Biosynthesis of Riboflavin: Cloning, Sequencing, and Expression of the Gene Coding for 3,4-Dihydroxy-2-Butanone 4-Phosphate Synthase of *Escherichia coli*

GERALD RICHTER,[1] RAINER VOLK,[1]† CORNELIA KRIEGER,[1] HANS-WERNER LAHM,[2]
URS RÖTHLISBERGER,[2] AND ADELBERT BACHER[1]*

*Lehrstuhl für Organische Chemie und Biochemie, Technische Universität München, Lichtenbergstrasse 4,
D-8046 Garching, Germany,[1] and Pharmaceutical Research New Technologies, F. Hoffmann-La Roche AG,
CH-4002 Basel, Switzerland[2]*

3,4-Dihydroxy-2-butanone 4-phosphate is biosynthesized from ribulose 5-phosphate and serves as the biosynthetic precursor for the xylene ring of riboflavin. The gene coding for 3,4-dihydroxy-2-butanone 4-phosphate synthase of *Escherichia coli* has been cloned and sequenced. The gene codes for a protein of 217 amino acid residues with a calculated molecular mass of 23,349.6 Da. The enzyme was purified to near homogeneity from a recombinant *E. coli* strain and had a specific activity of 1,700 nmol mg^{-1} h^{-1}. The N-terminal amino acid sequence and the amino acid composition of the protein were in agreement with the deduced sequence. The molecular mass as determined by ion spray mass spectrometry was 23,351 ± 2 Da, which is in agreement with the predicted mass. The previously reported loci *htrP*, "*luxH*-like," and *ribB* at 66 min of the *E. coli* chromosome are all identical to the gene coding for 3,4-dihydroxy-2-butanone 4-phosphate synthase, but their role had not been hitherto determined. Sequence homology indicates that gene *luxH* of *Vibrio harveyi* and the central open reading frame of the *Bacillus subtilis* riboflavin operon code for 3,4-dihydroxy-2-butanone 4-phosphate synthase.

6,7-Dimethyl-8-ribityllumazine (Fig. 1, compound 6), the direct biosynthetic precursor of riboflavin (compound 5), is formed by condensation of 5-amino-6-ribitylamino-2,4 (1*H*,3*H*)-pyrimidinedione (compound 2) with 3,4-dihydroxy-2-butanone 4-phosphate (compound 4). The enzyme catalyzing the formation of the novel carbohydrate (compound 4) from ribulose 5-phosphate (compound 3) has been purified to homogeneity from the flavinogenic yeast *Candida guilliermondii*, and its reaction mechanism has been studied in some detail (28–30). The yeast enzyme is a monomer of 24 kDa. The complex enzyme reaction involves the elimination of C-4 from ribulose 5-phosphate as formate via an intramolecular rearrangement as well as the conversion of the position 1 hydroxymethyl group to a methyl group. The catalytic process probably involves a sequence of tautomerization reactions (30). It is surprising that such a complex reaction can be performed by a single and relatively small protein.

A more detailed understanding of this intriguing reaction mechanism requires information on the structure of the enzyme. We have cloned and sequenced the gene for 3,4-dihydroxy-2-butanone 4-phosphate synthase from *Escherichia coli* by using a marker rescue strategy. The enzyme was purified to near homogeneity from a recombinant *E. coli* strain and characterized. Apparently, the gene for 3,4-dihydroxy-2-butanone 4-phosphate synthase is identical to the previously reported loci *htrP* (21), "*luxH*-like" (31), and *ribB* (3) located at 66 min on the *E. coli* chromosome. However, the function of the gene locus had not been established previously.

MATERIALS AND METHODS

Materials. Restriction enzymes and T4 DNA ligase were from Pharmacia and Bethesda Research Laboratories. The lumazine synthase-riboflavin synthase complex (heavy riboflavin synthase) of *Bacillus subtilis* was purified to a specific activity of about 1,000 nmol mg^{-1} h^{-1} (2, 24). Other enzymes were obtained from Boehringer Mannheim and Sigma. The T7 sequencing kit was from Pharmacia, and ^{35}S-Sequetide was from DuPont NEN. M13/pUC sequencing and reverse sequencing primers were purchased from Boehringer Mannheim. KS and SK primers were obtained from Stratagene. Other synthetic oligonucleotides were custom synthesized by Pharmacia.

Bacterial strains and plasmids. Strains of *E. coli* and plasmids used in this study are summarized in Table 1.

Culture media. Bacteria were grown on Luria-Bertani (LB) medium. Ampicillin (150 μg/ml) and riboflavin (400 μg/ml) were added as required.

Library construction. Chromosomal DNA of *E. coli* RR28 was isolated by the method of Godson and Vapnek (8) and partially digested with *Sau*3AI restriction endonuclease. A fraction encompassing 6- to 15-kb restriction fragments was isolated by agarose gel electrophoresis and ligated into the *Bam*HI site of the plasmid pBluescript II SK−. For amplification, the recombinant DNA was transformed into *E. coli* DH1 cells by the high-efficiency procedure described by Hanahan (10). The cells were plated on agar containing ampicillin (150 μg/ml), and 50,000 colonies were harvested. Plasmid DNA was isolated by the Triton-lysozyme method of Davis et al. (7).

Transformation and screening procedures. Mutant cells were made competent and transformed as described by Hanahan (10). Cells were allowed to express the phenotype in LB medium containing 10 mM Mg^{2+} and riboflavin (400

