JOURNAL OF BACTERIOLOGY, Jan. 1989, p. 473–482
0021-9193/89/010473-10$02.00/0

Vol. 171, No. 1

Structural and Functional Analysis of Transcriptional Control of the *Rhodobacter capsulatus puf* Operon

CAMELLIA W. ADAMS,[1] MARY E. FORREST,[2] STANLEY N. COHEN,[1,3] AND J. THOMAS BEATTY[2,4]*

Departments of Genetics[1] and Medicine,[3] Stanford University School of Medicine, Stanford, California 94305, and Departments of Microbiology[2] and Medical Genetics,[4] University of British Columbia, Vancouver, British Columbia, Canada V6T 1W5

Received 18 July 1988/Accepted 18 October 1988

We report data indicating that the *Rhodobacter capsulatus puf* operon promoter and the site for its oxygen regulation are located more than 700 base pairs upstream from the previously identified *puf* genes and have identified the nucleotide sequences that constitute these control signals. A model is proposed in which a polycistronic transcript at least 3.4 kilobases in length is initiated near the O_2-regulated promoter and is processed posttranscriptionally by endonucleolytic cleavage at multiple sites, yielding discrete mRNA segments that are degraded at different rates. A newly identified gene (*pufQ*), which includes a hydrophobic domain having some similarity to domains of the products of the *pufL* and *pufM* genes, begins 313 nucleotides into the *puf* transcript and is located entirely within the most rapidly degraded segment of the transcript. A previously identified *puf* transcript segment encoding structural proteins for photosynthetic membrane complexes persists after degradation of the most 5′ region of the transcript and is itself subject to segmentally specific degradation. Our results suggest a model in which differential expression of the multiple genes encoded by the *puf* operon is at least in part attributable to major differences in the rates of decay of the various segments of *puf* mRNA.

Organisms capable of growth under both aerobic and anaerobic conditions have evolved the ability to respond to the concentration of molecular oxygen in their environment by modulating the level of expression of relevant genes. Anaerobically induced genes have been identified in a variety of bacterial species, including *Escherichia coli*, *Salmonella typhimurium*, *Klebsiella pneumoniae*, and the photosynthetic bacteria (1, 7, 9, 24, 34, 41). In facultatively photosynthetic bacteria such as *Rhodobacter capsulatus*, alteration of gene expression by environmental factors, such as oxygen concentration and light intensity, affects the amount of pigment-protein complexes in photosynthetic membranes (16) and consequently affects the ability to initiate electron transport and produce ATP.

The photosynthetic membranes of *R. capsulatus* include the β and α subunits of the light-harvesting I (LHI or B870) complex and the L and M subunits of the reaction center (RC), all of which are encoded by a polycistronic operon (7) previously known as *rxcA* but recently renamed *puf* (22). Although the inductive effect of oxygen deprivation on synthesis of the photosynthetic apparatus has been well documented (12, 16), it has been shown only recently that the steady-state amounts of *puf* mRNAs increase with decreased culture aeration. This has been interpreted as indicating that expression of the genes of the *puf* operon is controlled at the level of transcription (7, 11, 44, 45). In addition, differential expression of the LHI and RC genes is accomplished at least in part by differential degradation of different transcript segments (7, 10, 23), and the mRNA species encoded by the various genes within the *puf* operon are not present in stoichiometric amounts.

During investigations of *puf* gene expression, we observed that sequences immediately upstream from the *pufB* gene, which is the most 5′ of the previously identified *puf* operon genes, were unable to initiate mRNA transcription in *R. capsulatus* as measured by β-galactosidase expression from *puf-lacZ* fusion constructs, whereas sequences located several hundred nucleotides upstream from *pufB* could accomplish O_2-regulated transcription (5). This finding suggested that the *puf* promoter may be separated by a great distance from the structural genes encoded by its transcript. To determine the actual location of the *puf* promoter, to identify the site of oxygen regulation of *puf* gene expression, and to elucidate certain other aspects of genetic control within the *puf* operon, we have used *puf-lacZ* fusions in conjunction with deletion analysis, site-directed mutagenesis, and mRNA mapping. Our experiments have localized *cis*-active DNA sequences involved in the initiation of *puf* mRNA synthesis and its regulation by oxygen and have revealed a complex pattern of processing of *puf* mRNA.

MATERIALS AND METHODS

Bacterial strains, growth conditions, and β-galactosidase assays. The wild-type strain B10 (27, 37) and the *puf* deletion mutant ΔRC6 (10) were grown under conditions of high and low oxygen as previously described (7). For β-galactosidase assays (30), cultures were inoculated at a cell density of 0.8×10^8 cells per ml and grown to a cell density of 3.2×10^8 cells per ml.

Construction of the promoter-probe vector pXCA601. A partial restriction map of the promoter-probe vector used in these studies, pXCA601, is given in Fig. 1. The fragment containing the *lacZ* gene was obtained from plasmid pMC1403*lac* (20). The 3′ end of fusion transcripts was stabilized by insertion of the *ompA* terminator (6). To prevent transcription initiated within the vector from continuing into the DNA inserts being tested for promoter

* Corresponding author.

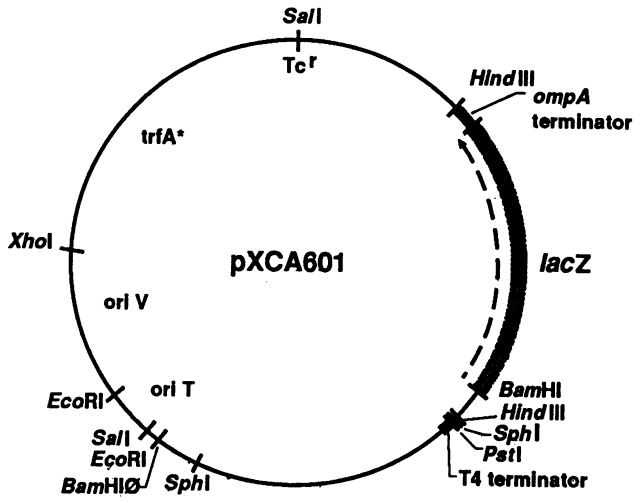

FIG. 1. The promoter-probe vector pXCA601. Insertion of DNA fragments at the *Bam*HI site of the vector creates in-frame translational fusions to the β-galactosidase gene, which is oriented as shown by the dashed arrow. The locations of the bacteriophage T4 (32) and *E. coli ompA* (6) terminators are indicated by bold lines.

activity, a 142-base-pair *Sau*3A fragment containing the T4 phage translation-transcription termination signals (32) was introduced upstream of the insertion site. The replicon for this vector was derived from pTJS133 (33), which can be mobilized into *R. capsulatus* and is stably maintained in this host as well as in *E. coli*.

Construction of *lacZ* fusion plasmids. The plasmid pJAJ21 (21) contains the region of *R. capsulatus* DNA extending from the *Xho*II site shown in Fig. 2 to the *Sph*I site in the *pufB* gene. pJAJ21 was cleaved with *Xho*I, and deletions were generated by treatment with exonuclease III as described previously (19). The 5′-end-shortened fragments were released by *Bam*HI cleavage and subcloned into pXCA601 to create pΔ4, pΔ14, and pΔ24 (Fig. 3). The 935 construct was made by digestion of pJAJ21 with *Pst*I and *Bam*HI and subcloning of the insert into pXCA601; the *Sal*I-to-*Bam*HI fragment of pJAJ21 was introduced into pXCA601 to yield the pΔ32 construct. The pΔMSP plasmid contains the *Mnl*I-to-*Bam*HI fragment of pJAJ21, while pΔAM contains the *Pst*I-to-*Acc*I fragment of pJAJ21 fused to

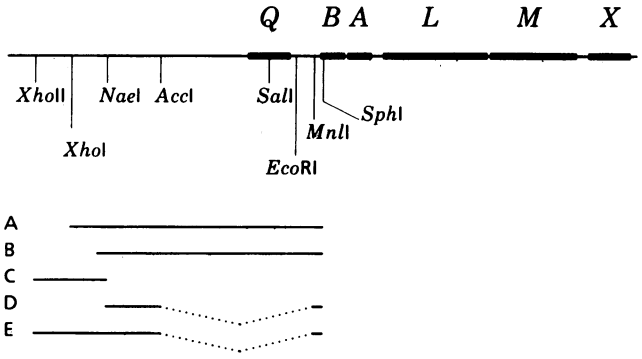

FIG. 2. Genetic and partial restriction map of the *puf* operon. Filled bars indicate the structural genes for the LHI and RC subunits, a newly identified open reading frame, Q, and an additional open reading frame (X) of unknown function. The lines below the map correspond to DNA fragments used as probes in S1 mapping and mRNA-capping experiments.

