Bst polymerase enhancement
A bioinformatics approach to improve *Bst* polymerase characteristics

Jonathan, Ernest Suryadjaja, Sulistyo Emantoko D. Putra*

Faculty of Biotechnology, University of Surabaya (UBAYA), Surabaya 60293, Indonesia
**Corresponding author. Email: emantoko@staff.ubaya.ac.id*

ABSTRACT
DNA polymerase is a remarkably incredible invention in the biotechnology field. Since its discovery, molecular genetic-based research has been growing rapidly. Various methods for molecular-based diagnostics have been developed since. One of which is Loop-Mediated Isothermal Amplification; this method offers simplicity, sensitivity, and a faster result than conventional PCR. However, the availability of the materials supporting the LAMP reaction is still challenging. LAMP materials are often hard to access in a developing country, and the need to import from other countries significantly increases the price. Here we offer a *Bst* polymerase design that can be produced with the same or even higher quality than the commercial one. Based on bioinformatics analysis, sto7d fusion to the C-terminus of *Bst* polymerase (Br512g3.1) shows higher stability and solubility. Sto7d fusion to the C-terminus improves processivity when used in Gss (*Bst*-like polymerase). Lastly, Sto7d fusion to Br512g3.1 offers higher stability and processivity that can be used to overcome the cost problem of *Bst* polymerase.

Keywords: *Polymerase, Bst, POCT, Protein Modification.*

1. INTRODUCTION

DNA polymerase has become widely used for genomic applications such as amplification, detection, and cloning. Many types of DNA polymerase from different species have been used for this application. Some are *Taq*, *Phi*29, *Pfu*, KOD, *Gss*, Phusion, and *Bst* polymerases. Different DNA polymerases have different characteristics for specific needs. As an example, *Taq* DNA Polymerase from Thermus aquaticus is widely used DNA polymerase for conventional PCR due to its stability in high temperatures allowing high-temperature cycling, which is suitable for DNA amplification. Some DNA polymerase enzymes, such as Phi29 able to amplify DNA at 30°C, amplify lengthy DNA products and have tight binding to DNA which causes high processivity [1]. On the other hand, *Bst* polymerase is a DNA polymerase enzyme mainly used in isothermal amplification since it has high strand displacement activity. Strand displacement activity is essential in isothermal amplification since there is no denaturation step.

LAMP (Loop-Mediated Isothermal amplification) is one of the isothermal DNA amplification methods that were found by Notomi *et al.* (2000) and that has now been popularly used in the diagnostic assay, namely SARS-CoV-2 [3], Malaria [4], *Mycoplasma genitalium* [5], and many more. The LAMP-based diagnostic method is preferred because its more sensitive than conventional PCR (due to the use of 2-3 pairs of primers), cheaper (no thermal cycler required), has higher resistance towards inhibitors, and gives faster results which make LAMP the suitable candidate for POCT. However, there are difficulties for developing countries to access LAMP material. This problem is caused by patents on the commercial kit and the complexity of direct material import. This further causes minimum ability to independently produce the materials needed for LAMP reaction. The need to develop one of the LAMP materials, specifically *Bst* polymerase, the primary enzyme for the reaction with the same or higher activity than the commercial one.

Protein modification to achieve better characteristics can be done by a few methods, such as mutation and fusion. We now analyze a few modification designs of *Bst* polymerase to support the independent production of *Bst* polymerase with identical or even better characteristics than the commercial one. Trustfully able to help people independently produce their *Bst* polymerase that also can boost molecular-based (LAMP-based) POCT.

J. Sukweenadhi and F. Setiawan (eds.), *Proceedings of the Conference on Natural Resources And Life Sciences 2022 (NRLS-BIO 2022)*, Advances in Biological Sciences Research 38,
https://doi.org/10.2991/978-94-6463-322-1_11

2. MATERIALS AND METHOD

The sequence of *Bst* polymerase (without 5'-3' exonuclease activity) was picked from previous research by Li et al. (2017) (*Bst*pol), which was taken *Geobacillus stearothermophilus* ATCC 12980 with modification (N- terminus 6x His – Thrombin site – *Bst*) and Br512g3.1 by Paik et al., (2021) with modification and few mutations for activity and stability improvement (N-terminus 8x His – vHP47 – *Bst*). For the modification, we choose alanine addition between T82 and A83 for *Bst*pol and fusion of a DNA binding protein from *Sulfolobus tokodaii* – Sto7d for Br512g3.1.

