Supporting Information

Structural Insights into DNA Replication Without Hydrogen-Bonds

Karin Betz,[†,§] Denis A. Malyshev,[‡,§] Thomas Lavergne,[‡] Wolfram Welte,[†] Kay Diederichs,[†] Floyd E. Romesberg,[*,‡] and Andreas Marx[*,†]

[†]*Departments of Chemistry and Biology, Konstanz Research School Chemical Biology, Universität Konstanz, Konstanz, Germany, and* [‡]*Department of Chemistry, The Scripps Research Institute, La Jolla, California, 92037 USA*

*Corresponding Author: floyd@scripps.edu, *andreas.marx@uni-konstanz.de

[§]These authors contributed equally.

Table of Contents

Sequences of the oligonucleotides:

template-I1-**5SICS**, 5'-d(AAC **5SICS** GGC GCC GTG GTC)-3'
template-E1-**NaM**, 5'-d(AG **NaM** GCG CCG TGG T)-3'
template-E2-**NaM**, 5'-d(TTC **NaM** GCG CCG TGG C)-3'
template-E2-**5SICS**, 5'-d(TTC **5SICS** GCG CCG TGG C)-3'
template-E3-**NaM**, 5'-d(TTG **NaM** GCG CCG TGG T)-3'
primer-I1: 5'-d(GAC CAC GGC GC)-3'
primer-E1: 5'-d(ACC ACG GCG C)-3'
primer-E2: 5'-d(GCC ACG GCG C)-3'
primer-E2-**5SICS**: 5'-d(GCC ACG GCG C **5SICS**)-3'

Crystallization conditions

KTQ$_{d5SICS}$: The binary complex KTQ$_{d5SICS}$ was obtained by incubating purified KlenTaq with primer-I1/template-I1-**5SICS** and ddCTP in a molar ratio of 1:1.2:5 in the presence of 20 mM $MgCl_2$. The final concentration of KlenTaq was 5 mg/ml. Setups were made using the sitting drop vapor diffusion method by mixing protein/primer-template/ddCTP and reservoir solution in a 1:1 ratio. Crystals were obtained with the following reservoir solution: 20% PEG 8000, 0.1 M Tris pH 8.0, 0.2 M magnesium formate, 20% glycerol. The N-terminal amino acid 293 and the loop between residues 647 and 659 were not modeled due to disorder.

KTQ$_{d5SICS\text{-}dNaMTP}$: The ternary complex KTQ$_{d5SICS\text{-}dNaMTP}$ was obtained by incubating KlenTaq with primer-I1/template-I1-**5SICS** and ddCTP in a molar ratio of 1:1.2:2 in the presence of 20 mM $MgCl_2$. The final concentration of KlenTaq was 6.5 mg/ml. Setups were made using the sitting drop vapor diffusion method by mixing protein/primer-template/ddCTP and reservoir solution in a 1:1 ratio. Crystals were obtained with the following reservoir solution: 15% PEG 8000, 0.1 M Tris pH 8.0, 0.2 M magnesium formate. For soaking, crystals were transferred into a drop containing the same amounts of protein/primer-template/ddCTP and reservoir solution and 2 mM d**NaM**TP. Crystals were soaked for 10 days. Before cryo cooling, crystals were consecutively soaked in reservoir solutions containing 10% and 20% glycerol.

KTQ(E1)$_{dNaM\text{-}d5SICS}$: The binary complex KTQ(E1)$_{dNaM\text{-}d5SICS}$ was obtained by incubating KlenTaq with primer-E1/template-E1-**NaM** and d**5SICS**TP in a molar ratio of 1:1.2:5 in the presence of 20 mM $MgCl_2$. Setups were made using the sitting drop vapor diffusion method by mixing KlenTaq/primer-template/d**5SICS**TP and reservoir solution in a 1:1 ratio. The final concentration of KlenTaq was 6.2 mg/ml. Crystals were obtained with the following reservoir solution: 0.2 M ammonium sulfate, 0.1 M MES pH 6.5, 28% PEG 5000MME. Crystals were cryoprotected in reservoir solution containing 20% glycerol before freezing.

