Appl Biochem Biotechnol (2011) 165:1188–1200
DOI 10.1007/s12010-011-9337-5

Cloning and Sequence Analysis of Novel DNA Polymerases from Thermophilic *Geobacillus* Species Isolated from Hot Springs in Turkey: Characterization of a DNA Polymerase I from *Geobacillus kaue* Strain NB

Melike Çağlayan · Neş'e Bilgin

Received: 17 March 2011 / Accepted: 2 August 2011 /
Published online: 26 August 2011

Abstract The complete coding sequences of the *polA* genes from seven thermophilic *Geobacillus* species, isolated from hot springs of Gönen and Hisaralan in Turkey, were cloned and sequenced. The *polA* genes of these *Geobacillus* species contain a long open reading frame of 2,637 bp encoding DNA polymerase I with a calculated molecular mass of 99 kDa. Amino acid sequences of these *Geobacillus* DNA polymerases are closely related. The multiple sequence alignments show all include the conserved amino acids in the polymerase and 5′-3′ exonuclease domains, but the catalytic residues varied in 3′-5′ exonuclease domain of these *Geobacillus* DNA polymerases. One of them, DNA polymerase I from *Geobacillus kaue* strain NB (*Gkaue* polI) is purified to homogeneity and biochemically characterized in vitro. The optimum temperature for enzymatic activity of *Gkaue* polI is 70 °C at pH 7.5–8.5 in the presence of 8 mM Mg^{2+} and 80–100 mM of monovalent ions. The addition of polyamines stimulates the polymerization activity of the enzyme. Three-dimensional structure of *Gkaue* polI predicted using homology modeling confirmed the conservation of all the functionally important regions in the polymerase active site.

Keywords *Geobacillus* sp. · DNA polymerase I · *Geobacillus kaue* strain NB polI · DNA polymerization in vitro · Homology modeling

Introduction

DNA polymerases are the essential enzymes in all organisms mainly responsible for the replication and repair of DNA [1]. DNA polymerases are classified into five families: A, B, C, X, and a newly identified Y family [2]. Family A DNA polymerases, including human DNA polA and *Escherichia coli* DNA polymerase I, share fundamental structural features in their polymerase active sites [3].

DNA polymerase I (DNA deoxynucleotidyltransferase, EC 2.7.7.7) is encoded by the *polA* gene, which consists of a single polypeptide chain comprising three distinct domains.

M. Çağlayan · N. Bilgin (✉)
Department of Molecular Biology and Genetics, Bogazici University, Bebek, 34342 Istanbul, Turkey
e-mail: bilginne@boun.edu.tr

These three domains are the C-terminal domain with DNA polymerase activity, the central domain with 3′-5′ exonuclease (or proofreading) activity and the N-terminal domain with 5′-3′ exonuclease activity [4]. The C-terminal domain includes the nucleotide and amino acid conservation in the three highly conserved regions (known as motifs A, B, and C) residing in three distinct subdomains designated as thumb, palm, and fingers [5]. The palm subdomain has the catalytic center and contains the conserved carboxylate residues, while the thumb and fingers subdomains are responsible for the binding of DNA and the incoming dNTP, respectively [6].

A few moderately thermostable DNA polymerase I have been purified and characterized from the thermophilic members of the family *Bacillaceae* [7], especially from the genus *Bacillus* and from a recently identified new genus *Geobacillus*: *Bacillus stearothermophilus* [8–10], *Bacillus caldotenax* [11, 12], *Geobacillus* sp. MKK [13], and *Geobacillus caldoxylolyticus* TK4 [14]. The crystal structure of a family A member DNA polymerase I from the genus *Bacillus* has only been identified for *B. stearothermophilus* DNA polymerase large fragment, known as *Bacillus* fragment (BF) [15, 16]. BF is a segment of the polymerase gene that encodes the C-terminal 592 amino acids of the full-length *B. stearothermophilus* DNA polymerase I and has the polymerase activity but lacks the 3′-5′ exonuclease and the 5′-3′ exonuclease activities of the full-length protein [17, 18].

In this study, we report seven new members of family A DNA polymerases from thermophilic *Geobacillus* species, collected from Gönen and Hisaralan hot springs in Turkey. DNA polymerase I sequences from seven thermophilic *Geobacillus* species were determined using PCR-based cloning and subsequent sequencing of the *polA* genes. One of them, DNA polymerase I from *Geobacillus kaue* strain NB (*Gkaue* polI) was expressed in *E. coli* and purified and characterized biochemically in vitro. The three-dimensional structure of *Gkaue* polI were predicted via homology modeling using the crystal structure of BF as a template.

Materials and Methods

Growth Conditions

Geobacillus species were cultivated at 70 °C in a growth medium containing 4 g yeast extract, 8 g tryptone, 3 g NaCl, 0.1 g nitrilotriacetic acid, 0.06 g $CaSO_4 \cdot 2H_2O$, 0.1 g $MgSO_4 \cdot 7H_2O$, 0.1 g KNO_3, 0.69 g $NaNO_3$, 0.1 g Na_2HPO_4, 0.28 mg $FeCl_3$, 2.2 mg $MnSO_4 \cdot H_2O$, 0.5 mg $ZnSO_4 \cdot 7H_2O$, 0.016 mg $CuSO_4$, 0.025 mg $Na_2MoO_4 \cdot 2H_2O$, 0.046 mg $CoCl_2 \cdot 6H_2O$, and 0.005 ml concentrated H_2SO_4 in 1 l at pH 7.5. *G. kaue* strain NB, whose DNA polymerase I was characterized in this study, was isolated from a hot spring in Gönen, and deposited in the Bacillus Stock Center at the University of Ohio (accession no. BGSC105A2).

