

Employment of Catalytic Analysis and Computational Modeling in the Development of a Novel Competitive Inhibitor of Cdc14 Phosphatases in C. zeina

Charlene Miciano, Hieu D. Nguyen, Rohan Sanda

From the Summer Science Program in Biochemistry, University of California San Diego, 9500 Gilman Drive, La Jolla, CA, 92093

Keywords: *C. zeina*, Cdc14, phosphatases, *in silico* inhibitor design, pNPP

Running Title: *Developing a Novel Inhibitor of Cdc14s in C. zeina*

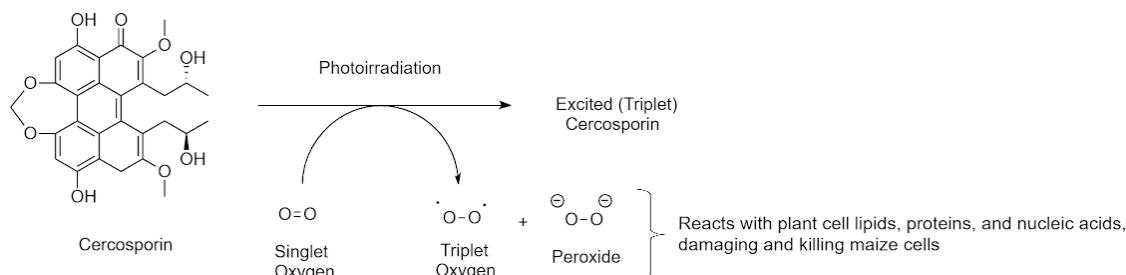
ABSTRACT

In the developing world, maize is an important staple crop that has the potential to not only mitigate food insecurity but also to decrease poverty levels in Sub-Saharan African countries. At least 900 million people who earn less than 2 USD per day rely on maize as a staple crop. With population levels increasing rapidly in recent years, the world has been faced with the epidemic of undernutrition. One specific contributor to this issue is the abundance of pathogenic fungi. In this paper, we analyze the enzymatic properties of a cell division cycle (Cdc) protein tyrosine phosphatase (PTP) found in the fungal pathogen *Cercospora zeina*, a propagator of a devastating maize disease known as Grey Leaf Spot (GLS) Disease. We empirically show that CzCdc14 possesses a low binding affinity for the substrate *para*-Nitrophenyl phosphate (pNPP). Multiple substrates and inhibitors were screened to identify potential inhibitor targets. IC₅₀ values and reversibility mechanisms for CzCdc14 with pNPP were examined for optimal inhibitors as well. *In silico* comparative modeling was employed to develop a homology model of CzCdc14. It was shown that CzCdc14 possesses a highly conserved active site domain to Cdc14s from five diverse species, prefers binding to a large hydrophobic-aromatic moiety, and contains the cystine, arginine, and aspartic acid residues that are critical in the development of ligand/receptor interactions. We performed docking and optimization of a novel competitive inhibitor that could form the foundation for future fungicide development and enhancement of global food security.

INTRODUCTION

Grey leaf spot (GLS) is a widespread crop-yield limiting fungal disease that primarily affects maize and corn plants in Southern Africa and the Americas. GLS is caused by the fungus *Cercospora zeina* and is characterized by the yellow-tinged lesions it forms along the leaves of host plants. Typically these lesions can be approximately 1.5 to 2 inches in length, however, as GLS reaches its late phase, lesions may merge and further destroy plant cells. As a result of the irreversible damage to maize plants, GLS is the most significant disease that decreases maize yield decreasing disease in the world [1]. Its large-scale biomass destruction has spurred work to determine the causal factors of GLS.

Recent investigations have revealed two closely related species of the *Cercospora* fungus genus to be responsible for the dissemination of GLS: *C. zeina* and *C. zeae maydis*. *C. zeina* has been found to be primarily responsible for GLS in Southern Africa and *C. zeae maydis* has been found to be more prevalent in Northern and Southern America [2]. While the neurotrophic effects of both fungi are similar, *C. zeae maydis* has been found to produce more cercosporin (the active toxin in *Cercospora* which can be photo-irradiated to produce reactive oxygen radicals that can damage cells, see Scheme 1) and propagate faster [3]. Currently, the high genetic diversity of both fungi has made fungicide development difficult [4]. In addition, *C. zeina* affects primarily the agrarian economies of developing nations in Southern Africa where maize production constitutes an important food source and export [5]. Therefore, research into the development of targeted inhibitors for *Cercospora* is necessary to safeguard the world's corn supply and provide regional stability.



Scheme 1

In this paper, we analyze the enzymatic properties of the Cdc14 protein phosphatase from *C. zeina* to explore possibilities for targeted inhibitor development. Cdc14 protein tyrosine phosphatases are a group of enzymes that dephosphorylate phosphoserine, phosphotyrosine, and phosphothreonine from target proteins, thus regulating critical signal transduction pathways. Current study of the Cdc14 in *Saccharomyces cerevisiae* has revealed that substrate binding affinity is influenced by the minimum motif of pSer-Pro-x-Lys on the substrate [6]. While the specific function of the target Cdc14 protein phosphatase in *C. zeina* is unknown, previous studies have found Cdc14s to be important in the dephosphorylation of cyclin-dependent kinases (CDK) in the human mitotic cell cycle. As a result of this role in cell proliferation, Cdc14s have been the subject of modern oncogenesis research [7]. It is hypothesized that Cdc14s in *C. zeina* may have a similar role in mitosis and could be a viable target for fungicide development.

Using an obtained protein nucleotide sequence that codes for a Cdc14 in *C. zeina* (CzCdc14), protein was amplified and purified. The kinetics and thermodynamics data of the protein with different substrates and inhibitors were then studied using spectrophotometry. These findings were used in the development of an *in silico* inhibitor of CzCdc14. This paper will describe the protein's enzymatic properties, along with a potential pesticide that could be utilized to combat Grey Leaf Spot disease in Southern Africa.

METHODS

Interspecies protein Comparisons

The Basic Local Alignment Search Tool (BLAST) was used to identify similar Cdc14 DNA sequences to the obtained nucleotide sequence of CzCdc14 (see Supplementary Figure 1A). *Homo Sapiens*, *Drosophila melanogaster*, *Caenorhabditis elegans*, *Saccharomyces cerevisiae*, and *Schizosaccharomyces pombe*, and *C. zeina* were the analyzed species. Result FASTA sequences were then

Developing a Novel Inhibitor of Cdc14s in *C. zeina*

inputted into the Clustal Omega program from EMBL-EBI. A sequence alignment was performed and conserved domains were recorded among species.

Molecular Modeling of CzCdc14

To gain a preliminary understanding of the structure of CzCdc14, comparative modeling algorithms were applied to its obtained amino acid sequence. We matched the query to the Cdc14 5XW5 through the Protein Data Bank (PDB). The known crystal structure of 5XW5 was used as the template in the comparative modeling software available on the Molecular Operating Environment (MOE) program. Once a target conformation of the query was developed, similarity and identity scores were calculated to determine the accuracy of the model. Root-mean-square deviation (RMSD) was also calculated to determine both the relative energetic stability and similarity of the target compared to the template protein. A Ramachandran plot detailing the unfavorable ϕ/ψ (phi/psi) angles was created and unfavorable bond lengths, dihedral bond angles, and rotamer energies were also computed. To determine the active site on the target structure, we aligned and superposed our target protein to the binding pocket of Cdc14 1OHE containing a bound phosphopeptide ligand in the active site. Using the MOE software, the ligand and active site pocket could be selected from the template and then identified on the target template. The amino acids that composed the active site of the template were then selected and highlighted on the target, demonstrating the active site within CzCdc14.

Transformation of the Protein

An obtained sequence of the CzCdc14 DNA was inserted into a 6x His-Cdc14 pET15b plasmid between the NdeI and BamHI sites on the LCK gene. The presence of the 6x His-tag at the N-terminus of the translated protein would allow for the purification of our protein via Ni-NTA column chromatography in later steps.

To transform *Escherichia coli* (*E.coli*) BL21 A1 cells with pET15b plasmid, we used a heat shock method. 1 μ L of plasmid was mixed with 100 μ L of competent bacteria and then placed on ice for 30 minutes. The plasmid and bacteria were then placed in a 42°C water bath for one minute and then transferred to ice for two minutes. 450 μ L of Luria Broth (LB) media was added to the cell mixture. We then incubated the plasmid and bacteria at 37°C for 30 minutes. A negative control consisting of 101 μ L of competent bacteria was also tested. Transformed bacteria were selected by Ampicillin-containing LB agar plates and amplified for approximately 15 hours (overnight).

A single isolated colony was transferred to 5 mL of the transformed bacteria to 1 L of 2xYT media (NaCl 0.5 % m/m, tryptone 1.56 % m/m, and yeast extract 1 % m/m) and added 1 mL of ampicillin. 3 mL of solution was then removed for reading on a spectrophotometer at a wavelength of 600 nm. The remaining solution was then incubated at 37°C and shaken at 180 revolutions per minute (rpm) for two hours. Multiple 3 mL samples of media were removed over a period of three hours and absorbances were measured using the same conditions. Once an OD₆₀₀ of 0.7 was obtained, a 100 μ L sample was removed from the sample and stored in SDS Loading Buffer. 10 g of L-Arabinose (20 % m/v) was then added to the 1 L flask to induce the sample. The sample was left to incubate overnight at 95° C. Next, the final OD₆₀₀ value was recorded from a 3 mL sample and an additional 100 μ L aliquot sample was removed and stored in SDS Running Buffer (Tris Base 0.025 M, glycine 0.19 M, sodium dodecyl sulfate (SDS) 3.46 M). The sample was then centrifuged at 5000 rpm for 30 minutes and the pellet removed.

Purification of Protein

Purification of the protein began by adding 30 mL of lysis buffer solution (25 mM HEPES pH 7.5, 500 mM NaCl, 0.1% (v/v) Triton X-100, 10 mM imidazole, 10% (v/v) glycerol). We then added a half of a Protease Inhibitor Tablet, 50 mg of Lysozyme, and 8% (v/v) of Universal Nuclease to the thawed cell pellet. A crushed cell pellet was added to the lysis buffer solution and was left to incubate on ice for an additional 30 minutes.

Next, 0.15 mL of phenylmethylsulfonyl fluoride (PMSF) Protease Inhibitor was added and the mixture was transferred to a 40 mL conical centrifuge tube. After centrifugation for 200,000 g for 45 minutes at 4°C, 24 mL of the supernatant was removed and an additional 11 mL of lysis buffer was added. The mixture was filtered using a 30 mL syringe and a 0.45 mm syringe filter and placed on ice. 75 μ L of the filtered solution was removed and stored for a future SDS-PAGE (sodium dodecyl sulfate –polyacrylamide gel electrophoresis). Another 25 μ L of the filtered solution was reserved for a future Bradford Assay to determine the concentration of the protein.

The CzCdc14 enzyme was isolated using the affinity chromatography technique. The Histidine tag located on CzCdc14 allowed for the binding to the Ni^{2+} – chelate column within the ÄKTA High-Performance Liquid Chromatography (HPLC) machine. Imidazole, contained in solutions of Nickel Buffer A (25 mM HEPES pH 7.5, 500 mM NaCl, 10% (v/v) glycerol) and Nickel Buffer B (25 mM HEPES pH 7.5, 500 mM NaCl, 250 mM imidazole, 10% (v/v) glycerol) was used to elute the protein from the column to be collected. Another 75 μ L and 25 μ L of the flow through was reserved for the SDS-PAGE and Bradford Assay, respectively. The protein solution was then dialyzed overnight.

Protein Concentration Determination

The concentration of the purified protein was then measured using a Nanodrop instrument and confirmed using a Bradford Assay. Protein dilutions of 10x and 25x were used for the Nanodrop reading and a dilution of 25x was used for the Bradford Assay at a wavelength of 595 nm. The nanodrop readings were conducted at a wavelength of 280 nm. The measured absorbances were converted to concentrations using the extinction coefficient using the ProtParam software available on the Expert Protein Analysis System (ExPASy): $69,790 \text{ M}^{-1} \text{ cm}^{-1}$. After correcting for the dilution factor the two concentrations for the Nanodrop readings were averaged to find a final concentration. The Bradford assay absorbance was then found and analyzed to ensure that it fell within the linear range of the Bradford standard curve (see Supplemental Figure 2A). The concentration that was found, after converting to molarity using the molecular weight calculated from ExPASy (58450.17 g/mol).

