impact that substitutions at position 202 can have on substrate and product binding. In summary, these mutations suggest that a polar interaction in this position is not important for activity and that an increased charge is detrimental.

Role of Phe 199. Phe 199 is located in the α-helix one helical turn away from Gln 202 and lies at the top of the active site as shown in Figure 4 and makes a substantial contribution to the primarily hydrophobic acceptor binding pocket. The side chain is about 5 Å above the hydrophobic trichothecene backbone. Changing this residue to glutamine maintains a residue with approximately the same volume but with much greater polarity. The introduction of the polar glutamine in the place of the hydrophobic phenylalanine eliminates enzymatic activity as measured in the coupled-continuous enzymatic assay with DON as a substrate. This, along with the fact that this residue is a conserved Phe in all UGTs that have activity towards DON, highlights the importance of this residue as a component of the acceptor binding pocket.

Role of Gln 143. Gln 143 is located on a loop in the lower left side of the acceptor binding pocket as it is seen in Figure 4. The Oε oxygen of Gln 143 is 3.5 Å from the C6 oxygen of the glucose moiety on U2F. Even though this is somewhat on the long-side for a substantial hydrogen bond this side chain appears to play an important role in substrate binding. The Q143A substitution does not demonstrate activity in the coupled-continuous enzymatic assay with DON as a substrate. This observation is consistent with the suggestion by Hiromoto et al. that the structurally equivalent residue in UGT78K6 (Asn 137) is important in donor recognition (*52*).

Role of Ser 203. Ser 203 is located on the bottom of the α-helix in the back of the acceptor binding pocket adjacent to Gln 202. This side chain was changed to alanine to investigate the role of a polar residue in this position. The K_M value of the S203A mutant is similar to that of WT Os79. In order to see if a small residue is important at this position, the S203L substitution was created. There is a10-fold increase in the K_M of the S203L mutant but no change in k_{cat} (Table 2) indicating that a bulkier residue at this position might hinder DON association or disassociation from the active site. The hydroxyl of S203 is 6.7 Å away from the closest carbon of DON (C4) and given that the change to an alanine has little effect it appears unlikely that it interacts directly with the substrate. It is more likely that a small residue is required at position 203 to allow Gln 202 to maintain a productive orientation. The 5-fold decrease in K_M as a result of the Q202L mutation illustrates the importance of that side chain. This supports the possibility that the S203L mutation might prompt a change in the position of Gln 202 that could increase the K_M for DON.

Role of Ala 384. Ala 384 is positioned on a loop in the upper left side of the acceptor binding pocket about 5 Å from C3 of DON (Figure 4). To investigate whether a hydrophobic residue at this position is important for activity this residue was changed to a serine. The K_M

value of the A384S mutant is similar to that of WT Os79 (Table 2). Similar to Ser 203, this appears to be a second sphere residue, where changes in polarity appear to have little influence on activity towards DON.

Structure of Os79 Q202A and Os79 H122A/L123A proteins. The kinetic measurements with a variety of trichothecene accepters discussed previously in this study revealed that the active site of Os79 is unable to accommodate the C4 acetyl group of trichothecene substrates. The preceding mutations indicate the importance of the helix that carries Phe 199, Gln 202, and Ser 203 where this α-helix is opposite and slightly above C4, C15, and C8 of the trichothecene. Inspection of the acceptor binding pocket revealed that His 122 and Leu 123 are opposite but below Phe 199, Gln 202, and Ser 203, with His 122 positioned only 4.2 Å away from C4. It is unlikely that 4.2 Å leaves sufficient space to accommodate an acetyl group. Although Leu 123 is positioned further from C4, its close proximity to C15 and C8 was viewed as a potential issue if the substrate needed to shift slightly in the active site to accommodate the acetyl moiety on C4. Consequently, His 122 and Leu 123 were simultaneously targeted for mutagenesis. Initially four amino acid substitutions provided Os79 the ability to glycosylate T-2 toxin as measured by an endpoint assay and analyzed by LC-MS/MS as described in the methods. These were the single substitutions of Q202L, Q202A, and Q202V, and the double substitution of H122A/L123G. No further work was performed on Q202V because it was intermediate between Q202L and Q202A. The catalytic efficiency and K_M values of Os79 H122A/L123G are 5.08 x 10^2 s^{-1} M^{-1} and 512 μM with T-2 toxin as a substrate and 1.99 10^2 s^{-1} M^{-1} and 2512 μM with DON as a substrate (Table 2). Activity for Os79 Q202 with T-2 toxin as a substrate was not detected with the assay used in this study, possibly due to a very high K_M. Os79 H122A/L123G was largely insoluble. Based on these findings, Os79

H122A/L123A was expressed and purified. Changing Leu 123 to an alanine instead of glycine improved the solubility of the protein. The catalytic efficiency of the H122A/L123A mutant with with T-2 as a substrate is 1.17 x 103, which is a 3.5-fold increase over WT. The K_M value of 926 μM is similar to the H122A/L123G mutant. To improve the K_M, the H122A/L123A substitutions were combined with the Q202A substitution to make Os79 H122A/L123A/Q202A. The K_M of the triple substitution with T-2 as a substrate is 89 μM and the catalytic efficiency is 2.84 x 10^4, these values represent a 10-fold decrease in K_M and a 20-fold increase in catalytic efficiency compared to the H122A/L123A double mutant. Similar kinetic parameters are observed with DAS as a substrate. The K_M of the H122A/L123A/Q202A triple mutant with 4-ANIV as a substrate is 501 μM, which is a 5.5-fold increase compared with T-2 toxin and DAS. Interestingly, the K_M value of the H122A/L123A/Q202A triple mutant with DON as a substrate is 1202 μM, and the catalytic efficiency is 7.4 x 10^2, these values represent a 20-fold increase and 23-fold decrease respectively compared to WT. It is clear that although Os79 H122A/L123A/Q202A is capable of glycosylating a broader range of substrates than WT Os79, the triple substitution comes at the cost of decreasing the catalytic efficiency with DON as a substrate. Given the observation that the Q202L substitution decreases the K_M for WT Os79 with respect to DON as a substrate, Os79 H122A/L123A/Q202L was made in an attempt to produce an enzyme with intermediate K_M values for both DON and T-2 toxin. The K_M values for Os79 H122A/L123A/Q202L are 118, 261, and 88 μM for T-2 toxin, DON, and 4-ANIV respectively. (Table 2).

The mutations all suggest that the volume of the active site of Os79 is an important determinant in broadening specificity by allowing the acceptor binding pocket to accommodate the C4 acetyl group. However, this does not exclude the possibility that the mutations cause a

structural change in the acceptor binding pocket. To address this question, the structures of Os79 H122A/L123A (PDB: 6BK2) and Os79 Q202A (PDB: 6BK0) were determined in the presence of UDP to a resolution of 1.47 Å and 1.29 Å respectively (Table 1). These structures show that the mutations result in very little change in the overall structures of the protein. The 50 residues that line and surround the acceptor binding pocket of the Os79 Q202A and H122A/L123A structures aligned to the corresponding residues of the Os79 WT structure with an rmsd of 0.12 and 0.16 Å respectively for structurally equivalent α-carbons. This indicates that the framework of the acceptor binding pockets of these proteins are almost identical to the WT enzyme. The main difference is the size of the trichothecene binding pocket. As seen in Figure 5, the H122A/L123A and Q202A mutations increase the volume of the active site. Specifically, these substitutions appear to open up the side of the active site that would need accommodate the C4 acetyl group. This is consistent with the hypothesis that the broad specificity is sanctioned by the general hydrophobicity and volume of the acceptor cavity which affords the C4 acetyl group enough space to allow toxins such at 4-ANIV, DAS, and T-2 to bind.

Conclusions

Broad specificity while maintaining a high catalytic efficiency is a puzzle in enzymology. This is particularly true for members of the plant UGTs that adopt the GT-B protein fold and have evolved to glycosylate a wide range of xenobiotics (*7-10*). As shown here, the wild-type trichothecene UDP-glucosyltransferase from rice, Os79, has a broad specificity that can modify substrates that differ in molecular weight by a factor of 1.8 (Isotrichodermol and HT-2 toxin; 250.3 and 424.5 respectively) with catalytic efficiencies over 1 x 10^4 s^{-1} M^{-1}, but even here there are enigmatic observations. Foremost, the wild-type enzyme is unable to glycosylate T-2 toxin and yet it readily modifies HT-2, which compared to T-2 is deacetylated at the C4-position

(Figure 1). Indeed, the wild-type enzyme is unable to accommodate substrates that are substituted at the C4 position.

The three-dimensional structure of the product complex (Os79·UDP·D3G) in combination with the previously determined structure of trichothecene bound to Os79 revealed that the acceptor pocket is mostly hydrophobic and includes only a few residues capable of forming hydrogen bonds. Mutagenesis of these polar residues that might interact with the trichothecene substrate had small effects on k_{cat} and K_M, whereas mutagenesis of Phe 199 to a glutamine eliminated activity. These observations suggested that the hydrophobicity and volume of the active site are primary factors in substrate specificity. Based on this structural knowledge, the volume of the active site was increased by mutagenesis with the consequence that Os79 H122A/L123A showed excellent activity towards T-2 toxin but reduced activity towards DON. Addition of the Q202L substitution created an enzyme that is a compromise which allows essentially equivalent activity towards both DON and T-2 toxin. This demonstrates that rational design of plant UGTs is a promising approach for the glycosylation of new small molecule substrates not only as a means of detoxifying a wider range of xenobiotics such as trichothecenes but also for the glycosylation of small molecules in medicinal biochemistry. The broad specificity of Os79 H122A/L123A/Q202L makes it an attractive gene to be incorporated in transgenic plants that are susceptible to infection of both T-2 toxin and DON producing *Fusarium* species, like maize (*F. sporotrichioides* and *F. graminearum*), oat (*F. langsethiae* and *F. culmorum*), or potatoes (*F. sambucinum* and *F. graminearum*).

Acknowledgements

We gratefully acknowledge Dr. Norma E. C. Duke and Dr. Stephan L. Ginell for assistance during the X-ray data collection at Argonne. We also thank Dr. James Thoden for, among other things, maintenance of our X-ray diffraction facility, Dr. George Reed for useful discussions of kinetic procedures, and Haley Brown and Michael Andreas for help and support in the laboratory.

References

1. Ross, J., Li, Y., Lim, E., and Bowles, D. J. (2001) Higher plant glycosyltransferases, *Genome biology 2*, REVIEWS3004.

2. Wang, X. (2009) Structure, mechanism and engineering of plant natural product glycosyltransferases, *FEBS Lett 583*, 3303-3309.

3. Caputi, L., Malnoy, M., Goremykin, V., Nikiforova, S., and Martens, S. (2012) A genome-wide phylogenetic reconstruction of family 1 UDP-glycosyltransferases revealed the expansion of the family during the adaptation of plants to life on land, *The Plant journal : for cell and molecular biology 69*, 1030-1042.

4. Tiwari, P., Sangwan, R. S., and Sangwan, N. S. (2016) Plant secondary metabolism linked glycosyltransferases: An update on expanding knowledge and scopes, *Biotechnol Adv 34*, 714-739.

5. Lairson, L. L., Henrissat, B., Davies, G. J., and Withers, S. G. (2008) Glycosyltransferases: structures, functions, and mechanisms, *Annu Rev Biochem 77*, 521-555.

6. Bock, K. W. (2016) The UDP-glycosyltransferase (UGT) superfamily expressed in humans, insects and plants: Animal-plant arms-race and co-evolution, *Biochemical pharmacology 99*, 11-17.

7. Vogt, T., Zimmermann, E., Grimm, R., Meyer, M., and Strack, D. (1997) Are the characteristics of betanidin glucosyltransferases from cell-suspension cultures of *Dorotheanthus bellidiformis* indicative of their phylogenetic relationship with flavonoid glucosyltransferases?, *Planta 203*, 349-361.

8. Taguchi, G., Imura, H., Maeda, Y., Kodaira, R., Hayashida, N., Shimosaka, M., and Okazaki, M. (2000) Purification and characterization of UDP-glucose: hydroxycoumarin 7-*O*-glucosyltransferase, with broad substrate specificity from tobacco cultured cells, *Plant Sci 157*, 105-112.

9. Lim, E. K., Doucet, C. J., Li, Y., Elias, L., Worrall, D., Spencer, S. P., Ross, J., and Bowles, D. J. (2002) The activity of *Arabidopsis* glycosyltransferases toward salicylic acid, 4-hydroxybenzoic acid, and other benzoates, *J Biol Chem 277*, 586-592.

10. Kramer, C. M., Prata, R. T., Willits, M. G., De Luca, V., Steffens, J. C., and Graser, G. (2003) Cloning and regiospecificity studies of two flavonoid glucosyltransferases from *Allium cepa*, *Phytochemistry 64*, 1069-1076.

11. Jones, P., and Vogt, T. (2001) Glycosyltransferases in secondary plant metabolism: tranquilizers and stimulant controllers, *Planta 213*, 164-174.

12. Modolo, L. V., Blount, J. W., Achnine, L., Naoumkina, M. A., Wang, X., and Dixon, R. A. (2007) A functional genomics approach to (iso)flavonoid glycosylation in the model legume *Medicago truncatula*, *Plant Mol Biol 64*, 499-518.

13. Osmani, S. A., Bak, S., and Moller, B. L. (2009) Substrate specificity of plant UDP-dependent glycosyltransferases predicted from crystal structures and homology modeling, *Phytochemistry 70*, 325-347.

14. Modolo, L. V., Escamilla-Trevino, L. L., Dixon, R. A., and Wang, X. (2009) Single amino acid mutations of *Medicago* glycosyltransferase UGT85H2 enhance activity and impart reversibility, *FEBS Lett 583*, 2131-2135.

15. Hansen, E. H., Osmani, S. A., Kristensen, C., Moller, B. L., and Hansen, J. (2009) Substrate specificities of family 1 UGTs gained by domain swapping, *Phytochemistry 70*, 473-482.

16. He, X. Z., Wang, X., and Dixon, R. A. (2006) Mutational analysis of the Medicago glycosyltransferase UGT71G1 reveals residues that control regioselectivity for (iso)flavonoid glycosylation, *J Biol Chem 281*, 34441-34447.

17. Bai, G. H., Desjardins, A. E., and Plattner, R. D. (2002) Deoxynivalenol-nonproducing *fusarium graminearum* causes initial infection, but does not cause disease spread in wheat spikes, *Mycopathologia 153*, 91-98.

18. McCormick, S. P., Stanley, A. M., Stover, N. A., and Alexander, N. J. (2011) Trichothecenes: from simple to complex mycotoxins, *Toxins 3*, 802-814.

19. Windels, C. E. (2000) Economic and social impacts of *fusarium* head blight: changing farms and rural communities in the northern great plains, *Phytopathology 90*, 17-21.

20. Starkey, D. E., Ward, T. J., Aoki, T., Gale, L. R., Kistler, H. C., Geiser, D. M., Suga, H., Toth, B., Varga, J., and O'Donnell, K. (2007) Global molecular surveillance reveals novel *Fusarium* head blight species and trichothecene toxin diversity, *Fungal Genet Biol 44*, 1191-1204.

21. Evans, E., Holtom, A.M., and Hanson, J.R. (1973) Biosynthesis of 2-cis-farnesol, *J. Chem. Soc. Chem. Commun. 1973*, 465a.

22. Pestka, J. J. (2010) Deoxynivalenol: mechanisms of action, human exposure, and toxicological relevance, *Arch Toxicol 84*, 663-679.

23. Cundliffe, E., Cannon, M., and Davies, J. (1974) Mechanism of inhibition of eukaryotic protein synthesis by trichothecene fungal toxins, *Proc Natl Acad Sci U S A 71*, 30-34.

24. Garreau de Loubresse, N., Prokhorova, I., Holtkamp, W., Rodnina, M. V., Yusupova, G., and Yusupov, M. (2014) Structural basis for the inhibition of the eukaryotic ribosome, *Nature 513*, 517-522.

25. Poppenberger, B., Berthiller, F., Lucyshyn, D., Sieberer, T., Schuhmacher, R., Krska, R., Kuchler, K., Glossl, J., Luschnig, C., and Adam, G. (2003) Detoxification of the *Fusarium* mycotoxin deoxynivalenol by a UDP-glucosyltransferase from *Arabidopsis thaliana*, *J Biol Chem 278*, 47905-47914.

26. Payros, D., Alassane-Kpembi, I., Pierron, A., Loiseau, N., Pinton, P., and Oswald, I. P. (2016) Toxicology of deoxynivalenol and its acetylated and modified forms, *Arch Toxicol 90*, 2931-2957.

27. Shin, S., Torres-Acosta, J. A., Heinen, S. J., McCormick, S., Lemmens, M., Paris, M. P., Berthiller, F., Adam, G., and Muehlbauer, G. J. (2012) Transgenic *Arabidopsis thaliana* expressing a barley UDP-glucosyltransferase exhibit resistance to the mycotoxin deoxynivalenol, *J Exp Bot 63*, 4731-4740.

28. Li, X., Shin, S., Heinen, S., Dill-Macky, R., Berthiller, F., Nersesian, N., Clemente, T., McCormick, S., and Muehlbauer, G. J. (2015) Transgenic Wheat Expressing a Barley UDP-Glucosyltransferase Detoxifies Deoxynivalenol and Provides High Levels of Resistance to *Fusarium graminearum*, *Mol Plant Microbe Interact 28*, 1237-1246.

29. Li, X., Michlmayr, H., Schweiger, W., Malachova, A., Shin, S., Huang, Y., Dong, Y., Wiesenberger, G., McCormick, S., Lemmens, M., Fruhmann, P., Hametner, C., Berthiller, F., Adam, G., and Muehlbauer, G. J. (2017) A barley UDP-glucosyltransferase inactivates nivalenol and provides *Fusarium* Head Blight resistance in transgenic wheat, *J Exp Bot*.

30. Schweiger, W., Boddu, J., Shin, S., Poppenberger, B., Berthiller, F., Lemmens, M., Muehlbauer, G. J., and Adam, G. (2010) Validation of a candidate deoxynivalenol-inactivating UDP-glucosyltransferase from barley by heterologous expression in yeast, *Mol Plant Microbe Interact 23*, 977-986.

31. Schweiger, W., Pasquet, J. C., Nussbaumer, T., Paris, M. P., Wiesenberger, G., Macadre, C., Ametz, C., Berthiller, F., Lemmens, M., Saindrenan, P., Mewes, H. W., Mayer, K. F., Dufresne, M., and Adam, G. (2013) Functional characterization of two clusters of *Brachypodium distachyon* UDP-glycosyltransferases encoding putative deoxynivalenol detoxification genes, *Mol Plant Microbe Interact 26*, 781-792.

32. Kluger, B., Bueschl, C., Lemmens, M., Berthiller, F., Haubl, G., Jaunecker, G., Adam, G., Krska, R., and Schuhmacher, R. (2013) Stable isotopic labelling-assisted untargeted metabolic profiling reveals novel conjugates of the mycotoxin deoxynivalenol in wheat, *Anal Bioanal Chem 405*, 5031-5036.

33. Michlmayr, H., Malachova, A., Varga, E., Kleinova, J., Lemmens, M., Newmister, S., Rayment, I., Berthiller, F., and Adam, G. (2015) Biochemical Characterization of a Recombinant UDP-glucosyltransferase from Rice and Enzymatic Production of Deoxynivalenol-3-*O*-beta-D-glucoside, *Toxins 7*, 2685-2700.

34. Wetterhorn, K. M., Newmister, S. A., Caniza, R. K., Busman, M., McCormick, S. P., Berthiller, F., Adam, G., and Rayment, I. (2016) Crystal Structure of Os79 (Os04g0206600) from *Oryza sativa*: A UDP-glucosyltransferase Involved in the Detoxification of Deoxynivalenol, *Biochemistry-Us 55*, 6175-6186.

35. Rocco, C. J., Dennison, K. L., Klenchin, V. A., Rayment, I., and Escalante-Semerena, J. C. (2008) Construction and use of new cloning vectors for the rapid isolation of recombinant proteins from *Escherichia coli*, *Plasmid 59*, 231-237.

36. Chen, G. J., Qiu, N., Karrer, C., Caspers, P., and Page, M. G. (2000) Restriction site-free insertion of PCR products directionally into vectors, *BioTechniques 28*, 498-500, 504-495.

37. Blommel, P. G., and Fox, B. G. (2007) A combined approach to improving large-scale production of tobacco etch virus protease, *Protein Expr Purif 55*, 53-68.

38. Gulick, A. M., Horswill, A. R., Thoden, J. B., Escalante-Semerena, J. C., and Rayment, I. (2002) Pentaerythritol propoxylate: a new crystallization agent and cryoprotectant induces crystal growth of 2-methylcitrate dehydratase, *Acta Crystallogr. D Biol. Crystallogr. 58*, 306-309.

39. Minor, W., Cymborowski, M., Otwinowski, Z., and Chruszcz, M. (2006) HKL-3000: the integration of data reduction and structure solution--from diffraction images to an initial model in minutes, *Acta Crystallogr D Biol Crystallogr 62*, 859-866.

40. 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 Appl Crystallogr 40*, 658-674.

41. Murshudov, G. N., Vagin, A. A., and Dodson, E. J. (1997) Refinement of macromolecular structures by the maximum-likelihood method, *Acta Crystallogr D Biol Crystallogr 53*, 240-255.

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

43. Adams, P. D., Afonine, P. V., Bunkoczi, G., Chen, V. B., Davis, I. W., Echols, N., Headd, J. J., Hung, L. W., Kapral, G. J., Grosse-Kunstleve, R. W., McCoy, A. J., Moriarty, N. W., Oeffner, R., Read, R. J., Richardson, D. C., Richardson, J. S., Terwilliger, T. C., and Zwart, P. H. (2010) PHENIX: a comprehensive Python-based system for macromolecular structure solution, *Acta Crystallogr. D Biol. Crystallogr. 66*, 213-221.

44. Brazier-Hicks, M., Offen, W. A., Gershater, M. C., Revett, T. J., Lim, E. K., Bowles, D. J., Davies, G. J., and Edwards, R. (2007) Characterization and engineering of the bifunctional *N*- and *O*-glucosyltransferase involved in xenobiotic metabolism in plants, *Proc Natl Acad Sci U S A 104*, 20238-20243.

45. Modolo, L. V., Li, L., Pan, H., Blount, J. W., Dixon, R. A., and Wang, X. (2009) Crystal structures of glycosyltransferase UGT78G1 reveal the molecular basis for glycosylation and deglycosylation of (iso)flavonoids, *J Mol Biol 392*, 1292-1302.

46. Offen, W., Martinez-Fleites, C., Yang, M., Kiat-Lim, E., Davis, B. G., Tarling, C. A., Ford, C. M., Bowles, D. J., and Davies, G. J. (2006) Structure of a flavonoid glucosyltransferase reveals the basis for plant natural product modification, *EMBO J 25*, 1396-1405.

47. Marasas, W. F., Yagen, B., Sydenham, E., Combrinck, S., and Thiel, P. G. (1987) Comparative yields of T-2 toxin and related trichothecenes from five toxicologically important strains of *Fusarium sporotrichioides*, *Appl Environ Microbiol 53*, 693-696.

48. Rocha, L. O., Laurence, M. H., Proctor, R. H., McCormick, S. P., Summerell, B. A., and Liew, E. C. (2015) Variation in type A trichothecene production and trichothecene biosynthetic genes in *Fusarium goolgardi* from natural ecosystems of Australia, *Toxins 7*, 4577-4594.

49. Rabie, C. J., Sydenham, E. W., Thiel, P. G., Lubben, A., and Marasas, W. F. (1986) T-2 toxin production by *Fusarium acuminatum* isolated from oats and barley, *Appl Environ Microbiol 52*, 594-596.

50. Yli-Mattila, T., Ward, T. J., O'Donnell, K., Proctor, R. H., Burkin, A. A., Kononenko, G. P., Gavrilova, O. P., Aoki, T., McCormick, S. P., and Gagkaeva, T. Y. (2011) *Fusarium sibiricum* sp. nov, a novel type A trichothecene-producing *Fusarium* from northern Asia closely related to *F. sporotrichioides* and *F. langsethiae*, *Int J Food Microbiol 147*, 58-68.

51. Imathiu, S. M., Edwards, S. G., Ray, R. V., and Back, M. A. (2013) *Fusarium langsethiae* - a HT-2 and T-2 Toxins Producer that Needs More Attention, *J Phytopathol 161*, 1-10.

52. Hiromoto, T., Honjo, E., Noda, N., Tamada, T., Kazuma, K., Suzuki, M., Blaber, M., and Kuroki, R. (2015) Structural basis for acceptor-substrate recognition of UDP-glucose: anthocyanidin 3-*O*-glucosyltransferase from *Clitoria ternatea*, *Protein Sci 24*, 395-407.

53. DeLano, W. L. (2002) The PyMOL Molecular Graphics System.

Tables

Table 1. Data Collection and Refinement Statistics

Protein	Os79 Q202A UDP	Os79 H122A/L123A UDP	Os79 T291V UDP	Os79 D3G·UDP
PDB entry	6BK0	6BK2	6BK1	6BK3
space group	$P2_12_12_1$	$P2_12_12_1$	$P2_12_12_1$	$P3_221$
Unit cell dimensions, Å	a = 59.4 b = 83.2 c = 99.1	a = 59.3 b = 82.9 c = 99.0	a = 59.4 b = 83.2 c = 98.7	a = 104.5 b = 104.5 c = 98.3
Data collection site	Argonne 19ID	Argonne 19ID	Argonne 19BM	Argonne 19BM
Wavelength (Å)	0.979	0.979	0.979	0.979
resolution range (Å)	50–1.47 (1.5-1.47)[a]	50–1.29 (1.31-1.29)	50-1.58 (1.61-1.58)	50–2.17 (2.21-2.17)
reflections:	1095493	1504096	568452	382690
reflections: unique	79234	115974	63979	33082
redundancy	13.1 (12.6)	12.3 (8.5)	8.4 (6.2)	11.6 (8.6)
completeness (%)	99.2 (84.1)	99.2 (96.9)	99.8 (95.2)	100 (100)
average I/σ	41.3 (9.2)	34.9 (4.0)	83.6 (11)	29.1 (4.4)
R_{merge} (%)[b]	5.1 (23.1)	10.2 (0)	5.2 (18.0)	8.6 (55.5)
R_{work}	15.1	16.0	16.0	19.6
R_{free}	17.9	18.0	19.2	21.0
protein atoms	3499	3464	3407	3353
ligand atoms	36	36	36	56
water molecules	608	398	557	197
average B factors	19.3	17.9	27.3	35.1
Ramachandran	--	--	--	--
most favored	96.8	97.2	97.5	96.0
allowed	2.98	2.8	2.5	4.0
disallowed	0.23	0	0.0	0.0
rms deviations	--	--	--	--
bond lengths (Å)	0.022	0.021	0.013	0.023
bond angles (deg)	1.889	1.755	1.573	2.234

[a]Values in parenthesis are for highest resolution shell

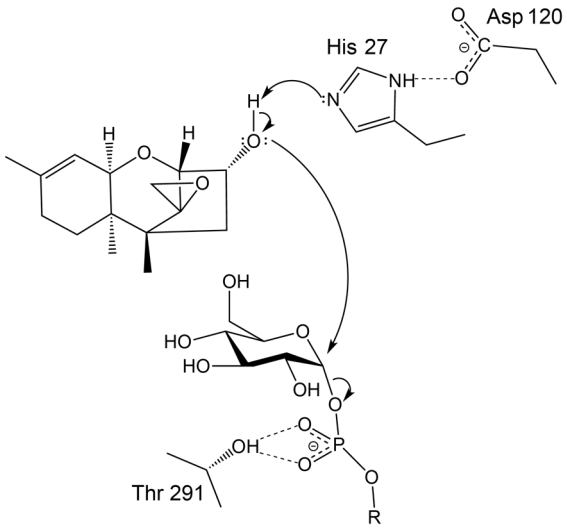
[b]R_{merge}= Σ| $I_{(hkl)}$ - I | x 100 /Σ| $I_{(hkl)}$ |, where the average intensity I is taken over all symmetry equivalent measurements and $I_{(hkl)}$ is the measured intensity for a given observation.

[c]R_{factor}= Σ| $F_{(obs)}$ - $F_{(calc)}$ | x 100 /Σ| $F_{(obs)}$ |, where R_{work} refers to the R_{factor} for the data utilized in the refinement and R_{free} refers to the R_{factor} for 5% of the data that were excluded from the refinement.

Table 2. Steady-state kinetic constants for WT Os79 and selected amino acid substitutions with a variety of trichothecene substrates. Steady-state kinetic constants were determined using a coupled-continuous enzymatic assay.

Os79 Variant	Substrate	k_{cat} (s^{-1})	K_M (μM)	k_{cat}/K_M (s^{-1} M^{-1})
wild-type	DON	1.07 ± 0.04	61 ± 6	1.75×10^4
	T-2 toxin	n/a	n/a	n/a
	HT-2 toxin	0.85 ± 0.02	22 ± 2	3.86×10^4
	DAS	n/a	n/a	n/a
	IsoT	0.79 ± 0.02	< 1.5	5.27×10^5
	NIV	0.38 ± 0.02	35 ± 6	1.09×10^4
	4-ANIV	n/a	n/a	n/a
	4,15,diANIV	n/a	n/a	n/a
A384S	DON	1.6 ± 0.07	78 ± 13	2.05×10^4
F199Q	DON	n/a	n/a	n/a
Q143A	DON	n/a	n/a	n/a
Q202A	DON	0.88 ± 0.04	77 ± 16	1.14×10^4
Q202E	DON	2.4 ± 0.3	1072 ± 257	2.24×10^3
Q202L	DON	0.8 ± 0.01	13 ± 1.7	6.15×10^4
S203L	DON	3.8 ± 0.2	612 ± 67	6.21×10^3
S203A	DON	0.77 ± 0.03	57± 10	1.33×10^4
H122A/L123G	DON	0.50 ± 0.05	2512 ± 361	1.99×10^2
	T-2 toxin	0.26 ± 0.02	512 ± 90	5.08×10^2
H122A/L123A	T-2 toxin	1.58 ± 0.11	926 ± 116	1.71×10^3
H122A/L123A /Q202A	DON	0.89 ± 0.08	1202 ± 220	7.40×10^2
	4-ANIV	1.06 ± 0.04	501 ± 48	2.11×10^3
	T-2 toxin	2.53 ± 0.10	89 ± 11	2.84×10^4
	DAS	1.78 ± 0.03	49 ± 3	3.63×10^4
H122A/L123A /Q202L	DON	0.85 ± 0.01	261 ± 9	3.26×10^3
	4-ANIV	0.96 ± 0.03	88 ± 11	1.09×10^4
	T-2 toxin	0.91 ± 0.02	118 ± 10	7.63×10^3
n/a, no measurable activity				

Schemes



Scheme 1. Proposed catalytic mechanism for the glycosylation of isotrichodermol by Os79. The His 27-Asp 120 catalytic dyad serves as a general base that accepts the proton from the 3-hydroxyl of the trichothecene acceptor as the nucleophilic oxygen attacks the anomeric C1′ of glucose. Thr 291 plays a critical role in positioning the phosphate and possibly protonating it before the release of UDP. It is possible that Thr 291 serves both of these roles.

Figures

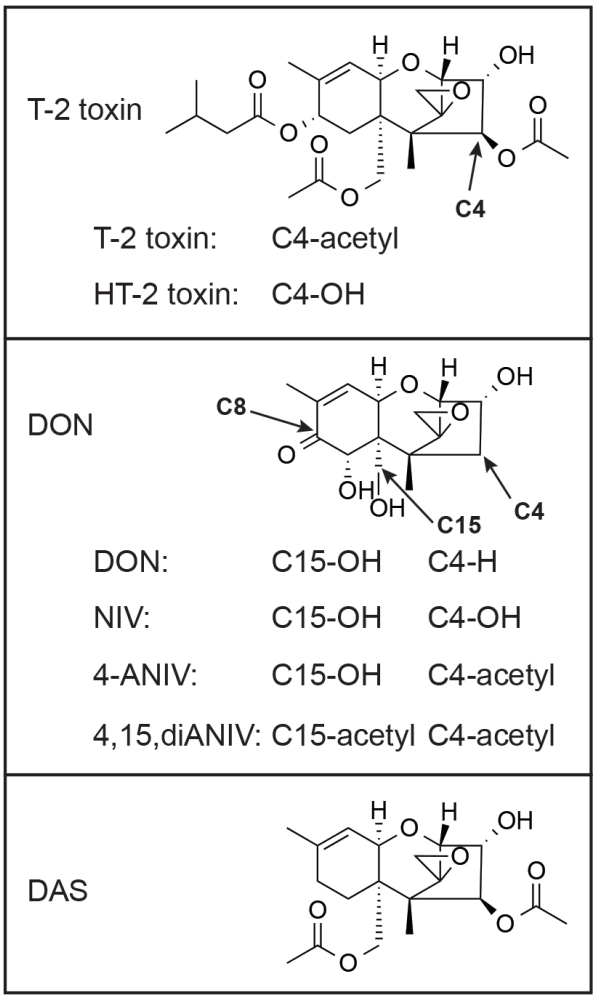


Figure 1. Trichothecene mycotoxins examined in this study. The structures of T-2 toxin, DON, and DAS are shown explicitly. NIV, 4-ANIV, 4,15,diANIV and HT-2 toxin differ from T-2 toxin or DON at the C15 and C4 positions as illustrated.

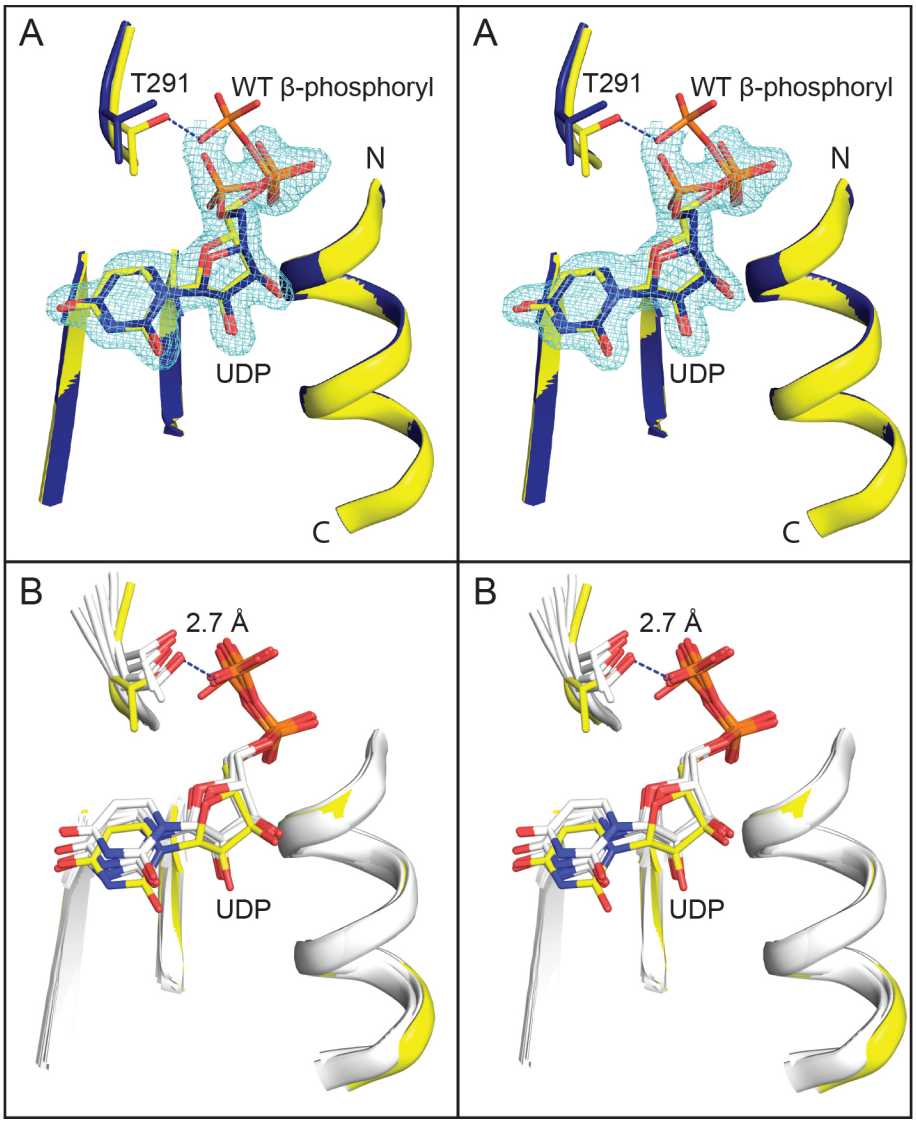


Figure 2. Stereo comparison of UDP binding in six plant UGTs and the role of Thr 291 in positioning the UDP phosphate. (A) The electron density for UDP in the T291V structure (PDB entry 6BK1) and comparison with the WT Os79 structure (PDB entry 5TME). Three segments of Os79 T291V (blue) were aligned to the WT Os79 (yellow) (rmsd = 0.06 Å). In the absence of the hydroxyl of Thr 291 the phosphate moiety adopts a different position in the active site. The density map was calculated with coefficients of the form Fo - Fc where UDP was omitted from the final refinement. The map was contoured at 3.5σ. (B) The same three segments surrounding the UDP binding site (shown in ribbon representation) from six plant UGTs (PDB entries 2C9Z (*46*), 2C1X (*46*), 3HBJ (*16, 45*), 3HBF (*45*), 2VCH (*44*), 2VG8 (*44*)) were aligned to the corresponding sections of wild-type Os79 (yellow) (PDB 5TME) with an average rmsd of 0.30 Å for the structurally equivalent α-carbons. A loop near the terminal phosphate is shown with the conserved serine or threonine residue in stick representation (Thr 291 in Os79). The hydroxyl of this conserved residue is on average 2.7 Å from the oxygen of the β-phosphate. Figures 2-5 were prepared with Pymol (*53*).

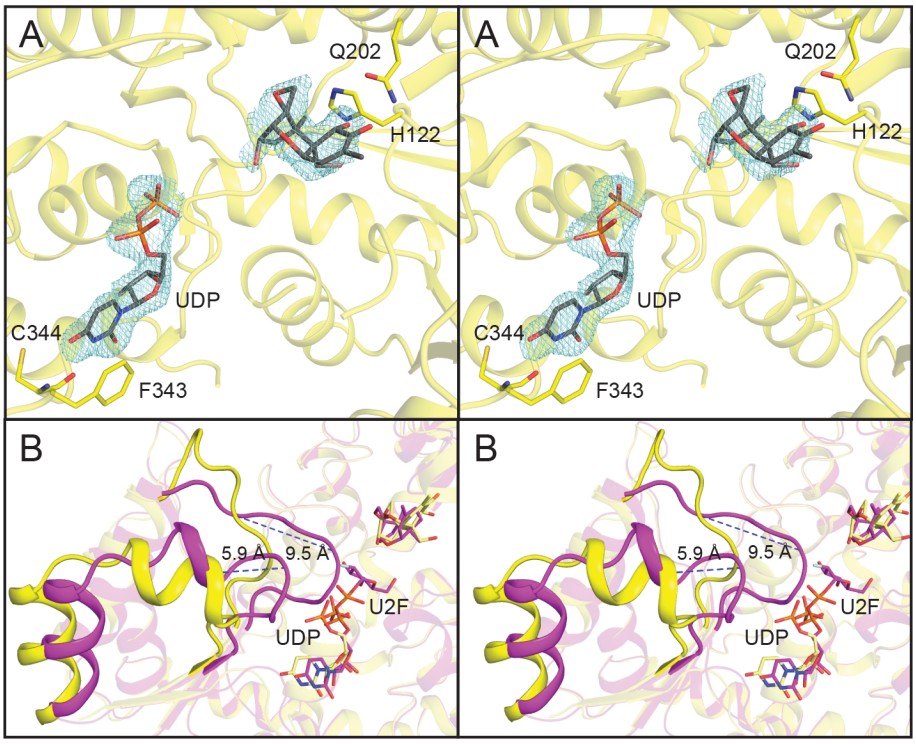


Figure 3. Structure of Os79 with the product D3G and UDP. (A) The active site of Os79 with UDP and D3G bound with DON and UDP shown in stick form (PDB entry 6BK3). The density maps were calculated with coefficients of the form Fo - Fc where the substrates were omitted from the final refinement. The map was contoured to 2.5σ. In both panels of the figure the glucose moiety of D3G was not included in the structure because there was insufficient electron density to determine its orientation. (B) Stereoview of Os79 with UDP and D3G bound (yellow) aligned to Os79 with U2F and TRI bound (magenta) with an rmsd of 0.7 Å for structurally equivalent α-carbons. UDP, U2F, TRI, and DON are shown in stick form and colored by element.

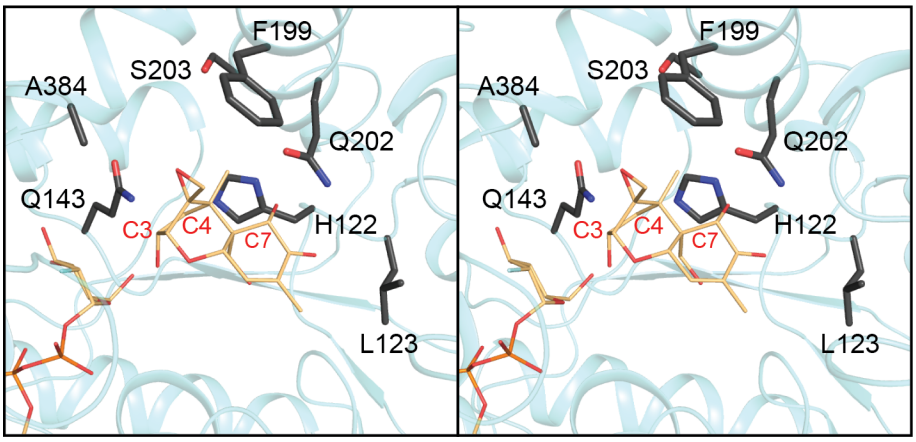


Figure 4. Active-site mutations in Os79. Stereoview of the active site of Os79 with U2F including a model of DON shown in stick form (*34*). DON was modeled into the active site by superimposing the backbone of the toxin onto the trichothecene in the active site in the previously solved Os79 structure (PDB 5TME). The orientation of DON observed in the complex with D3G was not used here because it is believed that the trichothecene backbone of DON rotates after glycosylation. Residues that were mutated in this study are shown in black stick representation and colored by element.

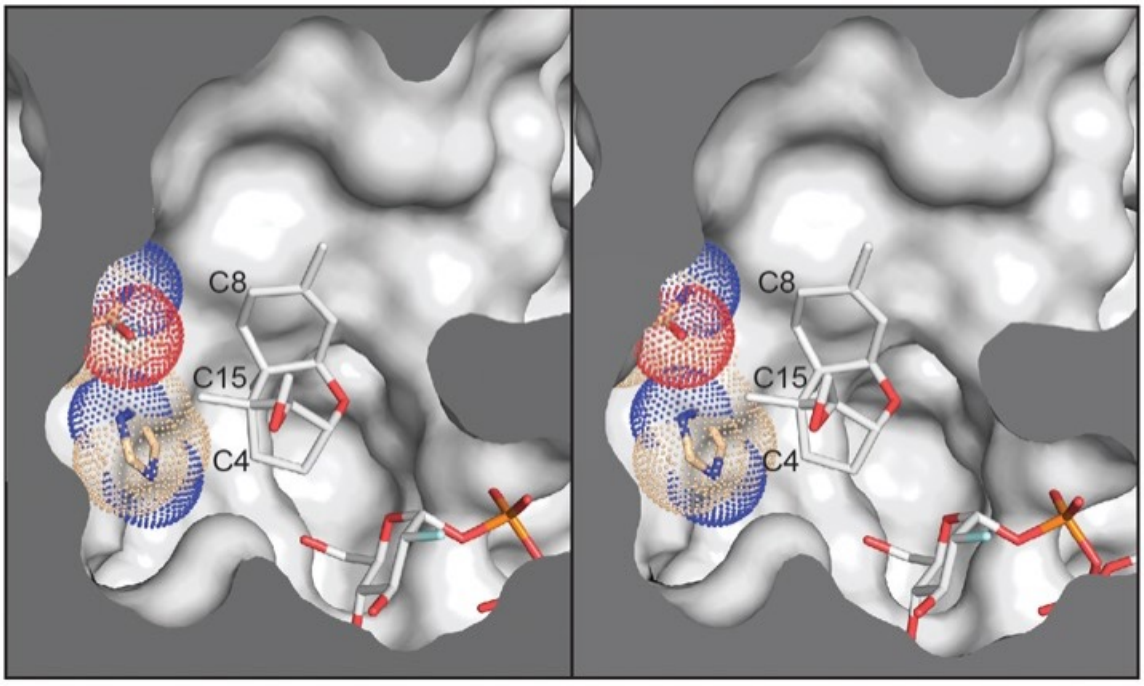


Figure 5. Comparison of Os79 WT and H122A/L123A and Q202A active sites. The structures of Os79 H122A/L123A (PDB entry 6BK2) and Q202A (PDB entry 6BK0) were aligned to WT Os79 with an rmsd of 0.17 and 0.16 Å respectively (PDB entry 5TMD). The cross section of the acceptor binding pocket is shown in a stereoview with the surface of the two variant proteins shown in white. The His 122, Leu 123, and Gln 202 residues from the WT protein are shown in stick form with wheat colored carbon atoms and dots to illustrate their size. Carbons 4, 8, and 15 of the trichothecene backbone are labeled. The 2-deoxy-2-fluoro glucose moiety from the U2F·trichothecene complex is also shown in the active site.

Chapter 4:

Development of a yeast selection protocol to identify trichothecene inactivating enzymes

This work was performed in collaboration with Dr. Susan McCormick and Dr. Mark Busman. Conceptual planning and development of the selection protocol was a collaboration between Dr. Ivan Rayment and Karl Wetterhorn. Trichothecenes were produced and purified by Dr. Susan McCormick and analysis of trichothecene products was performed by Dr. Susan McCormick and Dr. Mark Busman. All other laboratory work was performed by Karl Wetterhorn with the assistance of undergraduate researchers Rachell Caniza and Kaytlyn Gabardi.

Introduction

Fusarium head blight (FHB) is a fungal infection that is a world-wide agricultural problem, effecting cereal crops such as wheat and barley. Fungi that cause FHB relay on trichothecene mycotoxins as virulence factors (*1*). Trichothecenes are potent inhibitors of eukaryotic protein synthesis (*2*) and are toxic to humans and animals (*3*). Because they are toxic, accumulation of trichothecenes in crops poses a major health concern for humans and animals. Consequently, the ability to inactivate trichothecenes in harvested crops is desirable.

Trichothecenes can be inactivated by either substituting the C3 hydroxyl with a moiety such as glucose or acetyl (*4, 5*), or by breaking the epoxide bond (*6*). Inactivation of trichothecenes by glycosylation has led to some success in creating crops with resistance to *Fusarium* infection (*7*). The reversible nature of glycosylation means that it is not an effective method for permanent inactivation of trichothecenes in harvested crops. Because of this, it is highly desirable to discover or engineer an enzyme capable of breaking the essential epoxide bond in trichothecenes.

One approach to engineering an enzyme capable of breaking the epoxide bond of trichothecenes is to create a large and diverse gene library and select for enzyme variants with the desired trichothecene deepoxidation activity. Thanks to powerful molecular genetics tools like saturation mutagenesis, creating large and genetically diverse libraries is relatively easy, making the identification of proteins with the desired activity the primary limitation. As protein engineering has become a powerful tool for design and evolution of enzymes, many methods for identifying enzymes with the desired activities from large mutant libraries have been developed.

Procedures for assessing large gene libraries can be broadly categorized as either screening or selection methods. Screening methods such as fluorescence-activated cell sorting (FACS) and microtiter plate assays rely on the evaluation of every protein in the library (*8, 9*), while selection

methods rely on the elimination of nonfunctional variants by applying selection pressures on the mutant library (*10*). Selection methods can be further categorized as either display methods such as plasmid display and phage display, or compartmentalization methods, which rely on cells (*in vivo* compartmentalization) or man-made compartments such as emulsions droplets (*in vitro* compartmentalization). *In vivo* compartmentalization can be roughly broken down into growth complementation or reporter-based selection.

