Fall 5-13-2017

Identification & Evaluation of DNA Ligase Inhibitors: Predicting the Binding of Small Molecules to the DNA Binding Domain by Molecular Modeling

Timothy RL Howes
University of New Mexico

Timothy Richard Lloyd Howes
Candidate

Biomedical Sciences
Department

This dissertation is approved, and it is acceptable in quality and form for publication:

Approved by the Dissertation Committee

Alan E. Tomkinson, PhD
Chairperson

Mary Ann Osley, PhD

Montaser Shaheen, MD

Tudor Oprea, MD, PhD

IDENTIFICATION & EVALUATION OF DNA LIGASE INHIBITORS: PREDICTING THE BINDING OF SMALL MOLECULES TO THE DNA BINDING DOMAIN BY MOLECULAR MODELING

by

Timothy Richard Lloyd Howes

B.A.: Biochemistry and Molecular Biology, Drew University

DISSERTATION

Submitted in Partial Fulfillment of the
Requirements for the Degree of

DOCTOR OF PHILOSOPHY
BIOMEDICAL SCIENCES

The University of New Mexico
Albuquerque, New Mexico

May 2017

ACKNOWLEDGMENTS

I would like to thank Alan Tomkinson for taking me into his lab, and for remaining committed to my success as a PhD candidate. I would also like to acknowledge and thank the members of my committee, Mary Ann Osley, Montaser Shaheen, and Tudor Oprea, for their invaluable feedback on my work. I would also like to thank Darin Jones, for being my outside reader.

The current members of the Tomkinson lab, Nathanial Wiest, Ishtiaque Rashid, Rhys Brooks, and Krystal Henderson are all exceptionally talented individuals. I would like to single out the final member of our lab, Annahita Sallmyr, for her endless patience and advice. Our lab's former members, Yoshihiro Matsumoto, Zhimin Peng, Hui Yang, Julie Della-Maria, John O'Neil, and Dibyendu Banerjee have also helped me throughout my graduate work. Yoshi, whom I sat beside for years, has helped me more than words can possibly express. Feedback from members of our joint lab meeting group has also been invaluable, as well as input and suggestions from SBDR fellows and PIs. I would also like to thank my former teachers and professors who got me here, particularly my high school chemistry teachers Adrian Price and Christine Whitlock, and my undergraduate advisor Adam Cassano.

I could not have succeeded without the aid and companionship of friends, who were study partners, proofreaders, and critics. They encouraged and supported, shared drinks and hardships with me, they were friends at the table.

All research requires funding, of course. My research was supported by the University of New Mexico Comprehensive Cancer Center (P30 CA118100) and National Institute of Health Grants R01 GM57479 (to A.E.T.) and P01 CA92584. I would particularly like to thank Genevieve Phillips at the UNM Fluorescence Microscopy Shared Resource for her invaluable knowledge and assistance. Additionally, I would also like to thank my brother, David Howes, for creating software for me that greatly expedited the transcription of residue interaction data.

I would not be who or where I am today without my family. Thank you to my parents, Paul and Val Howes, and my brother David, for putting up with me. Thank you to my Aunt Sally, my Uncles Chris and Hefin, as well as my cousins James and Carol. My nana, Mollie Howes, thank you for being magnificent. Anyone who knows me will not be surprised to find that I have taken the time to thank my dogs, Zoe and Zonda, as well. You are both Good Dogs.

Finally, to my long-suffering partner, Meghan Kessler, whom I love with all my heart. Thank you for everything. Your support, even as I moved 2,000 miles away, has always been unwavering. You have always encouraged me to be my best.

Identification & Evaluation of DNA Ligase Inhibitors: Predicting the Binding of Small Molecules to the DNA Binding Domain by Molecular Modeling

by

Timothy Richard Lloyd Howes

B.A., Biochemistry & Molecular Biology, Drew University, 2008

Ph.D., Biomedical Sciences, The University of New Mexico, 2017

ABSTRACT

The phosphodiester backbone of DNA is maintained by DNA ligases. In human cells, there are three genes that encode DNA ligase polypeptides with distinct but overlapping functions. A series of small molecule inhibitors of human DNA ligases were previously identified using a rational structure-based approach. Three of these inhibitors, L82, a DNA ligase I selective inhibitor, and L67, an inhibitor of DNA ligases I and III, and L189, an inhibitor of all three human DNA ligases, have related structures. Here I present and characterize L82-G17 a next-generation ligase I specific inhibitor. L82-G17 is a potent ligase I uncompetitive inhibitor that is shown to act both biochemically and in cell culture models. Furthermore, the binding site for L82 and L82-G17 has been identified via molecular modeling, and verified through mutagenesis.

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

CHAPTER 1

Introduction

Structure and Function of DNA Ligases

Since DNA polymerases only synthesize DNA from 5' to 3', one of the two antiparallel strands of duplex DNA must be synthesized discontinuously as a series of short Okazaki fragments that are then be joined by a DNA ligase to generate an intact strand. In 1967, several laboratories identified DNA ligase activity in extracts from both uninfected *E. coli* cells and *E. coli* cells infected with bacteriophage T4 (Lehman 1974). The following year DNA ligase activity was described in extracts from mammalian cells (Soderhall and Lindahl 1976). Notably, the *Escherichia coli* DNA ligase is NAD^+-dependent whereas the bacteriophage and mammalian DNA ligases are ATP-dependent (Lehman 1974, Soderhall and Lindahl 1976). Subsequent studies have revealed the existence of both NAD^+- and ATP-dependent DNA ligases in prokaryotes. In contrast, eukaryotic and viral DNA ligases are almost exclusively ATP-dependent (Tomkinson, Vijayakumar et al. 2006, Ellenberger and Tomkinson 2008).

Apart from utilizing a different nucleotide co-factor, the reaction mechanisms of NAD^+- and ATP-dependent are identical. DNA ligases initially

react with the nucleotide co-factor to form a covalent DNA ligase-adenylate complex in which the AMP moiety is linked to a specific lysine residue via a phosphoramidite bond. When the DNA ligase-adenylate engages a DNA nick with 3' OH and 5' phosphate termini, it transfers the AMP group to the 5' phosphate, forming a covalent DNA adenylate intermediate. Finally, the non-adenylated DNA ligase interacts with the DNA-adenylate, catalyzing phosphodiester bond formation and release of AMP as a result of nucleophilic attack on the 5' DNA adenylate by the 3' OH group.

The first eukaryotic DNA ligase genes were identified in screens for cell division cycle mutants in the yeasts, *Saccharomyces cerevisiae* and *Schizosaccharomyces pombe* (Nasmyth 1977, Johnston and Nasmyth 1978). The DNA ligases encoded by the *CDC9* gene in *Saccharomyces cerevisiae* and the *CDC17* gene in *Schizosaccharomyces pombe* are required for cell viability because of their essential role in DNA replication. Biochemical and immunological characterization of DNA ligase activity in mammalian cell extracts provided the first evidence that eukaryotes contain more than one species of DNA ligase (Soderhall and Lindahl 1976). The presence of more than one species of DNA ligase suggested that these enzymes may have distinct cellular functions. My work, presented here focuses on the development of small molecule inhibitors of human DNA ligases, with a focus on human DNA ligase I.

Eukaryotic DNA ligase genes

As mentioned above, the DNA ligases encoded by the *CDC9* and *CDC17* genes of *Saccharomyces cerevisiae* and *Schizosaccharomyces pombe*, respectively, were the first eukaryotic DNA ligases to be identified (Nasmyth 1977, Johnston and Nasmyth 1978). Human cDNAs that complemented the temperature sensitive phenotype of a yeast *cdc9* strain were isolated from a human cDNA library (Barnes, Johnston et al. 1990). Subsequent DNA sequencing revealed that these cDNAs encoded a polypeptide that is highly homologous with the yeast DNA ligases and contained sequences that were identical to those of peptides from purified mammalian DNA ligase I (Barnes, Johnston et al. 1990). Thus, human DNA ligase I and the yeast DNA ligases are functional homologs that belong to the eukaryotic DNA ligase I family.

Nuclear DNA Metabolism	
DNA Replication	DNA Ligase I (primary) DNA Ligase III
Mismatch Repair	Unknown
Base Excision Repair	DNA Ligase III (short-patch) DNA Ligase I (long-patch)
Nucleotide Excision Repair	DNA Ligase III (constitutive) DNA Ligase I (S-phase)
Single-Strand Break Repair	DNA Ligase III (long-patch) DNA Ligase I (short-patch)
Homology Mediated Double-Strand Break Repair	Unknown
Non-Homologous End Joining	DNA Ligase IV DNA Ligase I (alternative) DNA Ligase III (alternative)
Mitochondrial DNA Metabolism	
DNA Replication	DNA Ligase III
DNA Repair	DNA Ligase III

Table 1. Functional redundancies of human DNA ligases

Two other mammalian genes that encode DNA ligases, *LIG3* and *LIG4*, have been identified (Chen, Tomkinson et al. 1995, Wei, Robins et al. 1995).

Homologs of the *LIG4* gene have been found in all eukaryotes, whereas *LIG3* has been found in some, but not all, eukaryotes (Ellenberger and Tomkinson 2008, Simsek and Jasin 2011). That *LIG1* null cells and cells lacking nuclear LigIII are both viable indicates a significant degree of functional redundancy between these two enzymes in nuclear DNA metabolism (**Table 1**) (Tomkinson, Howes et al. 2013). In contrast, DNA ligase III (LigIII) has been shown to be essential for mitochondrial DNA (mtDNA) replication and repair in an XRCC1-independent fashion. However, mitochondrial LigIII function can be replaced by mitochondrially-targeted heterologous DNA ligases, including LigI, or even the NAD^+-dependent LigA of *E. coli* (Simsek, Furda et al. 2011).

The DNA ligases encoded by the three *LIG* genes share a conserved catalytic region that is flanked by unrelated amino- and/or carboxyl-terminal regions. There is compelling evidence that interactions with specific protein partners mediated by these unique regions flanking the catalytic domain direct the participation of the DNA ligases in different DNA transactions (Ellenberger and Tomkinson 2008). An enzyme activity previously designated as DNA ligase II was later shown to be a proteolytic fragment of DNA ligase III (Chen, Tomkinson et al. 1995).

DNA ligase I; molecular genetics and cell biology

The gene encoding DNA ligase I, *LIGI*, which spans 58kb, is located on chromosome 19q13.2-13.3, and is made up of 28 exons

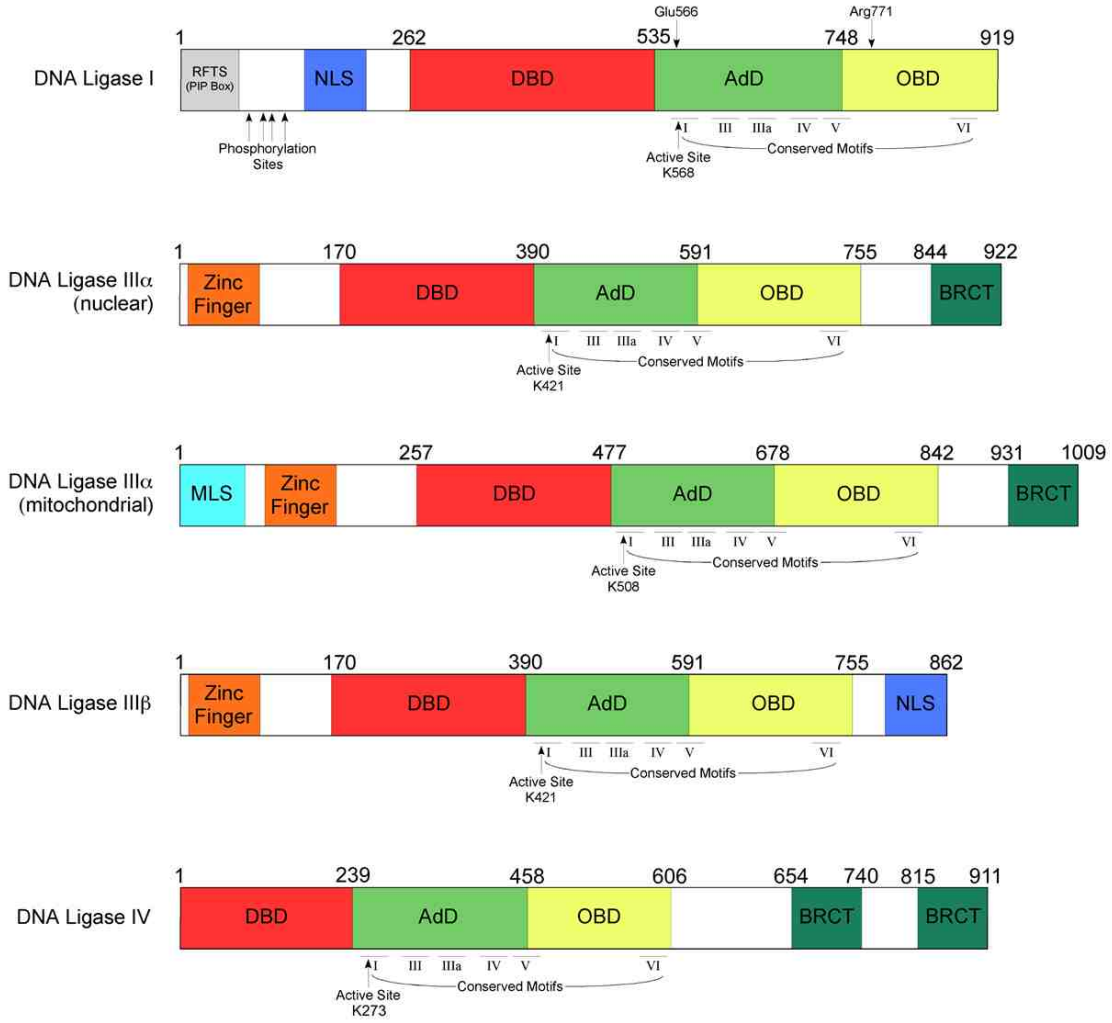


Figure 1. Human DNA ligases

The three human ligase genes encode several polypeptides. Each human ligase contains a DNA binding domain (DBD, red) as well as a catalytic core, which consists of an adenylation domain (AdD, green) and an oligomer binding domain (OBD, yellow). The catalytic core is common to all DNA ligases, as well as mRNA capping enzymes. Each DNA ligase III isoform possesses an N-terminal zinc finger (orange) that aids in DNA binding. Ligases IIIα and IV also have a "breast and ovarian cancer susceptibility protein-1 C-terminal" (BRCT, green) domain. Mitochondrial LigIIIα has a mitochondrial localization signal (MLS, cyan), while LigI and LigIIIβ have a nuclear localization signal (NLS, blue). In addition to the NLS, the unstructured N-terminal region of LigI also contains a replication factory targeting sequence (RFTS, grey), also known as the PIP box (PCNA-interacting peptide). Furthermore, several phosphorylation sites are located in the unstructured N-terminal region of LigI. The active site lysine, which binds AMP, has been indicated for each ligase, as well as each of the six conserved motifs. Finally, the two residues identified as mutated in the only reported case of LigI deficiency, Glu 566 and Arg 771, are indicated (Howes and Tomkinson 2012).

(Barnes, Tomkinson et al. 1992, Noguiez, Barnes et al. 1992). The increased expression of the *LIG1* gene when quiescent cells are induced to proliferate and the increased levels of DNA ligase I protein and activity in proliferating cells and tissues, implicated DNA ligase I in DNA replication (Soderhall and Lindahl 1976, Petrini, Huwiler et al. 1991). This linkage was strengthened by studies showing that DNA ligase I co-localized with replication foci in S phase cells (Lasko, Tomkinson et al. 1990). Distinct amino acid sequences within the non-catalytic N-terminal region of DNA ligase I function as nuclear localization (NLS) and replication foci targeting sequences (**Fig. 1**) (Montecucco, Savini et al. 1995, Cardoso, Joseph et al. 1997). The mechanism underlying the recruitment of DNA ligase I to replication foci is described below.

A single case of human DNA ligase I-deficiency has been described. This individual, whose symptoms included impaired growth, delayed development, recurrent ear and chest infections and lymphoma, died at age 19 as result of complications following a chest infection (Webster, Barnes et al. 1992). Sequencing of genomic DNA revealed the presence of two different mutant *lig1* alleles. The maternally inherited *lig1* allele encodes a DNA ligase polypeptide with reduced catalytic activity whereas the other mutant allele, whose origin is not known, encodes a DNA polypeptide with essentially no catalytic activity. In both mutant *lig1* alleles, the DNA sequence change results in a single amino acid substitution within the conserved catalytic region of DNA ligase I (Barnes, Tomkinson et al. 1992). The locations of the amino acid changes are described

in the section below. Primary (46BR) and SV40-immortalized (46BR.1G1) fibroblasts established from the DNA ligase I-deficient individual exhibit defective joining of Okazaki fragments and sensitivity to a wide range of DNA damaging agents, particularly DNA alkylating agents (Teo, Arlett et al. 1983). Both mutant *lig1* alleles are present in the primary fibroblasts, whereas, only the maternally inherited allele is present in the SV40-immortalized (46BR.1G1) fibroblasts. It appears that the maternal *lig1* allele is responsible for the patient's symptoms and the phenotype of the cell lines (Barnes, Tomkinson et al. 1992). As expected, both the DNA replication and repair defects of the 46BR.1G1 fibroblasts are complemented by expression of wild type DNA ligase I (Levin, McKenna et al. 2000). In addition, a DNA ligase I-deficient *Arabidopsis* plant cell line that showed severe growth defects, as well as delayed repair of single and double strand breaks has been described (Waterworth, Kozak et al. 2009).

Although the levels of DNA ligase I protein and activity are reduced by about 50% and 90%, respectively in the 46BR.1G1 fibroblasts compared with SV40-immortalized fibroblasts from a normal individual, there are no significant differences in cell cycle progression despite the defect in converting Okazaki fragments into high molecular weight DNA (Barnes, Tomkinson et al. 1992). In fact, results of pulse-labeling studies indicate that the majority of Okazaki fragments are degraded rather than ligated (Henderson, Arlett et al. 1985, Levin, Vijayakumar et al. 2004). Thus, it appears that, when DNA ligase I is not available to ligate the nick between adjacent Okazaki fragments, the downstream

fragment is displaced by DNA synthesis and then degraded. This model predicts that the lagging strand in 46BR.1G1 fibroblasts is synthesized as a series of longer fragments that are joined by either the defective DNA ligase I polypeptide or one of the other DNA ligases.

Based on the results of genetic studies in the yeasts *Saccharomyces cerevisiae* and *Schizosaccharomyces pombe* (Nasmyth 1977, Johnston and Nasmyth 1978), it was expected that the mammalian *LIG1* gene would be essential. In accord with this prediction, *lig1* null mouse embryonic stem cells could only be obtained when full length wild type DNA ligase I cDNA was ectopically expressed (Petrini, Xiao et al. 1995). Surprisingly, *lig1* null embryos generated by crossing heterozygous mice were detectable until day 16 (Bentley, Selfridge et al. 1996, Bentley, Harrison et al. 2002). Furthermore, it was possible to establish *lig1* null embryonic fibroblasts (MEFs) from these embryos, demonstrating that *LIG1* is not an essential gene in mouse somatic cells. Like the human 46BR.1G1 fibroblasts, the *lig1* null MEFs had a defect in converting Okazaki fragments into high molecular weight DNA but no defect in proliferation (Bentley, Selfridge et al. 1996, Bentley, Harrison et al. 2002). In contrast to the 46BR.1G1 fibroblasts, the *lig1* null MEFs have no apparent DNA repair defect (Teo, Arlett et al. 1983, Bentley, Selfridge et al. 1996, Bentley, Harrison et al. 2002). There are also *lig1* null mouse CH12F3 cells, in which exons 18-19 are removed, resulting in a frameshift in the *LIG1* which renders the protein non-functional. While these cells were also shown not to have an increased sensitivity

to DNA damaging agents such as cisplatin and camptothecin, they were more sensitive to methyl methane-sulfonate (Han, 2014).

The presence of additional DNA ligases in mammals encoded by the *LIG3* gene provides a possible explanation as to why the *LIG1* gene homolog, *CDC9*, is essential for cell viability in yeast but not in mammals. For example, DNA ligase IIIα and its partner protein XRCC1 are recruited to participate in the repair of DNA single strand breaks by an interaction with the poly(ADP-ribosylated) version of poly(ADP-ribose) polymerase 1 (PARP-1), an abundant nuclear protein that binds to and is activated by DNA single strand breaks (Okano, Lan et al. 2003, Okano, Lan et al. 2005). The defect in Okazaki fragment processing caused by DNA ligase I deficiency is likely to result in relatively long-lived, single-strand interruptions on the lagging strand. It is possible that these breaks are recognized and joined by the PARP-1/DNA ligase III single strand break repair pathway. The hyper sensitivity of human 46BR.1G1 fibroblasts (Lehmann, Willis et al. 1988) to a PARP inhibitor is consistent with the single strand break repair pathway joining single strand breaks between Okazaki fragments that remain after lagging strand DNA synthesis. It is important to note that recent publications have established that LigIII enable cells with reduced levels or absence of DNA ligase I to proliferate (Le Chalony, Hoffschir et al. 2012). However, LigIV cannot substitute for LigI to ligate Okazaki fragments, as *LIG1* and *LIG3* dual knockouts in chicken DT40 cells are synthetic lethal (Arakawa and Iliakis 2015).

A mouse knock-in model that reiterates the mutant allele of DNA ligase I that is expressed in the human SV-40 immortalized 46BR.1G1 cells has been generated. These animals are small and have hematopoietic defects (Harrison, Ketchen et al. 2002). Other notable features include increased genomic instability and an increased incidence of epithelial tumors (Harrison, Ketchen et al. 2002). As with the *lig1* null MEFs, the MEFs harboring the equivalent mutation to that in the human SV-40 immortalized 46BR.1G1 fibroblasts have a defect in replication but not repair, suggesting that the increased genome instability and cancer incidence is due to accumulation of abnormal replication intermediates (Harrison, Ketchen et al. 2002). Together, these studies indicate that DNA ligase I has a more important role in DNA repair in human than murine cells (Teo, Arlett et al. 1983, Harrison, Ketchen et al. 2002) and suggest that the relative contribution of DNA ligase IIIα-dependent repair may be greater in murine cells compared with human cells.

Structure of the Human DNA Ligase I Protein

The 919 amino acid polypeptide encoded by human DNA ligase I cDNA has a highly asymmetric shape, which results in abnormal behavior during density gradient sedimentation and gel filtration experiments (Tomkinson, Lasko et al. 1990). Using limited proteolysis, it was found that catalytic activity resides within a relatively protease-resistant C-terminal fragment of about 78 kilo Daltons (kDa) whereas the N-terminal fragment is extremely protease sensitive, indicative

of an unstructured region (Tomkinson, Lasko et al. 1990). Notably, this N-terminal region is likely to have an extended, flexible conformation because it contains a large number of proline residues (Barnes, Johnston et al. 1990). In addition, the high proline content of DNA ligase I (approximately 9%) results in abnormally mobility during SDS-polyacrylamide gel electrophoresis, such that DNA ligase I an apparent molecular mass of 125 kDa compared with the actual molecular weight of 102,000 (Tomkinson, Lasko et al. 1990).

The catalytic region of DNA ligase I contains six motifs that are conserved among the nucleotidyl transferase family, including mRNA capping enzymes, RNA ligases, and DNA ligases (Shuman and Schwer 1995). Motif I contains the active site lysine residue to which the AMP (or GMP) residue is attached via a covalent phosphoramidite bond. This residue was initially identified by determining the sequence of an adenylylated tryptic peptide from bovine DNA ligase I (Tomkinson, Totty et al. 1991). Using this sequence, it was possible to predict the position of the putative active site lysine residues in DNA ligases, RNA ligases, and mRNA capping enzymes. As expected, substitution of Lys568, the lysine residue that is covalently linked to the AMP moiety in human DNA ligase I, prevents formation of the enzyme-AMP complex and consequently abolishes enzymatic activity (Kodama, Barnes et al. 1991, Tomkinson, Totty et al. 1991). One of the mutant alleles in the DNA ligase I-deficient individual encodes a polypeptide with has a lysine residue in place of a glutamic acid at position 566 (Barnes, Tomkinson et al. 1992). This amino acid change two

residues away from the active site lysine markedly reduces formation of the enzyme-AMP intermediate (Kodama et al, 1991), indicating that this mutant allele encodes a polypeptide with very little or no activity.

The first nucleotidyl transferase structure to be determined was that of the DNA ligase encoded by bacteriophage T7 (Subramanya, Doherty et al. 1996). This was shortly followed by the characterization of the RNA-capping enzyme from PCBV-1 (Hakansson, Doherty et al. 1997). These structures revealed the existence of two domains, an adenylation/guanylation domain, which contains conserved motifs I through V, and an oligomer binding-fold (OB) domain (OBD) containing motif VI (**Fig. 1)**. In 2004, Pascal *et al.* successfully crystallized the catalytic C-terminal region of DNA ligase I (residues 233-919) bound to a nicked DNA substrate (Pascal, O'Brien et al. 2004). This structure revealed several novel features. Firstly, it showed that nicked DNA is encircled by DNA ligase I during catalysis, suggesting that the catalytic domain undergoes a large conformational change when it engages a nick. Secondly, it showed that the catalytic region of the larger eukaryotic DNA ligases contains a DNA binding domain in addition to the adenylation and OB-fold domains that make up the catalytic core. Thus, the adenylation/guanylation and OB-fold domains constitute the conserved catalytic core of nucleotidyl transferases with the DNA binding domain (DBD) being a characteristic feature of eukaryotic DNA ligases (Pascal, O'Brien et al. 2004).

The DNA binding domain of DNA ligase I, which spans residues 262 to 534, folds into 12 α-helices that exhibit a two-fold symmetry (Pascal, O'Brien et al. 2004). Due to the symmetry of the DBD, the interaction with DNA occurs via one tight reverse turn of two α-helices and an extended loop formation. This arrangement of the loop and helices creates a relatively flat surface of approximately 2000 $Å^2$ that interacts almost exclusively with the phosphodiester backbone of the DNA substrate. The DBD interacts with the minor groove of the DNA backbone on both sides of the nick, explaining how DNA ligase I binds to DNA in a sequence-independent binding manner and why chemicals that bind to the minor groove of DNA, such as distamycin, inhibit DNA ligase activity (Montecucco, Fontana et al. 1991). Notably, the DBD stimulates the weak DNA joining activity of the DNA ligase I catalytic core containing the AdD and OBD when added *in trans*, indicating that contacts between the DBD and both AdD and the OBD observed in the crystal structure stabilize the folding of the catalytic core around the DNA nick (Pascal, O'Brien et al. 2004).

The DNA ligase I adenylation domain, which spans from residue 535 to 747, contains conserved motifs I, III, IIIa, IV and V. These five motifs contribute to the surface of nucleotide binding pocket. Tryptophan 742, of motif V, provides co-factor specificity by sterically excluding GTP. Furthermore, Arg 573 and Glu 621, of motifs I and III, respectively, stabilize the hydroxyl groups on the ribose sugar of AMP via hydrogen bonding interactions (Pascal, O'Brien et al. 2004). As

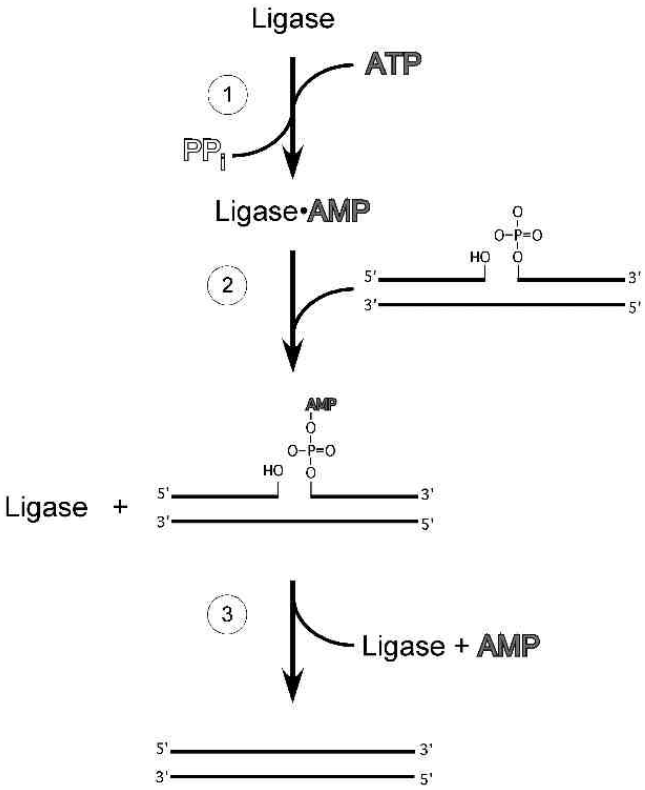


Figure 2. DNA ligation is a three-step process

First, the DNA ligase binds and hydrolyses ATP, pyrophosphate is released and a AMP is bound to DNA ligase. Second, the adenosine monophosphate is transferred to the 5' phosphate of the DNA at the nick. Finally, the ligase catalyzes the phosphodiester bond formation by nucleophilic attack by the 3' hydroxide. Both AMP and the ligase dissociate from the DNA (Howes and Tomkinson 2012).

mentioned above, one of the two mutant *LIG1* alleles identified in the individual with DNA ligase I deficiency encodes a polypeptide in which Glu 566, of motif I, is replaced by a lysine residue (Barnes, Tomkinson et al. 1992). From the structure of DNA ligase I, it is evident that Glu 566 contributes to the specific interaction with ATP by forming a hydrogen-bond with the N6 of the adenine moiety (Pascal, O'Brien et al. 2004). Replacement of Glu 566 with a positively charged lysine residues disrupts this, providing an explanation as to why the mutant polypeptide is defective in the first step of the ligation reaction, formation of the covalent enzyme-adenylate intermediate (**Fig. 2**).

The major structural feature of the OB-fold domain is a β-barrel and, similar to the other two domains, it also interacts with the minor groove of the DNA (Pascal, O'Brien et al. 2004). Contacts between AdD and OBD are critical for correctly positioning these domains when they engage a DNA nick. During catalysis, the AdD forms a salt bridge with the OBD via Asp 570 and Arg 871, stabilizing the ligase catalytic domains in a conformation in which they fully encircle the DNA nick. This positioning of the AdD and OBD creates a surface that binds to and distorts the nicked DNA. Notably, the phenylalanine residues at positions 635 and 872 of the AdD and OBD, are forced into the minor groove both 3' and 5' to the nick. As a result of these interactions, the DNA duplex upstream of the nick duplex assumes an A-form structure as the nick is opened up for ligation (Pascal, O'Brien et al. 2004). Notably, the DNA binding site downstream of the nick is specific for B-form DNA, explaining why ligase I is not active on nicks within A-form duplexes formed by RNA duplexes and RNA-DNA hybrids (Pascal, O'Brien et al. 2004). This ability to discriminate against duplexes containing ribonucleotides 5' to the nick presumably prevents premature joining of Okazaki fragments before the RNA primer has been removed. The maternally inherited mutant *LIG1* allele in the individual with DNA ligase I-deficiency encodes a polypeptide in which the arginine 771 within the OBD is replaced by a tryptophan residue (**Fig. 1**) (Barnes, Tomkinson et al. 1992). This mutant enzyme has markedly reduced catalytic activity and is, as expected, defective in step 2 of

the ligation reaction, transfer of the AMP moiety from the ligase to the 5' phosphate termini of the DNA nick (Prigent, Satoh et al. 1994).

Although eukaryotic DNA ligase I has not been crystallized in the absence of nicked DNA, others DNA ligases have been crystalized in the absence of DNA substrate; the structure of an ATP-dependent DNA ligase from the archaeal organism *Sulfolobus solfataricus* has been determined in the absence of DNA (Pascal, Tsodikov et al. 2006). The catalytic region of the archaeal enzyme has the same three domain organization as eukaryotic DNA ligases but, in the absence of DNA, the three domains are arranged in an extended conformation. The major difference between the extended and closed conformations of the three domains is the position of the OBD domain relative to the other two domains (Pascal, Tsodikov et al. 2006). The OBD undergoes a large change in conformation during the nicked DNA-dependent transition from the extended to the closed form with interactions between the OBD and DBD playing key roles in stabilizing the closed form. Based on structures of smaller DNA ligases, it appears likely that the OBD also undergoes conformational changes when the enzyme interacts with ATP to form the enzyme-adenylate, possibly reorienting the OBD to expose a DNA binding surface (Subramanya, Doherty et al. 1996, Odell, Sriskanda et al. 2000).

While the unstructured N-terminal region of DNA ligase I is dispensable for catalytic activity *in vitro*, it was presumed to be required for protein-protein

interactions *in vivo* (Petrini, Xiao et al. 1995, Mackenney, Barnes et al. 1997). This region contains a bipartite nuclear localization signal located between residues 111 and 179, and a sequence that is required for targeting to replication factories, residues 2 to 9 (Montecucco, Savini et al. 1995, Cardoso, Joseph et al. 1997). In addition, the N-terminal region is phosphorylated on several serine/threonine residues by casein kinase II and cyclin-dependent kinases during cell cycle progression (Prigent, Lasko et al. 1992, Frouin, Montecucco et al. 2002, Ferrari, Rossi et al. 2003) (**Fig. 1**). This results in a hyper-phosphorylated form of DNA ligase I in M phase cells. It appears likely that these phosphorylation events regulate the participation of DNA ligase I in DNA replication because phosphorylation site mutants fail to correct the DNA replication defect of DNA ligase I-deficient 46BR cells (Soza, Leva et al. 2009, Vijayakumar, Dziegielewska et al. 2009) (Peng et al.).

Human DNA Ligase I Protein-Protein Interactions

Proteins are directed to participate in complex DNA transactions such as DNA replication by specific protein-protein interactions. Proliferating cell nuclear antigen (PCNA), the eukaryotic homotrimeric DNA sliding clamp that functions as a processivity factor for the replicative DNA polymerases, was the first DNA ligase I-interacting protein to be identified (Levin, Bai et al. 1997). Residues 2 to 9 within the non-catalytic N-terminal region of DNA ligase I constitute the major PCNA binding site within DNA ligase I (Montecucco, Rossi et al. 1998). Notably,

this same sequence, which is homologous to a PCNA-interacting protein motif, or "PIP box" which has been identified in many proteins, is required for the recruitment of DNA ligase I to replication factories (Montecucco, Rossi et al. 1998). Amino acid changes that disrupt PCNA binding abolish both the recruitment of DNA ligase I to replication factories and the correction of the DNA replication defect in 46BR.1G1 cells, demonstrating the critical role of this interaction in the sub-nuclear targeting of DNA ligase I and the efficient joining of Okazaki fragments (Montecucco, Rossi et al. 1998, Levin, McKenna et al. 2000).

The PIP box motif binds to the interdomain connector loop of PCNA (Vijayakumar, Chapados et al. 2007), suggesting that, when the flexible N-terminal region of DNA ligase I initially binds to the interdomain connector loop of PCNA trimer, the catalytic region remains in an extended conformation (Pascal, Tsodikov et al. 2006). Notably, DNA ligase I stably interacts with PCNA trimers that are topologically linked to duplex DNA but only one molecule of DNA ligase I is bound per PCNA trimer (Levin, Bai et al. 1997). This suggests that the other potential binding sites are occluded either because of dynamic conformational changes in DNA ligase I as a consequence of its flexible, extended structure or because the initial docking of DNA ligase I with a PCNA trimer via the PIP box facilitates lower affinity interactions that extend the protein-protein interaction interface. In support of this latter idea, the DBD binds weakly to the subunit-subunit interface region of homotrimeric PCNA (Song, Pascal et al. 2009). Interestingly, the DBD also mediates the interaction with Rad9-Rad1-Hus1, a

heterotrimeric DNA sliding clamp involved in cell cycle checkpoints (Song, Pascal et al. 2009), and heterotrimeric PCNA from *Sulfolobus solfataricus* (Pascal, Tsodikov et al. 2006). Given the similarity in size and shape between the PCNA ring and the ring structure formed when the catalytic region of DNA ligase I engages a DNA nick, the PCNA ring may facilitate the transition of the catalytic region of DNA ligase I from the extended conformation to the compact ring structure. It has been proposed that the DBD, which provides the majority of the DNA binding affinity, serves as a pivot during this transition after initial docking via the PIP box. Unlike studies on the interaction between DNA ligase I and the heterotrimeric Rad9-Rad1-Hus1 checkpoint DNA sliding clamp (Wang, Lindsey-Boltz et al. 2006, Song, Levin et al. 2007), there are contradictory reports as to whether the interaction with PCNA stimulates nick-joining by DNA ligase I (Levin, Bai et al. 1997, Tom, Henricksen et al. 2001, Levin, Vijayakumar et al. 2004).