We performed SDS-PAGE using 180 V for 60 minutes and a running gel consisting of 30% solution of 29:1 Acrylamide/Bis-acrylamide. Coomassie Blue was used to stain the gel prior to being read using the Bio-Rad Gel Doc EZ imaging system.

Phosphatase Kinetics Assay

To determine the linear range and calculate the specific activity of the CzCdc14, a kinetic assay was performed using varying concentrations of CzCdc14 and 440 mM of pNPP. The concentrations that were tested were 10 μ M, 3 μ M, 1 μ M, 0.3 μ M, and 0.1 μ M. Assays were performed in 80% Enzyme Reaction Buffer (25 mM HEPES ((4-(2hydroxyethyl)-1-piperazineethanesulfonic acid) pH 7.5, 1 mM DTT (Dithiothreitol), 1 mM EDTA (Ethylenediamine tetracarboxylic acid disodium salt), 150 mM NaCl). The reaction was observed for 35 minutes and observed for the linear activity of the enzyme.

To confirm that the enzymatic activity of CzCdc14 was consistent with previously studied Cdc14 enzymes and to determine its behavior at steady state conditions, reactions were performed using the

artificial substrate, *para*-Nitrophenylphosphate (pNPP) disodium salt, and sodium orthovanadate (Na_3VO_4), a known protein tyrosine inhibitor. 0.3 μM of CzCdc14, 400 mM of pNPP, and 50 mM of sodium orthovanadate was used and the amount of dephosphorylated pNPP was calculated using a *p*-nitrophenol standard curve (see Supplemental Figure 3). The assays were performed in 79% (v/v) Enzyme Reaction Buffer and were read using a spectrophotometer at a wavelength of 405 nm. The reaction was quenched using 5 N NaOH.

Next, the steady-state Michaelis-Menten conditions were determined using pNPP assays read at the wavelength used previously. Previous literature was analyzed to find the K_m of Cdc14 in *Saccharomyces cerevisiae*. The assays were performed in 80% (v/v) Enzyme Reaction Buffer and 10% (v/v) CzCdc14 at 0.3 μM . The amount of pNPP that was used was 10% (v/v) and ranged from concentrations of 80 mM, 60 mM, 40 mM, 20 mM, 10 mM, 4 mM, and 1 mM in solution. A set of blank assays were also performed in 90% (v/v) Enzyme Reaction Buffer and 10% (v/v) pNPP. Initial velocity was calculated using the PNP standard curve (see Supplementary Figure 3). Non-linear Michaelis-Menten curve fitting was then applied to fit the data points. This was done using Equation 1 and Equation 2 (see below):

$$\text{Equation 1: Michaelis Menten Equation } V_0 = \frac{V_{\max}[S]}{K_m + [S]}$$

$$\text{Equation 2: } K_m = \frac{V_{\max}}{2}$$

V_0 = initial velocity (mM/s); [S] = substrate concentration (mM)

Due to the high activity of the CzCdc14 and laboratory material limitations, however, concentrations of substrate above 80 mM could not be achieved. Curve fitting was used to calculate the correct V_{\max} and K_m .

Substrate Assay

Using a Malachite Green standard curve (see Supplemental Figure 4), an enzyme concentration of 0.1 μM was determined to be sufficient for proceeding experiments due to hyperactivity observed at higher concentrations.

To determine the substrate selectivity of CzCdc14, assays were conducted at a wavelength of 405 nm using 24 phosphopeptides. The assays were performed in 80% Enzyme Reaction Buffer, 10% Enzyme at 0.1 μM , and 10% substrate at 0.1 mM in assay. Blanks were also run using 90% Enzyme Reaction Buffer and 10% substrate at 0.1 mM in assay. Each reaction was quenched using 200% (v/v) Malachite Green after 10 minutes reaction time. Color was allowed to develop for 20 minutes. Absorbances were measured using a spectrophotometer at a wavelength of 640 nm.

Initial rate parameters, specifically the slope of the curve, were measured and represented the specificity constant k_{cat}/K_m . This could be done due to the Michaelis-Menten assumption that substrate concentration was significantly less than K_m , allowing for $V_0 = V_{\max}/K_m$. A modified Michaelis-Menten equation was used to characterize the Michaelis-Menten graph: $V_0 = V_{\max}/K_m$.

Inhibitor Assays In silico

To develop an optimized *in silico* inhibitor, we began by docking, minimizing, and protonating the 1OHE ligand with the homology model of CzCdc14 created in previous steps. Docking occurred in using the Rigid Receptor function along with iteration settings of 30 poses with 5 solutions. The intermolecular interactions between the ligand and active site were then examined and recorded. These

interactions were then compared to the interactions that occurred between 1OHE and the phosphoserine ligand. We then uploaded the SDF file which contained structures of all 11 inhibitors screened in previous experiments. Docking was performed between the ligands and homology model using the Rigid Receptor model with 10 poses and 3 solutions per molecule. The S-values (docking scores) of all results were compared. Ligand interactions were compared to the ligand interactions between the phosphopeptide ligand and the CzCdc14 homology model. The top three models would exhibit highly similar interactions with the phosphopeptide and homology model as well as have a strong docking score. These top three inhibitors would then be docked again with the Pharmacophore algorithm. Initial parameters consisted of the coordinating intermolecular interactions between two oxygens of the phosphate group on the ligand and Arg349 on the protein. 30 poses and three results were run for each inhibitor. The interactions and S-scores of the resulting solutions were compared and the final inhibitor conformation chosen.

Optimization then began on the MOE software. The following criteria were followed throughout the optimization process: (1) the molecular weight would be less than 750 g/mol (2) the model would not be considered toxic to humans (3) the molecule would follow all of Lipinski's Rules of 5 (with the exception of the <500 g/mol rule) (4) no more than 6 changes could be made to the inhibitor molecule (5) the molecule had to remain energetically stable. The viability of each inhibitor was evaluated on the basis of its affinity score (kcal/mol) and efficiency. Once an optimized ligand was created, its interactions and properties were recorded.

Inhibitor Assays In vitro: Inhibitor Screenings

To determine the best inhibitors to use for proceeding experiments, 11 inhibitors were screened. Each assay was performed in 75% Enzyme Reaction Buffer, 10% 1 μ M enzyme, 10% 60 mM pNPP, and 5% of each inhibitor. A set of blanks was also run using 85% Enzyme Reaction Buffer, 10% 60 mM pNPP, and 5 μ L of inhibitor. All inhibitors, with the exception of H7 and I9, were tested at 0.1 mM in solution. Inhibitors H7 and I9 were each tested at 1 mM in solution. A spectrophotometer was used to read the assays at a wavelength of 405 nm after being quenched with 5 N NaOH.

Inhibitor Assays In vitro: Calculating IC_{50} values

The IC_{50} values were calculated experimentally using the two best (highest relative inhibition) inhibitors from the inhibitor assay. To determine the percent inhibition of the two inhibitors (G6 and G7), another set of positive and negative controls were performed to monitor the dynamic range. For each inhibitor, ten different inhibitor concentrations were used in each assay. Each assay was performed in 75% Enzyme Reaction Buffer, 10% 1 μ M Enzyme, 5% Inhibitor, and 10% 60 mM pNPP. A set of blanks were also read that contained 85% Enzyme Reaction Buffer, 5% of Inhibitor, and 10% 60 mM pNPP. Each of the reactions were allowed to react for 10 minutes and were then quenched with 5 N NaOH. Absorbances were read at a wavelength of 405 nm. The concentrations for G7 inhibitor that were used are 0.005 mM, 0.01 mM, 0.05 mM, 0.1 mM, 0.2 mM, 0.5 mM, 0.75 mM, 1 mM, 2 mM, and 4 mM. For G6, the concentrations 0.005 mM, 0.01 mM, 0.05 mM, 0.1 mM, 0.5 mM, 1 mM, 1.5 mM, 2 mM, 4 mM and 8 mM were used in the IC_{50} calculations.

Inhibitor Assays In vitro: Determining Reversible and Covalent Inhibitors

Inhibitor concentrations of 300 μ M, 600 μ M, and 1000 μ M and 30 μ M, 100 μ M and 300 μ M in the assay, were tested to determine a concentration yielding 30-50% inhibition for G6 and A9, respectively. Assays were then run under the same conditions as previous pNPP assays.

Once this concentration was determined, assays were conducted to determine the reversibility of each inhibitor. The assays were performed in 75% enzyme reaction buffer, 10% enzyme, 10% pNPP, and

Developing a Novel Inhibitor of Cdc14s in C. zeina

5% inhibitor. The enzyme was allowed to incubate with the inhibitor for time intervals of 0 minutes, 15 minutes, 30 minutes, 45 minutes, and 60 minutes. pNPP was added to start the reaction after the given pre-incubation time. Due to varying inhibitor concentration, the conditions for the assays differed as well. For G6, a concentration of 0.3 μ M enzyme, 20 mM pNPP, and 300 μ M inhibitor in stock were used for proceeding experiments. For A9, a concentration of 1 μ M enzyme, 600 mM pNPP, 300 μ M inhibitor in assay were used. Inhibitor was diluted using DMSO and pNPP with DEA to yield the correct concentrations stated above.

Reactions were then quenched with 50 μ L of 5 N NaOH. A set of blank assays was measured for absorbance as well. Each blank assay was performed in 85% enzyme reaction buffer, 10% pNPP, and 5% inhibitor at the corresponding concentrations. The absorbances were measured using a wavelength of 405 nm.

Positive and Negative Controls

Positive and Negative Controls were also performed to determine the dynamic range of CzCdc14 in the presence of an inhibitor for the inhibitor-screening assay, IC₅₀ experiments, and reversibility/irreversibility trials. A Z-factor was calculated for the initial 11 inhibitor screenings in order to ensure the statistical significance between 100% Activity of the enzyme (Positive Control) and 0% Activity (Negative Control). Positive Control Assays consisted of 75% Enzyme Reaction Buffer, 10% Enzyme, 5% Dimethyl Sulfoxide (DMSO), and 10% pNPP. Negative Control Assays contained 85% Enzyme Reaction Buffer, 10% pNPP, and 5% DMSO. Negative Control differed for the reversibility/irreversibility experiment enzyme assay once the final concentrations to use for each inhibitor as it contained 90% Enzyme Reaction Buffer and 10% pNPP.

