Proc. Natl. Acad. Sci. USA
Vol. 82, pp. 663–667, February 1985
Biochemistry

Cloned cDNA for rabbit erythrocyte carbonic anhydrase I: A novel erythrocyte-specific probe to study development in erythroid tissues

(nonglobin proteins/developmental control/recombinant DNA/synthetic oligonucleotides)

CHRISTOPHOROS P. KONIALIS, JONATHAN H. BARLOW, AND PETER H. W. BUTTERWORTH

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

Communicated by Max L. Birnstiel, September 17, 1984

ABSTRACT **Present understanding of gene expression in erythropoietic tissues is derived solely from studies of the globin genes. Of the three distinct carbonic anhydrase (carbonate dehydratase; carbonate hydro-lyase, EC 4.2.1.1) isozymes, carbonic anhydrase I is erythrocyte-specific and, in humans, is under developmental control. The appearance of carbonic anhydrase I in the erythrocyte late in fetal life follows closely the γ- to β-globin switch. In order to study the expression of this erythrocyte-specific nonglobin protein, we set out to isolate a cloned carbonic anhydrase I cDNA. A mixture of 17-base-long synthetic oligonucleotides was used as an *in situ* hybridization probe to screen a rabbit reticulocyte cDNA library. Two clones were isolated, and the complete nucleotide sequence of the clone with the largest insert was determined and shown to code for carbonic anhydrase I. This clone, designated pRCAI, is near full length and has provided the 40% of the amino acid sequence of rabbit carbonic anhydrase I, which was not known hitherto. The deduced primary structure has revealed potentially significant changes in the vicinity of the active site of the rabbit carbonic anhydrase I when compared with carbonic anhydrase I and II sequences from other species.**

Studies on gene expression in erythropoietic tissues have focused almost exclusively on the globin gene families (reviewed in ref. 1). The occurrence of a number of genetic defects in humans within the α- and β-globin gene clusters and the changes in the pattern of expression of the different globin genes during human development have provoked intensive efforts to unravel the underlying mechanisms. This has been particularly true for the clinically important "switch" from fetal to adult globin. These changes are presumably indicative of events involving activation/inactivation of a number of other genes whose functions are required at different stages of erythrocyte development. Owing to the lack of suitable cloned probes, nothing is known concerning the expression of nonglobin genes whose products are characteristic of the mature erythrocyte.

In mammals, there are three major isozymes of carbonic anhydrase (carbonate dehydratase; carbonate hydro-lyase, EC 4.2.1.1) known as carbonic anhydrase I, II, and III. All three are monomeric zinc metalloenzymes and catalyze the rapid and reversible hydration of CO_2. However, they vary considerably in the pattern of their tissue distribution and their respective activities. Carbonic anhydrase I, the "low activity" form, shows tissue specificity insofar as it is expressed exclusively in erythroid tissues. On the other hand, carbonic anhydrase II, the "high activity" form, has the highest turnover number known for any enzyme and is found in a wide variety of tissues and cell types including erythrocytes. In human erythrocytes, carbonic anhydrase I is the predominant form, being present at levels 5–10 times higher than carbonic anhydrase II. Carbonic anhydrase III exhibits unique kinetic characteristics and, in humans, is mainly found in red skeletal muscle. All three isozymes share extensive amino acid sequence homology, and there is a vast amount of information on their enzymology, structure, and physiology (reviewed in ref. 2). They are also polymorphic and a number of electrophoretic variants of erythrocyte carbonic anhydrase I and II have been reported (3). Very little is known, however, about the respective genes or the regulation of their expression. Although the genes for carbonic anhydrase I and II are probably closely linked (4) on chromosome 8 (5), the differences in tissue distribution suggest that they are likely to be independently regulated.