---

* Corresponding author.
† Present address: Institut für Prophylaxe und Epidemiologie der Kreislaufkrankheiten, D-8000 München 2, Germany.
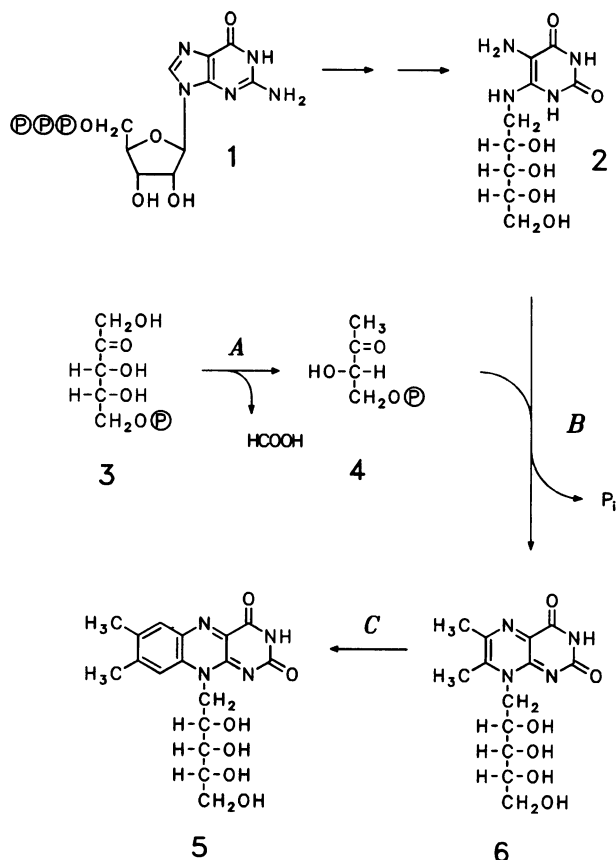
FIG. 1. Biosynthesis of riboflavin. (A) 3,4-Dihydroxy-2-butanone 4-phosphate synthase; (B) 6,7-dimethyl-8-ribityllumazine synthase; (C) riboflavin synthase.

μg/ml). Cells were washed twice with 0.9% NaCl and plated on LB agar plates containing 150 μg of ampicillin per ml.

Construction of subclones. Plasmid p10-16 was partially digested with endonuclease *Sau*3AI, and fragments were ligated into the *Bam*HI site of plasmid pBluescript II SK−. Ligation mixtures were transformed into the Rib10 mutant and screened for riboflavin prototrophy as described above. Colonies were isolated and purified. Plasmids from three clones (p10-1621, p10-1633, and p10-1664) were isolated and used for sequencing.

Plasmid p10-1664 was digested to completion with endonuclease *Sau*3AI. A band corresponding to a 464-bp fragment was excised from an agarose gel, purified with Geneclean (Bio 101) according to the manufacturer's instructions, and ligated into the *Bam*HI site of plasmid pBluescript II SK−. The ligation mixture was transformed into XL1-Blue cells. Several colonies were picked, and a plasmid (p10-16642) containing the 464-bp *Sau*3AI segment was used in subsequent studies.

DNA sequence analysis. Plasmid DNA from riboflavin-independent clones was isolated by the method of Holmes and Quigley (13) or by using Nucleobond AX columns from Macherey & Nagel (Düren, Germany) according to the manufacturer's instructions. DNA sequencing was performed with the T7 sequencing kit from Pharmacia and the ^{35}S-Sequetide labeling mix from DuPont NEN. After dideoxy sequencing, the products were separated on urea-acrylamide gels. The gels were subsequently dried and autoradiographed.

Assay of 3,4-dihydroxy-2-butanone 4-phosphate synthase activity. Assay mixtures contained 200 mM phosphate (pH 7.5), 10 mM ribose 5-phosphate, 20 mM $MgCl_2$, 0.1 U of pentose phosphate isomerase, and protein solution in a total volume of 50 μl. Samples were incubated at 37°C for 1 h. A solution (50 μl) containing 4 mM freshly prepared 5-amino-6-ribitylamino-2,4(1*H*,3*H*)-pyrimidinedione, 40 mM EDTA, 40 mM dithiothreitol, and 10 U of heavy riboflavin synthase from *B. subtilis* was added. The mixture was incubated for 2 h at 37°C. Riboflavin was determined by reversed-phase high-performance liquid chromatography (HPLC) (Nucleosil RP18 column, 4 by 250 mm; eluent, 40% methanol containing 100 mM ammonium formate). The effluent was monitored fluorometrically (excitation, 445 nm; emission, 516 nm). One unit of enzyme activity catalyzes the formation of 1 nmol of 3,4-dihydroxy-2-butanone 4-phosphate per h.

Enzyme purification. Frozen cells (40 g) of *E. coli* Rib10 carrying the plasmid p10-16 were thawed in 460 ml of 50 mM potassium phosphate (pH 7) containing 0.5 mM phenylmethylsulfonyl fluoride, 45 mg of lysozyme, and 4.5 mg of DNase I. The mixture was stirred at 37°C for 5 h. The suspension was centrifuged at 4,000 × g for 30 min.

The crude cell extract was applied to a column of DEAE-cellulose DE 52 (Whatman; 4.5 by 35 cm) which had been equilibrated with 50 mM potassium phosphate (pH 7). The flow rate was 100 ml/h. The column was developed with 1.5 liters of 80 mM potassium phosphate (pH 7) followed by 1.5 liters of 120 mM phosphate (pH 7). Fractions containing

TABLE 1. Bacterial strains and plasmids

Strain or plasmid	Relevant characteristics	Source and reference
E. coli strains		
RR28	*thi leu pro lac ara xyl endA recA hsd* r⁻ m⁻ *pheS supE44*	Hennecke (11)
DH1	*recA1 endA1 gyrA96 thi-1 hsdR17* (r_K^- m_K^+) *supE44 relA1*	Hanahan (9)
XL1-Blue	*recA1 endA1 gyrA96 thi-1 hsdR17 supE44 relA1 lac* [F′ *proAB lacI*�q*ZΔM15* Tn*10(tet*ʳ)]	Stratagene
Rib5	*thi leu pro lac ara xyl endA recA hsd* r⁻ m⁻ *pheS supE44 rib*	Katzenmeier (14)
Rib10	*thi leu pro lac ara xyl endA recA hsd* r⁻ m⁻ *pheS supE44 rib*	Katzenmeier (14)
Plasmids		
pBluescript II SK−	High-copy-number phagemid vector	Stratagene
p10-16	pBluescript II SK− with 6-kb *Sau*3AI fragment	This study
p10-1621	pBluescript II SK− with 1.7-kb *Sau*3AI fragment	This study
p10-1633	pBluescript II SK− with 1.0-kb *Sau*3AI fragment	This study
p10-1664	pBluescript II SK− with 1.5-kb *Sau*3AI fragment	This study
p10-16642	pBluescript II SK− with 464-bp *Sau*3AI fragment	This study

TABLE 2. Specific activities of 3,4-dihydroxy-2-butanone
4-phosphate synthase in cell extracts of *E. coli* strains

Strain	Plasmid	Sp act (U/mg)
RR28		1.3
Rib10	p10-16	110.5
Rib10	p10-1621	3.1
Rib10	p10-1633	3.7
Rib10	p10-1664	124.9

3,4-dihydroxy-2-butanone 4-phosphate synthase activity were collected.

