Mechanism of microhomology-mediated end-joining promoted by human DNA polymerase θ

Tatiana Kent1,2, Gurushankar Chandramouly1,2, Shane Michael McDevitt1,2, Ahmet Y Ozdemir1,2 & Richard T Pomerantz1,2

Microhomology-mediated end-joining (MMEJ) is an error-prone alternative double-strand break–repair pathway that uses sequence microhomology to recombine broken DNA. Although MMEJ has been implicated in cancer development, the mechanism of this pathway is unknown. We demonstrate that purified human DNA polymerase θ (Polθ) performs MMEJ of DNA containing 3’ single-strand DNA overhangs with ≥2 bp of homology, including DNA modeled after telomeres, and show that MMEJ is dependent on Polθ in human cells. Our data support a mechanism whereby Polθ facilitates end-joining and microhomology annealing, then uses the opposing overhang as a template in trans to stabilize the DNA synapse. Polθ exhibits a preference for DNA containing a 5’-terminal phosphate, similarly to polymerases involved in nonhomologous end-joining. Finally, we identify a conserved loop domain that is essential for MMEJ and higher-order structures of Polθ that probably promote DNA synapse formation.

Genome instability in the form of chromosome breaks, deletions and rearrangements is a hallmark of cancer cells and a driver of tumorigenesis. Mounting evidence has indicated that MMEJ, an error-prone alternative form of double-strand break (DSB) repair, promotes inter- and intrachromosome rearrangements associated with DNA deletions by using sequence microhomology to recombine broken DNA ends1–6. MMEJ has been documented in various eukaryotic organisms, and genetic data have demonstrated that this process is distinct from the classical nonhomologous end-joining (NHEJ) pathway1,2,4. For example, MMEJ functions in a Ku- and ligase IV–independent manner and is therefore referred to as an alternative end-joining (alt-EJ) pathway2,4. Considering that the majority of Ku- and ligase IV–independent end-joining occurs via a microhomology-mediated mechanism, MMEJ appears to be the major form of alt-EJ2,7,8.

Although the genetic requirements for MMEJ have not fully been identified, studies in mammalian cells have shown that MMEJ is promoted by PARP-1, ligase III, CtIP and Mre11 (refs. 1,3,4,9–12). Recent evidence has indicated that MMEJ is induced during S phase and G2 and that it functions in response to replicative stress, albeit at substantially lower levels than homology-directed repair (HDR), which is highly accurate and therefore important for preserving genome integrity3. Intriguingly, MMEJ and HDR use the same initial DNA-resection machinery, which includes Mre11 and CtIP and facilitates the formation of 3’ single-strand DNA (ssDNA) overhangs at DSBs3. MMEJ, however, requires only limited resection, whereas HDR depends on extensive resection performed by additional factors (i.e., exonuclease I, Bloom’s helicase and Dna2)3. Importantly, MMEJ uses sequence microhomology exposed by limited DNA resection to join DNA ends or stabilize spontaneously annealed end-joining intermediates1,2. This mechanism is in contrast to NHEJ, which does not require DNA resection, owing to its ability to join blunt-ended DNA and DNA containing short overhangs with little or no homology4. Recent comprehensive studies in human cells have demonstrated that the majority of chromosome rearrangements formed by NHEJ and MMEJ contain 0–2 bp and 2–6 bp of microhomology, respectively, at their junctions7,8. These and other studies have shown that MMEJ typically results in relatively large DNA deletions (~30–200 bp), whereas deletions due to NHEJ are substantially shorter in length (<10 bp)7,8,13.

Although MMEJ occurs infrequently in mammalian cells, it has been implicated in multiple processes involving chromosome rearrangements. For example, seminal studies have shown that MMEJ promotes class-switch recombination in NHEJ-deficient B cells6. A more recent study has demonstrated that MMEJ promotes VDJ recombination in NHEJ-proficient B cells containing mutations in RAG recombination genes5. Finally, studies have shown that MMEJ promotes error-prone replication repair and telomere fusions1,3,8.

Although the mechanism of MMEJ is unknown, genetic studies in invertebrates have suggested a central role for the atypical A-family DNA Polq13–15. For example, studies in Drosophila have indicated that Polq promotes MMEJ of DSBs induced by a sequence-specific endonuclease, whereas in Caenorhabditis elegans polq-I was shown to be required for MMEJ in response to replication-fork collapse at G quadruplexes13–15. Recent genetic studies in mice have also indicated the involvement of Polθ in MMEJ. For example, Polq has been shown to promote class-switch recombination and to confer resistance to DSB-inducing agents16.

1Fels Institute for Cancer Research and Molecular Biology, Temple University School of Medicine, Philadelphia, Pennsylvania, USA. 2Department of Biochemistry, Temple University School of Medicine, Philadelphia, Pennsylvania, USA. Correspondence should be addressed to R.T.P. (richard.pomerantz@temple.edu).

Received 29 October 2014; accepted 22 December 2014; published online 2 February 2015; doi:10.1038/nsmb.2961
The Polθ gene is highly unusual in that it encodes an N-terminal helicase-like domain, a large central domain and a C-terminal polymerase domain. Although the functions of the helicase and central domains are unknown, the polymerase domain encoded by human POLQ—referred to herein as Polθ—has been characterized as a highly promiscuous enzyme. For example, unlike most polymerases, Polθ promotes extension of ssDNA and partial ssDNA (pssDNA) substrates. Polθ also exhibits low-fidelity DNA synthesis, translesion synthesis and lyase activity, and it is implicated in interstrand-crosslink repair, base excision repair and DNA end-joining. Therefore, Polθ is among the A family of polymerases (i.e., Klenow fragment, Thermus aquaticus (Taq) polymerase and bacteriophage T7 polymerase), which normally perform high-fidelity synthesis and lyase activity, and it is implicated in interstrand-crosslink repair, base excision repair and DNA end-joining.

