

Supplementary Information

DNA-mediated coupling of ATPase, translocase and nuclease activities of a Type ISP restriction-modification enzyme

Mahesh Kumar Chand, Vanessa Carle, K. G. Anuvind, Kayarat Saikrishnan*

Division of Biology, Indian Institute of Science Education and Research, Pune, 411008, India.

**Corresponding author*

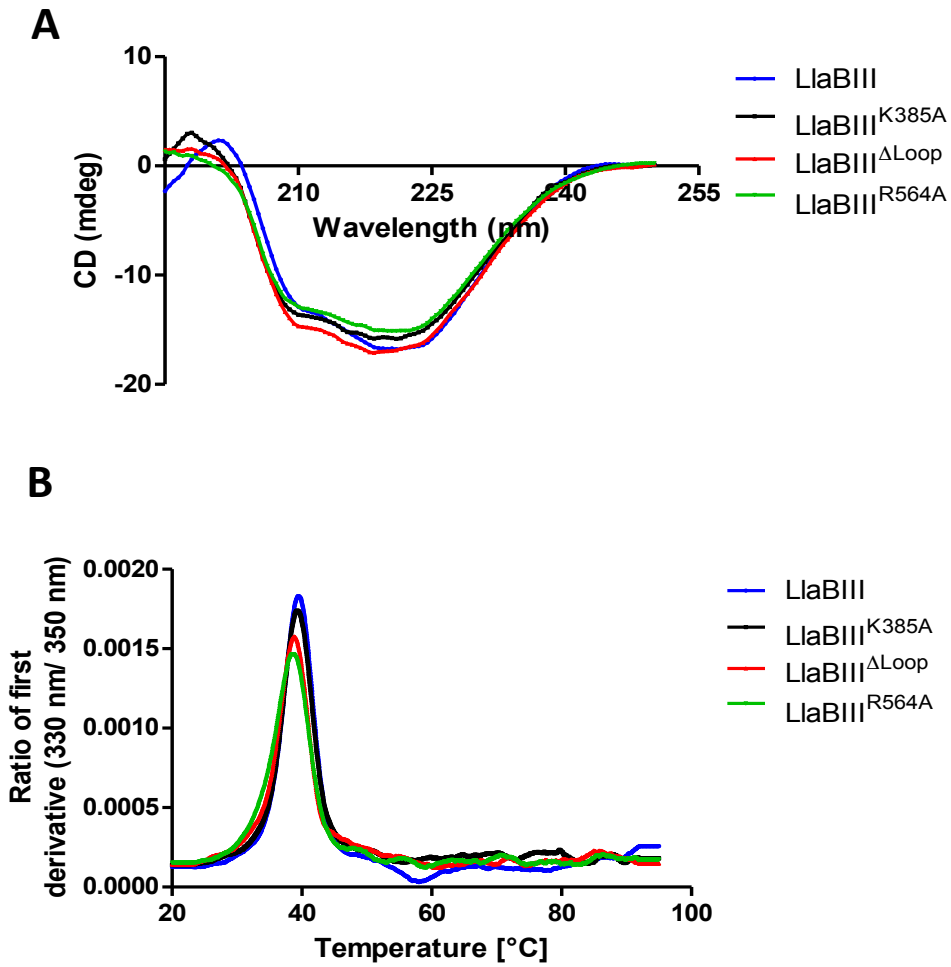
email: saikrishnan@iiserpune.ac.in

Phone: +91 2025908047

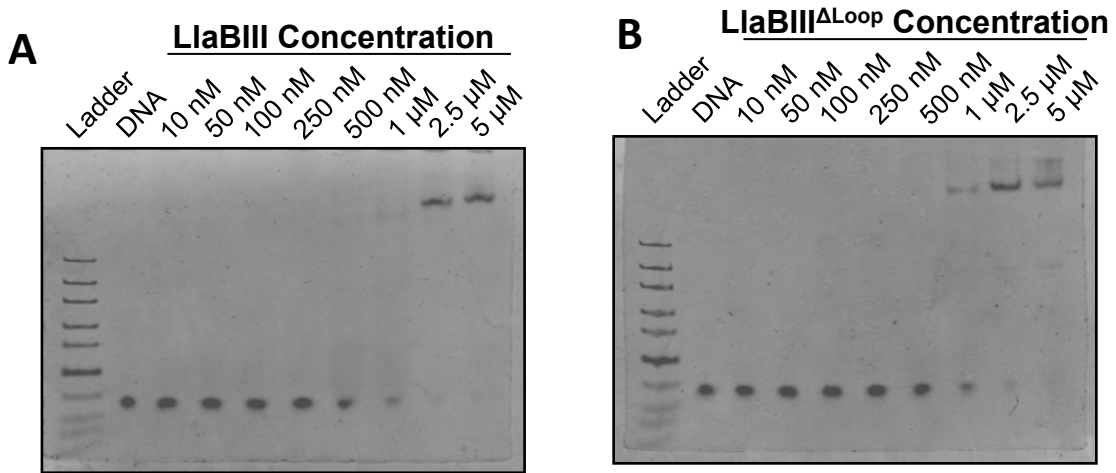
Fax: +91 2025908186

<i>LlaIII/1-29</i>	MYQTATPK	IYGESAKKNAK	DKSILLS	SMD
<i>LlaG1/1-29</i>	MYQTATPK	IYGESAKKNAK	DKSILLS	SMD
<i>AaaTCORF264P/1-29</i>	LYMTATPK	IYVQESKTKAA	ENDVKVFL	SMD
<i>AaaORF378P/1-29</i>	LYMTATPK	IFADKSKTKAN	EADAVLFL	SMD
<i>AaaK3ORFBP/1-29</i>	LYMTATPK	IYDDTKARAG	OKNAFLA	SMD
<i>Apo3283ORF18740P/1-29</i>	LYMTATPK	IYGDVAQEKAE	KEGAIIVY	GMN
<i>Acp180ORF1406P/1-29</i>	LYMTATPK	IFADKTKNTAS	EKDVLILT	SMD
<i>AspDJORF2890P/1-29</i>	LYMTATPK	IYGDNAKIMAE	SGEVTLCS	SMD
<i>Bce1271ORF56640P/1-29</i>	MYQTATPK	IYADSAKKNKAK	DKSILLS	SMD
<i>BfYORF1980P/1-29</i>	LYMTATPK	LYSDDSKSKAA	QGDAILLC	SMD
<i>BsBGRORF24100P/1-29</i>	LYMTATPK	IYAEASKTKAE	ESDIQVFL	SMD
<i>Bpr4ORF12830P/1-29</i>	LYMTATPK	IFSDNAKKQAN	EIDAVLA	SMD
<i>Bpr4ORF17060P/1-29</i>	LYMTATPK	IFSDNAKKQAN	EIDAVLA	SMD
<i>Bpr4ORF18960P/1-29</i>	LYMTATPK	IFGDNAKSRAN	EANVVLVA	SMD
<i>Bpr4ORF17840P/1-29</i>	LYMTATPK	IFTDKCLKKAD	VSDAVLA	SMD
<i>Bpr4ORF3060P/1-29</i>	LYMTATPK	IFTDDAKKNAD	EINAVLA	SMD
<i>BhoHORF15450P/1-29</i>	LYMTATPK	IFGDNAKSRAN	EANVVLVA	SMD
<i>BovEORF3337P/1-29</i>	LYMTATPK	LYGQSAKIKAS	EKDAILCS	SMD
<i>BscR1ORF40207P/1-29</i>	LYMTATPK	IFGDNVKSKQAN	AVGAVLA	SMD
<i>BtrCIPORF1006P/1-29</i>	LYMTATPK	IFSDTAKRRAD	EINAVLA	SMD
<i>BtrCIPORF1021P/1-29</i>	LYMTATPK	IFADSAKKQAH	EMNGILAS	SMD
<i>BtrCIPORF1035P/1-29</i>	LYMTATPK	IFSNHAKRRAD	EVDAVLA	SMD
<i>BtrCIPORF1064P/1-29</i>	LYMTATPK	IFADSAKKQAH	EMNGILAS	SMD
<i>BtrCIPORF1080P/1-29</i>	LYMTATPK	IFTDKCLKKAD	ASDAVLA	SMD
<i>BtrCIPORF1105P/1-29</i>	LYMTATPK	IFSDTAKRRAD	EINAVLA	SMD
<i>BtrCIPORF14P/1-29</i>	LYMTATPK	IFTDDAKKNAD	EINAVLA	SMD
<i>BtrCIPORF164P/1-29</i>	LYMTATPK	IFADSAKKQAH	EMNGILAS	SMD
<i>BtrCIPORF2491P/1-29</i>	LYMTATPK	IFGDNAKSRAN	EANAVLA	SMD
<i>BtrCIPORF455P/1-29</i>	LYMTATPK	IFSDKCLKRRAF	LSNNVLV	SMD
<i>BtrCIPORF541P/1-29</i>	LYMTATPK	IFSNHAKRRAD	EVDAVLA	SMD
<i>CehORF378EP/1-29</i>	LYMTATPK	RMYNVDARSQAA	KQAIPLWS	SMD
<i>CdpORF433P/1-29</i>	LYMTATPK	LFDEAVKGGKAA	EHS AFLA	SMD
<i>CglORF3009P/1-29</i>	LYMTATPK	LFDDSVKGGKAA	DHSAEVS	SMD
<i>Cms364ORFAP/1-29</i>	LYMTATPK	LFDDITKSKAE	EHSAEIVY	SMD
<i>Cps411ORF557P/1-29</i>	LYMTATPK	LFDDITVGGKAA	DHFAELAS	SMD
<i>Cps119ORF359P/1-29</i>	LYMTATPK	LFDDVGGKAA	DHFAELAS	SMD
<i>CspBORFGP/1-26</i>	LYMTATPK	VVYSVLIKAE	DMGKVIVY	SMD
<i>CspORF142P/1-29</i>	LYLTATPK	VFSDKTKKTKF	EKELNLF	SMD
<i>DaeORF402P/1-29</i>	IYMTATPK	IYGDTAKATAE	RDDIVLCS	SMD
<i>DvtRCH1ORF2026P/1-29</i>	LYMTATPK	LYAEOSKSKAK	ERDIAVFL	SMD
<i>EaeRM24P/1-29</i>	LYMTATPK	IYTDSAKSKAR	QSEFQLY	SMD
<i>GbrORF7642P/1-29</i>	LYMTATPK	IFADTVREKAE	EHS AELT	SMD
<i>Gen18890ORFHP/1-29</i>	LYMTATPK	IYGDVAREKAE	KEGAIIVY	GMN
<i>GspSXCC1ORF365BP/1-29</i>	LYMTATPK	IYGDVAREKAE	KEGAIIVY	GMN
