FEBS 26861 FEBS Letters 534 (2003) 75–81



# Isolation of a new dual functional caffeine synthase gene encoding an enzyme for the conversion of 7-methylxanthine to caffeine from coffee (*Coffea arabica* L.)[1]

Kouichi Mizuno[a,*], Akira Okuda[a], Misako Kato[b], Naho Yoneyama[b], Hiromi Tanaka[a], Hiroshi Ashihara[c], Tatsuhito Fujimura[a]

[a] *Institute of Agricultural and Forest Engineering, University of Tsukuba, Tsukuba Science City, Ibaraki 305-8572, Japan*
[b] *Graduate School of Humanities and Sciences, Ochanomizu University, Otsuka, Bunkyo-ku, Tokyo 112-8610, Japan*
[c] *Department of Biology, Faculty of Science, Ochanomizu University, Otsuka, Bunkyo-ku, Tokyo 112-8610, Japan*

Received 10 September 2002; revised 6 November 2002; accepted 18 November 2002

First published online 11 December 2002

**Abstract** **In coffee and tea plants, caffeine is synthesized from xanthosine via a pathway that includes three methylation steps. We report the isolation of a bifunctional coffee caffeine synthase (CCS1) clone from coffee endosperm by reverse transcription-polymerase chain reaction (RT-PCR) and rapid amplification of cDNA ends (RACE) technique using previously reported sequence information for theobromine synthases (CTSs). The predicted amino acid sequences of CCS1 are more than 80% identical to CTSs and are about 40% similar to those of tea caffeine synthase (TCS1). Interestingly, CCS1 has dual methylation activity like tea TCS1.**


*Key words:* Coffee; Caffeine; *N*-Methyltransferase; 7-Methylxanthine; *S*-adenosyl-L-methionine; Theobromine

## 1. Introduction

Caffeine is a major and important effective component of coffee and is used widely in beverages. It is important to clarify the caffeine biosynthesis pathway and to clone the genes related to the production of caffeine not only to determine the metabolism of the purine alkaloid but also to control the content of caffeine in the coffee plant.

The metabolic pathway of purines, which family caffeine is a member of, has not been elucidated in detail, although we recently proposed major biosynthetic pathways of the purine metabolites in leaves of coffee (*Coffea arabica*) and tea (*Camellia sinensis*) [1–4]. Details of minor pathways for theobromine and caffeine are still obscure especially in coffee plants, for example, Nazario and Lovatt proposed different pathways of caffeine biosynthesis in coffee leaves [5]. Since the caffeine biosynthetic pathway contains three *S*-adenosyl-L-methionine (SAM)-dependent methylation steps, *N*-methyltransferases play important roles. The available data indicate that a xanthosine → 7-methylxanthosine → 7-methylxanthine → theobromine → caffeine pathway is the major route to caffeine [3,4] with alternative entry to xanthosine 5′-monophosphate (XMP) → 7-methyl XMP → 7-methylxanthosine [6] (Fig. 1).

Recently we succeeded in highly purifying CS enzymes [7] and in cloning a caffeine synthase (CS) gene from young tea leaves [8]. We obtained homologous genes that encode caffeine biosynthetic enzymes from young coffee leaves [9] and independently Ogawa et al. also cloned similar genes using the sequence of our tea caffeine synthase (TCS) gene [10]. The latter genes encoded theobromine synthases named CTS1 and CTS2 [9] and CaMXMT by Ogawa et al. [10]. In contrast to tea TCS1 which catalyzes two final methylation steps of caffeine biosynthesis, both from 7-methylxanthine to theobromine and from theobromine to caffeine, CTSs only catalyze the conversion from 7-methylxanthine to theobromine and do not have 1-*N*-methylation activity (Fig. 1). Since the coffee plant generally contains small amounts of theobromine and accumulates high concentrations of caffeine, the coffee plant must have 1-*N*-methyltransferase that converts theobromine to caffeine. After efforts to isolate 1-*N*-methyltransferase genes from coffee, two new genes, tentatively named *CtCS6* and *CtCS7*, were isolated as homologs to CTS1 and CTS2. These clones were successfully characterized as CS cDNAs from young coffee beans and will be described here.

## 2. Materials and methods

### 2.1. Materials

Coffee plants (*C. arabica* L.) were grown in a greenhouse at the Agricultural and Forest Research Center, University of Tsukuba, Japan. Immature seeds and young leaves were harvested and then stored at −80°C until use. [Methyl-$^{14}$C]SAM (55 mCi/mmol) was purchased from Amersham Pharmacia.

### 2.2. Amplification of the coding sequences of coffee tentative caffeine synthase (CtCS) cDNA using rapid amplification of cDNA ends (3′-RACE and 5′-RACE)

Total cellular RNAs from immature coffee endosperm were extracted by cetyltrimethylammonium bromide (CTAB) solution according to the method of Chang et al. [11] with a slight modification. First-strand cDNAs were synthesized using a 3′-RACE core set and *oligo-dT 3 sites adaptor primer* (*oligo-dT 3SAP*) (TaKaRa, Japan) as a

*Corresponding author. Present address: Faculty of Bioresource Sciences, Akita Prefectural University, Akita City, Akita 010-0195, Japan. Fax: (81)-18-872 1677.
*E-mail address:* koumno@akita-pu.ac.jp (K. Mizuno).

[1] The nucleotide sequence data reported here will appear in the EMBL, GenBank and DDBJ Nucleotide Sequence Databases under the accession numbers AB086414 (CCS1) and AB086415 (CtCS7).

