The ϵ -subunit of ATP synthase from bovine heart mitochondria

Complementary DNA sequence, expression in bovine tissues and evidence of homologous sequences in man and rat

Octavi VIÑAS,* Steven J. POWELL,† Michael J. RUNSWICK, Vito IACOBazzi‡ and John E. WALKER§
M.R.C. Laboratory of Molecular Biology, Hills Road, Cambridge CB2 2QH, U.K.

The ϵ -subunit of ATP synthase from bovine heart mitochondria is assembled into the extrinsic membrane sector, F_1 -ATPase. The mature protein is 50 amino acid residues in length and its function is unknown. It is a nuclear gene product that is imported into the organelle. A mixture of 64 oligonucleotides 17 bases long, designed on the basis of the known protein sequence, was synthesized and used as a hybridization probe to isolate a cognate cDNA clone from a bovine library. The DNA sequence of this clone was determined, and the protein sequence of the ϵ -subunit deduced from it agrees exactly with that determined by direct sequence analysis of the protein isolated from bovine hearts. The bovine cDNA was used as a hybridization probe to examine the expression of the ϵ -subunit in various bovine tissues. mRNAs related to the cDNA are found in all of these tissues, and no evidence was obtained of the presence of mRNAs for the ϵ -subunit with similar coding sequences and dissimilar 3' non-coding regions. By hybridization experiments with digests of DNA from cow, man and rat it has been shown that sequences related to the bovine cDNA are present in the genomes of all three species. More than one related sequence was detected in all cases, indicating the presence in all three genomes of more than one gene and/or pseudogenes.

INTRODUCTION

The ATP synthases found in eubacteria, chloroplasts and mitochondria are related multi-subunit structures. They are membrane-bound enzymes that make ATP from ADP and P_i by harnessing the transmembrane electrochemical potential gradient for protons, $\Delta\mu_{H^+}$, to drive the phosphorylation of ADP (Mitchell, 1961; Nicholls, 1982). They contain an extrinsic membrane domain that is bound by a stalk to an intrinsic membrane structure, referred to as F_0 . This extrinsic domain can be dissociated from the membrane. It is a globular assembly known as F_1 -ATPase and consists of five different polypeptides, α , β , γ , δ and ϵ (reviewed by Senior, 1979). The smallest of these is the ϵ -subunit. In the enzyme isolated from bovine heart mitochondria this is a polypeptide of unknown function 50 amino acid residues long. Its protein sequence has been established by direct analysis of the protein isolated from heart tissue (Walker *et al.*, 1985) and shows that it is not evidently related to any of the subunits that have been described in bacterial and chloroplast ATP synthases. Somewhat confusingly, this lack of relationship includes subunits that carry the same name as the ϵ -subunit in bacterial and chloroplast enzymes (Walker *et al.*, 1982). It is also unclear whether mitochondrial ATP synthases in other species have a subunit that is related to the bovine ϵ -subunit. The F_1 -ATPases have been purified from mitochondria of *Saccharomyces cerevisiae* (Tzagoloff, 1969) and rat (Catterall & Pedersen, 1971), and both preparations contain five different polypeptides, including one known

as the ϵ -subunit. However, these ϵ -subunits appear to be considerably bigger than the bovine protein; that in *S. cerevisiae* is estimated to have an apparent molecular mass of 12 kDa (Velours *et al.*, 1987), and in rat mitochondria its size is about 7.5 kDa (Catterall *et al.*, 1973). So in the absence of sequence data there is doubt that these are equivalent to the bovine ϵ -subunit protein. Recently it has been demonstrated that the α -subunit of bovine F_1 -ATPase has at least two related isoforms with differences in protein sequence (Walker *et al.*, 1989). One form is expressed in heart tissue (Walker *et al.*, 1985, 1989) and the other in liver (Breen, 1988), and so the question arises of whether the same ϵ -subunit that has been characterized from heart is also expressed in other bovine tissues, and whether isoforms of this protein are also found. In order to explore these and other questions, as described in the present paper we have isolated and sequenced a cognate bovine cDNA clone for the bovine ϵ -subunit by using as a hybridization probe a synthetic mixture of 64 oligonucleotides designed on the basis of the known protein sequence. The sequence of the clone confirms the protein sequence of the ϵ -subunit. The clone has been employed as a hybridization probe to investigate the number of sequences related to the bovine ϵ -cDNA in the human, bovine and rat genomes, and also to examine the distribution of mRNAs for this protein in various bovine tissues. These studies show that all three mammals contain sequences related to the coding sequence of the bovine cDNA. In the genome of each species there are present more than one sequence related to the bovine cDNA, which indicates that there may be multiple genes,

