Kagan, H. M., Williams, M. A., Williamson, P. R., & Anderson, J. M. (1984) *J. Biol. Chem. 259*, 11203–11207.
Keech, M. I. (1961) *J. Biochem. Biophys. Cytol. 9*, 193–209.
Keller, K. M., Keller, J. M., & Kuhn, K. (1986) *Biochim. Biophys. Acta 882*, 1–5.
Koda, J. E., & Bernfeld, M. (1984) *J. Biol. Chem. 259*, 11763–11770.
Laemmli, U. K. (1970) *Nature (London) 227*, 680–685.
Lilja, S., & Barrach, H.-J. (1981) *Virchows Arch. A: Pathol. Anat. Histol. 380*, 325–338.
Lilja, S., & Barrach, H.-J. (1983) *Exp. Pathol. 23*, 173–181.
Mathews, M. B., & Decker, L. (1968) *Biochem. J. 109*, 517–526.
Oegema, T. R., Jr., Laidlaw, J., Hascall, V. C., & Dziewiatkowski, D. D. (1975) *Arch. Biochem. Biophys. 170*, 698–709.
Ohwiler, D. A., Jurkiewicz, M. J., Butcher, H. R., Jr., & Brown, J. B. (1959) *Surg. Forum 10*, 301–303.
Pasquali-Ronchetti, I., Bressan, G. M., Fornieri, C., Baccarini-Contri, M., Castellani, I., & Volpin, D. (1984) *Exp. Mol. Pathol. 40*, 235–245.
Pilcher, D. B., & Barker, W. F. (1970) *Am. J. Surg. 120*, 270–274.
Poole, A. R. (1986) *Biochem. J. 236*, 1–14.
Rosenberg, R. D. (1977) *Fed. Proc., Fed. Am. Soc. Exp. Biol. 36*, 10–18.
Siegel, R. C. (1974) *Proc. Natl. Acad. Sci. U.S.A. 71*, 4826–4830.
Siegel, R. C. (1979) *Int. Rev. Connect. Tissue Res. 8*, 73–118.
Smith, G. N., Jr., Williams, J. M., & Brandt, K. D. (1985) *J. Biol. Chem. 260*, 10761–10767.
Snowden, J. McK., & Swann, D. A. (1980) *Biopolymers 19*, 767–780.
Stinchfield, F. E., Sankaran, B., & Samilson, R. J. (1956) *J. Bone Jt. Surg., Am. Vol. 38-A*, 270–282.
Sullivan, K. A., & Kagan, H. M. (1982) *Methods Enzymol. 82*, 637–649.
Tollefsen, D. M. (1984) *Nouv. Rev. Fr. Hematol. 26*, 233–626.
Toole, B. P. (1969) *Nature (London) 22*, 872–873.
Toole, B. P. (1981) in *Cell Biology of Extracellular Matrix* (Hay, E. D., Ed.) pp 259–294, Plenum, New York.
Toole, B. P., & Lowther, D. A. (1968) *Biochem. J. 109*, 857–866.
Vogel, K. G., Paulsson, M., & Heinegard, D. (1984) *Biochem. J. 223*, 587–597.
Williams, M. A., & Kagan, H. M. (1985) *Anal. Biochem. 149*, 430–437.
Wood, G. C. (1960) *Biochem. J. 75*, 605–612.
Zimmerman, T. S., & Plow, E. F. (1983) in *Internal Medicine* (Stein, J. H., Ed.) pp 1607–1617, Little, Brown and Co., Boston, MA.

Complete Amino Acid Sequence of Ovine Salivary Carbonic Anhydrase†

Ross T. Fernley,* R. Douglas Wright, and John P. Coghlan
Howard Florey Institute of Experimental Physiology and Medicine, University of Melbourne, Parkville, Victoria 3052, Australia

Received October 19, 1987; Revised Manuscript Received December 17, 1987

ABSTRACT: The primary structure of the secreted carbonic anhydrase from ovine salivary glands has been determined by automated Edman sequence analysis of peptides generated by cyanogen bromide and tryptic cleavage of the protein and *Staphylococcus aureus* V8 protease, trypsin, and α-chymotrypsin subdigests of the large cyanogen bromide peptides. The enzyme is a single polypeptide chain comprising 307 amino acids and contains two apparent sites of carbohydrate attachment at Asn-50 and Asn-239. The protein contains two half-cystine residues at 25 and 207 which appear to form an intramolecular disulfide bond. Salivary carbonic anhydrase shows 33% sequence identity with the ovine cytoplasmic carbonic anhydrase II enzyme, with residues involved in the active site highly conserved. Compared to the cytoplasmic carbonic anhydrases, the secreted enzyme has a carboxyl-terminal extension of 45 amino acids. This is the first report of the complete amino acid sequence of a secreted carbonic anhydrase (CA VI).

