Nucleotide sequence of the gene for alkaline phosphatase of *Thermus caldophilus* GK24 and characteristics of the deduced primary structure of the enzyme

Taeshin Park, Jung-Ha Lee, Hyun-Kyu Kim, Hyang-Sook Hoe, Suk-Tae Kwon *

Department of Genetic Engineering, Sungkyunkwan University, 300 Chunchon-Dong, Jangan-Ku, Suwon 440-746, South Korea

Received 9 August 1999; received in revised form 9 September 1999; accepted 9 September 1999

Abstract

The gene encoding *Thermus caldophilus* GK24 (*Tca*) alkaline phosphatase was cloned into *Escherichia coli*. The primary structure of *Tca* alkaline phosphatase was deduced from its nucleotide sequence. The *Tca* alkaline phosphatase precursor, including the signal peptide sequence, was comprised of 501 amino acid residues. Its molecular mass was determined to be 54,760 Da. On the alignment of the amino acid sequence, *Tca* alkaline phosphatase showed sequence homology with the microbial alkaline phosphatases, 20% identity with *E. coli* alkaline phosphatase and 22% *Bacillus subtilis* (*Bsu*) alkaline phosphatases. High sequence identity was observed in the regions containing the Ser-102 residue of the active site, the zinc and magnesium binding sites of *E. coli* alkaline phosphatase. Comparison of *Tca* alkaline phosphatase and *E. coli* alkaline phosphatase structures suggests that the reduced activity of the *Tca* alkaline phosphatase, in the presence of zinc, is directly involved in some of the different metal binding sites. Heat-stable *Tca* alkaline phosphatase activity was detected in *E. coli* YK537, harboring pJRAP. © 1999 Federation of European Microbiological Societies. Published by Elsevier Science B.V. All rights reserved.

Keywords: Alkaline phosphatase gene; *Thermus caldophilus* GK24; *Tca* alkaline phosphatase; Sequence analysis

1. Introduction

Alkaline phosphatase is a non-specific phosphohomoesterase that functions through a phosphoseryl intermediate to produce free inorganic phosphate or to transfer the phosphoryl group to other alcohols [1]. Alkaline phosphatase has been most extensively studied in *Escherichia coli* [2]. The enzyme is a dimeric metalloenzyme (*m* = 94,000 Da) with 449 amino acid residues per monomer (*m* = 47,000 Da), containing two tightly bound atoms of zinc and one atom of magnesium in each of the two chemically identical subunits [3], and is located in the periplasmic space with neither a carbohydrate nor a fatty acid tail. The crystal structure of the Cd-substituted *E. coli* alkaline phosphatase complexed with inorganic phosphate, at 2.5 Å resolution, and the phosphate-free native *E. coli* alkaline phosphatase, at 2.8 Å resolution, have been refined [4]. The three
metal-ligand constellations (two Zn and one Mg) have been clarified.

Thermostable enzymes are valuable because of biotechnological applications. Thermostable alkaline phosphatases would be suitable for labeling primers and detection of PCR-amplified products. The properties of Thermus caldophilus GK24 alkaline phosphatase have been reported [5]. Tca alkaline phosphatase has an approximate molecular mass of 108,000 Da and consists of two subunits, each with a molecular mass of 54,000 Da. The enzyme is stable at temperatures above 80°C. Tca alkaline phosphatase was inhibited by Zn

$^{2+}$ ions and activated by Mg$^{2+}$ ions. However, zinc ions are essential for the activity of E. coli alkaline phosphatase [1-4]. In order to elucidate the relationship of the metal and its ligands to the active site region of Tca alkaline phosphatase, we attempted to clone its gene. In this paper we report (i) the cloning and nucleotide sequence of the Tca alkaline phosphatase gene, (ii) the characteristics of the deduced primary structure of the enzyme and (iii) expression of the gene in E. coli.

