1994

Molecular cloning of cDNAs and genes coding for beta-methylcrotonyl-CoA carboxylase of tomato

Xun Wang  
*Iowa State University*

Eve S. Wurtele  
*Iowa State University, mash@iastate.edu*

Greg Keller  
*Iowa State University*

Angela L. McKean  
*Iowa State University*

Basil J. Nikolau  
*Iowa State University, dimmas@iastate.edu*

---

Follow this and additional works at: [http://lib.dr.iastate.edu/bot_pubs](http://lib.dr.iastate.edu/bot_pubs)

Part of the [Biochemistry Commons](http://lib.dr.iastate.edu/bot_pubs), [Botany Commons](http://lib.dr.iastate.edu/bot_pubs), [Genetics Commons](http://lib.dr.iastate.edu/bot_pubs), [Molecular Biology Commons](http://lib.dr.iastate.edu/bot_pubs), and the [Molecular Genetics Commons](http://lib.dr.iastate.edu/bot_pubs)

---

**Recommended Citation**

Wang, Xun; Wurtele, Eve S.; Keller, Greg; McKean, Angela L.; and Nikolau, Basil J., "Molecular cloning of cDNAs and genes coding for beta-methylcrotonyl-CoA carboxylase of tomato" (1994). *Botany Publication and Papers. 68.*  
[http://lib.dr.iastate.edu/bot_pubs/68](http://lib.dr.iastate.edu/bot_pubs/68)

---

This Article is brought to you for free and open access by the Botany at Iowa State University Digital Repository. It has been accepted for inclusion in Botany Publication and Papers by an authorized administrator of Iowa State University Digital Repository. For more information, please contact digirep@iastate.edu.
Molecular Cloning of cDNAs and Genes Coding for β-Methylcrotonyl-CoA Carboxylase of Tomato*

(Received for publication, October 25, 1993, and in revised form, January 13, 1994)

Xun Wang, Eve Syrkin Wurtele, Greg Keller, Angela L. McKean, and Basil J. Nikolau

From the Department of Biochemistry and Biophysics and Department of Botany, Iowa State University, Ames, Iowa 50011.

Tomato cDNA and genomic clones were isolated by using as a probe a cDNA clone that had originally been identified by its ability to direct the synthesis of a biotin-containing polypeptide in Escherichia coli. The nucleotide sequences of the newly isolated cDNAs indicate that they are clones of a single mRNA molecule. However, one of the cDNA clones contains an insertion of a sequence which we identified as an unspliced intron. The amino acid sequence deduced from the nucleotide sequence of the cDNAs showed similarity to regions of previously sequenced biotin enzymes, indicating that the isolated cDNAs code for a biotin-containing protein. Portions of the cDNAs were expressed in E. coli as glutathione S-transferase or β-galactosidase fusion proteins. Each fusion protein was purified and used to immunize rabbits. The resulting antisera recognized a 78-kDa biotin-containing polypeptide in tomato leaf extracts. In addition, both antisera specifically inhibited β-methylcrotonyl-CoA carboxylase activity in extracts from tomato leaves. These characteristics have identified the isolated tomato cDNAs and genes as coding for the 78-kDa biotin subunit of β-methylcrotonyl-CoA carboxylase. Comparison of the deduced amino acid sequence of the biotin subunit of β-methylcrotonyl-CoA carboxylase with other biotin enzymes suggests that this subunit contains the biotin carboxylase and biotin carboxyl-carrier domains.

Biogenin is an essential cofactor for a set of enzymes involved in diverse metabolic processes, such as lipid metabolism, amino acid metabolism, and carbohydrate metabolism (Moss and Knowles, 1987; Wood and Barden, 1977; Moss and Lane, 1971). Although each enzyme has distinct metabolic functions, all biotin enzymes share the same catalytic mechanism. The biotin prosthetic group in these enzymes serves as an intermediate carrier of the carboxyl group being transferred from one substrate to another (Moss and Lane, 1971; Wood and Barden, 1977; Knowles, 1989). Knowledge of biotin-containing enzymes has been gathered mostly from the animal and microbial kingdoms (Dakshinamurti and Bhagavan, 1985; Wood and Barden, 1977; Moss and Lane, 1971). Recently, molecular cloning of genes or cDNAs coding for biotin-containing enzymes from these kingdoms has enabled the elucidation of the primary sequences of biotin-containing enzymes. Thus, the primary sequences of some biotin-containing enzymes have been deduced (Li and Cronan, 1992a, 1992b; Kondo et al., 1991; Lopez-Cassillas et al., 1988; Takai et al., 1987, 1988; Al-Feel et al., 1992; Zhang et al., 1993; Lim et al., 1988; Freytag and Collier, 1984; Browner et al., 1989; Lambonwah et al., 1986, 1989; Kraus et al., 1986; Murfit et al., 1985; Woelke et al., 1992; Schwarz and Oesterhelt, 1985; Schwarz et al., 1988; Laussermair et al., 1988; Genbaufe and Cooper, 1991).

