

Cloning and functional characterization of the *cis*-aconitic acid decarboxylase (CAD) gene from *Aspergillus terreus*

Shin Kanamasa · Lies Dwiarti · Mitsuyasu Okabe ·
Enoch Y. Park

Received: 21 March 2008 / Revised: 24 April 2008 / Accepted: 26 April 2008 / Published online: 27 June 2008
© Springer-Verlag 2008

Abstract A filamentous fungus *Aspergillus terreus* produces itaconic acid, which is predicted to be derived from *cis*-aconitic acid via catalysis by *cis*-aconitic acid decarboxylase (CAD) in the carbon metabolism of the fungus. To clarify the enzyme's function and a pathway for itaconic acid biosynthesis, we cloned a novel gene encoding the enzyme. The open reading frame of this gene (*CAD1*) consists of 1,529 bp encoding 490 amino acids and is interrupted by a single intron. Among the identified proteins in the database, the primary structure of the protein encoded by *CAD1* shared high identity with the MmgE/PrpD family of proteins, including a number of 2-methylcitrate dehydratases of bacteria. The cloned gene excluding an intron was introduced into the expression plasmid pAUR-CAD1 controlled by the *ADHI* promoter. The CAD activity in *Saccharomyces cerevisiae* was confirmed by directly detecting itaconic acid as a product from *cis*-aconitic acid as a substrate. This result reveals for

the first time that this gene encodes CAD, which is essential for itaconic acid production in *A. terreus*.

Keywords *cis*-Aconitic acid decarboxylase · *Aspergillus terreus* · Itaconic acid · Gene cloning · Gene expression

Introduction

Itaconic acid (methylenesuccinic acid, C₅H₆O₄) is a soluble unsaturated dicarboxylic acid mainly produced from sugars by several fungi. In 1931, Kinoshita reported that the filamentous fungus *Aspergillus itaconicus* produced itaconic acid in a medium with high sugar content as a carbon source (Kinoshita 1931). *Aspergillus terreus* is also an itaconic acid producer (Calam et al. 1939), and has been used for the industrial production of itaconic acid because of its high production rate. Itaconic acid is used as a monomer to form polymers that are widely used as raw materials for latex, synthetic resins, adhesives, paint, additives for acrylic resin fibers and paper, etc. It is also used as an acidulant and for the pH adjustment of foods. However, applications of this useful material in other industries have been limited because it is rather expensive.

Several groups have attempted to clarify the biosynthetic pathway of itaconic acid in *A. terreus*, with the aim of obtaining knowledge for metabolic engineering. Bentley and Thiessen (1957a,b) showed that *cis*-aconitic acid, produced in a tricarboxylic acid (TCA) cycle, could be a substrate of *cis*-aconitic acid decarboxylase (CAD, EC 4.1.1.6), forming itaconic acid. However, the primary structure of CAD was still unknown when we achieved the first purification and characterization. We purified a 55-kDa protein that had CAD activity from *A. terreus*

S. Kanamasa
JST Innovation Satellite Shizuoka,
Japan Science and Technology Agency,
3-5-1 Johoku, Naka-ku,
Hamamatsu 432-8561, Japan

L. Dwiarti · M. Okabe · E. Y. Park
Laboratory of Biotechnology,
Department of Applied Biological Chemistry,
Faculty of Agriculture, Shizuoka University,
836 Ohya, Suruga-ku,
Shizuoka 422-8529, Japan

E. Y. Park (✉)
Laboratory of Biotechnology, Integrated Bioscience Section,
Graduate School of Science and Technology, Shizuoka University,
836 Ohya, Suruga-ku,
Shizuoka 422-8529, Japan
e-mail: acypark@ipc.shizuoka.ac.jp

TN484-M1 to homogeneity and characterized the enzyme, concluding that the purified protein was essential for itaconic acid production in the fungus (Dwiarti et al. 2002). A gene coding for CAD, however, has never been cloned from any organism, despite the fact the enzyme directly participates in one of the major pathways of the TCA cycle, and the structure of the gene is still unknown. To date, some scientists have intended to clone the gene for CAD, but difficulties related to the low stability of the enzyme perturbed the enzyme purification, which is a necessary experimental step for its amino acid sequencing and its gene identification.

In this study, the *CAD1* gene encoding CAD of *A. terreus* was cloned, and its function was confirmed by its functional expression in the yeast *Saccharomyces cerevisiae*.

