Spring 2023

Structural, Mutational, and Kinetic Characterization of URA4, an Isocytosine Deaminase

Ashlee Hoffman
Western Washington University, ashlee_hoffman@outlook.com

Structural, Mutational, and Kinetic Characterization of URA4, an Isocytosine Deaminase

By

Ashlee Hoffman

Accepted in Partial Completion
of the Requirements for the Degree
Master of Science

ADVISORY COMMITTEE

Chair, Dr. P. Clint Spiegel

Dr. Jay McCarty

Dr. John Antos

GRADUATE SCHOOL

David L. Patrick, Dean

Master's Thesis

In presenting this thesis in partial fulfillment of the requirements for a master's degree at Western Washington University, I grant to Western Washington University the non-exclusive royalty-free right to archive, reproduce, distribute, and display the thesis in any and all forms, including electronic format, via any digital library mechanisms maintained by WWU.

I represent and warrant this is my original work, and does not infringe or violate any rights of others. I warrant that I have obtained written permissions from the owner of any third party copyrighted material included in these files.

I acknowledge that I retain ownership rights to the copyright of this work, including but not limited to the right to use all or part of this work in future works, such as articles or books.

Library users are granted permission for individual, research and non-commercial reproduction of this work for educational purposes only. Any further digital posting of this document requires specific permission from the author.

Any copying or publication of this thesis for commercial purposes, or for financial gain, is not allowed without my written permission.

Signature: *Ashlee Hoffman*
Date 5/24/2023

Structural, Mutational, and Kinetic Characterization of URA4, an Isocytosine Deaminase

A Thesis
Presented to
The Faculty of
Western Washington University

In Partial Fulfillment
Of the Requirements for the Degree
Master of Science

by
Ashlee Hoffman
May 2023

Abstract

Cytosine Deaminases (CD) are a class of enzymes found in prokaryotes and fungi that have been studied in the past due to their ability to deaminate the prodrug 5-fluorocytosine (5-FC) producing 5-fluorouracil (5-FU). 5-FU is a common anti-cancer drug that can inhibit DNA synthesis leading to cancer cell death. 5-fluorocytosine can interact with digestive bacteria leading to unwanted side effects for cancer patients. Isocytosine Deaminases (ICD) are enzymes that are of interest in the treatment of cancers using Gene Directed Enzyme Prodrug Therapy (GDEPT). ICDs can deaminate the prodrug 5-fluoroisocytosine (5-FIC) also producing the drug 5-FU. 5-FIC will likely not be recognized by digestive bacteria due to ICDs being less common than CDs resulting in reduced side effects for patients. In this study, the crystal structure of URA4, an isocytosine deaminase, was determined using X-ray crystallography. Using the crystal structure of URA4, ten point mutations were made in the active site to better understand the activity of URA4. Live cell assays have been used to determine the viability of human cancers cells after URA4 and the prodrug 5-FIC were added. The enzyme kinetic characteristics were determined for the mutants using both an absorbance assay and isothermal titration calorimetry assay.

Acknowledgements

The work described here was completed from Fall 2021 to Spring 2023 in Dr. P. Clint Spiegel's lab at Western Washington University. I want to thank Clint for the opportunity to join his research lab. I have gained so much experience in the last two years, and I will be forever grateful. Clint has fostered a lab that is collaborative and kind, for which I am incredibly thankful. Next, I want to thank Dr. Kenny Childers, a post-doc in the Spiegel lab. Kenny has provided so much help and support over the last two years. Kenny is always willing to answer questions and provide advice. I also need to thank the other graduate students in the lab, Jess Blair, Nathan Avery, and Jordan Vaughn. Jess trained me in the lab and helped me troubleshoot my project. I am thankful to Jess for helping me to be a better scientist. Thank you, Jordan and Nathan, for helping me with my project, and providing advice. I am grateful that I had their help and support throughout this project. I also want to say thank you to the rest of the Spiegel lab. Everyone has been so wonderful, and have created an amazing environment to work in.

I also want to thank my parents, Lorri and Bruce Hoffman, for their support as I continued my education. When I graduated with my bachelors in June 2020, I didn't think I would ever continue my education. Thank you both for encouraging me to apply for the Master's program. Once again thank you Clint, and the Spiegel lab as a whole.

Table of Contents

List of Figures

List of Tables

Abbreviations

5-FC 5-Fluorocytosine

5-FIC 5-Fluoroisocytosine

5-FU 5-Fluorouracil

AHS Amidohydrolase Superfamily

bCD Bacterial Cytosine Deaminase

BLI Bio-Layer Interferometry

CD Cytosine Deaminase

CDA Cytidine Deaminase

CDP Cytidine Diphosphate

CMPK Cytidine Monophosphate Kinase

CNT Na+ -Dependent SLC28 Family of Concentrative Nucleoside Transporter

CTP Cytidine 5'-Triphosphate

CTPS CTP Synthase

dC Deoxycytidine

dCDP Deoxycytidine Diphosphate

dCK Deoxycytidine Kinase

dCTP Deoxycytidine Triphosphate

DHFU Dihydrofluorouracil

DHOA Dihydroorotate

DHODH Dihydroorotate Dehydrogenase

DPD Dihydropyrimidine Dehydrogenase

dT Deoxythymidine

dTTP Deoxythymidine Triphosphate

dUMP Deoxyuridine Monophosphate

dUTP Deoxyuridine Triphosphate

ENT Na^+ -Independent SLC29 Family Equilibrative Nucleoside Transporter

FdUDP Fluorodeoxyuridine Diphosphate

FdUMP Fluorodeoxyuridine Monophosphate

FdUTP Fluorodeoxyuridine Triphosphate

FUDP Fluorouridine Diphosphate

FUDR Fluorodeoxyuridine

FUMP Fluorouridine Monophosphate

FUR Fluorouridine

FUTP Fluorouridine Triphosphate

GDEPT Gene-Directed Enzyme Prodrug Therapy

ICD Isocytosine Deaminase

ISOC Isocytosine

ITC Isothermal Titration Calorimetry

NDPK Nucleoside Diphosphate Kinase

OA Orate

OMP Orotidine Monophosphate

ORPT Orotate Phosphoribosyltransferase

PRPP 5-Phosphoribosyl-1-Phosphate

RNR Ribonucleotide Reductase

SEC Size Exclusion Chromatography

TK Thymidine Kinase

T_M Unfolding Temperature

TP Thymidine Phosphorylase

TS Thymidylate Synthase

UCK Uridine-cytidine Kinase

UDG Uracil-DNA-glycosylase

UK Uridine Kinase

UMP Uridine Monophosphate

UMPS UMP Synthase

UP Uridine Phosphorylase

UTP Uridine Triphosphate

yCD Yeast Cytosine Deaminase

Chapter 1: Introduction

Cancer

Cancer is the second leading cause of death in the United States.[1] Cancer is caused by the high proliferation of malignant cells. Healthy non-cancerous tissues strictly control the production and release of growth-promoting signals, preventing cells from proliferating. This control over the growth-promoting signals allows the healthy tissue to maintain healthy cell numbers and overall homeostasis. Cancer cells deregulate the growth-promoting signals allowing for the cancer cells to proliferate.[2] With the growth-promoting signals allowed to continue unregulated, the cancer cells can continue to divide and spread without any checks from the rest of the healthy tissue. Cancer cells will form tumors which are complex tissues that contain many different cell types. Cancer can eventually cause death to patients by spreading to healthy organs, and using the nutrients needed by the healthy organs. Tumors need nutrients and oxygen to survive, and during tumor progression angiogenesis is upregulated resulting in the formation of new blood vessels to transport nutrients to the tumor.[2]

Pyrimidine Synthesis

DNA and RNA contain the nucleobases; purines and pyrimidines. (Figure 1) Purines form hydrogen bonds with pyrimidines to establish the double stranded nature of DNA and RNA. Purines are made of two cyclic rings and include adenine and guanine, while pyrimidines contain one cyclic ring, and include cytosine, thymine, and uracil. Thymine is found only in DNA, and uracil is found in RNA. Guanine forms a proper Watson-Crick basepair with cytosine while adenine basepairs with thymine in DNA, and with uracil in RNA. One area of interest in cancer

treatment research is targeting pyrimidine synthesis pathways. Pyrimidines can be made using the *de novo* synthesis pathway or by the pyrimidine salvage pathway in mammals.[3] Healthy cells will preferentially use the *de novo* synthesis pathway to produce cytosine, thymine, and uracil for DNA and RNA synthesis. The *de novo* synthesis pathway can cost the cell between six and ten molecules of ATP.[4] This pathway costs the cells valuable energy when producing the necessary nucleobases for DNA and RNA synthesis.

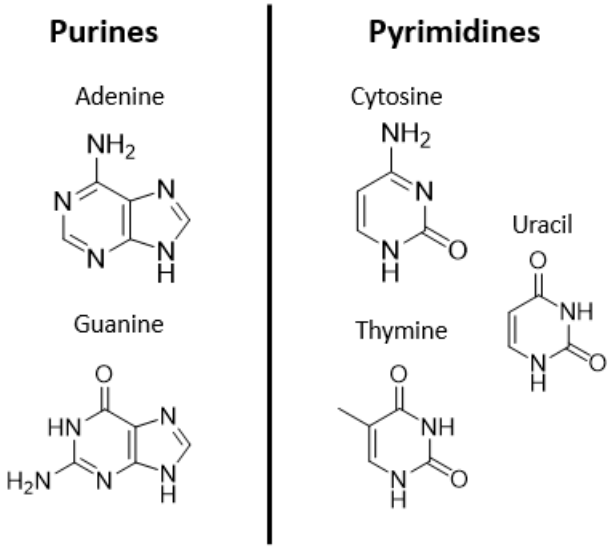

Figure 1: Structures of Purines and Pyrimidines.

Pyrimidine *de novo* Synthesis Pathway

The *de novo* synthesis pathway for pyrimidines begins with glutamine and aspartic acid, the two amino acids are converted to N-carbamoyl-aspartate and dihydroorotate (DHOA) at which point the pyrimidine ring is formed.[3] (Figure 2) Dihydroorotate dehydrogenase (DHODH) catalyzes the formation of orate (OA). OA will then become orotidine monophosphate (OMP) after 5-phosphoribosyl-1-phosphate (PRPP) is added. Glucose is the starting point for PRPP

synthesis and utilizes the pentose phosphate pathway. OMP is metabolized by UMP synthase (UMPS) forming uridine monophosphate (UMP). [3] UMP is the precursor for the formation of the pyrimidines; deoxycytidine triphosphate (dCTP) and deoxythymidine triphosphate (dTTP).

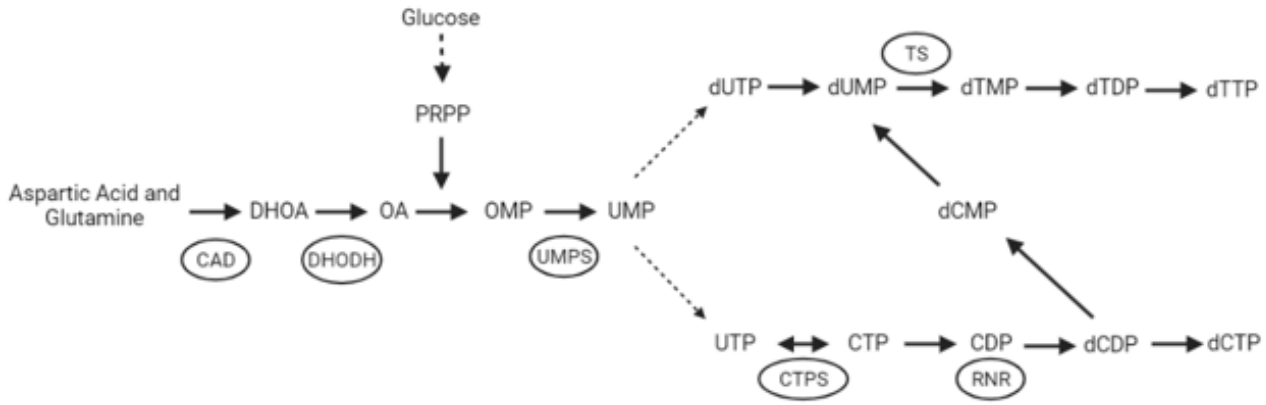

Figure 2: The Pyrimidine de novo synthesis pathway. The circles represent enzymes. Dashed lines represent multiple steps. Adapted from Walter and Herr, 2022.

dCTP is synthesized from UMP which is phosphorylated to uridine triphosphate (UTP) by cytidine monophosphate kinase (CMPK) and nucleoside diphosphate kinase (NDPK). Cytidine 5'-triphosphate (CTP) is formed by CTP synthase (CTPS). NDPK will dephosphorylate CTP to cytidine diphosphate (CDP). Ribonucleotide reductase (RNR) will then reduce CDP to deoxycytidine diphosphate (dCDP). dCTP is formed from dCDP by NRPK. dCTP can then be incorporated into DNA.

dTTP is also synthesized from UMP. Deoxyuridine triphosphate (dUTP) formation is catalyzed by dUTPases, and then dUTP is dephosphorylated to form deoxyuridine monophosphate (dUMP).[3] dUMP can also be formed from dCMP catalyzed by deoxycytidylate deaminase. dTMP is generated following thymidylate synthase (TS) converting dUMP. dTDP is

formed from dTMP by deoxythymidine monophosphate kinase. Finally, NDPK phosphorylates dTDP forming dTTP. dTTP can finally be used in DNA formation.

Pyrimidine Salvage Pathway

Cancer cells produce the nucleobases at a higher rate than healthy cells due to the higher rate of proliferation cancer cells undergo.[4] Cancer cells can use the *de novo* synthesis pathway as well as the pyrimidine salvage pathway to produce the necessary pyrimidines for DNA and RNA synthesis. Cancer cells will use the salvage pathway to avoid using the additional ATP that is needed for the *de novo* synthesis pathway.

