1. Introduction

DNA polymerases play essential roles in replication, repair, and recombination of nucleic acid. During each of these biological processes, the polymerase extends a primer using a DNA (or RNA template) to guide each incorporation event. Even during the bypass of lethal forms of DNA damage, the presence of a templating strand is absolutely essential for polymerase activity. However, the requirement for using a template is not universal as there exists a unique enzyme, denoted as terminal deoxynucleotidyl transferase (TdT), that possesses the unusual ability to incorporate nucleotides into a growing primer without a DNA or RNA template as a guide to direct each incorporation event. There is, however, one unique DNA polymerase designated terminal deoxynucleotidyl transferase that performs DNA synthesis using only single-stranded DNA as the nucleic acid substrate. In this chapter, we review the biological role of this enigmatic DNA polymerase and the biochemical mechanism for its ability to perform DNA synthesis in the absence of a templating strand. We compare and contrast the molecular events for template-independent DNA synthesis catalyzed by terminal deoxynucleotidyl transferase with other well-characterized DNA polymerases that perform template-dependent synthesis. This includes a quantitative inspection of how terminal deoxynucleotidyl transferase binds DNA and dNTP substrates, the possible involvement of a conformational change that precedes phosphoryl transfer, and kinetic steps that are associated with the release of products. These enzymatic steps are discussed within the context of the available structures of terminal deoxynucleotidyl transferase in the presence of DNA or nucleotide substrate. In addition, we discuss the ability of proteins involved in replication and recombination to regulate the activity of the terminal deoxynucleotidyl transferase. Finally, the biomedical role of this specialized DNA polymerase is discussed focusing on its involvement in cancer development and its use in biomedical applications such as labeling DNA for detecting apoptosis.
of the polymerase that include the utilization of various metal ion cofactors, nucleic acids, and nucleotide substrates. Finally, the biomedical importance of TdT will be discussed with emphasis on its potential role in the development of certain forms of leukemia as well as its utilization as a biochemical marker for apoptosis.

2. The role of TdT in V(D)J recombination

Most organisms possess sophisticated defense mechanisms to protect them against the invasion of foreign agents such as viruses, bacteria, and parasites. Simple prokaryotes use a complementary system involving DNA methylation of the host genome and endonuclease degradation of foreign genomic material to differentiate self from non-self [18]. Eukaryotes have developed more sophisticated systems to thwart off the invasion of foreign substances. Indeed, the mammalian immune system is arguably one of the most intricate and ingenious methods for actively seeking out and killing a wide variety of invaders.

The vertebrate immune system is divided into two subcategories, the innate and adaptive immune systems, that differ in their specificity. The innate immune system is generally considered to be less specific due to the promiscuous ability of the immune receptors to recognize a limited number of molecules that are common features to many infectious agents including polysaccharides, peptidoglycans, non-methylated CpG DNA, and double-stranded RNA [19–22]. This promiscuous activity allows the innate immune system to act as the first line of defense against infection by rapidly recognizing and responding to pathogens. If the defensive line of the innate system is breached, then a more specific and highly specialized offense system, the adaptive immune response, is mobilized to its full potential.

Adaptive immunity came into existence in vertebrates roughly 500 million years ago [23]. The cells of the adaptive immune system, namely T- and B cells, have a diverse repertoire of antigen receptors and antibodies that can recognize any antigen encountered throughout life [24]. After the adaptive immune cell receptors bind an antigen, they mount a rapid and robust protective response by a dramatic expansion in the number of pathogen-specific T cells [25–28]. Over the course of 1 week, thousands of clones are produced that possess effector functions [29–31]. Approximately 95% of these activated T cells undergo apoptosis [30,32]. However, a stable population of long-lived T cells resides in the lymphoid and non-lymphoid tissues [33,34] and patrol for these previously encountered pathogens. The immunological memory displayed by the adaptive immune system provides the vertebrate host with long-lasting protection against subsequent infection. For example, most individuals remain immune to measles for up to 75 years once exposed to an attenuated form of measles virus [35].

At the molecular level, the cells of the immune system have developed a strategy to increase acquired immunity against subsequent biological assaults [36](Fig. 2). This process, commonly known as V(D)J recombination, plays an essential role in abrogating these antigens. Rearrangement of the variable (V), diversity (D) and joining (J) gene segments creates versatility to a competent immune system by generating a diverse repertoire of antigen receptors with unique antibody specificities [37]. This transaction of breaking, rearranging, and rejoining of the V, D, and J regions of the germline immunoglobulin genes requires the collaborative efforts of the three distinct enzyme activities that include nucleases, polymerases, and ligases. Of the three major types of enzymatic activities, our understanding of how specific DNA polymerases function during V(D)J recombination is not yet firmly established. However, crucial information for understanding the role of specific polymerases in V(D)J recombination has started to emerge. The relative functions of the various members of the X family of DNA polymerases (TdT, pol μ, and pol λ) during the processing of DNA in V(D)J recombination are distinct [38] and non-overlapping [9] in vivo. Ramsden et al. have indicated that a “gradient” of weak to strong terminal deoxynucleotidyl transferase activity defines the distinct roles of pol λ, pol μ, and TdT in non-homologous end-joining (NHEJ), respectively [38]. Moreover, Bertocci et al. have shown that pol μ participates exclusively in light chain and not in heavy chain gene rearrangement [9,39]. In contrast, pol λ is reported to be recruited only in the heavy chain junctions during V(D)J recombination and precedes the action of TdT [9], which is primarily involved in the random addition of nucleotides to unpaired primer termini [38]. While pol μ
and pol λ play important roles in the immune system, the role of TdT in V(D)J rearrangement is showcased in this review. A more thorough description of pol β, pol μ and pol λ is provided in a chapter by Joann Sweasy and colleagues in this Special Issue.

