

# Cloning, Overexpression, and Mutagenesis of the *Sporobolomyces salmonicolor* AKU4429 Gene Encoding a New Aldehyde Reductase, Which Catalyzes the Stereoselective Reduction of Ethyl 4-Chloro-3-Oxobutanoate to Ethyl (*S*)-4-Chloro-3-Hydroxybutanoate

KEIKO KITA,<sup>1\*</sup> TAKANOBU FUKURA,<sup>1</sup> KOH-ICHI NAKASE,<sup>1</sup> KENJI OKAMOTO,<sup>1</sup> HIDESHI YANASE,<sup>1</sup> MICHIIHIKO KATAOKA,<sup>2</sup> AND SAKAYU SHIMIZU<sup>2</sup>

*Department of Biotechnology, Tottori University, 4-101 Koyama, Tottori 680-8552,<sup>1</sup> and Division of Applied Life Sciences, Graduate School of Agriculture, Kyoto University, Sakyo-ku, Kyoto 606-8502,<sup>2</sup> Japan*

Received 19 July 1999/Accepted 6 September 1999

We cloned and sequenced the gene encoding an NADPH-dependent aldehyde reductase (ARII) in *Sporobolomyces salmonicolor* AKU4429, which reduces ethyl 4-chloro-3-oxobutanoate (4-COBE) to ethyl (*S*)-4-chloro-3-hydroxybutanoate. The *ARII* gene is 1,032 bp long, is interrupted by four introns, and encodes a 37,315-Da polypeptide. The deduced amino acid sequence exhibited significant levels of similarity to the amino acid sequences of members of the mammalian 3 $\beta$ -hydroxysteroid dehydrogenase–plant dihydroflavonol 4-reductase superfamily but not to the amino acid sequences of members of the aldo-keto reductase superfamily or to the amino acid sequence of an aldehyde reductase previously isolated from the same organism (K. Kita, K. Matsuzaki, T. Hashimoto, H. Yanase, N. Kato, M. C.-M. Chung, M. Kataoka, and S. Shimizu, *Appl. Environ. Microbiol.* 62:2303–2310, 1996). The ARII protein was overproduced in *Escherichia coli* about 2,000-fold compared to the production in the original yeast cells. The enzyme expressed in *E. coli* was purified to homogeneity and had the same catalytic properties as ARII purified from *S. salmonicolor*. To examine the contribution of the dinucleotide-binding motif G<sub>19</sub>-X-X-G<sub>22</sub>-X-X-A<sub>25</sub>, which is located in the N-terminal region, during ARII catalysis, we replaced three amino acid residues in the motif and purified the resulting mutant enzymes. Substrate inhibition of the G<sub>19</sub>→A and G<sub>22</sub>→A mutant enzymes by 4-COBE did not occur. The A<sub>25</sub>→G mutant enzyme could reduce 4-COBE when NADPH was replaced by an equimolar concentration of NADH.

Aldehyde reductase (EC 1.1.1.2), aldose reductase (EC 1.1.1.21), and carbonyl reductase (EC 1.1.1.184) catalyze NADPH-dependent reduction of a variety of carbonyl compounds and are widely distributed in mammalian and plant tissues. These enzymes are members of the aldo-keto reductase superfamily (4, 8); however, their physiological functions are not well understood. The amino acid sequences of aldose reductases and aldehyde reductases exhibit significant levels of similarity, but the amino acid sequences of carbonyl reductases do not (32).

In previous papers, we described purification and characterization of three NADPH-dependent aldehyde reductases (ARI, ARII, and ARIII) of the red yeast *Sporobolomyces salmonicolor* AKU4429 (9, 14, 34). ARI is the most abundant aldehyde reductase in this yeast and catalyzes asymmetric reduction of ethyl 4-chloro-3-oxobutanoate (4-COBE) to ethyl (*R*)-4-chloro-3-hydroxybutanoate (4-CHBE) {enantiomeric excess for (*R*) = [(*R* - *S*)/(*R* + *S*)] × 100 and vice versa}, a promising chiral building block for organic synthesis. In contrast, ARII is produced in considerably smaller amounts but reduces 4-COBE to the (*S*) enantiomer (92.7% enantiomeric excess), which is also a useful chiral building block for chemical synthesis of pharmaceuticals. In addition to the stereoselectivity of activity against 4-COBE, the N-terminal amino acid

sequences of these two aldehyde reductases are quite different. Based on the amino acid sequence deduced from the cDNA sequence, ARI belongs to the aldo-keto reductase superfamily (13). Recently, an NADPH-dependent aldehyde reductase (S1), which reduces 4-COBE to the (*S*) enantiomer (100% enantiomeric excess), was purified from *Candida magnoliae* AKU4643 (31). The substrate specificities, subunit structures, and N-terminal amino acid sequences of ARII and S1 are not similar. This indicates that the two enzymes belong to the different groups.

In this study, we cloned and analyzed a cDNA clone of the aldehyde reductase gene (*ARII*) in order to compare the catalytic mechanisms of ARI and ARII and to understand the molecular basis of the stereospecific reduction of 4-COBE.

