RESEARCH COMMUNICATION

Cloning, characterization and mapping of the human *ATP5E* gene, identification of pseudogene *ATP5EP1*, and definition of the ATP5E motif

Qiang TU, Long YU[1], Pingzhao ZHANG, Min ZHANG, Honglai ZHANG, Juxiang JIANG, Chiyuan CHEN and Shouyuan ZHAO
Institute of Genetics, Fudan University, No. 220, Handan Road, Shanghai 200433, People's Republic of China

A cDNA encoding the ϵ subunit of human ATP synthase, ATP5E, was isolated from heart, skeletal muscle and spleen cDNA libraries respectively. Its genome structure was characterized as comprising three exons and two introns within a stretch of 5 kb, according to the genomic sequence AL109840. The gene was mapped to human chromosome 20q13.3 between marker D20S173 and 20qter using the radiation hybrid GB4 panel. Northern blot analysis showed that the *ATP5E* gene was expressed as a single 0.6 kb transcript in all 16 human tissues tested, with a high level present in heart and skeletal muscle. A new conserved motif composed of 24 residues, termed the ATP5E motif [$W(R/K)X_5YX_2(Y/F)X_3(C/A)X_4RX_3K$], was defined on the basis of sequences of ATP synthase ϵ subunits from ten different organisms. In addition, a pseudogene *ATP5EP1* was also identified on the basis of genomic sequence AC004066, localized on human chromosome 4q25. By analysing these results combined with the Southern blot patterns of human DNA hybridized with bovine *ATP5E* cDNA reported previously [Vinas, Powell, Runswick, Iacobazzi and Walker (1990) Biochem. J. **265**, 321–326], we provide evidence of yet further homologous sequences (either gene or pseudogene) of *ATP5E*, in addition to *ATP5E* and *ATP5EP1* in the human genome.

Key words: ATP synthase epsilon subunit, genome structure, tissue expression patterns.

INTRODUCTION

ATP synthases are membrane-bound enzymes that produce ATP from ADP and inorganic phosphate (P_i) by using the transmembrane potential gradient for protons to drive the phosphorylation of ADP [1]. ATP synthases found in bacteria, chloroplasts and mitochondria are generally similar in subunit structure. They consist of two sectors: (i) the extrinsic membrane sector, called F_1 and composed of hydrophilic subunits, is the catalytic site of ATP synthesis; and (ii) the intrinsic membrane sector, called F_0 and composed of hydrophobic subunits, is responsible for energy transduction [2,3]. The F_1 sectors of bacteria, chloroplasts and mitochondria all contain five different subunits, named α, β, γ, δ and ϵ in decreasing order of their molecular masses, of stoichiometry $\alpha_3\beta_3\gamma_1\delta_1\epsilon_1$. α, β and γ subunits from different species are homologous; however, the δ subunit from *Escherichia coli* is homologous with the mitochondrial oligomycin-sensitivity conferral protein ('OSCP'), and the ϵ subunit from *E. coli* is homologous with the mitochondrial δ subunit. Furthermore, the mitochondrial ϵ subunit has no counterpart in subunits from other species [4,5].

Since the first mitochondrial ϵ subunit was isolated from a bovine source [4], those from yeast [6], rat [7], sweet potato [8], *Caenorhabditis elegans* [9], mouse, fission yeast, *Zea mays* and *Arabidopsis thaliana* have been isolated and sequenced in succession. The crystal structure of bovine F_1 did not disclose the exact structure or position of the ϵ subunit, since it was not well ordered [10], but this subunit was located in the stalk region of F_1, as determined from phosphorescence and fluorescence energy-transfer measurements [11]. In a yeast strain with a null mutation of the ϵ subunit, oligomycin-sensitive ATPase activity could not be detected, and an uncoupling of the mitochondrial respiration rate that was highly sensitive to oligomycin was promoted [12]. Modification of the ϵ subunit decreased the inhibitory activity of inhibitor protein IF_1 [13]. These studies suggested a possible role for the ϵ subunit in the mechanism of regulation of the mitochondrial ATP synthase. Interestingly, the yeast strain with a null mutation in ϵ grew slowly on medium containing glycerol as the carbon source [14], which indicated that the absence of the ϵ subunit has a critical, but not lethal, effect on the function of the ATP synthase. In order to explore the possible role of ϵ subunit in human disease, it is necessary to study the subunit of the human enzyme in more detail.

