Molecular organization of genes constituting the virulence determinant on the *Salmonella typhimurium* 96 kilobase pair plasmid

Suvi Taira and Mikael Rhen

Molecular Biology Unit, National Public Health Institute, Mannerheimintie 166, 00300 Helsinki, Finland

The ability of intracellular growth is plasmid-dependent in *Salmonella typhimurium*. Only a small portion of this 96 kilobase pair plasmid appears essential for intracellular growth. The genetic organization of this region (the essential virulence determinant) was resolved. Fragments of the virulence determinant were cloned from the 96-kb plasmid pEX102 and transformed into minicell-producing *E. coli*. Plasmid-directed protein synthesis was investigated in metabolically labeled minicells. This analysis indicated the presence of at least four genes, *mkaA*, *mkaB*, *mkaC* and *mkaD*, within the virulence determinant encoding proteins of 70, 31, 30 and 29 kDa, respectively. The genes were positioned on the restriction map of the 96-kb virulence plasmid and the map locations confirmed by nucleotide sequence analysis of two new virulence genes (*mkaB* and *mkaC*)

Intracellular growth; Gene organization; Virulence gene

1. INTRODUCTION

The genus *Salmonella* exemplifies enteric bacteria with the ability to grow within eucaryotic host cells [1]. The infection pathogenesis of serovars causing invasive forms of infection contains a phase of several days' duration during which the bacteria grow within macrophages of the liver and spleen [1,2]. This growth phase is essential for the complete infection process; strains incapable of intracellular growth are avirulent [3–5].

A number of factors contribute to the virulence in *Salmonella*. These include intact lipopolysaccharide, the ability to synthesize aromatic amino acids and nucleotides [5,6] and invasion of eucaryotic cells [7]. These properties are chromosomally encoded. The ability of several serovars of *Salmonella* to grow within macrophages is additionally dependent on a large virulence plasmid [3,4,8–13].

The plasmid genes responsible for intramacrophage growth appear clustered in a region less than 7 kb in size in *S. typhimurium* [13–16]. Fine mapping of this region has so far revealed two closely linked virulence genes; *mkaA* encoding a 70 kDa protein [17] and *mkfA* encoding a 28 kDa protein [16–19]. Insertional inactivation of either gene has been shown to prevent net growth of bacteria within the liver and spleen resulting in complete loss of virulence. In addition to this 'essential' virulence determinant, Norel and co-workers [16] have presented evidence for another virulence determi-

nant some 15 kb apart encompassing at least two genes. Inactivation of these genes resulted in a partial decrease in virulence.

In this communication we carried out further fine mapping of the essential virulence determinant in order to gain more information of its structure and function. The determinant contains at least four genes, all of which are transcribed from the same DNA strand.

2. MATERIALS AND METHODS

2.1. *Bacterial strains, phages and cultivation*

S. typhimurium virulence plasmids were isolated from TML R66 strains SL2965, SH8224 (pEX102, zzx-2556::Tn5) and SH8269 (pEX102, zzx-2558::Tn5) [15]. *E. coli* HB101 [20] (*leu, pro, lac, thi, hsds, strA, recA*) was used as the cloning host. These strains were cultivated in L-broth [21] or on L-agar plates supplemented with appropriate antibiotics (100 μg/ml ampicillin; 30 μg/ml chloramphenicol; 20 μg/ml kanamycin; 10 μg/ml tetracycline). All antibiotics were from Sigma Chemical Co., St. Louis, USA. Minicells were generated from *E. coli* strain PK251 [22] (*minA, minB*) grown in Brain Heart Infusion Broth. M13 phages were propagated on *E. coli* JM101. The M13 phages used were M13 mp8 and mp9 [23].

2.2. *Recombinant DNA techniques, construction of recombinant plasmids*

Plasmid isolations, restriction analyses, ligations and transformations were performed by standard procedures essentially as described [24,25]. Restriction endonucleases and T4-DNA ligase were obtained from Amersham International (Buckinghamshire, England).

