# Imidazole Acetol Phosphate Aminotransferase in *Zymomonas mobilis*: Molecular Genetic, Biochemical, and Evolutionary Analyses

WEI GU, GENSHI ZHAO,† CHRIS EDDY,‡ AND ROY A. JENSEN\*

Department of Microbiology and Cell Science, University of Florida, Gainesville, Florida 32611

Received 6 October 1994/Accepted 13 January 1995

hisH encodes imidazole acetol phosphate (IAP) aminotransferase in Zymomonas mobilis and is located immediately upstream of tyrC, a gene which codes for cyclohexadienyl dehydrogenase. A plasmid containing hisH was able to complement an Escherichia coli histidine auxotroph which lacked the homologous aminotransferase. DNA sequencing of hisH revealed an open reading frame of 1,110 bp, encoding a protein of 40,631 Da. The cloned hisH product was purified from E. coli and estimated by sodium dodecyl sulfate-polyacrylamide gel electrophoresis to have a molecular mass of 40,000 Da. Since the native enzyme had a molecular mass of 85,000 Da as determined by gel filtration, the active enzyme species must be a homodimer. The purified enzyme was able to transaminate aromatic amino acids and histidine in addition to histidinol phosphate. The existence of a single protein having broad substrate specificity was consistent with the constant ratio of activities obtained with different substrates following a variety of physical treatments (such as freeze-thaw, temperature inactivation, and manipulation of pyridoxal 5'-phosphate content). The purified enzyme did not require addition of pyridoxal 5'-phosphate, but dependence upon this cofactor was demonstrated following resolution of the enzyme and cofactor by hydroxylamine treatment. Kinetic data showed the classic ping-pong mechanism expected for aminotransferases. K<sub>m</sub> values of 0.17, 3.39, and 43.48 mM for histidinol phosphate, tyrosine, and phenylalanine were obtained. The gene structure around hisH-tyrC suggested an operon organization. The hisH-tyrC cluster in Z. mobilis is reminiscent of the hisH-tyrA component of a complex operon in Bacillus subtilis, which includes the tryptophan operon and aroE. Multiple alignment of all aminotransferase sequences available in the database showed that within the class I superfamily of aminotransferases, IAP aminotransferases (family IB) are closer to the I $\gamma$  family (e.g., rat tyrosine aminotransferase) than to the I $\alpha$  family (e.g., rat aspartate aminotransferase or E. coli AspC). Signature motifs which distinguish the IAP aminotransferase family were identified in the region of the active-site lysine and in the region of the interdomain interface.

Histidine biosynthesis requires imidazole acetol phosphate aminotransferases (IAP aminotransferases) to catalyze the formation of histidinol phosphate by transamination between imidazole acetol phosphate and glutamate (40). Although a number of aminotransferases are not essential for growth because of the enzymatic backup attributed to the overlapping specificities of the intracellular repertoire of aminotransferases (18), IAP aminotransferase is essential in all organisms studied. Thus, mutants which lack IAP aminotransferase activity have been shown to be auxotrophic for histidine in *Escherichia coli* (13), *Bacillus subtilis* (38), *Corynebacterium glutamicum* (1a), *Halobacterium volcanii* (6), and *Streptomyces coelicolor* (21).

Genes responsible for histidine biosynthesis form a single operon in Salmonella typhimurium and E. coli (40). In B. subtilis, all genes of histidine biosynthesis also map together, except for hisH (28). The latter gene (called hisC in E. coli) encodes IAP aminotransferase and is flanked upstream by the tryptophan operon and downstream by tyrA (encoding prephenate dehydrogenase) and aroE (encoding enolpyruvylshikimate-5-phosphate synthase). The arrangement is 5'-aroFaroB-aroH-trpE-trpD-trpC-trpF-trpB-trpA-hisH-tyrA-aroE-3' (2, 18). A "supraoperon" was originally proposed (32) to explain the coordinate derepression of *hisH* and *tyrA* genes following derepression of the tryptophan operon. Later it was found that complex transcription with multiple transcription start sites and multiple readthrough events occurs within the entire *aroF*-*aroE* segment (2, 16). The latter system illustrates features of a complex operon which can be defined as a multicomponent transcriptional unit whose encoded proteins function in different biochemical pathways. Internal promoters which can facilitate the discoordinate synthesis of operon gene products commonly exist.

B. subtilis IAP aminotransferase can provide a backup function in tyrosine and phenylalanine biosynthesis (26), a role normally fulfilled by an aromatic aminotransferase specified by aroJ. Thus, hisH mutants are auxotrophic for L-histidine, while aroJ mutants remain prototrophic. However, hisH aroJ double mutants are auxotrophic for histidine, phenylalanine, and tyrosine. As expected from the in vivo results, purified IAP aminotransferase from B. subtilis was shown to transaminate phenylpyruvate and p-hydroxyphenylpyruvate in vitro (38). In view of the foregoing physical and functional relationship between hisH and tyrA in B. subtilis, it is intriguing that tyrC in Zymomonas mobilis was found to be adjacent to hisH (42). Does HisH function also overlap both histidine and aromatic amino acid biosynthesis in Z. mobilis? Z. mobilis is a gramnegative organism belonging to the beta subdivision of the proteobacteria (11). At the outset, two evolutionary differences were already known to distinguish Z. mobilis from B. subtilis with respect to aromatic amino acid biosynthesis. B. subtilis possesses a narrow-specificity prephenate dehydrogenase (specified by tyrA) which is feedback inhibited by L-tyrosine, while Z. mobilis possesses a broad-specificity cyclohexadienyl

<sup>\*</sup> Corresponding author. Mailing address: Department of Microbiology and Cell Science, P.O. Box 110700, University of Florida, Gainesville, FL 32611-0100. Phone: (904) 392-9677. Fax: (904) 392-8479.

<sup>&</sup>lt;sup>†</sup> Present address: Department of Microbiology, University of Texas Medical School, Houston, TX 77030.

<sup>‡</sup> Present address: ZeaGen, Fort Collins, CO 80525.

Strain or plasmid	Genotype or description	Source or reference
E. coli K-12		
JM83	ara $\Delta$ (proAB-lac) rpsL thi-1 $\phi$ 80dlacZ $\Delta$ M15	Gibco-BRL
AT2471	thi-1 tyrA4 relA1 $\lambda^-$ spoT	CGSC <sup>a</sup> 4510
UTH780	hisC780 malA1 ( $\lambda^{r}$ ) xyl-5 rpsL145 $\lambda^{-}$	CGSC 5954
DH5a	supE44 ΔlacU169 (φ80dlacZΔM15) hsdR17 recA1 endA1 gyrA96 thi-1 relA1	Gibco-BRL
Z. mobilis CP4	Prototroph	29
Plasmids	1	
pFBA9	Cosmid carrying Z. mobilis trpFBA in a 27.4-kb EcoRI fragment	9
pC2	Cosmid carrying Z. mobilis trpC in a 29.4-kb EcoRI fragment	9
pC5	Cosmid carrying Z. mobilis trpC in a 29.1-kb EcoRI fragment	9
pF5.8	pUC18 carrying Z. mobilis trpBA in a 5.8-kb EcoRI fragment	9
pF20	Cosmid carrying Z. mobilis trpF in a 34.6-kb EcoRI fragment	9
pUC18	Ap <sup>r</sup> <i>lacI</i> 'POZ	Gibco-BRL
pUC19	Ap <sup>r</sup> <i>lacI</i> 'POZ	Gibco-BRL
pGEM-5Zf(+)	Ap <sup>r</sup> lacI'POZ	Promega
pJZ5	Original clone of <i>tyrC</i> isolated from Z. <i>mobilis</i> CP4 library	42
pJZ5b	Derivative of pJZ5 generated by removal of a 3.2-kb <i>Hin</i> dIII fragment	42
pJG1	1.8-kb <i>Eco</i> RI- <i>Nsi</i> I fragment of pJZ5 subcloned into pUC18 at <i>Eco</i> RI- <i>Pst</i> I site	This study
pJG2	1.1-kb <i>NcoI-StuI</i> fragment of pJG1 subcloned into pGEM-5Zf(+) at <i>NcoI-Eco</i> RV site	This study
pJG3	2.5-kb <i>NcoI-NcoI</i> fragment of pJZ5 subcloned into pGEM-5Zf(+) at <i>NcoI</i> site	This study
pJG4-a	6-kb <i>Eco</i> RI- <i>Hin</i> dIII fragment of pJZ5b subcloned into pUC18 at <i>Eco</i> RI- <i>Hin</i> dIII site	This study
pJG4-b	6-kb <i>Eco</i> RI- <i>Hin</i> dIII fragment of pJZ5b subcloned into pUC19 at <i>Eco</i> RI- <i>Hin</i> dIII site	This study

<sup>a</sup> CGSC, Escherichia coli Genetic Stock Center, Yale University.

dehydrogenase (specified by *tyrC*) which is insensitive to feedback inhibition (42). Second, in contrast to *B. subtilis*, *Z. mobilis* tryptophan pathway genes are not organized as a single operon (9). Therefore, if the *hisH tyrC* region is part of a larger, *B. subtilis* type of complex operon, only one or several of the *trp* genes could be present.

