Virology **283,** 93–109 (2001)
doi:10.1006/viro.2001.0871, available online at http://www.idealibrary.com on IDEAL®

Analysis of the Complete DNA Sequence of the Temperate Bacteriophage TP901-1: Evolution, Structure, and Genome Organization of Lactococcal Bacteriophages

Lone Brøndsted,*[,1,2] Solvej Østergaard,† Margit Pedersen,†[,3] Karin Hammer,* and Finn K. Vogensen†

*Department of Microbiology, Technical University of Denmark, DK-2800 Lyngby; †Department of Dairy and Food Science, The Royal Veterinary and Agricultural University, Rolighedsvej 30, DK-1958 Frederiksberg C, Denmark

Received October 5, 2000; returned to author for revision November 14, 2000; accepted February 14, 2001

A complete analysis of the entire genome of the temperate lactococcal bacteriophage TP901-1 has been performed and the function of 21 of 56 TP901-1-encoded ORFs has been assigned. This knowledge has been used to propose 10 functional modules each responsible for specific functions during bacteriophage TP901-1 proliferation. Short regions of microhomology in intergenic regions present in several lactococcal bacteriophages and chromosomal fragments of *Lactococcus lactis* are suggested to be points of exchange of genetic material through homologous recombination. Our results indicate that TP901-1 may have evolved by homologous recombination between the host chromosome and a mother phage and support the observation that phage remnants as well as prophages located in the *Lactococcus* chromosome contribute significantly to bacteriophage evolution. Some proteins encoded in the early transcribed region of the TP901-1 genome were more homologous to proteins encoded by phages infecting gram-positive hosts other than *L. lactis*. This protein homology argues for the occurrence of horizontal genetic exchange among these bacteriophages and indicates that they have access to a common gene pool.

INTRODUCTION

Lactic acid bacteria are extensively used as starter cultures for a range of fermented dairy products. The major reason for fermentation failures in the dairy industry is bacteriophage attack on the fermenting bacteria (Klaenhammer, 1991; Forde and Fitzgerald, 1999). Furthermore, lysogeny is known to be widespread in many species of lactic acid bacteria, and it has been speculated that lytic phages may evolve from temperate phages released from lysogenic strains (Jarvis, 1989; Davidson *et al.,* 1990). During fermentation, a limited number of lactic acid bacteria species, which are potential hosts for closely related bacteriophages, are propagated in large numbers. In this situation exchange of genetic material in the production environment may occur and could be the reason for the rapid evolution among bacteriophages infecting lactic acid bacteria.

By comparison of complete genome sequences of related bacteriophages, knowledge of their evolution is obtained. The first class of bacteriophages subjected to this type of analysis was the group of temperate lambdoid phages (Botstein, 1980). It was concluded that these phages are related in ways not easily accounted for by standard ideas of evolution along branching trees of linear descent. Instead Botstein suggested a theory of modular evolution, defined as "the joint evolution of sets of functionally and genetically interchangeable elements, each of which carries out a particular biological function" (Botstein, 1980). Later work divided the genomes of lambdoid phages into 11 major segments of functionally clustered genes. Each of the regions contains more than one gene and could be considered as a functional module (Casjens *et al.,* 1992). During the last five years DNA sequences of bacteriophages infecting the lactic acid bacterium *Streptococcus thermophilus*, which is used for production of yogurt and hard cheese varieties, have been accumulated and subjected to extensive analyses with an emphasis on evolution among streptococcal bacteriophages (Lucchini *et al.,* 1998, 1999; Desiere *et al.,* 1998). Essentially the results support the modular theory of Botstein, with the comment that a module might only be a single gene (Neve *et al.,* 1998). Recently, comparison of DNA sequences of bacteriophages and prophages spanning a broad phylogenetic range of host bacteria has lead to the proposal of a model for the genetic structure and dynamics of the global phage population. This model suggests that the phage genomes have access, by horizontal exchange, to a large common genetic pool; however, access to the gene pool is not uniform for all phages (Hendrix *et al.,* 1999).

[1] To whom correspondence and reprint requests should be addressed at Department of Dairy and Food Science, Rolighedsvej 30, The Royal Veterinary and Agricultural University, DK-1958 Frederiksberg C, Denmark. Fax: +45 35 28 32 31. E-mail: lobr@kvl.dk.

[2] Present address: Department of Dairy and Food Science, The Royal Veterinary and Agricultural University, Rolighedsvej 30, DK-1958 Frederiksberg C, Denmark.

[3] Present address: Department of Microbiology, Technical University of Denmark, DK-2800 Lyngby, Denmark.

0042-6822/01 $35.00

The temperate lactococcal bacteriophage TP901-1 belongs to the *Siphoviridae* family containing an isometric head, a noncontractile tail, and a genome with terminally redundant ends (Christiansen *et al.,* 1994). The latent period is 65 min and the burst size is 40 ± 10 phage particles (Madsen and Hammer, 1998). TP901-1 belongs to the P335 phage species, which includes both virulent and temperate bacteriophages. DNA sequence data from other members of the P335 phage species are available, including the complete genome sequence of the temperate phage r1t, and a partial genome sequence of the virulent phage Φ31, as well as the temperate bacteriophages Tuc2009 and ΦLC3 (Lillehaug and Birkeland, 1993; Arendt *et al.,* 1994; Birkeland, 1994; van de Guchte *et al.,* 1994a,b; van Sinderen *et al.,* 1996; Dinsmore and Klaenhammer, 1997; Lillehaug *et al.,* 1997; Walker *et al.,* 1998; Walker and Klaenhammer, 1998; McGrath *et al.,* 1999).

In this paper the entire DNA sequence of the TP901-1 genome is presented and extensively analyzed. Based on the comparison of the TP901-1 genome with other bacteriophage genome sequences, aspects of bacteriophage evolution are discussed.

RESULTS

Identification of ORFs encoded by phage TP901-1 and genome organization

The TP901-1 genome consists of 37,667 bp with a G + C content of 35.4%, which is similar to the G + C content of the host *Lactococcus lactis* (34.8 to 35.6%) (Schleifer *et al.,* 1985). An analysis of the genome revealed 56 open reading frames (ORFs) based on several criteria. First, an ORF consists of at least 40 codons preceded by a potential Shine–Dalgarno sequence at an appropriate distance (2–9 bp) from the initiation codon (AUG, GUG, or UUG) (Table 1). Alternatively, in the absence of a potential Shine–Dalgarno sequence the initiation codon should be located close to the stop codon of the upstream gene, thereby achieving translation initiation by translational coupling. This is the case for ORF31, ORF37, and ORF41. Second, a codon usage table was created on the basis of genes with proven biological function and used to predict *orf*s in the entire TP901-1 sequence (data not shown). All initially predicted *orf*s except *orf24* were confirmed by this analysis, suggesting that *orf24* may not be expressed. Most ORFs initiate translation at an AUG codon (49 ORFs), while GUG (2 ORFs) and UUG (5 ORFs) are used less frequently. Comparison of the putative Shine–Dalgarno sequences revealed a TP901-1 consensus Shine–Dalgarno sequence ($A^A/_G$AAAGGAGG$^A/_T$), essentially identical to that of the host *L. lactis* (AGAAAGGAGGT) (Ludwig *et al.,* 1985).

The 56 open reading frames of TP901-1 are organized in two divergently oriented clusters consisting of 4 and 52 genes, respectively (Fig. 1). The four genes in the smaller cluster are involved in establishment and maintenance of lysogeny, whereas the larger cluster contains genes involved in lytic growth. During the lytic cycle, sequential clusters of the TP901-1 genome are temporally transcribed and the genome can be divided into early (*orf1* to *orf29*), middle (*orf24* to *orf29*), and late (*orf30* to *orf56*) regions (Madsen and Hammer, 1998).

Twelve extended noncoding areas ranging in size from 55 to 580 bp are distributed over the entire genome. Downstream of *orf2* lies a stem-loop structure that could function as a potential factor-independent terminator (Christiansen *et al.,* 1996). The divergently oriented promoters (P_R and P_L), responsible for early transcription of the TP901-1 genome, are located between *orf4* and *orf5*. In the 580-bp large intergenic region between *orf29* and *orf30*, a promoter, active in the late phase of the lytic cycle, has been identified (L. Brøndsted, unpublished data). The *pac* site of TP901-1 is located upstream of *orf30* (Johnsen, 1995).

Upstream of both *orf27* and *orf28* lies a 44-bp identical region that covers the SD sequence of these genes. Between *orf55* and *orf56* 4.5 copies of a 40-bp direct repeat are present; in the first part of each direct repeat a 9-bp inverted repeat is found. If this intergenic region is transcribed, large stable mRNA structures, as calculated by the M-fold program, may potentially form. However, no data concerning the functions of these sequences have been reported.

Homologies of ORFs encoded by the TP901-1 genome

The deduced proteins of the entire TP901-1 genome were compared with databases using Blast version 2.0.4 (Altschul *et al.,* 1997). In Table 1 homologies selected on the basis of three criteria are shown: either high significance of homology (low *e* value), homology to phage-encoded proteins, or homology to proteins with identified function. Of the 56 deduced proteins, 45 showed significant similarity to proteins in the databases. In most cases homologies were found to phages infecting gram-positive bacteria, primarily *L. lactis* and *S. thermophilus* (Table 1). Homologies to *Lactobacillus gasseri*, *Lactobacillus casei, Enterococcus faecalis, Staphylococcus aureus*, *Bacillus subtilis*, *Listeria monocytogenes*, and *Streptococcus pneumoniae* phages were also noted (Table 1). Unfortunately, the biological functions of many of these proteins have not been assigned. Only in seven cases could the function of TP901-1-encoded proteins be predicted with high probability, using significant homology to proteins with known functions as the basis (Table 1). These are the single-stranded DNA binding protein (ORF12), Holliday junction resolvase (ORF15), dUTPase (ORF21), terminase small subunit (ORF30), portal protein (ORF32), holin (ORF52), and lysin (ORF53).

