JOURNAL OF BACTERIOLOGY, Nov. 1996, p. 6158–6165 Vol. 178, No. 21
0021-9193/96/$04.00+0

Gene Organization and Transcription of a Late-Expressed Region of a *Lactococcus lactis* Phage

RICARDO PARREIRA, RUDY VALYASEVI,† ALDA L. S. LERAYER,‡
S. DUSKO EHRLICH, AND MARIE-CHRISTINE CHOPIN*

Institut National de la Recherche Agronomique, Laboratoire de Génétique Microbienne, 78352 Jouy-en-Josas Cedex, France

Received 8 April 1996/Accepted 15 August 1996

The lactococcal phage bIL41 belongs to the small isometric-headed phages of the 936 quasi-species and is resistant to the abortive infection determined by *abiB*. A 10.2-kb segment from this phage, in which late transcription is initiated, has been sequenced. Thirteen open reading frames (ORFs) organized in one transcriptional unit have been identified. The location of two of them and the structural features of the proteins they code for are evocative of terminase subunits. Five other ORFs specify proteins which are highly homologous to structural proteins from the closely related phage F4-1. By comparing the phage bIL41 sequence with partial sequences available for four related phages, we were able to deduce a chimerical phage map covering the middle- and a large part of the late-expressed regions. Phages from this quasi-species differ by the insertion or deletion of either 1 to about 400 bp in noncoding regions or an entire ORF. Transcription was initiated 9 min after infection at a promoter with a −10 but no −35 consensus sequence. Synthesis of a phage activator protein was needed for initiation of transcription. A large 16-kb transcript covering all of the late-expressed region of the genome was synthesized. This transcript gave rise to smaller units. One of these units most probably resulted from a RNase E processing.

Lactococci are widely used for the manufacture of dairy products. During the fermentation processes, they are challenged with high concentrations of a variety of phages. These highly selective environmental conditions are most probably responsible for the selection of strains with diverse phage defense mechanisms, including adsorption interference, restriction and modification, and abortive infection (38). Lactococci represent, therefore, a good model for studying bacteria or phage coevolution and identifying phage defense mechanisms. A variety of phage abortive infection mechanisms have been described in *Lactococcus lactis* (2, 24, 25, 30, 32, 40, 54). All of them are active against phages of one of the prevailing groups in the dairy environment, designated species 936 (42). A model has been proposed for two of these mechanisms, designated AbiB (24) and AbiD1 (2). AbiB was shown to prevent phage growth by promoting a dramatic degradation of phage transcripts 10 min after infection (57). It was proposed that phage infection either induces the synthesis or stimulates the activity of an RNase in AbiB cells (57). AbiD1 was proposed to prevent phage growth by decreasing the amount of an essential phage product encoded by middle-expressed genes (9). However, the molecular basis of phage-Abi interactions has not yet been elucidated, not only because of their complexity but also because of the scarcity of data available on lactococcal phages of the 936 species.

The lactococcal 936 phage species will be designated as a quasi-species in this paper, as suggested by Casjens et al. (15) for the lambdoid phages. Isometric-headed phages from the 936 quasi-species have a double-stranded DNA genome of about 30 kb and carry cohesive ends (42). The genome has, most probably, a modular organization. Such organization, originally described for lambdoid phages (13), was suggested by heteroduplex studies (43, 49) which revealed homologous regions covering from 81 to 94% of the genome and distinct regions of nonhomology. Recently, the overall transcription map of phage sk1 established that this phage fits the usual subdivision into temporal units designated as early, middle, and late (8, 18). The middle-expressed region has been characterized on phage bIL66 (9). It codes for four open reading frames (ORFs) organized in an operon transcribed about 10 min after infection. This operon, involved in sensitivity to the abortive infection mechanism AbiD1, codes for a product essential for phage growth (9). In contrast, there are no data on the early-expressed region of the phage, and the late one is poorly characterized. Two genes from phage P008 have been proposed to code for terminase subunits (10). A block of genes coding for four minor structural proteins (46) and the major capsid protein (23) has been characterized in phage F4-1. The lysin gene has been characterized in phage US3 (58).

In this paper, we present the organization and the transcription map of a late-expressed segment of a phage, from the quasi-species 936, designated bIL41. Thirteen ORFs have been identified on a 10.2-kb segment. Two of them might code for terminase subunits homologous to those of P008 (10). Five others specify proteins highly homologous to the major capsid protein (23) and to minor structural proteins identified in the closely related phage F4-1 (46). Upstream of these ORFs, a region homologous to those from bIL66 (9) and sk1 (19) that overlaps the *cos* site has been localized. Therefore, the sequence of the bIL41 segment established a link between several scattered sequences from related phages, thus providing insight into genome organization and evolution of this phage quasi-species. Data on genome organization allowed the determination of some features of the transcription of late-expressed genes. This transcription, initiated 9 min after infection, requires a phage activator protein to be effective. It yields

* Corresponding author. Phone: 33 01 34 65 25 30. Fax: 33 01 34 65 25 21. Electronic mail address: mchopin@biotec.jouy.inra.fr.

† Present address: National Center for Genetic Engineering and Biotechnology, National Science and Technology Department Agency, Bangkok 10400, Thailand.

‡ Present address: Instituto de Tecnologia de Alimentos, Campinas S.P., Brazil.

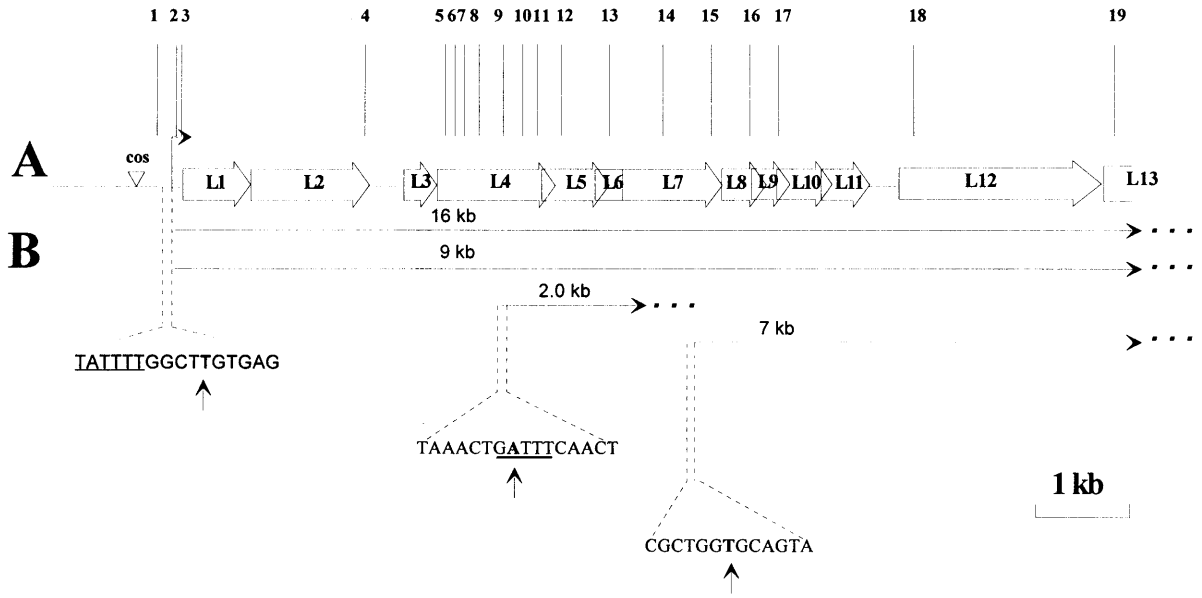

FIG. 1. (A) Genetic organization of a phage bIL41 sequenced segment. The promoter is indicated by a bent arrow. cos, cohesive site. (B) Transcription map of the sequenced segment. The transcript size is indicated. Vertical bars, identified by numbers, indicate the positions of oligonucleotides used as probes. The start sites of the transcripts are indicated by arrows below the corresponding sequences. The −10 box of the promoter and a consensus sequence for the RNase E cleavage site are underlined.

a large 16-kb transcript which covers all of the late-expressed region of the genome and is processed in smaller units.

