ppc*, the gene for phosphoenolpyruvate carboxylase from an extremely thermophilic bacterium, *Rhodothermus obamensis*: cloning, sequencing and overexpression in *Escherichia coli

Ken Takai,[†] Yoshihiko Sako and Aritsune Uchida

Author for correspondence: Ken Takai. Tel: +81 468 67 3894. Fax: +81 468 66 6364.
e-mail: kent@jamstec.go.jp

Laboratory of Marine Microbiology, Division of Applied Bioscience, Graduate School of Agriculture, Kyoto University, Kyoto 606-8502, Japan

The *ppc* gene, which encodes phosphoenolpyruvate carboxylase (PEPC) of an extremely thermophilic bacterium, *Rhodothermus obamensis*, was directly sequenced by the thermal asymmetric interlaced (TAIL) PCR method. An ORF for a 937 amino acid polypeptide was found in the gene. The *ppc* gene had a high G+C content (66.2 mol%) and the third position of the codon exhibited strong preference for G or C usage (85.0 mol%). The calculated molecular mass was 107 848 Da, which was consistent with the molecular mass of the enzyme as determined by SDS-PAGE (100 kDa). The amino acid sequence of *R. obamensis* PEPC was closely related to that of PEPC from another thermophile, a *Thermus* sp., and from a mesophile, *Corynebacterium glutamicum*, exhibiting 45.3% or 37.7% identity and 61.5% or 56.5% similarity, respectively. By Southern analysis, the *ppc* gene was found to be present in a single copy in the genomic DNA of this organism. The cloned gene was expressed in *Escherichia coli* using a pET expression vector system and a thermostable recombinant PEPC was obtained. Comparison of the deduced amino acid sequences of the thermophilic and mesophilic PEPCs revealed distinct or common preferences for specific amino acid composition and substitutions in the two thermophilic enzymes.

Keywords: extreme thermophile, gene cloning, phosphoenolpyruvate carboxylase, *Rhodothermus obamensis*

INTRODUCTION

Recently, there has been increasing interest in thermophilic organisms due to their novel biochemical machinery which sustains them under extraordinarily high temperatures (Adams, 1993; Adams *et al.*, 1995). A number of proteins have been isolated from various thermophiles and studied with respect to their unusual thermostability (Adams 1993, 1994; Adams *et al.*, 1995; Cowan, 1995; Daniel, 1996; Jaenicke, 1996a, b). To

elucidate the mechanisms of thermostability, protein intrinsic thermostability has been extensively probed by comparing amino acid sequences between thermophilic and mesophilic homologous proteins (Böhm & Jaenicke, 1994; Britton *et al.*, 1995), by analyses of protein folding (Schultes & Jaenicke, 1991; Cavagnero *et al.*, 1995), by using molecular engineering techniques (Arias & Argos, 1989; Cannio *et al.*, 1994; Tamakoshi *et al.*, 1995; Kotsuka *et al.*, 1996) and by determining protein three-dimensional structures (Day *et al.*, 1992; Chan *et al.*, 1995). However, the mechanisms involved have not yet been resolved.

Another approach to an understanding of these mechanisms is to explore extrinsic factors, which lie not in the protein itself but outside it, that might stabilize it under high temperature. We have studied the role of extrinsic factors in the thermostabilization of phos-

[†]**Present address:** Deep-sea Microorganisms Research Group (DEEP-STAR), Japan Marine Science and Technology Center (JAMSTEC), 2-15 Natsushima-cho, Yokosuka 237, Japan.

Abbreviations: PEPC, phosphoenolpyruvate carboxylase; TAIL PCR, thermal asymmetric interlaced PCR.

The EMBL accession number for the *R. obamensis pepC* sequence reported in this paper is X99379.

phenolpyruvate carboxylase (PEPC) (EC 4.1.1.31) from the recently isolated extreme thermophile *Rhodothermus obamensis*, which grows optimally at 80 °C and in the temperature range 50–85 °C (Sako *et al.*, 1996a). PEPC catalyses the reaction that fixes HCO_3^- on phosphoenolpyruvate to form oxaloacetate and inorganic phosphate using Mg^{2+} as the cofactor and primarily plays an anaplerotic role in bacteria by replenishing C_4 -dicarboxylic acids to the citric acid cycle (Utter & Kolenbrander, 1972). Most of the enzymes that have been isolated and purified to date are 400 kDa homotetramers consisting of 100 kDa subunits, but archaeal PEPCs from *Methanothermobacter sociabilis* and *Sulfolobus acidocaldarius* are about 240 kDa homotetramers with 60 kDa subunits (Sako *et al.*, 1996b, 1997). The PEPC purified from *R. obamensis* exhibited molecular and enzymological properties typical of bacterial PEPCs except for its extreme thermophilicity and thermostability (Takai *et al.*, 1997a). This implies that the thermophilicity and thermostability of *R. obamensis* PEPC are distinct from those of other bacterial PEPCs and that this enzyme is a good model for the comparative investigation of thermostability. In addition, the thermostability of *R. obamensis* PEPC was strongly enhanced by substrate, cofactor, salts and positive allosteric effectors, and a possible function of these extrinsic thermostabilization factors was to maintain quaternary structure (Takai *et al.*, 1997a, b). However, the molecular basis of the interaction between these extrinsic factors and enzyme, and structural changes induced by the factors is still unknown. Furthermore, the three-dimensional structure of this enzyme has been not yet been determined. To resolve these problems, the cloning, sequencing and expression of the gene for *R. obamensis* PEPC is essential.

We report here the cloning, sequencing and over-expression in *Escherichia coli* of the gene (*ppc*) for *R. obamensis* PEPC. The deduced amino acid sequence of the enzyme was compared with that of other thermophilic and mesophilic counterparts. Some preferred amino acid compositions and substitutions were found in the thermophilic enzymes, and the possible contribution of these to thermostability is discussed. Based on a phylogenetic analysis of the enzyme, its evolution is also discussed.

METHODS

Bacterial strain and growth conditions. The bacterial strain used in this study was *R. obamensis* OKD7 (JCM 9785), which was isolated from a shallow marine hydrothermal vent at Tachibana Bay, Nagasaki Prefecture, Japan (Sako *et al.*, 1996a). For culture of *R. obamensis*, Jx medium was used (Sako *et al.*, 1996a). The bacterium was grown at 76 °C and collected during the late-exponential growth phase. The cell pellet was frozen at –90 °C prior to genomic DNA extraction.

