Title

Organization of the {ital Rhodobacter Capsulatus} Carotenoid Biosynthesis Gene Cluster

Authors

Armstrong, G.A.
Alberti, M.
Leach, F.
et al.

Publication Date

1990

LBL-29677
Preprint

Lawrence Berkeley Laboratory
UNIVERSITY OF CALIFORNIA

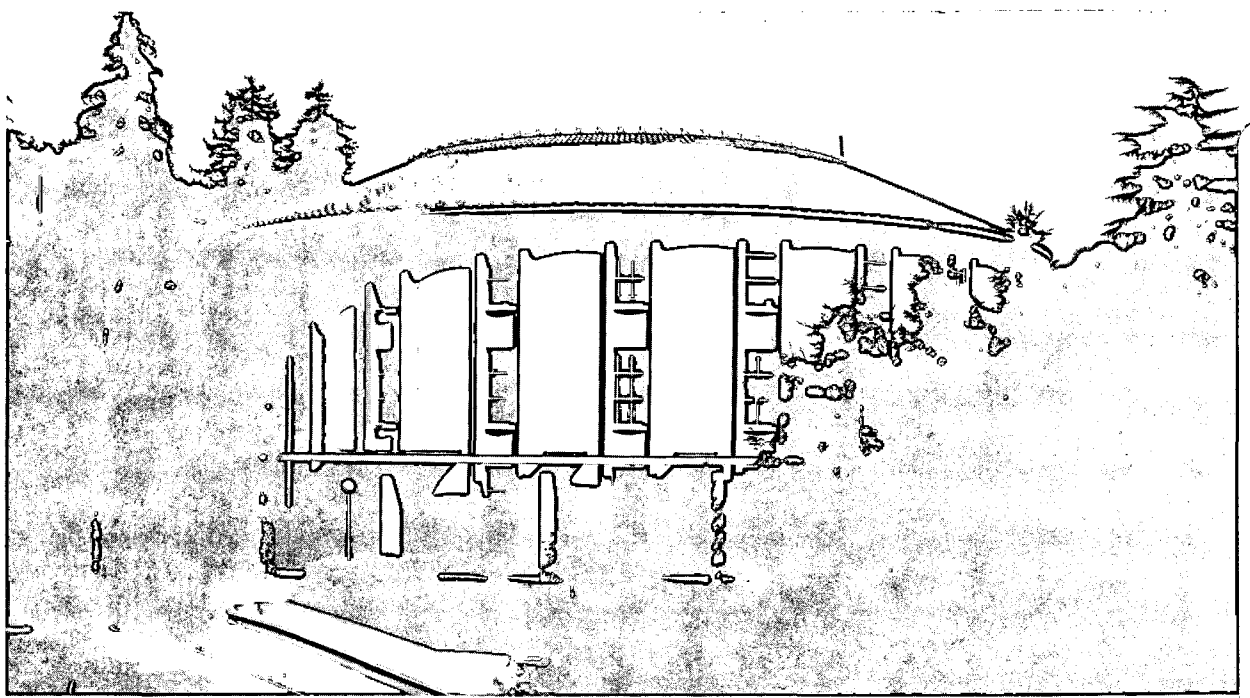
CHEMICAL BIODYNAMICS DIVISION

To be published as a chapter in **Molecular Biology of Membrane-Bound Complexes in Phototrophic Bacteria**, G. Drews and E.A. Dawes, Eds., Plenum Press, New York and London, 1990

Organization of the *Rhodobacter Capsulatus* Carotenoid Biosynthesis Gene Cluster

G.A. Armstrong, M. Alberti, F. Leach, and J.E. Hearst

August 1990



Prepared for the U.S. Department of Energy under Contract Number DE-AC03-76SF000

DISCLAIMER

LBL-29677

ORGANIZATION OF THE *RHODOBACTER CAPSULATUS* CAROTENOID BIOSYNTHESIS GENE CLUSTER

Gregory A. Armstrong, Marie Alberti, Francesca Leach and John E. Hearst

Lawrence Berkeley Laboratory

University of California

Berkeley, CA 94720

This work was supported by the Office of Basic Energy Science, Biological Energy Division, Department of Energy Contract DE-AC03O-76S00098

ORGANIZATION OF THE *RHODOBACTER CAPSULATUS* CAROTENOID BIOSYNTHESIS GENE CLUSTER

[1,2]Gregory A. Armstrong, [2]Marie Alberti, [2]Francesca Leach and [1,2]John E. Hearst

[1]Department of Chemistry, University of California and
[2]Division of Chemical Biodynamics, Lawrence Berkeley Laboratory
Berkeley, CA 94720, USA

