Mapping of replication initiation site in *Mycoplasma capricolum* genome by two-dimensional gel-electrophoretic analysis

Makoto Miyata, Ken-Ichi Sano, Ryosuke Okada and Takashi Fukumura

Department of Biology, Faculty of Science, Osaka City University, Sumiyoshi-ku, Osaka 558, Japan

Received May 28, 1993; Revised and Accepted August 17, 1993

DDBJ accession nos D14982 and D14983

ABSTRACT

The homolog of the *dnaA* gene, which has been reported to be present in the vicinity of the initiation site of replication in the genome of *Mycoplasma capricolum* (M.Miyata, L.Wang, and T.Fukumura, *J. Bacteriol.* 175: 655 – 660, 1993) was mapped precisely. A 9540-bp region containing the *dnaA* gene was cloned and the entire region was sequenced with the exception of a previously reported region of 2517 bp (Fujita, M.Q., Yoshikawa, H. and Ogasawara, N. *Gene* 93: 73 – 78, 1992). The organization of the 9540-bp region was compared with that of corresponding regions in other bacteria. The arrangement and directions of *rnpA*, *rpmH*, *dnaA*, *dnaN* were conserved, but no other open reading frames were found that were homologous to those that are commonly found around *dnaA* genes in other bacteria. The directions of movement of the replication fork around the *dnaA* gene were analyzed by neutral/alkaline two-dimensional gel electrophoresis. The forks developed in a 1569-bp region that consisted of the *dnaA* structural gene and its downstream non-coding region, and then they proceeded bidirectionally.

INTRODUCTION

Mycoplasmas are the simplest, wall-less, self-replicating, single-celled organisms. They consist of only a minimum set of the sub-cellular structures that are required for growth and reproduction (1), and their genomes are extremely small (2). Mycoplasmas are, therefore, attractive for use in the study of essential cellular mechanisms, and various molecular biological studies of these organisms have been reported (3,4). However, little is known about the replication of the mycoplasmal genome (5).

In studies of bacteria, the identification of the origin region that contains the site for initiation of replication has provided much valuable information about genome replication (6,7,8,9,10). Such origin regions include controlling motifs such as the DnaA protein-binding sequence (DnaA-box), a repeated 13-meric sequence, an AT-rich region, and a deoxyadenosine methylation site (6). In some bacteria, but not in enteric bacteria, *dnaA* genes, the products of which control initiation, can be found in the vicinity of the origin regions. The genetic organization around the *dnaA* genes is highly conserved and genes are

organized in the following order, *gyrB-recF-dnaN-dnaA-rpmH-rnpA*, in *Escherichia coli*, *Bacillus subtilis*, and *Pseudomonas putida* (11).

A few studies designed to identify the origin regions of mycoplasma genomes have been reported (9,13,14). Fujita *et al.* found the homolog of the *dnaA* gene in *M.capricolum* and showed that the *dnaN*, *dnaA* and *rpmH* genes are conserved and arranged in that order (13). We previously localized the *dnaA* gene and the origin region within the 46-kb *Bam*HI fragment (Bm8) that is bordered by two of the total of nine *Bam*HI sites on the physical map (14,15). However, clarification of the genetic organization of the *dnaA* flanking regions and precise mapping of the origin region have not yet been achieved.

Neutral/alkaline two-dimensional (2D) gel-electrophoretic analysis was developed as a method for obtaining crucial information about the direction of movement of replication forks (16), and it has been applied to chromosomal replication in eukaryotes but not yet in prokaryotes.

In this study, we clarified the genetic organization of the flanking regions of the *dnaA* gene in the genome of *M.capricolum*, and we mapped the initiation site by two-dimensional gel-electrophoretic (2D-gel) analysis.

MATERIALS AND METHODS

Strains and plasmids

M.capricolum ATCC 27343 was grown in modified Edward medium (MEM) (17) at 37°C for preparation of DNA for mapping and cloning. For preparation of replicating intermediates, cultivation was carried out at 39°C to allow frequent formation of replicating intermediates at a growth rate 20% higher than that at 37°C. The host strain for plasmids was *E.coli* XL1-Blue. pUC118/119 (20) were used for cloning of fragments of the *M.capricolum* genome. pKS(+)-BluescriptII (20) was also occasionally used for making deletion mutants. A derivative of M13 harboring the *dnaA* homolog from the *M.capricolum* genome (mpDD2.1) was kindly supplied by Dr M.Q.Fujita of Osaka University Medical School (13).