Growth complementation selection methods couple the desired enzyme activity to the fitness of the cell (*11*), allowing cells that contain enzymes with the desired activity to survive the selective pressure, while eliminating cells that are lacking such enzymes. Growth complementation *in vivo* compartmentalization selection methods are limited to detoxifying enzymes or enzymes that synthesize an essential nutrient for cell survival. It was hypothesized that trichothecene inactivating enzymes would lead to a selective advantage in cells that were exposed to trichothecenes, so growth complementation *in vivo* compartmentalization was chosen as the selection method for identification of trichothecene resistance genes.

One of the simplest growth complementation *in vivo* compartmentalization selection methods is agar plating. This strategy involves plating cells transformed with a mutant library onto agar plates to select for the desired enzyme activity. Cells can be auxotroph where the gene library restores an essential enzyme, or the library can contain a resistance gene which allows the cells to grow on agar plates containing a toxin. Agar plating selection strategies can select for enzymes in gene libraries as large as 1×10^5 mutants (*10*) and have been successfully used to identify enzymes with novel and low activity for almost four decades (*12*). In several studies, plating assays were able to select for enzymes with efficiencies in the low 1×10^2 s^{-1} M^{-1} (*13, 14*).

In this chapter, we describe the development of a yeast selection strategy that relies on the toxicity of the trichothecene isotrichodermol (IsoT) to create a selection advantage for cells containing a trichothecene resistance gene. As with any growth complementation *in vivo* compartmentalization selection strategy, one of the most critical elements was creating an advantage for cells containing an enzyme with the desired activity. For a selection strategy to be effective, cells must be susceptible to the selective pressure (in this case trichothecenes) and must not have the ability to acquire resistance to the selective pressure that is independent of the library enzymes. Throughout the development of the selection strategy presented in this chapter, one of the major challenges was creating a yeast strain which lacked the ability to acquire resistance to trichothecenes independently of resistances enzymes in the library. In order to understand the source of developed resistance to trichothecenes in yeast, an understanding of the mechanism of trichothecene toxicity as it relates to the ribosome is needed.

The yeast ribosome is comprised mainly of rRNA, but also contains 79 proteins (*15*). The largest of these proteins is ribosomal protein L3 (RPL3), which is highly conserved across all three kingdoms (*16*). RPL3 is one of the few ribosomal proteins that is required for peptidyl transferase activity (*17*). RPL3 is a 43.8 kDa protein comprised of a globular domain and two finger-like structures that penetrate into the rRNA core of the ribosome all the way to the A site of the peptidyl transferase center. In fact, the tryptophan at position 255 of one of these finger-like extensions comes the closest to the peptidyl transferase center of any amino acid in the ribosome. Substitutions of this tryptophan cause conformational changes in RPL3 which lead to significant rearrangements of the peptidyl transferase center (*18*). These changes lead to increased affinity to aa-tRNA, decreased affinity for eEF2, and decreased rates of peptidyl transfer (*18*).

Interestingly, the substitutions W255C, W255R, and H256Y in RPL3 all lead to yeast resistance to trichothecenes (*19*). The structural basis for this resistance is not known, even though the structures of the 80S *Saccharomyces cerevisiae* ribosome in complex with the trichothecenes T-2 toxin and deoxynivalenol have been determined to 3.1 and 3.3 Å respectively (*20*). In both of these structures the trichothecenes are bound in the A site of the peptidyl transferase center of the ribosome, but due to the low resolution and poor data quality (R_{meas} of 40 and 31 respectively) specific interaction between the rRNA of the ribosome and the trichothecenes cannot be determined. It is possible that the conformational changes to the peptidyl transferase center resulting from substitutions of tryptophan 255 eliminate the ability of trichothecenes to bind in the peptidyl transferase center, causing the trichothecene resistance phenotype in yeast with the W255C, W255R, and H256Y substitutions.

Regardless of the molecular mechanism that leads to trichothecene resistance in yeasts containing W255C, W255R, and H256Y substitutions in RPL3, it is clear that mutations in *RPL3* contribute significantly to the development of yeast resistance to trichothecenes even in the absence of an exogenous resistance gene (*19*). This method of acquired resistance was eliminated in the yeast strain that was developed for the selection protocol described in this study, as described in the methods section.

The yeast strain used in this study was further sensitized to trichothecenes by deleting *AYT1*, a trichothecene acetyltransferase gene, and by starting with the Green Monster yeast strain, which is lacking 17 multi drug transporters (*21*). The utilization of this sensitized yeast strain, coupled with the addition of a stress selection step, lead to the development of a selection protocol capable of selecting for trichothecene resistance genes in libraries where resistance genes were outnumbered 1,000,000:1 with non-resistance genes. The final selection protocol presented in this

study is a powerful tool for the selection of trichothecene resistance genes from large mutant libraries.

Experimental Procedures

Construction of the p426GPD-Tri101 plasmid. In order to construct the p426GPD-*Tri101* plasmid, the *Tri101* gene from *F. graminearum* (AB000874) was PCR amplified from genomic DNA of strain PH-1 (a gift from Dr. Nancy Keller, University of Wisconsin, Madison, WI) using forward primer

5-TAGTTTTAAAACACCAGAACTTAGTTTCGAGCTAGCATGGCTTTCAAGATACAGCTCGACACCCTCGG-3

and reverse primer

5-AGGTCGACGGTATCGATAAGCTTGCTGAGCTTAACCAACGTACTGCGCATACTTGGTCC-3.

The resulting product was introduced into the p426GPD plasmid (*22*) using an enzyme-free "QuikChange" method (*23*). The plasmid sequence was verified by DNA sequencing using BigDye protocols (ABI PRISM). Reaction mixtures were resolved by the University of Wisconsin-Madison Biotechnology Center.

Yeast electrocompetent cell preparation. All centrifugation steps were carried out at 4 °C for 7 min at 3500 *g*. All buffers and water used for washing cells were ice cold, and all steps were performed at 4 °C whenever possible. Four hundred mL of yeast minimal media with the appropriate selection was inoculated at an OD_{600} of ~0.03. The culture was incubated at 30 °C with 200 rpm shaking until the OD_{600} was 1.4 (5-18 hours depending on yeast strain and media used). Cells were harvested by centrifugation and washed twice with ice cold sterile MilliQ water followed by pelleting by centrifugation. The cell pellet was washed with Buffer A (1 M sorbitol, 1 mM $CaCl_2$) and the cells were pelleting by centrifugation. The cell pellet was resuspended in 100 mL Buffer B (0.1 M Lithium Acetate, 10 mM DTT) and the cells were incubated at 30 °C

with 200 rpm shaking for 1 h in a 500 mL sterile baffled flask. Cells were pelleted by centrifugation and washed twice with ice cold Buffer A. The final pellet was resuspended in ice cold 600 μL Buffer A, giving a total volume around 2 mL electrocompetent cells.

Construction of p425GPD-*RPL3*. *RPL3* was amplified from yeast genomic DNA by PCR using the forward primer

5-TTCGACGGATTCTAGAACTAGTGGATCCATGTCTCACAGAAAGTACGAAGCACCA-3

and reverse primer

5-TCGAGGTCGACGGTATCGATAGCTTTTACAAGTCCTTCTTCAAAGTACCCATGAAAGC-3

The resulting product was introduced into the p425GPD (*22*) plasmid using an enzyme-free "QuikChange" method (*23*). The plasmid sequence was verified by DNA sequencing using BigDye protocols (ABI PRISM). Reaction mixtures were resolved by the University of Wisconsin-Madison Biotechnology Center.

Generating RY0521 Green Monster *AYT1*Δ with p425GPD-*RPL3* yeast strain. An *AYT1*Δ yeast strain was generously provided by the Craig Lab at the University of Wisconsin-Madison. Genomic DNA from this yeast strain was isolated as described previously (*24*). The KANMX cassette was PCR amplified from the genomic DNA using the following primers:

Forward primer: 5-ATCAAAAAATACGTGTACCCCCTTGC-3

Reverse primer: 5-CAGGTGGCAGTTTCAATTGTAGTAAA-3

RY0521 Green Monster yeast (*21*) were transformed with the PCR product, recovered in YPD for 1 h at 30 °C, and plated onto YPD plates. The plates were incubated at 30 °C for 2 days and then replica plated onto YPD plates containing 500 μg/mL G418. Colonies that grew on the G418 plate were re-struck on G418 plates. Genomic DNA was isolated from colonies that grew after re-streaking and PCR amplification with the above two primers was used to confirm KANMX

cassette integration into the genomic DNA. PCR amplification with the reverse primer above and the internal sequencing primer 5-ACCTCTAGAAATACAAGTGGAGAGG-3 from *AYT1* was performed to confirm that *AYT1* had been replaced with the KANMX cassette.

Trichothecene Inactivating Enzyme Selection Protocol 1. An overview of this protocol is presented in Figure 1. JRY8012 yeast (BY4741 pdr5Δ::kanMX snq1Δ::kanMX yor1Δ::kanMX) were transformed with p426GPD-*Tri101* as a positive control and empty p426GPD as a negative control by electroporation at 2.6 kV and 25 μF. Cells were recovered in 8 mL 1:1 YPD:1M sorbitol at 30 °C for 1 h. Cells were pelleted by centrifugation (4 °C, 7 min, 3500 *g*) and washed once with 10 mL URA- media. Pellets were resuspended in 5.5 mL URA- media and 150 μL of cells were plated onto each of 72 URA- plates. Plates were incubated at 30 °C for 3 days. All plates were replica plated onto URA- plates containing 10 μM IsoT and incubated at 30 °C. Plates were checked daily for colonies. Plasmids from colonies that grew were isolated by miniprep and JRY8012 yeast were transformed with the purified plasmid. Transformed cells were then plated onto URA- plates, incubated at 30 °C for 3 days, and replica plated onto URA- plates containing 10 μM IsoT and incubated at 30 °C. Plates were checked for colonies daily.

Trichothecene Inactivating Enzyme Selection Protocol 2. Yeast demonstrated an amazing ability to develop resistance to trichothecenes independently of resistance genes in the library being screened. In order to reduce this background, a second iteration of the selection protocol was developed. An overview of this protocol is presented in Figure 1. This protocol was performed as described in Protocol 1 with the exception that instead of JRY8012 yeast, RY0521 Green Monster *AYT1*Δ with p425GPD-*RPL3* yeast were used (RY0521: MAT**alpha** adp1Δ snqΔ ycfΔ pdr15Δ yor1Δ vmr1Δ pdr11Δ nft1Δ bpt1Δ ybt1Δ ynr070wΔ yol075cΔ aus1Δ pdr5Δ pdr10Δ

pdr12Δ can1Δ::GMToolkit-**alpha** (CMVpr-rtTA NATMX4 STE3pr-LEU2) his3Δ leu2Δ0 ura3Δ0 met15Δ0).

Trichothecene Inactivating Enzyme Selection Protocol 3. Protocol 2 successfully eliminating the background, however very few colonies grew on the positive control plates. To address this issue, a liquid selection step was added to enrich for cells with a resistance gene. The updated selection protocol is described here, and an overview of the protocol is presented in Figure 1. RY0521 Green Monster *AYT1*Δ yeast cells with p425GPD-*RPL3* were transformed by electroporation with p426GPD-*Tri101* as a positive control, empty p426GPD as a negative control and any desired library as described in Protocol 1. Cells were immediately added to 250 mL URA- LEU- media in 500 mL sterile baffled flasks and incubated for 1 h at 16 °C with 200 rpm shaking. IsoT was added to the culture to a final concentration of 15 μM and the culture was incubated overnight 16 °C with 200 rpm shaking. Cells were pelleted by centrifugation (4 °C, 7 min, 3500 *g*) and washed once with ice cold sterile MilliQ water. The cell pellet was resuspended in 10 mL of URA- LEU- media and 150 μL of the cell solution was plated onto each of 10 URA-LEU- plates containing 15 μM IsoT. Plates were incubated at 30 °C and checked daily.

Trichothecene Inactivating Enzyme Selection Protocol 4. Selection Protocol 3 was unable to select for trichothecene resistance genes in libraries of 1x 10^4 and 1 x 10^6, so a more effective selection step was needed in the selection protocol. The updated selection protocol is described here, and an overview of the protocol is presented in Figure 2. RY0521 Green Monster *AYT1*Δ yeast cells with p425GPD-*RPL3* were transformed by electroporation with p426GPD-*Tri101* as a positive control, empty p426GPD as a negative control and any desired library as described in Protocol 1. Cells were immediately added to 250 mL URA- LEU- media in 500 mL sterile baffled flasks and incubated overnight at 16 °C with 200 rpm shaking. IsoT was added to

the culture to a final concentration of 15 μM and the culture was incubated for 1 h at 16 °C with 200 rpm shaking. NaCl was added to the culture to a final concentration of 1 M and the culture was incubated for 3 h at 30 °C with 200 rpm shaking. Amphotericin B was added to the culture to a final concentration of 2 μM and the culture was incubated overnight at 30 °C with 200 rpm shaking. Cells were pelleted by centrifugation (4 °C, 7 min, 3500 *g*) and washed once with ice cold sterile MilliQ water. The cell pellet was resuspended in 10 mL of URA- LEU- media and 150 μL of the cell solution was plated onto each of 10 URA-LEU- plates containing 15 μM IsoT. Plates were incubated at 30 °C and checked daily.

Trichothecene Inactivating Enzyme Selection Protocol 5a. In protocol 5, the selection steps were optimized to allow the identification of trichothecene inactivating enzymes with low activity. Protocol 5a is identical to Protocol 4, with the exception that in the liquid selection steps IsoT was added to final concentration of 10 μM instead of 15 μM, as shown in Figure 2.

Trichothecene Inactivating Enzyme Selection Protocol 5b. In protocol 5, the selection steps were optimized to allow the identification of trichothecene inactivating enzymes with low activity. Protocol 5b is identical to Protocol 4, with the exception that in the liquid selection steps NaCl was added to a final concentration of 0.5 M instead of 1 M, as shown in Figure 2.

Results and Discussion

The infection of cereal crops such as wheat and barley by *Fusarium* fungi leads to *Fusarium* head blight and the accumulation of toxic trichothecene mycotoxins in the harvest grain. The permanent enzymatic inactivation of these trichothecenes is an important step in producing grain that is safe for human and animal consumption (*3*). It is possible that enzymes capable of inactivating trichothecene mycotoxins by breaking the essential epoxide bond may be engineered

from existing enzymes by the creation of large mutant libraries. If such an approach is taken, a reliable selection protocol must be developed for the identification of the gene that encodes the resistance enzyme. In an effort to develop a selection protocol capable of identifying enzymes with low efficiency from large gene libraries, a yeast selection protocol was developed using an iterative approach of design, test, and optimize. Throughout the development of the selection protocol, empty p426GPD vector was used as a negative control and p426GPD-*Tri101* was used as a positive control. Tri101 is a trichothecene acetyltransferase which acetylates the C3 position of trichothecenes (*25*). It is believed that this reversible detoxification works in a similar manner to glycosylation, where the addition of the moiety at the C3 position prevents trichothecene binding to the ribosome.

Trichothecene Inactivating Enzyme Selection Protocol 1. In order to identify trichothecene resistance genes from a large gene library, a yeast plating selection protocol was developed. This method was chosen as the initial approach because such protocols have been successful at selecting for enzymes with low activity form large gene libraries in the past (*13*).

Based on the success of reported plating selection strategies, a yeast plating protocol was developed where JRY8012 yeast were transformed with p426GPD-*Tri101* as a positive control and empty p426GPD as a negative control. Yeast were then recovered and plated on minimal media agar plates, incubated, and replica plated on minimal media agar plates containing 15 μM IsoT. An overview of this protocol is presented in Figure 1. The JRY8012 yeast strain was used because it is lacking three multidrug transporters, pdr5, snq1, and yor1, making it more susceptible to toxins, and thus a better strain for selecting resistance genes. The only selection step in this protocol was plating on minimal media agar plates containing 15 μM IsoT. It was determined that plating more than 150 μL of cells per plate could lead to inconsistent exposure to IsoT and dilute

the effect of the toxin, meaning ~70 plates were needed for each selection protocol. After two days of incubation at 30 ºC, thousands of colonies had grown on the plates from both the empty p426GPD negative control and the p426GPD-*Tri101* positive control. The goal of this screen was to identify yeast that contained exogenous resistance gene. The presence of thousands of colonies on the negative control plates indicates that yeast are able to develop resistance to IsoT independently of any exogenous resistance gene. Several steps were taken in order to reduce the ability of yeast to develop resistance to trichothecenes.

Trichothecene Inactivating Enzyme Selection Protocol 2. The RY0521 Green Monster yeast strain (*21*) was selected for the second iteration of the selection protocol because 17 multi drug transporters have been knocked out of this strain as opposed to just three in JRY8012. Lacking 17 multi drug transporters sensitizes the yeast to trichothecenes as their ability to export toxin is greatly reduced.

A literature investigation into the causes of yeast resistance to trichothecenes revealed that the yeast acetyltransferase AYT1 is capable of acetylating some trichothecenes, including IsoT (*26*). Acetylation of trichothecenes reduces their toxicity significantly, much like glycosylation (*26*). In order to eliminate this source of trichothecene resistance, *AYT1* was deleted from the Green Monster RY0521 yeast strain as described in the methods section.

Trichothecene toxicity is mainly the result of inhibition of protein synthesis caused by trichothecenes binding to the eukaryotic ribosome (*20, 27*). Several studies have been published revealing that mutations in *RPL3* lead to trichothecene resistance in yeast (*28, 29*). This resistance is likely the result of decreased trichothecenes binding to ribosomes containing the altered RPL3 protein (*28, 29*). In order to ensure that the majority of the RPL3 present in the cell was wild-type even if resistance mutations occurred in genomic *RPL3*, around 30 copies of *RPL3* were added to

each cell by means of a high copy number plasmid containing *RPL3*. Briefly, *RPL3* was PCR amplified from yeast genomic DNA and inserted into p425GPD as described in the methods section. The Green Monster RY0521 yeast strain was then transformed with p425GPD-*RPL3*. The p425GPD plasmid is a high copy number plasmid, where a typical cell containing the plasmid will have around 30 copies (*22*). With around 30 copies of *RPL3* per cell, even if resistance mutations occur in one or two *RPL3* copies, the vast majority of RPL3 protein in the cell will be wild-type. This means that trichothecenes will be capable of binding to, and inhibiting, the vast majority of ribosomes in the cell.

In order to test whether the newly developed yeast strain demonstrated an increased sensitivity to trichothecenes and decreased ability to develop resistance to trichothecenes, a control selection protocol was performed as described in the methods section and presented in Figure 1. Briefly, RY0521 Green Monster *AYT1*Δ p425GPD-*RPL3* yeast were transformed with p426GPD-*Tri101* as a positive control, and empty p426GPD as negative control. This selection protocol produced very little background, however fewer than one yeast colony per plate were observed on the positive control plates. The low number of colonies growing on the positive control plates highlighted the need for a liquid selection step that would enrich for cells containing a resistance gene.

Trichothecene Inactivating Enzyme Selection Protocol 3. In the third iteration of the yeast selection protocol, a liquid selection step was added to enrich for cells containing a resistance gene as described in the methods section and presented in Figure 1. Briefly, RY0521 Green Monster *AYT1*Δ with p425GPD-*RPL3* were transformed with p426GPD-*Tri101* as a positive control, empty p426GPD as negative control and plasmid mixtures of 1:10,000 and 1:1,000,000 p426GPD-*Tri101* to empty p426GPD plasmid to determine the ability of the selection protocol to

recover a resistance gene from an excess of plasmids without a resistance gene. Cells were grown overnight in the presence of 15 μM IsoT and plated onto each of 10 URA-LEU- agar plates containing 15 μM IsoT. Plates were incubated at 30 °C. After three days, no yeast colonies were visible on the plates from the negative control (Figure 3). As seen in Figure 3, there were ~20 colonies per plate on the positive control, ~5 colonies per plate on the 1:10,000 and no colonies on the 1:1,000,000.

The yeast selection protocol was being designed with the goal of selecting for a resistance gene in a library of ~1 million variants. Protocol 3 was only capable of selecting for a resistance gene from ~10,000 plasmids not containing a resistance gene. In order to increase the ability of the protocol to select for resistance genes from libraries larger than 1 x 10^4, an additional step was needed.

Trichothecene Inactivating Enzyme Selection Protocol 4. In the fourth iteration of the selection protocol, a stress step was added in order to increase the ability of the protocol to select for resistance genes from libraries larger than 1 x 10^4. The mechanism of trichothecene toxicity dictates that an effective selection step will take advantage of the fact that, in the presence of trichothecenes, cells without a resistance gene are unable to synthesize protein.

Exposing yeast to different stresses induces large changes in the gene expression profile of the cell (*30*). Causton *et al.* investigated the changes in the expression profile of yeast when exposed to heat, acid, alkali, H_2O_2, salt, and sorbitol. All six stresses led to an increase in the expression of common stress induced genes, with salt and sorbitol causing the most substantial increase in expression, as seen in Figure 4. Acidity and salt were selected as the two stresses to be tested in the development of the selection protocol. These two stresses were chosen because salt induces one of the greatest increases in expression while acid induces one of the smallest increases

in expression and we wanted to investigate how different changes to the expression profile would affect our selection protocol. We hypothesized that exposing transformed yeast to stress would induce expression of a large number of stress response genes in cells with a resistance gene, while those lacking a resistance gene would be unable to properly adapt to the new environment, making them susceptible to low levels of a fungicide like amphotericin B.

To determine whether stress from salt or acidity would be more efficient at selecting for trichothecene resistance genes, RY0521 Green Monster *AYT1*Δ cells with p425GPD-*RPL3* were transformed with p426GPD-*Tri101* as a positive control and empty p426GPD as negative control. Cells were recovered and grown overnight followed by the addition of IsoT. After incubating for 1 h, NaCl was added to 1 M in one culture and the pH of the second culture was dropped from 6.0 to 4.0 by the addition of succinic acid. After a 3 h incubation, 150 µL of each culture was plated onto a minimal media plate. Amphotericin B was added to each culture to a final concentration of 2 µM and the cultures were incubated overnight. After washing, 150 µL of each culture was plated onto a minimal media plate.

As seen in Figure 5, the cell density in all cultures was the same after addition of stress and incubation for 3 h. This indicates that the stress did not kill any cells, even cells lacking a resistance gene. This result in not unexpected because trichothecenes are fungistatics, not fungicides. In the short term, the inhibition of protein synthesis caused by trichothecenes simply limits culture growth, rather than killing the cells. It was hypothesized that the addition of the stress would make cells lacking a trichothecene resistance gene more susceptible to a fungicide like amphotericin B than those with a trichothecene resistance gene.

As seen in Figure 5, amphotericin B preferentially killed cells that were exposed to a high salt concentration and contained a resistance gene, whereas acid did not make cells without a

resistance gene susceptible to amphotericin B. The change in expression that results from acid treatment is significantly less drastic than the change induced by high salt (Figure 4). This may be why cells without a resistance gene are more susceptible to amphotericin B after exposure to high salt in the presence of IsoT. Based on these results, salt was chosen as the stress step to add to the selection protocol.

To test how well the updated selection protocol could select for resistance genes in an excess of plasmids not containing a resistance gene, a control protocol was conducted as outlined in Figure 2. Briefly, RY0521 Green Monster *AYT1*Δ cells with p425GPD-*RPL3* were transformed with p426GPD-*Tri101* as a positive control, empty p426GPD as negative control, and plasmid mixtures of 1:10,000 and 1:1,000,000 p426GPD-*Tri101* to empty p426GPD plasmid. Cells were recovered and grown overnight, followed by the addition of IsoT to each culture. After a 1 h incubation, NaCl was added to 1 M and the cultures were incubated for 3 h. Amphotericin B was added to each culture to a final concentration of 2 μM and the cultures were incubated overnight. After washing, 150 μL of each culture was plated onto URA-LEU- plates containing 15 μM IsoT.

This selection protocol is capable of selecting for one resistance gene in 1,000,000 plasmids not containing a resistance gene with almost no background (Figure 6). Given that Tri101 is efficient at providing yeast resistance to trichothecenes, the fact that Protocol 4 can recover cells expressing Tri101 does not guarantee that the selection protocol is capable of recovering a less efficient enzyme. This means that the selection protocol could eliminate cells expressing inefficient enzymes that might be present in the library being analyzed. In order to overcome this issue, the IsoT and NaCl concentrations in the liquid selection were optimized.

Trichothecene Inactivating Enzyme Selection Protocol 5a and 5b. The goal of optimizing the IsoT and NaCl concentrations in the liquid selection step was to facilitate the

identification of enzymes with very low efficiency by reducing the concentrations of these two reagents to just above those which lead to high background. Reducing the IsoT concentration would decrease the amount of toxin that a resistance enzyme would have to inactivate to allow the cell to express proteins in response to the NaCl stress. It was hypothesized that decreasing the NaCl concentration would reduce the increase in protein expression needed for a cell to respond adequately to the stress, thus increasing the chances that a cell containing an enzyme with a low level of activity could survive the amphotericin B. In order to optimize these concentrations, 10 and 15 μM IsoT were each tested with 0.25, 0.5 and 1 M NaCl for a total of six combinations.

These selection protocols were performed as described in Protocol 4 of the methods section. After 150 μL of each culture was plated on URA-LEU- agar plates containing 15 μM IsoT, the plates were incubated at 30 ºC and colonies were counted after 2, 3, and 5 days and the average number of colonies per plate for each protocol was recorded. After two days, the Tri101 positive control plates for all six protocols were overgrown with colonies (over 2000). When 10 μM IsoT was used in the selection protocol along with 0.25 or 0.5 M NaCl, there were over 2,000 colonies on each plate after 3 and 5 days respectively (Figure 7). When using 10 μM IsoT and 1 M NaCl, there were 13, 15, and 41 colonies on the plates after 2, 3, and 5 days respectively (Figure 7). When 15 μM IsoT was used in the selection protocol along with 0.25, 0.5 and 1 M NaCl after 5 days there were 185, 46, and 10 colonies on the plates respectively (Figure 8).

The ideal selection protocol is one that challenges the yeast just enough to eliminate most cells not expressing a trichothecene inactivating protein but not enough to kill cells that are expressing very inefficient trichothecene inactivating enzymes. For this reason, the 10 μM IsoT with 1 M NaCl and the 15 μM IsoT with 0.5 M NaCl were selected as the best options because they introduce the least amount of selective pressure while still eliminating most cells that were

not expressing Tri101. It was hypothesized that 15 μM IsoT with 0.5 M NaCl inhibits protein synthesis more due to the higher concentration of IsoT, but the smaller increase in NaCl means that cells do not have to increase expression of stress induced genes as drastically. It was also thought that 10 μM IsoT with 1 M NaCl inhibits protein synthesis less but requires a greater increase in expression of stress induced genes in order for the cells to survive the amphotericin B.

It is not apparent which of these two screening options is more likely to select for inefficient enzymes, and it is possible that the best screen option may depend on their enzymatic characteristics like K_M and *k*cat. For example, it could be argued that 15 μM IsoT with 0.5 M NaCl could be more effective at selecting for enzymes which are inefficient primarily due to a high K_M, while 10 μM IsoT with 1 M NaCl might be more effective at selecting for enzymes which are inefficient primarily due to a low *k*cat. In the absence of information on the kinetic characteristics of possible resistance enzymes, it was decided that the best course of action would be to utilize both 10 μM IsoT with 1 M NaCl for Protocol 5a and 15 μM IsoT with 0.5 M NaCl for Protocol 5b, as illustrated in Figure 2.

Conclusions

Identifying an enzyme capable of breaking the essential epoxide bond of trichothecenes is a critical step in the permanent inactivation of trichothecene mycotoxins from harvested crops. To our knowledge, a high-throughput selection protocol capable of identifying trichothecene inactivating enzymes from large libraries has not been reported. Through an iterative approach of designing, testing, and modifying, we have designed a selection protocol with this capability. Key to the success of our selection protocol was the development of the RY0521 Green Monster *AYT1*Δ yeast strain with p425GPD-*RPL3*. The deletion of *AYT1* and addition of exogenous copies of *RPL3* drastically reduced the selection protocol background. The addition of liquid selection steps in the

form of trichothecene exposure followed by NaCl stress and the fungicide amphotericin B allowed our selection protocol to select for trichothecene inactivating enzymes even when their plasmids were outnumbering 1:1,000,000 with plasmids not containing resistance genes. This powerful selection protocol will be used as a tool to identify trichothecene inactivation enzymes from large mutant libraries and could play a significant role in combating *Fusarium* head blight.

References

1. Proctor, R. H., Hohn, T. M., and Mccormick, S. P. (1995) Reduced Virulence of *Gibberella-Zeae* Caused by Disruption of a Trichothecene Toxin Biosynthetic Gene, *Molecular Plant-Microbe Interactions 8*, 593-601.

2. Cundliffe, E., Cannon, M., and Davies, J. (1974) Mechanism of Inhibition of Eukaryotic Protein-Synthesis by Trichothecene Fungal Toxins, *P Natl Acad Sci USA 71*, 30-34.

3. Pestka, J. J. (2010) Deoxynivalenol: mechanisms of action, human exposure, and toxicological relevance, *Arch Toxicol 84*, 663-679.

4. Payros, D., Alassane-Kpembi, I., Pierron, A., Loiseau, N., Pinton, P., and Oswald, I. P. (2016) Toxicology of deoxynivalenol and its acetylated and modified forms, *Archives of Toxicology 90*, 2931-2957.

5. Poppenberger, B., Berthiller, F., Lucyshyn, D., Sieberer, T., Schuhmacher, R., Krska, R., Kuchler, K., Glossl, J., Luschnig, C., and Adam, G. (2003) Detoxification of the *Fusarium* mycotoxin deoxynivalenol by a UDP-glucosyltransferase from *Arabidopsis thaliana*, *J Biol Chem 278*, 47905-47914.

6. Binder, J., Horvath, E. M., Heidegger, J., Ellend, N., Danner, H., Krska, R., and Braun, R. (1997) A bioassay for comparison of the toxicity of trichothecenes and their microbial metabolites, *Cereal Res Commun 25*, 489-491.

7. Li, X., Shin, S., Heinen, S., Dill-Macky, R., Berthiller, F., Nersesian, N., Clemente, T., McCormick, S., and Muehlbauer, G. J. (2015) Transgenic Wheat Expressing a Barley UDP-Glucosyltransferase Detoxifies Deoxynivalenol and Provides High Levels of Resistance to *Fusarium graminearum*, *Mol Plant Microbe Interact 28*, 1237-1246.

8. Mack, M., Burger, M., Pietschmann, P., and Hock, B. (2008) A high-throughput microtiter plate-based screening method for the detection of full-length recombinant proteins, *Protein Expression and Purification 61*, 92-98.

9. Yang, G. Y., and Withers, S. G. (2009) Ultrahigh-Throughput FACS-Based Screening for Directed Enzyme Evolution, *Chembiochem 10*, 2704-2715.

10. Leemhuis, H., Kelly, R. M., and Dijkhuizen, L. (2009) Directed Evolution of Enzymes: Library Screening Strategies, *Iubmb Life 61*, 222-228.

11. Xiao, H., Bao, Z., and Zhao, H. (2015) High Throughput Screening and Selection Methods for Directed Enzyme Evolution, *Ind Eng Chem Res 54*, 4011-4020.

12. Hall, B. G. (1981) Changes in the Substrate Specificities of an Enzyme during Directed Evolution of New Functions, *Biochemistry 20*, 4042-4049.

13. Miller, B. G., and Raines, R. T. (2004) Identifying latent enzyme activities: Substrate ambiguity within modern bacterial sugar kinases, *Biochemistry 43*, 6387-6392.

14. Yang, K. C., and Metcalf, W. W. (2004) A new activity for an old enzyme: *Escherichia coli* bacterial alkaline phosphatase is a phosphite-dependent hydrogenase, *Proceedings of the National Academy of Sciences of the United States of America 101*, 7919-7924.

15. Woolford, J. L., Jr., and Baserga, S. J. (2013) Ribosome biogenesis in the yeast *Saccharomyces cerevisiae*, *Genetics 195*, 643-681.

16. Brodersen, D. E., and Nissen, P. (2005) The social life of ribosomal proteins, *Febs Journal 272*, 2098-2108.

17. Schulze, H., and Nierhaus, K. H. (1982) Minimal Set of Ribosomal Components for Reconstitution of the Peptidyltransferase Activity, *Embo Journal 1*, 609-613.

18. Meskauskas, A., and Dinman, J. D. (2007) Ribosomal protein L3: Gatekeeper to the a site, *Molecular Cell 25*, 877-888.

19. Mitterbauer, R., Poppenberger, B., Raditschnig, A., Lucyshyn, D., Lemmens, M., Glossl, J., and Adam, G. (2004) Toxin-dependent utilization of engineered ribosomal protein L3 limits trichothecene resistance in transgenic plants, *Plant Biotechnology Journal 2*, 329-340.

20. Garreau de Loubresse, N., Prokhorova, I., Holtkamp, W., Rodnina, M. V., Yusupova, G., and Yusupov, M. (2014) Structural basis for the inhibition of the eukaryotic ribosome, *Nature 513*, 517-522.

21. Suzuki, Y., St Onge, R. P., Mani, R., King, O. D., Heilbut, A., Labunskyy, V. M., Chen, W., Pham, L., Zhang, L. V., Tong, A. H., Nislow, C., Giaever, G., Gladyshev, V. N., Vidal, M., Schow, P., Lehar, J., and Roth, F. P. (2011) Knocking out multigene redundancies via cycles of sexual assortment and fluorescence selection, *Nat Methods 8*, 159-164.

22. Mumberg, D., Muller, R., and Funk, M. (1995) Yeast vectors for the controlled expression of heterologous proteins in different genetic backgrounds, *Gene 156*, 119-122.

23. Chen, G. J., Qiu, N., Karrer, C., Caspers, P., and Page, M. G. P. (2000) Restriction site-free insertion of PCR products directionally into vectors, *Biotechniques 28*, 498-+.

24. Looke, M., Kristjuhan, K., and Kristjuhan, A. (2011) Extraction of genomic DNA from yeasts for PCR-based applications, *Biotechniques 50*, 325-328.

25. Kimura, M., Kaneko, I., Komiyama, M., Takatsuki, A., Koshino, H., Yoneyama, K., and Yamaguchi, I. (1998) Trichothecene 3-*O*-acetyltransferase protects both the producing organism and transformed yeast from related mycotoxins - Cloning and characterization of Tri101, *J Biol Chem 273*, 1654-1661.

26. Alexander, N. J., McCormick, S. P., and Hohn, T. M. (2002) The identification of the *Saccharomyces cerevisiae* gene *AYT1*(ORF-YLL063c) encoding an acetyltransferase, *Yeast 19*, 1425-1430.

27. Rocha, O., Ansari, K., and Doohan, F. M. (2005) Effects of trichothecene mycotoxins on eukaryotic cells: a review, *Food Addit Contam 22*, 369-378.

28. Adam, G., Mitterbauer, R., Raditschnig, A., Poppenberger, B., Karl, T., Goritschnig, S., Weindorfer, H., and Glossl, J. (2001) Molecular mechanisms of deoxynivalenol resistance in the yeast *Saccharomyces cerevisiae*, *Mycotoxin Res 17 Suppl 1*, 19-23.

29. Mitterbauer, R., and Adam, G. (2002) *Saccharomyces cerevisae* and *Arabidopsis thaliana*: Useful model systems for the identification of molecular mechanisms involved in resistance of plants to toxins, *European Journal of Plant Pathology 108*, 699-703.

30. Causton, H. C., Ren, B., Koh, S. S., Harbison, C. T., Kanin, E., Jennings, E. G., Lee, T. I., True, H. L., Lander, E. S., and Young, R. A. (2001) Remodeling of yeast genome expression in response to environmental changes, *Molecular Biology of the Cell 12*, 323-337.

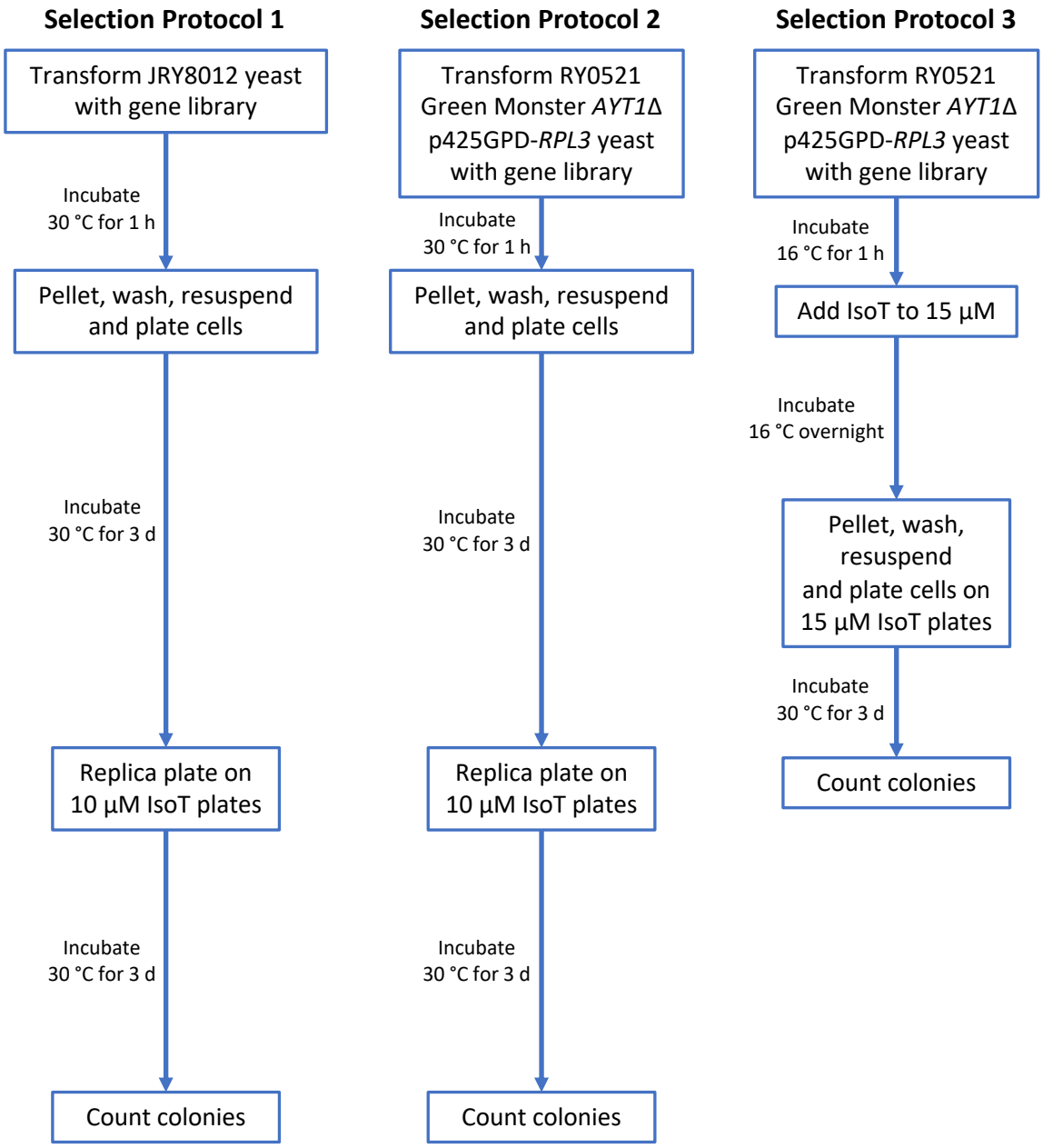


Figure 1. Outline of selection protocols 1, 2, and 3. Selection protocol steps are shown in blue boxes with incubation times and temperatures given in between. All incubations were carried out with 200 rpm shaking. Protocol details can be found in the Experimental Procedures section.

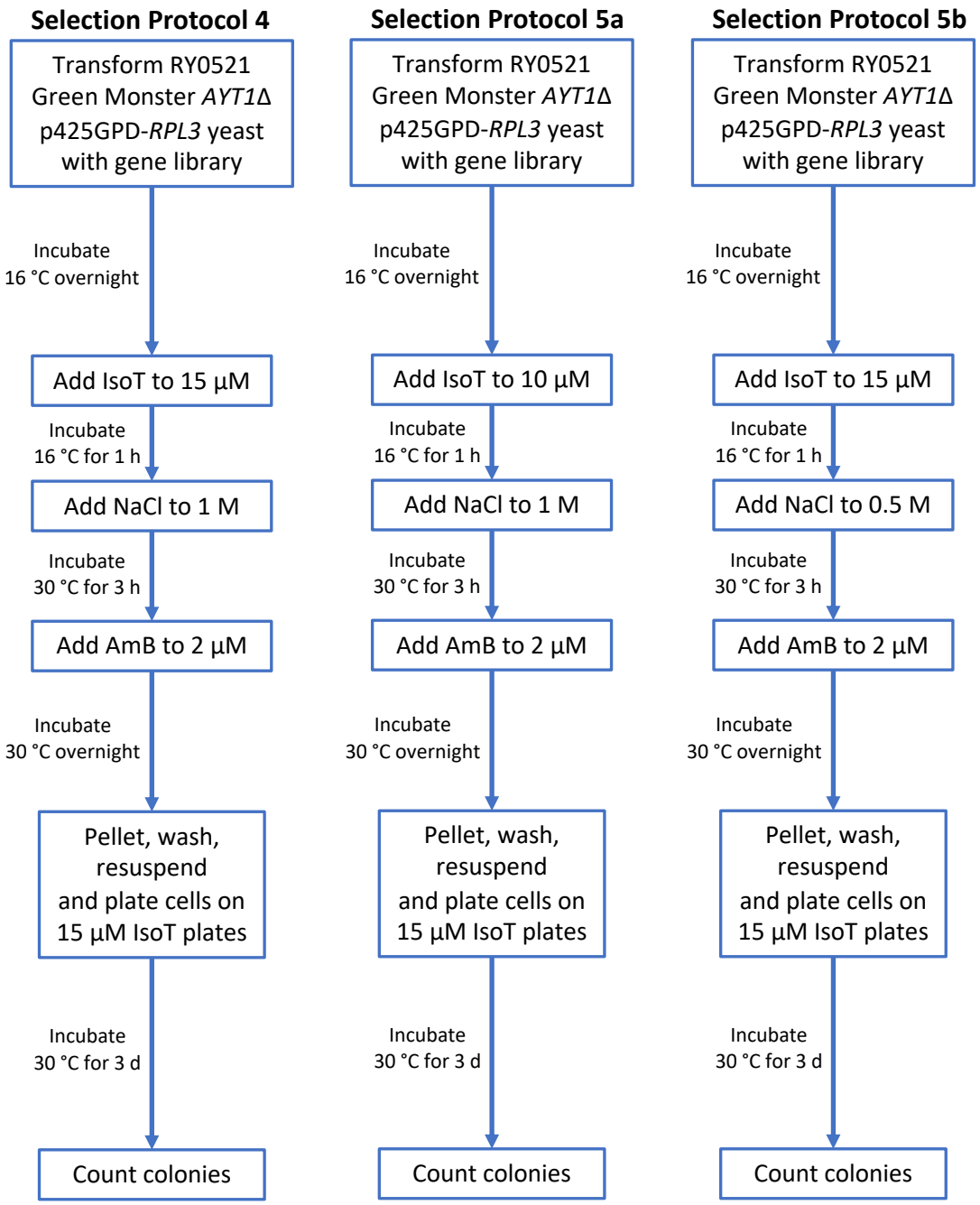


Figure 2. Outline of selection Protocols 4, 5a, and 5b. Outline of selection protocols 1, 2, and 3. Selection protocol steps are shown in blue boxes with incubation times and temperatures given in between. All incubations were carried out with 200 rpm shaking. Protocol details can be found in the Experimental Procedures section.

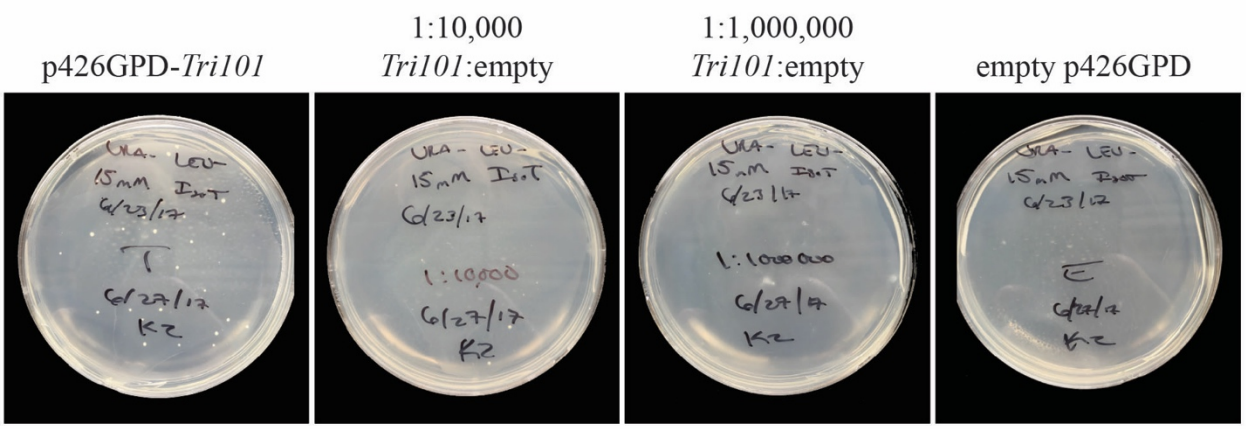


Figure 3. Trichothecene inactivating enzyme selection Protocol 3 control. RY0521 Green Monster *AYT1*Δ with p425GPD-*RPL3* yeast were transformed with p426GPD-*Tri101* as a positive control, empty p426GPD as negative control a mixture of 1:10,000 p426GPD-*Tri101* to empty p426GPD plasmid and a mixture of 1:1,000,000 p426GPD-*Tri101* to empty p426GPD plasmid. Following cell recovery, IsoT was added to the cultures to a final concentration of 15 μM and the cultures were incubated overnight 16 °C. Cells were pelleted and washed, resuspended in 10 mL of URA-LEU- media and 150 μL of the cell solution was plated onto URA-LEU- agar plates containing 15 μM IsoT. Plates were incubated at 30 °C and pictures were taken after 3 days.

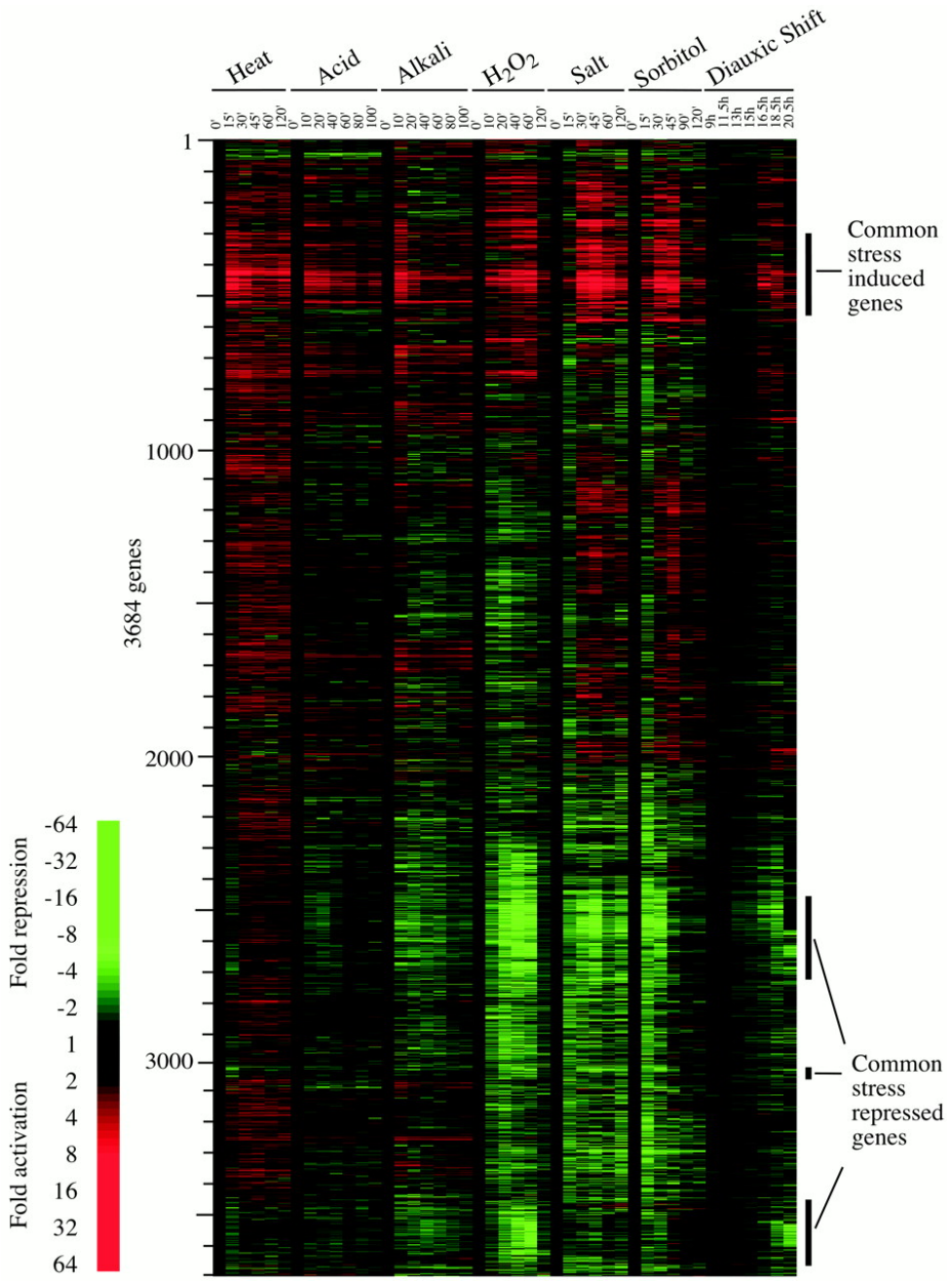


Figure 4. Changes in expression profile for yeast exposed to environmental stresses. Adopted from Causton *et al.* (*30*).