## MATERIALS AND METHODS

**Microorganisms and culture conditions.** *S. salmonicolor* AKU4429 was used as the DNA donor. This organism was cultivated at 30°C in YPG medium containing 5% glucose, 1% peptone, and 1% yeast extract. *Escherichia coli* JM109 [*recA supE endA hsdR gyrA relA thi* Δ(*lac-proAB*)F' (*traD proAB<sup>+</sup> lacI<sup>q</sup> lacZΔM15*)], *E. coli* LE392 [*supE supF hsdR galK galT metB trpR lacY*], and *E. coli* MV1184 [*ara* Δ(*lac-proAB*) *rpsL thi* (Δ80/*lacZΔM15*) Δ(*srl-recA*)306::Tn10 (*ter<sup>r</sup>*)F' (*traD proAB<sup>+</sup> lacI<sup>q</sup> lacZΔM15*)] were used as host strains. *E. coli* cells were grown at 37°C in Luria-Bertani medium containing 1% Bacto-Tryptone (Difco Laboratories, Detroit, Mich.), 0.5% Bacto-Yeast Extract (Difco Laboratories), and 1% NaCl (pH 7.0). When necessary, ampicillin (100 μg/ml), streptomycin (30 μg/ml), and tetracycline (10 μg/ml) were added to the medium.

**Enzymes and chemicals.** Restriction enzymes were purchased from Takara Shuzo Co., Ltd. (Kyoto, Japan) and Toyobo (Osaka, Japan). *ExTaq* DNA polymerase and *LA Taq* DNA polymerase were purchased from Takara Shuzo Co., Ltd.

**Amino acid sequencing.** ARII was purified from cells of *S. salmonicolor* as described elsewhere (14) and was incubated with lysyl endopeptidase (Wako

\* Corresponding author. Mailing address: Department of Biotechnology, Tottori University, 4-101 Koyama, Tottori 680-8552, Japan. Phone: 81-857-31-5277. Fax: 81-857-31-0881. E-mail: kita@bio.tottori-u.ac.jp.

Pure Chemicals, Osaka, Japan) in 50 mM Tris-HCl (pH 9.0) containing 3 M urea for 12 h at 37°C at a substrate/enzyme molar ratio of 200:1. The resulting peptides were separated by high-performance liquid chromatography by using a Cosmosil 5C<sub>18</sub>-P column (4.6 by 150 mm; Nacalai Tesque, Kyoto, Japan) that previously had been equilibrated in 0.05% trifluoroacetic acid and was eluted with a linear acetonitrile gradient (0 to 100% acetonitrile in trifluoroacetic acid) at a flow rate of 0.6 ml/min. The amino acid sequences of peptides were determined with a model PPSQ-10 protein sequencer (Shimadzu, Kyoto, Japan). The phenylthiohydantoin amino acid derivatives were separated and identified with an on-line phenylthiohydantoin analyzer (model C-R7A; Shimadzu) as recommended in the instruction manual.

**Screening of a genomic DNA library.** To amplify an ARII DNA fragment from *S. salmonicolor* chromosomal DNA by PCR, upstream and downstream primers were designed on the basis of the N terminus (14) and the internal amino acid sequence LYS, respectively. The sequences of the primers used were as follows: primer N1, 5'-GCIAA(A/G)AT(A/C/T)GA(C/T)AA(C/T)GCIGTI(C/T)T-3'; and primer LYS, 5'-TT(A/C)TC(A/G/T)ATICA(C/T)TCIA(A/G)(A/G)TTC CA-3'. Chromosomal DNA extracted from *S. salmonicolor* as described previously (13) was used as a template for amplification. The PCR mixture (100 µl) contained 50 pmol of each primer, each deoxynucleoside triphosphate (dNTP) at a concentration of 200 µM, 10 mM Tris-HCl (pH 8.3) (at 25°C), 1.5 mM MgCl<sub>2</sub>, 0.01% (wt/vol) gelatin, 500 ng of template DNA, and 10 U of *ExTaq* DNA polymerase. The reaction mixture was overlaid with mineral oil, and the reaction was carried out by using a Perkin-Elmer Cetus thermal cycler. The initial template denaturation step consisted of 2 min at 94°C. The amplification profile (1 min at 47°C, 1 min at 72°C, and 1 min at 94°C) was repeated for 35 cycles. The PCR product was purified and cloned into pGEM-T (Promega, Madison, Wis.).

The PCR product derived from the genomic DNA was then labeled by using a DIG DNA labeling kit (Boehringer Mannheim), and the resulting preparation was used to screen an *S. salmonicolor* genomic DNA library (24,000 plaques) (13).

**Cloning of ARII cDNA.** First-strand cDNA was synthesized from mRNA isolated from *S. salmonicolor* (28) by using random hexamers and a GeneAmp RNA PCR kit (Perkin-Elmer). To amplify ARII cDNA, two primers, primer N2 (5'-ATGGCCAAAATCGACAACGCTGTG-3') and primer C1 (5'-GGTTTCG GAGCCGACGAGGTC-3'), were synthesized based on the genomic DNA sequence. The PCR mixture (100 µl) contained 20 pmol of each primer, each dNTP at a concentration of 200 µM, 10 mM Tris-HCl (pH 8.3) (at 25°C), 1.5 mM MgCl<sub>2</sub>, 0.01% (wt/vol) gelatin, first-strand cDNA, and 2.5 U of *ExTaq* DNA polymerase. The reaction mixture was overlaid with mineral oil, and the reaction was carried out by using a Perkin-Elmer Cetus thermal cycler. The initial template denaturation step consisted of 2 min at 94°C. The amplification profile (1 min at 65°C, 1 min at 72°C, and 1 min at 94°C) was repeated for 35 cycles. The PCR product was purified and cloned into pGEM-T to obtain pGEM-AR2.

**Subcloning and DNA sequencing.** Phage particles were purified with LambdaSorb phage adsorbent (Promega), and DNA was extracted with phenol-chloroform. The subfragments generated were cloned into pUC118 and pUC119 to provide templates. DNA sequencing was performed by the dideoxy chain termination method (20, 24) by using an automated DNA sequencer (model 373A; Applied Biosystems). The sequencing reaction was carried out as recommended in the manuals for *Taq* dye terminator cycle sequencing kits (Applied Biosystems). The sequences were deduced from the data obtained for both strands.