Preparation and cloning of DNA

Genomic DNA was prepared by the agarose-block method (15) for field-inversion gel-electrophoresis (FIGE) and by the phenol method (18) for other purposes. The Bm8 fragment was isolated

by FIGE as described previously (15), and recovered by use of GeneCleanII (Bio101). For the polymerase chain reaction (PCR), primer 2 (GTATAATCAAAATACTTACTGGC) and primer 4 (GGAAGGGGAGCAAATGAACCAG) were designed on the basis of the sequences of p8SU and p9RUL, respectively. Primer 1 (CATGGTTTTAGAGCAAGAATGGCTAC) and primer 3 (GCTATGTGTCTTGCTGTTAC) were designed on the basis of previously reported sequences (13). The primers were synthesized chemically with an automated synthesizer (Gene Assembler Plus; Pharmacia). The reaction mixture for PCR was composed of 25 μ M primers, 1 μ g/ml DNA, all four deoxynucleotide triphosphates at 0.2 mM each, 50 mM KCl, 10 mM Tris-HCl (pH 8.3), 1.5 mM MgCl₂, 0.01% gelatin and 0.025 unit/ μ l Taq polymerase (19). The mixture was subjected to 30 cycles of a reaction that consisted of a 1-min denaturation period at 95°C, a 1-min annealing period at 42°C, and extension periods of appropriate duration at 72°C. The extension period was set at about 1 min per kilobase of desired fragment-length. Other procedures for cloning were performed by following published protocols (20).

Preparation of templates for probes

p9dnAHP and p9dnAPHD were prepared from pUC119 and a mycoplasma fragment of mpDD2.1 using a *Pvu*II site located in the *dnaA* gene (13). Templates for mapping were prepared by enzymatic digestion and purification on agarose gels. Templates for probes 2, 3, 4 and 7 were prepared by subcloning with appropriate restriction sites. Templates for probes 1 and 6 were prepared by PCR which amplified the end regions of 619 bp and 570 bp in length of clones p9RE and p9RUL, respectively. The sequences of primer 5 (GTTGGAAAGATAACGTACAGYC) and primer 6 (GGCGGAACTAATGTTTCTTATCG), which were used as probes 1 and 6, were designed on the basis of the sequences of nucleotides (nt) 598–619 and nt 7155–7177, respectively. Nucleotide number one was taken as the first base of the left *Cla*I site. The template for probe 5 was amplified from genomic DNA by PCR with primers 3 and 4, digested by *Dde*I, and purified by agarose gel electrophoresis.

Southern hybridization

Southern transfer (21) was performed by the vacuum method (22) using nylon filters (Biodyne A; Pall). The DNA was cross-linked to the filter by UV irradiation at 0.15 J/cm² and 312 nm (20) after transfer from the gel. The DNA probes were labeled by the random primer method (23) with [α -³²P]-dATP for the 2D-gel analysis, or with digoxigenin-linked dUTP (DIG-dUTP; Boehringer) for other purposes. The specific activity of radiolabeled probes was about 2×10^8 cpm/ μ g. Hybridization (24) was carried out in a rolling bottle for 16 h. The concentration of the radiolabeled probe in the hybridizing solution was adjusted to 2.5×10^6 cpm/ml. Dehybridization was carried out as described elsewhere (25). The filters with radiolabeled probes were exposed to a Fuji imaging-plate for analysis with a Bio Image-Analyzer (BAS 2000; Fuji Photo Film). The filters with DIG-labeled probes were treated according to the protocol of Southern Light (Tropix), and exposed to X-ray film.

Sequencing

Nested deletion mutants were made with exonuclease III and mung bean nuclease (20). The cloned fragments of p9RE, p9RUM, p9RUL and p8SU were occasionally subcloned into

pUC119 or pKS(+)-BSII. The polymerase reaction was carried out with denatured double-stranded plasmids by the dideoxy method (26) with T7 bacteriophage polymerase (Pharmacia) or BcaBEST (Takara) DNA polymerase. The entire sequence was determined at least once for each strand. The nucleotide and amino acid sequences were analyzed with GENETYX Ver 8 (Software Development Co., Tokyo, Japan). Protein sequences were compared with all data in the Swiss-Prot (Release 23) data base shipped in a compact disk issued in January 1993.

