Figure S1. Kinetic analysis of 5´-3´ exonuclease activity. The Lineweaver-Burk plots of Spn PolI-(his) derivatives are depicted. Velocities of conversion of $v_{15,14}$ (●) and $v_{14,13}$ (■) were measured at different substrate concentrations as indicated in Experimental Procedures using 22 nM Spn PolI-(his) (WT), 9 nM Spn PolIE88K-(his), 17 nM Spn PolIE114G-(his) or 41 nM Spn PolID10G-(his). The figures are the average of at least three independent experiments in which exonucleolytic rates were tested at four time points 2.5, 5.0, 7.5 and 10 min for WT and the E88K mutant, and 10, 15, 20 and 30 min for the E114G and D10G mutants. The double reciprocal plots were fitted by a linear least-squares regression analysis.

Figure S2. Alignment of the amino acid sequences of the 5´nucleolytic domain of 38 prokaryotic and eukaryotic proteins. The alignment of 18 sequences from Xu et al. (7), with minor adjustments, was used with the program PileUp (GCG software package) to compare the 5' exonuclease domains of the DNA polymerase I analogs from: *Streptococcus pneumoniae* (Spn_Pol; EMBL accession no. P13252), *Aquifex aeolicus* (Aae_Pol; EMBL accession no. O67550), *Chlamydia trachomatis* (Ctr_Pol; EMBL accession no. Q9Z7U2), *Anaerocellum thermophilum* (Ath_Pol; EMBL accession no. Q59156), *Deinococcus radiodurans* (Dra_Pol; EMBL accession no. P52027), *Mycobacterium tuberculosis* (Mtb_Pol; EMBL accession no. Q07700), *Thermus aquaticus* (Taq_Pol; EMBL accession no. P19821), *Borrelia burgdorferi* (Bbu_Pol; EMBL accession no. O51498), *Escherichia coli* (Eco_Pol; EMBL accession no. P00582), *Thermus aquaticus* subsp. *caldophilus* (Tca_Pol; EMBL accession no. P80194), *Bacillus caldotenax* (Bca_Pol; EMBL accession no. Q04957), *Thermus filiformis* (Tfi_Pol; EMBL accession no. O52225), *Haemophilus influenzae* (Hin_Pol; EMBL accession no. P43741), *Rickettsia prowazekii* (Rpr_Pol; EMBL accession no. O05949), *Thermus aquaticus* subsp. *flavus* (Tfl_Pol; EMBL accession no. P30313), *Helicobacter pylori* (Hpy_Pol; EMBL accession no. P56105), *Treponema pallidum* (Tpa_Pol; EMBL accession no. P74933), *Bacillus subtilis* (Bsu_Pol; EMBL accession no. O34996), *Lactococcus lactis* subsp. *cremoris* (Lla_Pol; EMBL accession no. O32801), *Thermus aquaticus* subsp. *thermophilus* (Tth_Pol; EMBL accession no. P52028), *Chloroflexus aurantiacus* (Cla_Pol; EMBL accession no. O08307), *Chlamydia pneumoniae* (Cpn_Pol; EMBL accession no. Q9Z7U2), *Mycobacterium leprae* (Mle_Pol; EMBL accession no. P46835), *Rhodothermus* sp. (Rhodothermus_Pol;

EMBL accession no. Q9ZIG3), and *Synechocystis* sp. (Synechocystis_Pol; EMBL accession no. Q55971), and the DNA polymerases I from *Bacillus stearothermophilus* gene *polG*1-encoded (Bst_Pol; EMBL accession no. Q45458) and gene *polA*-encoded (Bst_PolA; EMBL accession no. P52026). Nine additional 5' nucleases were added to the alignment: the gene 6-encoded exonuclease from bacteriophage T3 (T3_exo; EMBL accession no. P20321), the RNasaH from bacteriophage T4 (T4_exo; EMBL accession no. P13319), the 5' nuclease gene *D*15-encoded from bacteriophage T5 (T5_exo; EMBL accession no. P06229), the exonuclease gene 6-encoded from bacteriophage T7 (T7_exo; EMBL accession no. P00638), the exonuclease from *Mycobacterium tuberculosis* (Mtb_exo; EMBL accession no. Q07700), the exonuclease IX from *Escherichia coli* (Eco_exo; EMBL accession no. P38506), and 3 potential 5'-3' exonucleases from *Bacillus subtilis* (Bsu_exo; EMBL accession no. P54161), *Mycoplasma genitalium* (Mge_exo; EMBL accession no. Q49406) and the gene *polA*-encoded from *Mycoplasma pneumoniae* (Mpn_exo; EMBL accession no. P75403). Additionally, the eukaryotic flap type I endonucleases from human (*h*FEN-1; EMBL accession no. P39748) and arquibacteria *Methanococcus jannaschii* (*Mj*FEN-1; EMBL accession no. Q58839) were included. Residues identical in at least 80% of the sequences are highlighted on a black ground and those similar in at least 80 % are on a gray shaded ground. The resulting consensus sequence is shown.

Figure S3. Hill plots of the data presented in Fig. 5. Plots were generated by a linear least-squares fitting of the data. The Hill coefficient *n* were obtained for each protein from the slope and, the values are shown in the figure.
