A 26-Base-Pair Repetitive Sequence Specific for Neisseria gonorrhoeae and Neisseria meningitidis Genomic DNA

FREDERICK F. CORREIA, SUMIKO INOUYE, AND MASAYORI INOUYE*

Department of Biochemistry, State University of New York at Stony Brook, Stony Brook, New York 11794

Received 19 February 1986/Accepted 9 May 1986

Two-dimensional heteroduplex mapping of Neisseria gonorrhoeae genomic DNA revealed a number of spots, indicating the existence of repetitive sequences. When one of the spots was extracted and used as a probe for Southern blot analysis, two HindIII bands (11.0 and 3.6 kilobases [kb]) of the genomic digest hybridized with approximately equal intensity. The 3.6-kb fragment was cloned and found to contain two different types of repeated sequence. One type was approximately 1.1 kb in length and was found at least twice in the entire genome. The other consisted of a 26-base-pair family GT(CA)C(Py)G(Pu)TTTTGTTAAT(Py)C(Pu)CTATA (Py, pyrimidine; Pu, purine) that was repeated at least 20 times in the entire genome. This repetitive sequence was found also in Neisseria meningitidis but not in various other gram-negative bacteria.

We recently reported the development of a two-dimensional S1 nuclease heteroduplex mapping protocol, with which DNA undergoing rearrangements as well as DNA with repeated sequences could be detected and eventually cloned (23). In this study we report the discovery, using this methodology, of a remarkably well-conserved 26-nucleotide-long repeated sequence that occurs in the genomes of Neisseria gonorrhoeae and Neisseria meningitidis.

There have been few previous reports of such a family of sequences in the eubacteria (3, 19). One of these by Stern et al. (19) described a repetitive extragenic palindromic (REP) sequence that was approximately 35 nucleotides long with 35 copies in 16 independent operons of Escherichia coli and Salmonella typhimurium. The function of the REP sequence has not been established.

Other examples of repetitive sequences in E. coli include the rhs sequences (7), insertion sequences, the rRNA genes, and some tRNA genes. These repetitive sequences are distinguished from REP-type sequences in that their copy number is small and their regions of homology tend to be more extensive.

We chose to study N. gonorrhoeae with the two-dimensional heteroduplex mapping technique because during the course of infection this organism can alter several of its surface proteins, thus implying the existence of an unusual system of regulation at the DNA level. Among these alterations are the reversible expression of pili and opacity proteins (11, 14, 18). These phenomena are termed phase variation. Extensive antigenic variation of these two categories of protein also occurs and in the case of pilus genes has been correlated with their phase variation (5, 14, 15). In N. gonorrhoeae, direct repeats, which are thought to play a role in the recombinations that mediate pilus phase and antigenic variation, have been reported (4, 15). They are different in sequence from the repetitive family that we found.

MATERIALS AND METHODS

DNA preparation. Genomic DNA was prepared by the method of Yee and Inouye (22) from cell pellets of the following strains: N. gonorrhoeae 5019 received from K. Johnston of Enzo Biochem Inc., Neisseria meningitidis strain Y received from Enzo Biochem Inc.; DNA from strain MS11 was kindly provided by E. Segal of the Research Institute of Scripps Clinic. All plasmid DNA was isolated by the method of Birnboim and Doly (1).

Plasmid constructions. All cloning procedures were performed as described by Maniatis et al. (9). The actual plasmid constructions are outlined below. Restriction digestion of the DNA was as recommended by the manufacturers (Bethesda Research Laboratories, Gaithersburg, Md.; New England Biolabs, Inc., Beverly, Mass.).

Two-dimensional S1 nuclease heteroduplex mapping. The two-dimensional S1 nuclease heteroduplex mapping procedure of Yee and Inouye (23) was followed. Fifty micrograms each of isogenic, opaque, and translucent strain 5019 genomic DNA was digested with Alul. The DNA was denatured and then renatured at 15°C for 18 h, after which it was electrophoresed on a 5% polyacrylamide gel. The gel lane containing separated heteroduplex as well as homoduplex DNA species was cut out and soaked in S1 buffer and then reacted with S1 nuclease in situ. The gel strip was loaded onto a second-dimension 5% polyacrylamide gel and electrophoresed, resolving the S1-digested heteroduplex DNA as spots beneath a prominent diagonal of homoduplex DNA.

