Volume 212, number 1, 98–102 FEB 04426 February 1987

# Primary structure of the biotin-binding site of chicken liver acetyl-CoA carboxylase

Toshiyuki Takai, Kenji Wada and Tadashi Tanabe

*Laboratory of Cell Biology, National Cardiovascular Center Research Institute, Fujishiro-dai, Suita, Osaka 565, Japan*

Received 8 December 1986

Limited proteolysis of chicken liver acetyl-CoA carboxylase by staphylococcal serine proteinase yielded a fragment of 31 kDa which contained the biotinyl active site. This polypeptide was purified by preparative polyacrylamide gel electrophoresis and characterized. The complete amino acid sequence of this polypeptide has been deduced from the nucleotide sequence of cloned DNA complementary to the chicken liver acetyl-CoA carboxylase mRNA. A highly conserved sequence of Met-Lys-Met was found in the biotin-binding site. Appreciable homology was observed among the sequences in close vicinity of the biotin sites of chicken liver acetyl-CoA carboxylase and other biotin-dependent carboxylases including biotin carboxyl carrier protein of *Escherichia coli* acetyl-CoA carboxylase.

Acetyl-CoA carboxylase; Biotin-binding site; cDNA cloning; Nucleotide sequence; Primary structure

## 1. INTRODUCTION

Acetyl-coenzyme A carboxylase, a biotin enzyme, plays a critical role in controlling the rate of fatty acid biosynthesis [1,2]. This enzyme from animal species is composed of one kind of subunit with a large $M_r$ of 220 000–260 000 [3–7] and carries three catalytic functions including biotin carboxyl carrier protein, biotin carboxylase and carboxyltransferase as well as the regulatory functions [3,8]. Studies on the structural organization of the catalytic and regulatory domains of acetyl-CoA carboxylase are important to resolve the reaction mechanisms and evolution of the enzyme. Here, the primary structure of the biotin-binding site of chicken acetyl-CoA carboxylase, which possibly corresponds to the biotin carboxyl carrier protein moiety, has been deduced by cloning and sequencing cDNA encoding the enzyme.

Correspondence address: T. Takai, Laboratory of Cell Biology, National Cardiovascular Center Research Institute, Fujishiro-dai, Suita, Osaka 565, Japan

*Abbreviations:* HPLC, high-performance liquid chromatography; PTH, phenylthiohydantoin

## 2. MATERIALS AND METHODS

Acetyl-CoA carboxylase was purified from frozen chicken liver as described [6]. 50 mg purified enzyme was boiled for 2 min in 50 mM Tris-HCl buffer, pH 8.0, containing 0.5% SDS and 1 mM EDTA, and then incubated with 0.5 mg staphylococcal serine proteinase at 30°C for 1.5 h. Polypeptides were isolated by electrophoresis on 0.1% SDS/12.5% polyacrylamide gels in a preparative gel electrophoresis apparatus (model CD-50, Advantec Toyo) using the system of [9]. Alternatively, native acetyl-CoA carboxylase was digested with *Achromobacter* proteinase I followed by separation of the fragments with reverse-phase HPLC. Amino acid sequence analysis was carried out with a gas-phase sequencer, and PTH-amino acid analyzer (model 470A and 120A, Applied Biosystems). The polypeptide fragment containing biotin was identified according to [10] (see fig.1).



*Published by Elsevier Science Publishers B.V. (Biomedical Division)*
00145793/87/$3.50 

Total RNA was extracted from liver of an egg-laying hen as in [11], and poly(A)$^+$ RNA isolated as in [12]. A library of cDNA clones was prepared by the method of Okayama and Berg [13], using 19.2 $\mu$g poly(A)$^+$ RNA and 4.2 $\mu$g vector-primer DNA. The procedures for transformation and screening were as in [14]. Six oligodeoxyribonucleotide probes were synthesized on an automated DNA synthesizer (model 380A, Applied Biosystems): 5′-CCCCANAC$^A_G$TT$^T_C$TT-3′ (probe a; N denotes a mixture of A, T, G and C), 5′-GG$^A_G$TC$^A_G$TT$^T_C$TC$^T_C$TT-3′ (probe b), 5′-GC$^A_G$T-A$^T_C$TGNGCCATCAT$^T_C$TC-3′ (probe c), 5′-T-A$^A_G$TG$^A_G$TCNGCCAT-3′ (probe d), 5′-TCNG-C$^A_G$AA$^T_C$TGCAT-3′ (probe e) and 5′-C$^{TA}_{GG}$TT$^T_C$T-G$^T_C$TG$^A_G$AA-3′ (probe f). Probes a–f, labelled with $^{32}$P at the 5′-end, were used for hybridization at 37°C for probes a, b and f, 38°C for probes d and e and 50°C for probe c. DNA sequencing was carried out according to Maxam and Gilbert [15] and by the dideoxy chain-termination method [16] as described [17].

