Volume 257, number 2, 451–456 FEB 07813 November 1989

Multiple GF-1 binding sites flank the erythroid specific transcription unit of the human carbonic anhydrase I gene

Hugh J.M. Brady, Jane C. Sowden, Mina Edwards, Nicholas Lowe and Peter H.W. Butterworth

Department of Biochemistry, University College London, Gower Street, London WC1E 6BT, England

Received 14 September 1989

Six potential GF-1 sites which bind an erythroid factor are present in the 5′ and 3′ regions flanking the erythroid-specific transcription unit of the human carbonic anhydrase I (HCAI) gene. When two of these sites are placed upstream of a minimal eukaryotic promoter they confer up-regulated expression in erythroid over non-erythroid cells. The presence of the erythroid factor in TPA-treated HEL cells in which the level of HCAI transcript has greatly decreased and in non-HCAI-expressing K562 cells suggests that in these cases the presence of the factor is not sufficient for HCAI expression.

GF-1; Erythroid specific transcription factor; Human carbonic anhydrase I; *Trans*-acting protein

1. INTRODUCTION

Certain conserved DNA sequence elements to which transcription factors bind are necessary for the expression of most eukaryotic genes by RNA polymerase II; other elements, binding specific *trans*-acting protein factors have been shown to determine cell-specific expression [1]. In erythroid cells, the promoters of globin genes contain the conserved 'TATA' or 'CATA', 'CAAT' and 'CACCC' sequence cassettes and recent work has identified a sequence motif 'GATAAG' (or closely related variants thereof) which binds an erythroid-specific protein [2–5]. This sequence element is conserved across species and is found in either orientation in the regulatory regions of erythroid-specific genes. A cDNA encoding this factor has recently been cloned and designated GF-1 [6]. This paper defines the erythroid-specific transcription unit of the HCAI gene, the expression of which is characteristic of erythroid cells of the adult phenotype [7], and examines the binding of the erythroid-specific factor to sequences flanking it.

2. MATERIALS AND METHODS

2.1. *Transcription unit mapping*

Total human reticulocyte RNA was prepared by the guanidinium hydrochloride/caesium chloride method [8]. Primer extension analysis [9] of the 5′-end of HCAI mRNA used a single-stranded DNA oligonucleotide primer (3′-CACCAGGACAGACCGTCG-GA-5′) complementary to a sequence in the 5′-leader region of the HCAI gene (from +33 to +52). The primer was 5′-end labelled with T4 polynucleotide kinase and [γ-^{32}P]ATP and hybridised to 25 μg of RNA followed by extension with reverse transcriptase.

S_1-nuclease mapping used single-stranded DNA probes generated from M13 templates [10]. For 5′-end mapping, a *Pvu*II-*Hin*dIII fragment containing the first HCAI exon was subcloned into M13mp18 and the complementary strand synthesised using the above primer. For 3′-end mapping, a *Hin*dIII-*Mbo*I fragment from the 3′-untranslated region (221–572 bp downstream from the stop codon) was subcloned into M13mp19 and synthesis of the complementary strand initiated using the 17mer Amersham sequencing primer. 3×10^5 cpm of 5′-end probe was hybridised to 15 μg total human reticulocyte RNA at 62 or 70°C for 3 h. 1.4×10^5 cpm of 3′-end probe was hybridised to 9 μg of RNA for 1 h at 50 or 58°C. S_1-nuclease digestion was carried out for 2 h at 20°C.

Correspondence address: H.J.M. Brady, Department of Biochemistry, University College London, Gower Street, London WC1E 6BT, England

2.2. *Cell lines and tissue culture*

The erythroid cell lines used were K562 [11], K562-SA1 [12], HEL (92.1.7) [13] and mouse erythroleukemic (MEL) cells F412B2 (TK$^-$) [14]. K562 have an embryonic/foetal phenotype, the others an adult phenotype. All were grown in Dulbecco's MEM (DMEM) with 10% foetal calf serum (Gibco) plus penicillin (100 U/ml), streptomycin (100 μg/ml) and amphotericin B (2.5 μg/ml). Another MEL cell line C88 (APRT$^-$) [15] was also used and grown in α-MEM supplemented with 10% foetal calf serum and 50 μg/ml diaminopurine. HeLa cells were grown in DMEM as above. HL-60 (myeloid) [16] and HUT-78 (lymphoid) [17] were grown in RPMI 1640 plus 10% foetal calf serum with antibiotics and amphotericin B as above. All media were supplemented with 2 mM glutamine.

