Biochemical, Structural, and Physiological Characterization of Terminal Deoxynucleotidyl Transferase

Jason D. Fowler and Zucai Suo*

Department of Biochemistry, The Ohio State University, Columbus, Ohio 43210

Received September 20, 2005

Contents

* To whom correspondence should be addressed. Mailing address: Department of Biochemistry, The Ohio State University, 740 Biological Sciences Building, 484 West 12th Ave., Columbus, OH 43210. Phone: 614-688-3706. Fax: 614-292-6773. E-mail: suo.3@osu.edu.

1. Introduction

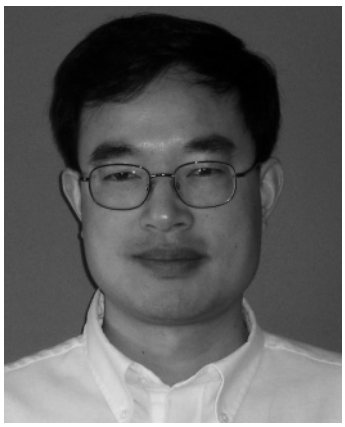
It is of paramount importance to biological organisms that their genetic information be preserved in an intact, replicable state to maintain and perpetuate their existence. However, perfect conservation of the genome is neither possible nor desirable, because those infrequent and tiny changes in the molecules of life provide the basis for evolution and adaptation to an ever-changing and frequently hostile environment. Fortunately, there has come to exist an ensemble of machinery to allow for both the faithful maintenance and the subtle, random change that has laid the foundation for life itself. At the very core of this machine are the DNA polymerases, the caretakers of the genome. These polymerases are responsible for DNA replication and recombination, repair of DNA lesions, and even tolerance of potentially lethal DNA damage through unique mechanisms of lesion bypass. Most polymerases are highly accurate when performing the tasks of genomic replication and repair. However, in those circumstances when "making a mistake is the only way to get ahead",[1] a lesser known group of low-fidelity polymerases can be brought to bear. Members of this group have greatly enhanced flexibility with respect to what substrates they can utilize. This reduced degree of fidelity possessed by these enzymes is what allows them to replicate patches of DNA that are severely damaged or even completely noninformative. However, DNA maintenance is not the only task these unusual polymerases perform. Terminal deoxynucleotidyl transferase (TdT), an enzyme that is unique even in the context of low-fidelity polymerases, is responsible for the generation of the random genetic information that is essential for the efficacious function of the vertebrate adaptive immune system. What follows is a review of the most current knowledge regarding TdT including the enzymatic activity, structural features, and physiological functions that make it one of the most intriguing DNA polymerases in the biological world.

2. Classification of DNA Polymerases

To best appreciate the role and significance of TdT, it is first necessary to understand the classification of all DNA polymerases and where TdT stands in relation to its peers.

10.1021/cr040445w CCC: $59.00
Published on Web 04/29/2006

Jason D. Fowler received his B.S. in Biochemistry from The Ohio State University in 2003. Currently he is pursuing a Ph.D. in biochemistry in the laboratory of Dr. Zucai Suo (Department of Biochemistry) at The Ohio State University and is a member of the Ohio State Biochemistry Program. Jason's research focus is on the characterization of recently discovered DNA polymerases using pre-steady-state kinetic methodologies.



Zucai Suo received both his B.S. and M.S. degrees in physical chemistry from Fudan University, Shanghai, China. Later, he moved to the U.S. and obtained his Ph.D. in chemistry at The Pennsylvania State University in 1997, investigating the effect of RNA secondary structure on the kinetics of HIV-1 reverse transcriptase and learning the art of pre-steady-state kinetics with Dr. Kenneth A. Johnson. As a Jane Coffin Childs Memorial Fund Postdoctoral Fellow, he worked on the biosynthetic mechanism of yersiniabactin catalyzed by yersiniabactin synthetase from *Yersinia pestis* in the laboratory of Dr. Christopher T. Walsh at Harvard Medical School. After working as a senior biochemist at Eli Lilly & Company for about a year, Dr. Suo started his independent research career as an Assistant Professor in the Department of Biochemistry at The Ohio State University in October, 2001. His research interests include kinetic, dynamic, and structure–function relationship studies of DNA lesion bypass catalyzed by low-fidelity DNA polymerases, mechanisms of DNA polymerization catalyzed by X-family DNA polymerases, and antiviral and anticancer drug discovery. Dr. Suo won the National Science Foundation Career Award in 2005.

Since Arthur Kornberg et al. discovered *Escherichia coli* DNA polymerase I in the 1950s,[2,3] many DNA polymerases performing a diverse repertoire of biological functions have been identified. These DNA polymerases have been grouped into six families, A, B, C, D, X, and Y, based on their phylogenetic relationships.[4,5] With the exception of the highly conserved carboxylate residues found within the polymerase active sites (see section 6.1), little sequence similarity is shared between members of different families. Indeed, within each polymerase family, many distinct biological functions can be found. Except for the Y-family, no DNA polymerase family has yet been found that is universally conserved among the three domains of life (Archaea, Bacteria, and Eukaryota). Not surprisingly, evolution of the DNA polymerase families is very complex and is likely to involve multiple gene exchanges between cellular and viral proteins.[6]

2.1. Family A

Family A DNA polymerases can be found in bacteria, metazoa, plants, mitochondria, and viruses.[6] In addition to their template-dependent polymerase activity, members of the A family possess 3′ → 5′ exonuclease activity and possibly 5′ → 3′ exonuclease activity. The representative member of family A is *E. coli* DNA polymerase I, which possesses all three of the activities mentioned above and is involved in DNA repair and recombination.[7] Mitochondrial DNA polymerase γ, another member of family A, is a heterodimeric protein that exhibits both polymerase and 3′ → 5′ exonuclease activities.[8] While primarily functioning in the replication of mitochondrial DNA, mitochondrial DNA polymerase γ also takes part in the repair of mitochondrial DNA through its 3′ → 5′ exonuclease activity.[9] Eukaryotic DNA polymerase θ[1] is also a member of family A and helps to replicate specific templates containing abasic lesions, via its 3′ → 5′ exonuclease functionality.[10] Viral replicative DNA polymerases in family A, such as vaccinia virus DNA polymerase,[11,12] catalyze template-dependent viral genome replication.

2.2. Family B

Family B is mainly composed of the eukaryotic replicative polymerases,[5] which are homologous to *E. coli* DNA polymerase II,[1] the prototype of family B. Family B members can be also found in Archaea, proteobacteria, phages, and viruses.[6] DNA polymerases α, δ, ϵ, and ζ[13,14] are typical family B members. Aside from template-dependent polymerase activity, most family B members, such as DNA polymerases δ and ϵ, also possess 3′ → 5′ exonuclease activity. Although lacking associated 3′ → 5′ exonuclease activity, DNA polymerase α contains both polymerase and primase activities and plays a significant role in eukaryotic replication.[15] DNA polymerase δ possesses 93% conservation from mouse to human[16] and functions in elongation of the leading and lagging strands during DNA replication. DNA polymerase ϵ is involved in DNA repair.[17] Analysis of the N-terminal and C-terminal regions of polymerase ϵ indicate that this enzyme serves as a way of quality control in the cell, while δ exclusively polymerizes extended DNA chains.[5] DNA polymerase ζ (Polζ), a recently discovered enzyme, is likely involved in DNA lesion bypass[13,18,19] and somatic hypermuation (see section 8.4).

2.3. Family C

Family C DNA polymerases are found exclusively in bacteria.[6] Family C is a high-fidelity family, and each member possesses both template-dependent polymerase and 3′ → 5′ exonuclease activities. The prototype of family C is *E. coli* DNA polymerase III, which replicates the genomic DNA of *E. coli*.[1]

2.4. Family D

Family D polymerases are found in the Euryarchaeota subdomain of Archea,[19−21] not in Bacteria or Eukaryota.[6] Each family D DNA polymerase exhibits both template-dependent polymerase activity and 3′ → 5′ exonuclease

activity.[20,22−24] The high-fidelity family D polymerases catalyze DNA replication in Euryarchaea.[21] One of the best known members of family D is from a hyperthermophilic archaeon *Pyrococcus furiosus* (*Pfu*).[21] Two *Pfu* proteins, DP1 and DP2, encoded by tandem genes form a polymerase complex: the former is a small accessory subunit, while the latter is the large catalytic subunit. The two *Pfu* proteins are highly conserved in the Euryarchaeota subdomain. The homologues of DP2 share more than 50% amino acid conservation, while the DP1 homologues possess more than 30% identity. However, family D polymerases generally share little sequence homology to polymerases from other families.

2.5. Family Y

Family Y DNA polymerases have been found in Archaea, Bacteria, and Eukaryota, but not in viruses. In humans, four Y-family members including DNA polymerases eta (Polη), iota (Polι), kappa (Polκ), and REV1 have been identified.[19] Y family members known as translesion polymerases have the ability to bypass DNA lesions, which stop replicative DNA polymerases.[1,5] For example, Polκ has demonstrated an ability to perform translesion synthesis on several aberrant primer−templates including those substrates containing abasic sites, *N*-2-acetyl aminofluorene (AAF) adducts, 8-oxoguanine lesions, and (−)-*trans*-*anti*-benzo(*a*)pyrene-N^2-dG adducts.[25] All the Y-family polymerases that have thus far been biochemically characterized are devoid of intrinsic proof-reading exonuclease activities and catalyze template-dependent DNA synthesis with low fidelity and poor processivity.[19,25−30] The fidelity of polymerization catalyzed by Polι differs with respect to the template base, with an error rate of 10^{-2} to 10^{-4} opposite a template A, G, or C.[31−33] Interestingly, Polι preferentially selects misincorporation of G opposite a template T,[29,31,32] possibly by Hoogsteen base pairing.[34] When a Y-family polymerase encounters a DNA lesion, it can bypass the lesion either in an error-free or in an error-prone manner.[1] For example, human DNA polymerase η has been shown to faithfully replicate through *cis*−*syn* thymine dimers.[19] Mutational inactivation of human Polη leads to a cancer-prone syndrome, which is a variant form of xeroderma pigmentosum (XPV).[19,25,26] In contrast, human Polη bypasses an 8-oxoguanine lesion by incorporating either base A or base C, an abasic site by inserting base A and less frequently base G, a (+)-*trans*-*anti*-benzo[*a*]pyrene-N^2-dG adduct by incorporating base A and less frequently base T, a 1,N^6-Ethenodeoxyadenosine lesion by inserting base T and less frequently base A, and an O^6-methylguanine lesion by incorporating base C or T.[27,29,32]

2.6. Family X

The X family of DNA polymerases is a subdivision of a larger superfamily of nucleotidyl transferases.[6] Members of this family can be found in Achaea, Bacteria, Eukaryota, and viruses. In addition to TdT, DNA polymerase β (Polβ), DNA polymerase λ (Polλ),[35−37] DNA polymerase μ (Polμ),[35,38] African swine fever virus DNA polymerase X (ASFV PolX),[39] yeast DNA polymerase IV (Pol IV),[40] and yeast DNA polymerase σ (Polσ)[41−43] are also members of family X.[44] TdT is known to catalyze nontemplated, random nucleotide addition at the V(D)J junctions thereby increasing antigen receptor diversity (see section 8). In vivo, TdT expression is thought to be restricted to primary lymphoid tissues (thymus and bone marrow),[45−49] although other theories do exist (see section 9). Polβ removes the 5′-deoxyribose phosphate moiety[50,51] and catalyzes gap-filling synthesis[50] during base excision repair (BER). ASFV PolX plays a role in BER analogous to the function of its mammalian counterpart, Polβ.[39] Polσ couples DNA replication to the establishment of sister chromatid cohesion.[41−43]

Nonhomologous end joining (NHEJ), a major pathway for repair of DNA double-strand breaks introduced by exogenous sources including oxidation and ionizing radiation, exists in all cell types. Yeast PolIV functions in NHEJ of double strand breaks[52] and possibly in BER.[53]

So far, the biological role(s) of the recently discovered Polλ have not been established. It is plausible that Polλ contributes to BER since it is related to Polβ and possesses two key enzymatic activities (gap-filling polymerase and a 5′-2-deoxyribose-5-phosphate lyase) required by BER. The gene encoding Polλ is mapped to mouse chromosome 19. Like Polβ,[54] Polλ is expressed at high levels in the developing mouse testes, suggesting a possible function of Polλ in DNA repair pathways, especially BER, associated with meiotic recombination.[37] In an in vitro BER reconstitution reaction, recombinant human Polλ and Polβ can replace each other to efficiently repair uracil-containing DNA in the presence of human uracil−DNA glycosylase, human AP endonuclease, and human DNA ligase I.[55] The role of Polλ in DNA repair is further supported by the following observations: (i) mouse embryonic fibroblast Pol$\beta^{-/-}$ cell extract contains substantial amounts of active Polλ, which can replace Polβ in reconstituted and uracil-initiated short-patch BER, and monoclonal antibodies against Polλ in this cell extract strongly reduce in vitro BER;[56] (ii) Polλ is the only X-family DNA polymerase found in higher plants and its expression is induced by DNA-damaging treatments;[57] (iii) Polλ protects mouse fibroblasts against oxidative DNA damage and is recruited to oxidative DNA damage sites.[58] Thus, Polλ may complement or support the function of Polβ in BER in vivo. On the basis of the current biochemical data, the second proposed biological role of Polλ is to repair double-stranded breaks (DSBs) through NHEJ pathways.[59,60] This hypothesis is supported by the results from immunodepletion studies suggesting that Polλ, rather than other X-family polymerases, is primarily responsible for the gap-filling synthesis associated with NHEJ in human nuclear extracts.[59] The last proposed role of Polλ in vivo is to bypass DNA lesions. This hypothesis is based purly on its ability to bypass an abasic site in the presence of Mn^{2+} in vitro.[4,61] So far, the generation of knock-out mice through deletion of exons 5−7 of the Polλ gene has not yet confirmed the involvement of Polλ in this or any other biological process.[62] These Polλ knock-out experiments are likely complicated by the existence of other DNA polymerases, especially Polβ,[63] that could fill in and compensate for the loss of functions of Polλ.

Similarly, the biological role(s) of another novel X-family member, Polμ, have yet to be identified. Preferential expression in secondary lymphoid tissues as well as the observed low fidelity of Polμ have led to the hypothesis that this enzyme is an error-prone mutase active in somatic hypermutation.[64] The presence of Polμ and the absence of TdT in germinal center B cells, the low levels of Polμ expression in thymus and bone marrow, and the intrinsic terminal transferase activity possessed by Polμ in the presence of Mn^{2+} all suggest that this enzyme may play a role in V(D)J recombination, thereby complementing the biological func-

tions of TdT.[64] Moreover, the basal expression of Polμ in most tissues suggests a potential role in NHEJ for general repair of DNA double-strand breaks. The proposed role of Polμ in NHEJ and V(D)J recombination is substantiated by the following two in vitro observations: Polμ and TdT form essentially identical complexes with the end-joining factors Ku and the XRCC4-ligase IV complex[65] and Polμ promotes microhomology searching and pairing to realign primers with terminal mismatches by looping out any mismatched template nucleotides.[66] Recently, Polμ, like TdT, has been shown to incorporate both rNTPs and dNTPs using either DNA or RNA primers.[67−69] Additionally, Polμ can bypass several DNA lesions through a deletion mechanism.[70,71] The lesion bypass ability of Polμ indirectly supports the proposed role of this polymerase as a mutase during somatic hypermutation.

Other than Polσ, none of the Family X polymerases contain proof-reading exonuclease activity. Recently, recombinant Polσ purified from *E. coli* was shown to display a Mg^{2+}-dependent 3′ → 5′ exonuclease activity in vitro,[43] although it seems more experiments would need to be performed to exclude the possibility that this observed exonuclease activity was due to a contaminating *E. coli* enzyme.

3. Sequence Alignment Analysis of the X Family DNA Polymerases

Sequence alignment and three-dimensional structural modeling studies predict that the C-termini of all family X polymerases possess Polβ-like domains (Figure 1). Each

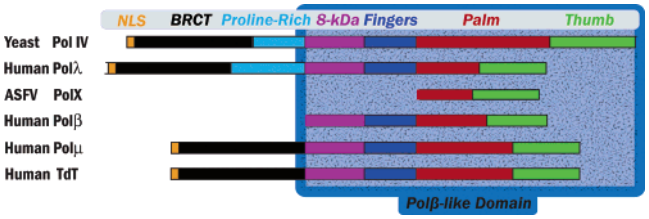


Figure 1. Domain organization of six X-family DNA polymerases. The protein sequence of each DNA polymerase is indicated by a bar, with domains differentiated using different colors.[64] NLS denotes a nuclear localization signal motif. BRCT indicates the BRCA1 carboxy terminus domain.