The DNA was electroeluted from spots, nick translated, and hybridized to nitrocellulose blots by procedures described previously (23).

Fractionation of strain 5019 DNA. A 0.7% cylindrical agarose gel with a length of 18.5 cm and a diameter of 2.5 cm was loaded with 50 µg of HindIII-digested genomic DNA. The buffer system used was 1 × T (40 mM Tris acetate [pH 8.0] containing 5 mM sodium acetate and 1 mM EDTA). Electrophoresis was for 19 h at 30V. The gel was sliced into 50 sections. DNA was extracted from each fraction by placing the gel slices into Eppendorf tubes, which were then frozen and subsequently centrifuged. The DNA-containing supernatant was saved and pooled with the supernatant from a repetition of the freeze-spin procedure. DNA from each fraction was extracted with phenol, followed by ethanol precipitation. The final pellet was dissolved in 60 µl of 0.1 × TE (1 × TE is 10 mM Tris hydrochloride [pH 8.0] and 1 mM EDTA).

* Corresponding author.
Hybridization analysis. All blots were made by the method of Southern (17). Plasmid probes were nick translated by the procedure of Rigby et al. (12). Hybridizations were performed overnight at 42°C by the general procedures of Davis et al. (2) with 50% deionized formamide-5× SSPE (1× SSPE is 180 mM NaCl, 10 mM NaH₂PO₄ [pH 7.4], 1 mM EDTA (pH 7.4))-1× Denhardt solution-0.3% sodium dodecyl sulfate containing 100 μg of denatured, sonicated pollock roe DNA (Sigma Chemical Co., St. Louis, Mo.) per ml. Blots were washed four times at room temperature in 10 mM NaH₂PO₄ (pH 6.8)-0.2% sodium dodecyl sulfate-1 mM EDTA for 15-min periods before autoradiography.

Colonies hybridizations were conducted by the pick-and-screen method described previously (6). Hybridization was conducted at 42°C overnight in 50% deionized formamide and 4× SSC (1× SSC is 750 mM NaCl plus 75 mM sodium citrate [pH 7.0]). Filters were washed three times at 42°C in hybridization fluid minus probe for 90-min periods and three times for 30 min in 2× SSC before autoradiography.

The 21-base oligomer (GATTTGTATATTACCTATA) was synthesized on a Systec Microsyn 1450 DNA synthesizer by the phosphoramidite method and purified by preparative gel electrophoresis as described previously (6). The oligomer was kinased with 30 μCi of [γ-³²P]ATP. Hybridization of the oligomer to blots was carried out as follows. Blots were prehybridized for 10 min in 6× NET (1× NET is 150 mM NaCl, 15 mM Tris hydrochloride [pH 7.5], 1 mM EDTA)-10× Denhardt solution-0.5% Nonidet P-40 containing 100 μg of denatured, sonicated pollock roe DNA per ml. The oligomer probe was then added to prehybridized blots and incubated at 42°C overnight. Filters were washed 4 times for 15 min in 6× SSC before autoradiography.

DNA sequencing. Fragments to be sequenced were cloned into pUC9 vectors and then into phage vectors M13 mp18 or M13 mp19, or cloned directly into the M13 phage cloning vectors M13 mp18 and M13 mp19 (21). Transformations and transfections were conducted with E. coli JM83 and JM105 (21). DNA sequencing was by the dideoxy chain-termination method described by Sanger et al. (13).

RESULTS

Two-dimensional S1 nuclease heteroduplex mapping of N. gonorrhoeae. We applied the two-dimensional S1 nuclease heteroduplex mapping technique on DNA from N. gonorrhoeae 5019 (Fig. 1). Beneath the highly ethidium bromide-stained diagonal of homoduplex DNA are numerous spots and streaks. These correspond to heteroduplex DNA, the unpaired strands of which were digested with S1 nuclease and, hence, migrated faster during electrophoresis in the second dimension. We excised several of the spots from the gel and electroeluted and nick-translated the DNA. In this study we focused on studies done on spot H DNA, which migrates at an approximate length of 1.1-kilobases (kb) in the second-dimension polyacrylamide gel (Fig. 1).

