Acta virol. **35**: 497-502, 1991

ANALYSIS OF STRAIN-SPECIFIC PLASMID SEQUENCES FROM *COXIELLA BURNETII*

M. F. MINNICK[1], C. L. SMALL[1], M. E. FRAZIER[2], L. P. MALLAVIA[1]*

[1]Department of Microbiology, Washington State University, Pullman, WA 99164-4233., [2]Biology Department, Battelle, Pacific Northwest Laboratories, Richland, WA 99352 U.S.A.

Received March 25, 1991

Summary. - Acute isolates of *Coxiella burnetii* possess a 36-kbp plasmid termed QpH1. DNA hybridizations show that QpH1 contains a ~6-kbp region of DNA which is not present in the QpRS plasmid from chronic isolates. This QpH1-specific region of DNA contains the contiguous *Eco*RI fragments G, E, and D. The GED region was found to possess seven open reading frames (ORF's) coding for proteins ranging from 5.5 to 42.3 kDa in molecular mass when subcloned and expressed *in vitro*. Summing the predicted ORF's accounts for 95 % of the GED coding potential. *E. coli* expression produced a stable 42.3-kDa protein from the pHIN19 subclone of GED. The ORF of the 42.3-kDa protein, termed *cbhE'*, has been localized on GED by both *in vitro* transcription/translation and DNA sequencing. The *cbhE'* gene is estimated as 1142 bp in length with a putative promoter region of TCAACT (-35)-N_{16}-TAAAAT (-10)-N_{14}-AGAAGGA (Shine-Dalgarno)--N_{10}-ATG.

Key words: Q fever; plasmids; gene expression; rickettsia

Introduction

A strain designation has been proposed which separates *Coxiella burnetii* into six strains or genomic groups based on restriction fragment-length polymorphisms, lipopolysaccharide, and plasmid content (Mallavia *et al.,* 1991). In humans, strains I, II, and III are associated with acute Q fever and possess the QpH1 plasmid (Samuel *et al.,* 1983), while strains IV and V are chronic endocarditis isolates which harbor a plasmid termed QpRS (Samuel *et al.,* 1985; Vodkin *et al.,* 1986) or chromosomally-integrated QpRS sequences (Savinelli and Mallavia, 1990). Strain VI contains a third plasmid, QpDG, and has yet to be found in humans (Hendrix *et al.,* 1991). Although greater than 90 % of all

* To whom correspondence should be addressed

plasmid DNA is conserved (Samuel and Mallavia, unpublished data) each plasmid type contains sequences which are unique. Since correlation exists between plasmid and disease manifestation, unique sequences are being analysed for their potential role in virulence.

Only one gene has been characterized from the plasmids of *C. burnetii*. This gene, termed *cbbE'*, is unique to the QpRS plasmid. The E' protein encoded by *cbbE'*, is surface-exposed and is unique to chronic strain IV *C. burnetiii* (Minnick *et al.*, unpublished data).

This report presents data on the analysis of a ~6.0 kbp region of DNA unique to the QpH1 plasmid from acute strains I, II, and III, and which contains a 1.1 kbp gene termed *cbhE'*. *cbhE'* is the first gene to be characterized from the QpH1 plasmid of *C. burnetii*.

Materials and Methods

C. burnetii strains. *C. burnetii* was grown and purified as described by Hendrix and Mallavia (1984). Plasmids and genomic DNA were isolated and purified from the strain I isolate Nine Mile RSA 493, strain II isolate M-44 RSA 459, strain III isolate Koka, strain IV isolates Priscilla Q177 and H WSU 101, strain V isolate Ko Q229, and strain V1 isolate Dugway 7E9-12; as described by Samuel *et al.* (1983).

Genetic manipulations. *C. burnetii* plasmids were subcloned into pUC19 (Yanisch-Perron *et al.*, 1985) by standard procedures (Maniatis *et al.*, 1982). The QpHI-pHK17 recombinant pQHI (Minnick *et al.*, 1990*b*) was also employed. DNA blotting, probe preparation, hybridizations, and washes at high stringency (~7% missmatch) were performed as previously described (Minnick *et al.*, 1990*a*).

