THE JOURNAL OF BIOLOGICAL CHEMISTRY
Vol. 279, No. 38, Issue of September 17, pp. 39767–39774, 2004

Printed in U.S.A.

The Gene Controlling Marijuana Psychoactivity

MOLECULAR CLONING AND HETEROLOGOUS EXPRESSION OF Δ^1-TETRAHYDROCANNABINOLIC ACID SYNTHASE FROM *CANNABIS SATIVA* L.*

Received for publication, April 2, 2004, and in revised form, June 4, 2004
Published, JBC Papers in Press, June 9, 2004, DOI 10.1074/jbc.M403693200

Supaart Sirikantaramas, Satoshi Morimoto, Yoshinari Shoyama, Yu Ishikawa, Yoshiko Wada, Yukihiro Shoyama, and Futoshi Taura‡

From the Graduate School of Pharmaceutical Sciences, Kyushu University, Fukuoka, 812-8582, Japan

Δ^1-Tetrahydrocannabinolic acid (THCA) synthase is the enzyme that catalyzes oxidative cyclization of cannabigerolic acid into THCA, the precursor of Δ^1-tetrahydrocannabinol. We cloned a novel cDNA (GenBank™ accession number AB057805) encoding THCA synthase by reverse transcription and polymerase chain reactions from rapidly expanding leaves of *Cannabis sativa*. This gene consists of a 1635-nucleotide open reading frame, encoding a 545-amino acid polypeptide of which the first 28 amino acid residues constitute the signal peptide. The predicted molecular weight of the 517-amino acid mature polypeptide is 58,597 Da. Interestingly, the deduced amino acid sequence exhibited high homology to berberine bridge enzyme from *Eschscholtzia californica*, which is involved in alkaloid biosynthesis. The liquid culture of transgenic tobacco hairy roots harboring the cDNA produced THCA upon feeding of cannabigerolic acid, demonstrating unequivocally that this gene encodes an active THCA synthase. Overexpression of the recombinant THCA synthase was achieved using a baculovirus-insect expression system. The purified recombinant enzyme contained covalently attached FAD cofactor at a molar ratio of FAD to protein of 1:1. The mutant enzyme constructed by changing His-114 of the wild-type enzyme to Ala-114 exhibited neither absorption characteristics of flavoproteins nor THCA synthase activity. Thus, we concluded that the FAD binding residue is His-114 and that the THCA synthase reaction is FAD-dependent. This is the first report on molecular characterization of an enzyme specific to cannabinoid biosynthesis.

Cannabinoids, which are found only in *Cannabis sativa*, are novel secondary metabolites consisting of alkylresorcinol and monoterpene groups. More than 60 cannabinoids have been isolated from marijuana or fresh *Cannabis* leaves, and their pharmacological properties have been extensively investigated (1). Among them, Δ^1-tetrahydrocannabinol (THC)[1] is the psychoactive component of marijuana (2). In addition, this cannabinoid is shown to exert a variety of therapeutic activities such as the relief of nausea caused by cancer chemotherapy (3) and the suppression of spasticity and pain associated with multiple sclerosis (4). Thus, because THC has attracted a great deal of attention, various chemical, pharmacological, and biosynthetic studies have been conducted on this cannabinoid (5, 6).

THC usually accumulates at a quite low level in the fresh leaves of *C. sativa* and is considered to be derived artificially from the acidic cannabinoid Δ^1-tetrahydrocannabinolic acid (THCA) by non-enzymatic decarboxylation during storage and smoking (7, 8) (Fig. 1). Our previous studies on the biosynthetic pathway of cannabinoids demonstrated that THCA, which had been believed to be formed through isomerization of cannabidiolic acid, was actually biosynthesized from cannabigerolic acid (CBGA) (9, 10) (Fig. 1). Furthermore, we identified an oxidoreductase named THCA synthase, which catalyzes THCA biosynthesis in rapidly expanding leaves of *C. sativa* (Mexican strain) (9). The activity of this enzyme was not detected in a nonpsychoactive *Cannabis* strain (CBDA strain) (10, 11), which contains cannabidiolic acid as a major cannabinoid and only a trace amount of THCA, indicating that THCA synthase is an important enzyme controlling the psychoactivity of *C. sativa*.

We also demonstrated that THCA synthase catalyzes a unique biosynthetic reaction. THCA synthase oxidatively cyclizes the monoterpene moiety of CBGA to form THCA (9). Similar reactions are often found in monoterpene biosynthesis, where biosynthetic enzymes (monoterpene cyclases) cyclize geranyl pyrophosphate into various monoterpenes (12), but these reactions are not accompanied by oxidation, contrary to the THCA synthase reaction. Because of the novel catalytic property, it is of interest to know the primary structure of THCA synthase, although we revealed no structural information on this enzyme except for the N-terminal sequence containing 15 amino acid residues (9). Moreover, we have not unequivocally determined which subclass this oxidoreductase belongs to, because the coenzyme or cofactor required for the oxidocyclization of CBGA was not identified (9). Thus, the mechanism of the THCA synthase reaction has remained largely unclear.

To obtain a more precise knowledge about the structure and reaction mechanism of THCA synthase, in the present study we cloned and characterized a THCA synthase cDNA (*THCAS*). Surprisingly, we found that the primary structure deduced from the nucleotide sequence of *THCAS* exhibited high homology to that of the berberine bridge enzyme (BBE) (13), which is involved in the biosynthesis of benzophenanthridine alkaloids. Based on the biochemical properties of THCA synthase, we established that this enzyme is a flavinylated oxidase that requires molecular oxygen for the oxidation of the substrate. We report here the structural characteristics and reaction mechanism of THCA syn-

* This research was partially supported by Grants-in-aid 14771317 (to F. T.) and 10672108 (to S. M.) from the Ministry of Education, Science, Sports and Culture of Japan. 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 1734 solely to indicate this fact.

The nucleotide sequence(s) reported in this paper has been submitted to the GenBank™/EBI Data Bank with accession number(s) AB057805.

‡ To whom correspondence should be addressed. Tel.: 81-92-642-6582; Fax: 81-92-642-6582; E-mail: taura@phar.kyushu-u.ac.jp.

[1] The abbreviations used are: THC, Δ^1-tetrahydrocannabinol; BBE, berberine bridge enzyme; CBGA, cannabigerolic acid; 6-HDNO, 6-hydroxy-D-nicotine oxidase; HPLC, high performance liquid chromatography; RACE, rapid amplification of cDNA ends; THCA, Δ^1-tetrahydrocannabinolic acid; *THCAS*, Δ^1-tetrahydrocannabinolic acid synthase cDNA.

