

Cloning, Sequence Analysis, and Characterization of the *astA* Gene Encoding an Arylsulfate Sulfotransferase from *Citrobacter freundii*

Jin-Wook Kang¹, Yeon-Joo Jeong¹, Ae-Ran Kwon¹, Hee-Jeong Yun¹, Dong-Hyun Kim², and Eung-Chil Choi¹

¹College of Pharmacy, Seoul National University, Seoul 151-742, Korea and ²College of Pharmacy, Kyunghee University, Seoul 131-701, Korea

(Received April 20, 2001)

Arylsulfate sulfotransferase (ASST) transfers a sulfate group from a phenolic sulfate ester to a phenolic acceptor substrate. In the present study, the gene encoding ASST was cloned from a genomic library copy of *Citrobacter freundii*, subcloned into the vector pGEM3Zf(-) and sequenced. Sequencing revealed two contiguous open reading frames (ORF1 and ORF2) on the same strand and based on amino acid sequence homology, they were designated as *astA* and *dsbA*, respectively. The amino acid sequence of *astA* deduced from *C. freundii* was highly similar to that of the *Salmonella typhimurium*, *Enterobacter amnigenus*, *Klebsiella*, *Pseudomonas putida*, and *Campylobacter jejuni*, encoded by the *astA* genes. However, the ASST activity assay revealed different acceptor specificities. Using *p*-nitrophenyl sulfate (PNS) as a donor substrate, α -naphthol was found to be the best acceptor substrate, followed by phenol, resorcinol, *p*-acetaminophen, tyramine and tyrosine.

Key words: Arylsulfate sulfotransferase, *astA* gene, *Citrobacter freundii*, Molecular cloning, Sequencing

INTRODUCTION

Phenol sulfation is considered to be a major metabolic pathway for the detoxification of endogenous and exogenous compounds bearing phenolic functional groups (Dodgson *et al.*, 1960). It is catalyzed by phenol sulfotransferase (PST), an enzyme produced by purification from several mammalian organs, including the liver, lung, brain, kidney, epithelial cells, and erythrocytes (Roy, 1981, Sekura *et al.*, 1981). PST catalyzes the transfer of the sulfate group from 3'-phosphoadenosine-5'-phosphosulfate (PAPS) to an acceptor compound, thus forming a sulfate ester (Sekura *et al.*, 1979, Spencer *et al.*, 1960).

Kim *et al.* isolated three arylsulfate sulfotransferase (ASST) - producing bacteria; *Eubacterium* A-44 (Kim *et al.*, 1986), *Klebsiella* K-36 (Kim *et al.*, 1992), and *Haemophilus* K-12 (Lee *et al.*, 1995) from human, rat, and mouse intestinal flora, respectively. Kwon *et al.* also isolated ASST from

Enterobacter amnigenus AR-37 (Kwon *et al.*, 1999). These ASST producing bacteria catalyze the stoichiometric transfer of a sulfate group from phenolic sulfate esters to phenolic compounds, whereas mammalian enzyme catalyzes the transfer of a sulfate group from PAPS to phenolic acceptor substrates (Kim *et al.*, 1991, 1994). ASST also catalyzes the sulfation of the tyrosine residues of peptides and proteins such as kyotorphin, enkephalin, cholecystokinin-8, trypsin inhibitor, and insulin (Kobashi *et al.*, 1986). In addition, sulfation by the ASST obtained from *Eubacterium* A-44 produced a recombinant hirudin whose antithrombin activity, was found to be increased by about 3.4-fold relative to that of unsulfated hirudin (Muramatsu *et al.*, 1994). In contrast to the numerous studies of mammalian tissue PST enzymes, little research has been conducted into the enzymatic sulfation by intestinal bacteria. As a preliminary experiment to determine the function of this enzyme, the distribution of bacteria exhibiting ASST activity was investigated by analyzing about 1,300 bacterial strains maintained in our laboratory. According to Baek *et al.*, only 29 bacterial isolates exhibit such activity, suggesting that this enzyme might not be essential for the bacterial viability (Baek *et al.*, 1998).

Correspondence to: College of Pharmacy, Seoul National University, San 56-1, Shillim-Dong, Kwanak-Gu, Seoul 151-742, Korea

E-mail: ecchoi@snu.ac.kr

Citrobacter freundii MB4-8242, which has high ASST activity, was selected for further genetic characterization. This paper describes the cloning and sequencing of the *C. freundii* *astA* gene, an arylsulfate sulfotransferase determinant.

MATERIALS AND METHODS

Bacterial strains, plasmids and growth conditions

C. freundii was grown in Luria-Bertani (LB) broth at 37 °C for 12 hr. *Escherichia coli* TH2 was used as the cloning host, and *Escherichia coli* NM522 was used for subcloning and DNA sequencing. These strains were also grown in LB broth at 37°C for 12 h. The plasmid pKF3 (Takara Shuzo, Tokyo, Japan) was used as the cloning vector, while the vector for subcloning and DNA sequencing was the plasmid pGEM3Zf(-) (Promega, Madison, WI). For the selection of *astA* transformants, 0.1 mM 4-methylumbelliferyl sulfate (4-MUS) was added to the LB medium. Since the acceptor substrate of ASST is supplied by the medium and the cellular components, the 4-MUS in the medium was cleaved to 4-methylumbelliferone, a fluorescent product, by ASST. Consequently, positive colonies were monitored by UV fluorescence (320 nm).

