

## The Calvin Cycle Enzyme Pentose-5-Phosphate 3-Epimerase Is Encoded within the *cfx* Operons of the Chemoautotroph *Alcaligenes eutrophus*

BERNHARD KUSIAN, JE-GEUN YOO, RALPH BEDNARSKI, AND BOTHO BOWIEN\*

*Institut für Mikrobiologie, Georg-August-Universität Göttingen, Grisebachstrasse 8,  
W-3400 Göttingen, Germany*

Received 13 July 1992/Accepted 16 September 1992

Several genes (*cfx* genes) encoding Calvin cycle enzymes in *Alcaligenes eutrophus* are organized in two highly homologous operons comprising at least 11 kb. One *cfx* operon is located on the chromosome; the other is located on megaplasmid pHG1 of the organism (B. Bowien, U. Windhövel, J.-G. Yoo, R. Bednarski, and B. Kusian, FEMS Microbiol. Rev. 87:445-450, 1990). Corresponding regions of about 2.7 kb from within the operons were sequenced. Three open reading frames, designated *cfxX* (954 bp), *cfxY* (765 bp), and *cfxE* (726 bp), were detected at equivalent positions in the two sequences. The nucleotide identity of the sequences amounted to 94%. Heterologous expression of the subcloned pHG1-encoded open reading frames in *Escherichia coli* suggested that they were functional genes. The observed sizes of the gene products CfxX (35 kDa), CfxY (27 kDa), and CfxE (25.5 kDa) closely corresponded to the values calculated on the basis of the sequence information. *E. coli* clones harboring the *cfxE* gene showed up to about 19-fold-higher activities of pentose-5-phosphate 3-epimerase (PPE; EC 5.1.3.1) than did reference clones, suggesting that *cfxE* encodes PPE, another Calvin cycle enzyme. These data agree with the finding that in *A. eutrophus*, PPE activity is significantly enhanced under autotrophic growth conditions which lead to a derepression of the *cfx* operons. No functions could be assigned to CfxX and CfxY.

When growing lithoautotrophically with hydrogen or organoautotrophically with formate as an energy source, the facultative chemoautotroph *Alcaligenes eutrophus* assimilates CO<sub>2</sub> via the reactions of the Calvin carbon reduction cycle (7). In strain H16, genes encoding enzymes of this cycle (*cfx* genes) are organized in two large, highly homologous *cfx* operons. One copy of the operon is located on the chromosome; the other is located on megaplasmid pHG1 adjacent to the hydrogenase gene cluster (9, 12, 20). Both operons, which possibly originate from a gene duplication event, are functional and expressed simultaneously. Each of them comprises at least 11 kb (48). The two promoter-proximal genes, *cfxL* and *cfxS*, encode the L and S subunits, respectively, of the CO<sub>2</sub>-fixing enzyme of the Calvin cycle, ribulose-1,5-bisphosphate carboxylase/oxygenase (RuBisCO), while the gene coding for glyceraldehyde-3-phosphate dehydrogenase (GAP), *cfxG*, is the most promoter-distal gene so far identified. Also located within the operons are the genes for fructose-1,6-bisphosphatase/sedoheptulose-1,7-bisphosphatase (FBP) (*cfxF*), phosphoribulokinase (PRK) (*cfxP*), and transketolase (TK) (*cfxT*) (Fig. 1). Transcription of the operons requires the activation by a regulatory protein, CfxR, the product of the *cfxR* gene, which is located immediately upstream of the chromosomal operon copy and oriented divergently to the latter (49). Full derepression or induction of the operons occurs only under autotrophic growth conditions, whereas complete or partial repression prevails during heterotrophic growth, depending on the organic substrate used (6, 29).

The information contained within the 2.7-kb DNA segment between the *cfxS* and *cfxF* genes of both operons was unknown and thus subjected to detailed analysis. Sequenc-

ing and heterologous gene expression revealed the existence of three genes in each of the two regions. The gene *cfxE*, upstream of *cfxF*, was identified to encode another Calvin cycle enzyme, pentose-5-phosphate 3-epimerase (PPE; D-ribulose-5-phosphate 3-epimerase; EC 5.1.3.1). To our knowledge, this represents the first report on the molecular cloning, sequencing, and identification of a PPE gene from any organism.

(A preliminary account of some of the data contained in this report has been presented elsewhere [9].)

### MATERIALS AND METHODS

**Strains, plasmids, and culture conditions.** The bacterial strains and plasmids used in this study are listed in Table 1. *A. eutrophus* H16 was grown in a mineral-salts medium at 30°C as described previously (48). The medium was supplemented with 0.2% (wt/vol) organic substrate for organoheterotrophic (fructose or pyruvate) or organoautotrophic (formate) growth. Lithoautotrophic cultivation of the organism was done by using a gas mixture of H<sub>2</sub>, CO<sub>2</sub>, and O<sub>2</sub> at a mixing ratio of 8:1:1 (vol/vol/vol) or 8.9:0.1:1 (vol/vol/vol) for CO<sub>2</sub>-limited growth. LB or XB medium (16) was used to propagate *Escherichia coli* at 37 or 30°C. The latter media contained antibiotics, when indicated, at the following concentrations: ampicillin, 50 or 200 µg/ml; kanamycin, 75 µg/ml; and tetracycline, 20 µg/ml.

**Preparation of cell extracts and assay of PPE.** Cell extracts were prepared at 0 to 4°C. *A. eutrophus* or *E. coli* cells harvested from mid-logarithmic-phase cultures were washed and resuspended in buffer (20 mM Tris-HCl [pH 7.6] containing 10 mM MgCl<sub>2</sub> and 1 mM dithioerythritol) at a density of about 20 mg of cell protein per ml. They were disrupted by either passage through a French pressure cell (*A. eutrophus*) or ultrasonication (*E. coli*). The supernatant resulting from a

\* Corresponding author.

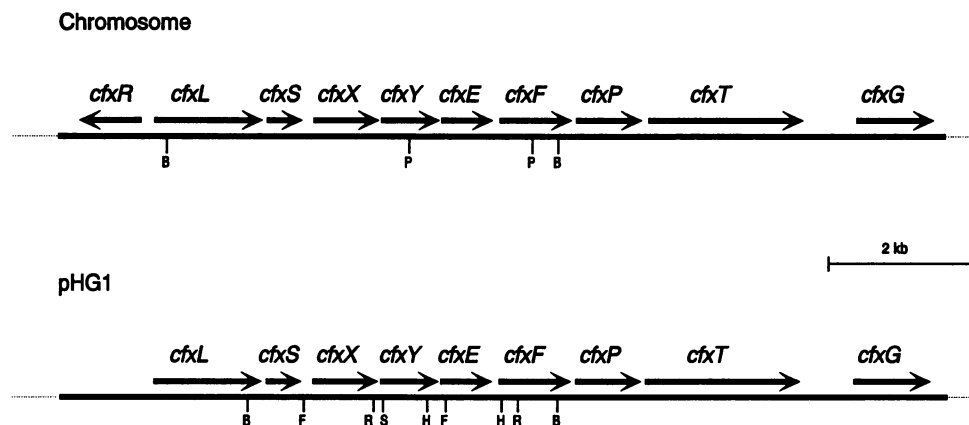


FIG. 1. Organization of the chromosomal and pHG1-encoded *cfx* gene clusters of *A. eutrophus* H16. The genes and their relative orientations are indicated by arrows. *cfxR*, activator gene; *cfxLS*, RuBisCO large- and small-subunit genes; *cfxX* and *cfxY*, genes of unknown functions; *cfxE*, gene for PPE; *cfxF*, gene for FBP; *cfxP*, gene for PRK; *cfxT*, gene for TK; *cfxG*, gene for GAP. The following restriction endonuclease sites were used for subcloning of genes: *Bam*HI (B), *Fok*I (F), *Hin*II (H), *Pst*I (P), *Rsa*I (R), and *Sal*I (S).

subsequent centrifugation of the homogenate at  $100,000 \times g$  for 1 h was used as the cell extract for assaying PPE activity and/or for polyacrylamide gel electrophoresis (PAGE). Protein concentrations were estimated by the method of Lowry et al. (32).

