Journal of Cellular Biochemistry 43:255–264 (1990)

Human Uncoupling Protein Gene: Structure, Comparison With Rat Gene, and Assignment to the Long Arm of Chromosome 4

Anne-Marie Cassard, Frédéric Bouillaud, Marie-Geneviève Mattei, Eliane Hentz, Serge Raimbault, Muriel Thomas, and Daniel Ricquier

Centre de Recherche sur la Nutrition, CNRS, 92190 Meudon (A.-M.C., F.B., E.H., S.R., M.T., D.R.) and Centre de Génétique Médicale, INSERM U-242, Hôpital de la Timone, 13385 Marseille Cedex (M.-G.M.) France

The uncoupling protein (UCP) gene encodes a unique mammalian mitochondrial proton carrier that induces heat production in brown adipocytes. Human UCP gene was isolated and its organization analyzed. A comparison was made with rat UCP gene. Human UCP gene spans 13 Kb and contains a transcribed region that covers 9 Kb of the human genome. All of the exons were also sequenced except the extreme end of the 3′ untranslated region. Two Kb DNA upstream the TATA box were also sequenced. This region contains several fragments that are highly homologous to the gene of rat UCP. Neither CCAAT sequence nor Sp 1 binding motif were detected. Human UCP gene is split into six exons. The complete amino acid sequence of the protein was determined. Human UCP has 305 amino acids and a molecular weight of 32,786. It has no N-terminal targeting sequence. It is 79% homologous to rat UCP both at nucleotidic and amino acid levels. The primary structure of UCP is significantly homologous to the primary structure of the human T_1 ADP/ATP carrier, particularly in the C-terminal extremity, which is supposed to contain a nucleotide-binding site in both proteins. Human UCP gene is single type, as it is in rodents. Two genomic fragments were used to detect a 1.9 Kb mRNA in human perirenal brown adipose tissue. Using in situ hybridization, UCP gene was assigned in humans to chromosome 4 in q31. Interestingly, the T1 gene encoding the heart-skeletal muscle ADP/ATP carrier has recently been shown to be on the same chromosome (Li et al. Biol Chem 264:13998, 1989).

Key words: mitochondria, thermogenesis, mitochondrial carrier, proton, adipocyte

The uncoupling protein (UCP) is a characteristic component of the inner membrane of brown adipose tissue mitochondria. It has the unique property to dissipate the

Abbreviations used: AAC, ADP/ATP carrier; CRE, cyclic AMP regulatory element; Kb, kilobase; UCP, uncoupling protein.

Received September 29, 1989; accepted January 31, 1990.

proton electrochemical gradient generated by respiration and thus to induce thermogenesis in brown adipocytes in appropriate physiological conditions [1–3]. Most studies on UCP and brown adipose tissue have been developed in rodents and cDNAs for rat and mouse UCP were isolated [4–6]. An interesting finding was that UCP is related to other ubiquitous carriers, such as the mitochondrial ADP/ATP-[7,8] and phosphate-[9] carriers. The expression of UCP gene, in both rats and humans, is known to be controlled by norepinephrine [10]. Recently, the organization of the mouse UCP gene was reported [11] as well as the complete sequence of rat UCP [12]. We report here the structure of the human UCP gene and its fine chromosomal localization. Previously we had reported the isolation of a genomic fragment useful as a probe to detect human UCP mRNA [13].

MATERIALS AND METHODS

Genomic Probes

The gene encoding human UCP was isolated from an amplified genomic library using rat UCP cDNA [13]. This genomic library was a gift from Dr. A. Kahn (Paris). It was made from a partial Sau 3AI digestion of human DNA ligated to the Bam HI site of EMBL4 vector. Restriction digestions of the isolated gene were analyzed using rat UCP cDNA [4] and subcloned in PUC or M13 vectors.

DNA Sequence Analysis

DNA was subcloned into the sequencing vectors M13mp18 or M13mp19 and subjected to the dideoxy chain termination method [14]. Sequencing reactions were primed with either 15-mer or 17-mer universal primers. Both strands were sequenced. The 3′-extremity of the fourth exon was determined using a synthetic oligonucleotide primer 5′-ACT CCC AAT CTG ATG AGA AGT GTC-3′, corresponding to amino-acids 177–184. Comparison of rat and human genomic sequences was done using a matrix homology (DNA inspector II).

Preparation of DNA, RNA, and Blot Hybridization

Total genomic DNA was extracted from human liver obtained at routine autopsy held within 24 h of death. DNA was digested with restriction endonucleases and fractionated by electrophoresis on 1% agarose gels as previously described [4]. DNA was transferred to nylon filters. Brown fat RNA was extracted and analyzed as described [13]. Hybridization conditions were as described [4].