Genomic DNA extraction and direct sequencing of the *ppc* gene. Genomic DNA was prepared as described previously (Sako *et al.*, 1996a). The purified DNA was used for Southern analysis and direct sequencing of the *ppc* gene. A partial fragment of the *ppc* gene was amplified by PCR. The primers used for amplification had the sequences 5'-TSACTGCYC-

AYCCAACSGA-3' (TAHPT primer) and 5'-GTCCAGGCS-ATSACCCASGGGATGGC-3' (SLRAIP primer), corresponding to the highly conserved amino acid sequences among bacterial PEPCs reported to date and positions 136–142 and 712–718, respectively, in the amino acid sequence of *E. coli* PEPC (Fujita *et al.*, 1984) (Fig. 1). The 1.9 kb PCR product was directly sequenced on both strands by the dideoxynucleotide chain-termination method (Sanger *et al.*, 1977) using a DNA sequencer model 373As (Applied Biosystems). Comparison of the translated amino acid sequence to that of other PEPCs showed significantly high similarity (Fig. 2). In particular, the sequence contained several highly conserved regions, including the putative substrate-binding site motif, FHGRGGXX-GRGG (Ishijima *et al.*, 1985), seen in most PEPCs. These results strongly supported the proposal that the amplified fragment was a part of the *ppc* gene. Based on the partial sequence of the *ppc* gene, thermal asymmetric interlaced (TAIL) PCR was carried out to determine the sequence of the unknown regions adjacent to the known sequence of the gene (Liu & Whittier, 1995). The long specific primers had the sequences 5'-CAGCAGGGTCTCCTGCTCCT-3' (LS1), 5'-CGACTTCGTCGCGAACGGTG-3' (LS2) and 5'-ACCGC-GACGGCAACCGGTAC-3' (LS3) for the upstream region of the known sequence and 5'-GGCCACGCGCAACCGTCTGA-3' (LS4), 5'-CGGCGCTGATCGATCGCGCC-3' (LS5) and 5'-TCAGCCGCTACCCATCGCC-3' (LS6) for the downstream region of the known sequence. For the TAIL PCR on both regions a short arbitrary degenerate primer was used, the sequence of which was 5'-GTCCASWGANAWGAAN-3' (AD1). With the primers LS3 and AD1 for the upstream region or LS4 and AD1 for the downstream region, the primary PCR was carried out by TAIL cycling. The products of the primary PCR were diluted and used as templates for secondary PCR by TAIL cycling with the primer sets LS2 and AD1 for the upstream region, and LS5 and AD1 for the downstream region. Finally, the secondary PCR products were diluted and used as templates for tertiary PCR by normal cycling with the primer sets LS1 and AD1, and LS6 and AD1. About 600 bp of fragment for the upstream region and about 800 bp of fragment for the downstream region were predominantly amplified and both fragments were then directly sequenced by the dideoxynucleotide chain-termination method. Both DNA sequences were successfully connected with the known sequence and their translated amino acid sequences revealed significantly high similarity with amino acid sequences of other PEPCs (Fig. 2).

Southern analysis. Equal amounts of genomic DNA (15 µg per lane) were digested with restriction endonucleases, electrophoresed in a 1.0% (w/v) agarose gel and transferred to a positively charged nylon membrane (Boehringer Mannheim). As a hybridization probe, a 1.3 kb RNA fragment labelled with digoxigenin (DIG)-11-UTP was used. For labelling, a 1.3 kb DNA fragment was amplified by PCR using the primers with the sequences 5'-CTTTGCAGATCGAAATCGAAGGC-3' (INIT1 primer) and 5'-CCCGGTAGTCGTTCTCGACA-3' (SPR2) indicated in Fig. 1, and subcloned into a site adjacent to the T7 RNA polymerase promoter in the pCR II vector (Invitrogen). The subcloned vector containing the 1.3 kb fragment was linearized with *Hind*III and used as a template for *in vitro* transcription. The DIG RNA labelling kit (Boehringer Mannheim) was used for labelling and the procedure used was as described in the manufacturer's manual. Hybridization was carried out overnight in 50% (v/v) formamide, 5 × SSC (1 × SSC is 150 mM NaCl, 15 mM sodium citrate, pH 7.0), 0.02% (w/v) SDS, 0.1% (w/v) sodium lauroylsarcosine and 2% (w/v) blocking reagent

Fig. 1. Scheme for sequencing the *ppc* gene. A partial fragment of the *ppc* gene was amplified by PCR using TAHPT and SLRAIP primers and directly sequenced on both strands. Based on the sequence of the fragment, TAIL PCR was carried out to determine the sequence of the unknown regions. The amplified fragments containing 5'- and 3'-flanking regions of the *ppc* gene were directly sequenced on both strands. The hatched box indicates a 1.3 kb RNA probe complementary to the fragment amplified with the INIT1 and SPR2 primers. The numbers above the box indicate the nucleotide positions in the gene.

(Boehringer Mannheim) at 60 °C. The filter was washed twice with $2 \times$ SSC, 0.1% SDS at room temperature for 5 min and then washed twice with $0.1 \times$ SSC, 0.1% SDS at 68 °C for 30 min. The detection of DIG-labelled RNA hybridized with homologous DNA sequence was carried out with the DIG luminescent detection kit for nucleic acids (Boehringer Mannheim).

Alignment and analysis of amino acid sequences. A multiple alignment of amino acid sequences from various organisms was constructed with the software package ODEN version 1.1.1 (National Institute of Genetics, Mishima, Japan). The sequence identity and similarity for each aligned pair were then calculated based on the alignment. Continuous gaps were regarded as a single substitution in the calculation and sequence similarity was defined as the percentage of sites occupied by residues sharing the same physico-chemical properties in the aligned pair. The molecular mass inferred from the sequence, codon usage in the *R. obamensis pepC* gene and amino acid compositions of various PEPCs were calculated by DNASIS software version 3.6 (Hitachi). Based on the multiple alignment, the evolutionary distances among various PEPCs known to date were calculated by the method of Kimura and the phylogenetic tree was constructed by the neighbour-joining method (Saitou & Nei, 1987) using the ODEN software. The secondary structures of three PEPCs, from *R. obamensis*, *Corynebacterium glutamicum* and a *Thermus* sp., were inferred from the amino acid sequences by the method of Chou & Fasman (1978) using the DNASIS software and compared to each other with respect to the location of the amino acid substitutions.

Cloning and overexpression of the *ppc* gene. The complete *ppc* gene of *R. obamensis* was amplified by PCR using the INIT2 primer with the sequence 5'-CATATGCTTCCCCCTTTGCAGATCGAA-3' as a primer for site-directed mutagenesis to replace the GTG initiation codon by ATG and to create a *Nde*I site, and the END primer with the sequence 5'-CTATCCGGTGCTCTGCATGGCGG-3'. The amplified gene was subcloned into the pCR II vector and the resulting plasmid was designated pCRP2.8. A 2.8 kb fragment con-

taining the *ppc* gene was cut out of pCRP2.8 with the restriction enzymes *Nde*I and *Bam*HI and ligated into the pET 11a expression vector (Stratagene). The resulting plasmid was designated pEPEP2 and contained a 2.8 kb ORF for *R. obamensis* PEPC.

E. coli BL21 cells containing pEPEP2 were cultured at 37 °C in LB medium (10 g tryptone l⁻¹, 5 g peptone l⁻¹, 10 g NaCl l⁻¹) containing 50 µg ampicillin ml⁻¹. In the middle of the exponential growth phase, 0.2 mM IPTG was added to the medium and growth at 37 °C continued for a further 5 h in the presence or absence of IPTG. Cells were harvested by centrifugation (8000 g, 20 min) at 4 °C and frozen at -90 °C prior to preparation of crude extracts.

