The Journal of Immunology

Mutational Analysis of Terminal Deoxynucleotidyltransferase- Mediated N-Nucleotide Addition in V(D)J Recombination

Jamie A. E. Repasky, Elizabeth Corbett, Cristian Boboila and David G. Schatz

J Immunol 2004; 172:5478-5488; ;
doi: 10.4049/jimmunol.172.9.5478
http://www.jimmunol.org/content/172/9/5478

The Journal of Immunology is published twice each month by
The American Association of Immunologists, Inc.,
1451 Rockville Pike, Suite 650, Rockville, MD 20852

Print ISSN: 0022-1767 Online ISSN: 1550-6606.

Mutational Analysis of Terminal Deoxynucleotidyltransferase-Mediated N-Nucleotide Addition in V(D)J Recombination[1]

Jamie A. E. Repasky, Elizabeth Corbett, Cristian Boboila, and David G. Schatz[2]

The addition of nontemplated (N) nucleotides to coding ends in V(D)J recombination is the result of the action of a unique DNA polymerase, TdT. Although N-nucleotide addition by TdT plays a critical role in the generation of a diverse repertoire of Ag receptor genes, the mechanism by which TdT acts remains unclear. We conducted a structure-function analysis of the murine TdT protein to determine the roles of individual structural motifs that have been implicated in protein-protein and protein-DNA interactions important for TdT function in vivo. This analysis demonstrates that the N-terminal portion of TdT, including the BRCA-1 C-terminal (BRCT) domain, is not required for TdT activity, although the BRCT domain clearly contributes quantitatively to N-nucleotide addition activity. The second helix-hairpin-helix domain of TdT, but not the first, is required for activity. Deletional analysis also suggested that the entire C-terminal region of TdT is necessary for N-nucleotide addition in vivo. The long isoform of TdT was found to reduce N-nucleotide addition by the short form of TdT, but did not increase nucleotide deletion from coding ends in either human or rodent nonlymphoid cells. We consider these results in light of the recently reported structure of the catalytic region of TdT. ***The Journal of Immunology*****, 2004, 172: 5478–5488.**

The enormous diversity in Ig and TCR Ag binding sites is created during assembly of their genes from individual V, sometimes D, and J gene segments in a process called V(D)J recombination (1). This diversity is generated at multiple levels. Combinatorial diversity arises from the use of different combinations of V, D, and J segments. Superimposed on this is junctional diversity, created by the imprecise joining of the gene segments during V(D)J recombination. Finally, Igs and TCRs are heterodimers of rearranged receptor genes, with both subunits contributing to Ag recognition, thus creating an additional layer of combinatorial diversity.

The first phase of V(D)J recombination begins with binding by the recombination-activating gene 1 (RAG1)[3]/RAG2 proteins to a pair of recombination signal sequences, followed by double-strand cleavage of the DNA to generate two hairpin sealed coding ends and two blunt signal ends. The second phase (end processing and joining) is performed primarily by the ubiquitously expressed nonhomologous ending joining machinery (DNA ligase IV, XRCC4, Ku70, Ku80, $DNAPK_{cs}$, and Artemis) and results in the formation of a coding joint and a signal joint. Coding joints are typically imprecise, exhibiting both nucleotide deletion and the addition of templated and nontemplated nucleotides. Signal joints are often precise fusions of the two recombination signal sequence, although frequent addition of nontemplated nucleotides to signal joints has been observed for several Ag receptor loci (2–4) and with extrachromosomal recombination substrates (5).

Nontemplated nucleotide (N-nucleotide) addition to coding and signal joints in V(D)J recombination is due almost entirely to the action of the enzyme TdT. TdT is a template-independent DNA polymerase that adds random nucleotides (with a GC bias) to the 3′-hydroxyl of single- or dsDNA (reviewed in Ref. 6). TdT-deficient mice are viable and perform V(D)J recombination normally, although virtually no N-nucleotides are found in the assembled Ig and TCR genes (7, 8). In addition, in the absence of TdT, coding joints show a significant increase in the use of short stretches of microhomology near the ends of the coding segments to mediate joining, suggesting that TdT also influences the joining process (9). How TdT interferes with microhomology-mediated joining is uncertain (reviewed in Ref. 10). TdT is not expressed before birth in the mouse; in the adult it is expressed at most stages of thymocyte development and in pro-B, but not pre-B, cells (reviewed in Ref. 11). Hence, adult TCR and IgH genes, but not Ig L chain genes, routinely contain N-nucleotides.

TdT belongs to the pol X family of polymerases (12), a subgroup of an ancient nucleotidyltransferase superfamily defined by homologies within nucleotide binding domains and active site motifs (13). Although DNA polymerases β (pol β), μ (pol μ), and λ (pol λ), preferentially add nucleotides complementary to those on a template strand, TdT is unique in that it adds nucleotides in an untemplated fashion. Homology among members of the pol X family is not limited to their catalytic center. Additional domains shared by several of these polymerases include BRCA-1 C-terminal (BRCT) domains as well as helix-hairpin-helix (HhH) domains (see Fig. 1*A* for schematic of the TdT protein). The BRCT domain is a phosphopeptide binding motif (14, 15) that mediates protein-protein interactions and is commonly found in proteins involved in DNA recombination and repair, such as BRCA-1, XRCC4, and

Howard Hughes Medical Institute, Section of Immunobiology, Yale University School of Medicine, New Haven, CT 06510

Received for publication June 26, 2003. Accepted for publication February 10, 2004.

The costs of publication of this article were defrayed in part by the payment of page charges. This article must therefore be hereby marked *advertisement* in accordance with 18 U.S.C. Section 1734 solely to indicate this fact.

[1] This work was supported by Grant AI32524 from the National Institutes of Health (to D.G.S.). D.G.S. is an investigator with the Howard Hughes Medical Institute.

[2] Address correspondence and reprint requests to Dr. David G. Schatz, Howard Hughes Medical Institute, Section of Immunobiology, Yale University School of Medicine, Box 208011, New Haven, CT 06520-8011. E-mail address: david.schatz@yale.edu

[3] Abbreviations used in this paper: RAG, recombination-activating gene; BRCT, BRCA-1 C-terminal; Cam, chloramphenicol; CHO, Chinese hamster ovary; $DNAPK_{cs}$, catalytic subunit of the DNA-dependent protein kinase; H2, second HhH domain; HhH, helix-hairpin-helix; Kan, kanamycin; N, nontemplated; NLS, nuclear localization signal; pol, polymerase; TdT-FL, full-length TdT; TdTL, long isoform of TdT.

0022-1767/04/$02.00

DNA ligase IV (16–18). One such BRCT domain was also identified in the N-terminal regions of TdT, pol μ and pol λ (19).

Support is growing for the possibility that TdT is recruited to coding ends during V(D)J recombination by components of the nonhomologous end-joining machinery. The lack of N-nucleotides in the rare CJs found in Ku80$^{-/-}$ and *scid* mice led to the suggestion that the Ku heterodimer and/or DNAPK$_{cs}$ plays this role (5, 20, 21). Furthermore, an in vitro study demonstrated that the Ku heterodimer interacts with the N-terminal portion of TdT in coimmunoprecipitation experiments (22). In contrast, Mickelsen et al. (23) reported that in gel shift assays, TdT was able to associate with DNAPK$_{cs}$ on DNA ends, but not with the Ku heterodimer. DNAPK$_{cs}$ could modulate the activity of TdT in vitro with respect to the length and composition of nucleotide addition (23), suggesting that DNAPK$_{cs}$ directs and controls TdT activity during V(D)J recombination. Perhaps most intriguing is a recent biochemical study demonstrating that TdT and pol μ can be stably recruited to DNA by virtue of interactions with Ku-XRCC4-DNA ligase IV, but that Ku alone or XRCC4-ligase IV alone is not sufficient (24). If active recruitment of TdT to coding ends does occur during V(D)J recombination, the domain(s) of TdT involved has not been defined, although the BRCT domain remains an attractive candidate.

HhH domains are nonsequence-specific DNA binding motifs that contact DNA by interactions of peptide backbone nitrogen atoms with phosphate groups of the DNA (25). The two putative HhH domains identified in TdT, pol β, pol μ, and pol λ, are thus expected to bind to the substrate DNA and to position it properly for nucleotide addition. Although both HhH domains in pol β match the consensus motif (GϕG, with ϕ being a hydrophobic residue, typically isoleucine or valine), in TdT, HhH1 differs from the consensus sequence (CϕG instead of GϕG). Because the two glycines are critical for the sharp hairpin characteristic of HhH domains, it is not clear whether this region of TdT can adopt this structure. Interestingly a similar variation of the HhH motif is also present in DNA pol μ and λ (19).

TdT from a number of species has been reported or predicted to undergo alternative splicing (26). In mice, alternative splicing results in the production of a long form of TdT (TdTL) containing a 20-aa insertion near the C terminus of the protein. TdTL was initially reported to display a thermolabile terminal transferase activity in vitro, but to be unable to add N-nucleotides to V(D)J coding joints (27, 28). More recently, however, TdTL was reported to possess a 3′ to 5′ exonuclease activity in vitro and to lack terminal transferase activity in vitro and in vivo (26). The finding of an exonuclease activity associated with TdTL provided an explanation for the increased nucleotide deletion from coding ends observed in transient V(D)J recombination assays (in Chinese hamster ovary (CHO) cells) containing TdTL compared with assays containing no TdT or containing the short form of TdT (which we will refer to in this report as full-length TdT (TdT-FL)) (26). It also explained the observation that TdTL was able to down-modulate the activity of the short isoform of TdT in transgenic mice (29) and in transient recombination assays (26). Because there remain some uncertainties concerning the catalytic activities associated with TdTL, we decided to re-examine its activities in vivo.