INTRODUCTION

Carotenoids are a major class of pigment molecules found in all photosynthetic organisms, and some nonphotosynthetic bacteria, fungi and yeasts (reviewed in Goodwin, 1980). In photosynthetic organisms carotenoids are not only essential physical quenchers of excited state triplet chlorophyll and bacteriochlorophyll (Bchl) and of singlet oxygen generated by these species, but also serve as accessory light-harvesting pigments (reviewed in Cogdell and Frank, 1987). The isolation of the R-prime plasmid pRPS404, containing a 46 kb region from the *Rhodobacter capsulatus* chromosome which complemented all known point mutation defects in photosynthesis, suggested that the genes encoding structural photosynthetic polypeptides and the enzymes of carotenoid and bacteriochlorophyll biosynthesis were clustered (Marrs, 1981). The genes encoding the reaction center and light-harvesting I polypeptides, flanking the pigment biosynthesis genes, were subsequently located and sequenced (Youvan et al., 1984a), as were the unlinked genes encoding the light-harvesting II antenna polypeptides (Youvan and Ismail, 1985). No DNA sequences were previously available for the genes encoding carotenoid biosynthetic enzymes from any carotenogenic organism. Thus, the determination of the nucleotide sequence and the organization of the *crt* genes from *R. capsulatus* is essential both to further studies of the gene products and of gene regulation. We have focused our attention on the characterization of the subcluster of *crt* genes within the photosynthesis gene cluster (for a description of the carotenoid biosynthesis pathway see Armstrong et al., 1989). Seven of the eight previously identified *R. capsulatus crt* genes were known to be clustered on the *BamH*I-H, -G, -M, and -J fragments of pRPS404 in the order *crtA, I, B, C, D, E, F* from left to right on the genetic-physical map (Fig. 1) (Taylor et al., 1983; Zsebo and Hearst, 1984; Giuliano et al., 1988). These studies established that mutations causing Bchl$^-$ phenotypes map within these four *BamH*I fragments, flanking both ends of the *crt* gene cluster. We have determined the nucleotide sequence of an 11039 bp region encompassing the *BamH*I-H, -G, -M, and -J fragments of pRPS404 (Armstrong et al., 1989). The nucleotide sequence reveals the presence of a new gene, *crtK*, not described in previous studies. We present here a comprehensive analysis of the DNA sequence and the gene organization, and discuss nucleotide sequences potentially involved in the initiation, regulation and termination of transcription within this region.

RESULTS

Alignment of the Nucleotide Sequence with Genetic-Physical Maps Identifies a New Gene, *crtK*

Sequencing across the *BamHI* sites (Fig. 1) demonstrated that the *BamHI*-J, -M, -G and -H fragments are indeed contiguous. Fig. 1 shows the genes located within the 11039 base pair (bp) sequenced region (Armstrong et al., 1989). Because pRPS404 carries the *crtD223* point mutation (Marrs, 1981), the nucleotide and deduced polypeptide sequences determined reflect this deviation from the *R. capsulatus* wildtype sequences. The sequenced region contains three additional ORFs, designated *crtK*, ORF H and ORF J, distinct from any of the previously described *crt* genes. Interposon mutations introduced at *SalI* (bp 5583) and *NruI* (bp 6723) sites (Fig. 1) have both been proposed to lie within *crtC* because they result in the accumulation of neurosporene, a CrtC$^-$ phenotype (Giuliano et al., 1988). Based on the DNA sequence, however, the interposons interrupt two distinct genes, which cannot be cotranscribed because of their convergent transcriptional orientations. Genetic-physical maps (Taylor et al., 1983; Zsebo and Hearst, 1984) have shown *crtC* to be bounded by *crtB* and *crtD*, with a gap left between crtB and *crtC*. On the basis of these studies, we designate the previously undetected gene found in this gap as *crtK* (Fig. 1).

Ribosome Binding Sites and Start Codons

The proposed amino acid sequences of the *crt* gene products and the ORFs (data not shown) correspond to the longest possible translations of ORFs possessing ribosome binding sites and typical *R. capsulatus* codon usage, located in the appropriate regions of the *crt* gene cluster. Translation of the nucleotide sequence in any of the alternative forward or reverse reading frames, with respect to a given gene, results in the frequent appearance of stop codons and the few alternative ORFs which do have ribosome binding sites show atypical codon usage (data not shown). ATA, CTA and TTA are never found among the 3038 predicted *crt* codons, while GTA (Val) appears only once, within *crtK* (Armstrong et al., 1989). All of the start codons proposed for the *crt* genes are preceded by purine-rich stretches containing possible ribosome binding sites (Fig. 2A) showing complementarity to the 3' end of the *R. capsulatus* 16 S rRNA (Fig. 2B) (Youvan et al., 1984b). An ATG start preceded by a ribosome binding site was not observed for ORFs in the region genetically mapped to *crtF* although a possible GTG start was found (Fig. 2A). We therefore propose that the coding

