Overexpression of soluble recombinant *Thermus thermophilus* (*Tth*) DNA polymerase in *Escherichia coli* BL21(DE3) using an MBP fusion tag as a solubility enhancer

Iman Permana Maksum1*, Diandra Firdiani Utami1, Eva Annisa Nurhakim1, Yosua1, Sriwidodo2, Muhammad Yusuf1, Muhammad Fadhlillah1, Ryan Adibagus Haryanto3

1Department of Chemistry, Faculty of Mathematics and Natural Sciences, Padjadjaran University, Sumedang, West Java, Indonesia.
2Department of Pharmaceutics and Pharmaceutical Technology, Faculty of Pharmacy, Padjadjaran University, Sumedang, West Java, Indonesia.
3PT. Bio Farma (Persero), Bandung, West Java, Indonesia.

**ARTICLE INFO**

Received on: 24/10/2021
Accepted on: 21/04/2022
Available Online: 04/09/2022

**Key words:** *Escherichia coli* BL21(DE3), pD861-His, pD861-MBP, RT-PCR, *Tth* DNA polymerase.

**ABSTRACT**

*Tth* DNA polymerase is a thermostable enzyme derived from *Thermus thermophilus* and acts as a DNA polymerase and reverse transcriptase. *Escherichia coli* is used for large-scale enzyme production because of its cost-effectiveness, rapid growth, and increased recombinant protein expression, but inclusion bodies can be formed during intracellular protein expression, so the maltose-binding protein (MBP) tag was used to improve the expression of soluble protein. The *Tth* DNA polymerase gene was optimized to have a codon adaptive index of 1.00% and 60.64% Guanine and Cytosine (GC) content, then inserted into *E. coli* BL21(DE3), which harbors pD861-His-*Tth* DNA polymerase and pD861-MBP-*Tth* DNA polymerase. The induction and postinduction incubation time were optimized to express pD861-His-*Tth* DNA polymerase and pD861-MBP-*Tth* DNA polymerase in the soluble form. The total protein concentration of His-*Tth* DNA polymerase is 3.9095 mg/ml while for MBP-*Tth* DNA polymerase it is 33.541 mg/ml; protein levels after optimization based on densitometry analysis show MBP-*Tth* DNA polymerase is seven times higher than His-*Tth* DNA polymerase. This indicates that MBP tag fusion increases the amount of soluble protein produced.

**INTRODUCTION**

*Tth* DNA polymerase is a thermostable DNA polymerase enzyme derived from the thermophilic bacterium *Thermus thermophilus*, which exhibits dual activity as a DNA polymerase and reverse transcriptase so that it can be used in Reverse transcription-polymerase chain reaction (RT-PCR). It catalyzes the polymerization of DNA in the presence of Mg2+ ions, thereby aiding the formation of cDNA from RNA (*Auer and Landre, 1995*; *Moreno et al., 2005*; *Myers and Gelfand, 1991*). *Tth* DNA polymerase has 5’-3’ exonuclease activity but does not have 3’-5’ exonuclease activity. *Tth* DNA polymerase has also been reported to have RNAse-H-like activity that plays a role in the degradation of DNA-RNA duplexes (*Auer and Landre, 1995*).

Recombinant protein technology is typically used for large-scale enzyme production (*Tripathi, 2016*). Among the many expression systems producing recombinant protein, *Escherichia coli* is commonly used due to its high density on an inexpensive substrate and relatively inexpensive cultivation costs. Moreover, its genetics has been well characterized, and there is a wide availability of mutant host strains (*Sriwidodo et al., 2019*). Common protocols for protein production in *E. coli* cells include cloning of target genes in expression plasmids, transformation in the host strain, and then induction of target protein synthesis (*Chae et al., 2017*). Protein production in *E. coli* can use two approaches, extracellular and intracellular (*Maksum et al., 2017; Silaban et al., 2019; Sriwidodo et al., 2017*). Intracellular protein expression...
occurs in the cytoplasm of *E. coli*, while extracellular expression will direct protein expression to the culture medium by adding signal peptides. The level of protein expression intracellularly has a higher rate than extracellularly (Choi and Lee, 2004; Su et al., 2006).