Materials and methods

Strains, media, plasmids, and cultivation

A. terreus IFO6365 (wild-type strain) and *A. terreus* TN484-M1 (an high-itaconic acid-producing strain derived from the IFO6365) were used for the isolation of *CAD1* (Dwiarti et al. 2002; Yahiro et al. 1995). *Escherichia coli* DH5 α and the plasmid pT7Blue (Merck KGaA, Darmstadt, Germany) were used as the host and vector, respectively, for deoxyribonucleic acid (DNA) manipulation. *S. cerevisiae* INVSc1 (*MAT α* , *his3- Δ 1*, *leu2*, *trp1-289*, *ura3-52*; Invitrogen, California, USA) and plasmid pAUR123 containing the alcohol dehydrogenase gene promoter of *S. cerevisiae* (Takara Bio, Kyoto, Japan) were used as constructs for the expression of *CAD1*. Fungal genomic DNA was extracted using a DNeasy Plant Mini kit (Qiagen GmbH, Düsseldorf, Germany). YPD medium (1% w/v Bacto yeast extract, 2% w/v Bacto peptone, 2% w/v glucose) and YPD medium containing 1 μ g/ml Aureobasidin A (Takara Bio) were used for the growth of *S. cerevisiae* cells and their transformants, respectively.

For the transcription analysis of *CAD1*, *A. terreus* strains were inoculated into a 500-ml Erlenmeyer flask containing 50 ml of the medium consisting of (per liter) 20 g of glucose, 35 g of corn steep liquor, and 20 g of NaCl (pH 3.0) and were grown at 30°C for 48 h using a rotary shaker set at 170 rpm. Mycelia of these strains were harvested on filter paper and washed twice with sterile water, and then a 1.0 g sample of wet mycelium was inoculated into a 200-ml Erlenmeyer flask containing 20 ml of the production medium consisting of (per liter) 140 g of glucose, 2.1 g of corn steep liquor, 2.9 g of NH₄NO₃, and 1.8 g of MgSO₄·7H₂O (pH 2.0). To investigate the effect of itaconic acid on *CAD1* transcription, 100 g/l of itaconic

acid was added to the production medium (pH 2.0). Cultivation was performed at 30°C for 12 h using a rotary shaker at 170 rpm.

Enzyme assay and measurement of protein concentration

The CAD activity was measured according to the method of Bentley (Bentley and Thiessen 1957c). Briefly, 0.1 ml of the enzyme solution was incubated with 0.4 ml of *cis*-aconitic acid solution (final concentration, 8.0 mM) and 2.5 ml of 0.2 M sodium phosphate buffer (pH 6.2) for 10 min at 37°C. The enzyme reaction was terminated by the addition of 0.1 ml of 12 M HCl. The released itaconic acid was measured by high-performance liquid chromatography (HPLC; JASCO, Tokyo). The samples were filtered through a 0.45- μ m filter and injected onto a column (Capcell Pak C18 MG, Lot BS14 Shiseido, Shiseido, Tokyo) of 4.6 mm diameter and 150 mm length. Itaconic acid was identified using a UV detector (UV-970, JASCO) at 210 nm. A solution of 2.5% acetonitrile and 0.1% phosphoric acid was premixed, degassed, and used as the mobile phase for analysis at a flow rate of 1.0 ml/min and temperature of 45°C.

The protein concentration was measured using the protein assay CBB solution (Nacalai Tesque) based on the Bradford method (Bradford 1976) with bovine serum albumin as a standard. The results were expressed as mean \pm standard deviation (SD) of three independent experiments.

Purification of CAD, amino acid sequence analysis, and isolation of a genomic clone encoding *A. terreus* CAD

Purified CAD was prepared from *A. terreus* mycelia as described previously (Dwiarti et al. 2002). The N-terminal and internal amino acid sequences of the protein were identified by Shimadzu (Kyoto, Japan). A draft *A. terreus* genome sequence and predicted protein data translated from the genome were obtained from the *A. terreus* genome database provided by the Broad Institute (<http://www.broad.mit.edu>). The protein data from the database were searched for a translated amino acid sequence datum that contained the analyzed amino acid sequences. To amplify a DNA fragment containing the positive locus from the *A. terreus* genome, oligo DNA primers targeted on 521 bp upstream (AT09971-F) and 213 bp downstream (AT09971-R) of the open reading frame (ORF) were synthesized (Table 1). Polymerase chain reaction (PCR) was done by KOD-plus-DNA polymerase (Toyobo, Osaka, Japan) under the following conditions: 95°C for 2 min, 30 cycles of 95°C for 30 s, 55°C for 40 s, 68°C for 2 min 20 s, and a final extension at 68°C for 5 min. The fragments amplified from

Table 1 Primers used in this study

Name	Sequence (5' to 3')
AT09971-F	CACCGATGGCTATCGCAATCGTCCCTG
AT09971-R	CTCCAGTCGCCGAATGTGGGAGTACGTG
Exon1-F	TCATAGGTACCATGACCAAGCAATCTGCG GACAGCAAC
Exon1-R	TGCTGCAACAGGCCCCAGTTTCTGCCATA TCCAATCACCCTGCAGGCT
Exon2-F	AAACTGGGGCCTGTTCAGCAGCCATG
Exon2-R	TGATCTCGAGTTATACCAGTGGCGATTTCAC GGGACAATTAAGCAG
Q-CAD1-F	CAGCCATGACCAATTCCGCTTTCA
Q-CAD1-R	AAGACCTCACTTGCTGCAAAGACC
Q-ACT-F	TCGTGACTTGACCGACTACCTCAT
Q-ACT-R	TGATGTCACGGACGATTCACGCT

