Designing an effective inhibitor for Cdc14 in fungal pathogen *Thielaviopsis punctulata*

Kelsey Chen, Cody Chou, Aranyo Ray

Author Affiliations: Summer Science Program

Running title: Developing a novel Cdc14 phosphatase inhibitor for *Thielaviopsis punctulata*

Keywords: *Thielaviopsis punctulata*, Cdc14 phosphatase (TpCdc14), inhibitor optimization, drug target, Molecular Operating Environment, homology modelling, enzyme kinetics

Due to the Covid-19 pandemic, the Summer Science Program (SSP) was conducted online in the Summer of 2020. Thus, all data used in this report were collected by participants who participated in previous years of the SSP Biochemistry program.

Abstract

Fungal pathogens are one of the leading causes of disease in agricultural crops. Due to Cdc14's absence in plants and conservation across fungi, TpCdc14 is a promising target of inhibition to reduce the pathogenicity of *T. punctulata*. The activity for the predicted TpCdc14 protein was hypothesized to play a critical role in cellular division as a tyrosine phosphatase. Through bioinformatics analysis of a phosphate assay, we demonstrated that TpCdc14 exhibits phosphatase activity. The kinetic constants were calculated by fitting the data from a phosphatase assay to the Michaelis-Menten equation. By analyzing an assay with 24 phosphopeptides, we confirmed that like other Cdc14s, TpCdc14 preferentially dephosphorylates substrates containing the pSer-Pro-x-Lys consensus motif, consistent with the highly conserved nature of the enzyme. The top three inhibitors of TpCdc14 were identified using a homology model based on a Cdc14 ortholog. Inhibitor G7, with lowest IC50 and highest percent inhibition, was docked and optimized to improve affinity. The final inhibitor followed the Lipinski Rule of Five. These findings could be tested in vitro and lead to the development of a novel antifungal compound to combat the spread of *T. punctulata*.

Introduction

The date palm (*Phoenix dactylifera* L.) is of great nutritional and economic value in the Middle East. It represents a major source of income for many farmers in the Arabian Peninsula [1] and worldwide date production has nearly quadrupled between 1971 and 2005.[2] *Thielaviopsis punctulata*, a soil-borne fungal pathogen that infects the date palm, is thus a potential threat to food security in these regions. At-risk nations primarily include Arab countries, who possess 70% of the 120 million date palms and account for 67% of the production globally.[3]

T. punctulata is commonly found in combination with *Thielaviopsis paradoxa*, both of which belong to the *Ceratocystis* family.[4] It has been linked to the black scorch disease in Oman, Qatar, Kuwait and the UAE, which is marked by the blackening of leaves, wilting, neck bending, and ultimately plant mortality and causes up to 50% loss in newly planted offshoots in humid areas.[5]

Current methods of management include chemical control such as Score [6] (difenoconazole) and biocontrol agents such as *Streptomyces globosus* UAE1 [7] and Mycostop, whose active micro-organism is *Streptomyces griseoviridis.*[8] However, increasing resistance to existing pesticides as well as indiscriminate action calls for discovering newer targets.

The Cdc14 phosphatase family is highly conserved across eukaryotes and plays multiple roles in regulating the cell cycle, most notably by dephosphorylating active cyclin-dependent kinases (Cdks) that promote cell division and signaling the completion of mitosis.[9] Among the best-characterized members of this family is ScCdc14 phosphatase, which has the ability to dephosphorylate pSer, pThr, and pTyr residues in budding yeast, preferentially dephosphorylating Cdk sites containing pSer.[10]

Thus, the infection rate and virulence of *T. punctulata* could be reduced by targeting Cdc14 activity. Cdc14s are an attractive target for antifungal treatment for many reasons. Firstly, Cdc14 inhibition has been shown to reduce the growth of other plant pathogens. In plant-pathogenic fungus *Fusarium graminearum*, Cdc14 regulates cytokinesis and hyphae formation, and cells without FcCdc14 experience reduced conidiation, ascosporogenesis, and pathogenesis.[11] Similar studies have shown that mutant Cdc14 in *Magnaporthe oryzae* and *Aspergillus flavus* reduced pathogenicity.[12,13] Second, the Cdc14 phosphatase family is highly conserved in other eukaryotes,[14] and its active site and substrate specificity has been studied in great detail.[9] Thus, inhibitors could be designed for many different fungal pathogens, including *Thielaviopsis punctulata* (TpCdc14). Third, Cdc14 has been shown to be highly specific for its substrates [15]; therefore, competitive inhibitors would be highly selective for Cdc14. Lastly, Cdc14 phosphatases are not present in plants [16] and are crucial to the infection cycle of many plant-pathogenic fungi. Therefore, inhibiting Cdc14s would not be harmful to the host plant.

In this report, we determined whether TpCdc14 would be a practical and effective pesticide target and subsequently designed candidate pesticide compounds to inhibit TpCdc14 to combat crop infection.

Experimental Procedures

Activity Hypothesis

Protein Sequence and Function Prediction

Using GenScan and the genome sequence from *T. punctulata*, the coding sequence and introns of our target protein were predicted. Similar ortholog sequences of TpCdc14 were obtained through a protein BLAST search. Then, the protein sequence was compared in a BLAST against *Saccharomyces cerevisiae* to analyze the potential functions. Other Cdc14 sequences were found in *Homo sapiens*, fruit flies (*Drosophila melanogaster*), worms (*Caenorhabditis elegans*), budding yeast (*Saccharomyces cerevisiae*) and fission yeast (*Schizosaccharomyces pombe*) with a BLAST search. By using Jalview, the alignments of the ortholog sequences were analyzed to determine conserved portions. The molecular weight and isoelectric point of the protein was found through the Compute pI/MW tool. GOR4 and NetSurfP-2.0 tools were used to predict the secondary structures of TpCdc14 (Table 1). All of the tools used in this procedure were from the ExPASy Bioinformatics Resource Portal.

Protein Expression and Purification

Constructing an Expression Vector for TpCdc14

TpCdc14 DNA was added to the pET15b plasmid sequence in SnapGene Viewer. Two restriction enzyme sites - BamHI and XhoI - in the pET15b multiple cloning sequence (MCS) were chosen for molecular cloning. The restriction enzymes were verified to be non-cutter for *Thielaviopsis punctulata* virulence locus X.

Designing Primers

Forward and reverse primers were designed for PCR cloning and checked for GC content and melting temperature. The coding sequence was amplified using these primers. The reverse complement of the sequence was generated and placed into the vector so that the 6xHis tag is incorporated at the N-terminus. To confirm the correct placement, the plasmid sequence was translated using GenScan and the protein of interest was found.

Protein Expression

The expression plasmid was transformed into competent *Escherichia coli* cells and incubated overnight at 37°C, and bacteria with the plasmid were selected by ampicillin. Large cultures were amplified at 37°C for 8 hours. L-arabinose was then added to trigger the production of the target protein from the plasmid and the cultures were incubated overnight. After centrifugation at 3,500 rpm for 20 minutes, the cell pellet at the bottom of the tube was used for purification.

Protein Purification and Evaluation

Recombinant protein harvested from *E. coli* colonies was purified using nickel affinity chromatography. SDS-PAGE was used to evaluate the purification of the protein, and the gel was set in Coomassie blue staining solution. The gel image obtained was analyzed using Image Lab (Bio-Rad).

TpCdc14 Kinetics

Specific Activity Calculations

To test if TpCdc14 exhibited tyrosine phosphatase activity, phosphate assays were conducted with pNPP (p-nitrophenyl phosphate). 0.33 μM TpCdc14 was mixed with 30mM pNPP in a buffer, and the absorbance of dephosphorylated pNPP (p-nitrophenyl, pNP) was measured at 405 nm each minute for 30 minutes. 3 trials were conducted in the absence of tungstate, a known inhibitor of tyrosine phosphatases. 2mM tungstate was then added to the same concentrations of TpCdc14 and pNPP, and the absorbance of pNP was measured at 405 nM for 3 reactions. A standard pNP curve was used to determine the amount of pNP produced in the six reactions, and a t-test was used to verify the statistical significance between the mean specific activities of TpCdc14 with and without the tungstate inhibitor.

Steady-State Assay

Then, a phosphate assay was conducted to determine steady-state parameters for future reactions. 1 μM TpCdc14 was added to substrate (pNPP) concentrations of 0, 1, 2, 5, 10, 20, 40, 60, and 80 mM, and endpoint absorbances of pNP were measured after 15 minutes for 3 trials. The velocity of each reaction was calculated, and the average velocity was used to determine the K_{cat}, K_m, and V_{max} of TpCdc14 using non-linear curve fitting to the Michaelis-Menten equation: $V_0 = V_{max}[S]/(K_m + [S])$.

Catalytic Specificity of TpCdc14

A steady-state endpoint assay measuring TpCdc14's ability to dephosphorylate 24 phosphopeptides was used to determine its substrate specificity. 0.75 μM enzyme was added to 0.1 mM of each substrate and 40 μL reaction buffers, and the reaction was incubated at 30°C for 10 minutes. Malachite green was then added to each well to stop reaction. After 20 minutes of incubation, the absorbance was measured, and the amount of phosphate in each well was calculated and average K_{cat}/K_m for each peptide was plotted on a bar graph.

