

The Unique Role of Polymerase μ in Nonhomologous End Joining

Bryan Joseph Davis

A dissertation submitted to the faculty of the University of North Carolina at Chapel Hill in partial fulfillment of the requirements for the degree of Doctor of Philosophy in the Department of Biochemistry and Biophysics.

Chapel Hill
2008

Approved by:

Dale A. Ramsden, Advisor

Stephen G. Chaney, Reader

Thomas A. Kunkel, Reader

Marila Cordeiro-Stone, Reader

Aziz Sancar, Reader

© 2008
Bryan Joseph Davis
ALL RIGHTS RESERVED

ABSTRACT

**BRYAN JOSEPH DAVIS: The Unique Role of Polymerase μ in Nonhomologous
End Joining**

(Under the direction of Dale A. Ramsden)

Nonhomologous end-joining (NEHJ) has the challenge of repairing DNA double strand breaks (DSBs) without using an intact template molecule to instruct synthesis. NHEJ solves this problem in part by recruiting polymerases from the pol X family to fill gaps that are present at some DSBs. If a gap has some complementary bases, multiple pol X family members can participate in NHEJ, but only pol μ is able to fill the gap if the DNA ends are noncomplementary. I determined that pol μ fills these gaps in a template-directed way, successfully using as template a DNA strand that is not annealed to the primer. My results show pol μ is able to do this because it interacts with DNA on both sides of the gap. Further, I demonstrate that the polymerase is most efficient when it participates in such an end-bridging complex. I then determined the roles of two amino acid residues in pol μ 's activity on noncomplementary ends. As a similar residue does in TdT, H329 helps pol μ stabilize the incoming nucleotide when the template strand does not. However, unlike TdT, I show pol μ is not promiscuously template independent. I argue this is in part because pol μ also possesses an element of the more canonical template dependent pol X family members. R175 is part of the 8 kDa domain and helps position the template opposite the primer terminus through an interaction with the downstream phosphate. Pol μ is the only polymerase that has both of these elements and the combination confers the unique ability to fill gaps between noncomplementary ends.

ACKNOWLEDGMENTS

I would first like to thank Dale for his support and encouragement during my graduate career. He makes his lab a great place to do science, and I am thankful I got to learn from him. In addition, his enthusiasm for college basketball fanned my interest in the sport. I hope in return that my explanations of the subtleties of NASCAR have had something to do with his new appreciation for that version of auto racing. I also would like to say thanks to Martin and Steve for their contributions to the atmosphere of the lab. I can't imagine I'll ever again work with people who understand me the way these guys do. I would be remiss not to express my gratitude to Eli as well, not only for his tremendous technical support, but also for being a superb roommate for so many years. I thank my parents Ed and Kathie, and brothers Trevor, Kevin and Casey for their love, encouragement and advice. They provide the foundation for my life, and I'm very lucky to have them. Finally, I thank my fiancée Meredith. She has shown me more patience and support in the last month than I could have expected in an entire year. I love her very much and I'm looking forward to sharing life with her more than she knows.

TABLE OF CONTENTS

LIST OF TABLES	viii
----------------	------

LIST OF FIGURES	ix
-----------------	----

1 Introduction	1
1.1 DNA Double Strand Break Repair	1
1.2 V(D)J recombination	2
1.3 Nonhomologous End Joining Core Factors	3
1.4 Nonhomologous End Joining Processing Factors	3
1.5 Pol X family	4
1.5.1 Structure	4
1.5.2 Non-mammalian pol X Family Members	5
1.5.3 Mammalian pol X Family Members	6
2 Efficient pol μ activity is template directed during NHEJ of noncomplementary ends	15
2.1 Introduction	15
2.2 Methods	17
2.2.1 Expression Constructs and Purified Proteins	17
2.2.2 <i>In Vitro</i> Assays	17
2.3 Results	19

2.3.1	Pol μ typically adds complementary nucleotides more efficiently than non-complementary nucleotides	19
2.3.2	End-bridging interactions help Pol μ fill short gaps between two ends	23
2.4	Discussion	24
2.5	Acknowledgements	26
3	Structural elements in pol μ important for its unique activity	31
3.1	Introduction	31
3.2	Materials and Methods	33
3.2.1	Expression Constructs and Purified Proteins	33
3.2.2	End Joining Assays	33
3.2.3	Polymerase activity assays	34
3.2.4	Analysis of Igk recombination in pol μ R175A over- expressing cells	35
3.3	Results	36
3.3.1	H329 is required for pol μ activity when the primer does not interact with the template	36
3.3.2	R175 is important for end-bridging interactions in pol μ	38
3.4	Discussion	41
3.5	Acknowledgements	42
4	Discussion	48
4.1	Pol μ 's activity in NHEJ	48
4.2	Polymerase choice	52
4.3	Concluding Remarks	54

LIST OF TABLES

2.1	Sequences of junctions from reactions noted with asterisks in Figure 2.3 A.	30
3.1	Steady-state kinetic analysis of nucleotide incorporation.	47
3.2	Polymerase activity on oligonucleotide substrates.	47

LIST OF FIGURES

1.1	Homologous Recombination vs. Nonhomologous End Joining.	9
1.2	V(D)J Recombination.	10
1.3	Domain Maps of Pol X Family Members.	11
1.4	Pol X family polymerases bend DNA 90°.	12
1.5	Gradient of Template Dependence in the Pol X Family.	13
1.6	Loop 1 in Pol X Family Polymerases.	14
2.1	Factors required for joining of noncomplementary ends.	27
2.2	Effect of overhang sequence on pol μ activity in NHEJ.	28
2.3	Effect of overhang length on pol μ activity in NHEJ.	29
3.1	Role of Pol μ His329 in template-dependent synthesis during nonhomologous end-joining.	44
3.2	Activity of pol μ R175A <i>in vitro</i>	45
3.3	Activity of pol μ R175A in cells.	46
4.1	H208 is predicted to stabilize the 5'-phosphate of the downstream strand.	56
4.2	Pol λ forms the more stable complex with NHEJ core factors.	57

Chapter 1

Introduction

1.1 DNA Double Strand Break Repair

DNA double strand breaks (DSBs) are the form of damage most dangerous to cells. Left unrepaired, these breaks can result in cell death or genomic instability that may ultimately lead to cancer. There are many causes of DSBs, including ionizing radiation and normal cellular processes such as meiosis and V(D)J recombination. Some DSBs result in noncompatible DNA ends that are damaged or contain gaps. Repair cannot be completed until these have been resolved.

Cells repair DSBs in two ways: homologous recombination (HR) and nonhomologous end joining (NHEJ) (Figure 1.1) (reviewed in [1]). In HR, an intact sister chromatid or homologous chromosome is used as a template for repair. A resection step helps ensure damaged ends are removed and subsequent DNA synthesis using the undamaged template molecule ensures restoration of the genetic information that may have been lost due to the DSB (reviewed in, for example [2, 3]). NHEJ, however, is the only repair process that operates without an intact template (reviewed in, for example, [4, 5]). The two DNA ends that result from the break are either minimally or not processed at all before they are religated, potentially leading to less accurate repair than would result from HR [5, 6]. Despite the fact that NHEJ activity is therefore potentially

mutagenic, it is active in all stages of the cell cycle, while HR is used primarily during S and G2 phases [7]. Because of its constant activity, it is important for NHEJ to make noncompatible ends ligatable, though the lack of an intact template strand makes this a difficult proposition. In this thesis I describe how I determined that a unique activity in polymerase μ helps NHEJ solve this problem and retain as much genetic information as possible.

1.2 V(D)J recombination

NHEJ is the DSB repair pathway used in V(D)J recombination (reviewed in [8]). Three pol X family polymerases (described below) have been implicated in the process and each has a different effect on the character of the repair [9, 10, 11]. During V(D)J recombination, the adaptive immune system generates its repertoire of diverse antigen receptors by rearranging multiple gene segments (reviewed in [12]). This rearrangement requires recombination activating genes 1 and 2. Their protein products, RAG1 and RAG2, cleave DNA segments at conserved recombination signal sequences (RSSs), and like any DSB, those they create must be repaired (Figure 1.2). The DNA ends which contain the RSSs (signal ends) have blunt ends after RAG cleavage and are typically circularized by blunt-end ligation and lost from the genome. The coding sequence ends have terminal hairpin structures that must be processed before ligation can occur. Activity by NHEJ core factors and processing factors repairs these DSBs. This repair process is flexible: my thesis specifically addresses differences in three polymerases that can be used by NHEJ, as well as how these differences impact the way the break is repaired.

1.3 Nonhomologous End Joining Core Factors

In order to function in mammals, NHEJ minimally requires several factors, which are capable, as a group, of repairing DSBs with fully complementary overhangs. These “core factors” are: Ku, the obligate heteroligomer XRCC4-Ligase IV (XL), and DNA Protein Kinase catalytic subunit (DNA-PKcs). These core factors align DNA ends to make the best use of any complementary nucleotides present and are capable of immediately ligating fully complementary overhangs (reviewed in [13]). Ku is a heterodimer of 70 and 80 kDa subunits that localizes to DNA ends and recruits other core factors [14, 15, 16, 17]. DNA-PKcs regulates access to the ends and helps Ku align the ends together [18, 19]. This regulation of access to the ends protects them from inappropriate activity of other enzymes which might degrade the ends. XL ligates the ends together [14, 20]. In mice, knock-outs of XRCC4 and LigaseIV are lethal during embryonic development, while loss of the other factors results in a severe combined immunodeficient and radiosensitive phenotype because of the inability to repair DSBs (reviewed in [13]). Some DSBs, however, cannot be repaired by the core factors alone.

1.4 Nonhomologous End Joining Processing Factors

DNA ends that are not compatible require the NHEJ core factors to recruit processing factors to modify the ends so they become ligatable (reviewed in [13, 21, 22]). The DNA ends may have damaged or missing bases, or termini with 3' phosphates (instead of 3' hydroxyl groups) or 5' hydroxyl groups (rather than 5' phosphates). Among the enzymes that can be recruited by the NHEJ core factors to make these structures amenable to ligation are polynucleotide kinase (PNK), the Artemis endonuclease, and polymerases of the pol X family [23, 24, 11, 5]. PNK has 3' phosphatase and 5' kinase activities and outside of NHEJ it functions in single-strand break repair, but it also

has been found to interact with XRCC4 [23]. Disruption of that interaction results in diminished NHEJ activity *in vitro* and slower repair of DSBs and higher sensitivity to ionizing radiation *in vivo*.

By itself, the Artemis nuclease exhibits 5'-3' exonuclease activity, but following phosphorylation by, and in a complex with, DNA-PKcs it is also capable of endonuclease activity [25]. This endonuclease activity is essential for proper resolution of V(D)J recombination induced DSBs because it opens the hairpin structures formed by the RAG proteins at coding ends. Absence of Artemis results in a radio-sensitive severe combined immunodeficient (RS-SCID) phenotype similar to that seen in the absence of DNA-PKcs in humans and in mice [24]. In general DSB repair, it may be that Artemis' nuclease activity is important for removing damaged bases before ligation takes place.

DNA overhangs that result either directly from damage or from Artemis activity may not be ligatable if the DNA sequence is not fully complementary. NHEJ is able to bypass this potential block to repair by the use of polymerases that can fill in gaps that remain after any complementary sequence has been aligned. These polymerases are members of the pol X family.

1.5 Pol X family

1.5.1 Structure

The pol X family consists of relatively small polymerases that lack 3'-5' exonuclease (proofreading) activity and have similar overall structural characteristics (reviewed in, for example, [26]). In addition to the “palm”, “fingers” and “thumb” domains common to many polymerases, each also has an “8 kDa” domain (so called because of the mass of the domain in the prototypical family member polymerase β (pol β)) which has lyase activity in mammalian pol β , pol λ , and *Saccharomyces cerevisiae* pol4 [27, 28, 29]

(Figure 1.3). In all the pol X family members that use a template for synthesis, this domain interacts with the downstream strand of a gapped DNA substrate that is to be repaired and helps with the overall positioning of the polymerase on the DNA substrate (Figure 1.3) [26]. While the pol X family polymerases contact the downstream strand with their 8 kDa domains, they have relatively few contacts with the upstream portions of the DNA duplex. In all of the family members, three aspartate residues position two magnesium ions which are essential for catalysis: they help position the attacking 3'-OH group on the primer strand next to the incoming nucleotide and further help activate the incoming nucleotide for phosphoryl transfer. Notably, the pol X family polymerases bend the template strand into a 90° angle within the polymerase, which results in a significant distance between the primer and downstream strand termini at the time of catalysis (Figure 1.4) [30, 31, 32]. Functionally speaking, these polymerases fill short gaps in DNA and are recruited for distinct cellular processes, including base excision repair (BER), V(D)J recombination and general DSB repair ([33] and reviewed in [34, 11]).

1.5.2 Non-mammalian pol X Family Members

At least one pol X family member has been found in all eukaryotes except protostomes (e.g. *Drosophila melanogaster* and *Caenorhabditis elegans*) [35]. Frequently these polymerases possess an N-terminal BRCT domain (Figure 1.3) and might therefore be considered pol λ orthologs though polymerases with similarities to all mammalian pol X family members can be found (reviewed in [35]). It has been proposed that these orthologs have their evolutionary root in a pol λ -like polymerase with functions in both BER and NHEJ. Subsequent evolution likely resulted in variants (pol μ , TdT and pol β) that are significantly more specialized with respect to both activity and biological role (detailed below).

S. cerevisiae have only one pol X family member, pol4, which is considered a pol λ ortholog. It associates with dnl4/lif1, the yeast counterpart to XL, and participates in gap filling during repair of DSBs [36]. Yeast lacking this polymerase show a marked decrease in NHEJ of ends that have 3' or 5' overhanging mismatches and increased sensitivity to DSBs induced by the HO endonuclease when homologous recombination has also been knocked out [37, 36].

1.5.3 Mammalian pol X Family Members

Mammals have not one but four pol X family members encoded in their genomes. While they share a great deal of structural and functional similarity, there remain significant differences between these family members, and this is the primary focus of my thesis. One family member, pol β , lacks an entire domain possessed by the other three family members, a BRCT protein-protein interaction domain (reviewed in [26]). This domain allows terminal deoxyribonucleotidyltransferase (TdT), pol λ and pol μ to associate with the NHEJ core factors; deletion of the BRCT domain mostly abolishes this association (reviewed in [11, 5, 38, 39, 40]). The BRCT domains of TdT and pol μ are highly similar, and mostly distinct from the BRCT domain of pol λ [41, 42, 26].

Among the polymerases recruited to NHEJ core factors, there are clear functional differences. TdT is a template independent polymerase expressed only during V(D)J recombination in developing lymphocytes [9]. Its template independent activity adds nucleotides to the forming antigen receptors, contributing to the extremely wide range of antigens recognized by the immune system. TdT knock-out mice have immature immune systems, lacking diverse antigen receptors in B- and T-cells [43].

Pol μ and pol λ each have roles at different points in V(D)J recombination, but they are also expressed in many cell types and facilitate general DSB repair [11, 44, 9, 39]. Pol μ promotes retention of overhanging sequence in $\nu\kappa$ junctions during V(D)J

recombination, and as a result mice lacking pol μ have a deficiency in B-cells, while pol λ has a role in heavy-chain recombination [45, 9, 46]. Mice without pol λ have B-cells with shorter coding segments from heavy chain rearrangement during V(D)J recombination than their wild-type counterparts. Mouse embryonic fibroblasts (MEFs) lacking pol λ exhibit an increased sensitivity to oxidative damage [47]. Mice deficient in both pol μ and pol λ are at most mildly sensitive to ionizing radiation [48]. Pol β , though not specifically recruited by the NHEJ core factors through a BRCT domain participates in base excision repair and may contribute to general DSB repair in certain contexts as well. MEFs with pol β knocked out are deficient in BER and more sensitive to oxidative damage than wt MEFs [49, 50]. Additionally, pol λ has been implicated as a potential back-up polymerase to pol β in BER [47].

