



Characterization of marine X-family DNA polymerases and comparative analysis of base excision repair proteins

Hirofumi Kodera^a, Ryo Takeuchi^a, Yukinobu Uchiyama^a, Yoichi Takakusagi^a, Kazuki Iwabata^a, Hidetsugu Miwa^b, Naoto Hanzawa^c, Fumio Sugawara^a, Kengo Sakaguchi^{a,*}

^a Department of Applied Biological Science, Tokyo University of Science, 2641 Yamazaki, Noda-shi, Chiba-ken 278-8510, Japan

^b Department of Biosciences, University of Helsinki, Viikinkaari, Helsinki FIN-00790, Finland

^c Graduate School of Science and Engineering, Yamagata University, Koshirakawamachi, Yamagata-shi, Yamagata 990-8560, Japan

ARTICLE INFO

Article history:

Received 4 October 2011

Available online 18 October 2011

Keywords:

DNA polymerase beta
X-family DNA polymerase
Marine POLX
Short-patch BER

ABSTRACT

While mammalian DNA polymerase β (Pol β), which is a member of the Pol X family, play important roles in base excision repair (BER) that efficiently removes DNA base lesions arising from both endogenous and exogenous agents, this protein has been found only a subset of animals. To understand natural evolution of this enzyme, we isolated and characterized Pol β from jellyfish *Aurelia* sp.1. (AsPol β). Despite of phylogenetic distance and environmental differences between jellyfish and mammals, *in vitro* assays showed biochemical characteristics of AsPol β were very similar to those of a mammalian counterpart. We also searched two other homologs of mammalian genes that were involved in short patch (sp) BER in the nucleotide sequence database, and found that both of these homologs were encoded in the genomes of a lineage from Cnidarians through mammals and Arthropods. This study suggests that a DNA repair mechanism resembling mammalian sp-BER may be largely limited to a subset of animals. On the basis of our findings and previous reports, we discuss possible evolutionary model of Pol β and the other members of the Pol X family.

© 2011 Elsevier Inc. All rights reserved.

1. Introduction

Sixteen types of DNA polymerases (Pols) and related genes including terminal deoxynucleotidyl transferase (TdT), which are categorized into four groups (families A, B, X and Y), have been found in the eukaryotic genomes [1–3]. While only a subset of Pols are involved in nuclear and organelle DNA replication processes [4,5], all except Pol α have been suggested to participate in highly diverse DNA repair pathways, which may be essential for maintenance of genome integrity in a wide variety of biological events such as development [1–3,6–10].

Among four types of eukaryotic Pol families, particularly the Pol X family genes show unique distribution across the kingdoms of eukaryotes [11]: plants and lower fungi such as slime mold, have only one X-family gene (Pol λ) [11,12], whereas higher fungi such as basidiomycetes contain two (Pols λ and μ) [13]. In metazoans, the most basal phylum that possesses three polymerases genes appears to be Echinodermata (Pol β , Pol λ , Pol μ /TdT). The Pol μ /TdT gene displays similar identity to both mammalian Pol μ and TdT genes,

and thus is likely an ancestor of Pol μ and TdT genes that are coded in the vertebrate genomes [11]. In contrast the mammalian genomes encode four types of the X-family Pols (Pol β , Pol λ , Pol μ and TdT). Since the X-family genes that share high identity with mammalian Pol λ gene appear to be widely spread across the kingdoms of eukaryotes (except a subset of protostomes, which appear to contain no Pol), the eukaryotic X-family genes might be diverged from a single gene that was homologous to Pol λ gene [11–13].

The X-family Pols have been well studied in mammals and known to participate in multiple DNA repair pathways. The three X-family proteins (Pol λ , Pol μ and TdT) are common repair pathways, immunoglobulin V(D)J recombination and nonhomologous end-joining (NHEJ), and disruption of each gene function results in a similar phenotype of knockout mice [9,14,15]. In contrast, Pol β is a major player in base excision repair (BER) consisting of short patch (sp) and long patch (lp) pathways by catalyzing removal of a 5'-deoxyribose phosphate (dRP) and gap filling, in addition to homologous recombination at early stage of meiosis [16]. In sp-BER, a damaged base is excised by a DNA glycosylase. The DNA backbone is then incised on 5' side of the apurinic/apyrimidinic (AP) site by an AP endonuclease, leaving a nicked intermediate, which is a substrate for 5'-deoxyribose phosphate (dRP) lyase and Pol activities of Pol β . Pol β fills the single-nucleotide gap, generating a nicked duplex that either DNA Ligase I (Lig1) or III (Lig3)/XRCC1 complex

* Corresponding author. Address: Department of Applied Biological Science, Faculty of Science and Technology, Tokyo University of Science, 2641 Yamazaki, Noda, Chiba-ken 278-8510, Japan. Fax: +81 471 23 9767.

E-mail address: kengo@rs.noda.tus.ac.jp (K. Sakaguchi).

seals. When the dRP is refractory to the removal by Pol β , the lp-BER pathway repairs the damaged site. Pol β display nucleotide incorporation to a 3'-primer end, displacing the downstream strand. The 5'-flapped structure is processed by a flap endonuclease, and the resulting nick is fixed by Lig1 [17]. Although we previously showed that putative Pol β homolog genes are likely spread in a subset of animals [11], it remains unclear whether all of them display biochemical activity similar to mammalian counterparts. Here, we describe the Pol β homolog found from a marine jellyfish (*Aurelia* sp.1.) that shows enzymatic properties similar to its mammalian counterparts. Following experimental verification, we searched Lig3 and XRCC1 homolog genes in the genomes of various organisms. Collectively, our study suggests that a repair pathway similar to sp-BER may be conserved only in a subset of animals: DNA-base lesions eliminated by spBER in mammalian cells may be repaired by a wide variety of repair proteins in organisms of the other phyla.

2. Materials and methods

2.1. Collection of jellyfish (*Aurelia* sp.1.) polyps

The polyp stage of *Aurelia* sp.1. was a gift from Mr. K. Okuizumi of Tsuruoka City Kamo Aquarium. The polyps were divided into small groups and immediately stored at -80°C until use.