Molecular cloning of *astA* from *Citrobacter freundii* MB4-8242

The chromosomal DNA from *C. freundii* MB4-8242 was isolated (Kim *et al.*, 1992) and partially digested with *Sau3AI* (Takara Shuzo, Tokyo, Japan). Fragments ranging from 4 to 9 kb were then fractionated by 0.8% agarose gel electrophoresis, purified, ligated into *Bam*HI-digested pKF3 with T4 DNA ligase (Gibco BRL, Gaithersburg, MD) and used to transform competent *E. coli* TH2 cells. The colonies carrying the subcloned plasmid were screened for enzyme activity by plating them on an LB plate containing 4-MUS. Plasmids from the positive colonies were prepared utilizing the alkaline extraction method, as previously described (Birnboim, 1983).

Nucleotide sequencing and sequence analysis

For sequencing, the small fragments which are digested with various enzyme were ligated into the pGEM3Zf(-) vector with T4 DNA ligase. These ligation mixtures were used to transform competent *E. coli* NM522 cells. By using the various subclones in pGEM3Zf(-), the sequence of the region containing the *astA* gene was determined by the dideoxy chain termination method (Sanger *et al.*, 1977) using an AccuPower sequencing kit (Bioneer, Taejon, Korea). Both of the DNA strands were sequenced by a combination of specific oligonucleotide primers (primer sequences: forward 5'-CACGACGTTGTAACGAC-3', reverse 5'-GGATAACAATTCACACAGG-3') on an automated DNA sequencer (Model 4000, Li-Cor, Lincoln, NE).

The deduced amino acid sequence of ASST was subjected to homology searches using the BLAST network service (Altschul *et al.*, 1990) at NCBI and the CLUSTAL W program (Thompson *et al.*, 1994).

ASST activity assay

Cell extracts from *C. freundii* MB4-8242 and *E. coli* NM 522 harboring the plasmid pJS1 were prepared as previously described (Baek *et al.*, 1996). The assay mixture (0.63 ml total volume) containing 30 µl of 50 mM *p*-nitrophenyl sulfate (PNS), 0.29 ml of 20 mM phenol (or occasionally alternative acceptors such as α -naphthol, resorcinol, *p*-acetaminophen, tyramine and tyrosine), 0.21 ml of 0.1 M Tris-HCl, pH 8.0, and 0.1 ml of the cell extract, was incubated at 37°C for 20 min, and the reaction was stopped by the addition of 0.4 ml of 1 N NaOH. The presence of ASST was detected by the appearance of a yellow color. Quantitative measurements were also performed by measuring the absorbance at 405 nm.

Nucleotide sequence accession number

The nucleotide sequence data reported in this paper will appear in the EMBL and NCBI nucleotide sequence databases under accession number AY029222.

RESULTS AND DISCUSSION

Cloning of *astA* from *Citrobacter freundii* MB4-8242

E. coli TH2 cells, transformed with the genomic library copy of *C. freundii* MB4-8242 chromosomal DNA by utilizing the pKF3 vector, were cultured on LB agar containing 4-MUS. The ASST activity-positive colonies which produced blue fluorescent light under UV illumination (320 nm) were isolated and purified. The smallest subclone harboring a 3.0-kb insert was designated as pJS1 and subjected to further analysis.

Nucleotide sequence analysis

The pJS1 restriction map and the ORF locations are depicted in Fig. 1. Various restriction enzymes were used