*Abbreviations:* CCS, coffee caffeine synthase; CtCS, coffee tentative caffeine synthase; CTS, coffee theobromine synthase; RACE, rapid amplification of cDNA ends; SAM, *S*-adenosyl-L-methionine; TCS, tea caffeine synthase; XMP, xanthosine 5′-monophosphate

0014-5793/02/$22.00 
doi:10.1016S0014-5793(02)03781-X

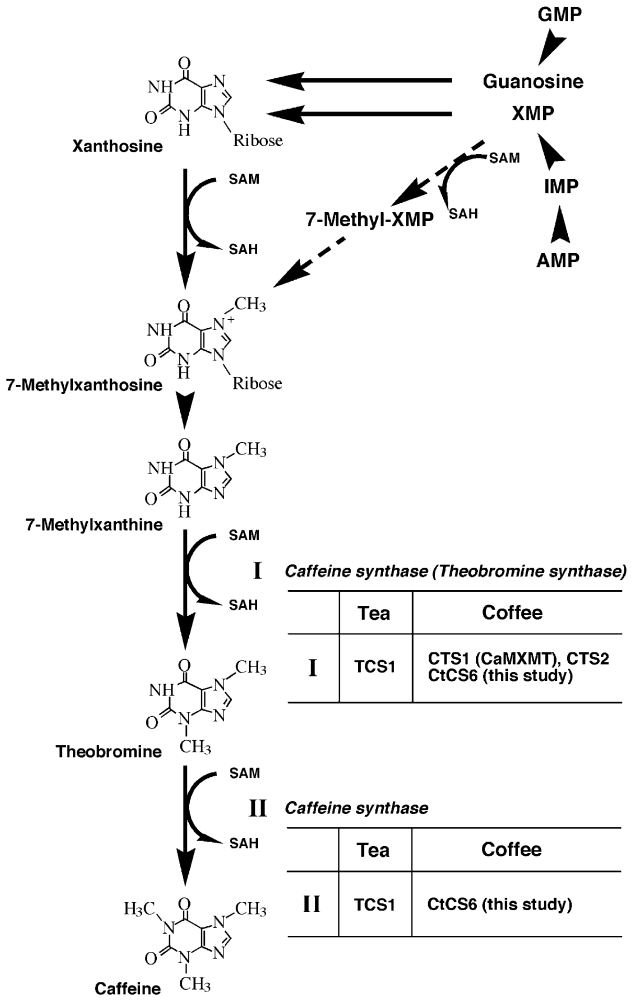


Fig. 1. Pathways for the biosynthesis of caffeine. Abbreviations: GMP, guanosine 5′-monophosphate; XMP, xanthosine 5′-monophosphate; IMP, inosine 5′-monophosphate; AMP, adenosine 5′-monophosphate; SAM, *S*-adenosyl-L-methionine; SAH, *S*-adenosyl-L-homocysteine. A XMP→7-methyl XMP→7-methylxanthosine pathway is operative in coffee leaves [6,12]. Caffeine synthase from tea leaves has broad substrate specificity and catalyzes the conversion of 7-methylxanthine to caffeine via theobromine (reactions I and II) [7].

primer. To obtain the 3′ regions of the CtCS cDNAs, we used two degenerated oligonucleotides, 5′-TICAYTGGYTIWCIMARGYICC-3′ (coding V/F HWLSQVP on CTS1 and CTS2 sequences) called *CS8f* and 5′-RTIGARRRRGAIGGIWSITT-3′ (coding VEEEGS C/F on CTS1 and CTS2 sequences) called *CS9f* in 3′-RACE. Polymerase chain reaction (PCR) was conducted in a PTC-200 thermal cycler (MJ Research) for 30 cycles (60 s at 94°C, 60 s at 55°C, and 90 s at 72°C) with the first-strand cDNAs and primers, *CS8f* and *3 sites adaptor primer* (*3SAP*) included in the 3′-RACE core set sequence. Subsequently nested second PCR was conducted for 30 cycles (60 s at 94°C, 60 s at 55°C, and 40 s at 72°C) with *CS9f* and *3SAP* as primers and the products of the first PCR as templates.

The nucleotide sequences of the 5′ regions of CtCS clones were obtained by 5′-RACE with the 5′ full RACE set (TaKaRa). *CS10r* (5′-CATATCTAAAAACAAACCAC-3′ coding 3′-untranslated region) was used for reverse transcription. The primer sets for 5′-RACE were as follows: *CS11f* (5′-CSGGCAAAGGCTTCTATGATAG-3′) and *CS12r* (5′-CGWAAATTGATCTAACGACA-3′). The PCR was conducted for 30 cycles with a PTC-200 thermal cycler (MJ Research) set to 94°C for 60 s, 62°C for 60 s, and 72°C for 90 s. The amplified DNA fragments were purified by polyacrylamide gel electrophoresis and were subcloned into the pGEM-T Easy vector (Promega).

### 2.3. Analysis of gene expression by semi-quantitative reverse transcriptase (RT)-PCR

For semi-quantitative RT-PCR, total RNAs extracted from developing endosperms, flower buds and young leaves of coffee were treated with RNase-free DNase I (TaKaRa). DNA-free total RNA (382 ng) from each tissue was used for first-strand cDNA synthesis in 20 μl reaction volume containing 2.5 units of AMV reverse transcriptase XL (TaKaRa) and 1 μM of *oligo-dT 3SAP*. The 20 μl PCR reaction mixture contained 0.1 μl of RT reaction mixture, 2.5 units of Gene Taq (Nippon Gene), 2.5 mM $MgCl_2$, 2 mM each of deoxyribonucleotide triphosphate and 0.5 μM of gene-specific primers. The gene-specific primers used were 5′-GGCTTCTATGATAGTG-3′ and 5′-CCTTTCCCCATATCTA-3′ for CtCS6 and CtCS7; 5′-GGAGGAAGGTTCTTTTG-3′ and 5′-ATCGCCGTATACTTGGA-3′ for CTS2. PCR run on a PTC-200 thermal cycler (MJ Research) using the following program: 94°C for 1 min, 57°C for 30 s, and 72°C for 10 s. The amplification was done for 18–30 cycles and reaction tubes were removed every two cycles. The amplicons were 115 (CtCS6 and CtCS7) and 117 (CTS2) bp lengths, respectively. The amplification for each gene showed a linear curve. The reaction products were visualized on 10% polyacrylamide gels stained by ethidium bromide and ultraviolet (UV) transillumination. Semi-quantitative analysis of the intensify of fluorescence was conducted by a Macintosh computer using 'NIH image' software (http://rsb.info.nih.gov/nih-image/).

