UNIVERSITY OF TARTU
Faculty of Science and Technology
Institute of Technology

Anastasia Kolosova

Purification of Bsm DNA polymerase mutant and characterization of its biological sample inhibitor tolerance

Bachelor's Thesis (12 ECTS)

Curriculum Science and Technology

Supervisor:
PhD Ilja Gaidutšik

Tartu 2020

Purification of Bsm DNA polymerase mutant and characterization of its biological sample inhibitor tolerance

Abstract:

DNA polymerases are widely used in various nucleic acid amplification techniques. The loop-mediated isothermal amplification is nowadays of a high interest because of method's sensitivity, specificity, single operation temperature and low-cost. The biological DNA polymerases are prone to various inhibiting substances found in probes, which may suppress or totally block the enzyme's activity and nucleic acid amplification. The current study aimed at generating a LF Bsm polymerase mutant from Bacillus smithii with an increased tolerance to biological inhibitor. It was proposed that a mutation in fingers domain will improve the polymerase's activity in the presence of inhibitor.

Keywords:

DNA polymerase, replication, A-family polymerase, chimeric DNA polymerases, loop mediated isothermal amplification, LAMP, inhibitors

CERCS:

T490 Biotechnology

Bsm DNA polümeraasi mutandi puhastamine ja selle inhibitor tolerantsete omaduste uurimine

Lühikokkuvõte:

DNA polümeraase laialdaselt kasutatakse paljudes nukleiinhapete amplifikatsioonil põhinevates meetodites. Silmusega vahendatud isotermiline amplifikatsioon (LAMP) on kaasaegne meetodika mis on väga huvipakkuv oma tundlikuse, spetsiifilisuse ja odavuse tõttu ning võimele amplifitseerida DNAt ühe muutumatu temperatuuri juures. Samas meetodis kasutatavad DNA polümeraasid on tundlikud igasuguste, proovides leitavate, inhibiitorite suhtes, mis võivad pärssida või täielikult blokeerida ensüümi aktiivsust ja seega nulkeiinhape amplifikatsiooni. Antud uurimistöö eesmärgiks oli luua Bacillus smithii Bsm polümeraasi suure fragmendi mutandi millel on võimendatud tolerantsus bioloogiliste inhibiitorite suhtes. Töös eeldati, et valgu sõrmede (fingers) domeenis tekitatud mutatsioon on võimeline parandada valgu aktiivsust inhibiitorite juuerolekul.

Võtmesõnad:

DNA polümeraas, amplifikatsioon, A-perekonda polümeraasid, kimäärsed DNA polümeraasid, LAMP, inhibitors

CERCS:

T490 Biotehnoloogia

TABLE OF CONTENTS

TERMS, ABBREVIATIONS AND NOTATIONS 6
INTRODUCTION 8
1 LITERATURE REVIEW 9
1.1 DNA polymerases 9
1.2 DNA polymerase families 10
1.3 Family A DNA polymerases 11
1.4 DNA polymerase structure 11
1.5 Strand displacement activity 14
1.6 Loop-mediated isothermal amplification 15
1.7 Nucleic acid amplification inhibitors 17
2 THE AIMS OF THE THESIS 20
3 EXPERIMENTAL PART 21
3.1 MATERIALS AND METHODS 21
3.1.1 Bacterial strains, plasmids and culture conditions 21
3.1.2 Transformation of expression vectors 23
3.1.3 Bacterial strain growth conditions 23
3.1.4 Purification of Bsm M496I 24
3.1.5 Protein concentration measurement 25
3.1.6 DNA polymerase assay for unit measurement 25
3.1.7 qLAMP 26
3.2 RESULTS 27
3.2.1 Optimization of protein expression conditions 27
3.2.2 Large-scale protein purification 30
3.2.3 Bsm M496I mutant unit measurement 32
3.2.4 qLAMP performance 33
3.3 DISCUSSION 35

SUMMARY 37

REFERENCES 38

NON-EXCLUSIVE LICENCE TO REPRODUCE THESIS AND MAKE THESIS PUBLIC 45

TERMS, ABBREVIATIONS AND NOTATIONS

Amp – ampicillin

Asp – aspartate

BIP – backward inner primer

BOP – backward outer primer

Cam – chloramphenicol

dNTP – deoxyribonucleoside triphosphate

dsDNA – double stranded DNA

EB – elution buffer

EDTA – ethylenediaminetetraacetic acid

FIP – forward inner primer

FOP – forward outer primer

INAAT – isothermal nucleic acid amplification testing

IPTG – isopropyl-β-D-thiogalactoside

KF – Klenow fragment

LAMP – loop mediated isothermal amplification

LB – Luria broth

LF – large fragment

NAAT – nucleic acid amplification technique

NHEJ – nonhomologous end joining

OD – optical density

PPi – pyrophosphate

RAM - ramification amplification

SD – strand displacement

sGRS – small gap-filling repair synthesis

ssDNA – single stranded DNA

STE – saline-tris-EDTA buffer

TLS – translesion synthesis

WB – washing buffer

WGA - whole genome amplification

INTRODUCTION

Nucleic acid amplification tests (NAATs) are widely used in research and medicine, as well as for detection and identification of various pathogenic organisms. Previously, NAATs relied on thermostable DNA polymerases able to withstand the thermal cycling needed for DNA denaturation. Nowadays, the strand displacing DNA polymerases with high tolerance to the presence of inhibiting substances in biological probes are of a high interest. Implemention of such polymerases is found to be attractive due to the provided ability to amplify the target nucleic acid sequences directly from contaminated samples.

Only the limited number of DNA polymerases from families A and B possess a natural strand displacing activity and inhibitor tolerance. Therefore, modified DNA polymerases could be produced to improve the properties of natural polymerases.

The current study involved the purification of constructed mutant LF of Bsm DNA polymerase of A-family with a mutation in a relatively conserved region and the characterization of its resistance to the inhibitor in blood.

1 LITERATURE REVIEW

1.1 DNA polymerases

The primary mission of DNA polymerases is an accurate replication of all living forms' genome which is crucial for maintaining life and passing the genetic information onto next generations (Garcia-Diaz and Bebenek, 2007). The synthesis of a new DNA molecule is performed in the following steps: binding of the primer, a short RNA sequence, to a single-stranded template DNA by primase, incorporation of incoming deoxynucleoside triphosphate (dNTP) according to the Watson-Crick base pairing, formation of phosphodiester bond as a result of nucleophilic attack, release of the pyrophosphate (PPi), and DNA polymerase's displacement to the following 3'OH primer terminus (Patel *et al*., 2001). Polymerization of nucleotides proceeds only in 5'-3' direction for both antiparallel strands. In addition to primers and a sufficient number of nucleotides, Mg^{2+} ions are also crucial for the process to take place (Patel and Loeb, 2001; Kucera and Nichols, 2008).

DNA polymerases vary in their capability of nucleotide incorporation prior to the dissociation from the DNA strand. The number of nucleotides incorporated to the 3' end of an RNA primer of the growing DNA strand before the enzyme's dissociation from the DNA is known as the processivity (Bloom and Goodman, 2001). Another essential feature of DNA polymerases is enzymes' fidelity, which refers to the DNA polymerase's ability to duplicate the genome with a high accuracy in order for the identical DNA to be passed from parental cell onto daughter cells (Bebenek and Ziuzia-Graczyk, 2018).

There is a number of different DNA polymerases found within a single cell accomplishing their specific tasks (Kornberg, 1984). According to the performed function, all DNA polymerases have been divided into two groups, replicative and non-replicative (Shanbhag *et al*., 2018). As the name suggests, replicative DNA polymerases replicate the genome and operate only during cell division (Albà, 2001; Shanbhag *et al*., 2018). On the other hand, the type of DNA polymerases not involved in the genetic material duplication operates throughout the cell's life cycle and accomplish the DNA repair, recombination, and translesion DNA synthesis for the replication of damaged DNA (Shanbhag *et al*., 2018; Berdis, 2017).

1.2 DNA polymerase families

DNA polymerases from all kingdoms of life have been divided into distinct individual families, notably A, B, C, D, X, Y, RT (reverse transcriptase), and PrimPol (primase and polymerase), based on the primary conserved amino acid sequences, crystal structure analyses and the function performed in a cell (Shanbhag, 2018; Yang and Gao, 2018). Except for only viruses, every living form has several different DNA polymerase types, each having its individual role in a cell (Chen, 2014). DNA polymerases sharing sequence homology with *E.coli* DNA pol I (polA), II (polB) and III (polC) have been grouped into families A, B and C, respectively. Polymerases from D-family are phylogenetically linked with Euryarchaeotic Pol II, family X with human Pol β, and Y-family with *E. coli* UmuC/DinB (damage-induced) and eukaryotic *RAD30*/xeroderma pigmentosum variant (Burgers, 2001). A-family polymerases are found within prokaryotes, eukaryotes and viruses (bacteriophages). The B-family polymerases belong to the same kingdoms as A-family, and archaea. Polymerases of C family are characteristic to prokaryotes (bacteria), whereas the D-family polymerases are common to archaea. The major part of A-D families polymerases are replicative DNA polymerases carrying 3'-5' exonuclease domain correcting replication errors, however, there are polymerases with low or no 3'-5' exonuclease activity and those are involved in TLS (translesion synthesis) and sGRS (small gap-filling repair synthesis). DNA polymerases without sequence homology with families A, B and C were assigned to the X-family, mostly belonging to Eukarya. Some family representatives can function with no template provided. These enzymes carry out the sGRS associated with two repair pathways: base excision repair (BER) and non-homologous end joining (NHEJ). The Y-family DNA polymerases are found among all kingdoms of life, sharing a low homology with previous families, and are responsible for TLS. Enzymes possess no proofreading exonuclease activity and have a more open active site to handle base damage. These polymerases are able to recognize and bypass various DNA lesions. The RT family synthesises DNA by using RNA sequence as a template and these are found in retroviruses and humans. This includes reverse transcriptases and eukaryotic telomerases. PrimPol is unique as it is able to synthesize DNA primers and carry DNA replication without RNA primer. The similarity between X, Y, RT and PrimPol families is the lack of a proofreading 3'-5' exonuclease activity (Shanbhag *et al*, 2018; Johansson and Dixon, 2013; Yang and Gao, 2018; Rothwell and Waksman, 2005; Laos *et al.*, 2014).

1.3 Family A DNA polymerases

The current work deals with the A-family DNA polymerase. Therefore, the properties and structure of this enzyme class will be described further. Family A polymerases received their name due to the homology with the polA gene of *E.coli* DNA pol I, which is said to typify this family (Braithwaite and Ito, 1993; Shanbhag *et al.*, 2018). It is the most understood class, and polymerases of this family are single-chain enzymes involved in both DNA replication and repair, namely nucleotide excision repair for processing of Okazaki fragments (Joyce *et al.*, 1981; Rothwell and Waksman, 2005).

