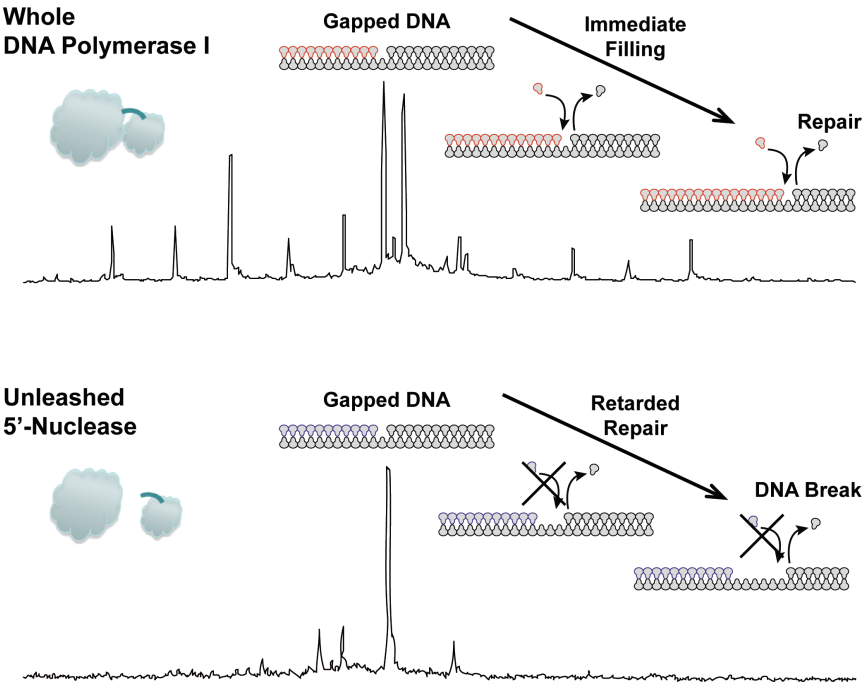
Analyst

Accepted Manuscript

ROYAL SOCIETY OF CHEMISTRY

www.rsc.org/analyst

Graphical Abstract



MALDI-TOF analysis elucidates the functions of two domains in pol I.

Analyst

RSC Publishing

ARTICLE

Mass Spectrometric Investigation of the Role of the Linking Polypeptide Chain in DNA Polymerase I

Cite this: DOI: 10.1039/xoxxoooooox

Taeho Yeom[a], Jungyoon Lee[a], SeongHyun Lee[a], Sunah Kang[a], Kyung Rok Kim[b], Byungwoo Han[b], Hyun Soo Lee[a] and Kyubong Jo[*a]

Received ooth January 2014,
Accepted ooth January 2014

DOI: 10.1039/xoxxoooooox

www.rsc.org/

DNA polymerase I offers great promise for a wide range of biotechnological applications due to its capability to add labeled nucleotides into double-stranded large DNA molecules by using both polymerase and nuclease domains. Accordingly, it is crucially important to thoroughly characterize this enzyme for further developments. Although the enzyme has been thus far characterized using mainly traditional analytical instruments, here we utilized an advanced and convenient means of mass spectrometry to elucidate enzymatic functions and mechanisms by measuring DNA oligomers generated by polymerase and nuclease reactions. Our analysis revealed several novel enzymatic features, including the observation that polymerase readily dissociates from the DNA molecule containing a wide single-stranded section. From this finding, we reasoned a serious situation of DNA break because polymerase domains cannot efficiently repair the wide single-stranded section, which is susceptible to DNA breaks. Furthermore, we deduced a plausible explanation for a paradoxical question as to why two domains of polymerase and 5′-nuclease are linked by a small and flexible polypeptide in polymerase I. The polypeptide link seems to prevent a 5′-nuclease from causing DNA breaks by locating a polymerase domain closely for immediate repair reaction. Here we present experimental evidence to prove our hypothesis *via* a set of mass spectrometric analyses as well as single DNA molecule observation and bacterial cell growth assay. Consequently, mass spectrometric analysis for DNA polymerase I provides a meaningful biological insight that a polypeptide link can be a molecular leash to control an aggressive domain in order to prevent unmanageable damages.

Introduction

DNA polymerase has become a versatile biochemical tool for a variety of bio-analytical systems. In the last decade, a number of next-generation sequencing platforms have employed DNA polymerases for high-throughput analysis, i.e., sequencing by synthesis.[1-3] Further, a single-molecule sequencing system consists of single DNA polymerases immobilized at the bottom of nano-sized wells to enable real-time sequencing of single DNA molecules.[4] A logical next advancement in this technology would be the development of a high-throughput sequencing platform to directly read sequences from individual genomic DNA molecules.[5] For this development, large double-stranded DNA molecules require nuclease activity to insert space for the addition of labeled nucleotides by polymerase. For this application, DNA polymerase I (pol I) is a strong candidate because it contains both a polymerase domain and a 5′-nuclease domain linked by a flexible polypeptide linkage. The unique function of the 5′-nuclease domain distinct from other polymerases makes pol I a useful tool for labeling DNA molecules with radioisotopes or fluorescent materials.[6] We had previously employed pol I-based nick translation for sequence-specific labeling of genomic DNA molecules.[7] In addition, we recently developed another single-molecule analytical method for the detection of ultraviolet-induced DNA damage by using pol I-based nick translation.[8] Nonetheless, although polymerases are known to be extremely accurate enzymes, stochastic random enzymatic errors occur and cause critical problems for single-molecule DNA analysis. For example, the use of single molecule sequencing is currently limited to re-sequencing analysis rather than *de novo* sequencing because of its relatively high error-rate, unlike that observed in amplification-based sequencing systems.[9] To overcome this shortcoming, it is necessary to thoroughly characterize the capabilities and the limitations of polymerases as well as their reaction mechanisms.