Our group previously hypothesized these polymerases collectively exhibit a gradient of template dependence (Figure 1.5) [45]. Pol β is the most template dependent and is generally unable to tolerate template strand breaks. Pol λ will tolerate a discontinuous template, but still requires base pairing interactions between the primer and template strands. Like pol λ , pol μ can tolerate a break in the template, but unlike pol λ , pol μ does not require pairing interactions between the primer and template. Importantly, I show in chapter 2 this polymerase nevertheless is mostly template instructed during synthesis, as canonical polymerases (including pol β and pol λ) are. In contrast, TdT is completely template independent and will disregard a template strand if one is present.

There are at least three structural features that contribute to this gradient in the pol X family. As previously described by our group and others, a loop region (termed Loop 1) is found in different lengths in each of the pol X family members (Figure 1.6) [45, 51]. The most template dependent, pol β , has only a short turn rather than a loop, while TdT, the least template dependent, has the longest Loop 1. It has been suggested this loop can sit in the space where a template strand should go in the polymerase active

site, either displacing the template strand (if any is present) in the case of TdT, or providing stabilizing interactions with the primer strand, in the cases of pol μ and pol λ [51, 52].

In Chapter 3 I describe contributions from two additional structural features. Both features are amino acid residues highlighted by the recently solved crystal structure of pol μ : H329 and R175. H329 was predicted by the structure to interact with the primer terminus and the incoming nucleotide [32]. R175 is located within the 8 kDa region and was predicted to interact with the 5' phosphate on the downstream strand [32]. This region is found in all pol X family members, but the residues in the region are more positively charged in pol β and pol λ , less positively charged in pol μ , and least positively charged in TdT (discussed in greater detail in Chapter 3). A pocket with stronger positive charge will interact more tightly with the 5' phosphate and therefore suggests a higher requirement for a template molecule.

Thus, I show here how pol μ helps NHEJ solve a problem integral to its central challenge: restoring chromosome integrity in the absence of an intact template that instructs replacement of lost DNA. I first perform a detailed examination of the substrates that define pol μ 's unique activity and then analyze structural features of this enzyme that help give it its unique activity.

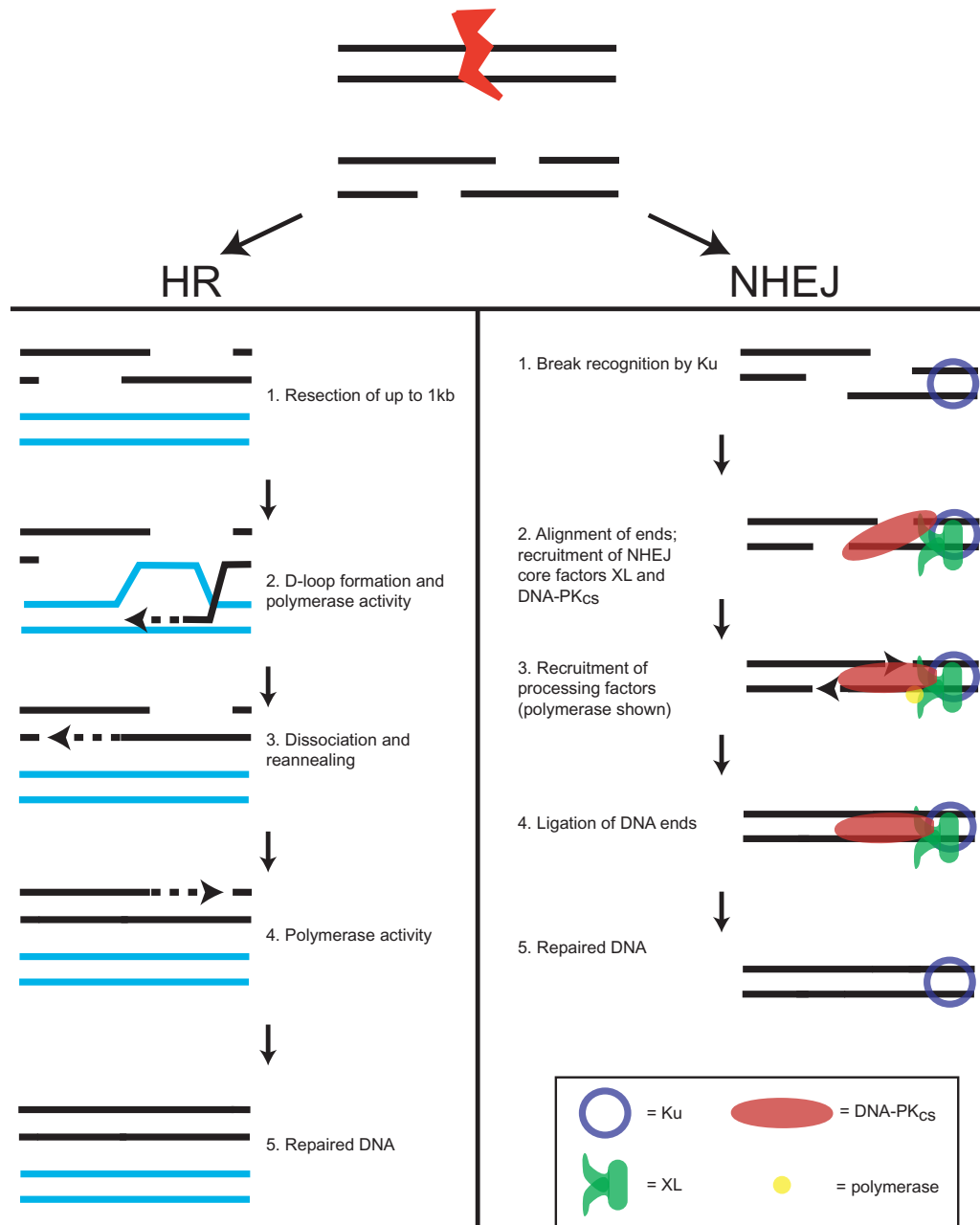


Figure 1.1: Homologous Recombination vs. Nonhomologous End Joining. DSBs can be repaired by either HR or NHEJ. In HR (left side) an intact DNA molecule is used as a template for repair. Resection of DNA followed by polymerase activity (dashed lines with arrowheads) ensures ligation can take place as any damaged ends will be removed and then replaced. NHEJ (right side) does not use an intact template, but it does recruit factors such as nucleases which can remove damaged end structures and polymerases (modeled in yellow) which can fill gaps to make ends ligatable. NHEJ proteins are omitted on the left side of the break for clarity.

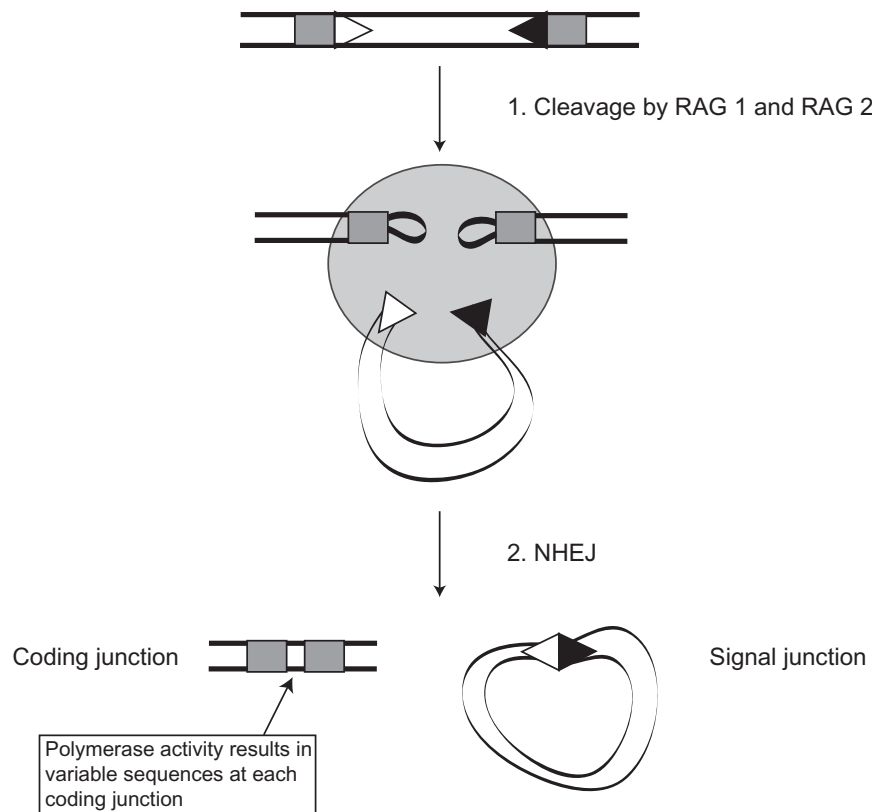


Figure 1.2: V(D)J Recombination. RAG1 and RAG2 proteins cut DNA at an immunoglobulin locus resulting in 2 DSBs. The 4 DNA ends are held together in a post-cleavage complex (grey circle). Blunt-end ligation between ends with recombination signal sequences (black and white triangles) results in a DNA circle that is lost from the genome. Hairpins at the coding ends (grey squares) are opened, processed and ligated by NHEJ. Polymerase activity during NHEJ results in variable sequences at the coding ends, and repeated recombination gives rise to diverse antigen receptors. Based on [53], page 21.

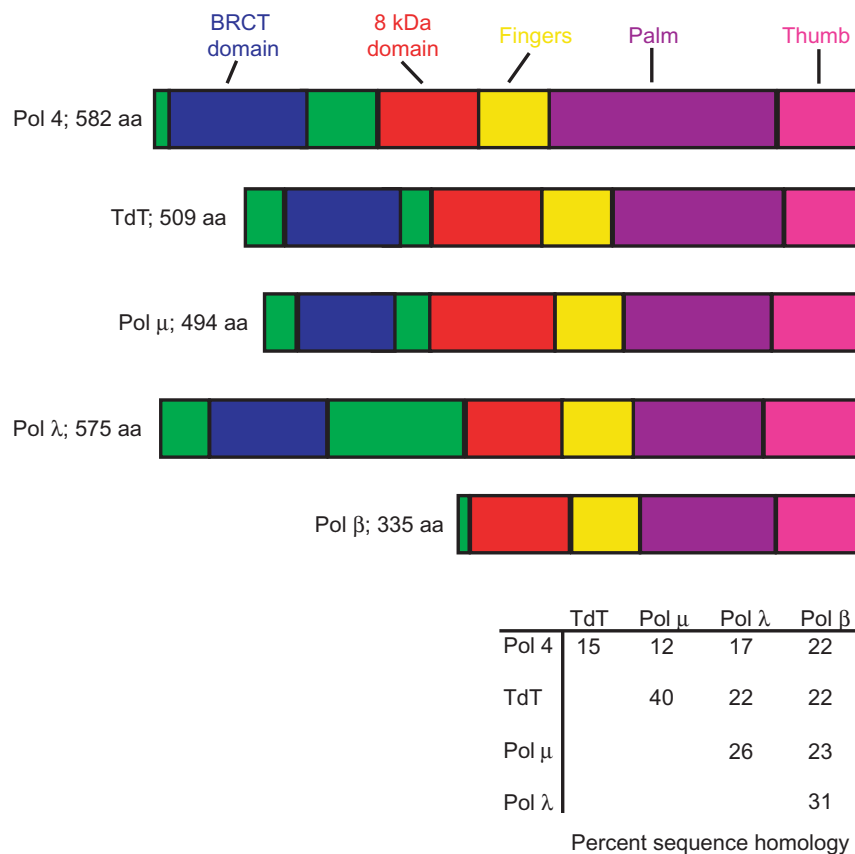


Figure 1.3: Domain Maps of Pol X Family Members. Top: box diagrams of domains in 5 pol X family members. The name of the polymerase and its length in amino acids are to the left of the diagram. Blue regions model BRCT-like domains, red the 8 kDa domains, yellow represents the “fingers,” plum the “palm,” and pink the “thumb.” Bottom: Table of pair-wise sequence homology based on clustalw2 analysis of the amino acid sequences for each polymerase. All polymerases except Pol4 (which is from *S. cerevisiae*) are human.



Figure 1.4: Pol X family polymerases bend DNA 90°. A model of pol μ bound to a gapped DNA duplex. Two bases from the template strand were omitted to simulate a noncomplementary overhang. DNA is shown in stick format with colored atoms and primer and template strands are labeled. The incoming dTTP is cyan and the templating base is magenta. Pol μ is a line drawing in tan, but two residues are in stick representation: H329 (purple) and R175 (red). Pol μ structure from coordinates found at the Protein Data Bank (PDB), accession code 2IHM. Model generated with PyMol.





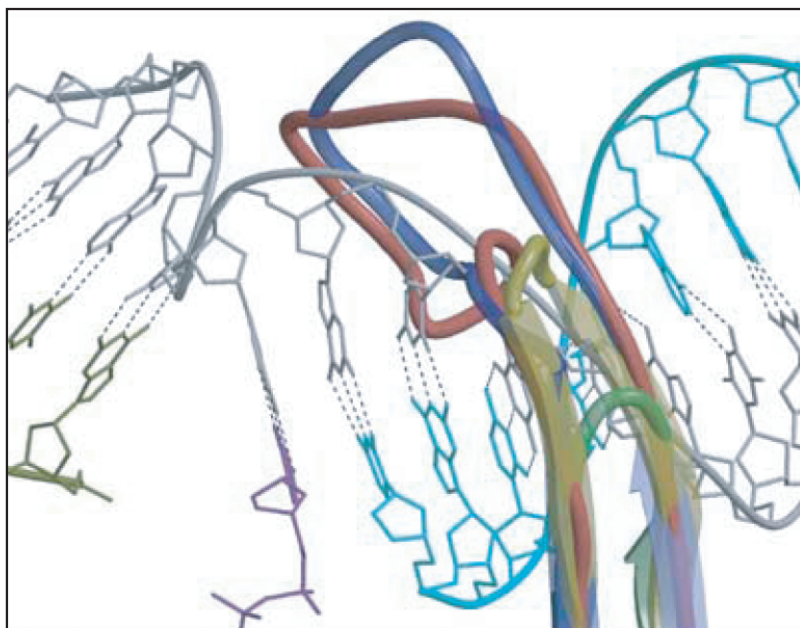
Repair substrates	Polymerases capable of filling the gap	Template is...
	pol β , pol λ , pol μ	Continuous
	pol λ , pol μ	Discontinuous
	pol μ	Discontinuous; not complementary to primer
	<u>TdT</u> , pol μ	Non-existent

Figure 1.5: Gradient of Template Dependence in the Pol X Family. On the left, diagrams of DNA substrates for pol X family members. Incoming nucleotides are red. The polymerases listed in the middle column are those capable of filling the illustrated gap; TdT is underlined because its activity is much higher than pol μ on the primer-extension substrate shown. The right column describes the template strand in the illustration. This dissertation focuses on pol μ 's activity on the discontinuous substrate that is not complementary to the primer (third from top). Figure based on [45].



```

pol β 223-FITDTLSKGE-----TKFMGVCQLP-242
pol λ 456-FLTDDLVSQEENGQ-----QQKYLGVCRLLP-480
pol μ 360-ILYHQHQHSCCESPTRLAQQ-SHMDAFERSFCIFRLP-395
TdT   374-LLYYDLVESTFEKLRLPSRKVDALDHFQKCFLIFKLP-410

```

Figure 1.6: Loop 1 in Pol X Family Polymerases. Top: a ribbon diagram showing Loop 1 for the four mammalian pol X family members, pol β (1BPY) is represented in green, pol λ (1RZT) in yellow, pol μ (modeled on TdT with SWISSMODEL) [54] in red and TdT (1JMS) in blue. These are modeled over a template strand (grey), primer (light blue), downstream strand (olive), and an incoming nucleotide (magenta) according to the pol β structure. The Loop 1 elements follow the backbone of the template strand. Bottom: A sequence alignment of the Loop 1 regions for each polymerase. Colored amino acids are parts of β -sheets 3 and 4 while black residues represent the Loop 1 residues between the β -sheets. Pol μ 's Loop 1 region is predicted based on the alignment with TdT. Figure based on figure 6(c) and (d) from [45]

Chapter 2

Efficient pol μ activity is template directed during NHEJ of noncomplementary ends

2.1 Introduction

DNA double strand break repair, generally speaking, helps the cell avoid translocations and restores genomic integrity. In HR this is done by using an intact copy of the areas affected by the DSB as a DNA template, resulting in accurate repair [2, 3]. NHEJ faces a challenge unique among DNA repair pathways because it does not use an intact copy of the DNA to facilitate repair. As a result, NHEJ must make use of whatever DNA remains after the break. The forces that broke the DNA strands could also have damaged or destroyed individual nucleotides, which are significant barriers to XL's ability to ligate the ends back together [55]. Further, lost nucleotides must be replaced because XL has a limited ability to ligate across a gap [56]. To overcome these barriers, nucleases are necessary to remove damaged bases and then polymerases must replace lost bases. The pol X family fills this role for NHEJ (reviewed in, for example, [26, 5]).