Fig. 3A illustrates the crucial steps of V(D)J recombination during the RAG cleavage phase. This process is initiated by introducing a double-strand break (DSB) at the edge of the selected gene segment by the RAG-1 and RAG-2 (Recombination-Activating Genes) proteins. These proteins selectively bind to specific and highly conserved recombination signal sequences (RSS) that contain heptamer and nonamer elements separated by a spacer region [40,41]. There are either 12 bp or 23 bp spacers between the heptamer and nonamer elements of the RSS commonly denoted 12-RSS or 23-RSS [40–43], respectively. Proper recombination demands that one 12-RSS and one 23-RSS be present for efficient cleavage in vivo [40,41,44], and this phenomenon is referred to as the "12/23-rule". After recognition of complementary RSS, the RAG complex introduces a nick between the D and J coding segments and the adjoining recombination signal sequence. The RAG complex also mediates the formation of hairpins at each coding end by using the 3′-OH moiety at each nick as the nucleophile.

Double-strand DNA breaks introduced during the RAG cleavage phase are repaired during the non-homologous end-joining phase (Fig. 3B). The nuclease, Artemis–DNA–PKcs, can trim the 5′ and 3′ overhangs [45] and open the hairpins at the coding ends generating palindromic nucleotide sequences (P-nucleotides). At this point, TdT randomly adds nucleotides to available 3′-OH ends increasing the variability of the non-templated nucleotide (N) region of the recombined gene segments [46]. The details of this process are provided in Fig. 4. While in vitro studies have shown that TdT can incorporate all four natural nucleotides onto single-stranded DNA (vide infra), there is a definitive bias for the incorporation of dGMP and dCMP versus dAMP or dTMP observed in vivo [8,47–50]. This preference may offer a plausible explanation for the high G+C content of the Ig and TCR N regions [48,49]. In addition, the average N-nucleotide segment length created in vivo is only 2–5 bp per coding joint [47]. This length appears optimal for allowing the DNA strands to undergo microhomology alignment using Watson–Crick base pair recognition patterns. Unpaired nucleotides are trimmed by an exonuclease such as the Artemis–DNA–PKcs complex [45]. The gaps are ultimately filled by template-dependent DNA polymerases, and the ligation of the coding ends is carried out by the XRCC4–DNA ligase IV complex [45].

The role of TdT in generating immunological diversity is to catalyze the random addition of small numbers of nucleotides at the N regions. This activity has been speculated by virtue of its catalytic properties in addition to its localization to primary lymphoid tissues such as thymus and bone marrow [15,51,52]. A wealth of information to support this hypothesis exists and include the correlation between the existence of N regions with TdT expression [13] or mRNA expression [53] and with the inclusion of the TdT gene into a cell line that continuously rearranges immunoglobulin genes but lacks the enzyme [7,8]. In addition, TdT-knockout mice and TdT-deficient lymphocytes [6,54] have also been used to validate the role of TdT in N-nucleotide addition during the V(D)J recombination process. As expected, TdT-knockout mice show a ten-fold reduction of T cell receptor (TCR) diversity compared to the wild-type mice and provides evidence that the addition of N-nucleotides by TdT is of paramount value to establishing a combinatorially diverse antigen receptor repertoire [54].

3. Biochemical properties of TdT

3.1. Isoforms of TdT

Baltimore and colleagues were amongst the first to show that the occurrence of short nucleotide insertions in the combinatorial junctions of DNA segments during the V(D)J recombination process correlates with the presence of TdT [13,55]. Over the course of several years, different mRNA splicing variants of the enzyme were identified and studied extensively in mice, bovines, and humans. Thus far, two splice variants of TdT have been observed in mice while bovines and humans each have three.

The two identified splice variants of TdT in mice are TdTS, a short form consisting of 509 amino acids [56], and TdTL, a long form with 529 residues [57]. Both isoforms share all domains that are essential for the binding of nucleotides, DNA, and metal ions [58,59]. However, the 20 amino acid difference between the two variants reportedly bestows distinct enzymatic behavior that remains controversial [60–62]. For example, Kearney et al. have shown that TdTS catalyzes the unbiased and non-templated addition of nucleotides to the combinatorial junctions of antigen receptor genes [60,62,63], while TdTL possesses 3′→5′ exonuclease activity that can delete nucleotides at the coding ends of the Ig and TCR gene segments [61]. In contrast, Papanicolaou et al. have argued that TdTL does not possess exonuclease activity and that it exhibits nearly identical enzymatic activity to TdTS in vitro [62,64]. This discrepancy may be attributed to the lower stability of TdTL compared to TdTS both in vitro [64] as well as in transfected cells [60]. In addition, TdTL can reportedly modulate the amount of nucleotide addition by TdTS [11,61]. While the mechanism for this regulation remains undefined, these data suggest that the two TdT murine isoforms function to balance the diversification of antigen repertoire assembly to preserve the integrity of the variable region of antigen receptors [61].

Human (h) and bovine (b) TdT isoforms are slightly more complicated as each have three alternative splice variants designated as TdTS (short), TdTL1 (long) and TdTL2 (long) [65–67]. hTdTL1 and hTdTL2 both localize in the nucleus [67]. However, hTdTL2 is expressed more abundantly in normal small lymphocytes compared to hTdTL1 which is readily detected in transformed lymphoid cell lines [67]. Both long isoforms of human TdT possess 3′→5′ exonuclease activity for nucleotide removal whereas the short isoforms perform nucleotide
Fig. 3. A simplified model for (A) the RAG cleavage phase generating double-strand breaks and (B) DNA repair through non-homologous end-joining pathway during DJ gene segment assembly of the V(D)J recombination mechanism.
elongation of the coding ends during V(D)J recombination [65,67]. Overexpression of *h*TdT5 or *h*TdT2 independently reduces the efficiency of V(D)J recombination greatly [67]. However, the simultaneous overexpression of *h*TdT5 and *h*TdT2 restores recombination frequencies back to normal levels. These observations suggest the possibility of a strong evolutionary selection for co-expressing one transferase (*h*TdT5) and one exonuclease (*h*TdT2) to regulate the length of single-stranded regions generated during NHEJ. In addition, there is evidence for prohibiting the coexpression of two exonucleases (*h*TdT1 and *h*TdT2) via regulation at the level of mRNA alternative
splicing during V(D)J recombination [67]. In fact, the presence of all three human TdT variants during antigen receptor gene rearrangement drastically diminishes recombination frequency. These results suggest that hTdT1 may serve to regulate hTdT2 or hTdT3 activity [67].

