Supporting Information

- Structural evidence for an *in trans* base selection mechanism involving Loop1 in Polymerase mu at an NHEJ double-strand break junction
-

- **Jérôme Loc'h[1], Christina A. Gerodimos[2], Sandrine Rosario[1], Mustafa Tekpinar[1], Michael R. Lieber[2] and Marc Delarue[1&]**
-
- [1]Unité de Dynamique Structurale des Macromolécules, Institut Pasteur; UMR 3528 du C.N.R.S., 25 rue du Dr Roux, 75015 Paris, France.
- [2]Departments of Pathology, Biochemistry & Molecular Biology, and Molecular Microbiology & Immunology and Department of Biological Sciences, Section of Molecular & Computational Biology, Norris Comprehensive Cancer Center, University of Southern California Keck School of Medicine, Los Angeles, California 90033.

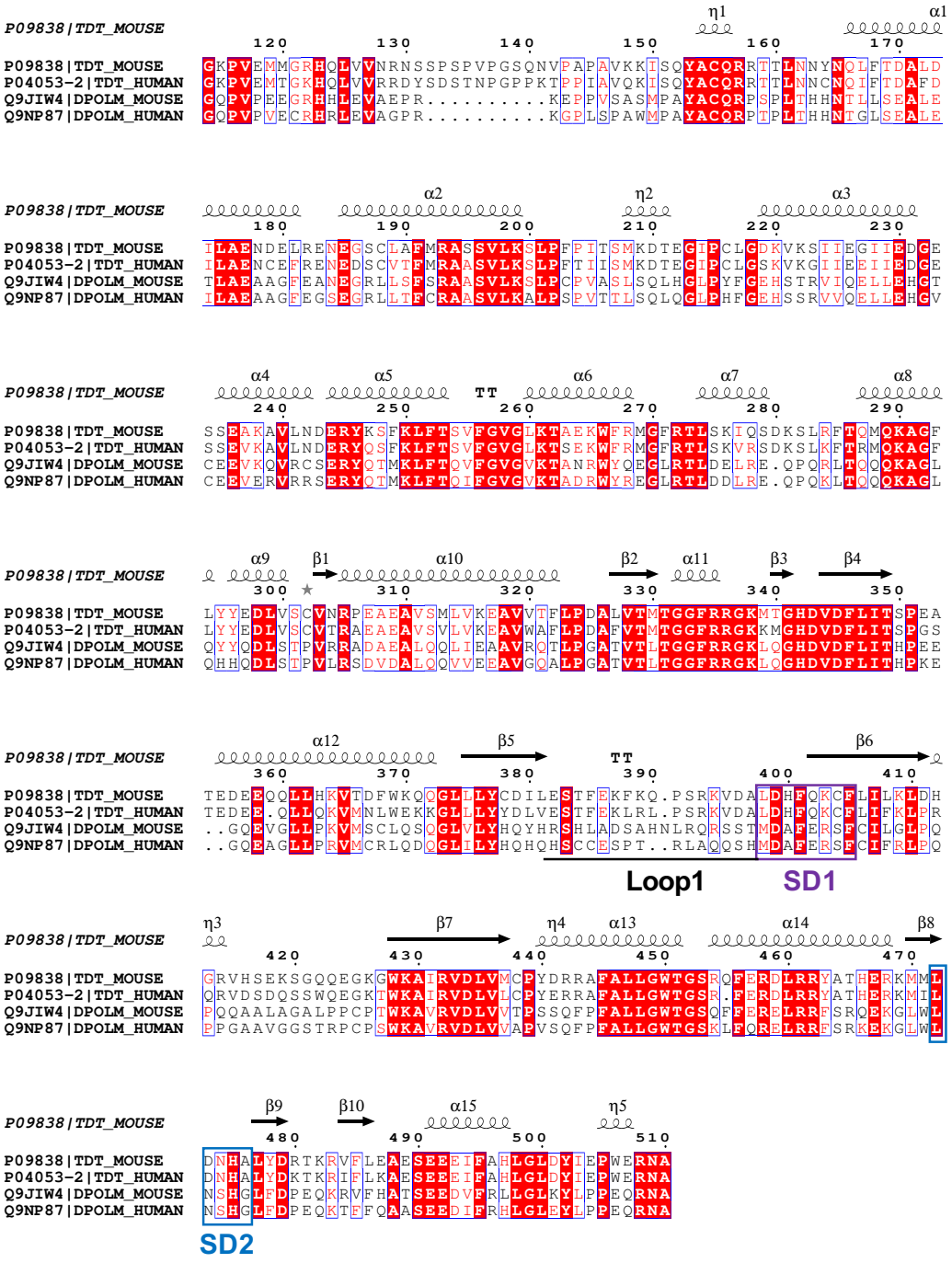


Figure S1. Sequence alignment of mouse and human TdT and Pol mu. SD1, SD2 and Loop1 regions are indicated. Conserved residues are boxed in red.

