Complete nucleotide sequence, molecular analysis and genome structure of bacteriophage A118 of *Listeria monocytogenes*: implications for phage evolution

Martin J. Loessner,[1]* **Ross B. Inman,**[2] **Peter Lauer**[3] **and Richard Calendar**[3]

[1]*Institut für Mikrobiologie, FML Weihenstephan, Technische Universität München, Weihenstephaner Berg 3, 85350 Freising, Germany.*
[2]*Institute for Molecular Virology, University of Wisconsin at Madison, 1525 Linden Drive, Madison, Wisconsin 53706, USA.*
[3]*Department of Molecular and Cell Biology, University of California at Berkeley, 401 Barker Hall, Berkeley, California 94720-3202, USA.*

Summary

A118 is a temperate phage isolated from *Listeria monocytogenes*. In this study, we report the entire nucleotide sequence and structural analysis of its 40 834 bp DNA. Electron microscopic and enzymatic analyses revealed that the A118 genome is a linear, circularly permuted, terminally redundant collection of double-stranded DNA molecules. No evidence for cohesive ends or for a terminase recognition (pac) site could be obtained, suggesting that A118 viral DNA is packaged via a headful mechanism. Partial denaturation mapping of DNA cross-linked to the tail shaft indicated that DNA packaging proceeds from left to right with respect to the arbitrary genomic map and the direction of genes necessary for lytic development. Seventy-two open reading frames (ORFs) were identified on the A118 genome, which are apparently organized in a life cycle-specific manner into at least three major transcriptional units. N-terminal amino acid sequencing, bioinformatic analyses and functional characterizations enabled the assignment of possible functions to 26 ORFs, which included DNA packaging proteins, morphopoetic proteins, lysis components, lysogeny control-associated functions and proteins necessary for DNA recombination, modification and replication. Comparative analysis of the A118 genome structure with other bacteriophages revealed

local, but sometimes extensive, similarities to a number of phages spanning a broader phylogenetic range of various low G+C host bacteria, which implies relatively recent exchange of genes or genetic modules. We have also identified the A118 attachment site *attP* and the corresponding *attB* in *Listeria monocytogenes*, and show that site-specific integration of the A118 prophage by the A118 integrase occurs into a host gene homologous to *comK* of *Bacillus subtilis*, an autoregulatory gene specifying the major competence transcription factor.

Introduction

Listeria monocytogenes is a non-spore-forming, opportunistic Gram-positive pathogen, responsible for severe infections in both animals and humans, which is almost exclusively transmitted via contaminated food. Recurrent outbreaks of Listeriosis (CDC, 1998; Slutsker and Schuchat, 1999) have emphasized the need for a better understanding of not only the molecular pathogenicity mechanisms, but also the possible phenotypic variability of the organism through interaction with specific bacteriophages and the environment. Also desirable is the availability of new tools for biomolecular research on this pathogen.

Information on *Listeria* bacteriophages is insufficient at the molecular level; very little is known about gene expression and the possible function of gene products. We have focused our present work on phage A118, a temperate bacteriophage specific for *Listeria monocytogenes* serovar 1/2 strains, which are often implicated in food-borne outbreaks of the disease (see Slutsker and Schuchat, 1999). The virus was induced by UV irradiation from a *L. monocytogenes* SV 1/2a strain (Loessner and Busse, 1990) that had been isolated from a Camembert soft cheese. A118 has a long, flexible, non-contractile tail of ≈300 nm and an isometric capsid with a diameter of 61 nm (Zink and Loessner, 1992). Therefore, it belongs to the Siphoviridae family of double-stranded DNA bacterial viruses in the order Caudovirales (B1 morphotype), and was taxonomically classified into *Listeria* phage species 2671. The virus adsorbs to the serovar-specific L-rhamnose and D-glucosamine substituents in the cell wall teichoic acids of host cells (Wendlinger *et al.*, 1996). At

30°C, the latent period of the lytic cycle is ≈65 min and is followed by a rise phase of 55 min. A118 is capable of general transduction; some particles package host DNA randomly and can transduce functional genetic markers into other susceptible cells (Hodgson, 2000). An average of 30 progeny virions is released from infected cells after lysis by the combined action of a holin and a unique endolysin (Ply) with L-alanoyl-D-glutamate specificity (Loessner *et al.*, 1995a). Cloning and expression of the *ply118* gene has enabled several biotechnological applications, such as rapid lysis of *Listeria* cells from without (Loessner *et al.*, 1995b) and programmed self-destruction of intracellular attenuated *Listeria* cells followed by release of antigen-encoding foreign DNA into the cytosol of macrophages (Dietrich *et al.*, 1998).

Lysogeny is widespread among *Listeria* strains. Despite the fact that most, if not all, strains carry prophages, many of which are apparently cryptic and are known as monocins (Zink *et al.*, 1995), the potential influence on the phenotype of lysogenized host cells has not been investigated.

Here, we report the entire nucleotide sequence of the A118 genome; the physico-molecular characterization of its linear, circularly permuted, terminally redundant double-stranded DNA molecule; the description and analysis of its open reading frames and genome organization; and functional assignments of gene products, based on N-terminal amino acid sequences and sequence similarities. We have also identified *attP* and *attB* and show that site-specific integration of the A118 prophage occurs into a *comK* homologue of the *L. monocytogenes* host strains.

Results

Nucleotide sequencing of the A118 genome

A shotgun sequencing strategy was used initially. The A118 DNA molecules were partially digested with an enzyme that cuts within a tetranucleotide sequence, and the fragments were cloned into an *E. coli* vector. Plasmids with different insert sizes (1–3 kbp) were selected and sequenced, and the overlapping sequences were assembled into several contiguous stretches. The remaining gaps were closed by sequencing both strands of genomic phage DNA, using individual upstream and downstream primers. A unit genome size of 40 834 bp was finally obtained, with an average sequence redundancy of approximately four. The experimentally obtained restriction maps and partial denaturation maps were in good agreement with those predicted from the sequence, indicating correct assembly of the sequences. The average GC content of A118 was calculated at 36.1 mol%, which is slightly lower than the ≈38 mol% reported for its *Listeria* host (Stuart and Welshimer, 1973). The sequence data presented here have been submitted to the EMBL/GenBank/DDBJ databases and appear under accession number AJ242593.
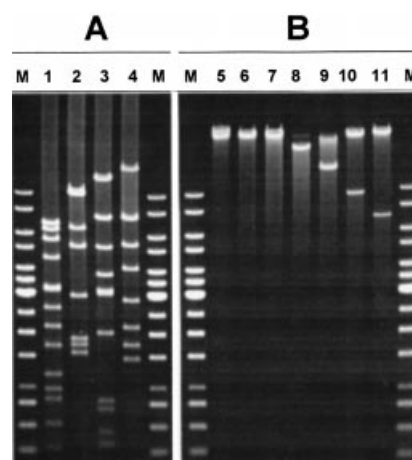
Fig. 1. A (lanes 1–4). The lack of a submolar DNA fragment after digestion of the circularly permuted, terminally redundant collection of A118 DNA molecules with selected restriction enzymes having 8–15 recognition sites (RS). Lanes: M, molecular size marker from 0.25 to 10 kpb; 1, *Eco*RI (15 RS, double band at 250 bp); 2, *Sca*I (8 RS, double band at 10 250 kbp); 3, *Pvu*II (12 RS, double bands at 3200 and 300 bp); 4, *Sfu*I (8 RS).
B (lanes 5–11). The patterns that result from digestion with enzymes having one or two RS: 5, undigested control DNA; 6, *Not*I (1 RS); 7, *Xba*I (1 RS); 8, *Xho*I (2 RS, double band); 9, *Bam*HI (2 RS); 10, *Kpn*I (2 RS); 11, *Sph*I (2 RS).

The A118 DNA is circularly permuted and terminally redundant

No evidence could be found for the presence of cohesive, protruding ends (*cos*) in the A118 DNA molecule; ligation of its DNA did not alter the restriction patterns (results not shown). Restriction endonuclease digestion of A118 DNA also failed to reveal a 'submolar' fragment (Fig. 1), which would probably contain the *pac* site for initial recognition and subsequent cutting of sequentially packaged phage concatemers by the terminase enzyme (Casjens *et al.*, 1987). Interestingly, fragments generated by enzymes that cut A118 DNA only once appear to be identical to uncut DNA. Enzymes that cut twice released one discrete fragment (we observed a little more smear in the background compared with uncut DNA, especially below the single, released bands). A second approach, which involved time-limited treatment of A118 DNA with the exonuclease *Bal*31, followed by complete digestion with restriction enzymes (Fig. 2), revealed that all fragments were simultaneously degraded, in contrast to the specific truncation of fragments observed in the control phage DNA (Fig. 2C). These results indicate that (i) there is probably no specific *pac* site involved in the recognition of viral DNA by the terminase; and (ii) there are no invariable ends in the mature A118 DNA molecules, i.e. the packaged DNA is circularly permuted.

In order to obtain further evidence for circular permutation, DNA molecules released from the phage were
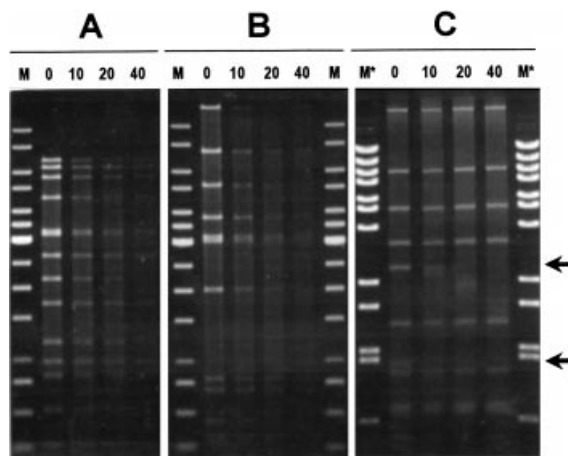
Fig. 2. Time-limited digestion with *Bal*31 exonuclease (0–40 min, as indicated) of A118 DNA (A and B) and control phage DNA (B025), followed by restriction enzyme digestion (A, *Eco*RI; B, *Pvu*II; C, *Bsp*HI). The arrows in (C) indicate the two sequentially degraded fragments of the linear control DNA. Lanes M contain a molecular size marker (0.25–10 kbp), and M* are marked lanes containing *Bst*EII-digested λ DNA (0.7–8.45 kbp).