## MATERIALS AND METHODS

**Bacterial strains, plasmids, and media.** Bacterial strains and plasmids used in this study are listed in Table 1. Luria-Bertani medium was used as enriched medium (33). The minimal medium used for *E. coli* strains was the M-9 formulation (22). Where indicated, ampicillin was added to media at 100  $\mu$ g/ml, tetracycline was added at 10  $\mu$ g/ml, thiamine was added at 17  $\mu$ g/ml, and amino acids were added at 50  $\mu$ g/ml. Medium was solidified with 2% (wt/vol) agar.

**DNA manipulation.** The subcloning techniques including plasmid purification, restriction enzyme digestion, ligation, and transformation were conducted by standard methods (22). Restriction enzymes, ligase, and calf intestine alkaline phosphatase were purchased from Gibco-BRL or Promega and were used according to manufacturer's instructions. Electroporation was conducted by using equipment from Invitrogen Corporation. Southern blot hybridization was carried

out under stringent conditions by use of a biotinylated probe according to the instructions of Promega.

**DNA sequencing and data analysis.** Plasmid pJG2 was purified by the method recommended by Applied Biosystems, Inc. (1). Primers were made by the DNA Synthesis Laboratory of the Interdisciplinary Center of Biotechnology Research (ICBR) at the University of Florida. Sequencing was performed by the DNA Core Facility of the ICBR. The nucleotide sequence and the deduced amino acid sequence were analyzed by using the updated version of sequence analysis software package offered by Genetics Computer Group, Inc. (7).

**RNA isolation and analysis.** Total RNA from cultures of *Z. mobilis* or *E. coli* UTH780 (pJG3) was isolated as described by Bialkowska-Hobrzanska et al. (3). Contaminating DNA was removed by treatment with RNase-free DNase I from Gibco-BRL. Northern (RNA) blot hybridization was carried out by using a biotinylated probe according to instructions provided by Promega. Determinations of RNA size were based on migration of known RNA standards (0.24- to 9.5-kb RNA ladder from Gibco-BRL) as described by Maniatis et al. (22).

The 5' termini of transcripts were determined by primer extension analysis essentially as described by Tsui et al. (36), with some modifications. Two 30-base oligonucleotide primers which anneal to hisH transcripts (5'-CCCCGGAAT GĂTAGGCGCÂATACTGTCGAT-3') at positions 366 to 395 (Fig. 2) and tyrC (5'-GGAACCGATCAGTCCTAATCCGATAATGGC-3') at positions 1461 to 1490 were synthesized by the DNA Synthesis Laboratory of the ICBR at the University of Florida. The primers were labeled on the 5' end with  $[\gamma^{-32}P]ATP$ (Amersham) and T4 polynucleotide kinase (Promega), and the unincorporated radioisotope was removed by two rounds of ethanol precipitation. The labeled primers were hybridized to various amounts of total RNA, and reverse transcriptase (Gibco-BRL) was used for primer extension. The cDNA obtained was dissolved in 2 µl of Tris-EDTA buffer and 4 µl of formamide loading buffer and was electrophoresed adjacent to a sequence ladder. The sequence ladder was generated by using the same primer,  $\left[\alpha^{-35}S\right]ATP$  (Amersham), purified recombinant plasmid DNA containing the region of interest, and a sequencing kit from Promega.

**Crude extract preparation and enzyme assays.** Bacterial cultures were grown at  $37^{\circ}$ C with vigorous shaking in minimal medium supplemented with ampicillin. The cells were harvested in the late exponential phase of growth by centrifugation and resuspended in 50 mM potassium phosphate buffer (pH 7.0). The cells were broken by sonication, using a Ultratip Labsonic System (Lab-Line Instruments, Inc., Melrose Park, III.). The suspension was centrifuged at  $150,000 \times g$  for 65 min at 4°C. The supernatant fraction was collected and passed through a DG-10 Sephadex column (1.5 by 5.5 cm). The resulting desalted preparation is designated the crude extract.

IAP aminotransferase was assayed by mixing enzyme,  $\alpha$ -ketoglutarate (final concentration, 10 mM), and histidinol phosphate (final concentration, 8.3 mM) in potassium phosphate buffer (50 mM, pH 7.0), incubating the mixture at 37°C for the designated amount of time, and then measuring the  $A_{280}$  of enolized imidazole acetol phosphate formed enzymatically (23). p-Hydroxyphenylpyruvate aminotransferase (HAT) and phenylpyruvate aminotransferase (PAT) activities were assayed in 50 mM potassium phosphate buffer (pH 7.0), by using 10 mM a-ketoglutarate and 10 mM L-tyrosine or 10 mM L-phenylalanine as substrates, and monitoring the  $A_{331}$  for p-hydroxyphenylpyruvate or the  $A_{320}$  for phenylpyruvate (39). Other aminotransferase activities were assayed by monitoring the formation of product amino acids by use of high-performance liquid chromatography (HPLC) (4). Prephenate dehydrogenase activity was assayed by monitoring the appearance of NADH on a spectrophotofluorometer (excitation at 340 nm and emission at 460 nm) at 37°C with saturating concentrations of NAD<sup>+</sup> (1 mM) and prephenate (2 mM) (42). Extinction coefficients of 5,310 for imidazole acetol phosphate (23), 17,500 for phenylpyruvate (39), and 32,850 for p-hydroxyphenylpyruvate (39) were used to calculate the concentration of each product. Protein concentration was determined by the method of Bradford (5).

Purification of the cloned IAP aminotransferase from E. coli UTH780. E. coli UTH780 carrying plasmid pJG3 was grown in 6 liters of minimal medium supplemented with ampicillin and thiamine at 37°C in a gyratory shaker up to the late exponential phase of growth. The cell pellet was harvested by centrifugation and washed once with buffer A (50 mM potassium phosphate buffer [pH 7.0], dithiothreitol, 1.0 mM, 1.4 mM β-mercaptoethanol, 0.05 mM pyridoxal 5'-phosphate [PLP]), then resuspended in the same buffer, and disrupted by sonication. After centrifugation (150,000  $\times$  g for 65 min at 4°C), the supernatant was dialyzed overnight against buffer A and then applied to a DEAE-cellulose column (2.5 by 55 cm) equilibrated with buffer A. The column was washed with 700 ml of buffer A and then eluted with 1,800 ml of buffer A containing a linear gradient of KCl from 0 to 200 mM. Fractions of 8 ml were collected, and those showing high IAP aminotransferase activity were pooled and concentrated by means of an Amicon PM-10 membrane. The concentrated preparation was dialyzed against buffer A overnight and applied to a hydroxylapatite column (1.5 by 51 cm) equilibrated with buffer A. After being washed with 400 ml of buffer A, the column was eluted with 1,000 ml of buffer A containing a linear gradient of potassium phosphate from 10 to 400 mM. Fractions of 5 ml were collected, and those showing high IAP aminotransferase activity were pooled and concentrated as described above.