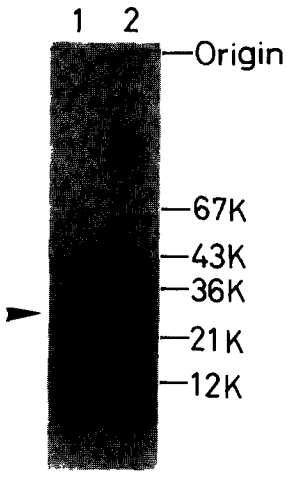


Fig.1. Identification of biotinyl peptide. Acetyl-CoA carboxylase was treated with staphylococcal serine proteinase. 4 $\mu$g of the peptide fragments were incubated with (lane 2) or without (lane 1) 10 $\mu$g avidin, and then subjected to SDS-polyacrylamide gel electrophoresis (7–15% polyacrylamide) [9]. The size markers used were bovine serum albumin, ovalbumin, lactate dehydrogenase, trypsin inhibitor and cytochrome *c*. The peptide of 31 kDa is indicated by an arrowhead.

## 3. RESULTS AND DISCUSSION

Acetyl-CoA carboxylase from chicken liver was subjected to limited proteolysis by staphylococcal serine proteinase. After incubation with avidin, the digest was subjected to electrophoresis. Fig.1 shows that the peptide of 31 kDa shifts to a higher molecular mass region (i.e. to 97–120 kDa) suggesting that it is complexed with avidin [10]. This peptide was isolated electrophoretically and subjected to amino-terminal sequence analysis (fig.2A). The peptide of 31 kDa isolated was further cleaved with *Achromobacter* proteinase I into fragments of 25 and 54 kDa. Each peptide was also isolated and subjected to automated Edman degradation (fig.2B,C). The peptide fragments generated by extensive digestion of the intact acetyl-CoA carboxylase with *Achromobacter* proteinase I were examined for their amino acid sequences. The sequences were (in one-letter code)

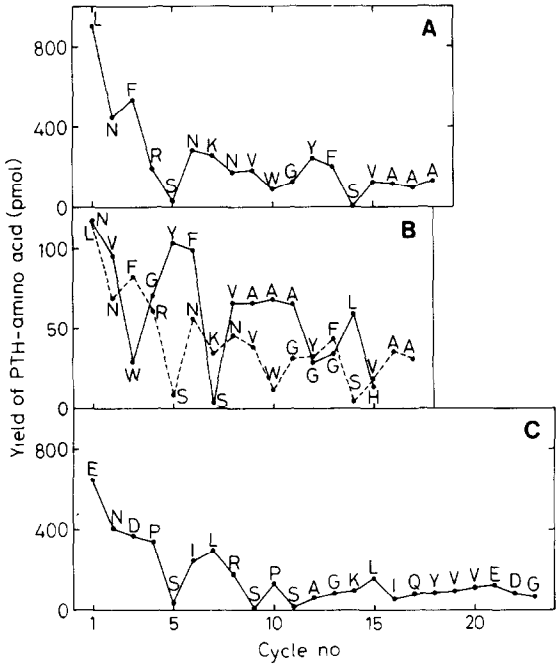


Fig.2. Partial amino acid sequence analysis of peptide fragments. The yield of PTH-amino acids at each cycle of Edman degradation of the peptide of 31 kDa (A), 25 kDa (B) and 54 kDa (C) is shown. The one-letter amino acid notation is used. In B, the preparation of the peptide of 25 kDa was a mixture of two peptides, the sequences of which were assigned as indicated.

EMMAQYASNITS, MADHYVPVPGGPMFR, EREEFLIPIYHQVAMQFAD and LLETESFQQNRIDTG.

