

## Inteins of *Thermococcus fumicolans* DNA Polymerase Are Endonucleases with Distinct Enzymatic Behaviors\*

(Received for publication, June 17, 1999, and in revised form, November 3, 1999)

Isabelle Saves†, Valérie Ozanne§¶, Jacques Dietrich||\*\*, and Jean-Michel Masson‡ §§ §§

From the ‡Institut de Pharmacologie et Biologie Structurale, IPBS/CNRS, 205 Route de Narbonne, F-31077 Toulouse Cedex, §Appligène-Oncor, Parc d'innovation, BP 72, F-67402 Illkirch Cedex, ||Ifremer, Laboratoire de Biotechnologie, Centre de Brest, BP 70, F-29280 Plouzané, and §§Institut National des Sciences Appliquées, Complexe Scientifique de Rangueil, F-31077 Toulouse Cedex, France

The DNA polymerase gene of *Thermococcus fumicolans* harbors two intein genes. Both inteins have been produced in *Escherichia coli* and purified either as naturally spliced products from the expression of the complete DNA polymerase gene or directly from the cloned inteins genes. Both recombinant inteins exhibit endonuclease activity, with an optimal temperature of 70 °C. The *Tfu* pol-1 intein, which belongs to the *Psp* KOD pol-1 allelic family, recognizes and cleaves a minimal sequence of 16 base pairs (bp) on supercoiled DNA with either Mn<sup>2+</sup> or Mg<sup>2+</sup> as cofactor. It cleaves linear DNA only with Mn<sup>2+</sup> and requires a 19-bp minimal recognition sequence. The *Tfu* pol-2 intein, which belongs to the *Tli* pol-2 allelic family, is a highly active homing endonuclease using Mg<sup>2+</sup> as cofactor. Its minimal recognition and cleavage site is 21 bp long either on linear or circular DNA substrates. Its endonuclease activity is strongly inhibited by the 3' digestion product, which remains bound to the enzyme after the cleavage reaction. According to current nomenclature, these endonucleases were named PI-*Tfu*I and PI-*Tfu*II. These two inteins thus exhibit different requirements for metal cofactor and substrate topology as well as different mechanism of action.

Since the first report of a protein splicing element as an in-frame insertion in the *VMA* gene of *Saccharomyces cerevisiae* (1, 2), over 80 putative inteins have been identified.<sup>1</sup> Inteins are widely distributed in a variety of eucarya, eubacteria, and archaea genes and tend to invade highly conserved and functionally important regions of their host genes. More than 30 different host proteins have been reported, including DNA and RNA polymerases, helicases, gyrases, RecA recombinase, ribonucleotide reductase, and metabolic enzymes.

Protein splicing is a rapid autocatalytic process occurring at the peptide level. The intein coding sequence is transcribed and translated in the proper reading frame with flanking regions of the extein to produce a large precursor peptide. The intein is then autocatalytically excised while the N and C extein fragments are ligated, yielding a native and functional

host protein. The molecular mechanism of splicing has been thoroughly investigated (3–7). The excised intein is also a stable protein. Like some group I intron open reading frames, some intein genes have been found to encode double-strand DNA endonucleases, which cleave the inteinless allele of the host gene. Moreover, Gimble and Thorner (8) demonstrated that the *VMA1*-derived endonuclease (VDE or PI-*Sce*I) of *Saccharomyces cerevisiae* introduces a staggered double-strand break, repair of which results in the invasion of the intein coding sequence in the *VMA1* gene *in vivo*. Thus, even if this homing event still remains to be demonstrated in eubacteria and archaea, it is clear that the endonuclease activity potentially confers highly specific mobility to its own coding sequence (9).

The majority of known inteins contain two conserved sequence motifs characteristic of the LAGLIDADG endonuclease family, called the dodecapeptide motifs, in their central region. In addition, two other motifs found in homing endonucleases and four motifs involved in splicing are generally observed in these inteins (10–13). While sequence alignment analyses tend to show that inteins containing dodecapeptide motifs probably exhibit endonuclease activity, only six inteins, *Sce* VMA, *Tli* pol-1, *Tli* pol-2, *Psp* GBD pol-1, *Psp* KOD pol-1, and *Psp* KOD pol-2, have been shown to be specific endonucleases and named PI-*Sce* I, PI-*Tli* II, PI-*Tli* I, PI-*Psp* I, PI-*Pko* I, and PI-*Pko* II, respectively, according to the current nomenclature (8, 12, 14, 15).

The coding sequence of the *Tfu* DNA polymerase gene from *Thermococcus fumicolans* (GenBank accession no. Z69882)<sup>2</sup> exhibits two in-frame insertions of 1080 and 1167 bp<sup>3</sup> at positions 1218 and 2700, respectively (Fig. 1a), corresponding to expected inteins of 41.4 kDa (*Tfu* pol-1) and 44.8 kDa (*Tfu* pol-2). While both *Tfu* polymerase inteins display the conserved motifs determined by sequence alignment of homing endonucleases, there is a marked divergence between their sequences (11). The *Tfu* pol-1 intein belongs to the allelic family of *Psp* KOD pol-1, while *Tfu* pol-2 intein is a *Tli* pol-2 allele. Thus, to compare the splicing and endonuclease activity of both *Tfu* DNA pol inteins, their coding sequences and the entire gene of the DNA polymerase were cloned in expression vectors. After the production and the purification of recombinant and spliced inteins, their putative endonuclease activities were investigated and characterized.

### EXPERIMENTAL PROCEDURES

**Production and Purification of the *Tfu* Polymerase Inteins**—The coding sequences of *Tfu* pol-1 and *Tfu* pol-2 inteins were amplified by PCR

\* The costs of publication of this article were defrayed in part by the payment of page charges. This article must therefore be hereby marked "advertisement" in accordance with 18 U.S.C. Section 1734 solely to indicate this fact.

† Present address: Valigène, Tour Neptune, 92086 Paris La Défense, France.

\*\* Supported by the Conseil Régional de Bretagne.

§§ Recipient of an incitative Ifremer grant. To whom correspondence should be addressed. Tel.: 33-5-61-17-54-76; E-mail: masson@ipbs.fr.

<sup>1</sup> This information is available via the World Wide Web (S. Pietrokowski's intein site).

<sup>2</sup> M. Cambon and J. Quérellou, unpublished data.

<sup>3</sup> The abbreviations used are: bp, base pair(s); pol, polymerase; PCR, polymerase chain reaction.