The pyrimidine salvage pathway recycles intracellular nucleic acids (UMP, CMP, TMP) and extracellular nucleobases and nucleosides from the bloodstream.[3] (Figure 3) The recycled components can be used to synthesize nucleotides for DNA replication. The nucleosides and nucleobases; deoxycytidine (dC), deoxythymidine (dT), cytosine, and uracil are transported into the cell by either Na^+-dependent SLC28 family of concentrative nucleoside transporter (CNT) or the Na^+-independent SLC29 family equilibrative nucleoside transporter (ENT).[5] The pyrimidines can be converted to dTMP or dCMP by deoxycytidine kinase (dCK) or by thymidine kinases (TKs).[3] dTMP and dCMP can be phosphorylated further forming the final products dTTP and dCTP. Cytidine deaminase (CDA) can catalyze the reaction from deoxycytidine to uracil. Uracil can be phosphorylated by uridine-cytidine kinase (UCK) forming UMP. At this point UMP will eventually become either dCTP or dTTP following the *de novo* synthesis pathway.

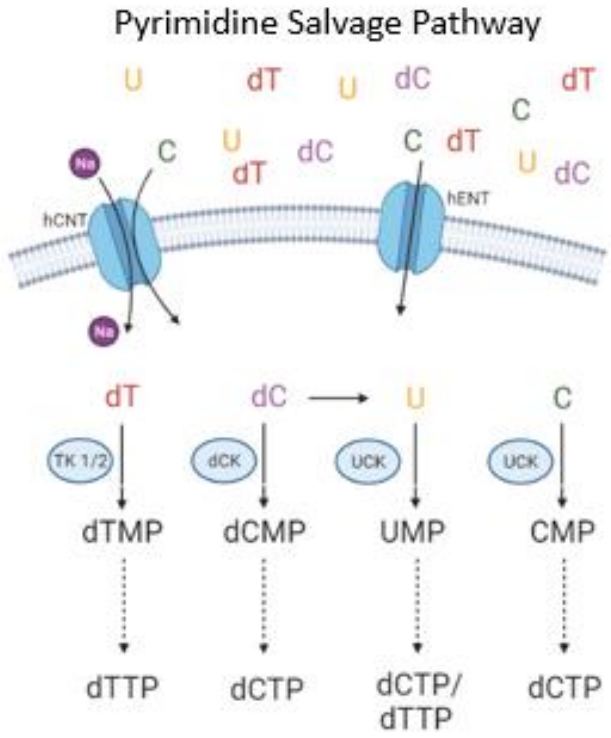

Figure 3: The pyrimidine salvage pathway. The circles represent enzymes. Dashed lines represent multiple steps. Adapted from Walter and Herr, 2022. Created with BioRender.com.

5-Fluorouracil

5-fluorouracil (5-FU) is an anti-cancer drug that takes advantage of the pyrimidine *de novo* synthesis pathway and leads to cancer cell death. 5-FU is an analog of uracil that contains a fluorine at position C-5 of the ring. (Figure 4)

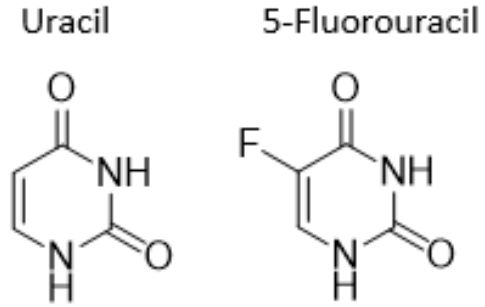

Figure 4: Structures of Uracil and 5-Fluorouracil

Both uracil and 5-FU can enter the cell through the transporter UraA.[6] 5-FU can then be converted to active metabolites including fluorodeoxyuridine monophosphate (FdUMP), fluorodeoxyuridine triphosphate (FdUTP), and fluorouridine triphosphate (FUTP).[7] (Figure 5) 5-FU can be metabolized through several pathways. In one pathway, 5-FU is converted to dihydrofluorouracil (DHFU) by

dihydropyrimidine dehydrogenase (DPD) and is broken down in the liver. 5-FU can also be converted to FUTP by being converted to fluorouridine monophosphate (FUMP). 5-FU can be converted to FUMP directly by orotate phosphoribosyltransferase (ORPT) with PRPP as a cofactor or can be converted to fluorouridine (FUR) by uridine phosphorylase (UP) and then catalyzed by uridine kinase (UK) forming FUMP. FUMP will then be phosphorylated forming fluorouridine diphosphate (FUDP). FUDP will be phosphorylated forming FUTP. FUTP can be incorporated into RNA and disrupt RNA function.

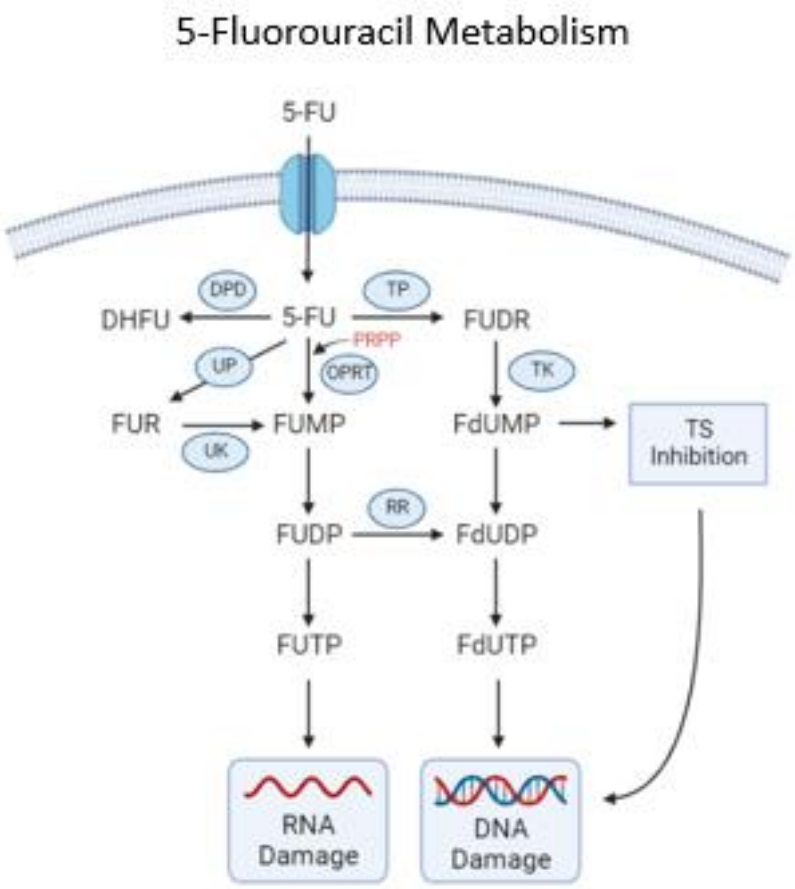

Figure 5: Metabolism of 5-Fluorouracil. Enzymes are represented by circles. 5-FU metabolism leads to RNA and DNA damage causing cell death. Adapted from Longley et al., 2003. Created with BioRender.com.

5-FU can disrupt DNA by being converted to fluorodeoxyuridine (FUDR) by thymidine phosphorylase (TP) and then to FdUMP by TK.[7] FdUMP can be phosphorylated to fluorodeoxyuridine diphosphate (FdUDP), FdUDP can also be formed from FUDP catalyzed by RNR. FdUDP will be phosphorylated to FdUTP. FdUTP can be misincorportaed into DNA. The cell will try to repair the DNA by using uracil-DNA-glycosylase (UDG), but high concentrations of FdUTP will cause further false-nucleotide incorporation. The repeated process of FdUTP being removed by UDG will lead to the DNA strands breaking, eventually causing cell death. The other route of DNA damage is by FdUMP inhibiting TS. Normally TS catalyzes the methylation of dUMP to dTMP in the *de novo* synthesis pathway allowing for dTMP to be used in DNA replication and repair. TS inhibition by FdUMP allows for DNA damage due to a decrease in dTMP being available for DNA replication, further allowing for FdUTP to be used in DNA replication.

Gene-Directed Enzyme Prodrug Therapy

Chemotherapeutic drugs are used extensively in the treatment of cancer in the United States, and around 650,000 patients will receive chemotherapeutic drugs to treat their cancer.[8] Unfortunately chemotherapeutic drugs can be toxic to non-malignant cells at therapeutic dose levels. The patient's heart, liver, kidney, gastrointestinal system, lungs, bone marrow, and nervous system can experience organ toxicity from the chemotherapeutic drugs.[8] Due to the damage chemotherapeutic drugs can cause to the patient, the therapeutic window, time range in which the treatment is effective without adverse side effects, can be limited.

Gene-directed enzyme prodrug therapy (GDEPT) is a treatment that can be used to target cancer cells without damaging healthy non-malignant cells. GDEPT requires a gene that

codes for an enzyme that can convert a prodrug to the active form of the drug, a prodrug, and a carrier that can deliver the gene to the cancer cell.[9] Figure 6 shows how GDEPT is used in the treatment of cancer. First, the gene coding for the enzyme is cloned into a vector and delivered to the cancer cell via a carrier. The gene coding for the enzyme is then transcribed into mRNA, and then translated forming the enzyme. The prodrug can then be administered to the patient systemically. The prodrug is an inactive form of the drug and should not cause cytotoxicity to healthy cells that do not contain the gene for the enzyme. Once the prodrug reaches the cancer cells, it can be converted to the drug by the enzyme. The drug can then target the cancer cell, selectively targeting the cancer cell and causing death. It is unlikely, however, that all cancer cells will receive the gene that codes for the enzyme. There is a mechanism that can compensate for inefficient gene transfer called the bystander effect.[9] The bystander effect is the passive diffusion of the drug between cancer cells. In order for enzymes to be good candidates for GDEPT, they need to have high catalytic activity in order to effectively convert the prodrug to active drug at low concentrations. The prodrugs need to be non-toxic prior to being converted into the drug and highly toxic after being converted into the drug. Since the toxic drug is only produced in the cancer cells, the healthy non-malignant cells will have very little exposure to the toxic drug. Resulting in an increase in the therapeutic window of GDEPT over chemotherapeutic drugs.

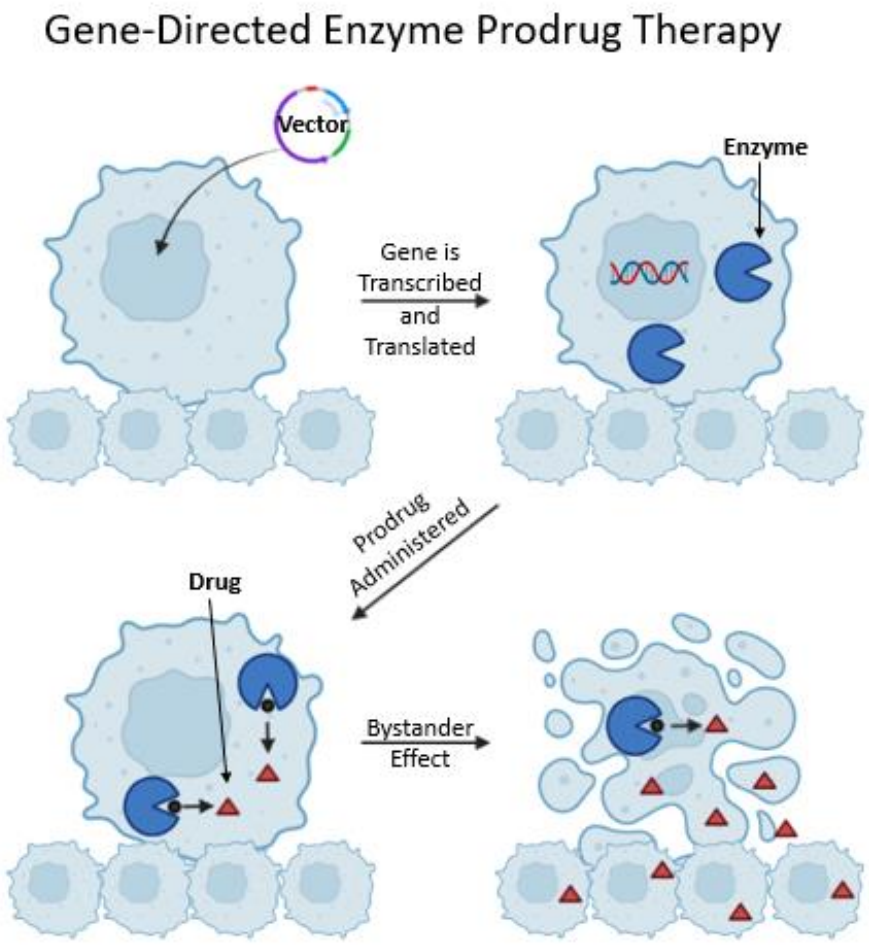

Figure 6: Mechanism of Gene-Directed Enzyme Prodrug Therapy. The vector that contains the gene coding for the enzyme is inserted into the cancer cell. The gene is transcribed and translated allowing for the enzyme to be produced by the cancer cell. The prodrug (black circle) is administered, and the enzyme converts the prodrug into the drug (red triangle). The drug will then cause the death of the cancer cell, and via the bystander effect, the drug can diffuse to other cancer cells that do not contain the gene coding for the enzyme. Adapted from Zhang et al., 2015. Created with BioRender.com.