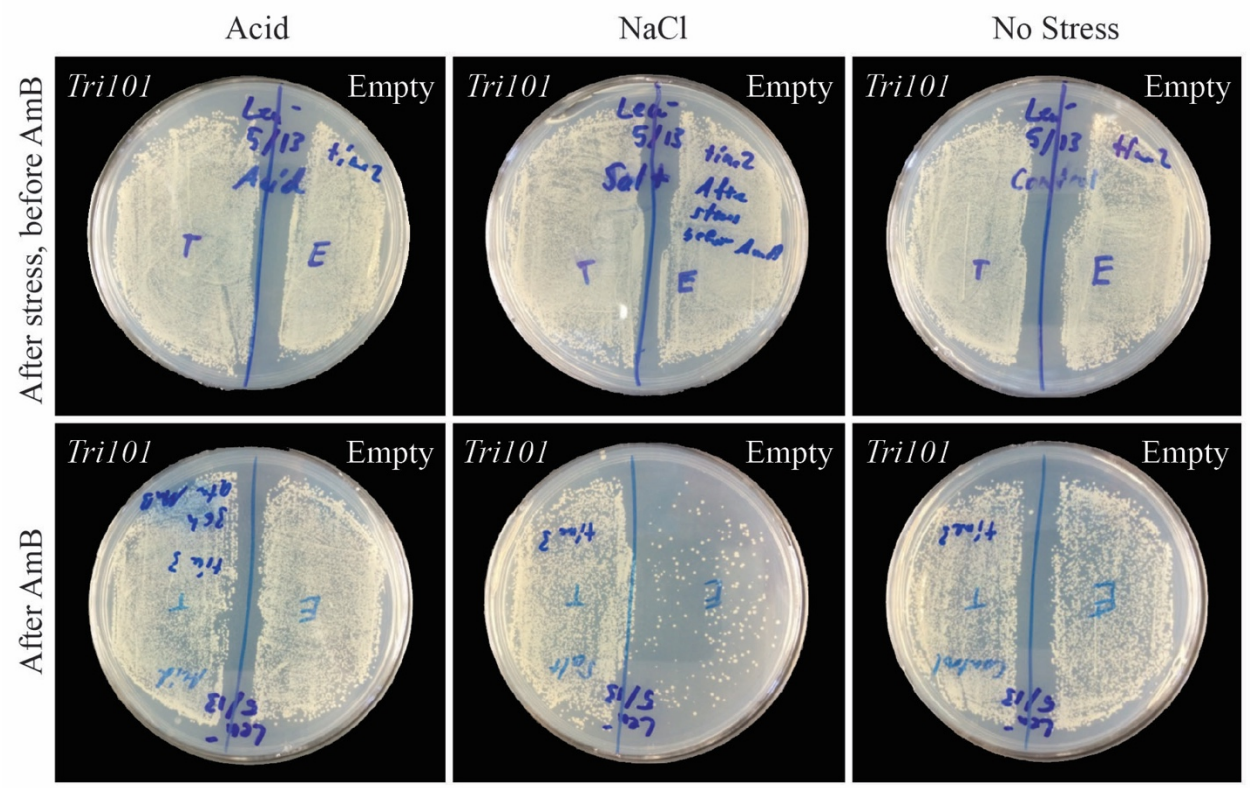


Figure 5. Effect of increasing NaCl in yeast liquid culture followed by treatment with amphotericin B selects for yeast with trichothecene resistance genes. Yeast were transformed with p426GPD-*Tri101* (left side of each plate) and empty p426GPD (right side of each plate). Cells were recovered and grown overnight followed by the addition of IsoT. After a 1 h incubation the culture was split into three equal cultures. NaCl was added to a final concentration of 1 M in one culture, the pH of the second culture was dropped from 6.0 to 4.0 by the addition of succinic acid, and no stress was introduced to the third culture. After a 3 h incubation 150 µL of each culture was plated onto minimal media plates (top row). Amphotericin B was added to each culture to a final concentration of 2 µM and the cultures were incubated overnight. After washing, 150 µL of each culture was plated onto minimal media plates and incubated for three day at 30 °C (bottom row).

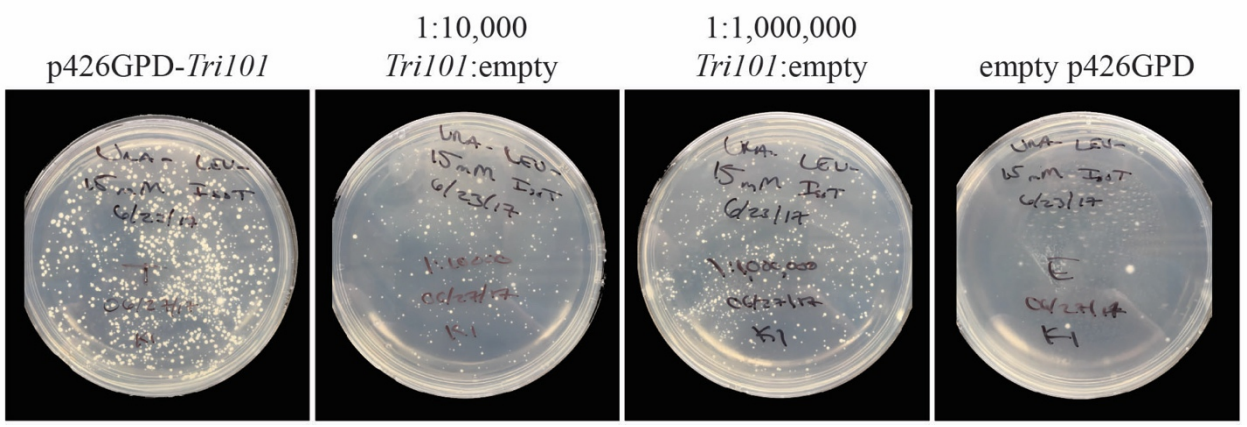


Figure 6. Trichothecene inactivating enzyme selection Protocol 4 control. RY0521 Green Monster *AYT1*Δ cells with p425GPD-*RPL3* were transformed with p426GPD-*Tri101* as a positive control, empty p426GPD as negative control a mixture of 1:10,000 p426GPD-*Tri101* to empty p426GPD plasmid and a mixture of 1:1,000,000 p426GPD-*Tri101* to empty p426GPD plasmid. Cells were incubated overnight 16 °C. IsoT was added to the cultures (15 μM) the cultures were incubated for 1 h, NaCl was added (1 M) and the cultures were incubated for 3 h, amphotericin B was added (2 μM) and the cultures were incubated overnight at 30 °C. Cells were pelleted and resuspended in 10 mL of URA- LEU- media and 150 μL of the cell solution was plated onto URA-LEU- agar plates containing 15 μM IsoT. Plates were incubated at 30 °C for 3 days.

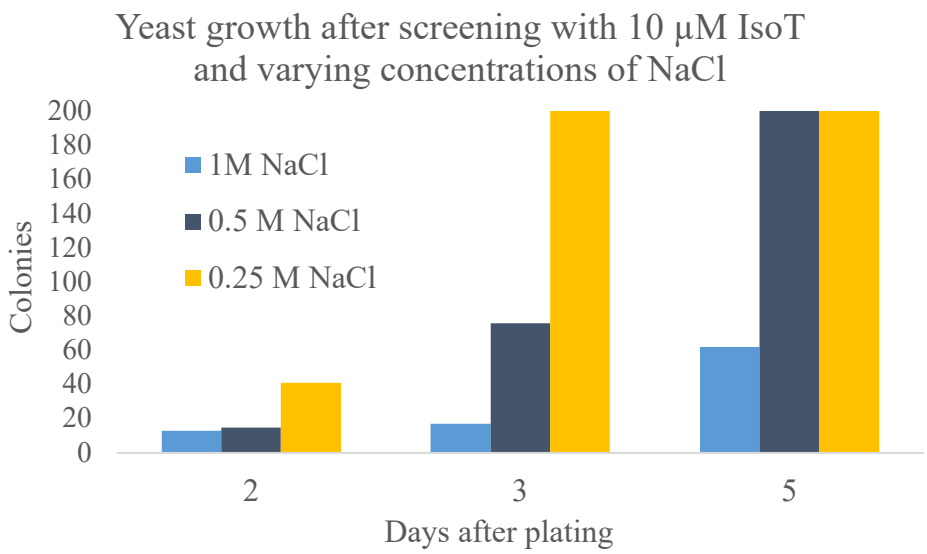


Figure 7. Yeast growth after selection with 10 µM IsoT and varying concentrations of NaCl. Yeast were transformed with a positive control p426GPD-*Tri101* plasmid or an empty p426GPD plasmid, incubated with 10 µM for 1 h, incubated with 1, 0.5, 0.25 M NaCl for 3 h, incubated with amphotericin B overnight and plated onto URA-LEU- agar plates with 15 µM IsoT. Colonies were counted after 2, 3, and 5 days. Bars of 200 colonies represent lawns (over 2000 colonies). Yeast transformed with the positive control p426GPD-*Tri101* plasmid produced lawns on the agar plates after two days. Colony counts shown above are only from plates with yeast containing the empty p426GPD plasmid.

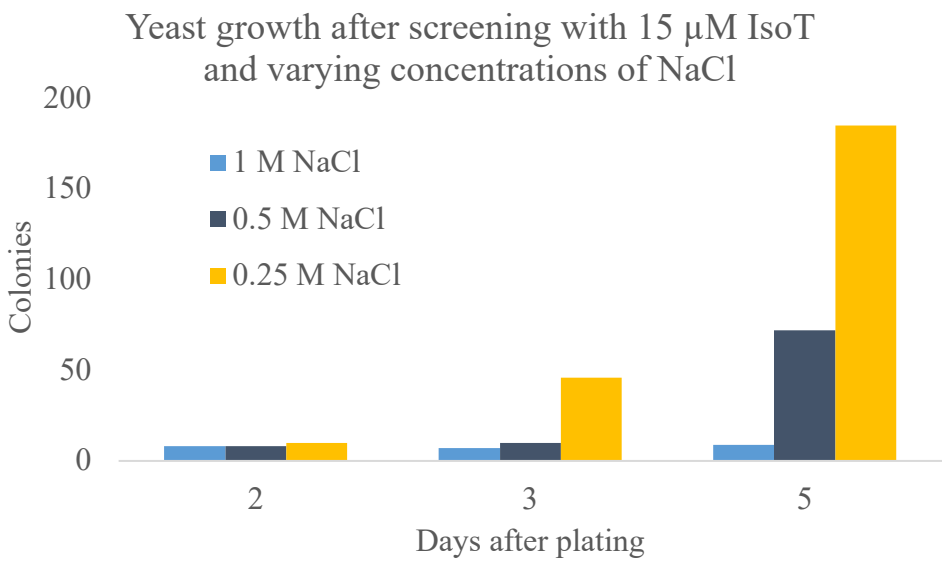


Figure 8. Yeast growth after screening with 15 µM IsoT and varying concentrations of NaCl. Yeast were transformed with a positive control p426GPD-*Tri101* plasmid and an empty p426GPD plasmid, incubated with 10 µM for 1 h, incubated with 1, 0.5, 0.25 M NaCl for 3 h, incubated with amphotericin B overnight and plated onto URA-LEU- plates with 15 µM IsoT. Colonies were counted after 2, 3 and 5 days. Yeast transformed with the positive control p426GPD-*Tri101* plasmid produced lawns on the agar plates after two days. Colony counts shown above are only from plates with yeast containing the empty p426GPD plasmid.

Chapter 5:

The PrpF protein of *Shewanella oneidensis* MR-1 catalyzes the isomerization of 2-methyl-cis-aconitate during the catabolism of propionate via the AcnD-dependent 2-methylcitric acid cycle

This work was performed in collaboration with Dr. Christopher Rocco, Dr. Graeme Garvey, Dr. Ivan Rayment, and Dr. Jorge Escalante-Semerena. Protein purification, crystallization and structural determination of *So*PrpF^{K73E} was performed by Karl Wetterhorn. All other experiments were performed by Dr. Christopher Rocco and Dr. Graeme Garvey.

A version of this work was previously published as:
Rocco CJ, Wetterhorn KM, Garvey GS, Rayment I, Escalante-Semerena JC. *The PrpF protein* of Shewanella oneidensis *MR-1 catalyzes the isomerization of 2-methyl*-cis-*aconitate during the catabolism of propionate via the AcnD-dependent 2-methylcitric acid cycle*. PloS one. 2017 Nov 16;12(11):e0188130.

Abstract

The 2-methylcitric acid cycle (2-MCC) is a common route of propionate catabolism in microorganisms. In *Salmonella enterica*, the *prpBCDE* operon encodes most of the 2-MCC enzymes. In other organisms, *e.g.*, *Shewanella oneidensis* MR-1, two genes, *acnD* and *prpF* replace *prpD*, which encodes 2-methylcitrate dehydratase. We showed that together, *S. oneidensis* AcnD and PrpF (*So*AcnD, *So*PrpF) compensated for the absence of PrpD in a *S. enterica prpD* strain. We also showed that *So*AcnD had 2-methylcitrate dehydratase activity and that PrpF has aconitate isomerase activity. Here we report *in vitro* evidence that the product of the *So*AcnD reaction is an isomer of 2-methyl-cis-aconitate (2-MCA], the product of the *Se*PrpD reaction. We show that the *So*PrpF protein isomerizes the product of the AcnD reaction into the PrpD product (2-MCA], a known substrate of the housekeeping aconitase (AcnB]. Given that *So*PrpF is an isomerase, that *So*AcnD is a dehydratase, and the results from *in vivo* and *in vitro* experiments reported here, it is likely that 4-methylaconitate is the product of the AcnD enzyme. Results from in vivo studies using a *S. enterica prp*D strain show that *So*PrpF variants with substitutions of residues K73 or C107 failed to support growth with propionate as the sole source of carbon and energy. High-resolution (1.22 Å) three-dimensional crystal structures of $PrpF^{K73E}$ in complex with *trans*-aconitate or malonate provide insights into the mechanism of catalysis of the wild-type protein.

Introduction

The 2-methylcitric acid cycle (2-MCC) (Figure 1) is widely distributed route of propionate catabolism in microorganisms. Originally identified in the fungus *Candida lipolytica* (*Yarrowia lipolytica*) (*1*), the 2-MCC has been characterized in Gamma-proteobacteria (*e.g*., *Salmonella enterica*, *Escherichia coli*) (*2-4*), actinobacteria (*e.g*. *Mycobacterium tuberculosis*, *Mycobacterium smegmatis*, *Corynebacterium glutamicum*) (*5-9*), and Beta-proteobacteria (*e.g*. *Ralstonia eutropha*, *Burkholderia sacchari*) (*10, 11*). Most of the enzymes that comprise the 2-MCC are encoded as an operon (*12*). In *S. enterica* (and many other bacteria), the operon consists of four genes, in the order *prpBCDE*. *prpB* encodes 2-methylisocitrate lyase (EC 4.2.1.99) (*2, 13, 14*); *prpC* encodes the 2-methylcitrate synthase (EC 2.3.3.5) (*2*); *prpD*, encodes the 2-methylcitrate dehydratase (EC 4.2.1.79) (*15*); and *prpE* encodes the propionyl-CoA synthetase (EC 6.2.1.17) (*16, 17*).

There is one notable variation in bacterial *prp* operons. That is, some *prp* operons lack the 2-methylcitrate dehydratase *prpD*, but rather contain two other genes, *acnD* and *prpF* (*15*). AcnD has aconitase-like activity (*18*), whilst PrpF has aconitate isomerase activity (*19*) (Figure 1). In *S. oneidensis* the sequence of *prp* genes is *prpR prpB prpC acnD prpF* compared to the *S. enterica* sequence *prpR prpB prpC prpD prpE*. In *S. oneidensis*, the *prpE* gene encoding propionyl-CoA synthetase is >1.5 Mbp away from the *prp* operon. AcnD and PrpF activities are necessary and sufficient to compensate for the lack of PrpD during growth with propionate of a *prpD* strain of *S. enterica* (*18*).

Structural and biochemical analyses of *Shewanella oneidensis* PrpF (hereafter *So*PrpF), revealed a role of *So*PrpF in the isomerization of *trans*-aconitate to *cis*-aconitate (*19*), leading us to propose that the role of *So*PrpF in the 2-MCC was to change the stereochemistry of the *S.*

oneidensis AcnD (hereafter *So*AcnD) reaction product. That is, *So*PrpF was proposed to have 2-methylaconitate isomerase activity (*19*).

Herein we specifically address the hypothesis that the *S. oneidensis* PrpF (*So*PrpF) protein isomerizes the product of the *So*AcnD dehydratase (putatively 4-methyl-cis-aconitate (4-MCA) to 2-methyl-cis-aconitate (2-MCA), the substrate of aconitase A, B. To facilitate this work, we used a *S. enterica* Δ*prpD* carrying the wild-type allele of the *S. oneidensis acnD* gene. Into the *S. enterica* Δ*prpD* / *pacnD*+ (*S. oneidensis*) strain we introduced a second plasmid encoding wild-type or variants of *S. oneidensis* PrpF (*So*PrpF). In this heterologous system, PrpF functionality was assessed *in vivo* under conditions that demanded propionate utilization as the sole source of carbon and energy, or a combination of propionate and succinate, which would allow us to assess the effect of the accumulation of 4-MCA on growth in the presence of functional or dysfunctional PrpF proteins.

Experimental Procedures

Chemicals and bacteria culture media. All chemicals were purchased from Sigma Chemical Co. unless otherwise stated. 2-Methylcitrate was purchased from CDN Isotopes (Pointe-Claire, Canada). Authentic 2-methyl-*cis*-aconitate was custom synthesized by AsisChem (Cambridge, MA). The experiments reported here were performed in well-characterized *S. enterica prp* mutant strains. A previous report from our laboratory showed that this approach allowed us to assess *S. oneidensis* AcnD and PrpF functions *in vivo* (*18*). *E. coli* strains used to overproduce recombinant proteins were grown in lysogeny broth (LB) (*20, 21*). No-carbon essential (NCE) medium (*22*) was used as minimal medium, and was supplemented with MgSO4 (1 mM) and methionine (0.5 mM). When added to rich medium, antibiotic concentrations were:

ampicillin (100 μg/ml), kanamycin (50 μg/ml), and chloramphenicol (25 μg/ml). BactoTM Agar (Difco, 1.5% w/v) was used as medium solidifying agent.

For cloning purposes, restriction endonuclease NdeI was purchased from Fermentas (Glen Burnie, MD), and BamHI was purchased from Promega (Madison, WI). Cloning was performed in *E. coli* strain DH5α/F' (New England Biolabs). Plasmids were introduced into *S. enterica* strains by electroporation. Cultures were grown in LB medium to an optical density (650 nm) ~0.6–0.8, 1.0 ml of culture was centrifuged at 18,000 x *g* using a Microfuge 18 Centrifuge (Beckman Coulter), cells were washed three times with 1.0 ml of ice-cold sterile water and re-suspended in 100 μl of water. Plasmids were electroporated into cells using a Bio-Rad Gene Pulser (Hercules, CA) according to the manufacturer's recommendations. Strains and plasmids used in this study are listed in Table 1.

Site-directed mutagenesis. All site-directed mutations were introduced into targeted genes using the QuikChange® II XL Site-Directed Mutagenesis Kit (Stratagene); all manipulations of plasmids carrying wild-type or mutant alleles of genes of interest were performed in XL10 Gold® Ultracompetent *E. coli* cells (Stratagene).

Polymerase chain reaction (PCR). Amplification conditions for site-directed mutagenesis were as follows: 95°C for 1 min, followed by 19 cycles of 95°C for 50 s, 60°C for 50 s, 68°C for 6 min 15 s, ending with 68°C for 7 min. All plasmids carrying mutant *S. oneidensis prpF* alleles were sequenced using two described primers (*19*). DNA sequencing reactions were performed using BigDye® (Applied Biosystems), were purified using CleanSEQ protocols (Agentcourt Biotechnology), and were resolved at the University of Wisconsin-Madison Biotechnology Center.

Plasmid pPRP195. The *S. oneidensis acnD*+ gene was amplified from plasmid pPRP141 (*18*) using primers 5'–GTT ATG AGC ACA CAT ATG AAC ACC CAA TAT C– 3' and 5' -GAT ATA GGC GGG ATC CAT GTC GGC ATT GC-3'. The resulting DNA fragment (~2.5 kb) was extracted from the gel using the QIAQuick Gel Extraction kit (Qiagen), the fragment was digested with NdeI and BamHI, and ligated into plasmid pET-15b (bla+) digested with the same enzymes; the resulting plasmid (pPRP195) was electroporated into *E. coli* DH5α/F', and cells were plated onto LB + ampicillin medium.

Plasmid pPRP205. Plasmid pPRP195 was digested with NdeI and BamHI; the fragment containing the *S. oneidensis acnD*+ allele was extracted from the gel as described above. The fragment was ligated into plasmid pTEV5 (*23*) digested with the same enzymes. The resulting plasmid (pPRP205) was electroporated into *E. coli* DH5α/F', and cells were plated onto LB + ampicillin medium.

Plasmids pPRP215-226. Plasmids pPRP153 (*18*) and pPRP196 (*19*) were used as templates to generate single-amino acid variants of the *So*PrpF protein with the QuikChange® II XL Site-Directed Mutagenesis kit (Stratagene). Other information pertinent to the construction of these plasmids is summarized in Table 2.

Isolation of proteins. *So*PrpF protein was purified as described (*19*). Variant *So*PrpF proteins were purified using the Maxwell™ 16 system (Promega). Cultures were grown in LB + ampicillin at 37°C and induced overnight with isopropyl-ß-D-thiogalactopyranoside (IPTG, 0.3 mM) at an OD650 of ~0.8. Recombinant PrpD, apo-AcnA and apo-AcnB proteins from *S. enterica* were purified from *E. coli* as His-tagged proteins as described (*15*). *So*AcnD protein was purified as follows. Plasmid pPRP205 (*S. oneidensis acnD*+) was introduced into *E. coli* strain BL21 (λDE3) by electroporation selecting for ampicillin resistance on LB agar + ampicillin. Single

colonies were used to inoculate 20 ml of LB + ampicillin; cultures were grown overnight at 37°C. Overnight cultures were used to inoculate two liters of LB + ampicillin, and cells were grown at 37°C until the culture reached an OD650 of ~0.7. At that point, expression of the plasmid-encoded *S. oneidensis acnD*+ gene was induced by the addition of IPTG (0.3 mM) to the medium, followed by an 18-h incubation period at 37°C. Cells were harvested by centrifugation at 8,000 x *g* at 4°C in 1-liter bottles using a JLA-8.l rotor and a Beckman/Coulter Avanti™ J-20 XPI centrifuge. Cells were broken by sonication (10 min, 50% duty, 5 s pulses, maximal setting) with a 550 Sonic Dismembrator (Fisher Scientific). Cell debris was removed by centrifugation at 39,000 x *g* for 20 min in a Beckman JA 25.5 rotor. The supernatant was filtered through a 0.45 μm filter (Thermo Fisher Scientific). Protein was isolated using Ni-chelate affinity chromatography on Novagen's His Bind® Resin following the manufacturer's protocols.

Crystallization and structural determination of *So*PrpF^{K73E}. *So*PrpF^{K73E} was screened for initial crystallization conditions in a 144-condition sparse matrix screen developed in the Rayment laboratory (unpublished information). Single, diffraction quality crystals were grown by hanging drop vapor diffusion by mixing 2 μL of 28 mg/mL *So*PrpF^{K73E} in 2-amino-2-(hydroxymethyl)propane-1,3-diol hydrochloride buffer (Tris-HCl, 10 mM, pH 7.6) containing NaCl (50 mM) with 2 μL well solution containing 2-(N-morpholino)ethanesulfonic acid buffer (MES, 100 mM, pH 6.0) containing sodium malonate (136 mM), polyethylene glycol 4000 (PEG 4K, 15% w/v) at room temperature. Hanging droplets were nucleated after 24 h from an earlier spontaneous crystallization event using a cat's whisker. Crystals grew to approximate dimensions of 200 X 200 X 400 μm within 3 days. The crystals were transferred directly to a cryoprotecting solution that contained MES buffer (100 mM, pH 6.0), sodium malonate (136 mM), PEG 4K (30% w/v) and vitrified by rapid plunging into liquid nitrogen. *So*PrpF^{K73E} crystallized in the space group

$P2_1$ with unit cell dimensions of a = 51.8 Å, b = 103.4 Å, c = 78.1 Å and two chains in the asymmetric unit.

X-ray diffraction data were collected on a Pilatus detector at SBC Beamline 19-ID (Advanced Photon Source, Argonne National Laboratory, Argonne, IL) The X-ray data were processed and scaled using the HKL-2000 program that integrates data collection, data reduction, phasing and model building (*24*). Relevant X-ray diffraction data collection statistics are presented in Table 3. The previously determined model for *So*PrpF apo structure (PDB ID: 2PVZ) was used as the search model to solve the *So*PrpF^{K73E} apo structure via molecular replacement with the program Phaser (*25*). Alternate cycles of manual model building and least squares refinement with the programs COOT (*26*) Refmac (*27*) and Phenix (*28*) reduced the R-factor to 16.5% for all X-ray data from 50–1.22 Å. Relevant refinement statistics are presented in Table 3.

Reactivation of aconitases. The Fe/S centers of *S. enterica* AcnA (hereafter *Se*AcnA), AcnB (hereafter *Se*AcnB), and *So*AcnD were reactivated using described protocols without modifications (*29, 30*). All solutions used were freed of dissolved O_2 by degassing as described (*31, 32*).

High-performance liquid chromatography (HPLC). Enzyme-dependent dehydration of citrate and 2-methylcitrate was performed in 1-ml reaction mixtures containing Tris-HCl buffer (90 mM, pH 8.0 at 25°C), citrate or 2-MC (5 mM), and *So*AcnD (10 μg of reactivated protein) or *S. enterica* PrpD (hereafter *Se*PrpD, 13 μg). Reaction mixtures were incubated for 1.5 h at 37°C. Reactions were stopped by the addition of 10 N H_2SO_4 to a final concentration of 5 mM. Particulate matter was removed from reaction products by filtration using a Spin-X® centrifuge tube filter (Costar), and products were resolved by HPLC using a Beckman/Coulter chromatograph equipped with an Aminex® HPX-87H HPLC organic acid analysis column (BioRad) equilibrated and

developed isocratically with H_2SO_4 (5 mM). Elution of compound off the column was detected by monitoring the absorbance at 210 nm.

Kinetic analysis of *S. enterica* aconitases. All reactions were performed in Tris-HCl buffer (50 mM, pH 8.0 at 25°C) containing KCl (100 mM). Enzyme was incubated in buffer for 3–5 min before reactions were initiated by the addition of either *cis*-aconitate or 2-methy-*cis*-aconitate as substrate. Reactions were performed in triplicate, and their progress was monitored using a Perkin Elmer (Norwalk, CT) Lambda 40 UV/Vis Spectrometer at 240 nm; temperature in the cuvettes was maintained with a circulating water bath set at 37°C. Data collection and analysis was performed with Perkin Elmer UV Kinlab software. An extinction coefficient of 3800 M^{-1} cm^{-1} was used for *cis*-aconitate (*33*), and 4690 M^{-1} cm^{-1} was calculated for the chemically synthesized 2-methyl-*cis*-aconitate. Kinetic curves were each repeated three times, with each substrate concentration tested in triplicate. The reported values are the median of all three experiments. Data did not deviate more than 15% from the median value.

***In vivo* assessment of activity associated with variant *So*PrpF proteins**. Growth curves of *S. enterica* strains were performed in NCE medium supplemented with succinate (30 mM, pH 7.0 at 25°C) or propionate (30 mM, pH 7.0 at 25°C) as carbon and energy source. In both cases a low concentration of glycerol (1 mM) was added to accelerate the catabolism of succinate or propionate. Strains were grown overnight in LB medium supplemented with the appropriate antibiotic. Two microliters of an overnight culture were used to inoculate 198 µl of NCE medium in a 96-well microtiter plate. Growth was monitored using an EL808TM microplate reader (Bio-Tek Instruments) with the incubation chamber set at 37°C. Absorbance readings were recorded every 15 min at 630 nm with 850 s of shaking between readings. All cultures were grown in triplicate. Growth curves were plotted using Prism v4.0 software (GraphPad Software).

***In vitro* assessment of activity associated with variant *So*PrpF proteins**. To analyze the activity of *So*PrpF variants, the product of the *So*AcnD enzyme was synthesized as follows. Reactivated *So*AcnD was incubated overnight at 30°C with 2-MC (1 mM) in Tris-HCl buffer (50 mM, pH 8.0) containing KCl (100 mM). *So*AcnD protein was removed from the mixture by the addition of His-MagTM Agarose Beads (Novagen) resuspended in Tris-HCl buffer (50 mM, pH 8.0) containing KCl (100 mM). His-MagTM beads were removed, and reaction mixtures were pooled. *So*PrpF protein (100 ng) was added to the reactions, samples were taken as a function of time, and were resolved by HPLC as described above.

Results

***So*AcnD and *Se*PrpD synthesize different methylaconitate isomers**. We previously showed that *So*PrpF has isomerase activity that can convert *cis*-aconitate to *trans*-aconitate (*19*). Given that *So*PrpF isomerase activity is needed to restore growth of a *S. enterica prpD* strain with propionate, we surmised that *Se*PrpD and *So*AcnD must synthesize two different methylaconitate isomers, and that *So*PrpF isomerizes the *So*AcnD product into the *Se*PrpD product, so that *Se*AcnB, the next enzyme in the 2-MCC, can synthesize 2-methylisocitrate (Figure 1).

To test this hypothesis, we determined whether *Se*PrpD and *So*AcnD proteins synthesized different isomers of aconitate and methylaconitate. To do this, we incubated *Se*PrpD and *So*AcnD with citrate or 2-methylcitrate (each at 5 mM) in 1-ml reaction mixtures. After a 1.5-hr incubation period, the reactions mixtures were acidified with H_2SO_4 to a final concentration of 5 mM, and a 100 µl sample of the reaction was resolved by HPLC. When incubated with citrate, *Se*PrpD (Figure 2A) and *So*AcnD (Figure 2B) converted citrate (retention time ~8 min) into *cis*-aconitate (retention time ~7.2 min). When incubated with 2-methylcitrate, *Se*PrpD synthesized a product that eluted

off the column as a broad peak centered at ~13.8 min (Figure 2C). On the other hand, *So*AcnD synthesized a product that eluted as a sharp peak at ~8 min (Figure 2D). The different retention times and chromatographic behavior indicated to us that the dehydration product of *Se*PrpD and *So*AcnD were different compounds. Analysis of chemically synthesized 2-methyl-*cis*-aconitate revealed a peak that matched the product of the *Se*PrpD reaction (Figure 2C), indicating that *Se*PrpD synthesized 2-methyl-*cis*-aconitate from 2-methylcitrate.

2-Methyl-*cis*-aconitate is a substrate of *Se*AcnA and *Se*AcnB. The HPLC chromatograms of the products of the *Se*PrpD and *So*AcnD reactions suggested that the *Se*PrpD product, 2-methyl-*cis*-aconitate, was the substrate that aconitases rehydrated to yield 2-methylisocitrate agreeing with previously reported data (*34, 35*). Since *So*AcnD cannot support growth with propionate in the absence of *So*PrpF, we surmised that the product of *So*AcnD was not a substrate of aconitases, but the product of the *So*PrpF reaction was.

Kinetic analysis of the *S. enterica* aconitases indicated that both, *Se*AcnA and *Se*AcnB, used 2-methyl-*cis*-aconitate as substrate to yield 2-methylisocitrate, albeit at a significantly slower rate than *cis*-aconitate, the TCA cycle intermediate (Table 4). The catalytic efficiencies of *Se*AcnA and *Se*AcnB were 40 and 60 lower when 2-methyl-*cis*-aconitate was the substrate than when *cis*-aconitate was the substrate, respectively. When 2-methyl-*cis*-aconitate was the substrate, K_M values ranged from 180 to 229 μM and V_{max} values ranged from 50 to 60 μM min^{-1}. The numbers reported here are representative of three kinetic studies.

***So*PrpF isomerizes the *So*AcnD reaction product to 2-methyl-*cis*-aconitate**. We tested whether *Se*AcnA and *Se*AcnB used the *So*AcnD reaction product as substrate (data not shown). The latter was synthesized from 2-MC using *So*AcnD. After the *So*AcnD enzyme was removed from the reaction mixture, either *Se*AcnA or *Se*AcnB was added, and the change in absorbance at

240 nm was monitored. Neither *Se*AcnA nor *Se*AcnB used the *So*AcnD reaction product as substrate. Purified *So*PrpF protein (100 μg) was added to the reaction mixture, and the reaction was allowed to proceed for one hour at 30°C. Addition of *Se*AcnA resulted in specific activities of approximately 14 ± 1 nmol min^{-1} μg^{-1} of protein, while addition of *Se*AcnB resulted in specific activities of 48 ± 4 nmol min^{-1} μg^{-1} of protein. Additionally, HPLC analysis of the products of the *So*PrpF reactions indicated that *So*PrpF converted the *So*AcnD reaction product into 2-methyl-*cis*-aconitate over time (Figure 3), which can then be used as a substrate by the aconitase enzymes.

Analysis of variants to gain insights into the mechanism of *So*PrpF catalysis. Previous crystallographic data identified residues C107 and K73 in the active site of *So*PrpF as likely to be involved in catalysis. Residue C107 is formally equivalent to the catalytic base of diaminopimelate epimerase (*19, 36, 37*), and residue K73 is in a structurally equivalent position to the catalytic glutamate of the phenazine biosynthetic protein PhzF from *Pseudomonas fluorescens* (*38*). We investigated the involvement of these two residues in *So*PrpF catalysis, using site-directed mutagenesis to introduce amino acids changes. Five *So*PrpF variants were constructed. Residue C107 was changed to either Ala or Ser, and residue K3 was changed to Ala, Glu, or Met. Plasmid pPRP153 (*prpF* in pBAD18-Kan) was used in *in vivo* complementation studies using a *S. enterica prpD* strain to assess the effect of specific substitutions on PrpF function. Substitutions at C107 (Figure 4A) and at K73 (Figure 4B) resulted in proteins that failed to support growth with propionate as a sole carbon and energy source in a *S. enterica* Δ*prpD* strain harboring a plasmid expressing *S. oneidensis acnD*+. However, when growing on medium containing succinate and propionate (30 mM each), a much less stringent test for propionate utilization, both proteins with substitutions at C107 supported propionate catabolism (Figure 4C), albeit at a slower growth rate than the wild-type *prpF* allele, suggesting the presence of proteins with lower activity that could

have been caused by misfolding. In contrast, none of the variants with substitutions at residue K73 supported growth under either condition (Figure 4D) suggestive of proteins that either lacked catalytic activity or were misfolded. We note that when grown on medium containing both succinate and propionate, the final cell density of the cultures was approximately 1.3–1.5 A_{630} units, which was substantially higher than the final density of approximately 1.0 A_{630} unit when growing with propionate alone, suggesting that the strains containing the mutant alleles encoding *So*PrpF variants used propionate and succinate as carbon and energy sources.

To assess the residual level of isomerase activity associated with variant *So*PrpF variants, appropriate mutations were introduced into plasmid pPRP196 (*prpF*+ in pTEV4); proteins containing the C107A or K73A substitutions were not overproduced and could not be purified. The remaining variants were isolated, and their activity was quantified. None of the variants had any detectable activity. While this may be expected for substitutions of K73, the proposed catalytic residue, it is unclear why the C107S mutant had no activity, especially since the variants were active *in vivo*. More studies may need to be undertaken to fully understand this result.

Analysis of the tertiary and quaternary structures of the catalytically inactive SoPrpF^{K73E} variant. To further understand the roles of residues K73 and C107, we solved the three-dimensional crystal structure of Apo-*So*PrpF^{K73E}, which crystallized in space group P21 with cell dimensions a = 51.8 b = 103.4 c = 78.1 Å and contained two monomers per asymmetric unit. The structure was determined at 1.22 Å resolution by molecular replacement using apo-*So*PrpF (PDB ID: 2PVZ) as a search model (Table 3). The two monomers in the asymmetric unit are related by a non-crystallographic twofold axis. The two monomers are highly similar where the rms difference between 381 α-carbon atoms is 0.13 Å. Given the similarity between the two

monomers, all of the discussion of the structure of a single protein chain is based on that of subunit A.

*So*PrpF assembles to form a homodimer and buries 2400 Å^2 of surface area per monomer, which represents 15% of each monomer's total surface area (Figure 5A). While the *So*PrpF^{K73E} variant structure can be superposed closely with both the apo-PrpFWT and *trans*-aconitate bound structures with rms differences of 0.6 Å and 0.4 Å over 385 α-carbon respectively, an alignment of the N-terminal domains from residues 5–185 reveals that the C-terminal domain of the *So*PrpF^{K73E} structure rotates into a more open conformation. The C-terminal domain rotates ~4.4° between the extremes provided by the apo (closed) and aconitate (open) bound forms. This rotation occurs around the segments that connect the N-terminal and C-terminal domain (E180-N183 and I379-M380) although there are negligible conformational changes associated with these residues since they represent the fulcrum point for the rotation. As such these residues cannot be viewed as a flexible hinge.

A comparison of the three structures showed that the *trans*-aconitate bound structure was in the most open conformation, while the *So*PrpF^{K73E} structure was intermediate and the apo-*So*PrpFWT structure was in the most closed conformation (Figure 5B). This was not surprising as a glycerol molecule, derived from the cryoprotectant used in the structural determination, was bound in the active site of the apo-*So*PrpFWT structure. Glycerol is a smaller ligand than the true substrate and thus allows the C-terminal domain to move closer to the N-terminal domain since it is not impeded by the larger *trans*-aconitate molecule. Similarly, a malonate molecule derived from the crystallization solution was bound in the active site of the *So*PrpF^{K73E} structure. Because malonate is larger than glycerol and smaller than *trans*-aconitate, the C-terminal domain adopted an intermediary conformation. Thus, the structure of the apo-*So*PrpF^{K73E} variant protein provided

an estimate of the conformational freedom available to the N- and C-terminal domains of PrpF. Attempts to obtain substrate complexed *So*PrpF^{K73E} crystals by co-crystallization or soaking into apo crystals with either trans-aconitate or 2-methylcitrate were not successful. This difficulty in co-crystallization was not surprising since binding of *trans*-aconitate into the *So*PrpF^{K73E} active site would position a substrate carboxyl moiety within 3 Å of the new glutamate carboxyl group of *So*PrpF^{K73E} variant, thus creating an unfavorable interaction (Figure 6).

Phylogenetic analysis of PrpF homologues. To provide some perspective of the wide distribution of PrpF homologues in nature, we performed a limited phylogenetic analysis of 70 microbes containing a total of 86 PrpF homologues using the MUltiple Sequence Comparison by Log-Expectation (MUSCLE) software. For this purpose, one hundred sequences were selected and aligned (Figure 7). Proteins displaying the shortest distances from the root (clusters at the top and bottom of the tree, respectively) were mostly found in operons with genes related to propionate catabolism. Proteins that are more divergent were more likely to be found in alternate genetic contexts. In Figure 7, proteins from Gram-positive species are indicated with squares and proteins from fungi are indicated with circles. When an organism possesses multiple copies of the gene, the additional copies that are not found to be associated with propionate catabolic genes are marked with triangles. As can be noted in the tree, proteins found in organisms that are more distant evolutionarily from the Gamma-proteobacteria, and proteins not involved in propionate catabolism tend to be more distant in the phylogenetic tree as well.

Discussion

In this paper, we show that the PrpF protein of *S. oneidensis* (*So*PrpF) has isomerase activity that converts the product of the *So*AcnD reaction into 2-methyl-*cis*-aconitate, which can then be converted into 2-methylisocitrate by aconitase. The *So*AcnD product is likely 4-methyl-

cis-aconitate. On the basis of the analysis of the crystal structures of *So*PrpFWT and *So*PrpF^{K73E}, we propose that *So*AcnD dehydrates 2-methylcitrate into 4-methyl-*cis*-aconitate. We established the order of the 2-MCC in *Shewanella oneidensis* and, by inference, in other bacteria that contain *acnD* and *prpF* homologues in their *prp* operons. That is, in organisms that use the AcnD/PrpF enzymes instead of PrpD, 2-methylcitrate is dehydrated by AcnD to 4-methyl-*cis*-aconitate, which is then isomerized to 2-methyl-*cis*-aconitate by PrpF and rehydrated to 2-methylisocitrate by aconitases (Figure 1).

Rationale for the use of the AcnD/PrpF system. It is unclear why some bacteria such as *S. oneidensis* and *V. cholera* use AcnD and PrpF to synthesize 2-methyl-*cis*-aconitate, while others use a PrpD homologue. This work does not address this interesting question, but it does provide experimental data to support the assignment of a biochemical activity for the *So*PrpF enzyme.

We note that a previous report suggested that the formation of 2-methyl-*cis*-aconitate must occur via a unique *syn* elimination of water (*3*). To date, a detailed analysis of the mechanism of *So*AcnD or any of its homologues has not been reported. Thus, for the sake of this discussion, we will assume that the *So*AcnD protein, like most aconitases, is limited to *anti* β-eliminations (*39*). Therefore, like other aconitases, *So*AcnD would remove the *pro-R* proton from the oxaloacetate-derived carbon of citrate (C-4), not from the carbon derived from acetate (C-2) (*40*). In 2-methylcitrate, carbons 3, 4, and 5 are the carbons derived from oxaloacetate, and as such, carbons 3 and 4 should be where *So*AcnD dehydrates the substrate. The C2 of 2-methylcitrate is derived from propionate, and should not participate in the *So*AcnD-catalyzed dehydration reaction. An *anti* ß-elimination of the *pro-R* proton at C4 should result specifically in the formation of 4-methyl-*cis*-aconitate by *Se*AcnD.

Since *So*AcnD is likely to yield 4-methyl-*cis*-aconitate, the presence of an isomerase, like *So*PrpF, becomes necessary, as the accumulation of 4-methyl-*cis*-aconitate could inhibit aconitase. Isomerization of 4-methyl-*cis*-aconitate to 2-methyl-*cis*-aconitate then allows aconitase to convert the latter to 2-methylisocitrate, which is cleaved by the 2-methylisocitrate lyase yielding succinate and pyruvate (Figure 1). The greater activity of *Se*AcnB on enzymatically-derived 2-methyl-*cis*-aconitate when compared to *Se*AcnA suggests AcnB is the primary aconitase utilized in propionate catabolism, which is in agreement with our previously reported data that indicated *Se*AcnB was the primary aconitase involved in propionate catabolism (*15*).

Modifications to the 2-methylcitric acid cycle. The identification of *So*PrpF as a 4-methyl-*cis*-aconitate isomerase resolves the issue of why *So*AcnD and *So*PrpF proteins are required to restore propionate catabolism in a Δ*prpD* strain in *S. enterica* (*18*). Figure 1 reflects the findings of this work, in that it shows the sequence of reactions catalyzed by AcnD and PrpF.

Why AcnD, and not any of the other aconitases? It is unclear why any organism would dedicate an aconitase to the 2-MCC, given that several aconitases already exist in the cell. Studies of the conformation of mitochondrial aconitases suggest that the methyl group of 2-methylcitrate may sterically interfere with this compound entering the active site correctly (*41*). We suggest that the active site of *So*AcnD may be different enough to use 2-methylcitrate as a substrate.

Structure-function analysis of the *So*PrpF active site. In our previous work we suggested that active site of PrpF utilized a single catalytic lysine (Lys 73) to catalyze a proton extraction and a conformational rearrangement around the C2-C3 bond (*19*). Data presented here indicates that the reaction instead involves an allylic rearrangement and bond migration mechanism of an aconitate isomerase from *Pseudomonas putida* described by Klinman and Rose that may have been a PrpF homologue (*42*). However, using an allylic rearrangement mechanism

would require two catalytic residues and PrpF appears to only have one. The most likely candidate for a second catalytic residue would be C107, which is formally equivalent to the catalytic residue of diaminopimelate epimerase. A recent publication suggested that 4-methyl-*cis*-aconitate would bind in a manner that would allow the C107 to function as a second catalytic residue (*43*). Our mutational analysis sheds light on this possibility. If C107 were required for function, complete loss of enzyme activity in variants with substitutions in C107 would be expected. In Figure 4A, we present *in vivo* evidence in support for a critical role for C107. When propionate was used as the sole source of carbon and energy, the *So*PrpF$^{C107S/A}$ variants failed to support growth of a Δ*prpD* strain expressing *So*AcnD. Although *So*PrpF$^{C107S/A}$ variants retain activity (Figure 4C), it is insufficient to support growth with propionate as a carbon and energy source. Further work is needed to better understand the catalytic mechanism of *So*PrpF.

Significance of the broad distribution of PrpF homologues. The widespread distribution of genes encoding PrpF homologues in prokaryotes and eukaryotes, reveals the importance of isomerization in cell physiology. While the majority of the bacterial homologues appear to be part of operons along with other propionate utilization genes, many of them are present in operons encoding genes of unknown function. These prpF homologues suggest the possibility of double-bond isomerization being involved in many metabolic pathways.

In summary, we report data in support the following conclusions: i) PrpF isomerizes 4-methyl-*cis*-aconitate into 2-methyl-*cis*-aconitate; ii) the product of the AcnD reaction with 2-methylcitrate is most likely 4-methyl-*cis*-aconitate; iii) the aconitase proteins of *S. enterica* (*i.e.*, AcnA, AcnB) only rehydrate 2-methyl-*cis*-aconitate to 2-methylisocitrate; and iv) the proposed catalytic lysine (Lys 73) is absolutely required for activity of PrpF, and while substitutions of C107 do not abolish enzyme activity, PrpF^{C107} variants do not support growth with propionate.

Conclusions

Work reported in this paper advances our understanding of the 2-methylcitric acid cycle responsible for the conversion of the short-chain fatty acid propionate to pyruvate. The best-characterized sequence of reactions of the pathway involves the PrpD enzyme, which dehydrates 2-methylcitrate to 2-methyl-*cis*-aconitate. However, some microorganisms have replaced PrpD with two proteins, namely AcnD and PrpF, whose functions are not well understood. *In vivo* and in vitro evidence presented in this paper support to the conclusion that in an AcnD/PprF-dependent 2-methylcitric acid cycle, AcnD likely generates 4-methyl-*cis*-aconitate, which is isomerized by PrpF into 2-methyl-*cis*-aconitate. Additional work is needed to confirm the identity of the *So*AcnD product. Structural analysis of an inactive variant of PrpF provides insights into its mechanism of function. A better understanding of PrpF activity will be of value to other investigators researching the function of PrpF-like isomerases, which are widely distributed among prokaryotes.