Various types of enzymes were successfully produced in the host *E. coli*, including the recombinant *Tth* DNA polymerase enzyme. The recombinant *Tth* DNA polymerase enzyme was successfully expressed in the host *E. coli* (Moreno et al., 2005). There have been previous studies on the recombinant DNA polymerase *Tth* enzyme successfully expressed in both *T. thermophilus* and *E. coli* hosts (Dabrowski and Kur, 1998; Moreno et al., 2005). In this study, protease-deficient strains, such as BL21, were used. BL21 is an *E. coli* strain lacking in ompT and lon, two proteases that can interfere with the isolation of intact recombinant proteins as an approach to increase target protein expression (Maksum et al., 2019). In addition, *E. coli* BL21 is also able to live on media with limited nutrition. Therefore, with the use of the *E. coli* BL21 strain, protein recombinants can be expressed with low protease levels (Jeong et al., 2009).

Although *E. coli* expression systems are cost-effective and suitable for large-scale recombinant protein production, the lack of molecular chaperons and the reducing environment of the *E. coli* cytoplasm can lead to the formation of insoluble aggregates. Their formation is often undesirable because they generally lack biological activity (Gomes et al., 2016; Singh et al., 2015). The formation of inclusion bodies was observed in the expression of the recombinant *Tth* DNA polymerase enzyme in *E. coli*, so efforts are needed to increase the acquisition of the active enzyme in the cytoplasm (Moreno et al., 2005). To address this issue, a solubility enhancer is required to improve the yield of soluble recombinant proteins, such as maltose-binding protein (MBP). Previously, it was shown that MBP-tagged fusion protein is more abundantly expressed in the soluble fraction than His-tagged and GST-tagged fusion protein, indicating that MBP is a very efficient solubility enhancer for the soluble expression of recombinant protein (Raran-Kurusu and Waugh, 2012). For these reasons, the *E. coli* expression system and MBP were chosen for this work to express recombinant *Tth* DNA polymerase in *E. coli* BL21(DE3).

In this report, we evaluated the expression of recombinant *Tth* DNA polymerase in *E. coli* BL21(DE3). The *Tth* DNA polymerase gene was optimized using the preference codon of *E. coli*, then constructed on the expression plasmid. *Escherichia coli* BL21(DE3) was transformed with pD861-His-*Tth* DNA polymerase and pD861-MBP-*Tth* DNA polymerase plasmids. *His-Tth* DNA polymerase and MBP-*Tth* DNA polymerase were expressed as soluble forms in *E. coli* BL21(DE3). Therefore, it is highly expected that the production of the recombinant *Tth* DNA polymerase enzyme in the host *E. coli* BL21(DE3) is expected to be developed as a component of the RT-PCR kit for the detection of SARS-CoV-2.

**MATERIALS AND METHODS**

**Bacterial strain, plasmids, and materials**

*Escherichia coli* BL21(DE3) was obtained from our laboratory stock. *Thermus thermophilus* (*Tth*) DNA polymerase I gene (GeneBank accession No. AP008226.1) was inserted into the pD861-His and pD861-MBP plasmids. Plasmids pD861-His-*Tth* DNA Pol and pD861-MBP-*Tth* DNA Pol were synthesized and purchased from ATUM (Newark, CA). A bacterial protein extraction kit was purchased from Thermo-Fisher Scientific (USA). Bacto agar and tryptone were from 1’st Base (Singapore). Bovine serum albumin, kanamycin sulfate, and *L*-rhamnose were purchased from Sigma-Aldrich (Singapore). Bradford reagent was purchased from HiMedia (India). Calcium chloride, hydrochloric acid, tris-base, urea, and *β*-mercaptoethanol were from Merck (Kenilworth, NJ), and yeast extract was purchased from Oxoid (UK).

