Vol. 96, No. 1, 1980 BIOCHEMICAL AND BIOPHYSICAL RESEARCH COMMUNICATIONS
September 16, 1980 Pages 135-142

THE AMINO ACID SEQUENCE OF CARBONIC ANHYDRASE I FROM THE RHESUS MACAQUE[1]

Dag Henriksson[2], Robert J. Tanis[3] and Richard E. Tashian[4]

Department of Human Genetics
University of Michigan Medical School
Ann Arbor, Michigan 48109

Received July 23,1980

Summary: The complete amino acid sequence of carbonic anhydrase I (CA I) isolated from the red cells of the rhesus macaque (Macaca mulatta) is presented. This sequence was obtained by aligning peptides derived from various fragmentation procedures with the fully characterized sequence of human CA I. When the peptides of rhesus CA I were ordered in this manner, 13 of the 260 residues were found to differ from the human CA I sequence. The known markedly higher specific esterase activity of rhesus CA I compared to human CA I could not be correlated with any changes in residues postulated to be within 10 Å of the single zinc ion at the active site.

INTRODUCTION

Mammalian carbonic anhydrases (EC 4.2.1.1.) are known to occur as three isozymes, designated CA I, CA II and CA III, which are the products of separate genetic loci (1-6). CA I is found in red cells and in regions of a limited number of tissues (e.g., colonic mucosa, rumen epithelium, pituitary gland), whereas CA II is found in red cells and a much wider variety of tissues (cf. 1). CA III appears to be found mainly in red skeletal muscle (7-9). These three isozymes which differ, sometimes markedly, in some of their physico-chemical and catalytic properties are all monomeric metalloenzymes (M.W. ~29,000) containing an essential zinc ion in the active site (cf. 3, 10, 11).

The complete amino acid sequence of eight carbonic anhydrase isozymes have been reported. These are: human CA I (12-14), horse CA I (15), ox CA I (3), human CA II (16,17), sheep CA II (18), ox CA II (19), rabbit CA II (20)

[1]Supported by grant GM-24681 of the U.S. Public Health Service.
[2]Present address: Department of Biochemistry, University of Göteborg and Chalmers Institute of Technology, Fack, S-402 20 Göteborg 5, Sweden.
[3]Present address: Department of Biochemistry, Michigan State University, East Lansing, Michigan 48824.
[4]To whom requests for reprints should be addressed.

0006-291X/80/170135-08$01.00/0

and ox CA III (3). Comparisons of the complete primary structure of all of these isozymes (including rhesus CA I) show that about 37% of the residues at homologous positions are invariant, indicating that these isozymes have evolved from a common ancestral gene. Here, we report the complete primary structure of CA I from the rhesus macaque monkey, Macaca mulatta. This represents part of our continuing studies to obtain comparative sequence data for the analysis of structure-function relationships of the carbonic anhydrase isozymes during their evolution.

MATERIALS AND METHODS

CA I was purified from rhesus hemolysates by affinity chromatography on a sulfonamide (Prontosil) column (21). Since rhesus red cells contain both CA I and CA II, CA I was separated from CA II by selective elution with KI. The purified protein was homogeneous as determined by starch gel electrophoresis (1).

TPCK-trypsin, chymotrypsin and carboxypeptidases A and B were obtained from Worthington Biochemical Corp., and thermolysin from Boehringer Mannheim Biochemicals. Carboxypeptidase Y, hydroxylamine-HCl, citraconic anhydride, cyanogen bromide and chemicals for the automatic sequencing were purchased from Pierce Chemical Co., and Sephadex gels from Pharmacia Fine Chemicals. All other chemicals were reagent grade quality.

Tryptic digestion of rhesus CA I was carried out on acid denatured material as previously described (20). After digestion, the pH was adjusted to 5.0, and peptides insoluble under these conditions were removed by centrifugation. The soluble peptides were then separated either by ion exchange chromatography on Technicon type P resin (22), or by a combination of gel filtration on Sephadex G-50 Fine columns and high voltage paper electrophoresis.

The insoluble material was repeatedly washed with H_2O to remove all soluble peptides. The insoluble peptides were then suspended in H_2O, solubilized by the addition of solid guanidine-HCl, fractionated on a 1 x 110 cm Sephadex G-50 Fine column, and eluted with 5 M guanidine-HCl, pH 5.