Recombinant plasmids pKTH3065, 3065-1 and 3065-2 were available from a previous work [17]. pKTH3067 was obtained by cloning the 0.9-kb *Eco*RI fragment from pKTH3065 harbouring the Tn1725 element [17] shown in fig.1a into the *Eco*RI site of pACYC184 [26], whereas pKTH3065-3 is a simple *Pst*I cutback derivative of pKTH3065. *Cla*I-*Eco*RI digests of pEX102 and pEX102 zzx-2558::Tn5 were cloned in the *Eco*RI site of pACYC184 in order to produce pKTH3068 and 3069, respectively. The desired inserts were identically positioned into pACYC184 with the use of a 0.3-kb

Correspondence address: S. Taira, Molecular Biology Unit, National Public Health Institute, Mannerheimintie 166, 00300 Helsinki, Finland
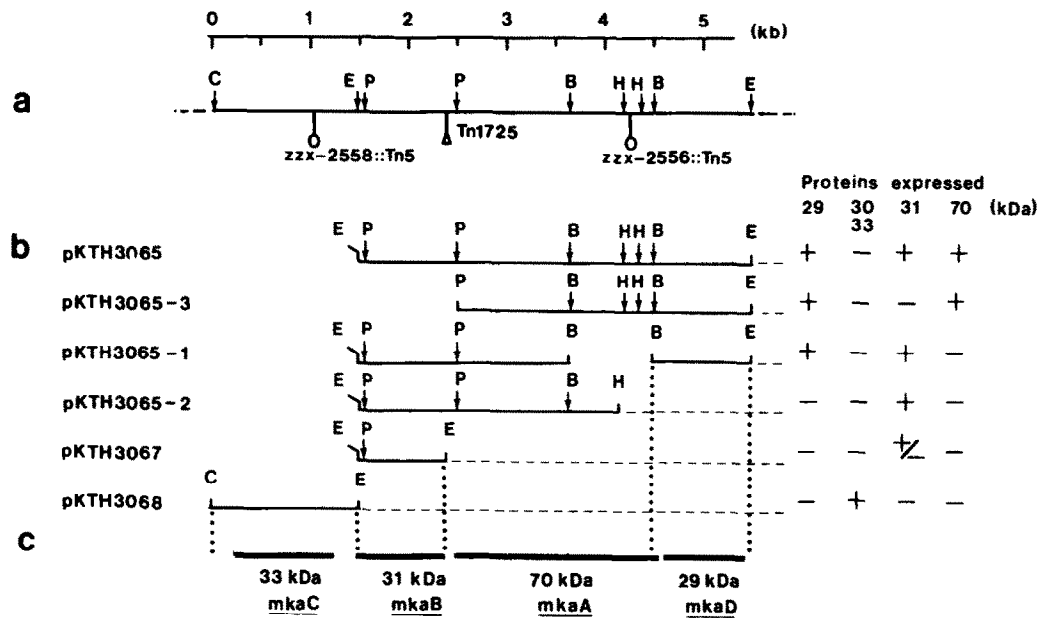
Fig.1. (a) The region of the 96-kb pEX102 plasmid involved in virulence. (b) Shows parts of the virulence region contained in the recombinant plasmids indicated. The table on the right summarizes the specific proteins encoded by the plasmids. (c) The suggested gene organization in this region. Restriction sites are abbreviated as follows: B, *Bam*HI; C, *Cla*I; E, *Eco*RI; H, *Hind*III; P, *Pst*I. Symbols: ⵱ Tn5 insertion; ⵛ Tn1725 insertion.

pEX102-derived *Cla*I-*Eco*RI fragment also ligated into the *Eco*RI site of the vector.