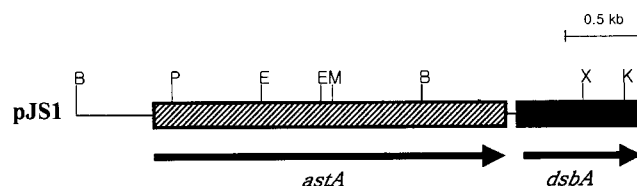


Fig. 1. Partial restriction map of pJS1: The solid box indicates the truncated *dsbA* gene, while *astA* is symbolized by the hatched box. The transcription direction of these different genes is shown by arrows. Restriction enzyme sites: B, *Bam*HI; E, *Eco*RI; K, *Kpn*I; M, *Mlu*I; P, *Pst*I; X, *Xho*I

```

1   GGATCCAATAATAATCCCAGCTATGGGGTGTATGTGGAATGTCGGTGATGGGATTAAGCCTCGTCATCCTCCAGTCAATGGGAATCGGATAATCCAGA
101  GGGCGGATGCCACCCTGTAAATTTACAGCATCCGCTCATCCATAGCAGAATATGAAGGACTAATCCTGTTCCTTAACCGCGCTGTATTCTCCAGAGC
201  GACGAAAGAGTTTGTTCGCGCTCAAGTAAGGCATCTCTGCTCGCGGAATGCCCGACAATAACGTTTCATATGGCAATAATAACCCACCGCAACGAG
301  TGCAATTAATACAATCAATACCAGGCAATAAACATTCTTCTACCCTTAAATTTAACATCATTAAACATTTAAACGTTAAATTTTACACAACCTTATTC
401  TGACATACTCATATAGTGGCATATTTCCAGAAACAAATATATGATTTGATCATATAAATCAGCGCTAAGCTATTAGCCGAACTCGTTATAACAATTAA
501  AACAAATGTCGACAAGATCACACTACGCAATAATAAATATTGTTCTTAAATATCTCACTCATACTATCAAAAAATAATAAAAAAAGCTCCTGAAGAAT
601  AATACATTTCAATATTAATTAATGCAATTAACAAAAGTAGTATTAGCAATACCAITTTATTTTATCAAACCGTAAAAACATAAATATATTCAGCAAGAT
                                     -35                               -10
701  TATATTTATAGGTAATATAAGGAGATATATCCATGTTTACACAATACCGAAAAACACTTCTGGCAGGACTCTGGCACTGACATTTGGATTAGCAGCAGG
      RBS           M F T Q Y R K T L L A G T L A L T F G L A A G
                    astA
801  AAATCTCTTGGCGTGGTTTTCAACCGGCCAACCTGCAGGAAAATGGGCGCAATAGTAGTTGACCCCTTATGAAAATGCCCGCTTACTGCGCTGGTT
      N S L A A G F Q P A Q P A G K L G A I V V D P Y G N A P L T A L V
901  GAATTAGATAGCCAGTTATTTTCAGATGTTAAAGTTACCGTCCATGGTAAAGGTGAAAAAGCGGTTCCGGTGACTTATACCGTGGCAAGAGTCACTCG
      E L D S H V I S D V K V T V H G K G E K G V P V T Y T V G K E S L A
1001 CCACGTACGATGGCATTCTCTATTTTGGCCTTTATCAAAGTTCGCAATAAAGTAAACCGTTGAATATAAAGAAAACCGTAAGCCATGAAAGATGATTA
      T Y D G I P I F G L Y . Q K F A N K V T V E Y K E N G K A M K D D Y
1101 TGTGGTCAAACCTCCGCCATCGTAAATCACTACATGGATAACCGCTCAATATCTGATTTACAGCAAAACAAAGTGATTAAGTAGCACCGGGTTTGAA
      V V Q T S A I V N H Y M D N R S I S D L Q Q T K V I K V A P G F E
1201 GATCGCTTTTATCGTGAATACCCACACCTTTACCCCGCAGGGTGCAGAATCCACTGGCATGGCGAAAAAGATAAAAAAGCCGATCCCTGGATGGCG
      D R L Y L V N T H T F T P Q G A E F H W H G E K D K N A G I L D A
1301 GTCTGCTGGCGGCGCACTCCCTTTGATATCGCACCCCTTACCTTTGCTCGATACGGAAGGTGAATATCGGTGGTGGCTGGATCAGGACACCTTCTA
      G P A G G A L P F D I A P F T F V V D T E G E Y R W W L D Q D T F Y
1401 TGACGGCCATGACATGGATATCAACAAGCGCGGTACCTGATGGGTATTCGTGAAACACCGGAGGCACCTTACCGCGCTCCAGGGCCAGCACTGGTAT
      D G H D M D I N K R G Y L M G I R E T P R G T F T A V Q G Q H W Y
1501 GAATTCGATATGCTGGGGCAAATTCGGCCGACCATAAGCTGCCGCGGATTCCTCGACGCTGCATGAGTCGGTCAAACAGTGAACGCGCACCGTAC
      E F D M L G Q I L A D H K L P R G F L D A S H E S V E T V N G T V
1601 TGCTTCGCGTGGGTAACCGGATTATCGCAAGAAGATGGCCTTACCTACATACCATCCGCGACCAGATTATTGAAGTCGATAAATCTGGACGTGTTGT
      L L R V G K R D Y R K E D G L H V H T I R D Q I I E V D K S G R V V
1701 TGACGTATGGGATCTGACACAAATTCGACCCGATGGGTGACCGCTGCTGGGTGCGTGGACGCGAGGCGGATGCGTCAACGTGGATCTCGCTCAC
      D V W D L T Q I L D P M R D A L L G A L D A G A V C V N V D L A H
1801 GCAGGCCAACAGGCAAACCTGGAGCCAGATACCCCTATGGTGACCCCTCGGCGTGGTGGCGGCGTAACTGGGCTCAGTCAACTCCATTGCCATG
      A G Q Q A K L E P D T P Y G D A L G V G A G R N W A H V N S I A Y
1901 ACGCTAAAGATGACTCCATCTTTCTTCCGCCATCAGGGTGTGCTCAAAATCGGTGCGGATAAACAGGTGAAATGGATCTGGCACCGTCTAAAGG
      D A K D D S I I L S S R H Q G V V K I G R D K Q V K W I L A P S K G
2001 CTGGAATAAGGCGCTGGCAAGCAAACCTGTTAAACCGGTTGATGATAAGGCAACCGCTGAAGTGTGATGAAAACGGTAAGTGGGAAATACTGATTTT
      W N K A L A S K L L K P V D D K G N A L K C D E N G K C E N T D F
2101 GATTTTACTACACCAACACACCGCTGGCTCTCCAGCAAAGGAACGCTAACCATCTTTGACAATGGCGATGGCGTGGACTCGAACACCGGCTCTAC
      D F T Y T Q H T A W L S S K G T L T I F D N G D G R G L E Q P A L