Preparation of crude extracts and enzyme assay conditions. Thawed cell pastes were suspended in 50 mM Tris/HCl (pH 8.0). The cells were broken by sonication and centrifuged at 24000 g for 20 min at 4 °C. The supernatants were used as crude extracts. These crude extracts were used both for electrophoresis and also in enzyme assays for PEPC activity at 65 °C. The PEPC activity of the crude extracts was routinely coupled to the malate dehydrogenase reaction and assayed by the previously described method (Takai *et al.*, 1997a, b).

Purification and characterization of recombinant PEPC. *E. coli* BL21/pEPEP2 cells were grown at 37 °C in LB medium containing 50 µg ampicillin ml⁻¹ to the mid-exponential growth phase. IPTG (0.2 mM) was then added to the medium and the cells were grown for a further 8 h at 37 °C. The cells were harvested as described above. The cell paste was suspended in 50 mM Tris/HCl (pH 8.0) and broken by sonication. The crude extract prepared as described above was dialysed against 100 vols 50 mM Tris/HCl (pH 8.0) twice and incubated at 75 °C for 30 min.

After heat treatment, the extract was centrifuged at 24000 g for 20 min at 4 °C and the supernatant was applied to a column of Q-Sepharose (bed volume 50 ml) (Pharmacia) equilibrated with 50 mM Tris/HCl (pH 8.0). The recombinant PEPC was eluted with a linear gradient of NaCl (0–0.6 M NaCl) in 50 mM Tris/HCl (pH 8.0). The active

R. obamensis	MLPPLQIEIEGTGISRPLSEHVNLGGLLQVQIEMAGPEMELVETLRRLLCKQAAQENR	60
C. glutamicum	-----MTDFLRDDIRFLGQILGEVIAEQEGQEVYELVEQARLTSFDIAK--G	45
Thermus sp.	-----MSDPFEALKAEVDLLGRLLGEAIRKVSGERFFALVEEVRLSKARRQ-GD	49
R. obamensis	PEFREQAYTRIHSATYDELLWLLRAYTAFFHLVNQAEQQEIIIRINRERAQQSTPERPRPE	120
C. glutamicum	NAEMDSLQVFDGITPAKATPIARAFSHFALLANLAED--LYDEELREQALDAGDTPPD	102
Thermus sp.	GAAAEVLSQRVERMPVEEMEALVRAFTHYFHLVNLAEERHVRVNRRLRTEGETLENPRPE	109
R. obamensis	-SIDEAILALKQQGRITLDDVLTLLERLDIQPTVTAHPTEARRRSILYKQQHIAQMLSQQR	179
C. glutamicum	STLDATWLKLNENGVGAEAVADVLRAEAPVLTAAHTPTERRRTVFDAQWITTHMRERH	162
Thermus sp.	-GFLALAKALKERGLSLEEAHLNRLALLTFTAHPTETRRRTLRRHHLERLQEELEGGD	168
R. obamensis	RCQ-----LTPEEQETLLLDLHNQITLLLGTAEVREERPTVRDEVEQGLYFIQSTIWEA	233
C. glutamicum	ALQSAEPTARTQSKLDEIEKNIRRRITILWQTALIRVARPRIEDEIEVGLRYYKLSLLEE	222
Thermus sp.	R-----ER-----LLARVLLYATEEVRKARPSVEDEIKGGLYLPTTLWRA	210
R. obamensis	VPRIEDVRRALRRYYGADVDFRFLRYRSWIGSDRDGNPYVTPEITRWALTQRRLALQ	293
C. glutamicum	IPRINRDVAVELRERFEGVPLKPVVKGSGWIGDHDGNPYVTAEVYESTHRAAETVLK	282
Thermus sp.	IPKVEGLEAALERVYKRPFLRSPVFRFSWMMGGDRDGNPYVTPEVTAFAGRYAREVAKG	270
R. obamensis	RYMEELRQLRRRLSLSDRYVAPPEELRRSLARDAREVSLPPHVLQRFHESFRLKISYIM	353
C. glutamicum	YYARQLHSLEHLSLSDRMKNVTPQLLA-----LADAGHNDVPSRVDEPYRRVAVHGR	335
Thermus sp.	RYLEELALVRDLSSLSEARIPVPKEVRE-----GGEG---VERFPGEYRRYFAALY	319
R. obamensis	GRLHGLLQALDDPTQPA-----PDYDADAFVEDLRLQLRCLEACGLERIARHDQLTR	405
C. glutamicum	GRILATTAELIGEDAVEGVWFVFTPYASPEEFLNDALTIDHSLRES-KDVLIADDRSLV	394
Thermus sp.	RALEG-----EALSTEGLARALKVAEKGLEGVGLAQVAQAFRLP-	358
R. obamensis	LLVLAQTFGFHLVTLDVQRHSSVHEAAVAELLRLAGVENDYRALPESRRQELLAEELSNP	465
C. glutamicum	LISAIESFGFNLYALDLRQNSSEYEDVLTelferaQVTANYRELSEAKEVLKELRSP	454
Thermus sp.	LEARLSAFGLELAPLDLREESGKLLAAEALLRLGGVHPDFLALSPEEKALLTEELKTA	418
R. obamensis	RPLLPGAR-VSEATRQVLETFVIRELVQLDP-RLVGSYIVSMHTVSDLLPEMLLAKE	523
C. glutamicum	RPLIPHSDEYSEVTDRELGI FRTASEAVKKFGPRMVPHCII SMASVTDVLEPMVLLKE	514
Thermus sp.	RPLLP-----VGEVPQGEALRVALGALRAWGD---KGAHVVSMTHTHPADLLAVFLLARE	469
R. obamensis	VGLWHYERDPRGTGKPGHVRCPIDFVPLFETIEDLEAAASRMEAILSHPVYRMQVAARGGF	583
C. glutamicum	FGLIAAN---GDNP---RGTVDPVPLFETIEDLQAGAGILDELWKIDLYRNYLLQRDNV	567
Thermus sp.	VGLYRP-----GKP---LPFDVVPVLFETIEDLERAPEVLRLLANPVRFAHAQGRGG-	518
R. obamensis	QEIMLGYSdstkdGGYWMANWALHRAQEQLAEVCLRHGVDFRLFHGRGGTVGRGGGRANQ	643
C. glutamicum	QEVMLGYSdsnkDGGYFSANWALYDAELQLVELCRSAGVKRLFLHGRGGTVGRGGGPSYD	627
Thermus sp.	VEVMIGYSdsnkDAGFLMANLALYQAEALHAGVGEAQGIPVFFHGRGTSTARGGGPAGR	578
R. obamensis	AILAMPPVHNGRIRFTEQGEVIFRYALPEIAHRHLEQIVNAMLRVVGLPAASTGDTGTD	703
C. glutamicum	AILAQPRGAVQGSVRITEQGEIISAKYGNPETARRNLEALVS-----ATLEASLLDVSE	681
Thermus sp.	AIAGLPPKSVGHRLRLTEQGEALADRYAHPDLAVRHLEQLL-----YHFAQAALGDGVE	632
R. obamensis	PATRNRLMD---ELAARSMRAYRRLIDAP-DFWSWYTRITPIDQISRLPIASRPVSRSSA	759
C. glutamicum	LTDHQRAYDIMSEISELSLKKYASLVHEDQGFIDYFTQSTPLQEIGSLNIGSRPSSR--K	739
Thermus sp.	PKAHWREALG--EAGERSMARYRALLSQE-GFFPFPEAFPIREIGELPIASRPVYR-HG	688
R. obamensis	REVDFESLRAIPVFWAFTQVRYLIPGWFGIGQALDELL---QTSPEHLETLRTWYRSWPF	816
C. glutamicum	QTSSVEDLRAIPVWLSWSQSRVMLPGWFGVGTALQWIGEGEQATQRIAEQLTNESWPF	799
Thermus sp.	RVRDIRDLRAIPVWMAWTQVRLLLPGWYGLS-----ALEGLPMLLREMYREWPF	738
R. obamensis	FRTVLQNAQREMVRARLEIAAYYDRLLGD--GPTAFHQMIEDYHRARTAILRITDQESL	874
C. glutamicum	FTSVLDNMAQVMSKAELRLAKLYADLIPDEVAERVYSVIREEYFLTKKMFVITGSDOL	859
Thermus sp.	FATTLESAAALAKADLGIAERYLKLVE--GLQGFYHHLAEYRRTVALLEAIFAP-L	795
R. obamensis	LDHDP IIRKSVQLRNPYTDVNLVQLELMRRIRSGAEADREPLRRALFLSINGIAAAMQS	934
C. glutamicum	LDDNPLLARSVQRRYPYLLPLNVIQVEMMRRYRK--DQSEQVSRNIQLTMNGLSTALRN	917
Thermus sp.	LHNQKTLERQIALRNPYVDFINFVQYELARYRAGGREDEGVRRALLLSLLGVAAGLRN	855
R. obamensis	TG	936
C. glutamicum	SG	919
Thermus sp.	AG	857

