Structure and evolution of a family of genes encoding antiseptic and disinfectant resistance in Staphylococcus aureus

(Recombinant DNA; ethidium bromide; plasmid evolution; membrane protein; transport)

Tim G. Littlejohn, Dario DiBerardino, Linda J. Messerotti, Samantha J. Spiers and Ronald A. Skurray

Department of Microbiology, Monash University, Clayton, Victoria 3168 (Australia)

Received by P.A. Manning: 29 August 1990
Revised: 23 November 1990
Accepted: 20 November 1990

SUMMARY

Resistance to antiseptics and disinfectants in Staphylococcus aureus, encoded by the qacC/qacD gene family, is associated with genetically dissimilar small, nontransmissible (pSK89) and large conjugative (pSK41) plasmids. The qacC and qacD genes were analysed in detail through deletion mapping and nucleotide sequence analysis, and shown to encode the same polypeptide, predicted to be 107 aa in size. Direct repeat elements flank the qacD gene, elements which also flank the qacC gene in truncated forms. These elements contain palA sequences, regions of DNA required for replication of some plasmids in S. aureus. The qacC gene is predicted to have evolved from the qacD gene, and in the process to have become reliant on new promoter sequences for its expression. The entire sequence of the 2.4-kb plasmid pSK89 (which contains qacC) was determined, and is compared with other plasmids from Gram+ bacteria.

INTRODUCTION

Resistance to antiseptics and disinfectants is widespread among multiresistant S. aureus and has been found to be specified by two families of determinants: the qacA/qacB gene family and the qacC/qacD gene family (Lyon and Skurray, 1987; T.G.L., I. Paulsen, M. Gillespie, J. Tennent, M. Midgley, M. Jones, A. Purewal and R.A.S., in preparation). These determinants have been identified in plasmid isolates from Australia, North America and Europe (Emslie et al., 1985; Townsend et al., 1985; Gillespie and Skurray, 1986; Lyon and Skurray, 1987). All of these determinants encode resistance to EtdBr and Qa (such as cetyltrimethylammonium bromide and benzalkonium chloride); the qacA/qacB family differs from the qacC/qacD family in that they confer a higher level of resistance to EtdBr. Additionally, qacA confers resistance to diamidines, such as propanidide isethionate, and to the biguanidine, chlorhexidine.

In addition to bestowing different levels of resistance to a wide variety of compounds, the various qac genes are associated with different S. aureus plasmid families. The qacA gene is predominantly detected on pSK1 family multiresistance plasmids, which range in size from 19–30 kb, and may also contain transposons that encode resistance to Gm, Tm and Km (Tn4001; Lyon et al., 1984; Rouch et al., 1987; Byrne et al., 1989), penicillin (Tn4002; Gillespie et al., 1988) and to Tp (Tn4003; Rouch et al., 1989). Furthermore, the qacA gene is likely to be present on the chromosome of clinical S. aureus isolates, possibly as an
integrated pSK1 family plasmid or part thereof (Gillespie et al., 1989). The qacB gene has been detected on large heavy-metal resistance plasmids such as pSK23 (Lyon and Skurray, 1987). The qacA/qacB family of genes encodes proton-dependent export proteins (Tennent et al., 1989; T.G.L., I. Paulsen, M. Gillespie, J. Tennent, M. Midgley, M. Jones, A. Purewal and R.A.S., in preparation) which show significant homology to other energy dependent transporters such as the tetracycline exporters from Gram + and Gram- bacteria (Rouch et al., 1990), alluding to a common mechanism of action by these proteins.

The qacC and qacD determinants have identical phenotypes and show restriction site homology. qacC is located on the small (2.4-kb) plasmid pSK89 which was isolated from an Australian hospital (Lyon and Skurray, 1987). Resistance to EtdBr is also encoded by the type-II nucleic acid binding resistance determinants, found on plasmids of size 2.24 to 2.88 kb such as pWG32 (Emslie et al., 1983), which have phenotypes equivalent to qacC (Lyon and Skurray, 1987). The qacD gene, on the other hand, resides on members of a family of large (50-kb), conjugative, multiresistance plasmids such as pJE1 (Evans and Dyke, 1988; Byrne et al., 1990b) and pSK41 (Lyon et al., 1987; Lyon and Skurray, 1987; Byrne et al., 1990a; T.G.L., I. Paulsen, M. Gillespie, J. Tennent, M. Midgley, M. Jones, A. Purewal and R.A.S., in preparation) from West German and North American isolates, respectively. pSK41 additionally encodes resistance to the aminoglycosides Gm, Tm, Km, Nm and Pm. These large plasmids were shown to share DNA homology with pSK89, suggesting that a similar or identical determinant was responsible for the QacA/QacD resistance phenotype in both groups of plasmids (Lyon and Skurray, 1987). Furthermore, like the qacA and qacB genes, the qacC and qacD genes also encode proton dependent efflux proteins (Jones and Midgley, 1985; T.G.L., I. Paulsen, M. Gillespie, J. Tennent, M. Midgley, M. Jones, A. Purewal and R.A.S., in preparation). qacC and qacD thus appeared to be very similar genes, yet they reside on very different replicons.