Cloning and Amino Acid Sequence Analysis of the *polA* Genes

We designed degenerate primers (Table 1) in order to amplify the previously unknown *polA* genes from seven *Geobacillus* species based on the most conserved regions of previously known *polA* genes from the *Bacillus* species [8–12]. The full-length *polA* gene sequences were determined by plasmid DNA sequencing of the genes after their PCR-based cloning into pETM-20 expression vector (Novagen). For this purpose, the oligonucleotide primers F-*Nco*I and R-*Kpn*I (Table 1) were designed. PCR reactions were in 25-μl reaction volume containing 1X *Pfu* buffer, 1.5 mM of $MgCl_2$, 0.2 mM of each dNTP, 1 μM of each primer, 50 ng of genomic DNA, and 0.25 unit of *Pfu* DNA polymerase. The thermal cycling conditions were as follows: 2 min and 30 s at 94 °C for initial denaturation, followed by 30

Table 1 Oligonucleotide primers used for PCR amplification and sequencing of the *polA* genes

Primer	5′-Sequence[a]
F-40	AGATTGAAGAAAAAACTCGT
F-272	GCRTGGTACAATAGRACAAGG A
F-339	GAYGGMARCAGYSTGGCDTA
F-639	GAAGCGGGACGAYATTAT
F-1300	ATG CCC CGATTGTCGGAATC
F-2175	GADCCCBAACYTGCACGARCTSATT
R-126	ACCAGCAAGTTGGGTCGGC
R-780	CGGTGATCCCTTTTTTCGTA
R-2030	TTCGACGATTTCATGGTGGG
R-2605	ACRTABCCTTTYTGTTTY
R-2950	TYTTATTTSGCRTCRCRTACCAY
F-poll	GCRTGGTACAATAGRACAAGGA
R-poll	TYTTATTTSGGRTCRTACCAY
F-NcoI	ATCACCATGGGAATGAGATTGAAGAAAAACTCGT
R-KpnI	ATTAGGTACCTTATTTGGGATCGTACCACGTCGG
F-T7	TAATACGACTCACTATAGGG
R-pRSET	TAGTTATTGCTCAGCGGTGG

[a] R=A/G, Y=T/C, M=A/C, S=G/C, N=A/G/T/C

cycles of 30 s at 95 °C for denaturation, 1 min and 30 s at 55 °C for annealing, and 3 min at 72 °C for extension. The amplified PCR product was inserted between the *Nco*I and *Kpn*I restriction enzyme cleavage sites into the multiple cloning region of pETM-20 vector; 1 μg PCR product and pETM-20 vector plasmid DNA were firstly digested with 1 unit of *Nco*I at 37 °C for 1.5 h and subsequently digested by 1 unit of *Kpn*I at 37 °C for an additional 1.5 h. Double-digested PCR product and pETM-20 vector were ligated overnight at 4 °C by T4 DNA ligase. Ligation mixture was then used for the transformation of *E. coli* TOP10F' cells. Purified plasmids harboring correctly inserted *polA* genes were then confirmed by restriction enzyme digestion and PCR amplification using pETM-20-specific primer pair F-T7 and R-pRSET (Table 1). The full-length *polA* gene sequences were determined by plasmid DNA sequencing using a set of primers (Table 1). The name of the plasmids carrying the cloned DNA polymerase I genes from seven *Geobacillus* species were given in Table 2. The nucleotide sequence analysis and homology search of the *polA* genes were performed using Basic Local Alignment Search Tool. For sequence alignments, multiple sequence alignment tool CLUSTAL W was used [19]. The phylogenetic and evolutionary analyses were conducted using Molecular Evolutionary Genetics Analysis (MEGA) version 4.0 [20] by neighbor-joining strategy [21].

Structural Homology Modeling

The three-dimensional (3D) models for the polymerase domain of *Gkaue* polI were generated using SWISS-MODEL [22] via homology modeling [23]. The crystal structure of DNA polymerase I large fragment from *B. stearothermophilus* (BF, pdb code 1XWL) was used as the template structure [15, 16]. The predicted models were visualized using PyMOL (DeLano Scientific).

Table 2 The name of the plasmids carrying the *polA* genes of *Geobacillus* species identified in this work

Source *Geobacillus* species	Name of the plasmid
Geobacillus anatolicus	pGanapolI
G. anatolicus strain C4	pGanaC4polI
Geobacillus bogazici	pGbogazicipolI
Anoxybacillus sp. NB	pAnoxypolI
Geobacillus kaue strain NB	pGkaueNBpolI
G. kaue strain MC	pGkaueMCpolI
G. kaue strain E1	pGkaueE1polI

Expression and Purification

Gkaue polI was expressed in *E. coli* JM109(DE3) cells. *E. coli* cells harboring pGkaueNBpolI were grown at 37 °C in LB medium containing 150 μg/ml ampicillin. When OD_{600} reached 0.6, the cells were induced by the addition of isopropyl-β-D-thiogalactopyranoside (IPTG) to a final concentration of 1 mM and growth was continued for another 4 h. Cells were harvested at 7,800×*g* at 4 °C for 15 min and stored at −80 °C until use. For protein purification, cells were suspended in 15 ml of ice-cold Ni^{2+}-affinity buffer (20 mM Na-phosphate buffer (pH 7.4), 0.5 M NaCl, 0.1 mM PMSF, 20 mM imidazole, 1 mM DTE) and disrupted using a bead-beater (three times for 30 s). After centrifugation at 30,000×*g* for 30 min at 4 °C to get rid of the cell debris, the clear supernatant was loaded directly to a 5 ml HisTrap FF column (GE Healthcare) equilibrated with Ni^{2+}-affinity buffer. A linear gradient from 20 to 500 mM imidazole was applied. Column fractions were analyzed by 10% sodium dodecyl sulfate-polyacrylamide gel electrophoresis (SDS-PAGE). The fractions containing the fusion protein were pooled and digested with TEV protease (1:300, TEV protease/fusion protein) overnight at 4 °C during dialysis against TEV protease digestion buffer (50 mM Tris–HCl (pH 7.9), 1 mM $MgCl_2$, 1 mM DTE, and 300 mM KCl). After digestion, the sample was reapplied to a second HisTrap FF column and the purified *Gkaue* polI was collected in the flowthrough. The purified protein was concentrated, dialyzed overnight against polymerase storage buffer (20 mM Tris–HCl (pH 7.9), 100 mM KCl, 0.1 mM EDTA, 1 mM DTE, 50% glycerol, and 0.1% Triton-X 100), and stored at −20 °C. Protein concentration was determined using the Bradford assay [24].