EXPERIMENTAL

The cDNA sequence of the bovine ATP synthase ϵ subunit (GenBank accession no. X16978) was used to search the GenBank Human EST (Expressed Sequence Tag) division. Homologous ESTs were assembled into a contig by the Assembly program of the Wisconsin Package (Version 10.0, Genetics Computer Group, Madison, WI, U.S.A.). Primers QT-E1 (5′-ATT GCC GGC GTC TTG GCG ATT CG-3′; nt 45–67) and QT-E2 (5′-ATT GTC AAT TTA TGA ACA AGA CAG G-3′; nt 357–381) were designed on the basis of the contig sequence, custom-made by ShengGong, Inc. (Shanghai, People's Republic of China) and were used to amplify several human tissue cDNA λgt11 libraries (Clontech, Palo Alto, CA, U.S.A.). PCR conditions were as follows: 1 μl of template [$> 10^8$ plaque-forming units (p.f.u.)/ml] was amplified in a 50-μl volume containing 5 μl of 10 × PCR buffer, 1 μl of 20 mM dNTPs, 1.5 μl of 2.5 mM $MgCl_2$, and 2 units of *Taq* polymerase (Promega, Madison, WI, U.S.A.) and 1 μl of 25 mM of each specific primer. PCR reactions were run on a PTC-200 DNA engine (MJ Research, Watertown, MA, U.S.A.) for 34 cycles (1 min at 93 °C, 45 s at 60 °C and 1 min at 72 °C) after an initial denaturation at 93 °C for 3 min, followed

Abbreviations used: EST, expressed sequence tag; GSS, Genome Survey Sequence; OMIM, Online Mendelian Inheritance in Man; RH, Radiation Hybrid.

[1] To whom correspondence should be addressed (e-mail longyu@fudan.edu.cn).

© 2000 Biochemical Society

by a 5-min incubation at 72 °C. The PCR products were subcloned into the pGEM-T vector, and sequenced by using a BigDye Terminator Sequencing Kit and an ABI377 sequencer (Perkin–Elmer Applied Biosystems, Foster City, CA, U.S.A.).

A BLAST search was performed by screening the homologous sequences in GenBank Genome Survey Sequence (GSS) division with the *ATP5E* cDNA sequence. Those GSSs corresponding to *ATP5E* were assembled into a genomic fragment, which was aligned with the cDNA sequence to identify exons. Radiation Hybrid (RH) mapping was subsequently performed by GB4 panel (GeneBridge 4 Human Hamster RH Panel; Research Genetics Inc., Huntsville, AL, U.S.A.) with PCR primers QTE-RH1 (5′-AAG ATC ATC CAT CCA AAC ACC AG-3′) and QTE-RH2 (5′-AAC TTA AGA TGC AAC TGT TCT TCT A-3′), which were designed on the basis of the flanking sequence of the exon. The PCR conditions used were those described above, and reactions were run for 30 cycles (1 min at 93 °C, 45 s at 55 °C and 1 min at 72 °C) after an initial denaturation at 93 °C for 3 min, followed by a 10-min incubation at 72 °C. The PCR products amplified with DNA samples from five Chinese individuals were sequenced to confirm that they were identical in sequence (395 bp) with the assembled genomic sequence. The PCR results were sent to the Sanger RH Mapping Server (http://www.sanger.ac.uk/RHserver/RHserver.shtml) for statistical analysis. The chromosome locus was determined from the GDB database (http://gdbwww.gdb.org/gdb/), and was searched in the register OMIM (Online Mendelian Inheritance in Man) of the National Center for Biotechnology Information (NCBI).