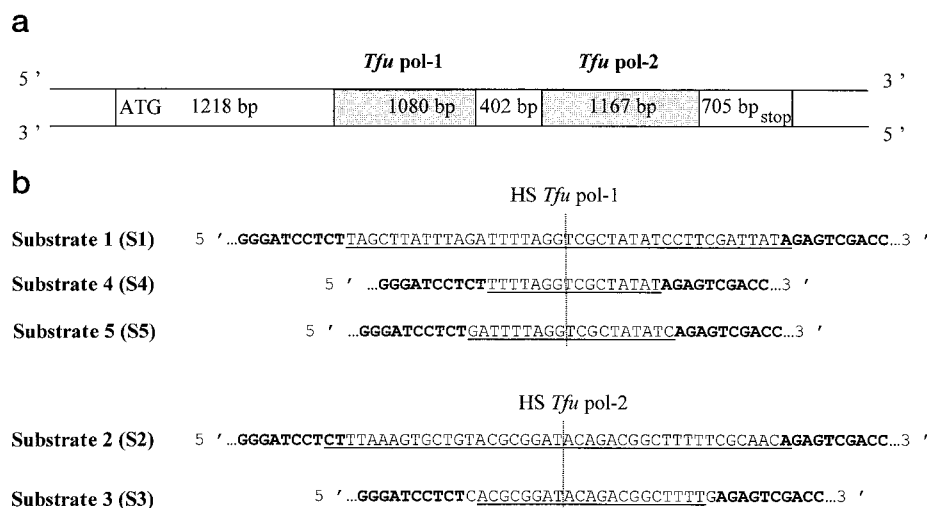


FIG. 1. *a*, structure of the DNA polymerase gene of *T. fumicolans* as described by Cambon and Quérellou (GenBank accession no. Z69882). Length in bp of the coding sequences of inteins and exons is indicated. Intein genes are shaded. *b*, DNA substrates for PI-*Tfu*I and PI-*Tfu*II. The sequences of PI-*Tfu*I and PI-*Tfu*II homing sites (HS) inserted in pUC19 (sequence of which appears in bold) are underlined.

from *Tfu* genomic DNA with oligonucleotides *Tfu* pol-1-ATG (5'-ttgatcatcatgtgcatctgcgcacact-3'), *Tfu* pol-1-TAG (5'-gatgtcgactaatgtgggtcagtatgcc-3'), *Tfu* pol-2-ATG (5'-ctgtacgcatatgagtggtacaggggacacag-3'), and *Tfu* pol-2-TAG (5'-agcgtcgacatgtgtggaacgagattccg-3'), respectively. These PCR products were a gift from J. Quérellou (Ifremer). They were digested by *Nde*I and *Sal*I and cloned into a *Nde*I-*Sal*I-digested pET26b vector (Novagen). The resulting constructs (pET26-*Tfu*1 and pET26-*Tfu*2) were sequenced by Isoprism (Toulouse, France). *Escherichia coli* BL21(DE3)[pLysS] bacteria transformed with pET26-*Tfu*1 or pET26-*Tfu*2 were grown at 37 °C in Luria broth culture medium supplemented with 50 µg/ml kanamycin (Sigma). A 2-h induction with 0.5 mM isopropyl-β-D-thiogalactopyranoside (Sigma) was performed, and cells were lysed in 20 mM sodium phosphate, pH 7.5 (measured at 25 °C), by six cycles of freezing-thawing. The lysate was centrifuged at 10,000 × *g* for 15 min, and the supernatant was chromatographed on a XK26 Sepharose Q Fast Flow column (Amersham Pharmacia Biotech) at room temperature. The proteins were eluted with a 0–1 M NaCl linear gradient in 20 mM sodium phosphate, pH 7.5, at a flow rate of 5 ml/min. *Tfu* pol-1 and *Tfu* pol-2 are eluted at 0.2 and 0.7 M NaCl, respectively, and are approximately 80% homogeneous. These fractions were dialyzed against 10 mM Tris-HCl, pH 7.5, 50% glycerol, 0.1 mM EDTA, 1 mM dithiothreitol, 200 µg/ml bovine serum albumin, and 50 mM NaCl for storage.

The *Tfu* pol gene was cloned in an expression vector by Appligène-Oncor (France). BL21(DE3)[pLysS] bacteria transformed with this plasmid were grown overnight at 25 °C. Extraction and purification procedures for spliced inteins were identical to those described for recombinant inteins, except that the protein extract was treated for 1 h at 70 °C before purification.

**Assay of Endonuclease Activity**—All the substrates were constructed by PCR as described by Weiner (16) using pUC19 as DNA matrix and different purified oligonucleotide pairs: S1–5' (5'-tcgtatatctctcgattatagagtcgacacgtcaggcatgc-3') and S1–3' (5'-cctaaatctaaataagagatcccccgggtaccgag-3'), S2–5' (5'-acagacggtcttttcgcaacagagtcgacacgtcaggcatgc-3') and S2–3' (5'-atccgcgtacagcactttaagaggatccccgggtaccgag-3'), S3–5' (5'-agagtcgacacgtcaggcatgc-3') and S3–3' (5'-caaaagccgtctgtatccggtgagaggatccccgggtaccgag-3'), S4–5' (5'-agagtcgacacgtcaggcatgc-3') and S4–3' (5'-atatagcagcactaaaagaggatccccgggtaccgag-3'), and S5–5' (5'-tcgtatatcatagagtcgacacgtcaggcatgc-3') and S5–3' (5'-cctaaatcagaggatccccgggtaccgag-3'), for substrates 1–5, respectively (Fig. 1b). The sequences in bold are necessary for the annealing to the pUC19 matrix at 60 °C.

The 41-bp DNA substrate 1 (S1) for *Tfu* pol-1 consisted of 20 bp of the 5' end of the homing site plus 21 bp of the 3' end, inserted in the *Xba*I site of pUC19. In the same way, a 43-bp substrate for *Tfu* pol-2 was constructed (substrate 2, S2). Substrates 3 and 4 (S3 and S4) corresponded to the minimal recognition and cleavage sites determined by the primer extension method (see below) for PI-*Tfu*II and PI-*Tfu*I, respectively. Substrate 5 (S5) corresponded to the minimal linear recognition site of PI-*Tfu*I.

