Genes from *Pseudomonas* sp. Strain BS Involved in the Conversion of L-2-Amino- Δ^2 -Thiazolin-4-Carbonic Acid to L-Cysteine

Toshikazu Shiba,* Kohji Takeda, Misako Yajima, and Makoto Tadano

Division of Molecular Chemistry, Graduate School of Engineering, Hokkaido University, Sapporo 060-8628, Japan

Received 1 October 2001/Accepted 5 February 2002

DL-2-amino- Δ^2 -thiazolin-4-carbonic acid (DL-ATC) is a substrate for cysteine synthesis in some bacteria, and this bioconversion has been utilized for cysteine production in industry. We cloned a DNA fragment containing the genes involved in the conversion of L-ATC to L-cysteine from *Pseudomonas* sp. strain BS. The introduction of this DNA fragment into *Escherichia coli* cells enabled them to convert L-ATC to cysteine via *N*-carbamyl-L-cysteine (L-NCC) as an intermediate. The smallest recombinant plasmid, designated pTK10, contained a 2.6-kb insert DNA fragment that has L-cysteine synthetic activity. The nucleotide sequence of the insert DNA revealed that two open reading frames (ORFs) encoding proteins with molecular masses of 19.5 and 44.7 kDa were involved in the L-cysteine synthesis from DL-ATC. These ORFs were designated *atcB* and *atcC*, respectively, and their gene products were identified by overproduction of proteins encoded in each ORF and by the maxicell method. The functions of these gene products were examined using extracts of *E. coli* cells carrying deletion derivatives of pTK10. The results indicate that *atcB* and *atcC* are involved in the conversion of L-ATC to L-NCC and the conversion of L-NCC to cysteine, respectively. *atcB* was first identified as a gene encoding an enzyme that catalyzes thiazolin ring opening. *AtcC* is highly homologous with L-*N*-carbamoylases. Since both enzymes can only catalyze the L-specific conversion from L-ATC to L-NCC or L-NCC to L-cysteine, it is thought that *atcB* and *atcC* encode L-ATC hydrolase and *N*-carbamyl-L-cysteine amidohydrolase, respectively.

L-cysteine has been widely used in medicines, cosmetics, and food additives. It has mainly been produced via acid or alkaline hydrolysis of human and animal hairs. However, these methods result in low yields and give rise to unpleasant odors and problems of waste treatment (11). In addition, products extracted from hair do not qualify for medical use because of sanitary problems.

A bioconversion, instead of hydrolysis of hairs, for producing L-cysteine has also been developed (14, 15). Some bacteria, especially in the genus *Pseudomonas* (14, 15, 21), have potent activities of asymmetrical hydrolysis of DL-2-amino- Δ^2 -thiazolin-4-carbonic acid (DL-ATC) to form L-cysteine. DL-ATC is a desirable substrate for the bioconversion since the method for chemical synthesis of DL-ATC is not complicated (14, 15). The total cost for L-cysteine synthesis using DL-ATC is favorable for the large-scale production in industry because the price of DL-ATC is less than that of L-cysteine. Cultured bacteria have been used for commercial production of L-cysteine. As summarized in Fig. 1, bacterial conversion of DL-ATC to L-cysteine consists of the following three successive steps (21): (i) enzymatic racemization of D-ATC to L-ATC; (ii) a ring-opening hydrolysis of L-ATC with an enzyme called ATC hydrolase, to produce either *N*-carbamyl-L-cysteine (L-NCC) or *S*-carbamyl-L-cysteine (L-SCC) as intermediates; and (iii) hydrolysis of these intermediates to L-cysteine with an enzyme. As a result, the whole racemic ATC is converted to L-cysteine.

In order to investigate the molecular mechanisms of the reactions catalyzed by *Pseudomonas* sp. strain BS, we cloned the genes involved in the conversion of L-ATC to L-cysteine.

L-ATC was hydrolyzed to L-NCC by a gene product of *atcB*, and subsequent hydrolysis from L-NCC to L-cysteine was catalyzed by L-NCC amidohydrolase encoded by *atcC*.

MATERIALS AND METHODS

Bacterial strains and plasmids. Bacterial strain BS (FERM P-16965, Fermentation Research Institute, Agency of Industrial Science and Technology, Japanese Ministry of International Trade and Industry), which is able to convert DL-ATC to L-cysteine, was isolated from soil and identified as a *Pseudomonas* sp. by the methods described by Sano et al. (15). *Escherichia coli* DH5 α (12) was used as the host strain for cloning and subcloning of DNA fragments isolated from strain BS. *E. coli* BL21(DE3) (19) and pT7-7 plasmids were used for overexpression of proteins under the control of the T7 promoter (19). pBlue-script SK+ (pBS-SK; Stratagene) was used as a cloning and an expression vector.

Media, culture conditions, and chemicals. Strain BS was grown in Luria-Bertani (LB) medium (12) at 30°C. *E. coli* DH5 α cells carrying the genomic DNA library of strain BS were grown in M9 agar plates containing 0.1% DL-ATC as the sole nitrogen source instead of NH₄Cl and supplemented with thiamine (1 μ g/ml) as a means of selection for gene cloning. For general purpose, *E. coli* DH5 α cells carrying recombinant plasmids were grown in LB medium containing ampicillin (100 μ g/ml). DL-ATC, L-ATC, and D-ATC were kindly donated by Nihon-Rikagaku-yakuhin Co. and Ajinomoto Co. (Tokyo, Japan). L-NCC and *N*-carbamyl-D-cysteine (D-NCC) were purchased from Wako Chemicals Co., Osaka, Japan.