2.3. Analysis of plasmid-directed protein synthesis in minicells

Minicells were isolated and labeled with ^{35}S-methionine (Amersham) at a concentration of 50 μCi/ml as described [27]. Washed, labeled minicells were then solubilized in SDS-polyacrylamide (PAGE) sample buffer and their protein content electrophoretically separated in 12 or 15% SDS-PAGE gels [28]. The gels were fixed in 10% (v/v) acetic acid, impregnated with Amplify^r (Amersham), dried and exposed on Kodak X-Omat AR films (Kodak, Rochester, NY). A ^{14}C-methylated protein molecular mass standard (Amersham) was utilized in order to estimate apparent molecular masses.

2.4. Nucleotide sequence analysis

Nucleotide sequence analysis was performed as a standard procedure [29] using the Sequenase kit (United States Biochemical Corporation, Cleveland, OH), and recombinant M13 8mp and 9mp phage DNA as template. Both synthetic heptadecamers, complementary to sequenced stretches, and commercially available universal primer were used as primers in the sequencing reactions. Both strands were sequenced for each gene. All sequence computer analyses were performed according to the PC/Gene (Genofit) program.

3. RESULTS AND DISCUSSION

3.1. Proteins encoded by the virulence determinant

The essential virulence determinant has, based on several investigations on the *S.typhimurium* [16–18] and *S.dublin* [14,30] plasmids, been mapped down to the apparently common region outlined in fig.1a. The proteins expressed by subclones constructed from this region in isolated, metabolically labeled *E.coli* minicells are shown in fig.1b and 2.

The virulence gene described by us, *mkaA* [17], en-

codes a 70 kDa protein (MkaA) expressed here by pKTH3065 (fig.1b; fig.2, lanes E and G). In addition to MkaA and the β-lactamase protein encoded by the cloning vector pBR325, pKTH3065 expressed two additional major proteins, apparently 29 and 31 kDa in size (fig.2, lanes B and G). Here, we have termed these proteins MkaD and MkaB, respectively.

Next we constructed a set of pKTH3065 deletions. Deletion of the region to the left of the *Pst*I site at 2.5 kb (as in pKTH3065-3 in fig.1) resulted in loss of the expression of MkaB and β-lactamase as a result of a concomitant deletion of this gene on the cloning vector pBR325 (fig.2, lanes E and H). Deletion of only the 0.8-kb *Bam*HI fragment (pKTH3065-1) abolished expression of MkaA, whereas deletion of the region to the right of the *Hind*III sites (pKTH3065-2) abolished also the expression of MkaD (fig.1b). These observations confirm previous mapping and sequencing experiments which positioned *mkaA* at map co-ordinates 2.5 and 4.4 kb [17]. MkaD would then be encoded by the region between co-ordinates 4.4 and 5.2 (fig.1a and b) and is therefore likely to correspond, based on available restriction maps, to the 28 kDa virulence protein [16,18].

The gene for MkaB (*mkaB*) would then logically map to the left of *mkaA*. To confirm this conclusion we constructed pKTH3067, a pACYC184-derivative containing the 0.9-kb *Eco*RI fragment created by the *Eco*RI site, a map co-ordinate 1.5 kb and the Tn1725 element [17] at 2.4 kb. Attempts to express MkaB from pKTH3067 resulted in a smear of bands starting at 31 kDa (fig.2, lane d) indicating possible proteolytic
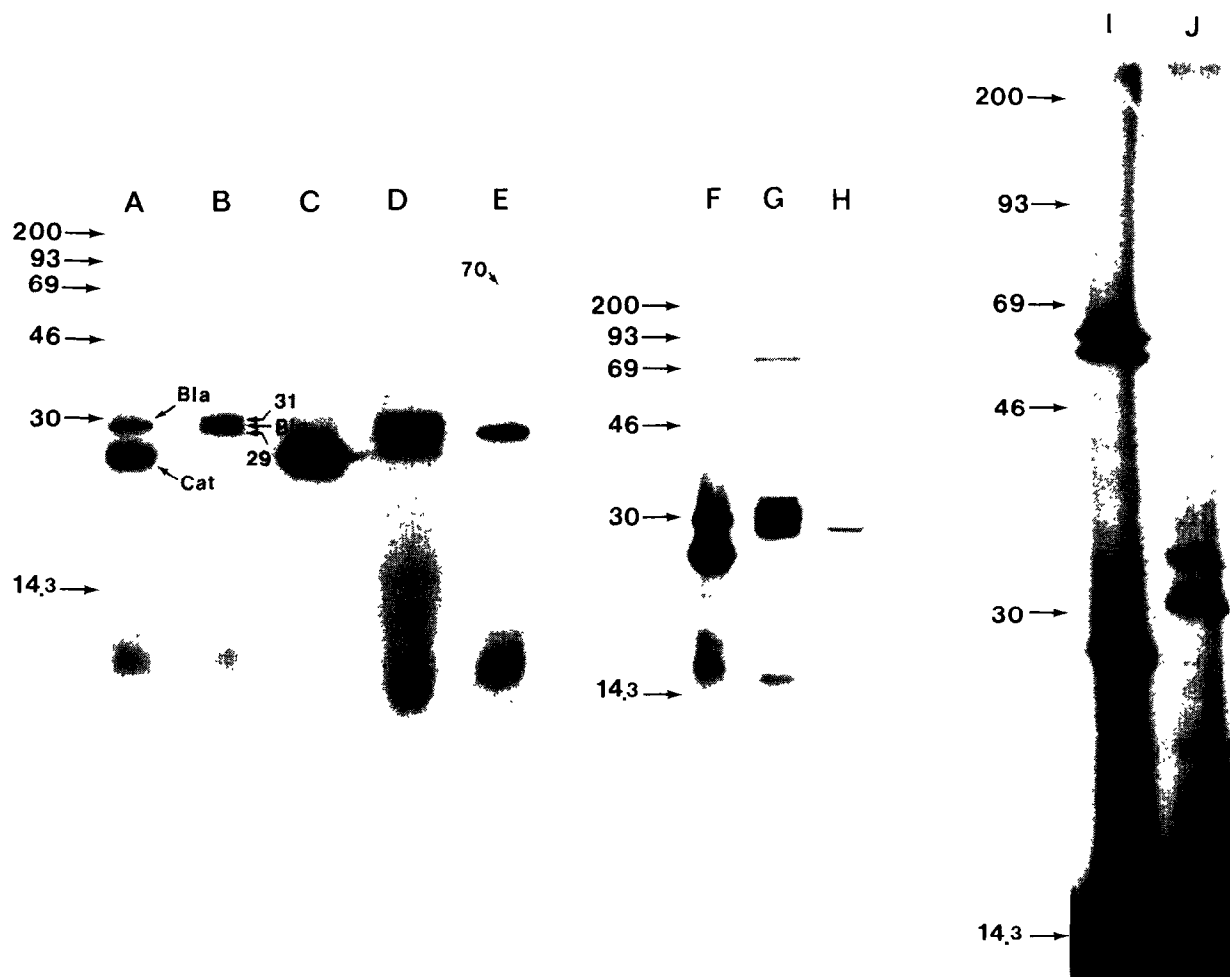
Fig.2. Fluorograms of ^{35}S-methionine labeled *E.coli* minicells run in 15% (lanes A–E) and 12% (F–J) SDS-PAGE. The plasmids contained are as follows: A and F, pBR322, B and G, pKTH3065; C, pACYC184; D, pKTH3067; E and H, pKTH3065-3. Numbers on the left (in kDa) indicate the positions of molecular mass markers.

degradation of MkaB in the absence of MkaA or MkaD.

Finally, the *Cla*I-*Eco*RI fragment in pKTH3068 representing map co-ordinates 0 to 1.5 kb (fig.1b) expressed apparently two novel proteins 30 and 33 kDa in size (fig.2, lane J).