**Expression of ARII in *E. coli*.** A PCR was used to subclone the *ARII* gene. The two synthetic primers used for this procedure were primer N3 (5'-CCGGA ATTCATAGGAGGGGATTATGGCCAAAATCGAC-3'), which contained a Shine-Dalgarno sequence (underlined nucleotides) (29) flanked by an *EcoRI* site, and primer C2 (5'-CGGAAGCTTTATCAAGCGGTTTCGGAGCCGAC GAGG-3'), which was flanked by a *HindIII* site. Plasmid pGEM-AR2 was used as the template. The PCR mixture (100 µl) contained 20 pmol of each primer, each dNTP at a concentration of 200 µM, 10 mM Tris-HCl (pH 8.3) (at 25°C), 1.5 mM MgCl<sub>2</sub>, 0.01% (wt/vol) gelatin, 1 ng of pGEM-AR2, and 5 U of *ExTaq* DNA polymerase. The initial template denaturation step consisted of 5 min at 94°C. The amplification profile (1 min at 60°C, 1 min at 72°C, and 30 s at 94°C) was repeated for 25 cycles. The PCR-generated DNA fragment was ligated into pUC118 cleaved with *EcoRI* and *HindIII* and then was transformed into *E. coli* JM109. After ampicillin selection, several clones were picked, and the nucleotide sequence of the plasmid DNA was examined.

**Purification of ARII expressed in *E. coli*.** *E. coli* JM109 cells carrying pUCAR2 were grown in 5 liters of Luria-Bertani medium containing ampicillin and 0.1 mM isopropyl-β-D-thiogalactopyranoside (IPTG) for 9 h at 37°C. The cells (15 g, wet weight) were then harvested by centrifugation, resuspended in 100 ml of 10 mM Tris-HCl (pH 8.0)-1 mM EDTA containing 500 µg of lysozyme per ml, and left on ice for 1 h. Then the cells were disrupted by sonication, and the debris was removed by centrifugation at 12,000 × g (30 min, 4°C). The supernatant was applied to a 200-ml DEAE-Sephacel (Pharmacia Biotech Inc., Tokyo, Japan) column (2.4 by 11 cm) that previously had been equilibrated with 10 mM Tris-HCl (pH 7.4)-0.1 mM dithiothreitol (buffer A). Proteins were eluted with a 500-ml linear 0 to 0.4 M NaCl gradient in buffer A. The fractions containing more than 10% of the maximum activity were combined. After the NaCl concentration was adjusted to 4 M with solid NaCl, the enzyme solution was applied to a Phenyl-Sepharose CL-4B (Pharmacia Biotech Inc.) column (2.4 by 11 cm) equilibrated with buffer A containing 4 M NaCl. After the column was washed

with the same buffer, the enzyme was eluted by using a linear 4 to 0 M NaCl gradient and a simultaneous linear increase in the ethylene glycol concentration from 0 to 70% in buffer A. ARII was purified to electrophoretic homogeneity by this procedure.

**Site-directed mutagenesis and purification of ARII variants.** Site-directed mutagenesis of the ARII gene was carried out by using the method described by Hashimoto-Gotoh et al. (5) and a Mutan-Express Km kit (Takara Shuzo Co., Ltd.) as recommended by the manufacturer. The ARII gene was excised from pUCAR2 with *EcoRI* and *HindIII* and then ligated to pKF19k cleaved by the same restriction enzymes. The following oligonucleotides were used for mutagenic primers: 5'-AGCCGTTGGCAGCGGTGACGAC-3' (G<sub>19</sub>→A), 5'-AAG CGACGAAAGCGTTGGCGCC-3' (G<sub>22</sub>→A), and 5'-GACGTGCGAACCGA CGAAGCC-3' (A<sub>25</sub>→G) (the mismatched nucleotides are underlined). The presence of the mutations and the absence of unwanted mutations elsewhere in the gene were confirmed by sequencing both strands of the entire gene of each mutant. Each mutant gene was recovered by digestion with *EcoRI* and *HindIII* and then ligated to pUC118 cleaved with the same restriction enzymes. The plasmid was transformed into *E. coli* JM109, and then the mutant enzyme was purified from the recombinant *E. coli* cells as described above for the wild-type enzyme.

**Analytical methods.** The activity of ARII and the optical purity of 4-COBE were determined as described elsewhere (34). The reductase activity was determined at 37°C by measuring the rate of decrease in the absorbance at 340 nm. The standard reaction mixture (1.0 ml) contained 0.2 mM 4-COBE, 125 µM NADPH, and 200 mM potassium phosphate buffer (pH 7.0). One unit of enzyme activity was defined as the amount of enzyme that catalyzed oxidation of 1 µmol of NADPH per min. Protein contents were measured by a Bio-Rad protein assay (Japan Bio-Rad Laboratories, Tokyo, Japan); bovine serum albumin was used as the standard. Sodium dodecyl sulfate (SDS)-polyacrylamide gel electrophoresis (PAGE) was performed by the method of Laemmli (16). Western blotting (immunoblotting) was performed as described previously (13). The standard methods described by Sambrook et al. (23) were used for DNA manipulation.

**Nucleotide sequence accession number.** The nucleotide sequence of the ARII gene has been deposited in the GenBank database under accession no. AF160799.

