Available online at www.sciencedirect.com

ScienceDirect

Current Opinion in Structural Biology

Terminal deoxynucleotidyltransferase: the story of an untemplated DNA polymerase capable of DNA bridging and templated synthesis across strands

Jérôme Loc'h and Marc Delarue

Terminal deoxynucleotidyltransferase (TdT) is a member of the polX family which is involved in DNA repair. It has been known for years as an untemplated DNA polymerase used during V(D)J recombination to generate diversity at the CDR3 region of immunoglobulins and T-cell receptors. Recently, however, TdT was crystallized in the presence of a complete DNA synapsis made of two double-stranded DNA (dsDNA), each with a 3′ protruding end, and overlapping with only one micro-homology base-pair, thus giving structural insight for the first time into DNA synthesis across strands. It was subsequently shown that TdT indeed has an *in trans* template-dependent activity in the presence of an excess of the downstream DNA duplex. A possible biological role of this dual activity is discussed.

Address
Unit of Structural Dynamics of Biological Macromolecules and UMR 3528 du CNRS, Institut Pasteur, 75015 Paris, France

Corresponding author: Delarue, Marc (marc.delarue@pasteur.fr)

Current Opinion in Structural Biology 2018, **53**:22–31

This review comes from a themed issue on **Protein–nucleic acid interactions**

Edited by **Eric Westhof** and **Dinshaw Patel**

https://doi.org/10.1016/j.sbi.2018.03.019

Introduction

Terminal deoxynucleotidyltransferase (TdT) was one of the first eukaryotic DNA polymerases purified in the early 1960s [1], from calf thymus extracts. However, instead of the expected classical *templated* polymerase activity, the biochemical characterization of TdT revealed an efficient *untemplated* polymerase (nucleotidyltransferase) activity [2,3], especially in the presence of divalent transition metal ions [4]. *In vivo*, the function of TdT was only fully understood in the eighties [5–7], after the discovery of V(D)J recombination [8–10]. During this process, TdT adds random nucleotides (N-segments) at the V-D and D-J junctions in heavy chains of immunoglobulins (Ig) and T-cell receptors, thereby contributing significantly to the diversity of the immune repertoire [11,12]. Subsequently, it was revealed that the V(D)J uses the same machinery [13,14] as the one of Non-Homologous End Joining (NHEJ) that repairs DNA double-strand breaks (DSB). This machinery includes a recognition complex (Ku heterodimer, DNA-PKcs), DNA end-processing enzymes such as a nuclease (Artemis or Metnase) and a DNA polymerase (pol X), as well as a ligation complex (Lig IV, XRCC4, XLF) [15]. The DNA polymerase is a member of the family polX that includes not only TdT, but also pol λ [16•] and pol μ [17•,18•], the last two participating to both NHEJ and V(D)J recombination [19–23]. All three polymerase domains X-ray structures have been determined to high resolution [24•,25•,26•] but the only one that was crystallized in a DNA-bridging context is TdT [27••,28••]. Here we focus on TdT and on the biological implications of these new structures.

Structural and biochemical features of pol X family polymerases

The X-family DNA polymerase (polX) is specialized in DNA repair. This family is composed of four different DNA polymerases: pol β, pol λ, pol μ and TdT. Only three members of the polX family possess an N-terminal BRCT (BRCA1 carboxy-terminal) domain (Figure 1a) that is essential for NHEJ activity both *in vitro* and *in vivo* [29•,30]. Pol β participates only in base excision repair (BER) [31,32] and is devoid of this small domain (around 11 kDa), which mediates protein–protein or protein–DNA interactions [33]. Three individual structures of a BRCT domain from the polX family were solved by NMR [34,35]. BRCT domain is typically an αβα sandwich made up of a central antiparallel β-sheet flanked by three α-helices [36–38]. Although the structures of all known BRCT domains are highly conserved, their interaction mode with ligands is greatly variable and their role in the sequential recruitment of the different proteins during NHEJ repair is not completely understood [39]. Ligase IV contains 2 BRCT domains whose peptide junction interacts with XRCC4, which otherwise makes filaments with XLF [40,41] and interacts with Ku heterodimer [42]. The BRCT domain of pol μ binds DNA [43] but the precise mode of binding of the BRCT domain of either pol μ or TdT to Ku heterodimer remains to be elucidated, as well as its orientation with respect to the catalytic polymerase domain.

Sequence analysis of the polymerase domain of polX polymerases allows to divide this family into two

subgroups [44•]. Pol β is closer to pol λ (34% of pairwise sequence identity) and TdT is closer to pol μ (44% of pairwise sequence identity). The structures of all four polymerases belonging to X-family have been solved by X-ray crystallography. Pol β was the first polX to be solved, alone and with different binary and ternary complexes [45•,46•,47••]. The overall structure of the catalytic domain shares the same general architecture (but not the topology) of all DNA polymerases, namely a finger domain, a palm domain and a thumb domain (Figure 1). The additional 8-kDa domain in pol β and pol λ contains a deoxyribose phosphate (dRP) lyase activity required in base excision repair (BER) of oxidative DNA damage (Figure 1). The amino acids necessary for dRP lyase activity are not conserved in TdT and pol μ, which do not participate in BER.

