Compartmentalized self-replication (CSR) selection of Thermococcus litoralis Sh1B DNA polymerase for diminished uracil binding

Agne Tubeleviciute and Remigijus Skirgaila

R&D, Fermentas UAB, Graiciuno 8, Vilnius LT-02241, Lithuania

1To whom correspondence should be addressed.
E-mail: remigijus.skirgaila@fermentas.lt

Received November 19, 2009; revised March 24, 2010; accepted May 3, 2010

Edited by Dan Tawfik

The thermostable archaeal DNA polymerase Sh1B from Thermococcus litoralis has a typical uracil-binding pocket, which in nature plays an essential role in preventing the accumulation of mutations caused by cytosine deamination to uracil and subsequent G–C base pair transition to A–T during the genomic DNA replication. The uracil-binding pocket recognizes and binds uracil base in a template strand trapping the polymerase. Since DNA replication stops, the repair systems have a chance to correct the promutagenic event. Archaeal family B DNA polymerases are employed in various PCR applications. Contrary to nature, in PCR the uracil-binding property of archaeal polymerases is disadvantageous and results in decreased DNA amplification yields and lowered sensitivity. Furthermore, in diagnostics qPCR, RT-qPCR and end-point PCR are performed using dNTP mixtures, where dTTP is partially or fully replaced by dUTP. Uracil-DNA glycosylase treatment and subsequent heating of the samples is used to degrade the DNA containing uracil and prevent carryover contamination, which is the main concern in diagnostic laboratories. A thermostable archaeal DNA polymerase with the abolished uracil binding would be a highly desirable and commercially interesting product. An attempt to disable uracil binding in DNA polymerase Sh1B from T. litoralis by generating site-specific mutants did not yield satisfactory results. However, a combination of random mutagenesis of the whole polymerase gene and compartmentalized self-replication was successfully used to select variants of thermostable Sh1B polymerase capable of performing PCR with dUTP instead of dTTP.

Keywords: archaeal polymerase/CSR/cytosine deamination/read-ahead/uracil binding

Introduction

The thermostable archaeal family B DNA polymerases are widely used in different PCR applications. Even though the properties of these highly homologous enzymes are different, many of them are successfully commercialized. The distinguishing feature of archaeal family B DNA polymerases is the ability to recognize and bind uracil base in template DNA during the amplification reaction (Lasken et al., 1996; Greagg et al., 1999; Grätz et al., 2003; Shuttleworth et al., 2004; Connolly, 2009). In addition to uracil, the archaeal family B DNA polymerases can interact with another deaminated base, hypoxanthine (Gill et al., 2007). Highly specific uracil binding is realized using a unique uracil-binding pocket localized in the N-terminal domain of archaeal DNA polymerases (Fogg et al., 2002). The N-terminal domain interacts with the incoming DNA template strand and provides a ‘read-ahead’ function when the DNA polymerase checks the parental DNA chain for the presence of uracil. Upon uracil binding, the polymerase stalls on the DNA template and polymerization reaction stops. Cytosine deamination to uracil caused by high temperatures results in the G–C base pair transition to A–T during DNA amplification (Lindahl, 1993). It is proposed that the uracil-binding pocket and ‘read-ahead’ function of archaeal family B DNA polymerases are a natural safeguard barrier against temperature-induced mutational drift. Uracil binding is an advanced feature in nature; however, it is undesirable in some PCR applications (Supphagh et al., 1993). Due to cytosine-to-uracil deamination during PCR cycling, trace amounts of uracil may decrease DNA amplification yields and lower the sensitivity in simple PCR, high-fidelity PCR and especially in long-range PCR, where long elongation times are required. Replacement of up to 100% of dTTP with dUTP is used in diagnostic PCR applications in order to prevent carryover contamination with the newly synthesized genetic material. This is achieved by treating PCR samples with uracil-DNA glycosylase, which hydrolyzes the N-glycosyl bond between uracil and sugar leaving an apyrimidinic site in the uracil-containing single- or double-stranded DNA, which later is degraded by heating. As a consequence, the wild-type archaeal family B DNA polymerases possessing ‘read-ahead’ function and the mutant variants of those enzymes with slightly decreased uracil binding are not suitable for diagnostic applications.