### 2.4. Construction of expression plasmids

Expression plasmids for CtCS clones were constructed in a pET32a or pET23d vector (Novagen). To prepare expression plasmids for CtCS cDNAs, we carried out PCR-directed mutagenesis to produce an *Nco*I site at the translation initiation site in the cDNA sequence of CtCS clones. The primer carrying *Nco*I restriction sites (bold-faced type) for cloning were as follows: N1, 5′-**CCATGG**AGCTCCAAGAAGTCC-3′. The pairs of primers were as follows: N1/CS10r for CtCS6 and CtCS7. The reaction program consisted of 30 cycles of 94°C for 60 s, 55°C for 60 s, and 72°C for 90 s. The PCR products were subcloned into pGEM-T Easy vector (Promega). The subclones of the PCR products for CtCS6 and CtCS7 were digested with *Nco*I and *Sac*I. The *Nco*I/*Sac*I fragments were introduced into the pET32a or pET23d vector at the *Nco*I and *Sac*I sites, respectively. The resulting expression plasmids were introduced into the expression host *Escherichia coli* BL21(DE3).

### 2.5. Production of recombinant enzymes in E. coli and assay for enzyme activity

A single colony of transformants was cultured at 37°C overnight in 3 ml of Luria broth containing 0.2 mg/ml ampicillin (LA) with constant shaking. A portion (50 μl) of the bacterial culture was added to 3 ml of fresh LA, and incubated at 37°C for 2 h with shaking. Production of the recombinant proteins was induced by adding 9 μl of 0.1 M isopropyl-β-D-thio-galactopyranoside (final concentration is 0.3 mM), and cell growth was continued for 8 h at 25°C. *E. coli* cells were harvested by centrifugation at 500×*g* for 5 min, and then washed with 50 mM Tris/HCl pH 8.5 containing 5 mM dithiothreitol (DTT), 2 mM ethylenediamine tetraacetic acid (EDTA)-$Na_2$, 50 mM NaCl and 20% glycerol (called TES-G buffer). This cell paste was suspended in 200 μl of TES-G buffer, frozen at −80°C, sonicated to disrupt cells and then centrifuged at 10 000×*g* for 10 min at 4°C. The supernatant was used in the enzyme assay for caffeine metabolism. Enzymatic activity of CS was determined as described previously [7]. The $K_m$ value was derived from Lineweaver–Burk plots analyzed with a Macintosh G3 computer using 'Enzyme Kinetics' software (Trinity Software, Campton, NH, USA). Thin layer chromatography (TLC) was performed as described previously with minor modification [12].

### 2.6. Analytical procedures

Protein concentrations were measured by the method of Bradford [13]. Polyacrylamide gel electrophoresis in the presence of 0.1% sodium dodecyl sulfate (SDS–PAGE) was carried out by the method of Laemmli [14]. Northern and Western blot analyses were carried out as described previously [15,16]. Nucleotide sequencing was carried out by the dideoxy chain-termination method of Sanger et al. [17] using a Perkin-Elmer genetic analyzer (model 310) and an ABI prism dye terminator cycle sequencing kit (Perkin-Elmer). Computer-aided analysis of nucleotide and protein sequences was performed using GENETYX software (Software Development Co., Tokyo, Japan).

## 3. Results and discussion

### 3.1. cDNA cloning of caffeine synthase, SAM-dependent N-methyltransferase from coffee

Though we have cloned five independent CtCS clones including CTS1 and CTS2 [9], the existence of more CtCS genes in a coffee genome has been predicted by Southern blot analysis [9]. The objective gene(s), CtCS(s), could be contained in those unisolated CtCS genes. That CtCS genes are expressed in coffee young leaves, flower buds and developing fruits was revealed by Northern blot analysis of total RNAs from those tissues, contain caffeine (Fig. 2). The CtCS cDNAs (CtCS1, CtCS3, CtCS4, CTS1 and CTS2) cannot be distinguished, because they are over 80% identical to each other at the nucleotide level. We carried out RT-PCR and RACE technique using total RNA from developing coffee endosperm as a template, because the strongest signals were detected in developing fruits (Fig. 2). The three conserved motifs (A, B and C) of the binding site of the methyl donor SAM have been reported in most plant SAM-dependent *O*-methyltransferases (OMTs)

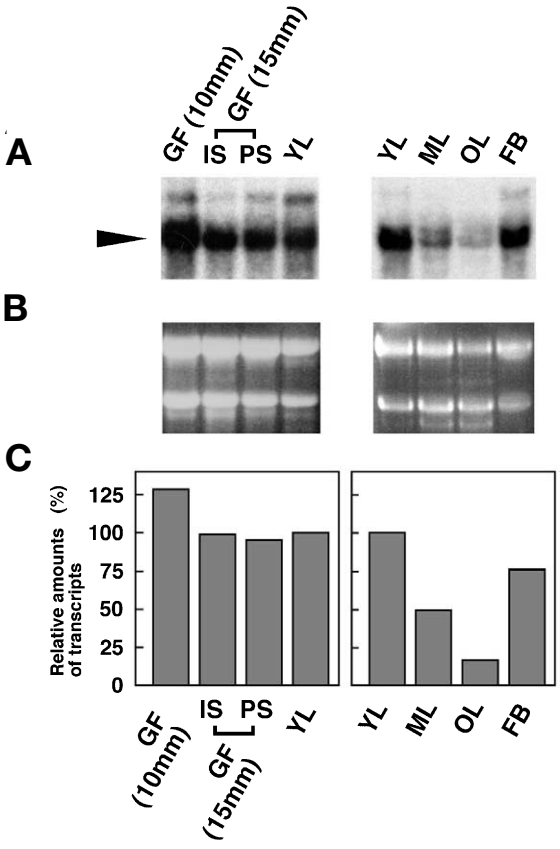


Fig. 2. Northern blot analysis of total cellular RNAs from coffee tissues. Total RNAs (10 μg each) of green fruit at two developing stages, leaves at various stages and flower buds were separated by agarose gel electrophoresis, transferred onto a Hybond $N^+$ nylon membrane, and probed with a $^{32}$P-labeled cDNA fragment encoding CTS2 (A). The amount of total RNA on the gel from each tissue was estimated by ethidium bromide staining (B). There are many transcripts of the genes for CtCS in developing fruits, young leaves and flower buds. The relative amounts of the transcripts are shown in C. Values were obtained by scanning the blot with a scanner and then analyzing the results with NIH image. GF, green fruits, the diameter of the fruits indicated in parentheses; IS, immature seeds; PS, pulps and outer skins; YL, young leaves; ML, mature leaves; OL, old leaves; FB, flower buds.

