The Mitochondrial Uncoupling Protein Gene
CORRELATION OF EXON STRUCTURE TO TRANSMEMBRANE DOMAINS*

Leslie P. Kozak†, James H. Britton, Ulrike C. Kozak, and James M. Wells
From The Jackson Laboratory, Bar Harbor, Maine 04609

The mitochondrial uncoupling protein, a protein essential for the thermogenic properties of brown fat in mammals, is inserted in the inner mitochondrial membrane by means of six α-helical hydrophobic transmembrane domains. We have sequenced a complete cDNA and parts of the gene to determine that the mitochondrial uncoupling protein gene is composed of six exons, each of which encodes a transmembrane domain. We also show that transcription of the uncoupling protein gene is from a single start site; however, the use of alternative poly(A) addition signal sequences results in two mRNAs, the major species of 1221 nucleotides, not including the poly(A) tail, and a minor species of about 1600 nucleotides. The 5′-untranslated region of the mRNA is composed of 231 nucleotides, and the 3′-untranslated region contains 81 nucleotides prior to addition of the poly(A) tail.

Three ion carrier proteins of the inner mitochondrial membrane, the ADP/ATP translocator, the uncoupling protein (Ucp) of brown fat, and the phosphate carrier protein, are characterized by the presence of six α-helical hydrophobic transmembrane domains by which they are inserted into the mitochondrial membrane (1-4). Similarities in amino acid sequence among the proteins suggest that they evolved from a common ancestral gene (5). All three proteins have the same basic repetitive structure. It is a tripartite structure consisting of three homologous repeating 100-amino acid segments, with each segment containing two transmembrane domains (2). We now report that the structure of the nuclear gene for the mitochondrial Ucp reflects the domain structure. Each of the six exons in the gene encodes one of the transmembrane domains. We also show that transcription of the uncoupling protein gene is from a single start site and that the use of alternative poly(A) addition signal sequences results in two mRNAs, the major species of 1221 nucleotides, not including the poly(A) tail, and a minor species of about 1600 nucleotides.

RESULTS AND DISCUSSION

We previously isolated a cDNA clone for the mitochondrial uncoupling protein (UCP) from mouse brown fat (6). By Northern blot analysis, this cDNA hybridized to two mRNAs, a major species of approximately 1300 nucleotides and a minor species of 1700 nucleotides. Both mRNAs were highly induced when mice were exposed to the cold at 5 °C, suggesting they were derived from the same transcriptional unit. In order to determine the molecular mechanisms controlling Ucp gene expression in brown fat, we have been investigating the structure of both the Ucp mRNA and gene. Additional screening of our brown fat cDNA library has provided a cDNA sequence which is complementary to the full length of the major Ucp mRNA species (Fig. 1A). The Ucp gene was isolated from a Charon 28 BALB/c embryo DNA library as two overlapping clones. We have sequenced the cDNA clones (Fig. 1A). We have also used synthetic primers to sequence the Ucp mRNA directly to define the 5′ end of the mRNA, and we have sequenced parts of the 5′ and 3′ flanking region and the exons of the Ucp gene (Fig. 1B).

The start site for transcription was determined by comparing the RNA sequence to the genomic sequence as illustrated in Fig. 2A. Reverse transcription of the RNA terminates with a G. Sequencing of an M13 clone of the 5′ flanking region of the mRNA sequence, and 81 bases of 3′-untranslated sequence. Sequencing of the 3′ flanking region of the Ucp gene suggested an explanation for the minor mRNA species of 1700 bases. Two additional AUAAA signal sequences are present, one is a single base downstream of the poly(A) addition site of the

MATERIALS AND METHODS

* This work was supported by National Institutes of Health Grant HD08431 (to L. P. K.). The Jackson Laboratory is fully accredited by the American Association for Accreditation of Laboratory Animal Care. The costs of publication of this article were defrayed in part by the payment of page charges. This article must therefore be hereby marked "advertisement" in accordance with 18 U.S.C. Section 1734 solely to indicate this fact.

† To whom correspondence should be sent.

‡ Portions of this paper (including "Materials and Methods," Footnote 2, and Fig. 1) are presented in miniprint at the end of this paper. Miniprint is easily read with the aid of a standard magnifying glass. Full size photocopies are included in the microfilm edition of the Journal that is available from Waverly Press.
Mitochondrial Uncoupling Protein Gene

A.

![Restriction enzyme map of overlapping Ucp cDNA clones](image)

**FIG. 1.** A, restriction enzyme map of overlapping Ucp cDNA clones. The dashed line at the right-hand end of p-Ucp 2 represents the cloning artifact described in the text. The encircled P represents restored Pst I sites at the insert/vector boundary. B, the sequencing strategy for the exons of the Ucp gene is shown. The size and position of the introns determined by partial sequencing and restriction enzyme mapping of the genomic clones is given. Sequencing reactions marked by a bar utilized synthetic deoxyoligonucleotide primers. The restriction enzyme sites are A, Aci I; Au, Aau I; B, Bam HI; Bg, Bgl II; Ba, BstNI; E, Eco RI; H, Hind III; Hc, Hinc II; M, Mbo I; N, Nar I; Ns, Nsi I; P, Pst I; Ps, Poul II; S, Sfa NI; and X, Xba I.

