Human muscle carbonic anhydrase: gene structure and DNA methylation patterns in fetal and adult tissues

Julie Lloyd, Carol Brownson, Susan Tweedie, Jillian Charlton, and Yvonne H. Edwards

MRC, Human Biochemical Genetics Unit, The Galton Laboratory, University College London, London NW1 2HE, UK; ~Department of Biological Sciences, City of London Polytechnic, London E1 7NT, UK

We report the isolation and analysis of genomic clones comprising the entire gene coding for the human muscle carbonic anhydrase, CAIII. The gene spans 10.3 kb and has a seven-exon/six-intron structure. A noncanonical TATA box, a CCAAT motif, and two CCGCCC elements are present in the sequences upstream of exon 1. Although the expression of CAIII shows strict tissue specificity, the gene exhibits a number of features normally associated with housekeeping enzymes. For example, there is 48% homology with a 25-bp consensus sequence between the TATA box and the cap site and there is a CpG-rich island spanning a 469-bp sequence near to the origin of transcription. Methylation studies suggest that some CCGG sites clustered in the CpG-rich island are undermethylated in DNA from fetal and adult muscle and in other tissues irrespective of CAIII expression. In contrast, several nonclustered CCGG sites show a methylation pattern that correlates with gene expression. However DNA from differentiated type II adult muscle fibers is undermethylated at these sites even though CAIII is not expressed.

[Key Words: Carbonic anhydrase; tissue-specific gene; CpG-rich islands; methylation; muscle development]

Received April 21, 1987; revised version accepted June 8, 1987.

Carbonic anhydrase III (CAIII, carbonate dehydratase EC 4.2.1) is a monomeric zinc metalloenzyme catalyzing the reversible hydration of carbon dioxide. In man, CAIII is an abundant muscle-specific protein (Carter et al. 1979) involved in the maintenance of ionic balance and acid–base homeostasis. CAIII is characteristic of slow-twitch, aerobic, type I muscle fibers. The structural gene for CAIII is part of a multigene family, other members of which show tissue- and organelle-specific expression. CAI activity is confined largely to erythrocytes. CAI is also a major protein in erythrocytes but is expressed in a wide variety of other tissues. Distinct mitochondrial and membrane isoforms of CA have also been described. Protein and cDNA nucleotide sequence analyses of CAI, CAII, and CAIII have shown that these genes are evolutionarily related [for a review of CA see Tashian and Hewett-Emmett 1984]. These three genes in man have been assigned to chromosome 8 (Venta et al. 1983; Edwards et al. 1986a,b), and more recently it has been shown that CAI and CAIII are in the same region of this human chromosome at q22 (Davis et al. 1987).

The CAIII gene appears to be differentially regulated both in adult tissues where high expression is confined to one cell type, and in fetal muscle at different stages of development [Jeffery et al. 1980, Lloyd et al. 1986]. In human limb muscle, trace amounts of CAIII protein and mRNA are detected at 10 weeks gestation. These levels increase to approximately 20% of mature adult muscle levels by 20 weeks and to about 50% by birth. CAIII expression can also be manipulated experimentally by muscle denervation [Wistrand et al. 1986] or electrostimulation [Edwards et al., in prep.], procedures that alter the contractile properties and fiber type distribution of the muscle. An unusual situation occurs in rat tissues where CAIII is expressed in liver as well as skeletal muscle. In the liver the CAIII protein levels exhibit a sexual dimorphism which is apparently regulated by growth hormone [Jeffery et al. 1986].

To provide probes for the analysis of the regulation of the CAIII gene, we have recently isolated a cDNA clone, pCA15, encoding the entire sequence of human muscle CAIII [Lloyd et al. 1986]. In the present study this cDNA clone was employed to isolate the CAIII gene and to elucidate its structure. Since methylation of cytosine residues has been implicated in gene regulation [see Razin et al. 1984; Bird 1986], we have investigated the pattern of methylation in DNA from CAIII expressing and nonexpressing tissues, using both cDNA and genomic DNA probes.

