Pol μ
Full-length

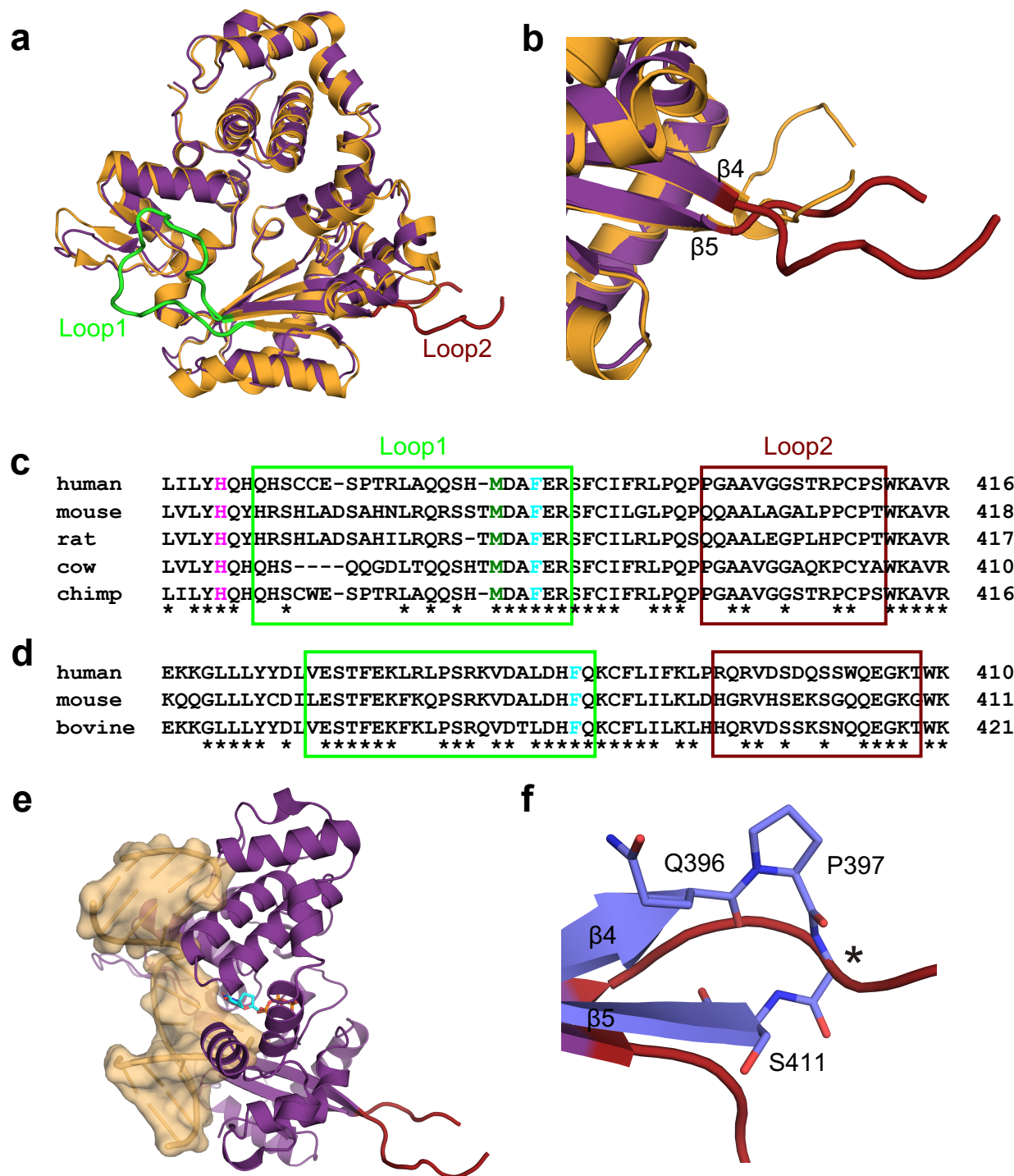Pol μ
cat. domain

C

17.5 30.1 15.5 32.5 % primer extended

Supplementary Figure 1: Comparison of wildtype full-length and catalytic domain (Pro138-Ala494) constructs of hPol μ in single-nucleotide gap-filling. A steady-state polymerase activity assay was performed in order to compare wildtype full-length or truncated Pol μ constructs (5nM), on a single-nucleotide gapped DNA substrate (no enzyme control, C). The percentage of primer extension was calculated for each construct, at two different time points. The two constructs behaved indistinguishably in this assay.

