UC San Diego

UC San Diego Electronic Theses and Dissertations

Title

A Kinetic Investigation of Molecular Mechanisms Underlying IDH1 and hPol "ε" Catalytic Activity

Permalink

https://escholarship.org/uc/item/6h31x5j7

Author

Luna, Lucas

Publication Date

2021

Peer reviewed|Thesis/dissertation

UNIVERSITY OF CALIFORNIA SAN DIEGO

SAN DIEGO STATE UNIVERSITY

A Kinetic Investigation of Molecular Mechanisms Underlying IDH1 and Human Pol ε

Catalytic Activity

A dissertation submitted in partial satisfaction of the requirements for the degree

Doctor of Philosophy

in

Chemistry

by

Lucas Luna

Committee in Charge:

    University of California San Diego

        Professor Elizabeth Komives
        Professor Kimberly Prather

    San Diego State University

        Professor Christal Sohl, Chair
        Professor Anca Segall
        Professor Manal Swairjo

2021

This dissertation of Lucas Luna is approved, and it is acceptable in quality and form for publication on microfilm and electronically:

_____

_____

_____

_____

_____

Chair

University of California San Diego

San Diego State University

2021

DEDICATION

Dedicated to little Ozzy and Qui-gon because they know all the answers already

TABLE OF CONTENTS

LIST OF ABBREVIATIONS

AML	Acute Myeloid Leukemia
TCA Cycle	Tricarboxylic Acid Cycle
GBM	Glioblastoma Multiforme
k_{cat}	Turnover Rate Constant
K_m	Michaelis-Menten Constant
k_{cat}/K_m	Catalytic Efficiency
˚C	Degree Celsius
mM	Millimolar
T_m	Melting Temperature
nM	Nanomolar
µM	Micromolar
FDA	Food and Drug Administration
Ni-NTA	Nickel Nitrilotriacetic Acid
psi	Pounds per square inch
TB	Terrific Broth
ND	Not Determined
Å	Angstrom
IC_{50}	Inhibitory Concentration at 50%
WT	Wild Type
MS	Mass Spectrometry
GC	Gas Chromatography
SEC	Size Exclusion Chromatography

K_d	Dissociation Constant
Da	Dalton
PDB	Protein Data Bank
DNA	Deoxyribose Nucleic Acid
ITC	Isothermal Calorimetry
NMR	Nuclear Magnetic Resonance
HDX-MS	Hydrogen/Deuterium Exchange – Mass Spectrometry
LB	Luria Broth
k_{obs}	Observed Rate

Small Molecules

ICT	Isocitrate
αKG	α-ketoglutarate
D2HG	D-2-Hydroxyglutarate
NAD^+	Nicotinamide Adenine Dinucleotide Oxidized
NADH	Nicotinamide Adenine Dinucleotide Reduced
$NADP^+$	Nicotinamide Adenine Dinucleotide Oxidized
NADPH	Nicotinamide Adenine Dinucleotide Reduced
DTT	Dithiothreitol
IPTG	1-thiol-β-D-galactopyranoside
β-ME	β-mercaptoethanol
DMSO	Dimethyl sulfoxide
PEI	Polyethyleneimine
dNTP	Deoxynucleotide Triphosphate

dTTP	Deoxythymidine Triphosphate
dATP	Deoxyadenosine Triphosphate
dNTPαS	Deoxynucleotide-5'-(α-thio)-triphosphate
T	Thymine
A	Adenine
C	Cytosine
G	Guanine
ESOM	Esomeprazole Sodium Salt
Proteins	
IDH1	Isocitrate Dehydrogenase 1
IDH2	Isocitrate Dehydrogenase 2
IDH3	Isocitrate Dehydrogenase 3
Pol ε	DNA Polymerase Epsilon
yPol ε	Yeast DNA Polymerase Epsilon
hPol ε	Human DNA Polymerase Epsilon
Pol δ	DNA Polymerase Delta
Pol α	DNA Polymerase Alpha
RB69	Bacteriophage RB69 DNA Polymerase
T7	T7 Bacteriophage DNA Polymerase
GDH1	Glutamate Dehydrogenase 1
GDH2	Glutamate Dehydrogenase 2
αKGDH	α-Ketoglutarate Dehydrogenase

BSA   Bovine Serum Albumin

FBS   Fetal Bovine Serum

LIST OF FIGURES

Chapter 3 Figures

LIST OF TABLES

ACKNOWLEDGEMENTS

VITA

Education

2013   B.S Biochemistry
               University of California Santa Barbara

2021   Ph.D. Chemistry
               University of California San Diego and San Diego State University

Employment History

2014-2015   Genentech: A Member of the Roche Group
               South San Francisco, CA

2015   NASA Ames Research Center
               Moffett Field, CA

2015-2021   Teaching Assistant
               San Diego State University San Diego, CA

Select Publications

Luna, Lucas A., Lesecq, Zach, White, Katharine A., Hoang, An, Scott, David A., Zagnitko, Olga, Bobkov, Andrey A., Barber, Diane L., Schiffer, Jamie M., Isom, Dan G., Sohl, Christal D. 2020 "An Acidic Residue Buried in the Dimer Interface of Isocitrate Dehydrogenase 1 (IDH) Helps Regulate Catalysis and pH Sensitivity." *Biochemical Journal,* **2020**, 477(16), 2099-3018.

Avellaneda Matteo, Diego, Wells, Grace A., **Luna, Lucas A**., Grunseth, Adam J., Zagnitko, Olga, Scott, Scott, David A., Hoang, An, Luthra, Amit, Swairjo, Manal, Schiffer Jamie M., Sohl, Christal D. "Inhibitor potency varies widely among tumor-relevant human isocitrate dehyrodgenase 1 Mutants." *Biochemical Journal,* **2018***,* 475 (20): 3221-3238.

Select Awards and Fellowships

2021   Rees-Stealy Research Foundation Fellowship

2019   University Graduate Fellowship

2018, 2019   Prebys Biochemical Research Endowed Scholarship

2018   Achievement Rewards for College Scientists (ARCS) Foundation A

ABSTRACT OF DISSERTATION

A Kinetic Investigation of Molecular Mechanisms Underlying IDH1 and Human Pol ε

Catalytic Activity

by

Lucas Luna

Doctor of Philosophy in Chemistry

University of California San Diego, 2021

San Diego State University, 2021

Professor Christal D. Sohl, Chair

Malignant tumors are distinguished from normal cells by a number of hallmarks, including metabolic reprogramming and genome instability. Metabolic reprogramming in cancer cells was first observed by Otto Warburg in the early 20th century. Recently metabolic enzymes have been shown to be drivers of cancer. Isocitrate dehydrogenase 1 (IDH1) is a metabolic enzyme that has altered activity in cancer and can drive tumorigenesis. Mutations in IDH1 are frequently observed in low grade gliomas and

secondary glioblastomas. However, wild-type IDH1 is also overexpressed in primary glioblastomas and has been linked to the growth of these tumor types. The products of the IDH1 reaction are essential for many metabolic processes. Changes in the cellular environment, such as pH fluctuations, substrate concentration levels, and oxygen levels can reroute metabolism by modifying the activity of metabolic enzymes like IDH1. Amino acid residues that sense changes in pH will typically have a shifted pK_a towards physiological pH levels. The change in pK_a stores the potential energy required to drive a structural/functional modification that alters catalytic activity. A goal of this dissertation was to identify and characterize pH-dependent activity in IDH1 and provide a mechanism for how it sensed changes in pH to regulate its catalytic activity.

Accurate genome replication is essential to ensure the survival of offspring. Considering the size of the human genome and its constant exposure to environmental and endogenous damage, this is not an easy task, and is further complicated by a damage associated with cancer. A number of DNA polymerases have evolved to handle DNA synthesis, repair, and overall genome maintenance. DNA Polymerase ε is responsible for the highly accurate and processive DNA replication on the leading strand. The high-fidelity of DNA Polymerase ε is maintained through a balance between its incorporation and proofreading activities. Mutations altering either of these abilities are frequently observed in endometrial cancer. The second goal of this thesis was to investigate mechanisms of (in)fidelity of polymerase domain mutants in DNA Polymerase ε.

1. Introduction

1.1 What is Cancer

Cancer is the uncontrolled and unregulated growth that arises from an accumulation of mutations in the genome and ranks as the second leading cause of death globally.[1,2] The tumor itself is very complex, as individual cells within the tumor population possess physical and genetic variations that give rise to heterogeneity within the tumor cell population.[3,4] The differences between cancer cells can include changes in extracellular pH, intracellular pH, oxygen levels, substrate concentration levels, protein activity, protein expression levels, possible genetic differences, and other changes that promote growth and metastasis.[5-7] Tumor heterogeneity also makes treatment challenging because individuals of the tumor population cells may respond differently or be treatment-resistant, which can lead to recurrence and metastasis. Despite tumor cell heterogeneity, researchers have observed several distinguishable characteristics common in all cancers.

1.2 Hallmarks of Cancer

The six hallmarks of cancer in Hanahan and Weinberg's first groundbreaking review were described as production of self-sufficient growth signals, insensitivity to anti-growth signals, limitless replicative potential, sustained angiogenesis, evading apoptosis, and tissue invasion and metastasis.[8] As normal cells grow and divide, neoplastic cells need to adopt these traits in order to ultimately transition to malignant cells. In 2011, the number of hallmarks has expanded to include metabolic reprogramming and evading the immune system.[9] Underscoring these hallmarks are genome instability and inflammation, which expedites the acquisition of these traits and promotes the function of these

hallmarks.[9] Together, these hallmarks involve disruption or ablation of normal cellular pathways and normal protein-protein or protein-ligand interactions in order to satisfy the energetic needs of the cancer cells and drive growth and metastasis.[9] Elucidating the mechanisms of how cancer cells alter normal cellular metabolism and protein function widens the knowledge on cancer biology and can identify new therapeutic targets and pathways.

Tumor sequencing efforts have identified enzymes playing significant roles in altered metabolic pathways and genome instability. Isocitrate Dehydrogenase 1 (IDH1) is important metabolic enzyme whose catalytic activity may be regulated by the cellular environment. DNA Polymerase Epsilon (Pol ε) is an important replicative polymerase that is responsible for maintaining the integrity of the eukaryotic genome. IDH1 and Pol ε will be discussed in further in this dissertation.

1.2.1 Cancer Metabolism

Cells utilize energy largely in the form of ATP, aptly dubbed the energy currency of the cell. They typically gain their energy by metabolizing one molecule of glucose in the cytoplasm to two molecules of pyruvate, two molecules of ATP, and two molecules of NADH.   Under anaerobic conditions, pyruvate is converted lactate by lactate dehydrogenase to regenerate NAD^+ for glycolysis. Under aerobic conditions, pyruvate can be converted to acetyl-CoA and transferred to the mitochondria where the tricarboxylic acid (TCA) cycle occurs. This process produces the high energy molecules NADH and $FADH_2$, while expelling CO_2. The electrons stored in NADH and $FADH_2$ are then carried through a series of redox reactions in the electron transport chain to generate

2

a proton gradient that drives the formation of ATP, and molecular oxygen serves as the final electron acceptor in this respiratory cascade and is reduced to water.

Otto Warburg noticed a shift in cellular metabolism in cancer cells. Cancer cells tend to utilize higher levels of glucose and produce lactate even under aerobic conditions.[10] Aerobic glycolysis, known as the Warburg Effect, is observed in various cancer types including colorectal cancer, lung cancer, and glioblastomas.[11-14] Warburg postulated that cancer cells must have dysfunctional mitochondria; however, more recent studies show that cancer cells usually have active and functional mitochondria.[15, 16] Interestingly, studies have shown that different tumor types, and even subpopulations within a tumor, have different metabolic alterations.[17] A tumor is likely to display metabolic flexibility, and a dynamic interplay between glycolysis and oxidative metabolism is observed in various cancers that is dependent on the real-time needs of the cell.[17]

An upregulation of glycolysis will demand an increased supply of NAD^+ to sustain itself. To fulfill this need, the cell typically overexpresses and increases lactate dehydrogenase activity to maintain NAD^+ pools. Interestingly, other metabolic pathways and enzymes have been linked to alterations in enzymatic activity. For example, under hypoxia, the low oxygen levels down-regulate the pyruvate dehydrogenase complex so cells will produce acetyl-CoA for the biosynthesis of palmitic acid from glutaminolysis instead of glycolysis.[17, 18] This process involves the coordination of many metabolic enzymes, including stabilizing HIF1-α and isocitrate dehydrogenase 1 (IDH1) that ultimately cause the TCA cycle to run in a reductive direction.[17, 18] When isocitrate dehydrogenase 1 (IDH1) encounters a change in pH or oxygen levels, or obtains a point mutation, a shift in the equilibrium position of its chemical reaction may occur. Mutations

in IDH1 produce products that may stabilize HIF1-α by inhibiting HIF1-α hydroxylases to support this altered metabolism.[20] IDH1 activity will be discussed further in this dissertation.

1.2.2. Tumor Invasion and Changes in the Cellular pH Gradient

During tumor development, the primary tumor may acquire adaptations necessary to travel to distant sites and invade other tissues. The distant settlement of tumor cells, or metastasis, is the cause of 90% of cancer deaths.[21] Invasion and metastasis are very complex processes, and their genetic and biochemical determinants remain incompletely understood.[8] Interestingly, a reversed cellular pH-gradient is observed in metastatic cancer cells and serves as a characteristic feature of most cancer types.[22] Normal differentiated adult cells will have an intracellular pH of ~7.2 and an extracellular pH of ~7.4.[22] Cancer cells will have this pH gradient reversed, with an intracellular pH ≥ 7.4 and an extracellular pH of 6.7 – 7.1.[22] The reversed pH gradient in maintained by changes metabolism and protein expression, and can create a perfect storm for metastatic progression.[22] An increased intracellular pH is required for cell growth and division, and an acidic extracellular pH promotes the degradation of the extracellular matrix that facilitates tumor migration.[22]

1.2.3 Role of pH in Regulation of Protein Activity

Changes in the cellular environment such as oxidative stress, substrate concentration levels, and pH can affect protein activity. Protein protonation is an often-overlooked modification that can affect protein-protein and protein-ligand interactions. Proteins that serve as pH sensors can sense changes in surrounding pH by changing the protonation state of amino acid residues, resulting in altered activity.[23] It is thought that

cancer cells can utilize the change in pH gradient to regulate protein activity. For example, glycolysis is promoted at alkaline intracellular pH levels and is mediated by pH-sensitive activity of rate-limiting glycolytic enzymes like phosphofructokinase-1.[22-26] An increased intracellular pH can also induce changes in the expression or localization of glycolytic enzymes, forming local acidic regions within the cytoplasm.[22] In these acidic regions, lactate dehydrogenase activity can be increased and replenish glycolysis with NAD^+, while the acidic products are transported out of the cell through the monocarboxylate transporter in the cell membrane.[22, 27] Depending on the needs of the cell and environmental conditions, cancer cells can produce acetyl-CoA through glutaminolysis, which also produces NADPH for biosynthetic pathways.[28]

Protein sensors can sense the change in surrounding pH and correspondingly change their activity or structure by mechanisms such as protonating buried ionizable amino acid residues.[22] A histidine residue is a natural candidate for modulating protein activity based on protonation state since its pK_a value is already close to a neutral pH, though any ionizable residue (glutamate, aspartate, cysteine, lysine, arginine) can exhibit a shifted pK_a value depending on its location and surrounding environment within the folded protein.[22, 23] The shift in pK_a value then provides the necessary free energy to drive a pH-dependent change in conformation and/or activity. Biophysical calculations determining pK_a values of amino acid residues within proteins are complicated by the folding of the protein that affects solvent accessibility, as well as the polarizability of nearby ionizable residues. Therefore, potential pH-sensing residues can be found using structural informatics algorithms. These algorithms account for spatial arrangements and protein topology. Through a collaboration with Prof. Dan Isom, we used an algorithm

called pHinder.[30]   This algorithm uses Delaunay Triangulation to locate non-redundant ionizable networks in proteins and calculates a surface to determine whether the amino acid is buried[30]   The results of the structural informatics analysis is then coupled to further biochemical and biophysical characterization of protein activity.[29-31]   The pH-dependent activity of IDH1 was assessed in this dissertation using steady-state kinetics and coupled with the pHinder algorithm to locate potential pH sensing residues.

1.3 Incidence of IDH1 in Cancer

Gliomas are a common form of brain cancer, with glioblastoma multiforme (GBM) serving as the most aggressive and lethal form of glioma.[37]   Extensive tumor sequencing efforts found that 80% of glioma patients and 88% of secondary glioblastoma patients expressed a mutation in IDH1 at position R132.[37]   The most common point mutation in IDH1 is R132H, followed by R132C, R132S, R132G, R132L, R132V.[38]   IDH1 R132H mutations are also observed in acute myeloid leukemias (AML), chondrosarcomas, and cholangiosarcomas.[39, 40]

WT IDH1 has also been observed to drive tumors.[41]   *In silico* analysis of the specimens profiled by The Cancer Genome Atlas (TCGA) revealed that WT IDH1 had elevated expression levels in 65% of primary glioblastoma as well as increased WT IDH1 mRNA levels.[41]   Consequently, these observations were reproduced in an investigation of glioblastoma tumor samples by Calvert and colleagues.[41]   They also reported that the WT IDH1 reactions play a role in regulating cellular processes such as lipid biosynthesis, redox homeostasis, and cellular differentiation by promoting tumor growth in cancer cells and decreasing survival times in mice with patient-derived xenografts.[38]   The high

frequency of WT IDH1 and various R132 IDH1 mutations has become a critical area of study in cancer research.

1.4 Structure of IDH1

The first structure of IDH1 showed that the protein exists as a dimer, with 414 residues per monomer (Figure 1).[42] The structure of IDH1 is mostly α-helical with few β-sheets.[42] The residues that form the active the site and interact with isocitrate are T77, S94, R100, R109, R132, Y139, and D275 in one monomer and K212', T214', and D252' in the other monomer.[42] The IDH1 structure features a regulatory domain spanning from residue from N271to G286, termed the α-10 helix.[42] D279 of this domain prevents isocitrate from binding by a combination of electrostatic repulsion and steric hinderance in the open, inactive state of IDH1 by forming hydrogen bonds with S94, T77, and N96.[42]
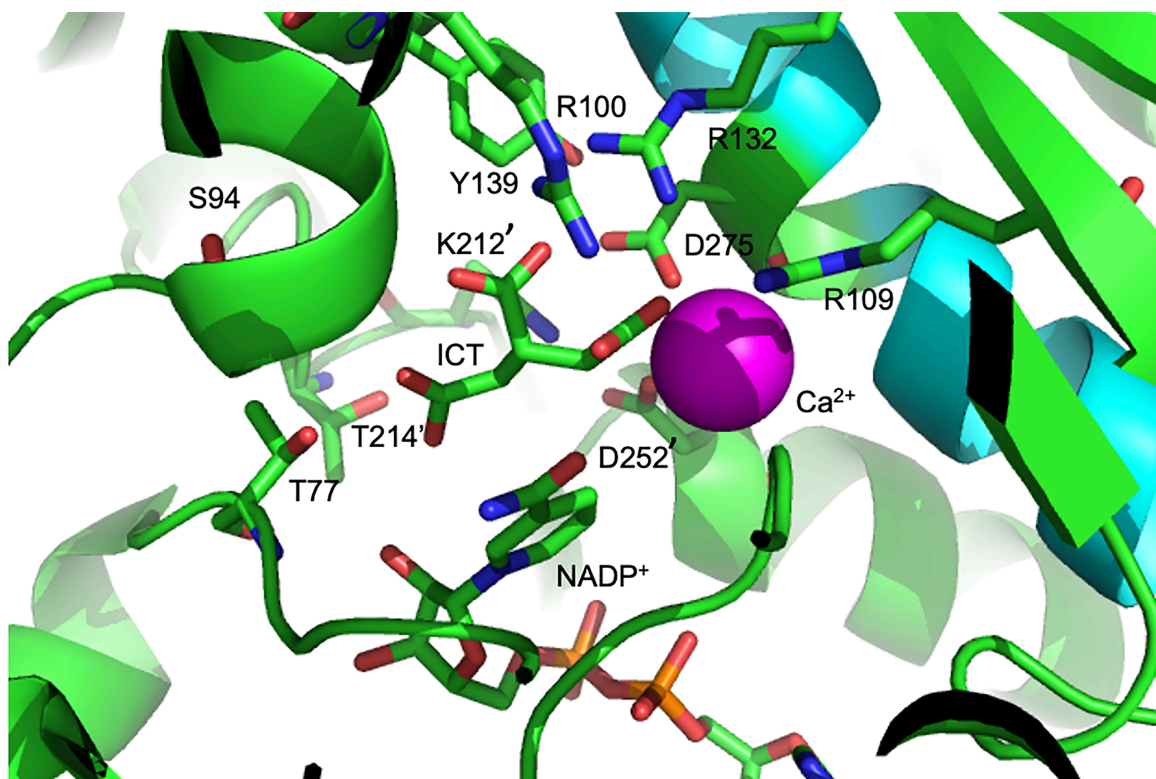
Figure 1: Zoomed in view of the IDH1 active site, with the α-10 shown in cyan (PDB code 1T0L).[42] Active site residues are labeled.

When the concentration of isocitrate and magnesium reach a certain level, the interaction between D279 and S94 is broken, opening the active site to catalysis.[42] This mode of competitive binding between isocitrate and magnesium with D279 triggers the refolding of the regulatory domain into an α-helix and a conformational change of the overall structure of IDH1 to the active, closed conformation occurs, followed by a reopening after chemistry.[42] The newly formed favorable interactions between isocitrate and active site residues likely compensate for any energy costs.[42]

1.4.1 IDH1 Activity and Function

The family of isocitrate dehydrogenase enzymes consists of three isoforms: IDH1, IDH2, and IDH3. All three catalyze the conversion of isocitrate to α-ketoglutarate (αKG) (Figure 2). IDH3 is a heterotetramer and catalyzes this reaction irreversibly, reducing the
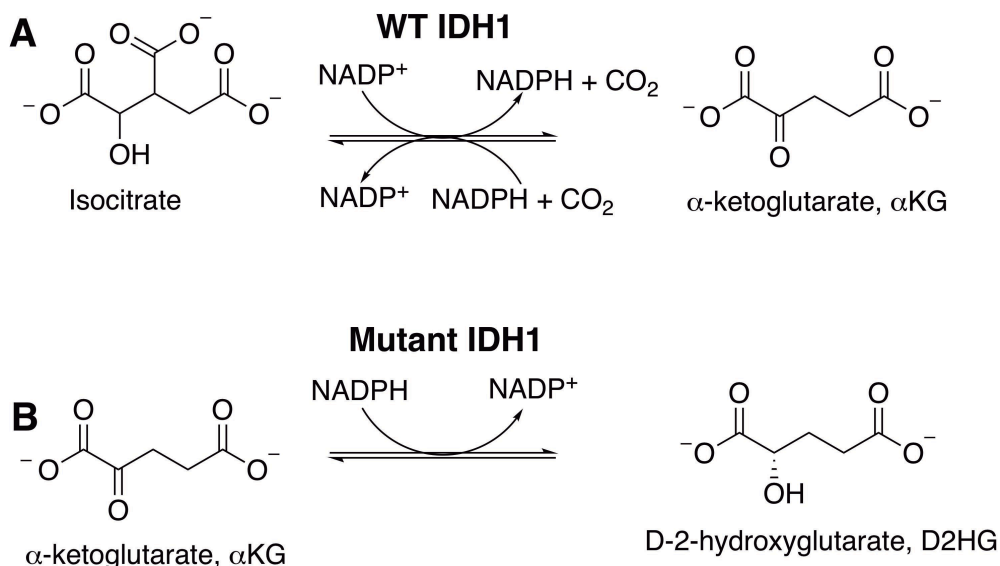
Figure 2: A. Reversible oxidative decarboxylation of isocitrate to αKG producing NADPH and CO_2. This is the reaction performed by WT IDH1. **B**. Reduction of αKG to D2HG depleting NADPH. This is the reaction performed by mutant IDH1.

cofactor NAD^+ to NADH in the process, and is localized to the mitochondria where it plays a critical role in the TCA cycle. Likely, given its role in the TCA cycle, it has not been found to be mutated in cancer. IDH2 is also found in the mitochondria and catalyzes the

same chemistry; however, unlike IDH3, it is a dimer, reduces $NADP^+$ to NADPH, and has been linked to cancer.[37] Similarly, IDH1 catalyzes the same reaction with an $NADP^+$ as a cofactor and is linked to cancer but is found in the cytosol and peroxisomes of the cell.[39-42] IDH2 and IDH3 will not be discussed in detail.

The forward and reverse reaction of IDH1 produces critical intermediates for metabolic pathways and gene regulation. The reverse reaction produces isocitrate that is important for glutamine metabolism; however, under normal cellular conditions the forward reaction is favored.[18, 19, 43] The NADPH produced is used in lipid and nucleic acid biosynthesis, as well as in regenerating glutathione to mitigate reactive oxygen species.[20, 44-46] The αKG produced by the forward reaction can help replenish the TCA cycle to generate ATP in the electron transport chain, combat reactive oxygen species, and is a cofactor used to signal HIF1-α degradation.[20] Furthermore, αKG-dependent enzymes are involved in DNA repair and DNA demethylation.[47,48] Normal activity is lost when IDH1 is mutated at position R132 in cancer, and this originally led researchers to consider mutant IDH1 to be a tumor suppressor.[49] Surprisingly, not only does R132 IDH1 lose normal activity, it catalyzes a neomorphic reaction: the NADPH dependent conversion of αKG to D-2-hydroxyglutarate (D2HG).[50] An oncometabolite is a small molecule metabolic product whose accumulation deregulates normal cellular metabolism and consequently primes cells for progression to cancer. D2HG is classified as an oncometabolite because it inhibits αKG-dependent histone demethylases, ten eleven translocation (TET) methylcytosine dioxygenases, and can competitively inhibit HIF1-α hydroxylases.[20, 51, 52] As a result, patients harboring IDH1 mutations have hypermethylated DNA that leads to cell de-differentiation.[52]

1.4.2 Selective Inhibition of IDH1 Mutants

Anti-cancer therapies must be highly specific for their target to minimize toxicity. As such, inhibitors for mutant IDH are highly specific, exhibiting poor binding to WT IDH1 and WT IDH2.[53, 54] Agios Pharmaceuticals has developed two FDA approved selective inhibitors that reduce the amount of D2HG produced by mutant IDH1 and IDH2, allowing the cell to resume differentiation pathways.[55, 56] Since D2HG levels are reduced, αKG dependent histone demethylases can resume their proper function in regulating gene expression.[57] Other commercially available compounds exist that have a similar binding mechanism as the FDA approved drugs, including ML309 and AGI-5198 (Figure 3).[53, 57] The specificity of these inhibitors for mutant IDH1 over WT IDH1 is thought to take advantage of the structural dynamics of the regulatory α10 helix.[58] A structural investigation of selective inhibitors and mutant IDH1 showed that mutations at the R132 position interrupt a crucial interaction between R132 and N271 that prevents folding of the helix and open a plastic allosteric binding site for inhibtiors.[58] However, the full mechanism is still under investigation.
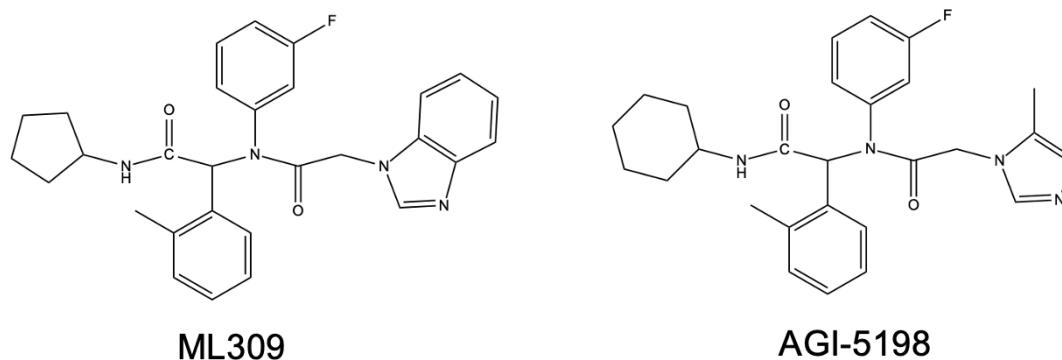
ML309   AGI-5198

Figure 3: Structure of ML309 (left). This inhibitor has the same scaffold as of AGI-5198 (right).

1.5 Genome Instability

The extraordinary accuracy of replicative polymerases combined with their ability to repair and maintain the human genome ensures a low rate of spontaneous mutations during cells division. Obtaining the required hallmarks of cancer depends largely on the accumulation of genomic alterations and is achieved through increased exposure to mutagens and/or a breakdown in one or more genome maintenance pathways.[9] The defects in genome maintenance genes, referred to as "caretaker genes", include repairing damaged DNA. Some of the most striking examples of defective caretaker genes responsible for a loss of genome stability in cancer are tumor protein 53 (TP53) and telomerase.[9] Mutations in replicative and repair DNA polymerases can also be included in this category, as they play active roles in replicating and repairing the genome as well. For example, DNA polymerase Epsilon (Pol ε), a eukaryotic DNA polymerase responsible for DNA replication on the leading strand, is frequently mutated in colorectal and endometrial cancer that causes a decrease in replicative accuracy and drives tumorigenesis.[32-36, 59, 60] Pol ε activity, mutations, and mechanisms of (in)fidelity will be discussed further in this dissertation.

1.6 Structure of DNA Polymerase Epsilon

DNA polymerases can be divided into seven families (Family A, B, C, D, X, Y, and RT) based on sequence homology, structure, and function. Family B consists of replicative DNA polymerases from various domains of life, as well as the eukaryotic replicative polymerases: DNA polymerase Alpha (Pol α), DNA polymerase Delta (Pol δ), and Pol ε.[61, 63, 64] Pol α contains both primase and polymerase activity, and initiates DNA synthesis by synthesizing a short RNA primer and extending the primer by ~ 20 bases

with its polymerase activity.[62, 63] DNA synthesis is then resumed by Pol δ and Pol ε on the lagging strand and leading strand, respectively.[62]
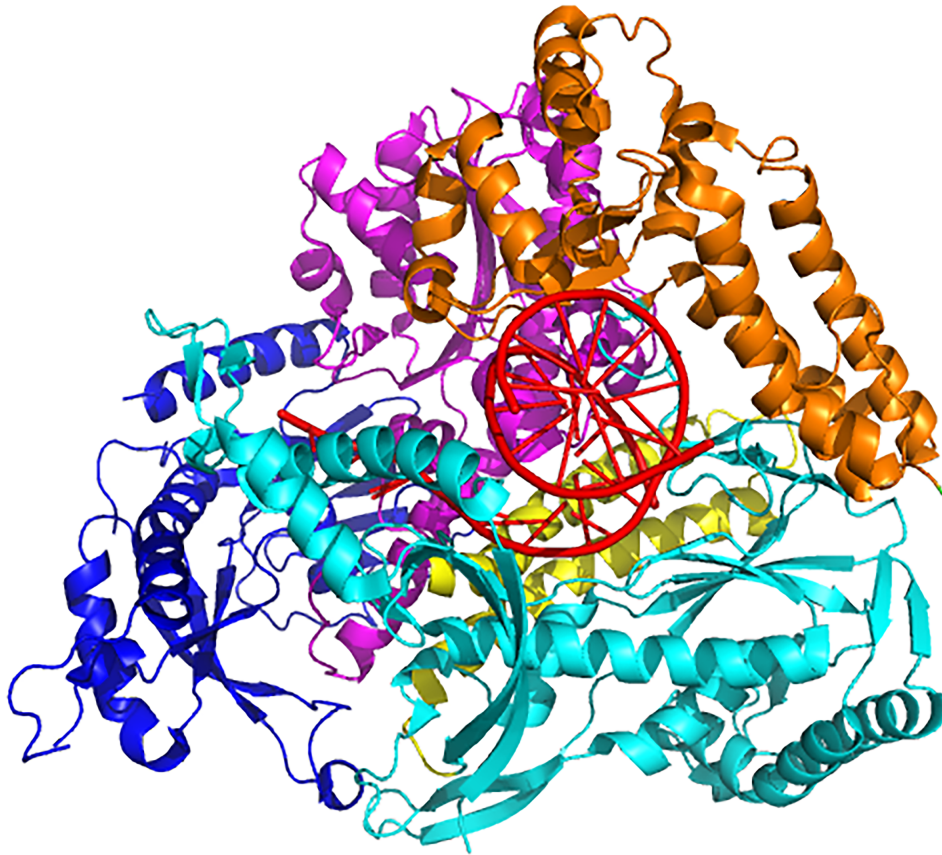
Figure 4: Structure of DNA Polymerase epsilon (PDB code 4M8O) colored by domain.[62] The *N*-terminal domain is colored dark blue, exonuclease domain in magenta, palm domain in cyan, fingers domain in yellow, thumb domain in orange, and DNA is colored red.

Characterization of Pol ε is limited by difficulties in heterologous expression and purification.[65-67] To overcome these issues, the catalytic domain of human Pol ε and yeast Pol ε have been successfully expressed and purified from *E. coli* and yeast, respectively.[61, 63, 64] The catalytic domain of Pol ε is structurally conserved among other B family polymerases and consists of an exonuclease and polymerase domain.[63] The exonuclease domain contains the catalytic residues responsible for proofreading, and the polymerase domain is responsible for incorporating the next dNTP.[63] The polymerase

domain can be divided further into the fingers, palm, and thumb subdomains (Figure 4).[68] The fingers domain undergoes a conformational changes during catalysis from an "open" to a "closed" state that brings the dNTP into position for catalysis.[63] The palm domain, which is highly conserved among all polymerases, binds DNA and contains the catalytic residues responsible for incorporating a dNTP.[68] The thumb domain makes additional contacts with the DNA and is believed to play a role in positioning DNA, translocating the DNA between exonuclease and polymerase active sites, and processivity.[68] Currently, high-resolution crystal structures exist for only the yeast Pol ε catalytic subunit, as well as a lower resolution structure of the yeast heterotetramer solved by cryogenic electron microscopy (cryo-EM).[61, 63, 64, 69]

A major difference between Pol ε and other B-family polymerases is the novel P-domain in within the palm domain of Pol ε.[63] The P domain extends outward from the palm domain toward the primer-template DNA, allowing Pol ε to encircle newly synthesized DNA.[63] This domain may represent a completely novel type of protein fold.[63] Mutagenesis coupled with primer-extension assays in the presence of a heparin trap revealed that the P domain is essential for the high processivity of Pol ε.[63]

1.6.1 Nucleotide Incorporation Mechanism

The model for the dNTP incorporation mechanism was derived from studies on T7 DNA polymerase.[70] Genetic and biochemical analyses of T7 DNA polymerases and related Family A polymerases have clarified the roles of conserved amino acids responsible for DNA, synthesis, nucleotide selection, and fidelity.[71] Features of the active sites of RB69, the DNA polymerase of RB69 bacteriophage, and the mammalian DNA polymerase β, are remarkably similar to Family A type polymerases showing that they

employ closely related strategies for replicating DNA.[71] Thus, T7 DNA polymerase, DNA

polymerase β, and RB69 DNA polymerase and can serve as working models for

constructing the catalytic mechanisms of other polymerases due to sequence similarities

in their highly conserved regions, ease of crystallization, and extensive kinetic studies.[70,][71]
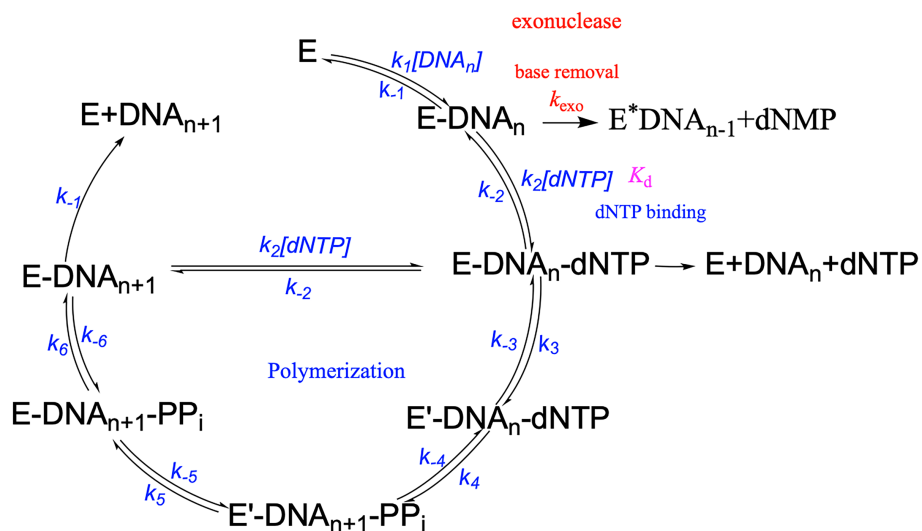
Figure 5: Minimal kinetic mechanism of DNA Polymerase Epsilon.

The pathway for DNA polymerization is shown in Figure 5. First, the polymerase

binds to the DNA tightly to form the binary enzyme-DNA complex. Then, the correct dNTP

binds to form the enzyme-DNA-dNTP ternary complex. dNTP binding leads to a

conformational change from an open state to a closed state that is then followed by the

nucleotide incorporation reaction. The enzyme changes conformation to the open state

and is followed by the release of pyrophosphate and translocation of the polymerase to

the next unpaired nucleotide.[70]

The incorporation reaction proceeds by a nucleophilic attack of the 3'OH of the DNA primer-template on the α-phosphate of the incoming dNTP. The fingers, palm, and thumb domains of the polymerase position the DNA primer-template next to the nucleotide binding site, and the nascent base pair between the incoming dNTP and primer-template fits nicely in the groove between fingers domain and the 3'-end of the DNA primer-template. This reaction uses a universal "two metal ion" mechanism in which the magnesium ions bound in the polymerase active site lower the entropic cost associated with the nucleophilic addition reaction. This is done by properly aligning the incoming dNTP for nucleophilic attack by the 3'-hydroxyl group of the DNA primer-template.[70, 71] However, two metals ions are not sufficient to complete the nucleotide incorporation reaction, and therefore additional catalytic residues are required. Conserved aspartate residues align the metal ions in the proper orientation.[70, 71] In addition, there is no residue that acts as a general base catalyst to deprotonate the hydroxyl group; however, the decreased pK_a and proper orientation of the metal-bound hydroxyl group is sufficient to drive the incorporation reaction.[70] The metal ions and nearby positively charged amino acid residues then stabilize the negative charges on the α-phosphate when the pentavalent intermediate is formed and participate in proton transfer reactions. Then pyrophosphate is released.[68, 70, 71]

1.6.2 Factors Affecting Polymerization Fidelity

A number of electrostatic, steric, and kinetic factors contribute to the nucleotide insertion efficiency and base selectivity of a replicative polymerase. Favorable secondary interactions that affect incorporation efficiency include hydrogen bonds between nucleobases, electrostatic interactions formed between non-catalytic amino acid residues

and the DNA minor groove, and base stacking interactions.[70] The insertion efficiency of

a polymerase is also affected by the geometry and mobility of the nascent base pair in

the polymerase active site.[70] This allows for less restricted movement of the template

base and lowers polymerization accuracy.[70] The accuracy of the incorporation reaction

is affected by a conformational change of the fingers domain and partitioning of the primer

terminus between polymerase and exonuclease active sites.[70] The following description

of the factors are general mechanisms of replicative polymerase fidelity and do not

specifically refer to Pol ϵ.

Pre-steady-state kinetic studies have shown that replicative polymerases insert the

correct nucleotide with a higher enzymatic efficiency compared to non-polar nucleoside

analogues because of the formation of crucial hydrogen bonds in canonical Watson-Crick

base pairing.[72] The formation of hydrogen bonds between the minor groove of the primer-

template DNA and the replicative DNA polymerase are also important for base selectivity

and reaction efficiency because the lone pairs on the nucleobases form crucial

interactions with the enzyme.   Structural and kinetic studies reveal that when

deoxyadenosine is replaced with 3-deaza-2'-deoxyadenosine, which is an adenosine

analogue where the N^3 of the purine ring is replaced with a carbon, enzymatic activity is

decreased.[72] This effect is more dramatic if the adenosine analogue is placed further

upstream in the primer-template strand. In that case, a hydrogen bond between a nearby

water molecule and the N^3 of adenosine is removed such that steric clashes compromise

catalytic efficiency.[72] In addition to hydrogen bonding between nucleobases, stacking

interactions between the aromatic purine and pyrimidine rings also stabilize the DNA

double helix.[72]   When the template base is replaced with tetrahydrofuran (THF) to

generate an abasic site, bases can no longer form hydrogen bonds. The efficiency and fidelity depend largely on base stacking interactions between the incoming dNTP and the penultimate base pair.[72] As a rule of thumb, a replicative polymerase will insert a dATP most efficiently opposite this abasic site to gain the most favorable stacking interactions.[72.] Strict adherence to this rule varies amongst polymerases and also depends on the composition of the DNA sequence preceding the abasic site.[72]

Hydrogen bonding and stacking interactions alone cannot account for the high accuracy of replicative polymerases.[70, 71] The vastly lower error rates of DNA polymerases (even without intrinsic proofreading activity) suggest that the polymerase domain active site residues exaggerate the energetic penalty of a mismatch.[71] This effect highlights the importance of the geometry of a correctly placed dNTP.[71] Tyrosine, lysine, and arginine residues at the polymerase active site form interactions with the template base to hold it strictly in place so that the incoming nucleotide can correctly pair with the template base.[72] Mutagenesis studies that replace these residues with alanine generate large cavities above and below the template base to provide the base with greater mobility.[72] The less restricted movement allows for a greater number geometric conformations that the template base can assume, often stabilizing the formation of a mismatch.

The conformational change that follows nucleotide binding provides another geometric rationale for the high fidelity of DNA polymerases. The conformational change from an open to a closed state that occurs after the nucleotide binds to the polymerase/DNA complex was initially thought to be rate-limiting.[70] However, fluorescence-based kinetic studies have confirmed that the confirmational change is

17

faster than the chemistry of the polymerization reaction of a correct dNTP.[70] This conformational change serves as an important checkpoint in the catalytic mechanism of a replicative polymerase that ensures fidelity. When inserting a correct nucleotide, the rate of the conformational change leading to the incorporation reaction (k_3, Figure 5) is much faster compared to the conformational change in the reverse direction (k_{-3}, Figure 5).[70, 71] When the incorrect nucleotide is inserted, the conformational change in the reverse direction (k_{-3}, Figure 5) is much faster than the conformational change in the forward direction (k_3, Figure 5).[68-72] Structurally, the incorporation of an incorrect nucleotide is thought to require an unfavorable rearrangement of active site residues. This results in the very slow incorporation reaction (k_4, Figure 5) and a very fast shift back to a "pre-chemistry" conformation (k_{-3}, Figure 5).[72]

The partitioning between exonuclease and polymerase active sites is another factor that ensures high fidelity in replicative polymerases. Incorporating the incorrect nucleotide favors relocating the DNA primer-template from the polymerase active site to the exonuclease active site to excise the misincorporated nucleobase. The fidelity of a polymerase is increased 100-1,000-fold with the presence of a properly active exonuclease domain.[72] Initial studies exploring how mismatched DNA is transferred between active sites in replicative polymerases involve Föster Resonance Energy Transfer (FRET) experiments and the use of heparin traps, but ensemble kinetic experiments may not be sufficient to elucidate polymerase-active site dynamics in detail.[70, 72]

1.7 Steady-State and Pre-steady State Kinetics

Protein biochemists and enzymologists use steady-state and pre-steady-state kinetics to determine the catalytic mechanisms of enzymes. Steady-state kinetics methodology was described by Leonor Michaelis and Maude Menten. They proposed a simplified model to describe substrate binding to enzyme through rapid equilibrium to form the Michaelis complex, followed by the irreversible formation of product and release.[73] In typical steady-state experiments, there is an excess of substrate to enzyme so that both the concentration of substrate and enzyme bound to substrate remain unchanged.[72] In addition, the reaction time is long enough to allow for multiple turnovers.[73] Steady-state experiments are often performed by utilizing intrinsic absorbent or fluorescent properties of a reactant or product in the enzyme's catalytic mechanism. The increase or decrease in absorbance or fluorescence is then measured as a function of time to calculate a substrate-dependent observed rate. The observed rate increases until the enzyme is fully saturated and cannot perform catalysis any faster. Because multiple turnovers occur during the time course of the reaction, the observed rate is dictated by the rate of the slowest step in the enzyme's catalytic mechanism.[74] Using several assumptions, Michaelis and Menten derived a hyperbolic equation (Eq 1) describing the substrate-dependent rate of product formation for a particular enzyme.

$$k_{obs} = (k_{cat}*[S])/(K_m+[S]) \qquad \text{Eq 1}$$

This equation provides three important enzymatic parameters: k_{cat}, the catalytic rate constant; K_m, the substrate concentration at half of the k_{cat}; and k_{cat}/K_m, which describes the probability of substrate turnover when substrate is bound to enzyme. These parameters can be very useful when comparing the parameters of mutational variants or different enzymes. For example, a hypothetical enzyme may be very fast at forming

product and have a correspondingly high k_{cat} value, but also have a high K_m. This means

that the enzyme needs a high concentration of substrate to achieve its maximum catalytic

$$E + S \xrightleftharpoons[k_{-1}]{k_1} ES \xrightarrow{k_2} E + P$$

k_{obs} vs [Substrate]
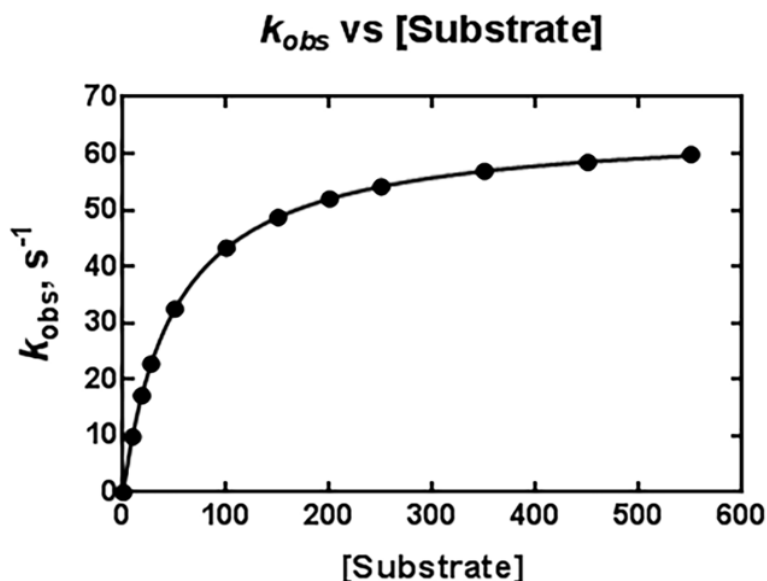
Figure 6: Simplified model of enzyme catalysis. Excess substrate binds to enzyme unchanged before catalysis. The data are fit to the hyperbolic function derived by Michaelis and Menten. Kinetic constants, k_{cat} and K_m, can be obtained from the hyperbolic fitting of the data. The catalytic efficiency (k_{cat}/K_m) is a second order rate constant that is a tangent line to the linear portion of the hyperbolic function.

rate and thus is not a very efficient enzyme.

