The worldwide increase in the rates of erythromycin resistance in streptococci since the 1990s has prompted extensive epidemiological as well as molecular investigations that in the last few years have led to substantial progress in the knowledge of the mechanisms, determinants, and genetic elements involved (35). The increase in the rate of resistance is partly due to the further spread among streptococci of conventional erm-class gene-encoded methylases, the presence of which usually leads to co-resistance to macrolide, lincosamide, and streptogramin B (MLS\textsubscript{B}) antibiotics (the MLS\textsubscript{B} phenotype) (36). To an even greater extent, however, it has been due to the emergence of an active efflux-mediated mechanism, encoded by mef\textsubscript{class} genes and associated with a pattern of low-level resistance affecting, among the MLS\textsubscript{B} antibiotics, only 14- and 15-membered macrolides (the M phenotype) (32). As a rule, a second efflux gene—an msr\textsubscript{class} determinant usually designated msr\textsubscript{D}—is located immediately downstream of the mef gene.

mef\textsubscript{A} was discovered in *Streptococcus pyogenes* (16) and is by far the most common mef\textsubscript{class} variant responsible for efflux-mediated erythromycin resistance in this species. However, the first mef\textsubscript{A}-carrying element, designated Tn\textsubscript{1207.1}, was detected in *Streptococcus pneumoniae* (30). Tn\textsubscript{1207.1} (7,244 bp) contains eight open reading frames (ORFs), of which mef\textsubscript{A} and msr\textsubscript{D} are the fourth and the fifth; it is integrated at a specific site of the pneumococcal chromosome into cel\textsubscript{B}, a late competence gene; and it has been reported to be transferable by transformation but not by conjugation. Whereas in *S. pneumoniae* Tn\textsubscript{1207.1} is the sole recognized element carrying mef\textsubscript{A}, in *S. pyogenes* it is not detected as such but is detected as part of larger and mobile composite elements. These elements differ depending on the tetracycline susceptibility or resistance of M-phenotype isolates of *S. pyogenes* (11, 22); and they are all chimeric in nature; i.e., they result from a transposon (identical or related to Tn\textsubscript{1207.1} inserted into a prophage (6, 22). One of two closely related elements, Tn\textsubscript{1207.3} (52,491 bp) (29) or \(\Phi\text{10394.4} (58,761 \text{ bp})\) (5, 6), is found in tetracycline-susceptible isolates integrated into the same chromosomal gene (comEC) and inserted into the same prophage (11), with the only difference being that in Tn\textsubscript{1207.3}, Tn\textsubscript{1207.1} represents the left end of the element, whereas in \(\Phi\text{10394.4} \) there is an additional left-hand region of 6 kb. Since this region has been reported to be quite variable in size (6), Tn\textsubscript{1207.3} could represent the left end of the variability range in which it is completely lacking (22). In tetracycline-resistant isolates, tetracycline resistance is consistently mediated by the tet\textsubscript{O} determinant and mef\textsubscript{A} is linked to tet\textsubscript{O} in a mobile, plasmage-like element (21, 22). In fact, there are a variety of related tet\textsubscript{O}-mef\textsubscript{A} plasmage-like elements, in which mef\textsubscript{A} is contained in a range of changeable and defective variants of Tn\textsubscript{1207.1} (11, 22). The most common such element, of which \(\Phi\text{m46.1}\) is the typical representative, is not integrated within the comEC gene and has been transferred to an *S. pyogenes* recipient in mating experiments (21). An 12-kb region of \(\Phi\text{m46.1}\) encompassing the tet\textsubscript{O} gene and the Tn\textsubscript{1207.1}-related transposon has been sequenced (11).

In the study described here, we determined the site of inte-
of the individual primer pairs. The Ex 

eX-Tag DNA polymerase (Amersham 

Biosciences). ORF analysis was performed by using the software ORF 

Finder available online (http://www.ncbi.nlm.nih.gov/projects/orf/). The criteria 

used to designate a potential ORF were the existence of a start codon and a 

minimum coding size of 50 amino acids. Sequence similarity and conserved 

domain searches were carried out by using the tools (BLAST and CDART) 

available at the National Center for Biotechnology Information of the 

National Library of Medicine (Bethesda, MD) (http://www.ncbi.nlm.nih.gov/). 

Electron microscopy. For phage induction, S. pyogenes strain m46 was treated 

with 0.2 μg ml⁻¹ mitomycin C (Sigma Chemical Co., St. Louis, MO) for 4 h at 

37°C. The bacteria were then centrifuged at 8,000 × g for 15 min, the supernatant 

was filtered through a 0.45-

m-pore-size nylon membrane (Millipore) and cen-

trifuged at 141,000 

g for 15 min, the supernatant 

degenerated into the chromosome, the complete genome sequence and 

organization, and the ultrastructure of Φm46.1. The genome displayed the distinctive modular arrangement of 
tailed bacteriophages, and electron microscopic analysis confirmed that it has the distinctive morphology of 

Siphoviridae family bacteriophages. Phage DNA (55,172 bp) was present in the 

host cell both as a prophage and as free circular DNA. The sequences of the ORFs of Φm46.1 were compared with those 

from protein databases. 

MATERIALS AND METHODS 

Bacterial strain. S. pyogenes m46 was used in this study. The strain, originally 

collected as a throat clinical isolate belonging to M type 4 (34), is resistant to 

tetracycline [MIC, 64 μg ml⁻¹; M phenotype; mef(A) genotype] and was used in all previous 

experiments that led to the identification of the genetic linkage between the 

tet(O) and the mef(A) genes in a mobile, prophage-associated element eventually 

designated Φm46.1 (3, 11, 22). 