Biochemical Characterization Assays

The polymerase activity of *Gkaue* polI was measured under standard DNA polymerizing assay conditions: The reaction mixture (in 50 μl) contained 0.2 mM of each dCTP, dTTP, and dGTP; 0.2 mM [^{3}H]dATP (specific activity of 400 cpm/pmol); and 0.5 μg activated calf thymus DNA (having 65% of single-stranded form) in polymerase reaction buffer 920 mM Tris–HCl (pH 8.0), 50 mM KCl) including 1.5 mM $MgCl_2$. Reactions were started by the addition of 2 pmol of *Gkaue* polI into prewarmed reaction mixture at 70 °C. Reactions were terminated by the addition of ice-cold 10% TCA containing 10 mM sodium pyrophosphate and the samples were left on ice until filtration. Precipitates were collected on glass fiber filters (Whatman, GF/C), washed three times with ice-cold 10% TCA then with ice-cold isopropanol, dried, and counted in a liquid scintillation counter.

The optimum temperature of *Gkaue* polI was determined by measuring the polymerase activity at a temperature range from 40 to 100 °C for 15 min under standard DNA polymerizing

assay conditions. In order to test thermostability, the enzyme was portioned and incubated at a temperature range from 50 to 90 °C for time intervals from 1 to 60 min. After this incubation, the samples were left on ice until each sample was prewarmed to 70 °C and their polymerase activity was assayed at 70 °C for 15 min under standard DNA polymerizing assay conditions. When added, 0.1 mg/ml BSA was present in the reaction mixtures as a stabilizer.

The pH optimum of *Gkaue* polI was determined by assaying polymerase activity at 70 °C for 15 min in a triple buffer system containing 50 mM of each of Bis-Tris propane, *N*-cyclohexyl-2-aminoethanesulfonic acid (CHES), and 2-(*N*-morpholino)ethanesulfonic acid (MES). All three of these buffers were mixed together before adjusting the pH to a series of pH values between 5.0 to 9.5 (with HCl) so that changes in the activity of *Gkaue* polI that might be caused by different buffer components were eliminated.

The optimum concentrations of divalent ions (0 to15 mM of $MgCl_2$ or $MnCl_2$) and monovalent ions (0 to 300 mM of KCl, NaCl, or NH_4Cl) were determined under standard DNA polymerizing assay conditions at 70 °C for 15 min. When indicated, polymix buffer (5 mM potassium phosphate (pH 7.5), 95 mM KCl, 5 mM NH_4Cl, 5 mM Mg^{2+}-acetate, 0.5 mM $CaCl_2$, 8 mM putresine, 1 mM spermidine, and 1 mM dithioerythritol) was used as the polymerization reaction buffer. Polymix buffer or DNA polymerization buffer were prepared without Mg^{2+} when the enzyme activity was compared in two buffer systems as a function of Mg^{2+} concentrations (from 1 to 10 mM).

Nucleotide Sequence Accession Numbers

The GenBank accession numbers for the *polA* genes from seven *Geobacillus* species identified in this work are presented in Table 3.

Results and Discussion

Sequence Analysis of *Geobacillus* DNA Polymerase I

In this study, we report new members of family A DNA polymerases from seven recently identified thermophilic *Geobacillus* species. After PCR-based cloning, the entire nucleotide sequences of the *polA* genes, including the Shine–Dalgarno sequence, translational start and stop sites of the proteins were identified and submitted to GenBank at NCBI (Table 3). The nucleotide sequences of the *polA* genes from all seven *Gebacillus* species contain 2,637 bp

Table 3 The list of *Geobacillus* species with their GenBank accession numbers for the *polA* genes identified in this work

Geobacillus species	Source	GenBank accession numbers of *polA* genes
Geobacillus anatolicus	Hisaralan, 98 °C	DG810290
G. anatolicus strain C4	Hisaralan, 80 °C	FJ215760
Geobacillus bogazici	Hisaralan, 96 °C	FJ215758
Anoxybacillus sp. NB	Hisaralan, 85 °C	FJ215762
Geobacillus kaue strain NB	Gönen, 68 °C	FJ215757
G. kaue strain MC	Gönen, 68 °C	FJ215759
G. kaue strain E1	Gönen, 77 °C	FJ215761

coding a protein with 877 amino acids (excluding the stop codon) with a calculated molecular mass of 99.3 kDa.

The phylogenetic analysis based on *polA* genes demonstrated that *Geobacillus* species isolated from the same or geographically nearby hot springs, from Gönen, (*G. kaue*) and from Hisaralan (*Geobacillus anatolicus*) diverged together in the phylogenetic tree, as expected (Fig. 1). The phylogenetic analysis based on 16S rRNA sequences confirmed that these bacteria are indeed different from each other (Fig. 2).

We named one of the strains as an *Anoxybacillus* (*Anoxybacillus* sp. strain NB) rather than as a *Geobacillus* because the closest relative to *Anoxybacillus* sp. strain NB is *Anoxybacillus flavithermus* strain WK1 showing 98% sequence similarity based on 16S rRNA sequence analysis (GenBank: CP000922). In fact, the genus *Anoxybacillus* has recently been branched out from the genus *Geobacillus* [25, 26]. However, when the *polA* gene sequences were compared, the closest relative to *Anoxybacillus* sp. strain NB is still a member of genus *Geobacillus* (*Geobacillus* sp. MKK, GenBank: DQ244056) having 98% sequence similarity (Table 4). It seems *Anoxybacillus* sp. strain NB is in the borderline between these two closely related genus.