Several different isozymes of carbonic anhydrase (CA)[1] (carbonate dehydratase, EC 4.2.1.1) have been reported from mammalian sources. The best characterized are the cytoplasmic isozymes termed CA I, CA II, and CA III (Tashian et al., 1983). Another type of isozyme has been isolated from bovine lung membranes (Whitney & Briggle, 1982) and human kidney membranes (Wistrand, 1984). These enzymes appear to be intrinsic membrane proteins and have been called CA IV. A fifth isozyme (CA V) has been isolated from the mitochondria of guinea pig hepatocytes (Dodgson et al., 1980) and its amino-terminal amino acid sequence determined (Hewett-Emmett et al., 1987). In 1979, we described an unusual carbonic anhydrase from the ovine parotid gland and saliva (Fernley et al., 1979) which does not seem to fit into any of these categories. Subsequently, this enzyme has been purified and characterized more thoroughly (Fernley et al., 1984, 1988). The equivalent enzyme has been purified also from rat saliva (Feldstein & Silverman, 1984) and more recently from human saliva (Murakami & Sly, 1987). These enzymes have apparent subunit molecular masses of 42 000–

† This work was supported by the National Health and Medical Research Council of Australia.
* Address correspondence to this author.

[1] Abbreviations: TFA, trifluoroacetic acid; CA, carbonic anhydrase; HPLC, high-performance liquid chromatography; SDS, sodium dodecyl sulfate; PTH, phenylthiohydantoin; Tris·HCl, tris(hydroxymethyl)-aminomethane hydrochloride; TPCK, L-1-(tosylamido)-2-phenylethyl chloromethyl ketone; DTNB, 5,5′-dithiobis(2-nitrobenzoic acid); EDTA, ethylenediaminetetraacetic acid.

0006-2960/88/0427-2815$01.50/0

46 000 daltons and are glycoproteins (Fernley et al., 1988; Feldstein & Silverman, 1984; Murakami & Sly, 1987). Following enzymatic removal of the N-linked carbohydrate residues, the apparent subunit molecular masses fall to 36 000 daltons for both the ovine (Fernley et al., 1988) and human (Murakami & Sly, 1987) enzymes. Antibodies raised against ovine CA II and salivary CA do not cross-react with one another's antigens (Fernley et al., 1988) nor does the rat salivary enzyme cross-react with anti-human CA II antibody (Feldstein & Silverman, 1984). However, Murakami and Sly (1987) found that human CA II did cross-react to some extent with human salivary CA antibody. This suggests that the two isozymes are distinct but may have some amino acid sequence in common. The HPLC profiles of the cyanogen bromide peptides of the two ovine isozymes show no similarity to one another (Fernley et al., 1988), indicating the two are not closely homologous. In order to demonstrate that this salivary enzyme represents a new isozyme of carbonic anhydrase, it needs to be characterized more completely, and to this end, we have determined the complete amino acid sequence of the ovine enzyme. This isozyme has been designated CA VI.

Materials and Methods

Salivary carbonic anhydrase was prepared from ovine saliva by a modification of the procedure described for the ovine parotid enzyme (Fernley et al., 1988). Briefly, saliva was obtained by cannulation of the parotid duct of Merino sheep. The saliva was collected on ice into a container with benzamidine and aprotinin (final concentrations 1.5 $g \cdot L^{-1}$ and 1 $mg \cdot L^{-1}$, respectively). The saliva was filtered and concentrated by ultrafiltration to about 100 mL and then loaded onto a sulfonamide–Sepharose affinity column. The subsequent steps were as described previously (Fernley et al., 1988). TPCK-treated trypsin was from Worthington, and endoproteinase Glu-C (*Staphylococcus aureus* V8 protease) and α-chymotrypsin were from Boehringer-Mannheim. Cyanogen bromide and 4-vinylpyridine were from Aldrich Chemical Co. Sequencer reagents were from Applied Biosystems and HPLC solvents from Mallinckrodt. All other chemicals were of reagent grade.

Generation of Peptides. Enzymatic and chemical cleavages were performed after reductive alkylation of the enzyme with 4-vinylpyridine essentially as described by Andrews and Dixon (1981). The modified protein was precipitated with methanol at −20 °C overnight; 2.5 mg of alkylated protein was taken up in 40 mM Tris·HCl (pH 8.0)/10 mM $CaCl_2$, and 25 μg of TPCK-treated trypsin was added. Digestion occurred at 37 °C for 10 h. Another 2.5 mg of protein was dissolved in 200 μL of 70% formic acid, and 25 mg of cyanogen bromide was added. After being flushed with N_2, the reaction was allowed to proceed in the dark at room temperature for 24 h. After addition of 3 mL of water, the reaction mixture was lyophilized. *Staph. aureus* V8 protease digests were carried out on the larger cyanogen bromide peptides; 1–10 nmol of peptide was dissolved in 200 μL of 0.1 M ammonium acetate (pH 4.1) and 1% (w/w) of enzyme added. Digestion occurred over 48 h at 37 °C. CB3 (2 nmol) was digested with α-chymotrypsin (1% w/w) for 1 h at 37 °C in 40 mM Tris·HCl (pH 8.0)/10 mM $CaCl_2$. All peptides were purified by HPLC. For amino acid analysis, samples were hydrolyzed in 6 M HCl in vacuo for 24 h and analyzed on a Beckman 6300 amino acid analyzer.

Sequencing Procedure. The amino-terminal sequences of the unmodified protein and the peptide fragments were determined by automated Edman degradation in an Applied Biosystems 470 A gas-phase sequencer with an on-line 120 A PTH-amino acid analyzer. Each sample (100–200 pmol) was sequenced twice.