Recently, a crystal structure of TdT was reported that confirms many of the predictions presented above (30). The portion of TdT crystallized (aa 130–510) excludes the putative nuclear localization signal (NLS) and BRCT domain, but retains catalytic activity (31). The structure of TdT closely resembles that of other DNA polymerases (30), with a shape similar to a right hand, consisting of an N-terminal finger domain, a central palm containing the active site, and a C-terminal thumb domain. TdT exhibits a particularly strong structural similarity to pol β. Although the primary amino acid sequence identity between the two proteins is only 22%, almost all secondary structure elements have been conserved. Pol β can adopt two different conformations, open and closed (32). The catalytic core of TdT adopts a closed conformation, mediated by an interaction of the N-terminal portion of the core protein with the C terminus. The active site of TdT contains three conserved aspartate residues (D343, D345, and D434), in keeping with a previous mutagenesis study (33), and mirrors that of pol β, suggesting an identical two-divalent metal ion mechanism for catalysis (34).

Although TdT is well characterized structurally, virtually nothing is known about the contributions of its defined or proposed structural domains to function in vivo. Therefore, we generated mutant forms of murine TdT to assess the relevance of proposed domains and motifs for the only known biological activity of TdT, the addition of N-nucleotides to coding joints in V(D)J recombination. The results raise questions about the relevance of the proposed N-terminal NLS, reveal a nonessential role for the BRCT domain, and define essential roles for the second HhH motif and the entire C terminus of TdT. We were unable to obtain evidence for an interaction between TdT and Ku by coimmunoprecipitation from cell extracts or for a nuclease activity associated with TdTL in human or rodent nonlymphoid cell lines.

Materials and Methods

Cloning of TdT mutants

The untagged TdT mutants were PCR-amplified using *Pfu* polymerase (Stratagene, La Jolla, CA) from the pDTS TdT cDNA (27) and cloned into the pEBB-RAG1 vector after digestion with *Eco*RI to remove the RAG1 cDNA. Constructs generated included TdT-FL (primers MUT11 and MUT3prim), TdT-ΔN (MUT10 and MUT3prim), TdT-ΔB (MUT11 and MUT3), NLS alone, MUT5 and MUT3prim (digested with *Xba*I and ligated before *Eco*RI cloning), TdT-ΔN+B (MUT19 and MUT3prim), TdT-232 (MUT22 and MUT3prim), TdT-280 (MUT23 and MUT3prim), TdT-421 (MUT11 and MUT18), TdT-453 (MUT11 and MUT21), and TdT-481 (MUT11 and MUT20). TdTL sequences were PCR-amplified from the mTdTL cDNA (provided by T.-H. Thai and J. Kearney, University of Alabama, Birmingham, AL). Mutations were confirmed by DNA sequencing.

Site-directed mutagenesis

The TdT cDNA was subcloned into the pBluescript II KS^{+} vector and mutagenized using the QuikChange kit and *Pfu* polymerase (Stratagene) according to the manufacturer's instructions. The mutants included TdT-ASM (primers D343/5E-F and D343/5E-B, template TdT-FL), TdT-ΔH1 (HHH1TOP and HHH1BOT, template TdT-FL), TdT-ΔH2 (HHH2TOP and HHH2BOT, template TdT-FL), and TdT-ΔH1/2 (HHH2TOP and HHH2BOT, template TdT-ΔH1). Mutations were confirmed by DNA sequencing. The mutants were PCR-amplified and cloned into pEBB.

Transient transfection V(D)J recombination assay in 293T and CHO-K1 cells

293T cells were transfected by a calcium phosphate method as described previously (35), using 5 μg of each plasmid: recombination substrate pSF290 or pSF200 (35), pEBB-RAG1, and pEBB-RAG2, with or without pEBB-TdT mutants. After 48 h at 37°C, extrachromosomal DNA was harvested by rapid alkaline lysis as described previously (35). DNA was then transformed into MC1061 bacteria and plated on agar containing kanamycin (Kan) and chloramphenicol (Cam). When transformed into bacteria, pSF290 and pSF200 confer resistance to Kan in both the unrearranged and the rearranged form, and after successful V(D)J recombination they confer resistance to Cam. To analyze the structure of the joints, individual Kan^{R}+Cam^{R} colonies were grown, and the portion of the plasmid containing the relevant junction was sequenced.

CHO-K1 cells were transfected using FuGene 6 (Roche, Indianapolis, IN) according to the manufacturer's instructions. Briefly, FuGene 6 (3 μl/μg DNA) was mixed with serum-free medium to a final volume of 100 μl, and DNA was then added (2.5 μg of the recombination substrate pJH290 (36) and 2 μg each of pEBB-RAG1, pEBB-RAG2, and, as appropriate, pEBB-TdT-FL and/or pEBB-TdTL). The mixture was added to the cells, and extrachromosomal DNA was harvested 2 days after transfection

and transformed into bacteria as described above. Recombined substrates were isolated from Amp^R+Cam^R bacterial colonies, and coding joints were analyzed by sequencing.

Immunoprecipitation

Cells were washed with PBS, and the cell pellet was resuspended with 0.5–1 ml of lysis buffer (1 mM PMSF, 20 μM leupeptin, 8 μM pepstatin, 1 mM benzamidine, and 1% Nonidet P-40 in TBS, pH 7.5). The cells were sonicated three times for 10 s each time on ice. The solution was rotated for 30 min at 4°C and spun at 14,000 rpm for 15 min at 4°C. Five micrograms of the Abs (TdTβ, Ku70 (clone N3H10; Kamiya Biomedical, Thousand Oaks, CA), Ku80 (clone 111; Kamiya Biomedical), DNA-PK_{cs} (clone 18-2; Neomarker), or IgG (Sigma-Aldrich, St. Louis, MO)) were added to each sample and rotated overnight at 4°C. The samples were spun again at 14,000 rpm for 10 min, and 30 μl of a washed protein G slurry was added to the supernatant and rotated for 2 h at 4°C. Each sample was washed five times with 1 ml of lysis buffer. The pellet was resuspended with 30 μl of 2× loading buffer, boiled, and loaded on an SDS-polyacrylamide gel.

Immunofluorescence

Forty-eight hours after transfection, cells were split onto coverslips in a 24-well plate (1×10^5 cells/200 μl) and incubated overnight at 37°C. Cells were washed with PBS, fixed in methanol for 5 min at 4°C, washed with PBS four times, and permeabilized with PS+ buffer (RPMI 1640, 10% newborn calf serum, 0.05% saponin, 10 mM HEPES, and 10 mM glycine) for 5 min at room temperature. The anti-TdT Ab (Supertech) was diluted 1/10 in 30 μl of PS+ buffer, and the cells were incubated upside down for 30 min at room temperature in a moisture chamber. The coverslip was returned to the 24-well plate, and the cells were washed four times in PS+. The secondary Abs (1/300 goat anti-rabbit, 1/5000 SYTOX) were diluted in 30 μl of PS+ and incubated on the coverslips for 30 min at room temperature in the moisture chamber. Coverslips were washed twice with PS+ PBS, dipped briefly in water, vacuum-dried, and mounted in 5 μl of mounting solution.

Chromosomal V(D)J recombination in 3TGR cells

3TGR cells were seeded in a 24-well plate in 0.5 ml of medium and incubated overnight at 37°C. Lipofectamine 2000 (3 μl; Invitrogen, Carlsbad, CA) was diluted in 50 μl of Opti-MEM medium (Life Technologies, Gaithersburg, MD) and incubated for 5 min at room temperature. A solution of pEBB-RAG1 and pEBB-RAG2, with or without pEBB-TdT mutants (3 μg of DNA total), was added to the Lipofectamine/Opti-MEM solution, incubated at room temperature for 20 min, added to the cells, and incubated for 48 h at 37°C. Cells were harvested, washed with PBS, resuspended in 500 μl of lysis buffer (50 mM Tris (pH 8), 20 mM EDTA, 1% sodium lauryl sulfate, and proteinase K), and incubated at 55°C overnight. The genomic DNA was phenol/chloroform-extracted twice, precipitated in ethanol, and resuspended in 50 μl of TE (10 mM Tris, pH 8.0, 1 mM EDTA). CJs were amplified using nested PCR and *Taq* polymerase (Life Technologies). First-round amplification (25-μl reaction) was performed for 15 cycles using primers V2 (5′-CTCCTCATCTATCGTGCATCCAACC-3′) and BR-1 (5′-GGAAGC GAGAAGAATCATAATGGG-3′). One microliter of the first-round reaction was added to a fresh mix containing primers BR-3 clone (5′-GCGGAAT TCGGAAGGCCATCCAGCCTCGCGTCG-3′) and V5 clone (5′-GCG GAATTCCAACCTAGAATCTGGGATCCC-3′) and amplified for an additional 25 cycles.

Statistical analysis

Statistical analysis of N-nucleotide addition, coding end deletion, and microhomology use was conducted using a two-sample *t* test with DataDesk version 5.0 (Data Description, Ithaca, NY).