DNA polymerase I is a well-characterized enzyme consisting of two domains.[10] The polymerase domain, called the Klenow fragment, adds new nucleotides to DNA chains and also has proofreading capability *via* 3′→5′ exonuclease activity. The

[a] *Department of Chemistry and Interdisciplinary Program of Integrated Biotechnology, Sogang University, Seoul, 121-742, Republic of Korea,* [b] *School of Pharmacy, Seoul National University, Seoul, 151-742, Republic of Korea,*
Tel.: +82 (2) 705-8881; E- mail: jokyubong@sogang.ac.kr

other functional domain of 5′-nuclease, originally called the 5′→3′ exonuclease domain, removes damaged nucleotides as well as RNA primers.[10-12] Early pol I studies were conducted during the 1960s and 1970s with the use of classical analytical techniques such as liquid chromatography and gel electrophoresis by using radioisotope labels.[13, 14] In the 1980s and 1990s, crystallographic analysis allowed for the determination of the molecular details of the enzymatic functions of pol I.[15-17] Genetic engineering has enabled the characterization of the roles of individual amino acids by substitution with different amino acids.[18] Recently, single-molecule fluorescence resonance energy transfer (FRET) has been used to characterize molecular motions.[19, 20] Alternatively, convenient but informative mass spectrometric measurements have replaced many assays that were previously based on gel electrophoresis employing radioisotope labels.[21] The primary advantage of mass spectrometry is that radioisotope probes are not necessary. Thus, matrix-assisted laser desorption ionization time of flight (MALDI-TOF) mass spectrometry, instead of sequencing gel electrophoresis, has been utilized for the analysis of DNA oligomers generated by DNA polymerase and exonuclease.[22, 23] In a mass spectrum of the DNA oligomers, the distance between the ladder peaks allows for the determination of four nucleotides since each base has a distinct molecular mass. These ladder-like patterns have been utilized to develop a novel DNA-sequencing platform as an alternative to the Sanger method.[24-26] On the other hand, these ladder-like patterns could also be very useful in the characterization of the properties and the reaction mechanisms of pol I.

In this paper, we investigated various characteristics of pol I by using MALDI-TOF mass spectrometry. Our characterization revealed an interesting novel feature of pol I that a polymerase domain readily dissociates from the DNA duplex containing a wide-gapped single-stranded portion, unlike that in the case of a narrow-gapped portion, while the 5′-nuclease domain digests the 5′ end of DNA without differentiating upon the gap sizes. From these findings, we have uncovered an interesting biochemical insight to explain the role of the polypeptide chain connecting the 5′-nuclease domain to the polymerase domain, which has remained unanswered in studies of pol I.[27, 28] This molecular linkage likely reduces the probability of DNA breaks by an immediate polymerase reaction right after a 5′-nuclease reaction. Here we present experimental evidence obtained through a set of mass spectrometric analyses, single DNA molecule analysis, and a bacterial cell growth assay.

Results and Discussion

Mass Spectrometric Analysis of Pol I

Fig. 1 shows straightforward ladder-like mass spectrometric patterns that depict both polymerase additions and nuclease removals of nucleotides. The mass distance between the ladder peaks indicates nucleotide bases, as the four bases have distinctly different molecular masses. In addition, the peak heights indicate the relative quantities of oligonucleotides of different lengths. Although most other MALDI-TOF studies used the chain termination reaction method employing dideoxy-nucleotides (ddNTP) for monitoring polymerase reactions like the Sanger method,[29, 30] we had difficulty in using the chain termination reaction because we observed that the mass spectra showed complicated peaks from incomplete polymerized products and chain-terminated products. Instead, we used only

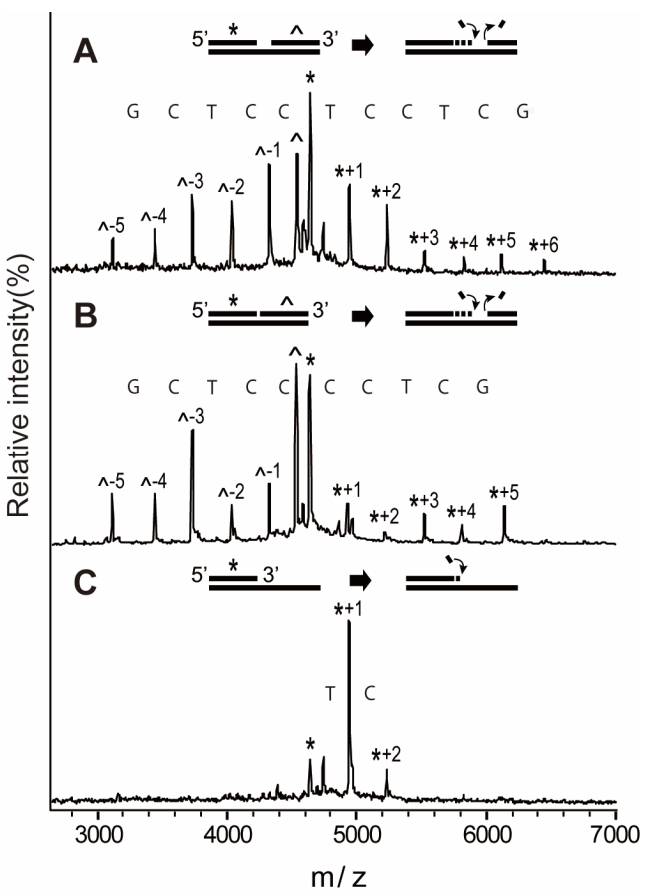


Fig. 1 Ladder-like mass spectra demonstrating pol I activity on oligonucleotide duplexes with a (A) one-nucleotide gap, (B) nick, and (C) without a downstream sequence. The upstream primers are denoted by asterisks (*) and increase with polymerase activity (*+number). In the same manner, the down stream strands are denoted by circumplexes (^) and decrease with 5′-nuclease activity (^-number). Each nucleotide is determined by the mass difference between two neighboring peaks. (See Table 1 for sequence information)

deoxy-nucleotide (dNTP) without ddNTP to generate simple ladder-like mass spectrometric patterns by optimizing the concentrations of the four dNTPs (20 μM), in concentrations similar to those observed in living bacteria; *E. coli* K-12 was reported to have 26 μM dATP, 22 μM dTTP, 35 μM dGTP, and 70 μM dCTP, though dNTP concentrations are dynamically controlled depending on various physiological conditions.[31, 32] The other reaction components consisted of DNA substrate (10 μM), and pol I (0.13 μM), and the reaction was performed at 37 °C for six minutes. From the reaction condition, it can be expected that the peaks in mass spectra are primarily the result of the processive activity of one polymerase molecule, because the number of polymerase molecules added in the solution is only 1.3% of the DNA substrates, and the peak for the unreacted substrates is relatively large. Processivity is defined as the average number of nucleotides added by a single DNA polymerase for every association and disassociation with the DNA template. Several previous studies have reported the processivity of pol I; for example, Bambara *et al.* reported that pol I had a processivity of 15~20 nucleotides, but many other studies have shown that processivity is not a fixed value.[33, 34] Das and Fujimura reported that the processivity of pol I varies from 1 to 30 nucleotides depending on the experimental conditions.[35] In addition, we observed that an increase in dNTP