In contrast, knowledge of biotin-containing enzymes from plants is limited. Indeed, until recently, the only biotin-containing enzyme known in plants was acetyl-CoA carboxylase. However, additional biotin-containing polypeptides have been detected in extracts from plants (Nikolau et al., 1987, 1984a, 1984b, 1985; Wurtele and Nikolau, 1992; Nikolau et al., 1993). In addition to acetyl-CoA carboxylase, three other biotin-containing enzyme activities have been reported in plants, these being propionyl-CoA carboxylase, pyruvate carboxylase, and MCCase1 (Wurtele and Nikolau, 1990, 1992). However, the identity of the biotin-containing polypeptide associated with each of these activities is not clear. Because of the poor understanding of the plant biotin-containing proteins, the present study was undertaken, which resulted in the identification of a cDNA clone coding for the biotin-containing subunit of MCCase of tomato.

EXPERIMENTAL PROCEDURES

Materials—Restriction endonucleases and DNA modifying enzymes were purchased from Bethesda Research Laboratories, United States Biochemicals Corp., or Promega-Biotechnology and were used as suggested by the suppliers. Aridin-agarose affinity matrix was purchased from Stratagene Cloning Systems, La Jolla, CA. The radioisotopes were purchased from ICN Biomedicals. All other reagents and biochemicals were purchased from Sigma, United States Biochemical Corp., or Fisher Scientific.

Tomato (Lycopersicum esculentum var. Rutgers) were grown in a greenhouse at 18–22 °C under supplemented illumination, with a 15-h day length. Plants were watered daily and fertilized at weekly intervals. Leaves were harvested 60–80 days after planting.

Isolation of cDNAs—Two different cDNA libraries were employed. The first (a gift from Dr. G. T. King, Native Plants Inc., Salt Lake City, UT) was a Ag10 cDNA library prepared by oligo(dT) priming of mRNA isolated from tomato leaves (King et al., 1988). Approximately 2 × 10⁸ plaque-forming units were subjected to hybridization with a 384-base

1 The abbreviations used are: MCCase, β-methylcrotonyl-CoA carboxylase; EBr, ethidium bromide; IPTG, isopropyl-β-D-thiogalactopyranoside; PAGE, polyacrylamide gel electrophoresis; kb, kilobase(s).

* This work was supported in part by Research Grant DCB-90-04409 from the National Science Foundation (to B. J. N. and E. W.) and by the Iowa State University Office of Biotechnology, Ames, IA. This is Journal Paper No. J-15562 of the Iowa Agriculture and Home Economics Experiment Station, Ames, Iowa, Project No. 2913. The costs of publication of this article were defrayed in part by the payment of page charges. This article must therefore be hereby marked "advertisement" in accordance with 18 U.S.C. Section 1754 solely to indicate this fact.

† Present address: Agracetus, 8520 University Green, Middletown, WI 53562.

‡ To whom correspondence should be addressed: Dept. of Biochemistry and Biophysics, Iowa State University, Ames, IA 50011. Tel.: 515-294-9423; Fax: 515-294-0453.
pair tomato cDNA clone (designated TBP) (Hoffman et al., 1987). Seventeen hybridizing clones were detected using hybridization and washing conditions given in Sambrook et al. (1989). These clones could be classified into two groups according to the sizes of the cDNA inserts. A representative clone of the longest cDNA was subcloned into pUC19.

The second cDNA library (a kind gift from Dr. A. Bennett, University of California, Davis) was a pBluescript SK cDNA library, prepared by oligo(dT) priming of mRNA isolated from ripening fruits of tomato. This library was probed with the 5'-most end, EcoRI-RsaI fragment from pTMC-A (designated BX1; see Fig. 1 for the restriction map of the TMC-A cDNA). From 2 x 10^5 recombinant clones, a single clone was detected. The cDNA insert (1.7 kb in length) was subcloned into pBluescript SK, and the resulting plasmid was called pTMC-B.

Isolation of Genomic Clones—A tomato genomic library, in Charon 34 phage, was a gift of Dr. G. T. King, Native Plants Inc., Salt Lake City, UT. Approximately 2 x 10^9 recombinant bacteriophage were used to hybridize with the cDNA clone from pTBP. Two hybridizing clones, αTG1 and αTG12, were isolated.