Polβ-like domain is further divided into the following subdomains: 8-kDa, fingers, palm, and thumb (Figure 2A). Notably, we prefer the subdomain nomenclature of Polβ (Figure 2A), rather than the nomenclature initially proposed for *E. coli* DNA polymerase I,[72] to describe the domain structures of X-family polymerases in this review. The difference between these two nomenclatures is simply a reversal of the names "thumb" and "fingers" for the subdomains on either side of the palm domain. Interestingly, the full-length ASFV Pol X, the smallest known nucleotide polymerase (174 residues, 20 kDa), possesses only the palm and thumb subdomains as revealed by nuclear magnetic resonance spectoscopy (NMR).[73,74] In addition to the C-terminal Polβ-like domain, TdT, Polλ, Polμ, and Yeast PolIV all have nuclear localization signal (NLS) motifs and breast cancer susceptibility protein BRCA1 C-terminal (BRCT) domains on their N-termini. BRCT domains are known to mediate protein/protein and protein/DNA interactions in DNA repair pathways and cell cycle check point regulation upon DNA damage.[75] For example, the BRCT domain of TdT is thought to interact with Ku70/86,[76] a protein

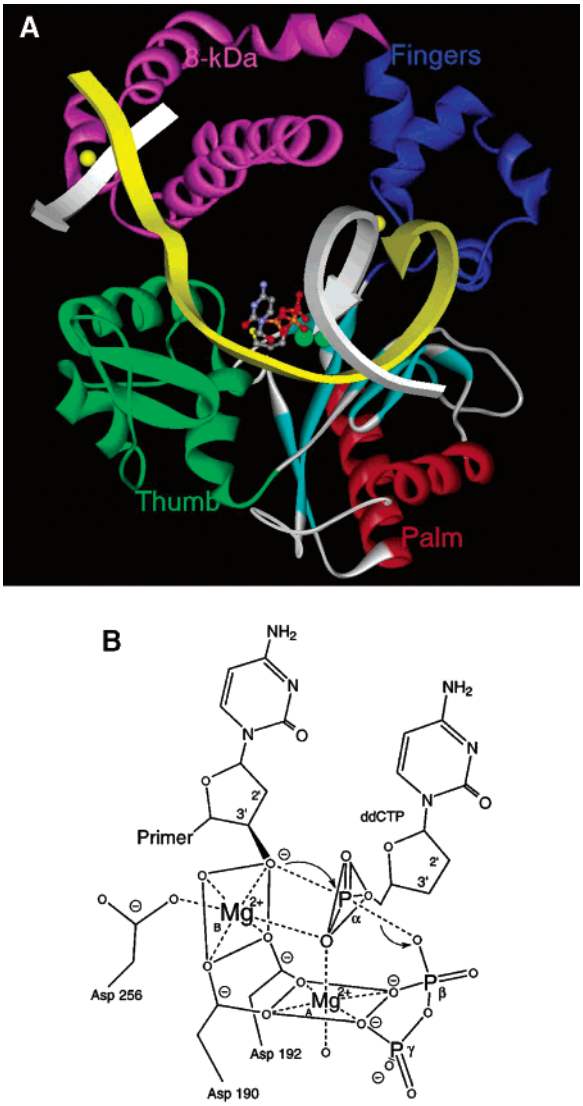


Figure 2. (A) Ternary structure of human DNA polymerase β·single nucleotide gapped DNA·ddCTP.[121] The 8-kDa (purple), fingers (blue), palm (mixed colors), and thumb (green) domains are shown in solid ribbon. The template 16-mer (yellow), upstream primer 10-mer (white), and downstream primer 5-mer (white) are depicted by arrows. The incoming ddCTP (mixed colors, ball-and-stick model), two Mg^{2+} ions (green CPK sphere), and two Na^{+} ions (yellow CPK sphere) are also shown. (B) Proposed "two-divalent-metal-ion" mechanism for nucleotide incorporation catalyzed by human DNA polymerase β.[121]

heterodimer involved in recognizing and binding free DNA ends during V(D)J recombination and double strand break repair.[77] In Polλ, a proline-rich domain can be found located between the BRCT and Polβ-like domains (Figure 1). Analysis of deletion mutants has suggested that the proline-rich domain may functionally suppress the polymerase activity of Polλ while the BRCT domain does not affect polymerase activity.[78] In contrast, the Polβ-like domain of Polμ is not active as a DNA polymerase in the absence of the BRCT domain.[69]

Sequence alignment analysis indicates that Polλ is most similar to Polβ, sharing 32% amino acid identity.[79] The C-terminal Polβ-like domain of Polλ is predicted to fold in a manner similar to Polβ. Analysis of the NMR structure of the 8-kDa domain of Polλ reveals a high degree of similarity to the corresponding domain in Polβ.[80] The X-ray crystal structure of the Polβ-like domain complexed with single-

nucleotide gapped DNA and an incoming nucleotide ddTTP[81] is also similar in many respects to the ternary structure of Polβ shown in Figure 2A.

Among all of the X family polymerases, Polμ has been found to be the closest relative of TdT,[1,5,70,82] sharing approximately 42% amino acid identity (Figure 1).[5,38] These two polymerases are predicted to possess an organization of domains as shown in Figure 1. Unfortunately, the crystal structure of Polμ has not yet been solved, and thus, its predicted domain organization cannot be confirmed.

4. Isoforms of Terminal Deoxynucleotidyl Transferase

TdT itself is highly conserved across the vertebrate phyla, from cartilaginous fish to birds and to humans.[83−89] For example, the TdT sequences of skate and shark share 70% identity at the amino acid level and over 50% nucleotide identity with mouse TdT.[89] So far, two mRNA splice variants have been reported in mice and three each in bovines and humans. The murine mRNA splice variants are translated into mature TdT isoforms designated TdT short isoform (TdTS) and TdT long isoform (TdTL). Murine TdTL (529 residues) differs from murine TdTS in that it contains an additional 20 amino acid residue insertion between the two β-sheets in the TdTS thumb subdomain (509 residues).[90] This addition is the result of the expression of an additional exon in the murine TdT gene.[91] Intriguingly, the effect of this insertion on the polymerase activity of murine TdT is somewhat controversial. Papanicolaou et al. have shown that this insertion decreases the thermostability of TdTS but does not affect its catalytic activity.[92] In contrast, murine TdTL is found by Kearney et al. to possess 3′ → 5′ exonuclease activity, rather than the template-independent polymerase activity associated with TdTS.[93] Although both murine TdTS and TdTL localize to the nucleus, it is believed that the long isoform may down regulate the polymerase activity of the short isoform in vivo.[94] However, this hypothesis has yet to be confirmed by experimentation. In transgenic mice deficient in TdT, the short isoform is sufficient to rescue N-addition activity[94] (see section 8.2) suggesting that the template-independent polymerase activity of TdTS, not TdTL, is required in vivo.

In vivo, the three mRNA splice transcripts in cattle[95] and humans[96] are also translated into three mature protein isoforms designated TdTS, TdTL1, and TdTL2.[97] In humans, the normal B and T lymphocytes express exclusively hTdTS and hTdTL2, whereas hTdTL1 expression appears to be restricted to transformed lymphoid cell lines. In in vitro recombination and primer extension/digestion assays, both human TdTL isoforms are shown to possess 3′ → 5′ exonuclease activity while human TdTS acts as a template-independent DNA polymerase.[96] Overexpression of hTdTS or hTdTL2 greatly reduced the efficiency of recombination, which was reverted to normal levels by the simultaneous expression of both enzymes. These data suggest that alternative mRNA splicing may prevent the adverse effects of unchecked elongation or diminution of coding ends during V(D)J recombination, thus affecting the survival of a B or T cell precursor during receptor gene rearrangements.[96]

5. Terminal Deoxynucleotidyl Transferase Expression and Purification

Chang and Bollum were among the first to attempt to purify TdT from calf thymus cell lysate.[98] Due to proteolysis, they incorrectly suggested that TdT was a heterodimer of α and β subunits. It was later discovered that purified full-length TdT is actually a single polypeptide with a molecular weight of approximately 60 kDa.[99] Because efforts to purify TdT from calf thymus were hindered by proteolysis, researchers started to investigate alternate means of obtaining homogeneous TdT.[100,101] Unfortunately, although several research groups have attempted to express recombinant human TdT in bacteria, none have succeeded in obtaining soluble and active protein.[102,103] It was in 1988 that TdT was first successfully expressed and purified using the baculovirus expression system.[103] A decade later, active murine TdTS and TdTL were successfully expressed in *E. coli* by lowering the bacterial growth temperature to 15 °C and overexpressing a rare arginyl tRNA. Those two isoforms obtained in this manner were successfully purified to apparent homogeneity through column chromatography.[92,104] This *E. coli* method allows for large scale production of those full-length murine TdTS and TdTL for enzymatic and structure−function relationship analysis.[92,104] Thanks to a high degree of sequence homology, this *E. coli* method may be applicable in the production of TdT's from other species as well.

6. Three-Dimensional Structures of DNA Polymerases

6.1. Crystal Structures of Template-Dependent DNA Polymerases

All template-dependent polymerases with known structures (both those crystallized and those in solution)[34,72,73,88,105−120] share a similar architecture at their polymerase catalytic domain. Intriguingly, these structures resemble the right-hand of a human being, with domains that resemble the palm, fingers, and thumb (and were so named). Domain nomenclature based on this observation was first proposed for *E. coli* DNA polymerase I[72] and has since been adopted for other polymerases as well. As an example, these domains can be seen in Figure 2A, which showcases the crystal structure of human DNA polymerase β complexed with a single-nucleotide gapped DNA substrate and an incoming nucleotide ddCTP.[121] In the polymerase active site, three aspartic acid residues, one water molecule, the 3′-OH of the upstream primer strand, and the triphosphate moiety of ddCTP act as ligands to bind two Mg^{2+} ions (Figure 2B). During polymerization, the first Mg^{2+} ion (B) promotes the deprotonation of the 3′-OH of the primer strand, facilitating the 3′ oxyanion's nucleophillic attack on the α-phosphate of the incoming nucleotide. The second Mg^{2+} ion (A) then stabilizes the pentacovalent transition state of the α-phosphate and assists the leaving of the pyrophosphate (Figure 2B).[121] In fact, two metal ions have been found in the active sites of all DNA polymerases with known crystal structures.[34,72,73,88,105−120] Based on this observation, T. A. Steitz proposed that perhaps all DNA polymerases might use a "two-divalent-metal-ion" mechanism to catalyze nucleotide incorporation.[122] These essential metal ions, which are likely to be Mg^{2+} in vivo, are bound by three carboxylates (aspartate or glutamate). Notably, these carboxylates are conserved across each of the six DNA polymerase families.[123]

6.2. Crystal Structure of the Polβ-like Domain of Murine Terminal Deoxynucleotidyl Transferase

The crystal structures of (i) the Polβ-like domain of murine TdTS (residues 130−510, resolution 2.35 Å), (ii) the binary

complex of the Polβ-like domain and a brominated DNA primer 9-mer (3 Å, Figure 3), and (iii) the binary complex of the Polβ-like domain and an incoming nucleotide ddATP−Co^{2+} (3 Å, Figure 4) have been solved.[88] The three-dimensional structure of the Polβ-like domain resembles a torus (Figure 3), with an α-helical N-terminal 8-kDa subdomain (residues 163−243), an α-helical fingers subdomain (residues 243−302), a central palm subdomain with a large antiparallel β-sheet (residues 302−450), and a C-terminal thumb subdomain (residues 450−510) containing a small antiparallel β-sheet. The thumb subdomain makes extensive contact with the 8-kDa domain to close the protein ring. Curiously, despite a low shared sequence identity of only 22−24%, the four subdomains in TdT are structurally homologous to the corresponding subdomains in Polβ (Figure 2A). In comparison to the protein sequence of Polβ, TdT has two insertions of 10−15 residues between β3 and β4 and between β4 and β5, which form loop 1 and loop 2 in the palm subdomain, respectively. The antiparallel β-sheets in the palm subdomain contain the three aspartate residues (Asp343, Asp345, and Asp434, Figure 4), which are highly conserved within the nucleotidyl transferase family. These residues have been demonstrated by site-directed mutagenesis experimentation to be essential for the binding of two divalent metal ions and for the catalytic activity of TdT.[102,124]

Figure 3. Binary crystal structure of the Polβ-like domain (residues 148−510) of murine TdT complexed with a brominated 9-mer at 3.0 Å.[88] The 8-kDa (purple), fingers (blue), palm (mixed colors), and thumb (green) subdomains are shown in solid ribbon. Three β-sheets, β3, β4, and β5, are labeled as 3, 4, and 5, respectively. Loop 1 and loop 2 are shown in yellow. Four 3′-nucleotides of the 9-mer primer are ordered and drawn in the ball-and-stick model. Mg^{2+} (green) and Na^{+} (purple) ions are shown as CPK spheres.

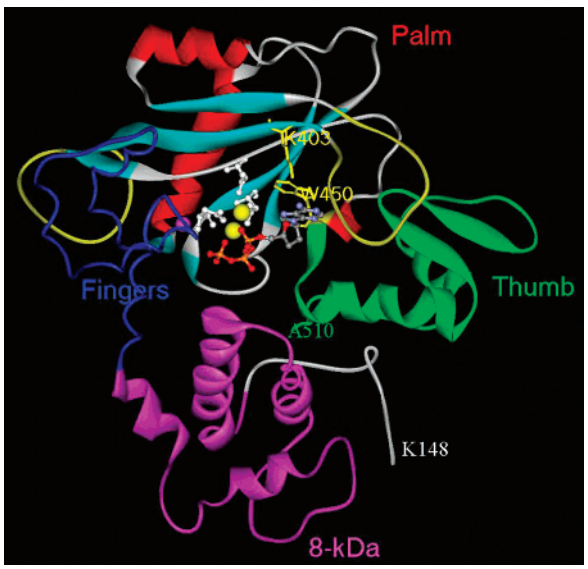


Figure 4. Binary crystal structure of the Polβ-like domain (residues 148−510) of murine TdT complexed with a ddATP−Co^{2+} at 3.0 Å.[88] All subdomains are depicted in the same manner as those in Figure 3. Three active site residues, Asp343, Asp 345, and Asp434 (white color, stick and ball model), as well as Lys403 (yellow color, stick model), Trp450 (yellow color, stick model), Co^{2+} (yellow CPK sphere), Na^{+} (purple CPK sphere), and an incoming ddATP (mixed colors, stick and ball model), are also shown.

