OPEN ACCESS Freely available online

ISSN: 2471-2663
Clinical & Medical Biochemistry

Short Communication

DNA Polymerases Drive in DNA Sequencing by Synthesis Technologies

Bibillo Hout*

Department of Protein Engineering, Illumina, San Diego, California, United States

DESCRIPTION

Next Generation Sequencing (NGS) technology has revolutionized modern biomedical and biological research. The engines responsible for this innovation are DNA polymerases; they catalyze biochemical reactions for deriving information about template sequences. In fact, DNA polymerase has been the cornerstone of DNA sequencing from the very beginning. The proteolytic fragment (Klenow) of *Escherichia coli* DNA polymerase I was originally used in Sanger di-deoxy chain-terminated DNA sequencing chemistry. From these humble beginnings, there has been an explosion of organism-specific genome sequence information accessible through public databases. Family A/B DNA polymerases from mesophilic/thermophilic bacteria/archaea were modified and tested in today`s standard Capillary Electrophoresis (CE) and NGS sequencing platforms. These enzymes were selected for their efficient incorporation of bulky dye-terminator and reversible dye-terminator nucleotides respectively. Third generation, real-time single molecule sequencing platform requires slightly different enzyme properties. Enterobacterial phage ϕ29 DNA polymerase copies long stretches of DNA and possesses a unique capability to efficiently incorporate terminal phosphate-labeled nucleoside polyphosphates. Furthermore, ϕ29 enzyme has also been utilized in emerging DNA sequencing technologies including nano-pore, and protein-transistor-based sequencing. DNA polymerase is and will continue to be an important component of sequencing technology.

Since the advent of enzymatic di-deoxy-DNA sequencing by Frederic Sanger [1], DNA/RNA sequencing has become standard practice in most molecular biology research. The proliferation of Next-Generation Sequencing (NGS) technologies has further transform modern biomedical and biological research. Today, large-scale whole-genome sequencing has become routine in life science research. Although technical advances in current NGS technology have dramatically changed the way nucleic acids are sequenced, the engine ultimately responsible for these modern innovations remains unchanged. Like Sanger sequencing, today's NGS technologies, with the exception of oligonucleotide-based ligation sequencing [2], still require DNA polymerase to carry out the biochemical reaction to replicate the sequence information of the substance. This unique polymerase-dependent sequencing method is often referred to as DNA Sequencing by Synthesis (SBS), because the subsequent sequencing reaction concurrently generates a newly synthesized strand of DNA.

However, unlike Sanger sequencing, DNA polymerases utilized in NGS technologies are more diverse and tailor made. The Klenow enzyme, a proteolytic fragment of *Escherichia coli* DNA polymerase I was originally utilized in Sanger`s di-deoxy chain terminating DNA sequencing chemistry. This enzyme was chosen for its efficient incorporation of 2′, 3′dideoxynucleotides (ddNTPs) that leads to chain termination of DNA synthesis [3]. From this humble beginning, followed by a robust sequencing chemistry improvement, the nucleotide substrates used for DNA sequencing became larger and bulkier. First, four fluorescent dyes with distinct, non-overlapping optical spectra were attached to both purine or pyrimidine bases, respectively and even the terminal gamma phosphate of four (A, T, C, and G) nucleotides for the ease of signal detection. Second, the 3′ hydroxyl groups on deoxyribose of four nucleotides were replaced with a larger, cleavable chemical group used to reversibly terminate DNA synthesis. As a result, the Klenow enzyme no longer efficiently incorporated these newly modified nucleotides. DNA polymerases with different enzymatic properties were required for improving the nucleotide incorporation reactions. Fortunately, the adoption of NGS sequencing in life science research allowed a rapid expansion of organism specific, genome sequence information accessible *via* public database. Various types of DNA polymerases from archaea viruses, bacteria, and thermophiles have been discovered and later screened for efficient incorporation of modified nucleotides into new DNA sequencing methods. A pool of new advantageous DNA polymerases from a wide variety of microorganisms has been selected and served as a protein backbone for further improvement through protein engineering or directed enzymatic evolution [4]. Evolutionary DNA polymerases with improved

Correspondence to: Bibillo Hout, Department of Protein Engineering, Illumina, San Diego, California, United States, E-mail: Bhout233@illumina.com

Received: 01-Apr-2022, Manuscript No. CMBO-22-16605; **Editor assigned:** 06-Apr-2022, PreQC No. CMBO-22-16605 (PQ); **Reviewed:** 27-Apr-2022, QC No CMBO-22-16605; **Revised:** 04-May-2022, Manuscript No. CMBO-22-16605 (R); **Published:** 11-May-2022, DOI: 10.35841/2471-2663.22.8.126.