In Situ Hybridization

In situ hybridization was carried out on chromosome preparations obtained from phytohemagglutinin-stimulated human lymphocytes cultured for 72 h, and 5-bromodeoxyuridine was added for the final 7 h of culture (60 μg/ml of medium) to ensure a posthybridization chromosomal banding of good quality. In situ hybridization was carried out using the H-UCP-3.8 clone. This plasmid, containing an insert of 3,800 base pairs in PUC, was tritium-labeled by nick-translation to a specific activity of 2.8×10^7 dpm.μg^1. The radiolabeled probe was hybridized to metaphase spreads at a final concentration of 100 ng per ml of hybridization solution as previously described [15]. After coating with nuclear track emulsion (Kodak NTB_2), the slides were exposed for 14

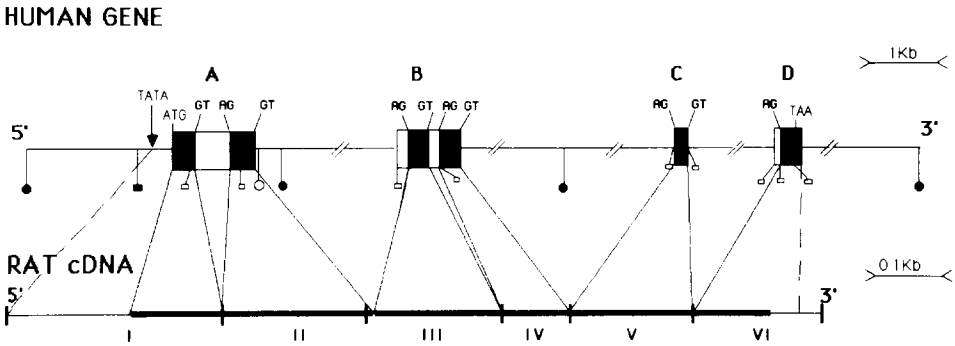

Fig. 1. Genomic organization of human uncoupling protein gene and alignment with rat cDNA [8]. Four genomic fragments A–D containing exonic (closed boxes) and intronic (open boxes) domains were sequenced. In rat cDNA, the thin line corresponds to 5′ and 3′ noncoding regions. The 3′ extremity of rat cDNA indicates the position of the first signal of poly A addition. The position in the 3′ end of the rat cDNA corresponding to the extreme 3′ nucleotide identified by sequencing of human gene is indicated by a slashed line. The noncoding domain of the first exon as well as 2 Kb of the 5′-upstream region were also sequenced, but the start site of transcription was not identified. A TATA box preceded by a GC-rich region was observed. ATG initiation codon and TAA stop codon are indicated. AG/GT consensus sequences for splice donor and acceptor sites are indicated. Restriction sites for Eco RI (●), Sau 3 AI (□), NcoI (○) and Sac I (■) are indicated. Roman numerals correspond to the six exons of rat UCP [12].

days at +4°C, then developed. To avoid any slipping of silver grains during the banding procedure, chromosome spreads were first stained with buffered Giemsa solution, and metaphases were photographed. R-banding then was performed by the fluorochrome-photolysis-Giemsa solution and metaphases rephotographed before analysis.

RESULTS AND DISCUSSION

Genomic Structure of Human UCP Gene

The isolation of human UCP gene using rat UCP cDNA has been previously reported [13]. In this preliminary study, the gene had been used only to subclone a genomic fragment H-UCP-0.5 able to detect UCP mRNA in patients with pheochromocytoma or hibernoma [13]. A partial restriction map of human UCP gene is given in Figure 1. Three Eco RI fragments (3.8, 4.2, and 5.2 Kb, from 5′- to 3′-end, respectively) form the gene. These three Eco RI fragments were detected in the cloned sequence (this study) and in a digest of human genomic DNA probed with rat UCP cDNA (not shown). The three Eco RI fragments were subcloned in PUC vectors and their SAU 3AI digestion probed with pUCP36 rat UCP cDNA [4]. This strategy led to the isolation of four genomic fragments containing exonic domains. These fragments A, B, C, and D are positioned in Figure 1. Among these fragments, B corresponds to the H-UCP-0.5 probe mentioned above. The four fragments, as well as other genomic fragments, were cloned in M13 vectors and sequenced. An alignment of genomic DNA sequences with rat UCP cDNA [8], combined with the identification of AG/GT consensus sequences for splice donor, and acceptor sites revealed that the human UCP gene at least contained six exons scattered over a 13-Kb genomic domain (Fig. 1). The first exon covers the 5′-untranslated region and the first 40 amino acids; the second exon extends from amino acids 41 to 107, the third from amino acids 108 to 173, the fourth from amino acids 174