The protein solution was dialyzed three times against a buffer containing 8.3 mM citric acid and 16.7 mM KH_2PO_4 (pH 4.8). The precipitate which formed during dialysis was removed by centrifugation. The solution was applied to a column of SP-Sephadex C-50 (Pharmacia; 4 by 32 cm) which had been equilibrated with the dialysis buffer. The flow rate was 80 ml/h. The column was developed with the equilibration buffer. Fractions containing 3,4-dihydroxy-2-butanone 4-phosphate synthase activity were collected, dialyzed against 50 mM potassium phosphate (pH 7), and concentrated by ultrafiltration with an Amicon PM10 membrane.

Aliquots containing 200 μg of protein were applied to an HPLC column (Superdex 75; 1 by 30 cm). The column was developed with 50 mM Tris hydrochloride (pH 8) containing 150 mM NaCl. The flow rate was 0.3 ml/min. The effluent was monitored photometrically (280 nm).

Sodium dodecyl sulfate-polyacrylamide gel electrophoresis was performed as described by Laemmli (15). Molecular weight standards were supplied by Sigma. Protein concentration was determined by the modified Bradford method (22). Bovine serum albumin was used as a standard.

Amino acid analysis. An aliquot of protein solution was concentrated and desalted by reversed-phase HPLC on Velosep RP-8 (Brownlee; 40 by 3.2 mm). The protein was hydrolyzed in 6 M hydrochloric acid at 110°C for 24 h essentially as described by Spackman et al. (25). The hydrolysate was analyzed with a Liquimat III amino acid analyzer (Kontron Instruments), employing ion exchange separation and postcolumn detection with ninhydrin.

Protein sequencing. Amino acid sequence analysis of the N terminus was performed by automated Edman degradation (12), using an ABI 475A protein sequencer equipped with an on-line phenyl thiohydantoin amino acid analyzer (model 120; Applied Biosystems).

Mass spectrometry. A sample aliquot was separated by reversed-phase HPLC (model 1090; Hewlett Packard) on an Aquapore RP-300 column (Brownlee; 100 by 1 mm) and detected on line by an ion spray mass spectrometer (model AP-III; SCIEX).

Data bank search. NBRF, EMBL, and Swissprot sequence data bases were searched by using the GCG Package (GCG Inc., Madison, Wis.).

RESULTS

Cloning. Two mutants (Rib5 and Rib10) obtained after mutagenesis of *E. coli* RR28 with methyl ethane sulfonate (14) could grow with riboflavin or diacetyl, respectively. On the basis of current knowledge of the biosynthetic pathway of riboflavin, it was assumed that these mutants should have a defect of 6,7-dimethyl-8-ribityllumazine synthase or 3,4-dihydroxy-2-butanone 4-phosphate synthase (Fig. 1). Both enzymes are required for the conversion of 5-amino-6-ribitylamino-2,4(1*H*,3*H*)-pyrimidinedione (compound 2) to 6,7-dimethyl-8-ribityllumazine (compound 6), but their action can be bypassed by a nonenzymatic reaction of the pyrimidine precursor with exogenous diacetyl (16, 20).

A gene bank was constructed by partial digestion of chromosomal DNA of *E. coli* RR28 with *Sau*3AI endonuclease followed by ligation of the fragments into the *Bam*HI site of the plasmid pBluescript II SK−. The gene bank was used to transform the mutants Rib5 and Rib10. A total of 94 colonies growing without riboflavin were obtained. They were found to harbor plasmids containing inserts with sizes of 6 kb or more. Plasmid p10-16 with a 6-kb insert was selected for further study. This plasmid complemented the riboflavin deficiency of both *rib* mutants (Rib5 and Rib10) under study. The recombinant strains contained high levels of 3,4-dihydroxy-2-butanone 4-phosphate synthase activity (about 90-fold higher than the wild-type levels) (Table 2), whereas the activity of 6,7-dimethyl-8-ribityllumazine synthase remained below the level of detection. This suggested that the plasmid harbors the gene for 3,4-dihydroxy-2-bu-
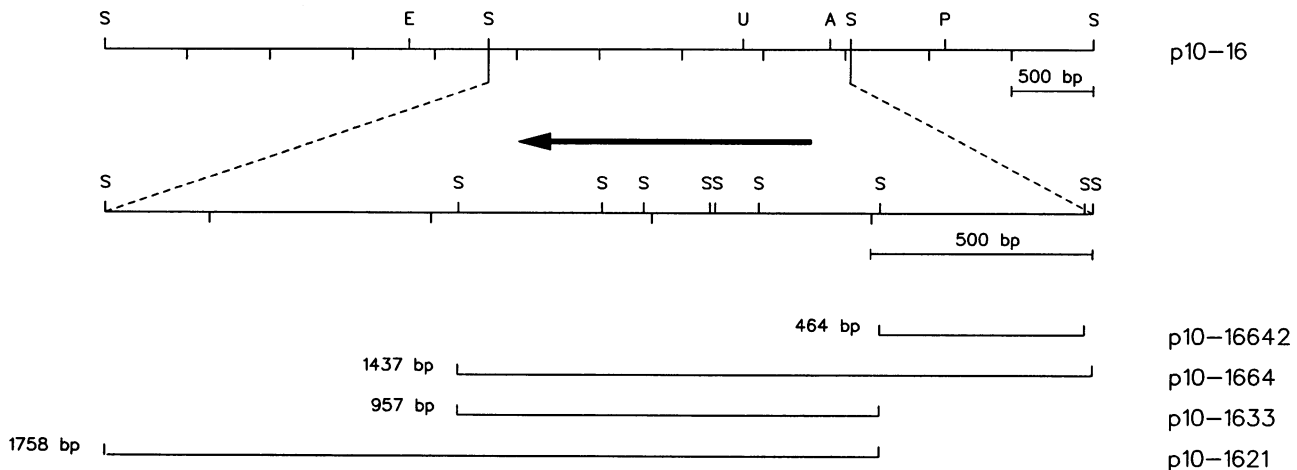