To investigate the expression patterns, Northern blot hybridization was performed with the above PCR products amplified from human heart cDNA library on multiple-tissue Northern (MTN) membranes (MTN I and MTN II; Clontech) carrying mRNA samples from 16 adult human tissues. The PCR products were labelled with [α-^{32}P]dATP (Amersham, Cleveland, OH, U.S.A.) using the PCR method, and purified by using Sepharose G50 column chromatography. The MTN membranes were prehybridized in hybridization/prehybridization solution {50 % (v/v) formamide/5 × SSPE [where 1 × SSPE is 0.15 M NaCl/10 mM NaH_2PO_4 (pH 7.4)/1 mM EDTA] /10 × Denhardt's solution (where 1 × Denhardt's solution is 0.02 % Ficoll 400/0.02 % polyvinylpyrrolidone/0.002 % BSA)/2 % (w/v) SDS/100 μg/ml calf-thymus DNA} at 42 °C for 16 h, followed by hybridization with the labelled probe for a further 24 h, shaking continuously in a hybridization oven (Hybaid U.K., Inc., Ashford, Middx., U.K.). The membranes were then washed several times with wash solution [0.1 × SSC (0.015 M NaCl/0.0015 M sodium citrate)/0.1 % (w/v) SDS] at 42 °C, before exposure to X-ray film at −80 °C for 5 days.

To facilitate the structural characterization of the putative protein, the sequence of ϵ subunit of the human enzyme was used to search GenBank by BLAST and search ProtoMap Cluster databases (http://www.protomap.cs.huji.ac.il/) [15]. Motif analysis was performed by using the MEME program (Multiple EM for Motif Elicitation; http://www.sdsc.edu/MEME/meme/website/) [16].

RESULTS AND DISCUSSION

More than 100 human ESTs homologous with X16978 were obtained from GenBank and assembled into one contig of 437 bp. cDNA fragments of 337 bp were obtained from heart, skeletal muscle and spleen cDNA libraries respectively, and sequenced to confirm that they were identical with the contig sequence. The full-length cDNA contains an open reading frame (nt 96–248) encoding 51 amino acid residues (Figure 1A). The deduced peptide sequence is 94 % identical with the ϵ subunit of the bovine form. The gene was named as ATP synthase, H^+-transporting, mitochondrial F1 complex ϵ subunit (*ATP5E*) by the Human Gene Nomenclature Committee.

Three GSSs (AQ312760, AQ387821 and AQ494137) were obtained when a search was made of the GenBank GSS division, and assembled into a genomic fragment of 1201 bp. By aligning this genomic fragment sequence and the *ATP5E* cDNA sequence, an exon was identified that enabled the primers for RH mapping to be designed. The PCR results of RH were recorded as 10000 10000 00011 01111 10010 10000 01010 00001 10001 00000 01100 01111 10001 01110 00110 11110 10000 11000 000 ('0' means a negative result, whereas '1' indicates a positive one). Statistical analysis with the Sanger RH Mapping Server demonstrated that this gene was located between marker D20S173 (lod score of 20.685) and 20qter. According to the GDB comprehensive gene map, D20S173 was located at 20q13.3. Thus *ATP5E* was mapped to 20q13.3 (see Figure 3A). According to the OMIM register, some interesting loci, such as obesity quantitative trait locus and low-voltage electroencephalogram, were mapped in this region. We noticed that a genomic DNA sequence of 154 kb (AL109840) was released in the GenBank HTGS division as a working draft sequence of chromosome 20 clone J543J19 on August 16 1999. The alignment of either *ATP5E* cDNA or the 1201 bp assembled genomic sequence described above with AL109840 proved that *ATP5E* gene was in this clone, and was composed of three exons. The exon amplified by PCR in the RH mapping is the second exon (see Figure 3A).