partially denatured and inspected using electron microscopy. The denaturation pattern delineates the position of the physical DNA ends with respect to the denaturation pattern. Figure 3C shows the denaturation maps of 50 molecules that have been aligned to give a good match of the denatured sites. The average length of the linear molecules was found to be 43.3 ± 0.96 kbp. Compared with the unit genome size of 40.8 kbp, the average terminal repeat region of a packaged DNA molecule is therefore 2.5 kbp (6% redundancy). It is obvious that molecular ends are not situated at unique positions, but are partially permuted. According to histograms (not shown) of the left end positions of the molecules shown in Fig. 3C, the permutations start at 10.0 kbp and fall off rapidly to end at about 32 kbp on the 0–80 kbp scale. These data therefore indicate a permutation range of 22.0 kbp within the 43.3 kbp molecule itself. According to the frequencies of the various permutations observed in the present experiments, it would appear that encapsidation moves in a left to right direction in the maps shown in Fig. 3C, and this was confirmed directly in an experiment described below.
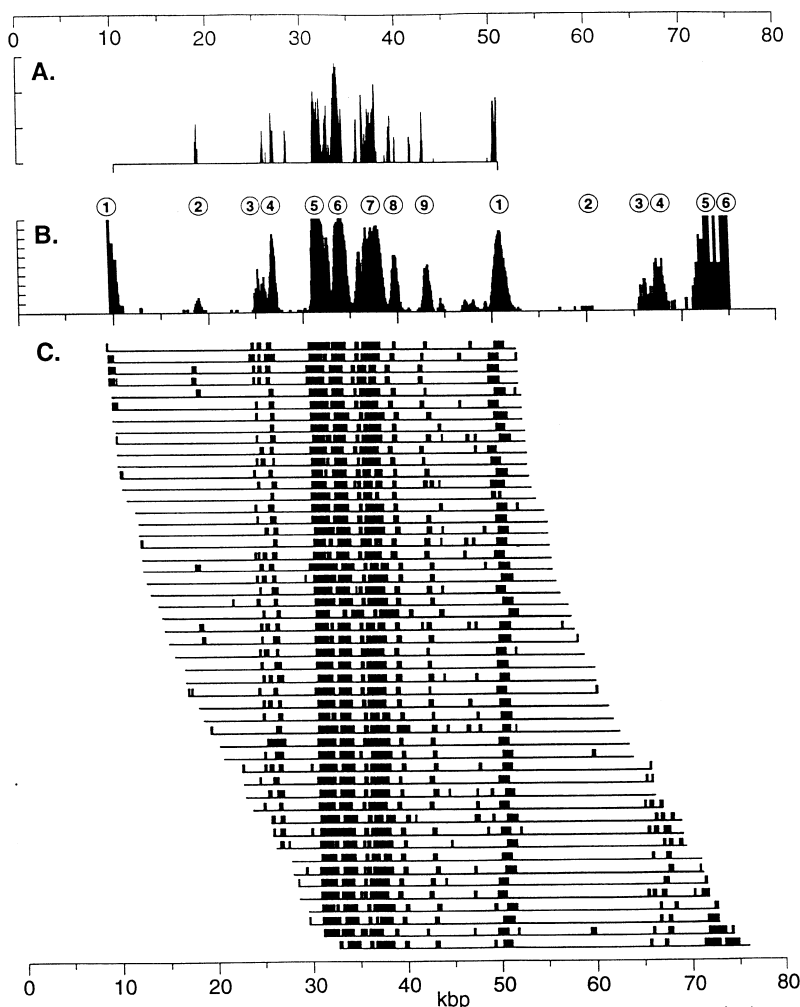

Fig. 3. This figure is described starting from the bottom panel.
C. Partial denaturation maps of 50 A118 DNA molecules. Each horizontal line represents a molecule, and black rectangles show the measured position of denatured sites arising from AT-rich regions. The molecules were manipulated (by reversal or by shifting to the left or right) such that the denatured sites could be aligned. They are shown on a 0–80 kbp scale chosen to encompass the various molecules, with an average size of 43.3 ± 0.9 kbp.
B. Histogram showing the position of AT-rich regions for the individual maps shown in (C). The vertical scale (0–100) shows the percentage of molecules with a denatured area at each of 1000 positions along the map. Peak numbers (1–9) shown above the histogram are discussed in the text.
A. Running average base composition determined from the unit sequence of A118 (40 834 bp). The average segment width was 400 bp. The running average is shown for those areas rich in AT (only areas whose average was above 67.5% AT are shown). The vertical scale runs from 67.5% to 75.0% AT. This range was chosen to yield about the same number of major peaks as in the experimental denaturation pattern shown in (B). The map was then aligned, on the horizontal axis, to give best fit with the experimental denaturation map.

The averaged denaturation map of the molecules in Fig. 3C is shown in Fig. 3B, and the major peaks have been numbered, starting from the left (1–9). It can be seen that, to the right of peak 9, the pattern begins to be repeated for peaks 1–6. In Fig. 3A, we show the running average base composition (Funnell and Inman, 1979) computed from the unit sequence. Only the AT-richest sequences are shown, so that they can be more easily compared with the experimental denaturation map (Fig. 3B). The major peaks in the predicted denaturation map agree quite well with the experimental map.

Direct evidence for terminal redundancy in permuted DNA molecules can be obtained by complete denaturation followed by renaturation (Lee *et al*., 1970). If redundancy is present, the renatured molecules will contain single-stranded regions protruding from double-stranded circles, and the circular path should represent the DNA length that does not include the redundancy. This exactly reflects our results using A118 DNA, which confirms the terminal redundancy. Measurements on 50 molecules yielded a double-stranded circular length distribution that was slightly asymmetric. The average length was found to be 39.9 ± 0.9 kbp, with the most frequent species occurring at 40.9 kbp. The linear DNA isolated from the phage was found to be 43.3 kbp, so the redundancy amounts to 2.4 kbp (5.5% of the phage DNA molecule). The length of the circular structures agrees well with the exact unit length of 40.834 bp.

DNA concatemers are sequentially packaged left to right with respect to the denaturation map and the genomic sequence

The first DNA end to enter a capsid during encapsulation can be predicted by treating intact phage with reagents that cross-link DNA and phage protein, followed by denaturation mapping. Such experiments have previously been used to determine which DNA end first enters the host during infection (Chattoraj and Inman, 1974). Because the DNA end that is first to emerge on infection should be in close proximity to (or partially inserted into) the proximal end of the phage tail, cross-linking should preserve this association. If the phage is now disrupted, this connection between a DNA end and the tail can be observed by electron microscopy. With phage A118, many examples are found with DNA ends attached to one end of phage tails; such a complex is shown in Fig. 4. Additionally, remnants of phage head protein can also be observed at the same end of the tail to which the DNA is attached. Therefore, as expected, the DNA is attached to the proximal end of the tail. Partial denaturation mapping of 47 molecules yielded maps quite similar to those presented in Fig. 3C; the average length was 43.1 ± 0.8 kbp, and the permutation range was similar. As judged by the partial denaturation
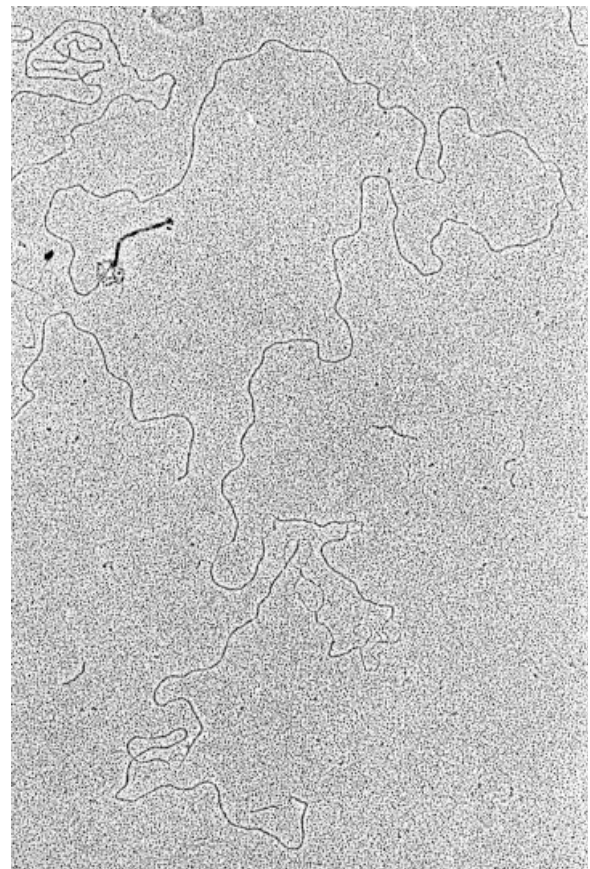

Fig. 4. Electron micrograph of A118 phage that had initially been treated to cross-link phage structural proteins and its DNA, then disrupted and the DNA partially denatured (see text). One DNA end is attached to one end of the phage tail, together with remnants of head protein (see top left of the micrograph). This DNA end is therefore attached to the proximal end of the tail, and partial denaturation mapping showed that this attachment occurs at the right ends of the DNAs.

maps, 46 out of 47 molecules had the proximal ends of phage tails attached to the right ends of the DNA (data not shown).

This result can be interpreted as follows. Within an intact phage, the DNA end first to emerge on infection is within, or in close proximity to, the proximal end of the tail and, in the above experiment, is cross-linked to it. Thus, the end to enter the phage head during encapsulation must be at the opposite DNA end. Encapsulation therefore proceeds from left to right with respect to the denaturation map or the genomic sequence.

Identification and organization of A118 genes

Bioinformatic analysis of the A118 sequence revealed the presence of 72 putative protein coding regions (Table 1), which are preceded by recognizable ribosome binding sites complementary to the 3′ end of *L. monocytogenes*

16S rRNA (Emond *et al.*, 1993). Most ORFs seem to initiate translation at an ATG codon; very few use GTG or TTG start codons. The A118 genome is apparently organized into three major gene clusters (Fig. 5): (i) the region from 1 to 21 906 (transcribed rightwards on the genetic map) probably encompasses the 'late genes', coding for structural and assembly proteins, DNA packaging proteins and the lysis proteins; (ii) the nucleotide region from 21 976 to 27 878 probably represents the 'lysogeny control' region (transcribed mostly leftwards), which encodes regulatory functions, such as the integrase and putative repressor proteins. This region also contains the *attP* site. (iii) The remainder of the A118 genome (co-ordinates 27 882–40 771) probably contains the 'early' genes, encoding products for the replication, recombination and modification of the phage DNA.