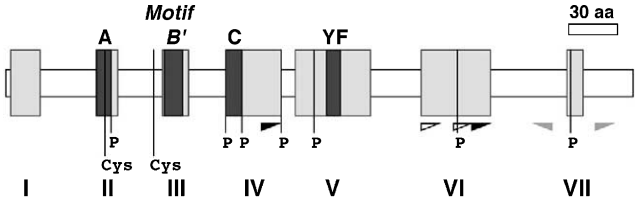


Fig. 3. Schematic diagram of homologous regions of motif B′ methyltransferase family. The conserved sequences in motif B′ methyltransferase family are indicated by open boxes (I, II, III, IV, V, VI, and VII). The proposed SAM-binding motifs (A, B′ and C) and the conserved region nominated as 'YFFF region' (YF) shown by closed boxes. Conserved amino acid residues, especially proline and cysteine, are indicated by P and Cys, respectively. The degenerated oligonucleotides designed for 3′-RACE primers are indicated by open and closed arrowheads, which were used for amplification of CtCS fragments. Consequently, *CS8f* and *CS9f* indicated by closed arrowheads were effective primers for the amplification of CtCS6 and CtCS7. Primers for 5′-RACE are shown by shaded arrowheads.

[18]. While TCS1 [8], SAMT [19], floral nectary-specific protein from *Brassica campestris* L. sp. *pekinensis* (NTR1) [20], BAMT [21] and jasmonic acid carboxyl methyltransferase from *Arabidopsis thaliana* (JAMT) [22] do not have motif B, the amino acid sequences of these five enzymes have another conserved region nominated here as motif B′ (Fig. 3). CTS1 and CTS2 also have three motifs (A, B′ and C), but lack motif B [9]. Here we call these methyltransferases 'motif B′ methyltransferase family'. From comparison of the amino acid sequences of motif B′ methyltransferases, three clones characterized previously and CtCSs cloned here, new conserved regions, other than SAM motifs, were identified. Four downstream degenerate oligonucleotide primers corresponding to the homologs regions were synthesized. The positions of the primers are shown in Fig. 3. We carried out 3′-RACE using total RNA from developing coffee endosperm as a template. Although a lot of CtCS clones already identified were amplified, two novel CtCS clones were isolated in the case of using CS8f and CS9f as a primer. These two cDNA sequences were tentatively named CtCS6 and CtCS7, respectively. The full-length cDNAs of CtCS6 and CtCS7 were isolated by the 5′-RACE method. CtCS6 and CtCS7 (accession numbers AB086414 and AB086415) have 1331 and 1344 nucleotides, respectively, and encode 384 amino acid residues.

### 3.2. Comparison of the sequences of CtCSs and other N-methyltransferases

When the amino acid sequences of CtCS6 and CtCS7 were compared with those of CtCS series, CtCS6 and CtCS7 shared a high degree of sequence identity (approximately 80%) with these series (Table 1). The amino acid sequence of CTS2 had the highest sequence identity (82.6%) to those of CtCS6 and CtCS7 (Table 1). The sequence identity between CtCS6 and CtCS7 was 94.5%, and 21 of the 384 amino acids are substituted. Recently, 10 kinds of sequences of CtCS homologs derived from *C. arabica*, *Coffea canephora*, and *Coffea liberica* have been submitted to the DDBJ sequence database. As comparison of the amino acid sequence of CtCS6 with those of these homologs showed that three sequences isolated from *C. liberica* (AF494418, AF494419 and AF494420), by Kretschmar and Baumann, indicated high sequence identity (approximately 93%) with CtCS6, it is most likely that these three proteins corresponded to CtCS6 or CtCS7 in *C. liberica*.

Table 1
Percentage of sequence identity among motif B′ methyltransferases

| | CtCS6 | CtCS7 | CTS1 | CTS2 | CtCS1 | CtCS3 | CtCS4 | TCS1 | SAMT | BAMT | JAMT |
|---|---|---|---|---|---|---|---|---|---|---|---|
| 1: CtCS6 | 100.0 | **94.5** | **80.2** | **82.6** | **81.0** | **82.3** | **80.8** | 38.9 | 38.4 | 37.7 | 37.1 |
| 2: CtCS7 | | 100.0 | **81.0** | **83.1** | **82.1** | **82.1** | **81.0** | 38.9 | 40.2 | 37.8 | 37.6 |
| 3: CTS1 | | | 100.0 | **93.2** | **84.4** | **80.8** | **80.0** | 39.5 | 39.7 | 37.1 | 38.6 |
| 4: CTS2 | | | | 100.0 | **85.2** | **82.9** | **82.3** | 38.4 | 39.1 | 37.2 | 37.8 |
| 5: CtCS1 | | | | | 100.0 | **82.9** | **81.6** | 39.5 | 40.6 | 37.4 | 37.8 |
| 6: CtCS3 | | | | | | 100.0 | **95.6** | 40.7 | 39.5 | 36.3 | 39.1 |
| 7: CtCS4 | | | | | | | 100.0 | 40.3 | 40.0 | 37.1 | 37.8 |
| 8: TCS1 | | | | | | | | 100.0 | 41.8 | 38.5 | 39.1 |
| 9: SAMT | | | | | | | | | 100.0 | 43.9 | 43.0 |
| 10: BAMT | | | | | | | | | | 100.0 | 40.5 |
| 11: JAMT | | | | | | | | | | | 100.0 |

1, 2: CtCS6 (AB086414) and CtCS7 (AB086415) from *C. arabica* tentative caffeine synthase. 3, 4: CTS1 (AB034700) and CTS2 (AB054841) from *C. arabica* theobromine synthase. 5–7: CtCS1 (AB034699), CtCS3 (AB054842) and CtCS4 (AB054843) from *C. arabica* tentative caffeine synthase. 8: TCS1 from *C. sinensis* caffeine synthase (AB031280). 9: SAMT from *C. breweri* salicylic acid carboxyl methyltransferase (AF133053). 10: BAMT from *A. majus* benzoic acid carboxyl methyltransferase (AF198492). 11: JAMT from *A. thaliana* jasmonic acid carboxyl methyltransferase (AF008434). The values over 80% are indicated in bold.

