

The Virion Proteins Encoded by Bacteriophage ϕ K and Its Host-Range Mutant ϕ KhT: Host-Range Determination and DNA Binding Properties¹

Ken-Ichi Kodaira,*[†] Masaya Oki,* Makiko Kakikawa,* Hisashi Kimoto,[†] and Akira Taketo[†]

*Molecular Biology Group, Chemical and Biochemical Engineering, Faculty of Engineering, Toyama University, 3190 Gofuku, Toyama, Toyama 930; and [†]Department of Biochemistry I, Fukui Medical School, Matsuoka, Fukui 910-11

Received for publication, December 15, 1995

The microvirid phage ϕ K, specific for *Escherichia coli* K12, contains a circular single-stranded (SS) DNA in the icosahedral virion, which comprises four phage gene products, F (capsid), G (major spike), H (minor spike), and J (core). ϕ KhT, a host-range mutant of ϕ K, can grow on *E. coli* C and B, besides K12, and is more thermosensitive than the parental phage ϕ K. Sequencing analysis revealed that the genome of ϕ K and ϕ KhT consists of 6,089 nucleotides (nt), and codes for eleven genes, whose sequences are similar to those of α 3, ϕ X174, and G4 infective to strain C. In ϕ KhT, two nt had changed: one is in the gene G, resulting in replacement of the 75th codon Ala with Ser, and the other is at 67th codon of the gene H: Val to Ala. Chemically synthesized gene J protein composed of 23 amino acids (aa) binds to ϕ K SS DNA more tightly than and preferentially over the host *E. coli* SS-DNA-binding protein (SSB). These results indicate that the two spike proteins G and H are involved in the determination of ϕ K host-range, and support a model in which the gene J protein functions in packaging the viral SS DNA into the virion vesicle.

Key words: bacteriophage, DNA-binding protein, host-range, ϕ K, virion proteins.

Various strains of microvirid (isometric) phages of *Escherichia coli* have been extensively used as probes for elucidation of various biochemical and genetic problems, including gene structure, DNA replication, transcription, and host-parasite relationship (1, 2). These phages contain a circular SS DNA in the icosahedral virion, which is composed of four phage gene products, F (capsid), G (major spike), H (minor spike), and J (core), and are classified into three major groups: (i) the first group consists of ϕ X174, S13, and G6, (ii) the second contains G4, ϕ C, and U3, and (iii) the third includes α 3, ϕ K, and St-1 (2, 3). ϕ X174 and G4 can grow only on *E. coli* C, and α 3 infects *E. coli* B as well as *E. coli* C, whereas ϕ K and St-1 propagate on *E. coli* K12 (3). Total genome sequencing has been accomplished with three species ϕ X174 (4), G4 (5), and α 3 (6). Comparative studies have indicated that the three phages closely resemble each other, and code for the same series of eleven genes A, A*, B, C, K, D, E, J, F, G, and H. However, the details of the mechanism of the host-range determination are still unknown (2): no host-range mutants have so far been isolated from the phages ϕ X174, G4, and α 3 (2), and it is unclear which of virion protein(s) F, G, H, and/or J is determinant for the host-range (or adsorption).

As in ϕ X174, G4, and α 3 (2), the life cycle of ϕ K consists of three successive stages (3, 7–9): (i) synthesis of parental double-stranded (RF) DNA (stage I), (ii) replication of progeny RF DNA by a rolling circle (RC) mechanism

(stage II), and (iii) RC-type synthesis of progeny viral SS DNA driven by packaging into a virion vesicle composed of ϕ K gene products F, G, and H (stage III or morphogenesis). The stages I and II depend upon host *E. coli* SSB (7, 10), whereas the stage III is thought to rely on ϕ K gene J protein, instead of SSB (11, 12). However, the function of the gene J protein remains to be clarified in detail: e.g., whether or not the gene J protein binds to SS DNA more preferentially over SSB.

In order to obtain further insight into role of the virion-forming proteins F, G, H, and J in determination of host-range and interaction with DNA, total genomes of ϕ K and its host-range mutant ϕ KhT were sequenced, and a 23-aa ϕ K gene J protein was chemically synthesized. The results obtained in this study clearly indicate that the ϕ K host-range is determined by the major (G) and minor (H) spike proteins, and that the synthetic J proteins bind to the viral SS DNA more tightly than and preferentially over SSB *in vitro*.

MATERIALS AND METHODS

Phages and Bacterial Strains—Bacteriophage ϕ K and *E. coli* K12 W3110, B, C, and their lipopolysaccharide-defective (LPS) mutants (Table I) were from our laboratory stock (3). ϕ K was multiplied in *E. coli* K12 W3110 (7). ϕ KhT, which infects *E. coli* C and B as well as K12 (see below), is a host-range mutant spontaneously derived from ϕ K, and was propagated on *E. coli* C or K12 W3110. Other phages, *E. coli* strains, and plasmids used in this study are

¹ Nucleotide sequence data have been registered with EMBL Data Library (accession number: X60323).

TABLE I. Strains of *E. coli*, phages, and plasmids.

Strain	Properties	Source/ Reference
<i>E. coli</i> bacteria		
K12 W3110	Delivative of K12	Lab. stock
C	Wild-type	Lab. stock
C61	LPS mutant of C	Lab. stock
C71	LPS mutant of C	Lab. stock
C23	LPS mutant of C	Lab. stock
BB	Wild-type	Lab. stock
BB1	LPS mutant of BB	Lab. stock
BB5	LPS mutant of BB	Lab. stock
XL1-blue	<i>supE44, hsdR17, recA1, endA1, gyrA46, thi, relA1, lac⁻, F' [proAB⁺, lacI^a, lacZΔM15, Tn10(tet^r)]</i>	(19)
JM109	<i>recA1, supE44, endA1, hsdR17, gyrA96, relA1, thi, Δ(lac-proAB), F' [traD36, proAB⁺, lacI^a, lacZΔM15]</i>	(20)
Phage		
ϕ K	Wild-type	Lab. stock
ϕ KhT	Host-range mutant of ϕ K	This study
	Wild-type	Lab. stock
M13mp18	M13-derived vector for sequencing	(21)
M13mp19	M13-derived vector for sequencing	(21)
M13KO7	M13-derived helper phage for sequencing	(22)
Plasmid		
pUC18	2.7-kb, <i>E. coli</i> vector, Ap ^r , <i>lacZ</i>	(21)
pUC118	3.2-kb, <i>E. coli</i> vector, Ap ^r , <i>lacZ</i>	(21)
pUC19	2.7-kb, <i>E. coli</i> vector, Ap ^r , <i>lacZ</i>	(21)
pUC119	3.2-kb, <i>E. coli</i> vector, Ap ^r , <i>lacZ</i>	(21)

LPS, lipopolysaccharide.

summarized in Table I.