Major structural proteins Cps and Tsh

The major structural components of the A118 capsid (Cps) and tail (Tsh) have been identified previously (Zink and Loessner, 1992). In this study, microsequencing of the isolated protein bands yielded the N-terminal amino acid sequences of the mature proteins, the major capsid protein (Cps, GFNPDTTTMQSAKTGSIPIN) and the major tail protein (Tsh, RIKNAKTKY), which permitted assignment of the proteins to the corresponding ORFs 6 and 12 respectively. Apparently, in both proteins, the N-terminal methionine is removed during maturation of the primary translation products. Cps has a predicted size of 32.8 kDa, which agrees reasonably well with the approximate size of 31 kDa determined by SDS−PAGE (Zink and Loessner, 1992), whereas we noted a discrepancy between the observed size (20.5 kDa) and the predicted size (15.7 kDa) for Tsh, which is similar to the situation reported for the major tail protein gpJ of *Mycobacterium* phage L5 (25.0 versus 21.5 kDa respectively; Hatfull and Sarkis, 1993). Interestingly, the nearest relatives to Cps found in database searches (38–42% similarity in 277–321 amino acid overlaps) were the major capsid subunits of mycobacteriophages L5, D29 and TM4 (Hatfull and Sarkis, 1993; Ford *et al.*, 1998a,b).
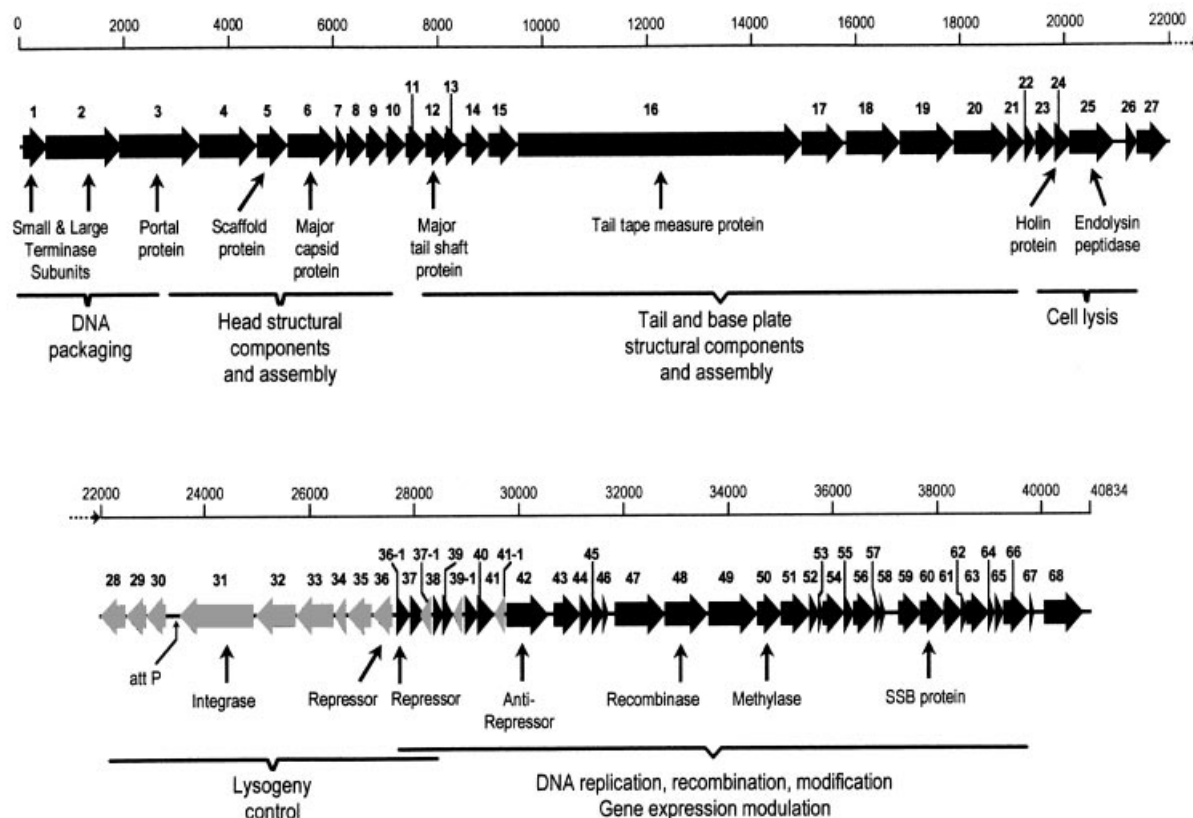
Fig. 5. Schematic representation of the A118 genome, with its assumed ORFs, some functional assignments and overall genetic organization. The ORFs are numbered consecutively (see Table 1) and are indicated by arrows or arrowheads that point in the direction of transcription. Black arrows indicate rightward transcription, and grey-shaded ORFs point leftwards. The ruler indicates their relative positions on the 40 834 bp A118 genome. The *attP* site is indicated by a bent arrow.

Table 1. Features of bacteriophage A118 ORFs, gene products (gp) and functional assignments.

ORF	Start	Stop	GP (kDa)	GP (pI)[a]	Functional assignments (homologies to other proteins are given in brackets)[b]
1	21	563	20.3	4.9	Terminase, small subunit [TerS] (φg1e, LL-H, SPP1 gp1, PBSX)
2	532	1863	51.4	6.1	Terminase, large subunit [TerL] (φg1e, SPP1 gp2, PBSX)
3[c]	1912	3366	55.3	4.6	Putative portal (φg1e gp504, LL-H gp61)
	1876		56.5		
4	3372	4511	43.1	9.6	Minor capsid (φg1e gp347, LL-H gp61)
5	4590	5159	20.9	4.6	Putative scaffold (φg1e gp204, LL-H gp 20, SPP1 gp11, *Streptococcus pyogenes* M proteins).
6	5183	6082	32.8	4.7	Major capsid protein [Cps] (TM4 gp9, L5 gp17, D29 gp17)
7	6082	6240	5.6	4.8	
8	6242	6637	14.6	5.0	
9	6637	6999	13.8	5.9	
10	6999	7337	12.8	9.8	Putative minor capsid (φg1e gp117b, LL-H gp113a)
11	7337	7744	15.1	4.1	(LL-H gp113b)
12	7747	8181	15.7	4.5	Major tail protein [Tsh]
13	8111	8443	11.5	5.5	
14[c]	8498	8920	16.5	4.8	
	8594		12.9	5.5	
15	8926	9531	23.3	4.2	Putative tail protein (φg1e gp198, LL-H gp75)
16	9542	14926	186.2	10.2	Tape measure protein [Tmp] (SPP1 gp18.1, φ105 gp37, O1205 gp40, Sfi11 gp1510, sk1 gp14, TM4 gp17, D29 gp26, L5 gp26, φCTX gp25, P2 gpT and others)
17	14928	15746	30.9	5.2	Putative tail or base plate (φ105 gp38; SPβc2 YomH)
18	15755	16780	39.4	4.8	Putative tail or base plate (φ105 gp40, φg1e gp372)
19	16781	17809	37.2	5.1	Putative tail or base plate (φ105 gp41)
20	17809	18882	39.2	4.8	Putative long tail fibre (φ105 gp42, SPβc2 Yom R, PBSX Xkdv)
21	18894	19211	12.4	4.2	Putative short tail fibre (φ105 gp43; SPβc2 YomQ, PBSX Xkdw)
22	19216	19374	6.1	4.8	
23	19402	19767	14.5	9.4	
24[c]	19771	20061	10.1	5.1	Holin [Hol]
	19780			9.8	
25	20061	20906	30.8	10.2	Endolysin [Ply], L-alanoyl-D-glutamate peptidase
26	21144	21320	6.5	10.2	
27	21307	21906	23.5	5.8	
28	22425	21976	17.4	5.6	
29	22780	22430	13.6	4.1	
30	23189	22812	13.9	7.2	
31	24876	23518	53.2	6.8	Integrase [Int] (phi-FC1 Int, TP901-1 Int, *Staphylococcus aureus* Mec, φ105 Int, TP21 gpOrf1, *E. coli* Pin, *L. monocytogenes* Tn5422 resolvase and many others)
32	25680	24940	25.7	4.4	
33	26424	25702	27.6	4.7	
34	26658	26440	7.9	6.0	
35	27172	26681	19.3	5.5	
36	27510	27205	11.8	5.1	Repressor [CI-like] (φg1e Cpg, PBSX Xre, TP-J34 gp121 and others)
36-1	27636	27878	8.7	10.3	Repressor [putative cro-like] (TP-J34 gp121, *Pseudomonas aeruginosa* pyocin and others)
37	27882	28148	9.9	4.9	70 nt complementary overlap with ORF 37-1

Similarities of the deduced A118 gene products to known sequences and functional assignments

The amino acid (aa) sequences of the products deduced from the 72 A118 open reading frames (ORFs) were screened for similarities with sequences from the available databases. The basic characteristics of the predicted gene products and the significant homologies found that permitted preliminary functional assignments are described below and are also listed in Table 1.

ORFs 1 and 2. The products of these two genes most probably represent the small and large subunits of the phage terminase respectively, which may mediate recognition of A118 DNA, ATP-dependent cleavage of the DNA concatemer and packaging of the terminally redundant molecule into the empty capsid shells. The homologies of gp1 (TerS) to the respective proteins from *Bacillus subtilis* phages SPP1 (Chai *et al*., 1992) and PBSX (McDonnell *et al*., 1994) and from *Lactobacillus* phages φg1e (Kodaira *et al*., 1997) and LL-H (Mikkonen and Alatossava, 1995) are convincing; they range from 52% to 62% similarity in overlaps from 89 to 178 aa. Moreover, A118 gp2 shows homologies to the TerL components from the same phages (except LL-H), with 49–63% similarity in overlaps ranging from 133 to 437 aa.

ORFs 3 and 4. Based on location in comparison with many other known bacteriophage genomes, it is possible

Table 1. *Continued.*

ORF	Start	Stop	GP (kDa)	GP (pI)[a]	Functional assignments (homologies to other proteins are given in brackets)[b]
37-1	28307	28077	8.6	6.5	
38	28372	28548	7.4	10.1	
39	28565	28720	5.7	7.6	
39-1	28889	28722	6.1	5.5	
40	28956	29192	8.9	4.6	
41	29189	29470	10.6	4.7	
41-1	29669	29472	7.2	8.6	
42	29733	30521	30.2	9.1	Antirepressor [Ant] (TP-J34 gp238, PVL gp34, P1 KilA, *Haemophilus influenzae* φflu and others)
43	30645	31178	19.6	6.5	
44	31175	31393	8.6	9.7	
45	31390	31578	7.3	7.2	
46	31578	31664	3.2	9.6	
47	31811	32770	36.7	5.9	Recombination (SPP1 gp34.1, skin element YqaJ)
48	32773	33588	30.6	5.2	Recombination [Rec] (*E. coli* RecT, SPP1 gp35, skin element YqaK, PVL gp43)
49	33608	34540	36.8	9.3	Replication initiation (PVL gp46, SSP1 gp38, *B. thuringiensis* plasmid gpOri60, *B. subtilis* DnaB)
50	34537	35001	18.2	9.2	Cyt-methylase (*Neisseria gonorrhoeae* Mod, φg1e gp161)
51	34998	35558	21.3	4.4	
52	35555	35701	5.5	4.2	
53	35717	35782	2.5	3.9	
54	35779	36180	15.7	4.5	
55	36177	36386	8.1	4.0	
56	36383	36781	15.1	9.5	
57	36787	36897	4.0	4.6	
58	36897	36989	3.3	7.3	
59	37225	37626	15.2	5.7	
60	37623	38105	17.8	5.1	ssDNA binding protein [SSB] (many phage and bacterial SSB proteins)
61	38137	38442	11.7	10.6	(*Listeria monocytogenes* LmaD; *Rickettsia prowazekii* gyrase subunit A)
62	38439	38579	5.6	5.0	
63	38545	38949	15.5	8.4	
64	38942	39067	4.6	10.6	
65	39078	39242	6.3	4.2	
66	39261	39695	17.1	9.0	(*Listeria monocytogenes* LmaC; T7 gp0.7 RNA polymerase kinase)
67	39741	39824	3.2	5.4	
68	40061	40771	27.4	5.3	

a. Predicted by computer analysis.
b. For references, see text.
c. Two possible translational starts.

that *orf3* encodes the A118 portal protein. It also exhibited convincing similarity (53% over 494 aa) to identified structural proteins from φg1e (gp504; Kodaira *et al.*, 1997) and LL-H (gp61; Mikkonen and Alatossava, 1994). The *orf4* product also shows significant homology to minor structural (capsid) components from the two aforementioned *Lactobacillus* phages, φg1e gp347 and LL-H gp61 (55–56% over 295–378 aa).