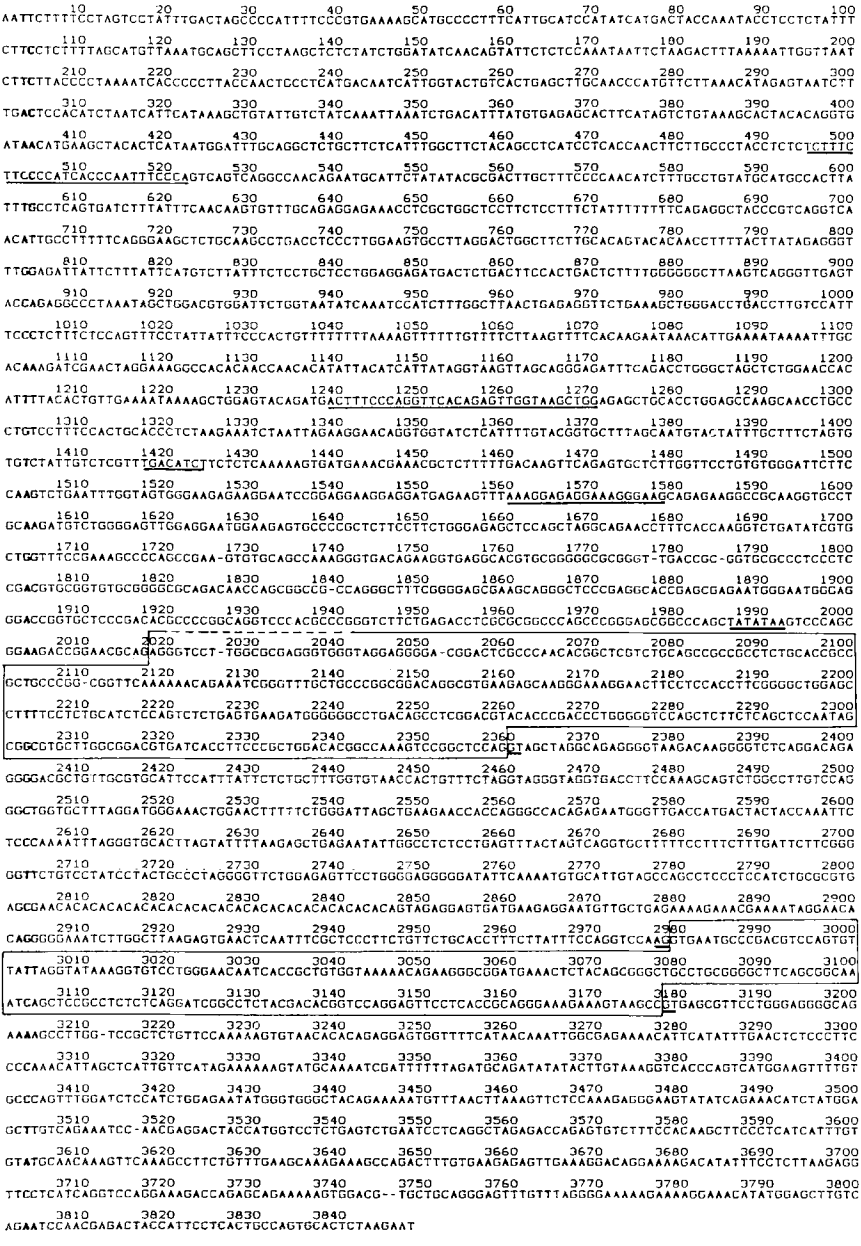

Fig. 2. Nucleotide sequence of the 3.8 Kb 5′ domain of human uncoupling protein gene. On the right-hand side of the figure, double arrows 1–4 indicate underlined sequences highly conserved in human and rat UCP genes. A presumed TATA sequence in underlined. The two first exons are boxed. The interrupted line above the 5′ extremity of the first exon corresponds to the presumed limit. Consensus splite sites (dinucleotides AG and GT) at the intron/exon boundaries are underlined. The dash at various positions indicates uncertainty that could not be removed in spite of repeated analysis. This H-UCP-3.8 genomic fragment was used to determine the chromosomal localization of the gene (Fig. 7).

to 207, the fifth from amino acids 208 to 267, and the last exon encodes the remaining 38 amino acids and the 3′ untranslated region. This structure is similar to that observed with mouse and rat UCP gene [11,12].