We have proposed a gradient of template dependence for the pol X family [45]. On one end of this gradient, pol β 's activity requires an intact template strand. On the other end, TdT will ignore a template strand if one is even present. Pol μ and pol λ fall in the middle, with pol λ closer to pol β and pol μ closer to TdT. Like pol λ , pol μ can add nucleotides to a primer in the presence of a template and like TdT, it can add nucleotides to a primer in the absence of a template [57]. Pol μ is not as efficient as TdT in this latter activity, however [44]. In NHEJ of partially complementary ends, both pol λ and pol μ are able fill the resulting gap and thereby allow XL to ligate the ends [45]. However, if the ends are not complementary, both pol μ and TdT are able to fill the gap and allow ligation to occur. However, TdT activity in this context remains template independent. Further, outside of V(D)J recombination, only pol μ is able to facilitate such end joining. Our group has shown this activity of pol μ to be critical for the retention of 3'-overhangs during $\nu\kappa$ rearrangement in V(D)J recombination [45]. Other groups and ours have argued it is a unique ability of pol μ to make use of a template (therefore distinct from TdT activity) that is not complementary to a primer (thus distinct from pol λ and pol β activity) to instruct addition of a nucleotide so the gap is filled and ligation can take place [32, 45, 57]. The possibility remains, however, that pol μ synthesis activity in the context of noncomplementary ends is random rather than template instructed [56]. If pol μ facilitates end joining of noncomplementary ends with random nucleotide additions, ligation would probably occur when the added nucleotide is fortuitously complementary to the template strand. Alternatively, under the mechanism we proposed, pol μ 's activity could be template directed: the polymerase preferentially adds a nucleotide complementary to the template, followed by ligation of the DNA ends by XL. In this chapter, we determined whether pol μ is template directed when adding a nucleotide to noncomplementary overhangs and defined some conditions that limit pol μ 's activity in this context.

2.2 Methods

2.2.1 Expression Constructs and Purified Proteins

Human recombinant Ku, XL and polymerases were expressed and purified as previously described [57, 58, 6, 59]. The Arg175>Ala (R175A) mutation was introduced independently in a human pol μ bacterial expression construct as well as a mouse pol μ retroviral construct by the Quickchange (Stratagene, Cedar Creek, TX) protocol, and the open reading frames for these constructs verified by sequencing as correct except for this mutation.

2.2.2 *In Vitro* Assays

Substrates for end-joining and synthesis assays were generated by amplification of a \approx 300-bp mouse genomic fragment including the $J\kappa 1$ coding region. Varied end structures were generated by using primers with different restriction enzyme sites appended 5' of a common sequence complementary to the upstream end of the \approx 300-bp kappa locus fragment (GTGGACGTTTCGGTGGAGGC; referred to as "U" below) and 5' of a sequence complementary to the downstream end of this fragment (GGCTACCCTGCTTCTTTGAGC; referred to as "D" below). We list in order 1) the structure of both fragment ends generated after digestion, 2) the restriction enzyme used to generate this structure, and 3) the sequences appended to the common 3' tails needed to generate the enzyme site: Blunt, Pvu II, CTGCCGCAGCTGTC-U and CGGACCAGCTG-D; 3'T, Ahd I, CTGCCGGACATTCAGTC-U and CGGCAGACCGTCAGTC-D; 3'TT, Bts CI, CTGCCGAACATCC-U and CGGCAGAACATCC-D; 3'TTT, Bgl I, TCATGGCCTAAACGGCT-U and CGGCAGCCTAAAT-D; 3'TTTT, Bst XI, CGGCACCATAAACT-U and TCATGCCATAAACTGGA-D; 3'GT, Bts CI, CTGCCGCACATCCG-U and CGGCAGCACATCC-D; 3' GTT, Bgl I,

TCATGGCCTCAACGGCT-U and CGGCAGCCTCAAT-D; 3' GTTT, Bst XI,
TCATGCCATCAAACCTGGA-U and CGGCACCATCAAACCT; 3' GG, Bts CI,
CTGCCGCCCATCC-U and CGGCAGCCCATCC-D; 3' AA, Bts CI,
CTGCCGTTTCATCC-U and CGGCAGTTCATCC-D; 3' CC, Bts CI,
CTGCCGGGCATCC-U and CGGCAGGGCATCC-D; 3'A, Ahd I,
CTGCCGGACATACAGTC-U and CGGCAGACCGACAGTC-D; 3'C, Ahd I,
CTGCCGGACATCCAGTC-U and CGGCAGACCGCCAGTC-D; 3' G, Ahd I,
CTGCCGGACATGCAGTC-U and CGGCAGACCGGCAGTC-D.

PCR products of these primer pairs were introduced into the TOPO-TA 2.1 vector (Invitrogen, Carlsbad, CA), and sequenced to verify the accuracy of the insert. Plasmid DNA for each insert was then amplified with the 5' biotinylated vector specific primers DAR470 (AGTGTGCTGGAATTCGCCCTT) and DAR471 (GTGATG-GATATCTGCAGAATTCGCCCT) in the presence of [α - 32 P]dCTP. After digestion with the appropriate enzyme, the fragment was purified by depletion of uncut or partially extended products with magnetic streptavidin beads (Roche biochemicals, Basel, Switzerland), as well as a Qiaquick PCR purification step (Qiagen; Valencia, CA) with a 35% guanadine isothiocyanate wash included.

End-joining assays were performed by preincubating 25 nM Ku, 50 nM XL, and 25 nM polymerase with 5 nM DNA substrate for ten minutes in our standard reaction buffer (25 nM Tris [pH 7.5], 1mM DTT, 150 mM KCl, 4% glycerol, 40 μ g/mL bovine serum albumin (BSA), 4% glycerol, 0.1 mM EDTA, and 10% (wt/vol) polyethylene glycol (MW > 8000 kDa) (PEG). Ligation was initiated by addition of dNTPs or individual ddNTPs to 25 μ M (or 25 μ M each), MgCl₂ to 5 mM and 200 ng supercoiled plasmid DNA (Litmus38; New England Biolabs, Ipswich, MA). Reactions were incubated at 37 °C for the noted times, stopped, and deproteinized. Reactions performed with dNTPs were analyzed by electrophoresis on a 5% native PAGE. Reactions performed with ddNTPs

were first digested with *Hinf*I prior to analysis on a denaturing 8% polyacrylamide gel (PAGE). All analysis of ddNTP reactions focuses on the substrate end fragment that is sufficiently small (55 nucleotides) to resolve single nucleotide additions to the substrate band. Head-to-tail junctions of *in vitro* reactions were sequenced for Table 2.1 by cloning products of 25 cycles of amplification of $1/1 \times 10^3$ of a reaction with the primers 5'GCTGGAATAGGCTAGACATG and 5'GCCACAGACATAGACAACGG.

2.3 Results

2.3.1 Pol μ typically adds complementary nucleotides more efficiently than non-complementary nucleotides

We determined how well two pol X family polymerases are able to help NHEJ join non-complementary ends by using an *in vitro* assay. A ^{32}P -labeled linear DNA fragment ($\approx 300\text{bp}$) with 3'TT overhangs (Figure 2.1 A) was incubated with Ku, XL, polymerases and dNTPs. This fragment's overhangs are not complementary, thus levels of joining in the presence of the NHEJ core factors Ku and XL alone are negligible (Figure 2.1 B, lane 4). As previously described [57], addition of pol μ to Ku and XL allowed for efficient generation of concatamer ligation products (Figure 2.1 B, lane 5). Pol λ , another pol X member that also interacts with Ku and XL, was unable to substitute for pol μ even when in 10-fold excess of pol μ (Figure 2.1 B, lanes 6 and 7).

XLF (XRCCIV-like factor) and DNA-PKcs have also been implicated as NHEJ core factors in cells. Under the conditions used here, further addition of purified XLF and DNA-PKcs to our *in vitro* reactions does not stimulate joining. However, we note that addition of XLF (but not DNA-PKcs) allows for more efficient joining of noncomplementary ends independent of the addition of pol μ , consistent with recent reports [60, 61]. Nevertheless, the pol μ -independent, XLF-dependent joining of this substrate was still

much less (≈ 6 fold) than that observed in the presence of pol μ (unpublished data). Moreover, neither XLF nor DNA-PKcs allowed pol λ to perform synthesis dependent joining. Ku and pol μ are thus both necessary and sufficient for efficient ligation of ends by XL under these conditions.

We next addressed if pol μ 's synthesis activity is dependent on both XL and Ku. To separate synthesis from ligation we repeated experiments as described above, but substituted ddNTPs for standard dNTPs. This assessed synthesis directly, rather than assessing only synthesis that gave rise to an end structure that permitted ligation. Interestingly, synthesis by pol μ was dependent on the presence of both Ku and XL (Figure 2.1 C, compare lanes 2 and 3 to lane 5).

How does pol μ promote NHEJ in this context? We previously proposed that pol μ was able to synthesize from a primer on one DNA end, but use as template the overhang sequence of a second DNA end [57]. However, only a limited number of substrate contexts have been explored, leaving open several alternate explanations.

We therefore devised the following assay to definitively resolve this issue (Figure 2.2 A) [62]. As shown previously (Figure 2.1 C; compare lanes 3 and 5), XL is essential for significant pol μ synthesis activity under our conditions, thus we continued to use ddNTP substrates to permit analysis of synthesis products in the presence of XL independently of whether synthesis can contribute to a joined product. Additionally, synthesis could conceivably only appear template-dependent under prior assay conditions if primer sequence affects which nucleotides can be added by pol μ . The ability to vary template end sequence independent of the primer end was achieved by mixing a small amount of radiolabeled (hot) fragment with an excess of unlabeled (cold) DNA fragment. Ends of the hot fragment will thus most frequently interact with ends of the cold fragment, allowing us to define the hot ends as the primer, and the overhangs of the cold ends as the primary template. The relative efficiency of synthesis for each of the 4 ddNTPs from

different priming end sequences can then be evaluated while independently varying the sequence of the template end.

We assessed activity using the previously described 3'TT overhang substrate as a primer, but varied the identity of potential template ends. For all three potentially noncomplementary template ends (G, T, and C), high activity (30-50% of substrate is converted to $n+1$ product) is observed only when the ddNTP that is complementary to the template end sequence is supplied (Figure 2.2 B, top panel, first three rows). In contrast, levels of activity with noncomplementary ddNTPs are typically 10-fold lower. Results were similar when standard dNTPs were used in one experiment (template T; our unpublished data), indicating inefficient use of noncomplementary ddNTPs is not a function of the missing 3' hydroxyl. Interestingly, when using dATP (complementary to template), the unligated products of extension from an end were detectable, but were less abundant (approximately one-third) than the levels of the synthesis-dependent ligations products involving the same end. Ligation is thus well-coupled to synthesis. We next addressed what the consequences would be of the absence of template under the same conditions. We substituted a blunt ended fragment (template "0") for the standard 3'-overhang containing 'template-end' fragments and failed to detect activity with any of the ddNTPs (Figure 2.2 B, top panel, bottom row), even after extended incubation (20 min). Pol μ activity using the TT overhanging primer is thus primarily both template-directed and template-dependent: template-directed because it is most efficient using the ddNTP complementary to the sequence of a template end overhang, and template dependent because activity is much lower when pol μ -containing NHEJ complexes interact with ends that lack a potential template (blunt ends).

We then repeated this analysis for the remaining nine different mismatching primer and template-end combinations. For 4 of these combinations (7/12 total), pol μ remains both efficient and accurate: over 20% of the primer is extended within 2.5 minutes when

using a ddNTP complementary to template sequence (reactions noted by filled triangles in Figure 2.2 B), while activity using noncomplementary ddNTPs was typically not detectable in the same time frame. We did not detect significant addition for any of the 4 ddNTPs after 2.5 min using GG, AA and CC primers when a blunt-ended fragment was substituted for template (bottom line in each section of Figure 2.2 B). Thus, as with the TT primer, we conclude pol μ is most active in the presence of a potential template (template dependent).

However, pol μ was not equally active for the remaining five combinations. For these combinations, activity was barely detectable over the same time frame regardless of the identity of the ddNTP (unpublished observations), and much longer reactions (10-20min) were needed to observe significant levels of activity. In these extended incubations, addition of at least one of the presumed noncomplementary ddNTPs (diamonds in Figure 2.2 B) can also approach that seen with the complementary ddNTP (open triangles in Figure 2.2 B). Significant activity can also be seen in the absence of added template (template '0'). It is important to note that when nucleotides added are not complementary to the intended template (or are observed using template '0'), they are often complementary to the primer (open diamonds in Figure 2.2 B), consistent with the use of a second primer molecule as template. This is probably the appropriate explanation for examples where ddC is added, since activity using this nucleotide is significant only in the presence of the GG primer or when G is the intended template. However, there remain several examples (most often additions of ddT) where the nucleotide added is not complementary to sequence in either the intended template or the primer overhang (filled diamonds in Figure 2.2 B). These represent definitive template-independent activity.

Thus, while high levels of pol μ activity are template-dependent and template-instructed, specific combinations of primer/template mismatches lead to much lower

template-instructed activity. Possible explanations for pol μ 's reduced activity in these circumstances will be addressed in the discussion section.

2.3.2 End-bridging interactions help Pol μ fill short gaps between two ends

Pol μ activity in NHEJ is thus affected by the sequence of the overhangs. We next determined if there were limitations on the length of the overhang. Comparison of joining activity on substrates with 1, 2, 3, and 4-nt 3'-overhangs indicates pol μ activity declines sharply as overhang length increases (Figure 2.3 A, lanes 3-6) (Table 2.1). For example, a single nucleotide overhang substrate was efficiently joined within 2.5 minutes, while a 4 nucleotide overhang substrate required 20 minutes for comparable levels of joining (Figure 2.3 A; compare lanes 3, 6, and 8). Additionally, sequences of the junctions from these reactions indicate that on the longer overhangs (3-4 nucleotides), one of the overhangs is entirely present, but 1-2 nucleotides of the second overhang were typically lost (Table 2.1). This is best explained if pol μ is unable to efficiently fill in gaps longer than 2 nucleotides.