Enzymes and Biochemicals—Restriction enzymes, phage T4 DNA ligase, *E. coli* DNA polymerase I (Klenow fragment), and reagents used for dideoxy sequencing were purchased from CalBiochem. (USA), New England Biolabs (USA), Takara Shuzo (Kyoto), and Nippon Gene (Toyama). Buffers for the enzymes were as recommended by the manufacturers. [α -³⁵S]dCTP was from NEN (USA). All other materials used in this study were prepared as described previously (6, 13, 14).

DNA Sequencing Analysis—Cloning and sequencing of ϕ K DNA was carried out essentially as described by Kodaira *et al.* (6). Restriction enzyme fragments of the ϕ K RF DNA were cloned into sequencing vectors (M13mp18, M13mp19, pUC118, or pUC119) using *E. coli* strain XL1-blue or JM109 (Table I). SS and RF DNAs as a direct template for sequencing were prepared from these clones by standard methods (15). Sequencing was performed by the chain termination method (16) with universal and reverse vector primers.

Biological Assay—Biological properties of phages were examined as reported previously (13). For determination of heat stability, ϕ KhT (or ϕ K) propagated on *E. coli* C (or W3110) were incubated at 60°C in dilution fluid (3), and the surviving fractions were periodically titrated, using C or W3110 as the indicator. For eclipse assay, cells of *E. coli* C (or W3110) exponentially growing at 37°C in LB broth (A_{550} : 0.25) were mixed with chloramphenicol (50 μ g/ml). After 10 min, phage was added (input multiplicity: 0.1) with CaCl₂ (final concentration: 5 mM). At intervals, an aliquot was removed, then treated with chloroform, and the remaining free phage was titrated on host strain C or W3110.

Chemical Synthesis of Protein—Gene J protein of ϕ K was chemically synthesized using an Applied Biosystems Peptide Synthesizer (Model 430A) with a standard solid-phase peptide technique, and then purified by high-performance liquid chromatography. The 23-aa alignment synthesized in this study was KKARRSPSRKRGAR-LWYVGGSSQF (Fig. 3A). The synthetic protein is soluble in H₂O or Tris-HCl buffer (pH 7.3), and can be stored stably at 4°C.

RESULTS AND DISCUSSION

Properties of the Host-Range Mutant ϕ KhT— ϕ K infects *E. coli* K12 derivatives but not strain C, differing in host-range from the related phages ϕ X174, G4, and α 3. Intact particles of ϕ K were unable to be adsorbed onto cells of *E. coli* C, but the DNAs (SS and RF) efficiently transfect this strain upon Ca²⁺-treatment or electroporation (data not shown). A host-range mutant ϕ KhT was spontaneously obtained from ϕ K at approximate frequency of 10⁻⁸ on *E. coli* C. Unlike the parental strain, ϕ KhT can infect *E. coli* B besides C and K12 at almost the same efficiency; the rate of eclipse of ϕ KhT on K12 W3110 was identical to that of ϕ K (Fig. 1a). As to host-range, ϕ KhT resembles α 3 (3), and can be adsorbed onto LPS mutants of *E. coli* C (e.g., C61 and C23) and B (BB1 and BB5). The plating efficiencies of ϕ KhT on these LPS strains are shown in Table II, compared with those of ϕ K and α 3.

ϕ KhT is somewhat unstable; the phage particles were more sensitive to heat treatment at 60°C in dilution fluid (3) than those of ϕ K (Fig. 1b). The virion particles of ϕ KhT are composed of four major proteins F, G, H, and J. By SDS polyacrylamide gel electrophoresis (PAGE), their molecular weights (MW) were estimated to be approximately 52.3, 35.8, 21.8, and 2.8 kDa, respectively. These values are identical to those of the parent ϕ K (14).

Genome Structures of ϕ K and ϕ KhT—To identify structural genes of the four virion-forming proteins (F to J), the total genomes of ϕ K and ϕ KhT were sequenced. After establishment of restriction enzyme cleavage maps (data not shown), appropriate restriction fragments of ϕ K (ϕ KhT) were cloned into the sequencing vector using *E. coli* strain XL1-blue or JM109 (Table I). SS and RF DNAs were prepared from the resulting recombinant clones, and their sequences were determined from both strands of the viral and complementary DNA by the chain termination method (16).

The circular SS DNA of ϕ K (ϕ KhT) consists of 6,089 nt (EMBL Data Library accession number X60323), and is 2, 53, and 703 nt longer than the genome of α 3 (6), G4 (5), and ϕ X174 (4), respectively. As in the three phages, ϕ K has eleven homologous genes (A, A*, B, C, K, D, E, J, F, G, and H), each of which is encoded in the viral strand DNA, preceded by a possible ribosome-binding sequence (17). The gene organization of ϕ K thus established is illustrated in Fig. 2, compared with that of α 3, ϕ X174, and G4. As a whole, ϕ K shows an extensive sequence homology (89.3%) with α 3, and moderately resembles ϕ X174 and G4 (approximately 50% homologies). In Fig. 3, the deduced aa alignments of ϕ K virion-forming proteins, F (capsid), G (major spike), H (minor spike), and J (core), are shown in comparison with those of α 3, ϕ X174, and G4.

Properties of the Virion Proteins F, G, H, and J—Like

$\alpha 3$ (6), ϕK gene F codes for a 431-aa protein, which is 4 residues longer than that of $\phi X174$ (4) and G4 (5). The capsid (F) protein is highly conserved among the four microvirid species (Fig. 3B): the degree of aa similarity is 91.2% between ϕK and $\alpha 3$, 70.8% between ϕK and $\phi X174$, and 63.1% between ϕK and G4 (66.0% between $\phi X174$ and G4). Despite the similarity, the three phages $\alpha 3$, $\phi X174$, and G4 can not substitute ϕK gene F protein for their own protein *in vivo* (6).