The Northern blot result revealed a single transcript of approx. 0.6 kb in all 16 human tissues (Figure 2). It is noteworthy that this gene is expressed at a high level in heart and skeletal muscle, which are the tissues that require a large amount of ATP, with intermediate or low levels of expression in brain, placenta, liver, kidney, pancreas, spleen, thymus, prostate, testis, small intestine, colon and peripheral blood leucocytes, and hardly any expression at all in lung and ovary.

Sequences of ATP synthase from nine other organisms were obtained by searching in GenBank and ProtoMap databases. A multiple alignment for these ϵ subunits was achieved with the GAP program of GCG, which showed that the amino acid sequences of this family of proteins are highly conserved, and a new motif, $W(R/K)X_5YX_2(Y/F)X_3(C/A)X_4RX_3K$ (where 'X' denotes 'any other residue'), was defined by the MEME program (Figure 1C). Each of the ten sequences has an e-value of much less than 10 (from 3.5e-37 to 7.9e-15). (The e-value of a sequence is the expected number of sequences in a random database of the same size that would match the motifs as well as the sequence does.) In this motif, seven residues are the most conserved, of which important content values (providing an idea of which positions in the motif are most highly conserved), are more than 3.0 bits. Four residues (Trp^1, Tyr^8, Arg^{20} and Lys^{24}) are identical among all the above organisms, and three other important residues are highly conserved: Arg^2 is identical in nine out of ten organisms, and Tyr^{11} and Cys^{15} are identical in eight out of ten. The high degree of conservation of the motif indicates that it might be an important functional site; for example, an active site or a binding site for interacting with other subunits.

Furthermore, another homologue of *ATP5E* was found in human genome sequence AC004066 (BAC clone B353C18 at 4q25), which was confirmed by sequencing the PCR products from five Chinese individuals with the primers QT-EP1 (5′-GCT TCA AGA TTT GAA CCA ATA AC-3′) and QT-EP2 (5′-AAC TGC CAT CTA TAT ATT CAT AG-3′). Since the homologous region in AC004066 is as long as the cDNA, whereas the *ATP5E*

© 2000 Biochemical Society

A

```
agtggtctgatcttcctgcggctgaaccgcccggctgagccgacattgccggcgtcttggcgattcggcccgacg 75
agctccgctttcgctacagcatggtggcctactggagacaggctggactcagctacatccgatactcccagatct 150
                    M  V  A  Y  W  R  Q  A  G  L  S  Y  I  R  Y  S  Q  I  C
gtgcaaaagcagtgagagatgcactgaagacagaattcaaagcaaatgctgagaagacttctggcagcaacgtaa 225
  A  K  A  V  R  D  A  L  K  T  E  F  K  A  N  A  E  K  T  S  G  S  N  V  K
aaattgtgaaagtaaagaaggaataatctaccctgactaaagcttgaaatgctacatttccaaggtgaagatgtg 300
  I  V  K  V  K  K  E  *     (51)
tgggcacatgttatggcagattgaaaaggatctcattccatgggaaaaaaaaaaatcctgtcttgttcataaatt 375
gacaatgtcaataaattgaaatatggttcactgttaaaaaaaaaaaaaaaaaaaaaaaaa                437
```

B

```
aggtatgagccactgcacccagcccaccattcaatgtagcttcaagatttgaaccaataacaaatcttaatctgc
75
tgtcatgacatcatgtggatgtattggtaacaggctgggctcagctatgtcaaaaactcccagatctgtgcaaaa 150
           M  W  M  Y  W  *  Q  A  G  L  S  Y  V  K  N  S  Q  I  C  A  K
gtagtaagaaatacactgatgactgaagtcaaagcaaatgctgaaaaagctcctggtagcagcataaaaatggtg 225
V  V  R  N  T  L  M  T  E  V  K  A  N  A  E  K  A  P  G  S  S  I  K  M  V
aatgtaaaaaaggagtaattgtatcctgactaaagcttgaaatgttacacttttaaggtggagatctatgaatat 300
N  V  K  K  E  *
atagatggcagtttgaaacgttctcatgccatgaaaaaaaaggtctaaattaaggtgtgatttttaaaaagtac 375
gtaattatttcctacattagaagaaatagaagaaacaaaggcat                              419
```

