Supporting information

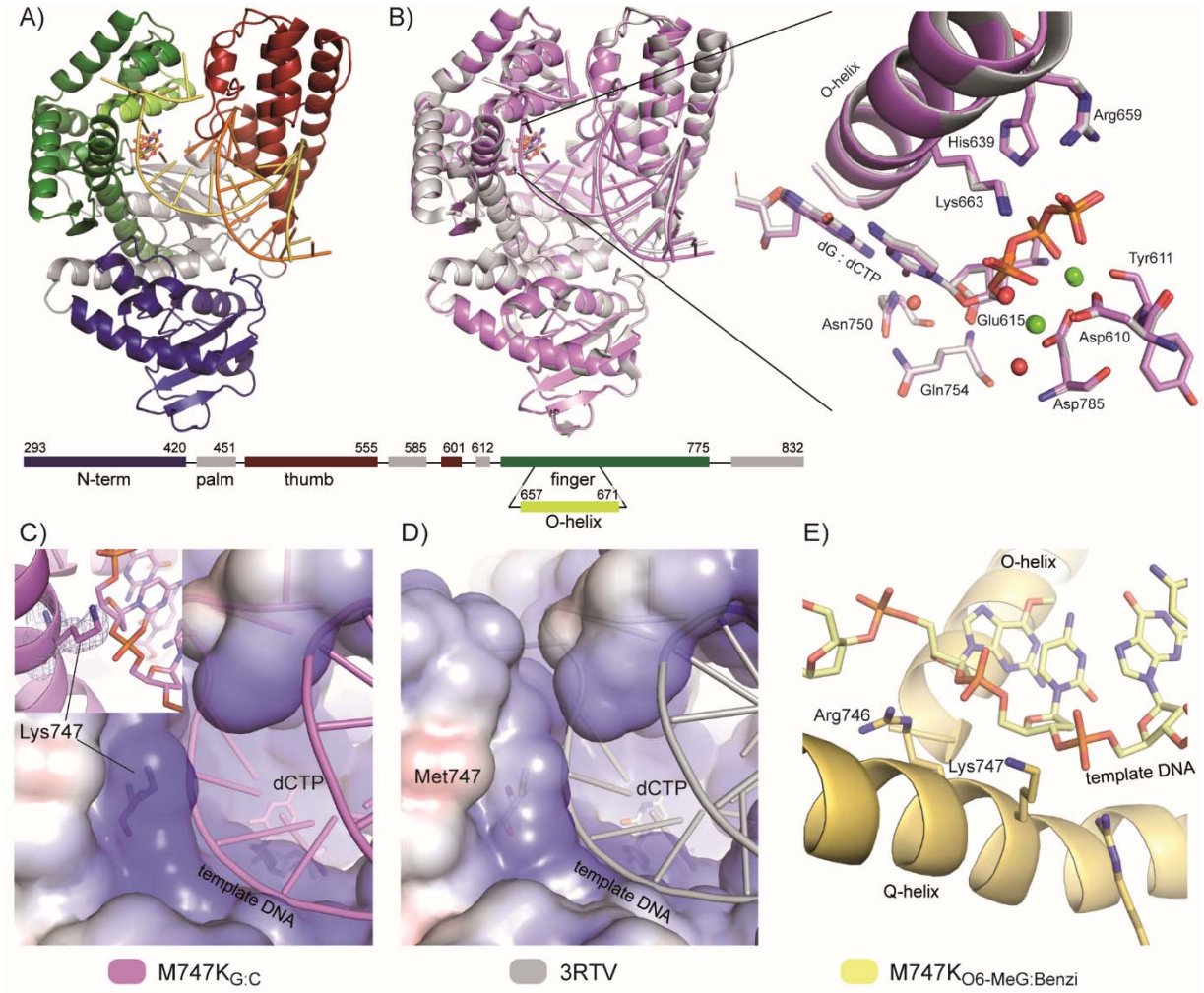

Figure S1: Structure of M747K$_{G:C}$ shown as cartoon with the different enzyme domains coloured as indicated below the structure. The DNA primer is coloured orange and the template is coloured yellow-orange. The substrate is shown as ball-and-stick model. B) Overlay of M747K$_{G:C}$ (violet) and the KlenTaq WT structure with the same p/t and substrate complexed (PDB ID: 3RTV, grey) zoomed into the active site, showing the nascent base pair and residues interacting with the triphosphate substrate and the metal ions as sticks. Magnesium ions are shown as green spheres and water molecules as red spheres. C) Electrostatic map of the protein surface of M747K$_{G:C}$ with the Lysine residue shown at the mutation site 747. Simulated