## Pseudomonas aeruginosa possesses homologues of mammalian phenylalanine hydroxylase and $4\alpha$ -carbinolamine dehydratase/ DCoH as part of a three-component gene cluster

Genshi Zhao\*, Tianhui Xia, Jian Song, and Roy A. Jensen<sup>†</sup>

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

Communicated by David B. Sprinson, October 22, 1993 (received for review July 12, 1993)

ABSTRACT Pseudomonas aeruginosa possesses a multigene operon that includes phenylalanine hydroxylase (PhhA; phenylalanine 4-monooxygenase, EC 1.14.16.1). phhA encodes PhhA ( $M_r = 30,288$ ), phhB ( $M_r = 13,333$ ) encodes a homologue of mammalian  $4\alpha$ -carbinolamine dehydratase/homeodomain protein transregulator, and phhC encodes an aromatic aminotransferase ( $M_r = 43,237$ ). The reading frames specifying phhB and phhC overlap by 2 bases. The P. aeruginosa PhhA appears to contain iron and is pterin dependent. Unlike the multimeric mammalian hydroxylase, the native P. aeruginosa enzyme is a monomer. The P. aeruginosa PhhA is homologous with mammalian PhhA, tryptophan hydroxylase, and tyrosine hydroxylase. Expression of PhhA from its native promoter required phhB. This may suggest a positive regulatory role for phhB, consistent with the dual catalytic and regulatory roles of the corresponding mammalian homologue.

Phenylalanine hydroxylase (PhhA; phenylalanine 4-monooxygenase, EC 1.14.16.1) catalyzes the irreversible conversion of phenylalanine to tyrosine, using tetrahydrobiopterin (BH<sub>4</sub>) as a reducing agent. In humans and other animals the enzyme catalyzes the rate-limiting step in the catabolism of dietary phenylalanine (1). BH<sub>4</sub> is oxidized to  $4\alpha$ -hydroxytetrahydrobiopterin, which is then dehydrated via  $4\alpha$ -carbinolamine dehydratase to form dihydrobiopterin (2). BH<sub>4</sub> is regenerated from dihydrobiopterin by dihydropteridine reductase.

PhhA has generally been considered to be rare in prokaryotes, where scattered reports of its existence have been limited to a single phylogenetic division of Gram-negative bacteria (3-7). Of these microbial pterin- and metal-dependent monooxygenases, only the PhhA proteins from *Pseudomonas acidovorans* (3, 7) and *Chromobacterium violaceum* (8, 9) have been purified and characterized. Recently the *C. violaceum* gene encoding PhhA has been sequenced (10), and PhhA was found to be a member of the mammalian hydroxylase gene family. *C. violaceum* PhhA differs from the mammalian enzymes in its smaller subunit size, its existence as a monomer (rather than as a homotetramer), and binding of copper (instead of iron) at the active site.

In this paper, we report the striking observation that PhhA from *Pseudomonas aeruginosa* is encoded by a gene organized within an apparent operon including genes producing homologues of mammalian  $4\alpha$ -carbinolamine dehydratase/DCoH and aromatic aminotransferase.<sup>‡</sup>

## **MATERIALS AND METHODS**

Strains, Plasmids, and Media. P. aeruginosa PA01 (11) and Pseudomonas stutzeri JM300 (12) are prototrophic strains. Escherichia coli strains used were JM83 [ara  $\Delta$ (proAB-lac) rpsL thi  $\phi$ 80 lacZ $\Delta$ M15] obtained from GIBCO/BRL, AT2471 (thi tyrA) obtained from Barbara Bachmann, JP2255 (aroF363 pheA361 pheO352 tyrA382 thi strR712 lacY1 xy15) (13), AB3257 (aroF, aroG, aroH) (14), and DG44 (thr1 leu6 thil proA2 argE3 hisG4 lacY1 galK2 ara14 xy15 mtll rpsL31 tsx33 supE44) (15). Plasmid pUC18 (lacZ Ap<sup>r</sup>) was obtained from GIBCO/BRL. LB and M9 (16) were used as growth media for E. coli and P. aeruginosa. Additions of ampicillin (50  $\mu$ g/ml), phenylalanine (50  $\mu$ g/ml), and thiamine (17  $\mu$ g/ml) were made when appropriate. Agar was added at 20 g/liter to solidify media.

**Molecular Cloning.** Construction of a *P. aeruginosa* gene library was performed as described (17). Recombinant plasmids were purified from the library (18) and used to transform *E. coli* AT2471 (*tyrA*), auxotrophic for tyrosine. Transformants able to grow on M9 minimal plates lacking tyrosine were selected. Plasmids were isolated from the transformants and used to retransform *E. coli* AT2471.

**DNA Manipulations.** Restriction-site mapping and subcloning were done by standard methods (16). Southern blot hybridization was carried out following the instructions of Promega with biotinylated probes.

DNA Sequencing and Data Analysis. Subcloned fragments, obtained after purification of plasmids by CsCl gradient centrifugation (18), were sequenced in both directions at the DNA Core Facility of the University of Florida. Data analysis was carried out with the Genetics Computer Group software package (19).

**Crude Extract Preparation and Enzyme Assay.** *P. aeruginosa* PA01 and *E. coli* JP2255 carrying various plasmids were grown in M9 and LB media (450 ml), respectively. The cells harvested at the late exponential phase of growth were disrupted by sonication, and the resulting extract was centrifuged at  $150,000 \times g$  for 1 hr at 4°C. The supernatant (crude extract) was collected and used for enzyme assay. PhhA was assayed by following tyrosine formation (8). Cyclohexadie-nyl dehydrogenase was assayed by a continuous spectrophotometric method (20).