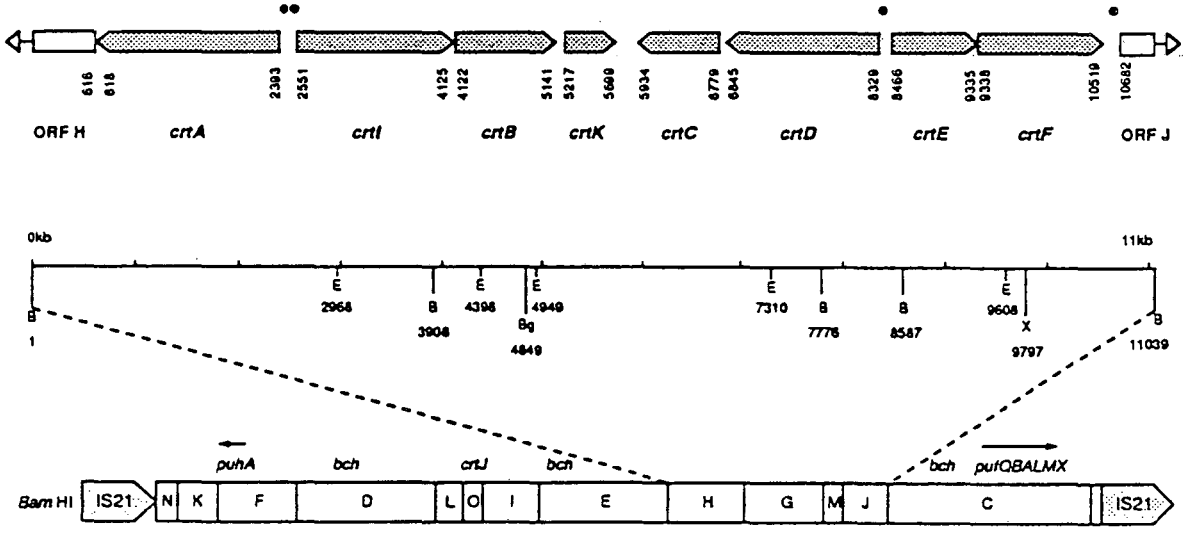


Fig. 1. Organization of the carotenoid biosynthetic gene cluster. Polarities of the *crt* genes (shaded) and ORFs (unshaded) are shown and putative transcriptional regulatory sites (Fig. 5) are indicated by (•). Numbers below the genes show the putative nucleotide positions of translational starts and stops. ORF H and ORF J extend beyond the region here, indicated by the detached arrowheads. A new start site has been assumed for *crtI* (Fig. 2A), replacing our previous proposal (Armstrong et al., 1989). Restriction sites referred to here or in previous genetic-physical mapping studies are indicated below the genes. B, *BamHI*; E, *EcoRI*; Bg, *BglII*; X, *XhoI*. Boxes containing a letter indicate specific restriction fragments from the photosynthesis gene cluster of pRPS404 (Zsebo and Hearst, 1984), while the IS21 elements derived from the vector are indicated to the left and right. The locations of photosynthesis genes outside the *crt* gene cluster are shown above the boxes. The *pufB, A, L and M* genes encode the LH-I β, α and the RC-L, RC-M polypeptides, respectively, while the *puhA* gene encodes the RC-H polypeptide (Youvan et al. 1984a). Regions containing Bchl biosynthetic genes are indicated by *bch*. The structure of *crtJ*, identified by a Tn*5*.7 insertion (Zsebo and Hearst, 1984) and separated from the other *crt* genes by about 12 kb (Fig. 1), is currently under study.

region of *crtF* begins with a GTG start codon. GTG start codons are used in about 8% of *Escherichia coli* genes (Stormo, 1986), and both the *fbcF* gene from *R. capsulatus* (Gabellini and Sebald, 1986), and the *pucB* gene of *Rhodobacter sphaeroides* have GTG starts. We originally proposed that the 5' end of the *crtI* coding region was a GTG codon found at bp 2650 (Armstrong et al., 1989), but more recent evidence suggests that the translation start corresponds to one of two upstream ATG codons located at bp 2551 and bp 2572, respectively (Armstrong G. A., and Hearst, J. E., unpublished data). We have assumed a *crtI* start at bp 2551 throughout the text and figures. Absolute confirmation of the deduced amino acid sequences will require the isolation of the gene products.