Nature of the Half-Cystine Residues. The concentration of free sulfhydryl groups was determined by using 5,5′-dithiobis(2-nitrobenzoic acid) (DTNB) (Ellman, 1959). Purified salivary CA (3 mg) was dissolved in oxygen-free 6 M guanidine hydrochloride, 0.1 M sodium phosphate buffer, and 1 mM EDTA (final pH 7.4), and DTNB (0.2 μmol) was added. The change in absorbance at 412 nm was measured against an appropriate blank. A molar absorption coefficient of 13 880 cm^{-1} was used for the thionitrobenzoate anion (Gething & Davidson, 1972). To determine the linkage of the half-cystine residues, freshly prepared CA (1 mg) was dissolved in 70% formic acid and cleaved with cyanogen bromide, and the peptides were separated by HPLC as described above. A portion of each peak was oxidized with performic acid (Hirs, 1967) and then hydrolyzed for amino acid analysis as described above. The cysteine-containing peak (as defined by its cysteic acid content) was sequenced as described previously.

Results

The strategy used for determining the primary structure of ovine salivary carbonic anhydrase is shown in Figure 1. Most of the amino acid sequence was derived from the cyanogen bromide peptides with the overlapping sequences provided from the tryptic peptides. *Staph. aureus* V8 protease digests of the larger cyanogen bromide peptides together with the tryptic peptides were used to establish the sequence of these peptides. Alignment of the sequences with those of the cytoplasmic CA's also assisted in establishing the complete sequence. The peptides have been designated according to the type of digest and the theoretical order in which they appear in the sequence.

The amino acid compositions of ovine salivary CA and its cyanogen bromide peptides and the composition derived from its sequence are shown in Table I, and they are in good agreement. The protein consists of 307 amino acids with a calculated molecular mass of 35 565 daltons. This is in excellent agreement with the value (36 000 daltons) obtained by SDS–polyacrylamide gel electrophoresis of the enzymatically deglycosylated enzyme (Fernley et al., 1988).

Sequence analysis of the intact protein yielded the sequence of the first 30 amino acids. Two sequences were evident, the minor sequence (20% of the total) representing the enzyme shortened by three amino acids at the amino-terminal end. This minor sequence was present (and in the same amount) in three preparations of the enzyme from saliva of three different sheep and also in enzyme prepared from the parotid gland. The complete amino acid sequence is shown in Figure 2.

Cyanogen Bromide Peptides. Most of the protein sequence data were derived from sequencing the cyanogen bromide peptides. Nine major peaks were isolated by HPLC (Figure 3). Seven of these contained the complete sequence for ovine CA VI, and the other two were derived from one peak, CB7. This arose because there was an allotypic replacement (isoleucine/methionine) at position 35 of CB7, and half of this sequence could be cleaved by cyanogen bromide. The peptides used here were from a CA VI preparation from one sheep. There was another allotypic replacement involving methionine, in CB3 (position 26), this time with threonine. This peptide also contained another two methionines, one followed by serine, the other by threonine. These bonds are not readily cleaved by cyanogen bromide, and the smaller peptides were not isolated. The sequence data are shown in the supplementary material (see paragraph at end of paper regarding supplementary material).