Among the four virion-forming proteins, the major spike protein (G) is most divergent (Fig. 3C). The ϕK gene G protein consists of 187 aa, like that of $\alpha 3$ (6), and is 12 and 10 residues longer than that of $\phi X174$ (4) and G4 (5), respectively. Concerning aa arrangement, ϕK is closely related to $\alpha 3$ with 88.2% similarity, and is relatively remote from $\phi X174$ and G4 with 31.6 and 28.3% coincidence, respectively (42.0% between $\phi X174$ and G4). The ϕK gene G product functions in $\alpha 3$ *in vivo*, but not in $\phi X174$ or G4 (6).

ϕK codes for a 332-aa gene H protein, which is 2 and 4 residues longer than that of $\alpha 3$ (6) and $\phi X174$ (4), respectively, and 5 residues shorter than that of G4 (5). The aa conservation among the four minor spike proteins is

TABLE II. Plating efficiency of ϕK hT. Cells of *E. coli* were infected with ϕK hT, ϕK , or $\alpha 3$, and plated at 37°C. *E. coli* strains used were as indicated in Table I. Relative plaque yield is presented.

Strain	Phage		
	ϕK	ϕK hT	$\alpha 3$
<i>E. coli</i> K12 W3110	1.0	8.2×10^{-1}	$< 1.0 \times 10^{-5}$
<i>E. coli</i> C			
C	6.3×10^{-8}	1.0	1.0
C61	$< 1.0 \times 10^{-8}$	1.0	5.8×10^{-1}
C71	$< 1.0 \times 10^{-8}$	$< 1.0 \times 10^{-8}$	$< 1.0 \times 10^{-8}$
C23	1.1×10^{-8}	6.7×10^{-1}	7.6×10^{-1}
<i>E. coli</i> B			
BB	4.3×10^{-8}	2.3×10^{-1}	8.9×10^{-3}
BB1	1.1×10^{-7}	2.4×10^{-1}	$< 1.0 \times 10^{-8}$
BB5	$< 1.0 \times 10^{-8}$	3.0×10^{-1}	1.1×10^{-2}

as follows (Fig. 3D): 97.3% between ϕK and $\alpha 3$, 77.6% between ϕK and $\phi X174$, and 71.7% between ϕK and G4 (65.5% between $\phi X174$ and G4). Complementation of the gene H proteins has not been detected between ϕK ($\alpha 3$) and $\phi X174$ (G4) (6).

In order to elucidate the molecular basis of host-range determination, the total genome sequence of ϕK hT was compared with that of ϕK ; ϕK hT has two nt changes (Fig. 4). One is from G (ϕK) to T (ϕK hT), located at the 223rd nt in the gene G. Consequently, the 75th aa of the ϕK hT

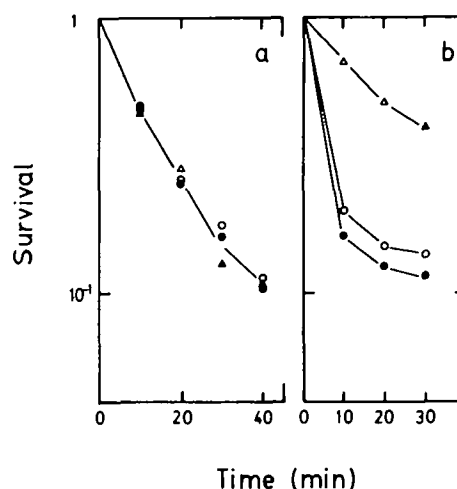


Fig. 1. Eclipse (a) and heat stability (b) of ϕK hT mutant. (a) *E. coli* C (or W3110) cells exponentially growing at 37°C in LB broth were mixed with chloramphenicol (50 μ g/ml). After 10 min, phage was added at a multiplicity of infection of 0.1, together with CaCl_2 (final concentration: 5 mM). At the indicated time, an aliquot was removed and treated with chloroform, and the remaining free phage was titrated on C (or W3110). \circ , ϕK hT on C; \bullet , ϕK hT on W3110; Δ , wt on W3110. (b) phage particles propagated on *E. coli* C were incubated at 60°C in dilution fluid and the surviving fractions were periodically determined, using C and K12 W3110 as the indicator. \circ , ϕK hT on C; \bullet , ϕK hT on W3110; Δ , wt on W3110.

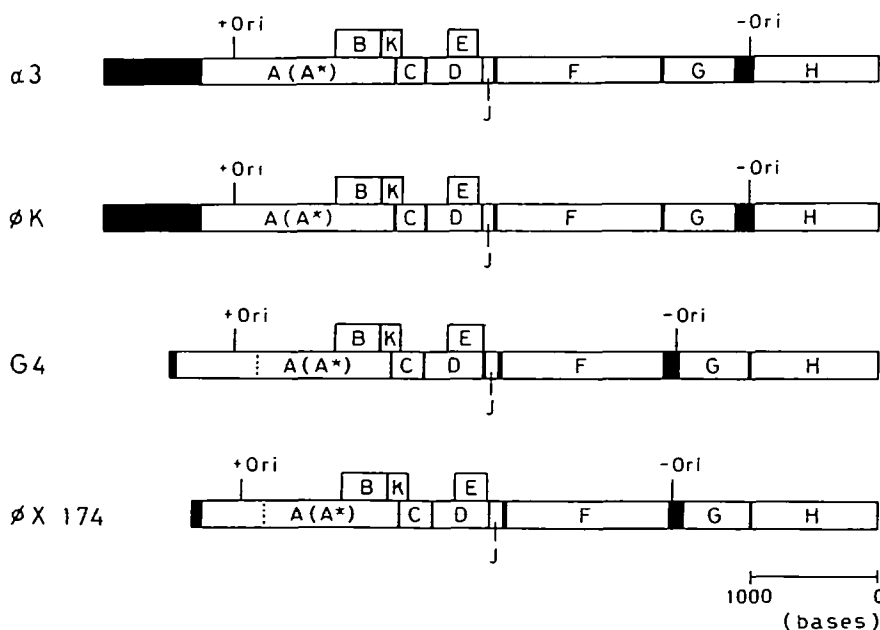


Fig. 2. Genetic maps of the genomes of $\alpha 3$, ϕK , G4, and $\phi X174$. The genetic maps of $\alpha 3$, ϕK , $\phi X174$, and G4 are drawn to scale from their nucleotide sequences: $\alpha 3$ from Kodaira *et al.* (6), ϕK from this study, G4 from Godson *et al.* (5), and $\phi X174$ from Sanger *et al.* (4). The two replication origins for viral and complementary strand are shown as +ori and -ori, respectively (6). The dotted lines in gene A of G4 (5) and $\phi X174$ (4) indicate the putative starting positions for gene A*.