a

b

β4

β5

Loop1

Loop2

c

<div>

Loop1   Loop2

```
human  LILYHQHQHSCCE-SPTRLAQQSH-MDAFERSFCIFRLPQPPGAAVGGSTRPCPSWKAVR  416
mouse  LVLYHQYHRSHLADSAHNLRQRSSTMDAFERSFCILGLPQPQQAALAGALPPCPTWKAVR  418
rat    LVLYHQYHRSHLADSAHILRQRS-TMDAFERSFCILRLPQSQQAALEGPLHPCPTWKAVR  417
cow    LVLYHQHQHS----QQGDLTQQSHTMDAFERSFCIFRLPQPPGAAVGGAQKPCYAWKAVR  410
chimp  LILYHQHQHSCWE-SPTRLAQQSH-MDAFERSFCIFRLPQPPGAAVGGSTRPCPSWKAVR  416
       * ****  *       * * *  ******* ***  ***  **   *  **  *****
```

</div>

d

<div>

```
human  EKKGLLLYYDIVESTFEKLRLPSRKVDALDHFQKCFLIFKLPRQRVDSDQSSWQEGKTWK  410
mouse  KQQGLLLYCDILESTFEKFKQPSRKVDALDHFQKCFLILKLDHGRVHSEKSGQQEGKGWK  411
bovine EKKGLLLYYDIVESTFEKFKLPSRQVDTLDHFQKCFLILKLHHQRVDSSKSNQQEGKTWK  421
       ***** * ****** *** ** ********* ** ** * * **** **
```