Anion-Exchange Chromatography. E. coli JP2255 (pJZ9-3a) was grown in 2 liters of LB broth supplemented with ampicillin (50  $\mu g/ml$ ) at 37°C and harvested by centrifugation during midexponential growth. The cell pellets were resuspended in 20 ml of 10 mM potassium phosphate buffer (pH 7.4) containing 50 mM KCl and 1 mM dithiothreitol (DTT) (buffer A) and were disrupted by sonication. The resulting suspension was centrifuged at 150,000  $\times g$  for 60 min. The supernatant was applied to a DEAE-cellulose column (2.5  $\times$  19 cm) that had been equilibrated with buffer A. The column was washed with 200 ml of buffer A and then eluted with a 500-ml linear salt gradient (50-300 mM KCl) in buffer A.

The publication costs of this article were defrayed in part by page charge payment. This article must therefore be hereby marked "*advertisement*" in accordance with 18 U.S.C. §1734 solely to indicate this fact.

Abbreviation: BH<sub>4</sub>, tetrahydrobiopterin.

<sup>\*</sup>Present address: Department of Microbiology, University of Texas Medical School at Houston, Houston, TX 77054.

<sup>&</sup>lt;sup>†</sup>To whom reprint requests should be addressed.

<sup>&</sup>lt;sup>‡</sup>The sequence reported in this paper has been deposited in the GenBank data base (accession no. M88627).

Proc. Natl. Acad. Sci. USA 91 (1994) 1367

Fractions of 5 ml were collected, and those exhibiting phhA activity were pooled and concentrated by use of an Amicon PM-10 membrane. This preparation had a specific activity of 153 nmol·min<sup>-1</sup>·mg<sup>-1</sup> and represented a purification of 4.3-fold with respect to crude extracts.

Gel-Filtration Chromatography. The  $M_r$  of the native PhhA was estimated by gel filtration. The concentrated preparation obtained after DEAE-cellulose fractionation was washed with buffer B (10 mM potassium phosphate, pH 7.4/0.1 M KCl/1 mM DTT) and was applied to a Sephadex G-100 column (2.8 × 56 cm) that had been equilibrated with buffer B and run at 4°C. Column fractions (3 ml) were collected. Molecular weight standards were bovine serum albumin ( $M_r$ , 66,000), ovalbumin ( $M_r$ , 45,000), and cytochrome c ( $M_r$ , 12,300) (Sigma). The fractions were monitored for protein at  $A_{280}$  and for PhhA activity. The peak fraction exhibited a specific activity of 1.32 µmol·min<sup>-1</sup>·mg<sup>-1</sup>. This purification procedure yielded a 37.2-fold purification with a 45% yield. Relative to the enzyme level present in wild-type *P. aeruginosa*, a purification exceeding 600-fold was obtained.

## RESULTS

Molecular Cloning and Identification of the P. aeruginosa phhA Gene. Since PhhA produces tyrosine, a reasonable strategy for cloning the phhA gene would utilize functional complementation of E. coli tyrosine auxotrophs. Two E. coli auxotrophs, AT2471 (tyrA) and JP2255 (tyrA pheA), were used. E. coli AT2471, used in the primary screening studies, would be expected to grow in the presence of plasmid-borne phhA or tyrC, the latter gene encoding cyclohexadienyl dehydrogenase in P. aeruginosa (21). We then used E. coli JP2255 to verify the identity of those plasmids carrying phhA. Plasmids carrying the phhA gene would be able to complement E. coli JP2255 in the presence of phenylalanine, whereas those carrying the tyrC gene would be unable to do so because of the absence of chorismate mutase activity in this strain (13). E. coli AB3257, lacking all three isoenzymes that catalyze the first step of aromatic biosynthesis and therefore unable to generate prephenate, was also used to confirm the presence of phhA and the absence of tyrC.

**Origin of the Cloned DNA Insert.** Restriction analysis showed that five randomly selected plasmids all shared a 2.5-kb *HincII* fragment, later demonstrated to contain *phhA*, *phhB*, and most of *phhC*. To confirm that the 5.7-kb DNA insert of plasmid pJ29

was derived from *P. aeruginosa*, the 2.5-kb *Hin*cII fragment was used as a probe in Southern hybridization with *P. aeruginosa*, *P. stutzeri*, and *E. coli* chromosomal DNAs after *Hin*cII digestion. The biotinylated probe hybridized with a 2.5-kb DNA fragment of *P. aeruginosa*, but not with the chromosomal DNAs of *E. coli* and *P. stutzeri*.

Subcloning Analysis and Identification of Genes Required for Hydroxylation in Vivo. Fig. 1 shows that the HincII fragment was the smallest insert that was capable of complementing the E. coli tyrA defect. This fragment carries phhA and phhB but not all of phhC. When the EcoRI/HincII fragment upstream of phhA is present (as in pJZ9 and pJZ9-1), the activities measured for PhhA were markedly depressed. The growth rates of the latter strains were also slow, apparently limited by the rate of tyrosine biosynthesis. In contrast, the strains carrying the pJZ9-3a and pJZ9-3b constructs exhibited high PhhA activity, and these strains grew at normal wild-type growth rates.

The 2.5-kb *HincII* fragment was cloned into plasmids with opposite polylinker orientations, with pJZ9-3a having the same orientation as the original insert. Since the orientation of pJZ9-3b precludes use of the *lacZ* promoter, the native promoter of the *phhA* gene must be recognized by *E. coli*.