2.2. cDNA cloning of X family DNA polymerases in *Aurelia* sp.1. and expression and purification of their recombinant proteins

Total RNA was extracted from polyps based on established procedure [18]. Total RNA from *Aurelia* sp.1. polyps was reverse transcribed using the SuperScript III First-Strand Synthesis System (Invitrogen) with an oligo-(dT)₂₀ primer.

The following degenerate primers were used: AsPol β , sense 5'-GTGGGGACATGGATGTTATA(A/C/T)T(A/C/G/T)AC(A/C/G/T)CA(C/T)CC and antisense 5'-GCCACCCTATATAATCGAAAA(C/T)(A/G)TC(C/T)TT(C/T)TC, which were designed on the basis of the sequence homology shared by the *Homo sapiens* (Hs) Pol λ (Accession No. NP_037406), HsPol β (NP_002681), and *Oryza sativa* (Os) Pol λ (BAD18976), obtained from the National Center for Biotechnology Information Entrez Protein Database (NCBI, <http://www.ncbi.nlm.nih.gov/protein/>), and AsPol λ , sense 5'-GGCAGATTATTACTGGACAAGGTGA(C/T)(A/C)A(A/G)TGG(A/C)G and antisense 5'-GATGGT AGTCTACTCATTAAATTTGTAG(A/C/G/T)T(A/G)(C/T)AG(A/C/G/T)GG, which were designed on the basis of the sequence homology shared by the HsPol λ , and *Aspergillus fumigatus* POL4 (EAL90376), obtained from NCBI. These primers were used to screen the cDNA library of the *Aurelia* sp.1. polyps by means of PCR. Sequence analysis of the PCR product revealed an incomplete open reading frame that encoded an amino acid sequence that was similar to the HsPol β and λ gene products. The cDNA from the mRNA library of the polyps were isolated using the GeneRacer™ Kit (Invitrogen) for the rapid amplification of 5' and 3' cDNA ends (5'-RACE and 3'-RACE). We named these genes AsPol β and AsPol λ . The nucleotide sequence data reported in this paper have been submitted to the DDBJ/EMBL/GenBank nucleotide sequence databases with the accession number AB636589 for AsPol β and AB636590 for AsPol λ .

AsPol β cDNA fragments, which contained a six-histidine C-terminal tag, were isolated by reverse transcription (RT)-PCR of *Aurelia* sp.1. cDNA using the following PCR primers: sense 5'-GGATCCATGAGCAAGAGAAAGGCACCAAG-3' and antisense 5'-CTCGAGCTTGTTACGATCAGTTGGTTCTACGTA-3'. The underlined portions of these sequences refers to the restriction enzyme sites, GGATCC (*Bam*HI), and CTCGAG (*Xho*I). The full-length AsPol β coding region was cloned into the pET-28a(+) expression vector (Novagen) and

transformed into *Escherichia coli* BL21(DE3) (Novagen). Overexpression and purification of AsPol β protein was accomplished as reported previously [13,19]. Recombinant AsPol β was purified using a HiLoad 16/60 Superdex 75 pg column (GE Healthcare Bio-Sciences KK). Aliquots of recombinant AsPol β were dialyzed against TEMG buffer containing 150 mM NaCl and flash-frozen in liquid nitrogen, and then stored at -80°C . DNA polymerase activity was measured as described previously [13,20].

2.3. Fidelity assay of AsPol β protein

A fidelity assay was performed using the methods described by Takeuchi et al. [4]. Synthetic oligonucleotides purified by PAGE were obtained from Sigma-Aldrich. Template-Primer molecules used for insertion assays were generated by annealing P1 primer (5'-CAGATCAGCAAGCTAG-3') to four T1 templates (5'-GTCGACCTGCTA C $\underline{\text{X}}$ CTAGCTTGCTGATCTG-3' where $\underline{\text{X}}$ is A, C, G or T). P1 primer was labeled at its 5'-end with [γ -³²P] ATP (3000 Ci/mmol, Amersham) and T4 polynucleotide kinase (New England Biolabs). These primers were then hybridized to template oligonucleotides to generate all different molecules in the presence of 20 mM Tris-HCl (pH 7.5), 150 mM NaCl, 1 mM EDTA, and heating the DNA mixtures for 5 min at 95°C followed by slow cooling to 25°C .

2.4. Terminal deoxynucleotidyl transferase assay

Terminal deoxynucleotidyl transferase (TdT) was assayed in a final volume of 20 μl containing 50 mM Tris-HCl (pH 7.5), 1 mM dithiothreitol, 2.5 mM MgCl₂, 25 mM dNTPs, 7.5% glycerol and HsTdT (0, 1.16, 0.98 pmol) (as positive control), AsPol β (0, 4.9, 2.5, 1.2, 0.6 pmol) or RnPol β (0, 4.9, 2.5, 1.2, 0.6 pmol) (as negative control). The oligonucleotide (5'-CAGCAACGCAAGCTTG-3') was 5'-³²P end-labeled by using polynucleotide kinase. After incubation at 37°C for 1 h, the DNA was analyzed by 10% denaturing PAGE in the presence of 8 M urea.

2.5. Other methods

Recombinant rat DNA polymerase β was overexpressed in *E. coli* BL21(DE3) (Novagen) harboring the expression plasmid pET-28a(+) and purified in the same way as AsPol β . The rat DNA polymerase β cDNA in pET-28a(+) had been constructed in our laboratory under the local ethical committees of the Tokyo University of Science [21] and we reused the existing sample in this study. In the case of the other vertebrate, we used only amino acid sequences.

We obtained the amino acid sequences of X-family DNA polymerases (Pol β , Pol λ , Pol μ and TdT), and sp-BER related proteins (XRCC1 and Lig3) from the National Center for Biotechnology Information Entrez Protein Database (<http://www.ncbi.nlm.nih.gov>).

Phylogenetic analysis was performed based on the amino acid sequence by the neighbor-joining method using three software applications (Clustalx2, BioEdit and NJplot) [22,23]. DNA polymerase domains of various X-family DNA polymerases were aligned, removed all gaps and used to produce the tree.