ORF 5. The product of *orf5* resembles several entries in the databases, i.e. minor phage structural (capsid) proteins: φg1e gp204 (58% over 188 aa; Kodaira *et al.*, 1997), LL-H gp20 (61% over 152 aa; Mikkonen and Alatossava, 1994); and SPP1 gp11 (55% over 169 aa; database accession number S58140). The last was found to be a scaffolding protein, which determines the size and shape of the viral capsid during particle morphogenesis. The putative designation of A118 gp5 as a scaffold is also

supported by its characteristic location immediately upstream of the major capsid protein gene (Hendrix and Duda, 1998). Interestingly, we also noted significant homology of gp5 to the M proteins of *Streptococcus pyogenes* (50–79% over 110–150 aa; Mouw *et al.*, 1988; Yung and Hollingshead, 1996).

ORFs 10 and 11. The deduced *orf10* product again shows convincing relatedness to products of the two *Lactobacillus* phages: φg1e gp117b (48% similarity over 108 aa; Kodaira *et al.*, 1997) and LL-H gp113a (49% over 83 aa; Mikkonen and Alatossava, 1994). Most interestingly, the location of this gene within the 'structural genes' cluster, two ORFs upstream of the main tail gene (*tsh*), seems to be conserved among A118 and the *Lactobacillus* phages. Therefore, it seems reasonable to assign to A118 gp10 a similar function to that shown for the other two phage proteins, which are minor structural components.

Moreover, the product of the downstream A118 ORF, gp11, is similar to the corresponding LL-H gp113b polypeptide (44% in 93 aa), which again points to the related organization of these phages.

ORFs 15 and 16.   The product specified by *orf15* is most probably involved in building the phage tail, based on its location and homologies to the corresponding gp198 from φg1e (50% over 201 aa) and to a product specified by LL-H *orf75*, as well as a previously undescribed coding region upstream of *orf75* and overlapping LL-H *orf125* (65% over 81 aa; Mikkonen and Alatossava, 1994), which was revealed using a T-BLAST-N database search. The predicted large protein encoded by A118 *orf16* should represent the tape measure protein (Tmp), which determines the length of the phage tail during virus morphogenesis. It shows a significant degree of similarity (more than 40% similarity in overlaps ranging from 100 to 1000 aa) to a number of other proven or suspected analogous proteins (see Table 1) from phages infecting *Bacillus* (SPP1; Alonso *et al.*, 1997; φ105, K. Kobayashi *et al.*, unpublished; database accession number AB016282), *Mycobacterium* (TM4, Ford *et al.*, 1998b; D29, Ford *et al.*, 1998a; L5, Hatfull and Sarkis, 1993), *Lactococcus* (sk1; Chandry *et al.*, 1997), *Streptococcus* (Sfi11, Lucchini *et al.*, 1998; O1205, Stanley *et al.*, 1997), *Pseudomonas* (φCTX; Nakayama *et al.*, 1999) and *E. coli* (P2; G. E. Christie, unpublished; database accession number AAD03293).

ORFs 17–21.   Based on genomic location, the products of this array of consecutive genes probably make up part of the phage tail, probably of the base plate and tail fibres. This interpretation is supported by the clear similarity of these proteins to those encoded by the corresponding genes present in several *Bacillus subtilis* phages; φ105 gp37 to gp43 (Kobayashi *et al.*, unpublished; database accession number AB016282) seem to be the homologues of A118 gp16 to gp21. This regional similarity to A118 proteins can also be found in phage SPβc2 YomI, YomH, YomR and YomQ (Lazarevic *et al.*, 1999), and in the functionally analogous gene products of the defective PBSX particle (Krogh *et al.*, 1996) (see Table 1).

ORFs 24 and 25.   These genes specify the A118 dual-lysis system, a holin (Hol) and a lysin (Ply). At the end of the phage morphogenesis, Hol is proposed to form unspecific lesions into the host cytoplasmic membrane, through which Ply can escape to the murein and specifically hydrolyse the L-alanine-D-glutamate peptide bonds (Loessner *et al.*, 1995a).

ORF 31.   The deduced protein was identified as the phage integrase, which we propose catalyses the site-specific integration and excision of the A118 prophage

genome (see below). It belongs to the invertase/resolvase family of enzymes, based on its convincing overall similarity to several other phage integrases, bacterial enzymes and transposon resolvases of this type, e.g. *Streptococcus faecalis* phage phi-FC1 integrase (57% over 476 aa; Y. W. Kim *et al.*, unpublished; database accession number AAD26564), *Lactococcus lactis* phage TP901-1 integrase (50% over 479 aa; Christiansen *et al.*, 1996), *Staphylococcus aureus* Mec protein (47% over 499 aa; Ito *et al.*, 1999), *Bacillus subtilis* phage phi-105 integrase (45% over 457 aa; Kobayashi *et al.*, unpublished; database accession number AB016282), *Bacillus cereus* phage TP21 gpOrf1 (50% over 367 aa; Loessner *et al.*, 1997), *E. coli* Pin (50% over 154 aa; Plasterk and van de Putte, 1985), *L. monocytogenes* Tn*5422* resolvase (51% over 154 aa; Lebrun *et al.*, 1994) and many others.

ORFs 36 and 36-1.   The gp36 protein shares extended similarity with repressors of the λ CI type, in particular with φg1e Cpg (57% over 93 aa; Kodaira *et al.*, 1997), the PBSX Xre protein (55% over 107 aa; McDonnell *et al.*, 1994) and with many others. This is in agreement with our finding that the A118 prophage is inducible by UV light, which should elicit a host SOS response and may eventually result in proteolytic cleavage of gp36. The second putative repressor-encoding gene (*orf36-1*) specifies a polypeptide that also shows resemblance to several phage-related transcription repressors and may represent a λ Cro analogue (see *Discussion*).

ORF 42.   Analysis of this gene reveals that the encoded protein may be an antirepressor, responsible for inactivation/bypass of the CI transcription repressor. Gp42 shows strong sequence similarity to designated antirepressors from a number of very different phages: gp238 of *Streptococcus thermophilus* phage TP-J34 (61% over 252 aa; Neve *et al.*, 1998), HI1422 of *Haemophilus influenzae* phage φflu (59% over 89 aa; Fleischmann *et al.*, 1995; Hendrix *et al.*, 1999), gp34 of *Staphylococcus aureus* phage φPVL (56% over 260 aa; Kaneko *et al.*, 1998), KilA of *E. coli* phage P1 (55% over 122 aa; Hansen, 1989) and several others.

ORFs 47–49.   The protein specified by the first of these three genes, *orf47*, is highly similar to YqaJ of the skin element, a phage-like element resident in the genome of *Bacillus subtilis* (75% over 314 aa; Takemaru *et al.*, 1995), and to gp34.1 of SPP1 (50% over 313 aa; Pedre *et al.*, 1994), which has been shown to be associated with DNA recombination. A118 gp48 shows strong relatedness to the respective products of the adjacent (downstream) genes of the two above-mentioned phages (76% over 264 aa to skin element YqaK, and 45% over 291 aa to SPP1 gp35) and also to φPVL gp43 (38% over 200 aa;

Kaneko *et al*., 1998). Its assignment as a recombinase (Rec) is supported by relatedness to *E. coli* RecT (45% over 229 aa), which was found to promote renaturation of complementary single-stranded DNA (Hall *et al*., 1993). Along these lines, A118 gp49 also seems to be involved in replication of the A118 genome. It is clearly related to φPVL gp46 (52% over 300 aa) and to SPP1 gp38 (55% over 168 aa). The latter was found specifically to recognize and bind to the *ori* sequence in the SPP1 replication region (Pedre *et al*., 1994) and, together with SPP1 gp39, forms the replisome organizer. Our functional assignment of A118 gp49 as DNA replication initiation factor is further supported by homologies to other proteins, such as the replication protein Ori60 from a large *Bacillus thuringiensis* plasmid (52% over 117 aa; Baum and Gilbert, 1991) and *B. subtilis* DnaB (50% over 94 aa; Hoshino *et al*., 1987).

ORF 50. The product deduced from *orf50* is very similar to two known proteins: an N-4 cytosine-specific methyltransferase (MTase) of *Neisseria gonorrhoeae* (77% over 150 aa; Radlinska and Piekarowicz, 1998) and gp161 of φg1e (75% over 154 aa; Kodaira *et al*., 1997). The convincing homology to the *Neisseria* enzyme justifies assignment of A118 gp50 as a modification methylase, which probably modifies the newly synthesized viral DNA at specific C-residues.

ORF 60. The deduced product of this gene strongly resembles single-stranded DNA binding (SSB) proteins, which are apparently highly conserved among phages and their bacterial hosts: more than 50 proteins with a high degree of sequence similarity (*P*-values in BLAST searches from 10^{-56} to 10^{-5}) were found in the homology searches. Moreover, A118 SSB exhibits around 90% homology to the corresponding proteins from *Listeria innocua* phage B056 and *Brochothrix thermosphacta*

phage A19 (M. J. Loessner and S. Scherer, unpublished). The proposed function of gp60 is that it binds to and stabilizes single-stranded DNA intermediates during genome replication and/or recombination (Pedre *et al*., 1994).

ORFs 61 and 66. Figure 6 illustrates a particularly interesting set of homologies: A118 gp61 is significantly similar to the LmaD antigen of *Listeria monocytogenes* (74% over 101 aa), and A118 gp66 revealed relatedness to the LmaC protein (75% in 144 aa). These two proteins are encoded by directly adjacent genes of the *lmaDCBA* operon (Schäferkordt and Chakraborty, 1997). Their function is not understood, but the operon is apparently restricted to pathogenic *Listeria* species. According to further sequence homologies found, these two phage proteins might be involved in DNA replication and gene expression modulation: gp61 is similar over its entire length to a DNA gyrase subunit A from *Rickettsia prowazekii* (51% over 101 aa; Wood and Waite, 1994), and gp66 shows some similarity to a kinase from *E. coli* phage T7, which phosphorylates the host RNA polymerase during early gene expression (44% over 119 aa; Dunn and Studier, 1983).

