Vol. 160, No. 1, 1989 BIOCHEMICAL AND BIOPHYSICAL RESEARCH COMMUNICATIONS
April 14, 1989 Pages 33-39

CLONING AND SEQUENCING OF A PHOSPHOLIPASE C GENE OF CLOSTRIDIUM PERFRINGENS

Akinobu Okabe, Tohru Shimizu, and Hideo Hayashi

Department of Microbiology, Kagawa Medical School, Kagawa 761 – 07, Japan

Received February 23, 1989

Summary: The gene encoding phospholipase C (α-toxin) of Clostridium perfringens was cloned into λ gt10. The maximal size of the coding region was 1.4 kb and the minimum was 1.1 kb as determined by subcloning into the vector pBR322 and testing for activity. The nucleotide sequence of this region contained a single open reading frame of 1194 bp corresponding to a protein of Mr 45473 with a possible N-terminal signal sequence of 28 amino acids which when removed, would give a mature protein of Mr 42521. This is in good agreement with the reported size of 43 kDa. The coding region has a dG + dC content of 33.7%, and the codon usage displays a pronounced preference for codons with the lowest dG + dC content.

The genus Clostridium comprises many species of obligatory anaerobic bacteria. The most striking feature of the Clostridial chromosome is an extremely high dA + dT content (1). Recently the complete nucleotide sequence of a 10 kb plasmid derived from Clostridium perfringens has been determined (2). The dA + dT content of this plasmid is 75%, which is very close to that of the host chromosome. Unusual codon usage has been suggested to be a cause for the difficulty in expression of Clostridial genes in Escherichia coli. Many efforts have been made to develop an E. coli shuttle vector (3) and a transformation procedure for Clostridial species (4). However, neither goal has been reached and little progress has been made in cloning Clostridial genes. One way to overcome these problems is to detect cloned genes with a DNA probe and λ gt10 is suitable for this purpose (5).

Our intial studies focused on constructing a gene library of C. perfringens in λ gt10. The first problem we encountered was isolating undamaged chromosome DNA because C. perfringens produces a highly active DNase (6). We developed a simple and rapid method for purification of chromosome DNA which is suitable for gene

0006-291X/89 $1.50

cloning (7). Using chromosome DNA prepared by this method we have succeeded in constructing a gene library in λ gt10 and cloning two hemolysins, α-toxin (phospholipase C, EC 3.1.4.3.) and θ-toxin (thiol-activated hemolysin). This paper describes the cloning and sequencing of the phospholipase C gene (plc). The aim of this work is to understand the genetic organization of the plc gene as an example of a gene in organisms with extremely high dA + dT content.

Materials and Methods

Materials: C. perfringens type A NCTC8237 (7), E. coli substrains HB101 (8), JM109 (9), and C600 hflA (10) were used as bacterial strains. Enzymes and chemicals were purchased from the following sources: restriction enzymes and T4 DNA ligase, 7-deaza-sequencing kit (klenow), 7-deaza sequencing kit (sequenase), reverse primer, dephosphorylated λ gt10 DNA arms and DNA in vitro packaging kit from Toyobo Co., Ltd. (Osaka, Japan); klenow fragment of DNA polymerase I, kilo-sequence deletion kit from Takara Shuzo Co., Ltd. (Kyoto, Japan); geneclean kit from Bio101 Inc. (La Jolla CA, U.S.A.); [α-^{35}S] dCTP (1000-1500 Ci/mmol) from Amersham International plc.; and gel bond from FMC Bioproducts (Rockland, ME, U.S.A.).

Construction of the DNA library: Whole cell DNA of C. perfringens was purified as described elsewhere (7). DNA was partially digested with EcoRI and 3-7 kb fragments were fractionated by sucrose gradient centrifugation (11). The fragments were ligated with λ gt10 and the ligation products were packaged in vitro to generate intact phage particles.

Cloning: Recombinant phages expressing hemolysis (Hly$^+$) were detected by plating the phage on E. coli C600 hflA at a density of Ca. 300 plaques per plate and the plaques were overlaid with 0.8% agarose in PBS containing 1 mM dithiothreitol and 7.5% packed washed sheep erythrocytes. After incubating the plates at 37℃ for 5-7 h, zones of hemolysis were detected.