3. Results

3.1. Identification and sequence analysis of the X-family DNA polymerase genes Pol β and Pol λ from *Aurelia* sp.1. (AsPol β and AsPol λ)

We isolated the full-length cDNA of AsPol β encoding 335 amino acid residues (Fig. 1). The nucleotide sequence was deposited in the DDBJ/EMBL/GenBank nucleotide sequence database with the accession number AB636589. Surprisingly, regardless of phylogenetic distance, AsPol β shares a similar level of amino-acid se-

| dRP lyase (8 kDa) Domain | |
|---------------------------------|--|
| AspPolβ | 1 MSKRKAPKONPSE-----YCDFLTELADYEKNVENSFKSKVYRKAASILANHPTKITSGAEARKLDGIGDKI |
| HspPolβ | 1 MSKRKAPQETLNGC-----ITDMLTELANFEKNVSCAIHKYNAYRKAASVIKYPKIKSGAEAKKLDGVCCKI |
| RnpPolβ | 1 MSKRKAPQETLNGC-----ITDMLTELANFEKNVSCAIHKYNAYRKAASVIKYPKIKSGAEAKKLDGVCCKI |
| MdpPolβ | 1 MSKRKAPQETLNGC-----ITDFLTELANYERNVSCAIHKYNAYRKAASVIKYPKIKSGAEAKKLDGVCCKI |
| GgpPolβ | 1 MSKRKAPQESLNGC-----ITDFLMELANYERNVSCAIHKYNAYRKAASVIRYPSRIKSGAEAKKLDGVCCKI |
| XtpPolβ | 1 MSKRKAPQESLNGC-----ITDFLVELANYERNVSCAIHKYNAYRKAASVIKYPKIKSGAEAKKLDGVCCKI |
| DrpPolβ | 1 MSKRKAPQESLNGC-----ITDFLVELANYERNVSCAIHKYNAYRKAASVIKYPKIKSGAEAKKLDGVCCKI |
| SppPolβ | 1 MSKRKAPDSNNPNQDIYVFHFFFRFSCCISELANYEKNVNRAMHKNAYRKAASVIKYPKIKSGAEAKKLDGVCCKI |
| TcpPolβ | 1 MCKRKAPSGENTPNSD-----FCEFLTELADYEKNVSRNIHKYNAYRKAASVLSHPTTRITSGAEAKKLDGVCCKI |
| HmpPolβ | 1 ----- |
| NvpPolβ | 1 MSKRKAPDQNPSE-----FCDFLTELADYEKNVTRQIHKYNAYRKAASVLAHKSHTKIKDGAERKLDGIGDKI |
| Boxed Region | |
| AspPolβ | 70 EKKIDEFLATGKLRKLEKIRDDTSS-SINFLTRVSGIGPAAARLVDDG-ICTSIEELK---KNNDKLNHHQKIGLKHYF |
| HspPolβ | 70 AEKIDEFLATGKLRKLEKIRDDTSS-SINFLTRVSGIGPAAARLVDDG-ICTSIEELK---KNNDKLNHHQKIGLKHYF |
| RnpPolβ | 70 AEKIDEFLATGKLRKLEKIRDDTSS-SINFLTRVSGIGPAAARLVDDG-ICTSIEELK---KNNDKLNHHQKIGLKHYF |
| MdpPolβ | 70 AEKIDEFLATGKLRKLEKIRDDTSS-SINFLTRVSGIGPAAARLVDDG-ICTSIEELK---KNNDKLNHHQKIGLKHYF |
| GgpPolβ | 70 AEKIDEFLATGKLRKLEKIRDDTSS-SINFLTRVSGIGPAAARLVDDG-ICTSIEELK---KNNDKLNHHQKIGLKHYF |
| XtpPolβ | 70 AEKIDEFLATGKLRKLEKIRDDTSS-SINFLTRVSGIGPAAARLVDDG-ICTSIEELK---KNNDKLNHHQKIGLKHYF |
| DrpPolβ | 70 AEKIDEFLATGKLRKLEKIRDDTSS-SINFLTRVSGIGPAAARLVDDG-ICTSIEELK---KNNDKLNHHQKIGLKHYF |
| SppPolβ | 81 AEKIDEFLATGKLRKLEKIRDDTSS-SINFLTRVSGIGPAAARLVDDG-ICTSIEELK---KNNDKLNHHQKIGLKHYF |
| TcpPolβ | 72 AEKIDEFLATGKLRKLEKIRDDTSS-SINFLTRVSGIGPAAARLVDDG-ICTSIEELK---KNNDKLNHHQKIGLKHYF |
| HmpPolβ | 1 ---IDEFLATGKLRKLEKIRDDTSS-SINFLTRVSGIGPAAARLVDDG-ICTSIEELK---KNNDKLNHHQKIGLKHYF |
| NvpPolβ | 70 EKKIDEFLATGKLRKLEKIRDDTSS-SINFLTRVSGIGPAAARLVDDG-ICTSIEELK---KNNDKLNHHQKIGLKHYF |
| AspPolβ | 144 EDPEKRIPREEMQMDIVLNEVKKVDSEYIATVCGSFRRGAESSGDMVLLTHPFTSSESAK---OPKLLHVVVEQLQ |