HEL cells were induced to undergo a macrophage-like shift by treatment with 12-*O*-tetradecanoyl-phorbol-13-acetate (TPA) [18]. TPA was dissolved in dimethylsulphoxide (DMSO) and used at 10^{-6} M for 4–8 days. Control cultures contained equivalent amounts of DMSO (0.01%).

2.3. *Protein preparation*

Whole cell extracts used for gel retardation were prepared by modification of the method of Dale et al. [19]: frozen cell pellets were made of total volume 0.2 ml containing $2–3 \times 10^7$ cells. 1.0 ml ice-cold extraction Buffer A (10 mM Hepes, pH 7.9, 0.4 M NaCl, 1.5 mM $MgCl_2$, 0.1 mM EGTA, 0.5 mM DTT, 0.5 mM PMSF and 5% glycerol) was added to a single pellet for lysis. The lysate was cen-

Published by Elsevier Science Publishers B.V. (Biomedical Division)
00145793/89/$3.50

trifuged at 100000 × g at 4°C for 15 min. The supernatant was desalted with Buffer B (same as Buffer A but with 50 mM NaCl) on a NAP-5 column (Pharmacia) and stored at −70°C.

Nuclear proteins for footprinting were prepared from 10^8 cells which were washed twice in phosphate-buffered saline and twice in Buffer I (0.05% Nonidet P-40, 10 mM Hepes pH 7.9, 10 mM NaCl, 3 mM $MgCl_2$). The lysate was resuspended in 10 ml Buffer II (10 mM Hepes, pH 7.9, 10 mM NaCl, 3 mM $MgCl_2$) and sedimented twice through 10 ml 30% sucrose in Buffer II at 1000 × g at 4°C for 5 min. Nuclei were resuspended in 3 ml Buffer III (20 mM Hepes pH 7.9, 0.42 M NaCl, 1.5 mM $MgCl_2$, 0.2 mM EDTA, 0.5 mM DTT, 0.5 mM PMSF, 25% glycerol) and the protein extract prepared as described by Wildeman [20] apart from the addition of 0.5 mM PMSF to the buffers used. Typically 5–10 mg · ml^{-1} protein was obtained from 10^8 cells and stored in aliquots at −70°C.

2.4. *Oligonucleotides and probe preparation*

The oligonucleotide 'αg2' derived from mouse α_1-globin was a gift from Dr M. Plumb (Beatson Institute, Glasgow). Other oligonucleotides (antisense strand shown below) were made as complementary single-stranded sequences and annealed before use:

Oligo A: 5′-GTATTTTTATTGATTATTGTGCTG-3′
Oligo B: 5′-ACCACTTCCCTTATCAGGTTCTC-3′
Oligo C: 5′-CCCACTCTAATCACCACAGGGCCA-3′
Oligo E: 5′-TGATCAAATGATTATCTTTTATAT-3′
Oligo F: 5′-CTATTTTTATCTTTAATTGACACA-3′
Oligo αg2: 5′-GATCCGGGCAACTGATAAGGATTCCC-AGATC-3′
Oligo CACCC: 5′-CTGATTAAATCCACACCCCA-3′

The oligonucleotides were 5′-end labelled as above and 5′-overhangs were filled in using excess dNTPs and Klenow fragment and purified by electrophoresis on a 10% polyacrylamide gel.

Fragment 'D', a 57 bp *Nhe*I-*Eco*RI fragment (195–251 bp downstream from the 'stop' codon) which lies between the two polyadenylation sites, was dephosphorylated using calf intestinal phosphatase and ^{32}P end-labelled as above.

2.5. *Gel retardation assay*

Gel retardation assays using whole cell extracts were carried out essentially as described by Dale et al. [19]. 10 μl of extract was preincubated with 1 μl of 5 mg · ml^{-1} poly(dI-dC) · poly(dI-dC) for 15 min at 20°C. Additional components were added to a final concentration of 0.5 × Buffer B, 2% Ficoll (w/v), 0.25 mg · ml^{-1} BSA, 10–20000 cpm end-labelled DNA and 100 ng competitor DNA where indicated in a final volume of 40 μl. The mixture was incubated for a further 15 min at 20°C. Samples were electrophoresed on a 5% nondenaturing polyacrylamide gel in 0.5 × TBE (89 mM Tris, 89 mM boric acid and 5 mM EDTA) at 150 V for 2 h.