Restriction of tryptic cleavage to arginyl bonds was accomplished by modifying the lysyl residues of purified rhesus CA I with citraconic anhydride as previously reported (20). Following dialysis to remove by-products and excess reagents, the modified enzyme was treated with 2% TPCK-trypsin at pH 8.5 for four hours. The digested material was fractionated using a 1.5 x 100 cm Sephadex G-100 column equilibrated with 0.1 M acetic acid.

Cleavage at both methionine and tryptophan residues with cyanogen bromide was performed essentially as described by Ozols and Gerard (23). After the reagents had been removed by lyophilization, the mixture of fragments was dissolved in 3 M NH_4OH at room temperature for one hour in order to improve the yield of the fragments. The mixture was then separated on a 1 x 110 cm Sephadex G-50 Superfine column in 5 M guanidine-HCl, pH 5. Cleavage at methionyl residues only was carried out with cyanogen bromide in 70% formic acid (24).

Cleavage of the only Asn-Gly bond (residues 24 and 25) was effected by hydroxylamine as described by Blumenfeld et al. (25). No separation of the cleaved material was attempted, and only a single sequence was seen when the mixture was sequenced.

Certain peptides were further digested by chymotrypsin, thermolysin or carboxypeptidases (26-29). Amino acid analysis was carried out as previously described (20).

The amino acid sequences were determined with the use of a Beckman 890B sequencer; details of the procedures employed have been previously reported (20). The Beckman sequencing program No. 111374 was used through most of this work. We also observed that the addition of 1-2 mg of human CA I to the sample being sequenced greatly promoted retention of the peptide during the routine extraction steps. Others have noticed a similar effect with other proteins (30). PTH-amino acids were identified using both gas chromatography (SP-400 or SE-30 coated supports) and polyamide-TLC plates (31). Arginine and histidine were identified by specific staining techniques (phenanthrenequinone and Pauly).

RESULTS

The results are summarized in Table I.

Peptides obtained from tryptic hydrolysis of either unmodified or citraconylated material covered the complete primary structure of rhesus CA I. Automatic sequencing procedures were straightforward and yielded data for all residues except 1-10, 24-34 and 212-227.