Gene expression and characterization. Subcloned fragments of QpH1 were analysed by *in vitro* transcription/translation (IVTT) as before (Minnick *et al.*, 1990a). The translational start site for *cbhE'* was identified by IVTT as previously described (Minnick *et al.*, 1990*b*). The *cbhE'* gene was expressed in *E. coli* DH5α *E. coli* were transformed by the methods of Chung *et al.* (1989). *E. coli* containing pHIN19 were grown to exponential phase ($O.D._{600}$~0.6) under selection with ampicillin (100 μg/ml). Cells were centrifuged, lysed in Laemmli sample buffer (Laemmli, 1970), and analysed by sodium dodecylsulphate-polyacrylamide gel electrophoresis (SDS-PAGE) (Laemmli, 1970). Protein bands were visualized by staining the gel with Coomassie blue-R. Double-stranded DNA sequencing was performed as before (Minnick *et al.*, 1990*a*) using the methods of Sanger *et al.* (1977).

Results and Discussion

DNA hybridization analyses identified a 6.0-kbp region of DNA which was unique to the QpH1 plasmid (Samuel and Mallavia, unpublished data). A partial restriction map of this sequence and three subclones used in its characterization (i.e. pXBA1, pQHG1, and pHIN19) are given in Fig. 1. This unique stretch of DNA, termed GED, contains the QpH1 *Eco*RI fragments previously designated G (1200 bp), E (2470 bp), and D (3844 bp) on the basis of relative size (Samuel *et al.*, 1985). Only 2.35 kbp of the D fragment adjacent to E

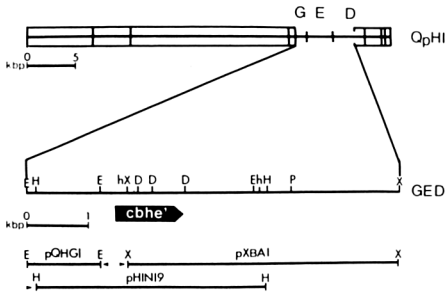

Fig. 1

*Eco*RI restriction map of the QpH1 plasmid of *C. burnetii,* showing regions conserved in all *C. burnetii* plasmids (boxed) and the QpH1-specific *Eco*RI fragments G, E, and D

A detailed restriction map of GED is also given, showing the position and direction of the *cbhE'* gene (large arrow). Three pUC19 subclones of GED are shown with the direction of the *lacZ'* gene indicated by small arrows. (Abbreviations: D, *Dra*I; E, *Eco*RI; h, *Hinc*II; H, *Hind*III; P, *Pst*I, *Xba*I).

is unique to QpH1. A 4.4-kbp *Xba*I subfragment of GED was cloned into pUC19 to produce pXBA1, and was used to analyse the core of the GED region. pXBA1 did not hybridize when probed with (^{32}P)-QpRS, confirming that it is also unique to QpH1 (Fig. 2).

Compilation of the IVTT analyses for pXBA1, pQHG1, and pHIN19 subclones of GED, suggests that seven proteins are encoded on GED and are expressed *in vitro* with apparent molecular masses (M_r) of ~5.5, 12.3, 34, 35.8, 40.9, and 42.3 kDa (Fig.3). A summary of the protein M_r's and their estimated

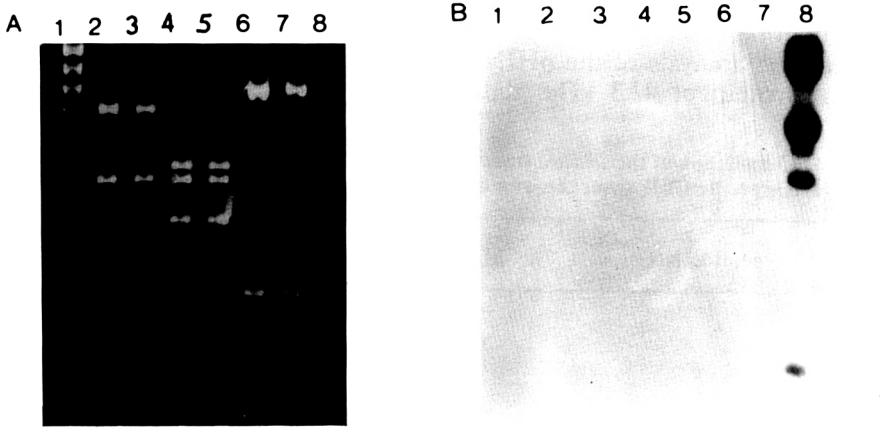

Fig. 2

DNA hybridization analysis of the pXBA1 subclone of GED

pXBA1 was digested with restriction enzymes, electrophoresed in a 1 % agarose gel (A) then blotted and probed with (^{32}P)-QpRS as previously described (Minnick *et al.,* 1990*a*). The autoradiograph is shown in (B). Lane 1, lambda *Hind*III, DNA; lanes 2 and 3, pXBA1 cut with *Xba*I; lanes 4 and 5, pXBA1 cut with *Eco*RI plus *Xba*I; lanes 6 and 7, pXBA1 cut with *Sph*I; and lane 8, QpRS cut with *Eco*RI.