A cDNA library derived from chicken liver poly(A)$^+$ RNA was screened by hybridization with each of oligodeoxyribonucleotide probes a–f (see section 2). Probe a corresponds to the sequence (in one-letter code) KNVWG contained in the peptides of 31 and 25 kDa (fig.2A,B). Probe b was synthesized on the basis of the sequence ENDP in the peptide of 54 kDa (fig.2C) and assuming that the preceding residue of the peptide is lysine because this peptide was generated (fig.2B,C) by digestion with *Achromobacter* proteinase I which cleaves the carboxyl side of a lysine residue [18]. Probes c, d, e and f correspond to the sequences, EMMAQYA, MADHY, MQFAD and FQQNR, respectively, all of which were contained in the peptide fragments generated by cleavage of the intact acetyl-CoA carboxylase with *Achromobacter* proteinase I (see above). From about $6 \times 10^5$ transformants, two clones (pACC33 and pACC34), hybridizing with probes b, c and f were isolated. No other transformants gave a hybridization signal with more than one probe. Restriction endonuclease mapping of pACC33 and pACC34 was unable to distinguish them suggesting that they are siblings.

The primary structure of chicken acetyl-CoA

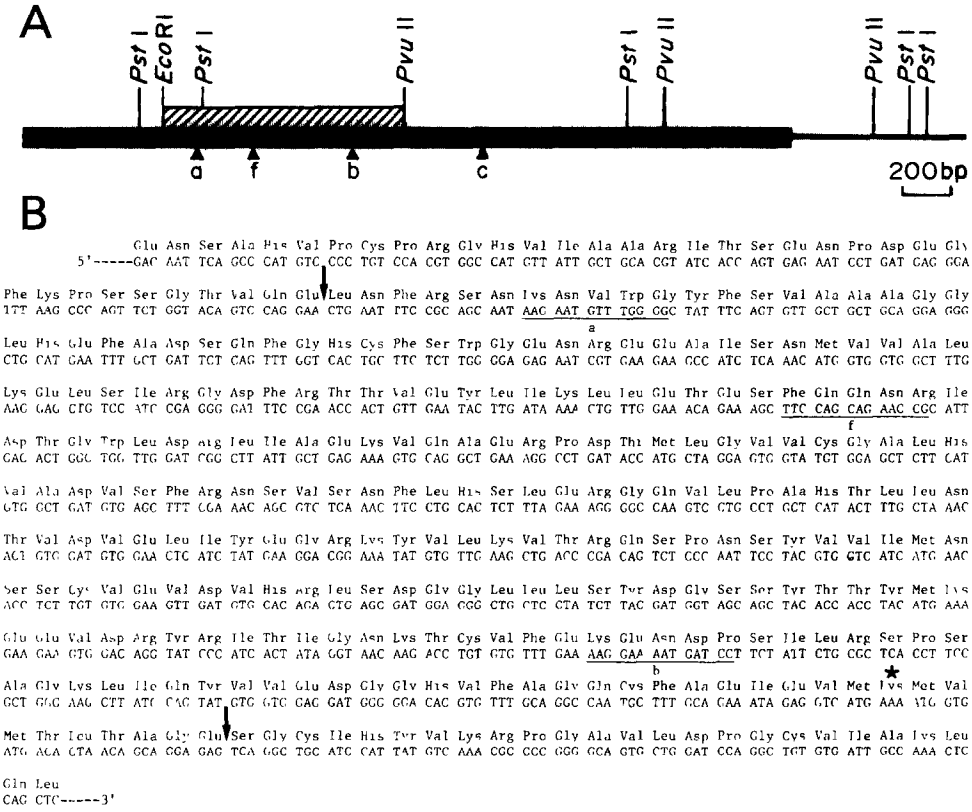


Fig.3. Restriction endonuclease map of the cDNA insert of clone pACC33 (A) and nucleotide sequence including the 978-bp *Eco*RI-*Pvu*II fragment (B). (A) The restriction map shows only the relevant restriction sites. The poly(dA)·poly(dT) tract and the poly(dG)·poly(dC) tail are not included in the map. The protein coding region is indicated by a closed box, and the 978-bp *Eco*RI-*Pvu*II fragment by a hatched box. The positions of the sequences corresponding to probes a, b, c and f are indicated by arrowheads. (B) The deduced amino acid sequence is given above the nucleotide sequence, and amino acid residues are tentatively numbered beginning with the first glutamic acid residue. The tentative number of the amino acid residue at the right-hand end of each line is shown. Underlined portions correspond to the sequences complementary to probes a, b and f. The amino-terminus and putative carboxy-terminus of the biotinyl peptide of 31 kDa (see text) are indicated by arrows. The biotin acceptor site is indicated by an asterisk.

carboxylase was deduced from the cDNA sequence by using the reading frame corresponding to the 4 partial amino acid sequences utilized for synthesizing probes a, b, c and f (fig.3). Although probe a apparently did not hybridize with clone pACC33 in the primary screening, the corresponding nucleotide sequence is contained in this clone. This may be due to inadequate hybridization conditions for the screening with probe a.