References

1. Tabuchi, T., and Hara, S. (1974) Production of 2-Methylisocitric Acid from N-Paraffins by Mutants of Candida-Lipolytica, *Agr Biol Chem Tokyo 38*, 1105-1106.

2. Horswill, A. R., and Escalante-Semerena, J. C. (1999) *Salmonella typhimurium* LT2 catabolizes propionate via the 2-methylcitric acid cycle, *J Bacteriol 181*, 5615-5623.

3. Brock, M., Maerker, C., Schutz, A., Volker, U., and Buckel, W. (2002) Oxidation of propionate to pyruvate in *Escherichia coli* - Involvement of methylcitrate dehydratase and aconitase, *Eur J Biochem 269*, 6184-6194.

4. Textor, S., Wendisch, V. F., DeGraaf, A., Muller, U., Linder, M. I., Linder, D., and Buckel, W. (1997) Propionate oxidation in *Escherichia coli*: evidence for operation of a methylcitrate cycle in bacteria, *Arch Microbiol 168*, 428-436.

5. Munoz-Elias, E. J., Upton, A. M., Cherian, J., and McKinney, J. D. (2006) Role of the methylcitrate cycle in *Mycobacterium tuberculosis* metabolism, intracellular growth, and virulence, *Mol Microbiol 60*, 1109-1122.

6. Upton, A. M., and McKinney, J. D. (2007) Role of the methylcitrate cycle in propionate metabolism and detoxification in *Mycobacterium smegmatis*, *Microbiol-Sgm 153*, 3973-3982.

7. Savvi, S., Warner, D. F., Kana, B. D., McKinney, J. D., Mizrahi, V., and Dawes, S. S. (2008) Functional characterization of a vitamin B(12)-dependent methylmalonyl pathway in *Mycobacterium tuberculosis*: Implications for propionate metabolism during growth on fatty acids, *J Bacteriol 190*, 3886-3895.

8. Huser, A. T., Becker, A., Brune, I., Dondrup, M., Kalinowski, J., Plassmeier, J., Puhler, A., Wiegrabe, I., and Tauch, A. (2003) Development of a *Corynebacterium glutamicum* DNA microarray and validation by genome-wide expression profiling during growth with propionate as carbon source, *J Biotechnol 106*, 269-286.

9. Claes, W. A., Puhler, A., and Kalinowski, J. (2002) Identification of two *prpDBC* gene clusters in *Corynebacterium glutamicum* and their involvement in propionate degradation via the 2-methylcitrate cycle, *J Bacteriol 184*, 2728-2739.

10. Bramer, C. O., and Steinbuchel, A. (2001) The methylcitric acid pathway in *Ralstonia eutropha*: new genes identified involved in propionate metabolism, *Microbiol-Sgm 147*, 2203-2214.

11. Bramer, C. O., Silva, L. F., Gomez, J. G. C., Priefert, H., and Steinbuchel, A. (2002) Identification of the 2-methylcitrate pathway involved in the catabolism of propionate in the polyhydroxyalkanoate-producing strain *Burkholderia sacchari* IPT101(T) and analysis of a mutant accumulating a copolyester with higher 3-hydroxyvalerate content, *Appl Environ Microb 68*, 271-279.

12. Horswill, A. R., and EscalanteSemerena, J. C. (1997) Propionate catabolism in *Salmonella typhimurium* LT2: Two divergently transcribed units comprise the *prp* locus at 8.5 centisomes, *prpR* encodes a member of the sigma-54 family of activators, and the *prpBCDE* genes constitute an operon, *J Bacteriol 179*, 928-940.

13. Grimek, T. L., Holden, H., Rayment, I., and Escalante-Semerena, J. C. (2003) Residues C123 and D58 of the 2-methylisocitrate lyase (PrpB) enzyme of *Salmonella enterica* are essential for catalysis, *J Bacteriol 185*, 4837-4843.

14. Grimm, C., Evers, A., Brock, M., Maerker, C., Klebe, G., Buckel, W., and Reuter, K. (2003) Crystal structure of 2-methylisocitrate lyase (PrpB) from *Escherichia coli* and modelling of its ligand bound active centre, *J Mol Biol 328*, 609-621.

15. Horswill, A. R., and Escalante-Semerena, J. C. (2001) *In vitro* conversion of propionate to pyruvate by *Salmonella enterica* enzymes: 2-methylcitrate dehydratase (PrpD) and aconitase enzymes catalyze the conversion of 2-methylcitrate to 2-methylisocitrate, *Biochemistry-Us 40*, 4703-4713.

16. Horswill, A. R., and Escalante-Semerena, J. C. (1999) The *prpE* gene of *Salmonella typhimurium* LT2 encodes propionyl-CoA synthetase, *Microbiol-Uk 145*, 1381-1388.

17. Horswill, A. R., and Escalante-Semerena, J. C. (2002) Characterization of the propionyl-CoA synthetase (PrpE) enzyme of *Salmonella enterica*: Residue Lys592 is required for propionyl-AMP synthesis, *Biochemistry-Us 41*, 2379-2387.

18. Grimek, T. L., and Escalante-Semerena, J. C. (2004) The acnD genes of *Shewenella oneidensis* and *Vibrio cholerae* encode a new Fe/S-dependent 2-methylcitrate dehydratase enzyme that requires *prpF* function *in vivo*, *J Bacteriol 186*, 454-462.

19. Garvey, G. S., Rocco, C. J., Escalante-Semerena, J. C., and Rayment, I. (2007) The three-dimensional crystal structure of the PrpF protein of *Shewanella oneidensis* complexed with *trans*-aconitate: Insights into its biological function, *Protein Sci 16*, 1274-1284.

20. Bertani, G. (1951) Studies on Lysogenesis .1. The Mode of Phage Liberation by Lysogenic *Escherichia coli*, *J Bacteriol 62*, 293-300.

21. Bertani, G. (2004) Lysogeny at mid-twentieth century: P1, P2, and other experimental, systems, *J Bacteriol 186*, 595-600.

22. Berkowit.D, Hushon, J. M., Whitfiel.Hj, Roth, J., and Ames, B. N. (1968) Procedure for Identifying Nonsense Mutations, *J Bacteriol 96*, 215-&.

23. Rocco, C. J., Dennison, K. L., Klenchin, V. A., Rayment, I., and Escalante-Semerena, J. C. (2008) Construction and use of new cloning vectors for the rapid isolation of recombinant proteins from *Escherichia coli*, *Plasmid 59*, 231-237.

24. Otwinowski, Z., and Minor, W. (1997) Processing of X-ray diffraction data collected in oscillation mode, *Method Enzymol 276*, 307-326.

25. 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 Appl Crystallogr 40*, 658-674.

26. Emsley, P., and Cowtan, K. (2004) Coot: model-building tools for molecular graphics, *Acta Crystallogr D 60*, 2126-2132.

27. Murshudov, G. N., Vagin, A. A., and Dodson, E. J. (1997) Refinement of macromolecular structures by the maximum-likelihood method, *Acta Crystallogr D 53*, 240-255.

28. Adams, P. D., Afonine, P. V., Bunkoczi, G., Chen, V. B., Davis, I. W., Echols, N., Headd, J. J., Hung, L. W., Kapral, G. J., Grosse-Kunstleve, R. W., McCoy, A. J., Moriarty, N. W., Oeffner, R., Read, R. J., Richardson, D. C., Richardson, J. S., Terwilliger, T. C., and Zwart, P. H. (2010) PHENIX: a comprehensive Python-based system for macromolecular structure solution, *Acta Crystallogr D 66*, 213-221.

29. Kennedy, M. C., Spoto, G., Emptage, M. H., and Beinert, H. (1988) The Active-Site Sulfhydryl of Aconitase Is Not Required for Catalytic Activity, *J Biol Chem 263*, 8190-8193.

30. Kennedy, M. C., and Beinert, H. (1988) The State of Cluster Sh and S2- of Aconitase during Cluster Interconversions and Removal. A Convenient Preparation of Apoenzyme, *J Biol Chem 263*, 8194-8198.

31. Gunsalus, R. P., Tandon, S. M., and Wolfe, R. S. (1980) A procedure for anaerobic column chromatography employing an anaerobic Freter-type chamber, *Anal Biochem 101*, 327-331.

32. Balch, W. E., and Wolfe, R. S. (1976) New approach to the cultivation of methanogenic bacteria: 2-mercaptoethanesulfonic acid (HS-CoM)-dependent growth of *Methanobacterium ruminantium* in a pressureized atmosphere, *Appl Environ Microbiol 32*, 781-791.

33. Kennedy, M. C., Emptage, M. H., Dreyer, J. L., and Beinert, H. (1983) The Role of Iron in the Activation-Inactivation of Aconitase, *J Biol Chem 258*, 1098-1105.

34. Schloss, J. V., Emptage, M. H., and Cleland, W. W. (1984) Ph Profiles and Isotope Effects for Aconitases from *Saccharomycopsis lipolytica*, Beef Heart, and Beef Liver - Alpha-methyl-*cis*-aconitate and *Threo*-Ds-alpha-methylisocitrate as Substrates, *Biochemistry-Us 23*, 4572-4580.

35. Aoki, H., Uchiyama, H., Umetsu, H., and Tabuchi, T. (1995) Isolation of 2-Methylisocitrate Dehydratase, a New Enzyme Serving in the Methylcitric Acid Cycle for Propionate Metabolism, from *Yallowia lipolytica*, *Biosci Biotech Bioch 59*, 1825-1828.

36. Cirilli, M., Zheng, R. J., Scapin, G., and Blanchard, J. S. (1998) Structural symmetry: The three-dimensional structure of *Haemophilus influenzae* diaminopimelate epimerase, *Biochemistry-Us 37*, 16452-16458.

37. Koo, H. M., and Kim, Y. S. (2000) Identification of active-site residues in *Bradyrhizobium japonicum* malonyl-coenzyme A synthetase, *Arch Biochem Biophys 378*, 167-174.

38. Blankenfeldt, W., Kuzin, A. P., Skarina, T., Korniyenko, Y., Tong, L., Bayer, P., Janning, P., Thomashow, L. S., and Mavrod, D. V. (2004) Structure and function of the phenazine biosynthetic protein PhzF from *Pseudomonas fluorescens*, *P Natl Acad Sci USA 101*, 16431-16436.

39. Hanson, K. R., and Rose, I. A. (1963) The Absolute Stereochemical Course of Citric Acid Biosynthesis, *Proc Natl Acad Sci U S A 50*, 981-988.

40. Gawron, O., Glaid, A. J., Fondy, T. P., and Bechtold, M. M. (1961) Stereochemistry of the succinic dehydrogenase system, *Nature 189*, 1004-1005.

41. Lauble, H., and Stout, C. D. (1995) Steric and conformational features of the aconitase mechanism, *Proteins 22*, 1-11.

42. Klinman, J. P., and Rose, I. A. (1971) Mechanism of the aconitate isomerase reaction, *Biochemistry-Us 10*, 2259-2266.

43. Velarde, M., Macieira, S., Hilberg, M., Broker, G., Tu, S. M., Golding, B. T., Pierik, A. J., Buckel, W., and Messerschmidt, A. (2009) Crystal structure and putative mechanism of 3-methylitaconate-delta-isomerase from *Eubacterium barkeri*, *J Mol Biol 391*, 609-620.

44. Krissinel, E., and Henrick, K. (2004) Secondary-structure matching (SSM), a new tool for fast protein structure alignment in three dimensions, *Acta Crystallogr D 60*, 2256-2268.

Tables

Table 1. Strains and plasmids used in this study.

Strain	Relevant genotype	Source
***E. coli* strains**		
BL21(λDE3)	F⁻ *ompT hsdS*$_B$ (r_b^- m_b^+) *dcm gal* I(DE3)	New England Biolabs
DH5α/F'	F'/*endA1 hsdR17* (r_k^- m_k^+) *supE44 thi-1 recA1 gyrA* (Nal^r) *relA1* Δ(*lacZYA-argF*)*U169 deoR* [f80d*lac*D (*lacZ*)*M15*]	New England Biolabs
XL10-Gold® Ultracompetent cells	TetR Δ(*mcrA*)183Δ(*mcrCB-hsdSMR-mrr*)173 *endA1 supE44 thi-1 recA1 gyrA96 relA1 lac* Hte [F' *proAB lacIqZΔM15* Tn*10* (TetR) Amy CamR]a	Stratagene
***S. enterica* strains**		
TR6583	*metE205 ara-9*	K. Sanderson via J. Roth
Derivatives of TR6583		
JE8250	*prpD126*::*cat*$^+$	This work
JE8256	JE8250 / pPRP138	This work
JE8429	JE8256 / pPRP153	This work
JE9373	JE8256 / pPRP215 *prpF255*	This work
JE9374	JE8256 / pPRP216 *prpF256*	This work
JE9592	JE8256 / pPRP218 *prpF257*	This work
JE9593	JE8256 / pPRP219 *prpF258*	This work
JE11597	JE8256 / pPRP223 *prpF259*	This work
Plasmids		
pPRP138	*S. oneidensis acnD*$^+$ in pBAD30 *bla*$^+$	[18]
pPRP153	*S. oneidensis prpF*$^+$ in pBAD18Kan *kan*$^+$	[18]
pPRP195	*S. oneidensis acnD*$^+$ cloned into pET-15b *bla*$^+$	This work
pPRP196	*S. oneidensis prpF*$^+$ cloned into pTEV4 *bla*$^+$	[19]
pPRP205	*S. oneidensis acnD*$^+$ cloned into pTEV5	This work
pPRP215	*S. oneidensis prpF*$^+$(*prpF255* C107A) in pBAD18Kan *kan*$^+$	This work
pPRP216	*S. oneidensis prpF*$^+$(*prpF256*; encodes $PrpF^{C107S}$) cloned into pBAD18-Kan *kan*$^+$	This work
pPRP217	*S. oneidensis prpF*$^+$(*prpF255*; encodes $PrpF^{C107A}$) cloned into pTEV4 *bla*$^+$	This work
pPRP218	*S. oneidensis prpF*$^+$(*prpF257*; encodes $PrpF^{K73A}$) cloned into pBAD18-Kan *kan*$^+$	This work
pPRP219	*S. oneidensis prpF*$^+$(*prpF258*; encodes $PrpF^{K73E}$) cloned into pBAD18-Kan *kan*$^+$	This work
pPRP220	*S. oneidensis prpF*$^+$(*prpF257*; encodes $PrpF^{K73A}$) cloned into pTEV4 *bla*$^+$	This work
pPRP221	*S. oneidensis prpF*$^+$(*prpF258*; encodes $PrpF^{K73E}$) cloned into pTEV4 *bla*$^+$	This work
pPRP223	*S. oneidensis prpF*$^+$(*prpF259*; encodes $PrpF^{K73M}$) cloned into pBAD18-Kan *kan*$^+$	This work
pET-15b	Cloning vector pBR origin of replication *bla*$^+$; fuses a thrombin-cleavable H_6 tag to the N-terminus of the protein of interest	Novagen
pTEV4	Cloning vector; F1 origin of replication *bla*$^+$; fuses a TEV-cleavable H_6 tag to the N-terminus of the protein of interest	[24]
pTEV5	Cloning vector; F1 origin of replication *bla*$^+$; fuses a TEV-cleavable H_6 tag to the N-terminus of the protein of interest	[24]
pBAD30	Cloning vector; pACYC184 origin of replication, *bla*$^+$; expression of the gene of interest under the control of the arabinose-inducible P_{araBAD} promoter	[46]
pBAD18-Kan	Cloning vector; pBR origin of replication, *kan*$^+$; expression of the gene of interest under the control of the arabinose-inducible P_{araBAD} promoter	[46]

Table 2. List of plasmids encoding PrpF variants used in this study.

Plasmid	Allele #	Primers 5'– 3'	Template	Protein Encoded
pPRP215	*prpF255*	GTGGATTGGAGTGGTAACGCGGGTAACTTAACAGCCGCCGGCGGCTGTTAAGTTACCCGCGTTACCACTCCAATCCAC	pPRP153	C107A
pPRP216	*prpF256*	GGATTGGAGTGGTAACAGCGGTAACTTAACAGCCGGCTGTTAAGTTACCGCTGTTACCACTCCAATCC	pPRP153	C107S
pPRP217	*prpF256*	GGATTGGAGTGGTAACAGCGGTAACTTAACAGCCGGCTGTTAAGTTACCGCTGTTACCACTCCAATCC	pPRP196	C107S
pPRP218	*prpF257*	CAACTTCTAGCACCAGCGCGACGGTTATACTGTCACGTGACAGTATAACCGTCGCGCTGGTGCTAGAAGTTG	pPRP153	K73A
pPRP219	*prpF258*	GCAACTTCTAGCACCAGCGAAACGGTTATACTGTCACACGTGTGACAGTATAACCGTTTCGCTGGTGCTAGAAGTTGC	pPRP153	K73E
pPRP220	*prpF257*	CAACTTCTAGCACCAGCGCGACGGTTATACTGTCACGTGACAGTATAACCGTCGCGCTGGTGCTAGAAGTTG	pPRP196	K73A
pPRP221	*prpF258*	GCAACTTCTAGCACCAGCGAAACGGTTATACTGTCACACGTGTGACAGTATAACCGTTTCGCTGGTGCTAGAAGTTGC	pPRP196	K73E
pPRP223	*prpF259*	CAACTTCTAGCACCAGCATGACGGTTATACTGTCACGTGACAGTATAACCGTCATGCTGGTGCTAGAAGTTG	pPRP153	K73M
pPRP225	*prpF255*	GTGGATTGGAGTGGTAACGCGGGTAACTTAACAGCCGCCGGCGGCTGTTAAGTTACCCGCGTTACCACTCCAATCCAC	pPRP196	C107A
pPRP226	*prpF259*	CAACTTCTAGCACCAGCATGACGGTTATACTGTCACGTGACAGTATAACCGTCATGCTGGTGCTAGAAGTTG	pPRP196	K73M

Table 3. Data collection and refinement statistics.

Data collection	**PrpF^{K73E}**
Space group	$P2_1$
Unit-cell	$a = 51.8$
parameters (Å)	$b = 103.4$
	$c = 78.1$
	$\beta = 104.47^{\circ}$
Wavelength	0.979
resolution range (Å)	50–1.22(1.24–1.22)[a]
reflections: measured	2973055
reflections: unique	227930
Redundancy	13.0 (6.1)
Completeness (%)	96.5 (84.5)
average I/	49.5 (2.4)
Rsym [b] (%)	5.3 (67.1)
*R*work [c] (%)	16.4 (23.2)
*R*free (%)	18.4 (28.6)
no. protein atoms	5907
no. water molecules	1013
Wilson B-value (Å^2)	14.5
Average B factors (Å^2)	
PrpF monomer A	17.2 for 2956
PrpF monomer B	19.1 for 2923
Ligand	16.5 for 28
Solvent	29.2 for 1013
Ramachandran (%)	
Most favored	98.0
Additionally	2.0
allowed	
Generously	0.0
allowed	
Disallowed	0.0
rms deviations	
Bond Lengths (Å)	0.006
Bond angles (Å)	1.066
Chiral	

[a]Data in parentheses represent highest resolution shell.

[b]*R*sym = Σ/I(hkl)–I/ / Σ/I(hkl)/, where the average intensity *I* is taken over all symmetry equivalent measurements and $I_{(hkl)}$ is the measured intensity for a given reflection.

[c]*R*factor = Σ/F(obs)–F(calc)/ / Σ/F(obs)/, where *R*work refers to the *R*factor for the data utilized in the refinement and *R*free refers to the *R*factor for 5% of the data that were excluded from the refinement.

Table 4. Kinetics of AcnA and AcnB with *cis*-aconitate and 2-methyl-*cis*-aconitate.

Isomer/Enzyme used	K_m (μM)	V_{max} (μM min^{-1})	k_{cat} (min^{-1})	k_{cat} / K_m (min^{-1} $μM^{-1}$)
***cis*-aconitate**				
AcnA	22 ± 2	86 ± 2	3739 ± 87	170
AcnB	17 ± 3	21 ± 1	2091 ± 100	123
2-methyl-*cis*-aconitate				
AcnA	210 ± 35	50 ± 3	708 ± 42	3.4
AcnB	208 ± 39	10 ± 1	396 ± 40	1.9

All kinetic reactions were performed in triplicate and individual kinetic curves were also repeated in triplicate. Reported values represent median values from all three experiments. No value differed more than 15% from the median value.

Figures

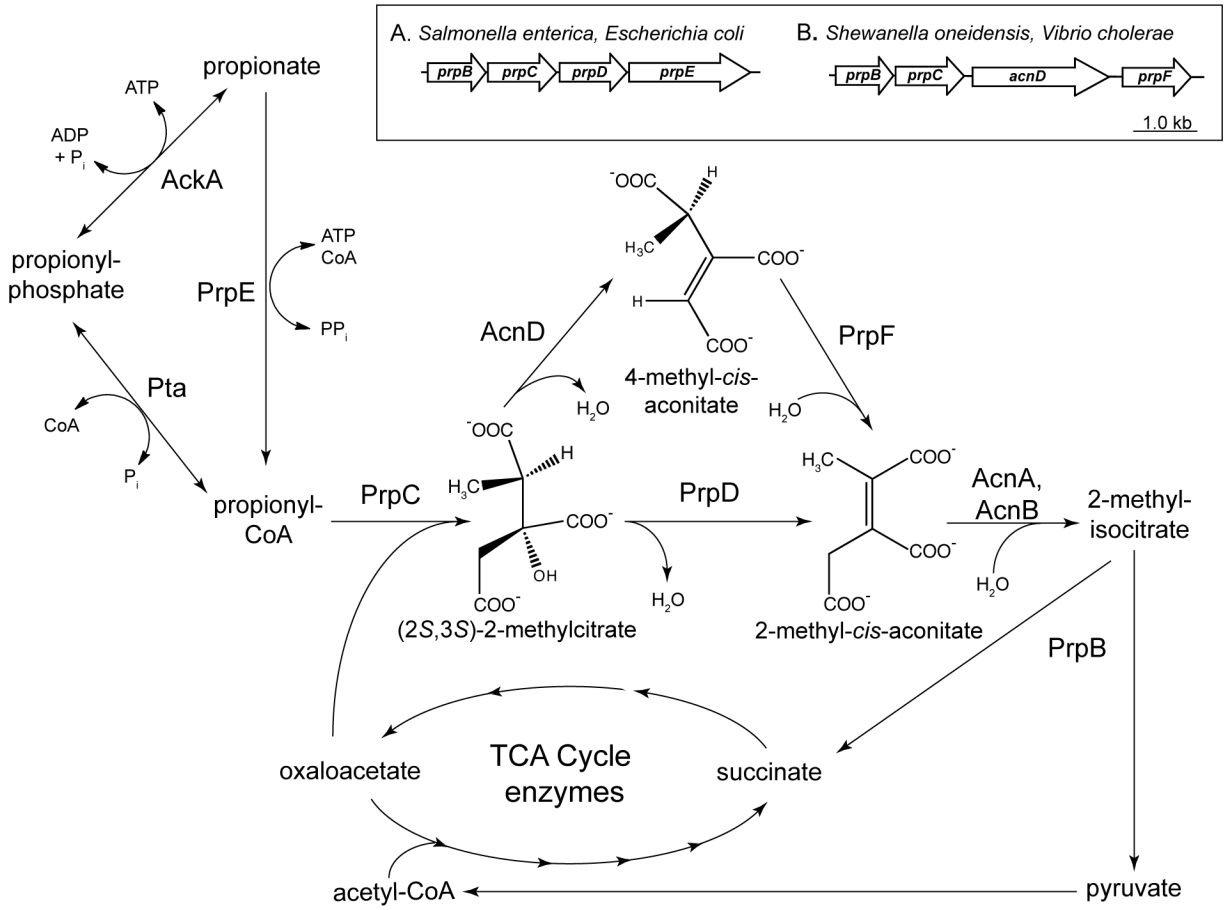


Figure 1. The 2-methylcitric acid cycle. In this metabolic pathway propionate is oxidized to pyruvate with succinate as byproduct. Inset box–difference in operon structure of *Salmonella enterica* and *Escherichia coli* (Inset box A) compared to *Shewanella oneidensis* and *Vibrio cholera* (Inset box B). AckA–acetate kinase; Pta–phosphotransacetylase; AcnA/B–aconitase A, aconitase B.

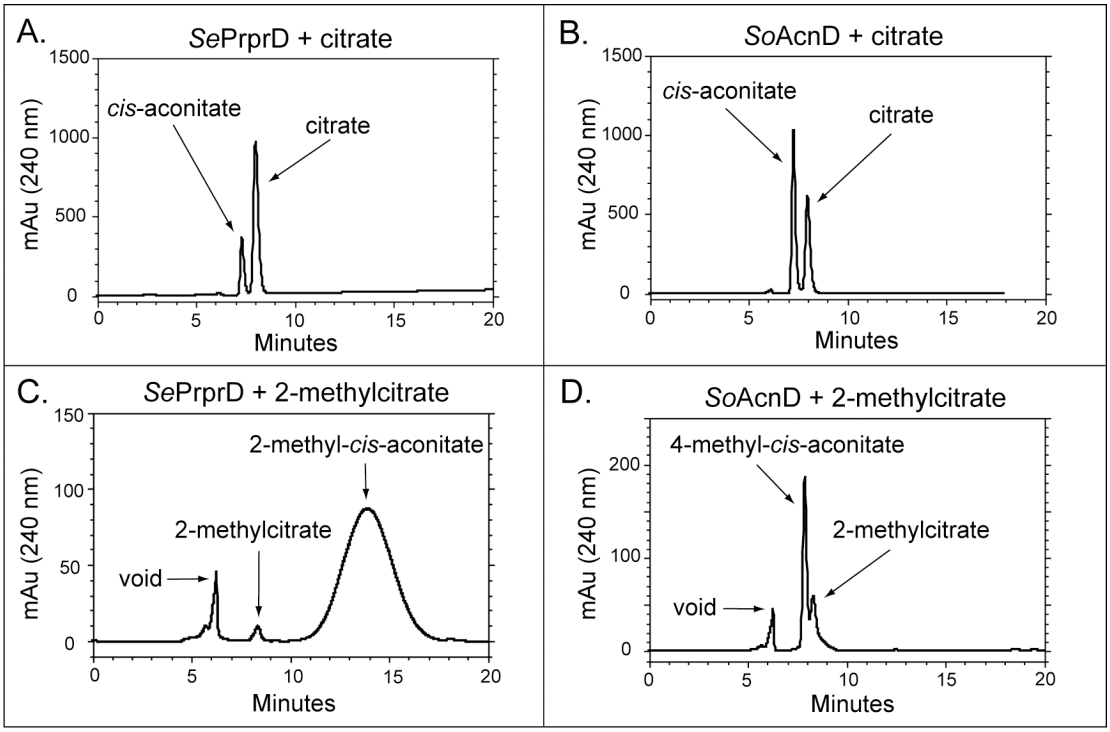


Figure 2. HPLC analysis of PrpD and AcnD reaction products. Reactions products when *Se*PrpD used citrate (panel A) or 2-methylcitrate (panel B) as substrate. Reaction products when *So*AcnD used citrate (panel B) or 2-methylcitrate (panel D) as substrate. Reaction mixtures were resolved by HPLC using a BioRad Aminex® HPX-87H Organic Acids Analysis column developed isocratically with 5 mM H_2SO_4 as the mobile phase. Elution was monitored at 240 nm.

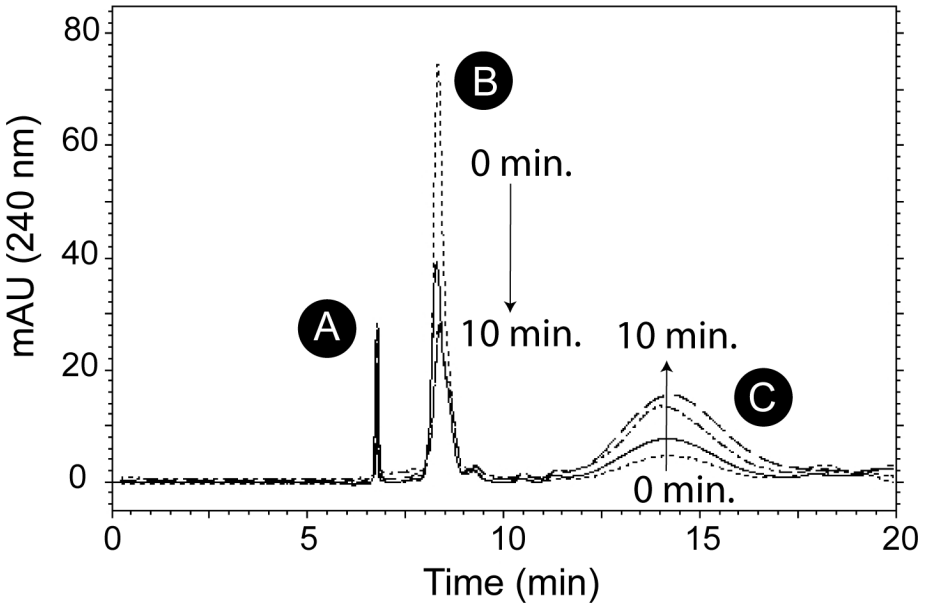


Figure 3. *So*PrpF isomerizes the *So*AcnD product into 2-methyl-*cis*-aconitate. Reactivated *So*AcnD was incubated with 2-MC, once *So*AcnD was removed from the reaction, PrpF was added (100 μg) and samples were removed at the indicated times stopped by the addition of H_2SO_4 to 5 mM. Reactions were resolved by HPLC using an Aminex HPX-87H Organic Acid column equilibrated and isocratically developed with 5 mM H_2SO_4; reaction products were monitored at 240 nm. A. void; B. 4-methyl-*cis*-aconitate C. 2-methyl-*cis*-aconitate.

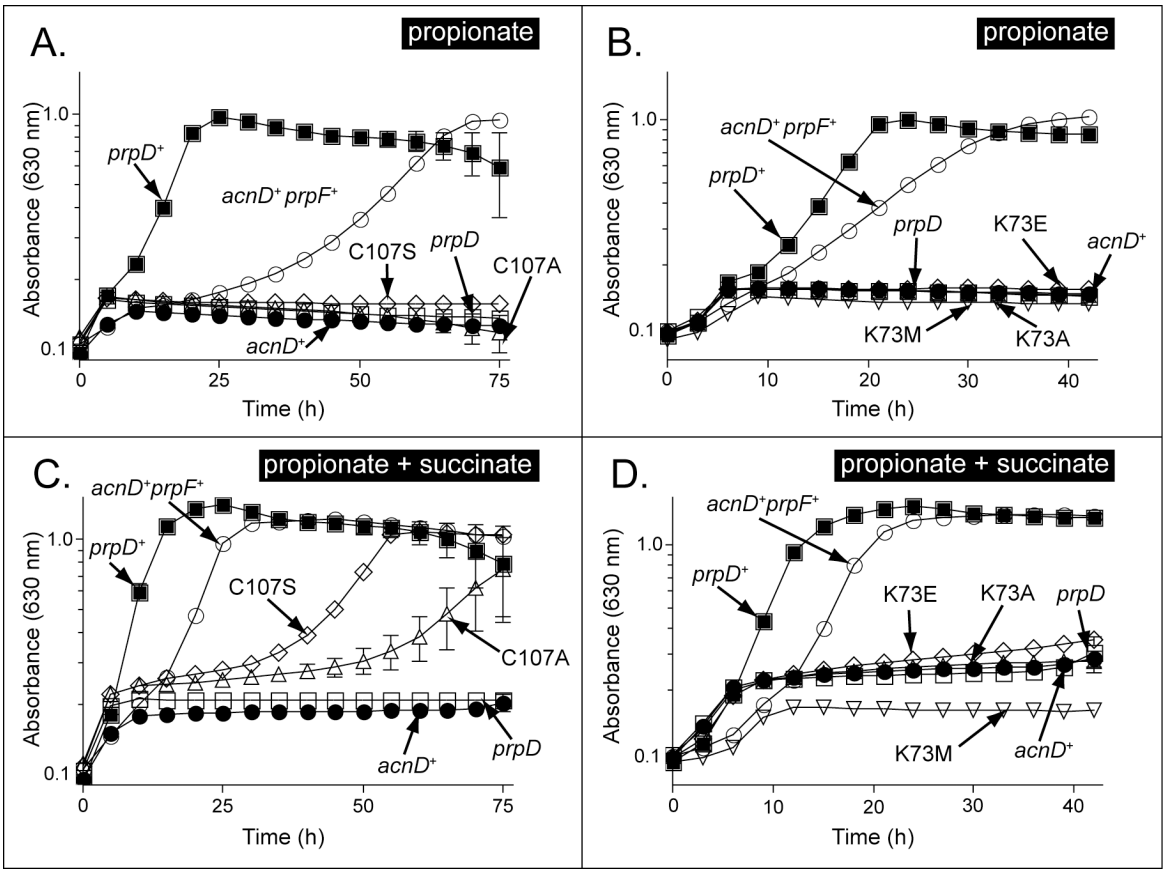


Figure 4. Growth behavior analysis of strains synthesizing variants of PrpF. Growth curves were performed in NCE minimal medium supplemented with propionate (30 mM) (Panels A, B) or propionate + succinate (30 mM ea.) (Panels C, D). Plasmids encoding PrpF variants with substitutions at position C107 failed to restore growth of strains JE9373 (*prpD* / pAcnDWT pPrpF^{C107A}) or JE9374 (prpD pAcnDWT pPrpF^{C107S}) with propionate (Panel A: PrpF^{C107S}, open diamonds; PrpF^{C107A}, open triangles, respectively), but did restore growth with propionate + succinate, albeit at a slower rate (panel C, diamonds, triangles, respectively). In contrast, substitutions at position K73 failed to compensate for the absence of PrpD on either propionate (panel B: PrpF^{K73A}, open triangles; PrpF^{K73E}, open diamonds, PrpF^{K73M}, open inverted triangles), or succinate + propionate (Panel D: PrpF^{K73A}, open triangles; PrpF^{K73E}, open diamonds; PrpF^{K73M}, open inverted triangles).

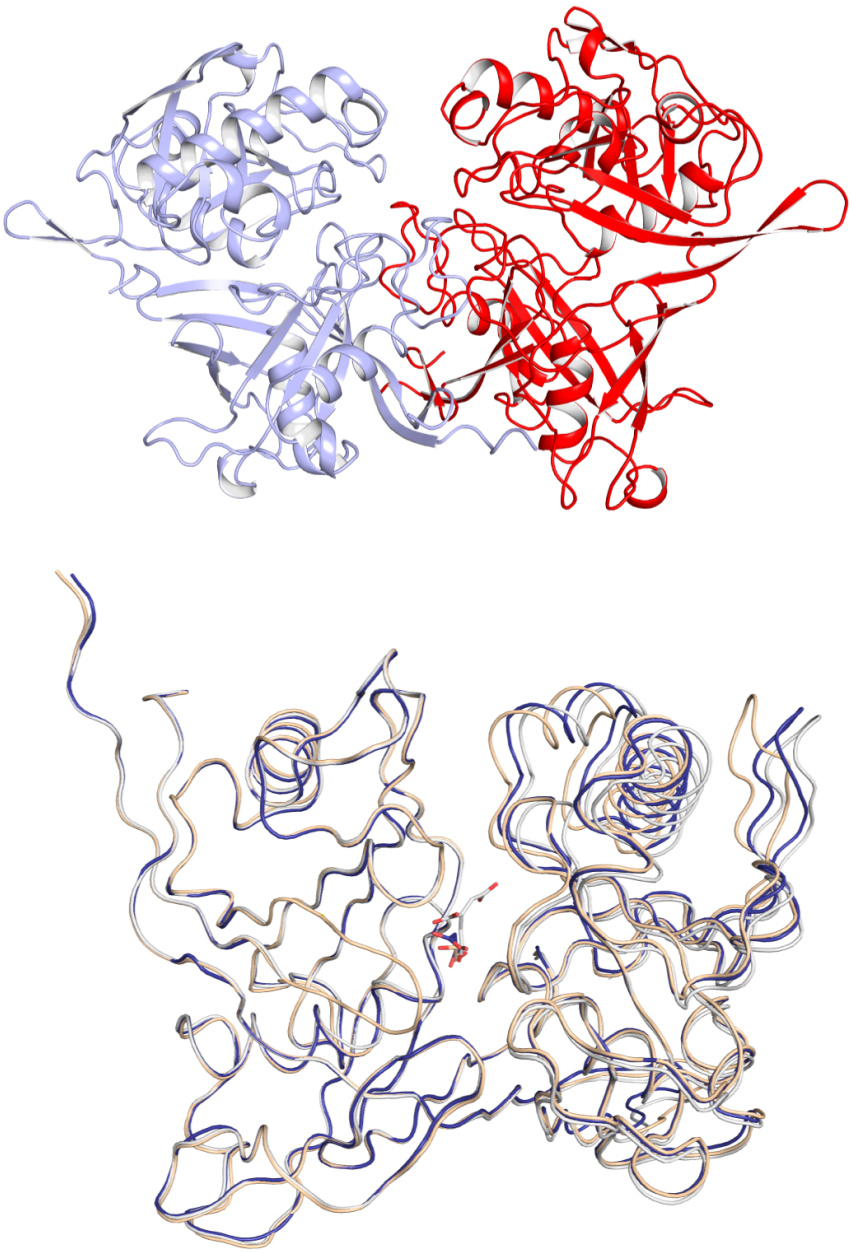


Figure 5. Crystal structure of *So*PrpF^{K73E} at 2.35 Å resolution. A. Ribbon representation of the *So*PrpF^{K73E} homodimer (light blue = monomer A, red = monomer B). B. Overlay of *So*PrpF^{K73E} (blue), *So*PrpF in complex with *trans*-aconitate (white, RCSB accession number 2PW0) and apo-*So*PrpFWT (tan, RCSB accession number 2PVZ). The superposition was performed by aligning residues 1–185 with the program Superpose (*44*). Figures 5 and 6 were prepared with the program Pymol; DeLano Scientific LLC, Palo Alto, CA.

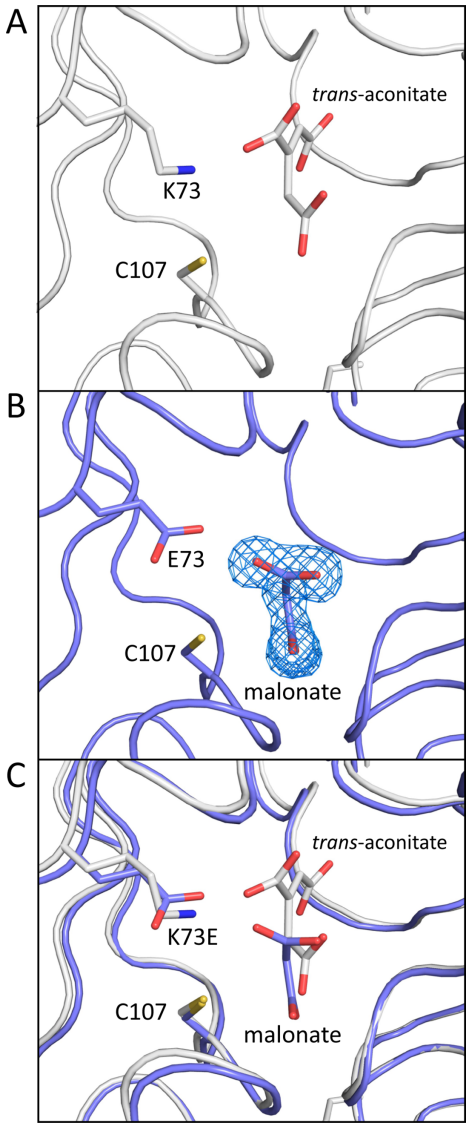


Figure 6. Active site comparisons of *So*PrpFWT and *So*PrpF^{K73E}. A. Detailed view of the catalytic active site residues, which contact *trans*-aconitate in the *So*PrpFWT protein determined in the presence of *trans*-aconitate (RCSB accession number 2pw0). B. Detailed view of the residues lining the active site in the *So*PrpF^{K73E} variant with malonate bound. The electron density map was calculated with coefficients of the form *Fo*−*Fc* where the ligand was omitted from the final phase calculation refinement and contoured to 4.5 σ. C. The superposition of active site residues with *trans*-aconitate and malonate included for reference.

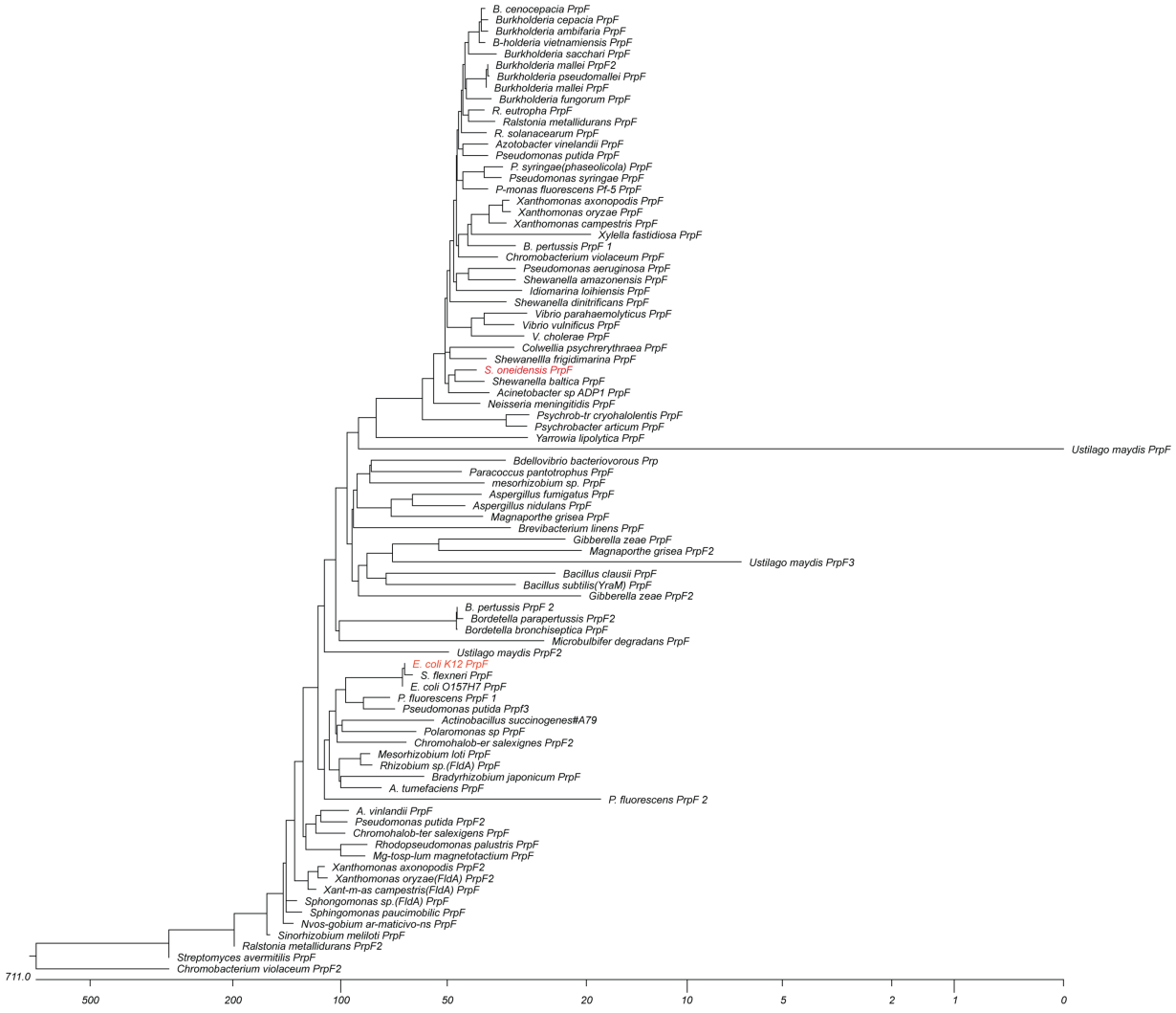


Figure 7. Phylogenetic tree of selected PrpF homologues in nature. Phylogenetic tree of selected PrpF homologues. Eighty-six sequences were selected and aligned using ClustalW software. Proteins towards the top of the tree are mostly found in operons with genes related to propionate catabolism. Proteins that are more divergent are more likely to be found in alternate genetic contexts. The PrpF protein of *Shewanella oneidensis* and the *E. coli* K12 PrpF homologue are highlighted in red. Distance is shown in substitutions per 100 residues.

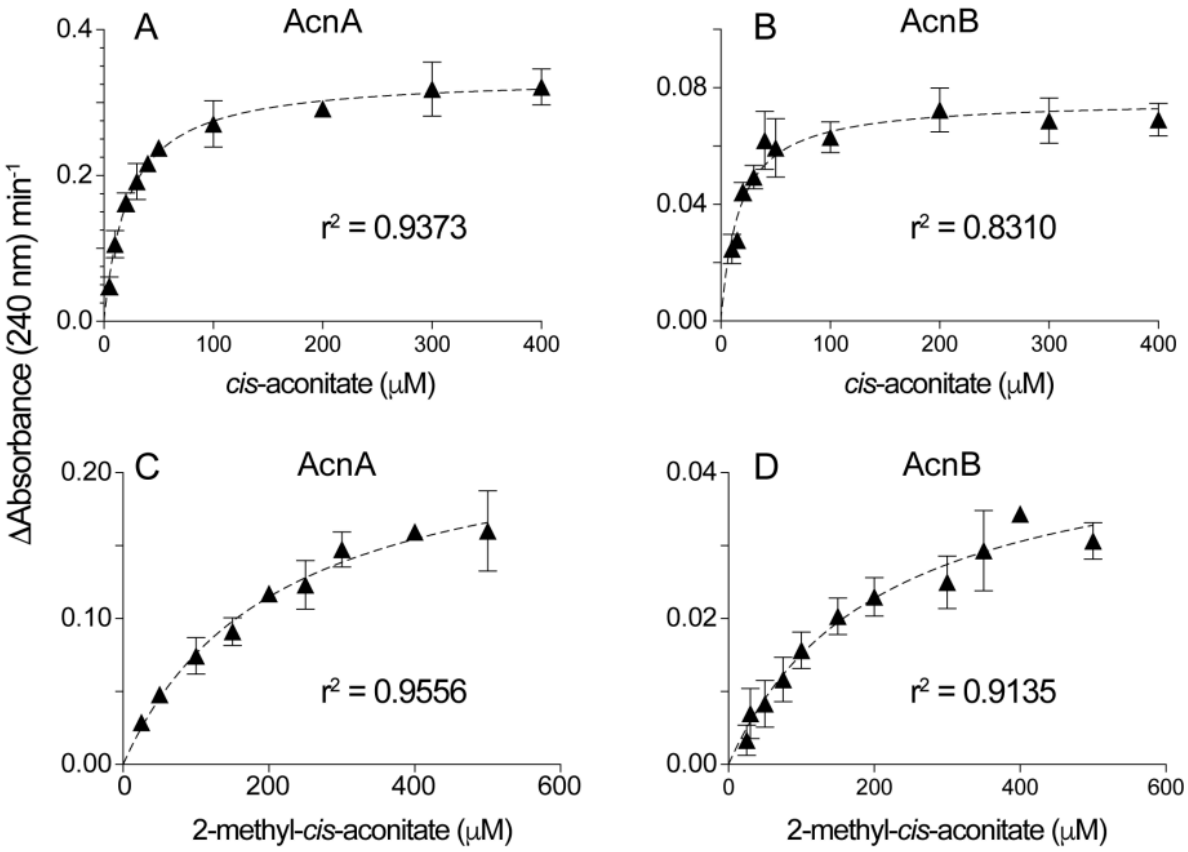


Figure S1. Kinetics of aconitase product formation. Reaction mixtures contained aconitase A (AcnA) and aconitase B (AcnB) with *cis*-aconitate (Panel A: AcnA; Panel B: AcnB) or 2-methyl-*cis*-aconitate (Panel C: AcnA; Panel D: AcnB) as substrate. Product formation (either isocitrate from *cis*-aconitate; 2-methylisocitrate from 2-methy-*cis*-aconitate) was monitored as a decrease in absorbance at 240 nm over time (min; y-axes) as a function of substrate concentration (μM; x-axes). The increase in the rate at which A_{240} decreased is what is plotted. Detailed assay conditions are described under *Materials and Methods*.

