Volume 297, number 1,2, 77–80 FEBS 10634 February 1992

A new gene of the *f1* operon of *Y. pestis* involved in the capsule biogenesis

A.V. Karlyshev, E.E. Galyov, O.Yu. Smirnov, A.P. Guzayev, V.M. Abramov and V.P. Zav'yalov

Institute of Immunology, State Concern 'Biopreparation', 142380 Lyubuchany, Moscow Region, USSR

Received 25 November 1991

The DNA sequence determination of the *f1* operon between the genes encoding the F1 subunit (*caf1*) and chaperone-like protein (*caf1M*) revealed a large open reading frame that codes for a polypeptide similar to some *E. coli* proteins involved in the biogenesis of fimbria. The deletion and *in trans* complementation analyses showed that this gene is not necessary for extracellular transport of the F1 subunit but plays a role in the capsule assembly.

Capsular antigen; Protein secretion; Nucleotide sequence; Complementation analysis; *Y. pestis*

1. INTRODUCTION

Y. pestis cells are known to form unusually large capsules coating the bacteria, the F1 protein subunit being its main (if not the sole) component [1]. We have previously cloned the *f1* operon in *E. coli* and found two genes, one of which codes for the capsule subunit (*caf1*) [2] and the other for a chaperone-like protein (*caf1M*) [3] responsible for the subunit secretion. A contiguous DNA region of about 2.5 kb with unknown function was located between these two genes. In order to elucidate the role of this intergenic region in the process of capsule formation, we carried out the DNA sequence determination and complementation analysis. The data obtained show that there is a sole large open reading frame between these two genes. This region was found to be of importance for capsule assembly. A possible role of the corresponding polypeptide is discussed.

2. MATERIALS AND METHODS

2.1. *Bacterial strains, antigen determination*

The *E. coli* strain JM103 was used in repression–induction experiments. For induction of the *lac*-promoter, the overnight culture grown in 1 ml LB broth at 37°C was one-tenth diluted with fresh medium, incubated for 1 h followed by the addition of IPTG to 1 mM and incubation at 37°C for 4 h. The bacterial cells were washed with 0.015 M NaCl, lysed by 20 μl of mixture (1 mg/ml lysozyme, 10 mM Tris-HCl, 1 mM EDTA, 0.1% Triton X-100, pH 8.0), diluted to 1 ml by buffer (0.015 M NaCl, 1 mM EDTA, 10 mM Tris-HCl, pH 8.0), and F1 antigen was tested by a passive hemagglutination reaction. Anti-plague agglutinating antiserum and erythrocytes conjugated with polyclonal anti-F1 antibodies were obtained from All-Union Anti-Plague Institute 'Microb' (Saratov, USSR).

Correspondence address: A.V. Karlyshev, Institute of Immunology, 142380, Lyubuchany, Chekhov District, Moscow Region, USSR

2.2. *DNA manipulations*

Plasmids p12R, pF18L and pF19R were described previously [2,3]. Unidirectional deletion derivatives of p12R were obtained by *Sma*I and *Kpn*I hydrolysis followed by treatment with *Exo*III and mung bean nucleases [4]. Plasmids pKM1 and pKM4 were obtained by cloning of the *Pvu*II fragment of the plasmid pF18L in opposite orientations into the *Hin*cII site of pACYC177. DNA cloning and sequencing by the Maxam–Gilbert and Sanger–Coulson methods were performed essentially as described by Maniatis et al. [4]. DNA and protein sequences analyses were made using the GENEPRO software package from Riverside Scientific Enterprises and PCGENE package from Genofit.

