Characterization of the genes encoding integrative and excisive functions of Lactobacillus phage og1e: cloning, sequence analysis, and expression in Escherichia coli

著者	KAKIKAWA Makiko, OKI Masaya, WATANABE Nobuyuki, YASUKAWA Hiroo, MASAMUNE Yukito, TAKETO Akira, KODAIRA Ken-Ichi
journal or publication title	Gene -Amsterdam-
volume	185
number	1
page range	119-125
year	1997-01
URL	http://hdl.handle.net/10098/1625

doi: 10.1016/S0378-1119(96)00648-8

Characterization of the genes encoding integrative and excisive functions of *Lactobacillus* phage øgle: cloning, sequence analysis, and expression in *Escherichia coli*

Makiko Kakikawa ^a, Masaya Oki ^a, Nobukatsu Watanabe ^a, Hiroo Yasukawa ^b,
Yukito Masamune ^b, Akira Taketo ^c, Ken-Ichi Kodaira ^a

^a Molecular Biology Group, Chemical and Biochemical Engineering, Faculty of Engineering, Toyama University, 3190, Gofuku, Toyama, Toyama 930, Japan

^b Department of Microbiology, Faculty of Pharmaceutical Sciences, Kanazawa University, Kanazawa, Ishikawa 920, Japan

^c Department of Biochemistry I, Fukui Medical School, Matsuoka, Fukui, Fukui 910-11, Japan

Abstract

øgle is a temperate phage of the *Lactobacillus* strain Gle. The phage-host junctions *attR* and *attL* cloned from the lysogen have a 24-bp common (core) sequence implicated in recombination. DNA sequencing analysis of a 5.2-kbp *SacI* fragment of the øgle phage genome (42.5 kbp) revealed two possible open reading frames (ORF), *xis* and *int*, and the phage attachment (recombination) site (*attP*), whose 24-bp sequence is identical to the core sequence detected in *attR* and *attL*. The deduced *int* product (Int) is a basic protein of 391 amino acids with an estimated *pI* of 9.70, and significantly resembles other presumed integrases encoded by the *Lactobacillus* and *Lactococcus* phages including øadh and øLC3, as well as the *Escherichia coli* phages such as λ. The predicted øgle *xis* protein (Xis) is small and very acidic (66 amino acids; *pI* 4.55), and shows a resemblance (32% overall identity) with a putative excisionase encoded by the *Staphylococcus* phage ø11. The øgle Int with a deduced molecular mass of 45.5 kDa was overproduced in *E. coli* cells, and electrophoretically analyzed.

Keywords: Bacteriophage øgle; Lactic acid bacteria; Integrase; Excisionase; DNA sequence

1. Introduction

Prevalence of lysogeny in various lactic acid bacteria has been reported (Davidson et al., 1990). The lysogenic

phages are thought to encode a set of proteins, which function in recombinational processes of its genome DNA, moving into and out of the host chromosome (integration and excision, respectively).

In coliphages λ and P2 (Leong et al., 1986; Landy, 1989; Campbell, 1994), their integration and excision into and out of the host genome, respectively, depend upon two phage-encoded proteins, besides host proteins such as an integration host factor (IHF). One of these phage proteins, known as an integrase, acts as a site-specific recombinase in both of the integrative and excisional processes, whereas the other protein (an excisionase) participates only in the excision process through a cooperation with integrase.

Several putative integrases have been reported from the *Lactobacillus* phage øadh (Fremaux et al., 1993) and mv4 (Dupont et al., 1995), and the *Lactococcus* phage øLC3 (Lillehaug and Birkeland, 1993), Tuc2009 (Van de Guchte et al., 1994), and BK5-T (Boyce et al., 1995).

Abbreviations: aa, amino acids; *attB*, bacterial attachment site; *attR* and *attL*, phage-host junctions (right and left, respectively); *attP*, phage attachment site; bp, base pair; CBB, Coomassie Brilliant Blue; *cpg*, gene encoding repressor (Cpg); IHF, integration host factor; *int*, gene encoding integrase (Int); IPTG, isopropyl β-D-thiogalactopyranoside; LB, Luria-Bertani (medium); *Lorf143*, ORF encoding 143-aa protein (gpLorf143); *Lorf304*, ORF encoding 304-aa protein (gpLorf304); *M_w*, molecular weight; ORF, putative open reading frame; *pI*, isoelectric point; PAGE, polyacrylamide gel electrophoresis; *plac*, promoter of *lacZ'* gene; *RB*, ribosomal binding site; *Rorf148*, ORF encoding 148-aa protein; *Rorf192*, ORF encoding 192-aa protein; *Rorf232*, ORF encoding 232-aa protein; *RS*, repeated sequence; SDS, sodium dodecyl sulfate; *Ti*, possible secondary structure; *xis*, gene encoding excisionase (Xis); YT, yeast-Tryptone (medium).