Based on the amino acid sequences of DNA polymerases, the evolutionary and similarity analyses indicated that all seven *Geobacillus* DNA polymerases in this work are closely related (Fig. 1) and show strong homology to family A DNA polymerases from other known members of genus *Geobacillus* rather than the genus *Thermus* or *E. coli* DNA polymerase I (Table 4). Most of the amino acid substitutions were within the similar functional groups, especially within uncharged amino acids. Such preferences have also been reported for some other thermophilic proteins other than DNA polymerases, suggesting the role of these residues in thermostability when compared with their mesophilic counterparts [27].

Amino acid sequence comparison at the C-terminal polymerase domain of the seven *Geobacillus* DNA polymerases of this work with 16 other well-known bacterial family A DNA polymerases confirmed the presence of highly conserved and crucial sequence motifs for polymerase [5, 28] and 5′-3′ exonuclease [29] activities (Fig. 3a, b, respectively). The amino acid variations were found mostly at the central part of the protein which is

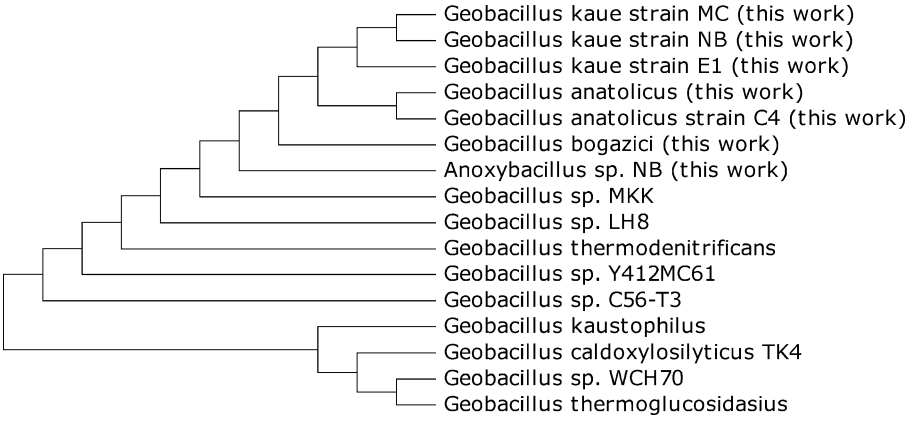

Fig. 1 Evolutionary relationship of known *Geobacillus* species together with seven *Geobacillus* species of this work based on DNA polymerase I sequences. Evolutionary analysis was conducted by MEGA [20] and inferred using the neighbor-joining method [21]. All positions containing gaps and missing data were eliminated. The GenBank accession numbers of the *polA* genes are presented in Tables 3 and 4

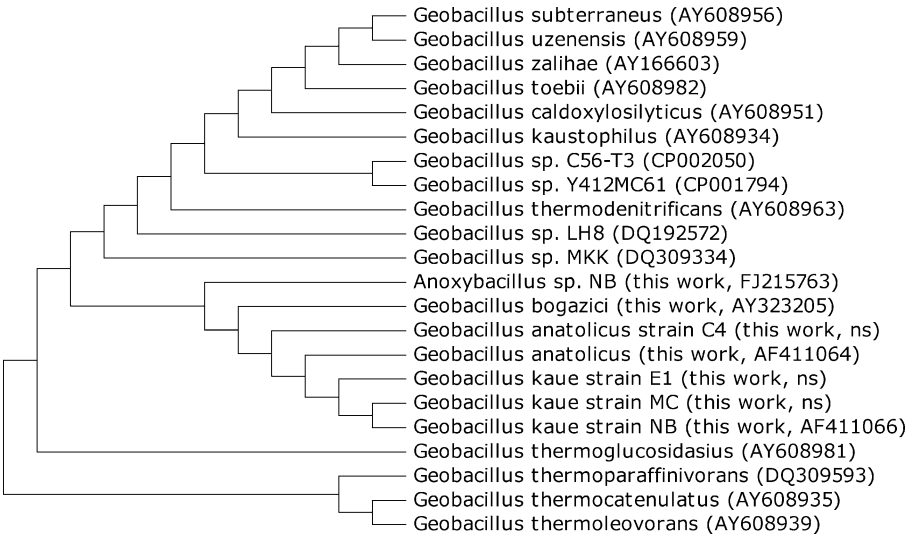

Fig. 2 Evolutionary relationship of some *Geobacillus* species together with seven *Geobacillus* species of this work based on 16S rRNA sequences. The GenBank accession numbers are presented in parenthesis. Evolutionary analysis was conducted similarly as depicted in Fig. 1

responsible for 3′-5′ exonuclease (proofreading) activity of the enzyme [30, 31]. The roles of the carboxylate groups (Asp^{355}, Glu^{357}, Asp^{424}, and Asp^{501}) in 3′-5′ exonuclease activity that are coordinated by divalent metal ions were reported before in *E. coli* DNA polymerase I [30]. These residues correspond to Val^{321}, Glu^{323}, Ala^{378}, and Lys^{452}, respectively, in all

Table 4 Similarity analysis of *Gkaue* polI with 16 well-known members of family A DNA polymerases

Species	Similarity to *Gkaue* polI (%)	Genbank accession numbers
Escherichia coli	55	V00317
Thermus thermophilus	59	D28878
Thermus aquaticus	59	D32013
Thermus filigormis	57	AF030320
Bacillus caldotenax	94	D12982
Bacillus stearothermophilus	93	BSU33536
Bacillus subtilis	82	U86402
Geobacillus sp. MKK	98	DQ244056
Geobacillus sp. LH8	98	DQ392964
Geobacillus thermodenitrificans	96	Z26928
Geobacillus sp Y412MC61	95	ACX77456
Geobacillus sp C56-T3	94	YP003670385
Geobacillus kaustophilus	94	YP148583
Geobacillus caldoxylosilyticus TK4	89	DQ340803
Geobacillus sp. WCH70	89	YP002950632
Geobacillus thermoglucosidasius	88	ZP06809476

Homology search was performed using Basic Local Alignment Search Tool based on amino acid sequences of the *polA* genes

Geobacillus DNA polymerases described in this work, indicating that three of the four critical metal-binding residues are different in these proteins (Fig. 3c). Due to these changes in metal-binding residues, the potential DNA and nucleotide-binding sites were altered, resulting in the loss of 3′-5′ exonuclease activity. Similar observations had been reported earlier for other thermophilic DNA polymerases from *B. stearothermophilus* [15] and from *Thermus aquaticus* [32].