DNA ligase I also functionally interacts with two other DNA replication proteins, replication protein A (RPA), a heterotrimeric complex that binds to single stranded DNA (Ranalli, DeMott et al. 2002), and replication factor C (RFC), a heteropentameric complex that loads PCNA onto DNA (Levin, Vijayakumar et al. 2004). Although a direct physical interaction between RPA and DNA ligase I has not been demonstrated, RPA specifically stimulates the rate of catalysis by DNA ligase I (Ranalli, DeMott et al. 2002). In contrast to RPA, RFC inhibits DNA joining by DNA ligase I (Levin, Vijayakumar et al. 2004). This interaction and inhibition, which involves the large subunit of RFC, p140, is

abolished by replacement of the four phosphorylation site serines in DNA ligase I with glutamic acid residues (Vijayakumar, Dziegielewska et al. 2009). More recently, phosphorylation of serine 51 was specifically identified as regulating the LigI-RFC interaction. The S51D phosphomimetic mutant LigI failed to stably interact with RFC. Furthermore, 46BR.1G1 cells expressing a S51A DNA ligase I were more resistant to alkylating agent MMS than either 46BR.1G1 cells that expressed S51D or cells transfected with an empty vector.

Notably, the inhibition of DNA ligase I by RFC can be alleviated by inclusion of PCNA in the reaction, providing that DNA ligase I has a functional PIP box (Vijayakumar, Dziegielewska et al. 2009). Unlike the interaction with RFC, DNA ligase I binding to PCNA is not modulated by phosphorylation (Vijayakumar, Dziegielewska et al. 2009). Thus, the failure of the phosphorylation site mutant of DNA ligase I to complement the replication defect in DNA ligase I-deficient cells may be due to the disrupted interaction with RFC (Vijayakumar, Dziegielewska et al. 2009). Although these studies indicate that physical and functional interactions among DNA ligase I, RFC and PCNA are critical for DNA replication, the mechanisms by which these interactions contribute to Okazaki fragment processing and joining are not fully understood.

DNA Ligase Inhibitors

Identifying Potential DNA Ligase Inhibitors

Inhibitors of DNA repair are emerging as a potent new class of chemotherapeutics; in recent years, several groups around the world have become interested in developing and using DNA ligase inhibitors as anti-cancer agents (Kotnis and Mulherkar 2014). While clinically relevant compounds, such as distamycin and its derivatives, have been shown to inhibit DNA ligases, they are highly toxic, and non-specific, as their action is mediated by binding to the minor groove of DNA, not DNA ligase itself (Montecucco, Fontana et al. 1991). Efforts to identify DNA ligase inhibitors through *in vitro* screens of natural products have met with limited success (Sangkook, Ik-Soo et al. 1996) compared with *in silico* screening methods. In 2008, ten novel small molecules (**Fig. S1**) that function to inhibit mammalian DNA ligases were reported. These were identified via a computer aided drug design (CADD) screen for compounds that were predicted to bind to the DBD of human DNA ligase I (Zhong, Chen et al. 2008). The DBD was chosen because the adenylation and oligo-binding domains are common to all nucleotidyl transferase enzymes (Doherty and Suh 2000, Pascal, O'Brien et al. 2004). Additionally, the DBD is less conserved than the AdD/OBD catalytic core, increasing the likelihood finding selective inhibitors for the human DNA ligases.

A potential binding pocket in the DNA ligase I DBD formed between three residues, His 337, Arg 449, and Gly 453 was selected as the target site for the *in silico* screening. The presence of a small molecule in this binding pocket, which is created by the loops between helices 5 and 6, and helices 12 and 13 on the DNA binding surface of the DBD, (Pascal, O'Brien et al. 2004, Zhong, Chen et al. 2008) was predicted to block binding to nicked DNA, thereby inhibiting ligation. More than one million compounds were screened for their ability to physically fit within the target binding pocket using the crystal structure of human DNA ligase I (PDB ID: 1X9N) (Pascal, 2004, Berman, 2000). Preliminary screening returned approximately 50,000 compounds that were predicted to bind in this pocket based on total interaction energy. This list was further refined by testing the molecules against four different potential conformations of the DNA ligase I DBD determined by molecular dynamic simulations. The 50,000 molecules were assessed based on the strength of their electrostatic interactions within the new conformations of the target binding pocket. Compounds that did not fit certain criteria called “Lipinski’s rule of 5” such as exceeding five rotatable bonds, or number of ring structures were eliminated (Lipinski 2000). Finally, to ensure maximal chemical diversity, these compounds were organized based on their chemical properties. A total of 233 compounds were identified and then assayed for their ability to inhibit DNA ligase I, yielding the 10 compounds with activity against DNA ligase I described in more detail below (Zhong, Chen et al. 2008).

Characterizing DNA Ligase Inhibitors

Of the 233 compounds identified by the *in silico* screen, named L1 through to L233, 192 were found to be commercially available. These were acquired and assayed for their ability to inhibit human DNA ligase I. To identify any compounds that were inhibiting ligation non-specifically by binding to DNA, the candidates were also tested against T4 DNA ligase, which lacks a DNA binding domain. From this process, ten compounds that specifically inhibit DNA ligase I were identified. While the majority of these compounds were, as expected, competitive inhibitors, one inhibitor, L82, was uncompetitive (Zhong, Chen et al. 2008). This was somewhat surprising, because the screen was designed to identify inhibitors that prevent DNA binding.

The DNA ligase I inhibitors were also screened for activity against the other human DNA ligases, revealing compounds that were also active against the other DNA ligases. These were placed in three groups; DNA ligase I selective; active against DNA ligase III as well as DNA ligase I; active against all three human DNA ligases. Representatives of each group, L82, L67 and L189, respectively were characterized further (Chen, Zhong et al. 2008). Interestingly L67 and L189 were highly cytotoxic, whereas L82 was not (Chen, Zhong et al. 2008, Ricci, Tedeschi et al. 2009). This high toxicity of L67 has recently been shown to be the result of inhibiting mitochondrial DNA ligase IIIα (Sallmyr, 2016). A separate group has described a derivative of L189, named SCR7, that appears

may be more selective for DNA ligase IV (LigIV) (Srivastava, Nambiar et al. 2012). However, recent publications have contradicted both the reported structure and selectivity of SCR7 (Greco, Conrad et al. 2016, Greco, Matsumoto et al. 2016).

DNA Ligases as Biomarkers of Abnormal DNA Repair

DNA Ligase I

DNA ligases perform essential functions in DNA replication, recombination, and repair. Therefore, increasing or decreasing the presence or enzymatic activity of any of the ligase proteins could have a sizable impact on these transactions. As mentioned above, there has only ever been a single DNA ligase I-deficient individual described. Mouse studies have shown that while *LIG1* null mice develop until about mid-way through embryogenesis before perishing, *LIG1* null cell lines can be established from these embryos (Bentley, Harrison et al. 2002). From this information, we can conclude DNA ligase I is essential for embryogenesis, but that one of the other DNA ligases can substitute for DNA ligase I, permitting cell viability. Indeed, it may be possible for steady state expression levels of DNA ligase I to fall significantly, without a noticeable phenotype. However, ligase I deficiency, as observed in 46BR cells, has been shown to trigger ubiquitylation of PCNA at lysine 107 (Das-Bradoo, Nguyen et al.

2010). Overexpression of LigI has been linked with genomic instability and decreased slipped-DNA repair, (Srivastava, Nambiar et al. 2012) it has also been observed that LigI is elevated in cancer cell lines (Sun, Urrabaz et al. 2001).

DNA Ligase III

The gene that encodes DNA ligase III (*LIG3*) is located on chromosome 17 (Chen, Tomkinson et al. 1995). The human *LIG3* gene encodes three different isoforms of DNA ligase III, nuclear and mitochondrial forms of LigIIIα and a form generated by alternative splicing, LigIIIβ (**Fig. 1**). Patients with chronic myelogenous leukemia (CML) may express the oncogene BCR-ABL, a fusion gene that is the result of the characteristic Philadelphia translocation, t(9;22)(q34;q11) (Hagemeijer, Bootsma et al. 1982). However, the presence of the Philadelphia chromosome is insufficient to successfully diagnose CML, as these features may also be present in acute lymphoblastic leukemia (Hermans, Heisterkamp et al. 1987) as well as acute myeloid leukemia (Paietta, Racevskis et al. 1998). Recent studies of DNA repair mechanisms in CML have shown that highly error-prone methods of double-strand break repair may contribute to disease progression (Sallmyr, Tomkinson et al. 2008). Double-strand DNA breaks are widely considered to be the most potentially dangerous form of DNA damage. The two major mechanisms of double strand break repair are homologous recombination (HR), which is the predominant mechanism during G2 and late S-phase, and non-homologous end-joining (NHEJ), which acts

primarily during G0, G1, and early S-phase. HR makes use of the undamaged sister chromatid present in G2 and late S-phase as a template for repair, and as such, is the ideal and least error prone mechanism of double-strand break repair. NHEJ does not have the luxury of a sister chromatid, and is therefore more error prone than HR. This can result in errors such as small insertions or deletions at the break site but may also join DNA ends that were previously not associated with each other, resulting in large chromosomal rearrangements. In terms of cell survival, this is still largely preferable to leaving double-strand breaks unrepaired (Rassool and Tomkinson 2010).

The major HR and NHEJ pathways are not the only mechanisms for repairing a double-strand break. Under certain circumstances, cells will resort to alternative-NHEJ pathways, which are far more error prone and frequently generate large deletions or chromosomal translocations (Nussenzweig and Nussenzweig 2007). CML patients have been observed to express higher than normal levels of the gene WRN (8p12) as well as DNA ligase IIIα, which is the major DNA ligase involved in Alt-NHEJ. WRN mutation is normally associated with Werner syndrome, a disease characterized by premature aging. While CML cells show elevated levels of DNA ligase IIIα, two other DSB repair proteins, Artemis and DNA ligase IV, are downregulated (Sallmyr, Tomkinson et al. 2008). Artemis has been clearly shown to be involved with both HR and NHEJ, but not Alt-NHEJ, (Rassool and Tomkinson 2010, De Ioannes, Malu et al. 2012) resulting

in hypersensitivity to ionizing radiation in Artemis-deficient cell lines (De Ioannes, Malu et al. 2012). The extent of these changes is increased in BCR-ABL expressing CML cells that have acquired resistance to imatinib (Tobin, Robert et al. 2013)**.** The consequence of the changes in steady-state expression levels of these proteins is that higher fidelity pathways of DNA double-strand break repair are downregulated, and Alt-NHEJ becomes a major means of sealing DSB's, resulting in an increase in genomic mutations that may serve to drive disease progression (Sallmyr, Tomkinson et al. 2008). Similar alterations, elevated DNA ligase IIIα and decreased DNA ligase IV, have been observed in breast cancers that are both estrogen receptor (ER) and progesterone receptor (PR) negative, as well as cancers that have developed resistance after long term exposure to either tamoxifen or aromatase inhibitors (Tobin, Robert et al. 2012).

Increased alt NHEJ activity in breast cancer and CML cells was also associated with elevated levels of Poly(ADP-ribose) Polymerase (PARP), another participant in Alt-NHEJ (Rassool and Tomkinson 2010). These results seem to indicate that the increased expression of proteins associated with Alt-NHEJ correlates with decreased expression of classical NHEJ proteins. This conclusion is further supported by experiments using siRNA knockdown of classical NHEJ component Ku70, which resulted in increased DNA ligase IIIα and PARP1 expression (Tobin, Robert et al. 2012). Elevated levels of PARP1 and DNA ligase IIIα, coupled with lowered DNA ligase IV and other factors

involved in the major NHEJ pathway, serve as biomarkers to identify cancer cells that rely on alternative non-homologous end-joining as their primary means of double-strand break repair (Sallmyr, Tomkinson et al. 2008, Tobin, Robert et al. 2012, Tobin, Robert et al. 2013).

DNA ligase IV

DNA ligase IV activity was first purified from HeLa cell extracts in 1996. The cDNA for DNA ligase IV was identified by the Lindahl laboratory, who showed that the *LIG4* gene is located on human chromosome 13 (Robins and Lindahl 1996). Unlike the other two human DNA ligases, the DBD of LigIV starts at the N-terminus of the DNA ligase IV polypeptide, which has two C-terminal BRCT domains, (**Fig. 1**). These two tandem BRCT domains as well as the space between these two domains has been identified as the region that interacts with the coiled-coil of X-ray repair cross-complementing protein 4 (XRCC4) (Grawunder, Zimmer et al. 1998). The yeast homolog of *LIG4* is *DNL4* (Doré, Furnham et al. 2006). While, based on the number of submissions, LigIV has the most crystal structures of the three human DNA ligases in the Protein Data Bank (PDB), (**Table 1**) most of these structures are of small LigIV fragments (usually a BRCT domain) and LigIV remains the only human DNA ligase yet to be crystallized in complex with DNA (Berman, Westbrook et al. 2000).

The DNA ligase IV/XRCC4 complex is not only a core factor in the major NHEJ pathway but is also critical for V(D)J recombination (Grawunder and Harfst

2001). In accord with these functions, human individuals with inherited mutations in the *LIG4* gene are immunodeficient and radiation sensitive (O'Driscoll, Cerosaletti et al. 2001). This condition, known as “LIG4 Syndrome,” (OMIM: 606593) (Hamosh, Scott et al. 2000) while not as rare as ligase I deficiency, is an uncommon disease. Patients can have impaired growth, as well as other physical dysmorphic features (Frank, Sharpless et al. 2000, Lee, Barnes et al. 2000, O'Driscoll, Cerosaletti et al. 2001, Girard, Kysela et al. 2004, IJspeert, Warris et al. 2013). Immunodeficiency, as well as radiation sensitivity is an unfortunate, if predictable, result of impaired NHEJ. While, as noted above, reduced levels of DNA ligase IV have been observed in cancers with increased alt NHEJ, elevated levels of DNA ligase IV occur in other cancers, recent evidence has shown that aberrant Wnt signaling in cancer results in increased LigIV expression and that the elevated levels of LigIV confer radioresistance (Jun, Jung et al. 2016).

Activity of DNA Ligase Inhibitors in Preclinical Models of Human Cancer

Inhibition of DNA ligases has a high potential for toxicity; inhibitors to NAD^+-dependent ligases are undergoing evaluation as anti-bacterial agents (Mills, Eakin et al. 2011, Stokes, Huynh et al. 2011). Moreover, in eukaryotes, the homologue of DNA ligase I is essential for survival in yeast, (Nasmyth 1977, Johnston and NASMYTH 1978) and deficiencies of various DNA ligases in plants have adverse consequences (Waterworth, Kozak et al. 2009, Waterworth,

Masnavi et al. 2010). In humans, DNA ligase III is essential for maintaining mitochondrial DNA, with deletion of *LIG3* resulting in lethality very early inembryogenesis, (Puebla-Osorio, Lacey et al. 2006, Gao, Katyal et al. 2011) and, as has been stated previously, DNA ligase I function is also necessary during embryo development (Bentley, Harrison et al. 2002). Mice lacking *LIG4* have been observed to die late in embryogenesis (Frank, Sekiguchi et al. 1998). Given the evidence that here is functional redundancy among the human DNA ligases and that, except for mitochondrial DNA ligase III, the nuclear function of any one of the DNA ligases is not required for cell viability, it seems likely that selective inhibition of single DNA ligase species will have limited effects on normal tissues and cells. The cancer predisposition of rare inherited DNA repair deficiency syndromes and the large number of mutations revealed by next generation sequencing of sporadic cancers, suggest that abnormalities in genome maintenance pathways is a common event and that these abnormalities may offer the opportunity to selectively target cancer cells. Since DNA ligation completes almost every DNA repair pathway, selective DNA ligase inhibitors may have utility in the development of therapeutic strategies that target cancer cells. This idea is supported by preclinical studies using a DNA ligase inhibitor to inhibit Alt-NHEJ in CML, breast cancer and, more recently, neuroblastoma (Newman, Lu et al. 2015).

Breast Cancer

Breast cancer is a complex, heterogeneous disease. Endocrine therapy has proven to be very effective in aiding the majority of patients afflicted with breast cancer, but it does have limitations. Roughly one quarter of all breast cancer patients express neither the estrogen receptor nor the progesterone receptor and so are unaffected by standard frontline endocrine therapies (Johnston and Dowsett 2003). Furthermore, these patients, who are usually treated with non-targeted chemotherapy agents, have poor outcomes. Studies have shown that breast cancers which are ER^-/PR^-, either naturally or due to long term anti-estrogen treatment, can be induced to (re)express both ER and PR with histone deacetylase (HDAC) inhibitors, thereby rendering them susceptible to tamoxifen and aromatase inhibitors (Chumsri, Sabnis et al. 2011, Sabnis, Goloubeva et al. 2011). However, HDACs regulate the expression of a large number of genes and there may be numerous side effects of HDAC use. One such documented effect is cardiac toxicity, and there may be others that are presently unknown. Currently, the FDA has only approved two HDAC inhibitors, for use treating advanced cutaneous T-cell lymphoma (Ververis and Karagiannis 2012).

The targeting of Alt NHEJ is a novel approach to selectively target ER^-/PR^- and triple negative breast cancers. Breast cancer cell lines that have acquired resistance to either tamoxifen or an aromatase inhibitor and ER^-/PR^- and triple

negative breast cancer cell lines have up-regulated alt NHEJ factors and down-regulated NHEJ factors, indicating that these cells are more dependent on alt NHEJ for DSB repair. Notably, these cell lines are hypersensitive to combination of PARP (ABT888) and DNA ligase III (L67) inhibitors that block alt NHEJ (Tobin, Robert et al. 2012). While biopsies are not routinely conducted in patients with tumors that have acquired resistance to endocrine therapy, this is not the case for tumors that are naturally ER^-/PR^-, Immunohistochemical analysis of ER^-/PR^- biopsies consistently revealed a decrease in Ku70, and in increase in PARP1, indicating that the alteration in DSB repair likely also occurs in tumors and that these tumors may be responsive to the repair inhibitor combination (Tobin, Robert et al. 2012).

Chronic Myeloid Leukemia

As has been stated previously, BCR-ABL1 positive cells, such as K562, exhibit elevated levels of DNA ligase IIIα and PARP1. Interestingly, both ligase IIIα and PARP1 were elevated to an even greater extent in an imatinib-resistant derivative of K562 (Tobin, Robert et al. 2013). In normal cells, and even leukemia cell lines that do not express BCR-ABL such as MO7e, derived from cells of myeloid lineage, DNA ligase IIIα is normally expressed. MO7e cells, which have been modified to express stably BCR-ABL do, however, upregulate DNA ligase IIIα (Sallmyr, Tomkinson et al. 2008). The standard treatment for Philadelphia chromosome positive CML is imatinib, a small molecule inhibitor of ABL, and it is

very effective. However, imatinib resistance does occur in these patients and it is a serious clinical problem. Novel therapeutic approaches are needed to tackle this problem (Soverini, Martinelli et al. 2005). As noted above, there is evidence that imatinib-resistant CML cell lines have in increased alt-NHEJ activity. In accord with this DSB alteration, CML cell lines with elevated levels of PARP1 and DNA ligase IIIα, are hypersensitive to the combination of a PARP inhibitor, NU1025, coupled with DNA ligase III inhibitor L67 (Tobin, Robert et al. 2013).

As with the breast cancer research, this Alt-NHEJ phenotype has also been observed in clinical samples. Bone marrow mononuclear cells from 19 CML patients that were either sensitive or resistant to Imatinib were analyzed for their protein expression levels. Just over half (10 out of 19) of these samples overexpressed both ligase IIIα and PARP1, with a further 21% overexpressing ligase IIIα but not PARP1. At present, the prognosis for patients that present with these symptoms is poor. These cells were assayed for their susceptibility to NU1025 in combination with L67, which produced some striking results. The patient samples were divided into three groups: sensitive, insensitive, and partially sensitive. All samples that had normal expression levels of both ligase IIIα or PARP1 were insensitive to the treatment, as measured by a colony formation assay. Of those that overexpressed both ligase IIIα and PARP1, 90% were sensitive to the treatment of both L67 and NU1025 (Tobin, Robert et al.

2013). This data provides promising pre-clinical evidence for the potential for DNA ligase inhibitors in the treatment of Alt-NHEJ dependent cancers.

Neuroblastoma

Neuroblastoma is the most common extracranial solid tumor in children, with the presence of *MYCN* amplification predicting a very poor prognosis (Cheng, Hiemstra et al. 1993, Attiyeh, London et al. 2005). Neuroblastomas also have chromosomal alterations, most frequent of which are 17q gain, and 1p or 11q loss of heterozygosity (Plantaz, Mohapatra et al. 1997, Attiyeh, London et al. 2005). Similar to both breast cancer and CML, neuroblastomas cell lines had decreased levels of LigIV and Artemis and elevated levels of Alt-NHEJ proteins PARP1 and LigIIIα as well as LigI (Newman, Lu et al. 2015). The neuroblastoma cell lines were hypersensitive to both a PARP inhibitor and L67 as single agents. These results suggest that increased activity of and dependence upon alt NHEJ may occur in a wide variety of cancers and that this repair pathway is a promising therapeutic target in cancers with increased alt NHEJ activity.

Reviewing the Catalog of Ligase Mutations in Cancers

The Catalogue of Somatic Mutations in Cancer (COSMIC) database maintains information about mutations identified in cancers. Data retrieved from the COSMIC database shows that, while there have been mutations in ligase encoding genes found in cancers, they are not very common.

Each human ligase gene was analyzed in just under 30,000 samples. Mutations in *LIG1, LIG3* and *LIG4* were identified at a rate of 0.69%, 0.49% and 0.58%, respectively. This represents a low rate of mutation, when compared to the rates of genes one would expect to find mutated in cancers, such as ABL1, which was mutated in 3.88% of 43,000 tested samples. p53, which is a protein mutated in many different tumors (Lehman 1974) was found to be mutated in the COSMIC database 31140 out of 119518 samples tested (26.1%). As for those mutations that have been catalogued, the DNA binding domain had the greatest percentage of mutations of the three catalytic domains (DBD, AdD and OBD), but the difference was less than 10%. In the case of LigI, the plurality of mutations (38%) were found in the N-terminal domain (Forbes, Beare et al. 2015). Thus, it appears that expression of DNA ligases is a much better metric by which to evaluate the potential involvement of DNA ligases as biomarkers, drivers and/or therapeutic targets in cancer, rather than mutational data.

Introductory Summary

The catalytic region of DNA ligase I encircles and ligates the nicks between adjacent Okazaki fragments during lagging strand DNA synthesis. Although there is compelling evidence that interactions with PCNA and RFC are critical for the specific participation of DNA ligase I in DNA replication, further studies are needed to delineate the precise molecular mechanisms by which

these interactions contribute to the coordinated processing and joining of Okazaki fragments and how these interactions are effected by phosphorylation of DNA ligase I. The predisposition to cancers exhibited by a mouse model of DNA ligase I-deficiency highlight the importance of this coordination and regulation in prevention of genome instability during DNA replication (Harrison, Ketchen et al. 2002). This dissertation has a hybrid structure; chapters two and three are papers that have been submitted for publication. In accordance with UNM guidelines, they have only minor edits, mostly for style purposes. All supplemental figures referenced in these chapters are included here as expansions on what has been submitted. In my research proposal, I put forward two specific aims for my graduate work; to identify compounds that improve on the current generation of ligase inhibitors, and to locate their binding site and mechanism of action. In the sections below, I describe significant progress towards achieving both goals

- **Specific Aim 1.** Identify determinants of structure, activity and specificity of DNA ligase inhibitors by characterizing derivatives of the DNA ligase I inhibitor, L82, and the DNA ligase I/III inhibitor L67. A number of derivatives have been created for the previously described small molecule inhibitors of DNA ligase I and DNA ligases I/III. Determine the activity and selectivity of these chemicals with the goal of identifying more selective inhibitors for DNA ligase I and DNA ligase III with activity in both biochemical and cell-based assays.

- **Sub-Aim 1a.** The lab currently has a number of compounds, intermediates between L82 and L67, purchased and synthesized. Determine if any of these have greater activity and/or specificity for DNA ligase I over DNA ligase III.
- **Sub-Aim 1b.** Upon identifying lead candidate(s), characterize them biochemically in order to determine structure-activity relationships. If no derivative compounds show improvement over parent inhibitors, use this information, as well as molecular modeling results from Specific Aim 2, to either conduct a new computer aided drug design screen, or synthesize new derivatives for testing.
- **Sub-Aim 1c.** Characterize the effects of the LigI selective inhibitor L82 and its derivative L82-G17 in cell culture. Using several cell types and complementary methods, such as colony assays, iPOND, and BrdU incorporation, elucidate the effects that L82 and L82-G17 have on cells. Compare and contrast these effects with those of L67.

- **Specific Aim 2.** Use molecular modeling approaches to predict the binding site(s) for the current pool of DNA ligase inhibitors and make specific amino acid substitutions to test these predictions. Identify residue(s) implicated as playing critical roles in inhibitor binding by computational approaches. In order to confirm the role of predicted key

residues, construct mutant versions of the DNA ligases with specific amino acid substitutions and then characterize their biochemical activity, in particular sensitivity to inhibition by DNA ligase inhibitors

- **Sub-Aim 2a.** L82 and L67 have different mechanisms of action (competitive vs uncompetitive) as well as functional target(s), we predict that they will bind to different locations within the DNA binding domain of DNA ligase I. To test this hypothesis, we will perform molecular modeling assays using the OpenEye suite of software to dock chemicals known to inhibit DNA ligase I on to the crystal structure. Cohorts of known DNA ligase inhibitors will predict the actual binding site.
- **Sub-Aim 2b.** Evaluate the predicted binding sites by determining effects of amino acid substitutions within the binding site by both *in silico* and biochemical approaches in order to identify residues that are critical for inhibitor activity. Express mutant versions of the DNA ligase that are resistant to inhibition in cell lines that lack either DNA ligase I or DNA ligase III to demonstrate that effects of DNA ligase inhibitors observed in cell-based assays are due to inhibition of the target protein rather than off-target effects.

CHAPTER 2

Characterization of an uncompetitive inhibitor of DNA Ligase I

Submitted for publication in DNA Repair.

Timothy R.L. Howes, Annahita Sallymr, Rhys Brooks, George E. Greco, Darin E. Jones, Yoshihiro Matsumoto, and Alan E. Tomkinson

Abstract

The integrity of the phosphodiester backbone of DNA is maintained by DNA ligases. In human cells, there are three genes that encode DNA ligase polypeptides with distinct but overlapping functions. A series of small molecule inhibitors of human DNA ligases were identified using a rational structure-based approach. Three of these inhibitors, L82, a DNA ligase I selective inhibitor, and L67, an inhibitor of DNA ligases I and III, and L189, an inhibitor of all three human DNA ligases, have related structures. Here we have performed an initial structure-activity analysis in order to identify determinants of activity and selectivity. Among the compounds evaluated, we identified L82-G17, a closely related derivative of L82, that exhibited increased activity against and selectivity for DNA ligase I in vitro. Notably. L82-G17 is an uncompetitive inhibitor that stabilizes complex

formation between DNA ligase I and nicked DNA. In accord with this mechanism of action and published evidence that the trapping of proteins on DNA by inhibitors correlates with increased cytotoxicity, cells expressing DNA ligase I were more sensitive to L82-G17 than isogenic *LIG1* null cells. Furthermore, the hypersensitivity of cells lacking nuclear DNA ligase IIIα, which can substitute for DNA ligase I in DNA replication, to L82-G17, provides further evidence that this compound inhibits the cellular functions of DNA ligase I. Together our studies provide a framework for the future design of DNA ligase inhibitors and describe a novel inhibitor with utility as a probe of the catalytic activity and cellular functions of DNA ligase I.

Introduction

DNA ligation is required to generate an intact lagging strand during DNA replication as well as in almost every recombination and DNA repair event. In human cells, this reaction is carried out by the DNA ligases encoded by the three human *LIG* genes (Ellenberger and Tomkinson 2008). Genetic analysis has revealed that there is considerable functional overlap among the DNA ligases encoded by the three *LIG* genes in nuclear DNA transactions (Frosina, Fortini et al. 1996, Audebert, Salles et al. 2004, Wang, Rosidi et al. 2005, Moser, Kool et al. 2007, Liang, Deng et al. 2008, Simsek, Brunet et al. 2011, Arakawa, Bednar

et al. 2012, Le Chalony, Hoffschir et al. 2012, Han, Masani et al. 2014, Oh, Harvey et al. 2014) whereas only the DNA ligase encoded by the *LIG3* gene, DNA ligase IIIα (LigIIIα), functions in mitochondrial DNA replication and repair (Lakshmipathy and Campbell 1999, Gao, Katyal et al. 2011, Simsek, Furda et al. 2011, Sallmyr, Matsumoto et al. 2016).

The steady state levels of DNA ligase I (LigI) are frequently elevated in cancer cell line and tumor samples (Sun, Urrabaz et al. 2001, Chen, Zhong et al. 2008). This presumably reflects the hyperproliferative activity of cancer cells since LigI is the predominant ligase involved in DNA replication (Lasko, Tomkinson et al. 1990, Barnes, Tomkinson et al. 1992, Levin, McKenna et al. 2000). Unexpectedly, many cancer cell lines exhibit both increased steady state levels of LigIIIα and reduced steady state levels of DNA ligase IV (LigIV), with these reciprocal changes indicative of alterations in the relative contribution of different DNA repair pathways between non-malignant and cancer cells (Chen, Zhong et al. 2008, Sallmyr, Tomkinson et al. 2008, Tobin, Robert et al. 2012, Tobin, Robert et al. 2013, Newman, Lu et al. 2015). The dysregulation of DNA ligases in cancer cells together with the involvement of these enzymes in the repair of DNA damage caused by agents used in cancer chemotherapy and radiation therapy suggests that LigI inhibitors may have utility as cancer therapeutics.

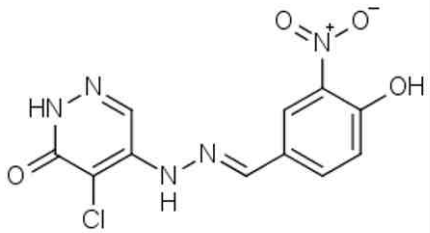
Name	Structure	Ligase Selectivity
L67		I & III
L82		I
L189		I, III & IV

Figure 3. Chemical Structures of DNA ligase inhibitors identified by computer aided drug design

Structures and selectivity of three previously described small-molecule inhibitors of human DNA ligases, L67, L82 and L189, are shown.

A set of small molecule LigI inhibitors were identified through an *in silico* structure-based screen, using the atomic resolution structure of LigI complexed with nicked DNA (Chen, Zhong et al. 2008). This screen yielded inhibitors that were selective for LigI (L82), inhibited both LigI and LigIII (L67) and inhibited all three human DNA ligases (L189) (**Fig. 3**). As expected, subtoxic levels of L67 and L82 enhanced the cytotoxicity of DNA damaging agents in cancer cell lines

(Chen, Zhong et al. 2008). Surprisingly, under similar conditions, non-malignant cell lines were not sensitized to DNA damage by the DNA ligase inhibitors, suggesting that there are alterations in genome maintenance pathways between non-malignant and cancer cells (Chen, Zhong et al. 2008). Further studies revealed that the repair of DNA double-strand breaks is abnormal in cancer cells with elevated levels of LigIIIα and PARP1 and that these cells are hypersensitive to inhibitors that target LigIIIα and PARP1 (Tobin, Robert et al. 2012, Tobin, Robert et al. 2013, Newman, Lu et al. 2015).

The DNA ligase inhibitors L82, L67 and L189 share some similarities in that they are each composed of two 6-member aromatic rings separated by different length linkers (**Fig. 3**). Here we have examined a series of related compounds in an attempt to identify determinants of activity and selectivity for LigI and LigIIIα. One of the compounds analyzed, L82-G17, is a selective, uncompetitive inhibitor of LigI. Furthermore, the activity of this compound in cell culture assays with genetically-defined cell lines indicates that it inhibits LigI function in cells.

Results

Biochemical activity of L82 derivatives

To gain insights into determinants of activity and selectivity, we either synthesized (L82-GXX) or purchased (L82-XX) compounds (**Fig. 4C**) whose structures were related to L82 but also exhibited similarities with the LigI/III inhibitor L67 and the inhibitor of all three human DNA ligases L189 (**Fig. 3**). In initial studies, we examined the effects of the L82 derivatives on LigI, LigIII and T4 DNA ligase activity. The DNA ligase IV/XRCC4 complex (LigIV/XRCC4) was not included in these assays as the purified enzyme, unlike LigI, LigIII and T4 DNA ligase, acts as a single turnover enzyme (Riballo, Woodbine et al. 2009, Chen and Tomkinson 2011). Any compounds that inhibited T4 DNA ligase, which lacks the DNA binding domain targeted in the structure-based screen (Chen, Zhong et al. 2008), were presumed to be non-specific inhibitors and excluded. The activities of the remaining compounds were compared with the previously described DNA ligase inhibitors, L82, L67 and L189 (Chen, Zhong et al. 2008) at 50 μM (**Fig. 4A and 4B**). Among the 10 compounds that inhibited LigI by at least 40%, 5 compounds preferentially inhibited LigI, 2 compounds preferentially inhibited LigIII and 3 compounds had similar activity against both LigI and LigIII (**Fig. 4B and 4C**). Among the 5 preferential inhibitors of LigI, two compounds, L82-30 and L82-G17 exhibited increased selectivity for LigI compared with L82 (**Fig. 4B**).

The linkers between the two aromatic rings of all the DNA ligase inhibitors can be grouped into three major types, vinyl, arylhydrazone, and acylhydrazone, linking the rings with 2, 3, and 4 atoms, respectively. None of the LigI selective inhibitors has a vinyl linker. In addition, all the LigI selective inhibitors except for L82-22 have a pyridazine ring whereas the LigI/LigIII and LigIII inhibitors do not. Comparing geometric shape coefficients, a value that represents a molecule's potential size based on the connections between atoms, the mean value for LigI selective inhibitors (7.4) was significantly different ($p < 0.05$) than that of LigI/III inhibitors (9.5). The mean geometric shape coefficient of LigI selective inhibitors was also significantly different than that of compounds that do not inhibit either LigI or LigIII. A further point of differentiation between LigI selective and LigI/LigIII inhibitors is their calculated partition coefficient (LogP). The calculated LogP is significantly lower for LigI selective inhibitors than inhibitors of both LigI and LigIII, 2.53 and 4.6 respectively ($p < 0.01$).

L82-G17 is an uncompetitive inhibitor of LigI

We chose to focus on characterizing L82-G17 because of its higher potency and increased selectivity. L82-G17 is more related to L82 than L67 with Tanimoto similarity scores of 78% and 32%, respectively, calculated using the Maximum Common Substructure method. The repositioning of a hydroxyl group on the non-pyridazine ring from para to meta, and the removal of a nitro group, appears to increase the selectivity of L82-G17 for LigI over LigIII. As was

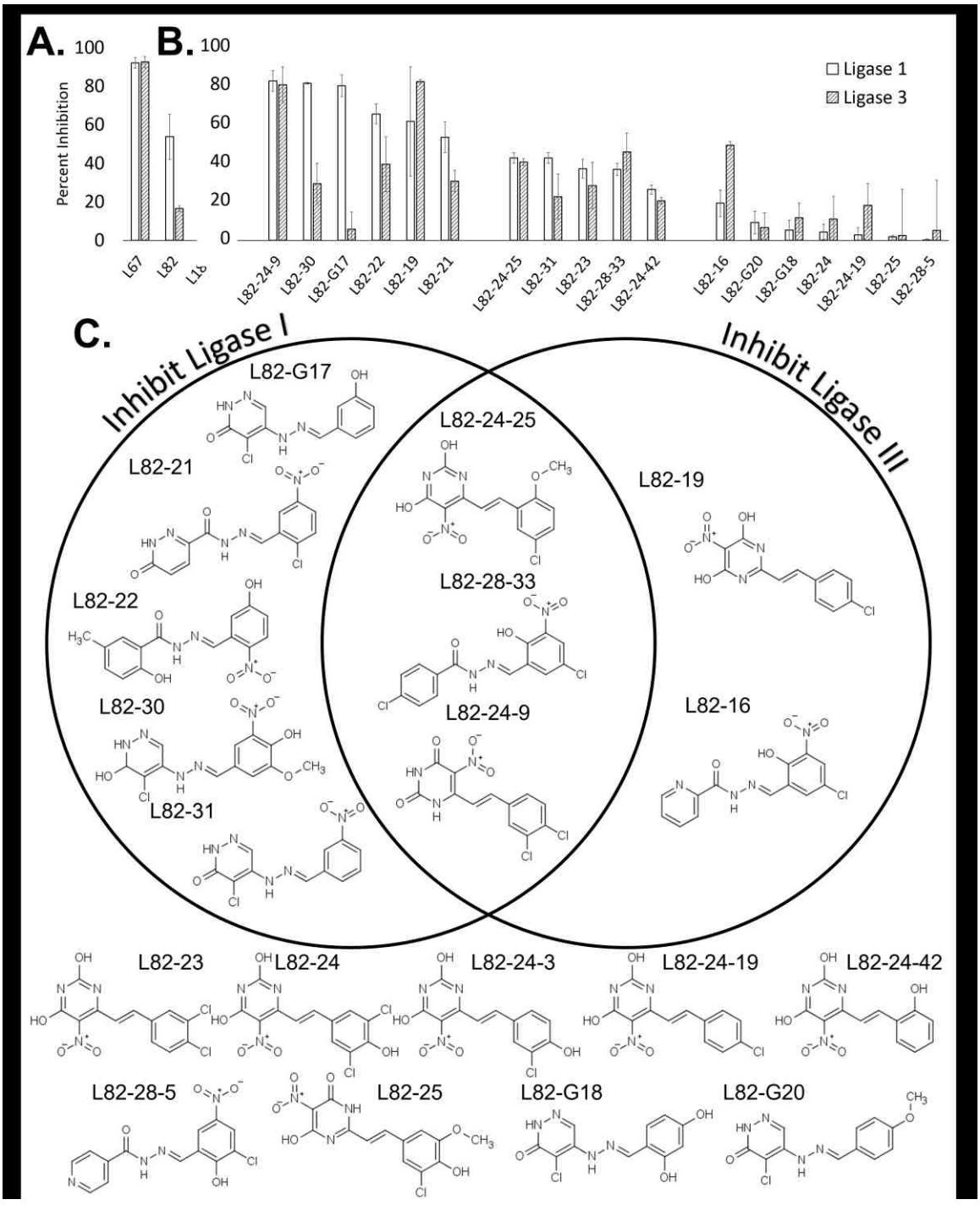


Figure 4. Activity and structures of compounds related to L67, L82 and L189

Figure 4. Activity and structures of compounds related to L67, L82 and L189.