We further explored this question by assaying synthesis activity while independently varying the length of the primer and template overhangs (Figure 2.3 B and C), using the experimental setup described above (Figure 2.1 C and Figure 2.2 B). Using a template 3'-overhang of fixed length (1 nucleotide), we determined that while pol μ is relatively inactive using a blunt end primer, 3'-overhang primers from 1-4 nucleotides in length can all be used with only a slight reduction in pol μ activity with increasing overhang length (Figure 2.3 B). We next addressed the impact of different length 3'-overhang templates, using a 2 nt 3'-overhanging primer. At the same time, we addressed two possible explanations why an increased length of template overhang might generate junctions with deletions (Table 2.1 and Figure 2.3 C). Pol μ might initiate synthesis correctly using an

overhang terminus as template, but because the overhangs were homopolymeric in these experiments, the primer terminus might have slipped and reannealed with an internal site in the template overhang. Alternatively, pol μ might be unable to use the terminal nucleotide effectively as a template on the longer (3-4 nt) overhangs: in this case, pol μ presumably initiates synthesis using an internal nucleotide as template instead. To distinguish these two possibilities, we used a series of substrates where the identity of the terminal template nucleotide (G) was different from internal template nucleotide(s) (T). As single nucleotide overhangs, both template nucleotides can be used efficiently with the 3' TT primer (Figure 2.3 C, top two rows; reproduced from Figure 2.2 B for comparison purposes). In contrast, on longer overhangs pol μ is progressively less able to add the terminal nucleotide (ddC), while activity with ddA is less affected (Figure 2.3 C, lower 3 rows). The loss of overhang nucleotides in the junctions described in Table 2.1 can thus be best explained if pol μ cannot efficiently use the terminal nucleotide as a template when the template overhang is over 2 nt long. Instead, pol μ uses internally located template nucleotides, presumably because these nucleotides are closer to where the template end is double stranded.

2.4 Discussion

We have shown here when pol μ efficiently adds nucleotides to a primer strand during NHEJ of noncomplementary ends, it does so in a template directed way. There remain, however, several cases where pol μ does not add nucleotides efficiently. Pol μ is inefficient when the template strand is blunt ended, when the template overhang is longer than two nucleotides, and when certain combinations of nucleotides are the terminal bases in the overhangs. The first two cases likely reflect an inability of pol μ to “see” the template. Blunt template ends do not present a template strand, and as we showed here, pol μ uses the interior nucleotides of a long template overhang as template; perhaps the length

of the overhangs makes it difficult for pol μ to interact with the double stranded portion of the template molecule, reducing activity.

This second exception to pol μ 's template instructed activity argues pol μ functions in NHEJ as part of a bridged-end complex. Put another way, it suggests pol μ requires interactions with both sides of a DNA gap to efficiently add a nucleotide when it functions in NHEJ. Therefore we propose pol μ can act in a mode intermediate to those of pol λ and TdT. Pol λ requires a template that interacts with a primer. TdT displaces a template if one is present, thus it requires a primer not interact with a template [52]. Pol μ uses a template (like pol λ), but the primer does not have to interact with it (like TdT). This hypothesis is explored more fully in chapter 3.

The third exception to pol μ 's template-instructed activity, where some primer/template combinations do not result in efficient activity, is more challenging to explain. G/T, A/C, C/A, A/G and C/T all result in inefficient and inaccurate pol μ activity. Three of the five pairings (G/T, A/C, and C/A) are pyrimidine/purine mismatches, which recent work has indicated can be aligned and joined directly by Ku and XL without synthesis. Pol μ 's reduced activity in those cases might then simply indicate Ku and XL had aligned the two DNA ends such that there was not a gap but rather a mismatch, leaving no substrate for pol μ [56]. The other inactive primer/template pairs (A/G and C/T) have not been shown to be ligated directly by Ku and XL without synthesis, but we note that pol μ is efficient and accurate on the reverse of those combinations (i.e. G/A and T/C). Therefore in a biological setting, pol μ could accurately work on the other strand, promoting ligation of that strand by XL and leave the "first" single strand gap to be resolved by a process other than NHEJ.

No matter the reason for pol μ 's inefficient activity, we point out that in those cases, the nucleotide it adds most of the time is ddT. As was demonstrated here, pol μ is most efficient and template directed when the terminal nucleotide in the primer is a T. This

may represent a compromise wherein the mutagenic potential of nontemplated additions is minimized (only one nontemplated base is added as opposed to potentially several), allowing repair to take place in a reasonable time so translocations can be avoided.

DNA damage that results in DSBs can also result in damaged bases, abasic sites, or even the loss of entire nucleotides. Damaged or lost bases have to be repaired by other processing factors, but DNA polymerases λ and μ of the pol X family are subsequently able to fill the gaps that result as long as bases on both sides of the gap are complementary. Pol μ , however, is uniquely capable of adding a templated base to a primer that does not interact with the template strand, as we showed here. We suggest this activity contributes to the ability of NHEJ to repair a wide range of DSBs, even in the absence of an intact template strand. The means by which pol μ is able to make this contribution to NEHJ lie in structural features of the polymerase, some of which are detailed in the next chapter.

2.5 Acknowledgements

This work was first published in May of 2008 in *Nucleic Acids Research* 36(9):3085-94 with the title *End-bridging is required for pol μ to efficiently promote repair of noncomplementary ends by nonhomologous end joining* authored by myself, Jody M. Havener and Dale A. Ramsden [63]. I have excluded some data from that article from this chapter and included it in the next. I am responsible for the *in vitro* ligations and the synthesis assays in Figure 2.1, Figure 2.2, Figure 2.3 B and two of the panels in Figure 2.3 c (TT/G and TT/TTG). D.A. Ramsden generated the remainder of the data presented in this chapter.

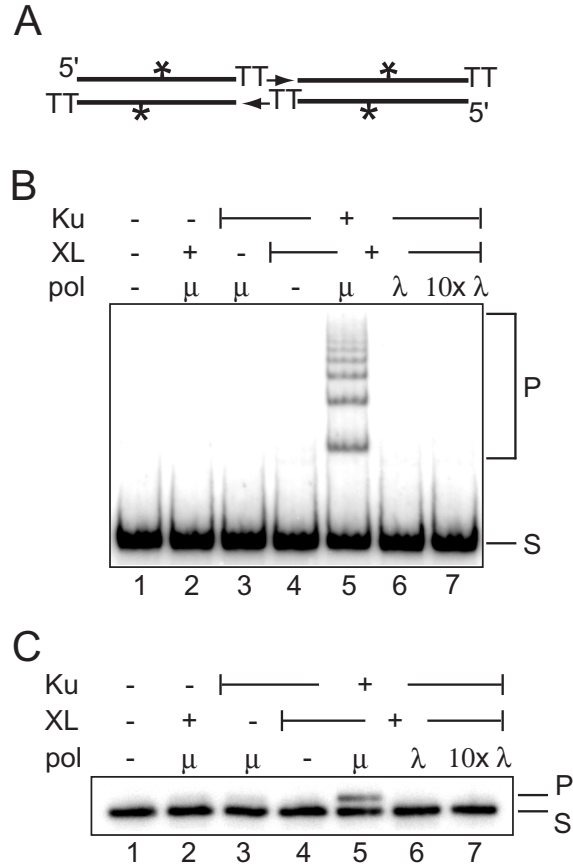


Figure 2.1: Factors required for joining of noncomplementary ends. A) A diagram of the standard 280 bp substrate labeled internally with ^{32}P (*), and possessing 3' TT overhangs. Arrows indicate direction of synthesis by pol μ after alignment of ends by core NHEJ factors Ku and XL. B) All reactions used 5 nM DNA substrate as illustrated in (A), and products analyzed after 5 minute reactions. 25 nM Ku and 50 nM XL were added as indicated (+). Polymerase μ or λ was added at 25 nM or 250 nM (10X). S - substrate; P - concatamer ligation products. C) Reactions performed as in B except ddNTPs substituted for dNTPs, and synthesis at one end analyzed by denaturing PAGE as described in methods. S - substrate; P - +1 synthesis product.

Appeared originally in [63]

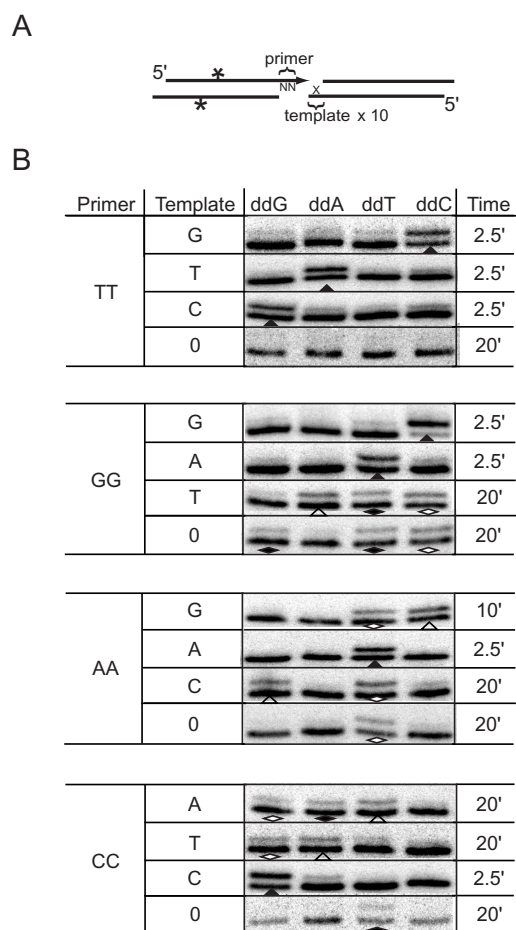


Figure 2.2: Effect of overhang sequence on pol μ activity in NHEJ. (A) Diagram of the 3'-overhang substrates, emphasizing how primer ends are distinguished from template ends by radiolabeling of the former, and inclusion of a 10-fold excess of the latter. (B) Column 1 lists the different primer overhang sequences and Column 2 the different template overhang sequences. The '0' in column 2 identifies reactions where a blunt-ended fragment was substituted for the 1 nt 3'-overhanging template ends used elsewhere. Synthesis with each mixture of primer and template end was assayed in the presence of pol μ , Ku and XL as in Figure 2.1C, except each of the four ddNTPs were included individually at 25 mM. Reaction times were varied as noted in the final column to allow for significant accumulation of product. Filled triangles mark combinations of primer, template and ddNTP that generated significant product within 2.5 min. For combinations of primer, template and dNTP that required 10-20 min for significant product, open triangles identify additions complementary to the intended template, open diamonds identify additions complementary to the primer, while filled diamonds identify additions that are definitively template independent (not-complementary to either primer or intended template).

Appeared originally in [63]

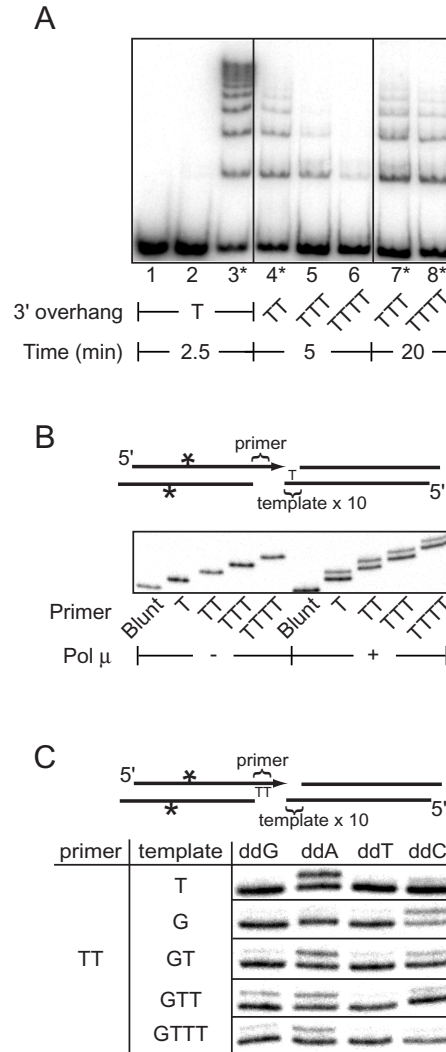


Figure 2.3: Effect of overhang length on pol μ activity in NHEJ. A) Substrates with the noted 3'-overhang sequences were incubated with dNTPs, Ku, XL, and pol μ as in Fig 2.1B, except control lanes 1 (no protein) and 2 (no pol μ), for the indicated time periods. Asterisks identify reactions where the products were amplified, sequenced, and reported in Table 2.1. B) A diagram of the substrates used in this panel. Primer molecules of varying length, as noted, were incubated as in Fig 2.2B, with Ku, XL, pol μ , 3' T overhang template and 25 μ M each ddNTP. C) A diagram of the substrates used in this panel. Assays were carried out as in Fig 2.2 B with a 3'-overhang TT primer, varying the template overhang sequence and ddNTP added as noted. Images of reactions with T and G templates are reproduced from Figure 2.2 B to aid in comparison.

Appeared originally in [63]

Substrate ¹	% Activity ²	Synthesis in Junctions ³	
T	100	A(17)	Accurate fill-in
TT	16	AA(13) C (1)	Accurate fill-in Mis-insertion
TTT	5	AA (8) AAA(4) GTAAA (1)	1 nt deletion Accurate fill-in N-addition+fill-in
TTTT	1	AA (9) AAA(4)	2 nt deletion 1 nt deletion

Table 2.1: Sequences of junctions from reactions noted with asterisks in Figure 2.3 A. ¹ 300 bp substrates with the noted 3'-overhangs. ² Joining activity, relative to the most active substrate (single T overhang; see Figure 2.3 A). ³ Nucleotides synthesized by pol μ in junctions. The number of examples in of each junction are in parentheses.

Appeared originally in [63]

Chapter 3

Structural elements in pol μ important for its unique activity

3.1 Introduction

NHEJ is one of two major DNA double strand break repair processes employed by mammals [2, 21]. It uses polymerases from the pol X family to fill gaps that result from the loss of nucleotides, thereby overcoming part of its major challenge, the lack of an intact template to instruct repair. Pol X family polymerases such as pol λ or pol μ can be recruited to fill the gaps that result from partially complementary ends [44, 57]. Pol μ , however, also has the unique ability to fill in gaps that result from noncomplementary DNA ends in a template-directed way, as we showed in chapter 2. Having defined this aspect of pol μ 's activity, we noted it shares activities associated with pol λ as well as TdT. Here we identified structural features that confer such activity upon pol μ .

The recent solution of the crystal structure of pol μ bound to a gapped DNA substrate provided new insights into the interactions between pol μ and its substrates [32]. In particular, the crystal structure allowed our group to form and test hypotheses about interactions important for pol μ 's unique activity. Experiments detailed in chapter 2 indicated pol μ must interact with the duplex region of DNA downstream of the gap in

order to efficiently add nucleotides. Based on our observation that pol μ 's activity shares qualities with the activities of pol λ and TdT, we sought protein-DNA interactions that were similar to those in pol λ and TdT. We identified two such residues.

Pol μ 's H329 sits between two phosphates: one phosphate is located in the backbone of the primer strand, while the other phosphate is in the incoming nucleotide (see Figure 1.4, purple residue). The Pederson group speculated it could help bridge the primer terminus and incoming nucleotide, and consequently could help support nucleotide addition when the incoming nucleotide is not stabilized by base-pairing with template [32]. Consistent with this hypothesis, a histidine is found at this position in both polymerases active in this context (pol μ and TdT).

R175 was highlighted by the structure as interacting with the 5'phosphate of the downstream strand, a potentially critical point for the polymerase to track because it is the end of the double-stranded portion of the template molecule (see Figure 1.4, red residue, Figure 3.2 A). We showed in the previous chapter that pol μ likely requires interactions with that DNA duplex for efficient activity on noncomplementary DNA ends.

We tested an alanine mutant of each of these residues and found H329 is critical for pol μ 's ability to add nucleotides to a primer in the absence of a template and that it is also required for NHEJ of noncomplementary overhangs. As predicted by the structure, R175 is an important residue for gap filling in the presence and absence of base pairs between the primer and template strands. This further indicates pol μ functions in a bridged-end complex, as we suggested in the previous chapter. Together these results indicate pol μ has structural elements similar to both TdT and pol λ , and the combination of those elements in a single polymerase results in an activity not seen in either of the other two.