Results and discussion

Isolation of the CAIII gene: mapping of exons and introns

The recombinant clones λCA1.1, 1.2, and 1.3 shown in Figure 1 were isolated by screening a genomic library in
Chargon 4A (Lawn et al. 1978) using the CAIII cDNA pCA15 (Lloyd et al. 1986) as a probe. The most 5' 2.2-kb EcoRI fragment from CAI.1 was subcloned into pUC8 (pCA 2.2) and used to screen the Charon 4A library and another human genomic library in x0001 (LeFranc et al. 1986) to derive the CAIII recombinants CA1.2 and CA2.2. An EcoRI, HindIII, and TaqI restriction enzyme map extending across a 22-kb region of genomic DNA and encompassing the whole of the CAIII gene was derived by a combination of single and double digests of DNA from the various CA genomic clones. All of the EcoRI and TaqI fragments from CA1.1 and CA2.2, shown in Figure 1B, were subcloned into pUC8. Exon sequences were localized by hybridization of restriction enzyme digests (HhaI, RsaI, Avall, Psrl) of these subclones to the 32P-labeled CAIII cDNA clone and to DdeI fragments derived from the 5' end (278 bp), the center (576 bp), and the 3' end (419 bp) of the cDNA clone (see Lloyd et al. 1986 for cDNA sequence).

In some cases further subclones were prepared which contained smaller DNA inserts cross-hybridizing to the cDNA. The DNA inserts from the various subclones (Fig. 1A) were sequenced from both ends to obtain the precise location of each of the exon–intron boundaries. In addition, the 2.2-kb 5'-terminal EcoRI fragment of CA1.1, containing exon 2, and the 2.8-kb 5'-terminal EcoRI fragment of CA2.1 were sequenced in full using the dideoxy method and 57 independent, randomly generated M13 subclones (hatched boxes in Fig. 1A).

These analyses showed that the gene for human CAIII is about 10.3 kb long and comprises seven exons. The sequences at the 5' and 3' junction of each intron are shown in Figure 2. All the introns, which vary in size from 750 bp (IVS 4) to 2360 bp (IVS 2), conform to the consensus eukaryotic structure beginning with the dinucleotide GT and ending with AG. Exon 1 contains 59 bp of 5' noncoding sequence and 34 bp of coding sequence. The most 3' exon contains 147 bp of coding sequence and 877 bp of 3' noncoding sequence. The other exons vary in size between 63 and 198 bp.

The overall architecture of the muscle carbonic anhydrase gene CAIII appears to be very similar to that described for the reticulocyte gene CAII isolated from the mouse (Venta et al. 1985). The CAII gene is also interrupted by six introns. The sizes of IVS 5 and 6 are similar in the two genes but the sizes of the other introns show considerable variation; in particular, IVS 2, which is 2.4 kb in the CAIII gene, is 7.2 kb in the CAII gene. The positions of five of the introns are identical to those found in CAII. However, IVS 4 which in the CAIII sequence occurs between codons 147 (Lys) and 148 (Ile) is found 14 bp upstream in CAII where it interrupts the codon for Gly-143. This difference between CAII and CAIII is unexpected and may represent a peculiarity of the mouse CAII sequence since there is some recent evidence that human CAI and CAII (Venta et al. 1987; R. Tashian, pers. comm.) and chicken CAII (Yoshihara et al. 1987) have IVS 4 in the same position as described here for CAIII. In general, intron positions within multigene families are conserved between homologous mammalian and avian genes. Although long evolutionary distances may lead to insertion or loss of introns, it is difficult to envisage a mechanism by which a shift in the position of an intron might occur.

3' End of the CAIII gene

The nucleotide sequence extending 270 bp downstream of the poly[A] addition site was determined. The sequence TTGTGGTTTTT is found 25 bp downstream of the AATAAA polyadenylation motif and this fits the consensus sequence YGTGTTYY discussed by McLauchlan et al. (1985), which is thought to be required for efficient formation of mRNA 3' termini. This sequence motif is
Lloyd et al.

**Figure 2.** Structural details of the CAIII gene. Each exon–intron junction is illustrated by giving on the left the 3' end of an exon and at the right the 5' sequences beginning the next exon (capital letters). The corresponding amino acids are given above numbered in terms of amino acid sequence. The beginning and ending of each intron are shown in lowercase lettering. The sizes of each exon and intron (IVS) are given in parentheses. The length of exon 1 is given as 93 bp based on the assessment of the 5' end of the mRNA (see text).

**Figure 3.** Nucleotide sequence at the 5' end of the CAIII gene. Numbering starts at the cap site. Exon 1 sequence is given in uppercase lettering. The ATG initiation codon and two CCCGCC elements are underlined. The noncanonical TATA sequence and the CCAAT motif are boxed. The MspI–HpalI restriction sites, CCGG, are overlaid and correspond to sites 1–8 referred to in the text and Fig. 5A.
been found in the 5' region of a number of eukaryotic genes and suggest that it is part of a regulatory system. This motif is thought to be part of the binding site for the Spl transcriptional factor. This motif has not accounted for these sequences in CAIII, since this gene is not expressed constitutively.