**Transformation of *E. coli* BL21(DE3)**

*Escherichia coli* BL21(DE3) was grown overnight in Luria Betani (LB) media at 37°C before 1% of the culture was inoculated into 50 ml LB media containing 75 μg/ml kanamycin. Cells were grown until the OD₆₀₀ reached 0.4–0.5 and then placed in an ice bath for 30 minutes before centrifugation at 4,000 g for 10 minutes at 4°C. The cell pellet was resuspended in 25 ml of 0.1 M cold CaCl₂ and incubated for 60 minutes before centrifugation at 4,000 g for 10 minutes at 4°C. The cell pellet was resuspended in 1 ml of 0.1 M cold CaCl₂. *Escherichia coli* BL21(DE3) was transformed with pD861-His-*Tth* DNA polymerase and pD861-MBP-*Tth* DNA polymerase by the heat shock method. The plasmid (1 μl) was added into a microtube containing 100 μl competent cells and incubated in an ice bath for 30 seconds and then placed in a water bath at 42°C for 45 seconds before cooling in an ice bath for 2 minutes. Subsequently, 900 μl of LB media was added to the microtube and incubated at 37°C for 1 hour. A 100 μl aliquot of transformed cells was spread on LB agar supplemented with 75 μg/ml kanamycin.

**Optimization of induction time**

*Escherichia coli* BL21(DE3) cells harboring pD861-His-*Tth* DNA Pol and pD861-MBP-*Tth* DNA Pol plasmids were grown overnight in LB media at 37°C. The starter culture was then transferred into three flasks containing 20 ml LB medium supplemented with 75 μg/ml kanamycin and incubated at 37°C with shaking at 200 rpm. Cells were grown until the OD₆₀₀ reached 0.4 (early log phase), 0.6 (mid-log phase), and 1.0 (late log phase). Protein expression was induced by the addition of 4 mM *L*-rhamnose and incubation at 37°C. Cells were harvested after 4 hours of incubation by centrifugation at 6,000 g for 20 minutes at 4°C. For each 1 g of pellet, a 4 ml bacterial protein extraction buffer containing 2 μl DNase and 2 μl lysozyme was added and the suspension was incubated for 25 minutes at room temperature before centrifugation at 15,000 g for 5 minutes at 4°C. The supernatant was collected as a soluble fraction, and 8 ml of solubilization buffer (8 M urea, 20 mM Tris-HCl, and 5 mM *β*-mercaptoethanol, pH 8.5) was added for each 1 g of pellet. The suspension was incubated at 100 rpm for 1 hour at room temperature and centrifuged at 12,000 g for 20 minutes at 4°C. The supernatant was collected as solubilized inclusion bodies and stored at −20°C for further analysis.

**Optimization of postinduction incubation time**

*Escherichia coli* BL21(DE3) cells harboring pD861-His-*Tth* DNA Pol and pD861-MBP-*Tth* DNA Pol plasmids were grown overnight in LB media at 37°C. The starter culture was inoculated into 50 ml LB media containing 75 μg/ml kanamycin. For His-*Tth* DNA polymerase, cells were grown until the OD₆₀₀ reached 0.4 and 1.0 for MBP-*Tth* DNA polymerase, and then 4 mM *L*-rhamnose was added to induce protein expression at 37°C.
After induction, cells were harvested at different times (4, 5, 6, 7, and 8 hours of incubation) by centrifugation at 6,000 g for 20 minutes at 4°C. Soluble protein and inclusion bodies were isolated following the previously described methods.

**Overexpression of recombinant His-Tth DNA polymerase and MBP-Tth DNA polymerase under optimized conditions**

*Escherichia coli* BL21(DE3) cells harboring pD861-His-Tth DNA Pol and pD861-MBP-Tth DNA Pol plasmids were grown overnight in LB media at 37°C. The starter culture was inoculated into a flask containing 50 ml LB media. For His-Tth DNA polymerase, cells were grown until the OD$_{600}$ reached 0.4 and then induced with 4 mM L-rhamnose at 37°C. Cells were harvested after 4 hours of incubation. For MBP-Tth DNA polymerase, cells were grown until OD$_{600}$ reached 1.0 and induced with 4 mM L-rhamnose at 37°C. Cells were harvested after 8 hours of incubation by centrifugation for 20 minutes at 6,000 g at 4°C. Protein extraction and inclusion bodies solubilization procedures were repeated to collect the soluble fraction and inclusion bodies.