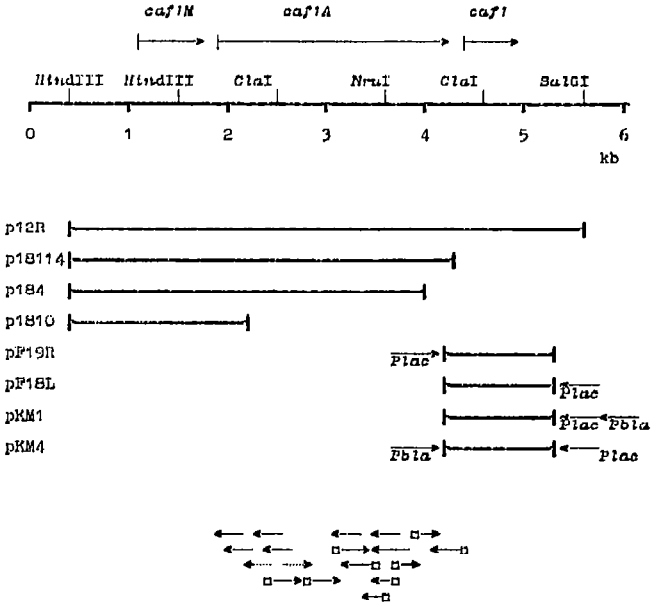

Fig. 1. Restriction map and the sequencing strategy of the *f1* operon. Solid bars indicate cloned fragments. Arrows above the restriction sites indicate open reading frames. Short arrows underneath indicate lengths and directions of the sequences determined; arrows with squares refer to the sequences obtained by internal primers; dotted arrows refer to the sequences obtained by the Maxam–Gilbert technique; other arrows refer to the sequences obtained from deletion derivatives by means of M13 direct primer.