Two Kb at the 5′-end of the gene were also sequenced, and the nucleotide sequence of the 3.8 Kb 5′ Eco RI fragment of the gene is given in Figure 2. A presumed TATA box was observed, but no CCAAT sequence was detected (Fig. 2). The TATA box is preceded by a 0.2 Kb G + C rich domain (71% nucleotides). However, no Sp 1 protein binding motif is present either in the 5′-upstream domain of the gene or in the first intron. The difficulty in obtaining high-quality poly (A^+) RNA from human brown adipose tissue prevented mapping the transcription start site as done for the mouse and rat UCP gene [11,12]. A comparison of the 5′-upstream domain of rat [12] and human genes was made (Fig. 3). A significant homology was found when 2 Kb of each gene were compared (left part of Fig. 3). The level of homology was particularly high when the region −1281 to −1253 of rat gene was compared with the region −1491 to −1463 of human gene (right part of Fig. 3). This small domain, of which the sequence is given in Figure 2 (arrows 1), could identify a sequence involved in the regulation of UCP gene in both species. Moreover, at position −565 in rat gene and −569 in human gene, the same 5′-TGACATCT-3′ sequence is present (arrows 3 in Fig. 2). Interestingly, this sequence is 85% homologous to the consensus CRE sequence. It suggests a role for cyclic AMP in the regulation of UCP gene expression in rats and humans. The identification of regulatory elements in rat and human UCP gene is presently under study.

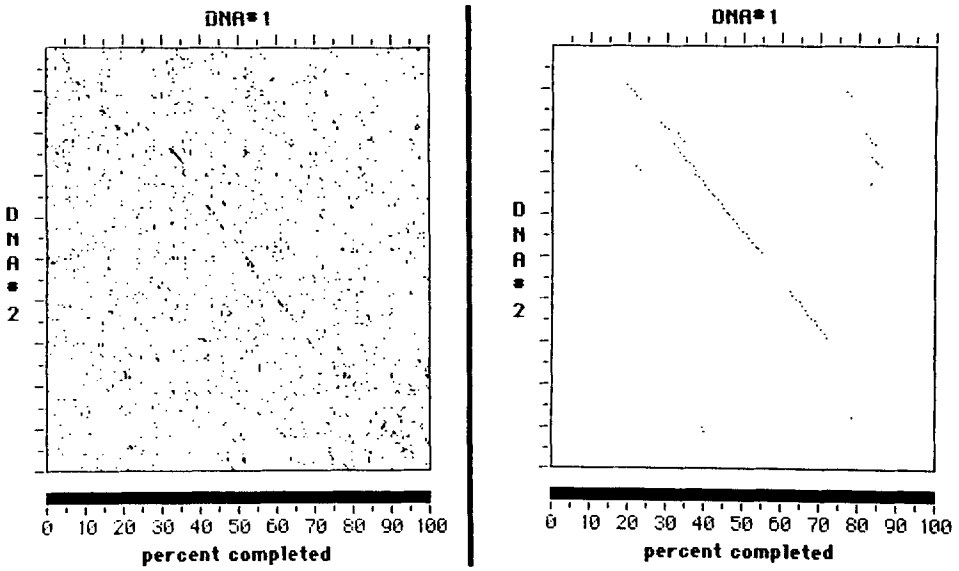

Fig. 3. Comparison of 5′ end of human and rat uncoupling protein genes. An homology matrix analysis was made (window: 10 nucleotides; maximum of mismatches: 2). DNA 1 and DNA 2 refer to rat and human sequence, respectively. The left part of the figure shows a study on 2 kb. The right-hand side of the figure corresponds to the region −1281 to −1253 of rat gene compared with the region −1491 to −1463 of human gene (the first T of the TATA box of both genes was taken as position +1). The sequence of this conserved domain is shown in Figure 2 (arrows 1).