PPE was assayed at 30°C in a reaction mixture containing, in a total volume of 0.6 ml, 50 mM Tris-HCl (pH 7.8), 10 mM MgCl<sub>2</sub>, 0.5 mM thiamine pyrophosphate, 0.25 mM NADH, 2 mM ribose-5-phosphate, 2 mM ribulose-5-phosphate, 1 U of TK, 3 U of GAP, 9 U of triosephosphate isomerase, and 0.005 to 0.02 mg of cell extract protein. The reaction was started by the final addition of the mixed pentose phosphates

and monitored in a spectrophotometer (Uvikon 810; Kontron, Eching, Germany) at 340 or 365 nm.

**Electrophoretic separations of proteins.** One-dimensional separation of proteins was carried out by sodium dodecyl sulfate (SDS)-PAGE (28); two-dimensional PAGE was performed as described by O'Farrell (35) as a combination of isoelectric focusing (Mini-IEF cell; Biometra, Göttingen, Germany) and SDS-PAGE (Minielectrophoresis cell; Biometra). Silver staining (5) was used to visualize proteins in gels. Radioactive proteins in gels were detected by autoradiography (Kodak X-Omat AR film; Kodak, Stuttgart, Germany).

TABLE 1. Bacterial strains and plasmids

Strain or plasmid	Relevant genotype or phenotype <sup>a</sup>	Source or reference
<b>Strains</b>		
<i>A. eutrophus</i> H16	Cfx Hox; pHG1	ATCC 17699
<i>E. coli</i> JW1	<i>ara strA thi Δ(lac-proAB) (Φ80 lacZΔM15) F'[proAB<sup>+</sup> lacI<sup>q</sup> lacZΔM15]</i>	25
<i>E. coli</i> K38	HfrC(λ) T2 <sup>r</sup> ; <i>phoA6 tonA22 garB10 ompF627 relA1 pit-10 spoT1 PO2A</i>	19
<b>Plasmids</b>		
pUC9	Ap <sup>r</sup> ; <i>lacPOZ'</i>	44
pUC18/19	Ap <sup>r</sup> ; <i>lacPOZ'</i>	51
pT7-7	Ap <sup>r</sup> ; T7 RNA polymerase promoter and translation start	S. Tabor
pGP1-2	Km <sup>r</sup> ; <i>P<sub>lac</sub>/cI857, P<sub>L</sub>/T7</i> gene 1	42
pAEC1180	Ap <sup>r</sup> ; <i>lacPOZ'</i> ; chromosomal 5.4-kb <i>Bam</i> HI fragment from <i>A. eutrophus</i> inserted into <i>pUC19</i> with <i>cfx</i> genes collinear to <i>lacPO</i>	This study
pAEC3010/3011	Ap <sup>r</sup> ; <i>lacPOZ'</i> ; chromosomal 1.7-kb <i>Pst</i> I fragment inserted into pUC9 with <i>cfxE<sub>c</sub></i> collinear/in divergent orientation to <i>lacPO</i>	This study
pAEP3050/3051	Ap <sup>r</sup> ; <i>lacPOZ'</i> ; 4.2-kb <i>Bam</i> HI fragment from pHG1 of <i>A. eutrophus</i> inserted into pUC19 with <i>cfx</i> genes collinear/in divergent orientation to <i>lacPO</i>	This study
pAEP9010/9011	Ap <sup>r</sup> ; <i>lacPOZ'</i> ; 1.1-kb <i>Fok</i> I- <i>Sal</i> I fragment from pHG1 inserted into pUC18 with <i>cfxX<sub>p</sub></i> collinear/in divergent orientation to <i>lacPO</i>	This study
pAEP9012	Ap <sup>r</sup> ; 1.1-kb insert (as <i>Eco</i> RI- <i>Bam</i> HI fragment) from pAEP9010 recloned into pT7-7 with <i>cfxX<sub>p</sub></i> collinear to <i>P<sub>T7</sub></i>	This study
pAEP9020/9021	Ap <sup>r</sup> ; <i>lacPOZ'</i> ; 0.9-kb <i>Rsa</i> I- <i>Fok</i> I fragment from pHG1 inserted into pUC18 with <i>cfxY<sub>p</sub></i> collinear/in divergent orientation to <i>lacPO</i>	This study
pAEP9030/9031	Ap <sup>r</sup> ; <i>lacPOZ'</i> ; 1.0-kb <i>Hin</i> II fragment from pHG1 inserted into pUC18 with <i>cfxE<sub>p</sub></i> collinear/in divergent orientation to <i>lacPO</i>	This study
pAEP9230	Ap <sup>r</sup> ; <i>lacPOZ'</i> ; 2.0-kb <i>Rsa</i> I fragment from pHG1 inserted into pUC18 with <i>cfxYE<sub>p</sub></i> collinear to <i>lacPO</i>	This study

<sup>a</sup> Cfx, ability to fix CO<sub>2</sub>; Hox, ability to oxidize H<sub>2</sub>; pHG1, megaplasmid pHG1 of *A. eutrophus* (13); Ap<sup>r</sup>, ampicillin resistant; Km<sup>r</sup>, kanamycin resistant; Tc<sup>r</sup>, tetracycline resistant; T2<sup>r</sup>, phage T2 resistant.

**Gene expression.** For gene expression experiments with *E. coli* strains harboring hybrid pUC plasmids, clones were grown in LB medium containing ampicillin until the cultures attained an optical density of 0.5 measured at 550 nm. After supplementation with 0.5 mM isopropyl- $\beta$ -D-thiogalactopyranoside (IPTG), incubation was continued for additional 4 h, and the cells were subsequently harvested for the preparation of extracts.

Expression of *cfxX<sub>p</sub>* in *E. coli* was achieved by using the phage T7 RNA polymerase promoter system (42) with *cfxX<sub>p</sub>* cloned into vector pT7-7. (In gene designations, subscripts "p" and "c" indicate plasmid pHG1 encoded and chromosomal, respectively.) Growth of the corresponding *E. coli* K38 transformants [*E. coli*(pGP1-2) and *E. coli*(pAEP9012)], induction of T7 RNA polymerase, and in vivo protein labeling with L-[<sup>35</sup>S]methionine were performed essentially as described previously (49).

**DNA preparation and manipulations.** Large-scale isolation of plasmid DNA was done by the alkaline-SDS lysis method (3). The rapid-boiling procedure (18) was used for plasmid minipreparations. DNA manipulations for cloning purposes were performed by standard protocols (2, 39), and enzymes were used under the conditions recommended by the commercial suppliers. Fragments were extracted from agarose gels by elution with glass milk (45) after electrophoretic separation.

**Construction of plasmids.** Plasmid vectors (pUC and pT7-7) were digested to completion with the appropriate restriction endonuclease(s) and dephosphorylated by alkaline phosphatase treatment. The various DNA fragments to be cloned (Table 1) were made blunt ended, if necessary, by using the Klenow fragment of DNA polymerase I and subsequently ligated to the vectors with T4 DNA ligase. Ligated DNA or isolated plasmids were transformed into *E. coli* strains as described by Mandel and Higa (33).

**DNA sequencing and computer analysis.** The sequence of double-stranded DNA was determined by the dideoxy-chain termination method (40) with labeling by [ $\alpha$ -<sup>35</sup>S]dATP and T7 phage DNA polymerase. To reduce formation of secondary structures, dGTP was substituted by 7-deaza-dGTP. Preparations of plasmids pAEC1180 and pAEP3050 were used as templates for primer-directed complete sequencing of both DNA strands. Synthesis of the oligodeoxynucleotide primers (17-mers) was accomplished with the Gene Assembler Plus DNA synthesizer (Pharmacia, Freiburg, Germany).

Sequence analyses were performed with the latest available versions of the GENMON programs (GBF, Braunschweig, Germany) and the GCG program package of the University of Wisconsin (10). The latter included the FASTA program (36) used for similarity searches against the GenBank (Los Alamos National Laboratories, Los Alamos, N. Mex.), EMBL/SwissProt (Heidelberg, Germany), and PIR (Georgetown University Medical Center, Washington, D.C.) sequence data bases.

