HHS Public Access

Author manuscript

DNA Repair (Amst). Author manuscript; available in PMC 2018 October 01.

Published in final edited form as:

DNA Repair (Amst). 2017 October ; 58: 47–51. doi:10.1016/j.dnarep.2017.08.006.

DNA polymerase ι: The long and the short of it!

Ekaterina G. Frank, **Mary P. McLenigan**, **John P. McDonald**, **Donald Huston**, **Samantha Mead**, and **Roger Woodgate***

Laboratory of Genomic Integrity, National Institute of Child Health and Human Development, National Institutes of Health, Bethesda, MD 20892-3371, USA

Abstract

The cDNA encoding human DNA polymerase ι (*POLI*) was cloned in 1999. At that time, it was believed that the *POLI* gene encoded a protein of 715 amino acids. Advances in DNA sequencing technologies led to the realization that there is an upstream, in-frame initiation codon that would encode a DNA polymerase ι (polι) protein of 740 amino acids. The extra 25 amino acid region is rich in acidic residues (11/25) and is reasonably conserved in eukaryotes ranging from fish to humans. As a consequence, the curated Reference Sequence (RefSeq) database identified polι as a 740 amino acid protein. However, the existence of the 740 amino acid polι has never been shown experimentally. Using highly specific antibodies to the 25 N-terminal amino acids of polι, we were unable to detect the longer 740 amino acid (ι-long) isoform in western blots. However, trace amounts of the ι-long isoform were detected after enrichment by immunoprecipitation. One might argue that the longer isoform may have a distinct biological function, if it exhibits significant differences in its enzymatic properties from the shorter, well-characterized 715 amino acid polι. We therefore purified and characterized recombinant full-length (740 amino acid) polι-long and compared it to full-length (715 amino acid) polι-short *in vitro*. The metal ion requirements for optimal catalytic activity differ slightly between ι-long and ι-short, but under optimal conditions, both isoforms exhibit indistinguishable enzymatic properties *in vitro*. We also report that like ι-short, the ι-long isoform can be monoubiquitinated and polyubiquitinated *in vivo*, as well as form damage induced foci *in vivo*. We conclude that the predominant isoform of DNA polι in human cells is the shorter 715 amino acid protein and that if, or when, expressed, the longer 740 amino acid isoform has identical properties to the considerably more abundant shorter isoform.

Keywords

DNA polymerase iota; Y-family DNA polymerase; Translesion DNA synthesis; Translation

*Corresponding author at: 9800 Medical Center Drive, Building C, Rm 320, Bethesda, MD 20892-3371, USA. woodgate@nih.gov (R. Woodgate).

Conflict of interest

The authors declare that they have no conflict of interest.

Appendix A. Supplementary data

Supplementary data associated with this article can be found, in the online version, at http://dx.doi.org/10.1016/j.dnarep.2017.08.006.

1. Introduction

The gene encoding human DNA polymerase ι (*POLI*) was originally annotated as encoding a 715 amino acid protein [1] (See GenBank Accession number: AF140501.1, along with independent submissions AF245438.1 and BC032662.1). However, with advances in DNA sequencing techniques and the abundance of whole genome analyses, it was realized that there is an in-frame upstream initiation codon in the genomic DNA that is missing in the originally identified cDNA sequence. Initiation from the upstream codon would produce a polι protein of 740 amino acids instead of 715 amino acids. As a consequence, the curated Reference Sequence (RefSeq) for human *POLI* (NCBI Reference Sequence: NM_007195.2) was modified to reflect that the *POLI* gene encodes as a 740 amino acid protein (NCBI Reference Sequence: NP_009126.2). However, as far as we are aware, there has never been any experimental evidence reported in the literature to indicate that the longer 740 amino acid polι actually exists in a living cell. We address that issue here by cloning and expressing both isoforms in human cells and comparing the size of the recombinant protein to that of the native, chromosomally expressed, polι from three different human cell lines. By using antibodies specific to both isoforms, we have been able to determine the relative abundance of each isoform expressed endogenously from the human genome. We have also purified and characterized the longer isoform and compared it to the well-characterized shorter, 715 amino acid isoform *in vitro*. Finally, we have also determined whether the longer 740 amino acid isoform can be ubiquitinated in response to DNA damage, as well as form damage-induced replication foci *in vivo*.

