



Cloning of DNA Polymerase I *Geobacillus thermoleovorans* SGAir0734 from a Batu Kuwung Hot Spring in *Escherichia coli*

Kenny Lischer^{1*}, Kevin Priyono Tansil¹, Mikael Januardi Ginting¹, Muhamad Sahlan¹, Anondho Wijanarko^{1**}, Masafumi Yohda²

¹Department of Chemical Engineering, Faculty of Engineering, Universitas Indonesia, Kampus UI Depok, Depok 16424, Indonesia

²Department of Biotechnology and Life Science, Tokyo University of Agriculture and Technology, 2-24-16 Naka-cho, Koganei-city, Tokyo 184-8588, Japan

Abstract. Access to biological engineering has become a critical point of modern science development through polymerase chain reaction (PCR). One of the main components in this process is DNA polymerase, which copies the main template DNA. However, there is a lack of studies on the production of DNA polymerase from indigenous thermophilic bacteria in Indonesia. To examine this process, DNA polymerase I gene (DNA pol I) from *Geobacillus thermoleovorans* (isolated from Batu Kuwung, Banten, Indonesia) was transformed into *Escherichia coli*. The gene was cloned by the cut and ligation method using NcoI and BamHI restriction enzymes, which were ligated with a pET23d vector. The recombinant gene was overexpressed in *E. coli* **BL21** and identified by using SDS-PAGE of 10% acrylamide gel, showing that the protein molecular weight was approximately **100 kDa**. This study successfully amplified the gene of interest, indicated by a high local similarity between the sequencing results and theoretical gene and positive intercellular protein expression. The results confirm that the study successfully cloned and synthesized recombinant DNA pol I of *Geobacillus thermoleovorans* from Batu Kuwung, Serang, Banten.

Keywords: DNA pol I; DNA polymerase; *Escherichia coli*; *Geobacillus thermoleovorans* SGAir0734; Thermophilic bacteria

1. Introduction

Polymerase chain reaction (PCR) is a widely known method to amplify a DNA strain and its complement. The PCR products are critical for biological engineering developments and breakthroughs. Many things can be achieved through PCR, including gene amplification and detection, disease detection, and the creation of genetically modified organisms (GMO) (Beyer et al., 2002). One of the components required for PCR reaction is DNA polymerase, also known as the main enzyme (Valones et al., 2009). Without DNA polymerase, the extension process as the main stage of PCR will not occur (Maddocks and Jenkins, 2017). In Indonesia, despite having access to PCR technology, almost all DNA polymerase is still imported.

DNA polymerase in both current global and local markets mostly derives from thermophilic bacteria due to its higher resistance to denaturation and aggregation at high

*Corresponding author's email: lischer.kenny@ui.ac.id, Tel.: +62-87888017811; Fax: +60-04 979 8636

**Corresponding author's email: anondho@che.ui.ac.id, Tel.: +60-04 979 8626; Fax: +60-04 979 8636

doi: [10.14716/ijtech.v11i5.4311](https://doi.org/10.14716/ijtech.v11i5.4311)

temperatures during conventional PCR (Maddocks and Jenkins, 2017; Gomes and Faísca, 2019). There are many papers regarding the finding that one of the first revolutionary thermophilic DNA polymerases comes from *Thermus aquaticus* (Engelke et al., 1990). There also other thermophilic bacterial-origin DNA polymerases such as *Pyrococcus furiosus* (Zheng et al., 2016), *Bacillus stearothermophilus* (Aliotta et al., 1996), or *Geobacillus thermoleovorans* species but from a different strain of ATCC 12980 (Li et al., 2017).

Thermophilic bacteria can be found in hot springs. Indonesia as a country is surrounded by the Pacific Ring of Fire and has many hot springs across the country. Therefore, the bioprospecting potential of indigenous thermophilic bacteria is very high. Although it seems promising, exploration of DNA polymerase from indigenous thermophilic bacteria in Indonesia remains low as there are still many thermophilic bacteria habitats that have no biodiversity studies published yet even the sites are considered public venues. So far, DNA polymerase I gene (DNA pol I) from *Geobacillus thermoleovorans* (from South Bandung hot springs) has been identified (Akhmaloka et al., 2008). There are still many unexplored thermophilic bacteria from different hot springs across the country. One of these sites is the Batu Kuwung hot spring in Banten. Neither DNA pol I nor thermophilic bacteria have been explored in this place to date. Therefore, the aim of this research is to clone the gene of DNA pol I from thermophilic bacteria in the Batu Kuwung hot spring. This research hypothesizes that thermophilic bacteria can be isolated from the hot spring and be used to clone its DNA pol I gene to *E. coli*. This research is the initial step for DNA pol I production from indigenous thermophilic bacteria from the Batu Kuwung hot spring in future research.