RESULTS

Human Polθ promotes MMEJ in vitro and in vivo

We examined whether Polθ promotes end-joining of substrates modeled after partially resected DSBs (pssDNA) containing variable microhomology lengths (Fig. 1a,b). We performed MMEJ by incubating Polθ with radiolabeled pssDNA in the presence of dNTPs for 30 min and then terminating the reaction by addition of EDTA and proteinase K, which degrades the polymerase. MMEJ of pssDNA is indicated by a product approximately twice the size of the substrate in a nondenaturing gel (indicated by a black asterisk in figures). Polθ converted pssDNA-6 into a double-size product (Fig. 1c) indicative of MMEJ, Polθ similarly promoted MMEJ of pssDNA-4 (Fig. 1c), but it was less efficient in joining pssDNA-2 (Fig. 1c), and it failed to join pssDNA-0, which lacks microhomology (Fig. 1c). Polθ completed MMEJ of pssDNA-4 in less than 20 min but joined only a fraction of pssDNA-2 substrates after 60 min (Supplementary Fig. 1a).

We next examined the ability of Polθ to perform MMEJ of ssDNA versions of the substrates used in Figure 1c (schematics in Fig. 1b). We found that, remarkably, ssDNA was a poor substrate for Polθ MMEJ, as compared to pssDNA. For example, Polθ failed to join ssDNA containing <4 bp of microhomology and performed very limited MMEJ of ssDNA-6 (Fig. 1d). Thus, the polymerase almost exclusively joins substrates modeled after partially resected DNA when homology is limited (i.e., microhomology ≤ 4 bp). We note that Polθ produced mostly lower-molecular-weight byproducts on substrates that do not support efficient MMEJ, such as ssDNA substrates and pssDNA containing <4 bp of microhomology (Fig. 1c,d). Previous studies have indicated that Polθ is proposed to extend paired 3’ overhangs at DNA synapses. (b) pssDNA and ssDNA substrates. Asterisk, 32P. (c,d) Nondenaturing gels showing MMEJ reactions with pssDNA (c) and ssDNA (d). Asterisk, MMEJ products. (e) Schematic of solid-phase MMEJ assay. SA, streptavidin. Left gel, denaturing gel showing DNA purified from pellet and supernatant fractions after MMEJ in the presence (lanes 2 and 4) or absence (lanes 1 and 3) of Polθ. Right gel, denaturing gel showing MMEJ reactions with pssDNA-4 (lanes 1 and 2) and ssDNA-4 (lanes 3 and 4). Gray asterisks, 32P; black asterisks, MMEJ products. (f) Nondenaturing gel showing MMEJ reactions with the indicated polymerases. Asterisk, MMEJ products. (g) MMEJ GFP reporter assay. Schematic of GFP reporter with microhomology, I-Scel site, stop codon and GFP gene indicated. Plot of percentage GFP-positive cells after transient expression of I-Scel and transfection with Polθ siRNA and scrambled siRNA control (error bars, s.d.; n = 9 independent experiments). Western blots of Polθ and β-actin control after transfection with Polθ siRNA (lane 2) and scrambled siRNA control (lane 1).
Figure 2 Polθ uses the opposing overhang as a template in trans to stabilize the DNA synapse. (a) Model of Polθ overhang extension of pssDNA-4. Microhomology is outlined. (b) Nondenaturing and denaturing gels showing MMEJ reactions in the presence of indicated dNTPs. Lane 3 in right gel represents a 48-nt marker based on model in a. Asterisk, MMEJ products. (c) Model of Polθ strand-displacement synthesis during MMEJ of pssDNA-4 conjugated with Cy3 and black hole quencher (BHQ). Plot of fluorescence intensity after MMEJ in the presence (gray) or absence (black) of dNTPs. RU, relative units. Error bars, s.d. (n = 3 independent experiments).

suggested that Polθ may extend ssDNA and pssDNA lacking microhomology via terminal transferase activity, which could account for these small byproducts. However, we demonstrate that Polθ lacks appreciable terminal transferase activity but instead extends substrates that do not support efficient MMEJ by performing ‘snap-back’ replication (Supplementary Fig. 2).

We devised a solid-phase experiment as a control to confirm MMEJ. Here, Polθ joining of radiolabeled to biotinylated pssDNA-4 substrates is detected by tethering of radiolabeled MMEJ products to streptavidin beads (Fig. 1e). Only the double-size product representative of MMEJ was retained on the beads in a Polθ-dependent manner (Fig. 1e, left gel); this product is due to MMEJ because it was generated exclusively from pssDNA-4 (Fig. 1e, right gel). We note that Polθ performed strand-displacement synthesis during MMEJ (Fig. 2c), and this explains the ~48-nt length of MMEJ products in the denaturing gel. MMEJ products also appeared in the supernatant as radiolabeled substrates joined to one another (Fig. 1e, left gel, lane 4). Smaller byproducts were exclusively in the supernatant; thus, they are not MMEJ products but are due to snap-back replication activity (Fig. 1e, left gel, lane 4, and Supplementary Fig. 2). These data along with restriction and length analysis of end-joining products unequivocally demonstrate that Polθ promotes MMEJ (Supplementary Fig. 1b,c).

We next investigated whether MMEJ is specific to Polθ by performing end-joining with polynucleotides from various families: Y family (Polɑ) and Polκ, X family (Polξ), B family (Polθ) and A family (Klenow fragment) (Fig. 1f). Remarkably, only Polθ promoted the MMEJ product. Klenow fragment produced a smaller product with low efficiency, thus suggesting that this related enzyme might exhibit a limited form of end-joining. Polθ degraded the DNA because of exonuclease activity. Because all the polynucleotides were active on a primer template (Supplementary Fig. 3), these data indicate that MMEJ is specific to Polθ.

We next examined whether MMEJ is dependent on Polθ in human cells. Using a previously characterized GFP MMEJ reporter system stably incorporated into U2OS cells21,24, we demonstrated that downregulation of Polθ expression via short interfering RNA (siRNA) resulted in suppression of MMEJ of an I-SceI–induced DSB, as indicated by a reduction in GFP-expressing cells (Fig. 1g and Supplementary Data Set 1). Together, these data demonstrate that human Polθ promotes MMEJ in vitro and in vivo.

