Microbiology (2005), **151**, 941–950 DOI 10.1099/mic.0.27632-0

Genetic structure and chromosomal integration site of the cryptic prophage CP-1639 encoding Shiga toxin 1

Kristina Creuzburg,[1] Bernd Köhler,[2] Helena Hempel,[1] Peter Schreier,[3] Enno Jacobs[1] and Herbert Schmidt[1]†

[1]Institut für Medizinische Mikrobiologie und Hygiene der TU Dresden, Germany

[2,3]Institut für Hygiene und Mikrobiologie[2] and Lehrstuhl für Lebensmittelchemie[3] der Bayerischen Julius Maximilians Universität Würzburg, Germany

Correspondence
Herbert Schmidt
hschmidt@uni-hohenheim.de

The sequence of 50 625 bp of chromosomal DNA derived from Shiga-toxin (Stx)-producing *Escherichia coli* (STEC) O111 : H$^-$ strain 1639/77 was determined. This DNA fragment contains the cryptic Stx1-encoding prophage CP-1639 and its flanking chromosomal regions. The genome of CP-1639 basically resembles that of lambdoid phages in structure, but contains three IS*629* elements, one of which disrupts the gene of a tail fibre component. The prophage genome lacks parts of the recombination region including integrase and excisionase genes. Moreover, a capsid protein gene is absent. CP-1639 is closely associated with an integrase gene of an ancient integrative element. This element consists of three ORFs of unknown origin and a truncated integrase gene homologous to *intA* of CP4-57. By PCR analysis and sequencing, it was shown that this integrative element is present in a number of non-O157 STEC serotypes and in non-STEC strains, where it is located at the 3′-end of the chromosomal *ssrA* gene. Whereas in most *E. coli* O111 : H$^-$ strains, prophages are inserted in this site, *E. coli* O26 strains contain the integrative element not connected to a prophage. In *E. coli* O103 strains, the genetic structure of this region is variable. Comparison of DNA sequences of this particular site in *E. coli* O157 : H7 strain EDL933, *E. coli* O111 : H$^-$ strain 1639/77 and *E. coli* K-12 strain MG1655 showed that the *ssrA* gene is associated in all cases with the presence of foreign DNA. The results of this study have shown that the cryptic prophage CP-1639 is associated with an integrative element at a particular site in the *E. coli* chromosome that possesses high genetic variability.

Received 17 September 2004
Revised 13 December 2004
Accepted 16 December 2004

INTRODUCTION

Shiga-toxin (Stx)-producing *Escherichia coli* (STEC) can cause sporadic cases and outbreaks of food-borne diseases such as diarrhoea, haemorrhagic colitis and the haemolytic–uraemic syndrome (HUS). Following their first description in 1983 (Riley *et al.*, 1983; Wells *et al.*, 1983), mainly STEC of serotype O157 : H7 were considered to be a serious health risk worldwide (Kaper, 1998). However, non-O157 *E. coli* serotypes with variations in virulence determinants have emerged during the last decade, especially in Europe (Beutin *et al.*, 2004; Bielaszewska *et al.*, 2004; Blanco *et al.*, 2004; Brunder *et al.*, 1999; Caprioli *et al.*, 1997; Friedrich *et al.*, 2002, 2003; Leung *et al.*, 2003; Misselwitz *et al.*, 2003; Schmidt *et al.*, 1999c; Tozzi *et al.*, 2003). Stx secreted from STEC cells can damage endothelial cells (Bitzan & te Loo, 2003; Jacewicz *et al.*, 1999; O'Loughlin & Robins-Browne, 2001; Zoja *et al.*, 2001) as well as renal tubular cells, which may result in acute renal failure (Kaneko *et al.*, 2001; Williams *et al.*, 1999). Thus, production of Stx is considered as the cardinal virulence factor of STEC. Stx comprise two main groups: Stx1 and Stx2 (Nataro & Kaper, 1998; O'Brien & Holmes, 1987). In both groups, a growing number of variants have been described. Whereas Stx1 represents a more homogeneous group with only two reported variants (Koch *et al.*, 2003; Zhang *et al.*, 2002), the Stx2 group consists of a number of variants including Stx2c, Stx2d, activatable Stx2d, Stx2e, Stx2f and Stx2g (Leung *et al.*, 2003;

†Present address: Department of Food Microbiology, Institute of Food Technology, University of Hohenheim, Garbenstrasse 28, 70599 Stuttgart, Germany.

Abbreviations: HUS, haemolytic–uraemic syndrome; STEC, Shiga-toxin-producing *Escherichia coli*; Stx, Shiga toxin.

The GenBank/EMBL/DDBJ accession numbers for the sequences determined in this work are AJ304858 (phage CP-1639 and chromosomal integration site) and AJ831374 (integrative element of *E. coli* O103 : H2 strain 2905/96).

Melton-Celsa *et al.*, 1996; Pierard *et al.*, 1998; Schmidt *et al.*, 2000; Schmitt *et al.*, 1991; Thomas *et al.*, 1994). Members of the Stx family are differentiated by their nucleotide and amino acid sequences and biological activity, and frequently by their cross-reactivity with Stx1 and Stx2 antisera. Whereas the production of Stx2 appears to be important for the development of serious disease, there is evidence that the presence of Stx2d and Stx2e is associated with milder disease (Friedrich *et al.*, 2002). The *stx* genes analysed up to now are located in the genomes of lambdoid bacteriophages, which have been termed Stx-converting bacteriophages or Stx phages (O'Brien *et al.*, 1984; Schmidt 2001; Unkmeir & Schmidt, 2000; Willshaw *et al.*, 1987). The specific location of *stx* 3′ to the late antiterminator gene *Q* brings its expression under the control of the late promoter pR′ (Neely & Friedman, 1998; Wagner *et al.*, 2001). Therefore, *stx* expression is influenced by the switch from the lysogenic state to lysis of lambdoid prophages and can be triggered by inducing agents such as UV light, mitomycin C and antibiotics (Schmidt, 2001). Stx phages can infect *E. coli* cells and insert their DNA into specific chromosomal sites, allowing the bacterial hosts to survive as lysogenic Stx-producing strains. Such transduction of *stx* genes has been shown *in vitro* (Schmidt *et al.*, 1999a; Scotland *et al.*, 1983; Strauch *et al.*, 2001) and *in vivo* (Acheson *et al.*, 1998; Toth *et al.*, 2003), and it may contribute to the spread of *stx* genes among different *E. coli* serotypes. Thus, integration of Stx phage DNA into *E. coli* chromosomes may play a key role in the emergence of new STEC pathotypes (Schmidt, 2001). The *stx* genes may also be present in defective prophage genomes as is the case for *stx* of *Shigella dysenteriae* and for stx_1 in *E. coli* O157 : H7 strain EDL933 (McDonough *et al.*, 1999; Perna *et al.*, 2001). Although the number of published phage genomes and chromosomal integration sites is growing, only a few such sites are known for Stx phages. Stx2 phage BP-933W is inserted into the *wrbA* gene, which encodes a tryptophan repressor binding protein (Plunkett *et al.*, 1999). The cryptic Stx1-encoding phage CP-933V and the corresponding Stx1 phage from *E. coli* O157 : H7 strain RIMD0509952 are located in the *yehV* gene (Hayashi *et al.*, 2001; Perna *et al.*, 2001). The integration site of Stx2e phage P27 and ϕ-297 is the *yecE* gene. Other genes such as z2755 or *sbcB* have also been described as insertion sites for Stx phages (Herold *et al.*, 2004). Interestingly, *E. coli* O157 : H7 strains EDL933 and RIMD0509952 harbour a number of prophages and prophage-like structures, that occupy roughly 5–7 % of the *E. coli* O157 : H7 genome (Hayashi *et al.*, 2001; Perna *et al.*, 2001). Since the integration of large DNA molecules alters the host's genome structure and properties, characterization of such sites is an interesting tool for population biology studies.