Organization of the Carotenoid Gene Cluster

The *crt* genes must form at least four distinct operons because of the inversions of transcriptional orientation which occur between *crtA-crtI*, *crtK-crtC*, and *crtD-crtE* (Fig. 1). *crtA* cannot be cotranscribed with the other *crt* genes because of its divergent orientation at one end of the gene cluster. An interposon insertion at an *EcoR*V site (bp 1303) and a transposon insertion (between bp 999-1244) cause Bchl$^-$ phenotypes, although both of these mutations lie within the 3' end of the *crtA* gene (Giuliano et al., 1988; Armstrong, G. A. and Hearst, J. E., unpublished data). Mutations at the 5' end of *crtA* cause a CrtA$^-$ phenotype but do not affect Bchl synthesis. The most likely explanation for these effects is the polar inactivation of ORF H or a downstream gene in the same operon (Fig. 1) in the 3' insertion mutants. This suggests that *crtA* is not cotranscribed with ORF H, although the promoter(s) for ORF H may overlap *crtA*. ORF H may thus be part of an operon required for Bchl biosynthesis. Groups of genes which could also form operons are *crtIBK*, *crtDC* and *crtEF*. A mutant bearing an interposon insertion at an *Apa*I site (bp 10713) within ORF J exhibits a Bchl$^-$ phenotype, suggesting that this ORF may also belong to an operon required for Bchl biosynthesis (Giuliano et al., 1988). ORF J, located downstream from *crtF* (Fig. 1), does not appear to be transcribed as part of an operon including *crt* genes (Armstrong, G. A. and Hearst, J. E., unpublished data). The coding regions of the *crt* genes are closely spaced. In the most extreme case, the TGA stop codon of *crtI* overlaps the putative ATG start of *crt*B, reminiscent of the overlap between the coding regions of the *R. capsulatus pufL* and *pufM* genes (Youvan et al., 1984a).

A

Gene		5'		3'	Residues	MW
crtA		tcacaggGGAGGactgag	ATG		591	64761
crtB		ccgggccAAGGcGGcgca	ATG		339	37299
crtC		ggcgaAAAGGccttctcg	ATG		281	31855
crtD		tgcgtgcgGGAGcgagcg	ATG		494	52309
crtE		gcagcGGAGGgctctgtc	ATG		289	30004
crtF		cgccgaGAGGgGctgact	GTG		393	43004
crtI	(I)	gaaactaccgaAGAAAcc	ATG		524	57974
crtI	(II)	tccaAGAAcacAGAAggt	ATG		517	57226
crtK		ccacaaccGGAGGccatg	ATG		160	17607

B

5' AGAAAGGAGGTGAT..3'

3' $_{HO}$UCUUUCCUCCACUA..5'

Fig. 2 A, B. Ribosome binding sites, start codons and the predicted gene products. (A) Sequences to the left of the ATG/GTG start codons contain purine-rich stretches (underlined), including nucleotides matching the predicted ribosome binding site (uppercase). The length of the gene product in amino acids and its calculated molecular weight are given to the right. The ribosome binding sites preceding two possible *crtI* start codons are shown, based on a revision of our original proposal for the 5' end of the *crtI* gene (Armstrong, G. A., Alberti, M., Leach, F. and Hearst, J. E., unpublished data). (B) shows the predicted ribosome binding site (above) as the DNA complement of the 3' end of the *R. capsulatus* 16 S rRNA (below) (Youvan et al. 1984b).

5' Non-coding Regions are A + T-rich

Fig. 3 illustrates the extreme asymmetry in % A + T content within the region encoding the *crt* gene cluster. Although the entire genome of *R. capsulatus* has an average A + T content of 34 %, the 5' flanking regions of *crtA-crtI*, *crtD-crtE* and ORF J are unusually A + T-rich, ranging up to 53 % in A + T content averaged over a 151 bp window. These A + T-rich regions contain DNA sequences which may bind transcription factors or serve as *E. coli*-like σ^{70} promoters (see below). The presence of A + T-rich islands in the 5' control regions of genes from an organism with a low average A + T content suggests a compelling selective pressure for the preservation of the nucleotide bias. Non-coding regions surrounding prokaryotic transcription initiation and termination points are A + T rich compared to the coding regions, as determined using a data base composed predominantly of *E. coli* genes (Nussinov et al., 1987). Specific structural features in A + T-rich regions of the chromosome may alert DNA-binding proteins to the presence of potential sites of action (Nussinov et al., 1987).

E. coli-like σ^{70} Promoter Sequences

We have located three sequences closely resembling the σ^{70} consensus promoter, TTGACA N_{15-19} TATAAT (N = any nucleotide) used by the major RNA polymerase of *E. coli* (McClure, 1985). The *R. capsulatus* sequences, found 5' to *crtI*, *crtD* and ORF J, are compared to the canonical *E. coli* promoter in Fig. 4. An optimal spacing of 17 bp is observed between the -35 and -10 regions in *E. coli* (McClure, 1985). The putative *crtD* and *crtI* promoters show spacings of 16 bp, while the putative ORF J promoter has a spacing of 17 bp. No other sequences with a total of nine or more nucleotide matches to the *E. coli* σ^{70} consensus promoter, including five of the six most conserved nucleotides (Fig. 4), were found within the *crt* gene cluster (Fig. 1). We allowed a variable spacing of N_{15-19} between the -35 and -10 regions for these homology searches.