MATERIALS AND METHODS

Bacterial strains, phage, and media. *L. lactis* subsp. *lactis* IL1403 (21) was grown at 30°C in M17 medium (63) in which lactose has been replaced by glucose (M17glc). *Escherichia coli* TG1 (33) was grown at 37°C in Luria-Bertani medium (51). When needed, ampicillin (100 μg/ml), erythromycin (5 μg/ml), or tetracycline (10 μg/ml) was added to the culture medium. Phage bIL41 was from our laboratory collection and was enumerated as described previously (63).

Phage DNA replication. Cells were grown at 30°C in M17glc to an optical density at 600 nm of 0.5 to 0.6 before they were infected with phage at a multiplicity of infection of 5. Samples of 2 ml were taken at different times throughout the multiplication cycle, and DNA was extracted as described by Hill et al. (39). The DNA was transferred to nylon Hybond-N^+ membranes (Amersham, Arlington Heights, Ill.) essentially as described by Maniatis et al. (51) and hybridized at high stringency with bacteriophage DNA labelled with [α-^{32}P] dCTP through the use of a random priming kit (Amersham).

Molecular cloning and DNA sequence analysis. Transformation of *L. lactis* was done by electroporation as described previously (41). DNA manipulation, transformation of *E. coli* cells, and cloning were essentially as described by Maniatis et al. (51). *E. coli* clones for sequencing were obtained by subcloning specific DNA fragments in the pBluescript $SKII^+$ plasmid (Stratagene, La Jolla, Calif.) and by using DNase I to generate a series of overlapping clones. The nucleotide sequence was determined on double-stranded DNA in a cycle extension reaction with *Taq* DNA polymerase and fluorescent dye-coupled primers on a 373 DNA sequencer (Applied Biosystems, San Jose, Calif.). A 1.1-kb region missing from the overlapping clones, the 1.2-kb segment located upstream from the *cos* site, and a 2-kb segment covering *orf12* were synthesized by a PCR and sequenced by a series of cycle extension reactions by using appropriate primers, *Taq* polymerase, and fluorescent dye-coupled dideoxynucleotides (Applied Biosystems). The DNA and protein sequences were analyzed with the Genetics Computer Group software (29) and the Blast program (1).

TABLE 1. General features of the phage bIL41 ORFs

Gene designation	Start[a]	Stop[a]	Product size (aa)	Similarities	Translation start[b]
orf1	1421 (AUG)	1945 (UAA)	174	Terminase[c]	AGAAAGGAcaau AUG CAA ACA
orf2	1946 (AUG)	3217 (UGA)	423	Terminase[c]	GAAAGGuuugaauaa AUG UAU UAU
orf3	3572 (AUG)	3856 (UAG)	94		AAAtGGcGGtgtcag AUG AAG UAU
orf4	3869 (UUG)	5005 (UAG)	378		AGAAAGGgaaaaa UUG AAC UUA
orf5	4986 (AUG)	5522 (UUA)	178		AGAAAGGAcguaacaagcacag AUG AAA CUA
orf6	5515 (AUG)	6696 (UAA)	393	Structural proteins[d]; P43	GAGGAuaucaaAAAGaaau AUG AAU AAA
orf7	5740 (AUG)	6696 (UAA)	318	P35	AAAGGAaaagcuaaaa AUG ACA AAC
orf8	6717 (AUG)	6980 (UGA)	87	P9.5	GGAGGgaauaa AUG AUA GAU
orf9	6980 (AUG)	7294 (UGA)	104	P11	GGAGGuauuug AUG AUA UUU
orf10	7284 (AUG)	7619 (UAA)	111		GAGGgagucguGAGca AUG GAA UUU
orf11	7613 (AUG)	7987 (UAA)	124		AGGuGGucaagc AUG GUA AUG
orf12	8326 (UUG)	10086 (UAA)	586		AGGAGuua UUG AAG GCG
orf13	10110 (AUG)	>10170	>19 (16S rRNA)	MCP[e] (3′ UCUUUCCUCCA 5′)[f]	AAAGGAaaacaaaaa AUG GAA UUA

[a] Numbers refer to positions in the sequence; start and stop codons are shown in parentheses.
[b] The ribosome binding sites and the beginnings of the ORFs are shown in capital letters.
[c] From P. Billard (10).
[d] From Chung et al. (23).
[e] Major capsid protein, from Kim and Batt (45).
[f] From Ludwig et al. (50).

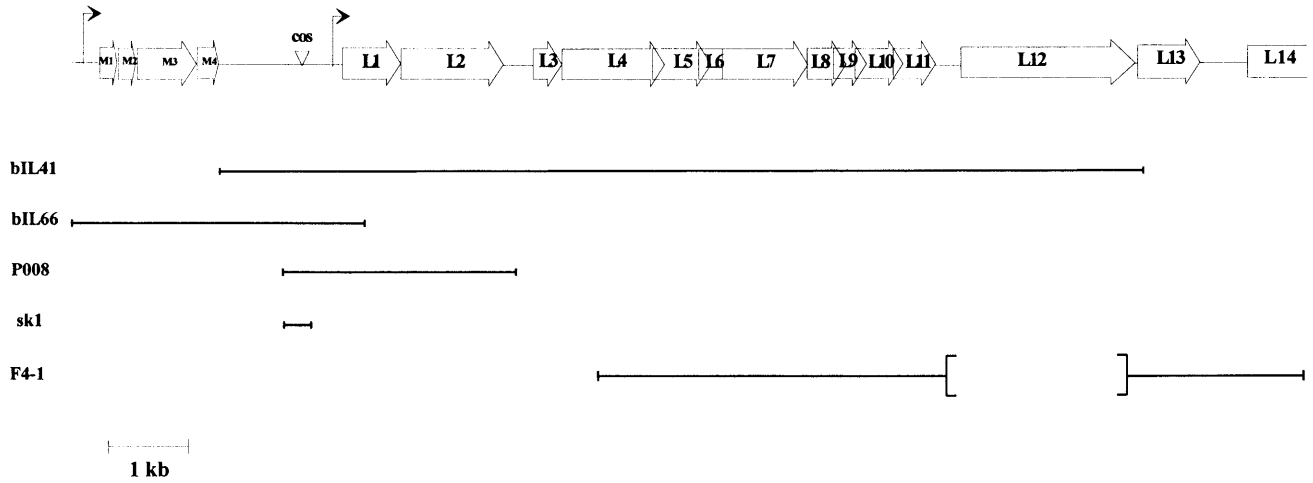

FIG. 2. Organization of a chimerical phage genome from the 936 quasi-species obtained by sequence comparison of five phages: bIL41 (L35061) (this study), bIL66 (L35175) (21), F4-1 (M37979) (23, 46), P008 (10), and sk1 (U01068) (19). The sequenced regions are represented by horizontal lines. The phage bIL41 sequence matched at positions 1 to 1319 with the phage bIL66 sequence (positions 1459 to 2385), at positions 443 to 831 with the phage sk1 sequence (positions 46 to 443), at positions 442 to 3144 with the phage P008 sequence (positions 1 to 2327), and at positions 4681 to 7989 and 10109 to 10173 with the phage F4-1 sequence (positions 1 to 3278 and 3287 to 3316, respectively). A DNA segment, absent on F4-1, is in brackets.