It is now well established that the virtual absence of carbonic anhydrase is characteristic of fetal erythropoiesis. Erythrocyte carbonic anhydrase levels rise sharply towards term and early infancy to reach adult levels by 3 years of age (6–8). A recent study (9) has now shown that this change in carbonic anhydrase concentration is in fact due to changes in the level of the erythrocyte-specific carbonic anhydrase I isozyme and that the timing and magnitude of this developmental change is similar to the observed change in adult hemoglobin (HbA) levels. The possibility of a parallel or even coordinate regulation of expression of the carbonic anhydrase I gene(s) with the "switch" from fetal to adult globin is particularly intriguing and is also supported by two other lines of evidence: (*i*) dimethyl sulfoxide induction of Friend erythroleukemia cells results in the parallel induction of carbonic anhydrase and globin synthesis (10), and (*ii*) a reversion to fetal erythropoiesis is consistently observed in a large proportion of juvenile chronic myeloid leukemias in which a marked increase in fetal hemoglobin (HbF) expression is accompanied by extremely low levels of carbonic anhydrase I (11, 12).

We have cloned a carbonic anhydrase I cDNA which provides the necessary probe to initiate a study of the various parameters that regulate the developmental expression of the carbonic anhydrase I gene in the erythroid compartment.

MATERIALS AND METHODS

RNA Isolations. Rabbit reticulocytes were obtained from anemic New Zealand White rabbits, and total RNA was extracted by homogenization in LiCl/urea (13), followed by guanidine·hydrochloride/CsCl (14) purification. Total RNA from human reticulocytes, obtained from a patient with hemolytic anemia, was extracted in a similar fashion. Poly(A)$^+$ RNA was isolated after two rounds of oligo(dT)-cellulose chromatography (15).

cDNA Cloning Procedures. Single-stranded cDNA (9 μg) was synthesized from 20 μg of rabbit reticulocyte mRNA by using 70 units of avian myeloblastosis virus reverse transcriptase (a gift from J. W. Beard) in the presence of 0.5 un-

The publication costs of this article were defrayed in part by page charge payment. This article must therefore be hereby marked "*advertisement*" in accordance with 18 U.S.C. §1734 solely to indicate this fact.

Abbreviation: bp, base pairs.

it/μl of placental RNase inhibitor (Bethesda Research Laboratories). The second strand was synthesized by a combination of avian myeloblastosis virus reverse transcriptase followed by *Escherichia coli* DNA polymerase I (Klenow fragment). The double-stranded cDNA was treated with 50 units of S1 nuclease (Sigma)/μg of cDNA at 23°C for 1 hr and then loaded onto a 3-ml Sepharose CL-4B column (Pharmacia) in 10 mM Tris·HCl, pH 8.0/1 mM EDTA/0.3 M NaCl. Fractions containing fragments greater than 400 base pairs were pooled and precipitated with ethanol. Of this double-stranded cDNA, 4 μg was further fractionated on a 1.5% low-gelling-temperature agarose gel (Marine Colloids, Rockland, ME), and cDNA in the region of 1000–1600 base pairs (bp) was recovered from the melted gel slice by extraction with phenol and precipitation with ethanol. Size-selected cDNA (10 ng) and unfractionated cDNA (0.4 μg) were dC-tailed at the 3′ end by using terminal transferase (P-L Biochemicals) and subsequently were hybridized to *Pst* I-cut, dG-tailed plasmid vector pAT153 (16). *E. coli* strain DH1 (17) was transformed with the recombinant plasmids, and tetracycline-resistant colonies were selected and grown on nitrocellulose filters (Schleicher & Schuell BA85). The "total cDNA" library consisted of ≈500,000 recombinants, while 500 colonies were obtained from the size-selected cDNA. Both libraries were replicated on nitrocellulose filters, and master filters were stored at −70°C (18).

Oligonucleotide Synthesis. A mixture of 16 17-base-long oligonucleotides, corresponding to all possible codons for amino acids 93–98 of the known amino acid sequence of rabbit carbonic anhydrase I (19), was synthesized by using a solid-phase phosphotriester method (20):

93 98	
-Phe-His-Phe-His-Trp-Gly-	Protein
5′-UU$^{U}_{C}$-CA$^{U}_{C}$-UU$^{U}_{C}$-CA$^{U}_{C}$-UGG-GG-3′	mRNA
3′-AA$^{A}_{G}$-GT$^{A}_{G}$-AA$^{A}_{G}$-GT$^{A}_{G}$-ACC-CC-5′	Oligo.

The product was purified by electrophoresis through a 20% polyacrylamide/urea gel followed by DE-52 column chromatography and desalting through a Bio-Gel P-2 column.