2. Materials and methods

2.1. Bacterial strains, enzymes and reagents

The E. coli strain XL1-blue [6] was used as the host for plasmid preparations. The E. coli strain YK537 [phoA8] [7] was used as the host for gene expression. T4 DNA ligase, T4 polynucleotide kinase, and other restriction enzymes were purchased from Boehringer Mannheim GmbH. Radioactive nucleotides were purchased from Amersham, and pBluescript SK$^{+/-}$ was purchased from Stratagene. Other reagents were obtained from Sigma.

2.2. Amino acid sequence analysis of peptides by BNPS-skatole treatments

Tca alkaline phosphatase was purified from the culture cells of T. caldophilus GK24 as described previously [5]. For isolation of internal peptides of Tca alkaline phosphatase, we performed a cleavage of BNPS-skatole according to the method of Crimmins et al. [8]. The cleaved sample was fractionated by SDS-polyacrylamide gel electrophoresis, then electrophoretically transferred to a polyvinylidene difluoride membrane [9]. Three of these peptides were cut out, and their amino acid sequences were analyzed by Edman degradation with an Applied Biosystems 470A gas-phase protein sequencer.

2.3. Alkaline phosphatase activity assays

Enzyme activity was assayed with p-nitrophenylphosphate according to Kim et al. [5]. One unit of activity was defined as the amount of enzyme which liberated 1 μmol of p-nitrophenol in 1 min at 80°C.

2.4. Molecular cloning and DNA hybridization techniques

Most of the methods used for molecular cloning and hybridization were based on those of Sambrook et al. [6]. Chromosomal DNA of T. caldophilus GK24 was isolated by the method of Marmur [10]. Oligodeoxyribonucleotides were synthesized using an automatic DNA synthesizer (Applied Biosystems).

2.5. DNA sequencing and computer-assisted analysis

The restriction fragments to be sequenced were cloned into appropriate restriction sites of pBluescript SK$^{+/-}$ vectors. DNA sequencing, by the chain-termination method, was performed according to Hattori and Sakaki [11]. Sequence data were analyzed using PCGENE and DNASIS.

2.6. Expression of Tca alkaline phosphatase gene

E. coli YK537, harboring recombinant plasmids, was cultured at 37°C in L-broth containing 50 μg ml$^{-1}$ ampicillin. When the OD$_{600}$ of the culture was about 0.8, the cultures were induced by the addition of isopropyl-β-D-thiogalactopyranoside to a final concentration of 0.2 mM and then incubated at 37°C for 6 h. The cells were collected by centrifugation, suspended in 10 mM Tris-HCl, pH 7.4 using 1 mM MgCl$_2$, and then disrupted by sonication. The sonicate was treated at 80°C for 30 min. After centrifugation, the supernatant was used as crude enzymes for activity assay.
3. Results and discussion

3.1. Molecular cloning of the Tca alkaline phosphatase gene

In the case of Tca alkaline phosphatase, the NH₂-terminal residue was blocked, as described previously[6]. Thus, we performed a cleavage of BNPS-skatole, and then determined the partial amino acid sequences of BNPS-skatole fragments. To clone the Tca alkaline phosphatase gene, two kinds of probes were synthesized on the basis of the internal amino acid sequence of Tca alkaline phosphatase determined as above; (i) 5'-TGGGACGTCC(or G)-TGG(or C)CGG(or C)CGG(or C)ACG-3' (Trp-Asp-Val-Leu-Ala-Ala-Asp-Glu), (ii) 3'-ACCGGG(or C)CTCCCG(or C)CAG(or C)G(or T)CG(or C)CTCCC-5' (sequence of the complementary strand for Trp-Pro-Glu-Gly-Val-Arg-Gln-Gly). Each codon was predicted from the nucleotide sequence of other Thermus genes, i.e., those from Tca DNA polymerase[12] and Tfr DNA ligase[13]. A Southern hybridization of restriction enzyme digests of T. caldophilus GK24 genomic DNA, performed at a high stringency, revealed 2.8-kb Apal and 2.1-kb BamHI fragments which hybridized with the 32P-labelled probes (results not shown). The 2.8-kb Apal and 2.1-kb BamHI fragments were separately ligated at the BamHI and Apal sites of pBluescript SK+ respectively, and then E. coli strain XL1-blue was transformed with the plasmids. A positive clone (named pGAPA), which hybridized with the probes, was identified by screening a genomic library containing the 2.8-kb Apal DNA fragment. A positive clone (named pGAPB), which hybridized with the probes, was also identified by screening a genomic library containing the 2.1-kb BamHI DNA fragment.