The first DNA polymerase to be discovered is the DNA polymerase I isolated from the model organism *Escherichia coli*. The detection of its ability to duplicate DNA sequences and later determination of the large (Klenow) fragment (LF) of Pol I crystal structure in 1985 by Steitz *et al* has aided in understanding of DNA replication mechanism as well as started an era of intensive research and discovery of new DNA polymerases, and study of their functions in the cell (Yang and Gao, 2018; Wu *et al.*, 2017; Shanbhag *et al.*, 2018). Unlike many other replicative enzymes, the Pol I is a single subunit polymerase. The enzyme is a monomer with three domains: 5'-3' polymerase, 5'-3' exonuclease and 3'-5' exonuclease (in *E.coli*) (Patel *et al.*, 2001). Polymerase is mainly responsible for the damaged DNA repair and the Okazaki fragments processing by displacing a primer on the Okazaki's 5' end generated during DNA synthesis (5'-3') and further gap filling between the fragments (Steitz *et al.*, 1987; Ollis *et al.*, 1985; Kornberg, 1984; Steitz and Yin, 2004; Patel and Loeb, 2001). Further studies have revealed that most pol I enzymes contain the 5'-3' exonuclease and 3'-5' proofreading activities. The 5'-3' exonuclease is much more essential as it removes RNA primers from Okazaki fragments generated during replicative DNA synthesis. The DNA polymerase activity is used to fill in the resulting gap, as well as gaps resulted after the repair of DNA lesions. Some replicative polymerases are known to cooperate with other proteins in order to achieve the accurate DNA duplication. For instance, T7 DNA pol interacts with thioredoxin which increases its processivity (Rothwell and Waksman, 2005).

1.4 DNA polymerase structure

All discovered DNA polymerases to date are found to share an extremely similar structure resembling a human right hand with three functional subdomains referred to as the palm,

thumb and fingers, with a deep cleft in the structure formed by the thumb (forming its base) and fingers domains (Rothwell and Waksman, 2005; Johansson and Dixon, 2013; Fitzgerald-Hayes and Reichsman, 2009). The crevice-containing domain has the DNA polymerase active site and is able to bind the B form dsDNA as practically all the positive electrostatic charge potential is located in the cleft, which is essential for binding the duplex DNA, and also due to size and shape of the crevice (Ollis *et al.*, 1985; Steitz, 1987; Tripathi, 2010). The catalytic site for the DNA polymerase activity is located at the bottom of the large crevice (Beese *et al.*, 1993). Enzymes bind the DNA through interactions with the negatively-charged phosphate backbone of incoming dNTP and highly conserved positive Arg754, Arg 682, Lys 758 and His 734 residues among polymerases, as shown in the studies of the Klenow fragment. There are also other conserved residues participating in the incoming dNTPs' incorporation by binding and directing it (Beese *et al.*, 1993;Tripathi, 2010; Patel *et al.*, 2001). Some replicative polymerases carry additional domains for 3'-5' exonucleolytic proofreading, polymerase's interactions with other proteins, or other enzymatic functions (Johansson and Dixon, 2013). The general structure of DNA polymerases is represented in Figure 1 with Gbst pol I LF. The palm subdomain is a catalytic unit of polymerases which catalyses the phosphoryl transfer reaction linking the α-phosphate (constituent of DNA backbone) of incoming nucleotide by nucleophilic attack to the 3'-OH group of a primer. DNA polymerases have sequences of highly conserved regions called motifs in the polymerization site which may have a devastating effect on enzyme's functioning if mutated (Perler et al., 1996). As palm domain is responsible for polymerization, it is the most conserved region among all polymerase families containing amino acids crucial for the catalysis, which bind two Mg^{2+} ions in enzyme's active site needed for the nucleophilic attack. Investigations in KF revealed the conserved Asp-882, Glu-883, and Asp 705 carboxylates (Steitz, 1999; Perler *et al.*, 1996; Franklin *et al.*, 2001; Hamilton et al., 2001). The first Mg^{2+} ion assists in deprotonation of the primer's 3'OH group (nucleophile), whereas the second ion interacts with the α-phosphate for the transition state stabilization and promotes the PPi release (Perler et al., 1996). The fingers subdomain poses an essential role in accurate nucleotide selection and acts by interacting with the both incoming dNTP and templating base in DNA strand. This subdomain possesses the highest structural diversity in a polymerase composition (Hamilton et al., 2001). Residues in the O- and O1-helix located in the fingers subdomain play an essential role for the enzyme's functioning. The dNTPs bind to the end of N-terminal region of O-helix (Beese *et al.*, 1993; Kaushik *et al.*, 1996; Rothwell and Waksman, 2005). Residues 759-775 in KF, of

the O- and O1-helices, contain the DNA template-primer binding domain, vital for polymerase activity (Kaushik *et al*., 1996). Moreover, residues of the O1-helix are important for the polymerase's strand-displacement activity (Singh *et al*., 2007; Piotrowski *et al*., 2019). The thumb domain binds the dsDNA's minor groove, contributes to its correct positioning in the polymerase active site, assists in DNA polymerase translocation onto the next templating base and interacts with incoming dNTP (Scott et al., 2001; Beese *et al*., 1993; Berdis, 2017; Patel *et al*., 2001; Rothwell and Waksman, 2005).

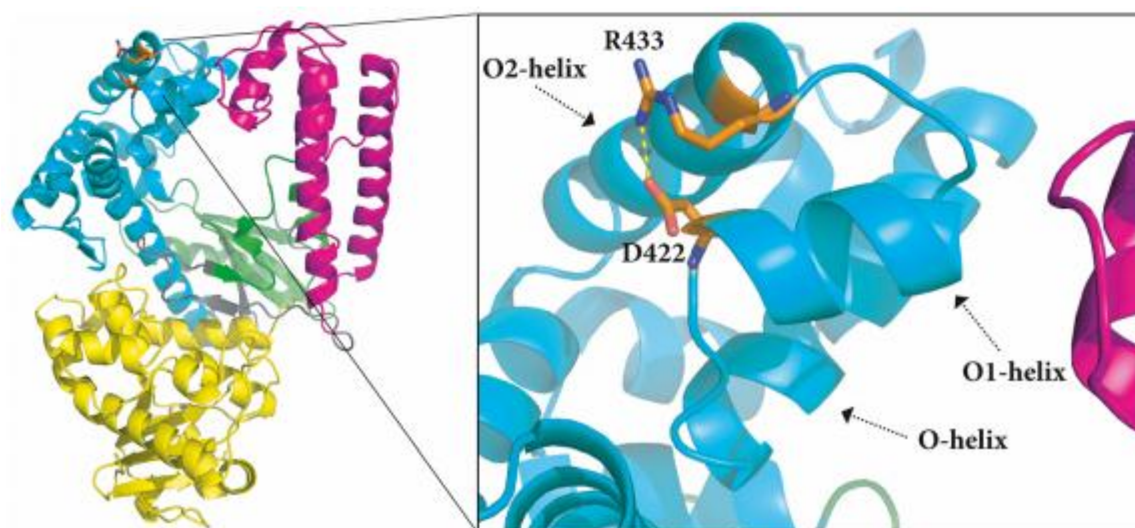


Figure 1. Overview of *Geobacillus stearothermophilus* (Gbst) pol I large fragment. On the left portion of the illustration the subdomains of the DNA pol LF are represented: palm (green), fingers (cyan) and thumb (magenta). On the right side there are O-, O1-, and O2-helices depicted with a close-up view of D422 residue on the O1-helix and R433 located on the O2-helix (both in orange), and the distance measurement between them (Piotrowski et al., 2019).

1.4.1 Bst and Bsm DNA polymerases from A family

A-family DNA polymerases are actively used for *in vitro* DNA amplification in biotechnology, molecular biology and diagnostics (Laos, 2014; Piotrowski *et al*., 2019). The majority of methods rely on thermostable DNA polymerases (Oscorbin *et al*., 2016). However, modern tools for isothermal amplification focus on moderately thermostable enzymes with strand displacement activity, such as Bst and other Bst-like DNA polymerases (Oscorbin *et al*., 2016). The large fragment of DNA pol I isolated from *Bacillus stearothermophilus*, Bst DNA pol LF, is the indispensable type of enzyme used in the

loop-mediated isothermal amplification (LAMP) due to its SD activity, also applied in Next Generation Sequencing (Kaneko *et al.*, 2006; Ma *et al.*, 2016). Bst polymerase itself consists of the N-terminal 5'-3' exonuclease and C-terminal 5'-3' DNA polymerase domains (Riggs *et al.*, 1996; Aliotta *et al.*, 1996). The LF is modified to lack the 5'-3' exonuclease activity by single amino acid substitution (Riggs *et al.*, 1996). Bst LF performs synthesis at optimal temperature of 65°C, requires low concentration of DNA substrate and posesses high processivity (Riggs *et al.*, 1996; Mead *et al.*, 1991).

Bsm pol LF is a fragment of DNA pol from *Bacillus smithii* catalyzing 5'-3' DNA synthesis similarly to Bst LF while lacking both 5'-3' and 3'-5' exonuclease domains, and demonstrating high SD activity at temperatures ranging between 30-63°C, with optimal of 60°C. Due to the high similarity with Bst DNA polymerase, it can be applied in most methods utilizing it, such as LAMP, whole genome amplification (WGA) and ramification amplification (RAM) [1].

1.5 Strand displacement activity

Only the limited number of A- and B-family DNA polymerases have a strong strand displacement activity (e.g., Bst LF and phi29 pol) enabling a protein to unwind the dsDNA without the assistance of cofactors, such as helicase and single-strand binding proteins, which is important for isothermal amplification (Singh *et al.*, 2006). SD enzymes have also been found to demonstrate high processivity (Hamilton et al., 2001; Milligan *et al.*, 2018; Blanco and Salas, 1984). The studies with two primers bound to template DNA revealed that once the SD polymerase reaches the 5' end of the second hindering primer during polymerization, the primer is displaced by an enzyme letting the synthesis to proceed. In case of polymerase without SD activity, there would be two possible outcomes. The polymerization would either stop at this point, or the hindering primer would be degraded by the 5′-3′ exonuclease (Hamilton et al., 2001).

The SD activity can be improved by producing recombinant polymerases with point mutations. LF of *Psychrobacillus* DNA polymerase I mutant with a substitution in fingers domain have demonstrated a 2.5-fold increase in SD activity. The identical residue change in two other thermostable polymerases of A-family leads to increased SD activity as well. The replacement of Asp residue at position 422 to other larger hydrophobic, polar and positively charged residues of varying lengths increased SD suggesting that the

substitution of negatively charged Asp is crucial for polymerase's SD activity (Piotrowski *et al*., 2019). Chimeric proteins may also obtain an improved SD. In the recent study Milligan *et al*. (2018) constructed a chimeric enzyme by the fusion of Bst LF with high SD activity and extremely thermostable Klentaq to perform the one-pot hot-start LAMP.