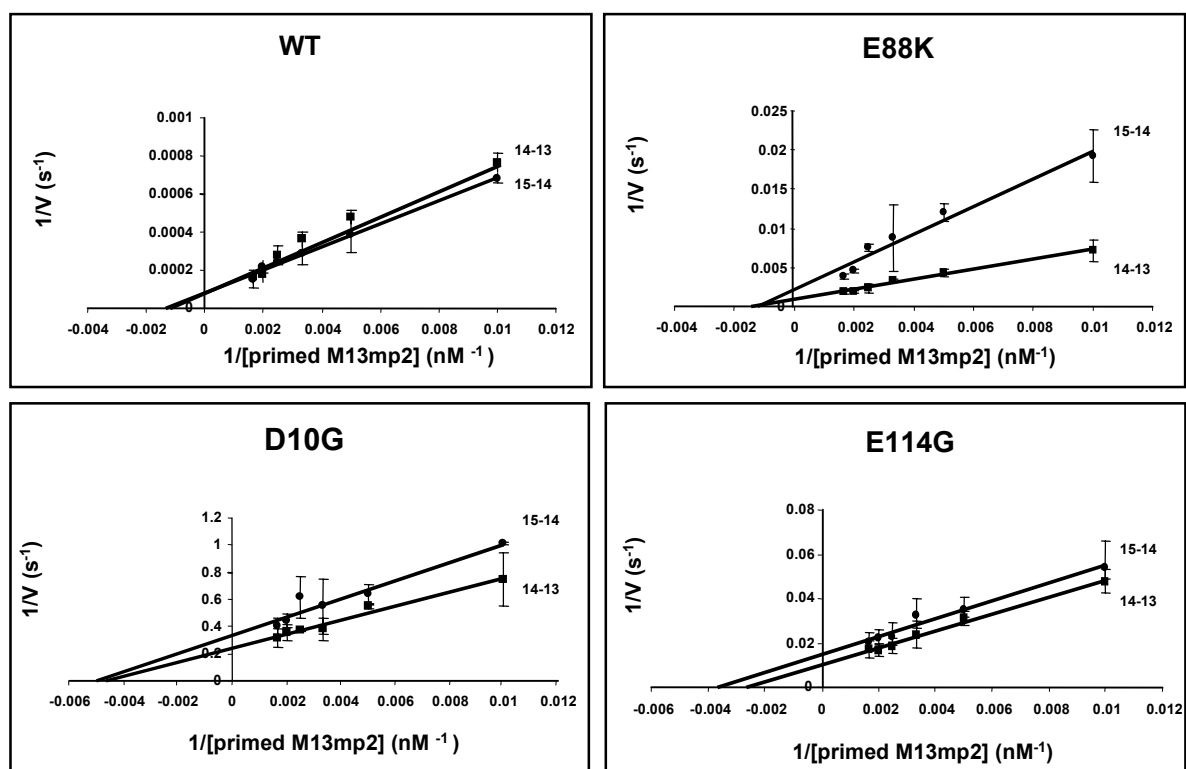
Figure S1

Figure S2

```
Spn_Pol            1 ........................MDKKK..DGSSVA.RA.F.L.YQQ
Aae_Pol            1 ........................MKT.Y..DGSSFV.RS.FA.....
Ctr_Pol            1 ........................MKK.F..DVSGFV.RA.FA.....
Ath_Pol            1 ........................MK..FDGNSIL.RA.FA.....
Dra_Pol            1 ...................MADASPDPSKPDA..DGHALA.RS.FA.....
Mtb_Pol            1 ...............MVTTASAPSEDRAKPT..DGNSLA.RA.YAL.PA.
Taq_Pol            1 ............MRGMLPLFEPKGR...DGHHLA.RT.HA.....
Bbu_Pol            1 ........................MKE.Y..DALNII.RN.HVMKNY.
Eco_Pol            1 .................MVQIPQNP..DGSSYL.RA.HAF....
Tca_Pol            1 ............MEAMLPLFEPKGR..DGHHLA.RT.FA.....
Bca_Pol            1 ....................MKKK..DGSSVA.RA.FA.....
Tfi_Pol            1 ...............MTPLFDLEEPPKR...DGHHLA.RT.YA.....
Bst_Pol            1 .....................MKNK..DGNSVA.RA.FA.....
Bst_PolA           1 .....................MKKK..DGNSVA.RA.FA.....
Hin_Pol            1 ................MAQIATNP..DGSSYL.RA.HAF....
Rpr_Pol            1 ...................MTKKNT...DGYGFV.RA.YA.....
Tfl_Pol            1 ...............MAMLPLFEPKGR...DGHHLA.RT.FA.....
Hpy_Pol            1 ..............MMEQPVIKEGT.A..DTFAYL.RS.YMSAKN.
Tpa_Pol            1 .................MQEKKT.Y.L.DAYGLI.RS.HAFARA.
Bsu_Pol            1 ...................MTERKK..DGNSLA.RA.FA.....
Lla_Pol            1 .................MEDKNR...DGSSVA.RA.F.L.YNQ
Tth_Pol            1 .............MEAMLPLFEPKGR..DGHHLA.RT.FA.....
Cla_Pol            1 ...................MAYPL.LV.DGHALA.RA.F.L.RE.
Cpn_Pol            1 ..................MKK.FV.DASGFI.RA.FA.....