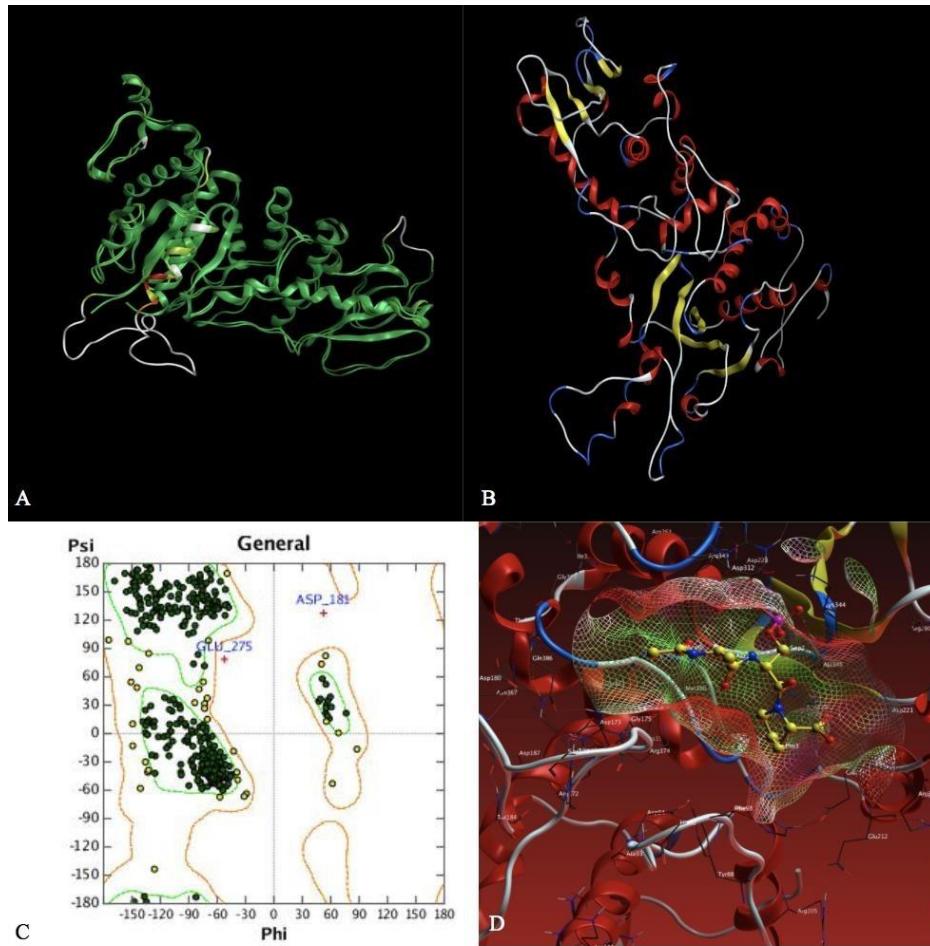


Figure 1: Homology Modeling of *CzCdc14* from Template 5XW5 Crystal Structure. *A.* *CzCdc14* homology model developed from crystal structure of 5XW5 crystal structure obtained from the Protein Data Bank (PDB). Polypeptide is shaded by root mean squared deviation (RMSD). Green regions refer to areas of low RMSD and red regions areas of high RMSD. Grey regions have no alignment with template structure. *B.* *CzCdc14* shaded by peptide structure. Red regions represent α helices, yellow regions represent β sheets, and blue regions represent turns. *C.* ϕ/ψ plot (Ramachandran Plot) of *CzCdc14* homology model. 8 unfavorable backbone angles were found. The favorable peptide conformations are beta-sheets and right-handed alpha helices. Left-handed alpha helices are less favorable but allowed. *D.* The active site pocket of *CzCdc14* after alignment and superpositioning with 1OHE. The phosphoserine ligand from 1OHE is positioned within the active site of *CzCdc14*.

RESULTS

Inter-Species Protein Comparisons

Prior to the empirical determination of its enzymatic properties, we sought to understand the properties of the active site of CzCdc14 using *in silico* models. The online resource Clustal Omega from EMBL-EBI was used to align multiple sequences of Cdc14s that were identified as similar from the Basic Local Alignment Search Tool (BLAST). Cdc14s from *Homo Sapiens*, *Drosophila melanogaster*, *Caenorhabditis elegans*, *Saccharomyces cerevisiae*, and *Schizosaccharomyces pombe* were analyzed. It was found that there was strong alignment between CzCdc14 and other Cdc14s indicating that the protein would likely perform similar functions to these proteins. The following amino acid motif was found to be strongly conserved across all species: 5'AVHCKAGLGRTG-3'. We believe that this constitutes the active site of the residues (see Supplemental Figure 1B).

Homology Model Development

Comparative modeling algorithms were applied to the obtained amino acid sequence for CzCdc14 using the Molecular Operating Environment (MOE). It was determined that the query was closely aligned with the 5XW5 enzyme from *Saccharomyces cerevisiae* using the Protein Data Bank (PDB). The percent identity was 38.4%, while the percent similarity was 50.8%. As a result of this degree of similarity, the comparative modeling algorithms developed a target protein structure that closely resembled the template (Figure 1A). Root-mean-square deviation (RMSD) was then calculated and displayed on the target protein structures (Figure 1B). In addition, a Ramachandran plot was created and unfavorable ϕ/ψ (phi/psi) angles were analyzed (Figure 1C). Unfavorable bond lengths, bond angles, and rotamer energies of the protein were also measured. The target and template models maintained relatively low RMSD values throughout the protein, however significant deviations from the template were observed from target residue numbers 239-265. No alignment between the target and template protein strands was found from target residue number 266-280. Eight residues had outlier dihedral backbone angles that did not fall within allowed regions of the computed Ramachandran Plot (see Supplemental Figure 1C). Eighteen outlier backbone dihedral angle values (at Z-score threshold of 4) and no outlier rotamer energies were computed (see Supplemental Figure 1C).

Active Site Identification

After obtaining a 3D structure of the protein, we sought to determine the location of the active site on CzCdc14. Using an additional PDB search, CzCdc14 was found to be closely aligned with the Cdc14 1OHE, a humanoid protein that contained a phosphopeptide (phosphoserine) docked at the active site. Using the MOE, the template and query sequences were aligned and then superimposed. The receptor site of 1OHE with its ligand was then identified to highlight the active site of CzCdc14 (Figure 1D). It was found that the active site of CzCdc14 occurred around the Cysteine 343 and Arginine 349 residues. It was also found that there are multiple hydrophobic residues within this active site. These residues were similar to those found conserved in the Clustal Omega alignment search.

Protein Purification via Nickel Affinity

Having established an *in silico* model, empirical analysis of the protein began with purification by nickel-affinity chromatography. Absorbance measurements were taken at a wavelength of 280 nm as the volume of product increased (Figure 2A). As shown in Figure 2A, the various stages of the chromatography process removed large amounts of extraneous material, isolating the protein by the final peak (region C). The high absorbance measured in peak B represents other protein fragments that were able to bind to the Nickel column. As a result of the graph produced from the HPLC start machine, fractions 8-13 contained our purified protein. The SDS-PAGE gel confirms the relative purity of the enzyme samples (Figure 2B). As shown in Figure 2B, at each step of the purification process, the quantity of constituent components decreased. The single band at column 5 represents the purified protein with no additional components. It was found that the protein had a molecular weight of approximately 66.2 kD or 66200 amu. The computational analysis of CzCdc14 on the database ExPASy revealed that CzCdc14 had a molecular weight of 58450.17 amu and an isoelectric point (pI) of 8.89. This is significantly different from the molecular weight determined through the SDS-PAGE gel electrophoresis. Because the isoelectric point of CzCdc14 is 8.89 and the pH of the SDS-Page buffer and stacking SDSPage gel were 8.3 and 6.3, respectively, it is likely that the purified protein traveled at different rates depending on different rates in the varying pH regions, if it was protonated or deprotonated.

Protein Concentration Determination

Developing a Novel Inhibitor of *Cdc14s* in *C. zeina*

To determine the concentration of filtered CzCdc14 protein, two independent procedures were pursued: Bradford Assay and Nanodrop reading. The A_{595} absorbance for the 25x enzyme dilution was found to be 0.194 AU. The corresponding concentration was found to be 6.4×10^{-5} M. Nanodrop readings were performed for 2 additional enzyme dilutions (10x and 25x) and A_{280} values calculated. The 10x dilution had an absorbance of 0.357 AU and the 25x dilution had a concentration of 0.161 AU. As these absorbances differed by a factor of approximately 2.2 and our dilutions a factor of 2.5, it was determined that these readings were valid. The average dilution-factor-corrected concentration was 5.46×10^{-5} M. Due to the similarity in concentration values from the Bradford assay and Nanodrop, it was determined that both readings were valid. The Nanodrop concentration was used as the concentration of CzCdc14 for the remaining experiments.

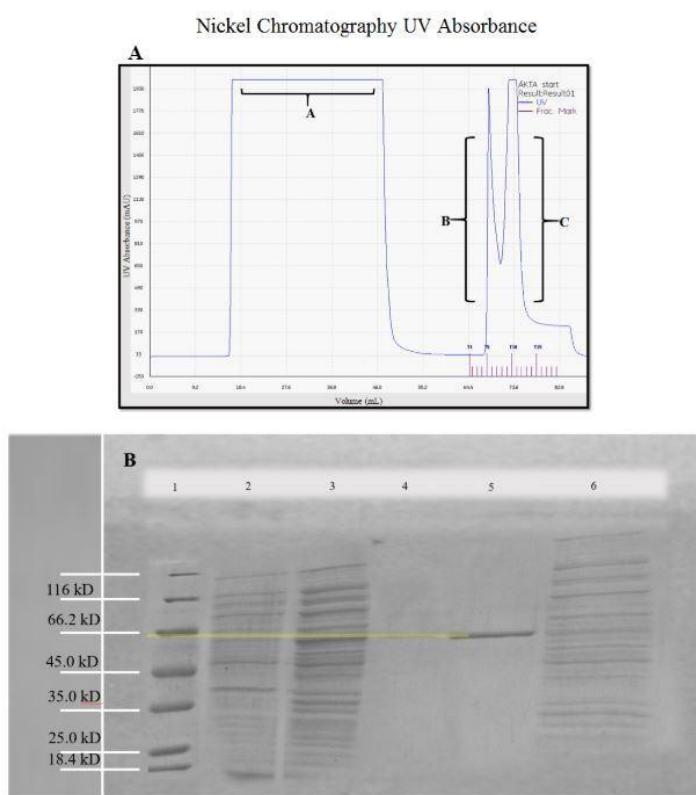


Figure 2. Methods of Protein Purification and Protein Concentration Measurement. *A*. A High Performance Liquid Chromatography (HPLC) machine was used to purify the protein. Readings were taken after Nickel Chromatography to isolate the enzyme at A_{280} . Region A corresponds to the elution of "flow through" proteins that do not contain Histidine. Region B corresponds to the elution of additional peptides containing Histidine residues to the Nickel Column. Region C displays the elution of *CzCdc14* from the Nickel Column, as well as the corresponding fractions into which enzyme was stored. *B*. An SDS-PAGE gel was used to determine the molecular weight of *CzCdc14*. Lane 1 designates the Molecular Weight Ladder of Bovine Serum Albumin (BSA). Lane 2 contains Pre-Induced Protein. Lane 3 contains the Pre-Column protein sample. Lane 4 contains the Pre-Dialysis sample. Lane 5 contains Purified Protein/Post Dialysis sample. Lane 6 contains the flow-through sample. The darkened band on Lane 3 band demonstrates the presence of *CzCdc14* post induction of *E. coli* cells by L-Arabinose. Dark lines in Lane 4 and 5 confirm the presence of protein both pre- and post- dialysis. The faint line in Lane 6 confirms that small quantities of enzyme were lost in the flow-through.