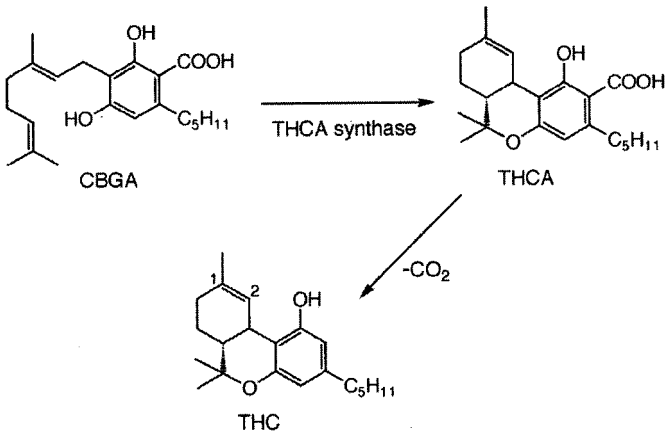

FIG. 1. **Scheme of THC formation.** THCA synthase catalyzes oxidative cyclization of the monoterpene moiety of CBGA to form THCA. THC is derived from THCA by non-enzymatic decarboxylation.

thase. Moreover, we attempted the expression of the active THCA synthase in tobacco hairy roots. Interestingly, the transformed tobacco hairy roots produced THCA on exogenous addition of the biosynthetic precursor CBGA. We also describe the development of THCA-producing tobacco.

EXPERIMENTAL PROCEDURES

Plant Materials and Reagents—C. sativa plants (Mexican strain) (11) were cultivated in the herbal garden of the Graduate School of Pharmaceutical Sciences, Kyushu University. Tobacco (*Nicotiana tabacum* cv. Xanthi) plants were cultured on Gamborg B5 medium (14) solidified with agar at 25 °C under 16 h of illumination. Reagents for molecular biological procedures were purchased from Amersham Biosciences and Toyobo (Osaka, Japan). Chemical and biochemical reagents were purchased from Sigma, Wako Pure Chemicals (Osaka, Japan), and Nacalai Tesque (Kyoto, Japan). THCA was purified from the dried leaves of *C. sativa* as described previously (15). CBGA was chemically synthesized by a two-step procedure, condensation of olivetol and geraniol to synthesize cannabigerol (16) followed by carboxylation of cannabigerol with methyl magnesium carbonate (17).

Purification of THCA Synthase—THCA synthase was purified from rapidly expanding leaves of *C. sativa* (Mexican strain) as reported previously (9). The purity of the enzyme was confirmed by SDS-PAGE analysis (18).

Microsequencing of THCA Synthase—Purified THCA synthase (20 μg) was resolved by SDS-PAGE (12.5% acrylamide gel) and electroblotted onto polyvinylidene difluoride membrane (Millipore). The N-terminal amino acid sequence was analyzed on an Applied Biosystems 473A protein sequencer. Peptide fragments to determine internal sequences were prepared by two different procedures (proteinase Glu-C digestion and cyanogen bromide cleavage). In-gel proteinase digestion (19) was carried out by incubating purified THCA synthase (20 μg) and endoproteinase Glu-C (1 μg) for 30 min in polyacrylamide gels. For cyanogen bromide digestion, purified enzyme (20 μg) was dissolved in 70% (v/v) formic acid containing 1% (w/v) cyanogen bromide and incubated at 30 °C for 10 h in the dark. The resulting peptide fragments were separated by SDS-PAGE (15% acrylamide gels), electroblotted onto polyvinylidene difluoride membranes, and sequenced as described above.

RNA Extraction and Reverse Transcription—Total RNA was extracted from rapidly expanding leaves of *C. sativa* by the acid-guanidinium-phenol-chloroform method (20), and three cDNA pools were prepared. cDNA pools 1 and 2 were synthesized by reverse transcription of the above RNA solution using an oligo-dT primer (5′-$(T)_{17}$-3′) and oligo-dT primer possessing an adapter (5′-GACTCGTCTAGAGGATCCCG$(T)_{17}$-3′), respectively. Both reactions were carried out with Moloney murine leukemia virus reverse transcriptase according to the manufacturer's protocol (Toyobo). cDNA pool 3 was prepared by attaching poly-dA tails to cDNA pool 1 using terminal deoxynucleotidyltransferase according to the manufacturer's protocol (Takara).

Cloning and Sequencing of THCAS—The following oligonucleotide primers were used in this study: degenerate primers a (5′-CGGAATTCTTYCCIGARYTIGGNAT-3′), b (5′-CGGGATCCGTRTACCAIARYTCRTA-3′), and c (5′-GGGGTACCAAYCCIMGIGARAAYTTYYTIAA-3′), designed from peptide sequences of THCA synthase, FPELGI, YELWYT, and NPREAFLK, respectively; gene-specific primers d (5′-CGGGATCCCACCACTGTAGAAGATGGTTG-3′), e (5′-GAAGATGTAGGAGCTGGGATG-3′), f (5′-GGGGTACCGGAGGAGATTTCAGAATCAGC-3′), g (5′-GGTTTTGGGGTTGTATCAGAG-3′), h (5′-GGGGTACCGTCGTGTTGAGTGTATACGAG-3′), i (5′-CGGGATCCATGAATTGCTCAGCATTTTC-3′), and j (5′-CTAGCTGAGCTCTTAATGATGATGCGGTG-3′) and adapter primers k (5′-GACTCGTCTAGAGGATCCCG$(T)_{17}$-3′) and l (5′-GACTCGTCTAGAGGATCCCG-3′). All cDNA fragments were amplified in 50-μl reaction mixtures consisting of primers (degenerate primer, 1 μM; gene-specific or adapter primer, 0.2 μM), 0.2 mM dNTPs, 1.5 mM $MgCl_2$, 50 mM KCl, 10 mM Tris-HCl (pH 8.3), and Taq DNA polymerase (1 unit).

Amplifications of cDNA fragments of THCA synthase were performed as follows. The core cDNA fragment (~350 bp) was obtained by PCR with degenerate primers a and b in the presence of cDNA pool 1 as template (30 cycles of 1 min each at 94, 37, and 72 °C). The cDNA fragment encoding the 5′-upstream region up to the protein N terminus (~1000 bp) was amplified with primers c and d (30 cycles of 1 min each at 94, 50, and 72 °C). The 3′-terminal and 5′-terminal regions were amplified by rapid amplification of cDNA ends (RACE) method (21). 3′-RACE was performed with a subsequent nested PCR. The first round of PCR was conducted with gene-specific primer e, adapter primer l, and cDNA pool 2 as template (30 cycles of 1 min each at 94, 55, and 72 °C). The ~650-bp 3′-RACE product was amplified with gene-specific primer f, adapter primer l, and the first round PCR product as template (30 cycles of 1 min each at 94, 55, and 72 °C). The 5′-RACE product (~200 bp) was obtained as follows. The first round of PCR was performed with gene-specific primer g and adapter primer k in the presence of template cDNA pool 3 (5 cycles of 1 min each at 94, 50, and 72 °C and then 30 cycles of 1 min each at 94, 55, and 72 °C). The second round of PCR (30 cycles of 1 min each at 94, 55, and 72 °C) was performed with gene-specific primer h, adapter primer l, and the PCR product from the first reaction. All PCR products were digested at restriction sites introduced via PCR primers, cloned into the vector pUC119, transformed into competent *Escherichia coli* Top10F′, and sequenced on an Applied Biosystems 310 genetic analyzer. The obtained sequences were combined using the SeqEd program (Applied Biosystems).