Nucleotide sequencing: The Hly$^+$ recombinant phage were isolated and the 4.4 kb insert they contained was subcloned into plasmid pBR322. Hly$^+$ subclones were detected by overlaying the colonies with blood containing agarose as described above. The smallest insert (1.4 kb) which gave a Hly$^+$ phenotype was ligated into plasmid pUC19. Clones to be sequenced were generated by exonuclease Ⅲ deletion. DNA sequencing was carried out by the dideoxy chain termination method (12) using 2'-7-deazaguanosine triphosphate instead of dGTP for better resolution (13).

SDS-polyacrylamide gel electrophoresis: E. coli cells were suspended in 1/10 of the original culture volume of 50 mM Tris-HCl (pH 7.0) and sonicated as described previously (14). The supernatants after centrifugation at 15,000 xg for 20 min at 4℃ were loaded on and electrophoresed through SDS-polyacrylamide gels as described previously (15). Prior to loading, each sample was diluted in an equal volume of 2x sample buffer lacking β-mecaptoethanol and incubated at 37℃ for 30 min. Approximately 260 μg of protein was loaded into the sample well. After electrophoresis at 120 V for 6 h, the gel was rinsed by the method of Focareta and Manning (16) to remove SDS from the gel. The gel was placed on an egg yolk agar sheet, which was prepared by pouring CW agar (Nissui SK, Tokyo, Japan) containing 5% egg yolk and 200 μg/ml of kanamycin on a sheet of gel bond. Phospholipase C was detected as a turbid spot on the egg yolk agar sheet after incubation for 4 h at 37℃.

Other procedures: Lecithinase assay on agar plates were carried out as follows. Clear phage lysates (15 μl) were added to a well on an egg yolk agar plate prepared as described above. The plate was incubated at 37℃ overnight. The well was surrounded by a wide (Ca 1 cm) turbid zone when a sample contained lecithinase. Phospholipase C was purified by the method of Yamakawa, et al. (17). The sequence of the N terminal amino acid was determined by a gas phase amino acid sequencer Model 470A (Applied Biosystems, Foster, CA, U.S.A.). Computer analysis was performed with the DNASIS program (Hitachi SK, Yokohama, Japan).

Results and Discussion

Molecular cloning: We screened 1260 plaques from the λ gt10 gene library for the Hly$^+$ phenotype and isolated 13 positive clones. Hemolysis by these clones became marked after transferring a plate from 37 to 4℃, which is characteristic of hemolysis by α-toxin (hot-cold hemolysis). Lysates of the positive phage were tested for phospholipase C activity by lecithinase assay on egg yolk agar. They all showed a positive reaction and the addition of 10 mM EDTA completely inhibited the reaction, indicating the hemolytic activity associated with the Hly$^+$ phage is due to the expression of the plc gene. The incidence of plc clones far exceeded that calculated by the formula of Clarke and Carbon (18). This may suggest that there are a large number of copies of the plc gene in a chromosome. However, the possibility that it is due to random variation in the cloning cannot be ruled out. The DNA of three Hly$^+$ recombinant phages were isolated and the cleavage sites for EcoRI and Hind Ⅲ were determined. The results showed that all three clones contained a 4.4 kb EcoRI fragment, which all have two Hind Ⅲ sites in the same positions. One of these inserts was ligated to the EcoRI site of pBR322. The resulting plasmid, named pKM1 was stably maintained in E. coli HB101. Cell lysates of E. coli HB101/pKM1 were electrophoresed on SDS-polyacrylamide gel and phospholipase C was detected by incubating on egg yolk agar. The cloned activity had the same mobility as purified C. perfringens phospholipase C (Fig. 1). No band was observed in a sample prepared from E. coli HB101 lacking the plasmid. The apparent Mr of the corresponding band is 42 kDa, being slightly less than the published value, 43 kDa. This difference is clearly due to the absence of β-mercaptoethanol in the sample. In fact, the Mr of reduced toxin was estimated as 43 kDa (Fig. 1). These results confirm that pKM1 expressed the plc gene.