3.2. Sources and purification of TdT

The calf thymus glands are an abundant source of TdT. In fact, TdT was first purified to apparent homogeneity from calf thymus cell lysate in 1971 [68]. At first, the isolated protein was reported as a proteolyzed form containing two polypeptides originally recognized as a heterodimer of α and β subunits [68,69]. However, subsequent studies established that indigenous TdT is truly a monomeric protein with a molecular weight of 58,000 to 60,000 Da [70–72]. Unfortunately, obtaining the enzyme in its relatively pure, intact, and active form is extremely laborious and difficult due to proteolysis. Copious amount of the enzyme can be purified from cultured cell lines propagated from patients with acute lymphoblastic leukemia [73]. However, this approach is too expensive and impractical for routine mass production of enzyme. Attempts to overexpress TdT in bacteria systems have generally failed due to three prominent factors: (1) mismatches in the codon frequencies and tRNA pools between E. coli and eukaryotes, (2) low solubility of the protein in these systems, and (3) lower levels of enzyme activity [59,74,75]. These complications can be alleviated by overexpressing a rare argU tRNA in the E. coli system and growing cultures at 15 °C [64,76] to boost the production of soluble and active forms of the enzyme. Purification of TdT to apparent homogeneity is accomplished by column chromatography which is facilitated by the presence of a hexa-histidine tag attached to the N-terminus that does not affect its activity [76]. Recombinant human TdT has also been overexpressed in the baculovirus expression system and purified to homogeneity in one step from Trichoplusia ni larvae using a monoclonal antibody affinity column [59,74]. This last advancement has played an instrumental role in generating significant amounts of protein required for structural characterization of TdT.

3.3. Primer requirement

The unique activity of TdT to incorporate nucleotides in a template-independent manner was reported simultaneously by Karkow and Kamen [77] and Bollum [78] using calf thymus extracts. The template-independent activity of TdT was distinguished from that of other template-dependent DNA polymerases by comparing levels of the incorporation of radiolabeled nucleotides using single-stranded versus double-stranded DNA [1]. Activity measured using single-stranded DNA is attributed only to that of TdT [2]. Subsequent biochemical studies confirmed that TdT requires a single-stranded initiator that is at least three nucleotides long with a free 5'-phosphate end and a free 3'-hydroxyl end for extension [78]. The replication of homopolymers by TdT requires an initiator chain of more than six nucleotides for poly(dA) and more than five nucleotides for poly(dT) [79]. The presence of a ribonucleotide 5'-monophosphate (rNMP) at the 3'-end of the primer does not inhibit the enzyme–primer complex formation [80]. However, further elongation occurs at a slower rate and the addition of more than two ribonucleotides to the single-stranded initiator does inhibit activity [80]. TdT is also unique for its ability to perform de novo synthesis of polynucleotides ranging in size from 2- to 15-mers when provided with dNTPs in the absence of a primer [81]. These DNA fragments are hypothesized to act as signals for DNA repair or recombination machinery [81]. There may indeed be credence to this provocative hypothesis since small RNAs, designated as microRNA, influence the activity of many biological pathways [82]. It will prove interesting to firmly test this hypothesis to evaluate if de novo DNA synthesis does indeed occur in vivo and that small DNA fragments can regulate biological processes such as DNA repair and recombination.

3.4. Metal utilization

All DNA polymerases require the presence of a divalent metal ion to catalyze the phosphorol transfer reaction associated with nucleotide incorporation. The general catalytic mechanism for phosphorl transfer utilized by DNA polymerases is provided in Fig. 5. In this

Fig. 5. The catalytic mechanism model for the nucleotidyl transfer reaction catalyzed by terminal deoxynucleotidyl transferase. See text for details.
mechanism, an aspartate residue near the deoxyribose sugar of the incoming dNTP serves as the general base to abstract the proton from the 3′-OH to generate a more reactive nucleophile. The electron rich 3′-oxygen then attacks the α-phosphate creating a trigonal-bipyramidal metal ion coordinated transition state. This step results in the inversion of the α-phosphate stereochemistry [83,84] and the release of the pyrophosphate leaving group coordinated to Zn2+. In addition, each metal ion has different effects on the kinetics of catalytic polymerization efficiency of the pyrimidines, dCTP and dTTP [90]. Zn2+ behaves as a unique positive effector for TdT since reaction rates with Mg2+ increase in the presence of divalent metal ions is ranked in the following order: Mn2+ > Zn2+ > Co2+ > Mg2+ [73]. In addition, each metal ion has different effects on the kinetics of nucleotide incorporation. For example, Mg2+ facilitates the preferential utilization of dGTP and dATP whereas Co2+ increases the catalytic polymerization efficiency of the pyrimidines, dCTP and dTTP [90]. Mn2+ behaves as a unique positive effector for TdT since reaction rates with Mg2+ are stimulated by the addition of micromolar quantities of Zn2+ [90]. This enhancement may reflect the ability of Zn2+ to induce conformational changes in TdT that yields higher catalytic efficiencies [90]. Polymerization rates are lower in the presence of Mn2+ compared to Mg2+, suggesting that Mn2+ does not support the reaction as efficiently as Mg2+ [73].