FIG. 2. DNA inserts of plasmids used in this study. The open reading frame coding for 3,4-dihydroxy-2-butanone 4-phosphate synthase is indicated (arrow). Selected restriction sites are also shown. A, *Aat*II; E, *Eco*RV; P, *Pst*I; S, *Sau*3AI; U, *Asu*II.

```
                 Sau3AI
GATCTGGCTGCTTTTTGATCTCAGTACTGCGATTTTCCAGTTCGCTGATTCTTTGTTCAA   60
                                        AatII
GCAGCGCCAGTTTTTCACGAGTCCGTAACAGGACTTGCGTTTGGACGTCGAACTCTTCAC  120
GGCTTACAAGGTCGAGGCGCGTCACGTGCGCTTGTAGGGTTTGGCGGATTTTTTTCTCCA  180
CATCTTCCCCGAACTCCCTGATTCCTTTAGGCATTGATTCGTGAACCTGGCGAGCGATTT  240
GCTCAATTTTTTTCGGGTCAATCATTGTAGTTTCCCTGTACTGATAGGTGTTGAGTGCCA  300
TTGTAGTGCGATAAGGGTAAGTCATAAACCAGAATTATGTGAAGCTATGCGTTCGTGCCG  360
CTAATCATTAGCGTTATAGTGAATCCGCTTATTCTCAGGGCGGGGCGAAATTCCCCACCG  420
GCGGTAAATCAACTCAGTTGAAAGCCGCGAGCGCTTTGGGTGCGAACTCAAAGGACAGCA  480
Sau3AI
GATCCGGTGTAATTCCGGGGCCGACGGTTAGAGTCCGGATGGGAGAGAGTAACGATTCTG  540
TCGGGCATGGACCCGCTCACGTTATTTTTGGCTATATGCCGCCACTCCTAAGACTGCCCTG  600
ATTCTGGTAACCATAATTTTAGTGAGGTTTTTTTACCATGAATCAGACGCTACTTTCCTC  660
                         S.D.        M  N  Q  T  L  L  S  S    8
        AsuII
TTTTGGTACGCCTTTCGAACGTGTTGAAAATGCACTGGCTGCGCTGCGTGAAGGACGCGG  720
   F  G  T  P  F  E  R  V  E  N  A  L  A  A  L  R  E  G  R  G   28

TGTAATGGTGCTTGATGATGAAGACCGTGAAAACGAAGGTGATATGATCTTCCCGGCAGA  780
   V  M  V  L  D  D  E  D  R  E  N  E  G  D  M  I  F  P  A  E   48

AACCATGACTGTTGAGCAGATGGCGCTGACCATTCGCCACGGTAGCGGTATTGTTTGCCT  840
   T  M  T  V  E  Q  M  A  L  T  I  R  H  G  S  G  I  V  C  L   68

GTGCATTACTGAAGATCGCCGTAAACAACTCGATCTGCCAATGATGGTAGAAAATAACAC  900
   C  I  T  E  D  R  R  K  Q  L  D  L  P  M  M  V  E  N  N  T   88

CAGCGCCTATGGCACCGGTTTTACCGTGACCATTGAAGCAGCTGAAGGTGTGACTACCGG  960
   S  A  Y  G  T  G  F  T  V  T  I  E  A  A  E  G  V  T  T  G  108

TGTTTCTGCCGCTGACCGTATTACGACCGTTCGCGCAGCGATTGCCGATGGCGCAAAACC 1020
   V  S  A  A  D  R  I  T  T  V  R  A  A  I  A  D  G  A  K  P  128

GTCAGATCTGAATCGTCCTGGCCACGTTTTCCCACTTCGCGCTCAGGCAGGTGGTGTACT 1080
   S  D  L  N  R  P  G  H  V  F  P  L  R  A  Q  A  G  G  V  L  148

GACGCGTGGCGGTCATACTGAAGCAACTATTGATCTGATGACGCTGGCAGGCTTTAAACC 1140
   T  R  G  G  H  T  E  A  T  I  D  L  M  T  L  A  G  F  K  P  168

GGCTGGTGTACTGTGTGAGCTGACTAATGACGATGGCACGATGGCGCGTGCACCAGAGTG 1200
   A  G  V  L  C  E  L  T  N  D  D  G  T  M  A  R  A  P  E  C  188

TATTGAGTTTGCCAATAAACACAATATGGCGCTCGTGACTATTGAAGACCTGGTGGCATA 1260
   I  E  F  A  N  K  H  N  M  A  L  V  T  I  E  D  L  V  A  Y  208

CCGTCAGGCACATGAGCGTAAAGCCAGCTGAAAACCGCTGCTTAATTTACTGCCTTAATC 1320
   R  Q  A  H  E  R  K  A  S                                    217

AAGAAACCGAAGTTGTAGCAGGCTTCGGTTTTTATTTTTCCCTGCTATGCAAGATGTTAA 1380
TTCCGCTTCTCTTTCTATGAGAAAATTTCATTAATATCAGGCATTCTTTTTTCATTAT   1437
```

FIG. 3. Nucleotide sequence of the gene coding for 3,4-dihydroxy-2-butanone 4-phosphate synthase and its flanking regions. The putative Shine-Dalgarno (S.D.) region is double underlined. The deduced amino acid sequence is indicated. The N-terminal peptide sequence determined by Edman degradation is underlined.

tanone 4-phosphate synthase. It should be noted that the specific activity in cell extracts from the recombinant strains showed some variation in different experiments.

Subcloning and sequencing. Three subclones (plasmids p10-1621, p10-1633, and p10-1664) complementing the riboflavin deficiency of the mutant were obtained as described in Materials and Methods (Table 1 and Fig. 2). Bacterial cells carrying plasmid p10-1664 expressed high levels of 3,4-dihydroxy-2-butanone 4-phosphate activity. The enzyme levels of cells carrying plasmids p10-1621 and p10-1633 were similar to the enzyme level of cells without plasmids (Table 2).