2. Methods

2.1. Template DNA and Plasmid

The template DNA was acquired from isolated *Geobacillus thermoleovorans* in the Batu Kuwung hot spring, Serang, Banten, which is highly similar to the *Geobacillus thermoleovorans* strain SGAir0734. Genome isolation was prepared and done by the Department of Chemical Engineering at Universitas Indonesia and the Department of Biotechnology and Life Science at the Tokyo University of Agriculture and Technology (data not shown). The plasmid used as a vector in this study is pET23d, which was provided by the Department of Biotechnology and Life Science at the Tokyo University of Agriculture and Technology.

2.2. Primer Preparation

The primers in the study consist of a forward primer, 5'-AAACCATGGGATTGAAAAAAAAAGCTTGTT-3', and a reverse primer, 5'-AAAGGATCCCGTTTCGCGTCATACCACGTCGAGCC-3'. The primers are determined based on local alignment between the novel native strain and the *Geobacillus thermoleovorans* strain SGAir0734 using BLAST (Basic Local Alignment Search Tool), as this strain is the most similar to the template DNA. The forward primer is designed to integrate the start codon into the NcoI restriction site. As for the reverse protein, the primer includes 2 redundant nucleotides just before the restriction site to shift the reading frame and therefore translate the His-tag sequence from the vector. The same primers are also used in sequencing.

2.3. Growth Condition

The transformed *E. coli* strains in both *DH5 α* and *BL21* are cultured in LB (Luria-Bertani) and 0.1% ampicillin in either broth or a 1% agar plate medium. For aeration, the incubator is set to 161 rpm for 20 hours (Hermansyah et al., 2018).

2.4. Gene Amplification by Polymerase Chain Reaction

The PCR reaction compositions were 1.5 μl for both 10 μM forward and reverse primers, 1 μl of purified template DNA, 8.5 μl of $d\text{H}_2\text{O}$, and 12.5 μl of the PCR mix 2X PrimeSTAR Max. For the temperature settings, the thermocycler was set for 32 cycles to the initial denaturation temperature of 98°C for 1 minute and 10 seconds per cycle. For annealing, the temperature was set to 55°C for 5 seconds, and for the extension phase, it was set to 72°C for 15 seconds. Lastly, for a final extension, it was set to 72°C for 5 minutes before keeping the mixture at 4°C indefinitely. After the PCR, the PCR product was checked with agarose gel electrophoresis consisting of 2 μl of the PCR product combined with 0.2 μl of Midori Green Direct Dye to stain the PCR product. The marker used in the study was 5 μl of λ marker stained with 0.5 μl of the same dye. The stained DNA was loaded in 1% agarose gel and run at the setting of 135 V for 20 minutes. After determining the success of the PCR according to the appropriate length of the amplified DNA fragment in the PCR product, the next step was to take out the non-DNA component from the PCR product by using a Monarch DNA PCR Cleanup (5 μg) Kit. This process aims to bind the amplicons, wash out PCR components, such as DNA polymerase and buffer, and finally elute the previously bound amplicons. For elution, 15 μl of $d\text{H}_2\text{O}$ was used.

2.5. Making the Recombinant Plasmid

The making of recombinant plasmid process consists of cutting the fragments with restriction enzymes. The mixture consisted of 5 μl PCR product, 0.5 μl of 2000 U/ml of both NcoI-HF and BamHI restriction enzymes, 2 μl of 10X CutSmart Buffer, and 12 μl of distilled water. The mixture was then incubated at 37°C for 2 hours. For the vector, the mixture consisted of 10 μl vector, 0.5 μl of 2000 U/ml of both NcoI-HF and BamHI restriction enzymes, 1 μl of 10X CutSmart Buffer, and 3 μl of distilled water. The mixture was then incubated at 37°C for 16 hours to ensure complete digestion. 10 μl of restricted inserts was then combined with 1 μl of Midori Green Direct Dye and loaded in 1% agarose gel. The apparatus was run at a setting of 100 V for 30 minutes. The process to acquire a specific length of the fragment started with cutting the band from the gel and purifying it with kit Monarch Gel Extraction Kit. For elution, 10 μl of $d\text{H}_2\text{O}$ was used.

