Biochimica et Biophysica Acta 933 (1988) 141–155
Elsevier

BBA 42741

Sequence and over-expression of subunits of adenosine triphosphate synthase in thermophilic bacterium PS3

Shigeo Ohta [a], Masafumi Yohda [a], Morio Ishizuka [b], Hajime Hirata [a], Toshiro Hamamoto [a], Yohko Otawara-Hamamoto [a], Kakuko Matsuda [a] and Yasuo Kagawa [a]

[a] *Department of Biochemistry, Jichi Medical School, Minamikawachi, Tochigi-ken and* [b] *Department of Industrial Chemistry, Faculty of Science and Engineering, Chuo University, Bunkyo-ku, Tokyo (Japan)*

(Received 21 September 1987)

Key words: ATP synthase; Amino acid sequence; Nucleotide sequence; Subunit purification; Gene expression; F_0F_1; ATP synthase stability

The primary structures of all the subunits of thermophilic ATP synthase were determined, and its α, β and γ subunits could be over-expressed in *Escherichia coli*, because these subunits were stable and reconstitutable. DNA of 7500 base pairs in length was found to contain a cluster of nine genes for subunits of ATP synthase. The order of their reading frames (size in base pairs) was: I(381):a(630):c(216):b(489):δ(537):α(1507):γ(858):β(1419):ε(396), I being a gene for a small hydrophobic, basic protein expressed in vitro. All the termini of TF_0F_1 subunits were confirmed by peptide sequencing. Large quantities of the overexpressed thermophilic α, β and γ subunits were prepared from the extract of *E. coli*, by a few purification steps.

Introduction

ATP synthase (F_0F_1) plays a central role in oxidative phosphorylation [1,2]. F_0F_1 consists of a catalytic portion, called F_1 [3], and a proton channel portion, called F_0 [4,5]. The F_1 portion consists of five different subunits α, β, γ, δ and ε [6]. Three subunits of F_1, α, β and γ, are required for reconstituting ATPase activity, while the δ and ε subunits bind F_1 to F_0 [6]. F_0 consists of three different subunits a, b and c [4]. An operon coding for F_0F_1 of *Escherichia coli* has been determined [7,8]. A gene coding for a protein of 14 kDa (subunit I) was found in the first open reading frame of the operon, and the order of the reading frames was found to be I, a, c, b, δ, α, γ, β and ε [8]. Thermophilic F_0F_1 (TF_0F_1) is highly stable without ATP-Mg [9], and can be reconstituted into a planar lipid bilayer to measure the proton current directly [10]. Because of the high reconstitutability of TF_1 [11], site directed mutagenesis may be applied to each subunit and activities can be measured after reconstitution of TF_1 from the subunits. Physicochemical analyses of subunits usually require large quantities of the gene product. Thus, studies were made on the cloning [12], sequencing and over-expression of an operon for TF_0F_1. The sequences of its β [12] and ε [13] subunits have been reported.

In this work, the sequences of the remaining genes for I, a, c, b, δ, α and γ were determined

Abbreviations: F_0F_1, ATP synthase; TF_0F_1, thermophilic F_0F_1; F_1, catalytic portion of F_0F_1; TF_1, thermophilic F_1; EF_1, *Escherichi coli* F_1, F_0, proton channel portion of F_0F_1; TF_0, thermophilic F_0; EF_0, *Escherichia coli* F_0; SDS, sodium dodecylsulfate.

Correspondence: Y. Kagawa, Department of Biochemistry, Jichi Medical School, Minamikawachi-machi, Kawachi-gun, Tochigi-ken, Japan 329-04.

and the major subunits (α, β and γ) were purified by expressing the genes connected with a tac or lac promoter.

Materials and Methods

Materials. Nucleotides and other reagents for assay and preparation of ATP synthase were obtained as described in previous reports [9–13]. [^{32}P]nucleotides, [^{35}S]methionine, and a prokaryotic DNA-directed translation kit were purchased from Amersham International Co., Bucks, England. Restriction endonuclease and related enzymes were obtained from Takara Shuzo, Kyoto, Japan, and Nippon Gene, Toyama, Japan [12,13]. Oligonucleotides used as sequence primers ('17-mers') were synthesized with an automatic DNA synthesizer (Applied Biosystems, Model 380 B, CA, U.S.A.) and purified with a high-performance liquid chromatograph (Spectra-Physics, Model 8700, California, U.S.A.).

Bacterial strains and plasmids. Thermophilic bacterium PS3 was originally isolated from a sample from Mine hot spring and since 1972 its ATP synthase [9–13] and its operon [12,13] has been analyzed. *Escherichia coli* strain HB101 [14] was obtained from Dr. G. Schatz (Biozentrum, Basel) and strain NM522, from Pharmacia, Sweden. Plasmids with a promoter, pKK223-3 (tac promoter), pTZ18R/19R (lac promoter) and pUC18/19 (lac promoter) [15] were all obtained from Pharmacia Japan, Tokyo.

Manipulations of thermophilic DNA. The TF_0F_1 operon was cloned by gene walking using the fragment cloned as described previously [12]. The Sau 3AI (partial digestion), Sal I-Hind III, Stu I-Sac I and Dra I-Bal I fragments were cloned into pTZ18R or pTZ19R as shown in Fig. 1 (lines with arrows). The sequencing of the β [12] and ε [13] genes has been described. The manipulations of the above *E. coli* strains and plasmids were as described by us [12,13,16], Maniatis et al. [14], and Davis et al. [15]. The sequencing was repeated several times from both directions by the Sanger chain termination method [17] and uncertain portions were confirmed using unique oligonucleotide primers, 17 bases in length. Prokaryotic DNA-directed translation in vitro was performed by a modification of the method of Sancar et al. [18].

Purification and sequencing of wild-type TF_0F_1 subunits. All subunits of TF_1 were prepared as described previously [11]. TF_0 (2 mg) [5] was dissolved in 2% sodium dodecylsulfate and the mixture was fractionated by gel filtration high performance liquid chromatography [13] using G3000 SW (Toyo Soda, 7.5 mm × 60 cm) and an eluting solution (0.1% sodium dodecylsulfate, 10 mM sodium phosphate and 0.1 M NaCl) at the flow rate of 1 ml/min and fraction volume of 1 ml. The peaks (monitored at 280 nm) of TF_0 a, b and c subunits appeared at the elution volumes of 30 ml, 34 ml and 36 ml, respectively. The fraction containing TF_0 a subunit (with small amount of the b subunit) was applied to a column of Hi-Pore C-4 reverse phase (RP 304, Bio-Rad, 4.6 mm × 25 cm) equilibrated with 0.1% trifluoroacetic acid [13]. The material was eluted with a linear gradient of isopropanol (0–60%) in 0.1% trifluoroacetic acid, and TF_0 a subunit was eluted with 60% isopropanol (at 72 ml, flow rate 1 ml/min, 1 ml/fraction). TF_0 a subunit was treated with 1 M HCl at 37 °C, for 1.5 h to remove the formyl group before its sequencing.

The amino acid sequence of the N-terminal and several other portions of the subunits were determined using a peptide sequencer (Applied Biosystems 470A, program for 470 A). Some thermophilic subunits were denatured by heating at 100 °C for 1 h and digested with trypsin (1 μg/mg subunit, in 50 mM Tris-HCl, pH 8.0) at 37° for 24 h, or were cleaved with 25 mM BrCN at room temperature for 16 h, and the resulting peptides were isolated by high-performance liquid chromatography as described previously [13]. Protein concentration was measured by the method of Bradford [19]. Polyacrylamide gel electrophoresis with or without sodium dodecylsulfate was performed as reported previously [13].

Over-expression of TF_1 subunits in Escherichia coli. The α subunit gene was cloned into the Sal I-Hind III site of pTZ18R. The insert fragment (2.3 kilobase pairs) contained the Shine-Dalgarno sequence and the reading frame for the α subunit. The resulting plasmid, named pTZ-α, was introduced into *E. coli* HB101. The DNA fragment containing the β subunit gene was excised with Sac I and Sma I and the resulting fragment (1.6 kilo base pairs) was subcloned into the Sac I and

TABLE I

SUMMARY OF PURIFICATION OF THE OVER-EXPRESSED α-SUBUNIT

The procedure is described in the text and Fig. 5.

Step	Fraction	Volume (ml)	Protein (mg/ml)	Total protein (mg)
1-1	extract 1	50	7.1	355
1-2	extract 2	50	5.3	267
2	DEAE-fr.	70	5.4	378
3	hydrophobic column fr.	30	8.7	261

Sma I site of pUC18. The resulting plasmid, named pUC-β, was introduced into *E. coli* HB101. Since the initiation codon for the thermophilic γ subunit was found to be GTG by sequencing both the DNA and protein, its over-expression system was constructed by introducing an oligonucleotide containing the initiation codon, ATG by the following steps. A fragment containing the gene for the γ subunit was excised with Sal I (the Sal I site of the 5′ end was originated from the cloning site of the vector), and the resulting fragment (1.3 kilobase pairs) was subcloned into the Sal I site of pUC19. The 5′ end of the insert was digested with Bal 31 exonuclease to remove the initiation codon. After the subcloning of the shortened inset, the nucleotide sequence was determined to know the junction point of the insert and the vector. The oligonucleotides,

```
5' AATT C ATG GCA TCG TTA CGC GAT ATT AAA 3'
    3'  G TAC CGT AGC AAT GCG CTA TAA TTT 5'
```

which contained Eco RI ligation site, ATG and the deficient nucleotides, was ligated into the plasmid. The 9th threonine residues of the γ subunit was replaced by an arginine residue to facilitate connection, so the product was named the γE subunit. A fragment containing the γE gene was excised again and introduced into the Eco RI, Sal I site of pKK223-3 (pKK-γE, which contains the Shine-Dalgarno sequence). The expression vector thus constructed, pKKγE, was introduced into *E. coli* NM522. The three kinds of *E. coli* cells, harboring pTZα, pUCβ and pKKγE, respectively, were preincubated in 2 ml of 2 × TY containing 100 μg/ml ampicillin [14,15]. When the cells reached the stationary phase, the culture medium was replaced by 100 ml of 2 × TY medium with or without an inducer, isopropyl β-D thiogalactoside, at the specified concentration described in Results. Cell growth was monitored by measuring the optical density at 600 nm of the culture media, and subunit production was monitored by disrupting the cells and analyzing the resultant supernatant by gel electrophoresis.