Putative promoters and terminators

The three major genetic functional regions (gene clusters) are apparently separated by intergenic regions containing promoter sequences and/or regions capable of forming fairly stable stem−loop structures. The latter probably represent transcription terminators and are mostly located at the end of operon-like clusters: (i) downstream of *orf25* (*ply*) at position 20 913−20 932 ($\Delta G = -13.2\,\text{kcal mol}^{-1}$); (ii) in the intergenic region between the two facing genes *orf27* and *orf28*, at position 21 910−21 928 ($\Delta G = -12.7\,\text{kcal mol}^{-1}$) and position 21 970−21 950 ($\Delta G =$

A

```
Gp61  (1) MSRKELRKKQWEVITMIEKSKTLTDRKNLIKKLETLEARGDKEKGLATPTQLLSIFTVTEYRRLSKKLTDTEIAEDMGISRSALIEFKRKNGLSIRQKVAT (101)
          :|||  |::|| ::|  |::::  ||:||::||  :|| ||:|||| |::| | |:| :||| ::: :|::|: :|:|||:| |:|||:||: ::
LmaD  (1) MDRKLLKEKQIQLIFQLEQEENRFIRKRLIEELEFFEALGDREKGLLTAEQKLLILTPSEYREYKRTKSDVQISRIIGVSRSSLAEWKRKKGLNRKKSQPV (101)
```

B

```
Gp66  (1) MGQLFNLPQVEDINYIQTVRAVRQFFKDYLTLRLMAGDRKFPNMTTMYKITPPNFSNEFHSKVEDAAIHNVDNVHAAQEAVKKYDAIMNQLEHIHRKILFE (101)
          |:: ||: :||||:||::|::||| |: :|::| :||   || :|:: :||:::||::::::|:||:|
LmaC  (1) MQVLVLPENKDINYIKTVQEVKRFFADFERFRMITGLSKKPHLLRNGFLEEP----QFEP-VAFSARHNKEVILEARWLVEKYTEMLNQMDDLYRTILME  (95)
```

```
Gp66 (102) KFIHNLQDRTIMLDIPYEERQYKREKRKAVIELATTLGIEVLN (144)
           :::: || ::|:|:||| |:|| |:||:||||| :|||  |
LmaC  (96) CYVERKQDVAVMMDLPYEIAQFKRIKKRAVLELATLMGILVRK (138)
```

Fig. 6. Alignment of the primary amino acid sequences of the deduced products of A118 *orf61* (gp61) (A) and *orf66* (gp66) (B) with *Listeria monocytogenes* LmaD and LmaC proteins of the *lmaDCBA* operon (Schäferkordt and Chakraborty, 1997) respectively. Identical residues (|) and conservative replacements (:) are indicated; a few gaps have been introduced to optimize alignment.

Promoter	Position	-35 region	-10 region	RBS	Start
P_R 26	21062	ATTTGAATTTAAAAAAAATG TTTACT TTTAAGCTAAATGTG	TATAGT ATATATTGTAAGGACTTAAAACTT	GGAGGG ATGAAA	**ATG**
P_L 30	23273	TTAATTTATAAAAACAC TTGACT ACTACGATTATTCGTAG	TATAGT GTAAGTATAGTAAAGATAACGAAACG	GAGG AATTTAAA	**ATG**
P_L 36	27581	ACTTTTTGTGATTT TTGTTT ATTTTTTTTTAAAATATCACT	TATAGT GATACTTAAGA	AGGAGGG AGATATT	**ATG**
P_R 36-1	27567	AAAATCACAAAAAGT GA TTGACA ATCACGTAAAGTGATAG	TATTAT TATCACATA	GAGTG ATAAAGGCG	**GTG**
P_L 37-1	28374	**CAT**GGTTTAGCCTCCTG-TT TTGGTT ACTCTCCAATCTGC	TATAAT TAGTTTGATT	GGAGGT GATAAT-	**ATG**
		‖‖‖ ‖ ‖‖‖‖‖‖‖‖‖ ‖‖ ‖‖ ‖‖			
P_L 39-1	28958	**CAT**TTTCTAGCCTCCTATTT TTGGTT ACTCTCCAATCTGC	TATAAT TAGTTTGATT	GGAGGT GATTATT	**TTG**
P_L 41-1	29938	TGA**CAT**TTGTTCTCCTT TCTGTT CGCCCTTCCACAGTGC	TATAGT TTTTATGAAG	GGAGGT GGGTAAA	**ATG**
P_R 47	31730	GTGCGTACTTAAGGGAT TTGAGA TATTAACTTAAGGAAAT	TATACC TCAGGTCCATTAAAAAAATCAAT	GGAGG TAACATAT	**ATG**
P_R 68	39992	TTTCCATCTGTTGAAAA TTGAGC AGCTGGTTTTTATTTGG	TATAGT GAAAAATAAAAGGGT	GGA TTATG	**ATG**

Fig. 7. The putative promoter sequences found in the A118 genome are shown, and their location upstream of the respective open reading frames is indicated by promoter number and nucleotide position. The letters R or L indicate the direction of transcription (right or left) on the genetic map (see Fig. 5). Putative −35 and −10 RNA polymerase recognition sequences are boxed; probable ribosome binding sites (RBS) are underlined; start codons are shown in bold letters; the overlapping complementary sequences of the adjacent promoters in between the two putative repressor genes (36 and 36-1) are shaded; and the vertical lines (‖) indicate the sequence repeats containing the two leftward promoters 37-1 and 39-1 (see text).

−5.8 kcal mol^{-1}); (iii) downstream of *orf31* (*int*) at position 23 445–23 395 ($\Delta G = -12.7$ kcal mol^{-1}); (iv) downstream of *orf68*, at position 40 799–40 816 ($\Delta G = -12.6$ kcal mol^{-1}). The putative promoter elements that could be identified upstream of (and in between) the transcriptional units are listed in Fig. 7. An unusual finding was the arrangement of the two directly adjacent, outward-facing, complementary overlapping promoters serving *orf36* and *orf36-1*, the two repressor-encoding genes. Also of note were the two 47 bp perfect sequence repeats with internal dyad symmetry, which apparently encompass conserved promoters and ribosome binding sites. The first repetition (P_L 37-1) is located at 28 311–28 357 and the second (P_L 39-1) at 28 894–28 940. This is reminiscent of the situation in *Lactococcus lactis* phage r1t, in which similar sequence repeats were found to act as operators regulating gene expression in the lysogeny control region (Nauta *et al.*, 1996). Also of interest is the putative leftward promoter P_L 30: after integration of A118 prophage (*attP* at 23 470, see below), this becomes an inward-facing promoter from the end of the phage genome. However, we do not know whether it is active in the prophage state or not, and what the function of the potential products from ORFs 28–30 might be.

Identification of attP *and* attB *revealed that A118 integrates into a* comK *gene homologue*

We began studying the integration systems of *Listeria* phages in U153, a close relative of A118. By comparing the restriction fragments of phage DNA with restriction fragments from lysogenic strains, we identified a phage fragment containing *attP*, as well as the corresponding prophage fragments containing the junctions of host and phage attachment sites. We cloned and sequenced the phage fragment and found an integrase gene, closely linked to *attP*. In order to determine the exact sequence in A118, we used polymerase chain reaction (PCR) technology. Two pairs of primers were used: the first (primers 1 and 2) crosses *attP* and only amplifies a product on purified phage DNA. The second (3 and 4) crosses *attB* and only amplifies a product on bacterial strains that do not have a prophage integrated. Using one primer specific for *attP* in conjunction with one specific for *attB* (primers 1 and 4), a 'hybrid' product was obtained, but only on lysogenized bacterial strains (see Fig. 8). This genomic organization was confirmed by the appropriate banding patterns in Southern blot experiments (results not shown). *attP* has the unusual property of a core sequence with only 3 bp of recognizable conservation between the phage and bacterial genomes (GGA) at the site of crossing over (see Fig. 8). As determined by nucleotide sequencing of the junction fragment PCR products and comparing them with *attP* and *attB* sequences, *attP* lies 48 bp downstream from *orf31*, which specifies the A118 integrase. The *attB* site was found to be located within an *L. monocytogenes* flanking sequence containing an ORF whose deduced product is highly similar (37% identity, 59% similarity) to the *Bacillus subtilis comK* gene product, a key regulatory transcription factor for competence development (Van Sinderen
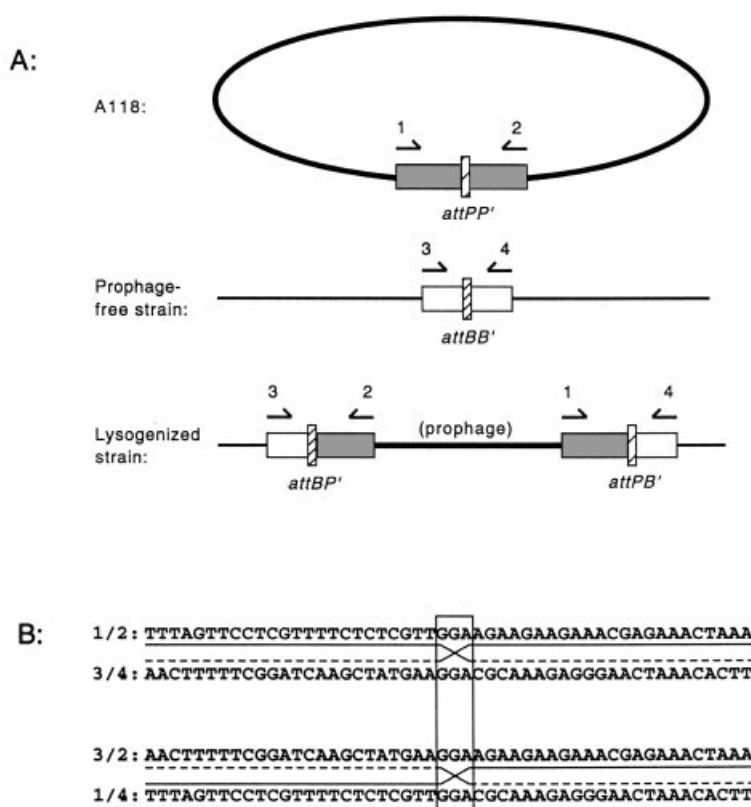
Fig. 8. Schematic diagram of A118 integration.
A. Location of the PCR primers used to identify and sequence the attachment sites are shown as arrows. The 861 bp *attP*-specific PCR product is shown as a grey box (top). The 417 bp *attB*-specific PCR product is fully contained within the *Listeria monocytogenes comK* gene and is shown as white box (middle). After integration, junction fragments are formed, and the hybrid PCR products used to diagnose these are shown as split grey and white boxes (bottom). Hatched boxes represent the core sequence of the attachment site (drawing is not to scale). A118 prophage is shown as a thick line in lysogenized strain *L. monocytogenes* WSLC 1118.
B. Partial sequence (50 bp) of junction fragments, including parts of the phage and bacterial attachment sites, which was obtained using PCR primers noted at the left. Integration of phage occurs within a 3 bp 'core' sequence (GGA), which is boxed. No recognizable homology could be detected on either side of this core sequence. The top two sequences are non-integrated phage (solid line) and bacterial sequences (dashed line); the bottom two are junction fragment sequences.

et al., 1995). The *attB* position is 187–189 in the 573 bp *comK* ORF, whose putative function is most probably inactivated upon integration of A118 prophage. As expected, lysogenized host cells (EGDe::A118) became resistant to superinfection with A118. These sequence data (*attB* flanking sequence) have been submitted to the EMBL/ GenBank/DDBJ databases under accession number AF174588.

Discussion

A118 is the first phage infecting *Listeria monocytogenes* for which the complete nucleotide sequence has been determined and whose genome has been analysed in detail. The A118 DNA represents a terminally redundant and circularly permuted collection of molecules without cohesive ends. Therefore, both arms must recombine by homologous recombination within the terminally redundant ends before the circularized molecule can initiate replication or integrate into the host chromosome. As a result of recombination, all the formerly permuted molecules are transformed into identical unit-length molecules, which is absolutely critical for maintaining genomic integrity. The genetic map (see Fig. 5) may therefore be drawn as a circle, which would properly reflect the mixture of permuted molecules contained in an A118 phage population.