RNA extraction, labelling, and primer extension analysis. Total RNA was extracted from phage-infected cultures of *L. lactis* subsp. *lactis* as described by Glatron and Rapoport (34), with previously described modifications (2). The quality of RNA preparations was checked by methylene blue staining (69).

Northern (RNA) blot experiments were carried out as described by Williams and Mason (71). Standard RNA weight markers were from GIBCO-BRL (Gaithersburg, Md.). The oligonucleotides used as probes were labelled on their 5′ end with [γ-^{32}P]dATP by using polynucleotide kinase (Boehringer GmbH, Mannheim, Germany). These oligonucleotides, complementary to the L35061 sequence, were as follows: 1, 5′ ATCAGATTGAACCAGCTA 3′ (positions 58 to 41); 2, 5′ CTCTTAGGTAACTCTGCTCT 3′ (positions 171 to 152); 3, 5′ ATG TGGTTATATCGTTTA 3′ (positions 243 to 226); 4, 5′ TCATAGCGCGCTG GGTCATA 3′ (positions 2118 to 1998); 5, 5′ TTGCCACGCTGTAACTCTTT 3′ (positions 2786 to 2767); 6, 5′ AGTATCAGAGCCAACATCAG 3′ (positions 2999 to 2889); 7, 5′ CTATTGCGTTCACCCTTAGA 3′ (positions 2977 to 2958); 8, 5′ TTGAATGCCTGCCAACGCAT 3′ (positions 3191 to 3172); 9, 5′ CCTG TGAAGCTGTACCAAGC 3′ (positions 3474 to 3455); 10, 5′ GCTATGTAAA CATCTCCACC 3′ (positions 3754 to 3735); 11, 5′ CATCTGTGCTTGTTACG TCC 5′ (positions 3828 to 3809); 12, 5′ CTCTAGCGAACTCGATACAG 5′ (positions 3970 to 3951); 13, 5′ TTAAGTTCTGCCAAGCGGTT 5′ (positions 4404 to 4385); 14, 5′ GTATCAATAGTGAGTGTGGC 5′ (positions 4911 to 4892); 15, 5′ AAGAGTTTCAACATGACCGC 5′ (positions 5506 to 5487); 16, 5′ TTTCGCCATTGTTCTGCTGG 5′ (positions 5762 to 5743); 17, 5′ ATTCC ATTGCTCACGACTCC 5′ (positions 6130 to 6111); 18, 5′ TGCGTCTAATA GCGACTT 3′ (positions 8676 to 8659); 19, 5′ CAGCTACGATTAGAGCTTC ATTAC 3′ (positions 10173 to 10150).

The 5′ end of RNAs was mapped by the reverse transcriptase-directed primer extension method described by Williams and Mason (71). Synthetic oligonucleotides 3, 20 (5′ ATATTTTCATGATACAAGTC 3′ [positions 4836 to 4817]), and 21 (5′ GGAATACCACAATAGACC 3′ [positions 6754 to 6737]) were end labelled with ^{32}P and hybridized to RNA. The hybridized primers were extended with reverse transcriptase from avian myeloblastosis virus (Appligene, Pleasanton, Calif.) for 60 min at 42°C. The sequencing reaction was performed by the dideoxy method with the commercial kit Sequenase version 2 (United States Biochemical, Cleveland, Ohio).

Nucleotide sequence accession number. The sequence described in this paper has been assigned GenBank accession number L35061.

RESULTS

Sequence analysis. Assignment of the late-expressed region of phage bIL41 was based on a comparison with phages sk1 and bIL66. In sk1, middle-expressed transcripts terminate before the cohesive ends and late-expressed transcripts start after them (18). The middle-expressed region was localized previously on phage bIL66 (9). Restriction maps of phages bIL66 (9) and bIL41 were oriented by hybridization of subfragments (data not shown), and a 10.2-kb segment covering the cohesive ends and a large part of the late-expressed region of phage bIL41 was sequenced. The organization of the sequence is presented in Fig. 1A. Thirteen ORFs have been identified and confirmed by use of a program developed for gene recognition in newly sequenced DNA (11) which is based on nonhomogeneous Markov chains. The general features of these ORFs are summarized in Table 1. Translational signals follow those already described for *L. lactis* bacterial (20, 66) and phage (62) genes.

Proteins specified by *orf1* and *orf2* have 97 and 96% identity, respectively, with the products of *orf1* and *orf2* located in the same part of the genome of the closely related phage P008 (10). In addition, the protein specified by *orf2* has 24% identity (41% similarity) over 129 amino acids with protein gp32 from the unrelated *L. lactis* phage bIL67 (62). gp1 and gp2 from P008 and gp32 from bIL67 have all been suggested to be terminase subunits (10, 62). This is in agreement with the presence of two helix-turn-helix motifs (detected by the Chou and Fasman algorithm [22]) in *orf1* from bIL41 at positions 56 to 91 and 100 to 123. In addition, putative ATP-binding domains (68) A (G-3X-GKS-6X-V; amino acids [aa] 100 to 114) and B (R-3X-G-3X-LNIFD; aa 189 to 201) were identified in *orf2*. The products of *orf6*, *orf7*, *orf8*, and *orf9* have 96, 98, 98, and 96% identity with structural proteins P43, P35, P9.5, and P11 of the closely related phage F4.1 (45, 46), respectively. The predicted 19 N-terminal amino acids, except one, of the protein specified by *orf13* were identical to those from the major capsid protein of phage F4-1 (23). The protein specified by *orf12* (gp12) has identity with that encoded by *orf47* (gp47) of the unrelated lactococcal temperate phage r1t (67). A putative start of *orf12*, localized at position 8047 (AUG), was missing a consensus ribosome binding site sequence. Therefore, the start of *orf12* was assigned to a downstream UUG at position 8329. However, the homology between the bIL41 and r1t gene products starts at position 8047 on bIL41. The homology between gp12 and gp47 varied along the protein: there are 61.6, 90.2, 27, and 91% identities for gp12 regions of aa 1 to 100 (aa 1 to 100), aa 101 to 417 (aa 101 to 417), aa 418 to 580 (aa 418 to 567), and aa 581 to 680 (aa 568 to 667), respectively (gp47 regions are given in parentheses). gp12 presents six hydrophobic domains. The proteins specified by the other ORFs have no

```
                                                                                R4
sk1    GCAGTACTTATCCTATGTGACTATGTATATAACATTCAAGAACGGAAAAGAAATAATA
bIL41  .............................................AAAAGAAAACAGA
P008   ....................................................
bIL66  ....................................................

sk1    AATAAAATAAATAAAGAGATAAAGATAATTCTTTCAGAGAAAAAATAAAAAT......
bIL41  AATAGAA.....ATAGAATTAAAGAAGA.....TAGGAGAAGAAATAAAAATTAAAGA
P008   AAATAAATAAATAAAGATATAAAGGTATATCTTTCAGAGAGGAAATAAAAAT......
bIL66  .............AAGCAATAAAG.TAATT......GTGAAAAGAAAATAAAAT....

                                                                          IR
sk1    AAAATAATTTTTTTATATATACCCCCCCCATTAATCGCTATGTTAAGGGAAATTTTCA
bIL41  AAAATAATTTTCCT.TAGGTACCGCCCC.AT.AATCGCTATGTTAAGGGAAATTTTCA
P008   ......ATTTTTCCATAGGTACCCGCCCCAT.AATCGCTATGTTTAAGGAGATTTTCA
bIL66  ..AAAAATTTTCCCATAGGTACCCGCCCCATTAATCGCTATGTT.AAGGAGATTTTAA

          cos                IR                  R3
sk1    GCACAAAGGACTCCTCTTTGAAAATCTCTCAAAAAATCAAAAAGAAAAGTTTTAGCTGA
bIL41  GCACAAAGGACTCCCCTTTACAAATCTCCCAAAAAATCAAAAAGAAAAGTTTTAGCTGG
P008   GCACAAAGGACTCCCCTTTACAAATCTCCCAAAAAATCAAAAAGAAAAGTTTTAGCTGG
bIL66  GCACAAAGGACTCCCCTTTACAAATCTCCCAAAAAATCAAAAAGAAAAGTTTTAGCTAG

sk1    TTTAATCTAAATTAATCTATTTTAATCTAGTAAAATCTTGCTAAATCTTAATAATCTCA
bIL41  TTCAATCTGATTTAATCTATTTTAATCTAGTACAATCTTGCTAAATCTAATAATCTCA
P008   TTTAATCTAAATTAATCTATTTTAATCTAGCATGATCTTGCTAAATCTTA..AATC...
bIL66  TTCAATCTAAATTAATCTATTTTAATCTAGTACGATCTTGCTAAATCTTA..AATC...

          R2
sk1    AATCAAAAGGAAATTATATGGCTATATGCTCATATTTTGGCTCGTATGAGCGTGTTTAT
bIL41  AATCAAAAGGAA.TTATAGGGCTATATGCTCATATTTTGGCTTGTGAGAGCAGAGTTAC
P008   ..............ATAGGGCTATATGCTCATATTTTGGCTCGTATGAG.GTGTTTAT
bIL66  ..............ATAGGGCTATATGCTCATATTTTGGCTCGTATGAGCGTGTTTAT

sk1    ATCATTACGATATAATTCAAGTCTAATTAACTTAACGCTCACAGGAACGAAAATATAGG
bIL41  CT.AAGAGGGTAAACTTATAAGCAAAAGTTATTGCTGTTCTTAGAAACGAAAATATAGG
P008   ATCATTGCGATATAATTCAAGTCTAATTAACTTAACGCTCACAGGAACGAAAATATAGG
bIL66  ATCATTGCGATATAATTCAAGTCTAATTAACTTAACGCTCACAGGAACGAAAATATAGG

                                  R1
sk1    CTTTAAACGATATAACCACATTTAAAAGAAAGGATAATA
bIL41  CTTTAAACGATATAACCACATTTAAAAGAAAGGACAATA
P008   CTTTAAACGATATAACCACATTTAAAAGAGAGGACAATA
bIL66  CTTTAAACGATATAACCACATTTAAAAGAAAGGACAATA
```

