Supporting Information

Bartlett et al. 10.1073/pnas.1302616110

A C DNAS

А	Mpa Ligase alignment
Mpa Lig	K M EQ P EW EE PM LAT LT - SKR FDDHG WA FGH K LDGAR C LAYR
Mtu LigD	K DQ K V FE FDN LA PM LATHGT VAG LKA SQ WA FEG KW DG YR LLVEA
<i>Msm</i> LigD	K NAR P ED FA PM LATEG SVAKYKAKQ WA FEG KW DG YR V I IDA
<i>Pae</i> LigD	K T PM P EW IAP E LA SL - V EQ PPH GQ WA YELK LDG YR LM SR I
Mpa Lig	D GH VR LM SRTRH EM NA SY PEVADA LM AQ KAND FIIDGE IVALDEN
Mtu LigD	H GAVR LR SR SG RD VTA EY PQ LRA LA ED LADHH VV LDG EAVVLD S
<i>Msm</i> LigD	H GQ LQ IR SRTGR EVTG EY PQ FKA LAAD LA EHH VV LDG EAVALDE S
<i>Pae</i> LigD	D GH VR LLTRNGH DW SER LPH LKKA LAG LG LER SW LDG ELVVLD E
Mpa Lig	GN SK FELLQ PRM GLEDPAEARRTG IPVHYYV FD ILFYDGYED TR
Mtu LigD	GVPSFSQMQNRGRDTRVEFWAFDLLYLDGRALLGT
<i>Msm</i> LigD	GVPSFGQMQNRARSTRVEFWAFDILWLDGRSLLRA
<i>Pae</i> LigD	GRPDFQALQNAFEEGHGENILYFLFDLPYHEGEDLRDV
Mpa Lig	P L L E R K H I L E K T L D Y G D T V R F L E D R L G G G L G L F D E A C K K G W E G
Mtu LigD	R Y Q D R R K L L E T L A N A T S L T V P E L L P G D G A Q A F A C S R K H G W E G
<i>Vism</i> LigD	K Y S D R R K I L E A L A D G G G L I V P D Q L P G D G P E AM E H V R K K R F E G
<i>Pae</i> LigD	A L E E R R A R L E A L L D G Q D E D P L R F S A T L A E N P R D L L A S A C R L G L E G
Mpa Lig	LIAKRAYG PYVH - GRSDEW LKFKCVQ NQEFVIGGYTEPHGKRI- E
Mtu LigD	VIAKRRDSRYQ PGRRCASWVKDKHWNTQEVVIGGWRAGEGGRSSG
Msm LigD	VVAKKWD STYQ PGRRSSSWIKDKIWNTQEVVIGGWRQ GEGGRSSG
Pae LigD	VIGKRLGSVYRS-RRSNDWIKLKCQLRQEFVIVGYTEPKGSR-H
Mpa Lig	FGA ILVSYYQDGKLMYAGKVGTGFDEKTLRSLMDKFRPLERK
Mtu LigD	VGSLLMGIPGPGGLQFAGRVGTCLSERELANLKEMLAPLHTD
<i>Msm</i> LigD	IGALVLGIPGPEGLQFVGRVGTGFTEKELSKLKDMLKPLHTD
<i>Pae</i> LigD	IGALLLGLYSPDEERRLRYAGKVGTGFTAASLKKVHERLAPLAIN
Mpa Lig	A SP FAEE VR EKEVHW LEPKLVABIG FEEW T D Y G K LRQ PRYD
Mtu LigD	E SP FD VP LPARDAKG IT Y VKPALVAEVRY SEW T P EG R LRQ S SW RC
<i>Msm</i> LigD	E SP FNAP LPKVDARGVT FVR PELVG EVRY SER T SD G R LRQ P SW R
<i>Pae</i> LigD	S SP LAKVPPARETGVVQ W VR PERLC EV SYAQM T RGG I I RQ AV FH
Mpa Lig	LR Y D K D P H S V IR E V P
Mtu LigD	LR P D K K P S E V V R
<i>Vism</i> LigD	LR P D K K P D E V V W
<i>Pae</i> LigD	LR E D K P A R E V T G E R P
В	Mpa Pol alignment
Mpa Pol	HD IKVTNPDKV LFPEDG ITKGELVDYYRR ISGVM VPLVRG RP
Mtu LigD	QRVT LTNADKV LYPATGTTK SD IFDYYAGVA EVM LGH IAGRP
Msm LigD	ERVR LTNPDKV LYPATGTTKA EVFDYYL SIAQ VM VPH IAGRP
Pae LigD	AGVR ISHPRR L IDPSIQASK LELA EFHARYAD LLLRD LR E <mark>RP</mark>
Mpa Pol	M TMQ R F P D G IGK EG F F Q K EA SDY - F P DW VHR - AT L E LGK G G I