The activity of L67, L82, L189 and L82 derivatives at 50 μM against LigI (0.625 nM) and LigIII (1.75 nM) were measured in assays with the radiolabeled DNA substrate as described in Materials and Methods. (A) L67, L82 and L189. (B) L82 derivatives. Results are shown graphically with inhibition expressed as a percentage of the values in assays with DMSO alone. Data shown is the result of at least three independent experiments. (C) Chemical structures of L82 derivatives that are grouped based on their selectivity for LigI (inhibit LigI more than LigIII by at least 20% and inhibit LigI by at least 40%), selective for LigIII (inhibit LigIII more than LigI by at least 20% and inhibit LigIII by at least 40%), have similar activity against LigI and LigIII, or have less than 40% activity against both enzymes (ungrouped).

observed with L82 (Chen, Zhong et al. 2008), L82-G17 did not inhibit LigIV at concentrations up to 200 μM (data not shown).

Among the LigI inhibitors identified by computer-aided drug design, L82 was unique in that it appeared to act as an uncompetitive rather than a competitive inhibitor (Chen, Zhong et al. 2008). This prompted us to examine the effects of L82 (**Fig. 5A**) and L82-G17 (**Fig. 5B**) on the kinetics of ligation by LigI. Under these reaction conditions, the V_{max} and K_m values for DNA ligase were 0.9 pmol ligations per min and 1.4 μM, respectively. Notably, the addition of L82 increased K_m and decreased V_{max}, whereas L82-G17 reduced both K_m and V_{max}. The Lineweaver-Burk plots obtained with L82 (**Fig. 5C**) indicate that this compound is in fact a mixed inhibitor that acts by both uncompetitive and competitive mechanisms. In contrast, the Lineweaver-Burk plots obtained with L82-G17 indicate that this compound is an uncompetitive inhibitor (**Fig. 5D**).

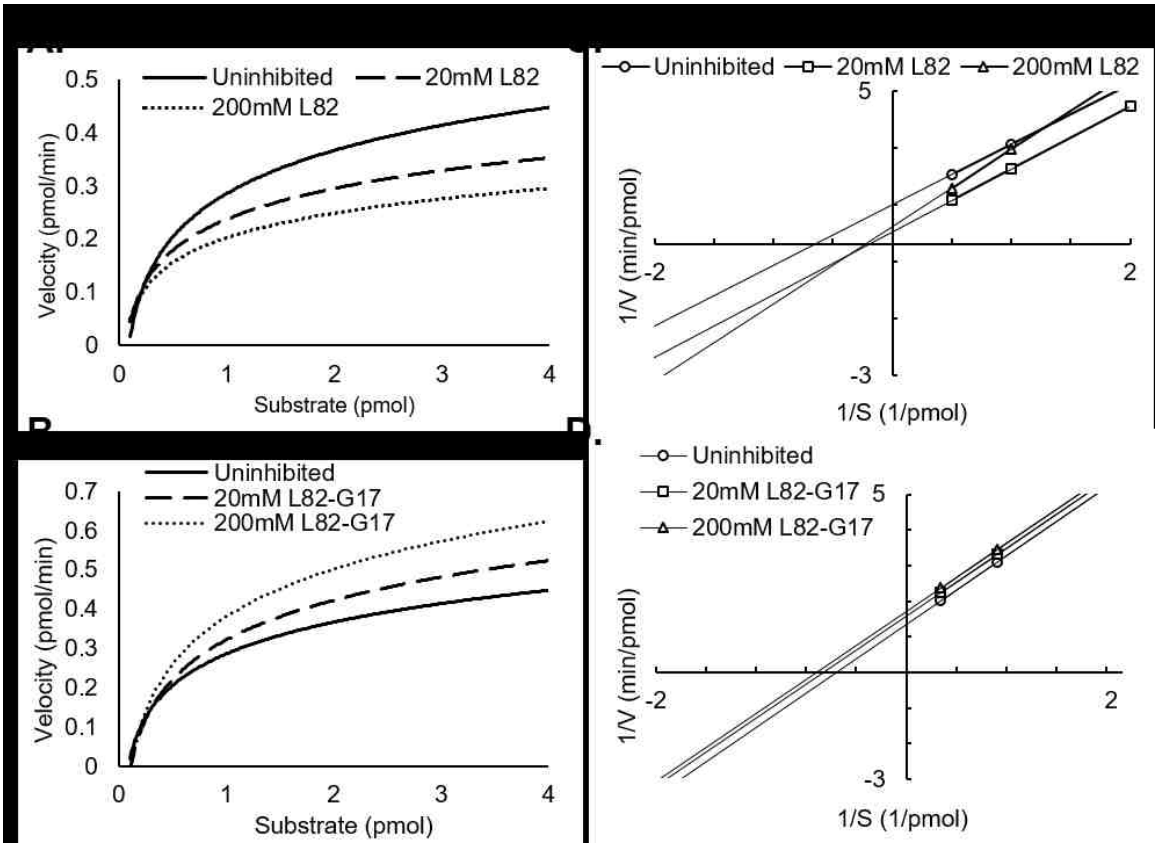


Figure 5. L82-G17 is an uncompetitive inhibitor

Michaelis-Menten saturation curve of at least three independent fluorescence-based ligation assays (Materials and Methods) depicting graphically the effect of increasing concentration of inhibitor on the affinity for substrate (μM) and velocity of the reaction (pmol/min) for compounds L82 (A) and L82-G17 (B). Lineweaver-Burk plot representation of the same data for compounds L82 (C) and L82-G17 (D), with standard error of mean and extrapolated trendline, are shown (uninhibited solid line, long dashes 20 μM, and short dashes 200 μM).

Electrophoretic mobility shift assays (EMSAs) and pulldown assays with a linear duplex containing a single non-ligatable nick were used to confirm the uncompetitive mechanism of L82 and L82-G17. As expected, L82 and, to a greater extent, L82-G17, increased the amount of LigI-DNA complex whereas the

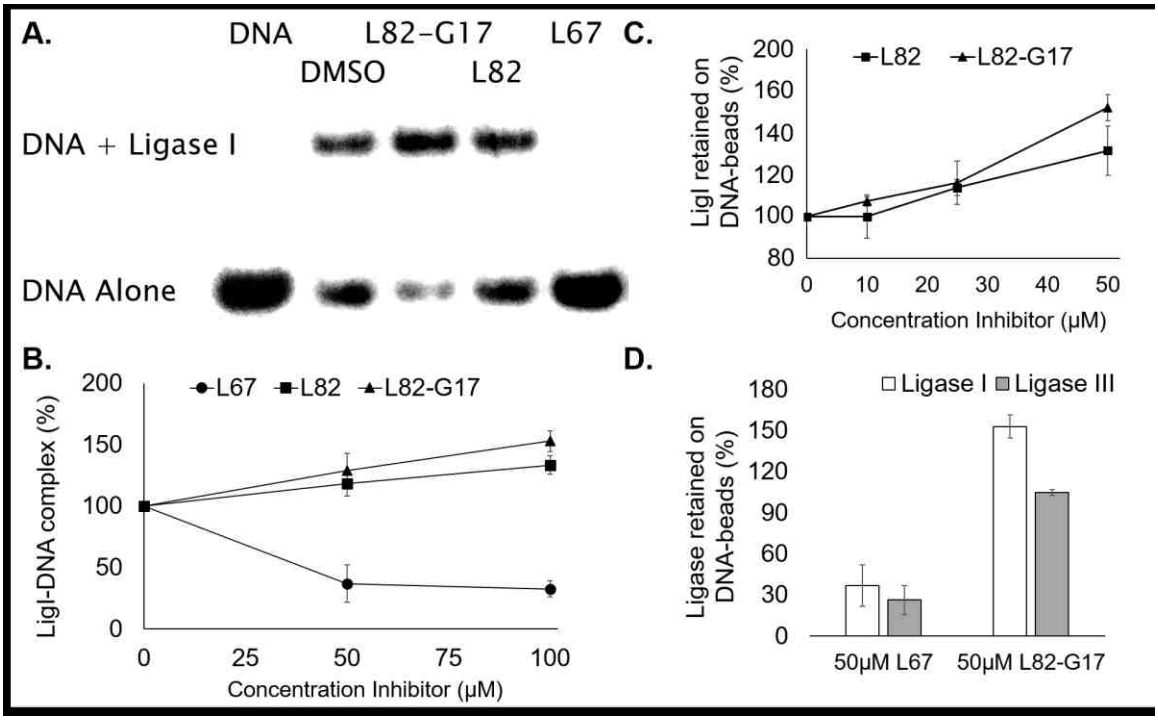


Figure 6. L82 and L82-G17 increase binding to nicked DNA whereas L67 decreases binding

(A) A representative gel showing the effect of the inhibitors, L67, L82 and L82-G17, at 100 μM and DMSO alone on DNA-protein complexes formed by LigI with nicked DNA. The position of the DNA substrate (DNA alone) and the LigI-DNA complex (DNA + Ligase I) are shown. LigI bound (upper band) and unbound (lower band) to radiolabeled DNA substrate. (B) Results of at least three independent EMSA assays are shown graphically. The effect of the inhibitors, L82 (closed squares), L82-G17 (closed triangles) and L67 (closed circles) on DNA-protein complex formation by LigI expressed as the ratio of bound and unbound DNA relative to uninhibited LigI. (C) The effect of L82 and L82-G17 on the amount of labeled LigI retained by streptavidin beads liganded by biotinylated nicked DNA. Results of three independent assays are shown graphically and expressed as a percentage LigI retained in assays with DMSO alone. (D) The effect of L67 and L82-G17 on the amount of labeled LigI and LigIII retained by streptavidin beads liganded by biotinylated nicked DNA. Results of three independent assays are shown graphically and expressed as a percentage LigI/LigIII retained compared with assays with DMSO alone.

competitive inhbitor L67 reduced the amount of LigI-DNA complex (**Fig. 6A and**

B). In pulldown assays with streptavidin beads liganded by a biotinylated linear

duplex containing a single non-ligatable nick, both L82 and L82-G17 increased the amount of labeled LigI retained by the beads in a concentration-dependent manner (**Fig. 6C**). In similar assays, 14% and 3.6% of LigI was retained by beads liganded by intact double-stranded DNA, and single-stranded DNA, respectively (data not shown) compared with the beads liganded by DNA with a single non-ligatable nick. Taken together these results demonstrate L82-G17 is an uncompetive inhibitor of LigI. To confirm the selectivity of L82-G17 for LigI, we performed pull down assays with both LigI and LigIII. As expected, the LigI/III competitive inhibitor L67 reduced the binding of both LigI and LigIII to the beads. In contrast, L82-G17 increased the retention of LigI by the beads but had vey little effect on LigIII binding (**Fig. 6D**).

Effects of L82-G17 on replicative DNA synthesis and cell proliferation

Since LigI is the DNA ligase predominantly responsible for joining Okazaki fragments during replicative DNA synthesis (Barnes, Tomkinson et al. 1992, Levin, McKenna et al. 2000, Bentley, Harrison et al. 2002), we examined the effect of L82-G17 and L82 on the incorporation of bromodeoxyuridine (BrdU) by an asynchronously proliferating popoulation of HeLa cells. While both L82 and L82-G17 reduced BrdU incorporation, their effects were modest compared with the LigI/III inhbitor L67 (**Fig. 7A**). This is consistent with studies showing that LigI is not essential for mammalian DNA replication becase of the ability of LigIII to act as a back-up (Bentley, Selfridge et al. 1996, Bentley, Harrison et al. 2002, Le

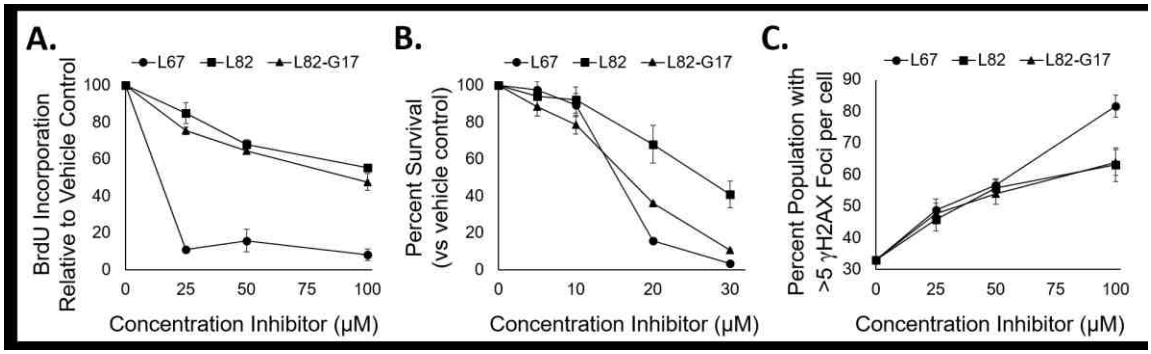


Figure 7. Effects of ligase inhibitors DNA synthesis, cell viability and DNA damage

The effects of L67, (closed circles) L82, (closed squares) or L82-G17 (closed triangles) on BrdU incorporation, cell viability and formation of γH2AX foci were determined as described in Materials and Methods. (A) Asynchronous HeLa cells were treated with the ligase inhibitors for 4 hours and were then assayed for BrdU incorporation. Results of three independent assays are shown graphically. (B) HeLa cells were incubated with the inhibitors for 5 days prior to the determination of cell viability using the MTT assay. Results of three independent assays are shown graphically. Data points at 20 and 30 μM are significant at $p < 0.005$. (C) HeLa cells were incubated with the ligase inhibitors for 4 hours prior to the detection of γH2AX foci by immunocytochemistry. Cells that contained at least 5 foci were counted as γH2AX positive. At least 100 cells counted per data point. The results of 3 independent assays shown.

Chalony, Hoffschir et al. 2012, Han, Masani et al. 2014). Since the DNA synthesis assay involved a relatively short incubation with the DNA ligase inhibitors, we examined their impact on cell proliferation over a 72 h time period using the MTT assay that quantitates the activity of NAD(P)H-dependent cellular oxidoreductase enzymes, an indicator of metabolic activity that correlates with the number of viable cells. L82-G17 was a more effective inhibitor of proliferation, reducing cell number by about 70% at 20 μM compared with a 30% reduction with 20 μM L82. Similar results were obtained with L82 and L82-G17 using a

CYQUANT assay that measures genomic DNA, another indicator of cell number (Jones, Gray et al. 2001) (data not shown).

Since L82 and L82-G17 had more severe effects on proliferation (**Fig. 7B**) than replicative DNA synthesis (**Fig. 7A**), we asked whether these compounds induced DNA damage that could activate cell cycle checkpoints, thereby reducing cell proliferation. Acute exposure to either L82 or L82-G17 for 4h, resulted in a concentration-dependent increase in formation of γH2AX foci, an indicator of DNA double-strand breaks (DSB)s (**Fig. 7C**) (Bonner, Redon et al. 2008). In accord with the cell proliferation (**Fig. 7B**) and DNA synthesis (**Fig. 7A**) assays, L67 was more effective at inducing γH2AX foci than either L82 or L82-G17 (**Fig. 7C**).

Cells lacking LigI are less suseptible to L82 and L82-G17

While the effects of L82-G17 and L82 on proliferation, BrdU incorporation and DNA damage are consistent with these inhibitors impacting DNA replication by inhibiting LigI, it is possible that they may be due to off-target effects. This prompted us to compare the effects of L82 and L82-G17 on the parental and LigI null derivatives of the mouse B cell line, CH12F3 (**Fig. 8A**) (Nakamura, Kondo et al. 1996, Han, Masani et al. 2014). Notably, both L82 and L82-G17 had a greater effect on the proliferation (**Fig. 8B**) and the survival (**Fig. 8C and 8D**) of the parental CH12F3 cells compared with the LigI null derivative, suggesting that, in

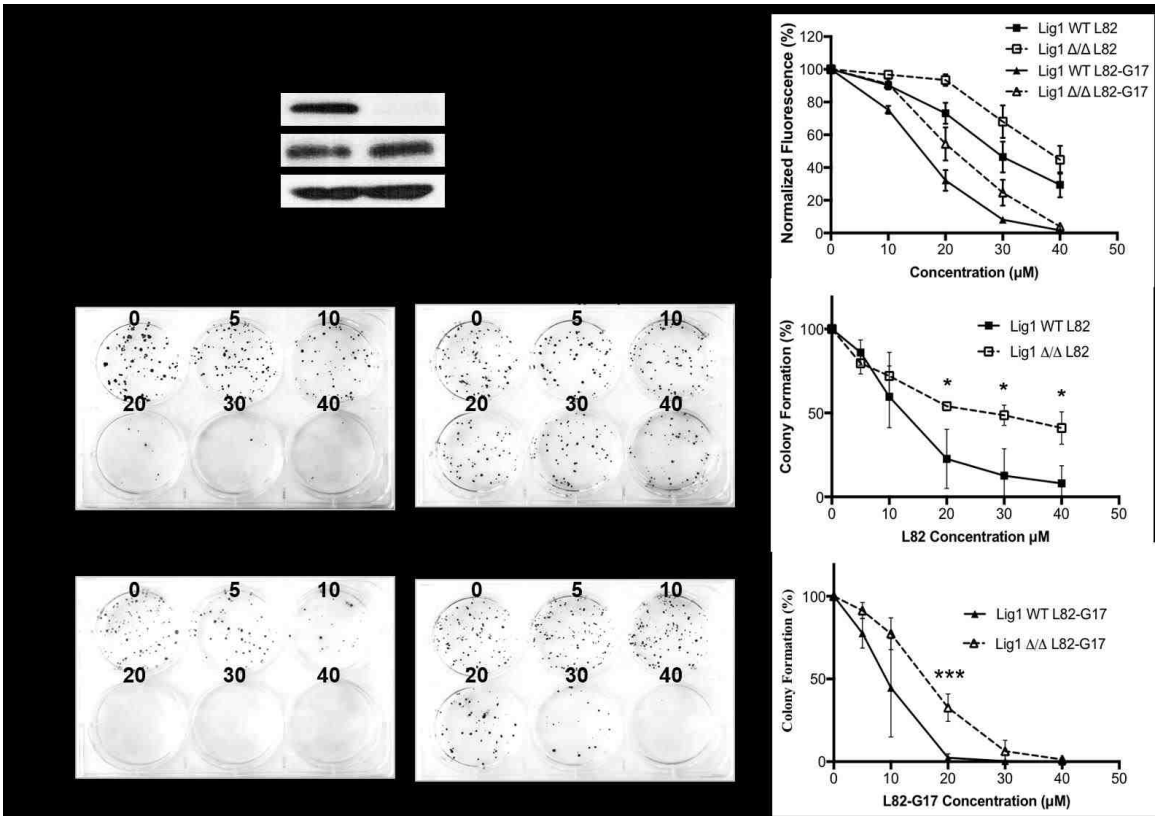


Figure 8. Cells lacking LigI are more sensitive to L82 and L82-G17

(A) LigI, LigIII and β-actin proteins were detected in extracts of CH12F3 WT and CH12F3 Δ/Δ cells by immunoblotting. (B) Effect of L82 (square) and L82-G17 (triangle) on the proliferation of CH12F3 WT (filled symbols) and CH12F3 Δ/Δ (empty symbols) cells was measured by the CyQUANT assay as described in Materials and Methods. Effect of L82 (C) and L82-G17 (D), colony formation by CH12F3 WT and CH12F3 Δ/Δ cells. Data shown graphically are the mean ±SEM of three independent experiments and are expressed as a percentage of the values for the untreated cells. * $p < 0.05$ and *** $p < 0.001$ using the unpaired two-tailed Student test.

the presence of LigI, these uncompetitive inhibitors cause DNA-protein adducts by trapping LigI on chromosomal DNA. Additional data is provided in **Figure S2**.

The absence of nuclear LigIIIα increases sensitivity to L82 and L82-G17.

Cells with both nuclear LigI and LigIIIα should be more resistant to a selective LigI inhibitor compared with cells with only nuclear LigI because of the functional redundancy between LigI and LigIIIα in DNA replication. To confirm this, we examined the effects of L82 and L82-G17 on the proliferation of a derivative of the human colorectal cancer cell line HCT116 that lacks nuclear LigIIIα and its parental cell line (**Fig. 9A**). The absence of nuclear LigIIIα markedly increased the inhibitory activity of L82-G17 (**Fig. 9C**) on cell proliferation whereas L82 had a smaller effect (**Fig. 9B**).Taken together, our results indicate that the activities of L82 and, in particular, L82-G17 in cell culture assays are due primarily to inhibition of LigI.

Discussion

There is emerging interest in the use of DNA repair inhibitors to exploit either cancer cell-specific alterations in or increased dependence upon genome maintenance pathways (Jackson and Helleday 2016). The abnormal expression of DNA ligases in cancer cell lines and samples from cancer patients suggest that DNA ligase inhibitors may have utility as anti-cancer agents either alone or in combination with DNA damaging agents (Sun, Urrabaz et al. 2001, Chen, Zhong et al. 2008, Tobin, Robert et al. 2012, Tobin, Robert et al. 2013, Newman, Lu et

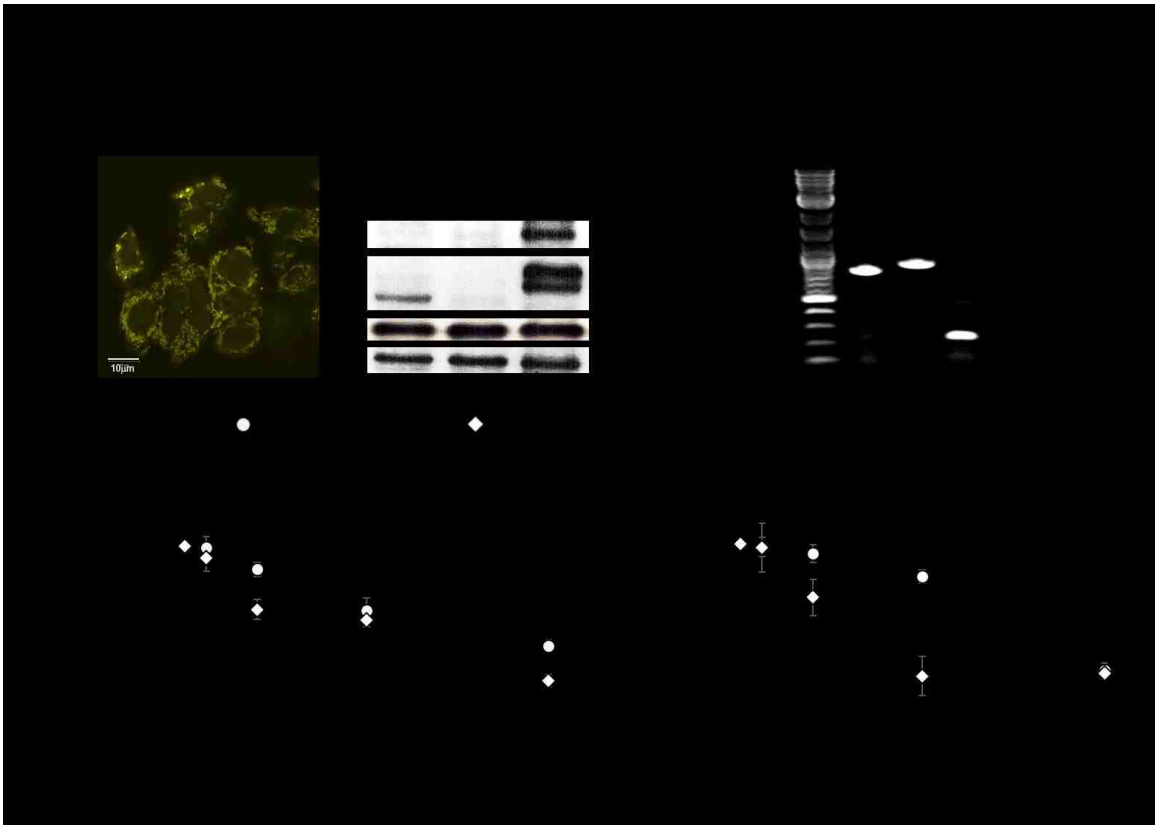


Figure 9. Cells lacking nuclear LigIII□ are more sensitive to L82 and L82-G17

(A) Fluorescent image of HCT116 Flox$^{-/-}$ expressing YFP tagged mito LigIIIα cells. Scale bars, 10 μm (left panel). Immunoblots with extracts of HCT116, HCT116 Flox+/- and HCT116 Flox$^{-/-}$ Mito Ligase 3α using antibodies against GFP, LigIII, LigI and β-actin. The positions of YFP-fusion protein (GFP), mito LigIIIα fused to GFP (Mito Ligase IIIα) and endogenous LigIIIα (Ligase IIIα), LigI (Ligase I) and β-actin are indicated (middle panel). Wild type and Floxed *LIG3* alleles and the integrated cDNA encoding YFP-tagged Mito LigIIIα were detected in genomic DNA from HCT116, HCT116 Flox+/- and HCT116 Flox$^{-/-}$ Mito Ligase IIIα by PCR as described in Materials and Methods (right panel). Proliferation of HCT116 cells (circles) and a derivative lacking nuclear LigIIIα (diamonds) incubated with (B) L82 or (C) L82-G17 for 5 days was measured by CyQUANT as described in Materials and Methods. Results of three independent assays are shown graphically.

al. 2015). Following the determination of the atomic resolution structure of LigI

complexed with nicked DNA, small molecule inhibitors with differing activities

against three human DNA ligases were identified by computer-aided drug design (Chen, Zhong et al. 2008, Zhong, Chen et al. 2008). The DNA ligase inhibitors were separated into three groups, LigI selective, LigI/III selective and inhibitors of all three human DNA ligases with L82, L67 and L189, respectively, serving as the representative compound for each of the groups. Here we examined the activity of compounds that are related to L82, L67 and L189 to gain insights into determinants of activity and selectivity. Each of the initial compounds, which differed in terms of their inhibitory activity against the three human DNA ligases, contained two aromatic rings but had linkers that differed in length and chemical composition (**Fig. 3**). Tested compounds were divided into groups by structure, vinyl (**Fig. 8A**), arylhydrazone (**Fig. 10B**), and acylhydrazone (**Fig, 10C**). The majority of LigI selective inhibitors had a 3-atom arylhydrazone linker (**Fig. 10B**). The greatest difference between active and inactive arylhydrazone class inhibitors occurs at the meta positions (8 and 10, **Fig. 10**), which must contain at least one polar group, such as the phenol in L82-G17. This observation also holds true for LigI selective acylhydrazone class inhibitors L82-21 and L82-22, which have either a nitro or a hydroxyl group at their position 10. Arylhydrazone linked inhibitors have the greatest potential for future development of LigI specific inhibitors, as no chemical in this category has activity against LigIII. However, the presence of the hydrazone linkage may hinder drug development with this scaffold as this functional group has been shown to be promiscuous and a metal

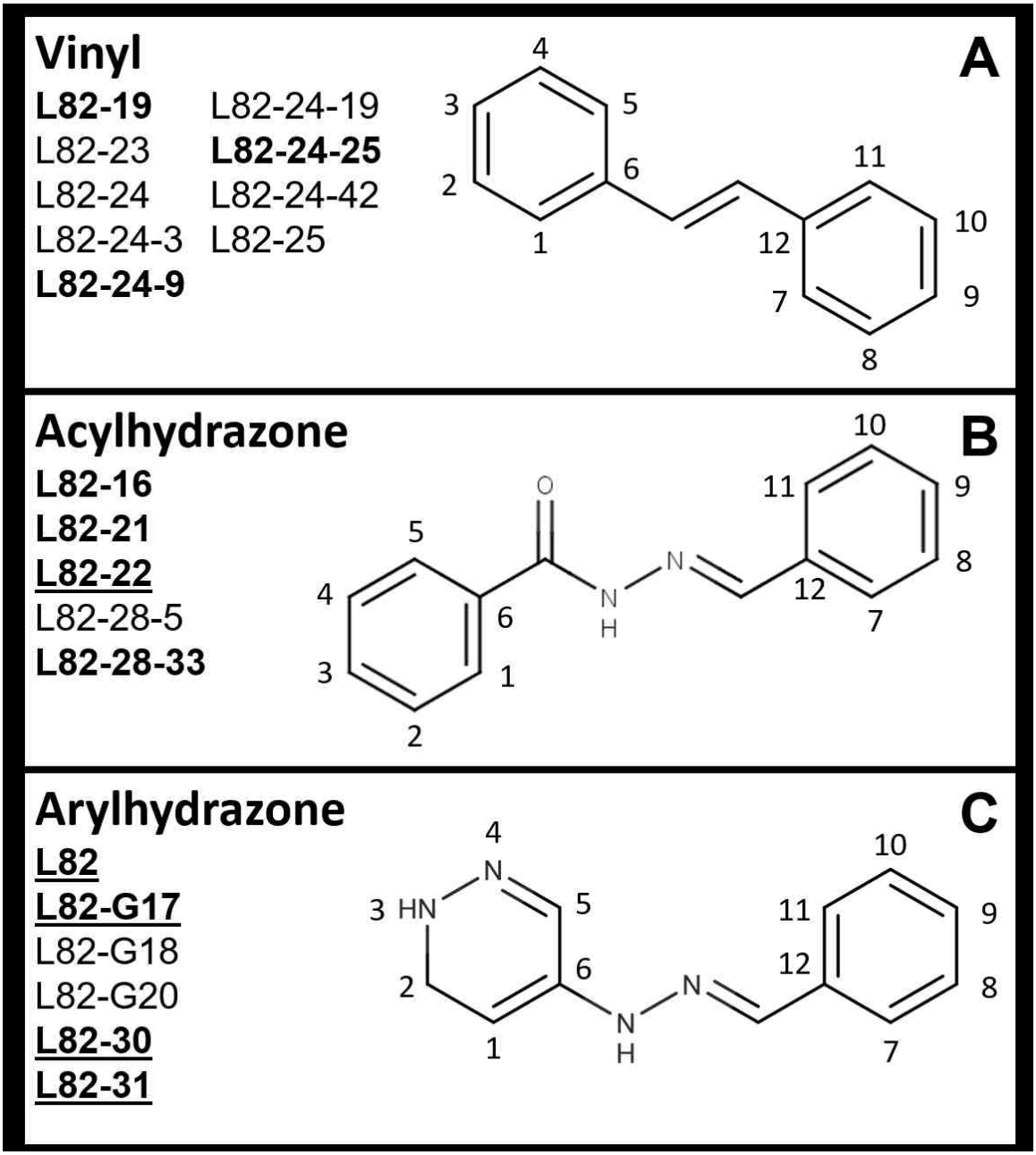


Figure 10. Grouping of active L82 derivatives based their chemical similarity

All L82 derivatives in this study fall into three structural groups: (A) vinyl linked, (B) hydrazide linked, (C) and hydrazone linked inhibitors. The members of each group are identified here. Compounds that inhibit either LigI or LigIII are shown in **bold** text, LigI specific inhibitors are also underlined.

scavenger which can lead to toxicity issues (Charkoudian, Pham et al. 2006, Baell and Holloway 2010, Peng, Tang et al. 2011).

Among the compounds related to L82, L67 and L189, we identified one compound, L82-G17 that exhibited increased activity against and increased selectivity for LigI compared with L82. Notably, it has the lowest molecular weight of all hydrazine inhibitors with activity against LigI, suggesting that it may represent the minimal requirements for this type of inhibitor. Further analysis revealed that L82-G17 is an uncompetitive inhibitor that stabilizes the LigI-nicked DNA reaction intermediate whereas L82 appears to act by both competitive and uncompetitive mechanisms. Thus, L82-G17 acts by the same mechanism as topoisomerase inhibitors, such as camptothecin, and a subset of PARP inhibitors that trap topo I-DNA and PARP1-DNA complexes, respectively (Staker, Hjerrild et al. 2002, Pommier, O'Connor et al. 2016). The increased cytotoxicity of L82-G17 compared with L82 is consistent with the studies showing that PARP inhibitors that trap PARP1-DNA complexes are more cytotoxic than PARP inhibitors that do not (Murai, Huang et al. 2012, Pommier, O'Connor et al. 2016). It is presumed that DNA-protein complexes, even when non-covalent, cause problems because of collisions with either the DNA replication or transcription machinery. In contrast to the trapped DNA-protein complexes formed by topoisomerases and PARPs, the replication machinery is unlikely to encounter trapped LigI-DNA protein complexes as these will predominantly formed behind the fork on the lagging strand.

Since the initial identification of DNA ligase inhibitors by a structure-based approach (Chen, Zhong et al. 2008, Zhong, Chen et al. 2008), there have been several reports describing LigI inhibitors using computer modeling and derivatives of the original DNA ligase inhibitors (Krishna, Singh et al. 2014, Shameem, Kumar et al. 2015, Pandey, Kumar et al. 2017). While these studies have shown that the inhibitors have activity against LigI in vitro, their affinity and selectivity appears to be less than L82-G17. Furthermore, there is no definitive evidence that these inhibitors target LigI in cells. Here we have shown that *LIG1* null cells are more resistant to L82-G17, presumably because there is no formation of trapped LigI-DNA complexes. Furthermore, cells that lack nuclear LigIIIα are more sensitive to L82-G17 as they lack a back-up activity for DNA replication.

The elevated levels of LigI in cancer cells (Sun, Urrabaz et al. 2001, Chen, Zhong et al. 2008) and the apparent viability of mammalian cells that lack LigI (Bentley, Selfridge et al. 1996, Bentley, Harrison et al. 2002, Han, Masani et al. 2014) suggest that LigI selective inhibitors may preferentially target cancer cells because of their high proliferation rate. Toxicity in normal cells is likely be limited because of the ability of LigIIIα to substitute for LigI in DNA replication and repair (Arakawa, Bednar et al. 2012, Le Chalony, Hoffschir et al. 2012) In the structure-activity studies described here, we have identified and characterized a novel uncompetitive inhibitor of LigI that selectively targets LigI both in *in vitro* and in

cells. Further work to develop improved uncompetitive inhibitors would be enhanced by the determination of atomic resolution structures of inhibitor-bound LigI-DNA complexes (Pascal, O'Brien et al. 2004). Alternatively, the predicted binding site for the LigI inhibitors could be validated using site-directed mutagenesis to generate versions of LigI that retain wild type catalytic activity but are resistant to the inhibitors.

Materials & Methods

Chemicals

L67 (IUPAC: 2-[(3,5-dibromo-4-methylphenyl)amino]-N'-[(2-hydroxy-5-nitrophenyl)methylidene]acetohydrazide), L82 (IUPAC: 4-chloro-5-{2-[(4-hydroxy-3-nitrophenyl)methylidene]hydrazin-1-yl-2,3-dihydropyridazin-3-one, and L189 (6-amino-5-[(phenylmethylidene)amino]-2-sulfanylpyrimidin-4-ol) were originally purchased from ChemDiv and Specs. Milligram quantities of each of the derivatives reported herein were also purchased from ChemDiv, Specs, or ChemBridge.