```

Fig. 2. Nucleotide and deduced amino acid sequences of the *astA* and truncated *dsbA* genes of *Citrobacter freundii* MB4-8242: The predicted 35 and 10 regions of the putative promoter and the putative ribosome binding site (RBS) are indicated. The stop codon of the *astA* gene is marked with an asterisk.

```

2201 CAACAATGAAATACTCCCGGTTTCGGAATATAAAATTGACGAGAAAAAGGAACCGTTTCAGCAAGTTTGGGAATACGGGAAGAGCGCGTTATGATTT
    P T M K Y S R F V E Y K I D E K K G T V Q Q V W E Y G K E R G Y D F
2301 CTACAGTCTATCAGTCGGTAATTGAATATCAAAAAGACCGGACACGATGTTGGATTGGTGGTTCTATTAATTTATTTGACGTTGGACAGCCAACC
    Y S P I T S V I E Y Q K D R D T M F G F G G S I N L F D V G Q P T
2401 ATCGGCAAAATCAACGAAATGACTACAAAACAAAAGAAGTCAAAGTCAAATGATGTCCTTTCGATAAACCAACCAACCACTACCGCGCATTAC
    I G K I N E I D Y K T K E V K V E I D V L S D K P N Q T H Y R A L
2501 TGGTTCGCCCAACAATAATGTTAAATTAATTCGTTAGATTAAGGATTAATTAATGTCATCTAAATGGATTACTTCATTATTTAAAAGCGTGGTATTAAGC
    L V R P Q Q M F K *           RBS           M S S K W I T S L F K S V V L T
                                           dsbA
2601 GCAGCGCTGGTATCACCATTTCAGCATCGGCTTTTACTGAAGTACTGACTATATGGTGCTGGAGAAAACCGATTCCCAATGCTGATAAAAACGTTGATCA
    A A L V S P F A A S A F T E G T D Y M V L E K P I P N A D K T L I
2701 AAGTATTCAGTACGCTGCGCATCTGCTACAAATATGATAAAGCCGTGACCGGCCCGGTGTCAGATAAAGTTGCCGATCTCGTCGCCTTCACACCGTT
    K V F S Y A C P F C Y K Y D K A V T G P V S D K V A D L V A F T P F
2801 TCATCTCGAGACCAAGGAGAATACGGCAACAGGCCAGTGAAGTATTGTCAGTAATGATTGCTAAGGATCAGGCCGAGGAATTCATTGTTTGTATGCC
    H L E T K G E Y G K Q A S E V F A V M I A K D Q A A G I S L F D A
2901 AAATCACAGTTCAAAAAGCTAAGTTTGCTTACTACACTGCCTACCACGATAAGAAAAGAACGCTGGTCTGACGGTAAAGATC
    K S Q F K K A K F A Y Y T A Y H D K K E R W S D G K D

```

Fig. 2. Continued

to generate this restriction enzyme map of the pJS1 insert. A total of 2982 bp was sequenced and the presence of two contiguous ORFs on the same strand was revealed. Fig. 2 shows the complete nucleotide sequences of the two ORFs, along with the corresponding deduced amino acid sequences. ORF1 is 1794 nucleotides long, beginning with an ATG start codon at position 733 and ending with a TAA stop codon at position 2526. A promoter-like sequence (Harley *et al.*, 1987) is located upstream from the start codon, with a -35 sequence of TTATCA followed by a 17 bp space and a -10 region of TATATT. A potential ribosome-binding site (AGGAGA) is present 7 bp upstream of the predicted ATG start codon, which is similar to that of *E. coli* (Shine *et al.*, 1974) (Fig. 2). A BLAST programs database search revealed that the deduced amino acid sequence shares many similarities with the ASSTs from *Salmonella typhimurium* (92% match) (Kang *et al.*, 2001), *Enterobacter amnigenus* (89% match) (Kwon *et al.*, 1999), *Klebsiella* (87% match) (Baek *et al.*, 1996), *Pseudomonas putida* (45% match) (Kahnert *et al.*, 2000), *Eubacterium rectale* (23% match) (Goldberg *et al.*, 2000), as well as with the arylsulfatase from *Campylobacter jejuni* (40% match) (Yao *et al.*, 1996) (Fig. 3). Therefore, ORF1 was designated as *astA*. The higher homology of the *C. freundii* enzyme with that of *S. typhimurium*, *E. amnigenus* and *Klebsiella*, as opposed to that of *P. putida* and *C. jejuni*, agrees well with the phylogenetic distance. The *Citrobacter*, *Salmonella*, *Enterobacter*, and *Klebsiella* genera belong to group 5 of the Enterobacteriaceae family, but the *Pseudomonas* and *Campylobacter* genera belong

to group 4 and 2, respectively (Holt *et al.*, 1994). The homology to the *E. rectale* arylsulfotransferase is very low. Furthermore, *Eubacterium A-44 astA* was not homologous to the DNA from *Klebsiella K-36* (Baek *et al.*, 1998). These findings allow us to predict that the gene encoding an ASST of Gram-positive bacteria is significantly different from that of Gram-negative bacteria. The alignment of the six apparent ASSTs from *C. freundii*, *S. typhimurium*, *E. amnigenus*, *Klebsiella*, *P. putida*, and *C. jejuni* indicates conserved regions, which will assist in the location of active sites. Such active sites might contain a tyrosine residue, which is the only amino acid containing the phenolic hydroxy group. At present, eleven conserved tyrosine residues have been identified, and the determination of possible active sites by site-directed mutagenesis is necessary.

Truncated ORF2 consists of 430 bp, starting with an ATG start codon at nucleotide 2553. The putative ribosome-binding site (AAGGA) is located 11 to 7 bp upstream of the start codon of the ORF2. A GenBank BLAST search revealed that the predicted protein encoded by ORF2 is significantly well matched to the *dsbA* genes (Bardwell *et al.*, 1991), which encode an enzyme involved in disulfide bond formation. The subclones containing *dsbA* did not exhibit ASST activity (data not shown). DsbA acts to facilitate the formation of disulfide bonds of many proteins in periplasmic space (Bardwell *et al.*, 1991).