**Enzymes and chemicals.** Restriction endonucleases were obtained from GIBCO BRL (Eggenstein, Germany), Pharmacia (Freiburg, Germany), or Boehringer (Mannheim, Germany). Pharmacia was also the supplier of T4 DNA ligase, Klenow fragment of DNA polymerase I, T7 DNA polymerase, and nucleotides and chemicals for oligodeoxynucleotide synthesis. Alkaline phosphatase, glycerol-3-phosphate dehydrogenase, triosephosphate isomerase, antibiotics, and some enzyme substrates (NADH and ribose-5-phosphate) were purchased from Boehringer. Reference proteins for SDS-PAGE, TK, thiamine pyrophosphate, and ribulose-5-phosphate came from Sigma Chemie (Deisenhofen, Germany).

Amersham Buchler (Braunschweig, Germany) supplied radiochemicals. Other chemicals were obtained from various sources.

**Nucleotide sequence accession numbers.** The nucleotide sequences presented in this report have been assigned accession numbers M64173 (chromosomal sequence) and M64172 (plasmid-encoded sequence) by the GenBank data base.

## RESULTS

**Sequence analysis of a subregion of the *cfx* operons.** Hybrid plasmids pAEC1180 and pAEP3050 carried subcloned regions of the chromosomal and plasmid *cfx* operons, respectively, that contained the segments between the *cfxS* and *cfxF* genes. These segments were sequenced by using the strategy of primer walking. They comprise 2,668 bp for the chromosomal sequence and 2,655 bp for the plasmid sequence, with the expected very high overall identity of 94% (Fig. 2). Relative insertions or deletions of nucleotides occur only outside potential open reading frames (ORFs) within the 150 bp downstream of *cfxS* and the 50 bp upstream of *cfxF*. Three closely linked ORFs oriented collinear with the known genes in the *cfx* operon were identified and designated *cfxX* (954 bp), *cfxY* (765 bp), and *cfxE* (726 bp). They are preceded by plausible ribosome-binding sites (Fig. 2) showing high homology to those of other *A. eutrophus* genes (22, 26, 37, 38, 49) and to the consensus site of *E. coli* (41). Their codon usage is also similar to that of other *A. eutrophus* genes (22, 26, 37, 38, 49). In agreement with an intraoperonal location of the analyzed sequence, no promoter-like structures were found. However, a potential stem-loop structure that could serve as a transcription termination signal might be present upstream of *cfxX* (Fig. 2).

The *M<sub>s</sub>* of the deduced protein gene products were calculated to be 35,059/34,954 (CfxX<sub>c</sub>/CfxX<sub>p</sub>), 27,065/27,063 (CfxY<sub>c</sub>/CfxY<sub>p</sub>), and 25,501/25,594 (CfxE<sub>c</sub>/CfxE<sub>p</sub>), with isoelectric points of pI 7.03/7.29, 5.39/5.95, and 5.54/6.17, respectively. The sequence identities of the corresponding protein pairs range between 95 and 98%. Hydrophobicity analyses revealed balanced distributions of hydrophilic and hydrophobic regions within the putative proteins (data not shown) characteristic of soluble proteins. Data base searches detected only two sequences with significant partial similarities to CfxX (see Discussion) and none similar to CfxY and CfxE; thus, no indications as to possible functions of the gene products were obtained.

**Heterologous expression of the *cfx* genes.** To identify potential products of the newly detected ORF, the putative genes from megaplasmid pHG1 and the chromosome (Fig. 1) were subcloned individually or in groups into pUC expression vectors. The resulting hybrid plasmids (Table 1) were used for heterologous expression of the genes in *E. coli* JW1. Plasmid pAEP3050 carried an insert that included the RuBisCO gene *cfxS<sub>p</sub>* together with the other three downstream genes. Two proteins corresponding in size to the S subunit (16 kDa) of RuBisCO and the predicted *cfxY<sub>p</sub>* product (27 kDa) were overproduced at different levels from this plasmid upon induction of the controlling *lac* promoter, but no *cfxX<sub>p</sub>* and *cfxE<sub>p</sub>* products were detected (Fig. 3, lane b). Nevertheless, this result is an indication for coexpression of these *cfx* genes.

Expression of *cfxX<sub>p</sub>* from pAEP9010 failed to provide evidence for the formation of CfxX<sub>p</sub> (Fig. 3, lane c). Definite overproduction of CfxY<sub>p</sub> and CfxE<sub>p</sub> (25.5 kDa) was directed by pAEP9020 and pAEP9030, respectively (Fig. 3, lanes d