Mle_Pol            1 .......MSAAMTAEVCEDHTKPML.LL.DGNSLA.RA.YAL.PT.
Rhodothermus_Pol   1 ..................MQR.Y..DAMALA.RAHYVF.IS.
Synechocystis_Pol  1 ...........MFMSAKSPLL..DVDGHSLAFRA.Y.FGLSK
Bsu_exo            1 .................MNNNK...DGMALLFRA.F.TAVHR
Mge_exo            1 ..................MKKA.L.DGNSLA.RA.FA.TWKQ
T3_exo             1 ........MALLDLKQFYELREGCDDKG.L.M.DGDWLV.QAMS.AEFDA
T4_exo             1 .............MDLEMMLDEDYKEG.CL.DFSQIALST.......A
T5_exo             1 ........SKSWGKFIEEEEAEMASRRNL.IV.DGTNLG.R........
T7_exo             1 ........MALLDLKQFYELREGCDDKG.L.M.DGDWLV.QAMS.AEFDA
Mpn_exo            1 ...................MKNA.L.DGNSLA.RA.YF..TWQQ
Mtb_exo           51 HPAPPRLRSPLDPTSGDPLHPAPPRLRSP.V.L.DCASMW.RS.FQVP...
Eco_exo            1 .......MRGLFPISHPAVACSGIECYPYRL.FK.VIVAVHLLIVDALNL
hFEN-1             1 MGIQGLAKLIADVAPSAIRENDIKSYFGRKVAI.ASMSI.QFLI.VR.QG
MjFEN-1            1 MGVQ.FGDFIPK...NIISFEDLK...GKKVAI.CMNAL.QFLTSIRLRD
consensus         51                     lllidg    fr  f  a
```

```
Spn_Pol             26  LDRFKNAAGLHTNAI..YGFQLMISHLIERVEPSHIL.....VAF.DAGK
Aae_Pol             20  LPPLSTSKGFPTNAI..YGFLRMIFSLIKKERPQYLV.....VVF.DAP.
Ctr_Pol             20  LPEMRGPNGESTQAV..FGFIRSIDKLIKDLSPEYVV.....AVF.DGPN
Ath_Pol             19  LPELTTSNNIPTNAI..YGFVNVILKYIEQEKPDYVA.....VAF.DKRG
Dra_Pol             30  LPPLNNSKGEMTDAI..VGFMKLILRLARQKSNQ.VI.....VVF.DPPV
Mtb_Pol             36  .ENFKTRGGLTTNAI..YGFTAMIINLIRDEAPTHIA.....AAF.DVSR
Taq_Pol             30  LKGLTTSRGEPVQAV..YGFAKSILKAIKEDG.DAVI.....VVF.DAKA
Bbu_Pol             24  ..PLLNTQGENVNAF..IGFFKTIFFIIKEKNPEHLI.....ITF.DSEV
Eco_Pol             26  .PPLTNSAGEPTGAV..YGVLNMIRSLIMQYKPTHAA.....VVF.DAKG
Tca_Pol             30  LKGLTTSRGEPVQAV..YGFAKSILKAIKEDGYKAVF.....VVF.DAKA
Bca_Pol             21  LPLLHNDKGIHTNAV..YGFTMMINKIIAEEEPTHML.....VAF.DAGK
Tfi_Pol             30  L.SLTTSRGEPVQMV..YGFARSILKAIKEDG.QAVV.....VVF.DAKA
Bst_Pol             21  LPLLHNDKGIHTNAV..YGFTMMINKIIAEEQPTHIL.....VAF.DAGK
Bst_PolA            21  LPLLHNDKGIHTNAV..YGFTMMINKIIAEEQPTHLL.....VAF.DAGK
Hin_Pol             26  .PSLTNAAGEPTSAV..YGVLNMIKSLISQVQPTHIA.....VVF.DAKG
Rpr_Pol             23  QKSLSSAKGEPVGAI..YGFTSMILKLISDFKPQYVA.....IVF.DSG.
Tfl_Pol             29  LKGLTTSRGEPVQAV..YGFAKSILKAIKEDG.DVVV.....VVF.DAKA
Hpy_Pol             32  .KPLTNDKGFPTGLI..TGLVGMVKKFYKDRK.NMPF.....IVF.ALES
Tpa_Pol             27  ..PLINDSGANVSAV..YGFFRSIHTLICHYRPRYFV.....AVF.DSLT
Bsu_Pol             23  LPLLSNDKGVHTNAV..YGFAMILMKMLEDEKPTHML.....VAF.DAGK
Lla_Pol             27  LDRFKAPNGLHTNAI..FAFHTMISSLMERIQPTHVL.....IAF.DAGK
Tth_Pol             30  LKGLTTSRGEPVQAV..YGFAKSILKAIKEDGYKAVF.....VVF.DAKA
Cla_Pol             25  .SGLRSSRGEPTYAV..FGFAQIILTAIAEYRPDYAA.....VAF.DVG.