Substrate specificity

The acceptor substrate specificity of ASST is presented

1	<i>Citrobacter</i>	---MFTQYRKTLLAG TLALTFGLAAGNSLA AGFQPAQAGKLGAI VDPYGNAPLTAIVE LDSHVISDVKYTVHG KGEKGVPTVTVGKE	87
2	<i>Salmonella</i>	---MFDQYRKTLLAG AVALTCGLTAASTFA AVFKPAQAGKLGAV VDPYGNAPLTAIVE LDSHVISDVKYTVHG KGEKGVPTVTVGKD	87
3	<i>Enterobacter</i>	---MFHPYRKTLLSG TVALALGLFATGATA AGFQPAQAGKLGAI VDPYGNAPLTAIE LDSHVISDVKYTVHG KGEKGVPSVSVGKQ	87
4	<i>Klebsiella</i>	---MFDKYRKTLLVAG TVAITLGLSASGVMA AGFKPAPPAGQLGAV IVDYGNAPLTAIVD LDSHVISDVKYTVHG KGEKGVETSYVPGQE	87
5	<i>Pseudomonas</i>	-----M NAKTETPALPEGACL TASVPPARDEALLGDV LVNPLYRLAPLTAIIR DGGRTLAAHVRLVLR RGERGVDIAYDADR	76
6	<i>Campylobacter</i>	MRI.SKTLCLMALLAGS TLLAPNVLMAMGGPS GAKIDWQIQQIGAI KMNPYGLSPLTAIIM DNGYVLSDIKVTIIVP K-PNGQITSYNVNSK	89
7	<i>Eubacterium</i>	-----MSVKYSFE DHIIVNRQYEAEGEML KEFEAGNYTIANPLV KYNAYLVNPLSAVVC FHTEKETAITVTVLGR KTPQGNISHTPPKAK	83
*			
1	<i>Citrobacter</i>	SLATYDGIPIFGLYQ KFKANKVTVYKENG- ---KAMKDDYVVQT SAIVN---HYMDNRS ISDLQQTQVIVKAVG FEDRLYLNVNTHT--F	167
2	<i>Salmonella</i>	SLETYDGIPIFGLYQ KFKANNVTVYKENG- ---KAMKDDYVVQT SAIVN---HYMDNRS ISDLQQTQVIVKAVG FEDRLYLNVNTHT--F	167
3	<i>Enterobacter</i>	SLATYDGIPIFGLYQ KHANKVTVYETENG- ---KAMKEDYVIQT SAIVN---RYMDNRS ISDLQKTQVIVKAVG FEDRLYLNVNTHT--F	167
4	<i>Klebsiella</i>	SLKTYDGVPIFGLYQ KFKANKVTVWVKENG- ---KVNKDDYVVHT SAIVN---NYMDNRS ISDLQQTQVIVKAVG FEDRLYLNVNTHT--F	167
5	<i>Pseudomonas</i>	SLWTYGGIPIFGLYQ DHVYQVEVTVKLDG- ---ERVREYQIYA PAVR---LPVVAQK TAALPQVEPIKAVG FEHRLYLFNLLGD-D	156
6	<i>Campylobacter</i>	MAKTYGGIPIFGLYQ SYLNTYKVSYTKTAN GKSQKVIDEYKITT PGVSIIEP-SGSTDQR GTFPFENVKVLKMDPK FSDRLYLVNNAQKQ	178
7	<i>Eubacterium</i>	---KHVLPVIGLYS DYQNRVETRAVYRGE- -----SNIITIDV PDVFDGKVEIYSMDT TPEYLQDNIILVSPA GEDLAVGFVDYAGDAR	161
*			
1	<i>Citrobacter</i>	TPQGAEFHWGKEDK NAGILLDAGPAGALP FDIAPFTFVVDTEGE YRWWLDQD--TFYDG HDMDINKRGYLMGIR ETPRGTFPTAVQQQHW	255
2	<i>Salmonella</i>	TPQGAEFHWGKEDK NAGILLDAGPAARALP FDIAPHTFVVDTEGE YRWWLDQD--TFYDG HDMNINKRGYLMGIR RKRPRGTFPTAVQQQHW	255
3	<i>Enterobacter</i>	TPQGAEFHWGKEDK NAGILLDAGPAAGALP FDIAPFTFVVDTEGE YRWWLDQD--TFYDG HDMNINKRGYLMGIR ETPRGTFPTAVQQQHW	255
4	<i>Klebsiella</i>	TAQGSDLHWGKEDK NAGILLDAGPATGALP FDIAPFTFVVDTEGE YRWWLDQD--TFYDG RDRDINKRGYLMGIR ETPRGTFPTAVQQQHW	255
5	<i>Pseudomonas</i>	IPGGRAFKWNK---- ----LC--G-AAE VQDQGNWVADSNGD YRWWLDIE--QIHDS NRKRD--LGGTMRGQ QTRDGLIHWGQGTY	229
6	<i>Campylobacter</i>	SGKGSQSVWNN---- ----PVG--G-AME VDENSVFVIDTKGE IRWYFDND--KLNWV DNIYN--RGIIMGFH QNKDGLTWGFGQRY	252
7	<i>Eubacterium</i>	WHMTIPCVDYKRLK NGNLIMGSHRVIQMP YVMSGLYEISPCGKI YKEFRFLPGGYHDEF EMEDGNLLSLTDLDT SETVEDMCLVIDRNT	251
*			
1	<i>Citrobacter</i>	YEFDMGLQILADHKL PRGFLDASHESVETV NGTVLLRVGKRDYRK EDGLHVHTIRDQIEE VDKS-GRVVDVWDLT QILDPMRDALLGALD	344
2	<i>Salmonella</i>	YEFDMGQILADHKL PRGFLDASHESVETV NGTVLLRVGKRDYRK EDGLHVHTIRDQIEE VDKS-GRVVDVWDLT QILDPMRDALLGALD	344
3	<i>Enterobacter</i>	YEFDMGQVLEADHKL PRGFLDATHESVETV NGTVLLRVGKRNRYRK EDGLHVHTIRDQIEE VDKS-GRVVDVWDLT KILDPRDLALLGALD	344
4	<i>Klebsiella</i>	YEFDMGQVLEADHKL PRGFADATHESVETV NGTVLLRVGKSNRYRK DDGHHVHTIRDHILE VDKS-GRVVDVWDLT KILDPRDLALLGALD	344