TABLE 1

General Features of the Putative ORFs of Bacteriophage TP901-1 Genome and Homology to Proteins in Databases

ORF	Gene name	RBS AGAAAGGAGGT	Start	Stop	Size of protein: aa	Size of protein: kDa	pI	Function and identity	*e*-Value	Identity[a] aa (%)	Reference or accession number
1	*int*	AGAAATGAGGT	1487	30	485	55,65	9,1	Integrase			Christiansen *et al.* (1996)
								Integrase in *E. faecalis* phage ΦFC	1e-58	494 (33)	AF124258
								Putative integrase in *L. monocytogenes* phage A118	9e-52	479 (30)	AJ242593
								Recombinase B of *S. aureus*	1e-47	479 (29)	AB014438
								Recombinase in *B. subtilis* phage Φ105	2e-39	295 (35)	AB016282
								Recombinase CisA of *B. subtilis*	5e-31	525 (28)	M29040
								Resolvase TnpX of *C. perfringens*	3e-20	393 (23)	U15027
								DNA invertase Pin of *E. coli*	2e-12	94 (44)	K03521
2		GATTAGGAGGA	2153	1608	181	20,53	8,9	ORF15 of *B. halodurans*	3e-03	34 (58)	AB013492
								ORF2 of *L. lactis* phage TPW22	1e-02	29 (65)	AF066865
3		ATAAAGGCGAT	2649	2212	145	17,24	4,7	ORF3 in *S. thermophilus* phage ΦO1205	4e-05	101 (33)	U88974
								Hypothetical protein in *L. gasseri* phage Φadh	4e-02	108 (25)	AJ131519
								ORF122 in *S. thermophilus* phage Sfi21	2e-01	95 (29)	AF115103
4	*cI*	AAAAAAGAGGT	3188	2646	180	20,77	5,8	Repressor			Madsen *et al.* (1999)
								Repressor in *L. lactis* phage Φ31	4e-41	167 (53)	AJ292531
								Repressor in *S. aureus* phage ΦPVL	1e-32	168 (50)	AB009866
								ORF4 in *S. thermophilus* phage ΦO1205	2e-22	138 (48)	U88974
								Repressor in *L. lactis* phage BK5-T	6e-09	124 (33)	L44593
								Repressor in *L. lactis* phage Tuc2009	5e-06	95 (33)	L26219
								ORF20 in *S. thermophilus* phage ΦO1205	2e-04	77 (36)	U88974
								Repressor in *L. lactis* phage r1-t	4e-04	97 (32)	U38906
5	*mor*	AGAAAGGAGAA	3357	3575	72	8,27	6,7	Modulator of repression			Madsen *et al.* (1999)
								ORF69 in *S. thermophilus* of phage Sfi19	4e-14	68 (51)	AF115102
								CRO in *L. lactis* phage Φ31	1e-12	71 (50)	AJ292531
								ORF31-a in *S. aureus* phage ΦPVL	1e-11	44 (75)	AB009866
								ORF5 in *S. thermophilus* phage ΦO1205	2e-11	69 (47)	U88974
6		AGAAAGGATTC	3612	4355	247	28,35	6,9	ORF238 in *L. lactis* phage Φ31.1	1e-66	134 (94)	AF208055
								Antirepressor of phage VT-Sa	1e-17	119 (42)	AP000363
								Mobilization protein of *B. thuringensis*	9e-03	101 (27)	X56204
7	*xis*	AGAAAGGAAAG	4368	4562	64	7,49	9,2	Excisionase			Breüner *et al.* (1999)
								No similarity found			
8		AGAAAGATTAA	4627	4875	82	9,45	5,3	ORF7 in *L. lactis* phage r1-t	2e-39	82 (100)	U38906
9		AAGTTGGCGGT	4872	5048	58	6,9	6,6	ORF8 in *L. lactis* phage r1-t	2e-26	58 (100)	U38906
								ORF61b in *L. lactis* and *L. lactis* phage ul36.1	1e-25	58 (94)	AF212844/AF212846
								ORF57 in *L. lactis* phage Φ31.1	1e-25	55 (100)	AF208055
								ORF61a in *L. lactis* phage ul36	3e-23	58 (86)	AF212845
								ORF9 in *L. lactis* phage TPW22	2e-12	55 (58)	AF066865
								CRO in *L. lactis* phage BK5-T	1e-10	54 (51)	L44593
10		GAAATGGAGAA	5147	5536	129	14,74	8,6	ORF131 in *L. lactis* phage ul36	2e-53	131 (80)	AF212845
11		ATGCGTGAGGT	5545	6168	207	24,08	5,4	ORF14 in *L. lactis* phage Tuc2009	1e-117	207 (97)	AF109874
								ORF35 in *L. lactis* phage sk1	1e-63	203 (61)	AF011378
								e12 in *L. lactis* phage bIL170	1e-61	203 (59)	AF009630

TABLE 1—*Continued*

ORF	Gene name	RBS AGAAAGGAGGT	Start	Stop	Size of protein aa	Size of protein kDa	pI	Function and identity	e-value	Identity[a] aa (%)	Reference or accession number
12	*ssb*	AGATTGGAGTA	6168	6620	150	16,69	9,1	Single stranded binding protein			Madsen and Hammer (1998)
								ORF15 in *L. lactis* phage Tuc2009	5e-81	150 (97)	AF109874
								ORF141 in *L. lactis* phage ul36	3e-47	151 (64)	AF212845
								ORF9 in *S. thermophilus* phage 7201	2e-45	151 (64)	AF145054
								SSB protein of *B. subtilis*	2e-40	172 (53)	D26185
								ORF45 in *S. aureus* phage ΦPVL	2e-38	157 (55)	AB009866
								Putative SSB in *L. monocytogenes* phage A118	2e-38	160 (51)	AJ242593
								GP36 in *B. subtilis* phage SPP1	2e-17	159 (34)	X97918
								SSB protein of *E. coli*	2e-13	158 (31)	J01704
13	*rep*	GGAAAGGAGAA	6747	7565	272	31,42	6,3	Replication protein			Østergaard *et al.* (2001)
								ORF235 in *L. lactis* and in *L. lactis* phage ul36.1	2e-13	168 (33)	AF212844/AF212846
								ORF269 in *L. lactis* phage Φ31.1	2e-13	168 (33)	AF208055
								ORF46 in *S. aureus* phage ΦPVL	5e-06	154 (25)	AB009866
								GP49 in *L. monocytogenes* phage A118	1e-05	152 (26)	AJ242593
								Replication protein ori60 *B. thuringensis*	2e-02	200 (24)	M60475
								DnaD of *B. subtilis*	9e-02	111 (29)	Z99115
14		GAAGTGGAGGC	7546	7767	73	8,5	5,3	ORF73 in *L. lactis* and in *L. lactis* phage ul36.1	2e-36	73 (98)	AF212844/AF212846
								ORF73 in *L. lactis* phage Φ31.1	2e-36	73 (98)	AF208055
15	*rus*	TATACGGAGAA	7757	8176	139	16,18	9,6	Holliday junction resolvase			This work
								ORF139a in *L. lactis* phage Φ31.1	1e-73	139 (97)	AF208055
								ORF109 in *L. lactis* and in *L. lactis* phage ul36.1	3e-57	109 (97)	AF212844/AF212846
								Holliday junction resolvase RusA *E. coli*	2e-02	115 (28)	X92587
16		ATCGAGGAGGT	8177	8416	79	9,22	9,3	ORF79 in *L. lactis* phage Φ31.1	9e-39	79 (96)	AF208055
								ORF79b in *L. lactis* and in *L. lactis* phage ul36.1	9e-39	79 (96)	AF212844/AF212846
								ORF16 in *L. lactis* phage r1-t	3e-04	60 (33)	U38906
								ORF129 in *L. lactis* phage ul36	4e-04	60 (33)	AF212845
								ORF20 in *L. lactis* phage Tuc2009	1e-03	46 (41)	AF109874
17		AGAATAAAGGC	8522	9046	174	20,15	9,2	ORF174 in *L. lactis* and in *L. lactis* phage ul36.1	2e-95	174 (100)	AF212844/AF212846
								ORF85 in *L. lactis* phage Φ31.1	2e-42	85 (100)	AF208055
18		TATAAGAAGGA	9060	9266	68	7,93	5,3	ORF68 in *L. lactis* phage Φ31.1	1e-33	68 (100)	AF208055
								ORF68c in *L. lactis* and in *L. lactis* phage ul36.1	1e-33	68 (100)	AF212844/AF212846
								ORF68b in *L. lactis* phage ul36	1e-13	68 (57)	AF212845
19		TATTTGGAGGA	9259	9891	210	24,26	4,7	ORF1 in *abiN* operon *L. lactis*	6e-71	214 (69)	Y11901
								ORF19 in *L. lactis* phage r1-t	8e-57	216 (58)	U38906
								ORF237 in *L. lactis* phage ul36	5e-49	221 (52)	AF212845
								ORF184 in *L. lactis* and in *L. lactis* phage ul36.1	1e-31	142 (57)	AF212844/AF212846
								ORF75 in *L. lactis* phage Φ31.1	1e-01	62 (37)	AF208055
								ORF24 in *L. lactis* phage Tuc2009	2e-01	69 (33)	AF109874
								GP51 in *L. monocytogenes* phage A118	3e-01	134 (27)	AJ242593
20		CCTGTGGAGGT	9888	10088	66	7,63	5,7	No similarity found			
21	*dut*	CCTGTGGAGGA	10085	10504	139	15,2	5,4	dUTPase			This work
								ORF20 in *L. lactis* phage r1-t (putative dUTPase)	7e-75	139 (100)	U38906
								ORF139b in *L. lactis* phage Φ31.1	2e-72	139 (97)	AF208055

TABLE 1—*Continued*

ORF	Gene name	RBS AGAAAGGAGGT	Start	Stop	Size of protein aa	Size of protein kDa	pI	Function and identity	e-value	Identity[a] aa (%)	Reference or accession number
								ORF139b in *L. lactis* and in *L. lactis* phage ul36.1	2e-72	139 (97)	AF212844/AF212846
								ORF3 of *abiN* operon of *L. lactis* (putative dUTPase)	1e-71	139 (96)	Y11901
								ORF139a in *L. lactis* phage ul36	4e-68	139 (92)	AF212845
								dUTPase of *E. coli*	1e-11	139 (33)	X01714
22		ACTGGGGAGGT	10507	10827	108	12,22	4,7	ORF118a in *L. lactis* phages ul36 and ul36.1	8e-26	84 (71)	AF212845/AF212846
								ORF118 in *L. lactis* phage Φ31.1	9e-04	81 (40)	AF208055
								ORF118b in *L. lactis*	9e-04	81 (40)	AF212844
								ORF21 in *L. lactis* phage r1-t	7e-02	60 (43)	U38906
23		CGATTGGAGGG	10820	11170	116	13,59	4,9	ORF120 in *L. lactis* phage Φ31.1	3e-36	116 (64)	AF208055
								ORF120a in *L. lactis* phage ul36	3e-36	116 (64)	AF212845
								ORF120b in *L. lactis*	3e-36	116 (64)	AF212844
24		GTGGGGGAGGG	11174	11350	58	6,59	6,3	No similarity found			
25		GGTGGGGAGGA	11403	11642	79	8,75	5,2	No similarity found			
26		GATTAGGAGTA	11639	11797	52	6,58	9,7	ORFX in *L. lactis* phage 712	6e-05	53 (42)	AF087814
								ORF66 in *L. lactis* phage Φ31.1	6e-03	74 (36)	AF208055
27		CAACTGGAGGA	11928	12221	97	11,04	5,9	ORF4 of *abiN* operon of *L. lactis*	2e-49	97 (100)	Y11901
28		CAACTGGAGGG	12344	12487	47	5,55	5,4	No similarity found			
29	*rlt*	CAGACGGAGAA	12563	12985	140	16,74	6,3	Regulator of late transcription			L. Brøndsted, unpublished
								ORF6 of *abiN* operon of *L. lactis*	8e-49	133 (71)	Y11901
30	*terS*	AGAAAGGAGAT	13562	13972	136	15,4	7,8	Terminase small subunit			This work
								ORF8 of *abiN* operon of *L. lactis*	3e-71	136 (100)	Y11901
								HI1411 terminase small subunit of *H. influenza*	1e-03	146 (26)	U32821
31	*terL*	nd[b]	13965	15353	462	53,1	6,3	Terminase large subunit			This work
								ORF9 in *abiN* operon of *L. lactis*	0	462 (100)	Y11901
32	*por*	TCAAAGGAGGA	15354	16712	452	51,82	4,6	Portal protein			This work
								ORF9 in *abiN* operon of *L. lactis*	1e-137	238 (100)	Y11901
								GP502 in *S. thermophilus* phage Sfi11	5e-35	428 (28)	AF057033
								ORF27 in *S. thermophilus* phage ΦO1205	3e-34	426 (27)	U88974
								GP6 in *B. subtilis* phage SPP1	3e-20	466 (24)	X56064
33		ACAAATGAGGT	16716	18410	564	64,28	9,1	ORF28 in *S. thermophilus* phage ΦO1205	2e-39	299 (33)	U88974
								GP284 in *S. thermophilus* phage Sfi11	3e-33	309 (31)	AF057033
34		AATATGGAGGT	18425	18652	75	9,27	6,1	No similarity found			
35	*sfp*	GAGTAGGAGGA	18767	19429	220	24,54	4,7	Scaffolding protein			This work
								Previously named ORFB1			Johnsen *et al.* (1996)
								No similarity found			
36	*mhp*	TAACAGGAGGC	19431	20249	272	28,73	5,2	Major head protein			Johnsen *et al.* (1996)
								No similarity found			
37		nd	20249	20446	65	7,5	9,6	Previously named ORFB2			Johnsen *et al.* (1996)
								No similarity found			
38		AATTAGGAGTT	20433	20765	110	12,83	4,7	Previously named ORFC1			Johnsen *et al.* (1996)
								ORF33 in *S. thermophilus* phage ΦO1205	2e-03	55 (38)	U88974
								GP113 in *S. thermophilus* phage Sfi11	4e-03	55 (38)	AF057033
39		GATAGGGAGGT	20762	21073	103	12,18	5,0	Previously named ORFB3			Johnsen *et al.* (1996)
								GP104 in *S. thermophilus* phage Sfi11	3e-03	93 (34)	AF057033
								ORF34 in *S. thermophilus* phage ΦO1205	6e-03	105 (32)	U88974