annealing mFo-dFc omit map of Lys747 is shown in the inlay. D) Electrostatic map of the protein surface of 3RTV with Methionine residue shown at position 747. E) Residues responsible for the positive electrostatic potential near the template are shown as sticks (including Lys747) for the structure M747K$_{O6\text{-}MeG:Benzi}$.

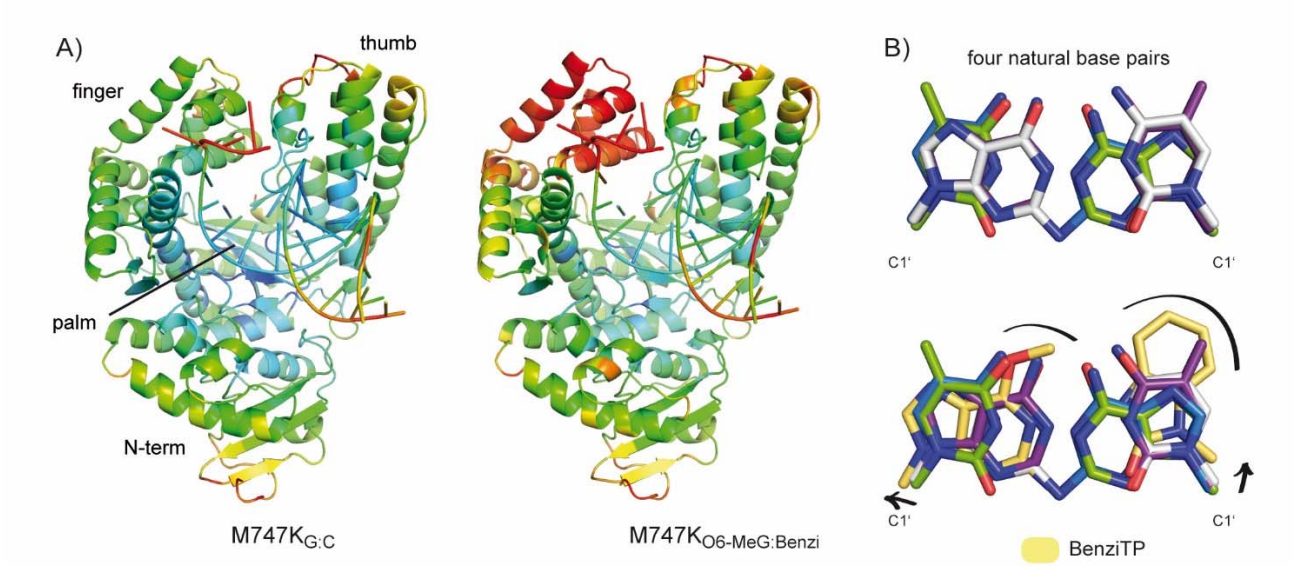

Figure S2: A) Structures of M747K$_{O6\text{-}MeG:Benzi}$ and M747K$_{G:C}$ coloured by B-factors with a colour range from dark blue (low flexibility) to red (high flexibility) B) Overlay of all four natural base pairs in the KlenTaq active site to visualize the consensus pocket as described by Kool et al.[1] Deviations of the O^6-MeG:Benzi pair from the consensus pocket are indicated with black lines and arrows.