Unlike the lambdoid phages, however, molecular details on integration and excision of the phages of lactic acid bacteria are still insufficient; structures and functions of their integration and excision genes mostly remain to be elucidated.

Recently, we isolated a new *Lactobacillus* temperate phage øgle, and identified several genes specifying the capsid proteins (Kakikawa et al., 1996), and the cyto-lytic function (Oki et al., 1996a).

In this report, we describe cloning and sequencing of the phage-host junctions between øgle and its host G1e, and øgle genes *xis* and *int* for its integration and/or excision, as well as the phage attachment (recombination) site (*attP*).

2. Materials and methods

2.1. Bacteria, phage and plasmids

The *Lactobacillus* strain G1e and its temperate phage øgle are from our laboratory stock, and were propagated in M17 broth as described previously (Kakikawa et al., 1996). The *E. coli* vector plasmids (pUC18, pUC19, pUC118 and pUC119) and their host XL1-Blue are from our laboratory stock, and were propagated using LB or 2YT broth (Oki et al., 1996a).

2.2. Analysis of DNA and protein

The øgle DNA was extracted from the virus particles (Kakikawa et al., 1996). The G1e chromosomal DNA was prepared by the method of Katty et al. (1989). Southern blot hybridization was performed as described previously (Oki et al., 1996b); ³²P-labeled probes were prepared by nick translation. Cloning of øgle DNA was carried out essentially as described by Oki et al. (1996a). Recombinant plasmids were introduced into *E. coli* XL1-Blue cells by Ca²⁺-dependent transformation or electroporation (Oki et al., 1996b). For DNA sequencing, various deletion clones were constructed from the restriction library of øgle DNA, by exonuclease digestion (Oki et al., 1996a). The DNA sequence was determined by the chain termination method (Sanger et al., 1982). SDS-polyacrylamide gel electrophoresis (PAGE) was performed as described by Kakikawa et al. (1996); proteins were visualized by staining with Coomassie Brilliant Blue (CBB) R-250.

All other procedures were performed as described previously (Kodaira et al., 1996).

2.3. Enzymes and biochemicals

Restriction enzymes, phage T4 DNA ligase, and alkaline phosphatase (calf intestine) were purchased from

Fig. 1. Physical map of the øgle *xis-int* region. A 5204-bp *SacI* fragment (*SacI*-II) sequenced in this study (see text and Fig. 4) was shown. The possible ORFs were indicated by boxes above the fragment. *Rorf232* has no putative ribosomal binding sites. Dr, *DraI*; Ec, *EcoRI*; Hd, *HindIII*; Kp, *KpnI*; Sl, *SalI*. Open circle, inverted repeat; closed circle, *p*-independent-like mRNA terminator. The four thick lines (probes 1–4) indicate restriction fragments used as probes for Southern blot analysis. The four recombinant plasmids constructed in this study (see text) were shown under the probes: dotted boxes, under the *lacZ'* promoter (small closed circle); open boxes, free from the *lacZ'* promoter.

Takara Shuzo (Kyoto) and Nippon Gene (Toyama). Buffers for each enzymes were as recommended by the manufacturers.

[α - 32 P]dCTP was from NEN (USA). All other materials were as described previously by Kodaira et al. (1996).

3. Results and discussion

3.1. Cloning and sequence of the *øgle* DNA

As a prophage, *øgle* genome is probably integrated within the host *Gle* chromosome between a phage attachment site (*attP*) and a host bacterial attachment site (*attB*), to generate two additional sites at the right (*attR*) and left (*attL*) junctions. For mapping the phage-host junctions, DNAs extracted from *øgle* and *Gle* were digested with several restriction enzymes, and subjected to Southern blot analysis with various *øgle* phage DNA fragments as probes (e.g., probes 1, 2, 3 and 4; see Fig. 1). When probe 2 (a 0.7-kbp *Hind*III fragment) was hybridized to *Sac*I digests of the *Gle* DNA (Fig. 2), two bands of about 4.3 and 24 kb, corresponding to the junction fragments (designated as *attL* and *attR*, respectively) (Fig. 2), were detected, in addition to the 5.2-kb *Sac*I *øgle* fragment containing

Fig. 2. Detection of the phage-host junctions. DNA extracted from *Gle* lysogen was digested with a restriction enzyme (*Sac*I and/or *Sal*I), and subjected to Southern blot hybridization (Section 2) probed to a 0.7-kbp *Hind*III fragment (probe 2) of *øgle* phage DNA (see Fig. 1). Lane 1, *Sac*I; lane 2, *Sac*I and *Sal*I (a few partial bands were located above a 7.8-kbp fragment). Arrows indicate the fragments containing the right (*attR*) or left (*attL*) junction (for the details, see text). Dotted arrows represent phage DNA fragments (see text).

attP (see below and also Fig. 1 and Fig. 4). Like the most related *Lactobacillus* phage mv4 (Dupont et al., 1995), an extrachromosomal form of *øgle* DNA seems to coexist with the integrated prophage form.