Cloning of genes encoding enzymes involved in L-cysteine production from DL-ATC. The experimental approach used for cloning of the genes involved in the conversion of DL-ATC to L-cysteine was to select the clones of the transformed *E. coli* cells that had the ability to utilize DL-ATC as a nitrogen source. Since *E. coli* does not grow on a medium containing DL-ATC as the sole nitrogen source, it was a feasible host for this experimental approach. The genomic DNA of strain BS was partially digested with *Sau*3AI, and 6- to 10-kb DNA fragments were separated by sucrose gradient centrifugation (12). The DNA fragments were ligated into the *Bam*HI site of pBS-SK. *E. coli* DH5 α was transformed by using ligated DNA, and the cells were spread onto an LB agar plate containing ampicillin (100 μ g/ml) and 5-bromo-4-chloro-3-indolyl- β -D(-)-galactopyranoside (X-Gal) (20 μ g/ml). White colonies were replicated to M9 selective plates, described above, and incubated at 37°C for a few days. Recombinant clones having the ability to utilize DL-ATC as a nitrogen source were expected to form

* Corresponding author. Present address: Fujirebio Inc., 51, Komiya-cho, Hachioji-shi, Tokyo 192-0031, Japan. Phone: 81-426-45-4740. Fax: 81-426-46-8325. E-mail: shiba@dove-mc.eng.hokudai.ac.jp.

FIG. 1. The process of bioconversion from DL-ATC to L-cysteine. Solid lines indicate a metabolic pathway in *Pseudomonas* species, including the BS strain used in this study (21). Dotted lines indicate an alternative pathway (14, 15).

colonies. This process yielded two clones, one of which was designated pTK42 and was used in further studies (Fig. 2).

Identification of gene products. Proteins encoded in the insert DNA fragment carried on pTK10, a deletion derivative of pTK42 (Fig. 2), were identified by overproduction of gene products and by the maxicell method (13).

For overproduction of proteins, DNA fragments containing open reading frame 1 (ORF1) and ORF2 were amplified by PCR and inserted into *Nde*I- and *Bam*HI-digested pT7-7 (19), which has the T7 promoter as described in the legend of Fig. 2. The resultant plasmid pTK-T7-30 or pTK-T7-40 has ORF1 or ORF2 downstream of the T7 promoter, and the gene products of these ORFs could be overproduced in *E. coli* strain BL-21(DE3), which expresses T7 RNA polymerase. Five-milliliter cultures of BL-21(DE3) harboring pTK-T7-30, pTK-T7-40, or pT7-7 (control vector) were grown to the mid-logarithmic phase (optical density at 600 nm = 0.6), and isopropyl- β -D-thiogalactopyranoside (IPTG) (final concentration, 0.4 mM) was added to induce the expression of T7 RNA polymerase. After 3 h of incubation at 37°C, cells were harvested by centrifugation and disrupted by sonication (Sonifier 450D; Branson) for 3 min at output 3. The disrupted cells were centrifuged at 10,000 $\times$ g for 30 min, and the resultant supernatants were collected. Proteins in the supernatant were separated by sodium dodecyl sulfate–14% polyacrylamide gel electrophoresis (SDS–14% PAGE) and visualized by Coomassie brilliant blue staining.

For the maxicell method, *E. coli* DH5 α harboring either pTK10 or pBS-SK was cultured in M9 medium and exposed to UV radiation at a lethal dose. After 12 h of incubation with D-cycloserine, Pro-mix L-[³⁵S] in vitro cell labeling mix (Amersham Pharmacia Biotech) was added to the culture, and plasmid-encoding proteins were specifically labeled. The labeled proteins were analyzed by SDS–12.5% PAGE and visualized using a radio image analyzer (BAS2000; Fujix).

Assay for L-cysteine synthetic activity, including activities of L-ATC hydrolase and L-NCC amidohydrolase. To prepare a crude extract containing L-ATC hydrolase and/or L-NCC hydrolase activities, *E. coli* cells harboring pTK10, pTK30, or pTK40 were collected by centrifugation and disrupted by sonication (Sonifier 450D; Branson). The disrupted cells were centrifuged at 10,000 $\times$ g for 30 min, and the supernatants were collected. The resultant supernatant was added to the standard reaction mixture containing buffer A (50 mM Tris-HCl [pH 8.0], 10 mM dithiothreitol, 2.5 mM hydroxylamine) and 0.5% DL-ATC as a substrate. Reactions were performed at 30°C and stopped by the addition of trichloroacetic acid (final concentration, 5%). Hydrolyzed products of DL-ATC in the reaction mixtures were separated by a cellulose thin-layer chromatography (TLC) plate (Merck) developed with a solvent of 1-butanol–acetic acid–water at the ratio of 2:1:1 (21) and visualized by iodine or ninhydrin staining. The Gaitonde ninhydrin method was used for quantitative analysis of cysteine synthetic activity (3). Cysteine concentration in the reaction mixture was colorimetrically measured by the absorbance at 560 nm.

Other procedures. The nucleotide sequence was determined by the dideoxy chain termination method with some specific primers using an ABI PRISM Big-Dye Terminator Cycle Sequencing Ready Reaction Kit (Applied Biosystems) and an ABI PRISM 310 Genetic Analyzer (Applied Biosystems).

Nucleotide sequence accession number. The nucleotide sequence data reported in this work will appear in the DDBJ, EMBL, and GenBank nucleotide sequence databases under accession number AB070707.

RESULTS

Cloning of genes encoding enzymes involved in L-cysteine production from DL-ATC. The experimental approach used for cloning of the genes involved in the conversion of DL-ATC to L-cysteine was to select the clones of the transformed *E. coli* cells that had the ability to utilize DL-ATC as a nitrogen source. *E. coli* transformants can grow on a medium containing DL-ATC if they have an activity to hydrolyze L-NCC or L-SCC, resulting in NH₃. A *Sau*3AI partially digested genomic DNA library of strain BS was constructed by insertions within the *Bam*HI site of pBS-SK. Approximately 4,000 recombinant clones were screened for their ability to form a colony on the selective plate containing DL-ATC. Two clones formed colonies, and they were also ampicillin resistant. Plasmid DNAs were prepared from these colonies and analyzed by restriction enzyme digestion. The fact that the patterns of the two plasmids were exactly the same indicated that both isolated insert DNAs represent the same region. Figure 2 shows the restriction map of one of the cloned DNAs (pTK42) and its deletion derivatives. Cysteine synthetic activity was detected by the Gaitonde ninhydrin method in the crude extract of *E. coli* harboring pTK-XB and pTK10 (Fig. 3). No cysteine synthetic activity was observed when D-ATC was used as a substrate (data not shown; see also Fig. 8).