FIG. 3. Sequence surrounding the *cos* site of the following phages of the 936 quasi-species: sk1 (U01068) (18), bIL41 (L35061) (this study), P008 (10), bIL66 (L35175) (9). The *cos* site, putative lambda-like R-sites, and direct and inverted repeats identified on sk1 (19) are in bold letters.

homology with known proteins, but a heme-binding signature (Cys-X-X-Cys-His [52]) was detected in *orf3.*

The protein homologies observed reflect a high level of sequence conservation between phages from the same quasi-species. Such sequence conservation allowed the joining of the middle-expressed region from bIL66 (9) and the late-expressed region from phage F4-1 (23, 46) to the sequenced bIL41 segment. The organization of a 14.5-kb chimerical phage genome region is shown in Fig. 2. The sequences of four phages, bIL41 (this study), bIL66 (9), P008 (10), and sk1 (19), all available around the *cos* site, were compared (Fig. 3). They differ by the insertion or deletion of short segments (1 to 400 nucleotides). The main structural features of this region, proposed to play a role in the packaging of phage sk1 (19), are conserved in the four phages. They are the *cos* site, two (R1 and R3) of the four putative terminase binding sites (lambda-like R-sites), five of the six AATCT sequences which may be involved in DNA bending, and the 10-bp inverted repeats located on each side of the *cos* site. In addition, phage bIL41 differs from phage F4-1 by the presence of *orf12* located between a gene cluster coding for minor structural proteins and the gene coding for the major capsid protein. The homology starts and stops at the ends of the ORF, and the breakpoints between segments are sharp, with no apparent gap or overlap.

Transcription map. Transcription of the sequenced region was monitored in *L. lactis* IL1403 infected with phage bIL41. Total cell RNA was extracted at different times after infection and analyzed by Northern hybridization with oligonucleotide 13 as a probe. Transcript number and size were determined on agarose gels. Three transcripts, of 16, 9.0, and 2.0 kb, were detected at about 20 min after infection, and their amounts were increased throughout the infection cycle (Fig. 4A). The position and form of the two other bands suggest that they correspond to a distortion of the electrophoretic migration of transcript breakdown products due to the presence of 23S and 16S rRNA. It is worth noting that the presence of the 16-kb transcript was unreliable and varied from one assay to another. In a previous study, late mRNAs from phage bIL170 were detected 20 min after infection on an agarose gel and about 5 min earlier on an acrylamide gel (57). Therefore, the time at which late-expressed transcripts from phage bIL41 start to be synthesized was determined with an acrylamide gel, which, however, cannot separate these transcripts because of their large size. The late-expressed transcripts appeared 9 min after infection (Fig. 4B). Since the middle-expressed transcripts from the related phage bIL66 were detected about 10 min after infection on an agarose gel (9), the appearance of bIL41 middle-expressed transcripts was also determined on an acrylamide gel. The probe used was complementary to the phage bIL66 sequence (9) since the sequence of the corresponding bIL41 region is not known. Middle-expressed transcripts were detected 6 to 7 min after infection (Fig. 4C). We concluded that the late bIL41 genes are expressed about 3 min later than the middle genes.

A detailed mapping of the transcripts carried out with 18

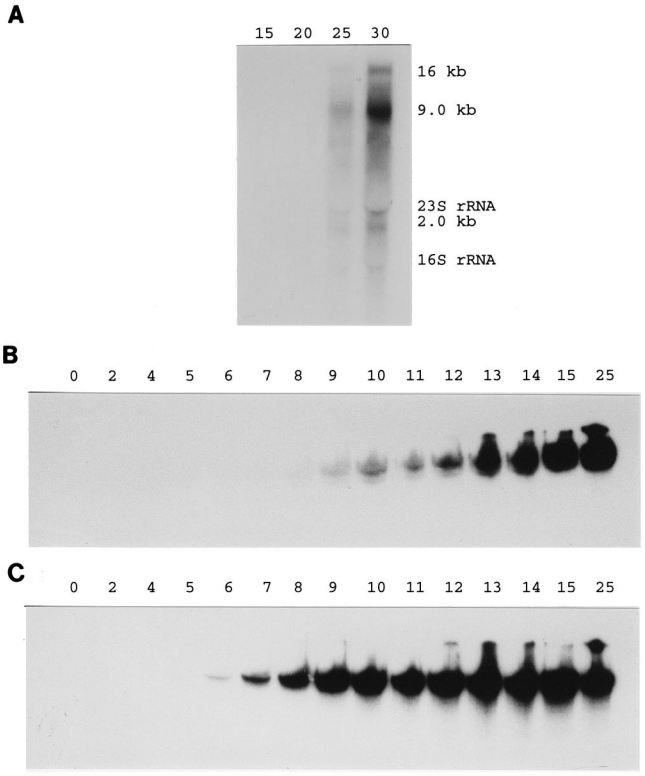

FIG. 4. Northern blot analysis of transcripts specific for phage bIL41. RNA extracted at different times after infection (minutes indicated at tops of lanes) of IL1403 with phage bIL41 was hybridized with oligonucleotide 13 specific for late-expressed transcripts (A and B) and with an oligonucleotide complementary to the sequence of phage bIL66 (AAATTAAAGCTGTATCAAGAT) (9) specific for middle-expressed transcripts (C). The same amount of total cell RNA was loaded in each well of an agarose gel (A) or an acrylamide gel (B and C).

