JOURNAL OF VIROLOGY, Apr. 1986, p. 142–151
0022-538X/86/040142-10$02.00/0

Vol. 58, No. 1

Nucleotide Sequence of the Small Double-Stranded RNA Segment of Bacteriophage φ6: Novel Mechanism of Natural Translational Control

TIMOTHY McGRAW,† LEONARD MINDICH,* AND BLAS FRANGIONE‡

Department of Microbiology, The Public Health Research Institute of the City of New York, Inc., New York, New York 10016

Received 6 October 1985/Accepted 27 December 1985

The lipid-containing bacteriophage φ6 has a genome composed of three segments of double-stranded RNA. We determined the nucleotide sequence of a cDNA copy of the smallest RNA segment. The coding sequences of the four proteins on this segment were identified. These sequences were clustered. Three of the genes had overlapping initiation-termination codons. All noncoding sequences were at the ends of the molecule. The genes of the small double-stranded RNA segment comprised two translational polarity groups. We propose that the translational coupling is the result of an inability of ribosomes to bind independently to two of the four genes. Translation of these genes occurred when ribosomes were delivered to them by translation of an upstream gene.

Phage φ6 is a lipid-containing bacteriophage of *Pseudomonas phaseolicola* (50). The genome is composed of three segments of double-stranded RNA (dsRNA), the sizes of which are 7.0, 4.0, and 3.0 kilobase pairs (kbp) (37). φ6 directs the synthesis of 12 proteins, 11 of which are found in the virion (42). The virion is composed of a polyhedral nucleocapsid that contains one copy of each dsRNA segment. Surrounding the nucleocapsid is an envelope composed of phosphatidylglycerol, phosphatidylethanolamine, and four phage proteins. Because of its relative simplicity, φ6 is a useful model system for studying the morphogenesis of membrane structures. Previous investigations from this laboratory have shown that at least one of the phage membrane proteins (P9) and a morphogenetic protein (P12) are required for the envelopment of the nucleocapsid (2, 29).

Three of the genes involved in the membrane acquisition are on the smallest RNA segment (segment S): P8, the major nucleocapsid protein; P9, the major membrane protein; and P12, the morphogenetic protein (21). The phage lysis protein P5 is also on this segment (21). These genes can be divided into two polarity groups: P8 nonsense mutations are completely polar on P12; P9 nonsense mutations are completely polar on P5 (40). There is no polarity between genes 8 and 9. The observed polarities are translational, rather than transcriptional, because one polycistronic mRNA molecule is made from segment S, and both P8 and P9 nonsense mutants make full-length segment S mRNA in vivo (41). Translational polarity has been documented in other bacterial systems, and this type of regulation has been termed translational coupling (32). A unique aspect of the translational coupling of the φ6 genes is the amount of proteins synthesized. P12 and P5 are made in approximately one-tenth the amounts of P8 and P9 (42).

To facilitate further studies of the particle assembly pathway, particularly the role that each protein plays in this process, cDNA clones of the φ6 RNA segments were prepared (27). Two cDNA fragments, designated F90 and F84, were found by Northern blot analysis to be homologous with segment S (27). Fragment F90 directs the synthesis of the four segment S proteins in a cell-free, coupled transcription-translation system (27). Furthermore, fragment F90, when carried on a vector in *P. phaseolicola*, complements mutations in the genes of segment S, thereby establishing that the cDNA cloned genes are biologically active (27). Here we report the complete nucleotide sequence of the S segment. The open reading frames (ORFs) of the four proteins are identified, and a model for the translational couplings is proposed.

MATERIALS AND METHODS

Bacterial strains, phage, and plasmids. *P. phaseolicola* HB10Y is the normal host for φ6 (50). *Pseudomonas pseudoalcaligenes* ERA(pLM2S4) is a suppressor strain used for the preparation of φ6 nonsense mutants (25). φ6*hls* is the phage strain that was used as the wild type; it carries a mutation in the gene for the absorption protein P3 that allows it to grow on plates of *P. pseudoalcaligenes* and a second mutation in an unidentified locus that improves plating efficiency on the suppressor strain. The phage mutants used in this study were derivatives of φ6*hls* which contained nonsense mutations in genes 8, 9, 12, or 5 (11). These phages also contained a second mutation, *h591*, which is a host range mutation mapped to the large segment (27).

Escherichia coli JM83 (a Δ*lacM15* strain [51]) and LM3 (a spectinomycin-resistant, streptomycin-sensitive derivative of JM83 [27]) were used as hosts for subclonings.

Plasmid pUC8 was purchased from Bethesda Research Laboratories, Gaithersburg, Md. (51). Plasmids pLMF90 and pLMF84 are pBR322 derivatives containing cDNA insert fragments of segment S (27). These inserts were designated F90 and F84. Plasmid pLM254, which can be maintained in both *E. coli* and *P. phaseolicola*, is an RSF1010 derivative containing the *lacZ*-polylinker region and the ampicillin resistance gene of pUC8 (27).

Preparation of DNA. Minilysate plasmid DNA, which was

* Corresponding author.
† Present address: Department of Pharmacology, New York University School of Medicine, New York, NY 10016.
‡ Present address: Department of Pathology, New York University School of Medicine, New York, NY 10016.

used for screening transformants, was prepared from 5 ml of overnight cultures by the procedure of Birnboim and Doly (3). Large plasmid preparations were prepared from 200 ml of overnight cultures by the cleared lysate method of Clewell (5). DNA was further purified by CsCl centrifugation or by passage over a Bio-Gel A50M (Bio-Rad Laboratories, Richmond, Calif.) column equilibrated with 0.5 M NaCl–10 mM Tris hydrochloride (pH 8.0).

Procedures for restriction digestion, 0.8% agarose electrophoresis, 5% polyacrylamide gel electrophoresis, and ligations have been described previously (23, 28). Transformation of *E. coli* and *P. phaseolicola* has been described previously (27).

DNA sequencing. Nucleotide sequence analysis was performed as described by Maxam and Gilbert (22). Fragments for sequencing were 5′ labeled, following dephosphorylation with bacterial alkaline phosphatase, by treatment with polynucleotide kinase and [γ-^{32}P]ATP (22) or 3′ labeled by treatment with terminal nucleotide transferase in the presence of [α-^{32}P]ddATP (22). The labeled fragments were secondarily cut, separated on 5% polyacrylamide gels, and electroeluted from gel slices at 120 V in 0.088 M Tris borate (pH 8.0) buffer. In some instances in which there were no convenient sites for secondary cutting, the fragments were strand separated on 8% gels (22) and then electroeluted. The purified, singly labeled fragments were chemically modified and cleaved by the procedures described by Maxam and Gilbert (22). The products of the reactions were displayed on 8 and 20% polyacrylamide–8.3 M urea sequencing gels.

The initial step in the sequence analysis was to subclone F90 DNA fragments into pUC8. The F90 insert fragment was liberated from a pBR322 clone (pLMF90) (27) by digestion with *Pst*I and then isolated from a 1% low-melting-point agarose gel. The purified fragment was digested with *Pvu*II, and the resulting six fragments were subcloned into pUC8. The *Pvu*II fragments (bound on both sides by *Pvu*II sites) were cloned into the *Hin*cII site, and the two terminal *Pst*I-*Pvu*II fragments were cloned into the *Hin*cII-*Pst*I sites of pUC8. This approach was adopted because DNA fragments inserted into the pUC8 cloning region can be sequenced from both sides by labeling at vector sites located within 20 base pairs (bp) of the insert-vector junction.

The 2.7-kb *Bam*HI-*Pst*I fragment of F90, which contains all of F90 sequences except for 62 bp at the 5′ terminus, was subcloned into the *Bam*HI-*Pst*I sites of pUC8. In this construct, designated pTM202, the cloned genes were oriented so that they were transcribed from the *lac* promoter of pUC8. Plasmid pTM202, containing all of the protein coding sequences of segment S, directs the synthesis of proteins P8, P12, P9, and P5 (P11) in a cell-free, coupled transcription-translation system (27). Sequencing across the intact *Pvu*II sites was accomplished by labeling at the *Dde*I, *Eco*RV, *Sal*I, and *Cla*I sites of the insert fragment of pTM202.