Part of the sequence at the 5' end of the CAIII gene [1074 bp] is shown in Figure 3. The transcription initiation site was identified using a 191-bp genomic probe, derived from a 600-bp EcoRI–HindIII fragment [Fig. 1A], which extended from the EcoRI site (−131 in Fig. 3) to a Ball site at the extreme 5' end of the coding sequence (+61 in Fig. 3). This fragment was cloned into pSP65 and the [32P]UTP-labeled antisense strand used to hybridize to poly(A)+ RNA from human and rodent skeletal muscle. An RNase-protected fragment of 63 bp was identified in RNA from human muscle [Fig. 4] and comparison with the Ball–EcoRI probe nucleotide sequence shows that the 5' end of this fragment corresponds to an A residue 16 bases upstream of the start of the cDNA. Protected fragments were not found in the rodent samples. Presumably the homology between the human and rodent 5' untranslated sequences is not sufficient to allow the formation of stable hybrid molecules. Protected fragments were not formed between human muscle RNA and the sense strand of the Ball–EcoRI probe (data not shown).

Two of the elements normally associated with genes transcribed by RNA polymerase II were identified. A putative TATA box, CATATAA, was found at position −22 relative to the CAP site and the consensus sequence CCAAT occurs at position −92.

The upstream sequences of the CAIII gene exhibit a number of features which are considered to be characteristic of housekeeping genes and which are not usually associated with genes expressed at high levels in restricted cell types. For example, there is fairly good alignment, 48% with the 25-bp consensus sequence found downstream of the TATA box in a number of housekeeping enzymes [for review, see Martini et al. 1986]:

**consensus** TATAAA\(\text{GGGGCCGCGCGCGCG}\)\(\text{CGGGTGGGCGCGCGG}\)\(\text{CAIII}\)\(\text{CAAAAGGCGCGGGCTCAGCGGACTTCTGCA}\)

In addition the hexanucleotide CCGCCC occurs twice within 170-bp upstream of the cap site of the CAIII gene and in one case is an exact inverted complement of the Sp1 binding decanucleotide \(\text{GGGGCCGGGCGG}\)\(\text{AAAT}\) (Fig. 3). The core sequence CCGCCC occurs twice in the SV40 genome and is thought to be part of the binding site for the Sp1 transcriptional factor. This motif has been found in the 5' region of a number of eukaryotic genes [for review, see Kadonga et al. 1986], and Martini et al. [1986] have noted the occurrence of multiple copies in the promoter regions of several housekeeping genes and suggest that it is part of a regulatory system that defines their constitutive expression. Such rationale does not account for these sequences in CAIII, since this gene is not expressed constitutively.

The sequences flanking the cap site of the CAIII gene are highly G/C rich. A 469-bp sequence extending from −317 to +152 contains 65% G + C [compared with 40% for bulk DNA] and contains 32 CpG and 49 GpC dinucleotides [Fig. 4]. This sequence has many of the characteristics of the HpaII-tiny-fragment (HTF) islands of nonmethylated CpG-rich DNA reported by Bird et al. [1985]. It has been proposed that the CpG-rich sequences are peculiar to housekeeping genes [Bird 1986]. However this view is gradually being eroded by the identification of CpG islands in tissue-specific genes such as the retinol-binding protein [d’Onofrio et al. 1985], the Thy-1 antigen [Kolsto et al. 1986], α-globin [Bird et al. 1987], and now CAIII.

**Methylation of CpG residues in and around the CAIII gene**

Several lines of evidence suggest a role for DNA methylation in eukaryotic gene regulation and, broadly speaking, two different patterns of methylation seem to be implicated. First, analyses of CpG sites that are widely spaced in and around a particular gene indicate that most tissue-specific genes are nonmethylated in tissues where they are expressed and heavily methylated in nonexpressing tissues and in germ line cells [for review, see Yisraeli and Szyf 1984]. However, for CpG...
sites which occur in clusters at the 5' end of genes near the origin of transcription (HTF islands), a rather different situation seems to be emerging (for review, see Bird 1986). Here it is predicted that the CpG clusters represent regions of the genome consistently protected from methylation by bound factors and that they will be nonmethylated in germ cell DNA. There is also evidence showing that methylation of the CpG clusters inactivates the associated gene. Methylation of cytosine in CCGG sequences can be investigated using the CpG methylation-sensitive enzyme HpaII and its CpG methylation-insensitive isoschizomer MspI. Published data derived using methylation-sensitive restriction enzymes are commonly confined to the widely spaced CpG sites. The clustered CpGs are characterized by closely spaced sites and the digested fragments are often small (50-150 bp) and more difficult to analyze by conventional Southern blot analysis (Bird et al. 1985).

