JOURNAL OF BACTERIOLOGY, Sept. 1989, p. 4914–4922
0021-9193/89/094914-09$02.00/0

Vol. 171, No. 9

# Genes Downstream from *pucB* and *pucA* Are Essential for Formation of the B800-850 Complex of *Rhodobacter capsulatus*

HANS VOLKER TICHY,[1] BÉATRICE OBERLÉ,[1] HARALD STIEHLE,[1] EMILE SCHILTZ,[2] AND GERHART DREWS[1]*

*Institut für Biologie II, Mikrobiologie, Schänzlestrasse 1,[1] and Institut für Organische Chemie und Biochemie,[2] Universität Freiburg, D-7800 Freiburg, Federal Republic of Germany*

Received 21 February 1989/Accepted 13 June 1989

**The formation of the light-harvesting complex B800-850 (LH-II) of *Rhodobacter capsulatus* requires, in addition to the synthesis of the polypeptides α and β (the gene products of *pucA* and *pucB*), the synthesis of bacteriochlorophyll and carotenoids and the expression of at least one gene localized downstream from the *pucBA* operon. This was concluded from the observation that a Tn*5* insertion downstream from *pucBA* inhibited the formation of the LH-II complex and the formation of the *pucBA* mRNA. The Tn*5* insertion point was mapped and found to be over 500 base pairs (bp) downstream from the end of the *pucA* gene, suggesting the presence of additional *puc* genes. A region of about 3,000 bp including the *pucB* and *pucA* genes and DNA downstream from *pucA* was sequenced and found to contain three open reading frames (ORFs C, D, and E). The polypeptide deduced from the first ORF (C) contains 403 amino acids with strongly hydrophobic stretches and one large and three small hydrophilic domains carrying many charged residues. The other two ORFs contain 113 (D) and 118 (E) codons. The amino acid sequences of the N terminus and two tryptic peptides of an alkaline-soluble $M_r$-14,000 subunit of the isolated LH-II complex were identical with the deduced amino acid sequence of ORF E.**

Purple nonsulfur bacteria such as *Rhodobacter capsulatus* are able to generate an electrochemical proton gradient across the membrane by photochemical reactions under anoxic conditions or by respiration in the presence of oxygen (7, 10). Lowering of oxygen tension in the medium induces the formation of three pigment-protein complexes of the photosynthetic apparatus. These are the photochemical reaction center and the light-harvesting antenna complexes B870 (LH-I) and B800-850 (LH-II). The light-harvesting complexes absorb photons of the visible and near-infrared spectrum and transfer the created excitation energy to the reaction center where charge separation occurs. The B800-850 light-harvesting complex contains two bacteriochlorophyll-binding polypeptides, α and β, with molecular weights of 7,322 and 4,579, respectively (35, 36), and a third polypeptide, γ ($M_r$ 14,000), which does not bind bacteriochlorophyll (5, 6). The genes *pucA* and *pucB* encoding the pigment-binding proteins have been cloned and sequenced (41).

The synthesis of the LH-II complex is regulated by oxygen partial pressure and/or light intensity (29, 30). The level of the *pucBA* mRNA increases after a shift from aerobic to semiaerobic conditions (17, 43, 44), followed by an increase in the synthesis of LH-II complex. In cells grown under high light intensity, the *puc* mRNA level is higher than in low-light-grown cells (47). Thus, a posttranscriptional mechanism of *puc* regulation has been postulated (47).

B800-850-negative mutants have been constructed by transposon mutagenesis (11, 16, 40, 46) and the introduction of deletions (42). Some of these mutants are defective in the synthesis of carotenoids, and this defect leads to loss of the B800-850 spectral complex.

In this communication, we will show that the transposon Tn*5*, which blocks formation of LH-II in the mutant strain NK3 of *R. capsulatus*, is inserted downstream from the *pucB* and *pucA* genes and that the genetic information in this region is essential for formation of the B800-850 complex. We have determined the DNA sequence of this region and will describe here the properties of open reading frames (ORFs) and deduced polypeptides of this new *puc* locus. In light of results from complementation experiments and protein sequence analyses, the functions of the new *puc* genes will be discussed.

## MATERIALS AND METHODS

**Bacterial strains and growth conditions.** Bacterial strains used are shown in Table 1. *Escherichia coli* strains were grown at 37°C in 2× TY medium (16 g of tryptone [Difco Laboratories, Detroit, Mich.], 10 g of yeast extract, 5 g of NaCl per liter) or H medium (10 g of tryptone [Difco], 8 g of NaCl per liter). *Rhodobacter* strains were grown at 32°C in a malate salt medium (RÄ) or RÄ supplemented with 0.5% yeast extract (Difco) (RÄH) (3). Antibiotics were used at the following concentrations (μg/ml): ampicillin, 100; kanamycin, 20; tetracycline, 20 (*E. coli*) or 2 (*R. capsulatus*). If needed, 5-bromo-4-chloro-3-indolyl-β-galactoside (0.002% final concentration) and isopropyl-β-D-thiogalactoside (0.1 mM final concentration) were added to solid media. Formation of the photosynthetic apparatus of *R. capsulatus* was induced by lowering the oxygen partial pressure as described previously (25).

**Plasmids and bacteriophages.** Plasmids used are shown in Table 1. The phages M13mp18 and M13mp19 (38) were used as vectors for DNA-sequencing purposes. Phage R408 (26) was obtained from Promega Biotec, Madison, Wis., and used as helper phage to prepare single-stranded DNA from pGEM-7Zf(+) derivatives.

**Plasmid transfer and selection of transconjugants.** Single colonies of the recipient were streaked onto RÄH plates and incubated overnight at 32°C. A 50-μl portion of an overnight culture of the plasmid-containing *E. coli* S17-1 derivative

* Corresponding author.

TABLE 1. Bacterial strains and plasmids

| Strain or plasmid | Relevant characteristic(s) | Source or reference |
|---|---|---|
| Bacterial strains | | |
| *R. capsulatus* | | |
| 37b4 | Wild type | DSM938[a] |
| Y5 | BChl$^+$ RC$^-$ B870$^-$ B800-850$^+$ Crt$^+$ | 5 |
| NK3 | BChl$^+$ RC$^+$ B870$^+$ B800-850$^-$ Crt$^+$ | 16 |
| *E. coli* | | |
| JM109 | *recA1* Δ(*lac pro*) *hsdR17 traD36 proAB*$^+$ *lacI*$^q$ *lacZ*ΔM15 F′ | Promega |
| S17-1 | RP4-2 (Tc::Mu)(Km::Tn*7*) integrated in the chromosome | 33 |
| Plasmids | | |
| pGEM-3 | Ap$^r$ | Promega |
| pGEM-7Zf(+) | Ap$^r$ f1 ori *lacZ* | Promega |
| pGSS33 | Tc$^r$ Ap$^r$ Cm$^r$ | 31 |
| pSUP202 | Tc$^r$ Ap$^r$ Cm$^r$ *mob* | 33 |
| pCB303b | Tc$^r$ *lacZ phoA* promoter-probe vector | 28 |
| pVK1 | Tc$^r$, carries *R. capsulatus* LH-II genes on an 8.5-kb insert in pRK290 | 4 |

[a] German collection of microorganisms, Braunschweig, Federal Republic of Germany.

was spotted onto the recipient cells. After an overnight incubation at 32°C, the cells were suspended in RÄ and plated on RÄ containing the appropriate antibiotics.