The DNA region containing *attL* (1.5-kb *Sac*I/*Sal*I fragment, Fig. 2) or *attR* (0.6-kb *Hha*I fragment, data not shown) was cloned into *E. coli* pUC118, and sequenced. As shown in Fig. 3, both *attR* and *attL* have an identical 24-bp stretch (5'-TGACCCACCCG-GACTCGAACCGG-3'), that is common to *attP* (see below and Fig. 4). The 24-bp sequence appears to function as a core domain for the strand exchange reaction in the *øgle* integration and/or excision pathways.

For analysis of *attP* and its surroundings, the 5.2-kbp *Sac*I-*II* fragment of *øgle* phage DNA (Fig. 1) was cloned into a *Sac*I site of pUC18, resulting in two recombinants, p18intScL and p18intScR (Fig. 1). As summarized in Fig. 1, sequencing analysis of the 5204-bp *Sac*I-*II* DNA (Section 2) revealed that one strand contains five possible ORFs in addition to a 24-bp *attP* sequence (see above), whose order is *cpg-Lorf143-Lorf304-xis-int-attP*, whereas the complementary strand (R-strand) has three potential ORFs *Rorf148-Rorf192-Rorf232* (the two putative ORFs *cpg* and *Rorf148* are truncated). These predicted ORFs (except *Rorf232*) are preceded by a possible ribosomal binding sequence (Table 1, see also Fig. 3), which is in good agreement with those of several genes found in *Lactobacillus* and *Lactococcus* species: 5'-AGGAGG-3' (Pouwels and Leer, 1993) and 5'-AGAAAGGAGGT-3' (Ludwig et al., 1985; Schouler et al., 1994), respectively.

In Fig. 4, a 2.8-kbp *øgle* DNA containing *xis-int-attP* was shown together with their deduced amino acid sequences. The predicted *int* product (termed Int) is a basic protein of 391 amino acids (for details, see below), whose estimated molecular mass and *pI* are 43 kDa and 9.70, respectively (Table 1). On the other hand, the putative *xis* product (Xis) (see below) is a small and very acidic protein (66 amino acids; 7.6 kDa; *pI* 4.55) (Table 1). Recently, we have cloned and sequenced *cpg* (Kodaira et al., unpublished results), demonstrating that its putative product of 132 amino acids contains a helix-turn-helix motif in its N-terminus, and resembles other repressors encoded by the lambdoid phages such as λ , P2 and P22 (Ljungquist et al., 1984). The roles of *Lorf143*, *Lorf304* and *Rorf192* (Fig. 1 and Table 1) in the *øgle* multiplication have not been ascertained. On

<i>attL</i>	5'-TACAATAATCAACCGTAAATAACGGTTTGA	TGACCCACCCGACTCGAACCGG	GATCAACCGCTTAGGAGGCGGGTGCCCTAT-3'
<i>attP</i>	5'-TACAATAATCAACCGTAAATAACGGTTTGA	TGACCCACCCGACTCGAACCGG	TGTTGAGAGAGCACTGACATATCAGTGTTT-3'
<i>attR</i>	5'-TATAGGACGGTTTTTGAAGAAATATTTAAC	TGACCCACCCGACTCGAACCGG	TGTTGAGAGAGCACTGACATATCAGTGTTT-3'

Fig. 3. DNA sequences of the phage-host junctions. The two phage-host junctions, *attR* and *attL* (see Fig. 2), were cloned (see text), and sequenced. The 24-bp common sequences among *attR*, *attL* and *attP* (see Fig. 4) were boxed.

End of Lorf304

S Q G N I I A E T S F T G N F K Y T G E *
 ATAGTCAAGGAAATATAATTGCAGAACTAGTTTCACTGGTAACITTAATATACCTGGTAATAAATAAGCCCTGTTTCAGGGCTTTACGCGAGTGT 2601
 <----->

AGTTTAGTGGTAAACGATAGCCTTCCAAGCTGTAGTCGCGGGTCCGATTCCCGTCACTCACTTAGTACCCCAITATTGGGGTATATATTTCAATCAAA 2701
 <----->

Xis

AAGAACATACGTTTGTTCATTAACTGAAAGGATGTGTTGTATGCCAGATGAAAGCTTGCTAAGCTTGATTCAATGATAGGATTGCCGTTGCCAATCA 2801
 R.B.