In the present study a total of 13 MspI–HpaII sites have been identified by a combination of sequence analysis and restriction enzyme mapping of subcloned fragments [Fig. 1]. Four sites are in the 3' half of the gene; two occur in exon 6 separated by 8 bp, another in intron 6, and the fourth in the 3'-flanking region 2 kb downstream of exon 7. There is a single site in the distal 5' region, 4 kb upstream of the cap site. The remaining eight sites are clustered in a 1074-bp region (sequence and MspI sites given in Fig. 3) which includes exon 1 and a region 5' to it [Fig. 5A]. The positions of these closely spaced sites were confirmed by MspI digestion of the 3.4-kb HindIII fragment, pCA3.4 [Fig. 1], and its constituent 2.6-kb and 0.6-kb EcoRI fragments [Fig. 5A]. The sizes of the constituent MspI fragments [Fig. 5B] agree exactly with those expected from the sequence [given in Fig. 3]. The MspI–HpaII sites are numbered 1–11 in Fig. 1 [excepting the most 5' site, which is not discussed] for easy reference. Since the two sites in exon 6 are very close, they are referred to collectively as 9.

**Methylation of clustered CpG sites**

Using a subcloned 3.4-kb HindIII fragment from λ2.2 pCA3.4 [Fig. 1C] as probe, a preliminary analysis has been made of the pattern of methylation in the CpG-rich cluster at the 5' end of the CAIII gene, MspI–HpaII sites 1–8. HindIII–MspI or HindIII–HpaII digestion of human DNA gives rise to a major fragment of 2.3 kb hybridizing to the probe that represents the HindIII–MspI fragment immediately adjacent and upstream to the MspI site 1 [Figs. 5A and 6A]. The size of this fragment does not vary among the adult tissues studied or during muscle development [Fig. 6A]. This implies that the most 5' site at position –622 relative to the cap site, is always unmethylated in tissue DNA irrespective of gene expression. Using the probe labeled to high specific activity [10⁹ cpm/μg] and long exposures [10 days] of autoradiographs, smaller fragments from the cluster [Fig. 6B] can be detected. The most prominent is about 350 bp and shows a size variation between MspI and HpaII digests in nonexpressing adult tissues and in fetal muscle DNA but not in adult muscle DNA. This size variation is very small, between 20 and 35 bp, and was best resolved after electrophoresis on 1.3% agarose gels. In addition, tissues that never express CAIII exhibit a fragment of about 600 bp in HindIII–HpaII digests that is more weakly seen or absent from the HindIII–MspI digests. This fragment is much more intense relative to the 350-bp fragment in placenta, liver, and kidney than in fetal or adult muscle. An invariant fragment of about 250 bp is seen in digests of all human DNA samples. DNA fragments smaller than 250 bp were not detected in these experiments, although very recent studies using denaturing acrylamide gels and an electroblotting procedure indicate that frag-

**Figure 5.** The MspI–HpaII cluster at the 5' end of the CAIII gene. (A) Map of the MspI–HpaII sites in the 3.4-kb HindIII fragment. The 0.6- and 2.8-kb fragments derived from the 3.4-kb fragment by EcoRI digest are shown below. (B) Restriction fragment patterns seen after MspI digestion of the 3.4-, 2.8-, and 0.6-kb fragments, 32P-end-labeling, and electrophoresis on a 6.5% nondenaturing acrylamide gel. The weak components seen in the 2.8- and 0.6-kb tracks are due to partial digestion. Lane M is a 32P-end-labeled MspI digest of pBR322 DNA.
ments of around 100 bp and 60 bp are present in most tissue DNA tested (data not shown). Firm identification of the various DNA components should emerge from a detailed analysis using fragment-specific probes derived from the 3.4-kb HindIII fragment. However, one possible explanation for the preliminary findings presented here is that the 350-bp fragment lies between sites 7 and 8 and the small variation in size seen in nonexpressing tissue DNA and in fetal muscle DNA arises by methylation of site 8 (+448 bp) and a concomitant increase in size of 28 bp [see Fig. 5A]. The 250-bp fragment could lie between sites 6 and 7 and the 600-bp fragment between site 6 and the 3′ HindIII site. If this is the case then it is proposed that site 6 (−112 bp) is also unmethylated in DNA from all tissues irrespective of CAIII expression. Site 7 (+118 bp) is at least partially methylated in nonexpressing tissue but is apparently less methylated in adult and fetal skeletal muscle. If sites 2, 3, 4, and 5 were methylated, a fragment of 509 bp would be expected in the HpaII digests. Since this could not be clearly identified in any DNA tested, it may be that at least some of these other sites are unmethylated in all tissues.