Expression of the Recombinant THCA Synthase in Tobacco Hairy Roots—Full-length *THCAS* was amplified using gene-specific primers i and j in the presence of template cDNA pool 1 (30 cycles of 1 min each at 94, 55, and 72 °C) with KOD DNA polymerase according to the manufacturer's instructions (Toyobo). The amplified cDNA was ligated into the vector pUC119 and sequenced as described above to confirm that its nucleotide sequence was identical to that of *THCAS*. The cDNA was excised from pUC119/*THCAS* at BamHI and SacI sites and recloned into the expression vector pBI121 (Clontech). The resulting construct was introduced into *Agrobacterium rhizogenes* strain 15834 (22) by tri-parental mating (23). The tobacco stems were infected directly with *A. rhizogenes* using needles and then cultured for 2 weeks at 25 °C on solidified Gamborg B5 media. Hairy roots generated from the infected sites were cultured on solidified B5 media containing cefotaxime and kanamycin to eliminate bacteria and select transformants. Adventitious shoots generated from hairy roots were cultured on the same media. Total soluble proteins were extracted from roots and shoots by homogenizing samples in buffer A (10 mM sodium phosphate buffer (pH 7.0) containing 3 mM mercaptoethanol). THCA synthase activity in protein samples was measured as described below.

For the bioconversion of CBGA into THCA, ~2-cm tips of the root expressing THCA synthase were inoculated in 30 ml of liquid B5 medium in 100-ml flasks and pre-cultured for 2 weeks at 25 °C with gentle shaking (90 rpm). Then 1 mg of CBGA was added to the medium, and the hairy roots were further cultured under the same conditions. The THCA produced in the hairy roots and culture medium was quantified using HPLC as previously described (9).

Expression of THCA Synthase in Insect Cell Culture—For baculoviral expression *THCAS* was excised from pUC119/*THCAS* at BamHI and SacI sites and recloned into the donor vector pFastBac1 (Invitrogen). Then *THCAS* was transposed into the bacmid DNA within *E. coli* DH10Bac (Invitrogen). The recombinant baculovirus harboring *THCAS* was prepared by transfection of *Spodoptera frugiperda* (Sf9) insect cells with the bacmid using Cellfectin reagent according to the manufacturer's instructions (Invitrogen). For protein production, cells cultured in Sf-900 II medium (Invitrogen) were infected with the recombinant baculovirus. Total cellular proteins were extracted with 50 mM Tris-HCl buffer (pH 8.5) containing 3 mM mercaptoethanol and 1% (v/v) Nonidet P-40. THCA synthase activity in the culture medium, and cell extract was measured as described below.

For purification of the recombinant THCA synthase, a 1000-ml suspension culture of Sf9 cells (~2.0×10^6 cells/ml) was infected with virus at a multiplicity of infection of 1.0. The culture was harvested 72 h after infection and centrifuged at 20,000 × *g* for 15 min. The supernatant was directly applied to a hydroxylapatite column (1.0 × 10.0 cm) equilibrated with buffer A. The column was washed with three column volumes of the same buffer, and bound proteins were eluted with a 600-ml linear gradient of buffer A to 0.5 M potassium phosphate buffer (pH 7.0) containing 3 mM mercaptoethanol. Fractions containing THCA synthase activity were concentrated and dialyzed against buffer A. The dialysate was applied to another hydroxylapatite column (1.0 × 10.0 cm). The column was washed with three column volumes of the same buffer, and THCA synthase was eluted with a 600-ml gradient of NaCl (0–2 M) in buffer A. The most active fractions were concentrated, dialyzed against buffer A, and used for the following biochemical characterization.

Spectroscopic Characterization of THCA Synthase and Its Peptide Fragment—For fluorescence analysis, the recombinant purified THCA synthase was dissolved in 100 mM sodium citrate buffer (pH 4.0) at a concentration of 1 mg/ml. The sample was irradiated at 450 nm, and the fluorescence emission spectrum was measured on a Hitachi Model F-2000 fluorometer. Flavin peptide was prepared and characterized as follows. 100 μg of THCA synthase was precipitated with 5% (v/v) trichloroacetic acid, resuspended in buffer A, and then incubated with trypsin and chymotrypsin (each 1 μg) at 37 °C for 5 h. Insoluble materials were removed by centrifugation after the addition of trichloroacetic acid (final concentration, 5%). The resulting flavin peptide solution was neutralized with 2 M $KHCO_3$ and treated with phosphodiesterase at 37 °C for 30 min. Alternatively, the peptide solution was heated at 110 °C for 10 min. Fluorescence emission spectra of peptide samples before and after phosphodiesterase or heat treatment were analyzed as described above. The release of AMP from flavin peptide was detected enzymatically by using adenylate kinase as described by Barzu and Michelson (24).

For absorption measurements, the recombinant purified THCA synthase was dissolved in buffer A (3.5 mg/ml). The absorption spectrum was measured on a Hitachi Model U-2001 spectrophotometer.

Construction of Mutant cDNAs—The mutant cDNAs were prepared by PCR-based mutagenesis as described by Higuchi *et al.* (25). Briefly, two overlapping PCR fragments were generated incorporating the codon for the new amino acid, and then the second reaction amplified the full-length mutant cDNA. PCRs were conducted with KOD DNA polymerase. This procedure was used in the construction of all site-directed mutants (H92A, R108A, R110A, H114A, and H208A). The mutant oligonucleotides (sense) were: H92A, 5′-AAATAACTCC**GCT**ATCCAAG-3′; R108A, 5′-GCTTGCAGATT**GCA**ACTCGAAG-3′; R110A, 5′-GATTCGAACT**GCA**AGCGGTGGC-3′; H114A, 5′-AAGCGGTGGC**GCT**GATGCTGAG-3′; R208A, 5′-ATTGATGCA**GCC**TTAGTCAA-3′. The prepared mutant cDNAs were introduced into pFastbac1, and sequenced. The mutant enzymes were expressed by the baculovirus expression system as described above.

THCA Synthase Assay—The standard reaction mixture consisted of 200 μM CBGA, 0.1% (w/v) Triton X-100 and 100 mM sodium citrate buffer (pH 5.0) in a total volume of 500 μl. The reaction was started by adding 100 μl of enzyme solution, and the mixture was incubated at 30 °C for 2 h. After termination of the reaction with 600 μl of methanol, a 50-μl aliquot was subjected to analytical HPLC as described previously (9). The K_m and V_{max} values for CBGA were calculated from Lineweaver-Burk double-reciprocal plots of the velocity curves of the reaction with increasing concentrations of CBGA.

To examine the molecular oxygen requirement of THCA synthase, the substrate and enzyme solutions were preincubated with glucose oxidase and catalase in the presence of glucose (26). The hydrogen peroxide generated after the THCA synthase reaction was quantified using a horseradish peroxidase reaction with 4-aminoantipyrin as substrate (27).

RESULTS

cDNA Cloning and Structural Characteristics of THCA Synthase—The cloning of *THCAS* was carried out by reverse transcription and polymerase chain reactions. To design degenerate PCR primers, we first analyzed the partial amino acid sequences of the purified enzyme. Along with the N-terminal sequence containing 15 amino acid residues as previously revealed (9), we determined a further 10 amino acid residues by optimizing the sequencing conditions (Table I). Concerning the internal sequences, the purified enzyme was treated with cyanogen bromide or endoproteinase Glu-C, and the resulting fragments (CNBr-1, CNBr-2, and Glu-C-1) were N-terminally sequenced (Table I).