L K N D L G Y A V Y D D F L I T D V P F P T N I A S Y H R F L A C
 ATTAAGAATGATTTGGGATATGCCGTCTATGATGACTTTCTGATAACCGATGTCCCTTTTCGACGAACATTTGCTTCTATCATCGATTTTTCAGCTGT 2901
 <----->

Int

D R P F F N K A I D S L E G * N A Y L Y K R A D G L W H W R I N R
 GACAGTTCTTTAAACAAAGCCATAGATAGTTTGGAGGATAAAAAACATGGCTTACTTGTATAAGCGTGCAGATGGATTGGCATTGGCGTATCAACCGC 3001
 R.B.

T I D G Q R V P I N S A G G F K L K S A A K E E A E E I E N Q P R H
 ACTATTGATGGTCAGCGTGTACCTATCAATTTCAGCAGGCGGTTTAACTAAAGAGTGCTGCCAAAGAAGAAGCTGAGGAGATAGAAAATCAGTTTCGAC 3101

G T Y V E T T D E S F A A Y Y K K W N E T F W I G K K G P D A D R
 ATGGCACTTACGTTGAAACCCGATGAAAGTTTGGCCATATTAAAAATGGATGGAACTTTTGGATTGGCAAAAAGGACGACGCTGATCG 3201

H Y R D A L K C I E H Y F P H A R M K D I T H D Q Y Q L F I N N F
 CCCTACTCGTGATGCCCTAAATGCATTGAACACTATTTTCGCGACGCTCGAATGAAGGACATCACTCATGACCACTACCACTATTATCAATAATTTT 3301

A K S H A K S T V H Q R H N Y I K K C L I E A F E E G I I K R N P A
 GCCAAAGCCATGCTAAGAGCAUTGTTATGACGCGCCATAACTACATCAAGAAATGCTTAATTGAAGCATTGAGGAAGGCATCATTAAACGAAATCCGG 3401

A R I N L T G N K N R E K K E E V K F N S L D D P K K I H N A A Y
 CTGCTAGAACATTTGACCGGTAAATAAGAACCCGGAGAAAAAGAAGTTAAATTTATGAGTTTGGATGACTTCAAAAAAATTTATGAATGCAGCTTA 3501

R K F D P N S P S T S M I I L H G A T G L R F E E A D G L T W D C
 TCGCAAGTTTGACCCAACTCCCTTCCACTTCTATGATCATTTCTAATGGGGGCACTGGGCTGCGGTTTGAAGAAGCTGATGGTCTAACCTTGGGATTGT 3601

V D F K N S S I T I N K T W D Y R H K H D F G P T K N P Q S N R T I
 GTGACTTTAAGAACAGTAGCATTACCATCAACAAGACCTGGGACTATCGTCATAAACAGGACTTTGGCCCCACTAGAACCACAGTCAATGCGAACTA 3701

K V D S T T H R V L K A L H A H Y A Q I K L V R H D W N K R K L V
 TTAAAGTAGATTGCAACCATGCGTGTCTTAAAGCGTTGACGCGCCACTACGCACAAATTAATTTAGTGGCGCATGACTGGAAACAAGCGTAAGCTAGT 3801

F S K S D G I P L S N N A I N K M L R S L C V Q N A V V A R N S S
 TTCTCAAAATCAGCGGCATCCCACTATCAACAACGCAATTAACAAATGTTTGGCAAGCCTGTGCGTTCAAAATGCAGTCTAGCAAGGAATCCAGC 3901

G V I Q K W Y T C H A L R H T H A S L L L Y E G R D I S Y V S K R L
 GGTGTAATTAAGAAATGGTACAGTGGCAGCCCTCGTCATACTACGCGATCACTGCTTTTATACGAGGGTGGGATATTTCTTACGTGTCAAGCGGT 4001

G H K D I N T T Y N Y T H V I Q E N S A R E D E A L D P T H S K
 TAGGCCATAAGGACATAATGACTACGTACAAACGATATACGACGTCATCCAGGAATGAGTGCCCGCAAGATGAAGCTCTTGATCCAACCATGAGCAA 4101

I F S K Q A *
 GATTTTTTCTAAGCAAGCTTAAATCTCGTGTACGTGTCTGTGTACCTCTGTGTACGTAATAATAGCAAACGGAAGGTTACAAAAACATCATGCGCACA 4201
 ===== RS-1 RS-2 RS-3