c	CGCGCGCGGGTGAACCG- -GCGCAGCGCGTCATGGCGCTGCCGCGCGCGATTTCCTGACTGTGCCAATCCCACGGTTGCGCCGCGCAAGGCCCGGACCACGGGGGAGCTGCGCCTC	118
d	CGCGCACCGGTGAACCGCGCGCGCGTCCACGACGCTGCCGCGCGCGATTTCCTGACTGTGCC-ATCCACGGTTGCGCCGCGCAGGTCTGCGGGAGCTGCGCCTC	119
	--> <i>cfxX</i>	
	M S A P E T T A P L Q P P A A P A A S L P G S L A	
c	GAATCTTTCGAGTCGGATGCTCGATTGATCTCACGGAGCCTGCCATGTCGCCACCTGAAAACGACCGCACCGCTGCAGCGCCAGCCGCTCCGGCCGATCGTTGCCCGGGTCCCTGGCC	238
d	GAATTTATCGAG-----GTGAAACGATCTCACGGAGCCTGCCATGTCGCCACCGAAAACGACCGCACCGCTGCAGCCACCGCTGCCAGGCCGATCGCTGCCCTGGATCGCTGGCC	231
	M S A P E T T A P L Q P P A A P A A S L P G S L A	
	E S L A S S G I T E L L A Q L D R E L I G L K P V K A R I R D I A A L L L V D K	
c	GAGTCGTGCGCAGCTCGGGATCACCGAGCTGTCGCCCAGCTTGACCGCGAGCTGATCGGGTGAAGCCGTTGAAAGCGCGCATCCGGATATCGCCGCTTGTCTGTGGTGGACAAG	358
d	GAATCGTGGCCAGCTCGGGATCACCGAGCTGTCGCCCAGCTGCACCGTGAACCTGATCGGACTGAAGCCGTTGAAAGCGCGCATCCGGATATTCGCCGCTTGTCTGTGGTGGACAAG	351
	E S L A S S G I T E L L A Q L D R E L I G L K P V K A R I R D I A A L L L V D K	
	L R A A R G F S A G A P S L H M C F T G N P G T G K T T V A M R M A Q I L H Q L	
c	CTGCGCGCGCCGCGCGCTTCAGCGCGCGTGCAGCCTGCATATGTCTTACCCGCAATGCCCGCAACCGGCAAGACCCAGTGGCAATGGCATGGCCAGATCTGCACCGCTG	478
d	CTGCGCGCGCCGCGCGCTTCAGCGCGCGTGCAGCCTGCATATGTCTTACCCGCAATGCCCGCAACCGGCAAGACCCAGTGGCAATGGCATGGCCAGATCTGCACCGCTG	471
	L R A A R G F S A G A P S L H M C F T G N P G T G K T T V A M R M A Q I L H Q L	
	G Y V R R G H L V A V T R D D L V G Q Y I G H T A P K T K E I L K K A M G G V L	
c	GGCTACGTGCGCGCGCCACTGGTGGCGGTGACCCGCGAGCACCTGCTGCGCCAGTACATCGGCCATACCGGCCAAGACCAGAGATCTGAAGAAGGCCATGGCGGGGTTGCTC	598
d	GGCTACGTGCGCGCGCCACTGGTGGCGGTGACCCGCGAGCACCTGCTGCGCCAGTACATCGGCCATACCGGCCAAGACCAGAGATCTGAAGAAGGCCATGGCGGGGTTGCTC	591
	G Y V R R G H L V A V T R D D L V G Q Y I G H T A P K T K E I L K K A L G G V L	
	F I D E A Y Y L Y R P E N E R D Y G Q E A I E I L L Q V M E N N R D D L V V I L	
c	TTCATCGAGCGGCTACTACCTTACCGCGGAGAACGACCGGACTAGCGCCAGGAGGCATCGAGATCTGCTGACGAGTGGAGAACAACCGCGCAGCCTGGTGGTGAATCTG	718
d	TTCATCGAGCGGCTACTACCTTACCGCGGAGAACGACCGGACTAGCGCCAGGAGGCATCGAGATCTGCTGACGAGTGGAGAACAACCGCGCAGCCTGGTGGTGAATCTG	711
	F I D E A Y Y L Y R P E N E R D Y G Q E A I E I L L Q V M E N N R D D L V V I L	
	A G Y K D R M D R F F E S N P G M S S R V A H H V D F P D Y Q L D E L R Q I A D	
c	GCCGCTACAAAGGACCGATGGACCGTTCCTTCGAGTCCAACCCGGCAGTTCCTCGCGCTGCGCCACCATGTCGACTTCCCGCAGTACCAGCTCGACGAGCTGCCAGATCGCCGAC	838
d	GCCGCTACAAAGGACCGATGGACCGTTCCTTCGAGTCCAACCCGGCAGTTCCTCGCGCTGCGCCACCATGTCGACTTCCCGCAGTACCAGCTCGACGAGCTGCCAGATCGCCGAC	831
	A G Y K D R M D R F F E S N P G M S S R V A H H V D F P D Y Q L D E L R Q I A D	
	L M L S E M Q Y R F D D E S R A V F A D Y L A R R M T Q P H F A N A R S V R N A	
c	CTGATGCTGTCGGAGATGCAATAACCGCTTCGACGAGAAAGCCGGCGGTGTTTGGCGACTACTGGCCCGCGCATGACACAGCCGACTTTGCCAATGCCCGCAGCTGCCAATGCG	958
d	CTGATGCTGTCGGAGATGCAATAACCGCTTCGACGAGAAAGCCGGCGGTGTTTGGCGATTACTGGCCCGCGCATGCGCGCAGCCGACTTTGCCAATGCCCGCAGCTGCCAATGCG	951
	L M L A E M Q Y R F D D E S R A V F A D Y L A R R M A Q P H F A N A R S V R N A	
	L D R A R L R H A S R L L D D A G T V V D D H T L T T I T A S D L L A S R V F S	
c	CTGACCGCGCGCGCTGCGCCATGCCCTTCGAGTCCAACCCGGCAGTTCCTTCGAGTCCAACCCGGCAGTTCCTCGCGCTGCGCCACCATGTCGACTTCCCGCAGTACCAGCTCGACGAGCTGCCAGATCGCCGAC	1078
d	CTGACCGCGCGCGCTGCGCCATGCCCTTCGAGTCCAACCCGGCAGTTCCTTCGAGTCCAACCCGGCAGTTCCTTCGAGTCCAACCCGGCAGTTCCTCGCGCTGCGCCACCATGTCGACTTCCCGCAGTACCAGCTCGACGAGCTGCCAGATCGCCGAC	1071
	L D R A R L R H A S R L L D D A G T V V D D H T L T T I T A S D L L A S R V F S	
	--> <i>cfxY</i>	
	K A A P D A R T P A K E M Q A L I F D V D G T L A D T E T A H L Q A F N A A	
c	AAGGCCGCGCGCGACGCGACCGCGCGCAAGGAGTAAAGCAATGCAAGCCCTGATTTCGATGTCGACCGCACCCCTGCGCGATACCGAAAGCGCGCACCTGCAAGCTTCAAACCGCGCC	1198
d	AAGGCCGCGCGCGACGCGACCGCGCGCAAGGAGTAAAGCAATGCAAGCCCTGATTTCGATGTCGACCGCACCCCTGCGCGATACCGAAAGCGCGCACCTGCAAGCTTCAAACCGCGCC	1191
	K A A P A Q T P A K E M Q A L I F D V D G T L A D T E T A H L Q A F N A A	
	F A E V G L D W Y W D A P L Y T R L L K V A G G K E R L M H Y W R M V D P E E A	
c	TTCCGCGAGGTCGCGCTGGACTGGTACTGGACCGCGCGCTTACACCGCGCTGCTCAAGTTGCCGCGCGCAAGGAGCCCTGATGCTATTCAGCGCATGGTGCAGCCGCAAGAGGCC	1318
d	TTCCGCGAGGTCGCGCTGGACTGGTACTGGACCGCGCGCTTACACCGCGCTGCTCAAGTTGCCGCGCGCAAGGAGCCCTGATGCTATTCAGCGCATGGTGCAGCCGCAAGAGGCC	1311
	F A E V G L D W H W D A P L Y T R L L K V A G G K E R L M H Y W R M V D P E E A	
	R G C K V K E T I D A V H A I K T R H Y A E R V G A G G L P L R P G I A R L I D	
c	CGCGCTGCAAGGTGAAGAAACCATCGACCGCTGCACGCCATCAAGACCCGCCACTATGCCGACCGCGCTCGGGCGGGCGCGCTGCCGCTGCCCGGGCATGGCCCGCTGATCGAC	1438
d	CGCGCTGCAAGGTGAAGAAACCATCGACCGCTGCACGCCATCAAGACCCGCCACTATGCCGACCGCGCTCGGGCGGGCGCGCTGCCGCTGCCCGGGCATGGCCCGCTGATCGCA	1431
	R G C K V K E T I D A V H A I K T R H Y A E R V G A G G L P L R P G I A R L I A	
	E A G E A G L P L A I A T T T T P A N L D A L L Q A P L G A D W R R R F A A I G	
c	GAGGCCGCGCGAGCGCGCTCCCGCTGGGATGTCACACCACACCCAGCCGCGCAACCTCGACCGCGCTGCTGCAGGCGCGCTTGGCGCGCATGGCCCGCTGCTTTGCGCGCATCGGC	1558
d	GAGGCCGCGCGAGCGCGCTCCCGCTGGGATGTCACACCACACCCAGCCGCGCAACCTCGACCGCGCTGCTGCAGGCGCGCTTGGCGCGCATGGCCCGCGCTTGTCTGCAATCGGC	1551
	E A G E A G L P L A I A T T T T P A N L D A L L Q A H L G A D W R G R F A A I C	
	D A G T T A I K K P A P D V Y L A V L E R L G L E G G D C L A I E D S A N G L R	
c	GACGCGCGCACCGGCAATCAAGAGCGCGCGCGTCTGCTACCTGGCGGTGTCGCGCGCGCGCGCTGGCGCGCGGCTGGAAGCCGCTGACTGCGCTGGCGATCGAGGACTCGGCGAACCGCGCTGCGC	1678
d	GACGCGCGCACCGGCAATCAAGAGCGCGCGCGTCTGCTACCTGGCGGTGTCGCGCGCGCGCGCTGGCGCGCGGCTGGAAGCCGCTGACTGCGCTGGCGATCGAGGACTCGGCGAACCGCGCTGCGC	1671
	D A G T T A I K K P A P D V Y L A V L E R L G L E A G D C L A I E D S G N G L R	