Large scale synthesis of L82-G17 (4-chloro-5-{2-[(3-hydroxyphenyl)methylidene]hydrazin-1-yl}-2,3-dihydropyridazin-3-one)

4,5-Dichloropyridazin-3(2H)one (660 mg, 4.0 mmol, 1.0 equiv) was dissolved in methanol (7 mL). Hydrazine monohydrate (380 mg, 7.6 mmol, 1.9 equiv) was dissolved separately in methanol (4 mL). Both solutions were combined and the reaction was heated at reflux for 1.5 hours. Additional methanol (3 mL) was added during heating, as the precipitate of product inhibited stirring. The reaction was cooled and the product was separated from the supernatant through vacuum filtration and dried overnight, yielding 5-chloro-4-hydrazidopyridazin-3(2H)one as a light yellow-brown powder (504 mg, 3.13 mmol, 78%). ^{1}H NMR (DMSO, 400 MHz) δ 12.46 (s, 1H, N-H), 8.07 (s, 1H, aryl), 7.96 (s, 1H, N-H), 4.57 (s, 2H, NH_2). This intermediate (100 mg, 0.62 mmol, 1.0 equiv) was dissolved into ethanol (1000 mL) with heat and 3-hydroxybenzaldehyde (91 mg, 0.75 mmol, 1.2 equiv) was added as a solid. The reaction mixture was heated under reflux overnight, then concentrated to 5 mL to precipitate a light brown solid. The product was collected, washed with ethanol, and dried under vacuum, yielding 78 mg, 0.295 mmol, 48%). ^{1}H NMR (DMSO, 400 MHz) δ 12.86 (s, 1H, N-H), 10.74 (s, 1H, N-H), 9.59 (s, 1H, O-H), 8.32 (s, 1H, imine C-H), 8.31 (s, 1H, Ar) 7.23 (t, 1H, Ar), 7.12 (s, 1H, Ar), 7.11 (d, 1H, Ar), 6.80 (d, 1H, Ar).

MarvinSketch v.14.9.8.0 (ChemAxon (http://www.chemaxon.com)) was used for drawing chemical structures. Tanimoto similarity scores were calculated using the Maximum Common Substructure algorithm in online tool ChemMine (Backman, Cao et al. 2011).

Proteins

After expression in *E. coli* BL21(DE3) cells, N-terminally poly(histidine)-tagged wild type human LigI was purified by HisTrap and HiTrap Q (GE Healthcare Life Sciences) column chromatography as described (Peng, Liao et al. 2016). In addition, a version of LigI with C-terminal poly(histidine), FLAG tags and PKA site was expressed in Rosetta2 cells and then purified by HisTrap HP and HiTrap Q column (GE Healthcare Life Sciences) column chromatography. Purified LigI fractions were pooled, concentrated using a 50 kDa MWCO centrifugal filter (EMD Millipore), and then stored at -80°C LigIIIβ and the LigIV/XRCC4 complex were purified from *E. coli* and insect cells, respectively as described previously (Chen, Pascal et al. 2006, Chen, Ballin et al. 2009).

DNA Ligation Assays

Nick joining was measured using radioactive- and fluorescence-based assays (Chen, Pascal et al. 2006, Chen, Ballin et al. 2009). The concentrations of oligonucleotides (Integrated DNA Technologies) DNA were measured using absorbance at 260 nm. In radioactive assays, the olignucleotide (5’

CAAATCTCGAGATCACAGCAACTAGACCA) was 5' end labeled with polynucleotide kinase and γ-^{32}P-ATP. After removing excess ATP using a P-30 polyacrylamide spin column (Bio-Rad), the labeled oligonucleotide was annealed with a longer complementary template oligonucleotide (5' CCATCTTAAGGAGTGCAGTAGAAGTCCAGATCAACGACACTAGAGCTCTAA ACGCAAGGCAACTGCATGGACGAACACTCGCGAGTGTTAATG) and another oligonucleotide (5'-GTAATTGTGAGCGCTCACAAGCAGGTACGTCAACGGAACG) that was also complementary to the template oligonucleotide, to generate a linear duplex with a single radiolabeled ligatable nick. DNA ligases and putative inhibitors were preincubated for 10 minutes at room temperature prior to the addition of the labeled DNA substrate (500 fmol) and incubation at 25°C for 30 minutes in a final volume of 20 μl containing 25 mM Tris-HCl pH 7.5, 12.5 mM NaCl, 6 mM $MgCl_2$, 0.4 mM ATP, 0.25 mg/mL BSA, 2.5% glycerol, 0.5 mM DTT, and 0.5% dimethyl sulfoxide (DMSO). Reactions were stopped by placing reaction tubes on ice followed by the addition of formamide dye. After heating at 95°C for 5 minutes, labeled oligonucleotides were separated by electrophoresis through a 15% polyacrylamide-urea gel and then quantitated by phosphorimager analysis on a Typhoon FLA 7000, using a 650nm laser (software version 1.2). Images were analyzed using the Fuji Film MultiGauge software (version 3.0). The substrate oligos were purified in house (**Fig. S3**).

In the fluorescence-based ligation assay (Chen, Pascal et al. 2006, Chen, Ballin et al. 2009), purified LigI (500 fmol) was incubated in the presence or absence of either L82 or L82-G17 with fluorescent nicked DNA (1 pmol) for 5 min in a final volume of 20 μl containing 60 mM Tris–HCl pH 7.4, 50mM NaCl, 10 mM $MgCl_2$, 5 mM DTT, 1 mM ATP, 50 μg/ml BSA, and 4% DMSO) at 25 °C. Following incubation, reactions were further diluted to 200 μL with a 30-fold molar excess of an unlabeled oligonucleotide (5'-TAGGAGGGCTTTCCTCCTCACGACCGTCAAACGACGGTCA) identical in sequence to the ligated strand in 10 mM Tris–HCl pH 7.4, 50 mM KCl, 1 mM EDTA and 5 mM $MgCl_2$ for 5 min, and then heated to 95 °C for 5 min. After cooling to 4 °C at a rate of 2 °C/min, fluorescence at 519 nm (excitation at 495 nm) was measured immediately using the Synergy H4 microplate reader (BioTek). Data, including Michaelis-Menten and Lineweaver-Burk equations and kinetic values, were calculated through GraphPad Prism software and graphed in Microsoft Excel.

DNA binding Assays

For electrophoretic mobility shift assays, a radiolabeled 73 bp linear duplex with a single nick was generated as above except that there was a dideoxycytosine residue at the 3' terminus of the nick (Pascal, O'Brien et al. 2004, Chen, Zhong et al. 2008). Putative inhibitors, 1500 fmol DNA ligase and the 500 fmol nicked substrate were incubated in 50 mM Tris-HCl pH 7.5, 15 mM

NaCl, 1 mM DTT, 0.2 mM ATP, 0.005 mg/mL BSA, 5 mM $MgCl_2$, 8.75% glycerol and 0.5% DMSO for 30 min. After electrophoresis through a 5% non-denaturing polyacrylamide gel using 0.5x TBE pH 8.3 as running buffer, labeled oligonucleotides were detected by phosphorimager and imaged as described above.

For DNA binding assays using magnetic beads, a non-radioactively labeled version of the DNA substrate used in the ligation assay was constructed with non-phosphorylated versions of the same three oligonucleotides except that a biotin molecule was attached to the 5' end of long template oligo. Purified LigI was ^{32}P-labeled by protein kinase A (New England Biolabs). Labeled LigI (1 pmol) and the biotinylated DNA substrate (1 pmol) bound to streptavidin magnetic beads were mixed in 20 mM HEPES-KOH pH 7.5, 20 mM NaCl, 10 mM $MgCl_2$, 5 μg/mL BSA, and 4% glycerol in a final reaction volume of 40 μL prior to incubation at 25°C for 30 minutes prior to capture of the beads using a magnetic tube rack. The beads were washed 3-times with a 20 mM HEPES-KOH pH 7.5, 100 mM NaCl, and 10 mM $MgCl_2$ before to heating at 95°C for 5 minutes with SDS sample buffer. Eluates from the beads were electrophoresed through a 10% denaturing polyacrylamide gel. Labeled LigI was detected by phosphorimager analysis and imaged as described above.

Cell Lines

Human cervical (HeLa) and colorectal (HCT116) cancer cell lines were acquired from the ATCC (Manassas, VA) and grown in the recommended media. The mouse B cell lines CH12F3 Lig1 WT and CH12F3 Lig1 Δ/Δ (Nakamura, Kondo et al. 1996, Han, Masani et al. 2014) were maintained in RPMI 1640 medium supplemented with 10% fetal bovine serum (FBS), 1% penicillin/streptomycin and 55 μM of β-mercaptoethanol. HCT116 cells were maintained in McCoy's 5A supplemented with 10% FBS and 1% Pen/Strep. HCT116 LIG3Flox$^{+/-}$ cells containing a conditional *LIG3* allele and a deletion allele (Oh, Harvey et al. 2014) were further engineered to express mitochondrially-targeted human LigIII. A full length human LigIIIα cDNA that encodes the mitochondrial leader sequence was mutated so that the internal ATG start codon for nuclear LigIIIα was altered to ATC and then fused in-frame to an EYFP gene in the pCAG-YFP-neo plasmid to generate pCAG-MitoLigIII alpha-YFP-neo. After verification by structure DNA sequencing, this plasmid was transfected into HCT116Flox$^{+/-}$ cells and YFP-positive cells were selected by flow cytometry. To delete the remaining conditional *LIG3* allele, cells were infected with an adenovirus type 5 (dE1/E3) virus encoding the Cre recombinase, (Ad-CMV-Cre #1045, Vector Biolabs). After 24 h, cells were washed and then cultured in fresh medium containing 0.5 mg/ml G418. Single cells were isolated in 96-well plates using an SY3200 cell sorter. Extracts of G418-resistant YFP-

positive clones were probed by immunoblotting with antibodies to human LigIII (GeneTEX #103172), LigI (Peng, Liao et al. 2012) and GFP (Santa Cruz #8334). The $Lig3^{-/-}$ genotype was confirmed using primers Lig3 Exon 5 F1: 5'-AAA GCA ACC CTC CTG TCT TCT CCT GCA AGT-3' and Lig3 Exon 5 R1: 5'-TGG TAC CAG GGA TAG AGT CAC GGA CAA ACC AA-3'.

Bromodeoxyuridine (BrdU) incorporation

Asynchronous populations of HeLa cells were incubated with the DNA ligase inhibitors for 4 h prior to pulse labeling with BrdU (10 μM) for 45 minutes. Immunofluorescent staining of cells was performed using a BD BrdU flow Kit (BD Pharmingen) according to the manufacturer's instructions and then quantitated by flow cytometry using an LSR Fortessa in the UNMCCC Shared Flow Cytometry and High Throughput Screening Resource. Total cellular DNA was stained with 7-AAD prior to analysis by flow cytometry. Data was analyzed using FlowJo (software version 10.1).

Cell Viability, Proliferation and survival assays

Cells were cultured in 96-well plates with either ligase inhibitors, or DMSO alone, for 5 days at 37°C. Cell viability was measured using the MTT assay, in which a tetrazolium dye, 3-(4,5-dimethylthiazol-2-yl)-2,5-diphenyltetrazolium bromide, is metabolized into (E,Z)-5-(4,5-dimethylthiazol-2-yl)-1,3-diphenylformazan in the mitochondria, resulting in a colour change from yellow to

purple. After incubation with the MTT reagent (Promega) for one hour at 37°C per the manufacturer's instructions, absorbance at 570 nm was measured using a PerkinElmer Victor 3V1420 Multilabel Counter. Cell viability is expressed as a percentage of the value obtained with DMSO-treated cells.

To quantitate genomic DNA as a measure of cell proliferation, cells were plated at density of 2000 per well in a 96-well in the presence or absence of L82 and L82-G17 and cultured for 72 h. After washing with 1 x PBS, the CyQUANT NF reagent (CyQUANT NF Cell Proliferation Assay Kit, Invitrogen) was added and incubation continued for 1h at 37^{o}C according to the manufacturer's instructions. Fluorescence intensities of triplicate samples were measured with a fluorescence microplate reader using excitation at 485 +/- 10nm and fluorescence detection at 530 +/- 15 nm. Cell number is expressed as a percentage of the value obtained with DMSO-treated cells.

Colony forming assays with CH12F3 Lig1 WT and CH12F3 Lig1 Δ/Δ cells were performed in methylcellulose-based media (Cat #HSC001 R&D Systems), which was diluted 1:3 with cell medium (RPMI Medium 1640, 10% FBS, 1% penicillin/streptomycin and 55 μM of β-mercaptoethanol) for approximately 30 minutes without disturbance. Cells were counted in order to have 300 cells per well of a 6-well plate. L82 and L82-G17 were added to cell suspensions and vortexed briefly, prior to the addition of 3 mL of methylcellulose-based medium and plating. After incubation for 10 days at 37˚C in 5% CO_2, colonies were

stained overnight with 1 mL of 1 mg/mL iodonitrotetrazolium chloride per well. Colonies were counted using ImageJ Cell Counter.

Formation of γH2AX

To detect γH2AX foci by immunocytochemistry**,** HeLa cells were grown on coverslips as described above in 12-well plates. Each well was seeded with 5000 cells that were allowed to adhere for at least 8 hours prior to incubation with inhibitors or DMSO alone for 4 hours. Cells were then incubated with anti-γH2AX FITC-conjugated antibodies using a kit purchased from BD Pharmingen and 4',6-Diamidino-2-Phenylindole (DAPI) to stain nuclei. After mounting of the coverslips onto slides, cells were imaged using Zeiss AxioObserver microscope in the UNMCCC Fluorescence Microscopy, with a Hamamatsu Flash 4 sCMOS camera and a 63x 1.4 NA objective. Images were viewed and analyzed using SlideBook (version 6.0.4).

Statistical analysis

Data are expressed as mean ± SEM. For comparison of groups, we used the Student two-tailed *t* test. A level of $P < 0.05$ was regarded as statistically significant.

Acknowledgements

We thank Dr. Jennifer Gillette for the use of her plate reader and Genevieve Phillips at the UNM Fluorescence Microscopy Shared Resource for her invaluable knowledge and assistance. This work was supported by the University of New Mexico Comprehensive Cancer Center (P30 CA118100) and National Institute of Health Grants R01 GM57479 (to A.E.T.) and P01 CA92584.

CHAPTER 3

PREDICTION & VALIDATION OF INHIBITOR BINDING POCKET ON DNA LIGASE I

In preparation for publication

Timothy R.L. Howes, Rhys Brooks, Darin E. Jones, Cristian Bologa, Jeremy Yang, Yoshihiro Matsumoto, Tudor I. Oprea and Alan E. Tomkinson

Abstract

With emerging evidence that abnormalities in DNA repair underlying the genetic instability in cancer cells can be therapeutically targeted, there is a need for inhibitors specific for different DNA repair enzymes and pathways. Since DNA ligases are involved in nearly all forms of DNA repair as well as the joining of Okazaki fragments during DNA replication, inhibitors of the three human DNA ligases, which have distinct but overlapping cellular functions, are likely to have utility as versatile probes of DNA repair in non-malignant and cancer cells and the potential for development as anti-cancer therapeutics. Previously we identified inhibitors with differing selectivity for the three human DNA ligases using a structure-based approach. Here we have used an unbiased in silico

screening approach of potential binding sites within the catalytic region of DNA ligase I combined with information about the activity of candidate inhibitors from the previous screen to predict the binding sites of a DNA ligase I-selective inhibitor L82 and DNA ligase I/III selective inhibitor L67. To validate the binding site, we showed that amino acid substitutions in DNA ligase I that are predicted to prevent L82 binding, abolish its inhibitory activity. The definitive mapping of the inhibitor binding will facilitate further rational design of inhibitors with increased affinity and selectivity.

Introduction

The three human DNA ligases share a related catalytic core that is composed of a DNA binding domain (DBD), a nucleotidyl transferase domain (NTD) and an oligonucleotide/oligosaccharide binding domain (ODB) (Ellenberger and Tomkinson 2008). Atomic resolution structures of the catalytic cores of DNA ligases I and III complexed with nicked DNA revealed that these domains, which adopt an extended conformation in the absence of DNA, encircle the nicked DNA duplex with each of the domains contacting the DNA (Pascal, O'Brien et al. 2004, Cotner-Gohara, Kim et al. 2010). While only the structure of the catalytic core of DNA ligase IV has only been determined in the absence of nicked DNA (De Ioannes, Malu et al. 2012, Ochi, Gu et al. 2013), the conservation of amino acid sequence and similarities in secondary structure

suggest that the DNA ligase IV catalytic core will adopt a similar clamp structure when it ligates DNA.

The atomic resolution structure of DNA ligase I complexed with nicked DNA was used to guide a structure-based screen for DNA ligase inhibitors (Chen, Zhong et al. 2008, Zhong, Chen et al. 2008). An initial *in silico* screen examined the predicted binding of small molecules to a DNA binding pocket defined by the residues histidine 337, arginine 449 and glycine 453 within the DNA ligase I DNA binding domain (DBD) (Pascal, O'Brien et al. 2004). Out of 1.5 million compounds, 233 candidate molecules were identified and then assayed for activity in DNA joining assays with DNA ligases I, III and IV. Among the candidates, 10 were active against DNA ligase I with some of these also exhibiting activity against one or both other human DNA ligases in biochemical and cell based-assays (Chen, Zhong et al. 2008, Zhong, Chen et al. 2008, Tobin, Robert et al. 2012, Tobin, Robert et al. 2013). However, it has not been demonstrated that these inhibitors do in fact bind to the targeted DNA binding pocket. The definitive identification of the binding site would greatly enhance efforts to design inhibitors with increased affinity and selectivity.

Here we have performed an unbiased *in silico* screen with the original 233 candidate molecules against binding pockets or receptors throughout the catalytic core of DNA ligase I. Binding pockets were ranked based on the predicted binding of molecules with activity in DNA joining activity. Using this

approach, we predicted that two DNA ligase inhibitors identified in the initial screen, L82 and L67 (**Fig. 11A**), have overlapping binding sites that are adjacent to the originally targeted binding pocket (Chen, Zhong et al. 2008, Zhong, Chen et al. 2008). The predicted binding site for L82 and its closely related derivative L82-G17 (**Fig. 11A**) was confirmed by demonstrating that substitution of a key residues within the binding site rendered the modified version of DNA ligase I resistant to the inhibitor.

Results and Discussion

Ab inititio molecular modeling indicates the DNA ligase inhibitors L67 and L82 have overlapping binding sites within the LigI DBD.

To ask whether one or more of the 10 DNA ligase I inhibitors identified in the initial structure-based screen (Chen, Zhong et al. 2008, Zhong, Chen et al. 2008) are likely to bind to the targeted site, we carried out an unbiased screen of over 1400 potential binding or receptor sites throughout the entirety of DNA ligase I (PDB: 1X9N). Initially, the 233 candidate inhibitors identified in the original *in silico* screen were docked onto approximately 300 potential binding or receptor sites within the atomic resolution structure of human DNA ligase I complexed with DNA (Pascal, O'Brien et al. 2004) using the OpenEye (OEChem 2012) suite of software. We hypothesized that, while it is unlikely that all 10 active inhibitors among the 233 candidates bind at the exact same location, it is likely that several compounds will inhibit DNA ligase I activity by interacting with

similar residues. Using the known data set of true positives (TPs) and true negatives (TNs) among the initial 233 candidate compounds (Chen, Zhong et al. 2008, Zhong, Chen et al. 2008), we determined the relative ranking of inhibitors using confusion matrices to evaluate potential binding sites. Specifically, the number of the known active inhibitors among the top 10 chemicals predicted to bind to a receptor, including the binding site used in the original computer aided drug design (CADD) screen defined by residues histidine 337, arginine 447, and glycine 453 within the DBD of DNA ligase I (Chen, Zhong et al. 2008, Zhong, Chen et al. 2008), was used to rank receptors as likely binding site for the inhibitors. Notably, the binding site used in the original screen defined by His 337, Arg 439, and Gly 453 (Chen, Zhong et al. 2008, Zhong, Chen et al. 2008) had a TP rate of 10%, putting it in the bottom 61% of potential receptor sites evaluated.

The highest scoring potential inhibitor binding sites were selected and then modified as described in Materials and Methods to create new candidate binding sites prior to re-evaluating *in silico* the 233 candidate compounds for their ability to fit into the new, second generation, putative binding sites. After several rounds of iteration, the receptor sites improved considerably with the average TP rate increasing from 7% to 32% for the final putative binding pockets (**Fig. 11B**). Residues involved in predicted enzyme-inhibitor interactions were identified for each receptor site and those with the highest frequency of predicted interactions

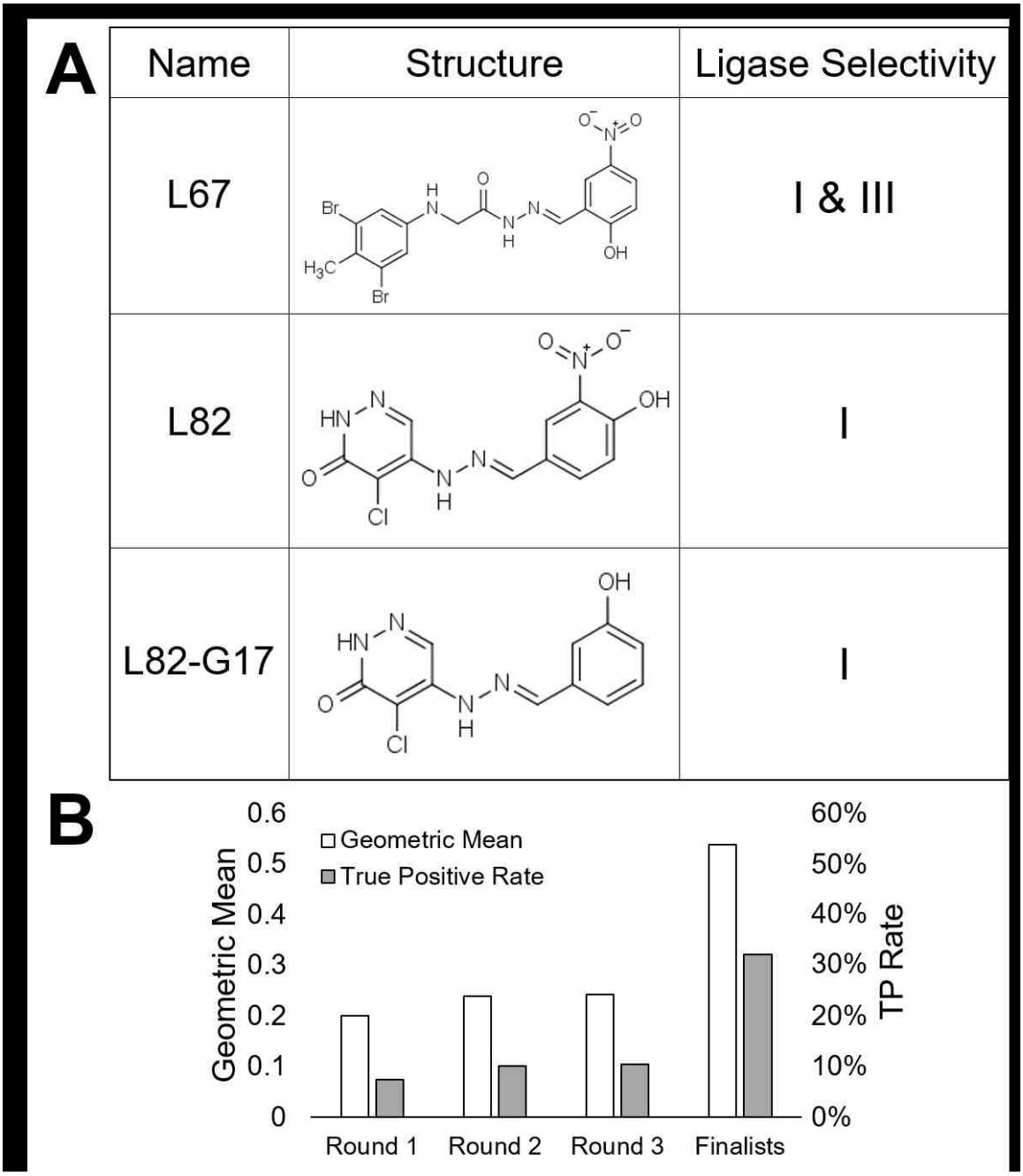


Figure 11. Chemical Structures of DNA ligase inhibitors

(A) Structures and selectivity of three previously described small-molecule inhibitors of human DNA ligases, L67, L82 and the L82 derivative, L82-G17. (B) Utilizing an iterative process, clusters of inhibitors that were predicted to bind to receptor/binding sites on DNA ligase I were identified. These were then scored using confusion matrices, the results of which are shown here to illustrate the improvement over four rounds of iteration.

were chosen for further study. All selected residues, except Lys 744 and Glu 592, were located within two regions of the DBD polypeptide, residues 336-351 and 445-448, that are in the same physical space (**Fig. 12A**). These residues are shown in green with select residues identified for clarity whereas residues involved in DNA binding, but not inhibitor binding residues are shown in pink (**Fig. 12A**). One of the two residues with the highest frequency of predicted interaction with inhibitors, His 337, formed part of the original binding site (Chen, Zhong et al. 2008, Zhong, Chen et al. 2008) whereas the other one, Gly 448, did not. Among the 10 DNA ligase I inhibitors (Chen, Zhong et al. 2008, Zhong, Chen et al. 2008), L67, a DNA ligase I/III inhibitor, and L82, a DNA ligase I selective inhibitor, are predicted to bind in each of the putative binding pockets used to generate the list of critical residues whereas L113 and L190 bind in some, but not all cases. Interestingly, there is significant overlap between the groups of residues that are predicted to interact with L82 and L67 (**Fig. 12B**) with Asn 336, His 337, Leu 338, Leu 347, and Ser 445 predicted to bind to both L67 and L82. In contrast, Gly 339, Pro 340, Pro 341, Leu 345, Phe 408, Glu 456 and Phe 507, are predicted to interact only with L67 whereas Val 349, Gly 350, Asp 351, Leu 446, Ser 447, and Gly 448 are predicted to interact only with L82 (**Fig. 12B**). In space filling models with common residues indicated in orange, L67 specific residues in purple and L82 specific residues in green, it appears that the binding sites have share a central common region (**Fig. 12C**, left panel) and that, while

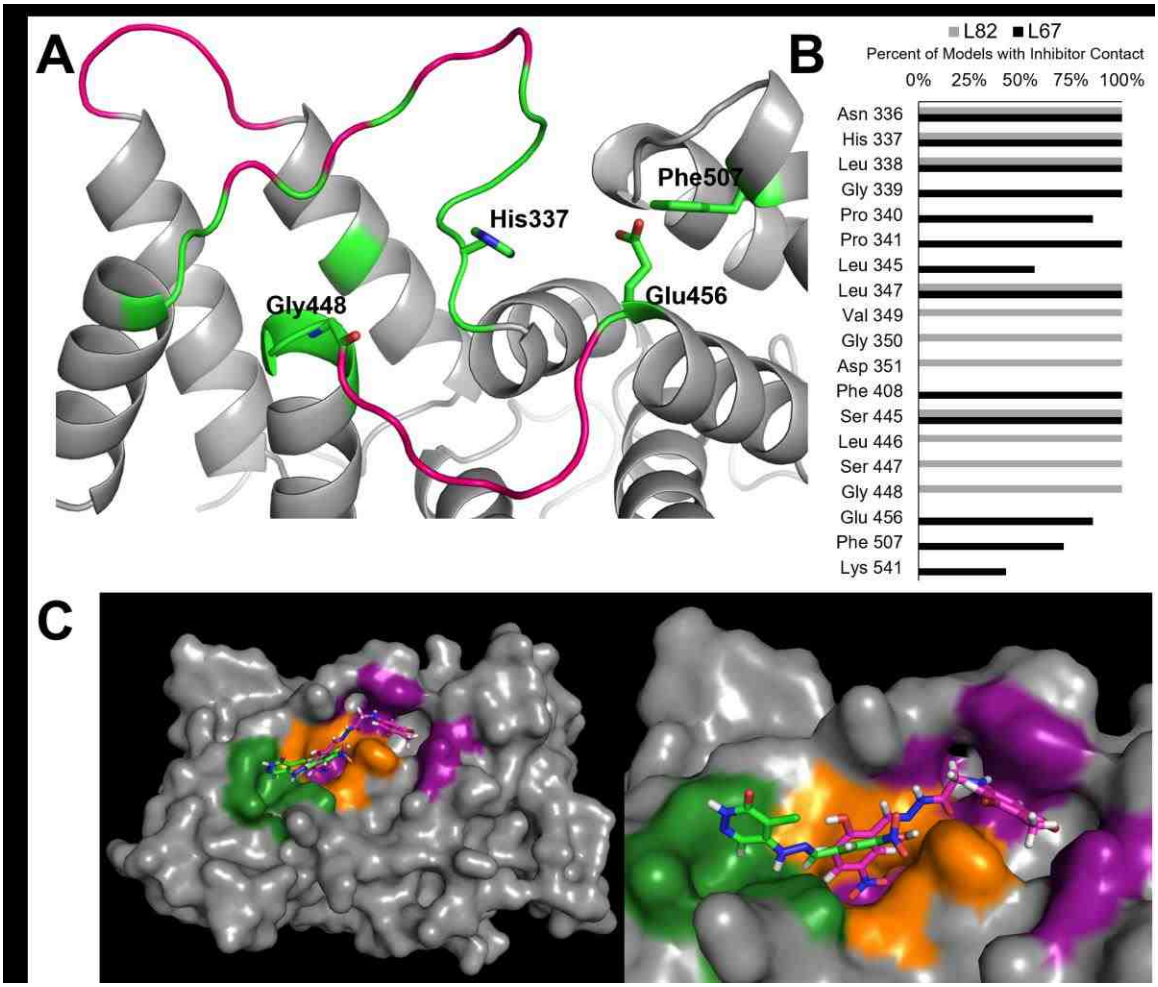


Figure 12. Putative binding pockets of L82 and L67 predicted by unbiased molecular modeling

(A) Ribbon diagram of the predicted inhibitor binding pocket of DNA ligase I (gray). Inhibitor-interacting amino acids are shown in green, while DNA-interacting amino acids are shown in pink. (B) Frequency of predicted interactions of amino acids from Asn336 to Lys541 with the inhibitors L82 and L67. (C) Space filling model of the DNA ligase I DBD showing the predicted conformations of L67 (magenta) and L82 (light green) in their overlapping binding pockets are shown. The location of the pocket relative to the rest of the DBD is shown on the left, while a magnified view of the pocket is shown on the right. DNA ligase I residues that are predicted to interact with L67 (purple), with L82 (green) or with both L67 and L82 (orange) are indicated.

the binding sites of L67 and L82 overlap in this common region, they extend out

in opposite directions, making unique contacts with different residues (**Fig. 12C**).

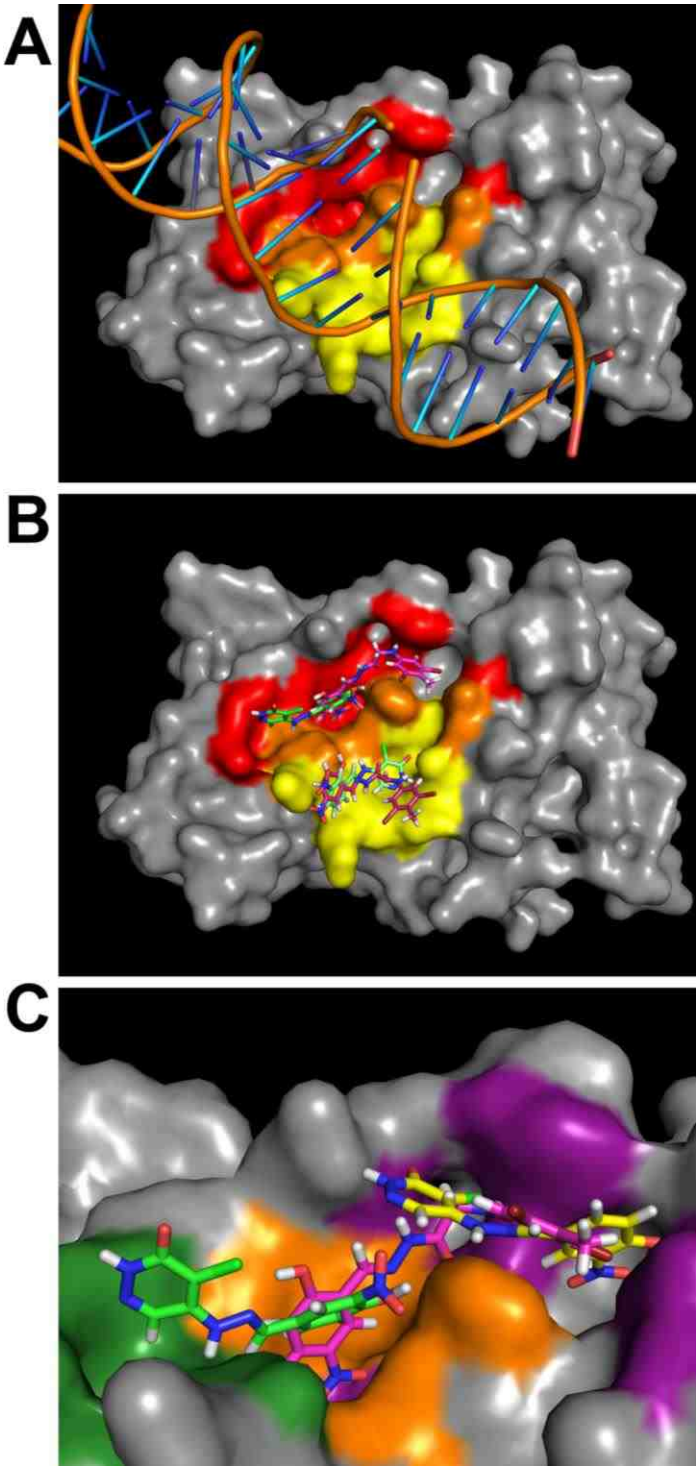


Figure 13. Comparing the new predicted binding pockets for L67 and L82 with the binding pocket targeted in the initial structure-based screen for DNA ligase inhibitors

(A) Space filling model of DNA ligase I (gray) and DNA (phosphodiester backbone, orange and nitrogenous bases, blue). Amino acids that are predicted to interact with either L67 or L82 have been highlighted based on which model they are implicated in. Residues constituting the binding pocket used in the initial structure-based screen (Chen, Zhong et al. 2008, Zhong, Chen et al. 2008) are shown in yellow, while residues constituting the new predicted binding pocket are shown in red. Resides contributing to both binding pockets are shown in orange. (B) Predicted binding locations and conformations of L67 (magenta) and L82 (green) in the two binding pockets. (C) Binding of L82 and L67 in the new predicted binding pocket in the presence and absence of DNA. DNA ligase I residues that are predicted to interact with L67 (purple), with L82 (green) or with both L67 and L82 (orange) are indicated. In the absence of DNA, the new model predicts L82 to have a new conformation (yellow), while L67's binding does not change in the absence of DNA (**Fig. S4**).

Effect of DNA on the predicted binding pockets of the DNA ligase I/III inhibitor L67 and the DNA ligase I-selective inhibitor L82.

In addition to differences in specificity, L67 and L82 also differ in their mechanism of inhibition with L67 acting as a competitive inhibitor and L82 as an uncompetitive inhibitor (Chen, Zhong et al. 2008, Zhong, Chen et al. 2008). In order to gain insights into the contribution of DNA to the predicted binding sites, additional, smaller receptor sites based on the predicted binding sites of L82 and L67 were tested. This reduced the noise from other compounds docking elsewhere. Initial studies focused on a receptor defined by Leu 335, Pro 340, and Gly 352, that, like previously tested pockets, also included the electron density contributed by the DNA that was co-crystallized with DNA ligase I (Pascal, O'Brien et al. 2004). Under these conditions, L67 and L82 were ranked as the first and second highest affinity compounds. This binding site (**Fig. 13A**, red) is immediately adjacent to the original one used for computer aided drug design (**Fig. 13A,** yellow). The predicted binding conformations of L67 (magenta) and L82 (green) in the original CADD screen pocket (yellow), as generated in this study, compared with the new pocket (red), are shown in **Figure 13B**. There is greater overlap of the binding sites in the original CADD pocket than in the new pocket (**Fig. 13B**). In accord with their different mechanisms of inhibition (Chen, Zhong et al. 2008, Zhong, Chen et al. 2008), L82 consistently contacts the DNA in the new binding pocket whereas L67 does not. The opposite is true in the original binding pocket where L67 has nearly double the number of DNA contacts

compared with L82. This is likely the result of its poor predicted binding at that location.