Gene detection and amplification experiments. The principal primer pairs 

used in the PCR experiments are listed in Table 1. DNA preparation and 
amplification and electrophoresis of the PCR products were carried out by 
established procedures and according to the conditions recommended for the use of the individual primer pairs. The Ex Taq system (TaKaRa Bio, Shiga, Japan) 

was used when the expected sizes of the PCR products exceeded 3 kb. 

Phage DNA sequencing and sequence analysis. Long PCR experiments, in-

verse PCR, and primer-walking techniques were used to obtain overlapping 

fragments of the Φm46.1 prophage with primers designed from Φ10394.4 ( 

EMBL accession no. NC_006086). Inverse PCR was carried out as described by 

Sambrook and Russell (28); HindIII-, Mbol-, or SexAI-digested genomic DNA 

(all endonucleases were from Roche Applied Science, Basel, Switzerland) was 
ligated and used as the template in the PCR assays (Table 1). All PCR products 

used for sequence analysis were purified by using Montage PCR filter units 

(Millipore Corporation, Bedford, MA). Amplicons were sequenced (bidirec-
tionally or by primer walking) with an ABI Prism sequencer (Perkin-Elmer Applied 

Biosystems, Foster City, CA) and dye-labeled terminators. The sequences were 

analyzed by using the Sequence Navigator software package (Perkin-Elmer Ap-
plied Biosystems). ORF analysis was performed by using the software ORF 

Finder available online (http://www.ncbi.nlm.nih.gov/projects/orf/). The criteria 

used to designate a potential ORF were the existence of a start codon and a 

minimum coding size of 50 amino acids. Sequence similarity and conserved 
domain searches were carried out by using the tools (BLAST and CDART) 

available at the National Center for Biotechnology Information of the 

National Library of Medicine (Bethesda, MD) (http://www.ncbi.nlm.nih.gov/). 

RESULTS AND DISCUSSION 

Bacteriophage Φm46.1 integration site. Use of the portion of 

Φm46.1 that included the tet(O) and mef(A) genes (11) as a 

TABLE 1. Principal oligonucleotide primer pairs used

| Procedure and gene | Primer designation | Sequence (5'-3') | Source or reference | Product size (bp) |
|--------------------|--------------------|-----------------|---------------------|------------------|
| Inverse PCR        |                    |                 |                     |                  |
| tet(O)             | TETO-INV1          | CTGGTTCGTCAATTGCGAACCA | This study         |                  |
| tet(O)             | TETO-INV2          | TATATTGTTTTTCTGCAAAGGATG | This study         |                  |
| orf57              | INV3               | TCTAGGGCTGTAATTTTCTC | This study         |                  |
| orf57              | INV4               | CATATAGGAAATTTTCTTAG | This study         |                  |
| orf54              | INV5               | TAATGCTGACCGAACCCACAC | This study         |                  |
| orf55              | INV6               | GTTTGGCTATAGGCGATATT | This study         |                  |
| Φm46.1 chromosome integration site | | | | |
| spy1198            | LYT-for            | GAAATGATAGCAGAGGAAGGAGTT | This study         | 4,425 |
| mef(A)             | MEFA2              | TCTTCTGTACTAAAGATGG | This study         | 31               |
| tet(O)             | TETO1              | AACTAGGATTCTTGGCTCAC | This study         | 6,573 |
| spy1195            | THIO-rev           | GTAAAGATGACGAAAGGAGTT | This study         |                  |
| Φm46.1 chromosome integration site | | | | |
| Phage sequencing   |                    |                 |                     |                  |
| mef(A)             | MEFA1              | AGTATCATTAATCATCAGTGC | This study         | 16,518 |
| orf22              | HELY-R             | CCCATGCTTAGGATCTAGTGC | This study         |                  |
| orf19              | PRI-for            | AAGCAGAAGAAGTTACAAAGG | This study         | 17,679 |
| orf46              | HP-rev             | GTGAAGATTTTCTTCCTACCTGC | This study         |                  |
| orf45              | MJT-for            | ATGTCCGTATTGAAAGGAGATT | This study         | 10,503 |
| orf54              | LYS-rev            | GGTGGAACCGGCTTGCGAGGG | This study         |                  |
| orf53              | HOL-rev            | TTGGTCATCTGATTTGACACGC | This study         | 5,893 |
| tet(O)             | TETO2              | TCCACATGTTGCATATCGTA | This study         |                  |

In the table, Φm46.1 indicates a bacteriophage with its chromosome integration site, which is important for understanding the genetic integration into the host cell's genome. The table lists the primer pairs used for inverse PCR, with their corresponding sequences and product sizes. The methodology for sequence analysis and integration site detection is also described, including the use of PCR techniques and electronic microscopy to study the phage's morphology.
An almost completely overlapping 18-bp sequence was found in the \( \Phi m46.1 \) and MGAS10750 genomes. By pairing primers TETO1 and THIO-rev, a 6,573-bp segment at the attR region was amplified from the lysogenic bacterial DNA template. Similar alignment assays revealed that positions 5,298 to 5,357 of the amplicon were the same as positions 1142,426 to 1142,393 of the MGAS10750 genome. The same 18-bp overlapping sequence found in the \( \Phi m46.1 \) and MGAS10750 genomes was also detected when the left junction was analyzed. This 18-bp sequence, shared by both the phage \([\text{between } tet(O) \text{ and } mef(A)]\) and the host bacterium (near the 3' end of the 23S rRNA uracil methyltransferase gene), should be the core site, i.e., the critical sequence where the site-specific recombination process presumably takes place. Definition of the core site enabled the integrative reaction and structure relationship between the phage genome and the host chromosome to be deduced according to the integration mechanism elucidated in bacteriophage lambda (12, 13). Interestingly, the 23S rRNA uracil methyltransferase gene identified as the chromosome integration site of \( \Phi m46.1 \) in \( S. pyogenes \) is the same as two integrated conjugative elements (ICEs), 2096-RD.2 and 6180-RD.1, in \( S. pyogenes \) strains MGAS2096 and MGAS6180, respectively (8).