2. Materials and methods

A detailed description of all Materials and Methods used in this manuscript can be found in supplementary material, S1_materials and methods.

3. Results

3.1. Phylogenetic analysis of polι from mammals

When the cDNA of human *POLI* was first subcloned, it was believed to encode a polι protein of 715 amino acids (polι-short) [1] (See GenBank Accession number: AF140501.1). This assumption was primarily based upon the fact that immediately upstream of the presumed methionine initiation codon are appropriately located sequences (GCCTGGGCCatgG) with a good match to the consensus “Kozak” ribosome binding site found in eukaryotes [2]. However, with the advent of better whole genome sequencing techniques, another potential in-frame Methionine initiation codon was identified that is 75 base-pairs upstream of the original initiation codon. It too, has appropriately spaced sequences (AGCGGCGGGatgG) with a satisfactory match to the Kozak ribosome-binding site, although the ideal cytosines at positions −2 and −1 are missing. This polι-long protein is predicted to encode a protein of 740 amino acids. The idea that the upstream initiation methionine codon may reflect the “full-length” polι protein is supported by the fact that similar upstream sequences are found in many eukaryotes ranging from fish to humans. An alignment of the N-terminus of the longer isoform of polι found in 24 mammals reveals that the region is reasonably well conserved (Supplementary Fig. 1A). An unrooted phylogenetic

tree (Supplementary Fig. 1B) of the proteins also indicates evolutionary conservation of these sequences. Of particular note, is the fact that the extra N-terminal amino-acids of the longer mammalian isoform are rich in acidic residues. Indeed, 11 of the extra N-terminal 25 amino acids (44%) in the longer human polι isoform are either Asp or Glu (Supplemental Fig. 1B) and these acidic residues could be potentially important for its regulation and/or function *in vivo*.

3.2. Detection of recombinant polι and comparison to endogenous polι

While the presence of ι-long has been annotated for some time in the RefSeq database, as far as we know, there has been no experimental evidence reported in the literature to support its existence. To address this question, we raised polyclonal antibodies to a KHL-conjugated peptide corresponding to the N-terminal 25 amino acids found in ι-long, but which are absent in ι-short. As seen in Fig. 1A, these antibodies are highly specific and only recognize recombinant ι-long overexpressed in the MRC5 fibroblast cell line and do not recognize the similarly overexpressed ι-short in the same cell line. In contrast, polyclonal antibodies raised to a KHL-conjugated peptide corresponding to the C-terminal 15 amino acids found in both ι-short and ι-long recognize both ι-long and ι-short proteins (Fig. 1B). Interestingly, when whole cell extracts of MRC5 cells that had not been transfected with either overproducing plasmid (i.e., expressing only endogenous polι), were probed with the same antibodies, the C-terminal specific antibodies were able to recognize a protein with the same electrophoretic mobility as ι-short, and no detectable signal was observed with the N-terminal antibodies. We conclude that in the commonly used MRC5 cell line, the major isoform of polι is, therefore, ι-short.

To determine if this conclusion might be extended to other human cell lines, we made extracts from three different *POLI*$^{+/+}$ or$^{-/-}$ cell lines (MRC5, HEK293T and BL2 derivatives) (Fig. 1C). Extracts were probed with the polyclonal C-terminal polι antibodies that recognize both ι-long and ι-short isoforms. As expected, no signal was observed in any of the *POLI*$^{-/-}$ cell lines. Although expressed at lower levels than in MRC5 cell line, the endogenous polι detected in whole cell extracts from the HEK293T and BL2 cell lines exhibited an electrophoretic mobility consistent with the size of ι-short, rather than ι-long.

To increase the sensitivity of our detection methods, we immunoprecipitated endogenous polι from the MRC5 extract with the monoclonal antibody (Abnova) raised to the C-terminal 100 amino acids found in both ι-short and ι-long, or with polyclonal antibodies raised to the twenty-five N-terminal amino acids found exclusively in ι-long. The eluate from the immunoprecipitate was then probed with polyclonal antibodies raised to the C-terminal fifteen amino acids of polι. As seen in Fig. 1D, the major isoform detected after immunoprecipitation is ι-short. However, after an extended exposure of the blot, a faint band corresponding to ι-long was also detected (Fig. 1D).