Results

TdT mutants

The first set of mutants was designed to inactivate individual predicted functional domains (all mutants are depicted schematically in Fig. 1*B*). The respective regions were either deleted entirely, or critical amino acids within these domains were mutated to abolish their function: TdT-FL (TdT-FL cDNA, short form; Fig. 1*A*), TdT-ASM (pol X active site motif mutant, D343E, D345E, identical with the catalytic mutant used by Yang et al. (33)), TdT-ΔN (deletion of putative NLS, aa 1–27 (37)), TdT-ΔB (deletion of the BRCT domain, aa 26–143), TdT-ΔH1 (mutated first HhH domain, $GIP_{213–215}$ AAA), TdT-ΔH2 (mutated second HhH domain, $GVG_{257–259}$ AAA), and TdT-ΔH1/2 (both HhH domains mutated).

The second panel of mutants comprised a series of C- and N-terminal truncations: TdT-ΔN+B (aa 1–143 deleted), TdT-232 (N-terminal truncation after the first HhH, lacking aa 1–232), TdT-280 (N-terminal truncation after the second HhH, lacking aa 1–280), TdT-421 (C-terminal truncation, deletion of aa 421–501), TdT-453 (C-terminal truncation, deletion of aa 453–501), and TdT-481 (C-terminal truncation, deletion of aa 481–501). All TdT mutants were expressed without an epitope tag, because in preliminary experiments the presence of even a short C-terminal Myc-epitope tag was found to dramatically reduce activity (data not shown). Finally, we also analyzed the long isoform of murine TdT, TdTL, which contains 20 additional aa near its C terminus as a result of alternative splicing (TdTL cDNA provided by T.-H. Thai and J. Kearney).

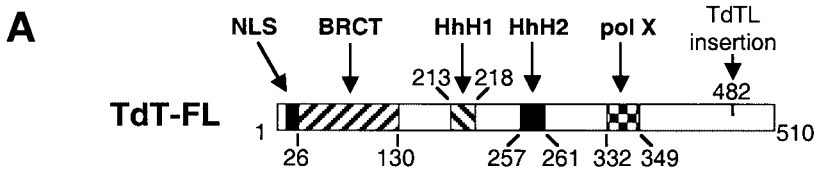
All TdT mutants were tested for their ability to add N-nucleotides to coding and signal joints. For selected TdT mutants, the assays were a performed with either full-length or truncated core RAG proteins to determine whether there exists a functional interaction between TdT and the noncore regions of the RAG proteins (which are dispensable for V(D)J recombination in such assays). For a subset of TdT mutants, including TdTL, activities were measured in both extrachromosomal and chromosomal V(D)J recombination assays. In all cases duplicate sequences obtained from a single transfection were counted only once. Histograms showing the distribution of coding end nucleotides deleted or N-nucleotides added for each mutant are available upon request.

N-nucleotide addition to coding joints using episomal recombination substrates

The activity of all TdT mutants was first assessed using a V(D)J recombination assay involving episomal substrates transfected into human 293T cells. Cells were transfected with expression vectors for RAG1, RAG2, and TdT and with artificial recombination substrate pSF290 or pSF200 to test for CJ or SJ formation, respectively. After 48 h the plasmids were harvested from the cells and transformed into bacteria, individual Kan^R+Cam^R colonies were grown, and the portion of the plasmid containing the relevant junction was sequenced. Coding junctions were analyzed for the frequency of addition of N-nucleotides, the number of N-nucleotides added, the number of nucleotides deleted from the coding ends, and whether microhomologies were present. Signal joints were analyzed for the frequency of addition of N-nucleotides. The results obtained in assays using the full-length RAG proteins are shown in Fig. 1*B*, whereas those obtained with the core RAG proteins are shown in Fig. 2. Western blotting demonstrated that all mutants of TdT were expressed at levels comparable to that of TdT-FL, with the exception of TdT-ΔN+B, which accumulated to lower levels (data not shown).

With full-length RAG1/RAG2, wild-type TdT (TdT-FL) added N-nucleotides to 59% of CJs and 44% of SJs (Fig. 1*B*, *line 2*), whereas in the absence of TdT, only a single, 1-bp addition was observed in the 36 junctions examined (*line 1*). For TdT-FL, the average number of N-nucleotides found in all CJs sequences was 2.0, whereas this value rose to 3.3 when only those CJs with N-nucleotides added were considered. The nucleotides added by TdT in the in vivo assay showed a preference for the addition of G and C nucleotides (60%), as observed in endogenous Ag receptor genes (6).

To confirm that the pol X catalytic core motif is required for enzymatic activity, an active site mutant, TdT-ASM, was examined. This TdT variant (DD343,345EE) is identical with the catalytic mutant analyzed in a previous biochemical study (33). These



B Full length RAG proteins in 293T cells

		Mutation	# Rx	CJ N-nt addition	Avg. Length N-nt addition	micro-homology use	SJ N-Nt addition	Avg Nt deletion
1	Full Length Rags alone		6	0% (0/24)	0 [0] **	15/24 (63%) *	8% (1/12)	6.1 (2.9)
2	TdT-FL		5	59% (19/32)	3.3 [2.0]	5/32 (16%)	44% (22/50)	5.7 (3.3)
3	TdT-ASM	D343E D345E	4	0% 0/22	0 [0] **	14/22 (64%) *	0% (0/12)	5.8 (2.8)
4	TdT- Δ N	del 1-27	2	80% (16/20)	2.7 [2.2]	1/20 (5%)	17% (2/12)	5.3 (2.5)
5	TdT- Δ B	del 26-143	3	55% (17/31)	2.0 [1.1]	5/31 (16%)	0% (0/12)	5.2 (2.2)
6	TdT-Δ N+ B	del 1-143	5	34% (12/35)	1.9 [0.7] *	10/35 (29%)	ND	5.4 (2.6)
7	TdT-Δ H1	GIP 213-5 AAA	4	63% (17/27)	2.2 [1.6]	7/27 (26%)	8% (1/12)	4.7 (2.7)
8	TdT-Δ H2	GVG 257-9 AAA	5	3% (1/32)	1 [0.03] **	20/32 (63%) **	0% (0/12)	5.4 (2.3)
9	TdT-Δ H1/2		2	0% (0/10)	0 [0] **	6/10 (60%) *	0% (0/18)	4.6 (2.6)
10	TdT-232	del 1-232	4	5% (1/21)	1 [0.05] **	10/21 (48%)	ND	6.0 (2.7)
11	TdT-280	del 1-280	3	9% (2/21)	1 [0.1] **	12/21 (57%)	ND	6.0 (3.5)
12	TdT-421	del 421-501	3	0% (0/20)	0 [0] **	12/20 (60%) *	8% (1/12)	6.0 (2.5)
13	TdT-453	del 453-501	2	0% (0/10)	0 [0] **	8/10 (80%) **	ND	5.3 (2.4)
14	TdT-481	del 481-501	4	4% (1/27)	1 [0.04] **	17/27 (63%)	ND	6.4 (2.8)
15	TdT-L	insert at 482 20aa	4	21% (5/24)	1.4 [0.3] **	11/24 (46%)	22% (4/18)	4.9 (2.7)
16	TdTL + TdTFL		3	65% (11/17)	1.6 [1.0] *	2/17 (12%)	35% (7/20)	6.5 (4.7)
17	TdT-ASM + TdT-FL		4	52% (12/23)	2.2 [1.1] *	6/23 (26%)	20% (3/15)	5.8 (2.7)

FIGURE 1. Activity of TdT mutants in 293T cells with full-length RAG proteins. *A*, Structure of TdT-FL. The domains of TdT are depicted as rectangles and labeled as follows: NLS, BRCA-1 C-terminal domain (BRCT), helix-hairpin-helix (HhH1 and HhH2), and pol X active site. Numbers refer to the murine TdT-FL sequence. The 20-aa insertion found in TdTL occurs after aa 482. *B*, 293T cells were transfected with vectors expressing the full-length RAG proteins, a recombination substrate (pSF290 or pSF200), and full-length (FL) or mutant TdT, as indicated. CJs and SJs were amplified and sequenced. Mutation, deleted or mutated amino acids; # Rx, number of independent recombination reactions; CJ N-nt addition, percentage with N-nt added (number of sequences with N-nucleotide added/total number of sequences); Avg Length N-nt addition, average number of N-nucleotides added to CJs based only those sequences with N-nucleotides (based on all sequences); statistical analysis of the N-nucleotide addition in comparison with N-nucleotide addition by TdT-FL is indicated as follows: *, $p < 0.05$; **, $p < 0.001$); microhomology use, number of sequences with at least one nucleotide of homology at the CJ (percentage; statistical analysis in comparison with microhomology use in presence of TdT-FL is indicated as above); SJ N-nt addition, percentage of SJs with at least one N-nucleotide added (number of sequences); Avg Nt-deletion, average total number of nucleotides deleted from the two coding ends (std dev); ND, not determined.

mutations maintain the charge and chemical nature of the side chains and thus are not expected to have a significant effect on the global structure of the protein. As expected, this TdT mutant was found to be inactive in vivo; no N-nucleotides were detected in the 22 joints obtained from four independent transfections (Fig. 1*B*, *line 3*). Thus, the active site identified within the pol X family catalytic core of TdT is essential for polymerase activity in vivo and in vitro.