* Present address: Department de Bioquímica i Fisiologia, Universitat de Barcelona, 08028 Barcelona, Spain.

† Present address: I.C.I. Pharmaceuticals Division, Northwich, Cheshire CW9 7RA, U.K.

‡ Present address: Dipartimento Farmaco-Biologica, Università di Bari, 70125 Bari, Italy.

§ To whom correspondence should be addressed.

These sequence data have been submitted to the EMBL/GenBank Data Libraries under the accession number X16978.

possibly including pseudogenes, for the ϵ -subunit. The hybridization experiments with samples of RNA isolated from several bovine tissues show that they all contain RNAs related to the bovine cDNA.

MATERIALS AND METHODS

Oligonucleotide synthesis

A mixture of 64 oligonucleotides with the sequence 3'-CTYAARTTYCGNTTRCG-5' was synthesized by a semi-manual solid-phase phosphotriester method (Gait *et al.*, 1982; Sproat & Bannwarth, 1983). This was based upon the protein sequence EFKANA, amino acid residues 29–34 in the protein sequence of the ϵ -subunit of ATP synthase from bovine mitochondria (Walker *et al.*, 1985). Its minimal dissociation temperature was estimated to be 44 °C from the 'rule of thumb' that each G·C and A·T base-pair contributes 4 °C and 2 °C respectively (Suggs *et al.*, 1981). The mixture of oligonucleotides was freed from salts and reaction by-products as described previously (Walker *et al.*, 1987a). Samples of the purified mixture were radiolabelled by the use of polynucleotide kinase and [γ -³²P]ATP (specific radioactivity approx. 3000 Ci/mmol). The procedure used for purifying the radio-labelled oligonucleotides has been described previously (Powell *et al.*, 1989).

Screening the cDNA library

The bovine cDNA library employed in experiments to isolate a cDNA for the ϵ -subunit of mitochondrial ATP synthase has been described previously (Gay & Walker, 1985a). It was prepared from mRNA isolated mostly from heart tissue, but a small proportion of the mRNA was derived from bovine liver. The cDNAs were cloned in the plasmid vector pUC 8 (Messing & Vieira, 1982). This library has provided cDNAs for other mitochondrial proteins (Gay & Walker, 1985b; Runswick *et al.*, 1987; Walker *et al.*, 1987a,b; Powell *et al.*, 1989; Dyer *et al.*, 1989), and procedures for screening it with mixtures of oligonucleotides have been described in detail in these earlier publications. In the case of the isolation of a cDNA for the ϵ -subunit, approx. 5×10^5 recombinant colonies were transferred to Pall nylon membranes. The DNA in them was fixed to the membrane and hybridized with a radiolabelled sample (about 50 pmol) of the mixture of synthetic oligonucleotides for 18 h at 39 °C, i.e. at 5 °C below the estimated minimum dissociation temperature. The membranes then were washed successively with 6×SSC at 41 °C, 43 °C and finally 45 °C. (1×SSC is 150 mM-NaCl/15 mM-sodium citrate.) They were autoradiographed at -70 °C for 48 h with pre-flashed Fuji RX 150 film. More than 20 positively hybridizing colonies were detected in this experiment, and ten of them were selected for further investigation. Plasmid DNA were prepared from each of these isolates. In order to release the inserted DNA, samples of the plasmids were digested with the restriction enzymes *Eco*RI and *Bam*HI, both together and separately. The products were analysed by electrophoresis in a 1.0% agarose minigel, and the restriction fragments that had arisen from the inserted DNAs were cloned into appropriately prepared pUC and M13 mp8 and M13 mp9 vectors.