FIG. 2. Nucleotide sequences of the chromosomal (c) and pHG1-encoded (p) 2.7-kb sections from within the two *cfx* operons of *A. entrophus* H16. They commence directly after the stop codon of *cfxS* and extend to the initiation codon of *cfxF*. The deduced amino acid sequences of the identified gene products CfxX, CfxY, and PPE are given in the one-letter code. Ribosome-binding sequences are underlined; →← indicates a region of dyad symmetry. Gaps (-) were introduced to optimize the sequence alignment.

```

A A R A A G I P T V V T P T A F S A Q D S F E G A L L V L P H L G D P G E P M P
c GCCGCGGGGGGGCGGCATTCCACCGTGGTACGCGCCACCGGTTTCAGCGCGCAGGACTTCCTCGAGGGGCGCGTGTGTCGCGCATCTTGGCGATCCGGCGAGCCCATGCCC 1798
|||||
p GCCGCGGGGGGGCGGCATTCCACCGTGGTACGCGCCACCGTTCAGCGCGCAGGACTTCCTCGAGGGGCGCGTGTGTCGCGCATCTTGGCGATCCGGCGAACCATGCCC 1791
A A R A A G I P T V V T P T T F S A Q D S F E G A L L V L P H L G D P A E P M P
--> cfxE
Q H V P G A A N R W A D L A A L R A W H H G T L I E A T M H A T E L N T G H G
c CAGCACGTGCCCGCGCGCAAAACCGCTGGGCGGACCTTGGCGCGTGGCGCGCTGGCACCCAGCGCACCTGATCGAGGCAACCTGACATGCATGCCACTGAACTCAACACCGGGCCATGG 1918
|||||
p CAGCACGTGCCCGCGCGCACACCGCTGGGCGGACCTTGGCGCGTGGCGCGCTGGCACCCAGCGCACCTGATCGAGGCGACCTGACATGCATGCCACCAGAAACCAACCGGGCCATGG 1911
Q H V P G A A H R W A D L A A L R A W H H G T L I E A T M H A T E P N T G H G
S Q R A I R L A P S I L S A D F A R L G E E V C A I E A G G A D L V H F D V M D
c CAGCCAGCGTGCATCCGCCTGGCGCCATCCATCCTGTCGGCCGATTTCGCGCGCTGGGCGAAGAGGTGTGCGCGATCGAGGCCGGCGGCGGACCTGGTGCACCTTCGATGTGATGGA 2038
|||||
p CAGCCAGCGCGCATCCGCCTGGCGCCATCCATCCTGTCGGCCGATTTCGCGCGCTGGGCGAAGAGGTGTGCGCGATCGAGGCCGGCGGCGGACCTGGTGCACCTTCGATGTGATGGA 2031
S Q R A I R L A P S I L S A D F A R L G E E V C A I E A G G A D L V H F D V M D
N H Y V P N L T I G P L V C E A I R P L V S I P I D V H L M V E P V D A L I P L
c CAACCACATATGTGCCAACCTGACCATTTGGCCCGCTGGTGTGCGAGGCGATCCGGCCGCTGGTCTCCATCCCCATCGACGTGCATCTGATGGTGGAAACCGGTCGATCGCGCTGATCCCGCT 2158
|||||
p CAACCACATATGTGCCAACCTGACCATCGGCCCGCTGGTGTGCGAGGCAATCCGGCCGCTGGTTCGATCCCATCGACGTGCACCTGATGGTGGAGCCGGTCGATCGCGCTGATCCCGCT 2151
N H Y V P N L T I G P L V C E A I R P L V S I P I D V H L M V E P V D A L I P M
F A K A G A N I I S F H P E A S R H V D R T I G L I R D H G C K A G L V L N P A
c GTTCGCCAAGCGGGCGCCAAACATCATCAGCTTCCATCCGAGGCGAGCGCCCATGTGACCGCACCATCGGCCCTGATCCGCGACCAACCGCTGCAAGGCGAGCCCTGGTCTCAACCCGGC 2278
|||||
p GTTCGCCAAGCGGGCGCCAAACCTCATCAGCTTCCATCCGGAAGCGAGCGCCCATGTGAGCCGACCAATCGGCCCTGATCCGCGACCAACCGCTGCAAGGCGAGCCCTGGTCTGAAACCCGGC 2271
F A K A G A N L I S F H P E A S R H V D R T I G L I R D H G C K A G L V L N P A
T P L G W L D H T L D Q L D L V L L M S V N P G F G G Q A F I P G V L D K V R Q
c CACGCGCTGGGCTGGCTGGACCATACGCTGGACAGCTCGACCTGGTGTCTGCTGATGAGCGTGAACCCGGGCTTGGCGCGCCAGGCCTTCATCCGGCGCTGCTGGACAAGGTGCGCCA 2398
|||||
p CACGCGCTGAGCTGGCTGGACCAACGCTGGACAAGCTCGACCTGGTGTCTGCTGATGAGCGTGAACCCGGGCTTGGCGCGCCAGGCCTTCATCCGGCGCTGCTGGACAAGGTGCGCCA 2391
T P L S W L D H T L D K L D L V L L M S V N P G F G G Q A F I P G V L D K V R Q
A R A R I D R Q V D A G G R P V W L E I D G G V K A D N I A A I A R A G A D T F
c GGCACGGCGCGCATCGACCGGAGGTGGACCGCGCGGGCGCGCGGCTGGCTGGAGATCGACGGCGCGCTCAAGGCCGACAACATTGCCCGCATCGCGCGAGCGGGCGCGACACCTT 2518
|||||
p GGCACGGCGCGCATCGACCGGAAAGTGGCCGCGCGGGCGCGCGGCTGGCTGGAGATCGACGGCGCGCTCAAGGCCGACAACATCACTGAGATCGCCCGTCCGGGGCGCGACACCTT 2511
A R A R I D R Q V A A G G R P V W L E I D G G V K A D N I T E I A R A G A D T F
V A G S A V F T G G A P D A D G G Y S S I L Y R L R E A A T V T
c CGTTCGGCGAGCGCGTGTTCGGCGCGCCGATGCGCGACCGCGGCTACTCGAGCATCCTTTACCGCTTGGCGGAGCGCGCCACCGTCAAGCGCGCGCCCGCCAGCCACCAACAAG 2638
|||||
p CGTTCGGCGAGCGCGTGTTCGGCGCGCCGATGCGCGACCGCGGCTACTCGGGCATCCTGCACCGCTTGGCGGAGCGCGCCACCATCAAGTAGCCCGC-----AC-GGCACCACACAAG 2631
V A G S A V F T G A P D A D G G Y R G I L H R L R E A A T I T
c AATGCATAGCCAATCTATAGGAGACCTGTC 2668
|||
p AATACATAGCCAATCTATAGGAGACCTGTC 2655

```

FIG. 2—Continued.

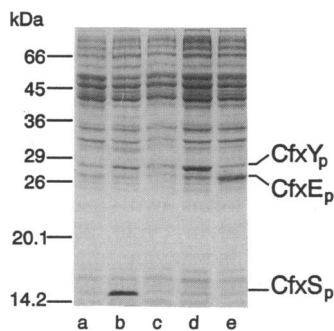


FIG. 3. Heterologous expression of *cfxSXYE<sub>p</sub>* from *A. eutrophus* H16 in *E. coli* JW1 harboring various hybrid plasmids, analyzed by SDS-PAGE of cell extracts. The cells were grown in LB medium plus ampicillin, and the *lac* promoter on the plasmids was induced by IPTG. Lanes: a, *E. coli*(pUC18) as a control; b, *E. coli*(pAEP3050); c, *E. coli*(pAEP9010); d, *E. coli*(pAEP9020); e, *E. coli*(pAEP9030). The overproduced gene products CfxY<sub>p</sub>, CfxE<sub>p</sub>, and CfxS<sub>p</sub> are indicated. Sizes of reference proteins are shown on the left.

and e), and overproduction of both proteins was directed by pAEP9230 (not shown). Overexpression of the chromosomal *cfxE<sub>c</sub>* gene encoded on the cloned insert of pAEC3010 was also achieved. Plasmids with inserts oriented in opposite direction to the *lac* promoter did not yield any overproduced proteins (not shown). Thus, heterologous expression of the genes depended on the vector promoter and, for unknown reasons, was much lower for *cfxX<sub>p</sub>* than for the other two genes.

Detectable expression of *cfxX<sub>p</sub>* required recloning of the gene into vector pT7-7 downstream the T7 RNA polymerase-dependent promoter. Labeling of proteins with L-[<sup>35</sup>S]methionine upon induction of the T7 promoter in pAEP9012 enabled the identification of a product exhibiting the predicted size of about 35 kDa (Fig. 4). The expression level of *cfxX<sub>p</sub>* (and probably of *cfxE<sub>c</sub>* as well) seemed to be much lower than that of the neighboring genes. Two-dimensional PAGE with cell extracts of the respective *E. coli* transformants confirmed the findings for *cfxY* and *cfxE* expression as well as the calculated isoelectric points of the CfxY and CfxE products (data not shown).