KTQ(E2)$_{dNaM\text{-}d5SICS}$: The binary complex KTQ(E2)$_{dNaM\text{-}d5SICS}$ was obtained by incubating KlenTaq with primer-E2-**5SICS**/template-E2-**NaM** in a molar ratio of 1:1.2 in the presence of 20 mM $MgCl_2$. Setups were made using the sitting drop vapor diffusion method by mixing KlenTaq/primer-template and reservoir solution in a 1:1 ratio. The final concentration of KlenTaq was 6.2 mg/ml. Crystals were obtained with the following reservoir solution: 0.2 M ammonium sulfate, 0.1M MES pH 6.5, 30% PEG 5000MME.

KTQ(E2)$_{d5SICS\text{-}dNaM}$: The binary complex KTQ(E2)$_{d5SICS\text{-}dNaM}$ was obtained by incubating KlenTaq with primer-E2/template-E2-**5SICS** and d**NaM**TP in a molar ratio of 1:1.2:5 in the presence of 20 mM $MgCl_2$. Setups were made using the sitting drop vapor diffusion method mixing KlenTaq/primer-template/d**NaM**TP and reservoir in a 1:1. The final concentration of

KlenTaq was 6.2 mg/ml. Crystals were obtained with the following reservoir solution: 0.2 M ammonium sulfate, 0.1 M MES pH 6.5, 30% PEG 5000MME and 20% glycerol.

KTQ(E3)$_{\text{dNaM-d5SICS}}$: The binary complex KTQ(E3)$_{\text{dNaM-d5SICS}}$ was obtained by incubating KlenTaq with primer-E1/template-E3-**NaM** and d**5SICS**TP in a molar ratio of 1:1.5:8 in the presence of 20 mM $MgCl_2$. Setups were made using the hanging drop vapor diffusion method mixing KlenTaq/primer-template/d**5SICS**TP and reservoir in a 1:1 ratio. The final concentration of KlenTaq was 6.2 mg/ml. Crystals were obtained with the following reservoir solution: 1.2 M lithium sulfate, 2% PEG 1000, 50 mM HEPES pH 7.5.

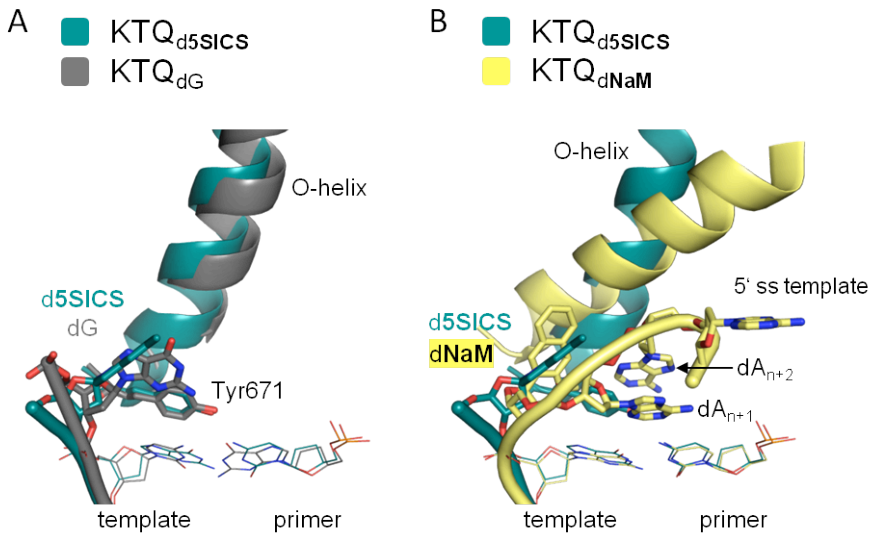

Figure S1.
Comparison of the templating nucleotide and 5' single-stranded template overhang in KTQ$_{\text{d5SICS}}$ (turquoise), KTQ$_{\text{dG}}$ (grey) and KTQ$_{\text{dNaM}}$ (yellow). (A) d**5SICS** and dG are positioned at a similar extrahelical position and the O-helix is in an open conformation. (B) The 5' single-stranded template overhang in KTQ$_{\text{dNaM}}$ is rotated towards the primer/template duplex and dA_{n+1} and dA_{n+2} stack on the last base pair. The O-helix is rotated towards the 5' single stranded template overhang.