Isolation and Analysis of DNA—Tomato genomic DNA was prepared from plant tissues as described (Dellaporta et al., 1983). Isolation of plasmid DNA and bacteria phage DNA, manipulation of the isolated DNAs, and transformation of Escherichia coli with plasmids were as described by Sambrook et al. (1989). Isolated DNA was digested with restriction endonucleases, and the resulting DNA fragments were separated by electrophoresis in 0.7%-1% agarose gels, transferred by capillary action (Southern, 1975) to nylon membranes by using 25 μg sodium phosphate buffer (pH 6.5). Southern blots were subjected to hybridization as described by Sambrook et al. (1989). After hybridization, the blots were washed, and the final most stringent wash was with 0.1 x SSC (1 x SSC = 0.15 μ NaCl, and 15 μ sodium citrate), 0.5% (w/v) SDS at 68 °C for 30 min.

DNA restriction fragments, fractionated by electrophoresis in agarose gels, were purified by adsorption to powdered glass particles which was based on the method developed by Vogelstein and Gillespie (1979).

^32P-Labeled DNA probes were generated in vitro by random primer-supported synthesis of DNA with the Klenow fragment of DNA polymerase I, in the presence of 1 x [γ-^32P]ATP (Fenberg and Vogelstein, 1984).

Restriction sites were mapped on DNA fragments from the ladder of partial digestion products and confirmed by complete digestion of the DNA fragment with restriction enzymes, singly and in combination.

Isolation and Analysis of RNA—Total cellular RNA was isolated from tomato leaves and roots as described by Logemann et al. (1987). RNA was fractionated by electrophoresis in 1.2% (w/v) agarose gels containing formaldehyde (Lehrach et al., 1977). The fractionated RNA was transferred to a nylon membrane with 10 x SSC, and the resulting blot was subjected to hybridization with ^32P-labeled DNA probe. The hybridization was done at 65 °C in a solution composed of 5 x SSPE (1 x SSPE is 0.15 μ NaCl, 1 μm EDTA, 10 μm sodium phosphate buffer (pH 7.4), 5 x Denhardt's solution, 0.5% (w/v) SDS, 100 μg of denatured, sheared salmon sperm DNA, 10% (w/v) dextran sulfate, and 5 ng/ml ^32P-labeled DNA probe. After hybridization, the blots were washed three times with 0.1 x SSC and 0.5% SDS at 65 °C.

DNA Sequencing—cDNAs were cloned in pBluescript SK vector in both potential orientations. Unidirectional deletions were generated by the procedure developed by Henikoff (1984). Single-stranded templates of the deletion mutant constructs were prepared as described (Viera and Messing, 1987). DNA sequencing was performed by using the Sequenase II kit purchased from United States Biochemical Corp., which is based on the dideoxy-mediated chain termination procedure developed by Sanger et al. (1977).

Computer Analysis of Nucleotide and Predicted Amino Acid Sequence—All computer-assisted analyses of nucleotide and predicted amino acid sequences were performed by utilizing the sequence analysis software package of the University of Wisconsin Genetics Computer Group (GCG, Madison, WI).

Expression of Proteins in E. coli—E. coli JM101 harboring expression vectors were grown overnight at 37 °C in 10 ml of LB media containing the appropriate antibiotic. The overnight culture was added to 1 liter of fresh LB media supplemented with antibiotic. After 4 h of growth, expression of the fusion protein was induced by the addition of IPTG to a final concentration of 1 μm. Four h after induction with IPTG, cells were collected by centrifugation at 2,000 x g for 10 min and washed twice with 50 μl Tris-HCl (pH 7.0). The resulting E. coli pellet was resuspended in 5 ml of 50 μl Tris-HCl (pH 7.0), 1 μm EDTA, and 2% (w/v) SDS and then incubated at 100 °C in a bath of boiling water for 15 min. The mixture was diluted 10-fold with 50 μl Tris-HCl (pH 7.0), and 1 μm EDTA before passing through a 2.5-ml avidin-agarose affinity column. The column was washed with 250 ml of 50 μl Tris-HCl (pH 7.0), 1 μm EDTA, and 0.2% (w/v) SDS. Biotin-containing polypeptides were eluted from the column with 3 ml of 100 μl Tris-HCl (pH 7.0), 1 μm EDTA, and 2% (w/v) SDS buffer, and the column was sealed at bottom and heated to 100 °C in a bath of boiling water for 20 min. The eluate was collected from the column by centrifugation at 4,000 x g for 10 min.