Cpn_Pol             20  LPEMKNHQGQATQAV..FGFIRSINKLIKEFSPEYMI.....SVF.DGPN
Mle_Pol             37  .ENFKTRGGLTTNAI..YGFTAMIINLIREEAPTHIA.....AAF.DVSR
Rhodothermus_Pol    23  .RPLVNSKGQNTSAA..YGFTTSILKLIEEHGMDYMA.....VVF.DAGG
Synechocystis_Pol   31  KGPLRTTAGIPTSVC..FGFLNSIMQVIESQKPAAIA.....IAF.DRRE
Bsu_exo             27  .NFMINDSCVPTNVI..NGFLKHIITAIETFQPTHVV.....CCM.DMGS
Mge_exo             24  VEYAKQNNLPFNNAI..RTMLLMCWNLIKANVYQYGI.....VSF.DTKA
T3_exo              42  SWEEEIWHRCCDHAKARQILEDSIKSYETRKKA..WV GAIVLAFTD...
T4_exo              29  LVNFPDKEKINLSMVRHLILNSIKFNVKKAKTLGYTK...IVLCI.DN..
T5_exo              33  ...FKHNNSKKPFAS...SYVSTIQSLAKSYSARTTI.....VLG.DKGK
T7_exo              42  SWEEEIWHRCCDHAKARQILEDSIKSYETRKKA..WA GAIVLAFTD...
Mpn_exo             24  VEFAKLHNLPFNNAI..RTMLMMCWNLIQSKQYDYGV.....ISF.DTKA
Mtb_exo             98  .SSITAPDGRPVNAV..RGFIDSMAVVITQQRPNRLA......VCL.DLD.
Eco_exo             44  IRRIHAVQGSP..CV..ETCQHAIDQLIMHSQPTHAV......AVF.DDEN
hFEN-1              50  GDVLQNEEGETTSHLMGM.FYRTIRMMENGIKPVYVFDGKPPQLKSGELA
MjFEN-1             44  GSPLRNRKGEITSAYNGV.FYKTIHLLENDITPIWVFDGEPPKLKEKTRK
consensus          101        g     ai      l    l                 f  d


Spn_Pol             68  ..TTFRTEMYADYKGGR...........AKTPDIFREGFPFIRELIDHM
Aae_Pol             61  .AKTKREKIYADYKKQR...........PKAPDPLKVGIPVIKEILKLA
Ctr_Pol             62  .NKQSRQELYADYKSNR...........DRQLGILPEGIRLIKQYCELL
Ath_Pol             61  ..REARKSEYEEYKANR...........KPMPDNLQVGIPYIREIIYAF
Dra_Pol             71  ..KTLRHEQYEGYKSGR...........AQTPEDIRGGINRIRALIDAL
Mtb_Pol             77  ..QTFRLQRYPEYKANR...........SSTPDIFAGGIDITKEVIGAL
Taq_Pol             71  ..PSFRHEAYGGYKAGR...........APTPEDFPRQLALIKELIDLL
Bbu_Pol             64  ..PTFRKQKYPSYKATR...........DLPPDDLIPGIGWIKEGILKA
Eco_Pol             67  ..KTFRDELFEHYKSHR...........PPMPDDLRAQIEPIHAMVKAM
Tca_Pol             72  ..PSFRHEAYEAYKAGR...........APTPEDFPRQLALIKELIDLL
Bca_Pol             63  ..TTFRHEAFQEYKGGR...........QQTPPILSEQFPLIRELIRAY
Tfi_Pol             70  ..PSFRHEAYEAYKAGR...........APTPEDFPRQLALIKRLIDLL
Bst_Pol             63  ..TTFRHETFQDYKGGR...........QQTPPILSEQFPLIRELIKAY
Bst_PolA            63  ..TTFRHETFQEYKGGR...........QQTPPILSEQFPLIRELLKAY
Hin_Pol             67  ..KTFRDEMFEQYKSHR...........PPMPDDLRKQIQPIHDMIRAL
Rpr_Pol             64  .GKNFRHKIYQNYKANR...........PLPPKDLIAQLPLVRDVASNF
Tfl_Pol             70  ..PSFRHEAYEAYKAGR...........APTPEDFPRQLALIKELIDLL
Hpy_Pol             72  QTKTKRAEKLGEYKQNR...........KDAPKIMLLQIPIALEWIQKM
Tpa_Pol             67  ..PTFRHVQYPAYKAKR...........DKTSALIYAQIPLIEEILCAL
Bsu_Pol             65  ..TTFRHGTFKEYKGGR...........QKTPPILSEQMPFIRELIDAY
Lla_Pol             69  ..TTFRTEMFADYKGGR...........SKTPDIFREQLPFIKEMIEKL
Tth_Pol             72  L.PSFRHEAYEAYKAGR...........APTPEDFPRQLALIKELIDLL
Cla_Pol             65  ..RTFRDDLYAEYKAGR...........AETPEFYPQFERIKQLIQAL
Cpn_Pol             62  .NKQSRQAIYADYKSNR...........QKKFEDIPPQIALIKEYCSLI
Mle_Pol             78  ..KTFRSECYAGYKANR...........SSIPAIFHGGIDITKEVIGAL
Rhodothermus_Pol    64  EEGTFREAIYEEYKAHR...........EPPPEDLLAMLPWIKEIVRAL
Synechocystis_Pol   73  ..PTFRHEADGAYKSNR...........QETPEDFAEDLSYIQQLIEAL
Bsu_exo             68  ..KTYRNDLFQDYKANR...........SAPPVELIPQFDLAKEAAAEL