Fig. 2. Alignment of PEPC sequences from *R. obamensis*, *Thermus* sp. and *C. glutamicum*. Amino acids that are identical in the three PEPCs are indicated by asterisks and those that are physico-chemically similar are indicated by dots. Dashes in the sequences represent gaps introduced to maximize alignment.

Fig. 3. Unrooted phylogenetic tree of representative bacterial and eukaryotic PEPCs. Numbers represent the bootstrap value on the branch (1000 replicates). The scale bar indicates substitutions per residue. The PEPC sequences in this figure are from GenBank under the following accession numbers: *Anabaena variabilis*, M80541; *Anacystis nidulans*, M11198; *Corynebacterium glutamicum*, M25819; *Thermus sp.*, D42166; *Rhodothermus obamensis*, X99379; *Escherichia coli*, X05903; *Haemophilus influenzae*, L46266; *Zea mays* C4, X03613; *Sorghum vulgare* C4, X55664; *Mesembryanthemum crystallinum* CAM, X13660; *Pisum sativum* root nodule, D64037; *Medicago sativa* root nodule, M83086; *Glycine max* C3, D13998; *Solanum tuberosum* C3, X67053; *Nicotiana tabacum* C3, X59016; *Flaveria trinervia* C4, X61304; *Flaveria australasica* C4, Z25853.

fractions eluted with 0.15–0.25 M NaCl were applied to a column of Phenyl-sepharose (bed volume 20 ml) (Pharmacia) equilibrated with 50 mM Tris/HCl (pH 8.0). The recombinant PEPC was eluted with the same buffer containing 40% (v/v) ethylene glycol.

The purified enzyme was dialysed against 100 vols 50 mM Tris/HCl (pH 9.0) containing 10% (v/v) glycerol twice and tested with respect to its enzymological properties and thermostability. The effect of temperature and pH on the activity, the requirement for divalent cations and the effect of positive or negative allosteric effectors on activity were examined under standard assay conditions as described previously (Takai *et al.*, 1997a, b) except for the modification of the pH value (from 8.0 to 9.0) in the reaction mixture.

The thermostability of the recombinant PEPC was also determined under the same conditions as the previous study on the wild-type *R. obamensis* PEPC (Takai *et al.*, 1997a, b) except for the pH value (9.0) of the incubation buffer. To test the sensitivity of the enzyme to the strong denaturants urea and guanidine hydrochloride, the same concentration (9 $\mu\text{g ml}^{-1}$) of wild-type and recombinant PEPCs was incubated in 50 mM Tris/HCl (pH 8.0 for wild-type and pH 9.0 for

recombinant enzyme) containing various concentrations of denaturants at 20 °C for 30 min. The residual activities were then measured under the assay conditions described above.

Other methods. PAGE of the crude extracts was performed with a 10% (w/v) polyacrylamide gel in the presence of SDS by the method of Laemmli (1970). Molecular mass markers for SDS-PAGE were from Bio-Rad. Protein concentrations were routinely estimated by the method of Bradford (1976) with bovine serum albumin as the standard.

RESULTS

Sequence of the *ppc* gene

The nucleotide sequence of the *ppc* gene of *R. obamensis* was successfully determined by using TAIL PCR methods as shown in Fig. 1. In the sequence, the termination codon TAG and the possible initiation codon GTG, which is often seen in genes from thermophiles (Itaya & Kondo, 1991), were found. The position of the initiation codon was quite similar to the initiation positions of other PEPCs (Fig. 2) and was preceded by a

Table 1. Preferred amino acid changes in PEPCs from *C. glutamicum*, *Thermus* sp. and *R. obamensis*

Analysis was carried out over the region of the sequence alignment shown in Fig. 2. The comparison between *C. glutamicum* and both extreme thermophiles indicates the amino acid changes from *C. glutamicum* PEPC to identical residues in both thermophilic PEPCs. (C) indicates the frequent substitutions commonly seen in both *Thermus* sp. and *R. obamensis* (thermophile-specific substitutions), and (R) or (T) indicate the frequent substitutions specifically seen in *R. obamensis* or *Thermus* sp. (*R. obamensis*-specific substitutions or *Thermus*-specific substitutions). Asterisks indicate the frequent simultaneous substitutions in the comparison between *C. glutamicum* and both extreme thermophiles and bold type indicates the possible thermophilic substitutions proposed or observed previously (Argos *et al.*, 1979; Privalov & Gill, 1988; Zuber, 1988; Arias & Argos, 1989).

