Yersiniophage ϕ R1-37 is a tailed bacteriophage having a 270 kb DNA genome with thymidine replaced by deoxyuridine

Saija Kiljunen,^{1,4} Kristo Hakala,^{2,3} Elise Pinta,^{1,4} Suvi Huttunen,^{1†} Patrycja Pluta,¹ Aneta Gador,¹ Harri Lönnberg² and Mikael Skurnik^{1,4}

Correspondence

Saija Kiljunen

saija.kiljunen@utu.fi

Department of Medical Biochemistry and Molecular Biology, Institute of Biomedicine¹, Department of Chemistry² and Department of Pharmacology and Clinical Pharmacology, Institute of Biomedicine³, University of Turku, Turku, Finland

⁴Department of Bacteriology and Immunology, Haartman Institute, University of Helsinki and Helsinki University Central Hospital Laboratory, Helsinki, Finland

Bacteriophage ϕ R1-37 was isolated based on its ability to infect strain YeO3-R1, a virulence-plasmid-cured O antigen-negative derivative of *Yersinia enterocolitica* serotype O:3. In this study, the phage receptor was found to be a structure in the outer core hexasaccharide of *Y. enterocolitica* O:3 LPS. The phage receptor was present in the outer core of strains of many other *Y. enterocolitica* serotypes, but also in some *Yersinia intermedia* strains. Surprisingly, the receptor structure resided in the O antigen of *Yersinia pseudotuberculosis* O:9. Electron microscopy demonstrated that ϕ R1-37 particles have an icosahedral head of 88 nm, a short neck of 10 nm, a long contractile tail of 236 nm, and tail fibres of at least 86 nm. This implies that the phage belongs to the order *Caudovirales* and the family *Myoviridae* in the ICTV (International Committee for Taxonomy of Viruses) classification. ϕ R1-37 was found to have a lytic life cycle, with eclipse and latent periods of 40 and 50 min, respectively, and a burst size of ~80 p.f.u. per infected cell. Restriction digestions and PFGE showed that the ϕ R1-37 genome was dsDNA and ~270 kb in size. Enzymically hydrolysed DNA was subjected to HPLC-MS/MS analysis, which demonstrated that the ϕ R1-37 genome is composed of DNA in which thymidine (T) is >99% replaced by deoxyuridine (dU). The only organisms known to have similar DNA are the *Bacillus subtilis*-specific bacteriophages PBS1 and PBS2. N-terminal amino acid sequences of four major structural proteins did not show any similarity to (viral) protein sequences in databases, indicating that close relatives of ϕ R1-37 have not yet been characterized. Genes for two of the structural proteins, p24 and p46, were identified from the partially sequenced ϕ R1-37 genome.

Received 14 June 2005

Revised 23 August 2005

Accepted 1 September 2005

INTRODUCTION

Yersinia enterocolitica is a Gram-negative bacterium of the family *Enterobacteriaceae*. The species includes over 30 serotypes, a number of which are human pathogens (Wauters *et al.*, 1991). While the major pathogenic serotypes in Europe, Canada, Japan and South Africa are O:3 and O:9; in the United States serotype O:8 is more prevalent. *Y. enterocolitica* is widely distributed in nature, swine being

the major reservoir of the pathogenic strains (Bottone, 1997).

Bacteriophages are the most abundant organisms on Earth, and it is estimated that for each microbial isolate at least 10 different phages can be found (Hendrix, 2002; Pedulla *et al.*, 2003). A number of phages infecting the members of the genus *Yersinia* have been isolated (Baker & Farmer, 1982; Popp *et al.*, 2000; Stevenson & Airdrie, 1984), but relatively few have been characterized in more detail. *Y. enterocolitica* O:3-specific phage ϕ YeO3-12 has been shown to be closely related to *Escherichia coli* phages T3 and T7 (Pajunen *et al.*, 2000, 2001). Likewise, *Yersinia pestis* phage ϕ A1122 was recently found to be a close relative of T7 (Garcia *et al.*, 2003a). Bacteriophage PY54 is a temperate phage isolated from *Y. enterocolitica* O:5 that maintains its lysogeny by replicating as a linear plasmid with covalently closed ends (Hertwig *et al.*, 2003a, b).

[†]Present address: Department of Virology, University of Turku, Tykistokatu 6 A, 20520 Turku, Finland.

Abbreviations: BER, base-excision repair pathway; DOC, deoxycholate; ICTV, International Committee for Taxonomy of Viruses; RT, room temperature; Ugi, uracil-DNA glycosylase inhibitor; Ung, uracil-DNA glycosylase.

The GenBank/EMBL/DDBJ accession numbers for the sequences reported in this paper are AJ972879 and AJ972880.

Recent interest in bacteriophages is based on the facts that (i) they are excellent targets for genome and evolution research, (ii) they are important vehicles in horizontal gene transfer, (iii) they are potential therapeutic agents, and (iv) they can themselves be used as tools in bacterial genetics and their gene products are used as tools in molecular biology. We have isolated several *Y. enterocolitica*-specific bacteriophages that use different parts of LPS as receptors and used them to study the molecular biology and genetics of LPS biosynthesis (Al-Hendy *et al.*, 1991; Skurnik *et al.*, 1995).

Bacteriophage ϕ R1-37 was isolated from sewage, based on its ability to infect strain YeO3-R1, which is an O antigen-negative derivative of *Y. enterocolitica* serotype O:3 strain 6471/76-c (=YeO3-c), a virulence-plasmid-cured derivative of wild-type strain 6471/76 (=YeO3) (Skurnik, 1984; Skurnik *et al.*, 1995). The host range of ϕ R1-37 and genetic data suggested that the LPS outer core hexasaccharide of *Y. enterocolitica* O:3 may function as the phage receptor (Skurnik *et al.*, 1995). In this work we characterized further the biological, structural and genomic features of ϕ R1-37. We also confirmed the role of YeO3 LPS outer core as the phage receptor, and provided evidence showing that this receptor structure resides in the O antigen of *Yersinia pseudotuberculosis* O:9.

METHODS

Bacterial strains, phages, plasmids and media. The bacterial strains used in this work are described in Table 1. All *Yersinia* and phage incubations were done at room temperature (RT); *E. coli* was cultured at 37 °C. Tryptic soy broth (TSB) medium (Oxoid) and Luria Broth (Sambrook & Russell, 2001) were used for bacterial liquid cultures. Soft-agar medium included additionally 0.4% (w/v) agar (Biokar Diagnostics). Luria agar (Sambrook & Russell, 2001) was used as solid medium for bacteria, and lambda agar (10 g Tryptone l⁻¹, 2.5 g NaCl l⁻¹, 15 g agar l⁻¹) for phage plates. Plates and liquid media were supplemented with ampicillin (100 µg ml⁻¹), kanamycin (100 µg ml⁻¹), nalidixic acid (100 µg ml⁻¹) or chloramphenicol (30 µg ml⁻¹), when required. Purification of ϕ R1-37 was done as described elsewhere for phage ϕ YeO3-12 (Pajunen *et al.*, 2000). The nalidixic acid-resistant derivative of *Y. pestis* D27 was obtained by culturing D27 on nalidixic acid-containing plates and selecting for resistant colonies.