4. Mechanism of template-independent polymerization

4.1. When does replication without a template occur?

A DNA polymerase is classically defined as an enzyme that catalyzes template-dependent addition of mononucleotides into a growing primer in the presence of four deoxynucleotides (dNTPs) [91]. Nevertheless, many archaeal, bacterial, viral, and eukaryotic DNA polymerases also catalyze non-template nucleotides additions to the 3′-termini of blunt-ended DNA [92–96]. In addition, numerous DNA polymerases have been demonstrated to incorporate nucleotides opposite non-instructional DNA lesions such as abasic sites via translesion DNA synthesis [96–103]. In both of these cases, however, these polymerases require the use of duplex DNA as a substrate for efficient catalysis to occur. In fact, the sole polymerase that can add nucleotides using only a single-stranded DNA initiator as a substrate is the template-independent DNA polymerase, TdT [1,2]. Terminal deoxynucleotidyl transferase activity has also been observed for pol μ [38] and pol λ [104,105], although DNA synthesis is primarily restricted to the use of a DNA template [38,104,105]. In the sections provided below, we compare and contrast the mechanism of nucleotide incorporation catalyzed by TdT with template-dependent DNA polymerases. Particular emphasis is placed on the replication of non-instructional DNA lesions as this represents the most frequent mode of template-independent DNA synthesis catalyzed by these enzymes.

4.2. Order of substrate binding

Theoretically, the binding of DNA and dNTP substrates to any DNA polymerase can be random, sequential, or strictly ordered. Efficient polymerization for a template-dependent polymerase, however, would be optimal through the strictly ordered binding of DNA substrate prior to dNTP since the converse order of dNTP binding prior to DNA would be correct only once out of four opportunities. Indeed, numerous steady-state and pre-steady-state studies have validated that all template-dependent polymerases obey this mechanism (reviewed in [106]).

Not surprisingly, TdT appears to be the lone exception to this rule. The order by which TdT binds DNA and dNTP is indeed random as determined through a series of initial velocity studies performed in the absence and presence of product inhibitors [73]. The double reciprocal plot of 1/rate versus 1/dNTP concentration at several fixed concentrations of single-stranded DNA intersected to the left of the y-axis, consistent with either an ordered or random kinetic mechanism. Inhibition by PPi at varying dNTP concentrations gave rise to a competitive inhibition pattern whereas inhibition by PPi at different concentrations of DNA yielded a mixed inhibition pattern. These results are consistent with rapid equilibrium random mechanism in which TdT forms the catalytic competent ternary complex via binding of dNTP prior to DNA or vice versa. It is currently unclear if the ability to randomly bind substrates plays a physiological role in generating random nucleotides during recombination. However, it is possible that the interactions of TdT with PCNA and Ku70/86 [107–109], proteins involved in replication and recombination, may influence its kinetic mechanism.

4.3. Mechanism of nucleotide selection

An important kinetic step for generating catalytic efficiency and fidelity with template-dependent polymerases occurs during the binding of dNTP to the polymerase:DNA complex. This process is highly influenced by the presence of a templating strand due to hydrogen bonding [110] and steric complementarity [111]. However, since TdT does not rely on a templating strand for polymerization activity, the molecular details regarding nucleotide binding and selection remain elusive.

Several laboratories have demonstrated that TdT displays an unequal bias in the kinetics of nucleotide incorporation. For example, studies from the Coleman laboratory reveal that the Km for dCTP is ~4-fold lower than the Km for dATP (compare Km values of 120 μM versus 540 μM, respectively) [59]. Similar results are reported by the Modak group [112]. More recently, we have demonstrated that the recombinant TdT utilizes dGTP, dCTP, and dTTP much more efficiently than dATP [113]. The preferential use of dGTP and dCTP argues that a biophysical bias for more efficient annealing of single-stranded DNA that are intermediates during NEHEJ. At the molecular level, the preference in nucleotide utilization could reflect favorable hydrogen-bonding interactions between the incoming dNTP and active site amino acids that guide nucleotide binding. This mechanism would resemble that displayed by the yeast Rev1 protein, an error-prone DNA polymerase that preferentially incorporates dCMP via direct interactions with an active site arginine as opposed to interactions with the templating base [114]. The discrimination against the utilization of dATP is also intriguing especially when compared to the preferential incorporation of dAMP by high fidelity DNA polymerases during the replication of non-coding DNA lesions such as abasic sites. Template-dependent DNA polymerases such as the E. coli Klenow fragment [101] and bacteriophage T4 polymerase [101,103] incorporate dAMP opposite an abasic site ~100-fold more efficiently than any of the other three natural nucleotides. The preferential incorporation of dAMP by these high fidelity DNA polymerases is often referred to as the “A-rule” of translesion DNA synthesis [115]. Thus, the preferential usage of dGTP and dCTP by TdT indicates that it does not follow the “A-rule” like the aforementioned well-characterized template-dependent DNA polymerases.

It is also interesting that TdT efficiently incorporates and extends a wide variety of 5-substituted indolyl deoxynucleotides that lack hydrogen-bonding functional groups (Fig. 6). TdT uses non-natural nucleotides such as 5-NITP with identical efficiencies as dGTP, the preferred natural nucleotide [113]. This result is intriguing since 5-NITP has been used as a chemical probe to validate the “A-rule” of translesion DNA synthesis with template-dependent DNA polymerases [101]. In fact, several template-dependent polymerases incorporate 5-NITP opposite an abasic site with remarkably high catalytic efficiencies that approach 10^9 M^–1 s^–1 [116]. Thus, TdT presents an interesting paradox by its ability to discriminate against
inserting dATP yet effectively incorporating non-natural nucleotides that are used as mechanistic probes for the “A-rule” of translesion DNA synthesis.

A naive mechanism to explain the utilization of these non-natural nucleotides is that the primary molecular determinant for nucleotide binding resides simply with interactions of the negatively-charged triphosphate moieties with positively-charged amino acids in the polymerase’s active site. This is unlikely to be the sole source for binding as it cannot account for the significant differences observed for the incorporation of various natural and non-natural nucleotide substrates. For example, 5-NIMP is incorporated and extended 10-fold more efficiently than the closely related analog, 5-AIMP, which is poorly incorporated even at concentrations greater than 500 μM [113].