C

```
HUMAN    M------VAYWRQAGLSYIRYSQICAKAVRDALKTEFKANA----EKTSGSNVKIVKVKKE      51
BOVIN    M------VAYWRQAGLSYIRYSQICAKAVRDALKTEFKANA----MKTSGSTIKIVKVKKE      51
  RAT    M------VAYWRQAGLSYIRFSQICAKAVRDALKTEFKANA----EKTSGTSIKTVKIKKE      51
MOUSE    M------VAYWRQAGLSYIRFSQICAKAVRDALKTEFKANA----EKTSGSSIKIVKVSKK-E    52
IPOBA    M-ASNAAVPFWRAAGMTYITYSNLCANMVRNCLKEPYRAEALS-REKVHFSFSKWVDGKPQ-KPAIRSDTGEE  70
ARATH    M-ASNAAVPFWRAAGMTYISYSNICANIVRNCLKEPHKAEALT-REKVHFSLSKWADGKPQ-KPVLRSDTPEV  70
MAIZE    MSATTAAVPFWRAAGMTYIGYSNICAALVRNCLKEPFKSEAAS-REKVHFSISKWTDGKQE-KPTVRTESDD  70
CAEEL    M------VA-WRAAGLNYVRYSQIAAQVVRQCTKGGANVKK----PQATLKTTAWENGKMV-SKSQ    54
YEAST    M------SA-WRKAGISYAAYLNVAAQAIRSSLKTELQTASVLNRSQTDAFYTQYKNGTAASEPTPITK  62
SCHPO    M-----AFA-WKKN-FSYSKYASICSQTVRQALKPEIKNEVKT-HGDAEFLYTRWKNGAQEKTESYNSAKSADKE  67

ATP5E Motif          WRQAGLSYIRYSQICAKAVRDALK
```

Figure 1 Nucleotide and deduced amino acid sequences of ATP5E (A) and ATP5EP1 (B), and multiple alignment of the ε subunits of ATP synthase from ten organisms with the ATP5E motif sequence (C)

(**A**) Primers for PCR amplification (QT-E1 and QT-E2) are shown underlined. The polyadenylation signal (aataaa) is enclosed in a box, and the stop codon is marked with an asterisk. (**B**) Shown are the nucleotide and deduced amino acid sequences for ATP5EP1, which correspond to nt 83900–84318 of AC004066. The asterisk indicates the stop codon; primers (QT-EP1, QT-EP2) are shown underlined. (**C**) The organisms used in the alignment are *Homo sapiens* (HUMAN; GenBank accession no. AF052955), *Bos taurus* (BOVIN; P05632), *Rattus norvegicus* (RAT; P29418), *Mus musculus* (MOUSE; P56382), *Ipomoea batatas* (IPOBA; Q06450), *A. thaliana* (ARATH; Q96253), *Z. mays* (MAIZE; Q41898), *C. elegans* (CAEEL; P34539), *Saccharomyces cerevisiae* (YEAST; P21306) and *Schizosaccharomyces pombe* (SCHPO; P87316). Residues that are identical among the ten organisms are shown in reversed-out white lettering on a black background, and the conserved residues are shown in bold. The most highly conserved seven residues of the ATP5E motif are also shown in reversed-out white lettering on a black background.

gene is a split gene and translation analysis showed there was an in-frame stop codon (Figure 1B), it was identified as a pseudogene and was designated with the abbreviation *ATP5EP1* by the Human Gene Nomenclature Committee.