The resulting plasmids were either linearized by *Sca*I (Appligène-Oncor) and purified or the supercoiled forms of the DNA substrates were purified from a 1% agarose gel in TBE (0.09 M Tris borate, 0.002 M EDTA) buffer. Finally, linear and circular forms of these substrates

were diluted in water to a concentration of 100 ng/µl. Endonuclease activity assays were performed in a final volume of 10 µl, in various reaction buffers and temperatures ranging from 37 °C to 80 °C (see "Results"). The reaction mixtures were analyzed on a 1% agarose gel in TBE buffer. The amount of undigested substrates and products were quantified with the ImageQUANT program (Molecular Dynamics Inc.).

One unit of PI-*Tfu*I or PI-*Tfu*II endonuclease was required to digest 1 µg of *Sca*I-linearized DNA substrate 1 or 2, respectively, in 1 h at 70 °C, in optimal buffers. Specific activities of PI-*Tfu*I and PI-*Tfu*II were measured by incubating known amounts of linear DNA substrates with known amounts of purified endonucleases.

**Definition of the Minimal Recognition and Cleavage Site of Endonucleases**—The endonuclease recognition sites were determined by a primer extension method as described by Wenzlau *et al.* (17). NaOH-denatured plasmids S1 and S2 were sequenced in both orientations using the T7 polymerase sequencing kit (Amersham Pharmacia Biotech) and universal primers SeqPuc (5'-gtaacgccagggtttcc-3') and M13Rev (5'-ggaaacagctatgacatg-3'). After chain termination reaction, the samples were split in two. The first half of the sample was overdigested with 100 units of PI-*Tfu*I or PI-*Tfu*II for 1 h at 70 °C, in a 50 mM Tris acetate, pH 8, buffer containing 100 mM NH<sub>4</sub>OAc and 25 mM MnSO<sub>4</sub> or 75 mM Mg(OAc)<sub>2</sub>, respectively. The cleavage reactions and the undigested half of the sequence reactions were ethanol-precipitated and resuspended in 5 µl of water. 4 µl of stop solution were added, and the samples were subjected to electrophoresis in a standard 6% denaturing polyacrylamide gel.

**Inhibition of Cleavage by Reaction Products**—4 µg of *Sca*I-linearized substrate 2 were overdigested by PI-*Tfu*II in optimal conditions by successive additions of enzyme. Both cleavage reaction products (940 and 1790 bp) were separated and purified from each other from a 1% agarose gel in TBE buffer. These two DNA fragments were then used as competitors in standard cleavage assays of substrate 2.

**Electrophoretic Mobility Shift Assay**—A 119-bp DNA probe containing the 43-bp recognition and cleavage site of PI-*Tfu*II was synthesized by PCR. The universal M13Rev oligonucleotide was first phosphorylated with [γ-<sup>32</sup>P]ATP (NEN Life Science Products) (18) and then used together with the pUC5'-Xba oligonucleotide (5'-agctcggtaccggggat-3') and 26 ng of substrate 2 DNA as matrix in a standard PCR amplification reaction. The 119-bp <sup>32</sup>P-labeled fragment was purified and diluted in water to a concentration of 25 nM. 7.5 fmol of this DNA probe were incubated in a 20 mM Tris acetate, pH 8, buffer containing 20 mM Mg(OAc)<sub>2</sub>, 20 mM NH<sub>4</sub>OAc, 0.1% Triton X-100, and 16% glycerol, either with or without PI-*Tfu*II, in a final volume of 20 µl, for 15 min at 37 °C or 70 °C. These reaction mixtures were then submitted to electrophoresis in a 7.5% nondenaturing polyacrylamide (30:1) gel in 0.25× TBE buffer at 220 V for 2 h. After electrophoresis, the gel was dried and radioactive bands were detected using a PhosphorImager (Molecular Dynamics Inc.). The specificity of the DNA-protein interaction was controlled using 7.5 pmol of herring sperm DNA as a nonspecific competitor in the binding reactions (data not shown).

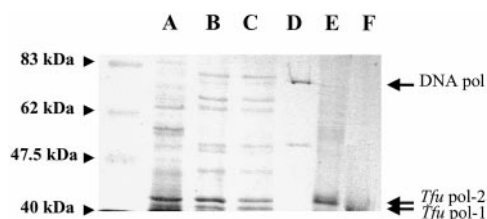


FIG. 2. **Expression and splicing of *Tfu* DNA polymerase in *E. coli*.** 10% SDS-polyacrylamide gel electrophoresis of crude extract of *Tfu* DNA pol unheated (lane A) or treated at 70 °C for 30 min (lane B) and 1 h (lane C), of purified *Taq* DNA pol (lane D) and recombinant *Tfu* pol-1 and *Tfu* pol-2 inteins (lanes E and F, respectively).

## RESULTS

***Tfu* Polymerase Inteins Were Spliced in *E. coli***—The complete *Tfu* polymerase gene was expressed in *E. coli*, at 25 °C. The SDS-polyacrylamide gel electrophoresis separation of the crude extract shows two bands of approximately 41 and 45 kDa, which comigrated with *Tfu* pol inteins (Fig. 2). At this temperature, a partial splicing reaction was observed since the DNA polymerase was not detected after staining the gel with Coomassie Blue, and we consistently obtained more *Tfu* pol-2 intein than *Tfu* pol-1.

Incubating this bacterial extract at 70 °C increased the yield of spliced proteins, the DNA polymerase being detected after heating for 30 min (Fig. 2). Even then, although no protein bands corresponding to the polymerase precursors were detected on the gel, splicing was probably not complete after 1 h of incubation at 70 °C because of the different amounts of each of the spliced products, *i.e.* the polymerase and both inteins.

Purified *Tfu* DNA pol is active in standard PCR reactions (Appligène-Oncor). Moreover, purified spliced *Tfu* pol-2 has the same specific endonuclease activity as the recombinant protein. In contrast, we could never detect any endonuclease activity for the purified spliced *Tfu* pol-1, unlike the recombinant protein (see below). We used the recombinant proteins to investigate further the endonuclease activity of both inteins.

**Endonuclease Activity of *Tfu* Inteins**—Inteins with known endonuclease activity recognize and cut a DNA sequence spanning the junction site created when the intein coding region is deleted from the host gene (homing site). We thus constructed substrate 1 for *Tfu* pol-1 and substrate 2 for *Tfu* pol-2 by cloning synthetic homing sites of 43 and 41 bp, respectively (Fig. 1b). The supercoiled and *Sca*I-linearized forms of these DNA substrates were purified and used to assay the endonuclease activity of both inteins.