Even though uracil recognition by replicative DNA polymerases is limited to archaea (Wardle et al., 2008), it is not an absolute feature of all archaeal family B DNA polymerases. The Neq DNA polymerase from Nanoarchaeum equitans successfully utilizes deaminated bases, such as uracil and hypoxanthine (Choi et al., 2008). The ‘read-ahead’ function of other commercially important archaeal polymerases, like Pfu or Tgo, can be easily switched off by site-specific mutagenesis of uracil-binding pocket (Fogg et al., 2002; Firbank et al., 2008). Point mutants, such as V93Q in Pfu polymerase, have significantly (more than 10-fold) reduced affinity for the DNA containing uracil (Fogg et al., 2002). The V93Q mutant of Pfu polymerase can efficiently perform PCR with 1% of dUTP, but shows considerably reduced efficiency of DNA amplification when 100% of dTTP is replaced with dUTP (Gill et al., 2007).
Thermococcus litoralis Sh1B DNA polymerase was isolated from the T. litoralis Sh1B strain found in deep-water hydrotherms of Shikotan Island (Kuril Islands) (Slobodkina et al., 2005). T. litoralis Sh1B DNA polymerase (GenBank accession number GQ891548) has 74% identical amino acids with the Pfua DNA polymerase (Fig. 1) and potentially can be used in many different PCR applications. Like other family B archaeal DNA polymerases, Sh1B DNA polymerase has a uracil-binding pocket and cannot perform PCR in the presence of dUTP (Gaidamaviciute et al., 2010). A version of Sh1B polymerase with the abolished uracil binding is highly desirable and could suite a broader range of PCR applications (including the long-range PCR, qPCR, RT-qPCR and end-point PCR in diagnostics). The uracil-binding pocket in Sh1B DNA polymerase is basically identical to the one in Pfua polymerase and can be identified using a simple amino acid sequence alignment (Fig. 1). An attempt to disable uracil-binding property of the DNA polymerase from T. litoralis by generating site-specific mutants analogous to Pfua polymerase resulted in polymerase variants in which decreased sensitivity to uracil, which tolerated 0.5–20 μM concentrations of dUTP in PCR. Consequently, a random mutagenesis of the whole polymerase gene and subsequent compartmentalized self-replication (CSR) were used for a successful selection of Sh1B DNA polymerase variants capable of performing PCR with dTTP completely replaced by dUTP.

Materials and methods

**Materials**

All reagents were from Fermentas, unless otherwise specified. The DNA restriction and nucleic acid modification enzymes (XbaI, Esp31, Eco31I, Cfr101 restriction endonucleases, Taq DNA polymerase, T4 DNA ligase, ExoI, SAP, uracil-DNA glycosylase (UDG), RNaseA) were supplied by Fermentas. The oligonucleotides were from Metabion. The pASK-IBA5 plasmid was from IBA and pUC19 and pUC57 plasmids from Fermentas. The oligonucleotides were from Metabion. The enzyme fused to the N-terminal Strep-tag was expressed in Escherichia coli JM109 strain. After the removal of plasmid, the strain were supplied by Fermentas.

**DNA manipulation and protein expression**

The T. litoralis Sh1B DNA polymerase coding gene was amplified from T. litoralis genome. After the removal of inton sequences, the remaining of the gene was cloned into pUC57 plasmid. The wild-type T. litoralis Sh1B DNA polymerase gene was amplified from pUC57Topol18(wt) plasmid with primers no. 1 (5'-CGA TGG CTC AGC GCC ATG ATA CTG GAC ACT GA-3') and no. 2 (5'-TAA AAA GGT CTC GTA TCA CCT TTT GAG CCA TGC-3') containing Eco31I restriction endonuclease recognition site, digested with Eco31I restriction endonuclease and cloned into the Eco31I-digested pASK-IBA5 protein expression vector. The pASK-IBA5-pol18(wt) plasmid (see online Supplementary data) coding for the wild-type Sh1B polymerase fused to the N-terminal Streptag was expressed in E. coli JM109 cells. Briefly, the E. coli cells carrying the expression plasmid were grown at 37°C until the OD600 of 1 was reached. The protein expression was induced by adding anhydrotetracycline (Acros Organics) to a final concentration of 0.2 μg/ml, and cells were grown for another 3 h.

**Library construction**

The initial library of genes coding for mutant Sh1B DNA polymerases was generated by error-prone PCR using a modified protocol described by Zaccolo et al. (1996). One hundred nanograms of pUC57Topol18(wt) plasmid was used to amplify the T. litoralis Sh1B polymerase gene with 0.5 μM of each of primers no. 1 and no. 2 (see above) in 100 μl of PCR mixture containing 1 × Taq buffer (75 mM Tris–HCl (pH 8.8 at 25°C), 20 mM (NH4)2SO4 and 0.01% (v/v) Tween 20), 0.2 mM of each standard dNTP, 0.1 μM dPTP, 5 μM 8-oxo-dGTP (modified nucleoside triphosphates were purchased from TriLink BioTechnologies), 1.5 mM MgCl2 and 2.5 U of Taq DNA polymerase. The temperature program used was 3 min at 94°C followed by 30 cycles of 30 s at 94°C, 30 s at 60°C, 2 min 30 s at 72°C and finished with 10 min at 72°C. The PCR product was digested with Eco31I restriction endonuclease and cloned into the Eco31I-digested pASK-IBA5 using T4 DNA ligase. Escherichia coli JM109 cells were transformed with the ligation mixture using electroporation.