Although we stress that other interpretations can be made of these preliminary data, it appears that the most 3′ of the CCGG sites in the CpG cluster [MspI 7 and MspI 8] are methylated in a variable manner associated with lack of gene activity. At least two sites in the cluster, MspI 1 and MspI 6, appear to be methylation free not only in DNA from tissues that express or have the potential to express CAIII, but in all tissues irrespective of whether CAIII is expressed or not. A similar observation has recently been reported by Kolsto et al. (1986) from studies of the tissue-specific gene Thy-1 in transgenic mice. Here, the CpG island of a microinjected mouse–human hybrid Thy-1 gene remained methylation free in all tissues, whether the hybrid gene construct was expressed or not. It has very recently been shown that the CpG cluster around the transcription start site of the human a-globin gene is nonmethylated in all tested tissues, including sperm, and furthermore sites upstream of the pseudo-a gene are also nonmethylated in sperm DNA (Bird et al. 1987). These findings prompt the proposal that while methylation-free CpG islands may be prerequisite for gene transcription they are, alone, not sufficient for gene transcription.

**Methylation of nonclustered CpG sites**

Using the CAIII cDNA as a probe, the methylation pattern of the widely spaced CCGG sites 9, 10, and 11 [Fig. 1] has been investigated. Comparison of the MspI and HpaII restriction fragment patterns [Fig. 7] with the CAIII genomic map provides data for each site. For example HindIII–HpaII digests of adult liver and kidney DNA do not exhibit a 3.2-kb fragment seen in the HindIII–MspI digests [Fig. 7B, kidney not shown]. This fragment is replaced in the HindIII–HpaII digests by one of 5.2 kb. This suggests that sites 10 and 11 are fully methylated at the inner C of the CCGG site in liver and kidney DNA. Similarly, the presence of the 4.2-kb fragment in the same HindIII–HpaII digests indicates that site 9 is at least partially unmethylated. The presence of an intact 9.4-kb HindIII fragment in the HindIII–HpaII digests and a 10.0-kb EcoRI fragment in the EcoRI–HpaII digests indicates that sites 9, 10, and 11 are fully methylated in white blood cell (wbc) DNA [the HindIII–MspI, HindIII–HpaII digests of wbc DNA are not shown in Fig. 7]. In contrast, patterns indicating no methylation at the inner C of sites 9, 10, and 11 were found in adult skeletal muscle. These observations are compatible with the view that there is an association between gene activity and reduced methylation and that at least two CpG dinucleotides in the 3′ half of the CAIII gene are implicated. The doublet of bands at the 3.4-kb position in the HindIII–MspI digests and the 5.5-kb band in the EcoRI–MspI digests may indicate that site 11 is partially methylated at the outer C of the CCGG, site, since MspI

![Figure 6](https://example.com/fig6.png)

**Figure 6.** Restriction fragment patterns seen after digestion of DNA from human adult muscle, liver, and kidney, placenta, and fetal muscle at various times during development (weeks). The DNA was digested with HindIII prior to digestion with MspI [lanes M] or HpaII [lanes H]. A 32P-labeled 3.4-kb HindIII fragment (see Fig. 5) was used as probe. (A) The 2.3-kb HindIII–MspI fragment seen after 2 days exposure of the autoradiograph. (B) Smaller fragments detected using the probe labeled to high specific activity [10⁹ cpm/μg] and after exposure of the autoradiograph for 10 days.
DNA from human adult muscle, liver, and white blood cells

Figure 7. Restriction fragment patterns seen after digestion of DNA from human adult muscle, liver, and white blood cells (wbc), placenta, and fetal muscle at various times during development (weeks). The DNA was digested with EcoRI (A) or HindIII (B), prior to digestion with MspI (lanes M) or HpaII (lanes H). 32P-labeled CAIII cDNA was used as probe.

is sensitive to methylation at this site. This methylation does not appear to correlate with expression since it occurs in DNA from all the tissues analyzed. Placental DNA is relatively undermethylated at the nonclustered CpG sites in the CAIII gene [Fig. 6A] and this is consistent with the low methylation estimated in bulk placental DNA [47% placenta, 70–80% other tissues] (for review, see Razin and Cedar 1984).