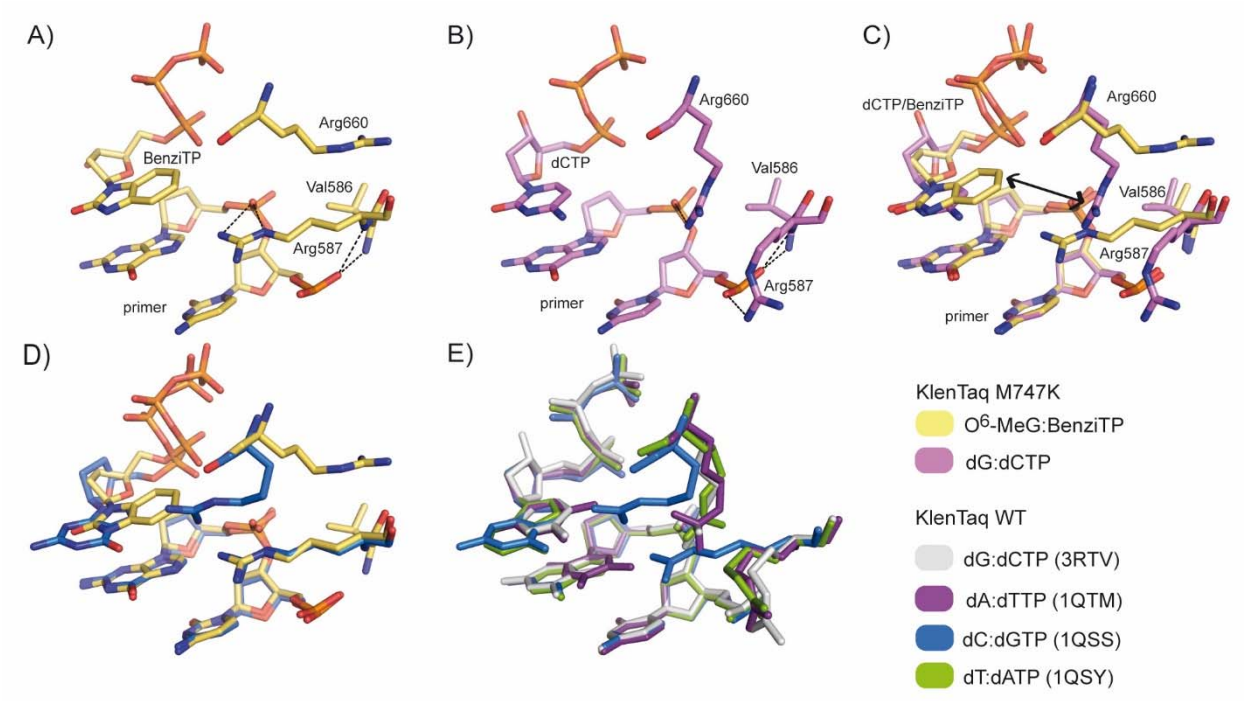

Figure S3: Interactions of the enzyme with the primer 3'-end. The last two nucleotides of the primer and the triphosphate substrate as well as relevant amino acids are shown as sticks for different structures[2, 3, 4]. Colour code is visible on the bottom right. Hydrogen bonds are shown as black dashed lines.

Table S1. Statistics for X-ray Diffraction Data Collection and Structure Refinement