This steady-state model is limited because it does not detail the steps of the

catalytic cycle of an enzyme. A typical catalytic cycle can include multiple substrate

binding events, conformational changes, chemistry step(s), and product release. These

additional steps may occur faster than the observed rate in a steady-state experiment,

and thus cannot be deconvoluted from the rate-determining step of catalysis. Therefore,

steady-state kinetic experiments cannot yield rates of the individual steps of a catalytic

cycle occurring at the active site. Instead, one must use pre-steady-state kinetics. Pre-steady-state kinetics can allow the enzymologist to define the sequence of reaction events occurring at the active site from substrate binding to product release. These steps account for the high-fidelity of DNA polymerases.[74] A typical pre-steady-state experiment involves the rapid mixing of enzyme with substrate to initiate the reaction for a specific time interval, followed by reaction termination by addition of a quenching reagent such as ethylenediaminetetraacetic acid (EDTA), strong acid, or other chemical denaturant (Figure 7).[74] The rate of a single turnover event can be obtained by analyzing the time-dependent amount of product formation on a rapid time scale.[74]

There are two common experiments performed under pre-steady-state conditions: a burst experiment where substrate concentration is in excess to enzyme, and a single turnover experiment where the concentration of enzyme is in excess to substrate. In an enzyme that displays burst kinetics, a plot of product formation versus time is biphasic. The single exponential first phase is the initial and rapid "burst" of product formation that is followed by the slower linear second phase representing steady-state turnover. The reaction time course can be fit to the burst equation (Eq 2)

$$[product]=A\left[\left(1-e^{-k_{obs}t}\right)+k_{ss}t\right] \qquad \text{Eq 2}$$

Analyzing the rate of the "burst" phase and its amplitude defines the rate of the reaction occurring at the active site (k_{obs}), the concentration of enzyme active in a single turnover (A), and analyzing the linear phase is the rate of the slower steps leading to product release (k_{ss}).[73] It should be noted that a "burst" phase is only observed if the release of product, or some other step *after* chemistry, is slower than catalysis.[74] If a "burst" is not observed, then chemistry, or a step *preceding* chemistry, is rate-limiting.[74]

The rate of the "burst" phase in a burst experiment should be equal to the rate in a single turnover experiment because both rates represent one turnover by the enzyme; however, the single turnover experiment provides more sensitivity because nearly 100% of the substrate is converted to product.[74] In a single turnover experiment, enzyme is in excess to substrate and only one turnover can occur. Therefore, the time-dependent formation of product is fit to a single exponential equation without a linear phase (Eq 3) because a steady-state cannot be reached.

$$[\text{product}]=A(1-e^{-k_{obs}t}) \qquad \text{Eq 3}$$

The observed rates of the exponential portion of the burst assays or observed rates in a single turnover assay are then plotted versus substrate concentration to obtain the maximum catalytic rate, the dissociation constant for a particular substrate (K_d) , and the efficiency of the enzyme similar to the Michaelis-Menten equation (Eq 1) with one exception. Due to a polymerase's processivity, the rate of nucleotide incorporation is much faster than the rate of dissociation from DNA (i.e product release). Therefore, the rapid equilibrium assumption is true for polymerases and K_d replaces K_m.

The experimental set-up involves a rapid chemical quench instrument (Figure 7) where reactants are loaded into small volume sample loops, then rapidly mixed together via force from the drive syringes in the reaction loop, and finally quenched with the quench reagent in the collection loop. An eight-way valve is used to select reaction loop volumes, and an electronic step motor is used to drive reactions at precise speeds. By varying the reaction loop volume, flow rate, and a utilizing a clever push-pause-push mode in the

motor, reaction times as fast as 2 milliseconds and up to 100 seconds (or longer) can be obtained.[74]
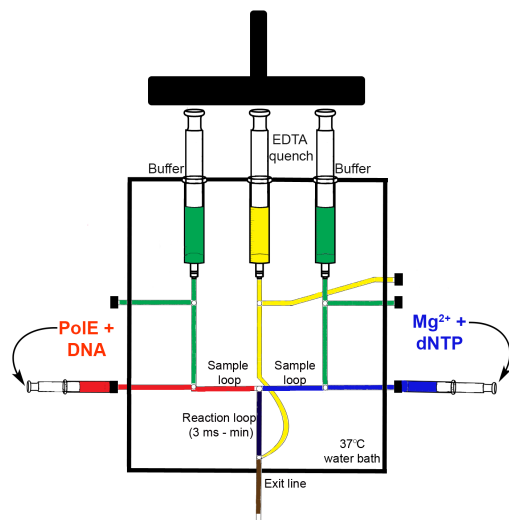
Figure 7: Schematic diagram of a rapid chemical quench. Syringes are loaded separately. One is loaded with Pol ε and labeled DNA. The other is loaded with magnesium and dNTP. An eight-way valve contains the reaction loops that each are a different volume. Based on this volume the computer calculates which reaction loop to use based on the entered reaction time. An electronic step motor forces the reaction volume from sample syringes and buffer from drive syringes into the reaction loop. The instrument then stops the reaction with quenching reagent at the correct time. Figure obtained from C. Sohl with permission.

In pre-steady-state kinetics, the enzyme concentration is present in stoichiometric amounts compared to reactant, and is not treated as a trace catalyst.[74] In addition, the pathway of events occurring at the active site can be established without the indirect interference of steady-state methods.[74] One needs to know the concentrations of total enzyme present, product formed over time, and substrate concentration in order to accurately interpret a kinetic experiment. In a number of laboratories, it is conventional to report moles of enzyme and moles of product formed as a function of time as opposed to concentration.[74] It is difficult to interpret these results because one must know the total reaction volume to calculate the concentration of total product formed.[74] It is better to

operate in concentration units and it greatly facilitates the quantitative analysis of the data.[74]

1.8 DNA Polymerase Epsilon Kinetic Mechanism

To simplify difficulties with expression, the kinetic mechanism of human Pol ε was first determined using the truncated catalytic subunit.[64-66, 75-77] The fragment consists of the residues 1-1189 of the *N*-terminal domain and has been characterized previously.[75-77] The *N*-terminal domain fragment contains all the conserved polymerase and exonuclease motifs and has been shown to be as active as the full-length catalytic subunit.[75] The *N*-terminal fragment possesses exonuclease activity that can complicate nucleotide incorporation kinetics experiments. To circumvent this, two highly conserved aspartates D275 and E277 in the exonuclease domain were mutated to alanine.

Early kinetic studies found that Pol ε extends DNA via an induced-fit mechanism.[75] Using a burst assay, Pol ε was shown to exhibit biphasic kinetics, demonstrating that the rate of nucleotide incorporation precedes dissociation from DNA.[75] Dissociation from DNA was confirmed to occur after nucleotide incorporation by directly measuring the steady-state rate constant.[75]

The rate of the conformational change in DNA polymerases preceding chemistry is usually a point of contention and varies among different polymerases families and even polymerase within the same family.[70, 72, 78] At least three different experiments exist to determine the position and/or rate relative to correct/incorrect nucleotide insertion for a polymerase: 1) examining the elemental effect of nucleotide incorporation, 2) pulse-chase and pulse-quench experiments, and 3) stopped-flow fluorescent experiments.[79, 80] When examining the elemental effect of nucleotide, a dNTPαS substrate is used. A dNTPαS is

a dNTP that contains a non-bridging phosphorothioate substitution at the α-phosphate. The rates of incorporation of dNTPαS and the natural dNTP are then compared and used as a diagnostic for determining whether the chemical step of polymerization is rate-limiting.[74, 80]. The effect of a sulfur elemental substitution is estimated to decrease reaction rate 4-10 fold based on the hydrolysis of phosphate diesters.[80] A small elemental effect of 0.9 was observed for Pol ε when inserting the correct nucleotide, suggesting that the rate of correct nucleotide incorporation was not limited by a conformational change.[74] However, when inserting an incorrect nucleotide, the elemental effect was calculated to be 167, suggesting that in this case the chemistry step is rate-limiting.[74] Substituting a sulfur for an oxygen may result in steric clashing between active site residues and polymerase-bound nucleotide, so these experiments were followed up with pulse-chase and pulse-quench experiments for diagnosing the rate-limiting step.[74] In pulse-quench experiments, enzyme and DNA are rapidly mixed with radiolabeled dNTP and incubated in the millisecond timescale, and then the reaction is rapidly quenched with acid.[74] In pulse-chase experiments, enzyme and DNA are rapidly mixed with radiolabeled dNTP for selected time points, and then incubated ("chased") with a large excess of unlabeled dNTP for an amount of time that allows for complete conversion to products.[74, 80] Reaction volumes following pulse-chase and pulse-quench assays are also normalized. During the chase period, tightly bound radiolabeled substrate is primarily converted to product that retains the radiolabel, while loosely bound substrate is exchanged with the excess unlabeled substrate to form unlabeled product.[74] Comparing the reaction kinetics obtained in pulse-quench and pulse-chase experiments provide evidence for the formation of a ternary complex that precedes chemistry.[80] In addition, these experiments

argue for the existence of an enzyme*-DNA-dNTP (where * indicates a conformational change) that is distinct from the ground state enzyme-DNA-dNTP ternary complex.[80, 81] The results of these experiments performed with Pol ε showed that the rates of product formation were equal, but more product was formed in the pulse-chase assay versus the pulse-quench assay.[74] This result suggests the accumulation of a stable intermediate complex prior to chemistry that is converted into product through the addition of excess unlabeled dNTP.[74] The intermediate can be a binary complex (enzyme-DNA), a ground-state ternary complex (enzyme-DNA-dNTP), or a ternary complex that has undertaken a conformational change (enzyme*-DNA-dNTP). The binary complex can be ruled since it would bind to unlabeled dNTP during the chase period, form undetectable product, and lead to no change in product formation between pulse chase and pulse quench assays.[74] The ground state ternary complex, enzyme-DNA-dNTP, can also be ruled out because the rate of dNTP dissociation was much faster than rate of product formation in the pulse-chase and pulse-quench experiments.[74] This suggests the intermediate is the accumulation of a distinct enzyme*-DNA-dNTP ternary complex that precedes chemistry.[74] However, some uncertainty still remains in the rate of this conformational change.[70, 81, 82] Though they have not been performed on Pol ε, stopped-flow fluorescent labeling experiments have also been used to measure rates of noncovalent conformational changes in polymerases. Results of these experiments have shown evidence that one or more fast conformational changes often precede nucleotide incorporation chemistry in replicative polymerase turnover.[70, 72, 78, 79, 81-84] Stopped-flow fluorescence will not be discussed in detail, but it does pose as an interesting path forward for elucidating Pol ε kinetics.

The contributing effects of the exonuclease domain, *C*-terminal domain (CTD), and additional subunits on incorporation rates, excision rates, DNA binding, and fidelity were then explored. With the truncated catalytic subunit, the exonuclease domain was found to enhance the accuracy of DNA replication from one error every $10^4 – 10^7$ incorporation events to one error every $10^6 – 10^{11}$ incorporation events.[76, 85] Zahurancik and colleagues successfully purified full-length Pol ε in 2015, including accessory subunits, and found the CTD and accessory subunits did not improve processivity. The CTD and accessory subunits did increase DNA binding affinity and decrease the rate of exicision.[77] A thorough kinetic investigation on full-length Pol ε was completed in 2020.[86] In this study, the minimal kinetic mechanism of Pol ε was extended to include three different modes of binding to DNA: a productive polymerization-ready state, a nonproductive state, and a productive exonuclease-ready state.[86] The nonproductive conformation is notable because it accounts for the observed low percent activity of Pol ε during the single turnover phase, and the isomerization to a productive conformation accounts for the double exponential curve fitting.[86] This work also reported a moderately increased rate of nucleotide incorporation. Overall, the full-length Pol ε followed the same kinetic mechanism as the truncated catalytic subunit. However, in this most recent study, the pulse-chase and pulse-quench experiments were not repeated to confirm the sub-rate-limiting conformation change.[86]

1.9 DNA Polymerase Epsilon Mutations in Cancer

Next-generation sequencing studies from The Cancer Genome Atlas (TCGA) network provided a new classification for endometrial tumors based on mutation spectra.[32] Pol ε mutated tumors account for 7-12% of all endometrial cancers and

possess an hypermutated phenotype (> 10 mutations per megabase).[32-36, 59, 60] Mutations in Pol ε tend to cluster in the exonuclease domain.[32-36, 59, 60] The most frequently observed mutation in Pol ε is P286R, which is located in the exonuclease domain.[59, 60] Initial investigations characterizing the yeast Pol ε variant of this P286R mutant (P301R in yeast) have provided evidence of a higher mutation rate compared to an exonuclease deficient Pol ε. This mechanism occurs due to increased activity of nucleotide incorporation coupled to deficient proofreading activity.[59, 60] A crystal structure of yeast P301R Pol ε coupled with molecular dynamics revealed that the arginine mutation prevents DNA from properly binding to the exonuclease active site.[61] Furthermore, the structural investigation revealed that in the low probability that the DNA did bind properly to the exonuclease active site, the magnesium ions would not be coordinated properly to facilitate exonuclease chemistry.[61] Mechanistically, the P286R mutant Pol ε is hypothesized to drive mutagenesis by not only preventing the exonuclease reaction from occurring, but also increasing the rate of nucleotide incorporation after mismatch.[61] The fidelity of DNA synthesis is partially maintained by a delicate balance between the forward polymerization and excision rates of reaction. Altering this balance will affect the propensity of the DNA polymerase to extend DNA synthesis from misincorporated nucleotides or bypass DNA lesions in an error-prone manner.[60] The rate of incorporating a correct nucleotide is ~ 300 s^{-1} for Pol ε.[61] If the enzyme attempts to extend past incorrectly matched dNTPs, the rate will be significantly slowed.[61] This slowed polymerization rate allows the polymerase to overcome the kinetic barrier to transfer the nascent single-stranded (ss) DNA to the exonuclease active site. The rates of separation of the DNA strands, transfer to the exonuclease site, and binding to the exonuclease

active site occur faster than the rate of incorporation past a mismatched nucleotide. Thus, this kinetically favors excision of the incorrectly incorporated nucleotide by the polymerase.[61] Due to steric clashes, the P286R mutation is expected to increase the kinetic barrier to form the necessary ssDNA-exonuclease domain complex. As a
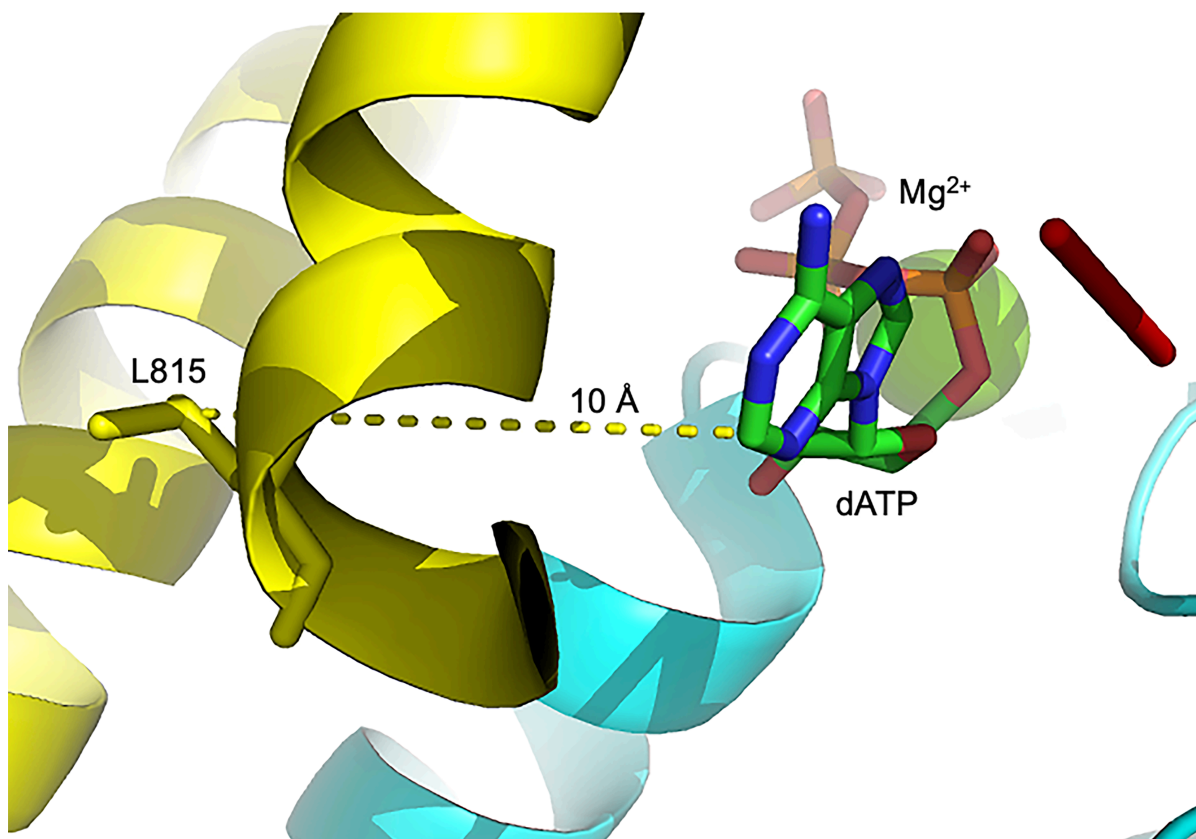
Figure 8: Energy minimized F815L (human numbering) was generated using the crystal structure of WT yeast Pol ε (PDB 4M8O).[63] The mutation is 10 Å away from the incoming dNTP and is thought to not directly interfere with the chemistry of the nucleotide incorporation reaction.

consequence, the ability to extend past mismatched nucleotides is enhanced and the rate excision decreases.[61]

In addition, a large number of Pol ε polymerase domain mutations have also been identified in endometrial cancer in regions of the enzyme critical for fidelity, implying that polymerase domain mutations could also confer a hypermutated phenotype that leads to cancer.[87] It is well established that mutations the polymerase domain of replicative

polymerases can alter fidelity and cause disease.[88, 89] For example, the degree of

catalytic dysfunction and infidelity of polymerase domain mutations in DNA polymerase γ

involved in neurodegenerative diseases such as Alpers syndrome has been previously

assessed.[90] In addition, the polymerase domain mutation R696W in human DNA

polymerase δ has been identified in colon and liver cancer and causes a hypermutated

phenotype in yeast.[91] Currently, the kinetic mechanisms and propensity for genome
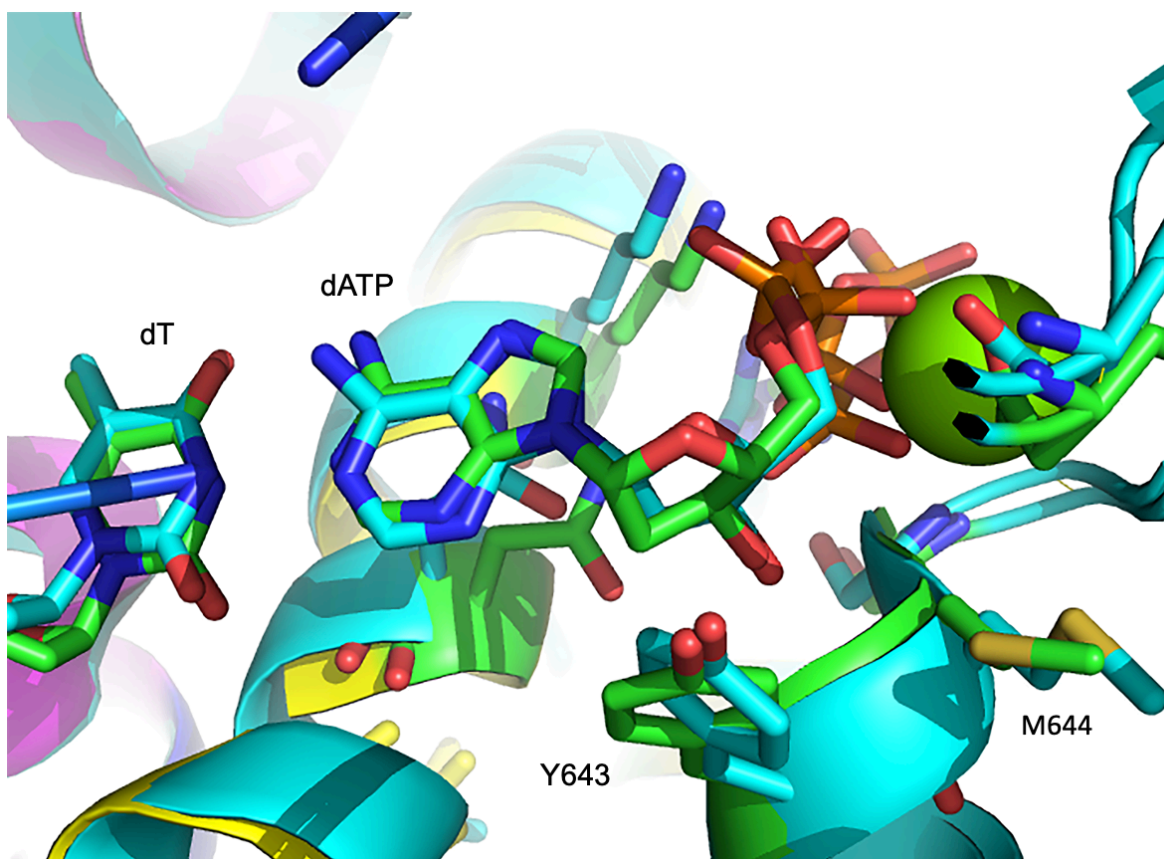
Figure 9: Aligned structures of F815L DNA Polymerase epsilon (colored in cyan) and WT yeast Pol ε (colored by domain) (4M8O).[63] Residues involved in the dNTP binding pocket are labeled. Notably Y643 and M644 are shifted in the F815L mutant. This causes increases the conformational flexibility of the incoming dNTP and can stabilize the formation of a mismatch.

instability of cancer-related polymerase domain mutations in Pol ϵ has not been

determined. As more Pol ϵ mutations are identified in tumors, it is vital to determine the

mechanistic consequences on replicative (in)fidelity that underscore genome instability.

We have chosen to focus on a mutation in the fingers subdomain, F815L, due to its identification in cancer patients and its distance from the active site to not directly interfere with the chemistry of nucleotide incorporation (Figure 8). From energy minimized structures of yeast Pol ε possessing the F815L (human numbering) mutant and its location in the protein, we hypothesize that this mutant interferes with proper dNTP selectivity and creates a larger binding pocket for DNA at the polymerase active site (Figure 9). By understanding the catalytic features of Pol ε mutations identified in tumors, we can elucidate the basic features of the enzyme's role in tumorigenesis and genome instability. Currently we have shown that the F815L mutant inserts the correct nucleotide with a lower catalytic efficiency that WT Pol ε. We hypothesize that the F815L mutant will be more efficient at incorporating incorrect nucleotides through higher affinity for the incorrect dNTP or faster rates for the incorrect dNTP, resulting in lowered fidelity.

1.10 Outstanding DNA Polymerase Mechanistic Questions

Structural and kinetic experiments provide a multitude of information and together form the strong basis of any polymerase investigation. Studies involving Pol ε studies and the role of tumorigenic mutants in the overall kinetic mechanism and tumorigenesis are still in the early stages. Traditional paths forward would be to obtain high-resolution structures of human Pol ε and mutant variants in an open and closed state, coupled with a complete kinetic characterization of the wild-type enzyme and mutants. This illuminates structure-function relationships and provides a rationale for the mechanism of (in)fidelity of the enzyme and mutants. Initial pre-steady state kinetic data for human Pol ε exist, however the data are limited to only the wild-type enzyme .[74-77, 86] High-resolution structures for Pol ε are limited to the catalytic subunit of the yeast homologue in a closed

state, in addition to a cryo-EM structures.[61, 63, 64, 69] Optimizing expression protocols to obtain high yields of active protein is essential for further kinetic and structural investigations. Probing for sensitivity to current FDA approved chemotherapeutic nucleoside analogs poses an interesting path forward for kinetic studies as it could direct potential treatment strategies.

Future research will require a transition from current structure-function relationships to studying polymerase dynamics. For example, investigating the pathway of the conformational change checkpoint that leads to catalysis, identifying the non-catalytic residues that affect the partitioning between the polymerase and exonuclease active sites and/or affect dNTP insertion kinetics, and identifying protein side chain rearrangements that occur during dNTP insertion can all play a role in base selectivity but are missed by traditional static structural studies.[72] The structures of the closed conformation help define the interactions that lead to catalysis, but do not shed light on the pathway from open to closed conformation.[72] Instead, these processes can be investigated by developing single molecule FRET (smFRET) and time-lapse X-ray crystallography. Time lapse X-ray crystallography has been successful in studying DNA polymerase η and DNA polymerase β, [92, 93] though these are both repair polymerases that have slower insertion rates. Therefore, Pol ε would have to be artificially kinetically slowed in order to observe the dynamics of dNTP insertion. Moreover, high-quality crystals may be difficult to obtain or gradually deform during the course of the reaction.[72] Performing Raman spectroscopy on the crystals would serve as a possible alternative because well-defined changes in the spectra may appear during the nucleotidyl transfer reaction to elucidate the kinetics of the appearance and disappearance of intermediate.[72,

[94] However, transition rates in solution that are important to understand the changes between different conformational states are lost. Thus, smFRET techniques using clever labeling techniques would help assign the role of conformational changes governing dNTP selectivity the transition from polymerase to exonuclease active sties.[72]

While these questions tackle local factors affecting fidelity, they do not take into account the potential global effects of coordinated efforts of full-length Pol ε and additional subunits, and accompanying proteins such as PCNA complex, RF-C, CMG-helicase complex at the replisome on fidelity, processivity, rates of nucleotide insertion and excision, cellular tumorigenic potential.

1.11 References

1.   Hassanpour, S. H., Dehghani, M. (2017). Review of cancer from perspective of molecular. *Journal of Cancer Research and Practice, 4*(4), 127-129.

2.   Siegel, R., Naishadham, D., Jemal, A. (2013). Cancer statistics, 2013. *CA Cancer J Clin, 63*(1), 11-30.

3.   Nowell, P. C. (1976). The Clonal Evolution of Tumor Cell Populations. *Science, 194*, 23-28.

4.   Charles, N., Ozawa, T., Squatrito, M., Bleau, A. M., Brennan, C. W., Hambardzumyan, D., Holland, E. C. (2010). Perivascular nitric oxide activates notch signaling and promotes stem-like character in PDGF-induced glioma cells. *Cell Stem Cell, 6*(2), 141-152.

5.   Balkwill, F. R., Capasso, M., Hagemann, T. (2012). The tumor microenvironment at a glance. *J Cell Sci, 125*(23), 5591-5596.

6.   Hao, G., Xu, Z. P., Li, L. (2018). Manipulating extracellular tumour pH: an effective target for cancer therapy. *RSC Advances, 8*(39), 22182-22192.

7.   Izumi, H., Torigoe, T., Ishiguchi, H., Uramoto, H., Yoshida, Y., Tanabe, M., Ise, T., Murakami, T., Yoshida, T., Nomoto, M., Kohno, K. (2003). Cellular pH regulators: potentially promising molecular targets for cancer chemotherapy. *Cancer Treatment Reviews, 29*(6), 541-549.

8.   Hanahan, D., Weinberg, R. A. (2000). The Hallmarks of Cancer. *Cell, 100*, 57-70.

9.   Hanahan, D., Weinberg, R. A. (2011). Hallmarks of cancer: the next generation. *Cell, 144*(5), 646-674.

10.   Warburg, O., Wind, F. and Negelein, E. (1927) The metabolism of tumors in the body. *J. Gen. Physiol*, 8, 519–530

11.   Sakashita, M., Aoyama, N., Minami, R., Maekawa, S., Kuroda, K., Shirasaka, D., Ichihara, T., Kuroda, Y., Maeda, S., Kasuga. (2001) Glut1 expression in T1 and T2 stage colorectal carcinomas: its relationship to clinicopathological features. *Eur. J. Cancer,* 37, 204–209

12.   Wu, M., Neilson, A., Swift, A.L., Moran, R., Tamagnine, J., Parslow, D, Armistead, S., Lemire, K., Orrell, J., Teich, J., Chomicz, S., Ferrick, D., A. (2007) Multiparameter metabolic analysis reveals a close link between attenuated mitochondrial bioenergetic function and enhanced glycolysis dependency in human tumor cells. *Am. J. Physiol. Cell Physiol*, 292, C125–C136

13.   Lai, J.-H., Jan, H.-J., Liu, L.-W., Lee, C.-C., Wang, S.-G., Hueng, D.-Y. et al. (2013) Nodal regulates energy metabolism in glioma cells by inducing expression of hypoxia-inducible factor 1α. *Neuro-Oncology,* 15, 1330–1341

14.   Michelakis, E.D., Sutendra, G., Dromparis, P., Webster, L., Haromy, A., Niven, E. et al. (2010) Metabolic modulation of glioblastoma with dichloroacetate. *Sci. Transl. Med*, 2, 31-34

15.   Warburg, O. (1956) On the origin of cancer cells. *Science,* 123, 309–314

16.   Warburg, O. (1956) On respiratory impairment in cancer cells. *Science,* 124, 269–270

17. Potter, M., Newport, E., Morten, K. J. (2016). The Warburg effect: 80 years on. *Biochem Soc Trans, 44*(5), 1499-1505.

18. Wise, D. R., Ward, P. S., Shay, J. E. S., Cross, J. R., Gruber, J. J., Sachdeva, U. M. (2011). Hypoxia promotes isocitrate dehydrogenase- dependent carboxylation of α-ketoglutarate to citrate to support cell growth and viability. *Proc. Natl. Acad. Sci. U.S.A, 108*(49), 19611-19616.

19. Filipp, F. V., Scott, D. A., Ronai, Z. A., Osterman, A. L., Smith, J. W. (2011). Reverse TCA cycle flux through isocitrate dehydrogenases 1 and 2 is required for lipogenesis in hypoxic melanoma cells. *Pigment Cell and Melanoma Res.*(25), 375-381.

20. Horbinski, C. (2013). What do we know about IDH1 / 2 mutations so far, and how do we use it? *Acta Neuropathol, 125*, 621-636.

21. Sporn, M., B. (1996). The war on cancer. *Lancet, 347*, 1377-1381.

22. Webb, B. A., Chimenti, M., Jacobson, M. P., Barber, D. L. (2011). Dysregulated pH: a perfect storm for cancer progression. *Nat Rev Cancer, 11*(9), 671-677.

23. Schonichen, A., Webb, B. A., Jacobson, M. P., Barber, D. L. (2013). Considering protonation as a posttranslational modification regulating protein structure and function. *Annu. Rev. Biophys, 42*, 289-314.

24. Al-Habori, M., Peak, M., Thomas, T., Agius, L. (1992). The role of cell swelling in the stimulation of glycogen synthesis by insulin. *Biochem. J., 282*, 789-796.

25. Dietl, K., Renner, K., Dettmer, K., Timischl, B., Eberhart, K., Dorn, C., Hellerbrand, C., Kastenberger, M., Kunz-Schughart, L. A., Oefner, P. J., Andreesen, R., Gottfried, E., Kreutz, M. P. (2010). Lactic acid and acidification inhibit TNF secretion and glycolysis of human monocytes. *J Immunol, 184*(3), 1200-1209.

26. Kuwata, F., Suzuki, N., Otsuka, K., Minoru, T., Sasai, Y., Wakino, H., Ito, M., Ebihara, S., Suzuki, K. (1991). Enzymatic regulation of glycolysis and gluconeogenesis in rabbit periodontal ligament under various physiological pH conditions. *J. Nihon Univ. Sch. Dent., 33*, 81-90.

27. Gray, J. A. (1966). Kinetics of enamel dissolution during formation of incipient caries-like lesions. *Arch. Oral Bio, 11*, 397-421.

28. Deberardinis, R. J., Sayed, N., Ditsworth, D., Thompson, C. B. (2008). Brick by brick: metabolism and tumor cell growth. *Curr Opin Genet Dev, 18*(1), 54-61.

29. Isom, D. G., Sridharan, V., Dohlman, H. G. (2016). Regulation of Ras paralog thermostability by networks of buried ionizable groups. *Biochemistry, 55*(3), 534-542.

30. Isom, D. G., Sridharan, V., Baker, R., Clement, S. T., Smalley, D. M., Dohlman, H. G. (2013). Protons as Second Messenger Regulators of G Protein Signaling. *Mol. Cell, 51*(4), 531-538.

31. Isom, D. G., Dohlman, H. G. (2015). Buried ionizable networks are an ancient hallmark of G protein-coupled receptor activation. *Proc Natl Acad Sci U S A, 112*(18), 5702-5707.

32. Heitzer, E., Tomlinson, I. (2014). Replicative DNA polymerase mutations in cancer. *Curr Opin Genet Dev, 24*, 107-113.

33. Kandoth, C., Schultz, N., Cherniack, A. D., Akbani, R., Liu, Y., Shen, H., Robertson, A. G., Pashtan, I., Shen, R., Benz, C. C., Yau, C., Laird, P. W., Ding, L., Zhang, W., Mills, G. B., Kucherlapati, R., Mardis, E. R., Levine, D. A. (2013). Integrated genomic characterization of endometrial carcinoma. *Nature, 497*(7447), 67-73.

34. Spier, I., Holzapfel, S., Altmuller, J., Zhao, B., Horpaopan, S., Vogt, S., Chen, S., Morak, M., Raeder, S., Kayser, K., Stienen, D., Adam, R., Nurnberg, P., Plotz, G., Holinski-Feder, E., Lifton, R. P., Thiele, H., Hoffmann, P., Steinke, V., Aretz, S. (2015). Frequency and phenotypic spectrum of germline mutations in POLE and seven other polymerase genes in 266 patients with colorectal adenomas and carcinomas. *Int J Cancer, 137*(2), 320-331.

35. Yoshida, R., Miyashita, K., Inoue, M., Shimamoto, A., Yan, Z., Egashira, A., Oki, E., Kakeji, Y., Oda, S., Maehara, Y. (2011). Concurrent genetic alterations in DNA

polymerase proofreading and mismatch repair in human colorectal cancer. *Eur J Hum Genet, 19*(3), 320-325.

36. Alexander, S., Ben-Shachar, S., Ling, S. C., Gallinger, S., Constantini, S., Dirks, P., Huang, A., Scherer, S. W., Grundy, R. G., Durno, C., Aronson, M., Gartner, A., Meyn, M. S., Taylor, M. D., Pursell, Z. F., Pearson, C. E., Malkin, D., Futreal, P. A., Stratton, M. R., Bouffet, E., Hawkins, C., Campbell, P. J., Tabori, U. (2015). Combined hereditary and somatic mutations of replication error repair genes result in rapid onset of ultra-hypermutated cancers. *Nat Genet, 47*(3), 257-262.

37. Parsons, D. W., Jones, S., Zhang, X., Lin, J. C.-h., Leary, R. J., Angenendt, P., Mankoo, P., Carter, H., Siu, I. m., Gallia, G. L., Olivi, A., McLendon, R., Rasheed, B. A., Keir, S., Nikolskaya, T., Nikolsky, Y., Busam, D. A., Tekleab, H., Jr, L. A. D., Hartigan, J., Smith, D. R., Strausberg, R. L., Kazue, S., Marie, N., Mieko, S., Shinjo, O., Yan, H., Riggins, G. J., Bigner, D. D., Karchin, R., Papadopoulos, N., Parmigiani, G., Vogelstein, B., Velculescu, V. E., Kinzler, K. W. (2008). An integrated genomic analysis of human glioblastoma mutliforme. *Science, 321*(5897), 1807-1812.

38. Yan, H., Williams, P., Jin, G., McLendon, R., Rasheed, B. A., Yuan, W., Kos, I., Ines, B-H., Jones, S., Riggins, G., Friedman., H., Friedman, A., Reardon, D., Herndon., J., Kinzler, K. W., Velculescu, V. E., Vogelstein, B., Bigner, D. D. (2009). IDH1 and IDH2 mutations in gliomas. *N Engl J Med, 360*(8), 765-773.

39. Larson, D. E., Ph, D., McLellan, M. D., Chen, K., Ph, D., Koboldt, D. C., Fulton, R. S., Delehaunty, K. D., McGrath, S. D., Fulton, L. A., Locke, D. P., Ph, D., Magrini, V. J., Ph, D., Abbott, R. M., Vickery, T. L., Reed, J. S., Robinson, J. S., Wylie, T., Smith, S. M., Carmichael, L., Eldred, J. M., Harris, C. C., Walker, J., Peck, J. B., Du, F., Dukes, A. F., Sanderson, G. E., Brummett, A. M., Clark, E., McMichael, J. F., Meyer, R. J., Schindler, J. K., Pohl, C. S., Wallis, J. W., Ph, D., Shi, X., Lin, L., Schmidt, H., Tang, Y., Haipek, C., Wiechert, M. E., Ivy, J. V., Kalicki, J., Elliott, G., Ries, R. E., Payton, J. E., Ph, D., Westervelt, P., Ph, D., Heath, S., Shannon, W. D., Ph, D., Nagarajan, R., Ph, D., Link, D. C., Walter, M. J., Graubert, T. A., Dipersio, J. F., Ph, D., Wilson, R. K., Ph, D., Ley, T. J. (2009). Recurring mutations found by sequencing an acute myeloid leukemia genome. *N Engl J Med, 361*, 1058-1066.

40. Borger, D. R., Tanabe, K. K., Fan, K. C., Lopez, H. U., Fantin, V. R., Straley, K. S., Schenkein, D. P., Hezel, A. F., Ancukiewicz, M., Liebman, H. M., Kwak, E. L., Clark, J. W., Ryan, D. P., Deshpande, V., Dias-Santagata, D., Ellisen, L. W., Zhu, A. X., Iafrate, J. A. (2012). Frequent mutation of isocitrate dehydrogenase (IDH)1

and IDH2 in cholangiocarcinoma identified through broad-based tumor genotyping. *Oncologist, 17*(1), 72-79.

41.   Calvert, A. E., Chalastanis, A., Wu, Y., Hurley, L. A., Kouri, F. M., Bi, Y., Kachman, M., May, J. L., Bartom, E., Hua, Y., Mishra, R., Schlitz, G. E., Dubrovskyi, O., Mazar, A. P., Peter, M. E., Zheng, H., James., C. D., Burant, C., F., Chandel, N. S., Davuluri, R. V., Horbinski, C., Stegh, A. H (2017). Cancer-associated IDH1 promotes growth and resistance to targeted therapies in the absence of mutation. *Cell Rep., 19*(9), 1858-1873.

42.   Xu, X., Zhao, J., Xu, Z., Peng, B., Huang, Q., Arnold, E., Ding, J. (2004). Structures of human cytosolic NADP-dependent isocitrate dehydrogenase reveal a novel self-regulatory mechanism of activity. *J Biol Chem, 279*(32), 33946-33957.

43.   Metallo, C. M., Gameiro, P. A., Bell, E. L., Mattaini, K. R., Yang, J., Hiller, K., Jewell, C. M., Johnson, Z. R., Irvine, D. J., Guarente, L., Kelleher, J. K., Vander, M. G. (2012). Reductive glutamine metabolism by IDH1 mediates lipogenesis under hypoxia. *Nature, 481*, 380-384.

44.   Leonardi, R., Subramanian, C., Jackowski, S., Rock, C. O. (2012). Cancer-associated isocitrate dehydrogenase mutations inactivate NADPH-dependent reductive carboxylation. *J Biol Chem, 287*(18), 14615-14620.

45.   Jo, S. H.; Son, M. K.; Koh, H. J.; Lee, S. M.; Song, I. H.; Kim, Y. O.; Lee, Y. S.; Jeong, K. S.; Kim, W. B.; Park, J. W.; Song, B. J.; Huh, T. L., Control of mitochondrial redox balance and cellular defense against oxidative damage by mitochondrial NADP+-dependent isocitrate dehydrogenase. *J Biol Chem***, *276* (19), 16168-76.

46.   Yi, C. Q.; Jia, G. F.; Hou, G. H.; Dai, Q.; Zhang, W.; Zheng, G. Q.; Jian, X.; Yang, C. G.; Cui, Q. A.; He, C. A. (2010) Iron-catalysed oxidation intermediates captured in a DNA repair dioxygenase. *Nature***, *468* (7321), 330-333.

47.   Figueroa, M. E.; Abdel-Wahab, O.; Lu, C.; Ward, P. S.; Patel, J.; Shih, A.; Li, Y.; Bhagwat, N.; Vasanthakumar, A.; Fernandez, H. F.; Tallman, M. S.; Sun, Z.; Wolniak, K.; Peeters, J. K.; Liu, W.; Choe, S. E.; Fantin, V. R.; Paietta, E.; Lowenberg, B.; Licht, J. D.; Godley, L. A.; Delwel, R.; Valk, P. J.; Thompson, C. B.; Levine, R. L.; Melnick, A. (2010**)** Leukemic IDH1 and IDH2 mutations result in

a hypermethylation phenotype, disrupt TET2 function, and impair hematopoietic differentiation. *Cancer Cell*, *18* (6), 553-67.

48. Chowdhury, R.; Yeoh, K. K.; Tian, Y. M.; Hillringhaus, L.; Bagg, E. A.; Rose, N. R.; Leung, I. K.; Li, X. S.; Woon, E. C.; Yang, M.; McDonough, M. A.; King, O. N.; Clifton, I. J.; Klose, R. J.; Claridge, T. D.; Ratcliffe, P. J.; Schofield, C. J.; Kawamura, A. (2011) The oncometabolite 2-hydroxyglutarate inhibits histone lysine demethylases. *EMBO Rep*, *12* (5), 463-9.

49. Nunez, F. J., Mendez, F. M., Kadiyala, P., Alghamri, M. S., Savelieff, M. G., Garcia-Fabiani, M. B., Haase, S., Koschmann, C., Calinescu, A., Kamran, N., Saxena, M., Patel, R., Carney, S., Guo, Z. M., Edwards, M., Ljungman, M., Qin, T., Sartor A. M., Tagett, R., Venneti, S., Brosnan-Cashman, J., Meeker, A., Gorbunova, V., Herting, J. C., Ross, J. L., Hambardzumyan, D., Hervey-Jumper, S., Figueroa, M. E., Lowenstein, P. R., Castro, M. G. (2019). IDH1-R132H acts as a tumor suppressor in glioma via epigenetic up-regulation of the DNA damage response. *Sci. Transl. Med., 11*, 1-13.

50. Dang, L., White, D. W., Gross, S., Bennett, B. D., Bittinger, M. A., Driggers, E. M., Fantin, V. R., Jang, H. G., Jin, S., Keenan, M. C., Marks, K. M., Prins, R. M., Ward, P. S., Yen, K. E., Liau, L. M., Rabinowitz, J. D., Cantley, L. C., Thompson, C. B., Heiden, M. G. V., Su, S. M. (2009). Cancer-associated IDH1 mutations produce 2-hydroxyglutarate. *Nature, 462*, 739-744.

51. Xu, W.; Yang, H.; Liu, Y.; Yang, Y.; Wang, P.; Kim, S. H.; Ito, S.; Yang, C.; Wang, P.; Xiao, M. T.; Liu, L. X.; Jiang, W. Q.; Liu, J.; Zhang, J. Y.; Wang, B.; Frye, S.; Zhang, Y.; Xu, Y. H.; Lei, Q. Y.; Guan, K. L.; Zhao, S. M.; Xiong, Y. (2011) Oncometabolite 2-hydroxyglutarate is a competitive inhibitor of alpha-ketoglutarate-dependent dioxygenases. *Cancer Cell*, *19* (1), 17-30.