Successful functional complementation corresponded to the presence of PhhA activity *in vitro* as shown in Fig. 1. SDS/PAGE profiles shown on the far right demonstrate the presence of prominent bands corresponding to PhhA and PhhB in some of the constructs. The synonymy of these bands with PhhA and PhhB were confirmed by N-terminal amino acid sequencing.  $4\alpha$ -Carbinolamine dehydratase activity was verified by direct enzyme assay in subclone pJZ9-3a (see *Discussion*).

Nucleotide Sequence of phhA, phhB, and phhC. The 2.5-kb HincII fragment required for functional complementation of *E. coli* JP2255 was sequenced. Most of this sequence, as well as some of the downstream flanking region, is presented in Fig. 2. The three open reading frames identified were designated phhA, phhB, and phhC.

*phhA* (789 bp) encodes a protein with 262 residues having a deduced  $M_r$  of 30,288 and an isoelectric point of 5.63. The G+C content of *phhA* was 64%, well within the range for the *P. aeruginosa* genome, and codon usage of *phhA* was also typical (22), showing a preference for G or C at the third base in 89.7% of the codons.



FIG. 1. Localization of *phhA*, *phhB*, and *phhC*. Linear maps of the original clone and its subclones are shown. Ability to complement the *tyrA* deficiency of *E. coli* AT2471 (or JP2255), together with specific activity values determined for PhhA, is also indicated on the right. Positions of restriction sites in the cloned *P. aeruginosa* DNA and in part of the pUC18 polylinker site (stippled bars) are shown at the top. Protein bands visualized by Coomassie blue staining after SDS/PAGE are also shown on the far right, and positions of bands corresponding to PhhA and PhhB are indicated. *Plac, E. coli lacZ* promoter.

← bhhA	
GGCAACGACAACAACAAAGACCCCGCGGCCACAAGCCGGTCATGGAGTCCGTATGAAAAGGACGACGACGACGACGACGACGACGACGACGACG	130
TGGAATACCCTGATCACCCGGCAACTGAAGGTGATCGAAGGCCGCGCCTGTCAGGAATACCTCGACGGCATCGACGGCCTGGCCCGCCGGGGGGGTCCCCCAGGTGGAGGATCAACAGGGTTC W N T L I T R Q L K V I E G R A C Q E Y L D G I E Q L G L P H E R I P Q L D E I N R V L	260
TCCAGGCCACCACCGGCTGGGCAGGGGTGGCAGGGGGGGG	390
CCTGCAGGAGCCGGACATCTTCCACGAGATCTTCCGCCACTGCCCACGCCGACCCAGCCCGGCCCGGCCTCGCCCGAGCCAAGGCGAGCAAGGAGGAACGCGTGTTC L Q B P D I F H B I F G H C P L L T N P W L A B F T H T Y G K L G L K A S K B B R V F	520
CTOSCCCCCCTTACTOSCATGACCATOSACTTOSCCCTOGTOGAGACOGACGAGACOGACAGCGCATCTACOGCOGCACCACTCTCCCGACGACGCCTCTACAGCCTCTCCGACGACGCCGCCG L A R L Y W M T I E P G L V E T D Q G K R I Y G G G I L S S P K B T V Y S L S D E P L H	650
ACCAGGCCTTCAATCOGCTGGAGGGGATGGGCACGCCCTACCGCATCGACATCGACGCCTCTATTTOGTCCTGCCCGGCCTGATCCGACTGGCCCAGGGAGGACATCATCGGCGCTGGT Q A F N P L E A M R T P Y R I D I L Q P L Y F V L P D L K R L F Q L A Q E D I M A L V	780
CCACGAGGCCATGGGCCTGGCGCCGGCGGCGGCGGCGGCGGCGGGGGGGG	910
GCCGCTCCCACATGCGACAACCAACGACGCACCACCACGCCCACGCCCAGGCCCACGCCGC	1040
GGATTGGAACATCGAAGGCGACGGCATCATGCAGCTAGAGAAGGTCTACCTGTTCAAGAACTTCAAGCATGCCCTGGCCTTCACCAATGCCGTCGGCGAGGGAGATATCCGAGGCCGAAGGCCGACGGTCG D W N I E V R D G I M Q L B K V Y L F K N F K H A L A F T N A V G E I S E A E G H H P	1170
GGCCTGCTGACCGAGTGGGGCAAAGTCACCGTGACCTGGTGGAGCCACTGATCAAGGGCCTGCACCGCAACGATTTCATCATGGCGGCGCGCACCGATGAGGTAGCGAAAACCGCCCAAGGGGCGCAAAT G L L T B N G K V T V T N N S H S I K G L H R N D F I M A A R T D E V A K T A B G R KM* -> DADC	1300
GASTCATTICGCCAAGGTCGCCGCGTACCGGGCCGACCCGGACCGCTGCTGCGGCGCGACGACGACGGCGGGGGGGG	1430
ACCCCGATCCTGGGTCGGTGAAACTCGCCGAGCAGGCAGG	1560
CCTOGCOGCTGTTGCTGGAGGAACGCCGACGGCAACCCAGACGCCCGGCGGGCACCGGCGG	1690
CTGGCCGATCCACGAGACCCTGTTCGCCGCCGCGCGGAGAGGTTTCCCACTACCCCTACGTCAGGCGACAACCGCCTGGATGTCGAGGCGATGCTTGGCGCTGGAGGCGATCCCCAGGGAGAC W P I H E T L F A A A G L K V S H Y P Y V S A D N R L D V E A M L A G L E R I P Q G D	1820
GTGGTGCTGCTGCTGCTGCCGCCACAACCCGACCGGTTTCGACCTGACCGACGACGGCGCGGGGGCGCGGGGGGGG	1950
TOGGCGACGGTCTOGAGGAAGACGCCTGGGCGGTACGCCTGTTCGCCCGGAGACTGCCGGAGGTGCTGGTCACCAGTTCCTGCTCGAGGAAGTTCGGCCGCGCGGCGGTCGGT	2080
CTGCGCGCAGAACGCCGAGAAGCTCACCGACCTGCGTAGCCAACTGGCCTACCCCGCCAACCTCTGGTCGACCCCGCGCGCAGGTGGTCGCGGCGACGCGCGACGCGAGTGG C A Q N A E K L T D L R S Q L A F L A R N L W S T P P A H G A E V V A A I L G D S E L	2210
AAGGACTCTGGCAGGAAGAGTCGAAGGCATGGCCTGGCGCATCGCCAGCCTGGCCTGGCCTGGCCGGCGCGGCGCGGCGCGCGC	2340
TGTTTTCCTATACGGACTGAGCCCGCAGCAGGTGCGCGGCGGCGACGGAGCAGCGGTTTACCTGGTCTCCAGGGGCCGAGCCAACGTCGCCGGCGCGGCGGCGGCGGCGGCGGCGGCGGCGGCG	2470
AGCCATCGCCCAGGTCTGCGCGGACTGAACCCTCACGGTTACCCTGAGCCGGCTGTGATGGCCGCCTTTTTTTCGCCCCAGCGAATGCCGCGCACTCCCGCCGAGGCGGAGGATAACCGCGTTGTGGTTA A I A Q V C A D * >>>>>> <<<<<<	2600

