Analysis of the EHEC hly operon and its location in the physical map of the large plasmid of enterohaemorrhagic Escherichia coli O157:H7

Herbert Schmidt, Christian Kernbach and Helge Karch

Almost all clinical enterohaemorrhagic Escherichia coli (EHEC) O157:H7 isolates harbour a large virulence plasmid designated p0157. In this study, p0157 of EHEC O157:H7 reference strain EDL 933 was characterized at the molecular level. A restriction map was constructed by using seven restriction enzymes, with appropriate gel electrophoretic and hybridization methods. The molecular size of p0157 was determined to be 93.6 kb. By sequencing the DNA region extending in the 3'-direction of the previously described EHEC hlyC and hlyA genes, two further genes were discovered and analysed; these were termed EHEC hlyB and EHEC hlyD. The newly discovered genes together with the EHEC hlyC and hlyA genes constitute a typical RTX (Repeats in Toxin) determinant (EHEC hly operon) with the gene order CABD. The map position of the operon was determined by hybridization experiments. Analysis of a DNA fragment carrying the downstream flanking region of the EHEC hly operon revealed an open reading frame which was highly homologous to orf1 of RepFIB, a basic replicon of incF plasmids. It was located close to the EHEC hly operon.

**Keywords:** enterohaemorrhagic Escherichia coli O157:H7, plasmid p0157, EHEC hly operon, hlyBD genes, RTX determinant

**INTRODUCTION**

Enterohaemorrhagic Escherichia coli (EHEC) O157:H7 is a major cause of serious outbreaks as well as sporadic cases of haemorrhagic colitis (HC) and classical haemolytic uraemic syndrome (HUS) (Tarr, 1995; Griffin & Tauxe, 1991). Since it was first described in 1983 (Riley et al., 1983; Wells et al., 1983), this organism has emerged as an important public health concern world-wide (O'Brien et al., 1992; Griffin & Tauxe, 1991). Shiga-like toxins are considered to be the major pathogenic factors of E. coli O157:H7 (Tesh & O'Brien, 1991; Tarr, 1995). The production of intimin is an important virulence characteristic of this bacterium. Intimin is a 97 kDa outer-membrane protein involved in the intimate bacterial attachment-effacement of the gut epithelium and actin aggregation in cells to which E. coli O157:H7 adheres (Yu & Kaper, 1992; Dytoc et al., 1993; Louie et al., 1993; Tzipori et al., 1995). Auxiliary virulence mechanisms like the production of fimbrial antigens (Karch et al., 1987) and exopolysaccharide (Fratamico et al., 1993), and tolerance of acidic environments (Benjamin & Datta, 1995), are only partially understood.

A large plasmid of approximately 90 kb (p0157) is present in virtually all clinical O157 isolates and several studies have been carried out to clarify the plasmid's function. Most of these studies have addressed the adherence properties of EHEC O157:H7 strains to cultured eukaryotic cells. Whereas some of the authors described a plasmid-dependent adherence to HEp-2 and Int407 cells, others could not confirm these observations (Karch et al., 1987; Toth et al., 1990; Fratamico et al., 1993; Sherman et al., 1987; Junkins & Doyle, 1989). Wadolkowski et al. (1990) reported that the plasmid-cured EDL933-cu could not compete with strain EDL933 for colonization of the streptomycin-treated mouse bowel. They suggested that the 90 kb plasmid might encode a required factor for E.
coli O157:H7 strains to colonize in the presence of other bacteria. In contrast, by using the gnotobiotic piglet model, plasmid pO157 could not be shown to promote proliferation or colonization of EHEC O157:H7 strain EDL933 in the gastrointestinal tract (Tzipori et al., 1986, 1987). Further studies on pO157 included the invasion characteristics of EHEC O157 in cultured bladder and colonic cells (Oelschlaeger et al., 1994), the influence on attaching and effacing adhesion (Toth et al., 1990), the regulation of the chromosomal aae.A gene (Gomez-Duarte & Kaper, 1995) and the expression of plasmid-encoded proteins (Fratamico et al., 1993). However, none of the studies provided definite proof of the importance of pO157 in the expression of EHEC virulence. As only single aspects of EHEC infections may be studied using the animal models currently available, efforts to elucidate the direct influence of plasmid pO157 on multisymptomatic HUS have been impeded.