5	<i>Pseudomonas</i>	SKYDLLGRRVWQRSL PDKFADPSHETRETA QCTYLLRVGTSYDRR PDGKRVRSIRDHIEE VNEA-GOVLDFWDLN QILDPRYRDLLETLG	318
6	<i>Campylobacter</i>	VKYDILGREIFNRKL PAAYIDFSHAMDNQK NGHLYLLRVASANTLR PDGKHVKTVDITVE VDEN-GNVVDVWRLY EILDYRSTIILKALD	341
7	<i>Eubacterium</i>	GEILKTDVYKFLDP KKVSKGSGVSDHDFW HINNAVYDKNTNSLT FSGRHIDSMVNI DVE TGEIWNIIIGDPEGW EEMQKYFFKPVGNPF	341
*			
1	<i>Citrobacter</i>	AGAVCVNVDLAHAG- ----QQAILEPDTF YGDALGVGAGRNVWAH VNSIAYDAKDDSIIL SSRHQG-VVKIGRDK QVKNWILAPSKGWNKA	427
2	<i>Salmonella</i>	AGAVCVNVDLAHAG- ----QQAILEPDTF YGDALGVGAGRNVWAH VNSIAYDAKDDSIIL SSRHQG-VVKIGRDK QVKNWILAPSKGWNKQ	427
3	<i>Enterobacter</i>	AGAVCVNVDLEHAG- ----QQAILEPDTF YGDALGVGAERNWAH VNPFIAYDAKDDSIIL SSRHQG-VVKIGSDK QVKNWILAPAKGWNKQ	427
4	<i>Klebsiella</i>	AGAVCVNVDLAHAG- ----QQAILEPDTF YGDALGVGPRNVWAH VNSIAYDAKDDSIIL SSRHQG-VVKIGRDK QVKNWILAPSKGWEKP	427
5	<i>Pseudomonas</i>	KAAIQLPDGVEKQDD RL---ANELAEGDLP FGDTPGVGTGRNVWAH VNAIDYDADDDSIIV SARHQG-VVKIGRDK AVKNWILAPSKGWPQR	404
6	<i>Campylobacter</i>	QGAVCLNIDASKAGK TLSDEELAKMDESOK FGDIAAGTIGRNVWAH VNSVDYDPSDDSIIL SSRHQSAVVKIGRDK KIKWILGAHKGWNKE	431
7	<i>Eubacterium</i>	GWQYEQHACVITPDG DVMCFDNIHYGSKNP EKYLAARDNYSRGRV YKINTDDMTIEQVWQ YGKURGAEEFFSPYIC NVEYVNEGHYVNVHSG	431
*			
1	<i>Citrobacter</i>	LASKLLKPVDDKQNA LKCDENG----KC ENTDFDFTYTQHTAW LSSKG-----TLTI FDNQDGRGLEQPALP TMKYSRFVEYKIDEX	505
2	<i>Salmonella</i>	LASKLLKPVDDHGNP LTCDENG----KC KDTDFDFTYTQHTAW LSSKG-----TLTV FDNQDGRGLEQPALP TMKYSRFVEYKIDEX	505
3	<i>Enterobacter</i>	LASKLLKPVDSKGNP LTCNENG----KC ENTDFDFTYTQHTAW LTKG-----TLTV FDNQDGRWLEQPALP TMKYSRFVEYKIDEX	505
4	<i>Klebsiella</i>	LASKLLKPVDSKGNP LTCNENG----LC ENSDFDFTYTQHTAW LSSKG-----TLTI FDNQDGRHLEQPALP TMKYSRFVEYKIDEX	505
5	<i>Pseudomonas</i>	LQDKVLPVYVGE-- ----FDNSWTQHTAW LTKG-----TLTV FDNQWGR-DFAPTKL TGNYSRAVEYRIDEA	466
6	<i>Campylobacter</i>	NKTVELQPVKNGKK LVCDYDYSKCPYEN DNGGFDFTWTQHTGW RIDSKSNKRYIYISV FDNQDARGAEGQAFPA SFQSGSRAYIYKIDQQ	521
7	<i>Eubacterium</i>	GIAVDTEGNPSEALG AFAKQDQGRLESITV EICDNKMLDLHVVP NYRGE-----KLKL YSDGINLELKGQKIL GEMGVTKFEDTEIPL	516
*			
1	<i>Citrobacter</i>	KGTVQQVWEYKERG YDFYSPITSVIEYQK DRDTMFGFGGSINLF DVGQPTIGKIN--E IDYKTEVKVEIDVL SDKPNQTHYRALLVYR	592
2	<i>Salmonella</i>	KGTVQQVWEYKERG YDFYSPITSVIEYQK DRDTMFGFGGSINLF DVGKPTVQKLN--E IDYKTEVKVEIDVL SDKPNQTHYRALLVH	592
3	<i>Enterobacter</i>	NGTVQPLWQYKERG YDFYSPITSVIEYQK DRDTIFGFGGSINLF EVGQPTIGKIN--E IDYKTKDVKVEIDVL SDKPNQTHYRALLVH	592
4	<i>Klebsiella</i>	KGTVQQVWEYKERG YDFYSPITSVIEYQA DRNTMFGFGGSINLF DVGQPTVQKLN--E IDYKTEVKVEIDVL SDKPNQTHYRALLVYR	592
5	<i>Pseudomonas</i>	KGTVEQVWEYKERG DEMYSPITSVAVYAP ETDQTFIYSASVNYL TPEKLTITVLN--E VRRGTQEVVVELKVH SRQPGSVGVRALVID	553
6	<i>Campylobacter</i>	NKTVELQPVKNGKK NEWFSPVTSITQYEP DKDISIMVYSATAGMA FDLKSGVSLGEPKPE IDEFNWGAKEPSVQI QFSGSGTGYQAMPFS	611
7	<i>Eubacterium</i>	EPSGEMLPESCNRARI EDEIDRFTFSPRFEK QQLVMLLEQGEVHV RYFISITAVPFLAMC CDTFLSDDRNTRTN INKTGLKGTYDLRV1	606
*			
1	<i>Citrobacter</i>	PQQMFK----- 598	
2	<i>Salmonella</i>	PTQMFK----- 598	
3	<i>Enterobacter</i>	PQQMFK----- 598	
4	<i>Klebsiella</i>	PQQMFK----- 598	
5	<i>Pseudomonas</i>	LAKAF----- 558	
6	<i>Campylobacter</i>	VDQAFNPKK----- 620	
7	<i>Eubacterium</i>	IDDKKYETGVITSC 620	