Southern blot analysis reveals a multicopy sequence family. Chromosomal DNA from strain 5019 digested with HindIII

FIG. 1. S1 nuclease heteroduplex mapping of strain 5019. DNA from the opaque derivative (50 μg) and DNA from the translucent derivative (50 μg) were digested with AluI and mixed, denatured, renatured, and electrophoresed on a 5% polyacrylamide gel. The gel strip containing the renatured DNA was then treated with S1 nuclease and electrophoresed in the second dimension as described in the text. The arrowhead indicates the heteroduplex spot H which was characterized further.

FIG. 2. (A) Hybridization analysis of the translucent derivative of strain 5019 digested with HindIII and probed with the following nick-translated fragments: lane 1, spot H; lane 2, 3.6-kb HindIII fragment; lane 3, 2.75-kb AvrI1 fragment; lane 4, 0.75-kb HindIII-AvrI1 fragment. (B) Restriction map of the 3.6-kb HindIII fragment in pUC9, indicating relationships of probes used in panel A. Abbreviations: A, AvrI1; H, HindIII.
was used for Southern blot hybridization. Employing spot H DNA as a probe, we obtained the hybridization pattern shown in Fig. 2A, lane 1, in which two bands at 11.0 and 3.6 kb hybridized with the probe.

To clone the 3.6-kb HindIII fragment, we digested chromosomal DNA with HindIII and fractionated it on a preparative 0.7% agarose cylindrical gel. One fraction which showed positive hybridization with spot H DNA probe was cloned into the HindIII site of pUC9. Colony hybridization with spot H DNA as probe allowed us to isolate the clone carrying the 3.6-kb HindIII insert. This plasmid was designated pNG273.

When nick-translated pNG273 was used as probe against blotted N. gonorrhoeae 5019 DNA digested with HindIII, two prominent hybridization bands of 11.0 and 3.6 kb were seen among a background of faint hybridization signals (Fig. 2A, lane 2). To delineate the hybridizing region we digested the 3.6-kb HindIII insert into three fragments with AvaII. The sizes of these fragments were 2.75, 0.75, and 0.1 kb (Fig. 2B). In Fig. 2A (lanes 3 and 4) are shown the results of using the 2.75- and 0.75-kb fragments as probes against HindIII-digested genomic pNG273 DNA. The 2.75-kb probe hybridized to the 5.0- and 11.0-kb fragments in the same manner as did the spot H DNA probe (Fig. 2A, lane 3). However, to our surprise, the 0.75-kb probe hybridized to multiple fragments, including the 3.6-kb fragment (Fig. 2A, lane 4). These results indicate that the 3.6-kb fragment contains two types of repetitive sequence. One type is located within the 2.75-kb AvaII fragment and occurs at least twice in the total genome. Because the hybridization density of the 11.0-kb band is almost identical to that of the 3.6-kb band in Fig. 2A, lane 3, much of the 2.75-kb fragment must be repeated in the 11.0-kb HindIII fragment. The second type of repetitive sequence is within the 0.75-kb HindIII-AvaII fragment, and it appears to be repeated as many as 20 times in the total genome, as judged by the number of hybridizing bands in Fig. 2A, lane 4.

