Production of Active Human Carbonic Anhydrase II in *E. Coli*

Cecilia Forsman, Gity Behravan, Annika Osterman and Bengt-Harald Jonsson

Department of Biochemistry, University of Umeå, S-901 87 Umeå, Sweden

Forsman, C., Behravan, G., Osterman, A. and Jonsson, B.-H., 1988. Production of Active Human Carbonic Anhydrase II in *E. Coli*. – Acta Chem. Scand., Ser. B 42: 314–318.

cDNA encoding human carbonic anhydrase II has been isolated and its nucleotide sequence determined. Expression of the isolated carbonic anhydrase gene in *Escherichia coli* from a plasmid containing the tac promoter yielded an active enzyme at a level of about 1 % of total protein.

Carbonic anhydrase is a zinc-containing metalloenzyme catalyzing the reversible hydration of carbon dioxide.[1] Three genetically distinct, cytosolic isoenzymes have been isolated from mammalian sources.[1] They have molecular weights of about 30000 and consist of a single polypeptide chain of 259 or 260 amino acid residues. The three-dimensional structures of the two human isoenzymes I (HCAI)* and II (HCAII) are known to 2 Å resolution.[1] The tertiary structures of the isoenzymes are very similar but they have markedly different kinetic properties.[1] The coordination of the essential Zn(II) ion is nearly identical in the two isoenzymes, and several amino acids in the active site are invariant in all known carbonic anhydrase sequences. The strikingly different catalytic efficiencies of different isoenzymes might depend on specific properties of the non-homologous amino acids in the active site or on subtle differences in the position of the homologous amino acids. Our aim is to investigate the role of the amino acids that constitute the surface of the active site crevice using site-directed mutagenesis. Thence we have isolated cDNA encoding human carbonic anhydrase II and determined its nucleotide sequence. We have also constructed an expression system for production of active HCAII in *Escherichia coli.*

*Abbreviations: HCAI, human carbonic anhydrase I; HCAII, human carbonic anhydrase II; IPTG, isopropyl-β-D-thiogalactopyranoside; SDS, sodium dodecylsulfate.

Experimental

Restriction endonucleases were obtained from Boehringer Mannheim, Gmbh, West Germany. Nitrocellulose filters and nucleotides labelled with ^{32}P and ^{35}S were obtained from Amersham International plc, UK. Synthetic oligonucleotides were obtained from KabiGen and from Syn-Tek AB, Sweden.

E. coli Y 1090[2] was used for plating of the λgt11 fetal liver cDNA library[2] and *E. coli* strain JM 103[3] was used for propagation of plasmids and for propagation of M13 vector derivatives. The human fetal liver cDNA library prepared in the expression vector λgt11, where the cDNA fragments were inserted into the unique *Eco*RI site, was obtained from Clontech Laboratories, Inc., USA. For the plasmid expression experiments, *E. coli* strain SG 20043 (*lon*$^-$)[4] and the plasmid pICB1, originally constructed for the expression of bovine intestinal calcium binding protein from a synthetic gene,[5] were used.

On the basis of the amino acid sequence Gln–Phe–His–Phe–His–Trp–Gly, a conserved sequence in carbonic anhydrase I and II,[1] 32 synthetic 20-meric oligodeoxyribonucleotides [5′-CCCCA(A/G)TG(A/G)AA(A/G)TG(A/G)AA (T/C)TG-3′] complementary to mRNA for carbonic anhydrase were synthesized.

The recombinant phages of the λgt11 human fetal liver cDNA library were screened for HCAII cDNA sequences by plating 10^5 phages on the indicator strain Y 1090. Transfer of

plaques to and treatment of nitrocellulose filters were done basically as described by Benton and Davies.[6] Prehybridization was performed for 2 h at 42 °C, and hybridization was performed for 15 h at 42 °C in the prehybridization solution supplemented with 10^6 cpm per ml of the mixture of ^{32}P end-labelled 20-meric probes described above.[6] After incubation, the filters were washed, dried to dampness, wrapped in cling film and exposed to Du Pont Cronex 4 X-ray film for 48 h. Positive plaques were analysed using a coupled immunoassay system essentially as described by de Wet *et al.*,[7] using monospecific antiHCAII IgG as first antibody followed by peroxidase-conjugated goat antirabbit IgG (BioRad) as second antibody.