3.2. Identification of mkaC as a virulence gene

We have previously described two Tn5 insertion mutations of the *S.typhimurium* virulence plasmid pEx102, zzx-2556::Tn5 and zzx-2558::Tn5, that abolish virulence in experimental mouse salmonellosis [15]. zzx-2556::Tn5 is situated within *mkaA* (fig.1a, [17]), and zzx-2558::Tn5 within a region of pEX102 that corresponds to the 1.5-kb *Cla*I-*Eco*RI fragment of PKTH3068 (fig.1a). We also cloned the corresponding *Cla*I-*Eco*RI fragment from the pEX102-derivative harbouring zzx-2558::Tn5 into pACYC184. This fragment (pKTH3069) did not express the 30 or 33 kDa proteins in minicells; instead a 28 kDa peptide appeared that

could represent a truncated form of either protein as a result of the insertion mutation (fig.2, lanes I and J). We termed the corresponding gene *mkaC*.

3.3. Nucleotide sequence analysis of mkaB and mkaC

In order to characterize *mkaB* and *mkaC* more closely and to confirm their map locations we sequenced separately the 1.5-kb *Cla*I-*Eco*RI and the 0.9-kb *Eco*RI fragments from pKTH3068 and pKTH3067, respectively. Each fragment contained only one open reading frame sufficient to code for the observed proteins.

The sequence of the 0.9-kb *Eco*RI fragment revealed a 765-bp open reading frame (ORF), with a coding capacity for a 255 amino acid protein here suggested to correspond to MkaB (fig.3). The methionine residue at 64 is also preceded by a potential ribosome-binding site (nucleotides 267–270), however, translation from this site would result in a much shorter protein. *mkaB* is read from the same strand as *mkaA*; the genes are separated by 183 bp.

mkaC is suggested to consist of the 892-bp ORF present in the 1.5-kb *ClaI-EcoRI* fragment (fig.4). The predicted MkaC protein would thus consist of 297 amino acid residues and have a calculated molecular mass of 33.8 kDa. The insertion zzx-2558::Tn5 is situated between nucleotides 604 and 605 obviously resulting in a truncated protein. Again, *mkaC* is read from the same strand as *mkaA* and *mkaB*. We could not define any other reading frame large enough to correspond to the 30 kDa protein. Possibly, one or the other protein band represents abnormally moving or posttranscriptionally modified forms of MkaC. MkaB and MkaC did not show any remarkable sequence homology with proteins listed in the Swiss protein sequence bank.

Taken together, our analysis of proteins expressed by pKTH3065, pKTH3068 and their derivatives (fig.1b) indicated the presence of at least four genes within the virulence determinant of the *S.typhimurium* virulence plasmid pEX102; *mkaC, mkaB, mkaA* and *mkaD* (fig.1c). The positions of *mkaC* and *mkaB* were verified by nucleotide sequence analysis (figs 3 and 4), whereas *mkaA* and *mkfA*, the counterpart for *mkaD* in the virulence plasmid pIP1350, have been sequenced previously [17,19]. Insertional inactivation of *mkaA* or *mkaC* abolishes virulence indicating that they are true virulence genes. The same is likely to hold for *mkaB* and *mkaD* as well, as several Tn5 insertions affecting virulence have been mapped on the virulence plasmids of *S.typhimurium* [16] and *S.dublin* [14] to positions

Fig.4. The nucleotide sequence representing the *mkaC* gene. The predicted amino acid sequence is shown underneath. Numbers on the right indicate the nucleotide vs amino acid residue at the end of the line. The arrowhead within the sequence indicates the position of zzx-2558::Tn5. The potential ribosome-binding site is underlined.

clearly corresponding to these two genes. Interestingly, all four genes are read from the same DNA strand (figs 3 and 4, [17,19]) suggesting that they may constitute a functional unit or possibly even a regulon.

Fig.3. The nucleotide sequence representing the *mkaB* gene. The predicted amino acid sequence is shown underneath. Numbers on the right indicate the nucleotide vs amino acid residue at the end of the line. The potential ribosome-binding site is underlined.

REFERENCES

[1] Carter, P.B. and Collins, F.M. (1974) J. Exp. Med. 139, 1189–1203.

[2] Heffernan, E.J., Fierer, J., Chimaki, G. and Guiney, D. (1987) J. Infect. Dis. 155, 1254–1259.