2.6. *Footprinting analysis*

The 255 bp *Pvu*II-*Ava*II fragment (−219 to +35) was subcloned into the *Sma*I site of Bluescript plasmid (KS+, Stratagene). Both strands were labelled for DNase I footprinting of the promoter region of HCAI: the coding and noncoding strands were 5′-end labelled at the polylinker *Hin*dIII site and the *Dde*I site at +14, respectively, and fragments were purified after secondary digestion with *Hae*II (−107) for the coding strand and with *Pst*I (in the polylinker) for the noncoding strand. Markers were prepared by Maxam-Gilbert sequencing of the 5′-end labelled fragments.

Nuclear protein (50–100 μg) was preincubated with 1 μg poly(dI-dC) · poly(dI-dC) in 40 μl binding buffer (50 mM KCl, 5 mM $MgCl_2$, 1 mM EDTA, 10 mM Tris-HCl pH 8.0, 1 mM DTT, 12.5% glycerol and 0.1% Triton X-100) at 4°C for 30 min. Labelled fragments (20000 cpm) were added and incubated at 4°C for 30 min. DNase I digestion at 0.5 μg · ml^{-1}, in the presence of protein and 0.01 μg · ml^{-1} in the absence of protein, was at 20°C for 2 min, followed by the addition of 0.1 vol. of 'stop' solution (1 mM EDTA, 10% SDS, 1 mg · ml^{-1} tRNA). DNA was purified by organic extraction and resolved on an 8% denaturing polyacrylamide gel.

2.7. *Transfection*

F412B2 MEL cells and HeLa cells were transfected using calcium phosphate/DNA precipitation as described by Rosenthal [21]. 1.5 × 10^6 F412B2 cells were plated on 100 mm Corning tissue culture dishes (Bibby) 20–22 h before transfection whereas 3 × 10^6 HeLa cells were plated on 75 cm^2 Falcon tissue culture flasks (Becton and Dickinson) at the equivalent time. The precipitate was left in contact with F412B2 cells for 24 h and with HeLa cells for 6 h, followed by glycerol shock for 2 min. Both types of cell were left for 48 h after adding the precipitate before harvesting. Cell lysates were then made by 3 cycles of freeze-thawing.

2.8. *Chloramphenicolacetyl transferase (CAT) and β-galactosidase assays*

β-Galactosidase assays were performed exactly as described by Herbomel et al. [22]. Equivalent amounts of β-galactosidase activity for each transfected plate or flask were then assayed for CAT activity exactly as described by Gorman et al. [23].