52. Lu, C.; Ward, P. S.; Kapoor, G. S.; Rohle, D.; Turcan, S.; Abdel-Wahab, O.; Edwards, C. R.; Khanin, R.; Figueroa, M. E.; Melnick, A.; Wellen, K. E.; O'Rourke, D. M.; Berger, S. L.; Chan, T. A.; Levine, R. L.; Mellinghoff, I. K.; Thompson, C. B. (2012) IDH mutation impairs histone demethylation and results in a block to cell differentiation. *Nature*, *483* (7390), 474-8.

53. Davis, M. I.; Gross, S.; Shen, M.; Straley, K. S.; Pragani, R.; Lea, W. A.; Popovici-Muller, J.; Delabarre, B.; Artin, E.; Thorne, N.; Auld, D. S.; Li, Z.; Dang, L.; Boxer, M. B.; Simeonov, A. (2014) Biochemical, cellular and biophysical characterization

of a potent inhibitor of mutant isocitrate dehydrogenase IDH1. *J Biol Chem*, *289*, 13717-13725.

54.   Popovici-Muller, J.; Saunders, J. O.; Salituro, F. G.; Travins, J. M.; Yan, S.; Zhao, F.; Gross, S.; Dang, L.; Yen, K. E.; Yang, H.; Straley, K. S.; Jin, S.; Kunii, K.; Fantin, V. R.; Zhang, S.; Pan, Q.; Shi, D.; Biller, S. A.; Su, S. M. (2012) Discovery of the first potent inhibitors of mutant IDH1 that lower tumor 2-HG in vivo. *ACS Med Chem Lett*, *3* (10), 850-5.

55.   Stein, E. M.; DiNardo, C. D.; Pollyea, D. A.; Fathi, A. T.; Roboz, G. J.; Altman, J. K.; Stone, R. M.; DeAngelo, D. J.; Levine, R. L.; Flinn, I. W.; Kantarjian, H. M.; Collins, R.; Patel, M. R.; Frankel, A. E.; Stein, A.; Sekeres, M. A.; Swords, R. T.; Medeiros, B. C.; Willekens, C.; Vyas, P.; Tosolini, A.; Xu, Q.; Knight, R. D.; Yen, K. E.; Agresta, S.; de Botton, S., Tallman, M. S., (2017) Enasidenib in mutant IDH2 relapsed or refractory acute myeloid leukemia. *Blood*, *130* (6), 722-731.

56.   Popovici-Muller, J.; Lemieux, R. M.; Artin, E.; Saunders, J. O.; Salituro, F. G.; Travins, J.; Cianchetta, G.; Cai, Z.; Zhou, D.; Cui, D.; Chen, P.; Straley, K.; Tobin, E.; Wang, F.; David, M. D.; Penard-Lacronique, V.; Quivoron, C.; Saada, V.; de Botton, S.; Gross, S.; Dang, L.; Yang, H.; Utley, L.; Chen, Y.; Kim, H.; Jin, S.; Gu, Z.; Yao, G.; Luo, Z.; Lv, X.; Fang, C.; Yan, L.; Olaharski, A.; Silverman, L.; Biller, S.; Su, S. M.; Yen, K. (2018) Discovery of AG-120 (Ivosidenib): a first-in-class mutant IDH1 inhibitor for the treatment of IDH1 mutant cancers. *ACS Med Chem Lett*, *9* (4), 300-305.

57.   Rohle, D.; Popovici-Muller, J.; Palaskas, N.; Turcan, S.; Grommes, C.; Campos, C.; Tsoi, J.; Clark, O.; Oldrini, B.; Komisopoulou, E.; Kunii, K.; Pedraza, A.; Schalm, S.; Silverman, L.; Miller, A.; Wang, F.; Yang, H.; Chen, Y.; Kernytsky, A.; Rosenblum, M. K.; Liu, W.; Biller, S. A.; Su, S. M.; Brennan, C. W.; Chan, T. A.; Graeber, T. G.; Yen, K. E.; Mellinghoff, I. K. (2013) An inhibitor of mutant IDH1 delays growth and promotes differentiation of glioma cells. *Science*, *340* (6132), 626-30.

58.   Xie, X.; Baird, D.; Bowen, K.; Capka, V.; Chen, J.; Chenail, G.; Cho, Y.; Dooley, J.; Farsidjani, A.; Fortin, P.; Kohls, D.; Kulathila, R.; Lin, F.; McKay, D.; Rodrigues, L.; Sage, D.; Toure, B. B.; van der Plas, S.; Wright, K.; Xu, M.; Yin, H.; Levell, J.; Pagliarini, R. A., (2017) Allosteric mutant IDH1 inhibitors reveal mechanisms for IDH1 mutant and isoform selectivity. *Structure*, *25* (3), 506-513.

59.   Barbari, S. R., Kane, D. P., Moore, E. A., Shcherbakova, P. V. (2018). Functional Analysis of Cancer-Associated DNA Polymerase epsilon Variants in Saccharomyces cerevisiae. *G3 (Bethesda), 8*(3), 1019-1029.

60.   Xing, X., Kane, D. P., Bulock, C. R., Moore, E. A., Sharma, S., Chabes, A., Shcherbakova, P. V. (2019). A recurrent cancer-associated substitution in DNA polymerase epsilon produces a hyperactive enzyme. *Nat Commun,(10)*

61.   Parkash, V., Kulkarni, Y., Ter Beek, J., Shcherbakova, P. V., Kamerlin, S. C. L., Johansson, E. (2019). Structural consequence of the most frequently recurring cancer-associated substitution in DNA polymerase epsilon. *Nat Commun, 10*(1), 373.

62.   Garcia-Diaz, M., Bebenek, K., . (2007). Multiple functions of DNA polymerases. *CRC Crit. Rev. Plant Sci., 26*(2), 105-122.

63.   Hogg, M., Osterman, P., Bylund, G. O., Ganai, R. A., Lundstrom, E. B., Sauer-Eriksson, A. E., Johansson, E. (2014). Structural basis for processive DNA synthesis by yeast DNA polymerase epsilon. *Nat Struct Mol Biol, 21*(1), 49-55. doi:10.1038/nsmb.2712

64.   Jain, R., Rajashankar, K. R., Buku, A., Johnson, R. E., Prakash, L., Prakash, S., Aggarwal, A. K. (2014). Crystal structure of yeast DNA polymerase epsilon catalytic domain. *PLoS One, 9*(4), e94835. doi:10.1371/journal.pone.0094835

65.   Syvaoja, J., Suomensaari, S., Nishida, C., Goldsmith, J. S., Chui, G. S. J., Jain, Sanjay, J., Linn, S., . (1990). DNA polymerases a, 6, and E- Three distinct enzymes from HeLa cells. *Proc. Natl. Acad. Sci. USA, 87*, 6664-6668.

66.   Li, Y., Asahara, H., Patel, V. S., Zhou, S., Linn, S. (1997). Purification, cDNA cloning, and gene mapping of the small subunit of human DNA polymerase epsilon. *J Biol Chem, 272*(51), 32337-32344. doi:10.1074/jbc.272.51.32337

67.   Chui, G., Linn, S. (1995). Further characterization of HeLa DNA polymerase epsilon. *J Biol Chem, 270*(14), 7799-7808. doi:10.1074/jbc.270.14.7799

68.   Steitz, T. A. (1999). DNA polymerases: structural diversity and common mechanisms. *J Biol Chem, 274*(25), 17395-17398. doi:10.1074/jbc.274.25.17395

69.   Yuan, Z., Georgescu, R., Schauer, G. D., O'Donnell, M. E., Li, H. (2020). Structure of the polymerase epsilon holoenzyme and atomic model of the leading strand replisome. *Nat Commun, 11*(1), 3156. doi:10.1038/s41467-020-16910-5

70.   Johnson, K. A. (2010). The kinetic and chemical mechanism of high-fidelity DNA polymerases. *Biochim Biophys Acta, 1804*(5), 1041-1048.

71.   Doublie, S., Ellenberger, T. (1998). The mechanism of action of T7 DNA polymerase. *Curr Opin Struct Biol, 8*, 704-712.

72.   Xia, S., Konigsberg, W. H. (2014). RB69 DNA polymerase structure, kinetics, and fidelity. *Biochemistry, 53*(17), 2752-2767.

73.   Michaelis, L.; Menten, M. L.; Johnson, K. A.; Goody, R. S. (2011) The original Michaelis constant: translation of the 1913 Michaelis-Menten paper. *Biochemistry*, *50* (39), 8264-9.

74.   Johnson, K. A. (1995). Rapid quench kinetic analysis of polymerases, adenosinetriphosphatases, and enzyme intermediates. *Methods Enzymol, 249*, 38-61.

75.   Zahurancik, W. J., Klein, S. J., Suo, Z. (2013). Kinetic mechanism of DNA polymerization catalyzed by human DNA polymerase epsilon. *Biochemistry, 52*(40), 7041-7049.

76.   Zahurancik, W. J., Klein, S. J., Suo, Z. (2014). Significant contribution of the 3'-->5' exonuclease activity to the high fidelity of nucleotide incorporation catalyzed by human DNA polymerase. *Nucleic Acids Res, 42*(22), 13853-13860.

77.   Zahurancik, W. J., Baranovskiy, A. G., Tahirov, T. H., Suo, Z. (2015). Comparison of the kinetic parameters of the truncated catalytic subunit and holoenzyme of human DNA polymerase varepsilon. *DNA Repair (Amst), 29*, 16-22.

78. Joyce, C. M., Benkovic, S. J. (2004). DNA Polymerase Fidelity- Kinetics, Structure, and Checkpoints. *Biochemistry, 43*(45), 14317-14324.

79. Joyce, C. M., Potapova, O., DeLucia, A. M., Huang, X., Basu, Vandana Purohit, Grindely, N. D. F. (2008). Fingers-Closing and Other Rapid Conformational Changes in DNA Polymerase I (Klenow Fragment) and Their Role in Nucleotide Selectivity. *Biochemistry, 47*, 6103-6116.

80. Johnson, K. A. (1992). Transient-State Kinetic Analysis of Enzyme Reaction Pathways. *The Enzymes, 20*, 1-61.

81. Johnson, K. A. (1992). Conformational Coupling in DNA Polymerase Information Transfer. *Philosophical Transactions of the Royal Society*.

82. Tsai, Y., Johnson, K. A. . (2006). A New Paradigm for DNA Polymerase Specificity. *Biochemistry, 45*, 9675-9687.

83. Zhang, H., Cao, W., Zakharova, E., Konigsberg, W., De La Cruz, E. M. (2007). Fluorescence of 2-aminopurine revelas rapid conformation changes in the RB69 polymerase-primer/template complexes upon binding and incorporation of matched deoxynucleoside triphosphates. *Nucleic Acids Research, 35*(18), 6052-6062

84. Dangerfield, T. L., Johnson, K. A. (2020). Conformational dynamics during high fidelity DNA replication and translocation defined using a DNA polymerase with a fluorescent artificial amino acid. *J Biol Chem*

85. Raper, A. T., Reed, A. J., Suo, Z. (2018). Kinetic Mechanism of DNA Polymerases: Contributions of Conformational Dynamics and a Third Divalent Metal Ion. *Chem Rev, 118*(12), 6000-6025.

86. Zahurancik, W. J., Suo, Z. (2020). Kinetic investigation of the polymerase and exonuclease activities of human DNA polymerase epsilon holoenzyme. *J Biol Chem, 295*(50), 17251-17264.

87. Cerami, E. Gao, J., Dogrusoz, U., Gross, B. E., Sumer, S. O., Aksoy, B. A., Jacobsen, A., Byrne, C. J., Heuer, M. L., Larsson, E., Antipin, Y., Reva, B., Goldberg, A. P., Sander, C., Schultz, N. (2012). The cBio cancer genomics portal:

78. Joyce, C. M., Benkovic, S. J. (2004). DNA Polymerase Fidelity- Kinetics, Structure, and Checkpoints. *Biochemistry, 43*(45), 14317-14324.

79. Joyce, C. M., Potapova, O., DeLucia, A. M., Huang, X., Basu, Vandana Purohit, Grindely, N. D. F. (2008). Fingers-Closing and Other Rapid Conformational Changes in DNA Polymerase I (Klenow Fragment) and Their Role in Nucleotide Selectivity. *Biochemistry, 47*, 6103-6116.

80. Johnson, K. A. (1992). Transient-State Kinetic Analysis of Enzyme Reaction Pathways. *The Enzymes, 20*, 1-61.

81. Johnson, K. A. (1992). Conformational Coupling in DNA Polymerase Information Transfer. *Philosophical Transactions of the Royal Society*.

82. Tsai, Y., Johnson, K. A. . (2006). A New Paradigm for DNA Polymerase Specificity. *Biochemistry, 45*, 9675-9687.

83. Zhang, H., Cao, W., Zakharova, E., Konigsberg, W., De La Cruz, E. M. (2007). Fluorescence of 2-aminopurine revelas rapid conformation changes in the RB69 polymerase-primer/template complexes upon binding and incorporation of matched deoxynucleoside triphosphates. *Nucleic Acids Research, 35*(18), 6052-6062

84. Dangerfield, T. L., Johnson, K. A. (2020). Conformational dynamics during high fidelity DNA replication and translocation defined using a DNA polymerase with a fluorescent artificial amino acid. *J Biol Chem*

85. Raper, A. T., Reed, A. J., Suo, Z. (2018). Kinetic Mechanism of DNA Polymerases: Contributions of Conformational Dynamics and a Third Divalent Metal Ion. *Chem Rev, 118*(12), 6000-6025.

86. Zahurancik, W. J., Suo, Z. (2020). Kinetic investigation of the polymerase and exonuclease activities of human DNA polymerase epsilon holoenzyme. *J Biol Chem, 295*(50), 17251-17264.

87. Cerami, E. Gao, J., Dogrusoz, U., Gross, B. E., Sumer, S. O., Aksoy, B. A., Jacobsen, A., Byrne, C. J., Heuer, M. L., Larsson, E., Antipin, Y., Reva, B., Goldberg, A. P., Sander, C., Schultz, N. (2012). The cBio cancer genomics portal:

an open platform for exploring multidimensional cancer genomics data. *Cancer Discov 2*, 401-404

88.   Copeland, W. C., Ponamarev, M. V., Nguyen, D., Kunkel, T. A. & Longley, M. J. (2003). Mutations in DNA polymerase gamma cause error prone DNA synthesis in human mitochondrial disorders. *Acta Biochim Pol, 50*, 155-167.

89.   Singh, K. K., Ayyasamy, V., Owens, K. M., Koul, M. S. & Vujcic, M. (2009). Mutations in mitochondrial DNA polymerase-gamma promote breast tumorigenesis. *J Hum Genet, 54*, 516-524

90.   Sohl, C. D., Kasiviswanathan, R., Copeland, W. C., Anderson, K. S. (2013). Mutations in human DNA polymerase gamma confer unique mechanisms of catalytic deficiency that mirror the disease severity in mitochondrial disorder patients. *Hum Mol Genet 22*, 1074-1085

91.   Daee, D. L., Mertz, T. M., Shcherbakova, P. V. (2010). A cancer-associated DNA polymerase delta variant modeled in yeast causes a catastrophic increase in genomic instability. *Proc Natl Acad Sci U S A 107*, 157-162

92.   Freudenthal, B. D., Beard, W. A., Shock, D. D., Wilson, S. H. (2013). Observing a DNA polymerase choose right from wrong. *Cell, 154*(1), 157-168.

93.   Nakamura, T., Zhao, Y., Yamagata, Y., Hua, Y. J., Yang, W. (2012). Watching DNA polymerase eta make a phosphodiester bond. *Nature, 487*(7406), 196-201.

94.   Espinoza-Herrera, S. J., Gaur, V., Suo, Z., Carey, P. R. (2013). Following DNA chain extension and protein conformational changes in crystals of a Y-family DNA polymerase via Raman crystallography. *Biochemistry, 52*(29), 4881-4890.

2. Kinetic and Cellular Consequences of pH on IDH1 Activity

2.1 Abstract

Isocitrate Dehydrogenase 1 (IDH1) is an important metabolic enzyme that is responsible for the reversible NADP$^+$ dependent oxidative decarboxylation of isocitrate to α-ketoglutarate (αKG) and NADPH in the cytosol and peroxisomes of the cell. This reaction is critical because it provides anapleurotic intermediates and reductive potential in the form of NADPH to drive lipid biosynthesis and combat reactive oxygen species. The reverse reaction, the conversion of αKG to isocitrate, is an important step glutamine metabolism. Changes in the cellular environment, such as altered pH, substrate concentration levels, and oxidative stress can reroute metabolism by altering the activity of activity of metabolic enzymes like IDH1. One way that protein activity is regulated is through protein protonation, a post-translation modification (PTM) that can play an important role in protein-protein interactions, protein-ligand interactions, and protein stability. Amino acid residues that sense changes in surrounding pH, are termed pH sensors, and typically have pK_a values shifted towards physiological pH levels. The change in pK_a stores the potential energy required to drive a structural/functional modification that alters catalytic activity. Here, we show evidence for IDH1 pH dependence in the conversion of isocitrate to αKG in that k_{cat} values decreased with more acidic pH values. The kinetic parameters of the reverse reaction were found to be largely buffer-dependent. We coupled our biochemical analysis with a structural informatics algorithm, pHinder, to identify potential pH sensor at residue D273 at the dimer interface of IDH1. Point mutations at D273 led to decreased catalytic activity, and a loss of pH

dependent activity. This work identified a new pH-sensitive residue that is important for IDH1 catalysis.

2.2 Introduction

Isocitrate dehydrogenase 1 (IDH1) is a metabolic enzyme that has three distinct isoforms – IDH1, IDH2, and IDH3. IDH3 is a heterotetramer and is to the localized to the mitochondria. It catalyzes the third step in the tricarboxylic acid (TCA) cycle, where it uses a NAD^+ cofactor to convert isocitrate to α-ketoglutarate (αKG) and the high-energy product NADH. IDH1 and IDH2 are homodimers, catalyzing the reversible conversion of ICT to αKG, while utilizing a $NADP^+$/NADPH cofactor to drive the reaction (Figure 1A). IDH2 is found in the mitochondria, while IDH1 localizes to the peroxisomes and cytosol of the cell. The reactions of IDH1 and IDH2 are important to the cell because
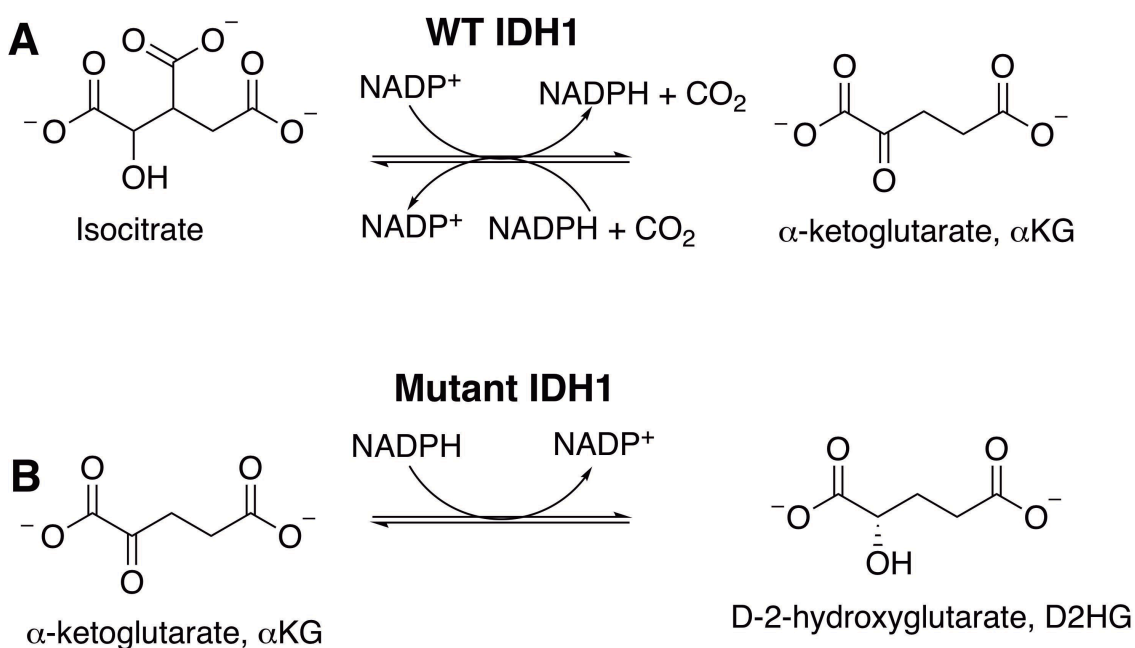
Figure 1: WT and mutant IDH1 catalytic activities. Shown are the **A.** normal oxidative decarboxylation reaction, and **B.** the mutant neomorphic reaction.

they provide metabolic intermediates to replenish the TCA cycle, provide NADPH as a cofactor for lipid and nucleic acid biosynthesis, and reductive power to combat reactive oxygen species.[1-3]

Tumor sequencing projects have found that IDH1 and IDH2 are frequently mutated in low grade gliomas, secondary glioblastomas; and in acute myeloid leukemia, chondrosarcomas, and cholangiocarcinomas to a lesser extent.[4-9] These studies found that IDH1 is commonly mutated at position R132, a residue responsible for coordinating the carboxylate moiety of isocitrate during catalysis. This position was most commonly mutated to a histidine or a cysteine, but tumors were also found to have glycine, leucine, and serine point mutations to a lesser extent.[4] These mutations drastically reduce normal IDH1 biochemical activity.[10, 11] Due to the loss of activity, it was originally thought that these mutations would simply lead to a loss of function in cancer. In addition, patients expressing R132H IDH1 mutations tended to be diagnosed at a younger age and live longer due to an epigenetic upregulation of the DNA damage response and greater resistance to radiation treatment.[12] However, the consequences on cellular metabolic activity still remained mysterious. In 2009, Dang and colleagues discovered that IDH1 is capable of converting αKG to *D*-2-hydroxyglutarate (D2HG) (Figure 1B).[13] D2HG is a proposed oncometabolite that inhibits histone demethylases and results in changes in gene expression and cell de-differentiation.[14, 15]

In addition, increased wild type (WT) IDH1 mRNA levels and protein expression levels are important features of some cancers.[4] In primary glioblastomas, glioblastomas arising *de novo*, IDH1 is not frequently mutated.[4, 15] However, wild-type (WT) IDH1 had increased expression and mRNA levels in 65% of primary glioblastoma in order to support

tumor progression.[15]   The upregulation of IDH1 is an example of metabolic reprogramming in cancer that increases lipid biosynthesis, alters the cellular redox state, promotes a more dedifferentiated cells state, and causes resistance toward receptor tyrosine kinase (RTK) targeted therapies.[15] Thus, IDH1 is a therapeutic target and the first selective small inhibitor for mutant IDH1 in newly diagnosed acute myeloid leukemias was approved by the FDA in 2019.[17, 18]

A change in the cellular environment brought on by hypoxia, changing metabolite concentrations, or disease can modulate protein activity in order to meet the cell's needs. For example, the conversion of isocitrate to αKG dominates at normal cellular conditions; however, in hypoxic conditions the reverse reaction of IDH is favored, though the mechanism remains unclear.[19-21]   Protein activity can be regulated through post-translational modifications (PTMs). Protein protonation is commonly overlooked as a PTM,[22] though many cellular processes are regulated by changes in intracellular pH (pH$_i$) including apoptosis,[23] oxidative stress,[24, 25] and cell differentiation.[26] Altered pH$_i$ is also associated with tumor metastasis.[22, 27, 28] The catalytic rates of some enzymes are pH-dependent. Notable examples include small GTPases,[29] the Na$^+$-H$^+$ proton pump,[30] and metabolic proteins such as phosphofructokinase,[31-33] and lactate dehydrogenase.[34, 35] Proteins can detect changes in pH based on the protonation state of ionizable amino acids residues that can have a biologically relevant change in protein-protein interactions, protein-ligand interactions, and protein stability.[22] Histidine residues are a natural candidate because their pK_a is already near physiological pH, but any ionizable residue can sense changes in intracellular pH (pH$_i$) if their pK_a values are shifted towards physiological pH in a folded protein.[22, 36] Previous work has shown that a large shift in

pK_a results in more potential energy to drive a conformational change to modify enzyme activity and fine-tune protein stability.[37, 38] Thus buried ionizable networks of amino acids play a critical role in protein stability and function.[39] There is some evidence that the forward and reverse reaction of IDH1 activity is pH dependent,[39-41] however, the steady-state parameters have not been reported.

In this work, we investigate IDH1 catalysis over a wide range of pH values and show evidence that the forward reaction rate is regulated by pH. The k_{cat} of the forward reaction shows the most consistent and reliable pH dependent trend in the tested conditions. The kinetic parameters of the reverse reaction are buffer-dependent. We also provide a possible mechanism for IDH1 pH dependent activity by identifying buried ionizable amino acids that have a shifted pK_a, D273 and K217, within the structure of IDH1. When the ionizability of IDH1 was removed by introducing point mutations at D273, there was a decrease in both activity and pH-sensitivity. This work overall combines experimental biochemical methods with structural informatics to identify pH dependent regulatory mechanism of IDH1 by D273.

2.3 Materials and Methods

2.3.1 Materials

Tris-hydrochloric acid, Tris Base, Bis-Tris, sodium hydroxide (NaOH), sodium chloride (NaCl), magnesium chloride (MgCl$_2$) hexahydrate, Dithiothreitol (DTT), β-Nicotinamide adenine dinucleotide phosphate (NADP$^+$) disodium salt, β-Nicotinamide adenine dinucleotide phosphate reduced (NADPH) trisodium salt, Potassium Phosphate (KPhos) monobasic and dibasic, potassium chloride (KCl), BL-21 Gold (DE3) Competent Cells, Luria-broth (LB)-Agar, kanamycin sulfate, Terrific Broth, 1-thio-β-D-

galactopyranoside (IPTG), EDTA-free Protease inhibitor Tablets, glycerol, Nickel-nitrilotriacetic acid (Ni-NTA) resin, Dulbecco's modified Eagle medium (DMEM), 2',7'-Bis-(2-carboxyethyl)-5-(and-6)-carboxyfluorescein acetoxymethyl ester (BCECF-AM), Dulbecco's phosphate buffered saline (DPBS), and Triton X-100 were all obtained from Fisher Scientific (Hampton, NH). β-Mercaptoethanol (β-ME) was obtained from MP Biomedicals in Santa Ana, CA. Isocitrate, alpha ketoglutarate (αKG), and imidazole were obtained from Acros Organics (Fisher Scientific Hampton, NH). Fetal bovine serum (FBS), was obtained from VWR (Radnor, PA). 5-(N-Ethyl-N-isopropyl)amiloride was obtained from Sigma-Aldrich (St Louis, MO). Esomperazole sodium salt was obtained from Apexbio (Houston, TX). Nigericin sodium was obtained from Tocris (Bristol, U.K). Stain free (4-20%) Sodium dodecyl sulfate polyacrylamide gel electrophoresis (SDS-PAGE) were obtained from Bio-Rad Laboratories (Hercules, CA).

2.3.2 Plasmid Mutagenesis

WT and R132H IDH1 clones were generously provided by Charles Rock (St. Jude's Hospital) for further study. All IDH1 cDNA constructs are in a pET-28b plasmid vector that contains an *N*-terminal hexa-histidine tag. Site-directed mutagenesis was used to generate single point mutations in WT IDH1 using the manufacturer's (Kappa Biosciences, Oslo, Norway) guidelines with the following primers: D273N (forward primer 5'-TTTATTTGGGCCTGCAAAAACTATAATGGTGATGTTCAGAGC, reverse primer 5'-GCTCTGAACATCACCATTATAGTTTTTGCAGGCCCAAATAAA); D273L (forward primer 5'-TGGTTTTATTTGGGCCTGCAAAAACTATCTGGGTGATGTTCAGAGCGA, reverse primer 5'-TCGCTCTGAACATCACCCAGATAGTTTTTGCAGGCCCAAATAAAACCA); D273S

50

(forward primer 5'-GTGGTTTTATTTGGGCCTGCAAAAACTATAGTGGTGATGTTCAGAG, reverse primer 5'-CTCTGAACATCACCACTATAGTTTTTGCAGGCCCAAATAAAACCAC); K217M (forward primer 5'-CTGAGCACCAAAAATACCATTCTGATGAAATACGATGGTCGCTTTAAAGATA, reverse primer 5'-TATCTTTAAAGCGACCATCGTATTTCATCAGAATGGTATTTTTGGTGCTCAG); K217Q (forward 5'-CTGAGCACCAAAAATACCATTCTGCAGAAATACGATGGTCGCTTTAAAGAT, reverse primer 5'-ATCTTTAAAGCGACCATCGTATTTCTGCAGAATGGTATTTTTGGTGCTCAG). Site-directed mutagenesis was used to generate single point mutations in R132H IDH1 background according to the manufacturer's guidelines (Kappa Biosciences, Oslo, Norway) using the following primers D273N (forward primer 5'-TTTATTTGGGCCTGCAAAAACTATAATGGTGATGTTCAGAGC, reverse primer 5'-GCTCTGAACATCACCATTATAGTTTTTGCAGGCCCAAATAAA); D273L (forward primer 5'-TGGTTTTATTTGGGCCTGCAAAAACTATCTGGGTGATGTTCAGAGCGA, reverse primer 5'-TCGCTCTGAACATCACCCAGATAGTTTTTGCAGGCCCAAATAAAACCA); D273S (forward primer 5'-GTGGTTTTATTTGGGCCTGCAAAAACTATAGTGGTGATGTTCAGAG, reverse primer 5'-CTCTGAACATCACCACTATAGTTTTTGCAGGCCCAAATAAAACCAC). All constructs were sequenced to confirm accuracy. Dr. Christal Sohl designed necessary

51

mutagenesis primers, and I performed all mutagenesis to generate single mutants and double mutants.

2.3.3 Protein Expression and Purification

IDH1 (WT and mutants) were transformed into BL-21 Gold (DE3) cells. The transformed cells were incubated at 37˚C in 0.5-2.0 L of Terrific Broth (TB) containing 30 μg/mL kanamycin at 200 RPM until reaching an optical density at 600 nm of 1.0-1.2. Then, the temperature was dropped 18.0˚C and shaking reduced to 130 RPM. The cultures were induced with a final concentration of 1mM IPTG. The cells were harvested by centrifugation at 8000 RPM for 15 minutes, cell pellets were resuspended in lysis buffer (20 mM Tris pH 7.5 at 4˚C, 500 mM NaCl, 0.1% Triton X-100, and supplemented with a crushed and dissolved EDTA-free protease inhibitor tablet). The cells were lysed via sonication and the supernatant was collected by centrifugation of the homogenates at 12000 RPM for 1 hour. The resulted supernatant was loaded into a Ni-NTA affinity column and washed with Ni-buffer 1 (50 mM Tris pH 7.5 at 4˚C, 500 mM NaCl, and 5 mM β-ME). Bound protein was eluted with Ni-Buffer 2 (50 mM Tris pH 7.5 at 4˚C, 500 mM NaCl, 500 mM imidazole at pH 7.5, and 5 mM β-ME). The elution was dialyzed in storage buffer (50 mM Tris pH 7.5 at 4˚C, 100 mM NaCl, 20% glycerol, and 1 mM DTT). Purity of IDH1 WT and mutants was confirmed to be over 95% via SDS-PAGE gels (4-20%). The concentration of purified proteins was calculated via obtaining the absorbance at 280 nm using a calculated molar extinction coefficient of 64143 $M^{-1}cm^{-1}$. The proteins were flash frozen in liquid nitrogen and stored at -80˚C for $\leq$ 1 month. I purified IDH1 WT, single mutants, and double mutants. To ensure that WT and mutant IDH1 dimerized to a similar degree and did not oligomeric aggregates in solutions, WT and D273L IDH1

were analyzed using size exclusion chromatography (SEC) and were both found to be primarily in the dimer form (Supplemental Figure S1). I purified all IDH1 WT proteins, IDH1 single mutants, and IDH1 double mutants.

2.3.4 Steady State Activity Assays

The activity of IDH1 homodimers were assessed at 37°C using steady-state methods described previously.[11] An 8452 diode array spectrophotometer (Olis, Atlanta, GA) was used to measure the steady state kinetic assays. The Tris-based assays were prepared as follows: for the conversion of isocitrate to αKG for pH > 7.0, a cuvette containing assay buffer (50 mM Tris buffer ranging from pH 8.0 to 7.2 at 37°C, 150 mM NaCl, 10 mM $MgCl_2$, 0.1 mM DTT), and 100 nM IDH1, were preincubated for 3 minutes at 37°C. For the conversion of isocitrate to αKG for pH $\leq$ 7.0 a cuvette containing assay buffer (50 mM bis-Tris ranging from pH 7.0 to 6.2 at 37°C, 150 mM NaCl, 10 mM $MgCl_2$, 0.1 mM DTT), and 100 nM WT IDH1 were preincubated for 3 minutes at 37°C. For the conversion of αKG to isocitrate for pH > 7.0,[39] a cuvette containing assay buffer (50 mM Tris ranging from 8.0 to 7.2 at 37°C, 150 mM NaCl, $MgCl_2$, 66 mM $NaHCO_3$), and 100 nM IDH1, were preincubated for 3 minutes at 37°C. For the conversion of αKG to isocitrate for $_{pH} \leq$ 7.0 a cuvette containing assay buffer (50 mM bis-Tris pH ranging 7.0-6.2 at 37°C, 150 mM 150 NaCl, 10 mM $MgCl_2$, 66 mM $NaHCO_3$) and 100 nM IDH1, were preincubated for 3 minutes at 37°C.

The potassium phosphate (KPhos) buffer-based assays were prepared as follows: for the conversion of isocitrate to αKG, a cuvette containing assays buffer (50 mM KPhos ranging from pH 8.0 to 6.2 at 37°C, 150 mM KCl, 10 mM $MgCl_2$, 0.1 mM DTT), and 100 nM IDH1 were preincubated for 3 minutes at 37°C. For the conversion of αKG to

isocitrate, a cuvette containing assay buffer (50 mM KPhos pH ranging from 8.0 to 6.2 at 37°C, 150 mM KCl, 10 mM MgCl$_2$, 66 mM NaHCO$_3$, and 0.1 mM DTT), and 100 nM IDH1 were preincubated for 3 minutes at 37°C.

The pH of αkG was adjusted to the pH of the assay before use. In reactions measuring isocitrate to αKG conversion, reactions were initiated by adding NADP$^+$ and isocitrate, with saturating NADP$^+$ and titrating in various concentrations of isocitrate. In reactions measuring αKG to isocitrate conversion, reactions were initiated by adding NADPH to αKG, with saturating NADPH and titrating in various concentrations of αKG. In the R132H background, reactions measuring the conversion of αKG to D2HG, reactions were initiated by adding NADPH and αKG, with saturating NADPH and titrating in various concentrations of αKG. In all cases change in absorbance due to changing NADPH concentrations was monitored at 340 nm.

For all reactions, the slope of the linear ranges of the assays were calculated to μM NADPH using the molar extinction coefficient for NADPH of 6.22 cm^{-1} mM^{-1} to obtain k_{obs} (i.e μM NADPH/μM enzyme s^{-1} at each substrate concentration). Results were fit to hyperbolic plots in GraphPad Prism (GraphPad Software, La Jolla, CA) to estimate k_{cat} and K_m values ± standard error (SE). I generated all final steady state kinetic data for the conversion of isocitrate to αKG in Tris, bis-Tris, and KPhos buffers from pH 8.0 – 6.2 for WT IDH1, D273L, D273N, D273S, K217M, K217Q IDH1 single mutants; the conversion of αKG to isocitrate in Tris, bis-Tris, and KPhos buffers from pH 8.0 – 6.2 for WT IDH1; and the conversion of αKG to D2HG for D273L, D273N, and D273S IDH1 R132H double mutants. Zach Lesecq performed preliminary steady-state experiments in Tris-based buffers and provided some training.

2.3.5 Thermal Stability Using Circular Dichroism

The stability of WT IDH1 in the pH range from 8.0 to 6.2, and D273L IDH1 at pH 7.5 were assessed using circular dichroism (CD). IDH1 was diluted to 5 μM in a buffer containing 10 mM KPhos at desired pH and 100 mM KCl. The thermal melt experiment was initiated at 20°C and the temperature increased to 70°C in 1°C increments. The secondary structure of IDH1 is rich in α-helices and was monitored via 222 nm peak, corresponding to α-helicity. Analysis was performed using GraphPad Prism (GraphPad Software, La Jolla, CA) with a Boltzmann Sigmoidal fit.[42] I performed the CD stability analysis for WT IDH1 at all tested pH values and IDH1 D273L at pH 7.5.

2.3.6 Cellular pH$_i$ Modulations

HT1080 cell lines were all cultured in DMEM supplemented with 10% FBS and incubated at 37°C with 5% CO_2 atmosphere. The effects of proton pump inhibitors on pH$_i$ was evaluated in triplicate by first adding 0.3×10^5 cells per well in a 24-well plate and incubated overnight. The following day the proton pump inhibitor, esomeprazole sodium salt, was dissolved in PBS and was added to the cells at a final concentration of 200 μM, and the cells were again incubated at 37°C and 5% CO_2. After 16-24 hours, the pH$_i$ was measured by first loading cells with 1 μM BCECF-AM dissolved in DMSO (final cellular concentration of DMSO at 0.1%) at 37°C and 5% CO_2 for 15 minutes. Excess dye was removed by washing the cells three times with bicarbonate buffer (25 mM bicarbonate, 115 mM NaCl, 5 mM KCl, 10 mM glucose, 1 mM $MgSO_4$, 1 mM $KHPO_4$ pH 7.4 at 37°C, and 2 mM $CaCl_2$) with esomeprazole to preserve pH$_i$ manipulation conditions. After the wash steps, the fluorescence intensity of the BCECF-AM probe was detected with dual excitation (490 nm and 440 nm) and a fixed emission wavelength of 535 nm.

Fluorescence ratios (490 nm/440 nm) were converted to pH_i values by calibrating each experiment with a standard curve ranging from pH 6.5 to 7.5 using 25 mM HEPES, 105 mM KCl, and 1 mM $MgCl_2$, and 10 μM nigericin to equilibrate the pH_i with the extracellular pH. Metabolites were derivatized and quantified using gas chromatography – mass spectrometry (GC-MS) as described previously by Dr. David Scott and Dr. Olga Zagnitko (Sanford Burnham Prebys).[42] Dr. Katharine White (Notre Dame) and Dr. Diane Barber (University of California San Francisco) aided in the design and training to perform them. I performed the intracellular pH modulation assays.

2.3.7 GC/MS Analysis of ICT formation and Cellular Metabolite Quantification

To confirm that WT IDH1 was performing the reverse reaction at all tested pH's GC-MS was used as described previously[43] (data not shown) to confirm the presence of isocitrate. Briefly, 100 nM IDH1 was incubated in KPhos reaction buffer at the desired pH, and the reaction was initiated by the addition of saturating NADPH and αKG. The absorbance was monitored at 340 nm until completion, and an aliquot of the reaction was removed. Then, the aliquot was lyophilized to dryness. The dried aliquots were redissolved in 50% methanol containing L-norvaline as an internal standard, and re-dried using a lyophilizer. L-norvaline, αKG, and ICT solutions were prepared to use as standards. The standards and incubations were derivatized with O-isobutylhydroxylamine hydrochloride and N-ter-butyldimethylsilytrifluoroacetamide and analyzed using GC/MS by Dr. David Scott and Dr. Olga Zagnitko (Sanford Burnham Prebys) .

For cellular metabolite quants, cells were initially seeded to a 6-well plate and grown to about 1.0×10^6 cells/well. Media was removed and cells were washed 3 times

with cold DPBS, followed by the addition of 0.45 mL cold methanol (50% v/v in water) with 20 μM L-norvaline as an internal standard to each well. Culture plates were then frozen on dry ice for 30 minutes. After thawing on ice, the methanol extract was transferred to a centrifuge tube, and cellular metabolites were analyzed via GC/MS at Sanford Burnham Prebys Core Facility by Dr. David Scott and Dr. Olga Zagnitko.

2.3.8 ITC Measurements

Isothermal titration calorimetry (ITC) was performed in Sanford Burnham Prebys Protein Analysis Core using a Low Volume Affinity ITC calorimeter (TA instruments). 3.0 – 6.0 μL aliquots of solution containing 0.15 mM AGI-5198 or ML309 were injected in the cell containing 0.03 – 0.05 mM protein. Twenty injections were made. The experiments were performed at 23°C in buffer containing 20 mM Tris pH 7.5, 100 mM NaCl, 2 mM β-mercaptoethanol, and 0.2 mM NADPH. Baseline control data were collected injecting ligand into the cell containing buffer only. ITC data were analyzed using Nanoanalyze software provided by TA instruments. ITC experiments were performed by Dr. Andrey Bobkov (Sanford Burnham Prebys).

2.3.9 pHinder Algorithm

The details of the pHinder algorithm have been described previously.[29, 33, 44] Briefly, protein structures were downloaded from the Protein Data Bank (PDB) and the algorithm was used to calculate ionizable residue networks. First, a Delaunay triangulation was calculated using the terminal side-chain atoms of all ionizable residues (aspartate, glutamate, histidine, cysteine, lysine, and arginine) and minimized by removing network edges longer than 10 Å. The results were further simplified by removing redundant connections. Second, each ionizable residue was assigned a

classification of either buried (> 3.0 Å below the surface), margin (< 3.0 Å below the surface < 1.0 Å above the surface), or exposed (> 1.0 Å above the surface). Depth of burial was determined by measuring the minimum distance between the ionizable group and the triangular facets of the pHinder-calculated surface. Dr. Daniel Isom performed the pHinder algorithm calculation and analyzed the results with Dr. Christal Sohl. Dr. Christal Sohl identified amino acids for further testing.

2.3.10 Computational Methods

PROPKA[45, 46] was used to predict the protonation states of amino acid residues in IDH1 using the PDB 1T0L structure.[47] Due to the instability of the α10 helix in mutant IDH1, and the structural similarity between mutant IDH1 and WT IDH1 holoenzyme forms, the WT background was selected for modeling. These structures (1T0L and 1T09)[47] were used to generate models of D273N, D273L, and D273S in the WT IDH1 with NADP$^+$ bound and WT IDH1 with isocitrate, NADP$^+$, and Ca^{2+} bound backgrounds. Mutations were made with Coot software[48] and energy minimized structures were made using Phenix software.[49] Five hundred maximum iterations and five macrocycles were used, with bond lengths, bond angles, nonbonded distances, dihedral angles, chirality, parallelity, and planarity considered. Dr. Jamie Schiffer performed this computational analysis to predict pK$_a$ values and performed ML309 docking, and Dr. Christal Sohl performed the energy minimization and wrote the paper. All authors contributed to editing.

2.4 Results

2.4.1 Effect of pH Modulation on WT IDH1 Activity

Human IDH1 was expressed and purified form *E. coli* and kinetically characterized under steady-state conditions to determine the effect of pH on catalytic activity. The k_{cat}

of the conversion of isocitrate to αKG (Figure 1A) showed the clearest dependence on

pH and increased with increasingly basic pH (Figure 2A, Table 1).
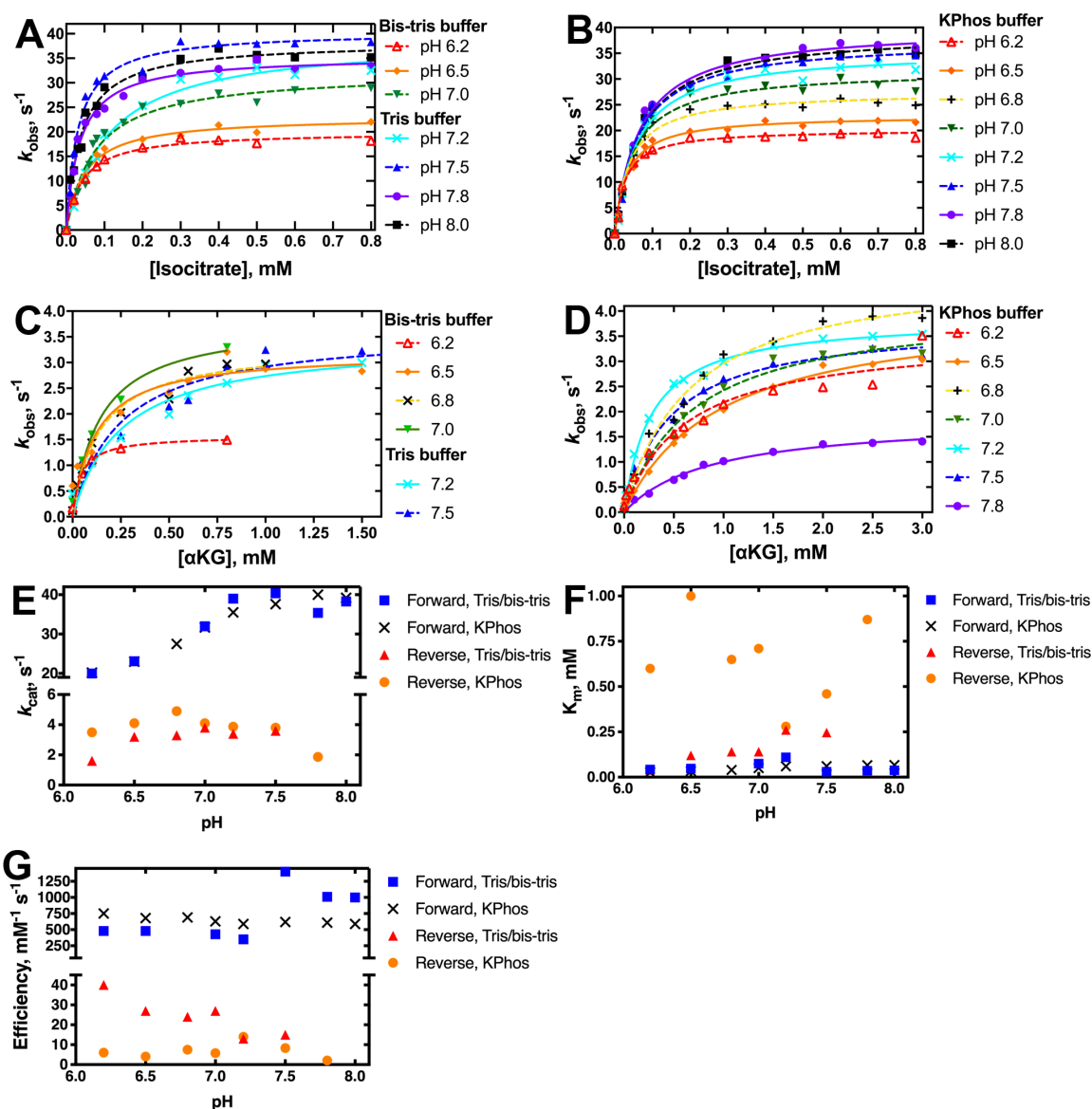
Figure 2: conversion of isocitrate to αKG, catalyzed by WT IDH1 were measured as function of varying pH using either Tris (pH 7.2-8.0) or bis-Tris (pH 6.2-7.0) buffers (A), or KPhos buffer (B). Steady-state rates of the reverse reaction, the conversion of αKG to isocitrate, catalyzed by WT IDH1 were measured as a function of varying pH using either Tris or bis-Tris buffers (C), or KPhos buffer (D). In (A-D), at least two protein preparations were used to measure the observed rate constants (k_{obs}) at varying substrate concentrations, which were calculated from determining the linear portion of plots of concentration versus time. Each point in the curve represents a single replicate. The variation in k_{cat} (E), K_m (F), and k_{cat}/K_m (efficiency) (G) due to change in pH from parts (A-D) is shown.