The terminase enzyme complex, responsible for DNA packaging, cutting and condensation into the viral capsid, often recognizes the DNA concatemer at a specific sequence designated as *pac*. However, using techniques that revealed such a *pac* site in other phages could not reveal it in A118. A possible explanation is that the DNAs are randomly permuted, and A118 uses a sequence-independent headful measuring cut depending solely on physical parameters, similar to the mechanism proposed for phages T4 or φ11 (Streisinger *et al*., 1967; Löfdahl *et al*., 1981). However, we cannot exclude the presence of multiple *pac* sequences, or extended migration of the terminase along the concatemer for cutting (Black, 1988), which could have prevented its detection. Nevertheless, the lack of *pac* may help to explain the high frequency of general transducing particles found in A118 (Hodgson, 2000), and also agrees with our results of partial denaturation mapping, which clearly showed a high degree of permutation among the individual A118 DNAs. However, permutation was not complete, as has been observed earlier with a number of other phages, such as P22 (Tye *et al*., 1974), and almost certainly arises from pure headful packaging of concatemeric DNA generated in the replication process (Streisinger *et al*., 1967). Consequently, and regardless of a *pac* site, the degree of permutation would then be dictated, in the simplest scenario, by the concatemeric length.

The results of our cross-linking partial denaturation experiment show clearly that encapsidation of the genome proceeds from left to right with respect to the genomic sequence, and it is the right end of the DNA molecule that first enters the cell upon infection. It is interesting to note that this direction is the same as the orientation of ORFs required for lytic development of A118 (Fig. 5) and that, in a majority of the A118 DNAs, the genetic region containing the early recombination genes is present on that portion of the (permuted) molecules that first enters the cells upon DNA injection (Fig. 3).

Out of 72 proteins deduced from A118 ORFs, 26 (36%) were found to have significant similarities to existing entries in the databases, and enabled assignment of a (putative) function, on the basis of homologies to proteins whose function is known or by direct experimental evidence. The remaining 46 deduced gene products more or less resembled new entries to which no convincing sequence homologies were detected.

An interesting observation is the homology of A118 gp5 to the M-proteins of *Streptococcus pyogenes* (Mouw *et al*., 1988; Yung and Hollingshead, 1996), which are IgG-interacting, fibrous surface proteins and important virulence determinants. This similarity may suggest that gp5 has some protein–protein interaction ability, which seems reasonable for a scaffolding protein that must manage to bind to and interact with the (major) capsid protein. This may also suggest that the scaffold is a fibrous array; the Sid external scaffold of phage P4 proheads does apparently have a fibrous appearance (Marvik *et al*., 1995).

The designated tape measure protein of the A118 tail (gp16, 1794 aa) is about twice the size of the respective protein from phage λ (gpH, 853 aa), and the A118 tail (300 nm; Zink and Loessner, 1992) is twice as long as the λ tail (150 nm). This observation is in line with the finding that the length of the λ tail shaft structure is directly dependent on the molecular size of gpH (Katsura, 1987), a protein ruler whose functional analogues appear to be widely distributed among the *Siphoviridae* viruses, independent of host taxonomic classification.

The order and organization of the two genes *orf36* and *orf36-1* (immediate neighbours, but opposite transcription direction with an overlapping promoter region) is similar to the situation in *Lactobacillus* phage φg1e (Kodaira *et al*., 1997), in which the CI analogue Cpg and the Cro analogue Cng are also oriented this way, and to the respective gene organization in *Streptococcus thermophilus* phages TP-J34 and Sfi21 (Bruttin *et al*., 1997; Neve *et al*., 1998). Additional evidence that supports our functional assignments is derived from using non-sequence alignment-based homology parameters, such as size and charge of the proteins (Chandry *et al*., 1997). Gp36 and gp36-1 have biochemical properties that are quite similar to the λ proteins, i.e. λ CI is acidic (predicted pI 4.9) similar to

gp36 (pI 5.1), and λ Cro is very basic (pI 10.2) as is gp36-1 (pI 10.3). The relative sizes of these two proteins are also analogous. These data taken together, it is tempting to speculate that gp36 and gp36-1 govern the genetic switch of A118.

The A118 genes clearly display a life cycle-specific organization. The opposite orientation and arrangement of gene clusters required for lytic growth compared with the designated lysogeny control region has also been found for several other siphoviral genomes analysed to date. Especially intriguing is the A118 putative genetic switch region, which apparently contains several small transcriptional units oriented in both directions, possibly reflecting the competing transcription events occurring during the early decision for lysis or lysogeny.

Bioinformatic analysis revealed that defined portions of the A118 genome resemble specific 'functional' regions of other bacteriophage genomes, in particular of phages infecting *Lactobacillus*, *Streptococcus thermophilus* and *Bacillus* hosts. These striking parallels in genome structure are similar to the situation within the lambdoid phages, which show conserved location of functional gene clusters but not necessarily extended sequence homology (Casjens *et al*., 1992). The data presented here support the proposed model that phage genomes are built essentially as mosaics from essential components, designated as genetic modules, or, in a more practical sense, as functional segments of varying size, ranging from whole genomic segments to single genes or even gene domains (Botstein, 1980; Highton *et al*., 1990). These are thought to be available (although with restrictions) from a large gene pool, which is accessible to the individual phage population through a variety of mechanisms, such as interchange of modules by homologous recombination and horizontal gene transfer (Botstein, 1980; Haggård-Ljungquist *et al*., 1992; Hendrix *et al*., 1999). Horizontal exchange in phages is entirely dependent on the genetic material carried in their individual hosts. However, the hosts of phages sharing homology to A118 in certain proteins (e.g. gp42 and the antirepressors from *Streptococcus thermophilus* φTP-J34, *Haemophilus influenzae* φflu, *Staphylococcus aureus* φPVL and *E. coli* φP1) obviously occupy different ecological niches, which should decrease the frequency of direct contact and genetic exchange. It is also evident from our data that most of the stronger homologies of A118 proteins are to phage proteins of hosts that are phylogenetically related to *Listeria*, i.e. *Lactobacillus*, *Bacillus*, *Staphylococcus*, *Lactococcus* and *Streptococcus*, which are all members of the low G + C sub-branch of Gram-positive eubacteria. This may be interpreted as evidence that the individual phages could have evolved from some ancestral phage able to infect the respective ancestral host, which would imply divergent evolution through conserved vertical passage.

One of the primary practical applications derived from

analysis of temperate phage regions controlling phage lysogeny is the construction of integration vectors for the host. The identification of the integrase gene (*orf31*) and adjacent *attP* site from A118 and a related phage (U153) enabled the design of a single-copy genomic integration vector for *L. monocytogenes*, which should be useful for specific gene expression studies (P. Lauer, R. Calendar and Daniel Portnoy, unpublished data). The small common core sequence of only 3 bp necessary for integrational recombination of A118 is unusual, and reminiscent of the situation in actinophage φC31, which also integrates via a 3 bp core sequence (Rausch and Lehmann, 1991). The overall relatedness of A118 gp31 to members of the large molecular mass subgroup of resolvase/invertase enzymes (e.g. φC31 Int; Thorpe and Smith, 1998) suggests that gp31 may have similar properties: (i) recognition of non-identical recombination sites with no requirement for topologically defined structures; and (ii) the small conserved 'core' sequence is critical for staggered strand breaks and subsequent rejoining (Thorpe and Smith, 1998).

It is important to note that integration into the coding region of the *comK* homologue is likely to knock out the corresponding gene product. However, the exact function of ComK in *L. monocytogenes* is currently unknown, and natural competence has not yet been demonstrated in *Listeria* species. Interestingly, we have noted that many of the commonly used laboratory strains (e.g. 10403S, EGDe, LO28 and others) carry an integrated functional or cryptic prophage at the *attB* within *comK* (unpublished data). Lysogenization with A118 leads to resistance against superinfection by A118 or related phages, but not to other phages of different immunity groups (Loessner *et al.*, 1991). The *comK* gene is apparently only one of several existing *attB* sites in the *Listeria* genome, as multiple lysogens can be created by subsequent challenge with different phages (Loessner *et al.*, 1991), and polylysogenic *L. monocytogenes* strains are also frequently naturally occurring. In conclusion, studies are needed to explore further the potential biological influence of temperate phage on the behaviour, ecology and virulence properties of this important pathogen, and are in progress in our laboratories.

Experimental procedures

Bacterial strains, media and culture conditions

Listeria monocytogenes strain EGDe (serovar 1/2a) was used for propagation of A118, and was grown in tryptose media at 30°C. *E. coli* strain DH5αMCR (Life Technologies) was used in recombinant DNA work, and cultures were passaged in Luria–Bertani (LB) medium at 37°C. For the selection of plasmid-bearing cells, ampicillin was added at 100 μg ml^{-1}.

A118 propagation and DNA purification

One plaque was picked from a soft agar overlayer plate of

A118 grown on EGDe and mixed with three A_{600} units of bacteria grown to stationary phase at 30°C in LB broth supplemented with 50 mM MOPS buffer, pH 7.3, and 0.4% glucose (LMG). The mixture was supplemented with 10 mM $CaCl_2$ and 10 mM $MgCl_2$ and incubated for 15 min at 30°C to promote adsorption. The culture was then added to 400 ml of LMG in a 2800 ml Fernbach flask and incubated further with shaking at 30°C. The A_{600} first rose to 0.6 and subsequently fell to 0.15 within ≈3 h. Cell debris was removed by centrifugation (6000 *g* for 10 min). Phage particles were precipitated in the presence of 10% (w/v) polyethylene glycol 8000 (Sambrook *et al.*, 1989), collected by centrifugation as described above and resuspended in 5 ml of phage buffer (1% ammonium acetate, 0.01 M $MgCl_2$, 50 mM Tris-HCl, pH 7.5). The suspension was extracted with chloroform (2 ml) and recentrifuged (10 000 *g* for 10 min). The titre in the supernatant was about 5×10^{11} plaque-forming units ml^{-1}. For electron microscopy, phage particles were purified by adding CsCl (2.2 g), followed by centrifugation in an SW50.1 rotor (40 000 r.p.m. for 40 h). The phage band was removed and dialysed against phage buffer. Recovery of plaque-forming units was 40%. In order to prepare RNA-free DNA for sequence analysis, the phage preparation (polyethylene glycol pellet, resuspended) was purified by passage over a 22 ml column of DEAE cellulose (Whatman DE52), equilibrated and washed with 50 mM Tris-HCl pH 7.5, 0.4 M NaCl, 10 mM $MgCl_2$. Fractions (2 ml) were collected, and about half of the applied phage titre was recovered in fractions 7–9. DNA was prepared by extracting the suspension once with phenol, twice with phenol/chloroform and once with chloroform. As this preparation was quite viscous (probably carbohydrate contamination), it was precipitated with ethanol and resuspended in 0.5 ml of 10 mg ml^{-1} egg white lysozyme, 10 mM EDTA, 25 mM Tris-HCl pH 8.0. After incubation for 30 min at 37°C, 60 μl of 10% SDS and 10 μl of proteinase K (20 mg ml^{-1}) were added, and the mixture was incubated for 1 h at 65°C. Then, 100 μl of 5 M NaCl and 80 μl of 10% CTAB in 0.7 M NaCl were added, and the solution was incubated for 10 min at 65°C. The mixture was extracted with chloroform/isoamyl alcohol (25:1, v/v), phenol/chloroform/isoamyl alcohol (50:48:2, v/v) and chloroform. Finally, the DNA was ethanol precipitated and redissolved in TE buffer at 0.5 μg μl^{-1}.