## RESULTS AND DISCUSSION

**Isolation of a cDNA clone.** Purified *S. salmonicolor* ARII was proteolytically digested with lysyl endopeptidase, and six peptides were obtained by reverse-phase high-performance liquid chromatography. The amino acid sequences of these peptides were VRGTARSASK, LANLQK, DLVGSETA, QGAYDE VIK, TEA, and SWNLESIDK. To clone the *ARII* gene on the basis of these sequences, two oligonucleotide probes, primer N1 derived from the N-terminal amino acid sequence and primer LYS derived from the internal amino acid sequence (SWNLESIDK), were synthesized. PCR performed with primers N1 and LYS yielded a single product that was approximately 0.6 kb long. This PCR product was subcloned into pGEM-T, and its nucleotide sequence was determined. We concluded that the 0.6-kb fragment was a portion of the ARII gene, because the amino acid sequences of five of the peptide fragments were found in the amino acid sequence deduced from the nucleotide sequence of this PCR product. The N-terminal amino acid sequence was interrupted by an inserted unrelated DNA sequence. These results strongly suggested that introns were present in the *ARII* gene. The PCR fragment was then used to screen a genomic DNA library constructed in λEMBL3, and five positive clones were obtained. One of these clones, SAL-27, was analyzed further. The nucleotide sequence of a 2.3-kb *Sali-PstI* fragment of this clone was determined. Open reading frames in this sequence encoded all of the peptides derived from direct amino acid sequencing of the protein.

To amplify ARII cDNA by reverse transcriptase PCR, two primers, one derived from the N-terminal amino acid sequence and the other derived from the C-terminal amino acid sequence, were synthesized by using the genomic DNA sequence. A 1.0-kb fragment was synthesized and cloned into pGEM-T, and its nucleotide sequence was determined and compared to the genomic DNA sequence. The genomic DNA sequence covering the region encoding ARII was 1,375 bp long and was interrupted by four introns. This analysis revealed a

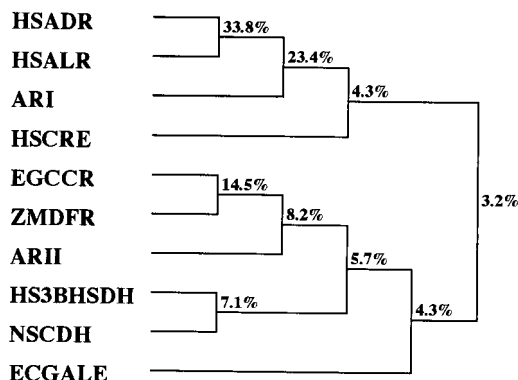


FIG. 1. Phylogenetic tree of ARII and related proteins. The tree was constructed by using the CLUSTAL V program (6). Abbreviations: HSADR, human aldose reductase (4); HSALR, human aldehyde reductase (4); ARI, *S. salmonicolor* ARI (13); HSCRE, human carbonyl reductase (32); ZMDFR, *Z. mays* dihydroflavonol 4-reductase (25); EGCCR, *E. gunnii* cinnamoyl-CoA reductase (15); ARII, *S. salmonicolor* ARII (this study); HS3BHSDH, human  $\beta$ -hydroxysteroid dehydrogenase (19); NSCDH, *Nocardia* sp. cholesterol dehydrogenase (7); ECGALE, *E. coli* UDP-galactose 4-epimerase (17).

single 1,032-bp open reading frame encoding a 344-amino-acid protein. The predicted molecular mass (37,184 Da, excluding the initial methionine) was nearly identical to the molecular mass estimated by SDS-PAGE (37,000 Da) (14).

**Comparison of the ARII amino acid sequence.** The deduced amino acid sequence of ARII was compared with other protein sequences in the GenBank database by using BLAST programs. The levels of identity with dihydroflavonol 4-reductases from higher plants (*Zea mays*, *Arabidopsis thaliana*, et al.), *Eucalyptus gunnii* cinnamyl alcohol dehydrogenase, and cinnamoyl coenzyme A (cinnamoyl-CoA) reductases were 30 to 34, 35, and 32 to 36%, respectively. Dihydroflavonol 4-reductase (EC 1.1.1.219) is required to convert dihydroflavonols to anthocyanin precursors (leucoanthocyanidins). Cinnamyl alcohol dehydrogenase (EC 1.1.1.195) catalyzes the conversion of *p*-hydroxycinnamaldehydes to the corresponding alcohols and is considered a key enzyme in lignin biosynthesis. Cinnamoyl-CoA reductase (EC 1.2.1.44) catalyzes the conversion of cinnamoyl-CoA esters to the corresponding cinnamaldehydes (i.e., the first specific step in the synthesis of lignin monomers). These enzymes are considered members of the mammalian  $\beta$ -hydroxysteroid dehydrogenase-plant dihydroflavonol 4-reductase superfamily (15). We compared the amino acid sequence of ARII with the amino acid sequences of other members of the superfamily by using CLUSTAL W programs. ARII exhibited significant levels of identity with bacterial UDP-galactose 4-epimerases (30%), *Nocardia* sp. cholesterol dehydrogenase (7) (30%), and mammalian  $\beta$ -hydroxysteroid dehydrogenases (30%). UDP-galactose 4-epimerase (EC 5.1.3.2) catalyzes the conversion of UDP-galactose to UDP-glucose through a mechanism involving transient reduction of NAD<sup>+</sup>.  $\beta$ -Hydroxysteroid dehydrogenase (EC 1.1.1.145) converts pregnenolone to progesterone, a key step in the synthesis of steroid hormones. NAD- or NADP-dependent cholesterol dehydrogenase from *Nocardia* sp. strain Ch 2-1 specifically oxidizes the  $\beta$ -OH group of cholesterol. In contrast, no similarity was found with carbonyl reductases or proteins belonging to the aldo-keto reductase superfamily, including ARI, which was isolated from *S. salmonicolor*. As shown in Fig. 1, ARII can be statistically considered a new member of the mammalian  $\beta$ -hydroxysteroid dehydrogenase-plant dihydroflavonol 4-reductase superfamily defined by Baker et al. (2) and Baker and

Blasco (1). From the standpoint of molecular biology, we found the first instance of multiple 4-COBE-reducing enzymes with various stereoselectivities occurring in the same strain but belonging to different superfamilies.