However, variability in K_m minimized this trend when considering enzymatic efficiency

(Figure 2E-G). This pH-dependent trend in k_{cat} was also observed in KPhos buffer (Figure

2B, Table 1). Again, however, the variability in K_m values ruined any trend in catalytic

efficiency (Figure 2E-G).

Table 1: Assessment of IDH1 catalysis upon varying pH. Kinetic parameters for the normal forward and reverse reactions catalyzed by WT IDH1 in varying pH and buffers are shown. Steady-state rates were derived from fiting plots to k_{obs} versus substrate concentration with the Michaelis-Menten equation, with the standard error of the measurement (S.E.M.) determined from the deviance from these hyperbolic fits. Data were obtained as described in Figure 2.

IDH1	Buffer	Reaction	k_{cat}, s^{-1}	K_m, mM	k_{cat}/K_m, mM^{-1}s^{-1}
WT	Bis-Tris, pH 6.2	Isocitrate → αKG	20.0 ± 0.4	0.042 ± 0.004	(0.48 ± 0.05) × 10^3
WT	Bis-Tris, pH 6.5	Isocitrate → αKG	23.1 ± 0.6	0.048 ± 0.006	(0.48 ± 0.06) × 10^3
WT	Bis-Tris, pH 7.0	Isocitrate → αKG	32 ± 1	0.075 ± 0.009	(0.43 ± 0.05) × 10^3
WT	Tris, pH 7.2	Isocitrate → αKG	39 ± 1	0.11 ± 0.01	(0.35 ± 0.03) × 10^3
WT	Tris, pH 7.5	Isocitrate → αKG	40.4 ± 0.8	0.030 ± 0.003	(1.4 ± 0.1) × 10^3
WT	Tris, pH 7.8	Isocitrate → αKG	35.4 ± 0.6	0.035 ± 0.003	(1.01 ± 0.09) × 10^3
WT	Tris, pH 8.0	Isocitrate → αKG	38.3 ± 0.9	0.038 ± 0.004	(1.0 ± 0.1) × 10^3
WT	KPhos, pH 6.2	Isocitrate → αKG	20.2 ± 0.4	0.027 ± 0.003	(0.75 ± 0.08) × 10^3
WT	KPhos, pH 6.5	Isocitrate → αKG	23.0 ± 0.4	0.034 ± 0.003	(0.68 ± 0.06) × 10^3
WT	KPhos, pH 6.8	Isocitrate → αKG	27.5 ± 0.6	0.040 ± 0.004	(0.69 ± 0.07) × 10^3
WT	KPhos, pH 7.0	Isocitrate → αKG	31.7 ± 0.8	0.050 ± 0.006	(0.63 ± 0.08) × 10^3

Table 1: Assessment of IDH1 catalysis upon varying pH, continued

IDH1	Buffer	Reaction	k_{cat}, s^{-1}	K_m, mM	k_{cat}/K_m, mM^{-1}s^{-1}
WT	KPhos, pH 7.2	Isocitrate → αKG	35.5 ± 0.8	0.060 ± 0.006	(0.59 ± 0.06) × 10^3
WT	KPhos, pH 7.5	Isocitrate → αKG	37.6 ± 0.7	0.061 ± 0.005	(0.62 ± 0.05) × 10^3
WT	KPhos, pH 7.8	Isocitrate → αKG	40.0 ± 0.7	0.066 ± 0.005	(0.61 ± 0.05) × 10^3
WT	KPhos, pH 8.0	Isocitrate → αKG	39.2 ± 0.7	0.067 ± 0.005	(0.59 ± 0.05) × 10^3
WT	Bis-Tris, pH 6.2	αKG → isocitrate	1.6 ± 0.1	0.04 ± 0.01	40 ± 10
WT	Bis-Tris, pH 6.5	αKG → isocitrate	3.2 ± 0.2	0.12 ± 0.04	27 ± 9
WT	Bis-Tris, pH 6.8	αKG → isocitrate	3.3 ± 0.2	0.14 ± 0.03	24 ± 5
WT	Bis-Tris, pH 7.0	αKG → isocitrate	3.8 ± 0.3	0.14 ± 0.04	27 ± 8
WT	Tris, pH 7.2	αKG → isocitrate	3.4 ± 0.3	0.26 ± 0.07	13 ± 4
WT	Tris, pH 7.5	αKG → isocitrate	3.6 ± 0.2	0.246 ± 0.06	15 ± 4
WT	KPhos, pH 6.2	αKG → isocitrate	3.5 ± 0.3	0.6 ± 0.2	6 ± 2
WT	KPhos, pH 6.5	αKG → isocitrate	4.1 ± 0.2	1.0 ± 0.1	4.1 ± 0.5
WT	KPhos, pH 6.8	αKG → isocitrate	4.9 ± 0.2	0.65 ± 0.08	7.5 ± 0.1

Table 1: Assessment of IDH1 catalysis upon varying pH, continued

IDH1	Buffer	Reaction	k_{cat}, s^{-1}	K_m, mM	k_{cat}/K_m, mM^{-1}s^{-1}
WT	KPhos, pH 7.0	αKG → isocitrate	4.1 ± 0.2	0.71 ± 0.09	5.8 ± 0.8
WT	KPhos, pH 7.2	αKG → isocitrate	3.87 ± 0.07	0.28 ± 0.02	14 ± 1
WT	KPhos, pH 7.5	αKG → isocitrate	3.8 ± 0.1	0.46 ± 0.05	8.3 ± 0.9
WT	KPhos, pH 7.8	αKG → isocitrate	1.87 ± 0.08	0.87 ± 0.09	2.1 ± 0.2

The reverse reaction of IDH1 was much less efficient, driven largely by a decrease in k_{cat} and by an increase in K_m, particularly in KPhos buffer (Figure 2, Table 1). Here, any pH-dependent trend was limited to studies in Tris/bis-Tris buffers (Figure 2C), as only a slight trend was observed in k_{cat} in KPhos (Figure 2D). A modest trend was seen for the reverse reaction in terms of catalytic efficiency (Figure 1G, Table 1) due to a small increase in K_m as the pH increased in Tris/bis-Tris buffers, and limited changes in k_{cat}. Though in general, pH-dependent trends observed for the reverse reaction seem to be small, inconsistent, and buffer-dependent.

To confirm that changes in IDH1 catalytic activity were due to pH and not changes in overall secondary structure, I used circular dichroism (CD) to show that IDH1 secondary structure remained stable at all tested pH values. These results (Supplemental Figure S2) showed that the T_m value did not change significantly across the range of pH values. Therefore, we concluded that the forward reaction exhibited the strongest and most consistent trends in the tested conditions, compared to the modest

trends observed in the reverse reaction. We thus moved forward with characterizing the forward reaction in KPhos buffer.

2.4.2 Modulation of pH in Cell Lines

In the cell, small changes in intracellular pH are noted to occur during a variety of processes including apoptosis,[23] hypoxia,[50] neurodegenerative diseases,[51, 52] cancer,[22, 27, 28, 53] and diabetes.[54]   Since an increase in k_{cat} was observed for the conversion of isocitrate to αKG by IDH1 with increasing pH in biochemical assays, we aimed to explore the effect of acidic cellular pH had on changes in IDH1-relevant metabolite levels.

We decreased the intracellular pH of cells by treating them with proton transport inhibitors,[55, 56] and then quantified metabolite levels using GC/MS.   Naturally, these experiments are limited because of the many redundant cellular pathways that produce and consume isocitrate and αKG. To minimize these effects as much as possible, we selected isogenic cell lines that had varying levels of WT IDH1: patient-derived HT1080 cells containing either an endogenous R132C mutation (R132C/+IDH1, where +IDH1 denotes IDH1 WT) or with a version of this cell line that is R132C-ablated and also stably overexpresses WT IDH1 (–/+++ IDH1).[57]   The R132C mutation results in the ability to produce D2HG and WT IDH1 activity is removed in protein expressed from this allele.[58-60] A variety of proton pump inhibitors, including esomeprazole (ESOM), which targets the V-ATPase proton pump, and other ion exchange inhibitors like 4,4'-diisothiocyano-2,2'-stilbenedisulfonic acid (DIDS) and 5-(N-ethyl-N-isopropyl)amiloride (EIPA), were tested at varying concentrations to find the effective concentrations that led to a decrease in intracellular pH without an observed increase in cell death (Supplemental Figure S3). For

the HT1080 cell lines, 200 μM ESOM treatment was found to be the most effective in

meeting those criteria (Table 2, Figure 2A).

Isocitrate levels in HT1080 cells dropped significantly upon treatment with ESOM

(Table 2, Figure 3B), though these values were near the limit of detection. αKG levels
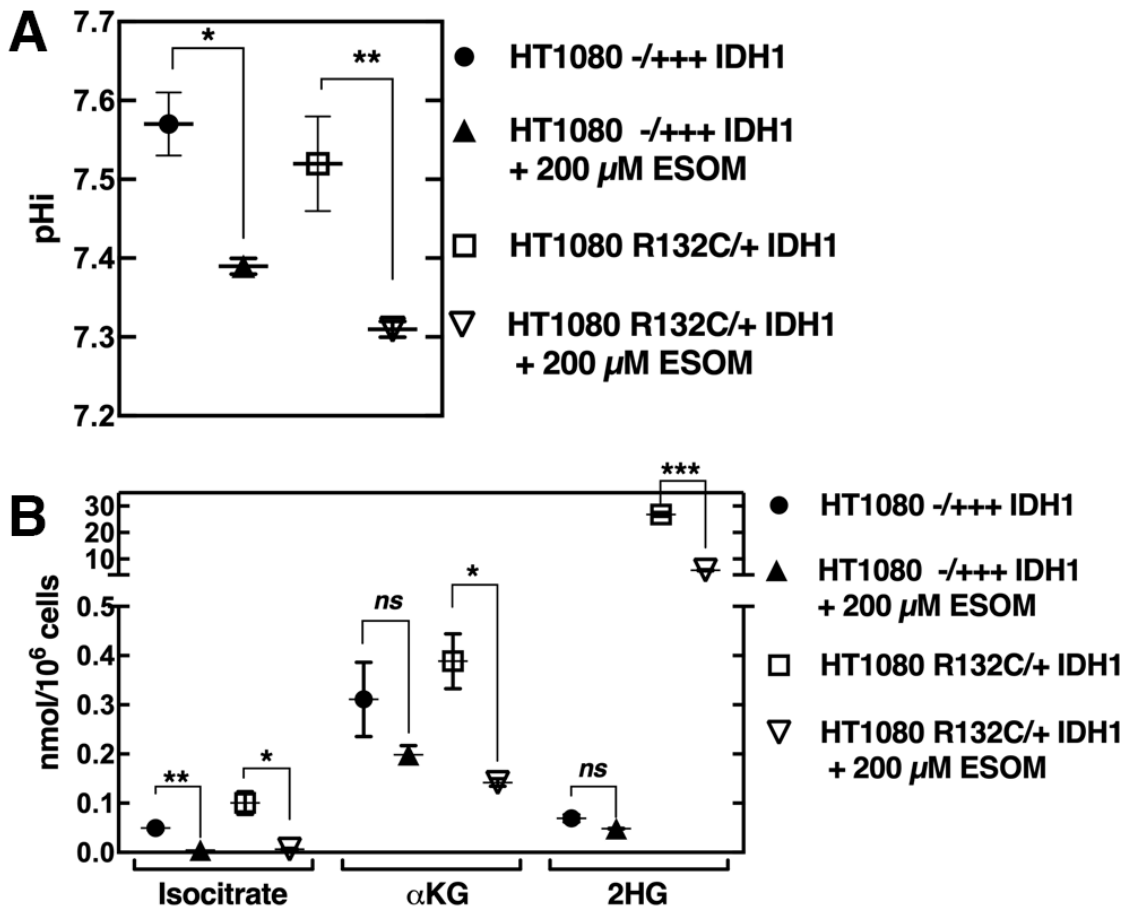
Figure 3: Change in IDH1-related metabolites upon pHi modification. Two sets of cell lines were used: patient-derived HT1080 cells, which contain an endogenous heterozygous R132C IDH1 mutation (R132C/+ IDH1), or HT1080 cells where the R132C IDH1 allele was ablated with a T77A mutation to destroy neomorphic activity (αKG to D2HG production) in the R132C allele, and made to stably overexpress WT IDH1 (-/+++ IDH1) [80]. Cell lines were either untreated or treated with 200 μM ESOM to decrease the pHi. The resulting change in pH are shown in (A). Experiments were performed as technical triplicates. (B) Metabolites were then quantified and changes in isocitrate, αKG, and 2HG (isomers unresolved) are highlighted (see Table S1 for additional metabolites). Experiments were performed as biological duplicates. Not significant (ns) $p > 0.05$, *$p ≤ 0.05$, **$p ≤ 0.01$, ***$p ≤ 0.001$.

also decreased upon a shift to an acidic intracellular pH in HT1080 cell lines (Table 2,

Figure 3B), though statistical significance was not achieved with HT1080 –/+++ IDH cells.

A decrease in αKG concentration in our cellular assays was not surprising because our biochemical assays showed a decrease in k_{cat} in more acidic pH values, but many proteins contribute to αKG levels, including the other IDH isoforms (IDH2 and IDH3). Interestingly, a global decrease in metabolite levels was not observed in cells upon ESOM treatment. Metabolite quantities both increased and decreased (Supplemental Table S1), thus minimizing the probability that ESOM was simply cytotoxic. 2HG levels in the mutant line (*D* and *L* isomers cannot be resolved) were also decreased upon ESOM treatment.

Table 2: Cellular quantitation of metabolites related to IDH1 activity at varying pH, biological replicates of two.

Cell line	Treatment	pH$_i$	[Isocitrate], nmol/10^6 cells	[αKG], nmol/10^6 cells	[2HG], nmol/10^6 cells
HT1080 (-/+++ IDH1)	No treatment + ESOM	7.57 ± 0.04 7.39 ± 0.01	0.05 ± 0 ≤ 0.01	0.31 ± 0.08 0.20 ± 0.02	0.07 ± 0.01 0.05 ± 0
HT1080 (R132C/+ IDH1)	No treatment + ESOM	7.52 ± 0.06 7.31 ± 0.01	0.10 ± 0.02 0.01 ± 0.0	0.39 ± 0.06 0.14 ± 0.01	26.9 ± 0.7 5.8 ± 0.2

2.4.3 Characterizing IDH1 Ionizable Networks

To identify a mechanism for IDH1 pH dependence in the conversion of isocitrate to αKG, we collaborated with Dr. Dan Isom to use a computational geometric algorithm known as pHinder[29, 33, 44] to find potential pH-sensing residues in IDH1. Briefly, this spatial algorithm uses Delaunay Triangulation to identify networks of ionizable residues (i.e aspartate, glutamate, histidine, cysteine, lysine, and arginine), calculates a surface based

on the volume of the folded protein, and classifies a residue as buried based on the distance to the calculated surface. The algorithm then removes redundant networks and ensures that the networks are continuous. Dr. Sohl examined the results of the structural informatics search and chose the most consistently buried residues for further study. These networks can then predict structure-function relationships. Buried stretches of basic or acidic amino acids within proteins tend to have their pK_a values shifted into the physiological pH range (5 – 8) and can sense small fluctuations in pH, triggering structural and functional changes.[33, 61, 62] Such residues would be termed pH-sensors.

A large of number of buried acidic or basic networks were found when analyzing previously solved structures of apoenzyme WT IDH1 with $NADP^+$ bound (1T09[47]); holoenzyme WT IDH1 with $NADP^+$, isocitrate, and Ca^{2+} (1T0L[47]); and mutant R132H IDH1 bound to NADPH, αKG, and Ca^{2+} (4KZO[63]) (Figure 4 and Supplemental Figure S4). The calcium ion mimics the magnesium cofactor required for catalysis. To identify amino acids residues within the network, we used the notation X###, where X represents the single letter code of the amino acid and ### denotes the residue number. Since IDH1 is a dimer, we distinguished between the two monomeric chains using X### and X###'. Point mutations were thus represented by X###Y, meaning amino acid X at position ### had been mutated to Y.

Given the large number of and complexity of the buried ionizable networks discovered after examining the three IDH1 structures, we eliminated residues that did not have a completely buried side chain and chose to focus only on two that appeared consistently throughout the analysis: a short acidic network running through the dimer interface and a longer basic network within each IDH1 monomer (Figure 4B and

Supplementary Figure 4). The critical component of the acidic network involves residues

D273 and D273' found at the dimer interface and are too distant from the active site to

form any direct interactions with catalytic substrates and cofactors (Figure 4B, 4D). The

D273 residues also participated in a an ionizable network in the structure of holo R132H

IDH1 (Figure 4E). Interestingly, D273 is close to where mutant IDH1 inhibitors bind

(Figure 4F). In contrast with the short acidic network, we also identified a longer basic

network in holoenzyme WT IDH1. This basic network crossed the active site, and some

residues were partially exposed. To avoid direct mutational effects on substrate binding

or chemistry, we chose to focus on the most deeply buried residues, K217', that was more

distant from the active site. The A-chain K217 also appeared in an additional network

involving many of the residues in the B-chain, though it was a bit longer (Figure 4C).

Thus, we selected D273 and K217 for further analysis, and proceeded to calculate

the pK_a value of each residue in the holoenzyme structure of WT IDH1[47] using PROPKA[45,
46] in collaboration with Dr. Jamie Schiffer. The goal of the calculations was to provide an

estimate of the D273 and K217 pK_a values (Supplemental Table S2). K217 had a

calculated pK_a value of 7.97, K217' had a calculated pK_a value of 8.24, consistent with

previous findings showing buried lysine residues have downshifted pK_a values that can

fine tune protein stability.[36, 38] D273 and D273' had their pK_a values upshifted to 6.59 and

6.4 respectively.

Based on these observations with computational models and pK_a calculations, we

hypothesized that K217 and D273 in IDH1 had features consistent with pH-sensing

residues. We sought to confirm the role of D273 and K217 in IDH1 catalysis biochemical

experiments.
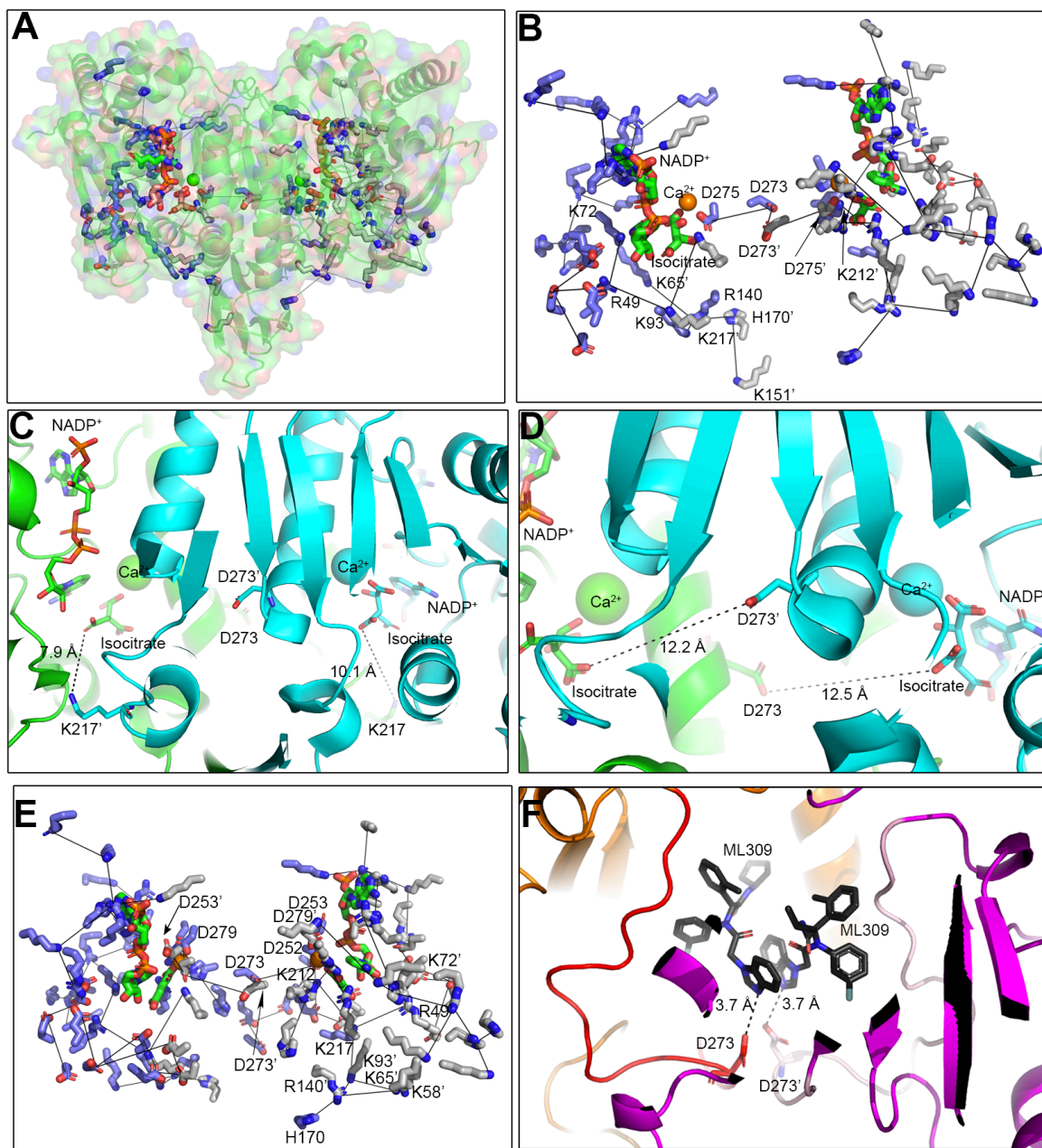
Figure 4: pHinder analysis of the holo IDH1 dimer. (A) Networks of buried ionizable residues were identified using pHinder in a previously solved structure of holo WT IDH1 bound to NADP+, isocitrate, and Ca2+ (1T0L).[47] These networks are shown as grey lines and with the residues making up this network shown as purple (chain A) or grey (chain B) sticks. (B) Only the residues involved in this network, as well as isocitrate and NADP+ (shown in green) are highlighted. (C) The location of K217 (chain A) and K217' (chain B) relative to the active sites. (D) The location of D273 (chain A) and D273' (chain B) relative to the active sites. (E) pHinder was also used to identify a network of buried ionizable residues in a previously solved structure of holo R132H IDH1 bound to αKG, NADP+, and Ca2+ (4KZO).[63] For clarity, only residues described in the text as being part of networks involving K212 and D273 are labeled in part B and E. (F) Two ML309 ligands were docked into the density present in a cryo-EM structure of R132C IDH1 incubated with ML309, and the localization of D273 relative to these inhibitors is shown.[59, 104]

68

2.4.4 Residue K217 in IDH1 Has a Modest Role in Catalysis

The K217 residue is found in a loop that lies 8-10 Å from isocitrate in the active site (Figure 4C). This residue is predicted to have a downshifted pK_a value in both chains, from ~10.7 in solution to 7.97 and 8.24 in IDH1. We selected point mutations that remove ionizability, while introducing minimal structural and steric perturbations to IDH1. Methionine, a nonpolar residue, and glutamine, a polar uncharged residue, were selected. Recombinant K217M and K217Q IDH1 were expressed and purified from *E. coli*, and the catalytic efficiency of the conversion of isocitrate to αKG was measured in steady-state kinetic conditions. K217M IDH1 had no effect on k_{cat} and only a modest increase in K_m, yielding a 2.2-fold decrease in catalytic efficiency (Figure 5A, Table 3). K217Q was only slightly more disruptive to normal IDH1 catalytic activity, and resulted in a 5.4-fold decrease in catalytic efficiency, largely driven by a 4-fold increase in K_m (Figure 5A, Table 3). Overall, ionizability at this residue did not appear vital for catalysis, though the mutation did affect activity.

2.4.5 Apparent Role of D273 in IDH1 pH Regulated Catalysis and Inhibitor Binding

Residue D273 was predicted to have an upshifted pK_a from a value of ~3.7 in solution to ~6.5. The residue is located in the α10 helix, an important regulatory domain that transitions from an ordered loop in the apoenzyme form of IDH1 to an α-helix in the holoenzyme form of IDH1.[47] This residue is ~12 Å from the nearest substrate, isocitrate, at the active site and is located at the dimer interface (Figure 4D). Again, mutations were selected that would remove ionizability without introducing major structural and steric changes to the overall IDH1 structure. Thus D273N, D273L, and D273S were created.
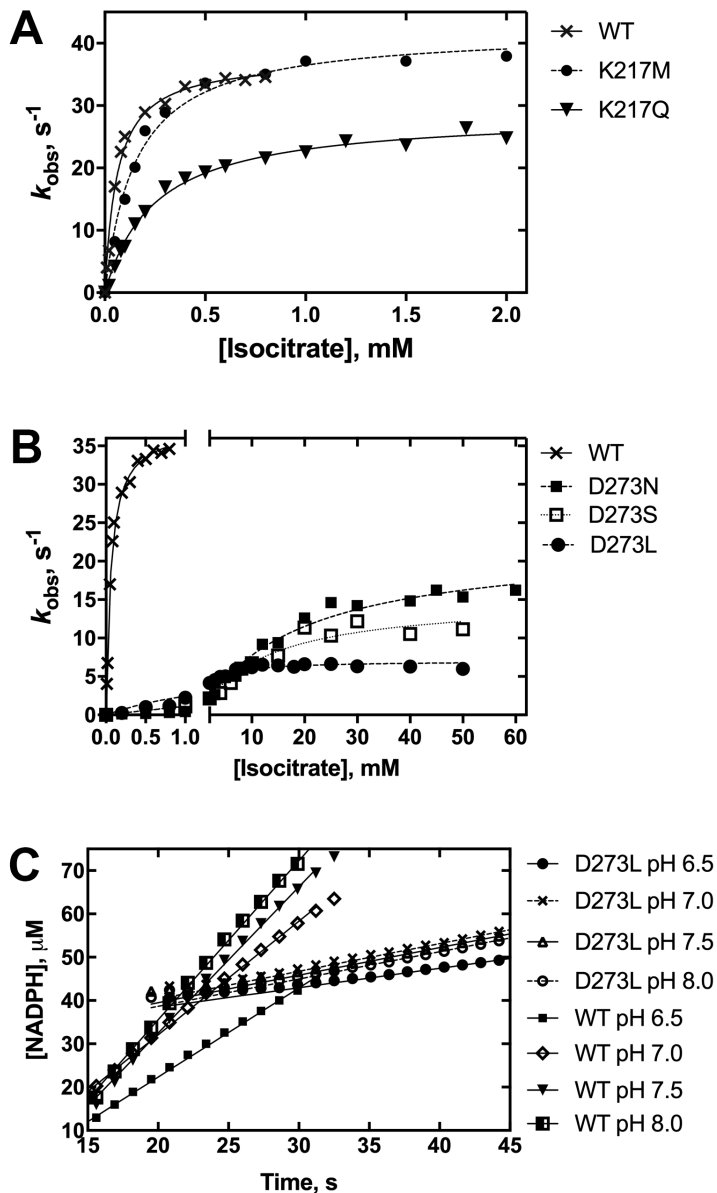
Figure 5: Kinetic characterization of K217 and D273 mutants in the WT IDH1 background. Steady-state rates of the normal forward reaction, the conversion of isocitrate to αKG, as catalyzed by (A) K217Q and K217M IDH1, and (B) D273N, D273S, and D273L IDH1 are shown relative to WT IDH1. k_{obs} values were obtained as described in Fig. 2. Replicates were performed for higher concentrations of isocitrate in the case of D273S IDH1 as these measurements proved more error prone. (C) Comparison of sensitivity to pH of the normal forward reaction in D273L and WT IDH1. Four representative pH values were tested, and loss of pH sensitivity was observed in D273L for most of the pH ranges tested. The average slope for the D273L mutants is 0.00042 ± 0.00006 µM s^{-1}, while the average slope is 0.04 ± 0.01 µM s^{-1}, indicating higher variability for the latter enzyme, and loss of pH dependence for the former. Each point in the curve represents a single replicate, and at least two different protein preparations were used to obtain replicates in the curves.

Table 3: Steady-state rates of k_{cat}, K_m, and k_{cat}/K_m were derived from fitting plots of k_{obs} versus substrate concentration with the Michaelis-Menten equation with the standard error determined from the deviance from these hyperbolic fits. Data were obtained as described in Figure 5. [a]From ref. [105]. [b]no detectable activity (NADPH consumption) was observed, so a limit of detection is listed. [c]Due to no detectable activity, K_m and k_{cat}/K_m could not be determined.

IDH1, reaction	k_{cat}, s^{-1}	K_m, mM	k_{cat}/K_m, mM^{-1}s^{-1}
WT, isocitrate → αKG	37.6 ± 0.7	0.061 ± 0.005	620 ± 50
K217M, isocitrate → αKG	42 ± 1	0.15 ± 0.02	280 ± 40
K217Q, isocitrate → αKG	28.7 ± 0.5	0.25 ± 0.02	115 ± 9
D273N, isocitrate → αKG	22 ± 1	19 ± 2	1.2 ± 0.1
D273S, isocitrate → αKG	15 ± 1	12 ± 3	1.3 ± 0.3
D273L, isocitrate → αKG	7.0 ± 0.2	1.9 ± 0.2	3.7 ± 0.4
R132H, αKG → D2HG	1.44 ± 0.05[a]	1.5 ± 0.2[a]	1.0 ± 0.1[a]
D273N/R132H, αKG → D2HG	≤ 0.02[b]	N.D.[c]	N.D.[c]
D273L/R132H, αKG → D2HG	≤ 0.02[b]	N.D.[c]	N.D.[c]
D273S/R132H, αKG → D2HG	≤ 0.02[b]	N.D.[c]	N.D.[c]

Energy minimized structures showed no major changes to overall IDH1 structure. To confirm no major changes occurred to the structure of IDH1, we used CD to investigate the secondary structure of D273L IDH1 and found no notable changes in T_m compared to WT IDH1 (Supplemental Figure S5). Recombinant D273N, D273L, and D273S were expressed and purified in *E. coli* and the kinetic profile of these mutants was determined under steady-state conditions (Figure 5B, Table 3). Insertion of a nonpolar residue led to a ~170-fold decrease in catalytic efficiency for D273L, driven a 5.4-fold change in k_{cat} and a 31-fold increase in K_m. Oddly, the more conservative mutations had an even greater decrease in catalytic efficiency. D273N IDH1 had a > 500-fold decrease in catalytic efficiency driven by a > 300-fold increase in K_m. D273S IDH1 had another ~500-fold drop in catalytic efficiency, driven by a ~200-fold increase in K_m and a ~2.5 fold decrease in k_{cat}. Thus, D273 mutants produced very large changed in WT IDH1 catalytic efficiency.

To determine whether the residue D273 conferred potential pH-sensitivity to WT IDH1, the rates of isocitrate to αKG conversion by D273L at saturating concentrations of substrates in KPhos buffer at pH 6.5, 7.0, 7.5, and 8.0 were measured and compared to values of WT IDH1 (Figure 5C) at the same pH values. Incubating WT IDH1 with 600 μM isocitrate and 200 μM NADP$^+$ yielded k_{obs} of 23.0, 27.5, 37.6, and 39.2 s^{-1} at a pH of 6.5, 7.0, 7.5, and 8.0, respectively, showing clear trends in pH-dependent activity. When examining D273L IDH1 catalysis, k_{obs} values were not altered by pH, except at the most acidic environment. The values for k_{obs} were determined to be 4.5, 6.4, 6.4, and 6.3 s^{-1} at pH 6.5, 7.0, 7.5, and 8.0, respectively (Figure 5C). These findings support our hypothesis that D273 is a critical mediator contributing to physiological pH sensitivity.

The α10 helix is suggested to be an important region within IDH1 that confers selectivity for mutant IDH1 inhibitor binding.[64] Since D273 is located in this domain, we asked if D273 mutants altered catalysis and inhibitor binding within the R132H IDH1 background. Double mutants (D273N/R132H, D273S/R132H, and D273L/R132H IDH1) were expressed and purified, and assessed via steady-state kinetic conditions to obtain the rate of conversion for αKG to D2HG. The double mutants were essentially catalytically inactive; thus, we only report a limit of detection, k_{obs} of $\leq$ 0.02 s^{-1} because measured rates fell below this value. Despite sampling a range of enzyme concentrations (up to 500 nM) across 4 different recombinant protein preparations with high substrate concentrations (Table 3) to improve detection of product formation, rates remained below the detectable limit. This low rate prevented us from determination steady-state kinetic constants and performing IC$_{50}$ measurements.

Instead, we decided to use isothermal titration calorimetry (ITC) to determine the affinity of D273L/R132H IDH1 for two commercially available and selective mutant IDH1 inhibitors AGI-5198[65] and ML309.[59] Both inhibitors showed that binding was affected. We obtained a K_d value of 3.3 $\pm$ 0.5 μM for AGI-5198 binding to D273L/R132H IDH1, compared to a previously reported IC$_{50}$ value of 0.07 μM for R132H IDH1.[65] Under our experimental conditions, ML309 did not bind to D273L/R132H, which implies a $K_d \geq$ 20 μM due to limits of detection, compared to a previously reported K_d value of 0.48 $\pm$ 0.05 μM for R132H IDH1.[66] These results indicate that D273 plays an important role in WT and mutant IDH1 catalysis, pH-sensing in catalysis of the conversion of isocitrate to αKG, and in mutant IDH1 inhibitor binding.

2.5 Discussion

In this work we show a comprehensive study of the effect of pH on WT IDH1 catalytic activity. During catalysis, IDH1 transitions from an inactive open conformation to an active closed conformation and the overall acid-base chemistry is catalyzed by an active site lysine-tyrosine pair.[2, 41, 47] Based on a number of structural, kinetic, and mutagenesis studies, the residues involved are Y139 and K212', where lysine functions as the general base to remove a proton from the hydroxy group of isocitrate and tyrosine functions as the general acid to protonate the enolate intermediate to form αKG.[41, 67] In the open state, the active site residues are solvent exposed and both the Y139 and K212' residues would be protonated.[2] Notably, K212' was found to be a buried ionizable residue by our pHinder data in the closed holoenzyme of IDH1. In addition, K212' has been suggested to be deprotonated in the closed active form of IDH1.[67] In a fully closed IDH1, K212' deprotonation would be facilitated by electrostatic repulsion from nearby arginine residues (R132, R100, R109), and nearby aspartate residues (D275, D279, or D252') that also aid in coordinating the divalent magnesium ion.[67] Recent studies have shown that D275 deprotonates K212 to facilitate catalysis, and thus a catalytic triad of an aspartate, lysine, and tyrosine is responsible for the conversion of isocitrate to αKG.[67] Our biochemical analysis found the rate of isocitrate to αKG conversion was pH sensitive in that k_{cat} values decreased with increasingly acidic environments. A more acidic environment may make it less favorable for the aspartate residue to deprotonate the lysine to act as a general base catalyst to form the oxalosuccinate intermediate, or more difficult to coordinate magnesium properly, and therefore decrease the rate of the forward reaction. In contrast, the trends for catalytic efficiency for the forward reaction were not pH-dependent because increasingly acidic pH led to varying K_m values. Isocitrate binds

preferentially to IDH1 when all three carboxylate moieties are deprotonated.[68-71] Thus, the variation in K_m values may be due to an altered ionic form of the substrates due to pH changes or changes in ionic strength of the bulk assay solution or at the active site of the enzyme.

The observed trends in the reverse reaction were found to be buffer dependent, had slower k_{cat} values, larger K_m values, and lower catalytic efficiencies compared to the forward reaction of IDH1. Although the IDH1 and IDH2 catalyzed reaction is reversible, the chemistry of the reverse reaction is complicated by a difficult carboxylation step. Physiological carboxylation reactions are typically accompanied by a cofactor, various ways of activating carbon dioxide, chaperones, environmental factors such as pH, or a combination of these. For example, in the Calvin Cycle, the energetically unfavorable carbon fixation reaction by RuBisCo is notably slow, and is therefore coupled with ATP hydrolysis through RuBisCo activase to facilitate carbamate formation between a lysine residue and carbon dioxide, in addition to a pH gradient and activating phosphate ions in solution .[14, 72] Moreover, other carboxylating enzymes such as acetyl-CoA carboxylase, methylcrotonyl-CoA, propionyl-CoA carboxylase, and pyruvate carboxylase are known to form carbamates between a covalently bound biotin cofactor only after bicarbonate is activated by ATP.[73-76] IDH1 and IDH2 are not regulated by ATP and do not have a covalent biotin cofactor.[2] Thus, another rationale is needed to explain the observation of the reverse reaction. Studies have shown that pH plays a role in the formation of a carbamate between amino acids and carbon dioxide.[77] Though neither lysine nor long polypeptide chains were studied, a bell-shaped pH dependent curve of carbamate formation was obtained due to maximized CO_2 solubility and nucleophilicity of the amino

group.[77] Our k_{cat} values for the reverse reaction follow a similar bell-shaped curve with pH, peaking at 6.8 in KPhos and 7.0 in Tris. Thus, pH 6.8–7.0 appears to satisfy the necessary conditions to drive the reverse reaction *in vitro*, while higher pH values would lower the concentration of CO_2 and lower pH values would decrease nucleophilicity of the potential amino acid residue forming the carbamate.[77] If a carbamate is not forming between CO_2 and IDH1, as in the enzymatic mechanism of phosphoenolpyruvate carboxylase, perhaps the CO_2 is activated by the phosphate ions from the KPhos buffer solution or reacting with the amines of Tris/bis-Tris buffer.[78, 79] The observance of the reverse reaction *in vivo* can be attributed to various mechanisms. For example, experimental evidence shows the reverse reaction of IDH2 is driven by a proton electrochemical gradient supplied by the H^+-transhydrogenase.[14] Furthermore, hypoxic conditions have been shown to drive the conversion of αKG to isocitrate; however, the exact conditions, and precise mechanism responsible are still under debate.[19-21] These conditions are associated with an acidic pH, but pH changes in the cell may be too small making our biochemical analysis not physiologically relevant.

In addition, our biochemical analysis is limited by the buffer systems we selected. In this investigation, we found some trends to be buffer dependent. Buffer dependent activity is observed for other proteins as well. For example, fumarase activity, has been documented to be dependent on phosphate ion concentration.[80, 81] In our studies, we chose to use Tris/bis-Tris and KPhos to study pH dependency in the IDH1 reaction. The components of the buffer may have stabilizing or destabilizing effects on protein structure. Destabilizing effects are frequently ascribed to Tris buffer.[82] Also, the buffer capacity range of Tris comprises $\pm$ 1.0 pH units and has a temperature-dependent pK_a. Since a

broad range of pH values were studied, Tris could not cover the entire range alone and we selected bis-Tris in order to study IDH1 biochemistry at more acidic pH values. This is an unsatisfactory procedure due to varying enzyme activities in different buffers.[82] Tris and bis-Tris buffers may also react with the CO_2 needed for the reverse reaction.[78] In such cases, more universal buffers that cover a broader pH range are applicable.[82] The enzyme reactions were repeated in KPhos buffer to explore any buffer dependent catalytic activity. KPhos buffer has the advantage of covering the entire pH range and has a pK_a that is less temperature dependent than Tris/bis-Tris, though the phosphate ion can sequester magnesium ions necessary for catalysis. Due to the interfering buffer effects, the IDH1 reaction should be investigated in zwitterionic buffers such as HEPES, MOPS, or TES that are additionally nonreactive with substrates in future experiments.[82]

One important point that is often overlooked in pH studies in the control of the ionic strength of reaction mixtures at the various pH values. Most enzymes accept ionic strength between 0.05 and 0.2 M, where the calculation takes into account each component of the reaction mixture and lower ionic strengths tend to destabilize protein structure.[83] The ionic strength of a solution provides a measure of the ionic environment that must influence the catalytic activity of enzymes because the ionizable amino acids undergo different protonation states when converting substrates to products. Additional effects of ionic strength are also observed with enzymes that utilize charged substrates like IDH1. Thus, when ionic strength is not controlled in pH-variation studies, a change in enzyme activity may be attributable to a change in ionic strength and not a change in an ionization process. It should be noted that a measurement of ionic strength reflects the bulk environment, and not the true environment within the active site of the protein.[83]

Future experiments involving IDH1 pH dependence should account for ionic strength if different buffer systems are used.

Our investigation is limited to steady-state kinetics, allowing us to measure the slow step in IDH1 catalysis. The rate-determining step is widely predicted to be hydride transfer.[84-87] Thus, by using only steady-state kinetics, we are missing the effect of pH on the faster steps of the catalytic mechanism of IDH1. To study the effect of pH on the faster steps, we would have to use pre-steady state kinetics. This would be an interesting avenue of exploration since previous studies have demonstrated a pH-sensitivity in the NADPH-dependent hydride transfer of dihydrofolate reductase.[88-92] These studies showed that changes in a hydrogen-bonding network serve as a mechanism to switch the rate determining step from product release at acidic pH levels to hydride transfer at higher pH levels.[88-92] Examining the effect of pH on each step of the reaction pathway could resolve the contributions of ionizable groups toward ground-state binding and transition state stabilization.

Using the pHinder algorithm we were able to identify residues K217 and D273 that have potential pH-sensing ability. Experimentally, IDH1 catalysis was largely unaffected by point mutations at K217, and this residue probably plays a role in maintaining IDH1 structural integrity like other lysine residues.[36, 38] In our inhibitor binding studies, we observed a decrease in inhibitor affinity in the D273 point mutations in the R132H IDH1 background. Energy minimized structures showed that the mutational variants at residue D273 are not predicted to affect global structure of IDH1 or local folding features (Supplemental Figure S6). Due to the location of D273 this finding implies a possible

resistance mechanism, though the double mutant did not perform the neomorphic reaction more experiments are necessary to clarify this potential role.[93, 94]

Interestingly, our biochemical analysis with D273 mutants showed that IDH1 catalysis and pH-dependent activity were affected. D273 is part of the α10 helix, a domain in that regulates IDH1 catalytic activity.[47] Throughout the catalytic cycle, IDH1 undergoes a conformational change from an open (inactive) conformation, a semi-open conformation, and finally a closed (active) conformation. Residue D279 of the α10 helix forms hydrogen bonds with S94 to stabilize the inactive conformation of IDH1.[47] At sufficient concentrations, isocitrate then out competes D279 for the active site of IDH1 and the newly formed favorable interactions between isocitrate and the active site residues stabilize the active conformation.[47] The α10 helix correspondingly goes through a conformational change as well, from an ordered loop in the inactive form of IDH1 to an α-helix in the active form of IDH1. Investigating if the transition of the α10 helix is pH-dependent provides an interesting path forward. Studying the conversion of isocitrate to αKG at lower pH values to provide further evidence of D273 as a pH sensor would lose physiological relevance and may affect protein stability. Instead, generating a D273E mutant would provide further evidence that D273 is a bona fide pH sensor since glutamate and aspartate differ only in a methylene group and pK_a values of the ionizable side chain should be remain more or less the same. I would be cautious with further mutagenesis experiments that removed or reversed the ionizability of D273 because the pK_a of the point mutation can be influenced by the surrounding environment. A major challenge when studying the effect of one titratable group in proteins is that a protein may contain many titratable groups in close proximity. If the titratable groups are located far apart

(~25 Å) then they may titrate independently and not effect each other or global protein structure; however, if two or more similarly titratable groups are spatially close more complicated behavior can result.[56] Put simply, the protonation state of each titratable group can affect the pK_a of all other proximal titratable groups, and these interactions can lead to non-sigmoidal titration curves.[56] That is, any amino acid substitution that eliminates or introduces a charged group can perturb the pK_a of nearby titratable residues, and this effect should be considered when interpreting the results of mutagenesis in studies investigating the pH-dependent properties of proteins.[56] Two-dimensional nuclear magnetic resonance spectroscopy (2D NMR) provides a working solution to this problem. 2D NMR is the one of the best tools to study pH-dependent conformational changes because the protein is in its native state, protein dynamics are visible in solution with NMR, and the pH of the solution can be easily modified.[33, 56] However, the major drawback to using NMR to study a pH-dependent conformational change in IDH1 is its size and dimer state in solution.[56] Since IDH1 is too big for NMR, Hydrogen-deuterium exchange (HDX) experiments would serve as a viable alternative to study pH-dependent conformational changes in the α10 helix; however, the effect of individual amino acids on secondary structure would not be observed. Protein crystallography and neutron diffraction at different pH values would help assign possible pH-dependent structure-function relationships and pinpoint hydrogen bond interactions, although electron density may be difficult to obtain in highly mobile domains of IDH1.[22, 56] In addition, molecular dynamics can be performed to study potential pH-dependent conformational changes in the α10 helix and the effect of point mutations on protein structure, as in a previous study on pH-sensing histidine mutations in EGFR.[27] Therefore, there are more biophysical

techniques available that complement our biochemical analysis to provide further evidence that D273 is a pH-sensing residue.