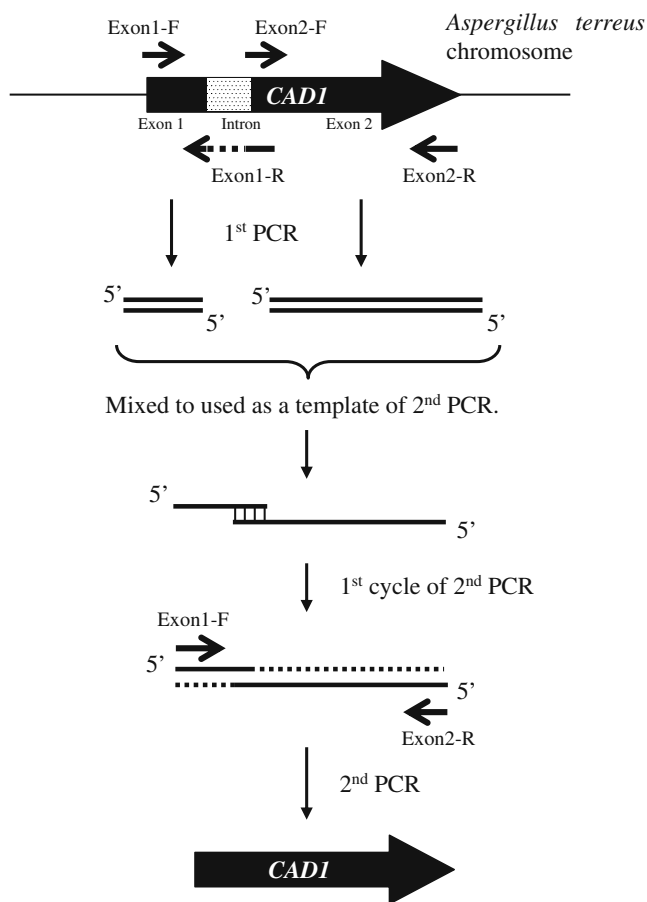
genomic DNAs of wild-type and the itaconic-overproducing strains were ligated into the *EcoRV* site in pT7Blue and were named pwCAD1 and pmCAD1, respectively. DNA sequencing analysis of these plasmids was performed by Operon Biotechnologies (Tokyo, Japan).

Construction of the CAD1 expression plasmid for *S. cerevisiae*

An intron that existed in the genomic sequence in *CAD1* was removed by two steps of PCR to express the gene in *S. cerevisiae* (Fig. 1). Information on the intron position was obtained from the *A. terreus* database. A primer set of Exon1-F and Exon1-R and a set of Exon2-F and Exon2-R were used to amplify the first and the second exons of *CAD1*, respectively (Table 1). Each first PCR was performed by KOD-plus-DNA polymerase under the following conditions: 95°C for 2 min, 30 cycles of 95°C for 30 s, 55°C for 40 s, 68°C for 2 min, and a final extension at 68°C for 5 min. The amplified fragments were mixed together and used as a template for the second PCR. The Exon1-F and Exon2-R primers were used for the second PCR, and the reaction was performed by KOD-plus-DNA polymerase under the following conditions: 95°C for 2 min, 30 cycles of 95°C for 30 s, 55°C for 40 s, 68°C for 2 min, and a final extension at 68°C for 5 min. The amplified fragment was ligated into the *EcoRV* site in pT7Blue, and the nucleotide sequence was confirmed using a DTCS Quick Start Master mix (Beckman Coulter, Fullerton, CA, USA) with M13-Forward and M13-Reverse primers (Invitrogen) according to the manufacturer's protocol. The prepared sequencing samples were analyzed using a capillary DNA sequencer CEQ2000 (Beckman Coulter). The resulting plasmid and pAUR123 were digested with *KpnI* and *XhoI* and ligated to obtain a plasmid named pAUR-CAD1.

Functional expression of CAD1 in *S. cerevisiae*

The plasmid pAUR-CAD1 was introduced into *S. cerevisiae* INVSc1 as reported previously (Elble 1992). Transformants were screened on a YPD medium plate containing Aureobasidin A, and grown colonies were analyzed by colony-direct PCR using primers Exon1-F and Exon2-R. The reaction was performed by *TaKaRa Ex Taq* (Takara Bio) under the following conditions: 95°C for 2 min, 35 cycles of 95°C for 20 s, 55°C for 30 s, and 72°C for 2 min. A positive transformant was used for further analysis. A transformant that introduced pAUR123 was also prepared as a control strain. These transformants were inoculated into 500-ml flasks containing 50 ml of YPD medium supplemented with Aureobasidin A and cultured at 30°C on a rotary shaker at 200 rpm for 3 days. Grown cells were harvested by centrifugation and homogenized by an ultrasonic processor (Sonics & Materials, Newtown, CT, USA) to prepare crude



The nucleotide sequence without the intron.