</div>

e

f

Q396   P397

β4

*

β5

S411

Supplementary Figure 2: Protein engineering of hPol μ Δ2. (a) Comparison of the X-ray crystal structures of mouse Pol μ (PDB ID code 2IHM[1], purple) and mouse TdT (PDB ID code 1JMS[2], orange). The relative positions of Loop 1 (green, from mTdT) and Loop 2 (red, from mPol μ) are displayed. (b) Ribbon diagram showing the ordered and disordered regions of Loop 2 in mTdT (orange) and mPol μ (red). (c) ClustalW[3] sequence alignments of Loop1 and Loop 2 in mammalian orthologs of Pol μ. Loop1 is boxed in green. Regions of decreased sequence conservation in Loop2 are boxed in red, and were subsequently deleted by site-directed mutagenesis. His363 (magenta), Met382 (green), and Phe385 (cyan) are clearly marked. (d) ClustalW sequence alignments of Loop 2 in mammalian orthologs of TdT. Loop1 is boxed in green, and regions that are structurally homologous to the deleted region of Loop 2 in hPol μ are boxed in red. Phe401 is marked in cyan. (e) Structure of mPol μ (purple), displaying the location of Loop 2 (red), distal from the active site (marked by the incoming nucleotide, cyan), and the DNA binding cleft (khaki ribbon and surface rendering). (f) Structure of engineered Loop 2 from hPol μ Δ2 binary complex (blue), compared to mPol μ (purple). Residues Pro398-Pro410 were deleted, and β-strands 4 and 5 fused by addition of a glycine residue (labeled Gly410 and marked by an asterisk). All structural figures were generated using PyMOL[4].
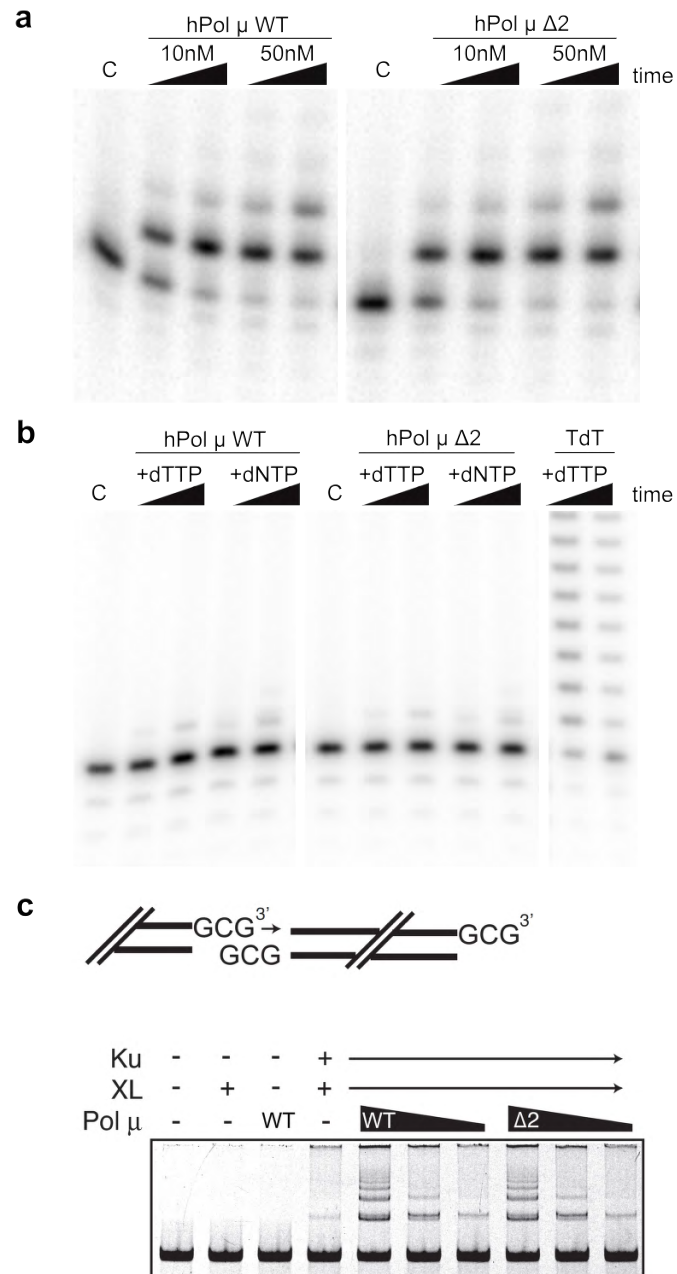
Supplementary Figure 3: Activity Assays with wildtype hPol μ and the hPol μ Δ2 variant. (a) Comparison of wildtype and Δ2 variant hPol μ template-dependent synthesis activity on a single nucleotide gapped DNA substrate. (b) Comparison of wildtype and Δ2 hPol μ template-independent synthesis activity on a single-stranded oligo-dT DNA substrate. (c) Comparison of wildtype and hPol μ Δ2 during in vitro NHEJ assays. Top, schematic diagram illustrating the structure of the DNA substrates used in the NHEJ assay. Bottom, ligation products of NHEJ synthesized by either wildtype or hPol μ Δ2.