Since *Geobacillus* species identified in this work were closely related, we only characterized one of them, *G. kaue* strain NB DNA polymerase I (*Gkaue* polI, GenBank: FJ215757), biochemically in vitro.

Structure of *G. kaue* DNA Polymerase I

The nucleotide sequence and deduced amino acid sequence of *Gkaue* polI was presented in Fig. 4. The 3D structure of *Gkaue* polI was predicted by homology modeling using BF as a template structure [15, 16] since it has a high sequence similarity (90%) with the corresponding region (between residues 299–878) of the full-length *Gkaue* polI. The structure for DNA polymerase domain (residues 471–878) of *Gkaue* polI was depicted resembling a right hand with fingers, palm, and thumb subdomains (Fig. 5a). These three subdomains define a common structural feature even in distantly related DNA polymerases and predicted to bind duplex DNA [4]. Root-mean-square deviation (RMSD) of equivalent α-carbon positions of *Gkaue* polI with respect to the template protein (BF) is 0.73 Å.

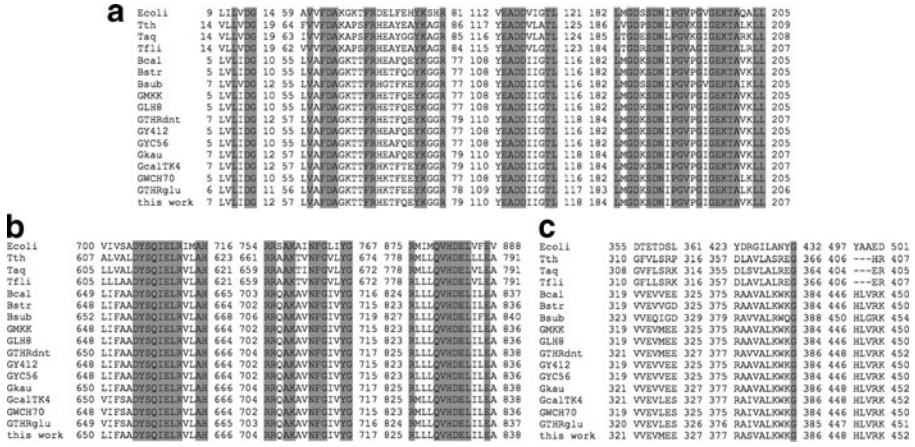

Fig. 3 The amino acid sequence alignment of 16 well-known members of family A DNA polymerases together with the consensus sequences of seven *Geobacillus* species in this work. The conserved amino acid residues involved in **a** 5′-3′ exonuclease, **b** polymerase, and **c** 3′-5′ exonuclease domains. The GenBank accession numbers of the *polA* genes and abbreviations of the species are as presented in Tables 3 and 4. The abbreviations of the species are as follows: *Escherichia coli* (Ecoli), *Thermus thermophilus* (Tth), *Thermus aquaticus* (Taq), *Thermus filiformis* (Tfli), *Bacillus caldotenax* (Bcal), *Bacillus stearothermophilus* (Bstr), *Bacillus subtilis* (Bsub), *Geobacillus sp. MKK* (GMKK), *Geobacillus* sp. *LH8* (GLH8), *Geobacillus thermodenitrificans* (GTHRdnt), *Geobacillus* sp. *Y412MC61* (GY412), *Geobacillus* sp. *C56-T3* (GYC56), *Geobacillus kaustophilus* (Gkau), *Geobacillus caldoxylosilyticus TK4* (GcalTK4), *Geobacillus* sp. WCH70 (GWCH70) *Geobacillus thermoglucosidasius* (GTHRglu)

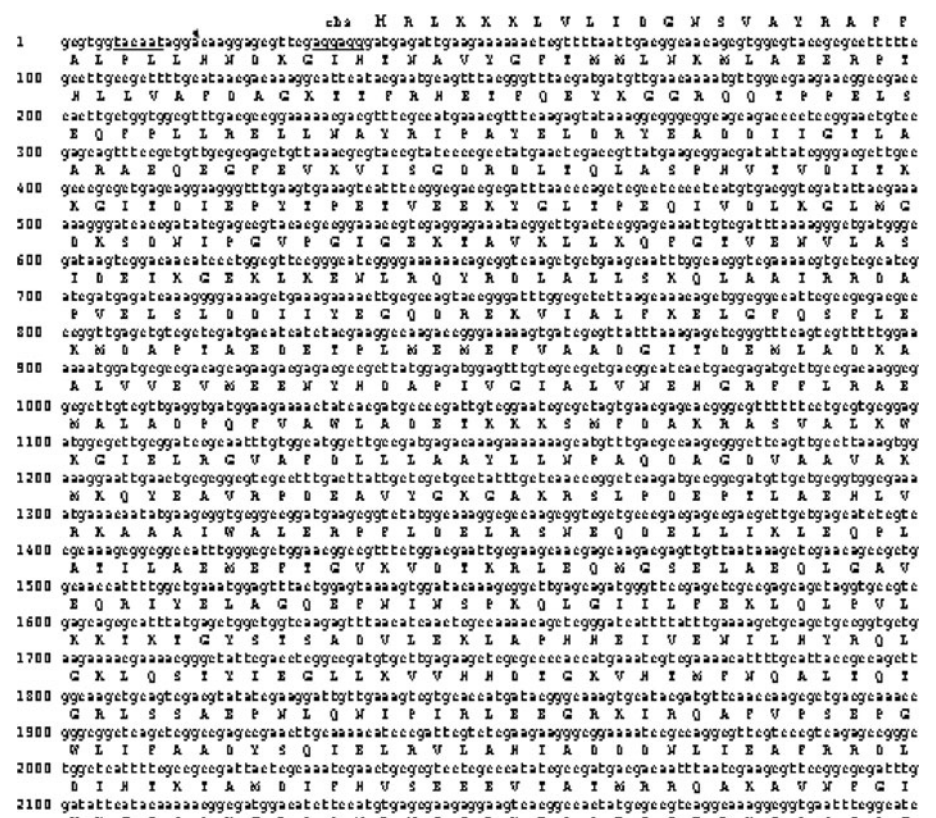

Fig. 4 Nucleotide sequence and deduced amino acid sequence of *Gkaue* polI. The predicted Shine–Dalgarno region located upstream from the ATG start codon (M) and −10 region prokaryotic promoter motif are *underlined*. The putative transcription start point is indicated by a *bent arrow*. The translation stop codon (UAA) is indicated by an *asterisk*

Furthermore, the polymerase active site residing on the palm subdomain at C-terminal polymerase domain of *Gkaue* polI (residues 617–834) revealed highly conserved motifs A, B, and C together with strictly conserved residues (Fig. 5b). For example, Tyr^{714} (Tyr^{716} in *Gkaue* polI) is positioned at the first nucleotide of 5′-template overhang and interacts with the template base. The location of the 3'-OH terminus is positioned for catalysis, which forms hydrogen bond to Asp^{830} (Asp^{832} in *Gkaue* polI), an essential and invariant residue in all polymerases.

Expression and Purification of *G. kaue* DNA Polymerase I

Gkaue polI was expressed into *E. coli* JM109(DE3) cells (Fig. 6a). The fusion protein, including the additional region (six histidine residues, thioredoxin (TrxA) coding sequence

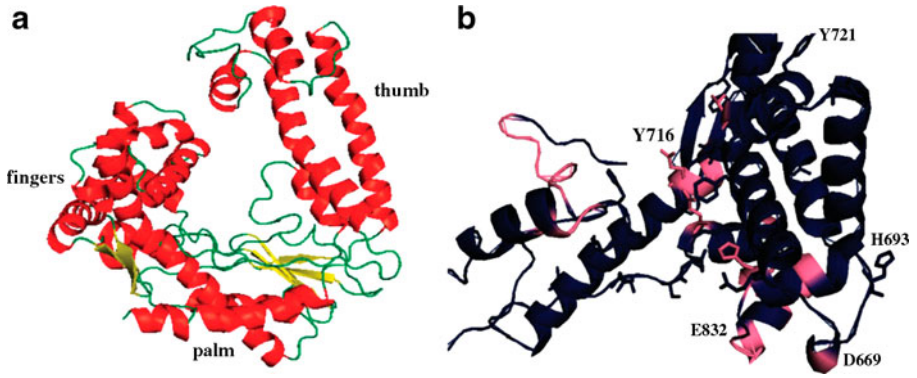

Fig. 5 Three-dimensional structures of *Gkaue* polI. **a** The predicted model for the C-terminal DNA polymerase domain of *Gkaue* polI. Helix, sheet, and loop were colored red, yellow, and green, respectively. **b** The predicted model for the active site of *Gkaue* polI. The highly conserved amino acid residues (motifs A, B, and C) in polymerase domains (shown in Fig. 2b) were colored *pink*. The positions of amino acid residues were depicted based on BF structure [15, 16]

and a TEV protease cleavage site), has a total molecular mass of 113.6 kDa. The fusion *Gkaue* polI was purified by Ni^{2+}-affinity chromatography and then digested with TEV protease. A second Ni^{2+}-affinity chromatography after TEV digestion enabled an efficient removal of the histidine-tag, as well as the contaminating cellular proteins carried over from the first chromatography step. The purified *Gkaue* polI (99.3 kDa) was homogeneous as judged by SDS-PAGE analysis (Fig. 6b); 2.6 mg of purified enzyme was obtained from 2.5 l of cell culture (Table 5).

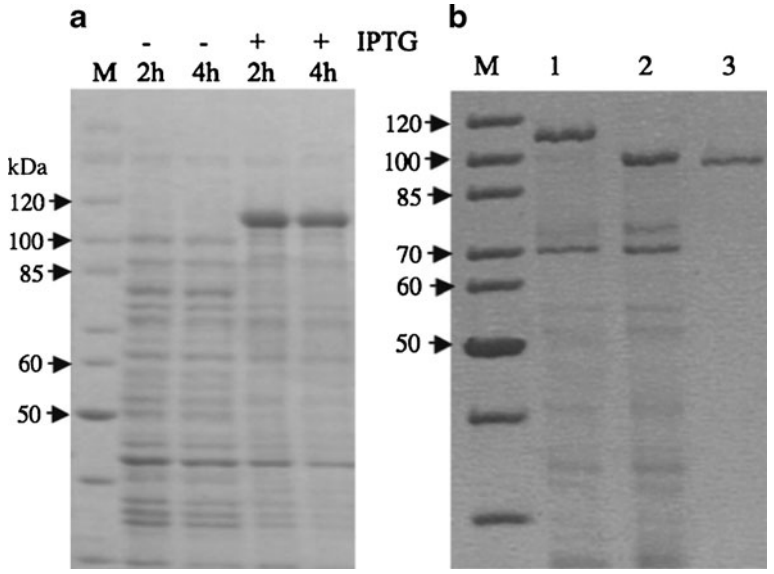

Fig. 6 SDS-PAGE analysis of the expression and the purification of *Gkaue* polI. **a** Analysis of the protein expression before and after IPTG induction. *Lane M*, molecular weight standard. **b** Analysis of the protein purification. *Lane 1*, undigested fusion protein; *lane 2*, TEV protease digested fusion protein; *lane 3*, purified enzyme after TEV protease digestion and after second Ni^{2+}-affinity chromatography

Table 5 Purification scheme of *Gkaue* polI

Step	Total protein (mg)	Total activity (U)	Specific activity (U/mg)	Yield (%)
Ni^{2+}-column purification	36	766	21	100
TEV digestion	13	350	27	46
Dialysis	2.7	100	37	14

Biochemical Properties of *Geobacillus* DNA Polymerase I

This is the second study, to our knowledge, a thermophilic family A DNA polymerase I from the genus *Geobacillus* was characterized biochemically. The first one was from DNA polymerase I from *Geobacillus caldoxylosilyticus* TK4 [14].

Optimum conditions for the enzymatic activity of *Gkaue* polI were measured in vitro using [^{3}H]dATP incorporation to a partially single-stranded DNA substrate (see "Materials and Methods"). The maximum polymerization activity was observed at 70 °C (Fig. 7a). The

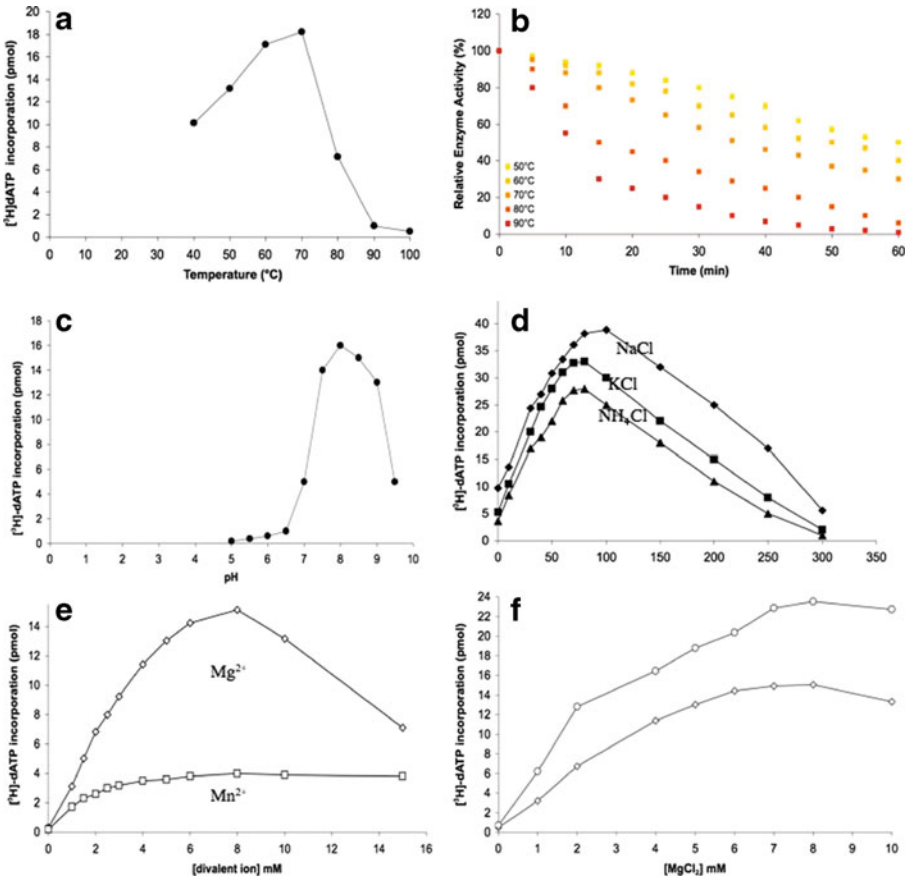

Fig. 7 The optimum conditions of *Gkaue* polI. **a** Optimum temperature. **b** Thermostability. **c** Optimum pH. Optimum concentrations for **d** monovalent ions and **e** divalent ions. **f** Optimum Mg^{2+} concentration in the presence of standard polymerase reaction buffer (*empty diamonds*) and polymix buffer (*empty circles*)

temperature optimum of the *Gkaue* polI agrees well with the temperature at which the source organism *G. kaue* strain NB was isolated originally (Gönen hot spring, 68 °C). *G. kaue* also grows optimally at 70 °C in the laboratory.

The thermostability of the enzyme was also investigated after the heat treatment of the protein over a temperature range (from 40 to 100 °C) for various incubation times (from 0 to 60 min) and subsequent polymerization assay at 70 °C. We have found that *Gkaue* polI lost 50% of its activity above 80 °C within 15 min (Fig. 7b). BSA did not improve the thermal stability of the enzyme (data not shown).

The pH optimum for *Gkaue* polI activity was measured in a triple buffer system including Bis-Tris propane, CHES, and MES hydrate, in the pH range from 5.5 to 9.0. All three of these buffers were used together in the same mixture in order to eliminate the effects of different buffer components on the enzyme activity. *Gkaue* polI activity was maximum at pH 7.5–8.5 and declines sharply after pH 9.0 (Fig. 7c).

Gkaue polI activity requires the presence of monovalent ions. The optimum concentrations for NaCl, KCl, or NH_4Cl are 100, 85, or 80 mM, respectively. Above 100 mM, monovalent salts become inhibitory for the activity (Fig. 7d). *Gkaue* polI activity was absolutely dependent on the presence of divalent cations Mg^{2+} and Mn^{2+}. The highest enzyme activity was observed at 8 mM Mg^{2+}. Mn^{2+} could not be replaced with Mg^{2+} completely (Fig. 7e).