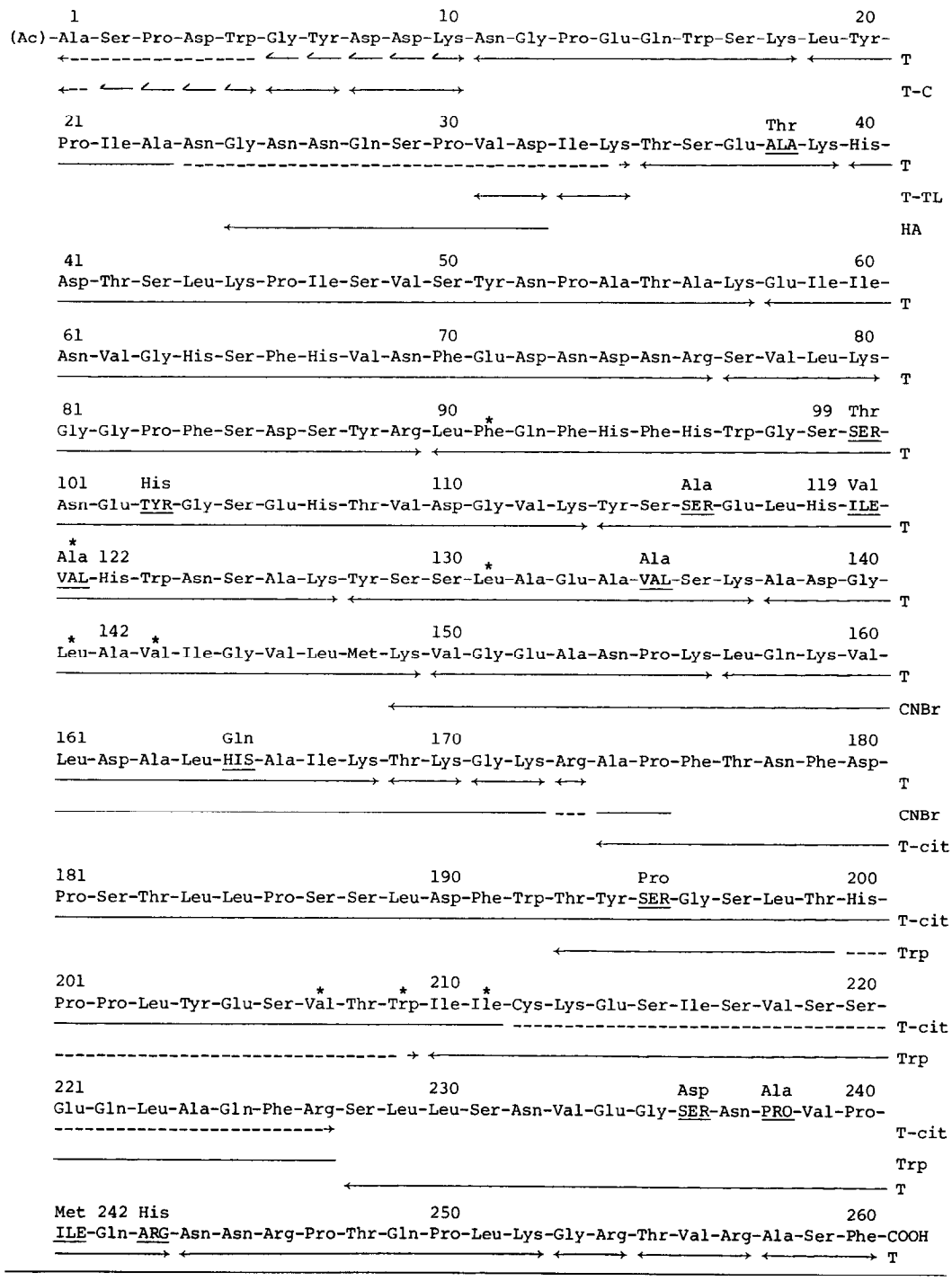
The N-terminal tryptic peptide (residues 1-10) was sequenced by analysis of its C-terminal residues with carboxypeptidases A and B, and sequence analyses or carboxypeptidase Y treatment of peptides produced by chymotryptic digestion of the intact tryptic peptide. This peptide was presumed to be acetylated at the N-terminus on the basis of its failure to undergo Edman degradation.

The sequence of residues 212-227 was obtained by automatic sequencing of fragment 210-260 obtained from CNBr cleavage at Trp-209.

Tryptic peptide 19-34 could only be sequenced for five steps after which no further residues were released. Thermolytic peptides were prepared and sequenced, leaving only residues 24-30 unresolved. Composition and electrophoretic mobility indicated the possibility of an Asn-Gly sequence which is known to frequently block Edman degradation by forming a β-aspartyl bond.

TABLE 1

Amino Acid Sequence of Rhesus Monkey Carbonic Anhydrase I[†]



The whole protein was therefore treated with hydroxylamine and, as no other Asn-Gly sequence had been found, the cleavage mixture was sequenced directly yielding a single sequence starting at residue 25.

DISCUSSION

The complete primary sequence of rhesus macaque carbonic anhydrase I is shown in Table 1. The sequenced fragments were ordered by aligning them with the fully characterized sequence of human CA I (12-14). Those 13 residues which differ from human CA I are indicated. With the exception of the Ser/Asp difference at position 236, which requires two base changes, all of the replacements can result from single base mutations.

When the specific esterase activities of rhesus and human CA I toward β-naphthyl acetate and p-nitrophenyl acetate are compared, the activity of rhesus CA I is about 25 and five times, respectively, that of human CA I (1). This high specific esterase activity also appears to be a characteristic of the CA I isozymes of all of the species of the genus Macaca that have been tested in our laboratory (cf. ref. 32). Because of this marked increase in esterase activity, it was of interest to examine the putative active site residues of the human and rhesus isozymes, especially those proposed by Arqos et al. (33) to be within binding distance of the aromatic structure of p-nitrophenyl acetate in human CA I. These residues are designated by an asterisk in Table 1. The only difference is at position 121 which is alanine; and which is, incidentally, also alanine in chimpanzee CA I (34) whose esterase activity is similar to that of human CA I (35). However, the fact that the