Mtu LigD	ATRKRW PN G VDQ PA F F E K Q LA LS - A P PW LSR - AT VAHR SGTT
Msm LigD	VTRKRW PN G VA E EA F F E K Q LA SS - A P SW LER - G S ITH K SGTT
Pae LigD	VSLV B D D G IG E LF F O K HAAP LK TP G IVO LDPA LDP G H P C
Mpa Pol Mtu LigD	
Pae LigD	Q HQ VVC D D A AT LVY LA SQ AM IT PHV FLSR IDKVH TYP-IID SATG LAW IAQQ AA LEVHVPQWR FVA EPG SG ELN PG TYP-IIN TR EG LAW VAQQ AS LEVHVPQWR FEDGDQ G LLQ IR SA EA LVGAVQM GS IE FHTW NA SLANL-E
Mpa Pol Mpa Pol Mtu LigD Msm LigD Pae LigD	Q HQ VVC DDAAT LVY LA SQ AM IT PHVFLSR IDKVH TYP-I ID SATG LAW IAQQAA LEVHVPQWRFVA EPG SG ELNPG TYP-I INTREGLAW VAQQAS LEVHVPQWRFEDGDQG
Mpa Pol Mtu LigD Msm LigD Pae LigD Mpa Pol Mtu LigD Msm LigD Pae LigD	Q HQ VVCDDAAT LVYLA SQ AM ITPHVFLSR IDKVH TYP-I IDSATGLAW IAQQAA LEVHVPQWRFVAEPGSGELNPG TYP-I INTREGLAW VAQQAS LEVHVPQWRFEDGDQG LLQ IRSAEALVGAVQMGSIEFHTWNASLANGDQG PATRLVFDLDPP-DNNFET-VRSAAKTIREALDAEGYPVYLM PATRLVFDLDPG-EGVMAQLAEVARAVRDLLADIGLVTFPV PATRIVFDLDPG-EGVMAQLAEVARAVRDLLADIGLVTFV RPDRFVLDLDPDALPWKR-MLEATQLSLTLLDELGIRAFLK TGSRGLHVVVPLDRSADFDTVRAFARGFGEKLTKKYPDRFTI SGSKGLHLYTPLDEPVSSRGATVLAKRVAQRLEQAMPKLVTS SGSKGLHLYYPLAEPISSRGASVLARRVAQLEQAMPKLVTA.
Pae LigD Mpa Pol Mtu LigD Msm LigD Pae LigD Mpa Pol Mtu LigD Msm LigD Pae LigD Mpa Pol Mtu LigD Mtu L	Q HQ V V C D D A A T L V Y LA SQ AM IT PH V F L SR I D K V - H TY P - I D S A T G LAW I A Q Q A A L E V H V P Q W R F V A E PG SG E LN P G TY P - I IN T R G LAW V A Q A S L E V H V P Q W R F V A E PG SG E LN P G TY P - I IN T R G LAW V A Q A S L E V H V P Q W R F F D G D Q G - L L Q IR SA E A L V G A V Q M G S I E F H TW N A S L A N L E Y P D R L I F D L D P P - D N N F E T - V R SA A KT I R E A L D A E G Y P V Y L M P A T R L V F D L D P G - E G V M M Q L A E V A R A V R D L L A D I G L V T F P V I P A T R I V F D L D P G - E G V T M Q L A E V A R A V R D L L A D I G L V T F P V I P A T R I V F D L D P G - E G V T M Q L A E V A R A V R D L L A D I D L E T Y P L I R P D R F V L D L D P D P A L P W K R - M L E A T Q L S L T L D E L G L R A F K Y SG SK G L H L Y T P L D E P V S SR G A T V L A K R V A Q R L E Q A M P A L V T S SG G SK G L H L Y T P L D E P V S SR G A T V L A K R V A Q L E Q A M P K L V T A. SG G K M H L L V P L E R H G W D E V K A F A Q A I S Q H A R L M P E R F S A ' E L SK E K R R G R L F L D Y L R N SY G Q T G V A P Y G V R A R SG A P V A T P I T T M T K S L R A G K V F V D W S Q N SG SK T T I A P Y S L R G R D H P T V A A P R T V SG P R N R V G R I F V D Y L R N S R G A ST V A A Y S V R A R G L P Y S V V