3.2. Nucleotide sequence of the Tca alkaline phosphatase gene and its deduced amino acid sequence

The restriction maps of the 2.1-kb BamHI and 2.8-kb Apal fragments are presented in Fig. 1. Each enzyme site of the restriction maps was used for the subcloning and DNA sequencing of cloned DNA fragments. The position of the Tca alkaline phosphatase gene in the cloned fragments is indicated by the open arrow. Fig. 2 shows the partial nucleotide sequence of the 2.1-kb BamHI fragment and the deduced amino acid sequence of Tca alkaline phosphatase. This coding sequence was preceded at a spacing of 9 bp by a potential ribosome-binding sequence (5'-GGAG-3'), which was homologous to the consensus Shine-Dalgarno sequence[14]. The promoter-like sequence in the −35 region and the −10 region, which can function in E. coli, was not found on the upstream flanking region of the Tca alkaline phosphatase gene. In the 3′-noncoding region of the gene, there was no potential transcriptional termination sequence able to form a stem-and-loop structure, followed by a pyrimidine-rich sequence (Fig. 2). The codon region had a G+C content of 69.3%, and showed biased codon usage. Although 48 of the 61 sense codons were used, there was a strong preference for the third codons in G+C (93.4%). Codon

Fig. 1. Restriction enzyme map of the Tca alkaline phosphatase gene and positions of the cloned DNA fragments in plasmid pGAPA (2.8-kb Apal) and pGAPB (2.8-kb BamHI). The restriction enzyme sites used for subcloning are shown: A, Apal; B, BamHI; SI, SacI; SII, SacII; X, XhoI. The open arrow indicates the coding region of the Tca alkaline phosphatase gene.
Fig. 2. Nucleotide and deduced amino acid sequences of the *Tca* alkaline phosphatase gene. Numbers on the left refer to the nucleotide (negative numbers refer to nucleotides upstream from the start codon, ATG, of *Tca* alkaline phosphatase) and the amino acid sequence. A putative Shine-Dalgarno sequence and the signal sequence for *Tca* alkaline phosphatase are underlined and the vertical arrow represents the cleavage site of *Tca* alkaline phosphatase. The amino acids determined by sequencing of the peptides derived from BNPS-skatole-cleaved *Tca* alkaline phosphatase are also underlined. Asterisks indicate the stop codon. The nucleotide sequence of the *Tca* alkaline phosphatase gene and its flanking regions was deposited in GenBank with accession number AF168770.
bias in Tca alkaline phosphatase is similar to that of Tca DNA polymerase [12].

The deduced amino acid sequence agreed with the determined amino acid sequences of the BNPS-skatole peptides, derived from Tca alkaline phosphatase (underlined in Fig. 2). The Tca alkaline phosphatase precursor comprised 501 amino acid residues, and its molecular mass was determined to be 54,760 Da. The amino acid residues and molecular mass of the mature Tca alkaline phosphatase cannot be calculated exactly because the blocking of the NH$_2$-terminal amino acid residue of Tca alkaline phosphatase [5].