Preparation of replication intermediates

The replication intermediates were prepared as described previously (27) with slight modifications. An exponential-phase culture (100 ml; OD_{600nm}, 0.4) was chilled by the addition of an equal volume of ice-chilled solution A (Sol. A), which consisted of 20 mM Tris-HCl (pH 7.6), 0.25 M NaCl and 10 mM EDTA, and incubation on ice for 5 min. The cells were harvested by centrifugation at $13,000 \times g$ for 10 min and suspended in Sol. A. The cells were collected again and resuspended in 20 ml of Sol. A. The suspension of cells was brought to 37°C and mixed with an equal volume of a 1% solution of molten low-melting-point agarose in Sol. A. The mixture was added to 80 ml of liquid paraffin at 37°C and shaken vigorously by hand for 30 sec. The emulsion was poured into 400 ml of ice-cold Sol. A and the mixture was stirred for several minutes on ice. Agarose beads were then collected by centrifugation at $18,000 \times g$ for 10 min. Since some of agarose beads floated just beneath the liquid-paraffin interface after the centrifugation, the supernatant was shaken vigorously and again subjected to centrifugation. Most of the beads remaining in the supernatant after the first centrifugation were recovered by the second centrifugation. The beads were combined, washed twice with Sol. A, suspended in 100 ml of 1% SDS in 25 mM EDTA, and then shaken at 37°C for 10 min. The beads were collected by centrifugation and suspended in 100 ml of Sol. B, which consisted of 1% Sarkosyl (potassium salt), 0.5 M Na-EDTA and 10 mM Tris (pH 9.5), and then shaken at 37°C for 10 min. The beads were again collected by the centrifugation, resuspended in 20 ml of Sol. B, and treated with 0.3 mg/ml proteinase K at 37°C for 60 min with shaking. The agarose beads were washed once with TE that contained 0.1 mM phenylmethylsulfonyl fluoride (PMSF) and then twice with TE, and then they were suspended in 5 ml of TE. The beads were stored at -20°C and melted just before use. The stability of the intermediates at -20°C was checked by monitoring the intensity of the signal from the hybridization image after 2D gel electrophoresis.

Agarose beads containing DNA from a 30-ml culture were treated with 500 units of *Pvu*II and *Xba*I in high buffer at 37°C for 1 h. The beads were washed once with 0.5 $\times$ TBE that contained 0.5% SDS, and twice with 0.5 $\times$ TBE. The DNA in the beads was recovered in a dialysis bag by electroelution at 5 V/cm in 0.5 $\times$ TBE buffer for 2 h at 4°C. After separation of the extracted beads from the solution of DNA by centrifugation at $16,000 \times g$ for 5 min, the beads were washed with 0.5 $\times$ TBE. The solutions of DNA obtained by the two centrifugations were combined and dialyzed against 0.8 M NaCl, 10 mM Tris (pH 8.0) and 1 mM EDTA. Selective adsorption of single-stranded DNA to benzoylated naphthoylated DEAE-cellulose (BND-cellulose; Sigma) was carried out as described previously (28). The DNA was eluted with caffeine buffer that consisted of 1.8% caffeine, 1 M NaCl, 10 mM Tris-HCl (pH 8.1) and 1 mM

EDTA and was then precipitated by addition of isopropanol. The precipitated DNA was washed with 70% ethanol, dissolved in 20 μ l of water, and loaded onto a lane of the first-dimension gel.

2D gel electrophoresis

2D gel electrophoresis was carried out essentially by the method of Huberman *et al.* (16). The DNA was loaded onto a 0.4% agarose gel, and electrophoresis was performed at 1.5 V/cm in TAE buffer for 12 h at 4°C. The lane containing the DNA was excised, placed at the top of a 1.5% agarose gel and allowed to equilibrate with 50 mM NaOH, 1 mM EDTA for 2 h. Electrophoresis in the second dimension was performed at 1 V/cm in 50 mM NaOH, 1 mM EDTA for 20 h at 4°C. The gel was soaked in TAE buffer that contained 0.5 μ g/ml ethidium bromide, and then it was treated in the standard manner for Southern transfer (21).