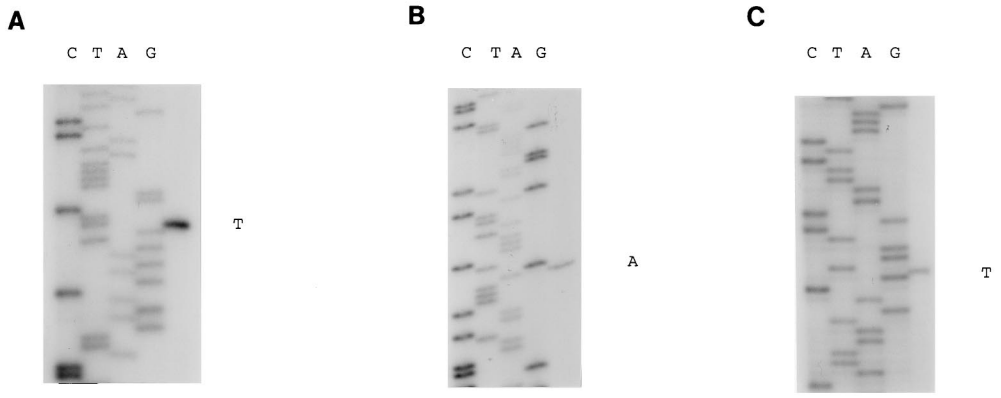

FIG. 5. Primer extension analysis carried out with transcripts from bIL41-infected IL1403 cells with oligonucleotides 3 (A), 20 (B), and 21 (C).

additional oligonucleotides as probes revealed an additional transcript of 7.0 kb (Fig. 1B). The 5′ ends of all transcripts were mapped by primer extension (Fig. 5). The 16- and 9.0-kb transcripts started at the same point (position 1311 on the L35061 sequence) upstream of *orf1* (Fig. 1B and 5A). A −10 consensus sequence (TATTTT) was detected upstream of this start, but there was no −35 consensus sequence. The 2.0-kb transcript started with a sequence presenting features of an RNase E cleavage site (Fig. 1B and 5B) (position 4715 on the L35061 sequence), which is G|ATTT followed 6 bp downstream by a hairpin structure (ACCGCAGACGAGTGGT; complement ary bases are underlined) (28). No particular features were detected in the starting region of the 7.0-kb transcript (Fig. 1B and 5C) (position 6678 on the L35061 sequence).

Taken together, these results indicate that late genes are transcribed as a 16-kb transcript initiated about 9 min after infection. This transcript is then processed to smaller units.

A phage product is needed for the transcription of late-expressed genes. Lack of a −35 consensus sequence upstream of the late transcription start site suggested that synthesis of a phage-encoded product is needed for transcription initiation. To check this hypothesis, protein synthesis was blocked by the addition of chloramphenicol at the time of phage infection. The effect of 5, 20, 50, 80, 120, and 150 μg of chloramphenicol per ml on phage multiplication, synthesis of early-, middle-, and late-expressed transcripts, and phage DNA replication was tested (data not shown). A concentration of 20 μg of chloramphenicol per ml was shown to allow the synthesis of early transcripts but to prevent those of middle and late transcripts (Fig. 6). However, early-expressed transcripts of 4.5 and 1.2 kb reached a peak after 10 min and ceased almost completely after 25 min in the absence of chloramphenicol. Their synthesis was slightly delayed in the presence of chloramphenicol. The concentration of 20 μg of chloramphenicol per ml was also shown to prevent phage DNA replication (Fig. 7). These results indicate that this concentration blocks phage protein synthesis needed for DNA replication and transcription of middle- and late-expressed genes. Therefore, as observed for other phages, transcription of early-expressed genes is carried out by the host RNA polymerase and that of middle- and late-expressed genes depends on phage-encoded proteins.

DISCUSSION

Sequence analysis of the bIL41 segment linked previously obtained data for the related phages bIL66 (9), F4-1 (23, 46), P008 (10), and sk1 (19) and provided insight into the organization of a segment covering the entire middle-expressed region and a large part of the late-expressed genome region. Data on genome organization allowed the study of the transcription of late-expressed genes.

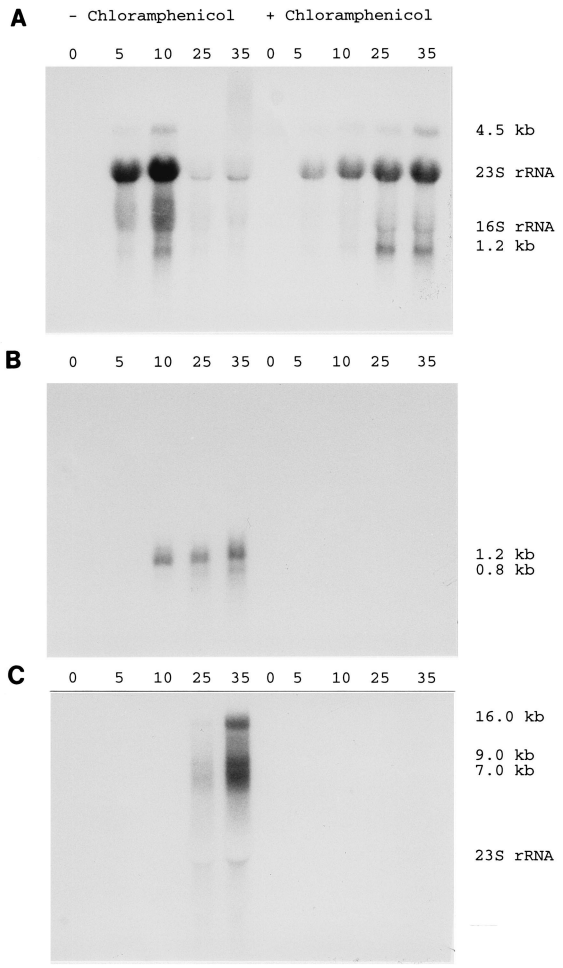

FIG. 6. Transcription of phage bIL41 DNA in the presence and absence of chloramphenicol (20 μg/ml). Early (A), middle (B), and late (C) transcripts are shown. The oligonucleotides used were complementary to the sequence of phage bIL170 (CGGAAGTCACGTCTGTT) (47a) and phage bIL66 (see Fig. 4) (9) for early and middle transcripts, respectively. Late transcripts were hybridized with oligonucleotide 18. The numbers at the tops of the lanes indicate the times (in minutes) after infection.