Recently, we described a novel pO157-encoded protein of EHEC O157:H7 which was designated EHEC haemolysin (Schmidt et al., 1995). This protein belongs to the RTX (Repeats in Toxin) family of pore-forming cytolyssins and shares high sequence similarity with E. coli α-haemolysin (Schmidt et al., 1995). RTX toxins represent a family of important virulence factors that are widely distributed among Gram-negative bacteria and are defined by a glycine-rich signature nonamer repeated several times in the C-terminal region of the protein (Coote, 1992; Welch, 1991, 1995). RTX proteins share some common features and may function as proteases, lipases, or pore-forming cytolyssins, the latter of which kill a variety of target cells by building pores due to their integration in target cell membranes (Coote, 1992; Welch, 1995).

The enterohaemolytic phenotype of EHEC O157:H7 is caused by EHEC haemolysin (Schmidt et al., 1995). This specific mode of haemolytic activity can be recognized by the appearance of small, turbid zones of haemolysis after 18–24 h incubation at 37 °C. In contrast, α-haemolysin causes big, clear zones of haemolysis after 4–6 h incubation at 37 °C on standard blood agar plates. When strains are grown in liquid culture, EHEC haemolysin remains cell-associated (Schmidt et al., 1995), whereas α-haemolysin is exported into the culture supernatant via a specific membrane translocator system consisting of HlyB, HlyD and the outer-membrane protein TolC (Fath & Kolter, 1993; Welch, 1991; Wandersman & Delepeleire, 1990). Furthermore, EHEC haemolysin triggers a specific immune response in patients suffering from O157:H7-associated HUS (Schmidt et al., 1995), indicating that it is produced during the infectious stage.

The purpose of this study was to provide a basic molecular foundation to further investigate plasmid pO157 and the EHEC haemolysin operon with respect to the biological functions and pathogenic properties of EHEC O157:H7. Data from restriction mapping and nucleotide sequence analyses should enable coordinated genetic studies including investigation of the regulation of gene expression, biological activity of plasmid-encoded proteins and the role of pO157 during E. coli O157:H7 infections.

**METHODS**

**Bacterial strains, plasmids and culture conditions.** The EHEC O157:H7 strain EDL933 was isolated from a patient with haemorrhagic colitis; it harbours the large virulence plasmid pO157 (~90 kb) and one smaller plasmid of 3-3 kb (Karch et al., 1987; O'Brien et al., 1983). E. coli HB101 [F vacB mcrB hsdS20 (r6 m5) recA13 luxB6 ara-14 proA.A lacY1 galK2 xyl-5 mtl-1 rpsL20 sapE54 Δ] was used as host for recombinant plasmids. Plasmid pSK3 is a derivative of pO157 obtained by transposon mutagenesis with the temperature-sensitive Tn501, which is a member of the Tn3 family (Karch et al., 1987; Heffron, 1983; Schmidt et al., 1994). For large-scale multiplication of plasmid DNA this strain was grown for at least 18 h in L-broth (for pSK3 at 30 °C, for all other plasmids at 37 °C) with vigorous shaking. Plasmid pK18 was used as cloning vector (Priddmore, 1987).

**General recombinant DNA techniques.** Large-scale purifications of pSK3 and pO157 were performed with the Qiagen plasmid kit according to the manufacturer's instructions for low-copy-number plasmids; pK18 derivatives were prepared by the regular protocol (Qiagen Plasmid Handbook). For maintenance of recombinant plasmids 100 μg ampicillin ml⁻¹ (pSK3) or 50 μg kanamycin ml⁻¹ (pK18 derivatives) was added to the culture media. For restriction endonuclease cleavage of pSK3, pO157 and subclones the six-cutters AccI, BamHI, ClaI, EcoRI, HindIII, PstI and SalI, and the rare-cutting enzymes RvuI, SfiI, SphI and XbaI (all from New England Biolabs) were selected in order to create suitable amounts and sizes of restriction fragments. Restriction endonuclease digestions were carried out according to the supplier's instructions. DNA fragments were purified from agarose gels using the Prep-a-Gene kit (Bio-Rad). Ligation and transformation experiments were carried out by standard methods (Sambrook et al., 1989).