Mge_exo             66  ..PTFRDQIYEGYKQKR...........VKTPVELLVQIPLIKQALVYL
T3_exo              86  SVNWRKELVDPNYKANR...........KATKKPVGYFEFIEALFERE
T4_exo              73  AKSGYWRRDFAYIYKKNRGKAREESTWDWEGYFESSHKVIDEIKAYMPYI
T5_exo              71  ..SVFRLEHLPEYKGNR.......DEKYAQRTEEIKALDEQFFEYLKDAF
T7_exo              86  SVNWRKELVDPNYKANR...........KAVKKPVGYFEFIDALFERE
Mpn_exo             66  ..PTFRDQLYSEYKSKR...........SKTPSELLVQIPVIKESIRHL
Mtb_exo            138  WRPQFRVDLIPSYKAHRVAEPEPNGQPDVEEVPDELTPQVDMIMELIDAF
Eco_exo             84  RSSGWRHQRLPDYKAGR...........PPMPEKLHDEMPAIRAAFEQR
hFEN-1              99  KRSERRAEAEKQLQQAQAAGAEQEVEKFTKRLVKVTKQHNDECKHLISLM
MjFEN-1             93  VRREMKEKAELKMIEAIKKEDFEEAAKYAKRVSYLTPKMVENCKYLISLM
consensus          151        r      yk  r                e    q    i    l
```

52

```
Spn_Pol              104 GI.....RHYELAQYEADD      DKLAEQDG..FD  V GD DLIQLT
Aae_Pol               98 GI.....PLLELPGYEADD      AEKFSQKG..FK K Y PD DLLQLV
Ctr_Pol               99 GI.....SCLEEKGVEADD      TKKAVADG..FE C C AD DLLQLV
Ath_Pol               97 NI.....PIIEFEGYEADD      VNQFKNTG..LD V I GD DTLQLL
Dra_Pol              107 GF.....PRLEEPGYEADD      TRMAEGKG..YE R V SD DAYQLL
Mtb_Pol              113 GI.....TVLSEPGFEADD      ATQAENEG..YR L V GD DALQLV
Taq_Pol              107 GL.....ARLEVPGYEADD      AKKAEKEG..YE R L AD DLYQLL
Bbu_Pol              100 KI.....PIFEMEGYEADD      FAKKAAKNN..YLTY I PD DLLQTM
Eco_Pol              103 GL.....PLLAVSGVEADD      AREAEKAGR..P L S GD DMAQLV
Tca_Pol              108 GF.....TRLEVPGYEADD      AKNPEKEG..YE R L AD DLDQLV
Bca_Pol               99 RI.....PAYELENYEADD      AARAEQEG..FE K I GD DLTQLA
Tfi_Pol              106 GL.....VRLEAPGYEADD      AKKAEREG..ME R L GD DFFQLL
Bst_Pol               99 RI.....PAYELDHYEADD      AARAEREG..FA K I GD DLTQLA
Bst_PolA              99 RI.....PAYELDHYEADD      AARAEQEG..FE K I GD DLTQLA
Hin_Pol              103 GI.....PLLVVEGIEADD      ALQASSLGK..K L S GD DMAQLV
Rpr_Pol              101 KF.....AILEKNGYEADD      FATKTVSLG..EE I I SD DLLQLV
Tfl_Pol              106 GL.....VRLEVPGFEADD      AKRAEKEG..YE R L AD DLYQLL
Hpy_Pol              110 GF.....VCVEVNGFEADD      SPATLSP.....YKTR Y KD DFNQLL
Tpa_Pol              103 GI.....TVLRHDGFEADD      AKRVAAEH..CH V I SD DVLQLV
Bsu_Pol              101 QI.....SRYELEQYEADD      AKSAEKDG..FE K F GD DLTQLA
Lla_Pol              105 GI.....RHYELANYEADD      DKMAEAPDVNFD T V GD DMIQLV
Tth_Pol              108 GF.....TRLEVPGYEADD      AKKAEKEG..YE R L AD DLYQLV
Cla_Pol              101 NI.....PIYTAEGYEADD      ARQATERGV..DTI L GD SDVLQLV
Cpn_Pol               99 GL.....AYLEKESVEADD      AKKAREEN..YK Y C AD DLLQLV
Mle_Pol              114 GI.....TVFAEAGFEADD      ATQAENEG..YR L V GD DALQLV
Rhodothermus_Pol     102 DI.....PVIEEPGVEADD      ARRAEAHGI..D V V PD DFLQLL
Synechocystis_Pol    109 NL.....QTITYAGYEADD      ACQGSDAG..YQ K L GD DLFQLV
Bsu_exo              104 GI.....MNIGFAGYEADDC     ADLFANEA...D T V GD DLLQLL
Mge_exo              102 GF.....LVCEKDGFEADD      YANLFTKQE..IT D Y SD DMLQLV
T3_exo               123 EFYCIREPML.....E DD   V ASNPSAFGARKA . I CD DFKT P
T4_exo               123 .......VMDIDKYEADDH   V VKKFSLEG..HK L I SDGDFTQL .