**PAGE.** Native polyacrylamide gel electrophoresis (PAGE), using the foregoing preparation, was performed as described by Orr et al. (27). After electrophoresis, a part of the gel was stained with Coomassie blue. The portion of the



FIG. 1. Physical map and subclone analysis of Z. mobilis hisH (shaded arrow) and tyrC (open arrow). Restriction site locations are indicated at the upper left. Plasmid DNA in the pJZ5 construct is depicted by heavy bars. The orientation of the *lac* promoter in the vector is shown for the pJG4a and pJG4b subclones, and that of all of the other constructs is the same as shown for pJG4a. The abilities of the various constructs to complement *tyrA* or *hisC* auxotrophs are shown at the right, as are specific activities (nanomoles of product per minute per milligram) for HisH in some of the constructs. ND, not done.

unstained gel which corresponded to the bands visualized on the stained gel were excised separately, minced in buffer A, and then incubated at  $37^{\circ}$ C for 1 h. The gel was removed by centrifugation at  $10,000 \times g$  by means of a microfilterfuge tube (Rainin Instrument Co., Inc., Woburn, Mass.). The preparation recovered could be used for further enzyme analysis.

Amino acid sequencing of the cloned IAP aminotransferase. The purified enzyme preparation was denatured by sodium dodecyl sulfate (SDS) and subjected to PAGE (19). The protein was transferred to a polyvinylidene difluoride membrane by a protein miniblotting apparatus and sequenced by using an Applied Biosystems model 407A protein sequencer with on-line 120A phenylthiohydantoin analyzer at the Protein Core Facility of the ICBR at the University of Florida.

**Molecular mass determination.** The molecular mass of the native enzyme was estimated by gel filtration using a Sephadex G-200 column (2.5 by 98 cm) which had been previously equilibrated with buffer A and eluted with the same buffer. The column was calibrated by chymotrypsinogen ( $M_{\rm r}$ , 25,000), bovine serum albumin ( $M_{\rm r}$ , 66,000), alcohol dehydrogenase ( $M_{\rm r}$ , 150,000), and  $\beta$ -amylase ( $M_{\rm r}$ , 200,000). SDS-PAGE was carried out to determine the subunit molecular mass of the enzyme.  $\alpha$ -Lactalbumin ( $M_{\rm r}$ , 14,400), soybean trypsin inhibitor ( $M_{\rm r}$ , 20,100), carbonic anhydrase ( $M_{\rm r}$ , 30,000), ovalbumin ( $M_{\rm r}$ , 43,000), bovine serum albumin ( $M_{\rm r}$ , 67,000), and phosphorylase ( $M_{\rm r}$ , 94,000) were used as molecular mass standards.

Biochemicals and chemicals. Histidinol phosphate,  $\alpha$ -ketoglutarate, PLP,  $\beta$ -mercaptoethanol, ampicillin, tetracycline, thiamine, Sephadex G-200, and amino acids were purchased from Sigma Chemical Company; dithiothreitol was from Research Organics, Inc.; DEAE-cellulose was from Whatman; hydroxylapatite was from Bio-Rad. Molecular mass standards for gel filtration and for SDS-PAGE were from Sigma and Pharmacia Fine Chemicals, respectively. Growth medium components and agar were from Difco.

Nucleotide sequence accession number. The nucleotide sequence reported in this communication has been assigned GenBank accession number L36343.

### RESULTS

**Identification of the gene upstream of** *tyrC***.** Zhao et al. (42) identified a partial open reading frame of 539 bp located directly upstream of *tyrC* (encoding cyclohexadienyl dehydrogenase) in *Z. mobilis*. Amino acid sequence comparison with other sequences in the database indicated that the upstream gene might encode IAP aminotransferase. The original *tyrC* clone (pJZ5) isolated (42) contained an intact *hisH* gene, since functional complementation (Fig. 1) of *E. coli* UTH780, a *hisC* mutant (lacking IAP aminotransferase activity), could be readily demonstrated. Plasmids purified from the ampicillinresistant transformants had the same size and banding pattern (supercoiled, circular, and linear) as pJZ5 on agarose gel, and

a second transformation of *E. coli* UTH780 with these purified plasmids again conferred histidine prototrophy.

Subcloning experiments shown in Fig. 1 localized *hisH* to the *NcoI-NcoI* fragment. The inability of plasmid pJG1 to complement the histidine auxotroph *E. coli* UTH780 indicated that cleavage at the *NsiI* site disrupted the integrity of *hisH*. Southern blot hybridization showed that a labeled 0.6-kb *StuI-StuI* fragment of pJG3 hybridized with a 0.6-kb fragment of *Z. mobilis* chromosomal DNA digested with *StuI* but did not hybridize with a similar digest of *E. coli* chromosomal DNA (data not shown).

Z. mobilis IAP aminotransferase expression in E. coli was measured directly (Fig. 1). Histidine auxotroph UTH780 exhibited the expected absence of IAP aminotransferase activity. Although the specific activity measured in UTH780 (pJZ5) progressively increased as insert size decreased in subclones UTH780 (pJG4a) and UTH780 (pJG3), the plasmids used have different promoters or cloning ends. In addition, specific activities have not been normalized to plasmid copy numbers. Since the insert of UTH780 (pJG4b) has an orientation opposite that of the *lac* promoter, a native promoter ahead of *hisH* appears to be recognized by *E. coli*.

**Nucleotide sequence.** The complete nucleotide sequence of *Z. mobilis hisH* along with the flanking upstream region is shown in Fig. 2. The structural gene, 1,110 bp in length, begins at codon ATG and terminates at codon TAA. The N-terminal amino acid sequence, obtained by using a sample cut from the major band on SDS-PAGE, was determined to be MTAAPE, a sequence which was identical to that deduced from the nucleotide sequence. The deduced amino acid sequence yields a protein of 370 residues with a molecular mass of 40,631 Da. The *hisH* terminator codon is immediately adjacent to the GTG start codon of *tyrC*, leaving no intercistronic space.

The G+C content of *hisH* (50.1%) falls into the range of 47 to 50% for the *Z. mobilis* genome (25). The 5' region corresponding to the untranslated transcript is exceedingly A+T rich. The codon usage of *hisH* exhibits a bias different from that of *Z. mobilis* genes which encode high-abundance proteins (29). For instance, six codons (GGG, AGG, AGA, ATA,