We extended our biochemical studies by investigating how changes in intracellular pH affected cellular αKG levels and isocitrate levels. Based on our *in vitro* biochemical analysis, we expected a decrease in αKG concentration in HT1080 cells based on biochemical analysis. However, our findings in the cell-based assays may not be solely attributable to IDH1 activity. Cells possess many redundant pathways; thus, other enzymes can consume and produce the same compounds including IDH2 and IDH3 isoforms of IDH1. Concentrations of αKG have been shown to decrease when intracellular pH is acidified and increased when intracellular pH is increased to more alkaline levels.[95] Two other enzymes that consume αKG are glutamate dehydrogenase (GDH) and α-ketoglutarate dehydrogenase (αKGDH). GDH is known to catalyze the reversible NAD(P)$^+$ dependent conversion of glutamate to αKG. GDH catalysis is affected by pH, substrate concentrations, ionic strength, and composition of the buffer.[96] Though human GDH1 is known to operate optimally at basic pH values (7.75 – 8.0), humans have evolved a distinct GDH2 isoenzyme that is regulated differently than GDH1 and can operate efficiently at more acidic pH values (7.25 – 7.0) and is expressed more, relative to GDH1 in rapidly dividing cells.[96] Early studies showed that reductive amination reaction would be favored *in vivo*; however, recent studies have shown that the equilibrium position of GDH is dependent on substrate concentration levels. More recent kinetic studies on both GDH1 and GDH2 show that the K_m for ammonia, a key substrate for the reductive amination reaction, increases as intracellular pH becomes more acidic.[98] The conditions favoring the reductive amination reaction can be met in the mitochondria;

81

however, functional studies have shown that the deaminating reaction is favored.[96, 98] Another metabolic enzyme, αKGDH, has been shown to possess pH-dependent enzyme activity. Whereas the K_m values for substrates for GDH1 and GDH2 increased as intracellular pH acidifies, the K_m values for αKGDH decrease as intracellular pH becomes more acidic. The acidic intracellular pH values of the HT1080 cells may have a greater effect on α-ketoglutarate dehydrogenase activity compared to IDH1.[99] Metabolic flux analysis would be ideal to elucidate the effect of pH on isocitrate and αKG concentrations in the cell due to IDH1 activity.[19-21, 100]

This work centrally focused on the reactions catalyzed by WT IDH1, though we do note in our cellular assays that the concentrations of 2HG dropped upon treatment with ESOM. This result was surprising as previous studies showed that 2HG levels increased upon intracellular acidification, although a different cell line was used and a different process of proton pump inhibition.[40] Thus we also consider the possibility that treatment with ESOM can affect cells beyond just lowering intracellular pH. While acting as a V-ATPase inhibitor, esomeprazole has also been shown to have anti-tumor properties and induce apoptosis.[101, 102] We also note that esomeprazole treatment increased the level of several amino acids (Supplemental Table S1), with the largest changes in occurring in branched-chain amino acids in the WT IDH1 expressing cells lines. Because αKG is important for the synthesis of branched-chain amino acids, the observed depletion of αKG upon treatment with esomeprazole may be simply due to increased branched-chain amino acid synthesis.[103]

Overall, our findings characterize the pH sensitivity of IDH1. We show that the k_{cat} of the forward reaction is the most sensitive to changes in pH. In addition, residue D273

plays a role in sensing changes in pH during IDH1 catalysis. Notably, removing the ionizability of this residue resulted in a loss of pH sensitivity, as this work combines structural informatics with biochemical experiments to identify a potential mechanism of regulation of IDH1 catalysis.

2.6 References

1.   Mailloux, R. J., Lemire, J., Singh, R., Che, D. R., Hamel, R. D., Appanna, V. D., Be, R. (2007). The tricarboxylic acid cycle , an ancient metabolic network with a novel twist. *PLoS One, 2*, e690-e690.

2.   Reitman, Z. J., Yan, H. (2010). Isocitrate Dehydrogenase 1 and 2 mutations in cancer : alterations at a crossroads of cellular metabolism. *J. Natl. Cancer Inst., 102*(13), 932-941.

3.   Bogdanovic, E. (2015). IDH1, lipid metabolism and cancer: shedding new light on old ideas. *Biochim. Biophys. Acta, 1850*(9), 1781-1785.

4.   Cerami, E., Gao, J., Dogrusoz, U., Gross, B. E., Sumer, S. O., Arman, B., Jacobsen, A., Byrne, C. J., Heuer, M. L., Larsson, E., Antipin, Y., Reva, B., Goldberg, A. P., Sander, C., Schultz, N. (2012). The cBio cancer genomics portal: an open platform for exploring multidimensional cancer genomics data. *Cancer Discov, 2*(5), 401-404.

5.   Parsons, D. W., Jones, S., Zhang, X., Lin, J. C.-h., Leary, R. J., Angenendt, P., Mankoo, P., Carter, H., Siu, I. m., Gallia, G. L., Olivi, A., McLendon, R., Rasheed, B. A., Keir, S., Nikolskaya, T., Nikolsky, Y., Busam, D. A., Tekleab, H., Jr, L. A. D., Hartigan, J., Smith, D. R., Strausberg, R. L., Kazue, S., Marie, N., Mieko, S., Shinjo, O., Yan, H., Riggins, G. J., Bigner, D. D., Karchin, R., Papadopoulos, N., Parmigiani, G., Vogelstein, B., Velculescu, V. E., Kinzler, K. W. (2008). An integrated genomic analysis of human glioblastoma mutliforme. *Science, 321*(5897), 1807-1812.

6.   Yan, H., Williams, P., Jin, G., McLendon, R., Rasheed, B. A., Yuan, W., Kos, I., Ines, B-H., Jones, S., Riggins, G., Friedman., H., Friedman, A., Reardon, D., Herndon., J., Kinzler, K. W., Velculescu, V. E., Vogelstein, B., Bigner, D. D. (2009). IDH1 and IDH2 mutations in gliomas. *N Engl J Med, 360*(8), 765-773.

7.   Larson, D. E., Ph, D., McLellan, M. D., Chen, K., Ph, D., Koboldt, D. C., Fulton, R. S., Delehaunty, K. D., McGrath, S. D., Fulton, L. A., Locke, D. P., Ph, D., Magrini, V. J., Ph, D., Abbott, R. M., Vickery, T. L., Reed, J. S., Robinson, J. S., Wylie, T., Smith, S. M., Carmichael, L., Eldred, J. M., Harris, C. C., Walker, J., Peck, J. B., Du, F., Dukes, A. F., Sanderson, G. E., Brummett, A. M., Clark, E., McMichael, J. F., Meyer, R. J., Schindler, J. K., Pohl, C. S., Wallis, J. W., Ph, D., Shi, X., Lin, L., Schmidt, H., Tang, Y., Haipek, C., Wiechert, M. E., Ivy, J. V., Kalicki, J., Elliott, G., Ries, R. E., Payton, J. E., Ph, D., Westervelt, P., Ph, D., Heath, S., Shannon, W. D., Ph, D., Nagarajan, R., Ph, D., Link, D. C., Walter, M. J., Graubert, T. A., Dipersio, J. F., Ph, D., Wilson, R. K., Ph, D., Ley, T. J. (2009). Recurring mutations found by sequencing an acute myeloid leukemia genome. *N Engl J Med, 361*, 1058-1066.

8.   Kosmider, O., Slama, L., Dreyfus, F., Quesnel, B., Al, E. (2009). Mutations of IDH1 and IDH2 genes in early and accelerated phases of myelodysplastic syndromes and MDS / myeloproliferative neoplasms. *Leukemia, 24*, 1094-1096.

9.   Borger, D. R., Tanabe, K. K., Fan, K. C., Lopez, H. U., Fantin, V. R., Straley, K. S., Schenkein, D. P., Hezel, A. F., Ancukiewicz, M., Liebman, H. M., Kwak, E. L., Clark, J. W., Ryan, D. P., Deshpande, V., Dias-Santagata, D., Ellisen, L. W., Zhu, A. X., Iafrate, J. A. (2012). Frequent mutation of isocitrate dehydrogenase (IDH)1 and IDH2 in cholangiocarcinoma identified through broad-based tumor genotyping. *Oncologist, 17*(1), 72-79.

10.   Pietrak, B., Zhao, H., Qi, H., Quinn, C., Gao, E., Boyer, J. G., Concha, N., Brown, K., Duraiswami, C., Wooster, R., Sweitzer, S., Schwartz, B. (2011). A tale of two subunits: how the neomorphic R132H IDH1 mutation enhances production of alphaHG. *Biochemistry, 50*, 4804-4812.

11.   Avellaneda Matteo, D., Grunseth, A. J., Gonzalez, E. R., Anselmo, S. L., Kennedy, M. A., Moman, P., Scott, D. A., Hoang, A., Sohl, C. D. (2017). Molecular mechanisms of isocitrate dehydrogenase 1 (IDH1) mutations identified in tumors: The role of size and hydrophobicity at residue 132 on catalytic efficiency. *J Biol Chem, 292*(19), 7971-7983.

12.   Nunez, F. J., Mendez, F. M., Kadiyala, P., Alghamri, M. S., Savelieff, M. G., Garcia-Fabiani, M. B., Haase, S., Koschmann, C., Calinescu, A., Kamran, N., Saxena, M., Patel, R., Carney, S., Guo, Z. M., Edwards, M., Ljungman, M., Qin, T., Sartor A. M., Tagett, R., Venneti, S., Brosnan-Cashman, J., Meeker, A., Gorbunova, V., Herting, J. C., Ross, J. L., Hambardzumyan, D., Hervey-Jumper, S., Figueroa, M. E., Lowenstein, P. R., Castro, M. G. (2019). IDH1-R132H acts as a tumor

suppressor in glioma via epigenetic up-regulation of the DNA damage response. *Sci. Transl. Med., 11*, 1-13

13.   Dang, L., White, D. W., Gross, S., Bennett, B. D., Bittinger, M. A., Driggers, E. M., Fantin, V. R., Jang, H. G., Jin, S., Keenan, M. C., Marks, K. M., Prins, R. M., Ward, P. S., Yen, K. E., Liau, L. M., Rabinowitz, J. D., Cantley, L. C., Thompson, C. B., Heiden, M. G. V., Su, S. M. (2009). Cancer-associated IDH1 mutations produce 2-hydroxyglutarate. *Nature, 462*, 739-744.

14.   Dang, L., Jin, S., Su, S. M. (2010). IDH mutations in glioma and acute myeloid leukemia. *Trends Mol Med., 16*(9), 387-397.

15.   Horbinski, C. (2013). What do we know about IDH1 / 2 mutations so far, and how do we use it? *Acta Neuropathol, 125*, 621-636.

16.   Calvert, A. E., Chalastanis, A., Wu, Y., Hurley, L. A., Kouri, F. M., Bi, Y., Kachman, M., May, J. L., Bartom, E., Hua, Y., Mishra, R., Schlitz, G. E., Dubrovskyi, O., Mazar, A. P., Peter, M. E., Zheng, H., James., C. D., Burant, C., F., Chandel, N. S., Davuluri, R. V., Horbinski, C., Stegh, A. H (2017). Cancer-associated IDH1 promotes growth and resistance to targeted therapies in the absence of mutation. *Cell Rep., 19*(9), 1858-1873.

17.   Popovici-muller, J., Saunders, O., Salituro, F. G., Travins, J. M., Yan, S., Zhao, F., Gross, S., Dang, L., Yen, K. E., Yang, H., Straley, K. S., Jin, S., Kunii, K., Fantin, V. R., Zhang, S., Pan, Q., Shi, D., Biller, S. A., Su, S. M. (2012). Discovery of the First Potent Inhibitors of Mutant IDH1 That Lower Tumor 2 - HG in Vivo. *ACS Med. Chem. Lett.*(3), 850-855.

18.   Zhang, V., Yen, K. E., Kapsalis, S. M., Yang, H., Dai, D., Fan, B., Goldwasser, M., Liu, H., Agresta, S., Wu, B., Attar, E. C., Tallman, M. S., Stone, R. M., Kantarjian, H. M. (2018). Durable Remissions with Ivosidenib in IDH1-mutated relapsed or refractory AML. *N. Engl J Med., 378*, 2386-2398.

19.   Metallo, C. M., Gameiro, P. A., Bell, E. L., Mattaini, K. R., Yang, J., Hiller, K., Jewell, C. M., Johnson, Z. R., Irvine, D. J., Guarente, L., Kelleher, J. K., Vander, M. G. (2012). Reductive glutamine metabolism by IDH1 mediates lipogenesis under hypoxia. *Nature, 481*, 380-384.

20. Wise, D. R., Ward, P. S., Shay, J. E. S., Cross, J. R., Gruber, J. J., Sachdeva, U. M. (2011). Hypoxia promotes isocitrate dehydrogenase- dependent carboxylation of α-ketoglutarate to citrate to support cell growth and viability. *Proc. Natl. Acad. Sci. U.S.A, 108*(49), 19611-19616.

21. Filipp, F. V., Scott, D. A., Ronai, Z. A., Osterman, A. L., Smith, J. W. (2011). Reverse TCA cycle flux through isocitrate dehydrogenases 1 and 2 is required for lipogenesis in hypoxic melanoma cells. *Pigment Cell and Melanoma Res.*(25), 375-381.

22. Schonichen, A., Webb, B. A., Jacobson, M. P., Barber, D. L. (2013). Considering protonation as a posttranslational modification regulating protein structure and function. *Annu. Rev. Biophys, 42*, 289-314.

23. Matsuyama, S., Llopis, J., Deveraux, Q. L., Tsien, R. Y., Reed, J. C. (2000). Changes in intramitochondrial and cytosolic pH : early events that modulate caspase activation during apoptosis. *Nat. Cell. Bio., 2*, 318-325.

24. Mulkey, D. K., Iii, R. A. H., Ritucci, N. A., Putnam, R. W., Dean, J. B. (2004). Oxidative stress decreases pH i and Na+ / H+ exchange and increases excitability of solitary complex neurons from rat brain slices. *Am. J. Physiol. Cell Physiol., 286*(4), 940-951.

25. Nakamura, U., Iawse, M., Uchizono, Y., Sonoki K., Sakai, N., Imoto, H., Goto, D., Iida, M. (2006). Rapid Intracellular acidification and cell death by H2O2 and alloxan in pancreatic β cells. *40*, 2047-2055.

26. Ulmschneider, B., Grillo-Hill, B. K., Benitez, M., Azimova, D. R., Barber, D. L., Nystul, T. G. (2016). Increased intracellular pH is necessary for adult epithelial and embryonic stem cell differentiation. *J Cell Biol, 215*(3), 345-355.

27. White, K. A., Ruiz, D. G., Szpiech, Z. A., Strauli, N. B., Hernandez, R. D., Jacobson, M. P., Barber, D. L. (2017). Cancer-associated arginine-to-histidine mutations confer a gain in pH sensing to mutant proteins. *Sci Signal, 10*(495).

28. Webb, B. A., Chimenti, M., Jacobson, M. P., Barber, D. L. (2011). Dysregulated pH: a perfect storm for cancer progression. *Nat Rev Cancer, 11*(9), 671-677.

29.   Isom, D. G., Sridharan, V., Dohlman, H. G. (2016). Regulation of Ras paralog thermostability by networks of buried ionizable groups. *Biochemistry, 55*(3), 534-542.

30.   Webb, B. A., White, K. A., Grillo-Hill, B. K., Schonichen, A., Choi, C., Barber, D. L. (2016). A histidine cluster in the cytoplasmic domain of the Na-H exchanger NHE1 confers pH-sensitive phospholipid binding and regulates transporter activity. *J Biol Chem, 291*(46), 24096-24104.

31.   Dobson, G. P., Yamamoto, E., Hochachka, P. W. (1986). Phosphofructokinase control in muscle: nature and reversal of pH-dependent ATP inhibition. *Am J Physiol, 250*(1 Pt 2), R71-76.

32.   Trivedi, B., Danforth, W. H. (1966). Effect of pH on the kinetics of frog muscle phosphofructokinase. *J Biol Chem, 241*(17), 4110-4112.

33.   Isom, D. G., Sridharan, V., Baker, R., Clement, S. T., Smalley, D. M., Dohlman, H. G. (2013). Protons as Second Messenger Regulators of G Protein Signaling. *Mol. Cell, 51*(4), 531-538.

34.   Gaspar, P., Neves, A. R., Shearman, C. A., Gasson, M. J., Baptista, A. M., Turner, D. L., Soares, C. M., Santos, H. (2007). The lactate dehydrogenases encoded by the ldh and ldhB genes in *Lactococcus lactis* exhibit distinct regulation and catalytic properties - comparative modeling to probe the molecular basis. *Febs j, 274*(22), 5924-5936.

35.   Suzuki, H., Ogura, Y. (1970). Effect of pH on the kinetic parameters of yeast L(+)-lactate dehydrogenase (cytochrome b2). *J Biochem, 67*(2), 291-295.

36.   Karp, D. A., Stahley, M. R., Garcia-Moreno, B. (2010). Conformational consequences of ionization of Lys, Asp, and Glu buried at position 66 in staphylococcal nuclease. *Biochemistry, 49*(19), 4138-4146.

37.   Bell-upp, P., Robinson, A. C., Whitten, S. T., Wheeler, E. L., Lin, J., Stites, Wesley. E., E Garcia-Moreno, Bertrand. (2011). Biophysical Chemistry Thermodynamic principles for the engineering of pH-driven conformational switches and acid insensitive proteins. *Biophysical Chemistry, 159*(1), 217-226.

38. Isom, D. G., Castaneda, C. A., Cannon, B. R., Garcia-Moreno, B. (2011). Large shifts in pKa values of lysine residues buried inside a protein. *Proc Natl Acad Sci U S A,108*(13), 5260-5265.

39. Leonardi, R., Subramanian, C., Jackowski, S., Rock, C. O. (2012). Cancer-associated isocitrate dehydrogenase mutations inactivate NADPH-dependent reductive carboxylation. *J Biol Chem, 287*(18), 14615-14620.

40. Nadtochiy, S. M., Schafer, X., Fu, D., Nehrke, K., Munger, J., Brookes, P. S. (2016). Acidic pH is a metabolic switch for 2-hydroxyglutarate generation and signaling. *J Biol Chem, 291*(38), 20188-20197.

41. Aktas, D. F., Cook, P., F. (2009). A Lysine-Tyrosine Pair Carries Out Acid-Base Chemistry in the Metal Ion-Dependent Pyridine Dinucleotide-Linked -Hydroxyacid Oxidative Decarboxylase. *Biochemistry, 48*(16), 3565-3577.

42. Orwig, S. D., Lieberman, R. L. (2011). Biophysical characterization of the olfactomedin domain of myocilin, an extracellular matrix protein implicated in inherited forms of glaucoma. *PLoS One, 6*(1), e16347.

43. Ratnikov, B., Aza-Blanc, P., Ronai, Z. A., Smith, J. W., Osterman, A. L., Scott, D. A. (2015). Glutamate and asparagine cataplerosis underlie glutamine addiction in melanoma. *Oncotarget, 6*(10), 7379-7389.

44. Isom, D. G., Dohlman, H. G. (2015). Buried ionizable networks are an ancient hallmark of G protein-coupled receptor activation. *Proc Natl Acad Sci U S A, 112*(18), 5702-5707.

45. Sastry, G. M., Adzhigirey, M., Day, T., Annabhimoju, R., Sherman, W. (2013). Protein and ligand preparation: parameters, protocols, and influence on virtual screening enrichments. *J Comput Aided Mol Des, 27*(3), 221-234.

46. Schrödinger Release 2018-3, Schrödinger Suite 2018-3 Protein Preparation Wizard.

47.   Xu, X., Zhao, J., Xu, Z., Peng, B., Huang, Q., Arnold, E., Ding, J. (2004). Structures of human cytosolic NADP-dependent isocitrate dehydrogenase reveal a novel self-regulatory mechanism of activity. *J Biol Chem, 279*(32), 33946-33957.

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

49.   Adams, P. D., Afonine, P. V., Bunkoczi, G., Chen, V. B., Davis, I. W., Echols, N., Headd, J. J., Hung, L. W., Kapral, G. J., Grosse-Kunstleve, R. W., McCoy, A. J., Moriarty, N. W., Oeffner, R., Read, R. J., Richardson, D. C., Richardson, J. S., Terwilliger, T. C., Zwart, P. H. (2010). PHENIX: a comprehensive Python-based system for macromolecular structure solution. *Acta Crystallogr D Biol Crystallogr, 66*(Pt 2), 213-221.

50.   Swenson, E. R. (2016). Hypoxia and Its Acid-Base Consequences: From Mountains to Malignancy. *Adv Exp Med Biol, 903*, 301-323.

51.   Harguindey, S., Reshkin, S. J., Orive, G., Arranz, J. L., Anitua, E. (2007). Growth and Trophic Factors, pH and the Na+:H+ Exchanger in Alzheimer´s Disease, Other Neurodegenerative Diseases and Cancer- New Therapeutic Possibilities and Potential Dangers. *Current Alzheimer Research, 4*(1), 53-65.

52.   Tracz, S. M., Abedini, A., Driscoll, M., Raleigh, D. P. (2004). Role of Aromatic Interactions in Amyloid Formation by Peptides Derived from Human Amylin. *Biochemistry, 43*, 15901-15908.

53.   Swietach, P., Vaughan-Jones, R. D., Harris, A. L., Hulikova, A. (2014). The chemistry, physiology and pathology of pH in cancer. *Philos Trans R Soc Lond B Biol Sci, 369*(1638), 20130099.

54.   Baldini, N., Avnet, S. (2018). The Effects of Systemic and Local Acidosis on Insulin Resistance and Signaling. *Int J Mol Sci, 20*(1).

55.   Grillo-Hill, B. K., Webb, B. A., Barber, D. L. (2014). Ratiometric imaging of pH probes. *Methods Cell Biol, 123*, 429-448.

56. Srivastava, J., Barber, D. L., Jacobson, M. P. (2007). Intracellular pH sensors: design principles and functional significance. *Physiology (Bethesda), 22*, 30-39.

57. Ma, S., Jiang, B., Deng, W., Gu, Z. K., Wu, F. Z., Li, T., Xia, Y., Yang, H., Ye, D., Xiong, Y., Guan, K. L. (2015). D-2-hydroxyglutarate is essential for maintaining oncogenic property of mutant IDH-containing cancer cells but dispensable for cell growth. *Oncotarget, 6*(11), 8606-8620.

58. Grassian, A. R., Parker, S. J., Davidson, S. M., Divakaruni, A. S., Green, C. R., Zhang, X., Slocum, K. L., Pu, M., Lin, F., Vickers, C., Joud-caldwell, C., Chung, F., Yin, H., Handly, E. D., Straub, C., Growney, J. D., Heiden, M. G. V., Murphy, A. N., Pagliarini, R., Metallo, C. M. (2014). IDH1 mutations alter citric acid cycle metabolism and increase dependence on oxidative mitochondrial metabolism. *Cancer Res., 74*, 3317-3332.

59. Davis, M. I., Gross, S., Shen, M., Straley, K. S., Pragani, R., Lea, W. A., Popovici-Muller, J., Delabarre, B., Artin, E., Thorne, N., Auld, D. S., Li, Z., Dang, L., Boxer, M. B., Simeonov, A. (2014). Biochemical, cellular and biophysical characterization of a potent inhibitor of mutant isocitrate dehydrogenase IDH1. *J Biol Chem, 289*, 13717-13725.

60. Jin, G., Pirozzi, C. J., Chen, L. H., Lopez, G. Y., Duncan, C. G., Feng, J., Spasojevic, I., Bigner, D. D., He, Y., Yan, H. (2012). Mutant IDH1 is required for IDH1 mutated tumor cell growth. *Oncotarget, 3*(8), 774-782.

61. Isom, D. G., Page, S. C., Collins, L. B., Kapolka, N. J., Taghon, G. J., Dohlman, H. G. (2018). Coordinated regulation of intracellular pH by two glucose-sensing pathways in yeast. *J. Biol. Chem., 293*(7), 2318-2329.

62. Harms, M. J., Castaneda, C. A., Schlessman, J. L., Sue, G. R., Isom, D. G., Cannon, B. R., Garcia-Moreno, E. B. (2009). The pK(a) values of acidic and basic residues buried at the same internal location in a protein are governed by different factors. *J Mol Biol, 389*(1), 34-47.

63. Rendina, A. R., Pietrak, B., Smallwood, A., Zhao, H., Qi, H., Quinn, C., Adams, N. D., Concha, N., Duraiswami, C., Thrall, S. H., Sweitzer, S., Schwartz, B. (2013). Mutant IDH1 enhances the production of 2-hydroxyglutarate due to its kinetic mechanism. *Biochemistry, 52*(26), 4563-4577.

64. Xie, X., Baird, D., Bowen, K., Capka, V., Chen, J., Chenail, G., Cho, Y., Dooley, J., Farsidjani, A., Fortin, P., Kohls, D., Kulathila, R., Lin, F., McKay, D., Rodrigues, L., Sage, D., Toure, B. B., van der Plas, S., Wright, K., Xu, M., Yin, H., Levell, J., Pagliarini, R. A. (2017). Allosteric mutant IDH1 inhibitors reveal mechanisms for IDH1 mutant and isoform selectivity. *Structure, 25*(3), 506-513.

65. Rohle, D., Popovici-Muller, J., Palaskas, N., Turcan, S., Grommes, C., Campos, C., Tsoi, J., Clark, O., Oldrini, B., Komisopoulou, E., Kunii, K., Pedraza, A., Schalm, S., Silverman, L., Miller, A., Wang, F., Yang, H., Chen, Y., Kernytsky, A., Rosenblum, M. K., Liu, W., Biller, S. A., Su, S. M., Brennan, C. W., Chan, T. A., Graeber, T. G., Yen, K. E., Mellinghoff, I. K. (2013). An inhibitor of mutant IDH1 delays growth and promotes differentiation of glioma cells. *Science, 340*(6132), 626-630.

66. Chambers, J. M., Miller, W., Quichocho, G., Upadhye, V., Matteo, D. A., Bobkov, A. A., Sohl, C. D., Schiffer, J. M. (2020). Water networks and correlated motions in mutant isocitrate dehydrogenase 1 (IDH1) are critical for allosteric inhibitor binding and activity. *Biochemistry, 59*(4), 479-490.

67. Neves, R. P. P., Fernandes, P. A., Ramos, M. J. (2015). Unveiling the Catalytic Mechanism of NADP+-Dependent Isocitrate Dehydrogenase with QM/MM Calculations. *ACS Catalysis, 6*(1), 357-368.

68. Ehrlich, R. S., Colman, R. F. (1987). Ionization of isocitrate bound to pig heart NADP+-dependent isocitrate dehydrogenase: 13C NMR study of substrate binding. *Biochemistry, 26*(12), 3461-3466

69. Ehrlich, Robert. S., Colman, Roberta. F. (1976). Influence of Substrates and Coenzymes on the Role of Manganous Ion in Reactions Catalyzed by Pig Heart Triphosphopyridine Nucleotide-Dependent Isocitrate Dehydrogenase. *Biochemistry, 15*(18), 4034-4041.

70. Colman, Roberta. F. (1972). Role of Metal Ions in Reactions Catalyzed by Pig Heart Triphosphopyridine Nucleotide-dependent Isocitrate Dehydrogenase II. Effect on Catalytic Properties and Residues. *Journal of Biological Chemistry, 247*(1), 215-223.

71. Bacon, C. R., Bednar, R. A., Colman, R. F. (1981). The reaction of 4-iodoacetamidosalicylic acid with TPN-dependent isocitrate dehydrogenase from pig heart. *Journal of Biological Chemistry, 256*(13), 6593-6599.

72. Portis, A. R. (2003) Rubisco activase – Rubisco's catalytic chaperone. *Photosynthesis Research 75*, 11-27.

73. Lee, C. K., Cheong, H. K., Ryu, K. S., Lee, J. I., Lee, W., Jeon, Y. H., Cheong, C. (2008). Biotinoyl domain of human acetyl-CoA carboxylase: Structural insights into the carboxyl transfer mechanism. *Proteins, 72*(2), 613-624.

74. Chu, C. H., Cheng, D. (2007). Expression, purification, characterization of human 3-methylcrotonyl-CoA carboxylase (MCCC). *Protein Expr Purif, 53*(2), 421-427.

75. Jitrapakdee, S., Vidal-Puig, A., Wallace, J. C. (2006). Anaplerotic roles of pyruvate carboxylase in mammalian tissues. *Cell Mol Life Sci, 63*(7-8), 843-854.

76. Stubbe, J., Fish, S., Abeles, R. H. (1980). Are carboxylations involving biotin concerted or nonconcerted? *Journal of Biological Chemistry, 255*(1), 236-242.

77. Yamamoto, Y., Hasegawa, J.-y., Ito, Y. (2012). Kinetic investigation on carbamate formation from the reaction of carbon dioxide with amino acids in homogeneous aqueous solution. *Journal of Physical Organic Chemistry, 25*(3), 239-247.

78. Persat, A., Chambers, R. D., Santiago, J. G. (2009). Basic principles of electrolyte chemistry for microfluidic electrokinetics. Part I: Acid-base equilibria and pH buffers. *Lab Chip, 9*(17), 2437-2453.

79. Kai, Y., Matsumura, H., Izui, K. (2003). Phosphoenolpyruvate carboxylase: three-dimensional structure and molecular mechanisms. *Archives of Biochemistry and Biophysics, 414*(2), 170-179.

80. Mescam, M., Vinnakota, K. C., Beard, D. A. (2011). Identification of the catalytic mechanism and estimation of kinetic parameters for fumarase. *Journal of Biological Chemistry, 286*(24), 21100-21109.

81.   Alberty, R. A., Massey, V., Frieden, C., Fuhlbrigge, A. R. (1954). Studies of the enzyme fumarase. III.1 The dependence of the kinetic constants at 25° upon the concentration and pH of phosphate buffers. *Journal of the American Chemical Society, 76*(9), 2485-2493.

82.   Bisswanger, H. (2014). Enzyme assays. *Perspectives in Science, 1*(1-6), 41-55.

83.   Ellis, K. J., Morrison, J. F., . (1981). Buffers of constant ionic strength for studying pH-dependent processes. *Methods Enzymol, 87*, 405-426.

84.   Colman, Roberta. F. (1975). Mechanisms for the oxidative decarboxylation of isocitrate: implications for control. *Adv Enzyme Regul, 13*, 413-433.

85.   Hurley, J. H., Dean, A. M., Koshland, D. E., Jr., Stroud, R. M. (1991). Catalytic mechanism of NADP(+)-dependent isocitrate dehydrogenase: implications from the structures of magnesium-isocitrate and NADP+ complexes. *Biochemistry, 30*(35), 8671-8678.

86.   Farrell, H. M., Jr., Deeney, J. T., Hild, E. K., Kumosinski, T. F. (1990). Stopped flow and steady state kinetic studies of the effects of metabolites on the soluble form of NADP+:isocitrate dehydrogenase. *J Biol Chem, 265*(29), 17637-17643.

87.   Dean, A. M., Koshland, D. E., Jr. (1993). Kinetic mechanism of Escherichia coli isocitrate dehydrogenase. *Biochemistry, 32*(36), 9302-9309.

88.   Fierke, C. A., Johnson, K. A., Benkovic, S. J. (1987). Construction and evaluation of the kinetic scheme associated with dihydrofolate reductase from Escherichia coli. *Biochemistry, 26*(13), 4085-4092.

89.   Miller, G. P., Wahnon, D. C., Benkovic, S. J. (2001). Interloop contacts modulate ligand cycling during catalysis by Escherichia coli dihydrofolate reductase. *Biochemistry, 40*(4), 867-875.

90.   Miller, G. P., Benkovic, S. J. (1998). Deletion of a highly motional residue affects formation of the Michaelis complex for Escherichia coli dihydrofolate reductase. *Biochemistry, 37*(18), 6327-6335.
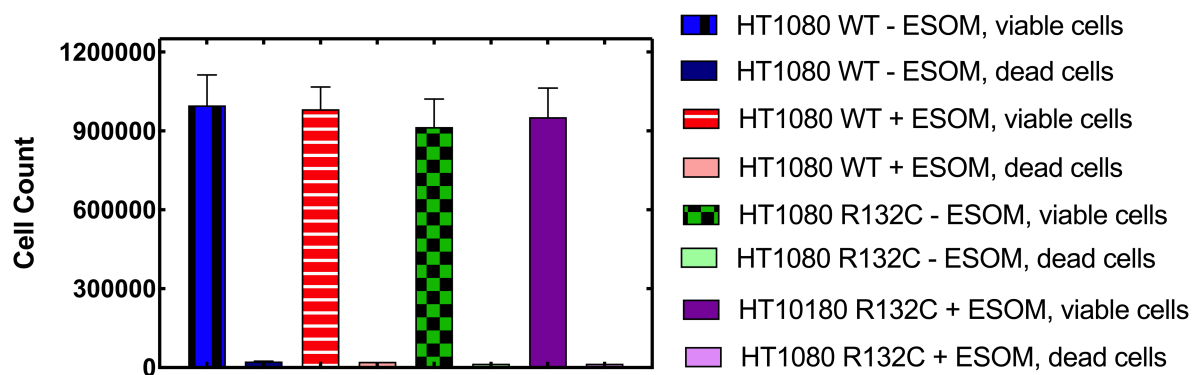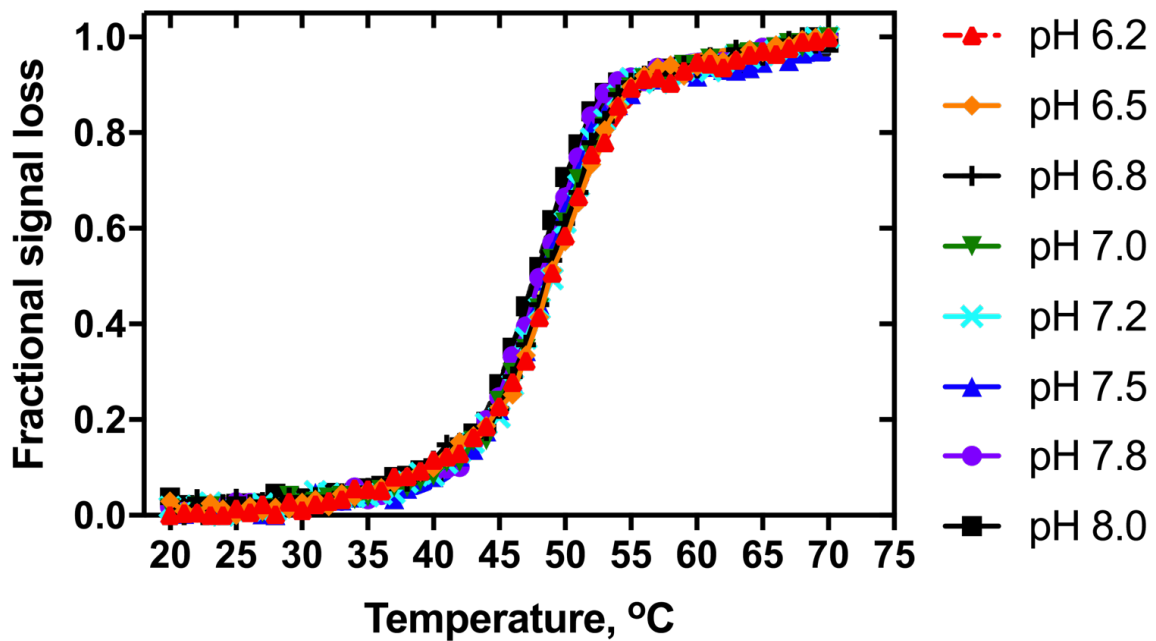
91. Miller, G. P., Benkovic, S. J. (1998). Strength of an interloop hydrogen bond determines the kinetic pathway in catalysis by Escherichia coli dihydrofolate reductase. *Biochemistry, 37*(18), 6336-6342.

92. Miller, G. P., Benkovic, S. J. (1998). Stretching exercises--flexibility in dihydrofolate reductase catalysis. *Chem Biol, 5*(5), R105-113.

93. Yen, K., Travins, J., Wang, F., David, M. D., Artin, E., Straley, K., Padyana, A., Gross, S., DeLaBarre, B., Tobin, E., Chen, Y., Nagaraja, R., Choe, S., Jin, L., Konteatis, Z., Cianchetta, G., Saunders, J. O., Salituro, F. G., Quivoron, C., Opolon, P., Bawa, O., Saada, V., Paci, A., Broutin, S., Bernard, O. A., de Botton, S., Marteyn, B. S., Pilichowska, M., Xu, Y., Fang, C., Jiang, F., Wei, W., Jin, S., Silverman, L., Liu, W., Yang, H., Dang, L., Dorsch, M., Penard-Lacronique, V., Biller, S. A., Su, S. M. (2017). AG-221, a first-in-class therapy targeting acute myeloid leukemia harboring oncogenic IDH2 mutations. *Cancer Discov, 7*(5), 478-493.

94. Intlekofer, A. M., Shih, A. H., Wang, B., Nazir, A., Rustenburg, A. S., Albanese, S. K., Patel, M., Famulare, C., Correa, F. M., Takemoto, N., Durani, V., Liu, H., Taylor, J., Farnoud, N., Papaemmanuil, E., Cross, J. R., Tallman, M. S., Arcila, M. E., Roshal, M., Petsko, G. A., Wu, B., Choe, S., Konteatis, Z. D., Biller, S. A., Chodera, J. D., Thompson, C. B., Levine, R. L., Stein, E. M. (2018). Acquired resistance to IDH inhibition through trans or cis dimer-interface mutations. *Nature, 559*(7712), 125-129.

95. Sahai, A., Laughrey, E., Tannen, R. L. (1990) Relationship between intracellular pH and ammonia metabolism in LLC-PK1 cells. *Am J Physiol*. 258, F103-108 *08*

96. Plaitakis, A., Kalef-Ezra, E., Kotzamani, D., Zaganas, I., Spanaki, C. (2017). The Glutamate Dehydrogenase Pathway and Its Roles in Cell and Tissue Biology in Health and Disease. *Biology (Basel), 6*(1).

97. Kanavouras, K., Mastorodemos, V., Borompokas, N., Spanaki, C., Plaitakis, A. (2007). Properties and molecular evolution of human GLUD2 (neural and testicular tissue-specific) glutamate dehydrogenase. *J Neurosci Res, 85*(5), 1101-1109.

98. Zaganas, I., Pajecka, K., Wendel Nielsen, C., Schousboe, A., Waagepetersen, H. S., Plaitakis, A. (2013). The effect of pH and ADP on ammonia affinity for human glutamate dehydrogenases. *Metab Brain Dis, 28*(2), 127-131.

99.   Qi, F., Pradhan, R. K., Dash, R. K., Beard, D. A. (2011). Detailed kinetics and regulation of mammalian 2-oxoglutarate dehydrogenase. *BMC Biochem, 12*, 53. doi:10.1186/1471-2091-12-53

100.   Badur, M. G., Muthusamy, T., Parker, S. J., Ma, S., McBrayer, S. K., Cordes, T., Magana, J. H., Guan, K. L., Metallo, C. M. (2018). Oncogenic R132 IDH1 mutations limit NADPH for de novo lipogenesis through (D)2-hydroxyglutarate production in fibrosarcoma cells. *Cell Rep, 25*(4), 1018-1026.e1014

101.   De Milito, A., Canese, R., Marino, M. L., Borghi, M., Iero, M., Villa, A., Venturi, G., Lozupone, F., Iessi, E., Logozzi, M., Della Mina, P., Santinami, M., Rodolfo, M., Podo, F., Rivoltini, L., Fais, S. (2010). pH-dependent antitumor activity of proton pump inhibitors against human melanoma is mediated by inhibition of tumor acidity. *Int J Cancer, 127*(1), 207-219.

102.   De Milito, A., Iessi, E., Logozzi, M., Lozupone, F., Spada, M., Marino, M. L., Federici, C., Perdicchio, M., Matarrese, P., Lugini, L., Nilsson, A., Fais, S. (2007). Proton pump inhibitors induce apoptosis of human B-cell tumors through a caspase-independent mechanism involving reactive oxygen species. *Cancer Res, 67*(11), 5408-5417.

103.   Ananieva, E. A., Wilkinson, A. C. (2018). Branched-chain amino acid metabolism in cancer. *Curr Opin Clin Nutr Metab Care, 21*(1), 64-70.

104.   Merk, A., Bartesaghi, A., Banerjee, S., Falconieri, V., Rao, P., Davis, M. I., Pragani, R., Boxer, M. B., Earl, L. A., Milne, J. L., Subramaniam, S. (2016). Breaking Cryo-EM resolution barriers to facilitate drug discovery. *Cell, 165*(7), 1698-1707.

105.   Avellaneda Matteo, D., Wells, G. A., Luna, L. A., Grunseth, A. J., Zagnitko, O., Scott, D. A., Hoang, A., Luthra, A., Swairjo, M. A., Schiffer, J. M., Sohl, C. D. (2018). Inhibitor potency varies widely among tumor-relevant human isocitrate dehydrogenase 1 mutants. *Biochem J, 475*(20), 3221-3238.

2.7 Supplemental Information

A

B
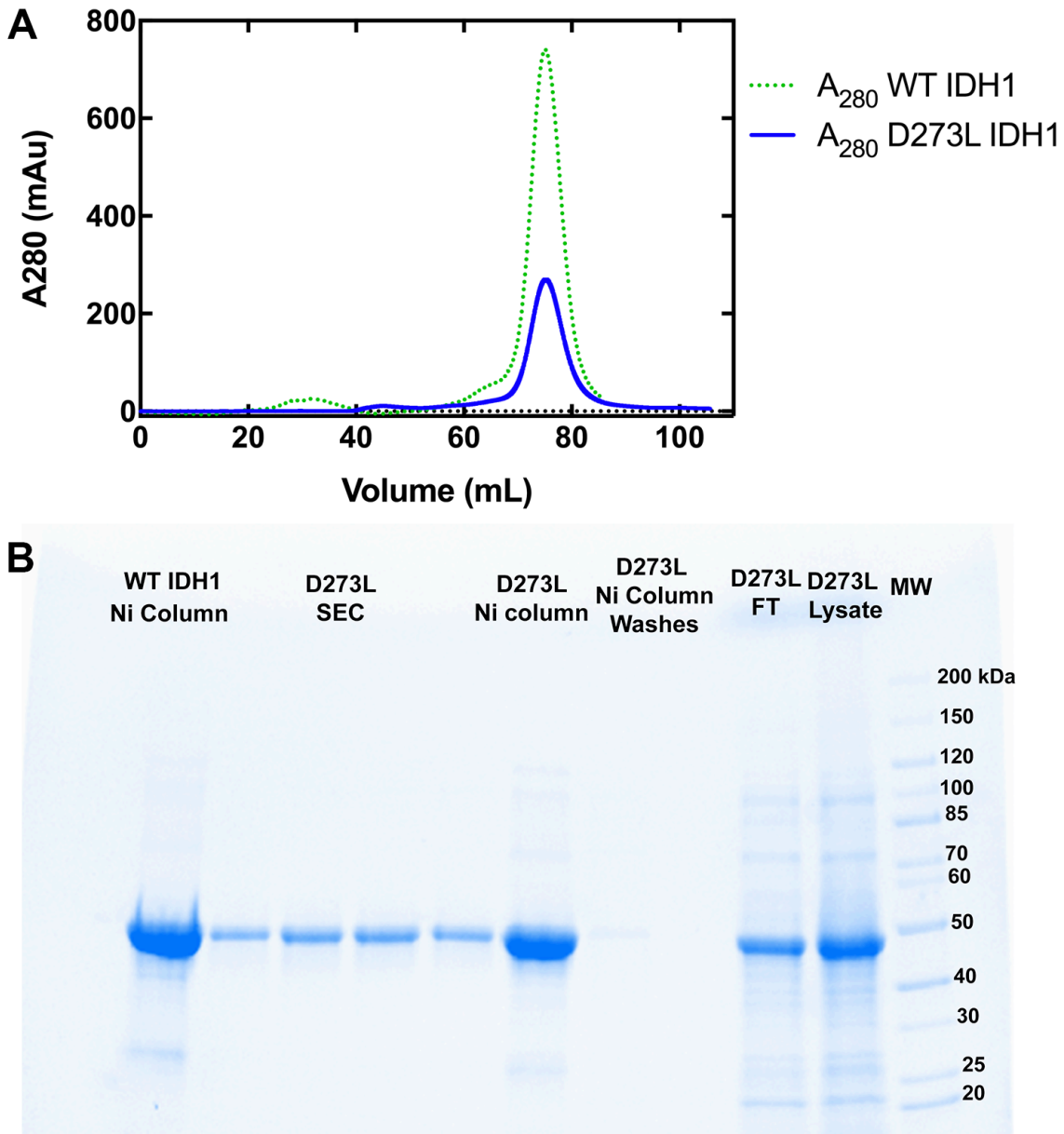
Supplemental Figure S1: (A) WT and D273L IDH1 purified from Ni-NTA affinity resin were assessed using size exclusion chromatography (SEC) to ensure protein was primarily found in the dimer form (~97 kDa dimer, elution peak at fraction 75 mL). Protein concentration for the WT IDH1 dimer is higher than the D273L IDH1 dimer because twice the volume of bacterial culture volume was used during heterologous expression of WT IDH1. (B) Sodium dodecyl sulfate-polyacrylamide gel electrophoresis (SDS-PAGE) analysis of IDH1 to assess protein purity (4-20% TGX Stain-free gel). FT: column flow-through.