T5_exo               112 ELCKTTFPTFTIRGVEADD   A Y VKLIGH..LYDH W I TDGDWDTLL
T7_exo               123 EFYCIREPML.....E DD   V ASNPSAFGARKA . I CD DFKT P
Mpn_exo              102 GF.....LVCEQDGFEADD      YARLMTQNN..VA D Y SD DLLQLV
Mtb_exo              188 GI.....AMAGAPGFEADD      ATRERRDP.... I V GD DLLQ V
Eco_exo              122 GV.....PCWSTSGNEADD   A  AVKVTQAG..HQAT V TD GYCQLL
hFEN-1               149 GI......PYLDAPSEA ASC A VKAGKVYA......AA ED MDCLTFG
MjFEN-1              143 GI......PYVEAPSEG AQASY AKKGDVWA...... V QDYDALLYG
consensus            201             Eaddiigtla        i i s Dkd  ql


Spn_Pol              147 DEHTV EI......SKKGVAEFEAFTPDY MEEM.G TPAQ....FI DLK
Aae_Pol              141 SENVL I..........NP.MNDEVFTKER IKKF.G EPQK....IP DYL
Ctr_Pol              142 SSRVS F..........NPWKEQE.IQYNE LSQF.G PPEQ....IA DYL
Ath_Pol              140 DKNVV KIV....STKFDKTVEDLYTVEN KEKY.G WANQ....VP DYK
Dra_Pol              150 DEHVK I...........ANDFSLIGPAQ EEKY.G TVRQ....WV DYR
Mtb_Pol              156 SDDVT LY......PRKGVSELTRFTPEA VEKY.G TPRQ....YP DFA
Taq_Pol              150 SDRIH L...........HPEGYLITPAW WEKY.G RPDQ....WA DYR
Bbu_Pol              143 SEYVK L.......KIE.NNSFIEMDNEY TKKF.G NSFQ....IK DYL
Eco_Pol              146 TPNIT I..........NTMTNTILGPEE VNKY.G PPEL....II DFL
Tca_Pol              151 SDRVA L.......HPEGHLITPEY WQKY.G KPEQ....WV DFR
Bca_Pol              142 SPHVT DI......TKKGITDIEPYTPEA VREKY.G TPEQ....IV DLK
Tfi_Pol              149 SEKVS L.......LPDGTLVTPKD QEKY.G PPER....WV DFR
Bst_Pol              142 SPQVT EI......TKKGITDIESYTPET VEKY.G TPEQ....IV DLK
Bst_PolA             142 SRHVT DI......TKKGITDIEPYTPET VREKY.G TPEQ....IV DLK
Hin_Pol              146 DDNIM I..........NTMNNSLLDRKG VIEKY.G PPEL....II DYL
Rpr_Pol              144 SKNIK Y.........DP.IKCKYITEDN VTKF.G TTPDK....LR EVM
Tfl_Pol              149 SERIA L.......HPEGYLITPAW IYEKY.G RPEQ....WV DYR
Hpy_Pol              150 SDKIA F............DGKTEFLAKDCVEKY.G LPSQ....FT DYQ
Tpa_Pol              146 CDTVQ L.......RLDIDHKWTCCDAAY QQRW.T MPTQ....LL DLF
Bsu_Pol              144 TDKTT AI......TRKGITDVEFYTPEH VKERY.G TPEQ....II DMK
Lla_Pol              150 DGNTR EI......SKKGVAEFEEFTPDY LLEKM.G TPSQ....FI DLK
Tth_Pol              151 SDRVA L...........HPEGHLITPEW LWEKY.G RPEQ....WV DFR
Cla_Pol              144 NDHVR ALA....NPYGGKTSVTLYDLEQ VRKRYDG EPDQ....LA DLR
Cpn_Pol              142 NDHVR AW.........NPWADQGVVGISE VIERY.G PPGN....IP DYL
Mle_Pol              157 SNDVT LY......PRKGVSELTRFTPEA VIEKY.G VTPAQ....YP DLA
Rhodothermus_Pol     145 S..PR SIY....KPARRGETFDLITIET FRETY.G EPHQ....FI DVL
Synechocystis_Pol    152 SPEKN SVLYLTRNPFSSNTGYDELDWQG VVDKM.G VTPAQ....IV DFK
Bsu_exo              146 TDKVS ALL......QKGIGNYKVYTKETFYEE.T GVMPKA....LI DIK
Mge_exo              145 NAFTN FL......CIKGTKEMVMYNNENFKSLFY GLAPYQ....VV EYK
T3_exo               167 N................CDFLWCTTGN LTQ.....TKETADWWHLFQ
T4_exo               162 ...............HKYPNVKQWSP..MHKKWVK IKSGSAEIDCMT.K
T5_exo               160 TDKVSRF.........SFTTRREYHLRD MYEHHNVDDVEQ....FISLK
T7_exo               167 N................CDFLWCTTGN LTQ.....TEESADWWHLFQ
Mpn_exo              145 DSMTS WL......CVKGTKEMKEYNTDNFAEQFF G TPKQ....VI YK
Mtb_exo              229 ADDPVPVRVLYLG...RGLAKATLFGPAE VAERY.G LPAHRAGAAYA LA
Eco_exo              165 SPTLR RDYFQ..........KRWLDAPFIDKEF G QPQQ....LP DYW
hFEN-1               187 SPVLMRHL....TASEAKKLPIQEFHLSRI LQEL.G LNQEQ....FV DLC
MjFEN-1              181 APRVVRNL....TT..TKEMP.ELIELNEV LEDL.R ISLDD....LI DIA
consensus            251       v                       1    gl        d
```

```
Spn_Pol             186  ALMGDKSDNIPGVTKVGEKTGIKLLLEHGSLEGIYE..N.....IDGMKT
Aae_Pol             176  ALVGDKVDNVPGLEGVGPKTAINLLKKYGSVENLLKN.......WEKF..