Library Screening. Colonies, replicated and fixed on nitrocellulose filters, were probed with 5′-^{32}P-labeled oligonucleotide mix (specific activity, 4–5 × 10^6 cpm/pmol) as follows: filters were prehybridized at 65°C for 4 hr in 6× SET buffer (1× SET = 30 mM Tris·HCl, pH 8.0/0.15 M NaCl/1 mM EDTA) containing 5× Denhardt's solution (1× Denhardt's solution = 0.02% Ficoll/0.02% polyvinylpyrrollidone/0.02% bovine serum albumin), 0.05% sodium pyrophosphate, and 100 μg of single-stranded salmon sperm DNA per ml. They were then hybridized at 38°C for ≈20 hr to labeled oligonucleotide (1 × 10^7 cpm/ml) in a medium containing 6× SET buffer, 2× Denhardt's solution, 0.05% sodium pyrophosphate, and 100 μg of tRNA per ml. Filters were then washed extensively with several changes of 6× NaCl/Cit (1× NaCl/Cit = 0.15 M sodium chloride, 0.015 M sodium citrate) containing 0.1% $NaDodSO_4$ and 0.05% sodium pyrophosphate at 38°C over a period of 90 min. They were then washed twice at 48°C for 10 min in 6× NaCl/Cit containing 0.05% sodium pyrophosphate, followed by a high-stringency wash at 52°C for 2 min in the same solution. Finally, the filters were dried and exposed to x-ray film with an intensifying screen at −70°C.

Isolation of Plasmid DNA. Plasmid DNA was extracted from bacteria as described by Ish-Horowicz and Burke (21), and restriction enzyme digests of plasmid DNA were analyzed on agarose gels.

DNA Sequence Analysis. DNA sequence determination was performed by the method of Maxam and Gilbert (22).

RNA Blotting and Hybridization. RNA was fractionated by electrophoresis through 1.2% agarose/2.2 M formaldehyde gels (23) and transferred onto nitrocellulose filters (24). Filters were hybridized to nick-translated (25) plasmid pRCAI.

RESULTS AND DISCUSSION

Screening of the cDNA Libraries and Isolation of a Carbonic Anhydrase I cDNA Clone. In order to obtain a probe for our studies on the human carbonic anhydrase I gene, we decided to isolate a rabbit carbonic anhydrase I cDNA clone from a rabbit reticulocyte cDNA library and to use this cDNA as a heterologous probe for the isolation of the corresponding human gene(s). Our approach was based on: (*i*) the fact that the partial rabbit carbonic anhydrase I amino acid sequence shows extensive homology with the human carbonic anhydrase I isozyme (19), which suggests that the rabbit carbonic anhydrase I cDNA could be used to detect human carbonic anhydrase I DNA sequences, and (*ii*) the availability of quantities of rabbit as opposed to human reticulocytes, particularly since large amounts of RNA might be required.

In our initial attempts to isolate a cloned carbonic anhydrase I cDNA from a rabbit reticulocyte "total" cDNA library, a mixed 14-mer oligonucleotide probe was constructed to known rabbit carbonic anhydrase I amino acid sequence; it contained the 14 bases at the 5′ end of the 17-mer whose sequence is shown in *Materials and Methods*. After screening ≈10,000 colonies with the ^{32}P-labeled 14-mer probe, we found that the signal obtained was extremely variable. A number of putative "positive" clones, which exhibit-