In addition, the replacement of an active site arginine residue (R336) involved in binding the triphosphate moiety of a dNTP with either glutamine (R336Q) or alanine (R336A) does not completely abolish nucleotide binding. Instead, these amino acid substitutions reduce the binding affinity of dGTP and dATP by only 10-fold [59]. These data indicate that ionic interactions between the active site arginine and the triphosphate group are important but not essential for nucleotide binding.

Fig. 6. Chemical structures of various 5-substituted indolyl nucleotides used to probe the activity of TdT.

In contrast, replacement of Co2+ with Mg2+ gives rise to a smaller elemental effect of ~3 (unpublished results, structural, and spectroscopic studies [93,98,117–121] and the reader is recommended to read the article by Kenneth Johnson (this issue) for further details on this subject. Regardless, this conformational change step is believed to impose discrimination against misinserting an incorrect nucleotide into DNA by altering the geometry of the polymerase’s active site to inhibit efficient phosphoryl transfer [122]. Since TdT uses only single-stranded DNA [12], the need for an obligatory conformational change step to ensure fidelity remains in doubt. Put another way, does the lack of “fidelity” displayed by TdT negate the need for a conformational change step? If so, does phosphoryl transfer then become the rate-limiting step for nucleotide incorporation as similarly reported for certain “error-prone” polymerases that have lower constraints in fidelity [123,124]? Indeed, it is established that with certain DNA polymerases such as pol β, the phosphoryl transfer step as opposed to a conformational change is involved in maintaining fidelity during the nucleotide incorporation step [125].

One way to evaluate the location of the rate-limiting step during polymerization is to measure the effects of thio-substituted nucleotides on the rate of nucleotide incorporation [126]. Since a non-bridging sulfur atom has decreased electronegativity, it is predicted to be less effective than oxygen at stabilizing electron density during the transition state of phosphoryl transfer. As a result, the magnitude of a “thio effect”, defined as the rate of incorporation using α-O-dNTP versus the rate using an identical concentration of α-S-dNTP, can be used as a diagnostic indication of whether or not chemistry is rate-limiting. If chemistry is the rate-limiting step, then a significant thio-elemental effect (>10) is typically observed while smaller thio-elemental effects of ~2 are observed if another kinetic step such as a conformational change preceding phosphoryl transfer is rate-limiting, and thus insensitive to this substitution. We have preliminary data defining the elemental effect for the incorporation of α-S-dGMP by TdT using different metal cofactors. With Co2+ as the metal cofactor, the rate of α-S-dGTP incorporation is 8-fold slower than that measured using an identical concentration of α-O-dGTP (unpublished results, Berdis, A.J.). The apparent elemental effect of 8 suggests that phosphoryl transfer is indeed the rate-limiting step and is consistent with structural data indicating that TdT exists in a “closed” conformation that alleviates the need for a conformational change for catalysis to occur [127]. In contrast, replacement of Co2+ with Mg2+ gives rise to a smaller elemental effect of ~3 (unpublished results, structural, and spectroscopic studies [93,98,117–121] and the reader is recommended to read the article by Kenneth Johnson (this issue) for further details on this subject. Regardless, this conformational change step is believed to impose discrimination against misinserting an incorrect nucleotide into DNA by altering the geometry of the polymerase’s active site to inhibit efficient phosphoryl transfer [122]. Since TdT uses only single-stranded DNA [12], the need for an obligatory conformational change step to ensure fidelity remains in doubt. Put another way, does the lack of “fidelity” displayed by TdT negate the need for a conformational change step? If so, does phosphoryl transfer then become the rate-limiting step for nucleotide incorporation as similarly reported for certain “error-prone” polymerases that have lower constraints in fidelity [123,124]? Indeed, it is established that with certain DNA polymerases such as pol β, the phosphoryl transfer step as opposed to a conformational change is involved in maintaining fidelity during the nucleotide incorporation step [125].

One way to evaluate the location of the rate-limiting step during polymerization is to measure the effects of thio-substituted nucleotides on the rate of nucleotide incorporation [126]. Since a non-bridging sulfur atom has decreased electronegativity, it is predicted to be less effective than oxygen at stabilizing electron density during the transition state of phosphoryl transfer. As a result, the magnitude of a “thio effect”, defined as the rate of incorporation using α-O-dNTP versus the rate using an identical concentration of α-S-dNTP, can be used as a diagnostic indication of whether or not chemistry is rate-limiting. If chemistry is the rate-limiting step, then a significant thio-elemental effect (>10) is typically observed while smaller thio-elemental effects of ~2 are observed if another kinetic step such as a conformational change preceding phosphoryl transfer is rate-limiting, and thus insensitive to this substitution. We have preliminary data defining the elemental effect for the incorporation of α-S-dGMP by TdT using different metal cofactors. With Co2+ as the metal cofactor, the rate of α-S-dGTP incorporation is 8-fold slower than that measured using an identical concentration of α-O-dGTP (unpublished results, Berdis, A.J.). The apparent elemental effect of 8 suggests that phosphoryl transfer is indeed the rate-limiting step and is consistent with structural data indicating that TdT exists in a “closed” conformation that alleviates the need for a conformational change for catalysis to occur [127]. In contrast, replacement of Co2+ with Mg2+ gives rise to a smaller elemental effect of ~3 (unpublished results,
5.1. The X-family of DNA polymerases

TdT is a member of the X-family of DNA polymerases that include DNA polymerase β (pol β), DNA polymerase λ (pol λ), and DNA polymerase μ (pol μ). Pol β removes the 5'-deoxyribose phosphate moiety and catalyzes gap-filling synthesis during base-excision repair. The mechanism of this polymerase is described in a chapter by Joann Berdis, A.J.). The change in the magnitude of the elemental effects suggests that phosphoryl transfer is no longer rate-limiting using Mg\(^{2+}\) as the metal cofactor. One implication of these findings is that metal ions directly influence the mechanism of template-independent polymerization by changing the location of the rate-limiting step. However, controversy exists regarding the use of thio-elemental effect to unambiguously define the rate-limiting step for nucleotide incorporation. Much of this controversy involves potential differences in the transition states between the enzyme-catalyzed and non-enzymatic reaction [128]. Therefore, further kinetic and spectroscopic studies on TdT are needed to further investigate this phenomenon.