(A)

ϕK	M	K	K	A-R	RS	---	P	---	SRR	KGARLWYVGG	S	Q	F
$\alpha 3$	M	K	K	A-R	RS	---	P	---	SRR	KGARLWYVGG	S	Q	F
ϕX	M	S	K	GKK	RS	G	A	R	PGRPQPLRGTKGR	KGARLWYVGG	S	Q	F
G 4	M	S	K	SIR	RS	G	---	---	GKS	KGARLWYVGG	S	Q	F
J-chen	M	K	K	A-R	RS	---	P	---	SRR	KGARLWYVGG	S	Q	F

(B)

ϕK	MSN	V	QT	S	A	E	R	EIV	DLSHL	A	F	DC	G	MI	GRL	K	T	V	S	W	TP	VI	AGDSFE	L	D	A	VGA	L	RLSPLRRGLA
$\alpha 3$	MSN	V	QT	S	A	E	R	EIV	DLSHL	A	F	DC	G	MI	GRL	K	T	V	S	W	TP	VI	AGDSFE	L	D	A	VGA	L	RLSPLRRGLA
ϕX	MSN	V	QT	S	A	E	R	EIV	DLSHL	A	F	DC	G	MI	GRL	K	T	V	S	W	TP	VI	AGDSFE	L	D	A	VGA	L	RLSPLRRGLA
G 4	MSN	V	QT	S	A	E	R	EIV	DLSHL	A	F	DC	G	MI	GRL	K	T	V	S	W	TP	VI	AGDSFE	L	D	A	VGA	L	RLSPLRRGLA

ϕK	I	DS	K	VD	F	F	T	F	I	PHRH	Y	YG	D	QWI	Q	FM	R	DGV	D	A	S	PLP	S	V	-	T	--TKYP	D	D	A	GVV	GTI	V
$\alpha 3$	I	DS	K	VD	F	F	T	F	I	PHRH	Y	YG	D	QWI	Q	FM	R	DGV	D	A	S	PLP	S	V	-	T	--TKYP	D	D	A	GVV	GTI	V
ϕX	I	DS	K	VD	F	F	T	F	I	PHRH	Y	YG	D	QWI	Q	FM	R	DGV	D	A	S	PLP	S	V	-	T	--TKYP	D	D	A	GVV	GTI	V
G 4	I	DS	K	VD	F	F	T	F	I	PHRH	Y	YG	D	QWI	Q	FM	R	DGV	D	A	S	PLP	S	V	-	T	--TKYP	D	D	A	GVV	GTI	V

ϕK	P	K-SNR	PK	F	L	H	Q	S	YLNINNYF	R	APW	MPER	T	E	ANP	S	N	---	LDR	D	DSR	Y	--GF	R	CCH	LK	T	I
$\alpha 3$	P	A-NNR	PK	F	L	H	Q	S	YLNINNYF	R	APW	MPER	T	E	ANP	S	N	---	INE	D	DAR	Y	--RF	R	CCH	LK	N	I
ϕX	P	D-TNK	PK	F	L	H	Q	S	YLNINNYF	R	APW	MPER	T	E	ANP	S	N	---	ELNQ	D	DAR	Y	--GF	R	CCH	LK	N	I
G 4	P	SSTLK	PK	F	L	H	Q	S	YLNINNYF	R	PPW	SDDL	T	Y	ANP	S	N	---	NPSE	D	DAR	Y	--GF	R	CCH	LK	N	I

ϕK	V	S	APLPP	E	T	KLAEQ	M	GIESN	S	-	IDINGLQAAYA	Q	LHT	E	QER	T	YFM	Q	RY	R	D	Y	I	SS-	FGG	S	TSY
$\alpha 3$	V	S	APLPP	E	T	KLAEE	M	GIESN	S	-	IDINGLQAAYA	Q	LHT	E	QER	T	YFM	Q	RY	R	D	Y	I	SS-	FGG	S	TSY
ϕX	V	S	APLPP	E	T	ELSRQ	M	TT-STT	S	-	IDINGLQAAYA	N	LHT	D	QER	D	YFM	Q	RY	R	D	Y	I	SS-	FGG	K	TSY
G 4	V	S	APLPP	E	T	RTSEN	M	TT-GT	S	T	IDINGLQAAYA	K	LHT	E	QER	D	YFM	Q	RY	R	D	Y	I	SS-	FGG	H	TSY

ϕK	D	A	DNRPLL	V	M	HTDF	WASGYDVGTDQ	S	SLGQFSGRVQQT	FK	H	S	VPRF	F	VPEH	GV-	M	N	TL	MLV-	RFPP	IS
$\alpha 3$	D	A	DNRPLL	V	M	HTDF	WASGYDVGTDQ	S	SLGQFSGRVQQT	FK	H	S	VPRF	F	VPEH	GV-	M	N	TL	MLV-	RFPP	IS
ϕX	D	A	DNRPLL	V	M	HTDF	WASGYDVGTDQ	S	SLGQFSGRVQQT	FK	H	S	VPRF	F	VPEH	GV-	M	N	TL	MLV-	RFPP	IS
G 4	D	A	DNRPLL	V	M	HTDF	WASGYDVGTDQ	S	SLGQFSGRVQQT	FK	H	S	VPRF	F	VPEH	GV-	M	N	TL	MLV-	RFPP	IS

ϕK	PL	E	HH	YL	-VGRN	LTYTD	L	A	G	DP	A	L	IG	NLPPRE	I	S	YQDL	F	RDG-RPGI	K	I	K	V	AE	SI	WYR	TH	P
$\alpha 3$	PL	E	HH	YL	-AGKSQ	LTYTD	L	A	G	DP	A	L	IG	NLPPRE	I	S	YQDL	F	RDG-RSGI	K	I	K	V	AE	SI	WYR	TH	P
ϕX	TK	E	IQ	YL	NA-KGA	LTYTD	L	A	G	DP	V	L	YG	NLPPRE	I	S	MKDV	F	RSQ-DSSK	K	F	K	I	AE	GQ	WYR	YA	P
G 4	EM	E	MH	YL	VG-KEN	LTYTD	L	A	C	DP	A	L	MA	NLPPRE	V	S	LKEF	F	HSSPDSA-	K	F	K	I	AE	GQ	WYR	TQ	P