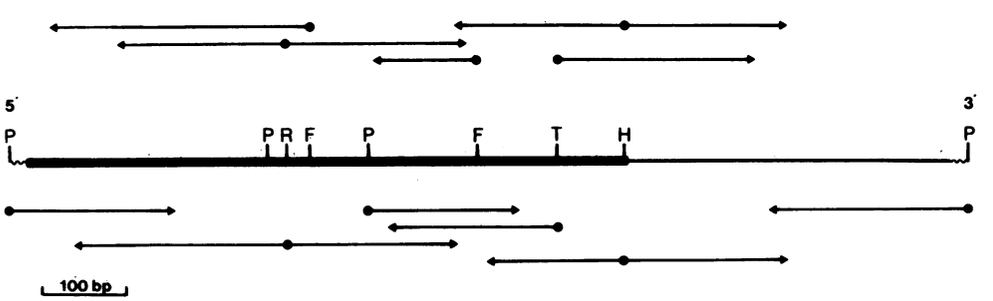

FIG. 1. Diagram of the cDNA insert of clone pRCAI and the strategy used to determine the DNA sequence. The heavy line identifies the protein coding region, the thin line denotes 3′ untranslated sequence, and the wavy line at the termini corresponds to G·C homopolymer tails. Only the restriction enzyme cleavage sites used to generate labeled fragments for sequencing are shown: P, *Pst* I; E, *Eco*RI; H, *Hin*dIII; F, *Hin*fI; and T, *Taq* I. The *Eco*RI and *Hin*dIII sites were labeled at both 5′ and 3′ ends. The *Taq* I site was 5′-labeled, while the *Pst* I and *Hin*fI sites were labeled at their 3′ ends. The poly(A) tail was identified from a partially sequenced fragment labeled at the *Pst* I site at the extreme 3′ end (data not shown). Arrows on the top indicate sequence readings of the sense (coding) strand while arrows at the bottom indicate sequence reading of the antisense (noncoding) strand.

ed relatively stable hybridization after stringent washing conditions, were subjected to limited DNA sequence analysis. None of them contained carbonic anhydrase I-coding sequences, neither did they contain rabbit globin cDNA inserts.

Although oligonucleotides of 14 bases long have been used for screening cDNA libraries, in this instance the signal obtained was not reproducible and resulted in the isolation of a number of "false positives." We also suspected that the abundance of carbonic anhydrase mRNA in reticulocytes may be considerably lower than we originally anticipated.

Therefore, we adopted a strategy that was designed to circumvent these problems. Double-stranded cDNA synthesized from rabbit reticulocyte poly(A)$^+$ RNA was fractionated by preparative agarose gel electrophoresis, and only sequences in the region of 1000–1600 bp were cloned, thereby excluding the vast majority (>90%) of globin transcripts. We also synthesized a 17-base-long mixed oligonucleotide (see *Materials and Methods*) and used this to screen the size-selected library. After hybridization and washing with 6× NaCl/Cit at 48°C, two positive colonies were clearly identified, both of which were positive on rescreening and washing at 52°C in 6× NaCl/Cit.

Plasmid DNA was isolated from the two positive clones and digested with *Pst* I. The inserts of both clones shared an internal 120-bp *Pst* I fragment containing an *Eco*RI site. The largest of the two clones contained an insert of 1120 bp from which an internal 120-bp *Pst* I–*Hin*fI fragment was isolated (Fig. 1) and subjected to DNA sequence analysis. Comparison of the deduced amino acid sequence with the partial amino acid sequence of rabbit carbonic anhydrase I (19) revealed that this clone contained carbonic anhydrase I-coding sequences and was designated pRCAI.