ϕ R1-37-resistant mutants were isolated by spreading a phage suspension on a bacterial lawn growing on agar plates and picking colonies grown within the lysis zone after 48 h incubation. Analysis of the LPS pattern of the resistant isolates was performed by deoxycholate (DOC)-PAGE gel electrophoresis, as described previously (Skurnik *et al.*, 1995).

Electron microscopy. The phage was sedimented using a Beckman Optima LE-80K ultracentrifuge and Ti50 rotor at about 85 000 g for 60 min, and suspended in TM (50 mM Tris/HCl, pH 7.8, 10 mM MgSO₄) (Sambrook & Russell, 2001). The phage particles were placed on 0.3% FormWar-coated grids (Agar Scientific), stained with 1% phosphotungstic acid and examined with a JEOL (Tokyo, Japan) JEM-1200 EX electron microscope. For infection samples, pelleted YeO3-R1 cells were suspended in TM and incubated with ϕ R1-37 for 2 min prior to placing on grids, staining and examination, as above. The phage dimensions were measured on photographs at a final magnification of 136 500-fold.

One-step growth curve. A one-step growth curve was done as previously described (Birge, 1988; Pajunen *et al.*, 2000). Briefly, a mid-exponential-phase culture (10 ml) of YeO3-R1 was harvested by centrifugation and resuspended in 1 ml TSB. ϕ R1-37 was added at an m.o.i. of 5×10^{-4} and allowed to adsorb for 2 min. The infected cells were centrifuged, resuspended in 10 ml TSB, and incubation was continued. Samples were taken at 10 min intervals. The first set of samples was immediately plated for phage titration, and the second set was treated with 1% chloroform to release the intracellular phage prior to titration.

DIG-labelling of ϕ R1-37 and immunodetection. Phage particles were labelled using the DIG Protein Labelling Kit (Roche). A 50 µl sample of ϕ R1-37 ($\sim 10^{11}$ p.f.u.) was diluted in 950 µl 50 mM potassium phosphate buffer, pH 8.5, containing 0.95% NaCl and 10 mM MgSO₄. DIG-ester (0.4 mg) dissolved in 10 µl DMSO was added and the reaction was allowed to proceed for 2 h at RT. The mixture was then dialysed against 1 l TM at 4 °C for 3 days. This preparation is referred to as DIG- ϕ R1-37.

For phage-blotting experiments, YeO3-R1 was grown for 16 h at RT in TSB. The bacteria were diluted to OD₆₀₀ 0.9–1.0, centrifuged, and resuspended to the same volume of SM (50 mM Tris/HCl, pH 7.5, 0.1 M NaCl, 10 mM MgSO₄, 0.01% gelatin) (Sambrook & Russell, 2001). The suspension was boiled at 95 °C for 20 min. Whole-cell lysates were prepared by resuspending bacteria in SDS-PAGE sample buffer and incubating the samples at 95 °C for 10 min. Total bacterial membranes were prepared from 5 ml overnight culture by sonication and differential centrifugation. The final membrane pellets were resuspended in 50 µl TM. Appropriate dilutions of bacterial preparations were applied to nitrocellulose membranes (Schleicher & Schuell) as 2–5 µl drops and allowed to dry. The filters were baked for 15 min at 60 or 80 °C. To study the effect of different treatments on the phage binding, some of the membranes were incubated with periodate (2 h at 22 °C in 10 or 100 mM periodic acid, 50 mM sodium acetate, pH 5.2) or with proteinase K (0.5 mg ml⁻¹, 60 °C, 60 min). The membranes were washed once in 0.9% NaCl, and incubated at RT for 60 min in 10 ml 3% BSA in TM. Half of the solution was then removed, 50 µl DIG- ϕ R1-37 was added, and the membranes were incubated for an additional 90 min. The membranes were then washed four times (5 min each) with TM and incubated in 3 ml anti-DIG-POD (polyclonal sheep anti-DIG Fab fragments, conjugated with horseradish peroxidase, Roche Molecular Biochemicals) for 60 min, washed four times with TM, and developed for peroxidase activity according to the manufacturer's instructions.

Construction of plasmid pRV16NP. Plasmid pRV7 (Table 1) containing the *Y. enterocolitica* O3 outer core gene cluster was digested with *Bam*HI and *Hind*III. The fragment containing the outer core gene cluster was gel-purified and ligated with *Bam*HI- and *Sph*I-digested pTM100 (Table 1) and a *Hind*III-*Sph*I adapter oligonucleotide (5'-AGCTCATG-3' (Biedzka-Sarek *et al.*, 2005). The ligation mixture was then transformed into *E. coli* C600 (Table 1). The plasmid having the cloned gene cluster transcribed from the tetracycline-resistance gene promoter was named pRV16NP and mobilized to different *Yersinia* strains by triparental conjugation (Gerhardt *et al.*, 1984) using helper strain HB101/pRK2013 (Table 1). The expression of *Y. enterocolitica* O:3 outer core hexasaccharide in these strains, as well as in *E. coli* C600, was confirmed by DOC-PAGE analysis, as above.

Phage DNA techniques. Phage DNA was extracted as described for bacteriophage lambda (Sambrook & Russell, 2001). Standard DNA techniques were performed (Sambrook & Russell, 2001), and enzymes were used as recommended by the suppliers.

The ϕ R1-37 genome size was determined by analysing undigested and *Pvu*II-, *Xho*I-, *Dra*III- and *Dra*I-digested phage DNA by PFGE