Purification of subunits. The α subunit was purified as summarized in Table I. pTZα/HB101 cells (4.14 g per 1.5 l culture medium) were disrupted by sonication (Tomy Model UP150P, 6A, 150 s) in the presence of 50 ml of buffer containing 20 mM Tris SO_4 (pH 8.0), 5 mM NaCl and 0.1 mM EDTA. The homogenate was centrifugated at 10000 × *g* for 15 min, and the supernatant was collected as extract 1. The precipitate was resuspended in the same buffer and centrifuged to give extract 2. The α subunit in the resulting extract (combined extracts 1 and 2) amounted to nearly 80% of the total protein. This extract was subjected to DEAE-Sephacel chromatography, and the material eluted with 200–280 mM NaCl (pH 8.0) was fractionated with ammonium sulfate (1.4–2.1 M). This fraction was purified by hydrophobic chromatography as described in the legend of Fig. 5 (see results).

The β subunit was purified as summarized in Table II. pUC-β/HB101 cells (1.7 g) were disrupted by sonication as described above, extracted three times with 50 ml volumes of the same buffer

TABLE II

SUMMARY OF PURIFICATION OF THE OVER-EXPRESSED β SUBUNIT

The procedure is described in the text and Fig. 6.

Step	Fraction	Volume (ml)	Protein (mg/ml)	Total protein (mg)
1-1	extract 1	50	3.2	160
1-2	extract 2	50	5.7	285
1-3	extract 3	50	3.1	155
2	hydrophobic column fr.	50	2.0	100
3	DEAE fr.	70	1.0	70

as for extraction of the α subunit and fractionated by hydrophobic chromatography (0.8–0.6 M ammonium sulfate) and DEAE-Sephacel chromatography as described in the legend of Fig. 6 (see Results). The over-expressed γE subunit was also purified to determine whether change in the N-terminal region affected the activity. The extract obtained by sonication of pKK233-3/NM522 in the presence of 100 ml of 20 mM Tris-Cl (pH 7.2) with 1 mM $MgSO_4$ was applied to a column of DEAE-cellulose (DE 52, Whatman, 2.5 cm × 22 cm, 108 ml), the γE subunit was eluted with a linear gradient of NaCl (0 to 1 M, 1600 ml) and fractions eluted with 461–578 mM NaCl were collected.

TABLE III

MOLECULAR WEIGHTS OF SUBUNITS OF THERMOPHILIC ATP SYNTHASE

Molecular weights are based on the processed protein, not only on the DNA sequence.

Name of subunit	Molecular weight
14 K	14595.39
a	23482.80
c	7333.77
b	17267.68
δ	19657.31
α	54589.89
γ	31778.26
β	51937.58
ε	14333.46

Results

Structure of the operon and subunits of TF_0F_1

The order of genes in the TF_0F_1 operon is summarized in Fig. 1. The catalytic activity is detected in the αβγ complex, which forms the core structure of F_1 [20]. The nucleotide sequence of the operon and the primary structure of all the subunits of TF_0F_1 was deduced from both the complete DNA sequence and the partial peptide sequences, as shown in Fig. 2. There are promoter structures in the upstream region of the I gene. The role of the I gene is controversial [7,8] because its product (14 K protein) has not been found in TF_0F_1, or *E. coli* F_0F_1 preparations. As shown in Fig. 3, the I gene was transcribed and translated in vitro as a radioactive band of 14 kDa (an arrow in Fig. 3). The bands of lactamase and another reading frame in the upstream region were also detected. The thermophilic promoter was thus active and the I gene was expressed. The intercistronic non-coding regions were all longer than those of *E. coli*. The stable structure of the terminator (stem size 13) was described previously [13]. The molecular weights of all the subunits are shown in Table III.

Over-expression of the thermophilic genes

The genes for the TF_1 core subunits α, β and γ (Fig. 1) were well expressed in large quantities in the soluble fraction of *E. coli* (Fig. 4). For the γ subunit, *E. coli* NM522 was used, because it did not express the genes inserted in the vector unless an inducer, isopropyl β-thiogalactoside, was added, and thus a harmful effect of over-expressed gene products on the growth of host cell could be avoided. For production of the α and β subunits,

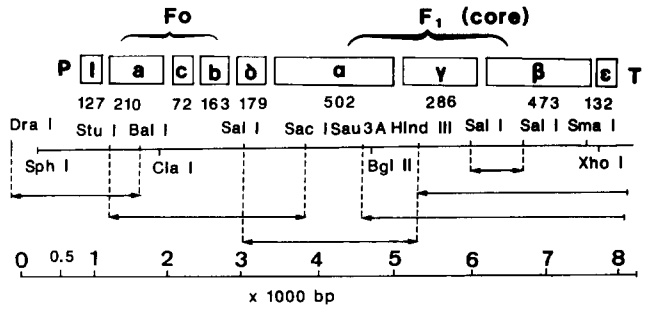

Fig. 1. Arrangements of genes encoding ATP synthase subunits in thermophilic bacterium PS3. The letters I, a, b, c, α, β, γ, δ and ε indicate the ATP synthase subunit encoded in the gene. P, promoter; T, terminator, Hind III, etc. are restriction sites.

```
*1                                                                    60*
GAA.AAG.GAA.AAT.AGG.CAT.AAA.AAT.AAA.ATA.TAT.CAT.TAT.ACA.TCA.TCA.TTG.TGA.GAA.GAA

*61                                                                   120*
AAT.GAG.GAA.ATT.GTC.ACA.AAT.TAT.TAA.TTT.GCC.CTA.TGT.GAA.TGA.GGT.TTT.ACG.TGG.AAA

*121                                                                  180*
CAC.ATT.GAC.ATT.GAC.CAG.GCC.GTA.TTG.TAT.CCT.TTA.CAA.GGG.GTA.CGA.TTG.AGA.AGG.TTT

*181                                                                  240*
TCA.TTA.TGT.TAA.AGG.GTG.TGT.TGG.CAT.GTC.CCC.AAA.ACA.CGC.CAC.CCG.TTT.CAA.GCC.ATG

*241                                                                  300*
GCG.TTG.CAT.GTC.CGC.CAT.CGT.CTC.CCA.ACT.TGT.TGG.TTC.GAT.TTT.GGT.CCG.GCG.TTT.TCG

*301                                                                  360*
GCG.GGA.GAT.GGA.TCG.ATG.ATC.GAT.TTG.GCA.CCG.AGC.CGA.TCT.TTT.TGA.TCG.TCG.GCC.TTT

*361                                                                  420*
TGC.TTG.CAC.TGG.CGG.CCG.GCG.TGT.ATG.CGA.TGT.TGC.GCT.TGA.TCC.GCC.AAT.ATT.TCT.CTG

*42S/D  I protein                                                     480*
GGG.AGC.AAT.AAA.ATG.GGA.AAT.CTT.CAA.GCG.ATG.TTT.TGG.CGG.CAA.GTT.CGA.TAC.ATA.TTG
                Met-Gly-Asn-Leu-Gln-Ala-Met-Phe-Trp-Arg-Gln-Val-Arg-Tyr-Ile-Leu
*481                                                                  540*
TAT.TTG.CTC.GCC.ATC.TAT.ACG.CTC.GGC.TTT.GGA.TTT.ACG.CCG.TAT.AAA.ACC.GTT.TTT.CTC
Tyr-Leu-Leu-Ala-Ile-Tyr-Thr-Leu-Gly-Phe-Gly-Phe-Thr-Pro-Tyr-Lys-Thr-Val-Phe-Leu
*541                                                                  600*
AGC.TTG.ATT.CTC.GGC.ACA.TCG.ATC.AGC.CTC.CTC.ATG.GTT.TGG.AAC.TTA.ACC.TGG.AAA.ATC
Ser-Leu-Ile-Leu-Gly-Thr-Ser-Ile-Ser-Leu-Leu-Met-Val-Trp-Asn-Leu-Thr-Trp-Lys-Ile
*601                                                                  660*
GAA.AAA.TTC.GGA.CAA.GCG.GTG.GCG.GCG.CGC.AAA.AAA.GTG.CGC.ACA.TTA.GGC.ACC.TTA.TCG
Glu-Lys-Phe-Gly-Gln-Ala-Val-Ala-Ala-Arg-Lys-Lys-Val-Arg-Thr-Leu-Gly-Thr-Leu-Ser
*661                                                                  720*
CGT.TTG.GCG.CTT.GCC.GCA.TTG.GCT.GCT.GTC.ATT.GTT.CTG.ACG.TAT.CCG.CAA.TAT.TTT.CAT
Arg-Leu-Ala-Leu-Ala-Ala-Leu-Ala-Ala-Val-Ile-Val-Leu-Thr-Tyr-Pro-Gln-Tyr-Phe-His
*721                                                                  780*
ATC.GTG.CCA.ACC.GTT.TTG.GGA.TTA.ATG.ACA.TCC.TAC.ATT.GTC.ATT.ATA.ATA.GAT.TTC.TTC
Ile-Val-Pro-Thr-Val-Leu-Gly-Leu-Met-Thr-Ser-Tyr-Ile-Val-Ile-Ile-Ile-Asp-Phe-Phe
*781                                                S/D     a subunit  839*
TTT.CAC.AAA.TGG.AAA.AAC.GAC.AAG.CTG.CAG.GCA.TGA. AA.GCG.AGG.TGA.GAG.AGG.ATG.GAG
Phe-His-Lys-Trp-Lys-Asn-Asp-Lys-Leu-Gln-Ala-***                      Met-Glu
*840                                                                  899*
CAT.AAA.GCG.CCG.CTT.GTC.GAA.TTT.TTA.GGC.CTT.ACG.TTT.AAT.TTA.TCT.GAT.ATG.TTG.ATG
His-Lys-Ala-Pro-Leu-Val-Glu-Phe-Leu-Gly-Leu-Thr-Phe-Asn-Leu-Ser-Asp-Met-Leu-Met
*900                                                                  959*
ATT.ACG.ATC.ACC.TGT.TTG.ATC.GTT.TTC.ATC.ATT.GCT.GTT.GCG.GCT.ACT.CGC.TCG.CTT.CAG
Ile-Thr-Ile-Thr-Cys-Leu-Ile-Val-Phe-Ile-Ile-Ala-Val-Ala-Ala-Thr-Arg-Ser-Leu-Gln
*960                                                                 1019*
CTG.CGC.CCG.ACG.GGG.ATG.CAA.AAC.TTT.ATG.GAG.TGG.GTC.TTT.GAC.TTT.GTT.AGA.GGC.ATC
Leu-Arg-Pro-Thr-Gly-Met-Gln-Asn-Phe-Met-Glu-Trp-Val-Phe-Asp-Phe-Val-Arg-Gly-Ile
*1020                                                                1079*
ATT.AAC.AGC.ACG.ATG.GAT.TGG.CAA.ACG.GGC.GGC.CGC.TTT.TTG.ACG.CTG.GGG.GTC.ACG.CTC
Ile-Asn-Ser-Thr-Met-Asp-Trp-Gln-Thr-Gly-Gly-Arg-Phe-Leu-Thr-Leu-Gly-Val-Thr-Leu
```