The process involves the addition of 15 μl of restricted vector, 1.5 μl of 3M sodium acetate, and 45 μl of 0 – 4°C 100% ethanol. The mixture was vortexed and then precipitated at a temperature of –80°C for 1 hour. The mixture was then centrifuged for 40 minutes in a setting of 4°C and 15000 rpm and then decanted. The mixture was washed twice using 0.5 ml of 0 – 4°C 75% ethanol and centrifuged at the same setting for 10 minutes. After the centrifugation, the pellet was air-dried and diluted in 14 μl of $d\text{H}_2\text{O}$.

To ensure the success of the restriction process, the vector was checked with electrophoresis agarose gel with the same composition of 0.2 μl Midori Green Direct Dye in 2 μl samples and 0.5 μl of stain in 5 μl of λ marker loaded in 1% agarose gel and run at a setting of 135 V for 30 minutes. The ligation composition was determined by using Cranenburgh's calculation with a target molar ratio of insert to vector of 7:1 (Cranenburgh, 2004). The process consists of mixing 1.76 μl per 17.3 $\mu\text{g}/\mu\text{l}$ insert and 1 μl of 6 $\mu\text{g}/\mu\text{l}$ plasmid with the addition of 3 μl of 2X Mighty Max ligation mix and 0.24 μl of distilled water, yielding 6 μl of the total volume mix. The mix was then incubated on a cool pad at 16°C for 30 minutes.

2.6. Amplifying of Recombinant Plasmid

The next step was to transform the recombinant plasmid to amplify its number by thawing both competent cells of *E. coli DH5 α* and the recombinant plasmid on ice. The volume ratio of competent cells to the plasmid was 100:1, which resulted in 0.5 μl of

plasmid being added to 50 μl of competent cells and incubated for 30 minutes on ice. The cell mixture then underwent a water bath at 42°C for 45 seconds. After the water bath, the cell was then incubated for 2 minutes on ice before being combined with 100 μl SOC (Super Optimal Broth) medium and incubated for an additional 45 minutes before spread on ampicillin-agar and left overnight. From the ampicillin agar, 4 random colonies were picked and diluted in 20 μl $d\text{H}_2\text{O}$. The next step was to make a PCR mix consisting of 2 μl of colony suspension, 0.5 μl of both 10 μM forward and reverse primers, 7 μl of $d\text{H}_2\text{O}$, and 10 μl of 2X Pfu PCR Mix. The thermocycler was set for a denaturation temperature of 94°C for 5 minutes, followed by a cycle of 30 seconds. For annealing, the temperature was set to 55°C for 30 seconds. For the extension setting, the temperature was set to 72°C for 6 minutes and the last cycle was set for 7 minutes. The cycles were set to repeat 32 times. 10 μl of the PCR product was combined with 1 μl of stain and loaded into 1% agarose gel and electrophoresis was run at the high setting for 30 minutes.

The corresponding colony suspension was moved to be cultured in LB broth. The culture consisted of a single colony in 20 ml and 20 μl of ampicillin. The colony was then incubated in a shaker incubator at 37°C with a speed of 161 rpm overnight. The 20 ml culture was then split into 5 parts, 4 ml each. Every tube was centrifuged at a speed of 11000 g for 1 minute and was then decanted. The purification process was completed using the NucleoSpin Transfection Plasmid Kit. Each part was separately lysed but bound in a single binding column to acquire a high concentration of plasmid. For elution, 54 μl of 50°C 1X TAE (Tris-acetate-EDTA) buffer was used.

2.7. Recombinant Plasmid Identification

After the purification, the acquired recombinant plasmid's concentration was determined using a nanodrop apparatus. 2 μl of TAE buffer was used as a blank sample and another 2 μl of recombinant plasmid as the main sample. The plasmid was identified with a concentration of 166.2 $\text{ng}/\mu\text{l}$. The plasmid was also digested by the same restriction enzymes that cut the restriction areas from the PCR product to confirm the plasmid's integration with the target insert. 0.5 μl of the plasmid was combined with 0.1 μl NcoI-HF, 0.1 μl BamHI, 10X CutSmart buffer, and 3.8 μl $d\text{H}_2\text{O}$ to obtain 10 μl mixture. The mixture was then incubated at 37 °C for 2 days. For sequencing, 5.44 μl plasmid was combined with 12.0 μl of either 10 μM forward or reverse primer and 3.56 μl $d\text{H}_2\text{O}$, yielding a final volume of 21 μl .

