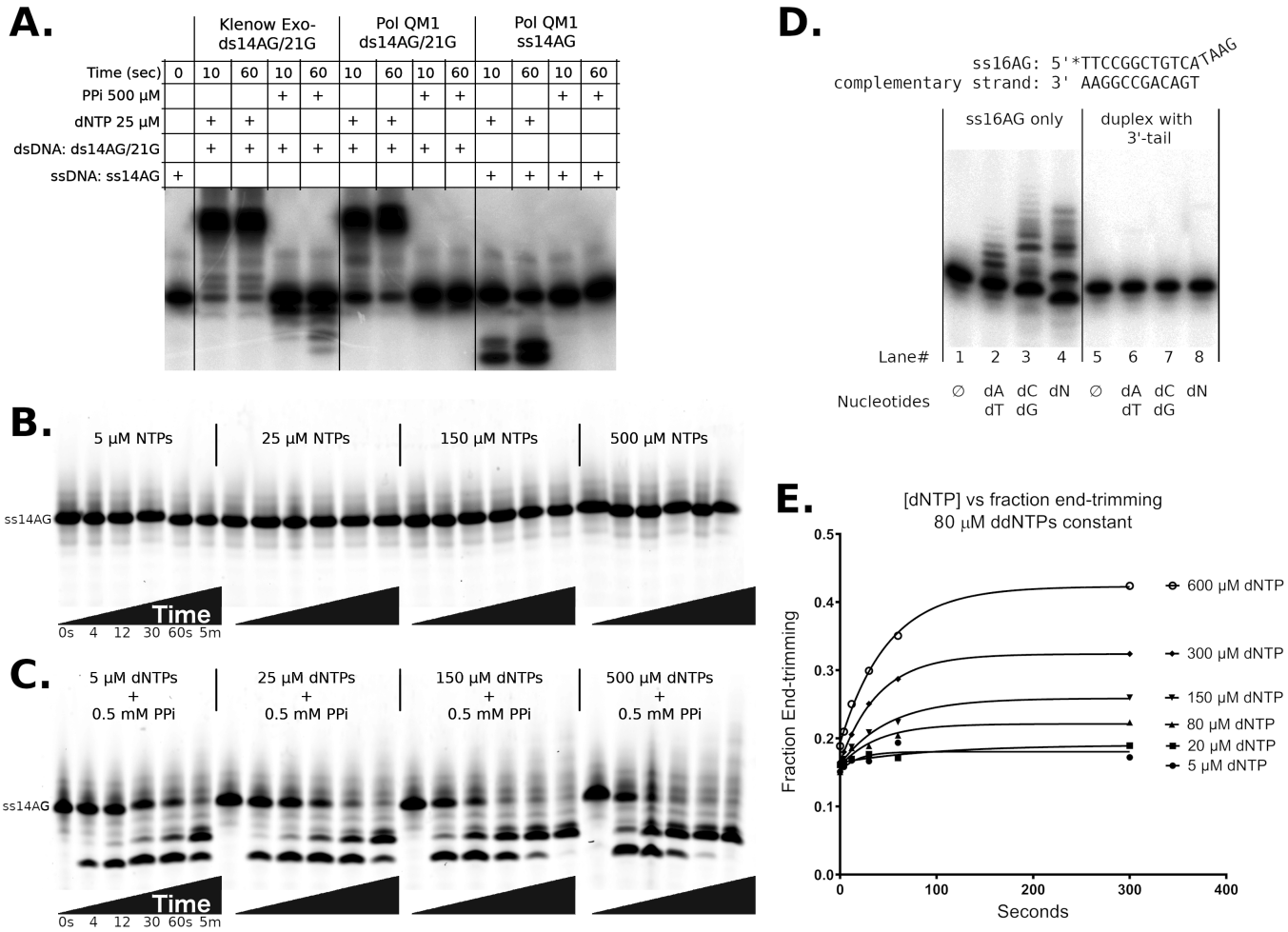



**Figure S1. End-trimming by Pol θ does not proceed by pyrophosphorolysis but is dNTP-dependent, Related to Figure 2**

(A) Left: Pol Klenow fragment exo- extends ds14AG/21C to full-length product with 25 µM dNTPs but degrades the primer strand by pyrophosphorolysis when nucleotides are substituted with 500 µM PPi. Middle: Pol QM1 extends the substrate when supplied with dNTPs but fails to degrade DNA by pyrophosphorolysis. Right: When provided with ss14AG, pol QM1 catalyzes end-trimming dependent on dNTPs but cannot degrade the ssDNA when provided with 500 µM PPi.