Since it is shown that among Coffea species *C. arabica* is the only amphidiploid cultivar ($2n = 44$), and that *C. liberica* has $2n = 22$ chromosomes [23], possibly unidentified genes that are very similar to CtCS6 or CtCS7 may be present in *C. arabica*. As compared with the motif B′ methyltransferases, TCS1, SAMT, NTR1, BAMT, JAMT, CtCS6 and CtCS7 shared a low amino acid sequence similarity (approximately 40%) (Table 1).

### 3.3. Expression of CtCS genes

As CtCS series are highly homologous with each other on the nucleotide level (over 80% identical), the amount of transcripts from these genes cannot be detected by Northern blot analysis. The expression of CtCS6 and CtCS7 was evaluated by semi-quantitative RT-PCR (Fig. 4) designed for specific amplification and detection of these transcripts. Although the transcripts of CtCS6 and CtCS7 were detected in all organs used in this experiment (developing endosperm, young leaf, and flower bud), the gene expression in the developing endosperm was the strongest of that in these organs. To compare the expression pattern, we also carried out RT-PCR for the transcripts of CTS2. The gene of CTS2 was expressed in all organs and the amount of transcripts in the developing endosperm was the highest of that in these organs like that of CtCS6 and CtCS7. It is likely that the differential expression of CtCS6 and CTS2 defined different caffeine synthetic pathways between young leaves and endosperms in the coffee plant.

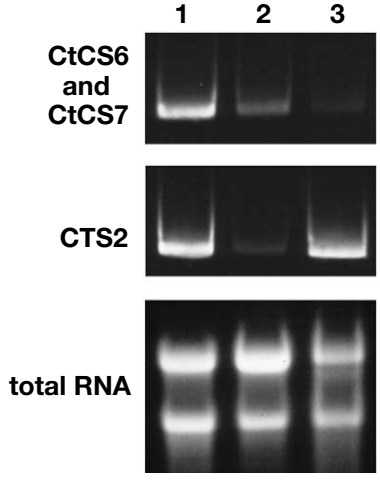


Fig. 4. Expression of CtCS6, CtCS7 and CTS2 in various tissues. RT-PCR for CCS6, CtCS7 and CTS1 were performed with 18–30 cycles respectively to determine the linear range of PCR amplifications. Here, the results obtained with 26 cycles for all templates and genes are shown. The quality and concentration of total RNA from each tissue were estimated by ethidium bromide staining. Lane 1, developing endosperms; lane 2, young leaves; lane 3, flower buds.

### 3.4. Properties of recombinant enzymes

The cDNAs were translated with an *E. coli* expression system, to elucidate whether the proteins encoded by the cloned cDNAs are related to the caffeine biosynthetic pathway. Expression plasmids for CtCS6 and CtCS7 cDNA were constructed in pET23d and pET32a (Novagen) vectors, respectively. Recombinant CtCS6 and CtCS7 proteins from both expression vectors were produced in *E. coli* BL21(DE3), respectively. SDS–PAGE and subsequent Western blot analysis using anti-CTS2 antibody showed that these recombinant proteins were produced in soluble protein fractions (Fig. 5). When the crude extracts of these *E. coli* were incubated with a variety of xanthine substrates of caffeine in the presence of [methyl-$^{14}$C]SAM as a methyl donor, CtCS6 catalyzed 1-*N*-methylation and 3-*N*-methylation of the purine ring of both mono- and dimethylxanthines. 7-*N*-methylation activity was not detected. When the methylated products from this assay were separated by TLC (Fig. 6), it was clear that CtCS6 catalyzed the conversion of both 7-methylxanthine to theobromine and theobromine to caffeine. Although recombinant CtCS7 was produced in *E. coli* soluble extract, only the thioredoxin (Trx) fusion protein from the pET32a system had weak activity to convert paraxanthine to caffeine in this assay system. As the activity was so weak, we abandoned further analysis of CtCS7 protein. Since CtCS6 catalyzes 1-*N*- and 3-*N*-methylation of the purine ring to convert 7-methylxanthine to theobromine and theobromine to caffeine, like tea TCS1, this clone (CtCS6) was renamed coffee caffeine synthase 1 (CCS1).

Hereafter we focused on the enzymatic properties of CCS1. The recombinant Trx fusion protein derived from the pET32a system and the native form protein derived from pET23d system had $K_m$ values for 7-methylxanthine of 75.1 and 125.6 μM, respectively. As indicated previously, the $K_m$ value of CTS1 (CaMXMT) for 7-methylxanthine was different from