Cleavage assays were performed in a wide range of experimental conditions, including Tris-HCl or Tris acetate buffers, pH 7–9 (measured at 25 °C), containing 0–100 mM Mg(OAc)<sub>2</sub>, 0–400 mM NH<sub>4</sub>OAc, 0–30% glycerol, and 0–100 mM NaCl, KCl, ZnCl<sub>2</sub>, CaCl<sub>2</sub>, MnSO<sub>4</sub>, or MnCl<sub>2</sub>. The assay temperature was varied from 37 to 80 °C.

The supercoiled substrate 1 was cleaved by *Tfu* pol-1 intein under a large set of conditions, at an optimal temperature of 70 °C. Both Mg<sup>2+</sup> and Mn<sup>2+</sup> were able to stimulate the cleavage reaction (Fig. 3a), and optimal digestion was obtained in a 50 mM Tris acetate, pH 8, buffer containing 25 mM MnSO<sub>4</sub> and 100 mM NH<sub>4</sub>OAc. Open circular DNA accumulated during the reaction leading to linear DNA product, indicative of a two-step reaction. When MnSO<sub>4</sub> was replaced by 50 mM Mg(OAc)<sub>2</sub>, cleavage efficiency dropped approximately 5-fold and the open circular intermediate accumulated (Fig. 3a).

In contrast, *Tfu* pol-1 showed no endonuclease activity on

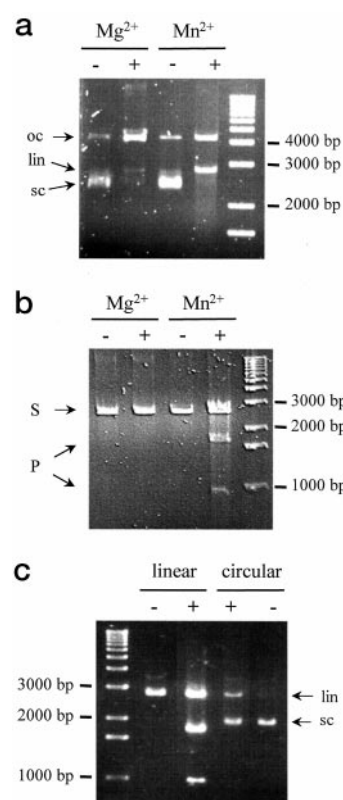


FIG. 3. **Standard cleavage assays for PI-*Tfu*I and PI-*Tfu*II.** a, cleavage assay for PI-*Tfu*I on circular substrate 1. 100 ng of purified supercoiled substrate 1 (Fig. 2) were incubated either with (+) or without (–) 40 ng of PI-*Tfu*I, for 10 min at 70 °C, in a 50 mM Tris acetate, pH 8, buffer containing 100 mM NH<sub>4</sub>OAc, and 50 mM Mg(OAc)<sub>2</sub> (Mg<sup>2+</sup>) or 25 mM MnSO<sub>4</sub> (Mn<sup>2+</sup>). Supercoiled (sc), open circular (oc), and linear (lin) forms of DNA were separated on a 1% agarose gel in TBE buffer. b, cleavage assay for PI-*Tfu*I on linear substrate 1. 100 ng of *Sca*I-linearized substrate 1 were incubated either with (+) or without (–) 40 ng of PI-*Tfu*I, for 10 min at 70 °C, in a 50 mM Tris acetate, pH 8, buffer containing 100 mM NH<sub>4</sub>OAc, and 50 mM Mg(OAc)<sub>2</sub> (Mg<sup>2+</sup>) or 25 mM MnSO<sub>4</sub> (Mn<sup>2+</sup>). Substrate (S, 2730 bp) and products (P, 940 and 1790 bp) were separated on a 1% agarose gel in TBE buffer. c, standard cleavage assay for PI-*Tfu*II. 100 ng of linear or circular substrate 2 (Fig. 2) were incubated with (+) or without (–) 0.1 ng of PI-*Tfu*II in a 50 mM Tris acetate, pH 8, buffer containing 100 mM NH<sub>4</sub>OAc, 75 mM Mg(OAc)<sub>2</sub>, and 10% glycerol, for 10 min at 70 °C. Circular (sc) and linear (lin) forms were separated on a 1% agarose gel in TBE buffer.

linear substrate 1 under all the conditions assayed, with up to 300 ng of protein/assay, except when manganese or zinc salts were added to the reaction buffer instead of magnesium salt (Fig. 3b). It turned out that MnSO<sub>4</sub> stimulated the reaction much more efficiently than MnCl<sub>2</sub> and ZnCl<sub>2</sub>. Optimal cleavage of linear substrate 1 was obtained, as with circular substrate 1, in a 50 mM Tris acetate, pH 8, buffer containing 25 mM MnSO<sub>4</sub> and 100 mM NH<sub>4</sub>OAc. Under these conditions, both linear and supercoiled substrates were cleaved with the same efficiency and specific activity of *Tfu* pol-1 (PI-*Tfu*I) was 4,800 units/mg.

In contrast, *Tfu* pol-2 exhibited endonuclease activity on substrate 2 under numerous assay conditions. No difference was observed between circular and linear forms of the DNA substrate, which were cleaved with the same optimal efficiency at 70 °C (Fig. 3c), in a 50 mM Tris acetate, pH 8, buffer containing 75 mM Mg(OAc)<sub>2</sub>, 100 mM NH<sub>4</sub>OAc, and 10% glycerol, with a specific activity of *Tfu* pol-2 (PI-*Tfu*II) of 748,000 units/mg. The cleavage efficiency was decreased 5–10-fold by replacing Mg<sup>2+</sup> by Mn<sup>2+</sup>. Otherwise, no cleavage was observed during the endonucleasic digestion of the supercoiled DNA (Fig.



TABLE I  
Cleavage properties of *PI-TfuI* and *PI-TfuII*

Relative activities of *PI-TfuI* and *PI-TfuII* on substrates 1 and 4, or 2 and 3, respectively. The cleavages of linear or circular forms of substrates are compared, and  $Mn^{2+}$  and  $Mg^{2+}$  are used as cofactor. +++ indicates that the cleavage is optimal; + indicate a 5–10-fold loss in cleavage efficiency; – indicates that no activity is detected.