Comparison	Total amino acid changes	Amino acid substitution	No. of changes
<i>C. glutamicum</i> and <i>R. obamensis</i>	506	Ile → Leu* (C)	13
		Glu → Ala (C)	12
		Ala → Leu (C)	11
		Leu → Ile (R)	11
		Ser → Thr*	10
		Val → Ile (R)	9
		Ala → Gln (R)	8
		Val → Leu* (C)	8
		Val → Arg	8
		Thr → Leu	8
<i>C. glutamicum</i> and <i>Thermus</i> sp.	496	Val → Leu* (C)	14
		Ser → Ala (C)	14
		Ile → Leu* (C)	13
		Asp → Glu (T)	10
		Ala → Arg*	9
		Ala → Glu (T)	9
		Ala → Leu (C)	8
		Glu → Ala (C)	8
		Leu → Val	7
		Val → Ala	7
		Met → Leu*	7
		Gln → Glu*	7
<i>C. glutamicum</i> and both extreme thermophiles	104	Ile → Leu* (C)	7
		Ile → Val*	4
		Ser → Thr*	4
		Ala → Arg*	4
		Val → Leu* (C)	3
		Met → Leu*	3
		Gly → Ala*	3
		Gln → Glu*	3
<i>Thermus</i> sp. and <i>R. obamensis</i>	448	Leu → Ile (R)	19
		Ala → Glu (T)	11
		Val → Ile (R)	11
		Gly → Ala*	11
		Glu → Arg	10
		Glu → Asp	10
		Ala → Val	9
		Ala → Arg*	9
		Val → Leu* (C)	8
		Glu → Gln	8
		Ala → Gln (R)	7
		Arg → Gln	7

putative rRNA binding site (Shine–Dalgarno sequence). These observations suggested that this GTG codon was the proper initiation codon of the *ppc* gene.

The ORF consisted of 2811 bp and encoded a polypeptide of 937 amino acid residues with a calculated molecular mass of 107 848 Da. This value matched the apparent molecular mass of the purified *R. obamensis* PEPC estimated by SDS-PAGE (Takai *et al.*, 1997a). The G + C content of the *R. obamensis ppc* coding region was high (66.2 mol%) and an extreme bias for G or C was found in the third position of the codons (85.0 mol%).

Southern analysis

To estimate the copy number of the *R. obamensis ppc* gene, a 1.3 kb RNA probe labelled with DIG-11-UTP, corresponding to the 5' half of the *ppc* gene (Fig. 1) was hybridized to *R. obamensis* genomic DNA digested with various restriction endonucleases. *Hind*III, *Bam*HI and *Sma*I restriction digests showed a single fragment that hybridized to the probe and the *Pst*I digest contained two fragments that corresponded to the presumed upstream and downstream fragments at the *Pst*I restriction site of the gene (Fig. 1). These results indicated that PEPC is encoded by a single copy of the gene.

Alignment of PEPCs and phylogenetic analysis

The deduced amino acid sequence of *R. obamensis* PEPC was aligned with that of 16 PEPCs from various bacteria and plants. The alignment indicated that the amino acid sequence of *R. obamensis* PEPC was closely related to that of PEPCs from a thermophile, a *Thermus* sp. (T_{opt} for growth 75 °C) (Nakamura *et al.*, 1995), and a mesophile, *C. glutamicum* (T_{opt} for growth 37 °C) (Eikmanns *et al.*, 1989; Regan *et al.*, 1989), showing 45.3% or 37.7% identity and 61.5% or 56.5% similarity, respectively. The alignment of the amino acid sequences of these three closely related PEPCs is shown in Fig. 2.

Based on the multiple alignment of 17 PEPCs, including that of *R. obamensis*, an unrooted phylogenetic tree was constructed by the neighbour-joining method (Fig. 3) (Saitou & Nei, 1987). The topology of the tree indicated that *R. obamensis* PEPC was most closely related to *Thermus* PEPC and that these extremely thermophilic PEPCs shared a common ancestor with the mesophilic enzyme (Fig. 3).

R. obamensis PEPC amino acid composition and substitutions

The amino acid composition of *R. obamensis* PEPC was compared with that of representative PEPCs from various sources. In a comparison of amino acids of three distinct physico-chemical types (neutral non-polar, neu-

tral polar, and acidic or basic), no apparent trends were found in *R. obamensis* PEPC, whilst the content of neutral nonpolar amino acids increased and the content of neutral polar amino acids decreased in *Thermus* PEPC. However, some preferences for specific amino acids were observed in the two thermophilic PEPCs. In *R. obamensis* PEPC, the content of Gly, Ser, Asn and Lys decreased and the content of Gln, His and Arg increased. Among these, the amino acid composition changes in Asn, Ser, Lys, His and Arg contents were consistent with the changes in *Thermus* PEPC. However, a lower content of Gly and a higher content of Gln were specific to *R. obamensis* PEPC.

In addition, amino acid substitutions were examined among the closely related PEPCs from *R. obamensis*, *Thermus* sp. and *C. glutamicum* to examine whether any bias for specific amino acid substitutions existed between the mesophilic and thermophilic PEPCs (Table 1). Three types of substitutions (*R. obamensis*-specific, *Thermus*-specific or thermophile-common substitutions) were found to be dominant in the thermophilic PEPCs (Table 1). Most of the thermophile-common substitutions were in good agreement with those previously observed with other thermophilic proteins (Argos *et al.*, 1979; Arias & Argos, 1989) but the substitutions of Ile for Leu, Ile for Val and Gln for Ala, dominant in *R. obamensis* PEPC (*R. obamensis*-specific substitutions) were not consistent with the general rule for thermophilic proteins. When the secondary structures of these PEPCs were compared, it was found that most of the substitutions of Ile for Leu and Ile for Val occurred in β -sheets whereas most of the *Thermus*-specific substitutions were found in α -helices.

The expression of the *R. obamensis ppc* gene in *E. coli* and characterization of the recombinant PEPC

E. coli BL21 cells were transformed with the recombinant plasmid pEPEP2 to express the cloned gene. The specific activity of PEPC in the crude extract from IPTG-induced *E. coli* BL21/pEPEP2 was 32.2 $\mu\text{mol NADH min}^{-1} (\text{mg protein})^{-1}$ at 65 °C. No PEPC activity was observed at 65 °C in the crude extract from *E. coli* BL21/pEPEP2 grown in the absence of IPTG. Fig. 4(a) shows the results of SDS-PAGE of crude extracts from *E. coli* BL21/pEPEP2 grown in the presence or absence of IPTG. A protein of 100 kDa was observed in extracts from IPTG-induced cells (lane 3 in Fig. 4a). The calculated molecular mass of the protein and its migration on SDS-PAGE gels were identical to those of the native *R. obamensis* PEPC. These results indicated that a thermophilic PEPC of *R. obamensis* was successfully expressed in *E. coli*.

The recombinant PEPC was purified to electrophoretical homogeneity by heat treatment at 75 °C for 30 min followed by Q-Sepharose and Phenyl-sepharose column chromatography (Fig. 4b and Table 2). Based on SDS-PAGE, the purified recombinant PEPC was electrophoretically identical to the native one (Fig. 4c). The purified recombinant PEPC was then characterized with

Fig. 4. SDS-PAGE of recombinant PEPC expressed and purified in *E. coli* BL21. (a) Approximately 20 μ g of protein was applied to each lane. After electrophoresis, the gel was stained with 40% (v/v) methanol/10% (v/v) acetate containing 0.2% (w/v) Coomassie brilliant blue R250. Lanes: 1, molecular mass markers (Bio-Rad); 2, crude extract from *E. coli* BL21/pEPEP2 grown for 5 h in the absence of 0.2 mM IPTG; 3, crude extract from *E. coli* BL21/pEPEP2 grown for 5 h in the presence of 0.2 mM IPTG. A 100 kDa protein corresponding to the recombinant PEPC is indicated by the arrow on the right. (b) Approximately 5 μ g of protein were applied to each lane. 1, Molecular mass markers; 2, crude extract from *E. coli* BL21/pEPEP2 grown for 8 h in the presence of 0.2 mM IPTG; 3, 4 and 5, samples after heat treatment at 75 °C for 30 min, Q-Sepharose column chromatography and Phenyl-sepharose column chromatography, respectively. A protein band corresponding to the recombinant PEPC is indicated by the arrow on the right. (c) Approximately 1 μ g of the purified native (lane 1) and recombinant (lane 2) PEPC were run on the gel side by side. The native PEPC used in this experiment had been purified and frozen in liquid nitrogen (Takai *et al.*, 1997a).