Cytosine Deaminases

Cytosine Deaminases (CD) are a class of enzymes found in prokaryotes and fungi. CDs convert cytosine into uracil and ammonia as part of the pyrimidine salvage pathway that is unique to bacteria and fungi.[10] (Figure 7) These enzymes are of interest in cancer therapies due to their ability to catalyze the reaction of the prodrug 5-fluorocytosine (5-FC) into the drug 5-FU. CDs have been studied as a potential GDEPT treatment for breast, colon, stomach,

pancreas, and brain cancers.[3] CDs have studied in several clinical trials including a trial targeting gliomas using human neural stem cells as the delivery system of the enzyme and a trial targeting prostate cancer using adenovirus to deliver CD.[11,12]

CDs are not found in mammalian cells which allows for the CDs to be used in GDEPT. There are two main CDs that have been studied, bacterial cytosine deaminases (bCD) from *E. coli* and yeast cytosine deaminase (yCD). bCD is a hexamer, and is roughly 300 kDa while yCD is a homodimer and is 34 kDa.[10] yCD is known to be more active than bCD, but yCD is less thermostable. (Table 1) The unfolding temperature (T_M) of yCD is 52°C.[13] The T_M of bCD is 85 ± 2°C. [14] Mutations have been made to both yCD and bCD that have increased activity and thermostability, but the enzymes can cause digestive issues for patients due to 5-FC being degraded by *E. coli* gut flora resulting in the *E. coli* dying. [15]

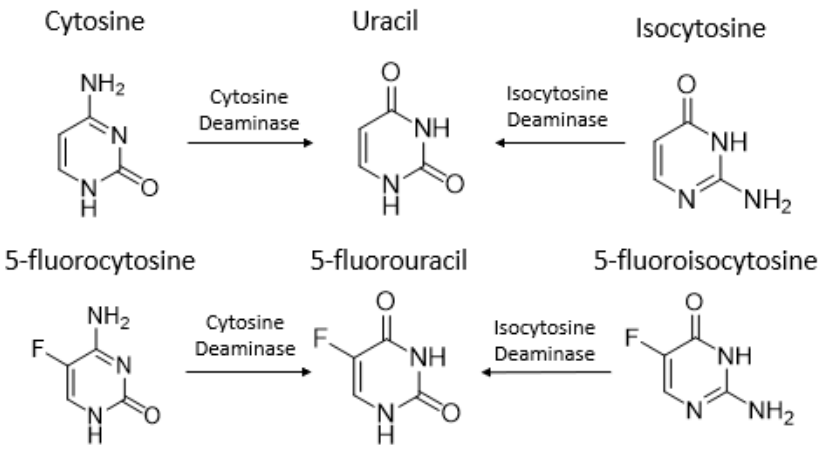

Figure 7: Deamination of cytosine derivatives by Cytosine Deaminase or Isocytosine Deaminase.

Table 1

Enzyme	bCD	yCD	URA3	VCZ
Substrate	Cytosine	Cytosine	Isocytosine	Isocytosine
K_{cat} (s^{-1})	49.68	170	2.00 ± 0.15	1.64 ± 0.06
K_M (µM)	460	1170	1085 ± 120	1860 ± 45
K_{cat}/K_M (M^{-1} s^{-1})	1.07 x 10^5	1.45 x 10^5	1800 ± 100	880 ± 30

Data for bCD[16], yCD[13], URA3 and VCZ[15]

Isocytosine Deaminases

One possible solution to the *E. coli* gut flora dying due to 5-FC is using Isocytosine Deaminases (ICD) for GDEPT instead. ICDs catalyze the deamination of isocytosine (ISOC) into uracil and ammonia as well the deamination of the prodrug 5-fluoroisocytosine (5-FIC) into 5-FU. [15] (Figure 7) Aucynaite et al. discovered three ICDs via screenings of metagenomic libraries from uncultivated soil bacteria. When screening for ICDs, the samples were grown in an *E. coli* strain that is unable to grow in the media without a source of uracil. The strains that could grow were ones that could make uracil using ISOC as the substrate. The three ICDs identified are called URA3, URA4, and VCZ. The kinetic parameters of URA3 and VCZ were similar to those of bCD (Table 1). URA4 has not yet been studied, and will be the subject of this thesis. URA4's structure will be determined using X-ray crystallography. Using the crystal structure model, point mutations were made in the active site of URA4 in an effort to increase kinetic activity.

URA4 will also be studied for cytotoxic effects to better understand how the enzyme will work within the cellular environment.

Chapter 2: Materials and Methods

Transformations

URA4 WT and 10 mutant constructs (G332S, G332S, G332N, S335A, S335T, S335N, S335D, Q79N, Y136F, and E246D) were ordered from Genscript in a pET-28a+ vector with *NdeI* and *XhoI* restriction sites. All constructs were transformed into BL21 (DE3) *E. coli* cells from New England Biolabs. The competent cells (100 µL) were thawed on ice for 15 minutes, and then 1 pg of plasmid DNA was added to the competent cells. Then incubated on ice for 30 minutes. The cell then were heat shocked for 45 seconds at at 40 ℃ and placed on ice for 3 minutes, and mixed with 1 mL of Lysogeny Broth (LB) (1% tryptone (w/v), 1% NaCl, 0.5% yeast extract (w/v)). The cells were then incubated in a shaker at 37 ℃ and 180 rpm for 1 hour before being plated on LB agar plates containing 35 mg/mL kanamycin. The plates were incubated overnight (16-18 hours) at 37 ℃. A single colony was selected for a cell stock and allowed to grow overnight (16-18 hours) in LB with 35 mg/mL Kanamycin in a shaker at 37 ℃ and 180 rpm. The cell stocks were stored at -80 ℃ after 20% (v/v) glycerol was added.

Overexpression and Purification of URA4 and Mutants

Overnight cultures were grown from cell stocks in 10 mL LB with 35 µg/mL Kanamycin while shaking at 210 rpm and 37 ℃. Then the overnight cultures were added to 1 L LB with 35 µg/mL Kanamycin while shaking at 180 rpm and 37 ℃ until the absorbance at 600 nm reached a value of 0.6-0.8. To induce protein growth 1 mM isopropyl B-D-thiogalactopyranoside (IPTG) was added and the temperature was lowered to 18 ℃ and incubated overnight for 18-20 hours. Cells were pelleted by centrifugation at 8327x g for 10 minutes at 4 ℃ (FIBERLite F10-6x500y

rotor Thermo Fisher Scientific). Cell pellets were resuspended in 6 mL of Lysis buffer (50 mM Tris-HCl pH 7, 200 mM NaCl, 10 mM imidazole pH 7, 2% (v/v) glycerol and 0.1% (v/v) Triton X-100) for every one gram of cell pellet. The cells were mechanically lysed using a Branson Sonifier 450 probe at power output 5 and duty cycle 50% for 30 seconds repeated four times with 30 second break between each cycle. Cell debris was removed by centrifugation at 35,003x g for 45 minutes at 4 ℃ (FIBERLite F20-12x50y rotor Thermo Fisher Scientific) followed by filtration with 5 µM sterile syringe filters. Immobilized metal affinity chromatography (IMAC) was performed using TALON cobalt immobilized affinity resin (TAKARA Bio). The column volume (CV) was 5 mL of TALON resin. The TALON resin was equilibrated with 4 CV of Lysis buffer. The filtered lysate was then added to the column. The resin was washed with 10-15 CV of Wash buffer (50 mM Tris-HCl pH 7, 200 mM NaCl, 10 mM Imidazole pH 7, and 2% (v/v) Glycerol); and eluted with Elution buffer (50 mM Tris-HCl pH 7, 200 mM NaCl, 250 mM Imidazole pH 7, and 2% (v/v) Glycerol). Eluted protein was dialyzed into storage buffer (50 mM Tris-HCl pH 7, 200 mM NaCl, and 2% (v/v) Glycerol). Purity was determined via SDS-PAGE using 12.5% bis-acrylamide gels. Protein was concentrated with a 10 kDa spin filter to a concentration between 5-10 mg/mL and stored at -80 ℃.

TEV Cleavage

URA4 has a 6x His-tag for protein purification. The His-tag was removed via TEV protease. The TEV protease concentration is 1 mg/mL, and TEV protease is added 1:50 ratio of protein. The protein and TEV protease were added to dialysis tubing and dialyzed overnight in storage buffer. The supernatant was then put over TALON resin, and cleaved URA4 was

collected in the flowthrough. Cleaved URA4 was concentrated with a 10 kDa spin filter to a concentration between 5-10 mg/mL and stored at -80 ℃.

Size Exclusion Chromatography

GE Healthcare S200 was used to determine the molecular weight of cleaved URA4 WT. The column was equilibrated with dialysis buffer at a flow rate of 0.3 mL/min. Peaks were collected and visualized on a 12.5% SDS page gel. A calibration curve was generated by injecting 50 µL of BioRad Gel Filtration Standard onto the column.

Kinetics Assay - Absorbance Method

The kinetic activity of URA4 WT and mutants were determined in triplicate by monitoring the decrease in absorbance at 285 nm due to the deamination of isocytosine to uracil. Kinetics assays were performed using a BioTek Synergy H1 Microplate Reader. URA4 protein concentration for the assay was 1 µM in the well. Substrate concentration was between 10-900 µM in the well. New isocytosine stocks were prepared each time to avoid possible degradation of isocytosine. Both the protein and substrate were prepared using dialysis buffer. The assay was done in a UV-Vis 96-well plate mixing 100 µL each of substrate and protein. The initial reaction rates were recorded and the kinetic parameters were determined by fitting the experimental data using a Michaelis-Menton equation by GraphPad Prism 8.

Kinetics Assay - Isothermal Titration Calorimetry Method

Isothermal Titration Calorimetry (ITC) instruments measure the thermal power resulting from a solution in the syringe being titrated into another solution in the sample well. [17] (Figure 8) The instrument can detect the heat difference between the sample well and the reference

well that contains only a buffer and will change the power to a feedback heater to keep both wells at the same temperature.

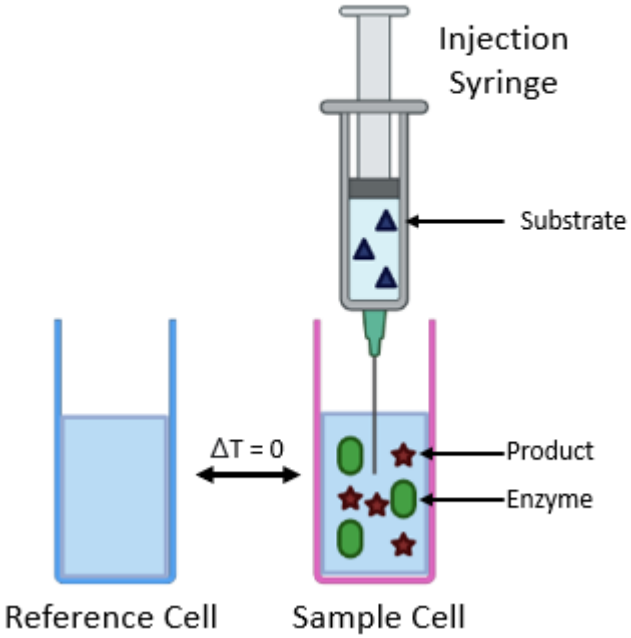

Figure 8: Isothermal Titration Calorimetry. The injection syringe contains substrate (black triangle). The substrate is injected into the sample cell containing the enzyme (green oval) producing the product (red star). The change in heat between the reference and samples cells resulting from the reaction is measured. Created with BioRender.com.

ITC experiments were performed using a MicroCal VP-ITC in high-feedback mode with a stirring rate of 351 rpm. All ITC experiments were performed at 25 ℃. Protein concentration was 2 µM in the sample well. The initial volume in the sample well was 1.4 mL. The substrate concentration in the syringe was 9.5 mM. Dialysis buffer is used to dilute both the protein and substrate. Dialysis buffer was used in the reference well. Before the protein, substrate and buffer are added to the wells and syringe, the components need to be degassed. There was a total of 13 injections of substrate into the sample well (2, 3, 3, 6, 6, 6, 10, 30, 30, 30, 50, 50, and

50 µL). The spacing in between injections for injections 1-7 was 400 sec, injections 8-10 was 600 sec, and injections 11-13 was 800 sec. The increased spacing between injections allows for the baseline to reach equilibrium. Spacing can be adjusted as needed for each enzyme. The kinetic assay is done in triplicate.

Calculations for Kinetics using ITC

The thermal power (dQ/dt), which is a result of one solution being titrated into another, is directly proportional to the reaction velocity (v_0 = d[P]/dt) and ΔH_{app}. ΔH_{app} is the experimentally determined molar enthalpy for the reaction. Thermal power is defined according to the equation below.[17]

$$\frac{dQ}{dt} = V_{cell}\,\Delta H_{app}\,\frac{d[P]}{dt}$$

V_{cell} is the volume of the sample cell. The equation can then be rearranged to give the equation for the rate.

$$Rate = \frac{d[P]}{dt} = \frac{1}{V_{cell}\,\Delta H_{app}} \cdot \frac{dQ}{dt}$$

dQ/dt can be determined by subtracting the average value from the baseline after each injection. (Figure 9)

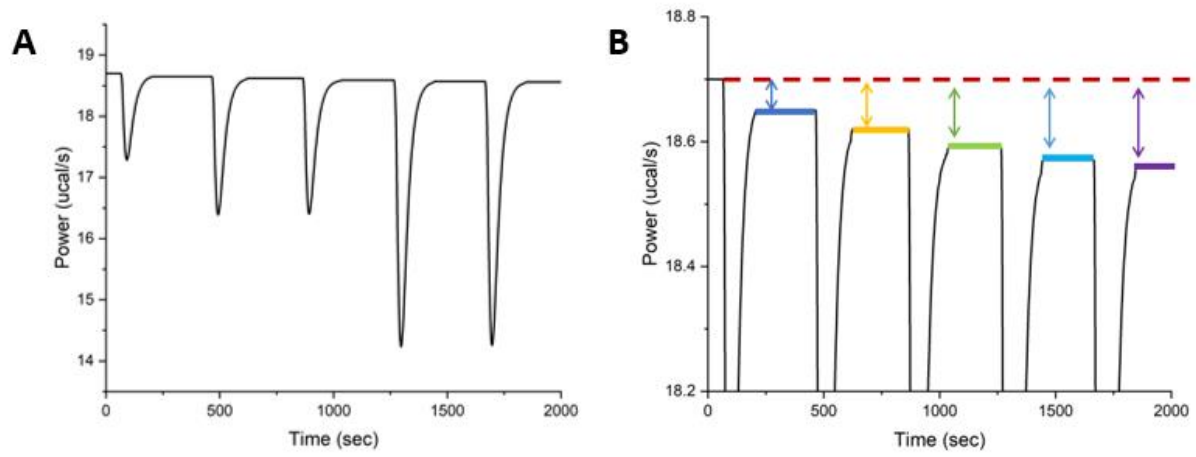

Figure 9: ITC injections. A: The first five injections of substrate. B: Representation of the initial baseline (red dashed line). Each injection's baseline is subtracted from the initial baseline resulting in the overall change in heat for each injection.