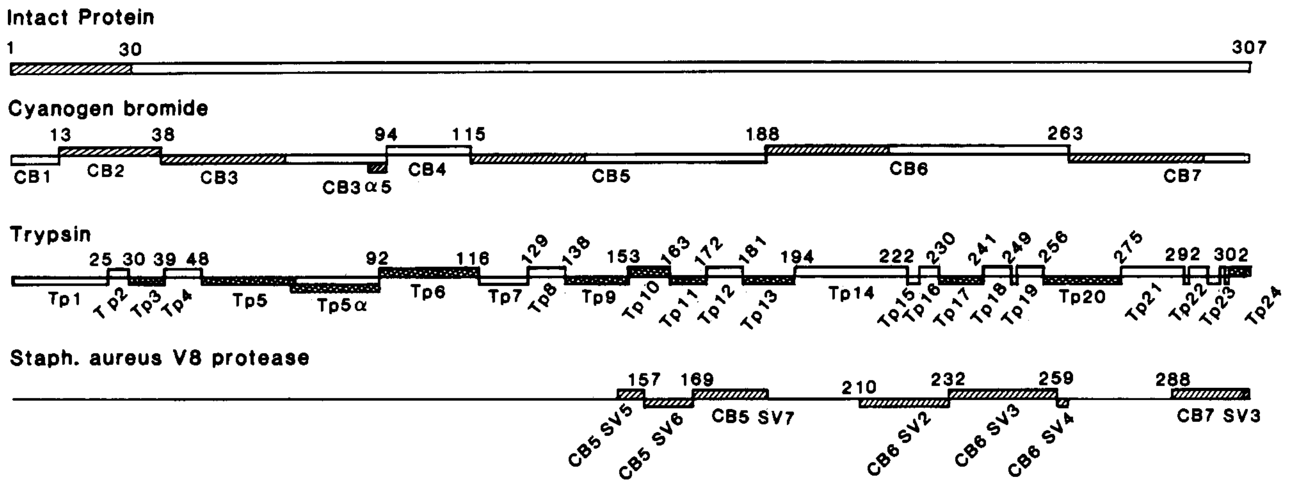

FIGURE 1: Strategy used to determine the amino acid sequence of ovine salivary carbonic anhydrase (CA VI). The peptides expected from each cleavage are indicated and are numbered continuously; the hatched areas represent the portion of the sequence determined by automated Edman degradation. The *Staph. aureus* V8 protease digests and α-chymotryptic digest were carried out on the cyanogen bromide peptides only, and this is indicated in the numbering system (e.g., CB6 SV3).

```
                                               25                                                 50
                                                                                                  Y
G H G V E W T Y S E G M L D E A H W P L E Y P K C G G R R Q S P I D L Q M K K V Q Y N P S L R A L N

                        M                     75                                                 100
L T G Y G L W H G E F P - T N N G H T V Q I S L P S T M S M T T S D G T Q Y L A K Q M H F H W G G A
                        V

                                               125                                               150
S S E I S G S E H T V D G M R Y V I E I H V V H Y N S K Y N S Y E E A Q K E P D G L A V L A A L V E

                                               175                                               200
V K D Y T E N A Y Y S K F I S H L E D I R Y A G Q S T V L R G L D I E D M L P G D L R Y Y Y S Y L G

                                               225                         Y                     250
S L T T P P C T E N V H W F V V A D T V K L S K T Q V E K L E N S L L N H Q N K T I Q N D Y R R T Q

                                               275                                        I      300
P L N H R V V E A N F M S R P H Q E Y T L A S K L H F Y L N N I D Q T L E Y L R R F I E Q K - P K R
                                                                                          M

K K Q E N W P
```

FIGURE 2: Complete amino acid sequence of ovine CA VI. The symbol Y represents the carbohydrate attachment sites at Asn-50 and Asn-235.

Table I: Amino Acid Composition of Ovine Salivary Carbonic Anhydrase and Its Cyanogen Bromide Peptides[a]

	residue								
	CA VI	CB1 (1–12),[g] 15.9 nmol	CB2 (13–37),[g] 14.0 nmol	CB3 (38–93),[g] 16.1 nmol	CB4 (94–114),[g] 8.2 nmol	CB5 (115–187),[g] 13.6 nmol	CB6 (188–262),[g] 10.9 nmol	CB7 (263–307),[g] 7.7 nmol	CB7-C (298–307),[g] 8.0 nmol
Asp	31.4 (30)	0.16 (0)	2.23 (2)	5.19 (5)	1.28 (1)	8.48 (8)	10.06 (10)	4.81 (4)	1.10 (1)
Thr[b]	20.4 (20)	1.14 (1)	0.18 (0)	5.70 (7)	1.04 (1)	2.36 (2)	7.06 (7)	2.40 (2)	0.13 (0)
Ser	23.4 (22)	0.98 (1)	1.01 (1)	4.95 (5)	3.80 (4)	5.35 (5)	4.12 (4)	2.25 (2)	0.27 (0)
Glu	39.0 (39)	2.10 (2)	4.04 (4)	4.85 (5)	2.26 (2)	10.1 (10)	8.60 (8)	7.45 (8)	2.17 (2)
Pro	13.4 (14)	0 (0)	2.93 (3)	3.14 (3)	0 (0)	1.34 (1)	4.09 (4)	2.08 (3)	2.15 (2)
Gly	20.4 (19)	3.01 (3)	1.96 (2)	5.21 (5)	4.12 (4)	3.24 (3)	2.27 (2)	0.60 (0)	0.24 (0)
Ala	14.0 (13)	0.05 (0)	1.13 (1)	2.06 (2)	1.10 (1)	5.64 (6)	2.22 (2)	1.60 (1)	0.18 (0)
$^1/_2$-Cys[c]	2.4 (2)	0 (0)	0.80 (1)	0 (0)	0 (0)	0 (0)	0.78 (1)	0 (0)	0 (0)
Val[d]	17.1 (18.5)	0.99 (1)	0 (0)	2.15 (2.5)[f]	0.8 (1)	6.82 (7)	6.41 (7)	0.97 (0)	0 (0)
Met (Hse)	7.0 (9)	0.6 (1)	0.22 (1)	0.64 (3.5)[f]	0.5 (1)	0.22 (1)	0.9 (1)	0 (0.5)[f]	0 (0)
Ile[d]	11.6 (11.5)	0 (0)	0.95 (1)	1.41 (1)	0.86 (1)	4.67 (5)	1.07 (1)	1.82 (2.5)[f]	0 (0)
Leu	30.8 (29)	0 (0)	2.91 (3)	7.27 (6)	0.38 (0)	5.68 (6)	9.11 (9)	4.81 (5)	0.14 (0)
Tyr	18.6 (21)	0.94 (1)	0.67 (1)	2.21 (3)	0.10 (0)	7.38 (8)	4.65 (5)	2.51 (3)	0 (0)
Phe	8.5 (7)	0 (0)	0 (0)	1.24 (1)	0.74 (1)	1.37 (1)	2.49 (2)	1.81 (2)	0 (0)
His	12.5 (15)	0.95 (1)	0.99 (1)	1.73 (2)	1.98 (3)	2.48 (3)	2.76 (3)	1.97 (2)	0 (0)
Lys	16.2 (17)	0 (0)	1.03 (1)	2.81 (3)	0.25 (0)	3.72 (4)	3.87 (4)	4.47 (5)	2.82 (3)
Arg	14.0 (14)	0 (0)	1.87 (2)	1.31 (1)	0.24 (0)	2.79 (3)	4.00 (4)	3.51 (4)	0.97 (1)
Trp[e]	4.1 (6)	ND[h] (1)	ND (1)	ND (1)	ND (1)	ND (0)	ND (1)	ND (1)	ND (1)