Table 2. Purification of recombinant PEPC from *E. coli* BL21/pEPEP2

Preparation	Volume (ml)	Total protein (mg)	Total activity (μ mol NADH min ⁻¹)	Yield (%)	Specific activity [μ mol NADH min ⁻¹ (mg protein) ⁻¹]	Purification (-fold)
Crude extract	55.0	228	8828	100	38.7	1.0
Heat treatment at 75 °C for 30 min	46.0	73.6	7590	85.9	103	2.7
Q-Sepharose	18.3	36.6	4233	47.9	116	3.0
Phenyl-sepharose	6.0	4.8	2430	27.5	506	13.1

respect to its enzymological properties and stability. The recombinant enzyme was purified 13.1-fold with a final specific activity of 506 μ mol NADH min⁻¹ (mg protein)⁻¹ under standard assay conditions and an overall yield of about 27.5% (Table 2). The enzymological

properties of the recombinant PEPC were basically similar to those of the wild-type (Table 3). The only apparent difference was a change in optimum pH (from 8.0 to 9.0) for activity. The recombinant PEPC was extremely thermostable and the times required for 50%

Table 3. Comparison of enzymological properties and thermostability between the wild-type PEPC from *R. obamensis* and recombinant PEPC expressed in *E. coli* BL21

The enzymological properties of the recombinant enzyme were determined under the same conditions as those of the wild-type (Takai *et al.*, 1997a) except for the modification of the pH value (from 8.0 to 9.0) in the reaction mixture. In the determinations of thermostability of the recombinant PEPC, thermoinactivation was carried out as described previously (Takai *et al.*, 1997a) except for the pH value (9.0) of the incubation buffer. $K_{Mg^{2+}}$, K_{PEP} , $S_{Acetyl-CoA}$ and $S_{F1,6-BP}$ indicate the concentrations which resulted in 50 % maximum activity or maximum effectiveness. Similarly, S_{L-Asp} and $S_{L-Malate}$ indicate the concentrations which resulted in 50 % inhibition of enzyme activity. Values in bold type are those found to be different between the wild-type and the recombinant PEPC.

	Wild-type PEPC from <i>R. obamensis</i>	Recombinant PEPC from <i>E. coli</i> BL21
Enzymological properties*		
T_{opt}	~ 75 °C	~ 75 °C
pH _{opt}	8.0	9.0
$K_{Mg^{2+}}$ (mM)	~ 0.5	~ 0.5
K_{PEP} (mM)	20.9	18.2
V_{max} [μ mol NADH min ⁻¹ (mg protein) ⁻¹]	378	409
$S_{Acetyl-CoA}$ (mM)	~ 0.35	~ 0.30
$S_{F1,6BP}$ (mM)	~ 1.2	~ 1.5
S_{L-Asp} (mM)	~ 0.7	~ 0.6
$S_{L-Malate}$ (mM)	~ 0.3	~ 0.3
Thermostability		
Half-life (min)	240 (90 °C) 60 (91 °C) 10 (93 °C)	120 (90 °C) 30 (91 °C) 15 (93 °C)
Effect of extrinsic factors		
(-fold decrease in inactivation rate)†		
PEP	3.14	3.02
MgSO ₄	4.37	0.20
Na ₂ SO ₄	5.58	4.56
Acetyl-CoA	8.81	10.4

* F1,6-BP, fructose 1,6-bisphosphate; PEP, phosphoenolpyruvate.

† Takai *et al.* (1997a).

loss of activity were about 120 min at 90 °C, 30 min at 91 °C and 15 min at 93 °C. This extreme thermostability of the recombinant PEPC was quite similar to that of the wild-type PEPC (Table 3) (Takai *et al.*, 1997a). Similarly, the stability of *R. obamensis* PEPC toward the strong denaturants, urea and guanidine hydrochloride was conserved in the recombinant enzyme from *E. coli*. In addition, the extrinsic thermostabilization systems were effective in the recombinant PEPC and the effect of most extrinsic thermostabilization factors on the thermostability of the recombinant enzyme was identical to that on the wild-type *R. obamensis* PEPC. Only MgSO₄, which was an effective thermostabilization factor for the wild-type *R. obamensis* PEPC, strongly decreased the thermostability of the recombinant enzyme (Table 3).

DISCUSSION

In this study the gene for PEPC from an extremely thermophilic bacterium, *R. obamensis* was successfully cloned, sequenced and expressed in *E. coli* (Fig. 1). The gene consisted of 2811 bp and encoded a polypeptide of 937 amino acid residues with a calculated molecular

mass of 107848 Da. The G+C content of the *R. obamensis* *ppc* coding region was 66.2 mol % and the content in the third position of the codons was 85.0 mol %. The G+C content of the genomic DNA and of the 16S rRNA gene (EMBL accession number X95071) of *R. obamensis* are 66.6 mol % and 60.2 mol %, respectively (Sako *et al.*, 1996a). The high G+C content found in the genomic DNA and the *ppc* gene was also observed in the genus *Thermus* (Brock & Brock, 1984; Nakamura *et al.*, 1995). The extreme bias for a G or C was probably associated with the preference for some amino acid compositions and substitutions in the thermophilic PEPCs. In fact, the comparison of amino acid composition revealed that some preferential amino acid compositions are present in thermophilic PEPCs. In *R. obamensis* PEPC, the content of Gly, Ser, Asn and Lys decreased and the content of Gln, His and Arg increased. Among these, the increased Gln content was unusual with respect to previously proposed rules for thermophilic adaptation (Argos *et al.*, 1979; Arias & Argos, 1989). The contribution of the increased Gln residues to thermostability is quite interesting but still unclear. In addition, the amino acid substitutions were

examined among the closely related PEPCs from *R. obamensis*, *Thermus* sp. and *C. glutamicum* (Fig. 2) to examine whether some bias for the specific amino acid substitutions existed between the mesophilic and thermophilic PEPCs (Table 1). From the phylogenetic analysis of PEPC described below, it is suggested that the two thermophilic PEPCs of *R. obamensis* and *Thermus* sp. share a common ancestor with the mesophilic enzyme from *C. glutamicum*. Hence, the alterations of amino acids between *C. glutamicum* and each thermophile were likely to reflect a possible strategy for intrinsic thermal adaptation of the thermophilic PEPCs (Table 1). As thermophile-common substitutions, the replacements of Ile by Leu, Val by Leu, Ala by Leu and Ala by Arg were found and these are the possible thermophilic substitutions previously proposed or observed (Argos *et al.*, 1979; Privalov & Gill, 1988; Zuber, 1988; Arias & Argos, 1989). However, the *R. obamensis*-specific replacements of Leu by Ile, Val by Ile and Ala by Gln were not consistent with such rules for thermophilic adaptation.