Crucially, identification of the core site provided the knowledge that the phage attachment sequence \( (\text{attP}) \) falls right in the region between \( tet(O) \) and \( mef(A) \) in the previously sequenced ~12-kb DNA fragment of \( \Phi m46.1 \). Therefore, both that sequence and the previously disclosed \( tet(O)-mef(A) \) linkage (21), with \( tet(O) \) being found ~5.5 kb upstream of \( mef(A) \), were from an unintegrated, circular form of \( \Phi m46.1 \), whereas in the integrated form, \( tet(O) \) is downstream of and far more distant from \( mef(A) \), with the former gene being close to the

**FIG. 1.** Bacteriophage \( \Phi m46.1 \) integration site. \( \Phi m46.1 \) was integrated into the chromosome of \( S. pyogenes \) \( m46 \) within a 23S rRNA uracil methyltransferase gene. This gene, detected in all \( S. pyogenes \) genomes sequenced to date, encodes a protein having the highest degree of amino acid similarity to the amino acid sequence of a protein encoded by the corresponding gene from \( S. pyogenes \) MGAS10750 (EMBL accession no. CP000262), which, like \( S. pyogenes \) \( m46 \), belongs to M type 4. Chromosomal ORF designations and aligned host genome sequences are thus from \( S. pyogenes \) MGAS10750. (A) Left junction. (Above) ORF map; chromosomal ORFs are indicated as black arrows, and \( \Phi m46.1 \) ORFs are indicated as white arrows, with \( mef(A) \) being indicated by a checkered arrow. (Below) Alignment of the amplified sequence (64 bp) at the \( attL \) region with the \( \Phi m46.1 \) and \( S. pyogenes \) MGAS10750 genomes. Numbers above and below the \( attL \) sequence refer to base positions in the amplicon obtained with primers LYT-for and MEF2A (4,425 bp). (B) Right junction. (Above) ORF map; chromosomal ORFs are indicated as black arrows and \( \Phi m46.1 \) ORFs are indicated as white arrows, with \( tet(O) \) being indicated by a striped arrow. (Below) Alignment of the amplified sequence (64 bp) at the \( attR \) region with the \( \Phi m46.1 \) and \( S. pyogenes \) MGAS10750 genomes. Numbers above and below the \( attR \) sequence refer to base positions in the amplicon obtained with primers TETO1 and THIO-rev (6,573 bp). In the almost completely overlapping 18-bp sequence shared by the bacteriophage and the host bacterium genomes and representing the core site, nucleotides are indicated with capital letters.

query in a BLASTN analysis of the ~12-kb sequence—previously determined in our laboratory (EMBL accession no. AJT15499)—yielded a 308-bp region common to the Streptococcus agalactiae A909 genome sequence (EMBL accession no. CP000114) (33). In our sequence, this region is upstream of the \( \text{Tn}1207.1 \)-like transposon. In the \( S. agalactiae \) A909 genome, the region (bases 652,826 to 653,124) carries the 3' end (36 bp) of the \( \text{rumA} \) gene, which encodes a 23S rRNA uracil methyltransferase, and part of the intergenic region (272 bp) between \( \text{rumA} \) (SAK_0718 locus) and the adjacent gene (SAK_0719 locus). It is worth noting that the \( S. agalactiae \) \( \text{rumA} \) gene is homologous to a 23S rRNA uracil methyltransferase gene detectable in all \( S. pyogenes \) genomes sequenced so far; the highest degree of homology (73%) was with the gene from \( S. pyogenes \) MGAS10750 (EMBL accession no. CP000262), a strain that also shares M type 4 with \( S. pyogenes \) \( m46 \). PCR experiments were carried out with two primer pairs, one for the left junction and one for the right junction, to explore the possibility that this region was the M46.1 integration site. The two primers for the left junction were LYT-for, internal to \( \text{Spy}1198 \), an ORF of the MGAS10750 genome just upstream of the 23S rRNA uracil methyltransferase gene (\( \text{Spy}1197 \)), and MEF2A, internal to the \( mef(A) \) gene (Fig. 1A). The two primers for the right junction were TETO1, internal to the \( tet(O) \) gene, and THIO-rev, internal to \( \text{Spy}1195 \) of the MGAS10750 genome, located downstream of \( \text{Spy}1197 \) (Fig. 1B). By pairing primers LYT-for and MEF2A, a 4,425-bp segment at the \( attL \) region was amplified from the lysogenic bacterial DNA template. This amplicon sequence was aligned with the \( \Phi m46.1 \) genome (see the next paragraph) and with the MGAS10750 genome. Positions 1,727 to 1,764 of the amplicon were found to be the same as positions 55,156 to 20 of the \( \Phi m46.1 \) genome, and positions 1,701 to 1,744 were found to be the same as positions 1,142,452 to 1,142,409 of the \( S. pyogenes \) MGAS10750 genome (Fig. 1A).
right end and the latter gene being close to the left end of the prophage.

**Organization of the Φm46.1 genome.** The complete genome sequence of Φm46.1 was determined. Its size was 55,172 bp. The G+C content was 40.0%. Genome sequence analysis revealed the presence of 63 ORFs, 51 of which were transcribed in the same direction and 12 of which were transcribed in the opposite direction. The Φm46.1 ORF map is shown in Fig. 2, and the major characteristics of the ORFs are detailed in Table 2. The ORF sequences were compared with sequences from protein databases by using the BLASTP program. On the basis of these comparisons, most of the ORFs could be assigned to different modules, according to a modular organization typical of tailed phages (14). Interestingly, both ends of the Φm46.1 genome were represented by nonphage DNA.