TpCdc14 Inhibition

Homology Modelling

The Molecular Operating Environment (MOE) software was used to create and evaluate a homology model of TpCdc14. PDB_5XW5 (*S. cerevisiae*) was used as a template as it had the largest query range, the lowest E-value, and matched the query across 2 domains. A homology model was generated using Amber 14:EHT. Each model was evaluated by identification of regions of poor structural similarity based on RMSD, a Ramachandran plot, and identification of unfavorable Phi-Psi angles, bond lengths, bond angles, and rotamer energies.

In vitro Inhibition

Next, 16 candidate inhibitors (Table 2) were screened *in vitro* to determine the optimal inhibitors of TpCdc14. First, the assay was conducted in absence of enzyme with 20 mM pNPP (negative control) and without the inhibitor with 20 mM pNPP and 1.44 μM TpCdc14 (positive control) to determine the Z-factor, using the equation:

$$\text{Z-factor} = 1 - \frac{3(\sigma_p + \sigma_n)}{|\mu_p - \mu_n|}$$

Next, the reactions were conducted in triplicate with 20 mM pNPP, 1.44 μM TpCdc14, and 50 μM of each inhibitor, except for I9 and H7, which were used at 1 mM. A t-test was used to determine the statistical significance of each inhibitor, and a bar graph of the percent inhibition was plotted to determine the three most effective inhibitors. To compare the potency of the top 3 inhibitors, G7, I1,

and I2, a dose response assay that tested a variety of inhibitor concentrations from 0.625 μM to 500 μM was analyzed to calculate the half-maximal inhibitory concentration (IC50).

In silico Inhibition

Ligand docking into the homology model was then conducted with the same 16 candidate inhibitor compounds in 20 poses using the rigid receptor ligand docking function in MOE. The S-values (scores that represent binding affinity) were then calculated by MOE for each inhibitor (Figure 10). Ligand interaction maps and S-values generated by MOE were considered to select the best inhibitors.

In silico Inhibitor Optimization

After G7 was selected for optimization, it was docked into the TpCdc14 homology model using induced fit to confirm the consistency of the model. The positions consistent with the rigid receptor model were chosen for optimization. Then, G7 was optimized in MOE by considering the similarity between G7 and the budding yeast ligand. Electrostatic interactions between TpCdc14 and the inhibitor were also used to modify the structure of G7 to minimize the S value.

Results

Activity Hypothesis

Protein Sequence and Function Prediction

The predicted protein sequence had 3 introns located. These 3 introns were spliced before inserting the protein into the vector plasmid. The sequence was then BLAST searched against the budding yeast *Saccharomyces cerevisiae*, a highly studied organism where an ortholog would most likely be characterized. In each organism, the ortholog belonged to the Cdc14 family (Table 3). Thus, we hypothesized that its biochemical function was to regulate cell division by dephosphorylating cyclin-dependent kinase (Cdk) targets and other mitotic kinases. To determine whether this sequence is conserved in other organisms, we found sequences of orthologs in humans, fruit flies, budding yeasts, worms, and fission yeasts through BLAST. Using Jalview and Clustal Omega, the alignments show the same conserved blocks across multiple organisms in Table 4.

Protein Expression and Purification

Protein Purification and Evaluation

After inserting the target gene into the pET-15b plasmid and inducing expression with L-arabinose, we analyzed the SDS-PAGE gel to determine if the target protein was produced (Figure 1). Using ImageLab, we found the molecular weight of the purified protein to be 52.7 kDa (Figure 2), which was close to the molecular weight of our predicted protein (53.507 kDa) generated by ExPASy. The pure protein pool lanes had a single thick band with lane percentage of approximately 50%. Thus, it confirmed that the purified protein was TpCdc14. Using ExPASy, we found the theoretical pI for TpCdc14 to be 8.63.

TpCdc14 Kinetics

Specific Activity Calculations

We showed that TpCdc14 exhibited phosphatase activity by analyzing data from a phosphatase assay with sodium tungstate, a known inhibitor of tyrosine phosphatases. The results demonstrated that the average product formation and the specific activity of TpCdc14 decreased with the addition of the 2mM sodium tungstate inhibitor (Figure 3, 4). The mean rate without tungstate was 7578 μM/min and with tungstate was 1816 μM/min which is a 73.8% inhibition. T-test analysis showed a statistically significant difference between the mean rates with and without tungstate, thereby confirming TpCdc14 phosphatase activity.

Steady-State Assay

To determine the K_{cat}, K_m, and V_{max} of TpCdc14, we then analyzed data from a phosphatase assay using non-linear curve fitting. The data was fit to the Michaelis-Menten equation (with no inhibition term), and the resulting equation showed that the K_{cat}, K_m, and V_{max} values were 0.46 sec^{-1}, 51.46 mM, and 27.62 μM/min, respectively (Figure 5).

Catalytic Specificity of TpCdc14

To determine the favorable substrate sites, an assay was conducted with 24 phosphopeptides (Figure 6). The average K_{cat}/K_m of the reaction with peptides with the pSer-Pro-x-Lys consensus motif was 19.27 $\mu M^{-1}min^{-1}$, while the average K_{cat}/K_m for peptides lacking the motif was 0.98 $\mu M^{-1}min^{-1}$. This showed that TpCdc14 preferentially dephosphorylated substrates containing the pSer-Pro-x-Lys

consensus motif. Furthermore, peptides #7 and #24, the two peptides that yielded the greatest velocity, contained basic residues at the +4 position relative to the phosphoserine, and the best substrate (peptide #24) contained another basic residue at +5, suggesting that basic residues downstream of pSer increase the catalytic efficiency of TpCdc14. The findings were consistent with those of previous studies,[17] which showed that ScCdc14 was highly selective for sites containing pSer-Pro-x-Lys. Our inhibitor was then designed to mimic the structure of the optimal substrates.

TpCdc14 Inhibition

Homology Modeling

To construct a model of the protein produced by TpCdc14, we used MOE to fit the protein to the budding yeast 5XW5 template. The structural and sequence alignment exhibited high similarity (RMSD across all 361 residues = 0.967), consistent with the highly conserved nature of Cdc14 (Figure 7. A). Since there were no unfavorable bond lengths and few unfavorable Phi-Psi angles, bond angles, and rotamer energies (Figure 7. B), we concluded that our homology model was accurate and could be used in further steps.

In vitro Inhibition

To find inhibitors for TpCdc14, a pNPP assay was used to screen 16 candidate inhibitors and the percent inhibition over 10 minutes was calculated. The Z-factor of our assay was 0.9273, indicating that it was an excellent assay. The top three inhibitors were G7, I1, and I2, which yielded 89, 76, and 75 percent inhibition respectively at 50 μM (Figure 8, 9.A, B, C). The IC50 values for G7, I1, and I2 were determined to be 21.85, 25.69, and 114.06 μM, respectively (Figure 9.D). Because G7 and I1 required a significantly lower concentration to reduce enzyme activity, we concluded that G7 and I1 were the most effective inhibitors.

In silico Inhibition

We conducted rigid receptor docking with the same 16 candidate inhibitors and the S-values (a representation of docking affinity) were calculated with MOE (Figure 10). While I1 exhibited the best docking affinity for the substrate, it did not interact with Cys330. However, G7 mimicked the binding patterns of the model peptide (Figure 11) by interacting with Cys330 and Arg336. Furthermore, after running an induced fit dock of G7, we found that the majority of the output positions were similar to

that of the rigid receptor fit. Thus, although initial docking outputs did not show that it had the greatest binding affinity for the site, we selected G7 to optimize.

In silico Inhibitor Optimization

To optimize the inhibitor, we altered the structure of G7 and minimized the S-value on MOE. The initial inhibitor position and structure exhibited an S-value of -5.63 kcal/mol. To mimic the functional groups of the ScCdc14 ligand, we added two alcohol groups to hydrogen acceptor regions. Furthermore, after considering the electrostatic map, we added an amine group in a hydrogen donor region, and an ethyl group to mimic substrate. Lastly, a cyanide group was added. This optimization followed the Lipinski Rule of Five and decreased the S-value to -7.74 kcal/mol (Figure 13). We hypothesize that these optimizations will work *in vitro*.

Discussion

Our initial hypothesis that the protein is a phosphatase of the Cdc14 family was supported by the BLAST analysis, steady state enzyme kinetics, specificity and phosphopeptide assays. Inhibitor screening helped narrow down the potential inhibitors which produced effective inhibition. These inhibitors were then used in a dose-response assay to pick the one with the lowest IC50 of 21.85 μM. Starting with the best molecule, we hypothesized that its inhibition could be increased by mimicking the interactions of the natural substrate of ScCdc14 (on which we designed our homology model of TpCdc14). Thus, we designed further modification in silico to optimize binding affinity with the aim of generating a more potent inhibitor of the enzyme. With reference to electrostatic map near the ligand as well as existing substrate recognition features, we added substituents to increase binding affinity towards our inhibitor. Thus, by computationally modelling the structure of TpCdc14 and determining its kinetic constants, specificity and activity, this study could successfully demonstrate that the activity of this enzyme can be blocked by a competitive inhibitor. The high docking affinity of the optimized structure of G7 (Figure 12) suggests that our design would be an effective inhibitor of TpCdc14, thereby limiting its pathogenicity. Although further research is required to determine whether the optimization works *in vitro*, G7 is a promising fungal inhibitor for *T. punctulata*.