When the docking assay was repeated with the new receptor in the absence of DNA, the predicted binding affinity of L82 was reduced and its predicted binding location shifted to a position that overlapped to a greater extent with the predicted binding site for L67 (**Fig. 13C**). In contrast, the predicted affinity and location of L67 binding was almost entirely unaffected by the absence of DNA (**Fig. S4**). The predicted DNA-dependent and DNA-independent binding modes of L82 are consistent with more recent kinetic data indicating that L82 is a mixed inhibitor acting by both competitive and uncompetitive mechanisms (Howes 2017).

Identifying key residues predicted to be involved in interactions with DNA ligase inhibitors by *in silico* mutagenesis.

To further evaluate the role of residues predicted to be involved in the interaction with L82, the homology modeling platform SWISS-MODEL (Bordoli, Kiefer et al. 2009, Biasini, Bienert et al. 2014, Bienert, Waterhouse et al. 2017) was utilized to generate three-dimensional structures of versions of DNA ligase I with altered amino acid sequences using wild-type DNA ligase I as a template. An identical receptor was then constructed for each mutant, defined by three residues, Leu 335, Pro 340 and Gly 352, which did not move as a result of *in silico* mutagenesis. More than 200 *in silico* mutants were evaluated, of which

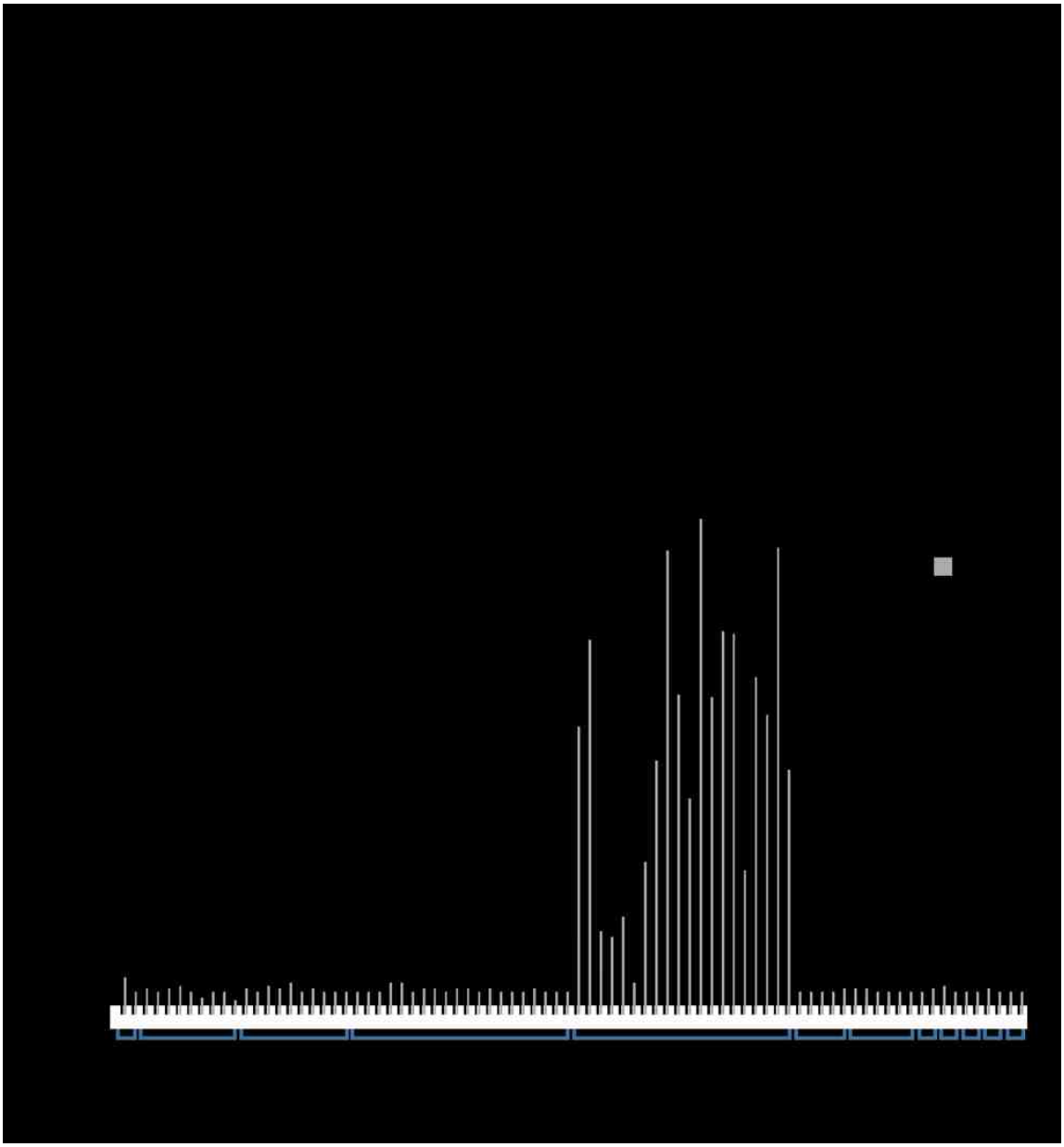


Figure 14. Utilizing in silico mutagenesis to prioritize biochemical assays

(A) Sequences of wild-type, consensus replacement, and single amino acid substitution mutants. Mutated residues have been underlined. (B) Large scale *in silico* single amino acid replacement mutagenesis shows that in this region, the only mutations predicted to effect inhibitor binding are mutations of glycine 448.

approximately 40% were single amino acid replacements. Other types of *in silico* mutations included the replacement of hairpin loops between regions of

secondary structure with consensus sequences from different DNA ligases. For example, hairpin loops from DNA ligase I (SLSG and GRLRLGLA, **Fig. 14A**) were replaced with the corresponding hairpin loop sequences from DNA ligase IV (MIIK and to KDLKLGVS, **Fig. 14A**), which is not inhibited by L82 and L67 (Chen, Zhong et al. 2008).

These amino acid replacements, which are located on the C terminal end of helix 9, as well as on the loop that connects helices 9 and 10, resulted in a decrease in the predicted binding of inhibitor L82, but no change for L67 (data not shown). In addition, the predicted effects of substituting every residue from Ser 445 to Ala 455 with at least four different amino acids were examined. Only changes of Gly 448 significantly impacted the predicted binding of L82, but not L67 (**Fig. 14B**) with replacement of Gly 448 with lysine resulted in a 10-fold decrease in the predicted binding of L82, relative to other small molecules in the known data set. The effect of this amino acid change is consistent with the models developed for the binding of L82 and L67 as the extra bulk contributed by the lysine residue occludes the L82 binding site (**Fig. 15A**), whereas this does not impact the L67 binding site which is further away (**Fig. 15B**).

The recently described derivative of L82, L82-G17, that is more selective for DNA ligase I than L82 and is an uncompetitive inhibitor (Howes et al, under review) is predicted to bind in the same location as L82 (**Fig. 15C**). The only differences between L82 and L82-G17 occur on the right-hand ring (**Fig. 11A**). In

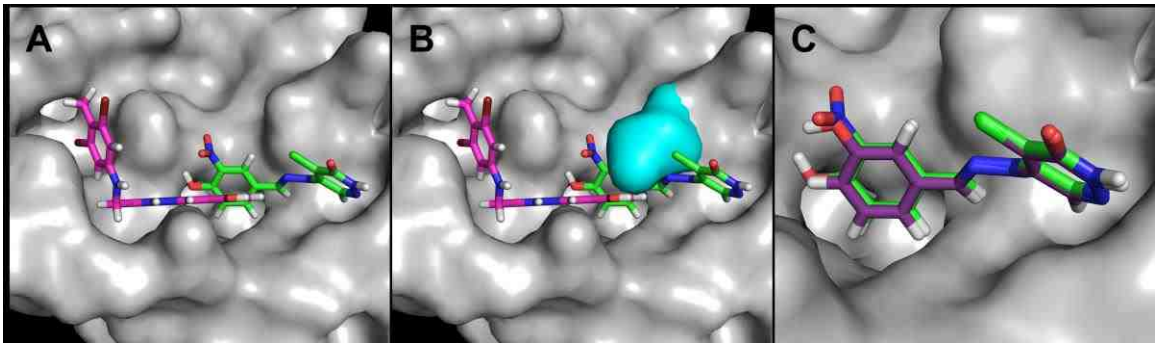


Figure 15. Replacement of glycine 448 with bulkier amino acids is predicted to block the L82 but not L67 binding site

Space filling models of DNA ligase I (grey) showing; (A) predicted binding sites of L67 (magenta) and L82 (green) in the wild type protein; (B) Effect of replacing glycine 448 with a lysine residue (cyan) on the predicted L82 and L67 binding sites. (C) Predicted binding sites of L82 and L82-G17 (purple).

L82-G17 the $-NO_2$ has been removed, and the hydroxyl group moves from a para to a meta position. While both L82 and L82-G17 are predicted to bind to DNA ligase I in the presence of DNA, L82-G17 is a predicted to bind better than L82, in accord with the activities of these compounds in biochemical assays. This mirrors the biochemical data presented in chapter 2 (Howes 2017).

Substitution of G448 with lysine or methionine confers resistance to inhibition by L82 and L82-G17.

To validate the predictions of the *in silico* modeling, we purified G448K and G448M versions of DNA ligase I. While the amino acid substitutions did not significantly alter catalytic activity, these versions of DNA ligase I were resistant to the inhibitory effects of L82 and L82-G17 whereas the inhibitory effect of L67 was not effected by either amino acid substitution (**Fig. 16A**). Since L82 and, to a

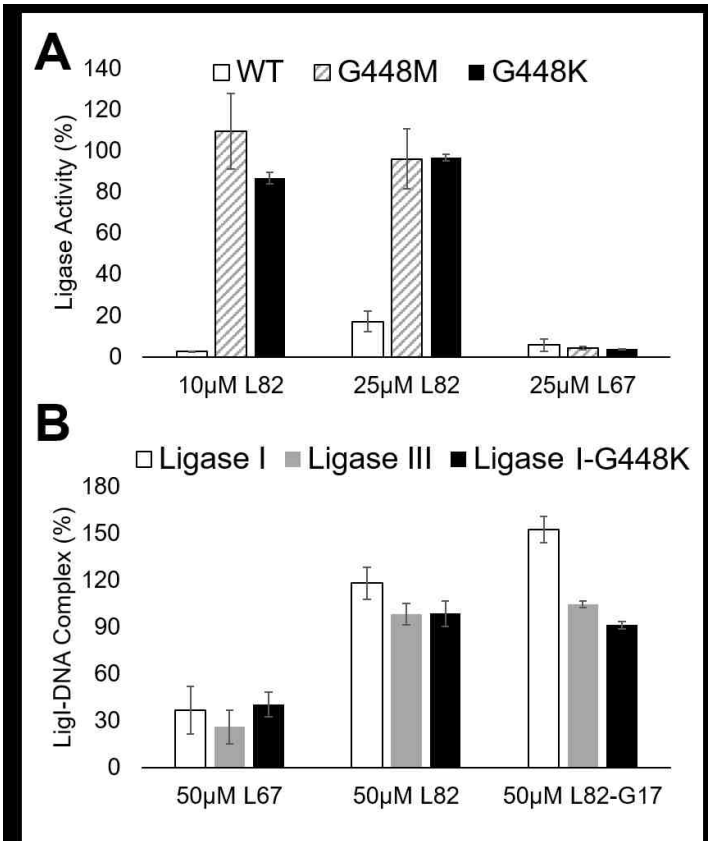


Figure 16. Substitution of glycine 448 with either methionine or lysine confers abolishes effects of L82 and L82-G17 on DNA ligase I

(A) Effect of L82 and L67 on ligation by wild type DNA ligase I and mutant versions in which glycine 448 is replaced with either methionine (G448M) or lysine (G448K). Results of three independent assays are shown graphically and expressed as a percentage of ligation in assays with DMSO alone. (B) Effect of L67, L82 and L82-G17 on the retention of labeled versions of wild type DNA ligase I, a mutant version in which glycine 448 is replaced with lysine (G448K) or DNA ligase III by streptavidin beads liganded by biotinylated DNA containing a single non-ligatable nick. Results of three independent assays are shown graphically and expressed as a percentage DNA ligase I retained in assays with DMSO alone.

greater extent, L82-G17 act as uncompetitive inhibitors, (Howes et al, under review) we examined the effect of the amino acid substitutions on the formation of DNA-protein complexes in the presence or absence of inhibitors by electrophoretic mobility shift assay (EMSA) (Hellman and Fried 2007). The competitive inhibitor L67 reduced the binding of wild-type and mutant versions of DNA ligase I to the DNA (**Fig. 16B**) whereas the uncompetitive inhibitors, L82-G17 and L82 promoted the presence of substrate bound ligase. In contrast, neither L82-G17 nor L82 increased the binding of the mutant versions of DNA

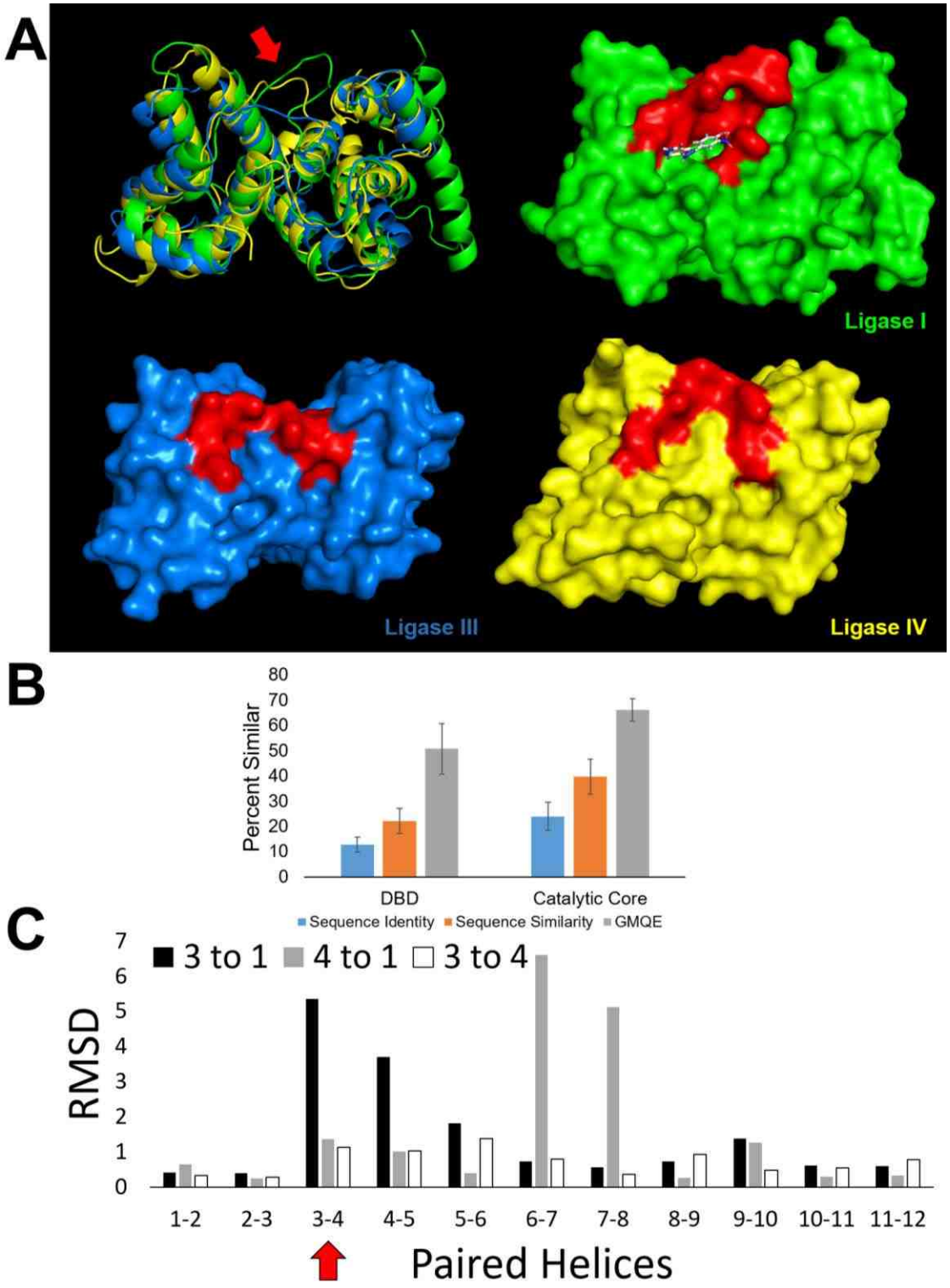


Figure 17. Structural differences between the DNA binding domains of human DNA ligases

Figure 17. Structural differences between the DNA binding domains of human DNA ligases

(A) Top left: Ribbon diagram showing the alignment of the DNA binding domains of DNA ligase I (green), DNA ligase III (blue), and DNA ligase IV (yellow). The red arrow highlights the area of greatest structural local dissimilarity between the three human DNA ligases. Top right: space filling model of DNA ligase I DBD, with docked L82 shown in white. Bottom left and right, space filling models of DNA ligase III (blue) and DNA ligase IV (yellow). Amino acids shown in red correspond to the loops indicated by the red arrow in the upper left panel. (B) Sequential and structural similarities of the DNA binding domains (DBD) and catalytic cores (AdD & OBD) of the three human DNA ligases. (C) Local RMSD of paired helices within the DBDs of the three human DNA ligases. The loops connecting helices 1-2, 3-4, 5-6, 7-8, 9-10, and 11-12 are located on the DNA binding side of the DBD, while all the others are on the exterior of the protein. Helix pair 3-4, indicated by the red arrow, represents the area of greatest structural dissimilarity in the DNA binding surface between DNA ligases I and III.

ligase I to DNA. Together these biochemical results validate the predictions from our *in silico* modeling of inhibitor binding.

Despite sequence divergences, the secondary structures of the DNA binding domains of the human DNA ligases are highly conserved.

The identification of the inhibitor binding site in DNA ligase I prompted us to examine the comparable regions of DNA ligase III and IV with the expectation that there will be differences that underlie the selectivity of existing inhibitors and may be exploited to guide the design of more selective inhibitors for each of the three human DNA ligases. While the DBDs of the mammalian DNA ligases are more recent, evolutionarily speaking, (Ellenberger and Tomkinson 2008) and

therefore more diverse than the adenylation (AdD) and oligosaccharide/oligonucleotide binding (OBD) domains, they do appear to be structurally similar (**Fig. 17A**). To better understand the differences and similarities between the catalytic regions of the three human ligases, each isoform was compared to the other two using protein structure homology modeling in conjunction with analysis of amino acid sequence homology using a BLOSUM62 substitution matrix. The OBD and AdD domains, which contain the conserved motifs that define the nucleotidyl transferase superfamily (Shuman, Liu et al. 1994, Shuman and Schwer 1995), have, as expected, a high degree of sequence identity and predicted structural homology. In contrast, the DBDs have only about 22% amino acid similarity. Nonetheless, these domains appear to still have a reasonably high level of structural homology, as shown by the global model quality estimation (GQME) (**Fig. 17B**). To corroborate this point, NCBI BLASTP searches limited to *Homo sapiens,* using either the DBDs or the AdD/OBD catalytic cores as the query sequences yielded very different results, with the catalytic cores but not the DBDs returning other DNA ligases as similar in sequence (data not shown).

While the human DNA ligase DBDs are structurally similar (**Fig. 17A**), they are some local differences. By analyzing the root mean square deviation (RMSD) between paired helices of different human DNA ligases, we identified the loop between the third and fourth helix as the part of the DBD that is most different between DNA ligases I and III (**Fig. 17C**). This loop, indicated with red in **Figure**

17A, is the major contributor to the L82 inhibitor binding pocket described above. While this loop is longer in DNA ligase I compared with the other DNA ligases, it forms a more compact structure in DNA ligase I compared with DNA ligases III and IV. These differences in local structure presumably underlie the selectivity of L82 for DNA ligase I and suggest that the comparable regions of DNA ligases III and IV are less amenable to small molecule binding.

In summary, we have used an unbiased modeling approach to predict the binding sites of DNA ligase inhibitors that were identified previously by an *in silico* structure-based approach (Chen, Zhong et al. 2008, Zhong, Chen et al. 2008). These studies predicted that the DNA ligase I selective inhibitor, L82, and the DNA ligase I/III inhibitor, L67, occupy overlapping binding pockets that are immediately adjacent to the binding pocket targeted in the initial structure-based approach (Chen, Zhong et al. 2008, Zhong, Chen et al. 2008). We have validated the binding pocket of L82 by demonstrating that replacing a glycine residue with bulkier residues that occlude the binding pocket abrogates the inhibitory activity of L82. The definitive identification of the inhibitor binding site will facilitate the rational design of more selective inhibitors. Furthermore, the construction of cell lines that express inhibitor-insensitive versions of DNA ligase I will play a critical role in demonstrating that cellular responses induced by exposure to the DNA ligase I selective inhibitors, L82 and L82-G17 are in fact due to inhibition of DNA ligase I rather than off-target effects.

Materials & Methods

Chemicals

The chemicals L67 (IUPAC: 2-[(3,5-dibromo-4-methylphenyl)amino]-N'-[(2-hydroxy-5-nitrophenyl)methylidene]acetohydrazide), L82 (IUPAC: 4-chloro-5-{2-[(4-hydroxy-3-nitrophenyl)methylidene]hydrazin-1-yl}-2,3-dihydropyridazin-3-one, and L82-G17 (4-chloro-5-{2-[(3-hydroxyphenyl)methylidene]hydrazin-1-yl}-2,3-dihydropyridazin-3-one) were done as described previously (Howes 2017).

Targeted Mutagenesis and Protein Purification

Oligonucleotides for DNA substrates, and mutagenesis were designed in Serial Cloner (version 2.6.1) and ordered from Integrated DNA Technologies (IDT, Iowa, USA). Annealing temperatures for mutagenesis using the polymerase chain reaction (PCR) were calculated with the aid of OligoCalc (Kibbe 2007). PCR was performed using a Bio-Rad MyCycler thermal cycler. and oligonucleotides

DNA Ligation Assays

In the fluorescence-based ligation assay (Chen, Pascal et al. 2006, Chen, Ballin et al. 2009), purified DNA ligase I (500 fmol) was incubated in the presence or absence of either L82 or L82-G17 with fluorescent nicked DNA (1 pmol) for 15 minutes in a final volume of 20 µL containing 60 mM Tris–HCl [pH 7.4], 50mM

NaCl, 10 mM $MgCl_2$, 5 mM DTT, 1 mM ATP, 50 µg/ml BSA, and 4% DMSO) at 25 °C. All controls and blanks contained an equivalent volume of vehicle DMSO. Following incubation, reactions were further diluted to 200 µL with a 30-fold molar excess of an unlabeled oligonucleotide (5'-TAGGAGGGCTTTCCTCCTCACGACCGTCAAACGACGGTCA) identical in sequence to the ligated strand in 10 mM Tris–HCl pH 7.4, 50 mM KCl, 1 mM EDTA and 5 mM $MgCl_2$ and then heated to 95 °C for 5 min. After cooling to 4 °C at a rate of 2 °C/min, fluorescence at 519 nm (excitation at 495 nm) was measured immediately using the Synergy H4 microplate reader (BioTek). Data, including Michaelis-Menten kinetics, was analyzed through GraphPad Prism software.

DNA Binding Assays

For electrophoretic mobility shift assays, a radiolabeled 73 bp linear duplex with a single nick was generated as above except that there was a dideoxycytosine residue at the 3' terminus of the nick (Pascal, O'Brien et al. 2004, Chen, Zhong et al. 2008). Putative inhibitors, purified DNA ligase I and the nicked substrate (amounts) were incubated in 50 mM Tris pH 7.5, 15 mM NaCl, 1 mM DTT, 0.2mM ATP, 0.005 mg/mL BSA, 5 mM $MgCl_2$, 8.75% glycerol and 0.5% DMSO for 30 minutes. After electrophoresis through a 5% non-denaturing polyacrylamide gel using 0.5 x TBE pH 8.3 as running buffer, labeled

oligonucleotides were detected by phosphorimager analysis and imaged as described above.

Compound Docking

The ability of small molecules to bind to DNA ligases was evaluated *in silico* using the OpenEye suite of software, VIDA, Make Receptor, Omega and OEDocking in collaboration with Dr. Tudor Oprea's group. Small molecules were generated in the SMILES format using MarvinSketch, (version 14.9.8.0). For any situation in which multiple stereoisomers were possible, all were used. Chemical binning and hierarchical clustering was done using the online tool ChemMine. Figures were rendered with PyMol (version 1.3). Small molecule structures were consolidated using VIDA (version 4.2.1), and multiconformers of every small molecule were generated using Omega (version 2.5.1.4), with a maximum of 500 conformations per molecule. These multiconformer files were used by the FRED (Fast Rigid Exhaustive Docking) function of OEDocking (version 3.0.1) to predict each compound's ability to interact with potential receptor sites on a given protein. Receptor sites were generated using Make Receptor (version 3.0.1), generally using 2-4 amino acids to define the potential receptor site. The results of the docking assays were exported as a spreadsheet, for combination and analysis in Microsoft Excel 2013 (version 15.0.4797.1003). Each receptor was evaluated using confusion matrices, allowing for a quick and iterative process of putative receptor site improvement. In addition to assessing true positives and

negatives, we also evaluated confusion matrices by the geometric mean, (**Fig. 11B**) a unitless measure that is useful when the number of negatives is much greater than the number of positives (Kubat, Holte et al. 1998).

Receptor iteration was done by performing any of the following methods: expansion, contraction, translation, rotation, and in the presence or absence of DNA. Additionally, receptors were created from individual domains, as well as the entire available DNA ligase I structure. The exact number of docking sites tested in **Figure 11B** for round 1, round 2 and round 3 are 334, 428 and 657 respectively. Figures including protein, DNA and small molecules were rendered with PyMol (version 1.3) (DeLano 2002).

Interaction Reports

Protein-inhibitor interaction reports were generating using a built-in function of the FRED software. All data generated from the Leu335, Pro340, Gly352 receptor is the result of testing seven different slight variations of the same receptor to reduce the impact of one-off interactions. Interaction reports were converted to text documents before conversion to spreadsheets.

Homology Modeling

Initial evaluation of putative critical residues was done by *in silico* mutagenesis. Specifically, the raw amino acid sequence of DNA ligase I was altered in a text document prior to utilizing the homology modeling platform

SWISS-MODEL (Bordoli, Kiefer et al. 2009, Biasini, Bienert et al. 2014) to quickly and easily generate a probable three-dimensional structure of the altered protein, using wild-type ligase DNA ligase I (PDB: 1X9N) as the template. NCBI (National Center for Biotechnology Information) BLASTP (Basic Local Alignment Search Tool – Protein) searches were performed using the NCBI web interface (https://blast.ncbi.nlm.nih.gov/Blast.cgi) and aligning the two amino acid sequences (Altschul, Gish et al. 1990).

Statistical analysis

Data are expressed as mean ± SEM. For comparison of groups, we used the Student two-tailed *t* test. A level of $P < 0.05$ was regarded as statistically significant.

Acknowledgements

This work was supported by the University of New Mexico Comprehensive Cancer Center (P30 CA118100) and National Institute of Health Grants R01 GM57479 (to A.E.T.) and P01 CA92584. We would also like to thank David Howes for creating software to greatly expedite the transcription of residue interaction data.

CHAPTER 4

SUMMARY, CONCLUSION & FUTURE DIRECTIONS

Summary

The work presented in this dissertation has identified the small molecule DNA ligase I-selective inhibitor L82-G17. L82-G17 represents a substantial improvement over its parent compound, L82. L82 was a first-generation ligase inhibitor published in 2008 by the Tomkinson lab (Chen, Zhong et al. 2008, Zhong, Chen et al. 2008). L82-G17 is an uncompetitive LigI-selective inhibitor, proven to reduce the survival of wild type cells to a significantly greater degree than isogenic *LIG1* null cells. Beyond identifying improved DNA ligase inhibitors, I have also identified the pocket to which these inhibitors bind in the LigI DBD. Furthermore, during my graduate work, I found it helpful to compile data about specific ligation mutations and important residues, as well as maintaining a master list of ligase crystal structures. These were invaluable to me during my work, and I have included some of those lists here.

The hybrid structure of this dissertation means that chapters two and three are presented in their original published form, only modified for format. However, these two papers do not reflect the entirety of my research on DNA ligase inhibitors. Additional research was done into characterizing not only the ligase

inhibitors, but also the inhibitor binding site that was identified in chapter three of this document.

Key structural elements of the LigI DBD confer sensitivity to L82.

As was the case with Gly 448, the sequence of the loops is less important than the shape that sequence imparts. Examination of each ligase's tertiary structure shows that the equivalent area to the binding pocket in LigI yields more insight into L82's inability to inhibit either DNA ligase III or IV. Assaying the L82 binding pocket sizes show that these differences result in significantly less potential inhibitor binding surface area in LigIII than in LigI. The ratio of examined volume to predicted pocket size predicts that an average of 19.7% $\pm$2.5 of space in LigI is open enough for small molecule binding, compared to only 15.6% $\pm$1.2 in LigIII (**Fig. 18A**). In accordance with the binding loop sizes (**Fig. 17A**), LigIV falls between these two, with a pocket area of 16.8% $\pm$2.1. However, we refrain from drawing too many conclusions from comparisons between the x-ray crystal structures of LigI/III and LigIV, as LigI/III were co-crystallized with DNA, while LigIV was not. Aligning each DBD via helix 4 reveals that this spatially equivalent site encroached upon by helix 9 in both LigIII and LigIV (**Fig. 18B**). Furthermore, the similar to how the G448K and G448M mutations both prevented L82 binding, (**Figures 15B and 16**) helix 9 of LigIII and LigIV have sidechains that protrude into what would be the L82 binding pocket (**Fig. 18C**). This provides potential justification for the L82's selectivity for DNA ligase I.

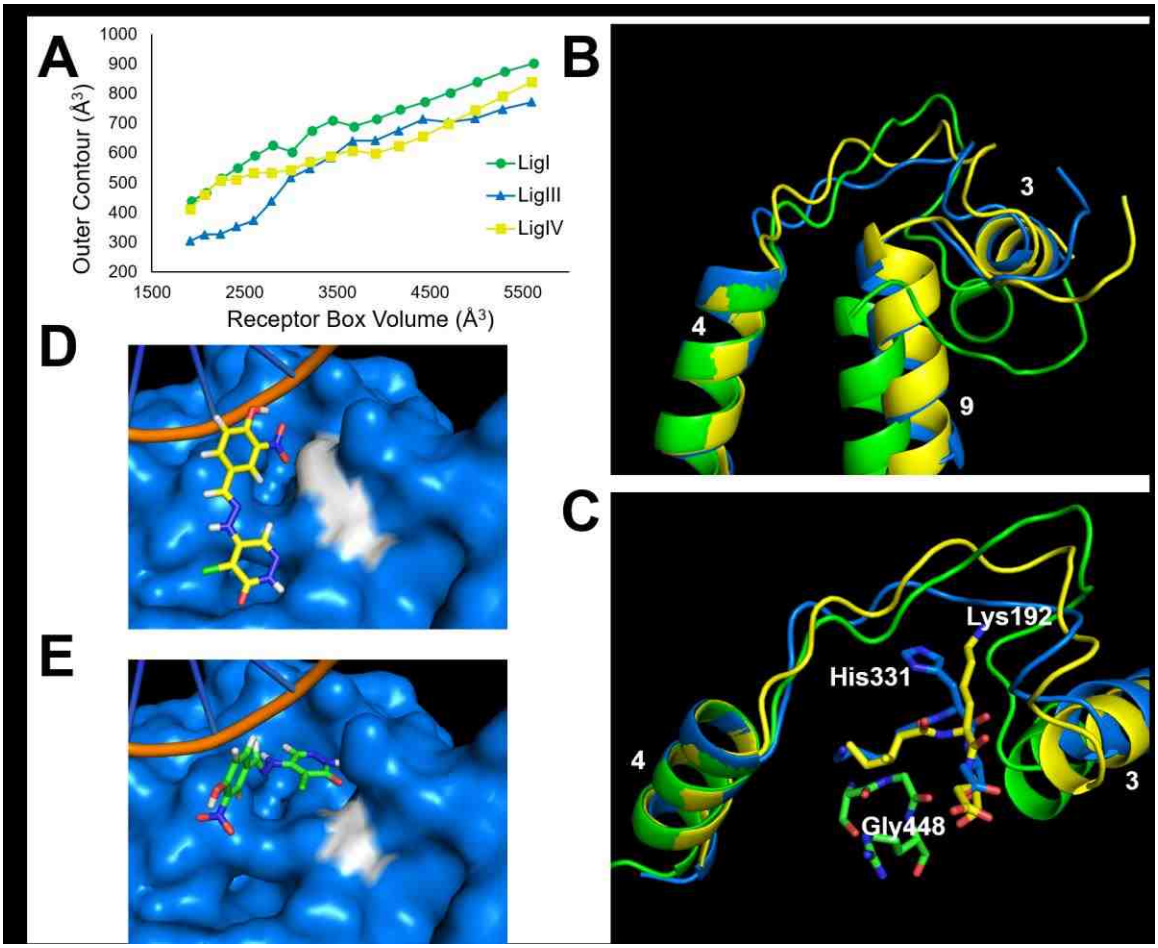


Figure 18. DNA Ligase III lacks the space for L82 to bind

(A) Over equivalent volumes of LigI and LigIII, LigI consistently has more surface area exposed for small molecules to fit in. This is true both in the presence and absence of DNA, data shown is in the absence of DNA. Green circles – LigI, blue triangles – LigIII, yellow squares – LigIV. (B) Human ligase DNA binding domains, aligned via the backbone of helix 4. Green – LigI, blue – LigIII, yellow – LigIV. The structure of helix 9 encroaches on the L82 binding pocket in both LigIII and LigIV. The effects of the tertiary structure are further exacerbated by the sequence, (C) in both LigIII and LigIV side chains protrude into the space that L82 binds in LigI. (D) His 331 (white) obfuscates the area equivalent to LigI's L82 binding pocket. L82 shown in yellow. (E) H331G *in silico* mutation (white) increases results in L82's predicted binding to the LigIII DBD. Bound L82 shown in green.

The binding loop of LigI was replaced with that of LigIII by both simple replacement and homology modeling. Furthermore, the LigIII equivalent to Gly

448, His 331, was mutated *in silico* to a glycine, removing the sidechain. While there was some negative effect on the predicted binding of L82, that effect was relatively small to the positive effect on L82 binding that removing the sidechain of His 331 had on L82's affinity for DNA ligase III (**Fig. 18 D&E**). *in silico* experiments show that deleting the histidine side chain, which takes up approximately 100 Å^3, opens the space needed for L82 binding, similar to what is observed in LigI. H331G *in silico* mutation results in L82 binding to LigIII in our predictive model. These two residues, His 331 and Lys192 in LigIII and LigIV respectively, appear to be the reason that L82 is ineffective, as they protrude into the cavity created by the L82 binding loop which connects helices 3 and 4 of the DBD. Studying the predicted mutant structures and the inhibitor binding, mutations of glycine 448 appear to obstruct L82's preferred binding site (**Fig.**

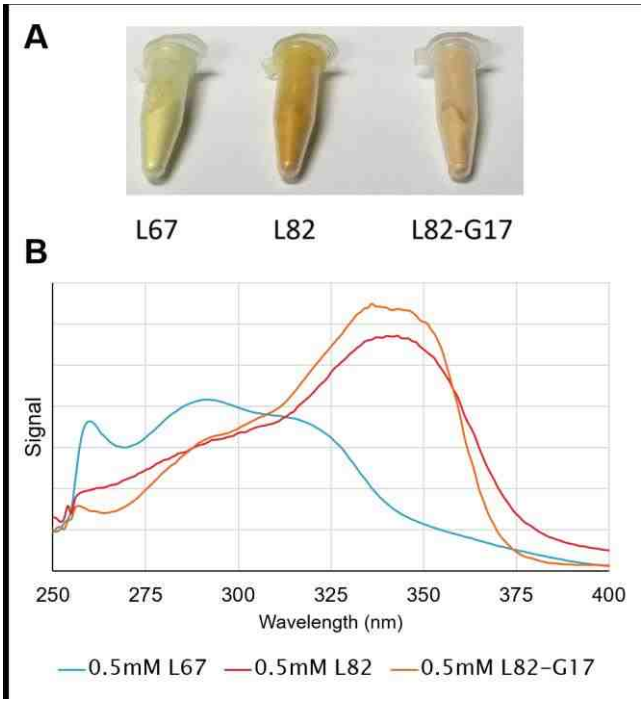


Figure 19. Spectroscopic profiles of LigI inhibitors

Analyzing the colour of L67, L82 and L82-G17. (A) The chemical differences between these compounds produce different colours. (B) UV spectra of L67, L82 and L82-G17 was used to confirm that the integrity of structures long after synthesis.