FIG. 2. Complete nucleotide sequence of region spanning *phh* structural genes. Deduced amino acid sequences encoded by *phhA*, *phhB*, and *phhC* are shown in single-letter code. Putative ribosome-binding sites are underlined and labeled RBS. Base residues forming the complementary stems of the putative terminator structure are marked with tandem arrowheads.

phhB (357 bp) encodes a  $M_r$  13,333 protein with a calculated isoelectric point of 6.39. The G+C content of phhB was 61.5%. The protein is unusual in having only one tyrosine residue (0.85 mol %), but four tryptophan residues (3.39 mol %). phhB was required for hydroxylation since the joint presence of phhA and phhB was required for functional complementation. Consistent with this, E. coli JP2255 harboring a plasmid construct with the intact phhA but truncated phhB (pJZ9-5) exhibited only marginal PhhA activity (Fig. 1).

phhC (1200 bp) encodes a protein having a deduced  $M_r$  of 43,237 and an isoelectric point of 6.57. The G+C content of phhC was 69.6%, and the codon usage of phhC was very similar to that of phhA. That phhC specifies an aminotransferase was apparent from its homology with E. coli aspC (47% identity) and with E. coli tyrB (45% identity). In confirmation of this, plasmid pJZ9 was able to transform E. coli DG44, an aspC tyrB mutant, to nutritional independence of aspartate and tyrosine.

Gene Organization. The three open reading frames identified are closely linked with no obvious terminator sequences between them. In fact, *phhB* and *phhC* exhibit a 2-bp overlap between the stop and start codons. A possible terminator sequence with a free energy of  $-13.4 \text{ kcal} \cdot \text{mol}^{-1}$  (1 cal = 4.184 J) was found to exist 17 bp beyond the stop codon of *phhC* (Fig. 2). No obvious promoter structure was observed in the 354-bp region upstream of *phhA*. Nevertheless, a promoter recognized by *E. coli* is implicated within this region since complementation was achieved by the *P. aeruginosa* DNA insert present in pJZ9-3a and pJZ9-3b (Fig. 1), these having opposite polylinker orientation with respect to the *lac* promoter of pUC18. Thus, it is highly probable that the three genes identified are organized as an operon.

The *EcoRI/HincII* region present in pJZ9 and pJZ9-1 upstream of *phhA* (see Fig. 1) may encode a trans-acting

repressor protein since excision of this region in pJZ9-3a and pJZ9-3b results in a dramatic increase of PhhA activity.

Initial Characterization of PhhA. Partially purified PhhA (see Materials and Methods) was used to characterize some of the properties of the enzyme. The enzyme is stable at  $-80^{\circ}$ C for at least 1 month. The requirement for pterin cofactor is absolute. Unlike the mammalian enzyme (23), precatalysis incubation with phenylalanine did not activate the P. aeruginosa enzyme when the natural pterin cofactor was used. Comparison of the calculated  $M_r$  of 27,000 after gel filtration with the deduced  $M_r$  indicates that PhhA is active as a monomer. The enzyme preparation used contained both PhhA and PhhB, as judged by SDS/PAGE (data not shown). Further purification steps (such as hydroxyapatite chromatography) separated PhhA and PhhB, thus demonstrating retention of catalytic competence by PhhA in the absence of PhhB. Since PhhB cofractionated with PhhA after gel filtration, the native form of PhhB is probably dimeric.

The PhhA metalloprotein probably contains iron. Treatment with EDTA abolished >99% of the activity, and this was completely restored by addition of  $Fe^{2+}$  after removal of the EDTA. In contrast, addition of copper failed to restore activity. These results were consistent with the abolition of all activity by the iron-selective chelator ethylenediamine di(*o*-hydroxyphenylacetic acid), but not by the copperselective chelator bathocuproine disulfonate.

