The human carbonic anhydrase I gene has two promoters with different tissue specificities

Hugh J. M. BRADY,* Nicholas LOWE, Jane C. SOWDEN, Mina EDWARDS and Peter H. W. BUTTERWORTH†

Department of Biochemistry, University College London, Gower Street, London WC1E 6BT, U.K.

Primer extension and S1 nuclease analysis of human carbonic anhydrase I (*HCAI*) mRNA transcripts in erythroid and non-erythroid tissues show that the *HCAI* gene has two promoters with different tissue specificities, separated by 36 kb of DNA. The promoter of the *HCAI* gene which is active in non-erythroid tissue shows strong sequence similarity with a similar region in the mouse *CAI* gene, but has two equally used transcription start sites.

INTRODUCTION

Carbonic anhydrase I (CAI) is the next most abundant protein after the haemoglobins in adult red blood cells. It is one of at least seven distinct isoenzymes (CAI–VII) which show varied tissue specificity and different CO₂ hydration activities [1]. As well as erythroid tissue, CAI has also been detected in intestinal epithelium, vascular endothelium, corneal epithelium and the lens of the eye [2,3].

We have previously shown that the human *CAI* (*HCAI*) gene product appears in a developmental, stage-specific, manner in red blood cells and that the expression of the *HCAI* gene is regulated by stage-specific *trans*-acting factors [4]. We have also isolated several overlapping clones containing the entire human *CAI* gene [5]. These encompass a region of 50 kb, some 36 kb of which consists of a large intron in the 5'-untranslated region. The huge intron separates an erythroid-specific promoter (upstream of exon 1a) from the start of the protein-coding sequence contained in exon 1c (see [6] and Fig. 2). Work on mouse *CAI* (*MCAI*) suggests that this intron contains another *CAI* promoter which is responsible for the transcription of the gene in non-erythroid tissues [7]. Long-range physical mapping of the human CA genes *CA1*, *CA2* and *CA3* [8] suggests a different gene order to that found for the equivalent mouse genes by linkage analysis [9]. Here we report on a detailed analysis of the *HCAI* gene which demonstrates that tissue-specific differential promoter utilization is, in fact, a common feature of the expression of the *CAI* gene.

METHODS

Northern analysis

Total human colon RNA was prepared using guanidinium thiocyanate/caesium chloride [10]. Total human reticulocyte RNA isolation, HEL cell RNA isolation and Northern analysis were as described previously [6].

Transcription unit mapping

Primer extension analysis was carried out essentially as described previously [6]. The two primers used were oligo-5 (3'-CACCAGGACAGACCGTCGGA-5'), complementary to +33 and +52 in the 5'-leader of *HCAI* erythroid-specific mRNA,

and oligo-3 (3'-GGTCTGACCCCTATACTACTG-5'), complementary to the region from codon 3 to codon 9 in the *HCAI* protein-coding sequence.

S1-nuclease analysis was also carried out as described previously [6]. The end-labelled single-stranded probe was generated using oligo-3 as primer on M13mp18 *EcoRI*–*PstI* template followed by *RsaI* restriction enzyme digestion (see Fig. 3b). The *EcoRI*–*PstI* fragment is a 1.8 kb region of the *HCAI* gene containing within it exon 1c of the *HCAI* gene. A total of 1.4×10^5 c.p.m. of probe was hybridized to 10 µg of total human colon RNA at 58 °C for 3 h.

RESULTS

The first step in investigating the non-erythroid transcript of the *HCAI* gene was to find a non-erythroid tissue which expresses *HCAI*. The work on *MCAI* used RNA isolated from whole mouse large intestine [7]. Human gut tissue was obtained as a by-product of intestinal resectioning operations and RNA was prepared from proximal and distal colon epithelia. Northern analysis of total colon RNA showed the presence of an *HCAI* transcript (Fig. 1) comparable with that present in the human erythroleukaemia cell line HEL 92.1.7 [6].

Fig. 1. Northern analysis of RNA from HEL cells (H) and two separate samples of human colon (C)

The blot has been hybridized with ³²P-labelled human *CAI* cDNA probe.

Abbreviations used: CA, carbonic anhydrase; PBGD, porphobilinogen deaminase.

* Present address: Laboratory of Gene Structure and Expression, National Institute for Medical Research, Mill Hill, London NW7 1AA, U.K.

† To whom correspondence should be sent, at present address: School of Biological Sciences, University of Surrey, Guildford, Surrey GU2 5XH, U.K.

Fig. 2. Diagram of the human *CAI* gene

The scale is numbered in kb from the transcription start point determined in Brady *et al.* [6]. Exons are indicated by black bars (coding sequence) or open boxes (non-coding sequence) and are not drawn to scale.

The structure of the *HCAI* gene is shown in Fig. 2. The 5'-untranslated region of the most abundant *HCAI* transcript present in erythroid tissue contains exons 1a and 1c continuous with one another. A minor species also contains exon 1b [5]. In the gene, exons 1a and 1c are separated by an intron of 36.5 kb. We have already shown that the region 5' to exon 1a contains an erythroid-specific promoter [6]. In *MCAI*, the region 5' to exon 1c contains a promoter which is active in intestinal tissue [7].