Developing a Novel Inhibitor of Cdc14s in *C. zeina*

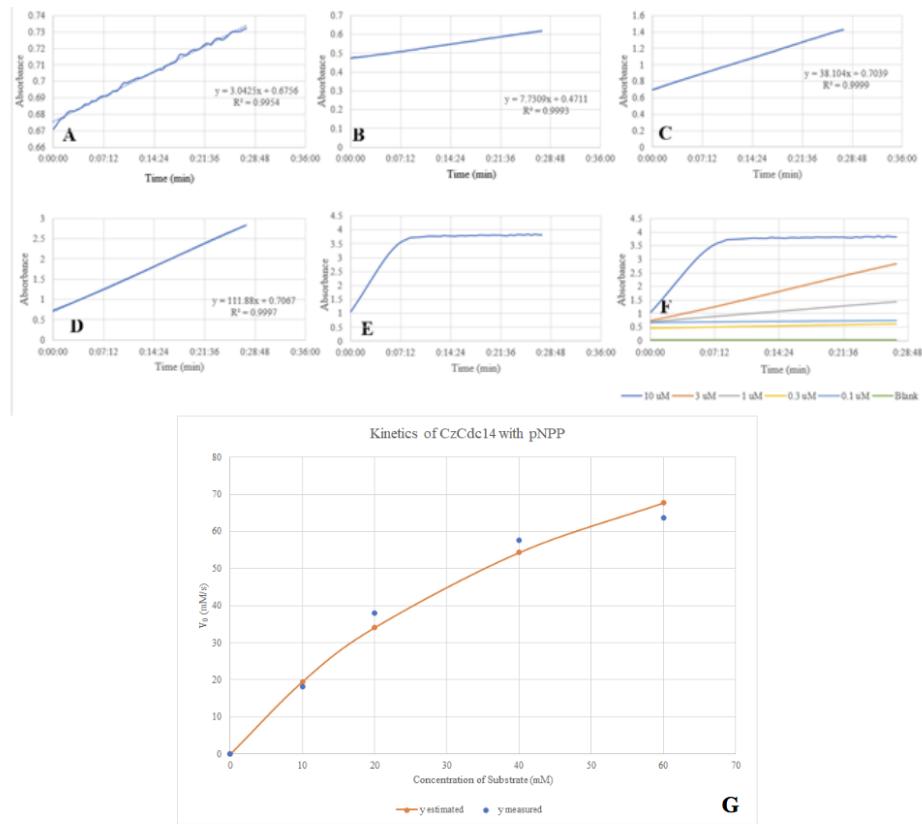


Figure 3. Determining Enzymatic Behavior at Steady State and Michaelis-Menten Kinetics Assay. Five concentrations of enzyme were tested to determine which kinetics assay yielded the strongest linear results. All pNPP kinetics assays were completed at A_{405} . Each trial was run in singlet and the reaction allowed to run for approximately 35 minutes, or until deviations from linearity were observed. *A*, 0.1 μ M: Relative linearity was observed throughout the trial and an R^2 value of 0.9964 was generated. It also contained a relatively high background absorbance at 0:00 min (0.675 AU). This was consistent throughout our trials, however, so it was attributed to default noise from the spectrophotometer. *B*, 0.3 μ M: A strong linear relationship was observed. *C*, 1 μ M: The strongest linear relationship was observed in this trial. It was not chosen for the final enzyme concentration, however, due to its comparatively smaller dynamic range. *D*, 3 μ M: This trial had a strong linear relationship and large dynamic range; it was chosen as our default enzyme concentration. *E*: 10 μ M: This trial yielded a strong digression from linearity and thus a linear regression was not calculated. *F*: All trials are shown for the relative comparison of linearity. A blank was run and is shown in red. *G*: Concentrations of pNPP substrate at 80 mM, 60 mM, 40 mM, 20 mM, 10 mM, 4 mM, and 1 mM in solution were combined with 0.3 μ M enzyme and run for 10 minutes. V_0 was calculated using the pNPP standard curve (see Supplemental Figures). Curve fitting techniques were applied and a V_{max} value of 133.29 μ mol/s and K_m of 58.23 mM were calculated. These values could not be empirically calculated due to a maximum concentration of stock pNPP at 800 mM (wells were diluted 10x). 1 trial was conducted.

Steady State Kinetics Assay

The results of the pNPP assay indicated that CzCdc14 exhibited behavior highly consistent with other Cdc14s. Concentrations of CzCdc14 as low as 0.1 μ M showed strong linear correlations with time and the production of PNP (Figure 3A). Regression values for the enzyme concentrations of 0.1 μ M, 0.3 μ M, 1 μ M, and 3 μ M were all above 0.99, indicating high enzymatic activity (Figure 3B-3D). The 10 μ M assay yielded a flattening curve after 8 minutes indicating that the enzyme had exhausted most of the available substrate, another indicator of strong Cdc14 activity (Figure 3E). From the experimental data, it was determined that 3 μ M would be the default enzyme concentration to use in future trials, as it produced a linear regression value of 1 and maintained a wide dynamic range. It should be noted that this concentration value would change depending on the day-to-day activity of the enzyme. A trial containing

blanks was also run and it yielded stable data, which ensured that previously observed enzymatic activity was not significantly affected by external factors (Figure 3F).

Michaelis-Menten Kinetics Assay

To more directly examine the enzymatic properties of CzCdc14, an additional kinetics assay was performed to determine how the initial rate of catalysis varied with substrate concentration. Specific activity was calculated and found to be 8.9876 pmol product/ pmol enzyme * min. Figure 3G represents the Michaelis-Menten curve for Cdc14 and pNPP. Curve fitting algorithms found the maximum catalytic velocity (V_{max}) to be 133.29 mmol*s⁻¹ and the Michaelis-Menten constant (K_m) to be 58.23 mM. It should be noted that the V_{max} and K_m that could be observed from Figure G are not correct. Due to the high activity of the CzCdc14 and laboratory material limitations, concentrations of substrate above 80 mM could not be achieved. Instead, curve fitting algorithms calculated these values. Future experiments used the calculated K_m and produced reliable results.

Determining Substrate Specificity of CzCdc14

In the trial involving 24 phosphopeptides, the quantity of product formed was measured from absorbance. A standard curve (Supplemental Figure 4A) was used to determine the concentrations of product from the measured absorbances. Initial rate parameters, specifically the slope of the curve, were measured and represented the specificity constant k_{cat}/K_m (Table A). This could be done due to the Michaelis-Menten assumption that substrate concentration was significantly less than K_m , allowing for $V_0 = V_{max}/K_m$. A modified Michaelis-Menten equation was used to characterize the Michaelis-Menten graph: $V_0 = V_{max}/K_m$. Overall, the specificity constants were relatively high, having an average value of around 30,000 M⁻¹s⁻¹. Phosphopeptides numbers 1, 7, 8, 9, 14, 15, 16, 17, 19, 20, 21, 24 were proven to be good substrate since they contain phosphoserine, lysine, and some surrounding basic residues. Specifically, substrate 1, 7, and 24 exhibit the highest substrate specificity (Figure 4A).

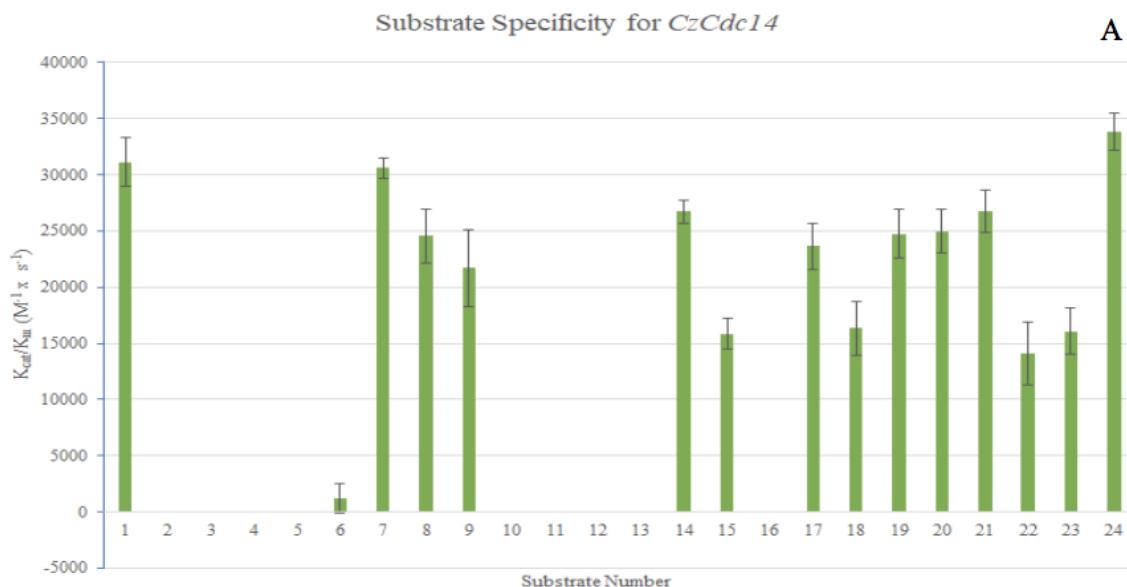


Figure 4. Inhibitor Screening Assay of the CzCdc14 with pNPP. A. 24 substrates were tested containing amino acid sequences shown in Table A. Readings were read at A_{640} for 1200 seconds. All substrates were at a concentration of 0.1 mM and run for 10 minutes. The enzyme was at 0.1 μ M. Error bars represent standard error of the mean. As shown in the bar plot of substrate number versus the specificity constant, substrates 1, 7, and 24 contained the highest specificity constants and thus have the highest binding affinity for our enzyme.

Developing a Novel Inhibitor of *Cdc14s* in *C. zeina*

Table 1: Substrate Sequences and Resulting Specificity. Column 2 represents the amino acid sequence for each substrate tested. Phosphopeptide 1 was the optimal inhibitor as it contained a sequence that was most conducive to binding to the predicted *CzCdc14* active site. Column 3 represents the average absorbance from the three trials done. Column 4 represents the amount of product formed (as a result of the amount of Biomol Green formed). Column 5 represents the final specificity constant calculated for each substrate. Negative concentrations of product were counted as 0.

Substrate	Sequence	Average Absorbance	[P] (μM)	K _{cat} /K _m (M ⁻¹ s ⁻¹)
1	HT{pS}PIKSIG	0.3660	373.6	31135.8
2	HT{pT}PIKSIG	0.001333	0	0
3	HT{pY}PIKSIG	0.004000	0	0
4	HT{pS}AIKSIG	0.001000	0	0
5	HT{pS}PIASIG	0.001000	0	0
6	HT{pS}PIRSIG	0.05167	14.80	1233.5
7	HT{pS}PIKKIG	0.3603	367.2	30596.8
8	HT{pS}PKKSIG	0.2967	294.5	24540.2
9	HT{pS}PIKSKG	0.2670	260.6	21718.0
10	HT{pS}PKASIG	0.001333	0	0
11	HT{pS}PIAKIG	0.001667	0	0
12	HT{pS}PIASKG	0.0006667	0	0
13	HT{pS}PIASIK	0.003333	0	0
14	AT{pS}PIKSIG	0.3197	320.7	26728.2
15	HA{pS}PIKSIG	0.2050	189.8	15820.0
16	HT{pS}KIKSIG	0	0	0
17	HT{pS}PEKSIG	0.2870	283.4	23620.6
18	HT{pS}PIKEIG	0.2103	195.9	16327.4
19	HT{pS}PIKSEG	0.2987	296.77	24730.5
20	HT{pS}PIKSIE	0.3013	299.8	24984.1
21	HT{pS}PIKSIK	0.3197	320.7	26728.2
22	KT{pS}PIKSIG	0.1867	168.9	14076.0
23	HK{pS}PIKSIG	0.2077	192.9	16073.7
24	HT{pS}PIKKRG	0.3943	406.0	33831.2

Docking and Optimization of Inhibitor In Silico

Using MOE and the homology model generated from 5XW5 and ligand 1OHE, minimization and protonation algorithms were applied to improve the binding of the phosphopeptide ligand to the homology model of *CzCdc14*. It was found that four key interactions occurred between active site amino acids and ligand. Sidechain acceptors included Cys343 and Asp312 and backbone acceptors included Arg 349 and Leu 347. After docking the 11 inhibitors, it was noted that three inhibitors could not be docked by MOE and were eliminated: G6, H7, and I2. Ligand interactions were also evaluated and similarities in interactions with the original phosphopeptide were looked for. Three inhibitors were found to have the greatest amount of similar interactions with the original 1OHE phosphoserine ligand. These included (in order from most similar to least) inhibitors I1 (S value: -5.7716), D3 (S value: -5.7600), and D7 (S value: -5.8390).

Using the phosphate group in the phosphopeptide in the pocket of the homology model, we again docked our three best inhibitors into the pocket using Pharmacophore function. Interactions between Arg349 and two oxygen atoms of the phosphate group of the ligand were inputted as criteria. After the docking process, we inspected the interactions between the drug molecule and nearby residues, especially the Arg 349 residue since the guanidino group in arginine is capable of making various hydrogen bonds and Coulombic interactions with other molecules. We found that I1 has the best interactions with the

surrounding residues around the pocket (Figure 5A, D). After minimization, its properties were recorded (see Table 2).

Optimization occurred on I1 following the criteria previously mentioned. The following groups were added to our drug molecule: a sulfonamide group, two amino groups, a cyclopropanecarbonyl group, a calicene carbonyl group, a cyclopentyl group. The resulting drug reached a peak affinity score of -12.00 kcal/mol, making it a potential drug for our fungal pathogen (see Figure 5B-C, E).