**Direct western blotting**

Whole-cell extracts were fractionated by SDS-PAGE and transferred to a nitrocellulose membrane using a transfer apparatus according to the manufacturer’s protocol (Bio-Rad). The membrane was immersed in 25 ml of blocking solution [5% nonfat dry milk in Tris Buffered Saline with Tween (TBST) (50 mM Tris-HCl pH 7.5, 150 mM NaCl, and 0.1% Tween 20)] and incubated for 15 minutes at room temperature. After incubation, the membrane was washed once with 10 ml of primary antibodies (penta-His antibody QIAGEN) for 10–20 minutes. The membrane was rinsed in 50 ml of wash buffer for 3 minutes before the addition of 10 ml of the substrate and incubated for 10–30 minutes.

**Total protein quantification**

The total protein yield was measured by the Bradford assay. A dilution series of standards were prepared using the Bovine Serum Albumin (BSA) solution (0.2–2 mg/ml). The Bradford reagent (1 ml) was added to 100 µl of the standards and samples, and the absorbance was measured at 595 nm after 10 minutes.

**RESULTS AND DISCUSSION**

**Construction of the recombinant plasmid**

The *Tth* DNA polymerase synthetic gene was obtained from GeneBank (Accession No. AP008226.1), consisting of 2,502 nucleotides. *Tth* DNA polymerase gene shares approximately 50% similarity to the codon preference of *E. coli*. A high percentage of rare codons on the gene of interest will decrease recombinant protein expression (*Gomes et al.*, 2016); therefore, the gene should be optimized to achieve a high codon adaptive index (CAI) and 40%–70% GC content (*Parret et al.*, 2016; *Silaban et al.*, 2019; *Sriwidodo et al.*, 2017). The *Tth* DNA polymerase gene was optimized using the OPTIMIZER software and the *E. coli* B preference codon (*Puigbò et al.*, 2007) to achieve a CAI of 1.000% and 60.64% GC content. The optimized *Tth* DNA polymerase gene was inserted into the expression plasmids pD861-His and pD861-MBP, then synthesized by ATUM. pD861-His-Tth DNA polymerase contains the N-terminal His-tag, while pD861-MBP-Tth DNA polymerase contains the N-terminal MBP tag. The His-tag was used as an affinity tag for purification purposes, and the MBP tag was used as a solubility enhancer. The plasmid map of pD861-His-Tth DNA polymerase and pD861-MBP-Tth DNA polymerase is shown in Figure 1.

![Figure 1. Detailed map of (a) pD861-His-Tth DNA polymerase and (b) pD861-MBP-Tth DNA polymerase. Both plasmids contain the rhaBAD promoter, kanamycin selection marker, ORI pUC, and strong RBS.](image)

---


The competent *E. coli* BL21(DE3) cells were chemically transformed using the heat shock method as it is a relatively inexpensive, simple method that does not require special equipment (Liu et al., 2014). *Escherichia coli* BL21(DE3) transformed cells were observed on LB agar containing kanamycin as a selection marker (Fig. 2). The plasmid contains gene encoding neomycin phosphotransferase II which is responsible for kanamycin resistance; therefore, the transformant cell harboring pD861-His-Tth DNA polymerase and pD861-MBP-Tth DNA polymerase will survive in the presence of kanamycin (Jang and Magnuson, 2013; Wegerer et al., 2008).

**Optimization of the induction time**

Induction time in different growth phases of *E. coli* plays a significant role in the production of soluble recombinant protein (Ahmad et al., 2018; Kaur et al., 2018). The results showed that His-Tth DNA polymerase and MBP-Tth DNA polymerase were expressed in soluble form when the induction was performed between the early exponential to late exponential growth phase (Fig. 3). To determine the optimal expression of the soluble fraction, semiquantitative densitometry analysis was carried out using ImageJ (Yosua et al., 2021), indicating that the highest yield of His-Tth DNA polymerase was achieved when the induction was carried out in the mid-log phase (Fig. 4a), whereas induction in the mid-log phase provided a high yield due to the rapid growth of *E. coli* leading to a higher rate of protein expression (Yosua et al., 2021). Meanwhile, the optimal induction of MBP-Tth DNA polymerase was in the late log phase (Fig. 4b), in line with a previous study that reported that the level of recombinant protein increased up to the late log phase (Noi and Chung, 2017).