2.8. Protein Test Expression

The next step was the transformation to *E. coli* BL21, which is conducted the same way as the previous *E. coli* DH5 α transformation but with a slight difference. The process consists of thawing competent cells of both *E. coli* BL21 and recombinant plasmid on ice. The composition of the recombinant plasmid and competent cells was 0.7 μl plasmid and then was added to 50 μl of competent cells and incubated for 30 minutes on ice. The cell mixture then underwent a water bath at 42°C for 45 seconds. After the water bath, the cell was then incubated for 2 minutes on ice before being combined with 100 μl of SOC medium and incubated for an additional 10 minutes. Then, it was spread on ampicillin-agar and a culture at 37°C and left overnight.

A colony from the plate was then cultured in 4 ml LB broth added to 4 μl ampicillin. The culture was incubated in an incubator at 37°C and a mixing speed of 150 rpm overnight. The mixture was separated through centrifugation at 11,000 g for 1 minute and then the pellet was decanted. The pellet was diluted in 1.5 ml of 1X PBS and sonicated with the setting of output 0.7 V , duty 50, and an interval duration of 2 minutes for 3 times. The mixture was centrifuged at the speed of 11,000 g at 4°C for 10 minutes to separate the

pellet and supernatant. 15 μl of the supernatant was combined with 15 μl of the SDS sample buffer and heated to 99 $^{\circ}\text{C}$ for 5 minutes. 15 μl of the solution was loaded into 10% acrylamide gel and run at the settings of 20 mA for stacking gel and 35 mA for running gel. The gel was stained in Coomassie Brilliant Blue for 30 minutes before being decolorized in water for 3 hours, with a water change every hour.

3. Results and Discussion

3.1. PCR Product

Geobacillus thermoleovorans SGAir0734 was successfully isolated from the Batu Kuwung hot springs (data not shown). To confirm whether the isolated bacteria contain the DNA pol I gene, a PCR was used. The PCR product after being electrophorized and observed under UV light is shown in Figure 1. According to the theoretical length, the amplified gene is 2637 bp long. Accordingly, lane number 3 shows a band with the desired size. There are 2 other dim bands, though these may indicate mispriming due to a slightly high concentration of primers used in the study to achieve relatively easy and distinguishable products (see lane 3 in Figure 1). The primers were likely to partially annealed in similar complement sequences and unintentionally amplified. Fortunately, due to lower affinity, the fragments were not exponentially amplified in addition to the main product. However, the results show that the PCR procedures effectively amplified the desired gene.

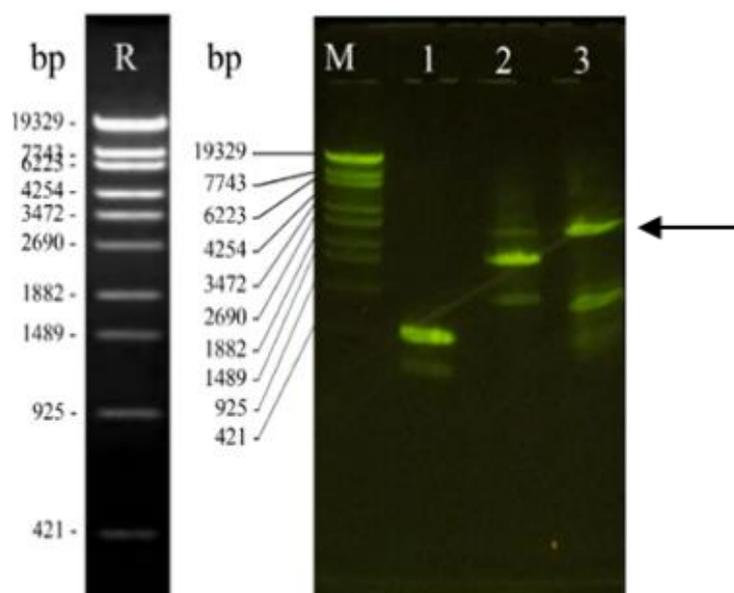


Figure 1 Amplified PCR product. (R) shows the reference of the $\lambda/\text{Sty1}$ marker; (M) shows the $\lambda/\text{Sty1}$ marker; (1) and (2) shows the positive control gene; and (3) shows the amplified genes. The gene of interest (**2637 bp**) is indicated by the arrow

3.2. Colony PCR

Figure 2 shows the results of the colony PCR on transformed colonies to select which colony had the right transformed plasmid. It can be seen that there is only one colony suspension that has the proper gene in lane number 4, as it confirms that the colony possesses plasmid that is integrated with the proper insert. The amplified gene in the gel also confirms successful ligation despite its relatively low ligation success percentage due to the large size of the insert to vector ratio (72.58%). The band also has the right length of the target DNA pol I gene, which is theoretically around 2637 bp ; the gel shows the band to be 2500 bp to 3000 bp .