**Plasmid constructions.** As a first step, the 4.5-kilobase-pair (kb) *Eco*RI-*Pst*I fragment of the plasmid pVK1 was subcloned in pGEM-3 (leading to plasmid pG3-EP4.5; Fig. 1) and the *Sma*I sites were mapped. This was done by *Bal*31 digestion of the linearized plasmid and restriction analysis of the digestion products at different time points. With the exception of the 1,266- and the 600-bp fragments mapping downstream from the 1,266-bp fragment, which we could not obtain as single inserts in pGEM-3, all *Sma*I fragments (Fig. 1) were cloned through pGEM-3 into M13mp18 and M13mp19. The 1,266-bp *Sma*I fragment was electroeluted from the gel using the Biotrap (Schleicher & Schuell, Dassel, Federal Republic of Germany) and digested with *Nru*I, and the subfragments were cloned through pGEM-3 into M13mp18 and M13mp19. *Sau*3AI subclones in M13 were prepared from three isolated fragments: the 776-bp *Nru*I-*Sma*I fragment, the 600-bp *Sma*I fragment mapping upstream of the 1,266-bp fragment, and the 1,065-bp *Sma*I fragment (Fig. 1). To obtain clones overlapping the *Sma*I sites, the 600-bp *Apa*I fragment containing *pucB* and *pucA* was cloned into pGEM-7Zf(+), and exonuclease III/nuclease S1-generated deletion subclones of the 3-kb *Nru*I-*Eco*RI fragment (Fig. 1) cloned in pGEM-3 (plasmid pG3-EN3.0; digestion starting at the *Nru*I site) or pGEM-7Zf(+) (plasmid pG7EN3.0; digestion starting at the *Eco*RI site) were isolated by the method of Henikoff (14). For the latter experiments, the Erase-a-Base system (Promega) was used. The inserts of deletion derivatives of pG3-EN3.0 were subcloned into M13 for sequencing.

**Preparation of single-stranded DNA.** Single-stranded DNA was prepared from M13 derivatives by standard methods (23) or from pGEM-7Zf(+) derivatives by using the helper phage R408 (26).

**Oligonucleotides.** The universal primer (17-mer) was used for sequencing M13 clones, and the T7 promoter primer (Promega) was used for pGEM-7Zf(+) derivatives. An oligonucleotide with the sequence CATGGAAGTCAGATCCA (specific for the ends of transposon Tn*5*) and a specific primer for sequencing were obtained as custom syntheses from Syn-Tek AB, Umeå, Sweden.

**DNA sequencing.** The chain termination method of Sanger et al. (27) was used throughout this work. For sequencing with DNA polymerase I (Klenow fragment), the Deaza nucleotide reagent kit from American Bionetics, Inc., Emeryville, Calif. was used. The reaction buffer supplied was replaced by low-salt buffer [10 mM Tris(hydroximethyl)-aminomethane hydrochloride (pH 7.5), 10 mM $MgCl_2$, 5 mM dithiothreitol] [$^{35}$S]dCTP (Amersham-Buchler, Braunschweig, Federal Republic of Germany) was used as the radioactive label. Sequencing with T7 polymerase was done with the Sequenase kit from U.S. Biochemical Corp., Cleveland, Ohio, and [$^{35}$S]dATP as the label. Wedge-shaped (0.1 by 0.4 mm), 5% polyacrylamide–7 M urea sequencing gels were run at 55°C and 20 mA constant current.

**Determination of amino acid sequences.** The B800-850 complex was isolated from *R. capsulatus* Y5 cells and purified as described previously (5). The $M_r$-14,000 protein was extracted from the complex with $Na_2CO_3$. An LH-II preparation (2.5 mg of protein per ml) was mixed 1:30 (vol/vol) with 210 mM $Na_2CO_3$ and stirred for 30 min at room temperature. After 16 h at 0°C, the suspension was centrifuged for 16 h at 46,000 rpm in an 60 Ti rotor (Beckman Instruments, Inc., Fullerton, Calif.) at 4°C. The clear supernatant was dialyzed against distilled water and concentrated 20-fold. The supernatant and the sediment were analyzed by electrophoresis on an 11.5 to 16.5% polyacrylamide-sodium dodecyl sulfate gel as previously described (21).

For amino acid sequence analysis, the extracted $M_r$-14,000 protein was electrophoresed on a polyacrylamide gel and electroblotted onto Immobilon membranes (Millipore, Eschborn, Federal Republic of Germany). The bands formed by the $M_r$ 14,000 protein were excised and sequenced by automated Edman degradation using a pulsed-liquid gas-phase sequencer (model 477A; Applied Biosystems, Inc., Foster City, Calif.) with an on-line phenylthiohydantoin derivative analyzer (model 120 A; Applied Biosystems). The $M_r$-14,000 protein remaining bound to the complex after extraction was also electroblotted and sequenced.

Fragments of the $M_r$-14,000 protein were prepared by digestion with trypsin as follows. After 45 μg of the $M_r$-14,000 protein was suspended in 120 μl of buffer (0.1% sodium dodecyl sulfate, 2% β-octylglycopyranoside, 100 mM Tris hydrochloride, pH 8.3), digestion was performed for 2 h at 37°C with 2 μg of trypsin (sequencer grade; Boehringer GmbH, Mannheim, Federal Republic of Germany). The reaction was stopped by boiling for 5 min.