Upon analysis of the crystal structure in Figure 3, the primer strand is observed to lie on the palm subdomain, perpendicular to the axis of the protein ring. Notably, only four nucleotides from the 3′ end of the 9-mer primer are ordered, suggesting tight association of these four nucleotides with the polymerase. It should be noted that these residues are in the B-type DNA conformation. The disordered five nucleotides from the 5′ terminus of the 9-mer are not in contact with the protein. Interestingly, the 3′-terminal nucleotide is located at the position of an incoming nucleotide found in the ternary complex of Polβ shown in Figure 2A.[121] Thus, the structure in Figure 3 is considered by Delarue et al. to mimic the ternary structure of TdT, DNA, and a nucleotide.[88] Alternatively, the binary structure in Figure 3 may represent the complex of TdT and a DNA product after TdT has incorporated the incoming nucleotide but before TdT has repositioned itself for the binding of the next incoming nucleotide. Nevertheless, the position of the primer in the active site of TdT shown in Figure 3 must change in the presence of an incoming nucleotide. Interestingly, the small number of ordered nucleotides in Figure 3 is consistent with the finding that TdT requires at least a 5′-phosphorylated trimer as a primer to act as an efficient polymerase.[125] However, as discussed below, in the presence of Mn^{2+}, TdT has been observed to catalyze DNA synthesis de novo.[82] In addition, no atoms can be identified between the amino acid residues of TdT and the primer nucleotides that reside in sufficient proximity to one another to be considered a polar interaction, thus indicating that binding must rely entirely upon interaction with the sugar−phosphate backbone. This observation could further explain the in vitro results indicating that TdT displays a low degree of specificity with respect to nucleotide selection.[88] TdT does, however, strongly prefer single-stranded DNA. Upon examination of the location of loop 1, it is likely that the presence of this lariat-shaped loop might preclude the accommodation of a template strand, thus making TdT an inefficient DNA polymerase in the presence of double-stranded DNA.[88]

Like several DNA binding proteins including Polβ and Polλ,[126] TdT possesses two DNA-binding helix−hairpin−helix (HhH) motifs (residues 208−231 and 244−267). The second HhH motif (residues 244−267), which interacts with the primer strand, uses the carbonyl groups of residues Thr253, Val255, and Val258 as ligands to chelate a Na^{+} ion (Figure 3). Similarly, these HhH motifs in Polβ[105,121] and Polλ[53,127] are found to coordinate K^{+} or Na^{+} ions and participate in sequence-independent interactions with the backbones of the template and primer strands. The HhH

motifs of Polβ are further shown to have a preference for cations in the order $K^+ > Na^+ > Mg^{2+} > Ca^{2+}$.[105]

The binary complex of an efficient template-dependent DNA polymerase and DNA usually binds to a correct incoming nucleotide tightly with a dissociation equilibrium constant (K_d) in the low micromolar range. This high ground-state binding affinity is partly achieved through base pairing interactions between the incoming nucleotide and the opposite template base. Since there is no template with which to base pair and anchor an incoming nucleotide in the active site of TdT, nucleotide binding must result exclusively from interaction with the polymerase active site. While it is true that the actual nucleotide binding site can only be revealed by observing the ternary structure of TdT, a primer, and a nucleotide (not yet reported), the nucleotide binding site of TdT can be estimated through analysis of the binary structure of the Polβ-like domain and ddATP−Co^{2+} (Figure 4).[88] In this structure, it can be seen that the aromatic ring of Trp450 is parallel to and partially stacked with the adenine ring of ddATP, with the CZ2 atom of Trp450 located 3.6 Å from the C8 atom of the adenine ring. The side chain of Lys403 is observed to point toward the adenine ring of the ddATP, with the ϵ amino group located 4 Å above the base. Additionally, one will notice that the anionic triphosphate moiety of ddATP is neutralized and stabilized by three positively charged residues (Arg336, Lys338, and Arg454). The sugar ring of the incoming ddATP is observed to bind Trp450 on one side and reside in close proximity to the *cis*-peptide bond between Gly452 and Ser453 on the other side. Notably, none of the amino acid residue side chains seem to reside in sufficiently close proximity to the 2′ or 3′ position of the ddATP ribose ring to sterically prohibit the accommodation of a ribonucleotide in the proposed active site. Therefore, one might hypothesize that TdT should be able to incorporate both deoxyribonucleotides and ribonucleotides with similar efficiency. This prediction is substantiated by the observed sugar selectivity values of TdT, which lie in the range of 2−5[124] or 2.6−8.9.[67] However, DNA polymerases such as *E. coli* DNA polymerase I[128] and T7 DNA polymerase[129] do not specifically bind a nucleotide in the free state, prior to the binding of DNA. This arouses the suspicion that the observed interactions between the active site residues of TdT and ddATP in Figure 4 might be altered in the presence of a single-stranded DNA primer. Furthermore, the N-terminal BRCT domain of TdT, which in this structure is absent, may also affect the nucleotide binding site through domain−domain interactions.

7. Enzymatic Activities of Terminal Deoxynucleotidyl Transferase

7.1. Template-Independent Polymerase Activity

Most DNA polymerases require a DNA template during replication of genomic DNA, while repairing DNA damage or bypassing DNA lesions. However, exceptions can be found in certain members of family X. Within this family, DNA polymerases possess either template-dependent or template-independent activity. Polβ, for example, catalyzes template-dependent gap-filling DNA synthesis during the process of base excision repair.[51] However, TdT will add random nucleotides to single-stranded DNA in a completely template-independent manner. In fact, TdT actually prefers single-stranded DNA over double-stranded DNA (recessed and blunt) and completely lacks the ability to copy a template.[130]

7.2. 5′-Deoxyribose-5-phosphate Lyase Deficiency in Terminal Deoxynucleotidyl Transferase

Cellular DNA is subject to a continuous assault by exogenous and endogenous DNA damaging agents. Under these conditions, DNA will accumulate a number of harmful and potentially lethal lesions. One of the major mechanisms by which these aberrations are corrected is the base excision repair pathway. A major player in this process, Polβ, catalyzes the following crucial enzymatic steps: it removes the 5′-deoxyribose phosphate moiety via its 5′-deoxyribose-5-phosphate lyase (dRPase) activity[131,132] after it has catalyzed gap-filling synthesis to replace the previously excised nucleobases.[132] The active site residues of the dRPase activity are located in the 8-kDa domain of Polβ. This domain will bind to the downstream primer of a gapped DNA substrate and increase the processivity and efficiency of polymerization.[121] The terminal 5′-phosphate on the downstream primer is buried in a positively charged pocket of Polβ consisting of His34, Lys35, Tyr39, Lys60, and Lys68, while Lys72 acts as the nucleophile for the dRPase reaction.[121,133] Although structurally TdT contains a similar 8-kDa subdomain (Figures 2 and 3), it lacks dRPase activity for the following two reasons: (i) the above residues essential for the lyase activity in Polβ are not conserved in TdT; (ii) the 8-kDa domain is highly basic in Polβ (net charge +10) but is acidic in murine TdT (net charge −6) and all other known TdT's (net charge −4 to −6).[88] Logically, a basic 8-kDa domain will bind more tightly to single-stranded DNA than an acidic 8-kDa domain simply due to charge−charge interactions. Lack of interaction between the 8-kDa and the 9-mer oligo in Figure 3 confirms this prediction.

7.3. Primer Requirement

To catalyze template-independent polymerase activity, TdT requires a primer at least as large as a trinucleotide, a free 3′-OH moiety for extension of that primer, and a free primer 5′-phosphate.[125] For example, pdApdApdA is an active primer but not dApdApdA. The minimum size of the primer is similar to the number of ordered nucleotides in Figure 3 (see section 6). Interestingly, when a primer contains a 3′-terminal β-L-nucleotide, TdT can only add one or two dNTPs to the primer.[134]

7.4. Metal Ion Dependence

As mentioned previously, it was T. A. Steitz who first proposed that perhaps all polymerases might use a "two-divalent-metal-ion" mechanism to catalyze nucleotide incorporation.[122] These essential metal ions, which are likely to be Mg^{2+} in vivo, are bound to DNA polymerases via three conserved carboxylates (aspartate or glutamate) found in the palm domain of the enzyme. TdT is no exception. It too requires the presence of divalent metal ions as cofactors. However, there is more than one species of ion including Mg^{2+}, Mn^{2+}, Zn^{2+}, and Co^{2+} that can be incorporated into the TdT active site in vitro. The binary structure of TdT complexed with ddATP−Co^{2+} (Figure 4) reveals that two bound Co^{2+} ions are next to ddATP at the active site,[88]

Table 1. Effect of Metal Ions on the Incorporation Rate of Each dNTP Catalyzed by TdT[a]

	relative incorporation rate[b]	
dNTP	Mg^{2+}	Co^{2+}
dATP	1.00	1.13
dGTP	1.63	2.30
dTTP	0.10	16.46
dCTP	0.13	14.39

[a] Reference 135. [b] Incorporation rates are relative to the rate of dATP (245 nmol/(mg TdT·h)).

indicating that, TdT too likely utilizes a "two-divalent-metal-ion" mechanism to catalyze nucleotide incorporation.[122] One should be cautioned, however, that biochemical studies have not yet revealed unequivocally the number or identity of the metal ions bound by TdT in vivo. Curiously, the efficiency of polymerization and the bias toward purines and pyrimidines can be significantly affected by the identity of the metal ion that TdT has chelated. For example, in the presence of Mg^{2+}, purine incorporation occurs at about a 10-fold faster rate than pyrimidine incorporation (Table 1);[135] however this substrate specificity is opposite in the presence of Co^{2+} with pyrimidines favored over purines by 10-fold (Table 1). In addition, micromolar amounts of Zn^{2+} added to the reaction along with Mg^{2+} increase the efficiency of polymerization of all nucleotides.[136]

7.5. Kinetic Mechanism of Template-Independent Polymerization

As yet, the kinetic mechanism for template-independent polymerization catalyzed by TdT has not been reported. A handful of template-dependent DNA polymerases that share the same minimal kinetic mechanism shown in Figure 5, have

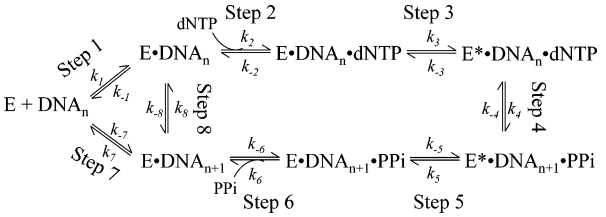


Figure 5. A minimal kinetic mechanism for polymerization catalyzed by DNA polymerases. E and E* denote a polymerase before and after conformational change, respectively. PPi represents pyrophosphate.

been kinetically characterized.[137] In this mechanism, DNA first binds to a polymerase to form E·DNA$_n$ (step 1). This binary complex then binds an incoming nucleotide dNTP to form a ground-state ternary complex E·DNA$_n$·dNTP (step 2) in which the polymerase is in an open conformation. In the following step, nucleotide binding energy is used to induce a change in protein conformation to form a tightly bound ternary complex E*·DNA$_n$·dNTP in which the polymerase is in a closed conformation (step 3). This open to closed conformational change is then followed by the incorporation of the dNTP into the growing DNA polymer and formation of pyrophosphate (step 4). After the second protein conformational change (step 5), pyrophosphate is released (step 6). At this point, if the polymerase is not processive, it will dissociate from the DNA product (step 7). However, if the polymerase is processive, it will translocate to the next template base (step 8) and start the incorporation cycle again. For DNA polymerases including the Klenow fragment of *E. coli* DNA polymerase I,[138] T7 phage DNA polymerase,[139] HIV-1 reverse transcriptase,[140,141] human mitochondrial DNA polymerase,[142] yeast DNA polymerase η,[143] and *Sulfolobus solfataricus* DNA polymerase IV,[30] a noncovalent step preceding phosphodiester bond formation (step 3) limits correct nucleotide incorporation, while phosphodiester bond formation itself (step 4) is rate-limiting for incorrect nucleotide incorporation. Thus, step 3 is considered as a critical fidelity check point.[137] Previous X-ray structural studies have suggested that step 3 may involve the swing and closure of the fingers domain of these polymerases once a correct nucleotide is bound.[107,144,145] However, recent stopped-flow fluorescence studies suggest that local structural rearrangement at the active site, rather than the fingers domain closing, may limit single nucleotide incorporation.[137,146,147] In contrast, the phosphodiester bond formation (step 4) is found to kinetically limit both correct and incorrect nucleotide incorporation catalyzed by Polβ.[121,148] Although, a significant protein conformational change in step 3 is observed when the binary and ternary structures of Polβ are compared.[121] In the case of TdT, our preliminary kinetic studies have indicated that it likely follows the same kinetic mechanism shown in Figure 5 to incorporate a single nucleotide, but with different microscopic rate constants (J. Fowler and Z. Suo, unpublished results). Whether the protein conformational change step (step 3) or the chemistry step (step 4) limits single nucleotide incorporation catalyzed by TdT is not clear at present. The binary structure of TdT shown in Figure 3 superimposes best with the ternary structure of Polβ*·DNA·dNTP (the closed conformation),[88] suggesting that TdT may not undergo significant conformational change in step 3. This seems to be reasonable because TdT, unlike template-dependent DNA polymerases, does not have to consider fidelity. Thus, a slow conformational change step (step 3), which is utilized by replicative DNA polymerases to perform a fidelity check, is not necessary. If this structural prediction is correct, step 4 rather than step 3 likely limits nucleotide incorporation catalyzed by TdT. However, if step 3 is limiting, local structural rearrangements at the active site of TdT must be involved.[137] Finally, because TdT catalyzes DNA synthesis in a strictly distributive mode,[124] TdT most likely dissociates from a DNA product (step 7), rather than progressing through step 8, to complete the incorporation cycle (Figure 5). This prediction is reasonable because there is no template involved. We suspect that the primer strand alone does not form a stable DNA α-helix as double-stranded DNA, leading to difficult translocation of TdT along the DNA after TdT incorporates a nucleotide.

7.6. Nucleotide Selectivity

The physiological role of TdT is to catalyze the addition of random dNTPs onto the 3′ hydroxyl terminus of single-stranded DNA, although in vitro it has also been shown that TdT can accept ribonucleotide triphosphates (rNTPs). Generally, sugar selectivity for DNA polymerases is very high. For example, Polβ has a sugar selectivity in the range of 2000−6000.[67] In comparison, TdT has only a vanishingly small preference (2−9-fold) for dNTPs over rNTPs.[124,125,149,150] Strikingly, TdT can incorporate a wide variety of nucleotide analogues such as *p*-nitrophenylethyl triphosphate,[151] *p*-nitrophenyl triphosphate,[134] d4TTP,[134] cordycepin 5′-triphosphate,[152] 2′,3′-dideoxynucleotides (ddNTPs),[153] α-D-dNTPs,[154] and dinucleoside 5′,5′-tetraphosphates (Figure 6).[155] TdT can also catalyze the transfer of phosphate ester groups and

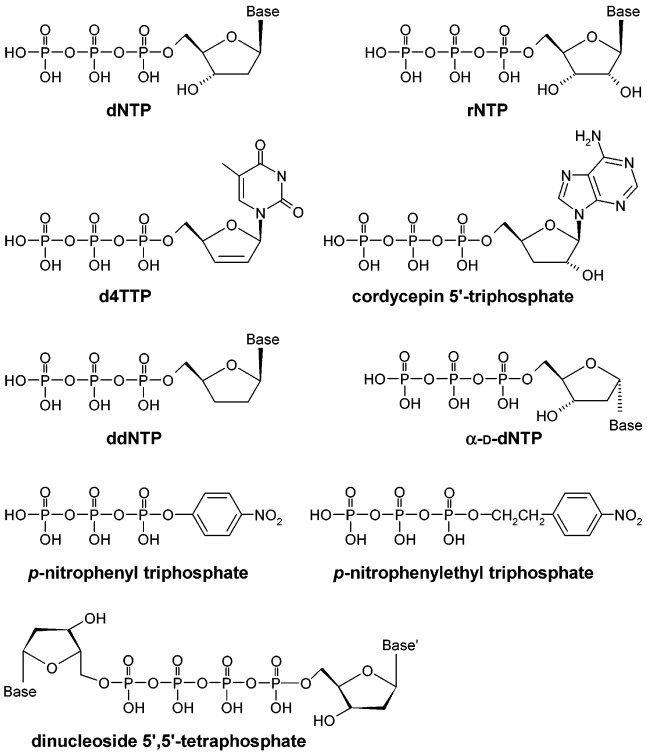


Figure 6. Chemical structures of nucleotide analogues.

phosphonate residues from their corresponding triphosphate derivatives onto the primer 3′-terminus (Figure 6).[155] The broad nucleotide substrate specificity of TdT suggests that the specific interaction between an incoming nucleotide and the TdT active site most likely occurs at the triphosphate moiety of the nucleotide, whereas the role of the base and sugar may be of lesser importance. If the incorporated analogue lacks a 3′-OH moiety, it will terminate primer extension, leading to drug-induced apoptosis.[156] For example, cordycepin has been used recently to target TdT-positive leukemia cells[155] because it is known that the level of TdT in the leukocytes of leukemia patients is very high.[157] Moreover, taking advantage of its broad nucleotide substrate specificity, TdT has been used to synthesize copolymers such as $(pdT)_6(dG)_{10}(dA)_{313}$,[158] label primer 3′-termini with highly radioactive nucleotides,[152] and attach biotinylated or fluoresceinated probes to synthetic oligonucleotides.[159]

7.7. Other Enzymatic Activities

In addition to template-independent polymerase activity, TdT can also catalyze pyrophosphorolytic dismutation of oligodeoxy-nucleotides by removing a 3′-nucleotide from one oligonucleotide and adding it to the 3′-end of another.[160] TdT further distinguishes itself from classical DNA polymerases through its ability to catalyze the creation of polynucleotides of 2-mer, 7-mer, 15-mer, and 21-mer de novo in the presence of Mn^{2+}, given only dNTPs. This ability has only been observed in two other polymerases to date, Polλ and Polμ.[82] It has been speculated that these DNA fragments might act as a recognition signal for DNA repair or recombination machinery.[82] However, the relevant metal ions for DNA polymerases in vivo are likely Mg^{2+}, rather than carcinogenic Mn^{2+};[161] therefore these observations may not be physiologically relevant. Interestingly, as mentioned earlier, it has recently been reported by Kearney et al. that the long isoforms of murine and human TdT's possess 3′ → 5′ exonuclease activities, rather than template-indepadant polymerase activities.[93,96]

8. Immune System and Terminal Deoxynucleotidyl Transferase

The main goal of the vertebrate adaptive immune system is to defend the organism from harmful foreign agents or "antigens". This objective is accomplished primarily through the recognition of antigens by antigen-binding proteins that the immune system produces. These proteins are divided into two major classes: the immunoglobulins (Ig's, Figure 7),