Fig. 2. Nucleotide sequence of thermophilic bacterium PS3 DNA encoding ATP synthase subunits. S/D: Shine-Dalgarno sequence (boxed). The underlined portions are the peptide sequences determined directly. The candidates for the promoter are shown by the doubled lines.

```
*1080                                                                1139*
ATC.ATG.TAT.GTG.TTT.GTG.GCC.AAT.ATG.CTT.GGA.CTG.CCG.TTT.TAT.GGT.CGT.CGC.CCT.CAC
Ile-Met-Tyr-Val-Phe-Val-Ala-Asn-Met-Leu-Gly-Leu-Pro-Phe-Tyr-Gly-Arg-Arg-Pro-His
*1140                                                                1199*
CCA.CTA.CTA.TGG.GGT.AAA.ATA.AAG.GGG.TTC.CAC.ATT.TTG.CGC.ATT.ACA.CCC.GTC.CAG.TCG
Pro-Leu-Leu-Trp-Gly-Lys-Ile-Lys-Gly-Phe-His-Ile-Leu-Arg-Ile-Thr-Pro-Val-Gln-Ser
*1200                                                                1259*
CCT.GGG.TCG.TTC.CCG.CTT.AAA.ATC.ATC.GAA.GAG.TTT.GCC.AAC.ACG.CTG.ACG.CTC.GGT.TTG
Pro-Gly-Ser-Phe-Pro-Leu-Lys-Ile-Ile-Glu-Glu-Phe-Ala-Asn-Thr-Leu-Thr-Leu-Gly-Leu
*1260                                                                1319*
CGT.CTT.TTC.GGG.AAC.ATT.TAC.GCC.GGG.GAA.ATT.TTG.CTT.GGC.TTG.CTC.GCC.AGC.CTT.GGC
Arg-Leu-Phe-Gly-Asn-Ile-Tyr-Ala-Gly-Glu-Ile-Leu-Leu-Gly-Leu-Leu-Ala-Ser-Leu-Gly
*1320                                                                1379*
ACG.CAT.TAC.GGT.GTG.CTT.GGC.GCT.GTT.GGC.GCG.AGC.CAG.TTT.CCG.ATC.ATG.GTG.TGG.CAA
Thr-His-Tyr-Gly-Val-Leu-Gly-Ala-Val-Gly-Ala-Ser-Gln-Phe-Pro-Ile-Met-Val-Trp-Gln
*1380                                                                1439*
GCG.TTC.AGT.ATT.TTT.GTC.GGA.ACG.ATT.CAG.GCG.TTC.ATT.TTT.ACG.ATG.TTA.ACC.ATG.GTT
Ala-Phe-Ser-Ile-Phe-Val-Gly-Thr-Ile-Gln-Ala-Phe-Ile-Phe-Thr-Met-Leu-Thr-Met-Val
*1440                                                                1499*
TAT.ATG.CCT.CAT.AAG.GTC.AAG.TCA.TGA.CCA.TTG.AGC.ATT.CCC.TTA.TCA.TTT.CAT.TAA.GGA
Tyr-Met-Pro-His-Lys-Val-Lys-Ser-***
                                  S/D          c subunit
*1500                                                                1559*
CTT.GAA.ACA.TTA.CTT.TAA.CAT.GGT.GAA.GGA.GGA.TCT.ATC.AAC.ATG.AGT.TTG.GGT.GTA.CTT
                                                 Met-Ser-Leu-Gly-Val-Leu
*1560                                                                1619*
GCA.GCT.GCG.ATT.GCG.GTA.GGT.TTG.GGG.GCG.TTA.GGC.GCC.GGC.ATC.GGC.AAC.GGG.TTG.ATC
Ala-Ala-Ala-Ile-Ala-Val-Gly-Leu-Gly-Ala-Leu-Gly-Ala-Gly-Ile-Gly-Asn-Gly-Leu-Ile
*1620                                                                1679*
GTC.AGC.CGT.ACC.ATT.GAA.GGG.ATT.GCT.CGT.CAA.CCA.GAA.TTG.CGT.CCG.GTT.TTG.CAA.ACG
Val-Ser-Arg-Thr-Ile-Glu-Gly-Ile-Ala-Arg-Gln-Pro-Glu-Leu-Arg-Pro-Val-Leu-Gln-Thr
*1680                                                                1739*
ACG.ATG.TTC.ATC.GGG.GTT.GCG.TTG.GTT.GAG.GCG.CTT.CCG.ATC.ATC.GGT.GTC.GTC.TTC.TCG
Thr-Met-Phe-Ile-Gly-Val-Ala-Leu-Val-Glu-Ala-Leu-Pro-Ile-Ile-Gly-Val-Val-Phe-Ser
*1740                                                                1797*
TTC.ATT.TAC.TTA.GGT.CGA.TAA.  A.TCG.ATA.TGG.AAA.GTG.ATG.GCG.AAG.TTC.GTG.TCA.ACG
Phe-Ile-Tyr-Leu-Gly-Arg-***
*1798                                                                1857*
ACC.ACT.CGC.CAT.TCC.TTT.ATG.TGC.GGT.TTT.GCG.AGA.CCG.CAT.CGG.TGC.GAT.ACG.TGT.CGA

*1858 S/D   b subunit                                                1917*
AAG.GAG.TGA.AAC.GCG.GTG.TTG.TGG.AAG.GCA.AAC.GTA.TGG.GTG.CTC.GGC.GAA.GCG.GCG.CAT
                    Met-Leu-Trp-Lys-Ala-Asn-Val-Trp-Val-Leu-Gly-Glu-Ala-Ala-His
*1918                                                  processing    1977*
GGC.ATC.AGC.GGC.GGC.ACG.ATT.ATT.TAC.CAA.TTG.CTG.ATG.TTC.ATT.ATT.TTG.CTG.GCC.TTG
Gly-Ile-Ser-Gly-Gly-Thr-Ile-Ile-Tyr-Gln-Leu-Leu-Met-Phe-Ile-Ile-Leu-Leu-Ala-Leu
*1978                                                                2037*
CTG.CGC.AAA.TTC.GCT.TGG.CAG.CCG.TTA.ATG.AAT.ATC.ATG.AAA.CAA.CGT.GAA.GAA.CAT.ATC
Leu-Arg-Lys-Phe-Ala-Trp-Gln-Pro-Leu-Met-Asn-Ile-Met-Lys-Gln-Arg-Glu-Glu-His-Ile
*2038                                                                2097*
GCG.ACG.AAA.TCG.ACC.AGG.CGG.AAA.AAC.GAC.CGT.CAA.GAA.GCA.GAA.AAA.CTT.CTT.GAA.GAA
Ala-Thr-Lys-Ser-Thr-Arg-Arg-Lys-Asn-Asp-Arg-Gln-Glu-Ala-Glu-Lys-Leu-Leu-Glu-Glu
*2098                                                                2157*
CAG.CGC.GAA.CTG.ATG.AAG.CAG.TCG.CGC.CAA.GAG.GCG.CAA.GCG.CTC.ATT.GAA.AAC.GCG.GCA
Gln-Arg-Glu-Leu-Met-Lys-Gln-Ser-Arg-Gln-Glu-Ala-Gln-Ala-Leu-Ile-Glu-Asn-Ala-Ala
```

```
*2158                                                                2217*
AGC.CTG.GCT.GAA.GAG.CAG.AAA.GAA.CAA.ATT.GTC.GCC.TCG.GCC.CGT.GCG.GAA.GCG.GAA.CGG
Ser-Leu-Ala-Glu-Glu-Gln-Lys-Glu-Gln-Ile-Val-Ala-Ser-Ala-Arg-Ala-Glu-Ala-Glu-Arg
*2218                                                                2277*
GTG.AAA.GAA.GCG.GCG.AAA.AAA.GAA.ATC.GAG.CGT.GAA.AAA.GAA.CAG.GCG.ATG.GCC.GCG.CTC
Val-Lys-Glu-Ala-Ala-Lys-Lys-Glu-Ile-Glu-Arg-Glu-Lys-Glu-Gln-Ala-Met-Ala-Ala-Leu
*2278                                                                2337*
CGC.GAA.CAA.GTG.GCC.TCG.TTG.TCT.GTC.TTG.ATC.GCC.TCA.AAA.GTG.ATT.GAA.AAA.GAG.CTG
Arg-Glu-Gln-Val-Ala-Ser-Leu-Ser-Val-Leu-Ile-Ala-Ser-Lys-Val-Ile-Glu-Lys-Glu-Leu
*2338                                                                2395*
ACC.GAG.CAA.GAT.CAA.GCG.GCA.AGC.TGA.  T.CGA.AGC.GTA.CAT.CAA.AGA.CGT.TCA.AGA.GGC
Thr-Glu-Gln-Asp-Gln-Ala-Ala-Ser-***
*2396 S/D      δ subunit                                             2455*
AGG.AGG.AGC.GCG.ATG.AAC.CAA.GAA.GTG.ATC.GCC.AAA.CGG.TAC.GCG.TCC.GCT.TTG.TTC.CAA
                Met-Asn-Gln-Glu-Val-Ile-Ala-Lys-Arg-Tyr-Ala-Ser-Ala-Leu-Phe-Gln
*2456                                                                2515*
ATC.GCG.CTC.GAA.CAA.GGA.CAG.CTT.GAT.CGA.ATC.GAA.GAA.GAT.GTT.CGC.GCC.GTC.CGC.CAG
Ile-Ala-Leu-Glu-Gln-Gly-Gln-Leu-Asp-Arg-Ile-Glu-Glu-Asp-Val-Arg-Ala-Val-Arg-Gln
*2516                                                                2575*
GCG.TTG.GCG.GAA.AAC.GGC.GAG.TTT.TTA.TCG.CTT.CTT.TCG.TAT.CCG.AAA.CTT.TCC.TTG.GAT
Ala-Leu-Ala-Glu-Asn-Gly-Glu-Phe-Leu-Ser-Leu-Leu-Ser-Tyr-Pro-Lys-Leu-Ser-Leu-Asp
*2576                                                                2635*
CAG.AAA.AAA.GCG.CTC.ATC.CGC.GAA.GCG.TTC.GCC.GGT.GTG.TCC.ACC.CCG.GTG.CAA.AAC.ACG
Gln-Lys-Lys-Ala-Leu-Ile-Arg-Glu-Ala-Phe-Ala-Gly-Val-Ser-Thr-Pro-Val-Gln-Asn-Thr
*2636                                                                2695*
CTT.CTC.CTT.CTT.CTG.GAG.CGC.CAT.CGC.TTC.GGC.CTT.GTG.CCC.GAG.CTG.GCT.GGA.ACA.GTT
Leu-Leu-Leu-Leu-Leu-Glu-Arg-His-Arg-Phe-Gly-Leu-Val-Pro-Glu-Leu-Ala-Gly-Thr-Val
*2696                                                                2755*
TCT.CGC.CCT.CGC.TCG.ACG.ACT.GCG.CGC.GGC.ATC.GCC.AAG.GCG.GTC.GCC.TAT.TCG.GGC.GCG
Ser-Arg-Pro-Arg-Ser-Thr-Thr-Ala-Arg-Gly-Ile-Ala-Lys-Ala-Val-Ala-Tyr-Ser-Gly-Ala
*2756                                                                2815*
GCG.TCG.ACG.GAC.GAA.GAA.CTG.CGG.GCG.CTT.TCC.GAC.GTC.TTT.GCC.CAA.AAA.GTC.GGC.AAA
Ala-Ser-Thr-Asp-Glu-Glu-Leu-Arg-Ala-Leu-Ser-Asp-Val-Phe-Ala-Gln-Lys-Val-Gly-Lys
*2816                                                                2875*
CAG.ACG.CTT.GAG.ATT.GAA.AAT.ATC.ATT.GAT.CCG.GAA.CTC.ATC.GGC.GGC.GTG.AAC.GTG.CGC
Gln-Thr-Leu-Glu-Ile-Glu-Asn-Ile-Ile-Asp-Pro-Glu-Leu-Ile-Gly-Gly-Val-Asn-Val-Arg
*2876                                                                2935*
ATC.GGC.AAC.CGC.ATT.TAC.GAC.GGC.AGC.GTC.AGC.GGG.CAG.CTT.GAA.CGG.ATT.CGG.CGG.CAG
Ile-Gly-Asn-Arg-Ile-Tyr-Asp-Gly-Ser-Val-Ser-Gly-Gln-Leu-Glu-Arg-Ile-Arg-Arg-Gln
*2936                                S/D          α subunit          2995*
CTG.ATC.GGC.TAA.CAT.TTG.AAG.ACA.GGG.GTG.AAA.GGC.ATG.AGC.ATT.CGA.GCG.GAA.GAA.ATC
Leu-Ile-Gly-***                                  Met-Ser-Ile-Arg-Ala-Glu-Glu-Ile
*2996                                                                3055*
AGC.GCG.CTC.ATT.AAG.CAG.CAG.ATT.GAA.AAC.TAT.GAA.TCG.CAA.ATC.CAA.GTG.AGC.GAC.GTC
Ser-Ala-Leu-Ile-Lys-Gln-Gln-Ile-Glu-Asn-Tyr-Glu-Ser-Gln-Ile-Gln-Val-Ser-Asp-Val
*3056                                                                3115*
GGC.ACC.GTC.ATC.CAA.GTC.GGC.GAC.GGG.ATC.GCG.CGC.GCT.CAT.GGG.CTC.GAT.AAC.GTC.ATG
Gly-Thr-Val-Ile-Gln-Val-Gly-Asp-Gly-Ile-Ala-Arg-Ala-His-Gly-Leu-Asp-Asn-Val-Met
*3116                                                                3175*
TCC.GGC.GAG.GCT.GTC.GAG.TTT.GCC.AAC.GCG.GTG.ATG.GGC.ATG.GCG.TTG.AAC.TTG.GAA.GAA
Ser-Gly-Glu-Ala-Val-Glu-Phe-Ala-Asn-Ala-Val-Met-Gly-Met-Ala-Leu-Asn-Leu-Glu-Glu
*3176                                                                3235*
AAC.AAC.GTC.GGT.ATC.GTT.ATT.TTA.GGA.CCG.TAC.ACC.GGC.ATT.AAA.GAA.GGG.GAC.GAA.GTG