FIGURES

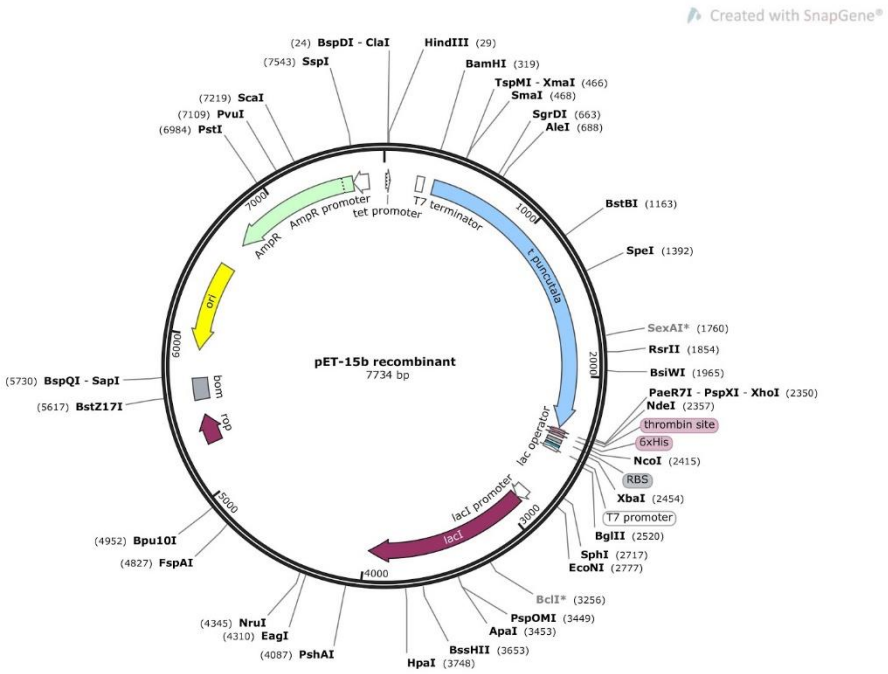


Figure 1: Recombinant pET-15b plasmid with TpCdc14 target protein. The TpCdc14 target protein, shown in blue, was inserted between the BamHI and XhoI restriction sites in the pET-15b plasmid. The forward and reverse primers used were 5'-ATCGGGATCCATGTCGCGCAC-3' and 5'-CGATCTCGAGTTACCGTAAGCGC-3', respectively.

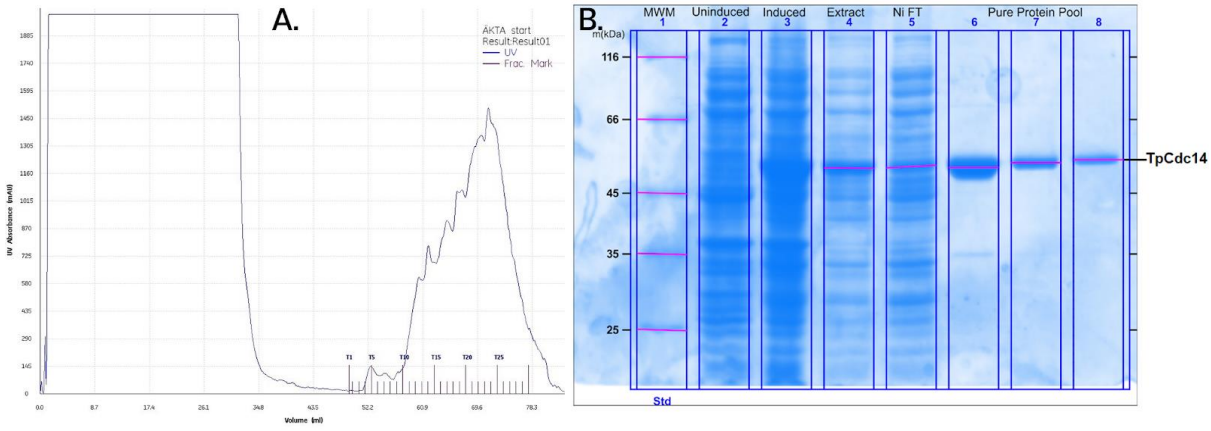


Figure 2: A. The UV/Vis absorbance curve from the chromatogram. The intense elution peak indicates that the sample that passed through the column contained a large quantity of purified protein. **B. SDS-polyacrylamide gel electrophoresis (SDS-PAGE) patterns of purified proteins.** Purified proteins were stained with Coomassie blue. ImageLab analysis tools indicated that the thick band was 53.5 kDa. It had a percent error of 1.5% compared to the theoretical value of 52.7 kDa.

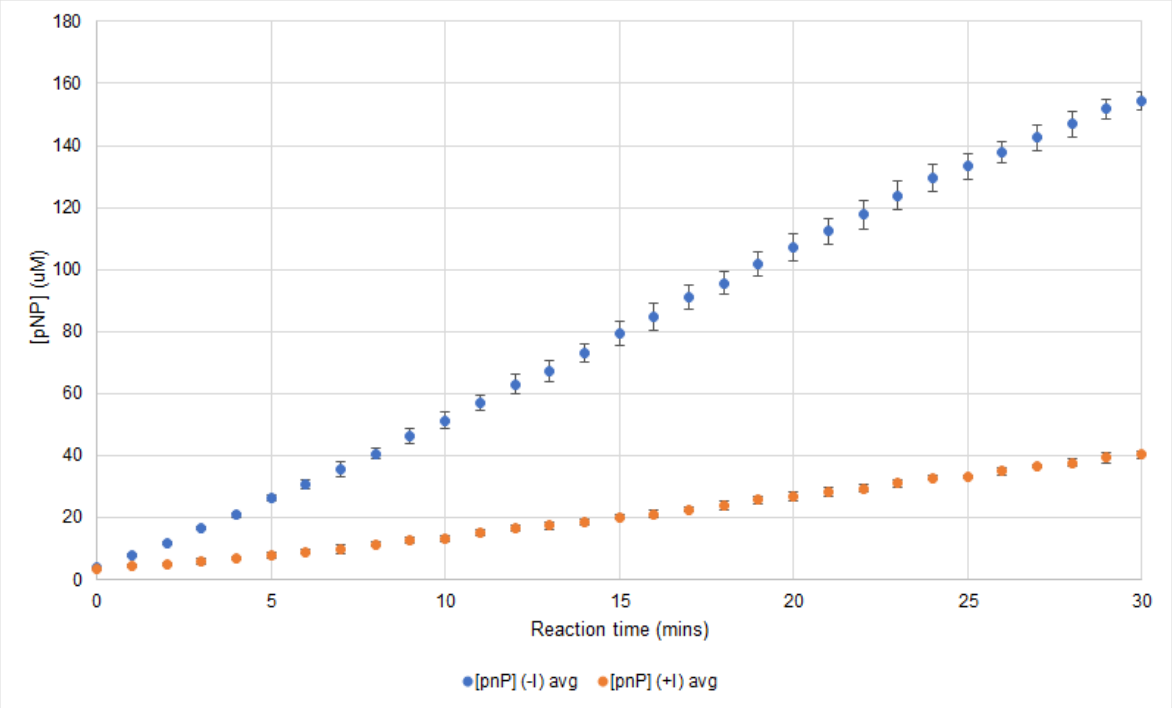


Figure 3: TpCdc14 phosphatase activity. Sodium tungstate, a known inhibitor of tyrosine phosphatases, was added to reactions containing pNPP and TpCdc14. -I and +I indicate the absence and presence of sodium tungstate, respectively. Product (pNP) formation was measured, and it was found that reaction rate decreased with the addition of sodium tungstate. Data is presented as mean±s.d., n=3.

Average specific activities		St. dev	Average Rates of formation (uM/min)	Average % inhibition
µM pNP per min per µM Cdc14 (-I)	15.5942	0.3105	5.1461	0
µM pNP per min per µM Cdc14 (+I)	4.0738	0.1181	1.3443	73.8757

Figure 4: Specific activity of TpCdc14 decreased with the addition of the inhibitor. The average specific activity over three trials was calculated and showed that sodium tungstate reduced the activity of TpCdc14 by a factor of 3.