---

†Solid lines indicate residues identified during automatic sequencing. Broken lines indicate that no positive identification was made. Fragments were obtained by cleavage with trypsin (T), chymotryptic cleavage of tryptic peptides (T-C), thermolytic cleavage of tryptic peptides (T-TL), tryptic cleavage after citraconylation (T-cit), cleavage of Met residues with cyanogen bromide (CNBr), cleavage of Met and Trp residues with cyanogen bromide (Trp), cleavage of Asn-Gly bonds with hydroxylamine (HA), and cleavage with carboxypeptidase (⟵ ⟵). See text for details. Residues which differ from human CA I are underlined and capitalized; residues in human CA I are designated above. Asterisks indicate residues postulated to be within contact distance of the aromatic structure of p-nitrophenyl acetate in human CA I (33).

homologous residue at this position is valine in all other carbonic anhydrase isozymes that have been sequenced through this position, i.e., the CA I isozymes of orangutan (1,34) horse (15), ox (3), and turtle (6,34), the CA II isozymes of human (16,17), sheep (18), ox (19), and rabbit (20), and the muscle CA III isozyme of ox (3), suggests that the Val/Ala difference is not directly responsible for the higher esterase activity of rhesus CA I. The other residues that have their side chains in the active site of human CA I, i.e., positions 7, 61, 64, 65, 67, 69, 92, 96, 145, 199, 201, 202, 204, and 206 (cf. ref. 36), are the same in rhesus CA I. Clearly, marked changes in enzyme activity can be produced by changes in residues which may not be located in the active site. In this respect, it is of interest to point out that the specific esterase activity of orangutan CA I is also considerably higher than that of human CA I (1), and as with rhesus CA I, no differences have as yet been found (with the exception of Val-121, see above) between the putative active site residues of human and orangutan CA I (orangutan CA I residues at positions 207, 209, and 211 have not been sequenced).

When the sequence of rhesus CA I is compared with the complete, or partially complete, sequences of the CA I isozymes of human [100% sequenced (16, 17)], chimpanzee [100% sequenced (34)], orangutan [90% sequenced (1,34)], rabbit [60% sequenced (1)], ox [100% sequenced (3)], horse [100% sequenced (15)], and turtle [93% sequenced (6,34)], an estimate of their evolutionary distances can be obtained. The sequence of rhesus CA I was found to differ from the CA I's of the above species by the following percentages: orangutan, 5%; chimpanzee, 5%; human, 5%; rabbit, 16%; horse, 18%; ox, 20%; and turtle, 40%.

ACKNOWLEDGEMENTS

We are very grateful to Dr. David Hewett-Emmett for his help in the preparation of the manuscript. We also thank Roy Morrison for his assistance in confirming the sequence of one of the fragments.