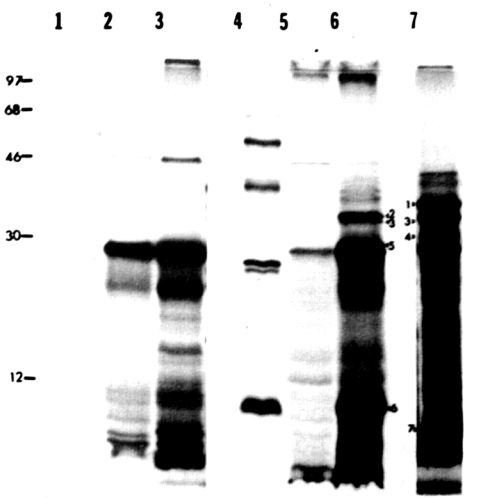

Fig. 3
In vitro transcription/translation (IVTT) analysis of subclones of the GED sequence of QpH1
Approximately 2 x 10^5 DPM of /^{35}S/-labelled proteins were separated by SDS-PAGE (12.5% acrylamide), dried and exposed to X-ray film overnight. The autoradiograph is shown. Lanes: 1 and 4 / ^{14}C/-M_r standards in kDa; 2 and 5, IVTT products of the cloning vector pUC19; 3, 6, and 7, IVTT products of pQHG1, pXBA1, and pHIN19, respectively. The insert-encoded polypeptides are numbered, with the CbhE' protein starred.

open reading frame (ORF) sizes is given in Table 1. Summation of the ORF sizes gives 5687 bp; a figure close to the GED size of 6.0 kbp. The ~313-bp discrepancy in the summed and actual coding potential of GED may be due to lack of compensating for individual gene promoter regions, rho-independent terminators, or noncoding regions of GED.

Maxicell analysis of the pHIN19 subclone gave a single, stable, insert-encoded protein of 42.3 kDa (data not shown). However, IVTT analysis of

Table 1. Compilation of the *in vitro* transcription/translation (IVTT) analysis on the subclones of the QpH1-specific GED sequence, and estimation[1] of the open reading frames (ORF's) of GED

Polypeptide No.	M_r by IVTT (kDa)	Predicted ORF (bp)
1*	42.3	1142
2	40.9	1104
3	39.8	1075
4	35.8	967
5	34.0	918
6	12.3	332
7	5.5	149

[1] Estimated by 27 bp DNA/kDa protein
* The cbhE' polypeptide
Polypeptide numbers correspond to those given in Fig. 3

pHIN19 suggests that the 42.3-kDa protein plus three other insert-specific proteins are encoded (Fig. 3). Following expression in *E. coli* maxicells the other proteins may be degraded, as has been seen in previous maxicell analyses on cloned *C. burnetii* DNA (Minnick *et al.*, 1990*a*). The 42.3-kDa protein was also detected in cell lysates of exponential-phase-harvested *E. coli* DH5α containing pHIN19 (Fig. 4). The approximate translational start site for the 42.3 kDa protein was mapped by IVTT and the position of its ORF in GED is given in Fig. 1. The gene is termed *cbhE'* for: *C. burnetii* (cb), hamilton strain (h) (Mallavia *et al.*, 1991), presence on the QpH1 E fragment (E), and plasmid origin (prime).

Because *cbhE'* is stably expressed in *E. coli* and is found in strains of *C. burnetii* associated with acute disease, we have begun characterizing the gene. The *cbhE'* gene is ~1142 bp in length and initial sequencing data yields a putative promoter region of TCAACT-(-35)-N_{16}-TAAAAT-(-10)-N_{14}-A-GAAGGA-(Shine-Dalgarno)-N_{10}-ATG. These data suggest that the *cbhE'* promoter regulatory region is more similar to the *E. coli* consensus promoter

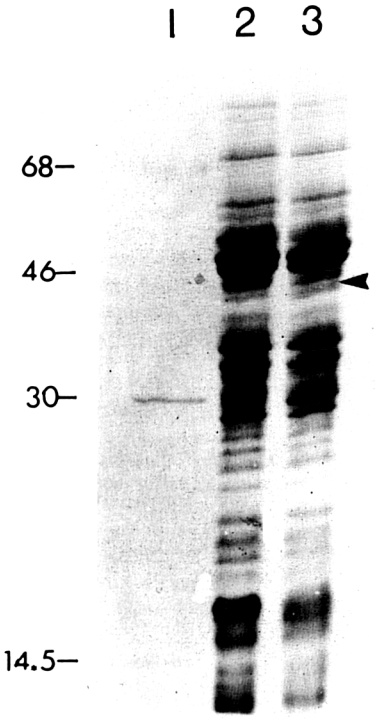

Fig. 4
Expression of the *cbhE'* gene in *E. coli* DH5α containing pHIN19
E. coli lysates (30 μg protein) were separated on 12.5 % SDS-PAGE gels, stained with Coomassie blue-R, and dried. Lane 1, M_r standards in kDa; lane 2, nontransformed DH5α; lane 3, DH5α containing pHIM19. The CbhE' protein is arrowed.