2.9. *Northern analysis*

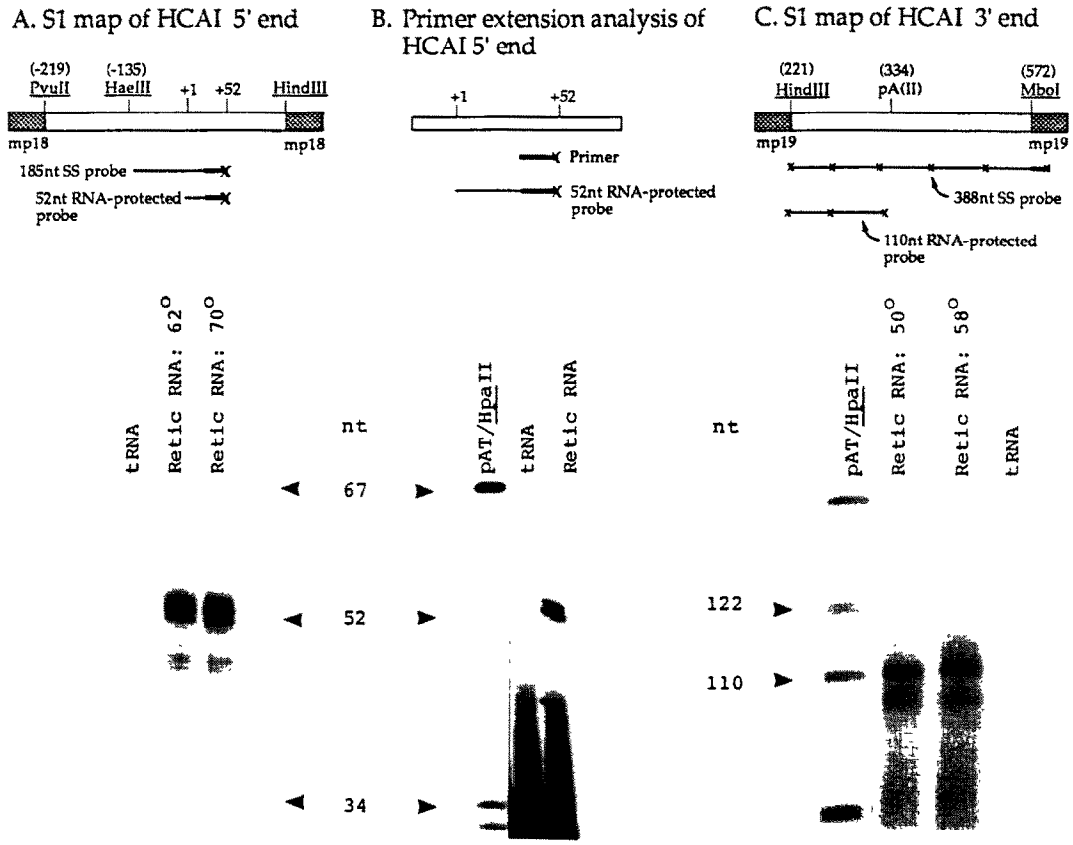
RNA was separated and transferred onto Gene Screen Plus [24]. HCAI mRNA was detected by hybridising the filter with 1 × 10^3 cpm HCAI cDNA [25] labelled with [γ-^{32}P]dCTP using random primers. Hybridisation and washing conditions were as recommended for Gene Screen Plus.

3. RESULTS AND DISCUSSION

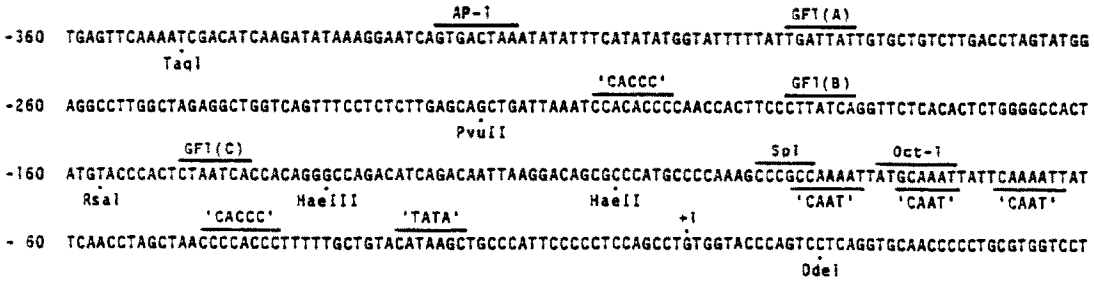
S_1-mapping and primer extension studies have defined the position of the 5′-end of the transcription unit (fig.1A,B). S_1-mapping has also identified the polyadenylation site, pA(II) at the 3′-end of the most abundant HCAI mRNA species (fig.1C) which lies 225 bp downstream from an alternative (yet rarely used) site of 3′-end maturation, pA(I), previously described from an analysis of cDNA clones [25]. Consensus sequences for the binding of general transcription factors are apparent (fig.1D). At −28 there is a globin-like 'CATA' motif [26] and three potential 'CAAT' box sequences [26] between −60 and −90. The flanking sequences also contain consensus binding sites for characterised transcription factors: for the 'CACCC'-binding factor [27] at −209 and −47; for AP-1 [28] at −324 and 801 bp downstream from the end of the protein-coding sequence; for Sp1 [29] at −93 and Oct-1 [30] at −81. Based on previously reported consensus sequences [3–5] for the binding of an erythroid-specific transcription factor, GF-1, six potential sites are found flanking the HCAI gene: sites A, B and C at −290, −190 and −149, respectively, and sites D, E and F located 223 bp, 581 bp and 833 bp downstream from the 'stop' codon. Site D lies between the two polyadenylation sites and site E has the sequence motif in two orientations.