J. E. Walker and co-workers [17] have inferred that there is more than one ε subunit gene or pseudogene of ATP synthase in humans, on the basis of patterns of Southern blot hybridization of the bovine ε subunit cDNA probe (probe 1 and probe 2 are from the 5′ and 3′ regions of the cDNA respectively) with human genomic DNA digested with *Bam*HI, *Hin*dIII, *Pst*I and *Sac*I restriction enzymes respectively. To demonstrate their inference, the restriction map of *ATP5E* (Figure 3A) was made from the genomic sequence AL109840, and that of *ATP5EP1* (Figure 3B) from AC004066. All predicted restriction fragments that should

© 2000 Biochemical Society

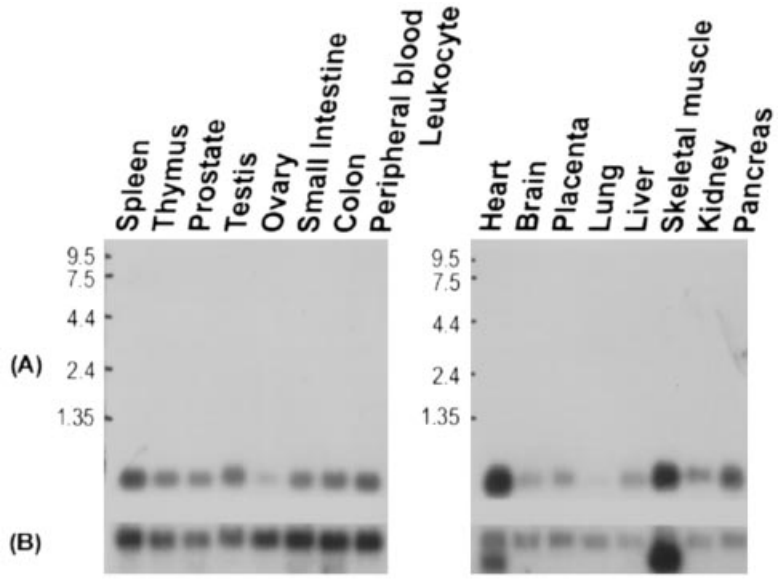

Figure 2 Northern blot analysis of *ATP5E*, showing the levels of expression in 16 human tissues

A multiple-tissue Northern blot membrane (Clontech) was probed with *ATP5E* (**A**), and β-actin (**B**) as a control. Molecular size (in kb) is indicated to the left of the blots.