Of the other 12 TdT mutants examined, four were found to be active in vivo: TdT-ΔN, TdT-ΔB, TdT-ΔN+B, and TdT-ΔH1 (Table I*B*, *lines 4–7*). Although three of these mutants, TdT-ΔN (80%), TdT-ΔB (55%), and TdT-ΔH1 (63%), demonstrated frequencies of N-nucleotide addition to CJ comparable to TdT-FL (59%), TdT-ΔN+B added N-nucleotide to only 34% of joints. As noted above, TdT-ΔN+B was not expressed as well as the other TdT proteins, and this could explain its reduced activity. The average length of the N regions added to CJs was, however, reduced in all four cases. Although the average number of nucleotides added by TdT-ΔN was 80% of the wild-type level (2.7 nt compared with 3.3 nt), for TdT-ΔB, TdT-ΔN+B, and TdT-ΔH1, N regions were, on the average, at least 1 bp shorter, containing only 1.9–2.2 N-nucleotides (Fig. 1*B*). This difference is statistically significant for TdT-ΔN+B ($p < 0.05$) and borderline significant for TdT-ΔB ($p = 0.057$). Whether this reduction is due to partial misfolding of the mutant proteins, reduced expression (in the case of TdT-ΔN+B), or a direct contribution of these deleted domains to the catalytic activity remains to be elucidated. In summary, neither the NLS, the

	Mutation	# Rx	CJ N-nt addition	Avg. Length N-nt addition	micro-homology use	SJ N-nt addition	Avg Nt deletion
1 Core Rags alone		2	5% (1/18)	1 [0.06] **	11/18 (61%) *	0% (0/12)	5.8 (3.2)
2 TdT-FL		3	68% (23/34)	3.9 [2.6]	5/34 (15%)	19% (5/27)	5.4 (2.6)
3 TdT-ASM	D343E D345E	4	6% (2/36)	1 [0.06] **	22/36 (61%) **	9% (1/11)	6.4 (3.4)
4 TdT-Δ B	del 26-143	3	39% (17/44)	2.6 [1.0] *	10/44 (23%)	13% (2/16)	5.3 (2.9)
5 TdT-421	del 421-501	4	0% (0/29)	0 [0] **	14/29 (48%) *	8% (1/12)	6.7 (4.0)
6 TdT-L	insert at 482 20aa	4	23% (7/30)	1.7 [0.4] **	12/30 (40%)	0% (0/12)	5.7 (2.7)
7 TdTL + TdT-FL		4	48% (14/29)	2.4 [1.2] *	6/29 (21%)	0% (0/16)	6.2 (2.3)
8 TdT-ASM + TdT-FL		2	67% (10/15)	3.9 [2.9]	1/15 (11%)	13% (2/16)	5.5 (4.2)

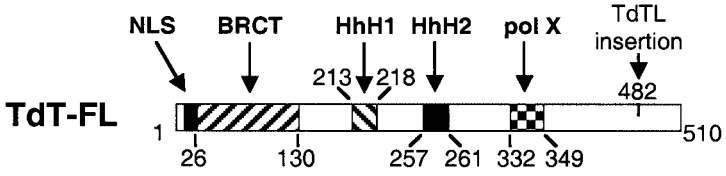


FIGURE 2. Activity of TdT mutants in 293T cells with core RAG proteins. 293T cells were transfected with vectors expressing the core RAG proteins, a recombination substrate (pSF290 or pSF200), and full-length (FL) or mutant TdT, as indicated. CJs and SJs were amplified and sequenced. Other symbols and abbreviations are explained in Fig. 1.

BRCT domain, nor the first HhH motif is essential for the addition of N-nucleotide to CJs by TdT. In particular, the first 27 aa, spanning the putative NLS, are largely dispensable for activity in this assay.

In contrast, all seven remaining mutants, TdT-ΔH2, TdT-ΔH1/2, TdT-232, TdT-280, and TdT-421, TdT-453, and TdT-481, show no appreciable activity in vivo. This indicates that the second HhH domain is critical for TdT activity, and that TdT cannot undergo even a 20-aa deletion at its C terminus without a complete loss of activity.

N-nt addition to signal joints

A somewhat different picture was seen when nucleotide addition to SJs was analyzed. The wild type protein, TdT-FL, was the only form that added N-nt to a significant number of SJs (44%). Three mutants with partial activity at CJs exhibited a much greater deficit in N-nt addition to SJs (Fig. 1B, lines 4, 5, and 7). Deletion of the putative NLS, which has almost no effect on activity at CJs, reduces activity at SJs by >60%, and deletion of the BRCT domain or mutation of HhH1 essentially eliminates N-nt addition to SJs. The other mutant forms of TdT analyzed for N-nt addition to SJs were also defective in this activity (lines 8, 9, and 12). As observed for CJs, the N-nucleotides added to SJs were G/C rich (57% with full length RAG proteins; 67% with core RAG proteins) (data not shown). Altogether, TdT is less active at signal ends as compared with coding ends, and the activity at signal ends is particularly sensitive to mutations in TdT. This is consistent with the previous observation that higher levels of TdT are required for addition of N-nucleotides to SJs than to CJs (38), and is in keeping with the relatively poor activity of TdT on blunt ends (39), such as signal ends.

N-nucleotide addition in the context of the core RAG proteins

The core RAG proteins appear to perform DNA cleavage in vivo efficiently, as measured by steady state levels of signal ends and coding ends, but they are significantly compromised for SJ and CJ formation relative to the full-length proteins (40–42). To determine whether the noncore regions of the RAG proteins are required for efficient recruitment of TdT to coding ends, TdT-FL, TdT-ASM, TdT-ΔB, and TdT-421 were analyzed using the core RAG proteins by transient transfection of 293T cells. As expected, TdT-ASM and TdT-421 were inactive (Fig. 2, *lines 3* and *5*). TdT-FL and TdT-ΔB both mediated N-nucleotide addition at CJs, but deletion of the BRCT domain resulted in a reduction in the frequency and average length of N-nucleotide addition (39%, 2.6 nt for TdT-ΔB; 68%, 3.9 nt for TdT-FL; $p = 0.046$; Fig. 2, *lines 4* and *2*). These results further support the conclusion that the BRCT domain of TdT is not essential for function, but contributes to the overall level of activity.

Interestingly, N-nt addition to SJs was reduced using the core RAG proteins compared with their full-length counterparts (compare Fig. 2 with Fig. 1). Previous results led to the suggestion that the nonessential portions of the RAG proteins, missing from the core proteins, facilitate remodeling of postcleavage complexes and thereby enhance signal joint formation (42). It may be that postcleavage complexes containing the core RAG proteins only inefficiently assume a conformation conducive to TdT activity

at signal ends, perhaps because the 3′-hydroxyl groups are not made accessible.

TdT- and microhomology-mediated coding joint formation

Previous results indicated that TdT reduces the use of microhomologies to mediate CJ formation (see introduction). Our results confirm this. The presence of TdT-FL significantly reduces the use of microhomologies in CJs relative to the situation where only the RAG proteins were expressed (16 vs 63%; Fig. 1*B*, *lines 2* and *1*). To separate possible effects of TdT on microhomology joining due to its enzymatic activity as opposed to the mere presence of TdT in the end-joining machinery, the CJs formed in the presence of the catalytic mutant, TdT-ASM, were analyzed. As noted above, the subtle mutations in this protein maintain the charge and chemical nature of the side chains and are not expected to have an effect on the global structure of the protein. Thus, putative interactions with the DNA substrate and other proteins should remain normal. Interestingly, TdT-ASM does not reduce microhomology-based joining compared with TdT-FL. Sixty-four percent of the CJs in the presence of TdT-ASM were created by the microhomology pathway, nearly identical with the 63% observed in the absence of TdT (Fig. 1*B*, *lines 1* and *3*). This suggests that the enzymatic activity of TdT, i.e., the addition of N-nucleotides, and not recruitment of the enzyme to the DNA ends, plays an important role in the choice of the joining pathway. Two additional findings provide further support for this model. First, all TdT mutants that lack polymerase activity show frequencies of microhomology-mediated CJ formation comparable to those found in the absence of TdT (Fig. 1*B*, compare *line 1* to *lines 3* and *8–14*). Second, all active mutants, TdT-ΔN, TdT-ΔB, TdT-ΔN+B, and TdT-ΔH1, just like the wild-type form TdT-FL, reduce microhomology use in CJs to less than half the value observed when expressing the RAG proteins alone (compare *line 1* to *lines 2* and *4–7*).

TdT and nucleotide deletion from coding ends

TdT is only one factor in a group of enzymes involved in processing of coding ends before joining. We were interested in determining whether TdT or mutants thereof have any effect on other end-processing activities. We thus analyzed all of our sequences, from reactions with full-length or core RAG proteins, for the number of nucleotides deleted from the coding ends. Even though the use of microhomology for joining is reduced by active forms of TdT, the average number of nucleotides deleted was not statistically affected by the expression of full-length or any mutant form of TdT (all $p > 0.1$ compared with the nucleotide deletion in the absence of TdT; Figs. 1*B* and 2). The most common junctions sustained a net loss of four, six, seven, or eight nucleotides, which corresponds to junctions generated by microhomology-mediated joining. Although coding ends with different sequences have not been analyzed, this indicates that end-processing activities that result in nucleotide deletion act independently of TdT. We also note that full-length and core RAG proteins gave rise to essentially indistinguishable patterns of nucleotide deletion from coding ends, indicating that nucleotide removal from coding ends is not significantly affected by the N-terminal region of RAG1 or the C-terminal region of RAG2.