Asn-Asn-Val-Gly-Ile-Val-Ile-Leu-Gly-Pro-Tyr-Thr-Gly-Ile-Lys-Glu-Gly-Asp-Glu-Val
```

```
*3236                                                                3295*
CGC.CGT.ACG.GGC.CGG.ATT.ATG.GAA.GTG.CCC.GTT.GGG.GAA.ACC.CTC.ATC.GGC.CGC.GTC.GTC
Arg-Arg-Thr-Gly-Arg-Ile-Met-Glu-Val-Pro-Val-Gly-Glu-Thr-Leu-Ile-Gly-Arg-Val-Val
*3296                                                                3355*
AAC.CCA.CTC.GGT.CAG.CCT.GTT.GAC.GGA.TTA.GGG.CCG.GTG.GAA.ACG.ACG.GAA.ACG.CGC.CCG
Asn-Pro-Leu-Gly-Gln-Pro-Val-Asp-Gly-Leu-Gly-Pro-Val-Glu-Thr-Thr-Glu-Thr-Arg-Pro
*3356                                                                3415*
ATT.GAA.AGC.CGT.GCG.CCG.GGC.GTT.ATG.GAC.CGG.AGA.TCG.GTG.CAT.GAG.CCG.CTG.CAA.ACC
Ile-Glu-Ser-Arg-Ala-Pro-Gly-Val-Met-Asp-Arg-Arg-Ser-Val-His-Glu-Pro-Leu-Gln-Thr
*3416                                                                3475*
GGG.ATT.AAA.GCG.ATC.GAC.GCG.CTC.GTG.CCG.ATC.GGC.CGC.GGG.CAG.CGC.GAG.CTC.ATC.ATC
Gly-Ile-Lys-Ala-Ile-Asp-Ala-Leu-Val-Pro-Ile-Gly-Arg-Gly-Gln-Arg-Glu-Leu-Ile-Ile
*3476                                                                3535*
GGC.GAC.CGA.CAA.ACG.GGG.AAA.ACG.TCC.GTC.GCC.ATT.GAC.ACG.ATC.ATC.AAC.CAA.AAA.GAC
Gly-Asp-Arg-Gln-Thr-Gly-Lys-Thr-Ser-Val-Ala-Ile-Asp-Thr-Ile-Ile-Asn-Gln-Lys-Asp
*3536                                                                3595*
CAA.AAC.ATG.ATT.TGT.ATT.TAT.GTC.GCC.ATC.GGG.CAA.AAA.GAA.TCG.ACG.GTC.GCA.ACC.GTT
Gln-Asn-Met-Ile-Cys-Ile-Tyr-Val-Ala-Ile-Gly-Gln-Lys-Glu-Ser-Thr-Val-Ala-Thr-Val
*3596                                                                3655*
GTC.GAA.ACG.CTC.GCA.AAA.CAC.GGC.GCG.CCT.GAC.TAT.ACG.ATC.GTC.GTC.ACC.GCT.TCG.GCG
Val-Glu-Thr-Leu-Ala-Lys-His-Gly-Ala-Pro-Asp-Tyr-Thr-Ile-Val-Val-Thr-Ala-Ser-Ala
*3656                                                                3715*
TCG.CAG.CCG.GCT.CCG.CTT.TTG.TTC.TTG.GCG.CCG.TAT.GCC.GGT.GTA.GCG.ATG.GGC.GAG.TAC
Ser-Gln-Pro-Ala-Pro-Leu-Leu-Phe-Leu-Ala-Pro-Tyr-Ala-Gly-Val-Ala-Met-Gly-Glu-Tyr
*3716                                                                3775*
TTC.ATG.ATA.ATG.GGC.AAG.CAC.GTT.TTG.GTT.GTG.ATC.GAC.GAT.TTA.TCG.AAG.CAG.GCC.GCG
Phe-Met-Ile-Met-Gly-Lys-His-Val-Leu-Val-Val-Ile-Asp-Asp-Leu-Ser-Lys-Gln-Ala-Ala
*3776                                                                3835*
GCA.TAC.CGG.CAA.TTG.TCG.CTC.TTG.CTT.CGC.CGT.CCG.CCG.GGC.CGT.GAA.GCG.TAT.CCG.GGG
Ala-Tyr-Arg-Gln-Leu-Ser-Leu-Leu-Leu-Arg-Arg-Pro-Pro-Gly-Arg-Glu-Ala-Tyr-Pro-Gly
*3836                                                                3895*
GAT.ATT.TTC.TAC.TTG.CAC.TCC.CGC.CTG.CTT.GAG.CGC.GCA.GCA.AAA.TTG.AGC.GAT.GCC.AAA
Asp-Ile-Phe-Tyr-Leu-His-Ser-Arg-Leu-Leu-Glu-Arg-Ala-Ala-Lys-Leu-Ser-Asp-Ala-Lys
*3896                                                                3955*
GGC.GGC.GGT.TCG.TTG.ACC.GCG.CTT.CCG.TTC.GTC.GAA.ACG.CAA.GCG.GGC.GAC.ATT.TCC.GCC
Gly-Gly-Gly-Ser-Leu-Thr-Ala-Leu-Pro-Phe-Val-Glu-Thr-Gln-Ala-Gly-Asp-Ile-Ser-Ala
*3956                                                                4015*
TAC.ATT.CCG.ACG.AAC.GTC.ATC.TCG.ATT.ACG.GAC.GGG.CAA.ATT.TTC.TTG.CAA.TCG.GAC.TTG
Tyr-Ile-Pro-Thr-Asn-Val-Ile-Ser-Ile-Thr-Asp-Gly-Gln-Ile-Phe-Leu-Gln-Ser-Asp-Leu
*4016                                                                4075*
TTC.TTC.TCC.GGC.GTC.CGA.CCA.GCG.ATC.AAC.GCA.GGG.TTG.TCC.GTT.TCG.CGC.GTC.GGT.GGG
Phe-Phe-Ser-Gly-Val-Arg-Pro-Ala-Ile-Asn-Ala-Gly-Leu-Ser-Val-Ser-Arg-Val-Gly-Gly
*4076                                                                4135*
GCA.GCG.CAA.ATC.AAA.GCG.ATG.AAA.AAA.GTA.GCC.GGG.ACG.CTT.CGT.TTG.GAC.TTG.GCC.GCT
Ala-Ala-Gln-Ile-Lys-Ala-Met-Lys-Lys-Val-Ala-Gly-Thr-Leu-Arg-Leu-Asp-Leu-Ala-Ala
*4136                                                                4195*
TAC.CGT.GAG.CTC.GAA.GCG.TTC.GCC.CAA.TTC.GGC.TCC.GAC.CTC.GAT.AAA.GCG.ACG.CAG.GCG
Tyr-Arg-Glu-Leu-Glu-Ala-Phe-Ala-Gln-Phe-Gly-Ser-Asp-Leu-Asp-Lys-Ala-Thr-Gln-Ala
*4196                                                                4255*
AAC.GTC.GCC.CGC.GGG.GCG.CGC.ACG.GTC.GAA.GTG.CTG.AAG.CAA.GAT.TTG.CAT.CAG.CCG.ATT
Asn-Val-Ala-Arg-Gly-Ala-Arg-Thr-Val-Glu-Val-Leu-Lys-Gln-Asp-Leu-His-Gln-Pro-Ile
*4256                                                                4315*
CCG.GTC.GAA.AAA.CAA.GTG.TTG.ATC.ATC.TAT.GCA.TTG.ACG.CGC.GGC.TTT.TTG.GAC.GAC.ATT
Pro-Val-Glu-Lys-Gln-Val-Leu-Ile-Ile-Tyr-Ala-Leu-Thr-Arg-Gly-Phe-Leu-Asp-Asp-Ile
```

```
*4316                                                                4375*
CCG.GTT.GAA.GAT.GTG.CGC.CGT.TTC.GAG.AAA.GAG.TTT.TAC.TTG.TGG.CTC.GAC.CAA.AAC.GGC
Pro-Val-Glu-Asp-Val-Arg-Arg-Phe-Glu-Lys-Glu-Phe-Tyr-Leu-Trp-Leu-Asp-Gln-Asn-Gly
*4376                                                                4435*
CAA.CAC.TTG.CTT.GAG.CAC.ATC.CGC.ACG.ACG.AAA.GAT.CTT.CCG.AAC.GAA.GAC.GAT.CTC.AAT
Gln-His-Leu-Leu-Glu-His-Ile-Arg-Thr-Thr-Lys-Asp-Leu-Pro-Asn-Glu-Asp-Asp-Leu-Asn
*4436                                                                4495*
CAA.GCG.ATC.GAA.GCG.TTC.AAG.AAA.ACG.TTT.GTC.GTT.TCT.CAA.TAA.GGC.CGG.CGG.GGG.CAT
Gln-Ala-Ile-Glu-Ala-Phe-Lys-Lys-Thr-Phe-Val-Val-Ser-Gln-***   γ subunit
*4496                                  S/D                           4555*
ATC.TGC.CCT.CAT.GCC.TGT.TAG.AAC.CCG.TGC.AGC.AAA.AAG.GTG.GTG.AAA.CCT.TTG.GCA.TCG
                                                  Met-Lys-Pro-Leu-Ala-Ser
*4556                                                  processing  4615
TTA.CGC.GAT.ATT.AAA.ACG.CGC.ATC.AAT.GCG.ACG.AAG.AAG.ACA.AGC.CAA.ATT.ACA.AAA.GCG
Leu-Arg-Asp-Ile-Lys-Thr-Arg-Ile-Asn-Ala-Thr-Lys-Lys-Thr-Ser-Gln-Ile-Thr-Lys-Ala
*4616                                                                4675*
ATG.GAA.ATG.GTC.CTC.ACG.TCG.AAG.CTG.AAC.CGC.GCG.GAA.AAA.CGC.GAA.ATC.GTT.CGT.CCA
Met-Glu-Met-Val-Leu-Thr-Ser-Lys-Leu-Asn-Arg-Ala-Glu-Lys-Arg-Glu-Ile-Val-Arg-Pro
*4676                                                                4735*
TAT.ATG.GAG.AAA.ATT.CAA.GAA.GTC.GTG.GCT.AAT.GTG.GCG.CTT.GCC.GCG.CGC.GCT.TCA.CAT
Tyr-Met-Glu-Lys-Ile-Gln-Glu-Val-Val-Ala-Asn-Val-Ala-Leu-Ala-Ala-Arg-Ala-Ser-His
*4736                                                                4795*
CCG.ATG.CTC.GTT.TCG.CGC.CCG.GTG.AAA.AAA.ACC.GGT.TAT.CTC.GTG.ATC.ACG.TCG.GAT.CGC
Pro-Met-Leu-Val-Ser-Arg-Pro-Val-Lys-Lys-Thr-Gly-Tyr-Leu-Val-Ile-Thr-Ser-Asp-Arg
*4796                                                                4855*
GGT.CTG.GCT.GGC.GCG.TAC.AAC.AGC.AAC.GTG.CTG.CGC.CTC.GTG.TAC.CAA.ACG.ATC.CAA.AAA
Gly-Leu-Ala-Gly-Ala-Tyr-Asn-Ser-Asn-Val-Leu-Arg-Leu-Val-Tyr-Gln-Thr-Ile-Gln-Lys
*4856                                                                4915*
CGC.CAT.GCT.TCT.CCG.GAC.GAA.TAT.GCG.ATC.ATC.GTC.ATC.GGC.CGC.GTC.GGG.TTG.AGC.TTT
Arg-His-Ala-Ser-Pro-Asp-Glu-Tyr-Ala-Ile-Ile-Val-Ile-Gly-Arg-Val-Gly-Leu-Ser-Phe
*4916                                                                4975*
TTC.CGC.AAG.CGG.AAC.ATG.CCG.GTC.ATT.CTC.GAC.ATC.ACC.CGC.TTG.CCG.GAC.CAG.CCG.TCG
Phe-Arg-Lys-Arg-Asn-Met-Pro-Val-Ile-Leu-Asp-Ile-Thr-Arg-Leu-Pro-Asp-Gln-Pro-Ser
*4976                                                                5035*
TTT.GCC.GAT.ATT.AAA.GAA.ATC.GCC.CGC.AAA.ACG.GTT.GGG.TTA.TTC.GCC.GAC.GGT.ACG.TTT
Phe-Ala-Asp-Ile-Lys-Glu-Ile-Ala-Arg-Lys-Thr-Val-Gly-Leu-Phe-Ala-Asp-Gly-Thr-Phe
*5036                                                                5095*
GAC.GAG.CTG.TAT.ATG.TAT.TAC.AAC.CAT.TAC.GTG.AGC.GCG.ATC.CAG.CAA.GAG.GTG.ACG.GAA
Asp-Glu-Leu-Tyr-Met-Tyr-Tyr-Asn-His-Tyr-Val-Ser-Ala-Ile-Gln-Gln-Glu-Val-Thr-Glu
*5096                                                                5155*
CGG.AAG.CTT.CTG.CCG.CTC.ACT.GAC.TTG.GCG.GAG.AAT.AAG.CAA.CGC.ACG.GTG.TAC.GAA.TTT
Arg-Lys-Leu-Leu-Pro-Leu-Thr-Asp-Leu-Ala-Glu-Asn-Lys-Gln-Arg-Thr-Val-Tyr-Glu-Phe
*5156                                                                5215*
GAA.CCG.TCG.CAA.GAA.GAA.ATT.TTG.GAC.GTC.TTA.TTG.CCG.CAG.TAT.GCG.GAA.AGC.CTC.ATT
Glu-Pro-Ser-Gln-Glu-Glu-Ile-Leu-Asp-Val-Leu-Leu-Pro-Gln-Tyr-Ala-Glu-Ser-Leu-Ile
*5216                                                                5275*
TAC.GGC.GCA.TTG.CTC.GAT.GCA.AAA.GCA.AGC.GAA.CAC.GCC.GCC.CGC.ATG.ACG.GCG.ATG.AAG
Tyr-Gly-Ala-Leu-Leu-Asp-Ala-Lys-Ala-Ser-Glu-His-Ala-Ala-Arg-Met-Thr-Ala-Met-Lys
*5276                                                                5335*
AAC.GCA.ACG.GAC.AAT.GCG.AAC.GAG.CTC.ATT.CGC.ACA.TTG.ACG.CTT.TCC.TAC.AAC.CGC.GCT
Asn-Ala-Thr-Asp-Asn-Ala-Asn-Glu-Leu-Ile-Arg-Thr-Leu-Thr-Leu-Ser-Tyr-Asn-Arg-Ala
*5336                                                                5395*
CGC.CAA.GCG.GCG.ATT.ACG.CAA.GAA.ATT.ACG.GAA.ATT.GTC.GCC.GGA.GCA.AAC.GCC.TTG.CAA
Arg-Gln-Ala-Ala-Ile-Thr-Gln-Glu-Ile-Thr-Glu-Ile-Val-Ala-Gly-Ala-Asn-Ala-Leu-Gln
```