[a] Determined by acid hydrolysis as described in the text. Values in parenthese are from the final sequence. [b] Serine and theronine determined by extrapolation to zero hydrolysis time. [c] Determined as cysteic acid after performic acid oxidation (Hirs, 1967). [d] Valine and isoleucine were determined after 72-h hydrolysis. [e] Tryptophan was determined spectrophotometrically by the method of Edelhoch (1967). [f] Sequencing the tryptic peptide covering this region showed an equal amount of valine and methionine at one position in CB3 and an equal amount of isoleucine and methionine at one position in CB8. [g] Residue number from sequence. [h] Not determined.

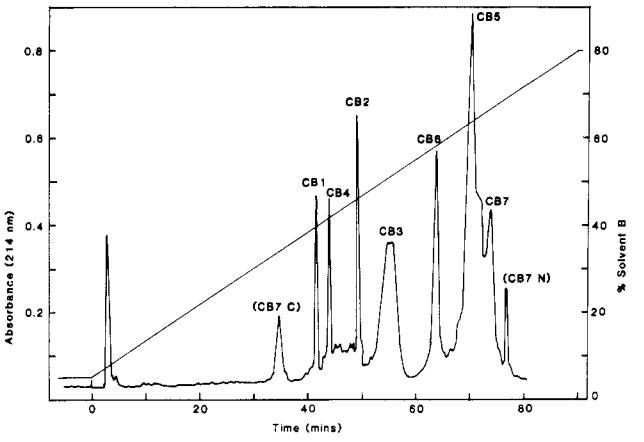

FIGURE 3: Fractionation of the peptides obtained by cyanogen bromide cleavage of reduced and pyridylethylated ovine CA VI on a Waters C_{18} μBondapak reverse-phase HPLC column (3.9 mm × 30 cm). The column was equilibrated in 0.1% TFA/14% acetonitrile, and the peptides were eluted with a linear gradient of 4% acetonitrile to 64% acetonitrile over 90 min. The flow rate was 1 mL·min^{-1}.

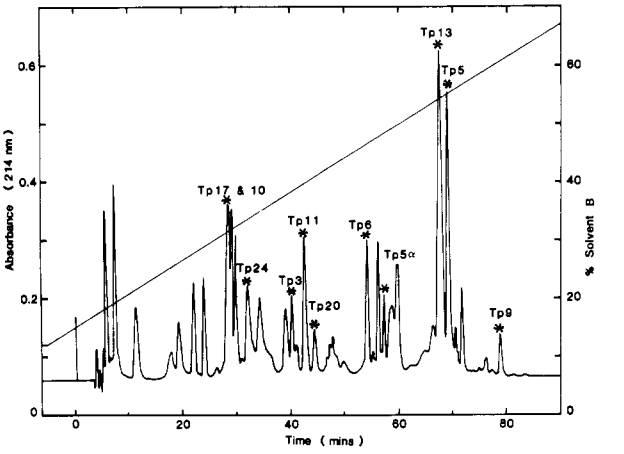

FIGURE 4: Fractionation of peptides obtained by tryptic digestion of pyridylethyl ovine CA VI. The column and conditions used were as described in Figure 3 except a gradient of acetonitrile from 10% to 56% was used.

Peak 1 resulted from cleavage at Met-297 (positions 298–307) and is part of CB7 (both Met and Ile occur in equal amounts at this position). The peptide was sequenced to cycle 10, and this accounted for all its amino acids. It contained no homoserine residues, and it represents the carboxyl terminus of CA VI.

Peak 2 correponded to CB1 (residues 1–12) whose sequence had been determined by sequencing the intact protein. Peak 3 was CB4 (residues 94–114) whose sequence was contained within the tryptic peptide Tp6. Peak 4 (CB 2: residues 13–37) was sequenced to cycle 21, and the remainder of the sequence was provided by sequencing Tp3. It contained one of the two cysteine residues of this protein. Peak 5 corresponded to CB3 (residues 38–93). The first 29 residues were determined by direct sequence analysis, and the rest of the sequence came from sequencing Tp5, Tp5α, and an α-chymotryptic peptide of CB3, α5. Cycle 13 of CB3 was blank, and the two subsequent cycles yielded Leu-Thr. Similarly, Tp5 showed a blank at this position (cycle 3). This is consistent with the consensus sequence (Asn-X-Thr/Ser) for an asparagine-linked carbohydrate side chain (Waechter & Lennarz, 1976). The amino acid composition of CB3 (Table I) shows there are four Asx residues which would tally with the sequence data if there was an Asx at cycle 13. Therefore, this position (50) has been identified as asparagine linked to a carbohydrate side chain.