Fig. S1. (Continued)

С	Mpa PE alignment
Mpa PE	EPLESYRSKRDFEKTPEPAGGEEKE – IKHP – IYV IQKHDASH
Mtu LigD	DRLTRYRRMRDASKTPEPIPTAKPVTGDGN – TFV IQEHHARR
<i>Msm</i> LigD	DKLTTYRSMRDASKTPEPVPKEIPKTGNND – KFV IQEHHARRI
<i>Pae</i> LigD	XPLAEYARKRDFRQ TPEP SGRKRRKDGDGLLRYC IQKHDASRI
Mpa PE	HYD LR LEM GGVLK SWAVPKGPSLDPKVKRLAM PTEDHPIGYAT
Mtu LigD	HYD FRLECDGVLVSWAVPKNLPDNTSVNHLAIHTEDHPLEYAT
<i>Msm</i> LigD	HYD LRLERDGVLVSFAVPKNLPETTAENRLAVHTEDHPIEYLA
<i>Pae</i> LigD	HYD FRLELDGTLKSWAVPKGPCLDPAVKRLAVQVEDHPLDYA
Mpa PE	FEGVIPEGQYGGGTVMVWDIGTYRNLREEKPEGSRMTIEQSYC
Mtu LigD	FEGAIPSGEYGAGKVIIWDSGTYDTEKFHDDPH
<i>Msm</i> LigD	FHGSIPKGEYGAGDMVIWDSGSYETEKFRVPEELDNPDDS
<i>Pae</i> LigD	FEGSIPKGHYGAGDVIVWDRGAWTPLDDPREGLE
Mpa PE	Q GK IEVFLEGKKLKG SYAL IRTGG I-EKRGWLFFKMKEPHEG S
Mtu LigD	TG EV IVNLHGGR ISGRYAL IRTNGDRW LAHRLKNOKDOR
<i>Msm</i> LigD	HGEIIVTLHGEKVDGRYAL IQTKGKNWLAHRMKDOKNAR
<i>Pae</i> LigD	KGHLSFLLDGEKLSGRWHLIRTNLRGKOPOWFLVKSKDDEAR S
Mpa PE Mtu LigD <i>Msm</i> LigD <i>Pae</i> LigD	Y ED IEK A A P D SV L V F E F
D	Mpa Ku alignment
Mpa Ku M	- QAIW SGNLAFGT IL IPVRMYAASELLRVP FHLVHAEDGG
Mtu Ku M	- RAIW TGSIAFGLVNVPVKVY SATADHD IR FHQVHAKDNG
<i>Msm</i> Ku M	- RSIW KGSIAFGLVNVPVKVY SATEDHD IK FHQVHAKDNG
<i>Pae</i> Ku M	ARAIW KGAISFGLVH IPV SLSAASSQGID PDW LDQR SMD
Mpa Ku F	VKYKKVC ELDGKELHM SD IHKGYE IGD-KM LQ FTDEELDE
Mtu Ku F	NRYKRVCEACGEVVDYRDLARAYESGDGQM VA ITDDD IAS
Msm Ku F	NRYKRVCEVCGEVVEYRD INKAFESDDGQM VV ITDEDIST
Pae Ku F	VGYKRVNKVTGKEIEREH IVKGVEYEKGRYVVLSEEEIRA
Mpa Ku I	R P V L SK S LK I LG F C EQ E E V P L LA L EQ P F Y LG T E S R K K G V A
Mtu Ku I	P E E R S R E I E V L E F V P A A D V D P M M F D R S Y F L E P D S K S S
<i>Msm</i> Ku I	P E E R S R E I E V V E F I P A EQ L D P L M Y D K S Y F L E P D S K S S
<i>Pae</i> Ku _A	H P K S T Q T I E I F A F V D S Q E I P L Q H F D T P Y Y L S P D R R G G
Mpa Ku Q	P FA LVKKAM EKSGKVA IV SW VTRASEQ IGMMM PYEK
Mtu Ku K	SY V LLAKT LA ET DRM A IVH FT LRNKTR LAALR VKD FGKR E
Msm Ku K	SY V LLAKT LA ET DR IA IVH FSLRNK SR LAALR VKD FSKR D
Pae Ku K	VYA LIR ET LEST GKVALANVV LHTRQHLALLR PLED
Mpa Ku G	FLIKAILYQQQIR PFDRVEVQESEVSEDLIDKGIML
Mtu Ku V	MMVHTLLWPDEIRDPDFPVLDQKVEIKPAELKMAGQV
<i>Msm</i> Ku V	MMIHTLLWPDEIRDPDFPILDKEVQIKPAELKMAGQV
<i>Pae</i> Ku A	LVLITLRWPSQVRSLDGLELDESVTAAKLDRRELEMARRL
Mpa Ku I	ERM T FK FDHAAYK E SY SQ A LR E I IEAKALGKEV - KV B PV -
Mtu Ku V	D SM ADD FN PDRYHDTY Q EQ LQ EL IDTK LEG G Q A FTA E DQ P
Msm Ku V	E SM TDD FK PD LYH DDY Q EQ LR E LVQ AK LEG G EA FSV E EQ P
Pae Ku V	EDM A SHWQ PDEYKD SFSDK IMKLV EEKAAKGQ LHAVE EEE
Mpa Ku Mtu Ku R Msm Ku A Pae Ku E	VK S - A ETR S L E A E L