(i) **Initial portion (1 to 248 bp).** The Φm46.1 prophage started with a short sequence (248 bp) that preceded the Tn10394-1-like transposon. The first 22 bp of this sequence restored the gene encoding the 23S rRNA uracil methyltransferase, whereas its first 222 bp displayed 88% homology with a sequence close to the left end of the pneumococcal mega element (EMBL accession no. AF274302).

(ii) **Tn10394-1-like transposon (249 to 7,267 bp).** As previously described in the same S. pyogenes isolate (11), the first two ORFs of the reference Tn10394-1 transposon (EMBL accession no. AF227520) are not found in the Tn10394-1-like transposon of Φm46.1. Thus, the acetyltransferase-encoding ORF originally designated orfD (11) was the first ORF (orf1) of the Φm46.1 genome. The following six ORFs, here renamed orf2 to orf7, include macrolide efflux genes mef(A) (orf3) and msr(D) (orf4).

(iii) **Phage modules (7,390 to 48,807 bp).** orf8 and orf9, which represent the beginning of the actual prophage-like region, probably corresponded to the left part (7,390 to 8,309 bp) of the lysogeny control module, which in Φm46.1 appears to be divided into two portions due to the insertion of a segment including a tet(O)-containing fragment from ICE 2096-RD.2 (8) and the mef(A)-containing Tn10394-1-like transposon. The DNA replication module (11,363 to 20,144 bp), spanning from orf14 to orf22, was highly conserved in the three mef(A)-carrying phage-like elements Φm46.1, Tn10394-1, and Φ10394.4. The DNA modification module (22,539 to 26,115 bp) spanned from orf27 to orf30. The DNA packaging and head morphogenesis module (26,993 to 33,651 bp) spanned from orf33 to orf42. Inside but apparently alien to the module, orf35 and orf36 seemed to be related to a toxin-antitoxin (TA) system (26) and might contribute to the stable maintenance of Φm46.1 in the bacterial population. It is noteworthy that the two putative TA-related genes partially overlapped. The tail morphogenesis module (34,336 to 39,108 bp) was formed by four ORFs (orf45 to orf48). Just downstream of this module, orf49 and orf50 (39,387 to 43,656 bp) encoded proteins exhibiting low levels of amino acid identity to distinct portions (for the presence of a stop codon) of a PblB protein. Proteins PblA and PblB act as phage-borne virulence factors by promoting bacterial binding to human platelets (7). The host cell lysis module (44,227 to 46,099 bp) was formed by three ORFs (orf53 to orf55), all of which were shared by ICE Sde3396, a newly described genetic element from Streptococcus dysgalactiae subsp. equisimilis (17). Downstream of the host cell lysis module, orf56 and orf57 (46,322 to 48,807 bp) encoded two site-specific recombinases (serine recombinases belonging to the resolvase family), i.e., enzymes that are usually found in the lysogeny control module (25). In Φm46.1, the two recombinase ORFs might have been separated from the rest of the lysogeny module by the insertion into the unintegrated phage DNA of a segment including the ICE 2096-RD.2 fragment and the Tn10394-1-like transposon.
### TABLE 2. Genome organization of Φm46.1