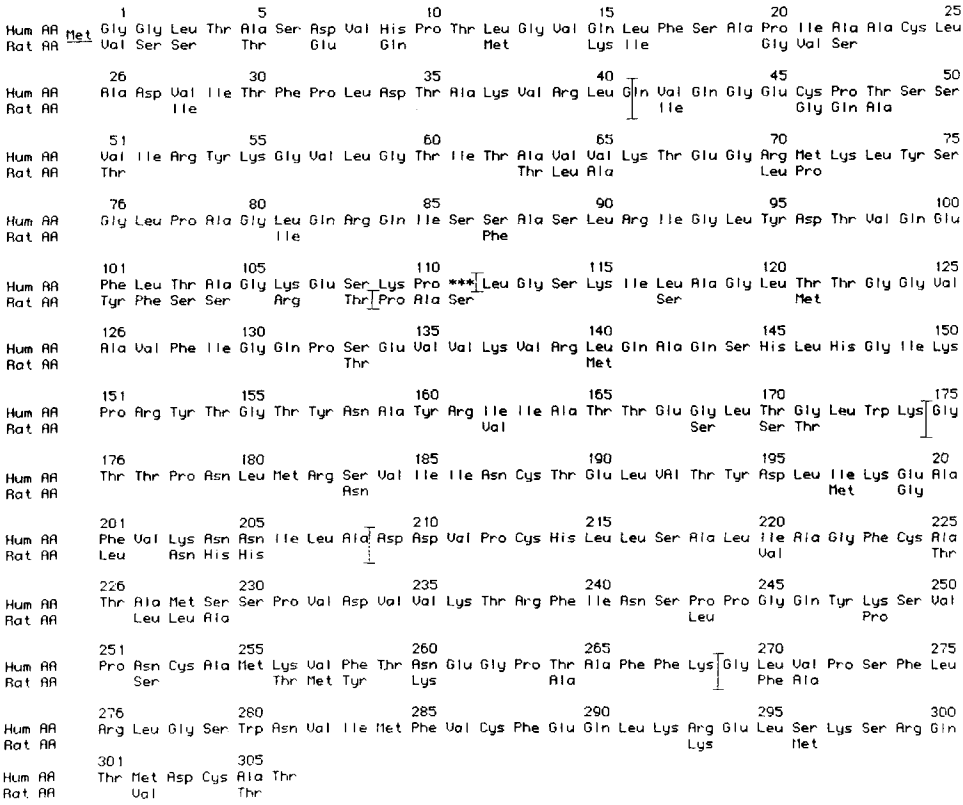

Fig. 4. Nucleotide-deduced amino acid sequence of human uncoupling protein and comparison with rat UCP primary structure [8]. For rat sequence, only the residues that differ from human UCP are indicated. The limits between exons are indicated by vertical and horizontal lines in rat and human sequences. The three stars at position 111 correspond to a deleted amino acid in human UCP.

Primary Structure of Human UCP and Comparison With Human ADP/ATP Carrier

The complete amino acid sequence of human UCP was derived from the nucleotidic sequence and the organization of exons (Fig. 4). The human UCP has 305 amino acids, a molecular weight of 32,786 and a polarity index of 41. The rat UCP has 306 amino acids, a molecular weight of 33,042, and a polarity index of 42% [8]. The human UCP has no N-terminal targeting sequence. The same observation was previously reported for rodents [3,5,7,8]. Rat and human exons share a 79% homology at both nucleotidic and amino acid levels.

It has been previously reported that rodent UCP and bovine or human mitochondrial ADP/ATP carrier (AAC) are partially homologous at the level of their primary structure [7,8,16–18] and that they share a similar organization made of three internal repeats [7]. A comparison of the three repeats of human UCP and human AAC [19] is shown in Figure 5. These data confirmed that the two proteins have the same organization and probably derive from a common ancestor. Interestingly, as previously noticed for rat [8] and bovine UCP [20,21], the highest degree of homology was found in the C-terminal region of the two proteins and particularly at the level of an ADP-binding site

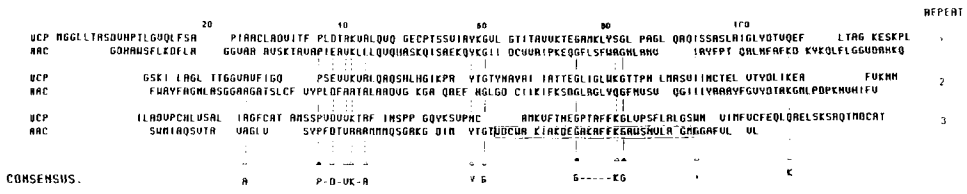

Fig. 5. Alignment of human uncoupling protein (UCP) and human skeletal muscle ADP/ATP carrier (AAC) [19,24] primary structures. Repeats 1, 2, and 3 are N-terminal, central, and C-terminal, respectively. Identities present in the six sequences are indicated by closed triangles and solid lines. Identities present in at least four of the six sequences are indicated by open triangles and dotted lines. The cytosolic-sided domain in AAC identified as an ADP-binding site [22] is boxed.

(boxed fragment in Fig. 5) recently identified in bovine AAC [22]. It is known that the proton translocating activity of UCP can be fully inhibited when purine nucleotides are bound to UCP [1]. This C-terminal region of UCP, which is highly homologous to AAC, could thus be implicated in the binding of nucleotides to UCP.

Southern and Northern Analysis of Human UCP

A preliminary study [13] of DNA from three individuals with H-UCP-0.5 probe (B probe in Fig. 1) suggested that there is a single gene for human UCP, as is known for rodents [4,6,11,12]. Southern analysis of human DNA from three other individuals is shown in Figure 6. The A genomic probe (Fig. 1) hybridized to the 3.8 Kb Eco RI band from which it is derived (lanes 1–3 in Fig. 6), whereas the 0.5 Kb B probe detected a 4.2 Kb Eco RI band (lanes 5 and 8 in Fig. 6). It can be concluded that the gene is single type in humans, and it exhibits no polymorphism, at least in this limited study. This is in contrast with what is known for human AAC, which is encoded by at least three different genes [19,23,24]. Another major difference between UCP and AAC genes is the existence of six exons in the mouse [11], rat [12], and human UCP genes compared with four exons in human AAC T1 gene [19,24] and AAC T2 gene [24]. Interestingly, in rat [12] and human UCP genes (Fig. 4) and in human AAC gene [19,24], the first intron interrupts the coding sequence at the same position (Leu G/ln Val Gln for UCP gene and Leu Gln/Val Gln for AAC gene). A similar observation had been made previously when comparing rat UCP gene with the AAC gene of *Neurospora crassa* [12].