RESULTS

Mapping of Bm8

It has been reported that there is a *Xho*I site near the *Bam*HI site that separates Bm8 and Bm4 (15). We isolated a 0.8-kb fragment by electrophoresis of genomic DNA that had been digested by *Xho*I and *Bam*HI, and inserted it into pUC118 to obtain clone p8XB0.8. The fragment in p8XB0.8 was localized at the right end of Bm8, as shown in Fig. 1, by hybridization analysis. *Bam*HI-digested genomic DNA was partially digested with *Eco*RI or *Pvu*II, fractionated on an agarose gel, and analyzed by hybridization with p8XB0.8 as probe. The positions of *Eco*RI and *Pvu*II sites were located on the Bm8 map (Fig. 1, upper line). A filter containing a *Pvu*II digest was probed with p9dnAHP and with p9dnAPHD. The results revealed the position and direction of the *dnaA* homolog. *Pvu*II digests were partially digested by *Hind*III, *Bgl*III, or *Xba*I, and each partial digest was subjected to hybridization analysis with p9dnAHP or p9dnAPHD as probe. A map of the 25-kb region bordered by two *Pvu*II sites was constructed (Fig. 1, lower line). The map was modified by reference to results of hybridization analysis with p826, p8NU, or p8SU (see below) as probes. The positions of *Cla*I and *Eco*RV sites were determined by hybridization analysis with the five probes.

Figure 1. Physical map of Bm8 and location of cloned fragments. Upper line: physical map of Bm8. The total length is 46 kb. The hatched square shows the position of p8XB0.8. Lower line: Magnification of the 25-kb region bordered by two *Pvu*II sites. The thick arrow shows the position and the direction of the *dnaA* gene. The open and hatched squares represent the clones used in this study. The clones indicated by hatched squares were used for mapping as probes. The small arrows indicate the primers for PCR. B, *Bam*HI; P, *Pvu*II; A, *Apa*I; E, *Eco*RI; Xh, *Xho*I; H, *Hind*III; X, *Xba*I; G, *Bgl*III; V, *Eco*RV; C, *Cla*I.

Cloning of the 9.5-kb region bordered by *Cla*I and *Bgl*III sites

The Bm8 fragment was digested by *Hind*III plus *Bgl*III, and the digested fragments were ligated to pUC118. Primary screening was performed with Southern hybridization using the Bm8 probe. The second screening was done by hybridization of probes prepared against the cloned fragments to the 9 bands of a *Bam*HI digest that had been separated by FIGE. Five separate clones were found to be derived from Bm8. Finally, 3 clones, namely, p826, p8NU and p8SU, were localized in the *Pvu*II-bordered 25-kb region (Fig. 1, lower line) by further hybridization analysis. p8SU had a *Hind*III site in its insert. The position and direction of the *Hind*III-bordered, small fragment of p8SU on the genome were confirmed by hybridization and PCR. Clone p9RE was isolated by genome walking using *Hind*III and *Cla*I sites and p9dnAHP as probe.

The other clones, namely, p9RUM and p9RUL, were isolated by PCR as follows. PCR for p9RUL and p9RUM was carried out with primers 1 and 2 and with primers 3 and 4, respectively. We inserted the products of PCR into pUC119 using *Hind*III sites and a *Hind*III site plus an *Nsp*I site, and we obtained clones p9RUL and p9RUM, respectively. The positions of p9RUM, p9RUL and p9RE on the genome were confirmed by hybridization analysis.

Sequencing of the 9.5-kb region bordered by *Cla*I and *Bgl*III sites

The sequence of the entire 9540 bp, apart from the sequence previously reported by Fujita *et al.* (13), was determined. The regions for which we determined the sequences in this study are shown as by the two solid portions of the restriction map in Fig. 2. The sequences of the C-terminal regions of *dnaN* (Fig. 3,B) and *rpmH* (Fig. 3,C) were determined. The extent of homology when *dnaN* and *rpmH* were compared to the corresponding genes (29,30) of *B. subtilis* was 24.3% and 61.4%, respectively. Eight additional ORFs were found in this region (Fig. 2). The directions of ORFs downstream of *dnaA* were the same as that of *dnaA*. The directions of ORFs upstream of *dnaA* were opposite to that of *dnaA*. A homolog of *rpmA* was located in the region adjacent to *rpmH*. The extent of homology of its sequence

Figure 2. Gene organization and position of the initiation site of replication. The top thin line shows sequence positions, with the first base of the left *Cla*I site taken as the start (nt1). The physical map is shown by a thick line. The solid portions of the map indicate the regions sequenced in this study. The open reading frames (boxes) and the non-coding regions (solid lines) are drawn to scale. The 500 bp region extending from the 5' end of *rpmH* to *dnaA* is magnified. The solid triangles indicate the putative DnaA-boxes. The genes homologous to any genes in other bacteria are hatched. The directions of genes are indicated by broken arrows. The region that contains the initiation site is indicated by a thick bar just below the physical map. The solid arrows extending from the ends of the bar indicate the movements of forks. X, *Xba*I; H, *Hind*III; C, *Cla*I; P, *Pvu*II; N, *Nsp*I; G, *Bgl*III; V, *Eco*RV.