Gel retardation assays show that all six GATAAG-like sequences flanking the HCAI transcription unit bind the same erythroid-specific protein (fig.2). Double-stranded oligonucleotides (23- or 24-mers) containing sites A, B, C, E and F and a 57 bp fragment containing site D were used. Each gives rise to a banding pattern containing a more abundant upper band

and a much less abundant lower band when incubated with a protein extract from erythroid (MEL) cells. In each case, competition using Oligo-B or Oligo-αg2 (in which the only common sequence is a GATAAG-like motif) shows binding to be specific to the GATAAG-like motif (lanes 15–26). Protein extracts from erythroid cell lines regardless of developmental phenotype (K562, K562-SA1, MEL and HEL, lanes 1,



D. (i) 5'-region flanking the transcription start site



(ii) 3' region flanking the translation stop signal

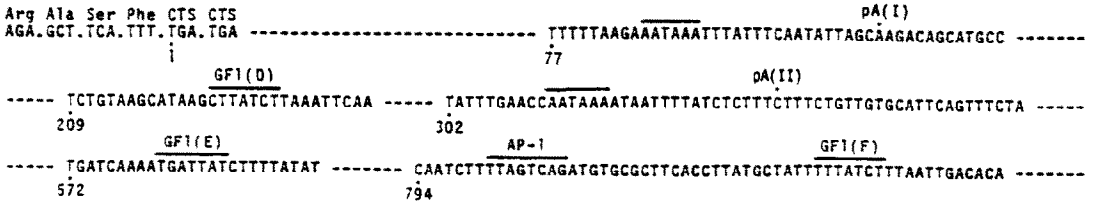


Fig.1. The HCAI transcription unit. (A) S_1-mapping and (B) primer extension analysis defining the transcription start site. (C) S_1-mapping of the most 3'-polyadenylation site. (D) DNA sequences flanking the 5'- and 3'-end of the HCAI transcription unit showing relevant restriction endonuclease cleavage sites and consensus sequences for binding of ubiquitous and cell-type specific transcription factors and for 3'-end maturation.

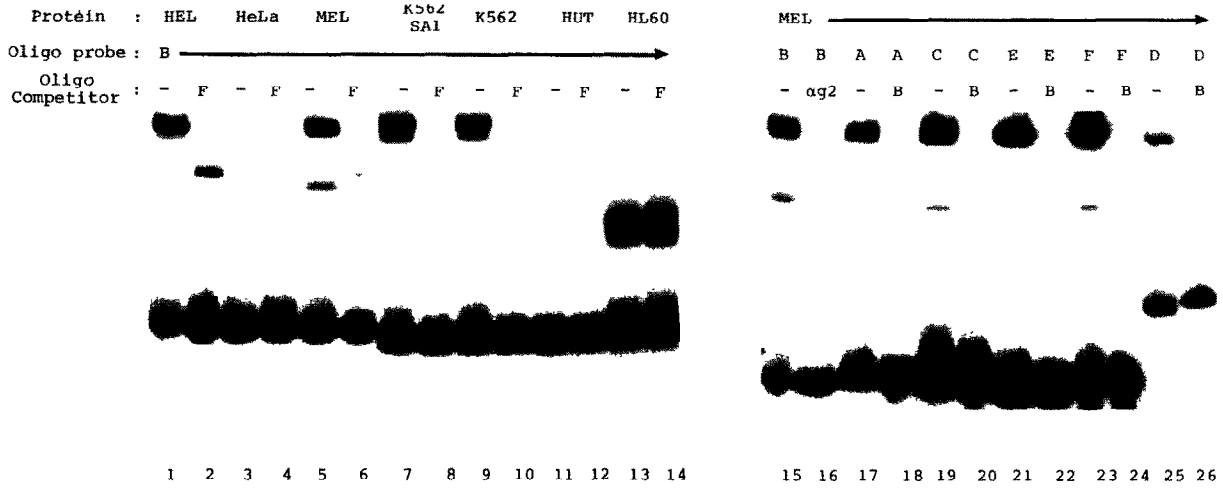


Fig.2. The binding of factors to GATAAG-like sequences flanking the HCAI gene. Gel retardation assays of 5′-end labelled double-stranded oligonucleotides with 50 μg protein extracts from erythroid and non-erythroid cell lines. Competition assays were performed with 150 ng unlabelled double-stranded oligonucleotides as specified. Those involving labelled oligo-B give rise to a characteristic band which does not occur with any of the other oligonucleotide probes used.