| HspPolβ | 144 EDPEKRIPREEMQMDIVLNEVKKVDSEYIATVCGSFRRGAESSGDMVLLTHPFTSSESAK---OPKLLHVVVEQLQ |
| RnpPolβ | 144 EDPEKRIPREEMQMDIVLNEVKKVDSEYIATVCGSFRRGAESSGDMVLLTHPFTSSESAK---OPKLLHVVVEQLQ |
| MdpPolβ | 144 DDPEKRIPREEMQMDIVLNEVKKVDSEYIATVCGSFRRGAESSGDMVLLTHPFTSSESAK---OPKLLHVVVEQLQ |
| GgpPolβ | 144 EDPEKRIPREEMQMDIVLNEVKKVDSEYIATVCGSFRRGAESSGDMVLLTHPFTSSESAK---OPKLLHVVVEQLQ |
| XtpPolβ | 144 DDPEKRIPREEMQMDIVLNEVKKVDSEYIATVCGSFRRGAESSGDMVLLTHPFTSSESAK---OPKLLHVVVEQLQ |
| DrpPolβ | 144 EDPEKRIPREEMQMDIVLNEVKKVDSEYIATVCGSFRRGAESSGDMVLLTHPFTSSESAK---OPKLLHVVVEQLQ |
| SppPolβ | 156 EDPEKRIPREEMQMDIVLNEVKKVDSEYIATVCGSFRRGAESSGDMVLLTHPFTSSESAK---OPKLLHVVVEQLQ |
| TcpPolβ | 146 EDPEKRIPREEMQMDIVLNEVKKVDSEYIATVCGSFRRGAESSGDMVLLTHPFTSSESAK---OPKLLHVVVEQLQ |
| HmpPolβ | 78 DKIKRSIPREEMQMDIVLNEVKKVDSEYIATVCGSFRRGAESSGDMVLLTHPFTSSESAK---OPKLLHVVVEQLQ |
| NvpPolβ | 144 DEFETRIPREEMQMDIVLNEVKKVDSEYIATVCGSFRRGAESSGDMVLLTHPFTSSESAK---OPKLLHVVVEQLQ |
| AspPolβ | 224 KANYITDTLS--LGEAKFMGVCKLPDNE-----SYRRLDLRLIPGDQYYCGVLYFTGSDIFNKNMRAHALEKGFNTINE |
| HspPolβ | 220 KVHFITDTLS--KGETKFMGVCKLPDNE-----SYRRLDLRLIPGDQYYCGVLYFTGSDIFNKNMRAHALEKGFNTINE |
| RnpPolβ | 220 KVHFITDTLS--KGETKFMGVCKLPDNE-----SYRRLDLRLIPGDQYYCGVLYFTGSDIFNKNMRAHALEKGFNTINE |
| MdpPolβ | 220 KVHFITDTLS--KGETKFMGVCKLPDNE-----SYRRLDLRLIPGDQYYCGVLYFTGSDIFNKNMRAHALEKGFNTINE |
| GgpPolβ | 220 KVHFITDTLS--KGETKFMGVCKLPDNE-----SYRRLDLRLIPGDQYYCGVLYFTGSDIFNKNMRAHALEKGFNTINE |
| XtpPolβ | 220 ECEFITDTLS--KGETKFMGVCKLPDNE-----SYRRLDLRLIPGDQYYCGVLYFTGSDIFNKNMRAHALEKGFNTINE |
| DrpPolβ | 220 SGEFITDTLS--KGETKFMGVCKLPDNE-----SYRRLDLRLIPGDQYYCGVLYFTGSDIFNKNMRAHALEKGFNTINE |
| SppPolβ | 232 DSKLVITDTIS--QGDTKFMGCIIRLPEKDAKEGEIYFRRRLDRLIPGDQYYCGVLYFTGSDIFNKNMRAHALEKGFNTINE |
| TcpPolβ | 220 NCGLITETIS--QGDTKFMGCIIRLPEKDAKEGEIYFRRRLDRLIPGDQYYCGVLYFTGSDIFNKNMRAHALEKGFNTINE |
| HmpPolβ | 155 EIGFIDKDSFSEDVSTKVMGVCKLPDNE-----MLRRIDIRFIPQESYHTALLYFTGSDIFNKNMRAHALEKGFNTINE |
| NvpPolβ | 222 KADFTVDTLS--LGEAKFMGVCKLPDNE-----SYRRLDLRLIPGDQYYCGVLYFTGSDIFNKNMRAHALEKGFNTINE |
| AspPolβ | 295 YCIRPVGSTGAPGEPIPVSEKDVDFDILGKEYVEPTDRNK |
| HspPolβ | 296 YTIRPLGVTGVAGEPLPVDSEKIDFYIOWKYREPKDRSE |
| RnpPolβ | 296 YTIRPLGVTGVAGEPLPVDSEKIDFYIOWKYREPKDRSE |
| MdpPolβ | 296 YTIRPLGVTGVAGEPLPVDSEKIDFYIOWKYREPKDRSE |
| GgpPolβ | 296 YTIRPLGVTGVAGEPLPVDSEKIDFYIOWKYREPKDRSE |
| XtpPolβ | 295 YTIRPLGVTGVAGEPLPVDSEKIDFYIOWKYREPKDRSE |
| DrpPolβ | 298 YTIRPLGVTGVAGEPLPVDSEKIDFYIOWKYREPKDRSE |
| SppPolβ | 310 YTIRPVGSTGAPGEPIPVSEKDVDFDILGKEYVEPTDRNK |
| TcpPolβ | 291 YTIRPVGSTGAPGEPIPVSEKDVDFDILGKEYVEPTDRNK |
| HmpPolβ | 228 YRILGN-----NEIFPKIMSEKIDFYIOWKYREPKDRSE |
| NvpPolβ | 297 YSIRPLGVTGVAGEPLPVDSEKIDFYIOWKYREPKDRSE |