[A] *k s g A*

Bs 1 VTEKGVIEIGPGIGALTEQLAKRAKKVVAFEIDQRLLPILKDLSPYENVTVIHQDVLKADVKSVEEQFQDCDEIMVVANLPYYVTTPIMKLL EEHLPLKGI VVMLQKEVAERMAAD
 Mc 1 MVEILKTKFNB-SNLEIQADVLEIDLKQLI-SKY-DYKNISII SNTPYYITSEILFKTLQISDLLTKAVFMLQKEVALRICSN
 Bs 121 PSSKEYGSLSI AVQFYTEAKTVNI VPKTVFVPQPNVD SAVIRLILRDGPAVDVENESFFFQLIKASFQRRKTLN LNLYNNLP EGKAQKSTIEQVLEETNIDGKRRGESLSIEEFAALSN
 Mc 82 KNENNYNNLSIACQFYSQRNFEFVVNKKMFYPIPKVDSAIISLTFNNIYKKQINDDKKFI EFVRTLFNNKRTILNNL-NNIIQNKNKALEY LKML--NISSNLRPEQLDIDEYIKLFN
 Bs 241 GLYKALF
 Mc 198 LIYISNF

[B] $d n a N$

Bs 1 MKFTIQNDILTKNLKKITRVLVKNISFPPILENILIQV--EDGTLSLTTTNLEI--ELISKIEIITKYIPGKTTISGRKILNICRTLSEKSKIKMLKNKKMYISSSENSNYILSTLSADTF
 Mc 1 MNFSINRIVLLDNLСКАAKVIDYKNVNPSLSGIYLVNLNDQVNVITTSGLSFKSLNQNQSDLEVQKEGKVLKPKYVLEMLRRLDDEFVVFMSVEDNELIKTNNSDFSIGVLNSEDY
 Bs 117 PNHQNFYISKFDISSNLKEMIEKTEFSMGQKQDVRYLVNGLLEKKDKFLRSVATDGYRLAISYTLKKDINF-FSIIIPNKAYMELLKLLNTQPLLNLIGSNSIRIYTKNLIFTQ
 Mc 121 PLIGFREKGIENLNPKEVKKTIYQVFYSMNENKKLILTGLNLKLNNQAIFSTDSFRISQKILEIQSDNNEDIDITIPFKTALELPKLLD-NAENLKIIVEGYITFIIDNVIFQSN
 Bs 236 LIEGEYPDYKSVLFKEKKNPIT-NSILLKSKLLRVAILAHEKFCGI-EIKIENGKFKVLSDNQEEETAEDLF-EIDYFGEK-IEISINVYLLDVINNIKSENIALFLNKSQSSIQIEA
 Mc 240 LIDGKFPNVQ-IAFPTKFETIITVKQKSILKVLRSFDLVADDGLPAIVNIKVNEDKIEFKSFISEVGKYEEDFDDFVIEGNKSLSISFNTRFLIDAIKTLEDRIELKLINSTKPIVINN
 Bs 352 ENNSSNAVYVMLLR
 Mc 359 VYDEHLKQVILPTFISN

[C] $r p m H$

Bs 1 MKRTFQPNNRKRKSKVHGFRSRMSSKNGRLVLARRRRKGRKVL SA
 |||| | | |||| | | |||| | | |||| | |
Mc 1 MKRTWQPSKLKHARVHGFRARMATKNGRKYIKARRAKGRYRL SA

[D] *г д р а*

Bs 1 MSHLKKRNRLKKNEFDQVKFHGTSVANRQFVLYTLDPENDELRVGLSVSKKIGNAVMRNRIRKLI RQAFLEEKERL--KEKDYIIIRKPKASQLTYEETKKSLLQHLFRKSSLYKKSSSK
 Mc 1 MKNKRVIKKNFQEIINYYKTIKNFCFVIY-YKDNEESLYKGISVGKKIGNAVIRNKVKRQIRMLKQNISEIGTVSKDIIILVRSVYLEIKYATISKILIKIIFEIK