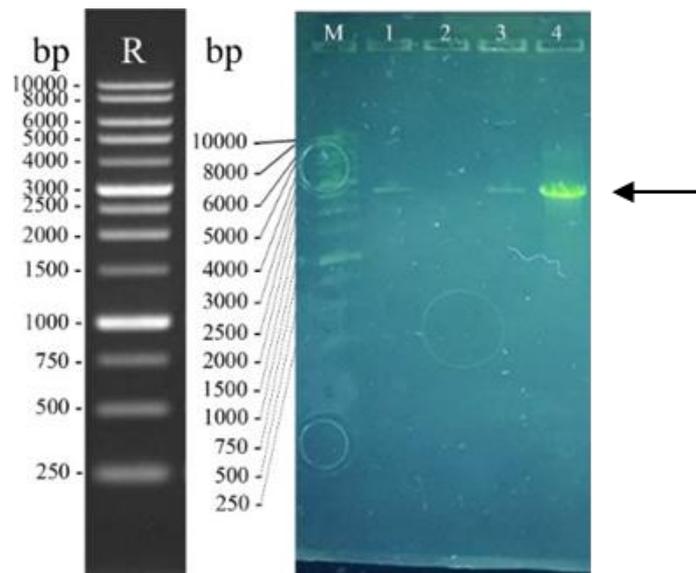


Figure 2 Colony PCR. (R) shows the reference 1 kb ladder; (M) shows the used 1 kb ladder marker; (1), (2), (3), and (4) show random-picked transformed colonies from the agar plate; and the arrow to (4) indicates the gene of interest (**2637 bp**)

3.3. Recombinant Plasmid Digestion

The (2) result showed in Figure 3, consists of 3 bands that can be identified in succession as a recombinant plasmid, empty vector, and gene of interest. This visual also confirms that the last two bands come from the same component due to their equal luminosity. There is a slight difference, though, between the proper position of the bands and the theoretical data. Fortunately, if all the bands in lane number 3 in Figure 3 are slightly shifted above together, the bands will match their theoretical lengths of around 6270 bp for the recombinant plasmid, 3633 bp for the cut vector, and 2637 bp for the gene of interest.

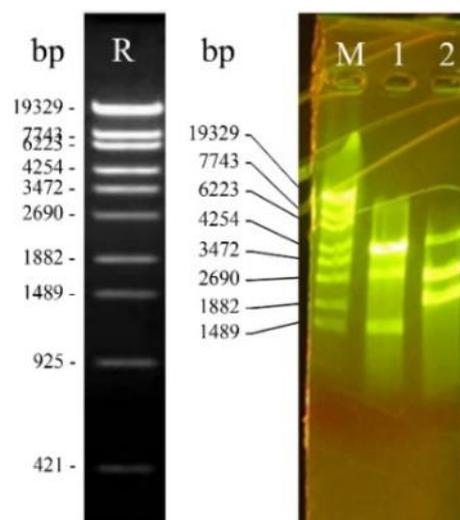


Figure 3 Digested Recombinant Plasmid. (R) shows the reference of λ /**Sty1** marker; (M) shows the λ /**Sty1** marker; (1) shows the positive control gene; and (2) shows the digested recombinant plasmid

A few factors may trigger this condition, such as slightly uneven overheated gel in the electrophoresis process causing the sample bands to migrate between columns slightly faster or a difference in solvent composition. This reading error happens due to different local ion concentrations in the sample; different columns experience different electromagnetic drive force, leading to different migration rates of the samples.

There are 2 bands instead of 3 on the gel. This shows that the digestion process was not completed despite the 2-day-long digestion process, which should have been more than enough time to fully digest all the recombinant plasmid into its precursor vector and the gene of interest. The situation is most likely caused by salt inhibition in addition to a very high concentration of the substrate. Since the total volume used in digestion a very small amount (5 μ l), a relatively significant amount of water was evaporated as a solvent from the mix and reduced the total volume, as several small condensed droplets around the main mix were visible after the digestion. This amount is relatively minute but can significantly raise the overall components' concentrations. Once a certain point of salt concentration was reached, the enzymes were then inactivated before they could digest all the recombinant plasmid in the solution.