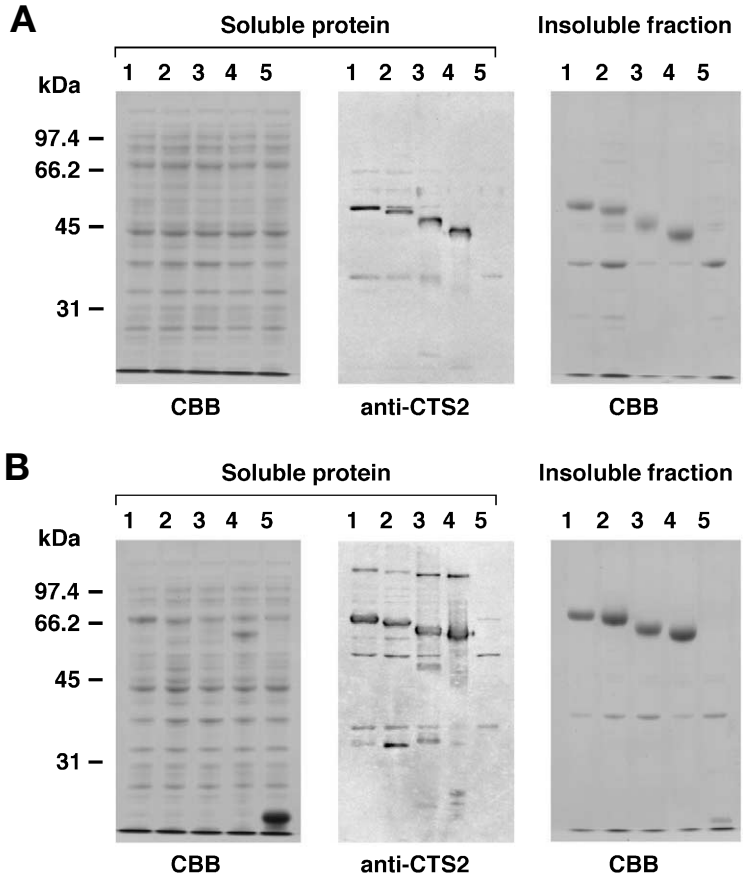


Fig. 5. SDS–PAGE and Western blot analysis of recombinant CtCS proteins. Soluble protein fractions and insoluble fractions from *E. coli* cells were subjected to SDS–PAGE, respectively, that were transformed by expression plasmids, derived from pET23d (A) or pET32a (B), for CtCS clones. Gels were stained with Coomassie brilliant blue R-250 (*CBB*). Soluble protein fractions were subjected to Western blot analysis (*anti-CTS2*). The blot was probed with anti-CtCS2 antibody. Lane 1, CtCS6; lane 2, CtCS7; lane 3, CTS1; lane 4, CTS2; lane 5, control (transformant of uninserted pET23d or pET32a).

that of the native form (873 μM) or that of GST fusion (50 μM) [10]. The difference in the $K_m$ values may be due to the different structure of these recombinant proteins [9]. Therefore, we estimated the enzymatic properties of CCS1 using the native form of recombinant protein from the pET23d system. Table 2 contains data on the substrate specificity of the recombinant CCS1. In contrast to CTS2, which catalyzes 7-methylxanthine to theobromine but does not have 1-*N*-methylation activity, CCS1 has dual methylation activities that convert 7-methylxanthine to theobromine and theobromine to caffeine like tea TCS1. Of three monomethylxanthines investigated, 7-methylxanthine was methylated the most readily. As shown in Table 2, this implies that the order of methylation of the purine base is $N\text{-}3 > N\text{-}1 \gg N\text{-}7$. The specific activity of crude extract of recombinant CCS1 was 275.4 pkat/mg proteins with 7-methylxanthine. The value was much higher than the 5.4 pkat/mg obtained with TCS1 activity from the *E. coli* crude extract [8] and the 10 fkat/mg obtained with NMT activity isolated from the liquid endosperm of *C. arabica* seed [24]. The $K_m$ values of CCS1 for paraxanthine, 7-methylxanthine, and theobromine were 30.8, 125.6, and 157.0 μM, respectively. The $K_m$ value of CCS1 for 7-methylxanthine (125.6 μM) is lower than those of CTS1 (873 μM) and CTS2 (171 μM). On the other hand, the $K_m$ values of tea TCS for paraxanthine, 7-methylxanthine, and theobromine were 24, 186, and 344 μM, respectively, as

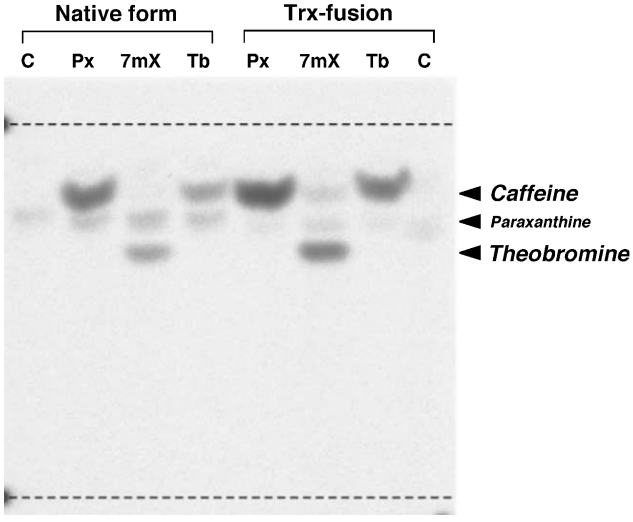


Fig. 6. Visualized reaction products derived from recombinant CtCS6. The autoradiograph of a TLC separation of the reaction products derived from recombinant enzymes of native form and Trx fusion. Products of the reaction mixture were extracted by chloroform and subjected to the TLC on a cellulose plate using a solvent of *n*-butanol/acetic acid/water (4/1/2, v/v/v). Autoradiography was conducted using an Image-Analyzer system (FLA-2000 Fuji-film Co.). Px, 7mX, and Tb indicated substrates. C, control (no substrate); Px, paraxanthine; 7mX, 7-methylxanthine; Tb, theobromine.

Table 2
Comparison of the substrate specificity of recombinant and native caffeine (theobromine) synthases

| *Plant* | Substrate/methylation position[a] | | | | | | References |
|---|---|---|---|---|---|---|---|
| | 7-mX/3*N* % | 3-mX/1*N* | 1-mX/3*N* | Tb/1*N* | Px/3*N* | Tp/7*N* | |
| *Coffee* | | | | | | | |
| Recombinant CCS1 (native form) | 100 | 3.4 | 2.2 | 102 | 416 | – | this study |
| Recombinant CCS1 (Trx fusion) | 100 | 3.5 | 2.7 | 142 | 622 | – | this study |
| Recombinant CTS1 (native form) | 100 | nd | nd | nd | 1.4 | nd | Mizuno et al. (2001) [9] |
| Recombinant CTS1 (GST fusion) | 100 | nd | nd | nd | 15 | nd | Ogawa et al. (2001) [10] |
| Endosperm (partially purified) | 100 | – | – | 185 | – | – | Mazzafera et al. (1994) [24] |
| Fruits (crude) | 100 | – | 5.7 | 127 | 175 | 4.6 | Roberts and Waller (1979) [25] |
| Leaves (crude) | 100 | – | – | – | – | – | Schulthess et al. (1996) [6] |
| *Tea* | | | | | | | |
| Recombinant TCS1 (native form) | 100 | 1 | 12.3 | 18.5 | 230 | tr | Kato et al. (2000) [8] |
| Leaves (purified) | 100 | 17.6 | 4.2 | 26.8 | 210 | tr | Kato et al. (1999) [7] |
| Methylated product | Tb | Tp | Tp | caffeine | caffeine | caffeine | |