To identify and map the cistron involved in plc expression, pKM1 was digested with restriction endonucleases. The liberated fragents were ligated to pBR322 and tested for expression of the Hly$^+$ phenotype. The plasmid containing a DraI/Hind Ⅲ fragment (1.1 kb) did not show the phenotype nor did the NdeI/Hind Ⅲ fragment (1.0 kb). On the other hand, the SspI/Hind Ⅲ fragment (1.4 kb), ligated into the SspI/Hind Ⅲ sites of pBR322 gave Hly$^+$ transformants. This plasmid, named pKM3 also expressed the plc gene (Fig.1). Therefore, 1.4 kb appeared to be the minimum size coding for

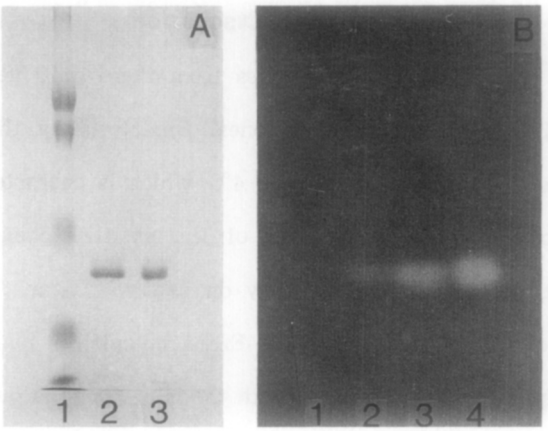

Fig. 1. Analysis of phospholipase C produced by the plc clones. (Panel A) SDS-polyacrylamide gel electrophoresis of the purified C. perfringens phospholipase C. Lane 1, the prestained Mr standards (Bio-Rad, Richmond, CA, U.S.A.) : phosphorylase b (130 kDa) ; bovine serum albumin (75 kDa) ; ovalbumin (50 kDa) ; carbonic anhydrase (39 kDa) ; soybean trypsin inhibitor (27 kDa). Lane 2, the purified toxin (8.8 μ g) boiled in sample buffer. Lane 3, the purified toxin incubated for 30 min at 37℃ in sample buffer lacking β-mercaptoethanol. After electrophoresis on 10% polyacrylamide gel (2 mm thick), protein was stained with coomassie brilliant blue R. (Panel B) Duplicate of the gel shown in panel A, which, after electrophoresis and washing to remove SDS, was overlaid onto egg yolk agar. Lane 1, HB101. Lane 2, HB101/pKM1. Lane 3, HB101/pKM3. Lane 4, the purified toxin (4.4 μ g).

phospholipase C. Cells harboring pKM3 grew on a tectracycline plate but not on an ampicillin plate, indicating that the plc promoter is proximal to the SspI end and can read through the plc gene to give tetracycline resistance.