Fig. 1 Construction of *CAD1* excluding its intron. Two PCR steps were used to prepare the *CAD1* nucleotide sequence without its intron, as described in “Materials and methods.” The resulting nucleotide sequence was used to construct the yeast expression plasmid pAUR-CAD1

enzyme solutions. The CAD activities of the transformants were analyzed as described above.

Estimation of the mRNA level of CAD1

The wild-type strain and the high itaconic acid producer strain of *A. terreus* were grown in the production medium with and without itaconic acid. After the incubation, cells were harvested on filter paper, and total ribonucleic acid (RNA) was extracted and purified using BioRobot EZ1 (Qiagen) and EZ1 RNA Tissue Mini Kit (Qiagen). A first-strand complementary DNA (cDNA) synthesis from the total RNA and estimation of the specific cDNA were performed using a real-time quantitative reverse transcription PCR (QRT-PCR) Mx3000P system (Stratagene, La Jolla, CA, USA) and FullVelocity SYBR Green QRT-PCR Master mix (Stratagene). The amount of the transcript level of *CAD1* was measured using a primer set of Q-CAD1-F and Q-CAD1-R and an actin gene (*ACT1*; Table 1), respectively. *ACT1* is a constitutively expressed gene and was used as an internal standard of the transcript level. The reaction and measurements were performed according to the protocol provided by the manufacturer.

Results

Cloning and sequencing analysis of CAD1 from a wild-type strain of *A. terreus*

The N-terminal and four internal amino acid sequences of the purified protein obtained by the amino acid sequencing are shown in Table 2. We searched for a gene that encoded these amino acid sequences from the database and found a positive translated polypeptide datum that contained these five amino acid sequences. The gene encoding this polypeptide was classified as ATEG_09971 in the database and is regarded as the *CAD1* gene. A fragment including the 5'- and 3'-untranslated regions of *CAD1* of the *A. terreus* wild-type strain was amplified by PCR, and the nucleotide sequence was analyzed (Fig. 2, DNA Data Bank of Japan accession number AB326105), and it showed

95.7% nucleotide sequence identity with that of ATEG_09971 ORF (data not shown). The difference between these sequences could be due to minor differences in the sources of *A. terreus* strains. The nucleotide sequences of *CAD1* around its initiation codon, intron, and termination codon were highly identical with ATEG_09971, and we predicted a coding region of *CAD1* based on the exon of ATEG_09971. The predicted *CAD1* is interrupted by 56 bp of the intron, and the exon consists of 1,470 bp encoding a polypeptide of 490 amino acid residues (Fig. 2), whose calculated molecular mass is 52,721 Da. This result was mostly consistent with our previous report that a molecular weight of CAD purified from *A. terreus* was estimated to be 55 kDa via sodium dodecyl sulfate polyacrylamide gel electrophoresis (Dwiarti et al. 2002).

Computing analyses of the primary structure of CAD1 and upstream and downstream untranslated regions of the gene

The amino acid sequence was analyzed by the BLASTP algorithm, and we found that the protein contained a conserved domain of the MmgE/PrpD family of proteins of bacteria and fungi, including a number of 2-methylcitrate dehydratases of bacteria that are involved in propionate catabolism (data not shown). The protein showing the highest identity (53%) with CAD was an unnamed protein product of *A. oryzae* that possessed a conserved region of the PrpD family (accession no. AP007175); nevertheless, this was not recognized as a high-yield organism for itaconic acid. The homology search algorithms found many proteins that had a high identity with CAD but none whose functions had been characterized. The WoLF PSORT (Horton et al. 2006) algorithms predicted this protein would be localized in the cytoplasm, suggesting that the CAD substrate of *cis*-aconitic acid, the TCA cycle intermediate, was transported from the mitochondria to the cytoplasm in *A. terreus* cells. The DIpro algorithm (Cheng et al. 2006) predicted that eight cysteine residues were involved in forming four disulfide bonds among the 12 cysteines in the protein, and the predicted disulfide bonds ordered by probability in descending order were as follows: Cys406 and Cys483, Cys19 and Cys48, Cys201 and Cys205, and Cys337 and Cys368.

In the 5'-untranslated region of *CAD1*, no typical sequence for the TATA box was detected, while consensus-binding motifs for the HAP complex (CCAAT), a global transcription activator identified in eukaryotes including filamentous fungi (Xing et al. 1993; Kato et al. 1998; Goda et al. 2005), were found in the regions from -60 to -56 and -400 to -395 of *CAD1* (Fig. 2). In addition, the consensus-binding motifs for AREA (HGATAR, H=not G, R=A

Table 2 Amino acid sequence analysis of CAD

Peptide	Sequence
N-terminal	(A/P)KSGVTS(E/Q/L)I(L) ^a
Internal no. 1	V(I/Q)GY(G/F)QK
Internal no. 2	VPWSEK
Internal no. 3	EEFDQ(D)(G/E)(N)
Internal no. 4	V(I/K)EIE(D/V/G)LVLGL(D/V)(Y)(A)(I)

^a The amino acid residues that could not be clearly determined are given in parentheses.