5. Structural insights into the mechanism of TdT

5.1. The X-family of DNA polymerases

TdT is a member of the X-family of DNA polymerases that include DNA polymerase β (pol β), DNA polymerase λ (pol λ), and DNA polymerase μ (pol μ). Pol β removes the 5'-deoxyribose phosphate moiety and catalyzes gap-filling synthesis during base-excision repair. The mechanism of this polymerase is described in a chapter by Joann Berdis, A.J.). The change in the magnitude of the elemental effects suggests that phosphoryl transfer is no longer rate-limiting using Mg\(^{2+}\) as the metal cofactor. One implication of these findings is that metal ions directly influence the mechanism of template-independent polymerization by changing the location of the rate-limiting step. However, controversy exists regarding the use of thio-elemental effect to unambiguously define the rate-limiting step for nucleotide incorporation. Much of this controversy involves potential differences in the transition states between the enzyme-catalyzed and non-enzymatic reaction [128]. Therefore, further kinetic and spectroscopic studies on TdT are needed to further investigate this phenomenon.

5.2. Features of TdT common with other X-family polymerases

Sequence alignment of the C-termini of the X-family DNA polymerases reveals that all these enzymes possess the fingers, palm, and thumb subdomains that are similar to those described first for pol β (Fig. 7). TdT, pol λ, and pol μ contain nuclear localization signal motifs as well as breast cancer susceptibility protein BRCA1 C-terminal (BRCT) domains in their N-termini [127]. BRCT domains function to mediate protein/protein and protein/DNA interactions in DNA repair and cell cycle checkpoint pathways that are activated by DNA damage [135]. As described later in this review, the BRCT domain of TdT, pol λ, and pol μ may interact with Ku70/80, a heterodimeric protein involved in recognizing and binding the ends of double-strand DNA breaks formed during V(D)J recombination.

Considering this information, it is quite surprising that members of this family show little primary amino acid sequence identity. For instance, TdT and pol μ share only 42% amino acid identity yet are considered the most closely related members of the X-family of polymerases [136]. Despite this low identity, it is proposed that TdT and pol μ arose from a common ancestor that was a template-dependent polymerase [137]. The strict template-independent activity of TdT appears to be a recent evolutionary event that coincides with the development of V(D)J recombination in mammals.

5.3. Tertiary structure of TdT

All template-dependent DNA polymerases characterized to date by structural methods show an overall molecular architecture that is analogous to a right hand containing thumb, fingers, and palm subdomains (Fig. 8, inset 1) (reviewed in [86]). Although these subdomains work synergistically during the polymerization process, it is easier to describe their functions independently. The palm subdomain is considered to be the heart of the polymerase as this is where the phosphoryl transfer reaction takes place. This subdomain contains at least two carboxylate residues that are highly conserved amongst all DNA polymerases [138] which function to coordinate metal ions that act as Lewis acids. These metal ions lower the activation energy barrier for the phosphoryl transfer reaction by stabilizing the build-up of the negative charge that accumulates on the α-phosphate and the oxyanion of the β- and γ-phosphate leaving group of the incoming dNTP substrate [88]. The function of the fingers subdomain is to coordinate interactions between the templating base and the incoming dNTP. These interactions are designed to align the nucleobases properly prior to catalysis. With template-dependent DNA polymerases such as pol β [87,139,140], the thumb subdomain serves a dual role by positioning duplex DNA for accepting the incoming dNTP as well as for translocation of the polymerase to the next templating base position after catalysis.

![Schematic representations of the different domains found in the four X-family DNA polymerases. Each domain is labeled and colored for clarity. NLS represents the nuclear localization signal motif, and BRCT indicates the BRCA1 carboxy terminus domain.](image)
Fig. 8. Crystal structure of terminal deoxynucleotidyl transferase showing the finger, thumb, palm and index finger (8 kDa) subdomains that work synergistically to catalyze nucleotide incorporation. Inset 1 shows the hand-like morphology of most DNA polymerases. Inset 2 shows a surface model for the ring-like structure of TdT in which the primer strand is located perpendicular to the hole where dNTPs presumably diffuse to enter the active site. The ternary complex structure was prepared using the available binary crystal structures of murine TdT (PDB ID codes 1KEJ [TdT-ddATP] and 1KD1 [TdT-ssDNA]) [116]). MOE (www.chemcomp.com) was used for all structural modeling.

Fig. 9. The active site of TdT (PDB ID code: 1KEJ) as defined by amino acids that exist within 6 Å of the bound nucleotide substrate, ddATP. The incoming nucleotide is shown in ball-stick representation in CPK color scheme. The two cobalt ions are colored as cyan. This figure was prepared using the UCSF Chimera package (http://www.cgl.ucsf.edu/chimera).
The structure of TdT shown in Fig. 8 shows similarities to template-dependent DNA polymerases by the presence of thumb, fingers, and palm subdomains. The catalytic site of TdT also shows the presence of two Co^{2+} ions in the palm domain [127]. This feature is consistent with the proposed “two divalent metal ion” mechanism [88] associated with all polymerases currently identified. Finally, these metal ions are coordinated by the oxygen atoms of the triphosphate moieties of the incoming dNTP as well as by three aspartate residues [127].