Dideoxy DNA sequencing. Two gene 8 deletion constructs, pTM248 and pTM249, were sequenced by the chain termination method of Sanger et al. (35). The *Eco*RI-*Xmn*I fragments from pTM248 and pTM249 (675 and 682 bp, respectively) were cloned into *Eco*RI-*Sma*I-cut phage M13mp11. The sequences of the *Sma*I-*Sca*I fusion site of pTM248 and the *Sma*I-*Eco*RV fusion site of pTM249 were derived from this analysis.

Dideoxy RNA sequencing. Regions of segment S were sequenced with dideoxynucleotides and avian myeloblastosis virus reverse transcriptase (Life Sciences, Inc., St. Petersburg, Fla.). In vitro-synthesized φ6 mRNA (33; provided by L. Lehman) was used as template. Three 17mer primers were synthesized with an Applied Biosystems 380A DNA synthesizer. The primers used were as follows: I, 3′-ACAAGGGTCTACAGTGC-5′ (base 2309 to base 2325; nucleotide position of the complete small segment with base 1 being the first base on the 5′ end; plus strand sense); II, 3′-AATCCGGCATTGGAAGG-5′ (base 2001 to base 2017); III, 3′-AGACTCTGGGCACGGCG-5′ (base 1701 to base 1717). The primers were gel purified from 20% polyacrylamide–8.3 M urea gels by the crush and soak method (22). The conditions for the sequence reactions were basically those of Hamlyn et al. (16); the optimal primer-template ratio was empirically determined for each primer.

Protein purification and amino acid sequencing. Purified P9 and P8 were isolated by the method of Sinclair et al. (42). Automated Edman degradation was achieved by using a Beckman 890c sequencer and a 0.1 M Quadril program (11). The amino acids were identified as their phenylthiohydantoin derivatives.

In vivo complementation assay. The biological activity of the cloned genes was assayed by spotting phage mutants on a lawn of *P. phaseolicola* carrying the cDNA cloned genes. Complementation was scored as the efficiency with which the mutants were able to form plaques on the clone-bearing strains (27).

In vitro protein synthesis. An S-30 extract was prepared from *E. coli* MRE600 cells by the procedure of Zubay (53). The in vitro protein synthesis reaction was performed in 25-μl reaction mixtures containing 2 to 3 μg of template DNA, 30 μCi of [^{3}H]leucine, 2 μl of S-30 extract, amino acids, and cofactors at the recommended concentrations (53). The samples were incubated at 37°C for 1 h with vigorous shaking. The protein was precipitated by the addition of 9 volumes of 90% acetone. The protein pellets were washed once with 90% acetone, air dried, suspended in 1 volume of running dye (48), placed in a boiling-water bath for 2 min, and stored at −20°C until use.

Protein samples were displayed on 10 to 20% linear polyacrylamide gels (42). Following fixation the gels were fluorographed with En^{3}Lighten (New England Nuclear Corp., Boston, Mass.), dried, and exposed at −70°C.

Computer analysis. The computer facilities at The Public Health Research Institute of the City of New York and The Los Alamos Sequence Analysis Center were used.

RESULTS

Determination of the nucleotide sequence. The strategy for sequencing the cDNA fragments F90 and F84 (27) by the chemical cleavage method of Maxam and Gilbert (22) is presented in Fig. 1 (see above). The entire sequence of the small RNA segment was determined. Approximately 97% of the sequence was determined on both strands. The nucleotide sequence is presented in Fig. 2. Segment S is 2,948 bp, which corresponds well with the reported size of 3.1 kb (37). The G+C content of segment S was 55%.

Iba et al. (17) have reported the terminal sequence of the three φ6 segments. F90 was found to contain the 5′ (plus strand sense) terminal sequence, preceded by a 27-bp deoxyguanosine tail which was added during the cDNA cloning (27). F90 was missing 143 bp at the 3′ terminus of segment S (Fig. 1). Clone F84 contained the 3′-terminal sequence, followed by a stretch of 12 deoxyadenosine and 10 deoxyguanosine residues which were added during the cloning procedure. F84 was missing approximately 700 bp at the 5′ end (Fig. 1). Fragments F90 and F84 therefore constitute a complete copy of the small segment. All the reported sequence was derived from F90, except for 143 bp at the 3′

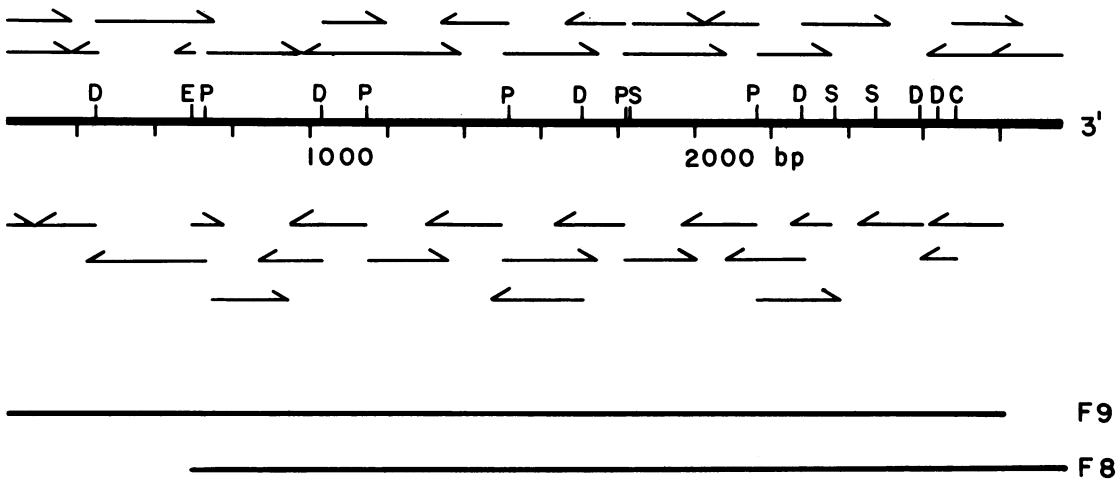


FIG. 1. Sequencing strategy of segment S cDNA clones. Left to right above the center line represents the sequence derived in the 5′ to 3′ direction from the plus strand; right to left above the center line represents the sequence derived from the plus strand in the 3′ to 5′ direction (fragments were 3′ end labeled). Below the center line, right to left is 5′ to 3′ on the minus strand. The tails of the arrows mark the labeling sites, and the length of the arrows represents the approximate sequence data read from each labeling. The majority of sequence information was derived from *Pvu*II subclones in pUC8. In these cases the arrows are marked as originating from the *Pvu*II sites, although the actual labeling sites were within the vector sequences. The sequence contained within the two cDNA clones pLMF90 and pLMF84 (F90 and F84) are shown below the sequencing strategy. pLMF90 contains all the sequence of segment S except for the 3′-terminal 143 base pairs. This information is contained in pLMF84. Only those restriction sites used in sequencing are shown: C, *Cla*I; D, *Dde*I; E, *Eco*RV; P, *Pvu*II; S, *Sal*I.

terminus which were determined from fragment F84. The ORFs of the four proteins of the small segment were identified. Each gene will be discussed separately.

P8, the major nucleocapsid protein. Protein P8 forms the surface on which the ϕ6 membrane is eventually assembled. The ORF coding for P8 began with an AUG at nucleotide 305 (N305) and terminated with a UAA at N752 (Fig. 2). Preceding the AUG codon by 8 bp was the sequence AAGGA, which is capable of base pairing to the 3′ terminus of the 16S RNA of the host pseudomonads of ϕ6 (52) and therefore is considered to be a Shine-Dalgarno (SD) sequence (38). The N-terminal sequence of purified P8 (44 amino acids) was determined by automated Edman degradation (Fig. 2). This result confirmed the reading frame assignment. This reading frame coded for a protein of 16,000 daltons (Da). Values of 10,500 Da (42) and 14,000 Da (49) were assigned to P8 by virtue of its migration in sodium dodecyl sulfate (SDS)-polyacrylamide gels.

P12, the morphogenetic protein. Protein P12 was the only ϕ6 protein not found in the virion. It plays an essential role in the formation of the viral membrane (21, 40). The ORF assigned to protein P12 began at N754 with an AUG codon and terminated at N1339 with a UAA codon (Fig. 2). This reading frame was assigned to P12 because it directly follows gene 8 and codes for a protein of the proper size (20,300 Da), compared with the size determined by SDS-polyacrylamide gel electrophoresis (PAGE) of 20,000 Da (42). This was the only ORF that directly followed gene 8 and had the proper size. Protein P12 was not purified, therefore precluding N-terminal amino acid sequence confirmation of this ORF assignment.