3.2 Materials and Methods

3.2.1 Expression Constructs and Purified Proteins

Human recombinant Ku, XL and polymerases were expressed and purified as previously described [45, 64, 65, 30]. The c. 522_524delCGTinsGCT, p. Arg175>Ala (R175A) mutation was introduced independently in a human pol μ bacterial expression construct as well as a mouse pol μ retroviral construct by the Quickchange (Stratagene, Cedar Creek, TX) protocol, and the open reading frames for these constructs verified by sequencing as correct except for this mutation.

The H329A mutant of human pol μ was generated using QuikChange mutagenesis (Stratagene), and was expressed and purified as described [65]. Polymerase activity was assayed by primer extension [32]. No exonuclease activity was detected.

3.2.2 End Joining Assays

Substrates were generated as described [32] for reactions involving pol μ H329A. For reactions with pol μ R175A, substrates were generated by amplification of a \approx 300-bp mouse genomic fragment including the Jk1 coding region. The end structures were generated with the restriction enzyme site for BtsCI appended 5' of sequence complementary to the upstream end of the \approx 300-bp kappa locus fragment and 5' of a sequence complementary to the downstream end of this fragment. Primers CTGCCGAACATCCGTGGACGTTTCGGTGGAGGC and CGGCAGAACATCCGGCTACCCTGCTTCTTTGAGC resulted. PCR products of these primer pairs were introduced into the TOPO-TA 2.1 vector (Invitrogen, Carlsbad, CA), and sequenced to verify the accuracy of the insert. Plasmid DNA for each insert was then amplified with the 5' biotinylated vector specific primers DAR470 (AGTGTGCTGGAATTCGCCCTT) and DAR471 (GTGATGGATATCTGCAGAATTCGCCCT) in the presence of

[α - 32 P]dCTP. After digestion with BtsCI, the fragment was purified by depletion of uncut or partially extended products with magnetic streptavidin beads (Roche biochemicals, Basel, Switzerland), as well as a Qiaquick PCR purification step (Qiagen; Valencia, CA) with a 35% guanadine isothiocyanate wash included.

End-joining assays were performed by preincubating 25 nM Ku, 50 nM XL, and 25 nM or 1 nM polymerase with 5 nM DNA substrate for ten minutes in our standard reaction buffer (25 nM Tris [pH 7.5], 1mM DTT, 150 mM KCl, 4% glycerol, 40 ug/mL bovine serum albumin (BSA), 4% glycerol, 0.1 mM EDTA, and 10% or 12.5% (wt/vol) polyethylene glycol (MW > 8000kDa) (PEG)). Ligation was initiated by addition of dNTPs or individual ddNTPs to 25 uM (or 25 uM each), MgCl₂ to 5 mM, 200 ng supercoiled plasmid DNA (Litmus38; New England Biolabs, Ipswich, MA) and 0.1 mM ATP in reactions with pol μ H329A. Reactions were incubated at 37°C for the noted times, stopped, and deproteinized. Reactions performed with dNTPs were analyzed by electrophoresis on a 5% native PAGE. End joining efficiencies were calculated as follows. The amount of each product species, determined by phosphorimaging, was multiplied by the number of ligation events per product species. The sum of these corrected product amounts was then divided by the total unreacted substrate plus product in the reaction. Reactions performed with ddNTPs were first digested with HinfI prior to analysis on a denaturing 8% polyacrylamide gel (PAGE). All analysis of ddNTP reactions focuses on the substrate end fragment that is sufficiently small (55 nucleotides) to resolve single nucleotide additions to the substrate band.

3.2.3 Polymerase activity assays

Oligonucleotide substrates for pol μ assays with R175A (Figure 3.2 and Table 3.2) were generated by annealing the primer Cy3-5'GCTTGAAGACTGGTGAAGACTTGAG (SNM34) to the template 5'CCATGAATCGACCTGTACCTCAAGTCTTCACCAGT-

CTTCA (SNM36), generating PE-1. A single nucleotide gapped substrate (“Gap”) was made by adding to annealing reactions the 5′ phosphorylated downstream strand 5′TACAGGTCGATTCATGGAGT (SNM35). A substrate to measure extension from a 3′ overhang (TdT-like primer extension; PE-2) was made by annealing Cy3-5′GTAGGG-CTCATGTTAGATCTATCGAGCAAGTGCATCTGCAGTACTCATATGGAATTCCC-AGCTGAG (DAR167) to 5′CAGCTGGGAATTCCATATGAGTACTGCAGATGCACT-TGCTCGATAGATCTAACATGAGCC (DAR166).

Gap filling and primer extension reactions comparing wt pol μ and pol μ R175A were incubated at 37°C with 25 mM Tris pH 7.5, 0.1 mM EDTA, 1.1 mM DTT, 25 mM KCl, 125 mM NaCl, 1% glycerol, 10 ug/mL BSA, and 5 mM MgCl₂ in the presence of 0.1 mM each dNTP and 5 nM DNA substrate. The reactions were then stopped by addition of formamide loading dye and analyzed by denaturing 10% PAGE. Extension from a 3′ overhang was analyzed by incubation at 37°C in a similar reaction buffer, except using 31.25 mM KCl, 118.75 mM NaCl, and 10% PEG. Reactions were stopped by addition of formamide loading dye and analyzed by denaturing 8% PAGE.

3.2.4 Analysis of Igk recombination in pol μ R175A overexpressing cells

Pol μ R175A over-expressing cells were made by infection of SP-9 with a pBABE-puro retroviral construct containing the pol μ R175A mutant cDNA as previously described for SP-9 cells overexpressing wild-type (wt) pol μ and pol μ Δ loop [57]. Western analysis was performed on whole cell extracts using an anti-myc antibody (#2272; Cell Signaling, Beverly, MA) and an anti-actin antibody (A2066; Sigma, St. Louis, MO) followed by visualization with an anti-rabbit-horseradish peroxidase secondary antibody and ECL+ chemiluminescence (GE biosciences, Piscataway, NJ). VJ κ 1 recombination junctions were generated and analyzed as previously described [57].

3.3 Results

3.3.1 H329 is required for pol μ activity when the primer does not interact with the template

The Kunkel/Pedersen group determined intrinsic polymerase activities of pol μ and the H329A mutant with steady-state kinetic assays. They first evaluated their abilities to fill gaps in DNA and found the mutation only mildly reduces that activity relative to wt pol μ (Table 3.1) [32]. Notably, wild type but not the H329A mutant could extend a primer in the absence of a template. These results are consistent with the H329 residue being important for stabilizing the incoming nucleotide relative to the primer strand, and also with that interaction being particularly important in the absence of a template strand. It is possible the template strand, when present and properly positioned by other interactions with the polymerase, provides a stable point with which pol μ , the primer strand, and the incoming dNTP can interact before catalysis, relegating the interactions involving H329 to a secondary role. It would then be that the interactions with H329 are only vital when the template strand is missing.

If that were the case, loss of H329 would have no effect on pol μ activity during NHEJ. To test that hypothesis, we used wt and mutant pol μ in NHEJ of partially and non-complementary overhangs. $\alpha^{32}\text{P}$ -labeled DNA substrate with partially complementary 3' overhangs was incubated with Ku, XL, dNTPs, and wt pol μ or the H329A mutant for the indicated times. In the absence of protein, Ku, or polymerase, these DNA ends cannot be ligated efficiently, so the products from these reactions are negligible (Figure 3.1 a, lanes 1-3). Addition of wt pol μ to the other reaction components results in the formation of DNA concatamers while addition of the mutant protein was 2-fold less able to stimulate NHEJ (Figure 3.1 a, compare lanes 4 and 5). However, when we similarly assayed NHEJ activity of the two polymerases on non-complementary DNA ends, the

mutant was about 30-fold less effective than the wild type (Figure 3.1 b, compare lanes 4 and 5). To determine the reason for this greatly decreased efficiency, the synthesis activities of pol μ and the mutant were examined for both kinds of DNA ends. In the presence of Ku, XL, DNA ends and a ddNTP substrate, pol μ has been shown to add a single ddNTP to the end of a labeled primer molecule, but XL is unable to ligate the DNA ends that terminate in a dideoxynucleotide. Thus, by examining the end of the primer for the addition of a nucleotide, the synthesis activity of the polymerase can be determined. Further, when each of the four ddNTPs is provided singly to four reactions, it can be determined whether pol μ is discriminating in its nucleotide addition. When provided with minimally complementary overhangs (3'TC and 3'GA) wt pol μ displays two different types of activity: it inserts both ddTTP and ddCTP (Figure 3.1 d, lanes 3 and 4). The incorporation of ddTTP is expected when the substrate ends are synapsed with base pairing interactions between the terminal C and G of the two DNA ends. Incorporation of ddCTP, however, is expected when the ends are synapsed without base pairs, leaving G as the first templating nucleotide. Importantly, pol μ H329A only displays evidence of one type of activity: it only inserts ddTTP, the nucleotide expected when the DNA ends form base pairs. The mutant polymerase incorporates no ddCTP, suggesting it cannot act as wt pol μ does on DNA ends without any base pairing interactions. To confirm this, we compared the synthesis of these polymerases on 3'TC overhangs, which cannot be aligned with base pairing. Pol μ primarily incorporated ddGTP as well as ddATP and ddCTP to a lesser extent (Figure 3.1 c, top panel). Pol μ H329A, however, was completely inactive when presented with such DNA ends, indicating it is unable to catalyze the addition of a nucleotide to a primer strand in the absence of base pairs between primer and template strands (Figure 3.1 c, bottom panel) in addition to being unable to add a nucleotide in the absence of a template strand. Therefore pol μ 's activity in NHEJ of noncomplementary overhangs relies on

H329. Because the template strand does not form base pairs with the primer strand, it may be that interactions between the template strand and the incoming nucleotide are too weak to use for proper positioning of the incoming nucleotide (while remaining strong enough to allow the polymerase to ensure the complementary nucleotide is being added). In that case, H329 might be required in this context to provide stabilizing interactions between the incoming nucleotide, the polymerase, and the primer terminus.

H329 thus appears to be a structural element similar to one in TdT that serves two purposes in pol μ : it allows pol μ 's weak transferase activity and facilitates a new activity, that is, gap filling of noncomplementary ends.

3.3.2 R175 is important for end-bridging interactions in pol μ

Having identified a TdT-like structural element that pol μ uses in its unique activity, we next sought an element similar to one in pol λ and pol β that pol μ similarly uses. Pol β and pol λ both require a template that interacts with the primer. Since we have shown pol μ uses a template strand to direct synthesis even in gap filling across noncomplementary overhangs, we targeted a region common to all three polymerases that interacts with the downstream strand. In NHEJ, it is through contacts with the downstream strand that the polymerases connect with the template strand, therefore this region is important for template dependence. Analysis of the 8 kDa domains of each of those polymerases indicated a residue the pol μ crystal structure suggests interacts with the downstream strand, in essence, the “other side” of the gap.

R175 is analogous to a positively charged residue in pol λ (R75) and in pol β (K35). It was predicted those residues interact with the 5' phosphate on the downstream DNA strand, which is the strand annealed to the template strand. Notably, template-independent TdT lacks a positively charged residue at this position. Mutation of R175 to alanine results in a specific defect in gap-filling. R175A is 4-fold less

active than wild type pol μ in primer extension in the presence of a 5' phosphorylated downstream strand (Figure 3.2 B, gapped substrate, compare lanes 1 and 3), but possesses activity more similar to wild type pol μ in primer extension when the downstream strand is missing (Figure 3.2 B, compare lanes 5 and 6; Table 3.2), or even when the downstream strand's 5' phosphate is missing (Table 3.2).

Much greater concentrations of enzyme and longer incubations are required to see significant levels of TdT-like extension under these conditions (Figure 3.2 B; Table 3.2). Importantly, R175A is as active as wild-type pol μ in this assay (Figure 3.2 B, compare lanes 7 and 8; Table 3.2). The R175A mutant thus has significantly reduced activity only on gapped substrates. However, we note that while R175A's defect in gap filling is significant in comparison to the wild-type protein, R175A is still much more active (10-fold) on gapped substrates than on those where the downstream strand is absent (Figure 3.2 B, compare lanes 3 and 4; Table 3.2). Pol μ is thus clearly able to interact in a functionally significant way with the downstream strand even when R175 is mutated (e.g. through the helix-hairpin-helix motif), making this mutation a partial loss of function only.

We next tested R175A's ability to promote NHEJ of noncomplementary ends. Strikingly, R175A is 4-fold less active than wild type pol μ in this context (Figure 3.2 C, compare lanes 3 and 4). Since this mutant has wild-type levels of activity on substrates without a downstream strand, we conclude gap recognition is an important component to pol μ 's ability to efficiently promote NHEJ of noncomplementary ends *in vitro*. However, we note R175A is still much more active than pol μ Δ loop (Δ 369-385) (Figure 3.2 C, compare lanes 5 and 7), a mutant previously defined as specifically defective in promoting NHEJ of noncomplementary ends [45]. Moreover, synthesis by R175A remains most efficient with ddATP, the nucleotide complementary to template (Figure 3.2 D). R175A thus possesses high residual template dependent activity in the NHEJ assay, con-

sistent with it retaining interactions with the downstream strand that are independent of R175 (see above).

Does mutation of R175 have a similar impact on NHEJ in cells? We previously used a pre-B cell line to assess relative activities of variant pol μ constructs in cellular NHEJ [45]. In mice, pol μ is essential for accurate resolution of intermediates in V(D)J recombination at Ig κ loci [9]. A pre-B cell line transformed with a temperature sensitive variant of the Abelson murine leukemia virus (ts-abl) can be induced to undergo high levels of Igk recombination in culture [58], but levels of deletion in Igk recombination junctions approach that seen in pre-B cells from pol μ deficient mice. Critically, stable overexpression of wild type pol μ , but not catalytically defective pol μ , is sufficient to correct this apparent defect [45].

We therefore generated a variant clone of this pre-B cell line that overexpresses R175A (Figure 3.3 A), and compared the accuracy of Igk recombination junctions to that seen in cells overexpressing wild type pol μ (Figure 3.3 B). Igk recombination in cells overexpressing R175A is less accurate than when wild type pol μ is overexpressed ($p < 0.05$; Mann-Whitney test), but more accurate than in the parental line, or in a line over-expressing pol μ Δ loop). Thus, the phenotype of the R175A mutant in cellular NHEJ is consistent with *in vitro* NHEJ results, where R175A's activity is similarly intermediate between wild type pol μ and the severely defective pol μ Δ loop mutant (Figure 3.2 C). Given that R175A is only partly defective in recognition of gapped substrates (Figure 3.2 B; Table 3.2) we conclude pol μ 's ability to bridge gaps between ends is an important component of its ability to promote NHEJ both *in vitro* and in cells.

3.4 Discussion

Previous work has shown that pol μ is uniquely able to facilitate repair of noncomplementary overhangs, and we have now identified some structural elements critical for pol μ 's ability to do so [45]. H329 interacts with the primer terminus and the incoming nucleotide. It is also conserved in TdT but not pol X family members pol λ and pol β , suggesting it is important for the transferase activities of pol μ and TdT. Indeed, it has been shown that mutation to alanine in both pol μ and in TdT results in a dramatic loss of template independent activity [32]. Pol μ 's gap filling activity was less than two-fold reduced by the mutation, however, indicating that in the presence of a template DNA molecule that interacts with the primer H329 is not critically important for synthesis. We examined the effect of H329 on end-joining in particular, and found that when the DNA ends are partially complementary, H329 is not critical for pol μ activity. However when the ends are noncomplementary, pol μ 's activity depends on H329. This supports the crystal structure's prediction that H329 is needed to properly position the incoming nucleotide and the primer terminus in the polymerase active site when there is no template or when it does not interact with the primer. It seems these elements are correctly positioned by other contacts when there are base pairing interactions between the two termini.