Activities of TdTL

Several observations emerged from the recombination assays we performed with TdTL. First, TdTL is capable of adding N-nucleotides to coding joints with both the full-length and core RAG proteins and exhibits some activity at signal joints as well (Fig. 1*B*, *line 15*, and Fig. 2, *line 6*). Second, the frequency and length of N-nucleotide additions to CJs are much reduced compared with those of TdT-FL. Third, TdTL appears to be less efficient than TdT-FL in blocking microhomology use, but still has some activity in this regard.

Benedict et al. (29) proposed that TdTL is involved in the regulation of TdT-FL (also referred to as the short isoform of TdT) activity in vivo. Specifically, TdTL was observed to reduce the average length of N-nucleotide additions when coexpressed with TdT-FL in transgenic mice (29) A plausible explanation for this was provided by the observation that TdTL is a 3′ to 5′ exonuclease, and hence is presumably able to delete nucleotides from coding ends (26). Evidence for this was provided by Thai et al. (26), who reported that the expression of TdTL resulted in increased nucleotide deletion from coding ends in a transient transfection assay in CHO cells.

Our data in 293T cells are consistent with a subset of these observations. Coexpression of TdTL with TdT-FL reduced the average length of N-nucleotide additions compared with TdT-FL alone: from 3.3 to 1.6 with full-length RAG proteins (Fig. 1*B*, *line 16*) and from 3.9 to 2.4 with core RAG proteins (Fig. 2, *line 7*). The frequency of CJs containing N-nucleotides was not reduced (full-length RAG proteins; Fig. 1) or was modestly reduced (core RAG proteins; Fig. 2). These results are consistent with those previously reported (26, 29) and support the model that TdTL acts to down-modulate the activity of TdT-FL without substantially affecting its recruitment to coding ends. Interpretation of these data is complicated by our observation that coexpression of TdT-ASM with TdT-FL resulted in a comparable reduction in N-nucleotide length, but only in reactions using the full-length RAG proteins (Fig. 1*B*, *line 17*), not with the core RAG proteins (Fig. 2, *line 8*). We conclude that TdTL is indeed capable of reducing the average length of N regions added in the presence of TdT-FL (discussed below).

In contrast with previously published results in CHO cells (26), we did not observe any increase in nucleotide deletion from coding ends when TdTL was expressed in 293T cells. This was true when TdTL was expressed by itself or when it was coexpressed with TdT-FL, and was true in reactions using either the full-length or core RAG protein (Fig. 1*B*, *lines 15* and *16*, and Fig. 2, *lines 6* and *7*). Hence, TdTL does not enhance nucleotide deletion from coding ends using an extrachromosomal substrate in 293T cells.

The difference between our results and those reported by Thai et al. (26) might be due to the use of different cell lines derived from different species. We therefore repeated a subset of these experiments in the CHO-K1 cell line using the coding joint recombination substrate pJH290 (36). The results were, for the most part, very similar to those we obtained in 293T cells. The expression of TdTL in CHO-K1 cells (confirmed in all experiments by Western blot) resulted in detectable N-nucleotide addition and did not cause an increase in deletion of nucleotides from coding ends (Table I). In contrast to results in 293T cells, we did not detect a decrease in N-nucleotide addition in transfections containing TdT-FL and TdTL compared with those containing only TdT-FL (Table I). Hence, using a hamster cell line similar to that used by Thai et al. (26), we did not detect evidence for an exonuclease activity associated with TdTL capable of removing nucleotides from coding ends. As noted in the next section, TdTL also did not enhance nucleotide deletion from coding ends using a chromosomal recombination substrate in mouse 3T3 fibroblasts. Hence, in both human and rodent nonlymphoid cells, with both extrachromosomal and chromosomal substrates, TdTL was not able to increase nucleotide deletion from coding ends. Possible explanations for the difference between our results and those obtained by Thai et al. (26) are discussed below.

Table I. *Activity of TdT-FL and TdT-L in CHO cells with full-length RAG proteins*

	No. Rx[a]	CJ N-nt addition[b]	Avg. Length N-nt addition[c] (seq. with N-nt/all seq.)	Avg. Nt deletion[d] (SD)
Full-length RAG alone	7	0% (0/25)	0/0	5.1 (3.1)
TdT-FL	4	57% (16/28)	2.3/1.3	5.3 (4.1)
TdT-L	7	19% (4/21)	1.5/0.3	5.2 (5.1)
TdTL + TdT FL	6	61% (14/23)	3.2/2.0	5.3 (4.4)

[a] Number of independent recombination reactions.
[b] Percentage of coding joints with N-nucleotides added (number of sequences with N-nucleotides added/total number of sequences).
[c] Average number of N-nucleotides added to CI based only on those sequences with N-nucleotides/based on all sequences.
[d] Average total number of nucleotides deleted from the two coding ends (SD).

TdT activities on a chromosomal substrate

Assays using extrachromosomal substrates in 293T or CHO cells have some limitations. First, the proteins, such as RAG1/2 and TdT, are typically expressed at supraphysiological levels. This could obscure an important role for nonessential domains, such as the BRCT domain of TdT. Second, the chromatin configuration of extrachromosomal substrates is probably not identical with that of a chromosomal substrate. Third, it is formally possible that some recombination of extrachromosomal substrates occurs outside the nucleus, and this would obscure the importance of certain motifs (e.g., NLS). To address some of these issues and further examine the activity of TdTL, we tested several of the TdT mutants in the murine 3T3 fibroblast cell line, 3TGR. 3TGR contains two copies of a retroviral V(D)J recombination substrate, each integrated at a different chromosomal location, and has been used previously for the study of V(D)J recombination (43, 44). Expression vectors encoding TdT-FL or TdT mutants were transiently transfected into 3TGR cells along with full-length RAG1/2 expression plasmids. After 48 h, genomic DNA was prepared from the cells, and DNA fragments containing coding joints were amplified by PCR, subcloned, and sequenced.

TdT-FL added N-nucleotide to 56% of the CJs, comparable to the frequency of 59% observed using an extrachromosomal substrate (Fig. 3, *line 2*), indicating that the polymerase is efficiently recruited to coding ends on both types of substrates. The average length of the additions, however, was smaller with the chromosomal substrate (1.6 vs 3.3 nt with the episomal substrate; Fig. 3, *line 2*). This could reflect differences in TdT expression levels in the two assays, or differences in the repair process and/or TdT activity with chromosomal vs episomal substrates. Both TdT-ΔN and TdT-ΔB were also able to add N-nucleotide to CJs with the integrated substrate, albeit at lower frequencies than in the transient assay (23 and 31%, compared with 80 and 55%, respectively; Fig. 3, *lines 3* and *4*). Despite this reduced frequency of N-nucleotide addition, the average number of nucleotides added to CJs by TdT-ΔN and TdT-ΔB was nearly identical with that for TdT-FL in this system (1.4 and 1.6 nt, respectively, vs 1.6 nt for TdT-FL). Thus, in the case of a chromosomal V(D)J break, both the putative NLS as well as the BRCT domain are more important for recruitment than for polymerase activity. As in the episomal assay, TdTL exhibited detectable terminal transferase activity in the chromosomal assay (Fig. 3, *line 6*).

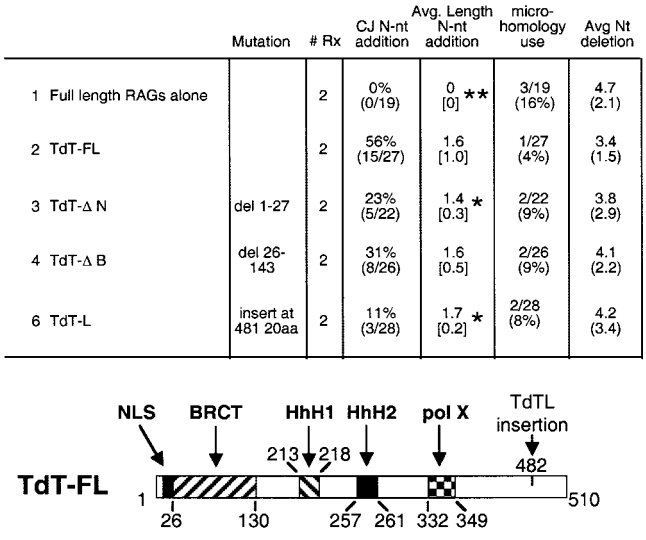


FIGURE 3. Activity of TdT mutants in 3TGR cells with full-length RAG proteins. 3TGR cells were transfected with vectors expressing the full-length RAG proteins and full-length (FL) or mutant TdT, as indicated. CJs were amplified from the recombination substrate stably integrated into the genome of 3TGR cells and sequenced. Other symbols and abbreviations are explained in Fig. 1.

We also analyzed the chromosomal CJs for nucleotide deletion from the coding ends. None of the TdT mutants analyzed, including TdTL, exhibited a substantial change in the extent of nucleotide deletion compared with reactions lacking TdT (Fig. 3). Therefore, under none of the conditions tested was any mutant of TdT, including TdTL, able to affect nucleotide deletion from coding ends.

Immunofluorescence localization of TdT

To address directly the question of whether TdT variants are properly transported to the nucleus, TdT proteins were expressed by transient transfection in 3TGR cells and localized by immunofluorescence with anti-TdT Abs. A nuclear stain (SYTOX) was used to delineate the nucleus (Fig. 4). As expected, no TdT was detectable in untransfected cells. TdT-FL, TdT-ΔN, TdT-ΔB, and TdT-421 all yielded a strong fluorescence signal with the anti-TdT Ab, and in all cases this signal colocalized with nuclear staining. This indicates that these TdT variants are transported efficiently into the nucleus, and that differences in N-nucleotide addition between different variants is probably not due to differences in subcellular localization. In particular, the putative NLS is not required for nuclear localization of TdT.