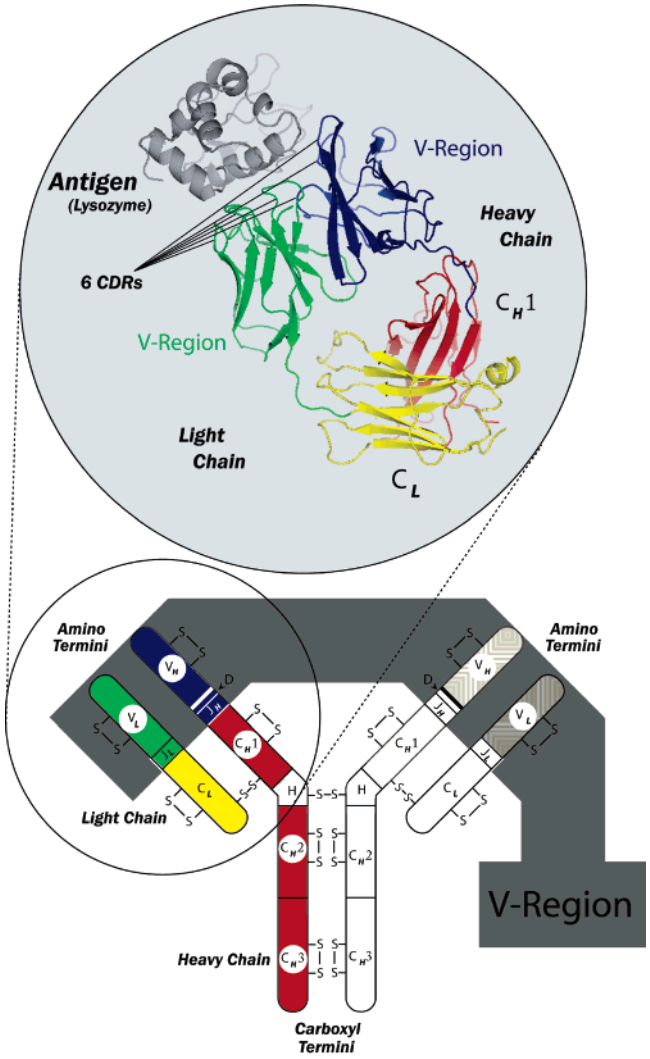


Figure 7. Crystal structure of the complex of an anti-lysozyme F_{ab} and the antigen, hen egg white lysozyme (top),[245] and the domain structure of an Ig molecule (bottom). The variable region in the heavy chain is composed of each of V, D, and J gene segments, while the variable region in the light chain possesses a V and a J gene segment. The intra- and interchain disulfide bonds are denoted as −S−S−.

which are either free glycoproteins present in the serum and tissue fluids (antibodies) or attached to the membrane of certain cells such as memory B cells, and the T cell antigen receptors (TCRs, Figure 8), which are very similar glycoproteins present on the surface of T cells.[162] In general, these proteins are remarkably specific for particular antigens that they recognize. Therefore, the number of these antigen-binding proteins must be tremendously large so that the immune system can respond to the maximum number of

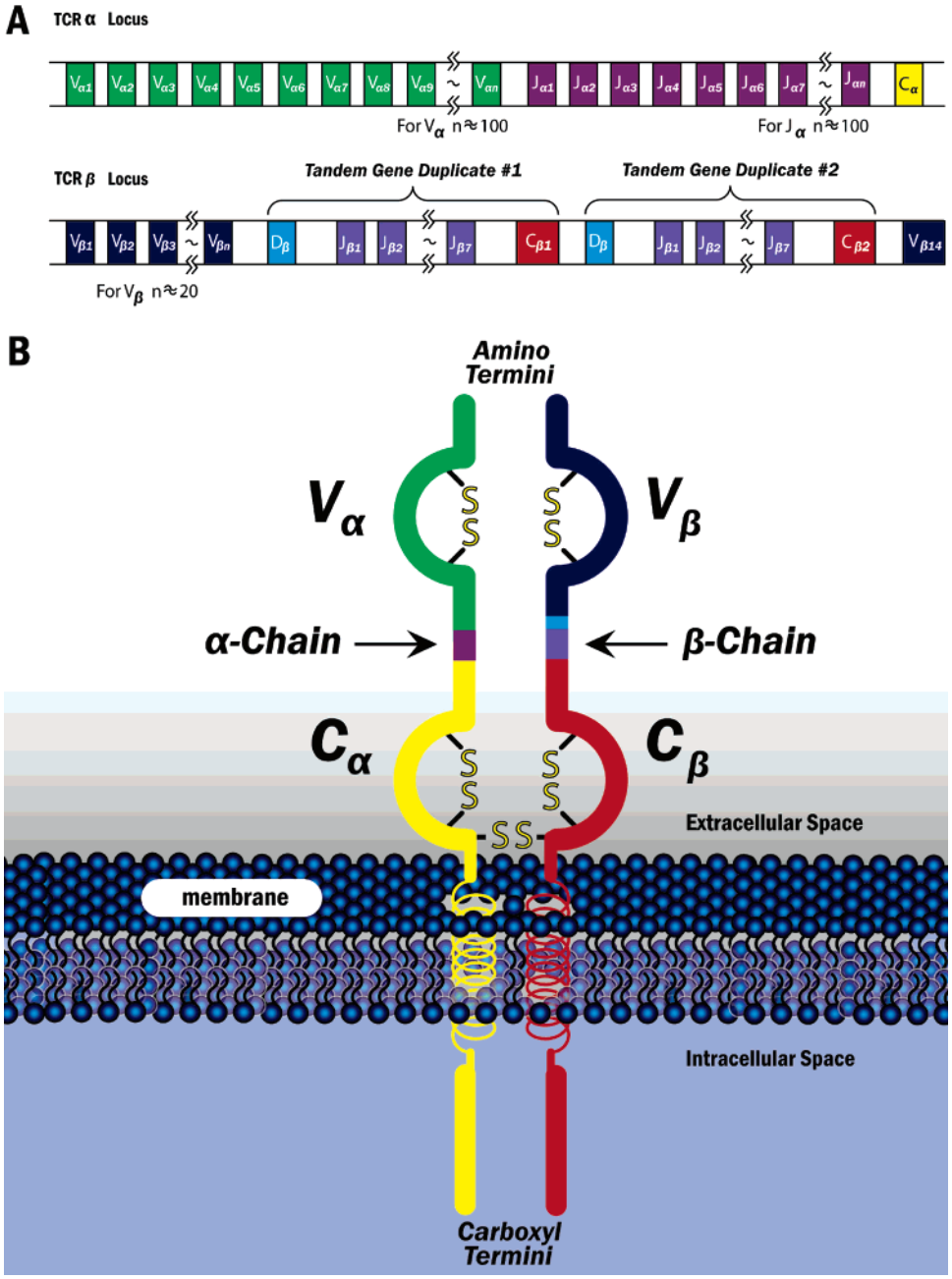


Figure 8. (A) The T cell receptor encoded by tandemly arranged clusters of V, D, and J gene segments. The constant region (C) gene segments follow the joining (J) gene segments. The TCR α chains do not possess the dependent (D) gene segments. (B) Schematic diagram of a T cell receptor. Each TCR chain is composed of a variable and a constant region.

antigens. In fact, some have estimated that in humans, there are approximately 10^{14} unique Ig's and around 10^{18} unique TCRs.[163,164] Given that there are somewhere between 30 000 and 150 000 genes in the entire human genome,[165] it would seem impossible that each of these antigen-binding proteins could be encoded by a unique gene. Even when one takes into account alternative gene processing pathways, the number of different gene products derived from a single gene cannot exceed a small number.[164] Therefore, to mount a maximally effective defense, the cells of the immune system have developed an interesting method to "shuffle" and "randomize" small elements of the germline immunoglobulin and T cell receptor genes. This process makes it possible for a relatively small number of genes to establish an astonishingly large body of unique proteins (see section 8.1).

A basic Ig protein consists of two large polypeptides and two small polypeptides (Figure 7). The two larger "heavy" chains are joined by several disulfide bonds at what is referred to as the "hinge" region. The two smaller "light" chains are each joined to one of the heavy chains by a disulfide bond. Each chain, heavy and light, has a conserved "constant" region and a "variable" region. The constant region of a heavy chain consists of equal thirds (C_H1, C_H2, and C_H3) that are similar in sequence. The constant region of the light chain (C_L) closely resembles C_H1, C_H2, and C_H3. A T cell antigen receptor, in contrast, consists of a 43-kDa α chain joined to a 43-kDa β chain by a disulfide bond (Figure 8B). Each chain of a T cell antigen receptor also possesses both constant and variable regions.[162] Although the overall structures of Ig's are conserved, the N-terminal variable regions of both the heavy and light chains each contain three hypervariable loops called complementarity determining regions (CDRs), which make specific contact with antigens and are responsible for the high antigen affinity

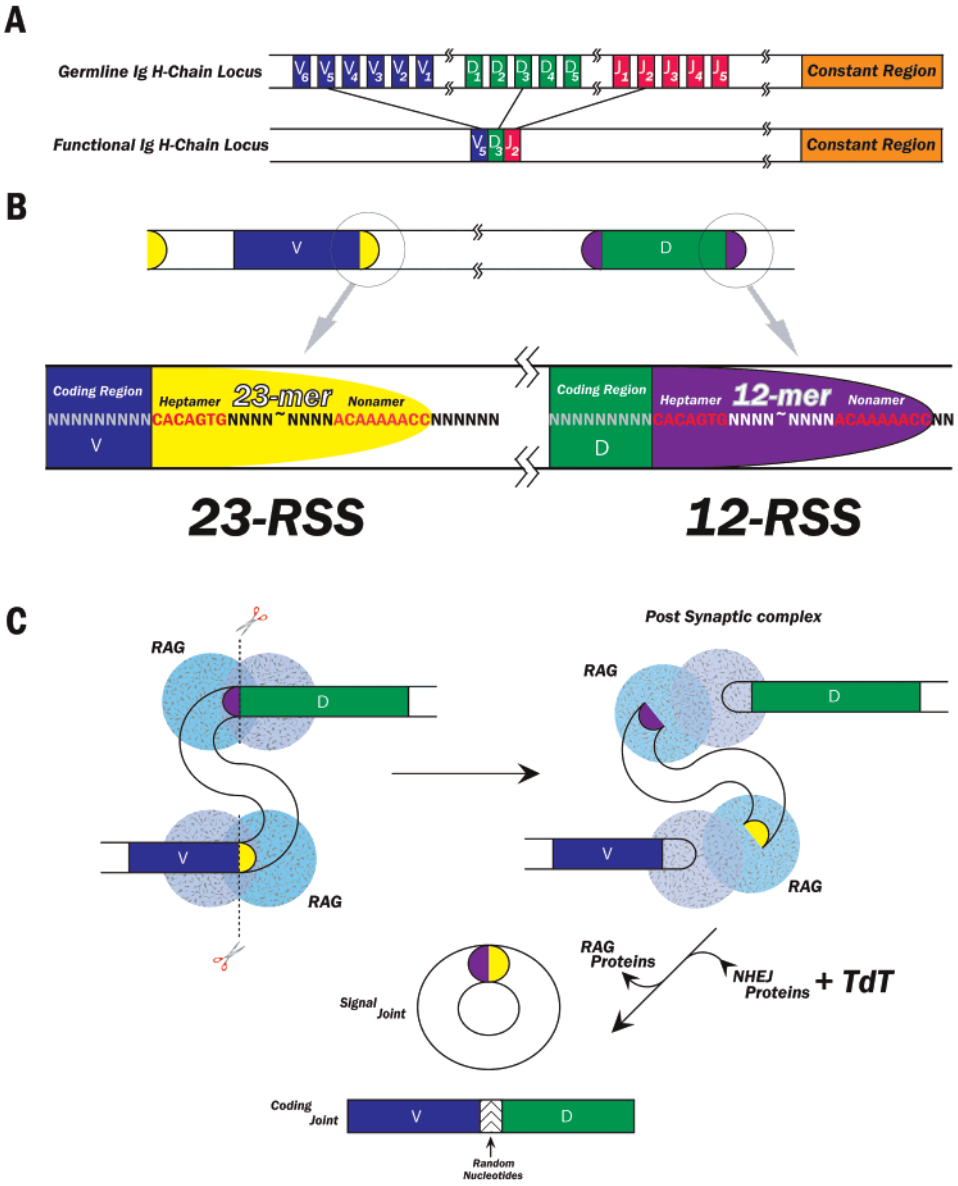


Figure 9. Scheme of V(D)J recombination. (A) A germline Ig heavy chain becomes a functional Ig heavy chain after V(D)J recombination. The colored boxes denote clustered coding segments. One of each of the V, D, and J segments are joined to form the variable region of a functional Ig. Random nucleotides (not shown) are added to the junctions between V, D, and J segments. The constant regions are not involved in V(D)J recombination. (B) Coding segments V and D are associated with recombination signal sequences 23-RSS and 12-RSS, respectively. The 23-RSS and 12-RSS are enlarged for clarity. (C) A V segment is joined by a D segment through cleavage by RAG proteins and processing by NHEJ proteins. TdT adds random nucleotides to the junction between V and D segments.

of Ig's (Figure 7). The binding affinity between an antibody and an antigen (protein, oligosaccharide, etc.) is generally in the picomolar to 0.1 μM range.[166,167] For example, anti-lysozyme antibody D1.3 binds to the antigen, hen egg white lysozyme, with an affinity of 2 nM (Figure 7).[168] It is believed that induced fit mechanistics play a role in the formation of many tight antibody−antigen complexes. Variable regions of Ig heavy chains and TCR β chains are assembled from three gene segments called variable (V), dependent (D), and joining (J) segments. In comparison, each corresponding variable region in the Ig light chain and TCR α chain is assembled from V and J gene segments only. In both mice and humans, germline V, D, and J gene segments are inherited as tandem clusters (Figure 8A and Figure 9A).

Low-affinity Ig's such as IgM circulate in the blood before encountering antigens, while high-affinity Ig's are necessary to disable viruses, toxins, and other foreign microorganisms. In the absence of the four classes of high-affinity Ig's including IgA, IgD, IgE, and IgG, an individual is unable to fight off infection and thus dies at a premature age.[169,170] Therefore, it becomes clear that a diverse pool of antibodies is of crucial importance to the ability of an organism to defend itself against the wide variety of antigens with which it may be presented. Currently, three mechanisms including V(D)J recombination, somatic hypermutation, and "class" or "isotope" switching are known to diversify Ig's in vivo. In comparison, only V(D)J recombination is involved in the diversification of TCRs in vivo. Here, we address only the first two mechanisms because of their relevance to TdT and other DNA polymerases. Class or isotope switching (not reviewed here) is a region-targeted recombination pathway to translocate a VDJ gene from a site near one "constant" gene segment (C gene) to a site near another C gene. Many recent reviews covering the class or isotope switching mechanism can be found in the literature.[171−175]

8.1. V(D)J Recombination

To achieve greater diversity in the variable regions of both Ig heavy chains and TCR β chains and maximize antigen-binding affinity, germline V, D, and J gene segments are recombined in a combinatorial manner to generate specific antigen-binding proteins (Figure 9). In germline DNA, V, D, and J genes are flanked by conserved DNA motifs called recombination signal sequences (RSS). An RSS consists of two conserved motifs: a heptamer (CACAGTG) and a nonamer (ACAAAAACC). One should note, however, that these motifs can and frequently do vary, thus affecting the efficiency of recombination.[176] These elements must be separated by a stretch of DNA that is either 12 or 23 base pairs in length, making the final size of an RSS either 28 (12-RSS) or 39 (23-RSS) base pairs long (Figure 9B).[177] In immunoglobulin genes and some T cell receptor genes, the V and J segments are associated with 23-RSSs, while the D segments have 12-RSSs flanking them on either side. Recombination events always occur between gene segments bordered by RSSs of different sizes, thus promoting the recombination of gene segments from different regions and making it unlikely that two segments of the same region will recombine.[177] This is known as the "12/23 rule".[178] Upon recognition of the RSS elements by a mixed tetramer containing two monomers of recombinase activating gene products 1 and 2 (RAG1 and -2), a nick is created 5′ to the heptamer element of the RSS (Figure 9C). The nicked strand 3′-OH then initiates a trans-esterification (mediated by RAG) on the phosphate of the complementary strand thereby forming a covalent hairpin structure on the coding end, leaving the other end of the double stranded break (the RSS) blunt.[177,179] Interestingly, although the presence of the RAG proteins (RAG1 and RAG2) is critical for the initiation of V(D)J recombination, the exact mechanism of their activity is still unclear. It has been noted, however, that the active site of RAG1 is similar in many ways to the active sites of retroviral integrases and transposases.[180−183] After cleavage, the DNA ends remain in association with the RAG proteins, which most likely aids in the subsequent joining steps.[183] These newly created DNA ends are processed and joined through the NHEJ pathway in which the ends are recognized by the Ku proteins and then processed and ligated through the actions of "Artemis", DNA-PKcs, TdT, X-ray cross complementing group 4 (XRCC4), and DNA ligase IV[179,184−186] and possibly other factors not yet identified.[176] These NHEJ proteins will be described further in the following section.