ΔH_{app} is experimentally determined by a separate ITC run. Protein concentration was 2 µM in the sample well. The substrate concentration in the syringe was 9.5 mM. Total of 5 injections, 23 µL each. The spacing between each injection is 800 seconds. The spacing between each injection needs to be long enough for the baseline to return to the same value as before the injection. (Figure 10) Integration of the peaks with respect to time is equal to the total heat produced. MicroCal VP-ITC instrument software calculates the total heat produced for each peak. The values can be averaged to give the total heat produced by the reaction. ΔH_{app} is the total heat produced by the reaction divided by the amount of substrate converted.

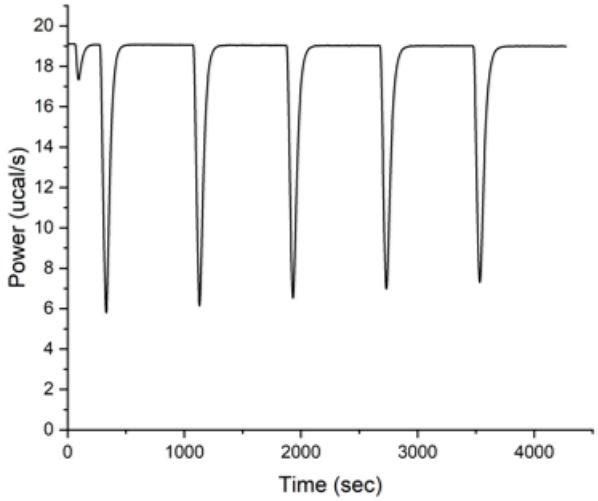

Figure 10: ΔH_{app} ITC. A: Injections 2-6 are 23 µL and spacing is long enough to reach initial baseline. ΔH is calculated for each injection by the software. ΔH is then divided by the injection volume to give $\Delta H_{app.}$

Biolayer Interferometry

Biolayer interferometry (BLI) was used to determine the binding affinity of one subunit of URA4 binding to another subunit of URA4. BLI was conducted with a Pall ForteBio BLITz instrument and BLITz Pro software. BLI was conducted in five steps: initial baseline, loading, baseline, association, and dissociation. A HIS1K Biosensor tip was attached to the instrument and 400 µL of dialysis buffer was added to a LiteSafe Black Microtube. The BLI run is then started, and the tip is submerged into the black microtube for the initial baseline. The tip is then moved into a 4 µL sample containing URA4 WT with 6x His-tag for the loading step. For the baseline step, the tip is then resubmerged into buffer. It is next placed into 4 µL of the second sample that contains cleaved URA4 WT allowing for the two samples to interact during the association step. The tip is then moved back into the black microtube for the dissociation step. A negative control was measured which follows the same procedure except instead the tip is placed into dialysis buffer during the loading step.

Live Cell Assay

Expi293F™ Human Embryonic Kidney cell line (ThermoFisher) was used for cell viability assays. The cells are stored in liquid nitrogen. Expi293™ Expression Medium (ThermoFisher) is heated to 37°C, and then 30 mL of the medium was added to a sterile, vented Erlenmeyer shaker flask. The cells are then thawed in a 37°C water bath for 1 to 2 minutes. The vial of cells is decontaminated by being wiped with 70% ethanol, and then opened in a laminar flow hood. The cells are then transferred to the pre-warmed Expi293™ Expression Medium. The cells are then incubated in a 37°C incubator with an orbital shaker platform at 225 rpm with 80% relative humidity and 8% CO_2. The cells are then allowed to grow for 3-4 days, and then viable cell density is determined. Cell density was determined using a hemocytometer. The viable cell density is typically 1 x 10^6 viable cells per mL. The cells can be subcultured every 3-4 days when the cells reach 3-5 x 10^6 viable cells per mL.

Once the cells have reached viable cell density they can be passaged. The cells are diluted to 0.4 x 10^6 viable cells per mL in 6 wells 3 mL each. The experiment was conducted by adding 1-5 mM 5-FIC and 2 μM URA4 WT and then incubated for four days at which point the cell density was determined. As a positive control 1-5 mM 5-FU was added to a well. For a negative control, water was added to the cell media.

Crystallization and Structure Determination

URA4 WT with TEV site cleaved was concentrated to 10 mg/mL using a 10 kDa MWCO spin column. Crystals were grown by hanging drop vapor diffusion in a 1:1 (v/v) ratio of URA4 WT protein and crystallization solution. The crystallization solution contained 0.2 M $MgCl_2$, 0.1 M BisTris (pH 5-7), and 21-35% (w/v) PEG 3350. The crystals were cryoprotected with a 1:1

addition of mother liquor containing 0.2 M $MgCl_2$, 0.1 M BisTris (pH 5-7), 21-35% (w/v) PEG 3350, and with either 30% (v/v) glycerol or 30% (w/v) PEG 100. X-ray diffraction was collected on the Advanced Light Source (ALS) Berkeley Center for Structural Biology (BCSB) beamline 5.0.1 (Berkeley, CA). Data collection and processing were performed with Adxv, XDS, and Aimless. Phasing of the URA4 WT crystal was accomplished using PHASER-MR with a previous URA4 WT structure. Model building and refinement were performed with WinCoot and PHENIX, respectively. All structure figures and structural alignments were generated with PyMOL Molecular Graphics System, Version 2.0 (Schrödinger, LLC).

Results and Discussion

Characterization of URA4

Protein Expression and Purification

Expression of URA4 was successful in BL21 (DE3) *E. coli* cells. Purification of URA4 via IMAC using TALON cobalt immobilized affinity resin was successfully completed. (Figure 11) Uncleaved URA4 WT is 54 kDa, and after TEV cleavage to remove the 6x His-tag cleaved URA4 WT is 52 kDa.

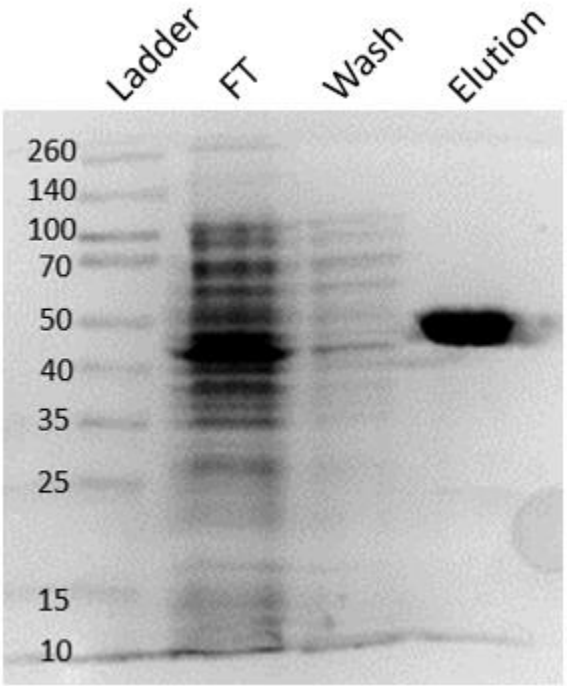

Figure 11: Purification of URA4 WT via TALON IMAC.

Size Exclusion Chromatography

The oligomerization state of URA4 WT is unknown. Size exclusion chromatography (SEC) was performed to determine the oligomerization state. SEC was performed with various concentrations of cleaved URA4 WT resulting in a large peak at a retention time of 15 mL.

(Figure 12) The resulting retention time corresponds to a mass of 52 kDa, and one monomer of URA4 is 52 kDa. A second smaller peak is present at a retention volume of 12.5 mL corresponding to a mass of 210 kDa. There is also a smaller peak present in the 7 mg/mL and 9 mg/mL SEC chromatograms at a retention volume of 11.4 mL corresponding to 380 kDa. This data suggests that URA4's oligomerization state is primarily a monomer, but can form to be a tetramer or possibly a hexamer.

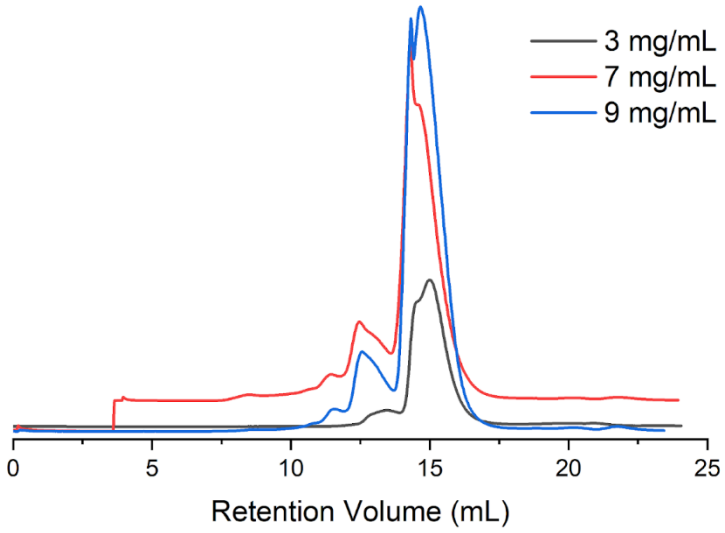

Figure 12: Size Exclusion Chromatograms of cleaved URA4 WT on a S200 column.

Biolayer Interferometry

Biolayer Interferometry (BLI) was used to determine the binding affinity of one subunit of URA4 WT to another subunit of URA4 WT. (Figure 13) BLI measures the interference pattern of white light that is reflected from a biomolecule that is immobilized on the surface of a sensor tip.[18] The sensor tip is coated with antibodies for 6x His-tag that allow for 6x His-tag to bind to the sensor tip. URA4 WT will bind to the sensor tip and be immobilized. Cleaved URA4 WT which has the 6x His-tag removed using TEV protease will be unable to bind to the tip. Any

resulting detection of binding will be a result of binding between two subunits of URA4 WT. The association and dissociation rate constants were 9.5 x 10^3 ± 1.2 x 10^9 $sec^{-1}M^{-1}$ and 6.6 ± 8.0 x 10^5 sec^{-1}, respectively. The K_D was subsequently calculated to be 690 ± 660 µM. The K_D is high with large error suggesting that URA4 WT has a low affinity for oligomerization of the subunits. URA4 WT must be in higher concentrations for the oligomerization states other than a monomer to be present. Overall, this data agrees with the SEC data that URA4 WT is primarily a monomer but can bind forming higher states in higher concentrations.

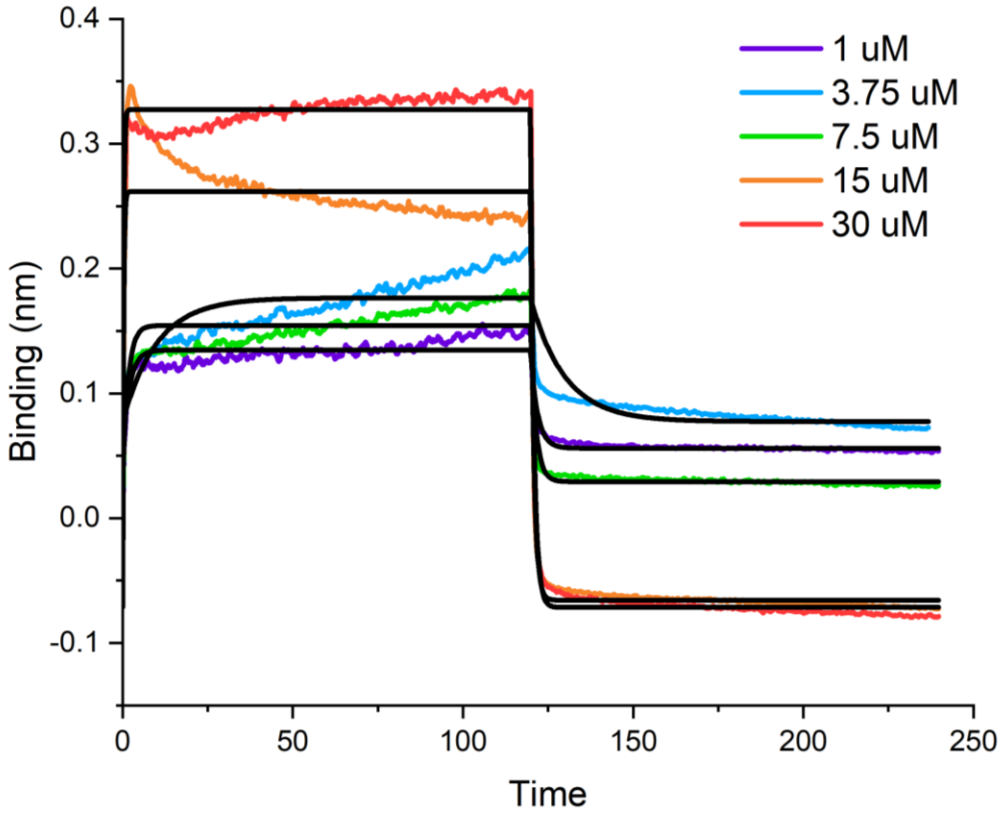

Figure 13: Association and disassociation binding rates for URA4 WT oligomerization. Baseline was established from uncleaved URA4 WT. Cleaved URA4 WT loaded onto the tips was step-corrected, averaged over three independent trials, normalized using Excel, plotted on GraphPad Prism, and fit using an association and dissociation non-linear regression setting.

Crystal Structure

URA4 WT was crystallized, and X-ray diffraction data was collected at the ALS beamline. The structure of URA4 WT was phased and refined to 2.4 Å resolution. (Table 2) The final model refinement statistics resulted in R_{factor} (R_{work}) and R_{free} of 0.1956 and 0.2606, respectively. R_{factor} value below 0.2 indicates a high-quality model.[19] The clashscore for this data set was 13.03. The clashscore is the number of serious clashes per 1000 atoms within a structure model.[20] A low clashscore indicates a well refined-model.