TABLE 1—*Continued*

ORF	Gene name	RBS AGAAAGGAGGT	Start	Stop	Size of protein aa	Size of protein kDa	pI	Function and identity	e-value	Identity[a] aa (%)	Reference or accession number
40		TTGCAGGAGGT	21070	21408	112	12,46	9,6	Previously named ORFA1			Johnsen *et al.* (1996)
								ORF35 in *S. thermophilus* phage ΦO1205	4e-06	111 (28)	U88974
								GP114 in *S. thermophilus* phage Sfi11	4e-06	111 (28)	AF057033
								GP16.1 in *B. subtilis* phage SPP1	4e-04	115 (26)	X97918
41		nd	21405	21794	129	14,76	6,6	Previously named ORFX			Johnsen *et al.* (1996)
								ORFX in *L. lactis* phage Tuc2009	2e-68	129 (99)	L31366
								ORF36 in *S. thermophilus* phage ΦO1205	1e-11	126 (29)	U88974
								GP128 in *S. thermophilus* phage Sfi11	1e-11	126 (29)	AF057033
42	*mtp*	AGATAGGAGAT	21805	22314	169	18,63	4,8	Major tail protein			Johnsen *et al.* (1996)
								ORFMP2 in *L. lactis* phage Tuc2009	1e-79	169 (87)	L31366
								ORF37 in *S. thermophilus* phage ΦO1205	2e-23	159 (44)	U88974
								GP168 in *S. thermophilus* phage Sfi11	4e-23	159 (44)	AF057033
								GP17.1 in *B. subtilis* phage SPP1	1e-07	155 (27)	X97918
43		TTTAAGGAGAA	22429	22764	111	12,47	5,0	Previously named ORFA2			Johnsen *et al.* (1996)
								GP117 in *S. thermophilus* phage Sfi11	7e-06	69 (32)	AF057033
								ORF38 in *S. thermophilus* phage ΦO1205	1e-05	97 (32)	U88974
44		AGTTCGGAGGA	22803	23123	106	12,82	9,4	Previously named ORFC2			Johnsen *et al.* (1996)
								ORF39 in *S. thermophilus* phage ΦO1205	5e-06	101 (30)	U88974
								GP105 in *S. thermophilus* phage Sfi11	2e-05	75 (34)	AF057033
45	*tmp*	GGAAAGGAGGA	23138	25951	937	100,3	8,8	Tape measure protein			Pedersen *et al.* (2000)
								ORF15 in *S. aureus* phage ΦPVL	2e-83	631 (30)	AB009866
								ORF36 in *B. subtilis* phage Φ105	1e-65	423 (36)	AB016282
								GPT in *E. coli* phage P2	5e-43	505 (28)	AF063097
								ORF14 in *L. lactis* phage sk1	2e-39	849 (23)	AF011378
								ORF42 in *L. lactis* phage r1t	2e-39	330 (30)	U38906
								ORFI16 in *L. lactis* phage bIL170	1e-32	670 (21)	AF009630
								GP18.1 in *B. subtilis* phage SPP1	1e-18	346 (22)	X97918
46		TTTTAGGAGGT	25961	26722	253	29,13	5,6	GP19.1 in *B. subtilis* phage SPP1	3e-08	131 (32)	X97918
47		AGAAAGGCGGT	26722	29478	918	102,1	5,4	Tail-host specificity			This work
								Putative glycyl-glycine endopeptidase in *D. radiodurans*	2e-18	118 (42)	AE002061
								GP13 in *B. subtilis* phage Φ29	2e-15	277 (29)	M14782
48		AGAAAGGGTCT	29491	30390	299	33,84	5,1	No similarity found			
49	*bpp*	AAAAAGGAGAA	30405	30896	163	17,15	7,7	Baseplate protein			Pedersen *et al.* (2000)
								ORFMP1 in *L. lactis* phage Tuc2009	1e-11	62 (56)	L31366
								ORF1904 in *L. lactis* phage BK5-T	1e-11	104 (31)	L44593
								ORF18 in *L. lactis* phage sk1	1e-06	111 (33)	AF011378
								ORF45 in *L. lactis* phage r1t	2e-04	54 (42)	U38906
50		AGAAAGTAGGG	30910	31134	74	8,65	6,3	ORF75 in *L. lactis* phage BK5-T	4e-35	73 (98)	L44593
51	*nps*	TAAAAGGAATA	31147	33150	667	71,57	5,0	Neck Passage Structure			Johnsen *et al.* (1995)
								ORF47 in *L. lactis* phage r1t	0	667 (94)	U38906
								ORFI12 in *L. lactis* phage bIL170	0	672 (71)	AF009630
								ORFL12 in *L. lactis* phage Φ41	0	594 (70)	L35061
								GP695 in *S. thermophilus* phage Sfi11	3e-10	216 (28)	AF057033
								ORFI15 in *L. lactis* phage c2	1e-09	247 (29)	L48605

TABLE 1—*Continued*

ORF	Gene name	RBS AGAAAGGAGGT	Start	Stop	Size of protein aa	Size of protein kDa	pI	Function and identity	e-value	Identity[a] aa (%)	Reference or accession number
								GP1276 in *S. thermophilus* phage Sfi21	3e-09	222 (31)	AF115103
								GP1291 in *S. thermophilus* phage Sfi19	6e-09	222 (31)	AF115102
52	*hol*	AAATAGGAGAG	33175	33441	88	9,64	9,1	Holin			This work
								ORFS in *L. lactis* phage Tuc2009	4e-44	88 (98)	L31364
								LYSA in *L. lactis* phage ΦLC3	1e-42	88 (96)	U04309
								HOL in *L. casei* phage A2	1e-19	77 (55)	AJ251790
								ORF2 in *S. pneumonia* phage EJ-1	1e-16	83 (49)	S43512
								Holin in *S. aureus* phage ΦPV83	5e-14	80 (43)	AB044554
								Hypothetical protein in *S. aureus*	1e-12	80 (41)	AB033232
								ORF24 in *S. thermophilus* phage DT1	9e-08	79 (36)	AF085222
								ORF141 in *S. thermophilus* phage Sfi11	1e-07	77 (35)	AF158600
								ORF141 in *S. thermophilus* phage Sfi19	2e-07	77 (35)	AF115102
								ORF141 in *S. thermophilus* phage Sfi21	2e-07	77 (35)	AF115103
53	*lys*	AAATCAGAGGA	33438	34727	429	46,2	6,2	Lysin			This work
								ORFA in *L. lactis* phage TPW22	0	429 (94)	AF066865
								LYSB in *L. lactis* phage ΦLC3	0	429 (95)	U04309
								ORFLYS in *L. lactis* phage Tuc2009	0	429 (93)	L31364
54		TTAGAGGAGAG	34912	35247	111	13,36	5,4	ORFDRA0366 in *D. radiodurans*	2e-05	45 (46)	AE001863
55		ACAATAAAGGT	35302	35973	223	26,42	5,1	No similarity found			
56		AAATTGGAGAA	36372	37628	418	49,03	5,4	Hypothetical protein in *N. meningitis*	8e-09	313 (24)	AL162756
								ORF261 in *S. aromaticivorans* plasmid pNL1	4e-06	125 (32)	AF079317
								Hypothetical protein in *H. pylori* plasmid pHPO100	6e-05	219 (23)	AF056496

[a] The number of amino acids, over which the percentage identity was determined, is indicated and followed by the percentage identity in parentheses.

Functional modules

A successful bacteriophage infection requires gene expression and regulation, DNA replication, formation of the phage capsid, and release of new phage particles from the infected host. In most bacteriophages the genes encoding these basic functions are clustered according to biological function, and the term "functional modules" was defined as a stretch of genes with related function (Botstein, 1980; Casjens *et al.,* 1992). The genomes of lambdoid phages have been divided into 11 major segments of functionally clustered genes. These are integration/excision, homologous recombination, early gene control, DNA replication, late gene control, head, tail, lysis, and three nonessential regions (Casjens *et al.,* 1992). Inspection of the identified functions encoded by the bacteriophage TP901-1 reveals that the genome can be divided into functional modules. The following functional modules are proposed and indicated in Fig. 1: site-specific recombination, early gene control, excision, DNA replication, late gene control, DNA packaging, head structural components and assembly, tail structural components and assembly, other structural components, and cell lysis.

The site-specific recombination module contains the TP901-1 integrase (ORF1) and the *attP* site necessary for site-specific integration during the establishment of lysogeny (Christiansen *et al.,* 1996; Brøndsted and Hammer, 1999). The integrase of TP901-1 shows homology to the putative integrase of *B. subtilis* phage Φ105 and to the integrase of *L. monocytogenes* phage A118 (Table 1) (Loessner *et al.,* 2000).

The module involved in early gene control contains the phage repressor encoded by *orf4*, the modulator of repression (ORF5), and the divergently located promoters P_L and P_R (Madsen and Hammer, 1998; Madsen *et al.,* 1999). ORF4 and ORF5 show homology to equivalently located proteins of *S. aureus* ΦPVL phage and *S. thermophilus* phage ΦO1205 (Table 1) (Stanley *et al.,* 1997; Kaneko *et al.,* 1998).

