A plasmid region encoding the active fragment and the inhibitor protein of colicin E3–CA38

Haruhiko Masaki and Takahisa Ohta

Department of Agricultural Chemistry, The University of Tokyo, Yayoi, Bunkyo-ku, Tokyo 113, Japan

Received 5 August 1982; revision received 5 October 1982

Colicin E3 *ColE3* *CloDF13* *Immunity* *Nucleotide sequence*

1. INTRODUCTION

Colicin E3 is a plasmid ColE3–CA38-directed antibacterial protein composed of two polypeptides, protein A and protein B of M_r 61000 and 10000, respectively [1,2]. After colicin E3 is adsorbed onto receptors in the outer membrane of sensitive cells, it finally inactivates ribosomes by cleaving the 16S RNA in the 30S subunit at a specific site [3,4]. This ribonuclease activity is exclusively attributed to the T2A domain, the C-terminal part of protein A. Tryptic digestion of intact colicin E3 (i.e., the AB complex) gives the T1 fragment and the T2 complex. This complex consists of the active fragment T2A and the inhibitor protein B [5,6]. Models similar to the above have been proposed for colicin E2 and cloacin DF13 [7,8]. The inhibitor protein acts specifically on each bacteriocin, and is produced in a greater quantity than bacteriocin, and in this way may protect the host cell from lethality. Consequently, the inhibitor protein is usually referred to as 'an immunity protein' or 'an immunity substance' [1,9–12]. Here, however, we simply refer to it as 'protein B'. The amino acid sequences of E3–T2A fragment [13], E3-protein B [14] and DF13-immunity protein [15] have been reported.

Once these bacteriocinogenic cells are treated by ultraviolet irradiation or with chemicals such as mitomycin C, 'SOS functions' are induced and the cells begin to produce a large number of bacteriocin molecules as well as free protein B molecules [9,16,17]. The production of protein B is thus inducible; nonetheless, immunity toward colicin E2 or E3 seems to function adequately in both the induced and non-induced states (unpublished). Our attention has been directed toward the mechanisms of expression and regulation of colicinogenicity and immunity. Although recent studies have shown some of the details involved in molecular structures of colicin E2 and E3, little is known regarding the genetic constructions of ColE2–P9 and ColE3–CA38 plasmids.

We now report the location and the nucleotide sequence of the DNA region encoding protein B and the T2A domain of protein A of colicin E3.

2. METHODS

All the plasmids were prepared by the method of [18]. ColE3–CA38 was isolated from *E. coli* RR1 (ColE3–CA38), which was obtained by transforming RR1 [19] with the plasmid fraction prepared from W3110Str^r (ColE3, ColI–CA38) [20].

The 2.8 kilobase *Pvu*II–*Hin*cII fragment derived from ColE3–CA38 was inserted between the *Pvu*II and *Hin*cII (in the *tet* gene) sites of pBR328 [19], giving rise to a recombinant plasmid pSH302. This plasmid confers both colicinogenicity and immunity on the host cell. That is, the intact colicin E3 gene must be carried on this *Pvu*II–*Hin*cII fragment (fig. 1a; details of derivation and characterization of pSH302 and the related plasmids are reported elsewhere).

The 5′-end-labelled DNA fragments were prepared and sequenced as in [21]. The fine restriction map was made according to [22].

Published by Elsevier Biomedical Press
001457593/82/0000—0000/$2.75

3. RESULTS AND DISCUSSION

Watson and Visentin reported the restriction maps of ColE2–P9 and ColE3–CA38, and showed that these plasmids have homologous restriction sites except for two non-homologous regions, *Bgl*I–*Pvu*I and *Pvu*II–*Sma*I [23]. The 2.8 kilobase *Pvu*II–*Hin*cII fragment of our plasmid pSH302 contains this *Pvu*II–*Sma*I segment and a short *Sma*I–*Hin*cII segment (fig. 1a). We found another restriction site *Cla*I at the center of the *Pvu*II–*Sma*I segment in both ColE2–P9 (not shown) and ColE3–CA38, so that the non-homologous region on *Pvu*II–*Sma*I was restricted to *Pvu*II–*Cla*I (fig. 1a). This finding is consistent with the result obtained from a heteroduplex experiment of ColE2–P9 and ColE3–CA38 [25]. We estimated the protein A-coding region of ColE3–CA38 to be 1.6–1.7 kilobase long (from M_r ~61 000) and the number of amino acids (~582) of protein A [2]. In serological and functional aspects, the protein A molecules of colicin E2 and colicin E3 were shown to be the same except in the C-terminal region [2,24]. Based on the above results, we assumed that the C-terminal part of ColE3–CA38 protein A might be encoded within the *Pvu*II–*Cla*I segment. Thus, we made a fine restriction map of this region (fig. 1a), and sequenced DNA fragments in order to identify the structural genes of protein A (gene A) and protein B (gene B) of colicin E3 (fig. 1b).