TABLE 3. Purification of 3,4-dihydroxy-2-butanone 4-phosphate synthase from *E. coli* Rib10 carrying plasmid p10-16

Procedure	Vol (ml)	Activity (U)	Amt of protein (mg)	Sp act (U/mg)
Cell extract	460	125,000	2,016	62
DEAE-cellulose	127	80,000	124	645
SP-Sephadex	100	65,000	65	1,000
Superdex 75[a]				1,700

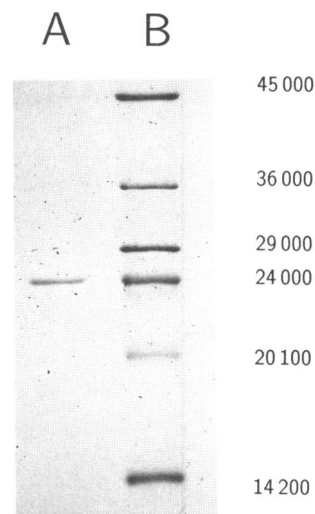[a] 200-μg aliquots were purified, with a yield of about 30%.

FIG. 4. Sodium dodecyl sulfate-polyacrylamide gel electrophoresis. Lane A, purified 3,4-dihydroxy-2-butanone 4-phosphate synthase; lane B, marker proteins. Molecular weights are indicated.

Plasmids p10-1664, p10-1621, p10-1633, and p10-16642 were sequenced by the dideoxy nucleotide method using primer walk strategies. An open reading frame consisting of 651 bp was present in all plasmids studied (Fig. 2 and 3). The deduced amino acid sequence codes for a peptide of 217 amino acid residues with a calculated mass of 23,349.6 Da.

Enzyme purification. 3,4-Dihydroxy-2-butanone 4-phosphate synthase was purified to near homogeneity from Rib10 cells carrying the plasmid p10-16 as described in Materials and Methods. Briefly, a sequence of anion exchange chromatography and cation exchange chromatography resulted in an approximately 17-fold enrichment. Subsequent FPLC on a gel permeation column yielded an enzyme with a specific activity of 1,700 nmol mg^{-1} h^{-1} at 37°C. A typical experiment is summarized in Table 3.

Enzyme properties. The purified enzyme gave a single band of 24 kDa in sodium dodecyl sulfate-polyacrylamide gel electrophoresis (Fig. 4). The accurate mass of the protein was determined by ion spray mass spectrometry (Fig. 5). Signals were observed for multiply charged species ranging from $(M + 8H)^{8+}$ to $(M + 29H)^{29+}$. An average molecular mass of 23,351 ± 2 Da was calculated from these data and corresponds to the predicted mass (23,349.6 Da) within the limits of experimental error.

An aliquot of the protein was hydrolyzed, and the amino acid composition was determined to be in close agreement with the predicted amino acid composition (data not shown). Automated Edman degradation gave a sequence of 39 N-terminal amino acids which was identical with the deduced amino acid sequence (Fig. 3).

DISCUSSION

We have cloned a DNA segment from *E. coli* which complements the metabolic defect of two *E. coli* mutants that could grow with riboflavin or diacetyl, respectively. On the basis of sequence analysis of three subclones, the complementation of the genetic defect of both mutants can be attributed conclusively to an open reading frame coding for a protein of 217 amino acids.

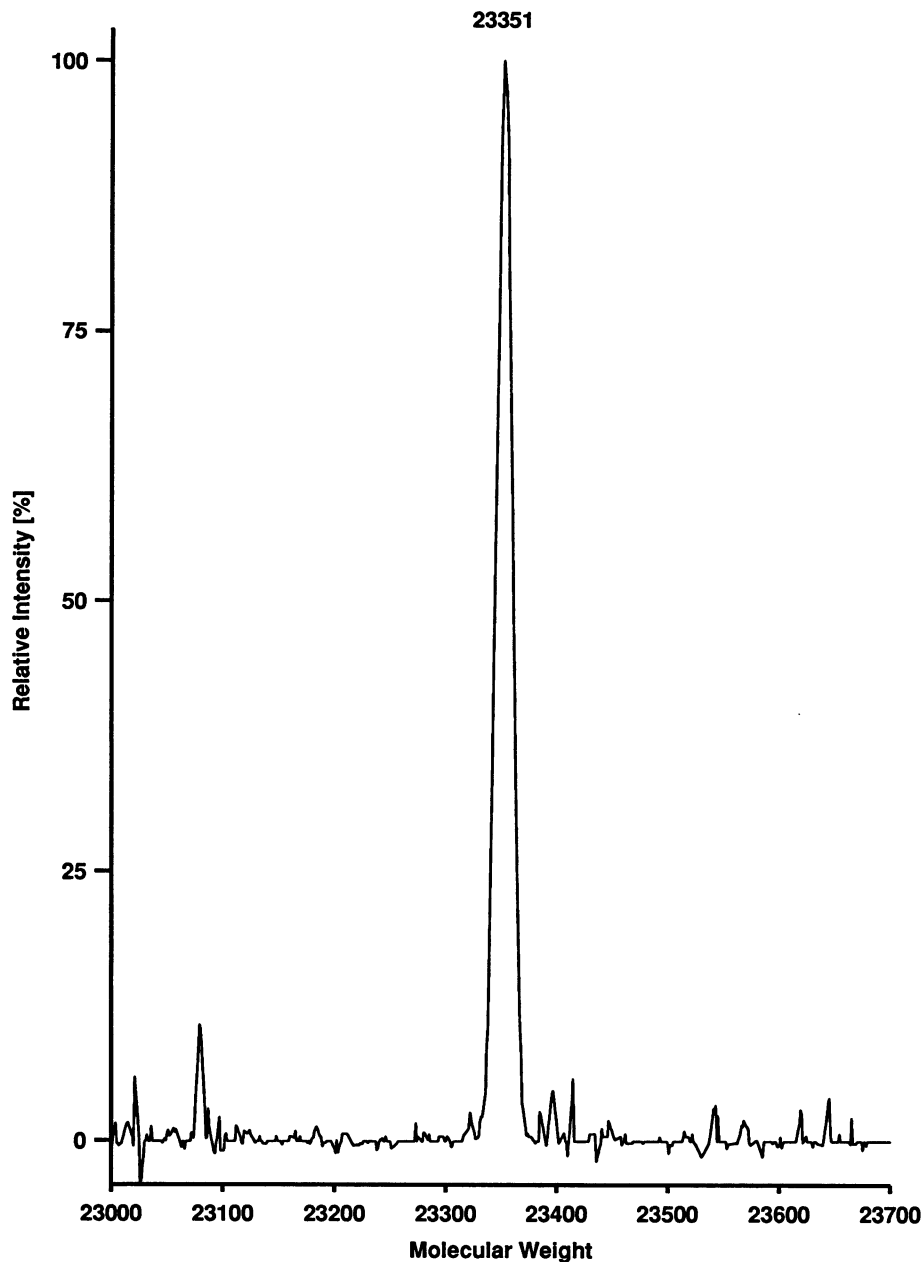The level of 3,4-dihydroxy-2-butanone 4-phosphate syn-