Having thus examined an element common to TdT and pol μ , we turned our attention to an element common to pol λ , pol β , and pol μ . We mutated R175, which was predicted by the pol μ structure to interact with the downstream strand. Our results show this interaction is important for gap filling ability, though it is clearly not the only residue that helps pol μ interact with the downstream or template strands, as the mutant enzyme retains some activity on such substrates. Importantly, this provides further support for the argument that pol μ in NHEJ must act on a bridged-end complex of DNA, as we proposed in the last chapter. Because the downstream strand is paired with the template

strand, pol μ 's interaction with the downstream strand helps position the template and allows the polymerase to "see" the distal side of a gap, consistent with our prediction that pol μ must interact with the duplex region of the template DNA.

The results of these experiments show pol μ has structural elements in common with polymerases from either side of the proposed gradient of template dependence. Pol μ is less template dependent than pol λ and more template dependent than TdT, with structural explanations found in the Loop 1 regions as well as the overall charge of the 8 kDa domains of the pol X family [26]. Here we present more structural evidence for pol μ 's placement in the gradient: pol μ has a residue found only in the polymerase less template-dependent than itself, and a residue found only in the pol X members more template-dependent than itself. More importantly, this work also describes how pol μ is able to catalyze a reaction unique among pol X polymerases, the filling of gaps between noncomplementary ends. We propose no other polymerase does this because no other polymerase has the combination of structural elements (including H329 and R175) that allow stable interactions with the two DNA ends, even when the ends do not stabilize each other through base pairs. This ability is important during $\nu\kappa$ rearrangement in V(D)J recombination, and is an example of a difficulty NHEJ faces in DSB repair. Since no intact template is used by the pathway, it must be able to overcome a wide variety of barriers to repair, and pol μ is one of the tools that gives it such flexibility.

3.5 Acknowledgements

The work detailed here has been previously published in two articles. Experiments involving pol μ H329A were published in *Nature Structural and Molecular Biology* 14 (1): 45-53 in January of 2007 as part of a larger work (*Structural insight into the substrate specificity of DNA Polymerase μ*) detailing the x-ray crystal structure of pol μ . The other authors were A. F. Moon, M. Garcia-Diaz, K. Bebenek, X. Zhong, D. A.

Ramsden, T. A. Kunkel, and L. C. Pedersen [32]. I contributed the analysis of NHEJ properties of pol μ and pol μ H329A, regarding both overall NHEJ activity and synthesis activity by each of the polymerases.

Experiments involving pol μ R175A were first published as part of a larger work in Nucleic Acids Research 36 (9): 3085-94 in May of 2008 (*End-bridging is required for pol μ to efficiently promote repair of noncomplementary ends by nonhomologous end joining*). The other authors were J. M. Havener and D. A. Ramsden [63]. I contributed the analysis of NHEJ properties of pol μ and pol μ R175A, including both overall NHEJ activity and synthesis activity of the polymerases. J. M. Havener contributed the analysis of the intrinsic activity of the two polymerases and the cellular effect of the overexpression of pol μ R175A. D. A. Ramsden also contributed to the analysis of both sets of experiments.

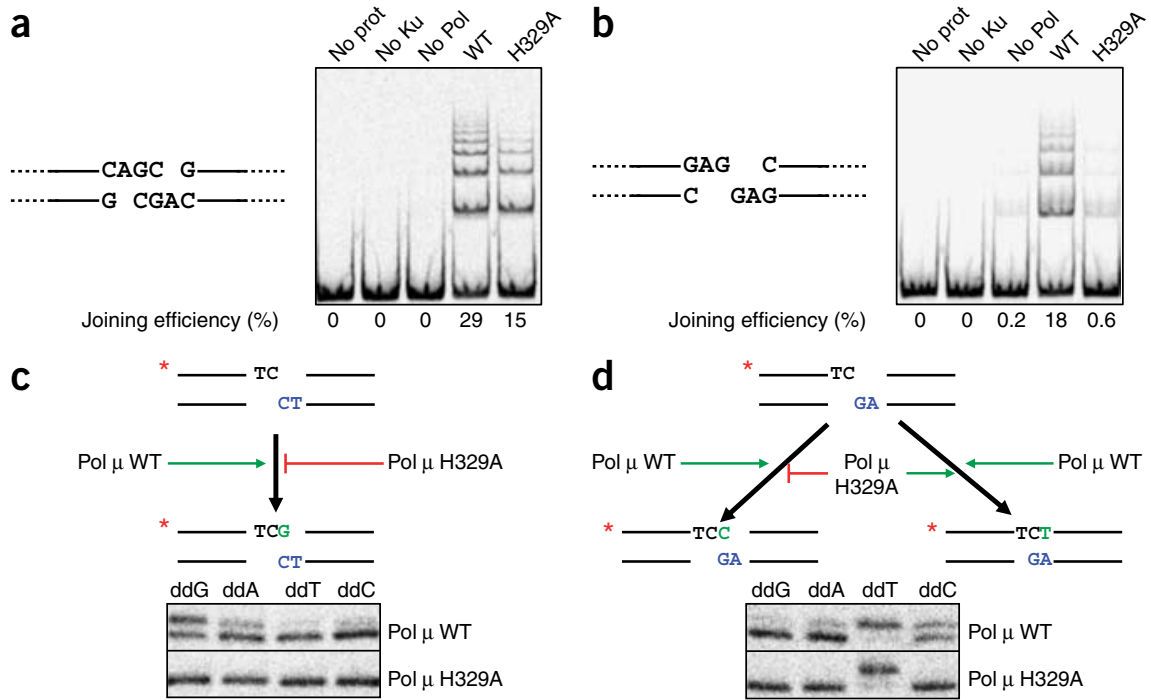


Figure 3.1: Role of Pol μ His329 in template-dependent synthesis during non-homologous end-joining. (a, b) NHEJ assays on DSB substrate with either two (a) or no complementary nucleotides (b). End structure of each 300-bp DNA substrate is depicted at left; spaces in DNA sequence indicate synthesis required for joining. End-joining efficiencies (see Methods) are noted below gel lanes. (c, d) Synthesis in the absence of joining. Substrates were labeled DNA with a TC 3'-overhang mixed with unlabeled DNA with either a TC 3' overhang (c) or an AG 3'-overhang (d). End structures of the 300-bp DNA substrates used for the proposed template-dependent synthesis activities are depicted; red asterisk marks labeled primer strand.

Originally appeared in [32]

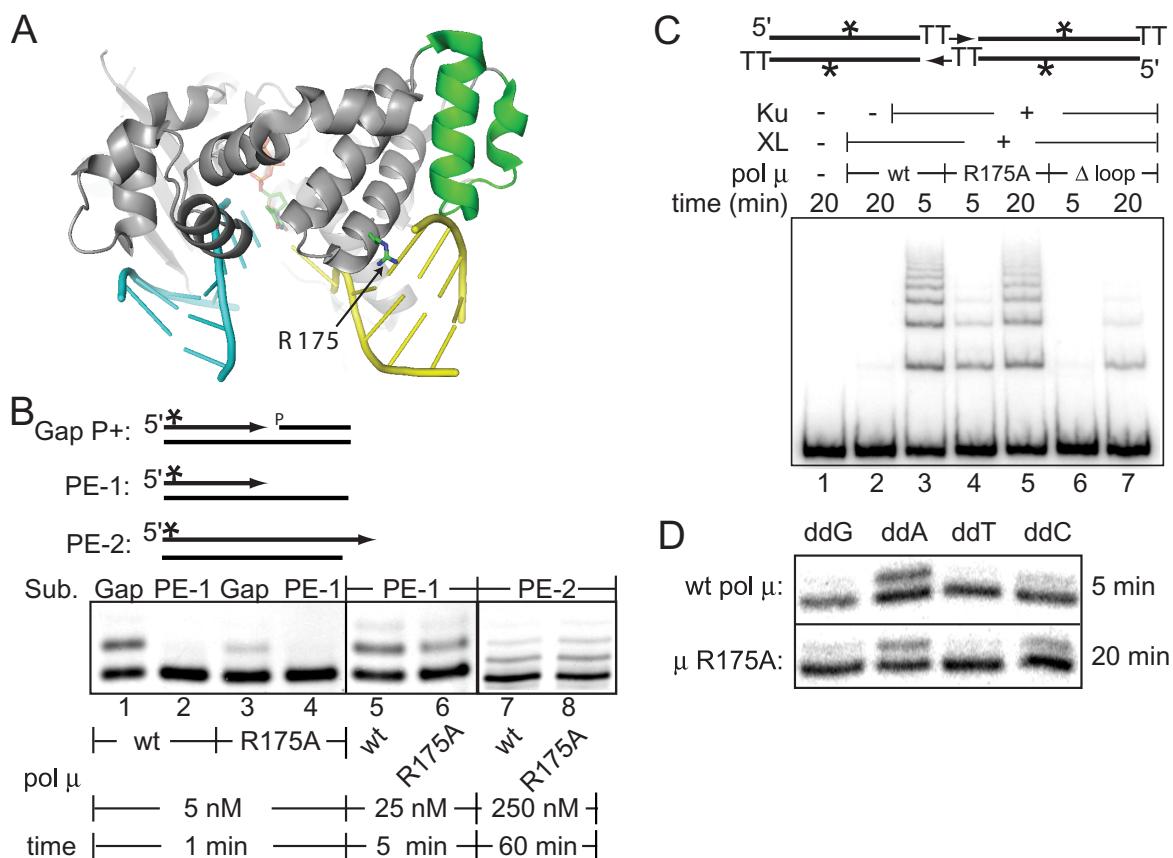


Figure 3.2: Activity of pol μ R175A *in vitro*. (A) An arrow locates R175 (in stick representation) in a structure of pol μ bound to a gapped DNA substrate with an incoming ddTTP (also in stick representation) (PDB code, 2IHM) [32]. The helix-hairpin-helix motif is in green. Upstream DNA duplex, including the primer, is in cyan, while downstream DNA duplex ('template end') is in yellow. Two nucleotides opposite the primer terminus (T6 and T7 in 2IHM) were omitted. (B) Gap and primer extension substrates (PE-1, template dependent; PE-2, template-independent) are shown, with 5' end labels noted by an asterisk. Assays were done with 5nM substrate and indicated amounts of wt pol μ or pol μ R175A and incubated for the lengths of time noted. Ku and XL were not included. (C) Reactions performed as in Figure 2.1 B, contained Ku, XL and wt pol μ , pol μ R175A or pol μ Δ loop as indicated. (D) Reactions contained Ku, XL, wt pol μ or pol μ R175A and 25 mM of the indicated ddNTP as in Figure 2.1 C for 5 or 20 min.

Originally appeared in [63]

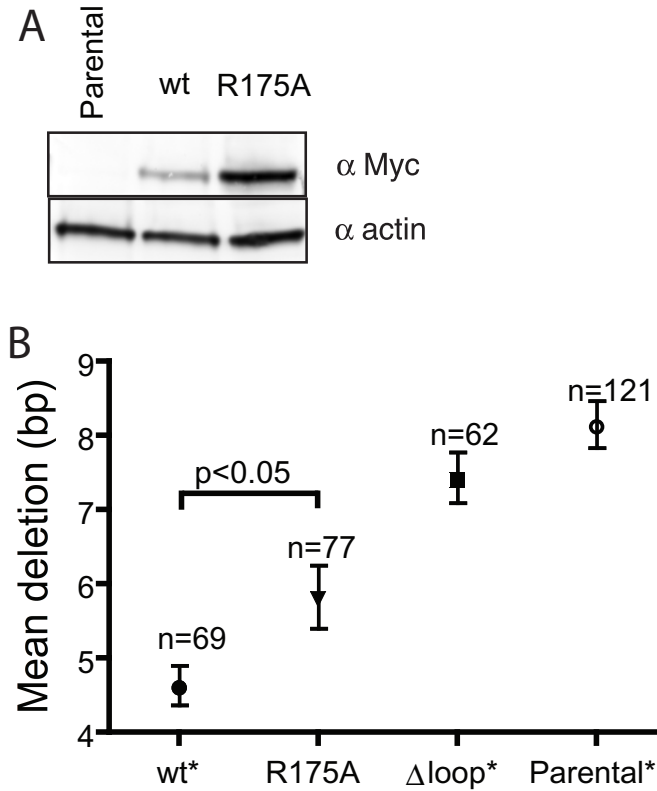


Figure 3.3: Activity of pol μ R175A in cells. (A) Western blot of whole cell extracts from the parental ts-abl pre-B-cell line (SP-9), a line stably over-expressing a myc-epitope tagged wild-type pol μ (wt), and a line stably over-expressing the myc epitope tagged R175A mutant pol μ (R175A). Extracts were probed with an anti-myc antibody or, to verify similar quantities of extract loaded, an antibody to actin. (B) Average number of deletions from recombined Igk loci in the pre-B-cell line for wt pol μ , pol μ R175A, pol μ Δ loop and the parental cell line. Error bars mark the SEM, and the number of Ig κ junctions sequenced for each cell line is noted. Asterisks identify data from ref. [45].

Originally appeared in [63]

Gap-filling synthesis on dsDNA gapped substrate :

Enzyme	k_{cat} (s^{-1})	K_m (μM)	k_{cat}/K_m ($s^{-1}\mu M^{-1}$)
Wildtype hPol μ	0.047 ± 0.024	11 ± 4.7	$(4.3 \pm 0.87) \times 10^{-3}$
hPol μ H329A	0.015 ± 0.007	6.0 ± 3.2	$(2.8 \pm 0.8) \times 10^{-3}$

Synthesis on ssDNA substrate:

Enzyme	k_{cat} (s^{-1})	K_m (μM)	k_{cat}/K_m ($s^{-1}\mu M^{-1}$)
Wildtype hPol μ	$(2.8 \pm 0.9) \times 10^{-3}$	83 ± 38	$(43 \pm 29) \times 10^{-6}$
hPol μ H329A	Not detected	Not detected	---

Table 3.1: Steady-state kinetic analysis of nucleotide incorporation. Errors shown are s.d.

Originally appeared in [32]

Substrate ¹	Polymerase activity ²	
	Wt	R175A
Gap filling P+ (GF P+)	1	0.26
Gap filling P-	0.28	0.17 (0.6)
Primer extension (PE-1)	3.8×10^{-2}	2.3×10^{-2} (0.6)
TdT-like (PE-2)	2.3×10^{-4}	3.3×10^{-4} (1.4)

Table 3.2: Polymerase activity on oligonucleotide substrates. ¹ GF P+, PE-1 and PE-2 refer to substrates described in Fig. 3.2B. The single nucleotide gap substrate was also tested without a 5'-phosphate on the down stream strand (P-). PE-2 assays required the presence of 10% polyethylene glycol to recover detectable activity. ² Specific activities are expressed relative to wt pol μ on a 1-nt gap with a 5' phosphate on the downstream strand (GF P+). The same substrate was also tested without this 5' phosphate (P-); other substrates as in Figure 3.2 B. For R175A, we also list in parenthesis the specific activity of pol μ R175A relative to wt pol μ on the same substrate. TdT-like activity is the average activity from reactions with each dNTP tested individually.

Originally appeared in [63]

Chapter 4

Discussion

4.1 Pol μ 's activity in NHEJ

NHEJ faces a challenge unique among DNA repair pathways. It must repair DSBs without an intact copy of DNA to use as a template to instruct repair [66]. When the DNA ends that result from the DSB are damaged or are missing nucleotides, NHEJ employs nucleases and pol X family polymerases to make the ends ligatable. A range of template dependence within the polymerase family allows NHEJ flexibility in the ends it can repair. In this dissertation, I examined the ability of pol μ to facilitate end joining when the ends are noncomplementary.

Two models for pol μ 's activity in this context had been proposed. One suggested pol μ 's template-independent activity was responsible for adding the nucleotides that filled the gap through cycles of random addition until a base complementary to the template was added, allowing ligation by XL [56]. The other model proposed pol μ 's template dependent activity was the one that was primarily responsible for promoting ligation [45]. Pol μ would use the identity of a base in the template strand to instruct its addition of a nucleotide to the primer, filling the gap and allowing XL to ligate the ends.