5396 S/D β subunit ⟶ 5455
TAG.GGC.ACT.AGC.AAG.TTA.GGA.GGG.AAA.ACG.ATG.ACA.AGA.GGA.CGC.GTT.ATC.CAA.GTC.ATG
*** Met-Thr-Arg-Gly-Arg-Val-Ile-Gln-Val-Met
5456 5515
GGT.CCG.GTT.GTA.GAC.GTC.AAG.TTT.GAG.AAC.GGC.CAC.TTG.CCG.GCG.ATC.TAC.AAC.GCC.CTG
Gly-Pro-Val-Val-Asp-Val-Lys-Phe-Glu-Asn-Gly-His-Leu-Pro-Ala-Ile-Tyr-Asn-Ala-Leu
5516 5575
AAA.ATT.CAA.CAT.AAA.GCG.CGC.AAC.GAA.AAC.GAA.GTC.GAC.ATC.GAC.TTG.ACA.TTG.GAA.GTC
Lys-Ile-Gln-His-Lys-Ala-Arg-Asn-Glu-Asn-Glu-Val-Asp-Ile-Asp-Leu-Thr-Leu-Glu-Val
5576 5635
GCC.TTG.CAC.CTT.GGC.GAT.GAT.ACA.GTA.CGG.ACG.ATC.GCG.ATG.GCG.TCC.ACA.GAC.GGC.CTC
Ala-Leu-His-Leu-Gly-Asp-Asp-Thr-Val-Arg-Thr-Ile-Ala-Met-Ala-Ser-Thr-Asp-Gly-Leu
5636 5695
ATC.CGC.GGC.ATG.GAA.GTT.ATC.GAT.ACC.GGT.GCA.CCG.ATT.TCG.GTG.CCG.GTC.GGG.CAA.GTC
Ile-Arg-Gly-Met-Glu-Val-Ile-Asp-Thr-Gly-Ala-Pro-Ile-Ser-Val-Pro-Val-Gly-Gln-Val
5696 5755
ACG.CTT.GGC.CGC.GTG.TTC.AAC.GTC.TTG.GGC.GAG.CCG.ATC.GAC.TTG.GAA.GGC.GAC.ATT.CCG
Thr-Leu-Gly-Arg-Val-Phe-Asn-Val-Leu-Gly-Glu-Pro-Ile-Asp-Leu-Glu-Gly-Asp-Ile-Pro
5756 5815
GCT.GAC.GCC.CGC.CGC.GAC.CCG.ATT.CAC.CGT.CCG.GCG.CCA.AAA.TTC.GAG.GAA.TTG.GCG.ACG
Ala-Asp-Ala-Arg-Arg-Asp-Pro-Ile-His-Arg-Pro-Ala-Pro-Lys-Phe-Glu-Glu-Leu-Ala-Thr
5816 5875
GAA.GTC.GAA.ATT.TTG.GAA.ACG.GGG.ATT.AAA.GTC.GTT.GAC.TTG.CTT.GCC.CCG.TAT.ATT.AAA
Glu-Val-Glu-Ile-Leu-Glu-Thr-Gly-Ile-Lys-Val-Val-Asp-Leu-Leu-Ala-Pro-Tyr-Ile-Lys
5876 5935
GGC.GGA.AAA.ATC.GGT.TTG.TTC.GGC.GGC.GCT.GGC.GTA.GGA.AAA.ACG.GTC.TTG.ATC.CAA.GAG
Gly-Gly-Lys-Ile-Gly-Leu-Phe-Gly-Gly-Ala-Gly-Val-Gly-Lys-Thr-Val-Leu-Ile-Gln-Glu
5936 5995
CTG.ATT.CAC.AAC.ATC.GCC.CAA.GAG.CAC.GGC.GGG.ATT.TCC.GTC.TTT.GCT.GGC.GTC.GGC.GAA
Leu-Ile-His-Asn-Ile-Ala-Gln-Glu-His-Gly-Gly-Ile-Ser-Val-Phe-Ala-Gly-Val-Gly-Glu
5996 6055
CGG.ACA.CGC.GAA.GGA.AAC.GAC.TTG.TAC.CAT.GAG.ATG.AAA.GAT.TCC.GGC.GTC.ATC.AGC.AAA
Arg-Thr-Arg-Glu-Gly-Asn-Asp-Leu-Tyr-His-Glu-Met-Lys-Asp-Ser-Gly-Val-Ile-Ser-Lys
6056 6115
ACG.GCC.ATG.GTG.TTC.GGA.CAA.ATG.AAT.GAG.CCG.CCG.GGG.GCG.CGG.ATG.CGC.GTC.GCC.TTG
Thr-Ala-Met-Val-Phe-Gly-Gln-Met-Asn-Glu-Pro-Pro-Gly-Ala-Arg-Met-Arg-Val-Ala-Leu
6116 6175
ACC.GGC.TTG.ACG.ATG.GCC.GAA.TAC.TTC.CGT.GAT.GAA.CAA.GGC.CAA.GAC.GGC.TTG.CTC.TTT
Thr-Gly-Leu-Thr-Met-Ala-Glu-Tyr-Phe-Arg-Asp-Glu-Gln-Gly-Gln-Asp-Gly-Leu-Leu-Phe
6176 6235
ATC.GAT.AAC.ATC.TTC.CGT.TTC.ACG.CAG.GCC.GGT.TCG.GAA.GTG.TCG.GCG.CTG.TTA.GGC.CGC
Ile-Asp-Asn-Ile-Phe-Arg-Phe-Thr-Gln-Ala-Gly-Ser-Glu-Val-Ser-Ala-Leu-Leu-Gly-Arg
6236 6295
ATG.CCG.TCG.GCC.ATT.GGT.TAC.CAA.CCG.ACG.TTG.GCG.ACG.GAG.ATG.GGT.CAA.TTG.CAA.GAG
Met-Pro-Ser-Ala-Ile-Gly-Tyr-Gln-Pro-Thr-Leu-Ala-Thr-Glu-Met-Gly-Gln-Leu-Gln-Glu
6296 6355
CGG.ATC.ACG.TCG.ACG.GCG.AAA.GGC.TCG.ATC.ACC.TCG.ATT.CAA.GCG.ATT.TAC.GTC.CCG.GCC
Arg-Ile-Thr-Ser-Thr-Ala-Lys-Gly-Ser-Ile-Thr-Ser-Ile-Gln-Ala-Ile-Tyr-Val-Pro-Ala
6896 6415
GAC.GAC.TAT.ACG.GAC.CCG.GCT.CCG.GCC.ACG.ACG.TTC.TCG.CAC.TTG.GAT.GCG.ACG.ACG.AAC
Asp-Asp-Tyr-Thr-Asp-Pro-Ala-Pro-Ala-Thr-Thr-Phe-Ser-His-Leu-Asp-Ala-Thr-Thr-Asn
6416 6475
CTG.GAG.CGG.AAG.CTC.GCG.GAG.ATG.GGG.ATT.TAT.CCG.GCC.GTT.GAC.CCG.CTC.GTC.TCG.ACA
Leu-Glu-Arg-Lys-Leu-Ala-Glu-Met-Gly-Ile-Tyr-Pro-Ala-Val-Asp-Pro-Leu-Val-Ser-Thr