The sequence of these exons and adjoining intron regions shows that the Ucp gene is composed of six exons (Fig. 1, Miniprint section). Also shown in Fig. 1 (Miniprint section) are differences in sequence found between the cDNA sequence of p-Ucp 2 and the genomic sequence. These differences are located at position 146 in the 5' untranslated region and at codons 247 and 299. Since these substitutions do not change the amino acid sequence, we think the differences are genuine and arise from differences in the genotype of the cDNA (C57BL/6J) and the genomic DNA (BALB/c).

The location of the introns relative to the domain structure of the uncoupling protein is clearly evident (Fig. 3). Introns II and IV interrupt the coding sequence within codons 108 and 209 to divide the coding region into the three 100-amino acid repetitive segments observed by sequence similarities for the inner membrane carrier proteins (2, 4). In addition, introns I, III, and V further subdivide the coding region so that introns I and III interrupt transmembrane domains A and B and C and D, respectively, while intron V is located in codon 269 slightly within the region which encodes transmembrane domain F. Intron I interrupts the gene in a region, codon 42, which encodes a β strand possibly associated with a membrane pore (2). Remarkably, each of the transmembrane domains of the UCP is encoded by a separate exon.

A relation between introns and exons encoding transmembrane domains was first observed with bovine rhodopsin which has seven transmembrane domains (16). Introns in the rhodopsin gene interrupt the coding regions at three positions which mark the boundaries of the hydrophobic regions. A more striking correlation was recently described for the 3-hydroxy-3-methylglutaryl coenzyme A reductase gene where each of the seven transmembrane domains, which enable the protein to be inserted in the endoplasmic reticulum, is separated by an intron (17). The single transmembrane domain of the H-2 molecule is also defined as a separate exon (18). A striking contrast to a gene organization where each transmembrane domain is encoded by a separate exon is found in the
Fig. 3. A schematic representation of the transmembrane domains in the inner mitochondrial membrane. The transmembrane domains are lettered A–F. The amino acid position at which the introns (arrows) interrupt the coding region is given. The dashed line represents the region of β-strand structure interrupted by intron I.

The entire protein is encoded by a single exon. The only G-protein-coupled receptor which has an exon/intron structure is rhodopsin and this is the first gene found to have a relationship between the transmembrane domains and intron/exon structure (16). It is unlikely that the presence or absence of introns in the coding region for transmembrane domains has any functional importance; however, insights into the evolutionary relationships among proteins with transmembrane domains may become evident when the structures of other genes encoding proteins with such domains are determined.

REFERENCES
1. Aquila, H., Miara, D., Eulitz, M., and Klingenberg, M. (1982) Hoppe-Seyler’s Z. Physiol. Chem., 363, 345–349
2. Aquila, H., Link, T. A., and Klingenberg, M. (1985) EMBO J. 4, 2369–2376
3. Boulaud, F., Weißenbach, J., and Riquier, D. (1986) J. Biol. Chem. 261, 14873–14890
4. Runswick, M. J., Powell, S. J., Nyren, P., and Walker, J. E. (1987) EMBO J. 6, 1367–1373
5. Klingenberg, M. (1985) in Membrane Transport Driven by Ion Gradients (Semenza, G., and Kinne, R., eds) pp. 279–288, New York Academy of Science, New York
6. Jacobsson, A., Stadier, U., Glotzer, M. A., and Kozak, L. P. (1985) J. Biol. Chem. 260, 16250–16254
7. Benton, W. D., and Davia, R. W. (1977) Science 196, 180–182
8. Sanger, F., Nicklen, S., and Coulson, A. R. (1977) Proc. Natl. Acad. Sci. U. S. A. 74, 5463–5467
9. Haltiner, M., Kenne, T., and Tjian, R. (1986) Nucleic Acids Res. 13, 1015–1025
10. Gelliber, J., Zeff, R. A., Melvold, R. W., and Nathenson, S. G. (1986) Proc. Natl. Acad. Sci. U. S. A. 83, 3791–3795
11. Aviv, H., and Leder, P. (1972) Proc. Natl. Acad. Sci. U. S. A. 69, 1408–1412
12. Dermann, E., Krauter, K., Walling, L., Weinberger, C., Roy, M., and Darnell, J. E., Jr. (1981) Cell 23, 731–738
13. Reinhart, A. P., and Vogelstein, B. (1983) Anal. Biochem. 132, 6–13
14. Corden, J., Wasylyk, B., Buchwalder, A., Sassone-Corsi, P., Keding, C., and Chambron, P. (1980) Science 209, 1406–1414
15. Austin, R. G., Patel, H. V., Gerber, G. E., Morton, R. C., and Freeman, K. B. (1986) Nucleic Acids Res. 14, 4025–4035
16. Nathans, J., and Hogness, D. S. (1980) Cell 24, 807–814
17. Liscum, L., Finer-Moore, J., Stroud, R. M., Luskey, K. L., Brown, M. S., and Goldstein, J. L. (1983) J. Biol. Chem. 258, 522–530
18. Kaufman, J. F., Aufrey, C., Koman, A. J., Shackelford, D. A., and Strominger, J. (1984) Cell 36, 1–13
19. Koblika, B. K., Matsui, H., Koblika, T. S., Yang-Feng, T. L., Franke, U., Caron, M. G., Letkowitz, R. J., and Regan, J. V. (1987) Science 238, 650–656
20. Koblika, B. K., Dixon, R. A. F., Friell, T., Dohman, H. G., Bolanowski, M. A., Sigal, I. S., Yang-Feng, T. L., Franke, U., Caron, M. G., and Letkowitz, R. J. (1987) Proc. Natl. Acad. Sci. U. S. A. 84, 46–50
21. Peralta, E. G., Winslow, J. W., Peterson, G. L., Smith, D. H., Ashkenazi, A., Ramachandran, J., Schimmerlik, M. I., and Capon, D. J. (1987) Science 236, 600–605
22. Koblika, B. K., Friell, T., Collins, S., Yang-Feng, T., Koblika, T. S., Franke, U., Letkowitz, R. J., and Caron, M. G. (1987) Nature 329, 75–79
Mitochondrial Uncoupling Protein Gene

The Mitochondrial Uncoupling Protein Gene

**Materials and Methods**

**Isolation of cDNA and Genomic Clones** — The isolation of a partial cDNA clone, p-Ucp 1, from a mouse brain cDNA library has been described (8). Recombining this library with hybridization with radioactively labeled p-Ucp 1 as a probe resulted in the isolation of additional cDNA clones, p-Ucp 2 and p-Ucp 3. The restriction enzymes map for the largest cDNA clone, p-Ucp 1, is illustrated in Fig. 1c.

Radioactively labeled p-Ucp 1 was also used to isolate genomic clones from the B10.BR mouse genome library (9) and the mouse Balb/c strain (10). Hybridizations of phage "gere" performed essentially as described by Davis and Davis (7). The entire B10.BR gene was contained in the overlapping fragments.

**Isolation of Radiolabeled p-Ucp**

The isolation of the radiolabeled p-Ucp gene was performed according to the procedure described by Leder et al. (9). The isolation of the genomic clones was performed by the method of Brown and Davis (7), using probes radiolabeled with random primers (10).

**DNA Sequencing Determination** — Coding regions within the genomic clones were mapped by double stranded radioactively labeled DNA sequencing methodology using the universal 1111 primer on sequencing gels. The clones were digested with restriction enzymes, and the fragments were hybridized to a Southern blot of HindII.

**Northern Blotting** — Total RNA was extracted from total RNA isolated by the guanidium-phenol procedure (11). Poly A+ RNA was isolated from total RNA by two passes through oligo d(T)-cellulose (11).

**Blot Analysis** — Northern blots were performed as described by Korak et al. (12) using probes radiolabeled with random primers (13).

**Phlippe Leder, personal communication.**

**Sequence Determination** — The sequence of the cDNA was determined from the sequence of the genomic clones. The nucleotide sequence of the cDNA was determined by the method of Maxam and Gilbert (14) from plaques hybridizing to cDNA (12). The isolation of a length fragment of 1.2, 1.3, and 1.6 kb as well as overlapping clones. 400 and last fragments were cloned into PUC vectors.

**Summary** — A detailed restriction enzyme map was made of the overlapping genomic clones. Restriction enzyme fragments found in the genomic clones were also observed on Northern blots of the cDNA. Restriction fragment patterns produced by Southern blotting with synthetic EcoRI fragments of 1.2, 1.3, and 1.6 kb as well as overlapping clones.

**FCS** — 2. The FCS sequence of the cDNA and genomic fragments of the mitochondrial uncoupling protein gene. The 17 splinting region is described in lower sequence letters except for the TATA box which is shaded and underlined. Box 1 starts at position 1. 17UTR sequence is sequenced until the coding region is reached, at which point the codons are numbered. The intron sequence is depicted in lower case. The poly A tail occurs at a site which has been highlighted with an asterisk.

**Poly A tail** — A poly A tail signal was identified using the poly A tail sequence and the genomic sequence. Differences between the cDNA sequence and the genomic sequence are shown at position 1 of the 17UTR region, at codons 247 and 299.