3.4. Recombinant Plasmid Protein Sequencing

To further confirm whether DNA pol I was successfully cloned, sequencing analysis was conducted to obtain the gene sequence from plasmid. The obtained sequence was analyzed by comparing the amino acid sequence with existing data (whole genome sequences done by Tokyo University of Agriculture and Technology). As shown in Figure 4., the sequences are 72.2% globally aligned identical. The comparison shows that the middle region of the recombinant gene sequence remained empty due to the sequencing capability reaching up to roughly 1000 *bp* while the original gene had 2637 *bp* and the sequencing only utilized the same primers as in the first step of the research to amplify the gene. The missing region causes a lack of insight to the whole genome and amino acid sequences, as an additional primer is required to fully sequence the recombinant gene and avoid mistranslation due to the possibility of frameshifting in reading. Although it is not fully sequenced, the results confirm that DNA pol I was successfully cloned.

3.5. Recombinant Protein Expression Test

After DNA pol I was successfully cloned, the next step was to determine the expression of recombinant DNA pol I in *E. coli*. To do so, DNA pol I from *E. coli* was expressed and analyzed by SDS-PAGE (sodium-dodecyl-sulfate polyacrylamide gel electrophoresis). The result shows proteins that are heavily expressed, with a molecular weight of around 100 *kDa*. This is consistent with the theoretical molecular weight of the gene, which is 99.65 *kDa*. This finding confirms that there is a high chance of the existence of recombinant plasmid in the solution. As for the many proteins in the mix, the process of obtaining the proteins was solely through sonication without any other addition purification method. This method explains why there are still many other proteins other than the desired recombinant protein of DNA Pol I exist in the sample. According to the theoretical data, the recombinant DNA Pol I roughly have a molecular weight of 100 *kDa*. As there is a significant tint on 100 *kDa* reading, it highly indicates that the recombinant protein was overexpressed.

Full	1	-MRLKKKLVLDGSSVAYRAFFALPLLHNDKGIHTNAVYGFMMLNKILA	49
Full_Seq_Merg	1	YGLK KKLVLDGSSVAYRAFFALPLLHNDKGIHTNAVYGFMMLNKILA	50
Full	50	EEEPHMLVAFDAGKTIIFRHEAFQEQYKGGQQTPPELSEQFPLLRELLRA	99
Full_Seq_Merg	51	EEEPHMLVAFDAGKTIIFRHEAFQEQYKGGQQTPPELSEQFPLLRELLRA	100
Full	100	YRIPAYELENYEADDIIGTLAARAEQEGFEVKVISGDRDLTQLASPHVTV	149
Full_Seq_Merg	101	YRIPAYELENYEADDIIGTLAARAEQEGFEVKVISGDRDLTQLASPHVTV	150
Full	150	DITKKGITDIEPYTPEAVREKYGLTPEQIVDLKGMGDKSINI PGVPGIG	199
Full_Seq_Merg	151	DITKKGITDIEPYTPEIVREKYGLTPEQIVDLKGMGDKSINI PGVPGIG	200
Full	200	EKTAVKLLKQFGTVENVLASIDEIKGKELKETLRQHREMASSKLLAIR	249
Full_Seq_Merg	201	EKTAVKLLKQFGTVENVLASIDEIKGKELKETLRQHREMASSKLLAIR	250
Full	250	RDAPVELSLDDIVYQGEDREKVVALFKELGFQSFLEKMSPSSEEEKPLA	299
Full_Seq_Merg	251	RDAPVELSLDDIVYQGEDREKVVALFKELGFQSFLEKMSPSSEEEKPLA	300
Full	300	KMAFTLADRVTEEMLADKAALVVEVEENYHDAPIVGIAVNEHGRFFLR	349
Full_Seq_Merg	301	KMAFTLADRVTEEMLADKAALVVEVEENYHDAPIVGIAVNEHGRFFLR	347
Full	350	PETALADPQFVAWLGDETKKSMFDSKRAAVALKWKVIELOGVSFDLLLA	399
Full_Seq_Merg	348	-----XWXXXXXXXX-----	355
Full	400	AYLLDPAQGVDDVAAAAMKQYEAURPDEAVYGGKAKRAVPDEFVLAEHL	449
Full_Seq_Merg	356	-----AKHX	359
Full	450	VRKAAAIWALERPFLDELRRNEQDRLLVELEQPLSSI LAEME FAGVKVDI	499
Full_Seq_Merg	360	XXXXXXXX-----	368
Full	500	KRLEQMGEELAEQLRTVEQRIYELAGQEFNINSPKQLGVILFEKLQLPVL	549
Full_Seq_Merg	369	-----XPKXXXXXXXXKXVX	386
Full	550	KKTETGYSTADVLEKLPYHEIVENILHYRQLGKLGSTYIEGLLKWRP	599
Full_Seq_Merg	387	EKTXTXKSADYKXKLPYHEIVXNKLHYRQLGKLGSTYIEGLLKWRP	436
Full	600	DTKKVHTIFNQALTQTGRLSSTEPNLQNIPIRLEEGRKIRQAFVPSSESW	649
Full_Seq_Merg	437	DTKKVHTIFNQALTQTGRLSSTEPNLQNIPIRLEEGRKIRQAFVPSSESW	486
Full	650	LI FAADYSQIELRVLAHIAEDDNLMEAFRRDLDIHTKTAMDI FQVSEDEV	699
Full_Seq_Merg	487	LI FAADYSQIELRVLAHIAEDDNLMEAFRRDLDIHTKTAMDI FQVSEDEV	536
Full	700	TPNMRQAKAVNFGIVYGISDYGLAQLNLSRKEAAEFIERVFESFPGVK	749
Full_Seq_Merg	537	TPNMRQAKAVNFGIVYGISDYGLAQLNLSRKEAAEFIERVFESFPGVK	586
Full	750	RYMENIVQEAQKQGYVITLLHRRRYLPDITSRNFMVRSFAERMAMNTPIQ	799
Full_Seq_Merg	587	RYMENIVQEAQKQGYVITLLHRRRYLPDITSRNFMVRSFAERMAMNTPIQ	636
Full	800	GSAADI IKKAMIDLNARLKEERLQARLLLQVHDELILEAPKEEMERLCRL	849
Full_Seq_Merg	637	GSAADI IKKAMIDLNARLKEERLQARLLLQVHDELILEAPKEEMERLCRL	686
Full	850	VPEVMEQAVILRVPLKVDYHYGSTWYDAK*----- 879	
Full_Seq_Merg	687	VPEVMEQAVILRVPLKVDYH-- KWLDVV*REXGS 718	

