ELSEVIER

Institut Pasteur

Research in Microbiology xx (2014) 1–4

www.elsevier.com/locate/resmic

Brief note

Insertions in the OCL1 locus of *Acinetobacter baumannii* lead to shortened lipooligosaccharides

Q2 Johanna J. Kenyon [a], Kathryn E. Holt [b], Derek Pickard [c], Gordon Dougan [c], Ruth M. Hall [a,*]

[a] *School of Molecular Bioscience, The University of Sydney, New South Wales, Australia*
[b] *Department of Biochemistry and Molecular Biology, and Bio21 Molecular Science and Biotechnology Institute, The University of Melbourne, Victoria, Australia*
[c] *Wellcome Sanger Trust Institute, Hinxton, Cambridge, United Kingdom*

Received 12 May 2014; accepted 14 May 2014

Abstract

Genomes of 82 *Acinetobacter baumannii* global clones 1 (GC1) and 2 (GC2) isolates were sequenced and different forms of the locus predicted to direct synthesis of the outer core (OC) of the lipooligosaccharide were identified. OCL1 was in all GC2 genomes, whereas GC1 isolates carried OCL1, OCL3 or a new locus, OCL5. Three mutants in which an insertion sequence (ISAba1 or ISAba23) interrupted OCL1 were identified. Isolates with OCL1 intact produced only lipooligosaccharide, while the mutants produced lipooligosaccharide of reduced molecular weight. Thus, the assignment of the OC locus as that responsible for the synthesis of the OC is correct.

Keywords: Lipooligosaccharide; Lipid A; Outer core oligosaccharide; OC locus; ISAba1; ISAba23

Lipooligosaccharide (LOS) is a phospholipid–carbohydrate surface structure that is associated with many pathogenic properties of Gram-negative species, including the important nosocomial pathogen, *Acinetobacter baumannii* [1,2]. It is composed of lipid A, which anchors the LOS in the outer leaflet of the outer membrane, and a core oligosaccharide that extends out from the cell surface. The core consists of several carbohydrates linked together, but is subdivided into inner core and outer core (OC) regions. Genes required for the synthesis of lipid A and the inner core are usually distributed throughout Gram-negative genomes, whereas genes for the OC are usually clustered [3]. In all Gram-negatives studied to date, the OC locus contains multiple genes encoding glycosyltransferase enzymes that catalyse the linkages between the sugars in the OC structure, and may contain genes for sugar synthesis or modification.

We recently reported the identification of two regions in the genomes of *A. baumannii* isolates that include a cluster of genes encoding enzymes for glycosyl transfer and the synthesis or modification of complex sugars [2]. These loci were the only regions found to contain multiple genes associated with surface carbohydrate biosynthesis and to show several different configurations in different genomes. On the basis of all available evidence, the larger locus was unambiguously identified as the gene cluster responsible for the synthesis of the polysaccharide capsule, and was designated the K locus [2,4,5]. The smaller gene cluster located between *ilvE* and *aspS* contained multiple glycosyltransferase genes, and must therefore direct the synthesis of the OC component of the LOS. It was designated the OC locus (OCL). Three different OCL forms (OCL1–OCL3) were found in the first ten completed genome sequences [2]. Each contained 9 genes and was between 11 and 12 kb in length.

There was good correlation between the resolved LOS structure of isolate ATCC 19606 [6] and the content of the OCL1 gene cluster that it carries. A second strain, SMAL, produced the same structure [7] and also carries OCL1

* Corresponding author.
E-mail addresses: johanna.kenyon@sydney.edu.au (J.J. Kenyon), ruth.hall@sydney.edu.au (R.M. Hall).

http://dx.doi.org/10.1016/j.resmic.2014.05.034
0923-2508/

(Kenyon and Hall, unpublished). OCL1 contains 5 genes predicted to encode glycosyltransferases that would form linkages in the OC1 structure. OC1 includes glucosamine (D-Glc*p*N) and galactosamine (D-Gal*p*N) sugars, and the *pda1* gene (Fig. 1) encodes an enzyme predicted to deacetylate UDP-*N*-acetyl-glucosamine (UDP-D-Glc*p*NAc) and/or UDP-*N*-acetyl-galactosamine (UDP-D-Gal*p*NAc), producing the UDP-linked form of these products. OCL1 also contains 3 genes without known relatives. One facing in the opposite direction to the remaining genes in OCL1 may encode a glycosyl hydrolase, and the remaining 2 have tentatively been assigned at glycosyltransferases [2].

Here, we have determined the genome sequences of *A. baumannii* isolates belonging to the clinically important global clones, global clone 1 (GC1) and global clone 2 (GC2) from our Australian collection. GC1 and GC2 correspond to CC1 and CC2 in Ref. [8]. Naturally occurring OCL mutants identified among them were used to examine the size of the LOS they produce. This provided experimental evidence confirming the assignment of the OC locus as the region that directs the synthesis of the OC of the LOS.

1. Distribution of OCL forms in GC1 and GC2

Whole genome sequences were determined for 82 multiply antibiotic resistant GC1 and GC2 *A. baumannii* isolates from Australian hospitals [9,10] using Illumina HiSeq. Paired-end reads of 100 bp were assembled using *Velvet*, as described previously [11]. This yielded a median of 116 contigs per genome (median N50, 147 kbp) with an average read depth of at least 70×.

The sequences were examined for the presence of the three reported forms of the OC locus [2]. All 61 GC2 isolates carried OCL1. However, among the 21 GC1 isolates, only 17 carried the OCL1 gene cluster. Three of the remaining 4 GC1 isolates, A85, RBH3 and 6772166 carried OCL3 (GenBank accession KC118540), and one isolate, D13, contained a novel locus designated OCL5 (GenBank accession HM590877) that will be described in detail elsewhere.

2. Identification of IS insertions in OCL1

In some genomes (2 GC1 and 11 GC2), OCL1 was not in a single contig, indicating that the locus may be interrupted by a repeated sequence. In each case, this was traced to the presence of an insertion sequence (IS). The assemblies were confirmed by PCR (as described in Ref. [9]) using primers specifically designed to amplify the region that includes the IS, and the products were sequenced for a representative of each group. Three different IS insertions were identified, and the location of these IS elements are shown in Fig. 1. The isolates that carry these interrupted forms of OCL1 are listed in Table 1 together with GenBank accession numbers for one representative.

Two GC1 isolates, D78 and D81, recovered at Royal North Shore Hospital, Sydney, Australia in the same year included a novel IS that interrupted *gtrOC6* (Fig. 1). This mutant form was designated OCL1b. The IS sequence was deposited in ISFinder (https://www-is.biotoul.fr//is.html) and assigned the name ISAba23. ISAba23 belongs to the IS*5* family, is 1249 bp in length and is bounded by 16 bp inverted repeats. The insertion of ISAba23 has created a duplication of the 5 bp target sequence.

A GC2 isolate, D1, also recovered at Royal North Shore Hospital in 2006, included OCL1c with ISAba1 in the *gtrOC2* glycosyltransferase gene. ISAba1 has previously been shown to increase expression of genes adjacent to its left end (when *insA*/*insB* genes are shown transcribed to the right). This is due to the presence of a strong outward-facing promoter (see Fig S1 in Ref. [12] for references and promoter location). The ISAba1 is oriented such that it directs transcription in the same direction as the majority of genes in the OC locus (Fig. 1). Consequently, it may alleviate polar effects on the transcription of the genes downstream of the IS in *gtrOC2*.

Ten GC2 isolates from Prince of Wales Hospital, Sydney, Australia that were all isolated in 2002 (A74–A82, A84), carried OCL1 with ISAba1 interrupting the *ghy* gene. This gene cluster was designated OCL1d (Fig. 1). ISAba1 is oriented such that the promoter it provides directly opposes the transcription of the preceding 5 genes in OCL1. It is therefore possible that the expression of these genes is prevented or substantially reduced.

3. Wild type strains produce only LOS and capsule

Of the 82 Australian isolates examined, five GC2 strains (A91, A93, A94, A96, A97) were found to be closely related to the GC2 isolates (A74-A82, A84) that carry OCL1d, which

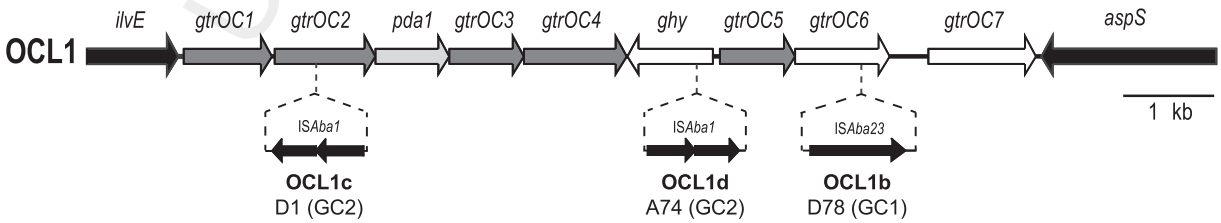

Fig. 1. Arrangement of the OCL1 gene cluster and position of IS insertions in Australian *A. baumannii* isolates. OCL1 has been described previously [2]. Arrows represent genes showing the direction of transcription, and gene names are above. Flanking genes are black, genes predicted to encode products required for nucleotide-linked sugar synthesis are light grey, and dark grey genes predict glycosyltransferases. White denotes genes that encode proteins of unknown function with possible functions indicated in the gene names. The *gtrOC6* and *gtrOC7* names were used because a total of 7 glycosyltransferases are needed to construct OC1. IS elements, insertion positions and OC locus names of the resulting OCL1 variants are indicated below.

Please cite this article in press as: Kenyon JJ, et al., Insertions in the OCL1 locus of *Acinetobacter baumannii* lead to shortened lipooligosaccharides, Research in Microbiology (2014), http://dx.doi.org/10.1016/j.resmic.2014.05.034