The condition that gene 12 must follow gene 8 was imposed because of the translational coupling of gene 12 to gene 8 (40). We found that these genes overlapped. The UAA termination codon of P8 overlapped with the AUG initiation codon of P12 (UAAUG). This construction alone was not sufficient to account for the translational coupling of these genes.

There was a purine-rich sequence (GGGA) 11 bp 5′ to the initiation codon of gene 12. This sequence could possibly serve as a ribosomal binding site (SD sequence); however, it would be a poor one, both in sequence and in its position relative to that of the initiation codon (47).

The possible mRNA secondary structures of the gene 8-gene 12 region were investigated by using the secondary structure algorithm of Zucker and Stiegler (54). No secondary structures which could be invoked to explain the translational coupling were found. A model to explain the translational coupling of these genes is presented below.

P9, the major membrane protein. P9, along with three other phage proteins and phospholipids, constitutes the phage envelope. P9 is the major phage membrane protein (7), and it is essential for viral membrane assembly (29). The ORF coding for P9 began with an AUG at N1341 and terminated with a UAA at N1611 (Fig. 2). The N-terminal sequence (66 amino acids) of purified P9 confirmed the reading frame assignment (Fig. 2). Also, the predicted amino acid composition was consistent with that previously calculated by Sinclair et al. (42). The size of P9 predicted by the sequence is 9,600 Da. The values reported previously were 12,600 Da (49) and 8,700 Da (42).

Gene 9 overlapped with gene 12 exactly as gene 12 did with gene 8 (UAAUG). Preceding the AUG codon by 9 bp was an SD sequence identical to the one which preceded gene 8 (Fig. 2). There was no coupling of the translation of gene 9 to that of gene 8 or 12, as mutations in P8 or P12 do not affect the level of P9 that is made (40).

P5, the phage lysozyme. Protein P5 is a phage-encoded lysozyme. P5 is found in the virion; it has been shown to be involved in both phage penetration and host cell lysis (26). The molecular mass of P5, as calculated by SDS-PAGE, is 24,000 Da. The synthesis of P5 is closely associated with the synthesis of another virion protein (P11) which has an apparent molecular mass of 25,000 Da. Nonsense mutations which block the synthesis of one also block the synthesis of the other (40). Mutations which alter the migration (SDS-

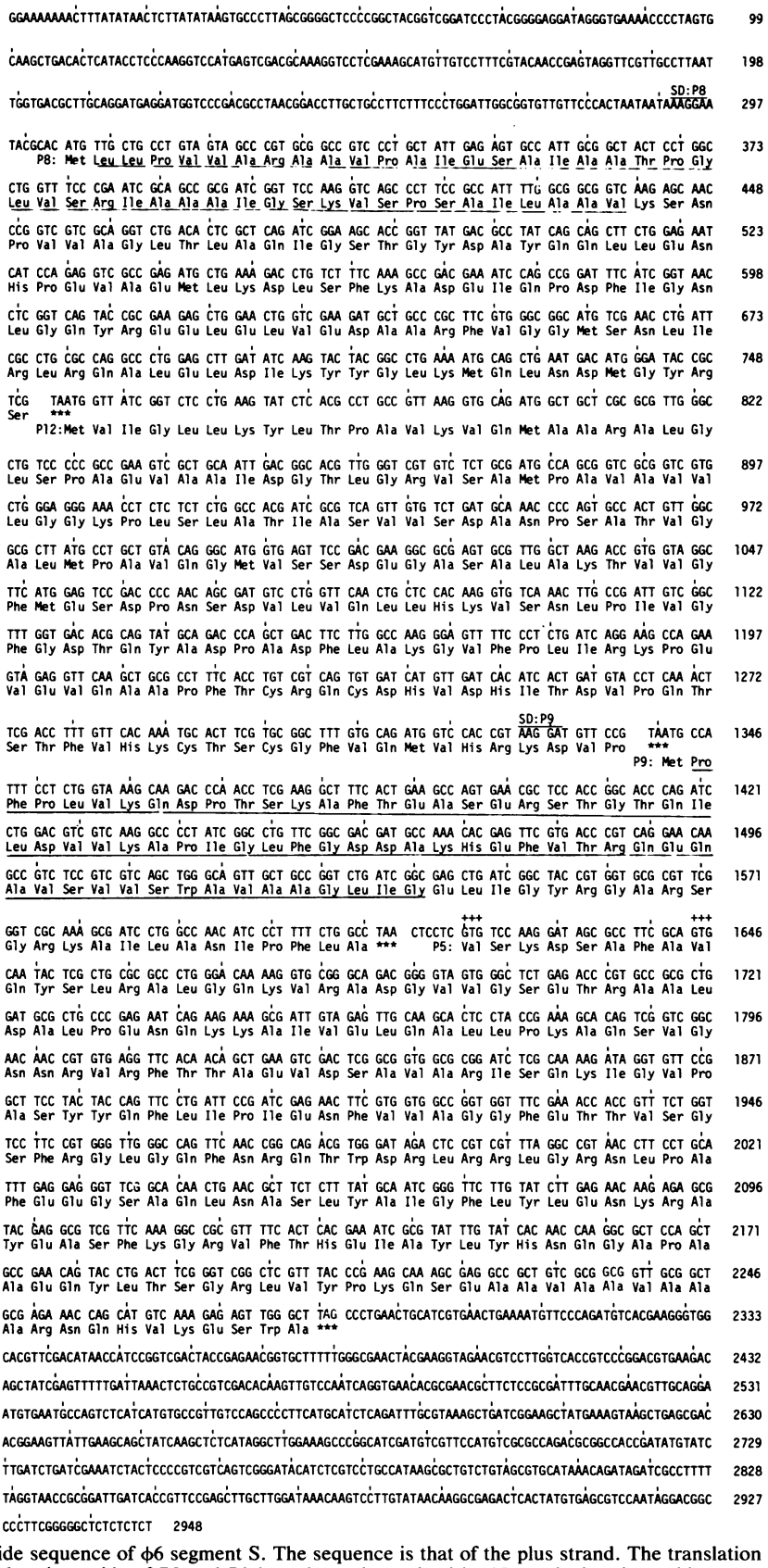


FIG. 2. cDNA nucleotide sequence of ϕ6 segment S. The sequence is that of the plus strand. The translation products of the four genes are shown. The underlined amino acids of P8 and P9 have been determined by N-terminal amino acid sequence analysis. SD sequences preceding genes 8 and 9 are indicated. Symbols: ***, termination codons of the four ORFs; +++, two possible initiation sites of P5(P11).

TABLE 1. Nucleotide differences between cDNA clone and mRNA of selected regions of ϕ6 segment S

Position	Nucleotide in:		Position	Predicted amino acid in:	
	cDNA	RNA		cDNA	RNA
1604	T	C	Codon 87 of P9	Phe	Phe
1643	A	G	Codon 8 of P5(P11)	Ala	Ala
1801	A	G	Codon 61 of P5(P11)	Asn	Ser
1877	C	T	Codon 86 of P5(P11)	Ser	Ser
1989	A	G	Codon 124 of P5(P11)	Arg	Gly

PAGE) of one also alter the migration of the other by an equivalent amount (29). Furthermore, results of radioactive amino acid incorporation studies have demonstrated that P5 and P11 do not contain methionine or cysteine (42; unpublished data). We believe that these two proteins are products of the same ORF.

Proteins P5 and P11 could result from translation initiation or termination at two different sites within the same ORF, or as a result of protein modification, for example, processing or a charge change due to derivatization of an amino acid. A precursor-product relationship between P11 and P5 is not observed in vivo (42). It is of interest to note that P11 is synthesized in vitro, whereas almost no P5 is made (27). Regardless of the relationship between these proteins, P11 is not an essential protein for the virus. In an alternative host for ϕ6, *P. pseudoalcaligenes*, only P5 is made (40).