<i>HaeSORF1335P/1-30</i>	LYMSATPK	IFSDNAKSKAKK	DVEVELY	SMD
<i>Hpy180ORF3395P/1-29</i>	LYMTATPK	KVYESSKAKAK	ESDNVIY	SMD
<i>Hpy2018ORF6581P/1-29</i>	LYMTATPK	KVYESSKAKAK	EKDNVIY	SMD
<i>Hpy20ORF3615P/1-29</i>	LYMTATPK	KVYESSKAKAK	ESDNAIY	SMD
<i>Hpy225ORF675P/1-29</i>	LYMTATPK	KVYESSKAKAK	ESDNAIY	SMD
<i>Hpy30ORF661P/1-29</i>	LYMTATPK	KVYESSKAKAK	ESDNAIY	SMD
<i>Hpy464ORF3445P/1-29</i>	LYMTATPK	KVYESSKAKAK	ESDNAIY	SMD
<i>Hpy4ORF3450P/1-29</i>	LYMTATPK	KVYESSKAKAK	ESDNVIY	SMD
<i>Hpy57ORF690P/1-29</i>	LYMTATPK	KVYESSKAKAK	ESDNAIY	SMD
<i>Hpy75ORFBP/1-29</i>	LYMTATPK	KVYESSKAKAK	EKDNVIY	SMD
<i>Hpy83ORF726P/1-29</i>	LYMTATPK	KVYESSKAKAK	ESDNVIY	SMD
<i>Hpy908ORF680P/1-29</i>	LYMTATPK	KVYESSKAKAK	EKDNVIY	SMD
<i>Hpy94ORF3450P/1-29</i>	LYMTATPK	KVYESSKAKAK	ESDNVIY	SMD
<i>Hpy99ORF612P/1-29</i>	LYMTATPK	KVYESSKAKAK	EKDNVIY	SMD
<i>HpyAORF663P/1-29</i>	LYMTATPK	KVYESSKAKAK	ESDNVIY	SMD
<i>HpyB8ORF370P/1-29</i>	LYMTATPK	KVYESSKAKAK	ESDNVIY	SMD
<i>HpyGORF629P/1-29</i>	LYMTATPK	KVYESSKAKAK	EKDNVIY	SMD
<i>HpyHORF652P/1-29</i>	LYMTATPK	KVYESSKAKAK	ESDNVIY	SMD
<i>HpyPORF680P/1-29</i>	LYMTATPK	KVYESSKAKAK	ESDNVIY	SMD
<i>HpySORF3515P/1-29</i>	LYMTATPK	KVYESSKAKAK	ESDNAIY	SMD
<i>Hso3336ORF1665P/1-26</i>	LYMTATPK	IYTEDAKK - -	SEGEVY	SMD
<i>KhpF74ORFGP/1-29</i>	LYMTATPK	RVEDGTRKKAPE	ENSVVVA	SMD
<i>LmPORF33P/1-29</i>	LYMTATPK	IFGDIAKASAE	KDNVTLCS	SMD
<i>Lga21881ORFAP/1-29</i>	LYQTATPK	IYVGOEAKNTAN	EOSVVA	SMD
<i>LncCORF21P/1-29</i>	MYQTATPK	IYGESAKKNAK	DKSILLS	SMD
<i>Lpo6026ORF338P/1-29</i>	LYQTATPK	IYDONAKKKAKE	ENSIVVS	SMD
<i>LsaSKORF6P/1-29</i>	MYQTATPK	IYGESAKKNAK	DKSILLS	SMD
<i>LmoDORFAP/1-29</i>	MYQTATPK	IYGESAKKNAK	DKSILLS	SMD
<i>LmtfGMORF20390P/1-29</i>	LYMTATPK	IFFTSIKDRAD	QHS AELV	SMD
<i>MboBCTORF2038P/1-29</i>	LYMTATPK	IFFTSIKDRAD	QHS AELV	SMD
<i>MpePMORF102P/1-29</i>	LYMTATPK	IYGEASKTKAG	EREAVLFL	SMD
<i>MspELB17ORFDP/1-26</i>	IYMTATPK	IYGEDAKQ - -	TENVTLCS	SMD
<i>MtuFORF12055P/1-29</i>	LYMTATPK	IFFTSIKDRAD	QHS AELV	SMD
<i>Mva6833447ORFDP/1-29</i>	LYMTATPK	IFGDAAKATAE	RDNVALCS	SMD
<i>MvaDKORF1808P/1-29</i>	LYMTATPK	IFGDAAKATAE	RDNVALCS	SMD
<i>NeuCORF2561P/1-26</i>	LYMTATPK	IYGDAAKQ - -	TENVTLCS	SMD
<i>NfrORF2220P/1-29</i>	LYMTATPK	IFDITVKKKAA	EHS AELS	SMD
<i>NhrORF2986P/1-29</i>	LYMTATPK	VFVPRLREKVE	EQGVEYFL	SMD
<i>P88505ORF1520P/1-29</i>	LYMTATPK	VYVNDNAKATAK	DKDLVLWS	MN
<i>PfCORF2950P/1-29</i>	LYMTATPK	IYGEDVKKKAD	EHS ALLTS	SMD
<i>Pma1322ORF4P/1-29</i>	LYMTATPK	IYTDHSKQKAD	DQNI GIYS	SMD
<i>PayTORF37P/1-29</i>	LYMTATPK	IYGDNAKIKAE	SGEVTLCS	SMD
<i>PayORF2066P/1-29</i>	LYMTATPK	LYTDEARKKAE	ENDAILCS	SMD
<i>Rso82ORF2304P/1-29</i>	LYMTATPK	IYGD TAKSAE	RDNVALCS	SMD
<i>RspKIDORF4062P/1-29</i>	LYMTATPK	IFGDNVRSKAD	EVGAFLAS	SMD
<i>RspSKORF2052P/1-29</i>	LYMTATPK	IFADTAKRKAD	DHDAKLA	SMD
<i>Snu2025ORF1069P/1-29</i>	LYQTATPK	IYGAEAKKKAKE	DLSVVVIS	SMD
<i>SroBORF1737P/1-29</i>	LYMTATPK	IYDEKAKGKAD	EHS AETA	SMD
<i>Ssp18395ORFBP/1-29</i>	LYMTATPK	RVFGDDARRKAD	DANAALAD	MD
<i>SspBhdORF1863P/1-29</i>	LYMTATPK	LYS EDSQKKAKE	DKEA TLCS	SMD
<i>TtnORF2915P/1-29</i>	LYMTATPK	IYGEVAKSKAD	RENIVLCS	SMD
<i>TpyZAS2ORF2024P/1-29</i>	IYMTATPK	LYS ESAKKAKE	EADALLCS	SMD
<i>TpyZAS2ORFAP/1-29</i>	IYMTATPK	LYS ESAKKAKE	EADALLCS	SMD