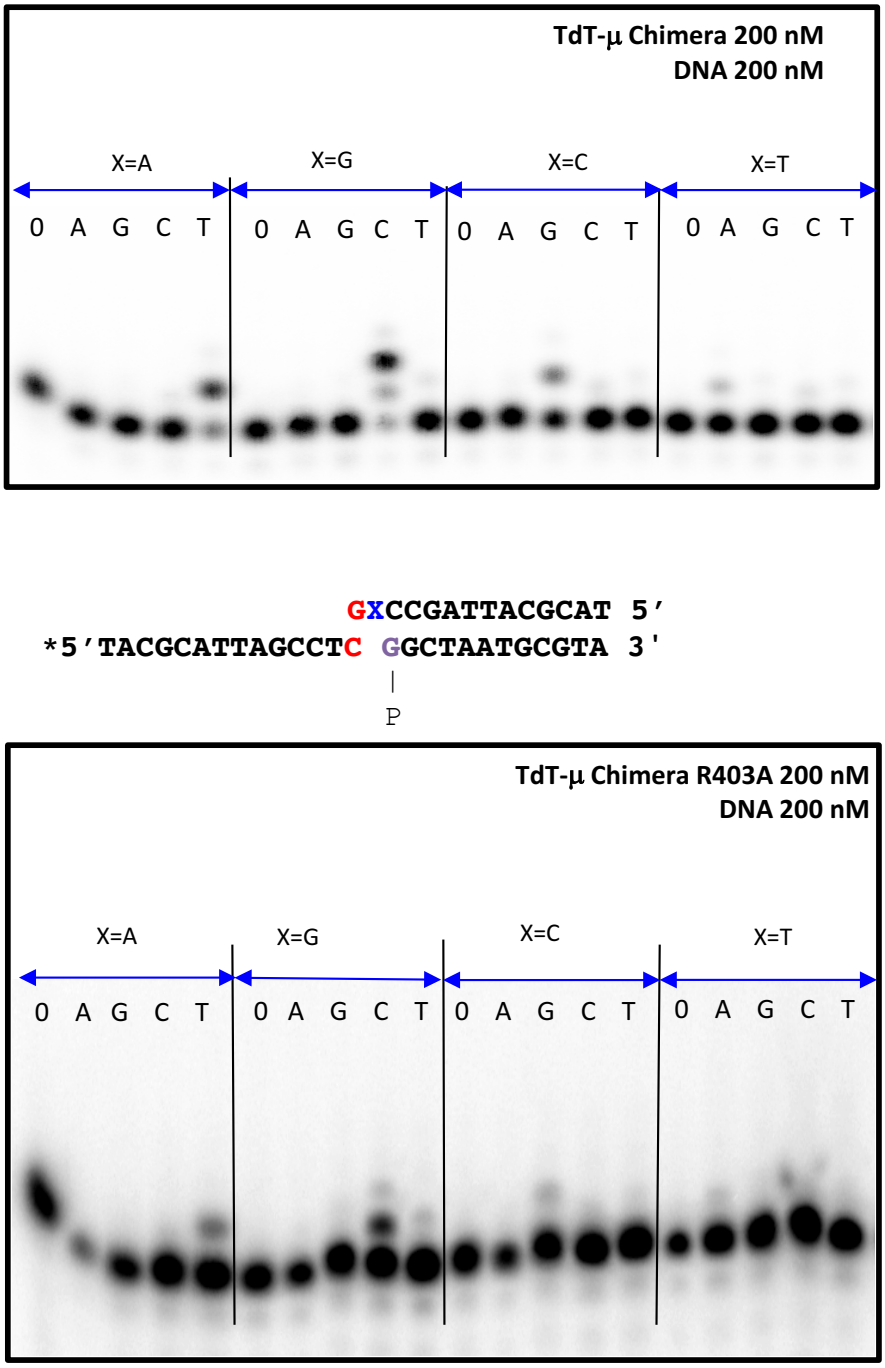


Figure S2. Functional tests of *in trans* templated activity. Top panel, TdT-μ chimera. Lower panel R403A mutant. The DNA substrate is shown in the middle panel, with the micro-homology base pair in red and the templating base *in trans* in blue (X). The concentration of both the enzyme and the DNA is 200 μM.