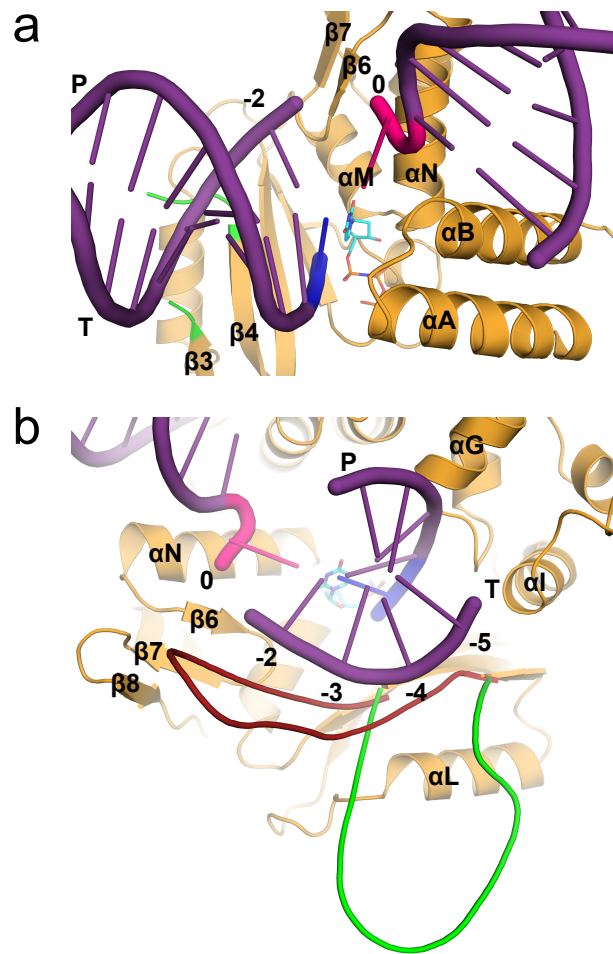
Supplementary Figure 4: Implications of Loop1 flexibility for substrate stabilization during DSB repair by NHEJ. (a) Model of a noncomplementary DSB substrate (purple) bound to the protein component of the hPol μ Δ2 pre-catalytic ternary complex (orange). The primer (strand P) terminus (blue) is unpaired, opposite the discontinuity in the template strand (strand T). The incoming nucleotide (cyan) is correctly paired opposite the templating base (magenta). Location of the disordered Loop1 is marked in green. (b) Hypothetical Loop1 conformations were manually generated for the single-nucleotide gapped ternary complex (green) and for a noncomplementary DSB substrate (red).
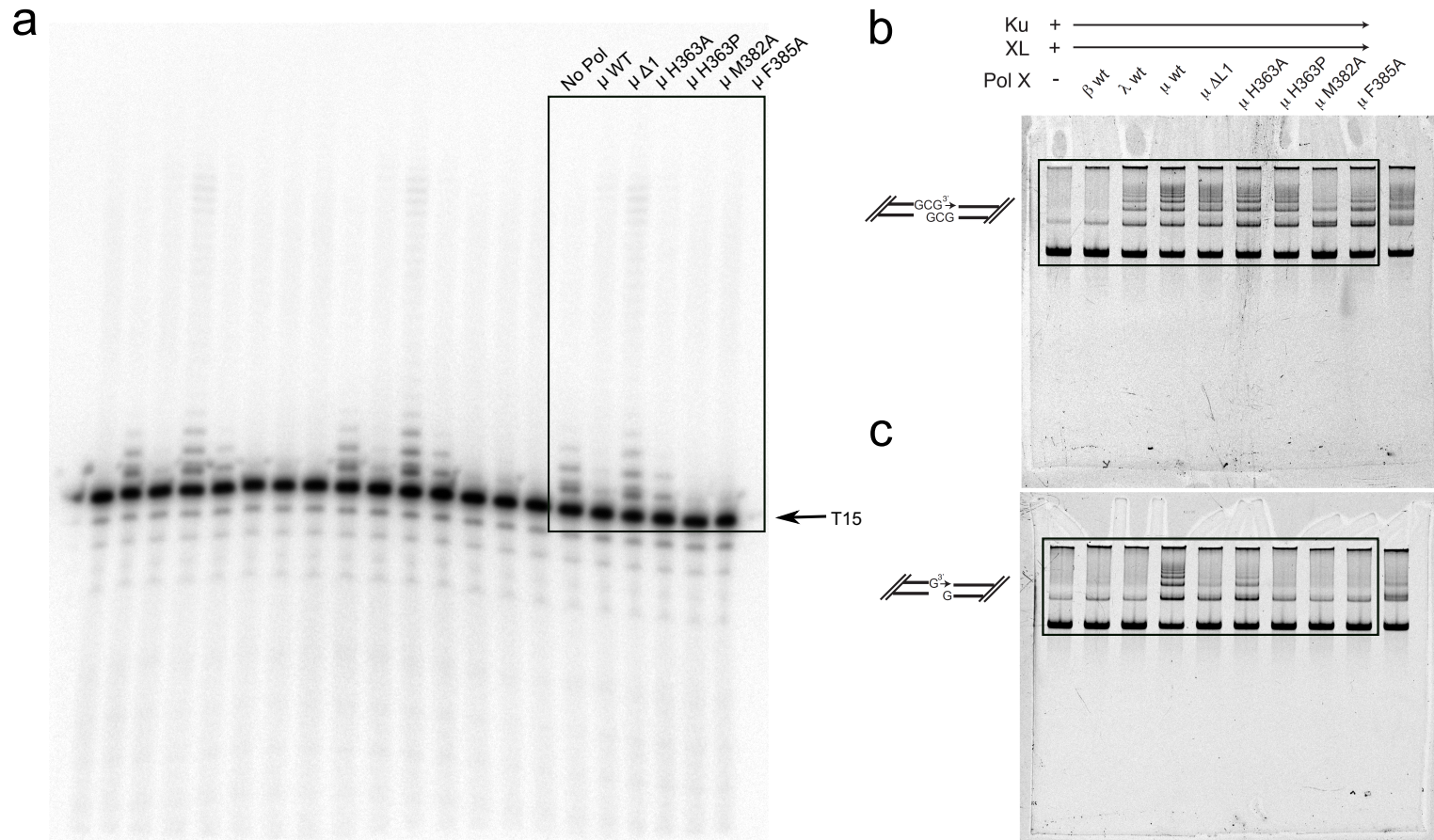
Supplementary Figure 5: Biochemical characterization of wildtype and Loop1 mutants of hPol μ. (a) Uncropped gel from Figure 5b. (b) Uncropped gel from Figure 5c (top panel). (c) Uncropped gel from Figure 5c (bottom panel). All wells in (b) and (c) also contain human Ku (10 nM) and XRCC4/Ligase IV complex (20 nM) in addition to the indicated Family X polymerase (0.5 nM).

```
TTTGCAGGGCTGGCAAGCCACGTTTGGTGGTGGCGACCATCCTCCAAAATCGGAT
        ──────────────────────────────────────────▶
                      5'pGEX


            Thrombin                        TEV
Leu Val Pro Arg Gly Ser Glu Asn Leu Tyr Phe Gln│Gly Ser
CTG GTT CCG CGT GGA TCG GAA AAC CTG TAT TTT CAG│GGC TCG
                    *                          cleavage


Ala Ala Ala Ala Ser Glu Phe Gly Ser Ser Lys Val Asp Leu
GCG GCC GCA GCT AGC GAA TTC GGA TCC TCT AGA GTC GAC CTG
   NotI         NheI    EcoRI    BamHI    XbaI     SalI   PstI


Gln Ala Ser Leu
CAG GCA AGC TTG ATCGTGACTGACTGACGATCTGCCTCGCGCGTTTCGGTG
        HindIII



ATGACGGTGAAAACCTCTGACACATGCAGCTCCCGGAGACGGTCACAGCTTGTCT
           ◀──────────────────────────
                      3'pGEX
```