Table 2

X-ray data statistics	
Wavelength (A)	1.00 A
Resolution range (A)	48.25 - 2.43 (2.517 - 2.43)
Space Group	P 2 2 21
Unit cell (A)	72.982 85.738 291.697 90 90 90
Total reflections	138281 (13056)
Unique reflections	69600 (6659)
Completeness (%)	99.28 (96.28)
Mean I/sigma (I)	7.15 (0.92)
R_{pim}	0.0732 (0.7415)
Model refinement statistics	
R_{factor}	0.1956 (0.3194)
R_{free}	0.2606 (0.3555)
Number of Atoms	14718
Protein	14481
Water	233
Protein residues	1924
RMS bonds (A)	0.011
RMS angles (°)	1.38
Ramachandran favored (%)	94.36
Ramachandran outliers (%)	0.89
Rotamer outliers (%)	0.13
Average B-factor	49.6
Protein	49.67
Solvent	44.9

The structure of URA4 is shown in Figure (14A). URA4 folds with a $(\beta/\alpha)_8$ barrel that contains a single metal ion bound in the active site believed to be a zinc ion. The divalent cation is coordinated by five histidine residues (H74, H76, H134, H280, H283), an aspartic acid (D331), and a water molecule. The substrate 5-fluoroisocytosine (5-FIC) was added to the model using Autodock Vina by David Cummins.[21] (Figure 14B) The computed free energy for 5-FIC docking to URA4 was an estimated -5.5 kcal/mol. Based on this model, the substrate is hydrogen bonding with Q79, E246 and S335. W96 may also be participating in pi stacking with the substrate's aromatic ring.

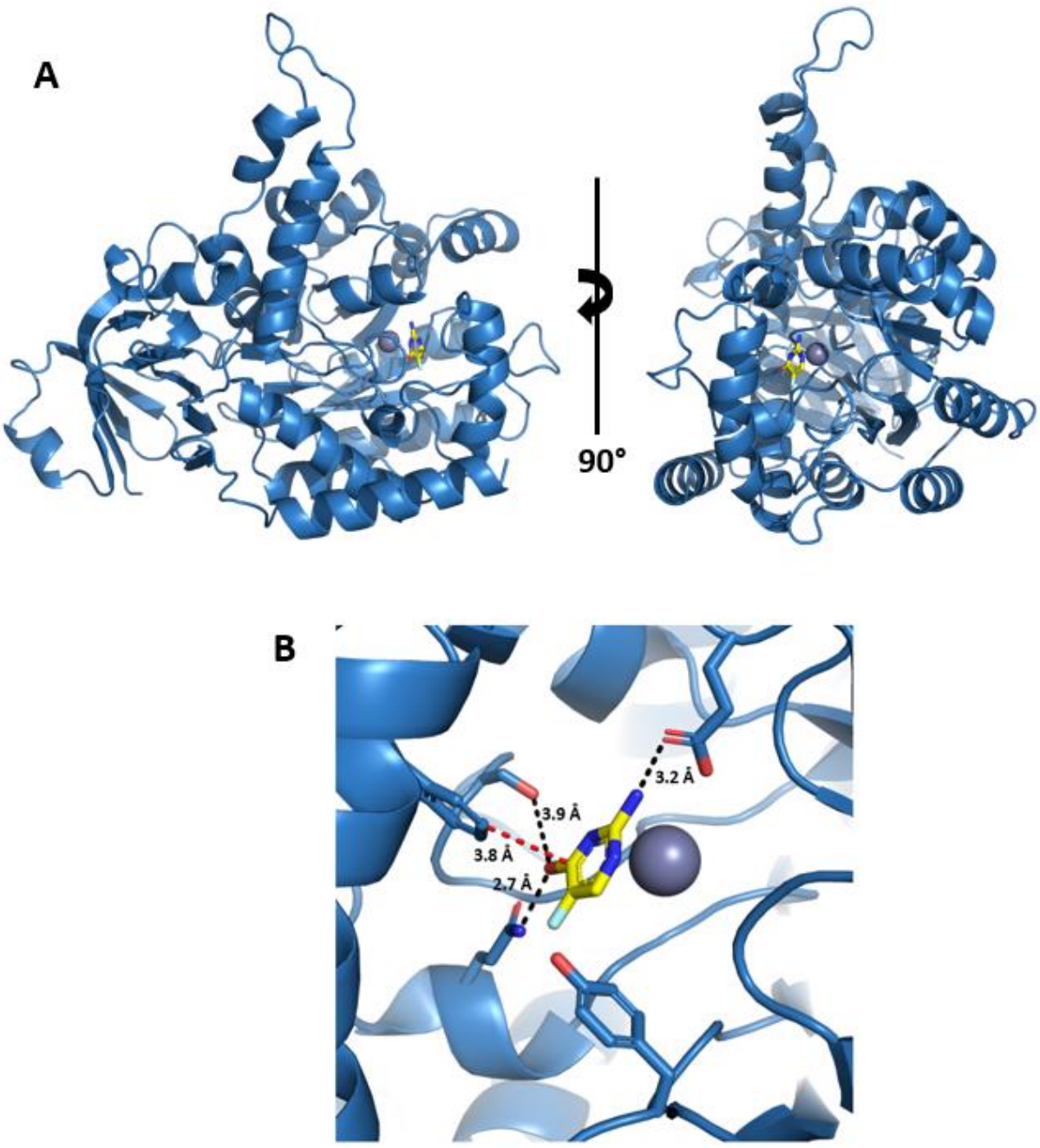

Figure 14: A: The crystal structure of URA4 WT represented in ribbon, with 5-FIC substrate docked in the active site using Autodock Vina. Rotated 90° depicting the active site. B: Magnification of the active site. The substrate is hydrogen bonding with Q79, E246 and S335. Pi stacking might be occurring between the aromatic rings of the substrate and W96.

Proposed Catalytic Mechanism

URA4 is a part of the amidohydrolase superfamily (AHS). Other enzymes that are a part of the AHS superfamily include enzymes that deaminate cytosine, guanine, isocytosine, and 8-oxoguanine. Three conserved amino acids that are found in all the enzymes in the AHS superfamily include an aspartic acid at the end of β strand 8, a histidine at the end of β stand 6, and a glutamic acid at the end of β-strand 5 of the β barrel. [22] In URA4 WT the glutamic acid is E246, the aspartic acid is D331, and the histidine is H280. (Figure 15) The putative mechanism for the deamination of isocytosine is found in Figure (16). The glutamic acid will deprotonate the metal bound water and then protonate the substrate. The histidine may assist in transferring a proton from the water to the glutamic acid. There is a nucleophilic attack by the metal-bound hydroxide, followed by protonation of the leaving group done by the Aspartic acid, and finally the C-N bond is cleaved.

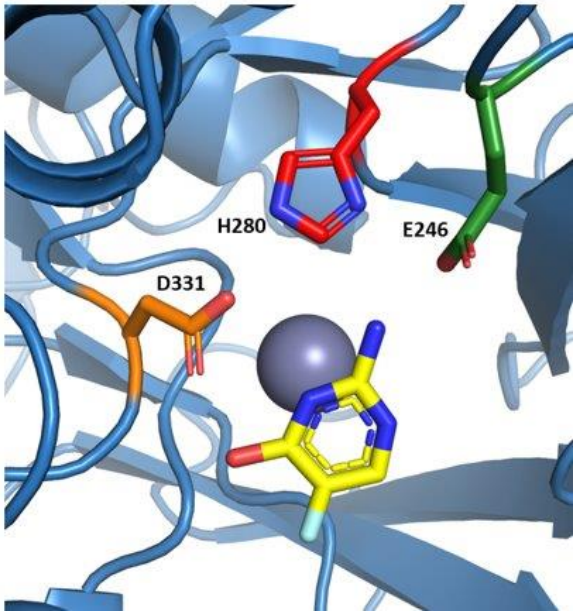

Figure 15: Amino acids involved in the deamination of the substrate, isocytosine or 5-fluoroisocytosine (pictured) D331 and H280 both help coordinate the zinc ion within the active site. E246 helps to coordinate a zinc bound water molecule used in the deamination reaction.

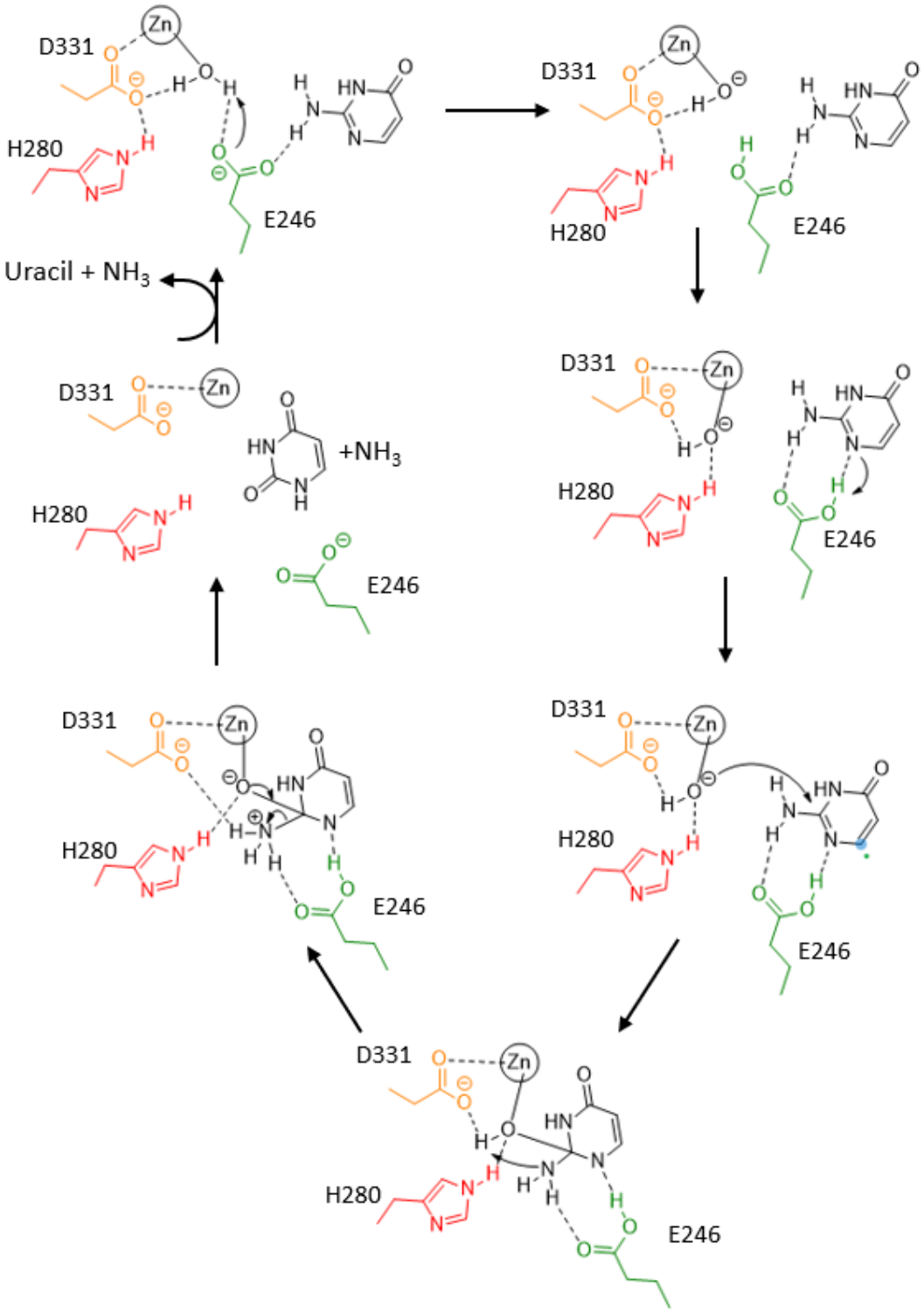

Figure 16: Putative mechanism for the deamination of isocytosine. E246 helps to coordinate a zinc bound water molecule and will deprotonate the water molecule, allowing for the substate to become protonated. A nucleophilic attack by the metal-bound hydroxide is followed by the leaving group becoming protonated by the D331. Finally, the C-N bond is cleaved producing the product uracil or 5-fluorouracil and ammonia. Mechanism is adapted from Manta et al.

Mutational Characterization

Based on the crystal structure, mutations were made within the active site to better understand substrate binding and catalysis. Seven of the point mutations (S335N, S335D, S335T, S335A, Q79N, Y136F, and E246D) were selected based on alignment of the crystal structure of bCD and the crystal structure of URA4 WT with the goal to increase kinetic activity and substrate specificity toward 5-FIC. Three mutations were made at position G332 (G332S, G332A, and G332N). These mutations were made based on a sequence pairwise alignment between URA4 and bCD to increase substrate specificity towards the fluorinated substrate. (Figure 17) The boxed amino acids in Figure (17) include G332 in URA4 and D314 in bCD. Mahan et al. found that the mutations D314A and D314S in bCD displayed a shift in substrate preference towards 5-FC allowing for a 2-4-fold decrease in IC_{50} for 5FC in vitro compared to wild-type bCD-expressing tumor cells.[16] Activity of the mutants was determined by an absorbance method and an ITC method. Of these ten point mutations only S335D and S335A catalyzed the reaction of isocytosine into uracil.

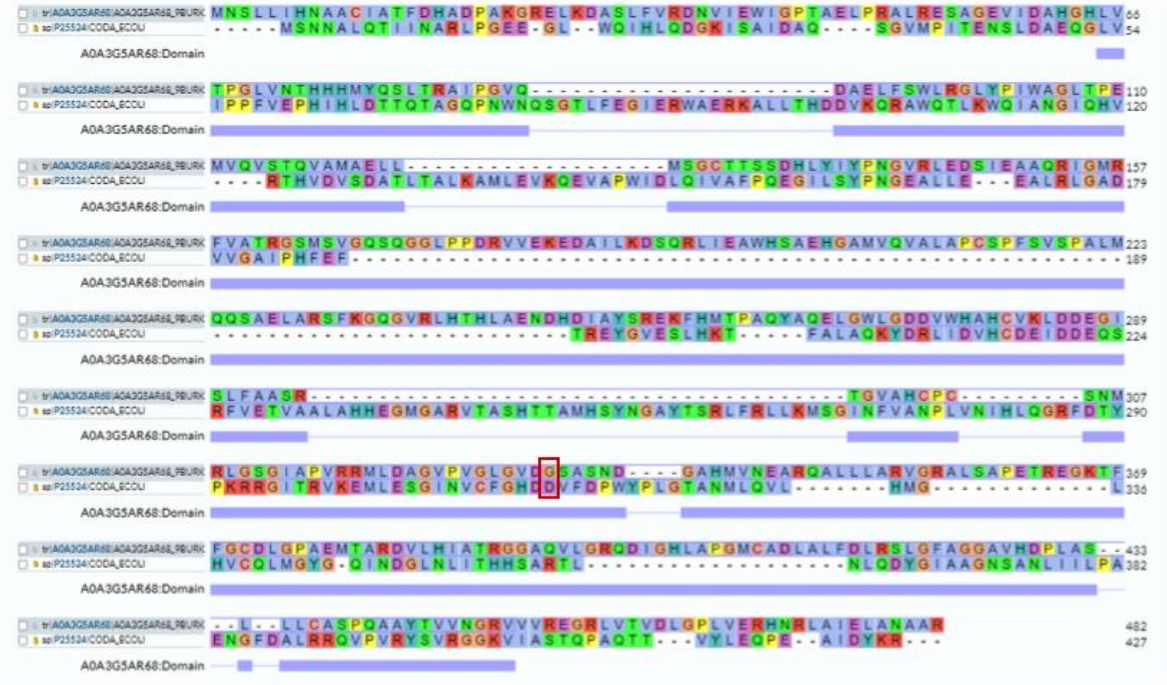

Figure 17: Sequence pair alignment between URA4 WT and bCD. The boxed amino acids represent G332 in URA4 and D314 in bCD. Mutating D314 in bCD to serine or alanine increased substrate specificity towards the fluorinated substrate.