One striking feature common to all polX is the high degree of conservation of the catalytic site, with three strictly conserved Aspartates (Figure 1b) that coordinate two essential metal ions, involved in the so-called two-metal ions mechanism, first described in [48] and later shown to be present also in DNA polymerases [47••]. Metal A activates the 3′OH of the last nucleotide to allow the attack of the alpha phosphate while Metal B that comes in with the incoming nucleotide triphosphate stabilizes the leaving group (PPi) (Figure 1b). In TdT, the coordination geometry of the divalent metal ions was studied in atomic detail during a full catalytic site, including transition metal ions such as Mn^{++}, Co^{++} and Zn^{++} which are known to be more efficient than Mg^{++} for the nucleotidyltransferase activity. It was concluded that Metal A has to leave and be replaced by Na^{+} in order to allow translocation of the newly extended primer strand into a catalytically competent position for a new addition [49•]. A movie of the reaction cycle, based on thirteen different structures was built [49•]. This scenario was also described in even greater details for pol β where it was also found that binding of Na^{+} in Metal A binding site, after nucleotide incorporation, is a key step for DNA translocation [50••]. In addition, time-lapse crystallography showed that in pol β, there is an additional divalent ion (a third Mg^{++}, Metal C) that comes during the reaction to counter-balance the apparition of a charge on the beta phosphate, and then leaves before the end-state is reached [50••,51•]. This situation was also observed for pol μ, in the case of Mn^{++} ions [52•].

Common features of TdT, pol μ and pol λ

TdT, pol μ and pol λ remain in a closed conformation throughout their catalytic cycle, contrary to pol β [24•,53,54•]. One possible explanation for this observation is that they have traded fidelity (which requires open-to-closed transition) for a very tight binding of the DNA synapsis, a very fragile structure. One unique feature revealed when the first X-ray structure of TdT was solved [24•], is a specific Loop (Loop1), composed of 20 amino acids (382-401), located between the β3 and β4 strands (Figure 1b), that prevents the binding of a 5′ overhang of the template strand. More than 30 structures of TdT (wild-type or mutants in different complexes) are available on PDB and in all of them, Loop1 adopts the same lariat-like conformation that prevents the binding of an *uninterrupted* template strand on TdT (Figure 2). Perhaps somewhat deceptively, in all known structures of pol μ, Loop1 (also about 20 amino acids long) is invisible in the electron density map, meaning that this region is disordered (Figures 2 and 3a). In pol λ, Loop1 is comparatively shorter (8 aa), but still longer than in pol β (Figure 3). However, pol λ has an additional Loop, called Loop3, that, interestingly, is located precisely where another form of TdT resulting from alternative splicing has an insertion of 20 additional residues [55,56] (Figure 3b) and where it is ideally placed to control bulges or insertions just before the *in trans* templating base [28••] (Figure 3a). Mutations experiments have consistently shown the importance of Loop1 for the substrate specificity not only in TdT [27••,57•] but also in pol μ [58•,59,60•], as well as pol λ [61•].

Early sequence comparisons in a structural context helped to define two important regions for the specificity of TdT versus pol μ [24•], later named SD1 and SD2 [57•] (Figure 3): they are located at the C-terminal border of Loop1, and in a β-turn-β structure close to Loop1, which can also bind an extra Zn^{++} ion [49•] but the precise role of this additional divalent ion is currently unknown. Mutations in these two regions profoundly affect the activity of both TdT [27••,28••] and pol μ [60•]. Mutation of only one amino acid in SD1 region (F401A) confers to Tdt an *in cis* templated polymerase activity, even in the presence of Co^{++} [57•].

A remaining puzzle concerns the conformation of Loop1 in pol μ, and its role in binding the DNA synapsis substrate. Experiments are currently underway in our lab using a TdT-pol μ chimera to determine the conformation of Loop1 of pol μ and they indeed suggest that Loop1 plays the crucial role of a gate that can be open or closed (disordered or ordered) when pol μ searches for a micro-homology region across a DNA synapsis (Loc'h et al., submitted).

Nucleotidyltransferase activity of TdT

Extensive biochemical experiments have demonstrated that TdT can add random deoxyribonucleotides (dNTPs) on a ssDNA primer, which has to be at least three nucleotides long, in a template-independent manner [2]. *In vitro* experiments show that TdT can use all four natural dNTPs with a preferential incorporation of dCTP and dGTP compared to dATP and dTTP [3]. Pol μ also has a significant nucleotidyltransferase activity in the presence of Mn^{++} [57•,62•,63•]. Interestingly, TdT nucleotide binding site can accommodate both deoxyribonucleotide and ribo-nucleotide triphosphates (dNTPs