Fig. 1. Alignment of *Aurelia* sp.1. Pol β amino acid sequences. Alignment of predicted amino acid sequences of Pol β from *Aurelia* sp.1 (As), *Homo sapiens* (Hs), *Rattus norvegicus* (Rn), *Monodelphis domestica* (Md), *Gallus gallus* (Gg), *Xenopus tropicalis* (Xt), *Danio rerio* (Dr), *Strongylocentrotus purpuratus* (Sp), *Tribolium castaneum* (Tc), *Hydra magnipapillata* (Hm) and *Nematostella vectensis* (Nv). The accession numbers are: AspPol β: AB636589, HspPol β: NP_002681, RnpPol β: NP_058837, MdpPol β: XP_001373104, GgpPol β: XP_424407, XtpPol β: NP_001006894, DrpPol β: NP_001003879, SppPol β: XP_787665, TcpPol β: AB639754, HmpPol β: XP_002165723 and NvpPol β: XP_001628358. Highly conserved residues are shaded, and alignment gaps are indicated by dashes. The dRP lyase domain is enclosed in a box.

quence identity (~55%) and similarity (~75%) with deuterostome Pol β proteins (Fig. 1). Although the X-family gene is absent in the genomes of well-characterized Ecdysozoans including *Caenorhabditis elegans* and *Drosophila melanogaster*, whole genome sequencing of *Tribolium castaneum* (the red flour beetle, a member of Arthropoda) [24] found an open reading frame (ORF) encoding Pol β homolog. This putative gene displays high identity with the homologs of deuterostome and jellyfish (Fig. 1). Whole genome sequencing of two Cnidarians, sea anemone (*Nematostella vectensis*) and hydra (*Hydra magnipapillata*) has been recently described [25,26]. Like other Pol β homologs, the amino acid sequences of the putative Pol β from these Cnidarians are well conserved, although *Hydra* Pol β gene lack its 5'-terminal region encoding the N-terminal 70 amino acid residues (Fig. 1). The absence of the start codon in the *Hydra* ORF sequence suggests that the deposited cDNA is partial or for a pseudogene (Fig. 1). These observations indicate that the Pol β homolog genes are broadly spread in the animal kingdom and have considerably high sequence identity.

We also isolated the partial cDNA of AsPol λ , which encoded a homologous protein to animal Pol λ enzymes (~40% identity and ~60% similarity; Fig. S1). Sequences encoding AsPol β and AsPol λ are obviously distal, suggesting these two individual X-family genes coded on the jellyfish genome. Pol β and Pol λ homolog genes have been found in two species of Cnidarians that were subjected to whole genome sequencing, but no gene homologous to the other two X-family Pol genes, suggesting that the genome of jellyfish harbors the two X-family genes.

3.2. Overexpression, purification and properties of the recombinant AsPol β protein

To compare biochemical activity of AsPol β with that of mammalian Pol β , we sought to obtain the purified recombinant protein. The ORF of AsPol β gene was cloned into a bacterial protein expression vector, and induced overexpression of the recombinant protein in

bacteria. The His-tagged AsPol β protein was purified to near homogeneity, using Ni²⁺-NTA agarose and HiLoad 16/60 Superdex 75 pg column chromatography (Fig. 2A). The elution pattern of the purified protein analyzed on a SDS-PAGE gel was in a good agreement with that of the DNA polymerase activity (Fig. 2B), indicating that the recombinant enzyme is substantially active.

We first tested impacts of canonical Pol inhibitors on the DNA synthetic activity of the recombinant AsPol β protein. Like mammalian (*Rattus norvegicus*) Pol β (RnPol β), AsPol β was sensitive to dideoxythymidine triphosphate, which is a potent inhibitor of the X-family Pols and a subset of the A-family Pols [13,20] (Fig. S2A). In contrast, the activity was little influenced by aphidicolin and N-ethylmaleimide (NEM), which inhibit primer extension of the replicative DNA polymerases α , δ , and ϵ (Fig. S2B and C). AsPol β was active in a broad range of temperatures, but substantially inactivated at 42 °C, where RnPol β retained its catalytic activity (Fig. 2C). Since jellyfish enzymes necessitate functioning at lower temperatures than mammalian ones, AsPol β might tolerate accumulation of nonsynonymous substitutions that reduced its thermal stability through evolution. The activity of RnPol β and AsPol β was similarly varied between pH 6 and pH 8.5, but the AsPol β activity was more severely reduced at higher pH (Fig. 2D). Both of RnPol β and AsPol β require divalent cations for their catalysis, and were much more active in the presence of manganese ions (Fig. S2D), as described previously [27,28]. These enzymes also showed a similar sensitivity to a monovalent salt (NaCl) (Fig. S2E).

We next compared preference of nucleotide incorporation catalyzed by these two Pols. The results shown in Fig. S3A and B suggest the enzymes display similar selectivity of nucleotides incorporated opposite to all four possible template bases. We also assayed template-independent polymerase activity that mammalian TdT and Pol μ possess [9,29,30], because jellyfish are unlikely to contain Pol μ /TdT gene as described above. Like the mammalian counterpart, AsPol β showed no activity under the conditions tested (Fig. S3C), whereas human TdT substantially extended a

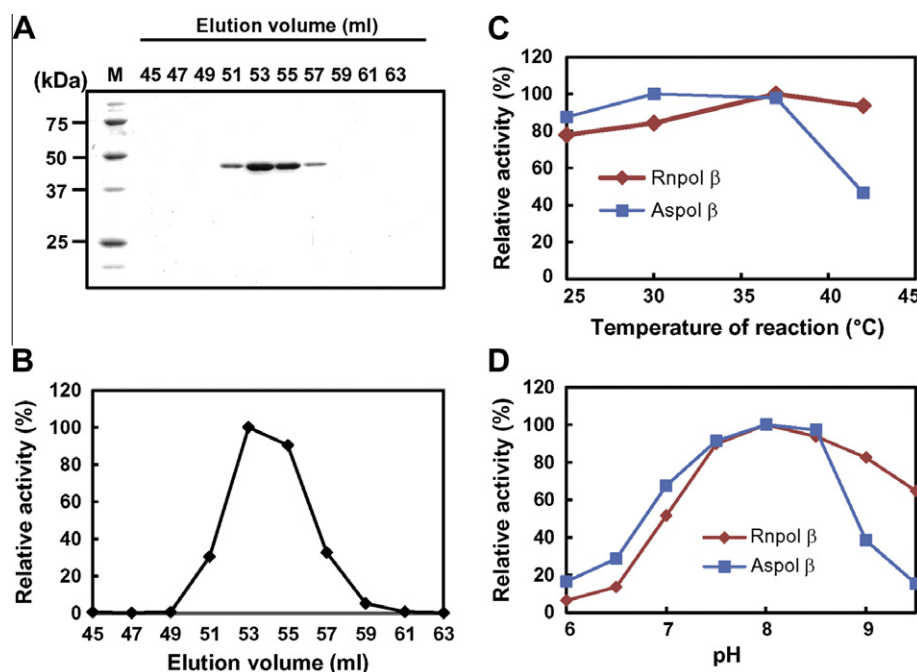


Fig. 2. Purification and characterized the recombinant AsPol β protein. (A) SDS-PAGE analysis of the AsPol β protein fractions purified by gel filtration chromatography. AsPol β was fractionated by 12.5% SDS-PAGE and stained with Coomassie brilliant blue. (B) Elution pattern with gel filtration chromatography. Fractions were tested under the standard assay conditions using poly(dA)/oligo(dT)₁₀₋₁ as template primer. (C) Effect of the reaction temperature on activity of AsPol β (squares) and RnPol β (diamonds). (D) Effect of pH with Tris-HCl buffers on activity of AsPol β (squares) and RnPol β (diamonds).