The synthesis of P5 and P11 [referred to as P5(P11), in which no distinction is made between the two peptides] is translationally coupled to the synthesis of P9. Nonsense mutations in P9 are completely polar on P5(P11) (40). Because of the translational polarity between P9 and P5(P11), we propose that the ORF for P5(P11) should be located 3′ to gene 9. An ORF of 220 amino acids directly followed the termination of gene 9, from N1614 to UGA at N2280 (Fig. 2). This was the only reading frame following P9 of the proper size which did not contain methionine or cysteine. The assignment of the exact translation initiation site is tentative. There are two candidates for translation initiation: GUG at N1620 and GUG at N1644. Initiation at the former would result in a protein of 23,300 Da, and initiation at the latter would result in a protein of 24,000 Da. Neither P5 nor P11 was purified; therefore, the amino acid sequence could not be used for confirmation of the ORF assignment. Neither of these initiation codons was preceded by properly positioned SD sequences. The GUG at N1620 was not preceded by a recognizable SD sequence. Preceding the GUG at N1644 by 13 bp was the sequence AAGGA, which was identical to the SD sequences preceding genes 8 and 9. This sequence is not ideally spaced to serve as a good SD sequence (47).

The possible mRNA secondary structures of the gene 9-gene 5(11) region were investigated by using the secondary structure algorithm of Zucker and Stiegler (54). As was the case with genes 8 and 12, we did not find any secondary structures which could be invoked to explain the translational coupling. A model to explain the translational coupling of these genes is presented below.

A comparison of the amino acid sequence predicted by this ORF to those of some of the other known lysozymes did not reveal any sequence similarities. This result was not surprising. The lysozymes of hen egg white, bacteriophage T4, and Embden goose have similar structures without detectable amino acid sequence similarity (14).

Dideoxy RNA sequence. Selected regions of segment S were sequenced with mRNA template, dideoxynucleotides, and avian myeloblastosis virus reverse transcriptase. A total of 680 bp were sequenced from three primers (see above). We found five single nucleotide differences between the RNA and cDNA sequences (Table 1). All changes were transitions. Two of these resulted in amino acid changes within P5(P11) coding sequences. Neither of these changes detectably altered the activity of P5(P11), because we have established that the cloned gene is biologically active (27). This limited analysis suggests an error rate of 7×10^{-3}, which is consistent with the predicted reverse transcriptase error rate of 10^{-3} (13). A further source of error was the possible heterogeneity of the single-stranded RNA population (as synthesized by ϕ6 replicase) which was used as template for the synthesis of the cDNA clones (27).

Other possible ORFs. The previously discussed genes are the only known genes of the small segment. To search for other possible proteins that code for ORFs, two algorithms were employed. The TESTCODE program identifies possible coding regions based on nucleotide periodicity differences between coding and noncoding ORFs and is independent of codon usage (10). This program predicted six possible coding frames with lengths greater than 50 amino acids. Included in this set were the genes discussed above. Of the other two possible ORFs only one was not completely contained within sequences assigned to the known genes. This ORF began at AUG at N2290 and terminated 111 amino acids later at UAA at N2593. This ORF overlapped with the terminal 20 bp of gene 5(11). There was no evidence for the product of this ORF.

The PERCEPTRON algorithm discriminates between coding and noncoding frames by searching for translation initiation signals (including ribosomal binding sites) within a predetermined region surrounding the beginning of an ORF (46). The PERCEPTRON algorithm, which examined regions of both 101 and 71 nucleotides in length, identified the coding frames assigned to genes 8 and 9. These were the only ORFs greater than 50 amino acids identified by this program. This result suggests that the ORFs of P12 and P5(P11) are not preceded by good ribosomal binding sites, which is consistent with our visual inspection of the regions 5′ to these genes. The ϕ6 genes are faithfully translated by *E. coli* ribosomes, thereby supporting the use of *E. coli* initiation signals as a guide for identifying the ϕ6 ORFs (6, 9, 27).

Codon usage. The codon usage of ϕ6 appeared to follow most of the trends observed in the other procaryotic organisms that have been studied (31). A distinct preference for *leu* codon CUG is found in the ϕ6 genes. A similar preference is seen in the enterobacteria but not in the bacteriophages.

There were no codon usage differences among the genes of segment S. This is an important observation in the context of understanding translational regulation because a possible mechanism for control could be the frequency with which codons that correspond to rare tRNA species are used (18, 19).

Gene organization. The genes of ϕ6 are tightly packed, as has been found in other bacteriophages (8, 34). Three of the four genes of segment S overlap: genes 8-12-9. In both cases the region of overlap has the same construction: UAAUG. As discussed above, the exact initiation site for P5(P11) was not identified, but the ORF which contains this gene begins immediately after the termination of gene 9. The 5′-terminal 300 bp and the 3′-terminal 600 bp do not code for any of the known proteins. The question must then be raised as to what

function these noncoding sequences are serving. The most obvious function would be an involvement in replication and genome packaging. The mRNA and the dsRNA genomic form of ϕ6 are synthesized by the phage replicase (20, 41). The terminal noncoding sequences could provide the signals for packaging, replicase recognition, and regulation of transcription.

Translational regulation. There are two aspects of translational regulation of the genes of segment S. The first is the previously discussed translational coupling of genes 8-12 and 9-5(11). The second is that the products of these genes are not made in equimolar amounts, as approximately 10-fold more P8 and P9 are made than P12 and P5(P11) (42). It is clear that regulation must be at the level of translation, because one polycistronic mRNA (single-stranded RNA) molecule is made from this segment, and none of the nonsense mutants result in truncated segment S single-stranded RNA molecules (41). Furthermore, there is no polarity between genes 8 and 9, which would be expected if the 8-12 polarity was a transcriptional coupling.

A model to account for the translational couplings is based on the lack of properly positioned SD sequences preceding the ORFs assigned to P12 and P5(P11). Both visual inspection of the sequences 5′ to these genes and the PERCEPTRON algorithm analysis support this proposal. It is proposed that ribosomes do not bind de novo to the mRNA at the initiation regions of genes 12 and 5(11); only those ribosomes which are already bound to the mRNA can translate these genes. Thus, translation of the upstream genes 8 and 9 provides a method of delivering ribosomes to genes 12 and 5(11), respectively. In the case of genes 8 and 12, the overlapping termination-initiation codons provide a method for delivery of ribosomes to the initiation codon of gene 12. Genes 9 and 5(11) do not overlap; the ORF for gene 5(11) begins either 6 or 30 bp from the termination codon of gene 9. In this case the ribosomes that terminate gene 9 translation are not delivered to the initiation codon of gene 5(11) but to the vicinity of gene 5(11) initiation. Premature termination of gene 8 or gene 9 translation would not allow ribosomes to contact the initiation regions of gene 12 or gene 5(11), thereby accounting for the observed polarity.

This model requires that a fraction of the ribosomes (or possibly their 30S subunits) remains associated with the mRNA after the termination of translation of the preceding gene. It is anticipated that the majority of these ribosomes would be released from the mRNA and thus that translation via reinitiation would be less efficient than translation proceeding by de novo ribosome binding. The relative amounts of the proteins made are consistent with this prediction.

Another possibility to account for the translational coupling of these genes is that the products of genes 8 and 9 interact with the mRNA to activate the translation of genes 12 and 5(11), respectively. To eliminate this possibility for the 8-12 coupling, two constructs were made in which gene 8 sequences were removed, leaving gene 12 intact.

In deleting the gene 8 sequences, it was necessary, in the conceptual construct of the proposed model for the translational coupling, to supply some method of ribosomal delivery to the start of gene 12. In accordance with this requirement, translation of the *lacZ′* peptide of pUC8 was used to deliver ribosomes to gene 12 in the absence of gene 8 sequences. The 3.1-kb *Sca*I fragment of cDNA fragment F90 was ligated to the 1.7-kb *Sca*I-*Sma*I fragment of pUC8 (Fig. 3A). In this construct, designated pTM248, the C-terminal 14 amino acids of gene 8 were fused in frame to the N-terminal 7 amino acids of the *lacZ′* peptide of pUC8 (Fig. 4A). The integrity of this construct was confirmed by sequence analysis (see above). In pTM248 ribosomes are delivered to gene 12 by the de novo binding of the mRNA at the ribosomal binding site of the *lacZ′* peptide, followed by translation of a 21-amino-acid peptide, the last 14 amino acids of which are the C-terminal amino acids of P8. Thus, in this construct ribosomes terminate translation at the same position relative to that of the gene 12 initiation codon, as in the wild-type construct (Fig. 4A), the difference being that in pTM248 all but the C-terminal 14 amino acids of P8 have been deleted. Plasmid pTM248 was used to program an *E. coli* cell-free, coupled transcription-translation system. This clone directed the synthesis of P12 in the absence of P8 (Fig. 4B, lane E). Plasmid pTM248 also contained intact sequences of genes 9 and 5(11), both of which were synthesized by pTM248 in vitro (Fig. 4B, lane E). As mentioned above, P11 was synthesized in the in vitro system and P5 was not. The P11 band was not visible in this exposure of the gel. Although the results of these in vitro experiments are not quantitative, pTM248 appears to synthesize P12 in vitro in amounts equivalent to that of the wild-type construct (compare the ratio of P12 to P9 in lanes D and E, Fig. 4B).