Peak 6 corresponded to CB6 (positions 188–262). The first 30 residues were sequenced, and then the peptide was subdigested with *Staph. aureus* V8 protease which resulted in 4 peptides being isolated by HPLC (supplementary material). All except the amino-terminal peptide were sequenced in their entirety (SV2, 22 residues; SV3, 27 residues; SV4, 4 residues). SV2 and SV3 were overlapped by Tp17, and SV3 and SV4 were overlapped by Tp20 which also linked CB6 to CB7. Cycle 8 of SV3 was blank and was followed by the sequence Lys-Thr. Similarly, cycle 10 of Tp17 was blank. The measured amino acid compositions of both these peptides would agree with compositions derived from the sequences if there was an Asx in this position. Consequently, this position (239) has been identified as the second asparagine with a carbohydrate side chain.

Peak 7 was CB5 (positions 115–187), and the first 27 amino acids were sequenced. The peptide was subdigested with *Staph. aureus* V8 protease and trypsin, and the two digests were subjected to HPLC. All except the amino-terminal peptides were sequenced in their entirety, and these provided an unambiguous sequence for the whole of CB5. Tp9 extended the sequence of CB5 to position 38 and was overlapped to Tp10 (fourth tryptic peptide of CB5) by CB5 SV5. CB5 SV6 overlapped Tp10 to Tp11 (fifth Tp of CB5), and CB5 SV7 overlapped Tp11, Tp12 (sixth Tp of CB5), and Tp13 (seventh Tp of CB5). Tp13 overlapped CB5 and CB6. Peak 8 corresponded to CB7 (positions 263–307), and the first 33 residues were sequenced. The peptide was subdigested with *Staph. aureus* V8 protease, and the peptides were separated by HPLC. The carboxyl-terminal SV3 peptide provided 18 residues and overlapped with Tp24.

Tryptic Digest. Tryptic digestion of the pyridylethyl-CA yielded a number of peptides which were separated by HPLC (Figure 4). These peptides provided sequence data to overlap the cyanogen bromide peptides and also to overlap *Staph. aureus* V8 protease peptides derived from the cyanogen bromide peptides. The peptides used are labeled in Figure 4. The sequence data are presented in the supplementary material. Tp3 (positions 30–38) provided the carboxyl-terminal sequence of CB2 and overlapped it by one residue to CB3. CB3 was the only cyanogen bromide peptide isolated having Lys at its amino terminus, and sequence homology with ovine CA II (Tanis et al., 1974) suggests this sequence placement is correct. Tp5 (residues 48–91) extends the sequence of CB3, 27 of the 44 residues being sequenced. Another peptide, Tp5α (residues 69–91) was isolated from the tryptic digest of CA VI. This was part of Tp5 resulting from an α-chymotryptic-like cleavage after a histidine residue. This peptide was sequenced to its carboxyl-terminal Lys (23 cycles) and completed the sequence of Tp5.

Tp5 also showed the presence of both Met and Val at cycle 16. Each PTH-amino acid was present at about half the amount of the PTH-amino acid in the preceding cycle. Tp6 (positions 92–115) contained CB4 and overlapped it by one residue to CB5. Again, CB5 was the only cyanogen bromide peptide with Arg at its amino terminus, and the homology around this region was strong with ovine CA II. All 24 residues of Tp6 were sequenced. Tp9, -10, -11, and -13 (positions 138–152, 153–162, 163–171, and 181–193) were present in CB5 and were all sequenced to their carboxyl-terminal basic residues. These peptides were used to overlap the *Staph. aureus* V8 peptides derived from CB5, thus giving its complete sequence. Tp13 also overlapped CB5 to CB6. Tp17 (positions 230–240) overlapped SV2 to SV3 within CB6, and Tp20 (positions 256–274) overlapped SV3 to SV4 and CB6 to CB7. Both were sequenced to their carboxyl-terminal residues. Tp24