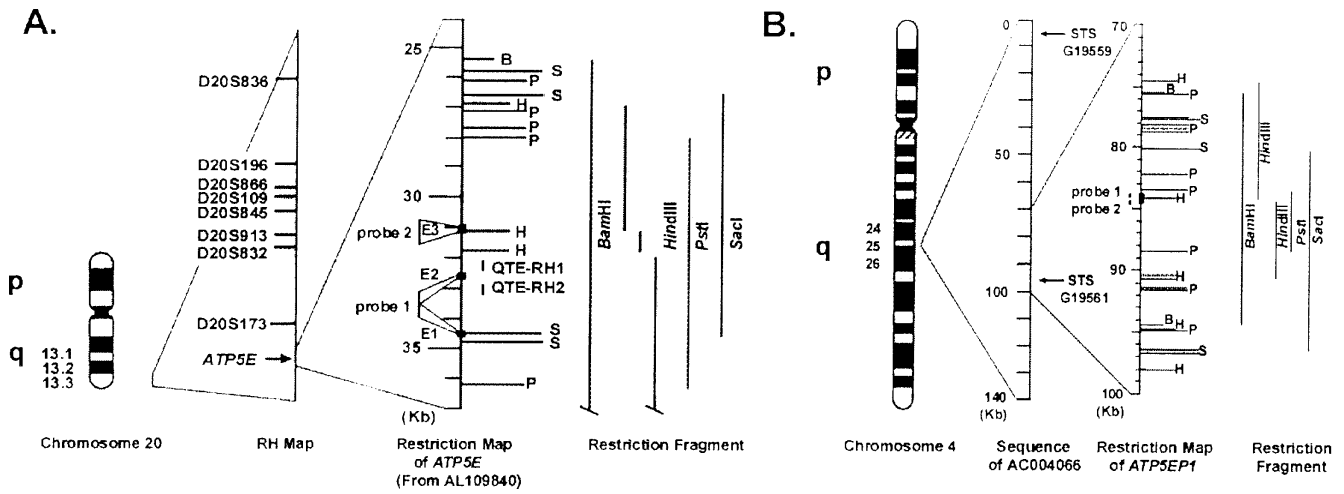

Figure 3 Mapping, genome structure and restriction map of *ATP5E* (A) and *ATP5EP1* (B)

ATP5E was mapped to chromosome 20q13.3, between markers D20S173 and 20qter. The *ATP5E* genome sequence was in HTGS AL109840, which was released as a working-draft sequence of clone J543J19. Alignment of genome sequence and *ATP5E* cDNA sequence showed that there are three exons, which are indicated by black boxes on the restriction map. RH primers (QTE-RH1 and QTE-RH2) are shown as short lines on the right-hand side of the restriction map. Positions corresponding to the bovine cDNA probes used by Walker and co-workers [17] are shown as short lines on the left-hand side of the restriction map. Restriction sites are indicated by horizontal lines on the right-hand side of the restriction map, with appropriate labelling (B, *Bam*HI; H, *Hin*dIII; P, *Pst*I; S, *Sac*I). Restriction fragments, which can hybridize with the bovine probes, are indicated by vertical lines to the right of the restriction map: *Bam*HI fragment ($>$ 11613 bp), *Hin*dIII fragment (probe 1: 4224 bp, 696 bp; probe 2: $>$ 5177 bp), *Pst*I fragment (8170 bp) and *Sac*I fragment (7978 bp). (**B**) *ATP5EP1* was identified in genome sequence AC004066 (clone B353C18, map 4q25) near 84 kb. The *APT5EP1* gene is indicated by the black box on the restriction map. Positions corresponding to the bovine cDNA probes are shown as short lines on the left-hand side of the restriction map. Restriction sites are indicated by horizontal lines on the right-hand side of the restriction map, as described for (**A**). Restriction fragments used in (**B**) and shown to the right of the restriction map are the *Bam*HI fragment (18694 bp), *Hin*dIII fragments (probe 1, 9389 bp; probe 2, 6200 bp), *Pst*I fragment (4965 bp) and the *Sac*I fragment (16215 bp).

be present in such a Southern blot were found as specific bands in the Southern hybridization experiment reported previously; however, a 4.2-kb band of the *Pst*I-digested fragment and a 6.0-kb band of the *Sac*I-digested fragment found in the aforementioned Southern hybridization cannot be interpreted in terms of the restriction maps of *ATP5E* and *ATP5EP1*. Therefore, in addition to *ATP5E* and *ATP5EP1* reported here, there might be other homologous sequences of *ATP5E* in human genome.

In conclusion, we have cloned, characterized and mapped the gene of the human ϵ subunit of F_1F_0 ATP synthase, *ATP5E*, to chromosome 20q13.3 and identified a pseudogene *ATP5EP1*

© 2000 Biochemical Society

at chromosome 4q25. We have also defined a motif of the ATP5E protein family in ten species, and provided suggestions of other sequences homologous with them in the human genome.

We thank Dr. John E. Walker (member of the Royal Society, worked at Medical Research Council Laboratory of Molecular Biology, Cambridge, U.K.) greatly for a presubmission review of this manuscript, and for further suggestions. This work was supported by the National 973 Program and 863 High Technology Program (863-Z-02-04-01) and the National Science Foundation of China (39525015,39680019).