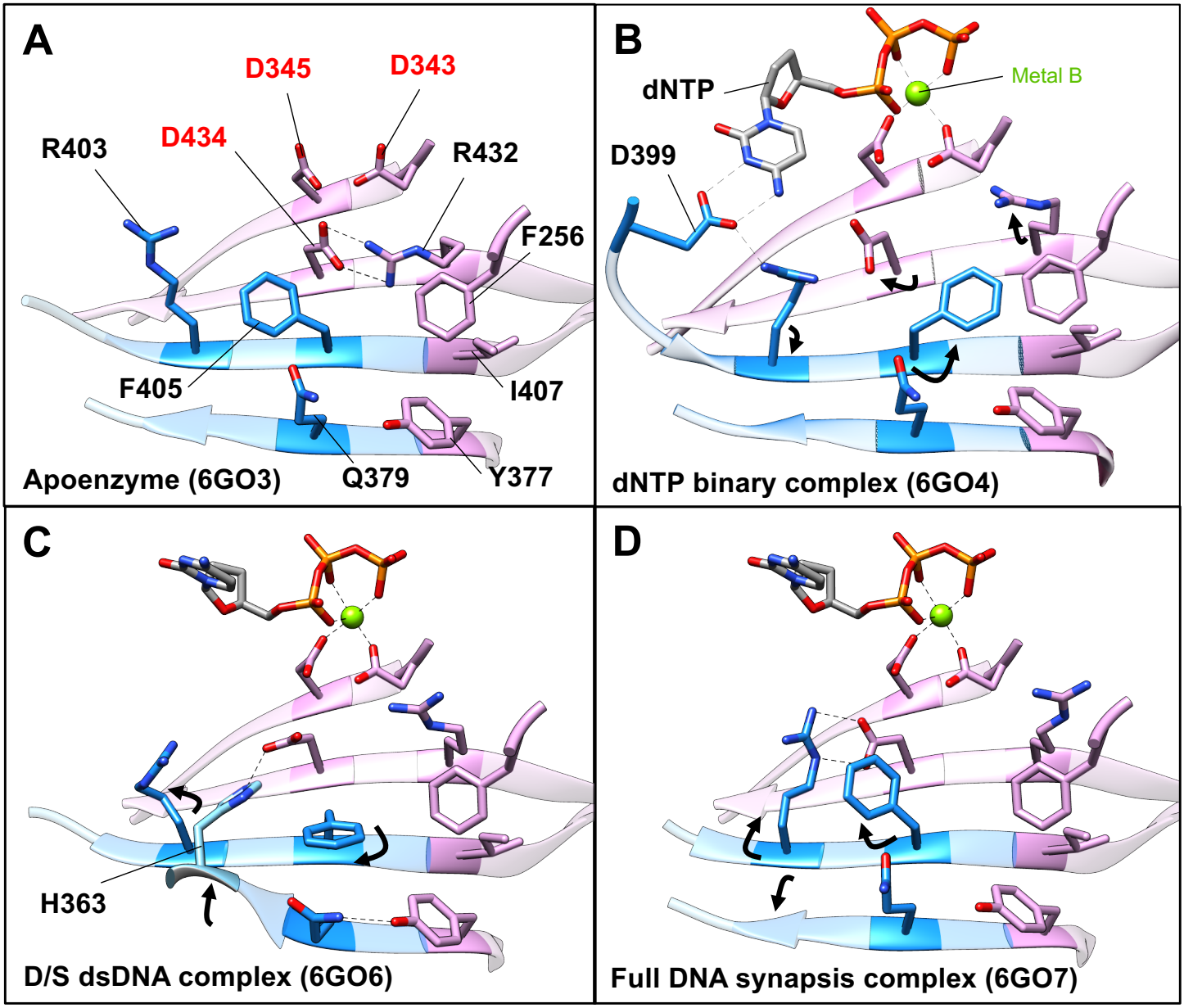


Figure S3. Reorganization of ion pairs in the catalytic site. (A) Apoenzyme **(B)** Binary complex with the incoming dNTP **(C)** Complex with the downstream DNA duplex **(D)** Full DNA synapsis complex. Residues are in pink for TdT and in cyan for Loop1 of Pol μ. The catalytic ion is in CPK sphere in green. The incoming dNTP is in ball-and-stick. Catalytic residues are labeled in red.