[3] Hackett, J., Kotlarsky, J., Mathan, V., Francki, K. and Rowley, D. (1986) J. Infect. Dis. 153, 1119–1125.

[4] Fields, P.I., Groisman, E.A. and Heffron, F. (1989) Science (NY) 243, 1059–1062.

[5] Fields, P.I., Swanson, R.W., Haidaris, C.G. and Heffron, F. (1987) Proc. Natl. Acad. Sci. USA 83, 5189–5193.

[6] Lindberg, A.A. (1977) Scand. J. Infect. Dis. (suppl.) 24, 86–91.

[7] Finlay, B.B., Starnbach, M.N., Francis, C.L., Stocker, B.A.D., Chatfield, S., Dougan, G. and Falkow, S. (1988) Mol. Microbiol. 2, 757–766.

[8] Hovi, M., Sukupolvi, S., Edwards, M.F. and Rhen, M. (1988) Microb. Pathogen. 4, 385–391.

[9] Gulig, P.A. and Curtiss, R., iii (1987) Infect. Immun. 55, 2891–2901.

[10] Helmuth, R., Stephen, R., Bunge, C., Hoog, B., Steinbeck, A. and Bulling, E. (1983) Infect. Immun. 48, 175–182.

[11] Terakado, N., Sekizaki, T., Hashimoto, K. and Naitoh, S. (1983) Infect. Immun. 41, 443–444.

[12] Barrow, A.A., Simpson, J.M., Lovell, M.A. and Binns, M.M. (1987) Infect. Immun. 55, 388–392.

[13] Kawahara, K., Haraguchi, Y., Tsichimoto, M., Terakado, N. and Danbara, H. (1988) Microb. Pathogen. 4, 155–163.

[14] Beninger, P.R., Chikami, G., Tanabe, K., Roudier, C., Fierer, J. and Guiney, D.G. (1988) J. Clin. Invest. 81, 1341–1347.

[15] Rhen, M., Virtanen, M. and Mäkelä, P.H. (1989) Microb. Pathogen. 6, 153–158.

[16] Norel, F., Coynault, C., Miras, I., Hermant, D. and Popoff, M.Y. (1989) Mol. Microbiol. 3, 733–743.

[17] Taira, S. and Rhen, M. (1989) Microb. Pathogen, in press.

[18] Gulig, P.A. and Curtiss, R., iii (1988) Infect. Immun. 56, 3262–3271.

[19] Norel, F., Pisano, M.R., Nicoli, J. and Popoff, M.Y. (1989) Res. Microbiol. 140, 263–265.

[20] Boyer, H.W. and Roulland-Dussoix, D. (1969) J. Mol. Biol. 41, 456–472.

[21] Lennox, E.S. (1955) Virology. 1, 190–206.

[22] Liljeström, P., Màattänen, P.L. and Palva, E.T. (1982) Mol. Gen. Genet. 188, 190–194.

[23] Messing, J. and Vieirera, A. (1982) Gene 19, 269–276.

[24] Maniatis, T., Fritsch, E.F. and Sambrook, J. (1982) in: Molecular Cloning: A Laboratory Manual. Cold Spring Harbour Laboratory, Cold Spring Harbour, NY.

[25] Rhen, M. (1985) J. Gen. Microbiol. 131, 571–580.

[26] Chang, A.C. and Cohen, S.N. (1978) J. Bacteriol. 134, 1141–1156.

[27] Dougan, G. and Kehoe, M. (1984) Methods. Microbiol. 7, 233–258, Academic Press.

[28] Laemmli, U.K. (1970) Nature (London) 227, 680–685.

[29] Sanger, F., Nicklen, S. and Coulson, A.R. (1977) Proc. Natl. Acad. Sci. USA 74, 5463–5467.

[30] Williamson, C.M., Pullinger, G.D. and Lax, A.J. (1988) Microb. Pathogen. 5, 469–473.

[31] Bolivar, F. (1978) Gene 4, 121–136.