3. RESULTS AND DISCUSSION

Bioinformatics is often used in preliminary research to simulate, predict, and determine the outcome of a research. We analyzed *Bst* polymerase modification using a few parameters, instability index, aliphatic index, probability of solubility, and flexibility. The instability index is a parameter used to analyze protein stability where the Instability value < 40 means the protein is relatively stable [7]. The result of each enzyme design shows a value above 40, meaning all the design isn't stable. Nevertheless, we can see the decreased value in alanine addition in *Bst*pol and Sto7d fusion in Br512g3.1.

C-terminus. Even if there is a decrease in the instability value of *Bst*pol with alanine addition, it's only a slight decline that can't be used to prove that alanine addition increases the stability of the protein. On the other hand, Sto7d fusion gives significant results compared to the modified *Bst*pol. The lower instability index value shows that the modified enzyme has higher stability than the non-modified one. Regarding the aliphatic index, it shows the relative value of the aliphatic side chain (Pro, Ala, Iso, and Val) in a protein. It is stated that the aliphatic index is found higher in the proteins from thermophilic bacteria, so this can imply that a higher aliphatic index indicates the broader range of thermal stability of a protein [8]; the modification in *Bst*pol didn't show any difference because the basis of aliphatic index computing is based on the volume occupied by aliphatic side chain in a whole protein so that little change (one addition of amino acid) won't show any significant differences. On the other hand, a fusion of Sto7d to Br512g3.1 shows a lower aliphatic index which means this fusion protein might have lower thermal stability; the main reason for Sto7d fusion is to increase the processivity of the Br512g3.1, and its resistances towards polymerization inhibitors as what has been done before [9]. The probability of solubility is used to predict protein solubility in an expression system.

Compared to the Sto7d fusion, *Bst*pol modification only show little increase of probability of solubility so that Sto7d fusion to Br512g3.1 is more favorable to achieve higher expressed protein yield than *Bst*pol+(A). The last parameter is protein flexibility, protein flexibility is important for its bioactivity. Other than flexibility, there is rigidity, both of flexibility and rigidity is important. Flexibility is used to make sure the active site of a protein can work properly to bind with substrate, in the other side rigidity is also important to compensate structural change caused by temperature fluctuation [10]. In the flexibility parameter we can see that there is only slight increase in Sto7d fusion and no change in *Bst*pol modification. Since Sto7d protein comes from *Sulfolobus tokodaii*, a thermophilic bacteria we can assume that this protein is generally stable in high temperature which conclude that this protein is rigid, the increase of flexibility in Sto7d might be caused by the flexible linker between Br512g3.1 and Sto7d. By that analysis, we suggest that Sto7d is a promising fusion protein to be fused with Br512g3.1 Since Br512g3.1 is already labeled as ultrafast, the addition of Sto7d is hopefully will increase the processivity and its resistance towards inhibitor so that Br512g3.1-Sto7d can be used as the enzyme for LAMP based POCT.

Table 1. Protein Characteristics analysis from ProtParam and SoDoPe

Parameters	*Bst*pol	*Bst*pol+(A)	Br512g3.1	Br512g3.1-Sto7d
Molecular Weight	68771.91	68842.99	73393.26	81283.43
pI	6.18	6.18	6.43	7.36
Instability Index	47.33	47.26	45.27	41.81
Aliphatic Index	96.31	96.31	92.11	89.25
Probability of Solubility	0.7747	0.7757	0.7881	0.8107
Flexibility	0.9998	0.9998	1.0025	1.004
GRAVY	-0.2823	-0.2789	-0.3837	-0.4177

*MW, pI, Instability Index, and Aliphatic Index were analyzed using ProtParam – ExPASy web server (https://web.expasy.org/cgi-bin/protparam)

*Probability of Solubility, Flexibility, and GRAVY were analyzed using SoDoPe – TISIGNER web server (https://tisigner.com/sodope)

4. CONCLUSION

In conclusion, the study highlights the significance of DNA polymerase in the field of biotechnology and the subsequent rapid growth of molecular genetic-based research. The Loop-Mediated Isothermal Amplification (LAMP) method has emerged as a valuable tool for molecular-based diagnostics, providing simplicity, sensitivity, and faster results compared to conventional PCR. However, the accessibility of materials supporting the LAMP reaction poses a significant challenge, especially in developing countries, leading to increased costs through imports.