## DISCUSSION

**Comparison of the Microbial PhhA Systems.** The *P. aeruginosa* PhhA amino acid sequence exhibited 35% overall identity with *C. violaceum* PhhA. However, homology was apparent only over the first 150 residues. Beyond this point the published *C. violaceum* sequence (10) appears to be incorrect due to a frameshift error. Correction by deletion of nucleotide 702 (G) in the published sequence increased the

Rno-Tyh Hsa-Tyh Cc0-Tyh Mmu-Trh Rno-Trh Rno-Trh Rno-Phh Hsa-Phh Dme-Phh Pae-Phh Cvi-Phh Rno-Tyh Hsa-Tyh Bta-Tyh Bta-Tyh Hsa-Ty	V P W V P W V L W V P W V V V V	P P R K V S E L D K F P R K V S E L D K F P R K V S E L D K F P R K V S E L D K F P R K I S D L D F F P K K I S D L D F F P K K I S D L D H F P R K I S D L D H F P R K I S D L D H F P R R I R D L D R G E P I P H V E Y T G D P I P R V E Y T C T C T P R V E Y T C T P	C H H L V T K F D P C H H L V T K F D P C H H L V T K F D P C H H L V T K F D P C H H L V T K F D P C A N R V L L Y G S C A N R V L L Y G S C A N R V L L Y G S C A N R V L M Y G S F A N Q I L S Y G A F A N Q I L S Y G A F A N Q I L S Y G S M N D •••••••••••••••••••••••••••••••••••	D L D L D H P G F S D L D L D H P G F S D L D L D H P G F S D L D L D H P G F S D L D L D H P G F S E L D A D H P G F K E L D A D H P G F K E L D A D H P G F K E L D A D H P G F K E L D A D H P G F K E L D A D H P G F K C L D A D H P G F K E L D A D H P G F K E L D A D H P G F K E L D A D H P G F K E L D A D H P G F K E L D A D H P G F K E L D A D H P G F K E L D A D H P G F K E L D A D H P G F K F C L C C C C C C C C C C C C C C C C C	O V Y R O R K L D Q V Y R Q R R K L D Q Q Y R Q R R K L D Q V Y R Q R R K L D Q V Y R R R K L D Q V Y R R R K L D Q V Y R R R K L D Q V Y R R R K Y D N V Y R R R K Y D N V Y R R R K Y D N V Y R R R K Y D P V Y R R R R Y D P V Y R R R R Y D D V Y R R R R Y D P V Y R R R R Y D P V Y R R R R Y D T T C X R C R E H L E G F H A C R E Y L R L R L H A C R E Y L N H H A C R E Y L N H H A C R E Y L N H H A C R E Y L N H H A C R Y L H H A C R Y H H H A C R Y H H H A C R Y H
Rno-Tyh Hsa-Tyh Bta-Tyh Cco-Tyh Mmu-Trh Rno-Trh Ocu-Trh Hsa-Trh Hsa-Phh Dme-Phh Pae-Phh Cvi-Phh	• 4 Q L L E R Y C G Y R B L L E R F C G Y R E L L B R F C G Y R P L L S K Y C G Y R P L L S K Y C G Y R P L L S K Y C G Y R P L L S K Y C G F R P L L E K Y C G F R P L L V D N C G F R P L L V D N C G L P E R L E V D	E     D     S     I     P     Q     L     E     V       E     D     S     I     P     Q     L     E     D     V       E     D     N     I     P     Q     L     E     D     V     E     D     N     I     P     Q     L     E     D     V     E     D     N     I     P     Q     L     E     D     V     E     D     N     P     Q     L     E     D     V     E     D     V     P     Q     L     E     D     V     E     D     V     P     Q     L     E     D     V     P     Q     L     E     D     V     D     P     Q     L     E     D     V     E     D     V     P     P     Q     L     E     D     V     D     P     D     N     P     P	• • • • • • • • • • • • • • • • • • •	Q L R P V A G L L S Q L R P V A G L L S Q L R P V A G L L S Q L R P V A G L L S S I R P V A G Y L S S I R P V A G Y L S S I R P V A G Y L S S I R P V A G Y L S R L R P V A G L L S T L R P V A G L L S T L R P V A G L L S T L R P V A G L L S T V A R V P A L I P K I V A V P G L I P	A R D F L A S L A F A R D F L A S L A F A R D F L A S L A F A R D F L A S L A F P R D F L A S L A F P R D F L S G L A F P R D F L S G L A F P R D F L S G L A F S R D F L G G L A F S R D F L G G L A F S R D F L G G L A F S R D F L G G L A F S R D F L A G L A F S R D F L A G L A S D V F F E H L A N
Rno-Tyh Hsa-Tyh Bta-Tyh Cco-Tyh Mmu-Trh Rno-Trh Hsa-Trh Hsa-Trh Hsa-Phh Hsa-Phh Dme-Phh Pae-Phh Cvi-Phh	• • • • • • • • • • • • • • • • • • •	pterin-binding     domain       + A     A     S     P     M     S     P       + A     S     P     M     S     P     E       + A     S     P     M     S     P     E     E       + A     S     P     M     S     P     E     E       + A     S     P     M     S     P     E     E     E     E     E     H     S     D     P     Y     T     E     E     H     S     S     D     P     Y     T     P     E     H     S     S     D     P     Y     T     P     E     H     S     K     D     P     Y     T     P     E     H     S     K     P     N     T     P     E     H     S     K     P     N     T     P     H     S     K     P     N     N     T	2     4     4     6     4     6     6       P     D     C     C     H     E     L     G     H       P     D     C     C     H     E     L     G     H       P     D     C     C     H     E     L     G     H       P     D     C     C     H     E     L     G     H       P     D     C     C     H     E     L     G     H       P     D     T     C     H     E     L     G     H       P     D     T     C     H     E     L     G     H       P     D     T     C     H     E     L     G     H       P     D     T     C     H     L     G     H     H     H     H     H     H     H     H     H     H     H     H	• • • • • • • • • • • • • • • • • • •	• • • • • • • • • • • • • • • • • • •