TABLE I
N-terminal and internal amino acid sequences of THCA synthase
Amino acid sequencing of THCA synthase and its peptide fragments was carried out as described under "Experimental Procedures." The letter *X* indicates that the identity of the residue was ambiguous.

Protein/Peptide	Sequence
N terminus	NPRENFLK*X*FSKHIPNNVANPKLV
CNBr-1	*X*KSFPELGIKK
CNBr-2	YELWYTASWEKQEDN
Glu-C-1	I*X*E*X*AIPFPHRA

When PCR was carried out using the degenerate primers designed from CNBr-1 and CNBr-2, a ~350-bp PCR product was obtained. We next amplified a cDNA fragment (~1000 bp) using respective degenerate and gene-specific primers from the N-terminal sequence and the above PCR product. Furthermore, the amplification of cDNA fragments containing 3′- and 5′-end regions was achieved by 3′- and 5′-RACE respectively, and finally, the full-length *THCAS* was obtained using gene-specific primers from 3′- and 5′-RACE products. *THCAS* is the first gene involved in cannabinoid biosynthesis to be cloned.

THCAS consisted of a 1635-nucleotide open reading frame encoding 545 amino acid residues (Fig. 2). The deduced primary structure contained all the amino acid sequences determined by N-terminal sequencing of the purified enzyme or its digestion products (CNBr-1, CNBr-2, and Glu-C-1). PSORT analysis of the deduced amino acid sequence indicated that the region containing 28 amino acid residues starting from the first methionine is a cleavable signal peptide, consistent with the result obtained by the N-terminal sequencing of the intact enzyme (Fig. 2). Hence, we concluded that mature THCA synthase consists of 517 amino acid residues. However, the theoretical molecular mass (58,597 Da) calculated from the amino acid composition of the deduced mature protein is apparently lower than the value (~75 kDa) obtained by SDS-polyacrylamide gel electrophoresis of the enzyme purified from *C. sativa* (9). We assumed that in plants THCA synthase undergoes post-translational modifications such as glycosylation, since eight possible Asn glycosylation sites were confirmed to be present in mature THCA synthase (Fig. 2).

An important finding regarding the structure of THCA synthase was obtained by comparing the primary structure with those of other proteins using the FASTA program. Surprisingly, THCA synthase catalyzing oxidocyclization of the geranyl moiety in CBGA had no similarity to the monoterpene cyclases, which cyclize geranyl pyrophosphate (28–34). In contrast, THCA synthase showed high homology (40.2% identity in a 535-amino acid overlap) to BBE in *Eschscholzia californica* (13) and ~49% overall identity to hypothetical proteins in various plants such as *Vigna unguiculata, Arabidopsis thaliana,* and *Helianthus annuus*. Among these proteins, BBE is a well investigated oxidase that catalyzes the oxidative cyclization of (*S*)-reticulin into (*S*)-scoulerine (35). Detailed biochemical characterization has demonstrated that BBE is a covalently flavinylated enzyme and that the flavin moiety is involved in the removal of electrons from the substrate (36). Interestingly, a bacterial covalently flavinylated enzyme, 6-hydroxy-D-nicotine oxidase (6-HDNO) in *Arthrobacter oxidans* (37, 38), also exhibited considerable homology (24.7% identity in a 466-amino acid overlap) with THCA synthase. These results implied an evolutionary relationship between THCA synthase and these flavoenzymes. In addition, motif analysis suggested that THCA synthase is also a flavinylated enzyme; the

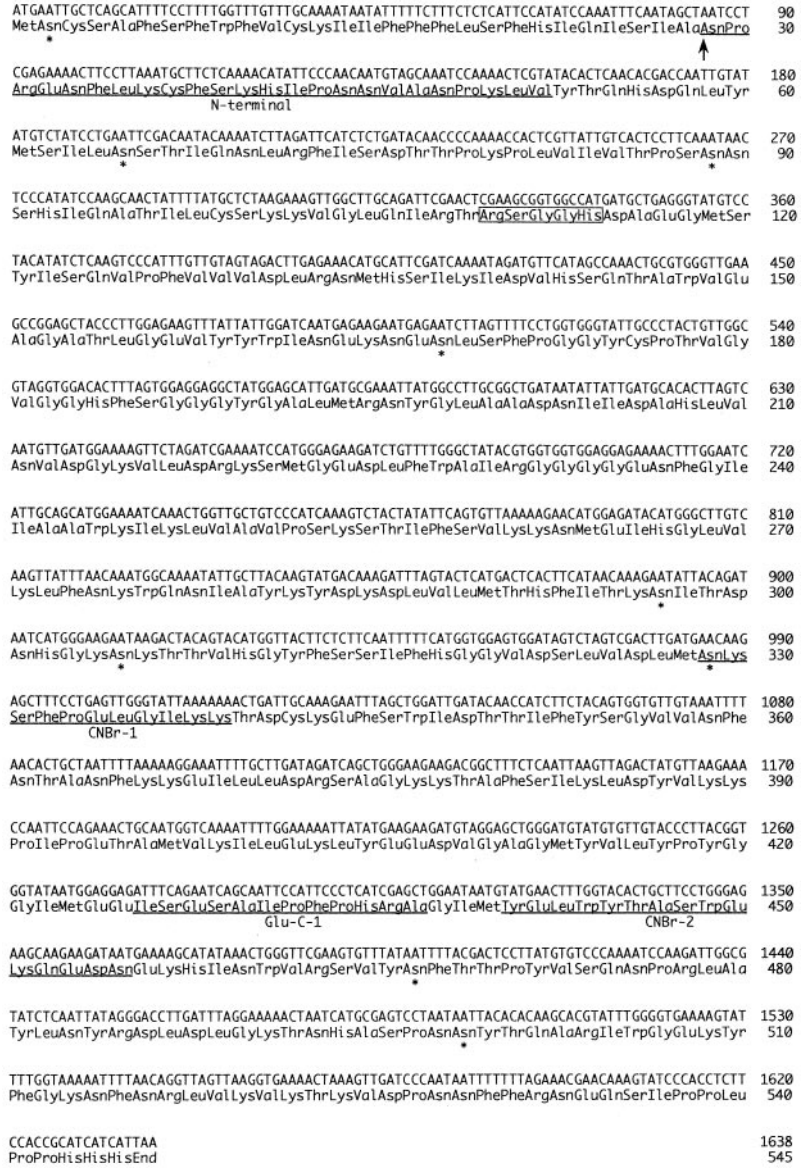

FIG. 2. **Nucleotide and deduced amino acid sequences of THCA synthase.** The amino acid sequences determined from the purified native enzyme are *underlined*. The putative signal peptide cleavage site is indicated by an *arrow*. The *asterisks* indicate potential Asn glycosylation sites. The putative flavin-binding site consensus sequence is *boxed*.

sequence (Arg-Ser-Gly-Gly-His), which is characteristic of flavin-binding sites in flavoproteins such as BBE and 6-HDNO (36, 39), was confirmed in THCA synthase (Fig. 2).