Production of Polyclonal Antibodies—Avidin antibody purified, expressed fusion proteins were fractionated by SDS-PAGE. After the gels were stained with Coomassie Brilliant Blue, protein bands were excised from the gel. Gel slices, which contained about 200 μg of protein, were mixed with Freund's complete adjuvant and macerated and emulsified by multiple passages through a syringe. The emulsified antigens were injected at 24–36 sites distributed over the back of New Zealand White female rabbits. Four weeks later, the animals were challenged at 2-week intervals by intramuscular injection into each of the two hind limbs of the animals with freshly prepared antigen emulsified in Freund's incomplete adjuvant. Blood was collected from such rabbits and allowed to coagulate. After centrifugation at 4,000 x g for 10 min, serum was collected, adjusted to 0.02% (w/v) NaN₃, and stored at −20 °C.

![Fig. 1. Restriction endonuclease maps and the strategies used to determine the nucleotide sequences of cDNAs TMC-B and TMC-A. Each DNA fragment was sequenced to the extent indicated by each of the arrows in the direction shown. The cDNA TMC-A was delineated into four parts by three RsaI restriction sites. The three 5'-most portions were subcloned into the SmaI site of pUC19 to generate pBX1, pBX2, and pBX3. Double-stranded DNA templates from these three plasmids were sequenced with the M13 universal primer and pUC reverse primer (arrows beginning at closed circles). In addition, all cDNAs were cloned in the two possible orientations into pBluescript SK, and a series of nested deletions of these plasmids was generated with exonuclease III and S1 nuclease. Single-stranded templates from these deleted plasmids were sequenced with the M13 universal primer and pUC reverse primer (arrows beginning at open circles). Restriction enzymes marked by an asterisk are sites obtained from cloning vectors or linkers. Identical sequences in each cDNA are indicated with identical shading patterns.](image-url)
Fig. 2. Nucleotide sequence of cDNA clones TMC-B, TMC-A, and TBP and the deduced amino acid sequence. The predicted amino acid sequence (single-letter code) is shown below the nucleotide sequence. Nucleotides are numbered in the 5′-3′ direction beginning with the first nucleotide at the 5′-end of TMC-B. The BTO sequence begins at position 1047; the TMC sequence begins at position 717; the TBP sequence begins at position 1567. The unspliced intron sequence is present only TMC-A and is shown in lowercase. The TMC-B cDNA does not contain this intron sequence; in this sequence, nucleotide 966 is juxtaposed to nucleotide 1535. The exon/intron consensus sequences are underlined, and the splicing sites are marked with dots underneath the sequences. The Real sites that divide the TMC-A sequence into four fragments, three of which were subcloned as BX1 (nucleotides 717–1038), BX2 (nucleotides 1039–1669), and BX3 (nucleotides 1670–1947), are indicated. The stop codon is labeled End. The putative polyadenylation signal sequence is boxed. The tetrapeptide Ala-Met-Lys-Met, the conserved biotinylation site with a dashed line.

SDS-PAGE and Western Blot Analysis—SDS-PAGE was carried out as described by Laemmli et al. (1970). Proteins were transferred from the gel to nitrocellulose filters with a semidry transfer apparatus (Kyhse-Andersen, 1984). Immunological detection of proteins was conducted with antisem diluted between 1:200 to 1:1000 in a 3% bovine serum albumin solution, after which, antigen-antibody complexes were detected with 125I-Protein A. Biotin-containing polypeptides were detected with 125I-streptavidin (Nikolau et al., 1985).

Purification of Biotin-containing Polypeptides from Plants—Expanding tomato leaves were harvested and frozen with liquid nitrogen, and while still frozen, the tissue was ground to a powder with a pestle, the mixture was transferred to a glass tube and heated with a boiling water for 15 min. The biotin-containing polypeptides were isolated with avidin-agarose affinity chromatography by the procedure described in the isolation of the expressed proteins from tomato extracts.

Enzyme Extraction—Expanding tomato leaves were harvested, frozen with liquid nitrogen, and pulverized to a powder in a mortar and pestle. Two volumes of SDS-containing extracting buffer (100 mm Tris-Cl (pH 7.0), 10 mm dithiothreitol, 5 mm EDTA, and 2% (w/v) SDS) was added to the powder. After additional homogenization with the pestle, the mixture was transferred to a glass tube and heated at 100°C in a bath of boiling water for 15 min. The biotin-containing polypeptides were isolated with avidin-agarose affinity chromatography by the procedure described in the isolation of the expressed proteins from tomato extracts.

Enzyme Assays—Biotin-dependent carboxylase activities were measured as the rate of incorporation of radioactivity from NaH14CO3 into the acid-stable products (Wurtz and Nikolau, 1990).