**Optimization of postinduction incubation time**

To investigate the effect of postincubation time on Tth DNA polymerase expression, five different incubation times (4, 5,
6, 7, and 8 hours) were evaluated, as a previous study suggested that postinduction incubation time can affect the recombinant protein yield (Fazaeli et al., 2019). The duration of postinduction incubation time is affected by several factors such as the strength of the promoter, concentration of inducer, solubility of recombinant protein, and intrinsic properties of the recombinant protein (Noi and Chung, 2017). The different incubation times after induction resulted in different yields of soluble protein (Fig. 5), with the highest expression of the His-Tth DNA polymerase reached within 4 hours of incubation, whereas MBP-Tth DNA polymerase reached maximal expression within 8 hours of incubation (Fig. 6).

**Expression of recombinant His-Tth DNA polymerase and MBP-Tth DNA polymerase under optimized conditions**

Recombinant His-Tth DNA polymerase and MBP-Tth DNA polymerase were successfully expressed as 95 and 135 kDa protein in E. coli, respectively. The expression of recombinant His-Tth DNA polymerase and MBP-Tth DNA polymerase was conducted under optimized induction and postinduction incubation time with the soluble fractions obtained before optimization used for comparison. As shown in Figure 7, His-Tth DNA polymerase was mainly expressed in the insoluble form when MBP-Tth DNA polymerase was expressed in soluble form, with the amount of both proteins increasing after optimization in Figure 8. Also, the
expression of MBP-Tth DNA polymerase produced more soluble protein than His-Tth DNA polymerase. This is in line with a previous study that reported that all MBP-tagged protein fusions yielded more soluble protein than His-tagged protein fusions (Raran-Kurussi and Waugh, 2012), as MBP serves as a “holdase” to maintain the passenger protein in a soluble form until it reaches its native conformation either by spontaneous folding or with the assistance of endogenous chaperones (Raran-Kurussi and Waugh, 2012). This mechanism may involve molecular interaction between MBP and hydrophobic amino acid residues existing on the incompletely folded passenger protein, thereby preventing aggregation (Guo et al., 2018; Raran-Kurussi and Waugh, 2012).

**Western blotting**

The expressions of pD861-His-Tth DNA polymerase and pD861-HMBP-Tth DNA polymerase were confirmed by Western blotting using Glyceraldehyde 3-phosphate dehydrogenase (GAPDH) as the internal loading control (Fig. 9).
Total protein quantification

The total protein concentration of the His-\textit{Tth} DNA polymerase soluble fraction before and after optimization was 3.7002 and 3.9095 mg/ml, respectively. Meanwhile, the total protein concentration of the MBP-\textit{Tth} DNA polymerase soluble fraction before and after optimization was 30.097 and 33.541 mg/ml, respectively. These results indicate that the optimization process increased the total protein concentration. The concentration of total protein in the soluble fraction of MBP was eight times higher than the concentration of total protein in the soluble fraction of His.

CONCLUSION

In summary, soluble \textit{Tth} DNA polymerase was expressed in \textit{E. coli} BL21(DE3), with MBP-\textit{Tth} DNA polymerase producing more soluble protein than His-\textit{Tth} DNA polymerase due to its ability as a solubility enhancer.

ACKNOWLEDGMENTS

This work was supported by Grant No. 74/FI/P-KCOVID-19.2B3/IX/2020 from the National Innovation Research Agency, Ministry of Technology Research and Higher Education.
CONFLICTS OF INTEREST
The authors have nothing to disclose.

AUTHORS’ CONTRIBUTIONS
All authors made significant contributions to the work reported, in the conception and design, acquisition and interpretation of data, statistical analysis, drafting, revising, or critically reviewing the article and gave final approval of the version to be published.

ETHICAL APPROVALS
This study does not involve experiments on animals or human subjects.

DATA AVAILABILITY
All data generated and analyzed are included within this research article.

PUBLISHER’S NOTE
This journal remains neutral with regard to jurisdictional claims in published institutional affiliation.

REFERENCES


Raran-Kurussi S, Waugh DS. The ability to enhance the solubility of its fusion partners is an intrinsic property of maltose-binding protein but their folding is either spontaneous or chaperone-mediated. PLoS One, 2012; 7(11):1–10.


How to cite this article: