Extremophiles (2000) 4:215–225 

ORIGINAL PAPER

**Marie-Anne Cambon-Bonavita · Philippe Schmitt
Montserrat Zieger · Jean-Michel Flaman
Françoise Lesongeur · Gérard Raguénès · Danièle Bindel
Nicolas Frisch · Zeinab Lakkis · Daniel Dupret
Georges Barbier · Joël Quérellou**

# Cloning, expression, and characterization of DNA polymerase I from the hyperthermophilic archaea *Thermococcus fumicolans*

Received: September 17, 1999 / Accepted: March 21, 2000

**Abstract** The DNA polymerase I gene of a newly described deep-sea hydrothermal vent Archaea species, *Thermococcus fumicolans*, from IFREMERS's collection of hyperthermophiles has been cloned in *Escherichia coli*. As in *Thermococcus litoralis*, the gene is split by two intervening sequences (IVS) encoding inteins inserted in sites A and C of family B DNA polymerases. The entire DNA polymerase gene, containing both inteins, was expressed at 30°C in *E. coli* strain BL21(DE3)pLysS using the pARHS2 expression vector. The native polypeptide precursor of 170 kDa was obtained, and intein splicing as well as ligation of the three exteins was observed in vitro after heat exposure. The recombinant enzyme was purified and some of its activities were characterized: polymerization, thermostability, exonuclease activities, and fidelity.

**Key words** Archaea · *Thermococcus fumicolans* · DNA polymerase · Exonuclease · Intein · Thermostability · Fidelity

Communicated by G. Antranikian

M.-A. Cambon-Bonavita · F. Lesongeur · G. Raguénès. · G. Barbier
Laboratoire de Caractérisation des Microorganismes Marins, DRV/VP, Ifremer centre de Brest, Plouzané, France

P. Schmitt · M. Zieger · D. Bindel · N. Frisch · Z. Lakkis
Appligene-Oncor, Parc d'innovation, rue Geiler de Kaysersberg, Illkirch cedex, France

J.-M. Flaman
Universite de Rouen, C.H.U Laboratoire de Génétique Moléculaire, 76031 Rouen, France

D. Dupret
Proteus, Nimes, France

J. Quérellou[1] (✉)
Laboratoire de biotechnologie des microorganismes hydrothermaux, DRV/VP, Ifremer Centre de Brest, Plouzané, France

[1] *Present address*: Ifremer, Centre de Brest, BP 70, 29280 Plouzané, France
Tel. +(33) 2 98 22 40 01; Fax +(33) 2 98 22 46 00
e-mail: jquerell@ifremer.fr

## Introduction

The discovery of a third Domain of life called Archaea (Woese and Fox 1977; Woese et al. 1990), opened up a new field of investigation. Since that time, more and more hyperthermophilic Archaea have been isolated, some of them collected from deep-sea hydrothermal vents. These organisms have raised considerable interest for both basic research and biotechnological applications. Among the different applications, the most outstanding is the polymerase chain reaction, or PCR (Mullis et al. 1986), based on the use of thermostable *Taq* DNA polymerase. In turn, this discovery raised interest in *Taq* and other thermostable and hyperthermostable DNA polymerases for commercial use and new applications that can advance molecular biology and related areas. As a result, DNA polymerases have been characterized mainly from three genera: *Thermus*, *Thermococcus*, and *Pyrococcus*. The two last genera amalgamate strictly hyperthermophile species and their enzymes, both native and recombinant, that are among the most hyperthermostable ever known (Baross and Holden 1996; Perler et al. 1996). Many *Taq* DNA polymerase variants have also been engineered (Ignatov et al. 1998; Kebelmann-Betzing et al. 1998; Park et al. 1997).

*Thermococcales* have been studied mainly for biotechnological applications and to a minor extent for phylogenetic considerations, especially since the discovery of inteins in their DNA polymerases (Bult et al. 1996; Perler et al. 1992, 1996). Inteins, polypeptide products resulting from transcription and translation of intervening sequences in-frame with their host genes, are removed from the polypeptide precursor by protein splicing (Xu and Perler 1996). Inteins have been found in the three domains: Bacteria, Archaea, and Eukarya (Perler 1999; Perler et al. 1997). Information related to inteins can be found in the intein database: InBase (*http://www.neb.com/neb inteins.html*) and the intein nomenclature used here is derived from InBase. Among Archaea, several *Thermococcales* species have inteins: in their DNA polymerase I genes, including *Pyrococcus* sp., strain GB-D (1 intein) (Xu

et al. 1993), *Thermococcus litoralis* (Perler et al. 1992), and *Pyrococcus kodakaraensis* KOD1 (2 inteins) (Takagi et al. 1997), whereas *Thermococcus aggregans* TY (Canganella et al. 1998) displays 3 putative inteins (Niehaus et al. 1997). The complete genome sequence of *Methanococcus jannaschii* has revealed (Bult et al. 1996; Gorbalenya 1998) 15 genes containing 19 putative inteins, including the 2 inteins in the DNA polymerase I gene, in the same location as those in *P. kodakaraensis* KOD1. More recently, a new DNA polymerase family was found in *P. furiosus* (Cann et al. 1998), and the corresponding subunit 2 genes in *P. horikoshii* also host an intein (Kawarabayasi et al. 1998).

Here we report the cloning, sequencing, and expression in *E. coli* of a new DNA polymerase gene as well as the preliminary characterization of the recombinant enzyme, isolated from *T. fumicolans*, a new deep-sea hydrothermal vent archaeon (Godfroy et al. 1996).

## Materials and methods

### Origin and isolation of archaeal strain

*Thermococcus fumicolans* was isolated from hydrothermal chimney wall fragments collected by the French research submersible *Nautile*, during the French-Japanese cruise "Starmer" (1989; 16°59′S, 173°55′W; depth, 2000 m) on the subduction area of the North Fiji Basin. Active chimneys are mainly composed of anhydrites and fluid temperatures that reach 288°C. Pure cultures were obtained on plates with medium containing solid sulfur (Godfroy et al. 1996). The isolate, based on microbiological considerations, 16*S* rDNA sequence analysis, and quantitative DNA–DNA hybridization (Meunier 1994), was identified as a new *Thermococcus* species. The strain was grown anaerobically at 90°C, without being shaken, in a broth containing 2 g/l peptone; 0.5 g/l yeast extract; 30 g/l sea salt (Sigma); 6.05 g/l PIPES buffer; 10 g/l elemental sulfur; and 1 mg/l rezasurin. pH was adjusted at 8.5 with 5 N NaOH at room temperature. Besides its thermophilic properties, this strain was selected for isolating the DNA polymerase gene because qualitative DNA–DNA hybridizations (dot-blot) with *T. litoralis* and *P. furiosus* produced a weak signal (Meunier 1994), and also because the DNA polymerase probes designed from *T. litoralis* and *P. furiosus* did not hybridize at medium stringency level (unpublished results). *Pyrococcus* sp. strain GE23 was provided by Dr. G. Erauso (CNRS, Station Biologique de Roscoff, France).

### *Escherichia coli* strains

*Escherichia coli* strain SURE (Stratagene, La Jolla, CA, USA) was grown in LB medium with appropriate antibiotics at 37°C with vigorous shaking. NovaBlue, BL21(DE3), and BL21(DE3)pLysS (Novagene, Madison, WI, USA) strains were cultured in 2xYT medium with appropriate antibiotics at 37°C or 30°C.

### Probe construction and DNA hybridization

Probe design was based on the availability of the DNA polymerase gene of *Pyrococcus* sp. GE23, obtained previously in the laboratory (accession number, Z54173). This gene, amplified by PCR, using primers containing the *Nde*I site in 5′ and *Sal*I site in 3′, was inserted in the pARHS2 expression vector. Two probes were prepared by enzymatic digestion of the GE23 DNA pol gene, one by *Cla*I–*Hind*III corresponding to position bp 8 to bp 1353 (GE23 "BEG.") and the second by *Xho*I–*Sal*I corresponding to bp 1879–2318 (GE23 "END."). An additional probe (TFU "MID.") was produced by PCR using *T. fumicolans* genomic DNA and the following primers: 5′-CACACCGAACCATAAGCTT-3′ and 5′-GGATTCCTCTACGCAAGCTT-3′. DNA hybridization was done according standard protocols (Southern 1975). *T. fumicolans* genomic DNA was digested with five restriction enzymes (Boehringer, Mannheim, Germany): *Bam*HI, *Hind*III, *Eco*RI, *Xba*I, or *Xho*I. Single or double digestion products were fractionated on a 0.8% agarose gel and vacuum transferred (VacuGene; Pharmacia, Upsala, Sweden) to Hybond N+ membrane (Amersham, Little Chalfont, UK). Probes were labeled with α-$^{32}$PdCTP by random priming (Amersham) and hybridized at 50°C and 55°C to the attached fractionated genomic DNA on the membrane according to standard procedures (Sambrook et al. 1989).

### Construction and screening of genomic sublibraries

*Thermococcus fumicolans* DNA was cut by the appropriate restriction endonuclease to obtain desirable fragments, previously identified by Southern hybridization. Digested DNA was separated by electrophoresis on a 0.8% agarose gel and eluted following dialysis by standard procedures (Sambrook et al. 1989) or by GeneClean purification (Bio 101, Vista, CA, USA). Fragments were then ligated into pUC18 vector (Appligene-Oncor, Illkirsh, France) and the recombinant plasmid was used to transform *E. coli* SURE cells. Transformed cells were plated on LB-Ampicillin or 2xYT-Ampicillion (100 μg/ml) agar plates and cultured at 37°C overnight. Colony hybridizations were carried out according to standard procedures (Grunstein and Hogness 1975), using the previously mentioned $^{32}$P-labeled probes. Positives clones were cultured at 37°C and DNA inserts were characterized following plasmid DNA extraction (Birnboim and Doly 1979). Appropriate constructions were sequenced using the dideoxynucleotide chain termination method (Sanger et al. 1977) on Applied Biosystems automatic sequencers (Europe Gene Sequence Services [ESGS], Paris, France).

### Sequence analysis

The multiple alignments of *Thermococcales* DNA polymerases were performed with the CLUSTAL W program (Thompson et al. 1994) using default parameters. Final alignment was refined manually to obtain putative intein

coding sequences properly inserted in the DNA polymerase coding section, on the basis of published results (Perler 1999; Perler et al. 1997).

## Subcloning for expression

The *T. fumicolans* DNA polymerase gene was amplified by PCR using *Taq* Extender (Stratagene) with the Stratagene 96 gradient thermocycler. Only 20 cycles were performed to limit errors. PCR was performed with 100 ng of genomic DNA and 100 pmol of each primer, which were added to the standard reaction mix in a final volume of 50 µl containing 200 µM of each dNTP, 100 µl of 10× *Taq* Extender reaction buffer, 2.5 U of *Taq* polymerase (Appligene), and 2.5 µl of *Taq* extender. PCR were performed with one denaturation step at 94°C for 1 min, followed by 20 cycles with temperature profile of 1 min at 94°C, 1 min at 52°C, 5 min at 72°C, and a final extension of 6 min at 72°C. The DNA polymerase gene was amplified by long-distance PCR using the whole ORF containing sequences coding for both inteins. A *Nde*I site was inserted at the initiation codon (primer 5′-GGTGGTGG**CATATG**ATCCTCGATACA-GACTAC-3′) and a *Bam*HI site at the termination codon (5′-AAGC**GGATCC**TTCATTTCTTCCCC-3′). The polymerase genes were cloned in pET12a vector (Novagen, Madison, WI, USA) and in pARHS2 (Eurogentec, Seraing, Belgium), a T7 promoter-based expression vector stabilized by the *par*B locus (De Moerlooze et al. 1992).

## Assay for DNA polymerase activity

Samples from expression cultures supposed to contain *Tfu* DNA polymerase were used for incorporation of radiolabeled $^{32}$P-dTTP into a DE81 absorbable DNA paper. One unit of DNA polymerase activity is defined as the amount required to incorporate 10 nmol of dNTP into an acid-insoluble form at 72°C in 30 min. Assays contained enzyme buffer (Tris-HCl, pH 9.0, 50 mM; KCl, 50 mM; $MgCl_2$, 2 mM; BSA, 0.2 mg/ml; and $(NH_4)_2SO_4$, 16 mM), 200 µM each dNTP, $^{32}$P-TTP, and 0.26 mg/ml activated calf thymus DNA (Appligene-Oncor).

## Purification of *Tfu* DNA polymerase from *E. coli* recombinants

The recombinant clones containing the DNA polymerase gene with both inteins, under control of the T7 promoter, were grown in *E. coli* BL21(DE3)pLysS containing T7 RNA polymerase for expression. This strain was grown in a 15-l fermenter at 30°C, in medium containing, per liter, 16 g tryptone, 10 g yeast extract, 5 g NaCl, 100 µg/ml ampicillin, and 15 µg/ml chloramphenicol. When the culture reached the OD (600 nm) of 0.5, induction was initiated by adding 1 mM of isopropyl-1-β-D-thiogalactopyranoside (IPTG). The culture was then incubated for 4 h, and cells were harvested by centrifugation and stored at −80°C. All subsequent steps were performed at 4°C.

Preparation of cell extract. After a 4-h induction with 1 mM IPTG, 142 g of cells were resuspended in 500 ml buffer (Tris-Cl, 20 mM, pH 7.5; NaCl, 10 mM; EGTA, 1 mM; Triton X-100, 1%; 0.2 mM PMSF) and lysed by sonication; 700 ml of a clear crude extract (fraction I) was obtained by centrifugation at 20000 rpm, 4°C, for 25 min (rotor type 45; Beckman). *E. coli* proteins contained in fraction I were partially thermodenaturated 10 min in total up to 70°C maximum (2 × 350 ml in 2 × 500-ml vials) and the lysate was cooled on ice. Insoluble material was removed by centrifugation, 40 min at 20000 rpm, 4°C (fraction II, 650 ml).

Heparine sepharose CL-6B-chromatography. Fraction II was applied to a 70-ml heparine sepharose column CL-6B (20 $cm^2$ × 3.5 cm) equilibrated with buffer A (10 mM Tris-HCl, pH 7.5; 0.5 mM EGTA; 10 mM mercaptoethanol; 0.2% Triton X-100; 10% glycerol) containing 40 mM NaCl. Thermostable DNA polymerase was eluted with a linear gradient of NaCl (0.1–0.8 M) in buffer A. The remaining pooled fractions were warmed at 70°C for 20 min and cooled on ice; insoluble material was removed by centrifugation at 9000 rpm, 4°C, for 30 min (SS-34 rotor; Sorvall) (fraction III, 200 ml).

Blue sepharose fast-flow chromatography. Fraction III was dialyzed against buffer A containing 0.1 M NaCl and applied to a 15-ml blue-sepharose-fast-flow column (5 $cm^2$ × 4 cm), equilibrated with buffer containing 40 mM NaCl. Thermostable DNA polymerase was eluted with a linear gradient of NaCl (0.1–0.7 M) in buffer A. Active fractions were analyzed in 7% SDS-PAGE, and fractions containing more than 95% of 90-kDa protein were pooled and dialyzed against buffer A containing 50 mM NaCl (fraction IV, 117 ml).

MonoQ anion-exchange column. Fraction IV was applied on a anion-exchange column (HR 10/10; Pharmacia) equilibrated with buffer A containing 50 mM NaCl and washed with the same buffer. The DNA polymerase was excluded, and some contaminants detrimental to PCR efficiency retained. Fractions were analyzed for activity and purity. Fractions with the highest activity were pooled (fraction V).

HiTrap SP cation-exchange column. Fraction V was dialyzed against buffer B (40 mM $KPO_4$, pH 7; 0.5 mM EGTA, 10 mM β-mercaptoethanol; 0.05% Triton X-100; 5% glycerol) containing 50 mM KCl. The fraction obtained after dialysis (fraction VI, 130 ml) was applied on a 5-ml cation-exchange column (HiTrap SP; Pharmacia) equilibrated with buffer B containing 40 mM KCl. Thermostable DNA polymerase was eluted with a linear gradient of KCl (0.04–0.5 M) at about 0.15 M KCl. The more active DNA polymerase fractions were pooled and dialyzed against storage buffer (10 mM Tris-HCl, pH 7.5; 100 mM KCl; 0.1 mM EDTA; 1 mM DTT; 0.1% Triton X-100; 50% glycerol) (fraction VII, 19 ml). This fraction VII was stored at −20°C.

## Polymerase assays

The enzyme was tested comparatively to *Tli* DNA polymerase with its own buffer (Tris HCl, pH 9.0, 20 mM; KCl, 10 mM; $MgSO_4$, 2 mM; $(NH_4)_2SO_4$, 10 mM; Triton X-100, 0.1%). This buffer was used as reference buffer for *Tfu* DNA polymerase assays, each component being successively modified during characterization of the DNA polymerase activity. The second buffer used in parallel was the *Taq* DNA polymerase buffer (Appligene-Oncor) (Tris HCl, pH 9.0, 10 mM; KCl, 50 mM; $MgCl_2$, 1.5 mM; Triton X-100, 0.1%; BSA, 0.2 mg/ml).

## Exonuclease assays

The enzyme was tested for exonuclease activity; 3′–5′ exonuclease assays quantified the release of $^{32}P$ from a λ DNA restricted with *Hin*dIII restriction enzyme (Appligene-Oncor) and end labeled using the Klenow enzyme with $^{32}P$-dATP and $^{32}P$-dTTP and unlabeled dCTP and dGTP. After 1 h at 37°C, the four unlabeled dNTPs were added in excess for 30 min. The DNA was purified, and Klenow and dNTP were removed by phenol extraction and ethanol precipitation. Exonuclease assays using enzyme buffer (Tris-HCl, pH 9.0, 10 mM; KCl, 50 mM; $MgCl_2$, 1.5 mM; Triton X-100, 0.1%; BSA, 0.2 mg/ml) and containing 0.02 mg/ml of the labeled DNA were incubated for 16 h at 72°C with 20 U of *Tfu* DNA polymerase in 100 µl final volume. The same assays were performed with *Tli* DNA polymerase in its own buffer as positive tests. Then, 10 µl of this assay was dropped on DE81 paper (Whatman), dried, and counted before and after washing with a 0.5 M $Na_2HPO_4$ solution to remove released nucleotides. The 3′–5′ exonuclease activity and thermostability were also tested for sensitivity to dNTP concentrations. The 5′–3′ exonuclease activity was also checked using polynucleotide kinase label of the restricted DNA on its 5′-end.

## Endonuclease and nickase assay

Endonuclease and nickase activity at 37°C and 70°C were tested for each active fraction. Supercoiled pBR322 DNA (1 µg) was incubated with 1–5 µl of the active fraction in different buffers (activity buffer or *Taq* buffer; Appligene-Oncor), in a final volume of 50 µl for 1–16 h. The relaxation of the supercoiled DNA was visualized on agarose gels.

## DNA polymerase thermostability

*Tfu* DNA polymerase was incubated up to 18 h at three different temperatures (92°C, 95°C, 100°C) in a 20-µl volume of the incubation buffer used for activity determination. The enzyme, diluted to 20 U/ml, was incubated under various temperatures and times under mineral oil, and then stored at −20°C. Polymerization activity was then measured as described.

## 3′–5′ Exonuclease thermostability

Experiments were performed as described for polymerization activity related to heat exposure (92°C, 95°C, 100°C) up to 4 h, by incubating 20 U of *Tfu* DNA polymerase in a 20-µl volume of PCR buffer 1×, under mineral oil. The retained 3′-exonuclease activity was measured as described.

## PCR assays

PCR assays to generate DNA fragments of 3.8 kb, 7.0 kb, or 10 kb were carried out using 10 ng of λ DNA as a substrate and 0.5 U of *Tfu* DNA polymerase, in the optimized PCR buffer; 40 µM of each dNTP was used, except for the 10-kb amplification, where 200 µM was needed. Addition of 10 mM $(NH_4)_2SO_4$ was also tested. Despite the high activity of its 3′–5′ exonuclease, no modified primers were required for an efficient PCR amplification. Primers were used at 50 pmol per PCR test. PCR programs were as follow: 3.8 kb (30 s at 94°C, 2 min at 65°C, 3 min at 72°C, ×20), 7.0 kb (20 s at 94°C, 2 min at 65°C, 6 min at 72°C, ×20), and 10 kb (10 s at 94°C, 2 min at 65°C, 10 min at 72°C, ×20). PCR products were electrophoresed on a 1% standard agarose gel. The molecular weight ladder used was Raoul (Appligene-Oncor).

## Fidelity assays

Error rates generated by *Tfu* DNA polymerase were measured according to the method of Flaman et al. (Flaman et al. 1994) except that the reporter gene HIS3 was replaced by the ADE2 gene. Yeast cells directly transformed with unpurified PCR products from p53 cDNA (Flaman et al. 1994) were cloned into a linearized expression vector by homologous recombination in vivo. A 1182-bp fragment containing the complete wild-type p53 open reading frame was amplified by using *Tfu* DNA polymerase, as well as *Taq* DNA polymerase. Mutations were scored according to the ability of p53 product to activate transcription of an ADE2 reporter gene. The error rate per nucleotide and per cycle (E) was calculated using the equation $E = mut/(d \times del \times 100)$, where mut is the percentage of mutated ade2 (−) measured, d is the amount of effective amplified fragment doublings, and del is the effective target size (542 bp) of the complete p53. d is expressed by the equation $d = \log X/\log_2$, X being the amount of PCR product versus the amount of effective input template.

PCR assays were carried out to amplify 20 pg of input DNA, with 3 U of DNA polymerase in its optimal incubation buffer, in a 50-µl volume. Different parameters were considered, such as the $Mg^{2+}$ concentration (1.5–3 mM), the dNTP concentration (40–300 µM each), the presence of $(NH_4)_2SO_4$ (10 mM) having an inhibitor effect on 3′–5′ exonuclease activity, and the presence of T4 bacteriophage gene 32 product (T4 gp32) (5–30 µg/ml), known to improve DNA polymerase proofreading (Sandhu and Keohavong 1994).

## Results

### Isolation of the *dna pol* gene

Among *Thermococcales* strains isolated from IFREMER's collection, *Pyrococcus* sp. GE23 (unpublished data) is a close relative of *P. abyssi*. Its DNA polymerase gene, previously isolated and sequenced (EMBL accession number Z54173), allowed us to use appropriate 5′- and 3′-fragments as probes. Two bands of *Thermococcus fumicolans* digested genomic DNA, immobilized on membranes, hybridized to the probe GE23 "BEG.": a 1.9-kb *Hind*III–*Hind*III fragment and a 5-kb *Xho*I–*Xho*I fragment. All other restriction endonucleases yielded hybridizing fragments larger than 12 kb. The 1.9-kb *Hind*III fragment (TfuMACa) was excised from a gel and cloned. Multiple alignment with already available DNA polymerase sequences demonstrated that TfuMACa (Fig. 1) encodes the 446-aa N-terminal region of a new DNA-dependent DNA polymerase with a putative 3′–5′ exonuclease. The polymerase sequence is split by an intervening sequence located in the SLYPSII conserved motif (also called motif A), similarly to *P. kodakaraensis* KOD1 (GenBank D29671) and *T. aggregans* TY (Genbank Y13030). However, this putative intein coding sequence was incomplete in TfuMACa. No positive hybridization signal could be obtained with colonies derived from *E. coli* transformants which contained the 5-kb *Xho*I–*Xho*I fragments, suggesting that we could be facing the same kind of cloning problems previously encountered with the *T. litoralis* DNA polymerase gene, or more precisely, with toxic inteins (Perler et al. 1992).

Membranes were hybridized again with probe GE23 "END." and revealed a positive 1.1-kb *Hind*III–*Hind*III fragment. This fragment was cloned into pUC18. Sequencing of one of these positive clones demonstrated (Fig. 1) that it encoded the C-terminal region of the DNA polymerase (TfuMACc), but it did not overlap with TfuMACa. It contains an intein closely related to Tli pol-2 and inserted in the YG.TDS motif (motif C). To clone the central part of the *Tfu* DNA polymerase gene, membranes were screened using the TFU "MID." probe. This probe hybridized to a 2-kb *Hind*III–*Hind*III fragment, which was also cloned in pUC18 and sequenced (TfuMACb). Fragments a, b, and c yielded a single ORF containing two inteins and three exteins, which comprise the polymerase (see Fig. 1). Amplification by PCR performed with oligonucleotides located at both ends of the DNA polymerase gene gave a unique fragment of 4572 bp, and its sequencing demonstrated that no sequence was missing across the *Hind*III sites. This fragment was unstable in pUC18 and in pET12a, but proved to be stable in the pARHS2 vector.

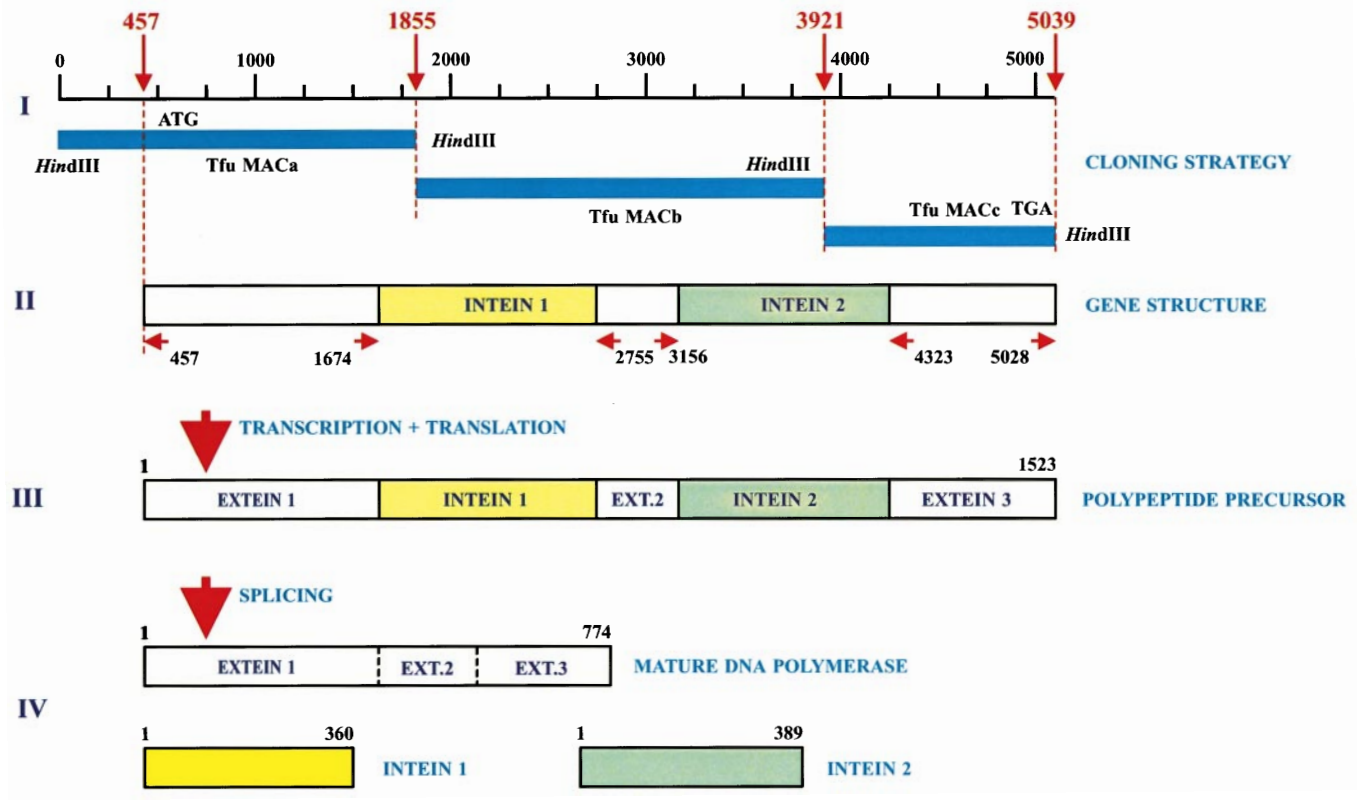


**Fig. 1.** Cloning strategy and gene structure of the DNA polymerase of *Thermococcus fumicolans*. Three fragments of genomic DNA digested with *Hind*III were cloned and sequenced independently. Those fragments encode a single open reading frame coding for one precursor. Protein splicing released the active DNA polymerase and the two inteins

## Sequence analysis

The three sequences, TfuMACa, -b, and -c, were assembled and encode the *Tfu* DNA-dependent DNA polymerase. This DNA polymerase belongs to class B (Ito and Braithwaite 1991). It displays characteristic motifs described by Blanco et al. (Blanco and Salas 1996). In addition, the motif $\mathbf{Y_{384}AGG}$ reported by Pisani (Pisani et al. 1998) to be involved in determining the processivity of the proofreading function and to interact directly with the nucleic acid substrate is highly conserved in *T. fumicolans*.

The gene is split by two intervening sequences. One is closely related to Psp KOD1 pol-1 intein and located at the same integration site (pol-a) corresponding to motif A (DFR*SLYP). The similarity between these two inteins or alleles at amino acids level is 75.3%. The second intein is related to Tli pol-2 and located in integration site pol-c corresponding to motif C (YAD*TDG), as in the *T. litoralis* DNA polymerase gene. Similarity between these two alleles is 62.2%. To our knowledge, this is the first time this kind of DNA polymerase gene organization has been found. Considering the criteria listed by Perler (Perler et al. 1997) for intein identification, these two intervening sequences encode putative inteins with homing endonuclease domains. First, they represent large in-frame elements. Second, they also display the same conserved sequences necessary for protein splicing (Xu and Perler 1996). Third, like inteins found in other *Thermococcales* DNA polymerase genes, they share the same dodecapeptide conserved motif (LAGLIDADG) repeated at about 100 amino acid residues apart.

The *Tfu* DNA polymerase sequence (after removal of inteins) is closely related to the *Thermococcus sp.* GE 8 DNA polymerase obtained in our laboratory (unpublished data), a new deep-vent hydrothermal species isolated by the CNRS/*Roscoff* (Marteinsson et al. 1995), sharing 91.9% of similarity at the amino acid level. The *Tfu* DNA polymerase is also 90.6% similar to the DNA polymerase of *Thermococcus* sp. 9°N (Southworth et al. 1996), but only 77.3% similar to *Tli* DNA polymerase, which appears to be rather distant from the others even though they share common inteins. A comparison of the two intein sequences of *T. fumicolans* demonstrates that the divergence level between them is greater than between inteins of the same allele class (i.e., integrated at the same site) from different isolates. This finding is in agreement with the results obtained by Perler and coworkers in compiling and analyzing intein sequences (Perler et al. 1997) and illustrates the independent phylogenetic relationship between inteins in the same gene at different locations.

## Expression

No information was available concerning the possible toxicity of Tfu pol-1 or Tfu pol-2 inteins, which could affect the stability of any construction including the intein coding regions because of possible endonuclease activity in *E. coli*. The trials to clone the whole *Tfu* DNA polymerase gene with both inteins, as described in Material and methods, gave negative results with pET12a but were successful with pARHRS2. One clone was selected and proved stable in the NovaBlue strain at 37°C. Direct expression using the BL21(DE3)pLysS strain was assayed. Cells cultured overnight at 37°C in LB were lysed whereas at 30°C lysis was limited. Expression of the proteins (the two inteins and the DNA polymerase) was improved with 2xYT medium supplemented with ampicillin and chloramphenicol at 30°C. Maximum expression was obtained using a clone that expressed the proteins after addition with 1 mM IPTG followed by 4 h of induction.

## *Tfu* recombinant DNA polymerase purification

Following cell lysis and centrifugation, purification of supernatant was assessed by electrophoresis on SDS-PAGE and fractions revealed by Coomassie blue staining. Protein profiles (Fig. 2) of partially heated fractions obtained from cells cultured at 30°C display several bands, a major one being located around 170 kDa (lane 2), which is approximately the size of the protein precursor containing the DNA polymerase with both inteins, and another of 89 kDa, possibly corresponding to mature DNA polymerase. Subsequent effective heating of 15-ml samples at 70°C for 10 min of previous fractions (partially heated) induced processing of most of the 170-kDa polypeptide and released

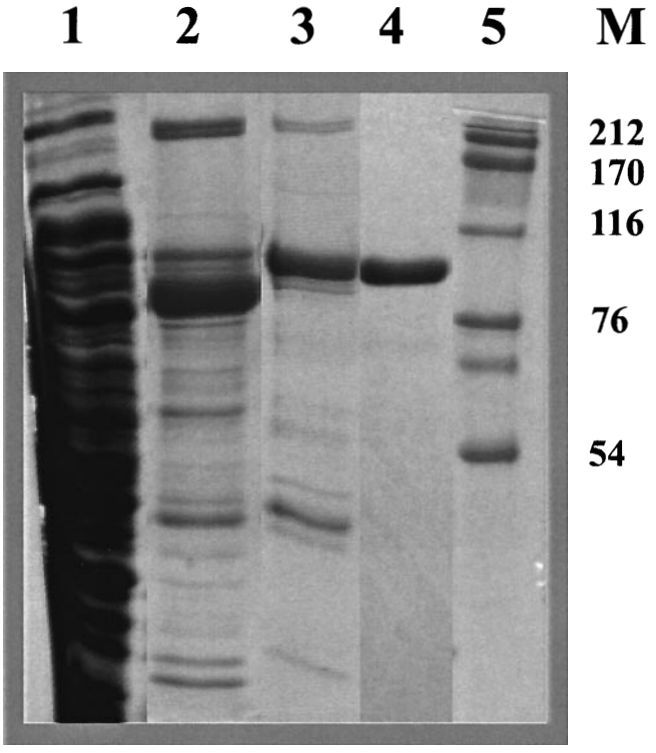


**Fig. 2.** *Tfu* recombinant DNA polymerase purification. *Lane 1*, crude extract without heating; *lane 2*, pooled fractions after partial heating and purification on Heparin column; *lane 3*, pooled fractions after additional heating at 70°C; *lane 4*, pooled fractions after purification on HiTrap column; *lane 5*, protein marker. *M* indicates the sizes in kDa

the 89-kDa DNA polymerase and inteins (lane 3). This result demonstrated that splicing did not significantly occur in *E. coli* during expression cultures at 30°C, which could explain the lack of toxic effects, if any, in cultures. In fact, splicing is thermosensitive and probably occurs in vitro above 40°C, during the first step of purification. Further purification (as described in Materials and methods) yielded a pure and active DNA polymerase with a relative molecular mass of 89 kDa (lane 4).

## Polymerase assay

*Tfu* DNA polymerase proved to be very sensitive to buffer composition, compared to the *Tli* DNA polymerase. Using the *Tli* DNA polymerase buffer or the *Taq* polymerase buffer (Appligene-Oncor), the *Tli* DNA polymerase activity was unchanged; this was not the case for *Tfu* DNA polymerase activity, which varied from 4 U/µl to about 60 U/µl, respectively, in the two buffers. In fact, KCl concentration in the buffer was proved to be important because an increase in KCl concentration from 10 to 75 mM results in a tenfold increase in polymerase activity (Fig. 3).

## Thermostability

Purified *Tfu* DNA polymerase thermostability was tested at three different temperatures: 92°C, 95°C, and 100°C. The *Tfu* DNA polymerase is more thermostable than *Tli* DNA polymerase (Kong et al. 1993) and much more thermostable than *Taq* polymerase (Lawyer et al. 1993), as the *Tfu* DNA polymerase retained 50% of its activity after 7 h incubation at 92°C, 3.3 h at 95°C, and 2 h at 100°C.

## Temperature optimum

Polymerization by *Tfu* DNA polymerase has been tested between 68.5° and 74°C on activated calf thymus DNA with an optimum activity at 72°C. This optimum probably reflects more a thermal denaturation of the substrate

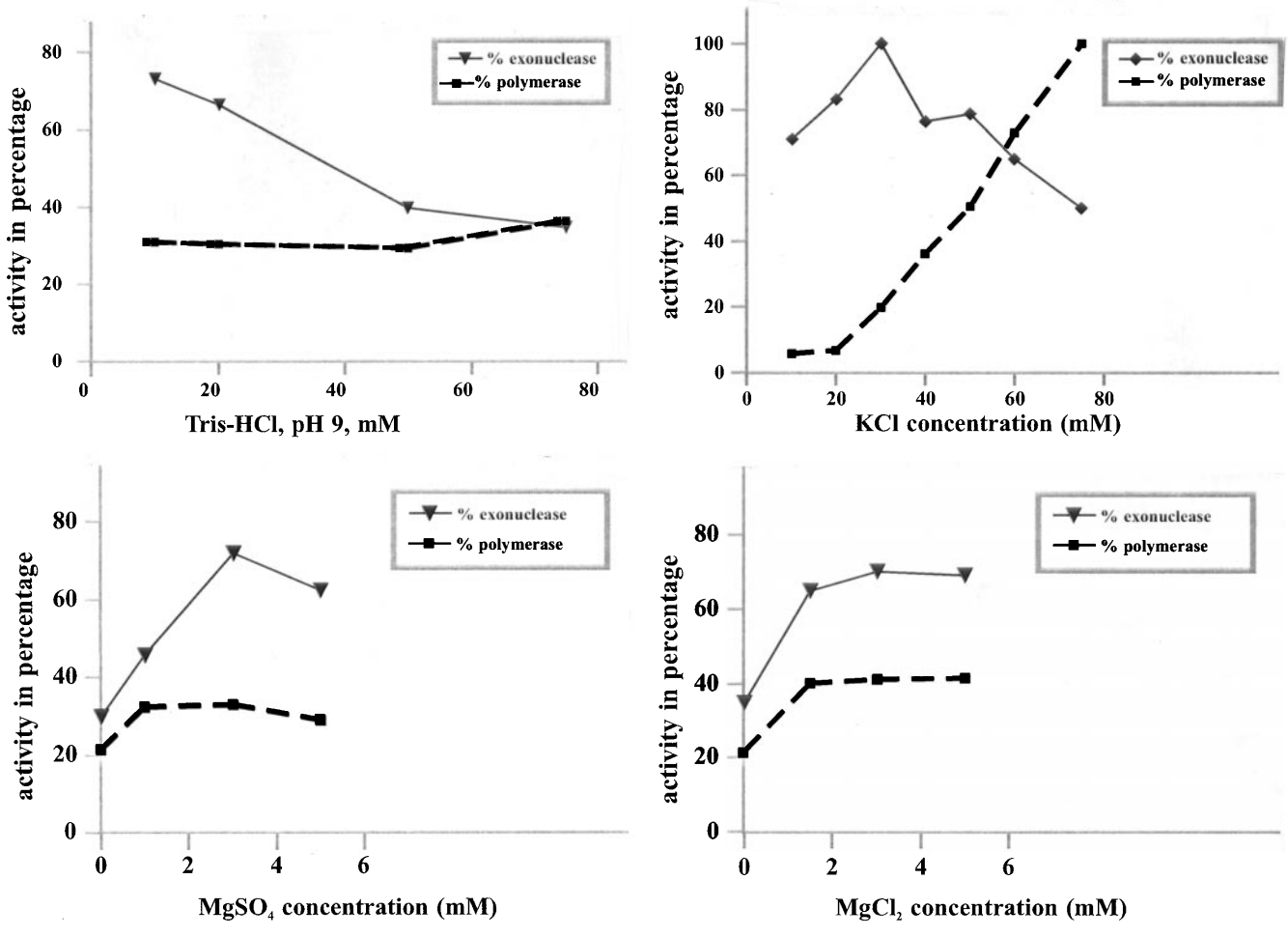


**Fig. 3.** DNA polymerase and 3′–5′-exonuclease assays using the *Tli* DNA polymerase buffer. Only one component was successively subjected to variation: Tris-HCl, KCl, $MgSO_4$, or $MgCl_2$; 100% of polymerase activity is the value obtained with the *Tli* DNA polymerase buffer in which the KCl concentration was increased from 10 mM to 75 mM. In the same way, 100% of exonuclease activity is the value obtained with the same buffer, but at 30 mM KCl concentration. Both activities need a higher concentration in KCl than *Tli* DNA polymerase

(activated calf thymus DNA) rather than a decrease in its activity, because this enzyme is till active in PCR experiments at 80°C in PCR (data not shown).

## Exonuclease activities

Incubation of a large amount of *Tfu* DNA polymerase with λ DNA cut with *Hin*dIII led to the degradation of this DNA (data not shown), indicating the presence of exonuclease activity. This activity was checked by assaying the $^{32}$P-labeled product released from an end-labeled DNA substrate. When the substrate was labeled on its 3′-end, about 75% of the $^{32}$P was released in 16 h. When the substrate was labeled on its 5′-end, no $^{32}$P was released. This result demonstrates that *Tfu* DNA polymerase possesses a 3′–5′ exonuclease activity, but no 5′–3′ exonuclease activity, which is consistent with its amino acid sequence. Unexpectedly, the exonuclease activity proved to be less sensitive to dNTP concentration in comparison with *Tli* DNA polymerase; 80% of *Tfu* exonuclease activity is retained with 0.75 mM dNTP concentration. Exonuclease activity of *Tli* DNA polymerase displays almost a linear decrease, but residual activity is 10% at 0.75 mM dNTP concentration. However, at the usual dNTP concentrations used for PCR and required to keep high-fidelity amplification, the responses are similar.

## Exonuclease thermostability

The 3′–5′ exonuclease activity of *Tfu* DNA polymerase was very thermoresistant as it was only slightly affected by 3 h of incubation at 92°C. The activity was slowly reduced after 80 min at 95°C, but decreased rapidly at 100°C, the 3′–5′ exonuclease being totally inhibited after incubation of 25 min at 100°C.

## PCR assays

DNA amplification in vitro using DNA polymerases from *Thermococcales* (*Tli* DNA pol, *Pfu* DNA pol, etc.) is already a common practice, and the properties of *Tfu* DNA pol in this field were analyzed. Using standard PCR protocols, fragments providing the expected size bands were obtained, up to 10 kb. Addition of $(NH_4)_2SO_4$ (Fig. 4, lane C) increases the signal intensity, but with a moderate loss in fidelity (Table 1). Nevertheless, *Tfu* DNA polymerase requires more strictly controlled thermal cycling conditions than DNA polymerases devoid of 3′–5′ exonuclease activity. For an optimum PCR product yield, an extension of 1 min/kb of expected amplified fragment is necessary if this fragment size is more than 2 kb. Under 2 kb, this rule cannot be respected. A minimum of 45 s of extension time is required to amplify a 280-bp fragment, and 75 s for a 400-bp fragment.

## Fidelity assays

In comparative standard conditions (polymerase buffer at 1.5 mM $MgCl_2$), at 200 μM of each dNTP (see Table 1), *Tfu*

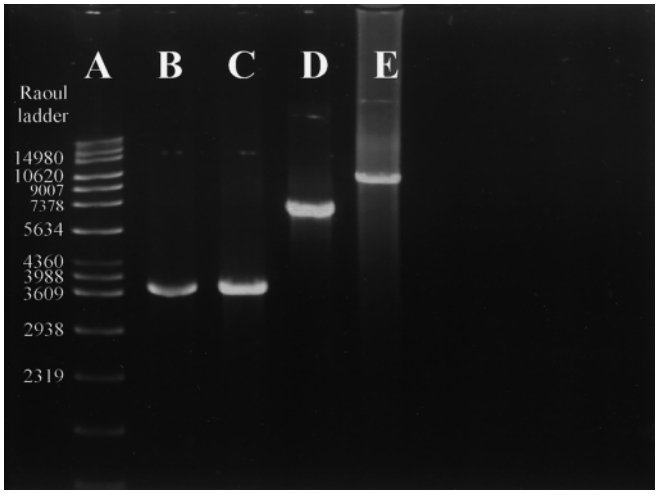


**Fig. 4.** Amplification of DNA fragments of various lengths by PCR using *Tfu* DNA polymerase. *Lane A*, Raoul ladder; *lanes B*, *C*, 3.8-kb amplifications; *lane D*, 7-kb amplification; *lane E*, 10-kb amplification. PCR of 3.8 and 7 kb was carried out with 40 μM of each dNTP; 200 μM of each dNTP was needed for 10-kb amplification; 0.5 U of *Tfu* DNA polymerase was used and 10 ng lambda DNA as template. *Lane C*, addition of 10 mM of $(NH_4)_2SO_4$ increases the signal but with a loss in fidelity. For each PCR, unmodified primers were used

**Table 1.** Tfu DNA polymerase fidelity under several conditions

| DNA polymerase | Incubation conditions | Input DNA (pg) | $Mg^{2+}$ concentration (mM) | Each dNTP molarity (μM) | Additives | Error rate (/bp/duplication) |
|---|---|---|---|---|---|---|
| *Tfu* | *Tfu* pol mix | 20 | 1.5 | 40 | | $0.9 \times 10^{-5}$ |
| | *Tfu* pol mix | 20 | 1.5 | 200 | | $1.6 \times 10^{-5}$ |
| | *Tfu* pol mix | 20 | 1.5 | 300 | | $1.9 \times 10^{-5}$ |
| | *Tfu* pol mix | 20 | 2 | 200 | | $2.0 \times 10^{-5}$ |
| | *Tfu* pol mix | 20 | 2.5 | 200 | | $2.55 \times 10^{-5}$ |
| | *Tfu* pol mix | 20 | 3 | 200 | | $5.3 \times 10^{-5}$ |
| | *Tfu* pol mix | 20 | 1.5 | 40 | $(NH_4)_2SO_4$, 10 mM | $0.95 \times 10^{-5}$ |
| | *Tfu* pol mix | 20 | 1.5 | 200 | $(NH_4)_2SO_4$, 10 mM | $1.85 \times 10^{-5}$ |
| | *Tfu* pol mix | 20 | 1.5 | 40 | gp32, 5 μg/ml | $0.45 \times 10^{-5}$ |
| | *Tfu* pol mix | 20 | 1.5 | 200 | gp32, 5 μg/ml | $1.57 \times 10^{-5}$ |
| | *Tfu* pol mix | 20 | 1.5 | 40 | gp32, 30 μg/ml | $1.05 \times 10^{-5}$ |
| *Taq* | *Taq* pol mix | 20 | 1.5 | 200 | gp32, 30 μg/ml | $2.4 \times 10^{-5}$ |

DNA polymerase exhibits a better fidelity ($1.6 \times 10^{-5}$ errors/bp/duplication) than *Taq* DNA polymerase ($2.4 \times 10^{-5}$ errors/bp/duplication). At a constant concentration of $MgCl_2$ (1.5 mM), the error rate of *Tfu* DNA polymerase decreases ($0.9 \times 10^{-5}$) by reducing the dNTP concentration to 40 μM each, a phenomenon previously observed (Cline et al. 1996) with *Pfu* DNA polymerase, whereas 300 μM of each dNTP increases the error rate ($1.9 \times 10^{-5}$). On another hand, an increase of $MgCl_2$ concentration drastically alters *Tfu* DNA polymerase fidelity, as the error rate measured (at 200 μM of each dNTP) varied from $1.6 \times 10^{-5}$ to $5.3 \times 10^{-5}$ whereas the $MgCl_2$ concentration varied from 1.5 to 3 mM. The loss of fidelity is more drastic when the $MgCl_2$ concentration increases from 2.5 to 3 mM (error rates from 2.55 to $5.3 \times 10^{-5}$). Surprisingly, no specific inhibition of the 3′–5′ exonuclease activity was observed when varying $MgCl_2$ concentration (data not shown). These results completely differ from those obtained with *Pfu* DNA polymerase (Cline et al. 1996) or with vent DNA polymerase (Ling et al. 1991).

The addition of 10 mM $(NH_4)_2SO_4$, in spite of its inhibitory effect on 3′–5′ exonuclease, barely affects the *Tfu* DNA polymerase error rate ($0.95 \times 10^{-5}$ compared to $0.9 \times 10^{-5}$ in absence of salt, using 40 μM of each dNTP). As described previously, T4 gp32 improves fidelity (Sandhu and Keohavong 1994). *Tfu* DNA polymerase fidelity is two fold better at gp32 concentration of 5 μg/ml ($0.45 \times 10^{-5}$) in optimum conditions (40 μM each dNTP, 1.5 mM $MgCl_2$). However gp32 has no effect on fidelity at 200 μM dNTP. At the concentration close to that used for T4 DNA polymerase (30 μg/ml) (Sandhu and Keohavong 1994), fidelity is unchanged ($1.05 \times 10^{-5}$ versus $0.90\ 10^{-5}$ without T4 gp32).

## Discussion

*Thermococcus fumicolans* is unique because its optimum growth temperature (90°C) makes it the most hyperthermophilic of all *Thermococcus* described, close to the temperature ranges of species belonging to the genus *Pyrococcus*. At the DNA polymerase level, *T. fumicolans* displays two interesting features. First, the intein insertion pattern in its DNA polymerase gene (motifs A and C) illustrates the fact that intein genes can be distributed in many possible combinations in family B DNA polymerase genes of *Thermococcales*. *T. fumicolans* displays the pattern A + C; *T. litoralis*, B + C; *T. aggregans* TY, A + B + C; *P. kodakaraensis* KOD1, A + B; *Pyrococcus* sp. strain GB-D, B; *P. furiosus* and *Thermococcus* sp. 9°N, none. Second, expression of the DNA polymerase gene including both inteins was successful in *E. coli* at 30°C and produced an unspliced and stable polypeptide precursor that could undergo splicing after heat exposure.

As was shown, similarities between intein-less *dna pol.* genes of *Thermococcales* are higher than similarities between intein coding sequences. Two hypotheses can be formulated about this result. First, inteins can accumulate more mutations than polymerases and still retain functions necessary for cell survival. Second, inteins displaying homing endonuclease motifs are more divergent because they evolved separately from DNA polymerases. This idea supposes the existence of horizontal transfer, already hypothesized in different species (Liu and Hu 1997; Perler et al. 1997; Pietrokovski 1998).

When cloning the *Tli* DNA polymerase, Perler et al. (Perler et al. 1992) noticed that when they tried to clone the entire gene in plasmids, recombinants grew poorly, plasmids rearranged, and no active polymerase could be detected. Inteins of *P. kodakaraensis* KOD1 (Nishioka et al. 1998) as well as those of *T. litoralis* recognize rare DNA sites (and specifically a unique site in intein-less host species), and we had no a priori insurance that inteins of *T. fumicolans* were not able to cut the genome of expression strains of *E. coli*[1]. With these previous results in mind, it would have been advisable to delete intein coding sequences at the gene level and to ligate in vitro the extein coding sections before expression experiments, as reported previously (Niehaus et al. 1997; Perler et al. 1992; Takagi et al. 1997). The toxicity of some inteins expressed in *E. coli* can result from the endonuclease action cutting the chromosome at the few sites recognized by the enzyme, possibly different in *E. coli* compared to the Archaea host strain due to specific cytoplasm characteristics, impairing the host genome and killing the cells.

Based on the hypothesis that the probability is high that *Tfu* inteins have no recognition sites in the whole *E. coli* genome, it was decided to clone the entire DNA polymerase gene including inteins. Despite negative results with the pET12a expression vector, suggesting a possible toxicity of the construction or its products in *E. coli*, attempts in pARHS2 vector at 30°C directly allowed expression of both inteins and the DNA polymerase after thermally induced splicing. Constructions used for these two types of vectors were similar (insertion between *Nde*I/*Bam*HI sites without signal sequence), and it appears that the vector used for cloning and expressing the recombinant polymerase in *E. coli* is a very serious limiting factor. This positive result could be due to the *parB* gene inserted in the pARHS2 vector (De Moerlooze et al. 1992). To our knowledge, this is the first report about a clear failure of splicing at 30°C of a thermostable intein in its native context in *E. coli* and subsequent heat-induced splicing of the precursor. Davis et al. (1992) reported that the wild-type *rec*A allele of *M. tuberculosis* produced a full-sized precursor 85-kDa protein as well as mature recA protein and intein in suboptimal conditions for splicing. However, attempts to show processing of the precursor of the wild-type *rec*A locus were unsuccessful.

Previous studies of protein splicing of purified precursor were done on intein embedded in foreign proteins (MBP–Psp GB-D pol-1–paramyosin Δ Sal), and allowed identification of a branched intermediate, modulation of splicing

[1] Characterization of the endonuclease properties of *Tfu* DNA polymerase inteins have been done recently by Masson and coworkers (Saves et al., 2000. J Biol Chem, 275(4):2335–2341)

by temperature and pH, and splicing mechanisms (Xu and Perler 1996; Xu et al. 1993, 1994). More recently, Klabunde et al. (1998) reported also the obtaining of a stable precursor by mutation (GFP–Mxe gyrA–GFP), but the N-terminal cysteine changed to alanine, preventing the first step of splicing and inactivating the precursor. The availability of a precursor in its native context might open the way to further investigation on splicing mechanisms and their structural basis, as up to now no stable native polypeptide precursor was available and studies on three-dimensional structure were performed on spliced inteins (Duan et al. 1997; Klabunde et al. 1998).

Considering its biochemical characteristics, exonuclease and polymerase domains can act independently, as shown previously (Kong et al. 1993; Perler et al. 1996; Pisani et al. 1998), and this independence can be seen in the temperature profiles of exonuclease and polymerase activities as well as in their relative thermostabilities. Thermostability of the *Tfu* polymerase and the 3′–5′ exonuclease domains were analyzed using the same protocol of exposure to high temperatures. However, besides the temperature, each experiment requires specific conditions and the proper ionic environment. Differences observed in retained activities of polymerase and exonuclease after exposure to high temperatures could reflect effective differences, a property observed previously for archaeal DNA polymerases (Hamal et al. 1990; Kong et al. 1993) or possibly to a minor extent the impact of different experimental conditions. The recently disclosed crystal structure of DNA polymerase of *T. gorgonarius* allowed Hopfner et al. (1999) to discuss adaptation to high temperatures. However, the different abilities of 3′–5′ exonuclease and polymerase domains to sustain high temperatures and remain active are still unexplained.

*Tfu* DNA polymerase proved to be an easy enzyme to use in PCR amplification and is available on the market (*Tfu* thermostable polymerase; Appligene-Oncor). It is very sensitive to buffer composition and has an optimal buffer containing 10 mM Tris-HCl, pH 9.0; 50 mM KCl; 1.5 mM $MgCl_2$; 0.1% Triton X100; and 0.2 mg/ml BSA. KCl concentration clearly enhances DNA polymerase activity but reduces the 3′–5′ exonuclease activity. This 3′–5′ exonuclease activity is much less sensitive to dNTP concentration than *Tli* DNA polymerase, which makes this enzyme suitable for PCR applications when high dNTP concentrations are required. Its sensitivity to buffer conditions allows the modulation of fidelity from $0.9 \times 10^{-5}$ to $5.3 \times 10^{-5}$ by changing the $MgCl_2$ and dNTP concentrations. Decrease of fidelity when $MgCl_2$ and dNTP concentrations increase is observed, although there is no specific inhibition of 3′–5′ exonuclease activity. This result might suggest that this phenomenon is primarily a consequence of a direct alteration of intrinsic fidelity of the polymerase domain, as already reported by Eckert (Eckert and Kunkel 1990). In optimum conditions, observed fidelity makes this new DNA polymerase attractive for applications requiring a strict limitation of errors in PCR. *Tfu* DNA polymerase is also convenient to use in PCR up to 10 kb.

**Acknowledgments** This work was supported by the Biotech Extremophiles program of U.E., DGXII, and by the Regional Council of Brittany, France. We thank Prof. P. Forterre and Dr. F.B. Perler for stimulating discussions and for her helpful advice and continuous support.

## References

Baross JA, Holden JF (1996) Overview of hyperthermophiles and their heat-shock proteins. In: Adams MWW (ed) Enzymes and proteins from hyperthermophilic microorganisms, vol 48. Academic Press, San Diego, pp 1–34

Birnboim HC, Doly J (1979) A rapid extraction procedure for screening recombinant plasmid DNA. Nucleic Acids Res 7:1503

Blanco L, Salas M (1996) Relating structure to function in phi 29 DNA polymerase. J Biol Chem 271:8509–8512

Bult CJ, White O, Olsen GJ, Zhou L, Fleischmann RD, Sutton GG, Blake JA, FitzGerald LM, Clayton RA, Gocayne JD, Kerlavage AR, Dougherty BA, Tomb JF, Adams MD, Reich CI, Overbeek R, Kirkness EF, Weinstock KG, Merrick JM, Glodek A, Scott JL, Geoghagen NSM, Weidman JF, Fuhrmann JL, Nguyen D, Mtterback TR, Kelley JM, Peter son JD, Sadow PW, Hanna MC, Cotton MD, Roberts KM, Hurst MA, Kame BP, Borodovsky M, Klenk HP, Fraser CM, Smith HO, Woese CR, Venter JC (1996) Complete genome sequence of the methanogenic archaeon, *Methanococcus jannaschii*. Science 273:1058–1072

Canganella F, Jones WJ, Gambacorta A, Antranikian G (1998) *Thermococcus guaymasensis* sp. nov. and *Thermococcus aggregans* sp. nov., two novel thermophilic archaea isolated from the Guaymas Basin hydrothermal vent site. Int J Syst Bacteriol 48:1181–1185

Cann IKO, Komori K, Toh H, Kanai S, Ishino Y (1998) A heterodimeric DNA polymerase: evidence that members of Euryarchaeota possess a distinct DNA polymerase. Proc Natl Acad Sci USA 95:14250–14255

Cline J, Braman JC, Hogrefe HH (1996) PCR fidelity of *Pfu* DNA polymerase and other thermostable DNA polymerases. Nucleic Acids Res 24:3546–3551

Davis EO, Jenner PJ, Brooks PC, Colston MJ, Sedgwick SG (1992) Protein splicing in the maturation of *M. tuberculosis* RecA protein: a mechanism for tolerating a novel class of intervening sequence. Cell 71:201–210

De Moerlooze L, Struman I, Renard A, Martial JA (1992) Stabilization of T7-promoter-based pARHS expression vectors using the *par*B locus. Gene (Amst) 119:91–93

Duan X, Gimble FS, Quiocho FA (1997) Crystal structure of PI-SceI, a homing endonuclease with protein splicing activity. Cell 89:555–564

Eckert KA, Kunkel TA (1990) High fidelity DNA synthesis by the *Thermus aquaticus* DNA polymerase. Nucleic Acids Res 18:3739–3744

Flaman J-M, Frebourg T, Moreau V, Charbonnier F, Martin C, Ishioka C, Friend SH, Iggo R (1994) A rapid PCR fidelity assay. Nucleic Acids Res 22:3259–3260

Godfroy A, Meunier JR, Guézennec J, Lesongeur F, Raguénès G, Rimbault A, Barbier G (1996) *Thermococcus fumicolans* sp. nov., a new hyperthermophilic archaeon isolated from a deep-sea hydrothermal vent in the North Fiji Basin. Int J Syst Bacteriol 46:1113–1119

Gorbalenya AE (1998) Non-canonical inteins. Nucleic Acids Res 26:1741–1748

Grunstein M, Hogness D (1975) Colony hybridization: a method for the isolation of cloned DNAs that contain a specific gene. Proc Natl Acad Sci USA 72:3961–3965

Hamal A, Forterre P, Elie C (1990) Purification and characterization of a DNA polymerase from the archaebacterium *Thermoplasma acidophilum*. Eur J Biochem 101:1–5

Hopfner KP, Eichinger A, Engh RA, Laue F, Ankenbauer W, Huber R, Angerer B (1999) Crystal structure of a thermostable type B DNA polymerase from *Thermococcus gorgonarius*. Proc Natl Acad Sci USA 96:3600–3605

Ignatov KB, Miroshnikov AI, Kramarov VM (1998) Substitution of Asn for Ser543 in the large fragment of *Taq* DNA polymerase

increases the efficiency of synthesis of long DNA molecules. FEBS Lett 425:249–250

Ito J, Braithwaite DK (1991) Compilation and alignment of DNA polymerase sequences. Nucleic Acids Res 19:4045–4057

Kawarabayasi Y, Sawada M, Horikawa H, Haikawa Y, Hino Y, Yamamoto S, Sekine M, Baba S, Kosugi H, Hosoyama A, Nagai Y, Sakai M, Ogura K, Otsuka R, Nakazawa H, Takamiya M, Ohfuku Y, Funahashi T, Tanaka T, Kudoh Y, Yamazaki J, Kushida N, Oguchi A, Aoki K, Kikuchi H (1998) Complete sequence and gene organization of the genome of a hyperthermophilic archaebacterium, *Pyrococcus horikoshii* OT3. DNA Res 5:55–76

Kebelmann-Betzing C, Seeger K, Dragon S, Schmitt G, Moricke A, Schild TA, Henze G, Beyermann B (1998) Advantages of a new *Taq* DNA polymerase in multiplex PCR and time-release PCR. BioTechniques 24:154–158

Klabunde T, Sharma S, Telenti A, Jacobs WR, Sacchettini JC (1998) Crystal structure of GyrA intein from *Mycobacterium xenopi* reveals structural basis of protein splicing. Nat Struct Biol 5:31–36

Kong HM, Kucera RB, Jack WE (1993) Characterization of a DNA polymerase from the hyperthermophile Archaea *Thermococcus litoralis* – vent DNA polymerase, steady state kinetics, thermal stability, processivity, strand displacement, and exonuclease activities. J Biol Chem 268:1965–1975

Lawyer FC, Stoffel S, Saiki RK, Chang SY, Landre PA, Abramson RD, Gelfand DH (1993) High-level expression, purification, and enzymatic characterization of full-length *Thermus aquaticus* DNA polymerase and a truncated form deficient in 5′ to 3′ exonuclease activity. PCR Methods Appl 2:275–287

Ling LL, Keohavong P, Dias C, Thilly WG (1991) Optimization of the polymerase chain reaction with regard to fidelity: modified T7, Taq, and vent DNA polymerases. PCR Methods Appl 1:63–69

Liu X-O, Hu Z (1997) A DnaB intein in *Rhodothermus marinus*: indication of recent intein homing accross remotely related organisms. Proc Natl Acad Sci USA 94:7851–7856

Marteinsson VT, Watrin L, Prieur D, Caprais JC, Raguénès G, Erauso G (1995) Phenotypic characterization, DNA similarities, and protein profiles of twenty sulfur-metabolizing hyperthermophilic anaerobic archaea isolated from hydrothermal vents in the southwestern Pacific Ocean. Int J Syst Bacteriol 45:623–632

Meunier JR (1994) Biodiversité et systématique des populations de microorganismes thermophiles isolés d'écosystèmes hydrothermaux océaniques abyssaux. Paris 6, Paris, pp 316–317

Mullis K, Faloona F, Saiki R, Horn G, Erlich H (1986) Specific enzymatic amplification of DNA *in vitro*: the polymerase chain reaction. Cold Spring Harbor Symp Quant Biol 51:263–273

Niehaus F, Frey B, Antranikian G (1997) Cloning and characterisation of a thermostable alpha-DNA polymerase from the hyperthermophilic archaeon *Thermococcus* sp. TY. Gene (Amst) 204:153–158

Nishioka M, Fujiwara S, Takagi M, Imanaka T (1998) Characterization of two intein homing endonucleases encoded in the DNA polymerase gene of *Pyrococcus kodakaraensis* strain KOD1. Nucleic Acids Res 26:4409–4412

Park Y, Choi H, Lee DS, Kim Y (1997) Improvement of the 3′–5′ exonuclease activity of *Taq* DNA polymerase by protein engineering in the active site. Mol Cell 7:419–424

Perler FB (1999) InBase, the New England Biolabs intein database. Nucleic Acids Res 27:346–347

Perler F, Kumar S, Kong H (1996) Thermostable DNA polymerases. In: Adams MWW (ed) Enzymes and proteins from hyperthermophilic microorganisms, vol 48. Academic Press, San Diego, pp 377–435

Perler FB, Comb DG, Jack WE, Moran LS, Qiang B, Kucera RB, Benner J, Slatko BE, Nwankwo DO, Hempstead SK, Carlow CKS, Jannasch H (1992) Intervening sequences in Archaea DNA polymerase gene. Proc Natl Acad Sci USA 89:5577–5581

Perler FB, Olsen GJ, Adam E (1997) Compilation and analysis of intein sequences. Nucleic Acids Res 25:1087–1093

Pietrokovski S (1998) Identification of a virus intein and a possible variation in the protein splicing reaction. Curr Biol 8:R634–R635

Pisani FM, DeFelice M, Rossi M (1998) Amino acid residues involved in determining the processivity of the 3′–5′ exonulcease activity in a family B DNA polymerase from the thermoacidophilic archaeon *Sulfolobus solfataricus*. Biochemistry 37:15005–15012

Sambrook J, Fritsch EF, Maniatis T (1989) Molecular cloning. A laboratory manual. Cold Spring Harbor, New York

Sandhu DK, Keohavong P (1994) Effects of the T4 bacteriophage gene 32 product on the efficiency and fidelity of DNA amplification using T4 DNA polymerase. Gene (Amst) 144:53–58

Sanger F, Nicklen S, Coulson AR (1977) DNA sequencing with chain terminating inhibitors. Proc Natl Acad Sci USA 74:5467–5473

Southern EM (1975) Detection of specific sequences among DNA fragments separated by gel electrophoresis. J Mol Biol 98: 503

Southworth MW, Kong H, Kucera RB, Ware J, Jannasch HW, Perler FB (1996) Cloning of thermostable DNA polymerases from hyperthermophilic marine Archaea with emphasis on *Thermococcus sp*. 9°N-7 and mutations affecting 3′–5′ exonuclease activity. Proc Natl Acad Sci USA 93:5281–5285

Takagi M, Nishioka M, Kakihara H, Kitabayashi M, Inoue H, Kawakami B, Oka M, Imanaka T (1997) Characterization of DNA polymerase from *Pyrococcus* sp. Strain KOD1 ant its application to PCR. Appl Environ Microbiol 63:4504–4510

Thompson JD, Higgins DG, Gibson TJ (1994) Clustal W: improving the sensitivity of progressive multiple sequence alignment through sequence weighting, position-specific gap penalties and weight matrix choice. Nucleic Acids Res 22:4673–4680

Woese CR, Fox GE (1977) Phylogenetic structure of the prokaryotique domain: the primary kingdoms. Proc Natl Acad Sci USA 74:5088–5090

Woese CR, Kandler O, Wheelis M (1990) Towards a natural system of organisms: proposal for the domains Archaea, Bacteria, and Eucarya. Proc Natl Acad Sci USA 87:4576–4579

Xu M-Q, Perler F (1996) The mechanism of protein splicing and its modulation by mutation. EMBO J 15:5146–5153

Xu MQ, Comb DG, Paulus H, Noren CJ, Shao Y, Perler FB (1994) Protein splicing: an analysis of the branched intermediate and its resolution by succinimide formation. EMBO J 13:5517–5522

Xu M-Q, Southworth MW, Mersha FB, Hornstra LJ, Perler FB (1993) *In vitro* protein splicing of purified precursor and identification of a branched intermediate. Cell 75:1371–1377