AAAAAAGAACGCCAATCGTTGATATAACAACGATTAGCGTTCTTCAATCTTACAATAATCAACCGTAATAACCGTTTGTATGACCCCAACCGGACTCGA 4301
 <----->

IHF

ACCGGTGTTTCAGAGAGCACTGACATATCAGTGTTTAATAGGTGTCACTCAAATATTCGTGTACCTCTCGTGTACCTTTTATATAAAATCAATGTAACGGC 4401
 <----->

IHF

CGTGTGCAAGATACGGCTTTTCAATATCTAATTGCCTAATTATATCATATAGAACATTTGAATTTACATTGACAATCAAAACAGGATTTTAAATTAT 4501
 <----->

Ti1

TTCCAAATAAAAAATGAAATGCGAGTCAACTTGGGATTTCAGATAGATGTGCGGGCCCCACCGGCTATCTGTTTAATTTACCCACGCTACCCCGGCTGG 4601
 <----->

Ti2

GTATTTTTTTCGGCAGAAATTTGAGTTTCGGCAGCTAACCGCGCTTCTTTAGCCTCTTCAATCGTATCAAAGTAGCCAAATACGCGTTTGAACGTGATTAAC 4701
 <----->

ATAAATCTGGCCAGGTACTTTGGCCATTTTTCGCTGAACCTCGGAACTCCCGTGACTCCAGTTTACTATTCTTAGCAACTTTTTTTCGGGTTGCTA 4801
 <----->

Ti3

TCAATCAAGCCCAAGGACTCCATCGACGAAAGCTTATCATAGTGCCTTCTCAATGCATTCAACCGCTCGGACGTTTTGAATCCATGACCGCAGC 4901

TTGTGCGTATGCCCTTTACCAGCGGGCTAGCGGGAACGTGAATTTGCTTGCACAGTGCCTCGCAAAGCCACAACTTTTGGGTGATTATCATCCAC 5001

CGGAATACGCTCAAGAACAGTTAAATGGCCAACTGACGCGCAGTCAAATCGCCATACCTTGATTTGTGGGGTGACGATACCTGTCAACCAAGTAGTT 5101

TGACCAACCGCTGTTACGCGTCCAGTGTCTACTTGTACCTGGTGGCACAAGTACACTGACACAGCCACATGTATCACCAGAGTGACTTCGGCCATCCA 5201

TCTTTTTTAGTACCGTCAGATGGCCGAAGTCTTCCACTCAAGTCTCGCGTCTTACTATTCTTTTTTTCATATCTTATAATTTGGCGCCTAGATT 5301
 * L K A G L N
 End of Rorf192

Table 1
Putative gene located in the surroundings of the *øgle attP* site

Gene	Ribosomal binding sequence	Predicted gene product			
		Protein	aa	$M_w (\times 10^{-3})$	pI
<i>Lorf143</i>	AaAAAGtAGGT-gtttacATG	gp <i>Lorf143</i>	143	16 333	4.35
<i>Lorf304</i>	AatttGGAGGa-gtaaataATG	gp <i>Lorf304</i>	304	33 578	8.75
<i>xis</i>	cGAAAGGAtGT-gtttgtATG	Xis	66	7557	4.55
<i>int</i>	AGtttGGAGGgataaaaacATG	Int	391	45 515	9.70
Consensus	AGAAAGGAGGT				

The consensus sequence for *Lactococcus* 16S rRNA is from Ludwig et al. (1985). Capital letters (consensus) indicate the nucleotides identical to those of the 16S rRNA. Dashes represent gaps inserted to align nucleotides for maximal homology. ATG, starting codon; aa, number of amino acid residues; M_w and pI, molecular weight and isoelectric point, respectively, estimated from DNA sequence.

Rorf148 (Fig. 1), present and previous (Oki et al., 1996a) studies have shown that its deduced product (a 17.0-kDa protein composed of 148 amino acids with a pI of 9.80) has an arginine-rich N-terminus, and somewhat resembles the arginine-rich motif found in lambdoid antiterminators (Burd and Dreyfuss, 1994).

3.2. Characterization of *øgle Int* and *Xis*

øgle Int exhibits significant similarity to other putative integrases encoded by the *Lactobacillus* phage *øadh* (Fremaux et al., 1993) and *mv4* (Dupont et al., 1995), and the *Lactococcus* phage *øLC3* (Lillehaug and Birkeland, 1993), *Tuc2009* (Van de Guchte et al., 1994), and *BK5-T* (Boyce et al., 1995), as well as the lambdoid phage including λ , $\phi 80$, *P2*, and *P22* (Campbell, 1994). For example, overall identity between *øgle* and *øLC3* is 37%.