**PPE activities in *E. coli* transformants and in *A. eutrophus*.** Cell extracts of various *E. coli* transformants examined previously for overproduction of proteins were assayed for

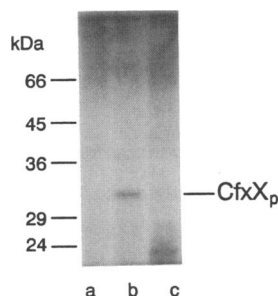


FIG. 4. Heterologous expression of *cfxX<sub>p</sub>* by means of the phage T7 RNA polymerase promoter system in *E. coli* K38(pGP1-2) harboring the additional hybrid plasmid pAEP9012, analyzed by autoradiography after SDS-PAGE of whole-cell lysates. Plasmid-encoded proteins were labeled with L-[<sup>35</sup>S]methionine under inducing or noninducing conditions. Lanes: a, pAEP9012 (not induced); b, pAEP9012 (induced); c, vector pT7-7 (induced) as a control. The *CfxX<sub>p</sub>* product is marked; sizes of reference proteins are shown on the left.

enhanced activities of those Cfx enzymes not previously known to be encoded within the *cfx* operon. These enzymes were fructose-1,6-bisphosphate aldolase, phosphoglycerate kinase, triosephosphate isomerase, pentose-5-phosphate isomerase, and PPE. Only the latter showed significantly enhanced activities in certain clones. PPE activities were increased between 7- and 19-fold above the background level in *E. coli* and occurred exclusively in strains harboring the *cfxE* genes in collinear orientation to the *lac* promoter of the expression vector (Table 2). The highest PPE activity was found in *E. coli*(pAEP9030), in which the hybrid plasmid contained only *cfxE<sub>p</sub>*, strongly suggesting that the gene encodes PPE, although a possible regulatory function of the gene product affecting PPE activity in *E. coli* cannot be discounted.

If *cfxE* is the PPE structural gene of the *cfx* operon, its synthesis should follow the same regulatory pattern as that observed for RuBisCO and PRK (13, 29). Indeed, this was the case, assuming that the activities reflect the synthesis rates as found for the two key Cfx enzymes. Autotrophic cells exhibited clearly derepressed activity levels, and CO<sub>2</sub> limitation during lithoautotrophic growth led to maximal derepression of the enzyme (Table 3). Partial derepression occurred under heterotrophic conditions with fructose as the carbon and energy source. The PPE level in pyruvate-grown cells which have completely repressed *cfx* operons (20) probably represents the basal activity of the enzyme in this organism.

TABLE 2. Activities of PPE in cell extracts of various transformants of *E. coli* JW1

Transformant <sup>a</sup>	Sp act of PPE (U/mg of protein)
<i>E. coli</i> (pUC18) .....	0.44
<i>E. coli</i> (pAEP3050).....	3.02
<i>E. coli</i> (pAEP3051).....	0.31
<i>E. coli</i> (pAEP9030).....	8.20
<i>E. coli</i> (pAEP9031).....	0.42
<i>E. coli</i> (pAEC3010).....	3.20
<i>E. coli</i> (pAEC3011).....	0.47

<sup>a</sup> Grown in LB medium and induction of *lacPO* by IPTG.

TABLE 3. Activities of PPE in cell extracts of *A. eutrophus* H16 grown on various substrates

Substrate	Sp act of PPE (U/mg of protein)
H <sub>2</sub> /CO <sub>2</sub> lim. <sup>a</sup> .....	13.24
H <sub>2</sub> /CO <sub>2</sub> .....	9.41
Formate <sup>b</sup> .....	5.75
Fructose.....	3.49
Pyruvate.....	2.56

<sup>a</sup> H<sub>2</sub>/CO<sub>2</sub> lim., lithoautotrophic growth under limiting CO<sub>2</sub> supply (1 vol%).

<sup>b</sup> Organoautotrophic growth on formate.

## DISCUSSION

In this work, we obtained evidence for the existence of three additional contiguous gene loci, *cfxXYE*, within the duplicated *cfx* operon of *A. eutrophus* H16. The genes are closely linked to and in the same orientation as are the other genes of the operon. Heterologous coexpression of *cfxSXYE* in *E. coli*, being dependent on the *lac* promoter of the vector plasmid, confirmed their status as constituent operon genes. Except for the 5'-terminal *cfxL* gene of the operon (20), all downstream genes require a foreign promoter for expression in *E. coli* (24, 47; unpublished results). The presently available data suggest that the promoter upstream of *cfxL* is the only functional promoter of the *cfx* operon (48).

Whereas *cfxX* and *cfxY* encode protein products of unknown functions, *cfxE*, like the remaining identified genes of the *cfx* operon, codes for a Calvin cycle enzyme. Two lines of evidence support this conclusion: (i) up to about a 20-fold increase of PPE activity in *E. coli* after expression of the *cfxE* gene and (ii) a pronounced increase (maximally about 5-fold) of PPE activity in *A. eutrophus* upon derepression of induction of the *cfx* operons under autotrophic growth conditions that correlates with the activity patterns of the other enzymes encoded in the operon (9). Definitive proof of the identity of PPE as the product of *cfxE* must come from N-terminal amino acid sequencing of the purified enzyme. In general, little information about the properties of PPE is available. The enzymes from bovine liver and human erythrocytes were described as homodimers of 23-kDa subunits (23, 43, 50), and the yeast enzyme exhibiting a native molecular mass of 46 kDa (46) may also have this quaternary structure. The deduced subunit mass of 25.5 kDa for the bacterial PPE from *A. eutrophus* is rather close to that of the eukaryotic enzyme. No PPE sequences from any source have been reported so far. Surprisingly, PPE does not have significant similarity to L-ribulose-5-phosphate 4-epimerase (EC 5.1.3.4; *araD* product) from *E. coli* (30) and *Salmonella typhimurium* (31), an enzyme which is involved in L-arabinose degradation.

Since *A. eutrophus* forms two special PPE isoenzymes that operate in the Calvin cycle, it has to be able to synthesize a third PPE isoenzyme functioning in heterotrophic carbon metabolism. When the Calvin cycle PPEs are not available, this isoenzyme is an essential catalyst in the organism's ribose biosynthesis (7). It is postulated to be the product of a separate PPE gene, *rpe*. This conclusion is based on the fact that (i) the *cfx* operons are completely repressed during growth on various organic acids and (ii) mutants with defective *cfx* operons are unaffected in heterotrophic growth (48). The same reasoning applies to the FBP, TK, and GAP isoenzymes. A chromosomally located *gap* gene has been detected (47).

Among the sequences listed in data bases, only the poten-

CfxX <sub>C</sub>	<b>MSAPETTAPLQPPAA</b> ----- <b>PAASLPGSLAESLASSGITE</b> - <b>LLAQLDRE</b> --- <b>LIGLKPVKARIRDIAALLLVDKL</b>	66
ORFC	<b>MLDVATSAPSALPA</b> ----- <b>EAAEGRDLGALFTESEVPE</b> - <b>FLAELDEG</b> --- <b>LIGLKPVKRRIREIAAHLVIGRA</b>	66
SpoVJ	<b>MLERAVTYKNGGQINIILNGQKQVLTNAEAEAEYQAALQKNEAKHGILKEIEKEMSALVGMEEEMKRNIKEIYAWIFVNQK</b>	80
CfxX <sub>C</sub>	<b>RAARGFSAGAPSLHMCFTGNP</b> <b>GTGKTTVAMRMAQILHQLGYVRRGHVAVTRDDL</b> <b>VGQYIGHTAPKTKEILKKAMGGVLF</b>	146
ORFC	<b>REKLGLTSGAPTLHMAFTGNP</b> <b>GTGKTTVALKMAQILHRLGYVRRGHVSVTRDDL</b> <b>VGQYIGHTAPKTKEILKKAMGGVLF</b>	146
SpoVJ	<b>RAEQGLKVGKQALHMMFKGNP</b> <b>GTGKTTVARLIGKLFEMNVLSKGHLEAERADLVGEYIGHTAQKTRDLIKKSLGGILF</b>	160
CfxX <sub>C</sub>	<b>IDEAYLYRPNENERDYGQEAIEILLQVMENNRDDL</b> <b>VVILAGYKDRMDRFFESNPGMSRVAHVDFPDYQLDEL</b> <b>RQIADL</b>	226
ORFC	<b>IDEAYLYRPNENERDYGQEAIEILLQVMENQRDDL</b> <b>VVILAGYKDRMDRFFESNPGFRSRIAAHIDFPDYEDAELVEIAKT</b>	226
SpoVJ	<b>IDEAYSLAR</b> - <b>GGEKDFGKEAIDTLVKHME</b> <b>DKQHEFILILAGYSREMDHFLSLNPLQSRFPISIDFPDYSVTQLMETIAKR</b>	240
CfxX <sub>C</sub>	<b>MLSEMQRFDDESRAVFADYL</b> -- <b>ARRMTQP</b> - <b>HFANARSVRNALDRARLRHASRLLDDAGTVVDDHTLTITASDLLASRV</b>	303
ORFC	<b>MAADADYTFSPAEVAIEEYV</b> -- <b>AKRRLQP</b> - <b>NFANARSIRNALDRMRLRQSLRRLFESGG</b> - <b>LADRAALSTISEGDVRSRV</b>	302
SpoVJ	<b>MIDEREYQLSQEAEWKLKDYLMTVKSTTSP</b> <b>IKFSNGRFVRN</b> <b>VIKESIRAQAMRLLMGDQYL</b> -- <b>KSDLMTIKSQDLSIKEE</b>	318
CfxX <sub>C</sub>	<b>FSK</b> -- <b>AAPDARTPAKE</b>	317
ORFC	<b>FAGGIDAPDYK</b> - <b>PQTE</b>	317
SpoVJ	<b>ASGSA</b>	323