Chapter 6:

Significance and Future Directions

This work was performed in collaboration with Dr. Ivan Rayment.

Significance

Fusarium head blight (FHB) and *Fusarium* ear rot are devastating plant diseases that affect small grain cereals and maize on a global scale (*1, 2*). Trichothecene mycotoxins produced by *Fusarium* fungi are virulence factors and accumulate in the grain of infected plants (*3-5*). They are potent inhibitors of eukaryotic protein synthesis, posing a significant threat to both animal and human consumers (*6, 7*).

To date, one of the most promising approaches to combating *Fusarium graminearum*, one of the most common *Fusarium* species responsible for FHB in the United States, has been the creation of transgenic crops that constitutively overexpress trichothecene UDP-glucosyltransferases (*8-10*). This resistance is the result of glycosylation of DON, the primary trichothecene produced by *Fusarium graminearum*, which reduces its toxicity (*11, 12*). Although there has been some success at creating wheat crops with resistance to *Fusarium graminearum*, the UGTs responsible for this resistance do not demonstrate broad specificity to the diverse group of over 200 trichothecene produced by *Fusarium* species (*11*). Consequently, these enzymes likely cannot provide resistance against other *Fusarium* species such as *Fusarium sporotrichioides*, which relies on T-2 toxin, among others, as a virulence factor. Identifying or engineering UGTs capable of glycosylating a broad range of trichothecenes including T-2 toxin would solve this issue.

Furthermore, identifying UGTs capable of glycosylating trichothecenes has proven difficult, as most UGTs identified by homology searches using known trichothecene UGTs as the search model do not demonstrate activity with trichothecene substrates (*12*). And identification of UGTs which are expressed in plants in response to DON treatment or *Fusarium* infection has also been met with limited success (*12*). Thus, detailed structural knowledge that establishes the basis

for trichothecene UGT specificity and substrate binding is needed to help identify and engineer UGTs with diverse and broad trichothecene specificity.

The work presented in this thesis represents the first structural information on trichothecene UGT specificity and mechanism. Prior to this work, the location and orientation of trichothecene substrates within UGT active sites was not known. The seven crystal structures of the rice UGT Os79 in complex with various trichothecene substrates and products that are presented here provide detailed information on substrate binding. This structural information, along with extensive kinetic analysis of Os79 variants, allowed us to identify residues that play important roles in trichothecene binding and UGT specificity. This work may help researchers identify UGTs capable of glycosylating trichothecenes from primary sequences and homology models.

In addition to structural information, we report substitutions that alter the substrate specificity of Os79, establishing the first step in the ability to predict which trichothecenes a given UGT will accept as substrates. We used this information to broaden the specificity of Os79, producing an enzyme capable of glycosylating all the trichothecenes that we tested. This enzyme has great potential in the creation of transgenic crops with resistance to a wide range of *Fusarium* species, where current transgenic crops have only demonstrated resistance to *Fusarium graminearum.*

In addition to providing structural information on UGT specificity, we also identified Thr 291 as a residue that plays a critical role in Os79 catalysis. We present structural comparisons with other plant UGTs which reveal that the hydroxyl of Thr 291 is conserved in all plant UGTs, highlighting the broad importance of this residue, which has previously not been reported as involved in plant UGT catalysis.

Although the establishment of the structural basis for specificity in trichothecene UGTs and the development of a UGT with broad trichothecene specificity is an enormous step forward in combatting FHB, the development of an enzyme capable of permanently detoxifying trichothecenes by breaking the epoxide bond is still desirable as a means of treating already infected and harvested grain. To facilitate the discovery or development of such an enzyme, we developed a selection protocol that is capable of selecting for a trichothecene resistance gene from a library of over 1 million variants. This method has enormous potential for screening large libraries of potential trichothecene resistance enzymes. Furthermore, the construction of a yeast strain that is fully sensitized to trichothecenes, as reported in this thesis, could have utility in the development of additional selection protocols designed to select for trichothecene resistance genes.

Future Directions

Although the work presented in this thesis establishes a structural basis for trichothecene UGT specificity, these findings have yet to be tested in a variety of other UGTs. Future work will focus on utilizing structural and kinetic studies to confirm that the active site residues that we identify as important to specificity in Os79 play similar roles in other plant UGTs.

The Os79 H122A/L123A/Q202L variant that we designed and tested, presented in Chapter 3 and Appendix I of this thesis, has enormous potential to provide resistance to a wide range of *Fusarium* species in transgenic crops. Future steps with this enzyme will involve generating a transgenic wheat crop and conducting field trials. It is possible that *in planta* factors will hinder the success of this enzyme, which makes the structural knowledge that we present in this thesis even more important, as it can be used to identify and design additions UGTs with broad specificity for trichothecene substrates and increased activity.

Utilizing the selection protocol that is presented in this thesis to identify trichothecene resistance genes in large genetic libraries is also a critical next step. As mentioned above, if development of additional selection protocols is needed, use of the yeast strain presented in this thesis, which is fully sensitized to trichothecenes, will serve as an excellent starting point for these studies.

In addition to using the selection protocol to identify enzymes capable of cleaving the essential epoxide bond of trichothecenes, future work will also focus on identifying such enzymes from natural sources. It is believed that sweet potato and possibly other plants express enzymes with this capability, and identifying them will be a priority. Once identified, these enzymes could serve as templates for the generation of mutant libraries to be screened with our selection protocol if they are not as active as desired.

References

1. Figueroa, M., Hammond-Kosack, K. E., and Solomon, P. S. (2018) A review of wheat diseases-a field perspective, *Mol Plant Pathol 19*, 1523-1536.

2. McMullen, M., Bergstrom, G., De Wolf, E., Dill-Macky, R., Hershman, D., Shaner, G., and Van Sanford, D. (2012) A Unified Effort to Fight an Enemy of Wheat and Barley: *Fusarium* Head Blight, *Plant Dis 96*, 1712-1728.

3. Proctor, R. H., Hohn, T. M., and McCormick, S. P. (1995) Reduced virulence of *Gibberella zeae* caused by disruption of a trichothecene toxin biosynthetic gene, *Mol Plant Microbe Interact 8*, 593-601.

4. Varga, E., Malachova, A., Schwartz, H., Krska, R., and Berthiller, F. (2013) Survey of deoxynivalenol and its conjugates deoxynivalenol-3-glucoside and 3-acetyl-deoxynivalenol in 374 beer samples, *Food Addit Contam A 30*, 137-146.

5. Walter, S., Nicholson, P., and Doohan, F. M. (2010) Action and reaction of host and pathogen during *Fusarium* head blight disease, *New Phytol 185*, 54-66.

6. Cundliffe, E., Cannon, M., and Davies, J. (1974) Mechanism of Inhibition of Eukaryotic Protein-Synthesis by Trichothecene Fungal Toxins, *P Natl Acad Sci USA 71*, 30-34.

7. Pestka, J. J. (2010) Deoxynivalenol: mechanisms of action, human exposure, and toxicological relevance, *Arch Toxicol 84*, 663-679.

8. Shin, S., Torres-Acosta, J. A., Heinen, S. J., McCormick, S., Lemmens, M., Paris, M. P. K., Berthiller, F., Adam, G., and Muehlbauer, G. J. (2012) Transgenic *Arabidopsis thaliana* expressing a barley UDP-glucosyltransferase exhibit resistance to the mycotoxin deoxynivalenol, *J Exp Bot 63*, 4731-4740.

9. Li, X., Shin, S., Heinen, S., Dill-Macky, R., Berthiller, F., Nersesian, N., Clemente, T., McCormick, S., and Muehlbauer, G. J. (2015) Transgenic Wheat Expressing a Barley UDP-Glucosyltransferase Detoxifies Deoxynivalenol and Provides High Levels of Resistance to *Fusarium graminearum*, *Mol Plant Microbe In 28*, 1237-1246.

10. Li, X., Michlmayr, H., Schweiger, W., Malachova, A., Shin, S., Huang, Y. D., Dong, Y. H., Wiesenberger, G., McCormick, S., Lemmens, M., Fruhmann, P., Hametner, C., Berthiller, F., Adam, G., and Muehlbauer, G. J. (2017) A barley UDP-glucosyltransferase inactivates nivalenol and provides *Fusarium* Head Blight resistance in transgenic wheat, *J Exp Bot 68*, 2187-2197.

11. Poppenberger, B., Berthiller, F., Lucyshyn, D., Sieberer, T., Schuhmacher, R., Krska, R., Kuchler, K., Glossl, J., Luschnig, C., and Adam, G. (2003) Detoxification of the *Fusarium* mycotoxin deoxynivalenol by a UDP-glucosyltransferase from *Arabidopsis thaliana*, *J Biol Chem 278*, 47905-47914.

12. Schweiger, W., Boddu, J., Shin, S., Poppenberger, B., Berthiller, F., Lemmens, M., Muehlbauer, G. J., and Adam, G. (2010) Validation of a Candidate Deoxynivalenol-Inactivating UDP-Glucosyltransferase from Barley by Heterologous Expression in Yeast, *Mol Plant Microbe In 23*, 977-986.

Appendix I:

APPLICATION FOR UNITED STATES LETTERS PATENT

FOR

COMPOSITIONS AND METHODS FOR T-2 TOXIN INACTIVATION

The following patent application was written in collaboration with the Wisconsin Alumni Research Foundation (WARF) and attorneys contracted by WARF.

Versions of this work have been previously published as:

Wetterhorn KM, Newmister SA, Caniza RK, Busman M, McCormick SP, Berthiller F, Adam G, Rayment I. *Crystal structure of Os79 (Os04g0206600) from* Oryza sativa*: A UDP-glucosyltransferase involved in the detoxification of deoxynivalenol.* Biochemistry. 2016; 55(44): 6175-86.

and

Wetterhorn KM, Gabardi K, Michlmayr H, Malachova A, Busman M, McCormick SP, Berthiller F, Adam G, Rayment I. *Determinants and Expansion of Specificity in a Trichothecene UDP-glucosyltransferase from* Oryza sativa. Biochemistry. 2017 Nov 30;56(50):6585-96.

ABSTRACT

Compositions and methods for increasing resistance to *Fusarium* infection in plants are provided herein. Polynucleotides, polypeptides, and expression constructs for expressing mutant UDP-glycosyltransferase proteins, plants comprising the polynucleotides, polypeptides or expression constructs, and methods of producing transgenic plants are also provided.

CROSS-REFERENCE TO RELATED APPLICATIONS

[0001] This application claims the priority of U.S. Provisional Appl. Ser. No. 62/470,726, filed March 13, 2017, the entire disclosure of which is incorporated herein by reference.

[0002]

STATEMENT REGARDING FEDERALLY SPONSORED RESEARCH OR DEVELOPMENT

[0003] This invention was made with government support under 59-0206-1-117 awarded by the USDA/ARS and 17-CRHF-0-6055 awarded by the USDA/NIFA. The government has certain rights in the invention.

FIELD OF THE INVENTION

[0004] The present disclosure relates to the field of biotechnology. More specifically, the disclosure relates to recombinant DNA molecules encoding proteins that inactivate T-2 toxin, as well as methods of producing T-2 toxin resistant plants and plants exhibiting increased T-2 toxin resistance.

BACKGROUND

[0005] Agricultural crop production often utilizes novel traits created using the methods of biotechnology. A recombinant DNA molecule encoding a modified polypeptide can be introduced into a plant to produce a novel trait. Expression of the recombinant DNA molecule in the plant confers a trait, such as T-2 toxin resistance, to the plant.

SUMMARY OF THE INVENTION

[0006] The present disclosure provides a polynucleotide comprising a DNA sequence encoding a modified UDP-glycosyltransferase polypeptide, the polypeptide comprising at least a first mutation relative to a wild-type UDP-glycosyltransferase polypeptide, wherein mutation renders the modified UDP-glycosyltransferase polypeptide capable of glycosylating T-2 toxin from *Fusarium*. In certain embodiments, the modified UDP-glycosyltransferase polypeptide comprises at least two mutations relative to a wild-type UDP-glycosyltransferase polypeptide. In particular embodiments the modified UDP glycosyltransferase polypeptide comprises a mutation at positions corresponding to amino acids 122 and 123 of SEQ ID NO:2. In some embodiments, the modified UDP glycosyltransferase polypeptide comprises the amino acid sequence of SEQ ID NO:4, SEQ ID NO:6 or SEQ ID NO:22. In other embodiments, the modified UDP-glycosyltransferase polypeptide comprises at least three mutations relative to a wild-type UDP-glycosyltransferase polypeptide. In yet other embodiments the modified UDP glycosyltransferase polypeptide comprises a mutation at positions corresponding to amino acids 122, 123 and 202 of SEQ ID NO:2. In further embodiments the modified UDP glycosyltransferase polypeptide comprises the amino acid sequence of SEQ ID NO:8 or SEQ ID NO:10.

[0007] In additional embodiments the polynucleotide is operably linked to a regulatory element. In various embodiments the regulatory element is a heterologous regulatory element. In certain embodiments the regulatory element is a promoter, for example a promoter that is functional in a plant or an inducible promoter.

[0008] The present disclosure further provides a polypeptide encoded by a polynucleotide comprising a DNA sequence encoding a modified UDP-glycosyltransferase polypeptide, the polypeptide comprising at least a first mutation relative to a wild-type UDP glycosyltransferase polypeptide, wherein mutation renders the modified UDP-glycosyltransferase polypeptide capable of glycosylating T-2 toxin from *Fusarium*. In certain embodiments the polynucleotide encodes the amino acid sequence of SEQ ID NO:4, SEQ ID NO:6, SEQ ID NO:8, SEQ ID NO:10, SEQ ID NO:12. SEQ ID NO:14, SEQ ID NO:16, SEQ ID NO:18, SEQ ID NO:20 or SEQ ID NO:22, or amino acid sequences comprising a T291V, T291S, Q202E, F199Q, Q143A, S203L or S203A mutation relative to wild-type Os79 (SEQ ID NO:2). In other embodiments the polynucleotide comprises the nucleic acid sequence of SEQ ID NO:3, SEQ ID NO:5, SEQ ID NO:7, SEQ ID NO:9, SEQ ID NO:11, SEQ ID NO:13, SEQ ID NO:15, SEQ ID NO:17, SEQ ID NO:19 or SEQ ID NO:21, or nucleic acid sequences encoding a T291V, T291S, Q202E, F199Q, Q143A, S203L or S203A mutation relative to wild-type Os79 (SEQ ID NO:2).

[0009] The present disclosure also provides a plant, plant part, cell, or seed comprising a polynucleotide comprising a DNA sequence encoding a modified UDP-glycosyltransferase polypeptide, the polypeptide comprising at least a first mutation relative to a wild-type UDP glycosyltransferase polypeptide, wherein mutation renders the modified UDP-glycosyltransferase polypeptide capable of glycosylating T-2 toxin from *Fusarium*. In some embodiments the plant, plant part, cell, or seed is defined as a monocotyledonous plant, seed, cell, or plant part. In

additional embodiments the monocotyledonous plant, seed, cell, or plant part is a corn, wheat, rice, barley, oats and sorghum plant, seed, cell, or plant part. In other embodiments the plant, plant part, cell, or seed is defined as a dicotyledonous plant, seed, cell, or plant part. In further embodiments the dicotyledonous plant, seed, cell, or plant part is a soybean, alfalfa, sunflower, cotton, canola, sweet potato, tomato, banana, curcubits, peppers and sugar beet plant, seed, cell, or plant part.

[0010] The present disclosure additionally provides a method of increasing the resistance of a plant to *Fusarium* infection, comprising expressing in the plant a polynucleotide comprising a DNA sequence encoding a modified UDP-glycosyltransferase polypeptide, the polypeptide comprising at least a first mutation relative to a wild-type UDP glycosyltransferase polypeptide, wherein mutation renders the modified UDP-glycosyltransferase polypeptide capable of glycosylating T-2 toxin from *Fusarium*. In various embodiments the plant is defined as a monocotyledonous plant or a dicotyledonous plant. In certain embodiments the method comprises transforming the plant with the polynucleotide and regenerating the plant therefrom. In other embodiments the method comprises crossing a parent plant comprising the polynucleotide with itself or a second plant to obtain the plant in which resistance to *Fusarium* infection is increased.

[0011] The present disclosure further provides a method of increasing the resistance of a plant to *Fusarium* infection, comprising modifying a UDP-glycosyltransferase polypeptide of the plant. In particular embodiments the modifying comprises mutating at least one amino acid of the UDP-glycosyltransferase polypeptide. In some embodiments the modifying comprises site-specific mutagenesis. In certain embodiments the site-specific mutagenesis comprises the use of single primer, zinc finger nucleases (ZFN), TALEN, or CRISPR technology.

BRIEF DESCRIPTION OF THE DRAWINGS

[0012] FIG. 1: Selected amino acid side chains and the secondary structural elements that carry them that define the active site architecture for the trichothecene binding site in Os79 (PDB accession number 5TMB). Members of the GT-B family of glucosyl transferase enzymes that exhibit activity towards DON will have an architecture (structure) that differs by less than an rms difference 2Å for the structurally equivalent alpha carbon atoms surrounding the active site from this model and thus will contain residues equivalent to H122, L123, and Q202 (others) that can be mutated to yield activity towards T-2 toxin.

[0013] FIG. 2: Clustal alignment of the amino acid sequences for UDP-glucosyl transferase 73C6 from *Arabidopsis thaliana* (AT_73C6_UGT_+we), DON-glucosyltransferase 1 from *Arabidopsis thaliana* (DOGT1_73C5_+ve_), UDP-glycosyltransferase superfamily protein from *Arabidopsis thaliana* (AT_73C4_+ve_Q9Z), predicted crocetin glucosyltransferase 2-like isoform X1 from *Brachypodium distachyon* (Bradi5g02780.1_), Os79 (Os_79_+ve_XP_01), hypothetical protein from *Sorghum bicolor* (sorghum) (Sb06g002180_wea), predicted UDP-glycosyltransferase 74E2-like from *Brachypodium distachyon* (Bradi5g03300.1_), and *Hordeum vulgare* subsp. vulgare UDP-glucosyltransferase HvUGT13248 from *Hordeum vulgare* subsp. vulgare (domesticated barley).

BRIEF DESCRIPTION OF THE SEQUENCES

[0014] SEQ ID NO:1: Os79 Codon Optimized DNA Sequence.

[0015] SEQ ID NO:2: Os79 WT Protein Sequence.

[0016] SEQ ID NO:3: Os79 H122A/L123G DNA Sequence.

[0017] SEQ ID NO:4: Os79 H122A/L123G Protein Sequence.

[0018] SEQ ID NO:5: Os79 H122A/L123A DNA Sequence.

[0019] SEQ ID NO:6: Os79 H122A/L123A Protein Sequence.

[0020] SEQ ID NO:7: Os79 H122A/L123A/Q202A DNA Sequence.

[0021] SEQ ID NO:8: Os79 H122A/L123A/Q202A Protein Sequence.

[0022] SEQ ID NO:9: Os79 H122A/L123A/Q202L DNA Sequence.

[0023] SEQ ID NO:10: Os79 H122A/L123A/Q202L Protein Sequence.

[0024] SEQ ID NO:11: Os79 Q202A DNA Sequence.

[0025] SEQ ID NO:12: Os79 Q202A Protein Sequence.

[0026] SEQ ID NO:13: Os79 Q202V DNA Sequence.

[0027] SEQ ID NO:14: Os79 Q202V Protein Sequence.

[0028] SEQ ID NO:15: Os79 Q202L DNA Sequence.

[0029] SEQ ID NO:16: Os79 Q202L Protein Sequence.

[0030] SEQ ID NO:17: Os79 A384S DNA Sequence.

[0031] SEQ ID NO:18: Os79 A384S Protein Sequence.

[0032] SEQ ID NO:19: HvUGT13248 WT DNA Sequence.

[0033] SEQ ID NO:20: HvUGT13248 WT Protein Sequence.

[0034] SEQ ID NO:21: HvUGT13248 H132A/L133A DNA Sequence.

[0035] SEQ ID NO:22: HvUGT13248 H132A/L133A Protein Sequence.

[0036] SEQ ID NO:23: AT_73C4_+ve_Q9Z DNA Sequence.

[0037] SEQ ID NO:24: AT_73C4_+ve_Q9Z Protein Sequence.

[0038] SEQ ID NO:25: DOGT1_73C5_+ve_ DNA Sequence.

[0039] SEQ ID NO:26: DOGT1_73C5_+ve_ Protein Sequence.

[0040] SEQ ID NO:27: AT_73C6_UGT_+we DNA Sequence.

[0041] SEQ ID NO:28: AT_73C6_UGT_+we Protein Sequence.

[0042] SEQ ID NO:29: Bradi5g02780.1_ DNA Sequence.

[0043] SEQ ID NO:30: Bradi5g02780.1_ Protein Sequence.

[0044] SEQ ID NO:31: Sb06g002180_wea DNA Sequence.

[0045] SEQ ID NO:32: Sb06g002180_wea Protein Sequence.

[0046] SEQ ID NO:33: Bradi5g03300.1_ DNA Sequence.

[0047] SEQ ID NO:34: Bradi5g03300.1_ Protein Sequence.

[0048] SEQ ID NO:35: Os79 forward primer.

[0049] SEQ ID NO:36: Os79 reverse primer.

[0050] SEQ ID NO:37: Os79: Q202A mutagenesis primer.

[0051] SEQ ID NO:38: Os79: Q202L mutagenesis primer.

[0052] SEQ ID NO:39: Os79: Q202V mutagenesis primer.

[0053] SEQ ID NO:40: Os79: A384S mutagenesis primer.

[0054] SEQ ID NO:41: Os79: H122A/L123A mutagenesis primer.

[0055] SEQ ID NO:42: Os79: H122A/L123G mutagenesis primer.

[0056] SEQ ID NO:43: HvUGT13248: H132A/L133A mutagenesis primer.

[0057] SEQ ID NO:44: HvUGT13248 forward primer.

[0058] SEQ ID NO:45: HvUGT13248 reverse primer.

[0059] SEQ ID NO:46: Os79 WT DNA Sequence.

[0060] SEQ ID NO:47: Os79: T291V mutagenesis primer.

[0061] SEQ ID NO:48: Os79: T291S mutagenesis primer.

[0062] SEQ ID NO:49: Os79: Q202E mutagenesis primer.

[0063] SEQ ID NO:50: Os79: F199Q mutagenesis primer.

[0064] SEQ ID NO:51: Os79: Q143A mutagenesis primer.

[0065] SEQ ID NO:52: Os79: S203L mutagenesis primer.

[0066] SEQ ID NO:53: Os79: S203A mutagenesis primer.

DETAILED DESCRIPTION

[0067] The present disclosure describes novel UDP-glycosyltransferase (UGT) polypeptides. The disclosed UGT polypeptides contain at least a first mutation compared to wild-type UGT, and in certain further embodiments comprise three mutations compared to wild-type UGT.

[0068] *Fusarium* head blight and *Fusarium* ear rot are devastating plant diseases that affect small grain cereals and maize on a global scale. Infection is caused by fungi of the genus *Fusarium*; members of the *Fusarium graminearum* species complex are the most prevalent agents (*1, 2*). The trichothecene mycotoxins produced by these fungi are virulence factors and accumulate in the

grain of infected plants. They are potent inhibitors of eukaryotic protein synthesis (*3*) posing a significant threat to both animal and human consumers (*4*).

[0069] Trichothecene mycotoxins are a highly diverse group of tricyclic sesquiterpenoid epoxides. More than 200 trichothecenes have been identified, all of which are characterized by a 12,13-epoxytrichothec-9-ene skeleton (*5*). Variations in the substitution pattern of the trichothecene backbone exist in different producing organisms. *F. graminearum* synthesizes either deoxynivalenol (DON) and its acetylated derivatives, 3-acetyl-deoxynivalenol and 15-acetyl-deoxynivalenol, or nivalenol (NIV) and acetylated derivatives, while T-2 toxin and HT-2 toxin are produced by *Fusarium sporotrichioides* (*5*). Differences in substitution can have a dramatic effect on the toxicity to both plants and animals, and have important implications in developing resistance strategies (*6*).

[0070] Previous research has been directed at understanding the role of the trichothecene mycotoxins in plant infection. Most efforts have focused on DON because it is the most prevalent mycotoxin associated with *Fusarium* head blight (*7*). There is strong evidence to suggest that DON is a virulence factor for plant pathogenesis (*8*). Disruption of trichothecene biosynthesis by knocking out trichodiene synthase (tri5) led to *Fusarium* that were still capable of causing infection, however with decreased virulence in wheat and a decreased ability to spread from the infection site (*8, 9*). The ability of DON to spread ahead of the fungus, causing bleaching and necrosis, is likely due to the inhibition of protein synthesis caused by the toxin and thereby facilitates the spread of infection in host tissue. As a result, DON resistance is considered an important component of *Fusarium* disease resistance (*10*).

[0071] One mechanism of DON resistance in plants is the ability to convert DON to deoxynivalenol 3-O-glucoside (D3G). It was shown that an increased ability to form D3G in

hexaploid wheat was responsible for an increased resistance to both the bleaching effects of DON and fungal spreading (*10*). Additionally, transcriptome analysis of DON-treated barley revealed the upregulation of several UGTs in a genotype that was shown to convert DON to D3G (*11*). D3G has a significantly decreased ability to inhibit wheat ribosomes *in vitro* (*12*). The first UGT capable of synthesizing D3G was cloned from *Arabidopsis thaliana* (DOGT1) and conferred increased tolerance to DON when constitutively overexpressed in *Arabidopsis* (*12*). Interestingly, DOGT1 overexpression did not confer protection against nivalenol and was accompanied by a dwarfism phenotype associated with conversion of the brassinosteroid brassinolide to the inactive brassinolide 23-O-glucoside (*13*). Subsequently, other UGT genes potentially associated with DON detoxification have been identified in *Arabidopsis*, wheat (*14*), and barley (*11*). When tested for their ability to confer DON resistance in sensitized yeast, only one (HvUGT13248) of four DON-induced barley UGT genes and two of six *Arabidopsis* UGT genes showed protection (*15*). Expression of the barley UGT gene (HvUGT13248) in wheat provided reduction in incidence and severity of *Fusarium* head blight, without obvious morphological effects, but the response was variable (*16*). This illustrates the difficulty in predicting which UGTs have the desired DON specificity among members of the very large gene family with 160−180 genes in diploid crop plants (*17-19*). An additional problem is that the UGT genes are frequently located in gene clusters that seem to undergo rapid evolution, so that even highly similar genes in clusters have different substrate specificities (*20*). Furthermore, none of the described UGTs have been reported to be effective against T-2 toxin or diacetoxyscirpenol (DAS).

[0072] Importantly, the presently disclosed UGT polypeptides are capable of detoxifying T-2 toxin, 4-acetyl-NIV (4-ANIV), 4,15-diacetyl-NIV (4,15-diANIV), and DAS, as well as DON, 15-acetyldeoxynivalenol (15-ADON), NIV, isotrichodermol (isoT) and HT-2. The present disclosure

surprisingly provides variants of UGT that are able to detoxify a number of different toxins produced by *Fusarium* species, in addition to DON, and therefore provide a broader range of protection against and resistance to *Fusarium* infections. Initially, three amino acid positions in the wild-type UGT enzyme isolated from rice (Os79; nucleotide sequence SEQ ID NO:1; amino acid sequence SEQ ID NO:2) were targeted for mutagenesis. Changing the histidine residue at position 122 of Os79 to alanine, the leucine residue at position 123 of Os79 to alanine, and the glutamine residue at position 202 of Os79 to alanine resulted in a modified UGT (nucleotide sequence SEQ ID NO:7; amino acid sequence SEQ ID NO:8) that was capable of glycosylating not only DON, NIV, isoT, and HT-2, but also T-2, 4-ANIV, 4,15-diANIV, and DAS. The following mutants also exhibited the ability to glycosylate T-2 toxin: Os79 Q202A (nucleotide sequence SEQ ID NO:11; amino acid sequence SEQ ID NO:12), Os79 Q202L (nucleotide sequence SEQ ID NO:15; amino acid sequence SEQ ID NO:16), Os79 Q202V (nucleotide sequence SEQ ID NO:13; amino acid sequence SEQ ID NO:14), Os79 H122A/L123G (nucleotide sequence SEQ ID NO:3; amino acid sequence SEQ ID NO:4), Os79 H122A/L123A (nucleotide sequence SEQ ID NO:5; amino acid sequence SEQ ID NO:6), and Os79 H122A/L123A/Q202L (nucleotide sequence SEQ ID NO:9; amino acid sequence SEQ ID NO:10). In addition, activity slightly above background was seen in Os79 A384S (nucleotide sequence SEQ ID NO:17; amino acid sequence SEQ ID NO:18). Some of the disclosed variants (for example Os79 A384S) have low activity that is only detectable by analyzing the samples for the glycosylated trichothecene product by LCMS.

[0073] The present disclosure therefore provides modified UGT polynucleotides and polypeptides that are capable of glycosylating, and therefore at least partially, or in certain embodiments completely, inactivating a range of toxins produced by different *Fusarium* species. The present

disclosure thus also provides methods of producing plants with increased resistance to a variety of toxins produced by different *Fusarium* species, by providing a modified UGT polynucleotide or polypeptide, or by modifying a naturally-occurring UGT polynucleotide or polypeptide existing in a plant, as described in detail herein.

[0074] The disclosure therefore permits increases in crop performance, grain yield and quality. In view of increasing concerns regarding food shortages in various areas of the world, this represents a significant advance to agriculture and the art in general.

I. Recombinant DNA Molecules

[0075] As used herein, the term "nucleic acid" or "polynucleotide" refers to a single or double-stranded polymer of deoxyribonucleotide bases or ribonucleotide bases read from the 5' to the 3' end. A nucleic acid or polynucleotide may also optionally contain non-naturally occurring or altered nucleotide bases that permit correct read through by a polymerase and do not reduce expression of a polypeptide encoded by that nucleic acid. The terms "nucleotide sequence" or "nucleic acid sequence" refer to both the sense and antisense strands of a nucleic acid as either individual single strands or in the duplex. The term "ribonucleic acid" (RNA) is inclusive of RNAi (inhibitory RNA), dsRNA (double stranded RNA), siRNA (small interfering RNA), mRNA (messenger RNA), miRNA (micro-RNA), tRNA (transfer RNA, whether charged or discharged with a corresponding acylated amino acid), and cRNA (complementary RNA). The words "nucleic acid fragment," "nucleotide sequence fragment", or more generally "fragment" will be understood by those in the art as a functional term that includes genomic sequences, ribosomal RNA sequences, transfer RNA sequences, messenger RNA sequences, operon sequences, and smaller engineered nucleotide sequences that express or may be adapted to express, proteins, polypeptides

or peptides. The nomenclature used herein is that required by Title 37 of the United States Code of Federal Regulations §1.822 and set forth in the tables in WIPO Standard ST.25 (1998), Appendix 2, Tables 1 and 3.

[0076] As used herein, the terms "nucleic acid" and "polynucleotide" refer to a deoxyribonucleotide, ribonucleotide, or a mixed deoxyribonucleotide and ribonucleotide polymer in either single- or double-stranded form, and unless otherwise limited, would encompass known analogs of natural nucleotides that can function in a similar manner as naturally-occurring nucleotides. The polynucleotide sequences include the DNA strand sequence that is transcribed into RNA and the strand sequence that is complementary to the DNA strand that is transcribed. The polynucleotide sequences also include both full-length sequences as well as shorter sequences derived from the full-length sequences. Allelic variations of the exemplified sequences also fall within the scope of the present disclosure. The polynucleotide sequence includes both the sense and antisense strands either as individual strands or in the duplex.

[0077] As used herein, the term "recombinant nucleic acid," "recombinant polynucleotide" or "recombinant DNA molecule" refers to a polynucleotide that has been altered from its native state, such as by linkage to one or more other polynucleotide sequences to which the recombinant polynucleotide molecule is not normally linked to in nature. Such molecules may or may not be present, for example, in a host genome or chromosome.

[0078] The present disclosure further provides polynucleotides that are complementary in sequence to the polynucleotides disclosed herein. Polynucleotides and polypeptides of the disclosure can be provided in purified or isolated form.

[0079] The subject disclosure also concerns oligonucleotide probes and primers, such as polymerase chain reaction (PCR) primers, that can hybridize to a coding or non-coding sequence of a polynucleotide of the present disclosure. Oligonucleotide probes of the disclosure can be used in methods for detecting and quantitating nucleic acid sequences encoding a mutant UGT polypeptide of the disclosure. Oligonucleotide primers of the disclosure can be used in PCR methods and other methods involving nucleic acid amplification. In a preferred embodiment, a probe or primer of the disclosure can hybridize to a polynucleotide of the disclosure under stringent conditions. Probes and primers of the disclosure can optionally comprise a detectable label or reporter molecule, such as fluorescent molecules, enzymes, radioactive moiety (*e.g.*, ^{3}H, ^{35}S, ^{125}I, *etc*.), and the like. Probes and primers of the disclosure can be of any suitable length for the method or assay in which they are being employed. Typically, probes and primers of the disclosure will be 10 to 500 or more nucleotides in length. Probes and primers that are 10 to 20, 21 to 30, 31 to 40, 41 to 50, 51 to 60, 61 to 70, 71 to 80, 81 to 90, 91 to 100 or more nucleotides in length are contemplated within the scope of the disclosure. Probes and primers of the disclosure can have complete (100%) nucleotide sequence identity with the polynucleotide sequence, or the sequence identity can be less than 100%. For example, sequence identity between a probe or primer and a sequence can be 99%, 98%, 97%, 96%, 95%, 90%, 85%, 80%, 75%, 70% or any other percentage sequence identity allowing the probe or primer to hybridize under stringent conditions to a nucleotide sequence of a polynucleotide of the disclosure. In one embodiment, a probe or primer of the disclosure has 70% or greater, 75% or greater, 80% or greater, 85% or greater, 90% or greater, or 95% to 100% sequence identity with a nucleotide sequence provided herein, including the complement thereof.

[0080] Because of the degeneracy of the genetic code, a variety of different polynucleotide sequences can encode polypeptides or mutant polypeptides disclosed herein. All possible triplet codons (and where U also replaces T) and the amino acid encoded by each codon is well-known in the art. In addition, it is well within the capability of one of skill in the art to create alternative polynucleotide sequences encoding the same, or essentially the same, mutant polypeptides of the subject disclosure. These variant or alternative polynucleotide sequences are within the scope of the subject disclosure. As used herein, references to "essentially the same" sequence refers to sequences that encode amino acid substitutions, deletions, additions, or insertions that do not materially alter the functional activity of the polypeptide encoded by the polynucleotides of the present disclosure. Allelic variants of the nucleotide sequences encoding a wild-type or mutant polypeptide of the present disclosure are also encompassed within the scope of the disclosure.

[0081] Amino acids can be generally categorized in the following classes: non-polar, uncharged polar, basic, and acidic. Conservative substitutions whereby a mutant UGT polypeptide of the present disclosure and/or a wild-type or mutant UGT polypeptide having an amino acid of one class is replaced with another amino acid of the same class fall within the scope of the subject disclosure so long as the polypeptide having the substitution still retains substantially the same functional activity (*e.g.*, enzymatic and/or increased toxin resistance of the described wild-type or mutant UGT enzyme) as the polypeptide that does not have the substitution. Functional activity may be determined as set forth in the Examples section below. Polynucleotides encoding a mutant UGT polypeptide or a wild-type UGT polypeptide having one or more amino acid substitutions in the sequence are contemplated within the scope of the present disclosure.

[0082] Glycine and cysteine are understood in the art to fall in both the nonpolar and uncharged polar classes. Substitution of amino acids other than those specifically exemplified or naturally present in the disclosed wild-type and/or mutant UGT polypeptides are also contemplated within the scope of the present disclosure. For example, non-natural amino acids can be substituted for the amino acids of a mutant UGT polypeptide, so long as the mutant UGT polypeptide having the substituted amino acids retains substantially the same functional activity as the mutant UGT polypeptide in which amino acids have not been substituted. Examples of non-natural amino acids include, but are not limited to, ornithine, citrulline, hydroxyproline, homoserine, phenylglycine, taurine, iodotyrosine, 2,4-diaminobutyric acid, a-amino isobutyric acid, 4-aminobutyric acid, 2-amino butyric acid, γ-amino butyric acid, e-amino hexanoic acid, 6-amino hexanoic acid, 2-amino isobutyric acid, 3-amino propionic acid, norleucine, norvaline, sarcosine, homocitrulline, cysteic acid, t-butylglycine, t-butylalanine, phenylglycine, cyclohexylalanine, β-alanine, fluoro-amino acids, designer amino acids such as β-methyl amino acids, C-methyl amino acids, N-methyl amino acids, and amino acid analogues in general. Non-natural amino acids also include amino acids having derivatized side groups. Furthermore, any of the amino acids in the disclosed proteins or polypeptides can be of the D (dextrorotary) form or L (levorotary) form. Allelic variants of the disclosed wild-type or mutant UGT polypeptides are also encompassed within the scope of the disclosure.

II. Methods of Modifying Nucleic Acids and Proteins

[0083] The subject disclosure also concerns variants of the polynucleotides of the present disclosure that encode functional wild-type or mutant UGT polypeptides of the disclosure. Variant sequences include those sequences wherein one or more nucleotides of the sequence have been substituted, deleted, and/or inserted. The nucleotides that can be substituted for natural nucleotides

of DNA have a base moiety that can include, but is not limited to, inosine, 5-fluorouracil, 5-bromouracil, hypoxanthine, 1-methylguanine, 5-methylcytosine, and tritylated bases. The sugar moiety of the nucleotide in a sequence can also be modified and includes, but is not limited to, arabinose, xylulose, and hexose. In addition, the adenine, cytosine, guanine, thymine, and uracil bases of the nucleotides can be modified with acetyl, methyl, and/or thio groups. Sequences containing nucleotide substitutions, deletions, and/or insertions can be prepared and tested using standard techniques known in the art.

[0084] As used herein, the term "percent sequence identity" or "percent identity" refers to the percentage of identical nucleotides in a linear polynucleotide sequence of a reference ("query") polynucleotide molecule (or its complementary strand) as compared to a test ("subject") polynucleotide molecule (or its complementary strand) when the two sequences are optimally aligned (with appropriate nucleotide insertions, deletions, or gaps totaling less than 20 percent of the reference sequence over the window of comparison). Optimal alignment of sequences for aligning a comparison window are well known to those skilled in the art and may be conducted by tools such as the local homology algorithm of Smith and Waterman, the homology alignment algorithm of Needleman and Wunsch, the search for similarity method of Pearson and Lipman, and preferably by computerized implementations of these algorithms such as GAP, BESTFIT, FASTA, and TFASTA available as part of the GCG® Wisconsin Package® (Accelrys Inc., Burlington, MA). Polynucleotides and polypeptides contemplated within the scope of the subject disclosure can also be defined in terms of identity and/or similarity ranges with those sequences of the disclosure specifically exemplified herein. In certain embodiments, the disclosure provides polypeptide sequences having at least about 70, 75, 80, 85, 90, 95, 99, or 99.5 percent identity to a polypeptide sequence provided herein, including, but not limited to, SEQ ID NO:2, SEQ ID

NO:4, SEQ ID NO:6, SEQ ID NO:8, SEQ ID NO:10, SEQ ID NO:12, SEQ ID NO:14, SEQ ID NO:16, SEQ ID NO:18, SEQ ID NO:20, SEQ ID NO:22, SEQ ID NO:24, SEQ ID NO:26, SEQ ID NO:28, SEQ ID NO:30, SEQ ID NO:32 or SEQ ID NO:34, as well as polypeptide sequences comprising a T291V, T291S, Q202E, F199Q, Q143A, S203L or S203A mutation relative to the wild-type Os79 protein sequence (SEQ ID NO:2). In certain embodiments, the disclosure provides polynucleotide sequences having at least about 70, 75, 80, 85, 90, 95, 99, or 99.5 percent identity to a polynucleotide sequence provided herein, including, but not limited to, SEQ ID NO:1, SEQ ID NO:3, SEQ ID NO:5, SEQ ID NO:7, SEQ ID NO:9, SEQ ID NO:11, SEQ ID NO:13, SEQ ID NO:15, SEQ ID NO:17, SEQ ID NO:19, SEQ ID NO:21, SEQ ID NO:23, SEQ ID NO:25, SEQ ID NO:27, SEQ ID NO:29, SEQ ID NO:31 or SEQ ID NO:33, as well as polynucleotide sequences encoding a T291V, T291S, Q202E, F199Q, Q143A, S203L or S203A mutation relative to the wild-type Os79 protein sequence (SEQ ID NO:2).

[0085] In certain embodiments, structural similarity is used to determine the amino acid residues that "correspond to" the amino acid residues of Os79 (SEQ ID NO:2). It is known in the art that enzymes that belong to the same fold-family as indicated by sequence similarity within a family of proteins and that demonstrate similar enzymatic activity and specificity will exhibit a topologically identical active site. The protein fold is conserved because it maintains a framework on which to carry out the enzymatic activity, even though the sequence identity between evolutionarily distant enzymes might be below 30%. This occurs because catalytic activity requires a finely tuned architecture in the active site of the enzyme where the organization of the side chains are arranged to enhance a chemical reaction. It is the stereo-chemistry of the catalyzed reaction that dictates the three-dimensional arrangement of the active site. An early example of this is seen in the conservation of catalytic residues in esterases and lipases (*21*). Thus,

glycosyltransferases that have the same activity are expected to have a closely related active site architecture as defined below. Likewise, changes in the active site architecture to yield a new activity in one enzyme can be expected to can be transferred to a structurally related enzyme with similar effect. This is accepted knowledge in the field of structural enzymology.

[0086] Structural models for GT-B fold family members disclosed in the present application can be readily calculated by those familiar with structural biology with publically available software packages including Robetta (*22*) and I-TASSER (*23-25*). These are proven algorithms for predicting a protein structure that only require the sequence of the enzyme that shows glucosyltransferase activity. These protocols are based on homology modeling, structural fragments, threading, molecular dynamics, and energy minimization, which are needed to yield a high-quality molecular model. The resultant models can be superimposed on the structure of Os79 (accession number 5TMB in the protein data bank) and the root mean square difference (rms) for the alpha carbon atoms for the residues that constitute the binding site for trichothecene (Table 2) can be calculated with the program LSQAB in the CCP4 program package (*26*) or within the graphic program Pymol for the selected secondary structural elements (DeLano, The PyMOL Molecular Graphics System, 2002). Visual inspection will demonstrate which amino acid residues are structurally equivalent to H122, L123, Q202, and A384 in Os79 in the new structural model and hence can be modified to generate activity towards, for example, T-2 toxin, 4-ANIV, 4,15-diANIV, and DAS.

[0087] The UDP-glucosyltransferase family of enzymes exhibit a GT-B fold (*27*). This fold consists of two β/α/β Rossmann-like domains that face each other with the active-site lying within the resulting cleft. These domains are associated with the donor (UDP-glucose) and acceptor (DON) substrate-binding sites. The trichothecene binding site (acceptor site) in Os79 is defined

by the residues given in Table 1 and their associated secondary structural elements (FIG. 1). The arrangement of ligands provided by the active site are consistent with the chemistry of glucosyl transfer and coordination of the trichothecene substrate. The corresponding structurally equivalent secondary structural elements in the trichothecene binding pocket will lie within an overall root mean square deviation (rms) 2Å (for the alpha carbon atoms) for sequences that are 20% identical or less and belong to the same fold family (*28*). Members of the GT-B glucosyltransferase superfamily that have activity towards DON will have an active site architecture as defined above that exhibits an rms difference of less than 2Å from that observed in Os79 (5TMB) for the structurally equivalent alpha carbon atoms surrounding the active site. Any GT-B member that falls within these specifications can be modified to yield activity towards T-2 toxin by changing the residues defined in Os79, and thus falls within the scope of the present disclosure.

[0088] In certain embodiments, the disclosure provides polynucleotides encoding polypeptides comprising the amino acid sequence provided herein, or a fragment or variant thereof. In certain embodiments, the polynucleotides encode polypeptides comprising a variant or mutant of the amino acid sequence provided herein, wherein the amino acid homologous or corresponding to position 122 ("at position 122") of SEQ ID NO:2 has been mutated, wherein the amino acid at position 123 of SEQ ID NO:2 has been mutated, wherein the amino acid at position 202 of SEQ ID NO:2 has been mutated, wherein the amino acid at position 384 of SEQ ID NO:2 has been mutated, or combinations, fragments, or variants thereof. As used herein, an amino acid homologous or corresponding to position 122, 123, 202 or 384 is an amino acid that is aligned with positions 122, 123, 202 or 384 of SEQ ID NO:2 when using the active site architecture methods to align polypeptides that are described herein. Thus, amino acids "corresponding to" amino acid positions 122, 123, 303 and/or 384 of SEQ ID NO:2 are those that lie within an overall

root mean square deviation (rms) of 2Å (for the alpha carbon atoms). In certain embodiments, the polynucleotides introduced into a plant encode one or more polypeptides comprising a variant wherein the amino acids corresponding to positions 122, 123, 202 and/or 384 of SEQ ID NO:2 have been mutated, or fragments or variants thereof. The disclosure further provides polynucleotides encoding polypeptides comprising a variant of the amino acid sequences provided herein wherein the amino acid corresponding to position 122 has been mutated to be an alanine (A), wherein the amino acid at position 123 has been mutated to be an A or glycine (G), wherein the amino acid at position 202 has been mutated to be an A, leucine (L) or valine (V), and/or wherein the amino acid at position 384 has been mutated to be a serine (S), or combinations thereof.

[0089] Fragments and variants of a mutant polypeptide of the present disclosure can be generated as described herein and tested for the presence of enzymatic activity as described herein, or using other standard techniques known in the art. Thus, an ordinarily skilled artisan can readily prepare and test fragments and variants of a mutant polypeptide of the disclosure and determine whether the fragment or variant retains functional activity relative to full-length or a non-variant mutant polypeptide. Fragments and variants of mutant polypeptides can be tested for UGT activity, for example using methods disclosed herein or by other methods well-known in the art.

[0090] The subject disclosure also concerns isolated mutant UGT polypeptides. In one embodiment, the mutant UGT polypeptide is an UGT polypeptide of *Oryza sativa*. In a specific embodiment, an UGT polypeptide of the disclosure has an amino acid sequence as shown in the sequence listing, or a functional fragment or variant thereof. An UGT polypeptide or enzyme of the disclosure can be purified using standard techniques known in the art. In one embodiment, a polynucleotide of the disclosure encoding an UGT polypeptide is incorporated into a

microorganism, such as *E. coli*, and the UGT polypeptide is expressed in the microorganism and then isolated therefrom.

[0091] In certain embodiments, polypeptides of the disclosure, and functional peptide fragments thereof, can be used to generate antibodies that bind specifically to a polypeptide of the disclosure, and such antibodies are contemplated within the scope of the disclosure. The antibodies of the disclosure can be polyclonal or monoclonal and can be produced and isolated using standard methods known in the art.

[0092] Polypeptide fragments according to the disclosure typically comprise a contiguous span of at least about 25 and about 463 amino acids of a sequence disclosed herein, including, but not limited to, SEQ ID NO:2, SEQ ID NO:4, SEQ ID NO:6, SEQ ID NO:8, SEQ ID NO:10, SEQ ID NO:12, SEQ ID NO:14, SEQ ID NO:16, SEQ ID NO:18, SEQ ID NO:20, SEQ ID NO:22, SEQ ID NO:24, SEQ ID NO:26, SEQ ID NO:28, SEQ ID NO:30, SEQ ID NO:32 or SEQ ID NO:34, , as well as polypeptide sequences comprising a T291V, T291S, Q202E, F199Q, Q143A, S203L or S203A mutation relative to the wild-type Os79 protein sequence (SEQ ID NO:2). In certain embodiments, polypeptide fragments comprise about 50, about 100, about 150, about 200, about 250, about 300, about 350, about 400 or about 450 amino acids of a sequence provided herein.