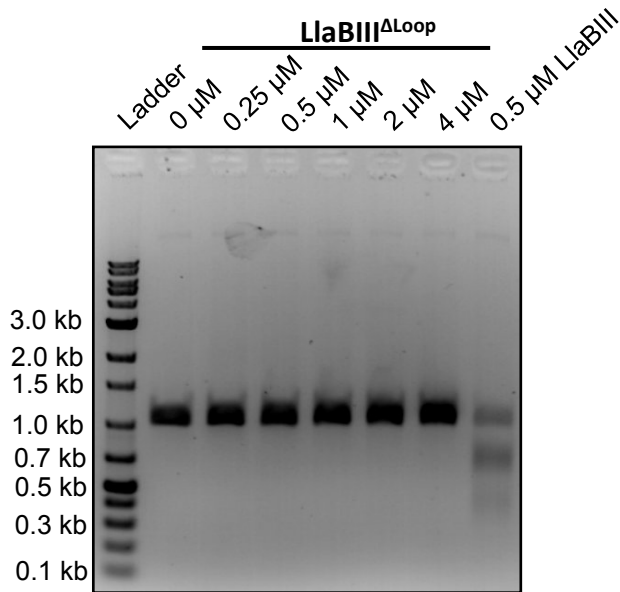
Supplementary figure 1: Amino acid sequence alignment of β -hairpin loop region in Type ISP R-M enzymes. The conserved lysine highlighted in red corresponds to LlaIII-Lys385.