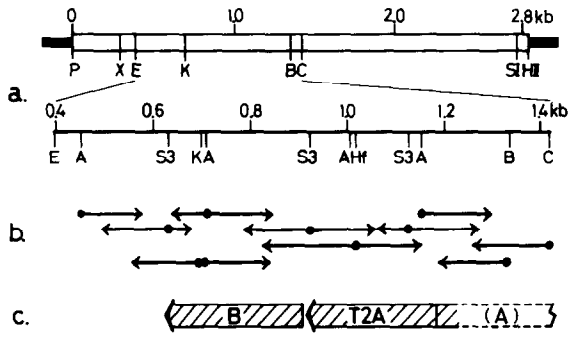

Fig. 1a) Restriction map of the 2.8 kilobase (kb) *Pvu*II–*Hin*cII fragment of pSH302. The DNA fragments derived from ColE3–CA38 and pBR328 are shown with an open box and solid lines, respectively. The restriction sites are indicated as follows: *Pvu*II(P), *Xho*I(X), *Eco*RI(E), *Kpn*I(K), *Bgl*I(B), *Cla*I(C), *Sma*I(SI), *Hin*cII(HII), *Alu*I(A), *Sau3A*I(S3), *Hin*fI(Hf). (b) The strategy used for sequencing. Arrows indicate the direction and extend of DNA sequences determined. (c) The arrangement of genes A and B identified in this work.

Fig. 2 shows the nucleotide sequence of 600 basepairs beginning at the position ~200 basepairs away from the *Cla*I site toward the *Pvu*II site. In the region from the position 34–321, we found an open reading frame compatible with the amino acid sequence of the T2A fragment determined in [13].

Moreover, another reading frame (334–588) was found to start at a point only 9 bases downstream from the termination codon of gene A. This frame was compatible with the amino acid sequence of protein B determined in [14] except for one amino acid residue (no. 45), so that this region was concluded to be gene B. The Asp at residue 45 proposed in [14] may have originated in *E. coli* cells from some post-translational modification of Asn-45 in our data.

Thus, genes A and B face the same direction and must be co-ordinately transcribed. We identified a candidate sequence for the promoter of gene A (and gene B) that possibly responds to 'SOS functions' near the *Sma*I site of pSH302 (experiments for verification of this now in progress). However, we also found possible promoter sequences at the center of the T2A-coding region; TTCTCA (198–203) and TATCGTG (223–229) may be 'the −35 region' and 'Pribnow box', respectively [26]. However, we have no evidence to show that they actually function.

The sequence GAGG (325–328) in the intervening space probably serves as a ribosome binding site [27]. This suggests that the translation of gene B does not depend on that of gene A. Alternatively, it is also possible that the translation continues into gene B immediately following the termination of gene A. These two possibilities seem to support the observation that protein B is produced in two different states in cells treated with mitomycin C: the AB complex and the free protein B [1,9,12].

The nucleotide sequence of CloDF13 immunity gene and the amino acid sequence of DF13-immunity protein was determined in [15]. The amino acid sequence of E3-protein B was also deduced from the partial sequence determination and the amino acid analysis of the tryptic peptides, and E3-