	Linear DNA		Circular DNA	
	$Mn^{2+}$	$Mg^{2+}$	$Mn^{2+}$	$Mg^{2+}$
<i>PI-TfuI</i>				
Substrate 1 (43 bp)	+++	–	+++	+
Substrate 4 (16 bp)	–	–	+++	+
<i>PI-TfuII</i>				
Substrate 2 (41 bp)	+	+++	+	+++
Substrate 3 (21 bp)	+	+++	+	+++

3c). The effects of DNA topology and divalent ions used as cofactor on endonuclease activity of *PI-TfuI* and *PI-TfuII* are summarized in Table I.

**Minimal Recognition Sequences**—A mixture of potential double-strand DNA substrates for *PI-TfuI* and *PI-TfuII* were generated by primer extension using circular substrate 1 or substrate 2 as templates, respectively. The extension products including all the bases needed for cleavage were digested and thus disappeared from the sequence reaction, while those lacking one or more bases of the minimal recognition site were not digested. Hence, the comparison of digested and undigested sequence reactions allowed the determination of the bases necessary for cleavage (Fig. 4a). This method yielded a 16-bp non-palindromic site for *PI-TfuI* and a 21-bp non-palindromic site for *PI-TfuII*. In both cases, cleavage generated non-identical 3' overhangs of 4 bases (Fig. 4b). Both cleaved substrates could be religated with T4 DNA ligase. Thus, cleavage by *PI-TfuI* or *PI-TfuII* generated 5'-phosphate and 3'-hydroxyl ends. To confirm the results obtained by primer extension, plasmids containing the 21-bp site for *PI-TfuII* (substrate 3) or the 16-bp site for *PI-TfuI* (substrate 4, Fig. 1b) were constructed and submitted to cleavage assays either as linear or supercoiled DNA.

Standard cleavage assays on linear and circular DNA substrates 3 and 4, under optimal conditions for each enzyme, highlight the different catalytic behaviors of the endonuclease inteins (Table I). While *PI-TfuII* cleaved with the same efficiency linear and circular substrates 2 and 3 in the magnesium-containing buffer, *PI-TfuI* cleaved linear substrate 1 but did not cleave linear substrate 4 even in presence of manganese (Fig. 5a). On the other hand, *PI-TfuI* cleaved with the same efficiency supercoiled substrates 1 and 4 (Fig. 5b). Cleavage assays on various linear substrates for *PI-TfuI* (data not shown), containing synthetic homing sites of different lengths, led to a minimal linear site of 19 bp (Fig. 1b), which is homologous to the one described for *PI-PkoI* (15).

Thus, *PI-TfuII* cleaved its homing site and made no difference between linear and supercoiled DNA containing the minimal or a larger recognition site. In contrast, *PI-TfuI* cleaved supercoiled DNA harboring its 16-bp recognition sequence but required a longer recognition sequence ( $\geq 19$  bp) and manganese as cofactor to cleave linear DNA. Moreover, under these conditions, linear DNA was cleaved as efficiently as circular DNA (Table I).

**Product Inhibition of *PI-TfuII***—While incubation of linear substrate 1 with *PI-TfuI* led to its complete cleavage, prolonged incubation of substrate 2 with *PI-TfuII* resulted in only partial cleavage (Fig. 6a). The same results were obtained with circular DNA substrates (data not shown). After 10 min, 80% of the cleavage was performed and the reaction stopped. Addition of

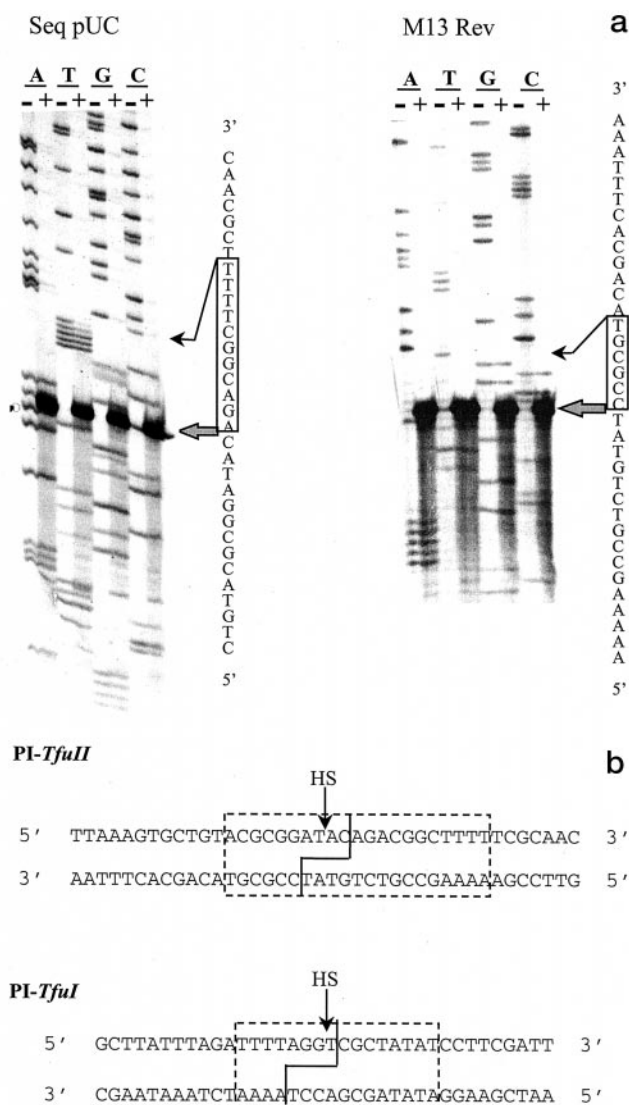


FIG. 4. Minimal recognition site of *PI-TfuI* and *PI-TfuII*, determined using the primer extension procedure. a, autoradiogram of the sequencing gel for *PI-TfuII* site determination. The sequencing reactions were performed in direct (SeqPuc) or reverse (M13Rev) orientations. *PI-TfuII*-digested (+ lanes) and undigested (– lanes) reactions were loaded side by side on a 6% denaturing polyacrylamide gel. A large arrow indicates the cleavage site on each DNA strand. Boxes represent bases belonging to the minimal site, in each direction. b, minimal nucleotide sequence necessary for recognition and cleavage by *PI-TfuI* and *PI-TfuII*. The arrows indicate the homing sites (HS) of the inteins in the DNA polymerase gene. The dashed boxes indicate the minimal recognition sequences.