Rno-Tyh Hsa-Tyh Bta-Tyh Cco-Tyh Mmu-Trh Rno-Trh Ocu-Trh Hsa-Trh Hsa-Phh Dme-Phh Pae-Phh Cvi-Phh	• • • • • • • • • • • • • • • • • • •	L S T V Y W F T V E L S T L S W F T V E L S T L S W F T V E L S T L Y W F T V E L A T L Y W F T V E L A T C Y F F T V E L A T C Y F F T V E L A T C Y F F T V E L A T C Y F F T V E L A T I Y W F T V E L A T I Y W F T V E L S T I F W F T V E L A R L Y W Y T V E	F G L C K Q N G E L F G L C K Q N G E V F G L C K Q N G E V F G L C K Q N G E V F G L C K Q N G I V F G L C K Q D G Q L F G L C K Q D G Q L F G L C K Q D G Q L F G L C K K Q G G S I F G L C K E G G D S I Y G L V K T P Q G K F G L U N T T P A G M	A     A     A     A       K     A     Y     G     A     L     L     S       K     A     Y     G     A     G     L     L     S       N     A     Y     G     A     G     L     L     S       N     A     Y     G     A     G     L     L     S       R     V     F     G     A     G     L     L     S       R     V     F     G     A     G     L     L     S       R     V     F     G     G     L     L     S       R     V     F     G     G     G     L     S       K     A     Y     G     A     G     L     S       K     A     G     G     G     L     S     S       R     I     G     G     G     I     L	Y G E L L H S L . S Y G E L L H S L . S Y G E L L H S L . S Y G E L L H S L . S I S E L K H A L . S I S E L K H A L . S I S E L K H A L . S I S E L K H A L . S F G E L Q Y C L . S Y G E L E Y C L . S Y G E L E Y C L . S Y G E L E Y C L . S
Rno-Tyh Hsa-Tyh Bta-Tyh Cco-Tyh Muu-Tch Rno-Trh Ocu-Trh Hsa-Trh Hsa-Phh Dme-Phh Pae-Phh Cvi-Phh	E E P E V R A F D P E E P E I R A F D P E E P E I R A F D P G H A K V K P F D P G H A K V K P F D P G H A K V K P F D P G H A K V K P F D P D K A K V K P F D P D K P K L L P L E L D K P K L L P L E L D K P Q L K D F E P A S P N R V G F D L	D T A A V Q P Y Q D E A A A V Q P Y Q D D A A A V Q P Y Q D D A A A V Q P Y Q D K I A C K Q E C L I K I T Y K Q E C L I K I T Y K Q E C L I E K T A C Q E Y S V E K T A C Q E Y S V E K T A C Q E Y S V E K T A C Q F Y S V E K T A C Q F Y S V E K T A C Q F Y S V E K T A C Q F Y S V E K T A C Q F Y S V E K T A C Q F Y S V E K T A C Q F Y S V E K T A C Q F Y S V E K T A C Q F Y S V E K T A C Q F Y S V E K T A C Y S V S V S V S T S V S V S V S V S V S T S V S V S V S V S V S V S V S V S V	Q T Y Q P V Y F V S Q T Y Q P V Y F V S Q T Y Q P V Y F V S Q T Y Q P V Y F V S T S F Q D V Y F V S T S F Q D V Y F V S T S F Q D V Y F V S T T F Q D V Y F V S T E F Q P L Y Y V A T E F Q P L Y Y V A T E F Q P L Y Y V A T E F Q P L Y Y V A	E S F N D A K D K L E S F S D A K D K L E S F S D A K D K L E S F S D A K N K L E S F E D A K E K K E S F E D A K E K K E S F E D A K E K K E S F E D A K E K K E S F E D A K E K K E S F E D A K E K K D S F E T A K E K V D S F E T A K E K K D S F K Q L	R N Y A S R I Q R P R S Y A S R I Q R P R S Y A S R I Q R P R S Y A S A I I K R P R E F A K T I K R P R E F A K T I K R P R E F K K T I I K R P R F A A T I P R P R T F A A T I P R P I K F A N S I P R P I K C A A P D F A F D A T A P D F A
Rno-Tyh Hsa-Tyh Bta-Tyh Cco-Tyh Mmu-Trh Rno-Trh Ocu-Trh Hsa-Trh Rno-Phh Hsa-Phh Dme-Phh Pae-Phh Cvi-Phh	•     •	• A I D V L D S P H T A I D V L D S P Q A A I D V L D S P Q T S V Q V L R D S K S S I Q I L K D T K S S I Q I L K D T K S S I Q I L K D T Q Q R I E V L D N T Q Q S V E V L D N K P Q Q A Q L A D A Q P W G A	I Q R S L E G V Q D V R R S L E G V Q D I R R A L D S V Q D I C H S L D S V R D I T S A M N B L R H I T S A M N B L R H I T S A M N B L R H I T S A M N B L R H L K I L A D S I N S I S N L M D N I N S A * A T S R R T T W C *	E L H T L A H A L S E L D T L A H A L S E M Q A L A H A L N E L H T L I N A L N D L D V I S D A L A D L D V V N D A L A D L D V V S D A L A D L D V V S D A L A E V G I L C N A L Q E F Q I L Q N A V A	A I S * A I S * A I S * V I S * R V T R W P S V * R V S R W P S V * K V S R Q L S V * K V S R K P S I * K I K S * K I K * K L R V *