A Conserved Palindromic Motif is Related to a Recognition Site for DNA-binding Regulatory Proteins

We have identified a highly conserved palindromic nucleotide sequence, found four times in 5' flanking regions within the *crt* gene cluster. This motif occurs twice in the *crtA-crtI* 5' flanking region, and once each in the *crtD-crtE* and ORF J 5' flanking regions (Fig. 1). A search among other published *R. capsulatus* nucleotide sequences also revealed the presence of

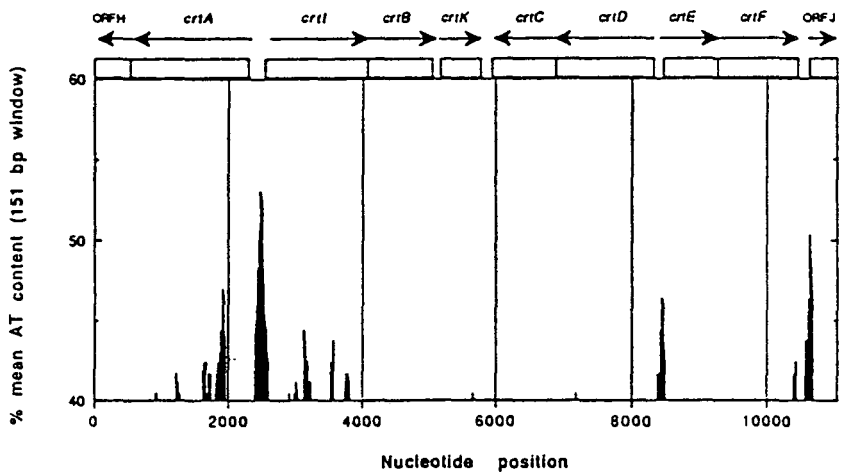


Fig. 3. Average A + T nucleotide content within the carotenoid gene cluster. Analysis of percentage A + T content throughout the DNA sequence was also performed using programs described by Pustell and Kafatos (1982). Locations and polarities of the *crt* genes are indicated at the top. ORF H and ORF J extend beyond the boundaries of the region shown. Nucleotide positions are as in Fig. 1. Percentage A + T content was calculated by averaging over a 151 bp window at 10 bp intervals, and values exceeding 40 % were plotted (average genomic percentage A + T content is 34 %). Note the unusually high mean A + T content in 5' flanking regions of genes.

this palindrome 5' to the coding region of the *puc* operon (Youvan and Ismail, 1985). Based on these five examples, the consensus sequence is TGTAART N_3 A N_2 TTACAC (R = purine) (Fig. 5B). The palindromes are centered anywhere from 162 bp (*pucB*) to 29 bp (*crtA*) from the start codon of the nearest gene (Fig. 5A). Each of the three putative *E. coli*-like σ^{70} promoters located within the *crt* gene cluster overlaps one of the palindromes (compare Figs. 4, 5). No additional palindromes were found when we required matches to each absolutely conserved nucleotide in the consensus (Fig. 5B) in a search of the coding and flanking regions of other published *R. capsulatus* nucleotide sequences encoding proteins (Armstrong et al., 1989), as well as from the 5' end of ORF J to the 5' end of the *pufQ* gene (Fig. 1) (M. Alberti, unpublished data). The *R. capsulatus* consensus palindrome shows strong similarity to a consensus sequence, TGTGT N_{6-10} ACACA, derived from the recognition sites of a collection of prokaryotic transcription factors containing examples of both positive and negative regulators (Fig. 5C) (Gicquel-Sanzey and Cossart, 1982; Buck et al., 1986). The *R. capsulatus* consensus palindrome is, in fact, very similar to the *E. coli* TyrR protein consensus recognition sequence TGTAAA N_6 TTTACA (Yang and Pittard, 1987). TyrR is known to be a transcriptional regulator of genes required for aromatic amino acid metabolism. Based on these sequence similarities to the sites of action of known DNA-binding proteins, we propose that the *R. capsulatus* palindromic motif represents the binding site for a transcription factor.

Rho-independent Transcription Termination Signals

The region shown in Fig. 1 was searched for regions of dyad symmetry with the potential to form stem-loop structures in RNA. Possible secondary structures found between *crtK* and *crtC* include two GC-rich stem-loops, one 3' to *crtK* and the other 3' to *crtC*, each followed by a run of three thymidines (Armstrong et al., 1989). Single regions of dyad symmetry followed by thymidine-rich stretches were found 3' to the *crtI*, *crtB* and *crtF* genes, respectively. The combination of a GC-rich dyad symmetrical region, followed by several thymidine residues is characteristic of rho-independent transcriptional terminators in bacteria (Platt, 1986). Possible rho-independent termination signals were previously noted close to the 3' ends of the *R. capsulatus puf* and *puc* operon mRNAs as mapped by nuclease protection experiments (Zucconi and Beatty, 1988; Chen et al., 1988).