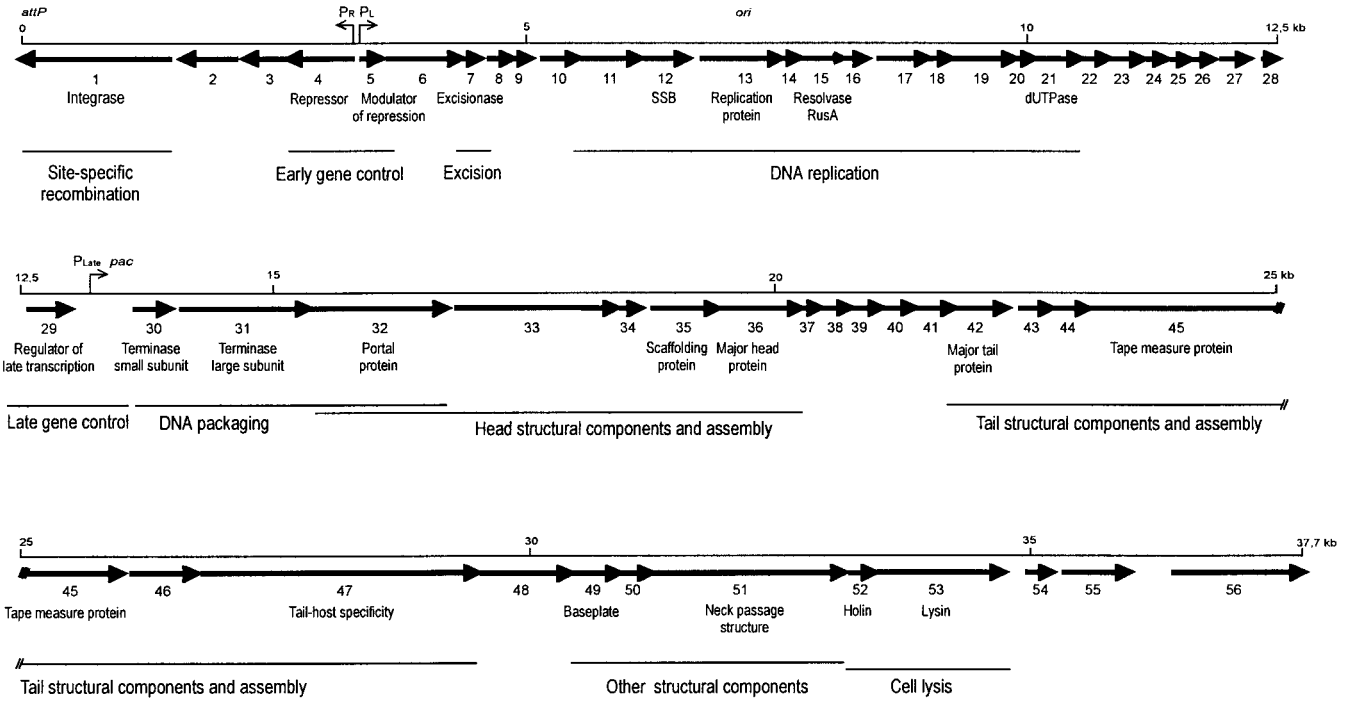

FIG. 1. Genetic organization of the bacteriophage TP901-1 genome. Open reading frames (ORFs) are numbered consecutively and arrows indicate the size and the direction of transcription of the ORFs. Above the arrows a ruler indicates the position on the 37,677-bp TP901-1 genome and above this DNA sites (*attP*, promoters, *ori*, *pac*) are shown. Assigned functions are written under the number of the ORFs. Thin lines show the extent of the suggested functional modules and the names of the functional module are indicated.

The excision module contains only the TP901-1 excisionase (ORF7). However, a functional integrase is required to excise the TP901-1 genome from the bacterial chromosome of a lysogenic strain (Breüner *et al.,* 1999).

We propose that the replication module covers *orf11* to *orf21* of the TP901-1 genome, since several functions related to DNA replication are located within this region. ORF11 was suggested to encode a topoisomerase I (Madsen and Hammer, 1998), while ORF12 shows high homology to single-stranded DNA binding proteins from a variety of bacteria and phages (Table 1). ORF13 was shown to be necessary for replication of the TP901-1 genome and repeats located within *orf13* were suggested to function as the origin of replication (Østergaard *et al.,* 2001). ORF12 and ORF13 are homologous to proteins encoded by *S. aureus* ΦPVL and by *L. monocytogenes* phage A118 (Kaneko *et al.,* 1998; Loessner *et al.,* 2000). ORF15 shows homology to the RusA endonucleases from *Escherichia coli* and coliphage 82, which have been shown to resolve Holliday junction intermediates formed during homologous recombination and DNA repair (Sharples *et al.,* 1994). Thus, the putative RusA endonuclease (ORF15) of TP901-1 could possibly resolve complex DNA structures formed during DNA replication. ORF21 shows high homology to dUTPases that catalyzes the hydrolysis of dUTP to dUMP and pyrophosphate. By containing a separate dUTPase enzyme, the phage may ensure that the intracellular concentration of dUTP is maintained at a low level, thereby preventing the synthesis of mutagenic uracil-substituted DNA and furthermore that a sufficient pool of dUMP, an essential precursor in the biosynthesis of dTTP, is present (Shlomai and Kornberg, 1978). The functions of the ORF14, ORF16, ORF17, ORF18, ORF19, and ORF20 are currently unknown.

It is not possible to assign ORF22 to ORF28 to any functional module since the biological function of these ORFs are unknown. These ORFs are expressed in the early and middle phase of the lytic cycle (Madsen and Hammer, 1999). ORF22 shows homology to ORFs of unknown function in the lactococcal phages ul36.1, Φ31.1 and r1-t, ORF23 to ORFs from Φ31.1 and ul36, ORF26 to ORFs in phage 712 and Φ31.1, and ORF27 to the ORF4 in the *abiN* operon (Table 1). For the remaining three ORFs, no similarities were found in the databases.

ORF29 is a positive regulator of late transcription and is required for initiation of transcription of the late promoter located upstream of *orf30* (L. Brøndsted, unpublished data). The regulator of late transcription and the late promoter constitute the late gene control module (Fig. 1).

Threading of DNA into procapsid shells during DNA packaging requires phage-specified DNA packaging proteins, called terminases, that interact directly with the substrate DNA (Bazinet and King, 1985). The DNA-interaction sites (*pac* or *cos*) of the terminases are located either within or close to the structural genes for the

terminase protein (Black, 1989). The *pac* site of TP901-1 is located upstream of *orf30* (Fig. 1) (Johnsen, 1995). Furthermore, the deduced amino acid sequence of ORF30 shows homology to the small terminase subunit of *Haemophilus influenzae* (Table 1), while ORF31 shows weak homology to the large terminase subunit of the *E. coli* phages T3 and T7 (data not shown). In addition, a putative DNA binding domain (helix-turn-helix motif) and an ATP binding domain could be found in ORF30, as well as an ATP binding domain in ORF31. Although not well conserved, some similarity to motifs present in *B. subtilis* phage SPP1 terminase subunits can be found (data not shown). This indicates that ORF30 and ORF31 encode the terminase of TP901-1. During DNA packaging, the portal protein associates with the terminase protein (Black, 1989). In TP901-1 ORF32 shows homology to the portal protein of the *B. subtilis* phage SPP1, suggesting that ORF32 is the TP901-1 portal protein involved in DNA packaging (Table 1). Thus, the DNA packaging module of TP901-1 contains the *pac* site, the two terminase subunits, and the portal protein.

One major head protein encoded by *orf36* has been identified in the TP901-1 genome (Johnsen *et al.,* 1995, 1996). Three *orf*s are located between the genes encoding the portal protein and the major head protein in TP901-1 genome (Fig. 1). In many bacteriophages one of these genes encodes the scaffolding protein. Since ORF35 has a molecular weight and isoelectric points (pI) value similar to scaffolding proteins of several bacteriophages (λ, SPP1, A118, *Sfi*11, and r1t), we propose that *orf35* encodes the scaffolding protein of TP901-1. ORF32 and ORF33 of TP901-1 show homology to *S. thermophilus* phage *Sfi*1 and ΦO1205 encoded proteins (Table 1). ORF33 and ORF34 may be involved in the earliest stages of procapsid assembly, since these functions are often located between the portal protein and the major head protein. The head structural components and assembly module may thus contain *orf32* to *orf36*.

In most bacteriophages the genes located between the major head (*orf36*) and major tail (*orf42*) are involved in formation and connection of the head and tail structures and in DNA packaging. In this region of the TP901-1 genome, seven proteins (ORF38 to ORF44) showed homology (26–44% identity) to *S. thermophilus* phage ΦO1205- and *Sfi*11-encoded proteins (Table 1). The homology and preservation of gene order indicate that these proteins may interact with each other and may be either a part of the head or tail structural components or involved in their assembly.

The functional module for tail structural components and assembly is proposed to cover *orf42* to *orf47*. The major tail protein (ORF42) and a tape measure protein (ORF45), important for tail length and assembly of the TP901-1 tail, have been identified (Johnsen *et al.,* 1996; Pedersen *et al.,* 2000). In many bacteriophages the genes located between the major tail and tape measure proteins are involved in the formation of a tail initiator complex onto which the major tail protein can polymerize. ORF46 shows homology to GP19.1 of *B. subtilis* phage SPP1, which is, however, of unknown function (Table 1). TP901-1 ORF47 shows sequence homology to a putative cell wall glycyl-glycine endopeptidase of the bacterium *Deinococcus radiodurans* R1, suggesting that ORF47 may be involved in host recognition.

A baseplate component (ORF49) and a neck passage structure (ORF51) that radiates from the connection between head and tail of TP901-1 are located in the functional module of other structural components (Johnsen *et al.,* 1995; Pedersen *et al.,* 2000). It has been shown that the baseplate and tail structures of TP901-1 assemble through a branched baseplate and tail assembly pathway (Pedersen *et al.,* 2000). ORF51 shows high homology to ORF47 from phage r1t and to ORFs encoded by phages belonging to the lactococcal 936 phage group (Table 1). Homology was also observed to ORFs described as putative antireceptors from the three *S. thermophilus* phages *Sfi*11, *Sfi*19, and *Sfi*21 (Table 1), suggesting that these genes all encode neck passage structures (Desiere *et al.,* 1999; Lucchini *et al.,* 1999).

The lysis module of TP901-1 consists of a holin protein (ORF52) and a murein hydrolase protein (ORF53). These proteins are almost identical to the holin and lysin of the lactococcal phages ΦLC3 and Tuc2009 (Table 1). The lysin of both phages has been shown to have lytic activity against *L. lactis* (Arendt *et al.,* 1994; Birkeland, 1994). Many phages are reported to have a dual-start motif for translation of their holin gene (Young and Bläsi, 1995). However, a dual start motif was not identified in either TP901-1 ORF52 or in phage ΦLC3 LysA (Birkeland, 1994). Based on sequence homology to the Tuc2009 Lys protein, ORF53 most probably has *N*-acetyl-muramidase activity (Arendt *et al.,* 1994; Sheehan *et al.,* 1996).

Some regions have not been assigned to any functional module due to lack of information regarding the biological functions of the encoded genes. These regions are primarily located in the early region of the TP901-1 genome, but also include the three genes located downstream of *orf53*. This latter region could, however, be nonessential because insertion of an erythromycin marker gene in *orf55* had no effect on phage viability (Koch *et al.,* 1997).