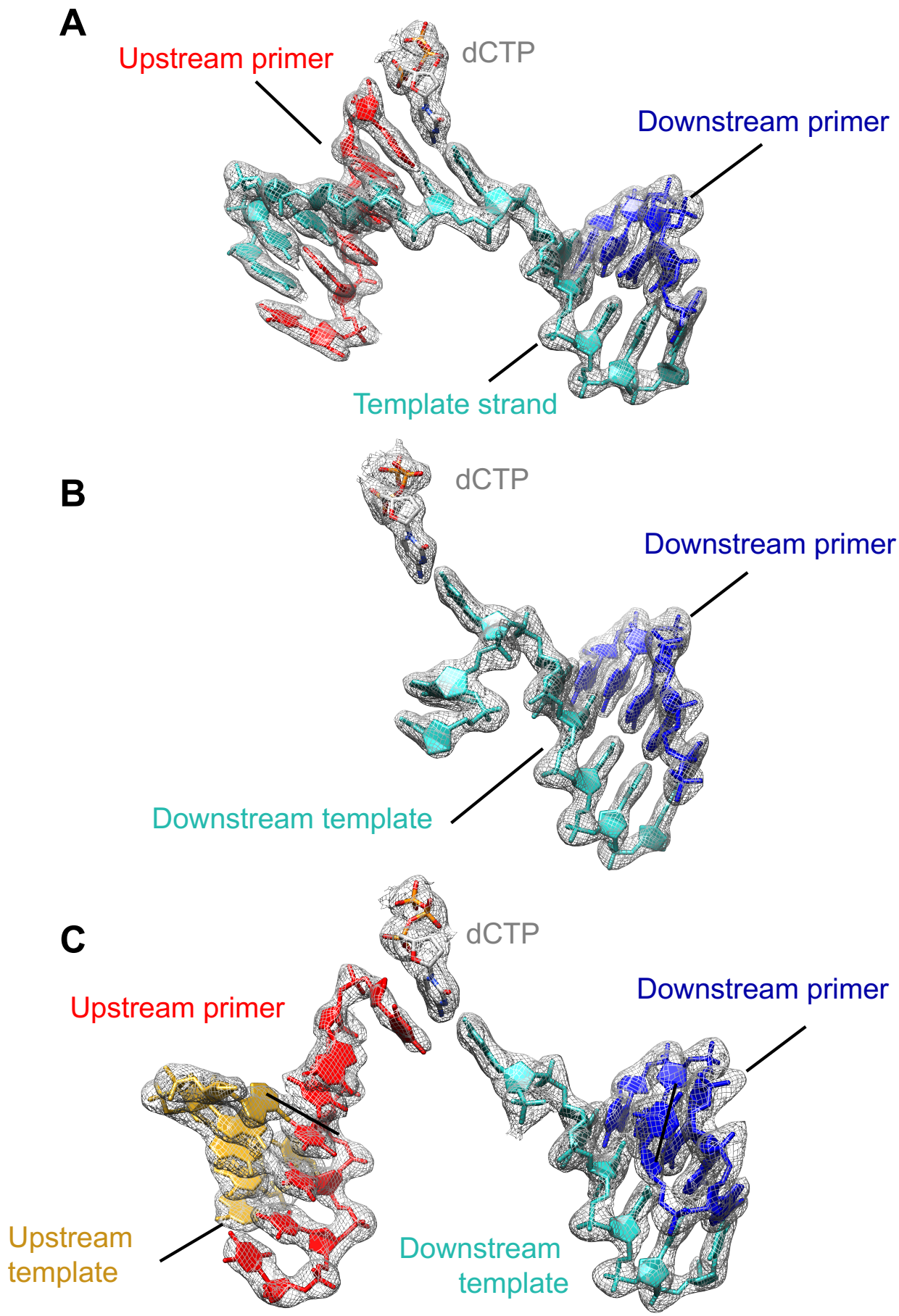


Figure S4. Electron density of DNA substrates in various TdT-μ chimera complexes
DNA and incoming nucleotide electron density in the 2Fo-Fc map (contoured at 1 σ in grey) of (**A**) One nucleotide gap-filling complex (**B**) Pre-ternary complex (**C**) Pre-catalytic full complex. Upstream and downstream primers are depicted in red and blue, respectively. The downstream template strand is represented in cyan, the upstream template strand is in yellow and the incoming nucleotide is in grey.