Nucleotide Sequence Analysis of pRCAI and Amino Acid Sequence of Rabbit Carbonic Anhydrase I. A detailed DNA sequence analysis of the pRCAI insert was undertaken using the strategy presented in Fig. 1, and the sequence is shown in Fig. 2. All of the sequence was determined on both strands, and each nucleotide was read on average 4.5 times. The pRCAI insert initiates 23 residues from the NH_2 terminus of rabbit carbonic anhydrase I, although the first 6 bp at the extreme 5′ end (corresponding to the known amino acids

```
                          30                                      40
Gly(Asn)Lys Gln Ser Pro Val Asp Ile Lys Ser Ser Glu Val Lys His Asp Thr Ser Leu
GGA ANT AAG CAG TCT CCA GTA GAT ATT AAA AGC AGC GAA GTG AAA CAT GAC ACC TCT CTG

          *               50                                    *   60
Lys Pro Phe Ser Val Ser Tyr Asn Pro Ala Ser Ala Lys Glu Ile Ile Asn Val Gly His
AAA CCT TTC AGT GTC TCC TAC AAT CCA GCC TCT GCC AAA GAA ATT ATC AAC GTG GGA CAT

 *                      70                                      80
Ser Phe His Val Asn Phe Glu Asp Asp Ser Gln Ser Val Leu Lys Gly Gly Pro Leu Ser
TCC TTC CAT GTC AAT TTT GAA GAT GAC AGC CAA TCA GTG CTG AAA GGC GGC CCT CTT TCT

    ***               90                                  *   100
Asp Asn Tyr Arg Leu Ser Gln Phe His Phe His Trp Gly Lys Thr Asp Asp Tyr Gly Ser
GAC AAC TAC CGA CTT TCT CAG TTC CAT TTT CAT TGG GGC AAG ACA GAT GAC TAT GGT TCT

                        110                                     120
Glu His Thr Val Asp Gly Ala Lys Phe Ser Ala Glu Leu His Leu Val His Trp Asn Ser
GAA CAC ACA GTG GAT GGA GCC AAA TTC TCT GCA GAG CTT CAC CTT GTT CAC TGG AAT TCT

                        130             ***                     140             ***
Gly Lys Tyr Pro Asn Ile Ala Asp Ser Val Ser Lys Ala Asp Gly Leu Ala Ile Val Ala
GGG AAG TAC CCC AAC ATC GCT GAC TCT GTT TCC AAG GCT GAT GGC TTG GCG ATT GTT GCT

    ***                 150  *       *                          160
Val Phe Leu Lys Val Gly Gln Ala Asn Pro Lys Leu Gln Lys Val Leu Asp Ala Leu Ser
GTT TTT CTG AAG GTT GGT CAG GCC AAC CCC AAG CTG CAG AAA GTT CTT GAT GCA CTG AGT

                        170       *       *                     180
Ala Val Lys Thr Lys Gly Lys Lys Ala Ser Phe Thr Asn Phe Asp Pro Ser Thr Leu Leu
GCA GTT AAA ACT AAG GGC AAA AAA GCC TCA TTC ACG AAC TTT GAC CCT TCC ACT CTG CTT

    ***                 190                                     200
Pro Pro Ser Leu Asp Tyr Trp Thr Tyr Ser Gly Ser Leu Thr His Pro Pro Leu His Glu
CCT CCG TCC TTG GAT TAC TGG ACC TAC TCT GGC TCT CTG ACT CAT CCC CCT CTT CAC GAG

                *** 210             ***                         220
Ser Val Thr Trp Leu Ile Cys Lys Asp Ser Ile Ser Ile Ser Ser Glu Gln Leu Ala Gln
AGT GTG ACC TGG CTC ATC TGT AAG GAC AGC ATC AGC ATC AGC TCA GAG CAG TTG GCA CAA

                        230                                     240
Phe Arg Ser Leu Leu Ser Asn Ala Glu Gly Glu Ala Ala Val Pro Ile Leu His Asn Asn
TTC CGC AGT CTT CTA TCG AAT GCT GAA GGT GAG GCT GCT GTC CCC ATA CTG CAC AAC AAC

          *             250
Arg Pro Pro Gln Pro Leu Lys Gly Arg Thr Val Lys Ala Ser Phe CTS
CGA CCA CCC CAG CCT CTG AAG GGC AGA ACT GTG AAA GCT TCA TTC TGATGGGTGCAAGAGGAGC

CTATCCTCTTAAGAAAGCAGCCTCGCCACCAACATGATCCAGTACAATAATCGCTTTAAGAAATAAGTTCATTTCCATT

TAATTTGTAGGACCAAAATCTTTAGTTTTAGTTCTTGATGTTTAACTAAAAATAGTATTCTGTAAGCTGCAAATAGCA
```

FIG. 2. Nucleotide sequence of clone pRCAI. The strategy shown in Fig. 1 was used to determine the nucleotide sequence by the method of Maxam and Gilbert (21). Most (>95%) of the sequence was determined from both strands, and each nucleotide was read on average 4.5 times. The deduced amino acid sequence of rabbit carbonic anhydrase I is also shown. Underlined portions identify regions where the amino acid sequence was previously unknown. *, Amino acid changes in previously known protein sequence at nonconserved (variant) positions (2); ***, amino acid changes at positions previously thought to be invariant for carbonic anhydrase I isozymes (2).