RESULTS

Biotin is attached to biotin enzymes via an amide bond between the carboxyl group of biotin and the ε-amino group of a lysine residue (Moss and Lane, 1971). The structural features that identify the lysine residue for biotinylation seem to have been highly conserved during evolution and are now being elucidated (Samols et al., 1987). The primary sequence flanking the lysine residue targeted for biotinylation is fairly well conserved in all biotin enzymes; that sequence is (Ala/Val)-Met-Lys-(Met/Ala). However, this sequence is not the only structural information that defines the biotinylation site because mutations away from this sequence motif affect biotinylation of this lysine residue (Murtil and Samols, 1987).

A consequence of the high degree of conservation at the site of biotinylation is that when a DNA sequence coding for a biotin enzyme is expressed in a heterologous host, the protein product can be biotinylated (Hoffman et al., 1987; Murtil and Samols, 1987; Cronan, 1990; Collins et al., 1987). Indeed, such biotinylation events have been reported as artifacts of immunological screening of expression cDNA libraries utilizing the "biotinavidin" system for detection (Hoffman et al., 1987). One example of these resulted in the isolation of a 384-base pair cDNA clone (Hoffman et al., 1987) that we call TBP (tomato biotin protein). This clone contains an incomplete translation open-reading frame of 70 amino acids that shows sequence similarity to regions of biotin enzymes, including the sequence Ala-Met-Lys-Met. The most striking sequence similarity is to the transcarboxylase of Propionibacterium.

Cloning and Sequencing cDNAs Coding for a Putative Biotin Enzyme from Tomato—The cDNA clone TBP was used as a probe to screen a tomato leaf cDNA library. This resulted in the isolation of several cDNA clones, the longest of which was approximately 1.5 kb (named TMC-A). In addition, the 5′-most end of the TMC-A cDNA (the fragment we call BX1) was used as a probe to screen a ripening tomato fruit cDNA library, resulting in the isolation of a 1.7-kb cDNA (named TMC-B). These cDNAs were subcloned into pBluescript SK, restriction maps prepared for each, and their nucleotide sequences determined by the strategies outlined in Fig. 1.

One end of both cDNAs contained a stretch of adenosine nucleotides, directly upstream of which was the TBP sequence (nucleotides 1687–2095 in Fig. 2). Thus, these ends of the cDNAs were identified as the 5′ end of the corresponding mRNA. The conserved eukaryotic polyadenylation signal 5′-AATAAA-3′ (Proudfoot and Brownlee, 1976) is not present in the 260 nucleotides of the 5′-noncoding sequence upstream of the poly(A) tail, instead, the sequence 5′-ACTCAAAA-3′, a proposed polyadenylation signal sequence found in the soybean
proteinase inhibitor cDNA (Hammond et al., 1984) is located 40 nucleotides upstream of the poly(A) tail. It has been reported that plant polyadenylation signals may diverge significantly from the animal consensus sequence (Hunt et al., 1987).

The TMC-B cDNA extends the translational reading frame of TMC-A between positions 717 and 1034. This indicates the carboxyl terminus of the protein. The biotin carboxylase domain of TMC-B is contained within the biotin carboxyl domain of the protein. The biotin carboxylase domain of TMC-B is contained within the biotin carboxylase domain of the protein (Muramatsu and Mizuno, 1984). For each of these sequences, residues are numbered as underlined in the cited reference.
Biomedical Research Foundation protein database indicates that the amino acid sequence predicted from the nucleotide sequence of TMC-B shows a high degree of similarity to many previously sequenced biotin enzymes (Fig. 3). These include propionyl-CoA carboxylases (Brower et al., 1989; Kraus et al., 1986; Lamhonwah et al., 1987, 1989), acetyl-CoA carboxylases (Lopez-Cassillas et al., 1988; Takai et al., 1987, 1988; Al-Fell et al., 1992), pyruvate carboxylases (Lim et al., 1988; Freytag and Collier, 1984; Zhang et al., 1993), transcarboxylase (Muritif et al., 1985), E. coli biotin carboxyl carrier protein of acetyl-CoA carboxylase (Muramatsu and Mizuno, 1988), and the biotin-containing subunits of oxaloacetate decarboxylases (Schwariz and Oesterhelt, 1985; Schwariz et al., 1988; Laussermair et al., 1989; Woelkh et al., 1992). The amino-terminal sequence predicted from TMC-B (TMC-B, residues 1-193, Fig. 3) also shows homology to the E. coli biotin carboxylase (Li and Cronan, 1992a, 1992b; Kondo et al., 1991). Fig. 3 shows the sequence alignment of rat acetyl-CoA carboxylase, human propionyl-CoA carboxylase, yeast pyruvate carboxylase, biotin carboxylase of E. coli, biotin carboxyl carrier protein of E. coli, and the Propionibacterium 1.3 S subunit of transcarboxylase with the amino acid sequence predicted from TMC-B cDNA clone.