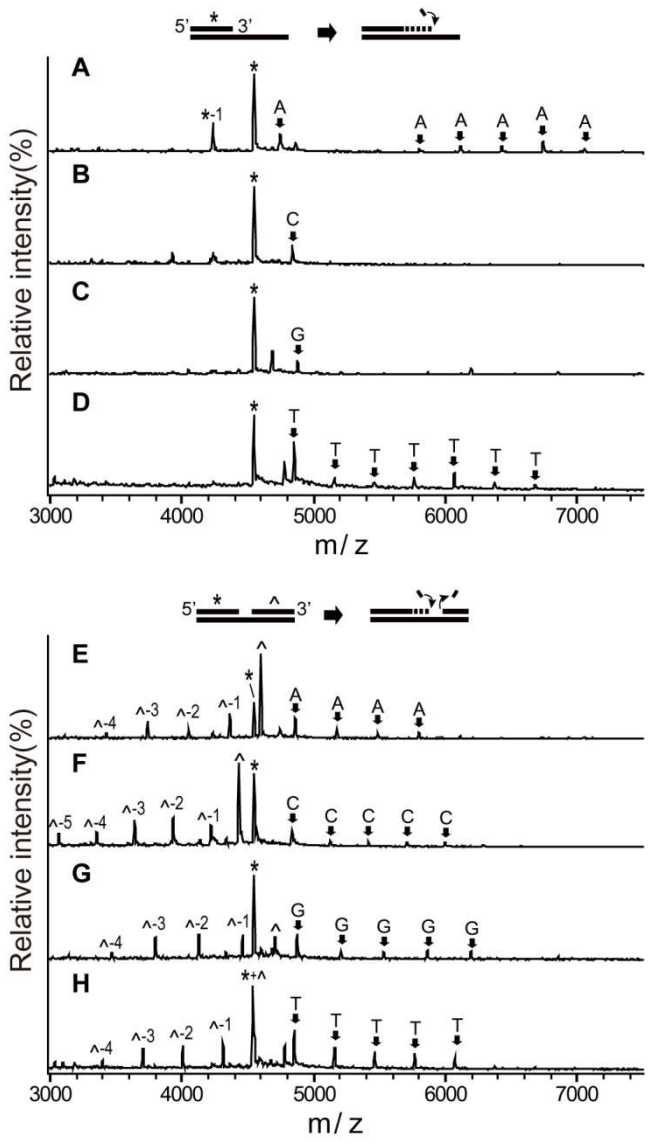


Fig.2 MALDI-TOF mass spectra demonstrating the differences in processivity on A, C, G, and T homopolymers. Pol I synthesizes (A) poly A on a poly T template, (B) poly C on a poly G template, (C) poly G on a poly C template, and (D) poly T on a poly A template. These DNA substrates do not have downstream strands. (E)-(F) can be compared with (A)-(D) respectively with the downstream strands

concentrations shifts the peaks to those reflecting a higher molecular mass, thus indicating an increase in the processivity.

Sequence Dependence of Pol I

An interesting finding of our study is that pol I cannot proceed further in the presence of single-stranded downstream DNA substrate. As shown in Fig. 1C, polymerase adds a thymine (T), but only a small portion of cytosine (C) before the polymerase reaction stops. We next prepared another template strand in which the first nucleotide to be added was cytosine (C) instead of thymine (T), and we observed that polymerase did not effectively add even the first cytosine. To obtain a more thorough understanding of sequence dependence, we utilized homopolymer templates. Fig. 2 A-D demonstrates the

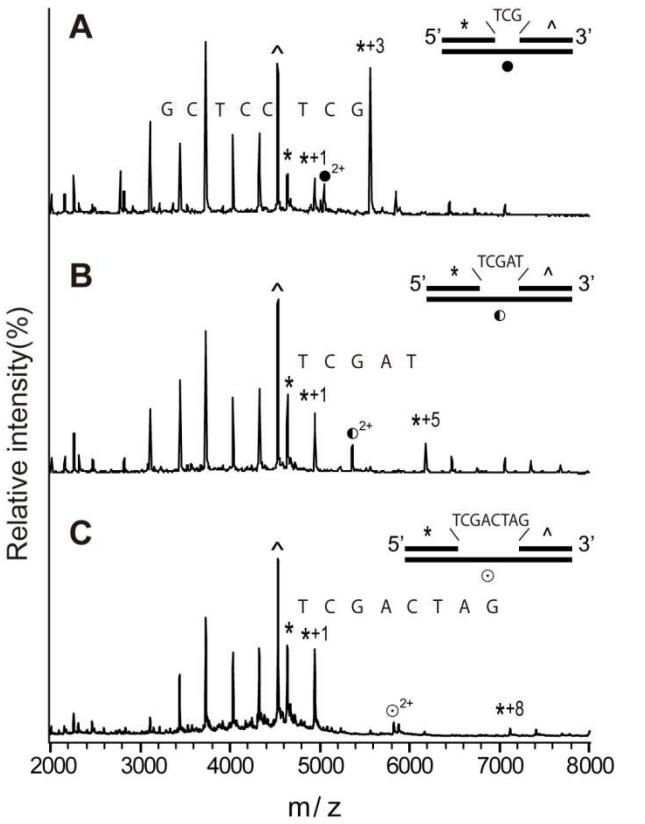


Fig. 3 MALDI-TOF mass spectra representing polymerase processivity for three different nucleotide (nt) gap sizes: (A) 3 nt gap, (B) 5 nt gap, and (C) 8 nt gap. The upstream primers are denoted by asterisks (*), and the downstream strands are denoted by circumflexes (^). Circles ($\bullet^{2+}$) show the doubly charged templates. (See Table 1 for sequence information)

sequence dependence of pol I on cytosine and guanine homopolymers templates, where the polymerase reaction halts on single-stranded downstream of poly-C and poly-G, while the polymerase reaction continues on poly-A and poly-T single-stranded templates. However, Fig. 2 E-F show contradictory results where the polymerase reactions do not halt on single nucleotide gaps with duplex downstream DNA for A, C, G, and T homopolymer templates. These data suggest that pol I is able to differentiate a gapped duplex from a single-stranded downstream region.