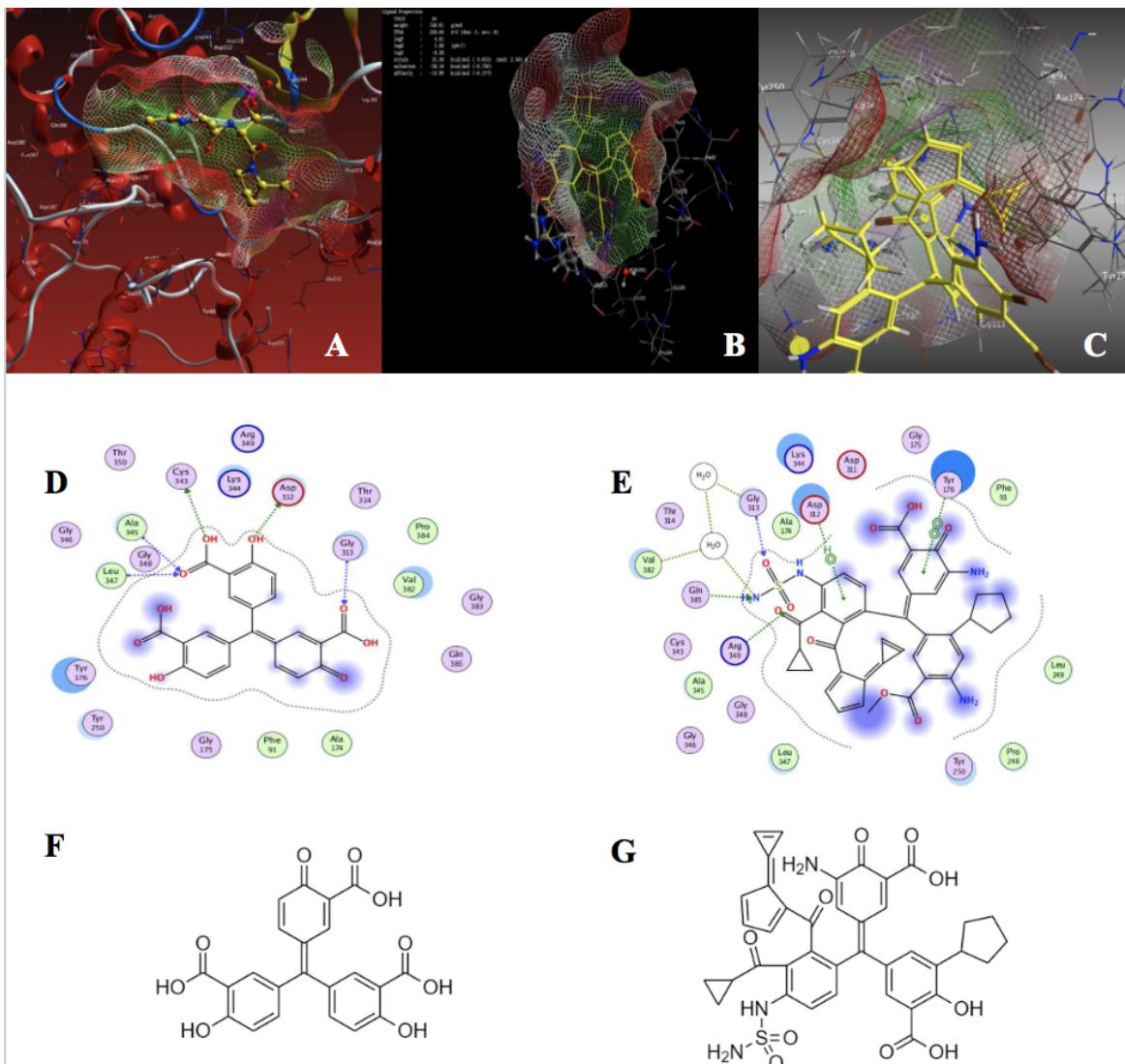


Figure 5. *In Silico* Docking and Optimization of Inhibitor to *CzCdc14*. *A*. Phosphoserine ligand shown docked to the active site pocket of *CzCdc14*. The pocket is shaded by electrostatic interactions. *B*. Final optimized I1 inhibitor bound to the pocket which is shaded by electrostatic interactions (mesh region). *C*. The optimized I1 inhibitor in the active site. Several intermolecular interactions can be observed. Hydrophobic pockets of receptor are filled by carbon rings on inhibitor. *D*. The intermolecular interactions between un-optimized I1 inhibitor and active site. *E*. Intermolecular interactions between optimized I1 inhibitor and *CzCdc14*. *F*. Molecular structure of I1 inhibitor. *G*. Molecular structure of optimized I1 inhibitor.

Relative Inhibition Screening Assay

After determining optimal substrates, screening for effective inhibitors were conducted. Eleven inhibitors were obtained and screened for percent inhibition on CzCdc14. A Z-factor of 0.876 was calculated from the positive and negative control experiments for a concentration of enzyme at 1 μ M. A dynamic range of approximately 0.84 AU₄₀₅ was produced, indicating that the inhibitor assay results would be statistically significant. Percent inhibition was compared across inhibitors and several effective inhibitors were identified (Figure 6A, Table 3). Inhibitors G7, I5, and I2 each had percent inhibitions above 60%. G6 was also a relatively effective inhibitor. H7 and I9 were not considered in the determination of an effective inhibitor due to an error that occurred with their solution preparations that resulted in each having a well concentration 10 times that of others. We found inhibitors G7, G6, and I2 to be the best inhibitors. While such error bars are relatively large, the 3 most effective inhibitors had higher average percent inhibitions, comparatively smaller error bars, and are thus more conclusively effective than other inhibitors.

Table 2: Properties of inhibitor II before and after modification. Row 1 represents II (Aurintricarboxylic acid) before our modification with MOE. Row 2 represents our designed inhibitor. Note that the affinity for our designed drug is not -12.00 kcal/mol since -9.60 is the common value for our drug. However, in certain conditions, the value can be as good as -12.00 kcal/mol.

	Molecular weight (g/mol)	Hydrogen bond donor	Hydrogen bond acceptor	logP	Affinity (kcal/mol)	Ligand Efficiency
II Pre-modification	422.34	5	9	2.97	-6.64	-0.214
II Post-modification	748.81	5	8	4.91	-9.60	-0.178

Determination of IC₅₀ values for G6 and G7 Inhibitors

IC₅₀ experiments were conducted to quantify the effectiveness of inhibitors G6 and G7, each at a concentration of 0.1 μ M. Due to a scarcity of stock inhibitor, the optimal inhibitors identified in previous experiments were not screened. To determine the concentrations of inhibitor to use, the concentration of inhibitor used and relative inhibitions from the previous inhibitor assay were referenced. G6 was tested at concentrations of 0.005 mM, 0.05 mM, 0.1 mM, 0.2 mM, 0.5 mM, 0.75 mM, 1 mM, 2 mM, and 4 mM. G7 was tested at inhibitor concentrations of 0.005 mM, 0.01 mM, 0.05 mM, 0.1 mM, 0.5 mM, 1 mM, 1.5 mM, 2 mM, 4 mM, 8 mM. Multiple trials were conducted to determine the IC₅₀ value of each inhibitor, however unpredictable enzyme behavior resulted in poor results being obtained. Conclusive data for both the G6 and G7 was eventually collected. Curve fitting algorithms were then applied to our data to determine the IC₅₀ values for G6 and G7. The IC₅₀ value for G6 was found to be 450.0 μ M and slope -2.00 (Figure 6D). The IC₅₀ value for G7 was found to be 343.49 μ M and slope -0.99 (Figure 6C). These values demonstrate that G6 and G7 are poor inhibitors at the tested conditions. Nonetheless, future experiments at more stable enzymatic conditions may corroborate our inhibitor screening assay results.

Determination of the Mechanism of Inhibition for A9 and G6 Inhibitors

Developing a Novel Inhibitor of *Cdc14s* in *C. zeina*

Reversibility assays were then conducted to determine the mechanism of inhibition for the A9 and G6 inhibitors. These inhibitors were investigated as part of a larger effort to chronicle the mechanisms of inhibition of all 9 inhibitors analyzed in the inhibitor assay with other investigators, and thus these inhibitors do not match the inhibitors previously studied. After determining the optimized conditions for each inhibitor assay, reversibility assays were conducted in singlet over a period of 60 minutes with 15 minute intervals (Figures 6E-F). It was found that G6 bound reversibly with our enzyme (Figure 6D) and that A9 bound irreversibly (Figure 6E). The non-straight line for G6 and a slight dip in the A9 reversibility graph seen at the 15 minute mark was attributed to pipetting error.

Table 3. Inhibitor Assay Relative Inhibition.

Inhibitor	Average Absorbance	Relative Inhibition (%)
A9	0.450	46.6
D3	0.411	51.2
D7	0.441	47.7
E1	0.396	53.1
I5	0.229	72.9
G6	0.393	53.4
G7	0.213	74.8
I1	0.370	56.1
I2	0.306	63.7
H7	0.145	82.8
I9	0.341	59.6