атааааататаааааасадтттатодстааааааатоддсааттттастстаттаддааатассааааатаааат	120
ттааттттттаталалаласалдалаласасалсссалталассалатсалтстаттттстдсттддассалаладстдталалтттттаталаттатдатасалалтсадтсастт	240
+1a +1b TGGTTCTAAAAAAAAAAAAAAAAAAAATATCGAGGATAGCCTTTAACGCGCGCCATTATTATCCGGATTTTTATC <u>AAAGAG</u> CTATTATGACTGCTGCACCTGAACTTCGCCCGAAAT RBS M T A P B L R P K S	360 11
CTTGGATCGACAGTATIGCGCCTTACATTCCGGGGTCATCAAAGACGCTGGGATGGCCGTCCGGCTGTCAAACTTTCATCTAATGAAAACCCTTTGGGAACCAGCCTAAAGGCGAAAGAAG WIDSIAPYIPGSSKTLDGRPAVKLSSNBNPLGTSLKAKEA	480 51
CCTATCGGGAAGCCATAGACTCCCTTTCCCTGTATCCTGACAGTGGGGCACCGCTCTGCGTGAGGCCATCGGGGCTTGCTATAATCTTGATCGGCTCGGATTATTCACGGCACCGGAT Y R E A I D S L S L Y P D S G A T A L R E A I G A C Y N L D P A R I I H G T G S	600 91
CGGATGAAATCCTGCATTTGGCCGCCGGGGCCTATGCCGGACAAGATGACGAGGGTTTATATCCTCGCTAGGTTTTCAGTTTACCCGCTGGCGGCGGCGCGGGGGGGG	720 131
TAGAAGCACCAGATGATGATGATGGTTGTTGTTGATGCCTTGTTGAAGGCTGTCACGCCACGGACACGGGTTGTCTTTATGCCAATGCCGAATAACCCGACGGGAACATGGATTACCC E A P D D D Y R C S V D A L L K A V T P R T R V V F I A N P N N P T G T W I T R Stur	840 171
GTGCAGAAGTTGAAAAACTGCATAACGGCCTTCCCCGGTAACTGCTTATTGGTTATCGATC <del>AGGCC</del> TATGCTGAAATATCTTGATCGTGAATGTGATGGCGCGTTAGCACTCGCCAAAA A E V E K L H N G L P R N C L L V I D Q A Y A E Y L D P E C D D G A L A L A K N	960 211
ACACAAAAAATGTGCTGGTAACGCGGACATTTTCTAAAATCTACGGATTGGCCGCGGAAAGAATGGCCTGGGCTTATGCCTGGCCTGAAATCATTGATGCCCTTAACGGGAACGGGCAC T K N V L V T R T F S K I Y G L A A E R I G W A Y A C P E I I D A L N R I R A P	1080 251
CCTTCAATGTCACGATTGCGGGACAAAAGGCAGCGGCGGCGCGCCTTTGGAAGATCAGGCGTTCATCCAGAACAGCTTCAAAAAAGGCGCGGGGCGGGC	1200 291
CGCTTTTAAGTAATGTTGGCATTCGTGTTATCCCGTCCTCGCCAATTTTACTCTACTGCTGTTGAAGGCAGCCTGACGGCTAAAACCGCCTATAAAGCCTTGATGGATCACGGCTATA L L S N V G I R V I P S S A N F T L L L F E G S L T A K T A Y K A L M D H G Y T Nsii RBS F	1320 331
CCACCCGTTGGTTGCCGGGACAGCGCCTTCCTC $\overline{ATGCAT}$ TACGTATCACTATCGCGCAGTGAAAAACATATGCAGGATGTCGCTGGTATTTTAACTTCCTTGGTT $\overline{AGGCAG}$ GCGCTCTAA $\dot{G}$ T R W L P G Q R L P H A L R I T I G S E K H M Q D V A G I L T S L V R Q A L * M $\longrightarrow tyrC$ StuI	1440 369/1
tgacegtetttaagcatattgccattateggattaggactgateggtteetetgeggcaegggcaacaaagggeetattgteetgatgtaaeggteagtetetatgaeaaaagegaattg TVFKHIAIIGLGLGLIGSSAARATKAYCPDVTVSLYDKSEFV	1560 41
gatattcacccttaaaatttgccagtattattcagccgtcagatatttttctgacggcttttttactgattaaagcgattgatcgcagcatgaacaagcgcttcggcttccttc	2425 293

FIG. 2. Nucleotide sequence of the Z. mobilis hisH region. Probable ribosome binding sites (RBS) just upstream of hisH and tyrC are underlined. Converging arrowheads mark a probable terminator structure following tyrC. The nucleotide sequence of tyrC (beginning at nucleotide 1440) is given in lowercase letters, and the dotted line indicates the portion of the tyrC sequence (nucleotides 1561 to 2305) not shown here. Nucleotide and amino acid numbers are given at the far right across from the appropriate line. Transcriptional start sites upstream of hisH indicated by primer extension analysis are labeled +1a and +1b. Possible transcriptional start sites for tyrC originating from an internal promoter within hisH are indicated with vertical arrowheads.

CGA, and CTA) are never used, and three codons (GTA, CGG, and ACA) are rarely (<0.1 mol%) used in high-abundance proteins but are used to a significantly greater extent in the low-abundance proteins. This is consistent with expectations that the low-abundance tRNA species in the cell may limit gene expression.

**Transcript analysis.** Northern blot analysis was performed as described in Materials and Methods. RNA was isolated from *E. coli* UTH780 (pJG3) and *Z. mobilis.* Two DNA probes were generated by PCR. One 511-bp probe spanned nucleotides 335 to 846 within *hisH* (Fig. 2), and another 441-bp probe spanned nucleotides 1874 to 2315 within *tyrC*. In either case, two transcripts of different sizes were detected (data not shown). One transcript of about 2 kb would be sufficient to read through both structural genes. A smaller, 1.1-kb transcript may correspond to *tyrC* driven from an internal promoter within *hisH*.

Primer extension analysis was carried out with a primer complementary to the 5' portion of hisH. Two possible transcription start sites for hisH were identified by using RNA from either E. coli UTH780 (pJG3) or Z. mobilis (as marked on Fig. 2). Since we observed numerous faint bands that indicated in vivo RNase processing, RNase protection experiments will be required to confirm positions marked +1a and +1b as transcription start sites. Primer extension analysis was also carried out with a primer complementary to the 5' portion of tyrC to determine possible transcription start sites for tyrC which originate from an internal promoter upstream of the NsiI restriction site within hisH. Four bands of nearly equal intensity correspond to the possible start sites indicated in Fig. 2. One or more of these bands could be transcript ends caused by in vivo RNase processing of the mRNA originating from the primary promoter ahead of hisH.

Inseparability of the cloned IAP aminotransferase and aromatic aminotransferase activities. The Z. mobilis enzyme was purified from E. coli UTH780 (pJG3) by fractionation using two chromatography steps (Table 2). One major protein band and one minor protein band were visualized by SDS-PAGE (Fig. 3). The molecular mass of the major band was estimated to be 40,000 Da, in good agreement with that deduced from DNA sequence. The molecular mass of the native enzyme is 85,000 Da, as determined by gel filtration on Sephadex G-200. The native enzyme is therefore a homodimer.

The fractions which contained IAP aminotransferase activity also exhibited a coinciding peak of aromatic aminotransferase activity (data not shown). The purified enzyme preparation from the hydroxylapatite column was loaded on a native polyacrylamide gel. The protein corresponding to the top band (see Materials and Methods) was eluted and found to exhibit IAP aminotransferase, PAT, and HAT activities. N-terminal amino acid sequencing also showed this band to correspond to IAP aminotransferase. A sample of the same eluate gave a single

TABLE 2. Purification of the cloned Z. mobilis IAP aminotransferase from E. coli UTH780 (pJG3)

Step	Total protein (mg)	Sp act ( $\mu$ mol min <sup>-1</sup> mg <sup>-1</sup> )		IAT <sup>a</sup> /PAT ratio	Purification factor
		IAT	PAT	Tatio	Idetoi
Crude extract	818	0.52	2.15	0.24	1
DEAE-cellulose	95.2	3.54	8.67	0.41	6.8
Hydroxylapatite	20.9	18.24	29.93	0.64	35

<sup>*a*</sup> IAT, IAP aminotransferase.



FIG. 3. SDS-PAGE of the protein samples obtained after the purification steps specified in Table 2 for the cloned IAP aminotransferase from *E. coli* UTH780. A 12% gel was used and stained with Coomassie blue. Lane 1, molecular weight standards; lane 2, crude extract; lane 3, after the DEAE-cellulose chromatography step; lane 4, after the hydroxylapatite chromatography step.

40,000-Da band on SDS-PAGE, thus indicating successful removal of the minor protein from the sample.

To confirm further that the minor protein was not responsible for the aromatic aminotransferase activity observed, N-terminal amino acid sequencing of the minor protein was carried out. The resulting sequence (MKVAV) did not correspond with that of TyrB (MFQKV), AspC (MFENI), or IlvE (MTTKK) of *E. coli*. These three *E. coli* enzymes would be the only extraneous sources of aromatic aminotransferase activity (14), and we have confirmed that a *tyrB aspC ilvE* triple mutant lacks all detectable aromatic aminotransferase activity. These data show that IAP aminotransferase possesses broad substrate specificity which includes aromatic keto and amino acids.

Heat stability. The purified HisH protein (Table 2) showed relatively good tolerance to elevated temperatures. Enzyme preparations were incubated at 5°C temperature increments over the range of 40 to 65°C for 2 min and then kept at 4°C until assay. Both IAP aminotransferase and PAT activities remained unchanged when the enzyme was treated at 50°C but decreased sharply and in parallel when incubated above 55°C for 2 min (data not shown). Under these conditions, 50 and 96% inactivation of both activities was obtained at 57.5 and 65°C, respectively.

**Freeze-thaw stability.** The purified enzyme (Table 2) is unstable to freeze-thaw treatment. After freezing at  $-20^{\circ}$ C overnight and thawing on ice, 98% of both the IAP aminotransferase and PAT activities were lost; after freezing at  $-80^{\circ}$ C overnight and thawing on ice, 58% of both activities were lost. In contrast, 100% of both activities remained after several months at  $-80^{\circ}$ C in 10% (vol/vol) glycerol.