1.6 Loop-mediated isothermal amplification

Since the development of PCR, the first nucleic acid amplification method, NAATs have been used in different fields, such as research, clinical medicine, diagnosis of infectious diseases, forensics, gene cloning, and food quality control (Becherer *et al*., 2019). Afterwards, alternative amplification methods have been invented, exemplified by rolling circle amplification (RCA) and nucleic acid sequence-based amplification (NASBA) (Notomi *et al*., 2000; Fakruddin *et al*., 2013; Becherer *et al*., 2019).

Notomi *et al.* (2000) presented a new technique for DNA amplification named loop-mediated isothermal amplification (LAMP). It is a rapid (~20-60 min), single-step, highly sensitive, specific, more resistant to inhibitors, low-cost and efficient method operating at a constant (60-65 °C, Becherer *et al*., 2019) temperature, thus avoiding the necessity of thermal cycling required in PCR (Notomi *et al*., 2000; Fakruddin *et al*., 2013; Xu *et al.*, 2019). The procedure involves a single Bst-like DNA polymerase with high strand displacement (SD) activity and two sets of inner and outer primers, four in total, that target six distinct regions on the target DNA (Notomi *et al*., 2000). The inner primers are the forward inner primer (FIP) and the backward inner primer (BIP), both containing two distinct sequences corresponding to the sense and antisense sequences of the target DNA, F2c and B2c, respectively. The outer primers, forward outer primer (FOP) and backward outer primer (BOP), contain the F3 and B3 sequences complementary to the F3c and B3c, respectively (Notomi *et al*., 2000).

The underlying mechanism behind the loop-mediated isothermal amplification is visualised in Figure 2. First, FIP hybridizes with its F2 region to F2c located on the target DNA initiating the polymerization of the second complementary strand by strand displacing DNA polymerase (extends the FIP and separates the DNA). Subsequently, F3 outer primer binds the F3c (upstream FIP binding site, F2c) thus initiating the strand synthesis and eventually displacing the strand with the bound FIP. This newly synthesized complementary strand with FIP is dissociated and forms a loop at the 5' end due to the

reverse complementary regions F1 and F1c. Next, the released ssDNA with a formed loop is used as a template for BIP, which binds the B2c region of the new released strand and is then extended to synthesize an another new strand. Afterwards, the B3 primer binds the B3c to displace the BIP-primed DNA strand. Once the BIP-primed strand is released, it forms the short dumbbell structure with looped structures on both ends which is used for further LAMP amplification cycling steps. The F1 region with the 3' end is then used for further amplification. The DNA polymerase slides along extending the strand and eventually opens the loop formed at the 5' end producing the stem-loop DNA. In order to finally initiate LAMP cycling, the FIP binds the loop in stem-loop structure in F2c, and F1 region binds F1c which continues the synthesis. Eventually, the single-stranded synthesised structure with regions complementary to the template dumbbell is displaced. The process continues by binding of BIP to the new dumbbell structure.

LAMP can be performed at a speed twice as high if an additional pair of primers, loop primers (forward and backward loop primers), which bind the stem-loop structures and provide additional priming sites, is utilised (Nagamine *et al*., 2002; Becherer *et al*., 2019).

The final products of LAMP cycling are a mixture of stem-loop DNA molecules of varying sizes (lengths) and cauliflower-resembling structures possessing multiple loops (Notomi *et al.*, 2000). Since it was developed, LAMP was used for detection and identification of viruses, bacteria and parasites (Shao et al., 2011; Sun *et al*., 2017).

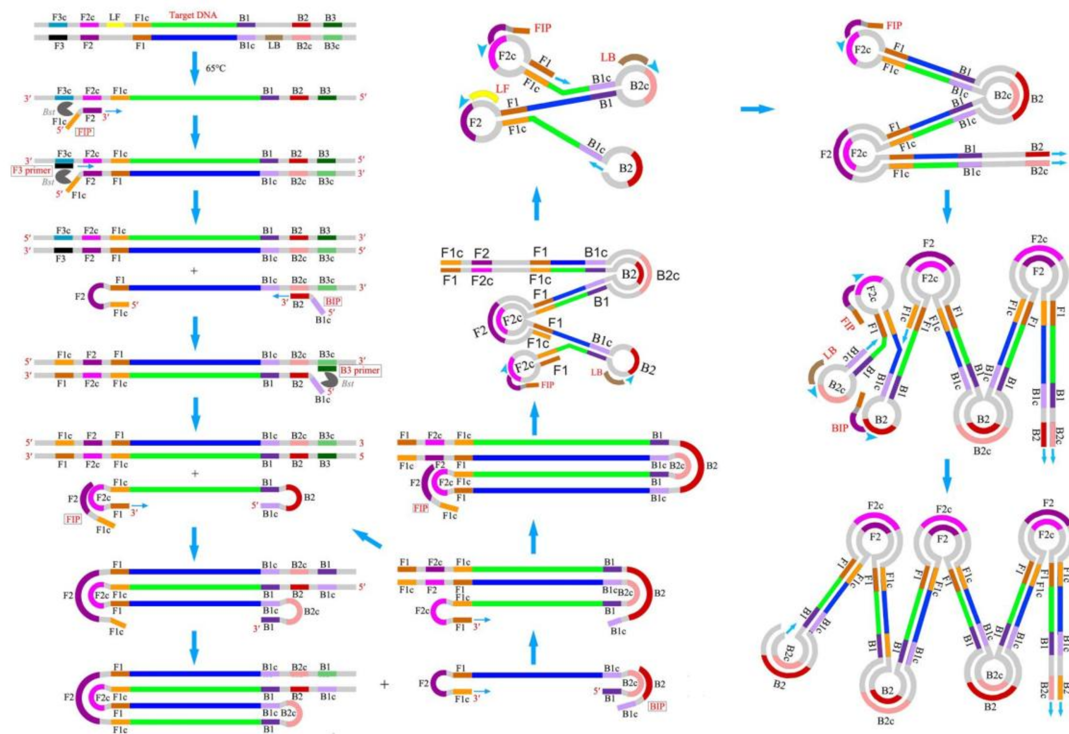


Figure 2. Schematic representation of step-by-step LAMP reaction demonstrating the product synthesis in the form of many dumbbell-structured DNA molecules. The reaction is carried out at 65 °C (Li *et al.*, 2016).

1.7 Nucleic acid amplification inhibitors

DNA polymerases are prone to various inhibitor substances found in complex biological probes which can suppress the enzyme's activity by protein's degradation or blocking its active site (Al-Soud and Radström, 1998; Śpibida *et al.*, 2016). Polymerase activity may be interrupted if enzyme's cofactors (e.g., Mg^{2+}) are inhibited. Inhibitors are also able to directly interfere with DNA. If an inhibitor binds the ss- or dsDNA, the amplification may not take place and the complex can be co-purified (Besetti, 2007).

PCR inhibitors accompanying blood samples may be either natural components (e.g., heme) or added anticoagulants (e.g., EDTA, heparin) (Al-Sound and Ra Radström, 2000). Amplification in whole blood can reduce the PCR sensitivity and cause false-negative results (Gosiewski *et al.*, 2013). Heparin is a natural anticoagulant, a member of complex carbohydrates' family, containing an uncommonly high level of sulphated disaccharide units. Heparin prevents blood clotting because of its unique pentasaccharide sequence which binds antithrombin III (inhibits blood clotting). Heparin inhibits the polymerase activity by competing DNA template binding while hybridising close to the enzyme's

DNA binding site and may also be copurified with DNA (Yokota et al., 1999; Ghadessy et al., 2001). EDTA (ethylenediaminetetraacetic acid) is another anticoagulant used for blood collection. It inhibits the nuclease activity of DNA pol by binding the divalent cation co-factors, usually Mg^{2+} or Ca^{2+}. Its presence can inhibit the amplification by lowering the Mg^{2+} concentration needed for the polymerase activity. In order to overcome the strong inhibition, the cofactor concentration should be increased (Kreader, 1996; Schrader *et al.*, 2012). Heme, a component of hemoglobin, is a strong inhibitor in blood acting by blocking the polymerase active site (Al-Soud and Radström, 1998; Gosiewski *et al.*, 2013). Heme groups contain iron, release of which affects the reaction's pH, inhibits the polymerase activity and quenches the fluorescence of dyes (Al-Sound and Radstrom, 2001; Kermekchiev et al., 2008; Sidstedt *et al.*, 2018).

Application of truncated forms of DNA polymerases has proven its effectiveness in nucleic acid amplification. For instance, Klentaq1 (derived from Taq) with an N-terminal (5'-3' exonuclease domain) deletion is approximately 10-100-fold resistant to whole blood, while the WT Taq is inhibited already by 0.1-1% blood. Klentaq1 also possesses higher fidelity and thermostability (Kermekchiev *et al.*, 2008).

As there is a growing demand in DNA polymerases able to amplify nucleic acids directly in contaminated probes, the production of chimeric enzymes with improved properties is a possible strategy. The engineered proteins are acquired via the fusion of natural polymerases and separate domains carrying the properties and functions of interest. As a result, DNA polymerases with inhibitor tolerance, strand displacement activity, increased thermostability and processivity can be produced (Oscorbin *et al.*, 2017). Another prominent strategy is the site-specific or random mutagenesis providing point mutations (substitutions) to generate mutant proteins (Yamagami *et al.*, 2014). Previously, Kermekchiev *et al* (2008) attempted to generate polymerases insensitive to inhibitors. Taq and Klentaq1 mutants with substitution in position 708 near the fingers domain demonstrated the tolerance of up to 20% of whole blood. In the other study, Arezi *et al* (2014) obtained three conserved point mutations in Taq mutants, L245M, E507K, and F749I for fast-cycling PCR. Substitutions E507K and F749I were located near the thumb and fingers subdomains, respectively. E507K mostly increased the fast-cycling, but has also been implicated in NaCl, blood and EDTA increased resistance. G59W, V155I, and F749I mutations showed a considerable tolerance of about 10-15-fold to blood inhibitors. Interestingly, the E708 mutants obtained an equal resistance to EDTA and heparinized blood. Besides the E708,

the F749I is of a high interest as it is located near the template strand and flanked with amino acids directly interacting with heparin or the template nucleoside.

2 THE AIMS OF THE THESIS

The aim of this study was to purify the mutant Bsm DNA polymerase and test its DNA polymerization activity in the presence of inhibitors.

3 EXPERIMENTAL PART

3.1 MATERIALS AND METHODS

3.1.1 Bacterial strains, plasmids and culture conditions

For expression of pET21a(+) Bsm M496I vector the competent *E.coli* BL21-CodonPlus-RIL, further denoted simply as RIL, and Rosetta, ROS, host strains were used. The bacterial strains used for the present study are described in Table 1. Competent cells were made by Ilja Gaidutšik.