Supplemental Figure S2: Circular dichroism (CD) was used to assess secondary structure and obtain Tm values. Tm values for WT IDH1 in KPhos buffer were determined to be as follows $48.9 \pm 0.1°C$ (pH 6.2), $49.0 \pm 0.1°C$ (pH 6.5), $48.7 \pm 0.1°C$ (pH 6.8), $48.61 \pm 0.08°C$ (pH 7.0), 47.8 ± 0.1 (pH 7.2), 48.35 ± 0.09 (pH 7.5), 47.99 ± 0.08 (pH 7.8), $47.66 \pm 0.08°C$ (pH 8.0).

Supplemental Figure S3: ESOM treatment was selected for its ability to robustly alter pH yet minimize cell death. Trypan blue was used to assess live versus dead cell 16 hours after ESOM treatment (200 µM final concentration). Comparisons of cell viability for WT (-/+++ IDH1) and R132C (R132C/+ IDH1) IDH1 cells, both ESOM treated (+ ESOM) and untreated (- ESOM), are shown. Precent viability is as follows: HT1080 WT – ESOM: 97.7%, HT1080 WT + ESOM: 97.8%, HT1080 R132C: – ESOM 98.4%, HT1080 R132C + ESOM: 98.4%.[1]

Supplemental Figure S4: pHinder analysis of the apo IDH1 dimer. (A) Networks of buried ionizable residues in a previously solved structure of apo WT IDH1 bound to NADP[+] (1T09) were identified. Residues involved in these networks are shown in stick form (blue in chain A, grey in chain B), and connected with grey lines. (B) An overlay of the networks identified in apo IDH1 (blue residues, with the protein also shown in cartoon, 1T09), and holo IDH1 (green residues, 1T0L). NADP[+] molecules are shown as lines. (C) Highlight of the ionizable residue networks found in apo WT IDH1 (1T09). For clarity, only the residues cited in the main text as being part of networks involving K212 and K273 are labeled.[2]
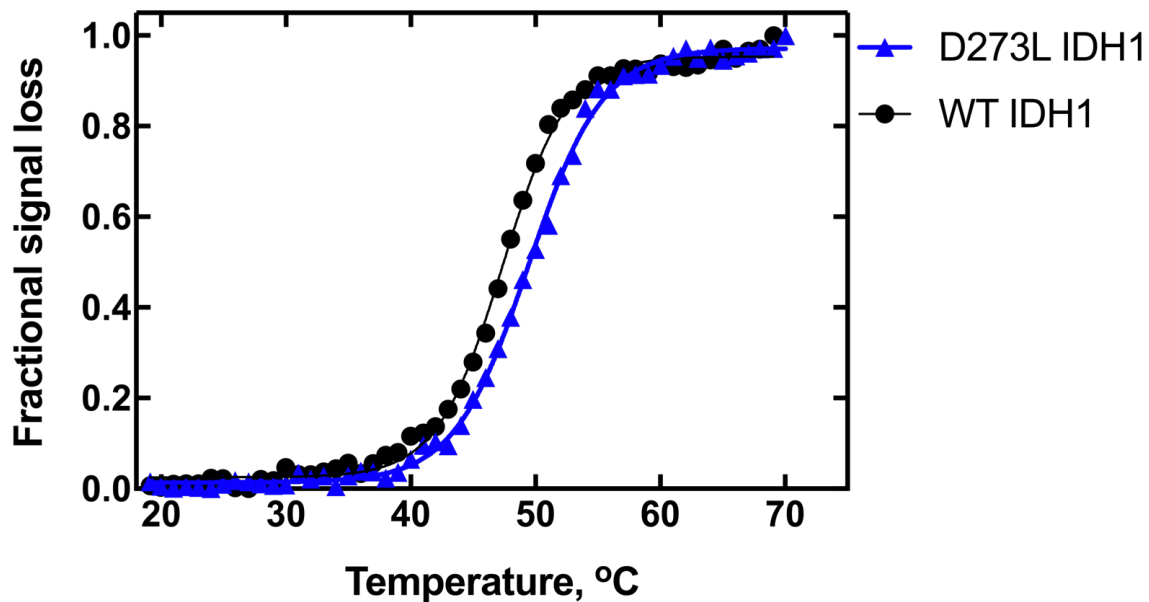
Supplemental Figure S5: Circular Dichroism (CD) was used to compare secondary structure and Tm values between WT and D273L IDH1. Tm values in KPhos buffer at pH 7.5 were determined to be as follows: 48.35 ± 0.09°C (WT IDH1), 49.44 ± 0.08°C (D273L IDH1).
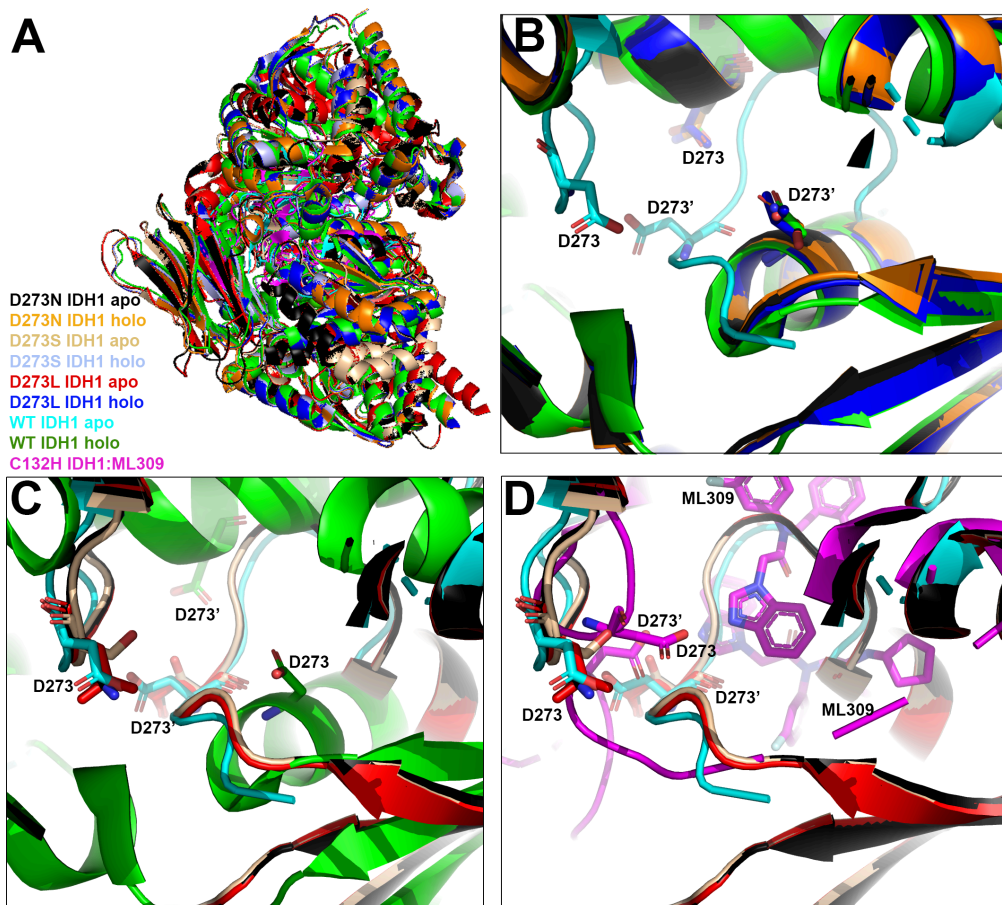
Supplemental Figure S6: Structural models of the D273 mutant series. (A) The orientation of IDH1 as shown in parts B-D, with color schemes indicated for the modeled D273N, D273S, D273L in the WT IDH1 holo (1T0L) and WT IDH1 apo (1T09) backgrounds. Previously solved structures of apo WT IDH, holo IDH1, and C132H IDH1 with ML309 docked (5K11) are also shown. We previously modeled the histidine mutation at residue 132 using a structure of R132C IDH1, and so denote this change as C132H IDH1. (B) Energy-minimized models of D273N, D273S, and D273L IDH1 generated in the WT IDH1 holo structure (αKG, NADP+, and Ca2+ bound, 1T0L) aligned with apo and holo WT IDH1. (C) Energy-minimized models of D273N, D273S, and D273L IDH1 generated in the WT IDH1 apo structure (NADP$^+$ bound, 1T09) aligned with apo and holo WT IDH1. (D) Energy-minimized models of D273N, D273S, D273L IDH1 generated in the WT IDH1 apo structure (NADP$^+$ bound, 1T09) aligned with apo WT IDH1 apo and C132H IDH1 with ML309 docked in the inhibitor binding pocket.[2, 3]

Supplemental Table S1. Absolute quantitation of cell metabolites in HT1080 cells. Patient-derivedHT1080 cells, which contain an endogenous heterozygous R132C IDH1 mutation were used as is (R132C/+ IDH1) or the R132C IDH1 allele was ablated with a T77A mutation to destroy neomorphic (αKG to D2HG) activity in the R132C allele, and made to stably overexpress WT IDH1 (-/+++ IDH1) [1]. To modulate pHi, cell lines received no treatment or were treated with 200 μM ESOM, which decreases the pHi from 7.58 ± 0.04 to 7.39 ± 0.01 (HT1080 -/+++ IDH1), or 7.52 ± 0.06 to 7.31 ± 0.01 (HT1080 R132C/+ IDH1). A norvaline standard was used to quantify each metabolite in the whole cell lysates. These experiments were repeated in duplicate, with averages +/- SD (standard deviation) shown.[4, 5]

Supplemental Table S1: Absolute quantitation of cell metabolites, Continued.

Metabolite	HT1080 (-/+++ IDH1) + no treatment, nmol/10^6 cells	SD	HT1080 (-/+++ IDH1) + ESOM, nmol/10^6 cells	SD	HT1080 (R132C/+ IDH1) + no treatment, nmol/10^6 cells	SD	HT1080 (R132C/+ IDH1) ESOM, nmol/10^6 cells	SD
2HG	**0.07**	**0.01**	**0.05**	**0.00**	**26.85**	**0.73**	**5.76**	**0.20**
3PG	0.62	0.18	0.67	0.19	1.40	0.45	1.20	0.06
GABA	0.06	0.09	0.05	0.01	0.22	0.02	0.05	0.01
Lactate	10.42	1.67	8.52	2.14	16.00	3.80	9.95	0.33
Ketoglutarate	**0.31**	**0.08**	**0.20**	**0.02**	**0.39**	**0.06**	**0.14**	**0.01**
Alanine	2.50	0.19	3.95	0.52	5.02	0.44	4.40	0.04
Asparagine	0.52	0.00	0.76	0.12	1.21	0.07	0.92	0.06
Aspartate	1.70	0.28	1.19	0.47	3.71	0.08	1.34	0.04
Beta-alanine	1.38	0.18	1.09	0.30	2.06	0.15	1.02	0.02
Citrate	0.60	0.00	0.18	0.01	1.10	0.03	0.16	0.03
Cysteine	1.24	0.04	2.61	0.18	1.88	0.23	2.79	0.26
DHAP	60.63	35.21	75.33	41.88	111.42	34.16	149.14	19.31
Fumarate	0.34	0.03	0.11	0.03	0.61	0.03	0.13	0.00
Glutamate	39.05	1.24	41.76	7.41	66.41	5.03	44.46	0.81
Glutamine	34.22	4.75	56.07	5.95	57.20	9.11	48.61	7.41
Glycerol_P(1)	0.54	0.06	1.13	0.25	1.43	0.24	1.98	0.15
Glycerol_P(2)	0.37	0.07	0.73	0.15	0.92	0.22	1.25	0.10
Glycine	14.93	2.33	17.84	4.96	25.45	1.69	18.49	0.51
Hydroxyproline	0.20	0.01	0.24	0.02	0.37	0.01	0.35	0.05
Isocitrate	**0.05**	**0.00**	**0.00**	**0.00**	**0.10**	**0.02**	**0.01**	**0.00**
Isoleucine	4.54	0.04	10.49	0.02	9.30	0.09	10.72	0.05
Leucine	4.08	0.02	9.26	0.22	8.29	0.06	9.84	0.39
Lysine	0.94	0.06	1.94	0.18	1.89	0.11	2.26	0.03
Malate	0.70	0.01	0.28	0.07	1.06	0.13	0.25	0.02
Methionine	1.65	0.02	3.94	0.04	3.26	0.04	4.03	0.01
Ornithine	0.14	0.01	0.26	0.04	0.29	0.02	0.31	0.02
Phenylalanine	3.05	0.03	6.57	0.24	5.96	0.22	6.80	0.14
PEP	0.18	0.03	0.22	0.02	0.39	0.06	0.30	0.03
Proline	1.19	0.02	0.37	0.07	1.49	0.07	0.40	0.02
Pyruvate	1.44	0.68	1.26	0.48	2.58	0.41	0.84	0.10
Serine	3.44	0.46	6.18	1.29	6.25	0.55	7.04	0.04

Supplemental Table S1: Absolute quantitation of cell metabolites, Continued.

Succinate	0.05	0.02	0.04	0.00	0.09	0.00	0.04	0.01
Threonine	20.98	5.28	32.53	9.16	40.37	1.84	31.95	1.36
Tryptophan	1.61	0.17	3.23	0.46	3.12	0.35	3.38	0.16
Tyrosine	2.76	0.08	5.64	0.42	5.51	0.29	6.01	0.05
Urea	0.26	0.10	0.23	0.07	0.33	0.05	0.38	0.18
Valine	4.25	0.06	9.54	0.36	8.68	0.43	10.18	0.10

Supplemental Table S2: Calculated pK_a values for 1T0L holo (WT IDH1) using PROPKA.[2, 6, 7] This program uses the abbreviation "Hid" for histidine.

pH	pKa	pH	pKa	pH	pKa	pH	pKa	pH	pKa	pH	pKa
MET1A	7.91	GLU110A	5.72	LYS233A	10.8	GLU365A	4.32	GLU62B	3.73		
LYS3A	9.5	CYS114A	11.64	TYR235A	11.46	GLU368A	3.98	LYS65B	11.28		
LYS4A	10.43	LYS115A	10.77	LYS236A	10.75	LYS374A	10.55	LYS66B	10.41		
GLU12A	4.1	ARG119A	11.83	GLU240A	4.39	ASP375A	1.01	HID67B	5.31		
ASP16A	5.32	LYS126A	11.47	LYS243A	10.45	CYS379A	11.94	LYS72B	8.96		
GLU17A	3.51	ARG132A	14.5	TYR246A	13.99	LYS381A	10.59	CYS73B	11.89		
ARG20A	13.68	HID133A	3.12	GLU247A	4.93	ARG388A	13.96	ASP79B	1.9		
GLU24A	4.68	TYR135A	13.8	HID248A	2.64	ASP390A	3.24	GLU80B	4.84		
LYS27A	11.85	ASP137A	5.52	ARG249A	11.88	TYR391A	10.22	LYS81B	9.92		
GLU28A	4.56	TYR139A	9.84	ASP253A	4.66	GLU396A	3.47	ARG82B	12		
LYS29A	9.93	ARG140A	13.19	LYS260A	12.39	ASP399A	3.62	GLU84B	4.49		
TYR34A	11.24	ASP143A	5.03	GLU262A	3.53	LYS400A	10.22	GLU85B	4.54		
GLU36A	4.58	LYS151A	10.56	CYS269A	14.06	GLU403A	4.61	LYS87B	10.04		
ASP38A	3.16	GLU153A	4.62	LYS270A	12.69	LYS406A	10.64	LYS89B	10.14		
HID40A	6.5	TYR156A	10.85	TYR272A	9.84	LYS408A	12.05	LYS93B	9.36		
TYR42A	10.73	ASP160A	2.7	**ASP273A**	**6.59**	LYS413A	10.16	ARG100B	12		
ASP43A	3	LYS164A	10.49	ASP279A	5.23	LEU414A	3.39	ARG109B	10.71		

Supplemental Table S2: Calculated p*K*a values for IT0L Continued.

GLU47A	4.77	TYR167A	11.11	TYR285A	13.08	MET1B	7.33	GLU110B	5.66
ARG49A	11.49	HID170A	4.29	CYS297A	12.29	LYS3B	10.25	CYS114B	11.54
ASP50A	2.77	GLU173A	4.65	ASP299A	2.64	LYS4B	10.46	LYS115B	10.77
ASP54A	5.13	GLU174A	4.79	LYS301A	10.41	GLU12B	3.9	ARG119B	11.85
LYS58A	10.47	TYR183A	10.77	GLU304A	7.52	ASP16B	4.92	LYS126B	11.35
ASP59A	3.42	ASP186A	4.33	GLU306A	3.65	GLU17B	3.74	ARG132B	14.17
GLU62A	3.73	LYS187A	11.26	HID309A	3.75	ARG20B	14.41	HID133B	3.12
LYS65A	11.28	GLU190A	5.31	ARG314A	12	GLU24B	4.96	TYR135B	11.2
LYS66A	10.39	ASP191A	4.49	HID315A	7.63	LYS27B	11.86	ASP137B	5.36
HID67A	5.29	HID194A	6.35	TYR316A	11.16	GLU28B	4.59	TYR139B	9.84
LYS72A	8.85	LYS203A	10.35	ARG317A	12.87	LYS29B	10.48	ARG140B	11.9
CYS73A	12.07	TYR208A	10.57	TYR319A	10.33	TYR34B	11.41	ASP143B	5.55
ASP79A	2.5	LYS212A	14.15	LYS321A	10.45	GLU36B	4.31	LYS151B	10.3
GLU80A	4.7	**LYS217A**	**7.97**	GLU324A	4.13	ASP38B	2.81	GLU153B	4.92
LYS81A	10.07	LYS218A	10.49	ARG338A	12.71	HID40B	6.45	TYR156B	11.07
ARG82A	14.45	TYR219A	9.84	HID342A	5.74	TYR42B	10.82	ASP160B	2.62
GLU84A	4.48	ASP220A	3.99	ARG343A	12.65	ASP43B	3.47	LYS164B	10.21
GLU85A	4.03	ARG222A	13.19	LYS345A	10.38	GLU47B	6.94	TYR167B	10.99
LYS87A	10.44	LYS224A	10.76	ASP347A	3.23	ARG49B	11.8	HID170B	4.01
LYS89A	10.35	ASP225A	5.31	LYS350A	10.29	ASP50B	4.13	GLU173B	4.62
LYS93A	9.6	GLU229A	4.16	GLU351A	4.03	ASP54B	5.15	GLU174B	4.75
ARG100A	12	TYR231A	11.72	GLU360A	5.15	LYS58B	10.45	TYR183B	14.01
ARG109A	10.68	ASP232A	4.19	GLU361A	3.82	ASP59B	3.44	ASP186B	4.26

Supplemental Table S2: Calculated p*K*a values for IT0L Continued.

LYS187B	11.37	HID309B	3.74	ARG20C	14.55	HID133C	3.15	GLU247C	4.93
GLU190B	5.36	ARG314B	12	GLU24C	5.06	TYR135C	10.75	HID248C	2.72
ASP191B	4.52	HID315B	7.31	LYS27C	11.87	ASP137C	6.05	ARG249C	11.87
HID194B	6.33	TYR316B	13.77	GLU28C	4.39	TYR139C	9.84	ASP253C	4.47
LYS203B	10.34	ARG317B	12.83	LYS29C	10.16	ARG140C	12.66	LYS260C	12.01
TYR208B	10.54	TYR319B	12.71	TYR34C	11.12	ASP143C	5.08	GLU262C	3.59
LYS212B	14.13	LYS321B	10.4	GLU36C	4.55	LYS151C	10.5	CYS269C	14.04
LYS217B	**8.24**	GLU324B	5.16	ASP38C	3.31	GLU153C	4.71	LYS270C	12.71
LYS218B	10.5	ARG338B	12.93	HID40C	6.66	TYR156C	11.07	TYR272C	9.84
TYR219B	9.84	HID342B	5.72	TYR42C	10.75	ASP160C	2.57	ASP273C	6.4
ASP220B	3.98	ARG343B	12.38	ASP43C	2.94	LYS164C	10.13	ASP279C	5.24
ARG222B	13.21	LYS345B	10.3	GLU47C	6.52	TYR167C	11.16	TYR285C	13.09
LYS224B	10.5	ASP347B	3.66	ARG49C	11.71	HID170C	4	CYS297C	12.26
ASP225B	5.23	LYS350B	10.53	ASP50C	3.9	GLU173C	4.7	ASP299C	2.26
GLU229B	4.17	GLU351B	4.37	ASP54C	5.1	GLU174C	4.59	LYS301C	10.52
TYR231B	11.69	GLU360B	4.61	LYS58C	10.47	TYR183C	13.82	GLU304C	7.57
ASP232B	4.13	GLU361B	3.8	ASP59C	3.43	ASP186C	4.71	GLU306C	3.97
LYS233B	10.79	GLU365B	4.34	GLU62C	3.73	LYS187C	11.26	HID309C	3.78
TYR235B	11.45	GLU368B	3.86	LYS65C	11.3	GLU190C	5.33	ARG314C	12
LYS236B	10.77	LYS374B	10.55	LYS66C	10.4	ASP191C	4.5	HID315C	7.78
GLU240B	4.38	ASP375B	1.17	HID67C	5.24	HID194C	6.36	TYR316C	12.92
LYS243B	10.46	CYS379B	11.99	LYS72C	8.9	LYS203C	10.34	ARG317C	12.94
TYR246B	13.98	LYS381B	10.6	CYS73C	11.98	TYR208C	10.54	TYR319C	9.84

Supplemental Table S2: Calculated p*K*a values for IT0L Continued.

GLU247B	4.92	ARG388B	13.39	ASP79C	1.87	LYS212C	13.46	LYS321C	10.36
HID248B	2.64	ASP390B	3.22	GLU80C	4.8	LYS217C	8.16	GLU324C	4.94
ARG249B	11.86	TYR391B	10.36	LYS81C	10.03	LYS218C	10.48	ARG338C	12.29
ASP253B	5.72	GLU396B	3.69	ARG82C	14.88	TYR219C	9.84	HID342C	5.63
LYS260B	12.35	ASP399B	3.39	GLU84C	4.49	ASP220C	3.99	ARG343C	12.13
GLU262B	3.67	LYS400B	10.24	GLU85C	4.41	ARG222C	13.2	LYS345C	10.35
CYS269B	14.05	GLU403B	4.51	LYS87C	10.07	LYS224C	10.67	ASP347C	3.75
LYS270B	12.73	LYS406B	10.55	LYS89C	10.29	ASP225C	5.21	LYS350C	10.4
TYR272B	9.84	LYS408B	12.01	LYS93C	9.64	GLU229C	4.18	GLU351C	4.52
ASP273B	**6.4**	LYS413B	10.42	ARG100C	12	TYR231C	11.7	GLU360C	5.52
ASP279B	4.08	LEU414B	3.4	ARG109C	10.79	ASP232C	4.14	GLU361C	3.93
TYR285B	13.08	MET1C	7.84	GLU110C	5.71	LYS233C	10.8	GLU365C	4.3
CYS297B	12.21	LYS3C	10.55	CYS114C	11.72	TYR235C	11.46	GLU368C	3.98
ASP299B	2.24	LYS4C	10.37	LYS115C	10.78	LYS236C	10.76	LYS374C	10.55
LYS301B	10.39	GLU12C	3.79	ARG119C	11.82	GLU240C	4.39	ASP375C	1.19
GLU304B	8.21	ASP16C	4.84	LYS126C	11.42	LYS243C	10.46	CYS379C	11.98
GLU306B	3.76	GLU17C	3.57	ARG132C	14.27	TYR246C	13.97	LYS381C	10.61
ARG388C	13.8	ASP79D	2.31	LYS212D	13.61	LYS321D	10.34		
ASP390C	3.15	GLU80D	4.74	LYS217D	8	GLU324D	4.32		
TYR391C	10.25	LYS81D	10.09	LYS218D	10.5	ARG338D	13.75		
GLU396C	3.73	ARG82D	14.31	TYR219D	9.84	HID342D	5.71		
ASP399C	3.5	GLU84D	4.47	ASP220D	3.99	ARG343D	12.03		
LYS400C	11.2	GLU85D	3.62	ARG222D	13.23	LYS345D	10.3		

Supplemental Table S2: Calculated p*K*a values for IT0L Continued.

GLU403 C	3.62	LYS87D	10.43	LYS224 D	10.78	ASP347 D	3.86		
LYS406 C	10.41	LYS89D	10.34	ASP225 D	5.37	LYS350 D	10.58		
LYS408 C	12.19	LYS93D	9.68	GLU229 D	4.16	GLU351 D	3.87		
LYS413 C	9.99	ARG100 D	12	TYR231 D	11.69	GLU360 D	3.79		
LEU414 C	3.37	ARG109 D	10.96	ASP232 D	4.19	GLU361 D	3.88		
MET1D	7.91	GLU110 D	5.72	LYS233 D	10.81	GLU365 D	4.6		
LYS3D	10.52	CYS114 D	11.66	TYR235 D	11.44	GLU368 D	3.94		
LYS4D	10.41	LYS115 D	10.77	LYS236 D	10.75	LYS374 D	10.56		
GLU12 D	3.88	ARG119 D	11.83	GLU240 D	4.4	ASP375 D	1.3		
ASP16 D	5.38	LYS126 D	11.47	LYS243 D	10.46	CYS379 D	12.02		
GLU17 D	4.15	ARG132 D	14.22	TYR246 D	13.97	LYS381 D	10.59		
ARG20 D	13.55	HID133 D	3.07	GLU247 D	4.81	ARG388 D	14.14		
GLU24 D	4.42	TYR135 D	13.8	HID248 D	2.6	ASP390 D	3.27		
LYS27D	11.87	ASP137 D	5.29	ARG249 D	11.87	TYR391 D	10.35		
GLU28 D	4.74	TYR139 D	9.84	ASP253 D	6	GLU396 D	3.04		
LYS29D	10.57	ARG140 D	13.22	LYS260 D	12.33	ASP399 D	3.12		
TYR34 D	11.37	ASP143 D	5.26	GLU262 D	3.54	LYS400 D	10.41		
GLU36 D	4.47	LYS151 D	10.24	CYS269 D	14.03	GLU403 D	4.5		
ASP38 D	2.86	GLU153 D	4.88	LYS270 D	12.75	LYS406 D	10.38		
HID40D	6.51	TYR156 D	11.01	TYR272 D	9.84	LYS408 D	12.01		
TYR42 D	10.77	ASP160 D	2.63	ASP273 D	6.59	LYS413 D	9.74		
ASP43 D	3.26	LYS164 D	10.45	ASP279 D	5.46	LEU414 D	3.46		
GLU47 D	4.78	TYR167 D	10.56	TYR285 D	13.09				

ARG49D	11.54	HID170D	4.26	CYS297D	12.41				
ASP50D	2.78	GLU173D	4.66	ASP299D	2.27				
ASP54D	5.15	GLU174D	4.59	LYS301D	10.41				
LYS58D	10.48	TYR183D	10.83	GLU304D	8.14				
ASP59D	3.41	ASP186D	4.25	GLU306D	3.65				
GLU62D	3.74	LYS187D	11.26	HID309D	3.81				
LYS65D	11.27	GLU190D	5.32	ARG314D	12				
LYS66D	10.41	ASP191D	4.53	HID315D	7.72				
HID67D	5.27	HID194D	6.39	TYR316D	10.99				
LYS72D	8.97	LYS203D	10.33	ARG317D	12.92				
CYS73D	12.06	TYR208D	10.57	TYR319D	10.23				

2.7.1 Supplemental References

1. Ma, S., Jiang, B., Deng, W., Gu, Z. K., Wu, F. Z., Li, T., Xia, Y., Yang, H., Ye, D., Xiong, Y. and Guan, K. L. (2015) D-2-hydroxyglutarate is essential for maintaining oncogenic property of mutant IDH-containing cancer cells but dispensable for cell growth. *Oncotarget. 6,* 8606-8620.

2. Xu, X., Zhao, J., Xu, Z., Peng, B., Huang, Q., Arnold, E. and Ding, J. (2004) Structures of human cytosolic NADP-dependent isocitrate dehydrogenase reveal a novel self-regulatory mechanism of activity. *J Biol Chem. 279*, 33946-33957.

3. Merk, A., Bartesaghi, A., Banerjee, S., Falconieri, V., Rao, P., Davis, M. I., Pragani, R., Boxer, M. B., Earl, L. A., Milne, J. L. and Subramaniam, S. (2016) Breaking Cryo-EM resolution barriers to facilitate drug discovery. *Cell. 165*, 1698-1707.

4. Ratnikov, B., Aza-Blanc, P., Ronai, Z. A., Smith, J. W., Osterman, A. L. and Scott, D. A. (2015) Glutamate and asparagine cataplerosis underlie glutamine addiction in melanoma. *Oncotarget. 6*, 7379-7389.

5. Avellaneda Matteo, D., Grunseth, A. J., Gonzalez, E. R., Anselmo, S. L., Kennedy, M. A., Moman, P., Scott, D. A., Hoang, A. and Sohl, C. D. (2017) Molecular mechanisms of isocitrate dehydrogenase 1 (IDH1) mutations identified in tumors: The role of size and hydrophobicity at residue 132 on catalytic efficiency. *J Biol Chem. 292,* 7971-7983.

6. Sastry, G. M., Adzhigirey, M., Day, T., Annabhimoju, R. and Sherman, W. (2013) Protein and ligand preparation: parameters, protocols, and influence on virtual screening enrichments. *J Comput Aided Mol Des. 27*, 221-234.

7. Schrodinger Release 2018-3, Schrodinger Suite 2018-3 Protein Preparation Wizard.

2.8 Acknowledgements

3. Initial Characterization of Human DNA Polymerase ε and Tumorigenic Mutant

3.1 Abstract

DNA polymerases are responsible for the accurate and efficient replication of the entire genome and its faithful transmission through generations. Considering the size of the human genome and its constant exposure to endogenous and environmental damage, this is not an easy task. A number of different polymerases have evolved that play a role in overall genome maintenance, either through replication, DNA repair pathways, or both. One replicative polymerase, DNA polymerase ε (Pol ε), is responsible for the highly processive, highly accurate DNA replication of the leading strand. Its high fidelity is due to a combination of intrinsic base selectivity and proofreading exonuclease activity that when coupled with post-replication DNA repair pathway yields a mutation rate of about 1 error per 10^9 base pairs. Recently, sequencing studies from The Cancer Genome Atlas (TCGA) have shown Pol ε mutations are present in 7-12% of all endometrial cancers. The majority of the mutations in Pol ε occur in the exonuclease domain, though many polymerase domain mutations were also discovered. Polymerase domain mutations in human Pol γ and Pol δ are known to contribute to neurodegenerative diseases, such as Alpers syndrome, and colon and liver cancer, respectively. The minimal kinetic mechanism of WT Pol ε has been previously characterized; however, a detailed kinetic investigation of tumorigenic mutants has not yet been undertaken. We have selected a Pol ε fingers domain mutant, F815L, for kinetic study due to its proximity to the incoming nucleotide, which suggests that this residue plays a role in nucleotide incorporation and efficiency. Here we present initial findings on the mechanism of

incorporation, fidelity, and processivity of novel Pol ε mutation found in endometrial tumors.

3.2 Introduction

DNA polymerases are responsible for the highly accurate replication of the genome and/or genome maintenance. These functions are critical for successful proliferation and survival of the organism. DNA polymerases are arranged into seven families (A, B, C, D, X, Y, and RT) that are categorized based on structural and functional similarities. While separate families exist, all DNA polymerases share a structurally conserved polymerase core consisting of finger, palm, and thumb subdomains.[1] In addition to the conserved polymerase core, DNA polymerases may possess additional domains and structural features that add to their functional diversity *in vivo*. For example, replicative polymerases from Families A and B possess an exonuclease domain that contains conserved aspartate residues that coordinate divalent metal ions at the active site to excise incorrectly paired nucleotides.

Highly accurate DNA is important for the continued survival from generation to generation and is driven by the high-fidelity catalytic mechanism of replicative DNA polymerases. The intrinsic fidelity of a replicative polymerase in addition to genome maintenance pathways ensure that DNA is copied with an error rate of one nucleotide per 10^6-10^8 nucleotides.[2] It was originally thought that the free energy differences between incorrectly paired nucleotides and correctly paired nucleotides was sufficient to account for highly accurate DNA replication.[2,3] However, more recent findings suggest that there are more factors at play in the high-fidelity kinetic mechanism of DNA polymerases. Overall, nucleotide selectivity is governed by base stacking interactions, solvation effects,

polymerase interactions with DNA at the minor groove, template base pair geometry, and induced-fit conformational changes.[2, 3, 4] These factors contribute to an error rate of one error per 10^4-10^7 nucleotides, showing that the polymerase domain alone is not able to attain maximum accuracy.[5-9] 3' – 5' proofreading activity from the exonuclease domain,
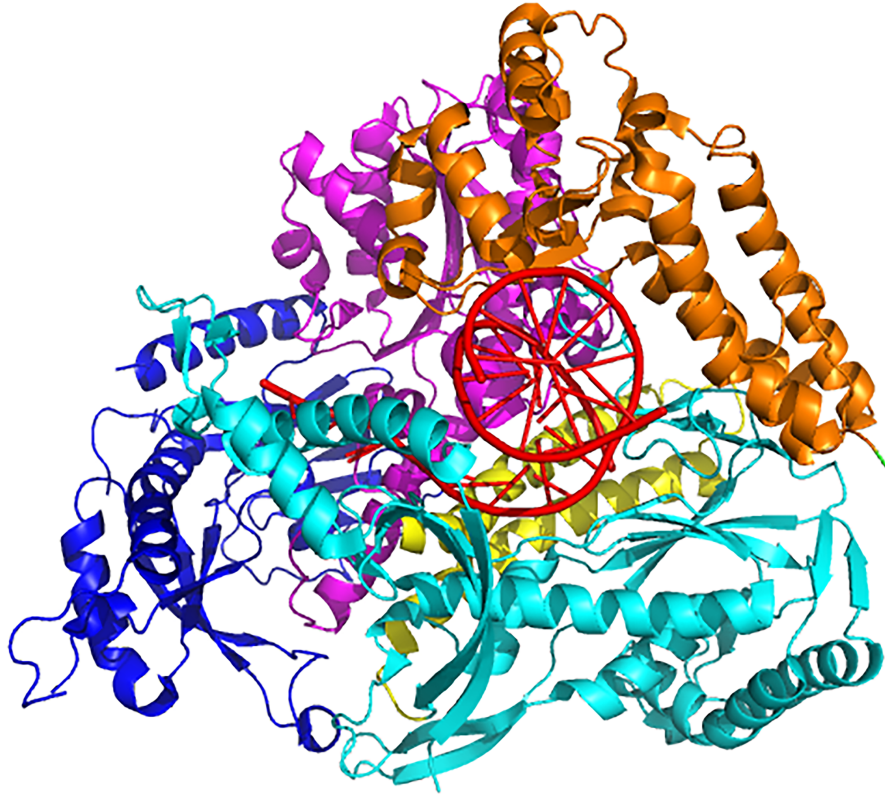

Figure 1: Structure of yeast DNA Polymerase epsilon (PDB 4M8O) colored by domain.[21] The *N*-terminal domain is colored in blue, the exonuclease domain is colored in magenta, the palm domain is colored in cyan, the fingers domain is colored in yellow, and the thumb domain is colored in orange. The DNA is colored in red.

found in A and B family polymerases, improves fidelity up to 200-fold in some polymerases.[5, 10, 11]

The main B family polymerases are responsible for the bulk of DNA synthesis in eukaryotes: Pol α, Pol δ, and Pol ε.[12] Pol α has both primase and polymerase activity, and initiates DNA synthesis by polymerizing a short primer. Genetic studies show that

the division of labor at the replication fork consists of Pol δ replicating the lagging strand and Pol ε replicating the leading strand during DNA synthesis.[13-16] Human Pol ε is a heterotetramer that consists of a catalytic subunit, p261, in addition to three smaller subunits p59, p12, and p17.[17] Barriers in heterologous expression of the four subunit Pol ε holoenzyme led us to express a truncated variant of the *N*-terminal domain that consists of all the conserved regions and motifs necessary for catalysis.[18-20] The structure of human Pol ε still remains unsolved, but the structure of the truncated catalytic subunit of yeast Pol ε is known (Figure 1).[21] Pol ε has two active sites: an exonuclease domain responsible for proofreading, and a polymerase domain that accurately inserts dNTPs.[21, 22] The polymerase domain structure shows the canonical right-hand configuration consisting of finger, palm, thumb domains, and the exonuclease domain.[21] The palm domain contains a unique P domain that plays a role in the highly processive DNA synthesis catalyzed by human Pol ε.[21]

An extensive sequencing analysis of human cancer exomes revealed an increased prevalence of mutations in the gene encoding human Pol ε in a sample of colorectal and endometrial tumors.[23-26] The most prevalent change was the replacement of P286 by arginine, though several additional mutations in the human Pol ε exonuclease and polymerase domains were also observed. In a yeast model, P301R Pol ε (corresponding to human P286R) had a 50-fold higher mutator effect than an exonuclease deficient Pol ε.[27] In mouse models, this mutation led to a hypermutated phenotype and resulted in cancer.[28, 29] Human Pol ε polymerase domain mutations have also been identified .[30] It is well established that mutations in the polymerase domain of replicative polymerases can alter fidelity and enzymatic efficiency to cause disease.[31, 32] For example, the degree

of catalytic dysfunction and infidelity of polymerase domain mutants in human mitochondrial DNA polymerase γ mirror the severity of mitochondrial disease seen in patients.[33] The polymerase domain mutant, R696W, in human Pol δ has been identified in colon and liver tumors and leads to a hypermutated phenotype in yeast.[34] As more human Pol ε mutations are identified in tumors, is it crucial to understand the kinetic
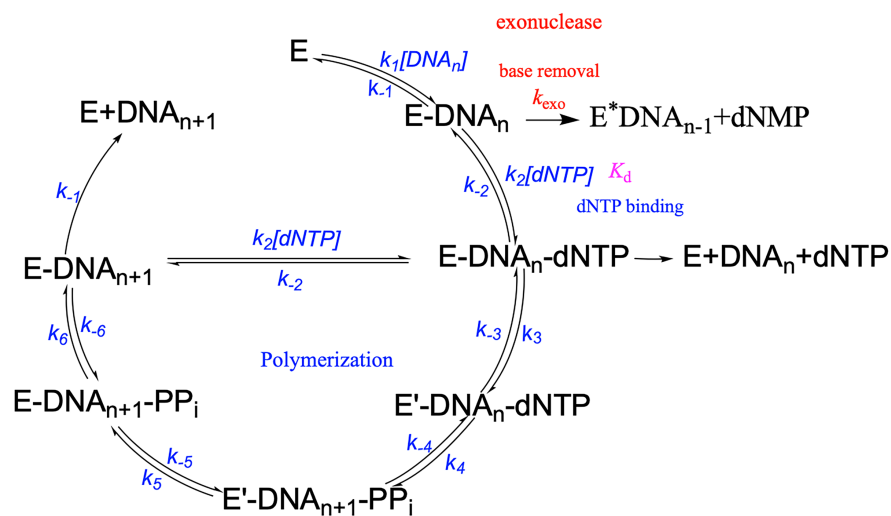

Figure 2: The minimal kinetic mechanism of Pol ε.

features of dNTP incorporation and excision as it helps determine mechanisms of genome instability.

The minimal kinetic mechanism for human Pol ε was recently established (Figure 2).[2, 35-37] The first study on WT human Pol ε used the truncated catalytic subunit, and revealed that human Pol ε inserts the correct nucleotide via an induced fit mechanism that is preceded by a conformational change, similar to many polymerases.[35] The replicative fidelity of truncated WT human Pol ε was shown to depend largely on the differences between incorporation rates and binding constants of correct and incorrect nucleotides.[2] In addition, this study confirmed that the exonuclease domain increases

114

fidelity, and that human Pol ε is highly efficient at removing mismatched nucleotides rather than extending past a mismatch.[2] The complete human Pol ε heterotetramer was successfully purified from an insect cell system, and subsequent investigations found that the additional subunits enhanced DNA binding, moderately improved nucleotide incorporation fidelity, and decreased exonuclease activity compared to the WT human
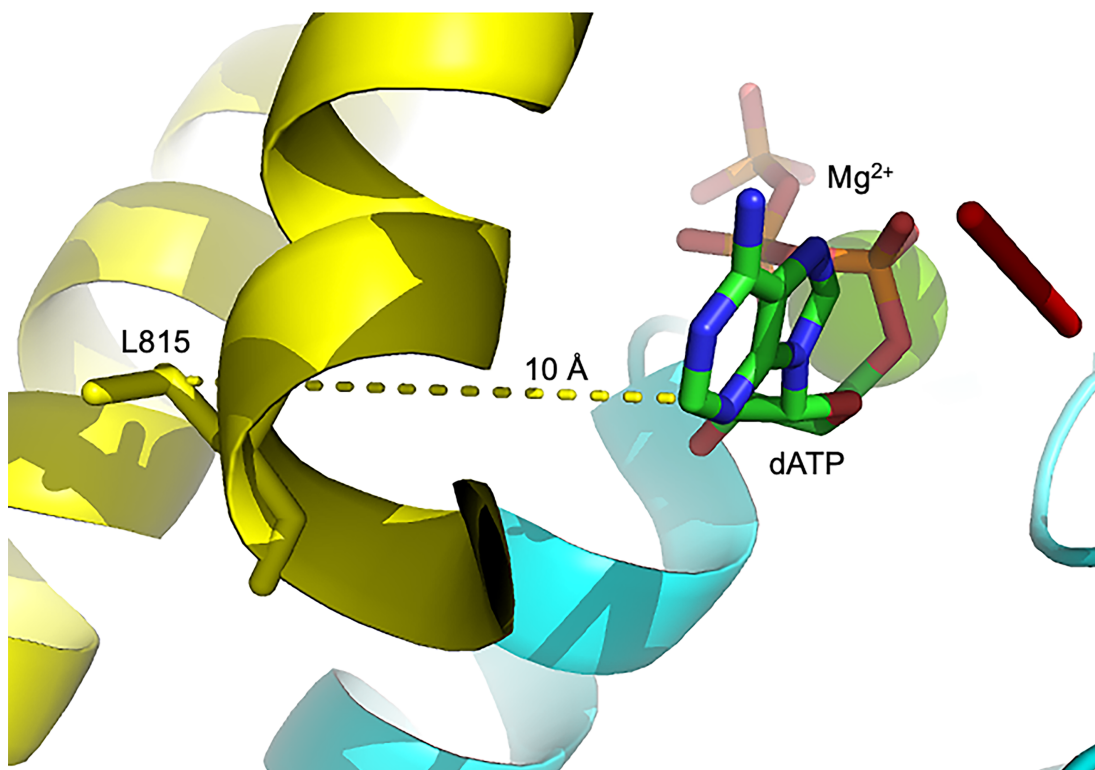

Figure 3: Energy minimized structure of yeast Pol ε featuring the F815L mutant (human numbering) (PDB code 4M8O).[21] The mutation lies 10 Å away from the incoming nucleotide.

Pol ε truncated catalytic subunit.[36, 37]

While many mutations in the polymerase domain of human Pol ε have been identified in The Cancer Genome Atlas (TCGA), none has been studied kinetically *in vitro* or *in vivo*. We chose the polymerase domain mutation, F815L, located in the fingers subdomain. Currently no structures exist for WT human Pol ε and of any mutants, so I we made energy minimized structures of the F815L mutant using the structure of yeast

Pol ε. Based on these structures, we hypothesize that the leucine mutant ablates a critical interaction that stabilizes the formation of an α-helix in the fingers domain of WT Pol ε. This residue is located 10 Å from the incoming dNTP, suggesting a role in nucleotide selection and incorporation efficiency (Figure 3). Here we report a preliminary investigation to distinguish the mechanistic effects that F815L human Pol ε has on correct nucleotide incorporation efficiency and substrate affinity using pre-steady-state kinetics. Ultimately, we will grow this work to a more thorough investigation of mechanisms of fidelity.