Despite these similarities, there are some notable structural variations in TdT that distinguish it from template-dependent polymerases. In particular, TdT possesses a “lariat-like” loop that prevents the enzyme from interacting with a template strand [127]. In addition, TdT contains an 8 kDa domain, also known as the “index finger” domain, which contacts the thumb subdomain to form a hole that may allow dNTPs to diffuse into the enzyme’s active site [127]. This feature defines the general morphology of TdT as a ring-like structure in which the primer lies perpendicularly to the axis of the hole on the palm domain (Fig. 8, inset 2).

Superimposing the structure of TdT (PDB ID code 1KEJ) [127] with the closed form of pol λ (PDB ID code 1XSN) [141] shows that the “lariat-like” loop of 16 amino acids in TdT crystal structure prevents the polymerase from interacting with duplex DNA (Fig. 9). This provides a physical explanation for the unique ability of TdT to catalyze the nucleotidyl transfer reaction in a template-independent fashion. Delarue et al. originally pointed out that TdT appears to be sealed in a closed conformation that is similar to pol β [127].

Fig. 10. Comparing the crystal structures of template-independent TdT enzyme (PDB code: 1KEJ) and template-dependent DNA polymerase, pol λ (PDB code: 1XSN). (A) Superimposed structures of TdT and pol λ. The superimposed structures show an incredibly high degree of similarity between the two polymerases of the family X despite the fact that they catalyze different modes of DNA polymerization (template-independent versus template-dependent). (B) The ribbon structure of TdT (left) with the putative model of primer-template duplex derived from the human pol λ ternary complex (right) to show the “lariat-like” loop (in magenta) that prevents the ability of TdT to accommodate a templating strand. (C) Molecular surface model of TdT (left) with the putative model of primer-template duplex (shown in ribbon) derived from the human pol λ ternary complex (right). The dNTP is shown in ball and stick model in CPK color scheme. The models were prepared using the MOE package (www.chemcomp.com).
suggesting that TdT can add nucleotides without undergoing an obligatory conformational change [127].

Insight into the mechanism of nucleotide selection comes from an examination of the available binary structures of the TdT·ddATP (PDB ID code 1KEJ) and TdT·ssDNA (PDB ID code 1KDH) complexes [127]. An inspection of amino acids that reside within 6 Å of the bound ddATP (Fig. 10) reveals the presence of three positively-charged residues (Lys338, Arg336, and Arg454) that point toward the triphosphate moiety to neutralize the negatively-charged phosphate molecules. In addition, three conserved aspartates (Asp343, Asp345, and Asp434) reside in the catalytic palm subdomain and function to position two Co$^{2+}$ ions for catalysis. The aromatic ring of Trp450 appears to provide favorable pi–pi stacking interaction as it lies parallel to the adenine ring of the incoming nucleotide and resides only 3.6 Å away. The positively-charged ε amino group of Lys403 side chain is only 4 Å away and can provide favorable pi–cation interactions with the adenine nucleobase. Although Arg454 interacts with the triphosphate moiety, it may also participate in pi–cation interactions with the aromatic adenine to provide additional stability. No other side chains are in close proximity with either the 2′ or the 3′ position of the sugar of ddATP. The absence of amino acids that could function as a “steric gate” could account for the rather promiscuous nature of TdT in its ability to incorporate both ribo- and deoxyribonucleotides [142].

6. Regulating the activity of TdT

Since V(D)J recombination is required for a competent immune system, it is imperative that the activities of the enzymes involved in this process be tightly regulated. TdT is no exception as it is regulated by transcriptional control, post-translational modifications, and protein–protein interactions [143].

Expression of TdT is confined to primary lymphoid tissues including thymus and bone marrow [15,51,52]. Transcriptional control is regulated by proteins such as AP-1 [144] as well as through the expression of the RAG genes [145]. The activity of TdT may also be regulated at the post-translational level by phosphorylation. It has been demonstrated that TdT is phosphorylated in lymphoblastoid cells when treated with $[^{32}P]$-phosphate [146]. In addition, recombiant human TdT is phosphorylated in vitro by protein kinase C [147] while calf thymus TdT can be phosphorylated by beef heart cAMP-dependent protein kinase [148]. In the latter case, calf thymus TdT is phosphorylated at multiple sites that correspond to amino acids Ser7 and Thr19 in human TdT [74,146]. While TdT can be phosphorylated at multiple sites, it is unclear as to how phosphorylation regulates its activity in vivo.

Protein–protein interactions between TdT and other DNA binding proteins have been shown to produce both positive and negative effects on the catalytic activity of TdT [108,149]. One group of proteins reported to regulate the activity of TdT are referred to as TdT-interacting factors (Tdif). Tdif1 is a protein that binds to the C-terminus of TdT and increases its polymerase activity by ~4-fold [108]. Tdif2 is another protein that binds TdT through its C-terminus, possibly through interactions with the proline-rich and poly β-like domains. However, the interaction of Tdif2 with TdT decreases the polymerase activity of TdT by ~2-fold [108].

PCNA is another protein that physically interacts with TdT [146]. By encircling DNA, PCNA acts as a general processivity factor for various template-dependent DNA polymerases including pol δ and ε [150]. In fact, PCNA orchestrates nearly every aspect of DNA synthesis ranging from chromosomal DNA replication to DNA repair and recombination [151]. PCNA decreases the polymerase activity of TdT by ~2-fold. This effect on TdT activity is similar to that observed by Tdif2 [146]. One possibility is that PCNA and Tdif2 compete with Tdif1 for binding to the C-terminus of TdT which provides a mechanism to reversibly regulate TdT activity.