[E] *lica*

```

Hi 1  MQSINQSINQSINQSINQSINQSINQSINQSINQIVGFVKTCYKPEEVFHFLLHQHSIPFSSIGGWTNQVLLNISGVKFLRIPNAVNLSSLINREYEAFFNAQAYRAGLNVE
Mc 1                                     MKMKITKGGTNVSYRIDNTFLQIKNYNSFNHQINYTELLKDFD

Hi 121 -TPVLDAKSGVKLTR-YLENSNTLSQIQLNEQSCLSQVWNLYRLHNSEFVFRNVFSVFDEFQRQYFSLLENKSAFYQADSRMDKLSAVFWQFEEINKDIILRPNCHNDLPVENMLLQDDRL
Mc 43  FVPKLISNDQKEIVWEYVEGNEPV--VDLNNIK---AITNQIKQLHNSNLNFPK--NNLKQRVQYYR--QKMVELNSGIEIIDKYANLIDDI--LDKMDHSTPLHNDLFPFNMIETKNKI

Hi 239 FFIDWEYSGLNDPLFDIATIIIEAHLSEKAADFLETYCNQTKYHKTEFQIAHKRLKIHFRFCQNVLWFLWTKVKEEHGENFGDYALKRLDAAFKLEELP
Mc 152  YFVDWEYATMGDKHFLAYLIETSNMNSECEKVFLDLY---SDYD--SYKLLLNKIFVN-YIV-ILW-IRTQTSAPYNTTF--FEQKIINYVTKLTN

```

Figure 3. Comparison of amino acid sequences between homologous genes. The amino acid sequences of *M. capricolum*, shown by the single-letter codes, were deduced from the nucleotide sequences of the cloned genes. The previously reported sequences from *M. capricolum* (13) are underlined. The references for the sequences from *B. subtilis* and *H. influenzae* are cited in the text. The conserved amino acids are marked by vertical lines. The locations of deletions of counterpart amino acids, shown by discontinuous lines, were determined by aligning amino acids with similar properties as far as possible. Bs, *B. subtilis*; Hi, *H. influenzae*; Mc, *M. capricolum*.

to that of the gene from *B. subtilis* (30) was 36.4% (Fig. 3,D). One of the ORFs exhibited 28.8% and 35.1% homology to *ksgA*s [S-adenosyl methionine-6-N',N'-adenosyl (rRNA) dimethyltransferase] of *E. coli* (data not shown) (31) and *B. subtilis* (Fig. 3,A) (Ogasawara, personal communication), respectively. Another ORF exhibited 21.4% homology to the LicA protein of *Haemophilus influenzae* (Fig. 3,E), which is supposed to mediate

biosynthesis of oligosaccharide of outer membrane (32). We searched for a DnaA-box consensus sequence (TTATCCACA) or its homolog with a single base change (Fig. 2) (13,33). No consensus sequence was found in the 9540 bp. One such sequence with a single base change was found in the *dnaA* gene near the N-terminus in the nt 4207–4215 region, in addition to those that have previously been reported (13).

Figure 4. 2D-gel analysis of the downstream region of *dnaA* bordered by *Xba*I and *Pvu*II sites. The horizontal line represents the physical map. The solid portion shows the 5.8-kb region that was analyzed in this experiment. The position and the direction of the *dnaA* gene are shown by a solid arrow. The open squares with numbers indicate the fragments used as probes for hybridizations. The numbers above the hybridization images correspond to the identification numbers of the probes. Position of size markers (denatured lambda DNA digested with *Hind*III) are shown in the left. The broken arrow indicates the direction of fork movement.

Analysis of the movement of the replication fork

The technique of neutral/alkaline 2D-gel analysis was developed by Huberman *et al.* (16). In the first dimension, DNA fragments are separated according to their mass. The DNA in the gel is then denatured in 50 mM NaCl and subjected to electrophoresis under alkaline conditions. As a consequence, the nascent strands dissociate from the parental strands and form a diagonal line. The lengths of the lines of nascent strands detected with various small probes provide information on the direction of fork movement (16).