In the open reading frame, the NH$_2$-terminus of the deduced amino acid sequence had the characteristics of a signal peptide for a periplasm protein (underlined in Fig. 2). Presumably, the first 27 of these amino acids represent a signal sequence like that found in other secreted proteins [15]. These amino acids appear to be a typical signal sequence, with charged amino acids, followed by a hydrophobic core, and a small side-chain amino acid at the cleavage point (between Ala-27 and Leu-28).

Bacillus subtilis (Bsu) alkaline phosphatase III contains a typical signal sequence of 32 amino acids and is cleaved between Ala-32 and Gln-33 [16]. Thus, the putative mature form of Tca alkaline phosphatase is comprised of 474 amino acid residues, and its molecular mass was estimated to be 52,006 Da.

Fig. 3. Comparison of the amino acid sequences of Tca alkaline phosphatase with those of other alkaline phosphatases. The sequence of Tca alkaline phosphatase (Tca) is shown compared with those of Bacillus subtilis alkaline phosphatase III (Bsu3) and IV (Bsu4) [16] and E. coli alkaline phosphatase (Eco) [3]. Identical amino acids between Tca alkaline phosphatase and others are indicated by dark-shaded boxes. The ligands to the three metals (*) and the phosphorylation site (@) are flagged.
3.3. Comparison of the amino acid sequence of Tca alkaline phosphatase with those of other alkaline phosphatases

Alignment of these alkaline phosphatase sequences (Fig. 3) was done following the strategy of Hulett et al. [16] which used signature sequences derived from comparison of E. coli and Bsu alkaline phosphatases sequences. The whole amino acid sequence of Tca alkaline phosphatase showed low homology to those of E. coli alkaline phosphatase [3] and Bsu alkaline phosphatases [16] (Fig. 3). Tca alkaline phosphatase shows 20% identity to the E. coli alkaline phosphatase, and 22% to Bsu alkaline phosphatases. High sequence identity was observed in the regions containing the amino residues important for metal binding, phosphate binding, and catalytic activity.

E. coli alkaline phosphatase is a dimeric metalloenzyme with two Zn\(^{2+}\) and one Mg\(^{2+}\) in each active site region [3,4]. The importance of these metals in both catalysis and structural stabilization of the enzyme has long been recognized. Zinc is absolutely required for activity, since it enhances the activity of E. coli alkaline phosphatase in the presence of zinc [1–4]. From the X-ray structure of E. coli alkaline phosphatase [4], Zn1 is penta-coordinated by the imidazole ring nitrogen of His-331 and His-412, both carboxyl oxygens of Asp-327 and one of the phosphate oxygens. Zn2 is coordinated tetrahedrally by the imidazole nitrogen of His-370, one of the carboxyl oxygens of Asp-327 and Asp-369, and one of the phosphate oxygens. All ligands to Zn1 and Zn2 in E. coli alkaline phosphatase are well conserved in Tca alkaline phosphatase (Fig. 3). However, His-412, one of the ligands to Zn1, is replaced by arginine (Arg-389) in Tca alkaline phosphatase (Fig. 3). Zinc binding is significantly reduced by mutations at His-412 in E. coli alkaline phosphatase [17]. These mutant enzymes exhibit not only reduced catalytic activity, but also increased K\(_m\) values. Tca alkaline phosphatase was actually activated by addition of Mg\(^{2+}\), but it was inhibited by addition of Zn\(^{2+}\) [5]. The activity of 2 mM EDTA-inhibited Tca alkaline phosphatase can also be restored by adding Mg\(^{2+}\) ions, but not by adding zinc ions. One possible explanation for the zinc inhibition of Tca alkaline phosphatase is that His-412, one of the ligands to Zn1, is replaced by arginine as described in mutant E. coli alkaline phosphatase (Fig. 3). Another possible explanation is that His-119 (corresponding to Asp-153 of E. coli alkaline phosphatase) in Tca alkaline phosphatase may provide an altered magnesium binding site with high affinity for zinc, similar to the mutant E. coli alkaline phosphatase in which Asp-153 is replaced by histidine (D153H) [18]. Asp-153 of E. coli alkaline phosphatase is an indirect ligand of the Mg\(^{2+}\), since it interacts with two of the water molecules that are direct ligands to the Mg\(^{2+}\) [4]. The crystal structure of mutant enzyme D153H reveals that the octahedral magnesium binding site has been converted to a tetrahedral zinc binding site with an imidazole ring nitrogen of His-153 as one of the ligands to zinc [18]. The structural data suggest that the D153H enzyme should be inhibited by addition of zinc, since the zinc could displace the Mg\(^{2+}\) at the Mg binding site. The binding of zinc at this site would result in an inactive enzyme. The addition of zinc to the D153H enzyme does indeed cause inactivation [18]. It is interesting to note that Asp-153 of E. coli alkaline phosphatase is replaced by histidine in Tca alkaline phosphatase (Fig. 3).