| ORF | Start position (bp) | Stop position (bp) | Size (no. of amino acids) | Predicted function | BLASTP analysis* | EMBL accession no. | % Amino acid identity (% amino acid similarity) |
|-----|---------------------|--------------------|---------------------------|-------------------|------------------|--------------------|-----------------------------------------------|
| orf1 | 558                | 1082               | 174                       | Acetyltransferase  | Ribosomal protein serine acetyltransferase (Bacillus thuringiensis ATCC 35646) | ZP_00740204.1 | 58 (77) |
| orf2 | 1113               | 1466               | 117                       | Hypothetical protein (Streptococcus dysgalactiae subspp. equisimilis) | CAJ45366.1 | 100 (100) |
| mef(A) | 2280             | 3497               | 405                       | Macrolide efflux protein |  |  |
| orf4 | 3617               | 5080               | 487                       | ABC transporter ATPase (Φm46.1 from S. pyogenes MGAS10394) |  |  |
| orf5 | 5500               | 5198               | 100                       | Hypothetical protein (Φm46.1 from S. pyogenes MGAS10394) |  |  |
| orf6 | 5855               | 5487               | 122                       | Hypothetical protein (Streptococcus salivarius) |  |  |
| orf7 | 7267               | 5852               | 171                       | ImpB/MucB/SamB family protein |  |  |
| orf8 | 7617               | 7390               | 75                        | Transcriptional repressor |  |  |
| orf9 | 8300               | 7620               | 229                       | Phage transcriptional repressor (Φm46.1 from S. pyogenes MGAS10394) |  |  |
| orf10 | 8568               | 8413               | 51                        | No significant homology found |  |  |
| orf11 | 8597              | 9250               | 217                       | No significant homology found |  |  |
| orf12 | 9333              | 10421              | 362                       | No significant homology found |  |  |
| orf13 | 10756             | 11325              | 189                       | Hypothetical protein (Φm46.1 from S. pyogenes MGAS10394) |  |  |
| orf14 | 13328             | 11363              | 651                       | DNA polymerase |  |  |
| orf15 | 15334             | 13367              | 55                        | Hypothetical protein (Φm46.1 from S. pyogenes MGAS10394) |  |  |
| orf16 | 14125             | 13553              | 190                       | Hypothetical protein (Φm46.1 from S. pyogenes MGAS10394) | AAR32303.1 | 90 (96) |
| orf17 | 15308             | 14106              | 400                       | Hypothetical protein (Φm46.1 from S. pyogenes MGAS10394) |  |  |
| orf18 | 15543             | 15220              | 107                       | Hypothetical protein (Φm46.1 from S. pyogenes MGAS10394) |  |  |
| orf19 | 15930             | 17762              | 610                       | Phage-associated DNA primase |  |  |
| orf20 | 17857             | 18186              | 109                       | No significant homology found |  |  |
| orf21 | 18506             | 18787              | 93                        | Hypothetical protein (Φm46.1 from S. pyogenes MGAS10394) |  |  |
| orf22 | 17878             | 20144              | 458                       | DNA helicase (Φm46.1 from S. pyogenes MGAS10394) |  |  |
| orf23 | 20137             | 20613              | 158                       | Hypothetical protein (Φm46.1 from S. pyogenes MGAS10394) |  |  |
| orf24 | 20775             | 20990              | 71                        | Hypothetical protein (Φm46.1 from S. pyogenes MGAS10394) |  |  |
| orf25 | 21038             | 22075              | 345                       | S-Adenosylmethionine synthetase |  |  |
| orf26 | 22108             | 22452              | 114                       | S-Adenosylmethionine synthetase (Φm46.1 from S. pyogenes MGAS10394) |  |  |
| orf27* | 22539            | 22877              | 112                       | Phage endonuclease (Φm46.1 from S. pyogenes MGAS10394) |  |  |
| orf28 | 23176             | 23637              | 153                       | Hypothetical protein (Clostridium thermocellum ATCC 27405) |  |  |
| orf29 | 23537             | 24865              | 442                       | DNA modification methylase |  |  |
| orf30 | 24862             | 26155              | 417                       | DNA methylase (Φm46.1 from S. agalactiae A909) |  |  |
| orf31 | 26183             | 26464              | 93                        | Methytransferase C-5 (Φm46.1 from S. agalactiae A909) |  |  |
| orf32 | 26902             | 26501              | 133                       |  |  |
| orf33 | 26993             | 27430              | 145                       | Phage terminase, small subunit (Φm46.1 from S. agalactiae CJB111) |  |  |
| orf34 | 27427             | 29019              | 530                       | Phage terminase, large subunit (Φm46.1 from S. agalactiae CJB111) |  |  |
| orf35 | 29590             | 29347              | 85                        | Phage terminase, large subunit (Φm46.1 from S. agalactiae CJB111) |  |  |
| orf36 | 29344             | 29715              | 123                       | Phage terminase, large subunit (Φm46.1 from S. agalactiae CJB111) |  |  |
| orf37 | 29856             | 31139              | 427                       | Phage portal protein |  |  |
| orf38 | 31132             | 31551              | 139                       | Cbp protease (Φm46.1 from S. agalactiae CJB111) |  |  |
| orf39 | 31643             | 31843              | 66                        | Phage capsid |  |  |
| orf40 | 31848             | 33053              | 401                       | Phage capsid |  |  |
| orf41 | 33056             | 33313              | 85                        | Hypothetical protein (Φm46.1 from S. agalactiae CJB111) |  |  |
| orf42 | 33313             | 33651              | 112                       | Phage head-tail adaptor |  |  |
| orf43 | 33644             | 34012              | 122                       | Phage head-tail adaptor (Φm46.1 from S. agalactiae CJB111) |  |  |

Continued on following page
described above highlights an intriguing correspondence between the short area of homology and the one (222 bp) sequence displayed 91% homology with a sequence at the orfC region of the orf53 gene (from approximately sequence m46. It is worth noting that inverse PCR assays were needed to first demonstrated (21) and further investigated (11) in strain tet (orf60, respectively, in the MGAS2096 genome.

(iv) ICE 2096-RD.2 fragment (49,017 to 52,098 bp). A ~3-kb fragment highly homologous (98%) to a region of 2096-RD.2, a 63-kb ICE-like element of S. pyogenes MGAS2096 harboring several antibiotic resistance genes (8), was found immediately downstream of the two site-specific recombinase-encoding ORFs. This region, which in the MGAS2096 genome spans from bases 1,101,217 to 1,103,555 (EMBL accession no. CP0000261), in the Em46.1 includes three ORFs, orf58, orf59, and orf60, corresponding to Spy1150, Spy1149, and Spy1148, respectively, in the MGAS2096 genome. orf59 was the tetracycline resistance gene tet(O), whose presence in S. pyogenes we first demonstrated (21) and further investigated (11) in strain m46. It is worth noting that inverse PCR assays were needed to sequence the unknown DNA region upstream of the tet(O) gene (from approximately orf53 to orf58).

(v) The last three ORFs (52,351 to 55,079 bp). The final region of the Em46.1 pro phage included the last three ORFs (orf61, orf62, and orf63), previously designated orfA, orfB, and orfC, respectively (11). Immediately downstream of orf63, a 76-bp sequence displayed 91% homology with a sequence at the right end of the pneumococcal mega element. Consideration of this short area of homology and the one (222 bp) described above highlights an intriguing correspondence between the ends of the Em46.1 pro phage and those of the mega element.

Comparative phage genomics. BLASTP analysis revealed that the proteins of Em46.1 had very high levels of amino acid sequence similarity to the amino acid sequences of proteins from other prophages, namely, 10394.4 of S. pyogenes and λSa04 of S. agalactiae. The two phages have quite a different prominence in the literature. 10394.4, detected in the genome of S. pyogenes MGAS10394 (EMBL accession no. CP000003), has been the subject of extensive specific investigations (5, 6, 11, 19, 20, 22) and is also found in GenBank under a separate accession number (GenBank accession no. AY445042). In contrast, λSa04, detected in the genome of S. agalactiae A909, is mentioned only in the deposited sequence alignment of the ORF map of Em46.1 and those of the mega element.
(i) Φm46.1 and Φ10394.4 prophage comparison. Φm46.1 and Φ10394.4 shared three major areas of homology, namely (from left to right), two clusters of closely related genes and a third cluster of moderately related genes. The first two clusters were separated by an area of nonhomology represented in Φ10394.4 by four ORFs, i.e., a restriction-modification cassette (orf16 to orf18) and orf19 (19), and in Φm46.1 by three unrelated ORFs (orf10 to orf12) (3). It is worth noting that while the DNA of tetracycline-resistant M-phenotype isolates, which typically carry Φm46.1 or a related tet(O)-mef(A) element, is usually digested by SmaI, tetracycline-susceptible M-phenotype isolates, which typically carry Φ10394.4 or Tn1207.3, are SmaI nontypable because a DNA-modifying methyltransferase encoded by thespyIM gene (orf16 in Φ10394.4) acts on the SmaI recognition sequence and makes the DNA refractory to cleavage by SmaI (3, 19, 20).