8.2. Role of Terminal Deoxynucleotidyl Transferase in V(D)J Recombination

While, the combinatorial selection of V, D, and J gene segments serves as a major source of diversity for Ig and TCR genes, there is yet another strategy employed by the adaptive immune system to enhance the antigen-binding repertoire. During V(D)J recombination, the joining of V, D, and J segments is highly imprecise and the coding joints between these segments are observed to lose or gain nucleotides before they are finally ligated (Figure 9C). This extra genetic information may arise via our proposed mechanism shown in Figure 10. This mechanism is based on the enzymatic properties of Artemis, TdT, exonucleases, template-dependent DNA polymerases, and ligases and on the statistical analysis of the sequences of 543 mouse Ig heavy chains.[187] In the first step, the covalently closed coding

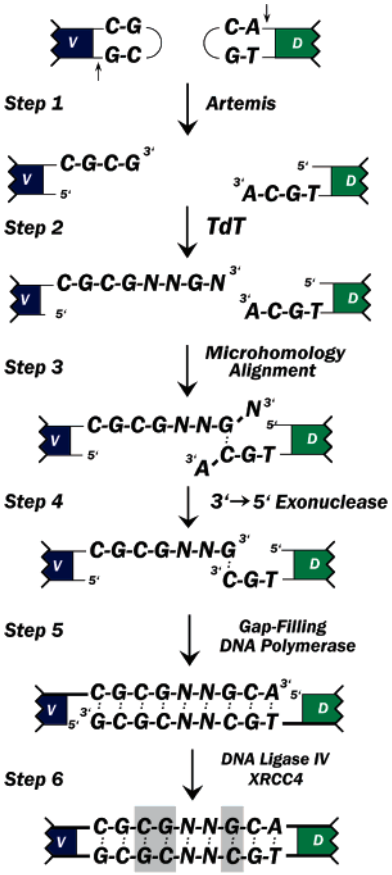


Figure 10. A proposed mechanism for the "N region" formation at the junction between a V and a D segment. The "P" nucleotides are shaded.

ends are opened via scission of the phosphodiester backbone at an imprecise location very near to the apex of the hairpin, most likely by the endonuclease Artemis.[188] This endonucleolytic cleavage can leave the DNA either blunt ended or with either a 3′ or a 5′ overhang. Following cleavage, a polynucleotide palindromic sequence is likely to exist due to the imprecision of the cleavage site. For example, Figure 10 shows that the terminal two nucleotides from the 5′ end of one of the gene segments are reversed and joined to the 3′ end of the other. The DNA ends created in this way are favorable substrates for TdT based on our preliminary finding that TdT catalyzes template-independent polymerization on this type of DNA as efficiently as on exclusively single-stranded DNA substrates (J. Fowler and Z. Suo, unpublished data). Interestingly, the statistical analysis of the sequences of 543 mouse Ig heavy chains indicates that the "N regions" are formed predominantly from DNA plus strand or from DNA minus strand polymerizations, rather than from both simultaneously.[187] Thus in step 2, TdT is shown to only extend the protruding plus strand at the V segment. The same statistical analysis has also suggested that homologous overlaps of as few as one nucleotide between gene segments may cause significant skewing of recombination sites.[187] Thus in step 3, the microhomology alignment between the two protruding strands is shown to occur through base pairing between two nucleotides. In step 4, the mismatched 3′-nucleotides on both strands are excised by an unknown 3′ → 5′ exonuclease. In step 5, an unknown template-dependent DNA polymerase catalyzes gap-filling DNA synthesis. It should be noted however that the fidelity of this gap-filling polymerase must not be too low because the sequence bias

in "N regions", which are rich in Gs and Cs, generated by TdT would be lost in the presence of an error-prone DNA polymerase. Also, double-stranded DNA containing any number of mismatched base pairs would not be stable enough for ligation in the next step. In addition, this gap-filling polymerase must be relatively processive to catalyze efficient V(D)J recombination. Additionally, to increase coordination of the V(D)J recombination reaction, the 3′ → 5′ exonuclease activity and the template-dependent DNA polymerase activity should reside on the same enzyme, as occurs in some repair DNA polymerases such as eukaryotic DNA polymerase ϵ. Finally, the nicked strands are ligated by DNA ligase IV in the presence of XRCC4. Notably, many steps in Figure 10 may require the addition of other NHEJ proteins, for example stabilizing of the base pairs in step 3 by the Ku proteins. Moreover, exonucleolytic activity may shorten the 3′ palindromic sequence by one or more nucleotides in step 4. However if any nucleotides from the palindromic sequence that were encoded by the germline DNA remain in the final coding joint, these bases are designated "P" nucleotides (for palindromic), which in Figure 10 are shaded.[189,190]

It has been demonstrated that TdT can only accomplish "N" nucleotide addition after recruitment to the site of recombination.[76,191] In addition, it has been shown that the factor responsible for the recruitment of TdT to the site of its activity is Ku. Perhaps one of the best understood functions of Ku is as a regulatory component of the DNA-dependent protein kinase (DNA-PK), which is a core component involved in the mammalian NHEJ pathway. However, Ku is also critical to the production of N-region diversity during the process of V(D)J recombination.[38,76,88] DNA-PKcs, the catalytic subunit of DNA-PK, is a nuclear serine/threonine kinase, which is recruited to DSBs by the Ku heterodimer and initiates repair of this potentially lethal damage via the phosphorylation of many downstream targets.[192] Ku is believed to regulate DNA-PK by acting as the primary recognition element of DSBs. Ku functions as a heterodimer composed of 70 and 80 kDa subunits having a high affinity for the ends of DNA duplexes.[76] After Ku binds to the ends of a DSB, the catalytic subunit of DNA-PK (DNA-PKcs) is recruited to initiate NHEJ, thereby repairing the damage. Therefore, it is reasonable to hypothesize that Ku is functioning in a similar capacity to recruit TdT to the ends of gene segments during the process of V(D)J recombination. The first convincing evidence to confirm that Ku and TdT are functional partners came in 1999 when these proteins were co-immunopreciptated from human Molt-4 lymphoid cell extracts.[76] This interaction was shown through mutational analysis to occur via the N-terminal BRCT domain of TdT. This revelation is not surprising considering that BRCT domains are commonly involved in mediation of protein−protein interactions between DNA repair components.[38,76,82,88] In addition, although TdT is stable in Ku80-deficient fibroblasts,[191] V(D)J junctions from these cells lack N regions, suggesting that Ku80 may play a crucial role in their formation. Ku70 is believed to stabilize Ku80, because neither of these proteins is stable by itself. It has also been demonstrated that Ku70 makes more intimate contact with DNA than does Ku80.[193] It should be noted however, that only the Ku70/80 dimer can act to recruit TdT.

8.3. Terminal Deoxynucleotidyl Transferase Regulation

As stated previously, it is of great importance that V(D)J recombination take place for the creation of a competent immune system. However, it is just as important that the process be tightly regulated. V(D)J recombination is observed to occur only very early in lymphocyte development and only at very well-defined loci.[179] While it has been demonstrated that TdT is regulated at the level of transcription by proteins such as AP-1,[194] regulation is thought to be achieved principally through the expression of the RAG genes.[179] However, the rate of random nucleotide addition catalyzed by TdT also seems to be under a complex system of both negative and positive control.

In addition to regulation at the level of expression, TdT is also regulated by proteins called TdT interacting factors (TdiFs). One of these proteins, TdiF1, is found to bind to TdT and enhance its polymerase activity by 1.5−4-fold.[195] Interestingly, TdiF1 interacts only with the C-terminus of TdT. If any portion of the 360 C-terminal amino acid residues of TdT is deleted, TdiF1 and the deletion mutant of TdT cannot bind to each other. Additionally, TdiF1 is expected to reside in the nucleus (where TdT is located) due to the presence of a nuclear localization sequence. The presence of TdiF1 in the nucleus is confirmed by both immunofluorescence microscopy and immunostaining.[195]

A second TdT interacting factor (TdiF2) has recently been identified and been shown to downregulate the polymerase activity of TdT in vitro.[185] Through immunoprecipitation, TdiF2 is demonstrated to bind TdT. However, for efficient binding, the entire C-terminus of TdT must be intact, including the proline-rich and Polβ-like domains. TdiF2 has also been shown to bind to single-stranded DNA. In the presence of increasing amounts of TdiF2, the polymerase activity of TdT is observed to drop by as much as 54% in vitro. Physically, TdiF2 is an acidic 82 kDa protein that is a member of a family of chromatin remodeling proteins.

Recently, TdT has been found to directly bind to TReP-132, which is involved in P450scc gene expression in steroid-hormone-producing cells or lymphoid cells.[196] The coexpression of TdT and TReP-132 in COS7 cells showed that these proteins are colocalized within the nucleus. TReP-132 reduces the N-addition activity of TdT to 2.5% of its maximum value in an in vitro polymerase assay in the presence of double-stranded DNA with a 3′ protrusion as a primer.[196] Thus, these results suggest that TReP-132 also downregulates the polymerase activity of TdT.

Proliferating cell nuclear antigen (PCNA) is another protein known to interact with TdT. Like TdiF2, PCNA downregulates TdT polymerase activity by as much as 83%.[195] It has been speculated that PCNA and TdiF1 may compete for the C-terminal region of TdT, representing a means of both positive and negative control for this enzyme.[195] Interestingly, upon binding to DNA, TdT has been demonstrated to lose its ability to bind both PCNA and TdiF2.[185,197]

Protein phosphorylation is another key regulatory mechanism of many cellular events. Experiments on labeling of human lymphoblastoid cells with [^{32}P]-phosphate has shown that TdT is phosphorylated in lymphoblastoid cells.[198] In vitro, recombinant human TdT has also been shown to be phosphorylated by protein kinase C (PKC).[199] PKC is found to localize to the nucleus in KM-3 cells.[199] In addition, fragments of calf thymus TdT are found to be independently phosphorylated by beef heart cAMP-dependent protein kinase, suggesting that calf thymus TdT may be phosphorylated at multiple sites.[101] These phosphorylation sites were later resolved to the TdT N-terminus, for example, Ser7 and

Thr19 in human TdT.[103,198] Although it is clear that TdT's are phosphorylated in vivo, how phosphorylation regulates the activity of TdT is not yet clear.

TdT activity at DNA ends may also be regulated indirectly by the highly homologous Polμ. Mahajan et al. have demonstrated that, at least in vivo, TdT and Polμ can efficiently compete for the same DNA substrate. If these proteins were both present during the process of V(D)J recombination, it is reasonable to assume that competition from Polμ may affect the activity of TdT during "N" region synthesis.[65]

The Ku proteins may also play a role in the regulation of TdT activity. TdT has been shown in vivo to add "N" nucleotides to double stranded DNA breaks generated through exonuclease activity. However, the "N" regions generated in the absence of Ku80 are unusually long.[200] This observation suggests that Ku80 may not only recruit TdT to the site of V(D)J recombination but also play a role in the regulation of its catalytic activity.

8.4. Somatic Hypermutation

As a B cell enters a germinal center of peripheral lymphoid tissue, it undergoes a second round of antibody diversification in a process called somatic hypermutation. In mice and humans, somatic hypermutation occurs at rates of 10^{-5} to 10^{-3} mutations per base pair per generation, which is about 10^6-fold higher than the spontaneous mutation rate in most other genes.[201] The somatic mutations are mainly single base substitions, with infrequent insertions and deletions. This process preferetially targets and mutates WRCY (W = A or T, R = A or G, Y = T or C) and WA motifs in the rearranged "variable" regions and its immediate flanking sequences, resulting in the generation of high-affinity antigen binding sites,[202] consequently developing the extensive and diverse immunoglobulin repertoire needed for survival. Variable regions that are not rearranged are rarely seen to undergo mutations.[203,204] The mature B cells resulting from this process are then selected for and become memory B cells, which produce antibodies for the recognition of pathogens.

So far, the mechanism of somatic hypermutation has not been established. Recently, somatic hypermutation was hypothesized to be initiated by a protein called activation-induced cytidine deaminase (AID)[205,206] based on the following facts: (i) AID, encoded by the *AICDA* gene, is expressed only in B lymphocytes;[207] (ii) mice deficient in AID are compeletely defective in somatic hypermutation and class switch recombination;[208] (iii) AID, which deaminates cytosine to uracil in DNA, peferentially targets WRC motifs in single-stranded DNA;[209] single-stranded DNA may arise transiently during gene transcription;[210] (iv) somatic hypermutation depends on the active transcription of antibody genes to create the target for cytosine deamination by AID.[211] Following initiation, the U:G mispairs in Ig DNA are either directly copied by a DNA polymerase to form C:G to T:A transition mutations, excised by a uracil−DNA glycosylase and then repaired through BER, or recognized by MSH2−MSH6 mismatch-recognition complex and then repired.[210] The latter two possiblities like the first also involve DNA polymerase(s), which may exhibit low fidelity and produce mutations.

At present, it is not clear how many or which DNA polymerase(s) catalyze somatic hypermutation. We speculate that TdT can be excluded as a candidate because it is not a template-dependent DNA polymerase nor is it expressed in the germinal centers of peripheral lymphoid tissues. This speculation is substantiated by the observation that somatic hypermutation can occur in the B cells of TdT-deficient mutant mice.[212] Preferential expression in secondary lymphoid tissues,[64] as well as the low template-dependent polymerization fidelity (10^{-3}−10^{-5}), of Polμ[69] has led to the hypothesis that this most homologous X-family DNA polymerase to TdT is an error-prone mutase, active in somatic hypermutation.[64] However, no alterations in the somatic hypermutation process have been found in Polμ knockout mice.[62]

Interestingly, four error-prone DNA polymerases including Polθ (A-family), Polζ (B-family), Polη (Y-family), and Polι (Y-family) have been implicated directly or indirectly in somatic hypermutation[5,213] based on the following evidence: (i) Polθ is highly expressed in lymphoid tissues including the spleen and germinal centers.[214,215] A complete deletion of the gene encoding Polθ in mice leads to a reduction of overall somatic hypermutation frequency by 60−80%, and the mutation spectrum is moderately shifted toward more transitions at both A:T and C:G basepairs.[216] (ii) Inhibition of the catalytic subunit of Polζ in human B cells by specific phosphorothioate-modified oligonucleotides impaires Ig and bcl-6 hypermutation by ~70%.[217] Expressing antisense RNA to a portion of mouse *REV3*, the gene encoding Polζ, in transgenic mice has been found to delay the generation of high-affinity antibodies and to decrease the accumulation of somatic mutations in the V_H gene segments of memory B cells.[218] (iii) Human Polι has an average misincorporation frequency of 10^{-2}−10^{-3}.[31−33,219] Strikingly, the fidelity of DNA incorporation by Polι is asymmetric, with a misincorporation rate of about 1×10^{-4} at a template base A, while the incorporation of G is favored by 3−11 times over A opposite a template base T.[31,32] This observed asymmetric fidelity is surprisingly similar to the strand bias found in Ig V regions,[220,221] where there are more mutations from A (due to misincorporations opposite template T) and fewer mutations from T (due to accurate incorporations opposite template A). (iv) Steady-state kinetic analyses have shown that both human Polη[222−226] and yeast Polη[226−228] incorporate nucleotides opposite both normal and UV-damaged DNA with a similarly low fidelity of about 10^{-2}−10^{-3}. When human Polη is mutated among xeroderma pigmentosum variant patients, they have normal immune systems and undergo somatic hypermutation, but they have altered mutation spectra.[229] (v) Expression of Polζ, Polη, and Polι in cultured Burkitt's lymphoma cells leads to a 5−10-fold increase in heavy chain V-region mutations if co-stimulated with T cells and IgM cross-linking, the presumed in vivo requirements for somatic hypermutation.[213] Together, these data suggest that more than one DNA polymerase is likely to be involved in somatic hypermutation.