**Agarose gel electrophoresis and restriction fragment size determination.** The size of pSK3 and pO157 (>90 kb) necessitates the separation of restriction enzyme digests by constant-field (CFGE) and pulsed-field gel electrophoresis (PFGE). For CFGE, 0.6–2% (w/v) submarine agarose gels (12 cm x 18 cm) were used. In order to create sharp and well-separated bands, a maximum electrical gradient of 5 V cm⁻¹ was applied. The concentration of DNA loaded on the gel was adjusted to approximately 10–20 ng per visible band. Under these conditions fragments ranging from 0.1 to 7 kb were properly separated in the logarithmic zone of the gel. Non-separable large bands located in the compressed zone of the constant-field gels could be separated by PFGE in a horizontal, CHEF-configured pulsed-field apparatus (Bio-Rad) with 0.5-fold concentrated Tris/borate/EDTA (TBE) buffer at 12 °C. The voltage was maintained at 170 V with a constant pulse time of 3 s. The restriction enzyme digests, containing 30–50 ng DNA, were loaded directly into the slots of a 1% (w/v) high-purity agarose gel (agarose III, PFGE grade, Amresco) and electrophoresed for 16 h. Large restriction fragments ranging from 5 to 100 kb could be separated under these conditions.

By comparing restriction enzyme digestions performed on different gels, it was possible to estimate fragment sizes between 0.1 and 98 kb with an error of less than 0.5%. Measurement of bands was performed on enlarged prints of the gels. Restriction fragment sizes were determined by comparing the migrations of the fragments with a calibration graph constructed with the
migrations of size markers separated on the same gel. Size markers were the High Molecular Weight DNA Marker (GibcoBRL) the Low-range PFGE Marker (New England Biolabs) and the 1 kb DNA Ladder (GibcoBRL).

**Southern blot hybridization.** DNA was transferred from agarose gels to Zeta-probe nylon-membranes (Bio-Rad) by standard methods (Sambrook *et al.*, 1989). For hybridization assays, the Boehringer Non-radioactive DNA Labelling and Detection Kit was used, according to the manufacturer's instructions. Hybridization was performed by incubating the membrane in hybridization solution for 18 h in a 60 °C water bath. The specific washing step was carried out twice at 60 °C for 5 min in 0.03 x SSC (45 mM sodium chloride, 0.45 mM sodium citrate) and 0.1% (w/v) SDS to effect a stringency of 95%. Probes were labelled with digoxigenin-11-dUTP either randomly or by PCR as described previously (Schmidt *et al.*, 1993).

**PCR.** PCR for creation of gene probes specific for Tn801 was performed in the GeneAmp PCR System 9600 (Perkin Elmer-Applied Biosystems). Amplification was carried out in a total volume of 50 μl containing each deoxynucleosidetriphosphate at 200 μM, 30 pmol of each primer (801-start (5'-CCG CAG TGA GAG CAG AGA TAG-3') and 801-stop (5'-GAG GTA GTC ACA GCC ATT CAG-3')), 5 μl tenfold-concentrated polymerase synthesis buffer containing 1.5 mM MgCl2, and 2.0 U AmpliTaq DNA polymerase (Perkin Elmer-Applied Biosystems). The DNA was denatured at 94 °C for 45 s, annealed at 53 °C for 60 s, and then extended for 120 s at 72 °C. After 30 cycles were completed, a final extension step of 5 min at 72 °C was conducted. PCR for preparing gene probes specific for EHEC blyA was carried out as described previously (Schmidt *et al.*, 1995).

**DNA sequencing.** This was carried out with universal and reverse primers for pUC/M13 vectors and customized primers as described previously (Schmidt *et al.*, 1995). Nucleotide sequence analyses and the searches for homologous DNA sequences in the EMBL and GenBank database libraries were performed with the program package HUSAR (Heidelberg UNIX Sequence Analysis Resources, German Cancer Research Centre, Heidelberg) and the DNAsis program package (Hitachi).