As shown in Fig. 2, the N-terminal region of ARII between position 14 and position 49 is quite similar to the analogous regions of enzymes belonging to this superfamily. Most of these enzymes require NADPH; the only exception is UDP-galactose 4-epimerase, which requires NAD(H). The conserved N-terminal region is probably involved in the cofactor binding sites of the enzymes, since the secondary structure of *E. coli* UDP-galactose 4-epimerase (3, 30) and the predicted secondary structure of *E. gunnii* cinnamoyl-CoA reductase (15) correspond to the  $\beta\alpha\beta$ -dinucleotide binding fold of NADP(H)- and NAD(H)-dependent reductases and dehydrogenases (26, 33). The roles of the conserved amino acid residues in the N-terminal region were analyzed further, as described below. The four-amino-acid motif IPKS, which is thought to be involved in the interaction with NADPH in NADPH-dependent aldo-keto reductases, was not present in ARII.

**Analysis of ARII purified from *E. coli*.** The ARII cDNA, to which a Shine-Dalgarno sequence was attached in the 5'-flanking region, was cloned into expression vector pUC118, and the resulting recombinant plasmid, pUCAR2, was transformed into *E. coli* JM109. Expression of the recombinant ARII was induced by adding IPTG; the enzyme was subsequently purified ninefold, and the yield was 61%. The purified enzyme produced a single band on SDS-PAGE gels (Fig. 3A) and on native PAGE gels (data not shown). The specific activity (15 U/mg),  $K_m$  (1.82 mM), and  $V_{max}$  (350  $\mu$ mol min<sup>-1</sup> mg<sup>-1</sup>) of the purified enzyme with 4-COBE were similar to the values obtained for the enzyme purified from *S. salmonicolor* (14). The sequence of the first 10 amino acids of the purified enzyme was determined by protein sequencing and was found to be identical to the sequence in the native enzyme. Also, the stereoselectivity [99.2% enantiomeric excess for (*S*)] for 4-COBE and the substrate specificity for typical aldehydes of the enzyme from *E. coli* were the same as those of the enzyme from *S. salmonicolor*.

We noted that the amount of ARII produced was 5.8 mg/g of *E. coli* JM109 cells; this value was 2,000 times the amount produced in yeast cells (0.003 mg/g of cells). Since *E. coli* does not exhibit a detectable level of reducing activity with 4-COBE,

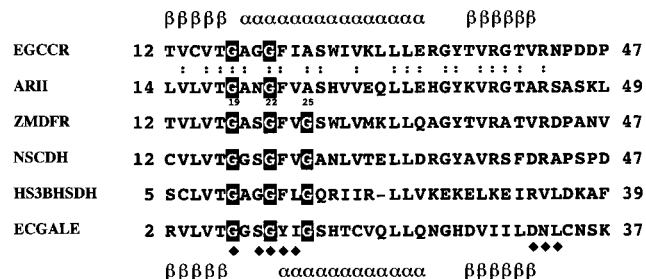


FIG. 2. Alignment of the N-terminal regions of ARII and related proteins. For an explanation of abbreviations see the legend to Fig. 1. The secondary structure of *E. coli* UDP-galactose 4-epimerase (30) and the calculated secondary structure of *E. gunnii* cinnamoyl-CoA reductase (15) are indicated below and above the amino acid sequence, respectively. The solid diamonds indicate amino acid residues which interact with NAD in *E. coli* UDP-galactose 4-epimerase. The numbers to the right and left of each sequence indicate the absolute locations within the whole sequence. The glycine residues in a putative nucleotide-binding motif are outlined in black. The positions conserved in cinnamoyl-CoA reductase and ARII are indicated by colons.  $\alpha$ ,  $\alpha$ -helix;  $\beta$ ,  $\beta$ -sheet.



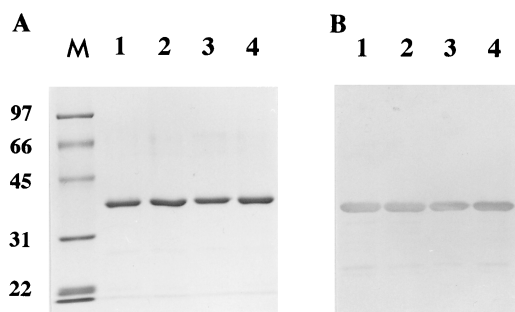


FIG. 3. SDS-PAGE of wild-type and mutant ARIIs purified from *E. coli*. Portions (1  $\mu\text{g}$ ) of wild-type (lane 1),  $G_{19}\rightarrow A$  (lane 2),  $G_{22}\rightarrow A$  (lane 3), and  $A_{25}\rightarrow G$  (lane 4) ARIIs were separated by 0.1% SDS-PAGE on a gel containing 10% polyacrylamide (A) and then analyzed by Western blotting (B) by using antibodies raised against wild-type ARII purified from *E. coli*. Lane M contained molecular mass standards, whose positions (in kilodaltons) are indicated on the left.

this system could be useful for producing large amounts of ARII. We recently developed a method for asymmetric reduction of 4-COBE to (*R*)-CHBE involving *E. coli* cells expressing the ARI gene as a catalyst (10–12). A recombinant *E. coli* strain expressing the ARII gene could also be an economical way to produce (*S*)-CHBE.