The relative activity is indicated as the percentage of the activity with 7-mX. Caffeine synthase activities of the recombinant enzymes, native form and Trx fusion, with 7-mX (100%) were 33.1 and 5.6 pkat/mg, respectively. nd, not detected; tr, trace; –, not determined.
[a]mX, methylxanthine; Tb, theobromine; Tp, theophylline; Px, paraxanthine.

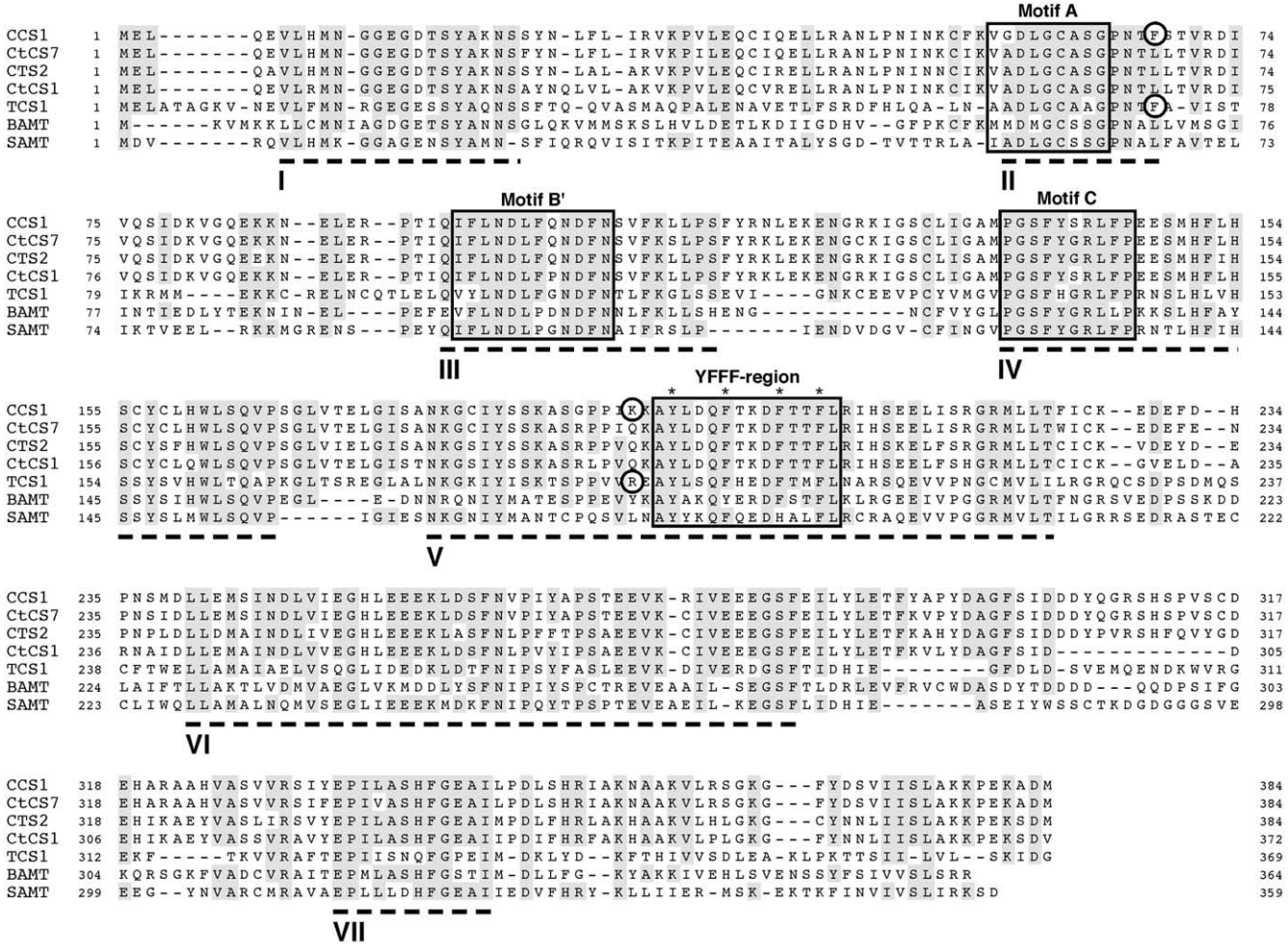


Fig. 7. Comparison of the amino acid sequences of CtCS proteins with related motif B′ methyltransferases. Alignment of the amino acid sequences for CCS1 (CtCS6), CtCS7, CTS2 and CtCS1 (accession numbers AB086414, AB086415, AB054841 and AB034699, respectively) from *C. arabica*, TCS1 from tea (accession number AB031280), BAMT from *Antirrhinum majus* (accession number AF198492), and SAMT from *Clarkia breweri* (accession number AF133053). Shaded boxes represent conserved amino acid residues, and dashes represent gaps that have been inserted for optimal alignment. The proposed SAM-binding motifs (A, B′ and C) and the conserved region nominated as 'YFFF region' are shown by open boxes. The conserved sequences in the motif B′ methyltransferase family are indicated by dotted lines (I, II, III, IV, V, VI, and VII). The postulated important amino acid residues for the substrate specificity of *N*-methylation are marked by open circles.

shown previously [7]. Though the sequence homology between tea TCS1 and CCS1 is almost 40%, these substrate specificities are only slightly similar.