DNA and amino acid sequence analyses. The nucleotide sequence of the 3.8-kb region of pTK10 (Fig. 2) was determined as shown in Fig. 4. Analysis of the nucleotide sequence revealed two putative ORFs, designated ORF1 and ORF2. Since these ORFs were in the same orientation and 4 bp, including the stop codon of ORF1 and the initiation codon of ORF2, were overlapped, they may constitute an operon. ORF1 and ORF2 are composed of 519 and 1,260 bp, respectively. ORF1 and ORF2 encode 173 and 420 amino acids, respectively, and each ORF was preceded by possible ribosome-

FIG. 2. Restriction map of a 10-kb insert DNA fragment of pTK42 and structures of its deletion derivatives. ORF1 and ORF2 are located downstream of the *lac* promoter in the pBS-SK vector (*lacOP*) in the same directions in pTK10, pTK30, and pTK40. In pTK-T7-30 and pTK-T7-40, ORF1 and ORF2 are located downstream of the T7 promoter in the pT7-7 expression vector (T7φ10) (19) in the same directions, respectively. Insert DNAs were prepared by PCR amplification with upstream primers p-30-U (5'-AATCTAGAGGAGCAGCATGAGTGATAAGCCAT-3') for pTK30 and p-30T7-U (5'-GGAATTCCATATGAGTGATAAGCCATTGCACCC-3') for pTK-T7-30 and with a common downstream primer, p-30-D (5'-AAGGATCCTTGGCGCTGGATGAGCGGCG-3'), for pTK30 and pTK-T7-30. For pTK40 and pTK-T7-40, insert DNAs were amplified with upstream primers p-40-U (5'-AATCTAGAGGTCAGCAATGAGCGTAATTGAA-3') and p-40T7-U (5'-GGAATTCCATATGAGCGTAATTGAAAACACGACCC-3'), respectively, and with a common downstream primer, p-40-D (5'-AAGGATCCATCGGCGTCGGC TGCAAAGT-3'). The amplified DNA fragments were inserted into *Xba*I- and *Bam*HI-digested pBS-SK and *Nde*I- and *Bam*HI-digested pT7-7 (19). Arrows indicate the region and direction of each ORF and promoters. Abbreviations for restriction enzymes: B, *Bam*HI; E, *Eco*RI; Sau, *Sau*3AI; Sc, *Sac*I; SI, *Sal*I; Sp, *Sph*I; X, *Xba*I.

binding sites (underlined in Fig. 4). Upstream of ORF1, sequences similar to the -10 and -35 consensus sequences for *E. coli* σ^{70} -dependent promoters were found (underlined in Fig. 4).

A computer-aided homology search using the FASTA program (10) revealed that ORF2 is homologous with *N*-carbamoyl-L-amino acid amidohydrolases (*L*-*N*-carbamoylase), including *N*-carbamoyl- β -alanine amidohydrolase, from *Arthrobacter aurescens* (23), *Bacillus stearothermophilus* (2, 9), *Pseudomonas* sp. NS671 (22), and *Saccharomyces kluyveri* (Z. Gojkovic, M. Sandrini, and J. Piskur, unpublished results [GenBank accession number AF333185]). ORF2 is also homologous with putative *L*-*N*-carbamoylase obtained in the genome sequence projects. Figure 5 shows the multiple alignment of five amino acid sequences of *L*-*N*-carbamoylase that have been identified by their enzymatic activity and seven amino acid sequences of putative *L*-*N*-carbamoylase ordered by the best homology scores (10). The sequence identities among these 10 sequences range from 31.7 to 47.2%.

Identification of gene products of the ORFs. As shown in Fig. 6, in the *E. coli* harboring pTK-T7-30 or pTK-T7-40, proteins with molecular masses of about 20 and 45 kDa, corresponding to the calculated masses of proteins encoded by ORF1 and ORF2, respectively, were overproduced. The maxi-

FIG. 3. Cysteine synthetic activity of *E. coli* carrying the deletion derivatives of pTK42. *E. coli* cells carrying pTK-XB, pTK10, or pBS-SK were used for preparing crude extracts. The cysteine synthetic reactions were performed at 30°C in 1-ml reaction mixtures each containing 0.5% DL-ATC. Synthesized cysteine was measured by the Gaitonde ninhydrin method (3).