	M747K$_{dG:dCTP}$	M747K$_{O6MeG:BenziTP}$
PDB ID	5O7T	5OXJ
Data collection		
Wavelength [Å]	0.999975	0.999987
Space group	$P3_121$	$P3_121$
Unit cell constants		
a, b, c [Å]	109.5, 109.5, 91.4	109.3, 109.3, 90.6
α, β, γ [°]	90, 90, 120	90, 90, 120
Resolution [Å]	47.4-1.80 (1.91-1.80)	47.3-2.00 (2.12-2.00)
Total reflections	582102 (87776)	429109 (66024)
Unique reflections	112020 (17660)	81385 (13020)
R_{meas} [%]	8.5 (287.9)	9.8 (187.8)
I/σI	11.1 (0.5)	10.7 (0.94)
Completeness [%]	98.2 (95.8)	99.3 (99.4)
Redundancy	5.2 (5.0)	5.2 (5.1)
$CC_{1/2}$ [%]	99.9 (16.9)	99.9 (26.0)
Refinement		
Resolution [Å]	47.4 – 1.80 (1.82-1.80)*	47.3-2.00 (2.02-2.00)
No. of reflections	112015	81358
R_{work} / R_{free} [%]	17.7 / 23.5 (39.0 / 40.3)	19.6 / 25.7 (40.0 / 44.7)
Coordinate error [Å]	0.33	0.35
No. of atoms		
Protein	8636	4231
Primer / Template / Substrate	377 / 507 / 40	243 / 343 / 47
Water	265	98
B-factors		
Protein	61.7	61.8
Primer / Template / Substrate	49.4 / 57.1 / 36.3	52.8 / 70.6 / 71.5
Water	46.3	47.8
R.m.s deviations		
Bond length [Å]	0.012	0.008
Bond angles [°]	1.241	0.958
Ramachandran [%]		
Favored	96.65	95.08
Allowed	3.16	4.73
Outlier	0.19	0.19

* Numbers in brackets refer to the highest resolution shell

Materials and Methods

Benzi triphosphates (BenziTP) was synthesized by previously reported procedures.[5, 6] The modified O^6-MeG oligonucleotide was purchased from Eurogentec (Seraing, Belgium).

The unmodified primer with the sequence 5'-d(GAC CAC GGC GC) as well as the unmodified template with the sequence 5'-d(AAA GCG CGC CGT GGT C) were purchased from MWG Eurofins. The O6-MeG containing template with the sequence 5'-d(AAA O^6-MeG CG CGC CGT GGT C) was purchased from Eurogentec (Seraing, Belgium). The KlenTaq mutant M747K was overexpressed and purified as described earlier for the WT enzyme. The protein was concentrated to approximately 10mg/ml and stored at 4°C in the buffer used for gel filtration (20mM Tris-HCl pH 7.5, 1mM EDTA, 0.15M NaCl, 1mM β-mercaptoethanol). The primer and the template were annealed.The protein was mixed with the primer/template duplex and ddGTP in a molar ration of 1:1.2:3 and the mixture was set to a final concentration of 20mM $MgCl_2$. Final concentration of the protein was 6.1-6.2 mg/ml. The solution was incubated at 30°C for 30min to ensure incorporation of ddCMP at the primer 3'-end. Best diffracting crystals yielding the structure $M747K_{dG:dCTP}$ grew in 20% PEG 3350, 0.1 M $Mg(OAc)_2$ and for $M747K_{O6\text{-}MeG:Benzi}$ crystals grew in 22% PEG3350 + 0.1M $Mg(OAc)_2$ + 0.1M HEPES pH 7.5. Crystals were cryoprotected in the reservoir solution containing 20% Glycerol before freezing in liquid nitrogen.

Data were collected at the beamlines PXI and PXIII at the Swiss Light Source (SLS), Paul-Scherrer Institute, Villigen, Switzerland. Data reduction was performed with the XDS package[7]. The structure was solved by difference Fourier techniques using KlenTaq wild-type (PDB 3M8S)[8] as model. Refinement was performed with PHENIX[9] and model rebuilding was done with COOT[10]. The high-resolution cutoff was established by paired refinement.[11] In the structure $M747K_{G:C}$ the complete enzyme (residues 293-832) was modelled whereas in $M747K_{O6\text{-}MeG:Benzi}$ residues 647-654 were not modelled due to weak electron density. In iterative rounds of refinement and model building geometry was validated using the MolProbity Server.[12, 13] Side chains without defined electron density were not deleted but modelled in a common rotamer conformation and high B-factors demonstrate their flexibility. The restraint file for BenziTP was generated using the grade webserver[14]. Figures of the complexes were generated with PyMol[15]. For comparison of enzymes, the complete complexes were superimposed in PyMOL if not stated otherwise. To determine rmsd values only Cα atoms were superposed in PyMOL with the default values (e.g. outlier rejection cutoff in RMS = 2, outlier rejection cycles = 5). To visualize electrostatic surfaces PQR files were generated using the PDB2PQR webserver with default settings (e.g. pH = 7).[16, 17] Before submission to the server, DNA, water molecules and ligands were deleted. Electrostatic surface representations were created in PyMOL using the APBS Tools plugin again with default settings. Omit maps for O^6-MeG and BenziTP are shown in Figure 2 and data collection and refinement statistics are shown in Table S1. The atomic coordinates and structure factors have been deposited in the Protein Data Bank with PDB codes 5O7T for $M747K_{G:C}$ and 5OXJ for $M747K_{O6\text{-}MeG:Benzi}$.