DNA sequence analysis

DNA sequences were determined by the dideoxy chain-

termination procedure (Sanger *et al.*, 1977) with the modifications described by Biggin *et al.* (1983). Data were compiled and analysed by using the computer programs DBUTIL (Staden, 1982) and ANALYSEQ (Staden, 1984).

Preparation of DNA and RNA from mammalian tissues

The preparation of human and bovine genomic DNAs have been described previously (Walker *et al.*, 1987a). Rat DNA was prepared from a liver by the same procedure. Total RNA was prepared from bovine tissues by the procedure of Chirgwin *et al.* (1979), and polyadenylated RNA was prepared from total RNA by using mAP (or messenger affinity paper) from Amersham International according to the manufacturer's instructions. About 10–20 μ g of polyadenylated RNA, depending upon the tissue, was obtained from 1 mg of total RNA.

Hybridization with RNA and genomic DNA

Samples (3 μ g) of polyadenylated RNA were fractionated by electrophoresis in a 1.4% high-gelling-temperature agarose gel, transferred to Hybond N membranes and hybridized with 'prime-cut' probes as described previously (Powell *et al.*, 1989). Then the membranes were washed once at room temperature with 6×SSC and twice for 15 min at 65 °C with 6×SSC. Then they were autoradiographed for 3 days at -70 °C with preflashed film in the presence of an intensifying screen. The probes were removed from the blots by washing the filters for 2 h at 65 °C in a solution containing 5 mM-Tris/HCl buffer, pH 8.0, 2 mM-EDTA and 0.1×Denhardt's solution. [Denhardt's solution contains 1 mg each of polyvinylpyrrolidone, bovine serum albumin (fraction V) and Ficoll/ml.] They were rehybridized in the presence of a probe for the mouse β -actin gene.

Hybridizations of 'prime-cut' probes with fractionated digests of genomic DNA were carried out at 65 °C for 16 h in the presence of a solution containing 6×SSC, 5×Denhardt's solution, sonicated salmon sperm DNA (100 μ g/ml), 0.5% Sarkosyl and 10% (w/v) dextran sulphate. The same probes were employed as in RNA hybridization experiments (see above). After hybridization filters were washed three times for 10 min at 65 °C in 2×SSC. For further details see the legend to Fig. 4.

Fig. 1. Sequence analysis of a cDNA clone encoding the precursor of the ϵ -subunit of bovine mitochondrial ATP synthase

The thick line represents the inserted DNA in recombinant pBovε1. The insert is flanked by an *Eco*RI site (▽) and a *Bam*HI site (▼) and also contains an internal *Eco*RI site. The horizontal arrows represent the DNA sequences that were determined.

Fig. 2. Nucleotide sequence of a cDNA encoding the ϵ -subunit of ATP synthase from bovine heart mitochondria

The position of the *N*-terminal valine residue of the mature protein isolated from heart mitochondria is shown. The mature protein is numbered from 1 to 50. A potential polyadenylation signal in the DNA sequence is boxed, and the sequence that hybridized with the synthetic oligonucleotide probe is underlined. An internal *Eco*RI site is found in the first six bases of the probe sequence and a *Bsm*I site is also marked. These sites were used in making probes for hybridization experiments.

RESULTS AND DISCUSSION

cDNA cloning and sequence analysis of the ϵ -subunit

By screening the bovine cDNA library with the mixture of 64 oligonucleotides designed on the basis of residues 29–34 in the known protein sequence of the mature ϵ -subunit, at least 20 positively hybridizing recombinants were detected. Restriction digestion of the recombinant plasmids purified from ten of them showed that they all contained an *Eco*RI–*Bam*HI fragment of about 0.3 kb, and in addition two of them, pBov ϵ 1 and pBov ϵ 2, also contained an *Eco*RI fragment of about 0.2 kb. Both fragments were cloned into M13 vectors and their DNA sequences were determined completely in both senses (see Fig. 1). Together with the known protein sequence of the ϵ -subunit, this showed that the two restriction fragments abutted at an *Eco*RI site, and that together the fragments encode the ϵ -subunit of bovine mitochondrial ATP synthase (see Fig. 2). No attempt has been made to overlap this *Eco*RI site by DNA sequence analysis, as the overlap is provided by the protein sequence.