Fig. 3. CLUSTAL W multiple alignment between the ASST from *Citrobacter freundii* MB4-8242 and the ASSTs from *Salmonella typhimurium* ATCC-13311, *Enterobacter amnigenus* AR-37, *Klebsiella* K-36, *Pseudomonas putida*, and *Eubacterium rectale*, as well as the arylsulfatase from *Campylobacter jejuni*: Gaps introduced to optimally align the protein sequences are indicated by dashes. 11 conserved tyrosine residues are shown by asterisks.

in Table I. The phenyl sulfate esters are unique donor substrates for ASST, but PAPS, the donor substrate of mammalian sulfotransferase, is ineffective. With PNS as a donor, the best acceptor substrate of *C. freundii* MB4-

8242 ASST is α -naphthol, followed by phenol, resorcinol, acetaminophen, tyramine, and tyrosine. The acceptor specificity of the present enzyme is quite different from that of the enzyme purified from *Eubacterium* A-44 (Kim

Table 1. Acceptor Substrate Specificity of ASST

Acceptor	Relative activity (%) [*]	
	<i>Citrobacter freundii</i>	Recombinant
Phenol	100	100
<i>p</i> -acetaminophen	26.3 ± 4.3	16.2 ± 2.3
Tyramine	13.4 ± 2.1	15.1 ± 1.8
Tyrosine	10.4 ± 1.9	10.3 ± 2.2
α-Naphthol	147.4 ± 13.3	155.3 ± 7.7
Resorcinol	81.5 ± 7.7	82.8 ± 3.0

^{*}Acceptor substrate specificity was measured by using PNS as a donor substrate. Absorbance of assay mixture (PNS+Phenol) at 405 nm was taken as 100%.

et al., 1986) and is similar to that from *E. amnigenus* AR-37, *Klebsiella* K-36, and *Haemophilus* K-12 (Lee *et al.*, 1995).

The *astA* gene encodes a protein that possesses a unique enzymatic activity, and hence its coding sequence could be used as a reporter gene. This experiment used the standard spectrophotometric assay, which is based on the production of PNP. Moreover, the ASST activity could also be measured accurately by fluorometric assay. The sulfate of the 4-MUS is cleaved by the ASST, and 4-methylumbelliferone, a fluorescent product, is produced. This fluorometric assay is very specific, extremely sensitive, inexpensive and rapid. Therefore, it could be possible to develop the *astA* coding region as a reporter gene system.

ACKNOWLEDGEMENTS

This study was supported in part by the 2000 and 2001 BK21 projects for Medicine, Dentistry, and Pharmacy.