Heterologous Expression of THCA Synthase in Tobacco Hairy Roots—To obtain further evidence that *THCAS* encodes an active THCA synthase, heterologous expression of this gene was attempted. We tried a heterologous plant expression to obtain the active recombinant enzyme because bacterial expression resulted in the accumulation of insoluble protein. For the expression in plants, *THCAS* was cloned into a pBI121 vector having the cauliflower mosaic virus ^{35}S promoter. The resulting construct was introduced into the tobacco (*N. tabacum* cv Xanthi) genome using *A. rhizogenes* (22). The transformants appeared as rapidly growing hairy roots from tobacco stem infected with *Agrobacterium* because of the co-integration of bacterial root-forming genes (22) (Fig. 3*A*). Several adventitious shoots were generated from hairy roots during subculture on selection media. We confirmed that the hairy roots (1.16 picokatal/mg of protein) and the shoots (0.27 picokatal/mg protein) possessed THCA synthase activity. Furthermore, we also investigated whether transformed tobacco can produce THCA. When the transgenic hairy roots were cultured in liquid B5 medium (30 ml) supplemented with 1 mg of CBGA, the maximum level of THCA (82 μg, 8.2% conversion from CBGA) was produced 2 days after the addition of CBGA (Fig. 3*B*). Although the hairy roots did not secrete THCA synthase, about half of the THCA produced was found in the culture medium (Fig. 3*B*), indicating CBGA uptake and THCA release by tobacco cells. These results demonstrated that *THCAS* can control THCA production not only in *C. sativa* but also in other plants such as tobacco.

Identification of Coenzyme Required for THCA Synthase Reaction—Previously, we characterized the native THCA synthase in a pure form but could identify neither a coenzyme nor a cofactor required for this reaction (9). However, in the present study the sequence analysis demonstrated the possibility that THCA synthase possesses a flavin that acts as a coenzyme. Therefore, we next attempted to identify the coenzyme that binds to this enzyme. Because it was impossible to purify enough THCA synthase from *C. sativa* or transformed tobacco cultures, we overexpressed the recombinant enzyme using a baculovirus-insect cell system. After insect cells (Sf9) were infected with baculovirus harboring *THCAS*, the THCA synthase activity in the cell extract and in the culture medium was measured. Consequently, much greater THCA synthase activity (10.5 picokatal/mg protein) was observed in the culture

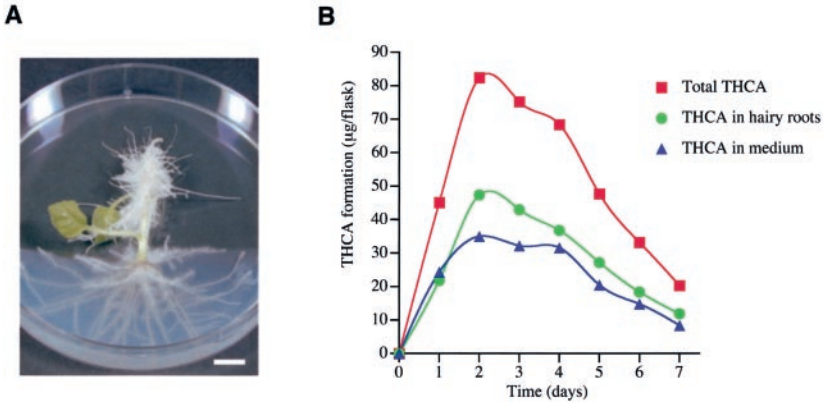

FIG. 3. **Development of THCA-producing tobacco hairy roots.** *A*, induction of hairy roots from tobacco stem infected with *A. rhizogenes* harboring *THCAS*. The *white bar* indicates 1 cm. *B*, production of THCA in suspension cultures of transgenic tobacco hairy roots expressing recombinant THCA synthase. The roots were inoculated in 30 ml of liquid Gamborg B5 medium in 100-ml flasks and pre-cultured for 2 weeks. Then 1 mg of CBGA was added to the culture, and the accumulation of THCA in the aliquot of hairy roots and medium was measured by HPLC. Data are the means of triplicate determinations.

medium than cell extract (0.1 picokatal/mg of protein). Thus, it is demonstrated that most of the recombinant enzyme was secreted outside the cells. The recombinant THCA synthase in the insect culture medium was readily purified to a homogeneous protein with a molecular mass of ~60 kDa by two consecutive hydroxylapatite chromatographies using different buffer systems (Fig. 4). The yield of the pure form was more than 1 mg from a 1-liter culture. The N-terminal amino acid sequence of the recombinant enzyme was identical to that of the native enzyme, indicating that the 28-amino acid signal peptide was correctly cleaved in insect cells. The purified enzyme was used for the biochemical analysis.

The concentrated solution of the recombinant THCA synthase gave a yellow coloration, and transillumination of the enzyme at 366 nm showed autofluorescence on SDS-PAGE. When the absorbance of THCA synthase was measured, absorbance maxima were confirmed at 365 and 450 nm (Fig. 5*A*). Furthermore, the fluorescence emission of THCA synthase exhibited a maximum at 515 nm, and treatment of the enzyme solution with the flavin-reducing reagent, sodium dithionite, resulted in a quenching of fluorescence (Fig. 5*B*). The release of the compound exhibiting autofluorescence from THCA synthase was not observed after trichloroacetic acid treatment. These properties were consistent with those of flavoproteins (40, 41), indicating that THCA synthase has a covalently bound flavin.

The majority of covalently flavinylated enzymes contain FAD, but several enzymes possessing FMN have been identified (42). We next attempted to identify the flavin that attaches to THCA synthase. The flavin-bound peptide was prepared by partial hydrolysis of THCA synthase with trypsin and chymotrypsin. Phosphodiesterase treatment of this peptide resulted in an ~1.5-fold increase of fluorescence emission at 515 nm along with the release of AMP, whereas heating the peptide at 110 °C gave similar results, indicating that FMN was formed by cleavage of the phosphodiester bond in FAD (41). These results showed that FAD covalently binds to the enzyme. Based on the comparison of the molar extinction coefficient of THCA synthase at 450 nm with that of authentic FAD, the molar ratio of FAD to the enzyme was concluded to be 1:1.