Table 1
Protein involved in sp-BER. Proteins involved in sp-BER were obtained using a BLAST search. If no match was found in an organism with a sequenced genome, this is shown by the entry “none”, and if no match was found in an organism with a genome that is not sequenced, this is shown by the entry “unknown”.

| Class | Organism | DNA polymerase β | | DNA ligase III | | XRCC1 | |
|-----------------|--------------------------------------|------------------------|-------------|------------------------|-------------|------------------------|-------------|
| | | Accession | Length (aa) | Accession | Length (aa) | Accession | Length (aa) |
| Mammalia | <i>Homo sapiens</i> | NP_002681 | 335 | NP_002302 | 949 | NP_006288 | 633 |
| Mammalia | <i>Pan troglodytes</i> | XP_001143904 | 335 | XP_511409 | 922 | XP_001156934 | 735 |
| Mammalia | <i>Canis lupus familiaris</i> | XP_532790 | 335 | XP_548265 | 991 | XP_533653 | 654 |
| Mammalia | <i>Rattus norvegicus</i> | NP_058837 | 335 | NP_001012011 | 943 | NP_445887 | 631 |
| Aves | <i>Gallus gallus</i> | XP_424407 | 335 | NP_001006215 | 902 | None | |
| Amphibia | <i>Xenopus laevis</i> | NP_001006894 | 334 | NP_001082183 | 988 | NP_001080711 | 651 |
| Actinopterygii | <i>Danio rerio</i> | NP_001003879 | 337 | NP_001025345 | 752 | NP_001003988 | 615 |
| Echinoidea | <i>Strongylocentrotus purpuratus</i> | XP_787665 | 349 | XP_786357 | 875 | XP_001180271 | 512 |
| Insecta | <i>Drosophila melanogaster</i> | None | None | XP_650187 | 806 | NP_572217 | 614 |
| Insecta | <i>Tribolium castaneum</i> | BAK40156 | 330 | XP_967954 | 853 | XP_975029 | 533 |
| Chromadorea | <i>Caenorhabditis elegans</i> | None | | None | | None | |
| Hydrozoa | <i>Hydra magnipapillata</i> | XP_002165723 (partial) | 262 | XP_002155088 (partial) | 134 | XP_002165367 (partial) | 292 |
| Scyphozoa | <i>Aurelia sp.1</i> | BAK40153 | 334 | Unknown | | Unknown | |
| Anthozoa | <i>Nematostella vectensis</i> | XP_001628358 (partial) | 336 | XP_001630058 (partial) | 894 | XP_001633392 (partial) | 199 |
| Agaricomycetes | <i>Coprinopsis cinerea</i> | None | | Unknown | | Unknown | |
| Saccharomycetes | <i>Saccharomyces cerevisiae</i> | None | | None | | None | |
| Liliopsida | <i>Oryza sativa Japonica Group</i> | None | | None | | BAF46899 | 347 |
| Magnoliopsida | <i>Arabidopsis thaliana</i> | None | | None | | NP_178158 | 353 |
| Liliopsida | <i>Zea mays</i> | None | | None | | NP_001149899 | 351 |
| Chlorophyceae | <i>Chlamydomonas reinhardtii</i> | None | | None | | None | |
| Bryopsida | <i>Physcomitrella patens</i> | None | | None | | XP_001777620 | 367 |

single-stranded primer. Taken together, a series of biochemical assays suggest that AsPol β possesses enzymatic properties similar to mammalian Pol β enzymes, as expected from sequence alignment.

3.3. Comparison of DNA Ligase III and XRCC1 genes from representatives of various taxa

The amino acid sequences of the dRP lyase domain are highly conserved among organisms shown in Fig. 1, suggesting that Pol β homologs are involved in a mechanism resembling mammalian sp-BER in these animal species. This prompted us to search for other homologs of mammalian genes that participate in the same repair pathway. Pol β , Lig3 and XRCC1 homolog genes that were deposited in the NCBI GenBank nucleotide database are listed in Table 1 (also see Figs. S4 and S5 for sequence alignment). In mammalian cells, Lig3 forms a constitutive complex with XRCC1, which mediates efficient ligation by Lig3 and binds multiple proteins involved in BER [31]. All of the three genes appear to be conserved in vertebrates and marine animals, although only partial cDNAs are found in some marine samples. This suggests that prototypic genes of the three mammalian sp-BER proteins arose in common ancestor of Cnidarians.

4. Discussion

The genomes of plants, fungi and protozoa are likely to lack two (Pol β and Lig3) or all of the three putative homolog genes for mammalian sp-BER proteins. Our recent study suggests the plant homolog of XRCC1, which lacks the N-terminal domain required for binding to Pol β (Fig. S4), may function in DNA repair mechanism different from mammalian sp-BER [32]. Comparison of the putative sp-BER proteins in three Ecdysozoans (Table 1; *C. elegans*, *D. melanogaster*, and *T. castaneum*) suggests that significant diversity of these genes in this group of phyla. The three genes are coded in the genome of the red flour beetle, while none is found in the nematode. Fruit flies contain Lig3 and XRCC1 homolog genes that share high identity with their mammalian counterparts (Figs. S4 and S5), but lack the X-family gene. We previously found that Rrp1, the Drosophila homolog of AP endonuclease 1 interacted with REV7, an accessory subunit of DNA polymerase ζ , suggesting this Pol may be involved in a lp-BER-like pathway [10]. Taken together, analysis of the three putative BER genes suggests that molecular mechanisms underlying repair of DNA base lesions such as AP sites are considerably diverse in a majority of eukaryotes.