Mutants

S335

Four mutations were made at S335 (S335T, S335N, S335A, and S335D). (Figure 18) The WT S335 may participate in substrate binding. The distance between S335 and the substrate is 3.9 Å, which may be too far for hydrogen bonding. Both the threonine and asparagine mutations caused the enzyme to no longer be active. These mutations may have caused steric clashes with neighboring amino acids preventing substrate binding or proper protein folding. The S335N mutation may have clashed with L309 as shown in Figure (18C). Both S335A and S335D are active. Alanine is a smaller amino acid than serine and is unlikely to experience steric hindrance with any nearby amino acids. The alanine mutation had similar kinetic values to

URA4 WT. The aspartic acid mutant may be able to hydrogen bond with the substrate. The result of the S335D mutation was increasing catalytic efficiency.

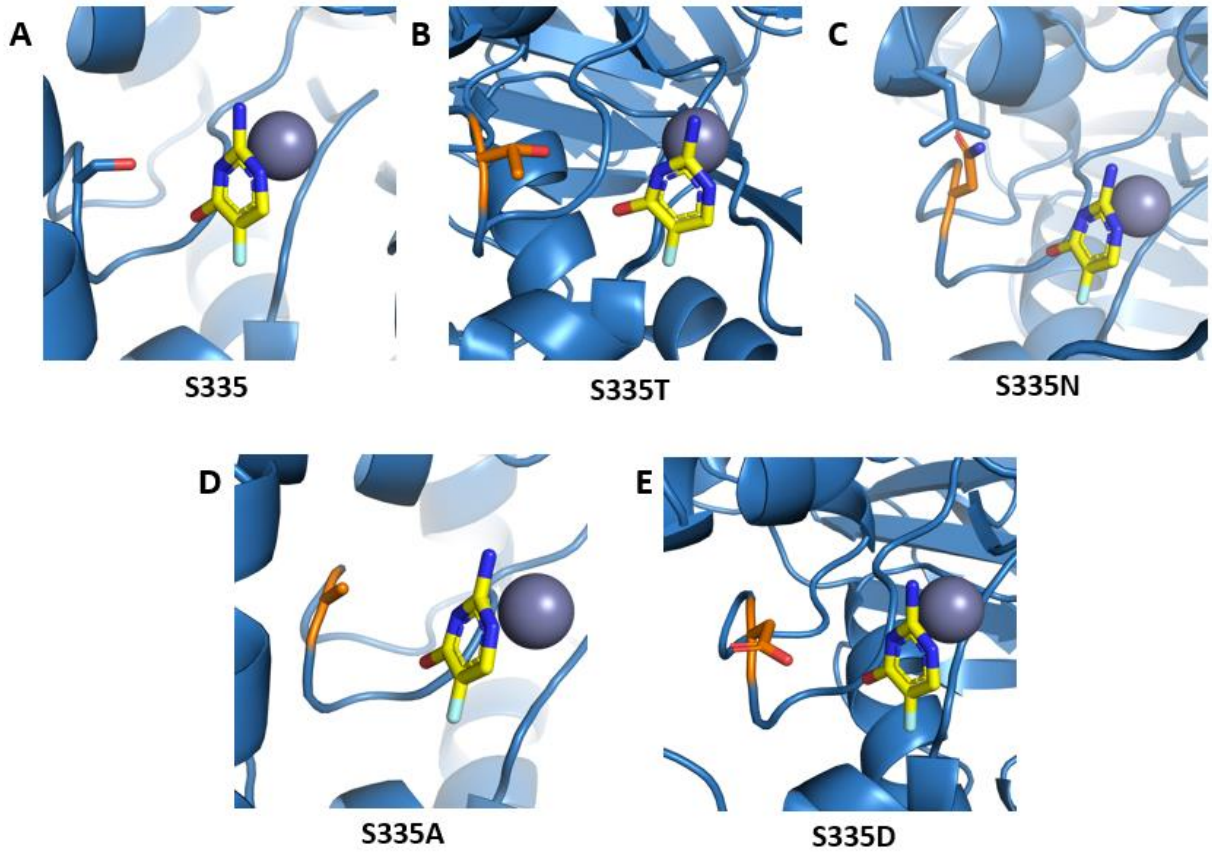

Figure 18: Serine 335 Mutations. A: S335 WT. B. S335T C. S335N. Introduced steric clashes with nearby amino acids. D. S335A. Similar kinetic activity as URA4 WT. E. S335D. Increased kinetic activity over URA4 WT. May hydrogen bond with the substrate.

E246

E246 is an important amino acid for binding the substrate and participating in the catalytic mechanism for the deamination of isocytosine. By mutating the glutamic acid to an aspartic acid, the distance between the substrate and amino acid is increased, and results in loss of catalytic activity. (Figure 19) The aspartic acid is too far away to deprotonate the metal bound water and protonate the substrate. The glutamic acid is highly conserved in the AHS family due to it being a vital amino acid in the catalytic mechanism.

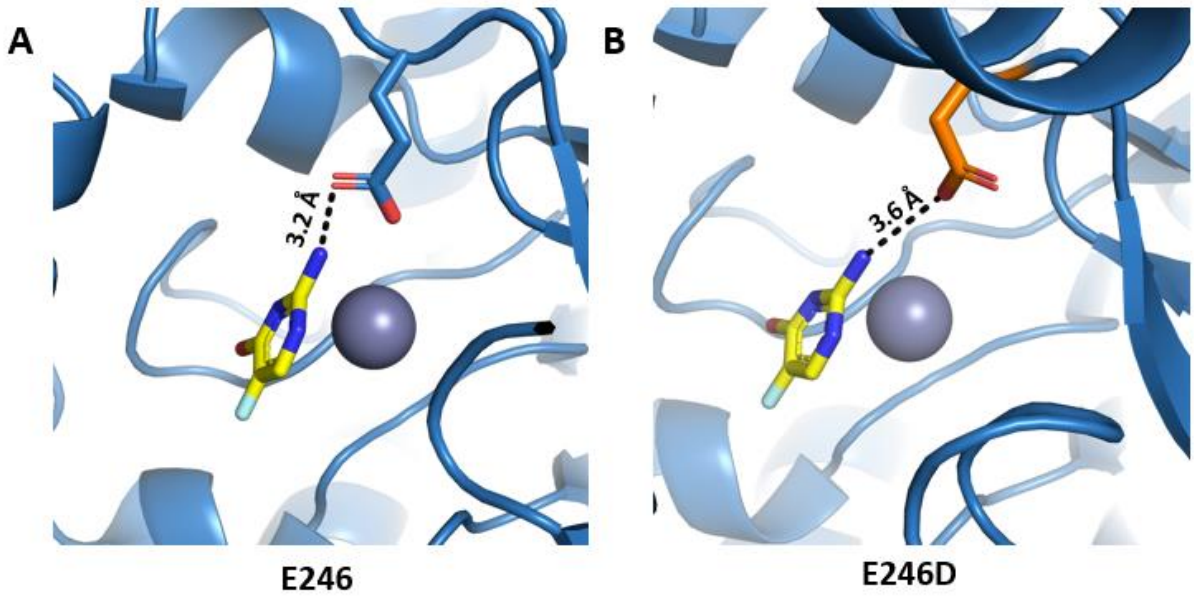

Figure 19: E246 Mutations. A. E246 WT is 3.2 Å away from the substrate in the model. B. E246D mutation increases the distance of the proposed hydrogen bond resulting in loss of substrate binding.

Y136

The tyrosine at position 136 does not appear to be participating in hydrogen bonding in this model. (Figure 20) However, by mutating the tyrosine to a phenylalanine and losing the -OH group results in complete loss of activity for the enzyme. The -OH group may be needed for hydrogen bonding with other nearby amino acids within the protein. More work is needed to explore substrate binding.

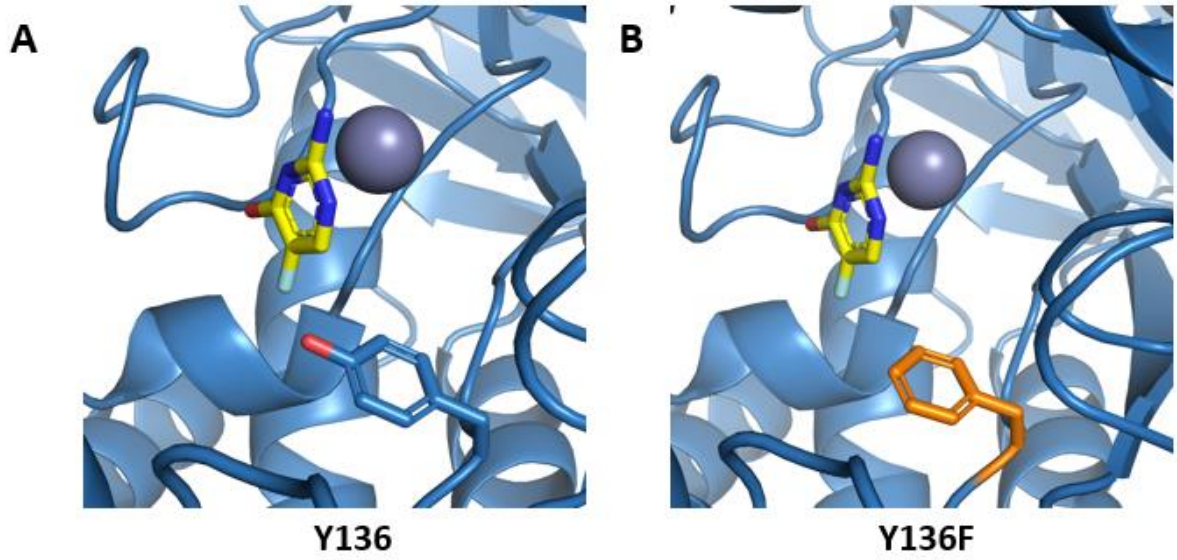

Figure 20: Mutation made at Y136. A. Y136 WT. B. Y136F mutation results in loss of function for the enzyme.

Q79

The glutamine at position 79 is coordinating the substrate by hydrogen bonding with the oxygen of 5-FIC at a distance of 2.7 Å. (Figure 21) The result of mutating the glutamine to an asparagine is loss of enzyme activity. The distance between the asparagine mutant and the substrate is 3.4 Å which is too far for hydrogen bonding. This likely results in the substrate being unable to bind.

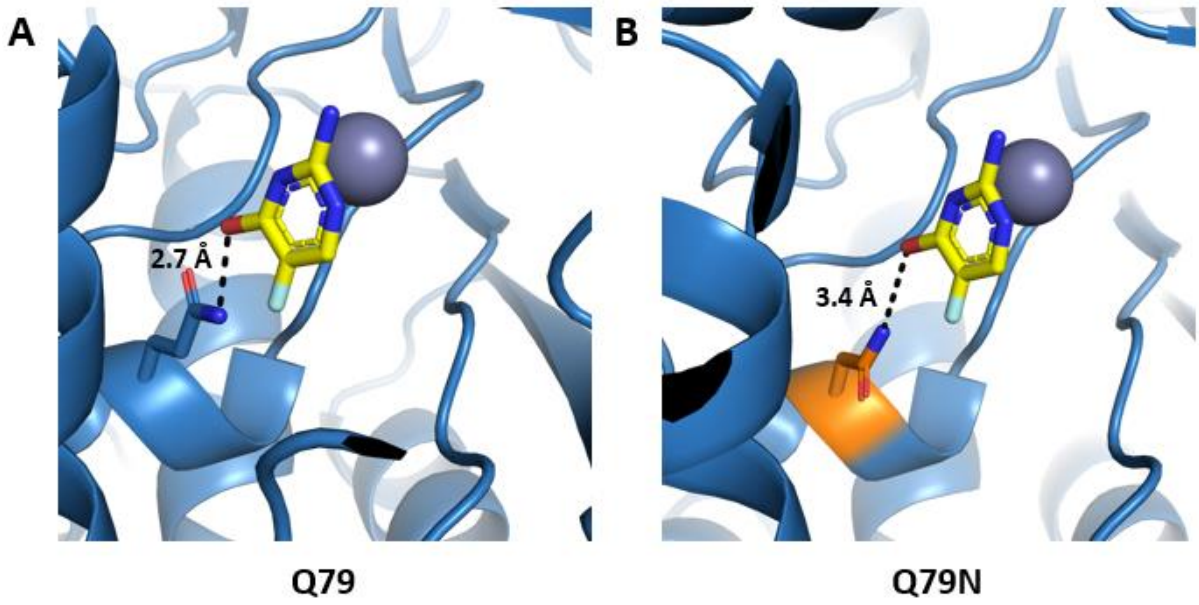

Figure 21: Mutation made at Q79. A. Q79 WT. 2.7 Å distance for possible hydrogen bond. B. Q79N. Distance between the amino acid and the substrate is increased to 3.4 Å. Resulting in a loss of activity for the mutant.