Citation: Hout B (2022) DNA Polymerases Drive in DNA Sequencing by Synthesis Technologies. Clin Med Bio Chem. 8:126.

biochemical performance were ultimately used for every unique sequencing technology.

In contrast to current SBS approaches, emerging DNA sequencing methods rely on unconventional applications of DNA polymerase. These techniques use DNA polymerase as a traditional incorporating enzyme, and alternatively as a molecular motor, responsible for the controlled translocation of DNA across the protein nanopore. Traditional, nanopore-based, SBS uses the commercial DNA polymerase Therminator γ, a 9°N variant of DNA pol, to incorporate terminal, γ-phosphate-labeled nucleoside tetraphosphates. These modified nucleotides are paired with four PEG-coumarin tags of different lengths corresponding to bases A, T, C and G [5]. The DNA sequence information can be ascertained by measuring the current (amps) fluctuations of the PEG-coumarin tags orderly released through the α-hemolysin nano-pore following DNA polymerase incorporation. A related, but fundamentally different, approach involves mutant Porin A (MspA) nanopores of Mycobacterium smegmatis, ϕ29 DNA polymerase, and natural dNTPs [6]. In this approach, the enzyme functions as both a DNA replicative enzyme and a molecular motor, controlling the speed of DNA translocation through the MspA nanopore.

Besides the nano-pore-based sequencing approach, a transistor-based protein sequencing method that utilizes the conductivity measurement of the ϕ29 DNA polymerase reaction has been reported [7]. Unfortunately, this study is currently being challenged, and the merits of this particular approach must be reevaluated.

CONCLUSION

Since Frederic Sanger introduced the first method of DNA sequencing in the mid-1970s, decades of scientific research on different DNA polymerases, starting with Arthur Kornberg's discovery of the enzyme in the mid-1950s, has provided the possibility to understand how these enzymes work which further strengthens the foundation for improved enzyme properties and applications in current and future DNA polymerase-based sequencing technologies. Large-scale organism-specific genomics study reveals the intrinsic diversity and unique characteristics of DNA polymerases found in all kingdoms of life, including their viruses. Diverse DNA polymerases with distinct functions and properties provide a large variety of natural protein variants that can be tested and subsequently used for continuously evolving sequencing chemistry. Tailor-made protein variants engineered through protein engineering or enzymatic evolution have created the powerful protein engines that have propelled the progression of DNA sequencing technology over the past few decades. Without a doubt, DNA polymerase was, is and will remain an important part of future sequencing technologies.

REFERENCES

1. Sanger F, Nicklen S, Coulson AR. DNA sequencing with chain-terminating inhibitors. Proc Natl Acad Sci. 1977;74:5463-5467.
2. Drmanac R, Sparks AB, Callow MJ, Halpern AL, Burns NL, Kermani BG, et al. Human genome sequencing using unchained base reads on self-assembling DNA nanoarrays. Science. 2010;327:78-81.
3. Atkinson MR, Deutscher MP, Kornberg A, Russell AF, Moffatt JG. Enzymatic synthesis of deoxyribonucleic acid. XXXIV. Termination of chain growth by a 2′,3′-dideoxyribonucleotide. Biochemistry. 1969;8:4897-4904.
4. Patel PH, Loeb LA. Getting a grip on how DNA polymerases function. Nat Struct Biol. 2001;8:656-659.
5. Kumar S, Tao C, Chien M, Hellner B, Balijepalli A, Robertson JW, et al. PEG-labeled nucleotides and nanopore detection for single molecule DNA sequencing by synthesis. Sci Rep. 2012;2:684.
6. Manrao EA, Derrington IM, Laszlo AH, Langford KW, Hopper MK, Gillgren N, et al. Reading DNA at single-nucleotide resolution with a mutant MspA nanopore and phi29 DNA polymerase. Nat Biotechnol. 2012;30:349-353.
7. Chen YS, Lee CH, Hung MY, Pan HA, Chiou JC, Huang GS. DNA sequencing using electrical conductance measurements of a DNA polymerase. Nat Nanotechnol. 2013b;8:452-458.