The entire cDNA sequence presented in Fig. 2 is 354 bases in length. It encodes a protein sequence that is identical with that of the mature ϵ -subunit of ATP synthase, which was determined by direct sequence analysis of protein isolated from the bovine heart mitochondrial F_1 -ATPase (Walker *et al.*, 1985). This sequence is followed by a 3' non-coding sequence that terminates with A-9, which is presumably derived from the poly(A) tail of the mRNA. At 14 bases to the 5' side of this run of A residues is found the sequence ATTAAG, which can serve as a signal for polyadenylation of transcripts (Proudfoot & Brownlee, 1976). The codon for the *N*-terminal valine residue of the mature protein is immediately preceded in the cDNA sequence by an additional 31 nucleotide residues. Translation of this sequence in the same phase of the DNA as that coding for the mature protein shows that it encodes the protein sequence ARRAPSRLDM. Since the 5' DNA sequence in the clone does not include an in-phase stop codon, it is unclear whether this protein sequence represents part of a mitochondrial import sequence or whether the ATG codon immediately preceding the codon for the *N*-terminal valine residue of the mature protein is the translational initiator. The protein sequence ARRAPSRLDM has features that are characteristic of other mitochondrial import sequences: it has a net basic

charge, and contains an arginine residue close to the start of the mature protein. The presence of an acidic amino acid residue in a mitochondrial import sequence would be rather unusual, but rare examples are known (Gay & Walker, 1985b). If the second interpretation were correct, and the ATG codon was indeed the translational initiator, the protein sequence specifying mitochondrial import would have to be contained within the mature ϵ -subunit, which has a net charge of +8. The absence of a cleavable import sequence in nuclear coded mitochondrial proteins is somewhat unusual, but examples are provided by the ADP/ATP translocase (Arends & Sebald, 1984; Baker & Leaver, 1985; Adrian *et al.*, 1986; Powell *et al.*, 1989; Cozens *et al.*, 1989) and the d-subunit of ATP synthase in bovine mitochondria (Walker *et al.*, 1987b).

Fig. 3. Hybridization of polyadenylated RNA from bovine tissues with probes derived from the cDNA for the ϵ -subunit of ATP synthase

Key to lanes: L, liver; K, kidney; B, brain; H, heart. The samples of fractionated RNA were hybridized with 'prime-cut' probes consisting of nucleotide residues 209–354 (left-hand panel) and nucleotide residues 1–120 (right-hand panel). See the Materials and methods section for further details.

Fig. 4. Hybridization of restriction digests of DNA from man, rat and cow with probes derived from the cDNA for the ϵ -subunit of ATP synthase

Samples of DNA (200 μ g) from man [(a) and (b)], cow [(c) and (d)] and rat (e) were digested with the restriction enzymes indicated. They were hybridized under the conditions described in the Materials and methods section with 'prime-cut' probes containing either nucleotide residues 1-120 [(a), (c) and (e)] or nucleotide residues 209-354 [(b) and (d)]. The positions of markers are indicated.

Expression of the ϵ -subunit in bovine tissues

In order to investigate the distribution of mRNAs for the ϵ -subunit in different bovine tissues, polyadenylated RNA was prepared from heart, liver, brain and kidney. These samples were fractionated by agarose-gel electrophoresis and then hybridized in separate experiments with nucleotide residues 1–120 and 209–354 of the cDNA sequence; the former contains the coding sequence for the *N*-terminal 29 amino acid residues of the protein and the short 5' sequence preceding it, and the latter is derived from the 3' non-coding sequence. These hybridization probes were selected because with other multiple expressed genes for mitochondrial proteins such as the dicyclohexylcarbodi-imide-reactive proteolipid (Gay & Walker, 1985b; Dyer *et al.*, 1989; Dyer & Walker, 1989) and the ADP/ATP translocase (Powell *et al.*, 1989; Cozens *et al.*, 1989) it has been found that the coding sequences are highly conserved whereas the 3' non-coding sequences have diverged extensively. Therefore the probe containing the coding region could be expected to hybridize with mRNAs for different but closely related isoforms of the ϵ -subunit (if present), and the probe derived from the 3' non-coding region would probably detect only the mRNAs from which the cDNA clone is derived. These experiments (see Fig. 3) show that mRNAs for the ϵ -subunit are present in all the tissues that were examined, and since hybridization signals of similar intensity were obtained with the 5' and the 3' probes with each tissue, and given that the two probes are of similar lengths, it appears that the major mRNA species for the ϵ -subunit of ATP synthase in all of the tissue examined is that corresponding to the isolated cDNA. Identical amounts of RNA were used in these experiments. This was confirmed by removing the probes and re-probing the same filters with a probe for the mouse β -actin gene (K. Willison & J. Rogers, unpublished work). This gave signals of approximately equal intensity. So it appears that there is considerable variation in the concentrations of mRNA for the ϵ -subunit from tissue to tissue, with the highest concentrations in brain.