Pol I Dissociates from the Extended Gaps

The next question we asked is how large of a gap pol I can recognize. We prepared various gap sizes up to twelve nucleotides in lengths, of which Fig. 3 presents three representative cases for different polymerase processivities. A three-nucleotide gap (gap 3) showed the largest peak immediately after gap filling, which implies that the polymerase activity is sufficient to fill the gap (Fig. 3A). In contrast, for an eight-nucleotide gap (gap 8), the single-addition peak is dominant, and the peak after gap filling is negligible in intensity, which implies that the polymerase activity is not sufficient to fill the gap (Fig. 3C). A five-nucleotide gap (Fig. 3B) shows a mixture of both the extreme cases. This enzymatic behavior can be explained by the structure of the Klenow fragment (PDB ID: 1KLN).[16] In the crystal structure, a thumb domain and a palm domain form a cleft, and the length from the active site to the end of the cleft is about 25 Å. Therefore, pol I

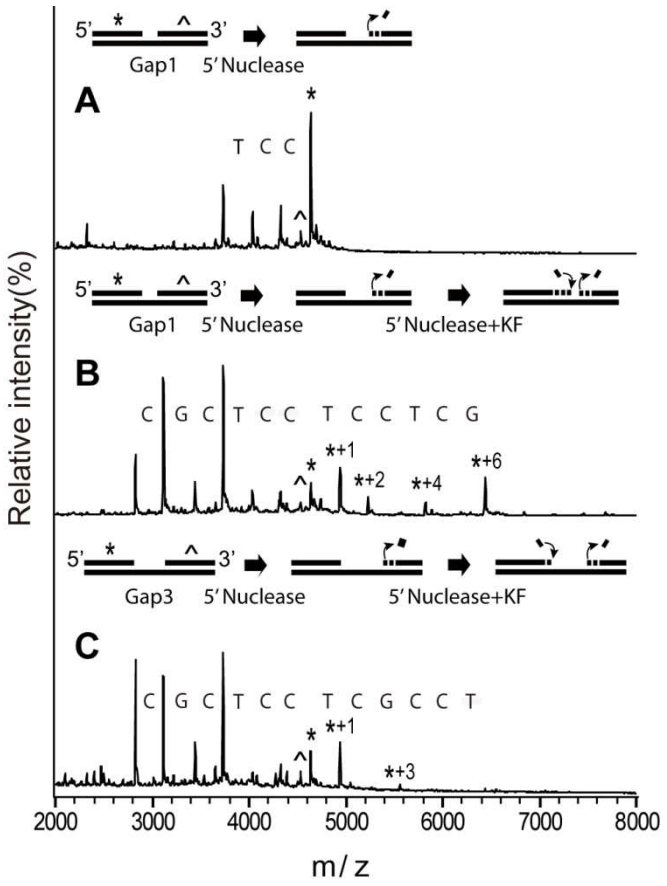


Fig. 4 (A) 5′-nuclease reaction on a one-nucleotide gap for 20 min. (B) 5′-nuclease reaction on a one-nucleotide gap for 20 min; Klenow fragment (KF) was added subsequently with an additional incubation of six minutes. (C) Three-nucleotide gap step reactions. 5′-nuclease reaction for 20 min, with KF added for an additional incubation of six minutes. There was a significant reduction in processivity on the widened gap.

can hold at most seven nucleotides because each nucleotide has a length of 3.4 Å. If a double-stranded portion existed after the single-stranded portion, it would affect the conformation of the cleft by exerting force on the thumb. Thus, the cleft can recognize three- and five-nucleotide single-stranded portions, but an eight-nucleotide gap is too long to be fit in the cleft formed by the thumb and palm domains.

Extended Single-Stranded Gap by 5′-Nuclease Domain

On the other hand, mass spectra in Fig. 1, 2, and 3 show that an individual binding of 5′-nuclease digests one to five nucleotides regardless of the gap sizes. From these data, we can envision a situation where the 5′-nuclease domain digests the same DNA multiple times. If the gap were widened due to multiple digestions, polymerase processivity would not be as efficient as that observed in the case of a small nucleotide gap. Therefore, after multiple attacks by 5′-nucleases, the DNA molecule may temporarily have a wide single-stranded portion, which is susceptible to various physicochemical attacks such as shear force, oxygen radicals, and radiation, resulting in DNA breakage. As mentioned earlier, we previously developed an approach to label large DNA molecules employing pol I-based nick translation.[8] During the development of this approach, we observed that overnight incubation of bacteriophage λ DNA (48.5 kb) with pol I generates a considerable amount of DNA fragments. Thus, we set the reaction time to less than an hour, or immobilized the DNA within an agarose gel plug for labeling reactions to prevent DNA breakage. Although multiple successive attacks may not occur frequently, even a single event can cause a large DNA molecule to break, which is particularly crucial for genomic DNA. In order to mimic this rare event of DNA breakage, we prepared 5′-nuclease by molecular cloning and performed the experiment shown in Fig. 4 in which 5′-nucleases were added as the first step to create an extended single-stranded portion before the addition of polymerase domains (Klenow fragments). As expected, the polymerase was only able to add a single thymine before dissociating from its DNA substrate due to the widened single-stranded portion generated by 5′-nuclease (Fig. 4C). Accordingly, this result confirms that multiple 5′-nuclease attacks not only increase the size of the single-stranded regions, but also reduce polymerase activity, thus increasing the lifetime of these extended single-stranded regions and the probability of occurrence of DNA breaks.

Coexistence of Two Domains in One Polypeptide

Our results suggest a plausible answer for the long-standing question of why a 5′-nuclease domain is linked to a polymerase domain. This question was initially raised by Setlow and Kornberg.[28] They reported that two proteolyticcally cleaved fragments of pol I showed an enzymatic activity similar to that of the complete pol I. Their results imply that the two linked domains may work independently without any cooperation. Therefore, they asked why two distinct enzymes, i.e. the Klenow fragment and 5′-nuclease, exist together in one polypeptide chain. Their question became more interesting with the discovery that eucaryotes and archaebacteria lack nuclease-linked polymerases.[36] Only eubacteria and some bacteriophages have polymerases linked with a 5′-nuclease, known as the pol I family. It is widely accepted that connected domains can work together at the same time. Thus, previous studies focused on the elucidation of a coordinated function of the two linked domains during DNA repair.[27] However, the polymerase and 5′-nuclease domains seem too bulky to be brought into close proximity at a damaged spot on double-stranded DNA. The size of a gap in the damaged DNA may be approximately 10 Å, but the distance between active sites of the two domains is known as 70 Å according to the crystal structure.[17] Therefore, this structural data implies that the reason to link the 5′-nuclease domain to the polymerase domain may not be for the domains to work together at the same time.