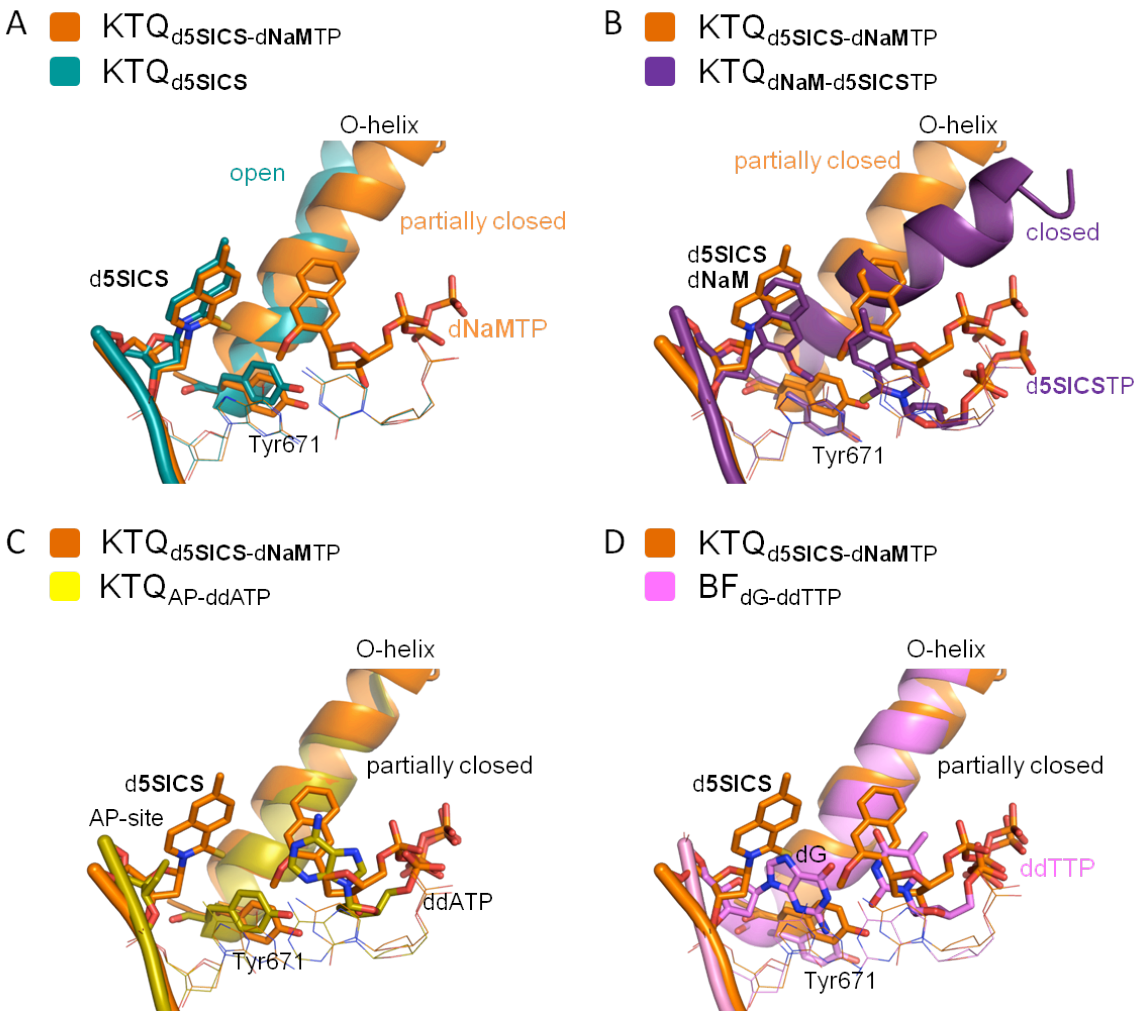

Figure S2.
Comparison of $KTQ_{\mathbf{d5SICS\text{-}dNaMTP}}$ with open, closed, and partially closed polymerase structures, respectively. The templating nucleotide, bound triphosphate, and Tyr671 are shown as sticks, the O-helix is shown as ribbon, and the base pair in the post-insertion site is shown as lines. (A) Superposition of $KTQ_{\mathbf{d5SICS\text{-}dNaMTP}}$ with $KTQ_{\mathbf{d5SICS}}$. Tyr671 stacks on top of the last base pair in both structures. (B) Superposition of $KTQ_{\mathbf{d5SICS\text{-}dNaMTP}}$ with $KTQ_{\mathbf{dNaM\text{-}d5SICSTP}}$ (PDB ID: 3SV3). In the closed structure the templating nucleotide is moved inwards and displaces Tyr671. (C) Superposition of $KTQ_{\mathbf{d5SICS\text{-}dNaMTP}}$ with KTQ in complex with an AP site in the template and ddATP bound at the O-helix (PDB ID: 3LWL).