[0093] Fragments of a mutant UGT polypeptide of the disclosure can be obtained by cleaving the polypeptides of the disclosure with a proteolytic enzyme (such as trypsin, chymotrypsin, or collagenase) or with a chemical reagent, such as cyanogen bromide (CNBr). Alternatively, peptide or polypeptide fragments can be generated in a highly acidic environment, for example at pH 2.5. Peptide or polypeptide fragments can also be prepared by chemical synthesis or using host cells transformed with an expression vector comprising a polynucleotide encoding a fragment of an

UGT polypeptide of the disclosure, for example, a mutant polypeptide that is a fragment of an amino acid sequence provided herein.

III. Expression Constructs

[0094] Polynucleotides useful in the present disclosure can be provided in an expression construct. Expression constructs of the disclosure generally include regulatory elements that are functional in the intended host cell in which the expression construct is to be expressed. Thus, a person of ordinary skill in the art can select regulatory elements for use in bacterial host cells, yeast host cells, plant host cells, insect host cells, mammalian host cells, and human host cells. Regulatory elements include promoters, transcription termination sequences, translation termination sequences, enhancers, and polyadenylation elements. As used herein, the term "expression construct" refers to one or a combination of nucleic acid sequences that provides for transcription of an operably linked nucleic acid sequence. As used herein, "operably linked" means two DNA molecules linked in manner so that one may affect the function of the other. Operably-linked DNA molecules may be part of a single contiguous molecule and may or may not be adjacent. For example, a promoter is operably linked with a polypeptide-encoding DNA molecule in a DNA construct where the two DNA molecules are so arranged that the promoter may affect the expression of the DNA molecule.

[0095] As used herein, the term "heterologous" refers to the relationship between two or more items derived from different sources and thus not normally associated in nature. For example, a protein-coding recombinant DNA molecule is heterologous with respect to an operably linked promoter if such a combination is not normally found in nature. In addition, a particular

recombinant DNA molecule may be heterologous with respect to a cell, seed, or organism into which it is inserted when it would not naturally occur in that particular cell, seed, or organism.

[0096] An expression construct of the disclosure can comprise a promoter sequence operably linked to a polynucleotide sequence encoding a modified polypeptide of the disclosure. Promoters can be incorporated into a polynucleotide using standard techniques known in the art. Multiple copies of promoters or multiple promoters can be used in an expression construct of the disclosure. In a preferred embodiment, a promoter can be positioned about the same distance from the transcription start site in the expression construct as it is from the transcription start site in its natural genetic environment. Some variation in this distance is permitted without substantial decrease in promoter activity. A transcription start site is typically included in the expression construct.

[0097] Embodiments of the disclosure further provide a recombinant DNA molecule encoding a modified UGT polypeptide, wherein the modified UGT polypeptide comprises the amino acid sequence of SEQ ID NO:4, wherein the recombinant DNA molecule is further defined as operably linked to a heterologous regulatory element. In specific embodiments, the heterologous regulatory element is a promoter functional in a plant cell. In further embodiments, the promoter is an inducible promoter.

[0098] If the expression construct is to be provided in or introduced into a plant cell, then plant viral promoters, such as, for example, a cauliflower mosaic virus (CaMV) 35S (including the enhanced CaMV 35S promoter (*see*, *e.g.*, U.S. Patent No. 5,106,739)), a CaMV 19S promoter or a cassava vein mosaic virus promoter can be used. Other promoters that can be used for expression constructs in plants include, but are not limited to, zein promoters including maize zein promoters, prolifera promoter, Ap3 promoter, heat shock promoters, T-DNA 1′- or 2′-promoter of *A.*

tumefaciens, polygalacturonase promoter, chalcone synthase A (CHS-A) promoter from petunia, tobacco PR-1a promoter, ubiquitin promoter, actin promoter, alcA gene promoter, pin2 promoter (*29*), maize WipI promoter, maize trpA gene promoter (U.S. Patent No. 5,625,136), maize CDPK gene promoter, and RUBISCO SSU promoter (U.S. Patent No. 5,034,322) can also be used. Constitutive promoters (such as the CaMV, ubiquitin, actin, or NOS promoter), developmentally-regulated promoters, and inducible promoters (such as those promoters than can be induced by heat, light, hormones, or chemicals) are also contemplated for use with polynucleotide expression constructs of the disclosure.

[0099] Expression constructs of the disclosure may optionally contain a transcription termination sequence, a translation termination sequence, a sequence encoding a signal peptide, and/or enhancer elements. Transcription termination regions can typically be obtained from the 3′ untranslated region of a eukaryotic or viral gene sequence. Transcription termination sequences can be positioned downstream of a coding sequence to provide for efficient termination. A signal peptide sequence is a short amino acid sequence typically present at the amino terminus of a protein that is responsible for the relocation of an operably linked mature polypeptide to a wide range of post-translational cellular destinations, ranging from a specific organelle compartment to sites of protein action and the extracellular environment. Targeting gene products to an intended cellular and/or extracellular destination through the use of an operably linked signal peptide sequence is contemplated for use with the polypeptides of the disclosure. Classical enhancers are *cis*-acting elements that increase gene transcription and can also be included in the expression construct. Classical enhancer elements are known in the art, and include, but are not limited to, the CaMV 35S enhancer element, cytomegalovirus (CMV) early promoter enhancer element, and the SV40 enhancer element. Intron-mediated enhancer elements that enhance gene expression are also

known in the art. These elements should be present within the transcribed region and are orientation dependent. Examples include the maize shrunken-1 enhancer element (*30*).

IV. Transformation Methods

[00100] One aspect of the disclosure includes plant cells, plant tissues, plants, and seeds that comprise the recombinant DNA provided by the disclosure. These cells, tissues, plants, and seeds comprising the recombinant DNA molecules exhibit resistance to diseases caused by *Fusarium*. Suitable methods for transformation of host plant cells for use with the current disclosure include virtually any method by which DNA can be introduced into a cell (for example, where a recombinant DNA construct is stably integrated into a plant chromosome) and are well-known in the art. Two effective methods for cell transformation are *Agrobacterium*-mediated transformation and microprojectile bombardment-mediated transformation. Microprojectile bombardment methods are illustrated, for example, in U.S. Patent Nos. 5,550,318, 5,538,880, 6,160,208, and 6,399,861. *Agrobacterium*-mediated transformation methods are described, for example in U.S. Patent No. 5,591,616. Transformation of plant material is practiced in tissue culture on nutrient media, for example a mixture of nutrients that allow cells to grow *in vitro*. Recipient cell targets include, but are not limited to, meristem cells, shoot tips, hypocotyls, calli, immature or mature embryos, and gametic cells such as microspores and pollen. Callus can be initiated from tissue sources including, but not limited to, immature or mature embryos, hypocotyls, seedling apical meristems, microspores and the like. Cells containing a transgenic nucleus are grown into transgenic plants.

[00101] In transformation, DNA is typically introduced into only a small percentage of target plant cells in any one transformation experiment. Marker genes are typically used to provide an efficient

system for identification of those cells that are stably transformed by receiving and integrating a recombinant DNA molecule into their genomes. Preferred marker genes provide selective markers which confer resistance to a selective agent, such as an antibiotic or an herbicide. Any of the herbicides to which plants of this disclosure can be resistant is an agent for selective markers. Potentially transformed cells are exposed to the selective agent. In the population of surviving cells are those cells where, generally, the resistance-conferring gene is integrated and expressed at sufficient levels to permit cell survival. Cells can be tested further to confirm stable integration of the exogenous DNA. Commonly used selective marker genes include those conferring resistance to antibiotics such as kanamycin and paromomycin (*nptII*), hygromycin B (*aph IV*), spectinomycin (*aad*A) and gentamycin (*aac3* and *aac*C4) or resistance to herbicides such as glufosinate (*bar* or *pat*), dicamba (DMO) and glyphosate (*aroA* or EPSPS). Examples of such selectable markers are illustrated in U.S. Patent Nos. 5,550,318, 5,633,435, 5,780,708 and 6,118,047. Markers that provide an ability to visually screen transformants can also be employed, for example, a gene expressing a colored or fluorescent protein such as a luciferase or green fluorescent protein (GFP) or a gene expressing a *beta*-glucuronidase or *uid*A gene (GUS) for which various chromogenic substrates are known.

[00102] The present disclosure provides methods and constructs for regenerating a plant from a cell with modified genomic DNA resulting from genome editing. The regenerated plant can then be used to propagate additional plants.

V. Genome Editing

[00103] Targeted modification of plant genomes through the use of genome editing methods can be used to create improved mutant or transgenic plant lines through modification or insertion of

plant genomic DNA. In addition, genome editing methods can enable targeted insertion of multiple nucleic acids of interest (a trait stack) into a plant genome. Exemplary methods for introducing recombinant DNA constructs into a plant or modifying genomic DNA of a plant include the use of zinc-finger nucleases, engineered or native meganucleases, TALE-endonucleases, or an RNA-guided endonuclease (for example a Clustered Regularly Interspersed Short Palindromic Repeat (CRISPR)/Cas9 system). Methods of genome editing to modify, delete, or insert genomic DNA are known in the art.

[00104] In exemplary methods provided by the disclosure, a CRISPR/Cas9 system is used to modify or replace an existing coding sequence within a plant genome, such as a sequence encoding an UGT polypeptide. In further embodiments, transcription activator-like effectors (TALEs) are used for modification or replacement of an existing coding sequence within a plant genome, such as a sequence encoding an UGT polypeptide. Modification or replacement of an endogenous UGT-encoding sequence according to the methods provided herein results in a polypeptide comprising a modified UGT enzyme, for example wherein the amino acid corresponding to position 122 of SEQ ID NO:2 has been mutated to A, wherein the amino acid at position 123 of SEQ ID NO:2 has been mutated to A or G, wherein the amino acid at position 202 of SEQ ID NO:2 has been mutated to A, L or V, and/or wherein the amino acid at position 384 of SEQ ID NO:2 has been mutated to S, or combinations thereof. The disclosure therefore provides DNA constructs capable of recognizing a specific nucleotide sequence of interest, such as an UGT sequence, within a genome of a plant to allow for mutation or integration at that site.

[00105] In certain embodiments, genome editing methods provided by the disclosure may introduce single nucleotide mutations, or alterations to a number of nucleotides within a target sequence, such as an UGT-encoding sequence. Modifications to an UGT-encoding sequence, for

example a sequence provided herein as SEQ ID NO:2, may result in a sequence encoding an UGT polypeptide, for example wherein the amino acid corresponding to position 122 of SEQ ID NO:2 has been mutated to A, wherein the amino acid at position 123 of SEQ ID NO:2 has been mutated to A or G, wherein the amino acid at position 202 of SEQ ID NO:2 has been mutated to A, L or V, and/or wherein the amino acid at position 384 of SEQ ID NO:2 has been mutated to S, or combinations thereof as described herein, capable of conferring to a plant improved resistance to toxins produced by various species of *Fusarium*, which confers resistance to diseases caused by *Fusarium* species.

[00106] In further embodiments, a DNA sequence, such as a transgene or expression cassette, may be inserted or integrated into a specific site or locus within the genome of a plant or plant cell via site-directed integration. Recombinant DNA constructs and recombinant DNA molecules provided herein may thus include a donor template sequence comprising at least one transgene, expression cassette, or other DNA sequence for insertion into the genome of the plant or plant cell. In certain embodiments of the disclosure, a donor template sequence comprises a nucleotide sequence encoding a polypeptide wherein of SEQ ID NO:2 has been altered to have improved resistance to toxins produced by various species of *Fusarium*, for example the polypeptide of SEQ ID NO:4, SEQ ID NO:6, SEQ ID NO:8, SEQ ID NO:10, SEQ ID NO:12, SEQ ID NO:14, SEQ ID NO:16 or SEQ ID NO:18, or a sequence encoding a fragment of such a polypeptide, or a nucleotide sequence encoding a polypeptide wherein of SEQ ID NO:20 has been altered to have improved resistance to toxins produced by various species of *Fusarium*, for example the polypeptide of SEQ ID NO:22. A donor template for site-directed integration may further include one or two homology arms flanking an insertion sequence (*i.e.*, the sequence, transgene, cassette, *etc.*, to be inserted into the plant genome). The recombinant DNA construct(s) of this disclosure

may further comprise an expression cassette(s) encoding a site-specific nuclease and/or any associated protein(s) to carry out site-directed integration. These nuclease expressing cassette(s) may be present in the same molecule or vector as the donor template (in *cis*) or on a separate molecule or vector (in *trans*). Several methods for site-directed integration are known in the art involving different proteins (or complexes of proteins and/or guide RNA) that cut the genomic DNA to produce a double strand break (DSB) or nick at a desired genomic site or locus. As understood in the art, during the process of repairing the DSB or nick introduced by the nuclease enzyme, the donor template DNA may become integrated into the genome at the site of the DSB or nick. The presence of the homology arm(s) in the donor template may promote the adoption and targeting of the insertion sequence into the plant genome during the repair process through homologous recombination, although an insertion event may occur through non-homologous end joining (NHEJ). Examples of site-specific nucleases that may be used include zinc-finger nucleases, engineered or native meganucleases, TALE-endonucleases, and RNA-guided endonucleases (*e.g.*, Cas9 or Cpf1). For methods using RNA-guided site-specific nucleases (*e.g.*, Cas9 or Cpf1), the recombinant DNA construct(s) will also comprise a sequence encoding one or more guide RNAs to direct the nuclease to the desired site within the plant genome.

[00107] The disclosure further provides plants produced by the methods disclosed herein. Plants of the present disclosure may be monocots or dicots, and may include, for example, rice, wheat, barley, oats, rye, sorghum, maize, soybean, alfalfa, sunflower, cotton, canola, sugar beet, sweet potato, tomato, tobacco, banana, curcubits and pepper plants.

VI. Toxins Produced by *Fusarium*

[00108] One aspect of the disclosure includes plant cells, plant tissues, plants, and seeds that comprise the recombinant DNA provided by the disclosure. These cells, tissues, plants, and seeds comprising the recombinant DNA molecules exhibit resistance to one or more of the toxins produced by various species of *Fusarium*. Non-limiting examples include: DAS, which is produced by *Fusarium acuminatum*, *Fusarium equiseti*, *Fusarium langsethiae*, *Fusarium poae*, *Fusarium sambucinum*, *Fusarium semitectum*, *Fusarium sporotrichioides*, *Fusarium venenatum*; T-2 toxin, which is produced by *Fusarium acuminatum*, *Fusarium armeniacum*, *Fusarium equiseti*, *Fusarium kyushuense*, *Fusarium langsethiae*, *Fusarium oxysporum, Fusarium sambucinum*, *Fusarium sporotrichioides*; DON, which is produced by *Fusarium asiaticum*, *Fusarium austroamericanum*, *Fusarium boothii*, *Fusarium culmorum*, *Fusarium graminearum*, *Fusarium mesoamericanum*, *Fusarium pseudograminearum*; HT-2, which is produced by *Fusarium acuminatum*, *Fusarium armeniacum*, *Fusarium langsethiae*, *Fusarium oxysporum, Fusarium sambucinum*, *Fusarium sporotrichioides*; and NIV, which is produced by *Fusarium asiaticum*, *Fusarium austroamericanum*, *Fusarium cortaderiae*, *Fusarium crookwellense*, *Fusarium culmorum*, *Fusarium equiseti*, *Fusarium graminearum*, *Fusarium kyushuense*, *Fusarium meridionale*, *Fusarium mesoamericanum*, *Fusarium pseudograminearum*. Examples of these and other toxins produced by these species and other *Fusarium* species are well-known to those of skill in the art.

[00109] As used herein, "resistance" or "*Fusarium* resistance" means a plant, seed, or cell's ability to resist the toxic effects of one or more toxin when infected or contacted with *Fusarium*. The resistance of a plant, seed, plant tissue, plant part, or cell may be measured by comparing the plant, seed, plant tissue, plant part, or cell to a suitable control. For example, the resistance may be

measured or assessed by applying *Fusarium* to a plant comprising a recombinant DNA molecule encoding a modified polypeptide capable of conferring resistance to *Fusarium* (the test plant) and a plant not comprising the recombinant DNA molecule encoding the modified polypeptide capable of conferring resistance to *Fusarium* (the control plant) and then comparing the plant injury or toxin levels of the two plants (or parts thereof), where resistance of the test plant is indicated by a decreased injury rate or toxin level as compared to the injury rate or toxin level of the control plant. A *Fusarium* resistant plant, seed, plant tissue, plant part, or cells exhibits a decreased response to the toxic effects of *Fusarium* infection, or decreased levels of one or more *Fusarium* toxin, when compared to a control plant, seed, plant tissue, plant part, or cell. As used herein, a "*Fusarium* resistance trait" is a transgenic trait imparting improved resistance to a plant as compared to the wild-type plant.

EXAMPLES

Example 1: Cloning and Expression of Os79

[00110] The gene for Os79 was amplified by polymerase chain reaction (PCR) using the forward primer 5′-ATGGGCTCTATGTCCACTCCTGC-3′ (SEQ ID NO:35) and the reverse primer 5′-ATTGGAATACTTTGCTGCAAACTC- 3′ (SEQ ID NO:36) from plasmid pWS57 that contained a codon optimized sequence for yeast (*31*). The resulting product was introduced into plasmid pKLD116, a pET31b derivative containing His6-tagged maltose-binding protein (MBP) followed by a TEV protease cleavage site (*32*), using an enzyme-free "Quikchange" method (*33, 34*). Os79 was overexpressed in *E. coli* strain BL21 Codon Plus (DE3). Cultures from a single colony were used to inoculate 6 L of lysogeny broth (LB) supplemented with 100 μg/mL ampicillin and 30 μg/mL chloramphenicol. Expression of Os79 was induced with 1 mM isopropyl β-D-thiogalactopyranoside when cultures reached an OD600 of ~1.0. Induction was conducted at 16°C

for 20 hours. Cells were harvested by centrifugation at 2500 g, washed with buffer containing 10 mM 4-(2-hydroxyethyl)-1-piperazineethanesulfonic acid (HEPES) (pH 7.6), and 100 mM NaCl, and flash-frozen in liquid nitrogen. Cells were stored at −80°C until they were used.

Example 2: Site-Directed Mutagenesis of Os79

Mutants of Os79 were generated using a single primer, PCR based method based on "Quikchange" mutagenesis (*33, 34*). The following mutagenesis primers were used for Os79: Q202A: 5'-TCCTGCCTTCTGTGAGGCGTCTATCGAGCAGTTTGCT-3' (SEQ ID NO:37); Q202L: 5'-TCCTGCCTTCTGTGAGCTGTCTATCGAGCAGTTTGCT-3' (SEQ ID NO:38); Q202V: 5'-TCCTGCCTTCTGTGAGGTGTCTATCGAGCAGTTTGCT-3' (SEQ ID NO:39); A384S: 5'-GCTATGCCTCACTGGTCTGATCAACCTACTATTAGCAAGTATG-3' (SEQ ID NO:40); H122A/L123A: 5'-AGAGTTTTAGTCTACGATCCAGCGGCG CCATGGGCTAGAAGAGT-3' (SEQ ID NO:41); H122A/L123G: 5'-AGAGTTTTAGTCT ACGATCCAGCGGGCCCATGGGCTAGAAGAGT-3 (SEQ ID NO:42); T291V: 5'-GGTTTT GGTGTCATACGGAGTGGTTTCTACTTTTGATGTTGCTAAAC-3' (SEQ ID NO:47); T291S: 5'-GGTTTTGGTGTCATACGGAAGCGTTTCTACTTTTGATGTTGCTAAAC-3' (SEQ ID NO:48); Q202E: 5'-TCCTGCCTTCTGTGAGGAATCTATCGAGCAGTTTGCT-3' (SEQ ID NO:49); F199Q: 5'-CCTGAATTAACTCCTGCCCAATGTGAGCAATCTATC GAGC-3' (SEQ ID NO:50); Q143A: 5'-CAGCCGCATTTCTAAGTGCGCCATGTGCTG TGGAC-3' (SEQ ID NO:51); S203L: 5'-GCCTTCTGTGAGCAATTGATCGAGCAGTTT GCTG (SEQ ID NO:52); and S203A: 5'- GCCTTCTGTGAGCAAGCTATCGAGCAGTTT GCTG (SEQ ID NO:53). The following mutagenesis primer was used for HvUGT13248: H132A/L133A: 5'-CCCGTGCGCGTGCTGGTGTACGACGCTGCGGCGGCGTGGGCAC GGCGGGTGGCACA-3' (SEQ ID NO:43). All mutations were verified by DNA sequencing

using BigDye® protocols (ABI PRISM). Multiple mutations were introduced by adding all relevant primers to the Quikchange reaction mixture. For example, the H122A/L123A/Q202L mutant was generated adding both the H122A/L123A and Q202L primers.

Example 3: Expression and Purification of Os79 Mutants

[00111] Os79 mutants were overexpressed in *E. coli* strain BL21 Codon Plus (DE3). Cultures from a single colony were used to inoculate 6 L Lysogeny broth (LB) supplemented with 100 μg/mL ampicillin and 30 μg/mL chloramphenicol. Expression of Os79 was induced with 1 mM isopropyl-β-D-thiogalactopyranoside when cultures reached OD_{600} ~0.8. Induction was carried out at 16°C for 20 hours. Cells were harvested by centrifugation at 3,000 x g, washed with buffer containing 10 mM HEPES pH 7.6 and 100 mM NaCl, and flash-frozen in liquid nitrogen. Cells were stored at -80°C until use.

[00112] All Os79 mutants were purified in a similar manner. All purification steps were carried out on ice or at 4°C. Around twenty grams of *E. coli* cells expressing His_6-MBP-Os79 were resuspended in 100 mL buffer containing 20 mM HEPES pH 7.6, 50 mM NaCl, 0.2 mM tris(2-carboxyethyl)phosphine (TCEP), 1 mM PMSF, 50 nM leupeptin (Peptide International), 70 nM E-65 (Peptide International), 2 nM aprotinin (ProSpec), 2 mM AEBSF (Gold BioTechnology) and 50 mg lysozyme. Cells were lysed with 5 pulses (45 s) using a Qsonica Q700 sonicator, and the lysate was clarified by centrifugation at 40,000 rpm in a Ti 45 rotor (Beckman-Coulter) for 30 min. The concentration of NaCl and imidazole were raised to 300 mM and 20 mM, respectively, by the addition of 4 M stock solutions and loaded onto an 5 mL nickel-nitrilotriacetic acid (NiNTA; Qiagen) column equilibrated with NTA buffer: 50 mM HEPES, pH 7.6, 300 mM NaCl, 20 mM imidazole, and 0.2 mM TCEP. After loading, the column was washed with 120 mL of NTA buffer. Os79 was eluted with 40 mL NTA buffer containing 300 mM

imidazole. His_6-tagged Tobacco etch virus (TEV) protease was added at 1:40 molar ratio to cleave the His_6-MBP from Os79. The mixture was dialyzed overnight in a buffer containing 10 mM HEPES pH 7.6, 50 mM NaCl, 0.5 mM EDTA, 0.2 mM TCEP. The NaCl and imidazole concentrations were brought up to 300 mM and 20 mM, respectively, and the solution was passed over a 5 mL NiNTA column equilibrated with NTA buffer. The flow-though contained Os79, while the column retained TEV protease, His_6-MBP, and undigested His_6-MBP Os79. Purified Os79 was concentrated using a centrifugal filter (Amicon) with a 30,000 nominal molecular weight limit to a final concentration of 10-20 mg/mL, estimated using calculated molar extinction coefficient of 57870 $M^{-1} \cdot cm^{-1}$ at 280 nm. Proteins were dialyzed against a storage buffer containing 10 mM HEPES pH 7.6, 0.2 mM TCEP, drop frozen in 30 µL aliquots in liquid nitrogen, and stored at -80°C.

Example 4: Crystallization

[00113] Crystallization of Os79·deoxynivalenol-3-*O*-glucoside (D3G)·UDP. Os79 was screened for initial crystallization conditions with a 144-condition sparse-matrix screen developed by the inventors. Single, diffraction quality crystals were grown at 4°C by hanging drop vapor diffusion by mixing 2 µL of Os79 at 12 mg/mL in 10 mM HEPES pH 7.6, 50 mM NaCl, 30 mM D3G, 5 mM UDP with 2 µL well solution containing 100 mM sodium acetate pH 5.0, 40% pentaerythritol propoxylate 426, 320 mM NaCl (Gulick, *et al.*, *Acta Crystallogr. D Biol. Crystallogr.* 58:306-309, 2002). Hanging droplets were nucleated after 24 h from an earlier spontaneous crystallization event using a cat's whisker. Crystals grew to approximate dimensions of 100 x 100 x 300 µm within 15 days. The crystals were transferred to a cryoprotecting solution that contained 100 mM sodium acetate pH 5.0, 40% pentaerythritol propoxylate 426, 320 mM NaCl, 30 mM D3G, 5 mM UDP and vitrified by rapid-plunging into liquid nitrogen. Os79

crystallized in the space group $P3_221$ with unit cell dimensions of a = 104.5 Å, b = 104.5 Å, c = 98.3 Å and one chain in the asymmetric unit.

[00114] Crystallization of Os79 Q202A·UDP. Os79 Q202A was screened for initial crystallization conditions as described above. Single, diffraction quality crystals were grown at 4°C by hanging drop vapor diffusion by mixing 2 µL of 11 mg/mL Os79 Q202A in 10 mM HEPES pH 7.6, 5 mM UDP-glucose with 2 µL well solution containing 100 mM 3-(*N*-Morpholino)propanesulfonic acid pH 7.5, 15% methyl ether polyethylene glycol 5000 (MEPEG 5K), 160 mM potassium glutamate. Hanging droplets were nucleated after 24 h from an earlier spontaneous crystallization event using a cat's whisker. Crystals grew to approximate dimensions of 100 x 100 x 300 µm within 8 days. The crystals were transferred stepwise to a cryoprotecting solution that contained 100 mM 3-(*N*-Morpholino)propanesulfonic acid pH 7.5, 15% MEPEG 5K, 15% glycerol, 5 mM UDP-glucose and vitrified by rapid plunging into liquid nitrogen. Os79 Q202A crystallized in the space group $P2_12_12_1$ with unit cell dimensions of a = 59.4 Å, b = 83.2 Å, c = 99.1 Å and one chain in the asymmetric unit.

[00115] Crystallization of Os79 T291V·UDP. The Os79 protein complex T291V·UDP was screened for initial crystallization conditions as described above. Single, diffraction quality crystals were grown at 23°C by hanging drop vapor diffusion by mixing 2 µL of 10 mg/mL T291V Os79 in 10 mM HEPES pH 7.6, 50 mM NaCl, 5 mM UDP-glucose with 2 µL well solution containing 50 mM HEPES pH 7.5, 32% MEPEG 5K. Hanging droplets were nucleated after 24 h from an earlier spontaneous crystallization event using a cat's whisker. Crystals grew to approximate dimensions of 75 x 75 x 300 µm within 11 days. The crystals were transferred to a cryoprotecting solution that contained 50 mM HEPES 7.5, 32% MEPEG 5K, 5 mM UDP-glucose and vitrified by rapid plunging into liquid nitrogen. Os79 crystallized in the space group $P2_12_12_1$

with unit cell dimensions of a = 59.4 Å, b = 83.2 Å, c = 98.7 Å and one chain in the asymmetric unit.

[00116] Crystallization of Os79 H122A/L123A·UDP. Os79 was screened for initial crystallization conditions as described above. Single, diffraction quality crystals were grown at 23°C by hanging drop vapor diffusion by mixing 2 µL of 11 mg/mL Os79 in 10 mM HEPES pH 7.6, 50 mM NaCl, 5 mM UDP with 2 µL well solution containing 50 mM HEPES pH 7.0, 18% methyl ether polyethylene glycol 2000. Hanging droplets were nucleated after 24 h from an earlier spontaneous crystallization event using a cat's whisker. Crystals grew to approximate dimensions of 75 x 75 x 300 µm within 4 days. The crystals were transferred stepwise to a cryoprotecting solution that contained 50 mM HEPES 7.0, 20% methyl ether polyethylene glycol 2000, 15% glycerol, and 5 mM UDP and vitrified by rapid plunging into liquid nitrogen. Os79 crystallized in the space group $P2_12_12_1$ with unit cell dimensions of a = 59.3 Å, b = 82.9 Å, c = 99.0 Å and one chain in the asymmetric unit.

[00117] Data collection and Refinement. X-ray data for the Os79 structures were collected at 100 K. Diffraction data were integrated and scaled with HKL3000 (*35*). Data collection and refinement statistics are given in Table 3. The structures were determined by molecular replacement using coordinates from the RCSB (accession number 5TME) as the molecular replacement search model in the program Phaser (*36, 37*). Final models were generated by alternating cycles of manual building and least-squares refinement using Coot, Phenix and Refmac (*38-40*).

Example 5: Glucosyltransferase Enzyme Assays

[00118] Steady-state kinetic analyses of wild-type Os79 (SEQ ID NO:2) and mutant Os79 (SEQ ID NO:4) with trichothecene substrates were performed in a coupled−continuous enzymatic assay at 23°C in a 1 cm path length cuvette. Reactions were initiated by the addition of varying volumes of trichothecenes to a master mix containing Os79 (58 nM final concentration), 3 units rabbit muscle lactic dehydrogenase and 2 units rabbit muscle pyruvate kinase (Sigma-Aldrich, buffered aqueous glycerol solution), 1.5 mM phosphoenolpyruvate, 100 μM β-NADH, 50 mM KCl, 10 mM $MnCl_2$, and 100 mM glycylglycine pH 8.0 to yield a final volume of 100 μL. Lactic dehydrogenase, pyruvate kinase, phosphoenolpyruvate, β-NADH, Os79, and UDP-glucose were added to a master-mix containing the remaining reaction components prior to the initiation of each reaction. Reaction progress was followed by monitoring the decrease in A_{340} caused by the oxidation of β-NADH. The rates of reaction were determined at various trichothecene concentrations and fit by non-linear regression to the Michaelis-Menten equation using GraphPad Prism software. The final mutant enzyme concentrations were 58 nM except for H122A/L123G which was 588 nM.

[00119] Os79 glycosylates DON (k_{cat} = 0.57 s^{-1}, K_m = 0.23 mM (*31*)), while the results of the assay shows that Os79 H122A/L123A/Q202L glycosylates DON (k_{cat} = 0.85 s^{-1}, K_M = 261 μM), T-2 toxin (k_{cat} = 0.9 s^{-1}, K_M = 118 μM) and 4-ANIV (k_{cat} = 0.96 s^{-1}, K_M = 88 μM). Os79 H122A/L123A/Q202A glycosylates DON (k_{cat} = 0.9 s^{-1}, K_M = 1.2 mM), T-2 toxin (k_{cat} = 2.5 s^{-1}, K_M = 89 μM), DAS (k_{cat} = 1.8 s^{-1}, K_M = 49 μM), and 4-ANIV (k_{cat} = 1.1 s^{-1}, K_M = 501 μM). Os79 H122A/L123A glycosylates T-2 toxin (k_{cat} = 1.6 s^{-1}, K_M = 926 μM). Os79 H122A/L123G glycosylates T-2 toxin (k_{cat} = 0.26 s^{-1}, K_M = 512 μM), and DON (k_{cat} = 0.5 s^{-1}, K_M = 2.5 mM). Any mutants or substrates for which kinetic parameters were not included were found to glycosylate/be

glycosylated by the observation of activity at a single high substrate concentration and thus k_{cat} and K_M were not determined in these cases.

Example 6: Role of the conserved Thr 291 in Os79

[00120] To address the structural role of Thr 291, the structure of Os79 T291V in complex with UDP was determined to 1.61 Å resolution (PDB: 6BK1) (Table 3). When the hydroxyl of Thr 291 is replaced with a methyl group in Os79 T291V the phosphate of UDP adopts a substantially different position in the active site compared to that seen in WT Os79. In the WT Os79 structure there is a hydrogen bond between the threonine hydroxyl and a phosphate oxygen atom of length about 2.5 Å. This is clearly lost when threonine is replaced by a valine. The change in the position of UDP is likely a direct result of the substitution since crystals of Os79 T291V and WT protein structures were grown at the same pH and under very similar crystallization conditions.

[00121] The importance of this interaction is confirmed as this interaction is present in all plant GT-1 glycosyltransferases. Comparison of the structures of six plant UGT structures that have been determined with UDP bound in their active sites reveals that the presence of a hydroxyl moiety within hydrogen bonding distance of a β-phosphate oxygen is a conserved characteristic in all of these enzymes. The proteins were aligned by superimposing an α-helix and two beta strands that surround the UDP binding site, because the overall architecture of these proteins varies substantially at the periphery of the protein, away from the UDP binding site. These three secondary structural elements that underlie the framework for the UDP binding site are very similar across the six proteins and align with an average rmsd (root mean square deviation) of 0.30 Å for the alpha carbon atoms. In every case, there is a threonine or serine hydroxyl positioned within 3.0 Å of the phosphate oxygen. In every structure the phosphate adopts almost exactly the same orientation. This provides further evidence that the change in the orientation of the phosphate

of UDP in the T291V structure is caused by the lack of a hydroxyl group and the inability to form a hydrogen bond. This serine or threonine residue has not been previously identified as a catalytically extremely important residue. The present disclosure demonstrates that it is extremely important for orienting the phosphate in the active site. While this does not preclude it from participating as a catalytic acid, it does confirm a structural role for this side chain. This likely applies to many, if not all, plant UGTs.

Example 7: Steady-state kinetic parameters for WT Os79

[00122] Many plant UGTs demonstrate broad substrate specificity, a useful characteristic when considering enzymes that glycosylate xenobiotics. Os79 exhibits broad specificity and can glycosylate DON, HT-2, IsoT, and NIV but not T-2, 4-ANIV, 4,15,diANIV or DAS. T-2 is commonly produced by *Fusarium* species in Europe, Asia, Africa, and Australia. Given the importance of detoxifying T-2, expanding the specificity of Os79 to include this trichothecene and others would be of great benefit. Given that the only difference between T-2 and HT-2 toxin is the C4 acetyl group that is present on T-2 in place of the hydroxyl on HT-2 the inventors reasoned that this acetyl group prevented T-2 from binding in the active site of Os79. By this reasoning, 4-acetyl nivalenol (4-ANIV), commonly known as FUS-X, will not be a substrate for Os79. 4-ANIV is acetylated at the C4 position but lacks the C15 acetyl and C8 isovaleryl groups of T-2. Indeed, glycosylation could not be detected with the coupled assay with 4-ANIV, 4,15,diANIV or DAS. This confirms that the 4-acetyl group is responsible for precluding T-2, 4-ANIV, 4,15,diANIV and DAS from the acceptor binding pocket. To further examine the structural aspects of the acceptor binding pocket that contribute to its inability to accommodate an acetyl group on the C4 position, the structure of Os79 with the glycosylated trichothecene product D3G in the active site was determined.

Example 8: Structure of Os79 in complex with UDP and D3G

[00123] Os79 was crystallized in the presence of the product D3G to further examine the structural components responsible for substrate specificity and further understand the nature of trichothecene binding in the acceptor pocket. The structure of Os79 in complex with UDP and D3G was determined to 2.17 Å resolution (PDB: 6BK3) (Table 3). The overall structure is very similar to the previously solved structure of Os79 with trichothecene (TRI) and UDP-2-fluoro-2-deoxy-D-glucose (U2F) bound in the active site with an rmsd of 0.7 Å for structurally equivalent α-carbons. Electron density corresponding to DON and UDP was observed, however there was no clear electron density corresponding to the glucose moiety of D3G and the moiety was not modeled. This appears to be the result of the flexibility of the glucose moiety.

[00124] There are two major conformational changes in Os79 with D3G bound compared to the structure with trichothecene bound in the active site. A loop from Ser 288 to Val 297 is shifted 9.5 Å away from the active site and the region that extends from Trp 316 to Lys 336, which is composed of a loop and two short α-helices, is shifted 5.9 Å away from the active site. These conformational changes highlight the flexibility of the acceptor binding region of the protein, and help to show the structural basis for UDP release after the donor sugar has been transferred. Aligning the structures with TRI or D3G bound in the active site reveals that the trichothecene skeleton of DON is rotated about 45 degrees compared to the backbone of TRI. This change is most probably the result of the presence of glucose on C3 and shows that the trichothecene backbone rotates in the acceptor binding pocket after the reaction is completed and before the glycosylated product is released. The orientation of trichothecene is more representative than D3G of the positioning of DON in the active site prior to glycosylation. Using the orientation of trichothecene as a reference, His 122, which is only 4.2 Å away from C4, was identified as a

residue that could clash with the C4 acetyl of the substrates that WT Os79 is unable to glycosylate. Based on the presently disclosed structural information, the role of His 122 and other residues identified as potential contributors to the specificity of Os79 were investigated by kinetic analysis.

Example 9: Steady-state kinetic assays on Os79 mutants

[00125] The remarkable substrate plasticity of Os79 prompted the inventors to investigate whether there are important residues in the trichothecene binding pocket that facilitate this ability. It also raised in the inventors the possibility of expanding the substrate specificity to accommodate trichothecenes with large substituents such as acetyl groups at the C4 position. Examination of the structures with trichothecene and D3G suggested seven residues that might influence binding. Three of these seven side chains, Phe 199, Gln 202, and Ser 203, are located on an α-helix that forms the back of the acceptor pocket. His 122 and Leu 123 are on a loop located in the binding pocket. Gln 143 and Ala 384 are on two separate loops in the binding pocket. Steady-state kinetic constants were determined for eight mutant proteins where these included single changes and combinations thereof. The effect of each of these mutations is discussed below.

[00126] Role of Gln 202. Gln 202 is located on an α-helix in the back of the acceptor binding pocket. In the presently disclosed model of DON binding in the acceptor pocket the carboxamide oxygen of Gln 202 is within hydrogen bonding distance (2.3 Å) of the C7 hydroxyl of DON. To examine the contribution of this residue to specificity it was changed to a glutamate, alanine, and leucine. The K_M value of Os79 Q202E is 17.5-fold higher than the K_M of WT with DON as a substrate where this is accompanied with a small increase in k_{cat}. The Q202E substitution maintains a similar residue size at the 202 position but adds a negative charge. The Q202E substitution may contribute a hydrogen bond to the enzyme substrate complex. To test whether there is a hydrogen bond that plays a role in DON binding between this hydroxyl and Gln 202 the

kinetic constants for Os79 Q202A were determined. There is no significant difference in the K_M value of Os79 Q202A compared to WT, indicating that either Gln 202 does not play a role in DON binding or that a water molecule can substitute the place of the side chain. Interestingly, the Q202L substitution decreases the K_M for DON by 4.7-fold without a major change in the value of k_{cat}, which is in contrast to the increase generated by the Q202E mutation. Together these substitutions emphasize the impact that substitutions at position 202 can have on substrate and product binding. In summary, these mutations suggest that a polar interaction in this position is not important for activity and that an increased charge is detrimental.

[00127] Role of Phe 199. Phe 199 is located in the α-helix one helical turn away from Gln 202 and lies at the top of the active site and makes a substantial contribution to the primarily hydrophobic acceptor binding pocket. The side chain is about 5 Å above the hydrophobic trichothecene backbone. Changing this residue to glutamine maintains a residue with approximately the same volume but with much greater polarity. The introduction of the polar glutamine in the place of the hydrophobic phenylalanine eliminates enzymatic activity as measured in the coupled-continuous enzymatic assay with DON as a substrate. This, along with the fact that this residue is a conserved Phe in all UGTs that have activity towards DON, highlights the importance of this residue as a component of the acceptor binding pocket.

[00128] Role of Gln 143. Gln 143 is located on a loop in the acceptor binding pocket. The Oε oxygen of Gln 143 is 3.5 Å from the C6 oxygen of the glucose moiety on U2F. Even though this is somewhat on the long-side for a substantial hydrogen bond this side chain appears to play an important role in substrate binding. The Q143A substitution does not demonstrate activity in the coupled-continuous enzymatic assay with DON as a substrate.

[00129] Role of Ser 203. Ser 203 is located on the bottom of the α-helix in the back of the acceptor binding pocket adjacent to Gln 202. This side chain was changed to alanine to examine the role of a polar residue in this position. The K_M value of the S203A mutant is similar to that of WT Os79. In order determine if a small residue is important at this position, the S203L substitution was created. There is a10-fold increase in the K_M of the S203L mutant but no change in k_{cat}, indicating that a bulkier residue at this position might hinder DON association or disassociation from the active site. The hydroxyl of S203 is 6.7 Å away from the closest carbon of DON (C4) and given that the change to an alanine has little effect it appears unlikely that it interacts directly with the substrate. It is more likely that a small residue is required at position 203 to allow Gln 202 to maintain a productive orientation. The 5-fold decrease in K_M as a result of the Q202L mutation illustrates the importance of that side chain. This suggests that the S203L mutation might prompt a change in the position of Gln 202 that could increase the K_M for DON.

[00130] Role of Ala 384. Ala 384 is positioned on a loop in the acceptor binding pocket about 5 Å from C3 of DON. To examine whether a hydrophobic residue at this position is important for activity this residue was changed to a serine. The K_M value of the A384S mutant is similar to that of WT Os79. Similar to Ser 203, this appears to be a second sphere residue, where changes in polarity appear to have little influence on activity towards DON.

[00131] Structure of Os79 Q202A and Os79 H122A/L123A proteins. The kinetic measurements with a variety of trichothecene accepters discussed above revealed that the active site of Os79 is unable to accommodate the C4 acetyl group of trichothecene substrates. The preceding mutations indicate the importance of the helix that carries Phe 199, Gln 202, and Ser 203 where this α-helix is opposite and slightly above C4, C15, and C8 of the trichothecene. Inspection of the acceptor binding pocket revealed that His 122 and Leu 123 are opposite but below

Phe 199, Gln 202, and Ser 203, with His 122 positioned only 4.2 Å away from C4. It is unlikely that 4.2 Å leaves sufficient space to accommodate an acetyl group. Although Leu 123 is positioned further from C4, its close proximity to C15 and C8 was viewed as a potential issue if the substrate needed to shift slightly in the active site to accommodate the acetyl moiety on C4. Consequently, His 122 and Leu 123 were simultaneously targeted for mutagenesis. Initially four amino acid substitutions provided Os79 the ability to glycosylate T-2 toxin as measured by an endpoint assay and analyzed by LC-MS/MS as described herein. These were the single substitutions of Q202L, Q202A, and Q202V, and the double substitution of H122A/L123G. The catalytic efficiency and K_M values of Os79 H122A/L123G are 5.08 x 10^2 s^{-1} M^{-1} and 512 μM with T-2 toxin as a substrate and 1.99 10^2 s^{-1} M^{-1} and 2512 μM with DON as a substrate. Activity for Os79 Q202 with T-2 toxin as a substrate was not detected with the assay used in this example, possibly due to a very high K_M. Os79 H122A/L123G was largely insoluble. Based on these findings, Os79 H122A/L123A was expressed and purified. Changing Leu 123 to an alanine instead of glycine improved the solubility of the protein. The catalytic efficiency of the H122A/L123A mutant with T-2 as a substrate is 1.71 x 10^3, which is a 3.5-fold increase over WT. The K_M value of 926 μM is similar to the H122A/L123G mutant. To improve the K_M, the H122A/L123A substitutions were combined with the Q202A substitution to make Os79 H122A/L123A/Q202A. The K_M of the triple substitution with T-2 as a substrate is 89 μM and the catalytic efficiency is 2.84 x 10^4, these values represent a 10-fold decrease in K_M and a 20-fold increase in catalytic efficiency compared to the H122A/L123A double mutant. Similar kinetic parameters are observed with DAS as a substrate. The K_M of the H122A/L123A/Q202A triple mutant with 4-ANIV as a substrate is 501 μM, which is a 5.5-fold increase compared with T-2 toxin and DAS. Interestingly, the K_M value of the H122A/L123A/Q202A triple mutant with DON as a substrate is 1202 μM, and the catalytic

efficiency is 7.4 x 10^2, these values represent a 20-fold increase and 23-fold decrease respectively compared to WT. It is clear that although Os79 H122A/L123A/Q202A is capable of glycosylating a broader range of substrates than WT Os79, the triple substitution comes at the cost of decreasing the catalytic efficiency with DON as a substrate. Given the observation that the Q202L substitution decreases the K_M for WT Os79 with respect to DON as a substrate, Os79 H122A/L123A/Q202L was made in an attempt to produce an enzyme with intermediate K_M values for both DON and T-2 toxin. The K_M values for Os79 H122A/L123A/Q202L are 118, 261, and 88 μM for T-2 toxin, DON, and 4-ANIV respectively.

[00132] The mutations all suggest that the volume of the active site of Os79 is an important determinant in broadening specificity by allowing the acceptor binding pocket to accommodate the C4 acetyl group. However, it could be possible that the mutations cause a structural change in the acceptor binding pocket. To address this question, the structures of Os79 H122A/L123A (PDB: 6BK2) and Os79 Q202A (PDB: 6BK0) were determined in the presence of UDP to a resolution of 1.47 Å and 1.29 Å respectively (Table 3). These structures show that the mutations result in very little change in the overall structures of the protein. The 50 residues that line and surround the acceptor binding pocket of the Os79 Q202A and H122A/L123A structures aligned to the corresponding residues of the Os79 WT structure with an rmsd of 0.12 and 0.16 Å respectively for structurally equivalent α-carbons. This indicates that the framework of the acceptor binding pockets of these proteins are almost identical to the WT enzyme. The main difference is the size of the trichothecene binding pocket. The H122A/L123A and Q202A mutations increase the volume of the active site. Specifically, these substitutions open up the side of the active site that accommodates the C4 acetyl group. This shows that the broad specificity is sanctioned by the

general hydrophobicity and volume of the acceptor cavity, which affords the C4 acetyl group enough space to allow toxins such at 4-ANIV, DAS, and T-2 to bind.

[00133] As shown here, the wild-type trichothecene UDP-glucosyltransferase from rice, Os79, has a broad specificity that can modify substrates that differ in molecular weight by a factor of 1.8 (isotrichodermol and HT-2 toxin; 250.3 and 424.5 respectively) with catalytic efficiencies over 1 x 10^4 s^{-1} M^{-1}. The wild-type enzyme is unable to glycosylate T-2 toxin and yet it readily modifies HT-2, which compared to T-2 is deacetylated at the C4-position. Indeed, as demonstrated herein, the wild-type enzyme is unable to accommodate substrates that are substituted at the C4 position.

[00134] The three dimensional structure of the product complex (Os79·UDP·D3G) in combination with the structure of trichothecene bound to Os79 revealed that the acceptor pocket is mostly hydrophobic and includes only a few residues capable of forming hydrogen bonds. Mutagenesis of these polar residues that might interact with the trichothecene substrate had small effects on k_{cat} and K_M, whereas mutagenesis of Phe 199 to a glutamine eliminated activity. These observations confirmed that the hydrophobicity and volume of the active site are primary factors in substrate specificity. Based on this structural knowledge, the volume of the active site was increased by mutagenesis with the consequence that Os79 H122A/L123A showed excellent activity towards T-2 toxin but reduced activity towards DON. Addition of the Q202L substitution created an enzyme that is a compromise which allows essentially equivalent activity towards both DON and T-2 toxin. The broad specificity of Os79 H122A/L123A/Q202L, as well as other mutations disclosed herein, makes it extremely useful for incorporation in transgenic plants that are susceptible to infection of both T-2 toxin and DON producing *Fusarium* species, like maize (*F. sporotrichioides* and *F. graminearum*), oat (*F. langsethiae* and *F. culmorum*), or potatoes (*F. sambucinum* and *F. graminearum*) (see Example 12, below).

Example 10: UGT Sequences From Additional Species

[00135] The gene for the wild-type (WT) HvUGT13248 was amplified by PCR from barley genomic DNA using the forward primer 5'- ATGGAGACCACGGTCACC-3' (SEQ ID NO:44) and the reverse primer 5'-TTATATTGACGAATACTTGGTAGCGAATT-3' (SEQ ID NO:45). The resulting product was introduced into the plasmid pKLD116 as described above for Os79. The H132A/L133A mutation was introduced to the WT gene using a single primer, PCR based method based on "Quikchange" mutagenesis (*33, 34*). The following primer was used: 5'-CCCGTGCGCGTGCTGGTGTACGACGCTGCGGCGGC

GTGGGCACGGCGGGTGGCACA-3' (SEQ ID NO:43).