Protein comparison to the *L. lactis* subsp. *lactis* IL1403 chromosome

TP901-1 encoded proteins were compared to available protein sequences of *L. lactis* subsp. *lactis* IL1403; homologies to proteins encoded by three different prophages were found (data not shown). Most pronounced is the similarity to prophage pi2 (Bolotin *et al.,* 1999), in which proteins nearly identical to ORF2 through ORF6 of TP901-1 are found. In addition, significant homologies

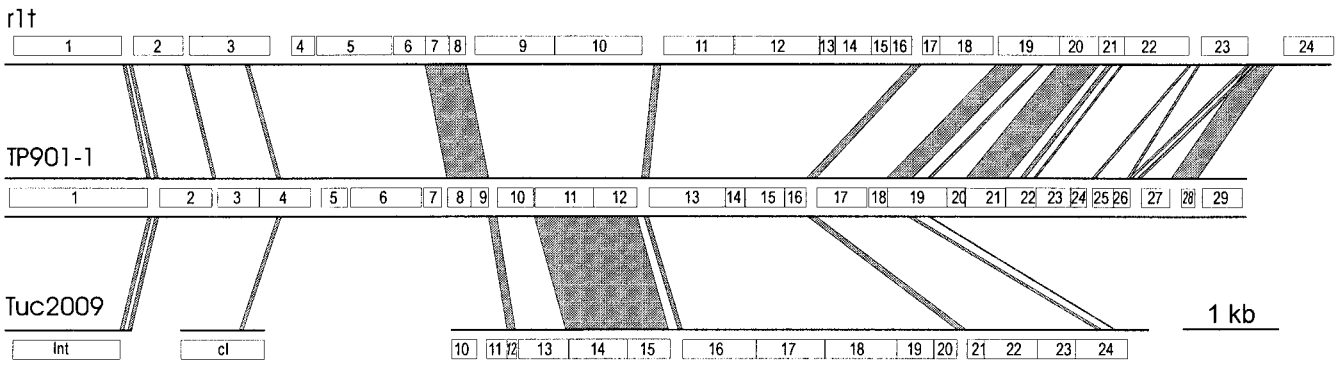

FIG. 2. Comparison of the early transcribed region of the TP901-1 genome with available genome sequences of temperate lactococcal bacteriophages r1t and Tuc2009. Boxes indicate open reading frames. Boxed regions in gray show sequences of high DNA identity (85–100%). For protein homologies, see Table 1.

are present interspersed all over the three prophage genomes. All three prophages also contain a dUTPase similar to TP901-1 ORF21, and prophage pi1 and pi3 both carry proteins highly similar to the TP901-1 regulator of late transcription (ORF29). Homologies to the following nonphage-encoded proteins in *L. lactis* subsp. *lactis* IL1403 were also found: a putative single-stranded DNA binding protein, a putative dUTPase, a putative replication protein, a putative *N*-acetylmuramoyl-L-alanine amidase, and an *N*-acetylmuramidase (data not shown). At present the complete DNA sequence of the *L. lactis* subsp. *lactis* IL1403 chromosome is not available nor is information available on whether the above prophages are inducible or are part of defective phage genomes.

DNA–DNA comparisons of phage genomes

DNA–DNA comparisons were made to obtain information about evolution of bacteriophages and particularly to identify regions with high probability of homologous recombination. Threshold levels for these comparisons were set to 85% identity.

The DNA sequence of the TP901-1 genome was compared to other *L. lactis* phage genomes. The genomes of lactococcal bacteriophages r1t (fully sequenced, van Sinderen *et al.,* 1996) and Tuc2009 (partial sequence, van de Guchte *et al.,* 1994a,b; McGrath *et al.,* 1999), which belong to the same P335 phage species as TP901-1, contain several regions with high homology (85–100%) to the early region of TP901-1 (Fig. 2). Both short (21–51 bp) and long (up to 2037 bp) regions of high homology were located in both coding and noncoding regions. Only a small part of the TP901-1 late region showed high homology to phage r1t: a short region of *orf49* (baseplate) and the entire *orf51* (neck passage structure). In contrast, significant homologies between Tuc2009 and TP901-1 were found in *orf42* (major tail protein) and in the region downstream of this gene. Furthermore, a 1752-bp TP901-1 region covering the last part of *orf51*, *orf52* (holin), and *orf53* (lysin) was highly homologous to the Tuc2009 genome (Arendt *et al.,* 1994).

Comparisons were also performed with available sequences of the virulent isometric-headed lactococcal bacteriophages Φ31 and ul36, which also belong to the P335 species. Several regions of the partly sequenced phage ul36 (Bouchard and Moineau, 2000) showed high homology to TP901-1 (Fig. 3A). Furthermore, the recombinant phage ul36.1, derived from ul36, showed extensive homology to TP901-1; a stretch of five genes (1824 bp) is almost identical (97% identity) in these two phages (Fig. 3A). By sequence analysis it was shown that phage ul36.1 has obtained 5.1 kb of the host *L. lactis* chromosome compared to its parent ul36 (Bouchard and Moineau, 2000). Comparison of TP901-1 with available sequences of the virulent phage Φ31 only revealed high DNA homology in the TP901-1 genes encoding the repressor (ORF4) and *orf6* as well as in two short regions (Dinsmore and Klaenhammer, 1997; Walker *et al.,* 1998; Walker and Klaenhammer, 1998; Accession No. AJ292531). In contrast, the recombinant phage Φ31.1 showed high DNA identity to the region covering *orf6* to *orf26* of the TP901-1 genome (Fig. 3B). Phage Φ31.1 contains 7.8 kb new DNA when compared to its parent Φ31, and it has been shown that Φ31.1 acquired host chromosomal DNA by homologous recombination (Durmaz and Klaenhammer, 2000). Homology to TP901-1 is located within the recombinant 7.8-kb region and in particular between TP901-1 *orf14* to *orf23*.

The DNA sequence of the *abiN* operon of *L. lactis* subsp. *cremoris* S114 shows extensive homology to the TP901-1 genome in the region covering *orf18* to *orf32*, but no homology to the *abiN* gene was found (Fig. 4). The *abiN* operon was isolated as a chromosomal fragment containing 10 genes, including *abiN*, and the presence of *abiN* was shown to give rise to an abortive infection phenotype (Prévots *et al.,* 1998).

In addition to conserved segments located in coding regions of the phage genomes, high homology to noncoding regions are found. Nine intergenic regions of the TP901-1 genome show significant homology to intergenic regions of other phages, and some of these se-

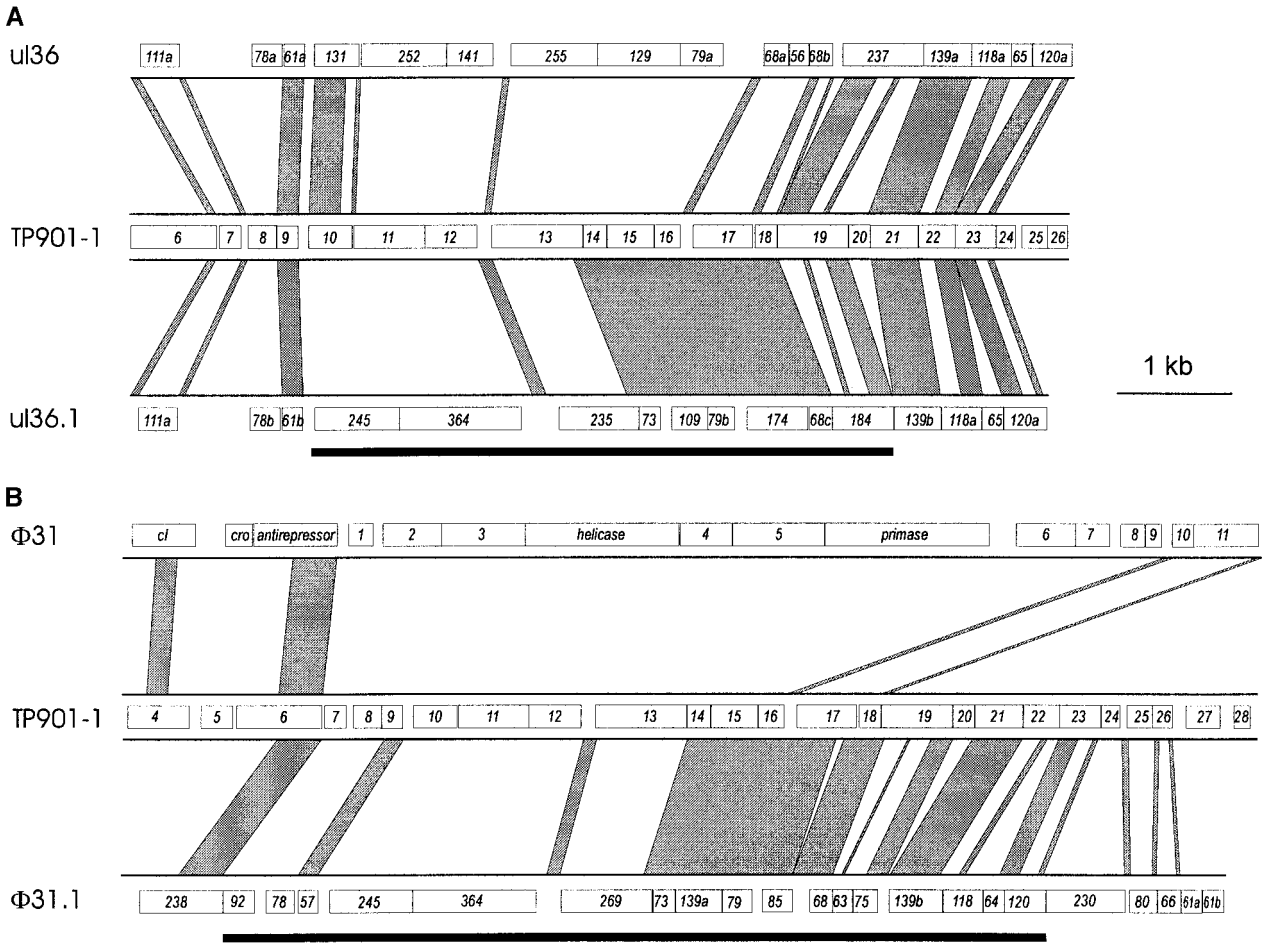

FIG. 3. Homology of the TP901-1 genome to the virulent lactococcal bacteriophages: (A) ul36 and ul36.1; (B) Φ31 and Φ31.1. Boxes indicate open reading frames. Boxed regions in gray show sequences of high DNA identity (85–100%). Black bars indicate the DNA region acquired from the *L. lactis* chromosome in phage ul31.1 and Φ31.1. For protein homologies, see Table 1.

quences are present in several phages (Table 2). These homologous regions are often located between genes with very little or no significant homology to each other. Only the intergenic regions between *orf12–13* and *orf16-17* are in some phages located next to homologous genes.

No homology or only limited homology at the DNA level was found to available sequences of prolate-headed lactococcal bacteriophages of the c2 and 936 species (bIL67, c2, bIL170, and sk1). In the genome of the partly sequenced temperate phage BK5-T (BK5-T phage species), a few short sequences of homology were found, as well as a 232-bp region with high homology to sequences in *orf50* of TP901-1 (data not shown) (Boyce *et al.,* 1995a,b).