The kinetic characterization of human Pol ε mutants is novel. First, will we determine the maximum insertion rate, dissociation constant, and enzymatic efficiency for each correct and incorrect dNTP with WT human Pol ε and F815L human Pol ε. This will serve as a baseline for comparison to the F815L mutant. By measuring the enzymatic efficiencies of the incorporation reaction for correct and incorrect nucleotides, we can compare the fidelities of WT human Pol ε and F815L human Pol ε. We will then compare DNA replication processivity between WT and F815L human Pol ε. Since F815L mutant is in the fingers domain of human Pol ε, which is responsible for selecting the correct incoming nucleotide,[1] we hypothesize that the F815L mutant will insert the correct nucleotide with a lowered enzymatic efficiency compared to WT human Pol ε. We predict this mechanism will occur either through an increased dissociation constant for dNTP binding, a decreased maximum correct dNTP incorporation rate, or both. In addition, we expect that the F815L mutant will insert an incorrect nucleotide with increased efficiency compared to WT human Pol ε, either through a decreased dissociation constant for dNTP binding, an increased maximum incorrect dNTP insertion rate, or both. We predict that

the F815L mutant will be less processive than WT human Pol ε in that it will take more time to fully extend a given DNA primer/template. Future experiments will include investigating the incorporation with chemotherapeutic nucleoside analogues to determine selectivity factors and probe for potential chemotherapeutic sensitivity.

3.3 Full Length DNA Polymerase Epsilon: Materials and Methods

3.3.1 Materials for Full-length Pol ε

The construct for the full-length human Pol ε was graciously donated by Dr. Sreerupa Ray (Yale University). The cDNA construct of the p261 subunit of human Pol ε was cloned into the mammalian expression vector phCMV1 with two consecutive repeats of a maltose binding protein (MBP) tag and a PreScission Protease (LEVLFQ/GP) cleavage site near the *N*-terminus of the protein to increase expression levels and solubility and facilitate purification (Figure 4). HEK293T cells, Dulbecco's modified eagle medium (DMEM), Dulbecco's phosphate buffered saline (DPBS), 0.25% trypsin-EDTA solution, 293 Freestyle (293F) cells, calcium chloride ($CaCl_2$), HEPES, sodium phosphate dibasic, sodium chloride (NaCl), glycerol, methanol, ethanol, DH5α cells, Luria broth (LB), LB-agar, kanamycin sulfate, chloroquine , polyethylenimine (PEI) 25 kDa, Tween-20, NP-40, dithiothreitol (DTT), bicinchoninic acid assay (BCA) Kit, rabbit polycolonal anti-maltose binding protein (anti-MBP) antibody, and phospho-buffered saline (PBS) premixed packets were obtained from Fisher Scientific (Hampton, NH). Goat anti-rabbit secondary antibody with horse radish peroxidase (HRP) was obtained from Sigma (St. Louis, MO). Fetal bovine serum (FBS) was obtained from VWR (Radnor, PA). Qiagen MagePrep Kit was obtained from Qiagen (Hilden, Germany). Pre-stained SDS-PAGE gels (4-20%) and Western Blot Kit were obtained from Bio-rad (Hercules, CA). Amylose

resin was obtained from New England Biolabs (Ipswich, MA), and Q-sepharose resin was obtained from GE Life Sciences (Boston, MA)

3.3.2 DNA Amplification and Purification of Full-Length Pol ε

Figure 4: Schematic representation of the cDNA construct for human full-length Pol ε.

The cDNA construct of the p261 subunit of human Pol ε was cloned in the phCMV1 mammalian expression vector and was amplified in DH5α competent cells, grown in LB media supplemented with 30 μg/mL kanamycin, and purified per manufacturer's instructions using the Qiagen Megaprep Kit (Hilden, Germany). Plasmid concentrations were quantified by UV spectroscopy at 260 nm and stored at -20 °C until further use. I purified the cDNA for transfection.

3.3.3 Cell Culture

Human embryonic kidney cells (HEK) 293T cells expressing a mutant version of the SV40 large T antigen were cultured in DMEM supplemented with 10% Fetal Bovine Serum, at 37°C, and 5% CO_2. 293 Freestyle (293F) were grown in suspension with 293F media at 8% CO_2 and 120 rpm. I performed all the cell culturing.

3.3.4 Transient Transfection of Full-Length Pol ε

The cell culture media was optimized by removing antibiotics to increase transfection efficiency. HEK293T were seeded at a density of 9.0×10^6 cells on a 15 cm dish. After 24 hours, the cells reached approximately 60-70% confluency and were treated with 25 μM chloroquine prior to transfection. Cell culture media was changed to remove chloroquine, and cells were transfected with 30 μg of plasmid DNA using the

calcium phosphate method. Briefly, plasmid DNA was added to a 2x HEPES buffered saline (HBS) (pH 7.05) solution and gently bubbled while a cold solution of 250 mM $CaCl_2$ was added. In this method, a precipitate forms with the DNA conjugated to the precipitate, the solution is added to the plated cells, and the cells take up the DNA through endocytosis. At 7-10 hours post-transfection, cell culture media was changed to remove the transfection reagent. At 36 hours post-transfection, the cells were washed with DPBS and lysed in lysis buffer (50 mM HEPES pH 7.5, 250 mM NaCl, 1% NP-40, 1 mM DTT) with gentle rocking at 4 °C. Extracts were clarified by centrifugation at 12,000 rpm for 10 minutes at 4 °C. The supernatants were then transferred to separate tubes for further analysis.

Transient transfection experiments were also performed in 293F cells in parallel. This cell line is derived from the HEK 293 cell line and can reach high densities in suspension cultures using serum-free media. These experiments also used a different transfection reagent, polyethylenimine (PEI), an inexpensive polymeric reagent that has been used for transfection in mammalian cells by forming PEI/DNA complexes that fuse to cell membranes and are taken up by the cell through endocytosis much more readily than calcium phosphate-based transfections. 293F cells were grown in suspension in 293F media until a density of 1.0×10^6 cells was reached. Cells were transfected with 1 μg of DNA per million cells using PEI. Briefly, 300 μg of DNA was added to 30 mL of PBS, followed by 0.6 mL of 1 mg/mL PEI. The solution was vortexed and allowed to incubate at room temperature for 10-20 minutes. The PEI/DNA solution was then added to the cells and incubated at 37 °C, 8% CO_2, and 120 rpm for 48 hours. The cells were then harvested by centrifugation and lysed in lysis buffer. Extracts were clarified via

centrifugation and stored until further processing. I performed the transient transfections in both cell lines.

3.3.5 Western Blot Analysis

A western blot was performed to confirm the expression of Pol ε. For each sample, total protein was determined using a BCA kit. Total protein (30 μg) was separated on an 4-20% pre-cast/pre-stained SDS-PAGE gel and transferred to a polyvinylidene difluoride (PVDF) membrane per manufacturer's instructions (Bio-rad Hercules, CA). After the transfer step, membranes were blocked with 3% bovine serum albumin (BSA) in phosphate buffered saline with 0.1% Tween (PBST) for 1 hour. Rabbit anti-MBP was diluted 1:1000 in PBST with 3% BSA, added to the membrane, and incubated overnight at 4 °C. The following day the membrane was washed 5 times with PBST, followed by incubation with a goat anti-rabbit horseradish peroxidase secondary antibody (1:10,000 dilution in PBST with 3% BSA) for 1 hour at room temperature. The membrane was then washed 5 times with PBST, and imaged using autoradiography film (Figure 5A, 5C). I verified the expression of full-length Pol ε via Western immunoblotting.

3.3.6 Purification Attempt of Full-Length Pol ε

The clarified lysate from the transient transfection was incubated with amylose resin (New England Biolabs) on a rotator overnight and passed through in a disposable plastic column. The initial flow-through was collected, and the remaining resin was washed with 100 mL Buffer W (50 mM HEPES pH 7.5, 10 mM NaCl, 0.5 mM EDTA, and 1 mM DTT). Bound protein was eluted using 20 mL maltose elution buffer (50 mM HEPES pH 8.2, 10 mM NaCl, 0.5 mM EDTA, 10% glycerol, 10 mM maltose, 1 mM DTT). Fractions were monitored for protein via absorbance at 280 nm. Protein-containing fractions were

analyzed via SDS-PAGE. However, peak fractions from the amylose column could only

be visualized with more sensitive silver staining. Pol ε in peak fractions was only visible

using Coomassie staining after pooling and concentrating peak fractions from 10 mL to

250 μL and applying large contrast to the image (Figure 5B). Presence of Pol ε in fractions

was determined by Western immunoblotting, but only with extended (9-10 minute)

exposure times.   Further purification steps using a Q-sepharose column resulted in

protein yields that were undetectable by Coommassie staining, silver staining, and

Western immunoblotting and unsuitable for any downstream analysis of Pol ε. Therefore,

we opted to express a truncated and catalytically active variant of Pol ε in bacteria. I

performed the purification of full-length Pol ε.
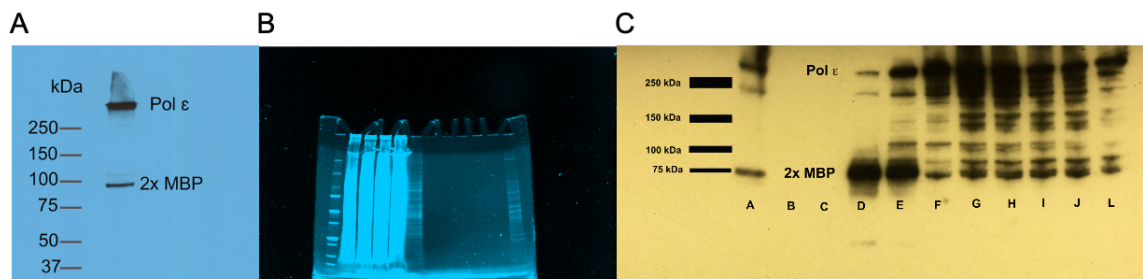
Figure 5: Expression and purification results of full-length human Pol ε. **A.** Western immunoblot analysis using an anti-MBP antibody confirmed the presence of DNA Polymerase epsilon in HEK 293T cells following calcium phosphate-based transfection. **B.** SDS-PAGE analysis of purification from amylose column. DNA Polymerase epsilon was only detectable after heavy computerized contrast was applied to the image. **C.** Western blot analysis using anti-MBP antibody confirmed the presence of DNA Polymerase epsilon in suspension 293F cells following transfection with PEI. Both methods resulted in low yields of protein and were not developed further.

3.4 Materials and Methods for First Pol ε Bacterial cDNA Construct

3.4.1 Materials

A truncated construct of the catalytic subunit of human Pol ε encoding for residues

1 – 1189 of the large p261 subunit was designed by Christal Sohl and purchased from

IDT.   The Pol ε catalytic domain construct was cloned in the pGEX-4T-3 vector to

generate pGEX4T3(thrombin)/Polε(1 – 1189)-His$_6$ (Figure 6). Tris-hydrochloric acid, Tris

base, sodium hydroxide (NaOH), sodium chloride (NaCl), magnesium acetate

(Mg(C$_2$H$_3$O$_2$)$_2$), DTT, BL21 (DE3) Rosetta competent cells, LB-Agar, rabbit polyclonal

anti-hexahistidine tag antibody, ampicillin sodium, Luria broth, 1-thio-β-D-

galactopyranoside (IPTG), ethylenediaminetetraacetic acid (EDTA), EDTA-free protease

inhibitor tablets, glycerol, ammonium persulfate (APS), tetramethylethylenediamine

(TEMED), boric acid, cobalt resin, and GSH resin were all obtained from Fisher Scientific

(Hampton, NH). β-Mercaptoethanol (β-ME) was obtained from MP Biomedicals in Santa

Ana, CA. Imidazole was obtained from Acros Organics (Fisher Scientific Hampton, NH).

Sequel Parts A and B were obtained from American Bio (Canton, MA). [γ-^{32}P]ATP was

obtained from Perkin Elmer, dNTPs were obtained from Thermofisher Scientific, and

HPLC-purified DNA substrates were purchased from Integrated DNA Technologies.

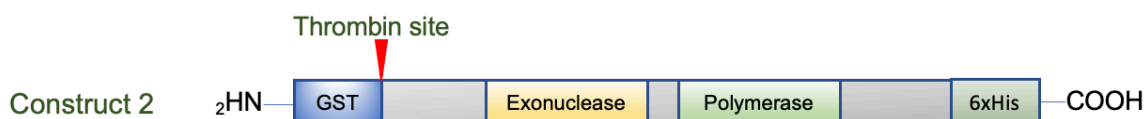### 3.4.2 Protein Expression and Purification

Figure 6: Schematic representation of the cDNA encoding for construct 2.

The cDNA construct coding for Pol ε was transformed into BL21-Gold Rosetta cells

(Figure 6). To generate an exonuclease-deficient mutant, two rounds of site directed

mutagenesis were used to make three amino acid substitutions (D275A/E277A/D368A)

with the following primers: D275A/E277A (forward primer 5' –

CCCAGTCGTCTTAGCATTTGCCATCGCAACAACAAAATTACCATTGAA – 3', reverse

primer 5' – TTCAATGGTAATTTTGTTGTTGCGATGGCAAATGCTAAGACGACTGGG –

3'), and D368 (forward primer 5' – CAATGGTGATTTCTTCGCCTGGCCCTTTGTAGAAG – 3', reverse primer 5' – CTTCTACAAAGGGCCAGGCGAAGAAATCACCATTG – 3'). A transformed colony was inoculated in LB media with 50 μg/mL ampicillin and grown at 37 ˚C until an optical density at 600 nm of 0.6-0.7 was reached. Then, the temperature was reduced to 20 ˚C and the cultures were induced with 0.5 mM IPTG and incubated for 16 hours. The cells were harvested by centrifuging at 8,000 rpm at 4 ˚C for 15 minutes. The pellets were resuspended in lysis buffer (150 mM Tris-HCl pH 7.8 at 4˚C, 500 mM NaCl, 10% glycerol, and 5 mM β-mercaptoethanol, and a protease inhibitor tablet) and lysed in a microfluidizer set to 16,000 psi. The lysate protein was separated by centrifugation at 18,000 rpm for 45 minutes at 4 ˚C. The supernatant containing the soluble protein was incubated with GSH resin for 1 hour at 4˚C with rocking, and then poured through a disposable plastic column. The resin was washed with lysis buffer and transferred to a separate tube. Then, 50-80 units of thrombin protease were added to cleave the GST tag, and the slurry was incubated for 10 hours at 4˚C, then transferred back to the column. The goal was to retain the GST tag on the column, while the cleaved Pol ε would be eluted from the column. Fractions were monitored for protein by absorbance at 280 nm. Peak fractions were combined and passed over a cobalt column. The column was then washed with cobalt wash buffer (150 mM Tris-HCl pH 7.8 at 4 ˚C, 500 mM NaCl, 10% glycerol, 5 mM β-mercaptoethanol, and 10 mM imidazole). Bound protein was eluted with a linear gradient of 10 mM imidazole to 200 mM imidazole in cobalt wash buffer. Fractions were monitored for protein by absorbance at 280 nm. Fractions containing protein were analyzed for > 95% purity on a 4-20% SDS-PAGE gel (Figure 7A). Fractions containing Pol ε were concentrated and dialyzed into storage buffer (50 mM Tris-Acetate pH 7.4, 1

mM DTT, 15% glycerol, 10 mM NaCl). Protein concentration was calculated by measuring the absorbance at 280 nm and using an extinction coefficient of 155,310 $M^{-1}cm^{-1}$. The protein was then flash frozen and stored at -80˚C until further use. An alumna of the Sohl lab, Anna Uvarova, designed and performed the mutagenesis, and optimized the purification protocol.

3.4.3 Primer Template Annealing

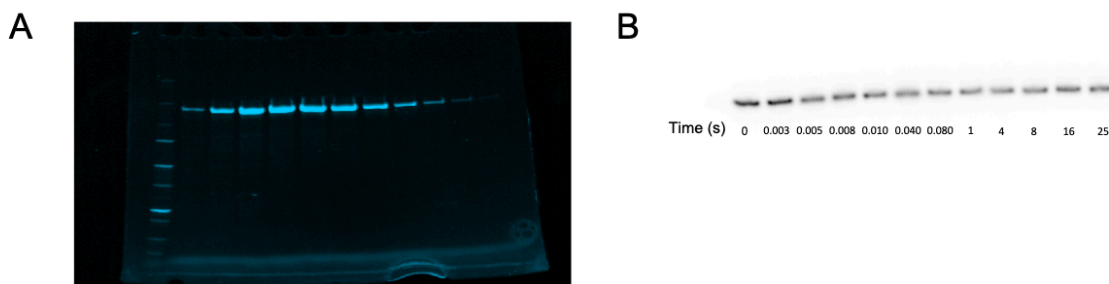HPLC purified DNA substrates (Table 1) were purchased from Integrated DNA

Figure 7: Construct 2 purification and activity results. **A.** SDS-PAGE analysis shows that purification was successful. **B.** Denaturing gel electrophoresis reveals that Pol ε was inactive in single nucleotide incorporation assays.

Technologies. The 21-mer primer strand was 5'-radiolabeled by incubation with $[\gamma-^{32}P]$ATP and T4 polynucleotide kinase for 30 minutes at 37 ˚C and then purified from free $[\gamma-^{32}P]$ATP by passing through a Bio-Spin column (Bio-rad). The 5'-radiolabled primers were annealed to 41-mer template by incubating the 21-primer with 1.4-fold molar excess template at 95 ˚C for 5 minutes before cooling slowly to room temperature. I performed the primer-template annealing.

3.4.4 Reaction Buffers

All assays using Pol ε were performed at 20 ˚C using reaction buffer (50 mM Tris-acetate pH 7.4 at 20 ˚C, 8 mM Mg(C$_2$H$_3$O$_2$)$_2$, 1 mM DTT, 10% glycerol, and 0.1 mg/mL

BSA). Fast reactions were carried out using a rapid chemical quench flow apparatus (KinTek). Notably, 37 ˚C was not used because the rate constant of product formation catalyzed by Pol ε was too fast (k_{pol} > 500 s^{-1}) to be accurately determined using a rapid chemical quench apparatus.[2, 35-39] Final concentrations are reported after 1:1 mixing in the rapid chemical quench.

Table 1: DNA Substrates used in this kinetic investigation.

D-1	5'–TTGGATGACAGAAACACTTT 3'–AACCTACTGTCTTTGTGAAAAGCTGTATCACACCACCACGG
D-2	5'–CGCAGCCGTCCAACCAACTCA 3'–GCGTCGGCAGGTTGGTTGAGTAGCTGCTAGGTTACGGCAGG

3.4.5 Pre-steady-State Kinetic Assays

In the burst assays, a pre-incubated of solution of 20 nM Pol ε and 100 nM 5'-radiolabeled DNA (D-1) in reaction buffer was rapidly mixed with varying concentrations of dCTP and 8 mM Mg^{2+}. All reactions were quenched with the addition of 0.37 M EDTA. All reactions were performed using a rapid chemical quench-flow apparatus (KinTek).

3.4.6 Product Analysis

Reaction products were separated on a 20% denaturing polyacrylamide gel containing 8 M urea and 1X Tris-Borate-EDTA running buffer (TBE). The gel was quantified with a Typhoon FLA 9500 scanner (GE Healthcare) and ImageQuant (Molecular Dynamics). There was no detectable activity with this construct (Figure 7B). Despite troubleshooting efforts of changing the concentration of dNTP, increasing reaction time, increasing the temperature to 37 ˚C, measuring the CD spectrum of the protein to ensure proper folding, using a size exclusion column to assess aggregation, and varying the pH and buffer component of the reaction buffer, there was still no detectable activity. Lack of catalytic activity in this construct forced us to use a second

construct that was successfully used for kinetic assays previously.[15, 38] The second construct is described below. I ran the activity assays and performed the product analysis.
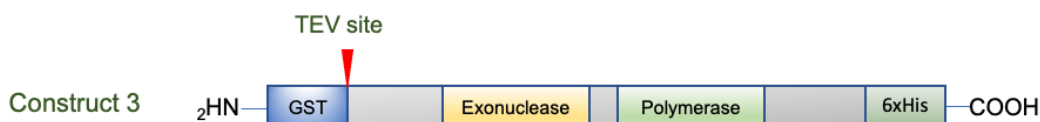

Figure 8: Schematic representation of cDNA construct 3.

3.5 Materials and Methods for Second Pol ε Bacterial cDNA Construct

3.5.1 Materials

A truncated construct of the catalytic subunit of human Pol ε encoding for residues 1 – 1189 of the large p261 subunit was graciously donated by Dr. Zachary Pursell (Tulane University). The Pol ε catalytic domain construct was cloned in the pET60-DEST vector and has a cleavage site recognized by TEV protease to generate GST(TEV)/Polε(1 – 1189)-His$_6$ (Figure 8). The vector pRK603 encoding for TEV protease was also donated by Dr. Zachary Pursell. Tris-hydrochloric acid, Tris base, NaOH, NaCl, EDTA, acetic acid, Mg(C$_2$H$_3$O$_2$)$_2$, DTT, BL21(DE3) Rosetta competent cells, LB-Agar, rabbit polyclonal anti-hexahistidine tag antibody, ampicillin sodium, kanamycin sulfate, potassium sulfate (K$_2$SO$_4$), Luria Broth, IPTG, EDTA-free protease inhibitor tablets, glycerol, ammonium persulfate (APS), tetramethylethylenediamine (TEMED), boric acid, sodium acetate (NaC$_2$H$_3$O$_2$), and cobalt resin were all obtained from Fisher Scientific (Hampton, NH). β-Mercaptoethanol (β-ME) was obtained from MP Biomedicals in Santa Ana, CA. Imidazole was obtained from Acros Organics (Fisher Scientific Hampton, NH). Sequel Parts A and

B were obtained from American Bio (Canton, MA). Deoxy nucleotide triphosphates (dNTPs) were obtained from Thermofisher Scientific, and HPLC purified and 5'-TET-(tetrachlorofluorescein) labeled DNA substrates were purchased from Integrated DNA Technologies. This fluorophore was chosen because it had been previously used in pre-steady kinetic measurements with yeast Pol ε and was compatible with our Typhoon FLA 9500 Imager (Figure 9).[39]
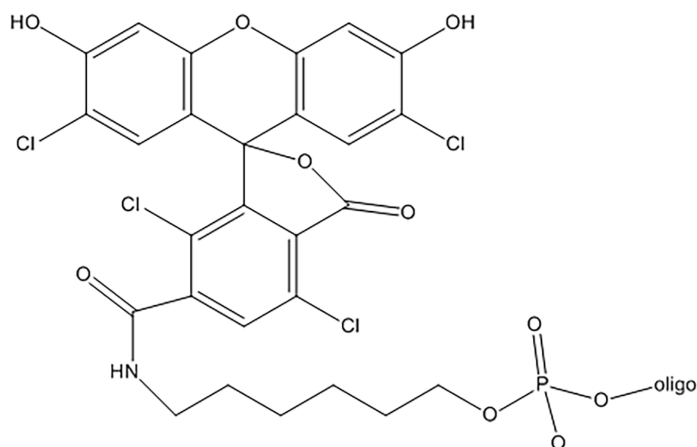
3.5.2 Plasmid Mutagenesis

Figure 9: Structure of TET, the fluorescent label.

An open reading frame encoding residues 1 – 1189 of the p261 subunit of human Pol ε was cloned into pET60-DEST. This construct contains a TEV site between the *N*-terminal GST tag and Pol ε, and a lysine linker between the *C*-terminal His$_6$ tag and Pol ε to generate GST(TEV)/Polε(1 – 1189)-His$_6$. To inactivate the proofreading activity of Pol ε, two highly conserved residues D275 and D277 were changed to alanine and previously shown to have negligible exonuclease activity (Pol ε exo-).[38] To generate the different polymerase domain mutants, directed mutagenesis (Kapa Bioscience,

Wilmington, MA) was used to generate F815L (forward primer 5' – CGCACAAATGTATCCTGAACAGTCTCTACGGCTACG – 3', reverse primer 5' – CGTAGCCGTAGAGACTGTTCAGGATACATTTGTGCG – 3'). Mutations were verified via sequencing to confirm accuracy. Notably a residual cloning artifact remained from previous cloning work. The vector contained an extra guanine-guanine-cytosine triplet (GGC) that codes for a glycine at the *N*-terminus. This cloning vestige is a common artifact from using a restriction enzyme like *Nde1* to create the start codon. When the extra GGC was deleted from the sequence we could no longer purify active Pol ε exo-. When the GGC was retained in the DNA sequence, the activity of Pol ε exo- matched the literature.[2, 35-39] I designed all primers and performed all the mutagenesis.

3.5.3 Expression and Purification of Human DNA Polymerase Epsilon

E. coli expression strain BL21(DE3) Rosetta were co-transformed with pET60-DEST encoding for Pol ε and vector pRK603 encoding for TEV protease. The transformed cells were incubated at 37 °C in 6 L of LB media in the presence of 50 μg/mL ampicillin and 25 μg/mL of kanamycin with shaking until the optical density at 600 nm (OD_{600}) of the cultures reached 0.6-0.7. Protein expression was induced by the addition of IPTG to a final concentration of 4 μM, and the temperature was lowered 20 °C for 15 hours. Cell cultures were pelleted by centrifugation at 8,000 rpm for 10 minutes at 4 °C. The pellets were resuspended in lysis buffer containing 300 mM Tris pH 7.8 at 4 °C, 100 mM NaCl, 0.5 mM EDTA, 20 mM K_2SO_4, 10% glycerol, and supplemented with EDTA-free protease inhibitor tablet. Cells were lysed by passing the resuspension through a microfluidizer at 16,000 psi. The cell lysate was clarified by centrifugation at 18,000 rpm for 45 minutes at 4 °C. Clarified cell lysates were loaded on a cobalt column and washed

with wash buffer containing 150 mM Tris pH 7.8 at 4 ˚C, 200 mM NaCl, 20 mM K_2SO_4, 5

mM β-mercaptoethanol, and 10% glycerol. Bound protein was eluted from the column

with elution buffer containing 150 mM Tris pH 7.8 at 4˚C, 200 mM NaCl, 5 mM β-

mercaptoethanol, 10% glycerol, and 300 mM imidazole. The cobalt elution was then

passed through a Superdex 16/600 size exclusion column in size exclusion buffer
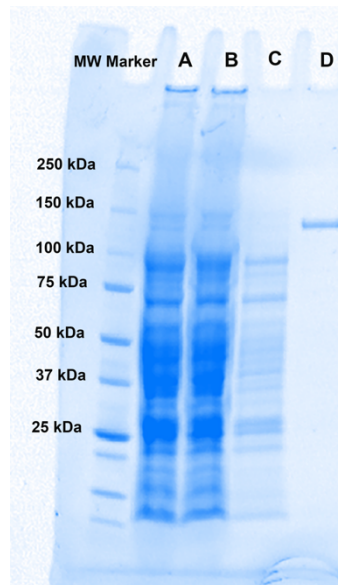

Figure 10: SDS-PAGE analysis to evaluate the purity of WT h Pol ε. The final protein sample was run on a 4-20% SDS-PAGE gel. Molecular weight (MW) marker is shown in the left most lane. Lane A is crude WT h Pol ε bacterial cell lysate, Lane B is the flowthrough from the cobalt column, Lane C is the wash from the cobalt column, and lane D is the purified protein.

containing 150 mM Tris pH 7.8 at 4˚C, 200 mM NaCl, 1 mM DTT, and 5% glycerol.

Fractions were collected and protein absorbance was monitored at 280 nm. Fractions

containing Pol ε were then pooled and dialyzed against storage buffer containing 50 mM

Tris pH 7.8 at 4 ˚C, 15% glycerol, 1 mM DTT, and 300 mM $NaC_2H_3O_2$. Pol ε was

determined to be ~95% pure by SDS-PAGE analysis (Figure 10). The concentration of

129

purified protein was measured by UV spectrophotometry at 280 nm using a calculated molar extinction coefficient at 155,510 $M^{-1}cm^{-1}$. Anna Uvarova and I optimized the purification of WT Pol ε. I performed the purification of both wild-type WT exonuclease deficient Pol ε and exonuclease deficient F185L Pol ε.

3.5.4 Reaction Buffers

All assays using Pol ε exo- and F815L Pol ε exo- were performed at 20 °C using reaction buffer (50 mM Tris Acetate pH 7.4 at 20°C, 0.1 mg/mL BSA, 10% glycerol, and 1 mM DTT). Reactions were carried out on a rapid chemical quench flow apparatus (KinTek). Notably, 37 °C was not used because the rate constant of product formation catalyzed by Pol ε is reportedly too fast at this temperature to be accurately determined using a rapid chemical quench apparatus (k_{pol} > 500 s^{-1}).[2, 35-39] Final concentrations are reported after 1:1 mixing. I tested various protein concentrations, DNA concentrations, reaction temperatures, reaction timepoints, polyacrylamide and crosslinker percentages in denaturing gels, running buffers used in gel electrophoresis, and the effect of additives such as sodium acetate and glycerol in purification and storage buffers in order to optimize protein solubility, stability, activity, and product analysis.

3.5.5 Primer Template Annealing

HPLC-purified DNA substrates (Table 1) were purchased from Integrated DNA Technologies. The 21-mer primer strand was purchased with a 5'-TET fluorescent label (D-2). The 5'-fluorescent primer was annealed to 41-mer template by incubating the 21-primer with 1.5-fold molar excess template at 95 °C for 5 minutes before cooling slowly to room temperature. For these studies, the DNA substrate D-2 (Table 1) was used. I performed the primer-template annealing experiments.

3.5.6 Pre-Steady State Burst Assays

In the burst assays, a pre-incubated solution of 100 nM Pol ε and 400 nM 5'-TET DNA in reaction buffer, described above, was rapidly mixed with varying concentrations of dNTP and 8 mM $Mg(C_2H_3O_2)_2$. All reactions were quenched with the addition of 0.37 M EDTA. All reactions were performed using a rapid chemical quench-flow apparatus (KinTek). I performed the burst assays with WT exonuclease deficient Pol ε with dTTP and dATP and I performed the burst assays with F815L exonuclease deficient Pol ε with dTTP.

3.5.7 Single Turnover Assays

The single turnover assays, a pre-incubated solution of WT Pol ε exo- and 5'-TET DNA substrate at various concentrations was rapidly mixed with 8 mM $Mg(CH_3COO)_2$ and varying concentration of dTTP in reaction buffer. In all cases, enzyme concentration was in excess to DNA substrate concentration. All reactions were quenched with the addition of 0.37 M EDTA. All reactions were performed using a rapid chemical quench-flow apparatus.

3.5.8 Processivity Assays

This assay was attempted three times. In the first attempt, a preincubated solution of 500 nM enzyme (WT Pol ε or F815L Pol ε mutant) and 50 nM DNA was rapidly mixed with 8 mM magnesium acetate and 200 µM of each dNTP (final concentrations reported) for various time points. In the second attempt, a preincubated solution of 800 nM Pol ε WT and 400 nM DNA was rapidly mixed with 8 mM magnesium acetate and 200 µM of each dNTP for various time points (final concentrations reported). In each case, the reactions were quenched by the addition of 0.37 M EDTA.

3.5.9 Product Analysis

10 μL of Formi dye containing 10 mL formamide, 10 mg xylene cyanol, and 10 mg bromophenol blue was added to the quenched incubations from the burst assays.   A total of 6 μL of this ~150 reaction was loaded on a 15% denaturing polyacrylamide gel containing 8 M urea and 1X Tris/borate/EDTA in lithium borate running buffer. The 2nd processivity assay was run in the dye described above without xylene cyanol and loaded on a 10% denaturing polyacrylamide gel containing 8 M urea and 1X Tris/borate/EDTA in lithium borate running buffer.   The gel was scanned with a Typhoon FLA 9500 (GE

WT h Pol ε exo- Burst Assay 200 μM dTTP
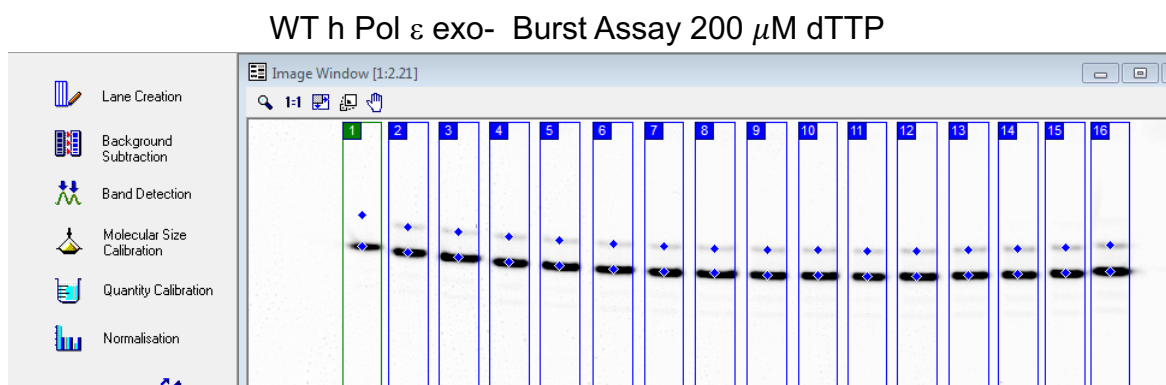
Figure 11: Annotated analysis of time-dependent formation of product, using a burst assay as an example. A preincubated solution of 100 nM WT h Pol ε and 400 nM D-2 DNA substrate was rapidly mixed with 8 mM Mg(CH$_3$COO)$_2$ and 200 μM dTTP and quenched with 0.37 M EDTA after various time intervals. One nucleotide extended product was separated from unreacted substrate by gel electrophoresis using a 15% acrylamide gel with 8 M urea with lithium borate running buffer. Time-dependent amount of product formation was calculated using ImageQuant by determining the total pixel intensity of each lane, then dividing the intensity of the product band by the total pixel intensity to calculate the percent of extended product formed at that time. The amount of product formed at the zero-time point was subtracted from the others, then multiplied by the total concentration of DNA substrate used in the assay (400 nM) to calculate the concentration of product formed at that time. The time-dependent formation of product was then fit to the burst equation (Eq 1) to determine kinetic constants k_{obs} and k_{ss}.

Healthcare) at the Alexa 532 nm setting to excite the TET fluorophore that was covalently bound to the 5'-end of the primer DNA. The band intensities were quantified with ImageQuant software (GE Healthcare) and relative intensities were calculated by dividing the intensity of a specific band by the total intensity of the observed bands (Figure 10). I optimized product analysis.

3.5.10 Data Analysis

All pre-steady-state kinetic data under burst conditions, DNA substrate in excess to enzyme, were fit by nonlinear regression using GraphPad Prism (GraphPad Software, La Jolla, California). Data for the burst assays were fit to Eq. 1

$$[\text{product}]=A[(1-e^{-k_{obs}t})+k_{ss}t] \quad (1)$$

Where A is amplitude of active enzyme, k_{obs} is the observed burst rate constant, and k_{ss} is the observed steady-state rate constant. When k_{ss} is divided by the A, it is converted a first-order rate constant that defines the rate of product release. The k_{obs} values were plotted against dNTP concentrations and the data were fit the following hyperbolic equation (Eq 2) using GraphPad Prism:

$$k_{obs}=\frac{k_{pol}*[\text{dNTP}]}{K_D^{dNTP}+[\text{dNTP}]} \quad (2)$$

where k_{pol} is the maximum rate constant of dNTP incorporation and K_d^{dNTP} is the equilibrium dissociation constant of the dNTP. Single turnover assays were fit to single exponential equation (Eq 3) using Graphpad Prism:

$$[\text{product}]=A[(1-e^{-k_{obs}t}) \quad (3)$$

The fidelity (*F*) of WT Pol ε was calculated as a ratio of catalytic efficiencies incorrect dNTP incorporation to correct dNTP incorporation using the following relationship:

$$F = \frac{(k_{pol}/K_d)_{incorrect}}{(k_{pol}/K_d)_{incorrect} + (k_{pol}/K_d)_{correct}} \qquad (4)$$

I performed data analysis of WT exonuclease deficient Pol ε incorporating the correct dTTP and incorrect dATP, exonuclease deficient F815L Pol ε with the correct dTTP, and processivity assays.

3.6 Results

3.6.1 Expression and Purification of Human WT Pol ε

The purification of the endogenous DNA Pol ε complete with 4 additional subunits from both yeast and human cells is very labor intensive and not practical for rapidly screening tumor-relevant Pol ε mutants.[18-20, 38] Various groups have circumvented this difficulty by expressing the *N*-terminal truncated fragment of the p261 catalytic subunit.[2, 35, 36, 38, 39] Importantly, this 140 kDa fragment contains all conserved polymerase and exonuclease motifs and has previously been shown to be as active as the full-length catalytic subunit and essentially retains the same processivity.[36, 38, 39] This fragment also possesses active exonuclease activity.[36, 38] The convention in nucleotide incorporation analysis is to ablate exonuclease activity so rates of correct and incorrect dNTP insertion can be measured without complication from nucleotide incision. Here, an exonuclease deficient mutant of the 140 kDa fragment was generated by substituting the two highly conserved catalytic residues D275 and D277 with alanine.[35, 38] Substitution of these

residues has been shown to generate a successful exonuclease deficient mutant.[35] In addition, the rate of incorrect nucleotide incorporation would be too slow to be observed with an active exonuclease domain. This is because the rate of incorrect excision is faster than the rate of incorrect nucleotide incorporation.[2] Therefore, the exonuclease deficient variant of Pol ε was generated in order to perform experiments in absence of proofreading activity.

Thus, we undertook a similar approach as previous groups to express and purify a fragment corresponding to the *N*-terminal 140 kDa catalytic subunit of Pol ε.[2, 35, 36, 38, 39] We co-expressed TEV protease with Pol ε during longer induction times to endogenously cleave the GST tag. We recovered protein suitable for pre-steady-state kinetic investigation through our purification process (Figure 9). These results are specific to the second bacterial construct donated from Zach Pursell (Figure 8). The full-length construct (Figure 4, Figure 5) and first bacterial expression construct (Figure 6, Figure 7) were unsuitable for kinetic analysis because yields were too low and purified enzyme was catalytically dead, respectively.

3.6.2 WT Pol ε exo- and Mutant Exhibit Biphasic Kinetics

Due to their processivity, replicative polymerases typically exhibit biphasic kinetics during single nucleotide incorporation.[35] Previous work has demonstrated that a single nucleotide incorporation event is preceded by a conformational change following nucleotide binding.[35] DNA dissociation from the enzyme-DNA binary complex limits additional incorporations (Figure 2).[35] Two phases, an initial burst of product formation and a slower linear phase, can be observed by performing a burst assay. In a burst assay, enzyme is pre-incubated with an excess of DNA substrate to form the stable enzyme-

DNA complex. We used a rapid chemical quench to examine the incorporation of a correct dTTP in our D-2 DNA substrate (Table 1) under pre-steady state conditions. The time course for product formation revealed biphasic kinetics characterized by an initial burst of product formation that is followed by a slower linear phase (Figure 11). The burst rate at 10 μM concentration of dTTP was determined to be $7 \times 10^1 \pm 1 \times 10^1$ s^{-1} and the
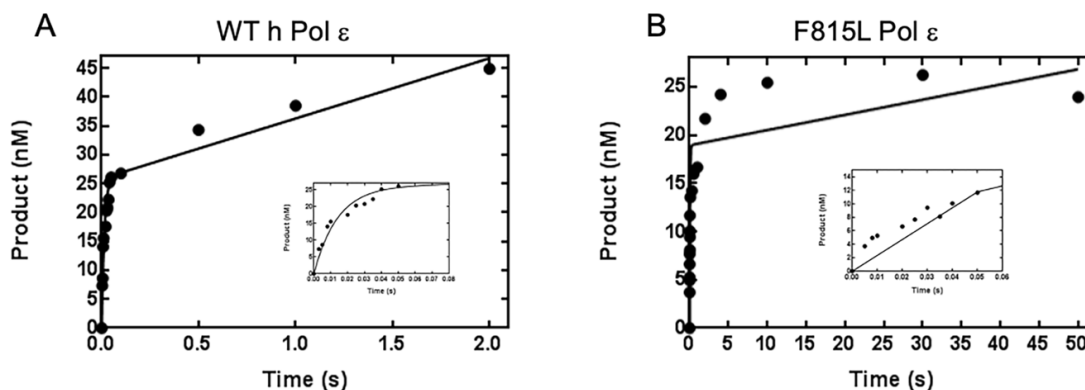

Figure 12: The burst kinetics of correct dTTP incorporation by WT human Pol ε exo- (A) and F815L Pol ε exo- (B) were measured by mixing a preincubated solution of either enzyme (100 nM) and 5'-TET D-2 DNA substrate (400 nM) with 200 μM dTTP and 8 mM Mg(CH$_3$COO)$_2$ and quenching with the addition 0.37 M EDTA at various time intervals. The F815L mutant was mixed with 1 mM dTTPThe data were fit to the burst equation (Eq 1) to yield a fast exponential rate constant of 7×10^1 $\pm 1 \times 10^1$ s^{-1} for WT h Pol ε and 20 ± 4 s^{-1} for F815L Pol ε. The linear steady state rate was found to be $2 \times 10^{-2} \pm 2 \times 10^{-3}$ s^{-1} for WT human Pol ε and $1 \times 10^{-2} \pm 4 \times 10^{-2}$ s^{-1} for F815L Pol ε exo-. The inset shows the exponential phase of each time course.

linear rate constant was $2.0 \times 10^{-2} \pm 2 \times 10^{-3}$ s^{-1} for WT exonuclease deficient Pol ε. For F815L exonuclease deficient Pol ε, the burst rate at 1 mM concentration of dTTP was determined to be 20 ± 4 s^{-1} and the linear rate constant was $1 \times 10^{-2} \pm 4 \times 10^{-2}$ s^{-1}. The fast, pre-steady state burst phase corresponds to the incorporation of a single nucleotide during the first enzyme turnover. The steady-state, linear phase corresponds to the incorporation of a nucleotide during subsequent enzyme turnovers. The observation of a burst shows that Pol ε likely follows the same kinetic mechanism as other characterized polymerases (Figure 2).[35] The observed rates obtained in the burst assay closely

matched the literature values.[2, 35-37, 39]   The presence of a burst phase provided an opportunity to examine nucleotide incorporation in the first enzyme turnover to assess the contributions of the nucleotide binding and incorporation steps toward the efficiency of Pol ε and mutant.

3.6.3 Single Turnover Assays

It is common to obtain dNTP dissociation constants and maximum incorporation rates using single-turnover assays.   These experiments are performed with excess enzyme to substrate to allow the direct observation of the nearly complete conversion of substrates to products. In a single turnover experiment, enzyme is in excess to substrate and only one turnover can occur. Unlike a pre-steady-state burst assay, steady-state formation of products does not occur. Recall, that the linear phase of the burst assay
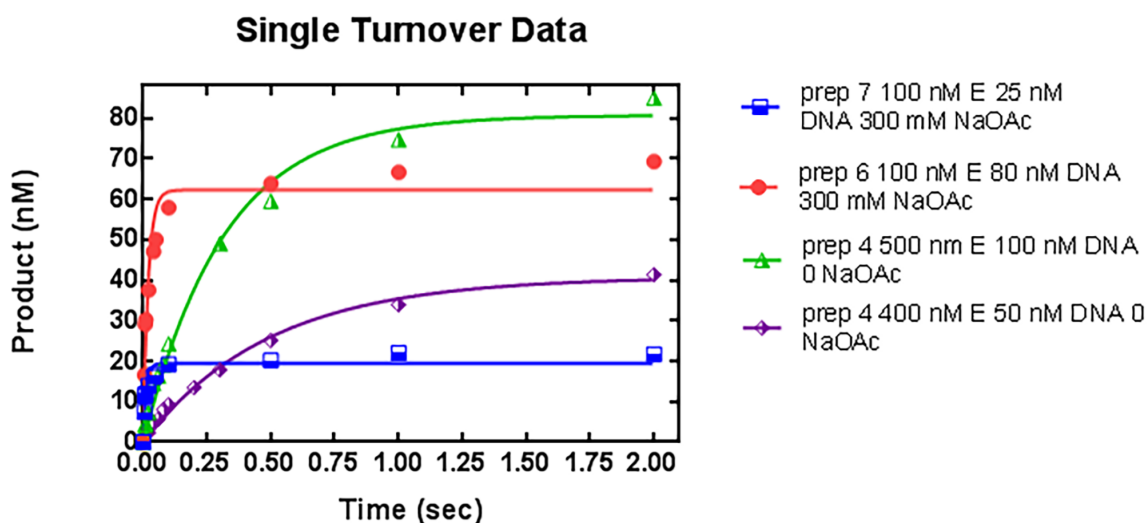

Figure 13: Select single turnover data. Active site concentrations (determined from the amplitude of the burst curve) were used to determine enzyme concentration in single turnover assays. Various concentrations of enzyme and DNA were rapidly mixed with 8 mM $Mg(CH_3COO)_2$ and a saturating concentration of dTTP. Reactions products were analyzed on via denaturing polyacrylamide gel electrophoresis and quantified as previously described.   Time-dependent amount of product concentration was fit to a single exponential equation (Eq 3) to determine observed rates of single nucleotide incorporation, k_{obs}.   Rates with enzyme preps that did not include $NaCH_3COO$ in the storage buffer were ~100-fold slower than observed rates obtained from the exponential phase of the burst equation. Rates did increase upon the addition of $NaCH_3COO$ to the storage buffer, and when the DNA concentration was about 4-fold lower than the K_d for DNA.[35]

represents the steady-state rate and multiple turnovers, and the initial burst phase represents the rate of a single turnover. Theoretically, the observed rates of single turnover experiment and the observed rate in the burst phase of a burst assay should be identical.   However, when we performed single turnover assays at saturating concentrations of dTTP our observed rates did not match those obtained in burst assays (Figure 12).