Figure 4 Comparison of amino acid sequences. “Full” shows the sequence from the DNA Pol I gene of the original *Geobacillus thermoleovorans* from Batu Kuwung, Serang, Banten; “Full_Seq_Merg” shows the recombinant DNA Pol I gene amino acid sequence

For purposes of comparison, Table 1 shows the molecular weights of recombinant DNA Pol I protein from the current study and other references. It can be seen that the recombinant DNA Pol I protein ranges from 69–110 kDa. The recombinant protein from the current study has a molecular weight of approximately 100 kDa and is therefore consistent with the data range conformity of DNA Pol I in many references.

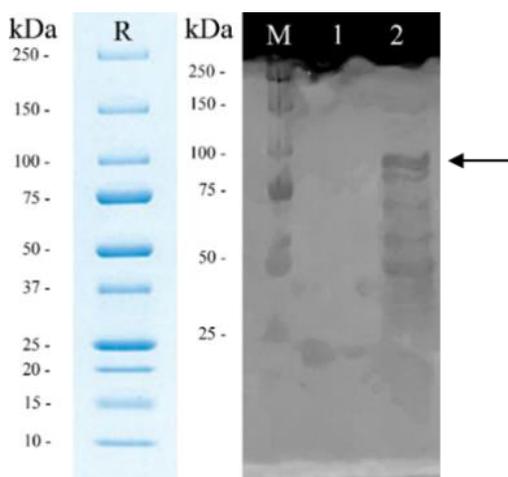


Figure 5 Protein expression test with SDS-PAGE. (R) shows the reference marker; (M) shows the SDS-PAGE marker; (1) shows the pellet waste; and (2) shows the intercellular supernatant protein content, with the overexpressed recombinant protein indicated by the arrow

Table 1 Molecular weight comparison

Thermophilic Bacteria	Molecular weight (<i>kDa</i>)	Reference
<i>Geobacillus caldoxylosilyticus</i>	±99	Sandalli et al., 2009
<i>Bacillus stearothermophilus</i> N3468	±97	Aliotta et al., 1996
<i>Geobacillus stearothermophilus</i> ATCC 12980	±69	Li et al., 2017
<i>Geobacillus thermoleovorans</i> (South Bandung)	±110	Akhmaloka et al., 2008
<i>Geobacillus thermoleovorans</i> (Batu Kuwung, Serang, Banten)	±100	Lischer et al., 2020