Previously, we hypothesize that a common ancestor of many eukaryotes contained only Pol λ homolog gene. After ancestors of the X-family gene-lacking species were branched from the others, this ORF might be duplicated and evolved, resulting in emergence of Pol β . However, in the present study, we found two types of the X-family genes (Pol β and Pol λ homologs) in one of Arthropods, suggesting that actual evolution of the X-family genes may slightly differ from our previous model: an ancestor of Ecdysozoans lacking the X-family gene (i.e. *C. elegans* and *D. melanogaster*) might have two X-family genes, which were subsequently lost though evolution, although it is possible that convergent evolution might give rise to the two X-family genes in the red flour beetle (Fig. 3). Further advance in whole genome sequencing of protostomes will provide new insights into this issue. We previously anticipated that Pol μ /TdT gene was present in a subset of Cnidarians [11]; no Pol μ /TdT gene has been found in the two Cnidarians and the Arthropod, of which whole genome sequencing was completed. Therefore, this gene might arise from duplication

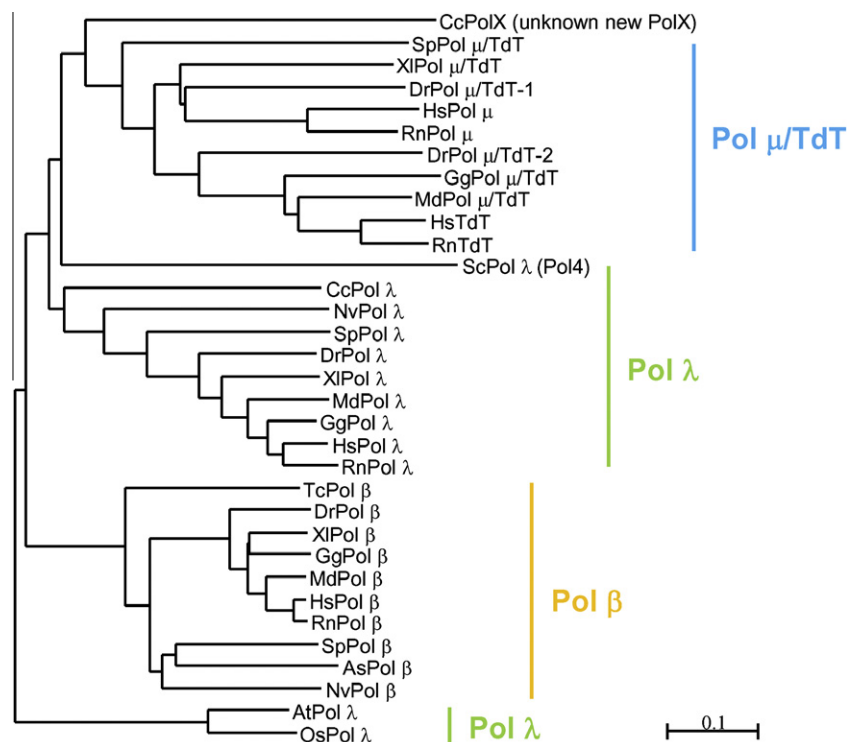


Fig. 3. Phylogenetic analysis based on the amino acids in the DNA synthesis active center of each Pol X conserved domain. For this analysis, we used the CLUSTALX program and NJplot. The bar corresponds to 0.1 amino acid substitution per residue. Hs, *Homo sapiens*; Rn, *Rattus norvegicus*; Md, *Monodelphis domestica*; Gg, *Gallus gallus*; XI, *Xenopus laevis*; Dr, *Danio rerio*; Sp, *Strongylocentrotus purpuratus*; Tc, *Tribolium castaneum*; Nv, *Nematostella vectensis*; As, *Aurelia* sp.1; Cc, *Coprinus cinereus*; Sc, *Saccharomyces cerevisiae*; Os, *Oryza sativa*; At, *Arabidopsis thaliana*. Pol λ homologs are encoded in the genomes of all eukaryotes except some protostomes. As Pol λ is the only member of the PolX family which shows such a wide distribution, it is likely that it is the most similar to the ancestor of the X-family DNA polymerases. We suggest that the Pol β diverged from a Pol λ-type ancestor before the emergence of Cnidaria and Pol μ/TdT diverged from Pol λ in Echinoderms.

of Pol λ homolog gene in an ancestor of Echinoderms including *S. purpuratus* (sea urchin) (Fig. 3).

Acknowledgments

We thank Drs. Yoshihiro Kanai and Tatsushi Ruike of Research Institute for Science & Technology, Tokyo University of Science for helpful discussions.

Appendix A. Supplementary data

Supplementary data associated with this article can be found, in the online version, at doi:10.1016/j.bbrc.2011.10.058.