**CSR selection**

The CSR strategy described by Ghadessy et al. (2001) was used to select dUTP resistant polymerases from the initial mutant library. Reaction emulsification was performed in 5 ml cryogenic vials (Corning, #430492) as described in Williams et al. (2006), except for few modifications. About 0.3 ml of CSR mix containing 1 × Pfua buffer (20 mM Tris–HCl (pH 8.8 at 25°C), 10 mM (NH4)2SO4, 10 mM KCl, 0.1% (v/v) Triton X-100, 0.1 mg/ml BSA), 0.5 μM of primers no. 3 (5'-CAG GAA ACA GCT ATG ACC AGC

![Fig. 1.](https://example.com/f1.png) The amino acid sequence alignment of the N-terminal domains of Thermococcus litoralis Sh1B (Sh1B, GQ891548) and Pyrococcus furiosus (Pfu, P80061) DNA polymerases. The amino acids depicted in grey are 100% identical; amino acids marked with asterisk comprise the uracil-binding pocket of Pfua polymerase (Firbank et al., 2008).
CAC CGG CAG TTC GAA AA-3') and no. 4 (5'-AGT AAG CGT GTA CCG ATC GTC AAG CTT AGT TAG ATA-3'), 0.2 mM dNTPs, 1.5 mM MgCl₂, 75 ng RNase A, 0.5 mg/ml BSA, dUTP (in the first CSR selection cycle of 0.2 μM, in subsequent four cycles of 0.5 μM, 2 μM, 4 μM and 8 μM, respectively) and 1 x 10⁷ of induced E. coli cells overexpressing mutant polymerases were added to 0.7 ml of oil phase containing 2% (vol/vol) ABIL EM 90 (Goldschmidt GmbH), 0.05% (vol/vol) Triton X-100 (Sigma) in mineral oil (Sigma, #M5904) under constant stirring (1714 rpm) at +4°C. After addition of the aqueous phase (gradually in a period of 2 min), stirring continued for 5 min more before thermocycling (94°C for 5 min, 20 times at 94°C for 1 min, 52°C for 1 min and 72°C for 5 min) was started. Then the aqueous phase was extracted using diethyl ether (Merck) and ethyl acetate (Sigma-Aldrich) as described in Williams et al. (2006), unincorporated primers were degraded with ExoI and SAP (70 μl of extracted aqueous phase was incubated with 10 U of SAP and 50 U of ExoI at 37°C for 15 min). After chlorophorm/phenol (Roth) extraction and ethanol precipitation, selection products were amplified with primers no. 5 (5'- CAG GAA ACA GCT ATG ACC -3') and no. 6 (5'-AGT AAG CGT GTA CCG ATC -3'). The PCR product was reamplified with primers no. 1 and no. 2, which contained an EcoIII recognition site, and recloned as above.

**Polymerase purification**

The E. coli cells overexpressing polymerase were harvested from 200 ml of LB culture by centrifugation at 4°C (5000 rpm for 10 min, Beckman J2-21M/E centrifuge, JA-10 rotor) and resuspended in buffer A (50 mM Tris–H₂PO₄ (pH 8.0), 2 mM EDTA, 150 mM NaCl, 0.1% Triton X-100, 5 mM 2-mercaptoethanol) with 1 mM phenylmethanesulphonyl-fluoride (Sigma). After sonication on ice (7.5 min), samples were centrifuged at 16 170 g for 20 min (Eppendorf 5417R). Next, the supernatant was heated at 75°C for 15 min to denature the most of E. coli mesophilic proteins. Precipitated proteins were removed by centrifugation (at 16 170 g for 20 min) and the supernatant was loaded onto a Strep-Tactin Superflow (IBA) minicolumn. To remove unspecifically bound proteins, the minicolumn was washed with buffer A and buffer B (buffer A without Triton X-100); the polymerase was eluted with buffer B containing 2.5 mM desthiobiotin (Sigma). Eluted polymerase was dialyzed against the storage buffer (20 mM Tris–HCl (pH 8.2), 0.1 mM EDTA, 1 mM DTT, 100 mM KCl, 0.1% Nonidet P-40, 0.1% Tween 20, 50% glycerol) and stored at -20°C.

**Activity assay for selected polymerases**

The DNA polymerase activity of selected purified enzymes was measured according to the following protocol. The enzyme was incubated in a reaction mixture (50 μl) consisting of 20 mM Tris–HCl (pH 8.8), 10 mM (NH₄)₂SO₄, 10 mM KCl, 0.1% (v/v) Triton X-100, 0.1 mg/ml BSA, 2 mM MgSO₄, 200 μM each of dATP, dCTP, dTTP and dGTP, 0.4 MBq/ml of [methyl-³²]thymidine 5'-triphosphate (Amersham), and 250 μg/ml of activated calf-thymus DNA at 72°C for 30 min. The reaction was stopped on ice, and an aliquot was spotted onto a DE-81 filter-paper disc. The disc was dried on a heat block, washed in 7.5% sodium phosphate buffer for 5 min three times and once in 70% ethanol for 2 min, then dried again. The incorporated radioactivity on the dried filter-paper disc was counted using a Beckman LS-1801 scintillation counter. One unit of Sh1B DNA polymerase catalyzes the incorporation of 10 nmol of deoxyribonucleotides into a polynucleotide fraction (adsorbed on DE-81) in 30 min at 72°C.