**Determination of the sequence responsible for the multiple hybridization pattern.** To further characterize the repetitive sequence, HindIII fragments from fractionated genomic DNA were shotgun cloned into the HindIII site of pUC9, and transformants of JM83 were selected with the 0.75-kb HindIII-AvaII probe. Among these clones, two that harbored a 5.0-kb HindIII insert (pNG268) and a 13.0-kb HindIII insert (pNG112) were further characterized. Digestion of the HindIII inserts of pNG273, pNG268, and pNG112 by Rsal, Thal, and AvaI, respectively, yielded a small unique hybridizing fragment when the 0.75-kb HindIII-AvaII probe was used (data not shown). These unique hybridizing species, a 493-base-pair (bp) Rsal fragment, a 383-bp HindIII-AvaI fragment, and a 418-bp HindIII-Thal fragment, were subcloned into M13 mp18 and M13 mp19 and sequenced by the Sanger dideoxy method (13) (Fig. 3A, B, and C, respectively). Within these sequences homologous sequences of 26 bp were found, as indicated in Fig. 3. This 26-bp common sequence was considered to be the cause of the hybridization between the 0.75-kb HindIII-AvaI probe and each of the subclones. It should be noted that within the 493-bp Rsal fragment (Fig. 3A) there were two such homologous sequences separated by 440 bp. These homologous sequences are listed in Fig. 4, together with the consensus sequence. The match between four independent sequences was perfect at 21 of 26 positions. At all but one of the remaining five positions, base changes were caused by transition substitutions, conserving a purine or pyrimidine at that position.

To prove that this sequence was responsible for the multiple hybridizations detected with the 0.75-kb probe, a 21-base oligomer (GATTTTTGTTAATTCACTATA) which was a subset of the 26-base consensus sequence was synthesized. It was used to probe genomic blots of strains 5019 and MS11 that had been digested with Clal. Clal was used to ensure complete digestion of MS11 DNA, because we previously found that strain MS11 DNA does not give complete digestion with HindIII, in spite of a vast excess of restriction enzyme or prolonged incubation. In Fig. 5A is shown strain 5019 genomic DNA (lanes 1 and 2) and strain MS11 genomic DNA (lanes 3 through 6) digested with Clal and probed with the 0.75-kb HindIII-AvaII fragment. In comparison, in Fig. 5B is shown strain 5019 genomic DNA (lanes 1 and 2) and MS11 genomic DNA (lanes 3 and 4) digested with Clal and probed with the 21-base oligomer. The results conclusively demonstrate that the 26-bp consensus sequence is responsible for the multiple hybridizations seen with the 0.75-kb probe.

**Existence of the consensus sequence in Neisseria spp. and other bacteria.** The Southern blot data presented in Fig. 5 show that interstrain (MS11 versus 5019) as well as interspecies (N. gonorrhoeae versus N. meningitidis) differences in the copy number and organization of the repetitive sequence exist. In N. gonorrhoeae 5019, as judged by the hybridization with the synthetic oligomer, at least 24 bands can be observed (Fig. 5B, lane 2), while in N. gonorrhoeae MS11, 22 bands are visible (Fig. 5B, lane 3). The genome of N. meningitidis Y yielded 21 bands (Fig. 5B, lane 5).

To examine whether pilus and protein II phase variation correlates with changes in organization and number of the repetitive sequence, we probed DNA from four phase vari-
FIG. 4. Complete nucleotide sequences of the 493-bp Rsal subclone of pNG273 (A). Single-letter amino acid codes are put on the first base of the codons. The amino acid sequences above the nucleotide sequence represent the open reading frame from left to right, and those below the nucleotide sequence represent the open reading frame from right to left. Hyphens represent the termination codons. (B) The 383-bp HindIII-AvaI subclone of pNG112. (C) The 418-bp HindIII-AvaI subclone of pNG268. In panels A through C, boxed areas contain the 26-bp homologies. (D) Comparison of four independent 26-bp repeated sequences. Asterisks indicate the positions where A was found only for pNG273(R).