Protein-protein interactions

Mahajan et al. (22) reported an interaction between TdT and Ku70/80 and proposed that this association was important for the recruitment of TdT to the V(D)J recombination complex. The first 131 aa of TdT were found to be essential for this interaction. We were interested in determining more precisely which region(s) of TdT was responsible, using coimmunoprecipitation to assess interactions. This effort was greatly facilitated by a novel panel of anti-TdT mAbs (a gift from J. Kearney, University of Alabama, Birmingham, AL). 293T cells were transfected with an Myc-tagged TdT-FL expression vector and lysed 48 h post-transfection, and proteins were immunoprecipitated using the anti-Myc, anti-

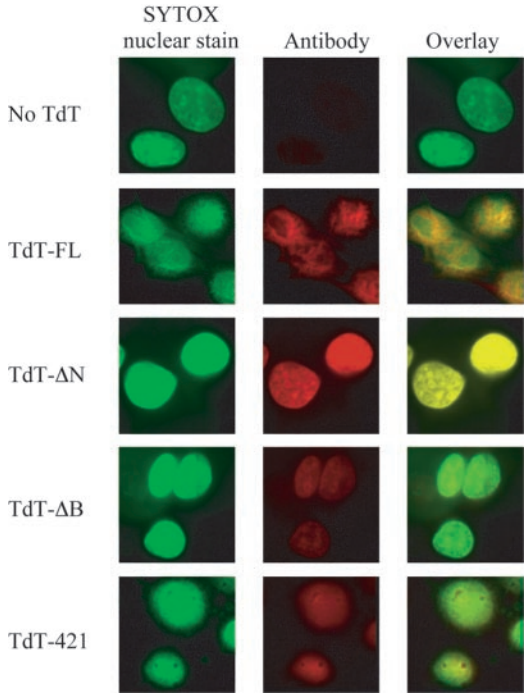


FIGURE 4. Immunofluorescence detection of TdT in transiently transfected 3TGR cells. 3TGR cells were transfected with the TdT expression plasmid indicated to the *left*. The *left panel* shows nuclear staining with SYTOX, the *middle panel* shows immunofluorescence with anti-TdT Abs, and the *right panel* shows the merging of the two signals, emphasizing the nuclear localization of TdT-FL and several mutant forms of TdT.

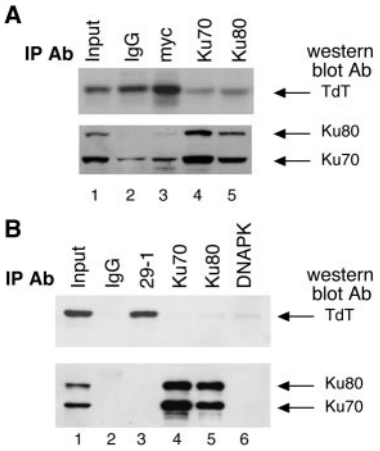


FIGURE 5. Immunoprecipitation (IP) of TdT, Ku, and DNAPKcs. Cell extracts from 293T cells transfected with TdT were immunoprecipitated with the Abs indicated above the lanes. Bead-bound proteins were analyzed by Western blots using anti-TdT or anti-Ku70/80 Abs, as indicated to the *right*. *A*, Extracts were prepared from 293T cells transfected with Myc-tagged TdT, immunoprecipitated with the indicated Abs and protein G beads, and then loaded on the gel without washes of the beads (22). *Lane 1*, 1/10 dilution of the input for the immunoprecipitation; *lane 2*, immunoprecipitation with control IgG; *lane* 3, immunoprecipitation with anti-Myc (TdT); *lane 4*, immunoprecipitation with anti-Ku70; *lane 5*, immunoprecipitation with Ku80. *B*, Extracts were prepared from 293T cells transfected with untagged TdT-FL, immunoprecipitated with the indicated Abs and protein G beads, and loaded on the gel after four washes of the beads. *Lane 1*, 1/20 dilution of the input; *lane 2*, immunoprecipitation with control IgG; *lane* 3, immunoprecipitation with 29-1 (anti-TdT); *lane 4*, immunoprecipitation with anti-Ku70; *lane 5*, immunoprecipitation with Ku80; *lane 6*, immunoprecipitation with anti-DNA-PKcs.

Ku70, anti-Ku80, and a control IgG Ab. The proteins were then analyzed by Western blot analysis using anti-TdT and anti-Ku70/80 Abs. Under mild conditions in which precipitates were not washed, TdT was clearly coprecipitated by the anti-Ku70 and anti-Ku80 Abs (Fig. 5*A*, *lanes 4* and *5*). This was not open to straightforward interpretation, however, because nonspecific IgG also efficiently precipitated TdT (Fig. 5*A*, *lane 2*).

We then repeated the immunoprecipitation experiment, expressing untagged TdT-FL in 293T cells and including several wash steps so as to reduce the nonspecific background. Under these conditions, TdT did not coprecipitate with Ku70 or Ku80 (Fig. 5*B*, *lanes 4* and *5*), nor did Ku70 or Ku80 coprecipitate with an anti-TdT Ab (*lane 3*). These results suggest that the interaction between the Ku complex and TdT in cell extracts does not occur to a significant extent or is of low affinity. It seems unlikely that both the anti-Ku70 and the anti-Ku80 Abs interfere with binding of Ku to TdT. To determine whether the catalytic subunit of the DNA-dependent protein kinase, $DNAPK_{cs}$, interacts with TdT, the immunoprecipitation experiment was repeated using anti-$DNAPK_{cs}$ Abs. Again, no significant interaction between TdT and $DNA\text{-}PK_{cs}$ was detectable (Fig. 5*B*, *lane 6*; the very faint band visible in this lane was not reproducible). To analyze whether the interaction of TdT with components of DNAPK occurs only in the context of V(D)J recombination, similar experiments were performed using cells that were cotransfected with an extrachromosomal recombination substrate. Again, no association of TdT with either the Ku proteins or $DNAPK_{cs}$ was observed (data not shown). Even though protein-protein interactions between TdT and Ku or between TdT and $DNAPK_{cs}$ were not detected, an interaction between these proteins may still be relevant to V(D)J recombination.

Discussion

Role of the HhH, putative NLS, and BRCT domains

This study represents the first extensive analysis of predicted functional domains within TdT. With the V(D)J recombination system used in this study, the putative NLS domain, the BRCT domain, and the first HhH domain were found to be nonessential for TdT activity in V(D)J recombination. Variants of TdT harboring deletions of or point mutations within these motifs (TdT-ΔN, TdT-ΔB, and TdT-ΔH1) show a reduction in the frequency of N-nucleotide additions compared with the wild type protein, suggesting that these mutants are recruited less efficiently to the coding ends within the recombination complex and/or that they associate less tightly with DNA ends. As both of our assays involved overexpression of TdT protein, the effects of inactivating these domains could be more dramatic in V(D)J recombination of the endogenous Ag receptor loci.

It is unclear whether the putative NLS (defined as a basic region at the N terminus) of TdT has any functional significance. Although overexpression in 293T cells might have masked the relevance of this sequence for nuclear transport in our transient transfection assay, TdT-ΔN was able to localize to the nucleus in 3TGR cells and was able to mediate N-nucleotide addition with a chromosomal recombination substrate. Our results indicate a nonessential or redundant role for this motif.

Nearly as enigmatic is the role of the BRCT domain of TdT. N-nucleotide addition by TdT-ΔB was easily detectable, but was reduced compared with TdT-FL in all of our assays. In the transient transfection assay, both the frequency and the length of N-nucleotide addition were reduced by deletion of the BRCT domain, whereas in the chromosomal assay only the frequency was decreased. BRCT domains are found in many DNA repair proteins

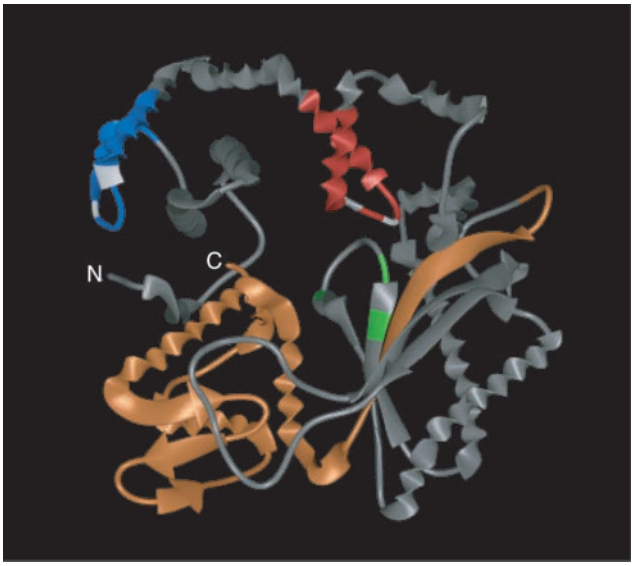


FIGURE 6. Crystal structure of murine TdT (aa 130–510). The two HhH motifs are marked in blue (HhH1, aa 206–227) and red (HhH2, aa 247–267), with the conserved glycine residues marked in white (aa 213, 218, 257, and 259). Active site residues (R336, D343, and D345) are marked in green. The C-terminal domain (aa 421–510) is marked in orange.

and are thought to mediate interactions between them. Our data are consistent with the hypothesis that the TdT BRCT domain is involved in nonessential interactions that help localize and/or retain TdT at the site of coding end repair. Previous findings indicated that the BRCT domain within TdT is required for the interaction between TdT and Ku70 (22). Using a coimmunoprecipitation assay, however, we could not find evidence for an interaction between Ku70 and TdT. Recent biochemical studies from the Ramsden laboratory (24) provide a means of reconciling our findings with those demonstrating a requirement for Ku in N-nucleotide addition (5, 20). Mahajan et al. (24) demonstrated that the combination of DNA ligase IV, XRCC4, and Ku, but not Ku by itself, is able to recruit pol μ or TdT to DNA ends. It is plausible, therefore, that the interaction between TdT and Ku is quite weak (consistent with our data) and that Ku plays a significant role in the recruitment of TdT to coding ends through interactions involving DNA ligase IV and XRCC4. Our results are also consistent with the idea that the BRCT domain of TdT is involved in this interaction and hence in the recruitment of TdT to coding ends, but they clearly demonstrate that this domain is not essential for N-nucleotide addition in V(D)J recombination, at least under the assay conditions used in this study.