[00136] HvUGT13248 WT (DNA sequence SEQ ID NO:19, protein sequence SEQ ID NO:20) and H132A/L1233A (DNA sequence SEQ ID NO:21, protein sequence SEQ ID NO:22) were overexpressed and purified using the same procedure as described above for Os79 with the exception that MBP was not cleaved from HvUGT13248. HvUGT13248 WT and H132A/L133A activity with different trichothecene substrates was determined with an end-point assay. The procedure for this activity assay was identical to the coupled assay performed with Os79 and described above with the exception that only a single, high concentration (~2 mM) of toxin was used and the final enzyme concentration was around 50 μM. HvUGT13248 WT exhibited no activity above the control with T-2 toxin as the substrate. HvUGT13248 H132A/L133A exhibited clear activity above the control with T-2 toxin as the substrate. Both enzymes exhibited activity with DON as the substrate.

Example 11: Sequence Alignments and Structural Predictions for UGTs that Exhibit DON UDP-glucosyltransferase Activity

[00137] Eight wild-type UDP-glucosyltransferases including Os79 have been shown to exhibit activity towards DON as demonstrated by introduced resistance in yeast and isolation of the reaction product for Os79 and HvUGT13248: UDP-glycosyltransferase superfamily protein from *Arabidopsis thaliana* ((*15*); DNA sequence SEQ ID NO:23; protein sequence SEQ ID NO:24), DON-glucosyltransferase 1 from *Arabidopsis thaliana* ((*12, 15, 20*); DNA sequence SEQ ID NO:25; protein sequence SEQ ID NO:26), UDP-glucosyl transferase 73C6 from *Arabidopsis thaliana* ((*15*); DNA sequence SEQ ID NO:27; protein sequence SEQ ID NO:28), predicted crocetin glucosyltransferase 2-like isoform X1 from *Brachypodium distachyon* ((*20*); DNA sequence SEQ ID NO:29; protein sequence SEQ ID NO:30), Os79 ((*20, 31, 37*); codon optimized DNA sequence SEQ ID NO:1; protein sequence SEQ ID NO:2, WT DNA sequence SEQ ID NO:46), hypothetical protein from *Sorghum bicolor* (sorghum) ((*20*); DNA sequence SEQ ID NO:31; protein sequence SEQ ID NO:32), predicted UDP-glycosyltransferase 74E2-like from *Brachypodium distachyon* ((*20*); DNA sequence SEQ ID NO:33; protein sequence SEQ ID NO:34), and *Hordeum vulgare* subsp. vulgare UDP-glucosyltransferase HvUGT13248 from *Hordeum vulgare* subsp. vulgare (domesticated barley;(*20*); DNA sequence SEQ ID NO:19; protein sequence SEQ ID NO:20). All of these enzymes belong to the glucosyltransferase family of proteins and based on sequence similarity are predicted to exhibit a GT-B fold (shown to be true for Os79). Molecular modeling shows that all of these enzymes will have an active site architecture closely related to that observed in Os79. This means that the active site architecture as defined above exhibits an rms difference of less than 2Å from that observed in Os79 for the structurally equivalent alpha carbons surrounding the active site. The amino acid side chains that

line the active site pocket in Os79 and these related enzymes are similar, which is consistent with their specificity towards DON. Even though the sequence identity might be as low as ~30%, the active site architecture (structural mechanism and specificity) is conserved. It is widely accepted that parallel changes within an enzyme family will (to a first approximation) produce the same results.

[00138] Clustal alignment of the amino acid sequences for these enzymes is shown in FIG. 2. The key for the sequence identifiers is shown in Table 4. This shows that the residues in those enzymes that are equivalent to H122 and L123 in Os79 are similar to those in Os79 and are inconsistent with activity towards T-2 in those wild-type enzymes. Thus the Clustal and structural alignment for H122 and L123 in Os79 agrees, even for those orthologs with low sequence similarity. The corresponding amino acids in these enzymes is shown in Table 5.

[00139] The Clustal and structural alignment for Q202 in Os79 do not completely agree for the orthologs with low sequence similarity such that prediction of the corresponding residue to Q202 based on sequence alone is inaccurate. The structurally equivalent positions are shown in Table 6 based on homology models.

Example 12: Transgenic Plants Incorporating the Disclosed Sequences

[00140] It has previously been shown that introduction of the barley glucosyltransferase gene (HvUGT13248) into wheat confers resistance to fungal head blight caused by *Fusarium graminearum* (*16*). However, as detailed herein above, the barley glucosyltransferase gene (HvUGT13248) does not have activity towards T-2 toxin or DAS. Thus introduction of a variant of HvUGT13248, or any of the other variant sequences disclosed herein that have been shown to glycosylate T-2 toxin into a plant will provide such a plant with resistance to fungal species that

synthesize T-2 and related toxins. Likewise, introduction of the Os79 variants described herein, as well as other variant sequences disclosed herein, into a plant, for example maize, oat, wheat, potatoes or barley, will result in the same or better protection against FHB caused by *Fusarium graminearum* and *Fusarium species* that synthesize T-2 toxin, due to the higher enzymatic activity of these variant sequences towards DON or T-2 toxin compared to that of HvUGT13248. Furthermore, the type II resistance correlates with the enzymatic activity, and hence Os79 variants, as well as the other variants disclosed herein, will provide more resistance due to their higher enzymatic activity. This is because trichothecene mycotoxins are known virulence factors for infection where reduction of levels of active toxin lead to lower fungal infection (*41-43*).

[00141] The term "about" is used herein to indicate that a value includes the standard deviation of error for the device or method being employed to determine the value. The use of the term "or" in the claims is used to mean "and/or" unless explicitly indicated to refer to alternatives only or the alternatives are mutually exclusive, although the disclosure supports a definition that refers to only alternatives and to "and/or." When not used in conjunction closed wording in the claims or specifically noted otherwise, the words "a" and "an" denote "one or more."

[00142] The terms "comprise," "have," and "include" are open-ended linking verbs. Any forms or tenses of one or more of these verbs, such as "comprises," "comprising," "has," "having," "includes," and "including," are also open-ended. For example, any method that "comprises," "has" or "includes" one or more steps is not limited to possessing only those one or more steps and also covers other unlisted steps. Similarly, any cell that "comprises," "has" or "includes" one or more traits is not limited to possessing only those one or more traits and covers other unlisted traits.

[00143] While the disclosure has been described in connection with specific embodiments thereof, it will be understood that the present disclosure is capable of further modifications by one of skill in the art. It is to be understood that, unless otherwise indicated, the present disclosure is not limited to particular materials, reagents, reaction materials, manufacturing processes, or the like, as such can vary. It is also to be understood that the terminology used herein is for purposes of describing particular embodiments only, and is not intended to be limiting. It is also possible in the present disclosure that steps can be executed in different sequence where this is logically possible. The present disclosure is therefore intended to encompass any variations, uses, or adaptations of the invention following, in general, the principles of the invention and including such departures from the present disclosure as come within known or customary practice within the art to which the invention pertains and as may be applied to the essential features herein before set forth.

[00144] All of the compositions and/or methods disclosed and claimed herein can be made and executed without undue experimentation in light of the present disclosure. While the compositions and methods of this disclosure have been described in terms of preferred embodiments, it will be apparent to those of skill in the art that variations may be applied to the compositions and/or methods and in the steps or in the sequence of steps of the method described herein without departing from the concept, spirit and scope of the invention. More specifically, it will be apparent that certain agents that are both chemically and physiologically related may be substituted for the agents described herein while the same or similar results would be achieved. All such similar substitutes and modifications apparent to those skilled in the art are deemed to be within the spirit, scope and concept of the invention as defined by the appended claims.

[00145] All publications, patents, patent publications, and nucleic acid and amino acid sequences cited are incorporated by reference herein in their entirety for all purposes.

What Is Claimed Is:

1. A polynucleotide comprising a DNA sequence encoding a modified UDP-glycosyltransferase polypeptide, said polypeptide comprising at least a first mutation relative to a wild-type UDP-glycosyltransferase polypeptide, wherein mutation renders the modified UDP-glycosyltransferase polypeptide capable of glycosylating T-2 toxin, diacetoxyscirpenol (DAS), 4-acetyl-NIV (4-ANIV), and 4,15-diacetyl-NIV (4,15-diANIV) from *Fusarium*.

2. The polynucleotide of claim 1, wherein said modified UDP-glycosyltransferase polypeptide comprises at least two mutations relative to a wild-type UDP-glycosyltransferase polypeptide.

3. The polynucleotide of claim 2, wherein said modified UDP glycosyl transferase polypeptide comprises a mutation at positions corresponding to amino acids 122 and 123 of SEQ ID NO:2.

4. The polynucleotide of claim 3, wherein said modified UDP glycosyltransferase polypeptide comprises the amino acid sequence of SEQ ID NO:4, SEQ ID NO:6 or SEQ ID NO:22.

5. The polynucleotide of claim 1, wherein said modified UDP-glycosyltransferase polypeptide comprises at least three mutations relative to a wild-type UDP-glycosyltransferase polypeptide.

6. The polynucleotide of claim 5, wherein said modified UDP glycosyltransferase polypeptide comprises a mutation at positions corresponding to amino acids 122, 123 and 202 of SEQ ID NO:2.

7. The polynucleotide of claim 6, wherein said modified UDP glycosyltransferase polypeptide comprises the amino acid sequence of SEQ ID NO:8 or SEQ ID NO:10.

8. The polynucleotide of claim 1, wherein said polynucleotide is operably linked to a regulatory element.

9. The polynucleotide of claim 5, wherein said regulatory element is a heterologous regulatory element.

10. The polynucleotide of claim 9, wherein said regulatory element is a promoter.

11. The polynucleotide of claim 10, wherein said promoter is functional in a plant.

12. The polynucleotide of claim 10, wherein said promoter is an inducible promoter.

13. A polypeptide encoded by the polynucleotide of claim 1.

14. The polypeptide of claim 13, wherein said polynucleotide encodes the amino acid sequence of SEQ ID NO:4, SEQ ID NO:6, SEQ ID NO:8, SEQ ID NO:10 or SEQ ID NO:22.

15. The polypeptide of claim 14, wherein said polynucleotide comprises the nucleic acid sequence of SEQ ID NO:3, SEQ ID NO:5, SEQ ID NO:7, SEQ ID NO:9 or SEQ ID NO:21.

16. A plant, plant part, cell, or seed comprising the polynucleotide of claim 1.

17. The plant, plant part, cell, or seed of claim 16, defined as a monocotyledonous plant, seed, cell, or plant part.

18. The plant, plant part, cell, or seed of claim 17, wherein said monocotyledonous plant is selected from the group consisting of: corn, wheat, rice, barley, oats and sorghum.

19. The plant, plant part, cell, or seed of claim 16, defined as a dicotyledonous plant, seed, cell, or plant part.

20. The plant, plant part, cell, or seed of claim 19, wherein said dicotyledonous plant is selected from the group consisting of: soybean, alfalfa, sunflower, cotton, canola, sweet potato, tomato, banana, curcubits, peppers and sugar beet.

21. A method of increasing the resistance of a plant to *Fusarium* infection, comprising expressing in said plant the polynucleotide of claim 1.

22. The method of claim 21, wherein said plant is defined as a monocotyledonous plant.

23. The method of claim 21, wherein said plant is defined as a dicotyledonous plant.

24. The method of claim 21, comprising transforming said plant with said polynucleotide and regenerating the plant therefrom.

25. The method of claim 21, comprising crossing a parent plant comprising said polynucleotide with itself or a second plant to obtain the plant in which resistance to *Fusarium* infection is increased.

26. A method of increasing the resistance of a plant to *Fusarium* infection, comprising modifying a UDP-glycosyltransferase polypeptide of said plant.

27. The method of claim 26, wherein said modifying comprises mutating at least one amino acid of said UDP-glycosyltransferase polypeptide.

28. The method of claim 26, wherein said modifying comprises site-specific mutagenesis.

29. The method of claim 28, wherein said site-specific mutagenesis comprises the use of single primer, zinc finger nucleases (ZFN), TALEN, or CRISPR technology.

References

1. Starkey, D. E., Ward, T. J., Aoki, T., Gale, L. R., Kistler, H. C., Geiser, D. M., Suga, H., Toth, B., Varga, J., and O'Donnell, K. (2007) Global molecular surveillance reveals novel *Fusarium* head blight species and trichothecene toxin diversity, *Fungal Genet Biol 44*, 1191-1204.

2. van der Lee, T., Zhang, H., van Diepeningen, A., and Waalwijk, C. (2015) Biogeography of *Fusarium graminearum* species complex and chemotypes: a review, *Food Addit Contam A 32*, 453-460.

3. Cundliffe, E., Cannon, M., and Davies, J. (1974) Mechanism of Inhibition of Eukaryotic Protein-Synthesis by Trichothecene Fungal Toxins, *P Natl Acad Sci USA 71*, 30-34.

4. Pestka, J. J. (2010) Deoxynivalenol: mechanisms of action, human exposure, and toxicological relevance, *Arch Toxicol 84*, 663-679.

5. McCormick, S. P., Stanley, A. M., Stover, N. A., and Alexander, N. J. (2011) Trichothecenes: From Simple to Complex Mycotoxins, *Toxins 3*, 802-814.

6. Anderson, D. W., R. M. Black, C. G. Lee, C. Pottage, R. L. Rickard, M. S. Sandford, T. D. Webber, and N. E. Williams. . (1989) Structure-activity studies of trichothecenes: Cytotoxicity of analogs and reaction products derived from T-2 toxin and neosolaniol, *Journal of Medicinal Chemistry 3*, 555-562.

7. Larsen, J. C., Hunt, J., Perrin, I., and Ruckenbauer, P. (2004) Workshop on trichothecenes with a focus on DON: summary report, *Toxicol Lett 153*, 1-22.

8. Proctor, R. H., Hohn, T. M., and Mccormick, S. P. (1995) Reduced Virulence of *Gibberella zeae* Caused by Disruption of a Trichothecene Toxin Biosynthetic Gene, *Mol Plant Microbe In 8*, 593-601.

9. Jansen, C., von Wettstein, D., Schafer, W., Kogel, K. H., Felk, A., and Maier, F. J. (2005) Infection patterns in barley and wheat spikes inoculated with wild-type and trichodiene synthase gene disrupted *Fusarium graminearum*, *P Natl Acad Sci USA 102*, 16892-16897.

10. Lemmens, M., Scholz, U., Berthiller, F., Dall'Asta, C., Koutnik, A., Schuhmacher, R., Adam, G., Buerstmayr, H., Mesterhazy, A., Krska, R., and Ruckenbauer, P. (2005) The ability to detoxify the mycotoxin deoxynivalenol colocalizes with a major quantitative trait locus for *fusarium* head blight resistance in wheat, *Mol Plant Microbe In 18*, 1318-1324.

11. Gardiner, S. A., Boddu, J., Berthiller, F., Hametner, C., Stupar, R. M., Adam, G., and Muehlbauer, G. J. (2010) Transcriptome Analysis of the Barley-Deoxynivalenol Interaction: Evidence for a Role of Glutathione in Deoxynivalenol Detoxification, *Mol Plant Microbe In 23*, 962-976.

12. Poppenberger, B., Berthiller, F., Lucyshyn, D., Sieberer, T., Schuhmacher, R., Krska, R., Kuchler, K., Glossl, J., Luschnig, C., and Adam, G. (2003) Detoxification of the *Fusarium* mycotoxin deoxynivalenol by a UDP-glucosyltransferase from *Arabidopsis thaliana*, *J Biol Chem 278*, 47905-47914.

13. Poppenberger, B., Berthiller, F., Bachmann, H., Lucyshyn, D., Peterbauer, C., Mitterbauer, R., Schuhmacher, R., Krska, R., Glossl, J., and Adam, G. (2006) Heterologous expression of *Arabidopsis* UDP-glucosyltransferases in *Saccharomyces cerevisiae* for production of zearalenone-4-*O*-glucoside, *Appl Environ Microb 72*, 4404-4410.

14. Ma, L. L., Shang, Y., Cao, A. Z., Qi, Z. J., Xing, L. P., Chen, P. D., Liu, D. J., and Wang, X. E. (2010) Molecular cloning and characterization of an up-regulated UDP-glucosyltransferase gene induced by DON from *Triticum aestivum* L. cv. Wangshuibai, *Mol Biol Rep 37*, 785-795.

15. Schweiger, W., Boddu, J., Shin, S., Poppenberger, B., Berthiller, F., Lemmens, M., Muehlbauer, G. J., and Adam, G. (2010) Validation of a Candidate Deoxynivalenol-Inactivating UDP-Glucosyltransferase from Barley by Heterologous Expression in Yeast, *Mol Plant Microbe In 23*, 977-986.

16. Li, X., Shin, S., Heinen, S., Dill-Macky, R., Berthiller, F., Nersesian, N., Clemente, T., McCormick, S., and Muehlbauer, G. J. (2015) Transgenic Wheat Expressing a Barley UDP-Glucosyltransferase Detoxifies Deoxynivalenol and Provides High Levels of Resistance to *Fusarium graminearum*, *Mol Plant Microbe In 28*, 1237-1246.

17. Achnine, L., Huhman, D. V., Farag, M. A., Sumner, L. W., Blount, J. W., and Dixon, R. A. (2005) Genomics-based selection and functional characterization of triterpene glycosyltransferases from the model legume *Medicago truncatula*, *Plant J 41*, 875-887.

18. Caputi, L., Malnoy, M., Goremykin, V., Nikiforova, S., and Martens, S. (2012) A genome-wide phylogenetic reconstruction of family 1 UDP-glycosyltransferases revealed the expansion of the family during the adaptation of plants to life on land, *Plant J 69*, 1030-1042.

19. Ross, J., Li, Y., Lim, E. K., and Bowles, D. J. (2001) Higher plant glycosyltransferases, *Genome Biol 2*.

20. Schweiger, W., Pasquet, J. C., Nussbaumer, T., Paris, M. P. K., Wiesenberger, G., Macadre, C., Ametz, C., Berthiller, F., Lemmens, M., Saindrenan, P., Mewes, H. W., Mayer, K. F. X., Dufresne, M., and Adam, G. (2013) Functional Characterization of Two Clusters of *Brachypodium distachyon* UDP-Glycosyltransferases Encoding Putative Deoxynivalenol Detoxification Genes, *Mol Plant Microbe In 26*, 781-792.

21. Cygler, M., Schrag, J. D., Sussman, J. L., Harel, M., Silman, I., Gentry, M. K., and Doctor, B. P. (1993) Relationship between Sequence Conservation and 3-Dimensional Structure in a Large Family of Esterases, Lipases, and Related Proteins, *Protein Sci 2*, 366-382.

22. Chivian, D., Kim, D. E., Malmstrom, L., Schonbrun, J., Rohl, C. A., and Baker, D. (2005) Prediction of CASP6 structures using automated Robetta protocols, *Proteins 61*, 157-166.

23. Roy, A., Kucukural, A., and Zhang, Y. (2010) I-TASSER: a unified platform for automated protein structure and function prediction, *Nat Protoc 5*, 725-738.

24. Yang, J. Y., Yan, R. X., Roy, A., Xu, D., Poisson, J., and Zhang, Y. (2015) The I-TASSER Suite: protein structure and function prediction, *Nat Methods 12*, 7-8.

25. Zhang, Y. (2008) I-TASSER server for protein 3D structure prediction, *Bmc Bioinformatics 9*.

26. Winn, M. D., Ballard, C. C., Cowtan, K. D., Dodson, E. J., Emsley, P., Evans, P. R., Keegan, R. M., Krissinel, E. B., Leslie, A. G. W., McCoy, A., McNicholas, S. J., Murshudov, G. N., Pannu, N. S., Potterton, E. A., Powell, H. R., Read, R. J., Vagin, A., and Wilson, K. S. (2011) Overview of the CCP4 suite and current developments, *Acta Crystallogr D 67*, 235-242.

27. Lairson, L. L., Henrissat, B., Davies, G. J., and Withers, S. G. (2008) Glycosyltransferases: Structures, functions, and mechanisms, *Annu Rev Biochem 77*, 521-555.

28. Russell, R. B., Saqi, M. A., Sayle, R. A., Bates, P. A., and Sternberg, M. J. (1997) Recognition of analogous and homologous protein folds: analysis of sequence and structure conservation, *J Mol Biol 269*, 423-439.

29. Xu, D. P., Mcelroy, D., Thornburg, R. W., and Wu, R. (1993) Systemic Induction of a Potato Pin2 Promoter by Wounding, Methyl Jasmonate, and Abscisic-Acid in Transgenic Rice Plants, *Plant Mol Biol 22*, 573-588.

30. Clancy, M., and Hannah, L. C. (2002) Splicing of the maize Sh1 first intron is essential for enhancement of gene expression, and a T-rich motif increases expression without affecting splicing, *Plant Physiol 130*, 918-929.

31. Michlmayr, H., Malachova, A., Varga, E., Kleinova, J., Lemmens, M., Newmister, S., Rayment, I., Berthiller, F., and Adam, G. (2015) Biochemical Characterization of a Recombinant UDP-glucosyltransferase from Rice and Enzymatic Production of Deoxynivalenol-3-*O*-beta-D-glucoside, *Toxins 7*, 2685-2700.

32. Rocco, C. J., Dennison, K. L., Klenchin, V. A., Rayment, I., and Escalante-Semerena, J. C. (2008) Construction and use of new cloning vectors for the rapid isolation of recombinant proteins from *Escherichia coli*, *Plasmid 59*, 231-237.

33. Chen, G. J., Qiu, N., Karrer, C., Caspers, P., and Page, M. G. P. (2000) Restriction site-free insertion of PCR products directionally into vectors, *Biotechniques 28*, 498-+.

34. van den Ent, F., and Lowe, J. (2006) RF cloning: A restriction-free method for inserting target genes into plasmids, *J Biochem Bioph Meth 67*, 67-74.

35. Minor, W., Cymborowski, M., Otwinowski, Z., and Chruszcz, M. (2006) HKL-3000: the integration of data reduction and structure solution - from diffraction images to an initial model in minutes, *Acta Crystallogr D 62*, 859-866.

36. 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 Appl Crystallogr 40*, 658-674.

37. Wetterhorn, K. M., Newmister, S. A., Caniza, R. K., Busman, M., McCormick, S. P., Berthiller, F., Adam, G., and Rayment, I. (2016) Crystal Structure of Os79 (Os04g0206600) from *Oryza sativa*: A UDP-glucosyltransferase Involved in the Detoxification of Deoxynivalenol, *Biochemistry-Us 55*, 6175-6186.

38. Adams, P. D., Afonine, P. V., Bunkoczi, G., Chen, V. B., Davis, I. W., Echols, N., Headd, J. J., Hung, L. W., Kapral, G. J., Grosse-Kunstleve, R. W., McCoy, A. J., Moriarty, N. W., Oeffner, R., Read, R. J., Richardson, D. C., Richardson, J. S., Terwilliger, T. C., and Zwart, P. H. (2010) PHENIX: a comprehensive Python-based system for macromolecular structure solution, *Acta Crystallogr D 66*, 213-221.

39. Emsley, P., and Cowtan, K. (2004) Coot: model-building tools for molecular graphics, *Acta Crystallogr D 60*, 2126-2132.

40. Murshudov, G. N., Vagin, A. A., and Dodson, E. J. (1997) Refinement of macromolecular structures by the maximum-likelihood method, *Acta Crystallogr D 53*, 240-255.

41. Muhitch, M. J., McCormick, S. P., Alexander, N. J., and Hohn, T. M. (2000) Transgenic expression of the TRI101 or PDR5 gene increases resistance of tobacco to the phytotoxic effects of the trichothecene 4,15-diacetoxyscirpenol, *Plant Sci 157*, 201-207.

42. Proctor, R. H., Hohn, T. M., and McCormick, S. P. (1995) Reduced virulence of *Gibberella zeae* caused by disruption of a trichothecene toxin biosynthetic gene, *Mol Plant Microbe Interact 8*, 593-601.

43. Wu, Q. H., Dohnal, V., Kuca, K., and Yuan, Z. H. (2013) Trichothecenes: Structure-Toxic Activity Relationships, *Curr Drug Metab 14*, 641-660.

Tables

Table 1. Listing of examples of amino acids belonging to each class.

Table 1 - Classes of Amino Acids	
Nonpolar	Ala, Val, Leu, Ile, Pro, Met, Phe, Trp, Gly, Cys
Uncharged Polar	Gly, Ser, Thr, Cys, Tyr, Asn, Gln
Acidic	Asp, Glu
Basic	Lys, Arg, His

Table 2.

Selected Trichothecene Contact residues in Os79 that define the binding site	**Secondary Structural Element carrying that residue**
F21	V16-F21 (β-strand and loop)
H27	H27-R36 (α-helix)
H122, L123	V116-A133 (β-strand, loop, and α-helix)
Q143	A136-Q143 (β-strand and loop)
Q202	P197-A208 (α-helix)
W383, A384	W383-S395 (Loop and α-helix)

Table 3.

Protein	Os79 Q202A·UDP	Os79 H122A/L123A·UDP	Os79 T291V·UDP	Os79·D3G·UDP
PDB ID	6BK0	6BK2	6BK1	6BK3
space group	$P2_12_12_1$	$P2_12_12_1$	$P2_12_12_1$	$P3_221$
Unit cell dimensions, Å	*a* = 59.4 *b* = 83.2 *c* = 99.1	*a* = 59.3 *b* = 82.9 *c* = 99.0	*a* = 59.4 *b* = 83.2 *c* = 98.7	*a* = 104.5 *b* = 104.5 *c* = 98.3
Wavelength (Å)	0.979	0.979	0.979	0.979
resolution range (Å)	50–1.47 (1.5-1.47)[a]	50–1.29 (1.31-1.29)[a]	50-1.58 (1.61-1.58)[a]	50–2.17 (2.21-2.17)[a]
reflections:	1095493	1504096	568452	382690
reflections:	79234	115974	63979	33082
redundancy	13.1 (12.6)	12.3 (8.5)	8.4 (6.2)	11.6 (8.6)
completeness (%)	99.2 (84.1)	99.2 (96.9)	99.8 (95.2)	100 (100)
average I/σ	41.3 (9.2)	34.9 (4.0)	83.6 (11)	29.1 (4.4)
R_{merge} (%)[b]	5.1 (23.1)	10.2 (0)	5.2 (18.0)	8.6 (55.5)
R_{work}	15.1	16.0	16.0	19.6
R_{free}	17.9	18.0	19.2	21.0
protein atoms	3499	3464	3407	3353
ligand atoms	36	36	36	56
water molecules	608	398	557	197
average B factors	19.3	17.9	27.3	35.1
Ramachandran	--	--	--	--
most favored	96.8	97.2	97.5	96.0
allowed	2.98	2.8	2.5	4.0
disallowed	0.23	0	0.0	0.0
rms deviations	--	--	--	--
bond lengths (Å)	0.022	0.021	0.013	0.023
bond angles (deg)	1.889	1.755	1.573	2.234

[a]Values in parenthesis are for highest resolution shell.

[b]$R_{merge}= \Sigma| I_{(hkl)} - I | \times 100 / \Sigma| I_{(hkl)} |$, where the average intensity I is taken over all symmetry equivalent measurements and $I_{(hkl)}$ is the measured intensity for a given observation.

Table 4.

Sequence Acronym in FIG. 1 and Table 3	Sequence Identifier	Location
Os_79_+ve_XP_01	XP_015635481.1 GI:1002262256	ncbi.nlm.nih.gov/protein/1002262256/
Sb06g002180_wea	XM_002447461.1	ncbi.nlm.nih.gov/gene/8067747
HvUGT13248_Barl	GU170355.1	ncbi.nlm.nih.gov/nuccore/289188049
Bradi5g03300.1_	XP_010239695.1	ncbi.nlm.nih.gov/protein/XP_010239695.1
Bradi5g02780.1_	XP_003581017.1	ncbi.nlm.nih.gov/protein/XP_003581017.1
AT_73C6_UGT_+we	NP_181217.1	ncbi.nlm.nih.gov/protein/NP_181217.1
DOGT1_73C5_+ve_	NP_181218.1	ncbi.nlm.nih.gov/protein/15228037?report=fasta
AT_73C4_+ve_Q9Z	NP_181215.1	ncbi.nlm.nih.gov/protein/15228033?report=fasta

Table 5.

Enzyme	**His122/L123**	**Percent Sequence Identity and Similarity to Os79**
Os_79_+ve_XP_01	H122 L123	100
Sb06g002180_wea	H128 L129	73, 83
HvUGT13248_Barl	H132 L133	73, 83
Bradi5g03300.1_	H123 V124	70, 80
Bradi5g02780.1_	H127 L128	70, 78
AT_73C6_UGT_+we	C132 L132	30, 47
DOGT1_73C5_+ve_	C130 L131	30, 47
AT_73C4_+ve_Q9Z	L132 L133	29, 46

Table 6.

Enzyme	**Residue No. Equiv. to Q202 in Os79 (Clustal)**	**Residue No. Equiv. to Q202 in Os79 (structural align. based on predicted models except for Os79)**	**Percentage Seq. Identity and Similarity to Os79**
Os_79_+ve_XP_01	Q202	Q202	100
Sb06g002180_wea	A206	A206	73, 83
HvUGT13248_Barl	V210	V210	73, 83
Bradi5g03300.1_	A203	A203	70, 80
Bradi5g02780.1_	A203	A203	70, 78
AT_73C6_UGT_+we	M213	I209	30, 47
DOGT1_73C5_+ve_	M213	I209	30, 47
AT_73C4_+ve_Q9Z	M214	F210	29, 46

Figures

Figure 1.

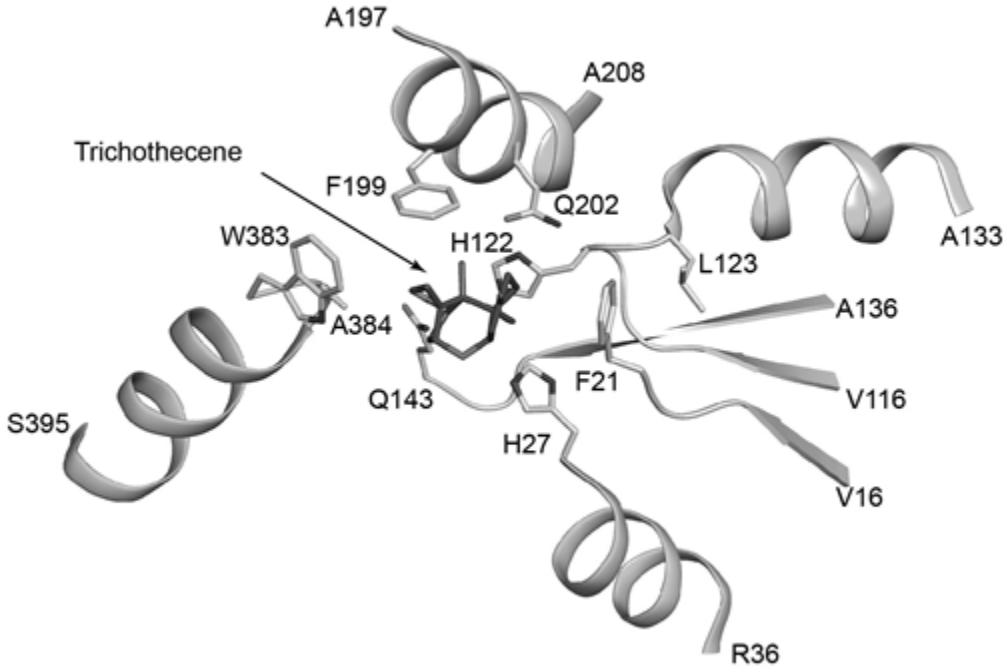


Figure 2.

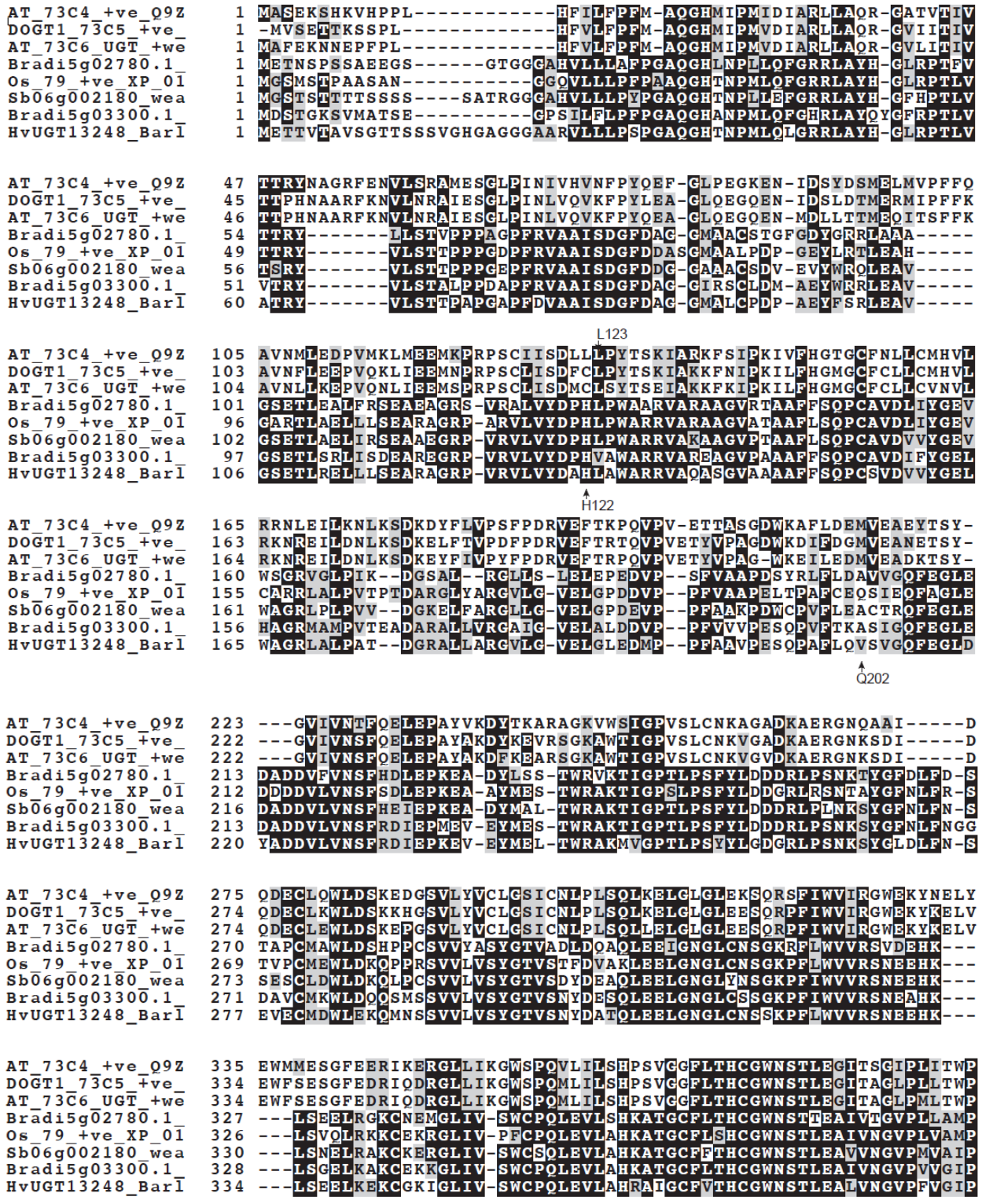
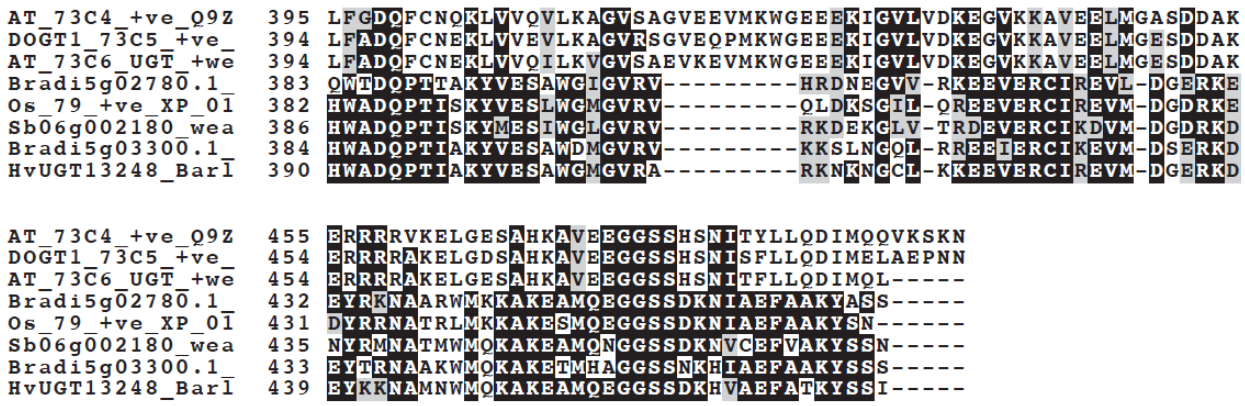


Figure 2 (con't)



Appendix II:

Functional characterization of a soluble NADPH-cytochrome P450 reductase from *Fusarium graminearum*

This work was performed in collaboration with Dr. Thomas Etzerodt, Karl Wetterhorn, Dr. Giuseppe Dionisio, and Dr. Ivan Rayment. Experiments were performed by Dr. Thomas Etzerodt with assistance from Karl Wetterhorn.

A version of this work was previously published as:

Etzerodt T, Wetterhorn K, Dionisio G, Rayment I. *Functional characterization of a soluble NADPH-cytochrome P450 reductase from* Fusarium graminearum. Protein expression and purification. 2017 Oct 1;138:69-75.

Abstract

Fusarium head blight is a devastating disease in wheat caused by some fungal pathogens of the *Fusarium* genus mainly *F. graminearum*, due to accumulation of toxic trichothecenes. Most of the trichothecene biosynthetic pathway has been mapped, although some proteins of the pathway remain uncharacterized, including an NADPH-cytochrome P450 reductase. We subcloned a *F. graminearum* cytochrome P450 reductase that might be involved in the trichothecene biosynthesis. It was expressed heterologously in *E. coli* as N-terminal truncated form with an octahistidine tag for purification. The construct yielded a soluble apoprotein and its incubation with flavins yielded the corresponding monomeric holoprotein. It was characterized for activity in the pH range 5.5–9.5, using thiazolyl blue tetrazolium bromide (MTT) or cytochrome c as substrates. Binding of the small molecule MTT was weaker than for cytochrome c, however, the rate of MTT reduction was faster. Contrary to other studies of cytochrome reductase proteins, MTT reduction proceeded in a cooperative manner in our studies. Optimum kinetic activity was found at pH 7.5–8.5 for both MTT and cytochrome c. This is the first paper presenting characterization of a cytochrome P450 reductase from *F. graminearum* which most likely is involved in mycotoxin biosynthesis or some primary metabolic pathway such as sterol biosynthesis in *F. graminearum*.

Abbreviations

CYP: Cytochrome P450

CPR: NADPH-cytochrome P450 reductase

FgCPR: *Fusarium graminearum* NADPH-cytochrome P450 reductase with N-terminal truncation (Δ1-28) and His8-tag

Cyt *c*: Cytochrome *c*

MTT: thiazolyl blue tetrazolium bromide

Introduction

Fusarium head blight is a devastating disease in wheat caused by infection fungal pathogens of the *Fusarium* genus, particularly *Fusarium graminearum*. The pathogen causes significant yield losses (*1*) due to the accumulation of trichothecene mycotoxins in the cereal heads with deoxynivalenol as the dominant trichothecene (*2*). In addition, trichothecenes acts as virulence factors promoting spread of the fungal pathogen within the cereal heads (*3, 4*).

The trichothecene biosynthetic pathway has been mapped extensively for *F. sporotrichioides* and *F. graminearum* (*5-8*). TRI5 converts farnesyl pyrophosphate to trichodiene (*6, 9*), which in turn is converted to isotrichodermol by TRI4, an NADPH-dependent cytochrome P450 (CYP) (*10, 11*). Other identified CYPs of the trichothecene biosynthetic pathway are TRI1, TRI11 and TRI13, which perform hydroxylation of C8, C15 and C4 positions on the trichothecene skeleton, respectively (*12-14*). Important CYPs characterized from *F. graminearum* not part of the trichothecene biosynthesis are three CYP51 paralogues (*15*), Fg08079 of the butenolide biosynthesis (*16*) and CLM2 of the culmorin biosynthesis (*17*). All of the above CYPs have been functionally characterized in fungi but have not yet been isolated and characterized *in vitro* so far.

CYPs require an NADPH-cytochrome P450 reductase (CPR) partner for electron transfer from NADPH to the CYP heme core. CPR proteins contain flavins, which mediate such electron transfer in a highly regulated fashion (*18, 19*). For *F. graminearum* only two CPR genes have been putatively identified (compared to 107 CYP genes) (*20*). Consequently, a single CPR must interact with numerous CYPs and therefore be involved in several diverse biosynthetic pathways. In eukaryotes, like *Fusarium* fungi, CYPs and their reductase partners are membrane bound via an N-terminal anchor, which is required for assembly and activity of CPR-CYP pairs in some cases

(*21*), even though some soluble truncated yeast and fungal CPRs retain their ability to reduce their CYP partners (*22-24*).

Comparison of kinetics for CPRs from different sources can be complicated since CPR-CYP complexes are highly variable and the specific CYP partner for the CPR of interest is not always available. Although not the natural substrate, *in vitro* reduction of Cyt *c* serves as a general electron acceptor to compare the activity of different CPR proteins (*22, 25, 26*). Other commonly used electron acceptors in activity studies of CPRs are ferricyanide (*27*), 1,1-diphenyl-2-picrylhydrazyl (*27*), and thiazolyl blue tetrazolium bromide (MTT), the latter having the best stability and highest molar absorptivity (*27, 28*).

The best studied reductase from the *Fusarium* genus is the nitric oxide reductase from *F. oxysporum* (*29-31*), but so far, no NADPH-cytochrome P450 reductases from *Fusarium* have been isolated for characterization. In this paper we present the characterization of a soluble N-terminally truncated CPR from *F. graminearum*.

Materials and methods

Selection of G-blocks for cloning of the *His8-FgCPRΔ1-28* fragment. A CPR gene has been identified for the *F. graminearum* PH-1 strain (*32*). Based on the ORF (~2200 bp), two G-block oligos were designed as C-t and N-t part of the CPR (see Table S 1 in suppl. mat.) with an internal overlap of 23 bp for initial PCR merging to get the full length *FgCPRΔ1-28* gene. Each G-block contained flanking sequences for restriction free (RF) cloning into a target vector. G-blocks, primers for cloning and Sanger sequencing were obtained from Integrated DNA Technologies (IDT, Coralville, IA, USA) and sequences are given in Table S 1 (of Suppl. Mat.).

Cloning and expression of the His8-FgCPRΔ1-28 construct. All cloning procedures were performed using RF cloning (*33, 34*). The *FgCPRΔ1-28* gene fragment was purified by gel

electrophoresis from a 1% agarose gel stained with SYBR Green I. The *FgCPRΔ1-28* fragment was cloned into a pTEV4 target vector with kanamycin resistance (from Novagen) downstream a His8-tag (for details see Suppl. Mat.). DH5α cells were transformed with the *FgCPRΔ1-28:pTEV4* construct. LB culture was grown from a single DH5α colony and plasmids were purified using a Spin Miniprep kit (Qiagen). The resulting *FgCPRΔ1-28* ORF was verified by Sanger sequencing.

CODON+ cells (Stratagene) were transformed with the *FgCPRΔ1-28:pTEV4* plasmid. Transformed *E. coli* CODON+ single colony was used to inoculate an overnight LB pre-culture which was used to inoculate a 10 L main expression LB culture. When OD_{600} reached 0.8–1 the 10 L culture was cooled to 16 °C. Protein expression was then induced with 1 mM IPTG and the culture was incubated at 16 °C for 24 h. Cells were harvested by centrifugation at 3000 *g* (4 °C). The pellet was frozen in liquid nitrogen and stored at −80 °C until lysis and protein purification.

Purification of apo His8-FgCPRΔ1-28 and incorporation of flavins. All purification steps were carried out on ice or at 4 °C. Buffer A was 50 mM Tris (pH 7.5), 0.1 mM EDTA, 1 mM TCEP, 10% glycerol. Frozen cells were re-suspended in 40 ml ice-cold buffer A plus 50 mM NaCl and 0.5 mg/ml lysozyme. The cell suspension was then lysed by sonication, clarified by centrifugation (at 80000 *g* and 4 °C for 30 min) and the NaCl concentration of the supernatant was brought to 300 mM prior to purification on a 6 ml NiNTA column (Qiagen). The unbound proteins (pass through) and the washed proteins (Buffer A with 300 mM NaCl and 20 mM imidazole) possessed only negligible amount of His-tagged proteins as evaluated by SDS-PAGE and western with anti-HIS5 antibodies (Qiagen, data not shown). The bound proteins were eluted in 30 ml elution buffer (30 mM Tris/HCl pH 7.5, 0.3 M NaCl and 300 mM Imidazole). The purified FgCPR apo protein was incubated with FMN and FAD (TCI Chemicals, Cambridge, MA) overnight at 4 °C in 10 and 5 M excess, respectively, based on the stoichiometry of a *S. cerevisiae* CPR (*35*).

Unbound flavins were removed by a second NiNTA purification and the protein solution was concentrated by ultrafiltration (Amicon Ultra-15 Centrifugal Filters, MWCO of 30 kDa; Merck Millipore Ltd, Billerica, MA) and then dialyzed against Buffer A without EDTA. Finally, the protein solution was flash frozen in liquid nitrogen (30 µl aliquots) and stored at −80 °C until analysis.

Purity, spectral analysis, and aggregation state of FgCPR. Protein concentration was determined spectrophotometrically from its absorbance at 280 nm (corrected for FAD and FMN absorbance) and confirmed by Bradford analysis. Flavin incorporation was assessed from spectral analysis (see Fig. 2) identical to CPRs in other studies (*24, 36*). Protein purity was assessed by SDS-PAGE. Briefly, FgCPR samples were boiled in 2xSDS buffer (0.1 M Tris (pH 8.0), 10 mM EDTA, 0.1 M DTT, 60% glycerol and 0.05% Bromo Phenol Blue) for 10 min, cooled to RT and 15 µl boiled sample was run on a NuPAGE 4–12% Bis-Tris Gel (ThermoFischer Scientific, Roskilde, Denmark).

Protein aggregation state was assessed by size exclusion chromatography (100 µl of 1.33 mg/ml protein) on an Amersham Superdex 200 HR 10/30 column (300 × 10 mm) using a 20 mM Tris, 1 mM TCEP, 10% ethylene glycol eluent (pH 7.5) with a flow rate of 0.5 ml/min. Apparent molecular weight was determined from comparison with protein markers (cytochrome *c*, carbonic anhydrase, bovine serum albumin, alcohol dehydrogenase and β-amylase from Sigma, St. Louis, MO).

Verification of primary sequence by LC-MS based proteomics. 50 µg FgCPR was denatured and reduced in 100 mM ammonium bicarbonate buffer, pH 8.0 (ABC) containing 3.84 M guanidinium chloride and 20 mM DTT, followed by alkylation with iodoacetamide (IAM, 23.8 mM). Excess IAM was then quenched with DTT. The samples were desalted by immobilization

on Supel-Tips C18 Micropipette tips (Sigma-Aldrich, Brøndby, Denmark) and the desalted fractions evaporated overnight. The dry pellet was redissolved in 100 mM ABC and incubated with 1 μg trypsin or chymotrypsin (Pierce, ThermoFisherScientific, Hvidovre, Denmark) at 37 °C for 2 h. Nano-LC-MS/MS analysis and data treatment was done as described in Dionisio *et al*. (*37*).

