Short Communication

HEAD FIBER PROTEIN GENE OF BACTERIOPHAGE Nf

SEUNG-KU YOO, DAN K. BRAITHWAITE, AND JUNETSU ITO*

*Department of Microbiology and Immunology, College of Medicine,
University of Arizona Health Sciences Center,
Tucson, Arizona 85724, U.S.A.*

(Received August 8, 1990)

Small *Bacillus* phages, such as $\phi 29$, $\phi 15$, Nf, M2, and GA-1, have been useful as model systems for the study of linear duplex DNA replication. They have been classified, based on immunological properties, genome size, terminal DNA sequences and electrophoretic patterns of their structural proteins, into the following three groups: group A ($\phi 29$, PZA, and $\phi 15$), group B (M2 and Nf) and group C (GA-1) (13, 14). All of these phages are morphologically similar and contain linear, non-redundant duplex DNA 16.7 to 19.7 Kb in length. The unique features of these small phage genomes are that they contain short inverted terminal repeat sequences and that the 5' ends of both strands are covalently linked to a terminal protein (1, 9).

M2 and Nf are the smallest among these phages and are very closely related (3, 6, 11, 13). The M2 genome is slightly smaller than the Nf genome and lacks a head fiber protein (6, 13). Restriction enzyme and heteroduplex analyses have revealed that a small different region is located approximately 40% of the way from the left end of the genome (6, 13). Compared to $\phi 29$, this region would correspond to gene 8.5, the structural gene of the head fiber protein of $\phi 29$ (8). Therefore, we suggested that M2 lacks the intact structural gene for the head fiber protein (13).

Recently, Mizukami et al. (7) determined the nucleotide sequences of this region of the M2 and Nf genomes. Their results showed that M2 is, indeed, a small deletion mutant of Nf. They also reported that a 14.7 Kdalton protein is encoded in this region. This protein's size was too small, however, to be Nf's head fiber protein which has been estimated to be 24 to 27 Kdaltons from the SDS-polyacrylamide gel electrophoresis (6, 13). Accordingly, Mizukami et al. postulated that the 14.7 K dalton protein is involved in the assembly of the head fiber protein. They suggested that the mature M2 phage particle lacks a head fiber protein because of the deletion

* Address reprint requests to: Dr. J. Ito, Department of Microbiology and Immunology, College of Medicine, University of Arizona Health Science Center, Tucson, Arizona 85724, U.S.A.

of the 14.7 Kdalton protein gene even though it has the intact structural gene for the head fiber protein. In support of this hypothesis, minicells infected with either M2 or Nf were stated to have produced a protein similar in size to the Nf head fiber protein (cited in ref. 7).

On the other hand, no such "head fiber assembly" gene has been found, even though the entire genomes of $\phi 29$ and PZA have been sequenced (12). Thus, if such a gene exists, it would represent a characteristic of phages belonging to group B and would provide a unique system to study phage morphogenesis. To confirm that Nf contains such a gene, we sequenced the corresponding region of the Nf genome and found that this region instead contains the structural gene for the head fiber protein of Nf. Our analyses further suggested that the structural gene for the Nf head fiber protein is flanked by gene 8, the major head protein, and gene 9, the tail protein. Thus, it appears that there is not a gene involved in the assembly of the head fiber protein and that the organization of genes for the head protein, the head fiber protein and the tail protein are exactly the same as $\phi 29$'s and PZA's.