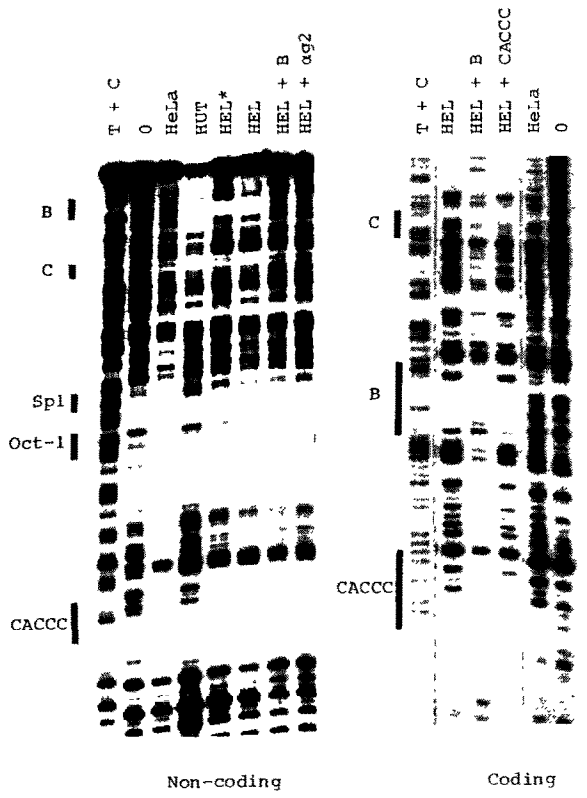


Fig.3. Footprint analysis of the −219 to +14 region of the HCAI gene. T+C indicates Maxam-Gilbert sequencing reactions; '0' indicates DNase I digestion without protein. The boxes to the left of the panels denote the footprint around each consensus: CACCC (−209 and −47), GF-1 binding sites B and C (−190 and −149, respectively), Sp1 (−93) and Oct-1 (−81). (A) Analysis of the non-coding strand from *Pvu*II (−219) to (+14) after binding with HEL* (50 μg) and HEL (100 μg), HeLa (100 μg) or HUT-78 (100 μg) nuclear extracts. Competition of the footprint over GF-I (site B) in HEL extracts was by the addition of 200 ng double-stranded oligonucleotide B or αg2. (B) Analysis of the coding strand from *Pvu*II (−219) to *Hae*II (−107) after binding with protein extracts from HEL (100 μg) and HeLa (100 μg); competitor for footprint at site B was 200 ng Oligo-B.

2, 5–10) all contain the factor (forming complexes with Oligo-B which are competed out by Oligo-F); thus the factor is present in erythroid cells even in embryonic cells in which HCAI is not expressed (see [6]). The factor is absent from non-erythroid haemopoietic cells (HL-60 and HUT-78, lanes 11–14). HeLa cells do not have the same factor; however, this non-erythroid cell line does contain a small amount of a protein which forms a lower molecular weight complex with the GATAAG motif in Oligo-B (competed by Oligo-F, lanes 3 and 4). Comparing the six binding sites with the other published consensus sequences [3–5] suggests a core recognition site of 3′-Py-$\frac{T}{A}$-A-T-C-$\frac{A}{T}$-5′.

DNase I footprinting of the HCAI promoter region by HEL, HeLa and HUT-78 nuclear proteins (fig.3) shows protection around the GATAAG motif at Site B, exclusively with proteins from erythroid cells. The region between −193 and −179 containing Site B is footprinted by HEL cell proteins with the induction of a hypersensitive site at −180. The footprint is specifically competed out by the addition of GATAAG motif-containing double-stranded oligonucleotides B and αg2 but not the 'CACCC' oligonucleotide. No footprint is evident on either DNA strand for Site C at −150 which suggests non-equivalence in function between the multiple GATAAG-like elements. Footprints over the Sp-1, Oct-1 and 'CACCC' consensus sequences are also observed which are not erythroid specific [26,28,29].