ϕK	DY	V	N-YK	Y	QL	L	E	GFPP	LDDA	P	GTT	S	-GDD	L	QKAI	L	IDHND	Y	NAC	FQS	Q	QL	LQ	WN	K	Q	ARY	N	VN
$\alpha 3$	DY	V	N-FK	Y	HD	L	E	GFPP	LDDA	P	GT	S	-GDD	L	QKAI	L	IDHND	Y	NAC	FQS	Q	QL	LQ	WN	K	Q	ARY	N	VN
ϕX	SY	V	S-PA	Y	HL	L	E	GFPP	LDDA	P	---	S	-GDD	L	QKAI	L	IDHND	Y	NAC	FQS	Q	QL	LQ	WN	K	Q	ARY	N	VN
G 4	DR	V	AFP-	Y	NA	L	D	GFPP	YSAL	P	---	S	-TE-	L	KDRV	L	VNTN	Y	DEI	FQS	Q	QL	LQ	WN	K	Q	TKF	N	IN

ϕK	VYR	HI	PT	V	RDSIMTS
$\alpha 3$	VYR	HM	PT	V	RDSIMTS
ϕX	VYR	NL	PT	T	RDSIMTS
G 4	VYR	HM	PT	T	RDSIMTS

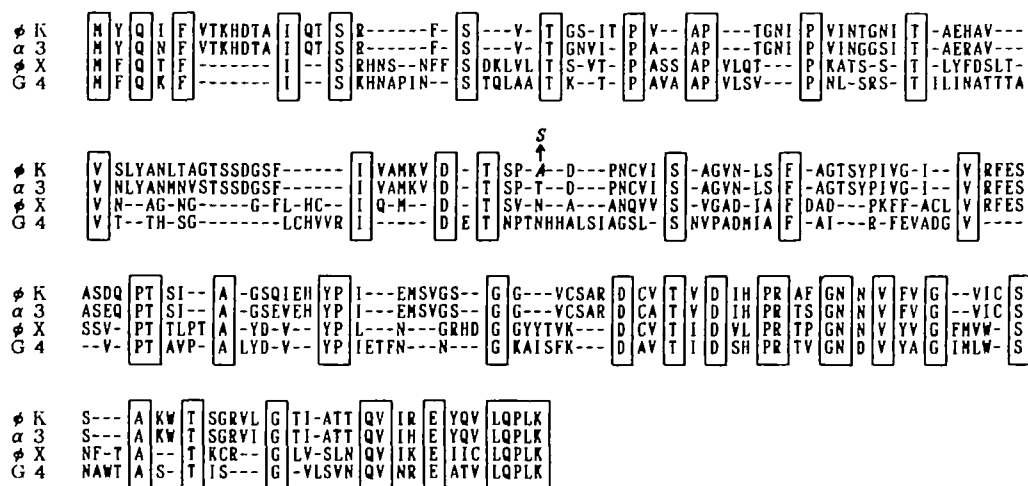
Fig. 3 (Continued on next page)

minor spike protein is altered to Ser from Ala in ϕK ; the other one, from T (ϕK) to C (ϕKhT), is at the 200th nt in the gene H: the aa replacement of minor spike protein (H) is from Val (ϕK) to Cys (ϕKhT). In the capsid (F) and core (J) proteins, ϕKhT has no such aa changes. These sequencing results clearly demonstrate that ϕKhT is a double mutant bearing two missense mutations in genes G and H, suggesting that, at least in ϕK , the two spike proteins G and H are involved in the host-range determination. In $\alpha 3$ as well (13), comparable results have been obtained; three missense strains of gene H (from Ile to Val at 56th aa, from Gly to Cys at 69th aa, and from Leu to Phe at 71st aa) showed altered heat sensitivities and eclipse rates. These facts suggest that the N-terminal domain of the gene H protein (see Fig. 3D) is important for functional and/or

structural interaction among virion proteins F, G, H, and J (see below). Certain host-range mutants of $\phi X174$ also have alterations in genes G and H (18).

In order to discern the contributions of gene G and H to the altered host range and heat stability, phage strains having single missense mutations were constructed by exchanging a *PpuM1* restriction fragment of 876 bp (from nucleotide 4059 to 4934, containing the mutated region of gene G), between ϕKhT and wild-type ϕK . Thus, each RFI DNA was cut into two fragments and purified by repeated agarose gel electrophoresis. Ligation of the wild-type large fragment with the ϕKhT G segment resulted in an RFI molecule having a single missense G mutation. Conversely, ligation of the ϕKhT large fragment with the wild-type G segment yielded an RF molecule with a single missense H

(C)



(D)

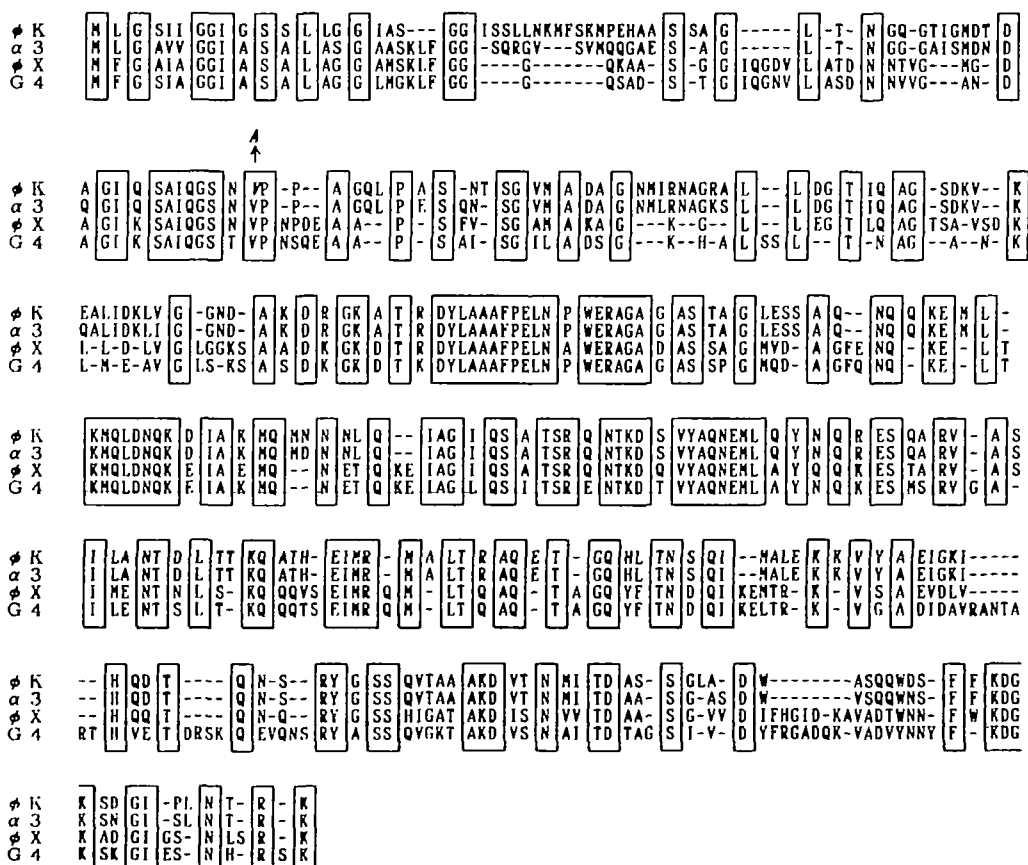


Fig. 3. Comparisons of the ϕ K virion proteins J, F, G, and H with those of α 3, ϕ X174, and G4. Amino acid sequences of ϕ K, α 3, ϕ X174, and G4 are from this study, Kodaira *et al.* (6), Sanger *et al.* (4), and Godson *et al.* (5), respectively: A, core (J) protein; B, capsid (F) protein; C, major spike (G) protein; D, minor spike (H) protein.

Dashes represent gaps inserted to optimize the protein alignment. Conserved aa residues through the four phage species are boxed. The aa alignment of synthetic gene J protein (J-chem, see text) is shown in (A). Amino acid changes found in the G and H proteins of ϕ KhT are indicated by \uparrow .

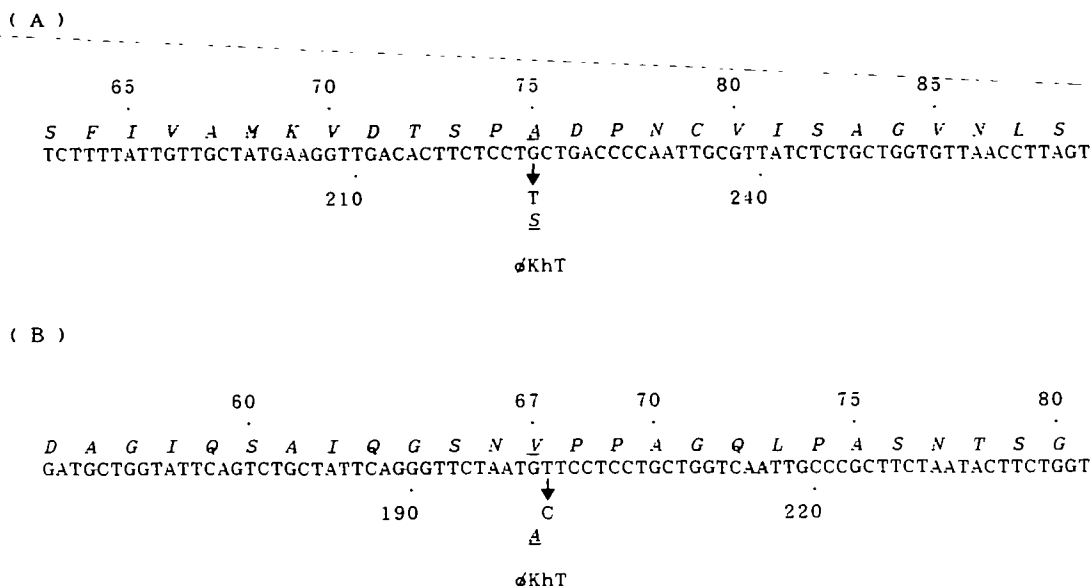


Fig. 4. Nucleotide sequences of the genes G and H of ϕK and $\phi K h T$. Relevant parts of nucleotide and amino acid sequences are presented. The nucleotide numbers of the genes G (A) and H (B) are from their starting codons, ATG. The sites of the changed nucleotides and amino acids in $\phi K h T$ are indicated by the arrows (\downarrow).

mutation. After Ca^{2+} -dependent transfection and single plaque isolation, the properties of the missense G (msG) strain and the missense H (msH) were tested.

In *E. coli* C and B, plating efficiency of the msG strain was, like that of the wild-type ϕK , negligibly low, whereas the msH mutant infected *E. coli* C at the relative efficiency of 10^{-3} of K12 W3110, and easily reached the level of $\phi K h T$ by further propagation on C cells. As to heat stability at $60^\circ C$, the msG strain was indistinguishable from wild-type ϕK . The heat inactivation profile of the msH mutant was complex: an initial slow decline was followed by a rapid loss of infectivity, suggesting that the msH phage preparation is somewhat heterogenous. When the msH mutant was repeatedly propagated on *E. coli* C, the progeny exhibited a heat inactivation profile similar to that of $\phi K h T$. [As can be seen in Fig. 1b, the $\phi K h T$ preparation contained a minor fraction relatively resistant to the heat treatment. The ratio of this fraction increased during storage at $4^\circ C$, accompanied with gradual loss of bulk infectivity. The nature of this change (conformational?) is presently unknown.]

These results suggest that plaques of ϕK initially formed on *E. coli* C were relatively heat-stable msH type with lower plating efficiency on this host, and $\phi K h T$, which is the msH msG double mutant thermosensitive and highly infective to C, was derived from the msH strain during propagation on C. At any rate, cells of *E. coli* C and B were rather insensitive to the single mutants of msH or msG type.

Like the three phages ($\alpha 3$, $\phi X174$, and G4), the ϕK gene J codes for a small and highly basic core protein of 24-aa identical to that of $\alpha 3$ (11), which is shorter than that of $\phi X174$ (4) composed of 38 aa and G4 of 25 aa (5). The four gene J proteins are structurally similar (Fig. 3D): (i) a basic N-terminal half rich in Lys and Arg (albeit diverged in aa sequence), (ii) a conspicuously conserved C-terminal half, whose identical decapeptide KGARLWYVGG is thought to

be a DNA-binding domain (R. McKenna *et al.*, personal communication), and (iii) a C-terminal aromatic residue (Phe in ϕK , $\alpha 3$, and $\phi X174$; Tyr in G4). As reported previously (6), $\alpha 3$ can use the ϕK gene J product *in vivo*, but not that of G4 or $\phi X174$, and *vice versa*.

These results on the four virion proteins (F to J) indicate that ϕK and $\alpha 3$ are functionally related, and remote from $\phi X174$ and G4, whereas $\phi X174$ and G4 are evolutionally closer.