From the X-ray structure of E. coli alkaline phosphatase [4], all ligands (Asp-51, Glu-322, Thr-155) to Mg\(^{2+}\) in E. coli alkaline phosphatase are well conserved in Tca alkaline phosphatase (Fig. 3).

Sequence identity was observed in the region containing the reactive serine residue, Ser-68, of Tca alkaline phosphatase, which corresponds to Ser-102 of E. coli alkaline phosphatase (Fig. 3). In the case of E. coli alkaline phosphatase, the replacement at Ser-102 by leucine produces an enzyme with 0.5% activity of the wild-type [19]. The reactive serine residue was found in many serine hydrolases including alkaline phosphatases.

### Table 1

<table>
<thead>
<tr>
<th>Species</th>
<th>Tca alkaline phosphatase activity (U ml(^{-1}) of culture medium)</th>
</tr>
</thead>
<tbody>
<tr>
<td>E. coli YK537 (control)</td>
<td>0.025</td>
</tr>
<tr>
<td>E. coli YK537/pGAPA</td>
<td>0.025</td>
</tr>
<tr>
<td>E. coli YK537/pGAPB</td>
<td>0.180</td>
</tr>
<tr>
<td>E. coli YK537/pJRAP</td>
<td>0.026</td>
</tr>
<tr>
<td>E. coli YK537/pJRAPR</td>
<td>0.025</td>
</tr>
</tbody>
</table>

One unit of activity was defined as the amount of enzyme which liberated 1 µmol of p-nitrophenol in 1 min at 80°C.
3.4. Expression of Tca alkaline phosphatase gene

We examined the expression directed by pGAPA and pGAPB in *E. coli* YK537, using an activity assay of the Tca alkaline phosphatase (Table 1). However, heat-stable alkaline phosphatase activity could not be detected in cells carrying the plasmids. Probably, the promoter of Tca alkaline phosphatase gene will have a different structure from that of *E. coli* alkaline phosphatase, as described above.

For the expression of Tca alkaline phosphatase gene, the 2.1-kb BamHI fragment of pGAPB was inserted into the same site of plasmid pJR [12] and then *E. coli* YK537 was transformed with the plasmid. Plasmids pJRAP and pJRAPR were obtained. pJRAP contained the Tca alkaline phosphatase gene, which was inserted in the same direction as the tac promoter. pJRAPR contained the Tca alkaline phosphatase gene, which was inserted in the opposite direction as the tac promoter. Heat-stable alkaline phosphatase activity was detected only in *E. coli* YK537 harboring pJRAP (Table 1). Low activity of Tca alkaline phosphatase in *E. coli* YK537 harboring pJRAP is supposed to be due to the long distance (100 bp) between the Shine-Dalgarno sequence of the tac promoter and the ATG start codon of the Tca alkaline phosphatase gene. We are now working on the high expression of the Tca alkaline phosphatase gene in *E. coli* using various expression vectors.

Acknowledgements

This work was supported by a Grant-in-Aid from the Korea Science and Engineering Foundation (KOSEF 98-2-04-0101-2).

References