The second HhH domain (H2) was found to be essential for TdT activity. In contrast to the dispensable first HhH motif, the second one matches the consensus sequence (GΦG) and was thus thought likely to represent the domain responsible for protein-DNA interactions near the active site. The recent crystal structure of TdT supports this model (30). In this study three carbonyl atoms in the second HhH domain were found to coordinate a sodium ion, and molecular modeling suggested that this sodium ion is directly involved in contacting and positioning of the second nucleotide of the primer strand. Our results support a critical role for the second HhH domain. Based on the crystal structure of TdT, no clear functional role can be assigned to the first HhH domain, and consistent with this, mutating critical residues of the motif did not have a clear deleterious effect on TdT activity. This region of TdT may not have a critical function, or alternatively, this region might perform a function that our assays do not measure. As a similar nonconsensus HhH domain is present in DNA pol μ and λ, it is tempting to speculate that this domain will also prove to be nonessential for their activity.

TdT and end joining involving microhomologies

Two general mechanisms can be envisioned to explain the inhibition of microhomology-mediated joining by TdT. The mere presence of TdT in the repair complex could inhibit the process, perhaps by interfering with some aspect of end alignment. Alternatively, the catalytic activity of TdT might be essential to direct the reaction into a pathway less able to make use of microhomologies. In addition, N-nucleotide addition could prevent the detection of homology-mediated joining, because microhomology-mediated joining involving N-nucleotides at one or both coding ends could not be identified as such. Our data reveal a tight correlation between N-nucleotide addition and suppression of microhomology use. Of particular note is the finding that mutation of the active site of TdT results in a protein that lacks catalytic activity and the ability to suppress microhomology-mediated joining, but presumably retains the ability to be recruited to sites of coding end repair. We conclude that it is likely to be the catalysis of N-nucleotide addition, and not the presence of TdT at coding ends, that suppresses microhomology-mediated joining

Role of the C-terminal domain of TdT

Any truncation of the C-terminal α domain of TdT abolishes polymerase activity. This could be a result of improper protein folding or a conformational change that renders the mutant proteins catalytically inactive. The crystal structure of TdT reveals a strong similarity to the closed conformation of DNA polymerase β. This closed conformation of TdT is stabilized by interactions between N-terminal as 148–163 and the C-terminal residues 498–506 and 508–510 (Fig. 6) (30). The inactivity of truncation mutants TdT-232, TdT-280, TdT-453, and TdT-481 could therefore easily be explained by a loss of exactly those interactions required for maintenance of the closed conformation (although in the case of TdT-232, residues in addition to those involved in this interaction were removed). In the case of pol β, the closed conformation is formed by an induced fit mechanism upon interaction of the incoming free nucleotide with the template strand (45) and thus represents the active state of the polymerase. TdT is not required to distinguish between different nucleotides and hence has been proposed to remain permanently in the active, i.e., closed, state. Our mutagenesis study provides the first experimental support for the model that the closed configuration is the one and only active state of TdT.

TdTL vs TdT-FL

Several lines of evidence support the idea that TdTL modulates the structure of V(D)J coding joints by virtue of an associated exonuclease activity. First, biochemical studies demonstrated a robust 3′ to 5′ exonuclease activity for TdTL, with no evidence of terminal transferase activity (26). Second, TdTL was able to reduce the length of N-nucleotide additions by TdT-FL in transfected CHO cells and in transgenic mice and to enhance nucleotide deletion from coding ends in the CHO cell assay (26, 29). There is, however, evidence for a thermolabile terminal transferase activity associated with TdTL (28), although in no case has N-nucleotide addition in vivo by TdTL been reported previously.

Consistent with previous reports, we found that coexpression of TdTL with TdT-FL reduced the average length of N-nucleotide additions in both 293T and 3T3 cell lines. Curiously, the catalytic mutant TdT-ASM had almost as large an effect on N-nucleotide addition by TdT-FL as did TdTL when the full-length (but not the core) RAG proteins were used. This raises the possibility that the

reduction in N-nucleotide addition caused by TdTL or TdT-ASM is due in part to competition with TdT-FL for the coding ends. Whether competition between TdTL and TdT-FL is an important mechanism regulating N-nucleotide addition during assembly of endogenous Ag receptor genes remains to be elucidated.

Two aspects of our results with TdTL were puzzling, however, given the previous findings mentioned above. First, we found that TdTL is apparently able to catalyze N-nucleotide addition to CJs formed on extrachromosomal plasmids in 293T and CHO-K1 cells, and, to a lesser degree, with integrated chromosomal substrates in 3T3 fibroblasts (although in all cases the level of N-nucleotide addition by TdTL was greatly reduced compared with that of TdT-FL). We speculate that the high levels of expression of TdT achieved in our experiments reveal a weak terminal transferase activity associated with TdTL, consistent with the findings of Boule et al. (28). Similarly, overexpression of TdT may help explain the relatively high frequency of N-nucleotide addition seen at signal joints with TdT-FL in 293 T cells (Fig. 1, *line* 2).

Our second puzzling finding was that the expression of TdTL, either by itself or together with TdT-FL, did not result in an increase in nucleotide deletion from coding ends in human 293T cells (with full-length or core RAG proteins), in CHO-K1 cells, or with the integrated substrate in murine 3TGR cells. This is different from the findings of Thai et al. (26), whose data were derived from transient recombination assays in CHO cells. One possible explanation for the discrepancy is a poorly understood difference in the cell lines or methodologies used in the two studies. Notably, although an average of 2.5 nt were deleted from CJs formed in the absence of TdT in the CHO cells used by Thai et al. (with core RAG proteins) (26), the equivalent values obtained in our experiments were higher: 6.1 and 5.8 nt in 293T cells with full-length and core RAG proteins, respectively; 5.1 nt in our CHO-K1 cells; and 4.7 nt in 3T3 fibroblasts. Hence, our cell lines appear to have a greater propensity to delete nucleotides from coding ends than the cells used by Thai et al. (perhaps reflecting a different balance of nuclease activities), and this could conceivably obscure the activity of TdTL in our cells. A mutation in our TdTL cDNA seems unlikely to provide an explanation, because its sequence was confirmed several times. It will be important to further analyze the activities of TdTL in lymphoid cells. The generation of gene-targeted mice unable to express TdTL might be needed to unambiguously resolve the function of TdTL in generation of the endogenous Ag receptor repertoire.

Acknowledgments

We thank S. D. Fugmann for his contribution to the mutational strategies of TdT and helpful suggestions during the sequence analysis; F. Balamuth for her help with the immunofluorescence experiments; J. Sweasy, P. Cresswell, S. Ghosh, C. Radding, and all members of the Schatz laboratory for helpful suggestions; J. Kearney for comments on the manuscript; and the W. M. Keck Facility at Yale Medical School for oligonucleotide synthesis and DNA sequencing services.