In line with studies on other tissue-specific genes (see, for example, Turcotte et al. 1986), it was expected that the increasing levels of CAIII during muscle development would be accompanied by changes in DNA methylation at the nonclustered CpG sites. However muscle DNAs at 10, 18, and 24 weeks gestation show methylation patterns virtually identical to each other and to the CAIII nonexpressing tissues [Fig. 7B]. There are at least two possible explanations for the apparent absence of demethylation in fetal muscle DNA. First, CAIII expression in fetal muscle may be confined to a minority population of fibers with relatively high CAIII activity, presumably the type I fiber precursors. Since there are only a few recognizable type I fibers (1.5–2.5%) at 20 weeks gestation (Colling-Saltin 1978), their contribution to the DNA sample might not be detected against a background of non-CAIII-expressing, methylated DNA. A second possibility is that poor DNA demethylation may be intrinsic to muscle and may reflect the low incidence of cell division in this tissue. The major wave of mitosis in the developing muscle is over before myoblast fusion and thereafter there are very few mitotic events. Since demethylation in eukaryotes is thought to be dependent on the prevention of methylation of newly replicated DNA, this low level of mitosis may lead to an unusually long delay between gene expression and demethylation. It is not clear how the donations of nuclei from undifferentiated satellite cells to the growing fiber (Schultz 1985) would affect this picture.

The apparent absence of methylation at sites 9, 10, and 11 in adult muscle DNA was also unexpected since it implies that DNA from the CAIII-expressing type I, slow, fibers (40–50% in adult muscle) and the non-CAIII-expressing type II, fast, fibers (50–60% in adult muscle; Dubowitz 1985) must be equally undermethylated. It is possible that this pattern of methylation is related to the fiber type plasticity that can be observed when muscle is exercised or artificially stimulated. In these conditions there appears to be some transformation of preexisting fibers from one type to another (for review, see Salmons and Henricksson 1981; Pette 1984) with accompanying changes in the expression of fiber type-specific genes (for example, see Pette 1984; Rubinstein et al. 1978, Williams et al. 1986). Thus, the undermethylation of CAIII DNA in type II fibers may reflect a potential for CAIII expression in this cell type.

Materials and methods

General methods

High-molecular-weight DNA was prepared by standard techniques (Maniatis et al. 1982) from 10 grams of flash-frozen tissue or from leukocytes. Poly(A)+ RNA was isolated from adult human limb muscle removed during amputation and from flash-frozen postmortem rodent muscle [Lloyd et al. 1986]. Southern blot hybridizations were performed in the presence of dextran sulfate using GeneScreen Plus membranes. DNA fragments used for hybridization probes were purified by agarose gel electrophoresis and electroelution. 32P Labeling of DNA fragments was performed using the oligonucleotide primer method [Feinberg and Vogelstein 1984].

DNA sequences were determined by the dideoxy chain-termination method of Sanger et al. [1977] and Biggin et al. [1983]. A 2.2-kb EcoRI fragment [pCA2.2 in Fig. 1C] containing exons 1 and 2 and a 3.4-kb HindIII fragment [pCA3.4 in Fig. 1C] were sequenced in full. These fragments were sheared by sonication (Deininger 1983), end-repaired using the Klenow fragment of DNA polymerase I, and fragments 400–900 bp long recovered by agarose gel electrophoresis. Fragments were ligated into the Smal site of M13mp8 or mp 18 RF DNA and transfected into E. coli JM101. Sequences across intron–exon junctions were de-
RNase protection assay

A 191-bp Ball–EcoRI fragment from pCA3.4 [Fig. 1C] was subcloned into the SalI site of pSP65 and the cloning orientation determined by sequence analysis. Five micrograms of double-stranded plasmid DNA was dissolved in 10 μl of 7 mM Tris [pH 7.5], 7 mM MgCl₂, and 50 mM NaCl. Two microliters (5 ng) of the pSP65 decanucleotide primer [Bethesda Research Labs] was added and the DNA-primer mix placed at 100°C for 5 min. After rapid cooling on ice, sequence reactions were carried out as described. Plasmids that transcribed to give an antisense version of the insert DNA were linearized at the HindIII site and a single-stranded RNA probe transcribed in the presence of [α-32P]UTP [Melton et al. 1984]. RNA (10 μg) in hybridization buffer was mixed with an excess, 2.5 × 10⁶ cpm, of probe. Hybridization and RNase protection were carried out as described [Melton et al. 1984]. RNase-resistant fragments were resolved on 7 M urea, 6% polyacrylamide gels.

Acknowledgments

The authors are grateful to Harvey Isenberg for the preparation of RNA samples, to Lee Pullen for his help with Southern analyses, and to the MRc Tissue Bank at the Royal Marsden Hospital for supplying fetal tissues.

References

Biggin, M.D., T.J. Gibson, and H.F. Hong. 1983. Buffer gradient gels and 35S label as an aid to rapid DNA sequence determination. Proc. Natl. Acad. Sci. 80: 3963–3965.