Presence of multiple sequences encoding the ϵ -subunit in the genomes of cow, man and rat

In order to look for the presence of sequences related to the bovine cDNA for the ϵ -subunit of ATP synthase in the human and rat genomes, samples of restriction digests were fractionated by agarose-gel electrophoresis, transferred to nylon membranes and then hybridized with the same radiolabelled probes as were employed in the Northern-blotting experiments described in the previous section. In both species strongly hybridizing bands were observed (see Fig. 4). So it seems that the genomes of both species are likely to contain a gene for a homologue of the bovine ϵ -subunit. Moreover, the presence of more than one hybridizing band in the various digests of the DNA not only in man and rat, but also in the bovine DNA, indicates that there are multiple genes (possibly including pseudogenes) for this protein in all three species. It is becoming apparent that this is true of many nuclear encoded proteins in mammalian mitochondria. For example, it has been shown that the proteolipid (Gay & Walker, 1985b; Dyer & Walker, 1989; Dyer *et al.*, 1989) and α -subunits (Walker *et al.*, 1989) of ATP synthase and the ADP/ATP translocase (Battini *et al.*, 1987; Walker *et al.*, 1987c; Neckelmann

et al., 1987; Houldsworth & Attardi, 1988; Powell *et al.*, 1989) have at least two, and in the last case at least three, expressed genes. The proteolipid has also been shown to have numerous spliced pseudogenes in both man and cow (Dyer & Walker, 1989; Dyer *et al.*, 1989). In all three examples the related genes are expressed differently in various bovine tissues. In the cases of other mitochondrial proteins, Southern-blotting experiments suggest that they also may have more than one human and one bovine gene (Walker *et al.*, 1987a; Dyer *et al.*, 1989; Pilkington & Walker, 1989).