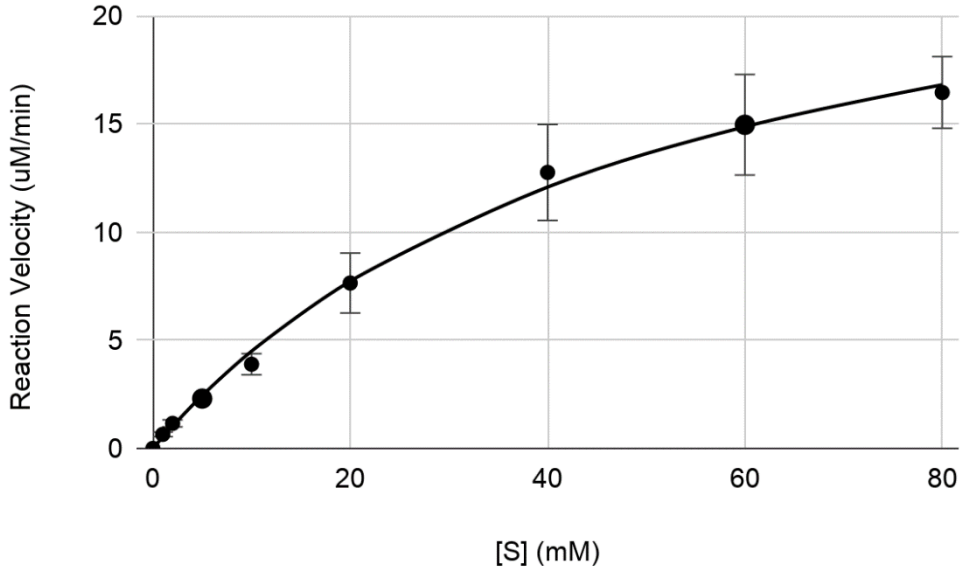


Figure 5: Michaelis-Menten Curve measuring the velocity of TpCdc14 with the Substrate pNPP. The average velocity of µM pNP produced per minute was measured for substrate concentrations of 0, 1, 2, 5, 10, 20, 40, 60, and 80 mM. The data was fitted to the Michaelis-Menten equation using non-linear regression to determine K_{cat} (0.46 sec^{-1}), K_m (51.46 mM), and V_{max} (27.62 µM/min). The R-squared value was 0.99. Data is presented as mean±s.d., n=3.

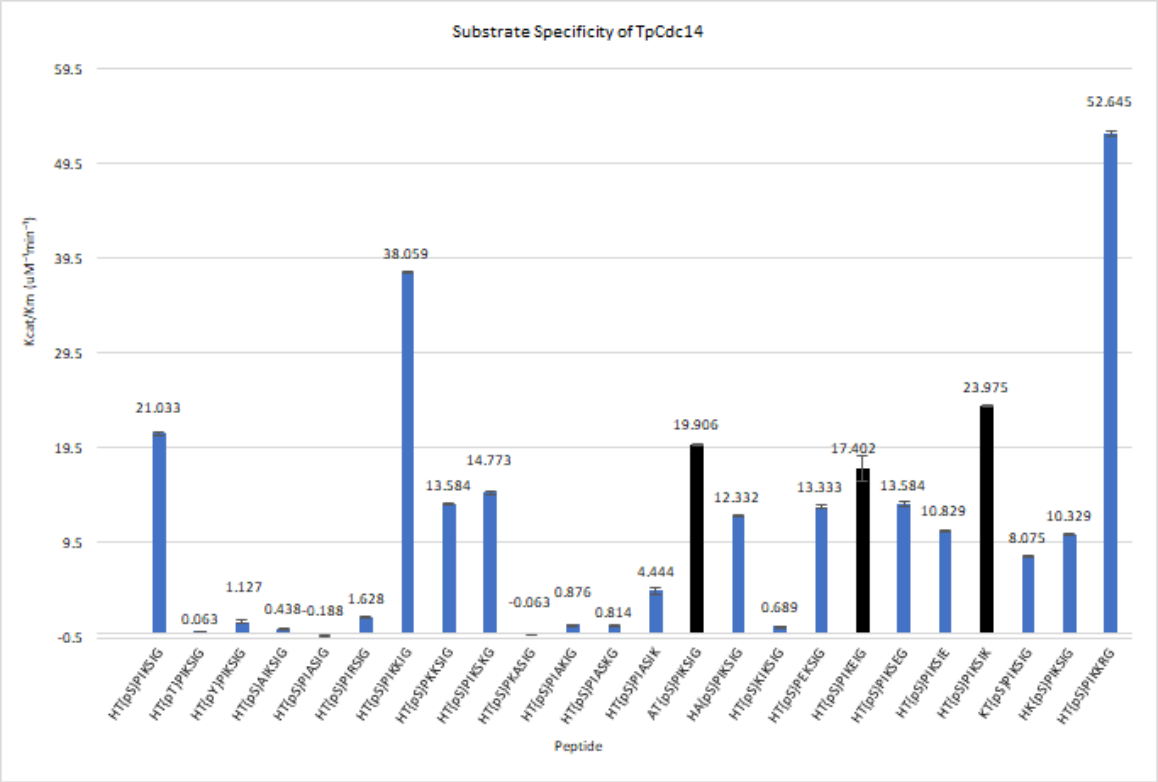


Figure 6: Substrate specificity of TpCdc14. The average velocity over 3 trials was calculated for each phosphopeptide and these values were used to determine the catalytic specificity of TpCdc14. Data is presented as mean±s.d. A two-tailed, unpaired t-test was conducted to compare velocities of each substrate against peptide 1. Blue bars represent significant data with p-values < 0.05.

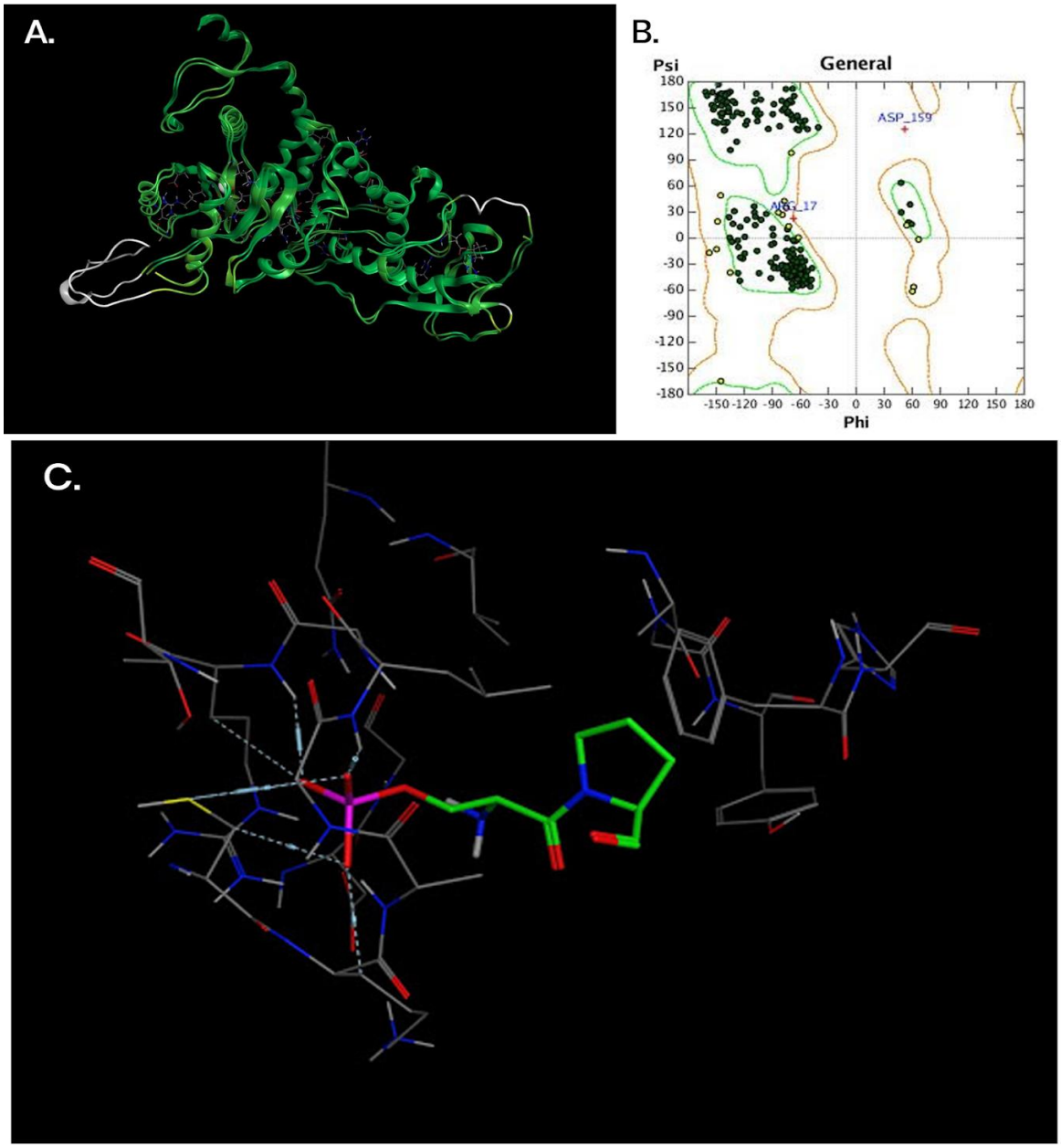


Figure 7: A. Homology model of TpCdc14 with the budding yeast 5XW5 template structure and the corresponding Ramachandran plot. A: The green color represents a low RMSD value. **B:** The Ramachandran plot showed few unfavorable Phi-Psi angles (represented by the red x's), thereby confirming the validity of our homology model. **C: Active site of the TpCdc14 homology model bound to the model peptide.** The 5XW5 model ligand SWI6 (shown in green) was fit into the active site of the TpCdc14 homology model, and ligand-receptor interactions were analyzed to assess the validity of the homology model. Since the ligand formed hydrogen bonds with the Cys330 and Arg336 residues on the active site, the model was determined to be valid.