FIG. 5. Hybridization analysis of Clal-digested N. gonorrhoeae phase variants and N. meningitidis DNA. (A) The 0.75-kb HindIII-AvaII fragment was used as a probe. Lane 1, opaque strain 5019; lane 2, translucent strain 5019; lane 3, piliated and translucent strain MS11; lane 4, nonpiliated and opaque strain MS11; lane 5, nonpiliated and opaque strain MS11; lane 6, piliated and opaque strain MS11. (B) The 21-bp oligomer was used as a probe. Lane 1, opaque 5019; lane 2, translucent 5019; lane 3, piliated and translucent MS11; lane 4, piliated and opaque MS11; lane 5, N. meningitidis Y. Numbers at the left represent the molecular masses in kilobases for panel A. Lines indicate the correlation between the different blots.
The pheno\textsc{\textnormal{t}}\textnormal{o}\textsc{\textnormal{p}}\textsc{\textnormal{e}}s occurs rearrangements its recombination. The HindIII total genomic digest of \textit{N. gonorrhoeae} (Fig. 2A, lane 3). It should also be noted that spot H DNA from the heteroduplex S1 mapping (Fig. 1) hybridized to both the 3.6- and the 11.0-kb HindIII fragments with almost equal intensity (Fig. 2A, lane 1). Therefore, it is considered that spot H (1.1 kb) is formed between identical sequences which exist within these two HindIII fragments and that this region of homology extends for at least 1.1 kb. This latter conclusion is based on the size estimate of spot H, which was resolved after S1 nuclease treatment. The 11.0-kb HindIII fragment is cloned and is being characterized.

The other repetitive family consists of only 26 bp, but it occurs in many copies in the \textit{Neisseria} genome. The most probable role of these small repetitive sequences is that of recombination sites. In \textit{E. coli} it has been reported that the minimum length necessary for recombination is 20 bp (20). The four independent sequences in Fig. 4D include mismatches among them that limit the longest stretch of exact match to 11 bp. These sequences that are only found in \textit{Neisseria} spp. may thus require specific protein factors if they are, in fact, involved in site-specific recombinations.

A distinctive property of \textit{N. gonorrhoeae} is the ability of its colonies to exhibit phase variation. This phenomenon includes switching among pilated and nonpiliated cells, as well as switching among opaque and translucent colonial phenotypes at a frequency high above the normal mutation rate. In the case of pilus switching, substantial genomic rearrangements (11), including deletion events at expression sites, appear to mediate antigenic variability as well as phase variation. These deletions are postulated to occur through recombination at directly repeated sequences in the expression loci (4, 15). The DNA sequence of the pilus structural gene, including the direct repeats, has been published (4, 5, 10). We were unable to find the 26-bp consensus sequence within the published pilus sequences. We did note an upstream noncoding sequence that showed 73\% similarity (GGCGCATTTcGTTAcTIttTATI) to the 26-bp consensus sequence (nucleotide positions 63 to 89 in the MS11 sequence published by Meyer et al. [10]). Southern blot experiments in which the synthetic oligomer was used as a probe against DNA isolated from the four possible phase variant phenotypes of MS11 failed to show banding pattern changes of the repetitive sequence. We conclude that the 26-bp repeat most likely is not involved in pilus or protein II switching.

It was intriguing that one of our clones contained two of the 26-bp sequences in inverted repeat orientation separated by 440 bp (Fig. 4A). This was reminiscent of a small insertion element (16). The 26-bp inverted repeat did not match any published inverted repeats of insertion elements. Two open reading frames existed on this 493-bp clone (Fig. 4A). Neither of the putative peptides from these open reading frames showed similarities with peptides in a protein library which included hypothetical insertion element-encoded proteins (8).

**ACKNOWLEDGMENTS**

We are grateful to R. Wilpan for synthesizing the DNA oligomer, E. Segal for strain MS11 DNA, and P. March and W. P. Yang for critical reading of the manuscript.

This work was supported by Public Health Service grants GM-19043 (to M.I.) and GM-26843 (to S.I.) from the National Institutes of Health.