ORF2	1	-----M-----SVIENT-TRT-PLIQPQ-----WDDLMSSEV-----GALGETM	36
A. aureus	1	-----M-----TLQKA-QAARIEKEI-----SRFSABCP-----RV	28
B. thermophilus	1	-----M-----IQ-G-BRIMORMEIGEV-----KQPSGV	25
P. sp NS671	1	-----M-----K-----T-VTISKERIRIHIHQLEGTCTKDKIV	29
P. aeruginosa PAO1*	1	-----M-----MSTARNVLQSTORHIDGRLQWGLMDLARIKATAGIV	38
M. loti (#1)*	1	-----M-----MA-----APGENLRINS-----DLSIMEMAKITPGIAGIN	33
M. loti (#2)*	1	-----M-----IAVDA-QRL-LGRIRE-----LGAVGRDGEGLRLALAA	32
M. loti (#3)*	1	-----M-----MSLTTPRIDGEL-----LGRLDAGAGIATPAGIV	30
B. halodurans*	1	-----M-----ILTMDINQKRLWRLMEL-----AEIGGTSPID-----VHV	33
S. solfataricus*	1	-----M-----MNPE-----RFLITTFHSLTN-----IGWTEGVL-----RL-----A-L-----NEY-DI-K-VRE-----EL	39
S. kluyveri	1	-----M-----MSKDVSSITITVSASPDGTLNLPAAAPLSIASGRNLTITLGTSGFGVARWQESHEFG	60
A. rhizogenes*	1	-----M-----LETGTMPRADFDATFD-----ISAIGAEAGI-----VTRLTA	35
ORF2	37	-----RFLALSEEDROVRDWFVACRAIL-----CSIEI-----QINIFATYFELLPNAPIAMSSHITOCBS	96
A. aureus	29	-----THLTUTPEHAAR-----RETLLIAAMKAALSVREDALENIIGRRRRTILFELBAIAVSSHFUSVRN	88
B. thermophilus	26	-----THLSFTAEEERRAKDLVASVMREALFVYEDAAENLIGRKECTINEDATVVLVSSHLISVYN	85
P. sp NS671	30	-----QHLALSKEDEHILLVSEWMREALLTVTHHFNENLIGRKECTINEATVVLVSSHLISVYN	89
P. aeruginosa PAO1*	39	-----CHLALSDDLDRCHDLFVQWCEAACTVSUVRVFNIFARRFCRNLDHPVMTGSHITOCIT	98
M. loti (#1)*	34	-----NQTVDTEDEGGRHLFKRWCDAALENGVDEMFMFARRRRTILFELHPVVTGSHITOCIT	93
M. loti (#2)*	33	-----S-DTGWGRDLFVGLRQAGLDVAIDRIGN-----IFGIWQAC-NAGEAPLLTISHTITVID	88
M. loti (#3)*	31	-----NQAFSVEDRIRRLLAELALGRHFTVFCGIANLFIRRECTDATLPLLISSHLISGFT	90
B. halodurans*	34	-----THLSLTKTLPRLRQYVIDLMKDS-----LSVQVDAVENIIGKLEGTDLLEBSVMTGSHISVH	93
S. solfataricus*	40	-----IKIL-----SSI-G-VH-----IMVD-DA-G-----NI-----IGEL-G-----GKLS-DAIAIGSHMSVHY	81
S. kluyveri	61	-----MRLAGTALDGMRDWFTNECESLGCKVKVDKGNMFVYPGKNGKGTATSSHLITOCHE	120
A. rhizogenes*	36	-----SAEDGMARGVFAIWEHAAVEVAHDAIGN-----QFACFTWAG-----BDAWIFAGSHITOCIR	90
ORF2	97	-----GGRFDGILGVLACILVIRLHALDSVRPAHITITLWTNTEGSRF-----APAMMGSGVYVHAHQ	155
A. aureus	97	-----GGRFDGILGVLACILVIRLHALDSVRPAHITITLWTNTEGSRF-----APAMMGSGVYVHAHQ	155
B. thermophilus	86	-----GGRFDGILGVLACILVIRLHALDSVRPAHITITLWTNTEGSRF-----APAMMGSGVYVHAHQ	146
P. sp NS671	90	-----GGRFDGILGVLACILVIRLHALDSVRPAHITITLWTNTEGSRF-----APAMMGSGVYVHAHQ	144
P. aeruginosa PAO1*	94	-----GGRFDGILGVLACILVIRLHALDSVRPAHITITLWTNTEGSRF-----APAMMGSGVYVHAHQ	157
M. loti (#1)*	94	-----GGRFDGILGVLACILVIRLHALDSVRPAHITITLWTNTEGSRF-----APAMMGSGVYVHAHQ	150
M. loti (#2)*	89	-----GGRFDGILGVLACILVIRLHALDSVRPAHITITLWTNTEGSRF-----APAMMGSGVYVHAHQ	142
M. loti (#3)*	91	-----GGRFDGILGVLACILVIRLHALDSVRPAHITITLWTNTEGSRF-----APAMMGSGVYVHAHQ	147
B. halodurans*	94	-----GGRFDGILGVLACILVIRLHALDSVRPAHITITLWTNTEGSRF-----APAMMGSGVYVHAHQ	151
S. solfataricus*	82	-----GGRFDGILGVLACILVIRLHALDSVRPAHITITLWTNTEGSRF-----APAMMGSGVYVHAHQ	138
S. kluyveri	121	-----GGRFDGILGVLACILVIRLHALDSVRPAHITITLWTNTEGSRF-----APAMMGSGVYVHAHQ	180
A. rhizogenes*	91	-----GGRFDGILGVLACILVIRLHALDSVRPAHITITLWTNTEGSRF-----APAMMGSGVYVHAHQ	148
ORF2	156	-----WQTVAAATCKSGISVAQALD-SICY-AG--NR-EPGFMQPSA-----VELHIEQGGVLEAGIE	210
A. aureus	147	-----ADRELDSLVDEGVSVRQAATAFGLKPGELQAARSAADLRHETELHIEQGGVLEAGIE	206
B. thermophilus	145	-----PEALBECRDAEGISLAEAMKQ-A-GLDPRLPQAARKPGTVKAVELHIEQGGVLEAGIE	202
P. sp NS671	149	-----PEDLQKVDNNVTRYEALKTGFGIDPDTHQSIREIGDIKHEFELHIEQGGVLEAGIE	208
P. aeruginosa PAO1*	158	-----LEETLAKRDADGVSVEGALD-AIGY-AG--AR-DCLGHPVGAIFELHIEQGGVLEAGIE	212
M. loti (#1)*	151	-----LDQADVYHTDONGKKGEELE-HI-GWKGT-E-KVGARQIHEFELHIEQGGVLEAGIE	207
M. loti (#2)*	143	-----AGGVDEAALAAVGTGDSLTGQELARIYAGEREPGFLKPHILHIEQGGVLEAGIE	202
M. loti (#3)*	148	-----D-TAAWESIUGADGALLAALEATTITLPAALRPTGMPISGMLHIEQGGVLEAGIE	206
B. halodurans*	152	-----TEITTF-SLADDKGVTYREAFANLNTPLYKQAIRSDQIKELHIEQGGVLEAGIE	216
S. solfataricus*	139	-----DKNY-VYSR-RDKDNTSFEEALRVSGFGDESNRLMMKPNYMLHIEQGGVLEAGIE	190
S. kluyveri	181	-----EEAYGLMSVGEKPEVSY-DSLNIGYIGDTFASVYKENEIHEFELHIEQGGVLEAGIE	239