9. Experience Dependent Memory Processing and Terminal Deoxynucleotidyl Transferase

In addition to V(D)J recombination, TdT has been hypothesized to play a role in the storage of memory. As early as 1965, it was suggested that long-term memory could be stored in the form of structural modification of synaptic connections within the brain.[230] In addition, these structural modifications are known to require protein synthesis.[231−235] It has been noted that the immune system and the nervous system are similar in many ways. Most notably, both systems

have the unique task of storing environmental information that is not genetically inherited.[235] If TdT were to play a role in memory storage, it would almost certainly have to be expressed in brain cells. Unfortunately, reports of TdT expression in the nervous system have been mixed. For example, in 1976 Viola et al.[236] reported that a cell lysate prepared from the cerebral cortex of the occipital lobe of a human with no cerebral pathology was demonstrated to exhibit TdT polymerase activity. But, in 1997 when rainbow trout cell lysates were analyzed for the presence of TdT and RAG1, reverse-transcriptase polymerase chain reaction analysis failed to detect the presence of TdT cDNA in brain tissue.[83] However, in 2003 TdT mRNA was in fact detected in the neurons of mouse brain tissue using an in situ hybridization screening.[235] Specifically, TdT mRNA was found in neuronal cells of the hippocampal formation, cerebellum, amygdala, and neocortex. These areas of the brain have all been implicated in the storage of memory.[235,237] In addition, the level of TdT mRNA in mice is found to differ as a function of the environment in which they are raised. Mice raised in enriched and highly stimulatory environments demonstrate enhanced spatial discrimination learning and memory. However, transgenic mice that have no TdT genes are found not to benefit in this way from an enriched environment during their development.[235] It may be noted that many of the components responsible for V(D)J recombination are also critical to neurogenesis.[235,238−242] Therefore, one might theorize that the TdT knockout mice used in this study may in fact be poor "learners" not because the TdT activity is missing but because of inhibited neurogenesis. However, the enrichment induced improvement of learning has been shown not to be dependent upon neurogenesis.[235,243]

10. Conclusions and Future Directions

Although TdT has been extensively studied by biochemists and immunologists since its discovery in 1960, many questions still remain to be answered about this unique template-independent DNA polymerase. First, a detailed kinetic mechanism including microscopic rate constants for each elementary step has not yet been published. Although such mechanisms have been established for template-dependent polymerases, the absence of a template strand would certainly alter the kinetic parameters for DNA and nucleotide binding, catalysis, and the product dissociation steps. Such a mechanism would provide insight into the kinetic and thermodynamic basis for random nucleotide additions during V(D)J recombination. Second, a ternary structure of TdT, a primer, and a nucleotide is not yet available. This structure not only would reveal how TdT interacts with a primer and an incoming nucleotide but also may provide a structural basis for protein−protein interaction between TdT and other NHEJ proteins and perhaps provide additional insight as to why TdT is a template-independent polymerase. Third, the identity and multiplicity of metal ions used by TdT in vivo is not yet clear. Free cellular Mg^{2+} concentrations are estimated to be in millimolar range, while concentrations of other transition metal ions are extremely low due to the tight binding and regulation by cellular proteins.[244] This fact, combined with the knowledge that TdT, like other DNA polymerases, does not possess high affinities for metal ions, suggests that the metal ion cofactor utilized by TdT in vivo is likely to be Mg^{2+}. However, although Mg^{2+} is recognized as the most relevant metal ion available to DNA polymerases in vivo, there is no direct evidence to single out Mg^{2+} as the metal ion cofactor for any polymerase including TdT. Fourth, TdT is known to have low sugar selectivity, and incorporation of ribonucleotides is known to lead to premature chain termination in vitro.[124] Thus, it is reasonable to expect that rNTP incorporation should affect V(D)J recombination in vivo. However, one is left to wonder what the consequences of such incorporation would be. Fifth, TdT alone is an exclusively distributive polymerase in vitro.[124] This is partly due to the absence of a DNA template. Nevertheless, "N regions" added by TdT in vivo are as long as 10 nucleotides and 3−4 nucleotides long on average.[187] This suggests that the processivity of TdT might in some way be slightly increased. It has been shown that Ku acts to recruit TdT, through interaction with its BRCT domain, to the site of V(D)J recombination.[76] It might be speculated that TdT could bind more tightly to DNA and become moderately processive in the presence of Ku. Sixth, the possibility still exists that TdT could interact with other NHEJ proteins thereby affecting their roles. Seventh, the reaction sequences in Figure 10 and the identity of the 3′ → 5′ exonuclease and the template-dependent gap-filling DNA polymerase are unknown. Those enzymes in Figure 10 have different preferences to double-stranded and single-stranded DNA substrates, and this may dictate the reaction sequences during random addition and deletion of nucleotides at the V, D, and J junctions. Finally, the possibility that TdT might be involved in a specific type of learning is intriguing. However, much more evidence is needed to confirm this potential in vivo role.

In this review, we have explored the world of DNA polymerases and have attempted to define TdT in relation to its peers. We have outlined the enzymatic functions, physical characteristics, and regulation of TdT, explored the process of V(D)J recombination, and even discussed some preliminary work suggesting that TdT may play a role in specific types of learning and memory storage. The value of TdT and of low-fidelity DNA polymerases as a whole is hard to overestimate. At least it is clear that without these invaluable enzymes, biological organisms and the precious genetic information they depend on would be devastatingly vulnerable to the harsh and ubiquitous assaults of the universe. Clearly, our fitness and indeed our survival depends on the tireless and unconventional functioning of these unique enzymes.

11. List of Frequently Used Abbreviations

AID	activation-induced cytidine deaminase
ASFV PolX	African swine fever virus DNA polymerase X
BER	base excision repair
BRCT	breast cancer susceptibility protein BRCA1 C-terminus
dRPase	5′-deoxyribose-5-phosphate lyase
CDR	complementarity determining region
C gene	"constant" gene segment
DSBs	double stranded breaks
HhH	helix−hairpin−helix
Ig	immunoglobulin
DNA-PK	DNA-dependent protein kinase
NHEJ	nonhomologous end joining
PCNA	proliferating cell nuclear antigen
Pol IV	DNA polymerase IV
NLS	nuclear localization signal
Polβ	DNA polymerase beta
Polη	DNA polymerase eta
Polι	DNA polymerase iota

Polκ	DNA polymerase kappa
Polλ	DNA polymerase lambda
Polμ	DNA polymerase mu
Polσ	DNA polymerase sigma
Polζ	DNA polymerase zeta
RAG	recombinase activating gene product
RSS	recombination signal sequences
TCR	T cell antigen receptor
TdiF	TdT interacting factor
TdT	terminal deoxynucleotidyl transferase
TdTL	TdT long isoform
TdTS	TdT short isoform
XRCC4	X-ray cross complementing group 4.

12. Acknowledgment

We would like to thank Dr. Kiran Mahajan at the Lineberger Comprehensive Cancer Center of the University of North Carolina and Mr. Kevin A. Fiala and Mr. Nikunj Bhatt at The Ohio State University for proof reading this manuscript.

13. References

(1) Rattray, A. J.; Strathern, J. N. *Annu. Rev. Genet.* **2003**, *37*, 31.
(2) Bessman, M. J.; Kornberg, A.; Lehman, I. R.; Simms, E. S. *Biochim. Biophys. Acta* **1956**, *21*, 197.
(3) Lehman, I. R.; Bessman, M. J.; Simms, E. S.; Kornberg, A. *J. Biol. Chem.* **1958**, *233*, 163.
(4) Ramadan, K.; Shevelev, I. V.; Maga, G.; Hubscher, U. *J. Biol. Chem.* **2002**, *277*, 18454.
(5) Hubscher, U.; Maga, G.; Spadari, S. *Annu. Rev. Biochem.* **2002**, *71*, 133.
(6) Filee, J.; Forterre, P.; Sen-Lin, T.; Laurent, J. *J. Mol. Evol.* **2002**, *54*, 763.
(7) Kornberg, A.; Baker, T. A. *DNA replication*, 2nd ed.; W. H. Freeman: New York, 1992.
(8) Johnson, A. A.; Tsai, Y.; Graves, S. W.; Johnson, K. A. *Biochemistry* **2000**, *39*, 1702.
(9) Graves, S. W.; Johnson, A. A.; Johnson, K. A. *Biochemistry* **1998**, *37*, 6050.
(10) Maga, G.; Shevelev, I.; Ramadan, K.; Spadari, S.; Hubscher, U. *J. Mol. Biol.* **2002**, *319*, 359.
(11) Jones, E. V.; Moss, B. *J. Virol.* **1984**, *49*, 72.
(12) Traktman, P.; Sridhar, P.; Condit, R. C.; Roberts, B. E. *J. Virol.* **1984**, *49*, 125.
(13) Nelson, J. R.; Lawrence, C. W.; Hinkle, D. C. *Science* **1996**, *272*, 1646.
(14) Gibbs, P. E.; McGregor, W. G.; Maher, V. M.; Nisson, P.; Lawrence, C. W. *Proc. Natl. Acad. Sci. U. S.* A. **1998**, *95*, 6876.
(15) Lehman, I. R.; Kaguni, L. S. *J. Biol. Chem.* **1989**, *264*, 4265.
(16) Hindges, R.; Hubscher, U. *Biol. Chem.* **1997**, *378*, 345.
(17) Syvaoja, J.; Suomensaari, S.; Nishida, C.; Goldsmith, J. S.; Chui, G. S.; Jain, S.; Linn, S. *Proc. Natl. Acad. Sci. U.S.A.* **1990**, *87*, 6664.
(18) Guo, D.; Wu, X.; Rajpal, D. K.; Taylor, J. S.; Wang, Z. *Nucleic Acids Res.* **2001**, *29*, 2875.
(19) Burgers, P. M.; Koonin, E. V.; Bruford, E.; Blanco, L.; Burtis, K. C.; Christman, M. F.; Copeland, W. C.; Friedberg, E. C.; Hanaoka, F.; Hinkle, D. C.; Lawrence, C. W.; Nakanishi, M.; Ohmori, H.; Prakash, L.; Prakash, S.; Reynaud, C. A.; Sugino, A.; Todo, T.; Wang, Z.; Weill, J. C.; Woodgate, R. *J. Biol. Chem.* **2001**, *276*, 43487.
(20) Cann, I. K.; Komori, K.; Toh, H.; Kanai, S.; Ishino, Y. *Proc. Natl. Acad. Sci. U.S.A.* **1998**, *95*, 14250.
(21) Cann, I. K.; Ishino, Y. *Genetics* **1999**, *152*, 1249.
(22) Ishino, Y.; Komori, K.; Cann, I. K.; Koga, Y. *J. Bacteriol.* **1998**, *180*, 2232.
(23) Gueguen, Y.; Rolland, J. L.; Lecompte, O.; Azam, P.; Le Romancer, G.; Flament, D.; Raffin, J. P.; Dietrich, J. *Eur. J. Biochem.* **2001**, *268*, 5961.
(24) Shen, Y.; Musti, K.; Hiramoto, M.; Kikuchi, H.; Kawarabayashi, Y.; Matsui, I. *J. Biol. Chem.* **2001**, *276*, 27376.
(25) Ohashi, E.; Bebenek, K.; Matsuda, T.; Feaver, W. J.; Gerlach, V. L.; Friedberg, E. C.; Ohmori, H.; Kunkel, T. A. *J. Biol. Chem.* **2000**, *275*, 39678.
(26) Masutani, C.; Kusumoto, R.; Yamada, A.; Dohmae, N.; Yokoi, M.; Yuasa, M.; Araki, M.; Iwai, S.; Takio, K.; Hanaoka, F. *Nature* **1999**, *399*, 700.
(27) Levine, R. L.; Miller, H.; Grollman, A.; Ohashi, E.; Ohmori, H.; Masutani, C.; Hanaoka, F.; Moriya, M. *J. Biol. Chem.* **2001**, *276*, 18717.
(28) Bebenek, K.; Tissier, A.; Frank, E. G.; McDonald, J. P.; Prasad, R.; Wilson, S. H.; Woodgate, R.; Kunkel, T. A. *Science* **2001**, *291*, 2156.
(29) Haracska, L.; Yu, S. L.; Johnson, R. E.; Prakash, L.; Prakash, S. *Nat. Genet.* **2000**, *25*, 458.
(30) Fiala, K. A.; Suo, Z. *Biochemistry* **2004**, *43*, 2116.
(31) Vaisman, A.; Tissier, A.; Frank, E. G.; Goodman, M. F.; Woodgate, R. *J. Biol. Chem.* **2001**, *276*, 30615.
(32) Zhang, Y.; Yuan, F.; Wu, X.; Wang, Z. *Mol. Cell. Biol.* **2000**, *20*, 7099.
(33) Tissier, A.; McDonald, J. P.; Frank, E. G.; Woodgate, R. *Genes Dev.* **2000**, *14*, 1642.
(34) Nair, D. T.; Johnson, R. E.; Prakash, S.; Prakash, L.; Aggarwal, A. K. *Nature* **2004**, *430*, 377.
(35) Aoufouchi, S.; Flatter, E.; Dahan, A.; Faili, A.; Bertocci, B.; Storck, S.; Delbos, F.; Cocea, L.; Gupta, N.; Weill, J. C.; Reynaud, C. A. *Nucleic Acids Res.* **2000**, *28*, 3684.
(36) Nagasawa, K.; Kitamura, K.; Yasui, A.; Nimura, Y.; Ikeda, K.; Hirai, M.; Matsukage, A.; Nakanishi, M. *J. Biol. Chem.* **2000**, *275*, 31233.
(37) Garcia-Diaz, M.; Dominguez, O.; Lopez-Fernandez, L. A.; de Lera, L. T.; Saniger, M. L.; Ruiz, J. F.; Parraga, M.; Garcia-Ortiz, M. J.; Kirchhoff, T.; del Mazo, J.; Bernad, A.; Blanco, L. *J. Mol. Biol.* **2000**, *301*, 851.
(38) Dominguez, O.; Ruiz, J. F.; Lain de Lera, T.; Garcia-Diaz, M.; Gonzalez, M. A.; Kirchhoff, T.; Martinez, A. C.; Bernad, A.; Blanco, L. *EMBO J.* **2000**, *19*, 1731.
(39) Oliveros, M.; Yanez, R. J.; Salas, M. L.; Salas, J.; Vinuela, E.; Blanco, L. *J. Biol. Chem.* **1997**, *272*, 30899.
(40) Prasad, R.; Widen, S. G.; Singhal, R. K.; Watkins, J.; Prakash, L.; Wilson, S. H. *Nucleic Acids Res.* **1993**, *21*, 5301.
(41) Carson, D. R.; Christman, M. F. *Proc. Natl. Acad. Sci. U.S.A.* **2001**, *98*, 8270.
(42) Wang, Z.; Castano, I. B.; De Las Penas, A.; Adams, C.; Christman, M. F. *Science* **2000**, *289*, 774.
(43) Wang, Z.; Castano, I. B.; Adams, C.; Vu, C.; Fitzhugh, D.; Christman, M. F. *Genetics* **2002**, *160*, 381.
(44) Friedberg, E. C. *DNA repair and mutagenesis*, 2nd ed.; ASM Press: Washington, DC, 2006.
(45) Coleman, M. S.; Hutton, J. J.; De Simone, P.; Bollum, F. J. *Proc. Natl. Acad. Sci. U.S.A.* **1974**, *71*, 4404.
(46) Goldschneider, I.; Gregoire, K. E.; Barton, R. W.; Bollum, F. J. *Proc. Natl. Acad. Sci. U.S.A.* **1977**, *74*, 734.
(47) Bollum, F. J. *Blood* **1979**, *54*, 1203.
(48) Gregoire, K. E.; Goldschneider, I.; Barton, R. W.; Bollum, F. J. *J. Immunol.* **1979**, *123*, 1347.
(49) Gregoire, K. E.; Goldschneider, I.; Barton, R. W.; Bollum, F. J. *Proc. Natl. Acad. Sci. U.S.A.* **1977**, *74*, 3993.
(50) Dianov, G. L.; Prasad, R.; Wilson, S. H.; Bohr, V. A. *J. Biol. Chem.* **1999**, *274*, 13741.
(51) Beard, W. A.; Wilson, S. H. *Mutat. Res.* **2000**, *460*, 231.
(52) Wilson, T. E.; Lieber, M. R. *J. Biol. Chem.* **1999**, *274*, 23599.
(53) Bebenek, K.; Garcia-Diaz, M.; Patishall, S. R.; Kunkel, T. A. *J. Biol. Chem.* **2005**, *280*, 20051.
(54) Hirose, F.; Hotta, Y.; Yamaguchi, M.; Matsukage, A. *Exp. Cell Res.* **1989**, *181*, 169.
(55) Garcia-Diaz, M.; Bebenek, K.; Kunkel, T. A.; Blanco, L. *J. Biol. Chem.* **2001**, *276*, 34659.
(56) Braithwaite, E. K.; Prasad, R.; Shock, D. D.; Hou, E. W.; Beard, W. A.; Wilson, S. H. *J. Biol. Chem.* **2005,** *280*, 18469.
(57) Uchiyama, Y.; Kimura, S.; Yamamoto, T.; Ishibashi, T.; Sakaguchi, K. *Eur. J. Biochem.* **2004**, *271*, 2799.
(58) Braithwaite, E. K.; Kedar, P. S.; Lan, L.; Polosina, Y. Y.; Asagoshi, K.; Poltoratsky, V. P.; Horton, J. K.; Miller, H.; Teebor, G. W.; Yasui, A.; Wilson, S. H. *J. Biol. Chem.* **2005**, *280*, 31641.
(59) Lee, J. W.; Blanco, L.; Zhou, T.; Garcia-Diaz, M.; Bebenek, K.; Kunkel, T. A.; Wang, Z.; Povirk, L. F. *J. Biol. Chem.* **2004**, *279*, 805.
(60) Fan, W.; Wu, X. *Biochem. Biophys. Res. Commun.* **2004**, *323*, 1328.
(61) Maga, G.; Villani, G.; Ramadan, K.; Shevelev, I.; Tanguy Le Gac, N.; Blanco, L.; Blanca, G.; Spadari, S.; Hubscher, U. *J. Biol. Chem.* **2002**, *277*, 48434.
(62) Bertocci, B.; De Smet, A.; Flatter, E.; Dahan, A.; Bories, J. C.; Landreau, C.; Weill, J. C.; Reynaud, C. A. *J. Immunol.* **2002**, *168*, 3702.
(63) Kobayashi, Y.; Watanabe, M.; Okada, Y.; Sawa, H.; Takai, H.; Nakanishi, M.; Kawase, Y.; Suzuki, H.; Nagashima, K.; Ikeda, K.; Motoyama, N. *Mol. Cell. Biol.* **2002**, *22*, 2769.
(64) Ruiz, J. F.; Dominguez, O.; Lain de Lera, T.; Garcia-Diaz, M.; Bernad, A.; Blanco, L. *Philos. Trans. R. Soc. London, Ser. B* **2001**, *356*, 99.