3.3. Comparison of the in vitro properties of ι-long versus ι-short

The western blots indicate that the major isoform of polι in three different human cell lines is the originally identified and well-characterized 715 amino acid polι-short protein. However, it is possible that even the trace amounts of ι-long that we detected after

immunoprecipitation might be of biological significance, especially if the protein has very different biochemical properties from ι-short. Indeed, Kim et al., compared the biochemical properties of the catalytic domains of ι-long and ι-short *in vitro* in the presence of manganese (0.15 mM), or magnesium (5 mM) and found ι-short to be 10-fold more catalytically active than ι-long on an undamaged template in the presence of magnesium (but not in the presence of manganese) and ι-short bound to DNA 20–29-fold better than ι-long [3]. However, we believe that such observations may possibly be influenced by the *in vitro* characterization of the catalytic domain of the polι protein compared to the full-length protein. For example, while the pI of ι-long increases from 6.78 in the full-length protein to 7.56 in the truncated protein, the effects are even greater with ι-short, where the pI of the full-length protein increases from 8.11 to 9.06 in the truncated version and this may facilitate better binding to DNA and concomitantly stimulate the catalytic activity of ι-short compared to ι-long [3].

We therefore characterized the *in vitro* properties of highly purified full-length ι-long to that of full-length ι-short. We have previously shown that the catalytic activity of polι-short is significantly influenced by the presence of low levels of manganese or magnesium ions [4]. We therefore first determined the optimal metal ion concentration for ι-long in the presence of manganese or magnesium (Fig. 2A). Our previous studies with ι-short revealed that it is most active in very low levels of manganese (50–150 μM), or low levels of magnesium (~250 μM) [4]. Analysis of ι-long revealed a similar, but slightly different metal ion requirement. Full-length polι-long is very active in the presence of 50–500 μM manganese, where it fully extends the primer, as well as exhibits template independent terminal transferase activity. The catalytic activity of the enzyme begins to drop significantly at concentrations of manganese > 1 mM and exhibits extremely poor activity at > 2 mM manganese. Polι-long dependent primer extension reactions in the presence of magnesium also displayed a different profile compared to ι-short. The greatest activity of ι-long was observed at 1 mM magnesium, with activity dropping significantly with either lower (< 200 μM), or higher concentrations (> 2 mM) of magnesium (Fig. 2A).

Given that ι-long and ι-short exhibit slightly different requirements for optimal activity *in vitro*, we decided to compare the activity of each protein under optimal conditions for both enzymes. In the case of ι-short, it was in the presence of 200 μM manganese or magnesium. For ι-long, we used 500 μM manganese and 1 mM magnesium. We first chose to compare their activity on the “T10AGC” template. This template was initially chosen as ι-short has greatest catalytic activity when incorporating T opposite template A [5]. As seen in Fig. 2B and C, both ι-long and ι-short exhibit considerable terminal transferase activity on the template in the presence of manganese and, in the presence of magnesium, both enzymes can extend to the end of the template in the presence of all four dNTPs or just T. Overall, while there are obvious differences in primer extension when comparing manganese to magnesium, there was little difference observed between ι-long and ι-short under optimal conditions *in vitro*.

Next, we compared the ability of ι-long and ι-short to extend a primer annealed to the “TTA” template. This template was chosen because of polι-short’s propensity to misincorporate dG opposite template T [5], and since extension from the mispair is

catalytically inefficient compared to a correct base pair [6], this leads to limited primer extension. These observations were recapitulated with both ι-long and ι-short in the presence of magnesium (Fig. 2B & C). As expected, manganese promoted relaxed fidelity and elongated replication products, but as observed previously with the 10TAGC template, ι-long and ι-short are largely indistinguishable under optimal primer extension conditions *in vitro*.