G332

The three point mutations made at G332 (G332A, G332S, G332N) resulted in the protein losing all activity. (Figure 22) These mutations all include larger amino acids than glycine that may have introduced steric clashes with neighboring amino acids. Serine 335 and Aspartic Acid 331 participate in substrate binding and coordinating the metal ion respectively. By mutating glycine 332, the larger amino acids may prevent S335 and D331 from binding to the substrate, resulting in the enzyme no longer being active. Interestingly G332N mutation does bind to the substrate, but it not catalytically active.

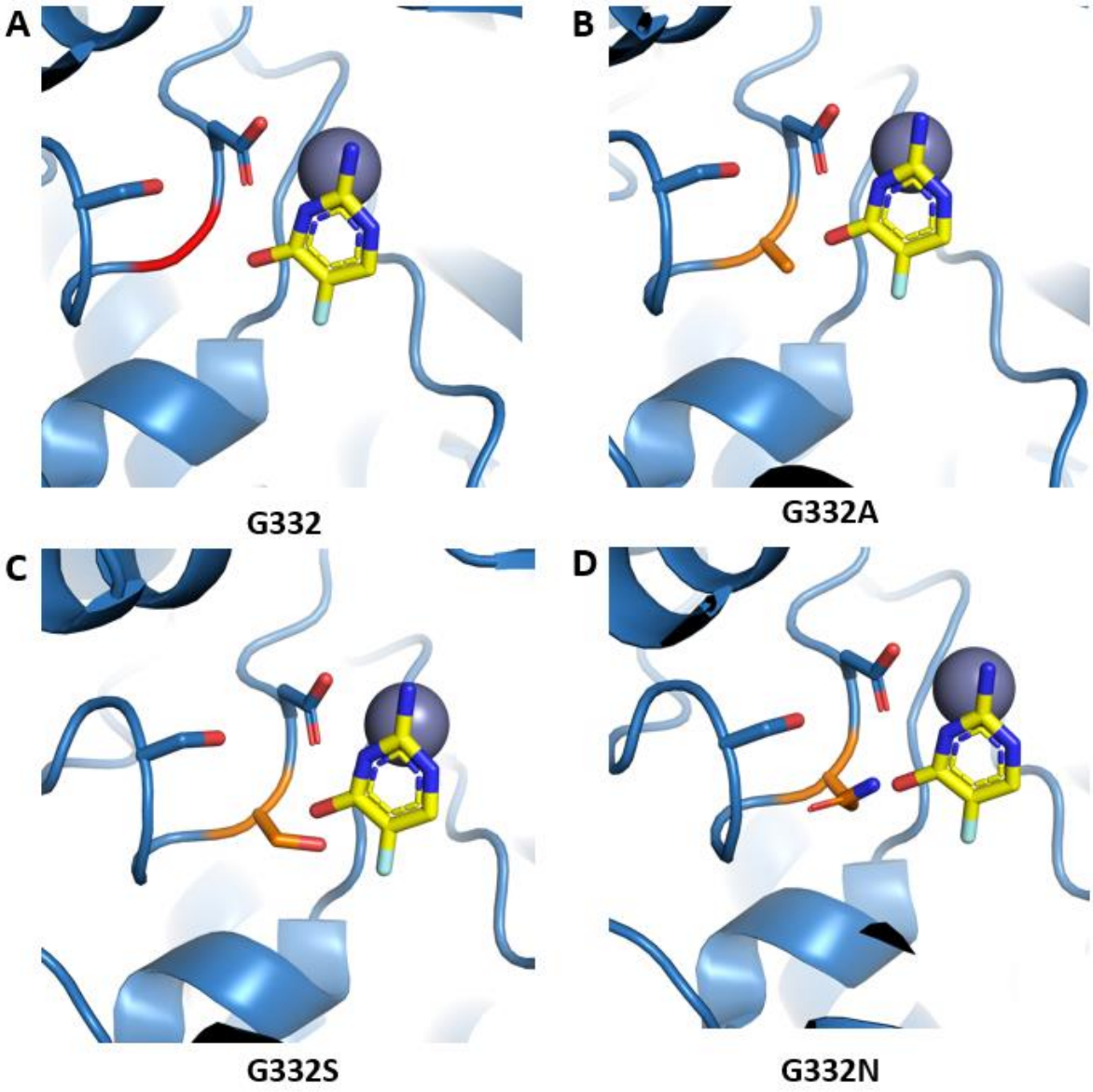

Figure 22: G332 Mutations. S335 and D331 are also shown. A. G332 WT (depicted in red). B. G332A (orange). C. G332S (orange) D. G332N (orange).

Substrate Binding via Isothermal Titration Calorimetry

In an effort to better understand the reason why the inactive mutants caused the enzyme to lose activity, the mutations were studied by isothermal titration calorimetry (ITC) to determine if binding was occurring. ITC can measure the change in heat that is a result of a ligand binding to a protein due to the difference in heat resulting from either an endothermic

or exothermic reaction. When determining if binding was occurring the substrate (ISOC) was titrated into the sample cell that contained the protein, and as a result the change in enthalpy was determined. (Figure 23) ISOC was titrated into buffer alone as a negative control resulting in a change of heat of -7.8 μcal. The titration of ISOC into buffer alone accounts for the heat of dilution. Any proteins that have a significant difference in the change of enthalpy may be binding to the substrate. URA4 WT, S335A, and S335D all have large changes in enthalpy due to their ability to bind and convert the substrate to the product. G332N is the only mutant that is not active and also shows a significant change in enthalpy compared to the negative control. The change in enthalpy for G332N was -194.0 μcal. G332N may be hydrogen bonding to the substrate while also disrupting the substrate from binding to other amino acids in the active site, resulting in the substrate binding but not being converted to the product. More work is needed to determine the binding affinity of ISOC to G332N.

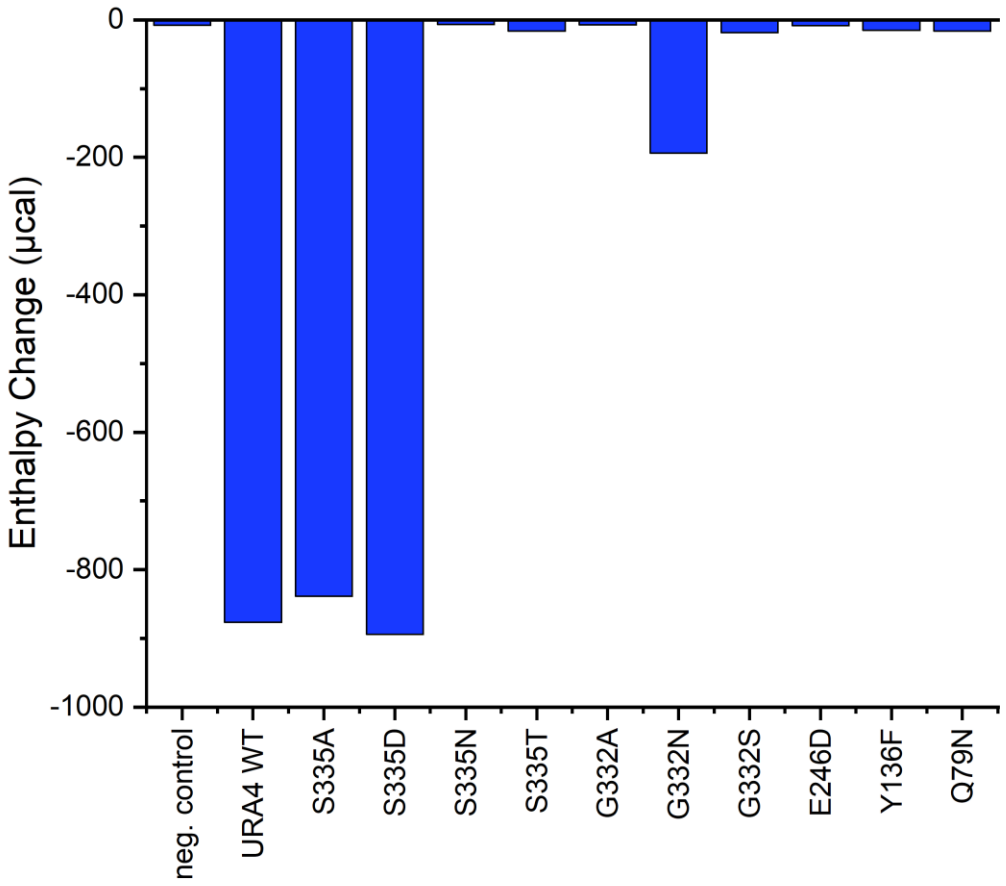

Figure 23: Change in enthalpy. Isocytosine is titrated into URA4 WT or mutants. A large decrease in enthalpy indicates substrate binding. URA4 WT, S335A, and S335D are active enzymes. G332N is an inactive enzyme but appears to bind the substrate.

Kinetics

The kinetic parameters of URA4 WT and the ten point mutations were determined by two methods. The first method observed the decrease in absorbance at 285 nm as isocytosine is converted into uracil by the enzyme. The change in absorbance after the first 15 seconds would be used to determine the reaction velocity. The second method used ITC to measure the change in heat as a result of catalysis. The difference between the initial baseline and subsequent baselines following each injection of substrate was used to determine the reaction

velocity. The kinetic parameters of URA4 WT, S335A, S335D, and VCZ are found in Table 3. VCZ is an isocytosine deaminase that has been previously studied by Aucynaite et al. and in our lab by David Cummins. Aucynaite et al. determined K_{cat}, K_m, and K_{cat}/K_m of VCZ to be 1.64 ± 0.06 sec^{-1}, 1.86 x 10^{-3} ± 4.5 x 10^{-5} M, and 880 ± 30 M^{-1} sec^{-1} respectively using the absorbance method.[15] The K_{cat}/K_m ratio (catalytic efficiency) for VCZ determined by our lab using the absorbance method is 9.49 X 10^3 ± 4.23 X 10^3 indicating a higher rate of catalysis than previously reported.

An issue with the absorbance method is that increasing the concentration of the substrate too high results in the instrument being unable to measure the absorbance in the linear range of detection. The highest concentration we can reasonably use during the absorbance method is 900 µM ISOC, resulting in Vmax not being reached. Figure 24 depicts the Michaelis–Menten curves of URA4 WT by the absorbance method and by the ITC method. The ITC method allows for higher concentrations of substrate to be used, potentially allowing for more accurate kinetic parameters to be determined. The ITC method for kinetics also allowed for 5-FIC as a substrate to be tested.

Lower K_m values correspond to higher affinity for a given substrate. Using the ITC data, the S335D mutation results in a lower K_m indicating that S335D mutation may have increased affinity for the substrate. The catalytic efficiency for S335D is also higher than URA4 WT indicating a higher rate of catalysis than URA4 WT. The S335D mutant's catalytic efficiency is also similar to both bCD (1.07 X 10^5 M^{-1} sec^{-1}) and yCD (1.45 X 10^5 M^{-1} sec^{-1}).

The K_m for URA4 WT with 5-FIC as the substrate is lower than the K_m for URA4 WT with ISOC as the substrate. Indicating that URA4 WT may have a stronger affinity for 5-FIC as the substrate over ISOC. Future work is needed to test 5-FIC with the other active mutants.

Table 3

Absorbance	K_m (M)	K_{cat} (sec^{-1})	K_{cat}/K_m (M^{-1} sec^{-1})
URA4 WT	$6.17 \times 10^{-4} \pm 1.37 \times 10^{-4}$	13.1 ± 1.46	$2.12 \times 10^{4} \pm 2.05 \times 10^{3}$
S335A	$7.08 \times 10^{-4} \pm 1.32 \times 10^{-4}$	4.03 ± 0.40	$5.69 \times 10^{3} \pm 4.48 \times 10^{2}$
S335D	$2.53 \times 10^{-4} \pm 2.76 \times 10^{-4}$	1.99 ± 0.71	$7.85 \times 10^{3} \pm 4.14 \times 10^{3}$
VCZ	$2.40 \times 10^{-4} \pm 4.09 \times 10^{-4}$	2.28 ± 1.91	$9.49 \times 10^{3} \pm 4.23 \times 10^{3}$
ITC	**K_m (M)**	**K_{cat} (sec^{-1})**	**K_{cat}/K_m (M^{-1} sec^{-1})**
URA4 WT	$1.01 \times 10^{-4} \pm 1.54 \times 10^{-5}$	3.86 ± 0.15	$3.82 \times 10^{4} \pm 4.37 \times 10^{3}$
S335A	$1.03 \times 10^{-4} \pm 1.81 \times 10^{-5}$	3.84 ± 0.19	$3.73 \times 10^{4} \pm 4.74 \times 10^{3}$
S335D	$2.05 \times 10^{-5} \pm 4.45 \times 10^{-6}$	6.24 ± 0.21	$3.05 \times 10^{5} \pm 5.83 \times 10^{4}$
VCZ	$2.09 \times 10^{-5} \pm 1.82 \times 10^{-5}$	1.10 ± 0.14	$5.28 \times 10^{4} \pm 4.76 \times 10^{4}$
URA4 WT w/ 5-FIC	$2.96 \times 10^{-5} \pm 1.05 \times 10^{-5}$	16.63 ± 0.93	$5.62 \times 10^{5} \pm 1.78 \times 10^{5}$

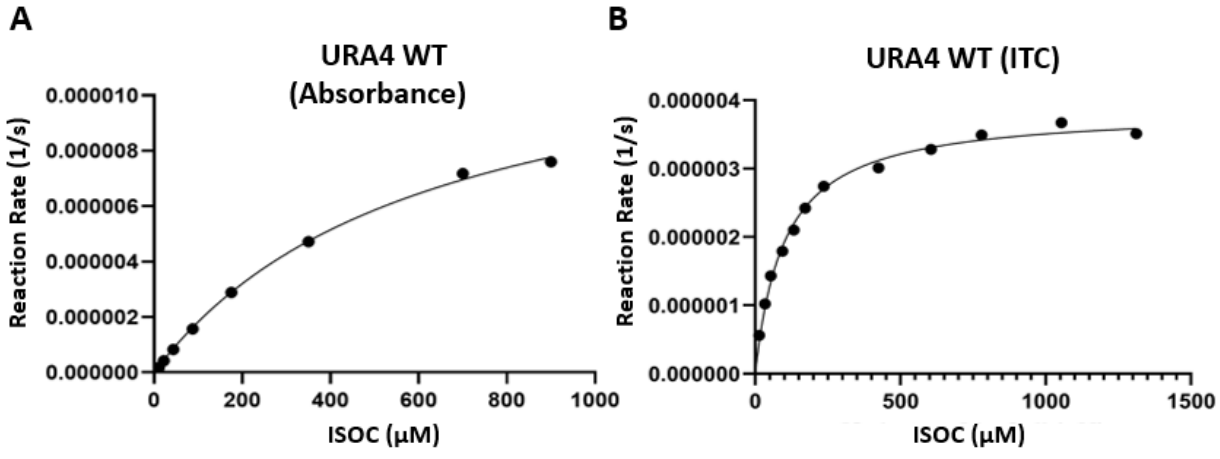

Figure 24: Michaelis–Menten curves of URA4 WT via (A) Absorbance Method and (B) ITC Method. The concentration of substrate is not high enough in the absorbance method to reach $V_{max.}$