Restriction fragments isolated after digestion with *Eco*RI, *Bam*HI, *Hind*III, *Pst*I and *Sau*3AI were subcloned into M13mp18 and M13mp19 and sequenced by the Sanger dideoxy chain-termination method using 2′-deoxyadenosine 5′-[α-^{35}S]thiotriphosphate as tracer.[8,9] Additional sequences were determined after generation of overlapping clones by the deletion subcloning method of Dale *et al.*[10]

The following protocol was used for SDS polyacrylamide gel electrophoresis of protein extracts: *E. coli* strain SG 20043 was grown in Luria broth medium to a turbidity of 0.5 at 660 nm, when IPTG was added to a final concentration of 0.5 mM. After another 2 h of growth, the cells were harvested by centrifugation and boiled in a SDS-containing polyacrylamide gel electrophoresis cocktail.[11] The proteins were then separated on a 15 % discontinuous SDS/polyacrylamide slab gel.[11] Gels were either stained with 0.25 % (w/v) Coomassie brilliant blue R-250 or used in electroblotting using the semi-dry technique and Immobilon polyvinylidenedifluoride transfer membrane (Millipore). Immunodetection of blotted proteins was done essentially as described above.

The carbon dioxide hydration activity was measured at 2 °C by the colorimetric method of Rickli *et al.*[12] A lysate of *E. coli* strain SG 20043 containing the expression plasmid pHCAII, grown under the same conditions as described above, was assayed without further purification. Inhibition of the activity was tested using 10 μM diamox (2-acetylamido-1,3,4-thiadiazole-5-sulfonamide), which gives virtually complete inhibition of human carbonic anhydrase II.

Results and discussion

Isolation and sequencing of the cDNA clone. The initial screening of the cDNA library by DNA hybridization yielded about 100 positive plaques. Of these plaques, 49 were replated and hybrid β-galactosidase fusion proteins were induced by transfer to IPTG-soaked nitrocellulose filters that were subsequently analysed using antiHCAII IgG and a coupled immunoassay system. Phage particles from an isolated plaque showing both positive hybridization reaction and positive reaction to HCAII antibodies were purified, and their DNA was extracted.[13] The use of two different screening methods in concert was found to effectively exclude the false positive clones appearing when either of the methods was used exclusively. The length of the cDNA insert was determined by agarose gel electrophoresis after cleavage of the phage DNA with *Eco*RI, since at the construction of the library the cDNA fragments were inserted into a unique *Eco*RI site of the vector λgt11 using *Eco*RI linkers. The recombinant phage was designated λCA24. The *Eco*RI cDNA fragment of λCA24 was subcloned into *Eco*RI digested pUC19. The new plasmid was named pCA24I and restriction analysis was carried out on pCA24I DNA (Fig. 1).