In chapter 2 of this dissertation I presented a systematic analysis of pol μ 's activity

in NHEJ of these kinds of ends by varying the identity of the terminal bases in the primer and template. Using a system that decoupled ligation from synthesis, I was able to determine the identity of the base pol μ added. The results indicated some cases when pol μ is neither template-directed nor efficient in adding a nucleotide. These cases included contexts when the “template” was blunt, when there were long ($>2\text{nt}$) gaps and when certain combinations of primer/template nucleotides were present. Nonetheless, in most cases pol μ was efficient and template directed; bases added within 2.5 minutes were preferentially complementary to the template base.

The length of the overhang was an important consideration for pol μ activity. With gaps longer than two nucleotides, it primarily used bases close to the duplex region of the template molecule as the template nucleotides, ignoring the terminal bases. Because template-dependent pol X family members interact primarily with the template strand through interactions with the downstream strand, this observation implied pol μ needs contacts with the distal side of the gap it is filling in order to function effectively. In turn, this provides further support for the model predicting pol μ uses its template-dependent activity to fill gaps resulting from noncomplementary overhangs.

We therefore sought to determine some structural features that contribute to pol μ 's ability to fill such gaps. The recently solved crystal structure of pol μ suggested two amino acid residues that might be important for this activity: H329 and R175. H329 was well-positioned to stabilize interactions between pol μ and the primer terminus as well as between pol μ and the incoming nucleotide. This residue is conserved in TdT but is not found in pol β or pol λ , suggesting it is important for template independence in pol X family members. Indeed, mutation of H329 in pol μ and its TdT counterpart abolished template independent activity. I showed, however, that in NHEJ of partially complementary overhangs, H329 is not critical for pol μ activity. Interestingly, it is essential for NHEJ of noncomplementary overhangs, suggesting elements important for

transferase activity enable pol μ to facilitate gap filling of this difficult substrate.

This might seem like a contradiction in experimental results: pol μ uses an element important for its template-independent activity to do a reaction instructed by a template. However, we also investigated a structural feature pol μ shares with exclusively template dependent pol X family members. Positively charged R175 has counterparts in pol λ and pol β and was predicted to interact with the distal side of a gapped DNA substrate at the 5'-phosphate. We mutated the residue to alanine and tested the mutant polymerase. Our results indicated not only that pol μ gap filling activity was defective, but also that NHEJ of noncomplementary overhangs was impaired. This indicates it is important for pol μ to interact with the template molecule regardless of the interactions between the DNA strands.

The structure of pol μ also suggested H208 interacts with the same point on the downstream strand as R175 (Figure 4.1) [32]. Since pol μ R175A exhibited a partial loss of function only, H208 might provide enough stability to form the bridged-end complex pol μ requires for gap filling. Similarly, a H208A mutant might be partially defective in gap filling the same way R175A is. It would be interesting to test both H208A and a double mutant in the same way we tested R175. We could then determine whether H208 or R175 contributes more to pol μ 's ability to form a bridged-end complex and whether those two residues together account for all of pol μ 's ability to track the 5'-phosphate of the downstream strand. If that is the case, we expect all the gap-filling activities of pol μ to be severely affected, but not its weak transferase activity.

These results lead us to suggest pol μ is uniquely suited to repair of noncomplementary ends because it has structural features found in two extremes of the pol X family: end-bridging elements from pol β and pol λ , and "template bypass" elements from TdT. Individually, these structural features account for pol μ 's template-independent and template-dependent activities, but the combination of them is required to facilitate

end joining of noncomplementary overhangs. With this view, pol μ can be seen as more than a weak transferase or a promiscuous version of pol λ . Rather, its can be viewed as a purpose-built polymerase designed to allow NHEJ to work in a difficult context. The transferase activity and ability to function in NHEJ of partially complementary overhangs might be seen more as byproducts of this activity than as the primary modes for pol μ .

The experimental observations described here provide more structural and functional support for pol μ 's position in our proposed gradient of template dependence between pol λ and TdT that was based on observations that pol μ 's Loop 1 is longer than pol λ 's but perhaps more flexible than TdT's. The length of the loop suggests it can stabilize interactions with a primer strand in the absence of a template, but the flexibility has been proposed to allow pol μ to switch between template-independent and template dependent modes of synthesis [52].

To more fully asses the implications of these experiments, it would be interesting to generate chimeras of pol X family members and test their activities. For example, if a residue equivalent to H329 and a Loop 1 like pol μ 's were added to pol λ , would it then gain the ability to facilitate NHEJ of noncomplementary overhangs? Does addition of positively charged residues, such as R175 and H208, to TdT confer the ability to use a template strand to instruct synthesis or does its Loop 1 have to be made flexible as well?

The range of polymerases that can be recruited by NHEJ core factors to help it process ends before ligation makes the pathway a remarkably flexible one. By mechanisms described here, pol μ activity allows the pathway to repair breaks that result in noncomplementary ends while still maintaining the genetic information present in these overhangs. By recruiting either pol μ or pol λ to DSBs where the ends are partially complementary, NHEJ can also repair those breaks in a similar way. Additionally, the

range of polymerases allows NHEJ to act in situations where diversity in DNA sequence, rather than conservation of it, is the desired result. TdT's template-independent activity causes NHEJ to repair DSBs generated during V(D)J recombination in such a way that the developing lymphocytes collectively gain a wide range of antigen receptor sequence. This in turn contributes to the adaptive immune system's ability to recognize a wide range of antigens. Thus NHEJ is able to repair DSBs when it is beneficial to maintain genetic information and when it is beneficial to diversify that information.

4.2 Polymerase choice

The polymerase selected by NHEJ can thus be seen as a critical component that determines whether DSB repair is conservative or not with regard to the overhanging DNA sequence. How is this choice controlled? Restricted expression clearly plays a role in keeping TdT out of general DSB repair [67] where its template-independent additions would be detrimental, but even if this is the only way the choice to use TdT is controlled, two questions remain about the other polymerases recruited to NHEJ core factors through BRCT domains: 1) how does NHEJ choose between pol λ and pol μ , and 2) given that pol μ does every *in vitro* reaction that pol λ does, why bother with pol λ ? Both polymerases are widely expressed [44, 46, 68, 28] and though different roles for each have been identified in V(D)J recombination, it is not yet clear whether a unique role exists for pol λ in general DSB repair or what such a role might be. One study shows an overexpression of catalytically inactive pol λ that might out-compete native, functional pol λ results in an increase of chromosomal aberrations and increased sensitivity to ionizing radiation compared to parental cells and cells overexpressing wt pol λ [69]. Unfortunately, this does not adequately address a unique role of pol λ . Data from our group (J.M. Havener; Figure 4.2) indicates pol λ is able to out-compete pol μ for complexes of the core NHEJ factors, even at a 10-fold lower concentration.

Thus, overexpression of a catalytically dead pol λ would out-compete both native pol λ and native pol μ . At best, then, this study does not rule out a role for pol λ in NHEJ in response to IR, leaving a unique role for pol λ undefined. Other studies have shown loss of pol λ and pol β in mice and MEFs results in an increased sensitivity to oxidative damage [47, 48], but this observation has not been explained at a biochemical level. Therefore I think it would be worthwhile to determine whether there is a unique role for pol λ in NHEJ, define it, and determine factors that allow the selection of one polymerase over another at a given DSB.

The fact that pol λ out-competes pol μ for the complex of core NHEJ factors provides the basis for a model that might explain how the selection of one polymerase over the other takes place. Pol λ 's higher affinity for core factors suggests it might be the first option for NHEJ. If the DNA ends are partially complementary, pol λ is likely able to help NHEJ resolve the DSB and no further polymerase recruitment is required. If the DNA ends are not complementary, pol λ will be unable to facilitate repair before disengaging. At some rate, pol μ will then bind the complex of Ku and XL and it will be able to fill the gap before it disengages, allowing XL to ligate the ends. Thus, NHEJ may “choose” between polymerases by trying pol λ first and using pol μ as a backup if pol λ is unable to fill the gap.

One way of testing this model requires being able to determine after the fact which polymerase filled a gap. *In vitro* this could be accomplished if it was found that pol λ was more accurate when filling a particular kind of gap than pol μ . I showed in this dissertation pol μ has low activity filling gaps longer than 2 nucleotides. If pol λ were able to accurately fill, for example, gaps of three nucleotides, the active polymerase could be determined by measuring the accuracy of the junctions. The model predicts increased accuracy of junctions in an end-joining experiment with both polymerases than with pol μ alone. Further, if this was true, it would also suggest a preferred role

for pol λ in NHEJ of gaps longer than two nucleotides, defining pol λ 's role in NHEJ.

I have conducted tests of these hypotheses, and unfortunately I have been unable to generate data which conclusively support or disprove them. *In vitro* activity assays like those just described have not yet shown a preference for pol λ ; if anything, pol μ appears more active on every substrate tried. Further, my efforts to define a particular substrate for which pol λ is more accurate than pol μ are so far inconclusive as well. Some factors that may be at play in confounding these results include the conditions of the particular activity assays and the overall activity of the polymerases. Nonetheless, these questions remain important for completing the functional portrait of the pol X family in mammals.

4.3 Concluding Remarks

In this dissertation I have focused on the way NHEJ is able to repair DSBs that result in noncomplementary ends. I first showed that when pol μ efficiently fills those gaps, it is template directed. I also presented evidence pol μ is inefficient on gaps longer than two nucleotides, suggesting a requirement for interactions with both strands of the downstream duplex. I then examined structural elements that lead to pol μ 's ability to facilitate end joining of noncomplementary overhangs.

One way to think about this substrate for repair is as an intermediate between two kinds of substrates: 1) partially complementary overhangs and 2) primers without templates. Noncomplementary overhangs and partially complementary overhangs are similar insofar as both present two ends that need to be religated. However, noncomplementary overhangs also resemble a primer without a template, since no DNA interacts with the primer terminus. Partially complementary overhangs could be repaired by pol λ while TdT is the polymerase best suited to act on primers lacking templates; the polymerase that helps repair "hybrid" overhangs is pol μ . It is not surprising, therefore, that

pol μ , the polymerase that works on the intermediate substrate, displays characteristics of both pol λ and TdT.

Thus have I shown how NHEJ is able to repair noncomplementary overhangs, which is one of the ways NHEJ overcomes the unique problem of repairing DNA in the absence of an intact copy to use as a template. I have added support for pol μ 's position in our proposed gradient of template dependence in the pol X family while detailing some limitations on pol μ 's activity in NHEJ. Further work with polymerase chimeras would indicate whether the interactions I described are the only ones that allow pol μ to hold its place in the gradient. Because of its unique structural set-up, I suggest pol μ can be seen as “purpose-built” for NHEJ of noncomplementary overhangs. With that view, and other evidence from another member of our group, I propose pol μ is the second choice for NHEJ in gap filling. I predict pol λ is the first choice and pol μ is only tried if pol λ is unable to facilitate repair.

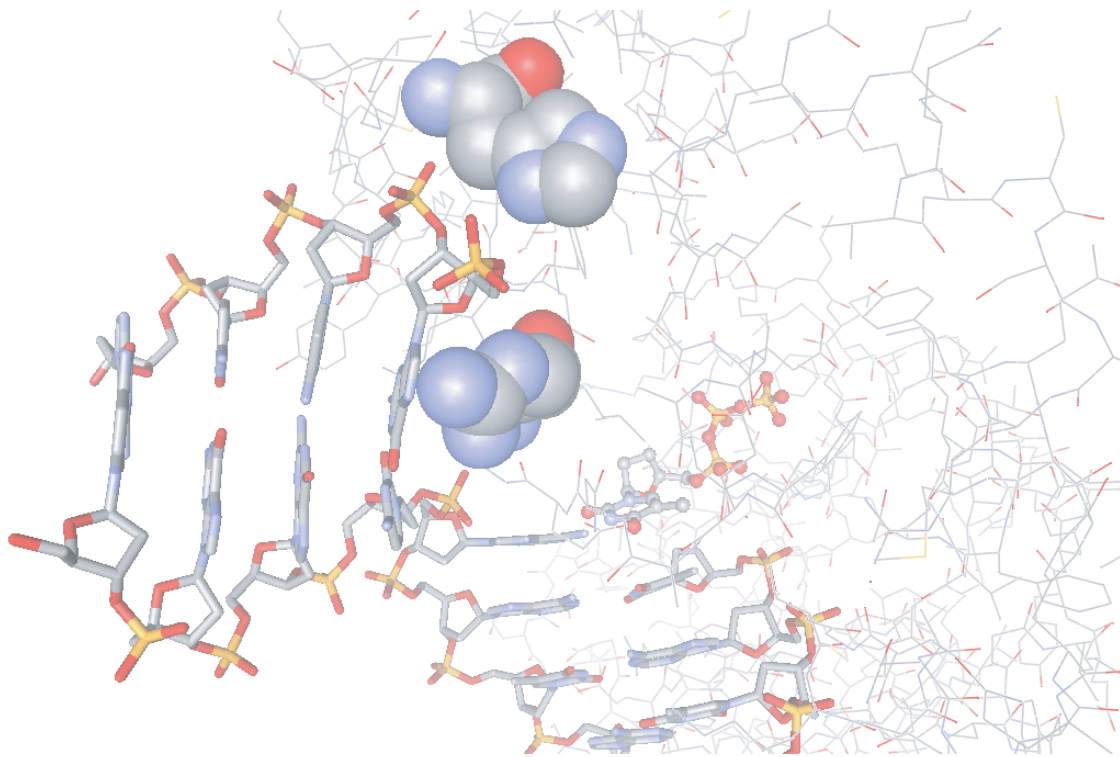


Figure 4.1: H208 is also predicted to stabilize the 5'-phosphate of the downstream strand. H208 and R175 are represented as space-filling models while the rest of pol μ is represented with lines. DNA molecules are represented as sticks and the incoming ddTTP is in ball-and-stick form. The 5'-phosphate from the downstream strand is between H208 and R175. 3D Molecule Viewer was used to generate the image based on the crystal structure of pol μ (PDB 2IHM).

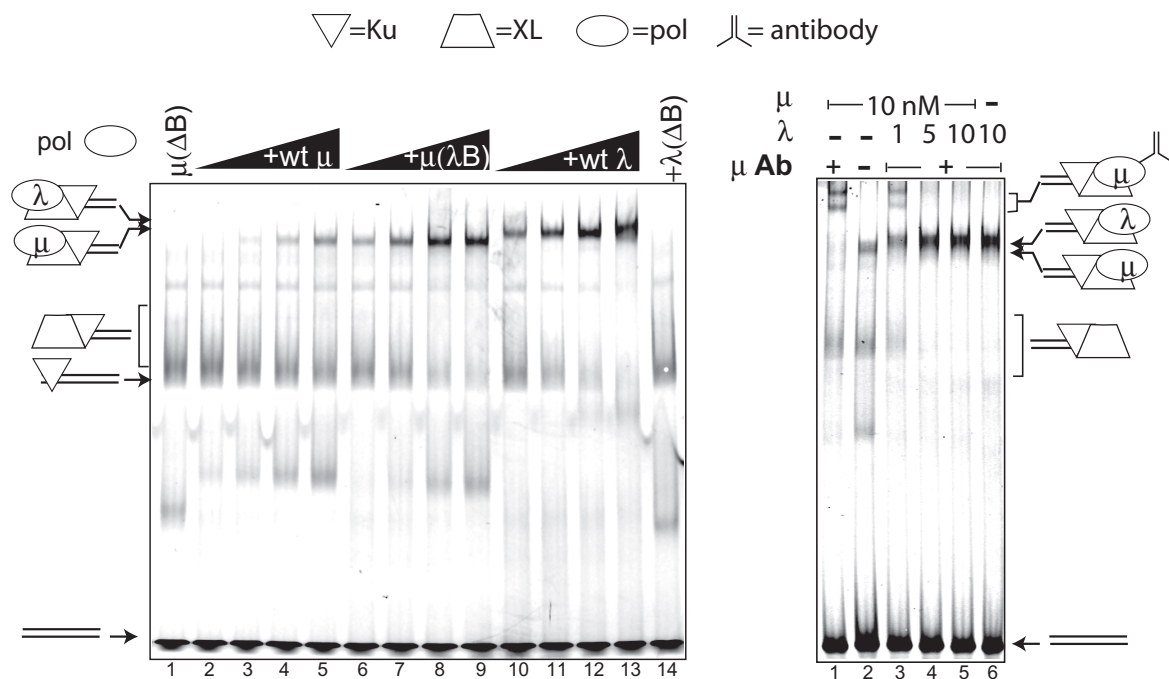


Figure 4.2: Pol λ forms the more stable complex with NHEJ core factors. Left side: an electromobility shift assay (EMSA) with NHEJ core factors and two pol X family polymerases. The short DNA duplex was incubated with 1 nM Ku, 10 nM XL, and increasing amounts of pol μ (lanes 2-5), pol λ (lanes 10-13) or a chimera of pol μ with the BRCT domain of pol λ (lanes 6-9). The complexes were run on a 3.5% polyacrylamide gel, and the diagrams of the substrate and proteins which make up the complexes at their appropriate mobilities are noted on the left side of the figure. Control reactions with polymerases having had their BRCT domains deleted are in lanes 1 and 14. Right side: an EMSA similar to the one on the right side of the figure. An antibody to pol μ is added as noted. Polymerases were added at the listed concentrations.