REFERENCES

1. Tashian, R.E., in: Isozymes: Current Topics in Biological and Medical Research, Vol. 2 (Rattazzi, M.C., Scandalios, J.G., and Whitt, G.S., eds.) Alan R. Liss, Inc., New York, pp. 21-62 (1977).
2. Tashian, R.E., Stroup, S.K., Yu, Y.-S.L., and Henriksson, D., Fed. Proc. 37, 1797 (1978).
3. Tashian, R.E., Hewett-Emmett, D., Stroup, S.K., Goodman, M., and Yu, Y.-S.L., in: Biophysics and Physiology of Carbon Dioxide (Bauer, C., Gros, G., and Bartels, H., eds.) Springer-Verlag, Berlin, Heidelberg, New York, pp. 165-176 (1980).
4. Holmes, R.S., Eur. J. Biochem. 78, 511 (1977).
5. Koester, M.K., Register, A.M., and Noltmann, E.A., Biochem. Biophys. Res. Commun. 76, 196 (1977).
6. Tashian, R.E., Hewett-Emmett, D., and Goodman, M., in: Protides of the Biological Fluids (Peeters, H., ed.) Pergamon Press, Ltd., Oxford (in press).
7. Holmes, R.S., Comp. Biochem. Physiol. 57B, 117 (1977).
8. Moynihan, J.B., Biochem. J. 168, 567 (1977).
9. Carter, N.D., Jeffery, S., Shiels, A., Edwards, Y., Tipler, T., and Hopkinson, D.A., Biochem. Genet. 17, 837 (1979).
10. Lindskog, S., Henderson, L.E., Kannan, K.K., Liljas, A., Nyman, P.O., and Strandberg, B., in: The Enzymes, Vol. 5 (Boyer, P.D., ed.) Academic Press, New York, pp. 587-665 (1971).
11. Pocker, Y., and Sarkanen, S., Adv. Enzymol. 47, 149 (1978).
12. Andersson, B., Nyman, P.O., and Strid, L., Biochem. Biophys. Res. Commun. 48, 670 (1972).
13. Lin, K.-T.D., and Deutsch, H.F., J. Biol. Chem. 248, 1885 (1973).
14. Giraud, N., Marriq, C., and Laurent-Tabusse, G., Biochemie 56, 1031 (1975).
15. Jabusch, J.R., Bray, R.P., and Deutsch, H.F., J. Biol. Chem. (in press).
16. Henderson, L.E., Henriksson, D., and Nyman, P.O., Biochem. Biophys. Res. Commun. 52, 1388 (1973).
17. Lin, K.-T.D., and Deutsch, H.F., J. Biol. Chem. 249, 2329 (1974).
18. Tanis, R.J., Ferrell, R.E., and Tashian, R.E., Biochim. Biophys. Acta 371, 534 (1974).
19. Sciaky, M., Limozin, N., Filippi-Foveau, D., Gulian, J.-M., and Laurent-Tabusse, G., Biochemie 58, 1071 (1976).
20. Ferrell, R.E., Stroup, S.K., Tanis, R.J., and Tashian, R.E., Biochim. Biophys. Acta 533, 1 (1978).
21. Osborne, W.R.A., and Tashian, R.E., Anal. Biochem. 64, 297 (1975).
22. Tanis, R.J., Ferrell, R.E., and Tashian, R.E., Biochim. Biophys. Acta 371, 534 (1974).
23. Ozols, J., and Gerard, C., J. Biol. Chem. 252, 5986 (1977).
24. Gross, E., in: Methods in Enzymology, Vol. 11 (Hirs, C.H.W., ed.) Academic Press, New York, pp. 238-255 (1967).
25. Blumenfeld, O.O., Rojkind, M., and Gallop, P.M., Biochemistry 4, 1780 (1965).
26. Smyth, D.G., in: Methods in Enzymology, Vol. 11 (Hirs, C.H.W., ed.) Academic Press, New York, pp. 214-231 (1967).
27. Ambler, R.P., and Meadway, R.J., Biochem. J. 108, 893 (1968).
28. Ambler, R.P., in: Methods in Enzymology, Vol. 11 (Hirs, C.H.W., ed.) Academic Press, New York, p. 155 (1967).
29. Hayashi, R., Moore, S., and Stein, W.H., J. Biol. Chem. 248, 2296 (1973).
30. Rochat, H., Bechis, G., Kopeyan, C., Gregoire, J., and van Rietschoten, J., FEBS Letters 64, 404 (1976).
31. Summers, M.R., Smythers, G.W., and Oroszlan, S., Anal. Biochem. 53, 624 (1973).
32. Tashian, R.E., Goodman, M., Headings, V.E., DeSimone, J., and Ward, R.H., Biochem. Genet. 5, 183 (1971).

33. Argos, P., Garavito, R.M., Eventoff, W., Rossmann, M.G., and Brändén, C.I., J. Mol. Biol. 126, 141 (1978).
34. Contel, E., Stroup, S.K., and Tashian, R.E. (unpublished data).
35. Tashian, R.E., Douglas, D.P., and Yu, Y.-S.L., Biochem. Biophys. Res. Commun. 14, 256 (1964).
36. Kannan, K.K., in: Biophysics and Physiology of Carbon Dioxide (Bauer, C., Gros, G., and Bartels, H., eds.) Springer-Verlag, Berlin, Heidelberg, New York, pp. 184-205 (1980).