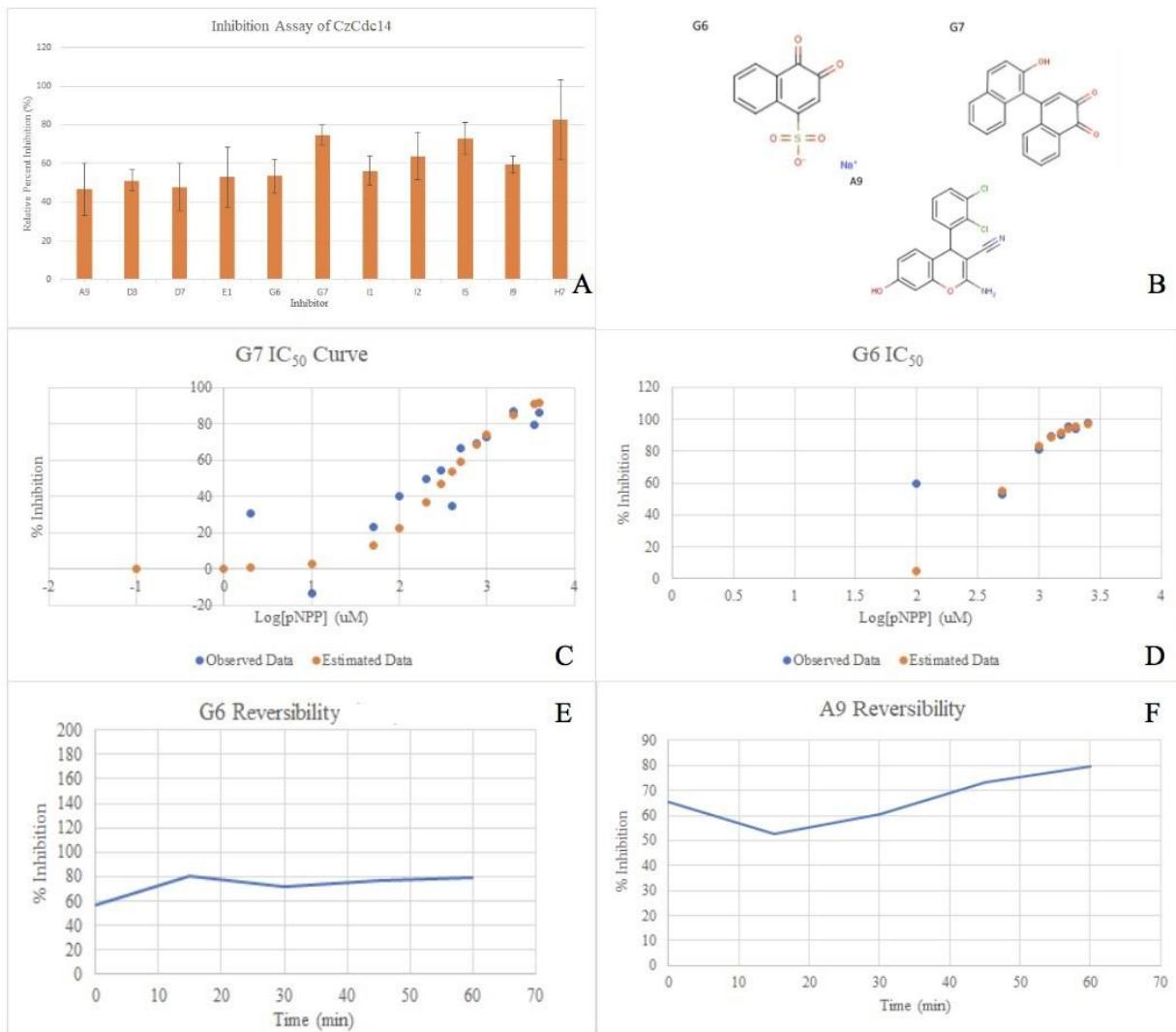


Figure 6. Determination of IC₅₀ values and Reversibility of Selected Inhibitors with CzCdc14: *A*. All assays containing pNPP were measured at a wavelength of 405 nm. An enzyme concentration of 0.3 uM and a pNPP concentration of 60 uM were used for three trials. All inhibitors, with the exception of H7 and I9, were tested at 0.1 mM in solution. Inhibitors H7 and I9 were each tested at 1 mM in solution. Error bars represent standard error of the mean. *B*. The structures of the top two inhibitors tested are shown to the right (G6 and G7) as well as inhibitor A9 which was used for reversibility experiments. *C*. The IC₅₀ curve for the G7 inhibitor trial was conducted with CzCdc14 (0.1 uM) at concentrations of 0.005 mM, 0.01 mM, 0.05 mM, 0.1 mM, 0.2 mM, 0.5 mM, 0.75 mM, 1 mM, 2 mM, and 4 mM. An IC₅₀ value for pNPP substrate at 60 mM was found to be 343.29 uM, indicating poor inhibition. *D*. The IC₅₀ curve for the G6 inhibitor trial was conducted at the same standard conditions as G7 with concentrations of 0.005 mM, 0.01 mM, 0.05 mM, 0.1 mM, 0.5 mM, 1 mM, 1.5 mM, 2 mM, 4 mM, and 8 mM. An IC₅₀ value for pNPP substrate at 60 mM was found to be 420.21 uM. *E*. A reversibility curve for G6 was generated using conditions empirically found to yield 30-50% inhibition in previous control trials. Assays were performed using 0.3 uM CzCdc14, 20 mM pNPP, and 300 uM inhibitor. The relatively horizontal curve indicates that G6 was reversible for CzCdc14. *F*. A reversibility curve of A9 at the following previously determined conditions was found to yield a irreversible curve: 1 uM CzCdc14, 60 mM pNPP, 300 uM inhibitor.

DISCUSSION

In this study, our team investigated the enzymatic properties of CzCdc14 in order to predict and design an effective competitive inhibitor to combat the dangerous impacts of fungal disease. We consistently found that the enzyme has a highly specific catalytic conditions, reacting differently with changes in pH and temperature. We confirmed that the enzymatic activity of CzCdc14 resembled other Cdc14 enzymes and determined the steady state kinetic parameters necessary for proceeding trials. Results concerning the specific residues in the active site from *in vivo* trials as well as *in silico* comparative modeling allowed for the design of an optimized inhibitor to CzCdc14. The results of the study contribute to the necessary future research endeavors that will combat crop disease caused by fungal pathogens.

Inter-Species Protein Comparisons

The results from the Clustal Omega search revealed that CzCdc14 possessed a highly conserved domain: 5'-AVHCKAGLGRTG-3'. This finding is hypothesized to be the active site domain of Cdc14s. This sequence contains the CX₅R motif which has been previously identified as a “hallmark of protein tyrosine phosphatases [8].” This correlation may be used as evidence to confirm that the isolated enzyme is indeed a protein tyrosine phosphatase. In addition, the high conservation of this active site may suggest that the function of Cdc14s is critical for biological function across many species. Thus, the function of regulating the mitotic cycle in other species may also be the function of the CzCdc14 in fungi. This critical function could make CzCdc14 a viable target for more fungicide development research.

Homology Model Development

The percent identity and percent similarity results that were calculated using the Molecular Operating Environment demonstrated that the 5XW5 *Saccharomyces cerevisiae* sequence produced a reliable homology model to demonstrate our target protein. The aligned and superimposed model was predominantly green except for a 26 amino acid long region (residues 239-265) and the lack of alignment in residues 266-280. These differences are hypothesized to be caused by the mobility of the protein, specifically around the A-loop. While the model was concluded to be viable, it should be noted that the unfavorable dihedral backbone angles, identified through the Ramachandran Plot indicate that the model likely does not represent the naturally occurring conformation of CzCdc14.

Active Site Identification

By aligning and superimposing the pocket of the homology model with the model of 1OHE, the intermolecular interactions between CzCdc14 and the phosphoserine ligand from the 1OHE X-ray crystallography structure could be identified. We hypothesize that Arg349 found within the active site functions to deprotonate the Cys343 residue. This will make Cys343 nucleophilic to the potential inhibitor. The presence of multiple hydrophobic residues is particularly important due to the fact that it has been found that hydrophobicity is critical in the substrate specificity within the family of Cdc14 phosphatases. The importance of the cysteine and arginine in CzCdc14 was observed in the ligand-receptor interactions between 1OHE and its ligand. These findings may suggest a mechanism for the dephosphorylation of phosphopeptide ligands among Cdc14s. The large number of hydrophobic residues in the active site may also corroborate previously reported findings that Cdc14s prefer ligands with two or more aromatic rings [9]. This finding could also explain why the single-ringed pNPP was a poor substrate for our enzyme.

HPLC Purification

Developing a Novel Inhibitor of Cdc14s in *C. zeina*

The results of our SDS-PAGE gel confirmed that our protein was purified, as a single band was shown in Column 5. We can thus conclude that experimental irregularities were likely not significantly influenced by impurities. The different molecular weights from the SDS-PAGE gel and ExPASy were likely caused by the broad pH gaps among buffers as these would cause the protein to travel at different rates depending if it was protonated or deprotonated. Protein Concentration Determination Our team decided to proceed with the concentration measured by the Nanodrop instrument due to the fact that the Bradford Assay had more possibility for human error by pipetting. After further investigation of this enzyme, it was observed that it had a tendency to precipitate, being highly sensitive to changes in pH and temperature. Our team had to adjust enzyme concentration day-to-day depending on the activity of the enzyme aliquot being used. As seen with the enzyme's significantly high K_m value compared to that found in previous literature, our group hypothesizes that the enzyme is not only sensitive to changes to its environment but as well as, is very hyperactive. Because of this, we decided that for every new experiment, we would perform 1:1 dilutions of the aliquots using Storage Buffer. This would ensure zero to very little protein precipitation and a homogenous concentration within the aliquot being used.

Steady-State Kinetics Assay/ Michaelis Menten Kinetics Assay

The inhibition of enzyme activity observed with the addition of sodium orthovanadate, a known protein tyrosine inhibitor, confirms that the enzyme functions to remove a phosphate group from a substrate. Because of this, we know that the enzymatic activity corresponds to that found in previously studied Cdc14 enzymes.

Although it is not clear that the Michaelis Menten Kinetics Assay graph levels off demonstrating an apparent V_{max} , we postulate that this occurred due to the enzyme's hyperactivity. The characteristics of pNPP cause it to be an ineffective substrate for CzCdc14. pNPP was used to measure the ability of the enzyme to cleave a phosphate group by producing a yellow-colored compound of *para*-Nitrophenol. Nevertheless, the small size and lack of phosphopeptides of pNPP cause it to not bind to the enzyme efficiently. This would explain the large K_m . We hypothesize that if had been able to test higher concentrations of pNPP, the enzyme would have been able to reach its true V_{max} and the Michaelis Menten graph generated from the data would begin to level off. It is still possible that the computed V_{max} and K_m are inaccurate and that higher concentrations of pNPP would reduce initial velocity. The computed K_m was significantly higher than the K_m of Cdc14s in budding yeast with pNPP (10.6 ± 0.9 mM) [8].

Substrate Screening Assay

The screening of 24 phosphopeptides with CzCdc14 allowed us to determine specific motifs within effective substrates as well as identify the residues that interact those motifs. Our protein CzCdc14 expressed a higher substrate specificity for our phosphopeptide when compared to pNPP substrate. Specifically, phosphopeptides 1, 7, 8, 9, 14, 15, 16, 17, 19, 20, 21, and 24 was shown to be good substrate for CzCdc14 (Table A). Among these phosphopeptides, substrates 1, 7, and 24 demonstrated the highest substrate specificity for CzCdc14, since they contain a phosphoserine residue which is also phosphorylated in a similar manner compared to tyrosine (Serine and tyrosine both contains a hydroxyl (-OH) group). Moreover, the 3 best substrates also have basic lysine residues, which likely deprotonates the cysteine in the active site, making it more nucleophilic. These findings are consistent with previously published literature that describes the optimal substrates for Cdc14s as possessing the following criteria: (1) phosphoserine (2) lysine in the +3 position relative to the p{S} (3) several basic residues surrounding the +3 K [9]. We found that the substrates that yielded the highest specificity efficiencies possessed all of these traits.

Determination of IC₅₀ values for G6 and G7 Inhibitors

The high IC₅₀ values found for the G6 and G7 inhibitors were caused by the erratic enzyme behavior – in which the enzyme would denature or become hyperactive sporadically. The use of a low enzyme concentration may also have contributed to these results. The exact causes for erratic enzyme behavior is unknown. It is hypothesized that the pH of the well solutions impaired enzymatic activity by denaturing the protein. Certain buffers and solvents used, such as diethanolamine, may have changed the pH in well. We hypothesize that the previously discussed poor pNPP binding affinity to the enzyme also contributes to the large IC₅₀ values. These values were significantly larger than those recorded by other investigators studying CzCdc14 and other Cdc14s. Thus, we do not believe that these results are completely conclusive.

In silico Inhibitor Design

A common moiety in our model inhibitor is a Michael acceptor, which consists of a C=C double bond in conjugation with an electron-withdrawing group (such as carbonyl C=O or nitrile C≡N). The cysteine residue in CzCdc14's active site has a thiol backbone, which under deprotonation by surrounding basic residues would become a thiolate anion, a soft nucleophile which would attack the soft carbon electrophile in the Michael acceptor in a 1,4 addition (Michael addition) reaction. When designing our inhibitor, we found that inhibitor I1 was an effective inhibitor for CzCdc14 for its low S value and multiple ligand interactions. As a result, we decided to use it as the starting core structure to further modify.

Reversibility

For our experiment on reversibility and irreversibility, we found out that G6 was a reversible inhibitor despite the fact that the carbon electrophile is bonded with a strong electron-withdrawing group (sulfonate group SO₃⁻). We hypothesized that G6 was such a small molecule that cannot fit tightly into the large pocket of CzCdc14, therefore it could not coordinate with the cysteine residue in the right orientation for the cysteine thiol to attack the carbon electrophile. We also discovered that G7 is an irreversible inhibitor, possibly due to its large size that helps to fit in the large pocket, and the electron stabilizing effect of the 2naphthol ring.

Our team believes that more research should be conducted to further analyze *Cercospora zeina*'s role as a fungal pathogen affecting the food security in the developing world. Successive teams should further research the characteristics of the active site of CzCdc14 and possibly explore other binding regions on the enzyme, as they may be more effective at inactivating the enzyme. Overall, follow up research and developments of inhibitors should be performed as world hunger continues to affect the lives of millions all over the world.