In this study, we investigated the genetic structure and chromosomal integration site of the Stx1-converting prophage CP-1639. PCR analysis was performed to study the distribution of this integration site among *E. coli* strains of serogroups O26, O103, O111, O128, O145 and O157.

METHODS

Bacterial strains and plasmids. *E. coli* O111 : H^- strain 1639/77 was originally isolated from a patient with bloody diarrhoea and has been described previously (Schmidt *et al.*, 1999b; Unkmeir & Schmidt, 2000). Most of the other *E. coli* strains of serogroups O26, O103, O111, O128, O145 and O157 used in this study were isolated at the Institute of Hygiene and Microbiology, University of Würzburg, by routine methods as described by Friedrich *et al.* (2002) and Schmidt *et al.* (1999c), and were generously made available by H. Karch (Münster, Germany). *E. coli* strains C124, 889, HBI/01, CA92-7164, E22, CB132 and CB8260, were provided by our colleagues E. Oswald (Toulouse, France), H. Ball (Belfast, Northern Ireland), J. Fairbrother (Montreal, Canada) and L. Beutin (Berlin, Germany), who were members of the EU project QLK-2000-0060. Serotypes, *stx* genotypes and disease associations of these strains are described in Table 2. Some of the strains have already been published (Rüssmann *et al.*, 1994a, b; Schmidt *et al.*, 1999b; Zhang *et al.*, 2000). *E. coli* O157 : H7 strain EDL933 (O'Brien *et al.*, 1983) was used as a reference strain, and the *E. coli* K-12 derivative C600 as control strain. *E. coli* laboratory strains DH5α (Invitrogen) and XL-1 Blue MR (Stratagene) were employed as hosts for recombinant plasmids and cosmids, respectively. *E. coli* strains were grown aerobically at 37 °C in Luria–Bertani (LB) medium. Cosmid SuperCosI (Stratagene) and plasmid pK18 (Pridmore, 1987) were used as cloning vectors. When required, antibiotics were added at the following concentrations: 50 μg kanamycin ml^{-1} and 100 μg ampicillin ml^{-1}.

Standard DNA techniques. Induction of prophages and isolation of phage DNA were basically performed as described earlier (Schmidt *et al.*, 1999a). Plasmid and cosmid DNA was isolated using Qiagen plasmid kits, and genomic DNA was prepared with a Nucleobond AXG 100 kit (Macherey–Nagel) according to the manufacturer's instructions. DNA was digested with restriction endonucleases purchased from New England Biolabs. Restriction fragments were analysed by electrophoretic separation on 0·5–2 % agarose gels in Tris/borate/EDTA or Tris/acetate/EDTA buffer and stained with ethidium bromide. DNA fragments were purified from agarose gels using commercial gel extraction kits (Qiagen). Transformation with plasmid DNA was performed according to standard protocols (Sambrook *et al.*, 1989). For Southern blot hybridization, DNA was prepared, separated electrophoretically and subsequently transferred from agarose gels to Zeta Probe GT blotting membranes (Bio-Rad). Hybridization assays were performed with digoxigenin-labelled PCR-generated $stxA_1$- and $stxB_1$-specific DNA probes (Schmidt *et al.*, 1994) under stringent conditions in accordance with the manufacturer's instructions (Roche). Specific washing steps were performed twice, each for 5 min at 60 °C in 2× SSC (0·3 M sodium chloride/0·03 M sodium citrate) and 0·1 % (w/v) SDS.

PCR. Amplification was carried out in a total volume of 50 μl containing 200 μM of each deoxynucleoside triphosphate, 30 pmol of each primer, 5 μl of 10-fold-concentrated polymerase synthesis buffer, 1·5 M $MgCl_2$ and 2·5 U AmpliTaq DNA polymerase (Applied Biosystems). Oligonucleotides were designed with OLIGO 4.0 (National Biosciences), and purchased from Sigma-ARK and TibMolBiol. The first step of the PCR consisted of 5 min at 94 °C for denaturation of DNA, followed by 30 cycles each with denaturing (94 °C; 30 s), annealing (for temperature see Table 1; 60 s) and extension (72 °C; for time-course see Table 1) steps. A final extension step was carried out for 5 min at 72 °C. For a detailed description of primer sequences and PCR conditions see Table 1.

Genome library. A whole-genome library was prepared in SuperCosI cosmid vector from genomic DNA of *E. coli* 1639/77, which was partially restricted with *Sau*3AI according to the manufacturer's instructions (Stratagene).

Table 1. PCR primers and conditions

Primer			PCR conditions		Length of PCR product (bp)	Reference
Target	Designation	Nucleotide sequence	Annealing temp. (°C)	Extension time (s)		
$stxA_1$	LP30	5′-CAGTTAATGTGGTGGCGAAGC-3′	57	60	348	Cebula *et al.* (1995)
	LP31	5′-CACCAGACAATGTAACCGCTG-3′				
$stxB_1$	KS7	5′-CCCGGATCCATGAAAAAAACATTATTAATA-3′	52	40	285	Schmidt *et al.* (1994)
	KS8	5′-CCCGAATTCAGCTATTCTGAGTCAACG-3′				
cuo	p23E35	5′-GGTTATCATAGCAAATTCTGGTC-3′	59	180	2194	This study
	p23F82-fwd-1	5′-GTTTATGCCTTTCACTTTCTCAG-3′				
Int 1	B2656	5′-TGGGGGTAGTGGTAAAAATGAC-3′	55	60	449	This study
	B2657	5′-TAGTATTGTGTAACCGCTCATT-3′				
Int 2	ouo1	5′-CGAGACCTTTTATTATTGATT-3′	49	120	1328	This study
	B2657	5′-TAGTATTGTGTAACCGCTCATT-3′				
Int 3	ouo2	5′-GAAAACGATGGAGATAGGATT-3′	53	120	1668	This study
	pint	5′-ACGAACAGCGGCAGAGGAGTC-3′				
Int 4	L6for	5′-AGAGTTGTGCTGTATATAAAGCC-3′	59	90	589	This study
	pint	5′-ACGAACAGCGGCAGAGGAGTC-3′				
Int 5	1639int-for	5′-CGAGTGGGCAAGCGCTGC-3′	57	120	1558	This study
	orfB-rev	5′GCTGATGGTCGTACGGACC-3′				
Int 6	1639int-for	5′-CGAGTGGGCAAGCGCTGC-3′	57	120	1549	This study
	ssrA-rev	5′-CTTTAGCAGCTTAATAACCTGC-3′				

DNA sequencing and sequence analysis. Nucleotide sequencing was performed on both strands with an automated DNA sequencer (model 377; Applied Biosystems) initially by using universal-forward (5′-ACG ACG TTG TAA AAC GAC GGC CAG-3′) and universal-reverse primers (5′-TTC ACA CAG GAA ACA GCT ATG ACC-3′) for pK18, and super fwd (5′-GCA TTT ATC AGG GTT ATT GTC-3′) and super rev primers (5′-GGA AGT CAA CAA AAA GCA GAG-3′) for SuperCosI and then subsequently using customized primers. Each nucleotide was determined at least twice. Contigs were aligned with PREGAP 4, GAP 4 (Staden & MacLachlan, 1982) and BIOEDIT (Hall, 2004). ORF analysis was performed with EDITSEQ and MAPDRAW (DNASTAR, V.08) using ATG and GTG as start codons and the GENEMARK.HMM software (Lukashin & Borodovsky, 1998). Results of both programs were compared with BLAST analyses (National Center for Biotechnology Information, NCBI, Bethesda, USA), and ORFs with concordance in at least two of the applications applied are listed in the database entries AJ304858 and AJ831374.

RESULTS

Genetic structure of the stx_1-flanking regions

Previous work of our group performed on the *Q*-stx_1-*S* region of *E. coli* O111 : H$^-$ strain 1639/77 provided evidence for a particular prophage DNA region downstream of stx_1 (Unkmeir & Schmidt, 2000). In that study, we found homologies to *Salmonella enterica* phage Fels-2, and in addition, there was a lack of the expected lysis genes *S* and *R*. Due to these observations, we wanted to investigate whether Stx1 of *E. coli* O111 : H$^-$ strain 1639/77 is encoded in a typical Stx phage or in another phage type not described to date.

Although a bacteriophage could be induced and isolated from *E. coli* strain 1639/77, this was not a Stx-converting phage, indicating that the suspected Stx_1 phage is defective. For characterization of the genetic background of the stx_1 gene, we sequenced its flanking regions. To achieve this, a genomic library was created from *E. coli* O111 : H$^-$ strain 1639/77 comprising 1200 cosmid clones. From eight of these clones we found PCR products after amplification with $stxA_1$- and $stxB_1$-specific primers, and four of these were processed further. These clones were digested with *Eco*RI and *Hin*dIII, and restriction fragments were subcloned in pK18. Using these subclones and the original recombinant cosmids as templates for sequencing reactions, we could determine a sequence of 50 625 bp of continuous chromosomal DNA of *E. coli* 1639/77 DNA, which flanks stx_1 on both sides. Sequence analysis of this DNA region with MAPDRAW, GENEMARK.HMM and NCBI BLAST revealed 70 ORFs (Fig. 1, GenBank accession no. AJ304858). Further ORF analysis demonstrated that stx_1 is obviously located in a defective prophage genome, which has been designated CP-1639. The genetic organization of the sequenced region is depicted in Fig. 1.

Chromosomal integration site and genetic structure of CP-1639

ORFs 1–4 are nearly identical to genes *ygaF*, b2659, b2658 and b2657, which occur consecutively in the *E. coli* K-12 chromosome. The functions of these genes are not yet completely understood. ORFs 5–7 have no homologues in the *E. coli* K-12 chromosome. ORF 5 putatively encodes a protein with an amino acid identity of 41 % to an ATPase-, histidine kinase-, DNA gyrase B-, and HSP90-like domain

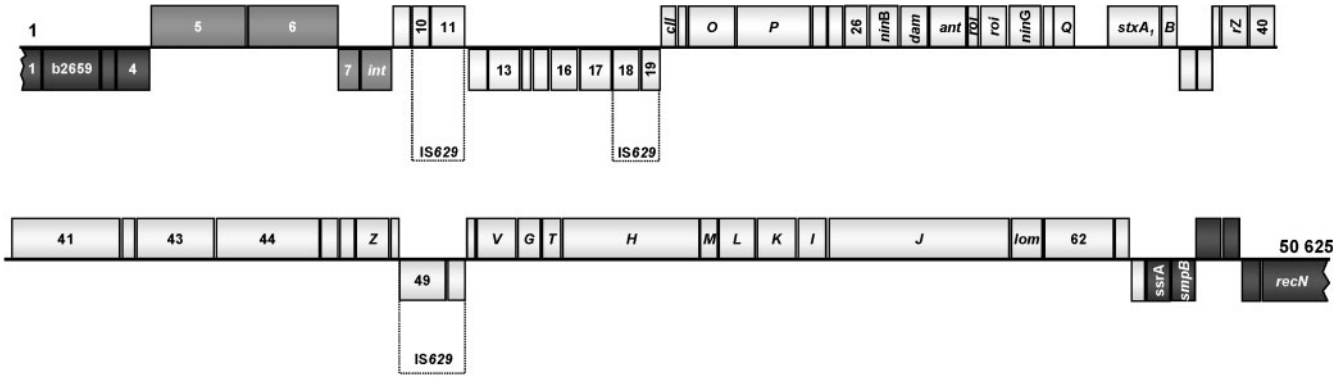

Fig. 1. Genetic map (to scale) of the analysed 50 625 bp chromosomal region of *E. coli* O111 : H$^-$ strain 1639/77. Bars above the black line indicate ORFs in rightward (5′–3′) transcription direction and bars below the line depict ORFs transcribed in leftward direction. Dark grey bars depict chromosomal genes, mid-grey bars an integrative element and light grey bars those ORFs which belong to CP-1639. ORFs are described in more detail in the database entry AJ304858. Numbers and gene designations correspond to those in the database, where the complete numberings are listed from EC_CP1639_01 to EC_CP1639_70, e.g. '5' in this figure corresponds to EC_CP1639_05 in the database entry. IS*629* elements are labelled.

protein of *Desulfovibrio vulgaris*. Lower sequence similarity was found to bacterial histidine sensory kinases of different micro-organisms. Proteins of this group belong to bacterial two-component regulatory systems, which transmit environmental signals into the bacterial cell in order to modulate gene activity (Gross *et al.*, 1989). ORF 6 also demonstrated weak homology to *D. vulgaris* DNA but ORF 7 was not similar to any protein present in the database. These three ORFs constitute a 4084 kb DNA region with a G+C content of 34 mol% in comparison to 50 mol% for the *E. coli* K-12 chromosome. The deduced 194 amino acids of ORF 8 were 55 % identical, by BLAST analysis, to the C-terminal 194 amino acids of the 410 amino acid prophage CP4-57 IntA integrase of *Yersinia pestis* strain KIM (Deng *et al.*, 2002), and showed lower homology with the C-termini of numerous other phage integrases. It appears that the N-terminus is truncated by 216 amino acids in the ORF 8 encoded protein. Further analysis by the NCBI CONSERVED DOMAIN SEARCH demonstrated that this truncated IntA protein belongs to the P4 family of integrase genes.

ORF 9 is related to ORFB of IS*3* and is followed by an intact IS*629* element covering base pair positions 7832–9141 (Fig. 1). The sequence is identical to a published IS*629* element (Matsutani & Ohtsubo, 1990). Subsequently, we found six phage-related ORFs (12–17), four of which appeared to be related to the corresponding region of phage P22-related phages Sf6 and HK022 (Casjens *et al.*, 2004). ORFs 18 and 19 represent genes of a defective IS*629* element, which is truncated by 423 bp at its 3′-end.

The area starting with *cII* (ORF 20) and ending with a *ninH*-like gene (ORF 32) is similar to the corresponding region of Stx2-encoding phage BP-933W. Interestingly, the stop codon of *ninG* and the start codon of *ninH* overlap as described for other Stx and lambdoid phages (Karch *et al.*, 1999). Analysis of ORFs 30 and 31, both with homology to different regions of the *roi* gene, indicates that two truncated *roi*-like ORFs are present in CP-1639.

The next region to consider is the *Q-stx*-lysis region. This region includes the antiterminator gene *Q* and a lysis-associated gene *rZ*, and is shorter than comparable regions in other Stx phages (Fig. 1). We did not find *S* and *R* homologues in this region. The Stx1A and B subunit genes of CP-1639 are identical to published stx_1 genes.

ORFs 40–62 are related to DNA packaging and head and tail morphology. These genes are highly homologous to the corresponding genes of other Stx phages. However, we could not identify a gene specifying a capsid protein. The order of tail fibre genes from gene Z to J is also found in the same direction in the cryptic phage CP-933U of *E. coli* O157 : H7 strain EDL933 (Perna *et al.*, 2001) and phage λ. The expected tail fibre gene *U* was not present in an intact form. It appears to be disrupted by a complete IS*629* element (Fig. 1). This region contains also a virulence-related ORF, encoding a *lom*-like outer-membrane protein precursor.

The last gene of the prophage genome was ORF 64, which also occurs in a similar form in Stx1 phage CP-933V (Perna *et al.*, 2001). This gene marks the junction of phage and chromosome, the latter of which starts with the *ssrA* gene. ORFs 65–70 are homologues to chromosomal genes which are present in *E. coli* K-12 (Blattner *et al.*, 1997).

Characterization of the chromosomal integration site of CP-1639 and identification of an integrative element

The particular structure of the left end of the prophage genome (Fig. 1) and the presence of genes of unknown

origin (ORFs 5, 6 and 7) between phage CP-1639 and the chromosome impeded the determination of the phage attachment site and thereby the length of the phage. In our initial investigations, we hypothesized that the *intA* gene (ORF 8) belonged to CP-1639. However, the presence of an *int* gene at the corresponding site in *E. coli* K-12, the homology to CP4-57 integrases and the orientation of *int* transcription suggest that the genes of unknown origin (ORF 5, 6 and 7) and the truncated *intA* belong together and form a separate genetic unit, hereafter designated an 'integrative element'.

If this hypothesis is true, we should be able to find *E. coli* strains harbouring such an integrative element consisting of ORFs 5, 6, 7 and *intA*. Moreover, it should be located close to *ssrA* without being interrupted by a prophage genome. We examined this hypothesis by PCR using primers for the detection of CP-1639 and this particular genetic element as depicted in Fig. 2. As templates for PCR we chose a number of STEC of different serotypes as well as Stx-negative *E. coli* (Table 2). The presence of the integrative element was assumed when Int 2, cuo, Int 3, Int 4 and Int 6 PCR reactions were positive, and Int 1 was negative (Fig. 2).

E. coli O111 : H$^-$ strains showed a uniform PCR pattern, indicating the same genetic structure as observed for *E. coli* 1639/77, with an integrative element connected to a prophage. Only strain 4556/99 was Int 5 PCR-negative, suggesting a prophage structure different from CP-1639 (Table 2).

Nineteen strains of serotypes O103, O26 and O128 showed an insertion of foreign DNA in this region (Int 1-PCR negative). In addition, all these strains were Int 5 negative, and 17 out of 19 strains were Int 6 positive (Table 2). This gene order indicates the presence of an integrative element which is not further associated with a phage-like structure. Here, the *intA* gene appeared to be close to *ssrA*. Two O103 strains were Int 5 and Int 6 negative. This could mean that the integrative element is associated with a further prophage which did not contain an IS*629* element at the corresponding site.

Some strains showed varying PCR results, with one or more PCR reactions, which did not fit the general pattern. Either there were sequence variations resulting in a lack of primer-binding sites, fragments of the suggested genes were missing or the presence of IS-elements caused enlarged PCR products. Int 1 PCR gave no evidence for the presence of foreign DNA between b2657 and *ygaR* in eight *E. coli* O157 : H7 and O145 : H$^-$ strains, suggesting the same genetic structure as in *E. coli* K-12-strain MG1655 in this region. In three *E. coli* O157 : H$^-$ strains, Int 1 and Int 4 were positive, indicating the presence of a short piece of foreign DNA related to ORF 7 and *intA*.

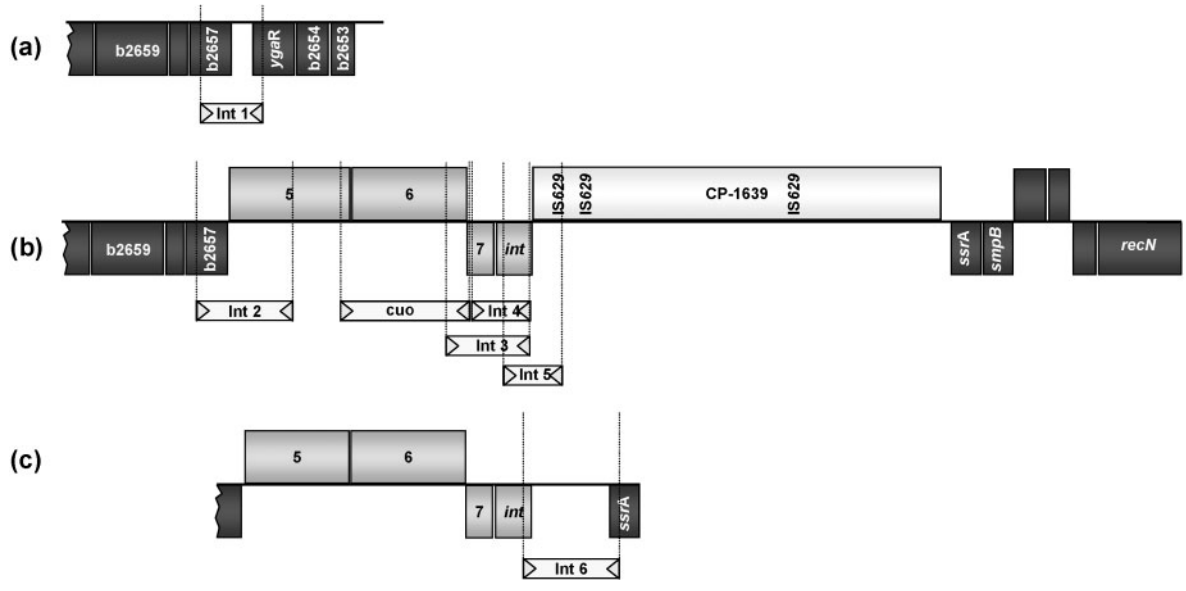

Fig. 2. PCR strategy for the detection of CP-1639, the association with an integrative element, the unoccupied integration site and the presence of an integrative element without prophage association. PCR primers and strategy are described in detail in the text. (a) The Int 1 PCR product is an indicator that the chromosomal region investigated is similar to that of *E. coli* K-12 and no foreign DNA is present. (b) Structure of the leftward end of CP-1639, the remnant of an integrative element and the junction to the chromosome. This structure can be assumed when Int 2, cuo, Int 3, Int 4 and Int 5 PCR products appear after amplification, and when Int 1 and Int 6 are negative. (c) The presence of an integrative element without prophage association is indicated when Int 2, cuo, Int 3, Int 4 and Int 6 PCR products appear, and Int 1 and Int 5 PCR are negative.