Polθ uses the opposing overhang as a template in trans

Because the melting temperature of the 4-bp microhomology sequence (GGCC, <10 °C) is substantially lower than the reaction temperature (37 °C), the DNA synapse is likely to be stabilized by overhang extension (Fig. 2a). To test this, we performed MMEJ with limiting mixtures of dNTPs (Fig. 2b, left and middle gels). Failure to produce a stable MMEJ product in reactions lacking necessary complementary dNTPs will indicate the dependence of the opposing overhang as a template in trans. Consistently with this model, Polθ failed to promote MMEJ with only one or two dNTPs present in the reaction (comparison of lanes 3–5 in left and middle gels). Instead, multiple products appeared as a smear, thus suggesting repetitive cycles of abortive snap-back replication (Supplementary Fig. 2). Polθ promoted MMEJ when only dGTP was omitted, conditions expected to allow extension by 12 nt (comparison of lanes 3 and 6, left and middle gels). The polymerase, however, extended the overhang in a similar manner to that when all four dNTPs were present (comparison of lanes 3 and 6, middle gel), a result that we attribute to the high frequency of nucleotide misincorporation and mismatch-extension activity exhibited by Polθ20,25. Using a denaturing sequencing gel, we show that the MMEJ product generated in the presence of all four dNTPs is approximately 1 nt longer than the expected 48-nt product (Fig. 2b, right gel). This result is probably due to the ability of Polθ to incorporate dAMP at the end of the template like Klenow fragment, as shown previously16.

We next examined whether Polθ exhibits strand-displacement activity during MMEJ. To test this, we performed the MMEJ reaction, using pssDNA-4 modified with Cy3 and a black hole quencher (BHQ; Fig. 2c). In this scenario, strand displacement is indicated by an increase in Cy3 fluorescence due to dissociation of the short strand conjugated with BHQ. We observed a substantial increase in Cy3 fluorescence only in the presence of dNTPs, thus demonstrating that Polθ performs strand-displacement synthesis during MMEJ. Together, the data show that Polθ facilitates MMEJ by using the opposing overhang as a template in trans to stabilize the DNA synapse.

Template preferences for Polθ MMEJ

We next examined the template preferences for Polθ MMEJ. Cellular studies previously indicated that MMEJ requires <20 nt of resection, which is performed by the resection initiation factors Mre11 and CtIP3.
Strikingly, we find that Polθ MMEJ requires a similar overhang length (<18 nt; Fig. 3a), a result suggesting that the polymerase acts exclusively on DNA after initial resection by Mre11, which was recently shown to be 15–20 nt in yeast. Considering that a 5′-terminal phosphate stimulates replication across gaps by X-family polymerases involved in NHEJ, we tested whether this is also the case for Polθ during MMEJ. Remarkably, a 5′-terminal phosphate increased the rate of MMEJ products generated by Polθ (Fig. 3b), thus indicating that Polθ, like NHEJ polymerases, exhibits an affinity for the 5′-terminal phosphate. Because microhomology annealing is likely to stabilize the DNA synapse, we examined whether MMEJ is more proficient in the presence of GC-rich microhomology. Indeed, we found that MMEJ was facilitated by GC-rich microhomology (Fig. 4a). This demonstrates that MMEJ is promoted by hydrogen bonds between overhangs, which appear to stabilize the end-joining intermediate.

Although previous studies have indicated that MMEJ requires limited resection promoted by Mre11 and CtIP, it remains unknown whether microhomology exposed by this process lies internally or at the 3′ terminus of the resulting overhangs. For example, the experiments presented above used exclusively pssDNA substrates that contain terminal microhomology. However, microhomology may also be positioned internally relative to the 3′ termini of overhangs. We therefore investigated whether Polθ can perform MMEJ of pssDNA containing internal microhomology. Indeed, Polθ promoted efficient MMEJ when the microhomology region was located 5 nt away from the 3′ terminus on one substrate (Fig. 4a). In this scenario, the polymerase probably extends only the overhang containing terminal microhomology (Fig. 4a). Next, we tested whether Polθ is capable of performing MMEJ when microhomology is located internally on both substrates. Indeed, the polymerase promoted efficient MMEJ when microhomology was located 3 nt away from the 3′ terminus on both substrates (Fig. 4b). The ability of Polθ to efficiently extend mismatched primers in previous studies suggests that the polymerase performs mismatch extension of both overhangs in this scenario (Fig. 4b). Polθ also performed MMEJ when microhomology was located 4 nt away from both 3′ termini, albeit with lower efficiency (Supplementary Fig. 4).

Because MMEJ facilitates telomere fusions, we examined whether Polθ joins substrates modeled after telomeres, which also contain internal microhomology (8, 9) (Fig. 4c). Again, we observed high-molecular-weight products indicative of MMEJ (Fig. 4c). The distribution of high- and low-molecular-weight products was nearly identical to that generated from pssDNA-2 (Fig. 1c), thus demonstrating low efficiency of MMEJ due to limited microhomology (2 bp).

As a result of inefficient MMEJ, the major lower-molecular-weight byproduct is due to snap-back replication (Supplementary Fig. 2d). We confirmed that the high-molecular-weight products are due to MMEJ by performing a solid-phase assay with the telomere substrate (Fig. 4d). The slightly longer than expected MMEJ products (i.e., >74 bp) are probably due to the ability of Polθ to promote nucleotide insertions along repetitive sequences because of primer slippage, as shown in previous studies. Together, the data presented in Figure 4 demonstrate the ability of Polθ to perform MMEJ when microhomology is positioned internally.

Polθ promotes DNA synapse formation and strand annealing

We next tested whether Polθ promotes DNA synapse formation separately from its replication activity. Here, we used fluorescence resonance energy transfer (FRET) to probe Polθ-dependent formation of DNA synapses in the absence of dNTPs (Fig. 5a). We found that fluorescence intensity increased as a function of Polθ concentration, whereas we observed no increase when the donor (Cy3) substrate was omitted (Fig. 5a). Hence, these data demonstrate that Polθ promotes DNA synapse formation separately from its replication activity. Next, we repeated the assay, using pssDNA substrates with or without microhomology to determine whether microhomology promotes synapse formation. Remarkably, Polθ promoted DNA synapses in the absence of microhomology, yet the presence of microhomology increased the extent of synapses, as indicated by higher fluorescence (Fig. 5a). These data indicate that the dissociation rate of DNA synapses or the distance between substrates is decreased by overhang base-pairing.