(65) Mahajan, K. N.; Nick McElhinny, S. A.; Mitchell, B. S.; Ramsden, D. A. *Mol. Cell. Biol.* **2002**, *22*, 5194.
(66) Zhang, Y.; Wu, X.; Yuan, F.; Xie, Z.; Wang, Z. *Mol. Cell. Biol.* **2001**, *21*, 7995.
(67) Nick McElhinny, S. A.; Ramsden, D. A. *Mol. Cell. Biol.* **2003**, *23*, 2309.
(68) Ruiz, J. F.; Juarez, R.; Garcia-Diaz, M.; Terrados, G.; Picher, A. J.; Gonzalez-Barrera, S.; Fernandez de Henestrosa, A. R.; Blanco, L. *Nucleic Acids Res.* **2003**, *31*, 4441.
(69) Roettger, M. P.; Fiala, K. A.; Sompalli, S.; Dong, Y.; Suo, Z. *Biochemistry* **2004**, *43*, 13827.
(70) Havener, J. M.; McElhinny, S. A.; Bassett, E.; Gauger, M.; Ramsden, D. A.; Chaney, S. G. *Biochemistry* **2003**, *42*, 1777.
(71) Zhang, Y.; Wu, X.; Guo, D.; Rechkoblit, O.; Taylor, J. S.; Geacintov, N. E.; Wang, Z. *J. Biol. Chem.* **2002**, *277*, 44582.
(72) Ollis, D. L.; Brick, P.; Hamlin, R.; Xuong, N. G.; Steitz, T. A. *Nature* **1985**, *313*, 762.
(73) Maciejewski, M. W.; Shin, R.; Pan, B.; Marintchev, A.; Denninger, A.; Mullen, M. A.; Chen, K.; Gryk, M. R.; Mullen, G. P. *Nat. Struct. Biol.* **2001**, *8*, 936.
(74) Showalter, A. K.; Byeon, I. J.; Su, M. I.; Tsai, M. D. *Nat. Struct. Biol.* **2001**, *8*, 942.
(75) Huyton, T.; Bates, P. A.; Zhang, X.; Sternberg, M. J.; Freemont, P. S. *Mutat. Res.* **2000**, *460*, 319.
(76) Mahajan, K. N.; Gangi-Peterson, L.; Sorscher, D. H.; Wang, J.; Gathy, K. N.; Mahajan, N. P.; Reeves, W. H.; Mitchell, B. S. *Proc. Natl. Acad. Sci. U.S.A.* **1999**, *96*, 13926.
(77) Zhu, C.; Bogue, M. A.; Lim, D. S.; Hasty, P.; Roth, D. B. *Cell* **1996**, *86*, 379.
(78) Shimazaki, N.; Yoshida, K.; Kobayashi, T.; Toji, S.; Tamai, K.; Koiwai, O. *Genes Cells* **2002**, *7*, 639.
(79) Maga, G.; Ramadan, K.; Locatelli, G. A.; Shevelev, I.; Spadari, S.; Hubscher, U. *J. Biol. Chem.* **2005**, *280*, 1971.
(80) DeRose, E. F.; Kirby, T. W.; Mueller, G. A.; Bebenek, K.; Garcia-Diaz, M.; Blanco, L.; Kunkel, T. A.; London, R. E. *Biochemistry* **2003**, *42*, 9564.
(81) Garcia-Diaz, M.; Bebenek, K.; Krahn, J. M.; Kunkel, T. A.; Pedersen, L. C. *Nat. Struct. Mol. Biol.* **2005**, *12*, 97.
(82) Ramadan, K.; Shevelev, I. V.; Maga, G.; Hubscher, U. *J. Mol. Biol.* **2004**, *339*, 395.
(83) Hansen, J. D. *Immunogenetics* **1997**, *46*, 367.
(84) Lee, A.; Hsu, E. *J. Immunol.* **1994**, *152*, 4500.
(85) Yang, B.; Gathy, K. N.; Coleman, M. S. *Nucleic Acids Res.* **1995**, *23*, 2041.
(86) Koiwai, O.; Yokota, T.; Kageyama, T.; Hirose, T.; Yoshida, S.; Arai, K. *Nucleic Acids Res.* **1986**, *14*, 5777.
(87) Peterson, R. C.; Cheung, L. C.; Mattaliano, R. J.; Chang, L. M.; Bollum, F. J. *Proc. Natl. Acad. Sci. U.S.A.* **1984**, *81*, 4363.
(88) Delarue, M.; Boule, J. B.; Lescar, J.; Expert-Bezancon, N.; Jourdan, N.; Sukumar, N.; Rougeon, F.; Papanicolaou, C. *EMBO J.* **2002**, *21*, 427.
(89) Bartl, S.; Miracle, A. L.; Rumfelt, L. L.; Kepler, T. B.; Mochon, E.; Litman, G. W.; Flajnik, M. F. *Immunogenetics* **2003**, *55*, 594.
(90) Bentolila, L. A.; Fanton d'Andon, M.; Nguyen, Q. T.; Martinez, O.; Rougeon, F.; Doyen, N. *EMBO J.* **1995**, *14*, 4221.
(91) Doyen, N.; d'Andon, M. F.; Bentolila, L. A.; Nguyen, Q. T.; Rougeon, F. *Nucleic Acids Res.* **1993**, *21*, 1187.
(92) Boule, J. B.; Rougeon, F.; Papanicolaou, C. *J. Biol. Chem.* **2000**, *275*, 28984.
(93) Thai, T. H.; Purugganan, M. M.; Roth, D. B.; Kearney, J. F. *Nat. Immunol.* **2002**, *3*, 457.
(94) Benedict, C. L.; Gilfillan, S.; Kearney, J. F. *J. Exp. Med.* **2001**, *193*, 89.
(95) Takahara, K.; Hayashi, N.; Fujita-Sagawa, K.; Morishita, T.; Hashimoto, Y.; Noda, A. *Biosci., Biotechnol., Biochem.* **1994**, *58*, 786.
(96) Thai, T. H.; Kearney, J. F. *J. Immunol.* **2004**, *173*, 4009.
(97) Thai, T. H.; Kearney, J. F. *Adv. Immunol.* **2005**, *86*, 113.
(98) Chang, L. M.; Bollum, F. J. *J. Biol. Chem.* **1971**, *246*, 909.
(99) Nakamura, H.; Tanabe, K.; Yoshida, S.; Morita, T. *J. Biol. Chem.* **1981**, *256*, 8745.
(100) Deibel, M. R., Jr.; Coleman, M. S. *Arch. Biochem. Biophys.* **1980**, *202*, 414.
(101) Chang, L. M.; Bollum, F. J. *J. Biol. Chem.* **1982**, *257*, 9588.
(102) Yang, B.; Gathy, K. N.; Coleman, M. S. *J. Biol. Chem.* **1994**, *269*, 11859.
(103) Chang, L. M.; Rafter, E.; Rusquet-Valerius, R.; Peterson, R. C.; White, S. T.; Bollum, F. J. *J. Biol. Chem.* **1988**, *263*, 12509.
(104) Boule, J. B.; Johnson, E.; Rougeon, F.; Papanicolaou, C. *Mol. Biotechnol.* **1998**, *10*, 199.
(105) Pelletier, H.; Sawaya, M. R.; Wolfle, W.; Wilson, S. H.; Kraut, J. *Biochemistry* **1996**, *35*, 12742.
(106) Gao, Y.; Sun, Y.; Frank, K. M.; Dikkes, P.; Fujiwara, Y.; Seidl, K. J.; Sekiguchi, J. M.; Rathbun, G. A.; Swat, W.; Wang, J.; Bronson, R. T.; Malynn, B. A.; Bryans, M.; Zhu, C.; Chaudhuri, J.; Davidson, L.; Ferrini, R.; Stamato, T.; Orkin, S. H.; Greenberg, M. E.; Alt, F. W. *Cell* **1998**, *95*, 891.
(107) Franklin, M. C.; Wang, J.; Steitz, T. A. *Cell* **2001**, *105*, 657.
(108) Johnson, S. J.; Taylor, J. S.; Beese, L. S. *Proc. Natl. Acad. Sci. U.S.A.* **2003**, *100*, 3895.
(109) Doublie, S.; Tabor, S.; Long, A. M.; Richardson, C. C.; Ellenberger, T. *Nature* **1998**, *391*, 251.
(110) Shen, Y.; Tang, X. F.; Yokoyama, H.; Matsui, E.; Matsui, I. *Nucleic Acids Res.* **2004**, *32*, 158.
(111) Silvian, L. F.; Toth, E. A.; Pham, P.; Goodman, M. F.; Ellenberger, T. *Nat. Struct. Biol.* **2001**, *8*, 984.
(112) Trincao, J.; Johnson, R. E.; Escalante, C. R.; Prakash, S.; Prakash, L.; Aggarwal, A. K. *Mol. Cell* **2001**, *8*, 417.
(113) Uljon, S. N.; Johnson, R. E.; Edwards, T. A.; Prakash, S.; Prakash, L.; Aggarwal, A. K. *Structure (Cambridge)* **2004**, *12*, 1395.
(114) Showalter, A. K.; Tsai, M. D. *J. Am. Chem. Soc.* **2001**, *123*, 1776.
(115) Kohlstaedt, L. A.; Wang, J.; Friedman, J. M.; Rice, P. A.; Steitz, T. A. *Science* **1992**, *256*, 1783.
(116) Kim, Y.; Eom, S. H.; Wang, J.; Lee, D. S.; Suh, S. W.; Steitz, T. A. *Nature* **1995**, *376*, 612.
(117) Kiefer, J. R.; Mao, C.; Hansen, C. J.; Basehore, S. L.; Hogrefe, H. H.; Braman, J. C.; Beese, L. S. *Structure* **1997**, *5*, 95.
(118) Korolev, S.; Nayal, M.; Barnes, W. M.; Di Cera, E.; Waksman, G. *Proc. Natl. Acad. Sci. U.S.A.* **1995**, *92*, 9264.
(119) Davies, J. F., 2nd; Almassy, R. J.; Hostomska, Z.; Ferre, R. A.; Hostomsky, Z. *Cell* **1994**, *76*, 1123.
(120) Georgiadis, M. M.; Jessen, S. M.; Ogata, C. M.; Telesnitsky, A.; Goff, S. P.; Hendrickson, W. A. *Structure* **1995**, *3*, 879.
(121) Sawaya, M. R.; Prasad, R.; Wilson, S. H.; Kraut, J.; Pelletier, H. *Biochemistry* **1997**, *36*, 11205.
(122) Steitz, T. A. *Curr. Opin. Struct. Biol.* **1993**, *3*, 31.
(123) Joyce, C. M.; Steitz, T. A. *Annu. Rev. Biochem.* **1994**, *63*, 777.
(124) Boule, J. B.; Rougeon, F.; Papanicolaou, C. *J. Biol. Chem.* **2001**, *276*, 31388.
(125) Kato, K. I.; Goncalves, J. M.; Houts, G. E.; Bollum, F. J. *J. Biol. Chem.* **1967**, *242*, 2780.
(126) Doherty, A. J.; Serpell, L. C.; Ponting, C. P. *Nucleic Acids Res.* **1996**, *24*, 2488.
(127) Garcia-Diaz, M.; Bebenek, K.; Krahn, J. M.; Blanco, L.; Kunkel, T. A.; Pedersen, L. C. *Mol. Cell* **2004**, *13*, 561.
(128) Eger, B. T.; Benkovic, S. J. *Biochemistry* **1992**, *31*, 9227.
(129) Wong, I.; Patel, S. S.; Johnson, K. A. *Biochemistry* **1991**, *30*, 526.
(130) Bollum, F. J. *Terminal deoxynucleotidyl transferase*; Academic Press: New York, 1974.
(131) Wong, D.; Demple, B. *J. Biol. Chem.* **2004**, *279*, 25268.
(132) Parsons, J. L.; Dianova, I. I.; Allinson, S. L.; Dianov, G. L. *Biochemistry* **2005**, *44*, 10613.
(133) Matsumoto, Y.; Kim, K.; Katz, D. S.; Feng, J. A. *Biochemistry* **1998**, *37*, 6456.
(134) Sosunov, V. V.; Santamaria, F.; Victorova, L. S.; Gosselin, G.; Rayner, B.; Krayevsky, A. A. *Nucleic Acids Res.* **2000**, *28*, 1170.
(135) Johnson, D.; Morgan, A. R. *Biochem. Biophys. Res. Commun.* **1976**, *72*, 840.
(136) Chang, L. M.; Bollum, F. J. *J. Biol. Chem.* **1990**, *265*, 17436.
(137) Joyce, C. M.; Benkovic, S. J. *Biochemistry* **2004**, *43*, 14317.
(138) Dahlberg, M. E.; Benkovic, S. J. *Biochemistry* **1991**, *30*, 4835.
(139) Patel, S. S.; Wong, I.; Johnson, K. A. *Biochemistry* **1991**, *30*, 511.
(140) Hsieh, J. C.; Zinnen, S.; Modrich, P. *J. Biol. Chem.* **1993**, *268*, 24607.
(141) Kati, W. M.; Johnson, K. A.; Jerva, L. F.; Anderson, K. S. *J. Biol. Chem.* **1992**, *267*, 25988.
(142) Johnson, A. A.; Johnson, K. A. *J. Biol. Chem.* **2001**, *276*, 38097.
(143) Washington, M. T.; Prakash, L.; Prakash, S. *Cell* **2001**, *107*, 917.
(144) Boudsocq, F.; Kokoska, R. J.; Plosky, B. S.; Vaisman, A.; Ling, H.; Kunkel, T. A.; Yang, W.; Woodgate, R. *J. Biol. Chem.* **2004**, *279*, 32932.
(145) Yang, W. *Curr. Opin. Struct. Biol.* **2003**, *13*, 23.
(146) Purohit, V.; Grindley, N. D.; Joyce, C. M. *Biochemistry* **2003**, *42*, 10200.
(147) Fidalgo da Silva, E.; Mandal, S. S.; Reha-Krantz, L. J. *J. Biol. Chem.* **2002**, *277*, 40640.
(148) Werneburg, B. G.; Ahn, J.; Zhong, X.; Hondal, R. J.; Kraynov, V. S.; Tsai, M. D. *Biochemistry* **1996**, *35*, 7041.
(149) Roychoudhury, R. *J. Biol. Chem.* **1972**, *247*, 3910.
(150) Nick McElhinny, S. A.; Snowden, C. M.; McCarville, J.; Ramsden, D. A. *Mol. Cell. Biol.* **2000**, *20*, 2996.
(151) Arzumanov, A. A.; Victorova, L. S.; Jasko, M. V.; Yesipov, D. S.; Krayevsky, A. A. *Nucleic Acids Res.* **2000**, *28*, 1276.
(152) Tu, C. P.; Cohen, S. N. *Gene* **1980**, *10*, 177.
(153) Ono, K. *Biochim. Biophys. Acta* **1990**, *1049*, 15.
(154) Semizarov, D. G.; Arzumanov, A. A.; Dyatkina, N. B.; Meyer, A.; Vichier-Guerre, S.; Gosselin, G.; Rayner, B.; Imbach, J. L.; Krayevsky, A. A. *J. Biol. Chem.* **1997**, *272*, 9556.