FIG. 5. Ion spray mass spectrum of 3,4-dihydroxy-2-butanone 4-phosphate synthase. Thirty-two scans over the HPLC peak of the protein were averaged, and the spectrum was reconstructed by the Fenn method (17).

thase activity is increased approximately 90-fold in strains carrying plasmid p10-16 or p10-1664 compared with that of the *E. coli* wild-type strain. The protein was purified to near homogeneity from a recombinant strain carrying the plasmid p10-16. The N-terminal sequence of the protein obtained by Edman degradation was identical to the predicted sequence of the cloned gene product. The total amino acid composition was in agreement with the predicted composition. Sodium dodecyl sulfate-polyacrylamide gel electrophoresis showed a single 24-kDa band. Mass spectroscopic analysis gave a molecular mass of 23,351 ± 2 Da, which is in excellent agreement with the predicted mass of 23,349.6 Da. Thus, it is unequivocally established that the cloned gene codes for 3,4-dihydroxy-2-butanone 4-phosphate synthase.

Whereas high levels of 3,4-dihydroxy-2-butanone 4-phosphate synthase were found in *E. coli* strains carrying plasmids p10-16 and p10-1664, Rib10 mutants carrying plasmid p10-1621 or p10-1633 had enzyme levels comparable to that of the *E. coli* wild-type strain despite the high copy number of the plasmids (Table 3). In the plasmids with low enzyme activity, the gene is preceded by a sequence of 158 bp. Apparently, this sequence allows for only a very low level of transcription. Thus, it appears that the promoter is located in the 464-bp *Sau*3AI segment, which is present in plasmids p10-16 and p10-1664 but not in plasmids p10-1621 and p10-1633. This implies a location of the promoter at a position more than 158 bp upstream from the start codon. Studies by Raina et al. (21) indicated the presence of an

```
E. coli      MNQTLLSSFGTPFERVENALAALREGRGVMVLDDEDRENEGDMIFPAETMTVEQMALTI
V. harveyi   MSSTSLLDEFGTPVQRVERAIEALKNGLGVLLMDDEDRENEGDLIFSAQHLTEAQMALMI
B. subtilis     MFHPIEEALDALKKGEVIIVVDDEDRENEGDFVALAEHATPEVINFMA

E. coli      RHGSGIVCLCITEDRRKQLDLPMMVENNTSAYGTGFTVTIEAAEGVTTGVSAADRITTVR
V. harveyi   REGSGIVCLCLTEERANWLDLPPMVKDNCSKNQTAFTVSIEAKEGVTTGVSAKDRVTTVK
B. subtilis  THGRGLICTPLSEEIADRLDLHPMVEHNTDSHHTAFTVSIDHRETK-TGISAQERSFTVQ

E. coli      AAIADGAKPSDLNRPGHVFPLRAQAGGVLTRGGHTEATIDLMTLAGFKPAGVLCELTNDD
V. harveyi   TATYFDAQPEDLARPGHVFPLVAKTNGVLARRGHTEGTIDLMYLANLVPSGILCELTNPD
B. subtilis  ALLDSKSVPSDFQRPGHIFPLIAKKGGLLKRAGHTEAVDLAEACGSPGAGVICEIMNED

E. coli      GTMARAPECIEFANKHNMALVTIEDLVAYRQAHERKAS
V. harveyi   GTMAKLPETIEFARRHGMPVLTIEDIVDYRTGIDLRNEYKSGLVREVSWS
B. subtilis  GTMARVPELIEIAKKHQLKMITIKDLIQYRYNLTTLVEREVDITLPTDFGTFKVYGYTNE
```
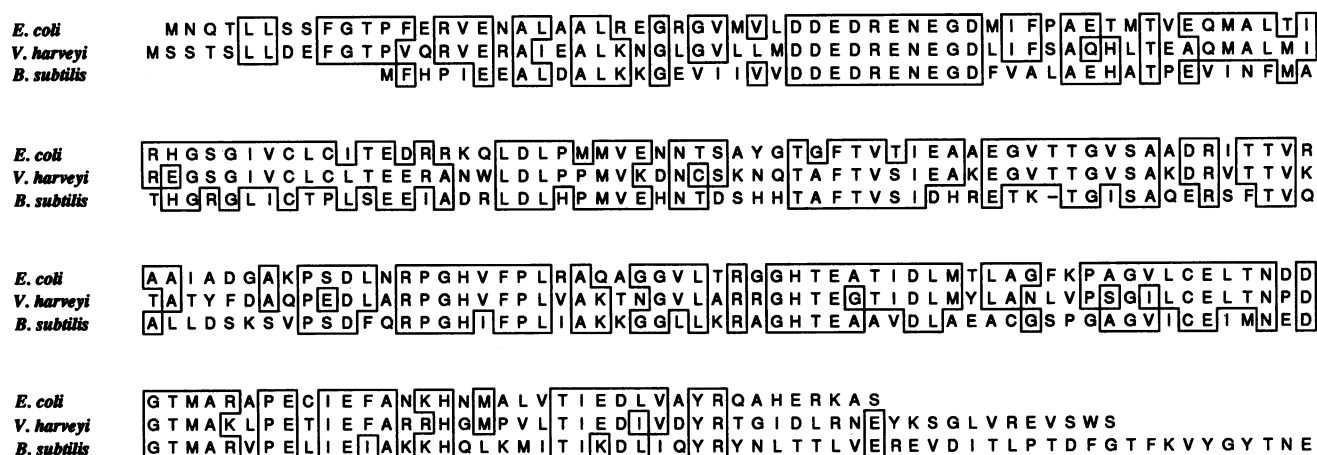
FIG. 6. Alignment of predicted peptide sequences of 3,4-dihydroxy-2-butanone 4-phosphate synthase from different microorganisms. Identical residues are boxed. Only the N-terminal segment of the *B. subtilis* gene is shown.

active promoter in the DNA segment between the *Aat*II and *Asu*II restriction sites shown in Fig. 3, i.e., between nucleotides 109 and 675 (for details, see below). Taken together, these data suggest that the promoter is located 148 to 529 bp before the start codon. Moreover, a weak promoter may be present in the 148 bp directly preceding the open reading frame. A search for sequence motifs related to the consensus promoter sequence gave no positive results.