Determination of the FAD-binding Site of THCA Synthase—The FAD moiety of flavoproteins such as BBE and 6-HDNO is shown to bind to the His residue in the consensus sequence (Arg-Ser-Gly-Gly-His) (36, 39). In addition, the Arg residue in this sequence is reported to assist binding of FAD to the His residue (36, 43). To identify the FAD-binding site in THCA synthase and to confirm whether FAD is essential for the enzymatic reaction, we conducted site-directed mutagenesis. Because THCA synthase also possesses a consensus sequence (Arg^{110}-Ser-Gly-Gly-His^{114}), we first constructed the mutants by changing the Arg-110 and His-114 residues of the wild-type

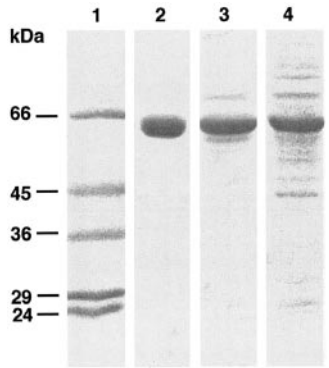

FIG. 4. **SDS-PAGE analysis of the recombinant THCA synthase.** Sf9 cells transfected with baculovirus harboring *THCAS* were cultured for 72 h. The recombinant enzyme was purified from the culture medium by two hydroxylapatite chromatography using different buffer systems as described under "Experimental Procedures." Protein samples at each purification step were resolved by SDS-PAGE (12.5% acrylamide gel) and stained with Coomassie Brilliant Blue R-250. *Lane 1*, molecular mass standards with the indicated molecular masses; *lane 2*, 5 μg of protein from the second hydroxylapatite column eluted with a NaCl gradient; *lane 3*, 5 μg of protein from the first hydroxylapatite column eluted with a phosphate ion gradient; *lane 4*, culture medium containing 10 μg of protein.

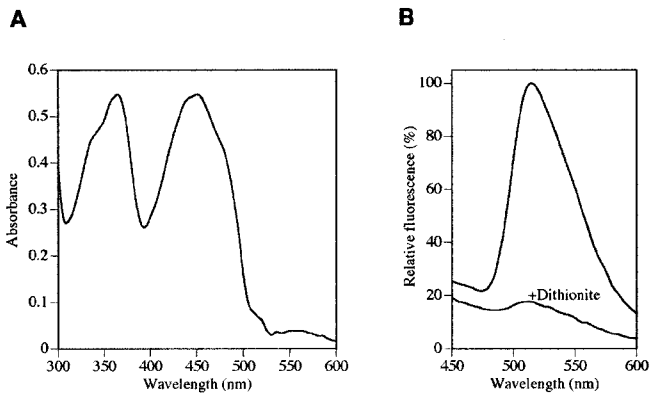

FIG. 5. **Spectral measurements of the purified recombinant THCA synthase.** *A*, the absorption spectrum of the enzyme (3.5 mg/ml) dissolved in 10 mM sodium phosphate buffer (pH 7.0) containing 3 mM mercaptoethanol. *B*, the fluorescence emission spectrum of the enzyme (1 mg/ml) dissolved in 10 mM sodium citrate buffer (pH 4.0). The sample was irradiated at 450 nm. The fluorescence emission was quenched by the addition of sodium dithionite.

enzyme to Ala-110 (R110A) and Ala-114 (H114A), respectively. For comparison, three control mutants (H92A, H208A, and R108A) were also prepared. All mutants were expressed by the baculovirus system, and the enzyme activities in culture media were measured. We confirmed the level of expression of all

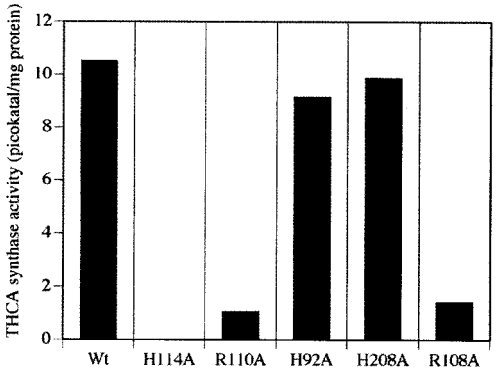

FIG. 6. **THCA synthase activities of wild-type and mutant enzymes.** Sf9 cells transfected with baculovirus harboring wild-type (*Wt*) or mutant cDNAs were cultured for 72 h. The culture media were dialyzed against 10 mM sodium phosphate buffer (pH 7.0) containing 3 mM mercaptoethanol and then used for THCA synthase assay as described under "Experimental Procedures." Data are the means of triplicate assays.

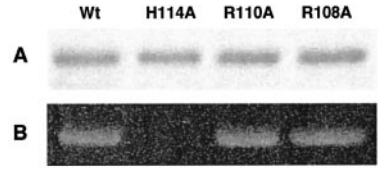

FIG. 7. **Wild-type and mutant enzymes analyzed by SDS-PAGE.** *A*, purified wild-type (*Wt*) and mutant enzymes (each 2.0 μg) stained with Coomassie Brilliant Blue R-250. *B*, the same gel as in *A* with the protein bands visualized by transillumination at 366 nm.

TABLE II
Kinetic parameters of wild-type and mutant enzymes

The K_m values were determined as described under "Experimental Procedures." The k_{cat} values were calculated using a subunit molecular mass of 60 kDa and one active site/subunit. Data are the means of triplicate assays.

Enzyme	K_m	k_{cat}	k_{cat}/K_m
	mM	s^{-1}	$mM^{-1}\ s^{-1}$
Wild type	0.540	0.300	0.556
R108A	0.446	0.043	0.096
R110A	0.783	0.075	0.096

mutant enzymes to be similar by SDS-PAGE analysis (data not shown). As shown in Fig. 6, H114A mutant enzyme exhibited no THCA synthase activity. On the other hand the R110A mutant showed considerably decreased but detectable enzyme activity (10% that of wild type). The THCA synthase activity of control mutants H92A and H208A was similar (87 and 94% activity, respectively) to that of the wild type, whereas interestingly, the R108A mutant showed appreciably reduced activity (13% that of wild type). Thus, we suspected that Arg-108 as well as Arg-110 affected the flavinylation of THCA synthase.

To confirm the presence of a FAD molecule in each mutant (H114A, R110A,and R108A), we purified these mutants and compared their fluorescence properties to that of wild-type enzyme. As shown in Fig. 7, the enzymatically inactive mutant H114A completely lost the in-gel autofluorescence property, indicating that binding of a FAD molecule to this His residue is crucial for the THCA synthase activity. On the other hand, the R108A and R110A mutants showed autofluorescence similar to that of the wild type (Fig. 7). In addition, the UV-visible absorption spectra of both mutants were also almost identical to that of the wild-type enzyme (data not shown). These results indicated that, like the wild-type enzyme, the R108A and R110A mutants, which have lower enzyme activity, are fully flavinylated. The residue corresponding to the Arg-110 in THCA synthase is conserved in other flavoproteins, and this Arg residue is considered to assist flavinylation of the FAD-binding site. However, it is evident that the Arg-110 as well as Arg-108 residues of THCA synthase are not involved in FAD binding.

To investigate the possible roles of these Arg residues, kinetic parameters for CBGA were determined and compared with those of wild-type enzyme. As shown in Table II, R108A and R110A mutants showed K_m values similar to that of the wild type, whereas the k_{cat}/K_m values of both mutants extensively decreased (~17% that of the wild type). Thus, these Arg residues would participate not in substrate binding but in the catalytic process of THCA synthase.

The Mechanism of Oxidation Catalyzed by THCA Synthase—The structural similarity between THCA synthase and BBE suggests that the oxidative cyclization of the substrate by THCA synthase proceeds through a mechanism similar to the BBE reaction. However, although BBE absolutely requires molecular oxygen (35), which acts as an electron acceptor, we previously reported that molecular oxygen is not essential for the THCA synthase reaction (9). In the previous study, molecular oxygen was removed by N_2 purge of the assay solution, but we suspected that the assay solution still contained enough molecular oxygen to activate the enzyme even after the purge. Therefore, we re-examined the effect of molecular oxygen on THCA synthase reaction using a glucose-glucose oxidase-catalase system, which completely consumes the molecular oxygen in solution (26). Consequently, the THCA synthase reaction was completely inhibited by this treatment, confirming that the reaction absolutely depends on molecular oxygen.