6476 6535
TCG.CGT.GCG.TTG.GCG.CCG.GAA.ATC.GTC.GGC.GAG.GAG.CAT.TAT.CAA.GTC.GCC.CGC.AAA.GTG
Ser-Arg-Ala-Leu-Ala-Pro-Glu-Ile-Val-Gly-Glu-Glu-His-Tyr-Gln-Val-Ala-Arg-Lys-Val
6536 6595
CAG.CAA.ACG.CTC.GAA.CGT.TAT.AAA.GAA.TTG.CAA.GAC.ATC.ATC.GCC.ATC.TTG.GGG.ATG.GAT
Gln-Gln-Thr-Leu-Glu-Arg-Tyr-Lys-Glu-Leu-Gln-Asp-Ile-Ile-Ala-Ile-Leu-Gly-Met-Asp
6596 6655
GAA.CTG.TCG.GAT.GAA.GAC.AAA.CTC.GTC.GTT.CAT.CGC.GCC.CGC.CGC.ATC.CAG.TTC.TTC.TTG
Glu-Leu-Ser-Asp-Glu-Asp-Lys-Leu-Val-Val-His-Arg-Ala-Arg-Arg-Ile-Gln-Phe-Phe-Leu
6656 6715
TCG.CAA.AAC.TTC.CAC.GTG.GCG.GAG.CAG.TTC.ACG.GGC.CAA.CCG.GGC.TCG.TAC.GTG.CCG.GTG
Ser-Gln-Asn-Phe-His-Val-Ala-Glu-Gln-Phe-Thr-Gly-Gln-Pro-Gly-Ser-Tyr-Val-Pro-Val
6716 6773
AAA.GAA.ACA.GTG.CGC.GGC.TTT.AAA.GAA.ATT.TTG.GAA.GGC.AAA.TAC.GAC.CAT.CTT.CCG.GAA
Lys-Glu-Thr-Val-Arg-Gly-Phe-Lys-Glu-Ile-Leu-Glu-Gly-Lys-Tyr-Asp-His-Leu-Pro-Glu
6774 6833
GAT.CGG.TTC.CGC.TTA.GTC.GGC.CGC.ATT.GAA.GAA.GTC.GTT.GAA.AAA.GCG.AAA.GCG.ATG.GGT
Asp-Arg-Phe-Arg-Leu-Val-Gly-Arg-Ile-Glu-Glu-Val-Val-Glu-Lys-Ala-Lys-Ala-Met-Gly
*6834 **S/D** **S/D** **ε subunit** → 6893*
GTC.GAA.GTG.TGA. C.CCG.GGA.TAG.GGG.GAT.TGG.ACA.ATG.AAA.ACG.ATC.CAC.GTG.AGC.GTC
Val-Glu-Val-*** Met-Lys-Thr-Ile-His-Val-Ser-Val
6894 6953
GTT.ACT.CCT.GAT.GGC.CCG.GTG.TAC.GAA.GAC.GAT.GTT.GAG.ATG.GTC.AGC.GTC.AAA.GCG.AAA
Val-Thr-Pro-Asp-Gly-Pro-Val-Tyr-Glu-Asp-Asp-Val-Glu-Met-Val-Ser-Val-Lys-Ala-Lys
6954 7013
AGC.GGC.GAG.CTC.GGC.ATT.TTG.CCG.GGG.CAC.ATT.CCG.CTT.GTC.GCC.CCG.CTC.GAG.ATC.AGC
Ser-Gly-Glu-Leu-Gly-Ile-Leu-Pro-Gly-His-Ile-Pro-Leu-Val-Ala-Pro-Leu-Glu-Ile-Ser
7014 7073
GCG.GCC.CGG.CTG.AAA.AAA.GGC.GGC.AAA.ACG.CAA.TAC.ATT.GCC.GTC.AGC.GGC.GGC.TTT.TTG
Ala-Ala-Arg-Leu-Lys-Lys-Gly-Gly-Lys-Thr-Gln-Tyr-Ile-Ala-Val-Ser-Gly-Gly-Phe-Leu
7074 7133
GAA.GTC.CGC.CCG.GAC.AAC.GTG.ACG.ATT.TTG.GCT.CAA.GCT.GCT.GAA.CGG.GCG.GAG.GAC.ATT
Glu-Val-Arg-Pro-Asp-Asn-Val-Thr-Ile-Leu-Ala-Gln-Ala-Ala-Glu-Arg-Ala-Glu-Asp-Ile
7134 7193
GAC.GTC.CTC.CGC.GCC.AAA.GCG.CGA.AAG.AGC.GGG.CGG.ACG.CCG.CTG.CAA.AGC.CAG.CAG.GAC
Asp-Val-Leu-Arg-Ala-Lys-Ala-Arg-Lys-Ser-Gly-Arg-Thr-Pro-Leu-Gln-Ser-Gln-Gln-Asp
7194 7253
GAC.ATC.GAC.TTC.AAA.CGG.GCC.GAA.CTG.GCG.TTA.AAA.CGC.GCC.ATG.AAC.CGT.TTG.AGC.GTT
Asp-Ile-Asp-Phe-Lys-Arg-Ala-Glu-Leu-Ala-Leu-Lys-Arg-Ala-Met-Asn-Arg-Leu-Ser-Val
7254 7313
GCG.GAA.ATG.AAG.TAA.AAA.GCA.GGG.GAG.TTT.TTC.CCC.TGC.TTT.TTT.TGA.TGA.AAA.CGG.AAG
Ala-Glu-Met-Lys-***
terminater
7314 7373
AAT.CGT.CCG.TCA.AGG.AGA.GAG.GCA.AAC.GCT.CTC.TCT.TTT.TTC.GTT.CCT.CGC.ATT.CCG.TCG