Table 1. Bacterial strains and plasmids used in the work

Strain or plasmid	Description/genotype	Source/reference
<i>Yersinia enterocolitica</i>		
YeO3-R1	Host strain for ϕ R1-37. A spontaneous rough derivative of virulence-plasmid-cured <i>Y. enterocolitica</i> serotype O:3 strain 6471/76-c	Al-Hendy <i>et al.</i> (1992)
467/73	Serotype O:9. Human stool isolate from patient	Skurnik (1985)
467/73- ϕ R1-37-R	Spontaneous phage ϕ R1-37-resistant derivative of 467/73	This work
3229	Serotype O:50. Human stool isolate	Skurnik & Toivanen (1991)
3229- ϕ R1-37-R	Spontaneous phage ϕ R1-37-resistant derivative of 3229	This work
14779/83	Serotype O:5. Human stool isolate	Skurnik (1985); Skurnik & Toivanen (1991)
14779/83- ϕ R1-37-R	Spontaneous phage ϕ R1-37-resistant derivative of 14779/83	This work
18425/83	Serotype O:25,26,44. Human stool isolate	Skurnik (1985); Skurnik <i>et al.</i> (1995)
18425/83- ϕ R1-37-R	Spontaneous phage ϕ R1-37-resistant derivative of 18425/83	This work
YeO3-c-OC	Serotype O:3. $\Delta(wzx-wbcQ)$. Outer core-negative derivative of 6471/76-c	Biedzka-Sarek <i>et al.</i> (2005)
YeO3-c-OC-R	Serotype O:3. $\Delta(wzx-wbcQ)$. Outer core- and O antigen-negative derivative of 6471/76-c	Biedzka-Sarek <i>et al.</i> (2005)
8081-c	Serotype O:8. Virulence-plasmid-cured derivative of 8081	Portnoy <i>et al.</i> (1981)
8081-c-R2	Serotype O:8. Virulence-plasmid-cured, O antigen-negative derivative of 8081	Zhang <i>et al.</i> (1997)
<i>Yersinia intermedia</i>		
821/84	Serotype O:52,54. Human stool isolate	Skurnik & Toivanen (1991)
821/84- ϕ R1-37-R	Spontaneous phage ϕ R1-37-resistant derivative of 821/84	This work
<i>Yersinia pseudotuberculosis</i>		
R708Ly	Serotype O:9 reference strain, isolated from a mole in Japan	Tsubokura & Aleksic (1995)
R708Ly-R	Spontaneous rough derivative of R708Ly	This work
R708Ly- ϕ R1-37-R	Spontaneous phage ϕ R1-37-resistant derivative of R708Ly	This work
PB1 Δ wb	$\Delta(ddhD-wzz)::KmR$ -GenBlock. O antigen-negative derivative of serotype O:1b strain PB1 (Porat <i>et al.</i> , 1995). KmR	J.-A. Bengoechea and M. Skurnik, unpublished results; Kukkonen <i>et al.</i> (2004)
<i>Yersinia pestis</i>		
KIM D27	Lcr ⁺ , pgm ⁻ , pst ⁺	Garcia <i>et al.</i> (1999)
KIM D27Nar	nalR; Lcr ⁺ , pgm ⁻ , pst ⁺ , spontaneous nalidixic acid-resistant derivative of KIM D27	This work
EV76-c	Virulence-plasmid-cured derivative of EV76	Ben-Gurion & Hertman (1958); Portnoy & Falkow (1981)
<i>Escherichia coli</i>		
DH10B	F ⁻ mcrA $\Delta(mrr-hsdRMS-mcrBC)$, $\phi 80lacZ\Delta M15 \Delta lacX74 deoR recA1 endA1 ara\Delta 139 \Delta(ara, leu)7697 galU galK \lambda^- rpsL nupG \lambda^- tonA$	Life Technologies
CJ236	F ⁺ LAM ⁻ dut-1 ung-1 thi-1 spoT1 relA1/pCJ105 (CmR)	<i>E. coli</i> Genetic Stock Center
C600	thi thr leuB tonA lacY supE	Appleyard (1954)
HB101	F ⁻ $\Delta(gpt-proA)62 leuB6 glnV ara-14 galK2 lacY1 \Delta(mcr-mrr) rpsL20$ (Str ^r) xyl-5 mtl-1 recA13	Boyer & Roulland-Dussoix (1969)
Plasmids		
pRK2013	kmR; helper plasmid for conjugation	Ditta <i>et al.</i> (1980)
pTM100	clmR tetR; mobilizable cloning vector	Michiels & Cornelis (1991)
pBR322	ampR tetR; cloning vector	Bolivar <i>et al.</i> (1977); Sutcliffe (1979)
pRV7	ampR; Ye O:3 outer core gene cluster cloned in pBR322	Skurnik <i>et al.</i> (1995)
pRV16NP	clmR; Ye O:3 outer core gene cluster cloned in pTM100	This work

(Sambrook & Russell, 2001). As a size standard, Lambda ladder PFGE Marker (BioLabs) was used, and the run was performed on 1 % PFGE gel, as recommended for the Lambda ladder by the producer.

To clone the ϕ R1-37 DNA into *E. coli* strain CJ236 (Table 1), the phage DNA was digested with *Eco*RI or *Pst*I, ligated into the *Eco*RI or *Pst*I site of plasmid pBR322 (Table 1), respectively, and electroporated to *E. coli*

CJ236 competent cells (Ausubel *et al.*, 1987). To confirm that the clones had inserts originating from the phage DNA, 1 µl samples of isolated plasmid DNA were blotted onto nylon membranes (Roche Molecular Biochemicals) and analysed by hybridization with DIG-dUTP-labelled phage DNA. Labelling and hybridization were performed with the DIG-High Prime DNA Labelling and Detection Starter Kit II (Roche Molecular Biochemicals).

Plasmid clones containing phage DNA were sequenced first with the vector-specific primers pBRPstFor1 (5'-CATCCAGTCTATTA-ATTGTTGC-3'), pBRPstBack1 (5'-GAACCGGAGCTGAATGAA-3'), pBREcoFor1 (5'-CCTGACGCTAAGAAACCATT-3') and pBREco-Rev1 (5'-GCGTTAGCAATTTAACTGTGA-3'), then using primer-walking with phage-specific primers (not shown). The sequencing reactions were performed with the ABI PRISM BigDye Terminator Cycle Sequencing v 2.0 Ready Reaction Kit and analysed with an ABI 377 DNA analyser, as recommended by the manufacturer. The sequence data were assembled with the Staden Package programs Pregap4 and Gap4, and the preliminary analysis was done with the European Molecular Biology Open Software Suite (EMBOSS) version 2.6.0 and the National Center for Biotechnology Information (NCBI) BLAST program (www.ncbi.nlm.nih.gov/blast/blast.cgi).

Analysis of modified nucleotides. The nucleotide composition of the ϕ R1-37 genome was determined by hydrolysing the phage DNA to deoxynucleosides, and analysing the hydrolysate for nucleosides and dinucleotides by LC-MS/MS. The hydrolysis was done by the method of Crain (1990), in which the DNA is treated with nuclease P1, snake venom phosphodiesterase I and *E. coli* alkaline phosphatase. The enzymes were purchased from ICN Biomedicals.

LC-MS/MS analyses for detecting nucleosides were performed on a PE Sciex API 365 triple quadrupole LC/MS/MS System equipped with two PE Series 200 Micro pumps and a PE Series 200 autosampler. The reverse-phase HPLC column used was Waters Symmetry C18, 2.1 × 100 mm, particle size 3.5 µm, and the guard column was Waters Symmetry C18, 2.1 × 10 mm, particle size 3.5 µm. As a mobile phase, methanol was used as eluent A, and 5 mM ammonium formate as eluent B. Gradient system 1 was as follows: from 0.5% A to 1.0% A in 0–0.5 min; from 1.0% A to 50% A in 0.5–22 min. The flow rate was 160 µl min⁻¹ and the injection volume was 20 µl. Detection was made by using positive ion spray ionization (ESI) in neutral loss scan mode. The resolutions for Q1 and Q3 quadrupoles were set to LOW and UNIT resolution, respectively. The neutral loss detected was deoxyribose, *m/z* 116.1. The needle potential was set to 5200 V, declustering potential 22 V, focusing potential 113 V, entrance potential 2.5 V and collision energy 13 V. Nebulizer gas (nitrogen) was set to value 10.0, curtain gas (nitrogen) to value 11.0, collision gas to value 2.0 and turbo ion spray gas (nitrogen, 300 °C) to 6000 ml min⁻¹.

LC-MS analyses for detecting dinucleotides were performed with the equipment described above. As the mobile phase, eluent A was 50% methanol and 0.1% formic acid in water (v/v), and eluent B was 0.1% formic acid in water (v/v). Gradient system 2 was as follows: from 10% A to 85% A in 0–18 min; 85% A in 18–20 min. The flow rate was 100 µl min⁻¹ and the injection volume was 70 µl. Detection was made by using negative ion spray ionization (ESI) in single quadrupole scan mode. The resolution for the Q1 quadrupole was set to UNIT resolution. The mass range scanned was *m/z* 150–900. The needle potential was set to –4200 V, declustering potential –36 V, focusing potential –175 V and entrance potential –10 V. Nebulizer gas (nitrogen) was set to value 10.0, curtain gas (nitrogen) to value 11.0 and turbo ion spray gas (nitrogen, 300 °C) to 6000 ml min⁻¹.