FIG. 5. Sequence comparison by alignment of the deduced amino acid sequences of CfxX<sub>C</sub> from *A. eutrophus* H16, ORF C from *X. flavus* H4-14, and SpoVJ from *B. subtilis*. The marked region is a potential nucleotide-binding site. Identical residues relative to CfxX<sub>C</sub> are in bold. Gaps (-) were introduced to optimize the alignment.

tial product of ORF C from another chemoautotroph, *Xanthobacter flavus* H4-14 (34), has high similarity to CfxX, with 65% of amino acid residues identical (Fig. 5). Although the resemblance extends throughout the proteins, it is particularly strong in their central parts. A sequence motif conforming to the consensus sequence (GNPGTGKTT) for a nucleotide-binding domain (17) was identified (Fig. 5). It is also present in the *spoVJ* product of *Bacillus subtilis* (11), which shows a significant overall similarity (39% residue identity) to CfxX (Fig. 5) and whose precise function in sporulation is still unclear. Like the *cfxX* gene in *A. eutrophus*, ORF C in *X. flavus* is located immediately downstream of the RuBisCO genes *cfxLS* within the *cfx* gene cluster of the organism, suggesting that *cfxX* and ORF C are homologous genes with the same, yet unknown function.

A gene homologous to *cfxX* may also be encoded in the 3'-flanking region of the *rbcLS* (= *cfxLS*) genes of the form I *cfx* gene cluster of the purple nonsulfur bacterium *Rhodobacter sphaeroides* (15). We detected 37% residue identity with the N-terminal portion of a potential gene product from an incomplete ORF starting 163 bp downstream of *rbcS* (data not shown). Even more interesting from an evolutionary point of view is the finding that the sequence from nucleotide positions 209 through 279 downstream of the *rbcLS* operon of the red alga *Antithamnion* sp. (27) can be translated (assuming a frameshift at position 235) into an amino acid sequence of 24 residues that has 83% identity with a corresponding region in the N-terminal part of CfxX. The possible partial conservation of a *cfxX*-like sequence in this eukaryote would support the fact that the RuBisCO sequences from chromophyte and rhodophyte plastids are more homologous to those from *A. eutrophus* and purple nonsulfur bacteria (form I enzyme; L<sub>8</sub>S<sub>8</sub>) than to those from chlorophyte plastids (1, 4, 21).

No function can yet be assigned to CfxX and CfxY. Although the *cfxX* gene is preceded by a plausible ribosome-binding site, its expression in *E. coli* was extremely low, a fact deserving attention in further studies on the function of the gene. The upstream *cfxS* and the downstream *cfxYE* genes were expressed much better, both individually and in combination. If *cfxX* expression in *A. eutrophus* is also low, a regulatory function of CfxX in autotrophic CO<sub>2</sub> fixation is

conceivable. The presence of a nucleotide-binding motif could indicate that the CfxX activity is energy requiring or regulated by a nucleotide. Site-directed mutagenesis of *cfxX* and *cfxY* is expected to provide more information on the metabolic roles of the respective gene products.

#### ACKNOWLEDGMENTS

This work was supported by a grant from the Deutsche Forschungsgemeinschaft.

We are grateful to Stanley Tabor for the generous donation of strain *E. coli* K38(pGP1-2) and vector plasmid pT7-7.

#### REFERENCES

- Assali, N.-E., W. F. Martin, C. C. Sommerville, and S. Loiseaux-de Goer. 1991. Evolution of the rubisco operon from prokaryotes to algae: structure and analysis of the *rbcS* gene of the brown alga *Pylaiella littoralis*. *Plant Mol. Biol.* 17:853-863.
- Ausubel, F., R. Brent, R. E. Kingston, D. D. Moore, J. G. Seidman, J. A. Smith, and K. Struhl (ed.). 1987. *Current protocols in molecular biology*. Wiley, New York.
- Birnboim, H. C., and J. Doly. 1979. A rapid alkaline extraction procedure for screening recombinant plasmid DNA. *Nucleic Acids Res.* 7:1513-1523.
- Boczar, B. A., T. P. Delaney, and R. A. Cattolico. 1989. Gene for the ribulose-1,5-bisphosphate carboxylase small subunit protein of the marine chromophyte *Olisthodiscus luteus* is similar to that of a chemoautotrophic bacterium. *Proc. Natl. Acad. Sci. USA* 86:4996-4999.
- Blum, H., H. Beier, and H. J. Gross. 1987. Improved silver staining of plant proteins, RNA and DNA in polyacrylamide gels. *Electrophoresis* 8:93-99.
- Bowien, B., B. Friedrich, and C. G. Friedrich. 1984. Involvement of megaplasmids in heterotrophic derepression of the carbon-dioxide assimilating enzyme system in *Alcaligenes* spp. *Arch. Microbiol.* 139:305-310.
- Bowien, B., and H. G. Schlegel. 1972. Der Biosyntheseweg der RNS-Ribose in *Hydrogenomonas eutropha* Stamm H16 und *Pseudomonas facilis*. *Arch. Mikrobiol.* 85:95-112.
- Bowien, B., and H. G. Schlegel. 1981. Physiology and biochemistry of aerobic hydrogen-oxidizing bacteria. *Annu. Rev. Microbiol.* 35:405-452.
- Bowien, B., U. Windhövel, J.-G. Yoo, R. Bednarski, and B. Kusian. 1990. Genetics of CO<sub>2</sub> fixation in *Alcaligenes eutrophus*. *FEMS Microbiol. Rev.* 87:445-450.
- Devereux, J., P. Haerberli, and O. Smithies. 1984. A comprehen-