**RESULTS**

**Construction of the physical map of recombinant plasmid pSK3**

Due to the presence of two plasmids in strain EDL 933 and the large amounts of plasmid DNA necessary to adjust gel-electrophoretic conditions, Tn801-tagged plasmid pSK3 (Schmidt *et al.*, 1994) was chosen for primary restriction analysis. This plasmid can be easily transferred and propagated in the non-pathogenic strain E. coli laboratory strain HB101. The three subclones were randomly labelled with digoxigenin and hybridized with several restriction enzyme digests of pSK3. The results of these hybridizations allowed us to determine the locations of the subclones in the Sfl–SpeI–XbaI map. In doing this, a large proportion of the restriction fragments created with the other enzymes could be arranged in a provisional restriction map. To determine the position of the remaining fragments in the map we performed cross-hybridizations: suitable restriction fragments were excised from agarose gels, randomly labelled with digoxigenin, and hybridized with restriction enzyme digests of pSK3 using all seven enzymes. By comparing the resulting hybridization signals we were able to detect related fragments, coordinate all restriction sites, and construct the physical map of the entire plasmid pSK3. One of the two XbaI sites was designated as the starting point of the map. The size for subcloning in pK18. Analysis of transformants revealed that two EcoRI fragments of similar sizes had been cloned. To map these EcoRI fragments of 11.4 kb (pE11, Fig. 1a) and 12.1 kb (pE12) EcoRI fragments are shown. The complete restriction map of the 12.0 kb (pB12) fragment carrying the EHEC hly determinants is shown in (b). ORFs are depicted by boxes. Arrows indicate the direction of transcription. The distances in kb indicated below the maps are in accordance with Fig. 2.

![Fig. 1. Detailed linear restriction enzyme maps of subclones derived from pSK3. In (a) and (c), the 11.4 kb (pE11) and 12.1 kb (pE12) EcoRI fragments are shown. The complete restriction map of the 12.0 kb (pB12) fragment carrying the EHEC hly determinants is shown in (b). ORFs are depicted by boxes. Arrows indicate the direction of transcription. The distances in kb indicated below the maps are in accordance with Fig. 2.](image-url)
H. Schmidt, C. Kernbach and H. Karch

Fig. 2. Circular restriction map of wild-type plasmid pO157. The restriction sites for XbaI (X), SpeI (S), SfiI (Sf), EcoRI (E), BamHI (B), Clal (C), and RsrlI (R) are indicated. These sites are positioned with respect to the XbaI site at coordinate 0, and the map is numbered by distance (in kb) from that site. The site of insertion of Tn807 is indicated at position 73.8 of the map. The positions of subclones pEll, pBl2 (pE040) and pEl2 are indicated by curved arrows. At position 7.8-15, the EHEC hly operon with genes hlyC, hlyA, hlyB and hlyD is depicted by bold arrows. EHEC orf1 spans the region from 4.7 kb to 5.6 kb.

Table 1. Sizes (kb) of BamHI, Clal, EcoRI, RsrlI, SfiI, SpeI and XbaI restriction fragments of wild-type plasmid pO157

<table>
<thead>
<tr>
<th>BamHI</th>
<th>Clal</th>
<th>EcoRI</th>
<th>RsrlI</th>
<th>SfiI</th>
<th>SpeI</th>
<th>XbaI</th>
</tr>
</thead>
<tbody>
<tr>
<td>25.9</td>
<td>31.4</td>
<td>32.4</td>
<td>42.8</td>
<td>50.3</td>
<td>58.0</td>
<td>61.8</td>
</tr>
<tr>
<td>23.5</td>
<td>21.3</td>
<td>18.3</td>
<td>20.1</td>
<td>21.9</td>
<td>21.2</td>
<td>31.8</td>
</tr>
<tr>
<td>12.0</td>
<td>10.6</td>
<td>12.1</td>
<td>15.5</td>
<td>21.4</td>
<td>7.7</td>
<td></td>
</tr>
<tr>
<td>8.6</td>
<td>9.8</td>
<td>11.4</td>
<td>8.5</td>
<td></td>
<td></td>
<td>6.6</td>
</tr>
<tr>
<td>3.7</td>
<td>7.8</td>
<td>3.3</td>
<td>3.7</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>2.8</td>
<td>5.9</td>
<td>3.2</td>
<td>2.0</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>2.7</td>
<td>3.0</td>
<td>2.8</td>
<td>0.9</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>2.0</td>
<td>2.4</td>
<td>2.8</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>1.8</td>
<td>0.74</td>
<td>2.7</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>1.8</td>
<td>0.6</td>
<td>2.2</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>1.7</td>
<td>0.8</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>1.7</td>
<td>0.3</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>0.95</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>0.8</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>0.8</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>0.7</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>0.7</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>0.4</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>0.35</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>0.25</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>0.15</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

of pSK3 (containing Tn801) was estimated to be 98.6 kb by summation of all fragment sizes from each digest.

Location of Tn801 and reconstruction of the physical map of pO157

To reconstruct the map of the wild-type plasmid pO157, it was necessary to determine the position of Tn801 in pSK3. Restriction of Tn801 with PstI should yield a 2.8 kb and a 0.7 kb PstI fragment (Grinsted et al., 1977; Heffron, 1983). We cloned the corresponding 2.8 kb PstI fragment derived from pSK3 in the vector pK18 to achieve a transposon-specific probe. The resulting clone was resistant to ampicillin, suggesting that it harboured the P-lactamase gene of Tn807. Following partial sequence analysis of the 2.8 kb insert, a search of the EMBL database library revealed >92% sequence similarity to transposase-encoding genes typically found in transposons of the Tn3 family (Heffron et al., 1979). The DNA sequences were used to design PCR primers for PCR-mediated probe-labelling. The resulting 2.5 kb Tn807 probe was hybridized with selected restriction enzyme digests of pSK3 and pO157. By analysing the hybridization results and comparing the restriction patterns of pSK3, pO157 and Tn801, we were able to determine the location of Tn801 as well as deduce the restriction map of the parental wild-type plasmid pO157 (Fig. 2). The fragment sizes resulting from digestion of pO157 with the seven enzymes mentioned above are shown in Table 1. The size of pO157 from strain EDL933 was estimated to be 93.6 kb.

DNA sequence analysis of the EHEC hlyB and hlyD genes and location of the EHEC hly operon in the pO157 map

The two genes EHEC hlyC and EHEC hlyA were recently described to be located on a 12 kb BamHI (pEO40) fragment of pO157 (Schmidt et al., 1995). Following analysis of that portion of pEO40 extending in the 3'-direction of the hlyA gene, we discovered another two open reading frames (ORFs), of 2121 bp and 1440 bp, capable of encoding proteins of 706 and 479 amino acids, respectively. Searching the EMBL database revealed high similarities of 69% and 62% to the genes hlyB and hlyD of the β-haemolysin operon; the genes were consequently termed EHEC hlyB and EHEC hlyD. Since two tandemly arranged ATG codons preceded the EHEC hlyB gene, the translational start point could not be clearly deduced from the sequence. A spacing region of 49 bp with the ability to build a stem-loop structure is located between the termination codon of the EHEC hlyA gene and the putative start codon of the EHEC hlyB gene. The EHEC hlyB and hlyD genes are separated by only three nucleotides. The putative proteins share high identity with other RTX proteins. Highest similarities were found with HlyB (72.9%) and HlyD (56.8%), encoded by the E. coli α-haemolysin operon, as well as with LktB (71.3%) and...
LktD (60%), encoded by the *Actinobacillus actinomyecetemcomitans* leukotoxin operon. HlyB and LktB are members of a family of ATP-binding transporter proteins found in prokaryotic and eukaryotic cell membranes (Fath & Kolter, 1993; Koronakis et al., 1993). Members of this family share conserved sequence motifs in their so-called ABC domain (Koronakis et al., 1995). The importance of single amino acid residues and motifs for the transport function has been studied in detail for HlyB of *E. coli* by Koronakis et al. (1995). To analyse the putative EHEC HlyB protein we compared its amino acid sequence with *E. coli* HlyB (Hess et al., 1986) and *A. actinomyecetemcomitans* LktB (Guthmiller et al., 1990a) (Fig. 3a).