A. rhizogenes*	149	-----SGTLGLEDALELLDGDGSVLSKALAAIGINGETRAVPPSPKAVELHIEQGGVLEAGIE	208
ORF2	211	-----TGVMEAVQVCLDDIKVPSVSAHAGGPMIMF-----ALVPAESKIVLAV-EAVASHLP-	265
A. aureus	207	-----TGVMEAVQVCLDDIKVPSVSAHAGGPMIMF-----ALVPAESKIVLAV-EAVASHLP-	262
B. thermophilus	209	-----TGVMEAVQVCLDDIKVPSVSAHAGGPMIMF-----ALVPAESKIVLAV-EAVASHLP-	257
P. sp NS671	209	-----TGVMEAVQVCLDDIKVPSVSAHAGGPMIMF-----ALVPAESKIVLAV-EAVASHLP-	265
P. aeruginosa PAO1*	213	-----TGVMEAVQVCLDDIKVPSVSAHAGGPMIMF-----ALVPAESKIVLAV-EAVASHLP-	268
M. loti (#1)*	208	-----TGVMEAVQVCLDDIKVPSVSAHAGGPMIMF-----ALVPAESKIVLAV-EAVASHLP-	263
M. loti (#2)*	203	-----TGVMEAVQVCLDDIKVPSVSAHAGGPMIMF-----ALVPAESKIVLAV-EAVASHLP-	258
M. loti (#3)*	207	-----TGVMEAVQVCLDDIKVPSVSAHAGGPMIMF-----ALVPAESKIVLAV-EAVASHLP-	262
B. halodurans*	211	-----TGVMEAVQVCLDDIKVPSVSAHAGGPMIMF-----ALVPAESKIVLAV-EAVASHLP-	265
S. solfataricus*	197	-----TGVMEAVQVCLDDIKVPSVSAHAGGPMIMF-----ALVPAESKIVLAV-EAVASHLP-	256
S. kluyveri	249	-----TGVMEAVQVCLDDIKVPSVSAHAGGPMIMF-----ALVPAESKIVLAV-EAVASHLP-	298
A. rhizogenes*	209	-----TGVMEAVQVCLDDIKVPSVSAHAGGPMIMF-----ALVPAESKIVLAV-EAVASHLP-	264
ORF2	266	-----GCTVGVFNAAFNRSNVTIGTTLTLEVLPHPPDAEISAL--ET-ELIAEI-KRISPAAEV	321
A. aureus	263	-----GCTVGVFNAAFNRSNVTIGTTLTLEVLPHPPDAEISAL--ET-ELIAEI-KRISPAAEV	322
B. thermophilus	258	-----GCTVGVFNAAFNRSNVTIGTTLTLEVLPHPPDAEISAL--ET-ELIAEI-KRISPAAEV	317
P. sp NS671	266	-----GCTVGVFNAAFNRSNVTIGTTLTLEVLPHPPDAEISAL--ET-ELIAEI-KRISPAAEV	325
P. aeruginosa PAO1*	269	-----GCTVGVFNAAFNRSNVTIGTTLTLEVLPHPPDAEISAL--ET-ELIAEI-KRISPAAEV	328
M. loti (#1)*	264	-----GCTVGVFNAAFNRSNVTIGTTLTLEVLPHPPDAEISAL--ET-ELIAEI-KRISPAAEV	323
M. loti (#2)*	259	-----GCTVGVFNAAFNRSNVTIGTTLTLEVLPHPPDAEISAL--ET-ELIAEI-KRISPAAEV	316
M. loti (#3)*	263	-----GCTVGVFNAAFNRSNVTIGTTLTLEVLPHPPDAEISAL--ET-ELIAEI-KRISPAAEV	322
B. halodurans*	266	-----GCTVGVFNAAFNRSNVTIGTTLTLEVLPHPPDAEISAL--ET-ELIAEI-KRISPAAEV	324
S. solfataricus*	257	-----GCTVGVFNAAFNRSNVTIGTTLTLEVLPHPPDAEISAL--ET-ELIAEI-KRISPAAEV	312
S. kluyveri	299	-----GCTVGVFNAAFNRSNVTIGTTLTLEVLPHPPDAEISAL--ET-ELIAEI-KRISPAAEV	357
A. rhizogenes*	265	-----GCTVGVFNAAFNRSNVTIGTTLTLEVLPHPPDAEISAL--ET-ELIAEI-KRISPAAEV	323
ORF2	322	-----RRVWRKPEVIFDSRIIMDAIDRTEAL-GYSARRMVGAGHI-AAHVAGIASAMTIFSY	379
A. aureus	323	-----RRVWRKPEVIFDSRIIMDAIDRTEAL-GYSARRMVGAGHI-AAHVAGIASAMTIFSY	380
B. thermophilus	318	-----RRVWRKPEVIFDSRIIMDAIDRTEAL-GYSARRMVGAGHI-AAHVAGIASAMTIFSY	375
P. sp NS671	326	-----RRVWRKPEVIFDSRIIMDAIDRTEAL-GYSARRMVGAGHI-AAHVAGIASAMTIFSY	383
P. aeruginosa PAO1*	329	-----RRVWRKPEVIFDSRIIMDAIDRTEAL-GYSARRMVGAGHI-AAHVAGIASAMTIFSY	386
M. loti (#1)*	324	-----RRVWRKPEVIFDSRIIMDAIDRTEAL-GYSARRMVGAGHI-AAHVAGIASAMTIFSY	381
M. loti (#2)*	317	-----RRVWRKPEVIFDSRIIMDAIDRTEAL-GYSARRMVGAGHI-AAHVAGIASAMTIFSY	378
M. loti (#3)*	323	-----RRVWRKPEVIFDSRIIMDAIDRTEAL-GYSARRMVGAGHI-AAHVAGIASAMTIFSY	380
B. halodurans*	325	-----RRVWRKPEVIFDSRIIMDAIDRTEAL-GYSARRMVGAGHI-AAHVAGIASAMTIFSY	382
S. solfataricus*	313	-----RRVWRKPEVIFDSRIIMDAIDRTEAL-GYSARRMVGAGHI-AAHVAGIASAMTIFSY	369
S. kluyveri	358	-----RRVWRKPEVIFDSRIIMDAIDRTEAL-GYSARRMVGAGHI-AAHVAGIASAMTIFSY	416
A. rhizogenes*	324	-----RRVWRKPEVIFDSRIIMDAIDRTEAL-GYSARRMVGAGHI-AAHVAGIASAMTIFSY	383
ORF2	380	-----NLSHNVLEBYSSPOCAQATVLLGAVLELDNLLTADDRNS	420
A. aureus	381	-----NLSHNVLEBYSSPOCAQATVLLGAVLELDNLLTADDRNS	412
B. thermophilus	376	-----NLSHNVLEBYSSPOCAQATVLLGAVLELDNLLTADDRNS	409
P. sp NS671	384	-----NLSHNVLEBYSSPOCAQATVLLGAVLELDNLLTADDRNS	414
P. aeruginosa PAO1*	387	-----NLSHNVLEBYSSPOCAQATVLLGAVLELDNLLTADDRNS	427
M. loti (#1)*	382	-----NLSHNVLEBYSSPOCAQATVLLGAVLELDNLLTADDRNS	416
M. loti (#2)*	376	-----NLSHNVLEBYSSPOCAQATVLLGAVLELDNLLTADDRNS	410
M. loti (#3)*	381	-----NLSHNVLEBYSSPOCAQATVLLGAVLELDNLLTADDRNS	414
B. halodurans*	383	-----NLSHNVLEBYSSPOCAQATVLLGAVLELDNLLTADDRNS	414
S. solfataricus*	370	-----NLSHNVLEBYSSPOCAQATVLLGAVLELDNLLTADDRNS	401
S. kluyveri	417	-----NLSHNVLEBYSSPOCAQATVLLGAVLELDNLLTADDRNS	455
A. rhizogenes*	384	-----NLSHNVLEBYSSPOCAQATVLLGAVLELDNLLTADDRNS	421