On the other hand, the coexistence of these two domains can be explained by the benefit of preserving unstable intermediates. It is known that either domain in pol I approaches a damaged DNA spot with about the same probability.[37] Thus, if one domain binds the gap and performs its reaction, the other domain finishes DNA repair immediately. The immediate reaction by the partner domain reduces the formation of 5′-nuclease induced extended single-stranded regions, unstable intermediates that are probably the primary cause for DNA breaks. The linkage of the two domains can also be explained by considering enzyme concentrations. In a bacterial cell, there are approximately 400 copies of pol I,[10] which corresponds to 0.95 μM because the total volume is 0.69 μm^3.[38] Thus, there are approximately 1 μM of polymerase and 1 μM of 5′-nuclease, respectively. In contrast, the two linked domains exist within 7 nm. If there is a particle within a sphere with a radius of 7 nm, the concentration is equivalent to 1 mM. Therefore, a covalent linkage provides a thousand-fold higher concentration of the

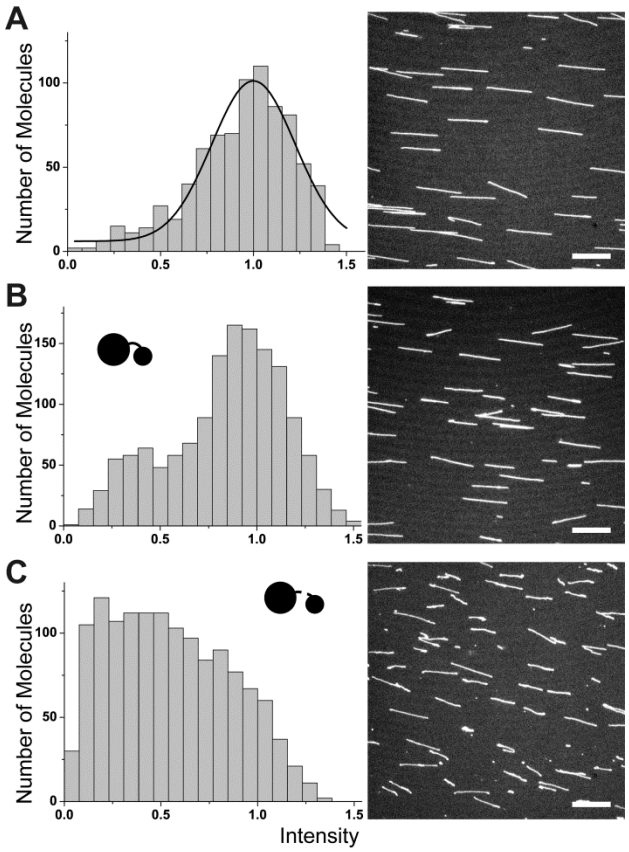


Fig. 5 Fluorescent micrographs show typical images of DNA molecules at each reaction condition. (A) Intact λ DNA molecules without any treatment as controls. The histograms of fluorescence intensity represent DNA size; intensity units are 1.0 at the center of the Gaussian curve of the intensity of the intact λ DNA molecules. (B) λ DNA molecules treated with complete polymerase I (0.26 μM) at 37 °C for 30 minutes. (C) λ DNA molecules treated with cloned 5′-nuclease (0.26 μM) and Klenow fragments (0.26 μM) at 37 °C for 30 minutes. The scale bar represents 10 μm.

domains, which determines the reaction order, of first one domain and then the other partner domain. In contrast, if the two domains were separated, the 5′-nuclease would first bind a gap and digest the DNA, and the following reaction would be performed by either a polymerase or another 5′-nuclease. In other words, a single gap site would have some probability to be digested twice by 5′-nucleases. In addition, the polymerase domain can readily dissociate from widened single-stranded gaps. Therefore, there is a greater chance for another 5′-nuclease to make the gap even wider. Therefore, we can conclude that the covalent linkage prevents DNA molecules from breaking through the immediate filling of a single-stranded gap.

Single DNA Molecule Breaks by 5′-Nucleases

Our hypothesis is that a polypeptide link is beneficial to preserve DNA from breaking. Unfortunately, the mass spectrometric analysis (Fig. 4) does not directly illustrate DNA breakage because the DNA molecules used in the experiment are only small oligonucleotides. In general, small DNA fragments are tolerant to shear force even when present as single-stranded DNA. In addition, broken DNA fragments are not readily detectable in MALDI-TOF and gel electrophoresis because the resulting fragments are of variable lengths. In

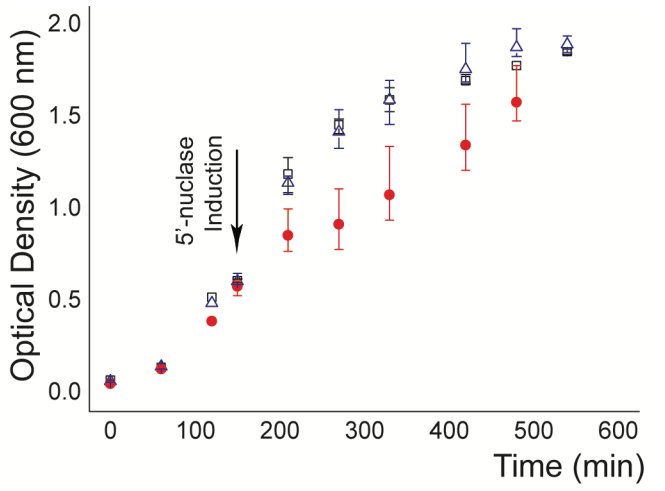


Fig. 6 Growth curves for cutres of three *E. coli* without a plasmid (□), with a plasmid expressing 5′-nuclease (●), and with a plasmid expressing pol I (△). IPTG (0.1 mM) induction selectively retarded the growth of the bacterial cells expressing 5′-nuclease without affecting cells expressing complete pol I. Error bars represent the maximum and minimal values on the means from 3–5 separate experiments.

contrast, large single DNA molecules are readily visible for monitoring DNA damage such as double and single-stranded breaks at the single-molecule level.[8] Therefore, as a next step, we performed an experiment using single DNA molecules to test our hypothesis. Fig. 5 demonstrates considerable DNA breakage over half an hour for λ DNA incubated in the presence of the separate domains. In comparison with intact λ DNA (Fig. 5A), DNA incubated with pol I exhibits some breaks (Fig. 5B), while the DNA breakage significantly increases when the DNA is incubated with separate polymerase and 5′-nuclease (Fig. 5C). Therefore, single DNA molecule analysis clearly indicates that the polypeptide linkage of pol I significantly reduces DNA degradation.

5′-Nuclease in Living Cells

We also asked if the 5′-nuclease is harmful in living *E. coli* cells. To answer this question, we monitored the bacterial growth of *E. coli* cells expressing 5′-nucleases. As a control, we also prepared transformed *E. coli* cells that express pol I by using the same plasmid. However, isopropyl β-D-1-thiogalactopyranoside (IPTG) induction is known to retard cell growths due to the production of a large amount of proteins. As expected, we observed inhibition of the growth of both cultures by typical IPTG concentration of 1.0 mM. Nonetheless, by reducing the IPTG concentration to 0.1 mM, we were able to differentiate the growth retardation effect of 5′-nuclease. Fig. 6 shows that *E. coli* cells containing a plasmid of 5′-nuclease grow more slowly than the control *E. coli* containing a plasmid for complete pol I. This data suggests that a large amount of 5′-nucleases causes harmful effects in living bacterial cells.