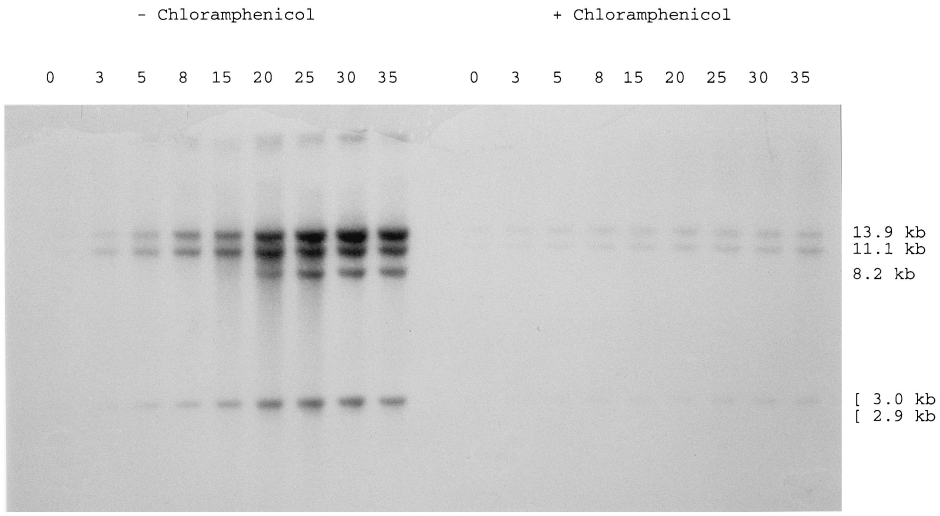

FIG. 7. DNA replication of phage bIL41 in the presence and absence of chloramphenicol (20 μg/ml). Total cell DNA was extracted at different times after infection (minutes indicated at the tops of the lanes), cleaved with *Hin*dIII, and analyzed by Southern hybridization with phage DNA as a probe. There are three *Hin*dIII sites on the phage bIL41 genome (56a) giving rise to two internal segments (13.9 and 3.0 kb) and two end segments (8.2 and 2.9 kb). The 11.1-kb segment represents the sum of the end fragments.

Thirteen ORFs have been identified. Two of them presented features evocative of terminase subunits. The gene organization of the terminase operon, highly conserved in different *E. coli* phages, was also observed on the *Bacillus subtilis* SPP1 phage (17). This operon, close to one end of the genome, codes for two subunits. In bacteriophage lambda, the small subunit (gpNu1; 181 residues) possesses a specific DNA-binding domain (47) followed by an ATP-binding site (5). The large subunit (gpA; 641 residues) carries an ATP-binding site and a basic leucine zipper involved in the endonuclease activity of the terminase (26). In *L. lactis* phages P008 (10) and bIL41, *orf1* and *orf2* are close to one end of the genome. gp1 (174 residues) has two helix-turn-helix motifs which could be involved in DNA binding. gp2 (403 residues) has two ATP binding motifs which could be involved in endonuclease activity. These results support the earlier proposal by P. Billard (10) that gp1 and gp2 could be terminase subunits.

The protein specified by *orf7* is highly homologous to the P35 protein encoded by the related phage F4-1 (46), which had been shown to be a structural protein (45). Presumably, the other proteins encoded by the adjacent ORFs are also structural proteins or play a role in phage morphogenesis or cell lysis. gp12, which has a high identity to gp47 from the unrelated phage r1t (67), might be involved in the last process, since it shares a certain level of homology (19% identity, 41% similarity) with IpaB, a protein expressed by *Shigella flexneri*, that acts as a membrane lysin (37). If they were able to complement, then, in the complex dairy environment, these proteins could contribute to phage species evolution.

Sequence comparison of phages bIL41, bIL66 (9), F4-1 (23, 46), P008 (10), and sk1 (19) revealed similar overall organizations. As already suggested by heteroduplex studies (43, 49), the phages are highly homologous in the major part of the sequenced region, differing only by the insertion or deletion of 1 to 400 bp in noncoding regions. At a specific locus, the sequence of phages bIL41 and F4-1, which differ by the insertion or deletion of an ORF (ca. 2 kb), diverge completely. Therefore, lactococcal phages from the 936 quasi-species seem to generally retain the same overall organization. Essential genes such as those coding for the terminase subunits, the major capsid protein, and several other late-expressed genes are also well conserved. This type of structural variation resembles that already described for the T-even phages (59) and differs from that of lambdoid phages, which retain the same overall organization but have few universally conserved segments (12).

Transcription of the bIL41 segment revealed a 16-kb transcript initiated close to the *cos* site, upstream from *orf1*. This transcript hybridizes with a probe specific for the lysin gene of phage US3 (58), which has been mapped about 15 kb downstream of the transcription start site (data not shown). The largest of the late transcripts detected on the closely related phage sk1 was 14 kb long and covered all of the late-expressed region of the genome (18). Therefore, we can presume that the 16-kb transcript covers all of the late-expressed genes of bIL41. Since different amounts of structural proteins are probably required for phage assembly, their rate of synthesis should be different, and as already hypothesized for other phages (16), regulation of synthesis presumably occurs at the translational level (53). This regulation could involve the unusual UUG start codon (*orf4* and *orf12*) (60), differences in the distance between ribosome binding sequence and the start codon (60), the presence of G residues 5′ of certain start codons (*orf3*, *orf5*, *orf9*, and *orf11*), and overlaps between some genes (*orf4* to *orf7* and *orf8* to *orf11*). This regulation could also result from trapping of the ribosome binding sequence either in RNA secondary structures (27, 28, 35, 36, 70) or by a repressor protein (3, 44, 61, 64, 65). The presence of the 16-kb transcript in variable amounts depending on the assays and of several shorter transcripts suggests that early termination, processing, and degradation of transcripts may also participate in gene regulation. Early termination, or processing at a specific site, could be involved in the formation of the 9.0-kb transcript initiated at the same point as the 16-kb one. Formation of the 2.0-kb transcript, shown to start at an RNase E consensus cleavage site, results most probably from processing. RNase E has not yet been described in *L. lactis*. However, this enzyme is a major determinant of bacteriophage T4 mRNA turnover and participates in the regulation of gene expression (55). The 3′ end of

the 2.0-kb transcript was not precisely mapped, but the hybridization results indicate that it could be adjacent to the 5′ end of the 7.0-kb transcript and would therefore result from endonucleolytic cleavage. A transcript covering the region located upstream of the 5′ end of the 2.0-kb transcript was not detected. It is most probably highly unstable and rapidly degraded. All other transcripts are stable and are present up to the end of the phage life cycle. Segmental differences in the stability of polycistronic mRNAs have been reported for a number of bacterial and phage operons (3, 4, 6, 7, 48, 56). In particular, a monocistronic T4 gene 32 transcript, produced by RNase E processing at the 5′ end, is more stable than the upstream portion of the transcript (7). It was proposed that one role of RNase E processing would be the reduction of upstream gene expression by degradation initiated at the RNase E cleavage site (14). This kind of regulation could play a role in phage bIL41, since the structural proteins which have to be synthesized in greater amounts than the terminase subunits are encoded by the downstream part of the transcript.

ACKNOWLEDGMENTS

We thank C. A. Batt for communication of the F4-1 sequence, W. M. de Vos for providing us with a lysin probe, and C. Anagnostopoulos for critically reading the manuscript.

R. Parreira was supported by Junta Nacional de Investigaçao Cientifica e Tecnologica (Lisbon, Portugal). A. L. S. Lerayer was supported by Conselho Nacional de Desenvolvimento Cientifico e Tecnologico (Brasilia, Brazil). This work was supported by BRIDGE contract Biot-CT91-0263 and BIOTECH contract BIO2-CT94-3055 of the Commission of the European Communities.