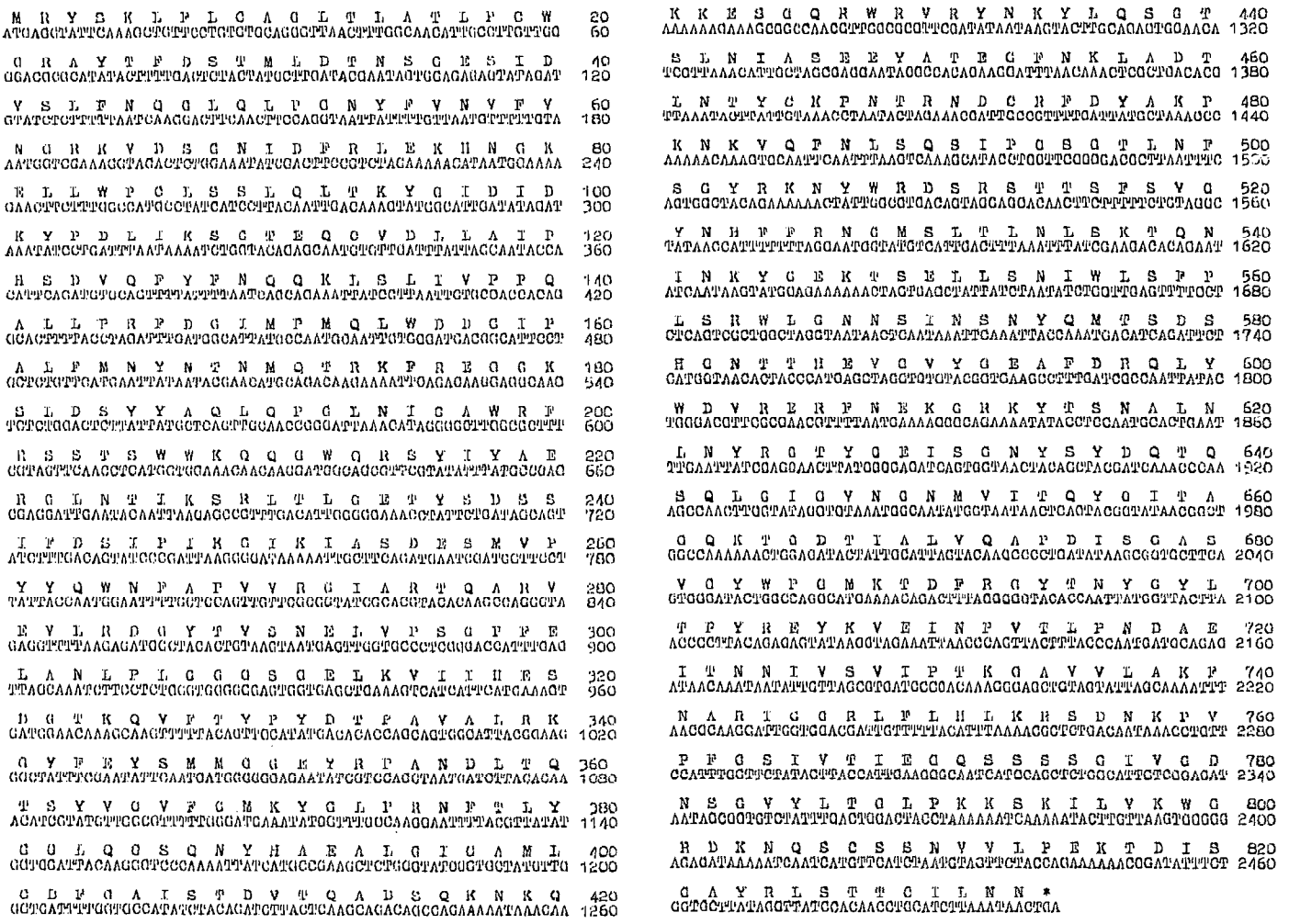

Fig. 2. Nucleotide sequence of *caf1A* gene and deduced amino acid sequence.

3. RESULTS AND DISCUSSION

The unidirectional deletion derivatives of p12R were used for DNA sequence determination and complementation analysis (Fig. 1). The sequencing of the *f1* operon between *caf1* and *caf1M* genes revealed a sole long reading frame encoding a polypeptide with M_r of 93,205 Da (Fig. 2). The hypothetical polypeptide was designated as Caf1A. A PIR'89 database homology search revealed stretches of homology between Caf1A and proteins FaeD [5], PapC [6] and FimD [7] involved in the assembly of *E. coli* fimbria. An overall homology between amino acid sequences of these polypeptides and Caf1A was found to be 18% for FaeD and FimD, and 13% for PapC. Fragments of alignments displaying maximum homology are shown in Fig. 3. The Caf1A protein like FaeD precursor contains a leader-like NH_2-terminal portion of 25 amino acid residues. This fragment has n, h and c regions typical of leader sequences and a potential cleavage site that conforms to the (−3,−1) rule [9]. The secondary structure prediction analysis revealed that all four proteins are mainly β structural, they have similar M_r and amino acid composition. These data allow us to suggest a similarity in their functions. To understand the role that Caf1A can play in capsular biogenesis we have carried out a complementation anal-

Table I

Production and secretion of F1 antigen by *E. coli* cells, containing corresponding plasmids.

Plasmids	Production	Secretion	Agglutination
p12R	+++	+++	+++
pF19R	++	−	−
pF18L	−	−	−
pKM1	−	−	−
pKM4	+	−	−
pKM1+p184	++	++	−
pKM4+p184	+++	+++	−
pKM4+p1810	+++	+++	−
pKM4+p18114	+++	+++	+++

```
CAF1A 127 YFNQQRLSLIVPPQALLPRFDGIMPNQLWDDGIPALFMNYNTNMQTRRFRSGGNS 181
FIMD  150 DVGQQRLNLTIPQAFMSNRARGYIPPELWDPGINAGLLNYNFSGNSVQNRIGGNS 204
PAPC  138 DTGAGVLRINMPQANLEYSDATNLPPSRWDDGIPGLMLDYNLNGTVSRNYQGGDS 192
FAED  136 DVSTQSLALSIPQKGLVRMPENVD----WDYGTSAFRVNYNANANTGRNNTSAFG 186

CAF1A 182 LDSYYAQLQPGLNIGAWRFRSSTSWN 207
FIMD  205 -HYAYLNLQSGLNIGAWRLRDNTTWS 229 —3.5·10^-26
PAPC  193 -HQFSYNGTVGGNLGPWRLRADYQGS 217 —1.4·10^-19
FAED  187 SADLRA------NIGSWVVSSSATAS 206 —2.1·10^-12

CAF1A 219 AERGLNTIRSRLTLGETYSDSSIFDSIPIKGIRIASDESMVPYYQNNFAPVVRGI 273
FIMD  250 LERDIIPLRSRLTLGDGYTQGDIFDGINFRGAQLASDDNMLPDSQRGFAPVIEGI 304
PAPC  239 LFRAIPRNRANLTLGENNINSDIFRSNSYTGASLESDDRMLPPRLRGYAPQITGI 294
FAED  222 ATRAIRALSADLAVGRTSTGDSLLGSTGTYGVSLSRNNSMKPGNL-GYTPVFSGI 275

CAF1A 274 ARTQARVEVLRDGYTVSNELVPSGPFELANLPL 306
FIMD  305 ARGTAQVTIKQNGYDIYNSTVPPGPFTINDIYA 337 —4.7·10^-22
PAPC  295 AETNARVVVSQQGRVLYDSMVPAGPFSIQDLDS 327 —1.1·10^-9
FAED  276 ANGPSRVTLTQNGRLLESEMVPAGPFSITDVPL 308 —3.0·10^-5

CAF1A 703 YREYKVEINPVTLPNDAEITNNIVSVIPTKGAVVLAKFNARIGGRLFLEGLKRSDN 757
FIMD  751 YRENRVALDTNTLADNVDLDNAVANVVPTRGAIVRAEFKARVGIKLLMTLTE-NN 804
PAPC  713 YYRNTTSVDLKRLPDDVEATRSVVESALTEGAIGYRKFSVLKGKRLFAILRLADG 767
FAED  697 YDNNTVTIDTGTLPLSTELTNTSQKVVPTDRAVVNMPFDALKVKRYLLQVKQRDG 751

CAF1A 758 KPVPFGSIVTIEGQSSSSGIVGDNSGVYLTGLPKKSKILVKWG 800
FIMD  805 KPLPFGAMVTSES-SQSSGIVADNGQVYLSGMPLAGKVQVKWG 846 —2.3·10^-30
PAPC  768 SQPPFGASVTSEK-GREGLGMVADEGLANLSGVTPGETLSVNWD 809 —8.8·10^-14
FAED  752 EFVP-GGTTAARDSRNTPLGFVANNGVLMINTVDAPGDITLGQC 793 —4.4·10^-11
```

Fig. 3. Alignment of the fragments of primary structures of the protein Caf1M of *Y. pestis* and FimD, PapC and FaeD proteins of *E. coli*. Amino acid residues coinciding in the sequences compared are indicated by large letters. The probability of random coincidence was estimated by the method in ref. [8].

ysis. For this, *caf1*-gene-lacking but *caf1M*-containing deletion derivatives of p12R (p184, p1810 and p18114) were used (Fig. 1). The *caf1A* gene was completely or partially removed from p12R yielding p1810 and p184, respectively; thus p184 encodes a truncated form of Caf1A. As shown in Table I, pKM4 and pKM1, containing only the *caf1* gene under or without the control of promoter P_{blu} respectively, do not determine the extracellular production of the F1 antigen. It is of interest that in both pKM4 and pKM1 the *caf1* gene is under the control of its own hypothetical promoter, but the F1 synthesis is not observed until p184 or p1810 plasmids have been added. Introduction of p184 and p1810 into cells containing pKM4 not only increases the antigen production but also leads to its secretion. It is important that even pF19R does not confer on bacteria the capability to produce a considerable amount of extracellular antigen despite the greater intracellular synthesis as compared to pKM4-containing cells. The question arises: what gene is responsible for the F1 antigen production and secretion? Since the induction takes place only after the addition of IPTG we searched for

open reading frames downstream from promoter P_{lac} in the smallest plasmid p1810. We found that the sole large open reading frame in this direction is that of the *caf1M* gene. From these data we can deduce the stimulatory role of the *caf1M* gene product for the F1 secretion. The Caf1M protein seems to interact with the subunit polypeptide and to prevent it from digestion by a protease, as in the case of K88 biogenesis [10]. It also followed from these experiments that the *caf1A* gene product is not necessary for the secretion.