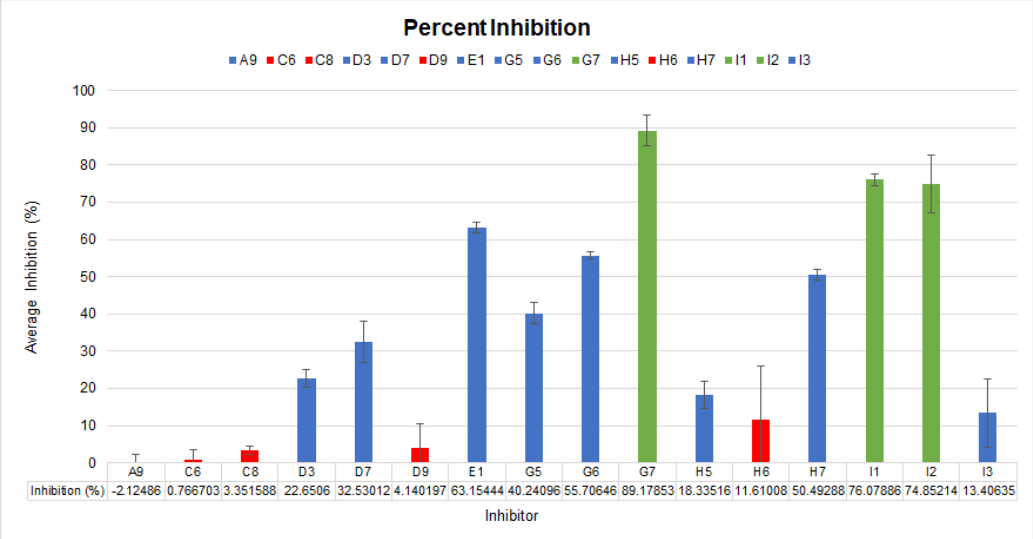
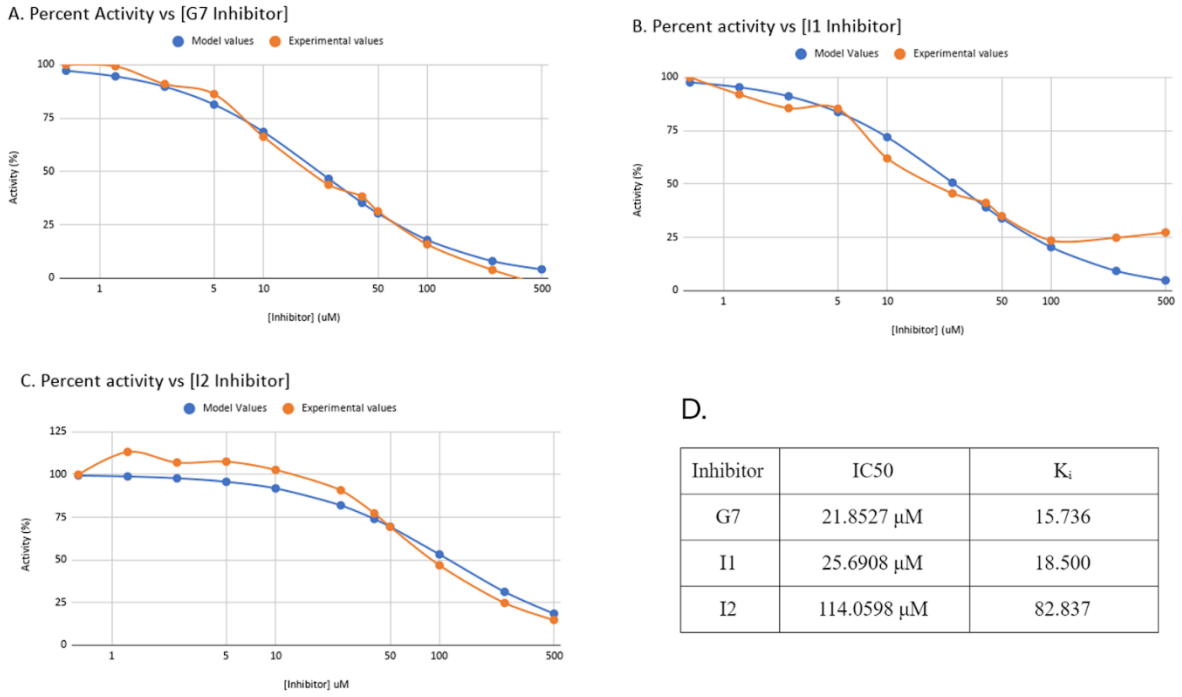


Figure 8: Percent inhibition for 16 candidate inhibitors. A phosphopeptide specific activity assay was performed to determine if our protein has the same specificity as ScCdc14. A t-test was performed to classify inhibitors as poor (p-value > 0.05), acceptable (p-value < 0.05, inhibition < 65%) and good (p-value < 0.05 and inhibition% > 65). Red bars represent poor inhibitors, blue bars represent acceptable inhibitors, and green bars represent the top 3 inhibitors. Data is presented as mean±s.d. Inhibitor structures are depicted in Table 4.



D.

Inhibitor	IC50	K_i
G7	21.8527 µM	15.736
I1	25.6908 µM	18.500
I2	114.0598 µM	82.837

Figure 9: Dose response curves for A. G7 inhibitor, B. I1 inhibitor, C. I2 inhibitor. D: IC50 calculations from dose response data. The scaled absorbance readings were used to calculate activity % at different inhibitor concentrations for our top 3 candidate inhibitors G7, I1 and I2. Model values were fitted against these activity % values using non-linear curve fitting to determine IC50.

Inhibitor	S	RMSD
I1	-6.3405	2.8749
C6	-5.9440	2.4891
A9	-5.8960	0.8107
C8	-5.8567	1.1394
D7	-5.8420	2.6412
E1	-5.7150	1.6724
G5	-5.6761	1.8424
G7	-5.6519	1.1360
I3	-5.5791	2.3499
G6	-5.4294	0.6123
D3	-5.3534	1.7528
H7	-5.0311	1.5861
D9	-5.0192	2.8490
H6	-4.8077	1.9418
H5	-4.5974	1.0589

Figure 10: Lowest S-values for 16 candidate inhibitors in MOE. The lowest S-value for each inhibitor computed by MOE was recorded to determine the top inhibitors.

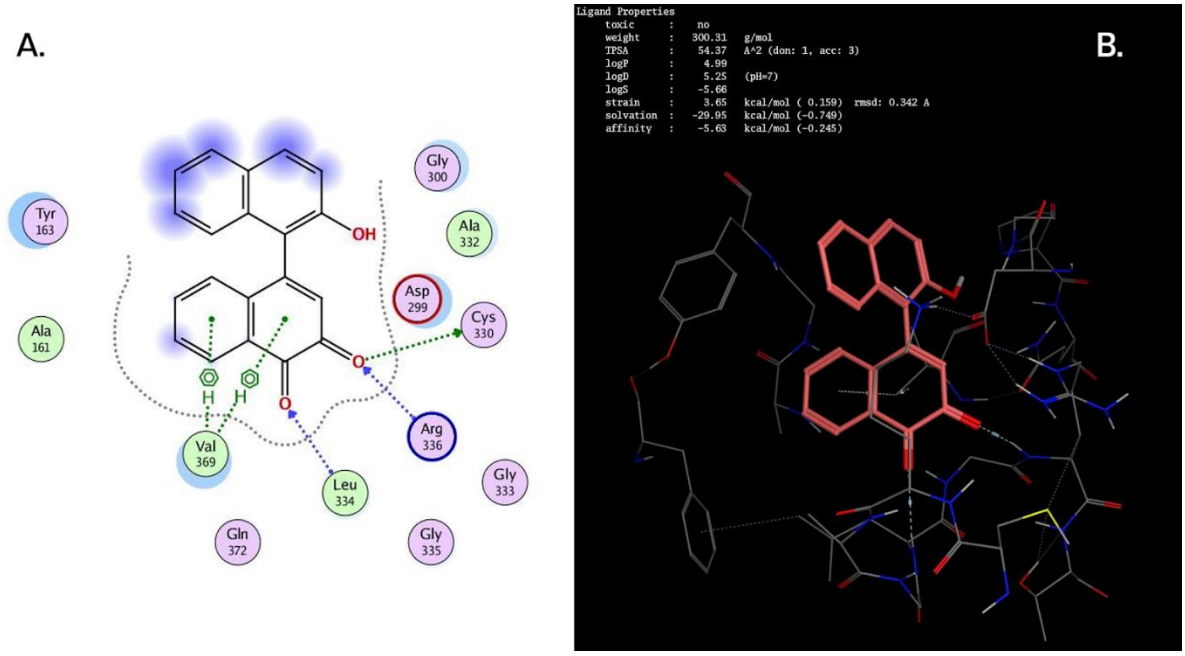


Figure 11: Initial inhibitor structure of G7 in the active site of the TpCdc14 homology model. A: The initial ligand interaction map showed that G7 interacted with Cys330 and Arg336, mimicking the mechanisms of the ligand of budding yeast. **B:** The binding site shows the inhibitor G7 colored in coral, which was found to have an initial docking score of -5.63 kcal/mol and logP of 4.99.