In addition, BBE was shown to produce hydrogen peroxide as well as (*S*)-scoulerine during oxidation of (*S*)-reticulin in the BBE reaction (35). When hydrogen peroxide in the assay solution was analyzed by a coloration method using 4-aminoantipyrine and peroxidase (27), we confirmed that the enzyme reaction produces hydrogen peroxide and THCA at a molar ratio of 1:1. Thus, we concluded that THCA synthase oxidizes CBGA via a mechanism similar to that of BBE and that the reaction can be summarized as follows; CBGA + O_2 → THCA + H_2O_2.

DISCUSSION

In the present study, we cloned a cDNA, *THCAS*, encoding THCA synthase by reverse transcription-PCR. *THCAS* consists of a 1635-bp open reading frame encoding a 545-amino acid polypeptide. The deduced amino acid sequence of THCA synthase had highest similarity to BBE from *E. californica* among functionally characterized proteins. Although the substrates are different, both enzymes catalyze the oxidative cyclization of their substrates to form new ring systems. Therefore, it may be reasonable that THCA synthase is structurally related to BBE. The high sequence identity overall between THCA synthase and BBE suggested an evolutionary relationship between these enzymes. It is of interest that homologous enzymes work in apparently distinct secondary pathways such as cannabinoid and alkaloid biosynthetic pathways. FASTA analysis demonstrated that a bacterial enzyme, 6-HDNO, also had considerable homology with THCA synthase. Therefore, oxidoreductases structurally related to THCA synthase may be widely distributed in various organisms.

We expressed enzymatically active recombinant THCA synthase in both insect cell and tobacco hairy root cultures. Interestingly, the recombinant enzyme with the same N-terminal sequence as that of the native THCA synthase accumulated in the insect culture medium, indicating that insect cells could cleave the signal sequence from the precursor to secrete the mature enzyme to the medium. In contrast, tobacco hairy roots

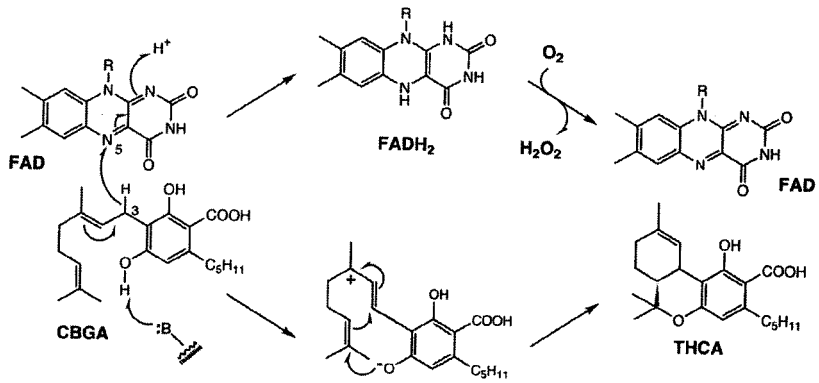

FIG. 8. **The reaction mechanism of THCA synthase.** Two electrons from the substrate are accepted by enzyme-bound FAD and then transferred to molecular oxygen to re-oxidize FAD. THCA is synthesized from the ionic intermediate via stereoselective cyclization by the enzyme. *R*, the rest of FAD molecule. *B*, the proposed basic residue of the enzyme.

harboring *THCAS* did not secrete the THCA synthase activity to the medium. Therefore, it is likely that the sorting mechanism of THCA synthase in plants is different from that in insect cells. In fact, similar results have been reported in studies on other plant enzymes. For example, BBE, which is localized in vacuoles in plant cells (44), is reported to be secreted into the medium when expressed in insect cells (45). We now investigate the subcellular localization of THCA synthase in *C. sativa* cells.

The expression system using the insect cells provided a sufficient amount of THCA synthase to characterize the structural and functional properties of THCA synthase. The most important structural feature is that THCA synthase is a flavoprotein. Site directed mutational analysis demonstrated that the H114A mutant enzyme exhibited neither spectral characteristics of flavoproteins nor THCA synthase activity. Thus, we concluded that the FAD binding residue is His-114, which is in the consensus sequence for flavinylation (Arg^{110}-Ser-Gly-Gly-His^{114}) and that the THCA synthase reaction is FAD-dependent. These results were consistent with those observed for BBE and 6-HDNO (36, 39). In the case of these flavoenzymes, the contributions of amino acid residues near the FAD binding His residue were also studied. For example, the Arg residue in the consensus sequence is shown to be essential for the attachment of FAD to BBE and 6-HDNO (36, 43). Mauch *et al.* (43) propose that interaction between the side chain of the Arg residue and isoalloxazine ring of FAD assists the covalent bond formation between His and FAD (43). Thus, we changed the corresponding Arg residue (Arg-110) of THCA synthase to Ala. However, unexpectedly, R110A mutant enzyme was shown to be flavinylated. Therefore, the flavinylation process for THCA synthase may be different from that for structurally related flavoenzymes. Furthermore, kinetic analysis of mutant enzymes also provided an interesting result; mutations of Arg-108 and Arg-110 considerably reduced the k_{cat} value of the enzyme without a remarkable change of the K_m value. These results suggested that these Arg residues are important for the optimal catalytic action of THCA synthase. To obtain further information on the structural basis for the THCA synthase reaction, we are now planning further mutational studies and a tertiary structural determination of the enzyme.

The THCA synthase reaction was apparently dependent on the attached flavin and absolutely required molecular oxygen and was coupled with a hydrogen peroxide release proportional to THCA formation. Based on these novel findings, we propose a mechanism for the THCA synthase reaction shown in Fig. 8. With this mechanism, the reaction is initiated by the hydride transfer from the C-3 position of CBGA to the reactive N-5 position of the isoalloxazine ring of FAD (46). This step produces a reduced flavin and an ionic intermediate, the configuration of which is competent for cyclization. The next step is the stereospecific and electrophilic ring closure to form THCA, and the hydride ion is transferred from the reduced flavin to molecular oxygen, resulting in hydrogen peroxide formation and re-oxidation of the flavin. The role of Arg-108 and Arg-110 residues remains to be investigated; however, the side chain of these residues might be involved in proton transfer as catalytic base in the reaction. In the proposed reaction mechanism, the cyclization step seemed to be similar to the partial reactions catalyzed by various monoterpene cyclases; *i.e.* cyclization of ionic intermediates derived from geranyl pyrophosphate to form cyclic monoterpenes (12). However, a detailed analysis did not demonstrate homologous regions between THCA synthase and monoterpene cyclases. Accordingly, amino acid residues different from those of monoterpene cyclases would compose the cyclization active site of THCA synthase. Although substrates are different, the reaction mechanism of THCA synthase is quite similar to that catalyzed by BBE (35). Taken together, it is evident that THCA synthase belongs to FAD-dependent oxidases.