In this paper, the cloning, transcriptional mapping and definition of the qacC and qacD genes through deletion mapping and DNA sequencing are described. In addition, the complete sequence of the qacC containing plasmid pSK89 is presented.

RESULTS AND DISCUSSION

(a) Cloning and deletion analysis of the qacC and qacD genes

The qacC and qacD genes were, respectively, localised to the larger HpaII-PstI fragment from pSK89 and to a 2.0-kb BglII fragment from pSK41 by Southern hybridisation (Fig. 1; Lyon and Skurray, 1987; data not shown). The locations of the qacC and qacD genes were more precisely determined through the construction of a series of deletion mutant derivatives of the plasmids pSK503 (qacC) and pSK534 (qacD) (Fig. 1). An equivalent fragment from the EtdBrR plasmid, pJE1, to the HpaI-BglII fragment in pSK534 has previously been shown to contain qacD (Jones and Midgley, 1985).

(b) Nucleotide sequence of the qacC and qacD genes

In addition to the deletion derivatives described above, various M13 subclones of pSK503 (qacC) and of pSK525 (qacD) were used to sequence the entirety of pSK89 (Fig. 2) and the BglII fragment from pSK41 containing qacD (Fig. 3). Analysis of these two sequences identified a region of nt sequence identity extending from nt 1457 in pSK89 (corresponding to nt 1127 in the BglII fragment from pSK41) to nt 2005 in pSK89 (corresponding to nt 1674 in the BglII fragment from pSK41), after which the homology between the nt sequences disappears. In this 548-bp region there are four changes in the pSK89 sequence relative to pSK41 (indicated by gaps in the heavy lines beneath the sequences in Figs. 2 and 3) and a single ORF of 107 aa in length that corresponds to the qacC gene on pSK89 and the qacD gene in the BglII fragment from pSK41. These ORFs correspond to the positions of the qacC and qacD genes as defined by deletion analysis (see section a above); supporting evidence for this is the presence of appropriately positioned RBSs (Figs. 2 and 3). Finally, the %G + C content of qacC/qacD is 29%, similar to the average value for Staphylococcus (Normore, 1976), suggesting that these genes arose in a Staphylococcus or closely related genus.

Recently, the nt sequence of an EtdBr resistance determinant from the conjugative plasmid pTZ20, which was isolated from an S. aureus strain in Japan, was reported (Susatsui et al., 1989). This sequence is identical to the qacD gene (corresponding to positions 995–1590 in the BglII fragment from pSK41), with the exception of an A to C (in pTZ20) change at position 1172 in pSK41. This result suggests that the qacD gene is present on large conjugative plasmids in Japan as well as in the United States and Europe.

Part of the IS257 element located S’ to the qacD gene on pSK41 (Byrne et al., 1990a) was sequenced in the process of obtaining the qacD nt sequence. The region of IS257 sequence shown in Fig. 3 is identical to that of an equivalent IS257 element on the related plasmid pSH6 (Byrne et al., 1990a) with the exception of a single bp change (T in pSK41 to C in pSH6) at position 585. This substitution does not change the aa at this position in the putative IS257 transposase (Rouch et al., 1989; Byrne et al., 1990a).
(c) Analysis of the qacC and qacD genes

The nt sequence analysis revealed that the qacC and qacD genes share an identical ORF, although they appear to be transcribed from different promoters (see f below). To gain some insight into the likely mode of action of the QacC and QacD transport proteins, the method of hydrophobicity analysis of Engleman et al. (1986) was applied to the putative polypeptide sequence.

The result of this analysis indicates that the QacC and QacD polypeptides have four large hydrophobic segments, all with the potential to traverse the cell membrane. This hydrophobicity profile, suggesting that the polypeptides are membrane-bound, is consistent with the role of QacC and QacD in energy dependent efflux of EtdBr from the cell (Jones and Midgley, 1985; T.G.L., I. Paulsen, M. Gillespie, J. Tennent, M. Midgley, M. Jones, A. Purewal and R.A.S., in preparation). Similarly, the qacA gene product (QacA) also relies on a proton gradient for mediating energised efflux of EtdBr from the cell (T.G.L., I. Paulsen, M. Gillespie, J. Tennent, M. Midgley, M. Jones, A. Purewal and R.A.S., in preparation). No significant homology could be identified between these two groups of transporters when the QacC/QacD inferred polypeptide sequence was compared to that of QacA (Rouch et al., 1990). The qacA gene...
encodes a protein predicted to be approx. fivefold larger than the QacC/QacD polypeptide, and shares significant homology to some tetracycline and sugar transport proteins (Rouch et al., 1990). Thus QacC/QacD probably use a mechanism that is quite distinct from QacA to energise the transport of their substrates.

Sequence similarity with well characterised proteins can be a useful aid in generating predictions about the function and mechanism of action of proteins (von Heijne, 1987). To determine if the predicted products of the qacC/qacD genes showed similarity with other polypeptides, their sequence was searched against the most recent GenBank, EMBL and Swissprot databases, courtesy of GenBank Pty. Ltd. (Davison and Chappell, 1990). Two sequences showing extensive identity (and similar lengths) to QacC/QacD were identified. The putative El gene from the E. coli TpR SuR plasmid pLMO20 (Sundström et al., 1988), and the ORF4 product from the Proteus vulgaris furamate reductase operon (Cole, 1987) show 41% and 33% identity with QacC/QacD, respectively. However, these two proteins have an unknown function.

(d) Direct repeats flank the qacD gene

Inspection of the nt sequences surrounding the qacD gene revealed the presence of near perfect DRs of 180 bp length flanking the gene. These are designated ‘DR1D’ on the 5’ side and ‘DR2D’ on the 3’ side of qacD (Figs. 1 and 3). Analysis of the DNA bordering the qacC gene revealed partial copies of the 180 bp DRs flanking qacD. The truncated repeats flanking qacC, which are 25 bp 5’ and 164 bp 3’ to qacC, appear in the same position relative to qacD. This is shown in Fig. 1 as the dashed lines between pSK89 and the BglIII fragment of pSK41. Taken together, the two portions of the DR sequence flanking qacC do not constitute a complete repeat. Possible roles of the DRs, and the sequence elements contained within them, are discussed in section e below.

(e) Other features of pSK89

Inspection of the pSK89 nt sequence revealed two additional ORFs which have significant identity with polypeptides encoded by plasmids from other Gram + bacteria. The first of these, between nt 1071–228 (labelled rep; Fig. 2), showed identity with the replication proteins encoded by the plasmids pC194 (Gros et al., 1987) and pUB110 (Maciag et al., 1988) from Staphylococcus; with ORFs located on the Lactobacillus plasmids pLP1 (Bouia et al., 1989) and pC30IL (Skaugen, 1989); and with ORFs from the Bacillus plasmids pTBD14 (Murai et al., 1987) and pRBH1 (Muller et al., 1986). The ORFs from Lactobacillus and Bacillus plasmids are presumed to encode replication proteins. The other ORF identified on pSK89, between positions 1003–1435 (labelled cop; Fig. 2) shows identity with the cop (control of replication) region of pC194 (Horninouchi and Weisblum, 1982; Novick, 1989). It is not known whether the cop gene is translated; rather it is thought to act to control plasmid replication through an antisense RNA (Novick, 1989).

A DNA database search of the pSK89 nt sequence revealed significant homology between the nt sequence next to qacC (between nt 1842–2002) and the pala regions of other small plasmids, including pC194 (Horninouchi and Weisblum, 1982), pC223 (Ehret and Matzura, 1988), pS194 (Projan et al., 1988), pC2GC (Brenner and Shaw, 1985), pIP855 (Brisson-Noel and Courvalin, 1986), pIP856 (Brisson-Noel et al., 1988) and pNS1CG (Noguchi et al., 1988).
The genetic composition of pSK89 appears to be very similar to that of pC194, containing genes of similar sequence and arrangement. This observation is consistent with the argument of Novick (1989) that many small S. aureus plasmids exchange resistance determinants between each other and even between plasmids from other Gram + organisms such as Bacillus. The palA sequence is located just 3' to a number of resistance genes in these plasmids, such as qacC, linA (lincomamide resistance, located on the plasmid pIP855; Brisson-Noel and Courvalin, 1986) and cat (chloramphenicol resistance) in pC194 (Horinouchi and Weisblum, 1982). The palindromic structure (characteristic of palA sequences; Novick, 1989), and high %G + C (41%), as opposed to 29% for the whole of pSK89) content, a characteristic of Rho-independent transcription terminators (Rosenberg and Court, 1979), argues that palA sequences could additionally play a role in transcription termination.