Using the dideoxy method, and M13mp18 and M13mp19 subclones we sequenced both DNA strands of pCA24I insert DNA in full (Fig. 1). The sequence includes 36 nucleotides of 5′ untranslated sequence up to the initiation codon, 780 nucleotides of coding sequence and 539 nucleotides of 3′ untranslated sequence. The DNA sequence found at the translation initiation codon, CGACCATG, is homologous to the postulated consensus sequence for eucaryotic initiation sites, $CC^{A}_{G}CCAUG$.[14] Three possible polyadenylation sites with the sequence AATAAA identical to the postulated consensus sequence are underlined in Fig. 1. Translation of the coding region gives an amino acid sequence that is identical to the sequence deduced by amino acid sequencing of the enzyme by Henderson *et al.*[15] After the completion of our sequence work, the nucleotide sequence of HCAII cDNA from liver has been published by Montgomery *et al.*[16] A comparison yields two differences. Firstly, at position 562 in the coding region we obtain T rather than C as the first nucleotide in a leucine codon. Secondly, we find an extra G at position 1076 in our se-

```
                                                MetSerHisHisTrpGlyTyrGlyLysHisAsn
GAATTCGGGCCGATTCCTGCCCTGCCCCGACCGCCAGCGCGACCATGTCCCATCACTGGGGGTACGGCAAACACAAC
EcoRI Linker          -20                       1                   20
GlyProGluHisTrpHisLysAspPheProIleAlaLysGlyGluArgGlnSerProValAspIleAspThrHis
GGACCTGAGCACTGGCATAAGGACTTCCCCATTGCCAAGGGAGAGCGCCAGTCCCCTGTTGACATCGACACTCAT
     40                   60                    80                  100
ThrAlaLysTyrAspProSerLeuLysProLeuSerValSerTyrAspGlnAlaThrSerLeuArgIleLeuAsn
ACAGCCAAGTATGACCCTTCCCTGAAGCCCCTGTCTGTTTCCTATGATCAAGCAACTTCCCTGAGGATCCTCAAC
         120                 140                Sau3AI              BamHI
AsnGlyHisAlaPheAsnValGluPheAspAspSerGlnAspLysAlaValLeuLysGlyGlyProLeuAspGly
AATGGTCATGCTTTCAACGTGGAGTTTGATGACTCTCAGGACAAAGCAGTGCTCAAGGGAGGACCCCTGGATGGC
              200                 220                 240
ThrTyrArgLeuIleGlnPheHisPheHisTrpGlySerLeuAspGlyGlnGlySerGluHisThrValAspLys
ACTTACAGATTGATTCAGTTTCACTTTCACTGGGGTTCACTTGATGGACAAGGTTCAGAGCATACTGTGGATAAA
                   280                 300                 320
LysLysTyrAlaAlaGluLeuHisLeuValHisTrpAsnThrLysTyrGlyAspPheGlyLysAlaValGlnGln
AAGAAATATGCTGCAGAACTTCACTTGGTTCACTGGAACACCAAATATGGGGATTTTGGGAAAGCTGTGCAGCAA
    340   PstI       360                 380                 400
ProAspGlyLeuAlaValLeuGlyIlePheLeuLysValGlySerAlaLysProGlyLeuGlnLysValValAsp
CCTGATGGACTGGCCGTTCTAGGTATTTTTTTGAAGGTTGGCAGCGCTAAACCGGGCCTTCAGAAAGTTGTTGAT
         420                 440                 460                 480
ValLeuAspSerIleLysThrLysGlyLysSerAlaAspPheThrAsnPheAspProArgGlyLeuLeuProGlu
GTGCTGGATTCCATTAAAACAAAGGGCAAGAGTGCTGACTTCACTAACTTCGATCCTCGTGGCCTCCTTCCTGAA
              500                 520                Sau3AI
SerLeuAspTyrTrpThrTyrProGlySerLeuThrThrProProLeuLeuGluCysValThrTrpIleValLeu
TCCTTGGATTACTGGACCTACCCAGGCTCACTGACCACCCCTCCTCTTCTGGAATGTGTGACCTGGATTGTGCTC
                   580                 600                 620
LysGluProIleSerValSerSerGluGlnValLeuLysPheArgLysLeuAsnPheAsnGlyGluGlyGluPro
AAGGAACCCATCAGCGTCAGCAGCGAGCAGGTGTTGAAATTCCGTAAACTTAACTTCAATGGGGAGGGTGAACCC
    640                 660                 680                 700
GluGluLeuMetValAspAsnTrpArgProAlaGlnProLeuLysAsnArgGlnIleLysAlaSerPheLys
GAAGAACTGATGGTGGACAACTGGCGCCCAGCTCAGCCACTGAAGAACAGGCAAATCAAAGCTTCCTTCAAA
         720                 740                 760     HindIII    780
TAAGATGGTCCCATAGTCTGTATCCAAATAATGAATCTTCGGGTGTTTCCCTTTAGCTAAGCACAGATCTAC  852
CTTGGTGATTTGGACCCTGGTTGCTTTGTGTCTAGTTTTCTAGACCCTTCATCTCTTACTTGATAGACTTAC  924
TAATAAAATGTGAAGACTAGACCAATTGTCATGCTTGACACAACTGCTGTGGCTGGTTGGTGCTTTGTTTAT  996
GGTAGTAGTTTTTCTGTAACACAGAATATAGGATAAGAAATAAGAATAAAGTACCTTGACTTTGTTCACAGC 1068
ATGTAGGGTGATGAGCACTCACAATTGTTGACTAAAATGCTGCTTTTAAAACATAGGAAAGTAGAATGGTTG 1140
AGTGCAAATCCATAGCACAAGATAAATTGAGCTAGTTAAGGCAAATCAGGTAAAATAGTCATGATTCTATGT 1212
AATGTAAACCAGAAAAAATAAATGTTCATGATTTCAAGATGTTATATTAAAGAAAAACTTTAAAAATTATTA 1284
TATATTTATAGCAAAGTTATCTTAAATATGAATTC
                             EcoRI
```