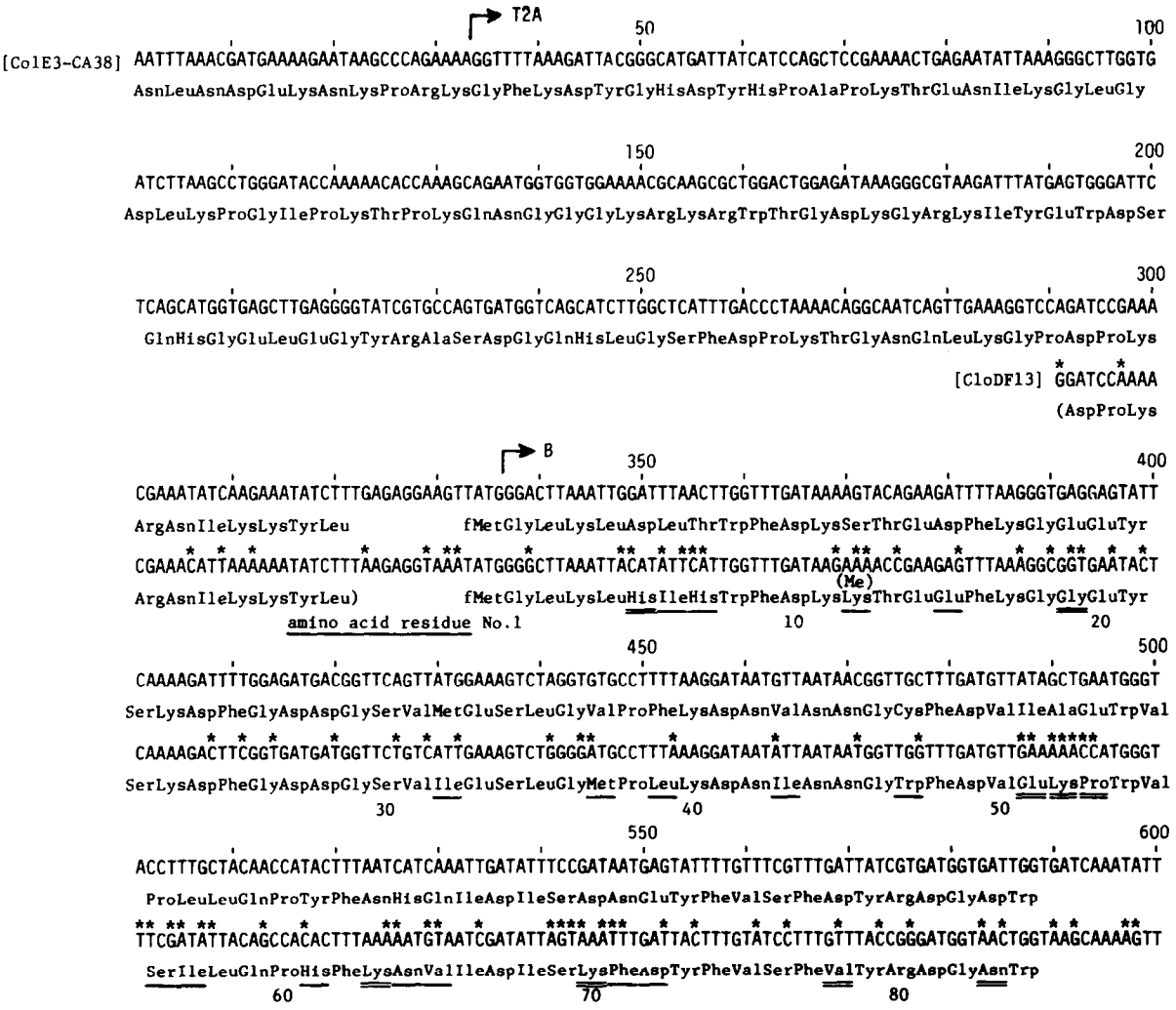

Fig. 2. Nucleotide and amino acid sequences of gene B and the latter part of gene A of ColE3–CA38, aligned with those of CloDF13 [15]. Replacements of nucleotides (*) and amino acids (== and ——) between the two plasmids are shown along the CloDF13 sequences; (==): replacements involving a change in charge; (——) the other replacements.

protein B compared with DF13-immunity protein [15]. Our sequence data on ColE3–CA38 gene B differ from those in [15] on E3-protein B in 24 out of 84 amino acid residues, and consequently, the gross profile of protein B assumed on the basis of our data is not the same as that proposed for E3-protein B in [15]:

(i) Residue 12 is not Lys but Ser in our data, and thus the methylation of this residue mentioned in [15] is not likely to occur in colicin E3.

(ii) According to [15], 18 differences were found between the amino acid sequences of E3-protein B and DF13-immunity protein, and these residues seem dispersed throughout all the molecules. Our data show that there are 25 differences, and that major replacements occur within the latter halves of the molecules.