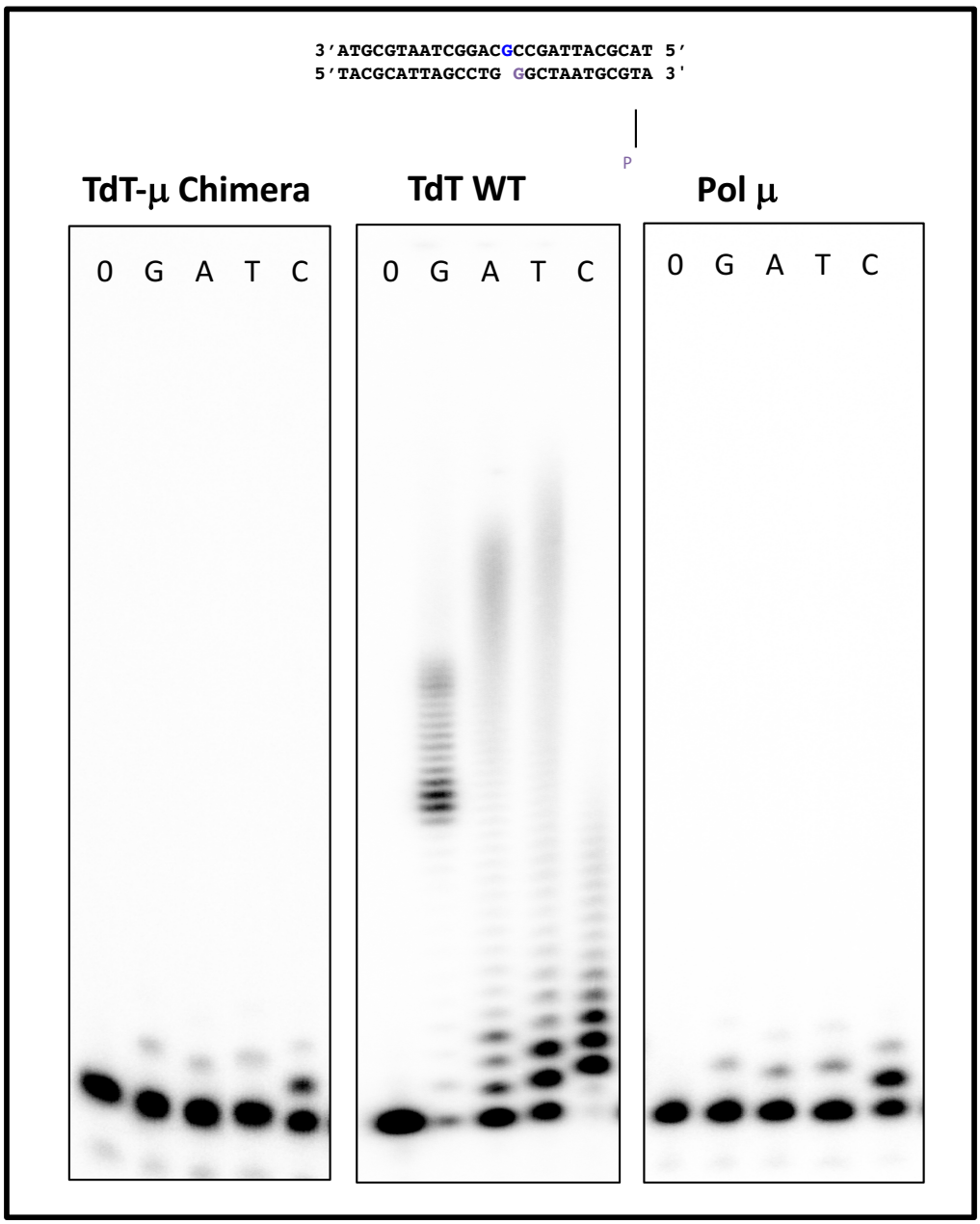


Figure S5. Gap-filling activity test for Tdt-μ chimera, TdT and Pol μ
The concentration of the enzyme was 200 μM, for DNA it was 50 nM for pol μ or TdT and 200 μM for Tdt-μ chimera. The templating base is in blue.