The homology between the TMC-B sequence and the sequences of other biotin enzymes confirm the initial identification of the TBP cDNA (and thus of TMC-B) as coding for a plant biotin-containing enzyme. However, because the similarity between TMC-B and other known biotin enzymes was not specifically higher for a particular biotin enzyme, we were unable to use the sequence data to predict the enzymatic function of the protein coded by TMC-B.

Expression of TMC-B in E. coli—To generate antibodies to the protein coded by the TMC-B cDNA, we expressed this sequence in E. coli. Two chimeric protein expression vectors were constructed. The TMC-B cDNA sequence from nucleotide 647 to 2095 was fused in-frame to the 3'-end of the glutathione S-transferase gene in the vector pGEX-2T (Smith and Johnson, 1989) (Fig. 4A). The resulting plasmid, pGSBTN was introduced into E. coli strains harboring pGSBTN identified the 57-kDa glutathione S-transferase fusion protein (Fig. 4B). Western analysis of such gels with 125I-streptavidin revealed that the GS-BTN fusion protein was biotinylated, as expected (Fig. 4C).

The COOH-terminal portion of TMC-B (the BX3 fragment, Fig. 2) was expressed as a β-galactosidase fusion protein using the pUR vectors (Rüther and Muller-Hill, 1983) (Fig. 5A). The chimeric protein expressed from this vector (GAL-BX3) was identified by SDS-PAGE (Fig. 5B) and Western analysis (Fig. 5, C and D). Western analysis with antibodies to β-galactosidase identified the overexpressed fusion protein as a protein band of about 10 kDa larger than β-galactosidase (cf., lanes 2 and 3 of Fig. 5C). The smaller polypeptide detected by the anti-β-galactosidase antibody is either a proteolytic degradation product or the result of premature termination during translation. Western analysis with 125I-streptavidin demonstrated that the β-galactosidase fusion protein was biotinylated (Fig. 5D).
**Fig. 5. Expression of the 3' end of the TMC-B cDNA in E. coli as β-galactosidase fusion protein.** An EcoRI-XbaI DNA fragment, from the 3' end of the TMC-B cDNA (nucleotides 1670-1914 in Fig. 2) was fused, in-frame to the 3' end of the β-galactosidase gene in the plasmid pUR289. Panel A is a schematic of the final construct of the β-galactosidase-BX3 chimeric gene that we called GAL-BX3. The nucleotide and predicted amino acid sequences at the junction between the β-galactosidase gene and BX3 is boxed. P$_\alpha$ is the E. coli lac promoter. The resulting expression vector pGAL-BX3 was introduced into E. coli, and expression was induced with IPTG. Aliquots of protein extracts were fractionated in triplicate by SDS-PAGE. The resulting gels were either stained with Coomassie Brilliant Blue (panel A) or subjected to Western blot analyses. One blot was sequentially incubated with mouse anti-β-galactosidase, anti-mouse IgG, and $^{125}$I-Protein A to identify β-galactosidase and β-galactosidase fusion proteins (panel B). The other blot was incubated with $^{125}$I-streptavidin to identify biotin-containing polypeptides (panel C). Proteins were extracted from the host E. coli strain, JM105, harboring no plasmid (lane 1), harboring the parent plasmid pUR289 (lane 2), and harboring the recombinant plasmid pGAL-BX3 (lane 3). The expressed β-galactosidase-BX3 chimeric protein was purified by affinity chromatography with immobilized avidin. The purified protein was analyzed by SDS-PAGE and stained with Coomassie Brilliant Blue (panel D).

As expected from the original isolation of the TBP cDNA, the expressed chimeric proteins were biotinylated in E. coli; therefore, we were able to purify both expressed fusion proteins to homogeneity by affinity chromatography with immobilized avidin (Fig. 4D shows the purified glutathione $S$-transferase fusion protein, and Fig. 5E shows the purified β-galactosidase fusion protein). Both purified chimeric proteins were further purified by preparative SDS-PAGE and were used to immunize rabbits for the generation of antisera.

**TMC-B Codes for the Biotin-containing Subunit of β-Methylcrotonyl-CoA Carboxylase**—The antisera that were generated to the GS-BTN and GAL-BX3 fusion proteins were utilized to confirm that the cDNA TMC-B codes for a biotin-containing protein in tomato and, subsequently, to identify the enzymatic function of that biotin protein. Young tomato leaves were harvested as the source for the purification of biotin-containing polypeptides. Proteins were extracted with a buffer containing SDS, and the extract was denatured by heating. Identical aliquots of this extract were subjected to affinity chromatography with avidin-agarose columns that were either pretreated with biotin to block biotin-binding sites or not so pretreated. After extensive washing of the columns to remove nonbinding proteins, the matrices were incubated at 100 °C in the presence of 2% (w/v) SDS. Such harsh conditions were required to elute biotin-containing polypeptides from the column owing to the high affinity between avidin and biotin (the dissociation constant is $10^{-15}$ M$^{-1}$; Green, 1975).