Fig. 1. Nucleotide sequence and deduced amino acid sequence of human carbonic anhydrase II cDNA. Possible polyadenylation sites are underlined and the restriction enzyme sites used in subcloning are indicated. Differences from the sequence published by Montgomery *et al.*[16] are indicated by dots at positions 562 and 1076.

quence. Our sequence ends at an internal *Eco*RI site at position 1319.

Construction of an expression vector for HCAII. The expression plasmid pICB1[5] was a kind gift from Dr. T. Grundström and was used for construction of an expression vector for the HCAII gene. This multicopy plasmid is a well-suited expression system because it carries the strong tac promoter, constructed from the −35 part of the *trp* promoter and the −10 part of the *lac* promoter.[17] It also contains a *lac* repressor overproducing gene, *lacIQ*, to keep the tac promoter in a repressed state until the inducer IPTG is added to the growth medium.

Before cloning into pICB1, the HCAII cDNA was modified in a three-step process as outlined in Fig. 2. Firstly, the long untranslated 3′ end was eliminated by digestion with *Hind*III, which cleaves the cDNA 13 bases before the stop codon. Secondly, two synthetic oligonucleotides containing the last 13 bases of the cDNA fol-

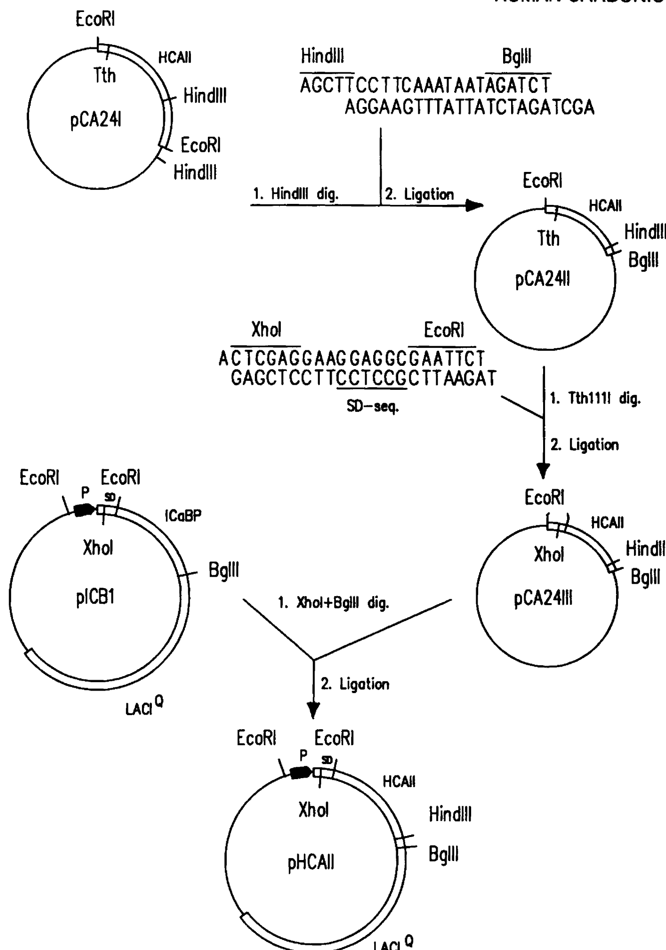

Fig. 2. Construction of the HCAII expression vector pHCAII. The untranslated 3′ end of the HCAII cDNA was eliminated by *Hind*III digestion and a *Bgl*II restriction site was introduced together with two stop codons by insertion of two oligonucleotides. A ribosome binding site, designated SD in the figure, along with a *Xho*I site was introduced by insertion of two oligonucleotides into the *Tth*111I site of pCA24II. The HCAII gene was subsequently inserted into pICB1 giving the plasmid pHCAII. The tac promoter is indicated with P and *lacIQ* is the *lac* repressor gene. *Tth*111I is abbreviated *Tth.*