3.4. Ubiquitination of polι in undamaged and plumbagin-treated cells

One could argue that the conditions utilized *in vitro* are not physiological. This may be the case with manganese, where cellular concentrations of the ion are likely to be in the low micromolar range (and certainly not 200 μM and higher). However, the *in vitro* reactions with magnesium were performed at the appropriate physiological concentration of free magnesium in a human cell (0.25–1 mM) [7]. Consequently, we wanted to compare the phenotypes of cells expressing ι-long to ι-short *in vivo*. Polι has been shown to be monoubiquitinated *in vivo* [8] and we have recently shown that the protein is polyubiquitinated after cells are exposed to naphthoquinones, especially plumbagin [9]. We introduced constructs expressing FLAG-tagged ι-long and ι-short into HEK293T cells and determined the extent of monoubiquitination and polyubiquitination in the absence and presence of plumbagin (Fig. 3A). We found that both ι-long and ι-short are monoubiquitinated in the absence of DNA damage and both exhibit a concentration dependent increase of polyubiquitination upon exposure to plumbagin. We conclude that the extra 25 amino acids encoded by ι-long do not affect the ability of polι to undergo either mono- or polyubiquitination.

3.5. Ability of ι-long to form UV-induced foci in vivo

Another property of polι-short, is its ability to re-localize into discrete replication foci after DNA damage [10]. To determine if polι-long exhibits the same traits, we generated an eGFP-polι-long construct and introduced the vector into MRC5 fibroblasts and assayed the ability of the eGFP-polι-long protein to form damage-induced foci (Fig. 3B). Similar to eGFP-polι-short, eGFP-polι-long was localized to the cell nucleus and formed foci in undamaged cells, with the extent of foci formation increasing after DNA damage. We conclude that the extra 25 amino acids encoded by ι-long do not affect the ability of polι to form replication foci in an undamaged cell and respond to the cellular signal that increases foci formation after DNA damage.

4. Discussion

The extended N-terminus of polι that is annotated in the RefSeq database is reasonably well conserved in higher eukaryotes suggesting that it may be physiologically important. However, comparison of recombinant forms of polι corresponding to ι-long (740 amino acids) and ι-short (715 amino acids) to endogenous levels of polι found in three human cell lines suggests that the predominant form of polι in these cells is ι-short (Fig. 1A–C), with only trace amounts of polι-long detected after immunoprecipitation and prolonged image exposure (Fig. 1D).

polι is a catalytic enzyme, meaning that trace amounts of the protein may nevertheless be biologically significant, especially if the extended N-terminus of polι changes the biochemical properties of the enzyme. However, *in vitro* primer extension reactions under optimal conditions for the respective full-length proteins failed to reveal any significant differences between ι-long and ι-short. Furthermore, similar to ι-short, ι-long is mono-ubiquitinated in HEK293T cells; is polyubiquitinated in response to exposure to the naphthoquinone, plumbagin; and forms UV-induced foci after UV-irradiation.

While we cannot formally exclude the possibility that under some conditions (such as after the exposure of human cells to certain types of DNA damage, or during different phases of the cell-cycle), the 740 amino acid isoform of polι might exhibit physical and functional properties different from the 715 amino acid polι protein, the evidence presented here, suggests that there are minimal differences between the two isoforms *in vitro* and *in vivo*.

Given its cellular abundance over the 740 amino acid isoform, it has been agreed upon with the curators of the RefSeq database, that the 715 amino acid polι protein be referred to as the predominant isoform found in human cells, as originally annotated in Genbank [1]. See: NM_001351632.1. During our discussions with the RefSeq curators, we realized that several other minor isoforms of polι may exist and these have also been given unique RefSeq identifiers (see Supplemental Table 1). At present, it is unknown if any of these additional protein isoforms physically exist, or have any functional relevance *in vivo*.

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

Acknowledgments

This work was supported in part by the National Institute of Child Health and Human Development/National Institutes of Health Intramural Research Program to RW. We thank Kyungjae Myung for the *POLI*$^{-/-}$ HEK293T cell line; Jean-Claude Weill for the BL2 (*POLI*$^{+/+}$) and BL2-267 (*POLI*$^{-/-}$) cell lines; Michael Murphy for help updating the RefSeq database; and Alan Lehmann, Patricia Kannouche and Tirzah Lajus for sharing unpublished data and helpful discussions.