In spite of the above replacements, there is considerable homology between colicin E3 and cloacin DF13 in both amino acid and nucleotide sequences. We found that this homologous region extends upstream from the termination of gene A through 30 nucleotides reported in [15]. Following termination of gene B, there are about 70 other

homologous nucleotides (not shown). The high homology in the amino acid sequence of gene B (71%) is comparable to that in the nucleotide sequence (73%) and implies that some requirement of amino acids might serve as a selective pressure in evolution. From this point of view, the codon-usage of Ser-69 is of interest; the TCC codon is used in ColE3–CA38 and the AGT in CloDF13.

Colicin E3 and cloacin DF13 are homologous bacteriocins with respect to both structure and mode of action, but protein B inhibits specifically the ribosome-inactivating activity of each protein A [17]. Here, we show that the homology between the two plasmids further extends to gene arrangement and nucleotide sequence. In addition, some differences between two protein B regions may possible provide an explanation for the highly specific inhibitor activities.

ACKNOWLEDGEMENTS

We thank Drs S. Ohno and T. Taniguchi for kindly instructing us in the use of the DNA sequencing method.

REFERENCES

[1] Jakes, K.S. and Zinder, N.D. (1974) Proc. Natl. Acad. Sci. USA 71, 3380–3384.
[2] Ohno-Iwashita, Y. and Imahori, K. (1980) Biochemistry 19, 652–659.
[3] Senior, B.W. and Holland, I.B. (1971) Proc. Natl. Acad. Sci. USA 68, 959–963.
[4] Bowman, C.M., Dahlberg, J.E., Ikemura, T., Konisky, J. and Nomura, M. (1971) Proc. Natl. Acad. Sci. USA 68, 964–968.
[5] Ohno, S., Ohno-Iwashita, Y., Suzuki, K. and Imahori, K. (1977) J. Biochem. 82, 1045–1053.
[6] Suzuki, K. and Imahori, K. (1978) J. Biochem. 84, 1021–1029.
[7] Yamamoto, H., Nishida, K., Beppu, T. and Arima, K. (1978) J. Biochem. 83, 827–834.
[8] Mooi, F.R. and De Graaf, F.K. (1976) FEBS Lett. 62, 304–308.
[9] Jakes, K., Zinder, N.D. and Boon, T. (1974) J. Biol. Chem. 249, 438–444.
[10] Schaller, K. and Nomura, M. (1976) Proc. Natl. Acad. Sci. USA 73, 3989–3993.
[11] De Graaf, F.K. and Klaasen-Boor, P. (1974) FEBS Lett. 40, 293–296.
[12] Sidikaro, J. and Nomura, M. (1974) J. Biol. Chem. 249, 445–453.
[13] Suzuki, K. and Imahori, K. (1978) J. Biochem. 84, 1031–1039.
[14] Mochitate, K., Suzuki, K. and Imahori, K. (1981) J. Biochem. 89, 1609–1618.
[15] Van den Elzen, P.J.M., Gaastra, W., Spelt, C.E., De Graaf, F.K., Veltkamp, E. and Nijkamp, H.J.J. (1980) Nucleic Acids Res. 8, 4349–4363.
[16] Tyler, J. and Sherratt, D.J. (1975) Mol. Gen. Genet. 140, 349–353.
[17] Kool, A.J., Pols, C. and Nijkamp, H.J.J. (1975) Antimicrob. Agents Chemother. 8, 67–75.
[18] Colman, A., Byers, M.J., Primrose, S.B. and Lyons, A. (1978) Eur. J. Biochem. 91, 303–310.
[19] Soberon, X., Covarrubias, L. and Bolivar, F. (1980) Gene 9, 287–305.
[20] Herschman, H.R. and Helinski, D.R. (1967) J. Biol. Chem. 242, 5360–5368.
[21] Maxam, A.M. and Gilbert, W. (1980) Methods Enzymol. 65, 499–560.
[22] Smith, H.O. and Birnstiel, M.L. (1976) Nucleic Acids Res. 3, 2387–2398.
[23] Watson, R. and Visentin, L.P. (1980) Gene 10, 307–318.
[24] Ohno-Iwashita, Y. and Imahori, K. (1979) FEBS Lett. 100, 249–252.
[25] Inselberg, J. and Johns, V. (1975) J. Bacteriol. 121, 381–389.
[26] Rosenberg, M. and Court, D. (1979) Annu. Rev. Genet. 13, 319–354.
[27] Steitz, J.A. and Jakes, K. (1975) Proc. Natl. Acad. Sci. USA 72, 4734–4738.