A search of the EMBL data bank (release 29, 3 October 1991) showed that the sequence of the open reading frame coding for 3,4-dihydroxy-2-butanone 4-phosphate synthase was identical to the sequence of a *luxH*-like gene of *E. coli* (accession number M77129). The *luxH*-like gene was one of five open reading frames present in a segment of 5,058 bp which has been sequenced by Yang and Depew (31). Nothing has been reported on the function of the *luxH*-like gene. It should be noted that the 5' region of the sequence reported in this paper shows minor differences with that of the *luxH*-like sequence, whereas the open reading frame is identical.

Moreover, the sequence was almost identical with the published sequence of the *htrP* gene of *E. coli*, presumed to code for a heat shock protein with a predicted mass of 27 kDa and a predicted inner membrane location. An *htrP* mutant was unable to grow at 37°C and could grow slowly at 30°C. The mutants Rib5 and Rib10 used in this study show no growth whatsoever at 30°C. The original *htrP* mutant (21) resulted from the insertion of transposon Tn*10* at a position 13 codons from the 3' end of the gene coding for 3,4-dihydroxy-2-butanone 4-phosphate synthase. It appears likely that this modification compromised the thermal stability and also led to reduced catalytic activity of the enzyme at 30°C. A strain with a presumed insertion of the ΩKan cassette in the central part of the chromosomal *htrP* gene displayed the same thermosensitive phenotype as the original *htrP* mutant (21). This observation remains unexplained. The evidence indicating that the gene codes for 3,4-dihydroxy-2-butanone 4-phosphate synthase is unequivocal, and the insertion of the ΩKan cassette into the central part of the gene, if successful, should have resulted in a phenotype with an absolute requirement for riboflavin or diacetyl.

The *htrP* gene has been located at 66.3 min of the *E. coli* chromosome, in the neighborhood of the *tolC* gene (21). Earlier, a mutation causing riboflavin deficiency in *E. coli*

had been mapped in the close vicinity of the *tolC* gene by transduction analysis (3, 27). This gene has been designated *ribB*, but its precise function has not been established. In light of the proposed map position, it appears likely that *ribB* is identical with the gene described in this study. In summary, each of the three published gene loci, *luxH*-like, *htrP*, and *ribB*, appear to be identical with the gene coding for 3,4-dihydroxy-2-butanone 4-phosphate synthase. However, the precise function of this gene had not been determined previously. On the basis of the results reported in this paper, we propose that *ribB* should be the preferred name for this gene.

The gene under study is homologous with the *luxH* gene of *V. harveyi* (26). The homology extends over the entire length of both genes (Fig. 6). The *luxH* gene is located at the 3' end of the luciferase operon, which contains seven open reading frames. The genes *luxA* to *luxE* code for luciferase and for enzymes involved in the formation of luciferase substrates. The roles of the genes *luxG* and *luxH* have not been hitherto determined. On the basis of the sequence homology, it is likely that *luxH* codes for 3,4-dihydroxy-2-butanone 4-phosphate synthase. The gene *luxG* is not homologous to known genes involved in the biosynthesis of riboflavin.

Bacterial luciferase utilizes reduced flavin mononucleotide as a substrate. Expression of the luciferase operon may therefore result in an increased demand for riboflavin. It is not clear why only one of the six enzymes required for riboflavin biosynthesis is under the control of the luciferase operator, and it has been shown that *luxH* is not required for bioluminescence activity (26). It is possible that the reaction step catalyzed by 3,4-dihydroxy-2-butanone 4-phosphate synthase is rate limiting for the biosynthesis of riboflavin in *V. harveyi*.

The gene under study is also homologous to one of the five open reading frames in the riboflavin operon of *B. subtilis* (18). This open reading frame (ORF3) codes for a predicted sequence of 398 amino acids. The homology to *ribB* of *E. coli* extends over the 198 N-terminal amino acid residues (Fig. 6). The C-terminal part of this open reading frame is homologous to the gene coding for GTP cyclohydrolase II of *E. coli*, i.e., the first enzyme of the riboflavin biosynthetic pathway (23).

Earlier genetic studies had attributed a gene (*ribA*) coding for GTP cyclohydrolase II to this area of the *B. subtilis* rib

operon (5, 6, 18, 19). More recently, a gene (*ribF*) was also tentatively located in this area. Earlier, this gene had been postulated to occur at various positions closer to the 5' end of the *rib* operon (for a review, see reference 1).

It has been claimed that *ribF* codes for a presumptive 6-methyl-7-dihydroxyethyl-8-ribityllumazine synthetase (4). The purported product of this hypothetical enzyme is in all likelihood not an intermediate in the biosynthesis of riboflavin, and the existence of the proposed enzyme appears doubtful. The observed sequence homology suggests that ORF3 of the *B. subtilis* operon codes for a bifunctional enzyme with GTP cyclohydrolase II and 3,4-dihydroxy-2-butanone 4-phosphate synthase activity.