Table 1. Bacterial strains used for the study.

Bacterial strain	Genotype	Source
BL21-CodonPlus-RIL	F- ompT hsdS (rB- mB-) dcm+ Tetr gal λ endA Hte (argU ileY leuW Camr)	*Agilent*
Rosetta (DE3)	F^- *ompT hsdS*$_B$(r_B^- m_B^-) *gal dcm* (DE3) pRARE (Cam^R)	*Novagen*

For the expression of mutant version of Bsm polymerase pET21a(+) Bsm M496I (5443 bp) DNA construct was used. It encoded the mutant LF Bsm DNA polymerase where Methionine (M) was substituted with Isoleucine (I) amino acid at position 496 (designed by Ilja Gaidutšik). Polymerase gene synthesis and its cloning into the pET21a(+) vector together with its verification by sequencing were conducted by *General Biosystems*.

The mutation introduced into sequence was expected to increase the tolerance of LF Bsm to blood-containing biological samples while it is located in a relatively conserved region of family A DNA polymerases. The amino acid change in the same position in Taq enzyme (F749I) sequence improved its inhibitor tolerance to EDTA-blood (<1% vs 2%) as well as heparinized blood (<1% vs 15%). The F749 is located in fingers subdomain near a DNA template strand and is surrounded by amino acids which contact directly with heparin or with template nucleoside on the opposite side of incoming nucleoside triphosphate (Arezi *et al*., 2014). The expression of Bsm M496I is controlled by inducible lacI promoter and could be induced by IPTG (isopropyl-β-D-

thiogalactoside). Bsm M496I contains N-terminal His-tag which could be removed by TEV protease cleavage (Figure 3). The construct was designed by Ilja Gaidutšik.

Figure 3. pET21a(+) Bsm M496I plasmid containing the amino acid substitution at position 496.

As a template for DNA synthesis in DNA polymerase assay (section 3.1.6) was used single-stranded M13 bacteriophage DNA M13mp18 (7249 b) (Messing, 1983) which was ordered from Bayou Biolabs. As a template for LAMP assays (section 3.1.7) pGL3 CDS2 plasmid was used. It contains the cloned region of CDS2 gene fragment from *Chlamydia trachomatis* into a pGL3-Promoter vector (Krõlov *et al*., 2014).

The primers used in this study for the qLAMP reaction were as in Jevtuševskaja *et al.* (2016). For DNA polymerase assay the M13 sequencing primer (5´-CGC CAG GGT TTT CCC AGT CAC GAC) was used.

3.1.1.1 Growing cultures

LB (*Luria broth*): 0.5 % yeast extract, 1.5% agar, 0.5% NaCl and 1% tryptone (solid LB medium) (MacWilliams and Liao, 2006).

LB: 0.5 % yeast extract, 0.5% NaCl and 1% tryptone (liquid medium)

Chloramphenicol (Cam) 25 µg/ml final and ampicillin (Amp) 100 µg/ml final, were added where antibiotic selection was required.

3.1.2 Transformation of expression vectors

The competent cells of RIL and Rosetta strains were thawn on ice. Approximately 50 ng of pET21a(+) Bsm M496I plasmid were added to 50 µl of competent cells, gently mixed and incubated on ice for 30 min. Then, cells were heat shocked at 42°C for 90 sec and immediately put on ice for 5 min. 300 µl of LB media (0,2% glucose and 20mM $MgSO_4$) were added to transformation mixture and incubated at 37°C for 45 min - 1h while being gently shacked. The cell mass was pelleted by centrifugation (1 min; 4000 rpm) on 5415R (*Eppendorf*) centrifuge. After the centrifugation the excess supernatant was removed. Cell pellet was resuspended in residual 100 µl of media cells were cultured on selective LB plates and incubated (upside-down) overnight at 37°C.

3.1.3 Bacterial strain growth conditions

For growth optimization conditions bacterial strains (RIL and Rosetta) were inoculated as overnight starters in 15 ml of LB Amp/Cam, and left in a shaker for overnight growth at 37°C. Optical density (OD) of overnight cultures was measured and cultures were diluted in appropriate volume to OD = 0.15 and grown at 37°C /30°C /25°C until OD = 0.6-0.7. Once the optical density of both cultures has reached 0.6, bacterial strains were induced with various concetrations of isopropyl-β-D-thiogalactoside or IPTG, (1mM, 0,5mM, 0,25mM, 0,1mM) for 2h/4h/6h/8h. After the induction, 15 ml of bacterial cells were pelleted by centrifugation at 4°C and 4000 rpm for 10 min on 5810R centrifuge (*Eppendorf*). After supernatant removal, cells were washed with 1/10 of culture volume of STE (saline-tris-EDTA) buffer (NaCl 100 mM, Tris [pH 8] 10 mM, EDTA 1mM) and centrifugated at 4°C and 4000 rpm for 5 min. Once the washed pellet was collected, the cells were frozen in a liquid nitrogen. For large scale purification 50 ml of overnight Rosetta bacterial culture was diluted in 100ml and 1000 ml with starting OD 0,15 and was grown until OD=0.7-0.8 and protein expression was induced by 0.25 mM IPTG at 37°C for 2h hours.

Polyacrylamide gel electrophoresis and sample preparation (PAGE)

For the analysis of collected cells of RIL and Rosetta strains and protein purification procedures the 10x10 cm 10% acrylamide was used. The probes were denatured in 1x SDS buffer (2% SDS, 4% glycerol, 80mM Tris-HCl [pH 6,8], 0,01% bromphenol blue (BPB), 10mM β-mercaptoethanol) while being heated at 95°C for 5 min. Gel electrophoresis was performed at 120V for 3h on the PowerPac Basic (*BioRad*). As a buffer for gel electrophoresis the protein run buffer (25mM Tris, 0,2M glycine, 0,1% SDS) was used. The control markers for gel electrophoresis were the 5-250 kDa PageRuler Unstained Broad Range Protein Ladder (*Thermo Scientific*) and the 10-250 kDa PageRuler Plus Prestained Protein Ladder (*Thermo Scientific*). Gels were washed 3x with mQ water for 10 min while being slightly shacked to remove SDS traces and stained with Page blue protein staining solution (*ThermoFischer*) overnight.

3.1.4 Purification of Bsm M496I

To obtain the purified Bsm M496I protein, the cell lysis needed to be performed first. Frozen bacterial cells were first thawn on ice and resuspended in 20 ml of LB (50 mM Tris-HCl 7.5, 100 mM NaCl, 5% glycerol, 1 mM EDTA, 10 mM imidazole pH 7.5, 0.1% Triton-X100, 1mM PMSF (*ex temp*), 2 mM β-mercaptoethanol (*ex temp*)) per 1l of culture. After, 2 mg/ml of lysozyme were added into solution, cells were incubated for 30 min on ice. Lyzed cells were then sonicated 2x for 20 sec 50% power and 50% cycle with Bandelin Sonoplus (60 W output) with a further centrifugation at 12000 g for 20 min and transfer into a clean tube. Then, resulted suspension was incubated for 20 min at 60°C in water bath. Subsequently, the lysate was centrifugated at 12000 g for 20 min and transferred into a new tube 2x in order to get a clear lysate. Two supernatants and resulted pellet were sampled. The mixture was filtered through 0.22 µm membrane. Further protein purification procedures were conducted on ÄKTA purifier machine (*GE Heathcare*) under the supervision of Ilja Gaidutšik.

The His Trap High Performance (*GE Healthcare*) 1 ml column was equillibrated with 5 CV (column volume) of washing buffer or WB (50 mM Tris-HCl [pH 7.5], 100 mM NaCl, 5% glycerol, 0.1 mM EDTA, 20 mM imidazol, 0.1 Triton-X100, 1mM PMSF, 2 mM β-mercaptoethanol). Afterwards, the protein was loaded into the column. Next, the column was washed with 10 CV of WB and eluted with 20 CV of continuous imidazole gradient (0%-100% of elution buffer, EB (50 mM Tris-HCl pH 7.5, 100 mM NaCl, 10% glycerol,

0.1 mM EDTA, 500 mM imidazol pH 7.5, 1 mM PMSF, 2 mM β-mercaptoethanol). The resulted fractions were analyzed on SDS PAGE.

The HisTrap peak fractions containing Bsm M496I were combined. The sample was diluted 2x with H0 and loaded onto the HiTrap Heparin HP (*GE Healthcare*) 1 ml column. Then, the column was washed with H50 until it reached the zero baseline. 20 ml (20 CV) gradient of H50 to H1000 was applied. The resulted fractions were analyzed on SDS PAGE All the solution used in ÄKTA purifier were filtered through the 0.22 µm membrane. Finally, the 3.5 ml of peak elution fractions were pipeted into Spectra/Por membrane (*Spectrum*^{Tm}) tube with 10K Mw cut off and dialysed against the Bsm storage buffer (10 mM Tris-HCl [pH 7.1], 50 mM KCl, 0 mM NaCl, 1 mM DTT, 0.1 mM EDTA, 50% glycerol, 0.1% Triton X-100) 3x 200 ml at +4 C with slow stirring, 2x 4h and overnight.

3.1.5 Protein concentration measurement

The protein concentration was measured according to the manufacturer's instructions provided in Pierce^{TM} BCA Protein Assay Kit (*Thermo Scientific*).

3.1.6 DNA polymerase assay for unit measurement

In order to assign the purified Bsm M496I polymerase a unit to test enzyme's activity, the DNA polymerase assay was performed with a M13 sequencing primer (section 3.1.1). The method was performed using the purified mutant Bsm DNA polymerase and commercially available Bsm polymerase (*Thermo Fischer*) as a control and single stranded circular M13mp18 (*Bayou Biolabs*) DNA which acted as a template for DNA amplification.

The steps of the DNA pol assay were as follows. First, the reagents (Bsm DNA pol 0.1U or appropriate dilution of Bsm M496I, 1X Bsm Buffer (20 mM Tris-HCl (pH 8.8), 10 mM KCl, 10 mM (NH4)2SO4, 2 mM MgSO4, 0,1% Tween 20), 10mM MgSO4, 400µM dNTP, 70nM M13 sequencing primer, 30nM M13mp18 ssDNA, 0,1 mg/ml BSA, 0,15 µl [3H]-dTTP (1µCi/µl) (Perkin Elmer)) were mixed on ice in a final volume 10 µl. The (isothermal) polymerization reaction was initiated by placing the tubes with samples in preheated PCR machine (*Applied Biosystems*) and proceeded at 60°C for 30 min. The reactions were stopped with a samples' transfer on ice and addition of 5 µl of 0.125M EDTA. The unincorporated dNTPs were removed using the Spin PCRapace nucleotide removal kit

(*Stratec*). Eventually, the eluted DNA was transferred into scintillation vials with an addition of 5 ml of Scintillation cocktail OPTIPHASE HISAFE 3 (Perkin Elmer). The counts per minute (CPM) in probes was measured by Tri-Carb 2810TR (Perkin Elmer liquid scintillation analyser using 3h_cpm program).