Cloning, nucleotide sequencing and computer analysis

For construction of the phage DNA library, A118 DNA was partially digested with *Tsp*509I (New England Biolabs). Fragments of the desired size (1–2 kbp) were recovered from agarose gels and ligated (T4 DNA ligase; Roche Molecular) into pBluescript-II SK– (Stratagene), which had been linearized with *Eco*RI (Roche Molecular) and treated with shrimp alkaline phosphatase (US Biochemical). Ligation products were electroporated into *E. coli*, followed by selection of insert-containing transformants on Xgal-containing agar plates. Plasmids were recovered from small-scale cultures (MiniPrep kit; Qiagen), and the individual plasmid inserts were compared after release from the vector by digestion with *Pau*I (MBI). A total of 35 plasmids was selected, and the nucleotide sequences of phage DNA inserts were determined using initial primers (SK and KS) complementary to sequences flanking

the multiple cloning site on pBluescript II, followed by sequential primer walking with synthetic oligonucleotides as new sequences became available. The gaps between the individual contigs were closed using A118 DNA directly as the template. The A118 sequence was completed by upstream and downstream sequencing the single last contig, until an overlap of the left and right ends of the DNA molecule was encountered. All sequencing was done using dye terminators on a ABI 373A automated DNA sequencer (Applied Biosystems).

The program DNASIS version 2.10 (Hitachi) was used for contig assembly and primary analysis of nucleotide and amino acid sequences. Homology searches were performed using the algorithms of the BLAST version 2.0.8 program (BLASTP and TBLASTN), available through the worldwide web (http://www.ncbi.nlm.nih.gov/cgi-bin/blast/nph-newblast) (Altschul *et al.*, 1997).

Determination of A118 genome physical structure

To determine the potential presence of a *pac* site within the A118 DNA molecule and the extent of permutation of the individual molecules, A118 DNA was digested with selected restriction endonucleases with either 8–15 predicted recognition sites (*Eco*RI, *Sca*I, *Pvu*II, *Sfu*I) or 1–2 predicted recognition sites (*Not*I, *Xba*I, *Xho*I, *Bam*HI, *Kpn*I, *Sph*I; purchased from Roche Molecular, New England Biolabs, MBI Fermentas, Amersham-Pharmacia).

In order to test whether the A118 genome has cohesive ends (*cos*), small amounts of purified DNA (0.5 μg) were incubated with T4 DNA ligase, followed by digestion with *Eco*RI or *Pvu*II, then heated (60°C for 10 min), and the fragments were separated by agarose gel electrophoresis. A non-ligated DNA sample served as a negative control.

To determine whether the A118 DNA molecule has invariable (specific) sequences at its two termini, phage DNA (5 μg) was incubated with the double-stranded DNA exonuclease *Bal*31 (20 units) for increasing time points (0–40 min), according to the instructions of the manufacturer (MBI Fermentas). The DNA was then ethanol precipitated, redissolved in TE buffer, digested with *Eco*RI or *Pvu*II and analysed by agarose gel electrophoresis. As a positive control, we used the DNA of *Listeria* bacteriophage B025 (Loessner *et al.*, 1994), which was previously known to have specific ends that can be progressively degraded (unpublished data).

Amino-terminal sequencing of A118 major structural proteins

This was performed as has been described earlier (Loessner *et al.*, 1994). Briefly, virion structural proteins were separated by SDS–PAGE, electroblotted onto a polyvinylidene difluoride (PVDF) membrane and stained with Coomassie blue. The two major protein bands representing Cps and Tsh (Zink and Loessner, 1992) were then excised from the membrane, and the sequence of their N-terminal amino acids was determined using an Applied Biosystems 477A automated protein sequencer.

Partial denaturation mapping of A118 DNA

DNA molecules were partially denatured by heating in high pH buffer and formamide and prepared for electron microscopy using a modified cytochrome *c* spreading technique, as has been described previously (Inman and Schnös, 1979). For A118 DNA, an adequate number of denatured sites occur after a 10 min incubation at 48°C in a solution consisting of 17.2 mM Na$_2$CO$_3$, 5.6 mM EDTA, 11.6 mM NaCl, 26.9 mM KOH, 8.0% formaldehyde and 15.8% formamide. Electron micrographs of partially denatured molecules together with a standard (M13mp18 DNA) were measured and processed as has been described earlier (Littlewood and Inman, 1982). The average length of the partially denatured molecules was 43.3 ± 0.9 kb, and all molecules were individually normalized to this length before alignment and production of the histogram average. The calculation of the running average AT content was determined with an average segment width of 400 bp and followed the procedure already described (Funnell and Inman, 1979).

In order to detect terminal redundancy by denaturation–renaturation experiments, A118 DNA was directly denatured in a solution containing 22 mM EDTA, 15 mM NaCl and 111 mM NaOH for 15 min at 20°C. The solution was then adjusted to 70 mM Tris-HCl (pH 7.0) and 35% formamide and, after 2 h at 20°C, dialysed against 20 mM NaCl, 4 mM EDTA before spreading for electron microscopy by the modified cytochrome *c* technique described above.

For the determination of the direction of packaging, intact phage particles were first cross-linked by incubation for 30 min at 37°C in a solution containing 2% glutaraldehyde, 20 mM NaCl and 5 mM EDTA. Samples were then spread by the cytochrome *c* method under the partial denaturation conditions described above.

Identification and nucleotide sequencing of attP and attB

Lysogenization of EGDe was carried out essentially as has been described previously (Loessner *et al.*, 1991). Four PCR primers were designed to be specific for phage sequences only, non-lysogenic host bacteria only or lysogenized bacterial strains and were used in various combinations. A118 PCR primers that cross *attP* were designed based on previous work with the related *Listeria* phage U153 (P. Lauer and R. Calendar, unpublished data). Primers 1 (5′-CTCATGAACTC-GAAAAATGCGG-3′) and 2 (5′-GTCTGTGTAACTTACCCA-TTCG-3′) specifically amplified *attP* from purified A118 DNA and yielded an 861 bp product. Primers 3 (5′-TGAAGT-AAACCCGCACACGATG-3′) and 4 (5′-TGTAACATGGA-GGTTCTGGCAATC-3′) yielded a 417 bp product from *L. monocytogenes* without a prophage inserted at *attB*. When PCR was performed on A118 lysogens, primers 1 and 4 yielded a 744 bp product and primers 2 and 3 yielded a 534 bp product. Standard conditions for all PCR reactions were as follows: 30 cycles of PCR were performed, using an annealing temperature of 55°C and ≈100 ng of template DNA. *AttB*, *attP* and junction fragment PCR products were gel purified (QIAquick gel extraction kit; Qiagen) according to the manufacturer's directions and sequenced directly. Southern blots were performed using standard techniques, using restriction enzyme-digested DNA from A118 phage and several strains of *L. monocytogenes* lysogenic for A118, including WSLC 1118, 10403::A118 and EGDe::A118, as well as phage-cured controls (10403). As probes, digoxigenin-labelled DNAs

from A118 and *attP* (PCR product using primers 1 and 2) were used, followed by standard chemiluminescent detection of the hybrids (data not shown).

References

Alonso, J.C., Luder, G., Stiege, A.C., Chai, S.H., Weise, F., and Trautner, T.A. (1997) The complete nucleotide sequence and functional organization of *Bacillus subtilis* bacteriophage SPP1. *Gene* **204:** 201–212.

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

Baum, J.A., and Gilbert, M.P. (1991) Characterization and comparative sequence analysis of replication origins from three large *Bacillus thuringiensis* plasmids. *J Bacteriol* **173:** 5280–5289.

Black, L.W. (1988) DNA packaging in dsDNA bacteriophages. In *The Bacteriophages*. Calendar, R. (ed.). New York: Plenum Press, pp. 321–373.

Botstein, D. (1980) A theory of modular evolution for bacteriophages. *Ann NY Acad Sci* **354:** 484–491.

Bruttin, A., Desiere, F., Lucchini, S., Foley, S., and Brüssow, H. (1997) Characterization of the lysogeny DNA module from the temperate *Streptococcus thermophilus* bacteriophage φSfi21. *Virology* **233:** 136–148.

Casjens, S., Huang, W.M., Hayden, M., and Parr, R. (1987) Initiation of bacteriophage P22 DNA packaging series. Analysis of a mutant that alters the DNA target specificity of the packaging apparatus. *Mol Biol* **194:** 411–422.

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

CDC (1998) Multistate outbreak of listeriosis – United States. *Morbid Mortal Weekly Rep* **47:** 1085–1086.

Chai, S., Bravo, A., Luder, G., Nedlin, A., Trautner, T.A., and Alonso, J.C. (1992) Molecular analysis of the *Bacillus subtilis* bacteriophage SPP1 region encompassing genes 1–6. The products of gene 1 and gene 2 are required for *pac* cleavage. *J Mol Biol* **224:** 87–102.

Chandry, P.S., Moore, S.C., Boyce, J.D., Davidson, B.E., and Hillier, A.J. (1997) Analysis of the DNA sequence, gene expression, origin of replication and modular structure of the *Lactococcus lactis* lytic bacteriophage sk1. *Mol Microbiol* **26:** 49–64.

Chattoraj, D., and Inman, R.B. (1974) Location of DNA ends in P2, 186, P4 and lambda bacteriophage heads. *J Mol Biol* **87:** 11–22.

Christiansen, B., Brondsted, 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.

Dietrich, A., Bubert, A., Gentschev, I., Sokolovic, Z., Simm, A., Catic, A., *et al.* (1998) Delivery of antigen-encoding plasmid DNA into the cytosol of macrophages by attenuated suicide *Listeria monocytogenes*. *Nature Biotechnol* **16:** 181–185.

Dunn, J.J., and Studier, F.W. (1983) Complete nucleotide sequence of bacteriophage T7 DNA and the location of T7 genetic elements. *J Mol Biol* **166:** 477–535.

Emond, E., Fliss, I., and Pandian, S. (1993) A ribosomal DNA fragment of *Listeria monocytogenes* and its use as a genus-specific probe in an aqueous-phase hybridization assay. *Appl Environ Microbiol* **59:** 2690–2697.

Fleischmann, R.D., Adams, M.D., White, O., Clayton, R.A., Kirkness, E.F., Kerlavage, A.R., *et al.* (1995) Whole-genome random sequencing and assembly of *Haemophilus influenzae* Rd. *Science* **269:** 496–512.

Ford, M.E., Sarkis, G.J., Belanger, A.E., Hendrix, R.W., and Hatfull, G.F. (1998a) Genome structure of mycobacteriophage D29: implications for phage evolution. *J Mol Biol* **279:** 143–164.

Ford, M.E., Stenstrom, C., Hendrix, R.W., and Hatfull, G.F. (1998b) Mycobacteriophage TM4: genome structure and gene expression. *Tubercle Lung Dis* **79:** 63–73.

Funnell, B.E., and Inman, R.B. (1979) Comparison of partial denaturation maps with the known sequence of SV40 and φX174 RF DNA. *J Mol Biol* **131:** 331–340.