References

- [1] P.M. Burgers, E.V. Koonin, E. Bruford, L. Blanco, K.C. Burtis, et al., Eukaryotic DNA polymerases: proposal for a revised nomenclature, *J. Biol. Chem.* 276 (2001) 43487–43490.
- [2] S. Kimura, Y. Uchiyama, N. Kasai, S. Namekawa, A. Saotome, et al., A novel DNA polymerase homologous to *Escherichia coli* DNA polymerase I from a higher plant, rice (*Oryza sativa* L.), *Nucleic Acids Res.* 30 (2002) 1585–1592.
- [3] F. Marini, N. Kim, A. Schuffert, R.D. Wood, POLN, a nuclear PolA family DNA polymerase homologous to the DNA cross-link sensitivity protein Mus308, *J. Biol. Chem.* 278 (2003) 32014–32019.
- [4] R. Takeuchi, S. Kimura, A. Saotome, K. Sakaguchi, Biochemical properties of a plastidial DNA polymerase of rice, *Plant Mol. Biol.* 64 (2007) 601–611.
- [5] T. Moriyama, K. Terasawa, M. Fujiwara, N. Sato, Purification and characterization of organellar DNA polymerases in the red alga *Cyanidioschyzon merolae*, *FEBS J.* 275 (2008) 2899–2918.
- [6] S. Kimura, K. Sakaguchi, DNA repair in plants, *Chem. Rev.* 106 (2006) 753–766.
- [7] U. Hubscher, G. Maga, S. Spadari, Eukaryotic DNA polymerases, *Annu. Rev. Biochem.* 71 (2002) 133–163.
- [8] Satya Prakash, Robert E. Johnson, Louise Prakash, Eukaryotic translesion synthesis dna polymerases: specificity of structure and function, *Annu. Rev. Biochem.* 74 (2005) 317–353.
- [9] A.F. Moon, M. Garcia-Diaz, V.K. Batra, W.A. Beard, K. Bebenek, et al., The X family portrait: structural insights into biological functions of X family polymerases, *DNA Repair (Amst)* 6 (2007) 1709–1725.
- [10] R. Takeuchi, T. Ruike, R. Nakamura, K. Shimanouchi, Y. Kanai, et al., *Drosophila* DNA polymerase zeta interacts with recombination repair protein 1, the *Drosophila* homologue of human abasic endonuclease 1, *J. Biol. Chem.* 281 (2006) 11577–11585.
- [11] Y. Uchiyama, R. Takeuchi, H. Kodera, K. Sakaguchi, Distribution and roles of X-family DNA polymerases in eukaryotes, *Biochimie* 91 (2009) 165–170.
- [12] Y. Uchiyama, S. Kimura, T. Yamamoto, T. Ishibashi, K. Sakaguchi, Plant DNA polymerase lambda, a DNA repair enzyme that functions in plant meristematic and meiotic tissues, *Eur. J. Biochem.* 271 (2004) 2799–2807.
- [13] A. Sakamoto, K. Iwabata, A. Koshiyama, H. Sugawara, T. Yanai, et al., Two X family DNA polymerases, lambda and mu, in meiotic tissues of the basidiomycete, *Coprinus cinereus*, *Chromosoma* 116 (2007) 545–556.
- [14] J. Yamtich, J.B. Sweasy, DNA polymerase family X: function, structure, and cellular roles, *Biochim. Biophys. Acta* 1804 (2010) 1136–1150.
- [15] B. Bertocci, A. De Smet, J.C. Weill, C.A. Reynaud, Nonoverlapping functions of DNA polymerases mu, lambda, and terminal deoxynucleotidyltransferase during immunoglobulin V(D)J recombination in vivo, *Immunity* 25 (2006) 31–41.
- [16] D. Kidane, A.S. Jonason, T.S. Gorton, I. Mihaylov, J. Pan, et al., DNA polymerase beta is critical for mouse meiotic synapsis, *EMBO J.* 29 (2010) 410–423.
- [17] J. Fan, D.M. Wilson 3rd, Nonoverlapping functions of DNA polymerases mu, lambda, and terminal deoxynucleotidyltransferase during immunoglobulin V(D)J recombination in vivo, *Immunity* 25 (2006) 31–41.
- [18] W. Schroth, A. Ender, B. Schierwater, Molecular biomarkers and adaptation to environmental stress in moon jelly (*Aurelia* spp.), *Mar. Biotechnol. (NY)* 7 (2005) 449–461.
- [19] T. Date, M. Yamaguchi, F. Hirose, Y. Nishimoto, K. Tanihara, A. Matsukage, Expression of active rat DNA polymerase beta in *Escherichia coli*, *Biochemistry* 27 (1988) 2983–2990.
- [20] H. Seto, M. Hatanaka, S. Kimura, M. Oshige, Y. Tsuya, et al., Purification and characterization of a 100 kDa DNA polymerase from cauliflower inflorescence, *Biochem. J.* 332 (1998) 557–563.
- [21] S. Murakami, S. Kamisuki, K. Takata, N. Kasai, S. Kimura, et al., Site-directed mutational analysis of structural interactions of low molecule compounds binding to the N-terminal 8 kDa domain of DNA polymerase beta, *Biochem. Biophys. Res. Commun.* 350 (2006) 7–16.
- [22] M.A. Larkin, G. Blackshields, N.P. Brown, R. Chenna, P.A. McGettigan, et al., Clustal W and Clustal X version 2.0, *Bioinformatics* 23 (2007) 2947–2948.

- [23] G. Perrière, M. Gouy, WWW-query: an on-line retrieval system for biological sequence banks, *Biochimie* 78 (1996) 364–369.
- [24] S. Richards, R.A. Gibbs, G.M. Weinstock, S.J. Brown, R. Denell, et al., The genome of the model beetle and pest *Tribolium castaneum*, *Nature* 452 (2008) 949–955.
- [25] N.H. Putnam, M. Srivastava, U. Hellsten, B. Dirks, J. Chapman, et al., Sea anemone genome reveals ancestral eumetazoan gene repertoire and genomic organization, *Science* 317 (2007) 86–94.
- [26] J.A. Chapman, E.F. Kirkness, O. Simakov, S.E. Hampson, T. Mitros, et al., The dynamic genome of Hydra, *Nature* 464 (2010) 592–596.
- [27] M.R. Sawaya, H. Pelletier, A. Kumar, S.H. Wilson, J. Kraut, Crystal structure of rat DNA polymerase beta: evidence for a common polymerase mechanism, *Science* 264 (1994) 1930–1935.
- [28] K.L. Menge, Z. Hostomsky, B.R. Nides, G.O. Hudson, S. Rahmati, et al., Structure-function analysis of the mammalian DNA polymerase beta active site: role of aspartic acid 256, arginine 254, and arginine 258 in nucleotidyl transfer, *Biochemistry* 34 (1995) 15934–15942.
- [29] F. Bollum, Terminal deoxynucleotidyl transferase, *The Enzyme* 10 (1974) 145–171.
- [30] O. Domínguez, J.F. Ruiz, T. Laín de Lera, M. García-Díaz, M.A. González, et al., DNA polymerase mu (Pol mu), homologous to TdT, could act as a DNA mutator in eukaryotic cells, *EMBO J.* 19 (2000) 1731–1742.
- [31] Y. Kubota, R.A. Nash, A. Klungland, P. Schär, D.E. Barnes, T. Lindahl, Reconstitution of DNA base excision-repair with purified human proteins: interaction between DNA polymerase beta and the XRCC1 protein, *EMBO J.* 15 (1996) 6662–6670.
- [32] Y. Uchiyama, Y. Suzuki, K. Sakaguchi, Characterization of plant XRCC1 and its interaction with proliferating cell nuclear antigen, *Planta* 227 (2008) 1233–1241.