Fig. 2 The nucleotide and deduced amino acid sequences of *CAD1*. The nucleotides that encode amino acids of CAD are given in *capital letters*. Initiation and termination (*asterisk*) codons are given in *bold letters*. Consensus-binding motifs for the HAP complex (CCAAT) are indicated by both *underlined* and *italic type*, and consensus-binding motifs for AREA (HGATAR, H=not G, R=A or G) in the complementary strand are indicated by *underlined type*

```

-509 caccgatggcatatcgcaatcgctccctgggtggtcttgaatcgatgccacacttgctcc
-449 ggatgaaacacattccggagcgcgcatcgatattgctacacagtagacccaatggtct
-389 gcagatgccctaaatggtagttctcactggcctgcattaagtcttggtgagatcattg
-329 tcggcctaacatcagtgtaggttacgggtgagatttacttgcataagaattccagacc
-269 acaaggttctagatcctttgacggcggactcccctcagaggtgccggggcgcgagctgtgc
-209 gttgctccgggatttgtaggacgcagctcggatacctagccggttatgggaatcggaggtt
-149 gtgacagcgtaaacacatggatagttaaatcggatgtacaccactggtggaatga
-89 cgggggctacaacacagagattatctgatccaattctggttcggtggcattctatcattc
-29 gcagcgaaaattgtctctattaaattgaccATGACCAAACAATCGCGGACGACCAACGCAA
      M T K Q S A D S N A K 11
32 AGTCAGGAGTTACGTCGCAAAATATGTCAATGGGCATCCAACCTGGCCACTGACGACATCC
   S G V T S E I C H W A S N L A T D D I P 31
92 CTTCCGACGTATTAGAAAGAGCAAAAATACCTTATTCTCGACGGTATTGCATGTGCCI TGGG
   S D V L E R A K Y L I L D G I A C A W V 51
152 TTGGTGAAGAGTGCCTTGGTCAGAGAAGTATGTTTCAGGCAACGATGAGCTTTGAGCCGC
   G A R V P W S E K Y V Q A T M S F E P P 71
212 CGGGGGCTCGAGGTTGATTTGGATATGGACAGgtaaattttattcactctagacgggtcca
   G A C R V I G Y G Q (Intron) 81
272 caaagtatactgacgatccttcgtatagAAACTGGGGCCTGTTCAGCAGCCATGACCAA
      K L G P V A A A M T N 92
332 TTCCGCTTTCATACAGGCTACGGAGCTTTCAGCACTTTCAGCACTTTCAGCACTTTCAGCACTC
   S A F I Q A T E L D D Y H S E A P L H S 112
392 TGCAAGCATTTCTTCTTCTGCGGTCTTTGACGCAAGTGAAGTCTTAGCCGAGCAGGGCAA
   A S I V L P A V F A A S E V L A E Q G K 132
452 AACAAATTCGGTATAGATGTTATTCTAGCCGCCATTGTGGGGTTTGAATCTGGCCACG
   T I S G I D V I L A A I V G F E S G P R 152
512 GATCGGCAAAGCAATCTACGGATCGGACCTCTTGAACAACGGCTGGCATTTGTGGAGCTGT
   I G K A I Y G S D L L N N G W H C G A V 172
572 GTATGGCCTCCAGCCGGTGCCTGGCCACAGGAAGCTCCTCGGTCTAATCTCAGACTC
   Y G A P A G A L A T G K L L G L T P D S 192
632 CATGGAAGATGCTCTCGGAATTGCGTGCACGCAAGCCTGTGGTTTAAATGTCGGCGCAATA
   M E D A L G I A C T Q A C G L M S A Q Y 212
692 CCGAGGCATGGTAAAGCGTGTGCAACACGGATTTCGACGCGTAATGGTCTTCTTGGGGG
   G G M V K R V Q H G F A A R N G L L G G 232
752 ACTGTTGGCCATGGTGGTACGAGGCAATGAAAGGTGTCCTGGAGAGATCTTACGGCGG
   L L A H G G Y E A M K G V L E R S Y G G 252
812 TTTCTCAAGATGTTCCCAAGGGCAACGGCAGAGCTCCCTACAAGAGGAGGAAGT
   F L K M F T K G N G R E P P Y K E E E V 272
872 GGTGGCTGGTCTCGGTTCAATCTGGCATACTTTACTATTTCGCATCAAGCTCTATGCGCTG
   V A G L G S F W H T F T I R I K L Y A C 292
932 CTGCGGACTTGTCCATGGTCCAGTTCGAGGCTATCGAAAACCTTCAGGGAGATACCCCGA
   C G L V H G P V E A I E N L Q G R Y P E 312
992 GCTCTGAATAGAGCCAACCTCAGCAACATTTCGCATGTTTATGTACAGCTTTCAACGGC
   L L N R A N L S N I R H V H V Q L S T A 332
1052 CTCGAACAGTCACTGTGGATGGATACCAGAGGAGAGACCCATCAGTTCAATCGCAGGGCA
   S N S H C G W I P E E R P I S S I A G Q 352
1112 GATGAGTGTGCATACATTCTCGCCGTCAGCTGGTCGACCAGCAATGTTTTCSCCA
   M S V A Y I L A V Q L V D Q Q C L L S Q 372
1172 GTTTTCTGAGTTTGATGACAACCTGGAGAGGCCAGAAGTTTGGGATCTGGCCAGGAAGGT
   F S E F D D N L E R P E V W D L A R K V 392
1232 TACTTCACTCAAAGCGAAGATTGATCAAGACGGCAACTGCTCAGTCCGGGTCCGGT
   T S S Q S E E F D Q D G N C L S A G R V 412
1292 GAGGATTGAGTCAACGATGGTTCTTATTACGGAAAGTGTGAGAAGCCTCTTGGTGT
   R I E F N D G S S I T E S V E K P L G V 432
1352 CAAAGAGCCATGCAAAACGACGGATTCTCCACAAAATACCGAACCCCTTCTGGTAGCGT
   K E P M P N E R I L H K Y R T L A G S V 452
1412 GACGGACGAAATCCCGGGTAAAAGAGATTGAGGATCTTGTCTCGGCCTGGACAGGCTCAC
   T D E S R V K E I E D L V L G L D R L T 472
1472 CGACATTAGCCATTTGCTGGAGCTGCTGAATTGCCCGTGAAATCGCACTGGTAAat
   D I S P L L E L L N C P V K S P L V * 490
1532 ggggaagcagatggaacatcttcatgtcacgggcacaatcttaggtcatatcgtaactg
    
```