REFERENCES

- Altschul, S. F., Gish, W., Miller, W., Myers, E. W., and Lipman, D. J., Basic local alignment search tool. *J. Mol. Biol.*, 215, 403-410 (1990).
- Baek, M. C., Kim, S. K., Kim, D. H., Kim, B. K., and Choi, E. C., Cloning and sequencing of the *Klebsiella* K-36 *astA* gene, encoding an arylsulfate sulfotransferase. *Microbiol. Immunol.*, 40, 531-537 (1996).
- Baek, M. C., Kwon, A. R., Chung, Y. J., Kim, B. K., and Choi, E. C., Distribution of bacteria with the arylsulfate sulfotransferase activity. *Arch. Pharm. Res.*, 21, 475-477 (1998).
- Bardwell, J. C., McGovern, K., Beckwith, J., Identification of a protein required for disulfide bond formation *in vivo*. *Cell*, 67, 581-589 (1991).
- Birnboim, H. C., A rapid alkaline extraction method for the isolation of plasmid DNA. *Methods Enzymol.*, 100, 243-255 (1983).
- Dodgson, K. S., and Tudball, N., The metabolite of the sulfate group of potassium *p*-nitrophenyl[35S]sulfate. *Biochem. J.*, 74, 154-159 (1960).
- Goldberg, S. L., Nanduri, V., Cino, P. M., and Patel, R., Purification, cloning and characterization of an arylsulfotransferase gene from the anaerobic bacterium *Eubacterium rectale* IIIH. *J. Ind. Microbiol.* 25, 305-309 (2000).
- Harley, C. B., and Reynolds, R. P., Analysis of *E. coli* promoter sequences. *Nucleic Acids Res.*, 15, 2343-2361 (1987).
- Holt, J. G., N. R. Krieg, P. H. A. Sneath, J. T. Staley, and S. T. Williams., *In Bergey's manual of determinative bacteriology*, The Williams & Wilkins Co., Baltimore, Md., (9th eds.), (1994).
- Kahnert, A., and Kertesz, M. A., Characterization of a sulfur-regulated oxygenative alkylsulfatase from *Pseudomonas putida* S-313. *J. Biol. Chem.*, 275, 31661-31667 (2000).
- Kang, J. W., Kwon, A. R., Kim, D. H., and Choi, E. C., Cloning and sequencing of the *astA* gene encoding arylsulfate sulfotransferase from *Salmonella typhimurium*. *Biol. Pharm. Bull.*, 24, 570-574 (2001).
- Kim, D. H., Konishi, L., and Kobashi, K., Purification, characterization and reaction mechanism of novel arylsulfotransferase obtained from an anaerobic bacterium of human intestine. *Biochem. Biophys. Acta.*, 872, 33-41 (1986).
- Kim, D. H., and Kobashi, K., The role of intestinal flora in metabolism of phenolic sulfate esters. *Biochem. Pharmacol.*, 35, 3507-3510 (1986).
- Kim, D. H., and Kobashi, K., Kinetic studies on a novel sulfotransferase from *Eubacterium* A-44, a human intestinal bacterium. *J. Biochem.*, 109, 45-48 (1991).
- Kim, D. H., Kim, H. S., and Kobashi, K., Purification and characterization of novel sulfotransferase obtained from *Klebsiella* K-36, an intestinal bacterium of rat. *J. Biochem.*, 112, 456-460 (1992).
- Kim, D. H., Kim, H. S., Imamura, L., and Kobashi, K., Kinetic studies on a sulfotransferase from *Klebsiella* K-36, a rat intestinal bacterium. *Biol. Pharm. Bull.*, 17, 543-545 (1994).
- Kobashi, K., and Kim, D. H., A novel sulfotransferase sulfates tyrosine-containing peptides and proteins. *Biochem. Biophys. Res. Commun.*, 140, 38-42 (1986).
- Kwon, A. R., Oh, T. G., Kim, D. H., and Choi, E. C., Molecular cloning of the arylsulfate sulfotransferase gene and characterization of its product from *Enterobacter amnigenus* AR-37. *Protein Expression Purif.*, 17, 366-372 (1999).
- Lee, N. S., Kim, B. T., Kim, D. H., and Kobashi, K., Purification and reaction mechanism of arylsulfate sulfotransferase from *Haemophilus* K-12, a mouse intestinal bacterium. *J. Biochem.*, 118, 796-801 (1995).
- Muramatsu, R., Nukui, E., Sukesada, A., Misawa, S., Komatsu, Y., Okayama, T., Wada, K., Morikawa, T., Hayashi, H., and Kobashi, K., Enzymatic O-sulfation of tyrosine residues in hirudins by sulfotransferase from *Eubacterium* A-44. *Eur. J. Biochem.*, 223, 243-248 (1994).
- Roy, A. B., *In Sulfotransferases, Sulfation of Drugs and*

- Related Compounds*, Mulder, G. J., (Eds.), CRC Press, Boca Raton, FL., pp. 131-185 (1981).
- Sanger, F., Nicklen, S., and Coulson, A. R., DNA sequencing with chain-terminating inhibitors. *Proc. Natl. Acad. Sci. USA.*, 74, 5463-5467 (1977).
- Sekura, R. D., and Jakoby, W. B., Phenol sulfotransferase. *J. Biol. Chem.*, 254, 5658-5663 (1979).
- Sekura, R. D., Duffel, M. W., and Jakoby, W. B., Arylsulfotransferase. *Methods Enzymol.*, 77, 197-206 (1981).
- Shine, J., and Dalgarno, L., The 3'-terminal sequence of *Escherichia coli* 16S ribosomal RNA: complementarity to nonsense triplets and ribosome binding site. *Proc. Natl. Acad. Sci. USA.*, 71, 1342-1346 (1974).
- Spencer, B., Endogenous sulfate acceptors in rat liver. *Biochem. J.*, 77, 294-304 (1960).
- Thompson, J. D., Higgins, D. G., and Gibson, T. J., CLUSTAL W: Improving the sensitivity of progressive multiple sequence alignment through sequence weighting, position-specific gap penalties and weight matrix choice. *Nucleic Acids Res.*, 22, 4673-4680 (1994).
- Yao, R., and Guerry, P., Molecular cloning and site-specific mutagenesis of a gene involved in arylsulfatase production in *Campylobacter jejuni*. *J. Bacteriol.*, 178, 3335-3338 (1996).