The deduced amino acid sequence presented in fig.3 contains the amino-terminal sequence of the peptide of 31 kDa as well as those of the peptides of 25 and 54 kDa. As staphylococcal serine proteinase preferentially cleaves the carboxyl side of glutamic acid residues [19], the carboxy-terminus of the biotinyl peptide of 31 kDa was tentatively assigned to glutamic acid at position 303 (fig.3; the numbering hereafter refers to the amino acid sequences as shown in this figure), because the calculated molecular mass of 29951 Da agrees with the 31 kDa estimated by SDS-polyacrylamide gel electrophoresis and because the amino acid composition of the peptide of 31 kDa deduced from the DNA sequence was consistent with the result of amino acid analysis of the peptide (not shown). On the other hand, the biotinyl prosthetic group of animal acetyl-CoA carboxylase as well as that of other carboxylases is linked to a lysine residue of the enzyme protein [1]. The known primary structures of bacterial and mitochondrial carboxylases [20–23] show high sequence homology in the vicinity of the specific lysine residue [24]. In particular, the Ala-Met-(biotinyl)Lys-Met sequence has been found in all the carboxylases. When the deduced sequence of the peptide of 31 kDa is compared with those of the other carboxylases (fig.4), appreciable sequence homology is found in the neighboring region of the lysine residue at position 294. Therefore, this lysine residue of the chicken liver enzyme was assigned as the biotin acceptor site. Although the typical sequence of Met-Lys-Met is also conserved in the chicken liver enzyme, a substitution of the valine at position 292 for alanine is found. It is noteworthy that the homology in the vicinity of the biotinyl lysine is high with respect to hydrophobic amino acids or to neutral amino acids. These amino acids are possibly important for the recognition of the biotin-binding site by biotin carboxylase [1] or by holoenzyme synthetase which biotinates the apocarboxylases [24].

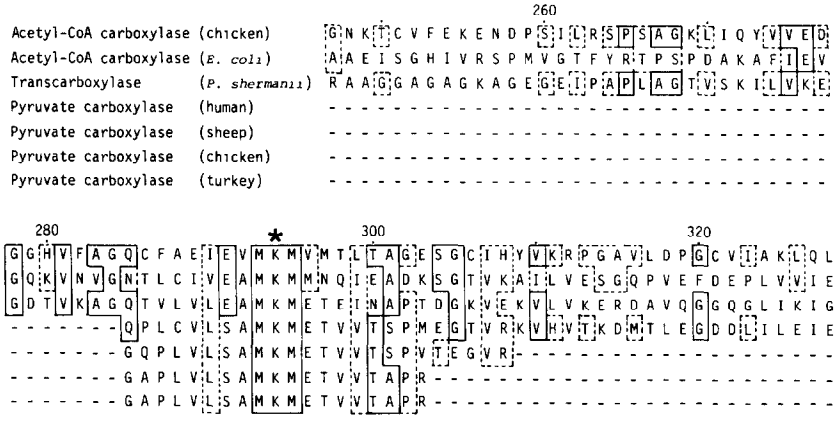


Fig.4. Comparison of the amino acid sequences around the biotin-binding site for carboxylases. The sequence of chicken acetyl-CoA carboxylase (top) was compared with those of biotin carboxyl carrier protein of *E. coli* acetyl-CoA carboxylase (second row) [20] and *Propionibacterium shermanii* transcarboxylase (third row) [21], and pyruvate carboxylase from human (fourth row) [22], sheep (fifth row), chicken (sixth row) and turkey (bottom) [23]. The one-letter amino acid notation was used. The peptides are aligned at the biotinyl lysine residue indicated by an asterisk. Residues identical with those of chicken acetyl-CoA carboxylase are enclosed with solid lines, and residues showing conservative substitution are enclosed with dotted lines. Conservative amino acid substitutions are defined as pairs of residues belonging to one of the following groups: S, T, P, A and G; N, D, E and Q; H, R and K; M, I, L and V; F, Y and W [25]. Hyphens (-) indicate the sequences not available. Position numbers tentatively assigned correspond to those in fig.3.