The reference substances for nucleoside analysis, 2'-deoxyadenosine (dA), 2'-deoxycytidine (dC), 2'-deoxyguanosine (dG), thymidine (T) and 2'-deoxyuridine (dU) were purchased from Sigma-Aldrich

and Merck. The reference substances used for dinucleotide analysis were 2'-deoxyadenylyl(3'→5')-2'-deoxycytidine (dA-dC) and 2'-deoxycytidylyl(3'→5')-deoxyguanosine (dC-dG) (Sigma-Aldrich).

SDS-PAGE and N-terminal sequencing of structural proteins.

Phage particles were sedimented as above, suspended in SDS loading buffer (Sambrook & Russell, 2001) and heated at 95 °C for 15 min. Electrophoresis was conducted using a Hoefer SE 600 device (Amersham Pharmacia Biotech). The phage proteins were run on a 10% SDS-PAGE gel (Sambrook & Russell, 2001) with a constant current of 25 mA, and stained with Coomassie blue R250 (Amersham Pharmacia Biotech), according to manufacturer's instructions.

For the determination of the N-terminal amino acid sequences of the structural proteins, SDS-PAGE was conducted as above. The proteins were blotted with a Semidry IX device (Bio-Rad) onto Problott membrane (Applied Biosystems), as recommended by the manufacturer. N-terminal amino acid sequences of the four major protein bands were determined with an Applied Biosystems 477A protein sequencer equipped with an Applied Biosystems 120A PTH (phenylthiodantoin)-amino acid analyser. For degradation cycles and amino acid analysis, standard parameters provided by the manufacturer were used. The sequence data were analysed with the NCBI BLAST program and EMBOSS.

RESULTS

ϕ R1-37 belongs to the *Myoviridae*

Electron microscopy of phosphotungstic acid-stained ϕ R1-37 particles revealed that the phage has an icosahedral head, the diameter of which is 88 ± 12 nm, a 10 nm-long neck, and a contractile tail, the dimensions of which are $236 \pm 15 \times 14$ nm (Fig. 1a, b). The contracted sheath of the tail had dimensions of 96×23 nm. The phage also had tail fibres whose length was at least 86 nm. According to the ICTV classification, ϕ R1-37 belongs to the order *Caudovirales* and the family *Myoviridae* (Ackermann, 2003; Maniloff & Ackermann, 1998). In Bradley's classification, ϕ R1-37 belongs to the A1 morphotype (Bradley, 1967). ϕ R1-37 is a relatively large bacteriophage; this is illustrated in Fig. 1(c), which shows a YeO3-R1 cell (length 1.7 µm) with the phage adsorbed on its surface.

ϕ R1-37 growth cycle

The infection cycle of ϕ R1-37 was studied by making a one-step growth curve of the phage propagating on YeO3-R1 (Fig. 2). The eclipse period was found to be 40 min, and the latent period 50 min. These were followed by a rise period of 20 min. The average burst size of ϕ R1-37 was ~ 80 p.f.u. per infected cell. The lytic life cycle of the phage was supported by the finding that DIG-dUTP-labelled ϕ R1-37 DNA did not hybridize to the genomic DNA of 128 bacterial strains tested, including both ϕ R1-37-sensitive and -resistant strains (not shown).

Characterization of ϕ R1-37 receptor

Bacteriophage ϕ R1-37 was isolated based on its ability to infect strain YeO3-R1, that is, a virulence-plasmid-cured, O antigen-negative derivative of *Y. enterocolitica* serotype O:3

Fig. 1. Electron micrographs of negatively stained ϕ R1-37 particles. Phages with non-contracted (a) and contracted (b) tails are shown. (c) ϕ R1-37 attached to *Y. enterocolitica* YeO3-R1 cell (length of cell, 1.7 μ m).

(Table 1). The virulence-plasmid-cured strain YeO3-c was resistant when grown at 22 °C and sensitive when grown at 37 °C, indicating that abundant LPS O side chain (O antigen) expression at 22 °C sterically blocked the phage receptor (Skurnik *et al.*, 1995). On the other hand, the wild-type strain YeO3 was resistant to phage ϕ R1-37 irrespective of growth temperature, suggesting that expression of YadA [*Yersinia* adhesin A, an outer-membrane protein which is expressed by the virulence plasmid at 37 °C (Bölin *et al.*, 1982, 1985; El Tahir & Skurnik, 2001)] combined with a

Fig. 2. One-step growth curve of ϕ R1-37 growing on *Y. enterocolitica* YeO3-R1. The p.f.u. per infected cell in chloroform-treated and untreated culture is shown by open and filled circles, respectively. Eclipse and latent periods were 40 and 50 min, respectively, and the burst size was ~80 p.f.u. per infected cell. Each data point is an average of two individual experiments. Error bars indicate range.

lower abundance of O antigen inhibited the phage infection. Rough forms of both virulence-plasmid-positive and -negative YeO3 were sensitive, regardless of the growth temperature. These results thus suggested that LPS outer core could function as the phage receptor.

To further characterize the phage receptor, we determined whether degradation of bacterial proteins or LPS inhibits the phage binding. To this end, a YeO3-R1 suspension was immobilized on nitrocellulose membranes and the membranes were subjected to different treatments, altering either protein or carbohydrate structures, prior to the detection of bacteria with DIG- ϕ R1-37. As shown in Table 2, drying, boiling, SDS sample buffer and Proteinase K had no effect on the phage binding. On the other hand, incubation of bacteria with periodate clearly destroyed the phage receptor.

Table 2. Detection of phage ϕ R1-37 receptor using DIG- ϕ R1-37: effect of different treatments

+, DIG- ϕ R1-37 binding to immobilized bacteria on nitrocellulose membranes; –, no binding.

Target	ϕ R1-37 receptor binding
Whole bacteria, spotted, dried	+
Boiled bacteria, spotted, dried	+
Boiled bacteria, spotted, dried, IO ₄ [−]	−
Boiled bacteria, spotted, dried, proteinase K	+
SDS sample buffer, spotted, dried	+
SDS-PAGE, phage blotting	−
Total membranes, spotted, dried	+

These results thus strengthened the conclusion that the outer core hexasaccharide structure (Radziejewska-Lebrecht *et al.*, 1998) is the phage receptor.

The host range of ϕ R1-37 showed that the phage can also infect other members of the genus *Yersinia* (Skurnik *et al.*, 1995), indicating that the hexasaccharide structure might be common between those bacteria. Indeed, phage-resistant derivatives of *Y. enterocolitica* strains 467/73 (serotype O:9), 3229 (serotype O:50), 14779/83 (serotype O:5), 18425/83 (serotype O:25,26,44) and *Y. intermedia* strain 821/84 (serotype O:52,54) (Table 1) showed a change in the LPS profile corresponding to a loss of outer core (data not shown). Surprisingly, two separate phage-resistant derivatives of *Y. pseudotuberculosis* serotype O:9 strain R708Ly, R708Ly-R and R708Ly- ϕ R1-37-R (Table 1), had lost the O antigen instead of the outer core (Fig. 3). Thus, unlike in other *Yersinia* strains tested, in this strain, the phage receptor structure was present in the O antigen.

In order to confirm that the YeO3 LPS outer core is the phage receptor, plasmid pRV16NP, in which the outer core gene cluster is cloned under a tetracycline promoter, was constructed (Table 1). The plasmid was then mobilized into different ϕ R1-37-resistant *Yersinia* strains: *Y. enterocolitica* YeO3-c-OC, YeO3-c-OC-R, 8081-c and 8081-c-R2; *Y. pestis* KIM D27Nar and EV76-c; and *Y. pseudotuberculosis* PB1 Δ wb (Table 1). The expression of the outer core hexasaccharide in the strains harbouring pRV16NP was verified with DOC-PAGE (Fig. 4). All the *Yersinia* and *E. coli* strains expressing YeO3 LPS outer core became ϕ R1-37 sensitive (not shown), confirming that the YeO3 outer core hexasaccharide forms the receptor for ϕ R1-37.

ϕ R1-37 DNA contains dU instead of T

Aberrant digestion of ϕ R1-37 DNA by *Acc65I* (the enzyme failed to digest the DNA, while its isoschizomer *KpnI*

Fig. 3. LPS profile of ϕ R1-37-sensitive *Y. pseudotuberculosis* O:9 strain R708Ly and its phage-resistant derivatives R708Ly- ϕ R1-37-R and R708Ly-R. LPS analysis was by DOC-PAGE and silver staining. LA, lipid A; C, core; O-Ag, O antigen.

Fig. 4. Expression of Ye O:3 outer core hexasaccharide in different *Yersinia* species. Analysis of LPS was by DOC-PAGE and silver staining. LA, lipid A; IC, inner core; OC, outer core.

digested it into > 10 fragments, not shown) and difficulties in cloning the phage genome in *E. coli* DH10B (see below) directed us to study the chemical composition of the phage DNA. When the hydrolysed DNA was analysed with LC-MS/MS (neutral loss scan mode), it was obvious that the phage genome contained the three normal deoxynucleosides, dA, dC and dG (Fig. 5). However, a highly interesting finding was that in the chromatogram the T peak was missing, and there was an additional peak whose molecular mass was identical to that of dU. Indeed, the peak was identified to be dU, since the retention times of the supposed dU and of dU in pure solvent were very near each other (Fig. 5), and the analysis method revealed that there was a deoxyribose moiety in the molecule.

The GC content of ϕ R1-37 DNA was calculated based on the areas of the peaks of individual nucleosides in the chromatograph. The estimation resulted in a GC content of 36 %, the proportions of individual nucleosides being 31 % for dA, 33 % for dU, 18 % for dG and 18 % for dC. A trace amount of T (0.5 %) was also detected (data not shown).

In order to verify that the composition analysis had been quantitative, the DNA hydrolysate was also studied for unhydrolysed dideoxynucleotides. The dinucleotides were not detected in the LC-MS/MS chromatograph (data not shown), indicating that the DNA hydrolysis had been complete and that the composition analysis was quantitative.

Size and partial sequencing of the ϕ R1-37 genome

The size of the ϕ R1-37 genome was analysed with restriction digestions and PFGE, and was estimated to be ~270 kb (not shown). This correlates well with the large size of the phage capsid.

All the preliminary attempts to clone ϕ R1-37 DNA to *E. coli* DH10B were unsuccessful (not shown). This was then found to be due to the presence of dU in the phage DNA, since dU-containing DNA is known to be degraded by bacterial enzymes (Duncan & Warner, 1977; Taylor & Weiss, 1982;

Fig. 5. LC-MS/MS analysis of ϕ R1-37 DNA after hydrolysis to nucleosides. (a) Reference 2'-deoxynucleosides dA, dC, dT and dG; (b) hydrolysed phage DNA; (c) reference deoxynucleoside dU. c.p.s., Counts per second.

Wang & Mosbaugh, 1988). The cloning of the phage DNA to *E. coli* CJ236, devoid of dUTPase (the product of the *dut* gene) and uracil *N*-glycosylase (the product of the *ung* gene), then proved successful.

The obtained plasmid clones were used as a starting material for sequencing of the ϕ R1-37 genome. The partial sequence (~ 76 kb, representing approximately 28 % of the genome, unpublished results) was subjected to preliminary sequence analysis, which showed that the GC content was 33 %. This is in reasonable agreement with the value (36 %) obtained by LC-MS/MS analysis.

Fig. 6. SDS-PAGE analysis of ϕ R1-37 structural proteins. Protein molecular weight marker is shown at the left; the four major phage particle proteins and their N-terminal amino acid sequences are indicated at the right.

Structural proteins of ϕ R1-37

SDS-PAGE of ϕ R1-37 structural proteins revealed four major protein bands (Fig. 6) that were named sp69, sp46, sp31 and sp24, according to their estimated sizes in kDa. Of these, sp46 was the most abundant, which makes it the likely major capsid protein.

The N-terminal amino acid sequences of the four major structural proteins were determined and searched for similarities to known viral proteins. The sequence analysis did not reveal significant similarities to any known protein sequences. Interestingly, however, the N-terminal sequences of sp69 and sp46 were almost identical: the only difference was the amino acid at the N-terminus, which was alanine for sp69 and serine for sp46.

In the preliminary genomic sequence, the ORFs encoding the structural proteins sp24 and sp46 were identified (accession numbers AJ972879 and AJ972880, respectively). The mature sp24 seems to be proteolytically processed; presumably 55 amino acids from the N-terminus are removed, as estimated by the location of the potential Shine–Dalgarno sequence and translational start codon. The deduced amino acid sequence of the processed protein predicts a small (205 aa, 21.4 kDa) and basic protein (isoelectric point 10.3). The sp24 sequence is not similar to known viral protein sequences and does not have easily recognizable motifs (e.g. DNA binding domains), and thus the exact role of sp24 in the phage capsid is not yet known.

The major capsid protein sp46 appears also to undergo proteolytic processing during its maturation, as the plausible translational start site is 43 amino acids before the N-terminus of the mature protein. The mature protein is 432 amino acids long, negatively charged (isoelectric point 5.0) and has a calculated size of 47.8 kDa. The deduced amino acid sequence shows no similarity to known proteins.

At the moment, it is not known whether sp69 is produced from the same or a different gene than sp46. In several different phages, sites for frameshifts that allow the translation of a longer product have been identified (Condron *et al.*, 1991; Garcia *et al.*, 2003b; Pajunen *et al.*, 2001), but no such site could be identified from the gene encoding sp46.

DISCUSSION

ϕ R1-37 has been shown to be a member of virus family *Myoviridae*. It is a very large bacteriophage: the size of its genome is roughly similar to that of the genome of the *Pseudomonas aeruginosa* phage ϕ KZ (280 kb), which is the largest bacteriophage DNA sequenced so far (Mesyanzhinov *et al.*, 2002). The large genome size and the presence of the receptor in many *Yersinia* strains make ϕ R1-37 a possible vehicle for horizontal gene transfer.

