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 φ29, φ15, 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 (φ29, PZA, and φ15), 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 φ29, this region would correspond to gene 8.5, the structural gene of the head fiber protein of φ29 (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 φ29 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 φ29’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 GATTCATCTTCTGGACTGCAAGCTGATAATTCTAGCTGACATGATGCAAAACCTCTG
2 DSSSAGCAAACTC MYSF TAY ANSD
3 ACAATTGGATGCTTATTcTGAATAATTGTATACCAATGGACTGACATGATGCAAAACCTCTG
4 IVAPHLKFLSSENNIEISY
5 ATCCGAGTCAAGCAAGCAACAGAAGCTGACATGATGCAAAACCTCTG
6 ADEDTIPEVSYSRDLKAGDK
7 AACACAACAATGGTTATATCCCAATTAGATGCTGGAAAGTGGCGATGAAACCTCTG
8 TIDLYPLVAWKVIAPQEDIT
9 CAATCGGGGAGCAGTGATGCTGGAAAGTGGCGATGAAACCTCTG
10 TGDRTSYSVKNGQVKJTDDL
11 GCAACATTCGCTGCTAGCCCTGATATACCAATGGACTGACATGATGCAAAACCTCTG
12 TAFGAYAUSLARAGQLVTVAI
13 TCTGACAGGTATTTGACGATATACCAATGGACTGACATGATGCAAAACCTCTG
14 STVFDTIITPDVGLNVDDDV
15 TAAAGGATCTTCTGATGCTCAATGCAACAGGAAATAGGATGCTGGATGCTAC
16 AELKSNATDANKLAKIRDL
17 TGATTCCAGTTCTGATGCTCAATGCAACAGGAAATAGGATGCTGGATGCTAC
18 VSNLDVKAFLNGSTSEDNKN
19 TAAATCTTAGAATTTACTGTTAGCAACTCTGGCTGTTATATTGGAAACC
20 NLRNMLVLVSNPAILAFINANP
21 CTGACACAGATACGCTGCAACAGCTAAAGGATGCTGGATGCTAC
22 TATTTTTGGGVRIFRAPGNS
23 CAATGCTGCTATCGCTGCAATGCAACAGGAAATAGGATGCTGGATGCTAC
24 TATDVAATLVKDFNLLLAAMR
25 GAAACAGTCTGCTAGATACCAATGGACTGACATGATGCAAAACCTCTG
26 TAGYILI*LSGSMAYVPLS
27 CACAGTGTTGATATCTGGCTGCGTTCTCCAAATGTATATTAAATAGCTAGGATGCT
28 DVFIFSAARSPMIIKILRWF
29 TCAAGCTGGGACGCTCATATTCTGATATCTGGACTGACATGATGCAAAACCTCTG
30 TCAADAQYSYFPNARNPVRVHVI

Fig. 1. Complete nucleotide sequence of Hind-Bell fragment of the Nf genome. The phage DNA and restriction enzymes were prepared according to Kawamura and Ito(5).
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 φ29 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 429 and PZA (12). Moreover, a 9-amino-acid long sequence just upstream of the head fiber protein is also highly homologous to φ29 and PZA, with 6 out of the 9 amino acids identical to the corresponding amino acids of the major head proteins of φ29 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 φ29 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

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).
12) Vlček, C. and Pačes, V., Nucleotide sequence of the late region of *Bacillus* phage φ29 completes the 19285-bp sequence of φ29 genome. Comparison with the homologous sequence of phage PZA. *Gene*, 46, 215–225 (1986).