BIBLIOGRAPHY

- [1] Helleday, T., Lo, J., vanGent, D. C., and Engelward, B. P. July 2007 *DNA Repair* **6(7)**, 923–35 PMID: 17363343.
- [2] Sonoda, E., Takata, M., Yamashita, Y. M., Morrison, C., and Takeda, S. July 2001 *Proceedings of the National Academy of Sciences of the United States of America* **98(15)**, 8388–94 PMID: 11459980.
- [3] West, S. C. June 2003 *Nature Reviews. Molecular Cell Biology* **4(6)**, 435–45 PMID: 12778123.
- [4] Featherstone, C. and Jackson, S. P. October 1999 *Current Biology: CB* **9(20)**, R759–61 PMID: 10531043.
- [5] Ma, Y., Lu, H., Tippin, B., Goodman, M. F., Shimazaki, N., Koiwai, O., Hsieh, C.-L., Schwarz, K., and Lieber, M. R. December 2004 *Molecular Cell* **16(5)**, 701–13 PMID: 15574326.
- [6] Budman, J., Kim, S. A., and Chu, G. April 2007 *The Journal of Biological Chemistry* **282(16)**, 11950–9 PMID: 17272270.
- [7] Rothkamm, K., Krger, I., Thompson, L. H., and Lbrich, M. August 2003 *Molecular and Cellular Biology* **23(16)**, 5706–15 PMID: 12897142.
- [8] Karran, P. April 2000 *Current Opinion in Genetics & Development* **10(2)**, 144–50 PMID: 10753787.
- [9] Bertocchi, B., Smet, A. D., Weill, J.-C., and Reynaud, C.-A. July 2006 *Immunity* **25(1)**, 31–41 PMID: 16860755.
- [10] Lieber, M. R. July 2006 *Immunity* **25(1)**, 7–9 PMID: 16860749.
- [11] McElhinny, S. A. N. and Ramsden, D. A. August 2004 *Immunological Reviews* **200**, 156–64 PMID: 15242403.
- [12] Gellert, M. (2002) *Annual Review of Biochemistry* **71**, 101–32 PMID: 12045092.
- [13] Weterings, E. and vanGent, D. C. November 2004 *DNA Repair* **3(11)**, 1425–35 PMID: 15380098.
- [14] McElhinny, S. A. N., Snowden, C. M., McCarville, J., and Ramsden, D. A. May 2000 *Molecular and Cellular Biology* **20(9)**, 2996–3003 PMID: 10757784.
- [15] Falzon, M., Fewell, J. W., and Kuff, E. L. May 1993 *The Journal of Biological Chemistry* **268(14)**, 10546–52 PMID: 8486707.

- [16] Rathmell, W. K. and Chu, G. July 1994 *Molecular and Cellular Biology* **14**(7), 4741–8 PMID: 7516471.
- [17] Paillard, S. and Strauss, F. October 1991 *Nucleic Acids Research* **19**(20), 5619–24 PMID: 1945839.
- [18] Reddy, Y. V. R., Ding, Q., Lees-Miller, S. P., Meek, K., and Ramsden, D. A. September 2004 *The Journal of Biological Chemistry* **279**(38), 39408–13 PMID: 15258142.
- [19] Gottlieb, T. M. and Jackson, S. P. November 1994 *Trends in Biochemical Sciences* **19**(11), 500–3 PMID: 7855895.
- [20] Chen, L., Trujillo, K., Sung, P., and Tomkinson, A. E. August 2000 *The Journal of Biological Chemistry* **275**(34), 26196–205 PMID: 10854421.
- [21] Lieber, M. R., Ma, Y., Pannicke, U., and Schwarz, K. September 2003 *Nature Reviews. Molecular Cell Biology* **4**(9), 712–20 PMID: 14506474.
- [22] Daley, J. M., Palmbo, P. L., Wu, D., and Wilson, T. E. (2005) *Annual Review of Genetics* **39**, 431–51 PMID: 16285867.
- [23] Koch, C. A., Agyei, R., Galicia, S., Metalnikov, P., O'Donnell, P., Starostine, A., Weinfeld, M., and Durocher, D. October 2004 *The EMBO Journal* **23**(19), 3874–85 PMID: 15385968.
- [24] Rooney, S., Sekiguchi, J., Zhu, C., Cheng, H. L., Manis, J., Whitlow, S., DeVido, J., Foy, D., Chaudhuri, J., Lombard, D., and Alt, F. W. December 2002 *Molecular Cell* **10**(6), 1379–90 PMID: 12504013.
- [25] Ma, Y., Pannicke, U., Schwarz, K., and Lieber, M. R. March 2002 *Cell* **108**(6), 781–94 PMID: 11955432.
- [26] Moon, A. F., Garcia-Diaz, M., Batra, V. K., Beard, W. A., Bebenek, K., Kunkel, T. A., Wilson, S. H., and Pedersen, L. C. (2007) *DNA Repair* **6**(12), 1709–1725.
- [27] Bebenek, K., Garcia-Diaz, M., Patishall, S. R., and Kunkel, T. A. May 2005 *The Journal of Biological Chemistry* **280**(20), 20051–8 PMID: 15778218.
- [28] Garca-Daz, M., Bebenek, K., Sabariego, R., Domnguez, O., Rodriguez, J., Kirchhoff, T., Garca-Palomero, E., Picher, A. J., Jurez, R., Ruiz, J. F., Kunkel, T. A., and Blanco, L. April 2002 *The Journal of Biological Chemistry* **277**(15), 13184–91 PMID: 11821417.
- [29] Garca-Daz, M., Bebenek, K., Kunkel, T. A., and Blanco, L. September 2001 *The Journal of Biological Chemistry* **276**(37), 34659–63 PMID: 11457865.
- [30] Garcia-Diaz, M., Bebenek, K., Krahn, J. M., Kunkel, T. A., and Pedersen, L. C. (2005) *Nature Structural & Molecular Biology* **12**(1), 97–8 PMID: 15608652.

- [31] Pelletier, H., Sawaya, M. R., Wolffe, W., Wilson, S. H., and Kraut, J. October 1996 *Biochemistry* **35**(39), 12742–61 PMID: 8841118.
- [32] Moon, A. F., Garcia-Diaz, M., Bebenek, K., Davis, B. J., Zhong, X., Ramsden, D. A., Kunkel, T. A., and Pedersen, L. C. (2007) *Nature Structural & Molecular Biology* **14**(1), 45–53 PMID: 17159995.
- [33] Gryk, M. R., Marintchev, A., Maciejewski, M. W., Robertson, A., Wilson, S. H., and Mullen, G. P. December 2002 *Structure (London, England: 1993)* **10**(12), 1709–20 PMID: 12467578.
- [34] Beard, W. A. and Wilson, S. H. August 2000 *Mutation Research* **460**(3-4), 231–44 PMID: 10946231.
- [35] Uchiyama, Y., Takeuchi, R., Kodera, H., and Sakaguchi, K. July 2008 *Biochimie* PMID: 18706967.
- [36] Tseng, H.-M. and Tomkinson, A. E. November 2002 *The Journal of Biological Chemistry* **277**(47), 45630–7 PMID: 12235149.
- [37] Wilson, T. E. and Lieber, M. R. August 1999 *The Journal of Biological Chemistry* **274**(33), 23599–609 PMID: 10438542.
- [38] Rooney, S., Chaudhuri, J., and Alt, F. W. August 2004 *Immunological Reviews* **200**, 115–31 PMID: 15242400.
- [39] Fan, W. and Wu, X. October 2004 *Biochemical and Biophysical Research Communications* **323**(4), 1328–33 PMID: 15451442.
- [40] Lee, J. W., Blanco, L., Zhou, T., Garcia-Diaz, M., Bebenek, K., Kunkel, T. A., Wang, Z., and Povirk, L. F. (2004) *The Journal of Biological Chemistry* **279**(1), 805–11 PMID: 14561766.
- [41] Mueller, G. A., Moon, A. F., Deroose, E. F., Havener, J. M., Ramsden, D. A., Pedersen, L. C., and London, R. E. August 2008 *DNA Repair* **7**(8), 1340–51 PMID: 18585102.
- [42] DeRose, E. F., Clarkson, M. W., Gilmore, S. A., Galban, C. J., Tripathy, A., Havener, J. M., Mueller, G. A., Ramsden, D. A., London, R. E., and Lee, A. L. October 2007 *Biochemistry* **46**(43), 12100–10 PMID: 17915942.
- [43] Gilfillan, S., Benoist, C., and Mathis, D. December 1995 *Immunological Reviews* **148**, 201–19 PMID: 8825288.
- [44] Aoufouchi, S., Flatter, E., Dahan, A., Faily, A., Bertocci, B., Storck, S., Delbos, F., Cocea, L., Gupta, N., Weill, J. C., and Reynaud, C. A. September 2000 *Nucleic Acids Research* **28**(18), 3684–93 PMID: 10982892.

- [45] McElhinny, S. A. N., Havener, J. M., Garcia-Diaz, M., Jurez, R., Bebenek, K., Kee, B. L., Blanco, L., Kunkel, T. A., and Ramsden, D. A. August 2005 *Molecular Cell* **19(3)**, 357–66 PMID: 16061182.
- [46] Bertocci, B., Smet, A. D., Berek, C., Weill, J.-C., and Reynaud, C.-A. August 2003 *Immunity* **19(2)**, 203–11 PMID: 12932354.
- [47] Braithwaite, E. K., Prasad, R., Shock, D. D., Hou, E. W., Beard, W. A., and Wilson, S. H. May 2005 *The Journal of Biological Chemistry* **280(18)**, 18469–75 PMID: 15749700.
- [48] Bertocci, B., Smet, A. D., Flatter, E., Dahan, A., Bories, J.-C., Landreau, C., Weill, J.-C., and Reynaud, C.-A. April 2002 *Journal of Immunology (Baltimore, Md.: 1950)* **168(8)**, 3702–6 PMID: 11937519.
- [49] Sobol, R. W., Horton, J. K., Khn, R., Gu, H., Singhal, R. K., Prasad, R., Rajewsky, K., and Wilson, S. H. (1996) *Nature* **379(6561)**, 183–6 PMID: 8538772.
- [50] Srivastava, D. K., Berg, B. J., Prasad, R., Molina, J. T., Beard, W. A., Tomkinson, A. E., and Wilson, S. H. August 1998 *The Journal of Biological Chemistry* **273(33)**, 21203–9 PMID: 9694877.
- [51] Delarue, M., Boul, J. B., Lescar, J., Expert-Bezanon, N., Jourdan, N., Sukumar, N., Rougeon, F., and Papanicolaou, C. February 2002 *The EMBO Journal* **21(3)**, 427–39 PMID: 11823435.
- [52] Jurez, R., Ruiz, J. F., McElhinny, S. A. N., Ramsden, D., and Blanco, L. (2006) *Nucleic Acids Research* **34(16)**, 4572–82 PMID: 16963491.
- [53] Ramsden, D. (1994) Recombination and repertoire the molecular biology of V(D)J rearrangement, and how V(D)J rearrangement affects B cell development and the B cell repertoire., National Library of Canada = Bibliotheque nationale du Canada, Ottawa.
- [54] Schwede, T., Kopp, J., Guex, N., and Peitsch, M. C. July 2003 *Nucleic Acids Research* **31(13)**, 3381–5 PMID: 12824332.
- [55] Ward, J. F. November 1994 *International Journal of Radiation Biology* **66(5)**, 427–32 PMID: 7983426.
- [56] Gu, J., Lu, H., Tippin, B., Shimazaki, N., Goodman, M. F., and Lieber, M. R. February 2007 *The EMBO Journal* **26(4)**, 1010–23 PMID: 17290226.
- [57] McElhinny, S. A. N. and Ramsden, D. A. April 2003 *Molecular and Cellular Biology* **23(7)**, 2309–15 PMID: 12640116.
- [58] Chen, Y. Y., Wang, L. C., Huang, M. S., and Rosenberg, N. March 1994 *Genes & Development* **8(6)**, 688–97 PMID: 7926759.

- [59] Brissett, N. C., Pitcher, R. S., Juarez, R., Picher, A. J., Green, A. J., Dafforn, T. R., Fox, G. C., Blanco, L., and Doherty, A. J. October 2007 *Science (New York, N.Y.)* **318(5849)**, 456–9 PMID: 17947582.
- [60] Gu, J., Lu, H., Tsai, A. G., Schwarz, K., and Lieber, M. R. (2007) *Nucleic Acids Research* **35(17)**, 5755–62 PMID: 17717001.
- [61] Tsai, C. J., Kim, S. A., and Chu, G. May 2007 *Proceedings of the National Academy of Sciences of the United States of America* **104(19)**, 7851–6 PMID: 17470781.
- [62] Daley, J. M., Laan, R. L. V., Suresh, A., and Wilson, T. E. August 2005 *The Journal of Biological Chemistry* **280(32)**, 29030–7 PMID: 15964833.
- [63] Davis, B. J., Havener, J. M., and Ramsden, D. A. May 2008 *Nucleic Acids Research* **36(9)**, 3085–94 PMID: 18397950.
- [64] Sawaya, M. R., Prasad, R., Wilson, S. H., Kraut, J., and Pelletier, H. September 1997 *Biochemistry* **36(37)**, 11205–15 PMID: 9287163.
- [65] Garcia-Diaz, M., Bebenek, K., Krahm, J. M., Blanco, L., Kunkel, T. A., and Pedersen, L. C. February 2004 *Molecular Cell* **13(4)**, 561–72 PMID: 14992725.
- [66] vanGent, D. C. and van derBurg, M. December 2007 *Oncogene* **26(56)**, 7731–40 PMID: 18066085.
- [67] Desiderio, S. V., Yancopoulos, G. D., Paskind, M., Thomas, E., Boss, M. A., Landau, N., Alt, F. W., and Baltimore, D. *Nature* **311(5988)**, 752–5 PMID: 6092963.
- [68] Domnguez, O., Ruiz, J. F., deLera, T. L., Garca-Daz, M., Gonzlez, M. A., Kirchhoff, T., Martnez-A, C., Bernad, A., and Blanco, L. April 2000 *The EMBO Journal* **19(7)**, 1731–42 PMID: 10747040.
- [69] Capp, J.-P., Boudsocq, F., Bertrand, P., Laroche-Clary, A., Pourquier, P., Lopez, B. S., Cazaux, C., Hoffmann, J.-S., and Canitrot, Y. (2006) *Nucleic Acids Research* **34(10)**, 2998–3007 PMID: 16738138.