**LITERATURE CITED**

1. Birnboim, H. C., and J. Doly. 1979. A rapid alkaline extraction procedure for screening recombinant plasmid DNA. Nucleic Acids Res. 7:1513–1523.
2. Davis, R. W., D. Butstein, and J. R. Roth. 1980. Advanced bacterial genetics. Cold Spring Harbor Laboratory. Cold Spring Harbor, N.Y.
3. Gilson, E., J.-M. Clement, D. Brutlag, and M. Hofnung. 1984. A family of dispersed repetitive extragenic palindromic DNA sequences in \textit{E. coli}. EMBO J. 3:1417–1421.
4. Haas, R., and T. F. Meyer. 1986. The repertoire of silent pilus genes in \textit{Neisseria gonorrhoeae}: evidence for gene conversion. Cell. 44:107–115.
5. Hahblom, P., E. Segal, E. Billiard, and M. So. 1985. Intragenic recombination leads to pilus antigenic variation in \textit{Neisseria gonorrhoeae}. Nature (London) 315:156–158.
6. Inouye, S., and M. Inouye. 1985. Oligonucleotide directed site-specific mutagenesis using double stranded plasmid DNA. In S. Narang (ed.), DNA and RNA synthesis. Academic Press, Inc., New York.
7. Lin, R.-J., M. Capage, and C. W. Hill. 1984. A repetitive DNA sequence, \textit{rhs}, responsible for duplications within the \textit{Escherichia coli} K-12 chromosome. J. Mol. Biol. 171:1–18.
8. Lipman, D. J., and W. R. Pearson. 1985. Rapid and sensitive protein similarity searches. Science 227:1435–1441.
9. Maniatis, T., E. F. Fritsch, and J. Sambrook. 1982. Molecular cloning: a laboratory manual. Cold Spring Harbor Laboratory. Cold Spring Harbor, N.Y.
10. Meyer, T. F., E. Billiard, R. Haas, S. Storzbach, and M. So. 1984. Pili genes of \textit{Neisseria gonorrhoeae}: chromosomal organization and DNA sequence. Proc. Natl. Acad. Sci. USA 81:6110–6114.
11. Meyer, T. F., N. Mlawer, and M. So. 1982. Pilius expression in \textit{Neisseria gonorrhoeae} involves chromosomal rearrangement. Cell 30:45–52.
12. Rigby, P. W. J., M. Dieckmann, C. Rhodes, and P. Berg. 1977. Labeling deoxyribonucleic acid to high specific activity in vitro by nick translation with DNA polymerase I. J. Mol. Biol. 113:237–251.
13. Sanger, F., S. Nicklen, and A. R. Coulson. 1977. DNA sequencing with chain terminating inhibitors. Proc. Natl. Acad. Sci. USA 74:5463–5467.
14. Schwabhe, R. S., P. F. Sparling, and J. G. Cannon. 1985. Variation of \textit{Neisseria gonorrhoeae} protein II among isolates from an outbreak caused by a single gonococcal strain. Infect. Immun. 49:250–252.
15. Segal, E., E. Billiard, M. So, S. Storzbach, and T. F. Meyer. 1985. Role of chromosomal rearrangement in \textit{N. gonorrhoeae} pili phase variation. Cell 40:293–300.
16. Shigetu, I., J. Meyer, and W. Arber. 1983. Prokaryotic IS elements, p. 159–206. In J. A. Shapiro (ed.), Mobile genetic elements—1983. Academic Press, Inc., New York, N.Y.
17. Southern, E. M. 1975. Detection of specific sequences among DNA fragments separated by gel electrophoresis. J. Mol. Biol. 98:503–517.
18. Stern, A., P. Nickel, T. F. Meyer, and M. So. 1984. Opacity determinants of \textit{Neisseria gonorrhoeae}: gene expression and chromosomal linkage to the gonococcal pili gene. Cell 37:447–456.
19. Stern, M. J., G. F. Ames, N. H. Smith, E. C. Robinson, and C. F. Higgins. 1984. Repetitive extragenic palindromic se-
quences: a major component of the bacterial genome. Cell 37:1015–1026.

20. Watt, V. M., C. J. Ingles, M. S. Urden, and W. J. Rutter. 1985. Homology requirements for recombination in Escherichia coli. Proc. Natl. Acad. Sci. USA 82:4768–4772.

21. Yanisch-Perron, C., J. Vieira, and J. Messing. 1984. Improved M13 phage cloning vectors and host strains: nucleotide sequences of the M13mp18 and pUC19 vectors. Gene 33:103–119.

22. Yee, T., and M. Inouye. 1981. Reexamination of the genome size of myxobacteria, including the use of a new method for genome size analysis. J. Bacteriol. 145:1257–1265.

23. Yee, T., and M. Inouye. 1984. Two-dimensional S1 nuclease heteroduplex mapping: detection of rearrangements in bacterial genomes. Proc. Natl. Acad. Sci. USA 81:2723–2727.