Requirement for PLP. When the purified enzyme preparation (Table 2) was dialyzed against 50 mM potassium phosphate buffer overnight to remove  $\alpha$ -ketoglutarate, PLP, and β-mercaptoethanol, both IAP aminotransferase and PAT activities were unaltered. Enzyme activity was neither dependent on nor stimulated by the presence of PLP in the reaction buffer. However, when the purified enzyme preparation was dialyzed against 20 mM potassium phosphate buffer containing 3 mM hydroxylamine and then dialyzed against the same buffer without hydroxylamine to remove the pyridoxamine generated, no enzyme activity could be detected in this PLP-resolved enzyme preparation. Incubation of the PLP-resolved enzyme preparation with 0.1 mM PLP for 15 min at 37°C restored about 17% of the original activities with either histidinol phosphate or phenylpyruvate. These results indicated that PLP binds tightly to the enzyme. Removal of the cofactor apparently triggers irreversible denaturation. The presence of 10% glycerol, which stabilizes the enzyme during storage, did not stabilize the apoenzyme resolved of cofactor.

TABLE 3. Substrate specificity of purified IAP aminotransferase<sup>a</sup>

Keto acid substrate <sup>b</sup>	Sp act (µmol of product min <sup>-1</sup> mg <sup>-1</sup> )	Amino acid substrate <sup>c</sup>	Sp act ( $\mu$ mol of product min <sup>-1</sup> mg <sup>-1</sup> )
a-Ketoglutarate	7.0	Phenylalanine	$ \begin{array}{c} 11.6\\ 11.6\\ 1.13\\ 0.06\\ 0\end{array} $
p-Hydroxyphenylpyruvate	7.11	Tyrosine	
Phenylpyruvate	0.56	Histidine	
Oxaloacetate	0.10	Leucine	
Pyruvate	0.25	Phosphoserine	

<sup>*a*</sup> See Materials and Methods for conditions used in enzyme assays. All activities were assayed by monitoring the formation of product amino acids, using HPLC.

<sup>b</sup> Used in combination with histidinol phosphate as amino donor substrate.

<sup>c</sup> Used in combination with  $\alpha$ -ketoglutarate as keto acid substrate.

Effect of pH on enzyme activity. The purified enzyme exhibited a broad pH range between pH 6.0 and 8.0 when different substrates were used (data not shown). Over this range, IAP aminotransferase, PAT, and HAT activities did not vary more than 25%, but pH optima varied: pH 6.5 for HAT, pH 7.1 for PAT, and pH 8.0 for IAP aminotransferase.

**Substrate specificity of the purified enzyme.** The specific activities of the purified enzyme when different substrates were used are listed in Table 3. In addition to high activities when histidinol phosphate, phenylalanine, and tyrosine were used as amino donors, histidine was also found to be an effective substrate. The enzyme has only very low activity as an aspartate aminotransferase or alanine aminotransferase, judging from results obtained with oxaloacetate or pyruvate in combination with histidinol phosphate. Phosphoserine, a phosphorylated intermediate of a biosynthetic pathway (like histidinol phosphate), was not utilized at all by IAP aminotransferase.

Kinetics of the purified enzyme. A series of kinetic analyses was carried out. Double-reciprocal plots of initial velocity as a function of varied concentration of one substrate when the other substrate was maintained at various fixed concentrations gave a family of parallel lines (Fig. 4). These experimental results conform with expectations for a ping-pong bi-bi type of reaction mechanism, as has been demonstrated for many other aminotransferases (34). The  $K_m$  and  $V_{max}$  values for different substrates were calculated from the secondary plots of intercept versus reciprocal concentrations of the fixed substrate (data not shown). The  $K_m$  and  $V_{\text{max}}/K_m$  values are listed in Table 4. The markedly higher affinity for histidinol phosphate than for any of the alternative substrates is consistent with its probable major role as an enzyme of histidine biosynthesis. Surprisingly, the  $K_m$  value determined for  $\alpha$ -ketoglutarate varied significantly depending upon the amino donor cosubstrate. Thus, the  $K_m$  value was 50-fold lower when histidinol phosphate was the amino donor than when aromatic amino acids were used. Perhaps imidazole acetol phosphate produced in the first half-step is not released before the binding of  $\alpha$ -ketoglutarate. If so, the bound imidazole acetol phosphate may induce a conformational change leading to increased affinity for  $\alpha$ -ketoglutarate. On the other hand, phenylpyruvate or 4-hydroxyphenylpyruvate may fail to induce an equivalent conformational change. This proposed mechanism is consistent with the ping-pong bi-bi kinetics shown in Fig. 4.

**Possible linkage of** *hisH-tyrC* with genes of tryptophan biosynthesis. Plasmid pF5.8 and cosmids pFBA9, pC2, pC5, and pF20 (Table 1) were introduced into *E. coli* UTH780 and AT2471 by electroporation, but none of the transformants obtained corrected the *hisC* and *tyrA* deficiencies. Thus, linkage of *hisH-tyrC* with *trpFBA* and *trpC* has been eliminated. Possible linkage to *trpD* or *trpE* has not been eliminated.



FIG. 4. Double-reciprocal plots of initial velocity as a function of one substrate concentration (millimolar) at several fixed concentrations (millimolar) of the second substrate. (A and B) Enzyme assayed as IAP aminotransferase. (A)  $\alpha$ -Ketoglutarate concentration ([ $\alpha$ -KG]) is varied, and histidinol phosphate concentration ([HP]) is fixed; (B) [HP] is varied, and [ $\alpha$ -KG] is fixed. (C and D) Enzyme assayed as PAT. (C) [ $\alpha$ -KG] is varied, and phenylalanine concentration ([Phe]) is fixed; (D) [Phe] is varied, and [ $\alpha$ -KG] is fixed. (E and F) Enzyme assayed as HAT. (E) [ $\alpha$ -KG] is varied, and tyrosine concentration [Tyr] is fixed; (F) [Tyr] is varied, and [ $\alpha$ -KG] is fixed.

Varied substrate	Variable	substrate		
varied substrate	$\overline{K_m (\mathrm{mM})}$	$V_{\rm max}/K_m$	Saturating substrate	
Histidinol phosphate	0.17	150.1	α-Ketoglutarate	
Tyrosine	3.39	4.3	α-Ketoglutarate	
Phenylalanine	43.48	1.1	α-Ketoglutarate	
α-Ketoglutarate	0.08	319.0	Histidinol phosphate	
α-Ketoglutarate	4.08	3.6	Tyrosine	
α-Ketoglutarate	3.85	12.1	Phenylalanine	

## DISCUSSION

Class I aminotransferase gene families. Multiple alignment clearly shows that Z. mobilis HisH belongs to the class I aminotransferase superfamily. This superfamily includes rat tyrosine aminotransferase, vertebrate aspartate aminotransferases, and the IAP aminotransferases (Fig. 5). The latter form a distinct cluster which we will refer to as family I $\beta$ . Z. mobilis HisH exhibits the highest amino acid identity (36%)



by the PILEUP multiple alignment program using the Genetics Computer Group software (9). The primary function of each aminotransferase in the biosynthesis of histidine (His), aromatic amino acids (Aro), aspartate (Asp), or branched-chain amino acids (Bch) are indicated by closed symbols. Secondary function in an additional pathway (when known) is indicated with open symbols. Abbreviations with accession numbers given in parentheses are as follows: Rme Tyr, Rhizobium meliloti aromatic aminotransferase (L05065); Zmo Asp, Z. mobilis aspartate aminotransferase (unpublished); Alf Asp1, alfalfa aspartate aminotransferase 1 (P28011); Alf Asp2, alfalfa aspartate aminotransferase 2 (12); Rat Asp(m), rat mitochondrial aspartate aminotransferase (M18467); Rat Asp(c), rat cytosolic aspartate aminotransferase (J04171); Eco AspC, E. coli aspartate aminotransferase (X03629); Pae PhhC, Pseudomonas aeruginosa aromatic aminotransferase (M88627); Eco TyrB, E. coli aromatic aminotransferase (X03628); Zmo HisH, Z. mobilis IAP aminotransferase (L36343); Bsu HisH, B. subtilis IAP aminotransferase (P17731); Hvo HisC, H. volcanii IAP aminotransferase (M33161); Eco HisC, E. coli IAP aminotransferase (P06986); Sty HisC, S. typhimurium HisC (J01804); Sce His5, Saccharomyces cerevisiae IAP aminotransferase (X05650); Sco HisC, Streptomyces coelicolor IAP aminotransferase (M31628); Lla HisC, Lactococcus lactis IAP aminotransferase (M90760); Rme Asp-a, R. meliloti aspartate aminotransferase a (L05064); Rme Asp-b, R. meliloti aspartate aminotransferase b (L12149); Bsp Asp, Bacillus species aspartate aminotransferase (P23034); Sso Asp, Sulfolobus solfataricus aspartate aminotransferase (X16505); Rat Tyr, rat tyrosine aminotransferase (X02741). Eco IlvE (E. coli branched-chain aminotransferase [X02413]) is shown (dashed lines) as an outlying member.

with *B. subtilis* HisH. The most divergent members of family I $\beta$  exhibit amino acid identities of about 25%.