We inserted the *lacZ*-polylinker region of pUC8 into plasmid RSF1010 (27). This plasmid, designated pLM254, can be propagated in the host of ϕ6, *P. phaseolicola*. We have shown that ϕ6 genes cloned into pLM254 (carried in *P. phaseolicola*) can complement ϕ6 mutations, thereby establishing the biological activity of the cloned genes (27). DNA fragments cloned into the polylinker region of pLM254 have the same structure relative to that of the *lacZ* peptide as in the pUC8 clones.

The result with pTM248 presented above establishes that P12 can be translated by an *E. coli* in vitro system in the absence of P8. To confirm that pTM248 makes biologically active P12 in the absence of *cis* P8, the pTM248 insert was transferred to pLM254. The 2,108-bp *Eco*RI (the site in the vector polylinker region)-*Pst*I insert of pTM248 was cloned into *Eco*RI-*Pst*I-cut pLM254. This clone was designated pTM260. Plasmid pLM254 contains the *lacZ′*-polylinker cloning region of pUC8; therefore, in pTM260 gene 12 has the same structure, relative to the *lacZ′* gene, as in pTM248. *P. phaseolicola* with plasmid pTM260 complemented P12 mutations with an efficiency of plating of 1 and, as expected, did not complement P8 mutations. Therefore, P8 is not required for the in vivo or in vitro synthesis of P12. Plasmid pTM260 also complemented mutations in P9 and P5(P11) with the same efficiency of plating as the wild-type construct (27).

To probe whether the overlap of gene 8 termination and gene 12 initiation codons is required for translation of gene 12, a construct was made in which gene 8 sequences were deleted and the *lacZ′* peptide was used for the delivery of ribosomes to gene 12, as in pTM248. However, in this construct ribosomes were not delivered to the AUG codon of P12 but 18 bp 5′ to it. The 4.8-kb *Eco*RV-*Sma*I fragment of cDNA fragment F90 was religated to itself (Fig. 3B). In this construct, designated pTM249, sequences 56 bp 5′ to the initiation codon of gene 12 were fused to the *Sma*I site of pUC8. The integrity of this construct was confirmed by sequence analysis (see above). In pTM249 the coding sequences for the N-terminal seven amino acids of the *lacZ′* peptide were fused to a reading frame that terminated 18 bp 5′ to gene 12 at UGA at N733 (Fig. 4A). In this construct, ribosomes bound to the mRNA at the SD sequence of *lacZ′* and translated a 17-amino-acid fusion peptide. Ribosomes were not delivered directly to the initiation region of P12, as

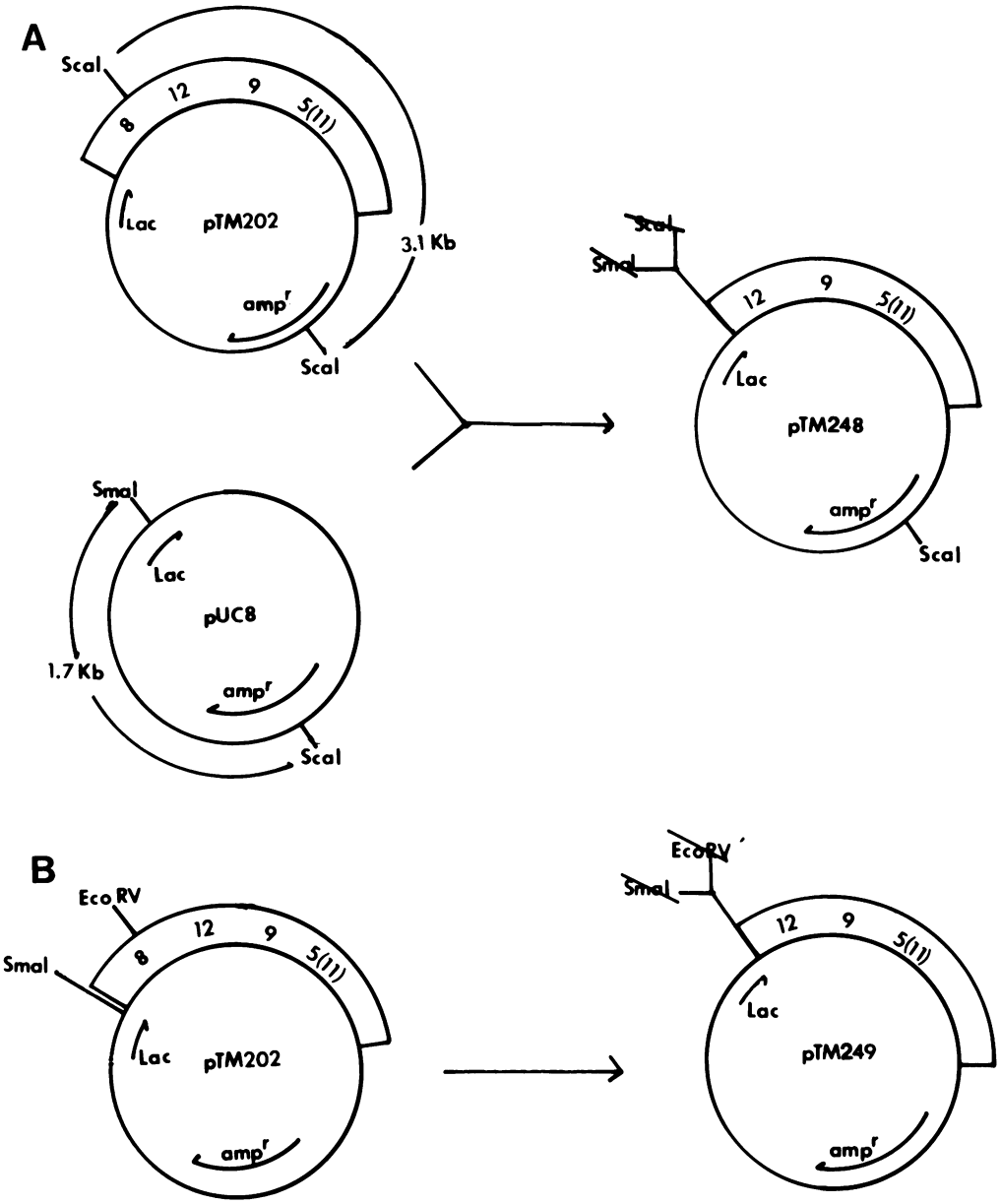
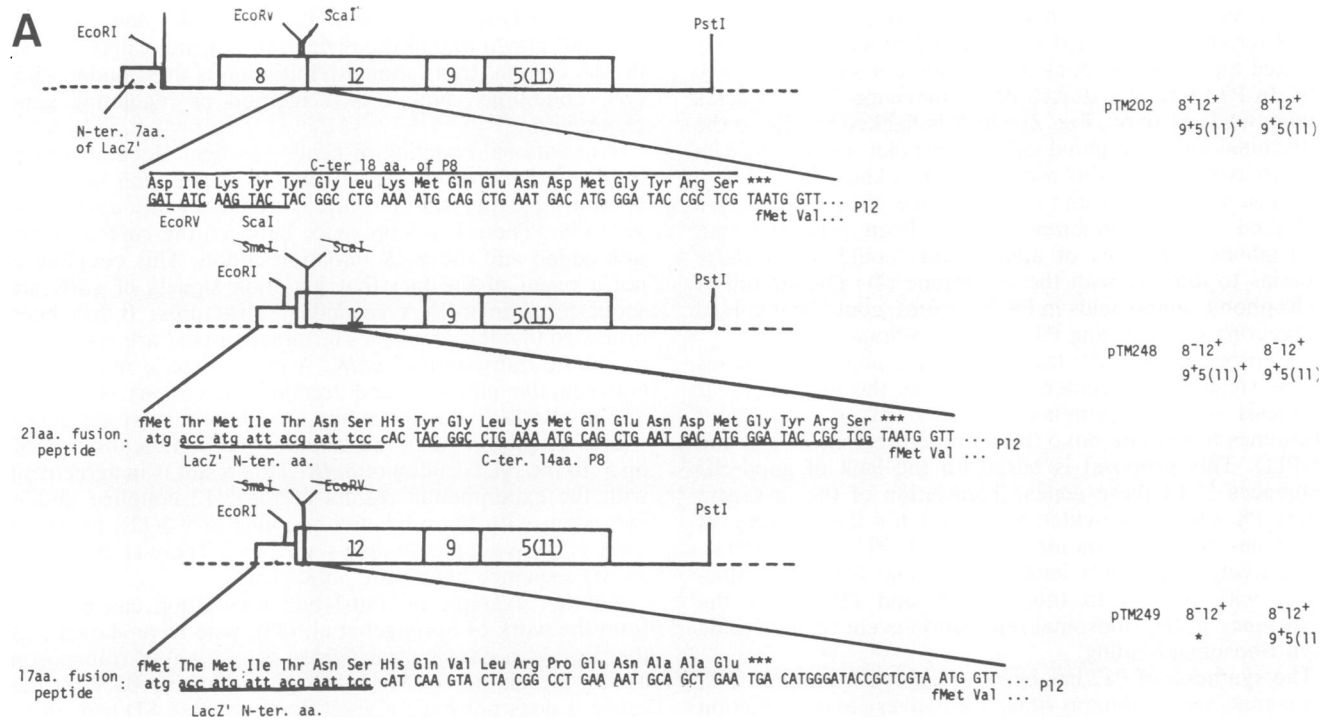