fresh *PI-TfuII* allowed the reaction to proceed, but a loss of activity rapidly followed (data not shown). As the enzyme was stable at 70 °C, this loss of activity could result from an inhibition of *PI-TfuII* by one of its reaction products. To test this hypothesis, the two DNA fragments resulting from the cleavage of *ScaI*-linearized substrate 2 were purified and added as potential inhibitors in cleavage reactions. Under the conditions used, 50 ng of substrate 2 were digested by *PI-TfuII*, when no inhibitor was added to the reaction. Fig. 6b shows that the 940-bp fragment, which includes the homing site 5' end, inhibited weakly the cleavage reaction. On the other hand, the 1790-bp fragment, which includes the homing site 3' end, was an efficient inhibitor of the reaction, a 1:1 ratio between substrate and DNA fragment resulting in 80% inhibition. Electro-

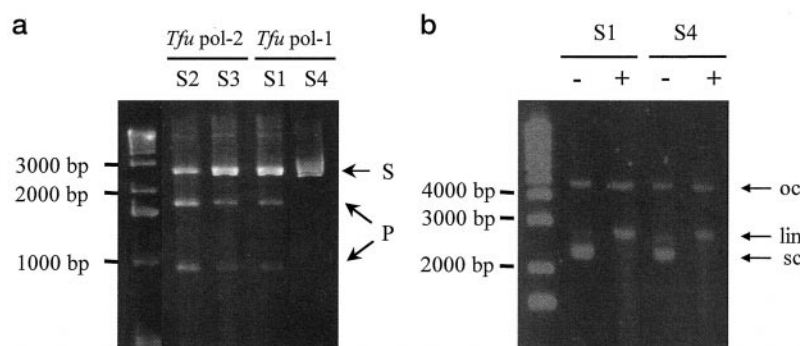


FIG. 5. Cleavage of minimal substrates of PI-*Tfu*I and PI-*Tfu*II. *a*, cleavage assay of linear substrates. 100 ng of *Sca*I-linearized substrates, S1 and S4 or S2 and S3, were incubated for 10 min at 70 °C, with 20 ng of PI-*Tfu*I or with 0.1 ng of PI-*Tfu*II, in a 50 mM Tris acetate, pH 8, buffer containing 100 mM NH<sub>4</sub>OAc, and 25 mM MnSO<sub>4</sub> or 75 mM Mg(OAc)<sub>2</sub>, respectively. Substrates (S, 2730 bp) and products (P, 940 and 1790 bp) were separated on a 1% agarose gel in TBE buffer. *b*, cleavage of supercoiled substrates by PI-*Tfu*I. 100 ng of purified supercoiled substrates 1 and 4 (Fig. 2) were incubated either with (+) or without (-) 200 ng of PI-*Tfu*I, for 10 min at 70 °C, in a 50 mM Tris acetate, pH 8, buffer containing 100 mM NH<sub>4</sub>OAc and 25 mM MnSO<sub>4</sub>. Circular (sc), open circular (oc), and linear (lin) forms of both substrates were separated on a 1% agarose gel in TBE buffer.

phoresis mobility shift assays confirmed the strong interaction between the 3' cleavage fragment and PI-*Tfu*II. Fig. 7 shows that this digestion product remained bound to the enzyme after cleavage was performed at 70 °C. Under these experimental conditions (large excess of enzyme over substrate), cleavage could also be observed at 37 °C but there was surprisingly no mobility shift for the resulting fragment.

#### DISCUSSION

Most DNA polymerase genes from archaeobacteria sequenced to date harbor one to three intein genes. These inteins are in-frame insertions at precisely the same locations, in the conserved motifs II, III, or I of the class B DNA polymerases (pol  $\alpha$ ). Thus, three allelic families of inteins have been defined, corresponding to these three different locations (Ref. 12; available via the World Wide Web from InBase, the New England Biolabs intein data base). Among the seven known intein-harboring polymerases, the *Tfu* polymerase is the only one that does not have a large intein within the conserved motif III. Its two inteins, *Tfu* pol-1 in motif II and *Tfu* pol-2 in motif I, belong to the *Psp* KOD pol-1 and *Tli* pol-2 allelic families, respectively. Both inteins exhibit endonuclease activity with an optimal temperature of 70 °C, which is in line with the thermophilic nature of *T. fumicolans*. Splicing of the inteins does occur in *E. coli*, but it is very inefficient. It is greatly enhanced by a prolonged incubation at 70 °C. Thus, unlike the splicing of PI-*Sce*I, splicing of the *Tfu* pol inteins seems quite temperature-dependent. One puzzling result was the fact that no endonuclease activity could be detected *in vitro* for the spliced PI-*Tfu*I, while the recombinant protein exhibits enzymatic activity. This suggests that, when produced in *E. coli*, the precursor protein fold is not fully correct, resulting in hindered splicing as well as inactive PI-*Tfu*I. Indeed, splicing of PI-*Tfu*II is much more efficient than splicing of PI-*Tfu*I and the spliced PI-*Tfu*II is fully active. When produced separately, the recombinant inteins and the DNA polymerase are active. High salt buffers are needed for optimal cleavage activity of both inteins. This feature could be linked to the fact that *T. fumicolans* is a marine organism, optimal growth of which is obtained with 30–40 g/liter salt (19). Indeed, the DNA polymerase activity is also stimulated by salts, while DNA polymerases from freshwater thermophilic bacteria, such as the *Taq* enzyme, do not need high salt concentrations.

*Tfu* pol-1 (PI-*Tfu*I) is 75.5% identical to *Psp* KOD pol-1, also known as homing endonuclease PI-*Pko*I (15). PI-*Tfu*I cleaves supercoiled DNA with magnesium or manganese as cofactor