[1] Both complexes are partially closed and Y671 is in the templating position. (D) Superposition of $KTQ_{\mathbf{d5SICS\text{-}dNaMTP}}$ with the mismatch dG-ddTTP complex of *Bacillus* Fragment (BF) (PDB ID: 3HP6).[2] Both complexes are partially closed but in $BF_{dG\text{-}ddTTP}$ Y671 is displaced by the templating dG.

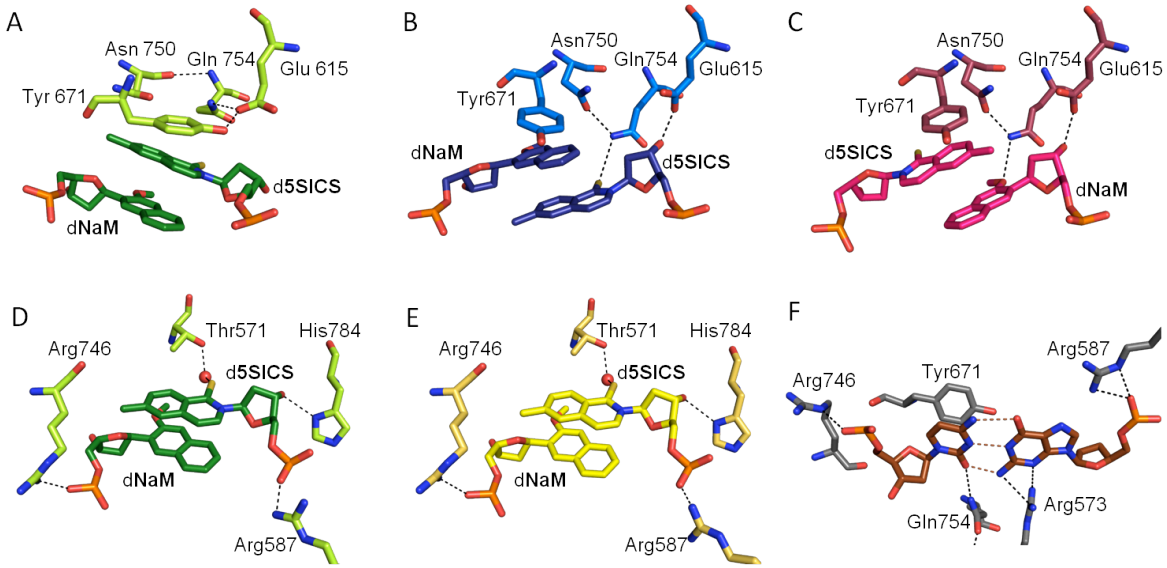

Figure S3.
Interaction networks near the intercalating base pair: (A,D) KTQ(E1)$_{\mathbf{dNaM\text{-}d5SICS}}$, (B) KTQ(E2)$_{\mathbf{dNaM\text{-}d5SICS}}$, (C) KTQ(E2)$_{\mathbf{d5SICS\text{-}dNaM}}$, (E) KTQ(E3)$_{\mathbf{dNaM\text{-}d5SICS}}$, (F) KTQ$_{dG}$ (PDB ID: 3SZ2). Protein side chains packing or hydrogen-bonding with the unnatural base pair in the post-insertion site are shown as sticks and are labeled. Hydrogen-bonds are shown as black dashes. Watson-Crick hydrogen bonds in the natural dC-ddG pair are shown in brown.