lowed by two stop codons were ligated into the *Hind*III site of the plasmid pCA24I. Thirdly, two other synthetic oligonucleotides were inserted into a *Tth*111I site immediately in front of the initiation codon, thus creating a ribosome binding site with high homology to the 3′ end of 16 S ribosomal RNA.[18] The modified HCAII cDNA was subsequently inserted in pICB1 by cleavage of pCA24III with *Xho*I and *Bgl*II, and ligation into pICB1 digested with the same restriction enzymes. The plasmid thus obtained was named pHCAII.

Expression of HCAII in E. coli. Total protein extracts from the *lon* protease deficient *E. coli* strain SG 20043 carrying the plasmid pHCAII, or pICB1 as a control, were separated by SDS-polyacrylamide gel electrophoresis. The gel was divided in two parts. One part was stained with Coomassie Brilliant Blue and the proteins of the other part were electroblotted onto Immobilon-filter for immunodetection with monospecific antiHCAII antibodies (Fig. 3). pHCAII-containing SG 20043 induced with IPTG (lanes 2, 7 in Fig. 3) gives expression of a protein with the same mo-

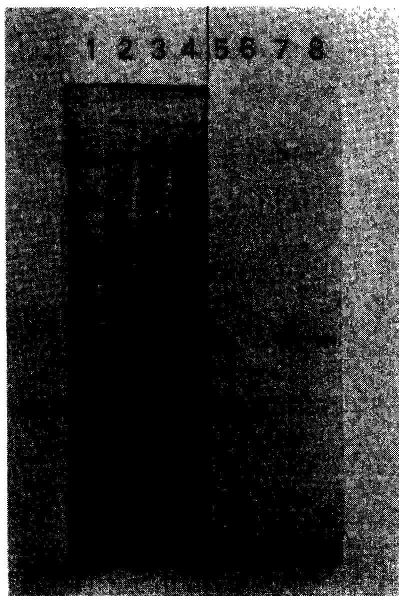

Fig. 3. SDS-polyacrylamide gel electrophoresis of protein extracts from *E. coli.* Lanes 1–4 are stained with Coomassie Brilliant Blue and lanes 5–8 are electroblotted and immunodetected. Lane 1: Pharmacia low-molecular weight standard. From top to bottom: M_r 94 000, 67 000, 43 000, 30 000, 20 100 and 14 400; Lanes 2,7: SG 20043 with pHCAII induced with IPTG; Lanes 3,6: SG 20043 with pHCAII without IPTG; Lanes 4,5: SG 20043 with pICB1 induced with IPTG; Lane 8: HCAII purified from red blood cells.

bility as HCAII purified from red blood cells (lane 8). For uninduced SG 20043 with pHCAII plasmid (lanes 3, 6) and for induced SG 20043 with pICB1 plasmid (lanes 4, 5), no reaction with HCAII antibodies could be detected.