## 4. General final discussion

CCS1 only has the activity of 1-*N*-methylation of theobromine. Although, all members of the proteins catalyzing *N*-methylation in the caffeine synthesis pathway from coffee, including this gene and other CtCSs, are highly homologous (>80%) with each other, they have no such activities except CCS1. On the contrary, TCS1 also has this enzymatic activity although with only an overall 40% homology to CCS1. These genes including TCS1 share four highly conserved regions, motif A, motif B′, motif C and YFFF region which will be discussed later (Fig. 7), with slight deviations. These regions may have important roles both in the common *N*-methylating catalytic reaction and in discriminating the positions of *N*-methylation of the purine ring. Motif A is conceived to be a SAM-binding site [18] which commonly plays an important role in these proteins. An amino acid residue flanked to this important region is Phe in CCS1 (F68) and TCS1 (F73). Both of them are 1-*N*-methylating enzymes, but the amino acid at the corresponding position of other proteins catalyzing *N*-methylation other than 1-*N* position is Leu. The physicochemical properties of the amino acid residues at this position may have some important role in the structural discrimination of *N*-methylation. These proteins and other proteins with motif B′ have a fourth conserved region downstream to motif C which was nominated as 'YFFF region' (Fig. 7). The consensus sequence of this region is 'AYXXQFXXDFXXFL' containing bulky amino acid residues. An amino acid residue upstream of this motif of CCS1 is Lys (K193) and the corresponding amino acid residue of TCS1 is Arg (R192). These were not the same, but had similar physicochemical properties. On the contrary, other proteins with *N*-methylating activity from coffee have Gln at the corresponding position. The conserved features and the difference of this region may be related to the positional specificity of *N*-methylation, although this must be confirmed by further experiments.

An amino acid sequence, Val159-His160-Tyr161 (VHW) in CaMXMT (=CTS1), also exists in TCS1, but not in other *N*-methylating enzymes from coffee. The importance of this sequence for the substrate specificity of *N*-methylation was postulated by Ogawa et al. [10]. As the corresponding sequence in CCS1 and other coffee enzymes is LHW, this region may not be important for the substrate specificity of these enzymes.

The caffeine-synthesizing genes cloned here may help to develop transgenic caffeine-deficient *C. arabica* plants through antisense mRNA technology or by gene silencing.

## References

[1] Ashihara, H., Monteiro, A.M., Gillies, F.M. and Crozier, A. (1996) Plant Physiol. 111, 747–753.
[2] Ashihara, H., Gillies, F.M. and Crozier, A. (1997) Plant Cell Physiol. 38, 413–419.
[3] Ashihara, H. and Crozier, A. (1999) in: Biosynthesis and Metabolism of Caffeine and Related Purine Alkaloids in Plants (Callow, J.A., Ed.), Advances in Botanical Research, Vol. 30, pp. 117–205, Academic Press, London.
[4] Ashihara, H. and Crozier, A. (2001) Trends Plant Sci. 6, 407–413.
[5] Nazario, G.M. and Lovatt, C.J. (1993) Plant Physiol. 103, 1203–1210.
[6] Schulthess, B.H., Morath, P. and Baumann, T.W. (1996) Phytochemistry 41, 169–175.
[7] Kato, M., Mizuno, K., Fujimura, T., Iwama, M., Irie, M., Crozier, A. and Ashihara, H. (1999) Plant Physiol. 120, 579–586.
[8] Kato, M., Mizuno, K., Crozier, A., Fujimura, T. and Ashihara, H. (2000) Nature 406, 956–957.
[9] Mizuno, K., Tanaka, H., Kato, M., Ashihara, H. and Fujimura, T. (2001) in: International Scientific Colloquium on Coffee, Vol. 19, ASIC, Paris, pp. 815–818.
[10] Ogawa, M., Herai, Y., Koizumi, N., Kusano, T. and Sano, H. (2001) J. Biol. Chem. 276, 8213–8218.
[11] Chang, S., Puryear, J. and Cairney, J. (1993) Plant Mol. Biol. Rep. 11, 113–116.
[12] Kato, M., Kanehara, T., Shimizu, H., Suzuki, T., Gillies, F.M., Crozier, A. and Ashihara, H. (1996) Physiol. Plant. 98, 629–636.
[13] Bradford, M.M. (1976) Anal. Biochem. 72, 248–254.
[14] Laemmli, U.K. (1970) Nature 272, 680–685.
[15] Mizuno, K., Kawasaki, T., Shimada, H., Satoh, H., Kobayashi, E., Okumura, S., Arai, Y. and Baba, T. (1993) J. Biol. Chem. 268, 19084–19091.
[16] Mizuno, K., Kobayashi, E., Tachibana, M., Kawasaki, T., Fujimura, T., Funane, K., Kobayashi, M. and Baba, T. (2001) Plant Cell Physiol. 42, 349–357.
[17] Sanger, F., Nicklen, S. and Coulson, A.R. (1977) Proc. Natl. Acad. Sci. USA 74, 5463–5467.
[18] Joshi, C.P. and Chiang, V.L. (1998) Plant Mol. Biol. 37, 663–674.
[19] Ross, J.R., Nam, K.H., D'Auria, J.C. and Pichersky, E. (1999) Arch. Biochem. Biophys. 367, 9–16.
[20] Song, J.T., Seo, H.S., Song, S.I., Lee, J.S. and Do, C.Y. (2000) Plant Mol. Biol. 42, 647–655.
[21] Dudareva, N., Murfitt, L.M., Mann, C.J., Gorenstein, N., Kolosova, N., Kish, C.M., Bonham, C. and Wood, K. (2000) Plant Cell 12, 949–961.
[22] Seo, H.S., Song, J.T., Cheong, J.J., Lee, Y.H., Lee, Y.W., Hwang, I., Lee, J.S. and Choi, Y.D. (2001) Proc. Natl. Acad. Sci. USA 98, 4788–4793.
[23] Raina, S.N., Mukai, Y. and Yamamoto, M. (1998) Theor. Appl. Genet. 97, 1204–1209.
[24] Mazzafera, P., Wingsle, G., Olsson, O. and Sandberg, G. (1994) Phytochemistry 37, 1577–1584.
[25] Roberts, M.F. and Waller, G.R. (1979) Phytochemistry 18, 451–455.