Nucleotide sequencing : The cloned 1.4 kb DNA was ligated into the SmaI site of plasmid pUC19 and the nucleotide sequence was determined (Fig. 2). Within this sequence is a single ORF of 1194 bp. As would be expected for an exported protein, there is a typical signal sequence of 28 amino acids at the N terminus. Cleavage of the signal sequence would give a mature protein of Mr 42521, which is in excellent agreement with the size of protein (43 kDa) obtained by SDS – polyacrylamide gel electrophoresis. Determination of the N terminal amino acid sequence of the purified toxin confirms the presence and length of the signal peptide (Fig. 2). The -10 and -35 regions of this promoter show good homology with the consensus sequence for gram positive promoters and their spacing is that of a functional promoter (19). There is a palindrome structure in the -10 region, suggesting expression of the plc gene may be negatively controlled by a repressor. A potential ribosome binding site is present 4 bp upstream of the first codon as is common for the translation initiation signals

```
      ATTCAAAAATTTAGTGAGCTTATGGTAATTATATGGTATAATTTCAGTGCAAGTGTTAATCGTTATCAAAAAAG

 75   GGGAGATTAATACTTGAAAAAAATTAACGGGGGATATAAAAAATGAAAAGAAAGATTTGTAAGGCGCTTATTTGT
                                                  M  K  R  K  I  C  K  A  L  I  C   11

150   GCCGCGCTAGCAACTAGCCTATGGGCTGGGGCATCAACTAAAGTCTACGCTTGGGATGGAAAGATTGATGGAACA
       A  A  L  A  T  S  L  W  A  G  A  S  T  K  V  Y  A  W  D  G  K  I  D  G  T   36

225   GGAACTCATGCTATGATTGTAACTCAAGGGGTTTCAATCTTAGAAAATGATCTGTCCAAAAATGAACCAGAAAGT
       G  T  H  A  M  I  V  T  Q  G  V  S  I  L  E  N  D  L  S  K  N  E  P  E  S   61

300   GTAAGAAAAAACTTAGAGATTTTAAAAGAGAACATGCATGAGCTTCAATTAGGTTCTACTTATCCAGATTATGAT
       V  R  K  N  L  E  I  L  K  E  N  M  H  E  L  Q  L  G  S  T  Y  P  D  Y  D   86

375   AAGAATGCATATGATCTATATCAAGATCATTTCTGGGATCCTGATACAGATAATAATTTCTCAAAGGATAATAGT
       K  N  A  Y  D  L  Y  Q  D  H  F  W  D  P  D  T  D  N  N  F  S  K  D  N  S  111

450   TGGTATTTAGCTTATTCTATACCTGACACAGGGGAATCACAAATAAGAAAATTTTCAGCATTAGCTAGATATGAA
       W  Y  L  A  Y  S  I  P  D  T  G  E  S  Q  I  R  K  F  S  A  L  A  R  Y  E  136

525   TGGCAAAGAGGAAACTATAAACAAGCTACATTCTATCTTGGAGAGGCTATGCACTATTTTGGAGATATAGATACT
       W  Q  R  G  N  Y  K  Q  A  T  F  Y  L  G  E  A  M  H  Y  F  G  D  I  D  T  161

600   CCATATCATCCTGCTAATGTTACTGCCGTTGATAGCGCAGGACATGTTAAGTTTGAGACTTTTGCAGAGGAAAGA
       P  Y  H  P  A  N  V  T  A  V  D  S  A  G  H  V  K  F  E  T  F  A  E  E  R  186

675   AAAGAACAGTATAAAATAAACACAGCAGGTTGCAAAACTAATGAGGCTTTTTATACTGATATCTTAAAAAACAAA
       K  E  Q  Y  K  I  N  T  A  G  C  K  T  N  E  A  F  Y  T  D  I  L  K  N  K  211

750   GATTTTAATGCATGGTCAAAAGAATATGCAAGAGGTTTTGCTAAAACAGGAAAATCAATATACTATAGTCATGCT
       D  F  N  A  W  S  K  E  Y  A  R  G  F  A  K  T  G  K  S  I  Y  Y  S  H  A  236

825   AGCATGAGTCATAGTTGGGATGATTGGGATTATGCAGCAAAGGTAACTTTAGCTAACTCTCAAAAAGGAACAGCG
       S  M  S  H  S  W  D  D  W  D  Y  A  A  K  V  T  L  A  N  S  Q  K  G  T  A  261

900   GGATATATTTATAGATTCTTACACGATGTATCAGAGGGTAATGATCCATCAGTTGGAAAGAATGTAAAAGAACTA
       G  Y  I  Y  R  F  L  H  D  V  S  E  G  N  D  P  S  V  G  K  N  V  K  E  L  286

975   GTAGCTTACATATCAACTAGTGGTGAGAAAGATGCTGGAACAGATGACTACATGTATTTTGGAATCAAAACAAAG
       V  A  Y  I  S  T  S  G  E  K  D  A  G  T  D  D  Y  M  Y  F  G  I  K  T  K  311

1050  GATGGAAAAACTCAAGAATGGGAAATGGACAACCCAGGAAATGATTTTATGACTGGAAGTAAAGACACTTATACT
       D  G  K  T  Q  E  W  E  M  D  N  P  G  N  D  F  M  T  G  S  K  D  T  Y  T  336

1125  TTCAAATTAAAAGATGAAAATCTAAAAATTGATGATATACAAAATATGTGGATTAGAAAAAGAAAATATACAGCA
       F  K  L  K  D  E  N  L  K  I  D  D  I  Q  N  M  W  I  R  K  R  K  Y  T  A  361

1200  TTCTCAGATGCTTATAAGCCAGAAAACATAAAGATAATAGCAAATGGAAAAGTTGTAGTGGACAAAGATATAAAC
       F  S  D  A  Y  K  P  E  N  I  K  I  I  A  N  G  K  V  V  V  D  K  D  I  N  386

1275  GAGTGGATTTCAGGAAATTCAACTTATAATATAAAATAATAAAAGTAAAAAAATAATTATTGGTTTTGGTGGTAT
       E  W  I  S  G  N  S  T  Y  N  I  K  *  *                                  398

1350  TTACAAAATAAAAGCT
```

Fig. 2. Nucleotide sequence of the plc gene and its deduced amino acid sequence. The nucleotide sequence is numbered at the left of the diagram, and the locations of possible promoter and ribosome-binding sites are denoted by lines above the sequence. The predicted amino acid sequence is numbered on the right of the diagram. The N terminal amino acid residues determined by sequencing the N-terminus of purified C. perfringens are underlined.

of genes from gram positive bacteria (20, 21). There is no typical hairpin structure like that present in most transcription terminators within the sequences determined. This explains expression of the tetracycline resistance as described above.