The integrase family of lambdoid phages has two highly conserved regions, known as domains 1 and 2 (Sherratt, 1993); the former, located in the central region, contains an Arg residue, whereas the later, situated in the C-terminus, carries a typical Tyr residue, which has been presumed to join to a recessed 3'-phosphate after the initial integrase-mediated phosphodiester cleavage (Sherratt, 1993; Han et al., 1994). Like other integrases, *øgle Int* also has two homologous domains 1 (from 204 to 218) and 2 (from 328 to 366), which contain Arg-207 and Tyr-364, respectively (Fig. 4). In Fig. 5, the two domains of *øgle* are compared with those of other phages *øadh* (Fremaux et al., 1993), *mv4* (Dupont et al., 1995), *øLC3* (Lillehaug and Birkeland, 1993), and λ (Sherratt, 1993).

øgle Xis shows a resemblance (32% identity and 46% similarity) with a presumed excisionase encoded by the *Staphylococcus* phage $\phi 11$ (Ye et al., 1990). Unlike other

excisionases of the lambdoid phages λ , *P22*, and $\phi 80$ (Leong et al., 1986) the two *Xis* proteins of *øgle* and $\phi 11$ are significantly acidic (Table 1), suggesting that *øgle Xis* functions via a mechanism different from the lambdoid excisionases, as in $\phi 11$ (Campbell, 1994; Ye et al., 1990).

3.3. Properties of the *øgle attP* region

The regions preceding and following *xis-int-attP* are rich in inverted sequence and secondary structure. An intergenic region between *Lorf304* and *xis* contains three successive secondary structures (Fig. 4), which may function in control of the *xis* and/or *int* expression. In coliphage *P2* (Yu et al., 1994), three secondary structures located just upstream of *int* has been predicted to be involved in control of the *int* expression. In *øgle xis*, there exists an imperfect secondary structure preceded by mRNA promoter-like sequences (see Fig. 4), which may function in the *int* expression during the integrative pathway, like the λ *int* system (Campbell, 1994).

DNA regions flanking *øgle attP* have also several inverted and/or direct repeats (Fig. 4). Three possible secondary structures (*Ti*) located downstream of *attP* end in a run of T residues, and somewhat resemble a ρ -independent mRNA terminator of lambda *int* (Guarneros et al., 1982). On the other hand, this DNA region of *øgle* contains five homologous repeated sequences (*RS*), each of which is composed of 11 bases with a consensus of 5'-TNGTGTACNTN-3': three of them (*RS1-RS2-RS3*) are located upstream of *attP*, and another two (*RS4-RS5*) are situated downstream of *attP*. In addition, *attP* is accompanied by three possible IHF-binding sequences: two of them are localized between *RS-3* and *attP*, and the other one is between *attP* and *RS-4*. These IHF-binding-like sequences of

Fig. 4. DNA sequence of the *øgle xis-int* region. A 2800-bp DNA of the 5204-bp *SacI*-II fragment (see Fig. 1) was shown together with deduced amino acid sequences of *Lorf304-xis-int-Rorf192*. The 24-bp core sequence (*attP*) was boxed (see text). Asterisks indicate stop codons. Bold italic lower letters indicate putative ribosomal binding sites (RB). Arrows show inverted repeats. *Ti*, possible ρ -independent mRNA terminator. Dotted double lines represent repeated sequences (*RS*). IHF, potential IHF-binding site.

	Domain 1	Domain 2
ϕ gle	TGLRF--EEADGLTWDCVDFK	HALRHTHASLLLYEG--RDISYVSKRLGHKDIMTTYNTYTH
ϕ Lc3	TGMRA--SEAFGLVWDDIDFN	HGLRHTHASVLLYHG--VDIMTVSKRLGHASVAITQQTYIH
ϕ adh	TGMRL--GEGCGLLVKNIFQN	HIFRHTHVSKLAEEG--YPLSLITDRVGHANSIDTRKIYLN
mv4	TGLRKRSTSPGLVGHRLRSK	HGFRHTFASLLITADPSIKPTDVQAILGHESIDITMEIYMH
λ	TGQRV--GDLCMKWSDIVDG	HELRSLSA-RLYEKQISDK--FAQHLLGHKS-DTHASQYRD
Consensus	TG-R---e--gl----i---	H--Rht-as-L---g-----v--rlGH----it---Y-h

Fig. 5. Comparison of ϕ gle Int with other integrases. The two conserved regions (domains 1 and 2) among the phage-encoded integrases were compared: ϕ gle (this study), the *Lactococcus* phage ϕ LC3 (Lillehaug and Birkeland, 1993), the *Lactobacillus* phage ϕ adh (Fremaux et al., 1993) and mv4 (Dupont et al., 1995), and the *E. coli* phage λ (Sanger et al., 1982). Domains 1 and 2 (boxed) were from Sherratt (1993). Dashes indicate positions with conservative substitutions. Consensus indicates the conserved aa residues; capital letters are conserved through the integrases.