Conclusions

In this study, we investigated the characteristics of pol I using MALDI-TOF mass spectrometry. The ladder-like mass spectrometric peaks provide the characteristic activities of polymerase and 5′-nuclease. From these results, we have

deduced an answer to the long-standing question of why the polymerase and 5′-nuclease domains coexist in a single enzyme. Instead of focusing on elucidating cooperative interactions, our study focused on how "unleashed" 5′-nucleases cause DNA breaks. Our mass spectrometric analysis provides a detailed mechanism indicating how fragile single stranded portions of DNA are susceptible to DNA breaks. Moreover, single-molecule DNA analysis and cell growth assays provide further evidence that DNA breaks are caused by 5′-nucleases. Therefore, we conclude that the primary usage of the polypeptide link is as a molecular leash to control an aggressive 5′-nuclease to prevent unmanageable DNA breaks rather than holding the two domains in close proximity for cooperative reactions.

Experimental

Enzymes

Pol I and Klenow fragments were obtained from New England Biolabs (Billerica, MA). In order to obtain 5′-nuclease, a custom-ordered expression plasmid was constructed for the N-terminal 323 amino acid residues (969 bases of *polA* gene) with a C-terminal six-histidine tag in a bacterial strain (*E. coli* BL21(DE3)) with a plasmid (pET-15b) obtained from Cosmo Genetech (Seoul, Korea). The expressed 5′-nuclease was purified in tandem by using Ni-HiTrapTM Chelating HP 5 mL (GE Healthcare, Piscataway, NJ), HiLoad 16/600 Superdex 200 (GE Healthcare, Piscataway, NJ), and HiTrap Q HP 5 mL columns (GE Healthcare, Piscataway, NJ). The purified protein was stored in 50 mM HEPES pH 7.0, and 300 mM sodium chloride. Then, the protein concentrations of pol I, the Klenow fragment, and the 5′-nuclease were determined by performing a bicinchoninic acid (BCA) assay employing the Pierce® BCA Protein Assay Kit (Thermo SCIENTIFIC, Rockford, IL). The stock concentration of pol I was 4.06 mg/mL, Klenow fragment was 8.75 mg/mL, and purified 5′-nuclease was 0.6 mg/mL, which were determined on the basis of a standard curve constructed using known concentrations of bovine serum albumin.

Mass Spectrometric Analysis

For MALDI-TOF analysis, oligonucleotides were used as DNA substrates for polymerase and 5′-nuclease reactions. For these reactions, the oligo-(deoxyribo)-nucleotides were dissolved in a TE buffer solution (10 mM Tris-HCl; 1mM EDTA; pH 8.0) at room temperature to a concentration of 100 pmol/μL. The template and primers were mixed equivalently and annealed at 80°C for 10 min, then slowly cooled to room temperature. The DNA sequences used for the experiments are shown in Table 1. Polymerase and nuclease reactions were performed in a 30-μL total volume: 10 μM hybridized DNA duplex, 0.13 μM pol I, 50 mM NaCl, 10 mM Tris-HCl (pH 7.9), 10 mM $MgCl_2$, 1 mM DTT, and 20 μM each dNTP. The reaction was incubated at 37 °C for six minutes. After an enzymatic reaction, 3 μL of 1 M triethylammonium acetate (TEAA) solution was added to the reaction buffer. Then, oligonucleotides were purified and concentrated on a micro-reverse-phase column integrated into the outlet of a P10 pipetting tip (ZIPTIP-C_{18}; Millipore, Billerica, MA), as per the instructions of the manufacturer. The bound molecules were washed twice with 80 μL of 0.1 mM TEAA, pH 7.0 and twice with 80 μL of distilled water and were eluted with 5 μL of 50% acetonitrile in water. The purified samples were mixed with matrix solution onto a Bruker MTP 384 polished steel MALDI sample support target plate. The matrix solution was 350 mM 3-hydroxypicolinic acid (Tokyo Chemical Industry, Tokyo, Japan), 80 mM pyrazinecarboxylic acid, and diammonium hydrogen citrate in 50% acetonitrile. MALDI-TOF analysis was performed using a Bruker Autoflex™ Speed (Bruker Daltonics Inc., Billerica, MA) in the linear positive ion mode. Mass spectra were obtained as sums of 1000 laser shots at 1000 Hz.

Table 1. Oligonucleotides Used in Mass Spectrometry

Gap ∞	5′-GTACGACTGCAGGGA 3′-CATGCTGACGTCCCTAGGAGCGCTTGTACCG
Gap 1	5′-GTACGACTGCAGGGA CCTCGCGAACATGGC 3′-CATGCTGACGTCCCTAGGAGCGCTTGTACCG
Gap 3	5′-GTACGACTGCAGGGA CCTCGCGAACATGGC 3′-CATGCTGACGTCCCTAGCGGAGCGCTTGTACCG
Gap 8	5′-GTACGACTGCAGGGA CCTCGCGAACATGGC 3′-CATGCTGACGTCCCTAGCTGATCGGAGCGCTTGTACCG
Nick	5′-GTACGACTGCAGGGATCCTCGCGAACATGGC 3′-CATGCTGACGTCCCTAGGAGCGCTTGTACCG

Single Molecule DNA Break Analysis

λ DNA molecules (0.02 pM) were treated with 0.26 μM of pol I, and a mixture of Klenow fragment and 5′-nuclease at 37 °C for 30 minutes (Fig. 5). Enzyme-treated DNA molecules were stained with YOYO-1. Subsequently, 3 μL YOYO-1-stained DNA was mounted on a positively charged glass.[39] λ DNA was imaged using a fluorescence microscope (Zeiss Observer A1) illuminated by a laser (Coherent Sapphire 488, Santa Clara, CA). The fluorescence images were captured using an EMCCD camera (Evolve, Roper Scientific, Tucson, AZ) and stored in 16-bit TIFF files generated by RS Image (Roper Scientific Tucson, AZ). For image processing, a homemade ImageJ plugin was utilized in the "FIJI" software.[40]

Cell Growth Curve

For cell growth assay, another plasmid was constructed to express pol I by using the same plasmid and the same bacteria described in the previous section of Enzymes. Instead of the 5′-nuclease gene, the complete pol I gene (*polA* gene) was inserted into a plasmid (pET-15b). These bacterial cells expressing pol I and 5′ nuclease were grown in the small culture media (5 mL LB) with ampicillin (0.3 mM) at 37°C until the optical density reached approximately 0.6. Subsequently, 0.1 mM IPTG was added as an inducer. The optical densities were measured every hour (Fig. 6). As controls, the same procedure followed for *E. coli* BL21 without plasmid in the same media lacking antibiotics.