Mehta et al. (24) first recognized rat tyrosine aminotransferase and IAP aminotransferases as homologs of the classical aspartate aminotransferase family. Since at that time only two members of family I $\beta$  and one member of family I $\gamma$  were available, the three family clusters shown in Fig. 5 were not yet apparent. The emergence of additional amino acid sequences representing each family shows that family I $\alpha$  diverged from a common ancestor of families I $\beta$  and I $\gamma$ .

All members of the class I superfamily utilize  $\alpha$ -ketoglutarate and glutamate very well, and each of the three families has at least some members which can utilize aromatic amino acids. Family I $\beta$  is unique for its IAP aminotransferase capability, and other aminotransferases are not known to be able to catalyze this activity. *Z. mobilis* HisH was found to be unable to use aspartate. Both *Z. mobilis* HisH and *B. subtilis* HisH can utilize aromatic amino acids; the breadth of substrate specificity for other IAP aminotransferases has not been studied. Both family I $\alpha$  and family I $\gamma$  possess members whose primary function is for aromatic amino acid biosynthesis or for aspartate biosynthesis.

Although *E. coli* IlvE is not a class I aminotransferase, it is included in Fig. 5 because it has been shown to exhibit functional overlap with *E. coli* TyrB and AspC (14, 18), which belong to family I $\alpha$ . The dendrogram in Fig. 5 suggests that there was a broad-specificity ancestor aminotransferase which transaminated both aspartate and aromatic amino acids. Divergent evolution led to the separation of two lineages, one being the common ancestor of family I $\alpha$  members. The other lineage diverged to yield family I $\gamma$  and a new specialized step of histidine biosynthesis (family I $\beta$ ).

The IAP aminotransferase gene family. A multiple alignment of all reported IAP aminotransferase amino acid sequences is shown in Fig. 6. A total of 28 residues are invariant within family I $\beta$ , and 11 of these are invariant within the entire superfamily. Amino acid sequences corresponding to the highly conserved interdomain interface and active-site lysine regions may be useful as signature motifs characteristic of family IB. These are  $192 \cdot P(x_1) \times P(x_2)$  and  $254 \cdot T(x_1)$ SK( $x_4$ )A( $x_2$ )R( $x_1$ )G-268, respectively. The two family I $\alpha$  arginine residues, R-292 and R-386, well known to interact with the distal and  $\alpha$ -carboxyl moieties of dicarboxylic substrates (10), are not aligned by computer program with arginine residues in family Iβ. Perhaps residues R-381 and R-393 of Z. mobilis HisH (invariant in family IB) perform an equivalent function. Indeed, residues equivalent to Z. mobilis R-393 have been manually aligned with pig cytosolic aspartate aminotransferase R-386 (24).

**Complex-operon function.** *hisH* and *tyrC* appear to be cotranscribed. Northern blot analysis using Z. mobilis RNA has shown a 2-kb transcript. Cotranscription is consistent with the apparent translational coupling of hisH and tyrC. As in other organisms, Z. mobilis HisH is undoubtedly critical for histidine biosynthesis. That the aminotransferase function may extend to the pathway of aromatic amino acid biosynthesis as well is consistent with the coexistence of *hisH* with *tvrC* in an apparent operon and also with the precedent in B. subtilis. The genes of tryptophan biosynthesis are not clustered within a single operon in Z. mobilis as they are in B. subtilis, but we considered the possibility that some tryptophan pathway genes might be located adjacent to the hisH-tyrC operon. When cosmid constructs containing Z. mobilis trpFBA or trpC were introduced by electroporation into E. coli hisC or tyrA mutants, no His<sup>+</sup> or Tyr<sup>+</sup> transformants were obtained. Thus, neither trpFBA nor *trpC* of Z. *mobilis* is closely linked to the *hisH-tyrC* region.

VOL.