or G), a global transcription repressor involved in nitrogen metabolite repression in *Aspergillus nidulans* (Kudla et al. 1990), were found in the regions from -500 to -495, -69 to -64, and -39 to -34 in the complementary strand (Fig. 2). In contrast, no polyadenylation signal typical of eukaryotes could be found in the 3'-noncoding region of *CAD1*.

Enzyme activity of the CAD1 product

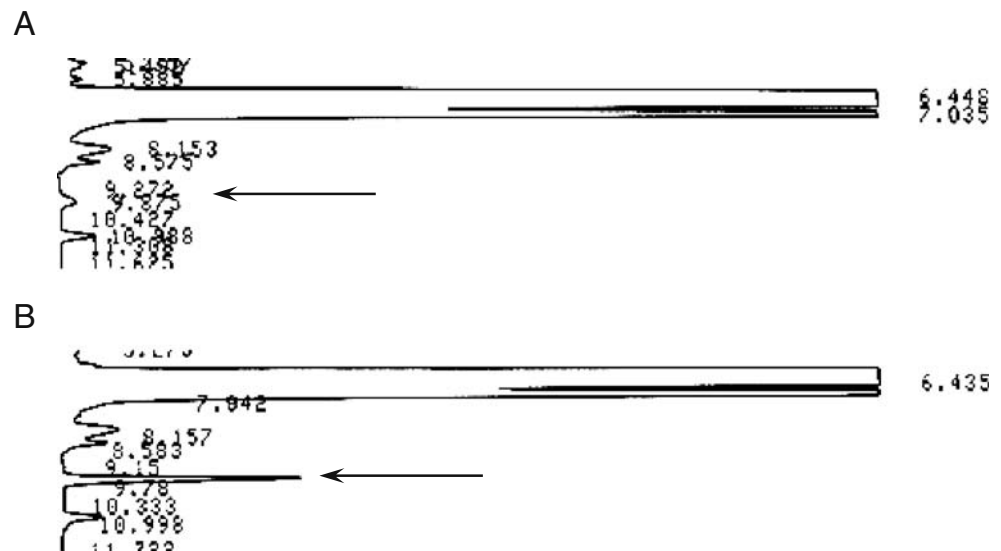
The transformation system for *A. terreus* has not been established; thus, we confirmed the function of the CAD protein by its gene expression in yeast. To investigate the actual activity of the CAD protein, the *CAD1* gene excluding an intron as shown in Fig. 1 controlled by the *ADHI* promoter was expressed in *S. cerevisiae*, and the CAD activity of the protein was measured. The plasmids

pAUR-CAD1 and pAUR123 were introduced into *S. cerevisiae*, and the total soluble proteins of these cells were extracted. Each CAD activity was measured by directly detecting itaconic acid as a product from *cis*-aconitic acid as a substrate. As expected, we found that the strain introduced by pAUR-CAD1 possessed 276.1±30.1 mU/mg protein of enzyme activity, while the strain introduced by pAUR123, as a control, showed little activity (37.9±18.4 mU/mg protein; Fig. 3). These results confirm that the *CAD1* gene encodes the CAD enzyme of *A. terreus*.