Other proteins such as the Ku proteins and template-dependent DNA polymerases may also regulate the activity of TdT. For example, the “N” regions generated by TdT are unusually longer when Ku80 is knocked-out compared to when it is normally expressed [152]. Since Ku80 aids in the recruitment of TdT to the site of V(D)J recombination, these knock-down experiments suggest that Ku80 also regulates the catalytic activity of TdT once it is bound to DNA [61,153]. TdT activity can also be regulated indirectly by other members of the X-family of polymerases. For example, TdT and Pol γ can efficiently compete for the same DNA substrate [152]. It has subsequently been argued that both proteins, if present during V(D)J recombination, could compete for DNA ends such that Pol γ could affect the activity of TdT during “N” region synthesis [132].

7. Biomedical applications of TdT

7.1. TdT as an anti-cancer target

There is mounting clinical evidence that alterations in TdT activity and/or its level of expression play significant roles in cancer initiation, progression, and response to chemotherapy. TdT is overexpressed in B- and T cell acute lymphocytic leukemias (AML) [154–156]. About 90% of patients with AML show variable levels of TdT expression as well as multiple isoforms of TdT in their blast cells [154–156]. While the frequency of TdT overexpression in AML is less (∼20%), it is still significantly higher than levels found in lymphoid malignancies such as chronic lymphocytic leukemia (CLL) [156]. In addition, higher levels of TdT activity correlate with a poor prognosis due to sub-optimal responses to chemotherapy that culminates in reduced survival times. Prognosis and survival studies show that remission rates are two-fold lower in leukemia patients that are TdT-positive compared to TdT-negative patients [157]. These findings have driven attempts to develop selective inhibitors against TdT that could be employed as chemotherapeutic agents against these forms of leukemia. One example is the nucleoside analog cordycepin (3′-deoxy adenosine) (Fig. 11A). This nucleoside analog is cytotoxic against TdT-positive leukemias, especially when used in combination with the adenosine deaminase inhibitor deoxycoformycin [158,159]. The cytotoxic activity of cordycepin correlates with the ability of the triphosphate form to inhibit single-stranded DNA synthesis by TdT in vitro [160]. Since cordycepin lacks a 3′-OH moiety, incorporation of this analog terminates primer extension and generates abortive intermediates along the recombination pathway that may induce apoptosis.

Unfortunately, cordycepin is not widely used in chemotherapy as it causes side effects that are associated with the inhibition of various enzymes involved in nucleos(t)ide metabolism [161]. As a consequence, there is significant effort in developing selectively inhibitors of TdT that do not resemble nucleos(t)ide analogs. Indeed, encouraging progress has been made by the groups of DiSanto and Maga as evident in their recent demonstration that certain arylidiketo hexenoic acids inhibit the catalytic activity of pol λ and TdT without acting as chain-terminators [162]. One analog, designated RDS 2119 (Fig. 11B), displays higher cytotoxicity against a TdT-positive leukemia cell line (Molt4) compared to a cell line derived from cervical cancer (HeLa) [162].

7.2. TdT as a biochemical tool

TdT utilizes a wide variety of nucleotide analogs such as 2′, 3′-dideoxynucleotides [163], p-nitrophenyl triphosphate [164], p-nitrophenyl triphosphate [164], 2′-deoxy-L-ribonucleoside 5′-triphosphates [165], and dinucleoside 5′, 5′-tetraphosphates [166]. The ability of TdT to tolerate bulky modifications to the nucleobase moiety has led to an effective method for in vivo and in vitro labeling of double-strand DNA breaks. One technique called TUNEL (TdT-
mediated dUTP-biotin nick end-labeling) is based on the ability of TdT to efficiently incorporate biotinylated dUMP on to 3'-ends of single-stranded DNA that occur at the sites of DNA breaks. The incorporated biotin dUMP is easily visualized by fluorescently labeled avidin or streptavidin and allows for direct conformation and quantitative measurements of the number and location of DNA breaks. This technology is widely used for the detection of apoptosis, a form of programmed cell death in eukaryotic cells [167]. In addition, TdT can be used as a biocatalyst to label the 3’-termini of synthetic oligonucleotides with a radioactive nucleotide or a variety of fluorescent probes [168]. These labeled primers can then be annealed to a complementary strand and used as radioactive substrates to monitor the activity of enzymes involved in nucleic acid metabolism that include restriction endonucleases, DNA glycosylases, and template-dependent DNA polymerases.

7.3. Conclusions

Amongst all DNA polymerases identified to date, TdT remains the most enigmatic. This remarkable polymerase can truly be called the “black sheep” of the polymerase family since it breaks (or at least bends) nearly every rule established for template-dependent polymerases. Most notably, TdT is defiant in its inability to perform polymerization with duplex DNA while incorporating nucleotides using single-stranded DNA as the substrate. The template-independence of TdT predicts that it should incorporate natural dNMPs with equal efficiencies. Again, TdT bends the rules by displaying an unexplained bias towards incorporating dGMP and dCMP. In fact, TdT even discriminates against the incorporation of dAMP, the preferred nucleotide for template-dependent polymerase during the replication of certain forms of damaged DNA. Finally, TdT is unique in its ability to randomly bind DNA and dNTP to form the catalytically competent ternary complex. It could be argued that TdT must be rebellious in order to perform its unique biological function for generating random coding information during V(D)J recombination.

Although TdT is unique, it does share many features that are common amongst most template-dependent DNA polymerases. From a structural perspective, TdT contains the fingers, thumb, and palm subdomains that are present in all DNA polymerases characterized to date. The most notable difference is the inclusion of the “lariat loop” that blocks the ability of TdT to utilize duplex DNA. Furthermore, TdT can interact with other replicative proteins such as PCNA that function to coordinate polymerase activity during replication, repair, and recombination. Although the functional outcome of these interactions remain poorly understood, future biochemical and cell-based studies will undoubtedly shed more insight into the biological role and molecular mechanism of this fascinating and enigmatic DNA polymerase.

References


L.M. Clark, F.J. Bollum, Multiple roles of divalent cation in the terminal deoxynucleotidyl transferase reaction, J. Biol. Chem. 265 (1990) 17436–17440.