Ctr_Pol             177  ALVGDSSDNIPGVSGCGPKKAQALLKEFQSVKELVAN.......TERLSG
Ath_Pol             181  ALVGDQSDNIPGVKGIGEKSAQKLLEEYSGLETYQ..N.....LDKIKS
Dra_Pol             184  ALTGDASDNIPGAKGIGPKTAAKLIQEYGTLKVYEAAH.....AGTLKP
Mtb_Pol             195  ALRGDPSDNIPGIPGVGEKTAAKWLAEYGPIRSLVDNVDAVR........
Taq_Pol             184  ALTGDEVSDNIPGVKGIGEKTARKLLEEWGSLEALLK..N.....LDRLKP
Bbu_Pol             180  AVVGDRSDNIPGIKGCAKGAANLIREFKLIDGIYSNLEIIN.....
Eco_Pol             181  AIMGDSSDNIPGVPGVGEKTAQALLQGLGGLDTIYAEPEKIA.GLSFRGA
Tca_Pol             185  ALVGDPSDNIPGVKGIGEKTALKLLKEWGSLENLLK..N.....LDRVKP
Bca_Pol             181  GLMGDKSDNIPGVPGVGEKTAVKLLRQFGTVENVLA..S.....IDEIKG
Tfi_Pol             183  ALTGDRSDNIPGVAGIGEKTALRLLKEWGSLENLLK..N.....LDRVKP
Bst_Pol             181  PLMGDKSDNIPGVPGVGEKTAVKLLKQFGTVENVLA..S.....IDEIKG
Bst_PolA            181  GLMGDKSDNIPGVPGVGEKTAVKLLKQFGTVENVLA..S.....IDEVKG
Hin_Pol             181  ALMGDSADNIPGVAGVGEKTALGLLQGIGSMAELYANLEKVA.ELPIRGA
Rpr_Pol             179  ALIGDRSDNIPGIPSIGPKTASSLITKFGSVENLFNS.......LDQISS
Tfl_Pol             183  ALAGDPSDNIPGVKGIGEKTAQRLIREWGSLENIFQ..H.....LDQVKP
Hpy_Pol             183  GIVGDSSDNYKGVKGIGSKNAKELLQRLGSIKKIYENLDLAK....NLLS
Tpa_Pol             184  SDMGDSSDNVPGVRGIGPKTAAHLLHCFGTLDGIYRHTYSLK........
Bsu_Pol             183  QLMGDSSDNIPGVPGVGEKTAIKLLKQFDSVAKLLE..S.....IDEVSG
Lla_Pol             189  ALMGDSSDNYPGVTKVGEKTGLKLLQEFGSIENDYE..N.....VETLKA
Tth_Pol             185  ALVGDPSDNIPGVKGIGEKTALKLLKEWGSLENLLK..N.....LDRVKP
Cla_Pol             186  GLKGDTSDNIPGVRGIGEKGAIALLKQFREIDALLEQIDAA........P
Cpn_Pol             178  ALVGDSSDNIPGIPGCGPKKAAALLKQFGSVEGILEN.......LDAVKG
Mle_Pol             196  ALRGDPSDNIPGIPGVGEKTAAKWIVDYGPQGIVDNVESVR........
Rhodothermus_Pol    184  ALMGDSSDNIPGIPGVGEKTAVQLLQQYGSVENLAHAEEVK.......G
Synechocystis_Pol   197  ALCGDKSDCIPGINGIGEKTAIKLLAEYEGEKLYENLAQIK........