3.1.7 qLAMP

To check the ability of purified protein to amplify DNA in LAMP assay, quantitative LAMP method was applied. Each reaction mixture (10 µl) contained 1 µM mixture of CDS2 primers, 1x Bsm pol buffer (*Thermo Fischer*), 1.4 mM dNTPs, 6 mM $MgSO_4$, 0.8 mM betaine, 0.2 mM evagreen, 0.3 mM rox, 0.32 U Bsm polymerase, EDTA (0 mM, 2 mM, 3 mM and 4 mM) and $10^7/10^6/10^5$ copies of linearized pGL3 CDS2 plasmid. The primers used for qLAMP method are described in section 3.1.1. The reactions were pipeted on 384 well plate and run in 7900HT qPCR machine (*Applied Biosystems*) using following program: 40 minutes 63ºC, 95ºC 10 min. Each reaction was presented as 4 technical replicates. The method was performed once.

3.2 RESULTS

3.2.1 Optimization of protein expression conditions

In order to select the optimal conditions providing the highest recombinant protein expression levels, several experiments were conducted to choose between the RIL and Rosetta *E.coli* strains, induction temperatures, induction time and IPTG titer. For each expression control the induction with 1 mM IPTG was used, which is considered as a stardard concentration ensuring the cell induction. The reason for selecting IPTG (isopropyl-β-D-thiogalactoside) as a chemical inducer of protein expression is that IPTG is not involved in *E.coli* metabolism. Therefore, the removal of a natural inducer (lactose) required for E.coli plasmid lac operon will result in the binding of lac repressor encoded by lacI promoter thus repressing the transcription. In a natural inducer's absence the IPTG is used to allosterically bind lac repressor (promoter) enabling the transcription (*Goldbio*).

First, the protein expression of two transformated RIL and Rosetta strains carrying the pET21a(+) Bsm M496I expression vector was tested in conditions of 37°C incubation and 2h induction with 1mM IPTG.

Since only the Rosetta strain demonstrated the protein expression under the treatment with 1 mM IPTG for 2h, it was chosen as a host strain for further optimization. The obtained length of expressed protein is 68.8 kDa. The data is visualized in Figure 4.

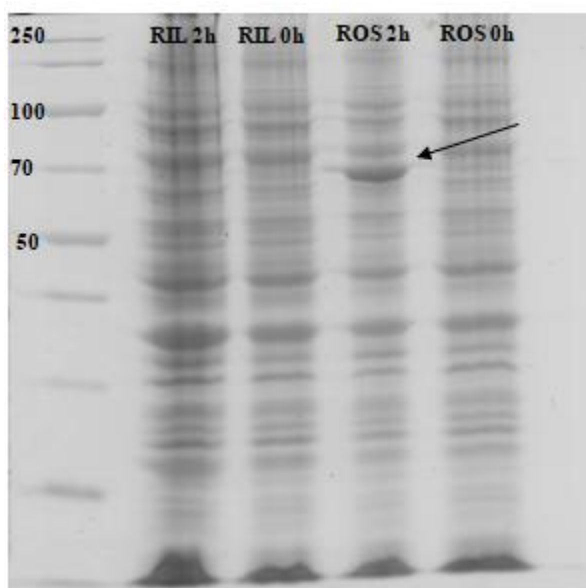


Figure 4. SDS PAGE analysis of Bsm M496I expression in the RIL and Rosetta *E.coli* protein expression strains. As a protein marker, the PageRuler unstained broad range protein ladder (5-250 kDa) was used. The expression of Bsm M496I (68.8 kDa) is pointed with an arrow.

Next, the optimal growth temperature has been selected for protein expression in Rosetta cells. The expression was tested at 37°C/30°C/25°C with 2h induction with 1 mM IPTG and analyzed on SDS PAGE. The highest level of expression can be seen at 37°C. Therefore, this incubation temperature has been selected for the testing of the rest growth conditions to be optimized. The data is represented in Figure 5.

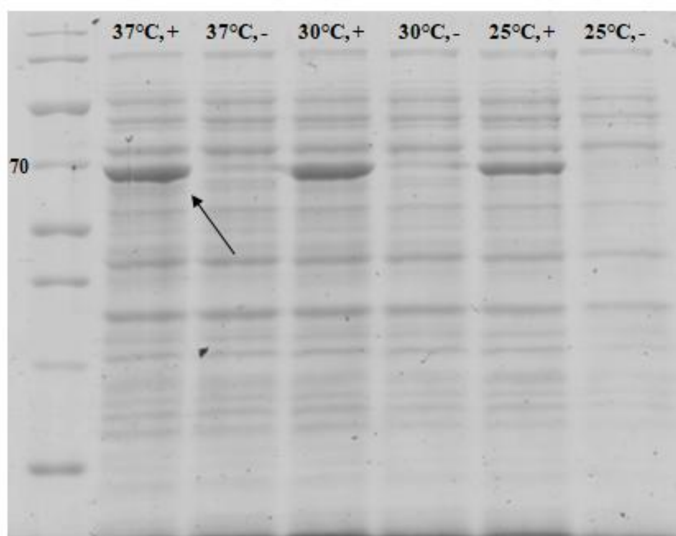


Figure 5. Analysis of pre-induction growth at three different temperatures (37°C/30°C/25°C). As a protein marker, the PageRuler unstained broad range protein lad-

der (5-250 kDa) was used. The lanes marked with “+” were subjected to the 2h induction. The band representing the best protein expression is pointed with an arrow.

Subsequently, the optimal induction time has been selected from a range of 2h, 4h, 6h and 8h of 1 mM IPTG treatment (Figure 6). After the SDS PAGE analysis it was concluded that 2h of IPTG induction will be sufficient for expression of mutant protein (Figure 6).

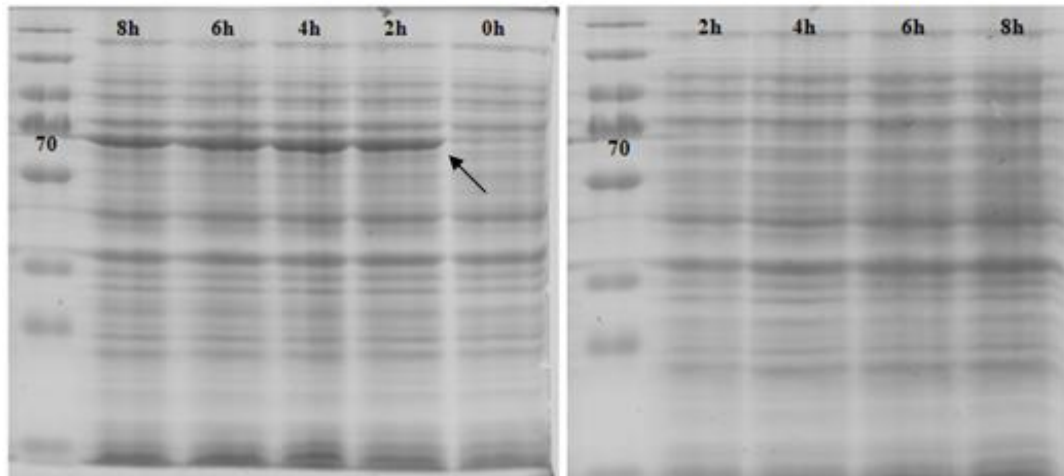


Figure 6. Analysis of 2h/4h/6h/8h IPTG treatment of Rosetta cells. As a protein marker, the PageRuler unstained broad range protein ladder (5-250 kDa) was used. Five lanes on the left side contain the induced cells. Four lanes on the right side were used in order to ensure no leaking at appropriate time during the expression.

Finally, after the induction temperature and time have been optimized for Rosetta strain, the optimal IPTG concentration has been selected. The cells were exprosed to four different concentrations, namely 1 mM, 0.5 mM, 0.25 mM and 0.1 mM IPTG. As can be seen from the SDS PAGE, the 0.25 mM IPTG is sufficient for strong Bsm M496I expression (Figure 7). The 0.25 mM IPTG has been selected as an optimal inducing concentration.

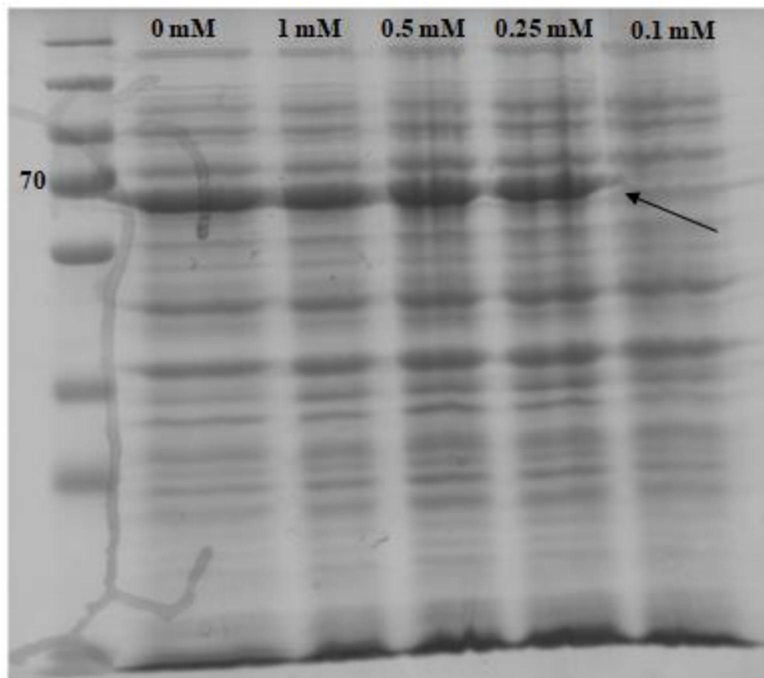


Figure 7. 1mM/0.5mM/0.25mM/0.1mM IPTG induction at 37°C for 2h. As a protein marker, the PageRuler unstained broad range protein ladder (5-250 kDa) was used. The band of interest is pointed with an arrow.

3.2.2 Large-scale protein purification

As one of the aims of the current work was to purify the recombinant protein, the Bsm M496I mutant was purified with the Ni Sepharose and Heparin HP columns.

Commonly, the recombinant proteins are purified by the fusion of an affinity tag to their N- or C-terminus to enable further purification. His tag is one of the most widely used fusion tags for protein purification and expression. It consists of six (or more) histidine residues. As it is a small molecule not posessing charge at a physiological pH, His tag usually does not influence either protein's structure or function. The method relies on immobilized metal ions, commonly Ni^{2+} or Cu^{2+}, on chomatographic matrix (Spriestersbach *et al.*, 2015).