Supplementary Figure 6: The multicloning site of the pGEXM expression vector. The vestigial thrombin protease cleavage site from the parental pGEX-4T3 expression vector is highlighted in blue. The Tobacco Etch Virus (TEV) protease cleavage site is shown in green, with a green line clearly delineating the actual site of cleavage. The BamHI restriction site between the two regions encoding the thrombin and TEV protease sites has been deactivated by site directed mutagenesis (blue asterisk), a silent mutation which does not affect the protein sequence. The multicloning site from the pMALX expression vector is shown[5], with the pertinent restriction sites marked. Three TGA stop codons (red) are situated immediately downstream of the multicloning site, one in each reading frame. Locations of sequencing primers 5'pGEX and 3'pGEX are shown with directional arrows.

SUPPLEMENTARY TABLES:

Supplementary Table 1: Single nucleotide incorporation kinetics for hPol μ WT and Δ2 variant

Enzyme	Km (μM)	kcat (1/s)	kcat/Km $s^{-1} \mu M^{-1}$
Pol μ WT	**1.5** $\pm$ 0.5	**0.026** $\pm$ 0.008	**0.019** $\pm$ 0.009
Pol μ Δ2	**1.6** $\pm$ 0.3	**0.024** $\pm$ 0.004	**0.016** $\pm$ 0.005

 The values for each enzyme are an average ($\pm$ standard deviation) of four independent determinations.