The proteins eluted by this treatment and an aliquot of the initial extract were subjected to SDS-PAGE and Western analysis (Fig. 6). Incubation of the blot with the antiserum against the glutathione $S$-transferase fusion protein (anti-GS-BTN) followed by $^{125}$I-Protein A revealed a 78-kDa polypeptide in the crude extract (Fig. 6, lane 1). This polypeptide was retained and purified by the nonbiotin-treated avidin affinity column (Fig. 6, lane 2) but was not bound by the avidin affinity column that had been preincubated with biotin. The same result was obtained with the antiserum against the β-galactosidase fusion protein (anti-GAL-BX3; data not shown). Therefore, independent of the glutathione $S$-transferase or β-galactosidase epitopes that each of these antibodies recognize, these antibodies identify the identical 78-kDa polypeptide, which, because of its tight binding to avidin, we conclude, contains biotin. These results demonstrate that the TMC-B cDNA codes for a 78-kDa biotin-containing polypeptide.

The following four biotin-dependent enzyme activities are known in the plant kingdom: acetyl-CoA carboxylase, propionyl-CoA carboxylase, MCCase, and pyruvate carboxylase (Wurtele and Nikolau, 1990). To determine which of these, are associated with the 78-kDa biotin-containing polypeptide, the anti-GS-BTN serum was used in experiments such as that illustrated in Fig. 7. Tomato leaf enzyme extracts were incubated on ice for 40 min with increasing amounts of either the antiserum or the preimmune serum. The resulting mixtures were subsequently assayed for acetyl-CoA carboxylase, propionyl-CoA carboxylase, MCCase, and pyruvate carboxylase activities. As shown in Fig. 7, the four carboxylase activities were not affected by increasing amounts of preimmune serum, and acetyl-CoA carboxylase, propionyl-CoA carboxylase, and pyruvate carboxylase activities were similarly unaffected by the introduction of the antiserum. MCCase activity, however, was...
Identification of the protein coded by the TMC-B cDNA.

A denatured protein extract (lane 1) from tomato leaves was subjected to affinity chromatography with immobilized avidin (lane 2) or with immobilized avidin pretreated with 2 mM biotin (lane 3). After extensive washing of the columns, biotin-containing polypeptides were eluted with boiling 2% SDS and collected. These three fractions were subjected to SDS-PAGE and Western blot analysis. The blot was sequentially incubated with anti-GS-BTN serum and \(^{125}\)I-Protein A.

gradually inhibited by the addition of increasing amounts of antiserum. Incubation of crude extracts with the anti-GAL-BX3 serum also specifically inhibited MCCase activity (data not shown). These results indicate that the 78-kDa biotin-containing polypeptide is the biotin-containing subunit of MCCase. The size of this subunit of MCCase from tomato is similar to the biotin-containing subunits of MCCases purified from animals (Lau et al., 1979), bacteria (Fall and Hector, 1977; Apitz-Castro et al., 1970), and other plants (Chen et al., 1993; Alban et al., 1993).

Northern Analysis of MCCase mRNA—To determine the size of the mRNA coding for the 78-kDa subunit of MCCase, RNA isolated from the expanding leaves of tomato plants was fractionated by electrophoresis in a formaldehyde-containing agarose gel and then transferred to a nitrocellulose filter. The blot was hybridized with the TMC-B cDNA (Fig. 8). A single RNA band of ~2.4 kb was detected; no additional bands were detected even upon long exposures of the blot. The size of this transcript (2.4 kb) is sufficient to code for a 78-kDa polypeptide.

MCCase Gene Organization in Tomato Genome—A tomato genomic library was screened by using the TBP cDNA sequence. This resulted in the isolation of two clones, AGT1 and AGT12 (Fig. 9). These clones were characterized in terms of their restriction maps and by Southern blot hybridization to position and orient the TMC-B sequence. That these two clones had different restriction maps indicates that they represent two different genes. To determine the number of genes in the tomato genome that code for the biotin subunit of MCCase, genomic DNA was isolated from tomato leaves and digested to completion with the restriction endonucleases HindIII, EcoRI, and BglII. The Southern blots were probed with the TMC-A cDNA. A simple banding pattern was observed, with two or three genomic restriction fragments hybridizing to the TMC-A cDNA (Fig. 10). This result and the isolation of two distinct genomic clones coding for the biotin subunit of MCCase suggest that the tomato genome contains at least two genes for this protein.