Abbreviations

pol	DNA polymerase
TLS	translesion DNA synthesis

References

1. McDonald JP, Rapic-Otrin V, Epstein JA, Broughton BC, Wang X, Lehmann AR, Wolgemuth DJ, Woodgate R. Novel human and mouse homologs of *Saccharomyces cerevisiae* DNA polymerase η. Genomics. 1999; 60:20–30. [PubMed: 10458907]
2. Kozak M. An analysis of 5'-noncoding sequences from 699 vertebrate messenger RNAs. Nucleic Acids Res. 1987; 15:8125–8148. [PubMed: 3313277]
3. Kim J, Song I, Jo A, Shin JH, Cho H, Eoff RL, Guengerich FP, Choi JY. Biochemical analysis of six genetic variants of error-prone human DNA polymerase ι involved in translesion DNA synthesis. Chem. Res. Toxicol. 2014; 27:1837–1852. [PubMed: 25162224]

4. Frank EG, Woodgate R. Increased catalytic activity and altered fidelity of DNA polymerase ι in the presence of manganese. J. Biol. Chem. 2007; 282:24689–24696. [PubMed: 17609217]
5. Tissier A, McDonald JP, Frank EG, Woodgate R. polι, a remarkably error-prone human DNA polymerase. Genes Dev. 2000; 14:1642–1650. [PubMed: 10887158]
6. Vaisman A, Tissier A, Frank EG, Goodman MF, Woodgate R. Human DNA polymerase ι promiscuous mismatch extension. J. Biol. Chem. 2001; 276:30615–30622. [PubMed: 11402031]
7. Grubbs RD. Intracellular magnesium and magnesium buffering. Biometals. 2002; 15:251–259. [PubMed: 12206391]
8. Bienko M, Green CM, Crosetto N, Rudolf F, Zapart G, Coull B, Kannouche P, Wider G, Peter M, Lehmann AR, Hofmann K, Dikic I. Ubiquitin-binding domains in Y-family polymerases regulate translesion synthesis. Science. 2005; 310:1821–1824. [PubMed: 16357261]
9. McIntyre J, McLenigan MP, Frank EG, Dai X, Yang W, Wang Y, Woodgate R. Posttranslational regulation of human DNA polymerase ι. J. Biol. Chem. 2015; 290:27332–27344. [PubMed: 26370087]
10. Kannouche P, Fernández de Henestrosa AR, Coull B, Vidal A, Gray C, Zicha D, Woodgate R, Lehmann AR. Localisation of DNA polymerases η and ι to the replication machinery is tightly co-ordinated in human cells. EMBO J. 2002; 21:6246–6256. [PubMed: 12426396]