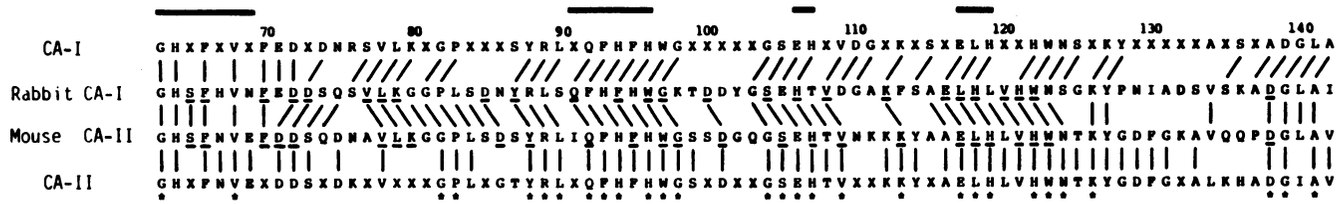

FIG. 3. Alignment of part of the protein-coding sequences of carbonic anhydrase I and II isozymes. The amino acid sequences are shown in the single-letter code (27), and the numbering of the residues shown for carbonic anhydrase I and II are taken from Tashian *et al.* (2) and only those amino acids that were thought to be invariant in all carbonic anhydrase I and II isozymes are presented; 'X' denotes a variable amino acid at that position. Rabbit carbonic anhydrase I and mouse carbonic anhydrase II sequences are the residues assigned from DNA sequence analysis of cDNA clones described in this paper and by Curtis *et al.* (26), respectively. Underlined amino acids in the rabbit carbonic anhydrase I and mouse carbonic anhydrase II sequences indicate absolute homology in the relevant codons; asterisks denote amino acids that are thought to be absolutely conserved between all carbonic anhydrase isozymes; the bold lines above the sequence identify regions that have been shown to be involved in the active site of the isozymes (see ref. 2 for a detailed review).

Ala-Asp) could not be identified unequivocally for technical reasons and are not presented. The insert contains all of the remaining coding region, giving a protein of 259 amino acids, the number of residues ascribed to carbonic anhydrase II but one less than other mammalian carbonic anhydrase I isozymes (see below). The cloned sequence also includes 360 bp of 3′ untranslated sequence extending to (but excluding) the poly(A) tail (data not shown).

Our deduced amino acid sequence provides the remaining 40% of the rabbit carbonic anhydrase I amino acid sequence, which was not known hitherto (2). The assignment is in general agreement where it overlaps with the 60% of the known sequence, although there are some differences identified in Fig. 2. However, an intriguing difference in amino acid sequence in the region of residues 70–80 is revealed when our assignment for the previously unknown sequence of rabbit carbonic anhydrase I is compared with the known sequences of other mammalian carbonic anhydrase I isozymes (2). There is a deletion at residue 73 (following the sequence Phe-Glu-Asp, see Fig. 3) in a region of the enzyme that separates two elements of the active site. Although errors during cDNA synthesis and/or cloning cannot be excluded, we believe that a deletion of exactly one codon is highly unlikely for a number of reasons. (*i*) The deduced amino acid sequence for rabbit carbonic anhydrase I (Asp-Ser-Gln) that follows the proposed deletion differs from the residues thought to be invariant in carbonic anhydrase I isozymes (Asp-Asn-Arg), but these amino acid changes are probably not random because the derived rabbit sequence shows homology with most carbonic anhydrase II isozymes in this region. In fact, the recently sequenced (26) mouse spleen carbonic anhydrase II cDNA exhibits homology with the rabbit carbonic anhydrase I cDNA at these residues, not only at the amino acid level but, more significantly, also at the nucleotide level (Fig. 3): the codons for the residues that straddle the proposed deletion are absolutely conserved. (*ii*) A detailed examination of the alignment of the known carbonic anhydrase I and II isozymes with our rabbit carbonic anhydrase I (Fig. 3) shows that the proposed deletion shifts the rabbit sequence out of phase with the arrangement of amino acids previously thought to be highly conserved in both sets of isozymes. It is particularly interesting that the rabbit carbonic anhydrase I gets back into register with the invariant residues in carbonic anhydrase II from position 127, the point at which there is an apparent deletion in all carbonic anhydrase II isozymes, which places the conserved amino acids in carbonic anhydrase II out of phase with carbonic anhydrase I isozymes. Whether these relative sequence displacements contribute in any way to the differences in catalytic properties of carbonic anhydrase I and II enzymes is not known. Previous attempts to maximize the amino acid sequence homology between the isozymes (2, 19) may have disguised real differences in sequence arrangement in the region of the active site. Nucleotide sequence homologies now have exposed subtle differences that could account for the distinctive catalytic properties of the carbonic anhydrase isozymes (illustrated in Fig. 3).