Flavin incorporation and stoichiometry. Samples (100 μl of 265 μg/ml FgCPR) were denatured in 1 M citric acid (pH 1.6) and ultrafiltrated to 25 μl in a VivaSpin 500 ultrafiltration units (GE Healthcare, Brøndby, Denmark). The concentrate was washed with citric acid followed by MQ and the flow-through (FT) fractions were pooled. The protein concentrate was transferred with 2 × 250 μl MQ and quantified by Bradford (using BSA as a standard). FTs (containing FAD and FMN) were purified by SPE (Sep-pak Vac C18, Waters, Hedehusene, Denmark) and analyzed (50 μl) by HPLC-DAD (from Agilent Technologies, Glostrup, Denmark) on a Hypersil BDS C18 column, 250 × 2.1 mm i.d., 5 μm (Fischer Scientific, Roskilde, Denmark) at 30 °C using a linear gradient of 10–37% B over 12 min with eluents A) 5 mM ammonium acetate (pH 6.2) and B) MeOH and a flow rate of 0.25 ml/min. FAD and FMN were detected at 450 nm and quantified using external calibration curves. Results were corrected for flavin losses via recovery analysis calculated from mixed flavin spikes (800 ng/ml final) in 1 M citric acid, analyzed as described above. All samples were performed in triplicates and repeated once.

FgCPR kinetics. Reduction of MTT and cytochrome *c* (Cyt *c*) was performed in 96-well plates and analyzed using a BioTek Instruments Epoch microplate spectrophotometer (Holm & Halby, Brøndby, Denmark). Both assays were performed in 150 mM buffer (citrate for pH 5.5, PIPES for pH 6.5, phosphate for pH 7.5, tricine for pH 8.5 and sodium carbonate for pH 9.5) at 23 °C with an NADPH regeneration system (2.1 mM glucose-6-P (G-6-P), 0.7 mM NADP+, 10 mM

MgCl2 and 0.7 U Glucose-6-P-Dehydrogenase (G6PDH), all obtained from Sigma-Aldrich, Brøndby, Denmark) which was charged for 30 min prior to initiation of the assay. For the MTT and Cyt *c* assays 20 nM FgCPR and 5 nM FgCPR were used, respectively. MTT or Cyt *c* substrate stock (10 times the final concentration) plus NADPH regeneration system in each buffer for the given pH were added to the wells and the plate was read prior to initiation of the assays to determine the initial absorbance (Aλ0c) for each concentration of substrate. Assays were initiated by addition of FgCPR to a final volume of 250 µl. Samples were read every 10 s for 10 min at 610 nm and 550 nm for MTT and Cyt *c* reduction, respectively. The difference in absorbance was calculated according to ΔA = Aλtc – Aλ0c where Aλtc is the absorbance measured at time t for substrate concentration c. Concentrations of reduced MTT and reduced Cyt *c* were calculated from the extinction coefficients, εred. MTT = 11.3 mM−1 and εred. Cyt *c* = 21.1 mM−1 and a sample path length of 0.7 cm. Non-linear regression analysis was performed using GraphPad Prism software v. 7.02 (San Diego, CA, USA). All experiments were performed in triplicate and repeated at least twice.

Results

Purity, spectral analysis, and aggregation state of FgCPR. NADPH-cytochrome P450 reductases from *Fusarium* sequence was selected from Uniprot.org (entry I1RZE7). This FgCPR sequence was chosen since it was the best characterized gene sequence and the only CPR sequence reviewed by experts for the Uniprot database. Submission of the full-length protein to TargetP server (http://www.cbs.dtu.dk/services/TargetP/) shows that the first 21 amino acids might belong to a leader peptide (LP) addressing the sequence to the secretion pathway. We removed this ER entry LP by PCR using the delta G-block synthetic gene construct as described in the Experimental section.

Forty milligrams of soluble N-terminally truncated (ΔN-t) protein was obtained from 16 g cells, which is considered a high expression level. Purity of the protein was assessed by SDS-PAGE analysis (Figure 1) with different amounts of protein loaded. Flavins were not incorporated into the apoform during expression in *E. coli* but were incorporated in vitro by incubation of the apoprotein with free flavins. UV/vis analysis (see Fig. 2) revealed associated flavins with the apo-proteins by spectra (in the region 300–500 nm) identical to that reported for other CPR flavoproteins (*18, 24*).

Tryptic and chymotrypic digestion of the FgCPR proved that protein sequence matched the expected sequence expressed from the *FgCPRΔ1-28* ORF (see Figure S 1 in Suppl. Mat.). The FgCPR eluted from the size exclusion column with a retention time of 26.5 min (Figure 3) corresponding to an apparent molecular weight of 89 kDa. The theoretical molecular weight of FgCPR is 77 kDa indicating that the protein exists in monomeric form.

Protein-flavin stoichiometry. Protein denaturation by heating or treatment with urea or guanidine did not released the bound flavins. Instead, citric acid (pH 1.6) was used for flavins release by temporary pH denaturation of the flavins binding pocket, and therefore, we could recover the protein quantitatively using VivaSpin ultrafiltration units. Recoveries of FAD and FMN, as assayed in the VivaSpin flow through, were determined to be 85% and 73%, respectively, by flavin quantifications using HPLC-DAD. Protein concentration by Bradford was determined to be 3.5 μM and FAD and FMN were quantified to 4.1 μM and 7.2 μM, respectively. This yields a final stoichiometry of 1:1:2 (FgCPR:FAD:FMN) in agreement with that reported for the yeast CPR by Lamb *et al.* (*35*).

FgCPR kinetics. Kinetics of the NADPH regeneration system for formation of NADPH from NADP+ was virtually unaffected at pH 6.5-9.5. We also tested pH 5.5 for the regeneration system, however, no activity was observed for FgCPR at this pH for reduction of MTT or Cyt *c* with the NADPH regeneration system (or free NADPH) (data not shown).

Cytochrome *c* assay. The kinetic data for Cyt *c* reduction by FgCPR were fitted by non-linear regression analysis (see Fig. 4, Table 1) using the regular Michaelis-Menten model of the GraphPad Prism software, from the equation:

$$V = \frac{V_{max} \cdot S}{K_m + S}$$

The assay was also tested at pH 6.5, however, no activity was observed below pH 7.5. Maximum binding (lowest K_M) and V_{max} was observed at pH 8.5. Since electrostatic interactions govern the binding between many cytochrome proteins to their reductase partners (*38-41*), the maximum rate and binding at pH 8.5 might be explained by an optimal electrostatic interaction. Increasing the pH further led to a decrease in both rate and binding of Cyt *c*.

Lamb *et al*. characterized full length and truncated yeast CPR using an NADPH regeneration system comparable to our studies except Lamb *et al*. did not include Mg2+ and they used 3 U of G6PDH (*22*) (we used 0.7 U G6PDH). The velocity was 9 times higher for their truncated *S. cerevisiae* CPR and the full-length reductase exhibited a Vmax comparable to our FgCPR for the reduction of Cyt *c*. Cyt *c* interaction with the *S. cerevisiae* reductase at pH 7.5 was much stronger (K_M of 1-1.5 μM) than in our studies (K_M of 44 μM). Furthermore, two full-length NADPH-cytochrome P450 reductase proteins from Capsicum annuum exhibited kinetic parameters similar to our studies (*42*), with a K_M of 81 μM in 100 mM K-phosphate buffer (pH 7.6).

Microsomal cytochrome P450 reductase from *Liza saliens* liver had strong binding of Cyt *c* with a K_M of 7.69 μM and a rate 10 times (V_{max} of 47.6 μmol/min/mg) (*43*) compared to our FgCPR. Activity of the *L. saliens* CPR was also studied as a function of pH and found an optimum in the range pH 7.4-7.8 with a rapid decrease in activity outside this range using free NADPH (final concentration 162 μM) as electron donor.

In our studies, FgCPR reduction of Cyt *c* by free NADPH (initial concentration 400 μM) was most active for pH 7.5, while only minor or no activity could be observed at other tested pH values (data not shown). When employing the NADPH regeneration system, the optimum pH was 8.5 and the FgCPR still retained considerable activity at pH 9.5. Reduction of Cyt *c* using free NADPH was best described by an allosteric sigmoidal model (goodness of fit 0.94) and resulted in a much stronger association between FgCPR and Cyt *c* (K_M of 16 ± 2.5 μM). This suggests that components of the regeneration system can alter binding of Cyt *c* to the FgCPR compared to free NADPH. A reduction in V_{max} from 6262 ± 274 nmol/min/mg with the NADPH regeneration system to 5198 ± 224 nmol/min/mg when free NADPH was used, most likely due to a lack of NADPH regeneration.

MTT assay. Kinetic data for CPR with MTT as a substrate yielded the best fit using an allosteric sigmoidal model of the GraphPad Prism software based on the equation

$$V = \frac{Vmax \cdot S^h}{{EC_{50}}^h + S^h}$$

where V is the CPR velocity, V_{max} is the maximum velocity for CPR (nmoles/min/mg), S is the substrate concentration (μM), h is the hill coefficient and EC_{50} is S resulting in 50% V_{max} (see Fig. 5, Table 2).

The sigmoidal fit with corresponding hill coefficients of 1.6–1.9 shows that the reduction of MTT by CPR proceeds in a cooperative manner even at pH 6.5 which was the lowest pH value where our FgCPR exhibited kinetic activity.

Yim *et al*. (*28*) observed a kinetic activity 3-4 times higher for rat recombinant CPR compared to our FgCPR at pH 7.5 using an NADPH regeneration system with components in concentrations comparable to our study except Mg2+. In our studies, K_M for MTT was higher than for Cyt *c*, but V_{max} for MTT was less affected by changes in pH. For pH 6.5 the velocity and binding to FgCPR for MTT was significantly decreased. This suggests that MTT binding is not governed primarily by electrostatic interactions as for Cyt *c*. V_{max} for MTT reduction, on the other hand, was affected by pH.

Discussion

FgCPR activity compared to other CPRs at pH 7.5. In general, we observed lower kinetic velocity (V_{max}) and weaker binding (higher K_M) compared to other truncated CPRs at pH 7.5. The reduced activity of our FgCPR compared to other truncated CPRs was initially suspected to stem from the N-terminal His8-tag, which was uncleavable in spite of the presence of a TEV cleavage site (data not shown). However, for the *S. cerevisiae* CPR, the N-terminal region is located distinctly from all co-factor and cytochrome binding sites (*44*), and the presence of the His8-tag should not directly interfere with NADPH or Cyt *c* binding. This is supported by studies of a tropinone reductase in which activity was strongly reduced when His-tagged at the C-terminus (the substrate binding site) but not with an N-terminal His-tag (opposite of the active site) (*45*). Thus, the N-terminal His8-tag should not influence activity of the FgCPR in our studies.

Since microsomal full-length reductase exhibits a V_{max} for Cyt *c* reduction comparable to that of our FgCPR, we speculate that the reduced rate of our FgCPR could be due to a non-optimal conformation of FgCPR. This would also explain the higher K_M (for Cyt *c*) for FgCPR compared to other CPRs. The comparison of a non-tagged to the FgCPR for activity and structure (e.g. X-ray) could reveal such information.

Another explanation might be that association with a membrane surface is more important for FgCPR for proper activity compared to other CPR proteins. Future studies with preparation of microsomal FgCPR could help clarify this.

Effect of substrate on kinetic activity. Interaction of CYP proteins and CPRs is governed by electrostatics. In addition, since membrane anchors are required for activity of some CPR-CYP pairs (*21*), restrictions for a proper co-alignment of such CPR-CYP pairs must apply. The small MTT only carries a single (permanent) positive charge, which explains why MTT binding to FgCPR is virtually unaffected by changes in pH. Furthermore, compared to Cyt *c*, fewer restrictions on alignment of MTT are required, which facilitates electron transfer and results in a higher V_{max}.

FgCPR reduction of Cyt *c* can be described using regular Michaelis-Menten kinetics, whereas for the smaller MTT an allosteric sigmoidal model yielded the best fit. The reason for this is not entirely clear, since MTT reduction by other CPRs was described using a regular Michaelis-Menten model (*28, 46*).

Our results from using Cyt *c* compared to MTT as substrates for FgCPR show that the reductase retains activity at alkaline pH but is sensitive to the type of substrate used.

Slightly alkaline pH yields optimal kinetics for the FgCPR. Theoretical pI values for Cyt *c* (equine heart) and FgCPR are 9.59 and 5.25, respectively, according to ExPASy (*47*). At pH

9.5, Cyt *c* will be neutrally charged and FgCPR negatively charged. This will abolish most electrostatic interactions between FgCPR and Cyt *c*. However, since FgCPR activity was still observed at pH 9.5, the proteins must be able to interact with each other. Perhaps the relatively high ionic strength (150 mM carbonate) partially shields electrostatic repulsions at pH 9.5 still allowing binding between FgCPR and Cyt *c*.

It is possible that the pH optimum for the membrane bound full length FgCPR will be different than for the truncated FgCPR, since kinetics differ for full length and truncated versions of yeast CPR (*22*). We also expressed and purified the full length FgCPR (data not shown) but were unable to incorporate flavins using the same approach as for the truncated FgCPR. Alternatively, full length CPR might be obtained from *F. graminearum* microsomes as for other CPRs (*43, 48, 49*). Kinetic studies of the full-length CPR from *F. graminearum* should be conducted with both detergent solubilization as well as reconstitution in liposomes mimicking the fungal membrane to evaluate the influence of the membrane system on protein activity. The sensitivity to substrate reduction illustrates that using only MTT or similar small molecule substrates might not be adequate for NADPH-cytochrome P450 reductase activity studies (where the natural substrate is a CYP) but might suffice for other types of reductases.

Intracellular pH changes with fungal changes and function, e.g. alkalinization of cytoplasm during cellular development of *C. albicans* (*50*) and during formation of macroconidia of *F. culmorum* (*51*). In addition, the pH of the endoplasmic reticulum in HeLa cells is permeable to H+ transfer to/from the cytosol (*52*). Menke *et al*. showed that *F. graminearum* cellular morphology changes *in vitro* upon induction of trichothecene biosynthesis (*53*), e.g. with development of toxisomes harboring CYPs and a CPR involved in trichothecene biosynthesis (*53*).

Except for plants only one or two CPR genes have been putatively identified in eukaryotic species. The FgCPR characterized in this paper could therefore behave as a redox partner to one of the TRI proteins of the trichothecene pathway. This could also explain its optimal kinetic activity at pH 8.5 in agreement with studies by Menke *et al*. (*53*). Alternatively, FgCPR might be a reductase partner involved in sterol or fatty acid biosynthesis in *F. graminearum*.

References

1. McMullen, M., Jones, R., and Gallenberg, D. (1997) Scab of wheat and barley: A re-emerging disease of devastating impact, *Plant Dis 81*, 1340-1348.

2. Foroud, N. A., and Eudes, F. (2009) Trichothecenes in Cereal Grains, *Int J Mol Sci 10*, 147-173.

3. Jansen, C., von Wettstein, D., Schafer, W., Kogel, K. H., Felk, A., and Maier, F. J. (2005) Infection patterns in barley and wheat spikes inoculated with wild-type and trichodiene synthase gene disrupted Fusarium graminearum, *P Natl Acad Sci USA 102*, 16892-16897.

4. Proctor, R. H., Hohn, T. M., and Mccormick, S. P. (1995) Reduced Virulence of Gibberella-Zeae Caused by Disruption of a Trichothecene Toxin Biosynthetic Gene, *Mol Plant Microbe In 8*, 593-601.

5. Kimura, M., Tokai, T., Takahashi-Ando, N., Ohsato, S., and Fujimura, M. (2007) Molecular and genetic studies of Fusarium trichothecene biosynthesis: Pathways, genes, and evolution, *Biosci Biotech Bioch 71*, 2105-2123.

6. Desjardins, A. E., Hohn, T. M., and Mccormick, S. P. (1993) Trichothecene Biosynthesis in Fusarium Species - Chemistry, Genetics, and Significance, *Microbiol Rev 57*, 595-604.

7. Desjardins, A. E., Plattner, R. D., and Vanmiddlesworth, F. (1986) Trichothecene Biosynthesis in Fusarium-Sporotrichioides - Origin of the Oxygen-Atoms of T-2 Toxin, *Appl Environ Microb 51*, 493-497.

8. McCormick, S. P., Stanley, A. M., Stover, N. A., and Alexander, N. J. (2011) Trichothecenes: From Simple to Complex Mycotoxins, *Toxins 3*, 802-814.

9. Hohn, T. M., and Vanmiddlesworth, F. (1986) Purification and Characterization of the Sesquiterpene Cyclase Trichodiene Synthetase from Fusarium-Sporotrichioides, *Arch Biochem Biophys 251*, 756-761.

10. Tokai, T., Koshino, H., Takahashi-Ando, N., Sato, M., Fujimura, M., and Kimura, M. (2007) Fusarium Tri4 encodes a key multifunctional cytochrome P450 monooxygenase for four consecutive oxygenation steps in trichothecene biosynthesis, *Biochem Bioph Res Co 353*, 412-417.

11. McCormick, S. P., Alexander, N. J., and Proctor, R. H. (2006) Fusarium Tri4 encodes a multifunctional oxygenase required for trichothecene biosynthesis, *Can J Microbiol 52*, 636-642.

12. Meek, I. B., Peplow, A. W., Ake, C., Phillips, T. D., and Beremand, M. N. (2003) Tri1 encodes the cytochrome P450 monooxygenase for C-8 hydroxylation during trichothecene biosynthesis in Fusarium sporotrichioides and resides upstream of another new Tri gene, *Appl Environ Microb 69*, 1607-1613.

13. Alexander, N. J., Hohn, T. M., and McCormick, S. P. (1998) The TRI11 gene of Fusarium sporotrichioides encodes a cytochrome P-450 monooxygenase required for C-15 hydroxylation in trichothecene biosynthesis, *Appl Environ Microb 64*, 221-225.

14. Lee, T., Han, Y. K., Kim, K. H., Yun, S. H., and Lee, Y. W. (2002) Tri13 and Tri7 determine deoxynivalenol- and nivalenol-producing chemotypes of Gibberella zeae, *Appl Environ Microb 68*, 2148-2154.

15. Liu, X., Yu, F., Schnabel, G., Wu, J. B., Wang, Z. Y., and Ma, Z. H. (2011) Paralogous cyp51 genes in Fusarium graminearum mediate differential sensitivity to sterol demethylation inhibitors, *Fungal Genet Biol 48*, 113-123.

16. Harris, L. J., Alexander, N. J., Saparno, A., Blackwell, B., McCormick, S. P., Desjardins, A. E., Robert, L. S., Tinker, N., Hattori, J., Piche, C., Schernthaner, J. P., Watson, R., and Ouellet, T. (2007) A novel gene cluster in Fusarium graminearum contains a gene that contributes to butenolide synthesis, *Fungal Genet Biol 44*, 293-306.

17. Bahadoor, A., Schneiderman, D., Gemmill, L., Bosnich, W., Blackwell, B., Melanson, J. E., McRae, G., and Harris, L. J. (2016) Hydroxylation of Longiborneol by a Clm2-Encoded CYP450 Monooxygenase to Produce Culmorin in Fusarium graminearum, *J Nat Prod 79*, 81-88.

18. Dawson, J. (1996) Cytochrome P450. Structure, mechanism, and biochemistry, 2nd edition - DeMontellano,PRO, *Science 271*, 1507-1508.

19. Hubbard, P. A., Shen, A. L., Paschke, R., Kasper, C. B., and Kim, J. J. P. (2001) NADPH-cytochrome P450 oxidoreductase - Structural basis for hydride and electron transfer, *J Biol Chem 276*, 29163-29170.

20. Lah, L., Krasevec, N., Trontelj, P., and Komel, R. (2008) High diversity and complex evolution of fungal cytochrome P450 reductase: Cytochrome P450 systems, *Fungal Genet Biol 45*, 446-458.

21. Black, S. D., French, J. S., Williams, C. H., and Coon, M. J. (1979) Role of a Hydrophobic Polypeptide in the N-Terminal Region of Nadph-Cytochrome P-450 Reductase in Complex-Formation with P-450lm, *Biochem Bioph Res Co 91*, 1528-1535.

22. Lamb, D. C., Warrilow, A. G. S., Venkateswarlu, K., Kelly, D. E., and Kelly, S. L. (2001) Activities and kinetic mechanisms of native and soluble NADPH-cytochrome P450 reductase, *Biochem Bioph Res Co 286*, 48-54.

23. Makovec, T., and Breskvar, K. (2002) Catalytic and immunochemical properties of NADPH-cytochrome P450 reductase from fungus Rhizopus nigricans, *J Steroid Biochem 82*, 89-96.

24. Venkateswarlu, K., Lamb, D. C., Kelly, D. E., Manning, N. J., and Kelly, S. L. (1998) The N-terminal membrane domain of yeast NADPH-cytochrome P450 (CYP) oxidoreductase is not required for catalytic activity in sterol biosynthesis or in reconstitution of CYP activity, *J Biol Chem 273*, 4492-4496.

25. Shen, A. L., Porter, T. D., Wilson, T. E., and Kasper, C. B. (1989) Structural-Analysis of the Fmn Binding Domain of Nadph-Cytochrome P-450 Oxidoreductase by Site-Directed Mutagenesis, *J Biol Chem 264*, 7584-7589.

26. Guengerich, F. P., Martin, M. V., Sohl, C. D., and Cheng, Q. (2009) Measurement of cytochrome P450 and NADPH-cytochrome P450 reductase, *Nat Protoc 4*, 1245-1251.

27. Yim, S. K., Yun, S. J., and Yun, C. H. (2004) A continuous spectrophotometric assay for NADPH-cytochrome P450 reductase activity using 1,1-dipheny1-2-picrylhydrazyl, *J Biochem Mol Biol 37*, 629-633.

28. Yim, S. K., Yun, C. H., Ahn, T., Jung, H. C., and Pan, J. G. (2005) A continuous spectrophotometric assay for NADPH-cytochrome P450 reductase activity using 3-(4,5-dimethylthiazol-2-yl)-2,5-diphenyltetrazolium bromide, *J Biochem Mol Biol 38*, 366-369.

29. Fuji, T., and Takaya, N. (2008) Denitrification by the fungus Fusarium oxysporum involves NADH-Nitrate reductase, *Biosci Biotech Bioch 72*, 412-420.

30. Kurakov, A. V., Nosikov, A. N., Skrynnikova, E. V., and L'vov, N. P. (2000) Nitrate reductase and nitrous oxide production by Fusarium oxysporum 11dn1 under aerobic and anaerobic conditions, *Curr Microbiol 41*, 114-119.

31. Morozkina, E. V., Kurakov, A. V., Nosikov, A. N., Sapova, E. V., and L'vov, N. P. (2005) Properties of Nitrate Reductase from Fusarium oxysporum 11dn1 Fungi Grown under Anaerobic Conditions, *Appl Biochem Micro+ 41*, 254-258.

32. Cuomo, C. A., Gueldener, U., Xu, J. R., Trail, F., Turgeon, B. G., Di Pietro, A., Walton, J. D., Ma, L. J., Baker, S. E., Rep, M., Adam, G., Antoniw, J., Baldwin, T., Calvo, S., Chang, Y. L., DeCaprio, D., Gale, L. R., Gnerre, S., Goswami, R. S., Hammond-Kosack, K., Harris, L. J., Hilburn, K., Kennell, J. C., Kroken, S., Magnuson, J. K., Mannhaupt, G., Mauceli, E., Mewes, H. W., Mitterbauer, R., Muehlbauer, G., Munsterkotter, M., Nelson, D., O'Donnell, K., Ouellet, T., Qi, W. H., Quesneville, H., Roncero, M. I. G., Seong, K. Y., Tetko, I. V., Urban, M., Waalwijk, C., Ward, T. J., Yao, J. Q., Birren, B. W., and Kistler, H. C. (2007) The Fusarium graminearum genome reveals a link between localized polymorphism and pathogen specialization, *Science 317*, 1400-1402.

33. Chen, G. J., Qiu, N., Karrer, C., Caspers, P., and Page, M. G. P. (2000) Restriction site-free insertion of PCR products directionally into vectors, *Biotechniques 28*, 498-+.

34. van den Ent, F., and Lowe, J. (2006) RF cloning: A restriction-free method for inserting target genes into plasmids, *J Biochem Bioph Meth 67*, 67-74.

35. Lamb, D. C., Kim, Y. C., Yermalitskaya, L. V., Yermalitsky, V. N., Lepesheva, G. I., Kelly, S. L., Waterman, M. R., and Podust, L. M. (2006) A second FMN binding site in yeast NADPH-cytochrome p450 reductase suggests a mechanism of electron transfer by diflavin reductases, *Structure 14*, 51-61.

36. Oprian, D. D., and Coon, M. J. (1982) Oxidation-Reduction States of Fmn and Fad in Nadph-Cytochrome-P-450 Reductase during Reduction by Nadph, *J Biol Chem 257*, 8935-8944.

37. Dionisio, G., Madsen, C. K., Holm, P. B., Welinder, K. G., Jorgensen, M., Stoger, E., Arcalis, E., and Brinch-Pedersen, H. (2011) Cloning and Characterization of Purple Acid Phosphatase Phytases from Wheat, Barley, Maize, and Rice, *Plant Physiol 156*, 1087-1100.

38. Bridges, A., Gruenke, L., Chang, Y. T., Vakser, I. A., Loew, G., and Waskell, L. (1998) Identification of the binding site on cytochrome P450 2B4 for cytochrome b(5) and cytochrome P450 reductase, *J Biol Chem 273*, 17036-17049.

39. Shen, S. J., and Strobel, H. W. (1993) Role of Lysine and Arginine Residues of Cytochrome-P450 in the Interaction between Cytochrome-P4502b1 and Nadph-Cytochrome P450 Reductase, *Arch Biochem Biophys 304*, 257-265.

40. Shen, S. J., and Strobel, H. W. (1994) Probing the Putative Cytochrome-P450- and Cytochrome-C Binding Sites on Nadph-Cytochrome-P450 Reductase by Antipeptide Antibodies, *Biochemistry-Us 33*, 8807-8812.

41. Shimizu, T., Tateishi, T., Hatano, M., and Fujiikuriyama, Y. (1991) Probing the Role of Lysines and Arginines in the Catalytic Function of Cytochrome-P450d by Site-Directed Mutagenesis - Interaction with Nadph-Cytochrome-P450 Reductase, *J Biol Chem 266*, 3372-3375.

42. Lee, G. Y., Kim, H. M., Ma, S. H., Park, S. H., Joung, Y. H., and Yun, C. H. (2014) Heterologous expression and functional characterization of the NADPH-cytochrome P450 reductase from Capsicum annuum, *Plant Physiol Bioch 82*, 116-122.

43. Sen, A., and Arinc, E. (1998) Purification and Characterization of Cytochrome P450 Reductase from Liver Microsomes of Feral Leaping Mullet (Liza saliens), *J Biochem Mol Toxic 12*, 103-113.

44. Wang, M., Roberts, D. L., Paschke, R., Shea, T. M., Masters, B. S. S., and Kim, J. J. P. (1997) Three-dimensional structure of NADPH-cytochrome P450 reductase: Prototype for FMN- and FAD-containing enzymes, *P Natl Acad Sci USA 94*, 8411-8416.

45. Freydank, A. C., Brandt, W., and Drager, B. (2008) Protein structure modeling indicates hexahistidine-tag interference with enzyme activity, *Proteins 72*, 173-183.

46. Park, S. H., Kang, J. Y., Kim, D. H., Ahn, T., and Yun, C. H. (2012) The Flavin-Containing Reductase Domain of Cytochrome P450 BM3 Acts as a Surrogate for Mammalian NADPH-P450 Reductase, *Biomol Ther 20*, 562-568.

47. Gasteiger, E., Gattiker, A., Hoogland, C., Ivanyi, I., Appel, R. D., and Bairoch, A. (2003) ExPASy: the proteomics server for in-depth protein knowledge and analysis, *Nucleic Acids Res 31*, 3784-3788.

48. Makovec, T., and Breskvar, K. (1998) Purification and characterization of NADPH-cytochrome P450 reductase from filamentous fungus Rhizopus nigricans, *Arch Biochem Biophys 357*, 310-316.

49. Menting, J. G. T., Cornish, E., and Scopes, R. K. (1994) Purification and Partial Characterization of Nadph-Cytochrome-C Reductase from Petunia-Hybrida Flowers, *Plant Physiol 106*, 643-650.

50. Stewart, E., Gow, N. A. R., and Bowen, D. V. (1988) Cytoplasmic Alkalinization during Germ Tube Formation in Candida-Albicans, *J Gen Microbiol 134*, 1079-1087.

51. Chitarra, G. S., Breeuwer, P., Rombouts, F. M., Abee, T., and Dijksterhuis, J. (2005) Differentiation inside multicelled macroconidia of Fusarium culmorum during early germination, *Fungal Genet Biol 42*, 694-703.

52. Kim, J. H., Johannes, L., Goud, B., Antony, C., Lingwood, C. A., Daneman, R., and Grinstein, S. (1998) Noninvasive measurement of the pH of the endoplasmic reticulum at rest and during calcium release, *P Natl Acad Sci USA 95*, 2997-3002.

53. Menke, J., Weber, J., Broz, K., and Kistler, H. C. (2013) Cellular Development Associated with Induced Mycotoxin Synthesis in the Filamentous Fungus Fusarium graminearum, *Plos One 8*.

Figures

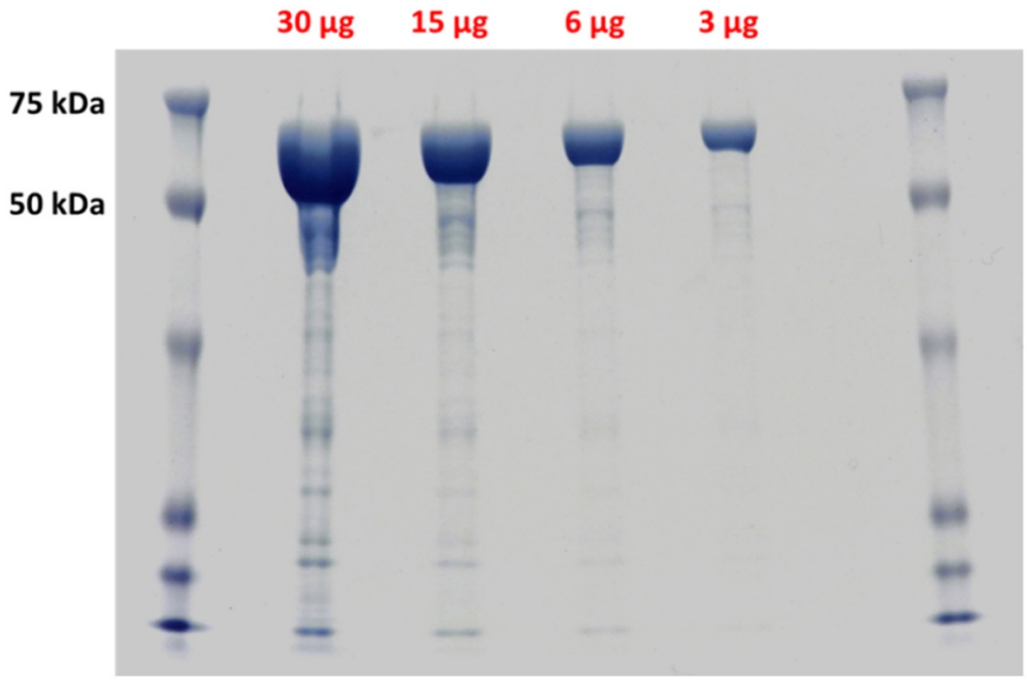


Figure 1. Analysis of FgCPR by SDS-PAGE and Coomassie R-250 staining of different amounts of protein loaded (as indicated in red). Reference bands are given for protein markers of 50 and 75 kDa from.

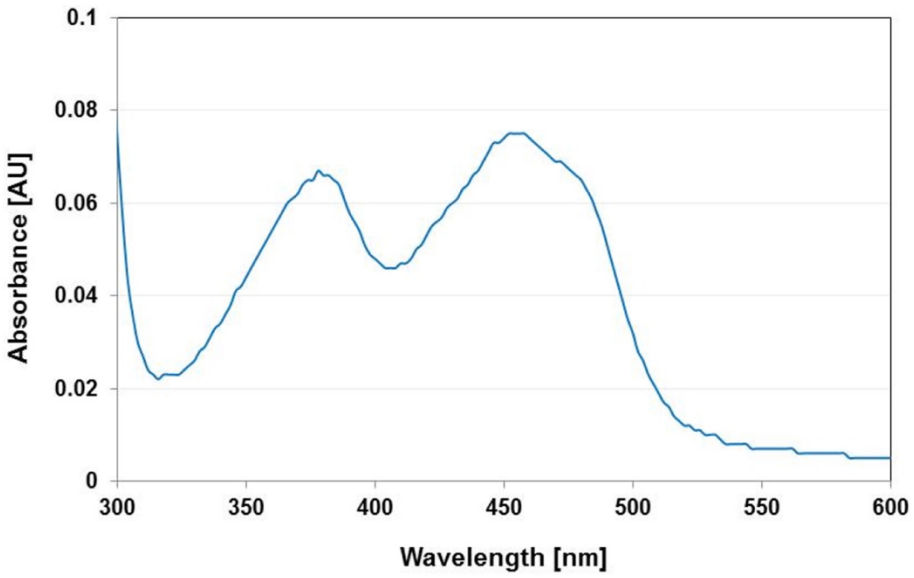


Figure 2. UV/vis spectrum of 8.2 μM FgCPR in 150 mM phosphate buffer (pH 7.5) confirming the presence of fully oxidized flavins from the peaks in the region 300–500 nm.

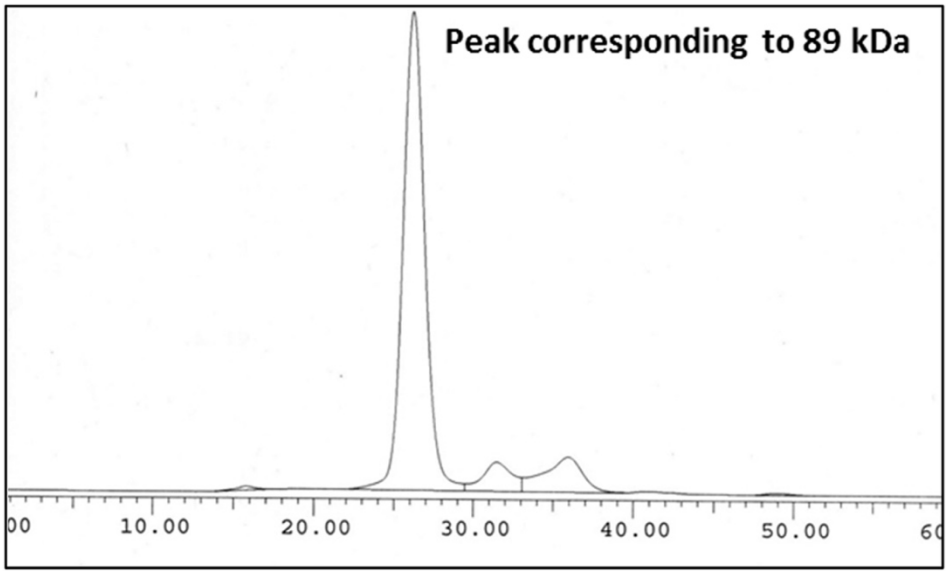


Figure 3. Size exclusion chromatography of FgCPR obtained under conditions described in the experimental section. The apparent molecular weight (compared to protein markers) was 89 kDa suggesting the protein to be monomeric.

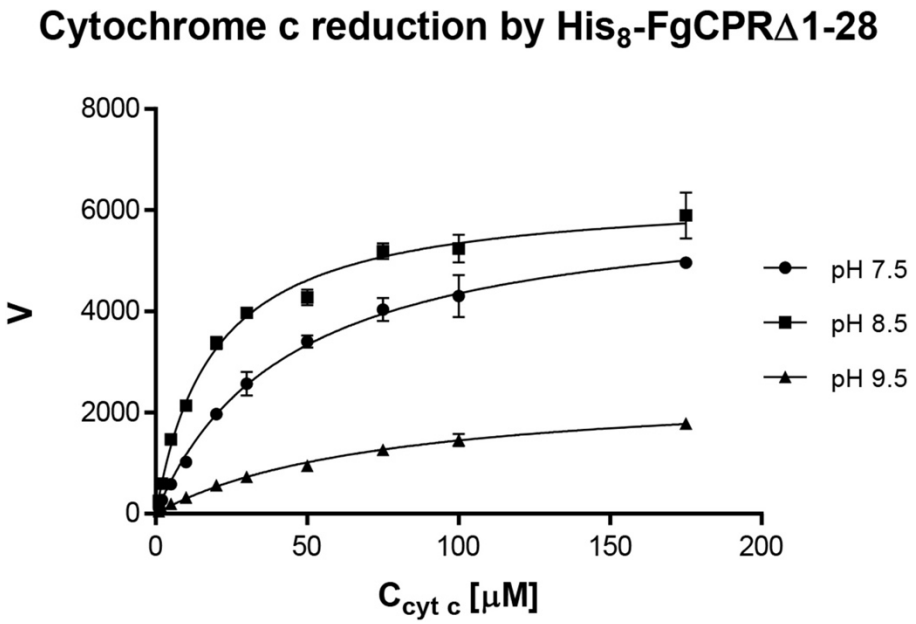


Figure 4. Kinetic activity of FgCPR with cytochrome c as substrate at pH 7.5 (circles), 8.5 (squares) and 9.5 (triangles). V is given in nmoles/min/mg. Error bars are standard errors.

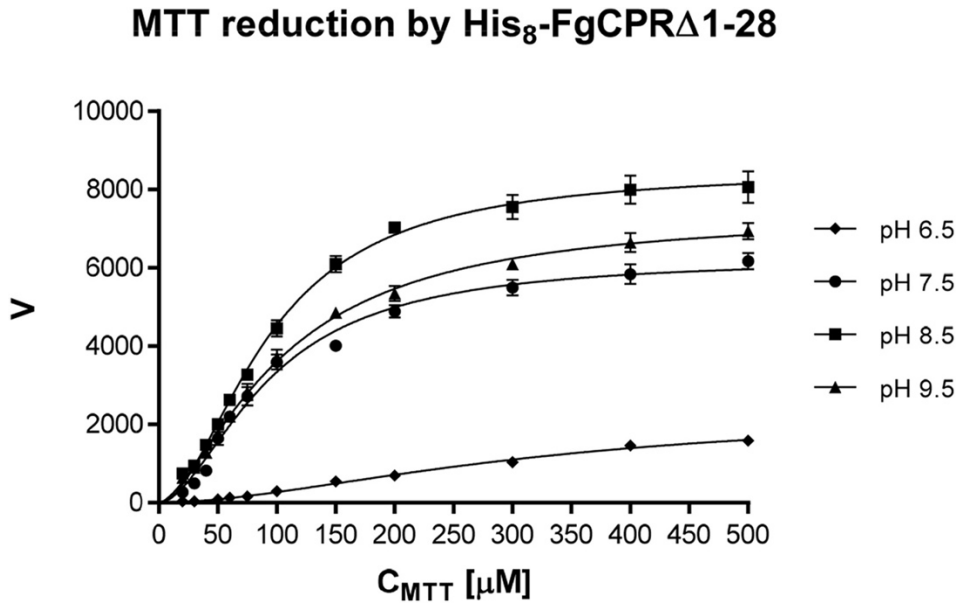


Figure 5. Kinetic activity of FgCPR with MTT as substrate at pH 6.5 (diamonds), pH 7.5 (circles), 8.5 (squares) and 9.5 (triangles). V is given in nmoles/min/mg. Error bars are standard errors.

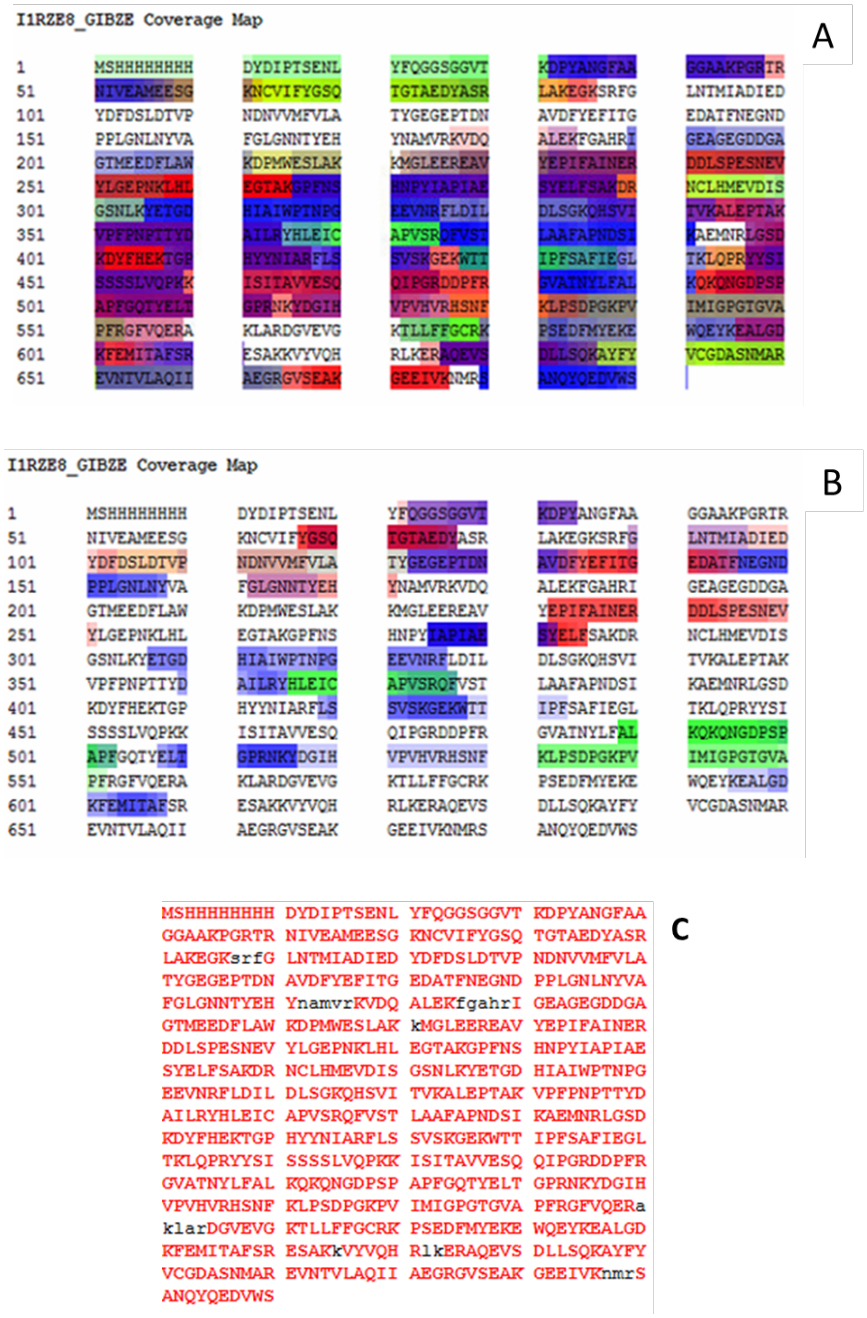


Figure S1. Trypsin (A) and chymotrypsin (B) peptide mapping of FgCPR for primary sequence verification. Combined trypsin and chymotrypsin coverage account up to 96.4 % of the total sequence (C). Peptide mapping (A and B panel) has been done using PLGS 2.5.3 (Waters, Milford, USA) and MSe data. Color code of the peptide map corresponds to the ion spectral counting abundance. The darker is the color, the more abundant was found the peptide. The overall coverage map has been done by using the software Protein Coverage Summarizer (https://omics.pnl.gov/software/protein-coverage-summarizer)

Table S1. Primers for restriction site-free cloning and sequencing of the FgCPRΔ1-28 gene fragment. For primers for merging of G-blocks, red segments are annealing to G-blocks and black segments are flanks for cloning into the pTEV4 target vector

G-block	**Sequence**	
N-t part (forward primer)	5’GGTGTTACCAAGGACCCTTATGCCAATGGCTTCGCTGCCGGCGGCGCTGCCAAGCCT GGTCGCACGAGAAACATCGTCGAGGCCATGGAGGAATCTGGCAAGAACTGTGTCATCT TCTATGGTTCTCAGACTGGTACCGCTGAAGATTATGCTTCCCGTCTCGCCAAGGAGGGT AAGAGTCGATTCGGACTCAACACCATGATCGCCGATATTGAAGACTACGATTTCGATAG CCTGGATACTGTCCCCAACGACAACGTTGTCATGTTTGTCCTCGCAACTTACGGTGAAG GCGAGCCTACCGACAACGCAGTCGATTTCTACGAGTTCATCACTGGTGAGGATGCCAC CTTCAACGAGGGCAATGACCCCCCTCTGGGCAACCTCAACTATGTCGCTTTCGGTCTCG GAAACAACACCTACGAACACTACAACGCCATGGTCCGTAAAGTTGACCAGGCTCTTGA GAAATTTGGTGCTCACCGCATCGGCGAAGCTGGTGAAGGTGACGATGGCGCTGGTAC CATGGAAGAGGACTTCTTGGCCTGGAAGGATCCCATGTGGGAGTCCCTTGCTAAGAAA ATGGGACTGGAAGAGCGCGAAGCCGTTTACGAGCCTATCTTTGCCATTAACGAGCGTG ACGACCTGAGCCCCGAATCCAACGAAGTCTATCTCGGTGAACCTAACAAGCTGCATCTC GAAGGCACTGCCAAGGGTCCCTTTAACTCTCACAACCCCTACATTGCCCCTATTGCCGA GTCATACGAGCTCTTCTCTGCCAAGGACAGGAACTGCCTTCACATGGAAGTCGACATCA GCGGTTCTAACCTCAAATACGAAACTGGTGATCACATTGCTATCTGGCCCACAAATCCT GGTGAGGAGGTTAACCGTTTCTTGGATATTCTCGATCTCTCCGGCAAGCAGCACAGCGT CATCACCGTTAAGGCTC-3’	
C-t part (reverse primer)	5’TTATGACCAAACATCCTCCTGATATTGATTTGCCGATCTCATGTTCTTCACAATCTCCT CACCCTTAGCCTCAGACACTCCACGTCCTTCGGCAATGATTTGCGCCAAAACAGTGTTG ACCTCGCGAGCCATGTTTGAAGCATCACCACAAACATAGAAGTAGGCCTTCTGGGAGA GCAGGTCGCTGACCTCTTGTGCTCTCTCCTTAAGTCGGTGCTGAACATAAACCTTCTTGG CGCTTTCTCGAGAAAAGGCTGTGATCATCTCGAACTTGTCGCCGAGAGCCTCCTTGTAT TCCTGCCATTCCTTCTCGTACATGAAATCTTCACTGGGCTTTCGGCAACCGAAGAACAAT AGAGTCTTTCCGACCTCAACACCATCGCGGGCCAACTTGGCACGTTCCTGGACGAAACC ACGGAAAGGAGCGACACCAGTACCAGGGCCAATCATGATGACAGGCTTTCCAGGGTC CGAAGGTAGCTTGAAGTTGGAGTGACGAACATGGACGGGGACATGGATACCATCGTA CTTGTTGCGGGGACCTGTAAGCTCGTATGTCTGACCAAAAGGTGCTGGGCTGGGGTCG CCGTTCTGCTTCTGCTTCAGGGCAAAGAGATAGTTGGTAGCGACGCCGCGGAAAGGGT CATCTCGGCCAGGAATTTGCTGGGACTCGACGACGGCGGTGATGGAAATCTTCTTAGG CTGAACCAGAGAAGAGGAGGAGATGGAGTAATAGCGGGGCTGGAGCTTGGTGAGAC CCTCGATGAAGGCGGAGAAGGGAATCGTGGTCCACTTTTCACCCTTGCTAACACTGGA GAGGAAACGAGCAATATTGTAGTAGTGAGGGCCGGTCTTCTCGTGGAAGTAGTCCTTA TCGCTACCGAGACGGTTCATCTCAGCCTTGATAGAGTCGTTGGGAGCGAAAGCAGCGA GAGTAGAAACGAATTGACGGGAAACAGGAGCGCAGATCTCGAGATGGTATCGCAGAA TGGCATCGTAGGTCGTAGGGTTAGGGAAGGGAACCTTGGCGGTAGGCTCAAGAGCCT TAACGGTGATGACGCTG-3’		
Primer	**PCR procedure**	**Sequence**
FgCPR Ext_F		5’CCTGTATTTTCAAGGTGGCTCTGGTGGTGTTACCAAGGA CCCTTATGCCAATG -3’
FgCPR Ext_R		5’AGTTATTGCTCAGCGCCCCATTTATGACCAAACATCCTCC TGATATTGATTTGC-3’