The genome sequence of TP901-1 was also compared to those of bacteriophages infecting bacteria other than *L. lactis*. Only a few short sequences of significant homology were found to *S. thermophilus* phages *Sfi*19, *Sfi*21, ΦO1205, to *S. aureus* phage ΦPVL, and to *L. monocytogenes* phage A118 (Stanley *et al.,* 1997; Desiere *et al.,* 1999; Kaneko *et al.,* 1998; Loessner *et al.,* 2000). These

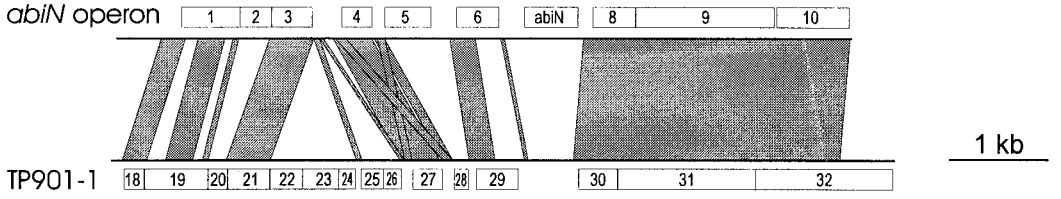

FIG. 4. Homology of the TP901-1 genome to the *abiN* operon of *L. lactis* subsp. *cremoris* S114. Boxes indicate open reading frames. Boxed regions in gray show sequences of high DNA identity (85–100%). For protein homologies, see Table 1.

TABLE 2

Homology[a] of TP901-1 Intergenic Regions to Lactococcal Bacteriophages and Chromosomal Fragments of *L. lactis*

TP901-1 region	Size (bp)	r1t	Tuc2009	ΦLC3	BK5-T	Φ31	Φ31.1	ul36	ul36.1	*abiN*
orf1-2	123	26 (96) 31 (90)	26 (96) 31 (90)	26 (96) 31 (90)	22 (95)	—[b]	—	—	—	—
orf2-3	61	41 (92)	—	41 (92)	33 (92)	—	—	—	—	—
orf7-8	67	—	—	—	26 (100) 28 (92)	—	—	56 (92)	56 (92)	—
orf9-10	101	—	95 (100)	—	—	—	—	—	—	—
orf12-13	129	118 (97)	39 (94)	—	—	—	121 (96)	39 (94)	121 (96)	—
orf16-17	108	73 (97)	89 (95)	—	—	89 (93)	108 (96)	89 (94)	108 (96)	—
orf24-25	54	30 (100)	—	—	30 (100)	—	52 (96)	—	—	54 (98)
orf26-27	129	44 (100) 30 (100)	—	—	—	—	—	—	—	59 (100) 45 (100) 29 (96)
orf27-28	124	121 (94)	—	—	—	—	—	—	—	49 (100) 44 (100) 34 (97) 29 (93)

[a] The number of bp over which the percentage identity was determined is indicated and followed by the percentage identity in parentheses.
[b] —, identity less than 85%.

were located in the early gene control module of TP901-1 (ΦO1205 and ΦPVL), in *orf12* encoding SSB (A118), and in *orf51* encoding the neck passage structure of TP901-1 (*Sfi*19 and *Sfi*21) (data not shown).

DISCUSSION

The biological functions of 11 bacteriophage TP901-1 genes were determined previously by experiment. Analysis of the entire genome has now allowed us to assign a probable function to 10 more genes, resulting in functions for 21 of the 56 proteins encoded by TP901-1 (Table 1 and Fig. 1). Based on related functions, the genome of TP901-1 has been subdivided into 10 modules (Fig. 1). The biological functions of the modules of TP901-1 are essentially those defined for phage λ, except for the absence of a recombination module in TP901-1 and the presence of a module containing the baseplate and neck passage structure (which λ lacks). The integrase and excisionase of TP901-1 are also located in separate modules; excisionase of TP901-1 is separated from integrase by five genes and is transcribed in the opposite direction. Interestingly, most of the proteins present in the early transcribed region of the TP901-1 genome that have not been assigned a function show homology primarily to other lactococcal bacteriophages (Table 1), suggesting that the genes are specifically required for growth in lactococci. No function has been identified for genes located in the one nonessential region of the TP901-1 genome, downstream of the holin and lysin genes.

The genetic organization of temperate lactococcal bacteriophages appears similar; they all have a small lysogenic operon containing at least integrase and repressor genes, and a large divergently located gene cluster involved in lytic growth (Figs. 1 and 2). Similar to TP901-1, these phages have a modular type of genome organization and the order of modules is essentially identical, but they may encode completely different proteins that perform analogous functions. Thus, genome structure is more highly conserved than sequence. A new phage arising after exchange of a functional module or gene is more likely to form a viable or even superior phage if the gene order is constant among the population of phages exchanging genetic material. It also seems necessary that genes encoding proteins that bind to *cis*-elements are located close to those elements. In TP901-1, the *attP* site is located next to integrase (but not excisionase); the early promoters P_L and P_R are located between the genes controlling promoter activity (repressor and modulator of repression); the origin of replication is located within the gene encoding the replication protein; the late promoter is located downstream of its regulator, and the *pac* site is upstream of both terminase subunits (Fig. 1). Thereby the chance of exchanging genes together with the corresponding *cis*-elements is increased. Furthermore, the location of *cis*-elements close to genes encoding proteins interacting with the *cis*-elements may have a role in phage physiology. Maintenance of a constant gene order may optimize the physiology of infection and at the same time facilitate evolution.

The organization of the head and tail structural and assembly modules seems to be conserved between TP901-1 and *S. thermophilus* phage ΦO1205 and *Sfi*11, since four genes in both phages separate the homologous proteins (data not shown). Similarity is less with the

B. subtilis phage SPP1, although the order of the homologous genes is preserved (data not shown). Thus, the lactococcal bacteriophage TP901-1 is more closely related to *S. thermophilus* phage ΦO1205 than to *B. subtilis* phage SPP1, which is also the case for the respective hosts. *L. lactis* is phylogentically closer to *S. thermophilus* than to *B. subtilis*. The presence of homologous genes and the preserved gene order in the head and tail morphogenesis modules suggest that these genes may have a common ancestor.

The genomes of lactococcal bacteriophages are mosaic in nature, with regions of obvious sequence similarity interspersed with regions that are apparently unrelated. Similar observations were made in lambdoid and *S. thermophilus* phages (Casjens *et al.,* 1992; Tremblay and Moineau, 1999). The homologies among lactococcal bacteriophages indicate that these phages have a common ancestor or that genetic material has been exchanged between these phages. It has been suggested that homologous recombination between short regions of microhomology, for example, conserved sites such as promoters or terminators, could be involved in genetic exchange between bacteriophages (Casjens *et al.,* 1992). Thus, the short homologous sequences in noncoding regions present in several lactococcal bacteriophages could be points of exchange of genetic material (Table 2). In some cases this could result in exchange of the intervening regions that lack homology. Recently, recombination in a small noncoding region has been shown to occur during the formation of a recombinant phage by recombination between phage ul36 and the host *L. lactis* chromosome (Bouchard and Moineau, 2000). This sequence is located between *orf12* and *orf13* in TP901-1 and is also present in other lactococcal bacteriophages (Table 2).

The two intergenic regions surrounding *orf13–orf16* of TP901-1 are present in phage r1t, Tuc2009, and ul36 (Table 2), indicating the possibility that homologous recombination could exchange the intervening regions (Figs. 2 and 3). It has been shown that *orf13* and *orf16* of TP901-1 and Tuc2009, respectively, encode the replication protein and that the phage origins are located within these genes (McGrath *et al.,* 1999; Østergaard *et al.,* 2001). Furthermore, results indicate that *orf11* of r1t and *orf235* of ul36 have a similar function (van Sinderen *et al.,* 1996; Bouchard and Moineau, 2000). Thus, by this exchange a recombinant phage could acquire only a new origin of replication and replication protein (the other *orfs* that would be exchanged may also have a role in DNA replication). Interestingly, by this hypothesis, the replication accessory proteins SSB (TP901-1 ORF12; Tuc2009 ORF15) and r1t DnaC (ORF10) would not be exchanged (van Sinderen *et al.,* 1996; McGrath *et al.,* 1999; Østergaard *et al.,* 2001). The hypothesis thus suggests that these accessory proteins may function with different initiation proteins.

Whether the identified homologous sequences in noncoding regions of the lactococcal phage genomes have been conserved for the ability to exchange modules or for function remains unknown. So far only the intergenic region between *orf1* and *orf2* has been suggested to encode a rho-independent terminator (Christiansen *et al.,* 1996). However, a 30-bp sequence present between *orf24* and *orf25* is identical to a region located immediately downstream of the −10 sequence of an identified BK5-T promoter (Lakshmidevi *et al.,* 1990). Functions for the remaining intergenic regions present in several lactococcal bacteriophages have not yet been proposed.

It has been suggested that homologous recombination may occur not only at the boundaries of genes but also at protein domain boundaries within genes (Neve *et al.,* 1998; Juhala *et al.,* 2000). TP901-1 *orf19* shows high DNA homology to several phages (Figs. 2, 3, and 4). However, these high-homology regions are either located in the first or in the last part of the gene, indicating that the *orf19* homologous genes have been created by recombination at protein domain boundaries within the gene.

In general, the TP901-1 genome seems to be more homologous to regions originating from the host chromosome than to previously isolated phages (r1t, Tuc2009, Φ31, and ul36). This suggests that TP901-1 may have evolved by homologous recombination between the host chromosome and a mother phage. TP901-1 shows extensive DNA homology to chromosomal fragments of three *L. lactis* strains (Figs. 3 and 4). These are the *abiN* operon, and the fragments acquired from the host chromosome by phages Φ31 and ul36 during formation of the recombinant phages Φ31.1 and ul36.1, respectively (Prévots *et al.,* 1998; Bouchard and Moineau, 2000; Durmaz and Klaenhammer, 2000). Our results further support the idea that prophages and prophage remnants in the *Lactococcus* chromosome contribute significantly to bacteriophage evolution. Recombination between an incoming phage and the host chromosome should occur more frequently than recombination between two phages, which requires simultaneous infection of the same cell. Since lysogeny is widespread in *Lactococcus*, this is likely to be a major cause of rapid evolution of lactococcal bacteriophages.

Homology to proteins encoded by phages infecting *S. thermophilus*, *S. aureus*, and *L. monocytogenes* was found in the early control module, indicating that this specific control module is present in bacteriophages infecting hosts other than *L. lactis* (Table 1) (Stanley *et al.,* 1997; Kaneko *et al.,* 1998; Loessner *et al.,* 2000). Furthermore, two important proteins of the replication module (SSB and the replication protein) were found in phages infecting *S. aureus* and *L. monocytogenes* (Table 1) (Kaneko *et al.,* 1998; Loessner *et al.,* 2000). In addition to this the integrase of TP901-1 showed homology to integrases of *B. subtilis* and *L. monocytogenes* phages and no homology to *L. lactic* phages is found (Table 1)

TABLE 3

Bacterial Strains and Plasmids

	Description	Antibiotic resistance	Source
Bacterial strains			
L. lactis subsp. *cremoris*			
901-1	Lysogenic for TP901-1		Braun *et al.* (1989)
3107	Indicator strain for TP901-1		Braun *et al.* (1989)
Escherichia coli			
XL1-Blue MRF′	Δ(*mcrA*)*183*Δ(*mcrCB-hsdSMR-mrr*)*173 endA1 supE44 thi-1 recA1 gyrA96 relA1 lac*[F′*proAB lacI*^{q}*Z*Δ*M15* Tn*10*]	Tc	Stratagene
Plasmids			
pBluescriptIISK+	*lacZ*α	Ap	Stratagene
pGEM7-Zf(+)	*lacZ*α	Ap	Promega
pGEM5-Zf(−)	*lacZ*α	Ap	Promega
pCI3340		Cm	Hayes *et al.* (1990)
pG7f1 to pG7f13	*Eco*RI library of TP901-1 cloned in pGEM7-Zf(+)	Ap	Christiansen *et al.* (1994)
pG5f1, pG5f3 to pG5f12	*Eco*RV partial library of TP901-1 cloned in pGEM5-Zf(−)	Ap	Christiansen *et al.* (1994)
pBf2-1 and pBf2-2	*Eco*RV-*Eco*RI fragments of *Eco*RV fragment 2 cloned in pBluescriptIISK+	Ap	Christiansen *et al.* (1994)
pBP1B, pBP2B, pBP3a/b to pBP7a/b	*Pst*I partial library of TP901-1 cloned in pBluescriptIISK+	Ap	This study
pLB82	*Acc*I fragment of pBf2-1 cloned in pGEM7-Zf(+)		
pCla3	*Cla*I fragment 3 of TP901-1 cloned in pBluescriptIISK+	Ap	M. G. Johnsen, unpublished
pCla4	*Cla*I fragment 4 of TP901-1 cloned in pBluescriptIISK+	Ap	M. G. Johnsen, unpublished
pS40C2	*Cla*I fragment 2 of TP901-1 cloned in pCI3340	Cm	This study
pS40C3	*Cla*I fragment 3 of TP901-1 cloned in pCI3340	Cm	Pedersen *et al.* (2000)

(Loessner *et al.,* 2000). In some of these cases the homology to lactococcal phage-encoded proteins is less pronounced than to proteins encoded by phages infecting hosts other than *L. lactis*. This protein homology argues for the occurrence of horizontal genetic exchange among these bacteriophages and supports the model by Hendrix *et al.* (1999), suggesting that phage genomes have access, by horizontal exchange, to a large common genetic pool.