4. Conclusions

From the results of this study, it can be concluded that the recombinant plasmid contains the desired gene of interest that is confirmed by digestion using the same restriction enzymes in making of the recombinant plasmid to identify the same restriction areas from gene cloning. The products of the digestion are also consistent with the theoretical gene size (2637 *bp*), despite the slight shift of the bands' positions. The result is further supported by the sequencing, where the approximately first and last 1000 *bp* sequences of the gene of interest inside the recombinant plasmid are consistent with the theoretical nucleotide sequence and correlate with the amino acid sequence. Additionally, the SDS-PAGE result protein expression assay suggests a significant overexpression of proteins at around 100 *kDa*, where the target protein has a molecular weight of 99.65 *kDa*. It is highly encouraged to scale up the protein expression and use a nickel column to purify the His-tagged recombinant protein for further characterization in future research.

Acknowledgements

We acknowledge the financial support from The Directorate of Research and Community Service (DRPM) Universitas Indonesia through Grant of Indexed Publication (Publikasi Internasional Terindeks - PIT 9) No. NKB- 0043/UN2.R3.1/HKP.05.00/2019.

References

- Akhmaloka, Pramono, H., Ambarsari, L., Susanti, E., Nurbaiti, S., Madayanti, F., 2008. Cloning, Homological Analysis, and Expression of DNA Pol I from *Geobacillus thermoleovorans*. *International Journal of Integrative Biology*, Volume 1(3), pp. 206–215
- Aliotta, J.M., Pelletier, J.J., Ware, J.L., Moran, L.S., Benner, J.S., Kong, H., 1996. Thermostable Bst DNA Polymerase I Lacks a 3' → 5' Proofreading Exonuclease Activity. *Genetic Analysis - Biomolecular Engineering*, Volume 12(5–6), pp. 185–195
- Beyer, P., Al-Babili, S., Ye, X., Lucca, P., Schaub, P., Welsch, R., Potrykus, I., 2002. Golden Rice: Introducing the β -Carotene Biosynthesis Pathway into Rice Endosperm by Genetic Engineering to Defeat Vitamin A Deficiency. *The Journal of Nutrition*, Volume 132(3), pp. 506S–510S
- Cranenburgh, R.M., 2004. An Equation for Calculating the Volumetric Ratios Required in a Ligation Reaction. *Applied Microbiol Biotechnol*, Volume 65, pp. 200–202
- Engelke, D.R., Krikos, A., Bruck, M.E., Ginsburg, D., 1990. Purification of *Thermus Aquaticus* DNA Polymerase Expressed in *Escherichia coli*. *Analytical Biochemistry*, Volume 191(2), pp. 396–400
- Gomes, C.M., Faisca, P.F.N., 2019. *Protein Folding: An Introduction*. Springer. doi: 10.1007/978-3-319-00882-0_1
- Hermansyah, H., Maresya, A., Putri, D. N., Sahlan, M. and Meyer, M. (2018) 'Production of Dry Extract Lipase from *Pseudomonas Aeruginosa* by the Submerged Fermentation Method in Palm Oil Mill Effluent', *International Journal of Technology*, 9(2), p. 325. doi: 10.14716/ijtech.v9i2.1511.
- Li, P., Amenov, A., Kalendar, R., Abeldenov, S., Khassenov, B., 2017. Cloning and Purification of Large Fragment of DNA Polymerase I from *Geobacillus stearothermophilus* and its Application in Isothermal DNA Amplification. *Biotechnology Theory and practice*, Volume March 2017, pp. 1-12. doi: 10.11134/btp.1.2017.6
- Maddocks, S., Jenkins, R., 2017. *Understanding PCR: A Practical Bench-top Guide*. Academic Press
- Sandalli, C., Singh, K., Modak, M.J., Ketkar, A., Canakci, S., Demir, I., Belduz, A.O., 2009. A New DNA Polymerase I from *Geobacillus caldxylosilyticus* TK4: Cloning, Characterization, and Mutational Analysis of Two Aromatic Residues. *Applied Microbiology and Biotechnology*, Volume 84(1), pp. 105–117
- Valones, M.A.A., Guimarães, R.L., Brandão, L.A.C., De Souza, P.R.E., De Albuquerque Tavares Carvalho, A., Crovela, S., 2009. Principles and Applications of Polymerase Chain Reaction in Medical Diagnostic Fields: A Review. *Brazilian Journal of Microbiology*, Volume 40(1), pp. 1–11
- Zheng, W., Wang, Q., Bi, Q., 2016. Construction, Expression, and Characterization of Recombinant Pfu DNA Polymerase in *Escherichia coli*. *Protein Journal*, Volume 35(2), pp. 145–153