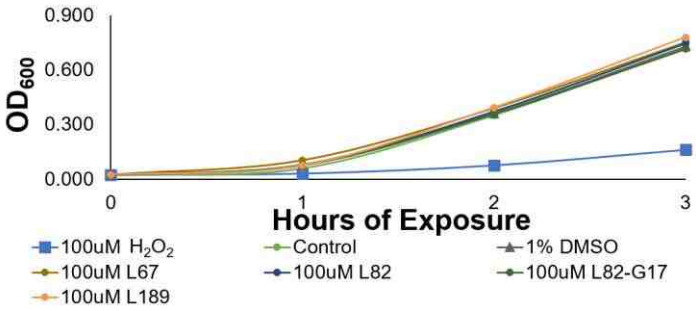


Figure 20. Ligase inhibitors do not kill or impede the growth of bacteria

Bacteria transformed with plasmid pUC19 were treated with 100 µM doses of ligase inhibitors (circles), an equivalent ratio of DMSO (triangles), or hydrogen peroxide (H_2O_2, squares). Only hydrogen peroxide made any difference to bacterial growth.

14A). Supporting this argument is that the equivalently located amino acid to glycine 448 on DNA ligases III covers the L82 binding site.

Colour and Spectroscopic Profile

Also investigated was the nature of the colour change observed in L82-G17 from L82 and L67. The UV-vis profile of L67, L82, and L82-G17 were analyzed. L67 has a visible spectrum peak of 434 nm, consistant with its more yellow colour. L82 and L82-G17 have absorbance peaks at 456 nm and 468 nm, respectively, which impart an orange colour (**Fig. 19A**). As expected, the UV spectra of L82 and L82-G17 are nearly identical, which follows their structural similarity (**Fig. 19B**). The main difference in the L82 and L82-G17 UV spectra occurs around 265 nm, at which point L82-G17 absorbs less than L82. This is consistant with the loss of a benzene bound -NO_2 group. The differences between L67 and L82 in the 250-300 nm range is due to the differences in the chain linking the two rings. The shoulder at 325 nm is caused by the two bromines on the left ring of L67 (**Fig. 3**) (Schirmer 1990, Linstrom and Mallard 2001).

Ligase Inhibitors Have No Effect on Prokaryotic Cells

Off target effects can be difficult to identify. One potential complication that could result from poor solubility is that the compounds themselves were effecting the cellular membrane by precipitating out of solution. No inhibitor had any effect on bacterial cell growth or proliferation (**Fig. 20**).

Modeling Data Supports Recent Findings on SCR7

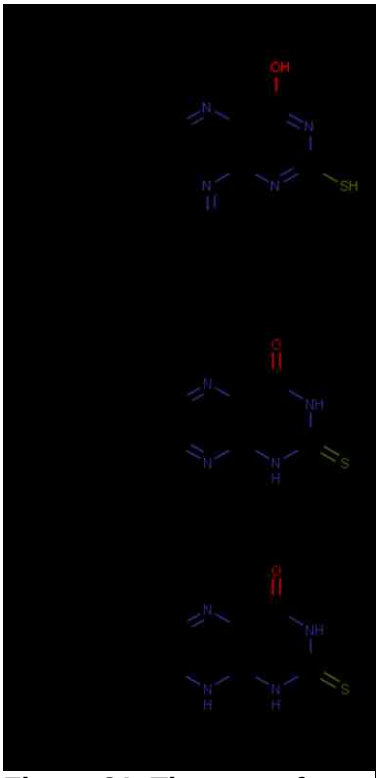


Figure 21. The many faces of SCR7

SCR7 is a reported, and debunked LigIV inhibitor with several reported structures (A) SCR7, (B) SCR7-G, and (C) SCR7-R.

Recently published data has called into question both the structure and the potency of SCR7 (Srivastava, Nambiar et al. 2012, Greco, Conrad et al. 2016, Greco, Matsumoto et al. 2016). Three structures of SCR7 have been identified, in addition to the original SCR7, (**Fig. 21A**) there is also SCR7-G (**Fig. 21B**) and SCR7-R (**Fig. 21C**). Each of the reported structures of alleged ligase inhibitor SCR7 were docked into the previously established binding sites. The calculated binding affinity was established relative to other known compounds, as described in chapter 3.

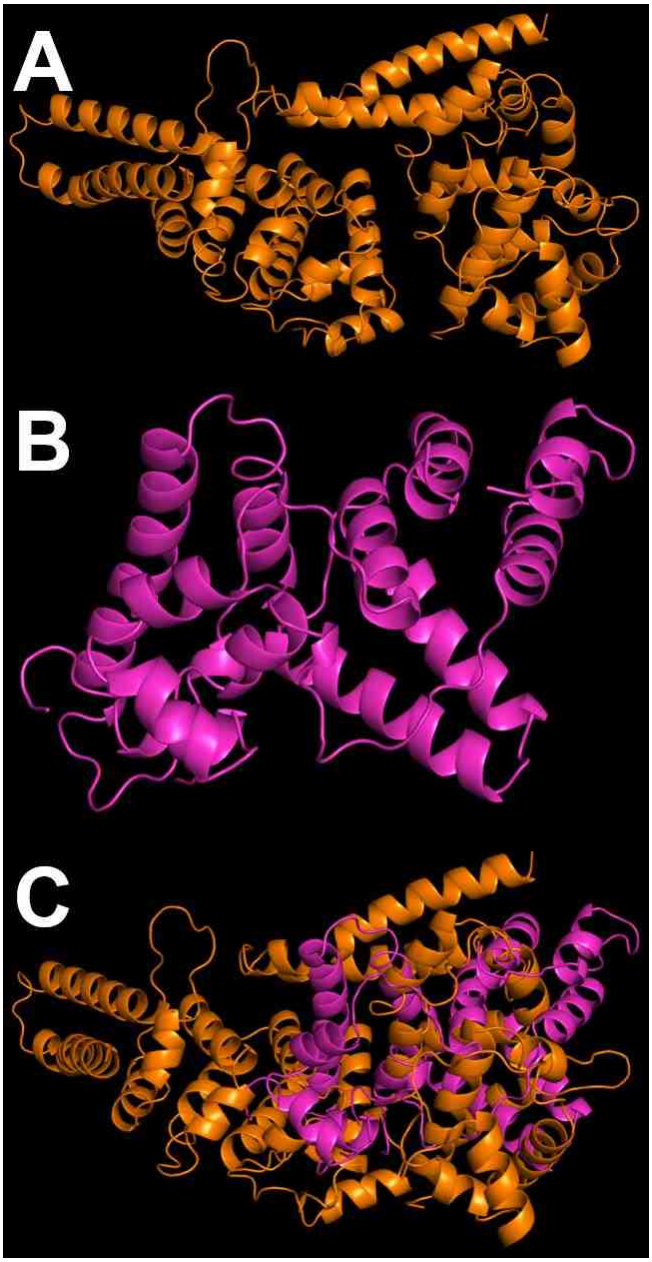


Figure 22. PatL1 is not at all similar to the DBD of LigIV

(A) PatL1 x-ray crystal structure, PDB: 2XEO. (B) DNA binding domain of DNA ligase IV, PDB: 3W1G. (C) Alignment of PatL1 and LigIV DBD yields an RMSD of over 20; these structures are not similar.

In addition, efforts were made to recreate structure that the original authors docked SCR7 onto. However, Srivastava *et al*. reference the PDB structure 2XEO as the basis for their homology modeling of DNA ligase IV: “The structure of DNA containing DSB was retrieved from PDB database (2XEO).” (Srivastava, Nambiar et al. 2012) 2XEO references a withdrawn structure of PatL1, a human mRNA decapping enzyme. I was able to find the paper that this structure was submitted with as well as the associated structure 2XES. 2XES is also an x-ray crystal structure and has the same title as 2XEO, “Human PatL1 C-terminal domain (loop variant)” and was submitted one

day after, May 16th, 2010. It does not contain DNA (Braun, Tritschler et al. 2010). As was expected, the structures of LigIV and PatL1 are highly dissimilar. A simple comparison between LigIV's DBD and 2XEO is 21.9 RMSD, which fits with how different the structures appear based on a visual comparison (**Fig. 22**). The LigIV DBD and the PatL1 structure are highly dissimilar, with an RMSD greater than 20. Initially beleiving "2XEO" to be a typo, a list of all published ligase structures was compiled (**Table 2**) (Subramanya, Doherty et al. 1996, Singleton, Håkansson et al. 1999, Lee, Chang et al. 2000, Odell, Sriskanda et al. 2000, Sibanda, Critchlow et al. 2001, Gajiwala and Pinko 2004, Pascal, O'Brien et al. 2004, Srivastava, Tripathi et al. 2005, Doré, Furnham et al. 2006, Nishida, Kiyonari et al. 2006, Pascal, Tsodikov et al. 2006, Nandakumar, Nair et al. 2007, Vijayakumar, Chapados et al. 2007, Pinko 2008, Han, Chang et al. 2009, Kim, Kim et al. 2009, Wu, Frit et al. 2009, Cotner-Gohara, Kim et al. 2010, Cuneo, Gabel et al. 2011, Mills, Eakin et al. 2011, De Ioannes, Malu et al. 2012, Ochi, Wu et al. 2012, Petrova, Bezsudnova et al. 2012, Surivet, Lange et al. 2012, Wang 2013, Murphy-Benenato, Wang et al. 2014, Unciuleac, Goldgur et al. 2017). No structure of any DNA ligase has a PDB ID similar to 2XEO, particularly none of the four structures that contain DNA: 1X9N, 2OWO, 3L2P, 4GLX.

In the interest of a thorough investigation, all SCR7 structures were docked onto multiple pockets of PatL1 crystal structure 2XES. SCR7 is not predicted to bind to PatL1, nor is L67, L82, L189 or any of the other previously described DNA ligase inhibitors (Zhong, Chen et al. 2008) or L82-G17.

PDB ID	Name	Species	Contains DNA	Portion	AA Length	Other PDB	Year
1A0I	T7 Ligase	*Enterobacteria Phage T7*	No	Full	348	*n/a*	1996
1B04	LigA	*Geobacillus stearothermophilus*	No	AdD	318	*n/a*	1998
1FVI	DNA Ligase	*Chlorella virus*	No	>90%	297	*n/a*	2000
1DGS	LigA	*Thermus filiformis*	No	Full	667	1DGT	2000
1IK9	Ligase IV	*Homo sapiens*	No	<5%	37	*n/a*	2001
1V9P	LigA	*Thermus filiformis*	No	>85%	584	*n/a*	2004
1TAE	LigA	*Enterococcus faecalis*	No	AdD	332	1TA8	2004
1X9N	Ligase I	*Homo sapiens*	**Yes**	DBD, AdD & OBD	668	*n/a*	2004
1ZAU	LigA	*Mycobacterium tuberculosis*	No	AdD	328	*n/a*	2005
1Z56	Lig4	*Saccharomyces cerevisiae*	No	BRCT domains	264	*n/a*	2006
2CFM	DNA Ligase	*Pyrococcus furiosus*	No	Full	561	*n/a*	2006
2HIV	DNA Ligase	*Sulfolobus solfataricus*	No	Full	621	2HIX	2006
2OD8	Cdc9	*Saccharomyces cerevisiae*	No	<5%	22	*n/a*	2007
2OWO	LigA	*Escherichia coli*	**Yes**	Full	671	*n/a*	2007
3BAA	LigA	*Enterococcus faecalis*	No	AdD	332	3BAB, 3BA9, 3BA8	2008
3BAC	LigA	*Haemophilus influenzae*	No	AdD	270	*n/a*	2008
3II6	Ligase IV	*Homo sapiens*	No	BRCT domains	263	*n/a*	2009
3JSN	LigA	*Staphylococcus aureus*	No	AdD	318	3JSL	2009
3GDE	DNA Ligase	*Archaeoglobus fulgidus*	No	Full	558	*n/a*	2009
3L2P	Ligase III	*Homo sapiens*	**Yes**	DBD, AdD & OBD	579	*n/a*	2010
3PN1	LigA	*Haemophilus influenzae*	No	AdD	318	*n/a*	2011
3QVG	Ligase III	*Homo sapiens*	No	BRCT domain	89	3PC6, 3PC7, 3PC8	2011
3VNN	Ligase IV	*Homo sapiens*	No	AdD	139	*n/a*	2012
4EQ5	DNA Ligase	*Thermococcus sibiricus*	No	Full	571	*n/a*	2012
4HTO	Ligase IV	*Homo sapiens*	No	DBD	240	4HTP	2012
4GLX	LigA	*Escherichia coli*	**Yes**	OBD & AdD	586	*n/a*	2012
4GLW	LigA	*Streptococcus pneumoniae*	No	AdD	305	*n/a*	2012
4EFB	LigA	*Enterococcus faecalis*	No	AdD	331	*n/a*	2013
3W1G	Ligase IV	*Homo sapiens*	No	DBD, AdD & OBD	610	3W1B, 3W5O	2013
4LH6	LigA	*Enterococcus faecalis*	No	AdD	323	4LH7	2014
5FPO	DNA Ligase	*Staphylococcus aureus*	No	AdD	318	5PFR	2016
5TT5	LigA	*Escherichia coli*	No	Full	691	*n/a*	2017

Table 2. Compilation of DNA ligase crystal structures

Summary Remarks

Compiling the collective knowledge of all ligase structure and LigI mutations in one place is also particularly useful for anyone beginning or continuing work on DNA ligases, particularly any structural biology work. The first crystal structure of a DNA ligase was published in 1996 (Subramanya, Doherty et al. 1996). There are now 45 x-ray crystal structures of DNA ligases published, (**Table 2**) covering a 19 different species (Subramanya, Doherty et al. 1996, Singleton, Håkansson et al. 1999, Lee, Chang et al. 2000, Odell, Sriskanda et al. 2000, Sibanda, Critchlow et al. 2001, Gajiwala and Pinko 2004, Pascal, O'Brien et al. 2004, Srivastava, Tripathi et al. 2005, Doré, Furnham et al. 2006, Nishida, Kiyonari et al. 2006, Pascal, Tsodikov et al. 2006, Nandakumar, Nair et al. 2007, Vijayakumar, Chapados et al. 2007, Pinko 2008, Han, Chang et al. 2009, Kim, Kim et al. 2009, Wu, Frit et al. 2009, Cotner-Gohara, Kim et al. 2010, Cuneo, Gabel et al. 2011, Mills, Eakin et al. 2011, De Ioannes, Malu et al. 2012, Ochi, Wu et al. 2012, Petrova, Bezsudnova et al. 2012, Surivet, Lange et al. 2012, Wang 2013, Murphy-Benenato, Wang et al. 2014, Unciuleac, Goldgur et al. 2017). as well as several nuclear magnetic resonance (NMR) structures of ligase fragments (PDB ID's: 1L7B, 1IN1, 1IMO, 1UW0, 2E2W, and 2LJ6) (Krishnan, Thornton et al. 2001, Kulczyk, Yang et al. 2004, Sahota, Goldsmith-Fischman et al. 2004, Nagashima 2006, Natarajan, Dutta et al. 2012). Of these, only four, 1X9N, 2OWO, 3L2P and 4GLX show the ligase complexed with DNA. 1X9N and

3L2P are structures of LigI and LigIII, respectively, while the other two are of *Escherichia coli* ligase, LigA (**Table 1**).

The active site lysine of human DNA ligase I was identified (Tomkinson, Totty et al. 1991) and then confrimed in 1991 by mutagenesis (Kodama, Barnes et al. 1991). Many other studies have examined the effects of various mutations on the function of DNA ligase I. I have endeavoured to catalog the effects of every published LigI mutation in **Table 3**. Negative data, such as glycine 363, were also included; these residues may or may not be critical in ways yet to be determined. Two residues, Gly 448 and Gly 571 were each found to be mutated once in the COSMIC database. “New data” indicates any data generated from my graduate work. Furthermore, several conclusions have been drawn from either observation of the three-dimensional structure of LigI, or by computational means. Every residue that is predicted to be important to LigI structure or function has been catalogued in **Table 4**. Two residues, Glu 621 and Arg 871 were each found to be mutated once in the COSMIC database (Forbes, Beare et al. 2015).

Residue	Function	COSMIC Mutation	COSMIC Source	Reference
Phe 8	PCNA Binding			Vijayakumar et al. 2009
Phe 9	PCNA Binding			Vijayakumar et al. 2009
Ser 51	Mediated RFC Interaction, phosphorylation site			Peng et al. 2012
Ser 66	Cell cycle, phosphorylation site			Peng et al. 2012
Ser 76	Cell cycle, phosphorylation site			Peng et al. 2012
Ser 91	Cell cycle, phosphorylation site			Peng et al. 2012
Gly 363	*(No effect on PCNA/DNA Binding)*			Song et al. 2009
Gln 365	*(No effect on PCNA/DNA Binding)*			Song et al. 2009
Leu 392	*(No effect on PCNA/DNA Binding)*			Song et al. 2009
Met 393	*(No effect on PCNA/DNA Binding)*			Song et al. 2009
Leu 394	*(No effect on PCNA/DNA Binding)*			Song et al. 2009
Lys 422	Impaired DNA Binding			Song et al. 2009
Lys 423	Impaired DNA Binding			Song et al. 2009
Ile 426	*(No effect on PCNA/DNA Binding)*			Song et al. 2009
Ile 427	*(No effect on PCNA/DNA Binding)*			Song et al. 2009
Lys 428	Impaired DNA Binding			Song et al. 2009
Leu 446	*(No effect on ligation or inhibition)*			*(new data)*
Ser 447	*(No effect on ligation or inhibition)*			*(new data)*
Gly 448	L82 inhibition	G448R	Squamous Cell Carcinoma	*(new data)*
Arg 449	*(No effect on ligation or inhibition)*			*(new data)*
Leu 450	*(No effect on ligation or inhibition)*			*(new data)*
Glu 566	Highly impaired AMP binding			Kodama et al. 1991
Tyr 567	*(No effect on ligation)*			Kodama et al. 1991
Lys 568	Prevents AMP binding			Kodama et al. 1991
Tyr 569	Impaired ligation			Kodama et al. 1991
Asp 570	Impaired ligation			Kodama et al. 1991
Gly 571	Prevents AMP binding	G571R	Adenocarcinoma	Kodama et al. 1991
Arg 573	Prevents AMP binding			Kodama et al. 1991
Gln 591	*(No effect on ligation or inhibition)*			*(new data)*
Arg 771	Impaired ligation			Barnes et al. 1992

Table 3. Biochemically tested DNA ligase I mutants

Residue	Function	COSMIC Mutation	COSMIC Source	Reference
Asn 336	Predicted L82 binding			*(new data)*
Leu338	Predicted L82 binding			*(new data)*
Asp 351	Predicted L82 binding			*(new data)*
Arg 589	Predicted L82 binding			*(new data)*
Thr 614	Mutation computed to be structurally damaging			Singh et al, 2011
Glu 621	Metal ion coordinating	E621K	Endometrioid Carcinoma	Pascal et al, 2004
Phe 635	Phenylalanine wedge			Pascal et al, 2004
Glu 720	Metal ion coordinating			Pascal et al, 2004
Arg 871	Salt bridge	R871C	Basal Cell Carcinoma	Pascal et al, 2004
Phe 872	Phenylalanine wedge			Pascal et al, 2004
Gln 892	Mutation computed to be structurally damaging			Singh et al, 2011

Table 4. Residues of DNA ligase I implicated by computational means

Conclusion

As stated at the end of chapter 1, I listed two specific aims in my thesis proposal, each with three sub-aims. I believe that both of these proposed research goals were achieved. During my time in graduate school, I have published one book chapter, three papers, with two that have been submitted for publication:

- Chumsri, S., **Howes, T.**, Bao, T., Sabnis, G., Brodie, A. Aromatase, aromatase inhibitors, and breast cancer. J Steroid Biochem & Mol Biol. 2011;**125**:13-22. (Chumsri, Howes et al. 2011)
- Chumsri, S., Sabnid, G.J., **Howes, T.**, Brodie, A.M.H. Aromatase inhibitors and xenograft studies. Steroids. 2011;**8**:730-735. (Chumsri, Sabnis et al. 2011)
- **Howes, T.R.** and Tomkinson A.E. DNA Ligase I, the Replicative DNA Ligase. Subcell Biochem. 2012;**62**:327-341. (Howes and Tomkinson 2012)
- Tomkinson, A.E., **Howes, T.R.L.**, Wiest, N.E. DNA ligases as therapeutic targets. Transl Cancer Res. 2013;**2**:203-214. (Tomkinson, Howes et al. 2013)
- **Howes, T.R.L.**, Sallmyr, A., Brooks, R., Greco, G.E., Jones, D.E., Matsumoto, Y., Tomkinson, A.E. Characterization of an

uncompetitive inhibitor of DNA ligase I. *(under review)* (Howes 2017)

- **Howes, T.R.L.**, Jones, D.E., Bologa, C., Yang, J., Matsumoto, Y., Oprea, T.I., Tomkinson, A.E. Prediction and Validation of Inhibitor Binding Pocket on DNA Ligase I. *(submitted)* (Howes 2017)

Aim 1

The first stated aim was to identify determinants of structure, activity and specificity of DNA ligase inhibitors by characterizing derivatives of the DNA ligase I inhibitor, L82, and the DNA ligase I/III inhibitor L67. This was to be achieved via assessing the chemical derivatives of L67 and L82 that the Tomkinson lab had in it's possession, for their ability to inhibit nick ligation by LigI or LigIII. The results of this step are shown in **Figure 4A** and **Figure 5**, and identified L82 derivative L82-G17 as superior to L82, being more effective against LigI and less effective against LigIII. This was followed by an a comprehensive examination of the properties of each inhibitor, which identified three subgroups of ligase inhibitors: vinyl-, hydrazine-, and hydrazone-linked (**Fig. 10)**. It was hypothesized that there would be chemical differences between active inhibitors and inactive inhibitors, and the analysis revealed that a meta polar group was crucial to inhibition (position 8 or 10 in **Figure 10C**). Other chemical properties of inhibitors were also determined, such as solubility and its UV-vis spectroscopic profile were also determined.

The final part of specific aim 1 was to examine the effects of L82-G17 in cell culture models. The proposal calls out a few techniques by name, such as colony assays (**Fig. 8 A&B**) and BrdU incorporation (**Fig. 7A**). The third technique specifically mentioned is iPOND, or isolation of proteins on nascent DNA, (Sirbu, Couch et al. 2011) a technique that allows for spatiotemporal analysis of replication fork dynamics. However, it was determined that while it is an interesting and powerful technique, it would be more beneficial to return to iPOND with a more potent DNA ligase I inhibitor.

My work has shown that L82-G17 specifically inhibits DNA ligase I both biochemically, and in cell culture models, and has identified the underlying structural commonalities behind those compounds that inhibit DNA ligase I.

Aim 2

The second specific aim in the thesis proposal for this work was proposed using molecular modeling approaches to predict a binding site for the current pool of DNA ligase inhibitors and make specific amino acid substitutions to test these predictions. The L82 binding site, shown in **Figure 12C**, was predicted and identified by screening the entirity of LigI *in silico* for potential ligase inhibitor binding sites. Potential small-molecule binding pockets were evaluating with confusion matrices, using the known data set of previously biochemically tested compounds to look for clusters of true positives. The best candidates were then selected and modified in an iterative process. The result of this was that a region

that returned L67 and L82 as the top two hits in the presence of DNA, and only L67 in the absence of DNA, mirroring what we knew to be biochemically true about the inhibitors.

Using this predicted binding pocket, I identified all of the amino acids that were predicted to be involved with inhibitor binding. In order to prioritize these for biochemical mutagenesis studies, I performed large scale single amino-acid replacements *in silico*, mutating not only those residues, but those around them as well. During this time, I also did consensus replacement mutations of LigI, replacing the loop connecting helices 3 and 4, with that of LigIV. These mutants were resistant to L82 inhibition. This supported the data generated from *in silico* mutagenesis, which identified glycine 448 as the residue that, when mutated, had the largest effect on L82 binding (**Fig. 14B**). Glycine 448 mutants were generated and purified by FPLC. These were also protected from inhibiton by L82 (**Fig. 16**).

Biochemically validating the computational model allowed me to use that model as the basis for further invesitation. I began by analyzing why L82 bound to LigI but not LigIII or LigIV. It is a long established fact that there is a large degree of amino acid sequence homology between all DNA ligases (Doherty and Suh 2000). It has also been qualitatively observed that the structures of ligase DNA binding domains are highly similar, despite fairly large differences in amino acid sequence. I was surprised to learn that this had not been backed up quantitatively, however. In order to determine what makes differences exist in the

DBDs, it is important to know just how similar they are. As expected, the catalytic core (AdD & OBD) of each ligase is both highly similar, both sequentially and structurally, however, the DBDs were also similar on a strucrual level, despite their differences in sequence (**Fig. 17B**). Most importantly, the ratio of sequence to structural homology meant showed that the DBDs were more structurally similar, relative to their sequnce identity, than the catalytic cores were. The loop between helices 3 and 4, which defines most of the L82 binding pocket, is the area of greatest structural dissimilarity on the DNA binding surface of the DNA binding domains (**Fig. 17 A&C**).

Further using the L82 binding model presented here, I was able to show that, in the equivalent region of LigIII, there was less space availiable for inhibitor binding. The ratio of calculated volume for small molecule binding in LigIII was significantly less than that of LigI, over a range of assayed sizes. Additionally, LigIII's equivalent to glycine 448, histidine 331, protruded into the area of that would be the L82 binding site. The difference in volume between these two equivalent areas being assayed was, on average, 126 $Å^3$, the volume of a histidine side chain is approxinately 100 $Å^3$. Based on this, I tested a H331G *in silico* mutant, and found that this was predicted to restore L82 binding (**Fig. 18**). Based on this, I am highly confident that we have located the site of L82 interaction and inhibition.

Future Directions

The data presented here represents significant steps forward in our understanding of ligase inhibition. However, there is always more research that can be done. In chapter one, I mention that *LIG1, LIG3* and *LIG4* were identified as mutated at a rather low rate of 0.69%, 0.49% and 0.58%, respectively (Forbes, Beare et al. 2015). Data on protein expression, however, is not as easy to come by, or to analyze. I believe that expression levels of ligases in cancer cells represents an area of investigation that would identify cancers and cancer cell lines that would be ideal targets for ligase inhibiton.

Complementing cells with the L82-resistant ligase mutant

The most immediate next step for continuing studies with L82 and L82-G17 would be to transfect LigI-resistant mutant G448K into LigI deficient cells. This would show both that L82 acts directly on L82 inside the cell, but also identify any toxicity that may be the result of off-target effects. I have previously attempted to show this using thermal titration of L82 treated cells, (Jafari, Almqvist et al. 2014) however, this would be a superior approach. I hypothesize that transfecting *LIG1-G448K* into DNA ligase I deficient cells would resore L82 sensitivity.

Attenuating DNA Ligase III to L82 inhibition

In **Figure 18** I propose a rational behind the ineffectiveness of L82 and L82-G17 on DNA ligase III. However this hypothesis, while backed up by *in silico* data, still requires biochemical testing. I would evaluate this mutation the same way that LigI mutations were tested in chapter three. Express and purify the H331G LigIII mutant from bacteria, and then assay its ability to both bind to and seal nicked DNA by ligation and electrophoretic mobility shift assays. I hypothesize, based on the data presented above, that removing the sidechain of histidine 331 would attenuate LigIII to inhibition by L82.

Sequential and Structural Conservation of DNA Ligases

In chapter three I touched briefly on how the structures of DNA ligases are more structurally similar than their amino acid sequences indicated. This is something that has been observed since the structures were determined, but had not been addressed quantitatively. Further investigation could be done into the structural similarity, not just between the 49 published structures of DNA ligases, but also the 8 full and 17 partial structures of mRNA capping enzymes availiable on the Protein Data Bank (Berman, Westbrook et al. 2000). These degrees of structural similarity would then be compared with the amino acid sequence similarity, such as in **Figure 17B**, but on a larger scale. This kind of knowledge has the potential to be extremely useful in both designing specific inhibitors, and evaluating inhibitor mechanisms of action.

Delivering Ligase Inhibitors

Both L67 and L82, as well as L82-G17, have limited solubility in acqueous solutions. While both readily dissolve into DMSO at 50-75 mM, DMSO is not ideal for measuring any biochemical activity. Several buffer systems were tested for their ability to solubilize these ligase inhibitors, including PBS, HEPES, and Tris, at different pH values. Inhibitors were more readily soluble in these, versus $_{d}H_2O$ alone. The solution that resulted in the greatest solubility of L82 and L67 was Tris pH 7.5 (data not shown). Also evaluated was the solubility of these inhibitors in ethanol as compared to DMSO. Both L82 and L67 were significantly more soluble in DMSO.

In collaboration with Dr. Eric Carnes at the University of New Mexico's department of chemical and nuclear engineering, we evaluated the effects of using nanocarriers to deliver L67. Three different types of nanocarriers, targeted- and untargeted-DOPC, as well as DoTAP (Butler, Durfee et al. 2016). These nanoparticles were effective at delivering L67 and increased its toxicity (data not shown). Ultimately, it was determined that L82 was the primary focus of my research, and this avenue of investigation was halted. It remains, however, an area of DNA ligase inhibitor research with a great deal of potential, and one that other research groups are also persuing (John, George et al. 2015).

Potential New Ligase Inhibitors

The small-molecule ligase inhibitors that we currently have are limited in several ways. They are poorly soluble in aqueous conditions, and are effective only in the low micromolar range. These solubility issues have already caused trouble for our lab, impairing both NMR and x-ray crystallographic studies. Using the ligase inhibitor binding site identified by molecular modeling, over a thousand derivatives of ligase inhibitors L67, L82 and L82-G17 have been docked into the inhibitor pocket to assess their potential for ligase inhibition. In addition, over six thousand small molecules between 300 and 400 molecular weight, were pulled from PubChem and modeled into the inhibitor binding pocket on DNA ligase I. Additionally, molecules of 75% of greater similarity to existing inhibitors, or *in silico* derivatives were pulled from the inventories of various chemical vendors such as ChemBridge. These chemicals were docked into the inhibitor binding site identified in chapter 3, and the top results were, in an iterative process, modified and re-docked into the pocket. The chemicals shown in **Figure 23A** are the results of this screening process. While several of these do not pass Lipinski's rule of 5 (Lipinski 2000) they are important in that they show it is theoretically possible to create an inhibitor, TH5-32-2-11 (**Fig. 23B**), that binds in both the L82 and L67 pockets. This compound, or one like it, has the potential to both be an extremely potent and a highly specific inhibitor designed for ligase I.

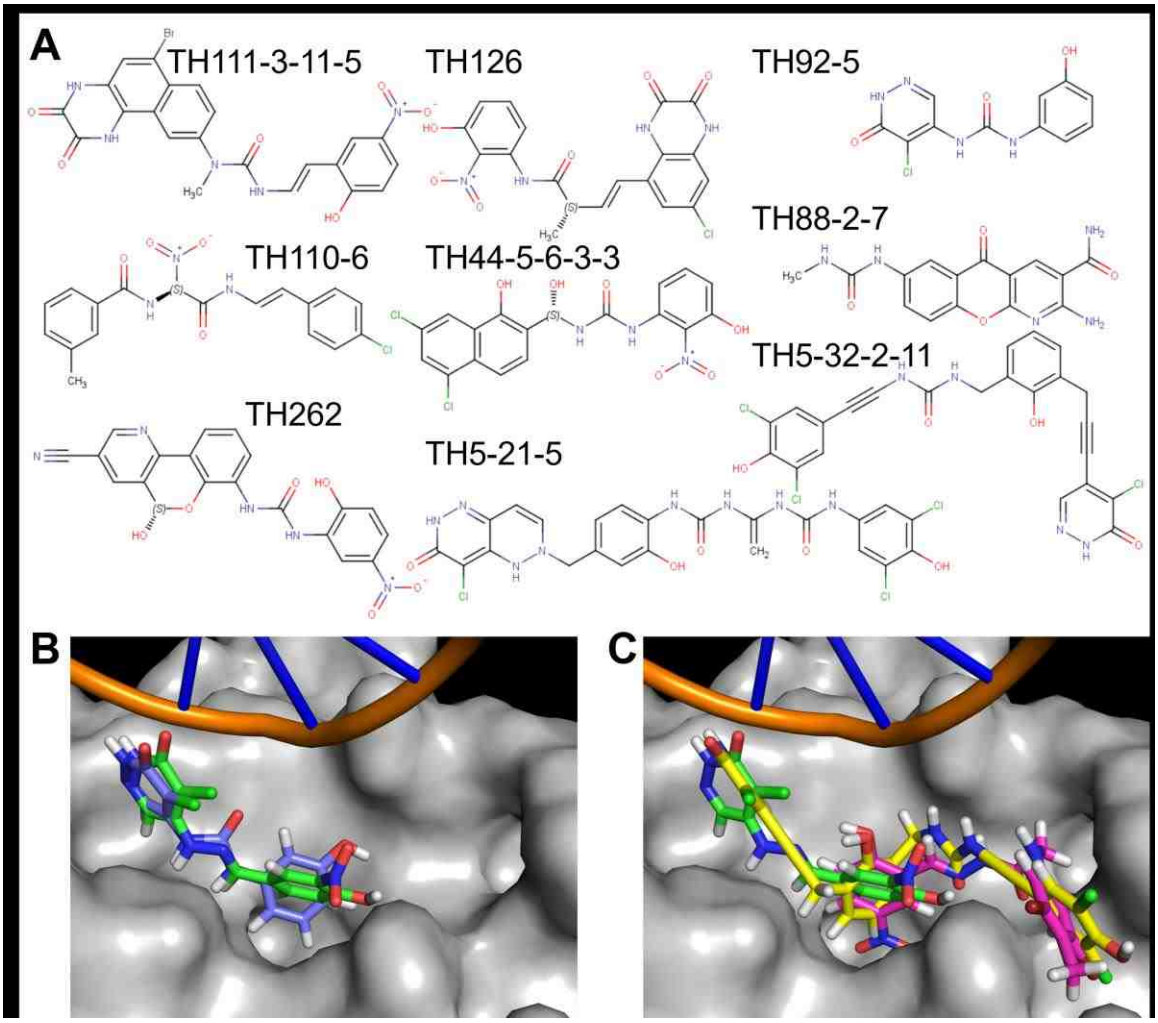


Figure 23. Potential DNA ligase I inhibiting compounds identified by in silico experiments

(A) Nine different potential ligase inhibitors with binding superior to L67 and L82, as determined by our computational model of inhibitor binding. (B) TH92-5 has the same structure as L82-G17, except the linker has been changed. (C) TH5-32-2-11 represents a construct that could theoretically overlap with the binding site and conformation of both L67 and L82.

From here, I believe the best course of action would be a new CADD screen against this verified inhibitor pocket. The current molecules have not insignificant solubility issues, and as mentioned previously, the N-N bond(s)

within L67 and L82 could present significant obstacles if these drugs are to be developed with an eye on therapeutic possibilities. To this end, proposed inhibitor TH95-2 (**Fig. 23C**) has had its hydrazone linker replaced with a urea based one, and may be a potential candidate for future examination. Another option is a pharmacophore based approach, which abstracts information about docking compounds, instead of docking actual chemicals, (Horvath 2011, Sanders, Barbosa et al. 2012) which is what I have done. Pharmacophore based screening has already been used to identify potential ligase inhibitors, (Krishna, Singh et al. 2014) but has not been utilized to examine the L82 binding site identified here (Howes 2017).

Final Remarks

My graduate work started with evaluating LigI phosphorylation by mass spectrometry, and testing Rad50 inhibitors. After two years in Baltimore, the lab moved to Albuquerque, New Mexico, where I transitioned into the work on DNA ligase inhibitors that is presented here. I would like to again thank Alan Tomkinson for accepting me into his lab, and for his support over these past years.