REFERENCES

1. **Altschul, S. F., W. Gish, W. Miller, E. W. Myers, and D. J. Lipman.** 1990. Basic local alignment search tool. J. Mol. Biol. **215:**403–410.
2. **Anba, J., E. Bidnenko, A. J. Hillier, D. Ehrlich, and M.-C. Chopin.** 1995. Characterization of the lactococcal *abiD1* gene coding for phage abortive infection. J. Bacteriol. **177:**3818–3823.
3. **Andrake, M. D., and J. D. Karam.** 1991. Mutational analysis of the RNA operator for T4 DNA polymerase. Genetics **128:**203–213.
4. **Baga, M., M. Goransson, S. Normark, and E. Uhlin.** 1988. Processed mRNA with differential stability in the regulation of *E. coli* pilin gene expression. Cell **52:**197–206.
5. **Becker, A., and M. Gold.** 1988. Prediction of an ATP reactive center in the small subunit, gpNu1, of the phage lambda terminase enzyme. J. Mol. Biol. **139:**219–222.
6. **Belasco, J. G., J. T. Beatty, C. W. von Gabain, and S. N. Cohen.** 1985. Differential expression of photosynthetic genes in *R. capsulata* results from segmental differences in stability within the polycistronic *rxcA* transcript. Cell **40:**171–181.
7. **Belin, D., E. A. Mudd, P. Prentki, Y. Yi-Yi, and H. M. Krisch.** 1987. Sense and antisense transcription of bacteriophage T4 gene 32. Processing and stability of the mRNAs. J. Mol. Biol. **194:**231–243.
8. **Beresford, T. P. J., L. J. H. Ward, and A. W. Jarvis.** 1993. Temporally regulated transcriptional expression of the genomes of lactococcal bacteriophages C2 and sk1. Appl. Environ. Microbiol. **59:**3708–3712.
9. **Bidnenko, E., S. D. Ehrlich, and M.-C. Chopin.** 1995. Phage operon involved in sensitivity to *Lactococcus lactis* abortive infection mechanism AbiD1. J. Bacteriol. **177:**3824–3829.
10. **Billard, P.** 1992. Etude moleculaire sur les bactéries lactiques: étude des régions terminales de l'ADN génomique de bactériophages infectant *Lactococcus lactis*. Ph.D. thesis. Université de Nancy I, Nancy, France.
11. **Bodorovsky, M., and J. McLninch.** 1993. Genmerk: parallel gene recognition for both DNA strands. Comput. Chem. **17:**123–133.
12. **Campbell, A.** 1994. Comparative molecular biology of lambdoid phages. Annu. Rev. Microbiol. **48:**193–222.
13. **Campbell, A., and D. Botstein.** 1983. Evolution of the lambdoid phages, p. 365–380. *In* R. W. Hendrix, J. W. Roberts, F. W. Stahl, and R. A. Weisberg (ed.), Lambda II. Cold Spring Harbor Laboratory, Cold Spring Harbor, N.Y.
14. **Carpoussis, A. J., E. A. Mudd, and H. M. Krisch.** 1989. Transcription and messenger RNA processing of upstream bacteriophage T4 gene 32. J. Mol. Biol. **219:**39–48.
15. **Casjens, S., G. Hattfull, and R. Hendrix.** 1992. Evolution of ds DNA tailed-bacteriophages genomes. Semin. Virol. **3:**383–397.
16. **Casjens, S., and R. Hendrix.** 1988. Control mechanisms in dsDNA bacteriophage assembly, p. 15–91. *In* R. Calendar (ed.), The bacteriophages. Plenum Press, New York.
17. **Chai, S., A. Bravo, G. Luder, A. Nedlin, T. A. Trautner, and J. C. Alonso.** 1992. Molecular analysis of the *Bacillus subtilis* bacteriophage SPP1 region encompassing genes 1 to 6. J. Mol. Biol. **224:**87–102.
18. **Chandry, P. S., B. E. Davidson, and A. J. Hillier.** 1994. Temporal transcription map of the *Lactococcus lactis* bacteriophage sk1. Microbiology **140:** 2251–2261.
19. **Chandry, P. S., S. C. Moore, B. E. Davidson, and A. J. Hillier.** 1994. Analysis of the cos region of the lactococcus bacteriophage sk1. Gene **138:**123–126.
20. **Chopin, A.** 1993. Organization and regulation of genes for amino acid biosynthesis in lactic acid bacteria. FEMS Microbiol. Rev. **12:**21–38.
21. **Chopin, A., M.-C. Chopin, A. Moillo-Batt, and P. Langella.** 1984. Two plasmid-determined restriction and modification systems in *Streptococcus lactis*. Plasmid **11:**260–263.
22. **Chou, P. Y., and G. D. Fasman.** 1978. Prediction of the secondary structure of proteins from their amino acid sequence. Adv. Enzymol. Relat. Areas Mol. Biol. **47:**45–148.
23. **Chung, D. K., J. H. Kim, and C. A. Batt.** 1991. Cloning and nucleotide sequence of the major capsid protein from *Lactococcus lactis* spp. *cremoris* bacteriophage F4-1. Gene **101:**121–125.
24. **Cluzel, P.-J., A. Chopin, S. D. Ehrlich, and M.-C. Chopin.** 1991. Phage abortive infection mechanism from *Lactococcus lactis* subsp. *lactis*, expression of which is mediated by an Iso-IS*S1* element. Appl. Environ. Microbiol. **57:**3547–3551.
25. **Coffey, A. G., G. F. Fitzgerald, and C. Daly.** 1991. Cloning and characterization of the determinant for abortive infection of bacteriophage from lactococcal plasmid pCI829. J. Gen. Microbiol. **137:**1355–1362.
26. **Davidson, A. R., and M. Gold.** 1992. Mutations abolishing the endonuclease activity of bacteriophage lambda terminase lie in two distinct regions of the A gene, one of which may encode a "Leucine Zipper" DNA-binding domain. Virology **189:**21–30.
27. **de Smit, M. H., and J. van Duin.** 1990. Control of prokaryotic translational initiation by mRNA secondary structure. Prog. Nucleic Acids Res. Mol. Biol. **38:**1–35.
28. **de Smit, M. H., and J. van Duin.** 1993. Translational initiation at the coat-protein gene of phage MS2: native upstream RNA relieves inhibition by local secondary structure. Mol. Microbiol. **9:**1079–1088.
29. **Devereux, J., P. Haeberli, and O. Smithies.** 1984. A comprehensive set of sequence analysis programmes for the vax. Nucleic Acids Res. **12:**387–395.
30. **Durmaz, E., D. L. Higgins, and T. R. Klaenhammer.** 1992. Molecular characterization of a second abortive phage resistance gene present in *Lactococcus lactis* subsp. *lactis* ME2. J. Bacteriol. **174:**7463–7469.
31. **Ehretsmann, C. P., A. J. Carpousis, and H. M. Krisch.** 1992. mRNA degradation in procaryotes. FASEB J. **6:**3186–3192.
32. **Garvey, P., G. F. Fitzgerald, and C. Hill.** 1995. Cloning and DNA sequence analysis of two abortive infection phage resistance determinants from lactococcal plasmid pNP40. Appl. Environ. Microbiol. **61:**4321–4328.
33. **Gibson, T. J.** 1984. Studies on the Eppstein-Barr virus genome. Ph. D. thesis. Cambridge University, Cambridge.
34. **Glatron, M. F., and G. Rapoport.** 1972. Biosynthesis of the parasporal reclusion of *Bacillus thuringiensis*: half-life of its corresponding messenger RNA. Biochimie **54:**1291–1301.
35. **Groeneveld, H., K. Thimon, and J. van Duin.** 1995. Translational control of maturation-protein synthesis in phage MS2: a role for the kinetics of RNA folding? RNA **1:**79–88.
36. **Hall, M. N., J. Gabay, M. Debarbouille, and M. Schwartz.** 1982. A role for mRNA secondary structure in the control of translation initiation. Nature (London) **295:**616–618.
37. **High, N., J. Mounier, M. C. Prevost, and P. J. S. Sansonetti.** 1992. IpaB of *Shigella flexneri* causes entry into epithelial cells and escape from the phagocytic vacuole. EMBO J. **11:**1991–1999.
38. **Hill, C.** 1993. Bacteriophage and bacteriophage resistance in lactic acid bacteria. FEMS Microbiol. Lett. **12:**87–108.
39. **Hill, C., I. J. Massey, and T. R. Klaenhammer.** 1991. Rapid method to characterize lactococcal bacteriophages genomes. Appl. Environ. Microbiol. **57:**283–288.
40. **Hill, C., L. A. Miller, and T. R. Klaenhammer.** 1990. Nucleotide sequence and distribution of the pTR2030 resistance determinant (*hsp*) which aborts bacteriophage infection in lactococci. Appl. Environ. Microbiol. **56:**2255–2258.
41. **Holo, H., and I. F. Nes.** 1989. High-frequency transformation, by electroporation, of *Lactococcus lactis* subsp. *cremoris* grown with glycine in osmotically stabilized media. Appl. Environ. Microbiol. **55:**3119–3123.
42. **Jarvis, A. W., G. F. Fitzgerald, M. Mata, A. Mercenier, H. Neve, I. B. Powell, C. Ronda, M. Saxelin, and M. Teuber.** 1991. Species and type phages of lactococcal bacteriophages. Intervirology **32:**2–9.
43. **Jarvis, A. W., and J. Meyer.** 1986. Electron microscopic heteroduplex study and restriction endonuclease cleavage analysis of the DNA genome of three lactic streptococcal bacteriophages. Appl. Environ. Microbiol. **51:**566–571.
44. **Karam, J., L. Gold, and B. S. Singer.** 1981. Translational regulation: identification of the site on bacteriophage T4rIIB mRNA recognized by the *regA* gene function. Proc. Natl. Acad. Sci. USA **78:**4669–4673.
45. **Kim, J. H., and C. A. Batt.** 1991. Molecular characterization of a *Lactococcus*