The tryptic digest was separated on a high performance

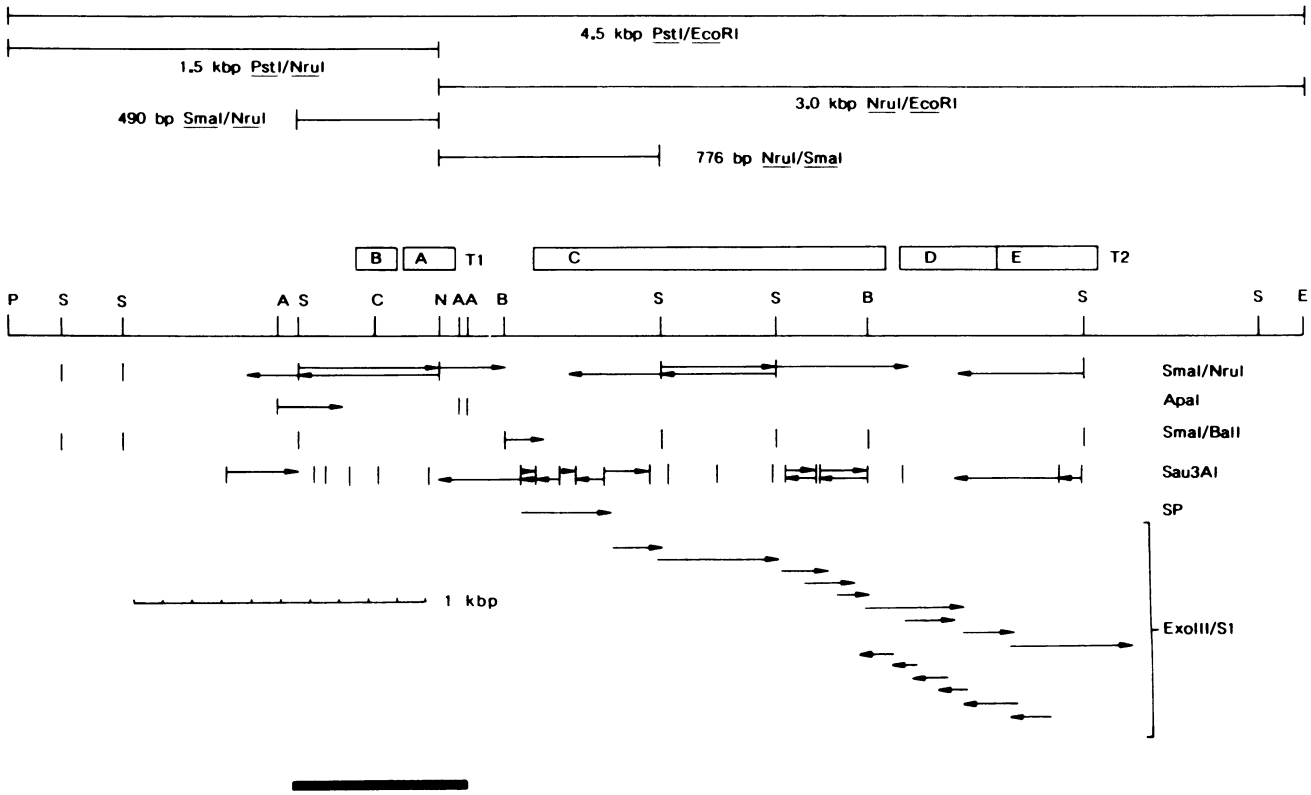


FIG. 1. Physical and genetic map of the insert of pG3-EP4.5. The 4.5-kb *Pst*I-*Eco*RI fragment of *R. capsulatus* DNA containing the genes *pucB*, *pucA*, and *pucE* and ORF C and ORF D (open boxes marked B, A, C, D, and E, respectively). The terminatorlike sequences are designated T1 and T2. With the exception of the *Sma*I and *Cla*I sites, all restriction sites (A, *Apa*I; B, *Bal*I; C, *Cla*I; E, *Eco*RI; N, *Nru*I; P, *Pst*I; S, *Sma*I) were mapped by using single and double digests. The *Sma*I fragments were ordered by linearization of the plasmid pG3-EP4.5 to the left or right of the insert, followed by *Bal*31 digestion for different times. After *Sma*I restriction of the *Bal*31 digests, the disappearance of *Sma*I fragments allowed their positioning on the physical map. Since *Cla*I is affected by *E. coli dam* methylation of the target sequence (9), the *Cla*I site present in the DNA sequence of *pucB* was not detected in the digests. The sizes of the internal *Sma*I fragments are (from left to right): 220, 600, 1,266, 400, 1,065, and 600 bp. At the top DNA fragments used in promoter tests and complementation experiments are shown together with their designations. The lower part of the figure shows the fragments used for DNA sequencing as well as the direction and extent of sequence information obtained with each fragment. SP indicates the use of a specific primer. Sequence information obtained with exonuclease III-nuclease S1-generated deletion subclones is denoted ExoIII/S1. The closed bar at the bottom shows the extent of DNA sequence information reported by Youvan and Ismail (41) for *R. capsulatus* SB1003.

liquid chromatography column [Bakerbond wide pore butyl (C4); 5-μm pore size; 4.6 by 250 mm] for 60 min with a gradient of 0.1% trifluoroacetic acid versus 60% acetonitril in 0.09% trifluoroacetic acid.

Elution was done at 50°C with a flow of 0.5 ml/min (high-performance liquid chromatography system; Knauer KG, Berlin). Peptides were detected at 220 nm with a sensitivity of 0.08 absorbance units full scale.

**Computer analyses.** The programs of the Kröger menu (18, 19; unpublished programs) available at the Universitäts-Rechenzentrum, Freiburg, Federal Republic of Germany, and the PC/GENE software package (Genofit, Geneva, Switzerland) were used to analyze DNA and amino acid sequences.

**Plasmid isolation.** Plasmids were isolated with the Triton lysis-boiling method (15). For small-scale isolations, 1.5 ml of cells were used. Derivatives of pCB303 were isolated from 10-ml cultures. For large-scale preparations, the protocol was scaled up for 50-ml culture volumes and the DNA was purified by CsCl density gradient centrifugation. Vector DNA of pCB303 was isolated by the method of Hansen and Olsen (13) from 1 liter of culture.

**Extraction of RNA and northern blotting (RNA blotting).** Total cellular RNA was extracted as described by Zhu and Kaplan (45). The RNA was denatured in formaldehyde, and RNAs of different lengths were separated by gel electrophoresis by using 6% formaldehyde–1.2% agarose gels. The RNA was transferred to nylon membranes (Amersham-Buchler) as recommended by the manufacturer of the membrane. Single-stranded DNA of an M13 recombinant phage carrying the 1,065-bp *Sma*I fragment of pG3-EP4.5 (Fig. 1) was used as the template to synthesize a $^{32}$P-labeled complementary strand which was subsequently used as the probe for the Northern blots.

**Enzyme assay.** Determination of β-galactosidase activity was done by the method of Miller (24).

## RESULTS

**Characterization of the Tn*5* mutant NK3.** The Tn*5* insertion mutant NK3 of *R. capsulatus* is LH-II negative but is not impaired in carotenoid biosynthesis and photosynthetic activities (16). To map the Tn*5* insertion to the nucleotide level, a subfragment of the plasmid pNK135 which contains the *Eco*RI fragment of the *R. capsulatus* NK3 genome carrying Tn*5* was cloned in M13 and sequenced by using a primer specific for the Tn*5* ends. The insertion point was localized 526 bp downstream from the *pucA* termination

codon (Fig. 2). This position is in agreement with our results of *Sma*I restriction mapping of the Tn*5* insertion in the plasmid pNK135. Since a Tn*5* insertion downstream from *pucA* affects the formation of the LH-II complex, sequences and/or genes downstream from *pucA* are necessary for wild-type LH-II formation. Klug et al. (17) have shown that in the mutant NK3, the level of B800-850-specific mRNAs is strongly decreased. Thus, the Tn*5* insertion in NK3 affects the formation or stability of mRNAs (including the *pucBA* mRNA) encoded by the *R. capsulatus* DNA cloned in pVK1.

**Restriction mapping and sequencing strategy.** The *R. capsulatus puc* operon has been cloned on the plasmid pVK1 (4). We subcloned in pGEM-3 the 4.5-kb *Pst*I-*Eco*RI fragment containing *pucB* and *pucA* for further analysis. The restriction map of this fragment is shown in Fig. 1. The strategy of the restriction mapping is described in the legend to Fig. 1. The DNA sequence of 3,169 bp of the 4.5-kb *Pst*I-*Eco*RI fragment was determined mainly by using defined restriction fragments in M13 or pGEM-7Zf(+) or exonuclease III-nuclease S1-generated deletion subclones (Fig. 1). *Sau*3AI subfragments from three individually isolated *Sma*I fragments were cloned in M13 and sequenced. A specific primer homologous to the positions 986 to 1002 (Fig. 2) was used to complete the DNA sequence of the 776-bp *Nru*I-*Sma*I fragment. The fragments used for sequencing and the extent of sequence information obtained with each fragment are shown in Fig. 1.