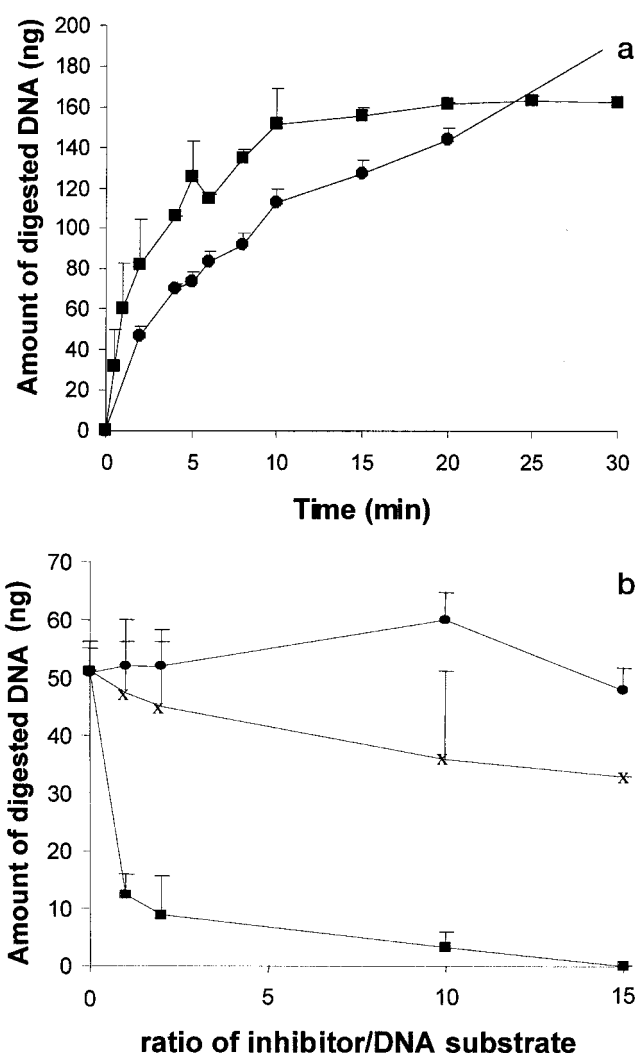


FIG. 6. Inhibition of PI-*Tfu*II. *a*, kinetics of cleavage by PI-*Tfu*I and PI-*Tfu*II. 200 ng of linearized substrates 1 and 2 (Fig. 2) were incubated with 80 ng of PI-*Tfu*I (●) or 0.3 ng of PI-*Tfu*II (■), respectively. The assays were performed at 70 °C, in their respective optimal reaction buffers and stopped at 4 °C after various incubation times. *b*, inhibition of PI-*Tfu*II activity by the reaction products. 100 ng of substrate 2 (Fig. 2) were incubated with 0.1 ng of PI-*Tfu*II for 10 min at 70 °C in optimal conditions. Various amounts of the 940-bp (×) and 1790-bp (■) cleavage products or poly(dI-dC) (●) were added to the reaction mixture to a molar ratio of inhibitor/substrate ranging from 1 to 15.

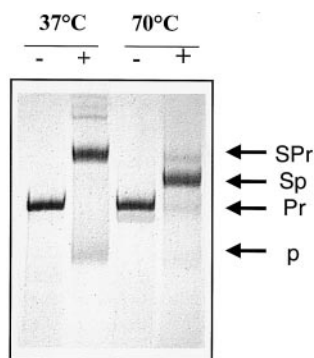


FIG. 7. **Electrophoretic mobility shift assay.** The DNA probe (7.5 fmol) was incubated for 15 min with (+) or without (-) 150 fmol of PI-TfuII at 37 °C, or with (+) or without (-) 75 fmol of PI-TfuII at 70 °C, in a 20 mM Tris acetate, pH 8, buffer containing 20 mM Mg(OAc)<sub>2</sub>, 20 mM NH<sub>4</sub>OAc, 0.1% Triton X-100, and 16% glycerol. The bands corresponding to the 119-bp DNA probe (Pr), the shifted probe (SPr), the 76-bp product (p), and the shifted product (SP) are indicated by arrows.

but requires manganese and a longer recognition site to cleave linear DNA. A difference in size requirement for minimal recognition sequence between linear and circular DNA has already been described for PI-SceI (20). The 16-bp minimal recognition sequence for PI-TfuI is 3 bp shorter than the one determined for PI-PkoI (15). In fact, on linear DNA, PI-TfuI and PI-PkoI both require a 19-bp sequence for binding and cleavage. The main difference between the two enzymes then seems to be metal requirements. Within the conserved sequence motifs that constitute the endonuclease signature, there are few differences between the two inteins PI-TfuI and PI-PkoI. The two acidic residues that are part of the catalytic triad (Glu-125 and Asp-225) are conserved, but the third residue, Lys-203, is replaced by an arginine in PI-TfuI. Such a substitution in PI-SceI (K301R) led to a 15-fold loss of activity and to a larger increase of activity when substituting magnesium with manganese (21). The equivalent substitution K92R totally inactivated *EcoRV* (22). On the other hand, an arginine is also found in the catalytic triad of PI-TliI, which requires Mg<sup>2+</sup> as cofactor (12), as does PI-PkoI. This change in residue therefore cannot account for the specific metal requirement of PI-TfuI.

Although manganese has been shown to stimulate the activity of magnesium requiring endonucleases, including PI-SceI (8, 21), it usually results in relaxed specificity, allowing cleavage of non-cognate sites. PI-TfuI is peculiar, for it has no detectable activity with magnesium on linear DNA while its activity with manganese as a cofactor is comparable to that of PI-PkoI. Thus, manganese relaxes the topological specificity of PI-TfuI, but not its requirement for cognate sequence, which is the same with either metal on supercoiled DNA. In this respect, it is reminiscent of the I91L mutant of *EcoRV*, where a single mutation within the active site of the enzyme results in a total shift in metal requirement while retaining specificity for the cognate sequence (23). The longer sequence needed for cleavage of linear DNA could translate the requirement of initial DNA bending for efficient substrate binding. On bent DNA, the enzyme takes a conformation that allows either magnesium or manganese to occupy the cofactor site and neutralize the phosphorus to trigger the cleavage of the phosphodiester bond. On linear DNA, the enzyme might take a slightly stretched conformation such that magnesium cannot occupy the cofactor site anymore while manganese, due to its 20% larger bonding radius, can still efficiently bind to its site on the protein and neutralize the phosphorus on DNA.

An open circular intermediate accumulates during the reac-

tion with supercoiled DNA, suggesting a two-step mechanism. A slower reaction rate with magnesium translates into more open circular intermediate. In a similar way, an open circular intermediate also accumulates during the magnesium-catalyzed cleavage by *EcoRV*, but disappears with manganese (either with the wild type *EcoRV* or with the I91L mutant) due to the enhanced cleavage rate.