**Screening and characterization**

A PCR assay was used to screen and rank selected mutant polymerases. About 0.5 U of purified polymerase were added to 20 μl of PCR mix and the efficiency of amplification of ~1 kb fragment from a plasmid DNA was compared under standard ([dNTP] = 0.2 mM) and selection conditions ([dNTP] = 0.2 mM plus increasing amounts of dUTP). In PCR mixtures 1 x Pfu buffer (the same as above), 1.5 mM MgCl₂, 0.5 μM of primers, 0.2 mM of each of standard dNTP were used. In the case of PCR with dUTP replacing dTTP, [dUTP] + [dTTP] = 0.2 mM. For the evaluation of dUTP incorporation by selected polymerases during PCR amplification, the PCR products obtained under standard PCR conditions (without dUTP) and under the selection conditions (with dUTP addition) were subjected to the uracil-DNA glycosylase (UDG) treatment (30 μl of reaction mixture containing ~150–300 ng of unpurified PCR product and 2 U of UDG were incubated at 37°C for 10 min) and after heating at 95°C for 15 min were analyzed in ethidium bromide-stained 1% agarose gel.

**Site-specific mutagenesis of T. litoralis Sh1B DNA polymerase**

Single site-specific mutations of T. litoralis Sh1B DNA polymerase mutants Y7A, V93Q, V93G, P36L and P36H were introduced into wild-type gene by the two-step PCR ‘maga-primer’ method (Barik, 1995). The mutant genes were cloned into pASK-IBA5 vector and expressed as above.

**Construction of the library of selected polymerases lacking mutations in the N-terminal domain**

The pool of pASK-IBA5 plasmids containing mutant polymerase genes obtained after final CSR selection round was subjected to XbaI and Esp3I digestion. The resulting 5042 bp DNA fragment was ligated with the 445 bp DNA fragment obtained after pASK-IBA5-pol18(wt) plasmid digestion with the same restriction enzymes.

**PCR fidelity assay**

The fidelity of Sh1B DNA polymerase and its mutant variants was determined using a modified lacZα-based PCR fidelity assay (Lundberg et al., 1991). This method is based on the amplification, circularization and transformation of the pUC19 plasmid, which contains a lacZα allele encoding the inactive N-terminal fragment of beta-galactosidase. This fragment, whose synthesis can be induced by IPTG, is capable of intra-allelic complementation with a defective form of beta-galactosidase encoded by the complementing E. coli strain. In the presence of IPTG, bacteria synthesize both fragments of the enzyme and form blue colonies on media with X-gal. The PCR-derived mutations in lacZα result in the synthesis of defective N-terminal fragment of beta-galactosidase and subsequent formation of an inactive beta-galactosidase enzyme, which results in the formation of white colonies on X-Gal indicator plates.
The pUC19 plasmid was linearized with Cfr10I restriction endonuclease, submitted to phenol/chloroform extraction and after purification with GeneJET™ PCR Purification Kit (Fermentas) was used as a template in PCR with primers Fid05 (5'-AAAAACCGCCTCCAGATTATCAGCAATAAACCAAG-3') and Fid06 (5'-AAAAAGCCTAGTGAGCGTGGGTCC-3'), which contain Cfr10I recognition site and anneal on the opposite ends of linearized plasmid. The PCR (100 μl) was performed in 1× Pfu buffer (see above) with 1.5 ng of template DNA, 0.5 μM of each primer, 0.2 mM dNTPs, 1.5 mM MgCl₂ and 5 U of enzyme. The cycle conditions were 45 s denaturation at 94°C, 45 s annealing at 63°C and 5.3 min elongation at 72°C repeated for 25 cycles. After PCR, the yield of amplification product was determined (in agarose gel) and amplified DNA was purified with GeneJET™ PCR Purification Kit. After digestion with Cfr10I and phenol/chloroform extraction, the DNA fragments were purified from a preparative agarose gel with GeneJET™ Gel Extraction Kit (Fermentas). The ligation reactions (50 ng of Cfr10I digested PCR fragment) were carried out with the Rapid DNA Ligation Kit (Fermentas). Competent E. coli JM109 cells (prepared by CaCl₂ method) were transformed with the resulting circular plasmids and plated on LB plates with 0.1 mg/ml ampicillin, 0.01 mg/ml X-Gal and 0.1 mM IPTG. After incubation overnight at 37°C, blue and white colonies were counted. The error rate f, was calculated with equation: 
\[ f = F(d \times b) \]
where F is the fraction of white colonies (white colonies/total colonies), d is the number of DNA duplications (2^n = output DNA/input DNA) and b is the number of nucleotides within the 159 nucleotide-long target region supposed to yield a mutant phenotype with a specific mutational change. It is known that for 250-nucleotide lacZα region in bacteriophage M13mp2 b = 110 or about 44% (Tindall and Kunkel, 1988). The approximate extrapolation to 159-nucleotide region results in b = 70.