DISCUSSION

We have isolated and characterized genes and cDNAs that were initially identified as coding for a biotin-containing protein of tomato (Hoffman et al., 1987). This initial identification was based on the observation that expression of these cDNAs in E. coli lead to the biotinylation of the protein product; however, the enzymatic function of the protein coded by these sequences was unknown. The characterizations undertaken in this study confirmed that the cloned tomato sequences code for a biotin protein in planta, a polypeptide of 78 kDa. This 78-kDa biotin-containing polypeptide was identified as the biotin subunit of the tomato MCCase.

MCCases were initially characterized from bacterial and ani-
MCCases have been purified to homogeneity (Chen et al., 1993) and more recently several plant MCCases have been isolated from plants. The presence of MCCase in plants was first deduced by the detection of its enzymatic activity in plant extracts (Wurtele and Nikolau, 1990), and more recently several plant MCCases have been purified to homogeneity (Chen et al., 1993; Alban et al., 1993; Dize et al., 1993). These plant MCCases, and the tomato MCCase, have a biotin containing subunit ranging between 74 and 78 kDa. Thus, the bacterial, animal, and plant MCCases have biotin subunits of similar molecular masses.

The identification and sequencing of the tomato cDNAs enabled us, for the first time, to deduce the primary structure of the biotin subunit of MCCase. The amino acid sequence of the biotin subunit of MCCase showed a high degree of homology to other biotin enzymes (see "Results"). In addition, we obtained notable sequence homology between MCCase and lipoamide transferases (Ali and Guest, 1990) and carbamoyl-phosphate synthetases (Nyuniyaya and Lusty, 1983). The lipoic acid carrier domain of lipoamide transferases has a sequence that is similar to the sequence of the biotin carboxyl carrier domain of MCCase (33% identical over 62 residues). As pointed out previously (Lim et al., 1988), this sequence similarity may be an indication of the homologous functions of the lipoyl acid-binding domains and the biotin carboxyl carrier domains in that both cofactors act as intermediate carriers of substrates between physically separate active sites.

The carbamoyl-phosphate synthetase sequences show some similarity to the biotin carboxylase domain of MCCase (over 194 residues, 17% are identical, and 36% are similar if conservative substitutions are allowed). This sequence similarity may be an indication that the reaction catalyzed by carbamoyl-phosphate synthetase may have mechanistic similarity to the biotin carboxylase reaction. Indeed, it is proposed that both reactions occur via the formation of a carboxy-phosphate intermediate, which for biotin enyzmes reacts with the N-1' of biotin (Knowles, 1989), and for carbamoyl-phosphate synthetase reacts with ammonia (Post et al., 1990).

The initial isolation of the MCCase cDNA clone was the result of the biotinylation of its protein product in E. coli (Hoffman et al., 1987). The finding that this sequence codes for a

![Identification of the mRNA coding for the biotin-containing subunit of MCCase.](image1)

![Identification of genes coding for the biotin subunit of MCCase by Southern blot analysis of tomato genomic DNA.](image2)
b-Methylcrotonyl-CoA Carboxylase

Lopez-Cassilas et al., (1988) Genbaufue and Cooper, 1991). Apart from this conserved tetrapeptide, the amino acid sequence surrounding the biotin is variable among different biotin-containing enzymes. Indeed, MCCase from tomato shares only 25% identity over an 87-amino-acid overlap with the E. coli BCCP tetrapeptide residues 350-356 (Fig. 3). However, the secondary structure surrounding the biotinylation site of biotin-containing enzymes may indeed be highly conserved. The secondary structure predicted by the method of Chou and Fasman (1978) indicates that the lysine residue for biotinylation by the holocarboxylase synthetase is in a-helix structure are not completely clear. Although the primary sequence immediately surrounding the lysine residue that is biotinylated is conserved (Ala/Val-Met-Lys-Met-Ala), this tetrapeptide does not on its own define a biotinylation site. The abbreviations used are as in Fig. 1.

REFERENCES

FIG. 11. Comparison of the predicted secondary structures surrounding the biotinylating sites of biotin enzymes. Protein secondary structures were predicted by the PEPTIDESTRUTURE program of the GCG software package. The predicted secondary structures are indicated at the top of the sequences. The sequences predicted to be in a-helix structure are underlined with a solid line. The sequences predicted to be in b-sheet structure are underlined with a dashed line. The conserved hydrophobic amino acid residues are marked with dots at the bottom of the sequences. MK is the conserved dipeptide found in all biotin enzymes, where the lysine residue is the biotinylation site. The abbreviations used are as in Fig. 3.