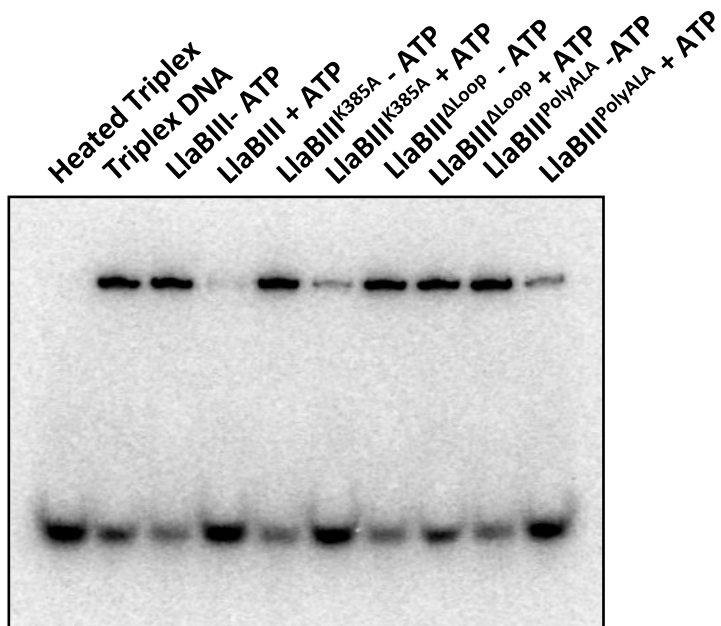
Supplementary figure 2: (A) Circular dichorism and **(B)** nanoDSF showing the conservation of secondary and tertiary structures of LlaBIII and its mutants.



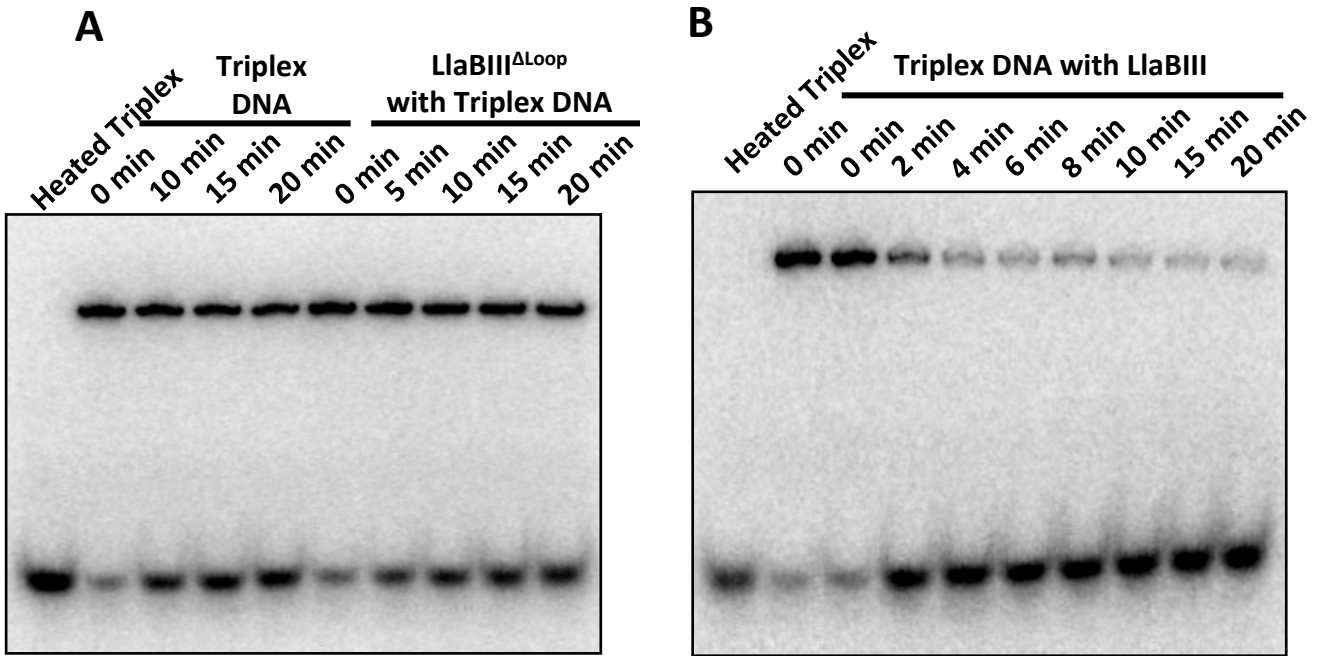
Supplementary figure 3: Electrophoretic mobility shift assay comparing the DNA binding affinities of **(A)** LlaBIII and **(B)** LlaBIII Δ Loop for a 28 bp non-specific DNA.



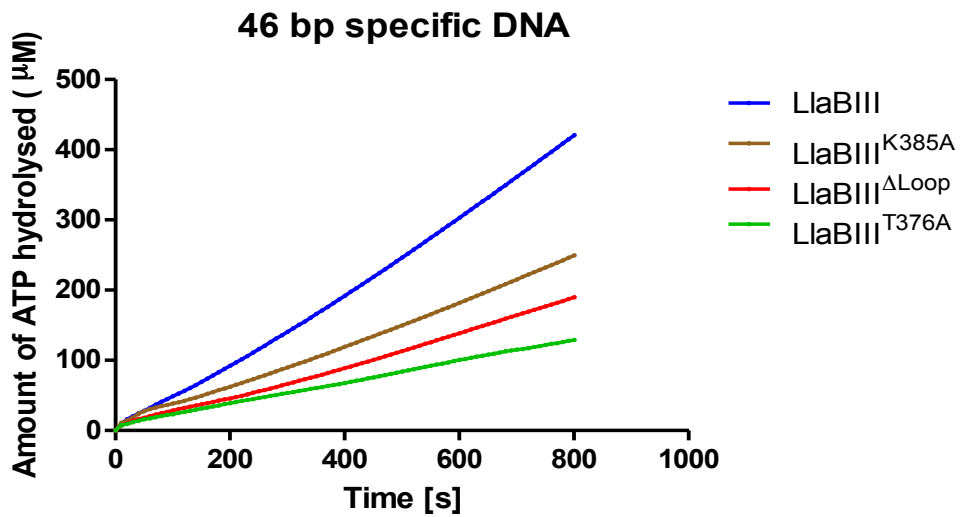
Supplementary figure 4: Concentration dependent DNA cleavage assay for LlaBIII^{ΔLoop} shows that this mutant is unable to cleave substrate DNA even at concentration of 4 μM. Last lane shows DNA cleavage by LlaBIII wild type at 500 nM concentration.