UCP genomic probes A and B could easily detect a 1.9 Kb mRNA in Northern analysis of RNA extracted from human perirenal brown adipose tissue (bottom of Fig. 6). In rodents, UCP has two mRNAs of 1.5 and 1.85 Kb length [3–6]. We have demonstrated that two polyadenylation sites in rat UCP gene explain the existence of two mRNAs [12]. We have also reported that newborn calves and lambs have only one 1.9 Kb mRNA encoding UCP and that the first polyadenylation site of rodents is lacking in bovine UCP gene [25]. Although the complete sequence of the 3′-end of the human UCP gene was not determined in the present study, it may be suggested that the presence of only one mRNA species in humans is also due to a single polyadenylation site in human UCP gene.

Human Gene Mapping by In Situ Hybridization

In situ hybridization was done with H-UCP-3.8 genomic probe (Fig. 7), which is the 3.8 Kb 5′ Eco RI fragment of human UCP gene (Fig. 1). Its complete sequence is shown in Figure 2. In genomic DNA analysis (data not shown), this probe gave exactly

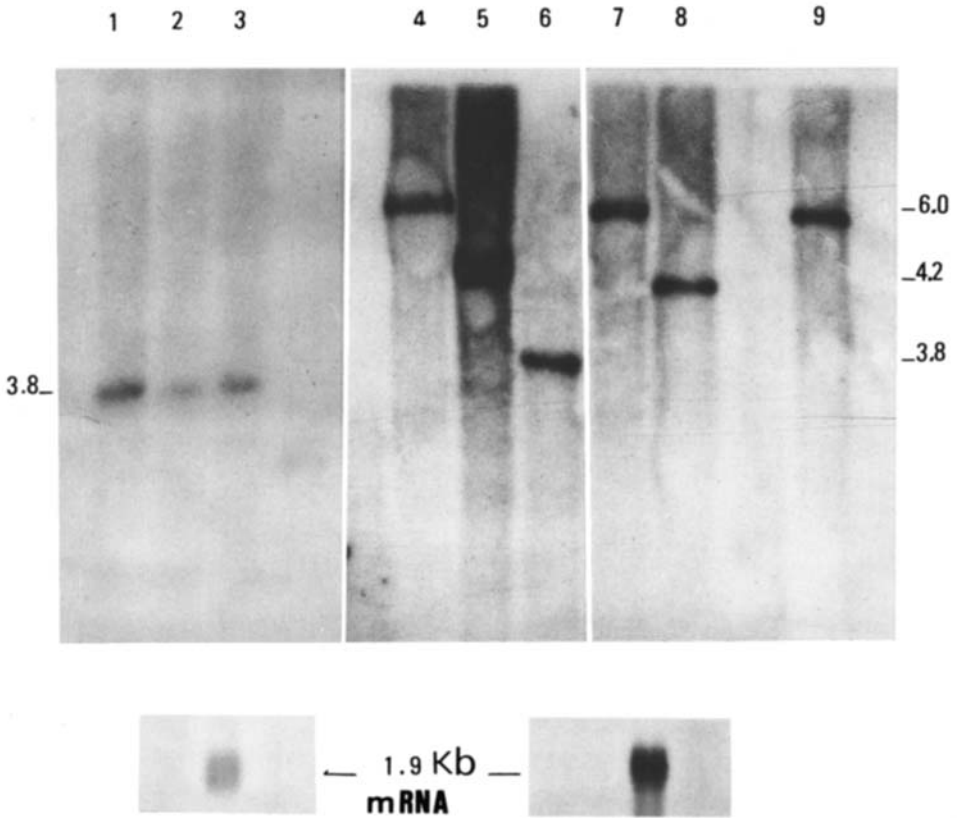

Fig. 6. Southern and Northern analysis of human uncoupling protein. Genomic DNA (20 μg) prepared from nongenetically-related individuals was cut by restriction enzymes Eco RI (lanes 1–3, 5, and 8), Hind III (lanes 4, 7 and 9) or TaqI (lane 6). Hybridization was made with genomic A probe (lanes 1–3) or B probe (lanes 4–9). These two probes correspond to fragments A and B in Figure 1. The sizes (in Kb) of the human genomic fragments are indicated at the left and right parts of the autoradiograph (estimated from Hind III digested λ DNA). At the bottom of the figure, the two inserts correspond to human UCP mRNA detected in Northern analysis of 5 μg of poly A^+-RNA extracted from human perirenal brown adipose tissue. No UCP mRNA was detected in liver poly A^+-RNA. Hybridization was made with A probe (left-hand side insert) or B probe (right-hand side insert).

the same signal as that obtained with the A probe (left-hand side of Fig. 6). In the 100 metaphase cells examined after in situ hibridization, there were 248 silver grains associated with chromosomes, and 41 of these (16.5%) were located on chromosome 4; the distribution of grains on this chromosome was not random: 73% mapped to the [q28–q31] region of chromosome 4 long arm with a maximum in the 4q31 band.