We also measured the DNA polymerization activity of *Gkaue* polI in polymix buffer and compared it in parallel with the conventional Tris-buffer. The polymix buffer, containing polyamines putrescine and spermidine, had been optimized for both rate and accuracy of protein synthesis in bacteria [33]. Polymix buffer provides the highest fidelity at the highest protein synthesis rates for ribosomes in vitro [34]. We observed 1.5- to 2-fold higher polymerase activity for *Gkaue* polI in the presence of the polymix buffer when compared with the conventional Tris-based polymerase buffer (Fig. 7f).

In addition to the improvement in the DNA polymerization activity that we observed in this work, the effect of the polymix buffer on the fidelity of DNA polymerization should now be investigated.

Acknowledgments This work was supported by Bogazici University Research Fund grants BAP-08B102D and BAP-03S102. We thank Dr. Ditlev Brodersen for the kind gift of TEV protease.

References

1. Bernad, A., Zaballos, A., Salas, M., & Blanco, L. (1987). *EMBO Journal, 6*, 4219–4225.
2. Filee, J., Forterre, P., Lin, T. S., & Laurent, J. (2002). *Journal of Molecular Evolution, 54*, 763–773.
3. Joyce, C. M., & Steitz, T. A. (1994). *Annual Review of Biochemistry, 63*, 777–822.
4. Patel, P. H., Suzuki, M., Adman, E., Shinkai, A., & Loeb, L. A. (2001). *Journal of Molecular Biology, 308*, 823–837.
5. Delarue, M., Poch, O., Tordo, N., Moras, D., & Argos, P. (1990). *Protein Engineering, 3*, 461–467.
6. Steitz, T. A. (1999). *Journal of Biological Chemistry, 274*, 17395–17398.
7. Nazina, T. N., Tourova, T. P., Poltaraus, A. B., Novikova, E. V., Grigoryan, A. A., Ivanova, A. E., et al. (2001). *International Journal of Systematic and Evolutionary Microbiology, 51*, 433–446.
8. Kaboev, O. K., Luchkina, L. A., Akhmedov, A. T., & Bekker, M. L. (1981). *Journal of Bacteriology, 145*, 21–26.
9. Phang, M. S., Teo, C. Y., Lo, E., & Wong, V. W. T. (1995). *Gene, 163*, 65–68.
10. Sellmann, E., Schröder, K. L., Knoblich, I. M., & Westermann, P. (1992). *Journal of Bacteriology, 174*, 4350–4355.
11. Burrows, J. A., & Goward, C. R. (1992). *Biochemistry Journal, 287*, 971–977.
12. Uemori, T., Ishino, Y., Fujita, K., Asada, K., & Kato, I. (1993). *Journal of Biochemistry, 113*, 401–410.
13. Kondori, K. M., Sadeghizadeh, M., Khajeh, K., Manesh, H. N., Ahadi, A. M., & Emamzadeh, A. (2007). *Applied Biochemistry and Biotechnology, 142*, 200–208.

14. Sandalli, C., Singh, K., Modak, M. J., Kethar, A., Canakci, S., Demir, I., et al. (2009). *Applied Microbiology and Biotechnology, 84*, 105–117.
15. Kiefer, J. R., Mao, C., Hansen, C. J., Basehore, S. L., Hogrefe, H. H., Braman, J. C., et al. (1997). *Structure, 5*, 95–108.
16. Kiefer, J. R., Mao, C., Braman, J. C., & Beese, L. S. (1998). *Nature, 391*, 304–307.
17. Aliota, J. M., Pelletier, J. J., Ware, J. L., Moran, L. S., Benner, J. S., & Kong, H. (1996). *Genetic Analysis, 12*, 185–195.
18. Lu, Y. Y., Ye, S. Y., & Hong, G. F. (1991). *Biotechniques, 11*, 464–466.
19. Thompson, J. D., Higgins, G., & Gipson, T. J. (1995). *Nucleic Acids Research, 22*, 4673–4680.
20. Tamura, K., Dudley, J., Nei, M., & Kumar, S. (2007). *Molecular Biology and Evolution, 24*, 1596–1599.
21. Saitou, N., & Nei, M. (1987). *Molecular Biology and Evolution, 4*, 406–425.
22. Schwede, T., Koop, J., Guex, N., & Peitsch, M. C. (2003). *Nucleic Acids Research, 31*, 3381–3385.
23. Nayeem, A., Sitkoff, D., & Krystek, S. (2007). *Protein Science, 15*, 808–824.
24. Bradford, M. M. (1976). *Analytical Biochemistry, 72*, 248–254.
25. Pikuta, E., Lysenko, A., Chuvilskaya, N., Mendrock, U., Hippe, H., Suzina, N., et al. (2000). *International Journal of Systematic and Evolutionary Microbiology, 50*, 2109–2117.
26. Saw, J. H., Mountain, B. W., Feng, L., Omelchenko, M. V., Hou, S., Saito, J. A., et al. (2008). *Genome Biology, 9*, R161.
27. Jaenicke, R., & Böhm, G. (1998). *Current Opinion of Structural Biology, 8*, 738–748.
28. Polesky, A. H., Steitz, T. A., Grindley, N. D., & Joyce, C. M. (1990). *Journal of Biological Chemistry, 265*, 14579–14591.
29. Gutman, P. D., & Minton, K. W. (1993). *Nucleic Acids Research, 21*, 4406–4407.
30. Beese, L. S., & Steitz, T. A. (1991). *EMBO Journal, 10*, 25–33.
31. Derbyshire, V., Grindley, N. D., & Joyce, C. M. (1991). *EMBO Journal, 10*, 17–24.
32. Korolev, S., Nayal, M., Barnes, W. M., DiCera, E., & Waksman, G. (1995). *Proceedings of the National Academy of Sciences, 92*, 9264–9268.
33. Jelenc, P. C., & Kurland, C. G. (1979). *Proceedings of the National Academy of Sciences, 76*, 3174–3178.
34. Ehrenberg, M., Bilgin, N., & Kurland, C. G. (1989). In Spedding G. (Ed), *Ribosomes and Protein Synthesis* (pp. 101–129). Oxford: IRL Press.