**DNA sequences of the *pucB* and *pucA* genes and upstream DNA of *R. capsulatus* 37b4.** The reported sequence from strain SB1003 (41) was compared with the sequence of strain 37b4. At position 304 (Fig. 2), a T instead of a C is present in 37b4. Thus, the direct repeat ACACTT implied in transcription start (47) is only a 5-bp repeat in 37b4. The stem of the terminatorlike sequence following *pucA* is two G · C and one A · T pair longer in 37b4 owing to the presence of two additional cytosine residues at positions 833 and 834 (Fig. 2). Otherwise, the sequences are identical.

The DNA sequence upstream of *pucB* shows the presence of several inverted repeats (Fig. 2) and five copies of the direct repeat CCGCA at positions 35, 75, 100, 156, and 264. The sequence CCGCA is part of the 10-bp direct repeat CCGCAAGGCG at positions 75 and 100 and part of the 8-bp direct repeat TGCCGCAG at positions 154 and 262. The 10- and 8-bp repeats contain 3-bp inverted repeats (CGCNGCG and $TGC(N)_3GCA$, respectively). The other occurences of the sequence CCGCA at positions 1137, 2718, and 2907 show that this sequence is much less frequent in and downstream from the *pucB* and *pucA* genes than upstream from these genes. The function of these inverted and direct repeats is not known at present.

**DNA sequences downstream from *pucBA*.** Since insertion of a transposon into the region downstream from the *pucA* gene blocks the formation of the B800-850 LH complex, the DNA sequence of that region was determined. We expected to obtain insight into the organization and properties of potential new *puc* genes. The sequence revealed three ORFs with codon preferences similar to the one described for five genes and seven ORFs of the photosynthetic gene cluster of *R. capsulatus* (39). The first of the ORFs, C, codes for a protein of 403 amino acids. It contains one large and three small hydrophilic domains containing numerous charged amino acid residues in a mainly hydrophobic surrounding (Fig. 3A). The two following ORFs, D and E, code for small proteins of 113 and 118 amino acids, respectively. The polypeptide deduced from ORF E is mainly hydrophilic and contains one hydrophobic domain close to the C terminus (Fig. 3B). A molecular weight of 13,493 was calculated for this polypeptide. The amino acid sequence analysis (see below) of the $M_r$-14,000 protein present in the LH-II complex assigns ORF E to the structural gene for the $\gamma$ subunit of the LH-II complex. ORF E is followed by a sequence similar to *rho*-independent transcription terminators. This terminator structure consists of a G+C-rich hairpin (calculated stability, −144.9 kJ) which is followed by seven uridine residues and one adenosine residue. This is remarkable because it has been shown that only four uridine residues after a G+C-rich hairpin structure lead to efficient termination (2).

**Partial sequence of the gamma subunit of the LH-II complex.** The amino acid sequence of the first 20 residues of the $M_r$-14,000 protein extracted with $Na_2CO_3$ from the complex (see Materials and Methods) and the first 10 residues of the $M_r$-14,000 protein remaining in the complex after extraction were determined. Residues 2 to 20 and 2 to 10, respectively, were identical to the N terminus of the polypeptide deduced from the DNA sequence of ORF E, the last ORF in the *puc* gene cluster (Fig. 1, 2, and 4). The amino acid sequence of tryptic peptides also corresponded to parts of the deduced polypeptide sequence (Fig. 4). ORF E can thus be assigned to the structural gene for the $M_r$-14,000 protein of the LH-II complex and should therefore tentatively be designated *pucE*. The amino acid composition of this $M_r$-14,000 protein, however, strongly differs from the amino acid composition of a $M_r$-14,000 polypeptide isolated from the LH-II complex of *R. capsulatus* by organic-solvent extraction (32, 37).

**Complementation experiments.** It has been shown previously that the plasmid pVK1 carrying an 8.5-kb *Eco*RI fragment of *R. capsulatus* can complement the Tn*5* mutation in the strain NK3 in *trans* (4). Since the Tn*5* insertion in this mutant is located over 500 bp downstream from the *pucA* termination codon (see above), we tried to complement this mutant with smaller fragments of that region. The 3.0-kb *Nru*I-*Eco*RI fragment (Fig. 1) was cloned in the wide-host-range vector pGSS33 (giving plasmid pGSS33-EN3.0) and in the narrow-host-range, mobilizable plasmid pSUP202 (leading to plasmid pSNF-EN3.0). After transfer of these plasmids into *R. capsulatus* NK3, no reconstitution of the LH-II-negative phenotype was observed, either with the plasmid pGSS33-EN3.0, which replicates in *R. capsulatus*, or with the plasmid pSNF-EN3.0, which integrates via a single crossover, at the DNA sequence homologous to that of the insert, into the chromosome.

The 4.5-kb *Pst*I-*Eco*RI fragment (Fig. 1), however, was able to reconstitute the mutant NK3 when located extrachromosomally (cloned in pGSS33; plasmid pGSS33-EP4.5) and when integrated (see above) into the chromosome (cloned in pSUP202; plasmid pSUP202-EP4.5).

**Transcription of *puc* genes downstream from *pucA*.** We used Northern hybridizations to characterize the transcription of genes downstream from *pucA* in the wild-type *R. capsulatus* 37b4 and the mutant NK3. As shown in Fig. 5, the wild-type synthesized an RNA species of about 1.1 kb hybridizing with the 1,065-bp *Sma*I fragment (Fig. 1) which contains sequences of ORF C, ORF D, and *pucE* (Fig. 1). The level of hybridizing RNA was induced in wild-type cells after a shift from aerobic to semiaerobic conditions but was not detectable in the mutant NK3 (Fig. 5).

**Promoter analyses.** To analyze the *pucBA* promoter region and to search for a possible promoter downstream from *pucA*, we cloned the 490-bp *Sma*I-*Nru*I fragment, the 1.5-kb *Pst*I-*Nru*I fragment, and the 776-bp *Nru*I-*Sma*I fragment