References

1. Tonegawa, S. 1983. Somatic generation of antibody diversity. *Nature 302:575*.
2. Iwasato, T., and H. Yamagishi. 1992. Novel excision products of T cell receptor γ gene rearrangements and developmental stage specificity implied by the frequency of nucleotide insertions at signal joints. *Eur. J. Immunol. 22:101*.
3. Shimizu, T., and H. Yamagishi. 1992. Biased reading frames of pre-existing DH–JH coding joints and preferential nucleotide insertions at VH–DJH signal joints of excision products of immunoglobulin heavy chain gene rearrangements. *EMBO J. 11:4869*.
4. Candeias, S., K. Muegge, and S. K. Durum. 1996. Junctional diversity in signal joints from T cell receptor β and δ loci via terminal deoxynucleotidyl transferase and exonucleolytic activity. *J. Exp. Med. 184:1919*.
5. Purugganan, M. M., S. Shah, J. F. Kearney, and D. B. Roth. 2001. Ku80 is required for addition of N nucleotides to V(D)J recombination junctions by terminal deoxynucleotidyl transferase. *Nucleic Acids Res. 29:1638*.
6. Gilfillan, S., C. Benoist, and D. Mathis. 1995. Mice lacking terminal deoxynucleotidyl transferase: adult mice with a fetal antigen receptor repertoire. *Immunol. Rev. 148:201*.
7. Komori, T., A. Okada, V. Stewart, and F. W. Alt. 1993. Lack of N regions in antigen receptor variable region genes of TdT-deficient lymphocytes. *Science 261:1171*.
8. Gilfillan, S., A. Dierich, M. Lemeur, C. Benoist, and D. Mathis. 1993. Mice lacking TdT: mature animals with an immature lymphocyte repertoire. *Science 261:1175*.
9. Gerstein, R. M., and M. R. Lieber. 1993. Extent to which homology can constrain coding exon junctional diversity in V(D)J recombination. *Nature 363:625*.
10. Lewis, S. M. 1994. The mechanism of V(D)J joining: lessons from molecular, immunological, and comparative analyses. *Adv. Immunol. 56:27*.
11. Feeney, A. J. 1992. Comparison of junctional diversity in the neonatal and adult immunoglobulin repertoires. *Int. Rev. Immunol. 8:113*.
12. Ito, J., and D. K. Braithwaite. 1991. Compilation and alignment of DNA polymerase sequences. *Nucleic Acids Res. 19:4045*.
13. Holm, L., and C. Sander. 1995. DNA polymerase β belongs to an ancient nucleotidyltransferase superfamily. *Trends Biochem. Sci. 20:345*.
14. Manke, I. A., D. M. Lowery, A. Nguyen, and M. B. Yaffe. 2003. BRCT repeats as phosphopeptide-binding modules involved in protein targeting. *Science 302:636*.
15. Yu, X., C. C. Chini, M. He, G. Mer, and J. Chen. 2003. The BRCT domain is a phospho-protein binding domain. *Science 302:639*.
16. Casey, G., S. Plummer, G. Hoeltge, D. Scanlon, C. Fasching, and E. J. Stanbridge. 1993. Functional evidence for a breast cancer growth suppressor gene on chromosome 17. *Hum. Mol. Genet. 2:1921*.
17. Critchlow, S. E., R. P. Bowater, and S. P. Jackson. 1997. Mammalian DNA double-strand break repair protein XRCC4 interacts with DNA ligase IV. *Curr. Biol. 7:588*.
18. Zhang, X., S. Morera, P. A. Bates, P. C. Whitehead, A. I. Coffer, K. Hainbucher, R. A. Nash, M. J. Sternberg, T. Lindahl, and P. S. Freemont. 1998. Structure of an XRCC1 BRCT domain: a new protein-protein interaction module. *EMBO J. 17:6404*.
19. Aoufouchi, S., E. Flatter, A. Dahan, A. Faili, B. Bertocci, S. Storck, F. Delbos, L. Cocea, N. Gupta, J. C. Weill, et al. 2000. Two novel human and mouse DNA polymerases of the polX family. *Nucleic Acids Res. 28:3684*.
20. Bogue, M. A., C. Y. Wang, C. M. Zhu, and D. B. Roth. 1997. V(D)J recombination in Ku86-deficient mice: distinct effects on coding, signal, and hybrid joint formation. *Immunity 7:37*.
21. Gu, Y. S., K. J. Seidl, G. A. Rathbun, C. M. Zhu, J. P. Manis, N. Vanderstoep, L. Davidson, H. L. Cheng, J. M. Sekiguchi, K. Frank, et al. 1997. Growth retardation and leaky scid phenotype of Ku70-deficient mice. *Immunity 7:653*.
22. Mahajan, K. N., L. Gangi-Peterson, D. H. Sorscher, J. S. Wang, K. N. Gathy, N. P. Mahajan, W. H. Reeves, and B. S. Mitchell. 1999. Association of terminal deoxynucleotidyl transferase with Ku. *Proc. Natl. Acad. Sci. USA 96:13926*.
23. Mickelsen, S., C. Snyder, K. Trujillo, M. Bogue, D. B. Roth, and K. Meek. 1999. Modulation of terminal deoxynucleotidyltransferase activity by the DNA-dependent protein kinase. *J. Immunol. 163:834*.
24. Mahajan, K. N., S. A. N. McElhinny, B. S. Mitchell, and D. A. Ramsden. 2002. Association of DNA polymerase μ (pol μ) with Ku and ligase IV: Role for pol μ in end-joining double-strand break repair. *Mol. Cell. Biol. 22:5194*.
25. Doherty, A. J., L. C. Serpell, and C. P. Ponting. 1996. The helix-hairpin-helix DNA-binding motif: a structural basis for non-sequence-specific recognition of DNA. *Nucleic Acids Res. 24:2488*.
26. Thai, T. H., M. M. Purugganan, D. B. Roth, and J. F. Kearney. 2002. Distinct and opposite diversifying activities of terminal transferase splice variants. *Nat. Immunol. 3:457*.
27. Bentolila, L. A., M. F. Dandon, Q. T. Nguyen, O. Martinez, F. Rougeon, and N. Doyen. 1995. The two isoforms of mouse terminal deoxynucleotidyl transferase differ in both the ability to add N regions and subcellular localization. *EMBO J. 14:4221*.
28. Boule, J. B., F. Rougeon, and C. Papanicolaou. 2000. Comparison of the two murine deoxynucleotidyltransferase terminal isoforms: a 20-amino acid insertion in the highly conserved carboxyl-terminal region modifies the thermosensitivity but not the catalytic activity. *J. Biol. Chem. 275:28984*.
29. Benedict, C. L., S. Gilfillan, and J. F. Kearney. 2001. The long isoform of terminal deoxynucleotidyl transferase enters the nucleus and, rather than catalyzing nontemplated nucleotide addition, modulates the catalytic activity of the short isoform. *J. Exp. Med. 193:89*.
30. Delarue, M., J. B. Boule, J. Lescar, N. Expert-Bezancon, N. Jourdan, N. Sukumar, F. Rougeon, and C. Papanicolaou. 2002. Crystal structures of a template-independent DNA polymerase: murine terminal deoxynucleotidyltransferase. *EMBO J. 21:427*.
31. Sukumar, N., J. B. Boule, N. Expert-Bezancon, N. Jourdan, J. Lescar, F. Rougeon, C. Papanicolaou, and M. Delarue. 2000. Crystallization of the catalytic domain of murine terminal deoxynucleotidyl transferase. *Acta Crystallogr. D 56:1662*.
32. Doublie, S., M. R. Sawaya, and T. Ellenberger. 1999. An open and closed case for all polymerases. *Structure 7:R31*.
33. Yang, B., K. N. Gathy, and M. S. Coleman. 1994. Mutational analysis of residues in the nucleotide binding domain of human terminal deoxynucleotidyl transferase. *J. Biol. Chem. 269:11859*.

34. Steitz, T. A., and J. A. Steitz. 1993. A general two-metal-ion mechanism for catalytic RNA. *Proc. Natl. Acad. Sci. USA 90:6498*.
35. Fugmann, S. D., and D. G. Schatz. 2001. Identification of basic residues in RAG2 critical for DNA binding by the RAG1-RAG2 complex. *Mol. Cell. 8:899*.
36. Lieber, M. R., J. E. Hesse, S. Lewis, G. C. Bosma, N. Rosenberg, K. Mizuuchi, M. J. Bosma, and M. Gellert. 1988. The defect in murine severe combined immune deficiency: joining of signal sequences but not coding segments in V(D)J recombination. *Cell 55:7*.
37. Koiwai, O., T. Yokota, T. Kageyama, T. Hirose, S. Yoshida, and K. Arai. 1986. Isolation and characterization of bovine and mouse terminal deoxynucleotidyltransferase cDNAs expressible in mammalian cells. *Nucleic Acids Res. 14:5777*.
38. Lieber, M. R., J. E. Hesse, K. Mizuuchi, and M. Gellert. 1988. Lymphoid V(D)J recombination: nucleotide insertion at signal joints as well as coding joints. *Proc. Natl. Acad. Sci. USA 85:8588*.
39. Chang, L. M., and F. J. Bollum. 1986. Molecular biology of terminal transferase. *CRC Crit. Rev. Biochem. 21:27*.
40. McMahan, C. J., M. J. Difilippantonio, N. Rao, E. S. Spanopoulou, and D. G. Schatz. 1997. A basic motif in the N-terminal region of RAG1 enhances recombination activity. *Mol. Cell. Biol. 17:4544*.
41. Roman, C. A. J., S. R. Cherry, and D. Baltimore. 1997. Complementation of V(D)J recombination deficiency in RAG-$1^{-/-}$ B cells reveals a requirement for novel elements in the N-terminus of RAG-1. *Immunity 7:13*.
42. Steen, S. B., J. O. Han, C. Mundy, M. A. Oettinger, and D. B. Roth. 1999. Roles of the "dispensable" portions of RAG-1 and RAG-2 in V(D)J recombination. *Mol. Cell. Biol. 19:3010*.
43. Schatz, D. G., and D. Baltimore. 1988. Stable expression of immunoglobulin gene V(D)J recombinase activity by gene transfer into 3T3 fibroblasts. *Cell 53:107*.
44. Schatz, D. G., M. A. Oettinger, and D. Baltimore. 1989. The V(D)J recombination activating gene (RAG-1). *Cell 59:1035*.
45. Sawaya, M. R., R. Prasad, S. H. Wilson, J. Kraut, and H. Pelletier. 1997. Crystal structures of human DNA polymerase β complexed with gapped and nicked DNA: evidence for an induced fit mechanism. *Biochemistry 36:11205*.