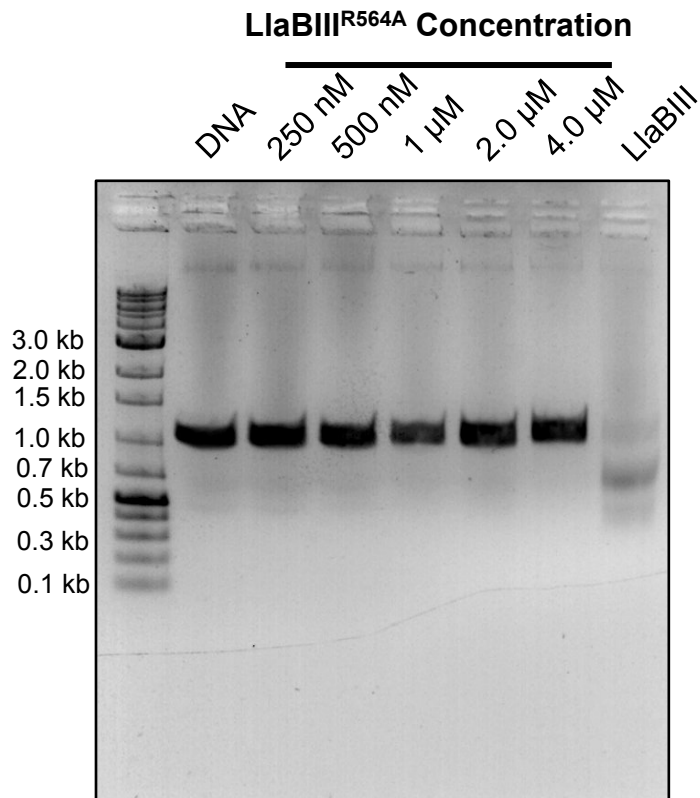
Supplementary figure 5: A representative gel of the triplex displacement assay showing the displacement of TFO by β -hairpin loop mutants in 10 minutes (see Figure 3B).



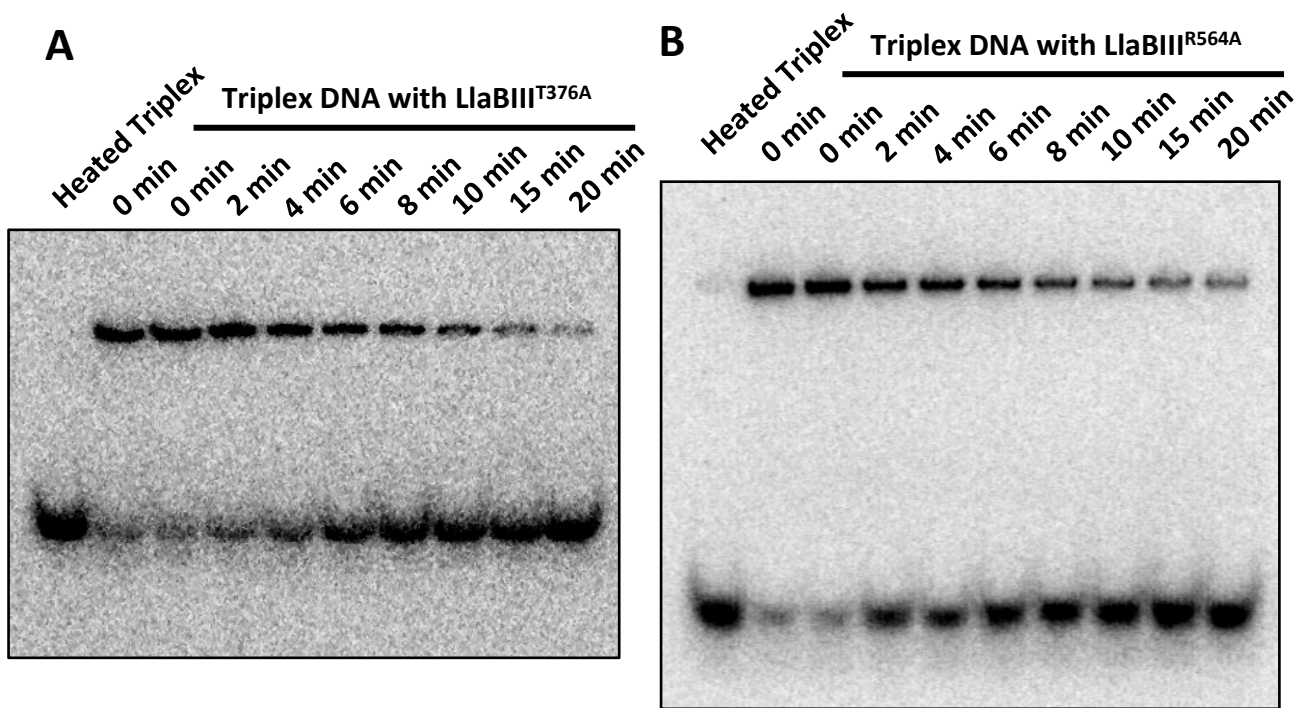
Supplementary figure 6: A representative gel of the triplex displacement assay showing time dependent displacement of TFO by LlaBIII and LlaBIII^{ΔLoop}. **(A)** Triplex DNA control and LlaBIII^{ΔLoop} **(B)** LlaBIII. (See Figure 3C)