DISCUSSION

The minimum four operons in the *crt* gene cluster are *crtA*, *crtIBK*, *crtEF* and *crtDC* based on the polarities of the genes, although the latter operon seems unlikely because of the phenotypes of polar Ω interposon insertions within *crtD* (Giuliano et al., 1988). Possible rho-independent transcriptional terminators have been found 3' to the *crtI, B, K, C* and *F* genes (Armstrong et al., 1989), suggesting that the former three genes could form separate operons. The *R. capsulatus crt* gene cluster (Fig. 1) is bounded by genetic loci required for Bchl

```
                -35                    -10

2489        TTGtaA atcggaattgac-gacc TATcAT.. 34bp..crtI
8434        TTGgcA ttcgcacctacctgtg- TAaAcT...77bp..crtD
10599       TTGACA gtcgggcgtgtaagttc aATgAT...54bp..ORF J
            ***  *  ***     *        *   *
            TTGACA       N15-19       TATAAT
```

Fig. 4. Comparison of sequences found 5' to *crtI*, *crtD* and ORF J with the *E. coli* σ^{70} consensus promoter. The numbers at left indicate the position of the 5' nucleotide of each sequence (Fig. 1). The distance in bp from each putative promoter to the start codon of the 3' gene is shown at the right. The consensus *E. coli* promoter is shown below with the six most highly conserved nucleotides underlined (McClure, 1985). Putative -35 and -10 regions (above) are indicated in boldface, and uppercase letters show matches to the *E. coli* consensus. Gaps (-) were placed between these regions to maximize the nucleotide alignment. Nucleotides absolutely conserved in all three *R. capsulatus* sequences are indicated by (*).

biosynthesis (Taylor et al., 1983; Zsebo and Hearst, 1984; Giuliano et al., 1988). The correspondence between these loci and specific ORFs has not yet been established, although mutations within ORF J and 5' to ORF H, in the 3' end of *crtA*, cause Bchl$^-$ phenotypes (Giuliano et al., 1988; Armstrong, G. A. and Hearst, J. E., unpublished data). We therefore propose that ORF J and ORF H are part of two operons which include genes required for Bchl but not carotenoid biosynthesis, and which are transcribed outwards away from the *crt* gene cluster. *crtK* was not identified previously by interposon mutagenesis, presumably because of the similarity of CrtC$^-$ and CrtK$^-$ phenotypes and the fact that *crtC* and *crtK* are adjacent (Giuliano et al., 1988).

E. coli-like σ^{70} promoters have never been observed in *Rhodobacter* (Kiley and Kaplan, 1988), nor have detailed data on any *R. capsulatus* promoters been available until recently. We have, however, found possible *E. coli*-like σ^{70} promoters (McClure, 1985) 5' to *crtI*, *crtD* and ORF J (Fig. 4). Within the constraints of our homology search (see Results), no other *E. coli*-like σ^{70} promoter sequences were found within the *crt* gene cluster, although the gene

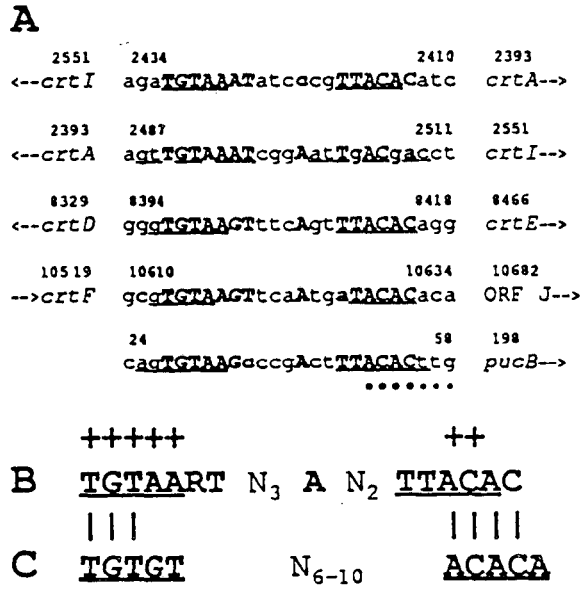


Fig. 5 A-C. Comparison of a palindromic motif found 5' to photosynthesis genes with a consensus regulatory protein binding site. (A) The genes flanking each palindrome are indicated to the left and right, respectively. Arrows show the directions of transcription. Numbers above each sequence show the nucleotide positions (as in (Fig. 1), except for *pucB* (see Youvan and Ismail 1985)) of the 5' or 3' ends of the flanking genes with respect to the location of the palindrome. Possible *puc* operon transcription initiation signals (Zucconi and Beatty, 1988) are indicated by (•). Complementary nucleotides in the two halves of the palindromes are underlined. (B) Nucleotides which match the *R. capsulatus* consensus are given in uppercase, while those that occur in positions defined by the consensus are shown in boldface. (+) indicates an absolutely conserved nucleotide in the palindrome. (C) The *R. capsulatus* consensus sequence is compared to a consensus derived from the recognition sites of the transcription factors NifA, AraC, CAP, LacI, GalR, LexA, TnpR, LysR and λ cII. Nucleotides conserved between the two consensuses are indicated by (|) between the sequences.