**Catalytic activity of ARII mutants.** The amino acid sequence of ARII contains the dinucleotide-binding motif  $G_{19}\text{-X-X-G}_{22}\text{-X-X-A}_{25}$  in its N-terminal region. Since  $G_7$  and  $G_{10}$  in the same motif in the UDP-galactose 4-epimerase have been found to interact with NAD (30), we assumed that  $G_{19}$  and  $G_{22}$  in the motif in ARII might be involved in binding of NADP. Position 25 is of particular interest since there is a glycine residue at this position in most sequences, whereas the cinnamoyl-CoA reductases of *E. gunnii* (15), *Saccharum officinarum* (27), *Z. mays* (22), and *Populus balsamifera* subsp. *trichocarpa* (18) and ARII have an alanine residue at position 25. To determine the contributions of  $G_{19}$ ,  $G_{22}$ , and  $A_{25}$  to the catalytic activity of ARII, we mutated these three amino acid residues and purified the mutant enzymes ( $G_{19}\rightarrow A$ ,  $G_{22}\rightarrow A$ , and  $A_{25}\rightarrow G$ ). The three mutant ARII enzymes were purified to homogeneity from *E. coli* lysates, and the yields were the same as the yield obtained for the wild-type enzyme (Fig. 3); they were then characterized by determining the NAD(P)H-dependent reductase activities. As expected, the  $G_{19}\rightarrow A$  and  $G_{22}\rightarrow A$  mutants both exhibited less than 20% of the activity of wild-type enzyme with NADPH in the presence of 0.2 mM 4-COBE, but the  $A_{25}\rightarrow G$  mutant exhibited the same level of activity as the wild-type enzyme. Initial velocity studies of 4-COBE reduction by the wild-type and mutant enzymes were performed by varying the concentration of 4-COBE and using a fixed concentration of NADPH (125  $\mu\text{M}$ ). The 4-COBE concentration was varied from 0.015 to 16 mM. The wild-type and  $A_{25}\rightarrow G$  mutant enzymes were strongly inhibited by excess concentrations of 4-COBE, but the  $G_{19}\rightarrow A$  and  $G_{22}\rightarrow A$  mutant enzymes were not (Fig. 4A). Wild-type ARII does not exhibit NADH-dependent reducing activity in the presence of 0.2 mM 4-COBE (14). We measured NADH-dependent 4-COBE reducing activity in the presence of a fixed concentration of NADH (125  $\mu\text{M}$ ). The 4-COBE concentration was varied from 0.015 to 16 mM. As shown in Fig. 4B, the wild-type and  $A_{25}\rightarrow G$  mutant enzymes exhibited significant activity with NADH in the presence of higher 4-COBE concentrations, but the  $G_{19}\rightarrow A$  and  $G_{22}\rightarrow A$  mutant enzymes did not (data not shown). The  $K_m$  for 4-COBE (2.7 mM) was equivalent to the

$K_m$  of the wild type (3.8 mM). The  $V_{\text{max}}$  ( $1.7 \mu\text{mol min}^{-1} \text{mg}^{-1}$ ) of the  $A_{25}\rightarrow G$  mutant enzyme for 4-COBE when NADH was a cofactor was three times greater than the  $V_{\text{max}}$  of the wild-type enzyme ( $0.66 \mu\text{mol min}^{-1} \text{mg}^{-1}$ ), which resulted in a 3.6-fold increase in the overall catalytic efficiency ( $k_{\text{cat}}/K_m$ ,  $21 \text{ min}^{-1} \text{mM}^{-1}$ ) compared to the catalytic efficiency of the wild-type enzyme ( $5.9 \text{ min}^{-1} \text{mM}^{-1}$ ). These results suggest that the amino acid residue at position 25 might be involved in recognition of the presence of a phosphate group at the 2' position of the adenine ribose ring in nicotinamide nucleotide coenzymes and in the interaction of 4-COBE as well; i.e., some interaction between the substrate and NADPH might occur at residue 25. Changes in coenzyme specificity from NADP dependent to NAD dependent and vice versa have been observed with many enzymes (21, 26, 35). From a practical viewpoint, it is important to alter the coenzyme specificity of ARII, which is a useful catalyst for production of (*S*)-CHBE. Changing cofactor specificity was not possible with the replacement of one amino acid residue that interacted with the coenzyme. Replacement of a few residues or a module is required because a specific conformational change is induced by the binding of cofactor and substrate. We purified ARII in amounts sufficient for crystallization. Analysis of the three-dimensional structure