In Φm46.1, the first cluster included part of the Tn1207.1-like transposon (orf2 to orf7) and the left portion of the lysogeny module (orf8 and orf9). In Φ10394.4, these ORFs corresponded to orf8 to orf15; the levels of amino acid identity were very high (mostly >90%). In Φm46.1 the second cluster spanned from orf13 to orf27, including the DNA replication module and the beginning of the DNA modification module. In Φ10394.4, these ORFs corresponded to orf20 to orf33; the levels of amino acid identity were >90% for most correlated ORFs. The third cluster of Φm46.1 spanned from orf29 to orf46, including part of the DNA modification module, the DNA packaging and head morphogenesis module, and the beginning of the tail morphogenesis module. In Φ10394.4, these ORFs corresponded to orf35 to orf51; the levels of amino acid identity were lower (<90%) than those for the first two clusters.

(ii) Φm46.1 and λSa04 prophage comparison. Φm46.1 and λSa04 shared three major areas of homology, namely (from left to right), a first cluster of moderately related genes, a second cluster of closely related genes, and a third cluster (in fact, a couple of ORFs) of moderately related genes.

In Φm46.1 the first cluster spanned from orf13 to orf27, corresponding in λSa04 to orf1 to orf14; the levels of amino acid identity were <90% for most correlated ORFs. The second cluster of Φm46.1 spanned from orf29 to orf46, corresponding in λSa04 to orf15 to orf31; the levels of amino acid identity were >90% for most correlated ORFs. Interestingly, in both Φm46.1 (orf35 and orf36) and λSa04 (orf20 and orf21), this second cluster of genes included two putative TA-related genes that were lacking in Φ10394.4. It is remarkable that in the two bacteriophages the two couples of TA-related genes, even though they encoded proteins with no significant amino acid sequence identities, were found in the same position, i.e., immediately downstream of an ORF (orf34 in Φm46.1, orf19 in λSa04) encoding a large terminase subunit. In Φm46.1, the third area of homology was represented by orf56 and orf57 (encoding the two site-specific recombinases). In λSa04, these two ORFs corresponded to orf40 and orf41, with the levels of amino acid identity being 70% and 76%, respectively.

Phage ultrastructure. After induction with mitomycin C, electron microscopic analysis revealed phage particles with the typical icosahedral head and tail morphology of the Siphoviridae (Fig. 3), the most common phage family in streptococci (1). That the phage particles were indeed Φm46.1 is consistent with the findings of pulsed-field gel electrophoresis experiments, which demonstrated that it is the only prophage carried by S. pyogenes m46 that is inducible by mitomycin C (data not shown). On the other hand, the modular organization of the Φm46.1 genome was that typical of tailed phages (14), and a similar ultrastructure has been reported for Φ10394.4 (6).

Concluding remarks. Φm46.1, whose complete sequence analysis and final characterization were the aim of this study, is the recognized representative of the most common variant of the so-called tet(O)-mef(A) elements, responsible for efflux-mediated erythromycin resistance in tetracycline-resistant S. pyogenes isolates (35). In an early study (21), genes mef(A) and tet(O) were detected in S. pyogenes m46, were cotransferred to a susceptible recipient of the same species, and were found to be linked, with mef(A) being detected ~5.5 kb downstream of tet(O); a single new DNA insertion into the transconjugants with the m(A) tet(O) genotype was consistent with a chromosomal location of the two genes. Subsequent investigations (11) demonstrated a variety of closely related tet(O)-mef(A) elements harboring a range of changeable and defective variants of Tn1207.1, of which the element detected and originally investigated in strain m46 was the most common. Mitomycin C induction experiments showed that the tet(O)-mef(A) elements were in fact prophages (22). The present study has conclusively clarified that the chromosomal integration site is within the 23S rRNA uracil methyltransferase gene (near its 3’ end). In the host cell, Φm46.1 exists not only as a prophage but also as free circular DNA. While in the latter form tet(O) is found ~5.5 kb upstream of mef(A), in the integrated form it is close to the right end of the prophage (orf59 of 63 ORFs), ~46.3 kb downstream of mef(A), which is close to the left end of the prophage (the third ORF). Accordingly, the designation “mef(A)-tet(O) elements” would be more appropriate than the one “tet(O)-mef(A) elements” (11, 22, 35), which has so far been used to indicate these genetic elements.

It is well established that each prophage is a unique entity that not only shares blocks of sequences with different prophages but that also possesses unique sequences with no known homologies in current databases. This genetic mosaicism is a
hallmark of tailed phages and reflects an unusually high degree of horizontal genetic exchange in phage evolution (23, 24). Genome mosaicism is characterized by the presence of novel sequence joints, in which the similarity between two phages abruptly ceases; other phages and bacterial hosts may be the sources for such new sequences (14). Φm46.1 and Φ10394.4, whose genomes are highly mosaic, appear to fit well into the general model. One explanation proposed for the origin of Φ10394.4 was that an erythromycin-susceptible S. pyogenes precursor strain containing a phage might have acquired the mef(A)-containing Tn1071.3 transposon, possibly from other Streptococcus species present in the upper respiratory tract (6). The latter surmise is supported by the short homologous sequences shared by the Φm46.1 prophage and the pneumococcal mega element at both their left and right ends. Starting from this mef(A)-carrying ancestor, the Φm46.1 and Φ10394.4 genomes may have diversified by independently exchanging genetic information and shuffling genetic modules. Major examples of such diversification (in Φm46.1 compared to Φ10394.4) are the lack of the initial 6-kb left-hand region, also lacking in Tn1071.3 orf1, in place of the initial portion of Tn1071.1; three new ORFs (orf10 to orf12) in place of the restriction-modification cassette; a putative TA system (orf35-orf36); the host cell lysis module (orf53 to orf55), shared by and possibly acquired from ICE Sde3396; the tet(O)-containing fragment (orf58 to orf60), shared by and possibly acquired from ICE 2096-RD.2; and the last three ORFs (orf61 to orf63), likely of chromosomal origin. In particular, as far as the two antibiotic resistance genes carried by Φm46.1 are concerned, the present findings support the hypothesis of the stepwise acquisition of mef(A) and tet(O) (11).