These data allow us to assign H-UCP-3.8 probe to the q31 band of the human chromosome 4. Jacobsson et al. [6] detected a Bam HI restriction fragment polymorphism for mouse UCP gene, and a study of this polymorphism in genomic DNA of recombinant inbred mouse strains allowed them to map the UCP gene to chromosome 8 of mouse. This observation, compared with the present assessment of UCP gene to chromosome 4 of humans, could indicate homologies between these two murine and human chromosomes. It is presently known that the proximal as well as the distal regions

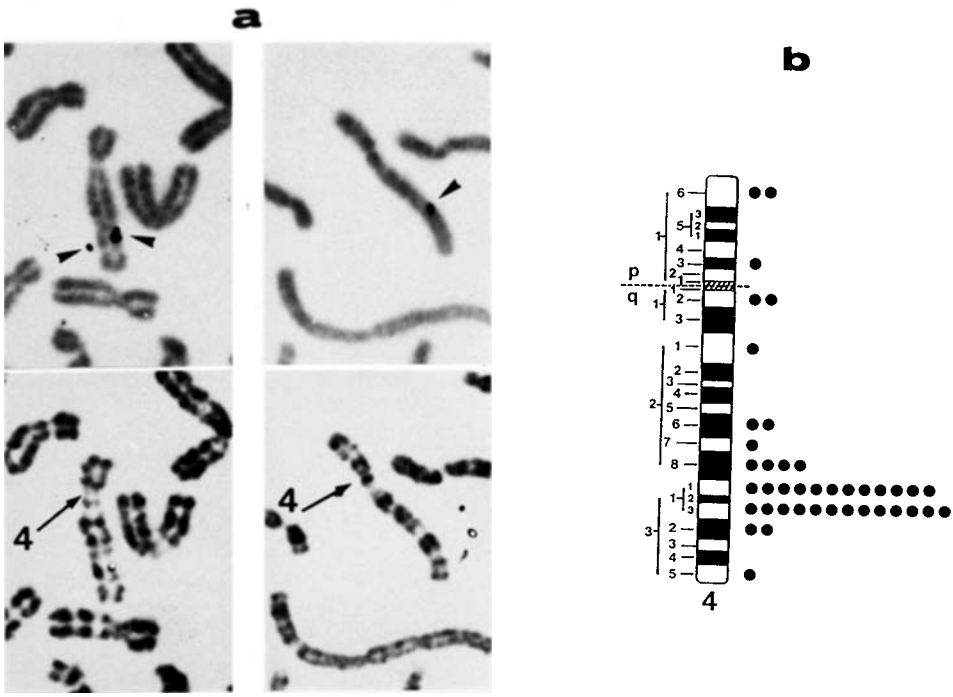

Fig. 7. Chromosomal localization of human uncoupling protein gene. **a:** Two partial human metaphases showing the specific site of hybridization to chromosome 4. Top: arrowheads indicate silver grains on Giemsa-stained chromosomes, after autoradiography. Bottom: chromosomes with silver grains were subsequently identified by R-banding (FPG technique). **b:** Ideogram of the human G-banded chromosome 4 illustrating the distribution of labelled sites for H-UCP-3.8 probe. The sequence of the genomic probe is given in Figure 2.

of the long arm of human chromosome 4 exhibit homologies to murine chromosomes 5 and 3, respectively [26]. Chromosome 4 is a large chromosome in humans and a region of human genome for which there are very few finely localized markers [27]. In addition to UCP gene, the only other gene markers that have been localized to or near 4q31 are fibrinogen alpha, beta, and gamma polypeptides, erythrocyte glycophorin, intestinal fatty acid binding protein FABP2, and lipocortin IIa [27]. In other respects, genetic defects involved in anterior segment mesenchymal dysgenesis and in sclerotylosis have been mapped to 4q28-q31 [28]. Recently, Blanquet et al. [29] have also reported the regional mapping to 4q32 of a DNA domain rearranged in human liver cancer. Whether genomic probes for human UCP will be helpful in studies on human genetic diseases remains to be investigated.