than that of the *cbbE'* gene from QpRS (Minnick *et al.*, 1990*a*). These data may explain why *cbhE'* expression in *E. coli* is apparently greater than was observed for *cbbE'* (Minnick *et al.*, 1990*a*).

The 6.0-kbp GED region of QpH1 is a contiguous stretch of plasmid DNA coding for at least seven polypeptides including CbhE'. These genes and gene products represent unique markers which could be used to distinguish between *C. burnetii* strains associated with acute or chronic disease, and may code for virulence determinants involved in the acute manifestations of Q fever.

References

Chung, C. T., Niemela, S. L., and Miller, R. H. (1989): One-step preparation of competent *Escherichia coli:* transformation and storage of bacterial cells in the same solution. *Proc. natn. Acad. Sci. U.S.A.* **86** 2172-2175.

Hendrix, L., and Mallavia, L. P. (1984): Active transport of proline by *Coxiella burnetii. J. gen. Microbiol.* **130**, 2857-2863.

Hendrix, L., Samuel, J. E., and Mallavia, L. P. (1991): Differentiation of *Coxiella burnetii* isolates by restriction endonuclease-digested DNA separated on SDS-PAGE. *J. gen. Microbiol.* **137**, 269-276.

Laemmli, U. K. (1970): Cleavage of structural proteins during the assembly of the head of bacteriophage T4. *Nature (London)* **227**, 680-685.

Mallavia, L. P., Samuel J. E., and Frazier, M. E. (1991): The genetics of *Coxiella burnetii:* etiologic agent of Q fever, pp. 259-284. In J. C. Williams, H. Thompson (Eds): *Q Fever, The Biology of Coxiella burnetii.* CRC Press, Boca Raton, Fl.

Maniatis, T., Fritsch, E. F., and Sambrook, J. (1982): *Molecular Cloning: a Laboratory Manual.* Cold Spring Harbor, NY: Cold Spring Harbor Laboratory.

Minnick, M. F., Heinzen, R. A., Frazier, M. E., and Mallavia, L. P. (1990*a*): Characterization and expression of the *cbbE'* gene of *Coxiella burnetii. J. gen. Microbiol.* **136**, 1099-1107.

Minnick, M. F., Heinzen, R. A., Douthardt, R., Mallavia, L. P., and Frazier, M. P. (1990*b*): Analysis of QpRS-specific sequences from *Coxiella burnetii. Ann. NY Acad. Sci* **590**, 514-522.

Minnick, M. F., Heinzen, R. A., Frazier, M. E., and Mallavia, L. P. (unpublished data).

Samuel, J. E., Frazier, M. E., Kahn, M. L., Thomashow, L. S., and Mallavia, L. P. (1983): Isolation and characterization of a plasmid from phase I *Coxiella burnetii. Infect. Immun.* **41**, 488-493.

Samuel, J. E., Frazier, M. E., and Mallavia, L. P. (1985): Correlation of plasmid type and disease caused by *Coxiella burnetii. Infect. Immun.* **49**, 775-779.

Samuel, J. E., and Mallavia, L. P. (unpublished data).

Sanger, F., Nicklen, S., and Coulson, A. R. (1977): DNA sequencing with chain-terminating inhibitors. *Proc. natn. Acad. Sci. U.S.A.* **74**, 5463-5467.

Savinelli, E. A., and Mallavia, L. P. (1990): Comparison of *Coxiella burnetii* plasmids to homologous chromosomal sequences present in a plasmidless endocarditis-causing isolate. *Ann. NY Acad. Sci.* **590**, 523-533.

Vodkin, M. H., Williams, J. C., and Stephenson, E. H. (1986): Genetic heterogeneity among isolates of *Coxiella burnetii. J. gen. Microbiol.* **132**, 455-463.

Yanisch-Perron, C., Vieira, J., and Messing, J. (1985): Improved M13 phage cloning vectors and host strains: nucleotide sequences of the M13mp18 and pUC19 vectors. *Gene* **33**, 103-119.