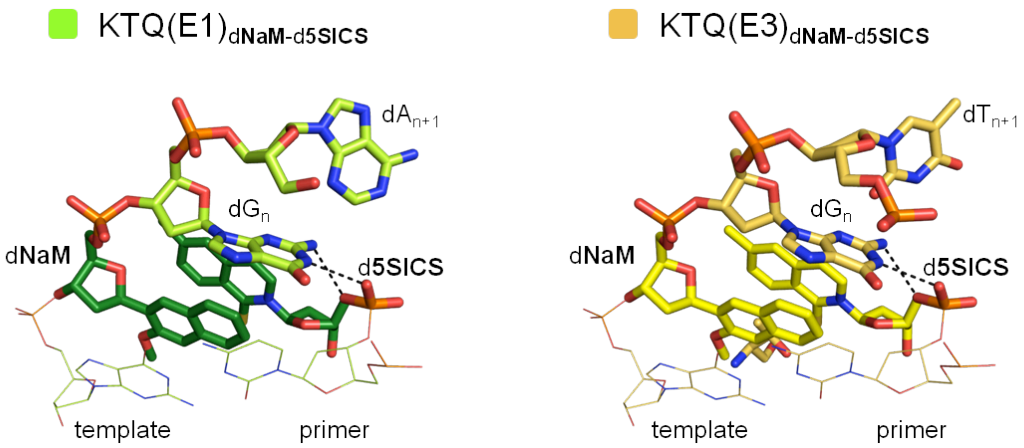

Figure S4.
Interaction of the 5' single stranded dG_n template nucleotide and the primer terminus in KTQ(E1)$_{\mathbf{dNaM\text{-}d5SICS}}$ and KTQ(E3)$_{\mathbf{dNaM\text{-}d5SICS}}$.

Table S1. Summary of data collection and refinement statistics

	KTQ_{d5SICS}	$KTQ_{d5SICS\text{-}dNaMTP}$	$KTQ(E1)_{dNaM\text{-}d5SICS}$
PDB ID	**4CCH**	**4C8K**	**4C8L**
Data collection			
Wavelength (Å)	1.0000	0.99987	1.0000
Space group	$P3_121$	$P3_121$	$C222_1$
Cell dimensions			
a, b, c (Å)	114.3, 114.3, 91.6	115.0, 115.0, 91.0	65.6, 100.5, 204.3
α, β, γ (°)	90, 90, 120	90, 90, 120	90, 90, 90
Resolution (Å)[a]	49.5-2.55 (2.7-2.55)*	48.6-2.16 (2.29-2.16)*	48.8-1.70 (1.80-1.70)*
Total reflections	205511 (25923)	227942 (14630)	492762 (76284)
Unique reflections	22906 (3650)	34375(3094)	74923 (11845)
R_{meas} (%)[b]	20.4 (144.2)	9.5 (217.4)	10.6 (275.8)
$I/\sigma I$	9.16 (0.65)	14.44(0.47)	11.57 (0.85)
Completeness (%)	99.9 (99.7)	91.1 (51.2)	99.7 (98.6)
Redundancy	9.0 (7.1)	6.6 (4.7)	6.6 (6.4)
$CC_{1/2}$ [1] (%)	99.2 (52.3)	99.8 (47.8)	99.9 (49.2)
Refinement			
Resolution (Å)	43.5-2.55	48.6-2.17	48.4-1.70
No. reflections	43559	33357	74501
R_{work}/R_{free}	20.7/26.1	18.1/22.1	18.9/21.4
Coordinate error (Å)[c]	0.42	0.25	0.37
No. atoms			
Protein	8458	8650	8698
DNA/triphosphate	795	826/49	772
Water	116	138	274
B-factors			
Protein	66.8	56.6	51.0
DNA/triphosphate	52.0	50.3/77.4	50.6
Water	38.3	45.6	43.4
R.m.s deviations			
Bond lengths (Å)	0.005	0.007	0.011
Bond angles (°)	0.770	0.870	1.239
Ramachandran[d] (%)			
Favored	95.04	97.77	97.23
Allowed	4.58	2.04	2.77
Outlier	0.38	0.19	0.00

Table S1, *continued*.