To show in vivo effects of erythroid specific factor binding, the 5′ *Taq*I-*Rsa*I fragment (−348 to −157) of

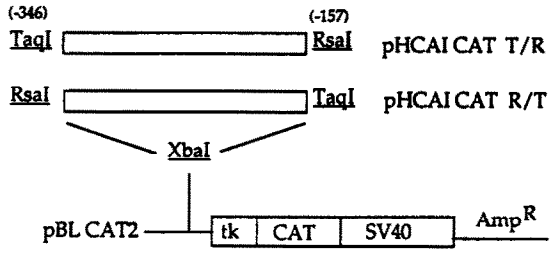


Fig.4. *Taq*I-*Rsa*I fragment of HCAI (−348 to −157) inserted (in either orientation) into the unique *Xba*I site of pBLCAT2, upstream of a minimal thymidine kinase (tk) promoter fused to the CAT reporter gene, the SV40 small t intron and polyadenylation site.

HCAI was placed in either orientation into an expression vector upstream of the minimal thymidine kinase promoter [31] fused to the CAT reporter gene illustrated in fig.4. Constructs containing the HCAI fragment, or vector alone, were cotransfected into cells with a plasmid containing the β-galactosidase reporter gene driven by the herpes simplex virus immediate early gene 4 promoter to normalise transfection efficiency. The transfected cells were MEL F412B2 cells which express mouse CAI and HeLa cells which do not express CAI. The plasmid containing the *Taq*I-*Rsa*I fragment of HCAI shows a 2.5–2.8-fold induction over the control plasmid in MEL cells but not in HeLa cells (table 1). This fragment which flanks the 5′-end of the HCAI gene contains two GATAAG-like motifs (Sites A and B), and consensus sequences for AP-1 and 'CACCC' binding proteins. However, 'CACCC'-box and AP-1 binding proteins are present in both HeLa [27,28] and MEL cells [4,32].

HEL cells constitutively express HCAI. When treated with the phorbol ester TPA, a shift takes place from erythroid to myeloid lineage as evidenced by the morphological, biochemical and functional changes they undergo [17]. Northern analysis (fig.5) shows that the treatment of HEL cells with TPA reduces the steady-state level of HCAI mRNA 7–8-fold compared with untreated HEL cells (from scanning densitometry). This is in contrast to the induction of CAII mRNA observed in TPA-treated HL60 cells [33] which

Table 1

Effect of an HCAI 5′-flanking region on minimal promoter function

	HeLa	F4
pBL CAT2	1.0	1.0
pHCAI CAT T/R	0.5	2.8
pHCAI CAT R/T	0.9	2.5

Each construct (illustrated in fig.4) was transfected separately into HeLa and MEL F412B2 (F4) cells. Normalised volumes of extracts from transfected cells (see section 2) were assayed for CAT activity and subsequently analysed by scanning densitometry. The data derived from each construct are given relative to pBL-CAT2 in each cell line

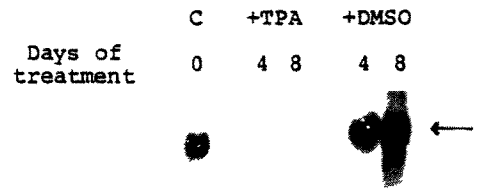


Fig.5. Northern analysis of equivalent amount of total RNA from control and TPA-treated HEL cells probed with ^{32}P-labelled HCAI cDNA.

indicates a difference in the regulation of CAI and CAII transcription. However, gel retardation assays with protein extracts from TPA-treated and control HEL cell cultures show no change in the binding pattern of the erythroid factor to Oligo-B (data not shown).

The presence of the erythroid factor in TPA-treated HEL cells in which the level of HCAI transcript has greatly decreased and in non-expressing K562 cells suggests that the presence of the erythroid factor (GF-1) is not sufficient for HCAI expression.

Acknowledgements: The authors are indebted to Drs Mark Plumb, Jon Frampton and Frank Grosveld for gifts of various cell lines, vectors and oligonucleotides and to Drs Yvonne Edwards, Ali Imam, Irving Johnston and David Linch for help and advice. This work was supported by a generous grant from the Wellcome Trust.