Zmo HisH Bsu HisH Hvo HisC Eco HisC Sty HisC Sce His5 Sco HisC	P K S W I D S I A P I K E H L K Q L K P P R D . L S A H A P A R E N V R N L T P A R E N V R N L V P V R P K I Y N L E P V R D E L R G K S P	Y Q P G K P I E A V Y V P G R G T E E V Y Q S A R Y Q S A R Y R C A R	K S E Y G L D K A R E L G M D P E D R L G G N G R L G G N G D . D F T E	A V K L S S N E N X L G T S L K A K E A Y R E A I D V V K L A S N E N Y G C S E A A K E A L H H E I Q L T K L S S N E N Y H G P S P K A V A A I E D A A P D V W L N A N E Y Y T A V E	57 59 57 50 53 49 57
Zmo HisH Bsu HisH Hvo HisC Eco HisC Sty HisC Sce His5 Sco HisC	* Q & A L Y P D S G A Q & A L Y P D G Y S T V S V Y P K T A H T & N R Y P E C Q P T & N R Y P E C Q P . X H R Y P D P H Q D & N R Y P D R D A	T A L R E A I A A L R T R L T D L T E R L L E F K T A M T K Y V E L	KAVIENYAQY KAVIENYAQY KAVIENYAQY RNKTSSYAND	GACYNLDPAR IIHGTGSDBI LHLAAGAYAG SKHLNVSETS LIFGNGSDBI EQIICRAFLN ADKWGLAPEQ VWVSPGAOGS EDVLTRAVLE AGVKPEQ VVVSRGADBG ELLVIRAFCE AGVKPEQ VLVSRGADBG ELLVIRAFCE AGVKPEQ VLVSRGADBG ELVIRAFCE PEVKPLTADN LCLGVGSDBS ELAIIRACCV TSGHPLDVSN VWAANGSNBV EQQLLQTFGG	104 107 104 97 100 109 110
Zmo HisH Bsu HisH Hvo HisC Eco HisC Sty HisC Sce His5 Sco HisC	Q D . D E V L Y P R D K . T N T V T A A P D . D R I L E P A P G K D A I L Y C P P G K D A I L Y C P R E . E K I L V L P P G R T A I G F . E	#         T         F         P         Q         Y         K         H         N         A           #         G         F         S         Y         S         M         S         A           #         T         Y         G         M         Y         S         V         S         A           #         T         Y         G         M         Y         S         V         S         A           #         T         Y         G         M         Y         S         V         S         A           #         T         Y         G         M         Y         S         V         S         A	V I E G A E V . R Y H H G D A V Q . E T I G V E S E T I G V E R N I N D I E V V Q C	VE.APDDDYR       CSVDALLKAV       TPRTRVVF         REIALRPDGS       HDLDAMLEAI       DEQTQVVV         YEVSKDDDFE       QTADLVLDAY       DGEYVV         RTVPTPDNWQ       LDLQGISDKL       DGVKAVY         RTVPALENWQ       LDLQGISDKL       DGTKVVF         PLIVSQC       MDTEAVLTIL       KNDSLIKLMF        PRHEDFT       IDVPAATRAI       DEHR.PDVVF	157 160 159 151 154 168 165
Zmo HisH Bsu HisH Hvo HisC Eco HisC Sty HisC Sce His5 Sco HisC	IAN         PNNPTGT           ICSPNNPTGT           LCSPNNPTGT           LTTPHNNPTGZ           VCSPNNPTGQ           VCSPNNPTGQ           VCSPNNPTGQ           VTSPGNPTGQ           ITTPNNPTGQ           ITTPNNPTGQ	W I T R A E V E K L Y T S E G E L L A F V L P R E E L V E L L I N P Q D L R T L L I N P Q D L R T L K I K F Q D L R T L A V P A E T V L A L	AESVEEHT LELTRGKA LELTRGKA LQ.NWDNG	* * * * * * * * * * * * * * * * * * *	215 218 214 206 209 221 222
Zmo HisH Bsu HisH Hvo HisC Eco HisC Sty HisC Sce His5 Sco HisC	UVTRTFSKAY AILRTFSKAY AILRTFSKAY VILRTLSKAF VILSKAF VILSKSF VVLOTLSKSF VVSRTMSKAF	* * * W A G & A A E R V G Y G G & A A L R V G Y G G & A A L R V G Y G G & A G L R I G Y A A & A G L R I G F T G & A G L R L G M T G & A G L R L G M T G & A G L R L G Y L	I A D E N L I R Q I C V P E A W A D A Y L A N E E V I N L L L A N E E V I N L L Y A T A E L A R I L	NRIRAPFNVT IAGQQKAAVAA LEDQAFIQ BPAREPFNTE RIGQAAAIAA LYDQAFIA ARVNTPFAAS EVACRAALAA LYDDQAFVB MKVIAPYPLS TPVADIAAQA LSPQGIVAMR LKVIAPYYPLS TPVADIAAQA SCPQESNAMR LKVIAPYYPLS STPVADIAAQA SCPQESNAMR QLVRLPYHLS AVTQATALAA LEHTDTL	266 269 265 259 262 274 273
Zmo HisH Bsu HisH Hvo HisC Eco HisC Sty HisC Sce His5 Sco HisC	N S F K H S K K V A S C V E Q N N A G . K S V E S A R W S R E R V A Q I I A E R D R V A Q T V Q E A T S K I I N E E K K Y V E Q L K T E R	E Q Q Y Y D F A D Y E R E H L D E Y E I A A L K E I Q Y E V N A L Q Q T M L E L K E L T A L	K T H G L K C Y P S A P T W E S P C V E . Q V F D S A C V E . H V F D S D Y V D . D Q Y V G	SANFTLLLFE       GSLTAKTAYK       ALM       D         QTNFVLIDFK       RPADELFQ       ALL       R         GNFVLVEVG       DADELFQ       ALL       R         EGNFVLVEVG       DADELFQ       ALL       R         GNFVLVEVG       DADELFQ       ALL       R        ETNYI.LA       RFFASSAVF.       KSLWDQ         GLDANFL.LI       RINGGDNVLA       KKLYYQLATQ         ESDANFVQFG       RFADSHATWR       KIL	324 322 314 315 318 337 326
Zmo HisH Bsu HisH Hvo HisC Eco HisC Sty HisC Sce His5 Sco HisC	HGYTTR. WLP KGYIVR. SGN EGVIVR. SGN .GIILRDQNK .GIILRDQNK SGVVVRFRGN RGVVVRDN	G Q R L P H A L R R I         A L G F P T S C I R R I         S F G L S G C L R R I         Q P S L S G C L R R I         E L G C S G C L R R I         E L G V P G W E	T I G T K E Q N E E S C G T E T Q T K R T V G T R E E S Q R T V G T R Q E N Q P T V G T H E E N T H	V A G I L T S L V R Q A L	369 363 361 356 359 385 370

FIG. 6. Multiple alignment of IAP aminotransferase (family I $\beta$ ). Invariant residues are boxed, and residues conserved in six of the seven amino acid sequences are shaded. Residues invariant throughout the entire class I superfamily are labeled with asterisks, except for the invariant active-site lysine, which is indicated with a solid circle. These invariant residues correspond to the following residues of porcine cytosolic aspartate aminotransferase: 70 (Y), 110 (G), 194 (N), 195 (P), 197 (G), 222 (D), 225 (Y), 258 (K), 266 (R), 268 (G), and 386 (R). Individual amino acid residue numbers are given at the far right. See the legend to Fig. 5 for abbreviations used at the left.

It is unknown whether HisH might be a major or sole enzyme of aromatic biosynthesis in Z. mobilis or whether it has a backup role, as appears to be the case in B. subtilis. An obvious rationale may not be readily apparent to explain the coexistence of genes from different pathways in operons. However, this is likely to reflect the coordination of interpathway relationships. Thus, in E. coli and S. typhimurium, serC and aroA exist as components of a complex operon (8, 17), and this has been explained as a mechanism for coordination of a response to iron limitation which requires output from both the serine and aromatic amino acid biosynthetic pathways to produce enterochelin. The serC-aroA operon is also involved in pyridoxine biosynthesis; serine, the aromatics, and pyridoxine converge at a number of places in metabolism (20, 31). The modern era of cloning in which sequenced genes of interest may be found to coexist with unexpected genes in complex operons has the potential to reveal novel interpathway relationships of regulation. For example, in E. coli (35) and S. typhimurium (41), a gene encoding an enzyme of heme biosynthesis coexists with a gene encoding a probable N-acetylmuramyl-L-alanine amidase, an enzyme of cell wall metabolism.

Other evolutionary and physiological considerations. E. coli

*hisC* (IAP aminotransferase) and *hisD* (histidinol dehydrogenase) are adjacent cistrons within the histidine operon (40). These genes have been fused in the laboratory to yield a remarkably functional bicatalytic protein (30). We considered the possibility that, like Z. mobilis hisH and E. coli hisC, Z. mobilis tyrC and E. coli hisD are homologs. However, this seems unlikely since (i) contemporary Z. mobilis TyrC has no histidinol dehydrogenase activity, (ii) a pairwise alignment of Z. mobilis TyrC and E. coli HisD did not exhibit obvious homology, and (iii) the mechanisms of the two dehydrogenase reactions are different.

Since purified Z. mobilis IAP aminotransferase can utilize histidine, it is possible that this activity is physiologically significant. Organisms capable of growth on imidazole lactate as a sole carbon source utilize a dehydrogenase to convert it to imidazolepyruvate, which is then transaminated to histidine. Histidine is then catabolized to yield glutamate (15). The possibility that HisH and TyrC function together as an aminotransferase-dehydrogenase unit able to convert imidazolelactate to histidine was considered since cyclohexadienyl dehydrogenase (TyrC) is already known to exhibit broad substrate specificity. However, Z. mobilis TyrC was found to be incapable of utilizing imidazolelactate as a substrate.

## ACKNOWLEDGMENTS

We acknowledge the most helpful suggestions and advice of John Gander of this department. We thank P. Gomez and L. O. Ingram for access to their unpublished sequence of Zmo Asp.