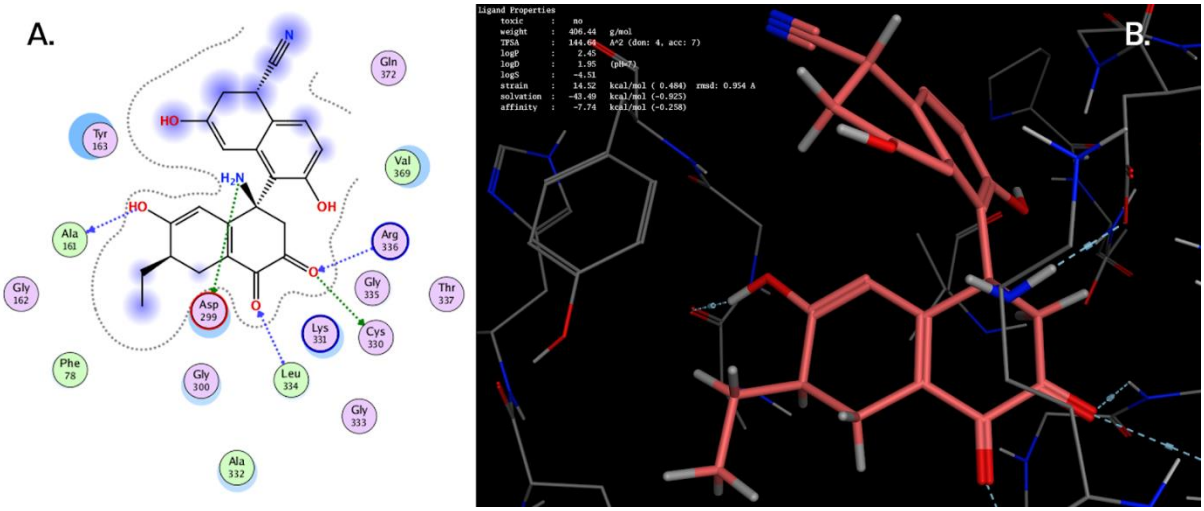


Figure 12: Final optimized inhibitor structure of G7 in the active site of the TpCdc14 homology model. A: The ligand interaction map shows G7 interacting with Cys330 and Arg336 residues on the homology model, similar to the Swi6 ligand. Furthermore, there are fewer than 5 hydrogen bond donors and 10 hydrogen bond acceptors, thereby following the Lipinski rule of Five. **B:** The binding site shows the optimized inhibitor colored in coral. Our optimization had an S-value of -7.74 kcal/mol, a decrease of 37% from the initial inhibitor.

References

1. Saeed, E. E., Sham, A., El-Tarabily, K., Abu Elsamen, F., Iratni, R., and AbuQamar, S. F. (2016) Chemical Control of Black Scorch Disease on Date Palm Caused by the Fungal Pathogen Thielaviopsis punctulata in United Arab Emirates. Plant Disease. 100, 2370–2376

2. Chao, C. T., & Krueger, R. R. (2007). The Date Palm (Phoenix dactylifera L.): Overview of Biology, Uses, and Cultivation, HortScience horts, 42(5), 1077-1082. https://journals.ashs.org/hortsci/view/journals/hortsci/42/5/article-p1077.xml (Accessed July 22, 2020)

3. El-Juhany, L. I. (2010) Degradation of date palm trees and date production in Arab countries: causes and potential rehabilitation. [online] /paper/Degradation-of-date-palm-trees-and-date-production-El-Juhany/37af83da0317113d82d31999ef311e18412936c1 (Accessed July 22, 2020)

4. Wingfield, M. J., De Beer, Z. W., Slippers, B., Wingfield, B. D., Groenewald, J. Z., Lombard, L., and Crous, P. W. (2012) One fungus, one name promotes progressive plant pathology. Mol. Plant Pathol. 13, 604–613

5. Abdelmonem, A. M., and Rasmy, M. R. (2007) Major diseases of date palm and their control. Commun. Inst. For. Bohem. 23, 9–23

6. Saeed, E. E., Sham, A., El-Tarabily, K., Abu Elsamen, F., Iratni, R., and AbuQamar, S. F. (2016) Chemical Control of Black Scorch Disease on Date Palm Caused by the Fungal Pathogen Thielaviopsis punctulata in United Arab Emirates. Plant Disease. 100, 2370–2376

7. Saeed EE, Sham A, Salmin Z, Abdelmowla Y, Iratni R, El-Tarabily K and AbuQamar S (2017) Streptomyces globosus UAE1, a Potential Effective Biocontrol Agent for Black Scorch Disease in Date Palm Plantations. Front. Microbiol. 8:1455. doi: 10.3389/fmicb.2017.01455

8. Suleman, P., AL-Musallam, A., and Menezes, C.A. The effect of biofungicide Mycostop on Ceratocystis radicicola , the causal agent of black scorch on date palm | SpringerLink [online] https://link.springer.com/article/10.1023/A:1014519726573 (Accessed July 22, 2020)

9. Bremmer, S. C., Hall, H., Martinez, J. S., Eissler, C. L., Hinrichsen, T. H., Rossie, S., Parker, L. L., Hall, M. C., and Charbonneau, H. (2012) Cdc14 phosphatases preferentially dephosphorylate a subset of cyclin-dependent kinase (Cdk) sites containing phosphoserine. J. Biol. Chem. 287, 1662–1669

10. Bremmer, S. C., Hall, H., Martinez, J. S., Eissler, C. L., Hinrichsen, T. H., Rossie, S., Parker, L. L., Hall, M. C., and Charbonneau, H. (2012) Cdc14 phosphatases preferentially dephosphorylate a subset of cyclin-dependent kinase (Cdk) sites containing phosphoserine. J. Biol. Chem. 287, 1662–1669

11. Li, C., Melesse, M., Zhang, S., Hao, C. F., Wang, C., Zhang, H., Hall, M. C. and Xu, J.-R. (2015). FgCDC14 regulates cytokinesis, morphogenesis, and pathogenesis in Fusarium graminearum. Mol. Microbiol. 98, 770-786. doi:10.1111/mmi.13157

12. Li, C., Cao, S., Zhang, C., Zhang, Y., Zhang, Q., Xu, J.-R., and Wang, C. (2018) MoCDC14 is important for septation during conidiation and appressorium formation in Magnaporthe oryzae. Molecular Plant Pathology. 19, 328–340

13. Yang, G., Hu, Y., Fasoyin, O. E., Yue, Y., Chen, L., Qiu, Y., Wang, X., Zhuang, Z., and Wang, S. (2018) The Aspergillus flavus Phosphatase CDC14 Regulates Development, Aflatoxin Biosynthesis and Pathogenicity. Front. Cell. Infect. Microbiol. 10.3389/fcimb.2018.00141

14. Zhao, W., Yang, X., Dong, S., Sheng, Y., Wang, Y., and Zheng, X. (2011) Transient silencing mediated by in vitro synthesized double-stranded RNA indicates that PsCdc14 is required for sporangial development in a soybean root rot pathogen. Sci China Life Sci. 54, 1143–1150

15. Bremmer, S. C., Hall, H., Martinez, J. S., Eissler, C. L., Hinrichsen, T. H., Rossie, S., Parker, L. L., Hall, M. C., and Charbonneau, H. (2012) Cdc14 Phosphatases Preferentially Dephosphorylate a Subset of Cyclin-dependent kinase (Cdk) Sites Containing Phosphoserine. J. Biol. Chem. 287, 1662–1669

16. Kerk, D., Templeton, G., and Moorhead, G. B. G. (2008) Evolutionary Radiation Pattern of Novel Protein Phosphatases Revealed by Analysis of Protein Data from the Completely Sequenced Genomes of Humans, Green Algae, and Higher Plants. Plant Physiology. 146, 351–367

17. Bremmer, S. C., Hall, H., Martinez, J. S., Eissler, C. L., Hinrichsen, T. H., Rossie, S., Parker, L. L., Hall, M. C., and Charbonneau, H. (2012) Cdc14 phosphatases preferentially dephosphorylate a subset of cyclin-dependent kinase (Cdk) sites containing phosphoserine. J. Biol. Chem. 287, 1662–1669

APPENDIX

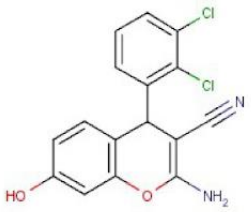
Table 1. Using GOR4 to predict secondary structure:

Alpha helix (Hh) : 150 is 22.26%
310 helix (Gg) : 0 is 0.00%
Pi helix (Ii) : 0 is 0.00%
Beta bridge (Bb) : 0 is 0.00%
Extended strand (Ee) : 131 is 19.44%
Beta turn (Tt) : 0 is 0.00%
Bend region (Ss) : 0 is 0.00%
Random coil (Cc) : 393 is 58.31%
Ambiguous states (?) : 0 is 0.00%
Other states : 0 is 0.00%

```
        10        20        30        40        50        60        70
         |         |         |         |         |         |         |
MSRTARTASVGQVVEYIPDRLYLAAYNKQPDALTFFPYGEPSSSLSISRSPARRRPGESQKKSPCYFTVD
ccccccccccceeeeeccchhhhhhhccccceeeeccccccccceecccccccccccccccccceeeecc
DSLLYNAFHHDFGPLHIGHLYRFALYFHDILAAKENKDRPIVFWSKPDAKSRANAACMLACYMVLIQNWP
cccchhhhhccccccccchhhhhhhhhhhhhhhhhhcccceeeeccccccchhhhhhhhhheeeeeecccc
PHLALAPIAQVDPPLMPFRDAGYSQADYGLTVQDVVYGVWKAKEEKCCDLENFDLDQYEKFERVEHGDFN
ccccccccccccccccccccccccccccccceeeeeeeeeeecccccccccccccccchhhhhheeeccccc
WITPDFLAFASPQQSPAIRFPEGSDMFKALPRTLDDIDAHPTLPQPFKNVLEHFSTNNIGLVVRLNSPLY
eecccccccccccccccceecccccchhhhccccccccccccccccccceeecccccccceeeeecccccc
SPSYFEAMGIQHIDMIFDDGTCPPLKTVRKFIRLAHEMITVKKKGIAVHCKAGLGRTGCLIGAYLIYRHG
ccceeeeeeeeeeeeecccccccccchhhhhhhhhhhhhhhhcceeeeeecccccccccccceeeeecc
FTANEVISFMRFMRPGMVVGPQQHWLHLNQGTFREWWIEERVEAKLRRELALANPPATPSRTQKSRGLSS
ccccceeeeeeeecceeeecccccceeeeeccccchhhhhhhhhhhhhhhhhhhcccccccccccccccccc
ASQTPTNRSPSNRTPLSEVDDRNNIGVQEDYLPVPTPGQPRKTARMGDRHYPYSRHGNANTNPPPVQKEP
ccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccc
EEVQTAAEVVASLKQARGASEPSDEEVHMRIKSQRKSSNQPSPSHPHAYPHAPTHSSHATHSVNAPSNAR
hhhhhhhhhhhhhhhhccccccccchhhhhhhhhccccccccccccccccccccccccccccccecccchhhhh
AAAVEQSRSRSISHNVTTTSTATGTTTSATTVYHTQYTTIDNDASQDIENLNSRPLAKVRSSKRINGPET
hhhhhhhccccceeceeeeeeeecccccceeccceeeeeeeeeeecccccchhhhhhccccchhhhhhccccccchh
ALRSKEGSGVRKTSGRVGSIGTGPTTSTSAAASATARKVSATGA
hhhhccccceeeecceeeeeccccccccchhhhhhhhheeeeeec
```

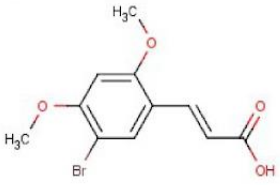
Table 2: Candidate inhibitor structures

A9



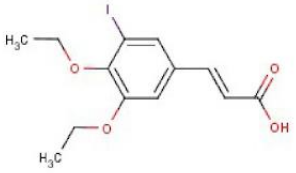
c1cc(c(c(c1)Cl)Cl)C2c3ccc(cc3OC(=C2C#N)N)O

C6



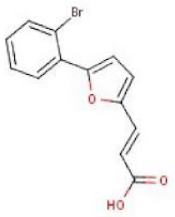
COc1cc(c(cc1C=CC(=O)O)Br)OC

C8



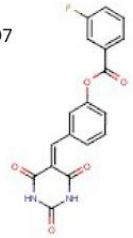
CCOc1cc(cc(c1OCC)I)C=CC(=O)O

D3



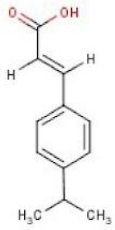
c1ccc(c(c1)c2ccc(o2)C=CC(=O)O)Br

D7



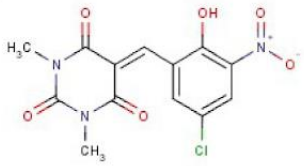
c1cc(cc(c1)OC(=O)c2cccc(c2)F)C=C3C(=O)NC(=O)NC3=O

D9



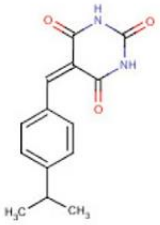
COC1=C(C=C(C=C1)Br)C=CC(=O)O

E1



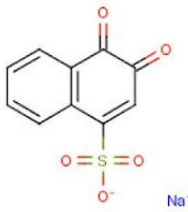
CN1C(=O)C(=CC2=CC(=CC(=C2O)[N+](=O)[O-])Cl)C(=O)N(C1=O)C

G5



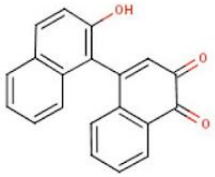
CC(C)c1ccc(cc1)C=C1C(=O)NC(=O)NC1=O

G6



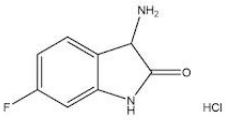
[Na+].[O-]S(=O)(=O)C1=CC(=O)C(=O)c2ccccc21

G7



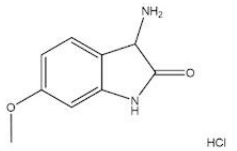
Oc1ccc2ccccc2c1C1=CC(=O)C(=O)c2ccccc21

H5



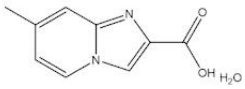
C1=CC2=C(C=C1F)NC(=O)C2N.Cl

H6



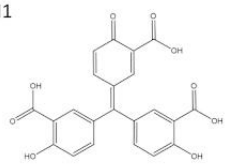
COC1=CC2=C(C=C1)C(C(=O)N2)N.Cl

H7



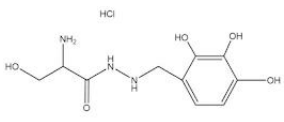
CC1=CC2=NC(=CN2C=C1)C(=O)O.O

I1



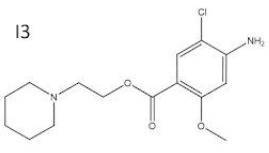
O=C(O)C(C1=O)=CC(C=C1)=C(C2=CC(C(O)=O)=C(O)C=C2)C3=CC=C(C(C(O)=O)=C3)O