Bsu_exo             185  ALMGDSSDNYPGVKGIGEKTAYKLLREYEGDRLLENLS........LLP
Mge_exo             185  GLVGDNSDNIAGIKGIGPIKGIELLQQYGIAGNIYTNFNNL........P
T3_exo              194  TIKGDMTDGYSGIPGWGD.TIEGFIN.....IPFIVEPVESVLKSGKNKG
T4_exo              193  ILKGDKDNIASVKVRSDFW.......KTRVIGERTPSMKTS.IVEAIAN
T5_exo              196  AIMGDLGDNIRGVEGIGAKRGYNIIREFGNVLDIIDQ.....LPLPGK
T7_exo              194  TIKGDITDGYSGIAGWGD.TAEDFIN.....NPFITEPKTSVLKSGKNKG
Mpn_exo             185  QLVGDNSDNITGLKGIGPKKGIDLLKQYGTIDNIFANFDKL........S
Mtb_exo             275  LIRGDPSDNIPGVPGVGEKTAATLLARHGSLDQIMAAADDRK....TTMA
Eco_exo             200  GLAGISSSKVPGVAGIGPKSATQLIVEFQSIENLDAV........A
hFEN-1              228  ILLG..SDYCESIRGIGPKRAVDLIQKHKSIEELVRRLDPNKYPVPENWL
MjFEN-1             219  IFMG.TDYNPGCVKGIGFKRAYELLVRSGVAKDVEKKEVE..........Y
consensus           301  al Gd sdni gv gvg k a  ll     sle i
```

```
Spn_Pol             229  SKMKEN.LINDKEQAFLSKTLATLDTK.A...PIAIGLEDLVYSGP..DV
Aae_Pol             217  ..QREF.PRAKKEDLELSYKLVKLYTD.L...DIELSEEDLKIKRP..DL
Ctr_Pol             220  KTK.QM.LEDQKETLLLSKRLATLHMD.L...AFPLTTEEFAFSPQAIDS
Ath_Pol             224  S.IREK.LEAGKDMAFLSKRLATLVCD.L...PLNVKLEDLRTKEW..NK
Dra_Pol             229  DGTRKK.LLDSEENVKFSHDLSCLVTD.L...PLDIEFG.VRRLPD..NP
Mtb_Pol             237  GKVGDA.LRANLASVVRNRELTDLVRD.V...PLAQTPDTLRLQPW..DR
Taq_Pol             227  .AIREK.LLAHMDDLKLSWDLAKVRTD.L...PLEVDFA..KRREP..DR
Bbu_Pol             222  KKHREL.LIKEKENAFLSYELVSLEEN........LKIPEIENFALKNFS
Eco_Pol             230  KTMAAK.LEQNKEVAYLSYQLATIKTD.V...ELELTCEQLEVQQP..AA
Tca_Pol             228  ENVREK.IKAHLEDLRLSLELSRVRTD.L...PLEVDLA..QGREP..DR
Bca_Pol             224  EKLKET.LRQHREMALLSKKLAAIRRD.A...PVELSLDDIAYQGE..DR
Tfi_Pol             226  DSLRRK.LEAHLEDLHLSLDLARIRTD.L...PLEVDFKALRRRTP..DL
Bst_Pol             224  EKLKEN.LRQYRDLALLSKQLAAICRD.A...PVELTLDDIVYKGE..DR
Bst_PolA            224  EKLKEN.LRQHRDLALLSKQLASICRD.A...PVELSLDDIVYEGQ..DR
Hin_Pol             230  KKLGEK.LLAEKNNADLSYTLATIKTD.V...ELNVTTDQLLLGES..QK
Rpr_Pol             222  IKQRKT.LQNAREAALISWKLIGLDSN.V...DLDFDLNNLKWSPP..NS
Tfl_Pol             226  .SLREK.LQAGMEALALSRKLSQVHTD.L...PLEVDFG..RRRTP..NL
Hpy_Pol             229  PKMYRA.LIHDKASAFLSKELATHERG.C...IKEFDFLSCAF.PSENPL
Tpa_Pol             226  EALRTK.IVCGKKDAFFSRSLIELRDD.V...PCVFSLEDSCCIPLDVTS
Bsu_Pol             226  KKLKEK.LEEFKDQALMSKELLATIMTD.A...PIEVSVSGLEYQGF..NR
Lla_Pol             232  SKMKDN.LIADKEMAFLSQQLATINTA...PLEIGLEDTLLKEK..NV
Tth_Pol             228  ENVREK.IKAHLEDLRLSLELSRVRTD.L...PLEVDLA..QGREP..DR
Cla_Pol             228  KRYQTL.LREQAEAARFSRQLATIVTD.V...PVQLDLEAARIGVY..DR
Cpn_Pol             221  LSQ.TM.LSERQETLKLSKRLALLDSN.I...PIPVPIESLTFPQHPVDE
Mle_Pol             238  GKVGEA.LRTHLASVVRNRELTELVKD.V...PLVQTSDTLRLQPW..DR
Rhodothermus_Pol    227  KRAREG.LLNHREEALLSKRLVTIRTD.V...PLRIRWEAFHRARP..DL
Synechocystis_Pol   239  GALKTR.LDNGKDDAMHSQMLARIVVD.V...PLPVTWEDLQLTGF..ST
Bsu_exo             227  KGQQGK.IQQGLSDLEMSRKLAEIHCS.V...PLACTLKDALFTLMEQA
Mge_exo             227  NQLQKL.LNNQKEIAKTFSFLAKIKTD.I...ELDQNIDLTGLKP..IQK
T3_exo              238  QTVTKWVKRAPDATETLWDCIKSIGAKAGMTEQEIIKQGQMARILRFEEY
T4_exo              235  DREQAKVLLTESEYNRYKENLVLLDFD.YIPDNIASNIVNYYNSYKLPPR
T5_exo              239  QKYIQN.LNASEELLFRNLIVDLPTY.C...VDAIAAVGQDVLDKF..T
T7_exo              238  QEVTKWVKRDPEPHETLWDCIKSIGAKAGMTEEDIIKQGQMARILRFNEY
Mpn_exo             227  KALQTI.LQGQIDTAKKFSFLASIKTD.I...KLNDDIVHAALKP..IDK
Mtb_exo             321  KGLRTK.LLAASAYIKAADRVVRVATD.A...PVTLSTPTDRFPLVAADP
Eco_exo             242  EKWRKK.LETHKEMAFLCRDIARLQTD.L...HIDGNLQQLRLVR.....
hFEN-1              276  HKEAHQLFLEPEVLDPESVELKWSEPNEEELIKFMCGEKQFSEERIRSGV
MjFEN-1             258  YDEIKRIFKEPKVTDNYSLSLKL..PDKEGIIKFLVDENDFNYDRVKKHV
consensus           351         l              l     l  i
```

54

Figure S3