FIG. 6. Identification of the proteins encoded by ORF1 and ORF2. The cell extracts of *E. coli* carrying pT7-7 (lane 1), pTK-T7-30 (lane 2), and pTK-T7-40 (lane 3) were analyzed by SDS-14% PAGE.

Functions of the gene products of ORF1 and ORF2. To identify the functions of each ORF, we constructed deletion derivatives of pTK10 as shown in Fig. 2. pTK30 and pTK40 contain a 575-bp DNA fragment including ORF1 and a 1,340-bp DNA fragment including ORF2, respectively, under the control of the *lac* promoter provided by the plasmid vector. These DNA fragments were amplified by PCR with specific primers containing the *Xba*I site (upstream primers) and *Bam*HI site (downstream primers). The primer sequences are presented in the legend of Fig. 2. Amplified DNA fragments were inserted into *Xba*I- and *Bam*HI-digested pBS-SK.

To determine the functions of the gene products of the ORFs, we examined ATC hydrolase activity and the hydrolyzed intermediate. A crude extract of *E. coli* cells carrying pTK30 or pTK40 was prepared as described in Materials and Methods. The activities to convert DL-ATC to a hydrolyzed intermediate and then to convert the intermediate to L-cysteine were analyzed by TLC. As shown in Fig. 7A, DL-ATC was converted to a hydrolyzed intermediate by the *E. coli* extract carrying pTK30 (lane 2), whereas no hydrolyzed product was observed in the reaction mixture containing *E. coli* extract carrying pTK40 (lane 3). The R_f value of the intermediate is 0.85, which is almost the same as that of authentic L-NCC (R_f = 0.84) and very different from that of L-SCC (R_f = 0.43) (21). As demonstrated in Fig. 3, cysteine was produced from DL-ATC with *E. coli* extract carrying pTK10 (lane 4). To further identify the intermediate, a reaction mixture (10 ml) contain-

ing 70 mM DL-ATC and the *E. coli* extract carrying pTK30 was prepared, and the reaction was performed at 30°C for 20 h. The intermediate was purified by the method described by Tamura et al. (21). Analyses of the purified intermediate by two-dimensional ^1H nuclear magnetic resonance indicated that the intermediate was NCC, not SCC (data not shown). In addition, the fact that no hydrolyzed intermediate was observed when D-ATC was used as a substrate in the reaction mixture containing the *E. coli* extract carrying pTK30 indicates that the product of ORF1 can only catalyze the L-specific conversion from L-ATC to L-NCC (Fig. 8A). The gene product of ORF1 encodes L-ATC hydrolase.

When the *E. coli* extract that produces a gene product of ORF2 was incubated with L-NCC, which was produced by incubating DL-ATC with a crude extract containing the product of ORF1, L-NCC was completely converted to L-cysteine (Fig. 7B, lanes 7 and 8). Since no conversion was observed when L-NCC was incubated with *E. coli* extract of a vector and pTK30 transformants (Fig. 7B, lanes 5 and 6, respectively), it seems reasonable to conclude that L-cysteine was produced by an ORF2 gene product, not by spontaneous hydrolysis of L-NCC. These results showed that ORF2 encodes L-NCC amidohydrolase. As well as the product of ORF1, L-NCC, but not D-NCC, can be hydrolyzed to L-cysteine by the product of ORF2 (Fig. 8B).

ORF1 and ORF2 were named *atcB* and *atcC*, respectively, since the genes were involved in ATC hydrolysis.

FIG. 7. Functions of gene products of ORF1 and ORF2. (A) Hydrolysis of DL-ATC by the gene product of *atcB*. Crude extracts of *E. coli* cells carrying pBS-SK (control vector) (lane 1), pTK30 (lane 2), pTK40 (lane 3), or pTK10 (lane 4) (20 μ g each) were mixed with 0.5% DL-ATC in 100- μ l reaction mixtures and incubated at 30°C for 3 h. Two microliters of the reaction mixture was spotted onto a TLC plate, and hydrolyzed products were analyzed and visualized by iodine staining. (B) Hydrolysis of L-NCC by the gene product of *atcC*. L-NCC was produced from DL-ATC in a 100- μ l reaction mixture by incubating 0.5% DL-ATC with a crude extract of *E. coli* cells carrying pTK30 (20 μ g) at 30°C for 12 h. The crude extract of *E. coli* cells carrying pBS-SK (control vector) (lane 1), pTK30 (lane 2), pTK40 (lane 3), or pTK10 (lane 4) (20 μ g each) was further added to the aliquoted reaction mixtures (20 μ l each), and the reactions were continued for a further 3 h. The hydrolyzed products were analyzed by TLC and visualized by iodine staining. L-NCC and L-cysteine are represented as NCC and Cys, respectively. Origin, origin of development of TLC.