I2



O=C(C(CO)N)NNCC1=CC=C(O)C(O)=C1O.Cl

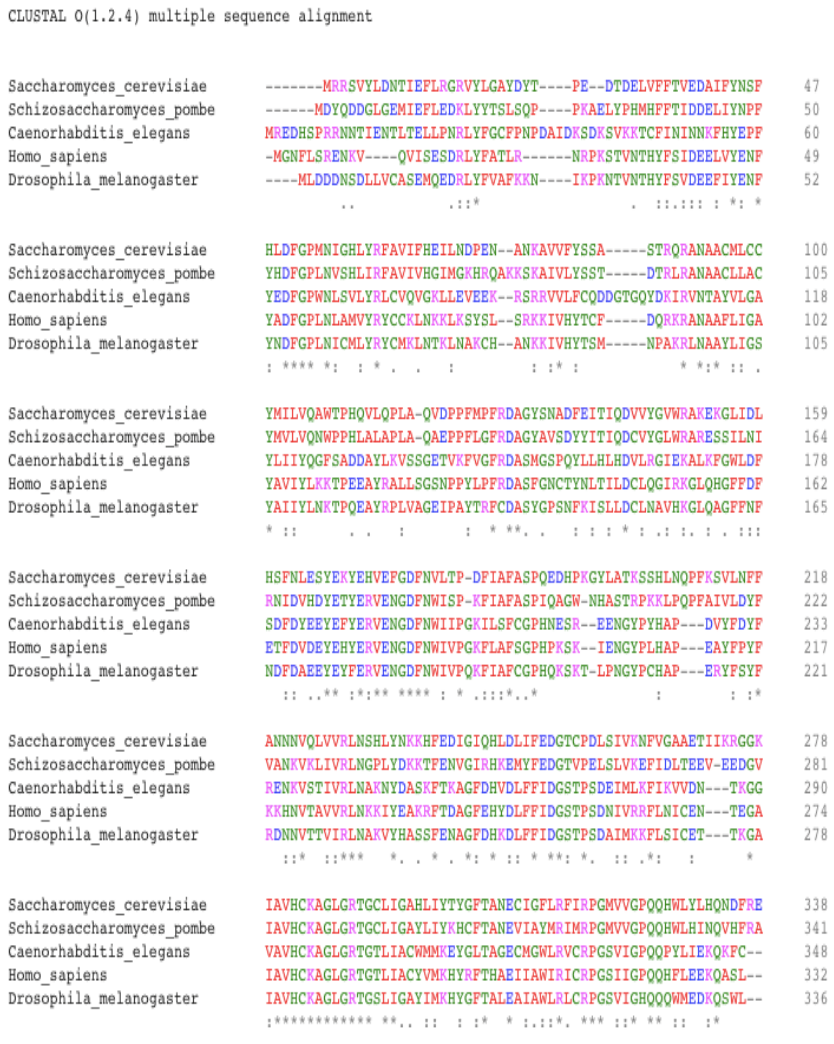
I3

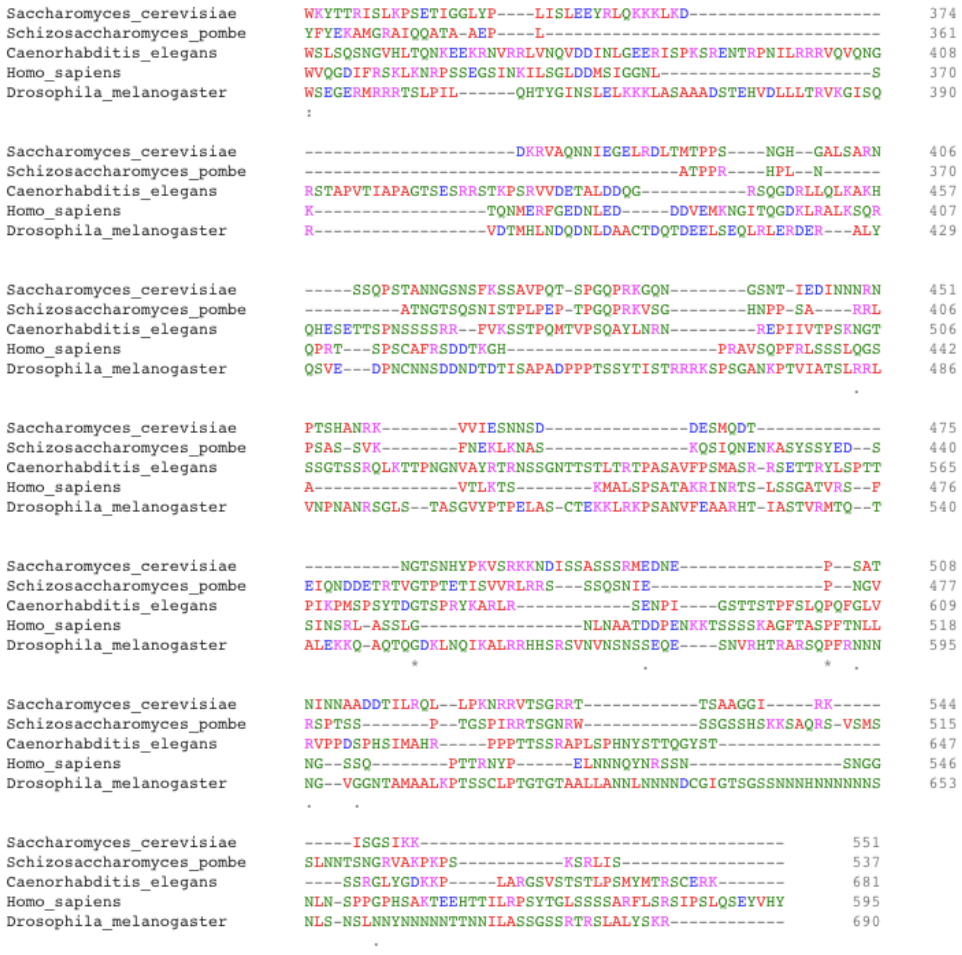


O=C(OCCN1CCCCC1)C2=C(OC)C=C(N)C(Cl)=C2

Table 3. Ortholog sequences			
Species	Protein ID	Percent Identity	Similarity (E-value)
Homo sapiens	XP_005271353.1	35.22%	5e-66
Caenorhabditis elegans	NP_001367521.1	31.27%	6e-57
Drosophila melanogaster	NP_001285722.1	33.42%	2e-64
Saccharomyces cerevisiae	NP_116684.3	52.56%	2e-145
Schizosaccharomyces pombe	NP_594716.1	46.51%	7e-139

Table 4. Alignments of the orthologs created with Jalview and Clustal Omega




Saccharomyces_cerevisiae WKYTTRISLKPSETIGGLYP----LISLEEYRLQKKKLKD-------------------- 374
Schizosaccharomyces_pombe YFYEKAMGRAIQQATA-AEP----L---------------------------------- 361
Caenorhabditis_elegans WSLSQSNGVHLTQNKEEKRNVRRLVNQVDDINLGEERISPKSRENTRPNILRRRVQVQNG 408
Homo_sapiens WVQGDIFRSKLKNRPSSEGSINKILSGLDDMSIGGNL----------------------S 370
Drosophila_melanogaster WSEGERMRRRTSLPIL------QHTYGINSLELKKKLASAAADSTEHVDLLLTRVKGISQ 390
Saccharomyces_cerevisiae ----------------------DKRVAQNNIEGELRDLTMTPPS----NGH--GALSARN 406
Schizosaccharomyces_pombe ---------------------------------------ATPPR----HPL--N------ 370
Caenorhabditis_elegans RSTAPVTIAPAGTSESRRSTKPSRVVDETALDDQG-----------RSQGDRLLQLKAKH 457
Homo_sapiens K------------------TQNMERFGEDNLED-----DDVEMKNGITQGDKLRALKSQR 407
Drosophila_melanogaster R------------------VDTMHLNDQDNLDAACTDQTDEELSEQLRLERDER---ALY 429
Saccharomyces_cerevisiae -----SSQPSTANNGSNSFKSSAVPQT-SPGQPRKGQN--------GSNT-IEDINNNRN 451
Schizosaccharomyces_pombe ----------ATNGTSQSNISTPLPEP-TPGQPRKVSG--------HNPP-SA----RRL 406
Caenorhabditis_elegans QHESETTSPNSSSSRR--FVKSSTPQMTVPSQAYLNRN---------REPIIVTPSKNGT 506
Homo_sapiens QPRT---SPSCAFRSDDTKGH----------------------PRAVSQPFRLSSSLQGS 442
Drosophila_melanogaster QSVE---DPNCNNSDDNDTDTISAPADPPPTSSYTISTRRRKSPSGANKPTVIATSLRRL 486
Saccharomyces_cerevisiae PTSHANRK--------VVIESNNSD---------------DESMQDT------------- 475
Schizosaccharomyces_pombe PSAS-SVK--------FNEKLKNAS---------------KQSIQNENKASYSSYED--S 440
Caenorhabditis_elegans SSGTSSRQLKTTPNGNVAYRTRNSSGNTTSTLTRTPASAVFPSMASR-RSETTRYLSPTT 565
Homo_sapiens A--------------VTLKTS--------KMALSPSATAKRINRTS-LSSGATVRS--F 476
Drosophila_melanogaster VNPNANRSGLS--TASGVYPTPELAS-CTEKKLRKPSANVFEAARHT-IASTVRMTQ--T 540
Saccharomyces_cerevisiae ----------NGTSNHYPKVSRKKNDISSASSSRMEDNE---------------P--SAT 508
Schizosaccharomyces_pombe EIQNDDETRTVGTPTETISVVRLRRS---SSQSNIE-----------------P--NGV 477
Caenorhabditis_elegans PIKPMSPSYTDGTSPRYKARLR------------SENPI----GSTTSTPFSLQPQFGLV 609
Homo_sapiens SINSRL-ASSLG-----------------NLNAATDDPENKKTSSSSKAGFTASPFTNLL 518
Drosophila_melanogaster ALEKKQ-AQTQGDKLNQIKALRRHHSRSVNVNSNSSEQE----SNVRHTRARSQPFRNNN 595
Saccharomyces_cerevisiae NINNAADDTILRQL--LPKNRRVTSGRRT------------TSAAGGI-----RK----- 544
Schizosaccharomyces_pombe RSPTSS-------P--TGSPIRRTSGNRW------------SSGSSHSKKSAQRS-VSMS 515
Caenorhabditis_elegans RVPPDSPHSIMAHR-----PPPTTSSRAPLSPHNYSTTQGYST----------------- 647
Homo_sapiens NG--SSQ--------PTTRNYP------ELNNNQYNRSSN---------------SNGG 546
Drosophila_melanogaster NG--VGGNTAMAALKPTSSCLPTGTGTAALLANNLNNNNDCGIGTSGSSNNNHNNNNNNS 653
Saccharomyces_cerevisiae -----ISGSIKK------------------------------------ 551
Schizosaccharomyces_pombe SLNNTSNGRVAKPKPS-----------KSRLIS----------------- 537
Caenorhabditis_elegans ----SSRGLYGDKKP-----LARGSVSTSTLPSMYMTRSCERK------- 681
Homo_sapiens NLN-SPPGPHSAKTEEHTTILRPSYTGLSSSSARFLSRSIPSLQSEYVHY 595
Drosophila_melanogaster NLS-NSLNNYNNNNNTTNNILASSGSSRTRSLALYSKR----------- 690