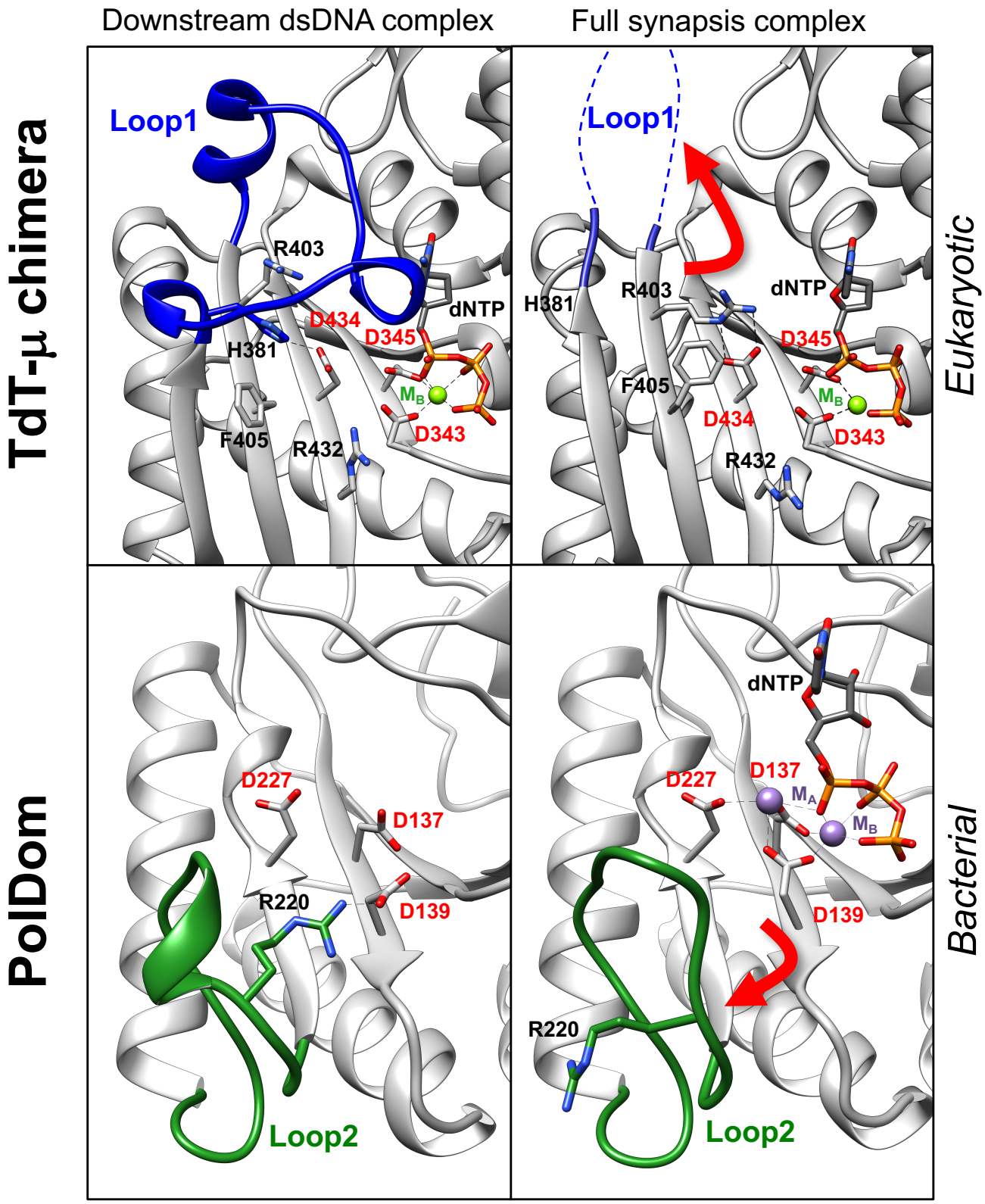
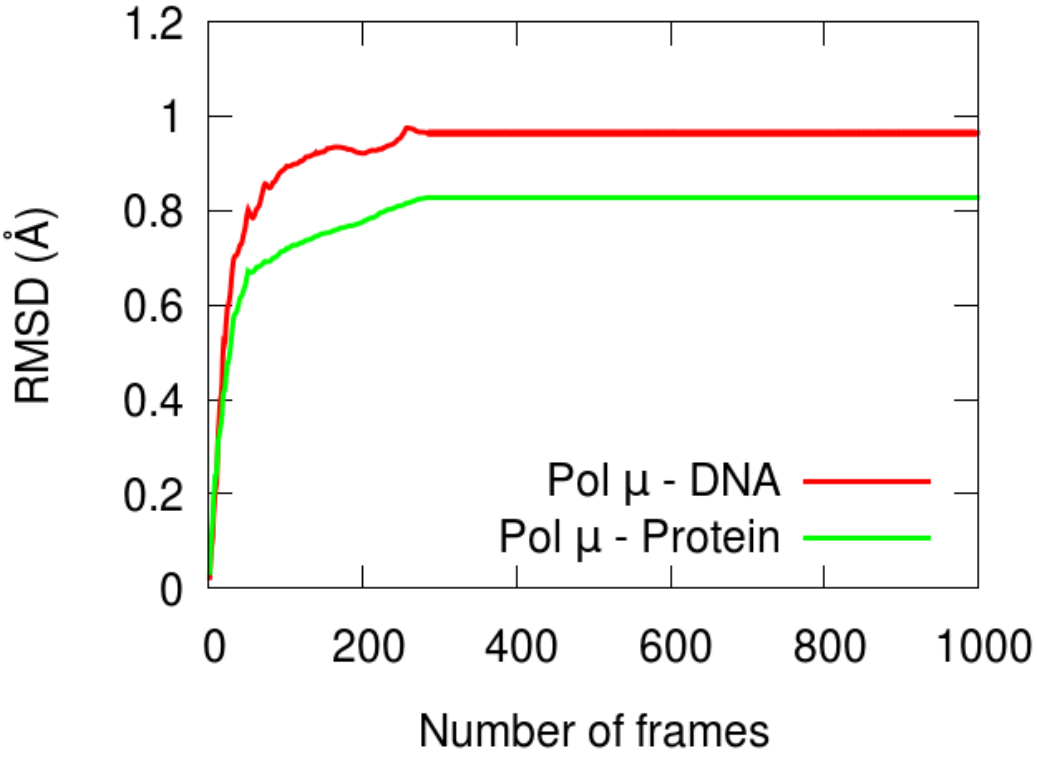