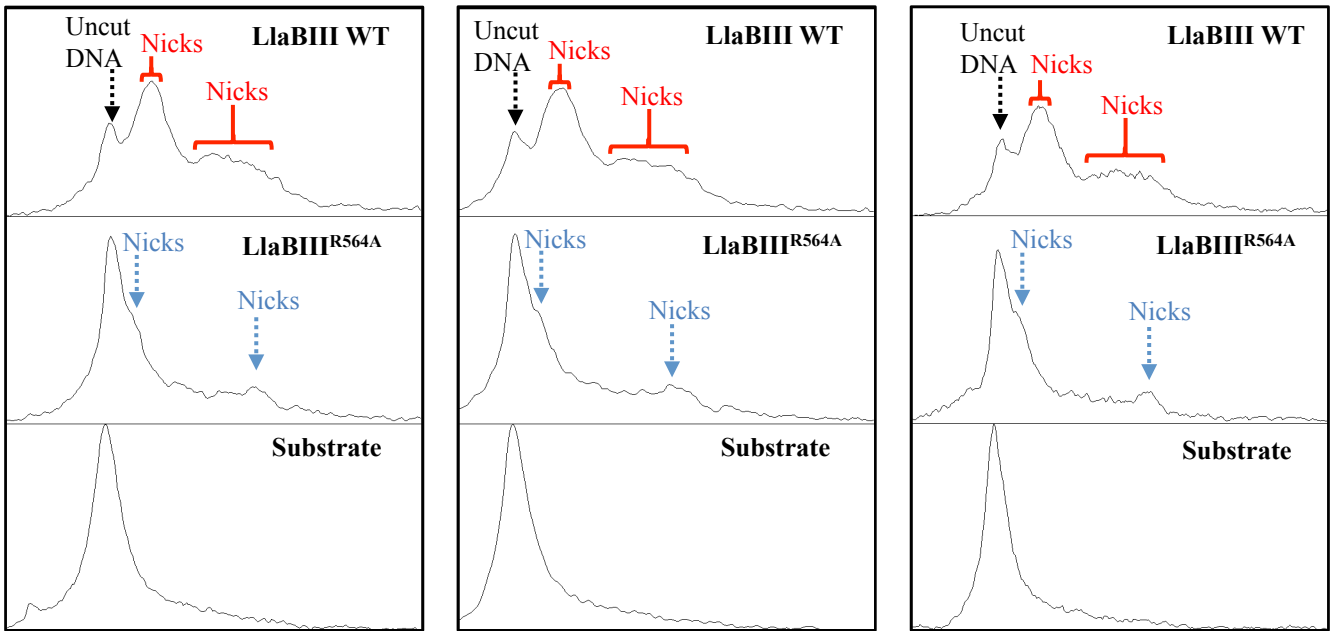
Supplementary figure 7: Comparison of the ATPase activity of LlaBIII^{T376A} with LlaBIII^{K385A} and LlaBIII^{ΔLoop} in presence of 46 bp specific DNA and 1 mM ATP carried out at 25°C.



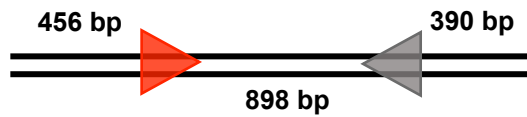
Supplementary figure 8: Concentration dependent DNA cleavage assay of LlaBIII^{R564A} shows that this mutant is unable to cleave substrate DNA even at concentration of 4 μ M. Last lane shows DNA cleavage by WT LlaBIII at 500 nM concentration.



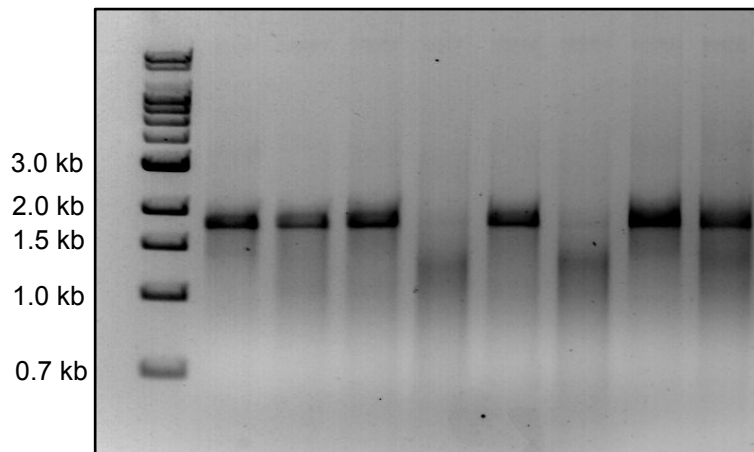
Supplementary figure 9: A representative gel of the triplex displacement assay showing time dependent displacement of TFO by (A) LlaBIII^{T376A} and (B) LlaBIII^{R564A} (see Figure 4E).



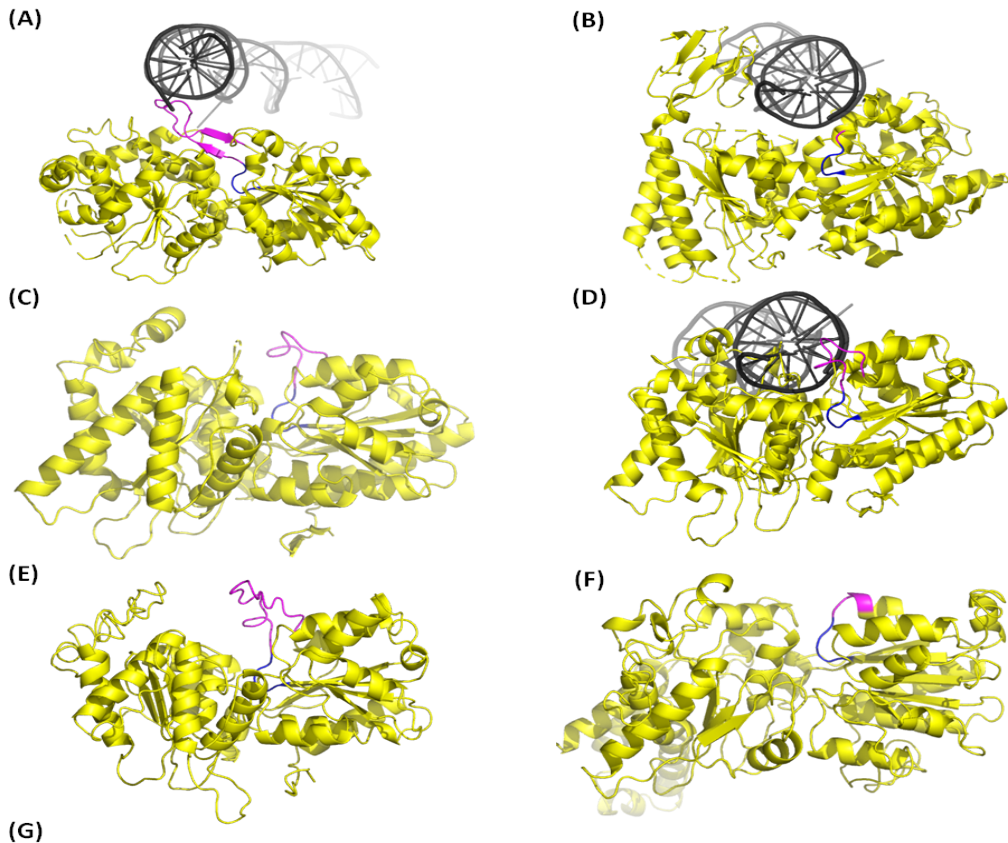
Supplementary figure 10: DNA nicking by LlaBIII and LlaBIII^{R564A}. Densitometry of the urea-formamide PAGE showing faint nicking by LlaBIII^{R564A} whereas LlaBIII nicks DNA with high efficiency. Black arrow shows the single-stranded uncut DNA. The red brackets shows the nicked product by LlaBIII whereas the blue arrows shows the nicked product by LlaBIII^{R564A}.