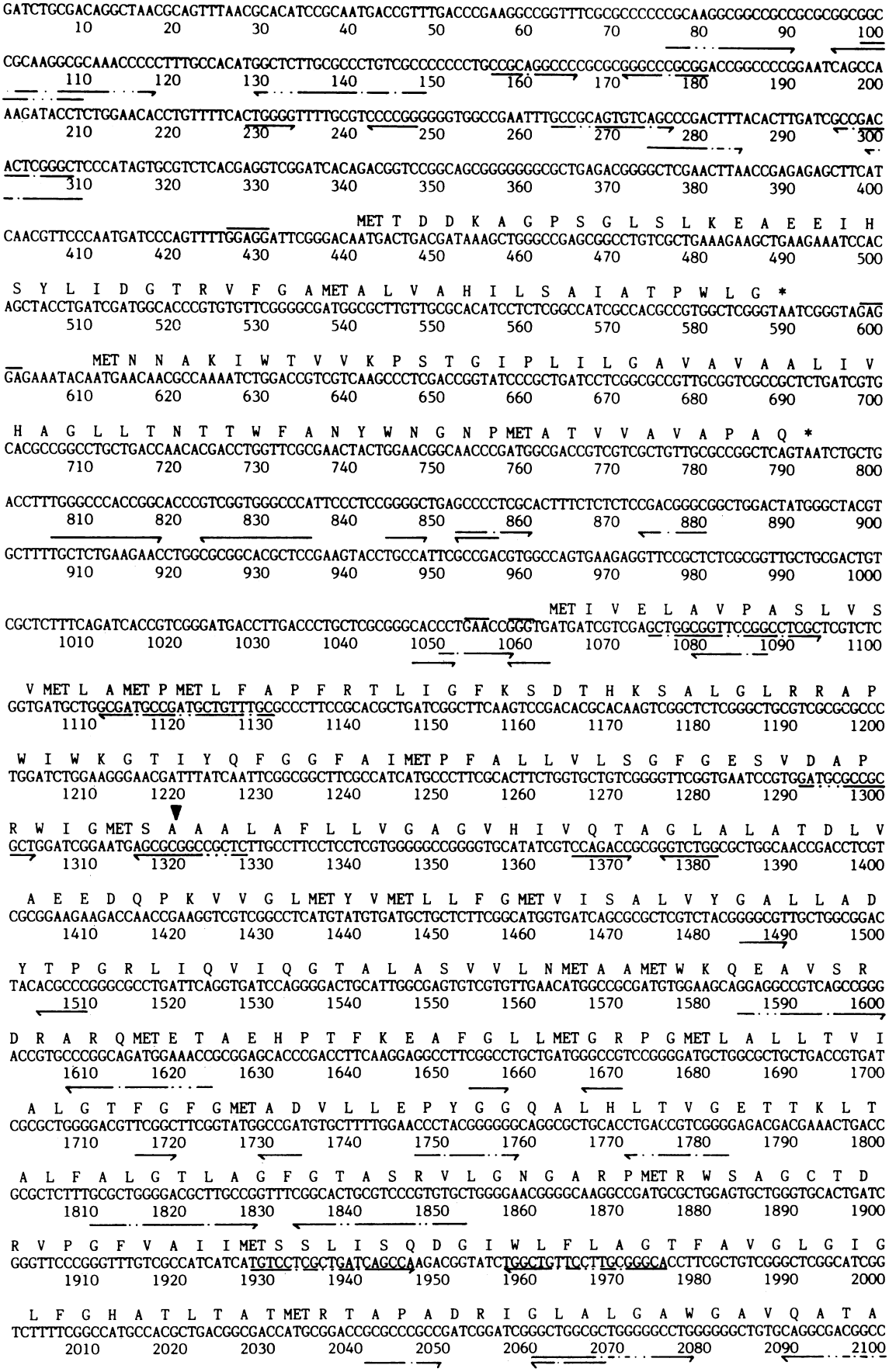

FIG. 2

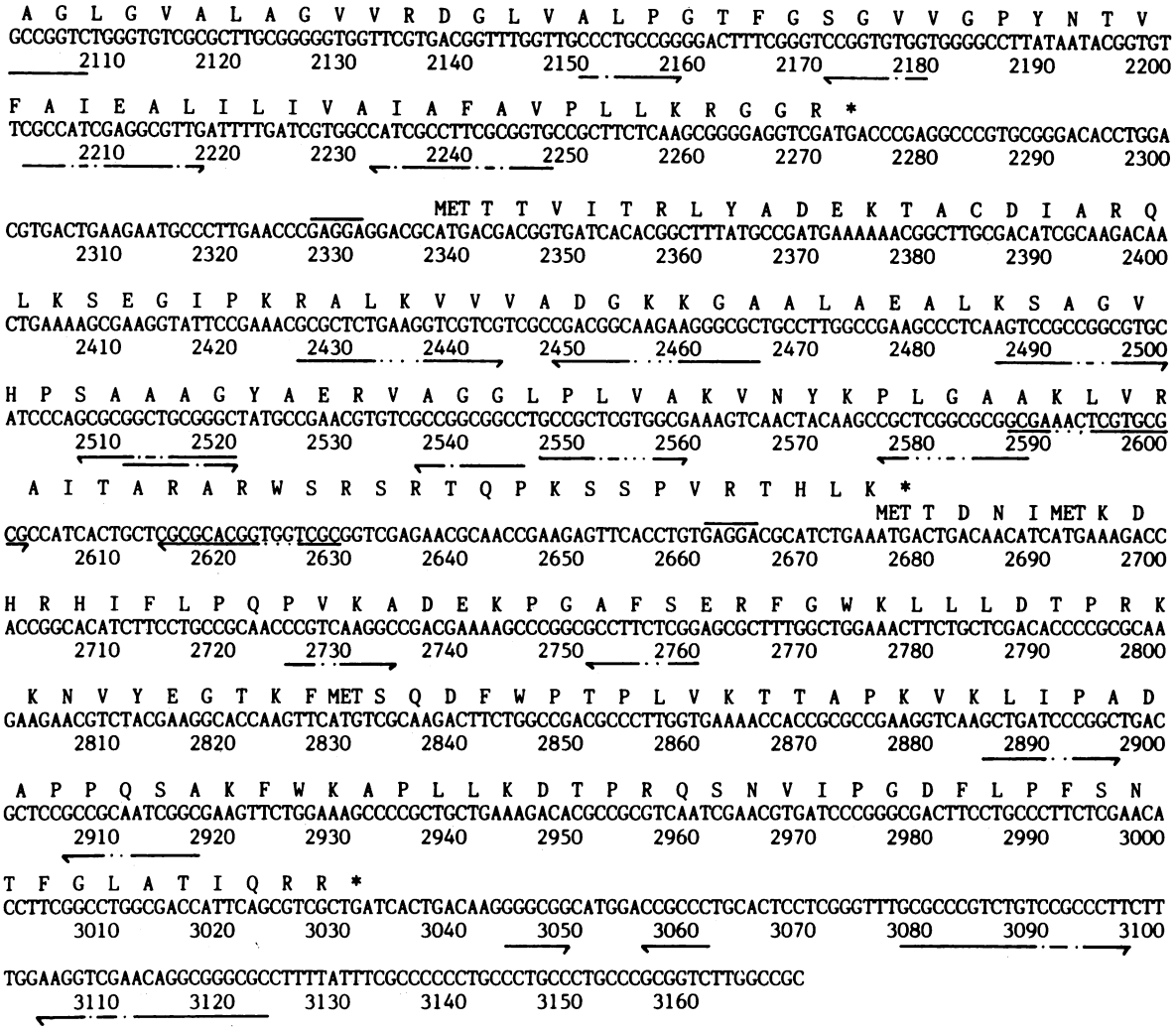


FIG. 2. DNA sequence of the *R. capsulatus puc* operon including upstream and downstream sequences. Amino acid sequences of proteins and deduced polypeptides are given above the DNA sequence. Arrows below the sequence designate inverted repeat sequences; putative Shine-Dalgarno sequences are indicated by lines above the sequence. The arrowhead points to the Tn*5* insertion point in the LH-II-negative mutant NK3.

(Fig. 1), the latter in both orientations, into the promoter-probe vector pCB303b. This vector is a derivative of pRK290 and allows the construction of transcriptional fusions to *lacZ* (or *phoA*) (28).

The 490-bp *Sma*I-*Nru*I fragment (plasmid pCB303b-SN0.5) showed high transcription of *lacZ*, whereas the larger *Pst*I-*Nru*I fragment (plasmid pCB303b-PN1.5) led to low *lacZ* expression (Table 2). Although activities obtained with anaerobically grown cells were significantly higher than values from aerobically grown cells, the *lacZ* expression did not reflect the kinetics of the formation of the *pucBA* mRNA after a shift from aerobic to semiaerobic growth conditions (17). This can be due to limitations of the test system or to the possibility that sequences downstream from the *Nru*I site have to be present in *cis* to allow wild-type transcription from the *puc* promoter. The overall lower *lacZ* expression from the larger fragment could be due to the presence of additional regulatory sequences upstream of *pucB*. The β-galactosidase activities determined for the fusions with the 776-bp *Nru*I-*Sma*I fragment (plasmids pCB303b-SN0.7 and pCB303-SN0.7b) were similar to the one determined for the vector control (Table 2).