Blot Hybridization of Rabbit and Human Carbonic Anhydrase I mRNA. In order to determine the actual size of the rabbit carbonic anhydrase I mRNA and establish whether clone pRCAI could be used to detect human carbonic anhydrase I mRNA sequences, rabbit reticulocyte mRNA and human reticulocyte total RNA were fractionated on a denaturing agarose gel and transferred to a nitrocellulose filter. Plasmid pRCAI was nick-translated and hybridized to the filter-bound RNA. After stringent washing conditions, only one band of identical size could be clearly observed in both samples (Fig. 4). The size of the mRNA corresponds to ≈1250 nucleotides. Since pRCAI contains the entire 360 bp of the 3′ untranslated region and 707 of the 777 bases of coding sequence, then we can conclude that, at least for rabbit carbonic anhydrase I, the 5′ untranslated sequence is ≈50

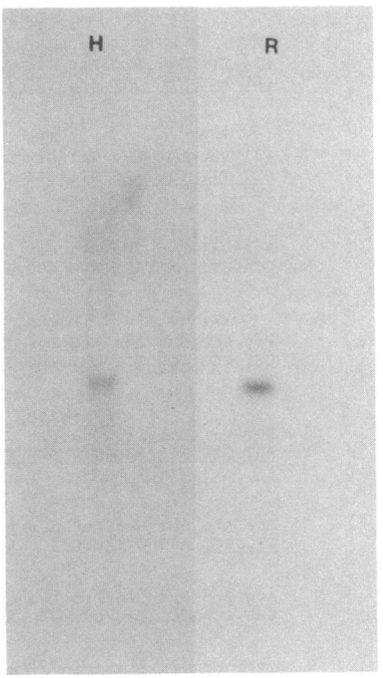

FIG. 4. Blot-hybridization analysis of rabbit and human carbonic anhydrase I mRNA. Rabbit reticulocyte poly(A)$^+$ RNA (4 μg; lane R) and human reticulocyte total RNA (20 μg; lane H) were analyzed on a 1.2% agarose/2.2 M formaldehyde gel and transferred to nitrocellulose (23). The hybridization probe was nick-translated plasmid pRCAI. The size was determined from known DNA restriction fragments analyzed in parallel tracks.

bases long [assuming an average poly(A) length of 80 bases]. These results demonstrate that the actual size of both rabbit and human carbonic anhydrase I mRNA is very similar and, more important, that pRCAI can indeed be used to detect and isolate human carbonic anhydrase I gene sequences. It is of interest that, although the 3′ untranslated region of rabbit carbonic anhydrase I mRNA is above average length, it is still considerably shorter than the 3′ untranslated region (670 bases) found in the recently sequenced (26) mouse spleen carbonic anhydrase II cDNA.

The availability of a cloned carbonic anhydrase I cDNA provides the probe needed to investigate the expression of this developmentally regulated, erythrocyte-specific protein. It is now possible to study in detail the expression of carbonic anhydrase I during human fetal development, with particular reference to the apparent coordinate regulation with the "switch" from fetal to adult hemoglobin. A considerable amount of data is accumulating regarding the controlled expression of the human β- and γ-globin genes (28). In this respect, information on the expression of the human carbonic anhydrase I gene, which appears to be coupled to the globin gene "switch," will provide further insight as to the (common ?) factors and events that may be involved. For this, it will be necessary to compare their expression *in vivo* with those of the γ- and β-globin genes following their introduction into erythroid and nonerythroid cells (28, 29).

We thank Dr. John Jenkins for many helpful suggestions during the course of this work and Hugh White for advice and help with the oligonucleotide synthesis. We also thank our colleagues Dr. Mike Rosemeyer, Dr. Steve Gregory, and Niall Dillon for many stimulating discussions and Steve Thomas for excellent technical assistance. We acknowledge the financial support of the Department of Biochemistry and the Wellcome Trust. P.H.W.B. is the recipient of a Medical Research Council fellowship under the RFAS scheme.