Fig. S1. Protein sequence alignments of Methanocella paludicola (Mpa) ligase (Lig), polymerase (Pol), phosphoesterase (PE), and Ku with Mycobacterium tuberculosis (Mtu), Mycobacterium smegmatis (Msm), Pseudomonas aeruginosa (Pae) ligase D (LigD), and Ku. (A) Mpa Lig protein sequence alignment with LigD ligase domains (LigDom) of Mtu, Msm, and Pae. (B) Mpa Pol protein sequence alignment with LigD Pol domains (PolDom) of Mtu, Msm, and Pae. (C) Mpa PE protein sequence alignment with LigD PE domains (PEDom) of Mtu, Msm, and Pae. (D) Mpa Ku protein sequence alignment with Ku of Mtu, Msm, and Pae. All alignments were made with TCoffee and edited in JalView.

Mpa NHEJ enzymes SDS-PAGE



Fig. S2. SDS/PAGE Mpa Lig, Pol, PE, and Ku. SDS/PAGE of gel filtration chromatography fractions of respective purifications of Mpa Lig, Pol, PE and Ku. Migration sizes of the proteins were Mpa Lig ~40; Mpa Pol ~34 kD; Mpa PE ~24 kD; and Mpa Ku ~30 kD (monomer) and ~60 kD (dimer). The gels demonstrate the high purity of the proteins.

Archaeo-eukaryotic primase (AEP) superfamily motifs



Fig. S3. Archaeo-eukaryotic primase (AEP) superfamily motif sequence alignment. Three key conserved motifs (I, II, and III) confer membership to the AEP superfamily. Here Mpa Pol is aligned with Mtu and Pae nonhomologous end-joining (NHEJ) Pols, and Homo sapiens, Mus musculus, Drosophila melanogaster, Saccharomyces cerevisiae, and Schizosaccharomyces pombe primase small subunits.



Fig. S4. Mpa PE ion preference. PE reaction mixtures contained 50 mM Tris (pH 7.5), 5 mM MnCl₂, MgCl₂, CoCl₂, CuCl₂, CaCl₂, NiCl₂, or Zn(O₂CCH₃)₂ where indicated, 30 nM 5'-fluorescein–labeled 16mer DNA/RNA (D14R2), and 300 nM Mpa PE protein. Reactions were incubated at 37 °C for 1.5 h and electro-phoresed on a 15% denaturing polyacrylamide gel. n-R-R, unmodified substrate; n-R-P, removal of the terminal ribonucleotide monophosphate to yield a terminal phosphate group; n-R-OH, product after phosphate removal leaving a terminal OH group.



Polynucleotide ligase and RNA capping superfamily motifs

Fig. S5. Polynucleotide ligase and mRNA-capping superfamily motifs. Six key conserved motifs (I, Ia, III, IIIa, IV, and V) confer membership in the polynucleotide ligase and mRNA-capping superfamily. Motif VI also is included with highly conserved residues. Here Mpa Lig is aligned with Mtu and Pae LigD; Enterobacteria phage T7, H. sapiens, and M. musculus DNA ligase; and H. sapiens, M. musculus, D. melanogaster, S. cerevisiae, and S. pombe mRNA-capping enzymes.