#### REFERENCES

- 1. Applied Biosystems, Inc. User bulletin 18. Applied Biosystems, Inc., Foster City, Calif.
- 1a.Araki, K., and K. Nakayama. 1974. A biochemical characterization of histidine auxotrophs of *Corynebacterium glutamicum*. Agric. Biol. Chem. 38: 2219–2225.
- Babitzke, P., P. Gollnick, and C. Yanofsky. 1992. The mtrAB operon of Bacillus sublifis encodes GTP cyclohydrolase I (MtrA), an enzyme involved in folic acid biosynthesis, and MtrB, a regulator of tryptophan biosynthesis. J. Bacteriol. 174:2059–2064.
- Bialkowska-Hobrzanska, H., C. A. Gilchrist, and D. T. Denhardt. 1985. Escherichia coli rep gene: identification of the promoter and N terminus of the Rep protein. J. Bacteriol. 164:1004–1010.
- Bonner, C., and R. A. Jensen. 1987. Prephenate aminotransferase. Methods Enzymol. 142:479–487.
- Bradford, M. M. 1976. A rapid and sensitive method for the quantitation of microgram quantities of protein utilizing the principle of protein-dye binding. Anal. Biochem. 72:248–254.
- Conover, R. K., and W. F. Doolittle. 1990. Characterization of a gene involved in histidine biosynthesis in *Halobacterium (Haloferax) volcanii*: isolation and rapid mapping by transformation of an auxotroph with cosmid DNA. J. Bacteriol. 172:3244–3249.
- Devereux, J., P. Haeberli, and O. Smithies. 1984. A comprehensive set of sequence analysis programs for the VAX. Nucleic Acids Res. 12:387–395.
- Duncan, K., and J. R. Coggins. 1986. The serC-aroA operon of Escherichia coli. Biochem. J. 234:49–57.
- Eddy, C. K., O. H. Smith, and K. D. Noel. 1988. Cosmid cloning of five Zymomonas mobilis trp genes by complementation of Escherichia coli and Pseudomonas putida trp mutants. J. Bacteriol. 170:3158–3163.
- Ford, G. C., G. Eichele, and J. N. Jansonius. 1980. Three-dimensional structure of a pyridoxal-phosphate-dependent enzyme, mitochondrial aspartate aminotransferase. Proc. Natl. Acad. Sci. USA 77:2559–2563.
- 11. Fox, G. E., and R. A. Jensen. Unpublished data.
- Gantt, J. S., R. J. Larson, M. W. Farnham, S. M. Pathirana, S. S. Miller, and C. P. Vance. 1992. Aspartate aminotransferase in effective and ineffective alfalfa nodules. Plant Physiol. 98:868–878.
- Garrick-Silersmith, L., and P. E. Hartman. 1970. Histidine-requiring mutants of *Escherichia coli* K-12. Genetics 66:231–244.
- Gelfand, D. H., and R. A. Steinberg. 1977. Escherichia coli mutants deficient in the aspartate and aromatic amino acid aminotransferases. J. Bacteriol. 130:429–440.
- Hacking, A. J., and H. Hassall. 1975. The purification and properties of L-histidine-2-oxoglutarate aminotransferase from *Pseudomonas testosteroni*. Biochem. J. 147:327–334.
- Henner, D. J., L. Band, G. Flaggs, and E. Chen. 1986. The organization and nucleotide sequence of the *Bacillus subtilis hisH*, *tyrA* and *aroE* genes. Gene 49:147–152.
- Hoiseth, S. K., and B. A. D. Stocker. 1985. Genes aroA and serC of Salmonella typhimurium constitute an operon. J. Bacteriol. 163:355–361.
- Jensen, R. A., and D. H. Calhoun. 1981. Intracellular roles of microbial aminotransferases: overlap enzymes across different biochemical pathways. Crit. Rev. Microbiol. 8:229–266.
- Laemmli, U. K. 1970. Cleavage of structural proteins during the assembly of the head of bacteriophage T4. Nature (London) 227:680–685.
- Lam, H. M., and M. E. Winkler. 1992. Characterization of the complex pdxH-tyrS operon of Escherichia coli K-12 and pleiotropic phenotypes caused by pdxH insertion mutations. J. Bacteriol. 174:6033–6045.
- 21. Limauro, D., A. Avitabile, M. Cappellano, A. M. Puglia, and C. B. Bruni.

1990. Cloning and characterization of the histidine biosynthetic gene cluster of *Streptomyces coelicolor* A3(2). Gene **90**:31–41.

- Maniatis, T., E. F. Fritsch, and J. Sambrook. 1989. Molecular cloning: a laboratory manual, 2nd ed. Cold Spring Harbor Laboratory, Cold Spring Harbor, N.Y.
- Martin, R. G., M. A. Berberich, B. N. Ames, W. M. Davis, R. F. Goldberger, and J. D. Yourno. 1971. Enzymes and intermediates of histidine biosynthesis in *Salmonella typhimurium*. Methods Enzymol. 17B:3–51.
- Mehta, P. K., T. I. Hale, and P. Christen. 1989. Evolutionary relationships among aminotransferases. Tyrosine aminotransferase, histidinol-phosphate aminotransferase, and aspartate aminotransferase are homologous proteins. Eur. J. Biochem. 186:249–253.
- Montenecourt, B. S. 1985. *Zymomonas*, a unique genus of bacteria, p. 261–289. *In* A. L. Demain and N. A. Soloman (ed.), Biology of industrial microorganisms. Benjamin-Cummings Publishing Co., Menlo Park, Calif.
- Nester, E. W., and A. L. Montoya. 1976. An enzyme common to histidine and aromatic amino acid biosynthesis in *Bacillus subtilis*. J. Bacteriol. 126:699– 705.
- Orr, M. D., R. L. Blakley, and D. Panagou. 1972. Discontinuous buffer systems for analytical and preparative electrophoresis of enzymes on polyacrylamide gel. Anal. Biochem. 45:68–85.
- Piggot, P. J., and J. A. Hoch. 1985. Revised genetic linkage map of *Bacillus subtilis*. Microbiol. Rev. 49:158–179.
- Pond, J. L., C. K. Eddy, K. F. Mackenzie, T. Conway, D. L. Borecky, and L. O. Ingram. 1989. Cloning, sequencing, and characterization of the principal acid phosphatase, the *phoC* product, from *Zymomonas mobilis*. J. Bacteriol. 171:767–774.
- Rechler, M. M., and C. M. Bruni. 1971. Properties of a fused protein formed by genetic manipulation. Histidinol dehydrogenase-imidazolylacetol phosphate:L-glutamate aminotransferase. J. Biol. Chem. 246:1806–1813.
- Roa, B. B., D. M. Conolly, and M. E. Winkler. 1989. Overlap between pdxA and ksgA in the complex pdxA-ksgA-apaG-apaH operon of Escherichia coli K-12. J. Bacteriol. 171:4767–4777.
- Roth, C. W., and E. W. Nester. 1971. Co-ordinate control of tryptophan, histidine and tyrosine enzyme synthesis in *Bacillus subtilis*. J. Mol. Biol. 62:577–589.
- Silhavy, T. J., M. L. Berman, and L. W. Enquist. 1984. Experiments with gene fusions. Cold Spring Harbor Laboratory, Cold Spring Harbor, N.Y.
- 34. Sung, M-H., K. Tanizawa, H. Tanaka, S. Kuramitsu, H. Kagamiyama, and K. Soda. 1990. Purification and characterization of thermostable aspartate aminotransferase from a thermophilic *Bacillus* species. J. Bacteriol. 172: 1345–1351.
- Troup, B., M. Jahn, C. Hungerer, and D. Jahn. 1994. Isolation of the *hemF* operon containing the gene for the *Escherichia coli* aerobic coproporphyrinogen III oxidase by in vivo complementation of a yeast HEM13 mutant. J. Bacteriol. 176:673–680.
- Tsui, H.-C. T., A. J. Pease, T. M. Koehler, and M. E. Winkler. 1994. Detection and quantitation of RNA transcribed from bacterial chromosomes and plasmids. Methods Mol. Genet. 3:179–204.
- Weigent, D. A., and E. W. Nester. 1976. Regulation of histidinol phosphate aminotransferase synthesis by tryptophan in *Bacillus subtilis*. J. Bacteriol. 128:202–211.
- Weigent, D. A., and E. W. Nester. 1976. Purification and properties of two aromatic aminotransferases in *Bacillus subtilis*. J. Biol. Chem. 251:6974– 6980.
- Whitaker, R. J., C. G. Gaines, and R. A. Jensen. 1982. A multispecific quintet of aromatic aminotransferases that overlap different biochemical pathways in *Pseudomonas aeruginosa*. J. Biol. Chem. 257:13550–13556.
- 40. Winkler, M. E. 1987. Biosynthesis of histidine, p. 395–411. In F. C. Neidhardt, J. L. Ingraham, K. B. Low, B. Magasanik, M. Schaechter, and H. E. Umbarger (ed.), *Escherichia coli* and *Salmonella typhimurium*: cellular and molecular biology. American Society for Microbiology, Washington, D.C.
- Xu, K., and T. Elliott. 1993. An oxygen-dependent coproporphyrinogen oxidase encoded by the *hemF* gene of *Salmonella typhimurium*. J. Bacteriol. 175:4990–4999.
- Zhao, G. S., T. H. Xia, L. O. Ingram, and R. A. Jensen. 1993. An allosterically insensitive type of cyclohexadienyl dehydrogenase from *Zymomonas mobilis*. Eur. J. Biochem. 212:157–165.