(B) Ribonucleotides do not support end-trimming of ss14AG.

(C) PPi (500 µM) has no effect on the nucleotide dependence of end-trimming (compare to **Fig 2D**).

(D) Pol QM1 is not a structure-specific nuclease that clips off 3' tails. Substrate ss16AG, labeled at the 5' end as shown by the asterisk, is subject to end-trimming and extension (lanes 1-4, 1 min reactions, 150 µM each nucleotide). A duplex substrate with a 3' tail was formed by annealing the same oligonucleotide to a complementary 12-mer. The 3' tail is not a substrate for end-trimming or extension (lanes 5-8)

(E) End-trimming of ss14AG was conducted by titrating dNTPs into reactions supplemented with 80 µM ddNTPs. The amplitude of the pre-steady state burst is limited to ~40% due to competition with ddNTPs.

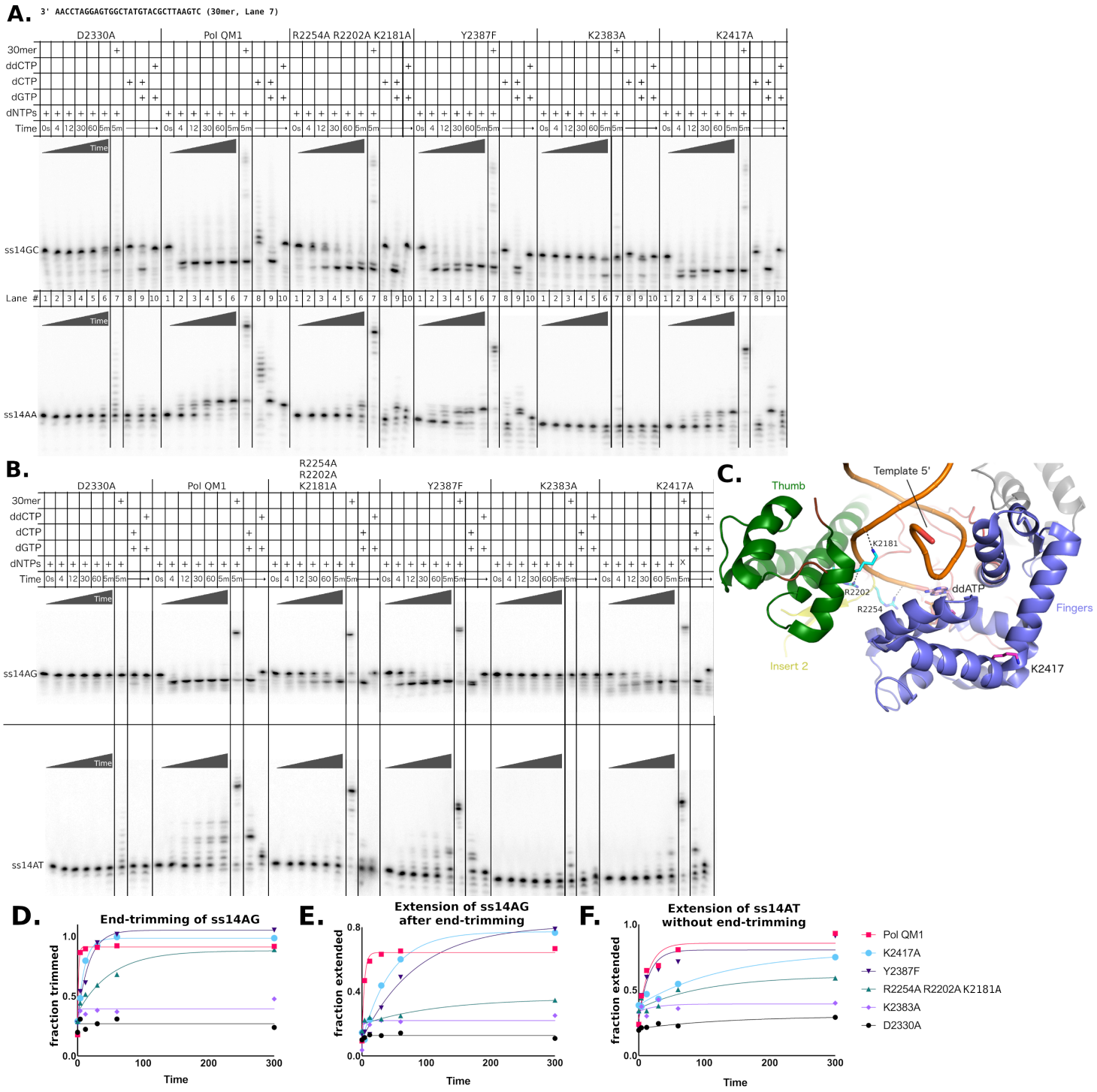



**Figure S2. End trimming and extension of additional sequences are catalyzed by Pol θ and variant polymerases, Related to Figure 3**

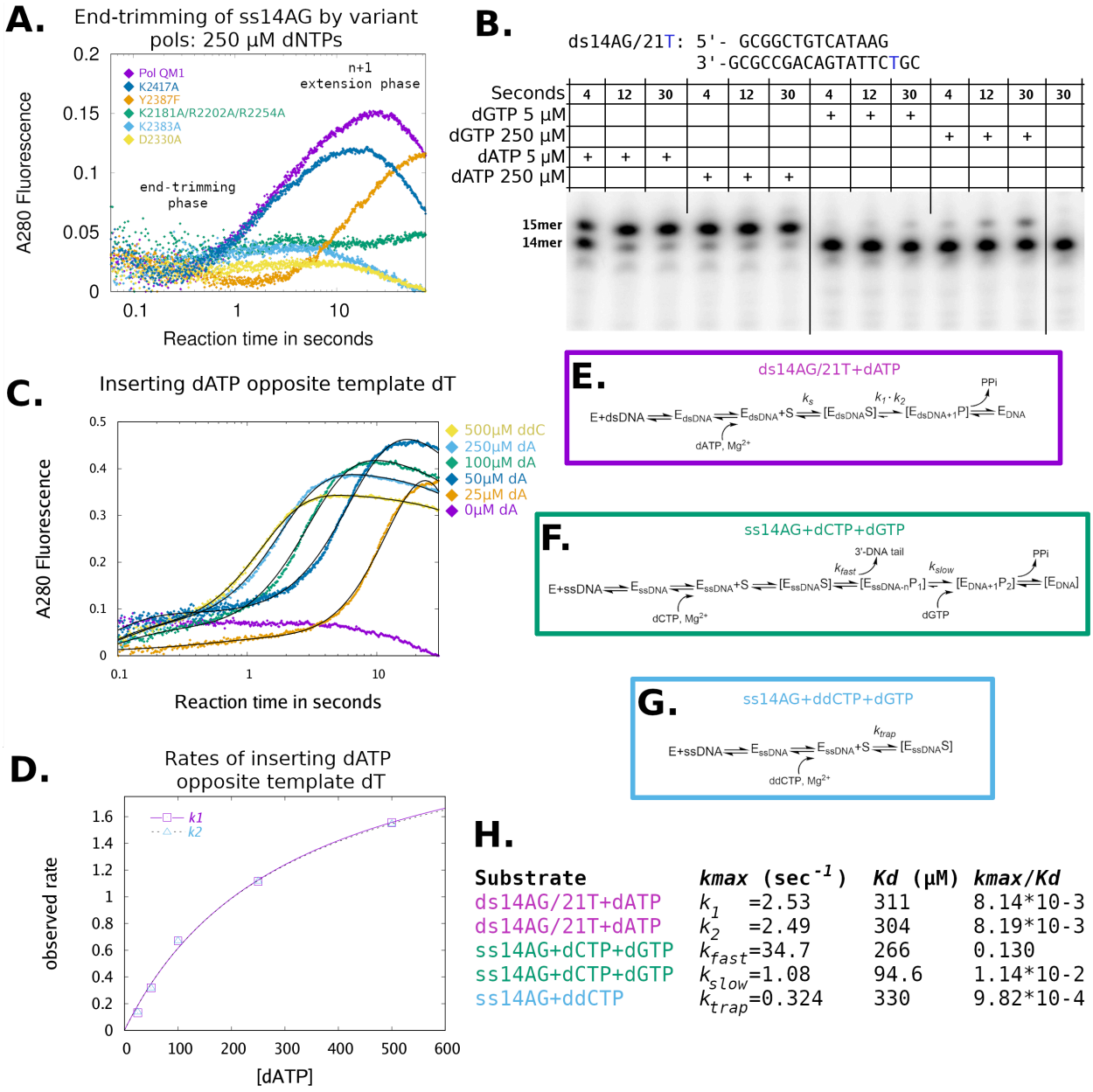
(A) Denaturing gels for additional variants of pol QM1, using ss14GC and ss14AA substrates