FIG. 3. Construction of two gene 8 deletion plasmids, pTM248 and pTM249. (A) The 3.1-kb *Sca*I fragment of pTM202 was ligated to the 1.7-kb *Sma*I-*Sca*I fragment of pUC8. This plasmid was designated pTM248. Cutting of both *Sca*I and *Sma*I resulted in blunt-ended DNA fragments; therefore, these ends could be joined by blunt-end ligation. The *Sma*I and *Sca*I sites with slashes are shown for orientational purposes only; the ligation does not re-form these sites. The integrity of the *Sma*I-*Sca*I junction was confirmed by sequence analysis (see text). (B) The 4.8-kb *Eco*RV-*Sma*I fragment of pTM202 was religated to itself. As indicated for panel A, both *Eco*RV and *Sma*I cutting resulted in blunt-ended DNA; therefore, these sites could be fused by blunt-end ligation. Sequence analysis confirmed the integrity of the *Eco*RV-*Sma*I junction. The *Sma*I and *Eco*RV sites with slashes are shown for orientational purposes only because these sites were not re-formed in the ligation. The integrity of the *Sma*I *Eco*RV junction was confirmed by sequence analysis (see text). In both panels the arrows indicate the direction of transcription from the *lac* promoter. The sites marked with slashes were destroyed in the construction process.

in the case of the wild-type construct and pTM248, but 18 bp 5′ to it. In the cell-free, coupled transcription-translation system this clone directed the synthesis of P12 in amounts equivalent to that of the wild-type construct (compare the ratio of P12 to P9 in lanes D and F, Fig. 4B).

The 1,133-bp *Eco*RI (the site in the vector polylinker region)-*Pst*I insert of pTM249 was transferred to pLM254. This construct, designated pTM261, when carried in *P. phaseolicola*, complements P12 mutations with an efficiency of plating of 1. Therefore, the requirement for the initiation of gene 12 translation in vivo and in vitro is the delivery of ribosomes to the vicinity of the start of gene 12, and the overlap of gene 8 termination and gene 12 initiation codons probably reflects a maximal use of sequence information. This result is consistent with the coupling of genes 9 and 5(11), in which there is no overlap of sequences.

Having established that the synthesis of P12 could be uncoupled from that of P8, it was possible to further confirm



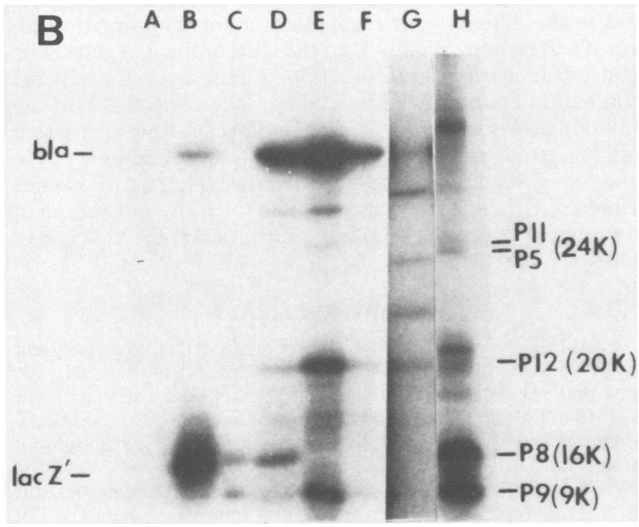


FIG. 4. Translation initiation in gene 8 deletion constructs pTM248 and pTM249. (A) Below the schematic illustrations of the wild-type (pTM202) and the gene 8 deletion constructs (pTM248 and pTM249), the nucleotide sequence of the gene 8-gene 12 overlap region is shown. In pTM202 the sequences coding for the C-terminal 18 amino acids (aa) of P8 and the first two amino acids of P12 are shown. Translation termination (ter) sites are indicated (***). In the deletion constructs, sequences of the *lacZ'* region are designated with lowercase letters. In both pTM248 and pTM249 a histidine codon was created at the fusion site. In vitro phenotypes were determined by in vitro translation (see panel B). In vivo phenotypes were determined by using the in vivo complementation test (27). For the complementation test the *Eco*RI-*Pst*I insert fragments of pTM248 and pTM249 were transferred to pLM254, in which they have the same structure, relative to the *lacZ'* peptide, as in the pUC8 constructs shown. Complementation of the wild-type construct was performed with the complete segment S cDNA clone pLM279 (27). The sites marked with slashes were destroyed in the construction process. (B) Autoradiogram of a 10 to 20% linear SDS-polyacrylamide gel of the *E. coli* in vitro-coupled, transcription-translation products of the following: translation system blank, no added DNA (lane A); pUC8 (lane B); ϕ6 in vivo-labeled protein standard pattern (lanes C and H); pTM202, a clone containing genes 8, 12, 9, and 5(11) (lane D); pTM248, a clone containing genes 12, 9, and 5(11) (lane E); pTM249, a clone containing genes 12, 9, and 5(11) (lane F); pTM265, a clone containing gene 12 (lane G). On this exposure the P11 band of pTM202, pTM248, and pTM249 is not visible. As has been previously described (27), almost no P5 is made in vitro. K, Molecular mass, in kilodaltons.

the assignment of the P12 ORF. To experimentally support the P12 ORF assignment, the 675-bp *Eco*RI-*Xmn*I (N1376; base position in complete segment S) fragment of pTM248 that contains the ORF assigned to P12 was subcloned into vector *Eco*RI-*Sma*I-cut pUC8. This fragment did not contain any ORFs, other than that assigned to P12, that were longer than 56 amino acids. The pUC8 subclone, designated pTM265, directed the in vitro synthesis of P12 (Fig. 4B, lane G). In pTM265 gene 12 had the same structure, relative to *lacZ*, as in pTM248. This result confirms the P12 ORF assignment.

DISCUSSION

The complete nucleotide sequence of dsRNA segment S of bacteriophage ϕ6 was determined. The ORFs corresponding to the small segment proteins were assigned.