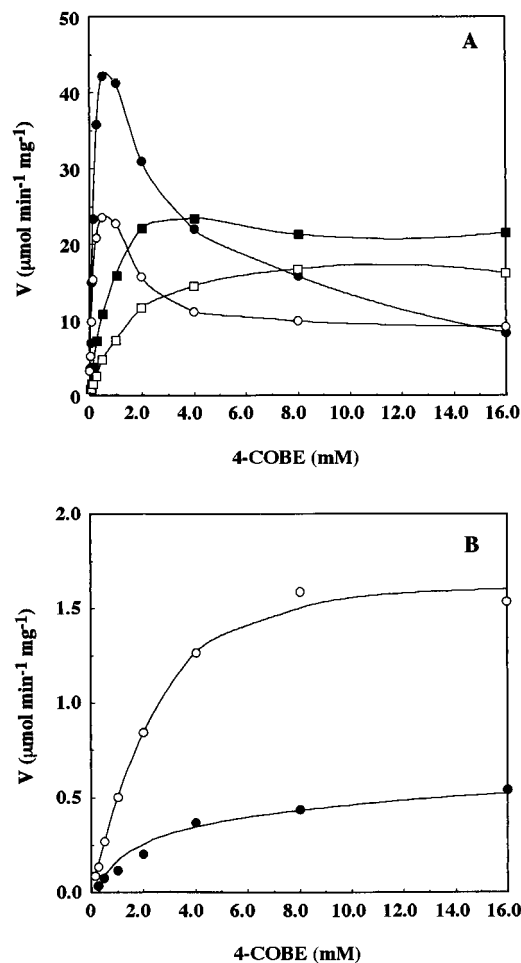


FIG. 4. Initial velocity ( $V$ ) of 4-COBE reduction by the wild-type and mutant ARIIs. Symbols: ●, wild type; ■,  $G_{19}\rightarrow A$  mutant; □,  $G_{22}\rightarrow A$  mutant; ○,  $A_{25}\rightarrow G$  mutant. (A) Experiments performed with NADPH as a cofactor. (B) Experiments performed with NADH as a cofactor.

of this enzyme should reveal a target region for replacement of amino acid residues for cofactor conversion.

#### ACKNOWLEDGMENTS

This work was supported in part by a grant-in-aid for scientific research from the Ministry of Education, Science and Culture of Japan and by grant JSPS-RFTF 97I00302 from the Research for the Future program of the Japan Society for the Promotion of Science.