In a comparison of amino acid compositions of D-glyceraldehyde-3-phosphate dehydrogenase (GAPDH) from mesophilic and thermophilic sources, it was shown that the Ile content was significantly increased and the Leu content was decreased in the GAPDH of a hyperthermophilic archaeon, *Pyrococcus furiosus*, and the increased Ile content was also seen in the GAPDH of a hyperthermophilic bacterium, *Thermotoga maritima* (Böhm & Jaenicke, 1994). In addition, Britton *et al.* (1995) indicated that the replacements of Leu by Ile and of Val by Ile were the most frequent substitutions in comparisons of the glutamate dehydrogenases from the mesophilic bacterium *Clostridium symbiosum* and the hyperthermophilic archaea *Thermococcus litoralis* and *P. furiosus*. It was suggested from the inferred three-dimensional structures of the hyperthermophilic glutamate dehydrogenases that these exchanges of amino acids involved the addition of an extra methyl group to the enzymes and the increased packing density caused by these exchanges was associated with the thermostability of the thermococcal enzymes (Britton *et al.*, 1995). In comparisons of the secondary structures of the closely related PEPCs from *R. obamensis*, *Thermus* sp. and *C. glutamicum* the *R. obamensis*-specific substitutions were found to occur in the β -sheets. It seems likely that the increased hydrophobicity arising from the exchange from Leu to Ile or from Val to Ile has an effect on stabilizing the β -sheet structure in *R. obamensis* PEPC. However, in the higher levels of structure, as suggested in the thermococcal glutamate dehydrogenases, these exchanges might be involved in the increased packing density at the molecular cavity or hydrophobic core in the enzyme. Further investigation of the specific amino acid composition and substitutions in *R. obamensis* PEPC will require structural data, including the three-dimensional structure of the enzyme and the structure-dependent molecular engineering of thermostability. As a first step, we established the overexpression of the *ppc* gene in *E. coli*.

The cloned *ppc* gene was successfully expressed in *E. coli* using a pET expression vector system (Fig. 4). This is the first example of an *R. obamensis* protein expressed in a mesophilic host. The enzymological properties and enzyme stability of the recombinant PEPC were basically similar to those of the wild-type enzyme. The minor changes observed were in optimum pH for activity and the effect of MgSO_4 on thermostability. From the sequence analysis of the recombinant plasmid pPEP2 several substitutions were found in the nucleotide sequence of the cloned *ppc* gene, but these were all synonymous substitutions and the amino acid sequence of the recombinant PEPC was completely conserved. This implied that the minor changes resulted not from the amino acid substitutions but from the mesophilic expression in *E. coli*, which may bring about the partially incomplete folding or incorrect local structure of the recombinant PEPC at relatively low temperatures. Another possible explanation is that some post-translational modifications of *R. obamensis* are involved in the correct folding and structure of native PEPC.

The phylogenetic analysis of PEPCs from various bacteria and plants indicated the interesting molecular evolution of the enzyme (Fig. 3). Based on the amino acid sequences of bacterial and eukaryotic PEPCs, Toh *et al.* (1994) and Lepiniec *et al.* (1994) have suggested a close relationship between proteobacterial and eukaryotic enzymes, and a bacterial origin of the eukaryotic PEPCs. In addition, recent phylogenetic studies based on central metabolic enzymes such as malate dehydrogenase, enolase and phosphoglycerate kinase have suggested a closer relationship between Bacteria and Eukarya than Archaea (Doolittle & Brown, 1994). Although no sequence information of archaeal PEPCs has yet been reported, we have studied several archaeal PEPCs from the hyperthermophilic methanogen *Methanothermobacter sociabilis* and thermoacidophilic *Sulfolobus* species (Sako *et al.*, 1996b; Sako *et al.*, 1997). In comparisons of molecular and enzymological properties, archaeal PEPCs were significantly different from bacterial and eukaryotic counterparts in their small size and allosteric properties. These results support the proposed close relationship between proteobacterial and eukaryotic enzymes and the bacterial origin of eukaryotic PEPCs. Sequence information for archaeal PEPCs is essential for understanding the molecular evolution of these enzymes and may also give new insight into their thermostability.

ACKNOWLEDGEMENTS

This work was supported in part by a Grant-in-Aid for Scientific Research (no. 07556048) from the Ministry of Education, Science and Culture of Japan.

REFERENCES

- Adams, M. W. W. (1993). Enzymes and proteins from organisms that grow near and above 100 °C. *Annu Rev Microbiol* **47**, 627–658.
- Adams, M. W. W. (1994). Biochemical diversity among sulfur-

dependent, hyperthermophilic microorganisms. *FEMS Microbiol Rev* 15, 261–277.

Adams, M. W. W., Perler, F. B. & Kelly, R. M. (1995). Extremozyme: expanding the limits of biocatalysis. *Bio/Technology* 13, 662–668.

Argos, P., Rosmann, M. G., Grau, U. M., Zuber, H., Frank, G. & Tratschin, J. D. (1979). Thermal stability and protein structure. *Biochemistry* 18, 5698–5703.

Arias, L. M. & Argos, P. (1989). Engineering protein thermal stability: sequence statistics point to residue substitutions in α -helices. *J Mol Biol* 206, 397–406.

Böhm, G. & Jaenicke, R. (1994). Relevance of sequence statistics for the properties of extremophilic proteins. *Int J Protein Res* 43, 97–106.

Bradford, M. M. (1976). A rapid and sensitive method for the quantitation of microgram quantities of protein utilizing the principle of protein–dye binding. *Anal Biochem* 72, 248–254.

Britton, K. L., Baker, P. J., Borges, K. M. M. & 7 other authors (1995). Insight into thermal stability from a comparison of the glutamate dehydrogenases from *Pyrococcus furiosus* and *Thermococcus litoralis*. *Eur J Biochem* 229, 688–695.

Brock, T. D. & Brock, M. L. (1984). Genus *Thermus* Brock and Freeze 1969, 295^{AL}. In *Bergey's Manual of Systematic Bacteriology*, vol. 1, pp. 333–337. Edited by N. R. Krieg & J. G. Holt. Baltimore: Williams & Wilkins.

Cannio, R., Rossi, M. & Bartolucci, S. (1994). A few amino acid substitutions are responsible for the higher thermostability of a novel NAD⁺-dependent bacillar alcohol dehydrogenase. *Eur J Biochem* 222, 345–352.

Cavagnero, S., Zhou, Z. H., Adams, M. W. W. & Chan, S. I. (1995). Response of rubredoxin from *Pyrococcus furiosus* to environmental changes: implications for the origin of hyperthermostability. *Biochemistry* 34, 9865–9873.

Chan, M. K., Mukund, S., Kletzin, A., Adams, M. W. W. & Rees, D. C. (1995). Structure of a hyperthermophilic tungstopterin enzyme, aldehyde ferredoxin oxidoreductase. *Science* 267, 1463–1469.