**Thermostability assay of polymerases**

About 50 μl of mixtures containing 1.25 U of polymerase in 1× Pfu buffer (see above) with 2 mM MgSO₄ and 0.2 mM dNTPs were incubated at 95°C for varied periods of time. At each time point, heated sample was transferred to an ice bath. Then, 5 μl of sample were mixed with 45 μl of the polymerase activity assay reaction mixture, and polymerase activity was measured as described above. The remaining activity was plotted versus time spent at 95°C. The data were fitted to the following equation: 
\[ \ln(A) = \ln(A₀) - kt \]
where A represents the remaining activity at time t, A₀ is the activity at t = 0 and k is the rate constant for enzyme inactivation. The half-life (t₁/₂) was determined from the following equation: 
\[ t₁/₂ = (\ln 2)/k \]

**Results**

**Polymerases generated by site-specific mutagenesis**

By performing site-directed mutagenesis of uracil-binding pocket of the thermostable archael DNA polymerase from *Pyrococcus furiosus* (*Pfu*), Connolly et al. identified key amino acid alterations that partially (Y7A) or completely (V93Q) abolished uracil-induced stalling of DNA replication (Fogg et al., 2002). Moreover, it was demonstrated that these mutations diminished the polymerase inhibition by dUTP in PCR, as both *Pfu*-Y7A and *Pfu*-V93Q mutant DNA polymerases were shown to be able to carry out PCR with dUTP completely replacing dTTP in the reaction mixture. It is, nevertheless, obvious that the PCR efficiency using mutant *Pfu* DNA polymerases and 100% of the dUTP is substantially decreased (Fogg et al., 2002; Gill et al., 2007). As the primary sequence of the archael Sh1B DNA polymerase is highly homologous to *Pfu* DNA polymerase (74% identity) and the amino acids comprising uracil-binding pocket are almost identical in these polymerases (Fig. 1), we expected that the introduction of Y7A or V93Q mutation would also decrease the uracil sensitivity of Sh1B DNA polymerase, yielding a polymerase able to perform PCR with dUTP instead of dTTP. However, the Y7A mutant of Sh1B DNA polymerase was inhibited in PCR by similar dUTP concentrations that are deleterious to the wild-type enzyme. The V93Q mutant of Sh1B DNA polymerase was able to perform PCR in the presence of increased amounts of dUTP (Fig. 2), but still could not amplify DNA when all dTTP was replaced by the dUTP (data not shown). These results show that neither V93Q nor Y7A mutation can completely disable the uracil binding by Sh1B DNA polymerase. Therefore, we assumed that probably other amino acid substitutions, different from those identified in *Pfu* polymerase (Fogg et al., 2002), had to be found.

**Polymerase selection by CSR**

The mutational scanning and site-saturation mutagenesis of the amino acids comprising uracil-binding pocket could be a successful way of trying to select the best variants of Sh1B DNA polymerase. However, the traditional rational design approach has a number of drawbacks and is quite laborious. Therefore, to find the crucial amino acid mutations, which could significantly reduce uracil inhibition of Sh1B DNA polymerase, we have employed a methodology of directed enzyme evolution known as the CSR (Ghadessy et al., 2001; Ghadessy et al., 2004; d’Abbadie et al., 2007; Loakes et al., 2009). A brief description of this simple selection scheme is as follows. A mutant library of the polymerase of interest...
(in our case Sh1B DNA polymerase) has to be expressed in *E. coli*. The bacterial cells, carrying plasmids encoding mutated variants of the enzyme—a thermostable polymerase by itself—are added to the PCR mixture (buffer, primers and dNTPs). The compartments of water in oil emulsion having no more than one cell per compartment are prepared and subjected to PCR cycling conditions. During the first denaturation step (94°C), the wall of the bacterial cell is disrupted and the molecules of thermostable DNA polymerase are released into the PCR mixture. Only the variants of PCR proficient DNA polymerase will be able to amplify its own coding gene during the PCR in emulsion and, therefore, will be enriched during each CSR selection step. In our case, as a selection pressure we have used the increasing concentration of dUTP in the PCR mixture. According to our preliminary studies, the wild-type Sh1B DNA polymerase, likewise the other family B thermostable archaeal DNA polymerases, misuses dUTP for dTTP in the process of DNA amplification and is noticeably inhibited in PCR with as little as 0.2 μM of dUTP, due to its inability to read through the uracil-containing DNA matrices.