The sequence of plasmid pWG32 (Emslie et al., 1985), a plasmid equivalent to pSK89, has been determined and shown to be identical to pSK89 in all but two places (L. Chien-Hsun and S.J. Projan, pers. commun.). Plasmid pWG32 has an additional T at position 1077 relative to pWG32. This additional bp increases the spacing between the proposed RBS and start codon of rep in pSK89. pWG32 has a C at position 1069, whereas pSK89 has a T, indicating that the proposed start codon for the rep gene is AUG (Met) in pWG32 but an AUA (Ile) in pSK89. AUA has been described recently as a start codon for a penP-lacZ gene fusion in the Gram + organism Bacillus (Peijnenburg et al., 1990). What effect these changes have on the expression of rep, and consequently on the copy number of pSK89 vs pWG32, is not clear.

(f) Transcriptional analysis of the qacC and qacD genes

The qacD gene was transcriptionally mapped using the primer extension method. The tsp was located to the C at position 1162, adjacent to -10 and -35 elements (Fig. 3). The BAL 31 deletion mutant in pSK543 removed the -35 sequence (Fig. 3) and resulted in the loss of the qacD phenotype, confirming the importance of this region in qacD expression. Primer extension failed to locate the start point

![Fig. 4. Alignment of sequences flanking qacD in pSK41 with sequences flanking qacC, the palA sequence from pC194 and the palA-like sequence found 3' to the linA gene in pIP855, using the Clustal multiples sequence alignment program (Higgins and Sharp, 1989). Dots represent identity, dashes represent gaps introduced into the sequence to optimise alignment. The unbroken line for DR1-qacC represents sequence lacking identity preceding the sequence shown and has been omitted for clarity. Shaded sequences contain the DRs flanking the qacD gene and the truncated portions of the DRs flanking the qacC gene. DR1-qacD, DR 3' to qacD; DR2-qacD, DR 3' to qacD; DR2-qacC, sequence 3' to qacC; DR1-qacC sequence on the 5' of qacC; palA-pC194, palA sequence from pC194 (Horinouchi and Weisblum, 1982); palA-pIP855, sequence found 3' to the linA gene in pIP855 assumed to have palA function (Brisson-Noel and Courvalin, 1986).]
of the qacC gene. This may have been due to the lack of a suitable −35 region 5′ to qacC; identity between qacC and qacD breaks down in the middle of this putative −35 region (Fig. 2; see also Fig. 4), suggesting that another promoter is used by qacC.

(g) Evolution of the qacC and qacD genes

Extensive regions of identity between pSK41 and pSK89, extending beyond the qac determinants, and including sequence features essential for replication of small plasmids, suggest that the qacC and qacD determinants share a common ancestry. Copies of the DRs that flank qacD (DR2_D, DR1_D) are found only in part flanking qacC on pSK89. Furthermore, the two truncated portions of the DR in pSK89 do not constitute a complete copy of the DR sequence, suggesting that the qacC determinant, and flanking sequences, have been derived from pSK41. This may have occurred via integration of sequences from pSK41 into a small cryptic or other S. aureus plasmid. The presence of palA sequences on a wide range of Gram + plasmids suggests that this integration event may have occurred by homologous recombination on this hypothetical plasmid and a DR flanking qacD on pSK41. The role of palA sequences in the replication of large conjugative plasmids such as pSK41 has not been established. It is unlikely that one or both of the palA sequences flanking the qacD gene are required for replication of pSK41, as natural derivatives of this plasmid such as pSH6 and pSK8 (McDonnell et al., 1983) lack the region that contains these palA sequences. Interestingly, the DR2_D sequence in pSK41 shows complete identity with sequences found 3' to the lna gene in pIP855 (Brisson-Noël and Courvalin, 1986). This identity extends 5 bp 5′ of the DR7_D sequence; it is tempting to speculate that the palA sequences on pIP855 also arose from pSK41.

(b) Conclusions

The sequence of the antiseptic and disinfectant resistance genes qacC and qacD was determined; the genes were shown to contain the same ORF of 107 aa in length. The involvement of this ORF in the qac phenotype was confirmed by deletion analysis. The product of the qacC/qacD gene is predicted to be a highly hydrophobic polypeptide that probably resides in the membrane. The qacD gene is flanked by DRs which contain within them sequences sharing extensive identity with palA sequences. palA serves as an origin for lagging-strand replication in small plasmids from Gram + organisms. The qacC gene appears to have arisen from qacD by an insertion event originating from the large conjugative plasmid pSK41 into a small plasmid to give rise to pSK89. In this process it appears that the qac promoter has been rendered nonfunctional so that a promoter different from that used by qacD is probably used by qacC. The entire sequence of the pSK89 plasmid was determined and shown to be homologous at the gene organisation and nt sequence level to other plasmids from Gram + hosts.

ACKNOWLEDGEMENTS

We would like to thank Dr. S.J. Projan, Public Health Research Institute, New York, NY for the pWG32 sequence, Dirk Truman for technical assistance and Mary Byrne and Sue Loh for critical reading of the manuscript. This work was supported by a Project Grant from the National Health and Medical Research Council (Australia).

REFERENCES