Figure S6. Role of Loop1 in TdT-μ chimera and Loop2 of *M. tuberculosis* PolDom in the catalytic site organization

Schematic representation of the active site and Loop1 (in blue) of TdT-μ chimera (top panel) or the active site and Loop2 (in green) of PolDom (bottom panel). The catalytic ions are represented as CPK spheres, the incoming nucleotide is in ball-and-stick and important catalytic residues are labeled and represented in sticks.



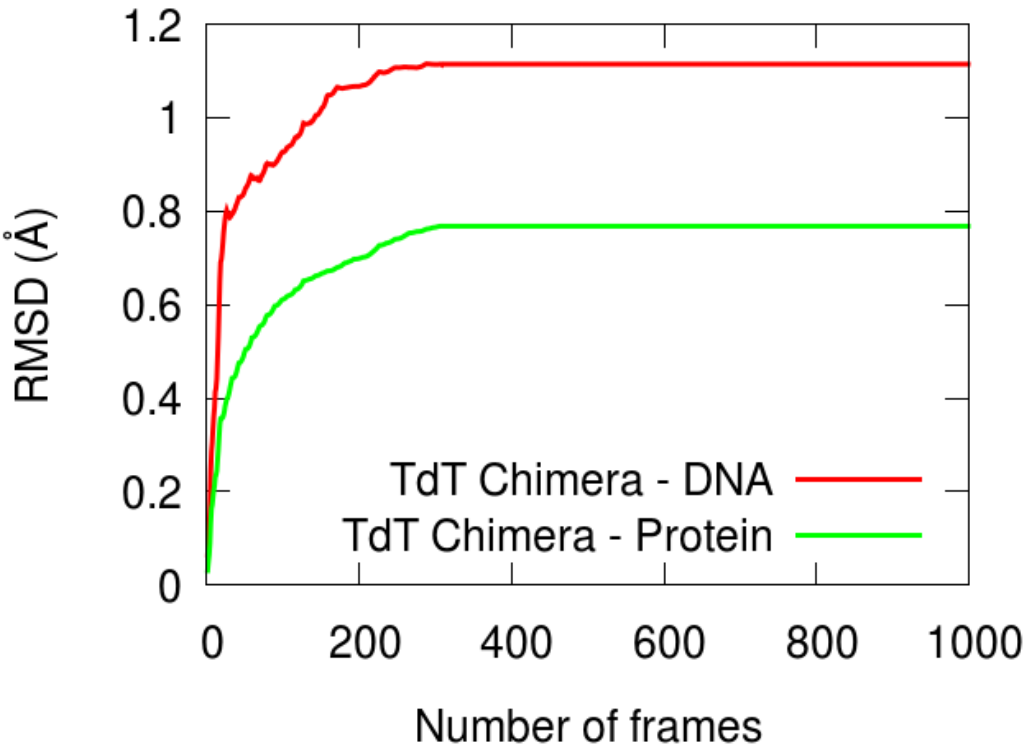


Figure S7. Rmsd during Energy minimization (Å). Top: pol μ homology model. Bottom: x-ray structure of TdT-μ chimera. The RMSD on DNA atoms is in red, the RMSD on protein atoms is in green.

Name	Apoenzyme	Binary complex	Gap-filling complex	Preternary complex	DSB DNA complex
Protein	Tdt-μ chimera	Tdt–μ chimera	Tdt-μ chimera	Tdt-μ chimera	Tdt-μ chimera
dNTP	/	ddCTP	dCpcpp	ddCTP	ddCTP
DNA	/	/	Gap filling DNA	Downstream dsDNA	DNA with DSB
Ions	/	1 Mg^{++}, 1 Na^{+}	1 Mg^{++}, 1 Na^{+}	1 Mg^{++}, 1 Na^{+}	1 Mg^{++}, 1 Na^{+}
PDB Code	6GO3	6GO4	6GO5	6GO6	6GO7
Wavelength (Å)	0.9785	0.9785	0.9762	0.9724	0.8728
Oscillation (°)	0.1	0.1	0.1	0.1	0.15
No. Frames	1-2000	1-2000	1-1500	1-1500	1-650
Reflections measured	179883 (28582)	242086 (37302)	110442 (17686)	246347 (37344)	32249 (5126)
Reflections Unique	25171 (3964)	34181 (5333)	39224 (6269)	45742 (6958)	16488 (2605)