Bird, A.P. 1986. CpG-rich islands and the function of DNA methylation. Nature 321: 209–213.

Bird, A., M. Taggart, M. Frommer, O. Miller, and D. McLeod. 1985. A fraction of the mouse genome that is derived from islands of nonmethylated, CpG-rich DNA. Cell 40: 91–99.

Bird, A., M. Taggart, R. Nichols, and D. Higgs. 1987. Nonmethylated CpG-rich islands at the human α locus: Implications for evolution of the α globin pseudogene. EMBO J. 6: 999–1004.

Carter, N., S. Jeffery, A. Shiels, Y. Edwards, T. Tipler, and D.A. Hopkinson. 1978. Characterisation of human carbonic anhydrase III from skeletal muscle. Biochem. Genet. 17: 837–854.

Colling-Saltin, A.-S. 1978. Enzyme histochemistry on skeletal muscle of the human foetus. J. Neurol. Sci. 39: 169–185.

Davis, M., L. West, J. Barlow, P.H.W. Butterworth, J. Lloyd, and Y.H. Edwards. 1987. Regional localisation of the carbonic anhydrase genes. Somat. Cell and Mol. Genet. [in press].

Deininger, P. 1983. Random sub-cloning of sonicated DNA: Application to shot-gun DNA sequence analysis. Anal. Biochem. 129: 216–233.

Duboweritz, D., ed. 1985. Muscle biopsy: A practical approach. Balliere Tindall, East Sussex, UK.

Edwards, Y.H., J.M. Barlow, C.P. Konialis, S. Povey, and P.H. Butterworth. 1986a. Assignment of the gene determining human carbonic anhydrase, CAI, to chromosome 8. Ann. Hum. Genet. 50: 123–129.

Edwards, Y.H., J. Lloyd, M. Parkar, and S. Povey. 1986b. The gene for human muscle specific carbonic anhydrase III [CAIII] is assigned to chromosome 8. Ann. Hum. Genet. 50: 41–47.

Feinberg, A.P. and B. Vogelstein. 1984. A technique for radiolabelled DNA restriction endonuclease fragments to high specific activity. Anal. Biochem. 137: 266–267.

Jeffery, S., Y.H. Edwards, and N. Carter. 1980. Distribution of CAIII in fetal and adult human tissue. Biochem. Genet. 18: 843–849.

Jeffery, S., C.A. Wilson, A. Mode, J.-A. Gustafsson, and N. Carter. 1986. Effects of hypophysectomy and growth hormone infusion on rat hepatic carbonic anhydrases. J. Endocrinol. 110: 123–126.

Kadonga, J.T., K.A. Jones, and R. Tjian. 1986. Promoter-specific activation of polymerase II transcription by SP1. Trends Biochem. Sci. 11: 20.

Kolsto, A.-B., G. Kollias, V. Giguere, K.-I. Isobe, H. Prydz, and F. Grosveld. 1986. The maintenance of methylation-free islands in transgenic mice. Nucleic Acids Res. 14: 9667–9675.

Lawn, R.M., E.F. Fritsch, R.C. Parkar, G. Blake, and T. Maniatis. 1978. The isolation and characterisation of a linked and β globin gene from a cloned library of human DNA. Cell 15: 1157–1174.

LeFranc, M.-P., A. Forster, R. Baer, M.A. Stenson, and T.H. Rabbitts. 1986. Diversity and rearrangement of human T cell rearranging γ genes: Nine germ-line variable genes belonging to two subgroups. Cell 45: 237–246.

Lloyd, J., M. McMillan, D.A. Hopkinson, and Y.H. Edwards. 1986. Nucleotide sequence and derived amino acid sequence of a cDNA encoding human muscle carbonic anhydrase. Gene 41: 233–239.

Maniatis, T., E.F. Fritsch, and J. Sambrook. 1982. Molecular cloning: A laboratory manual. Cold Spring Harbor Laboratory, Cold Spring Harbor, New York.

Martini, G., D. Toniola, T. Vulliamy, L. Luzzatto, R. Dono, G. Viglietto, G. Paonessa, M. Durso, and M.G. Persico. 1986. Structural analysis of the X-linked gene encoding human glucose-6-phosphate dehydrogenase. EMBO J. 5: 1849–1855.

McLauchlan, J., D. Gaffney, J.L. Whitten, and J.B. Clements. 1985. The consensus sequence YGTGTTYY located downstream from the AAATAA signal is required for efficient formation of mRNA 3′ termini. Nucleic Acids Res. 13: 1347–1368.