MATERIALS AND METHODS

Bacterial strains, plasmids, and bacteriophages

Bacterial strains and plasmids used are shown in Table 3. Lactococcal strains were grown at 30°C in M17-broth or -agar (Terzaghi and Sandine, 1975) containing 0.5% glucose (GM17). Bacteriophage titers were determined on GM17 containing 5 mM $CaCl_2$ as described by Terzaghi and Sandine (1975). *E. coli* strains were grown at 37°C with agitation in LB broth (Sambrook *et al.,* 1989). Solid media contained 15 g of agar per liter. When required, 20 μg/ml chloramphenicol (Cm), 100 μg/ml of ampicillin (Ap), 12.5 μg/ml of tetracycline (Tc), or 40 μg/ml of 5-bromo-4-chloro-3-indolyl-β-D-galactopyranoside (X-Gal), and 1 mM isopropyl-β-D-thiogalactoside (IPTG) were added. *E. coli* was made competent by the $CaCl_2$ procedure (Sambrook *et al.,* 1989).

Phage TP901-1 was induced from *L. lactis* subsp. *cremoris* 901–1 by UV light as previously described (Christiansen *et al.,* 1994). Bacteriophages were further purified by CsCl step gradients as described for bacteriophage λ (Sambrook *et al.,*1989).

DNA techniques

Phage DNA was isolated as described for bacteriophage λ (Sambrook *et al.,*1989). Plasmid DNA from *E. coli* strains was isolated by the alkaline lysis procedure (Sambrook *et al.,*1989) using the Pharmacia Flexiprep Kit (Amersham Pharmacia Biotech, Little Chalfont, U.K.) or the QIAGEN columns (QIAGEN GmbH., Hilden, Germany), as recommended by the suppliers. Restriction enzymes (New England Biolabs Inc.), T4-ligase (New England Biolabs or Amersham Pharmacia Biotech), calf intestine alkaline phosphatase (CIP) (New England Biolabs), and shrimp alkaline phosphatase (SAP) (Amersham Pharmacia Biotech) were used according to the recommendations of the suppliers. A partial *Pst*I library of TP901-1 (Table 3) was constructed by shotgun ligation of TP901-1 DNA digested with *Pst*I into *Pst*I-digested and SAP-treated pBluescript II SK+. The ligation mix was subsequently transformed into *E. coli* XL1-Blue MRF′. The plasmids pS40C2 and pS40C3 were constructed by shotgun ligation of purified TP901-1 DNA digested with

*Cla*I into *Cla*I-digested and CIP-treated pCI3340. Plasmid pLB82 was constructed by inserting a purified 1.3-kb *Acc*I fragment of pBf2–1 in the *Cla*I site of pGEM7-Zf(+). Nested deletions of library clones were conducted using available restriction sites in existing clones or by using the Double-Stranded Nested Deletion Kit according to the recommendations of the supplier (Amersham Pharmacia Biotech).

DNA sequencing

The complete sequence of TP901-1 is located in GenBank under the Accession No. AF304433. Previously published partial TP901-1 sequences can be found under the Accession No. X84706, X85213, Y14232, AF252967, AF252968, and AY007566 and were obtained as described (Johnsen *et al.,* 1996; Christiansen *et al.,* 1996; Madsen and Hammer, 1998; Pedersen *et al.,* 2000). The remaining sequence of TP901-1 was obtained by analyzing TP901-1 library clones and subclones (Table 3). DNA sequences were determined according to Sanger *et al.* (1977). Sequences were obtained using Cy5-labeled primers and the ThermoSequenase Fluorescent Labeled Primer Cycle Sequencing Kit (Amersham Pharmacia Biotech) or the USB Sequenase version 2.0 DNA sequencing kit and [α-^{33}P]-dATP (Amersham Pharmacia Biotech). Sequence reactions with Cy5-labeled primers were analyzed using an ALFexpress DNA Analysis System (Amersham Pharmacia Biotech). Sequences with the remaining primers were obtained using conventional denaturing polyacrylamide sequencing gels. Custom-made primers were obtained from DNA Technology (Århus, Denmark)

Sequence analysis

Sequences were assembled and analyzed using the Genetic Computer Group (GCG) sequence package from the University of Wisconsin version 9.1 (Devereux *et al.,* 1984). Putative ORFs were also identified using the GenMark program (http://genemark.biology.gatech.edu/GeneMark/webgenemark.html). Molecular weights and theoretical isoelectric points were calculated using the Expasy ProtParam program (http://www.expasy.ch/tools/protparam.html). Databases were searched using the BLAST 2.0 program (http://www.ncbi.nlm.nih.gov/BLAST/) (Altschul *et al.,* 1997) and DNA sequences from lactococcal phages were compared using the Blast 2 Sequences program (http://www.ncbi.nlm.nih.gov/gorf/bl2.html) (Tatusova and Madden, 1999). Comparison of the TP901-1 ORFs to the ORFs of *L. lactis* IL1403 chromosome (Bolotin *et al.,* 1999) was conducted by BLAST searches at http://spock.jouy.inra.fr/.

ACKNOWLEDGMENTS

We thank Mads G. Johnsen for donation of plasmid pCla3 and pCla4. This work was supported by grants from the Danish Government Program for Food Science and Technology (FØTEK) through the Danish Center for Advanced Food Studies, the EC BIOTECH program (BIO-96–0402), and the Carlsberg Foundation.

REFERENCES

Altschul, S. F., Madden, T. L., Schäffer, A. A., Zhang, J., Zhang, Z., Miller, W., and Lipman, D. J. (1997). Gapped BLAST and PSI-BLAST: A new generation of protein database search programs. *Nucleic Acids Res.* **25,** 3389–3402.

Arendt, E. K., Daly, C., Fitzgerald, G. F., and van de Guchte, M. (1994). Molecular characterization of lactococcal bacteriophage Tuc2009 and identification and analysis of genes encoding lysin, a putative holin, and two structural proteins. *Appl. Environ. Microbiol.* **60,** 1875–1883.

Bazinet, C., and King, J. (1985). The DNA translocating vertex of dsDNA bacteriophage. *Annu. Rev. Microbiol.* **39,** 109–129.

Birkeland, N.-K. (1994). Cloning, molecular characterization, and expression of the genes encoding the lytic functions of lactococcal bacteriophage ΦLC3: A dual lysis system of modular design. *Can. J. Microbiol.* **40,** 658–665.

Black, L. W. (1989). DNA packaging in dsDNA bacteriophages. *Annu. Rev. Microbiol.* **43,** 267–292.

Bolotin, A., Mauger, S., Malarme, K., Ehrlich, S. D., and Sorokin, A. (1999). Low-redundancy sequencing of the entire *Lactococcus lactis* IL1403 genome. *Antonie Leeuwenhoek* **76,** 27–76.

Botstein, D. (1980). A theory of modular evolution for bacteriophages. *Ann. N.Y. Acad. Sci.* **354,** 484–490.

Bouchard, J. D., and Moineau, S. (2000). Homologous recombination between a lactococcal bacteriophage and the chromosome of its host strain. *Virology* **270,** 65–75, doi:10.1006/viro.2000.0226.

Boyce, J. D., Davidson, B. E., and Hillier, A. J. (1995a). Sequence analysis of the *Lactococcus lactis* temperate bacteriophage BK5-T, and demonstration that the phage DNA has cohesive ends. *Appl. Environ. Microbiol.* **61,** 4089–4098.

Boyce, J. D., Davidson, B. E., and Hillier, A. J. (1995b). Identification of prophage genes expressed in lysogens of the *Lactococcus lactis* bacteriophage BK5-T. *Appl. Environ. Microbiol.* **61,** 4099–4104.

Braun, V., Hertwig, S., Neve, H., Geis, A., and Teuber, M. (1989). Taxonomic differentiation of bacteriophages of *Lactococcus lactis* by electron microscopy, DNA–DNA hybridization, and protein profiles. *J. Gen. Microbiol.* **135,** 2551–2560.

Breüner, A., Brøndsted, L., and Hammer, K. (1999). Novel organization of genes involved in prophage excision identified in the temperate lactococcal bacteriophage TP901-1. *J. Bacteriol.* **181,** 7291–7297.

Brøndsted, L., and Hammer, K. (1999). Use of the integration elements encoded by the temperate lactococcal bacteriophage TP901-1 to obtain chromosomal single-copy transcriptional fusions in *Lactococcus lactis*. *Appl. Environ. Microbiol.* **65,** 752–758.

Casjens, S., Hatfull, G., and Hendrix, R. (1992). Evolution of dsDNA tailed-bacteriophage genomes. *Semin. Virol.* **3,** 383–397.

Christiansen, B., Johnsen, M. G., Stenby, E., Vogensen, F. K., and Hammer, K. (1994). Characterization of the lactococcal temperate phage TP901-1 and its site-specific integration. *J. Bacteriol.* **176,** 1069–1076.

Christiansen, B., Brøndsted, L., Vogensen, F. K., and Hammer K. (1996). A resolvase-like protein is required for the site-specific integration of the temperate lactococcal bacteriophage TP901-1. *J. Bacteriol.* **178,** 5164–5173.

Davidson, B. E., Powell, I. B., and Hillier, A. J. (1990). Temperate bacteriophages and lysogeny in lactic acid bacteria. *FEMS Microbiol. Rev.* **7,** 79–90.

Desiere, F., Lucchini, S., and Brüssow, H. (1998). Evolution of *Streptococcus thermophilius* bacteriophage genomes by modular exchanges followed by point mutations and small deletions and insertions. *Virology* **241,** 345–356, doi:10.1006/viro.1997.8959.

Desiere, F., Lucchini, S., and Brüssow, H. (1999). Comparative sequence

analysis of the DNA packaging, head, and tail morphogenesis modules in the temperate *cos*-site *Streptococcus thermophilus* bacteriophage Sfi21. *Virology* **260,** 244–253, doi:10.1006/viro.1999.9830.