(B) Denaturing gels for additional variants of pol QM1, using ss14AG and ss14AT substrates

(C) The view in Fig 3A is rotated to provide a line-of-sight perpendicular to the DNA double-helix. The fingers subdomain cleft is occupied by K2417 and available for contacts to the 5′-end of the template DNA after conformational rearrangements.

(D) Quantification of end-trimming of ss14AG, for pol QM1 variants and controls

(E, F) Quantification of extension after end-trimming for ss14AG and ss14AT




H.

| Substrate | kmax (sec$^{-1}$) | Kd (μM) | kmax/Kd |
|---|---|---|---|
| ds14AG/21T+dATP | $k_1$ =2.53 | 311 | 8.14*10-3 |
| ds14AG/21T+dATP | $k_2$ =2.49 | 304 | 8.19*10-3 |
| ss14AG+dCTP+dGTP | $k_{fast}$=34.7 | 266 | 0.130 |
| ss14AG+dCTP+dGTP | $k_{slow}$=1.08 | 94.6 | 1.14*10-2 |
| ss14AG+ddCTP | $k_{trap}$=0.324 | 330 | 9.82*10-4 |

**Figure S3. Supporting analysis for stopped-flow spectroscopy of Pol θ, Related to Figure 4**

(A) Reaction fluorescence traces for variant polymerases reveal the predicted lag for variant Y2387F, especially during the second phase when the nascent primer terminus is extended by n+1 nucleotide.

(B) Primer-extension of ds14AG/21T shows that incorporation of 5 μM dATP is less than half complete by 4 secs. With 250 μM dATP, the reaction is complete by 4 secs, correlating the increase in fluorescence during the slow reaction phase with active primer elongation (i.e., fingers closing).

(C) The control substrate ds14AG/21T is analyzed by stopped-flow spectroscopy during incorporation of dATP. The curves are fit (black solid lines) to the traces as described in the extended methods.

(D) Rate-constants for the second rate-limiting step during incorporation of dATP into ds14AG/21T are plotted.

(E-G) The reaction models consistent with the three kinetics experiments are depicted, with measured rate-constants annotated.

(H) Table of rate and binding constants.

**A.**

| Substrate | Mapped Reads | Primer Reads (%) | Template Reads (%) | Sum of Primer- and Template-type Reads |
|---|---|---|---|---|
| ds14GA'/30A | 4,552,728 | 3' CCTAGGAGTGGCTATGTACGCTTAAG<br>5' GGATCCTCA**TA**GATACATGCGAATTC (48%) | 3' CCTAGGAGTGGCTATGTACGCTTAAG<br>5' GGATCCTCACCGATACATGCGAATTC (52%) | 221,751 |
| ds14AG'/30A | 4,980,185 | 3' CCTAGGAGTGG-CTATGTACGCTTAAG<br>5' GGATCCTCA**TA** **G**ATACATCGGAATTC (57%)<br>**A** | 3' CCTAGGAGTGGCTATGTACGCTTAAG<br>5' GGATCCTCACCGATACATCGGAATTC (43%) | 98,779 |
| ds28GA/54A | 19,308,460 | 3' CCTAGGTGCTCCGACAGTGGCTATGTACGACTGCTCTTAAG<br>5' GGATCCACGAGGCTGTCA**TA**GATACATGCTGACGAGAATTC (51%) | 3' CCTAGGTGCTCCGACAGTGGCTATGTACGACTGCTCTTAAG<br>5' GGATCCACGAGGCTGTCACCGATACATGCTGACGAGAATTC (49%) | 18,148,894 |
| ds28AG/54A | 16,363,210 | 3' CCTAGGTGCTCCGACAGTGG-CTATGTACGACTGCTCTTAAG<br>5' GGATCCACGAGGCTGTCA**TA** **G**ATACATGCTGACGAGAATTC (51%)<br>**A** | 3' CCTAGGTGCTCCGACAGTGGCTATGTACGACTGCTCTTAAG<br>5' GGATCCACGAGGCTGTCACCGATACATGCTGACGAGAATTC (49%) | 14,978,894 |

nascent DNA

**B.**

**Template position *vs* diversity of insertion sequences**

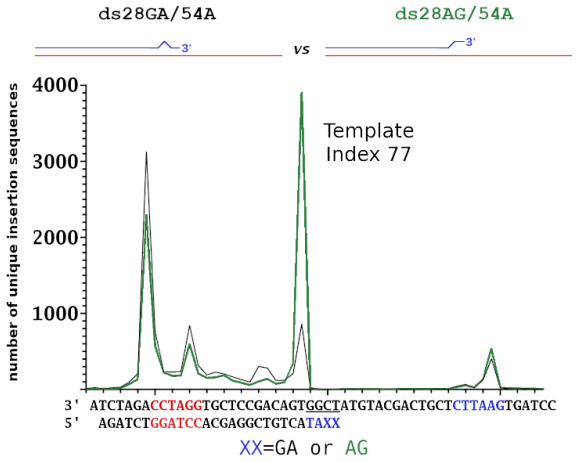



**Figure S4. Supporting and extended data for sequencing experiment with Pol QM1, Related to Figure 5**

(A) Major read sequences are summarized in table form. Approximately $20 \cdot 10^6$ reads were collected for each substrate. The reported sum of primer-type and template-type reads includes sequences mapping to the predicted insert sequence between the BamHI and EcoRI restriction site but does not include single nucleotide polymorphisms or cloning artifacts in the flanking regions.

(B) The number of different sequences inserted at the primer terminus (template index 77) is elevated ~4-fold for the substrate ds28AG/54A (green) *vs.* ds28GA/54A (black). At other sites, the two substrates show similar insertion diversity, suggesting that the 4 terminal mismatches at the 3′-primer end of ds28AG/54A are responsible for the elevated rate.

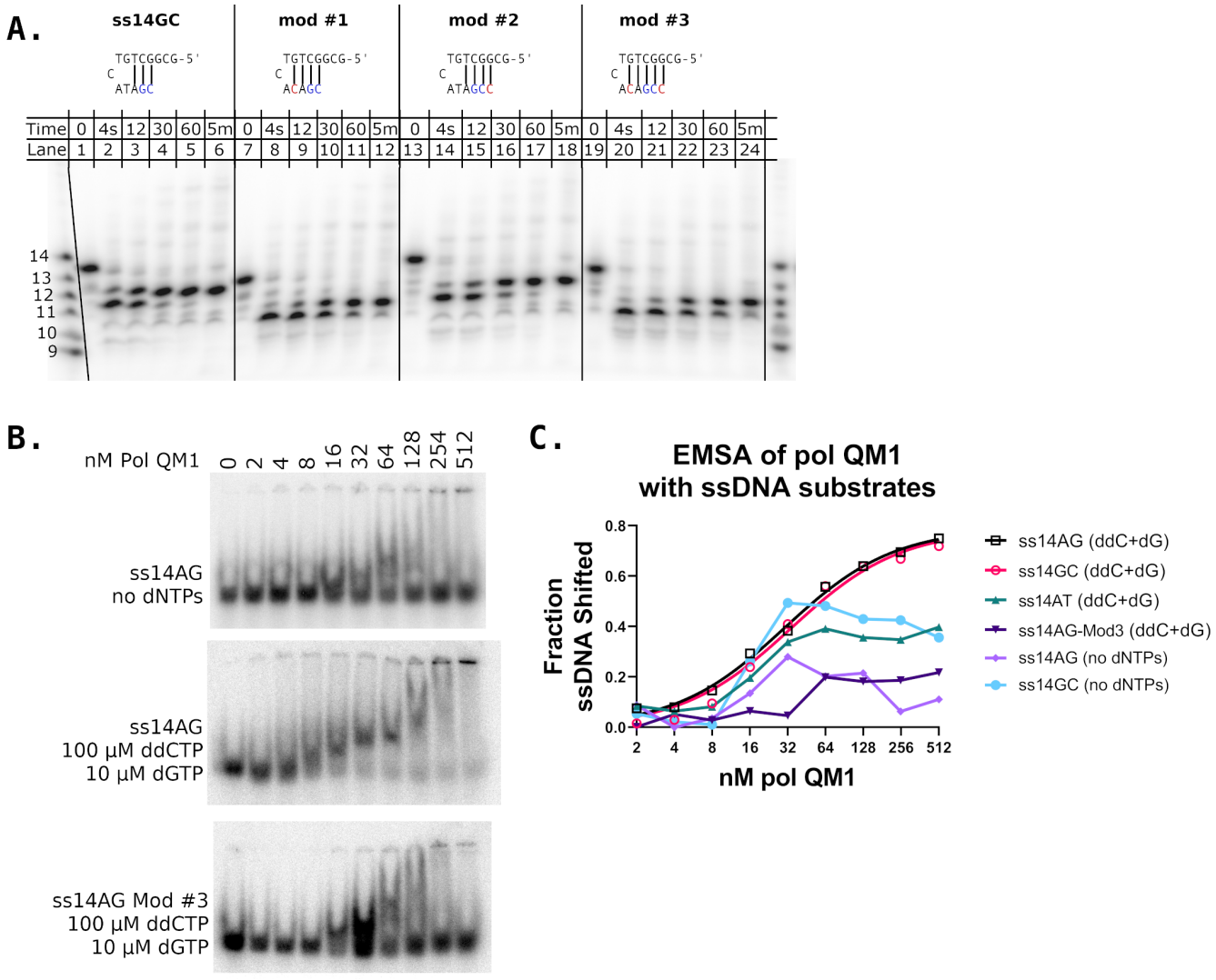



**Figure S5. Supporting analysis of ss14GC hairpin stem and additional EMSA data, Related to Figure 6**

(A) End-trimming assays with 50 µM dNTPs are depicted for ss14GC and three related DNA sequence modifications that are predicted to enhance stability of the putative hairpin.

(B) Examples of EMSA gels of pol QM1 complexes with the indicated DNA substrates (250 pM) and nucleotides. In reaction mixtures excluding nucleotides, or containing untrimmed sequences, the fraction of DNA complexes did not increase with higher concentrations of enzyme, possibly due to oligomerization of pol QM1 (Zahn et al., 2015) at the expense of ssDNA binding.

(C) The gels shown in **Fig S5B** and **Fig 6C** were quantified and plotted, omitting error bars for clarity. The data for ss14AG and ss14GC were fit to a hyperbola. Mean values for three replicate experiments are shown for each data set.

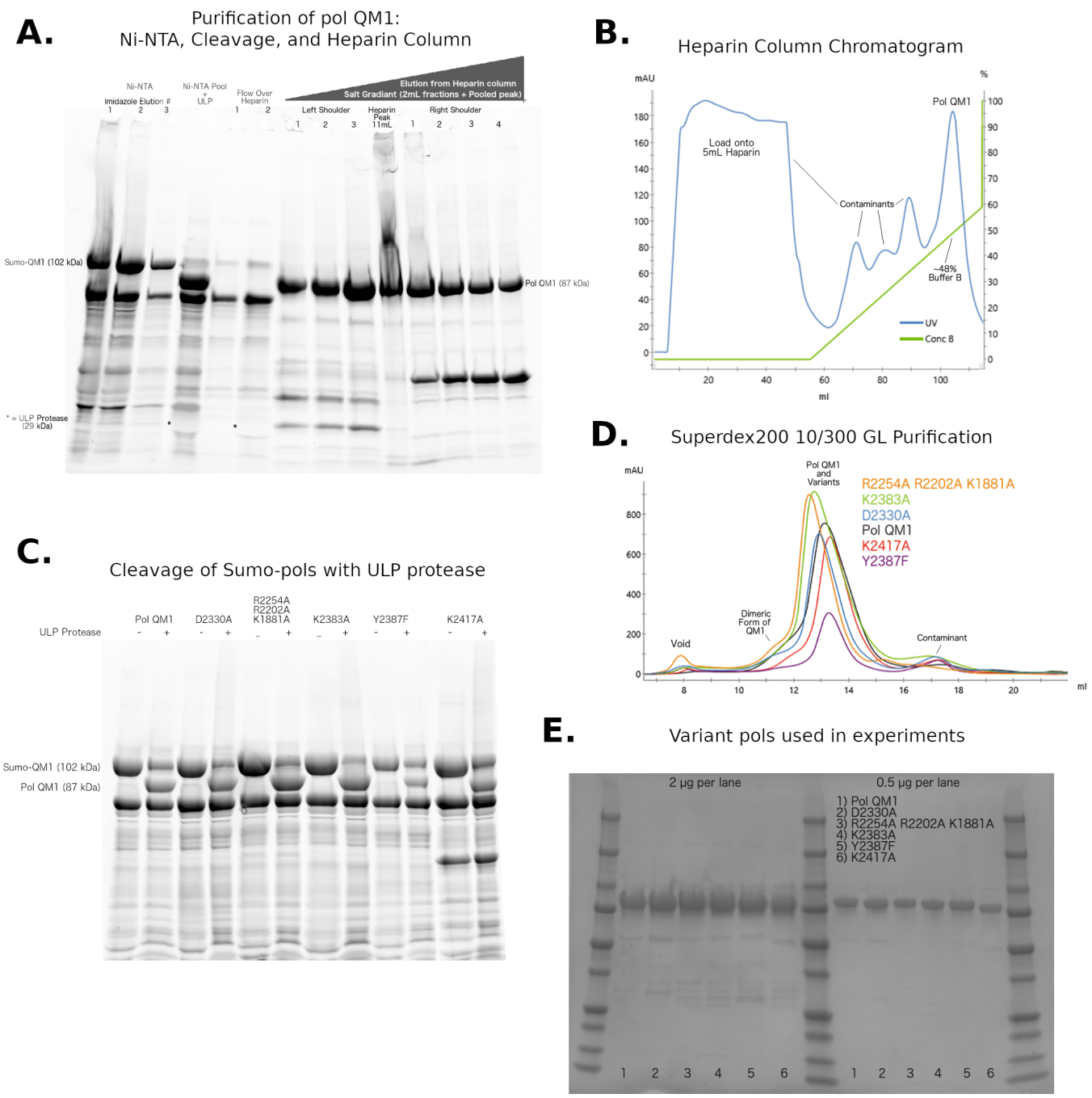



**Figure S6. Purification of pol QM1, showing gels and chromatograms from purification of parental and variant enzymes, Related to Figures 2 and 3**

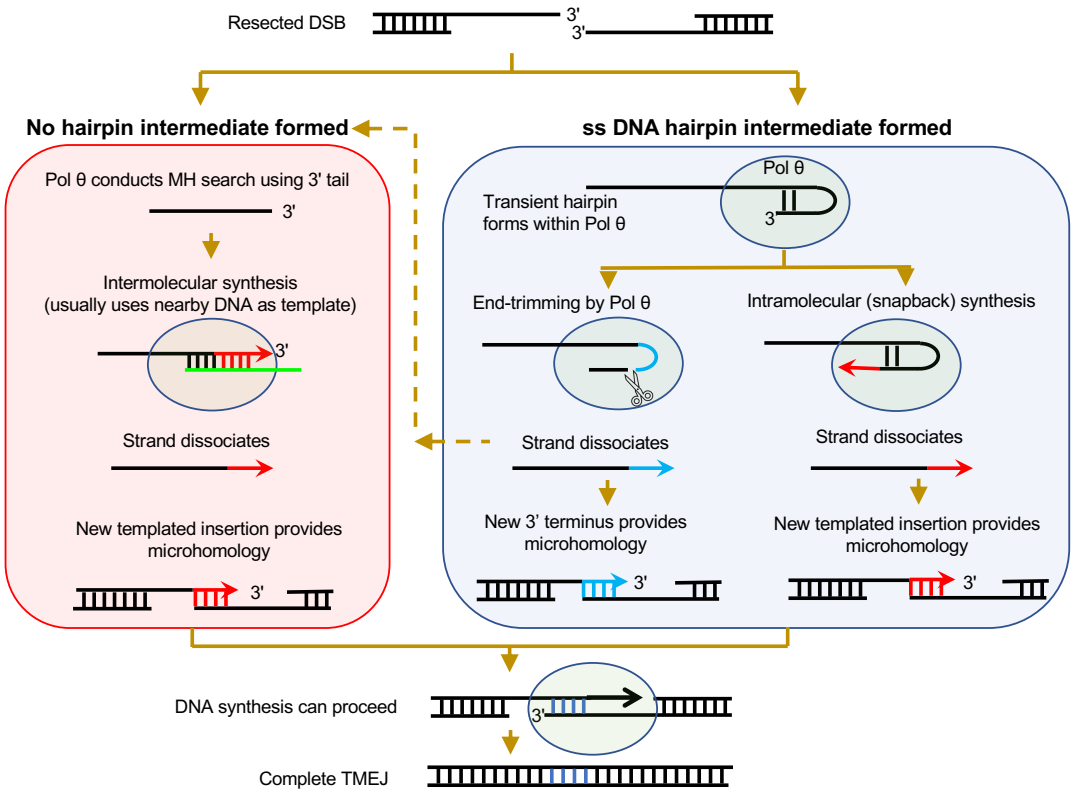



**Figure S7. Ways by which TMEJ products can arise, Related to Figure 7**

At the top, resection at a double strand break site produces two unpaired DNA tails with 3´ ends. Microhomologies (MH) of 3-6 bp at the repaired break junction are a hallmark of TMEJ, indicating that Pol θ utilizes MH but usually does not extend unpaired 3´ ends. The MH search is conducted bidirectionally from the 3´ ends, usually to a maximum of 15 nt inwards (Carvajal-Garcia et al., 2020). Experiments in multiple organisms, including human cells, show that joining products of TMEJ usually involve short deletions. Sequencing indicates that some products contain short insertions, templated by one or more cycles of self-templated (intramolecular) synthesis, or templating from another piece of DNA, usually nearby (intermolecular synthesis).

*Red-shaded box left:* If no hairpin can be formed within Pol θ (or if the hairpin has been cleaved off), the 3´ tail may engage in intermolecular DNA synthesis. This usually uses a terminally complementary sequence near the break (although more distant templates are sometimes used). This synthesis can generate new microhomology possibilities for a subsequent search.

*Blue-shaded box, right:* Pol θ may manipulate the 3´ end of a DNA tail into a transient hairpin, which is not stable in solution. Two possible outcomes are end-trimming by Pol θ, shown at left, and intramolecular synthesis, shown at right. End-trimming creates a different 3´ end, broadening the possibility of a successful microhomology search, which can then channel into the pathway at left as shown by the dotted line. Intramolecular (or snapback) synthesis produces a short insertion sequence in the form of an inverted repeat, which also provides a new 3´ end for a microhomology search. Multiple cycles of these events may occur.

**Table S1. Oligonucleotide sequences utilized in this work, Related to STAR Methods.**

| DNA Oligonucleotide | Sequence (5' to 3') |
|---|---|
| ss14AG | GCGGCTGTCATA**AG** |
| 21C (template) | GACCG**C**CTTATGACAGCCGCG |
| 21T (template) | GACCG**T**CTTATGACAGCCGCG |
| ss13AA | GCGGCTGTCAT**AA** |
| ss14CC | GCGGCTGTCATA**CC** |
| ss14GC | GCGGCTGTCATA**GC** |
| ss14G-ddC | GCGGCTGTCATA**G-ddC** |
| ss14AA | GCGGCTGTCATA**AA** |
| 30A, "30mer" (template) | CTGAATTCGCATGTATCGGTGAGGATCCAA |
| ss14GA' | TGGATCCTCATA**GA** |
| ss14AG' | TGGATCCTCATA**AG** |
| 54A (template) | CCTAGTGAATTCTCGTCAGCATGTA**TCGG**TGACAGCCTCGTGGATCCAGATCTA |
| ss14AG-mod1 | **A**CGGCTGTCATAAG |
| ss14AG-mod2 | ACG**A**CTGTCATAAG |
| ss14AG-mod3 | GCGG**A**TGTCATAAG |
| ss14CG | GCGGCTGTCATA**CG** |
| ss14AC | GCGGCTGTCATA**AC** |
| ss14AT | GCGGCTGTCATA**AT** |
| ss16AG | TTCCGGCTGTCATAAG |
| ss16AG complement | TGACAGCCGGAA |