Acknowledgements

This work was supported by the National Research Foundation of Korea grant (NRF-2011-0029414), the Converging Research Center Program (2013K000338), and Leading Foreign Research Institute Recruitment Program (2013K1A4A3055268) through the National Research Foundation of Korea(NRF) funded by the Ministry of Science, ICT and Future Planning, Korea.

Notes and References

1. D. R. Bentley, S. Balasubramanian, H. P. Swerdlow, G. P. Smith, J. Milton, C. G. Brown, K. P. Hall, D. J. Evers, C. L. Barnes, H. R. Bignell, J. M. Boutell, J. Bryant, R. J. Carter, R. K. Cheetham, A. J. Cox, D. J. Ellis, M. R. Flatbush, N. A. Gormley, S. J. Humphray, L. J. Irving, M. S. Karbelashvili, S. M. Kirk, H. Li, X. H. Liu, K. S. Maisinger, L. J. Murray, B. Obradovic, T. Ost, M. L. Parkinson, M. R. Pratt, I. M. J. Rasolonjatovo, M. T. Reed, R. Rigatti, C. Rodighiero, M. T. Ross, A. Sabot, S. V. Sankar, A. Scally, G. P. Schroth, M. E. Smith, V. P. Smith, A. Spiridou, P. E. Torrance, S. S. Tzonev, E. H. Vermaas, K. Walter, X. L. Wu, L. Zhang, M. D. Alam, C. Anastasi, I. C. Aniebo, D. M. D. Bailey, I. R. Bancarz, S. Banerjee, S. G. Barbour, P. A. Baybayan, V. A. Benoit, K. F. Benson, C. Bevis, P. J. Black, A. Boodhun, J. S. Brennan, J. A. Bridgham, R. C. Brown, A. A. Brown, D. H. Buermann, A. A. Bundu, J. C. Burrows, N. P. Carter, N. Castillo, M. C. E. Catenazzi, S. Chang, R. N. Cooley, N. R. Crake, O. O. Dada, K. D. Diakoumakos, B. Dominguez-Fernandez, D. J. Earnshaw, U. C. Egbujor, D. W. Elmore, S. S. Etchin, M. R. Ewan, M. Fedurco, L. J. Fraser, K. V. F. Fajardo, W. S. Furey, D. George, K. J. Gietzen, C. P. Goddard, G. S. Golda, P. A. Granieri, D. E. Green, D. L. Gustafson, N. F. Hansen, K. Harnish, C. D. Haudenschild, N. I. Heyer, M. M. Hims, J. T. Ho, A. M. Horgan, K. Hoschler, S. Hurwitz, D. V. Ivanov, M. Q. Johnson, T. James, T. A. H. Jones, G. D. Kang, T. H. Kerelska, A. D. Kersey, I. Khrebtukova, A. P. Kindwall, Z. Kingsbury, P. I. Kokko-Gonzales, A. Kumar, M. A. Laurent, C. T. Lawley, S. E. Lee, X. Lee, A. K. Liao, J. A. Loch, M. Lok, S. J. Luo, R. M. Mammen, J. W. Martin, P. G. McCauley, P. McNitt, P. Mehta, K. W. Moon, J. W. Mullens, T. Newington, Z. M. Ning, B. L. Ng, S. M. Novo, M. J. O'Neill, M. A. Osborne, A. Osnowski, O. Ostadan, L. L. Paraschos, L. Pickering, A. C. Pike, A. C. Pike, D. C. Pinkard, D. P. Pliskin, J. Podhasky, V. J. Quijano, C. Raczy, V. H. Rae, S. R. Rawlings, A. C. Rodriguez, P. M. Roe, J. Rogers, M. C. R. Bacigalupo, N. Romanov, A. Romieu, R. K. Roth, N. J. Rourke, S. T. Ruediger, E. Rusman, R. M. Sanches-Kuiper, M. R. Schenker, J. M. Seoane, R. J. Shaw, M. K. Shiver, S. W. Short, N. L. Sizto, J. P. Sluis, M. A. Smith, J. E. S. Sohna, E. J. Spence, K. Stevens, N. Sutton, L. Szajkowski, C. L. Tregidgo, G. Turcatti, S. vandeVondele, Y. Verhovsky, S. M. Virk, S. Wakelin, G. C. Walcott, J. W. Wang, G. J. Worsley, J. Y. Yan, L. Yau, M. Zuerlein, J. Rogers, J. C. Mullikin, M. E. Hurles, N. J. McCooke, J. S. West, F. L. Oaks, P. L. Lundberg, D. Klenerman, R. Durbin and A. J. Smith, Nature, 2008, 456, 53-59.
2. M. Margulies, M. Egholm, W. E. Altman, S. Attiya, J. S. Bader, L. A. Bemben, J. Berka, M. S. Braverman, Y. J. Chen, Z. T. Chen, S. B. Dewell, L. Du, J. M. Fierro, X. V. Gomes, B. C. Godwin, W. He, S. Helgesen, C. H. Ho, G. P. Irzyk, S. C. Jando, M. L. I. Alenquer, T. P. Jarvie, K. B. Jirage, J. B. Kim, J. R. Knight, J. R. Lanza, J. H. Leamon, S. M. Lefkowitz, M. Lei, J. Li, K. L. Lohman, H. Lu, V. B. Makhijani, K. E. McDade, M. P. McKenna, E. W. Myers, E. Nickerson, J. R. Nobile, R. Plant, B. P. Puc, M. T. Ronan, G. T. Roth, G. J. Sarkis, J. F. Simons, J. W. Simpson, M. Srinivasan, K. R. Tartaro, A. Tomasz, K. A. Vogt, G. A. Volkmer, S. H. Wang, Y. Wang, M. P. Weiner, P. G. Yu, R. F. Begley and J. M. Rothberg, Nature, 2005, 437, 376-380.
3. C. W. Fuller, L. R. Middendorf, S. A. Benner, G. M. Church, T. Harris, X. H. Huang, S. B. Jovanovich, J. R. Nelson, J. A. Schloss, D. C. Schwartz and D. V. Vezenov, Nat. Biotechnol., 2009, 27, 1013-1023.
4. J. Eid, A. Fehr, J. Gray, K. Luong, J. Lyle, G. Otto, P. Peluso, D. Rank, P. Baybayan, B. Bettman, A. Bibillo, K. Bjornson, B. Chaudhuri, F. Christians, R. Cicero, S. Clark, R. Dalal, A. Dewinter, J. Dixon, M. Foquet, A. Gaertner, P. Hardenbol, C. Heiner, K. Hester, D. Holden, G. Kearns, X. X. Kong, R. Kuse, Y. Lacroix, S. Lin, P. Lundquist, C. C. Ma, P. Marks, M. Maxham, D. Murphy, I. Park, T. Pham, M. Phillips, J. Roy, R. Sebra, G. Shen, J. Sorenson, A. Tomaney, K. Travers, M. Trulson, J. Vieceli, J. Wegener, D. Wu, A. Yang, D. Zaccarin, P. Zhao, F. Zhong, J. Korlach and S. Turner, Science, 2009, 323, 133-138.
5. A. Ramanathan, E. J. Huff, C. C. Lamers, K. D. Potamousis, D. K. Forrest and D. C. Schwartz, Anal. Biochem., 2004, 330, 227-241.
6. P. W. Rigby, M. Dieckmann, C. Rhodes and P. Berg, J. Mol. Biol., 1977, 113, 237-251.
7. K. Jo, D. M. Dhingra, T. Odijk, J. J. de Pablo, M. D. Graham, R. Runnheim, D. Forrest and D. C. Schwartz, Proc. Natl. Acad. Sci. USA, 2007, 104, 2673-2678.
8. J. Lee, H. S. Park, S. Lim and K. Jo, Chem. Commun., 2013, 49, 4740-4742.
9. S. Koren, M. C. Schatz, B. P. Walenz, J. Martin, J. T. Howard, G. Ganapathy, Z. Wang, D. A. Rasko, W. R. McCombie, E. D. Jarvis and A. M. Phillippy, Nat. Biotechnol., 2012, 30, 692-+.
10. A. Kornberg and T. A. Baker, in DNA Replication W.H. Freeman and Company, 2nd edn., 1992, ch. 4, pp. 113-164.
11. V. Lyamichev, M. A. D. Brow and J. E. Dahlberg, Science, 1993, 260, 778-783.
12. H. Klenow and I. Henningsen, Proc. Natl. Acad. Sci. USA, 1970, 65, 168-175.
13. C. C. Richardson, C. L. Schildkraut, H. V. Aposhian and A. Kornberg, J. Biol. Chem., 1964, 239, 222-232.
14. D. Brutlag, M. R. Atkinson, P. Setlow and A. Kornberg, Biochem. Biophys. Res. Commun., 1969, 37, 982-&.
15. D. L. Ollis, P. Brick, R. Hamlin, N. G. Xuong and T. A. Steitz, Nature, 1985, 313, 762-766.
16. L. S. Beese, V. Derbyshire and T. A. Steitz, Science, 1993, 260, 352-355.
17. Y. Kim, S. H. Eom, J. Wang, D.-S. Lee, S. W. Suh and T. A. Steitz, Nature, 1995, 376, 612-616.
18. Y. Xu, V. Derbyshire, K. Ng, X. C. Sun, N. D. F. Grindley and C. M. Joyce, J. Mol. Biol., 1997, 268, 284-302.
19. Y. Santoso, C. M. Joyce, O. Potapova, L. Le Reste, J. Hohlbein, J. P. Torella, N. D. F. Grindley and A. N. Kapanidis, Proc. Natl. Acad. Sci. USA, 2010, 107, 715-720.
20. C. M. Joyce, Biochimica et Biophysica Acta (BBA) - Proteins and Proteomics, 2010, 1804, 1032-1040.
21. S. Park, K. Jo and H. Bin Oh, Analyst, 2011, 136, 3739-3746.
22. H. Koster, K. Tang, D. J. Fu, A. Braun, D. vandenBoom, C. L. Smith, R. J. Cotter and C. R. Cantor, Nat. Biotechnol., 1996, 14, 1123-1128.
23. U. Pieles, W. Zurcher, M. Schar and H. E. Moser, Nucleic Acids Res., 1993, 21, 3191-3196.
24. F. Kirpekar, E. Nordhoff, L. K. Larsen, K. Kristiansen, P. Roepstorff and F. Hillenkamp, Nucleic Acids Res., 1998, 26, 2554-2559.