## DISCUSSION

In this work, we have shown that the *puc* locus of *R. capsulatus* contains more genes than the previously analyzed genes *pucB* and *pucA*. We hypothesize that at least one of the genes downstream from *pucA* has a regulatory function, since mutants defective in the expression of genes downstream from *pucA* showed a strongly decreased level of *puc* mRNA (17). Amino acid sequence studies suggest that *pucE* is the structural gene for the $M_r$-14,000 protein of the *R. capsulatus* LH-II complex. The discrepancy in the amino acid composition of $M_r$-14,000 polypeptides extracted by organic solvents (32, 37) and by alkaline solute (this work) indicates that more than one polypeptide of this $M_r$ exists. Both polypeptides are hydrophilic, but the alkaline-extracted $M_r$-14,000 polypeptide has a positive net charge, while the organic-solvent-extracted polypeptide has a negative net charge. The isolation of the *pucE* gene opens up the possibility of constructing mutants defective only in the gene for the $M_r$-14,000 protein, which does not bind pigments (6).

The structure of the polypeptide deduced from the ORF C sequence is similar to those of integral membrane proteins

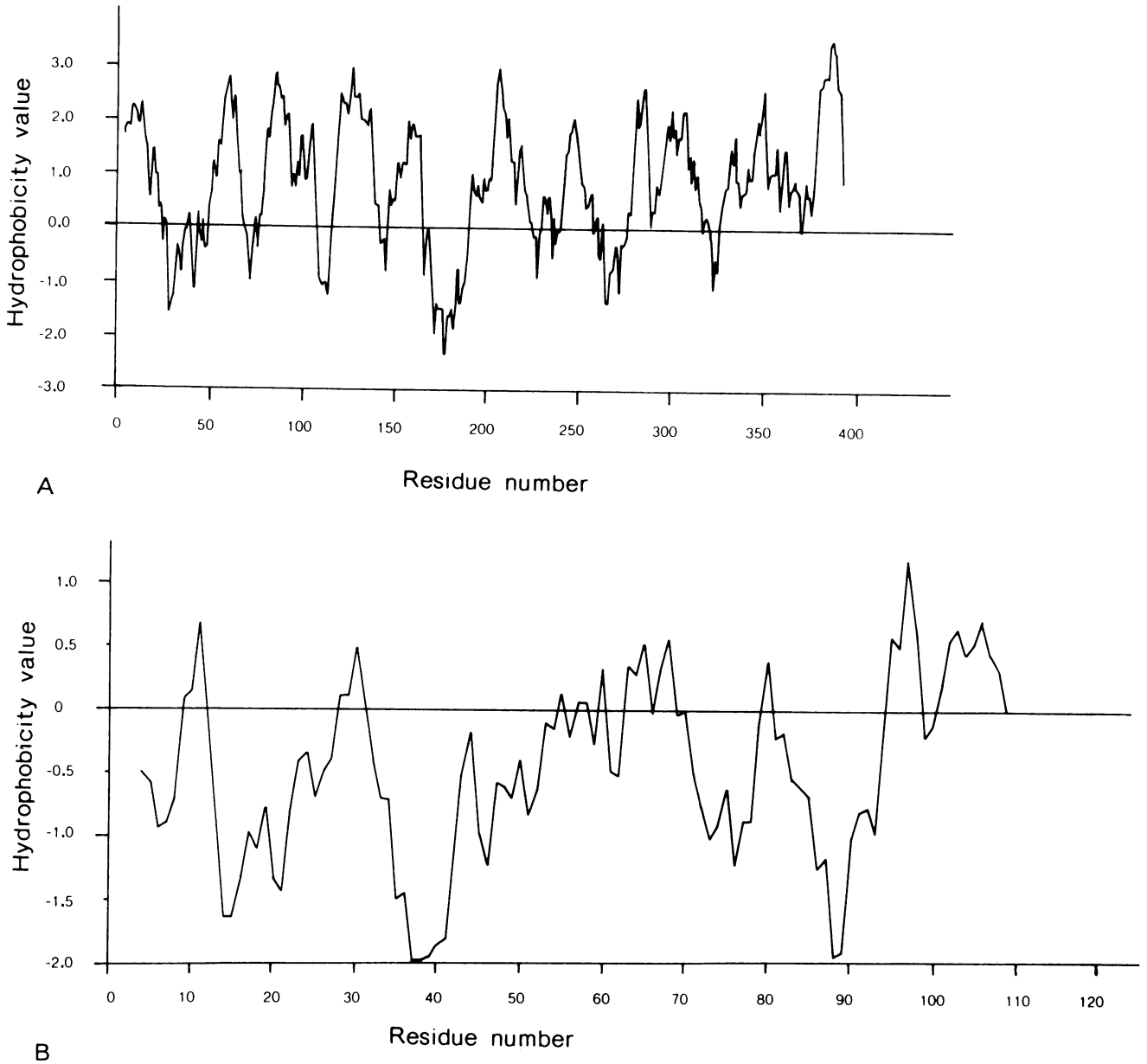


FIG. 3. Hydropathy plots of deduced polypeptides. The hydropathy values were calculated by the method of Kyte and Doolittle (20), with a moving window of 9 residues. (A) Hydropathy plot for the polypeptide deduced from ORF C. The polypeptide is mainly hydrophobic and contains one large and three smaller hydrophilic domains containing many charged amino acids. (B) Hydropathy plot for the polypeptide deduced from ORF E. The protein shows one hydrophobic domain close to the C terminus.

like FecC (34) or BtuC (8) involved in transport processes. Whether the similarity of the structure is a clue to the function of the hypothetical ORF C protein, however, remains to be determined, since no obvious sequence similarities between FecC or BtuC and the hypothetical ORF C polypeptide exist.

The experiments to characterize *puc* expression by using *lac* fusions on plasmids did not reflect the regulation of *puc* expression deduced from RNA studies (17, 43, 44). A possible reason for this result could be that the fragments used for the study do not contain all sequences necessary for wild-type transcription or that the plasmid location of the fragments affects inducibility of transcription. It has been reported that in *Bradyrhizobium japonicum* plasmid-borne *fixA*- or *fixB-lacZ* fusions only showed marginal activation by *nifA*, whereas fusions integrated in the chromosome showed *nifA*-dependent induction of the *fixA* and *fixBC* promoter under microaerobic conditions (12). A similar approach could be used to decide if *puc* expression can be characterized using *lacZ* fusions integrated in the chromosome. Alternatively, *lacZ* fusions to one of the genes downstream from *pucA* could be used, since an insertion of Mu dl(Ap$^r$ *lac*) in this region lead to oxygen-regulated β-galactosidase expression (17).