FIG. 3. Multiple alignment of aromatic amino acid hydroxylase amino acid sequences using the PILEUP program. Tyh enzymes were from rat (Rno), human (Hsa), cow (Bta), and quail (Cco) tissues. Trh enzymes were from mouse (Mmu), rat, rabbit (Ocu), and human tissues. PhhA enzymes were from rat, human, and *Drosophila melanogaster* (Dme) tissues. Microbial hydroxylases were from *P. aeruginosa* (Pae-Phh) and *C. violaceum* (Cvi-Phh). The latter sequence was corrected for a frameshift error as described in the text. Sequences are compared starting at a position that corresponds to an intron-exon boundary in Hsa-Phh. Residues 170-357, which correspond to exons VI-X in Hsa-Phh and have been suggested to comprise the essential hydroxylation domain (25), are marked with bent arrows. A potential pterin-binding region (26) is indicated. Residues conserved in all 13 sequences ( $\triangle$ ) or conserved in all 11 animal sequences ( $\bigcirc$ ) are shown. In addition, aligned residues corresponding to Hsa-Phh amino acids 280, 311, and 408, and whose alteration is associated with human phenylketonuria, are boxed.

overall identity to 46% and eliminated three gaps in pairwise alignment. The correction also extended the span of C. *violaceum* PhhA homology with mammalian aromatic amino acid hydroxylases (see below). Further evidence supporting a frameshift error in sequencing was the observation that a randomized codon usage in the distal portion of the original sequence was shifted following correction to the same bias of codon usage found in the N-terminal portion of C. *violaceum* 

*phhA*. The correction results in a stop codon at an earlier position to yield a *C. violaceum* peptide with 281 instead of 296 residues. The decreased  $M_r$  calculated from the corrected sequence corresponds more favorably with the  $M_r$  estimated directly with purified enzyme (8).

The two microbial enzymes thus far characterized at the molecular-genetic level are similar in size, reliance on pterin cofactor, and catalytic function in the monomeric state. Both Pae-PhhB MTALTQAHCEACRADAPHVSDEELPVLLRQIP..DWN.IEVRDGIMQLEKVYLFKNFKHALAPTNAVGEISEAEGHHPGLLTEWGKVTVTWWSHSIKGLHRNDFIMAARTDEVAKTAEGRK 118 :.l.::|.| . ||...:.:|| : |...::| : ||:|.:::: . || :|.::::.... Rno-DCOH .......MAGKAHRLSAEERDQLLPNLRAVGWNELEGRDAIF...KQFHFKDPNRAPGFWTRVALQAEKLDHHPEWFNVYNKVHITLSTHECAGLSERDINLASFIEQVAVSMT... 104

FIG. 4. Pairwise alignment (GAP program) of amino acid sequence corresponding to *P. aeruginosa phhB* (Pae-PhhB) and that of rat liver  $4\alpha$ -carbinolamine dehydratase/DCoH (Rno-DCoH). Identity is 32.7% and similarity is 57.4%, with three gaps.

hydroxylases are metalloproteins, but C. violaceum PhhA is unique in its use of copper while P. aeruginosa PhhA appears to use iron like all other eukaryotic and prokaryotic PhhA enzymes. Anti-PhhA antibody from C. violaceum did not cross-react with PhhA from P. aeruginosa.

Comparison of Microbial and Mammalian Pterin-Dependent Hydroxylases. PhhA, Tyh, and Trh from animals all possess an N-terminal extension of  $\approx 200$  amino acids, which is absent in the microbial hydroxylases. This region is known to be dispensable for catalytic function but possesses determinants for phosphorylation and phenylalanine activation (24), characteristics not known to exist for any of the bacterial phenylalanine (or tryptophan) hydroxylases.

Fig. 3 shows an alignment of *P. aeruginosa* PhhA (and *C. violaceum* PhhA) with 11 animal hydroxylases. The corresponding dendrogram (data not shown) indicates that substrate specificity evolved independently in the prokaryote and eukaryote lineages. The independent evolution of hydroxylase specificity for phenylalanine is consistent with the observation that of the 20 residues conserved in the three animal Phh enzymes but divergent in the animal Tyh and Trh enzymes none was conserved in the microbial Phh enzymes.

The P. aeruginosa and C. violaceum PhhA sequences aligned well throughout their entire sequences with the C-terminal portions of the animal hydroxylases. Each pairwise alignment exceeded 31% identity, sufficient to establish homology. Of 79 residues conserved in the 11 animal hydroxylases, 39 were also conserved in the 2 microbial hydroxylases. This implies that the latter residues are ancestral amino acids whose conservation has been most critical for retention of catalytic function. Two of the three residues known to be associated with phenylketonuria are not absolutely conserved in both Pae-Phh and Cvi-Phh. The entire length of the two microbial hydroxylase amino acid sequences align almost perfectly with few gaps with the highly conserved C-terminal portions of the three subfamilies of animal aromatic amino acid hydroxylases. This is consistent with the substantial evidence identifying the carboxyl region as the catalytic core region.

**Function of PhhB.** Pairwise alignment (Fig. 4) shows that *P. aeruginosia* PhhB is a homologue of mammalian  $4\alpha$ carbinolamine dehydratase. Given the homology and the genetic organization, it seemed likely that PhhB would also prove to function as  $4\alpha$ -carbinolamine dehydratase. This has now been demonstrated in a direct assay for dehydratase using chemically synthesized C4 $\alpha$ -hydroxy-6-methyltetrahydropterin as substrate (27). The turnover number and Michaelis constant are similar to those of rat liver enzyme (J. E. Ayling and R.A.J., unpublished data).