This year, 2017, is the 50th anniversary of the discovery of the first DNA ligase. Since then, there have been significant advances in our collective knowledge of their structure and function. Since then, over 50 structures of ligases have been published. DNA ligases are critical to cellular viability, during

DNA replication, a ligase interacts with PCNA and fully encircles DNA during ligation, a mockup of which can be seen in **Figure 24**, (Pascal, O'Brien et al. 2004) the exact orientation and mechanism of interaction has yet to be determined. The information that was gleaned from the crystal structure alone is remarkable, and further biochemical, biological and structural studies will invariably advance our understanding of how ligases operate. The predisposition to cancer in DNA ligase I-deficient mice highlights the importance of understanding how these proteins coordinate and regulate each other to maintain genome integrity. The ability to harness this information and design specific inhibitors represents a potentially powerful tool, both as a research tool, and as a therapeutic. I am confident improved inhibitors will be

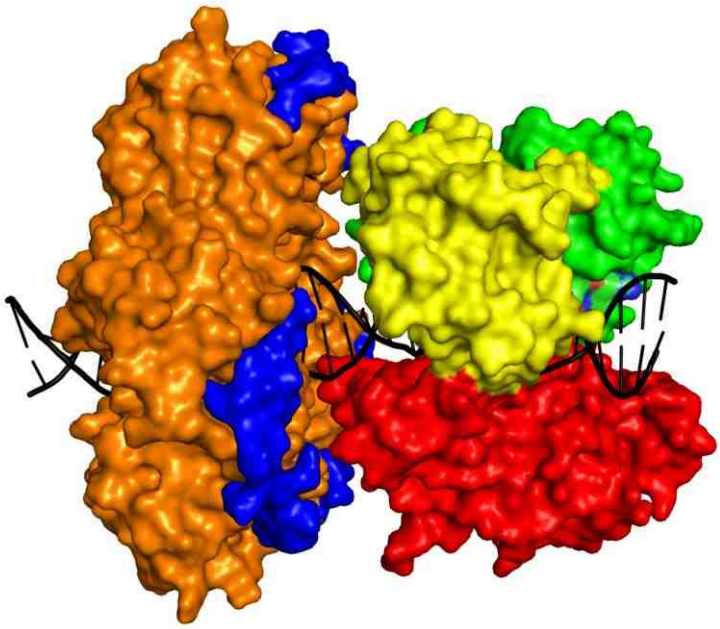


Figure 24. Ligase I and PCNA

A mockup of how PCNA and DNA ligase I might interact during ligation. LigI is shown in the same colours as in **Figure 1**, red DBD, greed AdD, and yellow OBD. PCNA is shown in orange, and p21, (Levin, Bai et al. 1997) which binds to PCNA at the same location as LigI's PIP box, is shown in blue. (PDB ID's used: 1AXC, 1X9N, 1BNA (Drew, Wing et al. 1981, Gulbis, Kelman et al. 1996, Pascal, O'Brien et al. 2004)).

developed, either by improving on the current generation of ligase inhibitors, or by using the identified L82 binding site to design novel ones.

Materials & Methods

Chemicals

The chemicals L67 (IUPAC: 2-[(3,5-dibromo-4-methylphenyl)amino]-N'-[(2-hydroxy-5-nitrophenyl)methylidene]acetohydrazide), L82 (IUPAC: 4-chloro-5-{2-[(4-hydroxy-3-nitrophenyl)methylidene]hydrazin-1-yl}-2,3-dihydropyridazin-3-one, and L82-G17 (4-chloro-5-{2-[(3-hydroxyphenyl)methylidene]hydrazin-1-yl}-2,3-dihydropyridazin-3-one) were done as described previously (Howes 2017). Nanoparticle protocells were developed and kindly provided by the lab of Dr. Eric Carnes.

MarvinSketch v.14.9.8.0 (ChemAxon (http://www.chemaxon.com)) was used for drawing chemical structures. Tanimoto similarity scores were calculated using the online tool ChemMine (Backman, Cao et al. 2011). All figures were made using a combination of Microsoft PowerPoint (version 16.0.6965.2117) and InkScape (version 0.92.0).

Cell Lines

Human liver (HEP3B) cancer cell lines were acquired from Dr. Walker Wharton at the University of New Mexico, and were grown in Dulbecco's Modified Eagle's Medium (DMEM), purchased from Corning Inc. supplemented with 10% fetal bovine serum (FBS) as well as the antibiotics penicillin and streptomycin. All cells were incubated at 37°C and 5% CO_2.

Protein Purification

Rosetta2 cells were transformed with plasmid DNA containing an ampicillin resistance gene, as well as C-terminal PKA-target, his- and flag-tagged LIG1 (**Fig. S5A**). An initial 2 mL culture in CircleGrow media (MP Biomedicals) containing ampicillin and chloramphenicol was scaled up to 250 mL and grown at 37°C. When the culture OD_{600} was measured at or around 0.7 cells were transferred to a 16°C. After 30 minutes at 16°C cells were induced with 0.2 mM iso-propyl-thio-galactooside (IPTG) for 16 hours. Bacteria were pelleted and lysed in 40 mM HEPES pH 7.5, 200 mM NaCl and 10% glycerol, plus protease inhibitors. Cell debris were removed by high-speed centrifugation, clarified lysate was loaded onto a 5 mL HisTrap HP column (GE Healthcare Life Sciences), and proteins were eluted by increasing concentration of imidazole stepwise. Ligase I containing fractions were purified using a HiTrap Q HP column (GE Healthcare Life Sciences), and eluted via NaCl. This procedure yields a high purity ligase with minimal degradation (**Fig. S5B**). Eluted ligase I fractions were pooled,

concentrated using a 50 kDa MWCO centrifugal filter (EMD Millipore), and stored at -80°C until needed. **Figure S5A** was generated using Serial Cloner, software version 2.6.1 (http://serialbasics.free.fr/Serial_Cloner.html).

Measuring Solubility

The solubility of L67 and L82 was measured in 100 mM solutions of Tris, HEPES, and PBS were made up at pH's of 7.2, 7.5, and 7.7, deionized water and 100% DMSO were used as controls. Precipitation was easily identifiable at 10x magnification. Each solution was examined under a microscope every 10 minutes for precipitation.

Compound Docking

The likelihood of small molecules to bind to DNA ligases was evaluated *in silico* using the OpenEye suite of software, VIDA, Make Receptor, Omega and OEDocking. The licenses for this software was used as part of collaboration with committee member Dr. Tudor Oprea's lab. Small molecules were generated in SMILES format using MarvinSketch (version 14.9.8.0), using an academic license. For any situation in which multiple stereoisomers were possible, all were used. Chemical binning and hierarchical clustering was done using the online tool ChemMine as well as calculation of Tanimoto similarity scores, and Figures were rendered with PyMol (version 1.3) (DeLano 2002). Small molecule structures were consolidated using VIDA (version 4.2.1), and multiconformers of

every small molecule were generated using Omega (version 2.5.1.4), with a maximum of 500 conformations per molecule. These multiconformer files were used by the FRED (Fast Rigid Exhaustive Docking) function of OEDocking (version 3.0.1) to predict each compound's ability to interact potential receptor sites on a given protein. Receptor sites were generated using Make Receptor (version 3.0.1), generally using 2-3 amino acids to define the potential receptor site. The results of the docking assay were exported as a spreadsheet, for combination and analysis in Microsoft Excel 2013 (version 15.0.4797.1003).

Cell Proliferation

HEP3B cells were cultured in 96-well plates with ligase inhibitors or 0.5% DMSO alone for five days at 37°C for use in an MTT assay. In the MTT assay, a tetrazolium dye, 3-(4,5-dimethylthiazol-2-yl)-2,5-diphenyltetrazolium bromide, is metabolized into (E,Z)-5-(4,5-dimethylthiazol-2-yl)-1,3-diphenylformazan in the mitochondria, during which its colour changes from yellow to purple. Cells were incubated with the MTT reagent (Promega) for one hour at 37°C according to the manufacturer's instructions. Absorbance at 570 nm was measured using a plate reader (PerkinElmer Victor 3V1420 Multilabel Counter). Cell viability is expressed as percentage of the value obtained with DMSO-treated cells.

Toxicity in bacteria was measured by growing in-house *E. coli* bacteria strain HI-1006 that had been transformed with the ampicillin containing plasmid pUC19. Plasmid DNA was generously provided by the lab of Dr. Osley of the

UNM Cancer Research Center. Cells were picked from a plate and grown in ampicillin containing circle grow media, before being divided into 1 mL cultures. The optical density at 600 nm (OD_{600}) was used to measure cell growth, and was measured hourly. DNA ligase inhibitors, L67, L82, L189, and L82-G17, were tested at 100μM; hydrogen peroxide and 1% DMSO (vehicle control) were used as position and negative controls, respectively.

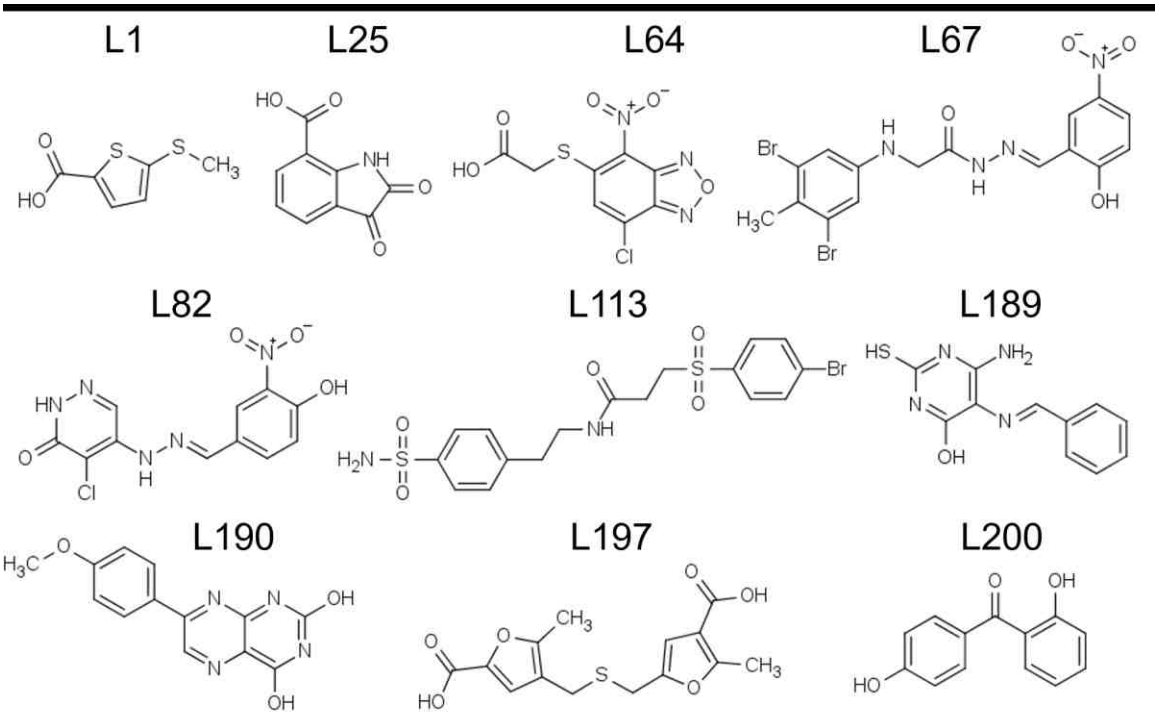
Cosmic Database

Data was accessed 1/28/2017 (Forbes, Beare et al. 2015).

Acknowledgements

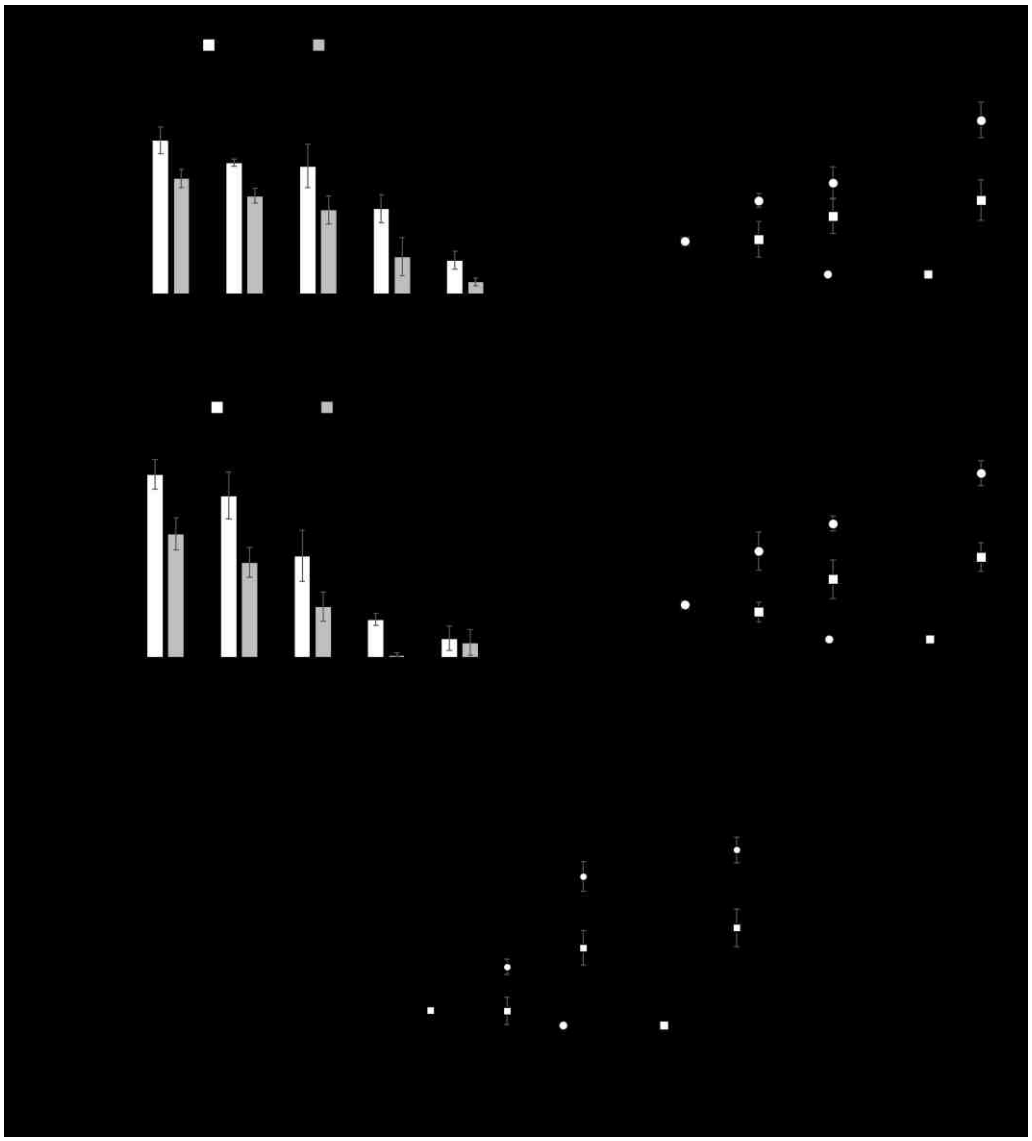
This work was supported by the University of New Mexico Comprehensive Cancer Center (P30 CA118100) and National Institute of Health Grants R01 GM57479 (to A.E.T.) and P01 CA92584.

Appendix – Supplemental Figures



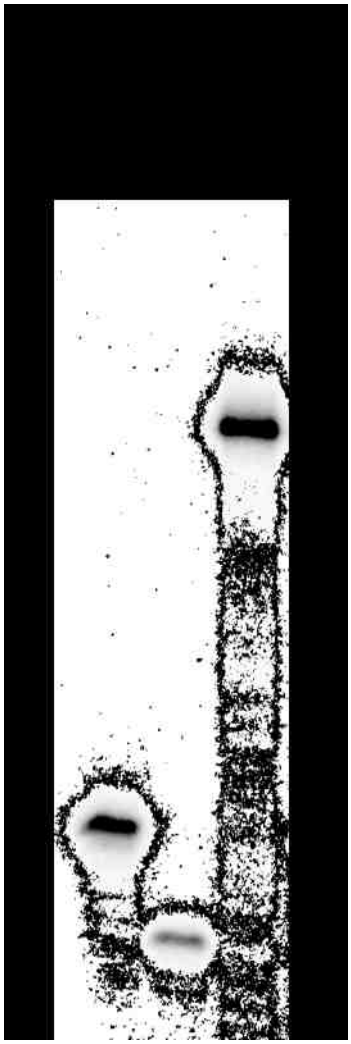
Supplemental Figure 1. Small-molecule inhibitors identified by computer aided drug design.

Previous work by the Tomkinson group published ten compounds that inhibit DNA ligase I, but not T4 ligase (Chen, Zhong et al. 2008, Zhong, Chen et al. 2008).



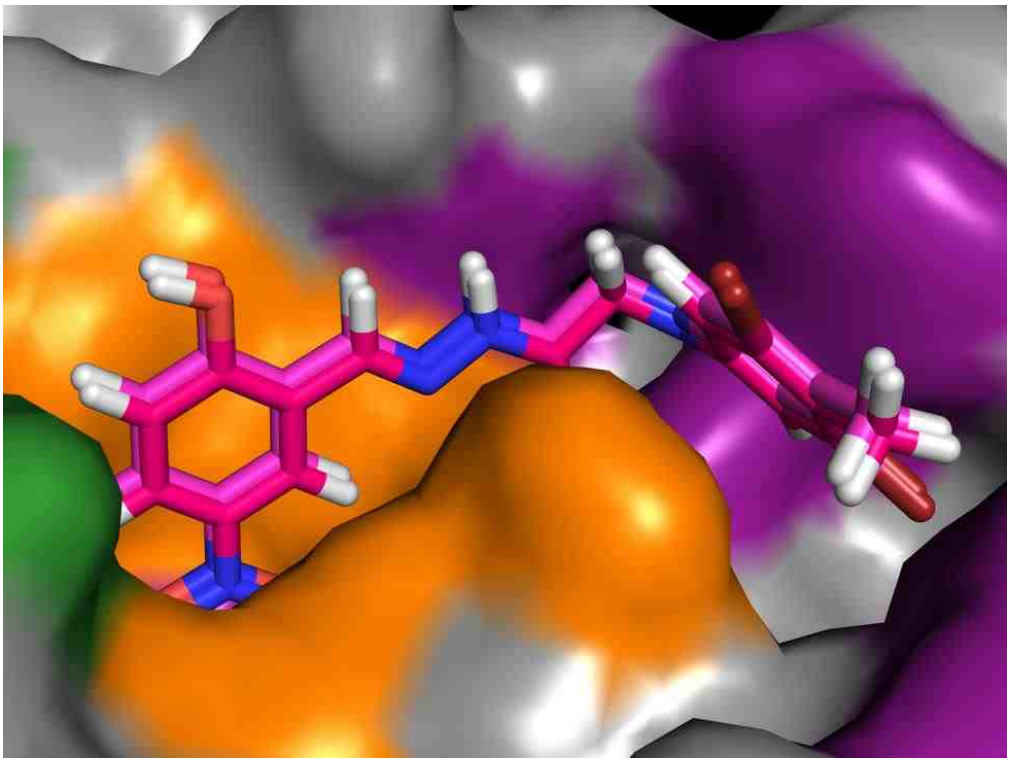
Supplemental Figure 2. Cells lacking LigI are more sensitive to L82 and L82-G17.

(Data in **Figure 8** was kindly generated by Dr. Annahita Sallmyr. The data presented here represents the data that I produced prior to that which was submitted for publication in DNA Repair.) Survival of wild-type (PF20) and *LIG1* null (PFL13) MEFs treated with (A) L82 or (B) L82-G17 for six days (*- $p<0.05$, **- $p<0.01$, ***- $p<0.005$) was measured as described in Materials and Methods. Results of at least three independent assays are shown graphically. γH2AX formation measured by flow cytometry in wild-type (PF20) and *LIG1* null (PFL13) MEFs incubated for 4 h (C) L82 or (D) L82-G17. Results of three independent assays are shown graphically. (E) The increase in γH2AX was also observed by immunofluorescence, following four hours of L82-G17 exposure.



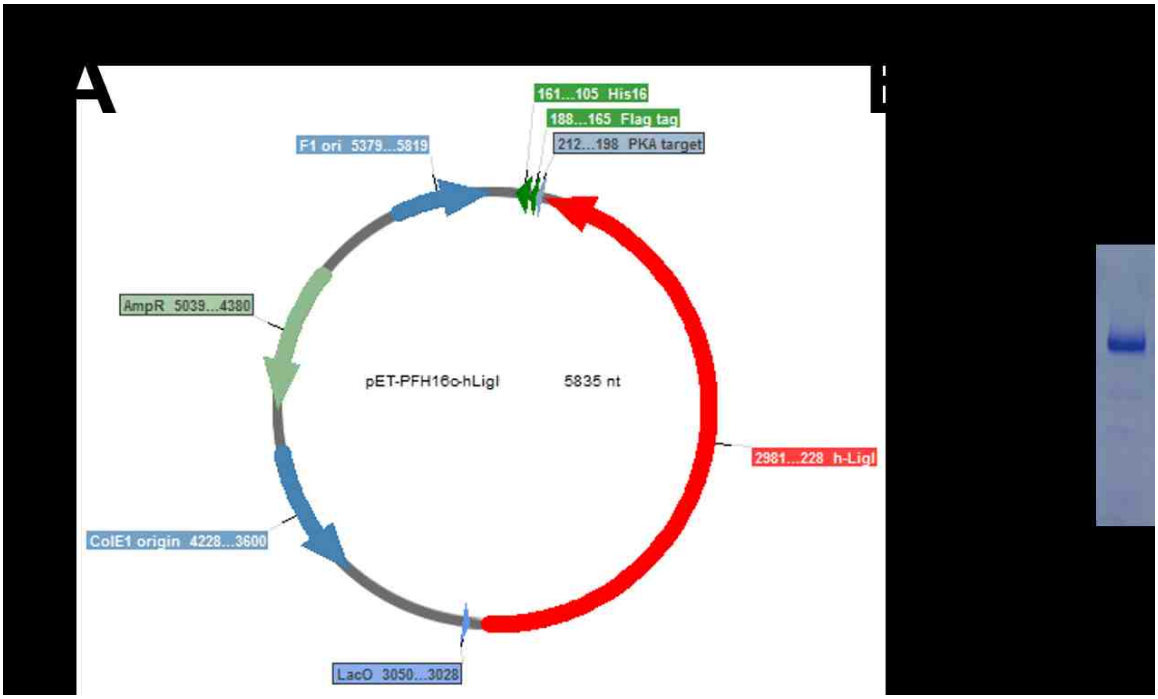
Supplemental Figure 3. Purity of ligation substrate oligos.

Each oligonucleotide, named TH_01, TH_02, and TH_03, were ordered from IDT as unpurified oligonucleotides and purified in house on a urea sequencing gel. Each oligo was then ^{32}P labeled and run again on a sequencing gel, to assess purity.



Supplemental Figure 4. L67 is not effected by DNA *in silico*

The predicted position and conformation of L67 binding does not change when modeled in the presence (lighter magenta) or absence (darker magenta) of DNA.



Supplemental Figure 5. ^{32}P labelable, His- and Flag-tagged DNA ligase I.

(A) Map of the plasmid used to express LigI that was labeled for DNA pulldown assays. (B) Results of two column purification of LigI.

REFERENCES

Altschul, S. F., W. Gish, W. Miller, E. W. Myers and D. J. Lipman (1990). "Basic local alignment search tool." J Mol Biol **215**(3): 403-410.

Arakawa, H., T. Bednar, M. Wang, K. Paul, E. Mladenov, A. A. Bencsik-Theilen and G. Iliakis (2012). "Functional redundancy between DNA ligases I and III in DNA replication in vertebrate cells." Nucleic Acids Res **40**(6): 2599-2610.

Arakawa, H. and G. Iliakis (2015). "Alternative Okazaki fragment ligation pathway by DNA ligase III." Genes **6**(2): 385-398.

Attiyeh, E. F., W. B. London, Y. P. Mossé, Q. Wang, C. Winter, D. Khazi, P. W. McGrady, R. C. Seeger, A. T. Look and H. Shimada (2005). "Chromosome 1p and 11q deletions and outcome in neuroblastoma." New England Journal of Medicine **353**(21): 2243-2253.

Audebert, M., B. Salles and P. Calsou (2004). "Involvement of poly(ADP-ribose) polymerase-1 and XRCC1/DNA ligase III in an alternative route for DNA double-strand breaks rejoining." J Biol Chem **279**(53): 55117-55126.

Backman, T. W., Y. Cao and T. Girke (2011). "ChemMine tools: an online service for analyzing and clustering small molecules." Nucleic Acids Res **39**(Web Server issue): W486-491.

Baell, J. B. and G. A. Holloway (2010). "New substructure filters for removal of pan assay interference compounds (PAINS) from screening libraries and for their exclusion in bioassays." J Med Chem **53**(7): 2719-2740.

Barnes, D. E., L. H. Johnston, K. Kodama, A. E. Tomkinson, D. D. Lasko and T. Lindahl (1990). "Human DNA ligase I cDNA: cloning and functional expression in Saccharomyces cerevisiae." Proc Natl Acad Sci U S A **87**(17): 6679-6683.

Barnes, D. E., A. E. Tomkinson, A. R. Lehmann, A. D. Webster and T. Lindahl (1992). "Mutations in the DNA ligase I gene of an individual with immunodeficiencies and cellular hypersensitivity to DNA-damaging agents." Cell **69**(3): 495-503.

Bentley, D., J. Selfridge, J. K. Millar, K. Samuel, N. Hole, J. D. Ansell and D. W. Melton (1996). "DNA ligase I is required for fetal liver erythropoiesis but is not essential for mammalian cell viability." Nat Genet **13**(4): 489-491.

Bentley, D. J., C. Harrison, A. M. Ketchen, N. J. Redhead, K. Samuel, M. Waterfall, J. D. Ansell and D. W. Melton (2002). "DNA ligase I null mouse cells show normal DNA repair activity but altered DNA replication and reduced genome stability." J Cell Sci **115**(Pt 7): 1551-1561.

Berman, H. M., J. Westbrook, Z. Feng, G. Gilliland, T. N. Bhat, H. Weissig, I. N. Shindyalov and P. E. Bourne (2000). "The protein data bank." Nucleic acids research **28**(1): 235-242.

Biasini, M., S. Bienert, A. Waterhouse, K. Arnold, G. Studer, T. Schmidt, F. Kiefer, T. Gallo Cassarino, M. Bertoni, L. Bordoli and T. Schwede (2014). "SWISS-MODEL: modelling protein tertiary and quaternary structure using evolutionary information." Nucleic Acids Res **42**(Web Server issue): W252-258.

Bienert, S., A. Waterhouse, T. A. de Beer, G. Tauriello, G. Studer, L. Bordoli and T. Schwede (2017). "The SWISS-MODEL Repository-new features and functionality." Nucleic Acids Res **45**(D1): D313-D319.

Bonner, W. M., C. E. Redon, J. S. Dickey, A. J. Nakamura, O. A. Sedelnikova, S. Solier and Y. Pommier (2008). "GammaH2AX and cancer." Nat Rev Cancer **8**(12): 957-967.

Bordoli, L., F. Kiefer, K. Arnold, P. Benkert, J. Battey and T. Schwede (2009). "Protein structure homology modeling using SWISS-MODEL workspace." Nat Protoc **4**(1): 1-13.

Braun, J. E., F. Tritschler, G. Haas, C. Igreja, V. Truffault, O. Weichenrieder and E. Izaurralde (2010). "The C-terminal α–α superhelix of Pat is required for mRNA decapping in metazoa." The EMBO journal **29**(14): 2368-2380.

Butler, K. S., P. N. Durfee, C. Theron, C. E. Ashley, E. C. Carnes and C. J. Brinker (2016). "Protocells: Modular Mesoporous Silica Nanoparticle-Supported Lipid Bilayers for Drug Delivery." Small.

Cardoso, M. C., C. Joseph, H. P. Rahn, R. Reusch, B. Nadal-Ginard and H. Leonhardt (1997). "Mapping and use of a sequence that targets DNA ligase I to sites of DNA replication in vivo." J Cell Biol **139**(3): 579-587.

Charkoudian, L. K., D. M. Pham and K. J. Franz (2006). "A pro-chelator triggered by hydrogen peroxide inhibits iron-promoted hydroxyl radical formation." J Am Chem Soc **128**(38): 12424-12425.

Chen, J., A. E. Tomkinson, W. Ramos, Z. B. Mackey, S. Danehower, C. A. Walter, R. A. Schultz, J. M. Besterman and I. Husain (1995). "Mammalian DNA ligase III: molecular cloning, chromosomal localization, and expression in spermatocytes undergoing meiotic recombination." Molecular and cellular biology **15**(10): 5412-5422.

Chen, J., A. E. Tomkinson, W. Ramos, Z. B. Mackey, S. Danehower, C. A. Walter, R. A. Schultz, J. M. Besterman and I. Husain (1995). "Mammalian DNA ligase III: molecular cloning, chromosomal localization, and expression in spermatocytes undergoing meiotic recombination." Mol Cell Biol **15**(10): 5412-5422.

Chen, X., J. D. Ballin, J. Della-Maria, M. S. Tsai, E. J. White, A. E. Tomkinson and G. M. Wilson (2009). "Distinct kinetics of human DNA ligases I, IIIalpha, IIIbeta, and IV reveal direct DNA sensing ability and differential physiological functions in DNA repair." DNA Repair (Amst) **8**(8): 961-968.

Chen, X., J. Pascal, S. Vijayakumar, G. M. Wilson, T. Ellenberger and A. E. Tomkinson (2006). "Human DNA ligases I, III, and IV-purification and new specific assays for these enzymes." Methods Enzymol **409**: 39-52.

Chen, X. and A. E. Tomkinson (2011). "Yeast Nej1 is a key participant in the initial end binding and final ligation steps of nonhomologous end joining." J Biol Chem **286**(6): 4931-4940.

Chen, X., S. Zhong, X. Zhu, B. Dziegielewska, T. Ellenberger, G. M. Wilson, A. D. MacKerell, Jr. and A. E. Tomkinson (2008). "Rational design of human DNA ligase inhibitors that target cellular DNA replication and repair." Cancer Res **68**(9): 3169-3177.

Cheng, J. M., J. L. Hiemstra, S. S. Schneider, A. Naumova, N.-K. V. Cheung, S. L. Cohn, L. Diller, C. Sapienza and G. M. Brodeur (1993). "Preferential amplification of the paternal allele of the n–myc gene in human neuroblastomas." Nature genetics **4**(2): 191-194.

Chumsri, S., T. Howes, T. Bao, G. Sabnis and A. Brodie (2011). "Aromatase, aromatase inhibitors, and breast cancer." The Journal of steroid biochemistry and molecular biology **125**(1): 13-22.

Chumsri, S., G. J. Sabnis, T. Howes and A. M. Brodie (2011). "Aromatase inhibitors and xenograft studies." Steroids **76**(8): 730-735.

Cotner-Gohara, E., I. K. Kim, M. Hammel, J. A. Tainer, A. E. Tomkinson and T. Ellenberger (2010). "Human DNA ligase III recognizes DNA ends by dynamic switching between two DNA-bound states." Biochemistry **49**(29): 6165-6176.

Cuneo, M. J., S. A. Gabel, J. M. Krahn, M. A. Ricker and R. E. London (2011). "The structural basis for partitioning of the XRCC1/DNA ligase III-α BRCT-mediated dimer complexes." Nucleic acids research: gkr419.

Das-Bradoo, S., H. D. Nguyen, J. L. Wood, R. M. Ricke, J. C. Haworth and A.-K. Bielinsky (2010). "Defects in DNA ligase I trigger PCNA ubiquitylation at Lys 107." Nature cell biology **12**(1): 74-79.

De Ioannes, P., S. Malu, P. Cortes and A. K. Aggarwal (2012). "Structural basis of DNA ligase IV-Artemis interaction in nonhomologous end-joining." Cell Rep **2**(6): 1505-1512.

DeLano, W. L. (2002). "The PyMOL molecular graphics system."

Doherty, A. J. and S. W. Suh (2000). "Structural and mechanistic conservation in DNA ligases." Nucleic acids research **28**(21): 4051-4058.

Doré, A. S., N. Furnham, O. R. Davies, B. L. Sibanda, D. Y. Chirgadze, S. P. Jackson, L. Pellegrini and T. L. Blundell (2006). "Structure of an Xrcc4–DNA ligase IV yeast ortholog complex reveals a novel BRCT interaction mode." DNA repair **5**(3): 362-368.

Drew, H. R., R. M. Wing, T. Takano, C. Broka, S. Tanaka, K. Itakura and R. E. Dickerson (1981). "Structure of a B-DNA dodecamer: conformation and dynamics." Proceedings of the National Academy of Sciences **78**(4): 2179-2183.

Ellenberger, T. and A. E. Tomkinson (2008). "Eukaryotic DNA ligases: Strucrural and Functional Insights." Annu. Rev. Biochem **In press**.

Ellenberger, T. and A. E. Tomkinson (2008). "Eukaryotic DNA ligases: Structural and functional insights." Annual Review of Biochemistry **In press**.

Ellenberger, T. and A. E. Tomkinson (2008). "Eukaryotic DNA ligases: structural and functional insights." Annu Rev Biochem **77**: 313-338.

Ferrari, G., R. Rossi, D. Arosio, A. Vindigni, G. Biamonti and A. Montecucco (2003). "Cell cycle-dependent phosphorylation of human DNA ligase I at the cyclin-dependent kinase sites." J Biol Chem **278**(39): 37761-37767.

Forbes, S. A., D. Beare, P. Gunasekaran, K. Leung, N. Bindal, H. Boutselakis, M. Ding, S. Bamford, C. Cole and S. Ward (2015). "COSMIC: exploring the world's knowledge of somatic mutations in human cancer." Nucleic acids research **43**(D1): D805-D811.

Frank, K. M., J. M. Sekiguchi, K. J. Seidl, W. Swat, G. A. Rathbun, H.-L. Cheng, L. Davidson, L. Kangaloo and F. W. Alt (1998). "Late embryonic lethality and impaired V (D) J recombination in mice lacking DNA ligase IV." Nature **396**(6707): 173-177.

Frank, K. M., N. E. Sharpless, Y. Gao, J. M. Sekiguchi, D. O. Ferguson, C. Zhu, J. P. Manis, J. Horner, R. A. DePinho and F. W. Alt (2000). "DNA ligase IV deficiency in mice leads to defective neurogenesis and embryonic lethality via the p53 pathway." Molecular cell **5**(6): 993-1002.

Frosina, G., P. Fortini, O. Rossi, F. Carrozzino, G. Raspaglio, L. S. Cox, D. P. Lane, A. Abbondandolo and E. Dogliotti (1996). "Two pathways for base excision repair in mammalian cells." J Biol Chem **271**(16): 9573-9578.

Frouin, I., A. Montecucco, G. Biamonti, U. Hubscher, S. Spadari and G. Maga (2002). "Cell cycle-dependent dynamic association of cyclin/Cdk complexes with human DNA replication proteins." Embo J **21**(10): 2485-2495.

Gajiwala, K. S. and C. Pinko (2004). "Structural rearrangement accompanying NAD+ synthesis within a bacterial DNA ligase crystal." Structure **12**(8): 1449-1459.

Gao, Y., S. Katyal, Y. Lee, J. Zhao, J. E. Rehg, H. R. Russell and P. J. McKinnon (2011). "DNA ligase III is critical for mtDNA integrity but not Xrcc1-mediated nuclear DNA repair." Nature **471**(7337): 240-244.

Girard, P.-M., B. Kysela, C. J. Härer, A. J. Doherty and P. A. Jeggo (2004). "Analysis of DNA ligase IV mutations found in LIG4 syndrome patients: the impact of two linked polymorphisms." Human molecular genetics **13**(20): 2369-2376.

Grawunder, U. and E. Harfst (2001). "How to make ends meet in V (D) J recombination." Current opinion in immunology **13**(2): 186-194.

Grawunder, U., D. Zimmer and M. R. Lieber (1998). "DNA ligase IV binds to XRCC4 via a motif located between rather than within its BRCT domains." Current biology **8**(15): 873-879.

Greco, G. E., Z. A. Conrad, A. M. Johnston, Q. Li and A. E. Tomkinson (2016). "Synthesis and structure determination of SCR7, a DNA ligase inhibitor." Tetrahedron Letters **57**(29): 3204-3207.

Greco, G. E., Y. Matsumoto, R. C. Brooks, Z. Lu, M. R. Lieber and A. E. Tomkinson (2016). "SCR7 is neither a selective nor a potent inhibitor of human DNA ligase IV." DNA repair **43**: 18-23.

Gulbis, J. M., Z. Kelman, J. Hurwitz, M. O'Donnell and J. Kuriyan (1996). "Structure of the C-terminal region of p21 WAF1/CIP1 complexed with human PCNA." Cell **87**(2): 297-306.

Hagemeijer, A., D. Bootsma, N. Spurr, N. Heisterkamp, J. Groffen and J. Stevenson (1982). "A cellular oncogene is translocated to the Philadelphia chromosome in chronic myelocytic leukemia." Nature **300**: 765.

Hakansson, K., A. J. Doherty, S. Shuman and D. B. Wigley (1997). "X-ray crystallography reveals a large conformational change during guanyl transfer by mRNA capping enzymes." Cell **89**(4): 545-553.

Hamosh, A., A. F. Scott, J. Amberger, D. Valle and V. A. McKusick (2000). "Online Mendelian inheritance in man (OMIM)." Human mutation **15**(1): 57.

Han, L., S. Masani, C. L. Hsieh and K. Yu (2014). "DNA ligase I is not essential for Mammalian cell viability." Cell Rep **7**(2): 316-320.

Han, S., J. S. Chang and M. Griffor (2009). "Structure of the adenylation domain of NAD+-dependent DNA ligase from Staphylococcus aureus." Acta Crystallographica Section F: Structural Biology and Crystallization Communications **65**(11): 1078-1082.