organization suggests that there must be promoters 5' to both *crtA* and *crtE* (Fig. 1). Whether these promoters have a weaker match to the σ^{70} consensus or perhaps have an entirely different structure remains to be determined. Neither Bchl nor carotenoids accumulate in *E. coli* strains harboring the *R. capsulatus* photosynthesis gene cluster carried on pRPS404 (Marrs, 1981). Our observation that the *R. capsulatus crtD* and *crtI* genes may have *E. coli*-like σ^{70} promoters, thus, was not anticipated. *E. coli* may fail to recognize at least one *R. capsulatus crt* promoter or lack the proper transcription factors required for *crt* gene expression. In addition, post-transcriptional regulation could also differ between the two species.

We have found five examples of a conserved nucleotide motif (Fig. 5A) in the 5' flanking regions of *R. capsulatus* photosynthesis genes. One example of the *R. capsulatus* palindromic motif occurs 5' to the *puc* operon (Fig. 5A), which encodes the LH-II antenna polypeptides. Zucconi and Beatty (1988) mapped the 5' triphosphate-containing ends of *puc* operon mRNAs and have suggested that a direct repeat of ACACTTG, located 5' to each of the two mapped mRNA start sites, may be involved in transcription initiation. The palindrome 5' to the *puc* operon overlaps the upstream ACACTTG sequence (Fig. 5A). and is located ~35 and ~50 nucleotides, respectively, upstream from the two 5' ends of the *puc* mRNAs.

Three other examples of the palindrome overlap the putative *E. coli*-like σ^{70} promoter sequences found 5' to *crtI*, *crtD* and ORF J (compare Figs. 4, 5). We propose a role for the palindromes in transcriptional regulation because of the extraordinary conservation of the motif and its sequence similarity to binding sites of known transcription factors, and because of its overlap with three putative *E. coli*-like σ^{70} promoters in the *crt* gene cluster. Overlap of the *R. capsulatus* palindromes with promoter sequences could be consistent with either positive or negative gene regulation. The regulatory sites may also be widely separated from the promoters with which they interact. Further experiments are in progress to define the interaction between the putative regulatory palindromes and sequences involved in transcription initiation.

The *puc* operon is highly regulated at the transcriptional level in response to oxygen tension (Klug et al., 1985). We have recently shown that expression of several *crt* genes is strongly induced during a shift from aerobic to photosynthetic growth (Armstrong, G. A. and Hearst, J. E., unpublished data), while Giuliano et al. (1988) have found an increase in the steady-state levels of 5' ends from *crtA*, *C* and *E* mRNAs in anaerobic versus aerobic cultures. The common feature of anaerobic gene induction would seem a reasonable explanation for the unexpected presence of identical transcriptional regulatory signals 5' to both the *puc* and *crt* operons. On the other hand, these regulatory sequences are not found close to the *puf* and *puh* operons (Fig. 1), whose expression is also induced by reduction of the oxygen tension (Clark et al., 1984; Klug et al., 1985). The palindromes (Fig. 5A) may thus bind a transcription factor involved in the regulation of a subset of the *R. capsulatus* photosynthesis genes. Whether a linkage exists between the expression of the *puc* operon and the regulated *crt* genes remains to be tested. We have determined the first nucleotide and deduced amino acid sequences of genes and genes products involved in carotenoid biosynthesis, and have also identified possible promoter, terminator and transcriptional regulatory signals which govern *crt* gene expression. Previous studies of *crt* gene regulation in *R. capsulatus* have been hampered by the lack of gene-specific probes (Clark et al., 1984; Klug et al., 1985; Zhu and Hearst, 1986; Zhu et al., 1986). The work presented here will facilitate an examination of the regulation of individual *crt* genes.

References

Armstrong, G. A., Alberti, M., Leach, F., and Hearst, J. E. (1989) Nucleotide sequence, organization, and nature of the protein products of the carotenoid biosynthesis gene cluster of *Rhodobacter capsulatus*, Mol. Gen. Genet., 216:254-268.

Buck, M., Miller, S., Drummond, M., and Dixon, R. (1986) Upstream activator sequences are present in the promoters of nitrogen fixation genes, Nature, 320:374-378.