Devereux, J., Haeberli, P., and Smithies, O. (1984). A comprehensive set of sequence analysis programs for the VAX. *Nucleic Acids Res.* **12,** 387–395.

Dinsmore, P. K., and Klaenhammer, T. R. (1997). Molecular characterization of a genomic region in a *Lactococcus* bacteriophage that is involved in its sensitivity to the phage defense mechanism AbiA. *J. Bacteriol.* **179,** 2949–2957.

Durmaz, E., and Klaenhammer, T. R. (2000). Genetic analysis of chromosomal regions of *Lactococcus lactis* acquired by recombinant lytic phages. *Appl. Environ. Microbiol.* **66,** 895–903.

Forde, A., and Fitzgerald, G. F. (1999). Bacteriophage defence systems in lactic acid bacteria. *Antonie Leeuwenhoek* **76,** 89–113.

Hayes, F., Daly, C., and Fitzgerald, G. F. (1990). Identification of the minimal replicon of *Lactococcus lactis* subsp. *lactis* UC317 plasmid pCI305. *Appl. Environ. Microbiol.* **56,** 202–209.

Hendrix, R. W., Smith, M. C. M., Burns, R. N., Ford, M. E., and Hatfull, G. F. (1999). Evolutionary relationships among diverse bacteriophages and prophages: All the world's a phage. *Proc. Natl. Acad. Sci. USA* **96,** 2192–2197.

Jarvis, A. W. (1989). Bacteriophages of lactic acid bacteria. *J. Dairy Sci.* **72,** 3406–3428.

Jarvis, A. W., Fitzgerald, G. F., Mata, M., Mercenier, A., Neve, H., Powell, I. B., Ronda, C., Saxelin, M., and Teuber, M. (1991). Species and type phages of lactococcal bacteriophages. *Intervirology* **32,** 2–9.

Johnsen, M. G. (1995). Virion proteins and genetics of lactococcal temperate bacteriophage TP901-1. Ph.D. thesis. The Royal Veterinary and Agricultural University, Frederiksberg, Denmark.

Johnsen, M. G., Neve, H., Vogensen, F. K., and Hammer, K. (1995). Virion positions and relationships of lactococcal temperate bacteriophage TP901-1 proteins. *Virology* **212,** 595–606, doi:10.1006/viro.1995.1517.

Johnsen, M. G., Appel, K. F., Madsen, P. L., Vogensen, F. K., Hammer, K., and Arnau, J. (1996). A genomic region of lactococcal temperate bacteriophage TP901-1 encoding major virion proteins. *Virology* **218,** 306–315, doi:10.1006/viro.1996.0199.

Juhala, R. J., Ford, M. E., Duda, R. L., Youlton A., Hatfull, G. F., and Hendrix, R. W. (2000). Genomic sequences of bacteriophage HK97 and HK022: Pervasive genetic mosaicism in the lambdoid bacteriophages. *J. Mol. Biol.* **299,** 27–51, doi:1006/jmbi.2000.3729.

Kaneko, J., Kimura, T., Narita, S., Tomita, T., and Kamio, Y. (1998). Complete nucleotide sequence and molecular characterization of the temperate staphylococcal bacteriophage ΦPVL carrying Panton–Valentine leukocidin genes. *Gene* **215,** 57–67.

Klaenhammer, T. R. (1991). Development of bacteriophage-resistant strains of lactic acid bacteria. *Biochem. Soc. Trans.* **19,** 675–681.

Koch, B., Christiansen, B., Evison, T., Vogensen, F. K., and Hammer, K. (1997). Construction of specific erythromycin resistance mutations in the temperate lactococcal bacteriophage TP901-1, and their use in studies of phage biology. *Appl. Environ. Microbiol.* **63,** 2439–2441.

Lakshmidevi, G., Davidson, B. E., and Hillier, A. J. (1990). Molecular characterization of promoters of the *Lactococcus lactis* subsp. *cremoris* temperate bacteriophage BK5-T and identification of a phage gene implicated in the regulation of promoter activity. *Appl. Environ. Microbiol.* **56,** 934–942.

Lillehaug, D., and Birkeland, N-K. (1993). Characterization of genetic elements required for site-specific integration of the temperate lactococcal bacteriophage ΦLC3 and construction of integration-negative ΦLC3 mutants. *J. Bacteriol.* **175,** 1745–1755.

Lillehaug, D., Nes, I. F., and Birkeland, N.-K. (1997). A highly efficient and stable system for site-specific integration of genes and plasmids into the phage ΦLC3 attachment site (*attB*) of the *Lactococcus lactis* chromosome. *Gene* **188,** 129–136.

Loessner, M. J., Inman, R. B., Lauer, P., and Calender, R. (2000). Complete nucleotide sequence, molecular analysis and genome structure of bacteriophage A118 of *Listeria monocytogenes*: implications for phage evolution. *Mol. Microbiol.* **35,** 324–340.

Lucchini, S., Desiere, F., and Brüssow, H. (1998). The structural gene module in *Streptococcus thermophilus* bacteriophage Φsfi11 shows a hierarchy of relatedness to siphoviridae from a wide range of bacterial hosts. *Virology* **246,** 63–73, doi:10.1006/viro.1998.9190.

Lucchini, S., Desiere, F., and Brüssow, H. (1999). Comparative genomics of *Streptococcus thermophilus* phage species supports a modular evolution theory. *J. Virol.* **73,** 8647–8656.

Ludwig, W., Seewaldt, E., Kilpper-Bälz, R., Schleifer, K. H., Magrum, L., Woese, C. R., Fox, G. E., and Stackebrandt, E. (1985). The phylogenetic position of *Streptococcus* and *Enterococcus*. *J. Gen. Microbiol.* **131,** 543–551.

Madsen, P. L., and Hammer, K. (1998). Temporal transcription of the lactococcal temperate phage TP901-1 and DNA sequence of the early promoter region. *Microbiology* **144,** 2203–2215.

Madsen, P. L., Johansen, A. H., Hammer, K., and Brøndsted, L. (1999). The genetic switch regulating activity of early promoters of the temperate lactococcal bacteriophage TP901-1. *J. Bacteriol.* **181,** 7430–7438.

McGrath, S., Seegers, J. F. M. L., Fitzgerald, G. F., and van Sinderen, D. (1999). Molecular characterization of a phage-encoded resistance system in *Lactococcus lactis*. *Appl. Environ. Microbiol.* **65,** 1891–1899.

Neve, H., Zenz, K. I., Desiere, F., Koch, A., Heller, K. J., and Brüssow, H. (1998). Comparison of the lysogeny modules from the temperate *Streptococcus thermophilus* bacteriophages TP-J34 and Sfi21: Implications for the modular theory of phage evolution. *Virology* **241,** 61–72.

Østergaard, S., Brøndsted, L., and Vogensen, F. K. 2001. Identification of a replication protein and repeats essential for DNA replication of the temperate lactococcal bacteriophage TP901-1. *Appl. Environ. Microbiol.* **67,** 774–781.

Pedersen, M., Østergaard, S., Bresciani, J., and Vogensen, F. K. (2000). Mutational analysis of two structural genes of the temperate lactococcal bacteriophage TP901-1 involved in tail length determination and baseplate assembly. *Virology* **276,** 315–328, doi:10.1006/viro.2000.0497.

Prévots, F., Tolou, S., Delpech, B., Kaghad, M., and Daloyau, M. (1998). Nucleotide sequence and analysis of the new chromosomal abortive infection gene *abiN* of *Lactococcus lactis* subsp. *cremoris* S114. *FEMS Microbiol. Letts.* **159,** 331–336.

Sambrook, J., Fritsch, E. F., and Maniatis, T. (1989). "Molecular cloning: G laboratory manual" (2nd ed.). Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY.

Sanger, F., Nicklen, S., and Coulson, A. R. (1977). DNA sequencing with chain-terminating inhibitors. *Proc. Natl. Acad. Sci. USA* **74,** 5463–5467.

Schleifer, K. H., Kraus, J. Dvorak, C., Kilpper-Bälz, R., Collins, M. D., and Fisher, W. (1985). Transfer of *Streptococcus lactis* and related Streptococci to the genus *Lactococcus* gen. nov. *Syst. Appl. Microbiol.* **6,** 183–195.

Sharples, G. J., Chan, S. N., Mahdi, A. A., Whitby, M. C., and Lloyd, R. G. (1994). Processing of intermediates in recombination and DNA repair: Identification of a new endonuclease that specifically cleaves Holliday junctions. *EMBO J.* **13,** 6133–6142.

Sheehan, M. M., García, J. L., López, R., and Garcia, P. (1996). Analysis of the catalytic domain of the lysin of the lactococcal bacteriophage Tuc2009 by chimeric gene assembling. *FEMS Microbiol. Lett.* **140,** 23–28.

Shlomai, J., and Kornberg, A. (1978). Deoxyuridine triphosphatase of *Escherichia coli*. Purification, properties, and use as a reagent to reduce uracil incorporation into DNA. *J. Biol. Chem.* **253,** 3305–3312.

van de Guchte, M., Daly, C., Fitzgerald, G. F., and Arendt, E. K. (1994a). Identification of *int* and *attP* on the genome of lactococcal bacteriophage Tuc2009 and their use for site-specific plasmid integration in

the chromosome of Tuc2009-resistant *Lactococcus lactis* MG1363. *Appl. Environ. Microbiol.* **60,** 2324–2329.

van de Guchte, M., Daly, C., Fitzgerald, G. F., and Arendt, E. K. (1994b). Identification of the putative repressor-encoding gene *cI* of the temperate lactococcal bacteriophage Tuc2009. *Gene* **144,** 93–95.

van Sinderen, D., Karsens, H., Kok, J., Terpstra, P., Ruiters, M. H. J., Venema, G., and Nauta, A. (1996). Sequence analysis and molecular characterization of the temperate lactococcal bacteriophage r1t. *Mol. Microbiol.* **19,** 1343–1355.

Stanley, E., Fitzgerald, G. F., Le Marrec, C., Fayard, B., and van Sinderen, D. (1997). Sequence analysis and characterization of ΦO1205, a temperate bacteriophage infecting *Streptococcus thermophilus* CNRZ1205. *Microbiology* **143,** 3417–3429.

Tatusova, T. A., and Madden, T. L. (1999). Blast 2 sequences—A new tool for comparing protein and nucleotide sequences. *FEMS Microbiol. Lett.* **174,** 247–250.

Terzaghi, B. E., and Sandine, W. E. (1975). Improved medium for lactic streptococci and their bacteriophages. *Appl. Microbiol.* **29,** 807–813.

Tremblay, D. M., and Moineau, S. (1999). Complete genomic sequence of the lytic bacteriophage DT1 of *Streptococcus thermophilus*. *Virology* **255,** 63–76, doi:10.1006/viro.1998.9525.

Walker, S. A., Dombroski, C. S., and Klaenhammer, T. R. (1998). Common elements regulating gene expression in temperate and lytic bacteriophages of *Lactococcus* species. *Appl. Environ. Microbiol.* **64,** 1147–1152.

Walker, S. A., and Klaenhammer, T. R. (1998). Molecular characterization of a phage-inducible middle promoter and its transcriptional activator from the lactococcal bacteriophage Φ31. *J. Bacteriol.* **180,** 921–931.

Young, R., and Bläsi, U. (1995). Holins: Form and function in bacteriophage lysis. *FEMS Microbiol. Rev.* **17,** 191–205.