## ACKNOWLEDGEMENTS

We thank Nikkaki Co. for determining some amino acid sequences and for synthesizing oligodeoxyribonucleotide probes. This investigation was supported in part by research grants from the Ministry of Education, Science, and Culture of Japan.

## REFERENCES

[1] Numa, S. and Tanabe, T. (1984) in: Fatty Acid Metabolism and Its Regulation (Numa, S. ed.) New Comprehensive Biochemistry, vol.7, pp.1–27, Elsevier, Amsterdam, New York.
[2] Wakil, S.J., Stoops, J.K. and Joshi, V.C. (1983) Annu. Rev. Biochem. 52, 537–579.
[3] Tanabe, T., Wada, K., Okazaki, T. and Numa, S. (1975) Eur. J. Biochem. 57, 15–24.
[4] Song, C.S. and Kim, K.-H. (1981) J. Biol. Chem. 256, 7786–7788.
[5] Beaty, N.B. and Lane, M.D. (1982) J. Biol. Chem. 257, 924–929.
[6] Wada, K. and Tanabe, T. (1983) Eur. J. Biochem. 135, 17–23.
[7] Hardie, D.G. and Cohen, P. (1978) Eur. J. Biochem. 92, 25–34.
[8] Obermayer, M. and Lynen, F. (1976) Trends Biochem. Sci. 1, 169–171.
[9] Laemmli, U.K. (1970) Nature 227, 680–685.
[10] Hunaiti, A.R. and Kolattukudy, P.E. (1982) Arch. Biochem. Biophys. 216, 362–371.
[11] Chirgwin, J.M., Przybyla, A.E., MacDonald, R.J. and Rutter, W.J. (1979) Biochemistry 18, 5294–5299.
[12] Aviv, H. and Leder, P. (1972) Proc. Natl. Acad. Sci. USA 69, 1408–1412.
[13] Okayama, H. and Berg, P. (1982) Mol. Cell. Biol. 2, 161–170.
[14] Takai, T., Noda, M., Furutani, Y., Takahashi, H., Notake, M., Shimizu, S., Kayano, T., Tanabe, T., Tanaka, K., Hirose, T., Inayama, S. and Numa, S. (1984) Eur. J. Biochem. 143, 109–115.
[15] Maxam, A.M. and Gilbert, W. (1980) Methods Enzymol. 65, 499–560.
[16] Sanger, F., Nicklen, S. and Coulson, A.R. (1977) Proc. Natl. Acad. Sci. USA 74, 5463–5467.
[17] Furutani, Y., Notake, M., Fukui, T., Ohue, M., Nomura, H., Yamada, M. and Nakamura, S. (1986) Nucleic Acids Res. 14, 3167–3179.
[18] Masaki, T., Fujihashi, T., Nakamura, K. and Soejima, M. (1981) Biochim. Biophys. Acta 660, 51–55.
[19] Houmard, J. and Drapeau, G.R. (1972) Proc. Natl. Acad. Sci. USA 69, 3506–3509.
[20] Sutton, M.R., Fall, R.R., Nervi, A.M., Alberts, A.W., Vagelos, P.R. and Bradshaw, R.A. (1977) J. Biol. Chem. 252, 3934–3940.
[21] Maloy, W.L., Bowien, B.U., Zwolinski, G.K., Kumar, K.G., Wood, H.G., Ericsson, L.H. and Walsh, K.A. (1979) J. Biol. Chem. 254, 11615–11622.
[22] Freytag, S.O. and Collier, K.J. (1984) J. Biol. Chem. 259, 12831–12837.
[23] Rylatt, D.B., Keech, D.B. and Wallace, J.C. (1977) Arch. Biochem. Biophys. 183, 113–122.
[24] Wood, H.G. and Barden, R.E. (1977) Annu. Rev. Biochem. 46, 385–413.
[25] Dayhoff, M.O., Schwartz, R.M. and Orcutt, B.C. (1978) in: Atlas of Protein Sequence and Structure, vol.5, suppl.3 (Dayhoff, M.O. ed.) pp.345–352, National Biomedical Research Foundation, Silver Spring, MD.