Melton, D.A., P.A. King, M.R. Rebagliati, T. Maniatis, K. Zinn, and M.R. Green. 1984. Efficient in vitro synthesis of biologically active RNA and RNA hybridization probes from plasmids containing a bacteriophage SP6 promoter. Nucleic Acids Res. 12: 7035–7056.

d’Onofrio, C., V. Calantuoni, and R. Cortese. 1985. Structure and cell-specific expression of a cloned human retinol binding protein gene: The 5′-flanking region contains hepatoma specific transcriptional signals. EMBO J. 4: 1981–1989.

Pette, D. 1984. Activity-induced fast to slow transitions in mammalian muscle. Med. Sci. Sports Exercise 16: 517–528.

Razin, A. and H. Cedar. 1984. DNA methylation in eukaryotic cells. Int. Rev. Cytol. 92: 159–185.

Razin, A., H. Cedar, and A.D. Riggs, eds. 1984. DNA methylation: Biochemistry and biological significance. Springer-Verlag, New York.

Rubinstein, N., K. Mabuchi, F. Pepe, S. Salmons, J. Gergely, and F. Sreter. 1978. Use of type-specific antimonyosins to demonstrate the transformation of individual fibers in chronically stimulated rabbit fast muscles. J. Cell Biol. 79: 252–261.

Salmons, S. and J. Henricksson. 1981. The adaptive response of skeletal muscle to increased use. Muscle Nerve 4: 94–105.

Sanger, F., S. Nicklen, and A.R. Coulson. 1977. DNA sequencing with chain-terminating inhibitors. Proc. Natl. Acad. Sci. 74: 5463–5467.
Schultz, E. 1985. Satellite cells in normal, regenerating and
dystrophic muscle. In Gene expression in muscle [ed. R.C.
Strohman and S. Wolf], pp. 73–85. Plenum Press, New York.
Tashian, R.E. and D. Hewett-Emmett, eds. 1984. Biology and
chemistry of the carbonic anhydrases. Ann. N.Y. Acad. Sci.
249.
Turcotte, B., M. Guertin, M. Chevrette, H. LaRue, and L. Be-
langer. 1986. DNase I hypersensitivity and methylation of
the 5'-flanking region of the α-fetoprotein gene during devel-
opmental and glucocorticoid-induced repression of its ac-
tivity in rat liver. Nucleic Acids Res. 14: 9827–9842.
Venta, P.J., J.C. Montgomery, D. Hewett-Emmett, K. Weibauer,
and R.E. Tashian. 1985. Structure and exon to protein do-
main relationships of the mouse carbonic anhydrase II gene.
J. Biol. Chem. 260: 12130–12135.
Venta, P.J., T.B. Shows, P.J. Curtis, and R.E. Tashian. 1983.
Polymorphic gene for human carbonic anhydrase II: A mo-
lecular disease marker located on chromosome 8. Proc.
Natl. Acad. Sci. 80: 4437–4440.
Venta P.J., J.C. Montgomery, and R.E. Tashian. 1987. Molecular
 genetics of carbonic anhydrase isozymes. Curr. Topics Biol.
Med. Res. 14: 59–71.
Williams, R.S., S. Salmons, E. Newsholme, R. Kaufman, and J.
Mellor. 1986. Regulation of nuclear and mitochondrial gene
expression by contractile activity in skeletal muscle. J. Biol.
Chem. 261: 376–380.
Wistrand, P., N. Carter, and H. Asmark. 1987. Induction of rat
muscle carbonic anhydrase by denervation demonstrated
with immunofluorescence. Comp. Biochem. Physiol. 86A:
177–184.
Yisraeli, J. and M. Szyf. 1984. Gene methylation patterns and
expression. In DNA methylation: Biochemistry and biolog-
ical significance [ed. A. Razin, H. Cedars, and A.D. Riggs]
pp. 353–378. Springer Verlag, New York.
Yoshihara, C., J.-D. Lee, and J. Dogson. 1987. The chicken car-
bonic anhydrase II gene: Evidence for a recent shift in intron
position. Nucleic Acids Res. 15: 753–770.
Human muscle carbonic anhydrase: gene structure and DNA methylation patterns in fetal and adult tissues.

J Lloyd, C Brownson, S Tweedie, et al.

Genes Dev. 1987, 1:
Access the most recent version at doi:10.1101/gad.1.6.594

References
This article cites 33 articles, 7 of which can be accessed free at:
http://genesdev.cshlp.org/content/1/6/594.full.html#ref-list-1

License

Email Alerting Service
Receive free email alerts when new articles cite this article - sign up in the box at the top right corner of the article or click here.