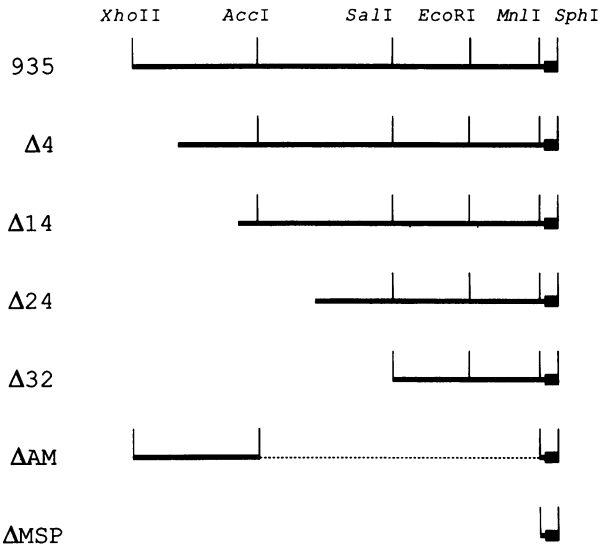

FIG. 3. Representation of deletion constructs used for *puf* promoter mapping. The designations assigned to each construct are given on the left. Restriction sites are given at the top and indicated by vertical lines, and the 5′ segment of the *pufB* gene is shown as a black box. The dashed line of the ΔAM fragment represents sequences that have been deleted.

the *Mnl*I-to-*Bam*HI fragment of pΔMSP and subsequently introduced into pXCA601.

Construction of oligonucleotide-directed mutants. The method used for site-directed mutagenesis was essentially that of Zoller and Smith (46). Single-stranded DNA from M13mp18Δ4, which contained the Δ4 deletion shown in Fig. 3 inserted between the *Hin*dIII and *Bam*HI sites of M13mp18, was used as the template. After annealing, extension and ligation, the DNA was introduced directly into *E. coli* JM101. Enrichment for mutations was achieved by use of *E. coli* RZ1032 (Ung^- Dut^-) for production of deoxyuracil-enriched template (25). The replicative forms of the M13 mutants were purified by standard techniques (29) and, after the entire fragment to be subcloned was sequenced, the mutants were digested with *Acc*I and *Pst*I to obtain the mutated regions, which were used to replace the unmutated *Pst*I-to-*Acc*I region of pJAJ21. The mutant derivatives of pJAJ21 were then digested with *Pst*I and *Bam*HI and subcloned into pXCA601. The Δ42 mutant was created with the oligonucleotide 5′-GAAGATTTATCTAGACGCTTCCTT-3′ so that the sequences 124 to 168 of Fig. 4 were replaced with 5′-AGA-3′ to generate an *Xba*I site. The Δ41 construct was created with the oligonucleotide 5′-CTTCCTTCTA-GACCCCCTTCAT-3′ so that sequences 179 to 208 were replaced with 5′-TAG-3′ to create an *Xba*I site, and the Δ44 construct was made with the oligonucleotide 5′-CATGGGTTGCGTGGGTAGCGTC-3′ so that the A residues at 225 and 227 of Fig. 4 were replaced by G residues.

RNA isolation. Cellular RNA was purified as described previously (36) from *R. capsulatus* B10 cells that had been grown under high-aeration conditions (7) to 80 Klett units (3.2×10^8 cells per ml) and then either shifted to low-aeration conditions (7) for 45 min (for capping experiments and high-resolution end mapping) or diluted to 20 Klett units and grown under low-aeration conditions for 8 h prior to cell harvest (for low-resolution end-mapping experiments). The same results were seen with either procedure.

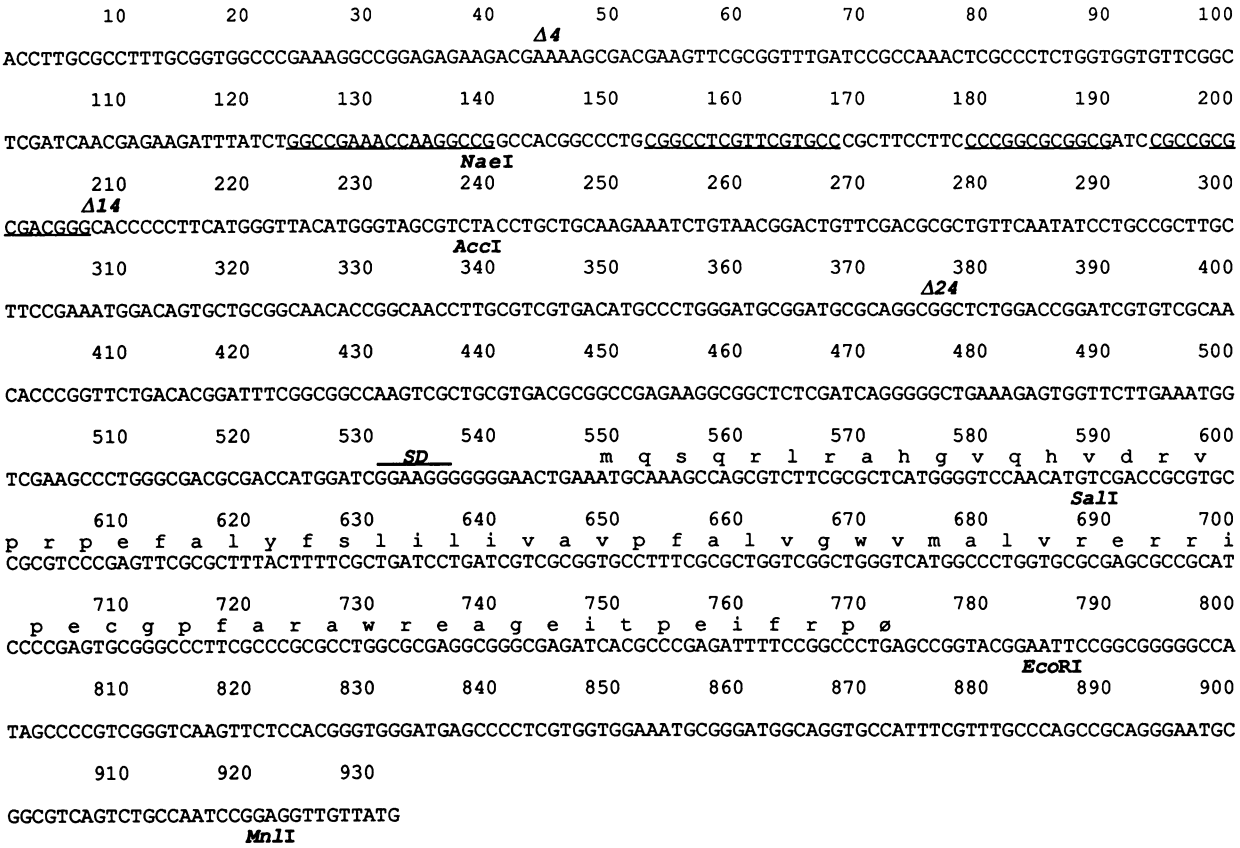

FIG. 4. Nucleotide sequence of the *puf* operon 5′ region. The amino acids corresponding to the translated sequence of the *pufQ* gene are indicated above the DNA sequence, as are a possible Shine-Dalgarno sequence (*SD*) and the end points of the Δ4, Δ14, and Δ24 deletions. The symmetrical sequences removed to create the Δ41 and Δ42 mutations are indicated by solid lines underneath the DNA sequence. The locations of the *Nae*I, *Acc*I, *Sal*I, *Eco*RI, and *Mnl*I restriction sites used in probe and deletion constructions are shown beneath the DNA sequence. The portion of the sequence from the *Eco*RI site to the *pufB* gene is from Youvan et al. (42). Although the sequence presented above differs from the one published by Bauer et al. (4) at positions 150, 333, and 355, an independently revised sequence obtained by the Marrs group is identical to this one (B. Marrs, personal communication).

S1 nuclease mapping of 5′ ends. Mapping of 5′ ends of mRNA was performed by the method of Berk and Sharp (8). Double-stranded DNA probes were end labeled by standard methods (26). Cellular RNA (6 μg) was hybridized with denatured end-labeled probe (10 ng) in the presence of carrier tRNA (5 μg) at 55°C overnight, and the hybrids were treated with S1 nuclease at 37°C for 3 h. Samples were then extracted with phenol and precipitated twice with ethanol to remove excess salt before electrophoresis under denaturing conditions. DNA sequence ladders were generated by the method of Maxam and Gilbert (28).

Primer extension. Primer extension mapping of 5′ RNA ends was performed essentially as described by Wilson et al. (40). Cellular RNA (10 μg) isolated from cells harvested after a 45-min shift to low-aeration conditions (7) was mixed with a 5′-end ^{32}P-labeled oligodeoxyribonucleotide primer and hybridized at 55°C for 3 h. After extension of the DNA strand by using Moloney murine leukemia virus reverse transcriptase (Bethesda Research Laboratories, Inc.), phenol extraction, and ethanol precipitation, the samples were denatured in formamide loading dye and electrophoresed on 6% polyacrylamide-urea gels with 0.5× Tris-borate-EDTA buffer (26).

Capping experiments. Cellular RNA was capped essentially as described previously (31, 47). RNA (40 μg) was mixed with 100 pmol of [α-^{32}P]GTP (3,000 Ci/mmol) and 12.5 U of guanylyltransferase (Bethesda Research Laboratories, Inc.) in buffer containing 25 mM Tris, pH 7.5, 2 mM $MgCl_2$, and 1 mM dithiothreitol. After incubation for 30 min at 37°C and phenol extraction, the RNA was precipitated three times by ethanol to remove unincorporated label and dissolved in RNA storage buffer (20 mM sodium phosphate [pH 6.5], 1 mM EDTA). An average of 4×10^4 dpm/μg of RNA was obtained in successful capping experiments.

Capped RNA was hybridized as for S1 experiments described above for 3 h to unlabeled DNA probes. After treatment with S1 nuclease (750 U), the samples were digested with 2.5 ng of boiled RNase A. The products were then denatured and electrophoresed as described above.

DNA sequencing and amino acid analysis. DNA sequencing was performed by use of both the Maxam and Gilbert (28) and Sanger dideoxy termination (3, 35) techniques. The DNA sequence data were analyzed by using a computer program obtained from A. Delaney (University of British Columbia, Vancouver, British Columbia, Canada). Alignments of amino acid sequences were done with the assistance of the computer program of Wilbur and Lipman (38).

RESULTS

DNA sequence of the *puf* operon 5′ to the *Eco*RI site. Northern (RNA) blot analysis of *puf* mRNA with DNA

TABLE 1. β-Galactosidase activities in extracts of cells containing *puf* 5′ segments fused to the *lacZ* gene of pXCA601

Construct[a]	β-Galactosidase activities[b] in cell extracts (mean [SD]) in:		Mean low O_2 values/mean high O_2 values
	Low O_2	High O_2	
935	3,181 (190)	494 (36)	6.4
ΔMSP	12 (5)	6 (9)	2.0
Δ4	2,092 (556)	327 (83)	6.4
Δ14	52 (5)	28 (2)	1.9
Δ24	64 (10)	29 (3)	2.2
Δ32	29 (5)	16 (2)	1.8
ΔAM	3,550 (771)	806 (209)	4.4
Δ42	1,983 (338)	431 (110)	4.6
Δ41	37 (8)	23 (4)	1.6
Δ44	111 (10)	33 (3)	3.4

[a] See Fig. 3.
[b] Activities are the means of three to seven independent assays and are expressed as nanomoles of *o*-nitrophenyl-β-D-galactopyranoside cleaved per min per mg of protein.

probes derived from the region encoding the LHI and RC genes has shown two mRNA species that share a common 5′ end and which map slightly downstream from the *Eco*RI site shown in Fig. 2. However, preliminary experiments (5) showed that the only promoter activity in the region is encoded by sequences located hundreds of base pairs upstream from the *Eco*RI site. As a step in understanding the role of this DNA segment in *puf* operon expression, the sequence of the DNA segment 5′ to the *Eco*RI site was determined.

One of the several notable features of the DNA sequence (Fig. 4) is an open reading frame capable of encoding a protein of 74 amino acids. This open reading frame is preceded by a sequence similar to the putative Shine-Dalgarno sequences of known *R. capsulatus* genes (13, 42, 43). Mutational analysis has indicated that the open reading frame encodes a biologically functional gene product necessary for the synthesis of bacteriochlorophyll-containing photosynthetic pigment-peptide complexes (3a; G. Klug, C. W. Adams, and S. N. Cohen, unpublished data, and M. E. Forrest, A. P. Zucconi, and J. T. Beatty, unpublished data). Because the region upstream from the *Eco*RI site is cotranscribed with the previously identified genes of the *puf* operon (see below), and because Marrs and co-workers have independently shown by gene fusion experiments that this open reading frame is actively translated (4), we have adopted their designation of *pufQ* for this gene.

Localization of the oxygen-regulated *puf* operon promoter. To determine the nature and location of the nucleotide sequences responsible for the induction or repression (or both) of *puf* operon expression, segments of DNA 5′ to the initiation codon for the B870β gene were joined to the promoterless *lacZ* indicator gene (Fig. 3). Each construction was assayed for β-galactosidase production in the *puf*-deleted *R. capsulatus* ΔRC6 under both high- and low-oxygen conditions (Table 1). Several constructions were also assayed in the wild-type host (B10), and equivalent results were obtained (data not shown). Control experiments utilized a segment (designated ΔMSP in Fig. 3) that lacks all *R. capsulatus* sequences 5′ to an *Mnl*I site that is located six nucleotides before the *pufB* ribosome-binding sequence (Fig. 4); the β-galactosidase activity enacted by this construct was taken as a baseline level.

As shown in Table 1, the ~1.2-kilobase (kb) *Xho*II-*Sph*I DNA segment contained in construct 935 resulted in *lacZ* expression in *R. capsulatus*. In addition, the amount of β-galactosidase activity in extracts of cells was reduced by aeration, suggesting that the transcription being assayed with this *lacZ* fusion plasmid was initiated at an O_2-regulated *puf* promoter. Deletion of approximately 180 nucleotides from the 5′ end of the 935 fragment, yielding the construct designated Δ4 in Fig. 3, had a minor effect on *lacZ* expression. However, removal of an additional 160 nucleotides, yielding the Δ14 fragment, caused about a 20-fold reduction in *lacZ* expression, down nearly to the baseline level obtained with the ΔMSP segment. Extension of the deletion to give the Δ24 construct, or to the *Sal*I site to give the Δ32 fragment, had little further effect on *lacZ* expression. Interestingly, an internal deletion of DNA sequences between the *Acc*I site and the *Mnl*I site shown in Fig. 4, yielding the ΔAM segment, resulted in fusions that retained both full activity and oxygen regulation, as had been observed with the entire 935 fragment. We conclude that sequences between the 5′ end of the DNA present in Δ4 and the *Acc*I site are both necessary and sufficient for oxygen-regulated initiation of transcription.

Examination of the DNA sequence within the transcriptional control region delimited by the Δ4 deletion and the *Acc*I site showed two inverted repeat segments of bidirectional rotational symmetry (underlined in Fig. 4). Replacement by oligonucleotide-directed mutagenesis of the more 5′ inverted-repeat sequence (i.e., from nucleotides 124 to 168 in Fig. 4) by the sequence 5′-AGA-3′ yielded the construct Δ42; the resulting *lacZ* fusion plasmid encoded β-galactosidase activity similarly high to that produced by the undeleted parental construct, Δ4 (Table 1). However, replacement of the proximal inverted repeat sequence (nucleotides 179 to 208 in Fig. 4) by the sequence 5′-TAG-3′, giving the Δ41 fragment, resulted in a reduction in *lacZ* expression to about the level observed for the control, ΔMSP.

The transcription initiation sites for *R. capsulatus puc* mRNAs (which encode the B800-850 light-harvesting peptides) map near a 7-nucleotide sequence, 5′-ACACTTG-3′, upstream of the *puc* structural genes (47). The *puf* operon DNA segment required for transcription includes the sequence 5′-ACATGGG-3′, centered at nucleotide 228 in Fig. 4. This sequence was changed to 5′-GCGTGGG-3′ by oligonucleotide-directed mutagenesis of the Δ4 construct, giving the Δ44 construct. Mutation of only two of the base pairs within this sequence reduced *lacZ* expression to only 5 to 10% of the level observed for the wild-type sequence (Table 1).

Low-resolution mapping of the 5′ end of *puf* transcripts. The results of the *lacZ* fusion experiments described above indicated that a locus required for O_2-regulated transcription of the LHI and RC genes is situated about 700 nucleotides upstream of the 5′ end of the *puf* mRNA species detected previously by using LHI or RC region probes. To investigate the nature of the transcript(s) initiated by this promoter, we analyzed the mRNA species protected from S1 nuclease digestion by a DNA fragment 5′ end labeled at the *Sph*I site shown in Fig. 2 and extending upstream to the *Xho*I site (Fig. 2, fragment A).

The most abundant RNA species protected by the *Xho*I-*Sph*I probe fragment correspond to the previously described *puf* transcripts encoding the LHI and RC gene products (Fig. 5A); the 5′ ends of these transcripts map to a position about 35 nucleotides downstream from the *Eco*RI site shown in Fig. 2 and 4. In addition, many other mRNA 5′ ends were observed when the *Xho*I-*Sph*I probe was employed; the

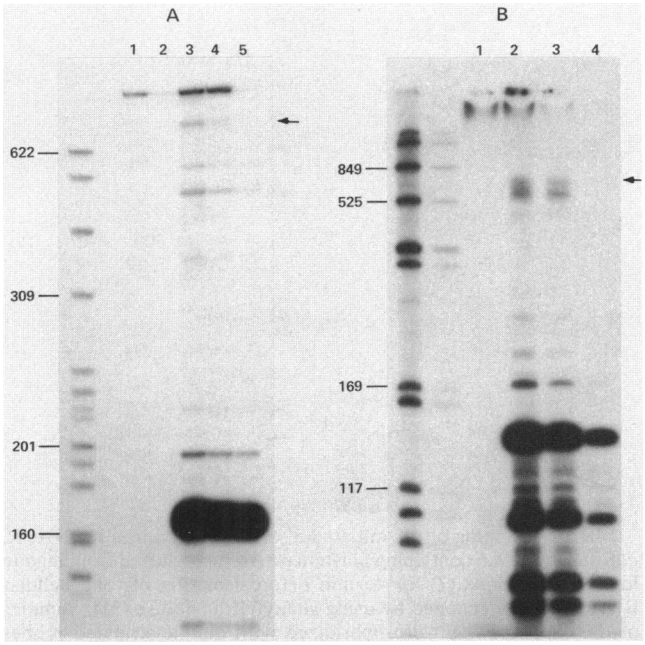

FIG. 5. Low-resolution S1 nuclease protection and primer extension mapping of the multiple 5′ ends of *puf* mRNA. (A) A double-stranded DNA probe (10 ng), 5′ end labeled at the *Sph*I site and extending to the *Xho*I site, was hybridized with 6 μg of total RNA extracted from *R. capsulatus* B10 (lanes 3, 4, and 5) or 6 μg of *E. coli* tRNA (lanes 1 and 2). Samples were treated with 300 (lane 3), 700 (lanes 1 and 4), or 1,400 (lanes 2 and 5) U of S1 nuclease. The top band represents reannealed probe. Arrow indicates the longest protected DNA fragment. The sizes, in nucleotides, of *Hpa*II fragments of pBR322 are given on the left. (B) An oligodeoxyribonucleotide primer end labeled by T4 polynucleotide kinase and having the sequence 5′-GAAGCTCAGGTCGTTCTTAT-3′ complementary to mRNA transcribed from bases 155 to 176 of the *puf* DNA sequence within the *pufB* gene as determined by Youvan et al. (42) was hybridized to 10 μg of *E. coli* tRNA (lane 1) or 10 μg of cellular RNA from strain B10 (lanes 2, 3, and 4). Experiments in lanes 2, 3, and 4 contained 1, 0.2, and 0.04 pmol of primer, respectively. Extension was carried out with Moloney murine leukemia virus reverse transcriptase. Molecular weight markers (far left lane) were derived from bacteriophage M13 single-stranded DNA digested with *Hae*III. Arrow indicates the longest extension product.

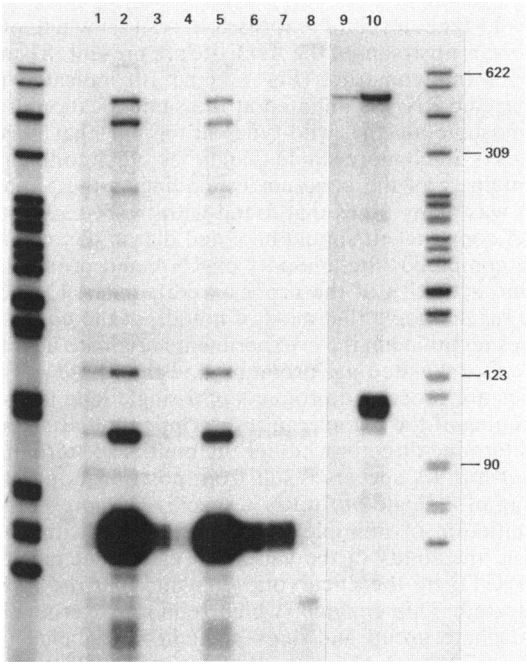

FIG. 6. S1 nuclease protection mapping of 5′ ends from mutant *puf* operon constructs. Total cellular RNA was extracted from *R. capsulatus* ΔRC6 (lane 8) or ΔRC6 bearing the following plasmids and grown under low-O_2 conditions. Lanes: 2, pΔ4; 3, pΔ14; 4, pΔ41; 5, pΔ42; 6, pΔ44; 7, pΔ32. Lanes 1 and 9 contain *E. coli* tRNA as a control. The probe in lanes 1 through 8 was fragment B (Fig. 2) derived from pΔ4; the probe in lanes 9 and 10 was fragment E (Fig. 2) derived from plasmid pΔAM; both probes were 5′ labeled at the *Bam*HI site in the polylinker region. All samples were treated with 700 U of S1 nuclease. An *Hpa*II digest of pBR322 DNA served as the molecular weight standard (flanking lanes); numbers at right indicate nucleotides.

three longest mRNA species had 5′ ends that mapped approximately 550, 360, and 300 nucleotides upstream from the location of the previously identified *puf* mRNA 5′ terminus. Since the labeled *Sph*I site of the DNA probe fragment used is located within the *pufB* gene, all of the transcripts containing the 5′ ends identified by this S1-mapping experiment must continue at least through the *Sph*I site in *pufB*. These results indicate that the 2.7-kb transcript, which is the longest *puf* mRNA species previously detected, results from processing of a *puf* operon transcript (or transcripts) initiated far upstream from the LHI and RC genes.

Primer extension studies (Fig. 5B) confirmed the presence of multiple 5′ ends for *puf* operon mRNA species extending through the *pufB* gene. In these experiments, a radioactively labeled synthetic oligonucleotide having a sequence homologous to a region in the *pufB* gene was hybridized to RNA isolated from *R. capsulatus* B10. After extension, the resulting DNA products were analyzed by electrophoresis on polyacrylamide gels and subsequent autoradiography. The multiple 5′ ends observed when a DNA primer that hybridized with *puf* mRNA in a region encoding the *pufB* gene was extended correspond to the 5′ ends detected by S1 analysis (Fig. 5A).

To determine the effect of removal of the putative *puf* promoter region on the occurrence of these various *puf* transcript termini, we analyzed the mRNA species synthesized by mutants having deletions at different locations upstream of the *Eco*RI site. For these experiments, total RNA was extracted from *R. capsulatus* ΔRC6 containing the mutant plasmids. Since ΔRC6 has a chromosomal deletion for the region between the *Sal*I site of the *pufQ* gene (Fig. 2) and the *Xho*II site downstream of the *pufX* gene (42), radioactively labeled probes (Fig. 2) consisting of plasmid-borne sequences derived from within the *Sph*I-*Xho*I segment detect only *puf* transcripts encoded by the plasmid.

The bands resulting from protection of labeled probes by transcript species derived from either the B10 chromosome or the Δ4 plasmid construct were indistinguishable (Fig. 5A and 6). In addition, removal of the most upstream inverted repeat, as in Δ42, also had no effect on the 5′ ends seen. However, deletion of the more 3′ of the two inverted repeats in the putative promoter region (i.e., the Δ41 construct) or deletion of *puf* sequences upstream from the 5′ end of the Δ14 construct resulted in disappearance not only of the most upstream of the 5′ termini identified for *puf* mRNA but also of almost all of the more downstream 5′ mRNA ends as well. Measurement of β-galactosidase activity in strains containing the ΔAM construction, in which sequences between the *Acc*I site and the *Mnl*I site had been removed, indicated that

the wild-type level of expression results when only the sequences upstream of the *Acc*I site are present. S1-mapping data for this construct (Fig. 6, lane 10) indicated that its *puf*-specific RNA is initiated at the same location as on the chromosome of the wild-type strain but that this RNA species is much more stable than in the other configurations; the intensity of the band corresponding to the 5′ end from ΔAM was many times that found with the 5′-most end from the Δ4 construct. It should be noted that in the experiments being compared, the amounts of RNA and probes and the specific activities of the probes were similar. Collectively, these results imply that most, if not all, of the *puf* transcript species identified in these experiments originate near the site of the O_2-regulated *puf* promoter and that these *puf* mRNA species are processed products of a single long transcript.

Capping of the *puf* operon transcription initiation product. To determine directly whether the multiple 5′ termini found for *puf* mRNA species result from posttranscriptional processing of a single *puf* mRNA species or represent instead the initiation of multiple transcript species within the *puf* operon, the ability of the various 5′ ends to be labeled with [α-^{32}P]GTP by the eucaryotic enzyme guanylyltransferase was tested. This enzyme, which requires a terminal di- or triphosphate group and does not add the labeled triphosphate to RNA termini generated by endonucleolytic cleavage, can thus recognize and cap only initiated 5′ ends of transcripts and not those resulting from processing (31). Initial experiments performed with RNA isolated from B10 plus a probe containing either the upstream region encompassing the three most 5′ ends or the downstream region that includes the locus of the most prominent 5′ end of *puf* mRNA showed no capped *puf* mRNA molecules of any length, even though capping was observed for *puc* operon mRNA included as a positive control (data not shown). However, when RNA from cells containing the ΔAM construction was treated with guanylyltransferase and was then hybridized to a 204-nucleotide DNA fragment homologous to the region extending from the *Nae*I site to the *Bam*HI site of plasmid pΔAM (Fig. 2, fragment D), a single, protected, 5′-end-labeled mRNA species was observed after treatment with S1 nuclease and RNase A (Fig. 7). The protected segment seen was about 115 nucleotides in size, thus locating this RNA terminus at or near the position of the most upstream of the 5′ ends identified by S1 mapping and primer extension analysis (i.e., the 5′ end adjacent to the promoter region).

High-resolution S1 mapping and primer extension. To map more precisely the transcription initiation site identified for *puf* mRNA, primer extension and S1-mapping experiments were carried out with RNA isolated from *R. capsulatus* ΔRC6 containing plasmid pΔAM and the results were analyzed by using sequencing gels. Primer extension experiments (Fig. 8A) identified the 5′ end of the *puf* mRNA as a G corresponding to the G at position 236 of Fig. 4, with a second 5′ end at the adjacent T. S1-mapping experiments (Fig. 8B) revealed a cluster of bands spanning the region AGCGTC (Fig. 4, nucleotides 233 to 238). The most intense band corresponded to the G identified by primer extension studies. This location is 5 nucleotides downstream from the *puf* transcript 5′ end reported by Bauer et al. (4) and 694 nucleotides upstream of the first codon of the *pufB* gene.

DISCUSSION

Previous work has shown that the polycistronic *puf* transcript of *R. capsulatus* includes a 0.5-kb mRNA segment,

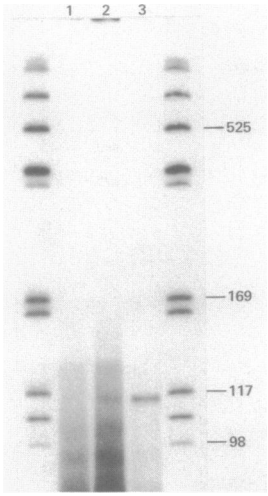

FIG. 7. Capping of the initiated 5′ end of *puf* operon mRNA. *R. capsulatus* ΔRC6 containing plasmid pΔAM was shifted from high to low O_2 conditions (7) for 45 min before isolation of total cellular RNA. RNA was capped by using guanylyltransferase (31). Samples of the capped RNA were hybridized with double-stranded probes and treated with S1 nuclease and RNase A. Lanes: 1, 10 μg of cellular RNA with no DNA probe; 2, 10 μg of cellular RNA and 40 ng of fragment C (Fig. 2) as the probe; 3, 10 μg of cellular RNA and 20 ng of fragment D (Fig. 2). Molecular weight markers (flanking lanes) were derived from an *Hae*III digest of bacteriophage M13 single-stranded DNA; numbers at right indicate nucleotides.

which encodes the LHI α and β subunits, and a 2.7-kb mRNA segment that includes the RC genes as well as the LHI genes (7, 10, 23). These *puf* mRNA species have common 5′ ends that map to a location about 35 base pairs downstream from the solitary *Eco*RI site in the operon. While the 0.5- and 2.7-kb *puf* mRNA molecules are the most abundant transcript species, S1 analysis with end-labeled probes that extend from within the *pufB* gene to the *Xho*I site (Fig. 2) has now revealed additional 5′ mRNA termini that map to various locations within the *puf* operon segment upstream from the *Eco*RI site, which was the 5′ terminus of the previously employed probes (7). The multiple 5′ termini observed in the region upstream from the *Eco*RI site mapped to the same locations in repeat experiments, and identical results were seen whether the *puf* RNA was isolated from the wild-type *R. capsulatus* strain B10 expressing the *puf* operon from the chromosome or from a strain of ΔRC6 which harbors a plasmid containing *puf* sequences.

The most 5′ *puf* mRNA terminus identified mapped near the hairpin loop region that contains the *puf* operon promoter. The multiple *puf* species thus appear to result from processing of a primary transcript at least 3.4 kb in length. The *puf* mRNA species having 5′ ends upstream from the *Eco*RI site are much less abundant than the more downstream *puf* transcript segments described previously. Earlier work has shown that differential stability of these downstream *puf* mRNA segments accounts for differential expression of the LHI and RC genes encoded by the polycistronic *puf* operon (7, 10, 23). The results reported here suggest that differential stability of different mRNA segments may be a prominent mechanism by which regulation of expression of genes within the entire *puf* operon is accomplished.

Sequences 5′ of the *Acc*I site (located about 540 nucleotides upstream of the *puf Eco*RI site) are both necessary and sufficient for the initiation of O_2-regulated transcription,

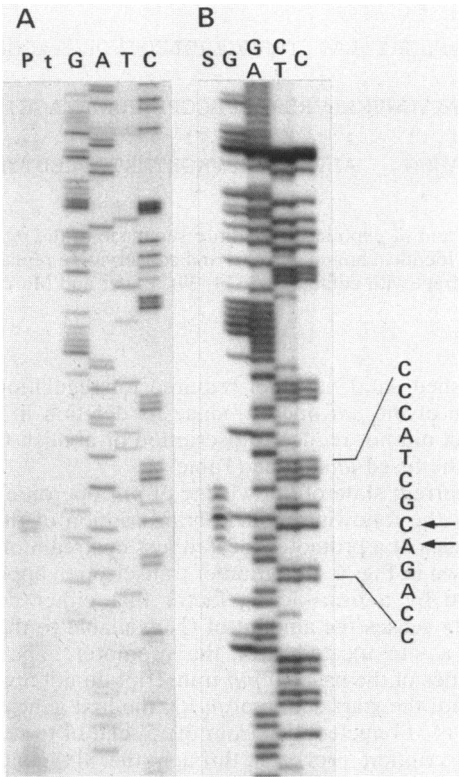

FIG. 8. Identification of the 5'-end nucleotide of the *puf* mRNA. (A) Primer extension. A primer with the sequence 5'-AATTC CAGTTCTTGCGCTTG-3' (complementary to mRNA transcribed from nucleotide 193 of the sequence from Youvan et al. [42] located in the *pufB* gene to the *Eco*RI site of the pUC polylinker region of pΔAM) was hybridized to RNA isolated from *R. capsulatus* ΔRC6 containing the plasmid pΔAM (lane P) or *E. coli* tRNA (lane t). After extension, products were electrophoresed beside a sequencing ladder obtained by using the same primer and double-stranded pUCΔAM DNA as a template in the Sanger dideoxy termination (35) technique. (B) S1 nuclease protection. A 0.54-kb DNA fragment extending from the *Bam*HI site in the polylinker region of the ΔAM construction to the *Xho*I site located upstream of the *puf* promoter region and labeled at the *Bam*HI site (Fig. 2, fragment E) was used as a probe. A portion of this labeled probe was hybridized to RNA from *R. capsulatus* ΔRC6 containing plasmid pΔAM and then treated with S1 nuclease (lane S), and part was cleaved chemically according to the sequencing technique of Maxam and Gilbert (28). The sequence notation on the right corresponds to the sequence of the coding strand, which is complementary to the DNA sequence shown in Fig. 4. Arrows indicate nucleotides most likely to correspond to the 5' end of the *puf* transcript.

whereas no significant promoter activity was detected in the region between the *Acc*I site and the protein-coding sequences of the *puf* operon. More precise localization of the *puf* promoter was accomplished by analysis of the effects of oligonucleotide-directed mutations, which implicated sequences between nucleotides 179 and 227 in Fig. 4 in the initiation of transcription. No significant homology with an *E. coli* consensus promoter (18) was detected within this region, and expression of the fused *lacZ* gene from this promoter did not occur in *E. coli* cells (unpublished data). Because of its homology with sequences located near *R. capsulatus puc* transcription start sites (47) and because of the results obtained with the Δ44 mutation (Table 1, Fig. 6), we suggest that the AT-rich sequence 5'-TTACAT-3', starting at position 223 in Fig. 4, functions as part of the *puf* operon promoter. A second component region of the promoter may be located in the vicinity of nucleotides 203 to 208 in Fig. 4. This notion is supported by the finding that the Δ14 deletion, which does not show significant promoter activity (Table 1), is missing all *R. capsulatus* sequences 5' of the cytidine nucleotide at position 210 in Fig. 4. While the *R. capsulatus* RNA polymerase seems similar in subunit composition to other procaryotic RNA polymerases (17), there is no evidence that it requires the same components and nucleotide spacing found in typical *E. coli* promoters (18).

Our data indicate that the sequences involved in O_2 regulation of *puf* operon expression are located very near the *puf* promoter sequence. It is possible that the region of bidirectional rotational symmetry extending from nucleotides 179 to 207 in Fig. 4 functions as an operator to bind proteins that regulate *puf* transcription initiation in response to changes in O_2 levels. In work published during preparation of this report, Bauer et al. (4) also identified an O_2-regulated *puf* operon promoter in this region. In addition, these authors have proposed the existence of a constitutively active second *puf* promoter that functions at approximately 3 to 4% of the induced level of activity found for the O_2-regulated promoter. While we also observed residual *lacZ* expression after deletion of the O_2-regulated *puf* promoter, the physiological significance of such expression is unclear, as the residual transcription was only 3% of the transcription obtained with the 935 and Δ4 constructs in cells grown under low aeration (Table 1). The mRNA 5' end associated by Bauer et al. with this putative second promoter is only one of several 5' ends that we attribute to cleavage of the primary transcript.

While the 5' end located 35 nucleotides downstream of the *Eco*RI site is many times more abundant than the 5' ends located further upstream, we were repeatedly unable to cap this more abundant species by using the enzyme guanylyltransferase under conditions in which controls indicated that capping of mRNA from another *R. capsulatus* operon occurred. However, addition of [α-^{32}P]GTP to the 5'-most end (Fig. 5A) occurred when DNA sequences corresponding to an mRNA segment that contained the multiple downstream 5' mRNA termini was deleted (i.e., the ΔAM construct) (Fig. 7). The ability to cap the most upstream 5' terminus of *puf* mRNA in the ΔAM construct correlated with an increased intensity of the band representing this transcript terminus on S1 analysis, consistent with the view that the deletion removed downstream processing sites and resulted in an

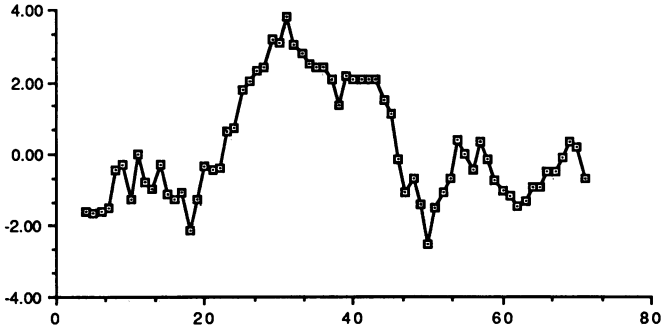

FIG. 9. Hydropathy plot of the amino acid sequence of the *pufQ* gene product. The amino acid sequence (numbered on the horizontal axis) was analyzed with the Kyte and Doolittle (15) algorithm by using a moving window of seven residues. Positive values indicate hydrophobic regions, and negative values indicate hydrophilic regions.

```
          180                 *                 *                 252
RCM       SHLDWTNQFSLDHGNLFYNPFHGLSIAALYGSALLFAMHGATILAV TRFGGERELEQIVDRGTASERAALFWR
          . .. .. ...: .  .: .: .. .  :. ::. : ..  : .:   : .  . . :.... ...:..
pufQ gp   MQSQRLRAHGVQHVDRVPRPEFALYFSLILIVAVPFALVGWVMALVRERRIPECGPFARAWREAGEITPEIFRP
          . ..... : . .  .:  : .::....: ..:. :  ::: .  : :  :. .... ....
RCL       THLDWVSNTGYTYGNFHYNPFHMLGISLFFTTAWALAMHG   ALVLSAANPVKGKTMRTPDHEDTYFRDLMGY
          153                 *                 *                 223
```

FIG. 10. Alignment of the *pufQ* amino acid sequence with segments of RC L and M peptides. The entire sequence of the *pufQ* product is given; the numbers denote the first and last residues of the L and M segments. Identical amino acids (:) and conservative replacements (·) are indicated, as are the histidines (*) of the RC peptides that are believed to interact with cofactors (2, 14, 39). Bauer and Marrs (3a) have recently published alignments similar to these.

increase in uncleaved *puf* mRNA. In contrast to the discrete decay intermediates revealed by S1 5′-end mapping of *puf* mRNA, attempts to map the 3′ ends of stable processing intermediates derived from the region upstream of the *pufB* gene yielded a smear or a ladder of nonspecific bands (data not shown). This result is consistent with the view that endonucleolytic cleavages within the *puf* mRNA segment encoded by the region upstream from the *Eco*RI site are followed by rapid 3′-to-5′ degradation within this segment. Taken together, our data suggest that the rate-limiting step in the degradation of *puf* mRNA is endonucleolytic cleavage between the nucleotide at the point of initiation of the transcript and the 5′ terminus corresponding to an mRNA location near the *Eco*RI site of the *puf* operon; a segment encoding this particularly labile region of *puf* mRNA appears to have been removed by the ΔAM deletion. These results indicate that the transcription initiation site for the *puf* operon is located in the region of DNA around nucleotide 236 in Fig. 4 and that the multiple *puf* mRNA species, including the one mapping immediately downstream from the *Eco*RI site, are probably processed products of transcripts originating at this furthest upstream location. Computer analysis showed several regions of possible secondary structure in the most 5′ region of the *puf* operon. It is unclear whether these play a role in processing of the rapidly cleaved segment of the *puf* transcript.

A hydropathy plot of the amino acid sequence of the putative *pufQ* gene product (Fig. 9) shows a central hydrophobic stretch of 21 amino acids and an average hydropathy value of 0.2 for the entire sequence. The products encoded by previously described genes of the *puf* operon are hydrophobic peptides, and homology has been noted between the *pufA* and *pufB* peptides and between the *pufL* and *pufM* gene products (42, 43). The predicted amino acid sequence of *pufQ* showed significant alignment (Fig. 10) with only the *pufL* and *pufM* gene products (the L and M peptides of the RC). These RC peptides are in turn homologous to the *R. sphaeroides* RC peptides, which have been shown to be involved in binding of quinone, iron, and bacteriochlorophyll pigments (2, 14, 39). However, the putative *pufQ* gene product (Fig. 10) lacks two histidines that have been implicated in the binding of RC peptides to cofactors (2, 14, 39). The relatively low abundance of the *pufQ*-encoding mRNA segments compared with that of the more stable LHI and RC-encoding mRNAs suggests that the *pufQ* gene product may be present at a lower concentration than the LHI and RC peptides. If this speculation is correct, the *pufQ* gene product may have a catalytic rather than a structural function. Although the *pufQ* gene is necessary for formation of mature photosynthetic pigment-peptide complexes (3a; G. Klug, C. W. Adams, and S. N. Cohen, unpublished data, and M. E. Forrest, A. P. Zucconi, and J. T. Beatty, unpublished data), it is not required for regulation of transcription of the *puf* operon, since its deletion in the ΔAM construct did not reduce transcription or abolish O_2 regulation of the fused *lacZ* gene (Table 1).

Our current state of knowledge of *puf* operon expression suggests the following model. Transcription of the *puf* operon begins at a promoter located just upstream of the *Acc*I site shown in Fig. 4. Initiation of transcription appears to be regulated by a *trans*-acting factor that either directly or indirectly senses the amount of O_2 available to the cell and acts at a site located near the promoter. The first 313 nucleotides of the nascent *puf* transcript do not appear to be translated; the start codon of *pufQ*, the first gene of the *puf* operon, is 313 nucleotides from the 5′ end of the transcript. As transcription proceeds through the six genes of the operon, and possibly before reaching the transcriptional terminators near the 3′ end of the *puf* transcriptional unit (7), cleavage and degradation of the highly labile 5′ segment begins. We speculate that the half-life of this relatively unstable portion of the transcript is on the order of tens of seconds. Following completion of transcription of the entire operon, degradation of the 3′-end segment begins, although at a slower rate than the region upstream from the *pufB* gene, so that its half-life is about 5 min (7, 10, 23). The most stable and therefore most abundant *puf* mRNA segment, which encodes the LHI peptides, persists after degradation of segments upstream and downstream from the LHI region. The proposed sequence of events would have the net effect of allowing differential expression of genes encoded by a polycistronic message over a very wide range, the limits of which would be influenced by different rates of decay of different segments of *puf* mRNA.

ACKNOWLEDGMENTS

We thank W. T. Tucker for providing plasmid pWTT552, the source of the T4 phage translation-transcription termination signals; D. Brenner for plasmid p*Plac-omp*, a derivative of plasmid pMC1403*lac*; T. J. Schmidhauser for plasmid pTJS133; and B. Marrs for exchange of sequence information prior to publication. We thank G. Klug for first making the observation that mutations in the region identified here as *pufQ* affect the formation of mature LH and RC complexes.

This research was supported by grants from the Canadian NSERC (A-2796 to J.T.B.) and by Public Health Service grant GM27241 to S.N.C. from the National Institutes of Health. C.W.A. was supported by National Institutes of Health training grant GM07790.

LITERATURE CITED

1. **Aliabadi, Z., F. Warren, S. Mya, and J. W. Foster.** 1986. Oxygen-regulated stimulons of *Salmonella typhimurium* identified by Mu d(Aplac) operon fusion. J. Bacteriol. **165**:780–786.
2. **Allen, J. P., G. Feher, T. O. Yeates, D. C. Rees, J. Deisenhofer, H. Michel, and R. Huber.** 1986. Structural homology of reaction

centers from *Rhodopseudomonas sphaeroides* and *Rhodopseudomonas viridis* as determined by X-ray diffraction. Proc. Natl. Acad. Sci. USA **83**:8589–8593.

3. **Barr, P. J., R. M. Thayer, P. Laybourn, R. C. Najarian, F. Seela, and D. R. Tolan.** 1986. 7-Deaza-2′-deoxyguanosine-5′-triphosphate: enhanced resolution in M13 dideoxy sequencing. Biotechniques **4**:428–432.

3a. **Bauer, C. E., and B. L. Marrs.** 1988. *Rhodobacter capsulatus puf* operon encodes a regulatory protein (PufQ) for bacteriochlorophyll biosynthesis. Proc. Natl. Acad. Sci. USA **85**: 7074–7078.

4. **Bauer, C. E., D. A. Young, and B. L. Marrs.** 1988. Analysis of the *Rhodobacter capsulatus puf* operon. Location of the oxygen-regulated promoter region and the identification of an additional *puf*-encoded gene. J. Biol. Chem. **263**:4820–4827.

5. **Beatty, J. T., C. W. Adams, and S. N. Cohen.** 1986. Regulation of expression of the *rxcA* operon of *Rhodopseudomonas capsulata*, p. 27–29. *In* D. C. Youvan and F. Daldal (ed.), Microbial energy transduction. Cold Spring Harbor Laboratory, Cold Spring Harbor, N.Y.

6. **Beck, E., and E. Bremer.** 1980. Nucleotide sequence of the gene *ompA* coding for the outer membrane protein II* of *Escherichia coli* K-12. Nucleic Acids Res. **8**:3011–3024.

7. **Belasco, J. G., J. T. Beatty, C. W. Adams, A. von Gabain, and S. N. Cohen.** 1985. Differential expression of photosynthesis genes in *R. capsulata* results from segmental differences in stability within the polycistronic *rxcA* transcript. Cell **40**:171–181.

8. **Berk, A. J., and P. A. Sharp.** 1977. Sizing and mapping of early adenovirus mRNAs by gel electrophoresis of S1 endonuclease digested hybrids. Cell **12**:721–723.

9. **Cannon, M., S. Hill, E. Kavanaugh, and F. Cannon.** 1985. A molecular genetic study of *nif* expression in *Klebsiella pneumoniae* at the level of transcription, translation and nitrogenase activity. Mol. Gen. Genet. **198**:198–205.

10. **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.

11. **Clark, W. G., E. Davidson, and B. L. Marrs.** 1984. Variation of levels of mRNA coding for antenna and reaction center polypeptides in *Rhodopseudomonas capsulata* in response to changes in oxygen concentration. J. Bacteriol. **157**:945–948.

12. **Cohen-Bazire, G., W. R. Sistrom, and R. Y. Stanier.** 1957. Kinetic studies of pigment synthesis by non-sulfur purple bacteria. J. Cell. Comp. Physiol. **49**:25–68.

13. **Davidson, E., and F. Daldal.** 1987. Primary structure of the bc_1 complex of *Rhodopseudomonas capsulata*. J. Mol. Biol. **195**: 13–24.

14. **Deisenhofer, J., O. Epp, K. Miki, R. Huber, and H. Michel.** 1985. Structure of the protein subunits in the photosynthetic reaction centre of *Rhodopseudomonas viridis* at 3 Å resolution. Nature **318**:618–624.

15. **Doolittle, R.** 1986. Of URFS and ORFS. A primer on how to analyze derived amino acid sequences. University Science Books, Mill Valley, Calif.

16. **Drews, G., and J. Oelze.** 1981. Organization and differentiation of membranes of phototrophic bacteria. Adv. Microb. Physiol. **22**:1–92.

17. **Forrest, M. E., and J. T. Beatty.** 1987. Purification of *Rhodobacter capsulatus* RNA polymerase and its use for *in vitro* transcription. FEBS Lett. **212**:28–34.

18. **Hawley, D. K., and W. R. McClure.** 1983. Compilation and analysis of *Escherichia coli* promoter DNA sequences. Nucleic Acids Res. **11**:2237–2255.

19. **Henikoff, S.** 1984. Unidirectional digestion with exonuclease III creates targeted breakpoints for DNA sequencing. Gene **28**:351–359.

20. **Hui, A., J. Hayflick, K. Dinkelspiel, and H. A. de Boer.** 1984. Mutagenesis of the three bases preceding the start codon of the β-galactosidase mRNA and its effects on translation in *Escherichia coli*. EMBO J. **3**:623–629.

21. **Johnson, J. A., W. K. R. Wong, and J. T. Beatty.** 1986. Expression of cellulase genes in *Rhodobacter capsulatus* by use of plasmid expression vectors. J. Bacteriol. **167**:604–610.

22. **Kaplan, S., and B. L. Marrs.** 1986. Proposed nomenclature for photosynthetic procaryotes. ASM News **52**:242.

23. **Klug, G., C. W. Adams, J. Belasco, B. Doerge, and S. N. Cohen.** 1987. Biological consequences of segmental alterations in mRNA stability: effects of deletion of the intercistronic hairpin loop region of the *Rhodobacter capsulatus puf* operon. EMBO J. **6**:3515–3520.

24. **Kranz, R. G., and R. Haselkorn.** 1986. Anaerobic regulation of nitrogen fixation genes in *Rhodopseudomonas capsulata*. Proc. Natl. Acad. Sci. USA **83**:6805–6809.

25. **Kunkel, T. A.** 1985. Rapid and efficient site-specific mutagenesis without phenotypic selection. Proc. Natl. Acad. Sci. USA **82**: 488–492.

26. **Maniatis, T., E. F. Fritsch, and J. Sambrook.** 1982. Molecular cloning: a laboratory manual. Cold Spring Harbor Laboratory, Cold Spring Harbor, N.Y.

27. **Marrs, B.** 1974. Genetic recombination in *R. capsulata*. Proc. Natl. Acad. Sci. USA **71**:971–973.

28. **Maxam, A. M., and W. Gilbert.** 1980. Sequencing end-labeled DNA with base-specific chemical cleavages. Methods Enzymol. **65**:499–650.

29. **Messing, J.** 1983. New M13 vectors for cloning. Methods Enzymol. **101**:20–78.

30. **Miller, J. H.** 1972. Experiments in molecular genetics, p. 352–355. Cold Spring Harbor Laboratory, Cold Spring Harbor, N.Y.

31. **Moss, B.** 1981. 5′-End labelling of RNA with capping and methylating enzymes, p. 253–266. *In* J. G. Chirikjian and T. S. Papas (ed.), Gene amplification and analysis, vol. 2. Elsevier/North-Holland Publishing Co., Amsterdam.

32. **Prentki, P., and H. M. Krisch.** 1984. *In vitro* insertional mutagenesis with a selectable DNA fragment. Gene **29**:303–313.

33. **Schmidhauser, T. J., and D. R. Helinski.** 1985. Regions of broad-host range plasmid RK2 involved in replication and stable maintenance in nine species of gram-negative bacteria. J. Bacteriol. **164**:446–455.

34. **Shaw, D. J., D. W. Rice, and J. R. Guest.** 1983. Homology between CAP and Fnr, a regulator of anaerobic respiration in *Escherichia coli*. J. Mol. Biol. **166**:241–247.

35. **Smith, A. J. H.** 1980. DNA sequence analysis by primed synthesis. Methods Enzymol. **65**:560–580.

36. **von Gabain, A., J. G. Belasco, J. L. Schottel, A. C. Y. Chang, and S. N. Cohen.** 1983. Decay of mRNA in *Escherichia coli*: investigation of the fate of specific segments of transcripts. Proc. Natl. Acad. Sci. USA **80**:653–657.

37. **Weaver, P. F., J. D. Wall, and H. Gest.** 1975. Characterization of *Rhodopseudomonas capsulata*. Arch. Microbiol. **105**:207–216.

38. **Wilbur, W. J., and D. J. Lipman.** 1983. Rapid similarity searches of nucleic acid and protein data banks. Proc. Natl. Acad. Sci. USA **80**:726–730.

39. **Williams, J. C., L. A. Steiner, and G. Feher.** 1986. Primary structure of the reaction center from *Rhodopseudomonas sphaeroides*. Proteins **1**:312–325.

40. **Wilson, H. R., P. T. Chan, and L. L. Turnbough, Jr.** 1987. Nucleotide sequence and expression of the *pyrC* gene of *Escherichia coli* K-12. J. Bacteriol. **169**:3051–3058.

41. **Winkelman, J. W., and D. P. Clark.** 1986. Anaerobically induced genes of *Escherichia coli*. J. Bacteriol. **167**:326–367.

42. **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.

43. **Youvan, D. C., and S. Ismail.** 1985. Light harvesting II (B800-850 complex) structural genes from *Rhodopseudomonas capsulata*. Proc. Natl. Acad. Sci. USA **82**:58–62.

44. **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 bacteriochlo-

rophyll and carotenoid biosynthesis in *Rhodobacter capsulatus* during the shift from anaerobic to aerobic growth. J. Bacteriol. **168**:1180–1188.

45. **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.
46. **Zoller, M. J., and M. Smith.** 1983. Oligonucleotide-directed mutagenesis of DNA fragments cloned into M13 vectors. Methods Enzymol. **100**:468–500.
47. **Zucconi, A. P., and J. T. Beatty.** 1988. Posttranscriptional regulation by light of the steady-state levels of mature B800-850 light-harvesting complexes in *Rhodobacter capsulatus*. J. Bacteriol. **170**:877–882.