Chromatograms below represent the process of Bsm M496I protein purification on ÄKTA purifier machine. The protein was first purified with the Ni Sepharose column binding His-tagged enzymes, and afterwards with the Heparin HP column which mimicks DNA thus binding the DNA-binding proteins.

The his-tagged protein which has bound HisTrap column was eluted mostly when the 20mM – 500mM imidazole gradient was applied. The BsmM496I mutant was eluted from

the HisTrap column at an imidazole concentration in the range of 150mM – 300mM. The fractions collected were (Figure 8).

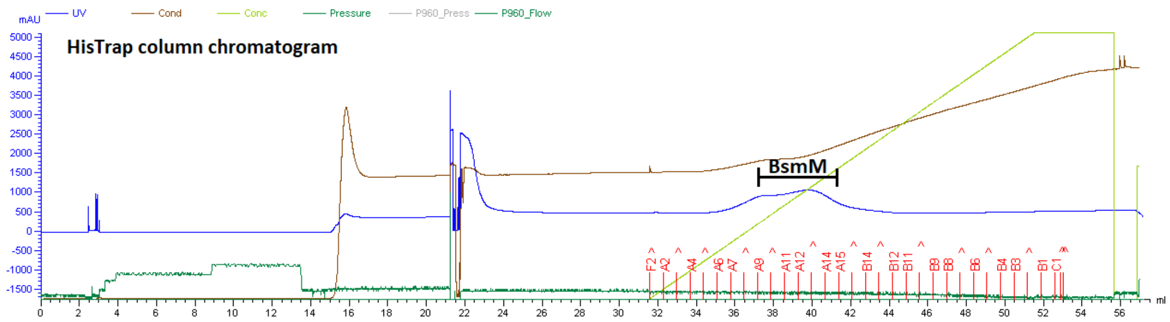


Figure 8. HisTrap column chromatography. The UV absorbance representing the amount of protein providing a signal is shown in blue. The imidazole gradient is shown in green. Collected fractions of protein can be seen in red. The highest amount of released protein is shown in the peak BsmM.

The pooled protein fractions from the HisTrap column were loaded into the Heparin column, proceeding with a protein elution at a NaCl concentration in the range of 510 mM - 620 mM (Figure 9). The single narrow peak with the highest purified protein release was obtained as a highly concentrated fraction. The fractions collected were A15, B14 and B13, each of volume of 0.7 ml.

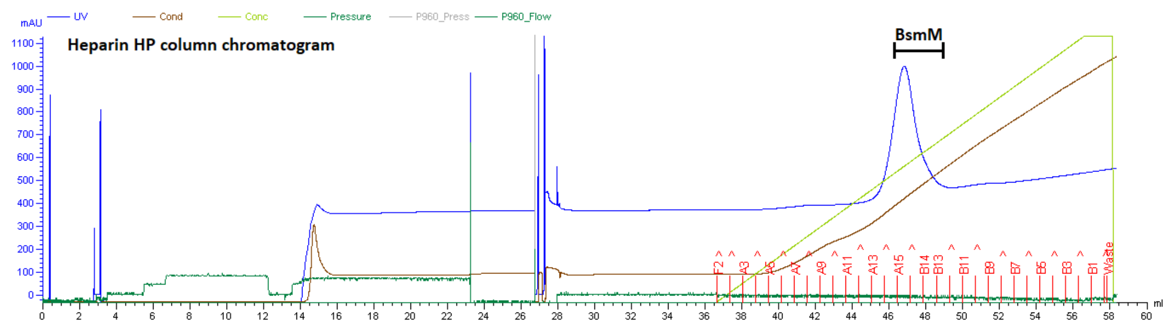


Figure 9. Heparin HP column chromatography. The UV absorbance and NaCl gradient are show in blue and green, respectively. Collected fractions are depicted in red. The peak on the graph indicates the largest portion of released protein.

As the final step of obtaining a purified protein for further use, the dialysis procedure was applied in order to change the purification buffer to the protein storage buffer and

moderately concentrate the protein fractions. The concentration of the final 1.5 ml BsmM496I-containing fraction obtained was 1.09 mg/ml. The total protein yield obtained from 1 liter is 1.635 mg.

PAGE analysis of collected fractions resulted from protein purification with HisTrap and Heparin HP columns is shown on Figure 10.

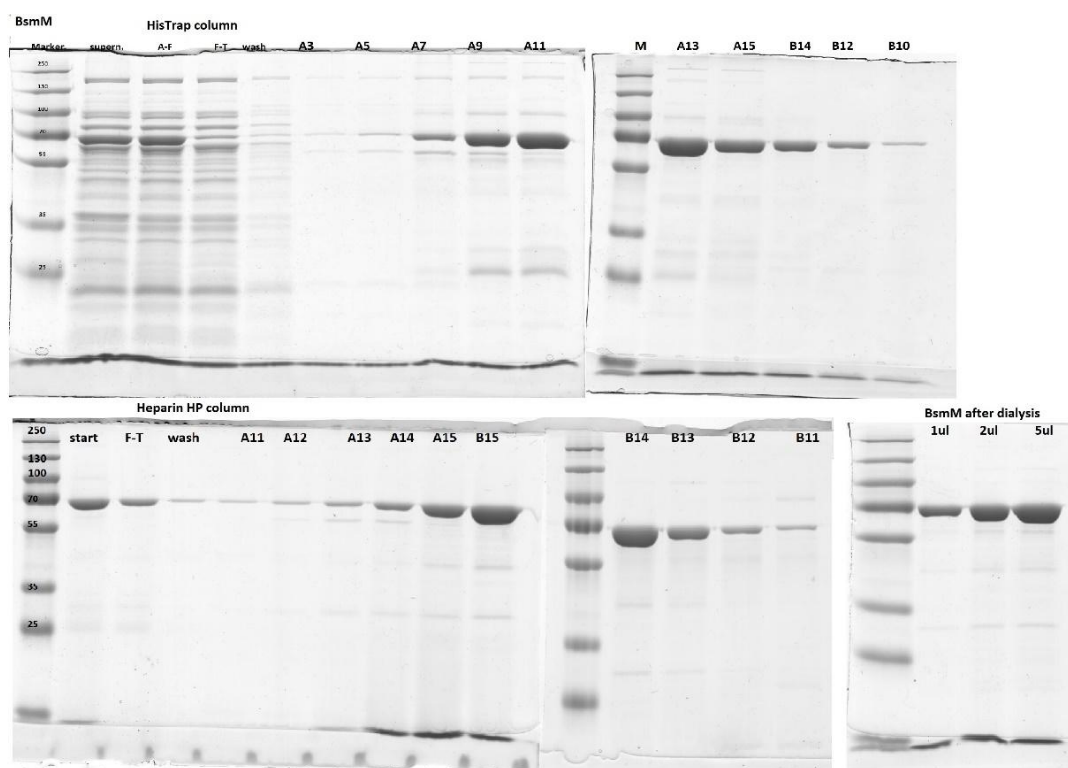


Figure 10. SDS PAGE gels representing the purification stages with the collected fractions from the HisTrap and Heparin column purification. The final result of Bsm M496I purificaton was obtained after dialysis, with the data represented on the last portion of the picture. The 10-250 kDa PageRuler Plus Prestained Protein Ladder (*Thermo Scientific*) was used as a marker.

3.2.3 Bsm M496I mutant unit measurement

The idea behind the DNA polymerase assay is in the separation of incorporated dNTPs into the M13mp18 ssDNA from unincorporated ones. The labelled [^{3}H]-dTTP acts as a radioactive marker for measurement of the total amount of dNTPs and the number of dNTPs incorporated into ssDNA, after the separation with unincorporated ones. There are two control probes not containing the Bsm M496I DNA polymerase. The first was measured in scintillation counter directly to estimate the total CPMs in all reactions (total), while in the second control the unbound dNTPs were separated in order to assess the level of unspecific

binding by calculating the CPMs (background). The number of incorporated dNTPs is calculated according to the following formula:

$$\frac{reaction - background}{total} * amount\ of\ dNTPs\ (nmol)$$

The protein activity is measured in units (U/µl). As a reference, the activity of commercial LF Bsm DNA polymerase from Thermo Fischer was used: 1U denoted as the amount of 10 nmol of dNTPs incorporated by an enzyme at 60°C in 30 min. Thermo Fischer's unit definition of 36.6 units (U/µl) was assigned to the purified protein.

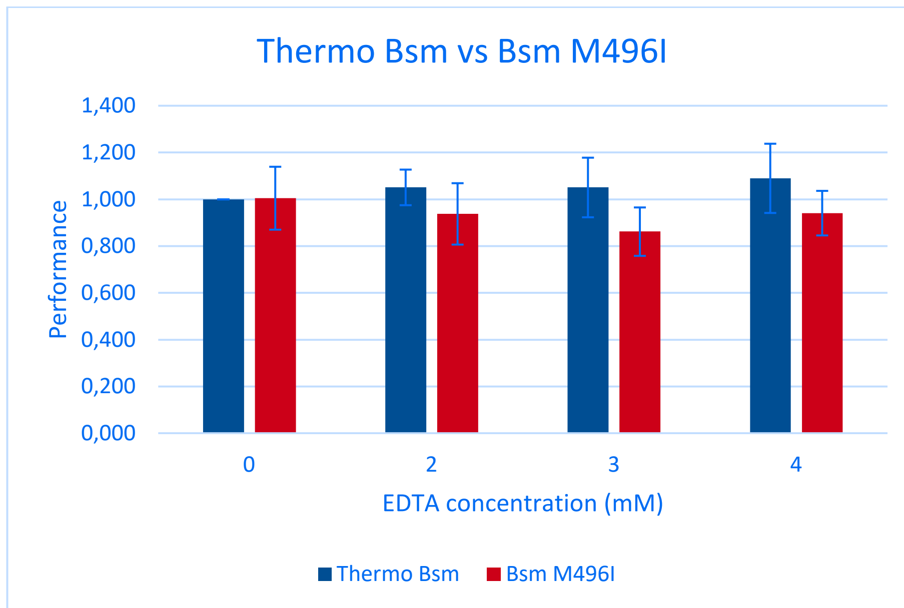
3.2.4 qLAMP performance

The final step of the current study was the quantitative loop-mediated isothermal amplification (qLAMP) in order to assess the designed mutant's activity in the presence of three different concentrations of blood collection related inhibitor, EDTA. As a reference, the commercially available Thermo Bsm polymerase was used. The performance of both polymerases was assessed as a time taken by a polymerase to amplify enough copies of DNA template of varying concentrations to provide a sufficient signal above the threshold, termed as C_t, or time to result.

The data from the qLAMP method demonstrated that the mutation at position 496 in fingers domain replacing the methionine amino acid with isoleucine resulted in slight decrease in time to result which may have lead to an increased tolerance to EDTA. As seen from the Table 2, the Bsm M496I mutant performed better in the presence of all tested EDTA concentrations, compared to commercial Bsm polymerase.