Harrison, C., A. M. Ketchen, N. J. Redhead, M. J. O'Sullivan and D. W. Melton (2002). "Replication failure, genome instability, and increased cancer susceptibility in mice with a point mutation in the DNA ligase I gene." Cancer Res **62**(14): 4065-4074.

Hellman, L. M. and M. G. Fried (2007). "Electrophoretic mobility shift assay (EMSA) for detecting protein–nucleic acid interactions." Nature protocols **2**(8): 1849-1861.

Henderson, L. M., C. F. Arlett, S. A. Harcourt, A. R. Lehmann and B. C. Broughton (1985). "Cells from an immunodeficient patient (46BR) with a defect in DNA ligation are hypomutable but hypersensitive to the induction of sister chromatid exchanges." Proc Natl Acad Sci U S A **82**(7): 2044-2048.

Hermans, A., N. Heisterkamp, M. von Lindern, S. van Baal, D. Meijer, D. van der Plas, L. M. Wiedemann, J. Groffen, D. Bootsma and G. Grosveld (1987). "Unique fusion of bcr and c-abl genes in Philadelphia chromosome positive acute lymphoblastic leukemia." Cell **51**(1): 33-40.

Horvath, D. (2011). "Pharmacophore-based virtual screening." Chemoinformatics and computational chemical biology: 261-298.

Howes, T. B., R; Jones, DE; Bologa, C; Yang, J; Matsumoto, Y; Oprea, TI; Tomkinson, AE (2017). "Prediction and Validation of an Inhibitor Binding Pocket on DNA Ligase I." (in preparation).

Howes, T. R. and A. E. Tomkinson (2012). DNA ligase I, the replicative DNA ligase. The Eukaryotic Replisome: a Guide to Protein Structure and Function, Springer**:** 327-341.

Howes, T. S., A; Brooks, R; Greco, GE; Jones, DE; Matsumoto, Y; Tomkinson, AE (2017). "Characterization of an uncompetitive inhibitor of DNA ligase I." (submitted).

IJspeert, H., A. Warris, M. Flier, I. Reisli, S. Keles, S. Chishimba, J. J. Dongen, D. C. Gent and M. Burg (2013). "Clinical spectrum of LIG4 deficiency is broadened with severe dysmaturity, primordial dwarfism, and neurological abnormalities." Human mutation **34**(12): 1611-1614.

Jackson, S. P. and T. Helleday (2016). "DNA REPAIR. Drugging DNA repair." Science **352**(6290): 1178-1179.

Jafari, R., H. Almqvist, H. Axelsson, M. Ignatushchenko, T. Lundbäck, P. Nordlund and D. M. Molina (2014). "The cellular thermal shift assay for evaluating drug target interactions in cells." Nature protocols **9**(9): 2100-2122.

John, F., J. George, S. V. Vartak, M. Srivastava, P. Hassan, V. Aswal, S. Karki and S. C. Raghavan (2015). "Enhanced efficacy of pluronic copolymer micelle encapsulated SCR7 against cancer cell proliferation." Macromolecular bioscience **15**(4): 521-534.

Johnston, L. H. and K. A. Nasmyth (1978). "Saccharomyces cerevisiae cell cycle mutant cdc9 is defective in DNA ligase." Nature **274**(5674): 891-893.

Johnston, L. H. and K. A. NASMYTH (1978). "Saccharomyces cerevisiae cell cycle mutant cdc9 is defective in DNA ligase." Nature **274**(5674): 891-893.

Johnston, S. R. and M. Dowsett (2003). "Aromatase inhibitors for breast cancer: lessons from the laboratory." Nature Reviews Cancer **3**(11): 821-831.

Jones, L. J., M. Gray, S. T. Yue, R. P. Haugland and V. L. Singer (2001). "Sensitive determination of cell number using the CyQUANT cell proliferation assay." J Immunol Methods **254**(1-2): 85-98.

Jun, S., Y.-S. Jung, H. N. Suh, W. Wang, M. J. Kim, Y. S. Oh, E. M. Lien, X. Shen, Y. Matsumoto and P. D. McCrea (2016). "LIG4 mediates Wnt signalling-induced radioresistance." Nature communications **7**.

Kibbe, W. A. (2007). "OligoCalc: an online oligonucleotide properties calculator." Nucleic Acids Res **35**(Web Server issue): W43-46.

Kim, D. J., O. Kim, H.-W. Kim, H. S. Kim, S. J. Lee and S. W. Suh (2009). "ATP-dependent DNA ligase from Archaeoglobus fulgidus displays a tightly closed conformation." Acta Crystallographica Section F: Structural Biology and Crystallization Communications **65**(6): 544-550.

Kodama, K.-i., D. E. Barnes and T. Lindahl (1991). "In vitro mutagenesis and functional expression in Escherichia coli of a cDNA encoding the catalytic domain of human DNA ligase I." Nucleic Acids Research **19**(22): 6093-6099.

Kodama, K., D. E. Barnes and T. Lindahl (1991). "In vitro mutagenesis and functional expression in Escherichia coli of a cDNA encoding the catalytic domain of human DNA ligase I." Nucleic Acids Res **19**(22): 6093-6099.

Kotnis, A. and R. Mulherkar (2014). "Novel inhibitor of DNA ligase IV with a promising cancer therapeutic potential." Journal of biosciences **3**(39): 339-340.

Krishna, S., D. K. Singh, S. Meena, D. Datta, M. I. Siddiqi and D. Banerjee (2014). "Pharmacophore-based screening and identification of novel human ligase I inhibitors with potential anticancer activity." J Chem Inf Model **54**(3): 781-792.

Krishnan, V. V., K. H. Thornton, M. P. Thelen and M. Cosman (2001). "Solution Structure and Backbone Dynamics of the Human DNA Ligase IIIα BRCT Domain†." Biochemistry **40**(44): 13158-13166.

Kubat, M., R. C. Holte and S. Matwin (1998). "Machine learning for the detection of oil spills in satellite radar images." Machine learning **30**(2-3): 195-215.

Kulczyk, A. W., J.-C. Yang and D. Neuhaus (2004). "Solution structure and DNA binding of the zinc-finger domain from DNA ligase IIIα." Journal of molecular biology **341**(3): 723-738.

Lakshmipathy, U. and C. Campbell (1999). "The human DNA ligase III gene encodes nuclear and mitochondrial proteins." Mol Cell Biol **19**(5): 3869-3876.

Lasko, D. D., A. E. Tomkinson and T. Lindahl (1990). "Mammalian DNA ligases. Biosynthesis and intracellular localization of DNA ligase I." J Biol Chem **265**(21): 12618-12622.

Le Chalony, C., F. Hoffschir, L. R. Gauthier, J. Gross, D. S. Biard, F. D. Boussin and V. Pennaneach (2012). "Partial complementation of a DNA ligase I deficiency by DNA ligase III and its impact on cell survival and telomere stability in mammalian cells." Cell Mol Life Sci **69**(17): 2933-2949.

Lee, J. Y., C. Chang, H. K. Song, J. Moon, J. K. Yang, H. K. Kim, S. T. Kwon and S. W. Suh (2000). "Crystal structure of NAD+-dependent DNA ligase: modular architecture and functional implications." The EMBO Journal **19**(5): 1119-1129.

Lee, Y., D. E. Barnes, T. Lindahl and P. J. McKinnon (2000). "Defective neurogenesis resulting from DNA ligase IV deficiency requires Atm." Genes & Development **14**(20): 2576-2580.

Lehman, I. R. (1974). "DNA ligase: structure, mechanism, and function." Science **186**(4166): 790-797.

Lehmann, A. R., A. E. Willis, B. C. Broughton, M. R. James, H. Steingrimsdottir, S. A. Harcourt, C. F. Arlett and T. Lindahl (1988). "Relation between the human fibroblast strain 46BR and cell lines representative of Bloom's syndrome." Cancer Res **48**(22): 6343-6347.

Levin, D. S., W. Bai, N. Yao, M. O'Donnell and A. E. Tomkinson (1997). "An interaction between DNA ligase I and proliferating cell nuclear antigen: implications for Okazaki fragment synthesis and joining." Proc Natl Acad Sci U S A **94**(24): 12863-12868.

Levin, D. S., W. Bai, N. Yao, M. O'Donnell and A. E. Tomkinson (1997). "An interaction between DNA ligase I and proliferating cell nuclear antigen: implications for Okazaki fragment synthesis and joining." Proceedings of the National Academy of Sciences **94**(24): 12863-12868.

Levin, D. S., A. E. McKenna, T. A. Motycka, Y. Matsumoto and A. E. Tomkinson (2000). "Interaction between PCNA and DNA ligase I is critical for joining of Okazaki fragments and long-patch base-excision repair." Curr Biol **10**(15): 919-922.

Levin, D. S., S. Vijayakumar, X. Liu, V. P. Bermudez, J. Hurwitz and A. E. Tomkinson (2004). "A conserved interaction between the replicative clamp loader and DNA ligase in eukaryotes: implications for Okazaki fragment joining." J Biol Chem **279**(53): 55196-55201.

Liang, L., L. Deng, S. C. Nguyen, X. Zhao, C. D. Maulion, C. Shao and J. A. Tischfield (2008). "Human DNA ligases I and III, but not ligase IV, are required for microhomology-mediated end joining of DNA double-strand breaks." Nucleic Acids Res **36**(10): 3297-3310.

Linstrom, P. J. and W. Mallard (2001). "NIST Chemistry webbook; NIST standard reference database No. 69."

Lipinski, C. A. (2000). "Drug-like properties and the causes of poor solubility and poor permeability." Journal of pharmacological and toxicological methods **44**(1): 235-249.

Mackenney, V. J., D. E. Barnes and T. Lindahl (1997). "Specific function of DNA ligase I in simian virus 40 DNA replication by human cell-free extracts is mediated by the amino-terminal non-catalytic domain." J Biol Chem **272**(17): 11550-11556.

Mills, S. D., A. E. Eakin, E. T. Buurman, J. V. Newman, N. Gao, H. Huynh, K. D. Johnson, S. Lahiri, A. B. Shapiro and G. K. Walkup (2011). "Novel bacterial NAD+-dependent DNA ligase inhibitors with broad-spectrum activity and antibacterial efficacy in vivo." Antimicrobial agents and chemotherapy **55**(3): 1088-1096.

Montecucco, A., M. Fontana, F. Focher, M. Lestingi, S. Spadari and G. Ciarrocchi (1991). "Specific inhibition of human DNA ligase adenylation by a distamycin derivative possessing antitumor activity." Nucleic acids research **19**(5): 1067-1072.

Montecucco, A., M. Fontana, F. Focher, M. Lestingi, S. Spadari and G. Ciarrocchi (1991). "Specific inhibition of human DNA ligase adenylation by a distamycin derivative possessing antitumor activity." Nucleic Acids Res **19**(5): 1067-1072.

Montecucco, A., R. Rossi, D. S. Levin, R. Gary, M. S. Park, T. A. Motycka, G. Ciarrocchi, A. Villa, G. Biamonti and A. E. Tomkinson (1998). "DNA ligase I is recruited to sites of DNA replication by an interaction with proliferating cell nuclear antigen: identification of a common targeting mechanism for the assembly of replication factories." EMBO J **17**(13): 3786-3795.

Montecucco, A., E. Savini, F. Weighardt, R. Rossi, G. Ciarrocchi, A. Villa and G. Biamonti (1995). "The N-terminal domain of human DNA ligase I contains the nuclear localization signal and directs the enzyme to sites of DNA replication." Embo J **14**(21): 5379-5386.

Moser, J., H. Kool, I. Giakzidis, K. Caldecott, L. H. Mullenders and M. I. Fousteri (2007). "Sealing of chromosomal DNA nicks during nucleotide excision repair requires XRCC1 and DNA ligase III alpha in a cell-cycle-specific manner." Mol Cell **27**(2): 311-323.

Murai, J., S. Y. Huang, B. B. Das, A. Renaud, Y. Zhang, J. H. Doroshow, J. Ji, S. Takeda and Y. Pommier (2012). "Trapping of PARP1 and PARP2 by Clinical PARP Inhibitors." Cancer Res **72**(21): 5588-5599.

Murphy-Benenato, K., H. Wang, H. M. McGuire, H. E. Davis, N. Gao, D. B. Prince, H. Jahic, S. S. Stokes and P. A. Boriack-Sjodin (2014). "Identification through structure-based methods of a bacterial NAD+-dependent DNA ligase inhibitor that avoids known resistance mutations." Bioorganic & medicinal chemistry letters **24**(1): 360-366.

Nagashima, T., Hayashi, F., Yokoyama, S. (2006). "PDB ID: 2E2W. Solution structure of the first BRCT domain of human DNA ligase IV.".

Nakamura, M., S. Kondo, M. Sugai, M. Nazarea, S. Imamura and T. Honjo (1996). "High frequency class switching of an IgM+ B lymphoma clone CH12F3 to IgA+ cells." Int Immunol **8**(2): 193-201.

Nandakumar, J., P. A. Nair and S. Shuman (2007). "Last stop on the road to repair: structure of E. coli DNA ligase bound to nicked DNA-adenylate." Molecular cell **26**(2): 257-271.

Nasmyth, K. A. (1977). "Temperature-sensitive lethal mutants in the structural gene for DNA ligase in the yeast Schizosaccharomyces pombe." Cell **12**(4): 1109-1120.

Nasmyth, K. A. (1977). "Temperature-sensitive lethal mutants in the structural gene for DNA ligase in the yeast Schizosaccharomyces pombe." Cell **12**(4): 1109-1120.

Natarajan, A., K. Dutta, D. B. Temel, P. A. Nair, S. Shuman and R. Ghose (2012). "Solution structure and DNA-binding properties of the phosphoesterase domain of DNA ligase D." Nucleic acids research **40**(5): 2076-2088.

Newman, E. A., F. Lu, D. Bashllari, L. Wang, A. W. Opipari and V. P. Castle (2015). "Alternative NHEJ Pathway Components Are Therapeutic Targets in High-Risk Neuroblastoma." Mol Cancer Res **13**(3): 470-482.

Nishida, H., S. Kiyonari, Y. Ishino and K. Morikawa (2006). "The closed structure of an archaeal DNA ligase from Pyrococcus furiosus." Journal of molecular biology **360**(5): 956-967.

Noguiez, P., D. E. Barnes, H. W. Mohrenweiser and T. Lindahl (1992). "Structure of the human DNA ligast I gene." Nucleic acids research **20**(15): 3845-3850.

Nussenzweig, A. and M. C. Nussenzweig (2007). "A backup DNA repair pathway moves to the forefront." Cell **131**(2): 223-225.

O'Driscoll, M., K. M. Cerosaletti, P.-M. Girard, Y. Dai, M. Stumm, B. Kysela, B. Hirsch, A. Gennery, S. E. Palmer and J. Seidel (2001). "DNA ligase IV mutations identified in patients exhibiting developmental delay and immunodeficiency." Molecular cell **8**(6): 1175-1185.

Ochi, T., X. Gu and T. L. Blundell (2013). "Structure of the catalytic region of DNA ligase IV in complex with an Artemis fragment sheds light on double-strand break repair." Structure **21**(4): 672-679.

Ochi, T., Q. Wu, D. Y. Chirgadze, J. G. Grossmann, V. M. Bolanos-Garcia and T. L. Blundell (2012). "Structural insights into the role of domain flexibility in human DNA ligase IV." Structure **20**(7): 1212-1222.

Odell, M., V. Sriskanda, S. Shuman and D. B. Nikolov (2000). "Crystal structure of eukaryotic DNA ligase-adenylate illuminates the mechanism of nick sensing and strand joining." Mol Cell **6**(5): 1183-1193.

Odell, M., V. Sriskanda, S. Shuman and D. B. Nikolov (2000). "Crystal structure of eukaryotic DNA ligase–adenylate illuminates the mechanism of nick sensing and strand joining." Molecular cell **6**(5): 1183-1193.

OEChem, T. (2012). "OpenEye Scientific Software." Inc., Santa Fe, NM, USA.

Oh, S., A. Harvey, J. Zimbric, Y. Wang, T. Nguyen, P. J. Jackson and E. A. Hendrickson (2014). "DNA ligase III and DNA ligase IV carry out genetically distinct forms of end joining in human somatic cells." DNA Repair (Amst) **21**: 97-110.

Okano, S., L. Lan, K. W. Caldecott, T. Mori and A. Yasui (2003). "Spatial and temporal cellular responses to single-strand breaks in human cells." Mol Cell Biol **23**(11): 3974-3981.

Okano, S., L. Lan, A. E. Tomkinson and A. Yasui (2005). "Translocation of XRCC1 and DNA ligase IIIalpha from centrosomes to chromosomes in response to DNA damage in mitotic human cells." Nucleic Acids Res **33**(1): 422-429.

Paietta, E., J. Racevskis, J. Bennett, D. Neuberg, P. Cassileth, J. Rowe and P. Wiernik (1998). "Biologic heterogeneity in Philadelphia chromosome-positive acute leukemia with myeloid morphology: the Eastern Cooperative Oncology Group experience." Leukemia **12**(12): 1881-1885.

Pandey, M., S. Kumar, G. Goldsmith, M. Srivastava, S. Elango, M. Shameem, D. Bannerjee, B. Choudhary, S. S. Karki and S. C. Raghavan (2017). "Identification and characterization of novel ligase I inhibitors." Mol Carcinog **56**(2): 550-566.

Pascal, J. M., P. J. O'Brien, A. E. Tomkinson and T. Ellenberger (2004). "Human DNA ligase I completely encircles and partially unwinds nicked DNA." Nature **432**(7016): 473-478.

Pascal, J. M., O. V. Tsodikov, G. L. Hura, W. Song, E. A. Cotner, S. Classen, A. E. Tomkinson, J. A. Tainer and T. Ellenberger (2006). "A Flexible Interface between DNA Ligase and PCNA Supports Conformational Switching and Efficient Ligation of DNA." Mol Cell **24**(2): 279-291.

Pascal, J. M., O. V. Tsodikov, G. L. Hura, W. Song, E. A. Cotner, S. Classen, A. E. Tomkinson, J. A. Tainer and T. Ellenberger (2006). "A flexible interface between DNA ligase and PCNA supports conformational switching and efficient ligation of DNA." Molecular cell **24**(2): 279-291.

Peng, X., X. Tang, W. Qin, W. Dou, Y. Guo, J. Zheng, W. Liu and D. Wang (2011). "Aroylhydrazone derivative as fluorescent sensor for highly selective recognition of Zn2+ ions: syntheses, characterization, crystal structures and spectroscopic properties." Dalton Trans **40**(19): 5271-5277.

Peng, Z., Z. Liao, B. Dziegielewska, Y. Matsumoto, S. Thomas, Y. Wan, A. Yang and A. E. Tomkinson (2012). "Phosphorylation of serine 51 regulates the interaction of human DNA ligase I with replication factor C and its participation in DNA replication and repair." J Biol Chem **287**(44): 36711-36719.

Peng, Z., Z. Liao, Y. Matsumoto, A. Yang and A. E. Tomkinson (2016). "Human DNA Ligase I Interacts with and Is Targeted for Degradation by the DCAF7 Specificity Factor of the Cul4-DDB1 Ubiquitin Ligase Complex." J Biol Chem **291**(42): 21893-21902.

Petrini, J. H., K. G. Huwiler and D. T. Weaver (1991). "A wild-type DNA ligase I gene is expressed in Bloom's syndrome cells." Proc Natl Acad Sci U S A **88**(17): 7615-7619.

Petrini, J. H., Y. Xiao and D. T. Weaver (1995). "DNA ligase I mediates essential functions in mammalian cells." Mol Cell Biol **15**(8): 4303-4308.

Petrova, T., E. Bezsudnova, B. Dorokhov, E. Slutskaya, K. Polyakov, P. Dorovatovskiy, N. Ravin, K. Skryabin, M. Kovalchuk and V. Popov (2012). "Expression, purification, crystallization and preliminary crystallographic analysis of a thermostable DNA ligase from the archaeon Thermococcus sibiricus." Acta Crystallographica Section F: Structural Biology and Crystallization Communications **68**(2): 163-165.

Pinko, C., Borchardt, A., Nikulin, V., Su, Y. (2008). PDB IDs: 3BAA, 3BAB, 3BAC, 3BA8, 3BA9. Structural Basis for the Inhibition of Bacterial NAD+ Dependent DNA Ligase. .

Plantaz, D., G. Mohapatra, K. K. Matthay, M. Pellarin, R. C. Seeger and B. G. Feuerstein (1997). "Gain of chromosome 17 is the most frequent abnormality detected in neuroblastoma by comparative genomic hybridization." The American journal of pathology **150**(1): 81.

Pommier, Y., M. J. O'Connor and J. de Bono (2016). "Laying a trap to kill cancer cells: PARP inhibitors and their mechanisms of action." Sci Transl Med **8**(362): 362ps317.

Prigent, C., D. D. Lasko, K. Kodama, J. R. Woodgett and T. Lindahl (1992). "Activation of mammalian DNA ligase I through phosphorylation by casein kinase II." Embo J **11**(8): 2925-2933.

Prigent, C., M. S. Satoh, G. Daly, D. E. Barnes and T. Lindahl (1994). "Aberrant DNA repair and DNA replication due to an inherited enzymatic defect in human DNA ligase I." Mol Cell Biol **14**(1): 310-317.

Puebla-Osorio, N., D. B. Lacey, F. W. Alt and C. Zhu (2006). "Early embryonic lethality due to targeted inactivation of DNA ligase III." Molecular and Cellular Biology **26**(10): 3935-3941.

Ranalli, T. A., M. S. DeMott and R. A. Bambara (2002). "Mechanism underlying replication protein a stimulation of DNA ligase I." J Biol Chem **277**(3): 1719-1727.

Rassool, F. V. and A. E. Tomkinson (2010). "Targeting abnormal DNA double strand break repair in cancer." Cellular and molecular life sciences **67**(21): 3699-3710.

Riballo, E., L. Woodbine, T. Stiff, S. A. Walker, A. A. Goodarzi and P. A. Jeggo (2009). "XLF-Cernunnos promotes DNA ligase IV-XRCC4 re-adenylation following ligation." Nucleic Acids Res **37**(2): 482-492.

Ricci, F., A. Tedeschi, E. Morra and M. Montillo (2009). "Fludarabine in the treatment of chronic lymphocytic leukemia: a review." Ther Clin Risk Manag **5**(1): 187-207.

Robins, P. and T. Lindahl (1996). "DNA ligase IV from HeLa cell nuclei." Journal of Biological Chemistry **271**(39): 24257-24261.

Sabnis, G. J., O. Goloubeva, S. Chumsri, N. Nguyen, S. Sukumar and A. M. Brodie (2011). "Functional activation of the estrogen receptor-α and aromatase by the HDAC inhibitor entinostat sensitizes ER-negative tumors to letrozole." Cancer research **71**(5): 1893-1903.

Sahota, G., S. Goldsmith-Fischman, B. Dixon, Y. Huang, J. Aramini, C. Yin, R. Xiao, A. Bhattacharya, D. Monleon and G. Swapna (2004). "Solution NMR structure of the BRCT domain from Thermus thermophilus DNA ligase: Surface features suggest novel intermolecular interactions." Proteins: Struct. Funct. Genetics.

Sallmyr, A., Y. Matsumoto, V. Roginskaya, B. Van Houten and A. E. Tomkinson (2016). "Inhibiting Mitochondrial DNA Ligase IIIalpha Activates Caspase 1-Dependent Apoptosis in Cancer Cells." Cancer Res **76**(18): 5431-5441.

Sallmyr, A., A. E. Tomkinson and F. Rassool (2008). "Up-regulation of WRN and DNA ligase IIIa in Chromic myeloid leukemia: Consequences for the repair of DNA double strand breaks." Blood **112**(4): 1413-1423.

Sanders, M. P., A. n. J. Barbosa, B. Zarzycka, G. A. Nicolaes, J. P. Klomp, J. de Vlieg and A. Del Rio (2012). "Comparative analysis of pharmacophore screening tools." Journal of chemical information and modeling **52**(6): 1607-1620.

Sangkook, L., L. Ik-Soo, C. Jingwen, P. LEITNER, J. M. BESTERMAN, D. A. KINGHORN and J. M. PEZZUTO (1996). "Natural-product inhibitors of human DNA ligase I." Biochemical Journal **314**(3): 993-1000.

Schirmer, R. E. (1990). Modern methods of pharmaceutical analysis, CRC press.

Shameem, M., R. Kumar, S. Krishna, C. Kumar, M. I. Siddiqi, B. Kundu and D. Banerjee (2015). "Synthetic modified pyrrolo[1,4] benzodiazepine molecules demonstrate selective anticancer activity by targeting the human ligase 1 enzyme: An in silico and in vitro mechanistic study." Chem Biol Interact **237**: 115-124.

Shuman, S., Y. Liu and B. Schwer (1994). "Covalent catalysis in nucleotidyl transfer reactions: essential motifs in Saccharomyces cerevisiae RNA capping enzyme are conserved in Schizosaccharomyces pombe and viral capping enzymes and among polynucleotide ligases." Proc Natl Acad Sci U S A **91**(25): 12046-12050.

Shuman, S. and B. Schwer (1995). "RNA capping enzyme and DNA ligase: a superfamily of covalent nucleotidyl transferases." Mol Microbiol **17**(3): 405-410.

Sibanda, B. L., S. E. Critchlow, J. Begun, X. Y. Pei, S. P. Jackson, T. L. Blundell and L. Pellegrini (2001). "Crystal structure of an Xrcc4–DNA ligase IV complex." Nature Structural & Molecular Biology **8**(12): 1015-1019.

Simsek, D., E. Brunet, S. Y. Wong, S. Katyal, Y. Gao, P. J. McKinnon, J. Lou, L. Zhang, J. Li, E. J. Rebar, P. D. Gregory, M. C. Holmes and M. Jasin (2011). "DNA ligase III promotes alternative nonhomologous end-joining during chromosomal translocation formation." PLoS Genet **7**(6): e1002080.

Simsek, D., A. Furda, Y. Gao, J. Artus, E. Brunet, A. K. Hadjantonakis, B. Van Houten, S. Shuman, P. J. McKinnon and M. Jasin (2011). "Crucial role for DNA ligase III in mitochondria but not in Xrcc1-dependent repair." Nature **471**(7337): 245-248.

Simsek, D. and M. Jasin (2011). "DNA ligase III: a spotty presence in eukaryotes, but an essential function where tested." Cell Cycle **10**(21): 3636-3644.

Singleton, M. R., K. Håkansson, D. J. Timson and D. B. Wigley (1999). "Structure of the adenylation domain of an NAD+-dependent DNA ligase." Structure **7**(1): 35-42.

Sirbu, B. M., F. B. Couch, J. T. Feigerle, S. Bhaskara, S. W. Hiebert and D. Cortez (2011). "Analysis of protein dynamics at active, stalled, and collapsed replication forks." Genes & development **25**(12): 1320-1327.

Soderhall, S. and T. Lindahl (1976). "DNA ligases of eukaryotes." FEBS Lett **67**(1): 1-8.

Song, W., D. S. Levin, J. Varkey, S. Post, V. P. Bermudez, J. Hurwitz and A. E. Tomkinson (2007). "A conserved physical and functional interaction between the cell cycle checkpoint clamp loader and DNA ligase I of eukaryotes." J Biol Chem.

Song, W., J. Pascal, T. Ellenberger and A. E. Tomkinson (2009). "The DNA bidning domain of human DNA ligase I inetracts with both nicked DNA and the DNA sliding clamps, PCNA and hRad9-hRad1-hHus1." DNA Repair (Amst): In press.

Soverini, S., G. Martinelli, G. Rosti, S. Bassi, M. Amabile, A. Poerio, B. Giannini, E. Trabacchi, F. Castagnetti and N. Testoni (2005). "ABL mutations in late chronic phase chronic myeloid leukemia patients with up-front cytogenetic resistance to imatinib are associated with a greater likelihood of progression to blast crisis and shorter survival: a study by the GIMEMA Working Party on Chronic Myeloid Leukemia." Journal of clinical oncology **23**(18): 4100-4109.

Soza, S., V. Leva, R. Vago, G. Ferrari, G. Mazzini, G. Biamonti and A. Montecucco (2009). "DNA ligase I deficiency leads to replication-dependent DNA damage and impacts cell morphology without blocking cell cycle progression." Mol Cell Biol **29**(8): 2032-2041.

Srivastava, M., M. Nambiar, S. Sharma, S. S. Karki, G. Goldsmith, M. Hegde, S. Kumar, M. Pandey, R. K. Singh and P. Ray (2012). "An inhibitor of nonhomologous end-joining abrogates double-strand break repair and impedes cancer progression." Cell **151**(7): 1474-1487.

Srivastava, S. K., R. P. Tripathi and R. Ramachandran (2005). "NAD+-dependent DNA ligase (Rv3014c) from mycobacterium tuberculosis crystal structure of the adenylation domain and identification of novel inhibitors." Journal of Biological Chemistry **280**(34): 30273-30281.

Staker, B. L., K. Hjerrild, M. D. Feese, C. A. Behnke, A. B. Burgin, Jr. and L. Stewart (2002). "The mechanism of topoisomerase I poisoning by a camptothecin analog." Proc Natl Acad Sci U S A **99**(24): 15387-15392.

Stokes, S. S., H. Huynh, M. Gowravaram, R. Albert, M. Cavero-Tomas, B. Chen, J. Harang, J. T. Loch, M. Lu and G. B. Mullen (2011). "Discovery of bacterial NAD+-dependent DNA ligase inhibitors: optimization of antibacterial activity." Bioorganic & medicinal chemistry letters **21**(15): 4556-4560.

Subramanya, H. S., A. J. Doherty, S. R. Ashford and D. B. Wigley (1996). "Crystal structure of an ATP-dependent DNA ligase from bacteriophage T7." Cell **85**(4): 607-615.

Subramanya, H. S., A. J. Doherty, S. R. Ashford and D. B. Wigley (1996). "Crystal structure of an ATP-dependent DNA ligase from bacteriophage T7." Cell **85**(4): 607-615.

Sun, D., R. Urrabaz, M. Nguyen, J. Marty, S. Stringer, E. Cruz, L. Medina-Gundrum and S. Weitman (2001). "Elevated expression of DNA ligase I in human cancers." Clin Cancer Res **7**(12): 4143-4148.

Surivet, J.-P., R. Lange, C. Hubschwerlen, W. Keck, J.-L. Specklin, D. Ritz, D. Bur, H. Locher, P. Seiler and D. S. Strasser (2012). "Structure-guided design, synthesis and biological evaluation of novel DNA ligase inhibitors with in vitro and in vivo anti-staphylococcal activity." Bioorganic & medicinal chemistry letters **22**(21): 6705-6711.

Teo, I. A., C. F. Arlett, S. A. Harcourt, A. Priestley and B. C. Broughton (1983). "Multiple hypersensitivity to mutagens in a cell strain (46BR) derived from a patient with immuno-deficiencies." Mutat Res **107**(2): 371-386.

Tobin, L. A., C. Robert, P. Nagaria, S. Chumsri, W. Twaddell, O. B. Ioffe, G. E. Greco, A. H. Brodie, A. E. Tomkinson and F. V. Rassool (2012). "Targeting abnormal DNA repair in therapy-resistant breast cancers." Mol Cancer Res **10**(1): 96-107.

Tobin, L. A., C. Robert, A. P. Rapoport, I. Gojo, M. R. Baer, A. E. Tomkinson and F. V. Rassool (2013). "Targeting abnormal DNA double strand break repair in tyrosine kinase inhibitor-resistant chronic meyloid leukemias." Oncogene **32**: 1784-1793.

Tom, S., L. A. Henricksen, M. S. Park and R. A. Bambara (2001). "DNA ligase I and proliferating cell nuclear antigen form a functional complex." J Biol Chem **276**(27): 24817-24825.

Tomkinson, A. E., T. R. Howes and N. E. Wiest (2013). "DNA ligases as therapeutic targets." Translational cancer research **2**(3).

Tomkinson, A. E., D. D. Lasko, G. Daly and T. Lindahl (1990). "Mammalian DNA ligases. Catalytic domain and size of DNA ligase I." J Biol Chem **265**(21): 12611-12617.

Tomkinson, A. E., N. F. Totty, M. Ginsburg and T. Lindahl (1991). "Location of the active site for enzyme-adenylate formation in DNA ligases." Proc Natl Acad Sci U S A **88**(2): 400-404.

Tomkinson, A. E., N. F. Totty, M. Ginsburg and T. Lindahl (1991). "Location of the active site for enzyme-adenylate formation in DNA ligases." Proceedings of the National Academy of Sciences **88**(2): 400-404.

Tomkinson, A. E., S. Vijayakumar, J. M. Pascal and T. Ellenberger (2006). "DNA ligases: structure, reaction mechanism, and function." Chem Rev **106**(2): 687-699.

Unciuleac, M.-C., Y. Goldgur and S. Shuman (2017). "Two-metal versus one-metal mechanisms of lysine adenylylation by ATP-dependent and NAD+-dependent polynucleotide ligases." Proceedings of the National Academy of Sciences **114**(10): 2592-2597.

Ververis, K. and T. C. Karagiannis (2012). "Overview of the classical histone deacetylase enzymes and histone deacetylase inhibitors." ISRN Cell Biology **2012**.

Vijayakumar, S., B. R. Chapados, K. H. Schmidt, R. D. Kolodner, J. A. Tainer and A. E. Tomkinson (2007). "The C-terminal domain of yeast PCNA is required for physical and functional interactions with Cdc9 DNA ligase." Nucleic acids research **35**(5): 1624-1637.

Vijayakumar, S., B. R. Chapados, K. H. Schmidt, R. D. Kolodner, J. A. Tainer and A. E. Tomkinson (2007). "The C-terminal domain of yeast PCNA is required for physical and functional interactions with Cdc9 DNA ligase." Nucleic Acids Res **35**(5): 1624-1637.

Vijayakumar, S., B. Dziegielewska, D. S. Levin, W. Song, J. Yin, A. Yang, Y. Matsumoto, V. P. Bermudez, J. Hurwitz and A. E. Tomkinson (2009). "Phosphorylation of human DNA ligase I regulates its interaction with replication factor C and its participation in DNA replication and DNA repair." Mol Cell Biol **29**(8): 2042-2052.

Wang, H., B. Rosidi, R. Perrault, M. Wang, L. Zhang, F. Windhofer and G. Iliakis (2005). "DNA ligase III as a candidate component of backup pathways of nonhomologous end joining." Cancer Res **65**(10): 4020-4030.

Wang, T., Charifson, P., Xu, W., Wei, Y. (2013). "PDB ID: 4EFB. Crystal structure of DNA ligase.".

Wang, W., L. A. Lindsey-Boltz, A. Sancar and R. A. Bambara (2006). "Mechanism of stimulation of human DNA ligase I by the Rad9-Rad1-Hus1 checkpoint complex." J Biol Chem.

Waterworth, W. M., J. Kozak, C. M. Provost, C. M. Bray, K. J. Angelis and C. E. West (2009). "DNA ligase 1 deficient plants display severe growth defects and delayed repair of both DNA single and double strand breaks." BMC plant biology **9**(1): 79.

Waterworth, W. M., G. Masnavi, R. M. Bhardwaj, Q. Jiang, C. M. Bray and C. E. West (2010). "A plant DNA ligase is an important determinant of seed longevity." The Plant Journal **63**(5): 848-860.

Webster, A. D., D. E. Barnes, C. F. Arlett, A. R. Lehmann and T. Lindahl (1992). "Growth retardation and immunodeficiency in a patient with mutations in the DNA ligase I gene." Lancet **339**(8808): 1508-1509.

Wei, Y. F., P. Robins, K. Carter, K. W. Caldecott, D. J. C. Papin, G.-L. Yu, R.-P. Wang, B. K. Shell, R. A. Nash, P. Schar, D. E. Barnes, W. A. Haseltine and T. Lindahl (1995). "Molecular cloning and

expression of human cDNAs encoding a novel DNA ligase IV and DNA igase III, an enzyme active in DNA repair and genetic recombination." Mol. Cell. Biol. **15**: 3206-3216.

Wu, P.-Y., P. Frit, S. Meesala, S. Dauvillier, M. Modesti, S. N. Andres, Y. Huang, J. Sekiguchi, P. Calsou and B. Salles (2009). "Structural and functional interaction between the human DNA repair proteins DNA ligase IV and XRCC4." Molecular and cellular biology **29**(11): 3163-3172.

Zhong, S., X. Chen, X. Zhu, B. Dziegielewska, K. E. Bachman, T. Ellenberger, J. D. Ballin, G. M. Wilson, A. E. Tomkinson and A. D. MacKerell, Jr. (2008). "Identification and validation of human DNA ligase inhibitors using computer-aided drug design." J Med Chem **51**(15): 4553-4562.