lactis bacteriophage F4-1. Food Microbiol. **8:**15–26.
46. **Kim, J. H., and C. A. Batt.** 1991. Nucleotide sequence and deletion analysis of a gene coding for a structural protein of *Lactococcus lactis* bacteriophage F4-1. Food Microbiol. **8:**27–36.
47. **Kypr, J., and J. Mrazek.** 1986. Lambda phage protein Nu1 contains the conserved DNA binding fold of repressors. J. Mol. Biol. **91:**139–140.
47a. **Kyriakidis, S.** Personal communication.
48. **Liu, J., W. O. Barnell, and T. Conway.** 1992. The polycistronic mRNA of the *Zymomonas mobilis glf-zwf-edd-glk* operon is subject to complex transcript processing. J. Bacteriol. **174:**2824–2833.
49. **Loof, M., and M. Teuber.** 1986. Heteroduplex analysis of the genomes of *Streptococcus lactis* "subsp." *diacetylactis* bacteriophages of the P008-type isolated from German cheese factories. Syst. Appl. Microbiol. **8:**226–229.
50. **Ludwig, W., E. Seewaldt, R. Klipper-Bälz, K. H. Schleifer, L. Magrum, C. R. Woese, G. E. Fox, and E. Stackebrandt.** 1985. The phylogenetic position of *Streptococcus* and *Enterococcus*. J. Gen. Microbiol. **131:**543–551.
51. **Maniatis, T. E., F. Fritsch, and J. Sambrook.** 1982. Molecular cloning: a laboratory manual. Cold Spring Harbor Laboratory, Cold Spring Harbor, N.Y.
52. **Mathews, F. S.** 1985. The structure, function and evolution of cytochromes. Prog. Biophys. Mol. Biol. **45:**1–56.
53. **McCarthy, J. E. G., and C. Gualerzi.** 1990. Translational control of prokaryotic gene expression. Trends Genet. **6:**78–85.
54. **McLansborough, L. A., K. M. Kolaetis, T. Requena, and L. L. McKay.** 1995. Cloning and characterization of the abortive infection genetic determinant *abiD* isolated from pBF61 of *Lactococcus lactis* subsp. *lactis* KR5. Appl. Environ. Microbiol. **61:**2023–2026.
55. **Mudd, E. A., A. J. Carpousis, and H. M. Krisch.** 1990. *Escherichia coli* RNase E has a role in the decay of bacteriophage T4 mRNA. Genes Dev. **4:**873–881.
56. **Newbury, F., N. H. Smith, and C. F. Higgins.** 1987. Differential mRNA stability controls relative gene expression within a polycistronic operon. Cell **51:**1131–1143.
56a. **Parreira, R., et al.** Unpublished data.
57. **Parreira, R., S. D. Ehrlich, and M. C. Chopin.** 1996. Dramatic decay of phage transcripts is observed in lactococcal cells carrying the abortive infection determinant AbiB. Mol. Microbiol. **19:**221–230.
58. **Platteew, C., and W. M. de Vos.** 1992. Location characterization and expression of lytic enzyme-coding gene, *lytA*, of *Lactococcus lactis* bacteriophage phi US3. Gene **118:**115–120.
59. **Repoila, F., F. Tetart, J.-Y. Bouet, and H. M. Krisch.** 1994. Genomic polymorphism in the T-even bacteriophages. EMBO J. **13:**4181–4192.
60. **Ringquist, S., S. Shinedling, D. Barrick, L. Green, J. Binkley, G. D. Stormo, and L. Gold.** 1992. Translation initiation in *Escherichia coli*: sequences within the ribosome-binding site. Mol. Microbiol. **6:**1219–1229.
61. **Romaniuk, P., P. Lowary, H. N. Wu, G. Stormo, and O. Uhlenbeck.** 1987. RNA binding site of R17 coat protein. Biochemistry **26:**1563–1568.
62. **Schouler, C., S. D. Ehrlich, and M.-C. Chopin.** 1994. Sequence and organization of the lactococcal prolate-headed bIL67 phage genome. Microbiology **140:**3061–3069.
63. **Terzaghi, B. E., and W. E. Sandine.** 1975. Improved medium for lactic streptococci and their bacteriophages. Appl. Environ. Microbiol. **29:**807–813.
64. **Tuerk, C. S., D. Parma, and L. Gold.** 1990. Autogenous translational operator recognized by bacteriophage T4 DNA polymerase. J. Mol. Biol. **213:** 749–761.
65. **Unnitham, S., L. Green, L. Morissey, J. Binkley, B. Singer, J. Karam, and L. Gold.** 1990. Binding of the bacteriophage T4 RegA protein to mRNA targets: an initiator AUG is required. Nucleic Acids Res. **18:**7083–7092.
66. **van de Guchte, M., J. Kok, and G. Venema.** 1992. Gene expression in *Lactococcus lactis*. FEMS Microbiol. Rev. **88:**73–92.
67. **van Sinderen, D., H. Karsens, J. Kok, P. Terpstra, M. H. J. Ruiters, G. Venema, and A. Nauta.** 1996. Sequence analysis and molecular characterization of the temperate lactococcal bacteriophage r1t. Mol. Microbiol. **19:** 1343–1355.
68. **Walker, J. E., M. Saraste, M. J. Runswick, and N. J. Gay.** 1982. Distantly related sequences in the alpha- and beta-subunits of ATP synthase, myosin, kinases and other ATP-requiring enzymes and a common nucleotide binding fold. EMBO J. **1:**945–951.
69. **Wilkinson, M., J. Doskow, and S. Lindsey.** 1990. RNA blots: staining procedures and optimization of conditions. Nucleic Acids Res. **19:**679.
70. **Wilkstrom, P. M., L. K. Lind, D. E. Berg, and G. R. Bjork.** 1992. Importance of mRNA folding and start codon accessibility in the expression of genes in a ribosomal protein operon of *Escherichia coli*. J. Mol. Biol. **224:**949–966.
71. **Williams, J. G., and P. J. Mason.** 1985. Hybridization in the analysis of RNA, p. 139–160. *In* B. D. Hames and S. J. Higgins (ed.), Nucleic acid hybridization—a practical approach. IRL Press, Oxford.