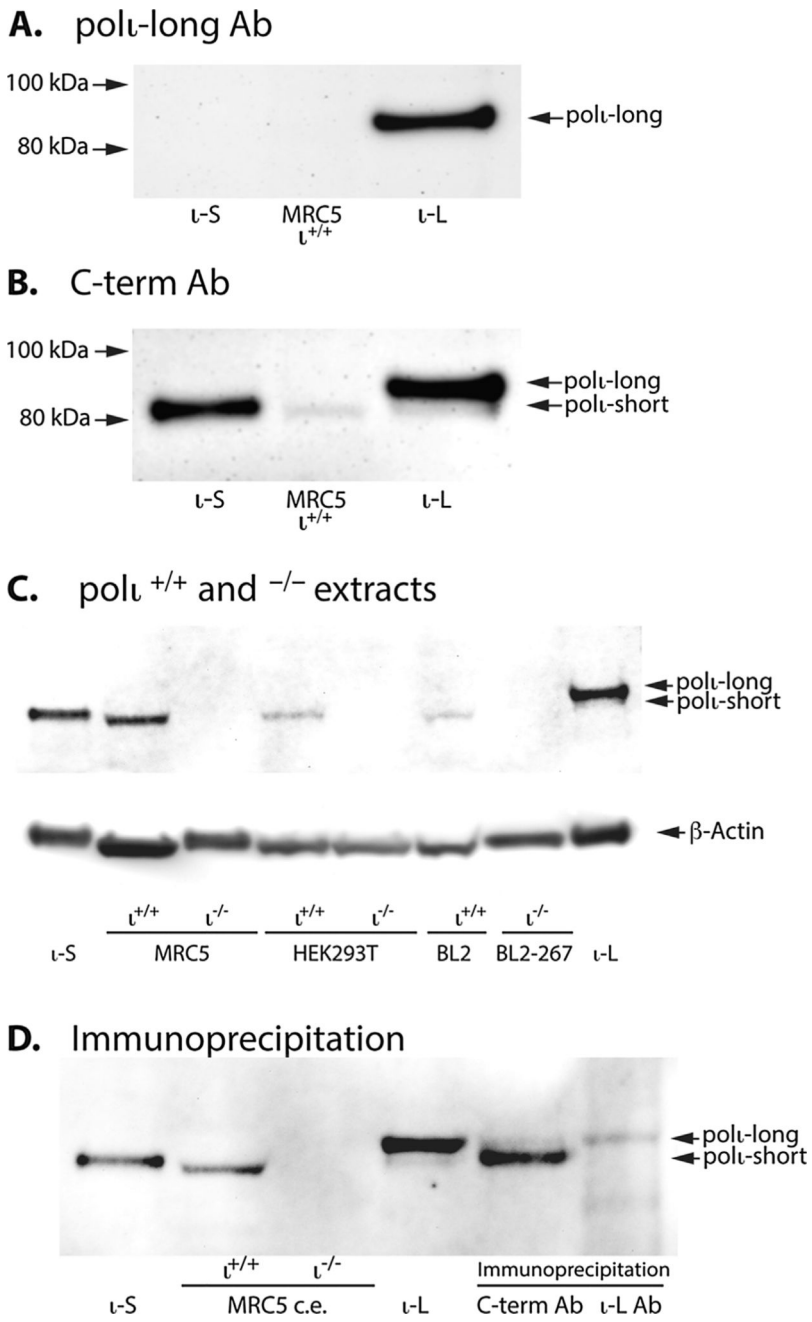

Fig. 1. Western blot analysis of polι-long and polι-short expressed in human cell extracts

A. MRC5 extracts probed with polι-long N-terminus antibodies that only detect polι-long. **B.** MRC5 extracts probed with polι C-terminus antibodies that detect both polι-long and polι-short isoforms. **C**. Western blot analysis using polι C-terminus antibodies to detect polι-long and polι-short expression in six *POLI*$^{+/+}$ or$^{-/-}$ human cell lines cell extracts and, as a controls, in MRC5 polι$^{-/-}$ cells expressing recombinant polι-short or polι-long. The lower part of the same gel was probed with β–Actin antibodies. **D.** Western blot of MRC5 polι$^{-/-}$ cells expressing recombinant polι-short (ι-S), polι-long (ι-L), MRC5 polι$^{+/+}$ and MRC5 polι$^{-/-}$ cell extracts after proteins were immunoprecipitated by polι C-terminal

monoclonal mouse antibodies (Abnova), or polyclonal antibodies raised to the N-terminus of polι-long. Immunoprecipitated polι was detected using polyclonal antibodies raised to the C-terminus of polι, but only after prolonged exposure of the blot.