Fig. S6. Mpa Lig preferentially ligates monoribonucleoside containing nicks in the presence of manganese, not magnesium. Lig reaction mixtures contained 50 mM Tris (pH 7.5), 5 mM MnCl₂ or Mg Cl₂, 30 nM 5'-fluorescein–labeled 16mer oligomer (D16, D15R, D14R2, D12R4, or D10R6 where indicated), 19mer 5'-phosphorylated DNA, and 300 nM Mpa Lig protein. Reactions were incubated at 37 °C for 2 h and electrophoresed on a 15% denaturing polyacrylamide gel.



Fig. 57. Mpa Pol gap filling and strand displacement at nicks is template dependent. Pol gap-filling and strand-displacement reaction mixtures contained 50 mM Tris (pH 7.5), 5 mM MnCl₂, 62.5 μM of the indicated NTP, 30 nM 5'-fluorescein–labeled 16mer DNA (*A* and *B*) or 47mer DNA (*C*), and 300 nM Mpa Pol protein. Reactions were incubated at 37 °C for 1.5 h and electrophoresed on a 15% denaturing polyacrylamide gel. A and *B* demonstrate template-independent incorporation of NTPs onto a substrate with a downstream strand. Although Mpa Pol is capable of incorporating NTPs to a DNA primer in a template-independent manner (Fig. S8), these assays indicate that base selection is performed primarily by template dependence even in the presence of an annealed sequence of downstream DNA. C demonstrates the template-dependent incorporation of NTPs onto a substrate with a self-annealing microhomology of four base pairs, creating a single-nucleotide gap on either side. Proficient incorporation of UTP required prior formation of the microhomology, because Mpa Pol can extend only from a 3'-hydroxyl, and template-independent incorporation function is limited to a single addition onto a double-stranded blunt DNA terminus. Furthermore, preferential incorporation of UTP indicates strong template dependence.



Fig. S8. Mpa PE phosphomonoesterase activity does not require 2'-OH. Phosphatase reaction mixtures contained 50 mM Tris (pH 7.5), 5 mM MnCl₂, 30 nM 5'-fluorescein–labeled 16mer DNA with 3' phosphate, and 250, 500, and 1,000 nM Mpa PE protein. Reactions were incubated at 37 °C for 1.5 h and electrophoresed on a 15% denaturing polyacrylamide gel.



Fig. S9. Mpa Pol incorporation of nontemplated NTP. Pol nontemplated extension reaction mixtures contained 50 mM Tris (pH7.5), 5 mM MnCl₂, 62.5 μM ATP, CTP, GTP, UTP, or 250 μM mixed nucleoside triphosphates (NTPs) and, where indicated, 30 nM 5'-fluorescein–labeled 36mer DNA and 300 nM Mpa Pol protein. Reactions were incubated at 37 °C for 1 h and electrophoresed on a 15% denaturing polyacrylamide gel.



Fig. S10. Mtu Pol domain (PolDom) strand displacement. Polymerase strand displacement reaction mixtures contained 50 mM Tris (pH 7.5), 5 mM MnCl₂, 250 μM NTPs, 30 nM 5'-fluorescein–labeled 16mer DNA, and 300 nM Mpa Pol or Mtu PolDom protein where indicated. Reactions were incubated at 37 °C for 1 h and electrophoresed on a 15% denaturing polyacrylamide gel.



Fig. S11. Mpa PE can rescue NHEJ intermediates formed by Mtu PolDom and Mpa Pol. Gap-filling, PE, and ligation reaction mixtures contained 50 mM Tris (pH 7.5), 5 mM MnCl₂, 250 μM NTPs, 30 nM 5'-fluorescein–labeled 16mer DNA, and 300 nM Mpa Pol or Mtu PolDom, 300 nM Mpa Lig, and PE protein where indicated. Reactions were incubated at 37 °C for 1 h and electrophoresed on a 15% denaturing polyacrylamide gel.

DNA C

S A



Fig. S12. Mpa PE resection of strand-displacement intermediates during 3' overhang NHEJ repair. Gap-filling, PE, and ligation reaction mixtures contained 50 mM Tris (pH 7.5), 5 mM MnCl₂, 250 μM NTPs, 30 nM 5'-fluorescein–labeled 42mer DNA (*A*) and 48mer DNA (*B*), and 300 nM Mpa Pol, Lig, and PE and 400 nM Ku protein where indicated above the lanes. Mpa Pol was preincubated with the substrate for 30 min where indicated above the lanes before the addition of Mpa Lig, PE and Ku. Electrophoresed on a 15% denaturing polyacrylamide gel.

SA