Preparation of Nf phage and DNA were carried out as described previously (4, 14). DNA was sequenced by the chain termination method (10) and by the DNA polymerase chain reaction method (2). DNA and protein sequences were analyzed with a Vax computer using the Genetics Computer Group programs of the University of Wisconsin.

```

1  GATTCATCTTCTGGAGGTGCGCAAGCGTAATGTATTCACTGACATATGCAAACCTCG
  D S S S G G A Q A * M Y S F T A V A N S D
61  ACATTGTCGCCTTTCATTATTGAAATTTCTATCGAGCGAAAACAACATAGAAATTTCTT
  I V A F H L L K F L S S E N N I E I S Y
120 ATGCAGATGAAGACACAATACCAGAATACGTTTCAATTTCGTGACCTAAAAGCTGGAGATA
  A D E D T I P E Y V S I R D L K A G D K
181 AAACAACAAATTGATTTATACCCATTAGTTGCGTGGAAAGTGATTGCTCAGGAGGCATAA
  T T I D L Y P L V A W K V I A Q E D I T
241 CAACTGGGGACAGAGTATCCGTTGGAAAGAATGGGCAAGTAAAAAGACAACTGACTTGC
  T G D R V S V G K N G Q V K K T T D L R
301 GCACAGCATTGCGGTACGCTGTTAGCCTTGCAAAAGCTGGACAACCTTGTACAGTCGCTA
  T A F G Y A V S L A K A G Q L V T V A I
361 TTTCAACGGTATTTGACACGATTATCACGCCAGATGATTTGGGAATGTTGACGATGATG
  S T V F D T I I T P D D L G N V D D D V
421 TAAAAGCATCTTGAAGTCAATGCAACAGACGCAATAAAGCAAAGTTGCGTGATCTAC
  K A F L K S N A T D A N K A K L R D L L
481 TCGTATCCAATCTTGATGTGAAAGCTTTTCTCAATGGTTCAACAAGTGAAGACAATAAGA
  V S N L D V K A F L N G S T S E D N K I
541 TAAATCTTAGAAATTTACTTGTAGCAACCTGCAATCCTACGCTTTCTTAATGCAAACC
  N L R N L L V S N P A I L A F L N A N P
601 CTGACACCGATACTCAGACGACATTAAGAACAATGATTGGCGCTGGTACACCATATACAC
  D T D T Q T T L R T M I G A G T P Y T L
661 TACCAGCCGCCACTACCAACAACACTAGGTGGTGTGAAACGAATACCAGCATTTGGAAACT
  P A A T T T T L G G V K R I P A F G N S
721 CAACTGCTACTGATGTTGCAACACTGGTAAAAGATTCAACAACCTACTTGGCGGCGATGA
  T A T D V A T L V K D F N N L A A M R
781 GAACAGCTGGTTACATTTATAGAGGTGAAATTTTCATGGCATATGTCCTTTATCTGGTT
  T A G Y I L * * M A Y V P L S G S
841 CCGACGTTTCGATTTTTCGGCGTTGCGTTCTCCAATGATTATAAAATACTAAGGTGGT
  D V R I F S A L R S P M I I K I L R W F
901 TCACAAGTCCGACGCTCAATATTCGTATTTTAACGCTAAACCACTGTGCGATGTGATCA
  T S A D A Q Y S Y F N A K P R V H V I .

```

Fig. 1. Complete nucleotide sequence of *Hinfl-BclI* fragment of the Nf genome. The phage DNA and restriction fragments were prepared according to Kawamura and Ito (5).

gene 8.5 homology					
ϕ 29	2	MVSFTARAKSNVMAYRLLAYSQGDDII	IEISHAAENTIPDYVAVKDV	DKGDLTQVNM	YPLA
NF	1	MYSFTAYANSDIVAFHLLKFLSSENNIE	ISYADEDTIPEYVSIRDLKAGDKTTID	LYPLV	
PZA	2	MVSFTARANSNILAYRLLAYSEGDDII	IEISHAAENTIPDYVAVKDV	DKGDLTQVNM	YPLA
ϕ 29	62	AWQVIAGSDIKVGDNLTTGKDGTAVPT	DDPSTVFGYAVEEAQEGQLVTLVISR	SKEISIE	
NF	61	ANKVIAQEDITTGDRVSVGKNGQVKKT	DLRTAFGYAVSLAKAGQLVTVAISTV	FDTIIT	
PZA	62	AWQVIAGSDIKVGDNLTTGKDGTAVPT	DDPSTVFGYAVEEAQEGQLVTLVISR	SKEISIE	
ϕ 29	122	VDDIKDAGDTGKRLKINTPSGARNII	ENEDAKALINGETTNTNKKNLQDLL	FS	DGNVK
NF	121	PDDLGNVDD.....	DVKAFKLSNATDANKAKLRDLLV	SNLDVK	
PZA	122	VDDIKDAGDTGKRLKINTPSGARNII	ENEDAKALINGETTNTNKKNLQDLL	FS	DGNVK
ϕ 29	182	AFLQATTDTDENKTALQQLLVSNADV	LGLLSGNPTSDNKINLRTMIGAGVP	YSLPAATTTT	
NF	159	AFLNGSTSEDNKNINLRNLLVSNPAIL	AFNLNANPDTDTQTTLRTMIGAGTP	YTLPAATTTT	
PZA	182	AFLQANTDTDENKAALQQLMVSNA	DVLGLLSGAPSSSENKVNIRSMIGAG	VPYTLPAATTTT	
ϕ 29	242	IGGVKKGAAVTASTATDVATAVKDL	NSLITVLKNAGIIIS		
NF	219	LGGVKRIPAFGNSTATDVATLVKDF	NNLLAAMRTAGYIL		
PZA	242	IGGVKKGAAVSASTATDVTA	VKDLNSLITVLKNAGIIIS		
gene 9 homology					
ϕ 29	1	MAYVPLSGTNRVILADVPF	SNDYKNTRWFTSSSNQYNWFNRK	SRVYEM	
NF	1	MAYVPLSGSDVRIFSA	LRSPMIKILRWFTSADAQYSYFNAK	PRVHVI	
PZA	1	MAYVPLSGTNRVILADVPF	SNDYKNTRWFTSSSNQYNWFNSK	TRVYEM	

Fig. 2. Alignment of ϕ 29, Nf, and PZA's amino acid sequences of genes 8.5 and 9.