Table 2. Serotypes, origin, *stx* genotype and PCR results of STEC and non-STEC strains used for analysis of the phage integration site

Strain	Serotype*	Origin†	PCR results‡								
			stx$_1$	*stx*$_2$	Int 1	Int 2	cuo	Int 3	Int 4	Int 5	Int 6
C600	ND	Lab.	−	−	+	−	−	−	−	−	+
1639/77	O111 : H$^-$	D	+	−	−	+	+	+	+	+	−
4417/96	O111 : H$^-$	HUS	+	−	−	+	+	+	+	+	−
5380/96	O111 : H$^-$	D	+	−	−	+	+	+	+	+	−
6037/96	O111 : H$^-$	HUS	+	+	−	+	+	+	+	+	−
5556/97	O111 : H$^-$	D	+	−	−	+	+	+	+	+	−
6366/97	O111 : H$^-$	HUS	+	+	−	+	+	+	+	+	−
4556/99	O111 : H$^-$	D	+	−	−	+	+	+	+	−	−
2905/96	O103 : H2	D	+	−	−	+	+	+	+	−	+
5714/96	O103 : H2	D	+	−	−	+§	+	+	+	−	−
3943/97	O103 : H2	HUS	+	−	−	+§	+	+	+	−	+
2491/99	O103 : H2	D	+	−	−	+§	+	+	+	−	+
C124	O103 : H2	HR	−	−	−	+§	+	+	+	−	+
889	O103 : H2	Chicken	−	−	−	+§	+	+	+	−	+
CA92-7164	O103 : [H7]	Dog	−	−	−	−	+	+	+	−	+
HBI/01	O103 : H2	Meat	+	−	−	+	+	+	+	−	+
2636/97	O103 : H2	HUS	−	+	−	+§	+	+	+	−	+
2791/97	O103 : H2	HUS	−	+	−	+§	−	+	+	−	+
1909/97	O103 : H18	D	+	−	−	+§	+‖	−	+	−	−
5577/96	O103 : H2	D	+	−	−	−	+	+	+	−	+
E22	O103 : H2	RD	−	−	−	+§	−	+	+	−	+
3967/96	O26 : H$^-$	AC	+	−	−	+	+	+	+	−	+
6105/96	O26 : H$^-$	D	+	−	−	+§	+	+	+	−	+
4740/97	O26 : H11	D	+	−	−	+	+	+	+	−	+
2569/98	O26 : H−	D	+	+	−	+§	+	+	+	−	+
CB132	O128 : H2	D	−	−	−	+§	−	+	+	−	+
CB8260	O128 : H2	AC	−	−	−	+§	+	+	+	−	+
EDL933	O157 : H7	HC	+	+	+	−	−	−	−	−	−
6790/96	O157 : H$^-$	HUS	−	−	+	−	−	−	+	−	−
431/97	O157 : H$^-$	D	−	−	+	−	−	−	+	−	+
659/97	O157 : H$^-$	D	−	−	+	−	−	−	+	−	+
1249/87	O157 : H7	HUS	−	+	+	−	−	−	−	−	−
19665/91	O157 : H7	D	−	−	+	−	−	−	−	−	−
4392/97	O145 : H$^-$	HUS	−	+	−	−	−	−	−	−	−
3485/99	O145 : H$^-$	HUS	−	+	+	−	−	−	−	+¶	+#
3517/99	O145 : H$^-$	HUS	−	+	+	−	−	−	−	+¶	+#
4557/99	O145 : H$^-$	HUS	−	+	+	−	−	−	−	+¶	+#
4672/99	O145 : H$^-$	HUS	+	+	+	−	−	−	−	+¶	+#

*ND, Not determined; [H7], contains the H7 *fliC* gene but is non-motile.

†AC, asymptomatic carrier; D, patients with diarrhoea; HC, haemorrhagic colitis; HR, healthy rabbit; HUS, patients with haemolytic–uraemic syndrome; RD, rabbit with diarrhoea.

‡+, PCR product obtained; −, no PCR product obtained.

§2·5 kb. ‖3·5 kb. ¶About 600 bp. #550 bp.

To investigate the sequence of the integrative element, we analysed the region between b2657 and *ssrA* in *E. coli* O103 : H2 strain 2905/96 (GenBank AJ831374). This showed that ORFs 5, 6 and 7 are present with sequence similarities to *E. coli* strain 1639/77 of >99 %, and a truncated *intA* identical to the *E. coli* 1639/77 *intA* gene was also identified (Fig. 2c). A short DNA region of 814 bp containing gene fragments with homology to an IS*3*-related gene and a Salmonella integrase was found between *intA* and *ssrA*. This suggests that this region was subject to intensive recombination. This region ends up with the 3′-end of *ssrA*. No intact phage sequences were found.

At the right side, the prophage sequence ends with ORF 64. ORF 65 represents the *ssrA* gene, encoding a small, stable tmRNA. The tmRNA is well known as an integration site for phages and other integrative elements, and serves also as an integration site for CP-1639. Williams (2003) analysed a number of enterobacterial integrative elements which have been inserted at the 3′-end of the tmRNA gene *ssrA*. We compared the rightward CP-1639 sequence including its junction to the bacterial chromosome with the corresponding sequences of *E. coli* O157 : H7 strain EDL933, *E. coli* K-12 strain MG1655, *E. coli* O103 : H2 strain 2905/96 and *Salmonella enterica* serovar Typhimurium strain LT2, according to an alignment published by Williams (2003). Seven-basepair crossover segments found in attachment sites were identified at the 3′-end of *ssrA*, which are present in all *E. coli* sequences but not in *S. enterica* serovar Typhimurium strain LT2. Downstream of the attachment sites, factor-independent terminators were characterized by Williams (2003). These sequences could also be identified in *E. coli* strains 1639/77, EDL933, MG1655, 2905/96 and *S. enterica* serovar Typhimurium LT2. From this alignment it can be concluded that the rightward attachment site of *E. coli* 1639/77 is identical to that of *E. coli* O157 : H7 strain EDL933 and is located at basepair positions 46 686–46 870.

Comparison of the integration region in different *E. coli* strains

To get more information on the site of CP-1639 integration, we compared this site in *E. coli* strain 1639/77 with published sequences of the corresponding regions of *E. coli* K-12 strain MG1655 (NC_000913) and *E. coli* O157 : H7 strain EDL933 (NC_002655). The prophage CP-1639 is inserted between the *E. coli* gene b2657 and the 3′-end of *ssrA* as described above. In *E. coli* K-12 strain MG1655, b2657 is located at position 2 785 000 and *ssrA* at 2 752 773. There is a region of about 32 kb which is not present in *E. coli* 1639/77. The *ssrA* gene is associated with a CP4-57 integrase, followed by a number of CP4-57-related genes. A number of genes designated *yfi* fills the space between the remnant of CP4-57 and *ypiA*. These genes are neither present in strain 1639/77 nor in strain EDL933. The function and origin of the *yfi* genes is not known. In *E. coli* O157 : H7 strain EDL933, O-island #108 (CP-933Y) is close to *ssrA*, encompassing a region of approximately 21·5 kb, containing phage-related genes. At the other side, *ypiA* is the first chromosomal gene, similar to *E. coli* K-12 but different from *E. coli* O111 : H^- 1639/77, where b2657 is the first chromosomal gene.

DISCUSSION

In a recent study, Unkmeir & Schmidt (2000) investigated the *stx*-flanking regions of a larger set of Stx-producing *E. coli* and *Shigella dysenteriae* type1 strains. A short region of the *E. coli* O111 : H^- strain 1639/77 *stx*-flanking region was analysed in that study and provided evidence for a particular prophage structure. All *E. coli* O111 strains included in those experiments carried a similar region downstream of the stx_1 gene, which is distinct from that of Stx phages present in other STEC serotypes (Unkmeir & Schmidt, 2000). Our findings confirm previous conclusions that although Stx phages share a similar morphology, a similar mode of replication and a similar genomic structure, they can be heterogeneous at the nucleotide and deduced amino acid level (Recktenwald & Schmidt, 2002).

STEC phages belong to the group of lambdoid phages, which displays a modular structure and a large extent of mosaicism (Campbell, 1994; Koch *et al.*, 2003; Makino *et al.*, 1999; Muniesa *et al.*, 2000; Neely & Friedman, 1998; Plunkett *et al.*, 1999; Sato *et al.*, 2003a, b). Although a number of Stx phages have genomes different from *E. coli* O157 : H7 phages BP-933W and H19J, the mode of *stx* expression seems to be uniquely phage growth-cycle-dependent (Neely & Friedman, 1998; Unkmeir & Schmidt, 2000; Wagner & Waldor, 2002).

It has been speculated that *stx* in some strains is expressed from a chromosomal locus, since Stx phages could not be isolated (O'Brien & Holmes, 1987). The current point of view is that it is more likely that *stx* is expressed from defective prophages in these cases. This seems also to be true for CP-1639, which can not be induced with standard methods. On a closer look on the molecular structure of CP-1639, it seems to be immobilized due to the lack of important phage genes.

The determination of the length of the CP-1639 prophage genome is difficult since the left prophage end could not be identified with certainty. Due to rearrangements and deletions in the recombination region, we can only speculate about the extent of CP-1639. However, the DNA region between the putative CP4-57 integrase remnant and the chromosomal *ssrA* represents most conclusively the cryptic prophage and encompasses 39 445 bp.

The presence of two IS*629* elements at the left side and a number of genes which seemed to be merged together from other phages let us suggest that this is a region of intensive recombination, which probably has caused a deletion of integrase and excisionase genes. The similarity of ORFs 12–17 to Salmonella phage P22 genes and the presence of *ant* indicate a connection between CP-1639 and Salmonella phages. Although single genes which are similar to the recombination and immunity region of Salmonella phage P22 are present, it seems that important recombination and immunity genes are missing. By inspection of the region 3′ of *stx*, only an *Rz* gene is present, which putatively encodes an endopeptidase. Moreover, no holin genes could be identified, and a gene encoding a putative head protein is absent also. The tail fibre components Z to J follow the gene order observed in phage λ and are most similar to the corresponding genes of *E. coli* O157 : H7 phage CP-933U (Perna *et al.*, 2001). All in all, three IS*629* elements have been found in the CP-1639. Since at least two of these

elements appeared functional, downstream effects can be envisaged that alter host transcription. The *stx* gene of *S. dysenteriae* type 1 is also encoded in a phage remnant. Here, the *stx* operon is flanked by a number of insertion elements and only a small number of phage genes remained (McDonough *et al.*, 1999) so that the prophage character of this element is archetypal. On the basis of sequence analysis, CP-1639 is considered to be a victim of a single or multiple recombination events, that left behind a handicapped mosaic phage genome.

The *ssrA* gene has been shown to be an integration site for a number of integrative elements. *ssrA* encodes a tmRNA, a molecule that possess tRNA and mRNA character. With its tRNA function it can transfer an attached amino acid residue to a growing protein in the ribosome, although this is not anticodon-mediated. With its mRNA properties it has a role in rescuing ribosomes from malfunction (Williams, 2003). In a recent review, Williams (2003) demonstrated that the *ssrA* site is occupied in different enterobacteriaceae with a number of different integrative elements which contain *ssrA*-specific integrases. One group is the P4 integrase family, and the truncated *intA* gene identified in *E. coli* 1639/77 belongs to this family. Since we have found an integrative element without associated Stx phages in a number of strains, this may be considered as an ancient integrative-element, which has served as integration site for CP-1639. The mechanism of insertion and deletion of important functions of the prophage can not be deduced clearly. However, arrays of such integrative elements have been reported to occur at *ssrA*. Particular integrase subfamilies can trigger integration at *ssrA* and may leave a reconstituted *ssrA* site, which can serve as a convenient integration site for a further integrative element. If this occurred in *E. coli* 1639/77, the tandem array was probably due to recombination of these arrayed elements and subsequent deletion events may have taken place. This could be the reason for that particular gene arrangement found in CP-1639.

The results of this study clearly show that the *E. coli* genome contains highly variable capturing sites that may be used by mobile genetic elements for stable entry in the chromosome of pathogens. Insertion, mutation and deletion events obviously led to a defective Stx-prophage. These occurrences can be considered as pathoadaptive mutations. It is not known what benefit the cell derives from immobilization of *stx*. It is not likely that plasmids, transposons or temperate phages are purely parasitic and can be maintained in the chromosome without paying for their dinner. The current view is that elements that are not purely parasitic bear genes that may be beneficial for the host cells, at least in certain cases (Levin & Bergstrom, 2000). Further work on the ecology and physiology of Stx phages and their hosts is needed to gain a better understanding of the mechanisms and evolutionary forces that extend the genetic spectrum of *E. coli* and other bacterial pathogens.

ACKNOWLEDGEMENTS

This work was supported by grants of Deutsche Forschungsgemeinschaft, grant Schm1360/1-3. We thank Stefanie Müksch for skilful technical assistance.

REFERENCES

Acheson, D. W., Reidl, J., Zhang, X., Keusch, G. T., Mekalanos, J. J. & Waldor, M. K. (1998). In vivo transduction with Shiga toxin 1-encoding phage. *Infect Immun* **66**, 4496–4498.

Beutin, L., Krause, G., Zimmermann, S., Kaulfuss, S. & Gleier, K. (2004). Characterization of Shiga toxin-producing *Escherichia coli* strains isolated from human patients in Germany over a 3-year period. *J Clin Microbiol* **42**, 1099–1108.

Bielaszewska, M., Fell, M., Greune, L., Prager, R., Fruth, A., Tschäpe, H., Schmidt, M. A. & Karch, H. (2004). Characterization of cytolethal distending toxin genes and expression in shiga toxin-producing *Escherichia coli* strains of non-O157 serogroups. *Infect Immun* **72**, 1812–1816.

Bitzan, M. & te Loo, D. M. (2003). Interaction of Shiga toxin with endothelial cells. *Methods Mol Med* **73**, 243–262.

Blanco, J. E., Blanco, M., Alonso, M. P., Mora, A., Dahbi, G., Coira, M. A. & Blanco, J. (2004). Serotypes, virulence genes, and intimin types of Shiga toxin (verotoxin)-producing *Escherichia coli* isolates from human patients: prevalence in Lugo, Spain, from 1992 through 1999. *J Clin Microbiol* **42**, 311–319.

Blattner, F. R., Plunkett, G., III, Bloch, C. A. & 14 other authors (1997). The complete genome sequence of *Escherichia coli* K-12. *Science* **277**, 1453–1474.

Brunder, W., Schmidt, H., Frosch, M. & Karch, H. (1999). The large plasmids of Shiga-toxin-producing *Escherichia coli* (STEC) are highly variable genetic elements. *Microbiology* **145**, 1005–1014.

Campbell, A. (1994). Comparative molecular biology of lambdoid phages. *Annu Rev Microbiol* **48**, 193–222.

Caprioli, A., Tozzi, A. E., Rizzoni, G. & Karch, H. (1997). Non-O157 Shiga toxin-producing *Escherichia coli* infections in Europe. *Emerg Infect Dis* **3**, 578–579.

Casjens, S., Winn-Stapley, D. A., Gilcrease, E. B., Morona, R., Kuhlewein, C., Chua, J. E., Manning, P. A., Inwood, W. & Clark, A. J. (2004). The chromosome of *Shigella flexneri* bacteriophage Sf6: complete nucleotide sequence, genetic mosaicism, and DNA packaging. *J Mol Biol* **339**, 379–394.

Cebula, T. A., Payne, W. L. & Feng, P. (1995). Simultaneous identification of strains of *Escherichia coli* serotype O157 : H7 and their Shiga-like toxin type by mismatch amplification mutation assay-multiplex PCR. *J Clin Microbiol* **33**, 248–250.

Deng, W., Burland, V., Plunkett, G., III & 18 other authors (2002). Genome sequence of *Yersinia pestis* KIM. *J Bacteriol* **184**, 4601–4611.

Friedrich, A. W., Bielaszewska, M., Zhang, W. L., Pulz, M., Kuczius, T., Ammon, A. & Karch, H. (2002). *Escherichia coli* harboring Shiga toxin 2 gene variants: frequency and association with clinical symptoms. *J Infect Dis* **185**, 74–84.

Friedrich, A. W., Borell, J., Bielaszewska, M., Fruth, A., Tschäpe, H. & Karch, H. (2003). Shiga toxin 1c-producing *Escherichia coli* strains: phenotypic and genetic characterization and association with human disease. *J Clin Microbiol* **41**, 2448–2453.

Gross, R., Arico, B. & Rappuoli, R. (1989). Families of bacterial signal-transducing proteins. *Mol Microbiol* **3**, 1661–1667.

Hall, T. A. (2004). BioEdit: a user-friendly biological sequence alignment editor and analysis program for Windows 95/98/NT. *Nucleic Acids Symp Ser* **41**, 95–98.

Hayashi, T., Makino, K., Ohnishi, M. & 19 other authors (2001). Complete genome sequence of enterohemorrhagic *Escherichia coli* O157 : H7 and genomic comparison with a laboratory strain K-12. *DNA Res* **8**, 11–22.

Herold, S., Karch, H. & Schmidt, H. (2004). Shiga toxin-encoding phages–genomes in motion. *Int J Med Microbiol* **294**, 115–121.

Jacewicz, M. S., Acheson, D. W., Binion, D. G., West, G. A., Lincicome, L. L., Fiocchi, C. & Keusch, G. T. (1999). Responses of human intestinal microvascular endothelial cells to Shiga toxins 1 and 2 and pathogenesis of hemorrhagic colitis. *Infect Immun* **67**, 1439–1444.

Kaneko, K., Kiyokawa, N., Ohtomo, Y., Nagaoka, R., Yamashiro, Y., Taguchi, T., Mori, T., Fujimoto, J. & Takeda, T. (2001). Apoptosis of renal tubular cells in Shiga-toxin-mediated hemolytic uremic syndrome. *Nephron* **87**, 182–185.

Kaper, J. B. (1998). Enterohemorrhagic *Escherichia coli*. *Curr Opin Microbiol* **1**, 103–108.

Karch, H., Schmidt, H., Janetzki-Mittmann, C., Scheef, J. & Kröger, M. (1999). Shiga toxins even when different are encoded at identical positions in the genomes of related temperate bacteriophages. *Mol Gen Genet* **262**, 600–607.

Koch, C., Hertwig, S. & Appel, B. (2003). Nucleotide sequence of the integration site of the temperate bacteriophage 6220, which carries the Shiga toxin gene stx_{1ox3}. *J Bacteriol* **185**, 6463–6466.

Leung, P. H., Peiris, J. S., Ng, W. W., Robins-Browne, R. M., Bettelheim, K. A. & Yam, W. C. (2003). A newly discovered verotoxin variant, VT2g, produced by bovine verocytotoxigenic *Escherichia coli*. *Appl Environ Microbiol* **69**, 7549–7553.

Levin, B. R. & Bergstrom, C. T. (2000). Bacteria are different: observations, interpretations, speculations, and opinions about the mechanisms of adaptive evolution in prokaryotes. *Proc Natl Acad Sci U S A* **97**, 6981–6985.

Lukashin, A. V. & Borodovsky, M. (1998). GeneMark.hmm: new solutions for gene finding. *Nucleic Acids Res* **26**, 1107–1115.

Makino, K., Yokoyama, K., Kubota, Y. & 15 other authors (1999). Complete nucleotide sequence of the prophage VT2-Sakai carrying the verotoxin 2 genes of the enterohemorrhagic *Escherichia coli* O157 : H7 derived from the Sakai outbreak. *Genes Genet Syst* **74**, 227–239.

Matsutani, S. & Ohtsubo, E. (1990). Complete sequence of IS*629*. *Nucleic Acids Res* **18**, 1899.

McDonough, M. A., Butterton, J. R., Macrina, F. L. & 14 other authors (1999). Spontaneous tandem amplification and deletion of the shiga toxin operon in *Shigella dysenteriae* 1. *Mol Microbiol* **34**, 1058–1069.

Melton-Celsa, A. R., Darnell, S. C. & O'Brien, A. D. (1996). Activation of Shiga-like toxins by mouse and human intestinal mucus correlates with virulence of enterohemorrhagic *Escherichia coli* O91 : H21 isolates in orally infected, streptomycin-treated mice. *Infect Immun* **64**, 1569–1576.

Misselwitz, J., Karch, H., Bielazewska, M., John, U., Ringelmann, F., Ronnefarth, G. & Patzer, L. (2003). Cluster of hemolytic-uremic syndrome caused by Shiga toxin-producing *Escherichia coli* O26 : H11. *Pediatr Infect Dis* **22**, 349–354.

Muniesa, M., Recktenwald, J., Bielaszewska, M., Karch, H. & Schmidt, H. (2000). Characterization of a shiga toxin 2e-converting bacteriophage from an *Escherichia coli* strain of human origin. *Infect Immun* **68**, 4850–4855.

Nataro, J. P. & Kaper, J. B. (1998). Diarrheagenic *Escherichia coli*. *Clin Microbiol Rev* **11**, 142–201.

Neely, M. N. & Friedman, D. I. (1998). Functional and genetic analysis of regulatory regions of coliphage H-19B: location of Shiga-like toxin and lysis genes suggest a role for phage functions in toxin release. *Mol Microbiol* **28**, 1255–1267.

O'Brien, A. D. & Holmes, R. K. (1987). Shiga and Shiga-like toxins. *Microbiol Rev* **51**, 206–220.

O'Brien, A. D., Lively, T. A., Chang, T. W. & Gorbach, S. L. (1983). Purification of *Shigella dysenteriae* 1 (Shiga)-like toxin from *Escherichia coli* O157 : H7 strain associated with haemorrhagic colitis. *Lancet* **2**, 573.

O'Brien, A. D., Newland, J. W., Miller, S. F., Holmes, R. K., Smith, H. W. & Formal, S. B. (1984). Shiga-like toxin-converting phages from *Escherichia coli* strains that cause hemorrhagic colitis or infantile diarrhea. *Science* **226**, 694–696.

O'Loughlin, E. V. & Robins-Browne, R. M. (2001). Effect of Shiga toxin and Shiga-like toxins on eukaryotic cells. *Microbes Infect* **3**, 493–507.

Perna, N. T., Plunkett, G., III, Burland, V. & 25 other authors (2001). Genome sequence of enterohaemorrhagic *Escherichia coli* O157 : H7. *Nature* **409**, 529–533.

Pierard, D., Muyldermans, G., Moriau, L., Stevens, D. & Lauwers, S. (1998). Identification of new verocytotoxin type 2 variant B-subunit genes in human and animal *Escherichia coli* isolates. *J Clin Microbiol* **36**, 3317–3322.

Plunkett, G., III, Rose, D. J., Durfee, T. J. & Blattner, F. R. (1999). Sequence of Shiga toxin 2 phage 933W from *Escherichia coli* O157 : H7: Shiga toxin as a phage late-gene product. *J Bacteriol* **181**, 1767–1778.

Pridmore, R. D. (1987). New and versatile cloning vectors with kanamycin-resistance marker. *Gene* **56**, 309–312.

Recktenwald, J. & Schmidt, H. (2002). The nucleotide sequence of Shiga toxin (Stx) 2e-encoding phage phiP27 is not related to other Stx phage genomes, but the modular genetic structure is conserved. *Infect* **70**, 1896–1908.

Riley, L. W., Remis, R. S., Helgerson, S. D. & 9 other authors (1983). Hemorrhagic colitis associated with a rare *Escherichia coli* serotype. *N Engl J Med* **308**, 681–685.

Rüssmann, H., Schmidt, H., Caprioli, A. & Karch, H. (1994a). Highly conserved B-subunit genes of Shiga-like toxin II variants found in *Escherichia coli* O157 strains. *FEMS Microbiol Lett* **118**, 335–340.

Rüssmann, H., Schmidt, H., Heesemann, J., Caprioli, A. & Karch, H. (1994b). Variants of Shiga-like toxin II constitute a major toxin component in *Escherichia coli* O157 strains from patients with haemolytic uraemic syndrome. *J Med Microbiol* **40**, 338–343.

Sambrook, J., Fritsch, E. F. & Maniatis, T. (1989). *Molecular Cloning: a Laboratory Manual*. Cold Spring Harbor, NY: Cold Spring Harbor Laboratory.

Sato, T., Shimizu, T., Watarai, M., Kobayashi, M., Kano, S., Hamabata, T., Takeda, Y. & Yamasaki, S. (2003a). Distinctiveness of the genomic sequence of Shiga toxin 2-converting phage isolated from *Escherichia coli* O157 : H7 Okayama strain as compared to other Shiga toxin 2-converting phages. *Gene* **309**, 35–48.

Sato, T., Shimizu, T., Watarai, M., Kobayashi, M., Kano, S., Hamabata, T., Takeda, Y. & Yamasaki, S. (2003b). Genome analysis of a novel Shiga toxin 1 (Stx1)-converting phage which is closely related to Stx2-converting phages but not to other Stx1-converting phages. *J Bacteriol* **185**, 3966–3971.

Schmidt, H. (2001). Shiga-toxin-converting bacteriophages. *Res Microbiol* **152**, 687–695.

Schmidt, H., Rüssmann, H., Schwarzkopf, A., Aleksic, S., Heesemann, J. & Karch, H. (1994). Prevalence of attaching and effacing *Escherichia coli* in stool samples from patients and controls. *Zentbl Bakteriol* **281**, 201–213.

Schmidt, H., Bielaszewska, M. & Karch, H. (1999a). Transduction of enteric *Escherichia coli* isolates with a derivative of Shiga toxin 2-encoding bacteriophage phi3538 isolated from *Escherichia coli* O157 : H7. *Appl Environ Microbiol* **65**, 3855–3861.

Schmidt, H., Geitz, C., Tarr, P. I., Frosch, M. & Karch, H. (1999b). Non-O157 : H7 pathogenic Shiga toxin-producing *Escherichia coli*: phenotypic and genetic profiling of virulence traits and evidence for clonality. *J Infect Dis* **179**, 115–123.

Schmidt, H., Scheef, J., Huppertz, H. I., Frosch, M. & Karch, H. (1999c). *Escherichia coli* O157 : H7 and O157 : H− strains that do not produce Shiga toxin: phenotypic and genetic characterization of isolates associated with diarrhea and hemolytic-uremic syndrome. *J Clin Microbiol* **37**, 3491–3496.

Schmidt, H., Scheef, J., Morabito, S., Caprioli, A., Wieler, L. H. & Karch, H. (2000). A new Shiga toxin 2 variant (Stx2f) from *Escherichia coli* isolated from pigeons. *Appl Environ Microbiol* **66**, 1205–1208.

Schmitt, C. K., McKee, M. L. & O'Brien, A. D. (1991). Two copies of Shiga-like toxin II-related genes common in enterohemorrhagic *Escherichia coli* strains are responsible for the antigenic heterogeneity of the O157 : H− strain E32511. *Infect Immun* **59**, 1065–1073.

Scotland, S. M., Smith, H. R., Willshaw, G. A. & Rowe, B. (1983). Vero cytotoxin production in strain of *Escherichia coli* is determined by genes carried on bacteriophage. *Lancet* **2**, 216.

Staden, R. & MacLachlan, A. D. (1982). Codon preference and its use in identifying protein coding regions in long DNA sequences. *Nucleic Acids Res* **10**, 141–156.

Strauch, E., Lurz, R. & Beutin, L. (2001). Characterization of a Shiga toxin-encoding temperate bacteriophage of *Shigella sonnei*. *Infect Immun* **69**, 7588–7595.

Thomas, A., Cheasty, T., Chart, H. & Rowe, B. (1994). Isolation of Vero cytotoxin-producing *Escherichia coli* serotypes O9ab : H- and O101 : H-carrying VT2 variant gene sequences from a patient with haemolytic uraemic syndrome. *Eur J Clin Microbiol Infect Dis* **13**, 1074–1076.

Toth, I., Schmidt, H., Dow, M., Malik, A., Oswald, E. & Nagy, B. (2003). Transduction of porcine enteropathogenic *Escherichia coli* with a derivative of a Shiga toxin 2-encoding bacteriophage in a porcine ligated ileal loop system. *Appl Environ Microbiol* **69**, 7242–7247.

Tozzi, A. E., Caprioli, A., Minelli, F. & 8 other authors (2003). Shiga toxin-producing *Escherichia coli* infections associated with hemolytic uremic syndrome, Italy, 1988–2000. *Emerg Infect Dis* **9**, 106–108.

Unkmeir, A. & Schmidt, H. (2000). Structural analysis of phage-borne *stx* genes and their flanking sequences in shiga toxin-producing *Escherichia coli* and *Shigella dysenteriae* type 1 strains. *Infect Immun* **68**, 4856–4864.

Wagner, P. L. & Waldor, M. K. (2002). Bacteriophage control of bacterial virulence. *Infect Immun* **70**, 3985–3993.

Wagner, P. L., Neely, M. N., Zhang, X., Acheson, D. W., Waldor, M. K. & Friedman, D. I. (2001). Role for a phage promoter in Shiga toxin 2 expression from a pathogenic *Escherichia coli* strain. *J Bacteriol* **183**, 2081–2085.

Wells, J. G., Davis, B. R., Wachsmuth, I. K., Riley, L. W., Remis, R. S., Sokolow, R. & Morris, G. K. (1983). Laboratory investigation of hemorrhagic colitis outbreaks associated with a rare *Escherichia coli* serotype. *J Clin Microbiol* **18**, 512–520.

Williams, K. P. (2003). Traffic at the tmRNA gene. *J Bacteriol* **185**, 1059–1070.

Williams, J. M., Boyd, B., Nutikka, A., Lingwood, C. A., Barnett Foster, D. E., Milford, D. V. & Taylor, C. M. (1999). A comparison of the effects of verocytotoxin-1 on primary human renal cell cultures. *Toxicol Lett* **105**, 47–57.

Willshaw, G. A., Smith, H. R., Scotland, S. M., Field, A. M. & Rowe, B. (1987). Heterogeneity of *Escherichia coli* phages encoding Vero cytotoxins: comparison of cloned sequences determining VT1 and VT2 and development of specific gene probes. *J Gen Microbiol* **133**, 1309–1317.

Zhang, W. L., Bielaszewska, M., Liesegang, A., Tschäpe, H., Schmidt, H., Bitzan, M. & Karch, H. (2000). Molecular characteristics and epidemiological significance of Shiga toxin-producing *Escherichia coli* O26 strains. *J Clin Microbiol* **38**, 2134–2140.

Zhang, W., Bielaszewska, M., Kuczius, T. & Karch, H. (2002). Identification, characterization, and distribution of a Shiga toxin 1 gene variant (stx_{1c}) in *Escherichia coli* strains isolated from humans. *J Clin Microbiol* **40**, 1441–1446.

Zoja, C., Morigi, M. & Remuzzi, G. (2001). The role of the endothelium in hemolytic uremic syndrome. *J Nephrol* **14**, 58–62.