The diversity of mutant library used in CSR selection is usually limited to $10^3–10^7$ compartments and is restricted by the volume of water phase used in emulsification and by the diameter of water compartments. As the diversity numbers of experimental setup are high enough, we have not restricted the mutant library of Sh1B DNA polymerase to the variations only in the N-terminal part of the protein, where uracil-binding pocket is localized. Instead, we have performed a random mutagenesis of the whole gene by error-prone PCR and have generated a population of random Sh1B mutants (L0 pool) with a diversity of $5 \times 10^5$ polymerase expressing clones, which on the average had one to four amino acid mutations per gene. In order to isolate the mutant Sh1B variants proficient in PCR and having a reduced or completely disabled uracil binding, we have subjected the initial mutant library (L0) to five rounds of CSR selection. The L1 pool of Sh1B polymerase mutant pools encoded by the initial library before selection (L0 pool) and after five selection rounds (L5 pool) were purified via the N-terminal Strep-tag by affinity chromatography on the streptactin column. As it was expected, the L0 pool of Sh1B polymerase mutants was able to perform PCR in a reaction mixture without dUTP and was inhibited by 0.5 μM concentration of dUTP used in addition to 200 μM of dTTP (Fig. 3A). Contrary to that, the L5 pool of Sh1B polymerase mutants was able to synthesize DNA fragment in the presence of 0.5–10 μM of supplemental dUTP. In general, it is possible that the increase in the tolerance to higher concentrations of dUTP in PCR has arisen during the selection process because of the increased $K_m$ for dUTP in comparison to dTTP. As a consequence, such Sh1B DNA polymerase variants would specifically use dTTP, which is present at 200 μM concentration, instead of dUTP, concentration of which was increased to 8 μM in the last selection round.

In order to prove that the L5 pool of Sh1B polymerase mutants is actually incorporating uracil, we have demonstrated that a PCR product synthesized with 200 μM of dTTP plus 10 μM of dUTP is completely degraded after the incubation with uracil-DNA glycosylase (UDG) and heat treatment (Fig. 3B).

**General analysis of evolved polymerases**

The L5 pool is a mixture of mutant Sh1B polymerases with the averaged enzymatic properties of all proteins comprising it. It is obvious that individual polymerases after five CSR selection rounds have acquired different levels of dUTP resistance (some polymerases should be less resistant to dUTP compared with L5 pool, some more). An individual clone analysis was performed on 17 randomly picked clones. All 17 mutant Sh1B polymerases encoding genes were sequenced and analyzed. A library of random mutants of the

![Fig. 3.](image-url) The properties of the L5 pool of selected Sh1B mutants. (A) The activity in PCR in the presence of dUTP. 1 kb DNA fragment was amplified with the initial (L0) and selected (L5) pool of mutant Sh1B variants in the presence of four standard dNTPs (each 200 μM) and the increasing amounts of dUTP. (B) The ability to read through the uracil-containing DNA. One-kilobase DNA fragment amplified with L5 pool in the presence of four standard dNTPs (each 200 μM) and the presence (+) or absence (−) of 10 μM of dUTP was incubated with (+) or without (−) uracil-DNA glycosylase (UDG) and heated. As L5 pool incorporates uracil-containing nucleotide into DNA and amplifies uracil-containing DNA, such PCR product is degraded after the treatment with UDG and heating. The PCR product synthesized without dUTP remains intact after the UDG treatment and heating step.
whole polymerase gene was used in CSR selection and as a consequence mutations all over the gene sequence could be found in selected clones. A more detailed analysis of mutation distribution revealed a much higher frequency of amino acid changes localized in the N-terminal domain of the protein where uracil-binding pocket is. In order to elucidate if the mutations, which are present outside the N-terminal region of the protein might be contributing to the uracil-resistance properties of selected enzymes or whether they are insignificant, we have replaced the N-terminal sequence (1–130 amino acids) in L5 pool of mutant polymerases with the corresponding N-terminal domain of the wild-type Sh1B polymerase (Fig. 4A). A new L5* pool of proteins containing mutant polymerases with the wild-type N-terminal domain was purified, and its ability to conduct PCR in the presence of increasing concentrations of dUTP was tested. The restoration of the wild-type N-terminal domain in evolved polymerases completely abolished the dUTP resistance feature of evolved mutants (Fig. 4B). The L5* pool of modified polymerases was inhibited in PCR by similar concentrations (0.5 μM) of dUTP that inhibited the wild-type Sh1B polymerase. These results clearly indicate that mutations located in the N-terminal domain of evolved Sh1B mutants are the key determinants of increased dUTP resistance. The types of amino acid alterations present in the N-terminal domain of evolved polymerases and their frequency of occurrence (the sequence data are from 17 randomly picked polymerases of L5 pool) are summarized in Table I. First of all, it is interesting to notice that in all 17 evolved polymerases that were chosen for sequence analysis at least one amino acid of uracil-binding pocket was found to be mutated (P36L, P36H, Y37H, D92G, V93G, R97M, R97W, F116L and R119H). The V93G, P36L and P36H amino acid alterations of uracil-binding pocket were found to be the most frequently occurring mutations in the N-terminal domain of evolved polymerases.