Space group	P212121	P212121	C2	P21212	P21
Cell parameters (Å)	48.87 85.28 119.71	46.97 85.62 114.91	234.18 69.25 59.69	86.42 189.83 46.48	47.00 90.31 65.90
Angles (°)	90. 90. 90.	90. 90. 90.	90. 95.20 90.	90. 90. 90.	90. 106.79 90.
Resolution (Å) (last shell)	46.81 – 2.20 (2.33 – 2.20)	47.78 – 1.96 (2.07 – 1.96)	44.86 – 2.35 (2.49 – 2.35)	47.46 – 2.09 (2.22 – 2.09)	44.99 – 2.55 (2.70 – 2.55)
Completeness (%)	99.7 (98.8)	99.7 (98.5)	98.2 (98.0)	98.9 (94.1)	94.7 (93.9)
Multiplicity	7.1	6.5	2.8	5.4	2.0
I/Sigma(I)	9.57 (1.14)	11.02 (1.02)	10.54 (1.08)	15.79 (1.22)	5.30 (1.05)
Rmerge (%)	13.5 (161.2)	10.1 (166.6)	6.5 (81.3)	5.7 (107.0)	13.5 (76.2)
CC ½ (%)	99.7 (51.2)	99.8 (49.8)	99.8 (57.5)	99.9 (61.2)	98.3 (55.4)
Rfactor (%)	20.1	20.1	21.9	20.2	19.2
Rfree (%)	25.1	24.4	26.5	23.2	22.0
No. of protein atoms	2762	2830	5218	2879	2665
No. of DNA atoms	0	/	935	425	772
No. of water molecules	253	359	327	329	180
B factor overall ($Å^2$)	63.64	63.66	75.99	68.63	53.67
B factor DNA ($Å^2$)	/	/	92.21	94.58	70.65
Ramachandran Favored (%)	98	99	96-98	99	99
Allowed (%)	2	1	4-2	1	1
Outliers (%)	0	0	0	0	0
RMSD bond lengths (Å)	0.010	0.010	0.009	0.010	0.010
RMSD bond angles (°)	1.05	1.01	1.05	0.97	1.05
Molprobity score	100^{th}	100^{th}	99^{th}	100^{th}	100^{th}

Table S1. Diffraction data and model refinement statistics.