These data are consistent with the current belief in the key role of bacteriophages in the evolution of important bacterial pathogens: on the one hand, by carrying a versatile range of genetic material, contributing to the extraordinary horizontal transfer of mobile genetic elements in this species. Such a formerly underestimated role of phages in clonal diversification has even led some investigators to question the validity of the conventional associations of certain M serotypes with unique combinations (10). This is particularly true of S. pyogenes, a species distinguished by a unique propensity to acquire and reshuffle phage-encoded virulence and resistance determinants (4). The highly mosaic genome of Φm46.1, in which different segments are related to distinct streptococcal phages, entails that phages of S. pyogenes continue to exchange genetic material, contributing to the extraordinary horizontal transfer of mobile genetic elements in this species. Such a formerly underestimated role of phages in clonal diversification has even led some investigators to question the validity of the conventional associations of certain M serotypes with specific clinical manifestations of S. pyogenes infections and to suggest the need for a new classification scheme that better represents the genetic bases of S. pyogenes virulence (2).

ACKNOWLEDGMENT

This work was partly supported by the Italian Ministry of Education, University and Research.

REFERENCES

1. Ackermann, H. W. 2003. Bacteriophage observations and evolution. Res. Microbiol. 154:245–251.
2. Aizai, R. K., R. A. Edwards, W. W. Taylor, D. E. Low, A. McGeer, and M. Koth, 2005. Mosaic phages with horizontally acquired genes account for the emergence and diversification of the globally disseminated M1T1 clone of Streptococcus pyogenes. J. Bacteriol. 187:3311–3318.
3. Bacciaglia, A., A. Brenciani, P. E. Varaldo, and E. Giovanetti. 2007. Small typeable and tetracycline susceptibility and resistance in Streptococcus pyogenes isolates with efflux-mediated erythromycin resistance. Antimicrob. Agents Chemother. 51:3042–3044.
4. Banks, D. J., S. B. Beretta, K. J. M. Musser. 2002. The fundamental contribution of phages to GAS evolution, genome diversification and strain emergence. Trends Microbiol. 10:515–521.
5. Banks, D. J., S. F. Porcella, D. Barbian, S. B. Beres, L. E. Philips, J. M. Vojvich, F. R. DeLeo, J. M. Martin, G. A. Somerville, and J. M. Musser, 2004. Progress toward characterization of the group A Streptococcus metagenome: complete genome sequence of a macrolide-resistant serotype M6 strain. J. Infect. Dis. 190:727–738.
6. Banks, D. J., S. F. Porcella, D. Barbian, J. M. Martin, and J. M. Musser. 2003. Structure and distribution of an unusual chimeric genetic element encoding macrolide resistance in phylogenetically diverse clones of group A Streptococcus. J. Infect. Dis. 188:1898–1908.
7. Bensing, B. A., J. R. Silio, and P. M. Sullivan. 2001. Proteins PhbA and PhbB of Streptococcus mitis, which promote binding to human platelets, are encoded within a lysisogenic bacteriophage. Infect. Immun. 69:6186–6192.
8. Beres, S. B., and J. M. Musser. 2007. Contribution of exogenous genetic elements to the Group A Streptococcus metagenome. PLoS One 2:e8000.
9. Beres, S. B., E. W. Richter, M. J. Nagiec, P. Sumby, S. F. Porcella, F. R. DeLeo, and J. M. Musser. 2006. Molecular genetic anatomy of inter- and intraserotype variation in the human bacterial pathogen group A Streptococc. Proc. Natl. Acad. Sci. U. S. A. 103:7055–7060.
10. Boyer, F. G., and H. Brissow. 2002. Common themes among bacteriophage-encoded virulence factors and diversity among the bacteriophages involved. Trends Microbiol. 10:521–529.
11. Brenciani, A., K. K. Ojo, A. Monachetti, S. Menzo, M. C. Roberts, P. Varaldo, and E. Giovanetti. 2004. Distribution and molecular analysis of mef(A)-containing elements in tetracycline-susceptible and -resistant Streptococcus pyogenes clinical isolates with ε-fluorescent erythromycin resistance. J. Antimicrob. Chemother. 54:991–998.
12. Campbell, A. M. 1986. Bacteriophage lambda as a model system. Bioessays 8:277–280.
13. Campell, A. 2003. Prophage insertion sites. Res. Microbiol. 154:277–282.
14. Casjens, S. R. 2005. Comparative genomics and evolution of the tailed-bacteriophages. Curr. Opin. Microbiol. 8:451–458.
15. Reference deleted.
16. Clancy, J., J. Petitts, F. Dib-Hajji, W. Yuan, M. Cronan, A. Kamath, J. Bergeron, and J. A. Retsema. 1996. Molecular cloning and functional analysis of a novel macrolide resistance determinant, mefA, from Streptococcus pyogenes. Mol. Microbiol. 22:867–879.
17. Davies, M. R., J. Shera, G. H. Van Domselaar, K. S. Sripirakasch, and D. J. McMillan. 2009. A novel integrative conjugative element mediates genetic transfer from group G Streptococcus to other β-haemolytic streptococci. J. Bacteriol. 191:2257–2265.
18. Reference deleted.
19. Euler, C. W., P. A. Ryan, J. M. Martin, and V. A. Fischetti. 2007. MSpIL, a DNA methyltransferase encoded on a mefA chimeric element, modifies the genome of Streptococcus pyogenes. J. Bacteriol. 189:1044–1054.
20. Figueiredo, T. A., S. I. Aguilar, J. Melo-Cristino, and M. Ramirez. 2006. DNA methylase activity as a marker for the presence of a family of phage-like elements conferring ε-fluorescent macrolide resistance in strepto- cocci. Antimicrob. Agents Chemother. 49:3639–3649.
21. Giovanetti, E., A. Brenciani, R. Lupidi, M. C. Roberts, and P. E. Varaldo. 2003. Presence of the tet(O) gene in erythromycin- and tetracycline-resistant strains of Streptococcus pyogenes and linkage with either the mef(A) or the erm(A) gene. Antimicrob. Agents Chemother. 47:2844–2849.
22. Giovanetti, E., A. Brenciani, M. Vecchi, A. Manzin, and P. E. Varaldo. 2005. Prophage association of mef(A) elements encoding ε-fluorescent erythromycin resistance in Streptococcus pyogenes. J. Antimicrob. Chemother. 55:445–451.
23. Hatfull, G. F. 2008. Bacteriophage genomics. Curr. Opin. Microbiol. 11:447–453.
24. Hendrix, R. W. 2003. Bacteriophage genomics. Curr. Opin. Microbiol. 6:506–511.
25. Lucchini, S., F. Desiere, and H. Brüssow. 1999. Similarly organized lysogenic modules in temperate Siphophilaion from low GC content gram-positive bacteria. Virology 263:427–435.
26. Magnuson, R. D. 2007. Hypothetical functions of toxin-antitoxin systems. J. Bacteriol. 189:6089–6092.
27. Olsvik, B., I. Olsen, and F. C. Tenover. 1995. Detection of tet(M) and tet(O) using the polymerase chain reaction in bacteria isolated from patients with periodontal disease. Oral Microbiol. Immunol. 10:87–92.
28. Sambrook, J., and D. W. Russell. 2001. Inverse PCR, p. 8.81–8.85. In Molecular cloning: a laboratory manual, 3rd ed. Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY.
29. Santagati, M., F. Ioannides, F. C. Cascone, F. Campanile, M. R. Oggoni, S. Stefani, and G. Pozzi. 2003. The novel conjugative transposon Tn1207.3 carries the macrolide ε-lux gene mef(A) in Streptococcus pyogenes. Microb. Drug Resist. 9:243–247.
30. Santagati, M., F. Iannelli, M. R. Oggioni, S. Stefani, and G. Pozzi. 2000. Characterization of a genetic element carrying the macrolide efflux gene \textit{mef}(A) in \textit{Streptococcus pneumoniae}. Antimicrob. Agents Chemother. 44: 2585–2587.