The region of the Nf genome sequenced (960 bp) is presented in Fig. 1. Our DNA sequence is very similar to that of Mizukami's, except we show only one A at position 401 and one T at position 799 instead of two As and two Ts (positions marked by arrows). The deduced amino acid sequences are also shown in Fig. 1. An open reading frame of 771 nucleotides which initiates with ATG at position 30 is evident (Fig. 1). This would encode for a protein of about 27,000 daltons which is very close to the molecular weight of Nf's head fiber protein, estimated from SDS-polyacrylamide gel migration analysis (13). When the deduced amino acid sequence was compared with those of ϕ 29 and PZA's head fiber proteins, highly conserved amino acid sequences were revealed (Fig. 2). These results strongly suggest that the deduced protein is indeed the head fiber protein of Nf. Figure 2 also shows that the first 39 amino acids of the next open reading frame are homologous to those of gene 9, the structural gene for the tail proteins, in the genomes of ϕ 29 and PZA (12). Moreover, a 9-amino-acid long sequence just upstream of the head fiber protein is also highly homologous to ϕ 29 and PZA, with 6 out of the 9 amino acids identical to the corresponding amino acids of the major head proteins of ϕ 29 and PZA (11). These results clearly indicate that the organization of gene 8 (major

head protein), gene 8.5 (head fiber protein) and gene 9 (tail protein) in Nf is exactly the same as in the genomes of $\phi 29$ and PZA and that there appears to be no gene involved in the assembly of the head fiber protein.

This research was supported in part by Public Health Service Grant GM28013 from the National Institute of General Medical Sciences and by the American Cancer Society Grant NP-704.

REFERENCES

- 1) Geiduschek, E. P. and Ito, J., Regulatory mechanisms in the development of lytic bacteriophages in *Bacillus subtilis*. In *The Molecular Biology of the Bacilli*, Vol. I, ed. by Dubnau, D. A., Academic Press, New York (1982), p. 203–245.
- 2) Gyllenstein, U., Direct sequencing of *in vitro* amplified DNA. In *PCR Technology: Principles and Applications for DNA Amplification*, ed. by Erlich, H. A., Stockton Press, New York (1989), p. 45–60.
- 3) Ikeda, Y., Saito, H., Miura, K., Takagi, J., and Aoki, H., DNA base composition, susceptibility to bacteriophages, and interspecific transformation as criteria for classification in the genus *Bacillus*. *J. Gen. Appl. Microbiol.*, **11**, 181–190 (1965).
- 4) Ito, J., Bacteriophage $\phi 29$ terminal protein: Its association with the 5' termini of the $\phi 29$ genome. *J. Virol.*, **28**, 895–904 (1978).
- 5) Kawamura, F. and Ito, J., Studies of *Bacillus subtilis* phage M2: A physical map of the M2 genome. *Virology*, **83**, 233–245 (1977).
- 6) Mizukami, Y. and Hirokawa, H., Evidence that *Bacillus* phage M2 is derived from Nf by deletion in a particular region of the genome. *J. Gen. Appl. Microbiol.*, **31**, 429–434 (1985).
- 7) Mizukami, Y., Sekiya, T., and Hirokawa, H., The nucleotide sequences of the heterologous region between the genomes of *Bacillus* phages M2 and Nf that indicate the two phages are originally identical. *FEBS Lett.*, **197**, 311–314 (1986).
- 8) Reilly, B. E., Nelson, R. A., and Anderson, D. L., Morphogenesis of bacteriophage $\phi 29$ of *Bacillus subtilis*: Mapping and functional analysis of the head fiber gene. *J. Virol.*, **24**, 363–377 (1977).
- 9) Salas, M., Phages with protein attached to the DNA ends. In *The Bacteriophages*, Vol. I, ed. by Calendar, R., Plenum Press, New York (1988), p. 169–191.
- 10) Sanger, F. S., Nicklen, S., and Coulson, A. R., DNA sequencing with chain terminating inhibitors. *Proc. Natl. Acad. Sci. USA*, **74**, 5463–5467 (1977).
- 11) Shimizu, N., Miura, K., and Aoki, H., Characterization of *Bacillus subtilis* bacteriophages II: Isolation and morphology of phage Nf and properties of its DNA. *J. Biochem.*, **68**, 277–286 (1970).
- 12) Vlček, Č. and Pačes, V., Nucleotide sequence of the late region of *Bacillus* phage $\phi 29$ completes the 19285-bp sequence of $\phi 29$ genome. Comparison with the homologous sequence of phage PZA. *Gene*, **46**, 215–225 (1986).
- 13) Yoshikawa, H., Elder, J. H., and Ito, J., Comparative studies on the small *Bacillus* bacteriophages. *J. Gen. Appl. Microbiol.*, **32**, 39–49 (1986).
- 14) Yoshikawa, H. and Ito, J., Terminal proteins and short inverted terminal repeats of the small *Bacillus* bacteriophage genomes. *Proc. Natl. Acad. Sci. USA*, **78**, 2596–2600 (1981).