Both reference and constructed mutant polymerases demonstrated an equal performance in the absence of inhibitor. At the 2 mM EDTA added in a sample, the Bsm M496I mutant needed 10.8% less time for reaching the C_t value, 18% less time at 3 mM EDTA and 13.65% at 4 mM EDTA. At the same time, at the 2-4 mM concentrations of EDTA the Thermo' alternative needed more time in order to reach the C_t value, thus indicating its sensitibility to the inhibitor.

Table 2. 0/2/3/4 mM of EDTA tolerance in large fragments of commercial Thermo Bsm polymerase and designed Bsm M496I mutant shown in blue and red, respectively. Error bars represent the standard deviation.



3.3 DISCUSSION

Nowadays, isothermal nucleic acid amplification techniques have gained a significant recognition in both science and medicine. The design of recombinant proteins with a high strand displacement activity and decreased sensitivity to various inhibitors found in biological probes is a current trend in medicine, diagnostics and research. This study aimed to produce an improved version of Bsm DNA polymerase LF able to withstand a type of anticoagulant used for blood collection, EDTA.

In this study, the designed Bsm M496I with a single amino acid substitution of methionine (M) with isoleucine (I) amino acid at position 496 in a relatively conserved region among A-family polymerases was expected to increase its tolerance to EDTA-blood.

In order to obtain the designed mutant, several experiments have been conducted to identify the optimal expression conditions of a protein. The temperature of 37°C has been selected as an optimal for the growth and expression of the selected strain with inserted pET21a(+) construct encoding the LF Bsm M496I. This result allowed to continue all further incubations at the same and unified temperature. As an optimal induction time the 2h induction has been chosen, providing with the time efficiency and exempting from the time-consuming incubations. Next, as a protein demonstrated a high level of expression at only 0.25 mM IPTG concentration, it was decided to induce the protein at this concentration thus providing with a cost-efficiency explained by the reagent's high cost.

The obtained protein grown at developed growth conditions was purified with the His-tagged and Heparin columns on ÄKTA purifier. The total yield of the purified protein was 1.635 mg which is sufficient enough for preliminary studies of protein inhibitor tolerance. Also, the analysis of SDS PAGE with the collected fractions from the HisTrap and Heparin column purification demonstrated that the obtained Bsm M496I mutant is not perfectly pure. The possible solution to obtain the absolutely pure enzyme is the usage of the third column. For example, size-exclusion chromatography column, which separates proteins based on their molecular weight.

The purified M496I mutant has demonstrated the ability to amplify the M13mp18 ssDNA from the M13 bacteriophage in the presence of EDTA, the anticoagulant interfering with DNA polymerase activity by decreasing the Mg^{2+} availability for DNA polymerase. The designed enzyme has amplified the single stranded DNA better at all tested EDTA concentrations in comparison with the commercial Bsm polymerase. Also, by comparing both enzymes' performance in the absence of inhibitor, it has been observed that the inserted

point mutation has not interfered with the LF Bsm M496I polymerase's performance as the enzyme's activity has not decreased in comparison with Thermo's Bsm alternative. These observations suggest that the mutant may be promising for nucleic acid amplification in EDTA-contaminated blood probes. However, there are other possible blood-related inhibitors, such as heme and heparin, which also may interfere with the polymerase's activity. Therefore, further investigations are required to fully characterize this enzyme.

SUMMARY

The isothermal nucleic acid amplification testing (INAAT) is a modern powerful tool for rapid detection and identification of various pathogenic organisms found in biological samples. The method relies on DNA polymerases with high strand displacement activity and biological inhibitor tolerance.

Based on the previous investigations by Arezi *et al.* (2014), in which the collection of designed Taq mutants with introduced substitutions has acquired an improved EDTA-blood and heparinized blood tolerance, it was attempted to create a Bsm mutant with a similar single point mutation at location 496 in a relatively conserved region of A-family polymerases located in fingers subdomain interacting with template nucleoside.

During the study, the protein's expression was optimized, the enzyme was purified with a total yield of 1.635 mg, units of the enzyme were measured and qLAMP method was performed.

In an attempt to generate a LF Bsm DNA polymerase mutant with higher resistance to blood-related inhibitor substances, the rather positive results have been obtained.

However, as in the current study the constructed mutant's tolerance has only been characterized for one inhibitor in blood, and due to the fact that the quantitative LAMP testing has been performed not sufficient amount of times, further investigations should be conducted in order to ensure the EDTA tolerance as well as to assess the full potential of Bsm M496I for applications in nucleic acid amplification.

REFERENCES

[1] Thermo Fischer Scientific. Bsm DNA polymerase, large fragment (8U/µl) https://www.thermofisher.com/order/catalog/product/EP0691#/EP0691?CID=BID_R01_PJT4624_DB_REG_OP_KT_1019 (20.05.2020)

Albà M. Replicative DNA polymerases. *Genome Biology*, 2001, vol. 2, no. 1.

Aliotta JM.; Pelletier JJ.; Ware JL.; Moran LS.; Benner JS.; Kong H. Thermostable Bst DNA polymerase I lacks a 3'-5' proofreading exonuclease activity. *Genet Anal,* 1996, vol. 12, no. 5-6, pp. 185-195.

Al-Soud WA.; Radström P. Capacity of nine thermostable DNA polymerases to mediate DNA amplification in the presence of PCR-inhibiting samples. *Appl Enviro Microbiol*, 1998, vol. 64, no. 10, pp. 3748-53.

Al-Sound WA.; Radström P. Purification and Characterization of PCR-Inhibitory Components in Blood Cells. *Journal of Clinical Microbiology*, 2001, vol. 39, no. 2, pp. 485-493.

Arezi B.; McKinney N.; Hansen C.; Cayouette M.; Fox J.; Chen K.; Lapira J.; Hamilton S.; Hogrefe H. Compartmentalized self-replication under fast PCR cycling conditions yields Taq DNA polymerase mutants with increased DNA-binding affinity and blood resistance. Front. Microbiol, 2014, vol. 5, no. 408.

Bebenek A.; Ziuzia-Graczyk I. Fidelity of DNA replication - a matter of proofreading. *Current Genetics*, 2018, vol. 64, pp. 985-996.

Becherer L.; Borst N.; Bakheit M.; Frischmann S.; Zengerle R.; von Stetten F. Loop-mediated isothermal amplification (LAMP) – review and classification of methods for sequencespecific detection. *Analytical methods*, 2020.

Beese LS.; Friedman JM.; Steitz TA. Crystal structures of the Klenow fragment of DNA polymerase I complexed with deoxynucleoside triphosphate and pyrophosphate. *Biochemistry*, 1993, vol. 32, no. 51, pp. 14095-14101.

Berdis AJ. Inhibiting DNA polymerases as a therapeutic intervention against canser. *Front. Mol. Biosci.*, 2017, vol.4, no. 78.

Berdis AJ. Inhibiting DNA polymerases as a therapeutic intervention against cancer. *Front. Mol. Biosci*, 2017, vol. 4, no. 78.

Blanco L.; Salas M. Characterization and purification of a phage phi 29-encoded DNA polymerase required for the initiation of replication. *Proc Natl Acad Sci U S A*, 1984, vol. 81, no. 17, pp. 5325-5329.

Bloom LB.; Goodman MF. Polymerase processivity: measurement and mechanisms. *ENCYCLOPEDIA OF LIFE SCIENCES*, 2001.

Braithwaite DK.; Ito J. Compilation, alignment, and phylogenetic relationships of DNA polymerases. *Nucleic Acids Res*, 1993, vol. 21, no. 4, pp. 787-802.

Brenner S.; Miller JH. Encyclopedia of genetics. 1st edition. 2001.

Burgers PMJ.; Koonin E.; Bruford E.; Blanco L.; Burtis KC.; Christman MF.; Copeland WC.; Friedberg EC.; Hanaoka F.; Hinkle DC.; Lawrencce W.; Nakanishi M.; Ohmori H.; Prakash L.; Prakash S.; Reynaud CA.; Sugino A.; Todo T.; Wang Z.; Weill JC.; Woodgate R. Eukaryotic DNA polymerases: proposal for a revised nomenclature. *The Journal of Biological Chemistry*, 2001, vol. 276, no. 47, pp. 43487-43490.

Chander Y.; Koelbl J.; Puckett J.; Moser M.; Kingele A.; Liles M.; Carrias A.; Mead D.; Schoenfeld T. A novel thermostable polymerase for RNA and DNA loop-mediated isothermal amplification (LAMP). *Front. Microbiol.*, 2014, vol. 5, no. 395.

Chen CY. DNA polymerases drive DNA sequencing-by-synthesis technologies: both past and present. *Front Microbiol.*, 2014, vol. 5, no. 305.

Fakruddin Md.; Mannan SB.; Chowdury A.; Mazumdar RM.; Hossain Md.; Islam S.; Chowdury Md. Nucleic acid amplification: Alternative methods of polymerase chain reaction. *Journal of Pharmacy and Bioallied Sciences*, 2013, vol. 5, no. 4, pp. 242-252.

Fitzgerald-Hayes M.; Reichsman F. DNA and Biotechnology. 2009.

Franklin MC.; Wang J.; Steitz TA. Structure of the replicating complex of a Pol α family DNA polymerase. *Cell*, 2001, vol. 105, pp. 657-667.

Garcia-Diaz M.; Bebenek K. Multiple Functions of DNA Polymerases. *Critical Reviews in Plant Sciences*, 2007, vol. 26, no. 2, pp.105–122

Ghadessy FJ.; Ong JL.; Holliger P. Directed evolution of polymerase function by compartmentalized self-replication. *Proc Natl Acad Sci U S A*, 2001, vol. 98, no. 8, pp. 4552-7.

Gosiewski T.; Szala L.; Pietrzyk A.; Brzychczy-Wloch M.; Heczko PB.; Bulanda M. Comparison of methods for isolation of bacterial and fungal DNA from human blood. *Current Microbiology*, 2013, vol. 68, pp. 149-155.

Jevtuševskaja J.; Uusna J.; Andresen L.; Krõlov K.; Laanpere M.; Grellier T.; Tulp I.; Langel Ü. Combination with antimicrobial peptide lyses improves loop-mediated isothermal amplification based method for Chlamydia trachomatis detection directly in urine sample. *BMC Infect Dis*, 2016, vol. 16, no. 329.

Johansson E.; Dixon N. Replicative DNA Polymerases. *Cold Spring Harb Perspect Biol.*, 2013, vol. 5, no. 6.

Joyce CM.; Kelley WS.; Grindley N. Nucleotide sequence of the Escherichia coli polA gene and primary structure of DNA polymerase I. *THE JOURNAL OF BIOLOGICAL CHEMISTRY*, 1981, vol. 257, no. 4, pp. 1958-1964.