P9 is an intrinsic membrane protein; its removal from the phage envelope requires treatment with 1% Triton X-100. Protease digestion of the intact virion (45) indicates that P9 does not span the lipid bilayer. These results are consistent with the iodination studies of Van Etten et al. (49). In light of these results, it is not surprising that P9 does not contain a signal sequence. The N-terminal amino acid sequence of P9 revealed that only the initiating fMet is removed during maturation. This reading frame is closed four codons 5′ to this AUG, precluding the possibility of translational initiation at a distal codon followed by protein processing to give

the N-terminal Pro found in mature P9 (Fig. 2). Furthermore, the N-terminal amino acid sequence of P9, containing many charged amino acids, does not resemble a signal sequence (39). In P9 there is a stretch of 16 noncharged amino acids (amino acids 51 to 66; Fig. 2) which is flanked by Glu to the N-terminal and C-terminal sides. The polar amino acid Gln appears twice within this region. Little is known about how proteins, which insert into only one face of the bilayer, are delivered to the membrane. It has been proposed that hydrophobic stretches of amino acids could direct these proteins to interact with the membrane (4). The stretch of hydrophobic amino acids in P9, therefore, could be involved in directing or anchoring P9 to the membrane.

It is proposed that the translational coupling of the genes of the small RNA segment, as well as the nonequimolar synthesis of these proteins, results from an inability of ribosomes to bind de novo the initiation regions of P12 and P5(P11). This proposal is based on the lack of good SD sequences 5′ to these genes. Translation of the upstream genes P8 and P9 provides a method for the delivery of ribosomes to the initiation regions of P12 and P5(P11), respectively. The lower levels of P12 and P5(P11) synthesized, with respect to those of P8 and P9, reflect the inefficiency of the ribosomal reinitiation event relative to de novo ribosomal binding.

The synthesis of P12 has been uncoupled from that of P8 in a manner that maintains ribosome delivery to the initiation region of gene 12. Furthermore, it has been shown that ribosomal delivery to the AUG codon of gene 12 is not a strict requirement for initiation of gene 12 translation. Rather, ribosomes delivered to within 17 bp of the AUG codon are capable of directing wild-type levels of P12 synthesis.

The question remains as to why ribosomes are not able to bind productively to gene 12. There is a purine-rich sequence, GGGA, preceding gene 12 which is capable of base pairing to the 3′ end of 16S RNA. This sequence is 11 bp from the initiating codon of gene 12 and therefore is not properly spaced to serve as a good SD sequence. However, there are examples of putative SD sequences which are more poorly spaced than the polypurine stretch preceding gene 12 (47). Furthermore, although the presence and spacing of an SD sequence is involved in controlling translation initiation, the exact relationship between these features and the level of initiation are ill defined (12).

Although the secondary structure of the 8-12 overlap region predicted by the algorithm described by Zucker and Stiegler (54) is not typical of those structures found to be involved in translational coupling (15), this is not a structureless region. It is possible that the polypurine stretch preceding gene 12 can serve as an SD sequence for de novo ribosomal binding and that mRNA secondary structure is involved in the coupling. The lower expression of gene 12, relative to that of gene 8, could be a result of the relatively poor gene 12 SD sequence. However, the simplest explanation for our results is that, in both P12 and P5(11), initiation of translation is due to ribosomes that arrive near the initiating codon through their involvement in the synthesis of the preceding gene products P8 and P9, respectively.

Translation reinitiation after premature termination at nonsense codons has been documented in *lacI*, *lacZ*, and bacteriophage T4 rIIB RNA (24, 30, 44). Although several initiation sites have been identified and there is a certain selectivity in choosing these sites, the sequences (or possibly, mRNA secondary structures) that signal reinitiation have not been determined. In these systems translational reinitiation is not the normal state of affairs and probably does not play a role in the regulation of gene expression. In the ϕ6 system, translational reinitiation is used under wild-type conditions, and it is a method of regulating gene expression.

Translational coupling of genes has been documented in the bacterial operons (1, 32, 36). In the *E. coli* galactose operon the expression of *galK* is translationally coupled to *galT* (36). There is a 3-bp space between the *galT* termination codon and the *galK* initiation codon. This coupling is not a result of the fact that initiation signals of *galK* are sequestered in mRNA secondary structures. It has been proposed that the ribosomes terminating *galT* are positioned to initiate translation of *galK*. A precise steric relationship between the initiation and termination codons is not required. Rather, the preceding translation must terminate within a region close to the initiation site, on the order of 50 bp 5′ to the AUG codon of *galK*. This result is in agreement with the experimental results of the 8-12 coupling and is consistent with the predicted coupling of 9-5(11), in which there is no overlap of coding sequences. There is, however, an SD sequence before the *galK* gene.

Another example of translation reinitiation can be seen from the work of Sprengel et al. (43), who have shown that the neomycin phosphotransferase gene (*neo*) of transposon Tn*5*, which is not translated in gram-positive bacteria because it does not have a good gram-positive SD sequence, can be translated in *Bacillus subtilis* if ribosomes are delivered to the vicinity of the initiation codon by prior translation of a gene immediately 5′ to the start of the *neo* gene. The most efficient translation of the *neo* gene occurs when the termination codon of the preceding gene overlaps with the initiation codon of the *neo* gene. These results are consistent with our proposed model for the translational coupling of the genes of ϕ6 segment S, and they demonstrate that ribosomes delivered to the initiation region of an ORF, by translation of an upstream gene, can reinitiate translation in the absence of an SD sequence.

ACKNOWLEDGMENTS

Parts of this work were supported by Public Health Service grant GM-31709 from the National Institutes of Health to L.M. T.M. was supported in part by Public Health Service training grant 5T32-AI-07180 from the National Institutes of Health awarded to the Department of Microbiology, New York University School of Medicine.

We thank J. Rocca-Serra for advice on RNA sequencing, B. Stitt and S. Projan for helpful discussions, J. Mindich for preparation of the purified P9, and A. Howard for preparation of the manuscript.

LITERATURE CITED

1. **Aksoy, S., C. L. Squires, and C. Squires.** 1984. Translational coupling of the *trpB* and *trpA* genes in the *Escherichia coli* tryptophan operon. J. Bacteriol. **157**:363–367.
2. **Bamford, D., and L. Mindich.** 1980. Electron microscopy of cells infected with nonsense mutants of bacteriophage ϕ6. Virology **107**:222–228.
3. **Birnboim, H. C., and I. Doly.** 1979. A rapid alkaline extraction procedure for screening recombinant plasmid DNA. Nucleic Acids Res. **7**:1513–1523.
4. **Blobel, G.** 1980. Intracellular protein topogenesis. Proc. Natl. Acad. Sci. USA **77**:1496–1500.
5. **Clewell, D. B.** 1972. Nature of ColE1 plasmid replication in *Escherichia coli* in the presence of chloramphenicol. J. Bacteriol. **110**:667–672.
6. **Cuppels, D. A., J. L. Van Etten, D. E. Burbank, L. C. Lane, and A. K. Vidaver.** 1980. In vitro translation of the three bacteriophage ϕ6 RNAs. J. Virol. **35**:249–251.