In this study we also obtained an interesting result with a heterologous plant expression system for THCA synthase. The liquid culture of the recombinant tobacco hairy roots harboring *THCAS* converted CBGA into THCA. This result provided not only direct evidence for the *in vivo* functionality of *THCAS* but also a potential biotechnological production system for THC, since CBGA is easy to synthesize (16, 17), and THCA is readily decarboxylized to THC by heating (8). Further molecular studies on cannabinoid biosynthesis may develop a biomimetic production system for THCA without the need for feeding precursors. In addition, *THCAS* would also contribute to the artificial control of THCA production in *C. sativa*. For example, overexpression of THCA synthase will produce THCA-rich plants with increased therapeutic potential. Because there has been no report on transgenic *Cannabis* plants, an efficient transformation system for this plant should be established before genetic modifications. However, cloning of *THCAS*, the gene controlling marijuana psychoactivity, is undoubtedly a key step in producing novel *C. sativa* plants suitable for medicinal applications.

Acknowledgments—We thank Dr. Yuji Ito for determining amino acid sequences. We also thank Dr. Kenji Hamase and Mr. Junzo Hirano for helpful advice regarding the fluorescence analysis.

REFERENCES

1. Turner, C. E., Elsohly, M. A., and Boeren, E. G. (1980) *J. Nat. Prod.* (*Lloydia*) **43,** 169–234
2. Gaoni, R., and Mechoulam, R. (1964) *J. Am. Chem. Soc.* **86,** 1946–1947
3. Guzman, M. (2003) *Nat. Rev. Cancer* **3,** 745–755
4. Baker, D., Pryce, G., Giovannoni, G., and Thompson, A. J. (2003) *Lancet Neurol.* **2,** 291–298
5. Mechoulam, R. (1970) *Science* **168,** 1159–1166
6. Mechoulam, R. (2000) *Curr. Pharm. Des.* **6,** 1313–1322
7. Yamauchi, T., Shoyama, Y., Aramaki, H., Azuma, T., and Nishioka, I. (1967) *Chem. Pharm. Bull.* **15,** 1075–1076
8. Kimura, M., and Okamoto, K. (1970) *Experientia* (*Basel*) **26,** 819–820
9. Taura, F., Morimoto, S., Shoyama, Y., and Mechoulam, R. (1995) *J. Am. Chem. Soc.* **117,** 9766–9767
10. Taura, F., Morimoto, S., and Shoyama, Y. (1996) *J. Biol. Chem.* **271,**

17411–17416
11. Shoyama, Y., Yagi, M., Nishioka, I., and Yamauchi, T. (1975) *Phytochemistry* **14,** 2189–2192
12. Croteau, R. (1987) *Chem. Rev.* **87,** 929–954
13. Dittrich, H., and Kutchan, T. M. (1991) *Proc. Natl. Acad. Sci. U. S. A.* **88,** 9969–9973
14. Gamborg, O. L., Miller, R. A., and Ojima, K. (1968) *Exp. Cell Res.* **50,** 151–158
15. Taura, F., Morimoto, S., and Shoyama, Y. (1995) *Phytochemistry* **39,** 457–458
16. Mechoulam, R., and Yagen, Y. (1969) *Tetrahedron Lett.* **10,** 5349–5352
17. Mechoulam, R., and Ben-Zvi, Z. (1969) *J. Chem. Soc. D.* **1969,** 343–344
18. Laemmli, U. K. (1970) *Nature* **227,** 680–685
19. Cleveland, D. W., Fischer, S. G., Kirschner, M. W., and Laemmli, U. K. (1977) *J. Biol. Chem.* **252,** 1102–1106
20. Chomczynski, P., and Sacchi, N. (1987) *Anal. Biochem.* **162,** 156–159
21. Frohman, M. A., Dush, M. K., and Martin, G. R. (1988) *Proc. Natl. Acad. Sci. U. S. A.* **85,** 8998–9002
22. White, F. F., and Nester, E. W. (1980) *J. Bacteriol.* **141,** 1134–1141
23. Van Haute, E., Joos, H., Maes, M., Warren, G., Van Montagu, M., and Schell, J. (1983) *EMBO J.* **2,** 411–417
24. Barzu, O, and Michelson, S. (1983) *FEBS Lett.* **153,** 280–284
25. Higuchi, R., Krummel, B., and Saiki, R. K. (1988) *Nucleic Acids Res.* **16,** 7351–7367
26. Fabian, J. (1965) *J. Bacteriol.* **89,** 921
27. Green, M. J., and Hill, H. A. O. (1984) *Methods Enzymol.* **105,** 3–22
28. Colby, S. M., Alonso, W. R., Katahira, E. J., McGarvey, D. J., and Croteau, R. (1993) *J. Biol. Chem.* **268,** 23016–23024
29. Yuba, A., Yazaki, K., Tabata, M., Honda, G., Croteau, R. (1996) *Arch. Biochem. Biophys.* **332,** 280–287
30. Bohlmann, J., Steele, C. L., and Croteau, R. (1997) *J. Biol. Chem.* **272,** 21784–21792
31. Wise, M. L., Savage, T. J., Katahira, E., and Croteau, R. (1998) *J. Biol. Chem.* **273,** 14891–14899
32. Bohlmann, J., Phillips, M., Ramachandiran, V., Katoh, S., and Croteau, R. (1999) *Arch. Biochem. Biophys.* **15,** 232–243
33. Hoelscher, D. J., Williams, D. C., Wildung, M. R., and Croteau, R. (2003) *Phytochemistry* **62,** 1081–1086
34. Phillips, M. A., Wildung, M. R., Williams, D. C., Hyatt, D. C., and Croteau, R. (2003) *Arch. Biochem. Biophys.* **15,** 267–276
35. Steffens, P., Nagakura, N., and Zenk, M. H. (1985) *Phytochemistry* **24,** 2577–2583
36. Kutchan, T. M., and Dittrich, H. (1995) *J. Biol. Chem.* **270,** 24475–24481
37. Brühmüller, M., Möhler, H, and Decker, K. (1972) *Eur. J. Biochem.* **29,** 143–151
38. Brandsch, R., Hinkkanen, A. E., Mauch, L., Nagursky, H., and Decker, K. (1987) *Eur. J. Biochem.* **167,** 315–320
39. Mauch, L., Bichler, V., and Brandsch, R. (1989) *FEBS Lett.* **257,** 86–88
40. Singer, T. P., and McIntire, W. (1984) *Methods Enzymol.* **106,** 369–378
41. Koziol, J. (1970) *Methods Enzymol.* **18,** 253–285
42. Decker, K. F. (1993) *Annu. Rev. Nutr.* **13,** 17–41
43. Mauch, L., Bichler, V., and Brandsch, R. (1990) *J. Biol. Chem.* **265,** 12761–12762
44. Bird, D. A., and Facchini, P. J. (2001) *Planta* **213,** 888–897
45. Kutchan, T. M., Bock, A., and Dittrich, H. (1994) *Phytochemistry* **35,** 353–356
46. Singer, T. P., and Edmondson, D. E. (1978) *Methods Enzymol.* **53,** 397–418