7374 7433
TTT.CAA.CAA.AGA.TGA.AGC.GCT.TAC.ATG.AAT.GGA.TAC.CGT.GTT.GAC.AGT.GTT.AAC.TTG.AAG

7434 7493
TTT.TGC.TGA.CAC.GGC.TGG.ATA.TAA.ATG.TTA.TAA.TGA.AAT.CGA.AGG.AAA.TTA.TTG.AAA.AAT

*7494
TTC.ATC.ATT.AAC.ACC.ATA.CGT.ATG.TAG.TAA.ATC.CTG.GAA.

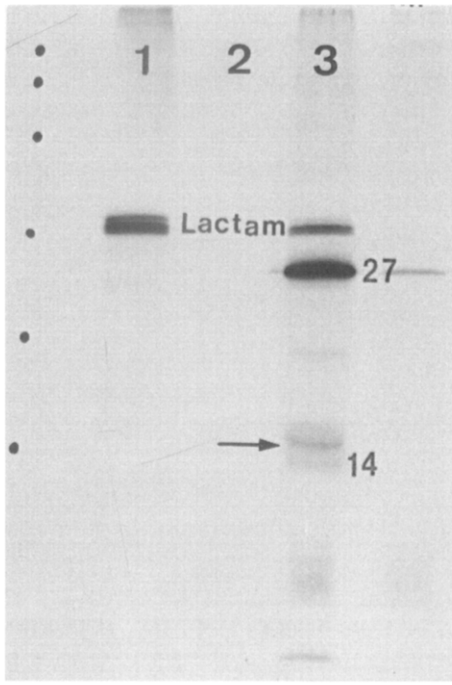

Fig. 3. Autoradiogram of the product of in vitro transcription-translation of the I gene of thermophilic ATP synthase by its own promoter. The Dra I-Bal I fragment of thermophilic DNA (2.3 kilobase base pairs) was inserted into pTZ19R. The plasmid DNA (5 μl, 500 μg/ml) was incubated in the mixture of an Amersham prokaryotic DNA-directed translation kit (N. 380Y) consisting of 5 μl of S-30 extract of *E. coli* (ribosomes and factors), 7.5 μl tRNA-nucleotide-energy source, 3 μl of amino acid mixture (without methionine), 2 μl of L-[^{35}S]methionine (500 Ci/mmol, 5 mCi/ml) and 7.5 μl of dilution buffer, at 37 °C for 1 h. Then 5 μl of methionine chase solution was added to the mixture and after 10 min, denaturating solution (8% sodium dodecylsulfate, 8% β-mercaptoethanol, 16% glycerol, 0.1 M Tris HCl (pH 6.8) and 1 mg/ml brom phenol blue) was added and the mixture was analyzed by acrylamide gel electrophoresis as described previously [13]. The manipulation of ^{35}S-proteins and autoradiography of the gel were essentially as reported by Sancar et al. [18]. Lane 1, control plasmid without inserted DNA (lactam means lactamase expressed); lane 2, no DNA as control; and lane 3, pTZ19R containing an insert of the promoter of TF_0F_1, the I gene and the gene for the N-terminal portion of subunit a. Dots on the left side indicate molecular mass standard proteins of 66, 45, 31, 22 and 14.4 kDa in succession.

which did not interfere with growth of the host, *E. coli* HB101 was used because the lac repressor of the cell did not inhibit transcription by the many copies of the pTZ system and thus the inducer was saved to express the gene.

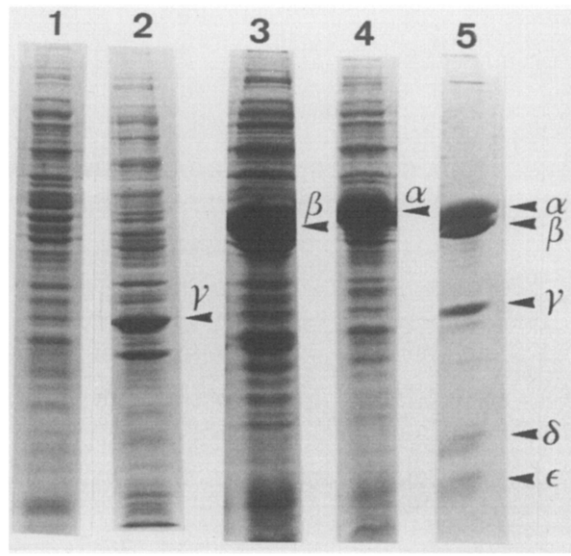

Fig. 4. Polyacrylamide gel electrophoresis of total *E. coli* proteins containing over-expressed subunits. The cells cultured as described in the text were dissolved in 1% SDS, 50 mM Tris-HCl (pH 6.8) and 1% 2-mercaptoethanol at 100 °C for 5 min, and then the total proteins (10 μl each) wre analyzed by electrophoresis in the presence of sodium dodecylsulfate as described previously [14]. Lane 1, extract of *E. coli* HB 101; lane 2, extract of *E. coli* NM522 harboring pKKγE after induction with isopropyl β-D thiogalactoside; lane 3, extract of *E. coli* HB 101 harboring pUCβ; lane 4, extract of *E. coli* HB101 harboring pTZα; lane 5, thermophilic F_1 (TF_1) 20 μg. Arrows with Greek letters indicate the positions of subunits of TF_1 with the following molecular weights: α, 54590; β, 51938; γ, 31778; δ, 19657 and ε, 13333.

E. coli HB101 haboring pTZα (pTZα/HB101) expressed the α subunit even at the beginning of the logarithmic phase (6 h), in the absence of the inducer (Fig. 4, lane 4). Over-expression of the α subunit did not interfere with cell growth, and the stationary phase was reached within 11 h. On the other hand, a large amount of the β subunit was produced only when the cells reached the stationary phase in the absence of the inducer (Fig. 4, lane 3). For obtaining the γE subunit, the appropriate addition of the inducer was necessary to express pKK-γE. The pKK-γE/NM522 cells were grown to an optical density of the culture of 1.0. Then 100 μg of the inducer was added and 6 h later, the cells were collected by centrifugation (Fig. 4, lane 2).

Thermophilic proteins are stable during purification. Thus the yields of the purified α (Fig. 5)

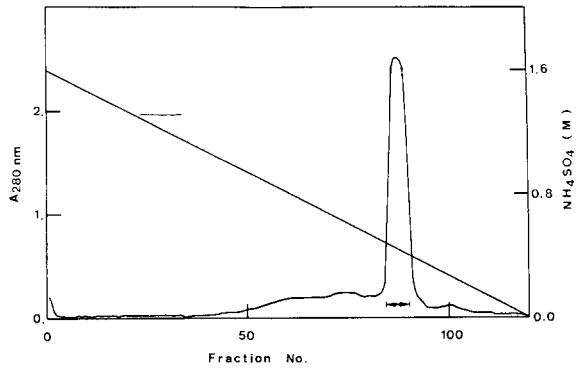

Fig. 5. Hydrophobic column chromatography of the α subunit. The DEAE Sephacel fraction described in the text (70 ml, in 1.4 M ammonium sulfate, 20 mM Tris SO_4 (pH 8.0), 0.1 mM EDTA) was applied to a hydrophobic column (2 cm×32 cm, 100 ml) of Toyopearl HW 65 fine (Toyo Soda, Tokyo, Japan) which had been equilibrated with a solution containing 1.6 M ammonium sulfate, 20 mM Tris SO_4 (pH 8.0). The column was washed with 100 ml of the same solution, and then proteins were eluted with a decreasing linear gradient of ammonium sulfate (1.6 M to 0 M) in 20 mM Tris-SO_4 (600 ml). Fractions of 5 ml were collected, and fractions nos. 85–90 (30 ml total volume) were pooled as indicated as the purified α subunit. The purity of this material was over 99% as shown by acrylamide gel electrophoresis [13].

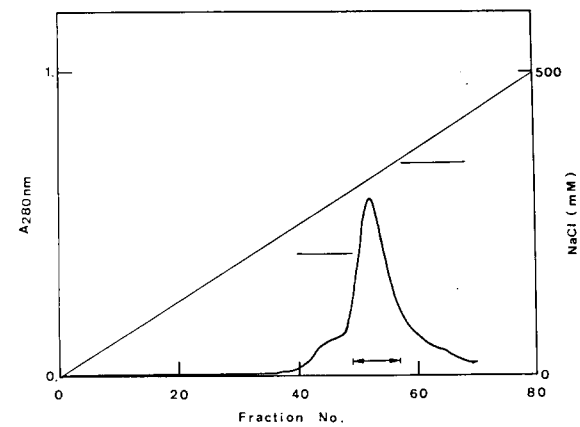

Fig. 6. DEAE Sephacel chromatography of the β subunit. The fraction obtained by hydrophobic chromatography as described in the text (50 ml) was dialyzed against a large excess of buffer consisting of 20 mM Tris-SO_4 (pH 8.0), 5 mM NaCl and 0.1 mM EDTA. Then the dialysate was applied to a column (2×25 cm, 78 ml) of DEAE Sephacel equilibrated with the same buffer, and washed with 200 ml of the same buffer. Proteins were eluted with a linear gradient (0 to 500 mM, 600 ml) of NaCl in the same buffer. Fractions of 7.5 ml were collected and fractions nos. 49–57 (67.5 ml) were pooled as purified β subunit. The purity of this material was 99% as shown by acrylamide gel electrophoresis [13].

and β (Fig. 6) subunits in the total extracts of transformants were about 42% (Table I) and 12% (Table II), respectively, of the total protein in the extract. In some cases, about 300 mg of the α subunit was purified from 1 l cultured cells. The over-produced purified α, β and γE subunits were reconstituted into an oligomer as described previously [20] and ATPase activity of the over-expressed αβγE complex was more than 95% of that of the αβγ complex of the wild type.

Discussion

Homology studies

The order of genes in the thermophilic bacterium TF_0F_1 operon was identical with that found for the corresponding gens in the *E. coli unc* operon [5,6]. However, the genes for *Rhodopseudomonas blastica* [21] and *Rhodospirillum rubrum* [22] differ from that of TF_0F_1 in that neither a homologue of the I gene nor genes for F_0 subunits are associated with the F_1 loci. In *Synechococcus* 6301 [23], the genes for the β and ε subunits are separated from other genes for F_0F_1, like those of chloroplasts F_0F_1 [24]. Homologies were detected with each of the five *E. coli* F_0F_1 proteins (Fig. 7). The homologies among the F_1 subunits of human (β) [25], bovine [26] and yeast (β) [27] mitochondria, chloroplasts [24] and *Synechococcus* [23] have been discussed in several reports [7,8,12,21–26]. The homology of ATP-binding sites in F_1 (α and β) and adenylate kinase [28] is important in site-directed mutagenesis studies.