The analysis of individual polymerases

Fifteen of 17 sequenced polymerases were purified individually and analyzed for the ability to perform PCR in the presence of increasing concentrations of dUTP. The best performing individual mutants (M4, M12, M8, M13, M16, M15) can tolerate 20 μM and higher concentrations of dUTP in addition to 200 μM of dTTP already present in the PCR mixture (Fig. 5A). Two selected polymerases M4 and M12, possessing the highest resistance to dUTP, were able to amplify DNA when dTTP was completely replaced by dUTP in the reaction mixture (Fig. 5B). Comparing the amino acid mutations of the six best performing polymerases, we noticed that all those mutants had an amino acid substitution in either 36th or 93rd position inside the uracil-binding pocket (Fig. 5A). To find if mutations in these positions determine an increased uracil resistance of selected polymerases, we constructed single P36H, P36L and V93G mutants of Sh1B DNA polymerase and evaluated their ability to conduct PCR in the presence of increasing amounts of dUTP. As expected, all three site-directed mutants of Sh1B DNA polymerase were able to amplify DNA in the presence of higher concentrations of dUTP compared with the wild-type enzyme (Fig. 6A). The Sh1B-P36L polymerase was able to synthesize 1-kb DNA fragment at 10–20 μM concentration of dUTP and was more resistant to dUTP than the Sh1B-V93G polymerase (5 μM of dUTP), but its uracil resistance was still comparable to that of Sh1B-V93Q polymerase (Fig. 2). The most striking uracil resistance was observed with the P36H mutant of Sh1B polymerase. P36H mutation localized in the uracil-binding pocket was found in both the best-performing mutants M4 and M12,

Table I. Mutations present in the N-terminal domains of evolved polymerases

<table>
<thead>
<tr>
<th>Mutation</th>
<th>Number of repeats</th>
</tr>
</thead>
<tbody>
<tr>
<td>V93Gb</td>
<td>5</td>
</tr>
<tr>
<td>P36Lb</td>
<td>5</td>
</tr>
<tr>
<td>P36Hb</td>
<td>3</td>
</tr>
<tr>
<td>V66I</td>
<td>2</td>
</tr>
<tr>
<td>K74Q</td>
<td>2</td>
</tr>
<tr>
<td>V82A</td>
<td>2</td>
</tr>
<tr>
<td>E130G</td>
<td>2</td>
</tr>
<tr>
<td>Q35R</td>
<td>1</td>
</tr>
<tr>
<td>Y37H</td>
<td>1</td>
</tr>
<tr>
<td>E57G</td>
<td>1</td>
</tr>
<tr>
<td>D92Gb</td>
<td>1</td>
</tr>
<tr>
<td>R97Wb</td>
<td>1</td>
</tr>
<tr>
<td>R97Mb</td>
<td>1</td>
</tr>
<tr>
<td>R97Wb</td>
<td>1</td>
</tr>
<tr>
<td>E102G</td>
<td>1</td>
</tr>
<tr>
<td>A105T</td>
<td>1</td>
</tr>
<tr>
<td>I107V</td>
<td>1</td>
</tr>
<tr>
<td>I109R</td>
<td>1</td>
</tr>
<tr>
<td>F116Lb</td>
<td>1</td>
</tr>
<tr>
<td>R119Hb</td>
<td>1</td>
</tr>
</tbody>
</table>

*bA number of evolved polymerases containing the indicated mutation. The data were derived from the analysis of amino acid sequences of 17 mutant polymerases randomly picked from the L5 pool (no identical sequences were found). *bMutations in the uracil-binding pocket.

![Fig. 4.](https://academic.oup.com/peds/article-abstract/23/8/589/1510260/fig4?download=true)
which were able to perform PCR in the reaction mixtures where dTTP was completely replaced by dUTP (Fig. 5B). The Sh1B-P36H mutant, in contrast to Sh1B-V93G and Sh1B-P36L, successfully synthesized PCR fragment at dUTP concentrations as high as 100 mM (Fig. 6A). Even more, the Sh1B-P36H mutant was able to perform PCR using only dUTP as a replacement of dTTP in dNTPs mixture (Fig. 6B) and, thus, far away outperformed rationally designed mutants Sh1B-Y7A and Sh1B-V93Q (Fig. 2).

Other features, like the thermostability and fidelity, are also very important for polymerases used in PCR. The thermostability (half-life \( t_{1/2} \approx 55 \text{ min at } 95^\circ \text{C} \)) and the fidelity \( (4.3 \times 10^4, \text{ an average number of correct nucleotides incorporated before an error occurs}) \) of the Sh1B-P36H mutant were determined and found to be very similar to the thermostability (half-life \( t_{1/2} \approx 59 \text{ min at } 95^\circ \text{C} \)) and fidelity \( (3.98 \times 10^4 \)) of the wild-type Sh1B DNA polymerase (see online Supplementary data).
Discussion