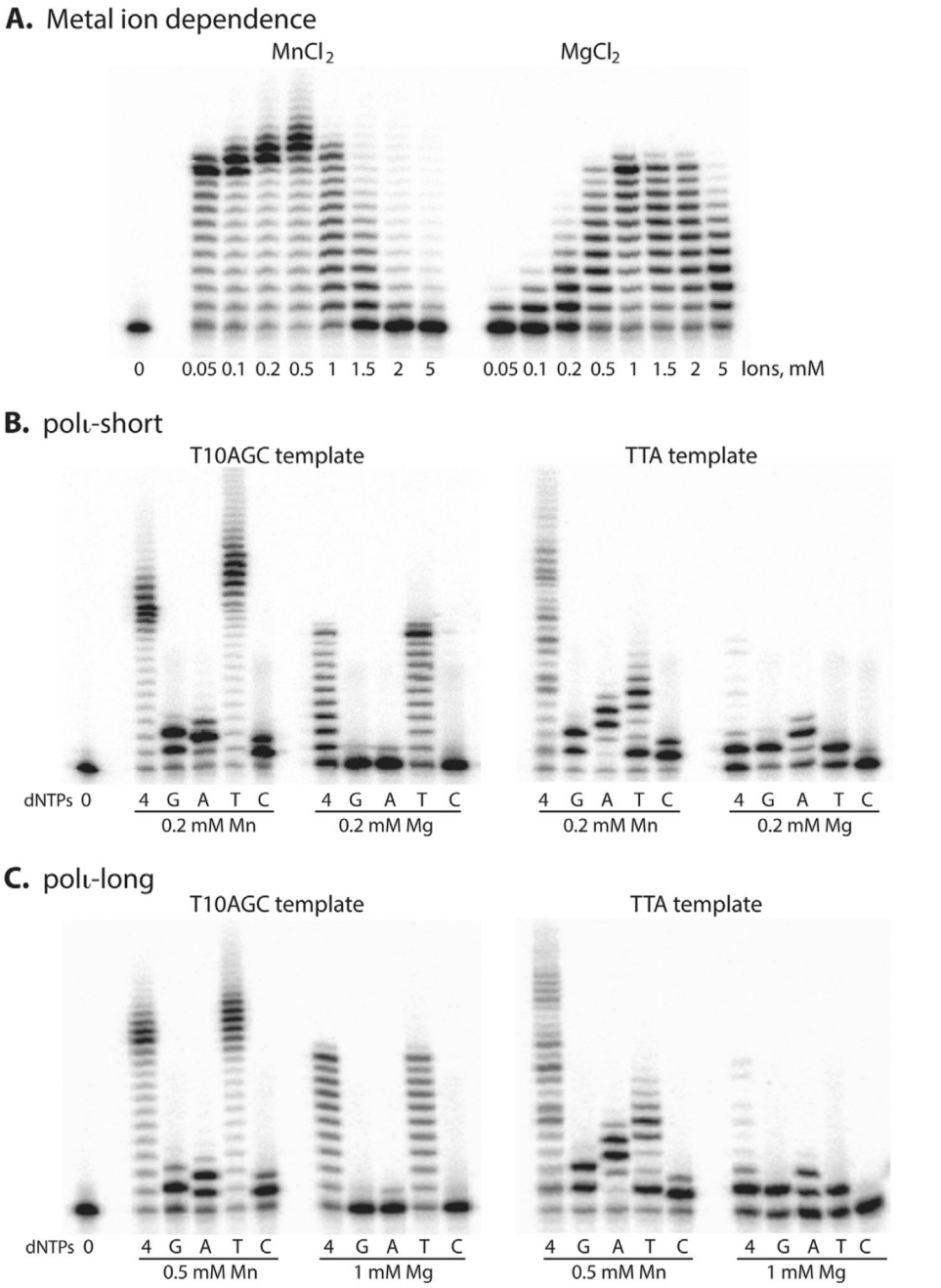

Fig. 2. *In vitro* analysis of polι

A. The ability of polι-long (expressed and purified from *E. coli*) to extend a radiolabeled primer annealed to the "T10AGC" template depends on the concentration of Mn^{2+} or Mg^{2+} ions. *In vitro* comparison of primer extension activity of polι-short (**B**), or polι-long (**C**), on two DNA templates. Replication reactions were performed in the presence of all four dNTPs (4) or individual dNTPs (G, A, T, C) on templates "T10AGC" or "TTA", under optimal concentrations of Mn^{2+} or Mg^{2+}. Under these conditions, the *in vitro* primer extension properties of polι-short and polι-long are largely indistinguishable.