31. Sutcliffe, J., T. Grebe, A. Tait-Kamradt, and L. Wondrack. 1996. Detection of erythromycin-resistant determinants by PCR. Antimicrob. Agents Chemother. 40:2562–2566.

32. Sutcliffe, J., A. Tait-Kamradt, and L. Wondrack. 1996. \textit{Streptococcus pneumoniae} and \textit{Streptococcus pyogenes} resistant to macrolides but sensitive to clindamycin: a common resistance pattern mediated by an efflux system. Antimicrob. Agents Chemother. 40:1817–1824.

33. Tettelin, H., V. Masignani, M. J. Cieslewicz, C. Donati, D. Medini, N. L. Ward, S. V. Angiuoli, J. Crabtree, A. L. Jones, A. S. Durkin, R. T. Deboy, T. M. Davidsen, M. Mora, M. Scarselli, I. Margarit y Ros, J. D. Peterson, C. R. Hauser, J. P. Sundaram, W. C. Nelson, R. Madupu, L. M. Brinkac, R. J. Dodson, M. J. Rosovitz, S. A. Sullivan, S. C. Daugherty, D. H. Haft, J. Selengut, M. L. Gwinn, L. Zhou, N. Zafar, H. Khouri, D. Radune, G. Dimitrov, K. Watkins, K. J. O’Connor, S. Smith, T. R. Utterback, O. White, C. E. Rubens, G. Grandi, L. C. Madoff, D. L. Kasper, J. L. Telford, M. R. Wessels, R. Rappuoli, and C. M. Fraser. 2005. Genome analysis of multiple pathogenic isolates of \textit{Streptococcus agalactiae}: implications for the microbial “pan-genome.” Proc. Natl. Acad. Sci. U. S. A. 102:13950–13955.

34. Varaldo, P. E., E. A. Debbia, G. Nicoletti, D. Pavesio, S. Ripa, G. C. Schito, G. Tempera, and the Artemis-Italy Study Group. 1999. Nationwide survey in Italy of treatment of \textit{Streptococcus pyogenes} pharyngitis in children: influence of macrolide resistance on clinical and microbiological outcomes. Clin. Infect. Dis. 29:869–873.

35. Varaldo, P. E., M. P. Montanari, and E. Giovanetti. 2009. Genetic elements responsible for erythromycin resistance in streptococci. Antimicrob. Agents Chemother. 53:343–353.

36. Weisblum, B. 2000. Resistance to the macrolide-lincosamide-streptogramin antibiotics, p. 694–710. In V. A. Fischetti, R. P. Novick, J. J. Ferretti, D. A. Portnoy, and J. I. Rood (ed.), Gram-positive pathogens. American Society for Microbiology, Washington, DC.