Human Uncoupling Protein and Human Muscle Adenine Translocator Are Localized on the Same Chromosome

In a very recent study, Wallace and his colleagues have reported the chromosomal mapping of the T1 gene encoding the human ADP/ATP carrier expressed in heart and skeletal muscle [19]. They could demonstrate, using both flow-sorted chromosomes and segregating human-mouse hybrid cells, that this gene is located on chromosome 4. Thus,

UCP and AAC genes belong to the same chromosome. The fine localization of AAC gene on the chromosome 4 has not been reported yet. Colocalization of the two genes on the same chromosome could confirm their common origin and stimulates further studies on the evolution of multigene family of mitochondrial carriers.

ACKNOWLEDGMENTS

The authors wish to thank Professors P. Callard and J. Bienaymé for help in obtaining human samples, N. Roeckel and S. Guidi for technical assistance in chromosomal mapping, Dr. O. Champigny for careful reading of the manuscript, and P. Méralli for secretarial assistance. This work was financially supported by CNRS, INSERM, Ministère de la Recherche et de l'Enseignement Supérieur and Direction des Recherches, Etudes et Techniques.

REFERENCES

1. Nicholls DG, Locke R: Physiol Rev 64:1, 1984.
2. Cannon B, Nedergaard J: Essays Biochem 20:110, 1985.
3. Ricquier D, Bouillaud F: In Trayhurn P, Nicholls DG (eds): "Brown Adipose Tissue." London: Edward Arnold, 1986, p 86.
4. Bouillaud F, Ricquier D, Thibault J, Weissenbach J: Proc Natl Acad Sci USA 82:445, 1985.
5. Ridley RG, Patel HV, Gerber GE, Morton RC, Freeman KB: Nucleic Acids Res 14:4025, 1986.
6. Jacobsson A, Stadler U, Glotzer A, Kozak LP: J Biol Chem 260:16250, 1985.
7. Aquila H, Link TA, Klingenberg M: EMBO J 4:2369, 1985.
8. Bouillaud F, Weissenbach J, Ricquier D: J Biol Chem 261:487, 1986.
9. Runswick RJ, Powell SJ, Nyren P, Walker JE: EMBO J 6:1367, 1987.
10. Ricquier D, Bouillaud F, Casteilla L, Cassard AM, Raimbault S, Klaus S, Champigny O, Hentz E: In, Bray G, Ricquier D, Spiegelman B (eds): "Obesity: Towards a Molecular Approach," UCLA Symposia Proceedings, 1989, in press.
11. Kozak LP, Britton JH, Kozak UC, Wells JM: J Biol Chem 263:12274, 1988.
12. Bouillaud F, Raimbault S, Ricquier D: Biochem Biophys Res Commun 157:783, 1988.
13. Bouillaud F, Villarroya F, Hentz E, Raimbault S, Cassard AM, Ricquier D: Clin Sci 75:21, 1988.
14. Sanger F, Nicklen S, Coulson AR: Proc Natl Acad Sci USA 74:5663, 1977.
15. Mattei MG, Philipp N, Passage E, Moisan JP, Mandel JL, Mattei JF: Hum Genet 69:268, 1985.
16. Battini R, Ferrari S, Kaczmarek L, Calabretta B, Chen S, Baserga R: J Biol Chem 262:4355, 1987.
17. Neckelman N, Li K, Wade RP, Shuster R, Wallace DC: Proc Natl Acad Sci USA 84:7580, 1987.
18. Houldworth J, Attardi G: Proc Natl Acad Sci USA 85:377, 1988.
19. Kang L, Warner CK, Hodge JA, Shinsei M, Kudoh J, Fukujama R, Masahiko M, Shimizu Y, Shimizu N, Wallace DC: J Biol Chem 264:13998, 1989.
20. Bouillaud F, Raimbault S, Casteilla L, Cassard AM, Ricquier D: In Azzi A, Nalecz A, Nalecz MJ, Wojkczd KL (eds): "Anion Carriers of Mitochondrial Membranes." Berlin: Springer-Verlag, 1989, p 251.
21. Casteilla L, Bouillaud F, Forest C, Ricquier D: Nucleic Acids Res 17:2131, 1989.
22. Dalbon P, Brandolin G, Boulay F, Hoppe J, Vignais PV: Biochemistry 27:5141, 1989.
23. Powel SJ, Medd SM, Runswick MJ, Walker SE: Biochemistry 28:866, 1989.
24. Cozens AL, Runswick MJ, Walker JE: J Mol Biol 206:261, 1989.
25. Casteilla L, Champigny O, Bouillaud F, Robelin J, Ricquier D: Biochem J 257:665, 1989.
26. Mouse News Letter: Oxford University Press 78:12, 1987.
27. Gusella JF, Wasmuth JJ: Human Gene Mapping 9. Cytogenet Cell Genet 46:132, 1987.
28. Mc Kusick VA: "Mendelian Inheritance in Man." Baltimore: John Hopkins University Press, 1987.
29. Blanquet V, Garreau F, Chenivesse X, Brechot C, Turleau C: Hum Genet 80:274, 1988.