DNA-Binding Properties of the Gene J Protein—The detailed structure and function of the gene J protein are still unknown, because of difficulty in preparation of the gene J protein from phage particles in sufficient amounts for analysis *in vitro*.

To investigate DNA-binding properties, the gene J protein of ϕK was chemically synthesized as described in "MATERIALS AND METHODS." The synthetic protein (termed J-chem) is 23 aa long, with the sequence shown in Fig. 3A; in the present synthesis, the N-terminal residue Met was omitted, according to a previous report (4) that the gene J protein isolated from $\phi X174$ particle was devoid of the N-terminal residue. The J-chem protein with a calculated MW of 2.7 kDa is soluble in H_2O as well in Tris-based buffers (e.g., 50 mM Tris-HCl, pH 7.3), and shows a random structural circular dichroism profile with no significant bands. Its electrophoretic mobility on a 17% SDS-polyacrylamide gel is the same as that of the natural gene J protein isolated from ϕK particle (data not shown).

Binding of the J-chem protein to ϕK circular SS DNA was examined by gel retardation analysis. SS DNA (0.2 μg) was mixed with various amounts of J-chem ranging from 0 to 0.2 μg in a reaction buffer (50 mM Tris-HCl, pH 7.3), incubated at $25^\circ C$ for 30 min, and then electrophoresed on a 1% agarose gel. As shown in Fig. 5 (lanes 1-4), SS DNA incubated with J-chem moved in front of the naked SS DNA, faster than expected from the mass of the complex, and its electrophoretic mobility was accelerated with

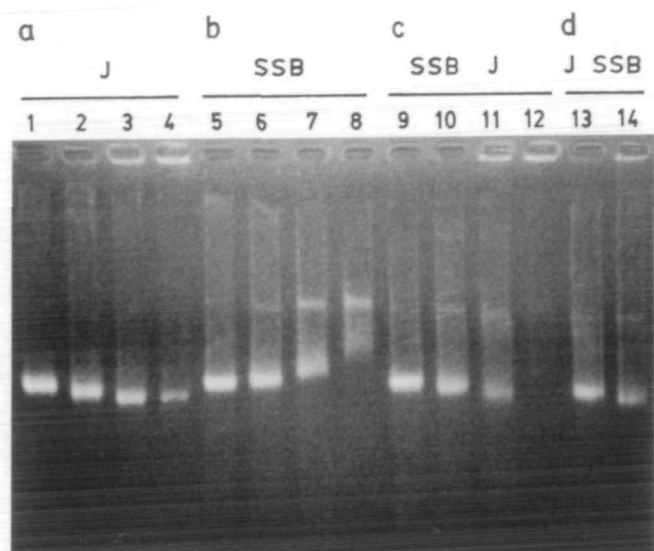


Fig. 5 Characterization of *in vitro* interaction between chemically synthesized gene J protein (J-chem) and ϕ K DNA. a: ϕ K SS DNA (0.2 μ g) was mixed with J-chem (lane 1, 0 μ g; lane 2, 0.05 μ g; lane 3, 0.1 μ g; lane 4, 0.2 μ g), incubated at 25°C for 30 min, and each mixture was subsequently electrophoresed on a 1% agarose gel. b: ϕ K SS DNA (0.2 μ g) was mixed with SSB (lane 5, 0 μ g; lane 6, 0.1 μ g; lane 7, 0.2 μ g; lane 8, 0.3 μ g) and analyzed as in (a). c: ϕ K SS DNA (0.2 μ g) was incubated with SSB (0.1 μ g) for 30 min at 25°C, and then J-chem (lane 9, 0 μ g; lane 10, 0.05 μ g; lane 11, 0.1 μ g; lane 12, 0.2 μ g) was added. After further incubation for 30 min, each mixture was analyzed as in (a). d: ϕ K SS DNA (0.2 μ g) was incubated with J-chem proteins (lane 13, 0.05 μ g; lane 14, 0.1 μ g) for 30 min at 25°C, and then SSB (0.1 μ g) was added. After further incubation for 30 min, each mixture was analyzed as in (a).

increasing J-chem concentration. Preliminary studies with fluorescence spectroscopy have shown that an SS DNA molecule of ϕ K seems to be saturated by the J-chem proteins at a weight ratio (J-chem/SS DNA) of 1.0, that is, approximately 600 molecules of J-chem cover one SS DNA. In addition, J-chem bound to ϕ K RF DNA tightly, but the complex moved more slowly than the naked RF DNA, as in *E. coli* SSB (data not shown). These results clearly indicate that the gene J proteins bind to both of SS and RF DNAs, and selectively condense SS DNA, but not RF DNA, into a compact form. Hamatake *et al.* (12) reported that gene J proteins purified from ϕ X174 particles have an ability to bind to RF as well as to SS DNA.

Binding of SSB to ϕ K SS DNA was examined under the same conditions as used for J-chem. SS DNA (0.2 μ g) in the presence of SSB (ranging from 0 to 0.3 μ g) moved more slowly than the naked SS DNA, and its mobility decreased with increasing SSB (Fig. 5, lanes 5-8), although two complexes appeared: their conformations are presently unknown. These results suggest that SSB binds tightly to SS DNA, but the complex is not so compact as to move faster than the naked SS DNA on the gel. To clarify whether SSB is replaced by gene J protein in morphogenesis, competition experiments between J-chem and SSB were performed. ϕ K SS DNA (0.2 μ g) was incubated with SSB (0.1 μ g) in the reaction buffer at 25°C for 30 min, and then J-chem (from 0.05 to 0.2 μ g) was added. After further incubation for 30 min, the mixtures were electrophoresed on a 1% agarose gel. As presented in Fig. 5 (lanes 9-11),

TABLE III Infectivity of ϕ K DNA treated with J-chem or SSB. Single- or double-stranded DNA of ϕ K was mixed with synthetic gene J protein (J-chem) or SSB in 50 mM Tris-HCl (pH 7.3), incubated at 25°C for 30 min, and transfected to Ca^{2+} -treated *E. coli* K12 W3110 (7).

Protein	(μ g)	Relative titer	
		SS-DNA (0.2 μ g)	DS-DNA (0.4 μ g)
None		1.0	1.0
J-chem	0.1	2.4×10^{-2}	1.0
	0.2	5.0×10^{-3}	1.0
	0.3	$< 1.0 \times 10^{-3}$	1.0
	0.4	—	1.0
	0.5	—	1.0
SSB	0.1	5.5×10^{-1}	1.0
	0.2	1.7×10^{-1}	1.0
	0.3	7.0×10^{-2}	1.0
	0.4	$< 3.0 \times 10^{-3}$	1.0
	0.5	—	1.0

SSB associated with SS DNA is substantially replaced by J-chem.