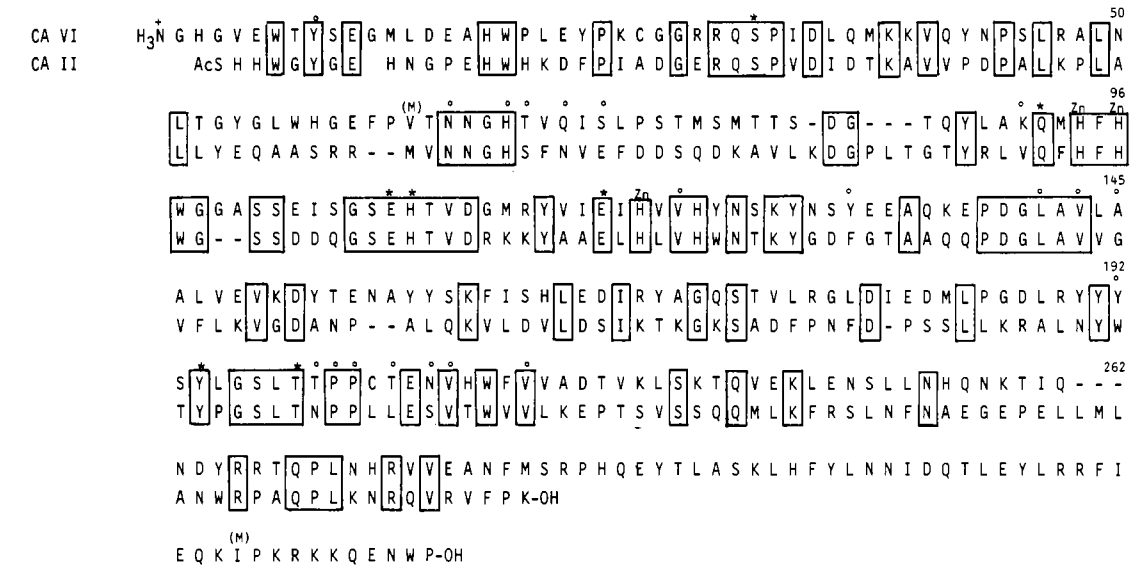

FIGURE 5: Comparison of the amino acid sequences of ovine CA VI and ovine CA II (Tanis et al., 1974). Identical residues in both sequences have been boxed, and gaps have been introduced into the sequences to maximize homology. The three histidine residues involved in binding the essential zinc are labeled Zn as are other residues in the active site hydrogen bonded to these three histidines or the zinc-bound solvent molecule (asterisks). Residues postulated to occur in the active-site cavity are also labeled (o). The numbering system used here is based on that of human CA I so comparisons of active-site residues can be made.

(positions 302–307) was sequenced for six cycles, and this accounted for all its amino acids and with data from sequencing peak 1 of the cyanogen bromide digest defined the carboxyl terminus of the protein.

Staphylococcus aureus V8 Protease Digest. CB peptides 5, 6, and 7 were digested with *Staph. aureus* V8 protease to complete the sequence of these peptides. The HPLC purifications and sequence data are shown in the supplementary material. Individual sequences have been discussed with the appropriate cyanogen bromide peptides.

Nature of the Half-Cystine Residues. After reaction of the enzyme in guanidine hydrochloride with excess DTNB, the absorbance increase at 412 nm was only 0.007, about 1% of the expected value if there was one free sulfhydryl group per subunit. To investigate whether the half-cystines were involved in a disulfide bond, the protein was digested with cyanogen bromide, and the peptides were separated. The CB2 peak disappeared from the HPLC profile, and the CB6 peak had broadened. This peak was the only one to contain cysteic acid (after performic acid oxidation) and when sequenced yielded the CB2 and CB6 sequences in equal amounts, indicating these two peptides are linked by a disulfide bond. Furthermore, reduction of this peptide with dithiothreitol resulted in the appearance of a new peak on HPLC, and this peak had the same amino acid composition as CB2 (data not shown).

Discussion

This paper presents for the first time the complete amino acid sequence of a secreted carbonic anhydrase (CA VI). The sequence has been derived by sequencing the intact protein, the cyanogen bromide peptides, the tryptic peptides, and the *Staph. aureus* V8 protease and α-chymotrypsin peptides derived from the cyanogen bromide peptides. Carbonic anhydrase VI is a single polypeptide chain of 307 amino acids.

This is considerably longer than the cytoplasmic carbonic anhydrases which have about 260 residues. The sequence of ovine CA VI is compared to that of ovine CA II in Figure 5. Most of the additional sequence lies at the carboxyl terminus of the protein, but there are several small insertions within the molecule, relative to CA I, II, or III, including two amino acids in the highly conserved active-site area. The carboxyl-terminal "tail" contains many hydrophilic amino acids and no stretch of hydrophobic residues likely to act as a stop-transfer sequence typical of membrane-bound proteins (Yost et al., 1983). As has been discussed previously (Feldstein & Silverman, 1984), the secreted CAs have very similar amino acid compositions to the membrane-bound CAs (CA IV) on a percentage basis, and both isozymes are glycosylated. It is possible that these two isozymes are closely related, the latter possessing a stop-transfer sequence to anchor it in the cell membrane. However, nothing is known about the structure of CA IV, and confirmation will await sequencing of a membrane-bound CA.

Overall, 33% of the residues in ovine salivary CA are identical with the ovine CA II sequence. This is a low degree of identity when compared with the degree of identity between cytoplasmic isozymes, typically 50–60%. To simplify comparisons, the numbering system is based on the human CA I sequence (Tashian et al., 1983). This is shown in Figure 5. In the amino-terminal region, the degree of homology between salivary CA and CA II is low, and only the residues involved in the aromatic cluster I (Nostrand et al., 1975) have been conserved. These include Trp-5, Tyr-7, Trp-16, and Tyr-20 (Phe in ovine CA II). The region at the active site, as expected, has been more highly conserved. His-94, His-96, and His-119, which bind the essential zinc ion, have been conserved as have Ser-29, Gln-92, Glu-106, His-107, Glu-117, Tyr-194, and Thr-199, which are hydrogen-bonded (directly or indirectly) to the three histidine residues or the zinc-bound solvent molecule. The remaining amino acid in this group, Asn-244, is not conserved, but it is possible a neighboring Asn residue could be involved in this binding. The histidine at position 64 has been conserved. It has been suggested that this His is involved in proton transfer and its absence in CA III may, at least partially, explain its low CO_2 hydrase activity (Carter & Jeffery, 1985). The salivary CA is a high-activity enzyme (Fernley et al., 1988; Feldstein & Silverman, 1984).