Supplementary Table 2. Root Mean Square Deviation (Å) of hPol μ Structural Superpositions

	hPol μ Apo	hPol μ Binary	hPol μ Ternary	hPol μ Nicked
hPol μ Apo	----	0.814/276[*]	0.802/274	0.791/271
hPol μ Binary	0.814/276	----	0.164/294	0.177/296
hPol μ Ternary (pre-catalytic)	0.802/274	0.164/294	----	0.118/287
hPol μ Nicked (post-catalytic)	0.791/271	0.177/296	0.118/287	----
hPol β Binary	4.407/271	3.302/247	3.246/249	3.290/251
hPol β Ternary (pre-catalytic)	1.936/226	1.620/218	1.678/218	1.650/220
hPol λ Binary	1.994/256	1.587/238	1.655/241	1.632/245
hPol λ Ternary (pre-catalytic)	1.779/257	1.365/238	1.422/242	1.379/243

[*]Number of Cα atoms used for R.M.S.D calculation.

Supplementary Table 3: Single nucleotide incorporation kinetics for hPol μ Loop1 substitution mutants

Enzyme	Km (μM)	kcat (1/s)	kcat/Km $s^{-1} \mu M^{-1}$	Catalytic Efficiency Relative to WT
Pol μ WT	**1.5** $\pm$ 0.5	**0.026** $\pm$ 0.008	**0.019** $\pm$ 0.009	1
H363A	**2.9** $\pm$ 0.7	**0.0085** $\pm$ 0.002	**0.0029** $\pm$ 0.0001	6x↓
H363P	**0.5** $\pm$ 0.1	**0.0024** $\pm$ 0.0008	**0.0047** $\pm$ 0.0004	4x↓
M382A	**5.7** $\pm$ 2.5	**0.013** $\pm$ 0.003	**0.0026** $\pm$ 0.001	7x↓
F385A	**2.1** $\pm$ 1.2	**0.008** $\pm$ 0.002	**0.0046** $\pm$ 0.002	3.5x↓

The values for each enzyme are an average ($\pm$ standard deviation) of 3-4 independent determinations.

Supplementary Table 4: Comparison of NHEJ activity using complementary or noncomplementary DSB substrates

Protein	Average Activity* (3′GCG Substrate)	3′GCG Activity Std Error	Average Activity* (3′G Substrate)	3′G Activity Std Error
Ku-XL (No Polymerase)	5.18	0.59	12.79	0.84
Pol β WT	9.26	0.57	26.49	3.96
Pol λ WT	83.40	13.86	21.18	3.99
Pol μ WT	100.00		100.00	
Pol μ ΔLoop	89.98	7.79	37.32	3.24
Pol μ H363A	92.74	3.68	56.87	3.90
Pol μ H363P	83.71	4.63	25.46	2.63
Pol μ M382A	40.58	2.82	16.14	0.27
Pol μ F385A	76.18	5.90	19.98	1.86

*Measurements portrayed as percentage of Pol Mu WT activity, and are the result of three independent measurements.

SUPPLEMENTARY REFERENCES:

1.   Moon, A.F. et al. Structural insight into the substrate specificity of DNA Polymerase mu. *Nat Struct Mol Biol* **14**, 45-53 (2007).
2.   Delarue, M. et al. Crystal structures of a template-independent DNA polymerase: murine terminal deoxynucleotidyltransferase. *EMBO J* **21**, 427-39 (2002).
3.   Thompson, J.D., Gibson, T.J. & Higgins, D.G. Multiple sequence alignment using ClustalW and ClustalX. *Curr Protoc Bioinformatics* **Chapter 2**, Unit 2 3 (2002).
4.   DeLano, W.L. The PyMOL Molecular Graphics System. DeLano Scientific (San Carlos, CA. USA). (2002).
5.   Moon, A.F., Mueller, G.A., Zhong, X. & Pedersen, L.C. A synergistic approach to protein crystallization: combination of a fixed-arm carrier with surface entropy reduction. *Protein Sci* **19**, 901-13 (2010).