As shown in Table 1, the deletion of the *caf1A* gene results in the incapability of bacteria to agglutinate with agglutinating anti-plague antiserum. It implies that bacteria do not form capsules. Thus, the Caf1A protein must be involved in the capsule assembly. It is likely that the Caf1A molecule located in the outer membrane, binds F1 antigen subunits during the extracellular secretion process. The data obtained suggest that the mechanisms of *E. coli* fimbria biogenesis and of *Y. pestis* capsular biogenesis are very similar. The proteins encoded by the *f1* operon have similar counterparts among the products of genes of fimbria operons (F1 subunit — fimbria subunit; Caf1M chaperone-like protein — PapD chaperone; Caf1A protein — FaeD assembly protein). This is a striking feature as there is no similarity between the filamentous shape of fimbria and a structure of capsule evenly coating the bacterium. It should be noted that, despite the similarity between the fimbria and capsular genes, there exists a clear-cut distinction in the organization of operons. For example, in *K99* and $f7_2$ fimbria operons, the genes of subunit and assembly proteins are placed in the closest proximity to each other [11] as in the case of the *f1* operon, but in *K88* and *fim* [11] some other genes are placed between these genes. Such differences might reflect both the evolutionary relationships between these operons and the relative mobility of genes composing them which have led to their rearrangement.

Acknowledgements: We are grateful to G.A. Zav'yalova for help with computer estimations.

REFERENCES

[1] Anisimov, P.I., Serdobintsev, L.N., Taranenko, T.M. and Naumov, A.V. (1987) Mol. Genet. Virol. (USSR) 2, 24–27.
[2] Galyov, E.E., Smirnov, O.Yu., Karlyshev, A.V., Volkovoy, K.I., Denesyk, A.I., Nazimov, I.V., Rubtsov, K.S., Abramov, V.M., Dalvadyanz, S.M. and Zav'yalov, V.P. (1990) FEBS Lett. 277, 230–232.
[3] Galyov, E.E., Karlyshev, A.V., Chernovskaya, T.V., Dolgikh, D.A., Smirnov, O.Yu., Volkovoy, K.I., Abramov, V.M. and Zav'yalov, V.P. (1991) FEBS Lett. 286, 79–82.
[4] Maniatis, T., Fritsch, E.F. and Sambrook, J. (1982) Molecular Cloning: A Laboratory Manual, Cold Spring Harbor Laboratory, Cold Spring Harbor, NY.
[5] Mooi, F.R., Claassen, I., Bakker, D., Knipers, H. and De Graaf, F.K. (1986) Nucleic Acids Res. 14, 2443–2457.
[6] Norgren, M., Boga, M., Tennent, J.M. and Normark, S. (1987) Mol. Biol. Rep. 12, 169–178.
[7] Klemm, P. and Christiansen, G. (1990) Mol. Gen. Genet. 220, 334–338.
[8] Finkelstein, A.V. (1977) FEBS Lett. 82, 169–171.
[9] Von Heijne, G. and Abrahamsen, L. (1989) FEBS Lett. 244, 439–446.
[10] Bakker, D., Vader, C.E.M., Roosendaal, B., Mooi, F.R., Oudega, B. and De Graaf, F.K. (1991) Mol. Microbiol. 5, 875–886.
[11] Klemm, P. (1985) Reviews in Infection Diseases 7, 321–340.