As is typical of many secreted (and membrane-bound) proteins, salivary CA is glycosylated (Fernley et al., 1988; Feldstein & Silverman, 1984; Murakami & Sly, 1987), and two N-glycosylation sites have been identified.

The amino acid composition of ovine CA VI shows there are two half-cystine residues per subunit. These are not ti-

tratable with DTNB even in the presence of 6 M guanidine hydrochloride, indicating there are no free thiol groups in the molecule as might be expected for a secreted protein. This is in contrast to some of the cytoplasmic CAs. For example, bovine CA III has five free thiol groups, two of which react readily with DTNB without any effect on enzyme activity (Engberg et al., 1985). Further modification of cysteine residues results in inactivation of the enzyme (Engberg & Lindskog, 1986). Ovine CA VI migrates as a M_r 45000 species in the presence or absence of reducing agents in SDS–polyacrylamide gels (Fernley et al., 1988), indicating (since there are no free thiol groups) that the two cysteines form an intramolecular disulfide bond. This was confirmed by isolating and sequencing the cysteine-containing peptide from the nonreduced protein. Assuming that the ovine CA VI has an overall three-dimensional structure which is similar to that of CA I and CA II (Nostrand et al., 1975), a model of CA VI would have Cys-25 and Cys-207 in close proximity.

This protein is a major constituent of saliva, at least in sheep and humans, and its structure appears to have been conserved between species (R. T. Fernley et al., unpublished experiments). As yet, no function has been proposed for this enzyme though it would be surprising if it was solely a carbonate dehydratase given the high levels of the CA II isozymes in the parotid gland and the high concentrations of bicarbonate in saliva. Further studies on the activities and properties of this unusual carbonic anhydrase may determine its role in saliva.

ACKNOWLEDGMENTS

We thank Michele McDonald and Marie John for skillful assistance with amino acid analyses and amino acid sequencing.

SUPPLEMENTARY MATERIAL AVAILABLE

Five figures showing HPLC profiles of *S. aureus* V8 digests of CA VI CB peptides 5, 6, and 7, the tryptic digest of CB5, and the α-chymotryptic digest of CB3 and four tables showing sequence data for the cyanogen bromide, tryptic, and *S. aureus* V8 peptides and the amino acid composition of tryptic peptides (10 pages). Ordering information is given on any current masthead page.

Registry No. CA, 9001-03-0.

REFERENCES

Andrews, P. C., & Dixon, J. E. (1981) *J. Biol. Chem. 256*, 8267–8270.

Carter, N., & Jeffery, S. (1985) *Biochem. Soc. Trans. 13*, 531–533.

Dodgson, S. J., Foster, R. E., II, Storey, B. T., & Mela, L. (1980) *Proc. Natl. Acad. Sci. U.S.A. 77*, 5562–5566.

Edelhoch, H. (1967) *Biochemistry 6*, 1948–1954.

Ellman, G. L. (1959) *Arch. Biochem. Biophys. 82*, 70–77.

Engberg, P., & Lindskog, S. (1986) *Eur. J. Biochem. 156*, 407–412.

Engberg, P., Millqvist, E., Pohl, G., & Lindskog, S. (1985) *Arch. Biochem. Biophys. 241*, 628–638.

Feldstein, J. B., & Silverman, D. N. (1984) *J. Biol. Chem. 259*, 5447–5453.

Fernley, R. T., Wright, R. D., & Coghlan, J. P. (1979) *FEBS Lett. 105*, 299–302.

Fernley, R. T., Congiu, M., Wright, R. D., & Coghlan, J. P. (1984) *Ann. N.Y. Acad. Sci. 429*, 212–213.

Fernley, R. T., Wright, R. D., & Coghlan, J. P. (1988) *Biochem. J. 249*, 201–207.

Gething, M. J. H., & Davidson, B. E. (1972) *Eur. J. Biochem. 30*, 352–353.

Hewett-Emmett, D., Cook, R. G., & Dodgson, S. J. (1987) *Isozyme Bull. 19*, 13.

Hirs, C. H. W. (1967) *Methods Enzymol. 11*, 197–199.

Murakami, H., & Sly, W. S. (1987) *J. Biol. Chem. 262*, 1382–1388.

Nostrand, B., Vaara, I., & Kannan, K. K. (1975) in *Isozymes: Molecular Structure* (Markert, C. L., Ed.) Vol. 1, pp 575–599, Academic, New York.

Tanis, R. J., Ferrell, R. E., & Tashian, R. E. (1974) *Biochim. Biophys. Acta 371*, 534–548.

Tashian, R. E., Hewett-Emmett, D., & Goodman, M. (1983) *Isozymes: Curr. Top. Biol. Med. Res. 7*, 79–100.

Waechter, C. J., & Lennarz, W. J. (1976) *Annu. Rev. Biochem. 45*, 95–112.

Whitney, P. L., & Briggle, T. V. (1982) *J. Biol. Chem. 257*, 12056–12059.

Wistrand, P. J. (1984) *Ann. N.Y. Acad. Sci. 429*, 195–206.

Yost, C. S., Hedgpeth, J., & Lingappa, V. R. (1983) *Cell (Cambridge, Mass.) 34*, 759–766.