7. **Day, L. A., and L. Mindich.** 1980. The molecular weight of bacteriophage φ6 and its nucleocapsid. Virology **103**:376–385.
8. **Dunn, J. J., and F. W. Studier.** 1983. Complete nucleotide sequence of bacteriophage T7 DNA and the locations of T7 genetic elements. J. Mol. Biol. **16**:477–535.
9. **Emori, Y., H. Iba, and Y. Okada.** 1980. Assignment of viral proteins to the three double-stranded RNA segments of bacteriophage φ6 genome: translation of φ6 messenger RNAs transcribed *in vitro*. Mol. Gen. Genet. **180**:385–389.
10. **Fickett, J. W.** 1982. Recognition of protein coding regions in DNA sequences. Nucleic Acids Res. **10**:5303–5318.
11. **Frangione, B., E. Rosenwasser, F. Prelli, and E. C. Franklin.** 1980. Primary structure of human γ3 immunoglobulin deletion mutant: γ3 heavy-chain disease protein Wis. Biochemistry **19**:4304–4308.
12. **Gold, L., D. Pribnow, T. Schneider, S. Shinedling, B. S. Singer, and G. Stormo.** 1981. Translational initiation in prokaryotes. Annu. Rev. Microbiol. **35**:365–404.
13. **Gopinathan, K. P., L. A. Weymouth, T. A. Kunkel, and L. A. Loeb.** 1979. Mutagenesis *in vitro* by DNA polymerase from an RNA tumour virus. Nature (London) **278**:857–859.
14. **Grutter, M. G., L. H. Weaver, and B. W. Matthews.** 1983. Goose lysozyme structure: an evolutionary link between hen and bacteriophage lysozymes. Nature (London) **303**:828–830.
15. **Gryczan, T. J., G. Grandi, J. Hahn, R. Grandi, and D. Dubnau.** 1980. Conformational alteration of mRNA structure and the posttranscriptional regulation of erythromycin-induced drug resistance. Nucleic Acids Res. **8**:6081–6097.
16. **Hamlyn, P. H., G. G. Brownlee, C. G. Cheng, M. J. Gait, and C. Milstein.** 1978. Complete sequence of constant and 3′ noncoding regions of an immunoglobulin mRNA using the dideoxynucleotide method of RNA sequencing. Cell **15**:1067–1075.
17. **Iba, H., T. Watanabe, Y. Emori, and Y. Okada.** 1982. Three double-stranded RNA genome segments of bacteriophage φ6 have homologous terminal sequences. FEBS Lett. **141**:111–115.
18. **Ikemura, T.** 1981. Correlation between abundance of *Escherichia coli* transfer RNAs and the occurrence of the respective codons in its protein genes. J. Mol. Biol. **146**:1–21.
19. **Ikemura, T.** 1981. Correlation between the abundance of *E. coli* tRNAs and the occurrence of the respective codons in its protein genes: a proposal for a synonymous codon choice that is optimal for the *E. coli* translation system. J. Mol. Biol. **151**:389–409.
20. **Kakitani, H., H. Iba, and Y. Okada.** 1980. Penetration and partial uncoating of bacteriophage φ6 particle. Virology **101**:475–483.
21. **Lehman, J. F., and L. Mindich.** 1979. The isolation of new mutants of bacteriophage φ6. Virology **97**:164–170.
22. **Maxam, A. M., and W. Gilbert.** 1980. Sequencing end-labeled DNA with base-specific chemical cleavages. Methods Enzymol. **65**:499–560.
23. **McGraw, T., H. Yang, and L. Mindich.** 1983. Establishment of a physical and genetic map for bacteriophage PRD1. Mol. Gen. Genet. **190**:237–244.
24. **Michaels, C., and D. Zipser.** 1969. Mapping of polypeptide reinitiation sites within the β-galacotosidase structural gene. J. Mol. Biol. **41**:341–347.
25. **Mindich, L., J. Cohen, and M. Weisburd.** 1976. Isolation of nonsense supressor mutants of *Pseudomonas*. J. Bacteriol. **126**:177–182.
26. **Mindich, L., and J. Lehman.** 1979. Cell wall lysin as a component of the bacteriophage φ6 virion. J. Virol. **30**:489–496.
27. **Mindich, L., G. MacKenzie, J. Strassman, T. McGraw, S. Metzger, M. Romantschuk, and D. Bamford.** 1985. cDNA cloning of portions of the genome of bacteriophage φ6. J. Bacteriol. **162**:992–999.
28. **Mindich, L., and T. McGraw.** 1983. Molecular cloning of bacteriophage PRD1 genomic fragments. Mol. Gen. Genet. **190**:233–236.
29. **Mindich, L., J. F. Sinclair, and J. Cohen.** 1976. The morphogenesis of bacteriophage φ6: particles formed by nonsense mutants. Virology **75**:224–231.
30. **Napoli, C., L. Gold, and B. S. Singer.** 1981. Translational reinitiation in the rIIB cistron of bacteriophage T4. J. Mol. Biol. **149**:433–449.
31. **Nichols, B. P., and C. Yanofsky.** 1979. Nucleotide sequence of the *trp*A of *Salmonella typhimurium* and *Escherichia coli*: an evolutionary comparison. Proc. Natl. Acad. Sci. USA **76**: 5244–5248.
32. **Oppenheim, D. S., and C. Yanofsky.** 1980. Translational coupling during expression of the tryptophan operon of *Escherichia coli*. Genetics **95**:785–795.
33. **Partridge, J. E., J. L. Van Etten, D. E. Burbank, and A. K. Vidaver.** 1979. RNA polymerase activity associated with bacteriophage φ6 nucleocapsid. J. Gen. Virol. **43**:299–307.
34. **Sanger, F., A. R. Coulson, G. F. Hong, D. F. Hill, and G. B. Petersen.** 1982. Nucleotide sequence of bacteriophage λ DNA. J. Mol. Biol. **162**:729–773.
35. **Sanger, F., S. Nicklen, and A. R. Coulson.** 1977. DNA sequencing with chain-terminating inhibitors. Proc. Natl. Acad. Sci. USA **74**:5463–5467.
36. **Schumperli, D., K. McKenney, D. A. Sobieski, and M. Rosenberg.** 1982. Translational coupling at an intercistronic boundary of the *Escherichia coli* galactose operon. Cell **30**:865–871.
37. **Semancik, J. S., A. K. Vidaver, and J. L. Van Etten.** 1973. Characterization of a segmented double-helical RNA from bacteriophage φ6. J. Mol. Biol. **78**:617–625.
38. **Shine, J., and L. Dalgarno.** 1975. Determinant of cistron specificity in bacterial ribosomes. Nature (London) **254**:34–38.
39. **Silhavy, T. J., S. A. Bensen, and S. D. Emr.** 1983. Mechanisms of protein localization. Microbiol. Rev. **47**:313–344.
40. **Sinclair, J. F., J. Cohen, and L. Mindich.** 1976. The isolation of suppressible nonsense mutants of bacteriophage φ6. Virology **75**:198–208.
41. **Sinclair, J. F., and L. Mindich.** 1976. RNA synthesis during infection with bacteriophage φ6. Virology **75**:209–217.
42. **Sinclair, J. F., A. Tzagoloff, D. Levine, and L. Mindich.** 1975. Proteins of bacteriophage φ6. J. Virol. **16**:685–695.
43. **Sprengel, R., B. Reiss, and H. Schaller.** 1985. Translationally coupled initiation of proteins in *Bacillus subtilis*. Nucleic Acids Res. **13**:893–909.
44. **Steege, D. A.** 1977. 5′-Terminal nucleotide sequence of *Escherichia coli* lactose repressor mRNA: features of translational initiation and reinitiation sites. Proc. Natl. Acad. Sci. USA **74**:4163–4167.
45. **Stitt, B. L., and L. Mindich.** 1983. The structure of bacteriophage φ6: protease digestion of φ6 virions. Virology **127**: 459–462.
46. **Stormo, G. D., T. D. Schneider, L. Gold, and A. Ehrenfreucht.** 1982. Use of the Perceptron algorithm to distinguish translational initiation sites in *E. coli*. Nucleic Acids Res. **10**: 2997–3011.
47. **Stormo, G. D., T. D. Schneider, and L. M. Gold.** 1982. Characterization of translational initiation sites in *E. coli*. Nucleic Acids Res. **10**:2971–2996.
48. **Studier, F. W.** 1973. Analysis of bacteriophage T7 early RNA's and proteins on slab gels. J. Mol. Biol. **29**:237–248.
49. **Van Etten, J., L. Lane, C. Gonzalez, J. Partridge, and A. Vidaver.** 1976. Comparative properties of bacteriophage φ6 and φ6 nucleocapsid. J. Virol. **18**:652–658.
50. **Vidaver, A. K., R. K. Koski, and J. L. Van Etten.** 1973. Bacteriophage φ6: a lipid-containing virus of *Pseudomonas phaseolicola*. J. Virol. **11**:799–805.
51. **Vieira, J., and J. Messing.** 1982. The pUC plasmids, an M13mp7-derived system for insertional mutagenesis and sequencing with synthetic universal primers. Gene **19**:269–276.
52. **Woese, C. R., P. Blanz, and C. M. Hahn.** 1984. What isn't a pseudomonad: the importance of nomenclature in bacterial classification. Syst. Appl. Microbiol. **5**:179–195.
53. **Zubay, G.** 1973. *In vitro* synthesis of protein in microbial systems. Annu. Rev. Genet. **7**:267–287.
54. **Zucker, M., and P. Stiegler.** 1981. Optimal computer folding of large RNA sequences using thermodynamics and auxiliary information. Nucleic Acids Res. **9**:133–148.