The ability of Polθ to promote DNA synapses in the absence of microhomology suggests that MMEJ includes two initial steps: DNA synapse formation, which brings the overhangs into proximity regardless of microhomology, and microhomology annealing, which promotes base-pairing between substrates that is necessary for overhang extension. We considered whether Polθ contributes to microhomology annealing or whether this process occurs spontaneously once the overhangs are in proximity. For the former case, the polymerase...
would have to exhibit strand-annealing activity. Indeed, Polθ facilitated annealing of ssDNA (Fig. 5b) and pssDNA (Fig. 5c). Together, these data demonstrate that Polθ promotes DNA synapse formation and strand annealing, both of which probably contribute to MMEJ.

**Figure 4** Polθ promotes MMEJ of DNA containing internal microhomology. (a,b) Left, schematic of pssDNA substrates with microhomology outlined; middle, nondenaturing gel showing MMEJ reactions with the indicated pssDNA; and right, model of MMEJ, for pssDNA-A and pssDNA-B (a) and pssDNA-C and pssDNA-D (b). (c) MMEJ of pssDNA modeled after telomeres. Schematic of pssDNA with telomere repeats underlined and microhomology outlined. Nondenaturing gel showing a time course of MMEJ. (d) Schematic of solid-phase MMEJ assay. Denaturing gel of DNA purified from supernatant and pellet fractions after MMEJ reactions in the presence (lane 2 and 4) or absence (lanes 1 and 3) of Polθ. 10% of supernatant was analyzed (right). Throughout figure, gray asterisks, 5′; black asterisks, MMEJ products.

**Insertion loop 2 is essential for MMEJ and Polθ multimers**
Polθ includes three insertion loops that are not present in other A-family polymerases but are highly conserved in vertebrate Polθ (Fig. 6a and Supplementary Note). Interestingly, loop 2 is necessary for Polθ ssDNA annealing and extension, thus suggesting that it may contribute to MMEJ (Supplementary Fig. 5b,c). Structural modeling with homologous Bacillus Pol I

**Figure 5** Polθ promotes DNA synapse formation and strand annealing separately from its replication function. (a) Schematic of DNA synapse assay. Plot of relative fluorescence intensity after Polθ synapse formation in the presence of Cy3 pssDNA with (gray) or without (black) Cy3 pssDNA. Plot of relative fluorescence intensity after Polθ synapse formation in the presence of Cy3 and Cy5 pssDNA with (gray) or without (black) 4 bp of microhomology, RU, relative units. Error bars, s.d. (n = 3 independent experiments). (b) Schematic of annealing assay, nondenaturing gel showing ssDNA (b) and pssDNA (c) annealing in the presence (lane 3) or absence (lane 2) of Polθ, as well as ssDNA (b) and pssDNA (c) marker (lane 1). Plot of percentage annealing (intensity of upper and lower bands). Error bars, s.d. (n = 3 independent experiments).
DISCUSSION

Although some of the genetic requirements for MMEJ have been determined, the central mechanism of this pathway has remained undefined, especially in mammalian cells. Recent genetic studies in mice, however, have confirmed a role for Polq in mammalian MMEJ, demonstrating that the polymerase promotes class-switch recombination and confers resistance to DSB-inducing agents. Considering that a consistent role for Polθ in MMEJ has been shown in mice, flies and worms, the polymerase appears to have evolved to perform an important alternative end-joining function in higher eukaryotes.

Here, we demonstrate that purified human Polθ is highly efficient in MMEJ of DNA containing 3' overhangs with microhomology, and we show that MMEJ is dependent on Polθ expression in human cells. Our data indicate that MMEJ activity is specific to Polθ, because multiple other polymerases from the X, Y, A and B families fail to perform this function in vitro. This suggests that Polθ possesses a unique structural configuration that has been selected to perform MMEJ. Indeed, Polθ includes three insertion-loop domains that are highly conserved among vertebrate Polθ enzymes but are not present in other A-family polymerases (Supplementary Note).

We found that loop 2 is essential for MMEJ in vitro, thus suggesting that this domain may interact with and potentially coordinate the positioning of 3' overhangs during DNA synapse formation and microhomology annealing (Fig. 7). Consistently with this, structural modeling predicts that loop 2 lies in proximity to the 3' terminus of the primer (Supplementary Fig. 6). Furthermore, the presence in this domain of multiple positively charged residues, such as lysine and arginine, suggests that it interacts with DNA (Supplementary Note). Indeed, we found that loop 2 promotes DNA binding, especially in the case of pssDNA; this demonstrates that it confers a structural configuration that favors pssDNA binding.

We found that loop 2 also promotes high-order structures of the polymerase. For example, wild-type Polθ formed dimers, and to a lesser extent multimers, as indicated by gel filtration.
Figure 7 Models of Polθ MMEJ. (a–c) After limited resection of a DSB by Mre11 and CtIP, Polθ dimers promoted by loop 2 facilitate DNA synapse formation (top). Polθ promotes annealing of terminal (a) or internal (b, c) microhomology after DNA synapse formation. Polθ extends the annealed overhang by using the opposing overhang as a template in trans to stabilize the DNA synapse. Overhang extension is facilitated by Polθ binding to the 5′ terminal phosphate on the opposing DNA, and it results in strand displacement. (a) Polθ extends both overhangs in the case of terminal microhomology. (b) Polθ extends only the terminally paired overhang when internal microhomology is located relatively far from the 3′ terminus on the opposing strand. (c) Polθ performs mismatch extension of overhangs that contain internal microhomology relatively close to their 3′ terminus. Finally, 5′ flap repair is presumably required before ligation in each case.