Kaneko H.; Kawana T.; Fukushima E.; Suzutani T. Tolerance of loop-mediated isothermal amplification to a culture medium and biological substances. J Biochem Biophys Methods, 2007, vol. 70, no. 3, pp. 499-501.

Kaushik N.; Pandey VN.; Modak MJ. Significance of the O-Helix Residues of Escherichia coli DNA Polymerase I in DNA Synthesis: Dynamics of the dNTP Binding Pocket. *Biochemistry*, 1996, vol. 35, pp. 7256-7266.

Kermekchiev MK.; Kirilova LI.; Vail EE.; Barnes WM. Mutants of Taq DNA polymerase resistant to PCR inhibitors allow DNA amplification from whole blood and crude soil samples. *Nucleic Acids Research*, 2009, vol. 37, no. 5.

Kornberg A. DNA replication. 1984, vol. 9.

Kreader CA. Relief of amplification inhibition in PCR with bovine serum albumin or T4 gene 32 protein. *Appl Environ Microbiol,* 1996, vol. 62, no. 3, pp. 1102-1106.

Krõlov K.; Frolova J.; Tudoran O.; Suhorutsenko J.; Lehto T.; Sibul H.; Mäger I.; Laanpere M.; Tulp I.; Langel Ü. Sensitive and Rapid Detection of Chlamydia trachomatis by Recombinase Polymerase Amplification Directly from Urine Samples. J Mol Diagn, 2014, vol. 16, no. 1, pp. 127-35.

Kucera RB.; Nichols NM. DNA-Dependent DNA Polymerases. *Current Protocols in Molecular Biology*, 2008, vol. 84.

Laos R.; Thomson JM.; Benner S. DNA polymerases engineered by directed evolution to incorporate non-standard nucleotides. *Front Microbiol*., 2014, vol. 5, no. 565.

Lawyer FC.; Stoffel S.; Saiki RK.; Myambo K.; Drummond R.; Gelfand DH. Isolation, characterization, and expression in Escherichia Coli of the DNA polymerase gene from Thermus Aquaticus. *J Biol Chem*, 1989, vol. 264, no. 11, pp. 6427-37.

Li J.; Xiong C.; Liu Y.; Liangand J.; Zhou X. Loop-Mediated Isothermal Amplification (LAMP): Emergence As an Alternative Technology for Herbal Medicine Identification. *Front Plant Sci*., 2016, vol. 7.

Ma Y.; Zhang B.; Wang M.; Ou Y.; Wang J.; Li S. Enhancement of polymerase activity of the large fragment in DNA polymerase I from Geobacillus stearothermophilus by site-directed mutagenesis at the active site. *BioMed Research International*, 2016, vol. 2016.

MacWilliams MP.; Liao MK. Luria broth (LB) and luria agar (LA) media and their uses protocol. 2006.

Mead DA.; Pey NK.; Herrnstadt C.; Marcil RA.; Smith LM. A universal method for the direct cloning of PCR amplified nucleic acid. *Bio/Technology*, 1991, vol. 9, pp. 657-663.

Messing J. New M13 vectors for cloning. *Methods Enzymol*, 1983, vol. 101, pp. 20-78.

Milligan JN.; Shroff R.; Garry DJ.; Ellington AD. Evolution of a thermophilic strand-displacing polymerase using high-temperature isothermal compartmentalized self-replication. *Biochemistry*, 2018, vol. 57, pp. 4607-4619.

Nagamine K.; Hase T.; Notomi T. Accelerated reaction by loop-mediated isothermal amplification using loop primers. *Molecular and Cellular Probes,* 2002, vol. 16, no. 3, pp. 223-229.

Notomi T.; Okayama H.; Masubuchi H.; Yonekawa T.; Watanabe K.; Amino N.; Hase T. Loop-mediated isothermal amplification of DNA. *Nucleic Acids Res*., 2000, vol. 28, no. 12.

Ollis DL.; Brick P.; Hamlin R.; Xuong NG.; Steitz TA. Structure of large fragment of Escherichia coli DNA polymerase I complexed with dTMP. *Nature*, 1985, vol. 313, pp. 762-766.

Oscorbin IP.; Belousova EA.; Boyarskikh UA.; Zakabunin AI.; Khrapov EA.; Filipenko ML. Derivatives of Bst-like Gss-polymerase with improved processivity and inhibitor tolerance. Nucleic Acids Res, 2017, vol. 45, no. 16, pp. 9595-9610.

Oscorbin IP.; Belousova EA.; Zakabunin A.; Boyarskikh UA.; Filipenko M. Comparison of fluorescent intercalating dyes for quantitative loop-mediated isothermal amplification (qLAMP). *Biotechniques*, 2016, vol. 61, no. 1, pp. 20-5.

Patel PH.; Loeb LA. Getting a grip on how DNA polymerases function. *Nature Structural Biology*, 2001, vol. 8, no.8, pp. 656-659.

Patel PH.; Suzuki M.; Adman E.; Shinkai A.; Loeb LA. Prokaryotic DNA polymerase I: evolution, structure, and "base flipping"" mechanism for nucleotide selection. *J. Mol. Biol.*, 2001, vol. 308, no. 5, pp. 823-837.

Piotrowski Y.; Gurung MK.; Larsen AN. Characterization and engineering of a DNA polymerase reveals a single amino-acid substitution in the fingers subdomain to increase strand-displacement activity of A-family prokaryotic DNA polymerases. *BMC Molecular and Cell Biology*, 2019, vol. 20, no.31.

Riggs MG.; Tudor S.; Sivaram M.; McDonough SH. Construction of single amino acid substitution mutants of cloned Bacillus stearothermophilus DNA polymerase I which lack 5'-3' exonuclease activity. *Biochim Biophys Acta*, 1996, vol. 1307, no. 2, pp. 178-86.

Rothwell PJ.; Waksman G. Structure and mechanism of DNA polymerases. *Adv Protein Chem*, 2005, vol. 71, pp. 401-40.

Saiki RK.; Gelfand DH.; Stoffel S., Scharf SJ.; Higuchi R.; Horn GT.; Mullis KB.; Erlich HA. Primer-directed enzymatic amplification of DNA with a thermostable DNA polymerase. *Science*, 1988, vol. 239, no. 4839, pp. 487-91.

Schrader C.; Schielke A.; Ellerbroek L.; Johne R. PCR inhibitors – occurence, properties and removal. *J Appl Microbiol*, 2012, vol. 113, no. 5, pp. 1014-26.

Shanbhag V.; Sachdev S., Flores JA; Modak MJ; Singh K. Family A and B DNA Polymerases in Cancer: Opportunities for Therapeutic Interventions. *Biology (Basel)*, 2018, vol. 7, no. 1.

Shao Y.; Zhu S.; Jin C.; Chen F. Development of multiplex loop-mediated isothermal amplification-RFLP (mLAMP-RFLP) to detect Salmonella spp. and Shigella spp. in milk. *International journal of food microbiology*, 2011, vol. 148, no. 2, pp. 75-9.

Sidstedt M.; Hedman J.; Romsos EL.; Waitara L.; Wadsö L., Steffen CR.; Vallone PM.; Rådström P. Inhibition mechanisms of hemoglobin, immunoglobulin G, and whole blood in digital and real-time PCR. *Analytical and Bioanalytical Chemistry*, 2018, vol. 410, pp. 2569-2583.

Singh K.; Srivastava A.; Patel SS.; Modak MJ. Participation of the Fingers Subdomain of Escherichia coli DNA Polymerase I in the Strand Displacement Synthesis of DNA. J Biol Chem, 2007, vol. 282, no. 14, pp. 10594-604.

Śpibida M.; Krawczyk B.; Olszewski M.; Kur J. Modified DNA polymerases for PCR troubleshooting. J Appl Genet., 2016, vol. 58, no. 1, pp. 133-142.

Spriestersbach A.; Kubicek J.; Schäfer F.; Block H.; Maertens B. Purification of His-tagged proteins. *Methods Enzymol*, 2015, vol. 559, pp. 1-15.

Steitz T.A.; Beese L.; Freemont P.S.; Friedman J.M.; Sanderson M.R. Structural studies of Klenow fragment: an Enzyme with two active sites. *Cold Spring Harb Symp Quant Biol*., 1987, vol. 52, pp. 465-471.

Steitz TA. DNA Polymerases: structural diversity and common mechanisms. *THE JOURNAL OF BIOLOGICAL CHEMISTRY*, 1999, vol. 274, no. 18, pp. 17395-17398.

Steitz TA.; Yin YW. Accuracy, lesion bypass, strand displacement and translocation by DNA polymerases. *Philos Trans R Soc Lond B Biol Sci*, 2004, vol. 359, no. 1441, pp. 17-23.

Sun Y.; Tian H.; Liu C.; Sun Y.; Li Z. One-step detection of microRNA with high sensitivity and specificity via target-triggered loop-mediated isothermal amplification (TT-LAMP). *Chem Commun (Camb)*, 2017, vol. 53, no. 80, pp. 11040-11043.

Tripathi G. Cellular and Biochemical Science. 2010.

Wu WJ.; Yang W.; Tsai MD. How DNA polymerases catalyse replication and repair with contrasting fidelity. *Nature Reviews Chemistry*, 2017, vol. 1, no. 9.

Yamagami T.; Ishino S.; Kawarabayasi Y.; Ishino Y. Mutant Taq DNA polymerases with improved elongation ability as a useful reagent for genetic engineering. *Front Microbiol*, 2014, vol. 5.

Yang W.; Gao Y. Translesion and Repair DNA Polymerases: Diverse Structure and Mechanism. *Annual Review of Biochemistry*, 2018, vol. 87, no. 1, pp. 239-261.

Yokota M.; Tatsumi N.; Nathalang O.; Tsuda I. Effects of heparin on polymerase chain reaction for blood white cells. *J Clin Lab Anal*, 1999, vol. 13, no. 3.

NON-EXCLUSIVE LICENCE TO REPRODUCE THESIS AND MAKE THESIS PUBLIC

I, Anastasia Kolosova,

1. herewith grant the University of Tartu a free permit (non-exclusive licence) to:

1.1. reproduce, for the purpose of preservation, including for adding to the DSpace digital archives until the expiry of the term of copyright, and

1.2. make available to the public via the web environment of the University of Tartu, including via the DSpace digital archives, under the Creative Commons licence CC BY NC ND 3.0, which allows, by giving appropriate credit to the author, to reproduce, distribute the work and communicate it to the public, and prohibits the creation of derivative works and any commercial use of the work from ***20/05/2025*** until the expiry of the term of copyright,

„Purification of Bsm DNA polymerase mutant and characterization of its biological sample inhibitor tolerance",

supervised by *Ilja Gaidutšik*.

2. I am aware of the fact that the author retains the rights specified in p. 1.

3. I certify that granting the non-exclusive licence does not infringe other persons' intellectual property rights or rights arising from the personal data protection legislation.

Anastasia Kolosova

20/05/2020