DISCUSSION

In this study, we isolated the gene encoding L-ATC hydrolase and L-NCC amidohydrolase from *Pseudomonas* sp. strain BS, which is able to convert DL-ATC to L-cysteine. Many bacteria, especially in the genus *Pseudomonas*, that have an activity to convert DL-ATC to L-cysteine have been isolated (14, 15, 21), and two biosynthetic pathways have been described (21). As an intermediate, one pathway produces L-NCC and another pathway produces L-SCC (Fig. 1). We showed the existence of the former pathway by identifying the gene involved in L-cysteine synthesis from DL-ATC in strain BS.

The 2.6-kb insert DNA fragment in pTK10 was sequenced, and two ORFs (*atcB* and *atcC*) were found. L-ATC hydrolase encoded by *atcB* did not share extensive homology with other enzymes whose functions have been reported in detail. However, recently, an *atcB* homologue has also been cloned from *Pseudomonas* sp., although the only report on this homologue is a Japanese patent (1). The homologue also has L-ATC hydrolase activity, and its amino acid sequence is 53.5% identical with AtcB (1). L-NCC amidohydrolase encoded by *atcC* shares strong homology with L-N-carbamoylase (23). L-N-Carbamoylase catalyzes the conversion of N-carbamyl-L-amino acids to L-amino acids (23). The gene for L-N-carbamoylase has been cloned in *Pseudomonas* sp. (22), *A. aurescens*, (23), and *B. stearothermophilus* (2, 9). In *Pseudomonas* sp. strain NS671,

hyuC constitutes an operon with *hyuA* and *hyuB*, which are involved in the conversion of D- and L-5-substituted hydantoin to corresponding N-carbamyl-D- and N-carbamyl-L-amino acids, respectively (22). Although it is not clear whether HyuC can hydrolyze L-NCC and whether AtcC can hydrolyze N-carbamyl-L-amino acids other than L-NCC, the similarity of the two enzymes could be due to the similarity of the substrates. Furthermore, the similarity of ATC-utilizing enzymes and hydantoin-utilizing enzymes suggests that these enzymes are evolutionarily oriented from the same origin and may have the same physiological functions in *Pseudomonas*. D-hydantoinase, the product of *hyuA*, may be involved in the catabolism of uracil and thymine (6), and it has gained considerable attention for its ability to synthesize the less commonly occurring D-amino acids by coupling with L-n-carbamoylase (6). ATC-utilizing enzymes may also be involved in amino acid synthesis in bacteria. Since a branched cyclic dodecylpeptide antibiotic bacteriocin that contains thiazolin is produced in some strains of *Bacillus* (7) and since some siderophores (bacterially produced iron-chelating compounds), such as yersiniabactin and anguibactin, also contain thiazolin (4), *atcB* and *atcC* may be involved in the synthesis or degradation of these thiazolin-containing metabolites. Further analysis of the metabolism of thiazolin-containing compounds is needed to elucidate the detailed biological functions of *atcB* and *atcC*.

FIG. 8. The L-specific conversion from L-ATC to L-NCC by AtcB (A) and from L-NCC to L-cysteine by AtcC (B). Crude extracts (20 μ g) of *E. coli* cells carrying pTK30 (A) or pTK40 (B) were mixed with DL-ATC (lane 1), L-ATC (lane 2), and D-ATC (lane 3) (A) or L-NCC (lane 1) and D-NCC (lane 2) (B). Reactions were performed in 100- μ l reaction mixtures containing buffer A with 0.5% substrates and incubated at 30°C for 1 h. Two microliters of the reaction mixture was spotted onto a TLC plate, and hydrolyzed products were analyzed and visualized by iodine staining (A) or ninhydrin staining (B).

Recently, we have also cloned an ATC racemase gene (17) and a regulator gene for ATC utilization in strain BS. Further research on these genes is now in progress.