(Supplementary Fig. 5d). In contrast, Polθ lacking loop 2 acted solely as a monomer (Supplementary Fig. 5d) and, as a potential result of this, failed to perform MMEJ (Fig. 6c). We therefore propose a model whereby Polθ dimers facilitate DNA synapse formation (Fig. 7, top). Consistently with this model, previous structural studies have demonstrated dimerization of an end-joining polymerase as a mechanism of DNA synapse formation. Considering that Polθ facilitated annealing of complementary overhangs separately from its replication activity, we propose that the polymerase promotes microhomology annealing after bringing the overhangs into proximity (Fig. 7). Because loop 2 was necessary for strand annealing (Supplementary Fig. 5c), it is probably involved in forming minimally paired overhangs. This function would explain why loop 2 is required for Polθ ssDNA extension, because such activity requires a minimally paired primer (Supplementary Fig. 2).

Although it is unclear how Mre11 and CtIP perform the limited resection step required for MMEJ, microhomology exposed by this activity may be positioned internally or at the 3′ terminus of overhangs (Fig. 7). We found that, remarkably, Polθ promotes MMEJ of pssDNA containing terminal or internal microhomology. In the case of terminal microhomology, our data support a mechanism whereby the polymerase extends each overhang by using the opposing overhang as a template in trans to stabilize the DNA synapse (Fig. 7a). Because the rate of overhang extension was increased in the presence of a 5′-terminal phosphate, Polθ probably possesses a 5′-terminal phosphate–interacting motif, as do X-family polymerases involved in NHEJ and base excision repair (Fig. 7). In the case of internal microhomology, we present two models based on our findings. In the first example, in which one end contains internal microhomology relatively far (>3 nt) from the 3′ terminus, Polθ is limited to extending the paired end containing terminal microhomology (Fig. 7b). In the second example, in which both ends contain internal microhomology relatively close (1–3 nt) to the 3′ termini, the polymerase performs mismatch extension of each end, potentially generating mutations in addition to a DNA deletion (Fig. 7c). Finally, because Polθ promotes a certain degree of strand-displacement synthesis, 5′ flap repair is likely to be required before ligation.

In summary, our data reveal a central mechanism of MMEJ promoted by Polθ. Given that POLQ also encodes an N-terminal helicase domain and a large central domain, it will be interesting to determine whether these components contribute to the mechanism or regulation of MMEJ. Finally, because upregulation of Polθ corresponds to a poor clinical outcome for patients with breast cancer, it will be important to determine whether MMEJ contributes to cancer progression and chemotherapy resistance in these tumors.

METHODS

Methods and any associated references are available in the online version of the paper.

Note: Any Supplementary Information and Source Data files are available in the online version of the paper.

ACKNOWLEDGMENTS

Research was funded by US National Institutes of Health grant (4R00CA160648-03) awarded to R.T.P. and Temple University School of Medicine start-up funds to R.T.P. We thank S. Wallace (University of Vermont) for wild-type Polθ and Polθ L2 expression vectors, I. Stark (Beckman Research Institute, City of Hope) for the U2OS cell line EJ2-GFP and A.K. Aggarwal (Mount Sinai Hospital) for Polk.

AUTHOR CONTRIBUTIONS

R.T.P., T.K., G.C., S.M.M. and A.Y.O. performed the experiments. R.T.P. designed the experiments and wrote the paper.
1. McVey, M. & Lee, S.E. MMEJ repair of double-strand breaks (director’s cut): deleted sequences and alternative endings. Trends Genet. 24, 529–538 (2008).

2. Bennardo, N., Cheng, A., Huang, N. & Stark, J.M. Alternative-NHEJ is a mechanistically distinct pathway of mammalian chromosome break repair. PLoS Genet. 4, e1000110 (2008).

3. Truong, L.N. et al. Microhomology-mediated end joining and homologous recombination share the initial end resection step to repair DNA double-strand breaks in mammalian cells. Proc. Natl. Acad. Sci. USA 110, 7720–7725 (2013).

4. Deriano, L. & Roth, D.B. Modernizing the nonhomologous end-joining repertoire: alternative and classical NHEJ share the stage. Annu. Rev. Genet. 47, 433–455 (2013).

5. Corneo, B. et al. Rag mutations reveal robust alternative end joining. Nature 449, 483–486 (2007).

6. Yan, C.T. et al. IgH class switching and transllocations use a robust non-classical end-joining pathway. Nature 449, 478–482 (2007).

7. Ghezraoui, H. et al. Chromosomal translocations in human cells are generated by canonical nonhomologous end-joining. Mol. Cell 55, 829–842 (2014).

8. Jones, R.E. et al. Escape from telomere-driven crisis is DNA ligase III dependent. Cell Reports 8, 1063–1076 (2014).

9. Zhang, Y. & Jasin, M. An essential role for CtIP in chromosomal translocation formation through an alternative end-joining pathway. Nat. Struct. Mol. Biol. 18, 80–84 (2011).

10. Simsek, D. et al. DNA ligase III promotes alternative nonhomologous end-joining during chromosomal translocation formation. PLoS Genet. 7, e1002080 (2011).

11. Audebert, M., Salles, B. & Calsou, P. Involvement of poly(ADP-ribose) polymerase-1 and XRCC1/DNA ligase III in an alternative route for DNA double-strand breaks rejoining. J. Biol. Chem. 279, 55117–55126 (2004).

12. Lee-Theilen, M., Matthews, A.J., Kelly, D., Zheng, S. & Chaudhuri, J. CtIP promotes microhomology-mediated alternative end joining during class-switch recombination. Nat. Struct. Mol. Biol. 18, 75–79 (2011).

13. Koole, W. et al. A Polymerase Theta-dependent repair pathway suppresses extensive genomic instability at endogenous G4 DNA sites. Nat. Commun. 5, 3216 (2014).

14. Chan, S.H., Yu, A.M. & McVey, M. Dual roles for DNA polymerase theta in alternative end-repairing of double-strand breaks in Drosophila. PLoS Genet. 6, e1001005 (2010).

15. Roenink, S.F., van Schendel, R. & Tijsterman, M. Polymerase theta-mediated end joining of replication-associated DNA breaks in C. elegans. Genome Res. 24, 954–962 (2014).