#### REFERENCES

- Baker, M. E., and R. Blasco. 1992. Expansion of the mammalian  $\beta$ -hydroxysteroid dehydrogenase/plant dihydroflavonol reductase superfamily to include a bacterial cholesterol dehydrogenase, a bacterial UDP-galactose 4-epimerase, and open reading frames in vaccinia virus and fish lymphocystis disease virus. *FEBS Lett.* **301**:89–93.
- Baker, M. E., V. Luu-The, J. Simard, and F. Labri. 1990. A common ancestor for mammalian  $\beta$ -hydroxysteroid dehydrogenase and plant dihydroflavonol reductase. *Biochem. J.* **269**:558–559.
- Bauer, A. J., I. Rayment, P. A. Frey, and H. M. Holden. 1992. The molecular structure of UDP-galactose 4-epimerase from *Escherichia coli* determined at 2.5Å resolution. *Proteins* **12**:372–381.
- Bohren, K. M., B. Bullock, B. Wermuth, and K. H. Gabbay. 1989. The aldo-keto reductase superfamily. *J. Biol. Chem.* **264**:9547–9551.
- Hashimoto-Gotoh, T., T. Mizuno, Y. Ogasahara, and M. Nakagawa. 1995. An oligodeoxyribonucleotide-directed dual amber method for site-directed mutagenesis. *Gene* **152**:271–275.
- Higgins, D. G., A. J. Bleasby, and R. Fuchs. 1992. CLUSTAL V: improved software for multiple sequence alignment. *Comput. Appl. Biosci.* **8**:189–191.
- Horinouchi, S., H. Ishizuka, and T. Beppu. 1991. Cloning, nucleotide sequence, and transcriptional analysis of the NAD(P)-dependent cholesterol dehydrogenase gene from a *Nocardia* sp. and its hyperexpression in *Streptomyces* spp. *Appl. Environ. Microbiol.* **57**:1386–1393.
- Jez, J. M., M. J. Bennett, B. P. Schlegel, M. Lewis, and T. M. Penning. 1997. Comparative anatomy for the aldo-keto reductase superfamily. *Biochem. J.* **326**:625–636.
- Kataoka, M., H. Sakai, T. Morikawa, M. Katoh, T. Miyoshi, S. Shimizu, and H. Yamada. 1992. Characterization of aldehyde reductase of *Sporobolomyces salmonicolor*. *Biochim. Biophys. Acta* **1122**:57–62.
- Kataoka, M., L. P. S. Rohani, K. Yakamoto, M. Wada, H. Kawabata, K. Kita, H. Yanase, and S. Shimizu. 1997. Enzymatic production of ethyl (*R*)-4-chloro-3-hydroxybutanoate: asymmetric reduction of ethyl 4-chloro-3-oxobutanoate by an *Escherichia coli* transformant expressing the aldehyde reductase gene from yeast. *Appl. Microbiol. Biotechnol.* **48**:699–703.
- Kataoka, M., L. P. S. Rohani, M. Wada, K. Kita, H. Yanase, I. Urabe, and S. Shimizu. 1998. *Escherichia coli* transformant expressing the glucose dehydrogenase gene from *Bacillus megaterium* as a cofactor regenerator in a chiral alcohol production system. *Biosci. Biotechnol. Biochem.* **62**:167–169.
- Kataoka, M., K. Yamamoto, H. Kawabata, M. Wada, K. Kita, H. Yanase, and S. Shimizu. 1999. Stereoselective reduction of ethyl 4-chloro-3-oxobutanoate by *Escherichia coli* transformant cells coexpressing the aldehyde reductase and glucose dehydrogenase genes. *Appl. Microbiol. Biotechnol.* **51**:486–490.
- Kita, K., K. Matsuzaki, T. Hashimoto, H. Yanase, N. Kato, M. C-M. Chung, M. Kataoka, and S. Shimizu. 1996. Cloning of the aldehyde reductase gene from a red yeast, *Sporobolomyces salmonicolor*, and characterization of the gene and its product. *Appl. Environ. Microbiol.* **62**:2303–2310.
- Kita, K., K. Nakase, H. Yanase, M. Kataoka, and S. Shimizu. 1999. Purification and characterization of new aldehyde reductases from *Sporobolomyces salmonicolor* AKU4429. *J. Mol. Catal. B Enzym.* **6**:305–313.
- Lacombe, E., S. Hawkins, J. Van Doorselaere, J. Piquemal, D. Goffner, O. Poeydomenge, A. E. Boudet, and J. Grima-Pettenati. 1997. Cinnamoyl-CoA reductase, the first committed enzyme of the lignin branch biosynthetic pathway: cloning, expression and phylogenetic relationships. *Plant J.* **11**:429–441.
- Laemmli, U. K. 1970. Cleavage of structural proteins during the assembly of the head of bacteriophage T4. *Nature (London)* **227**:680–685.
- Lemaire, H. G., and B. Muller-Hill. 1986. Nucleotide sequences of the *galE* gene and the *galT* gene of *E. coli*. *Nucleic Acids Res.* **14**:7705–7711.
- Leple, J. C., J. Grima-Pettenati, M. Montagu, and W. Boerjan. 1998. A cDNA encoding cinnamoyl-CoA reductase from *Populus trichocarpa*. *Plant Physiol.* **117**:1126.
- Lorence, M. C., B. A. Murry, J. M. Trant, and J. I. Mason. 1990. Human  $\beta$ -hydroxysteroid dehydrogenase/8 5-4 isomerase from placenta: expression in nonsteioidogenic cells of a protein that catalyzes the dehydrogenation/isomerization of C21 and C19 steroids. *Endocrinology* **126**:2493–2498.
- Messing, J., and J. Vieira. 1982. A new pair of M13 vectors for selecting either DNA strand of double-digest restriction fragments. *Gene* **19**:269–276.
- Nishiyama, M., J. J. Birktoft, and T. Beppu. 1993. Alteration of coenzyme specificity of malate dehydrogenase from *Thermus flavus* by site-directed mutagenesis. *J. Biol. Chem.* **268**:4656–4660.
- Pichon, M., I. Courbou, M. Beckert, A. M. Boudet, and J. Grima-Pettenati. 1998. Cloning and characterization of two maize cDNAs encoding cinnamoyl-CoA reductase (CCR) and differential expression of the corresponding genes. *Plant Mol. Biol.* **38**:671–676.
- Sambrook, J., E. F. Fritsch, and T. Maniatis. 1989. *Molecular cloning: a laboratory manual*, 2nd ed. Cold Spring Harbor Laboratory, Cold Spring Harbor, N.Y.
- Sanger, F., S. Nicklen, and A. R. Coulson. 1977. DNA sequencing with chain-terminating inhibitors. *Proc. Natl. Acad. Sci. USA* **74**:5463–5467.
- Schwarz-Sommer, Z., N. Shepherd, E. Tacke, A. Gierl, W. Rohde, L. Leclercq, M. Mattes, R. Berndtgen, P. A. Peterson, and H. Saedler. 1987. Influence of transposable elements on the structure and function of the A1 gene of *Zea mays*. *EMBO J.* **6**:287–294.
- Scrutton, N. S., A. Berry, and R. N. Perham. 1990. Redesign of the coenzyme specificity of a dehydrogenase by protein engineering. *Nature (London)* **343**:38–43.
- Selman-Housein, G., M. Lopez, D. Hernandez, L. Civardi, F. Miranda, J. Rigau, and P. Puigdomenech. 1998. Unpublished data.
- Sherman, F., G. R. Fink, and J. B. Hicks. 1986. *Laboratory course manual for methods in yeast genetics*. Cold Spring Harbor Laboratory, Cold Spring Harbor, N.Y.
- Shine, J., and L. Dalgarno. 1974. The 3'-terminal sequence of *Escherichia coli* 16S ribosomal RNA: complementarity to nonsense triplets and ribosome binding sites. *Proc. Natl. Acad. Sci. USA* **71**:1342–1346.
- Thoden, J. M., P. A. Frey, and H. M. Holden. 1996. Crystal structures of the oxidized and reduced forms of UDP-galactose 4-epimerase isolated from *Escherichia coli*. *Biochemistry* **35**:2557–2566.
- Wada, M., M. Kataoka, H. Kawabata, Y. Yasohara, N. Kizaki, J. Hasegawa, and S. Shimizu. 1998. Purification and characterization of NADPH-dependent carbonyl reductase, involved in stereoselective reduction of ethyl 4-chloro-3-oxobutanoate, from *Candida magnoliae*. *Biosci. Biotechnol. Biochem.* **62**:280–285.
- Wermuth, B., K. M. Bohren, G. Heinemann, J.-P. Von Wartburg, and K. H. Gabbay. 1988. Human carbonyl reductase. Nucleotide sequence analysis of a cDNA and amino acid sequence of the encoded protein. *J. Biol. Chem.* **263**:16185–16188.
- Wierenga, R. K., M. C. H. De Maeyer, and W. G. J. Hol. 1985. Interaction of pyrophosphate moieties with  $\alpha$ -helices in dinucleotide binding proteins. *Biochemistry* **24**:1346–1357.
- Yamada, H., S. Shimizu, M. Kataoka, H. Sakai, and T. Miyoshi. 1990. A novel NADPH-dependent aldehyde reductase, catalyzing asymmetric reduction of  $\beta$ -keto acid esters, from *Sporobolomyces salmonicolor*: purification and characterization. *FEMS Microbiol. Lett.* **70**:45–48.
- Yaoi, T., K. Miyazaki, T. Oshima, Y. Komukai, and M. Go. 1996. Conversion of the coenzyme specificity of isocitrate dehydrogenase by module replacement. *J. Biochem.* **119**:1014–1018.