The comparison of N-terminal or deduced amino acid sequences of ϕ R1-37 structural proteins to databases did not reveal similar protein sequences, indicating that close relatives of ϕ R1-37 have not been characterized. The probable major capsid protein sp46 had an almost identical N-terminal sequence to that of sp69, suggesting that sp69 may be a minor capsid protein. So far, only the gene encoding sp46 has been identified in the partial genomic sequence, and it is not yet known whether these two proteins are products of the same gene or a result of an ancient gene duplication and subsequent differentiation. Both systems have been shown to apply for bacteriophage capsid proteins. The members of the T7 group, T7, T3 and ϕ YeO3-12, produce the major and minor capsid proteins gp10A and gp10B, respectively, via a -1 frameshift that allows the translation of the longer product (Condron *et al.*, 1991; Pajunen *et al.*, 2001). Similarly, in *Lactobacillus casei* phage A2, the two different-sized capsid proteins are translated from the same gene (Garcia *et al.*, 2003b). On the other hand, in *E. coli* phage T4, the head proteins gp23 and gp24 have evolved via gene duplication (Kutter *et al.*, 1995; Miller *et al.*, 2003). Predicted from the nucleotide sequences of their genes, both structural proteins sp24 and sp46 are apparently proteolytically processed during phage maturation. Post-translational processing of structural proteins is a common strategy for morphogenesis of tailed bacteriophages, and has been described for phages such as T4 (Mathews *et al.*, 1983), ϕ KZ (Mesyanzhinov *et al.*, 2002) and A2 (Garcia *et al.*, 2003b).

Our results demonstrated conclusively that the phage receptor is present in the outer core hexasaccharide of the *Y. enterocolitica* O:3 LPS. The receptor was sensitive to periodate treatment, which is known to degrade carbohydrates containing a *cis*-1,2-diol motif in their structure, and the outer core hexasaccharide was lost in phage-resistant mutants. In addition, the receptor was expressed along with the outer core hexasaccharide in *Y. pestis*, *Y. pseudotuberculosis*, *Y. enterocolitica* and *E. coli* strains when the outer core gene cluster carrying plasmid pRV16NP was introduced

into these bacteria. A novel finding was that *Y. pseudotuberculosis* serotype O:9 strain R708Ly was phage-sensitive also. As phage-resistant derivatives of R708Ly had lost the O antigen, it is very likely that, unlike in *Y. enterocolitica*, in *Y. pseudotuberculosis* O:9, the phage receptor is part of the O antigen structure. Work is under way to determine the sugar-residue composition and structure of the *Y. pseudotuberculosis* O:9 O antigen in order to elucidate this question.

The most intriguing finding in this study was that the ϕ R1-37 genome is composed of DNA in which T is almost completely replaced by dU. The only organisms so far known to have dU in their DNA are the *Bacillus subtilis*-specific bacteriophages PBS1 and PBS2 (Takahashi & Marmur, 1963).

Uracil residues are normally introduced into DNA either by the accidental incorporation of dUMP in place of TMP by DNA polymerase, or by the spontaneous deamination of cytosine in DNA (Lindahl, 1993; Tye *et al.*, 1978). The latter reaction will yield a premutagenic U.G mispair and promotes a mutation in the next DNA-replication cycle. It is thus crucial for cells to be able to repair the uracil residues occurring in their DNA. A system known as the base-excision repair (BER) pathway is responsible for this function (Sung & Mosbaugh, 2003; Taylor & Weiss, 1982). BER is initiated by uracil-DNA glycosylase (Ung), which cleaves the *N*-glycosylic bond between the uracil base and the DNA backbone, leaving an apyrimidinic site (Lindahl *et al.*, 1977). This site is then incised by an apyrimidinic endonuclease, after which the gap is filled by DNA polymerase I. Finally, the repair is completed by DNA ligase. The first enzyme of the pathway, Ung, is highly specific for U in DNA, cleaving both U.A and U.G pairs, and even single-stranded DNA (Lindahl *et al.*, 1977). Ung is found to be highly conserved in evolution, with *E. coli* and human enzymes showing 55.7% sequence identity (Olsen *et al.*, 1989, 1991).

Bacteriophage PBS2 escapes the base-excision repair pathway by expressing a protein that inhibits Ung (Bennett *et al.*, 1993; Lundquist *et al.*, 1997; Wang & Mosbaugh, 1988). This protein, uracil-DNA glycosylase inhibitor (Ugi), abolishes Ung activity in both bacteria and humans (Acharya *et al.*, 2003; Olsen *et al.*, 1991; Wang & Mosbaugh, 1988), and has been shown to increase mutation frequency when expressed in human cells (Radany *et al.*, 2000). In addition to Ugi, PBS2 expresses a set of enzymes that function cooperatively to increase the intracellular dUTP pool and decrease the TTP concentration, thus resulting in the close to 100% incorporation of dU instead of T in the DNA (Wang & Mosbaugh, 1988). In the future, it would be interesting to know whether ϕ R1-37 and PBS2 have similar enzymic machinery for their DNA synthesis. The main question that remains to be answered is how these phages are related to each other, and whether they have acquired the genes responsible for their DNA synthesis through horizontal gene transfer, or have evolved independently.

In their review article, Poole *et al.* (2001) speculated that there might be a whole U-DNA world alive in viruses. The present finding that there are U-DNA phages infecting bacterial species as distant as *Y. enterocolitica* and *B. subtilis* now supports this hypothesis. Since the enzymes responsible for maintenance of uracil-containing DNA are proven to be mutagenic in cultured human cells (Radany *et al.*, 2000), the possibility of having a U-DNA genome should specially be considered when thinking about the therapeutic use of bacteriophages.

ACKNOWLEDGEMENTS

We thank Lauri Pelliniemi and Jouko Mäki at the Laboratory of Electron Microscopy, University of Turku, for their help with the electron microscopy pictures. Jukka Hellman at the Turku Centre for Biotechnology is thanked for the N-terminal amino acid sequencing, and Jari Jalava and Marianne Lindgren at the National Public Health Institute in Turku for the PFGE. Jukka Karhu is greatly appreciated for technical assistance. Academy of Finland projects 50441, 40932 and 203602, the European Commission (Project QRLT-1999-00780) and the Finnish National Technology Agency, TEKES, are thanked for financial support.