1. Maniatis, T., Fritsch, E. F., Lauer, J. & Lawn, R. M. (1980) *Annu. Rev. Genet.* **14,** 145–178.
2. Tashian, R. E., Hewett-Emmett, D. & Goodman, M. (1983) in *Isozymes: Current Topics in Biological and Medical Research*, eds. Rattazzi, M. C., Scandalios, J. G. & Whitt, G. S. (Liss, New York), Vol. 7, pp. 79–100.
3. Tashian, R. E., Kendall, A. G. & Carter, N. D. (1980) *Hemoglobin* **4,** 635–651.
4. Tashian, R. E. & Carter, N. D. (1976) in *Advances in Human Genetics*, eds. Harris, H. & Hirschhorn, K. (Plenum, New York), Vol. 7, pp. 1–56.
5. Venta, P. J., Shows, T. B., Curtis, P. J. & Tashian, R. E. (1983) *Proc. Natl. Acad. Sci. USA* **80,** 4437–4440.
6. Weatherall, D. J. & McIntyre, P. A. (1967) *Br. J. Haematol.* **13,** 106–114.
7. Sell, S. E. & Petering, H. G. (1974) *J. Lab. Clin. Med.* **80,** 369–377.
8. Eng, L.-I. L. & Tarail, R. (1966) *Nature (London)* **211,** 47–49.
9. Boyer, S. H., Siegel, S. & Noyes, A. N. (1983) *Dev. Biol.* **97,** 250–253.
10. Conscience, J.-F. & Meier, W. (1980) *Exp. Cell Res.* **125,** 111–119.
11. Sheridan, B. L., Weatherall, D. J., Clegg, J. B., Pritchard, J., Wood, W. G., Callender, S. T., Durrant, I. J., McWhirter, W. R., Ali, M., Partridge, J. W. & Thompson, E. N. (1976) *Br. J. Haematol.* **32,** 487–506.
12. Alter, B. P., Rappeport, J. M., Huisman, T. H. J., Schroeder, W. A. & Nathan, D. G. (1976) *Blood* **48,** 843–853.
13. Auffray, C. & Rougeon, F. (1980) *Eur. J. Biochem.* **107,** 303–314.
14. Chirgwin, J. M., Przybyla, A. E., MacDonald, R. J. & Rutter, W. J. (1979) *Biochemistry* **18,** 5294–5299.
15. Aviv, H. & Leder, P. (1972) *Proc. Natl. Acad. Sci. USA* **69,** 1408–1412.
16. Twigg, A. J. & Sherratt, D. (1980) *Nature (London)* **283,** 216–218.
17. Hanahan, D. (1983) *J. Mol. Biol.* **166,** 557–580.
18. Hanahan, D. & Meselson, M. (1980) *Gene* **10,** 63–67.
19. Tashian, R. E. (1977) in *Isozymes: Current Topics in Biological and Medical Research*, eds. Rattazzi, M. C., Scandalios, J. G. & Whitt, G. S. (Liss, New York), Vol. 2, pp. 21–62.
20. Gait, M. J., Matthes, H. W. D., Singh, M., Sproat, B. S. & Titmas, R. C. (1982) *Nucleic Acids Res.* **10,** 6243–6254.
21. Ish-Horowicz, D. & Burke, J. F. (1981) *Nucleic Acids Res.* **13,** 2989–2999.
22. Maxam, A. M. & Gilbert, W. (1980) *Methods Enzymol.* **65,** 499–560.
23. Seed, B. (1982) in *Genetic Engineering*, eds. Setlow, J. K. & Hollaender, A. (Plenum, New York), Vol. 4, pp. 91–102.
24. Thomas, P. S. (1980) *Proc. Natl. Acad. Sci. USA* **77,** 5201–5205.
25. Rigby, P. W. J., Dieckmann, M., Rhodes, C. & Berg, P. (1977) *J. Mol. Biol.* **113,** 237–251.
26. Curtis, P. J., Withers, E., Demuth, D., Watt, R., Venta, P. J. & Tashian, R. E. (1983) *Gene* **25,** 325–332.
27. IUPAC-IUB Commission on Biochemical Nomenclature (1968) *Eur. J. Biochem.* **5,** 151–153.
28. Chao, M. V., Mellon, P., Charnay, P., Maniatis, T. & Axel, R. (1983) *Cell* **32,** 483–493.
29. Joyner, A., Keller, G., Phillips, R. A. & Bernstein, A. (1983) *Nature (London)* **305,** 556–558.