Codon usage: The dA + dT content of the ORF is 66% which is slightly lower than that of the entire chromosome of C. perfringens (73-76%) (1). There is a pronounced preference for codons containing A and T. This is particularly apparent in the wobble position of the codons as 78% of the codons in the plc gene end in dA or dU. A choice between functionally related amino acids has also been shown to reflect the dA + dT content of the plasmid. This is also the case for the plc gene, for example, 40 Lys vs 10 Arg and 26 Asn vs 10 Gln. The presence of a functional promoter and a translational initiation signal support the prediction of Ganier and Cole (2) that unusual codon usage might limit the expression of extremely dA + dT rich genes in E. coli. However, this cannot be concluded definitely until the production of phospholipase C is quantitatively determined and compared to the level of transcription of the plc gene in E. coli.

Acknowledgments

We gratefully acknowledge the assistance of Miss Chiharu Shimizu in N terminal amino acid sequencing. We are also indebted to Professor David B. Wilson, Section of Biochemistry, Molecular and Cell Biology, Cornell University, Ithaca, New York, for invaluable assistance for preparing the manuscript.

References

1. Sneath, P.H.A. (1986) In Bergy's Manual Of Systematic Bacteriology (P.H.A. Sneath, N.S. Mair, M.E. Sharpe, and J.G. Holt, eds.), Vol. II, pp.1104 - 1207. Wiliams & Wilkins, Baltimore.
2. Ganier, T., and Cole, S.T. (1988) Plasmid 19, 134-150.
3. Squires, C.H.D., Heefner, D.L., Evans, R.J., Kopp, B.J., and Yarus, M.J. (1984) J. Bacteriol. 159, 465-471.
4. Mahony, D.E., Mader, J.A., and Dubel, J.R. (1988) Appl. Environ. Microbiol. 54, 264 - 267.
5. Huynh, T.V., Young, R.A., and Davis, R.W. (1985) In DNA Cloning (D.M. Glover, ed.), Vol. I, pp.49-78. IRL Press, Oxford.
6. Doll, W. (1973) Zbl, Bakt. Hyg., I. Abt. A224, 115-119.
7. Okabe, A., Yoshihara, K., Shimizu, T., and Hayashi, H. (1989) submitted to J. Bacteriol.
8. Maniatis, T., Fritsch, E.F., and Sambrook, J. (1982) In Molecular Cloning : A Laboratory Manual, Cold Spring Harbor Laboratory, Cold Spring Harbor.
9. Yanisch-Perron, C., Vieira, J., and Messing, J. (1985) Gene 33, 103-119.
10. Young, R.A., and Davis, R.W. (1983) Proc. Natl. Acad. Sci. USA 80, 1194-1198.
11. Kaiser, K., and Murray, N.E. (1985) In DNA Cloning (D.M. Glover, ed.), Vol. I, pp. 1-47. IRL Press, Oxford.

12. Sanger, F., Nicklen, S., and Coulson, A.R. (1977) Proc. Natl. Acad. Sci. USA 74, 5463-5467.
13. Mizusawa, S., Nishimura, S., and Seela, F. (1986) Nucleic Acids Res. 14, 1319-1324.
14. Okabe, A., Matsushita, O., and Hayashi, H. (1988) J. Antibio. 41, 667-674.
15. Okabe, A,, Matsushita, O., Katayama, S., and Hayashi, H. (1986) Antimicrob. Agents Chemother. 30, 82-87.
16. Focareta, T., and Manning, P.A. (1987) Gene 53, 31-40.
17. Yamakawa, Y., and Ohsaka, A. (1977) J. Biochem. 81, 115-126.
18. Clarke, L., and Carbon, J. (1976) Cell 9, 91-99.
19. Graves, M.C., and Rabinowitz, J.C. (1986) J. Biol. Chem. 261, 11409-11415.
20. McLaughlin, J.R., Murray, C.L., and Rabinowitz, J.C. (1981) J. Biol. Chem. 256, 11283-11291.
21. Moran, C.P.,Jr., Lang, N., Le Grice, S.F.J., Lee, G., Stephens, M., Sonenshein, A.L., Pero, J., and Losick, R. (1982) Mol. Gen. Genet. 186, 339-346.