16. Yousefzadeh, M.J. et al. Mechanism of suppression of chromosomal instability by DNA polymerase POLQ. PLoS Genet. 10, e1004654 (2014).

17. Seki, M., Marini, F. & Wood, R.D. POLQ (Pol theta), a DNA polymerase and DNA-dependent ATPase in human cells. Nucleic Acids Res. 31, 6117–6126 (2003).

18. Hogg, M., Seki, M., Wood, R.D., Doublet, S. & Wallace, S.S. Lesion bypass activity of DNA polymerase theta (POLQ) is an intrinsic property of the pol domain and depends on unique sequence insertions. J. Mol. Biol. 405, 642–652 (2011).

19. Hogg, M., Sauer-Eriksson, A.E. & Johansson, E. Premiscuous DNA synthesis by human DNA polymerase theta. Nucleic Acids Res. 40, 2611–2622 (2012).

20. Arana, M.E., Seki, M., Wood, R.D., Rogozin, I.B. & Kunkel, T.A. Low-fidelity DNA synthesis by human DNA polymerase theta. Nucleic Acids Res. 36, 3847–3856 (2008).

21. Yoon, J.H., Roy Choudhury, J., Park, J., Prakash, S. & Prakash, L. A role for DNA polymerase theta in promoting replication through oxidative DNA lesion, thymine glycol, in human cells. J. Biol. Chem. 289, 13177–13185 (2014).

22. Prasad, R. et al. Human DNA polymerase theta possesses 5’-dRP lyase activity and functions in single-nucleotide base excision repair in vitro. Nucleic Acids Res. 37, 1868–1877 (2009).

23. Gunn, A., Bennardo, N., Cheng, A. & Stark, J.M. Correct end use during end joining of multiple chromosomal double strand breaks is influenced by repair protein RAD50, DNA-dependent protein kinase DNA-PKcs, and transcription context. J. Biol. Chem. 286, 42470–42482 (2011).

24. Seki, M. & Wood, R.D. DNA polymerase theta (POLQ) can extend from mismatches and from bases opposite a (6–4) photoproduct. DNA Repair (Amst.) 7, 119–127 (2008).

25. Garcia-Diaz, M. et al. DNA polymerase theta (POLQ) is an intrinsic property of the pol domain and from bases opposite a (6–4) photoproduct. Nucleic Acids Res. 37, 28215–28226 (2012).

26. Schwede, T., Kopp, J., Guex, N. & Peitsch, M.C. SWISS-MODEL: an automated protein homology-modeling server. Nucleic Acids Res. 33, 389–391 (2005).

27. Gonzalez-Melero, J.M., Yebra, A., Guex, N. & Schwede, T. SWISS-MODEL: an automated protein homology-modeling server. Nucleic Acids Res. 39, 397–402 (2011).

28. Moon, A.F. et al. The X family portrait: structural insights into biological functions of X family polymerases. DNA Repair (Amst.) 6, 1709–1725 (2007).

29. Brissett, N.C. et al. Molecular basis for DNA double-strand break annealing and primer extension by an NHEJ DNA polymerase. Cell Reports 5, 1108–1120 (2013).

30. Schevitz, R.W. et al. An alternate DNA polymerase Theta-dependent alternative end-joining pathway. EMBO J. 33, 993–1004 (2004).

31. Wang, W., Wu, E.Y., Hellinga, H.W. & Beese, L.S. Structural factors that determine DNA polymerase theta selectivity for deoxy-, dideoxy- and ribonucleotides. J. Biol. Chem. 287, 28215–28226 (2012).

32. Brissett, N.C. et al. Structure of a NHEJ polymerase-mediated DNA synaptic complex. Science 318, 456–459 (2007).

33. Lemée, F. et al. DNA polymerase theta up-regulation is associated with poor survival in breast cancer, perturbs DNA replication, and promotes genetic instability. Proc. Natl. Acad. Sci. USA 107, 13390–13395 (2010).
ONLINE METHODS

MMEJ. 100 nM 5'-32P–radiolabeled psDNA or ssDNA preincubated at 37 °C in buffer A (25 mM Tris-HCl, pH 8.8, 10% glycerol, 1 mM DTT, 0.01% NP-40, 10 mM MgCl2, and 0.1 mg/ml BSA) was mixed with 100 nM Polθ for 15 min; this was followed by the addition of 500 µM dNTPs for another 30 min at 37 °C in a total volume of 20 µL. For analysis in non-denaturing gels, reactions were terminated by the addition of 4 µl of non-denaturing stop buffer (0.5 M Tris-HCl, pH 7.5, 10 mg/ml proteinase K, 80 mM EDTA, and 1.5% SDS); this was followed by a further 15–30 min incubation at 37 °C. Analysis in denaturing gels, reactions were terminated by the addition of 20 µl of 2× denaturing stop buffer (90% formamide and 50 mM EDTA). Radiolabeled DNA was resolved in non-denaturing or denaturing (urea) polyacrylamide gels and visualized by autoradiography. Concentrations are listed as final. Experiments in Figure 2c were performed as above; however, psDNA-4 was conjugated with Cy3 and BHQ, and strand displacement was determined by Cy3 fluorescence intensity with a Clariostar (BMG Labtech) plate reader.

MMEJ solid-phase assay. MMEJ was performed with or without Polθ as described above; however, for reaction volumes were pooled, equimolar concentrations (50 nM) of radiolabeled and biotinylated psDNA substrates were used, and the pooled reactions (80 µl) were incubated with 160 µl of magnetic streptavidin beads (Thermo Scientific Pierce) for 30 min rather than stop buffer being added. The supernatant containing soluble DNA was removed with a magnetic separation rack (NEB), and then the beads were washed three times with 200 µl of 200 mM NaCl. Finally, the pellet fraction containing biotinylated DNA was removed from the beads by the addition of 30 µl of 1× denaturing stop buffer (45 mM formamide and 25 mM EDTA); this was followed by boiling and removal of supernatant with a magnetic separation rack. DNA from the supernatant fraction was ethanol precipitated, then resuspended in 30 µl of 10 mM Tris-HCl, pH 8.0, and 1 mM EDTA). Pellet and 50% of supernatant fractions were analyzed in a denaturing urea polyacrylamide gel. For experiments in Figure 4d, ten reactions were pooled, the pellet was washed with 50 mM NaCl, and 10% of the supernatant was analyzed.