SUPPLEMENTARY FIGURES

FIGURE 1A. DNA Sequence for *CzCdc14*:

5' - ATGCCGCACCAGTCCCCTCCAGCGTGCAGGAAGCGTCTGCCATGACTTCGTTGGGCAG
GTCATTGAATAACATCCAAGATCGTCTGTACCTCGCATACACGCACACTCCGACGAG
AGCACGCCGTTCCATACCCAGCGAACAGAGGCCATGCCACACAAGTCCAGATCATCG
CGAGCACAAAGCAGGACCACAGTCTGCCGCAGCCGCCGCAGCACCCGCCGTTCTACTTT
TCCATCGACCACACTCTGCTGTACAATGCTTCCACGCCGACTTGGTCCGCTCCACATC
GGTCATCTATACCGCTTGCCGTCCAGCTTCACGAAGTGCTTGGCACCCGGATAATGAA
AATCGACCCGTAGTGTCTGGAGCCATGCAGATTGAGAAGTCGTGCCAACGCGGGCATGC

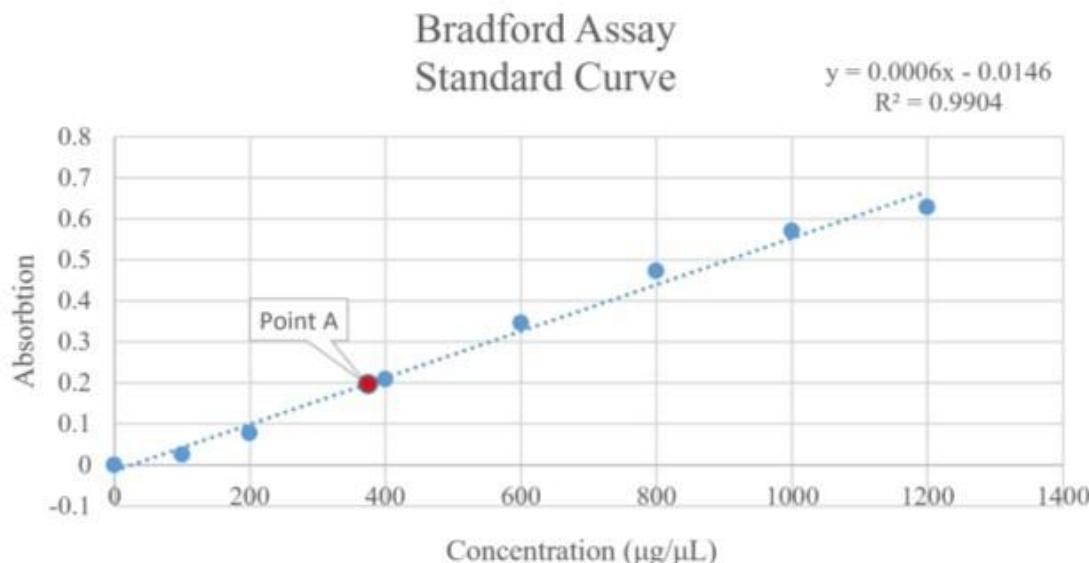
ATCCTTGCAACATACATGGTGCTGATCCAAAATTGGCCTCCCCATCTCGCGCTTGCAGCA
ATTGCGCAAATGGACCCCTCCATGCATGCCGTTCCGTGATGCGGGATATGCCAGGCCGAC
TATTCGCTCACCAATTCAAGGATGTTGTCTATGGCGTCTGGAAAGCAAAGGAAGAGAATCTT
GTGGGCTGAAGGAATTCAAGCCTGGAAGAGATGATGAAAAGTACGAGCAGGGTCGACATGGGT
GACTTCAACTGGGTTTCGCCGACTTCTCGCATTGCCCTGCCAACATCTGCCACT
CACGTTTCAGCCCCACCGAACCTTGATGCACAGCTTCCACGAACAGTGTGGCCGTG
AAACGCGATTCAAATCTTCAACGCCATTCAAGAATGTGCTACCCACTCGCGGAGCGT
GGCATGGGCTTGTGGTCCGACTCAACTCAGAGCTATACTCACCGAGCTATTCACTGCC
CTTGGGATCAAACACCTGGACATGATCTCGACGATGGAACCTGTCCGCCCTCTAAACCTG
GTCAAGAAGTTCATCAGCCTTGACACACCAGATGATCAACGAAAAGCGCAAGGGCATTGCA
GTGCACTGCAAAGCTGGTCTTGGCCGACTGGCTGTCTGATCGCGCGTACTGATCTAT
CGCTATGGCTTCACCGCAAACGAGGTGATAGCATTGATGCCGTTCATGCCGCTGGTATG
GTGGTGGTCCGCAACAAACATTGGCTTCAGTGAATCAAGGCACATTCCGAGAGTGGTGG
TTCGAGGACACCATGCGCGAGAAGCTTATGGCTTCGATGCCAGCGACTCCCACGAGA
TTGCAACAATCGAATAGCCGAAACTTGCAGCAATGGACAAACCTTCAGGCCCCAAAT
GGGGACCGCCACCAAGACCAGCCCACGCAGGGCATTAGGTGAGATCACCAACAATGAAGGA
GCTCCCGCATCGACGTACACCGATCACATGAAGAGTAGCGTTGGTGGCAGACGAGAAT
TTGCCAGCGCCCACTCCAGGTCAAGCCTCGGAAGACGAACAAGGTCTACGGTGGTCGTAGA
TGA - 3'

FIGURE 1B.

FIGURE 1C. Ramachandran Plot Outlier Values (Unfavorable ϕ/ψ (phi/psi) angles)

Template Structure	<table border="1"> <thead> <tr> <th>#</th><th>Chain</th><th>Residue</th><th>Psi</th><th>Phi</th><th>Score</th><th>Region</th></tr> </thead> <tbody> <tr> <td>1</td><td>1:[5XW5.A]</td><td>VAL 321</td><td>102.5</td><td>71.0</td><td>0.00026</td><td>outlier</td></tr> </tbody> </table>	#	Chain	Residue	Psi	Phi	Score	Region	1	1:[5XW5.A]	VAL 321	102.5	71.0	0.00026	outlier																																																																																																				
#	Chain	Residue	Psi	Phi	Score	Region																																																																																																													
1	1:[5XW5.A]	VAL 321	102.5	71.0	0.00026	outlier																																																																																																													
Homology Model Unfavorable Files	<table border="1"> <thead> <tr> <th>#</th><th>Chain</th><th>Residue</th><th>Psi</th><th>Phi</th><th>Score</th><th>Region</th></tr> </thead> <tbody> <tr> <td>1</td><td>2:[Target]</td><td>PRO 38</td><td>-140.1</td><td>-82.9</td><td>0.00023</td><td>outlier</td></tr> <tr> <td>2</td><td>2:[Target]</td><td>LYS 49</td><td>114.9</td><td>54.5</td><td>0.00006</td><td>outlier</td></tr> <tr> <td>3</td><td>2:[Target]</td><td>PRO 76</td><td>-64.6</td><td>-82.8</td><td>0.00023</td><td>outlier</td></tr> <tr> <td>4</td><td>2:[Target]</td><td>PRO 269</td><td>-74.8</td><td>-76.8</td><td>0.00000</td><td>outlier</td></tr> <tr> <td>5</td><td>2:[Target]</td><td>GLU 291</td><td>68.0</td><td>-61.2</td><td>0.00040</td><td>outlier</td></tr> <tr> <td>6</td><td>2:[Target]</td><td>VAL 381</td><td>120.2</td><td>59.2</td><td>0.00009</td><td>outlier</td></tr> <tr> <td>7</td><td>2:[Target]</td><td>PRO 415</td><td>68.2</td><td>31.0</td><td>0.00000</td><td>outlier</td></tr> </tbody> </table>	#	Chain	Residue	Psi	Phi	Score	Region	1	2:[Target]	PRO 38	-140.1	-82.9	0.00023	outlier	2	2:[Target]	LYS 49	114.9	54.5	0.00006	outlier	3	2:[Target]	PRO 76	-64.6	-82.8	0.00023	outlier	4	2:[Target]	PRO 269	-74.8	-76.8	0.00000	outlier	5	2:[Target]	GLU 291	68.0	-61.2	0.00040	outlier	6	2:[Target]	VAL 381	120.2	59.2	0.00009	outlier	7	2:[Target]	PRO 415	68.2	31.0	0.00000	outlier																																																										
#	Chain	Residue	Psi	Phi	Score	Region																																																																																																													
1	2:[Target]	PRO 38	-140.1	-82.9	0.00023	outlier																																																																																																													
2	2:[Target]	LYS 49	114.9	54.5	0.00006	outlier																																																																																																													
3	2:[Target]	PRO 76	-64.6	-82.8	0.00023	outlier																																																																																																													
4	2:[Target]	PRO 269	-74.8	-76.8	0.00000	outlier																																																																																																													
5	2:[Target]	GLU 291	68.0	-61.2	0.00040	outlier																																																																																																													
6	2:[Target]	VAL 381	120.2	59.2	0.00009	outlier																																																																																																													
7	2:[Target]	PRO 415	68.2	31.0	0.00000	outlier																																																																																																													
Homology Model Bond Angle Deviations (Z-score threshold = 4)	<table border="1"> <thead> <tr> <th>#</th><th>Chain</th><th>Residue</th><th>Angle</th><th>Z-Score</th><th>Type</th></tr> </thead> <tbody> <tr> <td>1</td><td>2:[Target]</td><td>ALA 48</td><td>120.9</td><td>4.043 <</td><td>N CA C</td></tr> <tr> <td>2</td><td>2:[Target]</td><td>LYS 49</td><td>129.7</td><td>5.220 <</td><td>pC N CA</td></tr> <tr> <td>3</td><td>2:[Target]</td><td>PRO 77</td><td>123.4</td><td>4.196 <</td><td>pCA pC N</td></tr> <tr> <td>4</td><td>2:[Target]</td><td>ASP 83</td><td>130.0</td><td>5.414 <</td><td>pC N CA</td></tr> <tr> <td>5</td><td>2:[Target]</td><td>SER 128</td><td>128.0</td><td>4.135 <</td><td>pC N CA</td></tr> <tr> <td>6</td><td>2:[Target]</td><td>ASP 173</td><td>128.6</td><td>4.519 <</td><td>pC N CA</td></tr> <tr> <td>7</td><td>2:[Target]</td><td>TYR 176</td><td>128.4</td><td>4.346 <</td><td>pC N CA</td></tr> <tr> <td>8</td><td>2:[Target]</td><td>SER 177</td><td>128.4</td><td>4.376 <</td><td>pC N CA</td></tr> <tr> <td>9</td><td>2:[Target]</td><td>GLU 289</td><td>129.2</td><td>4.900 <</td><td>pC N CA</td></tr> <tr> <td>10</td><td>2:[Target]</td><td>SER 264</td><td>128.1</td><td>4.192 <</td><td>pC N CA</td></tr> <tr> <td>11</td><td>2:[Target]</td><td>ASN 265</td><td>128.9</td><td>4.671 <</td><td>pC N CA</td></tr> <tr> <td>12</td><td>2:[Target]</td><td>GLU 291</td><td>128.9</td><td>4.702 <</td><td>pC N CA</td></tr> <tr> <td>13</td><td>2:[Target]</td><td>PHE 310</td><td>127.9</td><td>4.073 <</td><td>pC N CA</td></tr> <tr> <td>14</td><td>2:[Target]</td><td>LYS 335</td><td>128.7</td><td>4.594 <</td><td>pC N CA</td></tr> <tr> <td>15</td><td>2:[Target]</td><td>LYS 337</td><td>128.7</td><td>4.552 <</td><td>pC N CA</td></tr> <tr> <td>16</td><td>2:[Target]</td><td>VAL 381</td><td>127.9</td><td>4.063 <</td><td>pC N CA</td></tr> <tr> <td>17</td><td>2:[Target]</td><td>ARG 420</td><td>128.5</td><td>4.428 <</td><td>pC N CA</td></tr> <tr> <td>18</td><td>2:[Target]</td><td>SER 426</td><td>128.4</td><td>4.362 <</td><td>pC N CA</td></tr> </tbody> </table>	#	Chain	Residue	Angle	Z-Score	Type	1	2:[Target]	ALA 48	120.9	4.043 <	N CA C	2	2:[Target]	LYS 49	129.7	5.220 <	pC N CA	3	2:[Target]	PRO 77	123.4	4.196 <	pCA pC N	4	2:[Target]	ASP 83	130.0	5.414 <	pC N CA	5	2:[Target]	SER 128	128.0	4.135 <	pC N CA	6	2:[Target]	ASP 173	128.6	4.519 <	pC N CA	7	2:[Target]	TYR 176	128.4	4.346 <	pC N CA	8	2:[Target]	SER 177	128.4	4.376 <	pC N CA	9	2:[Target]	GLU 289	129.2	4.900 <	pC N CA	10	2:[Target]	SER 264	128.1	4.192 <	pC N CA	11	2:[Target]	ASN 265	128.9	4.671 <	pC N CA	12	2:[Target]	GLU 291	128.9	4.702 <	pC N CA	13	2:[Target]	PHE 310	127.9	4.073 <	pC N CA	14	2:[Target]	LYS 335	128.7	4.594 <	pC N CA	15	2:[Target]	LYS 337	128.7	4.552 <	pC N CA	16	2:[Target]	VAL 381	127.9	4.063 <	pC N CA	17	2:[Target]	ARG 420	128.5	4.428 <	pC N CA	18	2:[Target]	SER 426	128.4	4.362 <	pC N CA
#	Chain	Residue	Angle	Z-Score	Type																																																																																																														
1	2:[Target]	ALA 48	120.9	4.043 <	N CA C																																																																																																														
2	2:[Target]	LYS 49	129.7	5.220 <	pC N CA																																																																																																														
3	2:[Target]	PRO 77	123.4	4.196 <	pCA pC N																																																																																																														
4	2:[Target]	ASP 83	130.0	5.414 <	pC N CA																																																																																																														
5	2:[Target]	SER 128	128.0	4.135 <	pC N CA																																																																																																														
6	2:[Target]	ASP 173	128.6	4.519 <	pC N CA																																																																																																														
7	2:[Target]	TYR 176	128.4	4.346 <	pC N CA																																																																																																														
8	2:[Target]	SER 177	128.4	4.376 <	pC N CA																																																																																																														
9	2:[Target]	GLU 289	129.2	4.900 <	pC N CA																																																																																																														
10	2:[Target]	SER 264	128.1	4.192 <	pC N CA																																																																																																														
11	2:[Target]	ASN 265	128.9	4.671 <	pC N CA																																																																																																														
12	2:[Target]	GLU 291	128.9	4.702 <	pC N CA																																																																																																														
13	2:[Target]	PHE 310	127.9	4.073 <	pC N CA																																																																																																														
14	2:[Target]	LYS 335	128.7	4.594 <	pC N CA																																																																																																														
15	2:[Target]	LYS 337	128.7	4.552 <	pC N CA																																																																																																														
16	2:[Target]	VAL 381	127.9	4.063 <	pC N CA																																																																																																														
17	2:[Target]	ARG 420	128.5	4.428 <	pC N CA																																																																																																														
18	2:[Target]	SER 426	128.4	4.362 <	pC N CA																																																																																																														