Infectivity of SS or RF DNA complexed with J-chem was investigated by Ca^{2+} -dependent transfection assay (7). As shown in Table III, infectivity of ϕ K SS DNA was markedly reduced in the presence of J-chem as well as SSB, although the latter was less inhibitory than J-chem, whereas transfecting activity of ϕ K RF DNA was not prevented by J-chem or SSB.

The present *in vivo* and *in vitro* observations thus strongly indicate that ϕ K nascent SS DNA preferentially binds to gene J (core) protein, replacing host SSB, and hence is packaged into a virion vesicle composed of three gene proteins F (capsid), G (major spike), and H (minor spike). It is also evident that two spike proteins G and H determine the host-range of ϕ K, through still unidentified protein-protein and protein-LPS interaction(s). Taking advantage of the ϕ K host-range mutant and synthetic gene J protein, further studies on structural and functional interactions among the virion-forming proteins are in progress.

We thank Drs S. Aimoto and S. Ono for chemical synthesis of J-chem, and Dr. N. Shimamoto for a gift of *E. coli* SSB.

REFERENCES

- Kornberg, A. (1980) *DNA Replication*, WH Freeman, San Francisco
- Baas, P.D. and Jansz, H.S. (1988) Single-stranded DNA phage origins *Curr. Top. Microbiol. Immunol.* **136**, 31-70
- Taketo, A. and Kodaira, K.-I. (1978) Host-factor requirements and some properties of α 3 and related phages in *The Single-Stranded DNA Phages* (Denhardt, D.T., Dressler, D., and Ray, D.S., eds.) pp. 361-367, Cold Spring Harbor Laboratory, New York
- Sanger, F., Air, G.M., Barrell, B.G., Brown, N.L., Coulson, A.R., Fiddes, J.C., Hutchinson, C.A., III, Slocombe, P.M., and Smith, M. (1977) Nucleotide sequence of bacteriophage ϕ X174 DNA. *Nature* **265**, 687-695
- Godson, G.N., Fiddes, J.C., Barrell, B.G., and Sanger, F. (1978) Comparative DNA sequence analysis of the G4 and ϕ X174 genomes in *The Single-Stranded DNA Phages* (Denhardt, D.T., Dressler, D., and Ray, D.S., eds.) pp. 51-86, Cold Spring Harbor Laboratory, New York
- Kodaira, K.-I., Nakano, K., Okada, S., and Taketo, A. (1992) Nucleotide sequence of the genome of the bacteriophage α 3: Interrelationship of the genome structure and the gene products

- with those of the pages, $\phi X174$, G4, and ϕK . *Biochim. Biophys. Acta* **1130**, 277-288
7. Taketo, A. and Kodaira, K.-I. (1978) Growth of single-stranded DNA phages in replication mutants of *Escherichia coli*. *Mol. Gen. Genet.* **162**, 151-155
 8. Sims, J., Capon, D., and Dressler, D. (1979) *dnaG* (primase)-dependent origins of DNA replication: Nucleotide sequences of the negative strand initiation sites of bacteriophages St-1, ϕK , and $\alpha 3$. *J. Biol. Chem.* **254**, 12615-12628
 9. Kodaira, K.-I., Godson, G.N., and Taketo, A. (1995) Comparative studies of the minus origin mutants of *Escherichia coli* spherical single-stranded DNA phages. *Biochim. Biophys. Acta* **1260**, 191-199
 10. Meyer, R. and Laine, P. (1990) The single-stranded DNA-binding protein of *Escherichia coli*. *Microbiol. Rev.* **54**, 342-380
 11. Kodaira, K.-I. and Taketo, A. (1984) Isolation and some properties of bacteriophage $\alpha 3$ gene J mutant. *Mol. Gen. Genet.* **195**, 541-543
 12. Hamatake, R.K., Aoyama, A., and Hayashi, M. (1985) The J gene of bacteriophage $\phi X174$: *In vitro* analysis of J protein function. *J. Virol.* **54**, 345-350
 13. Kodaira, K.-I., Nakano, K., and Taketo, A. (1985) Function and structure of microvirid phage $\alpha 3$ genome. DNA sequence of H gene and properties of missense H mutant. *Biochim. Biophys. Acta* **825**, 255-260
 14. Kodaira, K.-I. and Taketo, A. (1984) Function and structure of microvirid phage $\alpha 3$ genome I. Electrophoretic characterization of proteins encoded by wild-type phage. *Biochim. Biophys. Acta* **788**, 333-338
 15. Sambrook, J., Fritsch, E.F., and Maniatis, T. (1989) *Molecular Cloning: A Laboratory Manual*, Cold Spring Harbor Laboratory, Cold Spring Harbor, NY
 16. Sanger, F., Nicklen, S., and Coulson, A.R. (1977) DNA sequencing with chain-terminating inhibitors. *Proc. Natl. Acad. Sci. USA* **74**, 5463-5467
 17. Shine, J. and Dalgarno, L. (1974) The 3'-terminal sequence of *Escherichia coli* 16S ribosomal RNA: Complementary to non-sense triplets and ribosome binding sites. *Proc. Natl. Acad. Sci. USA* **71**, 1342-1346
 18. Tessman, E.S. and Tessman, I. (1978) The genes of the isometric phages and their functions in *The Single-Stranded DNA Phages* (Denhardt, D.T., Dressler, D., and Ray, D.S., eds.) pp. 9-29, Cold Spring Harbor Laboratory, New York
 19. Bullock, W.O., Fernandez, J.M., and Short, J.M. (1987) XL1-blue: A high efficiency plasmid transforming *recA* *Escherichia coli* strain with beta-galactosidase selection. *Biotechnology* **5**, 376-378
 20. Yanish-Perron, C., Vieira, J., and Messing, J. (1985) Improved M13 phage cloning vectors and host strains: Nucleotide sequences of the M13mp18 and pUC19 vectors. *Gene* **33**, 103-119
 21. Messing, J. (1983) New M13 vectors for cloning. *Methods Enzymol.* **101**, 20-78
 22. Vieira, J. and Messing, J. (1987) Production of single-stranded plasmid DNA. *Methods Enzymol.* **153**, 3-11