Regulation of CAD1 in *A. terreus*

The inhibitory effect of itaconic acid on its production by *A. terreus* was reported by Lockwood and Reeves (1945). To clarify whether feedback inhibition by itaconic acid existed at the transcription level of *CAD1*, a transcription analysis

Fig. 3 Recombinant CAD activity detected by HPLC analysis. Activities of CAD from the introduced yeast strains pAUR123 (a) and pAUR-CAD1 (b) were analyzed by HPLC as described in “Materials and methods.” Peaks for itaconic acid, a product of the enzyme, are indicated by arrows



of the gene was performed. The *A. terreus* wild-type strain was grown in media with and without 10% (w/v) itaconic acid, and the amounts of *CAD1* transcripts were measured by QRT-PCR. The transcripts of *CAD1* that were prepared from the mycelia grown in each medium were present in similar quantities (Fig. 4). This fact suggests that the transcription of *CAD1* is not affected by the presence of itaconic acid.

Structure and regulation of *CAD1* in the high-itaconic acid-producing strain

A fragment including the 5'- and 3'-untranslated regions of *CAD1* of the high-itaconic acid-producing strain was amplified, and the nucleotide sequence was analyzed using the same procedure as for the *A. terreus* wild-type strain, as mentioned above. A comparison of the nucleotide sequences from the wild-type strain and the high-producing strain showed they were exactly the same (data not shown).

A transcription assay of *CAD1* in the high-producing strain indicated that there was no effect of itaconic acid on the transcription of *CAD1*, as seen in the wild-type strain. Moreover, the transcription of *CAD1* was approximately fivefold stronger in the high-producing strain compared to that in the wild-type strain (Fig. 4). These facts suggest that high itaconic acid productivity was not caused by the substitution of the amino acid sequence of CAD but by higher expression levels of *CAD1* in the high-producing strain compared to the wild-type strain.

Discussion

In this study, we succeeded in cloning *CAD1* and confirming the *CAD1* gene function for the first time and enhanced

our understanding of the roles of the CAD protein in *A. terreus*. Generally, CAD is an enzyme of low stability so that it is difficult to obtain enough from the amino acid sequence analysis to perform gene cloning. In this study, we prepared a sufficient amount of purified CAD to analyze its amino acid sequence and succeeded in gene cloning and measurement of its functional expression.

It is interesting that the primary structure of CAD shared high identity with proteins possessing conserved regions of the MmgE/PrpD family, although the catalytic function of 2-methylcitrate dehydratase belonging to this family is different from that of CAD (Brock et al. 2002). This fact would be due to amino acid residue(s) in CAD being

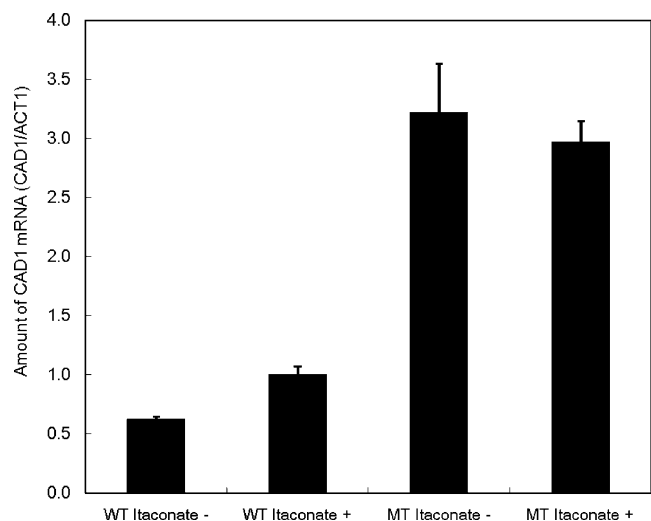


Fig. 4 Transcription of *CAD1* in *A. terreus* wild-type and the high-itaconic acid-producing strains. The amounts of transcripts of *CAD1* in the wild-type strain (*WT*) and the high-itaconic acid-producing strain (*MT*) cultured in media with (*Itaconate* +) or without (*Itaconate* -) itaconic acid were measured using QRT-PCR as described in “Materials and methods”

different from 2-methylcitrate dehydratase. This result also suggests that some uncharacterized proteins that have been classified into the MmgE/PrpD family according to homology might have CAD activity. Previously, CAD was found to be inactivated by reagents acting on cysteine residues including 5,5'-dithio-bis(2-nitrobenzoic acid), which suggested that the enzyme had cysteine residue(s) near the active site or that the residue(s) played an important role in the protein structure (Dwiarti et al. 2002). Computing the analysis on the deduced amino acid sequence of CAD indicated that 8 of the 12 cysteines were predicted to form four S–S bonds, which would be essential for stabilizing the protein structure. As far as the localization of CAD is concerned, there has been some discussion as to whether it exists in the mitochondria or the cytoplasm because *cis*-aconitic acid, the substrate of CAD, is produced in the TCA cycle, but itaconic acid is finally secreted into the cultured broth. The WoLF PSORT algorithm predicted that this protein was localized in the cytoplasm, and therefore it would be predicted that *cis*-aconitic acid was transported from the mitochondria to the cytoplasm in *A. terreus*. The existence of an element in the 5'-flanking region of *CAD1*, the consensus-binding sequences of the HAP complex (CCAAT; Xing et al. 1993; Kato et al. 1998; Goda et al. 2005), suggests that *CAD1* is a highly transcribed gene.