O. V. and V. I. were supported by an EMBO short-term Fellowship and a FEBS Fellowship respectively.

REFERENCES

- Adrian, G. S., McCammon, M. T., Montgomery, D. L. & Douglas, M. G. (1986) *Mol. Cell. Biol.* **6**, 626–634
- Arends, H. & Sebald, W. (1984) *EMBO J.* **3**, 377–382
- Baker, A. & Leaver, C. J. (1985) *Nucleic Acids Res.* **13**, 5857–5867
- Battini, R., Ferrari, S., Kaczmarek, L., Calabretta, B., Chen, S. T. & Baserga, R. (1987) *J. Biol. Chem.* **262**, 4355–4359
- Biggin, M. D., Gibson, T. J. & Hong, G. F. (1983) *Proc. Natl. Acad. Sci. U.S.A.* **80**, 3963–3965
- Breen, G. A. M. (1988) *Biochem. Biophys. Res. Commun.* **152**, 264–269
- Catterall, W. A. & Pedersen, P. L. (1971) *J. Biol. Chem.* **246**, 4987–4994
- Catterall, W. A., Coty, W. A. & Pedersen, P. L. (1973) *J. Biol. Chem.* **241**, 7427–7431
- Chirgwin, J. M., Przbyla, A. F., MacDonald, A. J. & Rutter, W. J. (1979) *Biochemistry* **18**, 5294–5299
- Cozens, A. L., Runswick, M. J. & Walker, J. E. (1989) *J. Mol. Biol.* **206**, 261–280
- Dyer, M. R. & Walker, J. E. (1989) *J. Mol. Biol.*, in the press
- Dyer, M. R., Gay, N. J. & Walker, J. E. (1989) *Biochem. J.* **260**, 249–258
- Gait, M. J., Matthes, H. W. D., Singh, M., Sproat, B. S. & Titmas, R. C. (1982) *Nucleic Acids Res.* **10**, 6243–6254
- Gay, N. J. & Walker, J. E. (1985a) *Biochem. J.* **225**, 707–712
- Gay, N. J. & Walker, J. E. (1985b) *EMBO J.* **4**, 3519–3524
- Houldsworth, J. & Attardi, G. (1988) *Proc. Natl. Acad. Sci. U.S.A.* **85**, 377–381
- Messing, J. & Vieira, J. (1982) *Gene* **19**, 269–276
- Mitchell, P. (1961) *Nature (London)* **191**, 144–148
- Neckelmann, N., Li, K., Wade, R. P., Shuster, R. & Wallace, D. C. (1987) *Proc. Natl. Acad. Sci. U.S.A.* **84**, 7580–7584
- Nicholls, D. G. (1982) *Bioenergetics: An Introduction to the Chemiosmotic Theory*, Academic Press, New York and London
- Pilkington, S. J. & Walker, J. E. (1989) *Biochemistry* **28**, 3257–3264
- Powell, S. J., Medd, S. M., Runswick, M. J. & Walker, J. E. (1989) *Biochemistry* **28**, 866–873
- Proudfoot, N. J. & Brownlee, G. G. (1976) *Nature (London)* **263**, 211–214
- Runswick, M. J., Powell, S. J., Nyren, P. & Walker, J. E. (1987) *EMBO J.* **6**, 1367–1373
- Sanger, F., Nicklen, S. & Coulson, A. R. (1977) *Proc. Natl. Acad. Sci. U.S.A.* **74**, 5463–5467
- Senior, A. E. (1979) in *Membrane Proteins in Energy Transduction* (Capaldi, R. A., ed.), pp. 233–278, M. Dekker, New York and Basel
- Sproat, B. S. & Bannwarth, W. (1983) *Tetrahedron Lett.* **24**, 5771–5774

- Staden, R. (1982) *Nucleic Acids Res.* **10**, 4731–4751
- Staden, R. (1984) *Nucleic Acids Res.* **12**, 505–519
- Suggs, S. V., Hirose, T., Miyake, T., Kawashima, E. H., Johnson, M. J., Itakura, K. & Wallace, R. B. (1981) in *Developmental Biology using Purified Genes* (Brown, D. D., ed.), pp. 683–693, Academic Press, New York and London
- Tzagoloff, A. (1969) *J. Biol. Chem.* **244**, 5020–5026
- Velours, J., Arselin de Chateaubodeau, G., Galante, M. & Guérin, B. (1987) *Eur. J. Biochem.* **164**, 579–584
- Walker, J. E., Runswick, M. J. & Saraste, M. (1982) *FEBS Lett.* **146**, 393–396
- Walker, J. E., Fearnley, I. M., Gay, N. J., Gibson, B. W., Northrop, F. D., Powell, S. J., Runswick, M. J., Saraste, M. & Tybulewicz, V. L. J. (1985) *J. Mol. Biol.* **184**, 677–701
- Walker, J. E., Gay, N. J., Powell, S. J., Kostina, M. & Dyer, M. R. (1987*a*) *Biochemistry* **26**, 8613–8619
- Walker, J. E., Runswick, M. J. & Poulter, L. (1987*b*) *J. Mol. Biol.* **197**, 89–100
- Walker, J. E., Cozens, A. L., Dyer, M. R., Fearnley, I. M., Powell, S. J. & Runswick, M. J. (1987*c*) *Chem. Scr.* **27B**, 97–105
- Walker, J. E., Powell, S. J., Viñas, O. & Runswick, M. J. (1989) *Biochemistry* **28**, 4702–4708

Received 23 June 1989/2 August 1989; accepted 8 August 1989