Haggård-Ljungquist, E., Halling, C., and Calendar, R. (1992) DNA sequences of the tail fibre genes of bacteriophage P2: evidence for horizontal transfer of tail fibre genes among unrelated bacteriophages. *J Bacteriol* **174:** 1462–1477.

Hall, S.D., Kane, M.F., and Kolodner, R.D. (1993) Identification and characterization of the *Escherichia coli* RecT protein, a protein encoded by the recE region that promotes renaturation of homologous single stranded DNA. *J Bacteriol* **175:** 277–287.

Hansen, E.B. (1989) Structure and regulation of the lytic replicon of phage P1. *J Mol Biol* **207:** 135–149.

Hatfull, G.F., and Sarkis, G.J. (1993) DNA sequence, structure and gene expression of mycobacteriophage L5: a phage system for mycobacterial genetics. *Mol Microbiol* **7:** 395–405.

Hendrix, R.W., and Duda, R.L. (1998) Bacteriophage HK97 head assembly: a protein ballet. *Adv Virus Res* **50:** 235–288.

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.

Highton, P.J., Chang, Y., and Myers, R.J. (1990) Evidence for the exchange of segments between genomes during the evolution of lambdoid bacteriophages. *Mol Microbiol* **4:** 1329–1340.

Hodgson, D.A. (2000) Generalized transduction of serotype 1/2 and serotype 4b strains of *Listeria monocytogenes*. *Mol Microbiol* **35:** 312–323.

Hoshino, T., McKenzie, T., Schmidt, S., Tanaka, T., and Sueoka, N. (1987) Nucleotide sequence of *Bacillus subtilis dnaB*: a gene essential for DNA replication initiation and membrane attachment. *Proc Natl Acad Sci USA* **84:** 653–657.

Inman, R.B., and Schnös, M. (1979) Partial denaturation of thymine- and BU-containing DNA in alkali. *J Mol Biol* **49:** 93–98.

Ito, T., Katayama, Y., and Hiramatsu, K. (1999) Cloning and nucleotide sequence determination of the entire mec DNA of pre-methicillin-resistant *Staphylococcus aureus* N315. *Antimicrob Agents Chemother* **43:** 1449–1458.

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.

Katsura, I. (1987) Determination of bacteriophage λ tail length by a protein ruler. *Nature* **327:** 73–75.

Kodaira, K.I., Oki, M., Kakikawa, M., Watanabe, N., Hirakawa, M., Yamada, K., *et al*. (1997) Genome structure of the *Lactobacillus* temperate phage φg1e: the whole genome sequence and the putative promoter/repressor system. *Gene* **187:** 45–53.

Krogh, S., O'Reilly, M., Nolan, N., and Devine, K.M. (1996) The phage like element PBSX and part of the *skin* element, which are resident at different locations on the *Bacillus subtilis* chromosome, are highly homologous. *Microbiology* **142:** 2031–2040.

Lazarevic, V., Düsterhöft, A., Soldo, B., Hilbert, H., Mauël, C., and Karamata, D. (1999) Nucleotide sequence of the *Bacillus subtilis* temperate bacteriophage SPβc2. *Microbiology* **145:** 1055–1067.

Lebrun, M., Audurier, A., and Cossart, P. (1994) Plasmid-borne cadmium resistance genes in *Listeria monocytogenes* are present on Tn*5422*, a novel transposon closely related to Tn*917*. *J Bacteriol* **176:** 3049–3061.

Lee, C.S., Davis, R.W., and Davidson, N. (1970) A physical study by electron microscopy of the terminally repetitious circularly permuted DNA from the coliphage particles of *Escherichia coli* 15. *J Mol Biol* **48:** 1–22.

Littlewood, R.K., and Inman, R.B. (1982) Computer-assisted DNA length measurements from electron micrographs with special reference to partial denaturation mapping. *Nucleic Acids Res* **10:** 1691–1706.

Loessner, M.J., and Busse, M. (1990) Bacteriophage typing of *Listeria* species. *Appl Environ Microbiol* **56:** 1912–1918.

Loessner, M.J., Goeppl, S., and Busse, M. (1991) The phagovar variability of *Listeria* strains under the influence of virulent and temperate bacteriophages. *Lett Appl Microbiol* **12:** 192–195.

Loessner, M.J., Krause, I.B., Henle, T., and Scherer, S. (1994) Structural proteins and DNA characteristics of 14 *Listeria* typing bacteriophages. *J Gen Virol* **75:** 701–710.

Loessner, M.J., Wendlinger, G., and Scherer, S. (1995a) Heterogeneous endolysins in *Listeria monocytogenes* bacteriophages: a new class of enzymes and evidence for conserved holin genes within the siphoviral lysis cassettes. *Mol Microbiol* **16:** 1231–1241.

Loessner, M.J., Schneider, A., and Scherer, S. (1995b) A new procedure for efficient recovery of DNA, RNA, and proteins from *Listeria* cells by rapid lysis with a recombinant bacteriophage endolysin. *Appl Environ Microbiol* **61:** 1150–1152.

Loessner, M.J., Maier, S.K., Daubek-Puza, H., Wendlinger, G., and Scherer, S. (1997) Three *Bacillus cereus* bacteriophage endolysins are unrelated but reveal high homology to cell wall hydrolases from different bacilli. *J Bacteriol* **179:** 2845–2851.

Löfdahl, S., Zabielski, J., and Philipson, L. (1981) Structure and restriction maps of the circularly permuted DNA of staphylococcal bacteriophage φ11. *J Virol* **37:** 784–794.

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.

McDonnell, G.E., Wood, H., Devine, K.M., and McConnell, D.J. (1994) Genetic control of bacterial suicide: regulation of the induction of PBSX in *Bacillus subtilis*. *J Bacteriol* **176:** 5820–5830.

Marvik, O.J., Dokland, T., Nokling, R.H., Jacobsen, E., Larsen, T., and Lindqvist, B.H. (1995) The capsid size-determining protein Sid forms an external scaffold on phage P4 procapsids. *J Mol Biol* **251:** 59–75.

Mikkonen, M., and Alatossava, T. (1994) Characterization of the genome region encoding structural proteins of *Lactobacillus delbrueckii* supsp. *lactis* bacteriophage LL-H. *Gene* **151:** 53–59.

Mikkonen, M., and Alatossava, T. (1995) A group I intron in the terminase gene of *Lactobacillus delbrueckii* supsp. *lactis* phage LL-H. *Microbiology* **141:** 2183–2190.

Mouw, A.R., Beachey, E.H., and Burdett, V. (1988) Molecular evolution of streptococcal M protein: cloning and nucleotide sequence of the type 24 M protein gene and relation to other genes of *Streptococcus pyogenes*. *J Bacteriol* **170:** 676–684.

Nakayama, K., Kanaya, S., Ohnishi, M., Terawaki, Y., and Hayashi, T. (1999) The complete nucleotide sequence of phi CTX, a cytotoxin converting phage of *Pseudomonas aeruginosa*: implications for phage evolution and horizontal gene transfer via bacteriophages. *Mol Microbiol* **31:** 399–419.

Nauta, A., van Sinderen, D., Karsens, H., Smit, E., Venema, G., and Kok, J. (1996) Inducible gene expression mediated by a repressor–operator system isolated from *Lactococcus lactis* bacteriophage r1t. *Mol Microbiol* **19:** 1331–1341.

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.

Pedre, X., Weise, F., Chai, S., Lüder, G., and Alonso, J. (1994) Analysis of *cis* and *trans* elements required for the initiation of DNA replication in the *Bacillus subtilis* bacteriophage SPP1. *J Mol Biol* **236:** 1324–1340.

Plasterk, R.H., and van de Putte, P. (1985) The invertible P-DNA segment in the chromosome of *Escherichia coli*. *EMBO J* **4:** 237–242.

Radlinska, M., and Piekarowicz, A. (1998) Cloning and characterization of a gene encoding a new methyltransferase from *Neisseria gonorrhoeae*. *Biol Chem* **379:** 1391–1395.

Rausch, H., and Lehmann, M. (1991) Structural analysis of the actinophage φC31 attachment site. *Nucleic Acids Res* **19:** 5187–5189.

Sambrook, J., Maniatis, T., and Fritsch, E.F. (1989) *Molecular Cloning: a Laboratory Manual*, 2nd edn. Cold Spring Harbor, NY: Cold Spring Harbor Laboratory Press.

Schäferkordt, S., and Chakraborty, T. (1997) Identification, cloning, and characterization of the *lma* operon, whose gene products are unique to *Listeria monocytogenes*. *J Bacteriol* **179:** 2707–2716.

Slutsker, L., and Schuchat, A. (1999) Listeriosis in humans. In *Listeria, Listeriosis, and Food Safety*. Ryser, E.T., and Marth, E.H. (eds). New York, NY: Marcel Dekker, pp. 75–95.

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* CNRZ 1205. *Microbiology* **143:** 3417–3429.

Streisinger, G., Emrich, J., and Stahl, M.M. (1967) Chromosome structure in the phage T4. III. Terminal redundancy and length determination. *Proc Natl Acad Sci USA* **57:** 292–295.

Stuart, S.E., and Welshimer, H.J. (1973) Intrageneric relatedness of *Listeria* Pirie. *Int J Syst Bacteriol* **23:** 8–14.

Takemaru, K., Mizuno, M., Sato, T., Takeuchi, M., and Kobayashi, Y. (1995) Complete nucleotide sequence of a skin element excised by DNA rearrangement during sporulation in *Bacillus subtilis*. *Microbiology* **141:** 323–327.

Thorpe, H.M., and Smith, M.C.M. (1998) *In vitro* site-specific integration of bacteriophage DNA catalyzed by a recombinase of the resolvase/invertase family. *Proc Natl Acad Sci USA* **95:** 5505–5510.

Tye, B.K., Huberman, J.A., and Botstein, D. (1974) Nonrandom circular permutation of phage P22 DNA. *J Mol Biol* **85:** 501–532.

Van Sinderen, D., Luttinger, A., Kong, L., Dubnau, D., Venema, G., and Hamoen, L. (1995) *comK* encodes the competence transcription factor, the key regulatory protein for competence development in *Bacillus subtilis*. *Mol Microbiol* **15:** 455–462.

Wendlinger, G., Loessner, M.J., and Scherer, S. (1996) Bacteriophage receptors on *Listeria monocytogenes* cells are the *N*-acetylglucosamine and rhamnose substituents of teichoic acids or the peptidoglycan itself. *Microbiology* **142:** 985–992.

Wood, D.O., and Waite, R.T. (1994) Sequence analysis of the *Rickettsia prowazekii* gyrA gene. *Gene* **151:** 191–196.

Yung, D.L., and Hollingshead, S.K. (1996) DNA sequencing and gene expression of the emm gene cluster in a M50 group A *Streptococcus* strain virulent for mice. *Infect Immun* **64:** 2193–2200.

Zink, R., and Loessner, M.J. (1992) Classification of virulent and temperate bacteriophages of *Listeria* spp. on the basis of morphology and protein analysis. *Appl Environ Microbiol* **58:** 296–302.

Zink, R., Loessner, M.J., and Scherer, S. (1995) Characterization of cryptic prophages (monocins) in *Listeria* and sequence analysis of a holin/endolysin gene. *Microbiology* **141:** 2577–2584.