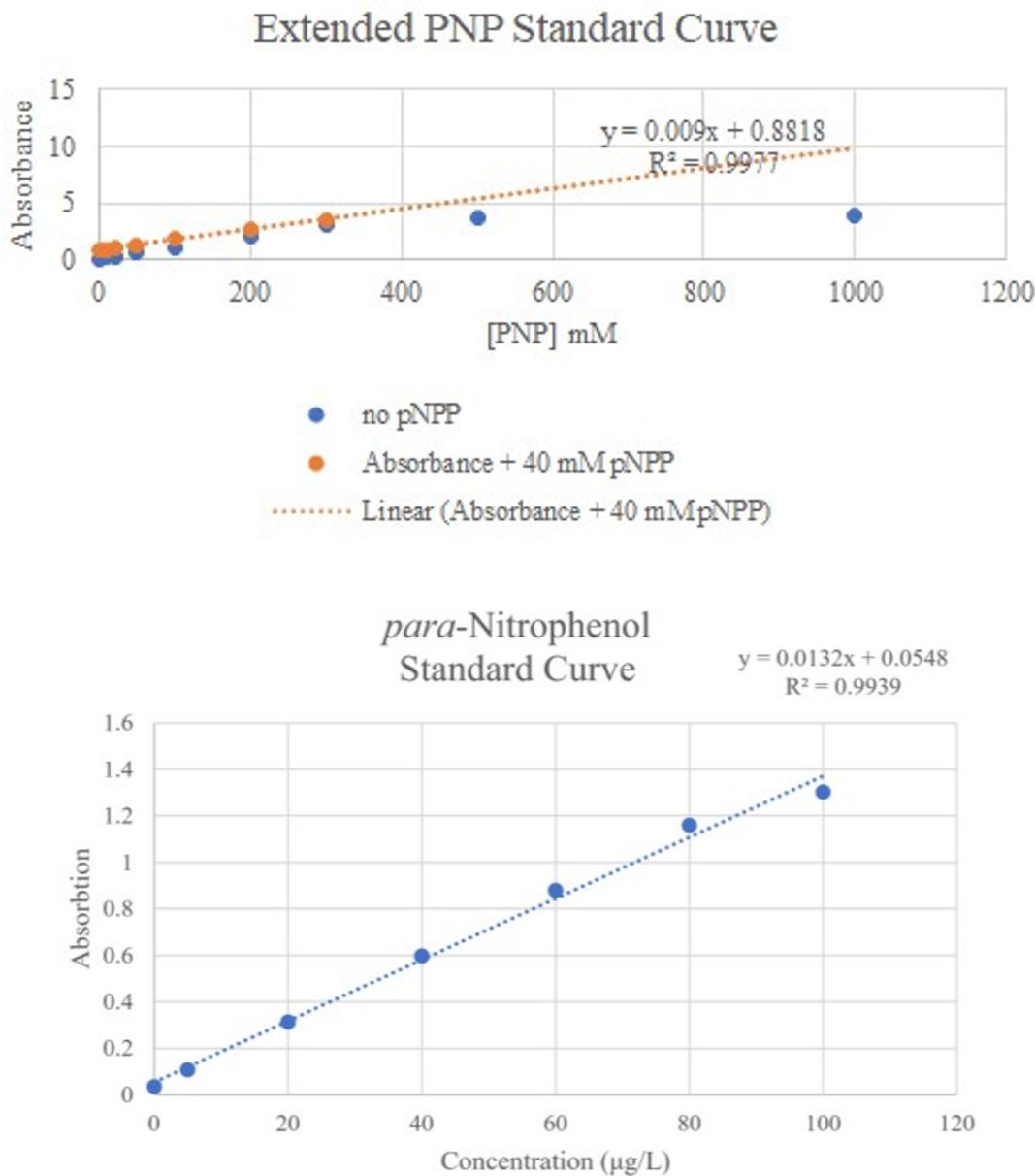
FIGURE 2A. Bradford Assay Standard Curve



Developing a Novel Inhibitor of *Cdc14s* in *C. zeina*

Above. A Bradford Assay Standard curve was generated using concentrations of Bradford Dye at 0, 100, 200, 400, 600, 800, 1000, and 1200 μ g/uL. The wavelength at which absorbances were measured at 595 nm. The Standard Curve yielded a strong regression value of 0.9904 and a broad linear range (0.635 AU). Point A delineates the 10x protein dilution point that was tested to determine the concentration of protein. The corresponding absorbance was 0.194 AU. A second point was not tested due to a limited amount of materials.

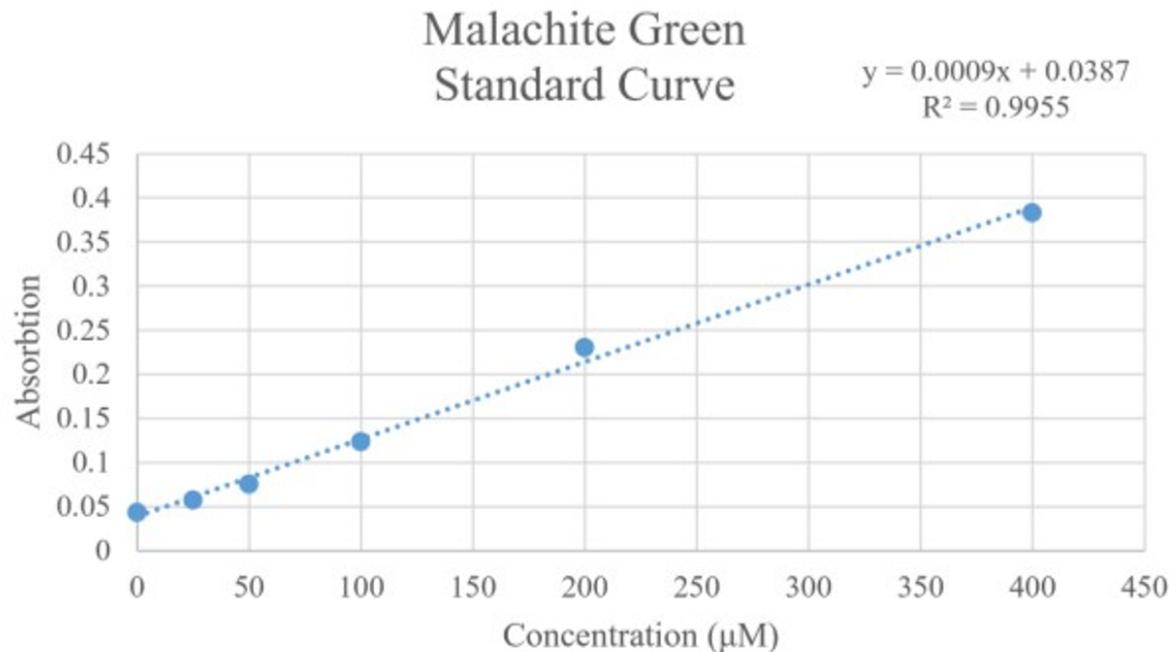
Figure 3. PNP Assay Standard Curve.



Developing a Novel Inhibitor of Cdc14s in *C. zeina*

Top. The blue points represent absorbance. Due to the possibility of observing higher absorbances in later experiments, the standard curve was extended to increase linear range. It was not used however. *Bottom.* A *para*-Nitrophenol (PNP) standard curve was generated for use in future experiments. It yielded a strong regression value.

Figure 4. Malachite Green Standard Curve with Phosphopeptide #1



Above. A standard curve was generated for the substrate specificity assay that was conducted using. Another standard curve was generated to determine an enzyme concentration within a linear range for a reaction with a phosphopeptide. Assays of Sodium Phosphate at concentrations of 160 μ M, 120 μ M, 80 μ M, 40 μ M, 20 μ M, 10 μ M, and 5 μ M in assay and 90% (v/v) Enzyme Reaction Buffer were used to generate this curve. A blank was also measured using 100% (v/v) Enzyme Reaction Buffer. Each of the assays were quenched with 200% (v/v) of Biomol Green after 10 minutes of reaction time and allowed to incubate for 20 minutes for the color to develop. The assays were read using a wavelength of 640 nm.

ACKNOWLEDGMENTS

We would like to thank all faculty at the Summer Science Program for their support in the completion of this project. It would not have been possible without their help.

REFERENCES

- [1] Gray Leaf Spot. (n.d.). Retrieved from <https://www.pioneer.com/home/site/us/agronomy/cropmanagement/corn-insect-disease/gray-leaf-spot>
- [2] Meisel, B., Korsman, J., Kloppers, F. J., & Berger, D. K. (2009, February 17). *Cercospora zeina* is the causal agent of grey leaf spot disease of maize in southern Africa.
- [3] Crous, P. W., Groenewald, J. Z., Groenewald, M., Caldwell, P., Braun, U., & Harrington, T. C. (2006). Species of *Cercospora* associated with grey leaf spot of maize.
- [4] Berger, D. K., Carstens, M., Korsman, J. N., Middleton, F., Kloppers, F. J., Tongoona, P., & Myburg, A. A. (2014, May 22). Mapping QTL conferring resistance in maize to gray leaf spot disease caused by *Cercospora zeina*.
- [5] Muller, M. F., Barnes, I., Kunene, N. T., Crampton, B. G., Bluhm, B. H., Phillips, S. M., . . . Berger, D. K. (2016, October). *Cercospora zeina* from Maize in South Africa Exhibits High Genetic Diversity and Lack of Regional Population Differentiation.
- [6] Powers, Brendan L., et al. "A Substrate Trapping Method for Identification of Direct Cdc14 Phosphatase Targets."
- [7] "Closing Mitosis: The Functions of the Cdc14 Phosphatase and Its Regulation." Annual Reviews, www.annualreviews.org/doi/full/10.1146/annurev.genet.38.072902.093051. Bembeneck, Joshua, et al.
- [8] "Wei-Qing Wang." Journal of Biological Chemistry, 16 July 2004, www.jbc.org/content/279/29/30459.short
- [9] Bremmer, S. C., Hall, H., Martinez, J. S., Eissler, C. L., Hinrichsen, T. H., Rossie, S., . . . Charbonneau, H. (2012, January 13). Cdc14 phosphatases preferentially dephosphorylate a subset of cyclin-dependent kinase (Cdk) sites containing phosphoserine.