Mammalian Cell Work

The prodrug (5-FIC) and enzyme (URA4) system was studied within the Expi293F™ Human Embryonic Kidney cell line to better understand how this system might work in cancer cells. The substrate, 5-FIC has been shown to not be toxic in human glioblastoma U87MG and colorectal adenocarcinoma Caco-2 cell lines at a concentration of 100 µM by Kazlauskas et al. [23] To determine cell viability 2 µM URA4 proteins and 5 mM 5-FIC was added to mammalian cell culture media. For a positive control 5 mM 5-FU was used. Figure 25 depicts the percent cell viability after 5 days of growth. After 5 days the percent of viable cells in the wells containing 2 µM URA4 WT, 2 µM URA4 S335A, and the well in which nothing else was added had percent cell viability of around 90%. The positive control contained 5 mM 5-FU which had a percent viability of 10%. The wells containing the enzyme combined with 5 mM 5-FIC resulted in percent viability of less than 10%. Indicating that both URA4 WT and URA4 S335A are able to convert 5-FIC into 5-FU resulting in similar cell death as 5 mM 5-FU alone. The sample containing URA4 S335D had a percent viability of 15%. The well containing the mammalian cell media and 2 µM URA4 S335D may have been contaminated with 5-FIC or 5-FU resulting in higher cell death than expected. This particular experiment needs to be repeated.

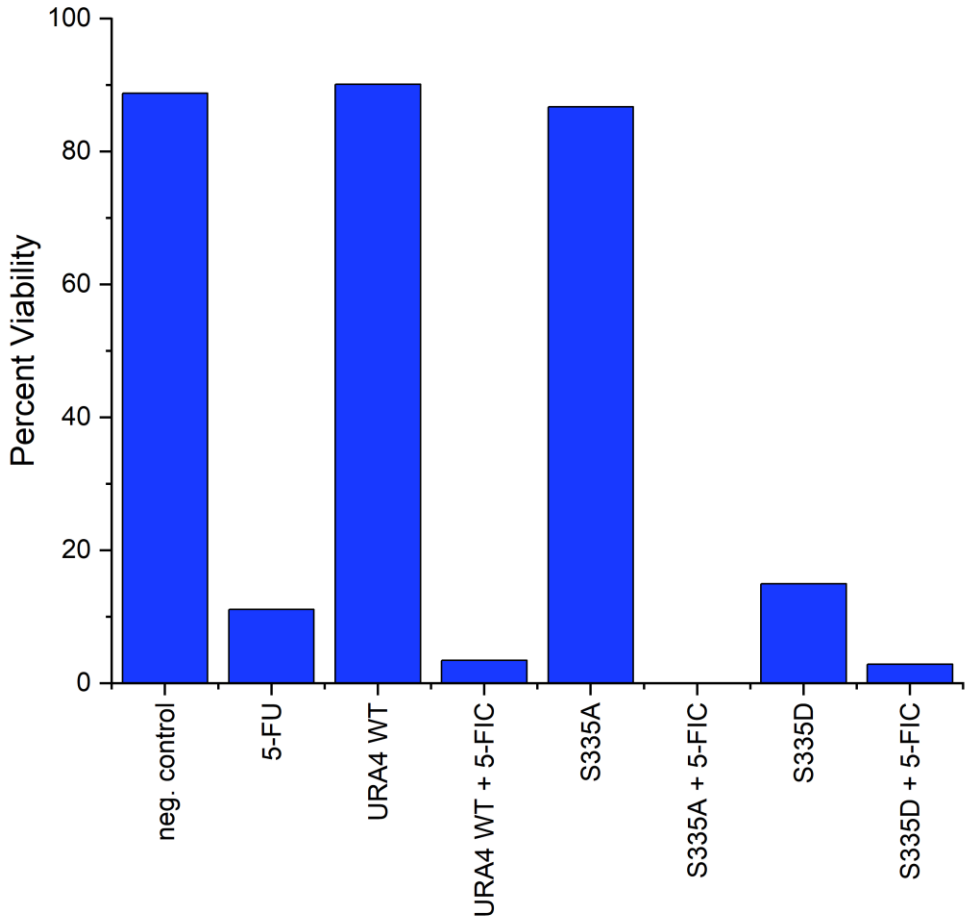

Figure 25: Cell Viability. The addition of 5-FIC to URA4 WT and S335A results in significant cell death.

Conclusions and Future Work

URA4 is an enzyme in the AHS superfamily that is able to deaminate both ISOC and 5-FIC, producing uracil and 5-FU respectively. Due to URA4's ability to deaminate 5-FIC producing the highly toxic 5-FU, it could potentially be used to kill cancer cells in GDEPT. 5-FIC has been shown in previous works to be nontoxic to mammalian cells when ICDS are not present. The main focus of this work has been to structurally characterize URA4 WT and introduce mutations to increase enzyme activity. The crystal structure of URA4 WT was resolved at 2.4 Å. The crystal structure of URA4 shares many of the characteristics found in other enzymes in the AHS

superfamily. Including a $(\beta/\alpha)_8$ barrel forming the active site with a divalent cation. The catalytic mechanism for URA4 likely follows the same mechanism as others in the AHS superfamily.

Ten point mutations were studied, resulting in two mutations that maintained catalytic activity. The eight inactive mutants provide more information on URA4's active site. E246 is a necessary amino acid for the deamination of isocytosine. A catalytic glutamic acid is highly conserved among the enzymes in the AHS superfamily, and by mutating the glutamic acid to an aspartic acid, substrate binding was lost. The mutations made at G332 also resulted in loss of kinetic activity likely due to the mutations being near D331. A larger amino acid at that loop may shift D331, preventing D331 from participating in the catalytic mechanism.

The point mutations were studied to determine if substrate binding was occurring. The only inactive mutant that also binds to substrate was G332N. The change in enthalpy for G332N was -194.0 µcal. The G332N mutation allows for the substrate to bind to the protein but prevents catalysis from occurring. The asparagine mutation likely hydrogen bonds with isocytosine but shifts D331 preventing coordination of the water molecule needed for catalysis. More work is needed to determine the binding affinity of G332N to the substrate.

The kinetic parameters of URA4 and the mutants were determined via an absorbance method and ITC method. The absorbance method prevented the use of higher concentrations of substrate due to the instrument being unable to detect the absorbance in the linear range of detection. The ITC method allowed for higher concentrations of substrate to be used, resulting in more accurate kinetic parameters. The catalytic efficiency via the ITC method of URA4, S335A, and S335D are $3.82 \times 10^4 \pm 4.37 \times 10^3$ M^{-1} sec^{-1}, $3.73 \times 10^4 \pm 4.74 \times 10^3$ M^{-1} sec^{-1}, and $3.05 \times 10^5 \pm 5.83 \times 10^4$ M^{-1} sec^{-1} respectively. S335D mutation resulted in an increase in kinetic

activity. The increase in kinetic activity might be a result of the aspartic acid mutation being able to hydrogen bond with the substrate.

The catalytic efficiency of URA4 WT with 5-FIC as the substrate was $5.62 \times 10^5 \pm 1.78 \times 10^5$ M^{-1} sec^{-1}, while the catalytic efficiency of URA4 WT with ISOC was $3.82 \times 10^4 \pm 4.37 \times 10^3$ M^{-1} sec^{-1}. URA4 WT has an increased substrate specificity toward 5-FIC. More work is needed to determine the kinetic parameters of S335A and S335D with 5-FIC as the substrate.

URA4 WT, S335A, and S335D were studied in a cell viability assay. The protein and 5-FIC were added to mammalian cells, and the percent cell viability was determined after five days. URA4 WT and S335A with 5-FIC resulted in similar cell death to the well containing 5-FU. Further cell assays are needed to confirm the results.

References

1. Siegel RL, Miller KD, Fuchs HE, Jemal A. Cancer statistics, 2022. *CA Cancer J Clin*. 2022;72(1):7-33. doi:10.3322/caac.21708

2. Hanahan D, Weinberg RA. Hallmarks of cancer: the next generation. *Cell*. 2011;144(5):646-674. doi:10.1016/j.cell.2011.02.013

3. Walter M, Herr P. Re-Discovery of Pyrimidine Salvage as Target in Cancer Therapy. *Cells*. 2022;11(4). doi:10.3390/cells11040739

4. Wang W, Cui J, Ma H, Lu W, Huang J. Targeting pyrimidine metabolism in the era of precision cancer medicine. *Front Oncol*. 2021;11:684961. doi:10.3389/fonc.2021.684961

5. Young JD, Yao SYM, Baldwin JM, Cass CE, Baldwin SA. The human concentrative and equilibrative nucleoside transporter families, SLC28 and SLC29. *Mol Aspects Med*. 2013;34(2-3):529-547. doi:10.1016/j.mam.2012.05.007

6. Lu F, Li S, Jiang Y, et al. Structure and mechanism of the uracil transporter UraA. *Nature*. 2011;472(7342):243-246. doi:10.1038/nature09885

7. Longley DB, Harkin DP, Johnston PG. 5-fluorouracil: mechanisms of action and clinical strategies. *Nat Rev Cancer*. 2003;3(5):330-338. doi:10.1038/nrc1074

8. Zeien J, Qiu W, Triay M, et al. Clinical implications of chemotherapeutic agent organ toxicity on perioperative care. *Biomed Pharmacother*. 2022;146:112503. doi:10.1016/j.biopha.2021.112503

9. Zhang J, Kale V, Chen M. Gene-directed enzyme prodrug therapy. *AAPS J*. 2015;17(1):102-110. doi:10.1208/s12248-014-9675-7

10. Ireton GC, McDermott G, Black ME, Stoddard BL. The structure of Escherichia coli cytosine deaminase. *J Mol Biol*. 2002;315(4):687-697. doi:10.1006/jmbi.2001.5277

11. Portnow J, Synold TW, Badie B, et al. Neural Stem Cell-Based Anticancer Gene Therapy: A First-in-Human Study in Recurrent High-Grade Glioma Patients. *Clin Cancer Res*. 2017;23(12):2951-2960. doi:10.1158/1078-0432.CCR-16-1518

12. Freytag SO, Khil M, Stricker H, et al. Phase I study of replication-competent adenovirus-mediated double suicide gene therapy for the treatment of locally recurrent prostate cancer. *Cancer Res*. 2002;62(17):4968-4976.

13. Stolworthy TS, Korkegian AM, Willmon CL, et al. Yeast cytosine deaminase mutants with increased thermostability impart sensitivity to 5-fluorocytosine. *J Mol Biol*. 2008;377(3):854-869. doi:10.1016/j.jmb.2008.01.002

14. Mahan SD, Ireton GC, Knoeber C, Stoddard BL, Black ME. Random mutagenesis and selection of Escherichia coli cytosine deaminase for cancer gene therapy. *Protein Eng Des Sel*. 2004;17(8):625-633. doi:10.1093/protein/gzh074

15. Aučynaitė A, Rutkienė R, Tauraitė D, Meškys R, Urbonavičius J. Discovery of Bacterial Deaminases That Convert 5-Fluoroisocytosine Into 5-Fluorouracil. *Front Microbiol*. 2018;9:2375. doi:10.3389/fmicb.2018.02375

16. Fuchita M, Ardiani A, Zhao L, Serve K, Stoddard BL, Black ME. Bacterial cytosine deaminase mutants created by molecular engineering show improved 5-fluorocytosine-mediated cell killing in vitro and in vivo. *Cancer Res*. 2009;69(11):4791-4799. doi:10.1158/0008-5472.CAN-09-0615

17. Wang Y, Wang G, Moitessier N, Mittermaier AK. Enzyme kinetics by isothermal titration calorimetry: allostery, inhibition, and dynamics. *Front Mol Biosci*. 2020;7:583826. doi:10.3389/fmolb.2020.583826

18. Müller-Esparza H, Osorio-Valeriano M, Steube N, Thanbichler M, Randau L. Bio-Layer Interferometry Analysis of the Target Binding Activity of CRISPR-Cas Effector Complexes. *Front Mol Biosci*. 2020;7:98. doi:10.3389/fmolb.2020.00098

19. Brünger AT. Free R value: a novel statistical quantity for assessing the accuracy of crystal structures. *Nature*. 1992;355(6359):472-475. doi:10.1038/355472a0

20. Chen VB, Arendall WB, Headd JJ, et al. MolProbity: all-atom structure validation for macromolecular crystallography. *Acta Crystallogr D Biol Crystallogr*. 2010;66(Pt 1):12-21. doi:10.1107/S0907444909042073

21. Trott O, Olson AJ. AutoDock Vina: improving the speed and accuracy of docking with a new scoring function, efficient optimization, and multithreading. *J Comput Chem*. 2010;31(2):455-461. doi:10.1002/jcc.21334

22. Manta B, Raushel FM, Himo F. Reaction mechanism of zinc-dependent cytosine deaminase from Escherichia coli: a quantum-chemical study. *J Phys Chem B*. 2014;118(21):5644-5652. doi:10.1021/jp501228s

23. Kazlauskas A, Darinskas A, Meškys R, Tamašauskas A, Urbonavičius J. Isocytosine deaminase Vcz as a novel tool for the prodrug cancer therapy. *BMC Cancer*. 2019;19(1):197. doi:10.1186/s12885-019-5409-7