The site of transcription initiation of the colon *HCAI* transcript was examined by primer extension and S1-nuclease analyses. For the primer extension analysis, two different oligonucleotide primers were used: oligo-5, which lies within exon 1a, and oligo-3, which lies within exon 1c (Fig. 3a). If the same situation

applies in the *HCAI* gene as in *MCAI*, then the 1c primer should result in an extended product in colon RNA but the 1a primer should not, since exon 1a would not be present in that transcript.

Fig. 3(a) shows the result of the primer extension analysis. Using the 1a primer (oligo-5), a correctly initiated 52 nt product is found in adult human reticulocyte RNA (lane 4), but no extension product is present using the colon RNA (lane 3). The 1c primer (oligo-3) produces the expected 122 nt extension using reticulocyte RNA (lane 2). However, using colon RNA, the 1c primer results in 110 nt and 93 nt extension products (lane 1). Therefore the *HCAI* gene has at least two different start sites, one used in erythroid tissues and possibly more than one in the colon.

S1-nuclease analysis was then carried out on colon RNA to confirm whether or not the *HCAI* colon mRNA start site was continuous with exon 1c as in the *MCAI* transcript. An end-labelled single-stranded DNA probe was generated from an M13mp18 *EcoRI*-*PstI* template using oligo-3 as primer (Fig. 3b). This *EcoRI*-*PstI* fragment contains exon 1c. When this probe was used in S1-nuclease analysis of total human colon RNA, two protected fragments of 93 nt and 110 nt were produced (Fig. 3b). Therefore the transcription start sites of the *HCAI* gene used in the colon are directly 5' to exon 1c in the region previously thought to be 'intron'. S1-nuclease mapping clearly

Fig. 3. Transcription start point analysis using colon (C) and reticulocyte (R) RNA

(a) Primer extension analysis of total human reticulocyte and colon RNA using primers specific for exon 1c (lanes 1 and 2) or exon 1a (lanes 3 and 4). No extension is seen from the primer to exon 1a using colon RNA. (b) S1-nuclease mapping of the transcription start site using colon RNA, confirming that the primer extension product seen in (a) results from transcripts starting upstream of the protein coding sequence (translation start point is indicated as ATG).

1. Tashian, R. E. (1989) *BioEssays* **10**, 186-192
2. Carter, M. J. & Parsons, D. S. (1971) *J. Physiol. (London)* **215**, 71-94
3. Spicer, S. S., Steward, P. J. & Tashian, R. E. (1979) *J. Histochem. Cytochem.* **27**, 820-831
4. Brady, H. J. M., Edwards, M., Linch, D. C., Knott, L., Barlow, J. H. & Butterworth, P. H. W. (1990) *Br. J. Haematol.* **76**, 135-142
5. Lowe, N., Brady, H. J. M., Barlow, J. H., Sowden, J. C., Edwards, M. & Butterworth, P. H. W. (1990) *Gene* **93**, 277-283
6. Brady, H. J. M., Sowden, J. C., Edwards, M., Lowe, N. & Butterworth, P. H. W. (1989) *FEBS Lett.* **257**, 451-456
7. Fraser, P., Cummings, P. & Curtis, P. J. (1989) *Mol. Cell. Biol.* **9**, 3308-3313
8. Lowe, N., Edwards, Y. H., Edwards, M. & Butterworth, P. H. W. (1991) *Genomics*, in the press
9. Beechey, C., Tweedie, S., Spurr, N., Ball, S., Peters, J. & Edwards, Y. H. (1990) *Genomics* **6**, 692-696
10. Maniatis, T., Fritsch, E. F. & Sambrook, J. (1982) *Molecular Cloning: A Laboratory Manual*, Cold Spring Harbor Laboratory, Cold Spring Harbor
11. Kadonaga, J. T., Courey, A. J., Ladika, J. & Tjian, R. (1988) *Science* **242**, 1566-1570
12. Jones, K. A., Yamamoto, K. R. & Tjian, R. (1985) *Cell* **42**, 559-572
13. Tsai, S. F., Martin, D. I. K., Zon, L. I., D'Andrea, A. D., Wong, G. G. & Orkin, S. H. (1989) *Nature (London)* **339**, 446-451
14. Martin, D. I. K., Zon, L. I., Mutter, G. & Orkin, S. H. (1990) *Nature (London)* **344**, 444-447
15. Wilson, D. B., Dorfman, D. M. & Orkin, S. H. (1990) *Mol. Cell. Biol.* **10**, 4854-4862
16. Chretien, S., Dubart, A., Beaupain, D., Raich, N., Grandchamp, B., Rosa, J., Goossens, M. & Romeo, P. H. (1988) *Proc. Natl. Acad. Sci. U.S.A.* **85**, 6-10
17. Le Van Kim, C., Colin, Y., Mitjavilia, M. T., Clerget, M., Dubart, A., Nakazawa, M., Vainchenker, W. & Cartron, J. P. (1989) *J. Biol. Chem.* **264**, 20407-20414

Vol. 277