ϕ gle are apparently similar to the *E. coli* IHF-binding consensus of 5'-(C/T)AANNNTTGAT(A/T)-3' (Landy, 1989).

3.4. Expression of the ϕ gle int product in *Escherichia coli*

As shown in Fig. 1, p18intScL (see above) carries *int* under *plac*, but this plasmid did not overproduce Int in *E. coli* XL1-blue cells, by treatment with IPTG (data not shown). Therefore, a DNA region of 1585 bp spanning from 2901 to 4485 (Fig. 4) was cloned into a *Sma*I site of pUC18, resulting in two plasmids, p18intL (under *plac* control) and p18intR (free from *plac*) (Fig. 1): the cloned ϕ gle DNA region contains a possible ribosomal binding site (5'-GGAGG-3'), *int*, *attP*, and one putative mRNA terminator *Til* (Fig. 4). Unlike p18intScL, *E. coli* XL1-blue cells carrying p18intL cannot grow stably at 37°C even without induction, but multiply normally at 25°C.

Total content of proteins extracted from XL1-Blue/p18intL was analyzed by SDS PAGE (Section 2). Upon induction by IPTG (1 mM) at 37°C, p18intL directed synthesis of a protein increasing steadily; its apparent molecular mass is 42.0 kDa (Fig. 6). This protein was not detected when IPTG was omitted or the cells harboring p18intR or pUC18 were used. The overproduced protein is probably Int, although the mass on an SDS gel was slightly lower than that (45.5 kDa) predicted from DNA sequence (Table 1). By gel mobility-shift assay, a cell-free extract from XL1-Blue/p18intL resulted in a lower mobility of a DNA fragment containing the *attP* region specifically (e.g., a 90-bp *Hind*III/*Esp*I fragment from nucleotide positions 4117 to 4196, see Fig. 4) (unpublished results).

The results in this study suggested that the two ϕ gle putative proteins, Int and/or Xis, are involved in the phage movement into or out of the host chromosome. Int probably acts as a site-specific recombinase in both excision and integration using two specific attachment

Fig. 6. Expression of ϕ gle Int in *E. coli*. *E. coli* XL1-blue carrying p18intL, p18intR or pUC18, was grown at 25°C in LB medium containing ampicillin (60 µg/ml). When the A_{600} of the culture had reached 0.5, the growth temperature was shifted to 37°C with concomitant addition of IPTG (final concentration 1 mM). At intervals, a portion of the culture was removed, disrupted, and subjected to 10% SDS PAGE (Section 2). Lane 1, markers; lane 2, no plasmid (after 3 h); lane 3, pUC18 (3 h); lane 4, p18intL (1 h); lane 5, p18intL (3 h); lane 6, p18intR (3 h). An arrow indicates the putative Int overproduced.

sites, *attP* and *attB*, locating on the phage genome and on the host chromosome, respectively. Like the ϕ 11 excisionase (Ye et al., 1990), ϕ gle Xis, an acidic protein, may participate in the excisive pathway, through a cooperation with Int. As in λ (Landy, 1989), the ϕ gle integrative and excisional pathways are probably under a complicated control. Further studies on structure, expression, and function of ϕ gle *int* and *xis* are now in progress.