AUTHORS' CONTRIBUTIONS

All authors designed the study, J collected, analyzed, interpreted data, and drafted the main manuscript. SEDP reviewed the manuscript. The authors read and approved the final manuscript.

ACKNOWLEDGEMENTS

This study was supported by Faculty of Biotechnology, University of Surabaya, Indonesia.

REFERENCES

[1] P. Li, A. Amenov, R. Kalendar, S. Abeldenov, and B. Khassenov, "Cloning and purification of large fragment of DNA polymerase I from Geobacillus stearothermophilus and it's application in isothermal DNA amplification," *Biotechnol. Theory Pract.*, no. October, 2017, doi: 10.11134/btp.1.2017.6.

[2] T. Notomi *et al.*, "Loop-mediated isothermal amplification of DNA," *Nucleic Acids Res.*, vol. 28, no. 12, p. e63, 2000, [Online]. Available: https://watermark.silverchair.com/2800e63.pdf?token=AQECAHi208BE49Ooan9kkhW_Ercy7Dm3ZL_9Cf3qfKAc485ysgAAAmYwggJiBgkqhkiG9w0BBwagggJTMIICTwIBADCCAkgGCSqGSIb3DQEHATAeBglghkgBZQMEAS4wEQQMy3AQPdcwVpshAVuIAgEQgIICGSfU5ehJRqbIy2A_4pd12_oNbrQem7FM_kTE9ZYQL40XC4l.

[3] A. Maranhao, S. Bhadra, I. Paik, D. Walker, and A. D. Ellington, "An improved and readily available version of *Bst* DNA polymerase for LAMP, and applications to COVID-19 diagnostics," *medRxiv*, no. 8, pp. 1–19, 2020, [Online]. Available: https://doi.org/10.1101/2020.10.02.20203356.

[4] N. W. Lucchi, D. Ljolje, L. Silva-Flannery, and V. Udhayakumar, "Use of malachite green-loop mediated isothermal amplification for detection of plasmodium spp. parasites," PLoS One, vol. 11, no. 3, pp. 1–9, 2016, doi: 10.1371/journal.pone.0151437.

[5] T. Edwards *et al.*, "Loop-mediated isothermal amplification (LAMP) for the rapid detection of Mycoplasma genitalium," *Diagn. Microbiol. Infect. Dis.*, vol. 83, no. 1, pp. 13–17, 2015, doi: 10.1016/j.diagmicrobio.2015.05.010.

[6] I. Paik *et al.*, "Multi-modal engineering of *Bst* DNA polymerase for thermostability in ultrafast LAMP reactions," 2021, doi: https://doi.org/10.1101/2021.04.15.439918.

[7] E. Gasteiger *et al.*, "Protein Identification and Analysis Tools on the ExPASy Server," *Proteomics Protoc. Handb.*, no. 52, pp. 571– 608, 2005, doi: 10.1385/1592598900.

[8] A. Ikai, "Thermostability and aliphatic index of globular proteins," *J. Biochem.*, vol. 88, no. 6, pp. 1895 1898, 1980, doi: 10.1093/oxfordjournals.jbchem.a133168.

[9] I. P. Oscorbin, E. A. Belousova, U. A. Boyarskikh, A. I. Zakabunin, E. A. Khrapov, and M. L. Filipenko, "Derivatives of *Bst*-like Gss-polymerase with improved processivity and inhibitor tolerance," *Nucleic Acids Res.*, vol. 45, no. 16, pp. 9595–9610, 2017, doi: 10.1093/nar/gkx645.

[10] M. Vihinen, "Relationship of protein flexibility to thermostability.," *Protein Eng.*, vol. 1, no. 6, pp. 477-480, 1987, doi: 10.1093/protein/1.6.477