Chou, P. Y. & Fasman, G. D. (1978). Prediction of the secondary structure of proteins from their amino acid sequences. *Adv Enzymol Relat Areas Mol Biol* 47, 45–148.

Cowan, D. A. (1995). Protein stability at high temperatures. *Essays Biochem* 29, 193–207.

Daniel, R. M. (1996). The upper limits of enzyme thermal stability. *Enzyme Microb Technol* 19, 74–79.

Day, M. W., Hsu, B. T., Joshuator, L., Park, J.-B., Zhou, Z. H., Adams, M. W. W. & Rees, D. C. (1992). X-ray crystal structures of the oxidized and reduced forms of the rubredoxin from the marine hyperthermophilic archaeobacterium *Pyrococcus furiosus*. *Protein Sci* 1, 1494–1507.

Doolittle, W. F. & Brown, J. R. (1994). Tempo, mode, the progenote, and the universal root. *Proc Natl Acad Sci USA* 91, 6721–6728.

Eikmanns, B. J., Follettie, M. T., Griot, M. U. & Sinskey, A. J. (1989). The phosphoenolpyruvate carboxylase gene of *Corynebacterium glutamicum*. *Mol Gen Genet* 218, 330–339.

Fujita, N., Miwa, T., Ishijima, S., Izui, K. & Katsuki, H. (1984). The primary structure of phosphoenolpyruvate carboxylase of *Escherichia coli*. Nucleotide sequence of the *ppc* gene and deduced amino acid sequence. *J Biochem* 95, 909–916.

Ishijima, S., Katagiri, F., Kodaki, T., Izui, K., Katsuki, H., Nishikawa, K., Nakashima, H. & Ooi, T. (1985). Comparison of amino acid

sequences between phosphoenolpyruvate carboxylases from *Escherichia coli* (allosteric) and *Anacystis nidulans* (non-allosteric): identification of conserved and variable regions. *Biochem Biophys Res Commun* 133, 436–441.

Itaya, M. & Kondo, K. (1991). Molecular cloning of a ribonuclease H (RNase HI) gene from an extreme thermophile *Thermus thermophilus* HB8: a thermostable RNase H can functionally replace the *Escherichia coli* enzyme *in vivo*. *Nucleic Acids Res* 19, 4443–4449.

Jaenicke, R. (1996a). Glyceraldehyde-3-phosphate dehydrogenase from *Thermotoga maritima*: strategies of protein stabilization. *FEMS Microbiol Rev* 18, 215–224.

Jaenicke, R. (1996b). Stability and folding of ultrastable proteins: eye lens crystallins and enzymes from thermophiles. *FASEB J* 10, 84–92.

Kotsuka, T., Akamura, S., Tomuro, M., Yamagishi, A. & Oshima, T. (1996). Further stabilization of 3-isopropylmalate dehydrogenase of an extreme thermophile, *Thermus thermophilus*, by a suppressor mutation method. *J Bacteriol* 178, 723–727.

Laemmli, U. K. (1970). Cleavage of structural proteins during the assembly of the head of bacteriophage T4. *Nature* 227, 680–685.

Lepiniec, L., Vidal, J., Chollet, R., Gadai, P. & Cretin, C. (1994). Phosphoenolpyruvate carboxylase: structure, regulation and evolution. *Plant Sci* 99, 111–124.

Liu, Y. G. & Whittier, R. F. (1995). Thermal asymmetric interlaced PCR: automatable amplification and sequencing of insert end fragments from P1 and YAC clones for chromosome walking. *Genomics* 25, 674–681.

Nakamura, T., Yoshioka, I., Takahashi, M., Toh, H. & Izui, K. (1995). Cloning and sequence analysis of the gene for phosphoenolpyruvate carboxylase from an extreme thermophile, *Thermus* sp. *J Biochem* 118, 319–324.

Privalov, P. L. & Gill, S. J. (1988). Stability of protein structure and hydrophobic interaction. *Adv Protein Chem* 39, 191–234.

Regan, M. O., Thierbach, G., Bachmann, B., Vileeval, D., Lepage, P., Viret, J. F. & Lemoine, Y. (1989). Cloning and nucleotide sequence of the phosphoenolpyruvate carboxylase-coding gene of *Corynebacterium glutamicum* ATCC13032. *Gene* 77, 237–251.

Saitou, N. & Nei, M. (1987). The neighbor-joining method: a new method for reconstructing phylogenetic trees. *Mol Biol Evol* 4, 406–425.

Sako, Y., Takai, K., Uchida, A., Ishida, Y. & Katayama, Y. (1996a). *Rhodothermus obamensis* sp. nov., a modern lineage of extremely thermophilic marine bacteria. *Int J Syst Bacteriol* 46, 1099–1104.

Sako, Y., Takai, K., Uchida, A. & Ishida, Y. (1996b). Purification and characterization of phosphoenolpyruvate carboxylase from the hyperthermophilic archaeon *Methanothermus sociabilis*. *FEBS Lett* 392, 148–152.

Sako, Y., Takai, K., Nishizaka, T., Uchida, A. & Ishida, Y. (1997). Biochemical relationship of phosphoenolpyruvate carboxylases (PEPCs) from thermophilic archaea. *FEMS Microbiol Lett* 153, 159–165.

Sanger, F., Nicklen, S. & Coulson, A. R. (1977). DNA sequencing with chain-terminating inhibitors. *Proc Natl Acad Sci USA* 74, 5463–5467.

Schultes, V. & Jaenicke, R. (1991). Folding intermediates of hyperthermophilic D-glyceraldehyde-3-phosphate dehydrogenase from *Thermotoga maritima* are trapped at low temperature. *FEBS Lett* 290, 235–238.

Takai, K., Sako, Y., Uchida, A. & Ishida, Y. (1997a). Extremely thermostable phosphoenolpyruvate carboxylase from an extreme thermophile, *Rhodothermus obamensis*. *J Biochem* 122, 32–40.

Takai, K., Sako, Y. & Uchida, A. (1997b). Extrinsic thermostabilization factors and thermodenaturation mechanism of phosphoenolpyruvate carboxylase (PEPC) from an extremely thermophilic bacterium *Rhodothermus obamensis*. *J Ferment Bioeng* **84**, 291–299.

Tamakoshi, M., Yamagishi, A. & Oshima, T. (1995). Screening of stable proteins in an extreme thermophile, *Thermus thermophilus*. *Mol Microbiol* **16**, 1031–1036.

Toh, H., Kawamura, T. & Izui, K. (1994). Molecular evolution of phosphoenolpyruvate carboxylase. *Plant Cell Environ* **17**, 31–43.

Utter, M. F. & Kolenbrander, H. M. (1972). Formation of oxaloacetate by CO₂ fixation on phosphoenolpyruvate. In *The Enzymes*, pp. 117–168. Edited by P. D. Boyer. New York: Academic Press.

Zuber, H. (1988). Temperature adaptation of lactate dehydrogenase: structural, functional and genetic aspects. *Biophys Chem* **29**, 171–179.

Received 10 September 1997; revised 21 January 1998; accepted 30 January 1998.