References

- Boyce, J.D., Davidson, B.E. and Hillier, A.J. (1995) 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.
- Burd, C.G. and Dreyfuss, G. (1994) Conserved structures and diversity of functions of RNA-binding proteins. *Science* 265, 615–621.
- Campbell, A. (1994) Comparative molecular biology of lambdoid phages. *Annu. Rev. Microbiol.* 48, 193–222.
- Davidson, B.E., Powell, I.B. and Hillier, A.J. (1990) Temperate bacteriophages and lysogeny in lactic acid bacteria. *FEMS Microbiol. Rev.* 87, 79–90.
- Dupont, L., Boizet-Bonhoure, B., Coddeville, M., Auvray, F. and Ritzenhaler, P. (1995) Characterization of genetic elements required for site-specific integration of *Lactobacillus delbrueckii* subsp. *bulgaricus* bacteriophage mv4 and construction of an integration-proficient vector for *Lactobacillus plantarum*. *J. Bacteriol.* 177, 586–595.
- Fremaux, C., de Antoni, G.L., Raya, R.R. and Klaenhammer, T.R. (1993) Genetic organization and sequence of the region encoding integrative functions from *Lactobacillus gasseri* temperate bacteriophage ϕ adh. *Gene* 126, 61–66.
- Guarneros, G., Montanz, C., Hernandez, T. and Court, D. (1982) Posttranscriptional control of bacteriophage λ *int* gene expression from a site distal to the gene. *Proc. Natl. Acad. Sci. USA* 79, 238–242.
- Han, Y.W., Gumport, R.I. and Gardner, J.F. (1994) Mapping the functional domains of bacteriophage lambda integrase protein. *J. Mol. Biol.* 235, 908–925.
- Kakikawa, M., Oki, M., Tadokoro, H., Nakamura, S., Taketo, A. and Kodaira, K.-I. (1996) Cloning and nucleotide sequence of the major capsid proteins of *Lactobacillus* bacteriophage ϕ gle. *Gene*, 175, 157–165.
- Katty, J., Trees, S., Frank, M., Christ, P., Patrick, S., Henk, J., Patrick, D., Mark, Z. and Jacques, M. (1989) Characterization of a Gram-positive broad-host-range plasmid isolated from *Lactobacillus hilgardii*. *Plasmid* 21, 9–20.
- Kodaira, K.-I., Oki, M., Kakikawa, M., Kimoto, H. and Taketo, A. (1996) The virion proteins encoded by bacteriophage ϕ K and its host-range mutant ϕ KhT: host-range determination and DNA binding properties. *J. Biochem.* 119, 1062–1069.
- Landy, A. (1989) Dynamic, structural, and regulatory aspects of λ site-specific recombination. *Annu. Rev. Biochem.* 58, 913–949.
- Leong, J.M., Nuncs-Duby, S.E., Oser, A.B., Lesser, C.F., Youderian, P., Susskind, M.M. and Landy, A. (1986) Structural and regulatory divergence among site-specific recombination genes of lambdoid phage. *J. Mol. Biol.* 189, 606–616.
- Lillehaug, D. and Birkeland, N.-K. (1993) Characterization of genetic elements required for site-specific integration of the temperate lactococcal bacteriophage ϕ LC3 and construction of integration-negative ϕ LC3 mutants. *J. Bacteriol.* 175, 1745–1755.
- Ljungquist, E., Kockum, K. and Bertani, L.E. (1984) DNA sequences of the repressor gene and operator region of bacteriophage P2. *Proc. Natl. Acad. Sci. USA* 81, 3988–3992.
- Ludwig, W., Seewaldt, E., Klipper-Balz, R., Schleiffer, 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.
- Oki, M., Kakikawa, M., Yamada, K., Taketo, A. and Kodaira, K.-I. (1996a) Cloning, sequence analysis, and expression of the genes encoded lytic functions of Bacteriophage ϕ gle. *Gene* in press.
- Oki, M., Kakikawa, M., Taketo, A. and Kodaira, K.-I. (1996b) Genome structure of a *Lactobacillus* plasmid pNMO belonging to the rolling-circle type replicon. *J. Gen. Appl. Microbiol.* 42, 271–283.
- Pouwels, P.H. and Lee, R.J. (1993) Genetics of *Lactobacilli*: plasmid and gene expression. *Antonie van Leeuwenhoek* 84, 85–107.
- Sanger, F., Coulson, A.R., Hong, G.F., Hill, D.F. and Petersen, G.B. (1982) Nucleotide sequence of bacteriophage λ DNA. *J. Mol. Biol.* 162, 729–773.
- Schouler, C., Ehrlich, S.D. and Chopin, M.-C. (1994) Sequence and organization of the lactococcal prolate-headed bIL67 phage genome. *Microbiology* 140, 3061–3069.
- Sherratt, D.J. (1993) Site-specific recombination and the segregation of circular chromosomes. *Nucleic Acids Mol. Biol.* 7, 202–216.
- Van de Guchte, M., Daly, C., Fitzgerald, G.F. and Arendt, E.K. (1994) 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.
- Ye, Z.-H., Buranen, S.L. and Lee, C.Y. (1990) Sequence analysis and comparison of *int* and *xis* genes from staphylococcal bacteriophages L54a and ϕ 11. *J. Bacteriol.* 172, 2568–2575.
- Yu, A., Barreiro, V. and Haggard-Ljungquist, E. (1994) Regulation of *int* gene expression in bacteriophage P2. *J. Bacteriol.* 68, 4220–4226.