ACKNOWLEDGMENTS

REFERENCES

1. **Bacher, A.** 1990. Biosynthesis of flavins, p. 215–259. *In* F. Müller (ed.), Chemistry and biochemistry of flavoenzymes, vol. 1. CRC Press, Boca Raton, Fla.
2. **Bacher, A., R. Baur, U. Eggers, H. Harders, M. K. Otto, and H. Schnepple.** 1980. Riboflavin synthases of *Bacillus subtilis*. Purification and properties. J. Biol. Chem. **255**:632–637.
3. **Bandrin, S. V., P. M. Rabinovich, and A. I. Stepanov.** 1983. Three linkage groups of genes for riboflavin biosynthesis in *Escherichia coli*. Genetika **19**:1419–1425.
4. **Bresler, S. E., D. A. Perumov, T. P. Chernik, and E. A. Glazunov.** 1976. Investigation of the operon of riboflavin biosynthesis in *Bacillus subtilis*. X. Genetic and biochemical study of mutants that accumulate 6-methyl-7-(1',2'-dihydroxyethyl)-8-ribityllumazine. Genetika **12**:83–91.
5. **Chikindas, M. L., E. V. Luk'yanov, P. M. Rabinovich, and A. I. Stepanov.** 1987. Study of the region 210° of the *Bacillus subtilis* chromosome with the aid of recombinant plasmids. Mol. Genet. Mikrobiol. Virusol. **2**:20–24.
6. **Chikindas, M. L., V. N. Mironov, E. V. Luk'yanov, Y. R. Borestskii, L. S. Arutyunova, P. M. Rabinovich, and A. I. Stepanov.** 1987. Establishment of the boundaries of the riboflavin operon of *Bacillus subtilis*. Mol. Genet. Mikrobiol. Virusol. **4**:22–26.
7. **Davis, L. G., M. D. Dibner, and J. F. Battey.** 1986. Basic methods in molecular biology. Elsevier, New York.
8. **Godson, G. N., and D. Vapnek.** 1973. A simple method of preparing large amounts of ϕX 174RF1 supercoiled DNA. Biochim. Biophys. Acta **299**:516–520.
9. **Hanahan, D.** 1983. Studies on transformation of *Escherichia coli* with plasmids. J. Mol. Biol. **166**:557–580.
10. **Hanahan, D.** 1985. Techniques for transformation of *Escherichia coli*, p. 109–135. *In* D. M. Glover (ed.), DNA cloning, vol. 1. IRL Press, Oxford.
11. **Hennecke, H., I. Günther, and F. Binder.** 1982. A novel cloning vector for the direct selection of recombinant DNA in *Escherichia coli*. Gene **19**:231–234.

12. **Hewick, R. M., M. W. Hunkapiller, L. E. Hood, and W. J. Dreyer.** 1981. A gas-liquid solid phase peptide and protein sequenator. J. Biol. Chem. **256**:7990–7997.
13. **Holmes, D. S., and M. Quigley.** 1981. A rapid boiling method for the preparation of bacterial plasmids. Anal. Biochem. **114**:193–197.
14. **Katzenmeier, G.** 1991. Ph.D. thesis. Technical University of Munich, Garching, Germany.
15. **Laemmli, U. K.** 1970. Cleavage of structural proteins during the assembly of the head of bacteriophage T4. Nature (London) **227**:680–685.
16. **Lingens, F., O. Oltmanns, and A. Bacher.** 1967. Über Zwischenprodukte der Riboflavin-Biosynthese bei *Saccharomyces cerevisiae*. Z. Naturforsch. Sect. B Chem. Sci. **22**:755–758.
17. **Mann, M., C. K. Meng, and J. B. Fenn.** 1989. Interpreting mass spectra of multiply charged ions. Anal. Chem. **61**:1702–1708.
18. **Mironov, V. N., M. L. Chikindas, A. S. Kraev, A. I. Stepanov, and K. G. Skryabin.** 1989. Operon organization of genes of riboflavin biosynthesis in *Bacillus subtilis*. Mol. Biol. (Moscow) **312**:237–240.
19. **Morozov, G. I., P. M. Rabinovich, S. V. Bandrin, and A. I. Stepanov.** 1984. Structure of *Bacillus subtilis* riboflavin operon. Mol. Genet. Mikrobiol. Virusol. **7**:42–46.
20. **Oltmanns, O., A. Bacher, F. Lingens, and F. K. Zimmermann.** 1969. Biochemical and genetic classification of riboflavin deficient mutants of *Saccharomyces cerevisiae*. Mol. Gen. Genet. **105**:306–313.
21. **Raina, S., L. Mabey, and C. Georgopoulos.** 1991. The *Escherichia coli htrP* gene product is essential for bacterial growth at high temperatures: mapping, cloning, sequencing and transcriptional regulation of *htrP*. J. Bacteriol. **173**:5999–6008.
22. **Read, S. M., and D. H. Northcote.** 1981. Minimization of variation in the response to different proteins of the Coomassie blue G dye-binding assay for protein. Anal. Biochem. **116**:53–64.
23. **Richter, G., H. Ritz, G. Katzenmeier, and A. Bacher.** Unpublished data.
24. **Schott, K., J. Kellermann, F. Lottspeich, and A. Bacher.** 1990. Riboflavin synthases of *Bacillus subtilis*. Purification and amino acid sequence of the α subunit. J. Biol. Chem. **265**:4204–4209.
25. **Spackman, D. H., W. H. Stein, and S. Moore.** 1958. Automatic recording apparatus for use in the chromatography of amino acids. Anal. Chem. **30**:1190–1206.
26. **Swartzman, E., C. Miyamoto, A. Graham, and E. Meighen.** 1990. Delineation of the transcriptional boundaries of the lux operon of *Vibrio harveyi* demonstrates the presence of two new lux genes. J. Biol. Chem. **265**:3513–3517.
27. **Teslyar, G. E., and G. M. Shavlovskii.** 1983. Location of the genes coding for riboflavin synthase and GTP-cyclohydrolase on the chromosome of *Escherichia coli* K12. Tsitol. Genet. **5**:54–56.
28. **Volk, R., and A. Bacher.** 1988. Biosynthesis of riboflavin. The structure of the 4-carbon precursor. J. Am. Chem. Soc. **110**:3651–3653.
29. **Volk, R., and A. Bacher.** 1990. Studies on the four carbon precursor in the biosynthesis of riboflavin. Purification and properties of L-3,4-dihydroxy-2-butanone 4-phosphate synthase. J. Biol. Chem. **265**:19479–19485.
30. **Volk, R., and A. Bacher.** 1991. Biosynthesis of riboflavin. Studies on the mechanism of L-3,4-dihydroxy-2-butanone 4-phosphate synthase. J. Biol. Chem. **266**:20610–20618.
31. **Yang, T. P., and R. E. Depew.** Unpublished data.