1-Sce-I–induced MMEJ. U2OS-EJ2-GFP cells with MMEJ reporter24 (1 × 105) were plated on six-well plates and transfected 24 h later with 2.5 µg pCMV-3×NLS–I-SceI (or control vector, GFP expression vector) per well. For siRNA experiments, cells were transfected with 100 pmol siRNA (Polθ vector (pCMV-3×NLS-GFP). For siRNA experiments, cells were transfected with

Western blotting. A portion of U2OS-Alt-EJ cells24 from the I-Sce-I–induced MMEJ solid-phase assay. MMEJ was performed with or without Polθ as described above; however, for reaction volumes were pooled, equimolar concentrations (50 nM) of radiolabeled and biotinylated psDNA substrates were used, and the pooled reactions (80 µl) were incubated with 160 µl of magnetic streptavidin beads (Thermo Scientific Pierce) for 30 min rather than stop buffer being added. The supernatant containing soluble DNA was removed with a magnetic separation rack (NEB), and then the beads were washed three times with 200 µl of 200 mM NaCl. Finally, the pellet fraction containing biotinylated DNA was removed from the beads by the addition of 30 µl of 1× denaturing stop buffer (45 mM formamide and 25 mM EDTA); this was followed by boiling and removal of supernatant with a magnetic separation rack. DNA from the supernatant fraction was ethanol precipitated, then resuspended in 30 µl of 10 mM Tris-HCl, pH 8.0, and 1 mM EDTA). Pellet and 50% of supernatant fractions were analyzed in a denaturing urea polyacrylamide gel. For experiments in Figure 4d, ten reactions were pooled, the pellet was washed with 50 mM NaCl, and 10% of the supernatant was analyzed.

Strand annealing. 5'-32P–radiolabeled ssDNA (RP40, 20 nM) or psDNA (RP40/RP364, 1 nM) substrates were incubated with or without 100 nM Polθ WT or Polθ L2 in 20 µl of buffer C (50 mM Tris-HCl, pH 7.5, 1 mM DTT, 0.1 mg/ml BSA, 10% glycerol, and 0.5 mM Mg acetate) at 37 °C for 5 min. Complementary ssDNA (RP40C, 20 nM) or psDNA (RP365/RP364, 1 nM) was then added for 30 s. Annealing was terminated by the addition of excess cold ssDNA (RP40, 200 nM) or psDNA (RP40/RP364, 30 nM) and then 4 µl of nondenaturing stop buffer (0.5 M Tris-HCl, pH 7.5, 10 mg/ml proteinase K, 80 mM EDTA, and 1.5% SDS); this was followed by a further 15–30 min incubation at 37 °C. DNA was resolved in non-denaturing 12% polyacrylamide gels and analyzed by phosphorimager (Fujiﬁlm FLA 7000).

Structural modeling. Modeling was performed with the Swiss Model server (http://swissmodel.expasy.org/)23 with Polθ residues 1890–2590. Bacillus Pol θ in complex with primer template in closed conformation (PDB 4DQQ) was used as a template for modeling23. SWISS-MODEL was used in first approach mode with default parameters. Structures were visualized with Swiss-PDBViewer24. Images were generated with PyMOL software27. For superposition of the Polθ model and the Bacillus Pol θ–DNA structure, the Ca2+–bound form of Pol θ was used as a template onto which residues 1944–2590 of the Polθ model were superimposed. With the least-squares ﬁtting option in SPDBV20, all carbon atoms superimpose with an r.m.s. of 0.22 Å.

Primer extension. The indicated Polθ (100 nM) was incubated with 5'-32P–radiolabeled primer template (RP25/RP266, 100 nM) along with 0.5 µM dNTPs in 10 µl of buffer A for 30 min at 37 °C. Reactions were terminated by the addition of 10 µl of denaturing stop buffer, and DNA was resolved in denaturing (urea) polyacrylamide gels and analyzed by autoradiography.

EMSA. 12 nM Cy3-conjugated primer template (RP25Cy3/RP266), psDNA (RP348Cy3/RP343P) or ssDNA (RP348Cy3) was mixed with the indicated amounts of Polθ WT or Polθ L2 in 20 µl of buffer (25 mM Tris-HCl, pH 7.5, 30 mM NaCl, 0.01% NP-40, 1 mM DTT, 0.1 µg/ml BSA, 0.5 mM MgCl2, and 10% glycerol) on ice (primer template) or at room temperature (psDNA) for 60 min. Reactions were resolved in non-denaturing 5% polyacrylamide gels containing 2.5% glycerol, and DNA was visualized with a Multimage III ﬂuorescence imager (Alpha Innotech) and FluorChem Q software by monitoring Cy3 fluorescence.

Gel filtration. Polθ in protein storage buffer (50 mM HEPES, pH 8.8, 10% glycerol, 1 mM NaCl, 1 mM DTT, 0.02% NP-40, and 10% (v/v) glycerol) was injected into a HiLoad 16/600 Superdex 200 pg column (GE Health Sciences) equilibrated with protein storage buffer with AKTA L1 Pure (GE Health Sciences) with an automated fraction collector and a multichannel peristaltic pump. Protein peaks were plotted versus elution volume with a UV monitor detecting at 280 nm.

Native gel analysis. 1 µM Polθ WT or Polθ L2 was incubated in 25 mM Tris-HCl, pH 8.8, 0.01% NP-40, 1 mM DTT, 10% glycerol, 676 mM NaCl, and 0.2% Tween-40 at room temperature for 1 h. Protein solutions were then resolved in a 4–15% Mini-PROTEAN TGX native gel (Bio-Rad) in Tris-glycine buffer, pH 8.9. Protein was then visualized by silver staining.

Sequence alignment. The indicated amino acid sequences of the polymerase domain of Polθ from the indicated vertebrates and other indicated A-family Pols were aligned with Clustal Omega (http://www.ebi.ac.uk/Tools/msa/clustalo/; European Bioinformatics Institute) default settings.

Cell lines and cell culture. The U2OS cell line (EJ-2-GFP)24 was a kind gift from J. Stark. U2OS cells were grown in Dulbecco’s Modified Eagle’s Medium (DMEM) with 10% fetal bovine serum (FBS).

Proteins. Polθ WT and Polθ L2 were purified as previously described18. Klenow fragment was purchased from New England BioLabs. Polθ was purified as described18. Polθ was provided by A.K. Aggarwal. Polθ was purchased from Enzymax. Polθ was purified as described19.

DNA. Templates are as follows. psDNA-6 (RP344/RP343), psDNA-4 (RP348/RP343), ssDNA-2 (RP346/RP343), psDNA-0 (RP347/RP343), ssDNA-6

NATURE STRUCTURAL & MOLECULAR BIOLOGY
doi:10.1038/nsmb.2961
(RP344), ssDNA-4 (RP348), ssDNA-2 (RP340), ssDNA-0 (RP347). **Figure 1e**, RP348/RP343, RP348B/RP343. **Figure 2a**, ssDNA-4 (RP348). **Figure 2b**, ssDNA-2 (RP340). **Figure 2c**, ssDNA-0 (RP347). **Figure 3a**, RP366/RP343, RP348/RP343, RP362/RP343, RP363/RP343. **Figure 3c**, RP340/RP343, RP357/RP343. **Figure 4a**, pssDNA-A (RP399/RP343); pssDNA-B (RP348/RP343). **Figure 4b**, pssDNA-C (RP396/RP343); pssDNA-D (RP397/RP343). **Figure 4c**, RP360, RP316 (b); RP348 (c). **Figure 5a**, pssDNA-4 (RP348/RP343), pssDNA-2 (RP340/RP343) (a); RP370/RP343 (b); pssDNA-4A (RP348/RP343), pssDNA-4B (RP356/RP353) (c). **Supplementary Figure 1**, pssDNA-4 (RP348/RP343), pssDNA-2 (RP340/RP343) (a); RP370/RP343 (b); pssDNA-4A (RP348/RP343), pssDNA-4B (RP356/RP353) (c). **Supplementary Figure 2**, RP360, RP316 (b); RP348 (c). **Supplementary Figure 3**, primer template (RP25/RP266). **Supplementary Figure 4**, pssDNA-E, RP396/RP343; pssDNA-F, RP402/RP343. **Supplementary Figure 5**, RP316Cy3 (a); ssDNA-4 (RP348) (b); RP40, RP40C (c).

All pssDNA templates were 5′-phosphorylated on the shorter strand with T4 polynucleotide kinase (NEB) and ATP or purchased with a 5′-phosphate, except for RP331/RP332 and where indicated in the text and figures. pssDNA substrates were annealed by mixture of a ratio of 1:1.5 of long to short strands then boiling and slow cooling to room temp. DNA was 32P-5′-radiolabeled with T4 polynucleotide kinase (NEB) and [\(\gamma\)-32P]ATP (PerkinElmer).

DNA oligonucleotides (Integrated DNA Technologies) are as follows (5′-3′).

RP348, CACTGTGAGCTTAGGGTTAGAGCCGG; RP348B, biotin-CACTGTGAGCTTAGGGTTAGAGCCGG; RP343P, P-CTAAGCTCACAGTG; RP344, CACTGTGAGCTTAGGGTTAGCCCGGG; RP343, CACTGTGAGCTTAGGGTTAGCCCGGG; RP368, CACTGTGAGCTTAGAGCCGG; RP366, CACTGTGAGCTTAGAGCCGG; RP362Cy3, Cy3-CACTGTGAGCTTAGGGTTAGAGCCGG; RP363, CACTGTGAGCTTAGATTCTAGGTTAGAGCCGG; RP343P-BkFQ, P-CTAAGCTCACAGTG-BkFQ; RP357, CACTGTGAGCTTAGGGTTAGAGATTACAT; RP40, CACTGTGAGCTTAGGGTTAGAGGAGGAGGGATGAGAATATT; RP40C, AATATTCTCATCCCTCCCTCCTCCCTCATTTAG; RP364, CACTGTGAGCTTAGATTCTAGGTTAGAGCCGG; RP331, ACTGTGAGCTTAGGGTTAGGGTTAGGGTTAGGGTTAG; RP331B, Biotin-ACTGTGAGCTTAGGGTTAGGGTTAGGGTTAGGGTTAG; RP332, CTAACCCTAACCCTAAGCTCACAGTG; RP396, CACTGTGAGCTTAGGGTTAGGGTTAGGGTTAG; RP397, CACTGTGAGCTTAGGGTTAGGGTTAGGGTTAG; RP316Cy3, Cy3-TTTTTTTTTTTTTTTTTTTTTTTTTTTTT; RP402, CACTGTGAGCCTTAGTTTACCCTCCTATT.

34. Schwede, T., Kopp, J., Guex, N. & Peitsch, M.C. SWISS-MODEL: an automated protein homology-modeling server. *Nucleic Acids Res.* **31**, 3381–3385 (2003).
35. Wang, W., Wu, E.Y., Hellinga, H.W. & Beese, L.S. Structural factors that determine selectivity of a high fidelity DNA polymerase for deoxy-, dideoxy-, and ribonucleotides. *J. Biol. Chem.* **287**, 28215–28226 (2012).
36. Guex, N. & Peitsch, M.C. SWISS-MODEL and the Swiss-PdbViewer: an environment for comparative protein modeling. *Electrophoresis* **18**, 2714–2723 (1997).
37. DeLano, W.L. Unraveling hot spots in binding interfaces: progress and challenges. *Curr. Opin. Struct. Biol.* **12**, 14–20 (2002).
38. Fazlieva, R. et al. Proofreading exonuclease activity of human DNA polymerase delta and its effects on lesion-bypass DNA synthesis. *Nucleic Acids Res.* **37**, 2854–2866 (2009).
39. Acharya, N. et al. Complex formation of yeast Rev1 with DNA polymerase eta. *Mol. Cell. Biol.* **27**, 8401–8408 (2007).