25. J. R. Edwards, H. Ruparel and J. Y. Ju, Mutat. Res.-Fundam. Mol. Mech. Mutag., 2005, 573, 3-12.
26. C. M. Qiu, S. Kumar, J. Guo, L. Yu, W. J. Guo, S. D. Shi, J. J. Russo and J. Y. Ju, Anal. Biochem., 2012, 427, 193-201.
27. Y. Xu, N. D. F. Grindley and C. M. Joyce, J. Biol. Chem., 2000, 275, 20949-20955.
28. P. Setlow and A. Kornberg, J. Biol. Chem., 1972, 247, 232-&.
29. F. Sanger, S. Nicklen and A. R. Coulson, Proc. Natl. Acad. Sci. USA, 1977, 74, 5463-5467.
30. M. C. Fitzgerald, L. Zhu and L. M. Smith, Rapid Commun. Mass Spectrom., 1993, 7, 895-897.
31. K. Suzuki, M. Miyaki, T. Ono, H. Mori, H. Moriya and T. Kato, Mutation Research Letters, 1983, 122, 293-298.
32. S. K. Das and L. A. Loeb, Mutat. Res./DNA Repair, 1984, 131, 97-100.
33. R. A. Bambara, D. Uyemura and T. Choi, J. Biol. Chem., 1978, 253, 413-423.
34. S. W. Matson and R. A. Bambara, J. Bacteriol., 1981, 146, 275-284.
35. S. K. Das and R. K. Fujimura, J. Biol. Chem., 1979, 254, 1227-1232.
36. Y. Liu, H. I. Kao and R. A. Bambara, Annu. Rev. Biochem, 2004, 73, 589-615.
37. R. C. Lundquist and B. M. Olivera, Cell, 1982, 31, 53-60.
38. H. E. Kubitschek, J. Bacteriol., 1990, 172, 94-101.
39. E. T. Dimalanta, A. Lim, R. Runnheim, C. Lamers, C. Churas, D. K. Forrest, J. J. de Pablo, M. D. Graham, S. N. Coppersmith, S. Goldstein and D. C. Schwartz, Anal. Chem., 2004, 76, 5293-5301.
40. J. Schindelin, I. Arganda-Carreras, E. Frise, V. Kaynig, M. Longair, T. Pietzsch, S. Preibisch, C. Rueden, S. Saalfeld, B. Schmid, J. Y. Tinevez, D. J. White, V. Hartenstein, K. Eliceiri, P. Tomancak and A. Cardona, Nat. Methods, 2012, 9, 676-682.