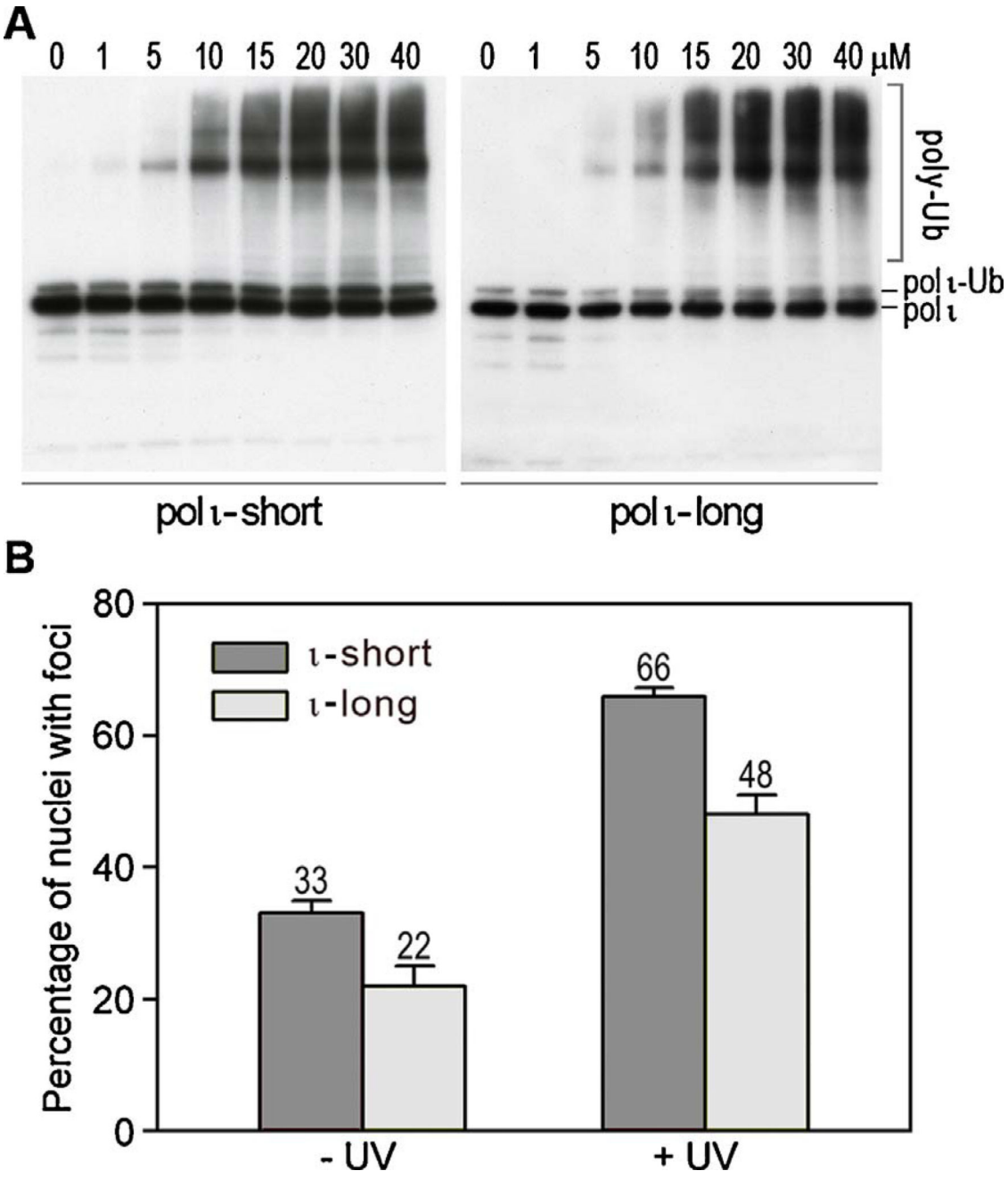

Fig. 3. *In vivo* analysis of polι

A. Comparison of the effect of plumbagin treatment on polι-short and polι-long ubiquitination in HEK293T cells. Cells transfected with either FLAG-tagged polι-short or FLAG-tagged polι-long were treated with the indicated amount of plumbagin for 1 h prior to harvesting. Both isoforms of polι were visualized in Western blots using monoclonal antibodies to the N-terminal FLAG epitope. In an undamaged cell, polι-short and polι-long exist as unmodified and mono-ubiquitinated forms. After plumbagin treatment, both isoforms are polyubiquitinated in a dose response manner. **B.** pol ι-short and pol ι-long form foci in undamaged and UV-irradiated cells. Though pol ι-short exhibits a slightly

better ability to form foci in undamaged and UV-treated cells, both isoforms of polι exhibit a similar ~2-fold increase in the number of foci in response to UV-irradiation.