In the industrial production of itaconic acid via *A. terreus*, its synthesis is inhibited by the itaconic acid produced. However, our findings clarified that the presence of itaconic acid was not the cause of gene repression.

The transcription of *CAD1* in the high-itaconic acid producer was higher than that in the wild-type strain, while the alteration of nucleotide sequences including the 5'- and 3'-untranslated regions did not occur in the high-itaconic acid producer, suggesting that the higher itaconic acid production by this strain may be caused by higher expression levels of the gene. The reason for this phenomenon cannot be determined at present, and further analyses are needed, although the release from the repressor of the gene or unusual activation of the transcription factor is assumed.

In this study, we cloned and characterized the gene coding for CAD of *A. terreus*, which could be useful for molecular breeding of an itaconic acid-hyperproducing organism.

Acknowledgment We acknowledge the financial support by the comprehensive support program for the creation of regional innovation in the Japanese Science and Technology Agency (JST).

References

- Bentley R, Thiessen CP (1957a) Biosynthesis of itaconic acid in *Aspergillus terreus*. I. Tracer studies with ¹⁴C-labeled substrates. *J Biol Chem* 226:673–687
- Bentley R, Thiessen CP (1957b) Biosynthesis of itaconic acid in *Aspergillus terreus*. II. Early stages in glucose dissimilation and the role citrate. *J Biol Chem* 226:689–701
- Bentley R, Thiessen CP (1957c) Biosynthesis of itaconic acid in *Aspergillus terreus*. III. The properties and reaction mechanism of *cis*-aconitic acid decarboxylase. *J Biol Chem* 226:703–720
- Bradford MM (1976) A rapid and sensitive method for the quantitation of microgram quantities of protein utilizing the principle of protein–dye binding. *Anal Biochem* 72:248–254
- Brock M, Maerker C, Schutz A, Volker U, Buckel W (2002) Oxidation of propionate to pyruvate in *Escherichia coli*. Involvement of methylcitrate dehydratase and aconitase. *Eur J Biochem* 269:6184–6194
- Calam CT, Oxford AE, Raistrick H (1939) Studies in the biochemistry of microorganisms. LXIII. Itaconic acid, a metabolic product of a strain of *Aspergillus terreus* Thom. *J Biochem* 33:1488–1495
- Cheng J, Saigo H, Baldi P (2006) Large-scale prediction of disulphide bridges using Kernel methods. Two-dimensional recursive neural networks, and weighted graph matching. *Proteins* 62:617–629
- Dwiarti L, Yamane K, Yamatani H, Kahar P, Okabe M (2002) Purification and characterization of *cis*-aconitic acid decarboxylase from *Aspergillus terreus* TN484-M1. *J Biosci Bioeng* 94:29–33
- Elble R (1992) A simple and efficient procedure for transformation of yeasts. *Biotechniques* 13:18–20
- Goda H, Nagase T, Tanoue S, Sugiyama J, Steidl S, Tuncher A, Kobayashi T, Tsukagoshi N, Brakhage AA, Kato M (2005) Nuclear translocation of the heterotrimeric CCAAT binding factor of *Aspergillus oryzae* is dependent on two redundant localising signals in a single subunit. *Arch Microbiol* 184:93–100
- Horton P, Park K, Obayashi T, Nakai K (2006) Protein Subcellular Localization Prediction with WoLF PSORT. In: Proceedings of the 4th Annual Asia Pacific Bioinformatics Conference APBC06, Taipei, Taiwan, pp 39–48
- Kato M, Aoyama A, Naruse F, Tateyama Y, Hayashi K, Miyazaki M, Papagiannopoulos P, Davis MA, Hynes MJ, Kobayashi T, Tsukagoshi N (1998) The *Aspergillus nidulans* CCAAT-binding factor AnCP/AnCF is a heteromeric protein analogous to the HAP complex of *Saccharomyces cerevisiae*. *Mol Gen Genet* 257:404–411
- Kinoshita S (1931) Über die produktion von itaconsäure und mannit durch einen neuen schimmelpilz. *A itaconicus*. *Acta Phytochim Jpn* 5:271–287
- Kudla B, Caddick MX, Langdon T, Martinez-Rossi N, Bennett, CF, Sibley, S, Davies, RW, Arst HN (1990) The regulatory gene *areA* mediating nitrogen metabolite repression in *Aspergillus nidulans*. Mutations affecting specificity of gene activation alter a loop residue of a putative zinc finger. *EMBO J* 9:1355–1364
- Lockwood LB, Reeves MD (1945) Some factors affecting the production of itaconic acid by *Aspergillus terreus*. *Arch Biochem* 6:455–469
- Xing Y, Fikes JD, Guarente L (1993) Mutations in yeast HAP2/HAP3 define a hybrid CCAAT box binding domain. *EMBO J* 12:4647–4655
- Yahiro K, Takahama T, Park EY, Okabe M (1995) Breeding of *Aspergillus terreus* mutant TN-484 for itaconic acid production with high yield. *J Ferment Bioeng* 79:506–508