REFERENCES

1 Mitchell, P. (1961) Coupling of phosphorylation to electron and hydrogen transfer by a chemi-osmotic type of mechanism. Nature (London) **191**, 144–148
2 Senior, A. E. (1990) The proton-translocating ATPase of *Escherichia coli.* Annu. Rev. Biophys. Biophys. Chem. **19**, 7–41
3 Weber, J. and Senior, A. E. (1997) Catalytic mechanism of F1-ATPase. Biochim. Biophys. Acta **1319**, 19–58
4 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. (1985) Primary structure and subunit stoichiometry of F1-ATPase from bovine mitochondria. J. Mol. Biol. **184**, 677–701
5 Walker, J. E., Lutter, R., Dupuis, A. and Runswick, M. J. (1991) Identification of the subunits of F1F0-ATPase from bovine heart mitochondria. Biochemistry **30**, 5369–5378
6 Arselin, G., Gandar, J. C., Guerin, B. and Velours, J. (1991) Isolation and complete amino acid sequence of the mitochondrial ATP synthase epsilon-subunit of the yeast *Saccharomyces cerevisiae.* J. Biol. Chem. **266**, 723–727
7 Higuti, T., Yoshihara, Y., Kuroiwa, K., Kawamura, Y., Toda, H. and Sakiyama, F. (1992) A simple, rapid method for purification of epsilon-subunit, coupling factor 6, subunit d, and subunit e from rat liver H(+)-ATP synthase and determination of the complete amino acid sequence of epsilon-subunit. J. Biol. Chem. **267**, 22658–22661
8 Morikami, A., Ehara, G., Yuuki, K. and Nakamura, K. (1993) Molecular cloning and characterization of cDNAs for the gamma- and epsilon-subunits of mitochondrial F1F0 ATP synthase from the sweet potato. J. Biol. Chem. **268**, 17205–17210
9 Wilson, R., Ainscough, R., Anderson, K., Baynes, C., Berks, M., Bonfield, J., Burton, J., Connell, M., Copsey, T. and Cooper, J. (1994) 2.2 Mb of contiguous nucleotide sequence from chromosome III of *C. elegans.* Nature (London) **368**, 32–38
10 Abrahams, J. P., Leslie, A. G., Lutter, R. and Walker, J. E. (1994) Structure at 2.8 Å resolution of F1-ATPase from bovine heart mitochondria. Nature (London) **370**, 621–628
11 Gabellieri, E., Strambini, G. B., Baracca, A. and Solaini, G. (1997) Structural mapping of the epsilon-subunit of mitochondrial H(+)-ATPase complex (F1). Biophys. J. **72**, 1818–1827
12 Guelin, E., Chevallier, J., Rigoulet, M., Guerin, B. and Velours, J. (1993) ATP synthase of yeast mitochondria. Isolation and disruption of the ATP epsilon gene. J. Biol. Chem. **268**, 161–167
13 Solaini, G., Baracca, A., Gabellieri, E. and Lenaz, G. (1997) Modification of the mitochondrial F1-ATPase epsilon subunit, enhancement of the ATPase activity of the IF1-F1 complex and IF1-binding dependence of the conformation of the epsilon subunit. Biochem. J. **327**, 443–448
14 Lai-Zhang, J., Xiao, Y. and Mueller, D. M. (1999) Epistatic interactions of deletion mutants in the genes encoding the F1-ATPase in yeast *Saccharomyces cerevisiae.* EMBO J. **18**, 58–64
15 Yona, G., Linial, N., Tishby, N. and Linial, M. (1998) A map of the protein space – an automatic hierarchical classification of all protein sequences. I.S.M.B. **6**, 212–221
16 Bailey, T. L. and Elkan, C. (1994) Fitting a mixture model by expectation maximization to discover motifs in biopolymers. I.S.M.B. **2**, 28–36
17 Vinas, O., Powell, S. J., Runswick, M. J., Iacobazzi, V. and Walker, J. E. (1990) The epsilon-subunit of ATP synthase from bovine heart mitochondria. Complementary DNA sequence, expression in bovine tissues and evidence of homologous sequences in man and rat. Biochem. J. **265**, 321–326

Received 6 January 2000/2 February 2000; accepted 9 February 2000

© 2000 Biochemical Society