[1] E. T. Kool, *Annu. Rev. Biochem.*, 2002, **71**, 191–219.

[2] Y. Li, S. Korolev, G. Waksman, *EMBO J.*, 1998, **17**, 7514–7525.

[3] Y. Li, G. Waksman, *Protein Sci.*, 2001, **10**, 1225–1233.

[4] K. Betz, D. A. Malyshev, T. Lavergne, W. Welte, K. Diederichs, T. J. Dwyer, P. Ordoukhanian, F. E. Romesberg, A. Marx, *Nat. Chem. Biol.*, 2012, **8**, 612–614.

[5] L. A. Wyss, A. Nilforoushan, F. Eichenseher, U. Suter, N. Blatter, A. Marx, S. J. Sturla, *J. Am. Chem. Soc.*, 2015, **137**, 30–33.

[6] A. Nilforoushan, A. Furrer, L. A. Wyss, B. van Loon, S. J. Sturla, *J. Am. Chem. Soc.*, 2015, **137**, 4728–4734, PMID: 25786104.

[7] W. Kabsch, *Acta. Crystallogr. Sect. D: Biological Crystallography*, 2010, **66**, 125–132.

[8] K. Betz, F. Streckenbach, A. Schnur, T. Exner, W. Welte, K. Diederichs, A. Marx, *Angew. Chem. Int. Ed.*, 2010, **49**, 5181–5184.

[9] P. D. Adams, P. V. Afonine, G. Bunkoczi, V. B. Chen, I. W. Davis, N. Echols, J. J. Headd, L.-W. Hung, G. J. Kapral, R. W. Grosse-Kunstleve, , *Acta. Crystallogr. Sect. D: Biological Crystallography*, 2010, **66**, 213–221.

[10] P. Emsley, B. Lohkamp, W. G. Scott, K. Cowtan, *Acta. Crystallogr. Sect. D: Biological Crystallography*, 2010, **66**, 486–501.

[11] P. A. Karplus, K. Diederichs, *Science*, 2012, **336**, 1030–1033.

[12] I. W. Davis, A. Leaver-Fay, V. B. Chen, J. N. Block, G. J. Kapral, X. Wang, L. W. Murray, W. B. Arendall, J. Snoeyink, J. S. Richardson, , *Nucleic Acids Res.*, 2007, **35**, W375–W383.

[13] V. B. Chen, W. B. Arendall, J. J. Headd, D. A. Keedy, R. M. Immormino, G. J. Kapral, L. W. Murray, J. S. Richardson, D. C. Richardson, *Acta Crystallogr Sect D: Biological Crystallography*, 2010, **66**, 12–21.

[14] O. S. Smart, T. O. Womack, A. Sharff, C. Flensburg, P. Keller, W. Paciorek, C. Vonrhein, G. Bricogne, *grade, version 1.2.9*, http://www.globalphasing.com, 2011.

[15] The PyMOL Molecular Graphics System, Version 1.8.5 Schrödinger, LLC.

[16] T. J. Dolinsky, J. E. Nielsen, J. A. McCammon, N. A. Baker, *Nucleic Acids Research*, 2004, **32**, W665.

[17] T. J. Dolinsky, P. Czodrowski, H. Li, J. E. Nielsen, J. H. Jensen, G. Klebe, N. A. Baker, *Nucleic Acids Research*, 2007, **35**, W522.