REFERENCES

- Asada, Y., T. Omachi, and Y. Tamura. March 2001. *Pseudomonas* D,L-2-amino- Δ^2 -thiazolin-4-carboxylic acid hydrolase and *N*-carbamoyl-L-cysteine amidohydrolase genes for L-cysteine biosynthesis. Japan Patent, Kokai Tokkyo Koho, JP2001069985.
- Batisse, N., P. Weigel, M. Lecocq, and V. Sakanyan. 1997. Two amino acid amidohydrolase genes encoding L-stereospecific carbamoylase and aminocyclase are organized in a common operon in *Bacillus stearothermophilus*. *Appl. Environ. Microbiol.* **63**:763–766.
- Gaitonde, M. K. 1967. A spectrophotometric method for the direct determination of cysteine in the presence of other naturally occurring amino acids. *Biochem. J.* **104**:627–633.
- Gehring, A. M., E. DeMoll, J. D. Fetherston, I. Mori, G. F. Mayhew, F. R. Blattner, C. T. Walsh, and R. D. Perry. 1998. Iron acquisition in plague: modular logic in enzymatic biogenesis of yersiniabactin by *Yersinia pestis*. *Chem. Biol.* **5**:573–586.
- Kaneko, T., Y. Nakamura, S. Sato, E. Asamizu, T. Kato, S. Sasamoto, A. Watanabe, K. Idesawa, A. Ishikawa, K. Kawashima, T. Kimura, Y. Kishida, C. Kiyokawa, M. Kohara, M. Matsumoto, A. Matsuno, Y. Mochizuki, S. Nakayama, N. Nakazaki, S. Shimpo, M. Sugimoto, C. Takeuchi, M. Yamada, and S. Tabata. 2000. Complete genome structure of the nitrogen-fixing symbiotic bacterium *Mesorhizobium loti*. *DNA Res.* **7**:331–338.
- Kim, G. L., D. E. Lee, and H.-S. Kim. 2000. Functional expression and characterization of the two cyclic amidohydrolase enzymes, allantoinase and a novel phenylhydantoinase, from *Escherichia coli*. *J. Bacteriol.* **182**:7021–7028.
- Konz, D., A. Klens, K. Schörgendorfer, and M. A. Marahiel. 1997. The bacitracin biosynthesis operon of *Bacillus licheniformis* ATCC 10716: molecular characterization of three multi-modular peptide synthetase. *Chem. Biol.* **4**:927–937.
- Moriguchi, K., Y. Maeda, M. Satou, M. Kataoka, N. Tanaka, and K. Yoshida. 2000. Analysis of unique variable region of a plant root inducing plasmid, pRil724, by the construction of its physical map and library. *DNA Res.* **7**:157–163.
- Mukohara, Y., T. Ishikawa, K. Watabe, and H. Nakamura. 1993. Molecular cloning and sequencing of the gene for a thermostable *N*-carbamyl-L-amino acid amidohydrolase from *Bacillus stearothermophilus* strain NS1122A. *BioSci. Biotechnol. Biochem.* **57**:1935–1937.
- Pearson, W. R., and D. J. Lipman. 1988. Improved tools for biological sequence comparison. *Proc. Natl. Acad. Sci. USA* **85**:2444–2448.
- Ryu, O. H., J. Y. Ju, and C. S. Shin. 1997. Continuous L-cysteine production using immobilized cell reactors and product extractors. *Proc. Biochem.* **32**:201–209.
- Sambrook, J., E. F. Fritsch, and T. Maniatis. 1990. *Molecular cloning: a laboratory manual*, 2nd ed. Cold Spring Harbor Laboratory Press, Plainview, N.Y.
- Sancar, A., R. P. Wharton, S. Seltzer, B. M. Kacinski, N. D. Clarke, and W. D. Rupp. 1981. Identification of the *uvrA* gene product. *J. Mol. Biol.* **148**:45–62.
- Sano, K., and K. Mitsugi. 1978. Enzymatic production of L-cysteine from D,L-2-amino- Δ^2 -thiazolin-4-carboxylic acid by *Pseudomonas thiazolinophilum*: optimal conditions for the enzyme formation and enzymatic reaction. *Agric. Biol. Chem.* **42**:2315–2321.
- Sano, K., K. Yokozeki, F. Tamura, N. Yasuda, I. Noda, and K. Mitsugi. 1977. Microbial conversion of DL-2-amino- Δ^2 -thiazolin-4-carboxylic acid to L-cysteine and L-cysteine: screening of microorganisms and identification of products. *Appl. Environ. Microbiol.* **34**:806–810.
- She, Q., R. K. Singh, F. Confalonieri, Y. Zivanovic, G. Allard, M. J. Awayez, C. C. Chan-Weiher, I. G. Clausen, B. A. Curtis, A. DeMoors, G. Erauso, C. Fletcher, P. M. Gordon, I. Heikamp-De Jong, A. C. Jeffries, C. J. Kozera, N. Medina, X. Peng, H. P. Thi-Ngoc, P. Redder, M. E. Schenk, C. Theriault, N. Tolstrup, R. L. Charlebois, W. F. Doolittle, M. Duguet, T. Gaasterland, R. A. Garrett, M. A. Ragan, C. W. Sensen, and J. Van Der Oost. 2001. The complete genome of the crenarchaeon *Sulfolobus solfataricus* P2. *Proc. Natl. Acad. Sci. USA* **98**:7835–7840.
- Shiba, T. April 2001. 2-Aminothiazolin-4-carboxylate racemase and gene encoding therefor. U.S. patent 6,214,590 B1.
- Stover, C. K., X. Q. Pham, A. L. Erwin, S. D. Mizoguchi, P. Warrenner, M. J. Hickey, F. F. Brinkman, W. O. Hufnagle, D. J. Kowalik, M. Lagrou, R. L. Garber, L. Goltry, E. Tolentino, S. Westbrook-Wadman, Y. Yuan, L. L. Brody, S. N. Coulter, K. R. Folger, A. Kas, K. Larbig, R. Lim, K. Smith, D. Spencer, G. K. Wong, Z. Wu, and I. T. Paulsen. 2000. Complete genome sequence of *Pseudomonas aeruginosa* PA01, an opportunistic pathogen. *Nature* **406**:959–964.
- Studier, F. W., A. H. Rosenberg, J. J. Dunn, and J. W. Dubendorff. 1990. Use of T7 RNA polymerase to direct expression of cloned genes. *Methods Enzymol.* **185**:60–89.
- Takami, H., K. Nakasone, Y. Takaki, G. Maeno, R. Sasaki, N. Masui, F. Fuji, C. Hirama, Y. Nakamura, N. Ogasawara, S. Kuhara, and K. Horikoshi. 2000. Complete genome sequence of the alkaliphilic bacterium *Bacillus halodurans* and genomic sequence comparison with *Bacillus subtilis*. *Nucleic Acids Res.* **28**:4317–4331.

21. **Tamura, Y., M. Nishino, T. Ohmachi, and Y. Asada.** 1998. *N*-carbamoyl-L-cysteine as an intermediate in the bioconversion from D,L-2-amino- Δ^2 -thiazoline-4-carboxylic acid to L-cysteine by *Pseudomonas* sp. ON-4a. *Biosci. Biotechnol. Biochem.* **62**:2226–2229.
22. **Watabe, K., T. Ishikawa, Y. Mukohara, and H. Nakamura.** 1992. Cloning and sequencing of the genes involved in the conversion of 5-substituted hydantoins to the corresponding L-amino acids from the native plasmid of *Pseudomonas* sp. strain NS671. *J. Bacteriol.* **174**:962–969.
23. **Wilms, B., A. Wiese, C. Syldatk, R. Mattes, J. Altenbuchner, and M. Pietzsch.** 1999. Cloning, nucleotide sequence and expression of a new L-*N*-carbamoylase gene from *Arthrobacter aureus* DSM 3747 in *E. coli*. *J. Biotechnol.* **68**:101–113.