- sive set of sequence analysis programs for the VAX. *Nucleic Acids Res.* **12**:387–395.
11. Foulger, D., and J. Errington. 1991. Sequential activation of dual promoters by different sigma factors maintains *spoJV* expression during successive developmental stages of *Bacillus subtilis*. *Mol. Microbiol.* **5**:1363–1373.
  12. Friedrich, B. 1990. The plasmid-encoded hydrogenase gene cluster in *Alcaligenes eutrophus*. *FEMS Microbiol. Rev.* **87**: 425–430.
  13. Friedrich, B., C. Hogrefe, and H. G. Schlegel. 1981. Naturally occurring genetic transfer of hydrogen oxidizing ability between strains of *Alcaligenes eutrophus*. *J. Bacteriol.* **147**:198–205.
  14. Friedrich, C. G. 1982. Derepression of hydrogenase during limitation of electron donors and derepression of ribulose biphosphate carboxylase during carbon limitation of *Alcaligenes eutrophus*. *J. Bacteriol.* **149**:203–210.
  15. Gibson, J. L., D. L. Falcone, and F. R. Tabita. 1991. Nucleotide sequence, transcriptional analysis, and expression of genes encoded within the form I CO<sub>2</sub> fixation operon of *Rhodobacter sphaeroides*. *J. Biol. Chem.* **266**:14646–14653.
  16. Hanahan, D. 1983. Studies on the transformation of *Escherichia coli* with plasmids. *J. Mol. Biol.* **166**:557–580.
  17. Higgins, C. F., I. D. Hiles, G. P. C. Salmond, D. R. Gill, J. A. Downie, I. J. Evans, I. B. Holland, L. Gray, S. D. Buckel, A. W. Bell, and M. A. Hermodson. 1986. A family of related ATP-binding subunits coupled to many distinct processes in bacteria. *Nature (London)* **323**:448–450.
  18. Holmes, D. S., and M. Quigley. 1981. A rapid boiling method for the preparation of bacterial plasmids. *Anal. Biochem.* **114**:193–197.
  19. Horiuchi, K., and N. D. Zinder. 1967. Azure mutants: a type of host-dependent mutant of the bacteriophage  $\phi 2$ . *Science* **156**: 1618–1623.
  20. Husemann, M., R. Klintworth, V. Büttcher, J. Salnikow, C. Weissenborn, and B. Bowien. 1988. Chromosomally and plasmid-encoded gene clusters for CO<sub>2</sub> fixation (*cfx* genes) in *Alcaligenes eutrophus*. *Mol. Gen. Genet.* **214**:112–120.
  21. Hwang, S.-R., and F. R. Tabita. 1991. Cotranscription, deduced primary structure, and expression of the chloroplast-encoded *rbcl* and *rbcS* genes of the marine diatom *Cylindrotheca* sp. strain N1. *J. Biol. Chem.* **266**:6271–6279.
  22. Jendrossek, D., A. Steinbüchel, and H. G. Schlegel. 1988. Alcohol dehydrogenase gene from *Alcaligenes eutrophus*: subcloning, heterologous expression in *Escherichia coli*, sequencing, and location of Tn5 insertions. *J. Bacteriol.* **170**:5248–5256.
  23. Karmali, A., A. F. Drake, and N. Spencer. 1983. Purification, properties and assay of D-ribulose 5-phosphate 3-epimerase from human erythrocytes. *Biochem. J.* **211**:617–623.
  24. Klintworth, R., M. Husemann, C. Weissenborn, and B. Bowien. 1988. Expression of the plasmid-encoded phosphoribulokinase gene from *Alcaligenes eutrophus*. *FEMS Microbiol. Lett.* **49**: 1–6.
  25. Kolmar, H., K. Friedrich, J. Pschorr, and H.-J. Fritz. 1990. Hybrids of circular DNA single strands as intermediates in DNA cloning, nucleotide sequence analysis, and directed mutagenesis. *Technique* **2**:237–245.
  26. Kossmann, J., R. Klintworth, and B. Bowien. 1989. Sequence analysis of the chromosomal and plasmid genes encoding phospho-ribulokinase from *Alcaligenes eutrophus*. *Gene* **85**:247–252.
  27. Kostrzewa, M., K. Valentin, U. Maid, R. Radetzky, and K. Zetsche. 1990. Structure of the rubisco operon from the multicellular red alga *Antithamnion* spec. *Curr. Genet.* **18**:465–469.
  28. Laemmli, U. K. 1970. Cleavage of structural proteins during the assembly of the head of bacteriophage T4. *Nature (London)* **227**:680–685.
  29. Leadbeater, L., K. Siebert, P. Schobert, and B. Bowien. 1982. Relationship between activities and protein levels of ribulose-biphosphate carboxylase and phosphoribulokinase in *Alcaligenes eutrophus*. *FEMS Microbiol. Lett.* **14**:263–266.
  30. Lee, N., W. Gielow, R. Martin, E. Hamilton, and A. Fowler. 1986. The organization of the *araBAD* operon of *Escherichia coli*. *Gene* **47**:231–244.
  31. Lin, H. C., S. P. Lei, G. Studnicka, and G. Wilcox. 1985. The *araBAD* operon of *Salmonella typhimurium* LT2. III. Nucleotide sequence of *araD* and its flanking regions, and primary structure of its product. *Gene* **34**:129–134.
  32. Lowry, O. H., N. J. Rosebrough, A. L. Farr, and R. J. Randall. 1951. Protein measurement with the Folin phenol reagent. *J. Biol. Chem.* **193**:265–275.
  33. Mandel, M., and A. Higa. 1970. Calcium dependent bacteriophage DNA infection. *J. Mol. Biol.* **53**:159–162.
  34. Meijer, W. G., A. C. Arnberg, H. G. Enequist, P. Terpstra, M. E. Lidstrom, and L. Dijkhuizen. 1991. Identification and organization of carbon dioxide fixation genes in *Xanthobacter flavus* H4-14. *Mol. Gen. Genet.* **225**:320–330.
  35. O'Farrell, P. H. 1975. High resolution two-dimensional electrophoresis of proteins. *J. Biol. Chem.* **250**:4007–4021.
  36. Pearson, W. R., and D. J. Lipman. 1988. Improved tools for biological sequence analysis. *Proc. Natl. Acad. Sci. USA* **85**: 2444–2448.
  37. Peoples, O. P., and A. J. Sinskey. 1989. Poly- $\beta$ -hydroxybutyrate (PHB) biosynthesis in *Alcaligenes eutrophus*. Characterization of the genes encoding  $\beta$ -ketothiolase and acetoacetyl-CoA reductase. *J. Biol. Chem.* **264**:15293–15297.
  38. Peoples, O. P., and A. J. Sinskey. 1989. Poly- $\beta$ -hydroxybutyrate (PHB) biosynthesis in *Alcaligenes eutrophus*. Identification and characterization of the PHB polymerase gene (*phbC*). *J. Biol. Chem.* **264**:15298–15303.
  39. Sambrook, J., E. F. Fritsch, and T. Maniatis. 1989. *Molecular cloning: a laboratory manual*, 2nd ed. Cold Spring Harbor Laboratory, Cold Spring Harbor, N.Y.
  40. Sanger, F., S. Nicklen, and A. R. Coulson. 1977. DNA sequencing with chain-terminating inhibitors. *Proc. Natl. Acad. Sci. USA* **74**:5463–5467.
  41. Shine, J., and L. Dalgarno. 1975. Determinants of cistron specificity in bacterial ribosomes. *Nature (London)* **254**:34–38.
  42. Tabor, S., and C. A. Richardson. 1985. T7 RNA polymerase/promoter system for controlled exclusive expression of specific genes. *Proc. Natl. Acad. Sci. USA* **82**:1074–1078.
  43. Terada, T., H. Mukae, K. Ohashi, S. Hosomi, T. Mizoguchi, and K. Uehara. 1985. Characterization of an enzyme which catalyzes isomerization and epimerization of D-erythrose 4-phosphate. *Eur. J. Biochem.* **148**:345–351.
  44. Vieira, J., and J. Messing. 1982. The pUC plasmids, an M13mp7-derived system for insertion mutagenesis and sequencing with synthetic universal primers. *Gene* **19**:259–268.
  45. Vogelstein, B., and D. Gillespie. 1979. Preparative and analytical purification of DNA from agarose. *Proc. Natl. Acad. Sci. USA* **76**:615–619.
  46. Williamson, W. T., and W. A. Wood. 1966. D-Ribulose 5-phosphate 3-epimerase. *Methods Enzymol.* **9**:605–608.
  47. Windhövel, U., and B. Bowien. 1990. Cloning and expression of chromosomally and plasmid-encoded glyceraldehyde-3-phosphate dehydrogenase genes from the chemoautotroph *Alcaligenes eutrophus*. *FEMS Microbiol. Lett.* **66**:29–34.
  48. Windhövel, U., and B. Bowien. 1990. On the operon structure of the *cfx* gene clusters in *Alcaligenes eutrophus*. *Arch. Microbiol.* **154**:85–91.
  49. Windhövel, U., and B. Bowien. 1991. Identification of *cfxR*, an activator gene of autotrophic CO<sub>2</sub> fixation in *Alcaligenes eutrophus*. *Mol. Microbiol.* **5**:2695–2705.
  50. Wood, T. 1979. Purification and properties of D-ribulose-5-phosphate 3-epimerase from calf liver. *Biochim. Biophys. Acta* **570**:352–362.
  51. Yanisch-Perron, C., J. Vieira, and J. Messing. 1985. Improved M13 phage cloning vectors and host strains: nucleotide sequences of M13mp18 and pUC19 vectors. *Gene* **33**:103–109.