The goal of this work was to abolish the ‘read-ahead’ uracil-binding function of Sh1B DNA polymerase and so to produce an enzyme capable of performing efficient PCR in both the presence and absence of dUTP. The simplest approach—a rational design employing mutations localized in the uracil-binding pocket, analogous to Pfu mutations, did not give satisfactory results. The Sh1B polymerase mutants Y7A and V93Q had phenotypes similar to Pfu polymerase mutants. Both enzymes had a decreased sensitivity to uracil and were able to use higher concentrations of dUTP in PCR compared with the wild-type enzyme (Y7A - 0.5 μM of dUTP, V93Q - 20 μM, Fig. 2), but were unable to perform PCR with dTTP completely replaced by dUTP. Most likely, it is not enough to diminish the uracil binding by introducing a steric hindrance in the uracil-binding pocket of the polymerase. It is much more important to find the appropriate amino acid substitution, which will block uracil binding, but will not affect other features of the polymerase important for PCR, like the thermostability, processivity, affinity for the substrate, etc. The solution for this complex task is quite problematic. An intelligent PCR enzyme evolution setup of CSR, which exploits the actual PCR cycling conditions and can deal with the library of a large number of mutants (10^5–10^7), is a proper tool to find a perfect DNA polymerase proficient in PCR. Sh1B polymerase mutant P36H selected after five CSR selection rounds with the increasing concentrations of dUTP can efficiently perform PCR in the reaction mixture, where dTTP is completely replaced by dUTP. In cells the ‘read-ahead’ function is part of a complex fidelity mechanism, which prevents the replication of deaminated cytosines. However, the fidelity of the P36H mutant of Sh1B DNA polymerase (4.3 × 10^4) in PCR is basically the same as of the wild-type Sh1B DNA polymerase (3.98 × 10^4). It is obvious, that in our case the polymerase fidelity in the PCR is determined by the combination of the general nucleotide incorporation accuracy and the 3′→5′ exonuclease proof-reading activity and does not depend on the uracil recognition and ‘read-ahead’ function.

The structure of Sh1B DNA polymerase is unknown; however, the structure of highly homologous Tgo DNA polymerase (79% of identical amino acids) in complex with a DNA primer–template containing uracil in the single-stranded region (Firbank et al., 2008) is available and can be exploited in Sh1B polymerase’s analysis (Fig. 7). In the case of Sh1B polymerase proline 36 mutation to histidine (mutant P36H), a bulky imidazole side chain of histidine should be localized in the uracil-binding pocket, filling the cavity and completely abolishing the ‘read-ahead’ function of polymerase. Proline residues typically are critical residues in maintaining the integrity of thermostable proteins at elevated temperatures (Sakaguchi et al., 2007). The distinctive cyclic structure of proline’s side chain ‘locks’ the protein’s Cα backbone, giving the proline and protein backbone an exceptional conformational rigidity. There is a high possibility that proline replacement by any other amino acid would decrease the polymerase’s rigidity and the thermostability at the same time. However, in the Sh1B DNA polymerase P36H mutant case, the imidazole (pentagonal heterocycle) side chain of histidine fitted perfectly in the uracil (hexagonal heterocycle) binding pocket, was immobilized in the cavity and as a consequence rescued the protein’s structure. Changes in the same position to other amino acids, such as P36L in Sh1B polymerase also selected in CSR (Table I, Fig. 6A) or P36A in Pfu polymerase (Firbank et al., 2008) constructed by the rational design could not match the P36H mutant in PCR, especially when dTTP is completely replaced by dUTP. Therefore, the proline exchange to histidine (P36H) is a proper mutation, which completely switches off the uracil binding and at the same time maintains the polymerase’s rigidity, structural integrity and thermostability (respectively half-life t1/2 ~55 min for P36H and t1/2 ~59 min for the WT polymerase at 95°C). The P36H mutant of Sh1B DNA polymerase selected by CSR outperforms rationally designed mutant V93Q in PCR where dUTP instead of dTTP is used and now can be employed in diagnostic PCR applications. Taking into account the close similarity of the uracil-binding pockets in Pfu and Sh1B and high correlation in the site-specific mutagenesis data (mutants Y7A and V93Q (Fogg et al., 2002); Fig. 2), it is very likely, that the behaviour of Sh1B mutants can be directly extrapolated to Pfu DNA polymerase and vice versa. Thus, there is a high probability, that

![Fig. 7. (A) Uracil-binding pocket of Thermococcus gorgonarius (Tgo) polymerase (2VWJ) in complex with a DNA primer–template containing uracil in the single-stranded region (Firbank et al., 2008). Uracil residue is displayed as spheres and adjacent amino acid residues (Y7, P36, Y37, P90, V93, I114, P115) as sticks. (B) Uracil-binding pocket of Tgo polymerase with modelled mutation P36H, which clashes with uracil. Mutation was modelled using DeepView Swiss-Pdb Viewer v. 4.0 and Swiss-model server.](https://academic.oup.com/peds/article-abstract/23/8/589/1510260/fig7)
the P36H mutant of *Pfu* DNA polymerase would behave similar to P36H-Sh1B enzyme and would also be able to outperform the previously designed *Pfu* mutant V93Q in PCR where dTTP is completely replaced by dUTP.

**Supplementary data**

Supplementary data are available at *PEDS* online.

**Acknowledgements**

We thank Edita Gaidamaviciute for the construction of Sh1B DNA polymerase mutants Y7A, V93Q, Gediminas Alzbutas for the help with structural modelling, professor Arvydas Janulaitis and Arunas Lagunavicius for fruitful discussions and Ramune Leipuviene for the help with manuscript preparation.

**References**