Table 2: Select single turnover assay rate constants

Single Turnover Assay conditions	k_{obs} (s^{-1})
100 nM WT Pol, 25 nM D-2, 300 uM dTTP, 8 mM MgOAc, 300 mM NaOAc	9.0×10^2 $\pm 2 \times 10^1$
100 nM WT Pol, 80 nM D-2, 200 uM dTTP, 8 mM Mg OAc, 300 mM NaOAc	5.0×10^2 ± 9
500 nM WT, Pol 100 nM D-2, 200 uM dTTP, 8 mM MgOAc	3.0×10^0 $\pm 3 \times 10^{-1}$
400 nM WT, Pol 50 nM D-2, 200 uM dTTP, 8 mM MgOAc	2.0×10^0 $\pm 1 \times 10^{-1}$

Initial rates were obtained to be~2 s^{-1} at saturating dNTP concentration. This is ~140-fold lower than the observed rate in the burst experiments. Altering the amount of excess enzyme in our single turnover assays also did not increase the observed rate. Only when sodium acetate was added to the storage buffer did the observed rates in the single turnover assays increase to ~90 s^{-1}. This still represents a 3-fold decrease from the observed rates in the burst assays at saturating conditions, but a 45-fold improvement over previous single turnover assays (Table 2). The observed rates of the single turnover assays also did not match the literature values.[2, 35-37, 39] I have not elucidated the reason behind the discrepancy in rates, but some ideas are mentioned in the discussion.

3.6.4 Determination of K_d^{dTTP} and k_{pol} for Nucleotide Incorporation by WT Pol ε exo- and F815L mutant

To examine rates of nucleotide incorporation, I started by measuring the rate of correct dTTP and incorrect dATP incorporation opposite dAMP in the DNA template. In these experiments, I measured the rate of product formation over time at varying dNTP concentrations under burst conditions (DNA substrate in excess to enzyme). The product is the primer/template DNA substrate extended by one nucleotide. The observed rates of the pre-steady-state burst phase were then plotted as a function of dNTP concentration and fit to a hyperbolic equation (Eq 2). The maximal rate of single nucleotide incorporation, k_{pol}, is a rate constant made up of dNTP binding, a conformational change, and polymerization chemistry . Due to rapid equilibrium between the enzyme-DNA binary complex and dNTP binding, the dNTP concentration at half k_{pol} is an apparent dissociation constant (K_d^{dNTP}). The efficiency of nucleotide incorporation can be measured by calculating k_{pol}/K_d. This parameter can be used to assess fidelity by the following relationship: $(k_{pol}/K_d)_{incorrect}/[(k_{pol}/K_d)_{correct} + (k_{pol}/K_d)_{incorrect}]$ (equation 4 in Methods section). The data for the correct dTTP incorporation are shown in Figure 13 and yielded a k_{pol} of 288 ± 8 s^{-1} for WT Pol ε exo- and a K_d^{dTTP} of 36 ± 3 μM. For the F815L exo- mutant, the k_{pol} was 21 ± 1 s^{-1} and the K_d^{dTTP} was 96 ± 30 μM (Figure 13, Table 3). Thus, the F815L Pol ε exo- mutant had a ~14-fold decrease in k_{pol}, a 3-fold increase in K_d^{dTTP}, and a ~37-fold decrease in enzymatic efficiency for correct dTTP incorporation opposite dAMP.

The rate of misincorporation of a dATP across a template dAMP for Pol ε was examined next. The experiments were performed under burst conditions (enzyme concentration limiting to DNA concentration) and fit to a burst equation (Eq 1) (Figure 13
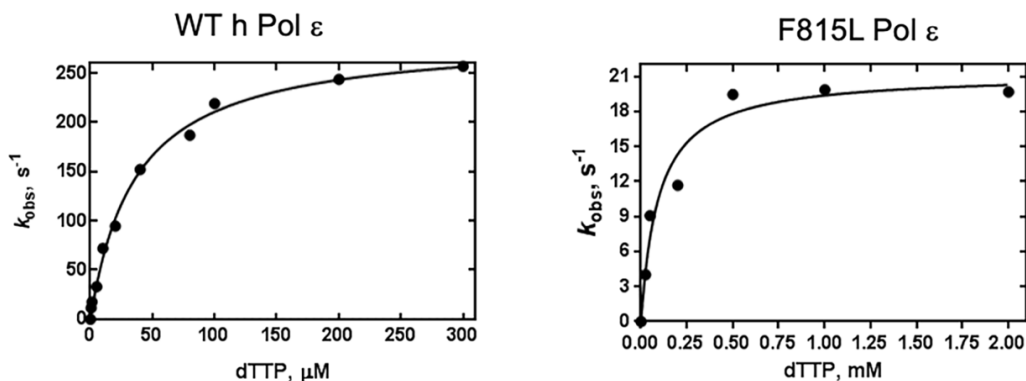
WT h Pol ε

F815L Pol ε

Figure 14: Kinetics of correct dTTP incorporation. A preincubated solution of WT h Pol ε exo- (left) or F815L Pol ε exo- (right) (100 nM) and 5'TET labeled D-2 DNA substrate was rapidly mixed with varying concentrations of dTTP and 8 mM Mg(CH$_3$COO)$_2$ for varying time intervals before quenching the reaction with 0.37 M EDTA. Product concentration was plotted against time and data were fit to the burst equation (Eq 1). The observed exponential rate constants (k_{obs} in Eq 1) at each dTTP concentration were plotted against their respective dTTP concentration. The plot of k_{obs} vs [dTTP] was fit to the hyperbolic equation (Eq 2) to yield a k_{pol} of 288 ± 8 s^{-1} and a K_d^{dTTP} of 0.036 ± 0.003 mM for WT h Pol ε exo- and a k_{pol} of 21 ± 1 s^{-1} and a K_d^{dTTP} of 0.10 ± 0.03 mM for F815L Pol ε exo-.
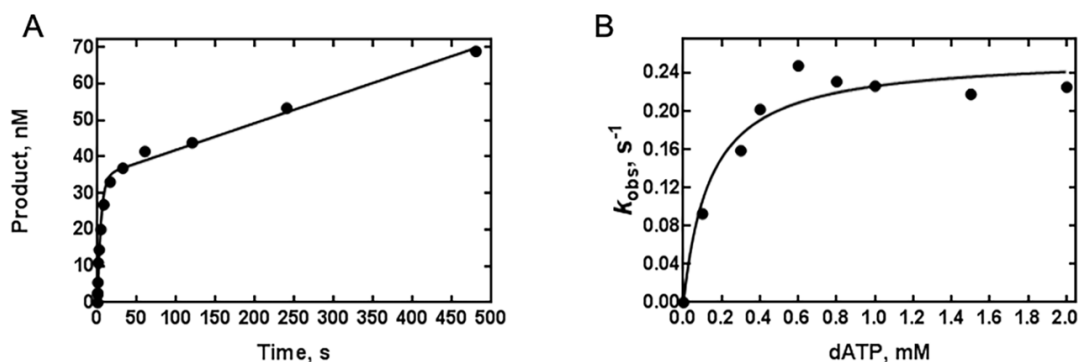

A

B

Figure 15: Nucleotide concentration dependence on the pre-steady-state kinetic parameters of incorrect dATP incorporation opposite a template deoxyadenosine catalyzed by WT h Pol ε exo-. **A.** An example of the burst kinetics of dATP misincorporation. A preincubated solution of100 nM WT h Pol ε exo- and 400 nM 5'TET labeled D-2 substrate was mixed with 800 µM of dATP(shown) and 8 mM Mg(CH$_3$COO)$_2$ for various time intervals before quenching with 0.37 M EDTA. The data of product formation versus time at each dATP concentration were fit to the burst equation (Eq 1) to yield k_{obs}. **B.** The resulting k_{obs} values were plotted versus dATP concentration and fit a hyperbolic equation (Eq 2) to yield a k_{pol} of 0.26 ± 0.02 s^{-1} and a K_d of 0.142 ± 0.044 mM.

A). The kinetics of the formation of a A:A mismatch by WT Pol ε are shown at various

nucleotide concentrations in (Figure 13B). The observed rates were plotted as a function

of the dATP concentration to fit a hyperbola to yield a K_d of 1.4 x10^2 ± 4 x 10^1 µM and a

140

k_{pol} of $2.6 \times 10^{-1} \pm 2 \times 10^{-1}$ s^{-1}. More points should be obtained for the F815L mutant to

reduce the error in the K_d value. The fidelity of WT Pol ε exo- base substitution was then

calculated as $(k_p/K_d)_{incorrect}/[(k_{pol}/K_d)_{correct}$ + $(k_{pol}/K_d)_{incorrect})$ to be 2.4 × 10^{-4}.

Misincorporation assays for the F815L mutant will be performed in the future. The results

of the kinetic investigation is shown in Table 3.

Table 3: Kinetic Parameters of Pol ε and F815L mutant.

Pol ε exo-	dNTP	k_{pol} (s^{-1})	K_d (mM)	k_{pol}/K_d (mM^{-1}s^{-1})	$F_{pol}{}^1$
WT	dTTP	288 ± 8	$3.6 \times 10^{-2} \pm 3 \times 10^{-3}$	$8.0 \times 10^3 \pm 7 \times 10^2$	
WT	dATP	0.26 ± 0.02	$1.4 \times 10^{-1} \pm 4 \times 10^{-2}$	$2 \pm 6 \times 10^{-1}$	2.4×10^{-4}
F815L	dTTP	21 ± 1	$1 \times 10^{-1} \pm 3 \times 10^{-2}$	$2.2 \times 10^2 \pm 7 \times 10^1$	

3.6.5 Processivity Assay

The small DNA binding subunits of the Pol ε tetramer have been shown to

positively affect processivity in both the human and yeast homologs.[36, 39]   To see if

processivity of Pol ε was affected by the presence of the F815L mutation, I compared the

processivities of WT Pol ε and F815L Pol ε on the DNA template.   Three different

conditions were studied. In the first, a pre-incubated solution of WT Pol ε or F815L Pol ε

(500 nM) and labeled 5'-TET labeled DNA primer/template (50 nM) was mixed with all

four dNTPs (200 uM) and Mg(CH$_3$COO)$_2$ for various time points. In the second, 800 nM

Pol ε WT and 400 nM DNA was rapidly mixed with all four dNTPs and Mg^{2+} for various

time points. In both cases, the products were separated by denaturing PAGE. In the first

case, WT Pol ε exo- and the F815L Pol ε exo- mutant was found to not extend past 2

nucleotides and did not restart extension even at long timepoints (Figure 15). This was

thought to be due to the high molar excess of enzyme to DNA and the low DNA concentration. At 10-fold excess enzyme to DNA and 50 nM DNA, substrate may be consumed too fast and not produce detectable amounts of fully extended product. In addition, the xylene cyanol dye may mask fully extended product since it migrates to a similar location as the product. Therefore, I ran the processivity assay again without
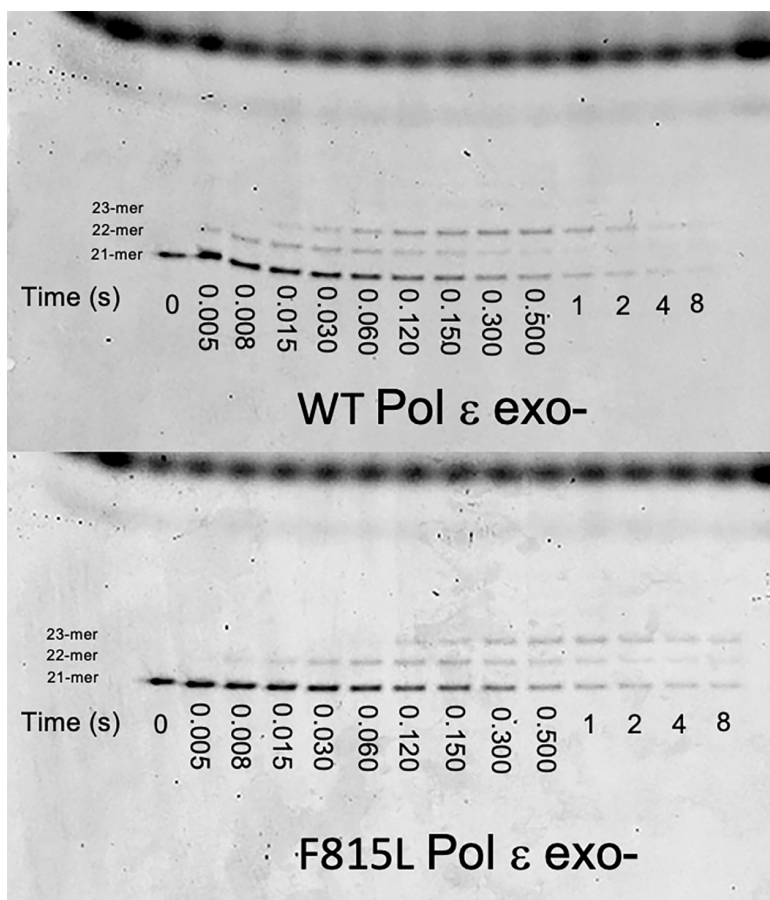

Figure 16: First processivity attempt with WT human Pol ε exo- (top) and F815L Pol ε exo- (bottom). A preincubated solution of 500 nM of either enzyme with 50 nM DNA was mixed with 8 mM Mg(CH$_3$COO)$_2$ and 200 μM of each dNTP. Both enzymes seemed to be non-processive and stopped incorporation chemistry after 2 insertions.

xylene cyanol in the loading dye, used 2-fold excess enzyme to DNA, analyzed the products on a 10% acrylamide gel instead of a 15% acrylamide gel, and included longer reaction times in order to visualize any potential fully extended product. When these

142

changes were made, the resolution was improved and a high molecular weight band was

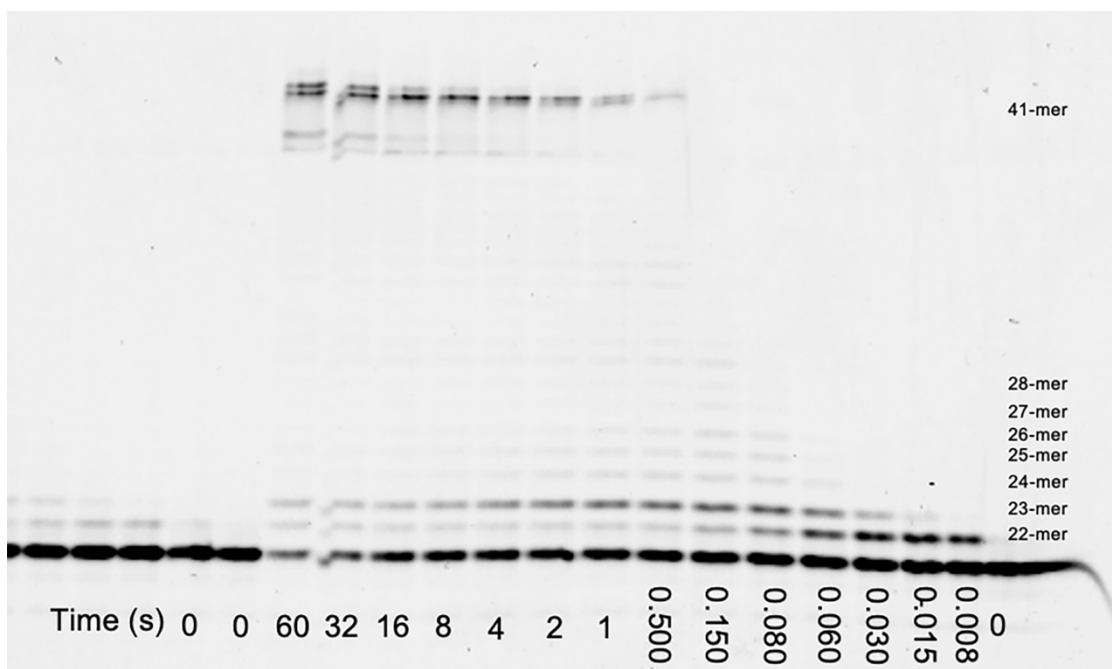visualized, presumed to be a fully extended DNA substrate (Figure 16).

Figure 17: Processivity assay with WT human Pol ε exo-. A preincubated solution of 800 nM WT h Pol ε exo- and 400 nM DNA was mixed with 8 mM Mg(CH₃COO)₂ and 200 µM of each dNTP and quenched with 0.37 M EDTA for the corresponding time interval. WT h Pol ε exo- was shown to form fully extended primer in between 150 and 500 ms.

3.7 Discussion

In this work, we carried out pre-steady-state kinetics to partially characterize the

mechanisms of fidelity between WT Pol ε and F815L mutant Pol ε, observed in

endometrial cancer. Using the truncated fragment of the *N*-terminal domain, the burst

assay revealed that both enzymes exhibited biphasic kinetics that were characterized by

a "burst" phase followed a slow linear phase. The initial "burst" phase represents the rate

of nucleotide incorporation and the linear rate represents the steady-state phase of

product formation during multiple turnovers. The F815L mutant was found to insert a

correct nucleotide ~13 fold slower than WT Pol ε, and exhibited a ~3-fold higher K_d^{dTPP}

of ~100 uM, resulting in in a 40-fold decrease in catalytic efficiency. The difference in

enzymatic efficiency is best determined through a structural analysis combined with pre-

steady-state kinetics. However, no structures of human Pol ε exist. Therefore, energy

minimized structures of the F815L mutant were created using the yeast homolog. This

mutation is 10 Å away from the enzyme active site and likely doesn't play a role in directly

selecting the incoming dNTP; however, it may perturb the geometry of the active site

(Figure 3). Interestingly, aligning the active site of the energy minimized F815L mutant
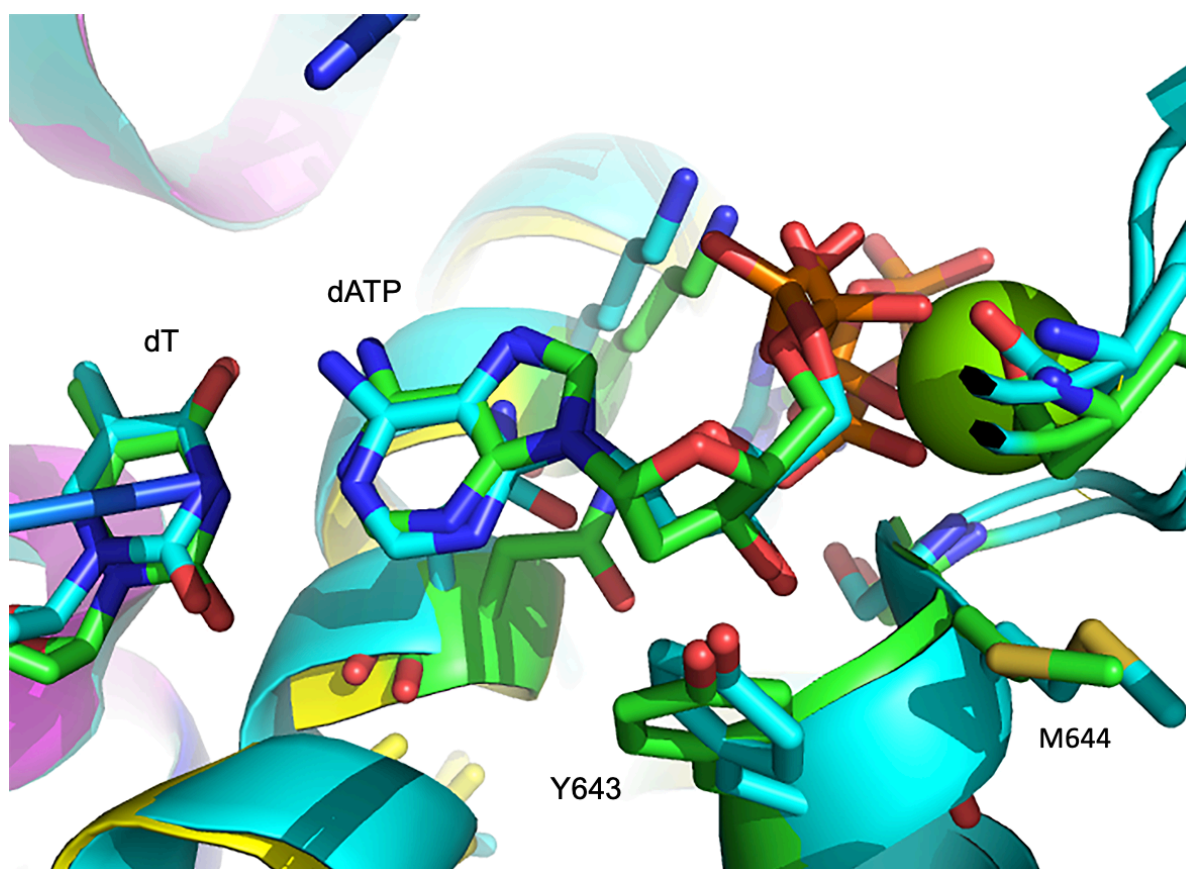

Figure 18: Structural alignment of WT yeast Pol ε (colored by element) and energy minimized F815L (cyan) structure (4M8O).[21] This highlights the dNTP binding pocket. Y643 and M644 are shown to have a modified arrangement in the mutant that leads to increased conformational freedom of the dNTP and templating base.

Pol ε (human numbering) structure with that of WT yeast Pol ε shows that Y643 and M644

are shifted from their normal position (Figure 17).

144

I hypothesize that this mutant increases the degrees of freedom that the template base can assume, leading to stabilization of incorrect dNTP binding compared to WT Pol ε, as previously shown with RB69 polymerase.[3] RB69 and Pol ε are both replicative polymerases, though more is known about RB69 due to ease of expression.[3] Polymerase domain mutations have been generated in RB69 and their kinetic and structural effects have been elucidated. Mutating the corresponding tyrosine to an alanine at the RB69 active site creates a large cavity below the incoming dNTP and templating base.[3] This cavity provides the templating base and dNTP with more degrees of freedom so that it can shift to pair with mismatched dNTPs, leading to a decreased incorporation rate and decreased replicative fidelity.[3] A similar effect is observed in T7 DNA Polymerase, a family A replicative polymerase.[3] Shifting of the templating base is a likely reason why the incorporation rate and fidelity of repair polymerases is so low compared to those of replicative polymerases.[3] Moreover, M644 has been previously shown to important for fidelity in yeast Pol ε.[38] Mutation to a glycine residue led to a decrease in fidelity due the increased space in the hydrophobic pocket as a result of the mutation.[38] The combined effects of the Y643 and M644 shift lead to an increase in conformational flexibility between the incoming dNTP and the template base that could stabilize the formation of an incorrect base pair. Additionally, the F815L mutation is not as frequently observed in cancer as the P286R mutation. Though highly speculative, the tumor may select for P286R mutation over the F815L because it affects both exonuclease and polymerase activity at the same time by preventing proper exonuclease function and promoting promiscuous polymerase activity.[26, 27, 41] The structure of the P286R mutant in yeast Pol ε showed that the arginine substitution prevents proper DNA binding to the exonuclease active site and

proper orientation of the metals necessary for catalysis.[41] Notably, when the P286R

mutation was studied in yeast Pol ε (P301R) it was found to have an increased ability to

extend past a mismatch and bypass unusual secondary structures in DNA, such as

hairpins that frequently occur when DNA unwinds during replication.[42] Furthermore, it

was more capable of fully extending a DNA primer/template with a combination of a

hairpin and a mismatch when compared to WT yeast Pol ε.[42] When WT yeast Pol ε

encountered these lesions, DNA replication would normally pause and resume after the

lesion was repaired properly.[42] The authors concluded that when P286R Pol ε encounters

a mismatch or an unusual secondary structure, DNA replication continues with the lesion

left unrepaired.[42] This can lead to a higher mutation rate, genome instability, and an

increased risk of developing cancer. Similarly, the F815L mutant may increase the ability

of Pol ε to accommodate these unusual DNA secondary structures at its active site.

Comparatively, the F815L mutant would still have a fully active exonuclease active site

and would theoretically be able to be able to excise an incorrectly placed nucleotide. If

exonuclease activity is affected, the mutation rate of the F815L may still not match that of

P286R, or other documented exonuclease domain mutants.[26, 27] A more likely reason is

that the F815L mutant impairs the enzymatic efficiency of the incorporation reaction. The

maximum incorporation rate for the F815L mutant was found to be ~10-fold slower than

the WT enzyme and the mutant possessed a 3-fold higher dissociation constant for dTTP.

At physiological conditions, this F815L mutant would be a poor choice for a rapidly

dividing cancer cell and would favor the selection of an exonuclease domain mutant.

Mechanistically, the presence of a burst suggests that product release or a step

after chemistry is rate limiting. When I performed a burst assay with Pol ε, I observed the

146

presence of a burst. This indicates that Pol ε catalyzes correct nucleotide incorporation by following a similar mechanism that has been established for numerous polymerases and reverse transcriptases.[35] Interestingly, the rate of single nucleotide incorporation, k_{pol}, in single turnover assays performed with WT Pol ε did not match the same parameter measured in burst assays at saturating dNTP concentrations (Figure 12, Table 2). The first possible rationale was that the amount of active enzyme in a single turnover was incorrectly estimated from the amplitude of the burst curve. The active site concentration of WT Pol ε in theory should be able to be determined from the amplitude of the burst assay. This is typically the case because the rate of nucleotide incorporation is much faster than product release. In addition, the rate of forming the enzyme-DNA binary complex is faster than its breakdown. Therefore, the amplitude of the burst represents the concentration of productive enzyme bound to substrate that is active in a single turnover. The percent of active enzyme in a single turnover can then be calculated by dividing the amplitude by the total enzyme concentration used in the burst assay. However, this model is somewhat inadequate due to the assumed irreversibility of the formation of the enzyme-DNA binary complex.[43, 44] This model can fail when the dissociation constant of DNA increases due to the rate of dissociation of the binary complex increasing. The truncated catalytic subunit of WT Pol ε exo⁻ was previously determined to bind to DNA with a dissociation constant of 79 nM, somewhat higher than seen with other replicative polymerases.[35] Therefore, the internal equilibrium arising from the breakdown of the binary complex may reduce the amplitude of the burst to a value that is less than that predicted by the burst equation. The concentration of DNA used in our burst assay was 400 nM. This concentration is about 5-fold higher than the

147

dissociation constant and would still allow for tight binding of enzyme and DNA. However, two assumptions underlie a burst experiment in practice: one is that the rate of DNA binding is much faster than catalysis, and the second is that the release of product is essentially irreversible.[43-45] In practice, these assumptions do not limit experiments as long as nonlinear regression is used for curve fitting in order to provide an initial estimate of the constants.[43, 44] These estimates can then be refined by a more complete analysis based on computer simulation.[43, 44] Because the rate of DNA dissociation is much slower than nucleotide incorporation, an active site titration assay can be used to determine the active site concentration of the enzyme with more rigor, as well as the K_d for DNA.[43, 44] In this experiment, enzyme is incubated with various concentrations of DNA and reacted with a saturating amount of dNTP on a fast time scale. The time dependent amount of product formation is then fit to the burst equation. Only those enzyme molecules with DNA bound at the polymerase active site will react to form product. The amplitude of the fast polymerization reaction is now a direct measurement of the concentration of the enzyme-DNA complexes at the start of the reaction.[43-45] By fitting the variation in burst amplitude versus DNA concentration to a quadratic equation, one can establish the concentration of active enzyme sites and measure the dissociation constant for DNA binding.[42, 43] This experiment could be performed, and then the active site concentrations derived from the burst experiment and this titration experiment could be compared. Any differences between the two would suggest a potential artifact in the less-rigorous burst experiment.

Another possibility is that WT Pol ε may be forming higher-order oligomeric aggregates at higher enzyme concentrations that prevent the enzyme active sites from

being fully saturated.[37]   Notably, WT Pol ε has been previously documented to form aggregates even with size-exclusion chromatography was included during the purification process.[37]   In addition, our purification strategy lacks an ion-exchange column or heparin column that is commonly used in Pol ε studies.[2, 35-37, 39]   This additional column may remove additional impurities, enzyme-bound DNA, or aggregates from enzyme preparations. I did not perform mobility shift assays or generate SEC-MALS data to test for the presence of aggregation, but I do not believe that Pol ε is co-purifying with contaminating DNA for a few reasons. First, we use EDTA in our lysis buffer, which precludes DNA binding to enzyme by chelating magnesium ions required for DNA binding out of the active site.[21] Second, we measured the ratio of absorbance at 260 nm and 280 nm of the purified Pol ε and determined an average value of 0.67, indicating 95-100% protein is present.[46] Any leftover impurities from the enzyme preparations may drive the formation of nonproductive enzyme states that would first have to first have to converted to a productive conformation before chemistry can occur.[37] This additional conformational change step could account for the lowered catalytic activity in the single turnover assays.[2, 35-37] It is also worth noting that when sodium acetate was included in the storage buffer, the observed rates in the single turnover assays increased substantially 45-fold (Figure 12, Table 2). Perhaps the solution is as simple as adding more sodium acetate to the storage buffer for protein used in the nucleotide incorporation assays in order to prevent aggregates from forming or increase the population of WT Pol ε in a productive conformation, as has been previously used with yeast Pol ε.[39]

The fidelity of WT Pol ε when inserting an incorrect dATP over a correct dTTP was determined to be 2.4×10^{-4} (Table 3). This is consistent with the fidelity rate of $1.2 \times 10^-$

[4] reported previously.[2] Misincorporation assays for the F815L mutant, as well as the fidelities of the other possible misincorporation events, will be completed in the future. Differences in k_{pol} and K_d will be the major determinants of base substitution fidelity; however, fidelity may also be sequence-dependent, as has been shown previously.[2] Pol ε mutants in cancer are classified by an ultrahigh mutation frequency dominated by C → A transversions, as well as C → T transitions.[15, 38] The transversion represents a pyrimidine-to-purine substitution, and the transition represents a pyrimidine-to-pyrimidine substitution.   There was also a degree of sequence specificity since the C → A transversions were flanked between two T residues and the C → T transitions were flanked by T and G residues.[15, 38] These substitutions also occurred more frequently at origins of replication.[15, 38] Thus, there is evidence of an error signature unique to Pol ε and a propensity to make certain errors at specific sites on the DNA strand. It would be interesting to compare the fidelity of WT Pol ε and F815L Pol ε in this physiological context to what we have found here. At the nucleotide insertion site, our DNA substrate contains a TAG that can be readily modified to the TAT recognition site. Previous pre-steady-state kinetic investigations also show that dATP is misincorporated with the highest fidelity relative to the other dNTPs using a similar DNA substrate and with other polymerases.[2, 3] Based on the unique error signature of tumorigenic Pol ε mutants, I would hypothesize the F815L mutant would possess an increased rate of dATP insertion in pre-steady-state assays on a physiologically-relevant DNA substrate. The exonuclease domain and post-replication mismatch repair pathways are known to dramatically improve fidelity and should be incorporated into future investigations on Pol ε fidelity.[2, 47]

The processivity assays showed that WT Pol ε and the F815L mutant seemingly exhibited the same poor ability to extend past two nucleotides. This result was alarming, as Pol ε is responsible for DNA synthesis on the leading strand as opposed to the discontinuous DNA synthesis of the Okazaki fragments on the lagging strand.[13-16] In addition, human Pol ε is assumed to possess the same P-domain found in the yeast homolog that increased processivity by allowing the enzyme to encircle DNA.[21] Processivity of both truncated yeast Pol ε and human Pol ε have been previously investigated.[35, 39] Global fitting algorithms used to calculate rates of processive DNA synthesis by human WT Pol ε show that it can incorporate at least 11 nucleotides before dissociating from DNA.[35] Global fitting algorithms performed with yeast Pol ε showed that it can incorporate 35 nucleotides before dissociating from DNA.[39] Processive polymerization experiments performed with yeast WT Pol ε used physiological concentration of each dNTP, whereas, processive polymerization experiments performed with human WT Pol ε used a dNTP concentration of 100 μM for all four dNTPs.[35, 39] I performed this experiment under single-turnover conditions to assess the processivity of our enzyme. I did not observe full extension with a 10-fold excess of enzyme to DNA. I then attempted a 2-fold excess of enzyme to DNA, in addition to using 10% acrylamide in the denaturing gel and removing xylene cyanol from the loading dye. These steps were taken in an attempt to gain better resolution of the further extended products. Figure 17 shows that these changes did help achieve fully extended product. To further improve resolution and investigate any potential pause points along the DNA strand, adjusting the voltages and temperature settings while the gel is running would help eliminate the smeariness of the bands. Even if the highest molecular weight band in Figure 17 is an

151

artifact, which is unlikely as it is not seen in the 0 time point, WT human Pol ε exo- is still

inserting 13 nucleotides before dissociating from the DNA strand. This feature matches

well with the literature.[35, 39]

In summary, our kinetic analysis demonstrates that WT Pol ε activity and fidelity

provides a critical foundation for understanding Pol ε mutants like the the F815L

polymerase domain mutant, and shows that kinetic parameters are consistent among

different primer/template substrates. We met our goal to establish robust expression and

purification protocols for both human WT Pol ε and F815L Pol ε. These protocols provide

sufficient amounts of protein required for kinetic studies and yield reproducible catalytic

parameters consistent with previously reported values.

3.8 References

1.   Steitz, T. A. (1999). DNA polymerases: structural diversity and common mechanisms. *J Biol Chem, 274*(25), 17395-17398.

2.   Zahurancik, W. J., Klein, S. J., Suo, Z. (2014). Significant contribution of the 3'-->5' exonuclease activity to the high fidelity of nucleotide incorporation catalyzed by human DNA polymerase. *Nucleic Acids Res, 42*(22), 13853-13860.

3.   Xia, S., Konigsberg, W. H. (2014). RB69 DNA polymerase structure, kinetics, and fidelity. *Biochemistry, 53*(17), 2752-2767.

4.   Petruska, J., Sowers, L. C., Goodman, M. F. (1986). Comparison of nucleotide interactions in water, proteins, and vacuum- model for DNA polymerase fidelity. *Proc. Natl. Acad. Sci. USA, 83*, 1559-1562.

5.   Zhang, H., Rhee, C., Bebenek, A., Drake, J. W., Wang, J., Konigsberg, W. (2006). The L561A Substitution in the Nascent Base-Pair Binding Pocket of RB69 DNA Polymerase Reduces Base Discrimination. *Biochemistry, 45*, 2211-2220.

6.   Zhang, L., Brown, J. A., Newmister, S. A., Suo, Z. (2009). Polymerization fidelity of a replicative DNA polymerase from the hyperthermophilic archaeon Sulfolobus solfataricus P2. *Biochemistry, 48*(31), 7492-7501.

7.   Dieckman L.M., Johnson R.E., Prakash S., Washington M.T. (2010). Pre-steady state kinetic studies of the fidelity of nucleotide incorporation by yeast DNA polymerase delta, *Biochemistry*, *49*, 7344-7350.

8.   Lee, H. R., Johnson, K. A. (2006). Fidelity of the human mitochondrial DNA polymerase. *J Biol Chem, 281*(47), 36236-36240.

9.   Wong, I., Patel, S. S., Johnson, K. A. (1991). An induced-fit kinetic mechanism for DNA replication fidelity- direct measurement by single-turnover kinetics. *Biochemistry, 30*, 526-537.

10.   Donlin, M. J., Patel, S. S., Johnson, K. A. (1991). Kinetic Partitioning between the Exonuclease and Polymerase Sites in DNA Error Correction. *Biochemistry, 30*, 538-546.

11.   Johnson, A. A., Johnson, K. A. (2001). Exonuclease proofreading by human mitochondrial DNA polymerase. *J Biol Chem, 276*(41), 38097-38107.

12.   Garcia-Diaz, M., Bebenek, K. (2007). Multiple functions of DNA polymerases. *CRC Crit Rev Plant Sci, 26*(2), 105-122.

13.   McElhinny, N. S. A., Gordenin, D. A., Stith, C. M., Burgers, P. M., Kunkel, T. A. (2008). Division of labor at the eukaryotic replication fork. *Mol Cell, 30*(2), 137-144.

14.   Pursell, Z. F., Isoz, I., Lundstrom, E-B., Johansson, E., Kunkel, T. A. (2007). Yeast DNA polymerase epsilon participates in leading-strand DNA replication. *Science, 317*(5834), 127-130.

15.   Shinbrot, E., Henninger, E. E., Weinhold, N., Covington, K. R., Goksenin, A. Y., Schultz, N., Chao, H., Doddapaneni, H., Muzny, D. M., Gibbs, R. A., Sander, C., Pursell, Z. F., Wheeler, D. A. (2014). Exonuclease mutations in DNA polymerase epsilon reveal replication strand specific mutation patterns and human origins of replication. *Genome Res, 24*(11), 1740-1750.

16. Miyabe, I., Kunkel, T. A., Carr, A. M. (2011). The major roles of DNA polymerases epsilon and delta at the eukaryotic replication fork are evolutionarily conserved. *PLoS Genet, 7*(12), e1002407.

17. Bermudez, V. P., Farina, A., Raghavan, V., Tappin, I., Hurwitz, J. (2011). Studies on human DNA polymerase epsilon and GINS complex and their role in DNA replication. *J Biol Chem, 286*(33), 28963-28977.

18. Li, Y., Asahara, H., Patel, V. S., Zhou, S., Linn, S. (1997). Purification, cDNA cloning, and gene mapping of the small subunit of human DNA polymerase epsilon. *J Biol Chem, 272*(51), 32337-32344.

19. Chui, G., Linn, S. (1995). Further characterization of HeLa DNA polymerase epsilon. *J Biol Chem, 270*(14), 7799-7808.

20. Syvaoja, J., Suomensaari, S., Nishida, C., Goldsmith, J. S., Chui, G. S. J., Jain, S., Linn, S. (1990). DNA polymerases a, 6, and E- Three distinct enzymes from HeLa cells. *Proc. Natl. Acad. Sci. USA, 87*, 6664-6668.

21. Hogg, M., Osterman, P., Bylund, G. O., Ganai, R. A., Lundstrom, E. B., Sauer-Eriksson, A. E., Johansson, E. (2014). Structural basis for processive DNA synthesis by yeast DNA polymerase epsilon. *Nat Struct Mol Biol, 21*(1), 49-55.

22. Jain, R., Rajashankar, K. R., Buku, A., Johnson, R. E., Prakash, L., Prakash, S., Aggarwal, A. K. (2014). Crystal structure of yeast DNA polymerase epsilon catalytic domain. *PLoS One, 9*(4), e94835.

23. The Cancer Genome Atlas Network. (2012). Comprehensive molecular characterization of human colon and rectal cancer. *Nature, 487*(7407), 330-337.

24. Kandoth, C., Schultz, N., Cherniack, A. D., Akbani, R., Liu, Y., Shen, H., Robertson, A. G., Pashtan, I., Shen, R., Benz, C. C., Yau, C., Laird, P. W., Ding, L., Zhang, W., Mills, G. B., Kucherlapati, R., Mardis, E. R., Levine, D. A. (2013). Integrated genomic characterization of endometrial carcinoma. *Nature, 497*(7447), 67-73.

25. Rayner, E., van Gool, I. C., Palles, C., Kearsey, S. E., Bosse, T., Tomlinson, I., Church, D. N. (2016). A panoply of errors: polymerase proofreading domain mutations in cancer. *Nat Rev Cancer, 16*(2), 71-81.

26. Barbari, S. R., Shcherbakova, P. V. (2017). Replicative DNA polymerase defects in human cancers: Consequences, mechanisms, and implications for therapy. *DNA Repair (Amst), 56*, 16-25.

27. Kane, D. P., Shcherbakova, P. V. (2014). A common cancer-associated DNA polymerase epsilon mutation causes an exceptionally strong mutator phenotype, indicating fidelity defects distinct from loss of proofreading. *Cancer Res, 74*(7), 1895-1901.

28. Albertson, T. M., Ogawa, M., Bugni, J. M., Hays, L. E., Chen, Y., Wang, Y., Treuting, P. M., Heddle, J. A., Goldsby, R. E., Preston, B. D. (2009). DNA polymerase epsilon and delta proofreading suppress discrete mutator and cancer phenotypes in mice. *Proc Natl Acad Sci U S A 106*, 17101-17104.

29. Goldsby, R. E., Hays, L. E., Chen, X., Olmsted, E. A., Slayton, W. B., Spangrude, G. J., Preston B. D. (2002). High incidence of epithelial cancers in mice deficient for DNA polymerase delta proofreading. *Proc Natl Acad Sci U S A 99*, 15560-15565.

30. Cerami, E. Gao, J., Dogrusoz, U., Gross, B. E., Sumer, S. O., Aksoy, B. A., Jacobsen, A., Byrne, C. J., Heuer, M. L., Larsson, E., Antipin, Y., Reva, B., Goldberg, A. P., Sander, C., Schultz, N. (2012). The cBio cancer genomics portal: an open platform for exploring multidimensional cancer genomics data. *Cancer Discov 2*, 401-404

31. Copeland, W. C., Ponamarev, M. V., Nguyen, D., Kunkel, T. A. & Longley, M. J. (2003). Mutations in DNA polymerase gamma cause error prone DNA synthesis in human mitochondrial disorders. *Acta Biochim Pol 50*, 155-167.

32. Singh, K. K., Ayyasamy, V., Owens, K. M., Koul, M. S. & Vujcic, M. (2009). Mutations in mitochondrial DNA polymerase-gamma promote breast tumorigenesis. *J Hum Genet 54*, 516-524.

33. Sohl, C. D., Kasiviswanathan, R., Copeland, W. C. & Anderson, K. S. (2013) Mutations in human DNA polymerase gamma confer unique mechanisms of catalytic deficiency that mirror the disease severity in mitochondrial disorder patients. *Hum Mol Genet 22*, 1074-1085.

34. Daee, D. L., Mertz, T. M. & Shcherbakova, P. V. (2010). A cancer-associated DNA polymerase delta variant modeled in yeast causes a catastrophic increase in genomic instability. *Proc Natl Acad Sci U S A 107*, 157-162.

35. Zahurancik, W. J., Klein, S. J., Suo, Z. (2013). Kinetic mechanism of DNA polymerization catalyzed by human DNA polymerase epsilon. *Biochemistry, 52*(40), 7041-7049.

36. Zahurancik, W. J., Baranovskiy, A. G., Tahirov, T. H., Suo, Z. (2015). Comparison of the kinetic parameters of the truncated catalytic subunit and holoenzyme of human DNA polymerase varepsilon. *DNA Repair (Amst), 29*, 16-22.

37. Zahurancik, W. J., Suo, Z. (2020). Kinetic investigation of the polymerase and exonuclease activities of human DNA polymerase epsilon holoenzyme. *J Biol Chem, 295*(50), 17251-17264.

38. Korona, D. A., LeCompte, K. G., Pursell, Z. F. (2011). The high fidelity and unique error signature of human DNA polymerase epsilon. *Nucleic Acids Research, 39*(5), 1763-1773.

39. Ganai, R. A., Osterman, P., Johansson, E. (2015). Yeast DNA polymerase catalytic core and holoenzyme have comparable catalytic rates. *J Biol Chem, 290*(6), 3825-3835.

40. Washington, M. T., Johnson, R. E., Prakash, L., Prakash, S. (2003). The mechanism of nucleotide incorporation by human DNA polymerase eta differs from that of the yeast enzyme. *Mol Cell Biol, 23*(22), 8316-8322.

41. Parkash, V., Kulkarni, Y., Ter Beek, J., Shcherbakova, P. V., Kamerlin, S. C. L., Johansson, E. (2019). Structural consequence of the most frequently recurring cancer-associated substitution in DNA polymerase epsilon. *Nat Commun, 10*(1), 373.

42.   Xing, X., Kane, D. P., Bulock, C. R., Moore, E. A., Sharma, S., Chabes, A., Shcherbakova, P. V. (2019). A recurrent cancer-associated substitution in DNA polymerase epsilon produces a hyperactive enzyme. *Nat Commun,* *(10)*

43.   Johnson, K. A. (1992). Transient-State Kinetic Analysis of Enzyme Reaction Pathways. *The Enzymes, 20*, 1-61.

44.   Johnson, K. A. (1995). Rapid quench kinetic analysis of polymerases, adenosinetriphosphatases, and enzyme intermediates. *Methods Enzymol, 249*, 38-61.

45.   Johnson, K. A. (2010). The kinetic and chemical mechanism of high-fidelity DNA polymerases. *Biochim Biophys Acta, 1804*(5), 1041-1048.

46.   Sambrook, J., (2001). Molecular Cloning: a laboratory manual. *Cold Spring Harbor Laboratory Press, 3.*

47.   Raper, A. T., Reed, A. J., Suo, Z. (2018). Kinetic Mechanism of DNA Polymerases: Contributions of Conformational Dynamics and a Third Divalent Metal Ion. *Chem Rev, 118*(12), 6000-6025.