(155) Krayevsky, A. A.; Victorova, L. S.; Arzumanov, A. A.; Jasko, M. V. *Pharmacol. Ther*. **2000**, *85*, 165.
(156) Koc, Y.; Urbano, A. G.; Sweeney, E. B.; McCaffrey, R. *Leukemia* **1996**, *10*, 1019.
(157) McCaffrey, R.; Harrison, T. A.; Parkman, R.; Baltimore, D. *N. Engl. J. Med.* **1975**, *292*, 775.
(158) Ratliff, R. L.; Hoard, D. E.; Ott, D. G.; Hayes, F. N. *Biochemistry* **1967**, *6*, 851.
(159) Kumar, A.; Tchen, P.; Roullet, F.; Cohen, J. *Anal. Biochem*. **1988**, *169*, 376.
(160) Anderson, R. S.; Bollum, F. J.; Beattie, K. L. *Nucleic Acids Res.* **1999**, *27*, 3190.
(161) Stoner, G. D.; Shimkin, M. B.; Troxell, M. C.; Thompson, T. L.; Terry, L. S. *Cancer Res.* **1976**, *36*, 1744.
(162) Roitt, I. M.; Brostoff, J.; Male, D. K. *Immunology*, 6th ed.; Mosby: Edinburgh, New York, 2001.
(163) Janeway, C. A., Jr. *Immunobiology: the immune system in health and disease*, 4th ed.; Current Biology Publications: London, 1999.
(164) Sadofsky, M. J. *Nucleic Acids Res.* **2001**, *29*, 1399.
(165) Makalowski, W. *Acta Biochim. Pol.* **2001**, *48*, 587.
(166) Foote, J.; Eisen, H. N. *Proc. Natl. Acad. Sci. U.S.A.* **1995**, *92*, 1254.
(167) Houk, K. N.; Leach, A. G.; Kim, S. P.; Zhang, X. *Angew. Chem., Int. Ed.* **2003**, *42*, 4872.
(168) Ward, E. S.; Gussow, D.; Griffiths, A. D.; Jones, P. T.; Winter, G. *Nature* **1989**, *341*, 544.
(169) Imai, K.; Slupphaug, G.; Lee, W. I.; Revy, P.; Nonoyama, S.; Catalan, N.; Yel, L.; Forveille, M.; Kavli, B.; Krokan, H. E.; Ochs, H. D.; Fischer, A.; Durandy, A. *Nat. Immunol*. **2003**, *4*, 1023.
(170) Revy, P.; Muto, T.; Levy, Y.; Geissmann, F.; Plebani, A.; Sanal, O.; Catalan, N.; Forveille, M.; Dufourcq-Labelouse, R.; Gennery, A.; Tezcan, I.; Ersoy, F.; Kayserili, H.; Ugazio, A. G.; Brousse, N.; Muramatsu, M.; Notarangelo, L. D.; Kinoshita, K.; Honjo, T.; Fischer, A.; Durandy, A. *Cell* **2000**, *102*, 565.
(171) Stavnezer, J.; Amemiya, C. T. *Semin. Immunol.* **2004**, *16*, 257.
(172) Diamant, E.; Melamed, D. *Autoimmun. Rev*. **2004**, *3*, 464.
(173) Chaudhuri, J.; Alt, F. W. *Nat. Rev. Immunol*. **2004**, *4*, 541.
(174) Yu, K.; Lieber, M. R. *DNA Repair (Amsterdam)* **2003**, *2*, 1163.
(175) Fenton, J. A.; Pratt, G.; Rawstron, A. C.; Morgan, G. J. *Hematol. Oncol.* **2002**, *20*, 75.
(176) Roth, D. B. *Nat. Rev. Immunol*. **2003**, *3*, 656.
(177) Bailin, T.; Mo, X.; Sadofsky, M. J. *Mol. Cell. Biol.* **1999**, *19*, 4664.
(178) Tonegawa, S. *Nature* **1983**, *302*, 575.
(179) McBlane, J. F.; van Gent, D. C.; Ramsden, D. A.; Romeo, C.; Cuomo, C. A.; Gellert, M.; Oettinger, M. A. *Cell* **1995**, *83*, 387.
(180) Fugmann, S. D.; Villey, I. J.; Ptaszek, L. M.; Schatz, D. G. *Mol. Cell* **2000**, *5*, 97.
(181) Kim, D. R.; Dai, Y.; Mundy, C. L.; Yang, W.; Oettinger, M. A. *Genes Dev.* **1999**, *13*, 3070.
(182) Landree, M. A.; Wibbenmeyer, J. A.; Roth, D. B. *Genes Dev.* **1999**, *13*, 3059.
(183) Bassing, C. H.; Swat, W.; Alt, F. W. *Cell* **2002**, *109* (Suppl), S45.
(184) Frank, K. M.; Sekiguchi, J. M.; Seidl, K. J.; Swat, W.; Rathbun, G. A.; Cheng, H. L.; Davidson, L.; Kangaloo, L.; Alt, F. W. *Nature* **1998**, *396*, 173.
(185) Fujita, K.; Shimazaki, N.; Ohta, Y.; Kubota, T.; Ibe, S.; Toji, S.; Tamai, K.; Fujisaki, S.; Hayano, T.; Koiwai, O. *Genes Cells* **2003**, *8*, 559.
(186) Grawunder, U.; Lieber, M. R. *Nucleic Acids Res.* **1997**, *25*, 1375.
(187) Kepler, T. B.; Borrero, M.; Rugerio, B.; McCray, S. K.; Clarke, S. H. *J. Immunol*. **1996**, *157*, 4451.
(188) Ma, Y.; Pannicke, U.; Schwarz, K.; Lieber, M. R. *Cell* **2002**, *108*, 781.
(189) Lafaille, J. J.; DeCloux, A.; Bonneville, M.; Takagaki, Y.; Tonegawa, S. *Cell* **1989**, *59*, 859.
(190) McCormack, W. T.; Tjoelker, L. W.; Carlson, L. M.; Petryniak, B.; Barth, C. F.; Humphries, E. H.; Thompson, C. B. *Cell* **1989**, *56*, 785.
(191) Purugganan, M. M.; Shah, S.; Kearney, J. F.; Roth, D. B. *Nucleic Acids Res.* **2001**, *29*, 1638.
(192) Collis, S. J.; DeWeese, T. L.; Jeggo, P. A.; Parker, A. R. *Oncogene* **2005**, *24*, 949.
(193) Jin, S.; Weaver, D. T. *EMBO J.* **1997**, *16*, 6874.
(194) Peralta-Zaragoza, O.; Recillas-Targa, F.; Madrid-Marina, V. *Immunology* **2004**, *111*, 195.
(195) Yamashita, N.; Shimazaki, N.; Ibe, S.; Kaneko, R.; Tanabe, A.; Toyomoto, T.; Fujita, K.; Hasegawa, T.; Toji, S.; Tamai, K.; Yamamoto, H.; Koiwai, O. *Genes Cells* **2001**, *6*, 641.
(196) Fujisaki, S.; Sato, A.; Toyomoto, T.; Hayano, T.; Sugai, M.; Kubota, T.; Koiwai, O. *Genes Cells* **2006**, *11*, 47.
(197) Ibe, S.; Fujita, K.; Toyomoto, T.; Shimazaki, N.; Kaneko, R.; Tanabe, A.; Takebe, I.; Kuroda, S.; Kobayashi, T.; Toji, S.; Tamai, K.; Yamamoto, H.; Koiwai, O. *Genes Cells* **2001**, *6*, 815.
(198) Elias, L.; Longmire, J.; Wood, A.; Ratliff, R. *Biochem. Biophys. Res. Commun*. **1982**, *106*, 458.
(199) Trubiani, O.; Bollum, F. J.; Di Primio, R. *FEBS Lett.* **1995**, *374*, 367.
(200) Sandor, Z.; Calicchio, M. L.; Sargent, R. G.; Roth, D. B.; Wilson, J. H. *Nucleic Acids Res.* **2004**, *32*, 1866.
(201) Rajewsky, K.; Forster, I.; Cumano, A. *Science* **1987**, *238*, 1088.
(202) MacLennan, I. C. *Annu. Rev. Immunol*. **1994**, *12*, 117.
(203) Weiss, S.; Wu, G. E. *EMBO J*. **1987**, *6*, 927.
(204) Selsing, E.; Storb, U. *Cell* **1981**, *25*, 47.
(205) Honjo, T.; Nagaoka, H.; Shinkura, R.; Muramatsu, M. *Nat. Immunol*. **2005**, *6*, 655.
(206) Neuberger, M. S.; Di Noia, J. M.; Beale, R. C.; Williams, G. T.; Yang, Z.; Rada, C. *Nat. Rev. Immunol*. **2005**, *5*, 171.
(207) Muramatsu, M.; Sankaranand, V. S.; Anant, S.; Sugai, M.; Kinoshita, K.; Davidson, N. O.; Honjo, T. *J. Biol. Chem*. **1999**, *274*, 18470.
(208) Muramatsu, M.; Kinoshita, K.; Fagarasan, S.; Yamada, S.; Shinkai, Y.; Honjo, T. *Cell* **2000**, *102*, 553.
(209) Pham, P.; Bransteitter, R.; Petruska, J.; Goodman, M. F. *Nature* **2003**, *424*, 103.
(210) Seki, M.; Gearhart, P. J.; Wood, R. D. *EMBO Rep.* **2005**, *6*, 1143.
(211) Barreto, V. M.; Ramiro, A. R.; Nussenzweig, M. C. *Trends Immunol.* **2005**, *26*, 90.
(212) Texido, G.; Jacobs, H.; Meiering, M.; Kuhn, R.; Roes, J.; Muller, W.; Gilfillan, S.; Fujiwara, H.; Kikutani, H.; Yoshida, N.; Amakawa, R.; Benoist, C.; Mathis, D.; Kishimoto, T.; Mak, T. W.; Rajewsky, K. *Eur J. Immunol*. **1996**, *26*, 1966.
(213) Poltoratsky, V.; Woo, C. J.; Tippin, B.; Martin, A.; Goodman, M. F.; Scharff, M. D. *Proc. Natl. Acad. Sci. U.S.A.* **2001**, *98*, 7976.
(214) Kawamura, K.; Bahar, R.; Seimiya, M.; Chiyo, M.; Wada, A.; Okada, S.; Hatano, M.; Tokuhisa, T.; Kimura, H.; Watanabe, S.; Honda, I.; Sakiyama, S.; Tagawa, M.; J, O. W. *Int. J. Cancer* **2004**, *109*, 9.
(215) Shima, N.; Munroe, R. J.; Schimenti, J. C. *Mol. Cell. Biol.* **2004**, *24*, 10381.
(216) Zan, H.; Shima, N.; Xu, Z.; Al-Qahtani, A.; Evinger Iii, A. J.; Zhong, Y.; Schimenti, J. C.; Casali, P. *EMBO J*. **2005**, *24*, 3757.
(217) Zan, H.; Komori, A.; Li, Z.; Cerutti, A.; Schaffer, A.; Flajnik, M. F.; Diaz, M.; Casali, P. *Immunity* **2001**, *14*, 643.
(218) Diaz, M.; Verkoczy, L. K.; Flajnik, M. F.; Klinman, N. R. *J. Immunol*. **2001**, *167*, 327.
(219) Frank, E. G.; Tissier, A.; McDonald, J. P.; Rapic-Otrin, V.; Zeng, X.; Gearhart, P. J.; Woodgate, R. *EMBO J*. **2001**, *20*, 2914.
(220) Smith, D. S.; Creadon, G.; Jena, P. K.; Portanova, J. P.; Kotzin, B. L.; Wysocki, L. J. *J. Immunol*. **1996**, *156*, 2642.
(221) Foster, S. J.; Dorner, T.; Lipsky, P. E. *Eur J. Immunol*. **1999**, *29*, 3122.
(222) Matsuda, T.; Bebenek, K.; Masutani, C.; Hanaoka, F.; Kunkel, T. A. *Nature* **2000**, *404*, 1011.
(223) Bebenek, K.; Matsuda, T.; Masutani, C.; Hanaoka, F.; Kunkel, T. A. *J. Biol. Chem*. **2001**, *276*, 2317.
(224) Johnson, R. E.; Washington, M. T.; Prakash, S.; Prakash, L. *J. Biol. Chem*. **2000**, *275*, 7447.
(225) Zhang, Y.; Yuan, F.; Wu, X.; Rechkoblit, O.; Taylor, J. S.; Geacintov, N. E.; Wang, Z. *Nucleic Acids Res.* **2000**, *28*, 4717.
(226) Washington, M. T.; Johnson, R. E.; Prakash, S.; Prakash, L. *J. Biol. Chem*. **2001**, *276*, 2263.
(227) Washington, M. T.; Johnson, R. E.; Prakash, S.; Prakash, L. *J. Biol. Chem*. **1999**, *274*, 36835.
(228) Yuan, F.; Zhang, Y.; Rajpal, D. K.; Wu, X.; Guo, D.; Wang, M.; Taylor, J. S.; Wang, Z. *J. Biol. Chem.* **2000**, *275*, 8233.
(229) Zeng, X.; Winter, D. B.; Kasmer, C.; Kraemer, K. H.; Lehmann, A. R.; Gearhart, P. J. *Nat. Immunol.* **2001**, *2*, 537.
(230) Hyden, H.; Egyhazi, E. *Proc. Natl. Acad. Sci. U.S.A.* **1962**, *48*, 1366.
(231) Albright, T. D.; Kandel, E. R.; Posner, M. I. *Curr. Opin. Neurobiol*. **2000**, *10*, 612.
(232) Chen, C.; Tonegawa, S. *Annu. Rev. Neurosci*. **1997**, *20*, 157.
(233) Elgersma, Y.; Silva, A. J. *Curr. Opin. Neurobiol.* **1999**, *9*, 209.
(234) Milner, B.; Squire, L. R.; Kandel, E. R. *Neuron* **1998**, *20*, 445.
(235) Pena De Ortiz, S.; Colon, M.; Carrasquillo, Y.; Padilla, B.; Arshavsky, Y. I. *Neuroreport* **2003**, *14*, 1141.
(236) Viola, M. V.; Cole, M. L.; Norton, P. *J. Neurochem.* **1976**, *27*, 1157.
(237) Christian, K. M.; Thompson, R. F. *Learn. Mem.* **2003**, *10*, 427.
(238) Barnes, D. E.; Stamp, G.; Rosewell, I.; Denzel, A.; Lindahl, T. *Curr. Biol.* **1998**, *8*, 1395.
(239) Gao, Y.; Chaudhuri, J.; Zhu, C.; Davidson, L.; Weaver, D. T.; Alt, F. W. *Immunity* **1998**, *9*, 367.
(240) Frank, K. M.; Sharpless, N. E.; Gao, Y.; Sekiguchi, J. M.; Ferguson, D. O.; Zhu, C.; Manis, J. P.; Horner, J.; DePinho, R. A.; Alt, F. W. *Mol. Cell* **2000**, *5*, 993.
(241) Gu, Y.; Sekiguchi, J.; Gao, Y.; Dikkes, P.; Frank, K.; Ferguson, D.; Hasty, P.; Chun, J.; Alt, F. W. *Proc. Natl. Acad. Sci. U.S.A.* **2000**, *97*, 2668.

(242) Lee, Y.; Barnes, D. E.; Lindahl, T.; McKinnon, P. J. *Genes Dev.* **2000**, *14*, 2576.
(243) Feng, R.; Rampon, C.; Tang, Y. P.; Shrom, D.; Jin, J.; Kyin, M.; Sopher, B.; Miller, M. W.; Ware, C. B.; Martin, G. M.; Kim, S. H.; Langdon, R. B.; Sisodia, S. S.; Tsien, J. Z. *Neuron* **2001**, *32*, 911.
(244) Finney, L. A.; O'Halloran, T. V. *Science* **2003**, *300*, 931.
(245) Fischmann, T. O.; Bentley, G. A.; Bhat, T. N.; Boulot, G.; Mariuzza, R. A.; Phillips, S. E.; Tello, D.; Poljak, R. J. *J. Biol. Chem.* **1991**, *266*, 12915.

CR040445W