The transcription of the ORFs downstream from *pucBA* is not yet clear. The 3′ ends of the *pucBA* mRNA have been mapped to the terminatorlike structure following *pucA* (47). Together with the characteristics of the mutant NK3, this

```
          5         10        15        20        25        30
          |         |         |         |         |         |
1  M T D N I M K D H R H I F L P Q P V K A D E K P G A F S E R
31 F G W K L L L D T P R K K N V Y E G T K F M S Q D F W P T P
61 L V K T T A P K V K L I P A D A P P Q S A K F W K A P L L K
91 D T P R Q S N V I P G D F L P F S N T F G L A T I Q R R
```

FIG. 4. Amino acid sequence determination of the $M_r$-14,000 protein. The polypeptide sequence deduced from the DNA sequence is shown with the regions verified by amino acid sequence determinations. The solid line beneath the sequence designates amino acids identical with the deduced polypeptide sequence and the amino acid sequence determined from the $M_r$-14,000 protein. One point designates amino acids not unambiguously determined; two points mark an amino acid obtained in low yield as PTH derivative.

would imply the presence of a transcription start between the *pucBA* terminator and the Tn*5* insertion point in NK3. Complementation studies with cloned fragments on plasmids, however, were not successful in expressing the genes downstream from *pucA* without the presence of *pucB*, *pucA*, and upstream DNA in *cis*. Experiments using a promoter-probe vector gave no evidence for a functional promoter within 500 bp downstream from the *pucBA* terminator structure. Possibly, the activity of a promoter downstream from *pucA* is dependent on sequences not present on the hybrid plasmids used for complementation and promoter test experiments. The extrachromosomal location of the *puc* genes itself did not prevent their expression, as shown by the complementation of the mutant NK3 with pGSS33-EP4.5. Another possibility could be that at least one gene downstream from *pucA* is transcribed from the *pucBA* promoter. Although S1 mapping experiments (47) gave no evidence for transcription proceeding beyond the *pucBA* terminator, readthrough of this signal could be below the detection limit or it could occur only at a specific stage of development of the photosynthetic membrane. In this case, the expression of the LH-II complex would be controlled by an autoregulatory circuit. Several bacterial genes, e.g. *araC* (1) and *lysA* (22), regulate their own expression.

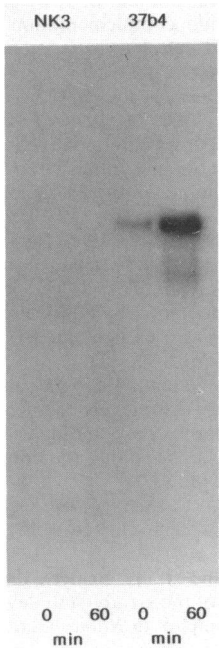


FIG. 5. Northern hybridization with a probe containing sequences of ORF C, ORF D, and *pucE*. Cells of the Tn*5* mutant *R. capsulatus* NK3 and of the wild-type *R. capsulatus* 37b4 were harvested from strongly aerated cultures (0 min) and 60 min after induction of the photosynthetic apparatus by lowering the oxygen partial pressure. The RNA was isolated, fractionated on an agarose gel, and hybridized with the probe. The autoradiograph was exposed for 6 h with an intensifying screen.

TABLE 2. Expression of extrachromosomal *puc-lacZ* fusions in *R. capsulatus* wild-type 37b4

| Plasmid | β-Galactosidase activity (U) under growth conditions[a]: | |
|---|---|---|
| | Aerobic | Anaerobic |
| pCB303b (control) | <5 | <5 |
| pCB303b-SN0.7 | <5 | <5 |
| pCB303b-SN0.7b | <5 | <5 |
| pCB303b-SN0.5 | 1,250 | 2,350 |
| pCB303b-PN1.5 | 75 | 200 |

[a] Values are means of at least three determinations.

## ACKNOWLEDGMENTS

This research was supported by grant Dr 29/31-3G of the Deutsche Forschungsgemeinschaft and Fonds der Chemischen Industrie.

We thank Nasser Gad'on for his contribution to the isolation of the $M_r$-14,000 polypeptide.

## LITERATURE CITED

1. **Casadaban, M. J.** 1976. Regulation of the regulatory gene for the arabinose pathway, araC. J. Mol. Biol. **104**:557–566.
2. **Chen, C.-Y. A., J. T. Beatty, S. N. Cohen, and J. G. Belasco.** 1988. An intercistronic stem-loop structure functions as an mRNA decay terminator necessary but insufficient for *puf* mRNA stability. Cell **52**:609–619.
3. **Drews, G.** 1983. Mikrobiologisches Praktikum, p. 62. Springer Verlag KG, Berlin.
4. **Drews, G., N. Kaufmann, and G. Klug.** 1985. The bacterial photosynthetic apparatus: molecular organization, genetic structure, and biosynthesis of the pigment-protein complexes, p. 211–222. *In* K. Steinback, S. Bonitz, C. J. Arntzen, and L. Bogorad (ed.), Molecular biology of the photosynthetic apparatus. Cold Spring Harbor Laboratory, Cold Spring Harbor, N.Y.
5. **Feick, R., and G. Drews.** 1978. Isolation and characterization of light harvesting bacteriochlorophyll-protein complexes from *Rhodopseudomonas capsulata*. Biochim. Biophys. Acta **501**: 499–513.
6. **Feick, R., and G. Drews.** 1979. Protein subunits of bacteriochlorophylls B802 and B855 of the light-harvesting complex II of *Rhodopseudomonas capsulata*. Z. Naturforsch. Teil C **34**:196–199.
7. **Frenkel, A. W.** 1954. Light induced phosphorylation by cell-free preparations of photosynthetic bacteria. J. Am. Chem. Soc. **76**:5568–5569.
8. **Friedrich, M. J., L. C. DeVeaux, and R. J. Kadner.** 1986. Nucleotide sequence of the *btuCED* genes involved in vitamin $B_{12}$ transport in *Escherichia coli* and homology with components of periplasmic-binding-protein-dependent transport systems. J. Bacteriol. **167**:928–934.
9. **Geier, G. E., and P. Modrich.** 1979. Recognition sequence of the dam methylase of *Escherichia coli* K12 and mode of cleavage of DpnI endonuclease. J. Biol. Chem. **254**:1408–1413.
10. **Geller, D. M.** 1962. Oxidative phosphorylation in extracts of *Rhodospirillum rubrum*. J. Biol. Chem. **237**:2947–2954.
11. **Giuliano, G., D. Pollock, H. Stapp, and P. A. Scolnik.** 1988. A genetic-physical map of the *Rhodobacter capsulatus* carotenoid biosynthesis gene cluster. Mol. Gen. Genet. **213**:78–83.

12. **Gubler, M., and H. Hennecke.** 1988. Regulation of the *fixA* gene and *fixBC* operon in *Bradyrhizobium japonicum*. J. Bacteriol. **170:**1205–1214.
13. **Hansen, J. B., and R. H. Olsen.** 1978. Isolation of large plasmids and characterization of the P2 incompatibility group plasmids pMG1 and pMG5. J. Bacteriol. **135:**227–238.
14. **Henikoff, S.** 1984. Unidirectional digestion with exonuclease III creates targeted breakpoints for DNA sequencing. Gene **28:** 351–359.
15. **Holmes, D. S., and M. Quigley.** 1981. A rapid boiling method for the preparation of bacterial plasmids. Anal. Biochem. **114:** 193–197.
16. **Kaufmann, N., H. Hüdig, and G. Drews.** 1984. Transposon Tn5 mutagenesis of genes for the photosynthetic apparatus in *Rhodopseudomonas capsulata*. Mol. Gen. Genet. **198:**153–158.
17. **Klug, G., N. Kaufmann, and G. Drews.** 1985. Gene expression of pigment-binding proteins of the bacterial photosynthetic apparatus: transcription and assembly in the membrane of *Rhodopseudomonas capsulata*. Proc. Natl. Acad. Sci. USA **82:**6485–6489.
18. **Kröger, M., and A. Kröger-Block.** 1984. Extension of a flexible computer program for handling DNA sequence data. Nucleic Acids Res. **12:**113–120.
19. **Kröger, M., and A. Kröger-Block.** 1984. Simplified computer programs for search of homology within nucleotide sequences. Nucleic Acids Res. **12:**193–201.
20. **Kyte, J., and R. F. Doolittle.** 1982. A simple method for displaying the hydropathic character of a protein. J. Mol. Biol. **157:**105–132.
21. **Laemmli, U. K.** 1970. Cleavage of structural proteins during the assembly of the head of bacteriophage T4. Nature (London) **227:**680–685.
22. **Martin, C., F. Borne, B. Cami, and J. C. Patte.** 1986. Autogenous regulation by lysine of the *lysA* gene of *Escherichia coli*. FEMS Microbiol. Lett. **36:**105–108.
23. **Messing, J.** 1983. New M13 vectors for cloning. Methods Enzymol. **101:**20–78.
24. **Miller, J. H.** 1972. Experiments in molecular genetics, p. 353–355. Cold Spring Harbor Laboratory, Cold Spring Harbor, N.Y.
25. **Nieth, K. F., and G. Drews.** 1975. Formation of reaction centers and light-harvesting bacteriochlorophyll-protein complexes in *Rhodopseudomonas capsulata*. Arch. Microbiol. **104:**77–82.
26. **Russel, M., S. Kidd, and M. R. Kelley.** 1986. An improved filamentous helper phage for generating single stranded plasmid DNA. Gene **45:**333–338.
27. **Sanger, F., S. Nicklen, and A. R. Coulson.** 1977. DNA sequencing with chain-terminating inhibitors. Proc. Natl. Acad. Sci. USA **74:**5463–5467.
28. **Schneider, K., and C. F. Beck.** 1987. New expression vectors for identifying and testing signal structures for initiation and termination of transcription. Methods Enzymol. **153:**452–461.
29. **Schumacher, A., and G. Drews.** 1978. The formation of bacteriochlorophyll-protein complexes of the photosynthetic apparatus of *Rhodopseudomonas capsulata* during early stages of development. Biochim. Biophys. Acta **501:**183–194.
30. **Schumacher, A., and G. Drews.** 1979. Effects of light intensity on membrane differentiation in *Rhodopseudomonas capsulata*. Biochim. Biophys. Acta **547:**417–428.
31. **Sharpe, G. S.** 1984. Broad host range cloning vectors for Gram-negative bacteria. Gene **29:**93–102.
32. **Shiozawa, J. A., P. A. Cuendet, G. Drews, and H. Zuber.** 1980. Isolation and characterization of the polypeptide components from light-harvesting pigment-protein complex B800-850 of *Rhodopseudomonas capsulata*. Eur. J. Biochem. **111:**455–460.
33. **Simon, R., U. Priefer, and A. Pühler.** 1983. A broad host range mobilization system for in vivo genetic engineering: transposon mutagenesis in gram negative bacteria. Biotechnology **1:**37–45.
34. **Staudenmaier, H., B van Hove, Z. Yaraghi, and V. Braun.** 1989. Nucleotide sequences of the *fecBCDE* genes and locations of the proteins suggest a periplasmic-binding-protein-dependent transport mechanism for iron(III) dicitrate in *Escherichia coli*. J. Bacteriol. **171:**2626–2633.
35. **Tadros, M. H., R. Frank, and G. Drews.** 1985. The complete amino-acid sequence of the small bacteriochlorophyll-binding polypeptide B800-850β from light-harvesting complex B800-850 of *Rhodopseudomonas capsulata*. FEBS Lett. **183:**91–94.
36. **Tadros, M. H., F. Suter, G. Drews, and H. Zuber.** 1983. The complete amino-acid sequence of the large bacteriochlorophyll-binding polypeptide from light-harvesting complex II (B800-850) from *Rhodopseudomonas capsulata*. Eur. J. Biochem. **129:** 533–536.
37. **Tadros, M. H., H. Zuber, and G. Drews.** 1982. The polypeptide components from light-harvesting pigment-protein complex II (B800-850) of *Rhodopseudomonas capsulata*. Solubilization, purification and sequence studies. Eur. J. Biochem. **127:**315–318.
38. **Yanisch-Perron, C., J. Vieira, and J. Messing.** 1985. Improved M13 phage cloning vectors and host strains: nucleotide sequences of the M13mp18 and pUC19 vectors. Gene **33:**103–119.
39. **Youvan, D. C., E. J. Bylina, M. Alberti, H. Begusch, and J. E. Hearst.** 1984. Nucleotide and deduced polypeptide sequences of the photosynthetic reaction-center, B870 antenna, and flanking polypeptides from *R. capsulata*. Cell **37:**949–957.
40. **Youvan, D. C., J. T. Elder, D. E. Sandlin, K. Zsebo, D. P. Alder, N. J. Panopoulos, B. L. Marrs, and J. E. Hearst.** 1982. R-prime site-directed transposon Tn7 mutagenesis of the photosynthetic apparatus in *Rhodopseudomonas capsulata*. J. Mol. Biol. **162:** 17–41.
41. **Youvan, D. C., and S. Ismail.** 1985. Light-harvesting II (B800-B850 complex) structural genes from *Rhodopseudomonas capsulata*. Proc. Natl. Acad. Sci. USA **82:**58–62.
42. **Youvan, D. C., S. Ismail, and E. J. Bylina.** 1985. Chromosomal deletion and plasmid complementation of the photosynthetic reaction center and light harvesting genes from *Rhodopseudomonas capsulata*. Gene **38:**19–30.
43. **Zhu, Y. S., D. N. Cook, F. Leach, G. A. Armstrong, M. Alberti, and J. E. Hearst.** 1986. Oxygen-regulated mRNAs for light-harvesting and reaction center complexes and for bacteriochlorophyll and carotenoid biosynthesis in *Rhodobacter capsulatus* during the shift from anaerobic to aerobic growth. J. Bacteriol. **168:**1180–1188.
44. **Zhu, Y. S., and J. E. Hearst.** 1986. Regulation of expression of genes for light-harvesting antenna proteins LH-I and LH-II; reaction center polypeptides RC-L, RC-M and RC-H; and enzymes of bacteriochlorophyll and carotenoid biosynthesis in *Rhodobacter capsulatus* by light and oxygen. Proc. Natl. Acad. Sci. USA **83:**7613–7617.
45. **Zhu, Y. S., and S. Kaplan.** 1985. Effects of light, oxygen, and substrates on steady-state levels of mRNA coding for ribulose-1,5-bisphosphate carboxylase and light-harvesting and reaction center polypeptides in *Rhodopseudomonas sphaeroides*. J. Bacteriol. **162:**925–932.
46. **Zsebo, K. M., and J. E. Hearst.** 1984. Genetic-physical mapping of a photosynthetic gene cluster from *R. capsulata*. Cell **37:** 937–947.
47. **Zucconi, A. P., and J. T. Beatty.** 1988. Posttranscriptional regulation by light of the steady-state levels of mature B800-850 light-harvesting complex in *Rhodobacter capsulatus*. J. Bacteriol. **170:**877–882.