Highly conserved sites of the HlyB ABC domain comprise the Walker 'A' (GRSGSGKST) and 'B' (ILIL) motifs, and the linker peptide (LSGGQ). These sites are identical in the putative EHEC HlyB protein, suggesting ATP-hydrolysing activity. Only two differences from the HlyB ABC domain core sequence were detected in regions of lower sequence conservation: K625R and G654N (see Fig. 3a). The first substitution, K625R, preceded the Walker B-site, whereas the second one, G654N, was not close to any of the highly conserved sites. The same amino acid substitutions were found in the LktB protein (see Fig. 3a).

Fig. 3. (a) Sequence comparison of the putative EHEC HlyB protein (top line) with the HlyB protein from a-haemolytic plasmid pHly152 (middle line) and LktB (bottom line) from the leukotoxin operon of *A. actinomyecetemcomitans*. The Walker A' motif (1), the linker peptide (2) and the Walker B' motif (3) are boxed and numbered below. Empty circles above the sequences indicate sites of low sequence conservation among HlyB-related proteins, whereas the filled circles represent highly conserved sites. Amino acid substitutions K625R and G654N are marked by arrows. Amino acids identical to the EHEC HlyB sequence are indicated by dashes. (b) Comparison of the amino acid sequence of the putative EHEC HlyD protein (top line) with HlyD (middle line) and LktD (bottom line). The C-terminal 10 amino acid residues of the putative EHEC HlyD protein share high similarity with the same region of other HlyD-related proteins (Fig. 3b).

Together with the previously described EHEC *hlyC* and *hlyA* genes, the EHEC *hlyB* and *hlyD* genes constitute an RTX operon with the gene order CABD. The ORFs for the EHEC *hlyC*, *hlyA*, *hlyB*, and *hlyD* genes range, respectively, from positions 74 to 580; 582 to 3578; 3628 to 5748; and 5752 to 7191. By comparing restriction fragment patterns and hybridizing specific gene probes with restriction enzyme digests of subclones and the entire p0157, the position of the EHEC haemolysin operon was determined to be from 7.8 kb to 15 kb in the physical map of p0157 (Fig. 2).
Replication-associated gene of pO157

Subclone pE11 (Figs 1a and 2), carrying the 11.4 kb EooR1 DNA fragment located directly downstream from the EHEC hly genes, was further subcloned. Subclones were subjected to DNA sequencing using M13/pUC18 universal and reverse primers. An overall similarity of approximately 96% was found with RepFIB, a basic replicon found in large plasmids of the IncF group (Gannon et al., 1990). By using subclone pBl-5 (Fig. 1a) and customized primers, we detected an ORF of 954 bp highly homologous to orf7 of the RepFIB region (Saul et al., 1989). This ORF has been shown to be essential for replication and highly homologous when investigated from several IncF plasmids. We therefore concluded that the origin of replication of pO157 is partially located on this clone. The entire origin of replication may be encoded by this clone, but this was not determined by this investigation. The EHEC orf1 is located at map position 4.7 kb to 5.6 kb, close to the EHEC hly operon (Fig. 2).

DISCUSSION

Although pO157 of E. coli O157:H7 strains has been the subject of several investigations, its role in pathogenicity and virulence has not yet been clearly defined. In contrast to virulence plasmids of other enteropathogenic bacteria like Shigella dysenteriae or Yersinia enterocolitica, no pathogenicity factors have been identified for pO157 since its discovery in 1982. However, we recently discovered a protein encoded by pO157 which possibly has a function in pathogenicity. This protein, EHEC haemolysin, is responsible for the enterohaemolytic phenotype and is a member of the RTX family of exoproteins (Schmidt et al., 1995). Most of the RTX proteins are transported through the cytoplasmic and outer membranes by a unique transport system which is specified by RtxB, RtxD and a genetically unlinked outer-membrane protein (Welch, 1991, 1995). This led us to the hypothesis that the structural basis for the enterohaemolytic phenotype is the result of a defective or inefficient membrane transport system necessary to export EHEC haemolysin.

The phenomenon of retaining an RTX protein on the cell surface was described for the Gram-negative periodontal pathogen Actinobacillus actinomycetemcomitans, which produces leukotoxin (LktA). Although the lktB and lktD genes, with high sequence similarity to the E. coli hlyB and D genes, are present, leukotoxin is not secreted from the cell and remains associated with the outer membrane. Mutational inactivation of the lktBD genes resulted in a significant decrease of leukotoxin, whereas the transcription of lktA was unaffected. The lktBD genes may be responsible for post-transcriptional effects on the leukotoxin relative to protein stability and the leukotoxin location on the cell surface (Guthmiller et al., 1995; Lally et al., 1991).

In this study we demonstrated the presence of EHEC hlyB and hlyD genes with high sequence similarity to other RTX transport proteins. From nucleotide sequence analysis data we concluded that no missense mutation disturbs the function of either EHEC HlyB or HlyD. However, the putative EHEC HlyB protein was shown to carry two amino acid substitutions in the ABC domain when compared with the HlyB protein. One of the two, K625R, is not thought to affect the translocation process. Koronakis et al. (1995) showed that HlyB mutations located at the same position did not influence the transport function. The second amino acid substitution was located in a region with low sequence conservation; no data are available concerning the region's importance. However, the substitution of the small, nonpolar amino acid glycine by the basic asparagine could strongly influence the protein structure and activity. Interestingly, the LktB protein shows the same amino acid substitutions as EHEC HlyB. Whether these amino acid residues correspond to the specific mode of export of EHEC haemolysin and leukotoxin remains to be elucidated.

A topological model for the translocator HlyD encoded by the \( \alpha \)-haemolysin operon has been proposed by Schülein et al. (1992). Analysis by PhoA and LacZ fusions and computer-based predictions revealed that the N-terminus (residues 1–60) of HlyD was located in the cytoplasm, followed by a transmembrane domain (residues 60–80). Residues 47–64 are thought to span the periplasmic space. The C-termini of HlyD and HlyD-related proteins are highly conserved and thought to play an essential role in the translocation process (Schülein et al., 1992). EHEC HlyD shows the same hydrophobicity pattern and C-terminal sequence conservation as HlyD and LktD, suggesting a function similar to these proteins.

The association of EHEC haemolysin with the bacterial cell may influence the biological activity of this toxin. Pore-forming cytolysins, e.g. E. coli \( \alpha \)-haemolysin and Staphylococcus aureus \( \alpha \)-toxin have been shown to damage target cells in a concentration-dependent manner. Bhakdi et al. (1994) showed that high concentrations of S. aureus \( \alpha \)-toxin bind in a non-specific manner to target cells whereas toxin in low concentrations binds specifically to cells expressing a high-affinity binding site. \( \alpha \)-Haemolysin, which is exported in high concentrations in the environment, is presumed to bind indiscriminately to lipid bilayers. Therefore, the association of EHEC haemolysin with the bacterial cell surface or its extremely low export levels may contribute to the biological function of this toxin in the intimate environment of infecting EHEC O157:H7.

The irregular distribution of restriction sites in pO157 is striking (Fig. 2). By in vitro transcription–translation experiments it was shown that approximately 50 proteins are encoded on this plasmid (Fratamico et al., 1993). A proportion of these proteins seem to be F-plasmid-specific, and therefore necessary for maintenance, replication control, and transfer functions. Hales et al. (1992) reported an approximately 23 kb BamHI fragment to be present in all investigated plasmids of EHEC strains, which is responsible for the incompatibility and replication functions. We found the fragment corresponded to the 23.5 kb BamHI fragment spanning the region from 72.2 kb to 2.1 kb in the physical map of pO157 (Fig. 2).
The occurrence of the replication-associated gene EHEC orf1 located on that fragment supported this finding. Interestingly, the EHEC hly operon is located in the region with numerous restriction sites. The information available from the physical map and the nucleotide sequence data determined here will enable further systematic investigation of pO157 and the function of EHEC haemolysin. Future studies should identify other possible virulence factors on the plasmid, investigate the regulation of plasmid encoded proteins, and clarify the role of EHEC haemolysin and the entire pO157 in the pathogenesis of HC and HUS.

ACKNOWLEDGEMENTS

We thank Jürgen Heesemann, Iaylo Gentschev and Werner Göbel, Würzburg, Germany, for helpful discussions, and Beatriz Henkel for excellent technical assistance. This work was supported by grant Ka 717-2/2 from the Deutsche Forschungsgemeinschaft.

REFERENCES


H. SCHMIDT, C. KERNBACH and H. KARCH


Received 24 August 1995; revised 1 November 1995; accepted 29 November 1995.