A**B**

LlaBIII	-	-	+	+	-	-	-	-
LlaBIII ^{T376A}	-	-	-	-	+	+	-	-
LlaBIII ^{ΔLoop}	-	-	-	-	-	-	+	+
LlaGI	-	+	-	+	-	+	-	+
DNA	+	+	+	+	+	+	+	+

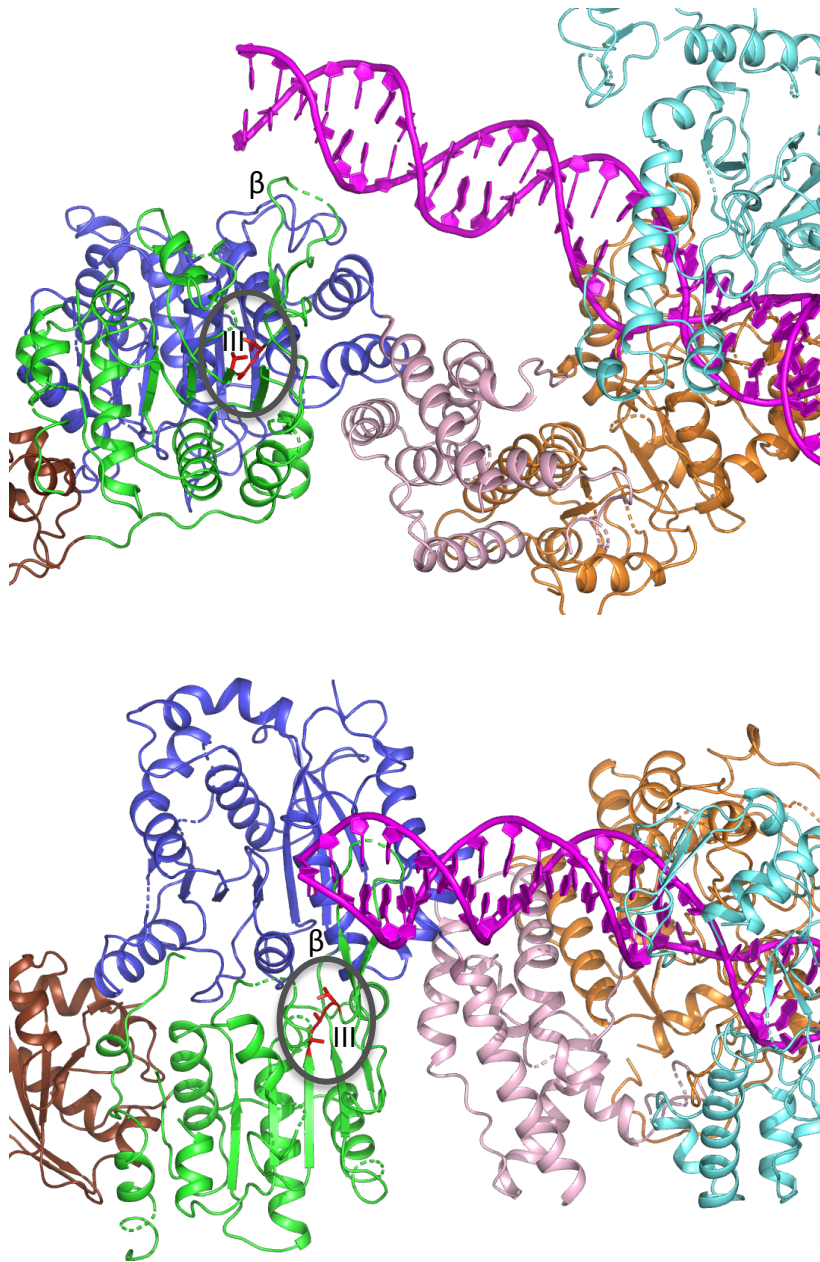


Supplementary figure 11: Heterologous cooperation assay. (A) Schematic of the DNA substrate used with LlaBIII site in red and the LlaGI site in grey. **(B)** DNA cleavage by LlaBIII or its mutants in cooperation with LlaGI. 500 nM of each enzyme was incubated with 10 nM of DNA and 4 mM of ATP for 30 minutes at 25 °C.