	KTQ(E2)$_{\text{dNaM-d5SICS}}$	KTQ(E2)$_{\text{d5SICS-dNaM}}$	KTQ(E3)$_{\text{dNaM-d5SICS}}$
PDB ID	**4C8O**	**4C8M**	**4C8N**
Data collection			
Wavelength (Å)	1.00001	1.0000	0.97793
Space group	$C222_1$	$C222_1$	$C222_1$
Cell dimensions			
a, b, c (Å)	64.8, 99.2, 203.6	65.4, 101.5, 204.5	65.6, 101.1, 204.3
α, β, γ (°)	90, 90, 90	90, 90, 90	90, 90, 90
Resolution (Å)[a]	48.2-1.75 (1.86-1.75) *	49.2-1.57 (1.66-1.57) *	49.1-1.88 (1.99-1.88) *
Total reflections	439052 (70744)	577762 (53959)	361542 (48980)
Unique reflections	66214 (10539)	94841 (14820)	55373 (8568)
R_{meas} (%)[b]	6.5 (216.4)	6.2 (155.1)	13.1 (200.5)
$I/\sigma I$	17.73 (0.99)	13.48 (0.78)	9.1 (0.80)
Completeness (%)	99.6 (99.2)	99.5 (97.1)	99.4 (96.6)
Redundancy	6.6 (6.7)	6.1 (3.6)	6.5 (5.7)
$CC_{1/2}$ [3] (%)	100.0 (52.7)	99.9 (53.1)	99.8 (47.6)
Refinement			
Resolution (Å)	47.9-1.75	49.3-1.57	49.1-1.88
No. reflections	66155	94768	55305
R_{work}/R_{free}	18.5/21.2	17.9/20.7	19.6/23.2
Coordinate error (Å)[c]	0.28	0.28	0.34
No. atoms			
Protein	8680	8728	8657
DNA	739	770	774
Water	230	338	203
B-factors			
Protein	57.7	49.0	51.7
DNA	71.5	64.9	50.0
Water	49.3	46.2	43.5
R.m.s deviations			
Bond lengths (Å)	0.005	0.011	0.007
Bond angles (°)	0.964	1.283	1.062
Ramachandran[d] (%)			
Favored	97.97	98.35	98.14
Allowed	1.84	1.28	1.48
Outlier	0.18	0.37	0.37

* Numbers in brackets refer to the highest resolution shell.

[a] Resolution value of the structures 4CCH, 4C8K, 4C8L, 4C8O, 4C8M and 4C8N at which I/σ = 2 is 2.90 Å, 2.28Å, 1.85Å, 1.90Å, 1.70Å and 2.05Å, respectively.

[b] for definition of R_{meas}, see Diederichs & Karplus, 1997[4]

[c] maximum likelihood based (as determined by PHENIX[5])

[d] as determined by MolProbity[6]

References

(1) Obeid, S.; Blattner, N.; Kranaster, R.; Schnur, A.; Diederichs, K.; Welte, W.; Marx, A. *EMBO J*. **2010**, *29*, 1738-1747.
(2) Wu, E. Y.; Beese, L. S. *J. Biol. Chem*. **2011**, *286*, 19758-19767.
(3) Karplus, P. A.; Diederichs, K. *Science* **2012**, *336*, 1030-1033.
(4) Diederichs, K.; Karplus, P. A. *Nat. Struct. Biol*. **1997**, *4*, 269-275.
(5) Adams, P. D.*, et al*. *Acta Crystallogr., Sect. D Biol. Crystallogr*. **2010**, *66*, 213-221.
(6) Davis, I. W.; Leaver-Fay, A.; Chen, V. B.; Block, J. N.; Kapral, G. J.; Wang, X.; Murray, L. W.; Arendall, W. B., 3rd; Snoeyink, J.; Richardson, J. S.; Richardson, D. C. *Nucleic Acids Res*. **2007**, *35*, W375-383.