Chen, C. Y., Beatty, J. T., Cohen, S. N., and Belasco, J. G. (1988) An intercistronic stem-loop structure functions as an mRNA decay terminator necessary but insufficient for *puf* mRNA stability, Cell, 52:609-619.

Clark, W. G., Davidson, E., and Marrs, B. L. (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.

Cogdell, R. J., and Frank, H. A. (1987) How carotenoids function in photosynthetic bacteria, Biochim. Biophys. Acta, 895:63-79.

Gabellini, N., and Sebald, W. (1986) Nucleotide sequence and transcription of the *fbc* operon from *Rhodopseudomonas sphaeroides*, Eur. J. Biochem., 154:569-579.

Gicquel-Sanzey, B., and Cossart, P. (1982) Homologies between different procaryotic DNA-binding regulatory proteins and between their sites of action, EMBO J., 1:591-595.

Giuliano, G., Pollock, D., Stapp, H., and Scolnik, P. A. (1988) A genetic-physical map of the *Rhodobacter capsulatus* carotenoid biosynthesis gene cluster, Mol. Gen .Genet., 213:78-83.

Goodwin, T. W. (1980) The Biochemistry of the Carotenoids: Plants, Chapman and Hall, Ltd., New York, New York.

Kiley, P. J., and Kaplan, S. (1988) Molecular genetics of photosynthetic membrane biosynthesis in *Rhodobacter sphaeroides*. Microbiol. Rev., 52:50-69.

Klug, G., Kaufmann, N., and Drews, G. (1985) Gene expression of pigment-binding proteins of the bacterial photosynthetic apparatus: transcription and assembly in the membrane of *Rhodopseudomonas capsulata*, Proc. Nat. Acad. Sci., USA, 82:6485-6489.

Marrs, B. (1981) Mobilization of the genes for photosynthesis from *Rhodopseudomonas capsulata* by a promiscuous plasmid, J. Bacteriol., 146:1003-1012.

McClure, W. R. (1985) Mechanism and control of transcription initiation in prokaryotes. Annu. Rev. Biochem., 54:171-204.

Nussinov, R., Barber, A., and Maizel, J. V. (1987) The distributions of nucleotides near bacterial transcription initiation and termination sites show distinct signals that may affect DNA geometry, J. Mol. Evol., 26:187-197.

Platt, T. (1986) Transcription termination and the regulation of gene expression, Annu. Rev. Biochem., 55:339-372.

Pustell, J., and Kafatos, F. (1982) A convenient and adaptable package of DNA sequence analysis programs for microcomputers, Nucl. Acids. Res., 10:51-59.

Stormo, G. D. (1986) Translation initiation, in: "Maximizing Gene Expression," W. Reznikoff W and L. Gold, eds., Butterworths, Stoneham, Massachusetts.

Taylor, D. P., Cohen, S. N., Clark, W. G., and Marrs, B. L. (1983) Alignment of the genetic and restriction maps of the photosynthesis region of the *Rhodopseudomonas capsulata* chromosome by a conjugation-mediated marker rescue technique, J. Bacteriol., 154:580-590.

Yang J., and Pittard, J. (1987) Molecular analysis of the regulatory region of the *Escherichia coli* K-12 *tyrB* gene, J. Bacteriol., 169:4710-4715.

Youvan, D. C., Bylina, E.J., Alberti, M., Begusch, H., and Hearst, J. E. (1984a) Nucleotide and deduced polypeptide sequences of the photosynthetic reaction-center, B870 antenna, and flanking polypeptides from *R. capsulata*, Cell, 37:949-957.

Youvan, D. C., Alberti, M., Begusch, H., Bylina, E. J., and Hearst, J. E. (1984b) Reaction center and light-harvesting genes from *Rhodopseudomonas capsulata*, Proc. Nat. Acad. Sci., USA, 81:189-192.

Youvan, D. C., and Ismail, S. (1985) Light-harvesting II (B800-B850 complex) structural genes from *Rhodopseudomonas capsulata*, Proc. Nat. Acad. Sci., USA, 82:58-62.

Zhu, Y. S., and Hearst, J. E. (1986) Regulation of the expression of the 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. Nat. Acad. Sci., USA, 83:7613-7617.

Zhu, Y. S., Cook, D. N., Leach, F., Armstrong, G. A., Alberti, M., and Hearst, J. E. (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.

Zsebo, K. M., and Hearst, J. E. (1984) Genetic-physical mapping of a photosynthetic gene cluster from *R. capsulata*, Cell, 37:937-947.

Zucconi, A. P., and Beatty, J. T. (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.

LAWRENCE BERKELEY LABORATORY
TECHNICAL INFORMATION DEPARTMENT
1 CYCLOTRON ROAD
BERKELEY, CALIFORNIA 94720