LlaBIII/198-239	YQIATPKIYGESAKKNAKDKSILLS	SMD	DESKYGEVFFRMG	-----
LlaGI/198-239	YQIATPKIYGESAKKNAKDKSILLS	SMD	DESKYGEVFFRMG	-----
Mva1261III/166-196	GLTATPVRYISRN	---	TFQMF	FGCESTDPTYAYG--
EcoAI/170-201	GLTATPKETHEVS	---	---	STDYFGDPVVVYSLKEGIE
EcoKI/194-221	ALTATPALH	---	---	TVQIFGEPVRYRYRTAVI
KpnAI/203-230	GFTGTPIIDA	---	---	TMEVFGNIVDSYTMTESVQ
BpuJo2/104-235	GFTGTPEVSTDRN	---	---	TPAVFGGEYIDVYDMSQAVE
PhaAI/205-236	GFTGTPIELDDKD	---	---	TQEVFGKYVSIYDFEVAVE
PacII/190-221	GLTGTPIISGIERN	---	---	TFKLFGAEDDPGRYMNRYS
DnuI/190-220	AFTGTPLLKDDK	---	---	TQNKFGPIIHAYTMQRAVE
FtnUI/177-210	GFTGTPLLKSQKSDG	---	---	SITKFNGLIHKYTI DQAIK
EcoOI/271/187-219	GFTGTPLMKKDKKK	---	---	SVEVFGPPYIHTYKFDEAVA
FclI/186-218	GFTGTPLLQKDKQT	---	---	IMDIFGRYIHTYKFNEAVT
CjeFII/209-236	AFTATPKPK	---	---	TLEMFGMPCDVNGEQRFI P
Lla82I/202-242	GLTGTPIFEANKKQENGT FART	---	---	TSQQYGPLLHSYTTKNAMD
MpuUI/197-230	GFSGTPIFEENNDRE	---	---	TQKIFGNEIDSYNMKDA I L
TelBI/171-215	GFTGTPIFEQNATYRTIEGQEARMVT	---	---	TADVFGKQLHAYTITHAID
EcoR124I/183-218	GFTGTPIFPENALGSET	---	---	TASVFGRELHSHYVITDAIR
Hov99XVI/189-227	GFTGTPIFAONCDKNNPLGT	---	---	TEOKFGRC LHOYTI I DAIR
EcoP15I/251-272	RYGATF	---	---	SEGYKNLVYRLTAVDA-
EcoPI/251-272	RYGATF	---	---	SEGYKNLVYRLTAVDA-

Supplementary figure 12: Structural comparison of the ATPase domains of (A) DNA-bound LlaBIII, (B) DNA-bound Type III RM enzyme EcoP15I (32), (C) Type I RM enzyme EcoR124I (33), (D) EcoR124I with a modeled DNA indicating the possibility of the interaction of hairpin loop (magenta) with the DNA, (E) Type I RM enzyme TelBI modeled using I-TASSER (34) and (F) EcoKI modeled using I-TASSER. The position of the β -hairpin loop is colored magenta. Blue color represents motif III. The β -hairpin loop appears to be present in EcoR124I and TelBI, but is absent in EcoKI and EcoP15I. (G) Sequence alignment of a part of the helicase domain of Type ISP (LlaBIII and LlaGI), Type I and Type III (EcoP15I and EcoPI) indicating the possible presence of an equivalent loop in Type I RM enzymes.



Supplementary figure 13: Structure of LlaBIII-DNA complex highlighting the position of motif III (red) with respect to the DNA and β -hairpin loop.

Supplementary Table 1: DNA oligomers used for the study.

Primer	Sequence
LB-pRSF-F	ACTTTAATAAGGAGATATAACCATGGTGGCATT TTTGGAAGGAATG
LB-pRSF-R	CGCAGCAGCGGTTTCTTTACCAGACTCGAGTTATAGTCCCTGTACTACTCT TG
K385A-R	GATACTCTTGTCTTTGGCATT TTTTGGCCGCACTTTCCCATATATCTTTGG
Δ Loop-2G-F	CACCAAAGATATATGGGGAAGGTGGTAAGAGTATCTTAC
Δ Loop-2G-R	GATGAAAGTAAGATACTCTTACCACCTTCCCATATATC
PolyALA-R	GATACTCTTGTCTGCGGCAGCTGCTGCCGCACTTCCCATATATC
T376A-R	CTTTCCCATATATCTTTGGTGCAGCCGTTTGGTACATTC
R564A-R	CGATTCCTTCCGTTAAGAAGGCGACATTAGAAACAATTCCG
LB-1074-F	GCAACGGACGCTCGCTGATCCAG
LB-1074-R	CCATCGCTTGGGAGACGGGGTTTTG
LJ1HISF	GAAGGAGATATACATATGGGTAAAATCGTCCTGCC
LJ1HISR1	GATGATGATGATGGGATCCTTATTCTTCCGTGGAC
200bp_FP	CTGTATGAAGCCCTGCAGAAC
1439bp_RP	TCTATTAATTGTTGCCGGGAAGC
Oligo23-F	TTAGCTAATAGACTGAGCCGAGG
Oligo23-R	TCCTCGGCTCAGTCTATTAGCTA
Oligo28-F	GCTCTAGCTAATAGACTGAGCCGAGGTG
Oligo28-R	CACCTCGGCTCAGTCTATTAGCTAGAGC
Oligo32-F	GCCTGCTCTAGCTAATAGACTGAGCCGAGGTG
Oligo32-R	CACCTCGGCTCAGTCTATTAGCTAGAGCAGGC
Oligo33-F	GACCTGCTCTAGCTAATAGACTGAGCCGAGGTG
Oligo33-R	CACCTCGGCTCAGTCTATTAGCTAGAGCAGGTC
Oligo34-F	GTACCTGCTCTAGCTAATAGACTGAGCCGAGGTG
Oligo34-R	CACCTCGGCTCAGTCTATTAGCTAGAGCAGGTAC
Oligo35-F	GTCACCTGCTCTAGCTAATAGACTGAGCCGAGGTG
Oligo35-R	CACCTCGGCTCAGTCTATTAGCTAGAGCAGGTGAC
Oligo46-F	GTCTTATGCAGGTCACCTGCTCTAGCTAATAGACTGAGCCGAGGTG

Oligo46-R	CACCTCGGCTCAGTCTATTAGCTAGAGCAGGTGACCTGCATAAGAC
TFO_1	TTCTTTTCTTTCTTCTTTCTTT
TFO_3	TTCTTTCTTTTCTTTCTTTCTT
NSP40-F	GTACTCAGCAGTATCCTGTATGCTACGTATTGCTATCGTG
NSP40-R	CACGATAGCAATACGTAGCATAACAGGATACTGCTGAGTAC

Supplementary References:

32. Gupta, Y.K., Chan, S.H., Xu, S.Y. and Aggarwal, A.K. (2015) Structural basis of asymmetric DNA methylation and ATP-triggered long-range diffusion by EcoP15I. *Nature communications*, **6**, 7363.
33. Lapkouski, M., Panjikar, S., Janscak, P., Smatanova, I.K., Carey, J., Ettrich, R. and Csefalvay, E. (2009) Structure of the motor subunit of type I restriction-modification complex EcoR124I. *Nature structural & molecular biology*, **16**, 94-95.
34. Yang, J., Yan, R., Roy, A., Xu, D., Poisson, J. and Zhang, Y. (2015) The I-TASSER Suite: protein structure and function prediction. *Nature methods*, **12**, 7-8.