REFERENCES

- Acharya, N., Kumar, P. & Varshney, U. (2003). Complexes of the uracil-DNA glycosylase inhibitor protein, Ugi, with *Mycobacterium smegmatis* and *Mycobacterium tuberculosis* uracil-DNA glycosylases. *Microbiology* **149**, 1647–1658.
- Ackermann, H. W. (2003). Bacteriophage observations and evolution. *Res Microbiol* **154**, 245–251.
- Al-Hendy, A., Toivanen, P. & Skurnik, M. (1991). Expression cloning of the *Yersinia enterocolitica* O:3 *rfb* gene cluster in *Escherichia coli* K12. *Microb Pathog* **10**, 47–59.
- Al-Hendy, A., Toivanen, P. & Skurnik, M. (1992). Lipopolysaccharide O side chain of *Yersinia enterocolitica* O:3 is an essential virulence factor in an orally infected murine model. *Infect Immun* **60**, 870–875.
- Appleyard, R. K. (1954). Segregation of new lysogenic types during growth of doubly lysogenic strain derived from *Escherichia coli* K12. *Genetics* **39**, 440–452.
- Ausubel, F. M., Brent, R., Kingston, R. E., Moore, D. D., Seidman, J. G., Smith, J. A. & Struhl, K. (1987). *Current Protocols in Molecular Biology*. New York: Wiley.
- Baker, P. M. & Farmer, J. J., III (1982). New bacteriophage typing system for *Yersinia enterocolitica*, *Yersinia kristensenii*, *Yersinia frederiksenii*, and *Yersinia intermedia*: correlation with serotyping, biotyping and antibiotic susceptibility. *J Clin Microbiol* **15**, 491–502.
- Ben-Gurion, R. & Hertman, I. (1958). Bacteriocin-like material produced by *Pasteurella pestis*. *J Gen Microbiol* **19**, 289–297.
- Bennett, S. E., Schimerlik, M. I. & Mosbaugh, D. W. (1993). Kinetics of the uracil-DNA glycosylase/inhibitor protein association. Ung interaction with Ugi, nucleic acids, and uracil compounds. *J Biol Chem* **268**, 26879–26885.
- Biedzka-Sarek, M., Venho, R. & Skurnik, M. (2005). Role of YadA, Ail, and lipopolysaccharide in serum resistance of *Yersinia enterocolitica* serotype O:3. *Infect Immun* **73**, 2232–2244.
- Birge, E. A. (1988). *Bacterial and Bacteriophage Genetics: an Introduction*, 2nd edn. New York: Springer.
- Bölin, I., Norlander, L. & Wolf-Watz, H. (1982). Temperature-inducible outer membrane protein of *Yersinia pseudotuberculosis* and *Yersinia enterocolitica* is associated with the virulence plasmid. *Infect Immun* **37**, 506–512.
- Bölin, I., Portnoy, D. A. & Wolf-Watz, H. (1985). Expression of the temperature-inducible outer membrane proteins of *Yersiniae*. *Infect Immun* **48**, 234–240.
- Bolivar, F., Rodriguez, R. L., Greene, P. J., Betlach, M. C., Heyneker, H. L. & Boyer, H. W. (1977). Construction and characterization of new cloning vehicles. II. A multipurpose cloning system. *Gene* **2**, 95–113.
- Bottone, E. J. (1997). *Yersinia enterocolitica*: the charisma continues. *Clin Microbiol Rev* **10**, 257–276.
- Boyer, H. W. & Roulland-Dussoix, D. (1969). A complementation analysis of the restriction and modification of DNA in *Escherichia coli*. *J Mol Biol* **41**, 459–472.
- Bradley, D. E. (1967). Ultrastructure of bacteriophage and bacteriocins. *Bacteriol Rev* **31**, 230–314.
- Condrón, B. G., Atkins, J. F. & Gesteland, R. F. (1991). Frameshifting in gene 10 of bacteriophage T7. *J Bacteriol* **173**, 6998–7003.
- Crain, P. F. (1990). Preparation and enzymatic hydrolysis of DNA and RNA for mass spectrometry. *Methods Enzymol* **193**, 782–790.
- Ditta, G., Stanfield, S., Corbin, D. & Helinski, D. R. (1980). Broad host range DNA cloning system for Gram-negative bacteria: construction of a gene bank of *Rhizobium meliloti*. *Proc Natl Acad Sci U S A* **77**, 7347–7351.
- Duncan, B. K. & Warner, H. R. (1977). Metabolism of uracil-containing DNA: degradation of bacteriophage PBS2 DNA in *Bacillus subtilis*. *J Virol* **22**, 835–838.
- El Tahir, J. & Skurnik, M. (2001). YadA, the multifaceted *Yersinia* adhesin. *Int J Med Microbiol* **291**, 209–218.
- Garcia, E., Nedialkov, Y. A., Elliott, J., Motin, V. L. & Brubaker, R. R. (1999). Molecular characterization of KatY (antigen 5), a thermoregulated chromosomally encoded catalase-peroxidase of *Yersinia pestis*. *J Bacteriol* **181**, 3114–3122.
- Garcia, E., Elliott, J. M., Ramanculov, E., Chain, P. S., Chu, M. C. & Molineux, I. J. (2003a). The genome sequence of *Yersinia pestis* bacteriophage ϕ A1122 reveals an intimate history with the coliphage T3 and T7 genomes. *J Bacteriol* **185**, 5248–5262.
- Garcia, P., Ladero, V. & Suarez, J. E. (2003b). Analysis of the morphogenetic cluster and genome of the temperate *Lactobacillus casei* bacteriophage A2. *Arch Virol* **148**, 1051–1070.
- Gerhardt, P., Murray, R. G. E., Wood, W. A. & Krieg, N. R. (1984). *Methods for General and Molecular Bacteriology*. Washington, DC: American Society for Microbiology.
- Hendrix, R. W. (2002). Bacteriophages: evolution of the majority. *Theor Popul Biol* **61**, 471–480.
- Hertwig, S., Klein, I., Lurz, R., Lanka, E. & Appel, B. (2003a). PY54, a linear plasmid prophage of *Yersinia enterocolitica* with covalently closed ends. *Mol Microbiol* **48**, 989–1003.
- Hertwig, S., Klein, I., Schmidt, V., Beck, S., Hammerl, J. A. & Appel, B. (2003b). Sequence analysis of the genome of the temperate *Yersinia enterocolitica* phage PY54. *J Mol Biol* **331**, 605–622.
- Kukkonen, M., Suomalainen, M., Kyllönen, P., Lähtenmäki, K., Lång, H., Virkola, R., Helander, I. M., Holst, O. & Korhonen, T. K. (2004). Lack of O-antigen is essential for plasminogen activation by *Yersinia pestis* and *Salmonella enterica*. *Mol Microbiol* **51**, 215–225.
- Kutter, E., Gachechiladze, K., Poglazov, A., Marusich, E., Shneider, M., Aronsson, P., Napuli, A., Porter, D. & Mesyanzhinov, V. (1995). Evolution of T4-related phages. *Virus Genes* **11**, 285–297.

- Lindahl, T. (1993). Instability and decay of the primary structure of DNA. *Nature* **362**, 709–715.
- Lindahl, T., Ljungquist, S., Siebert, W., Nyberg, B. & Sperens, B. (1977). DNA N-glycosidases: properties of uracil-DNA glycosidase from *Escherichia coli*. *J Biol Chem* **252**, 3286–3294.
- Lundquist, A. J., Beger, R. D., Bennett, S. E., Bolton, P. H. & Mosbaugh, D. W. (1997). Site-directed mutagenesis and characterization of uracil-DNA glycosylase inhibitor protein. Role of specific carboxylic amino acids in complex formation with *Escherichia coli* uracil-DNA glycosylase. *J Biol Chem* **272**, 21408–21419.
- Maniloff, J. & Ackermann, H. W. (1998). Taxonomy of bacterial viruses: establishment of tailed virus genera and the order *Caudovirales*. *Arch Virol* **143**, 2051–2063.
- Mathews, C. K., Kutter, E., Mosig, G. & Berget, P. B. (1983). *Bacteriophage T4*. Washington, DC: American Society for Microbiology.
- Mesyanzhinov, V. V., Robben, J., Grymonprez, B., Kostyuchenko, V. A., Bourkaltseva, M. V., Sykilinda, N. N., Krylov, V. N. & Volckaert, G. (2002). The genome of bacteriophage ϕ KZ of *Pseudomonas aeruginosa*. *J Mol Biol* **317**, 1–19.
- Michiels, T. & Cornelis, G. R. (1991). Secretion of hybrid proteins by the *Yersinia* Yop export system. *J Bacteriol* **173**, 1677–1685.
- Miller, E. S., Kutter, E., Mosig, G., Arisaka, F., Kunisawa, T. & Ruger, W. (2003). Bacteriophage T4 genome. *Microbiol Mol Biol Rev* **67**, 86–156.
- Olsen, L. C., Aasland, R., Wittwer, C. U., Krokan, H. E. & Helland, D. E. (1989). Molecular cloning of human uracil-DNA glycosylase, a highly conserved DNA repair enzyme. *EMBO J* **8**, 3121–3125.
- Olsen, L. C., Aasland, R., Krokan, H. E. & Helland, D. E. (1991). Human uracil-DNA glycosylase complements *E. coli ung* mutants. *Nucleic Acids Res* **19**, 4473–4478.
- Pajunen, M., Kiljunen, S. & Skurnik, M. (2000). Bacteriophage ϕ YeO3-12 specific for *Yersinia enterocolitica* serotype O:3 is related to coliphages T3 and T7. *J Bacteriol* **182**, 5114–5120.
- Pajunen, M. I., Kiljunen, S. J., Söderholm, M. E.-L. & Skurnik, M. (2001). Complete genomic sequence of the lytic bacteriophage ϕ YeO3-12 of *Yersinia enterocolitica* serotype O:3. *J Bacteriol* **183**, 1928–1937.
- Pedulla, M. L., Ford, M. E., Houtz, J. M. & 17 other authors (2003). Origins of highly mosaic mycobacteriophage genomes. *Cell* **113**, 171–182.
- Poole, A., Penny, D. & Sjöberg, B. M. (2001). Confounded cytosine! Tinkering and the evolution of DNA. *Nat Rev Mol Cell Biol* **2**, 147–151.
- Popp, A., Hertwig, S., Lurz, R. & Appel, B. (2000). Comparative study of temperate bacteriophages isolated from *Yersinia*. *Syst Appl Microbiol* **23**, 469–478.
- Porat, R., McCabe, W. R. & Brubaker, R. R. (1995). Lipopolysaccharide-associated resistance to killing of *Yersiniae* by complement. *J Endotoxin Res* **2**, 91–97.
- Portnoy, D. A. & Falkow, S. (1981). Virulence-associated plasmids from *Yersinia enterocolitica* and *Yersinia pestis*. *J Bacteriol* **148**, 877–883.
- Portnoy, D. A., Moseley, S. L. & Falkow, S. (1981). Characterization of plasmids and plasmid-associated determinants of *Yersinia enterocolitica* pathogenesis. *Infect Immun* **31**, 775–782.
- Radany, E. H., Dornfeld, K. J., Sanderson, R. J., Savage, M. K., Majumdar, A., Seidman, M. M. & Mosbaugh, D. W. (2000). Increased spontaneous mutation frequency in human cells expressing the phage PBS2-encoded inhibitor of uracil-DNA glycosylase. *Mutat Res* **461**, 41–58.
- Radziejewska-Lebrecht, J., Skurnik, M., Shashkov, A. S., Brade, L., Rozalski, A., Bartodziejska, B. & Mayer, H. (1998). Immunochemical studies on R mutants of *Yersinia enterocolitica* O:3. *Acta Biochim Pol* **45**, 1011–1019.
- Sambrook, J. & Russell, D. (2001). *Molecular Cloning: a Laboratory Manual*, 3rd edn. Cold Spring Harbor, NY: Cold Spring Harbor Laboratory.
- Skurnik, M. (1984). Lack of correlation between the presence of plasmids and fimbriae in *Yersinia enterocolitica* and *Yersinia pseudotuberculosis*. *J Appl Bacteriol* **56**, 355–363.
- Skurnik, M. (1985). *Studies on the virulence plasmids of Yersinia species*. PhD thesis, University of Oulu.
- Skurnik, M. & Toivanen, P. (1991). Intervening sequences (IVSs) in the 23S ribosomal RNA genes of pathogenic *Yersinia enterocolitica* strains. The IVSs in *Y. enterocolitica* and *Salmonella typhimurium* have a common origin. *Mol Microbiol* **5**, 585–593.
- Skurnik, M., Venho, R., Toivanen, P. & Al-Hendy, A. (1995). A novel locus of *Yersinia enterocolitica* serotype O:3 involved in lipopolysaccharide outer core biosynthesis. *Mol Microbiol* **17**, 575–594.
- Stevenson, R. & Airdrie, D. (1984). Isolation of *Yersinia ruckeri* bacteriophages. *Appl Environ Microbiol* **47**, 1201–1205.
- Sung, J. S. & Mosbaugh, D. W. (2003). *Escherichia coli* uracil- and ethenocytosine-initiated base excision DNA repair: rate-limiting step and patch size distribution. *Biochemistry* **42**, 4613–4625.
- Sutcliffe, J. G. (1979). Complete nucleotide sequence of the *Escherichia coli* plasmid pBR322. *Cold Spring Harb Symp Quant Biol* **43**, 77–90.
- Takahashi, I. & Marmur, J. (1963). Replacement of thymidylic acid by deoxyuridylic acid in the deoxyribonucleic acid of a transducing phage for *Bacillus subtilis*. *Nature* **197**, 794–795.
- Taylor, A. F. & Weiss, B. (1982). Role of exonuclease III in the base excision repair of uracil-containing DNA. *J Bacteriol* **151**, 351–357.
- Tsubokura, M. & Aleksic, S. (1995). A simplified antigenic scheme for serotyping of *Yersinia pseudotuberculosis*: phenotypic characterization of reference strains and preparation of O and H factor sera. *Contrib Microbiol Immunol* **13**, 99–105.
- Tye, B. K., Chien, J., Lehman, I. R., Duncan, B. K. & Warner, H. R. (1978). Uracil incorporation: a source of pulse-labeled DNA fragments in the replication of the *Escherichia coli* chromosome. *Proc Natl Acad Sci U S A* **75**, 233–237.
- Wang, Z. & Mosbaugh, D. W. (1988). Uracil-DNA glycosylase inhibitor of bacteriophage PBS2: cloning and effects of expression of the inhibitor gene in *Escherichia coli*. *J Bacteriol* **170**, 1082–1091.
- Wauters, G., Aleksic, S., Charlier, J. & Schulze, G. (1991). Somatic and flagellar antigens of *Yersinia enterocolitica* and related species. *Contrib Microbiol Immunol* **12**, 239–243.
- Zhang, L., Radziejewska-Lebrecht, J., Krajewska-Pietrasik, D., Toivanen, P. & Skurnik, M. (1997). Molecular and chemical characterization of the lipopolysaccharide O-antigen and its role in the virulence of *Yersinia enterocolitica* serotype O:8. *Mol Microbiol* **23**, 63–76.