Although there is less homology in the F_0 subunits of different species, the *oli*2 and *oli*4 sites of a subunit are quite homologous [7,8,23] (Fig. 7). The a subunit of TF_0, like mitochondrial subunit 6, is 61 amino acid residues shorter than that of EF_0, and the segment near the N-terminal of the a subunit is highly homologous to that of *Synechococcus* (from no. 1000 to no. 1070 of Fig. 2, –ATRSLQLRPTGMQNFMEWVFDF–). The I subunit of TF_0 is also homologous to the corresponding subunit of *Synechococcus* especially around No. 720 (–SRLAL–). The peptide sequence of the c subunit of TF_0 is exactly as reported by Hoppe and Sebald [4]. The b subunit consists of a hydrophobic N-terminal domain and hydrophilic α-helices as in the b subunit of EF_0 [7,8].

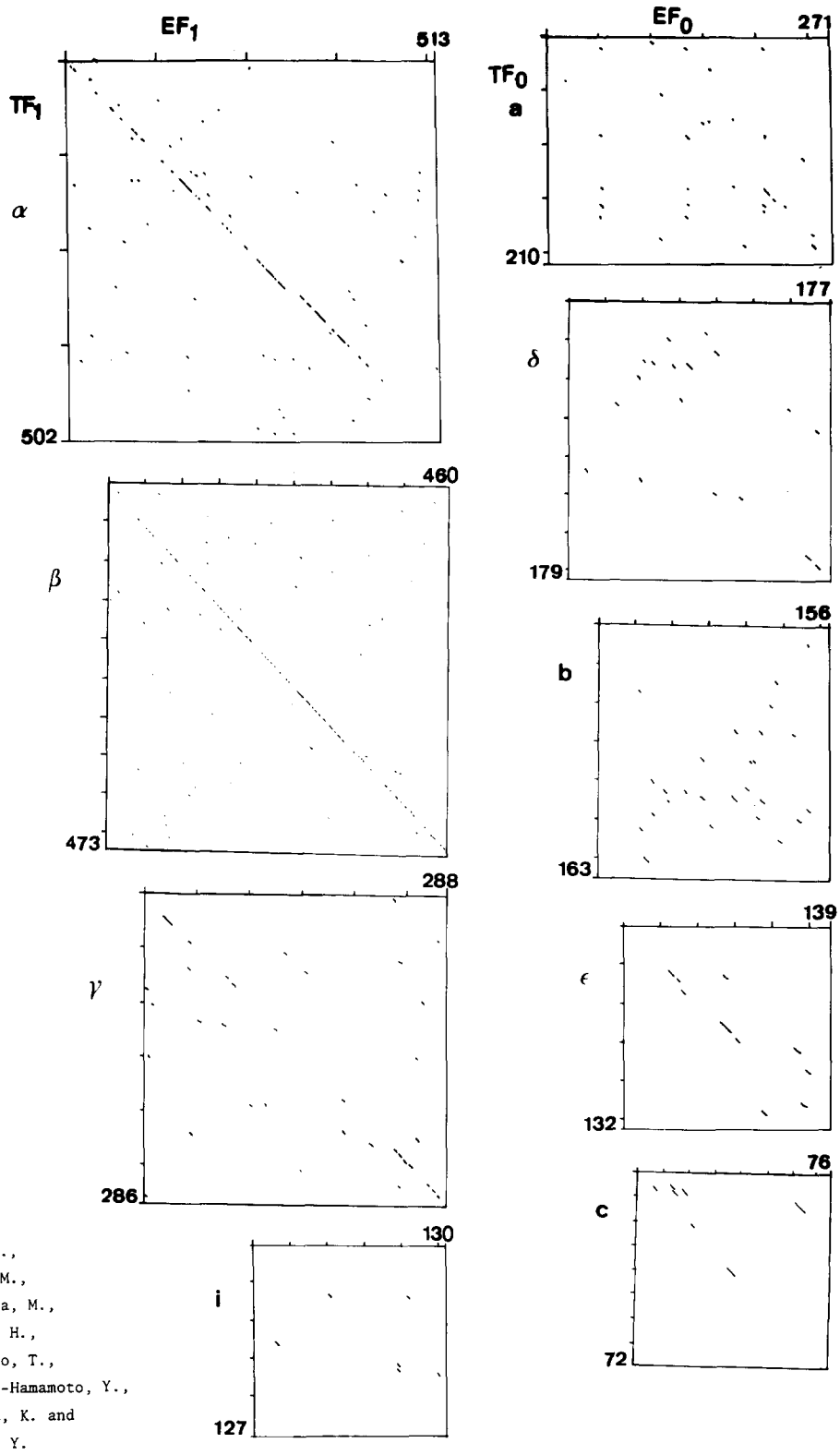

Fig. 7.
Ohta, S.,
Yohda, M.,
Ishizuka, M.,
Hirata, H.,
Hamamoto, T.,
Otawara-Hamamoto, Y.,
Matsuda, K. and
Kagawa, Y.

Fig. 7. Pairwise comparisons of protein sequences of the subunits of F_0F_1 obtained from thermophilic bacterium PS3 with their *E. coli* homologues, using the computer program Harr plot, at span 3, GENETYX, Software Development Co., Ltd. Tokyo. The ordinates are the sequences of thermophilic bacterium PS3, and the abscissae, those of *E. coli*. The scale on each axis is in amino acid residues.

Over-expression of thermophilic subunits

Differences among the codon usages of subunits of ATP synthase are suggested to be a determinant of the subunit stoichiometry by affecting the efficiency of gene expression [7]. The codon usage for TF_0F_1 is quite different from that for *E. coli* F_0F_1; Leu (TTG), Val (GTC), Thr (ACG), Ser (UCG), Arg (CGC), and Gly (GGC) are often used as codons in TF_0F_1 (Fig. 2), but thermophilic genes were well expressed in *E. coli* (Figs. 3 and 4). A gene of an extreme thermophile, *Thermus thermophilus* was also well expressed in *E. coli* [16]. It is noteworthy that a Shine-Dalgarno sequence of thermophilic genes for the α and β subunits was also used in *E. coli*. However, the rare initiation codon, GTG, of the thermophilic γ subunit was not used efficiently in *E. coli*, but the product of the γ subunit with an artificial N-terminus (γE subunit) was functionally equivalent to the wild-type γ subunit isolated from TF_1.

Acknowledgments

The authors thank Dr. Hiroshi Nojima for his discussion in this work. This work was supported by a Grant-in-Aid for Scientific Research on Priority Areas of 'Bioenergetics' to Y.K. from the Ministry of Education, Science and Culture of Japan.

References

1 Mitchell, P. (1985) J. Biochem. 97, 1–18.
2 Kagawa, Y. (1972) Biochim. Biophys. Acta 265, 297–338.
3 Pullman, M.E., Penefsky, H.S., Datta, A. and Racker, E. (1960) J. Biol. Chem. 235, 3322–3329.
4 Hoppe, J. and Sebald, W. (1984) Biochim. Biophys. Acta 768, 1–27.
5 Okamoto, H., Sone, N., Hirata, H., Yoshida M. and Kagawa, Y. (1977) J. Biol. Chem. 252, 6125–6131.
6 Kagawa, Y. (1979) Biochim. Biophys. Acta 505, 45–93.
7 Futai, M. and Kanazawa, H. (1983) Microbiol. Rev. 47, 258–312.
8 Walker, J.E., Saraste, M. and Gay, N.J. (1984) Biochim. Biophys. Acta 768, 164–200.
9 Ohta, S., Tsuboi, M., Oshima, T., Yoshida, M. and Kagawa, Y. (1980) J. Biochem. 87, 1609–1617.
10 Hirata, H., Ohno, K., Sone, N., Kagawa, Y. and Hamamoto, T. (1986) J. Biol. Chem. 261, 9839–9843.
11 Yoshida, M., Sone, N., Hirata, H. and Kagawa, Y. (1977) J. Biol. Chem. 252, 3480–3485.
12 Kagawa, Y., Ishizuka, M., Saishu, T. and Nakao, J. (1986) J. Biochem. 100, 923–934.
13 Saishu, T., Nojima, H. and Kagawa, Y. (1986) Biochim. Biophys. Acta 867, 97–106.
14 Maniatis, T., Fritsch, E.F. and Sambrook, J. (1982) Molecular Cloning: a Laboratory Manual, Cold Spring Harbor, New York.
15 Davis, L.G., Dibner, M.D. and Battey, J.F. (1986) Basic Methods in Molecular Biology, Elsevier, Amsterdam.
16 Kagawa, Y., Nojima, H., Nukiwa, N., Ishizuka, M. Nakajima, T., Yasuhara, T., Tanaka, T. and Oshima, T. (1984) J. Biol. Chem. 259, 2956–2960.
17 Sanger, F., Nicklen, S. and Coulson, A.R. (1977) Proc. Natl. Acad. Sci. USA 74, 5463–5467.
18 Sancar, A., Wharton, R.P., Seltzer, S., Kacinski, B.M., Clarke, N.D. and Rupp, W.D. (1981) J. Mol. Biol. 148, 45–62.
19 Bradford, M. (1976) Anal. Biochem. 72, 248–258.
20 Kagawa, Y. and Nukiwa, N. (1981) Biochem. Biophys. Res. Commun. 100, 1370–1376.
21 Tybulewicz, V.L.J., Falk, G. and Walker, J.E. (9184) J. Mol. Biol. 179, 185–214.
22 Falk, G. and Walker, J.E. (1985) Biochem. J. 229, 663–668.
23 Cozens, A.L. and Walker, J.E. (1987) J. Mol. Biol. 194, 359–383.
24 Shinozaki, K. and Sugiura, M. (1982) Nucl. Acids Res. 10, 4923–4934.
25 Ohta, S. and Kagawa, Y. (1986) J. Biochem. 99, 135–141.
26 Walker, J.E., Fearnley, I.M., Gay, N.J., Gibson, B.W., Northrop, F.D., Powell, S.J., Runswick, M.J., Saraste, M. and Tybulewicz, V.L.J. (1985) J. Mol. Biol. 184, 677–704.
27 Takeda, M., Vassarotti, A. and Douglas, M.G. (1985) J. Biol. Chem. 260, 15458–15465.
28 Fry, D.C., Kuby, S.A. and Mildvan, A.S. (1986) Proc. Natl. Acad. Sci. USA 83, 907–911.
29 Anderson, S., Banker, A.R., Barrell, B.G., De Bruijn, M.H.L., Coulson, A.R., Drouin, J., Eperon, I.C., Nirlich, D.D., Roe, B.A., Sanger, F., Schreier, P.H., Smith, A.J.H., Staden, R. and Young,I.G. (1981) Nature 290, 457–465.
30 Macreadie, I.G., Novitski, C.E., Maxwell, R.J., John, U., Ooi, B.G., McMullen, G.L., Lukins, H.B., Linnane, A.W. and Nagley, P. (1983) nucl. Acids Res. 11, 4435–4451.