REFERENCES

[1] Maniatis, T., Goodbourn, S. and Fischer, J.A. (1987) Science 236, 1237–1245.
[2] Kemper, B., Jackson, P.V. and Felsenfeld, G. (1987) Mol. Cell. Biol. 7, 2059–2069.
[3] Evans, T., Reitman, M. and Felsenfeld, G. (1988) Proc. Natl. Acad. Sci. USA 85, 5976–5980.
[4] DeBoer, E., Antonio, M., Mignotte, V., Wall, L. and Grosveld, F. (1988) EMBO J. 7, 4203–4212.
[5] Plumb, M., Frampton, J., Wainwright, H., Walker, M., MacLoed, K., Goodwin, G. and Harrison, P. (1989) Nucleic Acids Res. 17, 73–92.
[6] Tsai, S.-F., Martin, D.I.K., Zon, L.I., D'Andrea, A.D., Wong, G.G. and Orkin, S.H. (1989) Nature 339, 446–451.
[7] Boyer, S.H., Siegel, S. and Noyes, A.N. (1983) Dev. Biol. 97, 250–253.
[8] Chirgwin, J.M., Przybyla, A.E., MacDonald, R.J. and Rutter, W.J. (1979) Biochemistry 18, 5294–5299.
[9] Sutton, R.E. and Boothroyd, J.C. (1986) Cell 47, 527–535.
[10] Burke, J.F. (1984) Gene 30, 63–68.
[11] Lozzio, C.B. and Lozzio, B.B. (1975) Blood 45, 321–334.
[12] Spandidos, P.A. (1984) Mol. Biol. Med. 2, 167–175.
[13] Martin, P.J. and Papayannopoulou, T. (1982) Science 216, 1233–1235.
[14] Spandidos, D. and Paul, J. (1982) EMBO J. 1, 15–20.
[15] Deisseroth, A. and Hendrick, D. (1978) Cell 15, 55–63.
[16] Collins, S.J., Gallo, R.C. and Gallagher, R.E. (1977) Nature 270, 347–349.
[17] Gootenberg, S.E., Ruscetti, F.W., Mier, J.W., Gazdar, A. and Gallo, R.C. (1981) J. Exp. Med. 154, 1403–1418.
[18] Papayannopoulou, T., Nakamoto, B., Yokochi, T., Chait, A. and Kannagi, R. (1983) Blood 62, 832–845.
[19] Dale, T., Ali Imam, A.M., Kerr, I.M. and Stark, G.R. (1989) Proc. Natl. Acad. Sci. USA 86, 1203–1207.

[20] Wildeman, A.G., Sassone-Corsi, P., Grundstrom, T., Zenke, M. and Chambon, P. (1984) EMBO J. 3, 3129–3133.
[21] Rosenthal, N. (1987) Methods Enzymol. 152, 704–720.
[22] Herbomel, P., Bourachot, B. and Yaniv, M. (1984) Cell 39, 653–662.
[23] Gorman, C.M., Moffat, L.F. and Howard, B.M. (1982) Mol. Cell. Biol. 2, 1044–1051.
[24] Fourney, R.M., Miyokoshi, J., Day, R.S. and Paterson, M.C. (1988) BRL Focus 10, 5–7.
[25] Barlow, J.H., Lowe, N., Edwards, Y.E. and Butterworth, P.H.W. (1987) Nucleic Acids Res. 15, 2386.
[26] Proudfoot, N.J., Shander, M.H.M., Manley, J.L., Gefter, M.L. and Maniatis, T. (1980) Science 209, 1329–1336.
[27] Schule, R., Muller, M., Oksuka-Murakani, H. and Renkawitz, R. (1988) Nature 332, 87–90.
[28] Lee, W., Mitchell, P. and Tjian, R. (1987) Cell 49, 741–752.
[29] Jones, K.H. and Tjian, R. (1985) Nature 317, 179–182.
[30] Sive, H.L., Heintz, N. and Roeder, R.G. (1986) Mol. Cell. Biol. 6, 3329–3340.
[31] Luckow, B. and Schutz, G. (1987) Nucleic Acids Res. 15, 5490.
[32] Hirai, S.I., Ryseck, R.P., Mechta, F., Bravo, R. and Yaniv, M. (1989) EMBO J. 8, 1433–1439.
[33] Shapiro, L.H., Venta, P.J., Ya-Shion, L.Yu. and Tashian, R.E. (1989) FEBS Lett. 249, 307–310.