Tfu pol-2 (PI-TfuII), which is 65% identical to Tli pol-2 (PI-TliI), is a very active endonuclease. It recognizes and cuts a 21-bp site on linear or circular DNA, leaving a 4-base 3'-OH overhang. This site is 15 bp shorter than the PI-SceI homing site (24) and slightly larger than the sites recently determined for the PI-PkoI, PI-PkoII (15), or PI-TfuI (this work) enzymes. It only differs by 3 bases from the 21-bp sequence central to the PI-TliI recognition site. A more thorough comparison of these two enzymes is needed to determine if they truly have different cleavage sites. Unlike PI-TfuI, PI-TfuII requires Mg<sup>2+</sup> as a cofactor and is less active with Mn<sup>2+</sup> for cleavage of its cognate sequence.

PI-TfuII is subject to efficient product inhibition by one of the cleavage fragments. Our gel shift experiments indicate that the *K<sub>i</sub>* for the product should be of the same order of magnitude as the *K<sub>m</sub>* for the substrate. We also observe that binding of the substrate at 37 °C is still efficient, while cleavage is almost abolished as well as product inhibition. This result is reminiscent of an inhibition mechanism involving the catalytic residues. Similar product inhibition has been reported for other endonucleases, including PI-SceI, although in this latter case inhibition appears more drastic (20). Thus, the DNA recognition and cleavage mechanism is probably homologous for both endonucleases, with a strong binding site 3' to the cleavage site.

*T. fomicolans* DNA polymerase gene harbors two intein genes like many other archaea DNA polymerase genes, but it is the only one that does not have a Tli pol-1 type intein allele. Its two intein genes code for active endonucleases that greatly differ in specific activity, cognate sequence, as well as metal requirement, cleavage site topology, and sensitivity to product inhibition. Such a situation has not been yet described and provides a contrasted model for the study of archaeobacterial endonuclease inteins. Site-directed mutagenesis of the active site of PI-TfuI will provide more information on the role of its residues and on its unusual topological and cofactor specificity.

**Acknowledgements**—We thank Joël Quérellou for providing the PCR products of intein genes and Heidy Eleaume and Emmanuelle Laurine for technical assistance.

#### REFERENCES

- Kane, P. M., Yamashiro, C. T., Wolczyk, D. F., Neff, N., Goebl, M., and Stevens, T. H. (1990) *Science* **250**, 651–657.
- Hirata, R., Ohsumi, Y., Nakano, A., Kawasaki, H., Suzuki, K., and Anraku, Y. (1990) *J. Biol. Chem.* **265**, 6726–6733.
- Xu, M.-Q., Comb, D. G., Paulus, H., Noren, C. J., Shao, Y., and Perler, F. B. (1994) *EMBO J.* **13**, 5517–5522.
- Chong, S., Shao, Y., Paulus, H., Benner, J., Perler, F. B., and Xu, M.-Q. (1996) *J. Biol. Chem.* **271**, 22159–22168.
- Xu, M.-Q., and Perler, F. B. (1996) *EMBO J.* **15**, 5146–5153.
- Shao, Y. S., and Kent, S. B. H. (1997) *Chemistry and Biology* **4**, 187–194.
- Petrokovski, S. (1998) *Current Biology* **8**, 634–635.
- Gimble, F. S., and Thorner, J. (1992) *Nature* **357**, 301–306.
- Belfort, M., Reaban, M. E., Coetzee, T., and Dalgaard, J. Z. (1995) *J. Bacteriol.* **177**, 3897–3903.
- Mueller, J. E., Bryk, M., Loizos, N., and Belfort, M. (1993) *Nucleic Acids* **111**–143.
- Dalgaard, J. Z., Klar, A. J., Moser, M. J., Holley, W. R., Chatterjee, A., and Saira Mian, I. (1997) *Nucleic Acids Res.* **25**, 4626–4638.
- Perler, F. B., Olsen, G. J., and Adam, E. (1997) *Nucleic Acids Res.* **25**, 1087–1093.
- Petrokovski, S. (1998) *Protein Science* **7**, 64–71.
- Perler, F. B., Comb, D. G., Jack, W. E., Moran, L. S., Qiang, B., Kucera, R. B., Benner, J., Slatko, B. E., Nwankwo, D. O., Hempstead, S. K., Carlow, C. K. S., and Jannash, H. (1992) *Proc. Natl. Acad. Sci. U. S. A.* **89**,

- 5577–5581
15. Nishioka, M., Fujiwara, S., Takagi, M., and Imanaka, T. (1998) *Nucleic Acids Res.* **26**, 4409–4412
16. Weiner, M. P., Costa, G. L., Schoettlin, W., Cline, J., Mathur, E., and Bauer, J. C. (1994) *Gene* **151**, 119–123
17. Wenzlau, J. M., Saldanha, R. J., Butow, R. A., and Perlman, P. S. (1989) *Cell* **56**, 421–430
18. Sambrook, J., Fritsch, E. F., and Maniatis, T. (1989) *Molecular Cloning: a Laboratory Manual*, 2<sup>nd</sup> Ed., Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY
19. Godfroy, A., Meunier, J.-R., Guezennec, J., Lesongeur, F., Raguénès, G., Rimbault, A., and Barbier, G. (1996) *Int. J. Syst. Bacteriol.* **46**, 1113–1119
20. Wende, W., Grindl, W., Christ, F., Pingoud, A., and Pingoud, V. (1996) *Nucleic Acids Res.* **24**, 4123–4132
21. He, Z., Crist, M., Yen, H., Duan, X., Quiocho, F. A., and Gimble, F. S. (1998) *J. Biol. Chem.* **273**, 4607–4615
22. Vipond, I. B., and Halford, S. E. (1996) *Biochemistry* **35**, 1702–1711
23. Vipond, I. B., Moon, B. J., and Halford, S. E. (1996) *Biochemistry* **35**, 1712–1721
24. Gimble, F. S., and Wang, J. (1996) *J. Mol. Biol.* **263**, 163–180

**Inteins of *Thermococcus fumicolans* DNA Polymerase Are Endonucleases with Distinct Enzymatic Behaviors**

Isabelle Saves, Valérie Ozanne, Jacques Dietrich and Jean-Michel Masson

*J. Biol. Chem.* 2000, 275:2335-2341.

doi: 10.1074/jbc.275.4.2335

---

Access the most updated version of this article at <http://www.jbc.org/content/275/4/2335>

Alerts:

- [When this article is cited](#)
- [When a correction for this article is posted](#)

[Click here](#) to choose from all of JBC's e-mail alerts

This article cites 22 references, 10 of which can be accessed free at <http://www.jbc.org/content/275/4/2335.full.html#ref-list-1>