It is surprising that constructs possessing *phhA* but not *phhB* failed to exhibit PhhA activity since the dehydratase is not essential *in vitro* when BH<sub>4</sub> is supplied in excess. In this context, it is most intriguing that  $4\alpha$ -carbinolamine dehydratase has recently been shown (28, 29) to be synonymous with DCoH (a trans-acting regulator of a homeodomain transcription factor). Thus, it is possible that in addition to its catalytic activity PhhB may exert a positive regulatory effect, either directly or by interaction with another protein, on the expression of PhhA.

The Aminotransferase Component. phhC encodes a homologue of a well-known family of aminotransferases in which the animal aspartate aminotransferase member has been fully characterized by x-ray crystallography. The highest pairwise identities were with E. coli AspC (46.6%) and E. coli TyrB (45.3%). High levels of phenylalanine/ $\alpha$ -ketoglutarate activity have been found (data not shown) in a strain carrying plasmid construct pJZ9-2 (Fig. 1).

**Perspective.** The striking similarity of *P. aeruginosa* PhhA to human pterin-dependent iron-containing hydroxylases, as well as of PhhB to human  $4\alpha$ -carbinolamine dehydratase (DCoH), may be significant because of prospects for molecular-genetic characterization of catalytic and regulatory features that might contribute insight into the nature of the technically less assailable human system.

We thank S. J. Benkovic for antibody raised against C. violaceum PhhA. We are indebted to J. E. Ayling for carrying out the assays of carbinolamine dehydratase. This is Florida Agricultural Station Journal Series No. R-03595.

- 1. Kaufman, S. (1987) Methods Enzymol. 142, 3-27.
- Lazarus, R. A., Benkovic, S. J. & Kaufman, S. (1983) J. Biol. Chem. 258, 10960-10962.
- 3. Guroff, G. & Ito, T. (1963) Biochim. Biophys. Acta 77, 159-161.
- 4. Decicco, B. T. & Umbreit, W. W. (1964) J. Bacteriol. 88, 1590-
- 1594.
- 5. Friedrich, B. & Schlegel, H. G. (1972) Arch. Mikrobiol. 83, 17-31.
- Letendre, C. H., Dickens, G. & Guroff, G. (1974) J. Biol. Chem. 249, 7186-7191.
- Berry, A., Johnson, J. L. & Jensen, R. A. (1985) Arch. Microbiol. 141, 32–39.
- Nakata, H., Yamauchi, T. & Fujisawa, H. (1979) J. Biol. Chem. 254, 1829–1833.
- Pember, S. O., Villafranca, J. & Benkovic, S. J. (1986) Biochemistry 25, 6611-6619.
- Onishi, A., Liotta, L. J. & Benkovic, S. J. (1991) J. Biol. Chem. 266, 18454-18459.
- 11. Holloway, B. W. (1955) J. Gen. Microbiol. 13, 572-581.
- Carlson, C. A., Pierson, L. S., Rosen, J. J. & Ingraham, J. L. (1983) J. Bacteriol. 153, 93-99.
- 13. Baldwin, G. S. & Davidson, B. E. (1981) Arch. Biochem. Biophys. 211, 66-75.
- 14. Wallace, B. J. & Pittard, J. (1969) J. Bacteriol. 99, 707-712.
- 15. Gelfand, D. H. & Steinberg, R. A. (1977) J. Bacteriol. 130, 429-440.
- Sambrook, J., Fritsch, E. F. & Maniatis, T. (1989) Molecular Cloning: A Laboratory Manual (Cold Spring Harbor Lab. Press, Plainview, NY), 2nd Ed.
- Zhao, G. S., Xia, T. H., Fischer, R. S. & Jensen, R. A. (1992) J. Biol. Chem. 267, 2487-2493.
- Humphreys, G. O., Willshaw, G. A. & Anderson, E. S. (1975) Biochim. Biophys. Acta 383, 457-463.
- Devereux, J., Haeberli, P. & Marquess, P. (1987) The Program Manual for the Sequence Analysis Software Package (Genetics Computer Group, Madison, WI).
- 20. Bonner, C. A. & Jensen, R. A. (1987) Methods Enzymol. 142, 488-494.
- 21. Xia, T. H. & Jensen, R. A. (1990) J. Biol. Chem. 265, 20033-20036.
- 22. West, S. E. H. & Iglewski, B. H. (1988) Nucleic Acids Res. 16, 9323-9335.
- Davis, M. D., Ribeiro, P., Tipper, J. & Kaufman, S. (1992) Proc. Natl. Acad. Sci. USA 89, 10109-10113.
- 24. Dahl, H. H.-M. & Mercer, J. F. (1986) J. Biol. Chem. 261, 4148-4153.
- Morales, G., Requena, J. M., Jimenez-Ruiz, A., Lopez, M. C., Ugarte, M. & Alonzo, C. (1990) Gene 93, 213–219.
- Jennings, I. G., Kemp, B. E. & Cotton, R. G. H. (1991) Proc. Natl. Acad. Sci. USA 88, 5734–5738.
- Bailey, S. W., Boerth, S. R., Dillard, S. B. & Ayling, J. E. (1993) in *Chemistry and Biology of Pteridines and Folates*, eds. Ayling, J. E., Nair, M. G. & Bough, C. M. (Plenum, New York), pp. 47–54.
- Hauer, C. R., Rebrin, I., Thony, B., Neuheiser, F., Curtius, H-Ch., Hunzikar, P., Blau, N., Ghisla, S. & Heizmann, C. W. (1993) J. Biol. Chem. 268, 4828-4831.
- Citron, B. A., Davis, M. D., Milstien, S., Gutierrez, J., Mendel, D. B., Crabtree, G. R. & Kaufman, S. (1992) Proc. Natl. Acad. Sci. USA 89, 11891-11894.