Measurements of the CO_2 hydration activity of total protein extracts after repeated freezing-thawing and lysozyme treatment show an activity corresponding to a yield of active carbonic anhydrase at a level of about 1 % of total protein. Spectrophotometric scanning at 595 nm of the Coomassie-stained gel also indicates a yield of about 1 % of total protein. The activity of protein extracts from bacteria carrying pICB1 did not differ significantly from the uncatalyzed reaction, i.e. without protein extract added to the reaction medium. The activity measured from bacteria with pHCAII was completely inhibited by 10 μM of the sulfonamide diamox ($K_i = 6 \cdot 10^{-8}$ M for human CA II).

Acknowledgements. This investigation has been financially supported by a grant to Prof. Sven Lindskog from the Swedish Natural Science Research Council (K 2911). We are grateful to Ms. Katarina Wallgren for excellent technical assistance and to Prof. Sven Lindskog for stimulating discussions.

References

1. Lindskog, S. In: Spiro, T. G., Ed., *Zinc Enzymes,* Wiley, New York 1983, pp. 78–121.
2. Huynh, T. V., Young, R. A. and Davis, R. W. In: Glover, D., Ed., *DNA Cloning Techniques: A Practical Approach,* IRL Press, Oxford 1984, pp. 49–78.
3. Messing, J., Crea, R. and Seeburg, P. H. *Nucl. Acids Res. 9* (1981) 309.
4. Triesler, P. and Gottesman, S. *J. Bacteriol. 160* (1984) 184.
5. Brodin, P., Grundström, T., Hofmann, T., Drakenberg, T., Thulin, E. and Forsén, S. *Biochemistry 25* (1986) 5371.
6. Benton, W. D. and Davies, R. W. *Science 196* (1977) 180.
7. de Wet, J. R., Fukushima, H., Dewji, N. N., Wilcox, E., O'Brien, J. S. and Helinski, D. R. *DNA 3* (1984) 437.
8. Messing, J. and Vieira, J. *Gene 19* (1982) 269.
9. Sanger, F., Nicklen, S. and Coulson, A. R. *Proc. Natl. Acad. Sci. U.S.A. 74* (1977) 5463.
10. Dale, R. M. K., McClure, B. A. and Houchings, J. P. *Plasmid 13* (1985) 31.
11. Jergil, B. and Ohlsson, R. *Eur. J. Biochem. 46* (1974) 13.
12. Rickli, E. E., Ghazanfar, S. A. S., Gibbons, B. M. and Edsall, J. T. *J. Biol. Chem. 239* (1964) 1065.
13. Davis, R. W., Botstein, D. and Roth, J. R. *Advanced Bacterial Genetics,* Cold Spring Harbor Laboratory, Cold Spring Harbor, New York 1980.
14. Kozak, M. *Nucl. Acids Res. 12* (1984) 857.
15. Henderson, L. E., Hendriksson, D. and Nyman, P. O. *J. Biol. Chem. 251* (1976) 5457.
16. Montgomery, J. C., Venta, P. J., Tashian, R. E. and Hewett-Emmett, D. *Nucl. Acids Res. 15* (1987) 4687.
17. de Boer, H. A., Comstock, L. J. and Vasser, M. *Proc. Natl. Acad. Sci. U.S.A. 80* (1983) 21.
18. Shine, J. and Dalgarno, L. *Proc. Natl. Acad. Sci. U.S.A. 71* (1974) 1342.

Received December 14, 1987.