The replicating intermediate DNAs, prepared from an exponentially growing culture of mycoplasma in agarose beads, were digested by *Pvu*II plus *Xba*I, concentrated by BND-cellulose column-chromatography and subjected to 2D-gel analysis. The filter was probed successively with 7 small fragments as shown in Figs. 4 and 5. The results obtained with the probes that are on the left side of the *Pvu*II site in the *dnaA* gene are shown in Fig. 4. The positions of probes 1, 2 and 3 are defined by nucleotide numbers in Fig. 2, as nt 2–619, nt 2137–2683 and nt 2678–3560, respectively. A schematic diagram of the image obtained after 2D-gel electrophoresis (Fig. 4) is shown in Fig. 6. Each signal is interpreted as follows: a, nascent and parental strands from intermediates whose replication has nearly finished (double-mass position); b, parental strands from replicating intermediates; c, both strands from nonreplicating molecules (single-mass position); d, both strands from fragmented molecules; e, this line has not been reported previously (a detailed interpretation can be found below in the 'DISCUSSION'); f, nascent strands from intermediates; and g, nascent strands and nicked parental strands of intermediates with breaks at joints of

Figure 5. 2D-gel analysis of the upstream region of *dnaA* bordered by *Pvu*II and *Xba*I sites. The horizontal line represents the physical map. The solid portion shows the 5.0-kb region that was analyzed in this experiment. The position and the direction of the *dnaA* gene are shown by a solid arrow. The open squares with numbers indicate the fragments used as probes for the hybridizations. The numbers above the hybridization images correspond to the identification numbers of the probes. Position of size markers (denatured lambda DNA digested with *Hind*III) are shown in the left. The broken arrow indicates the direction of fork movement.

Figure 6. Schematic representation of results of 2D-gel electrophoresis and the interpretation of each signal. The molecules separated in the first dimension are shown above a horizontal broken line. Signals (b), (f) and (c) are composed of single-stranded DNAs that are represented within boxes. The parental and nascent strands are shown by broken and solid lines, respectively. See text for details.

Figure 7. Comparison of *dnaA* regions of various bacterial chromosomes. The genes that are homologous in the three species of bacteria are hatched. The direction of the mycoplasmal genes has been reversed relative to that in Fig. 4 to be the same as that in reference (11).

the fork. The signals due to the nascent strands (f) were detected as diagonal lines from the double-mass position (a), which extended for varying distances that depended on the positions of probes on the map. The longest curve was obtained with probe 3 (Fig. 4), and the length of the curve decreased with increasing distance of the position of probes from the right end.

The results for the right side of the *PvuII* site are shown in Fig. 5. The interpretation of the images is the same as in Fig. 4. The positions of probes 4, 5, 6 and 7 are defined by nucleotide numbers in Fig. 2 as nt 3561–4249, nt 4250–5313, nt 7155–7725 and nt 7923–8558, respectively. The longest curve was obtained with probe 4, and the length of the curve decreased with increasing distance of the position of probes from the left end. These results (Figs. 4 and 5) show that the replication forks arose in a region that corresponded to the locations of probe 3 (nt 2678–3560) and probe 4 (nt 3561–4246) and proceeded bidirectionally.

DISCUSSION

The origin regions of some bacteria are located near their *dnaA* genes (8,9,10,34). The origin region is generally supposed to include the initiation site of replication and the entire set of initiation-controlling *cis*-factors. In the genome of *M. capricolum*, the *dnaA* gene is also found within the 46-kb region that contains the initiation site.

In this study, we first attempted to isolate clones that harbored large fragments that included the *dnaA* gene, but a mycoplasma library constructed with lambda EMBL3 was extremely unstable. The instability of the library may possibly have been due to the repetitive sequences in the mycoplasma genome, which has a high AT-content (35). Therefore, we tried to isolate clones of the *dnaA*-flanking regions as small fragments using stable pUC plasmids.

The regions around the *dnaA* gene are highly conserved in *E. coli*, *B. subtilis* and *Pseudomonas putida* (11). The similarity is remarkable not only among individual genes but also when their organization on the respective chromosomes is compared. At least 6 genes, *gyrB*, *recF*, *dnaN*, *dnaA*, *rpmH*, *rnpA*, in that order, are conserved in the three bacterial species that have been examined to date (11). In *M. capricolum*, we found that the arrangement and the directions of *dnaN*, *dnaA*, *rpmH* and *rnpA* were conserved, but the neighboring genes were not identical

to any genes that are found in the corresponding regions in other bacteria (Fig. 7). Instead, an ORF that is homologous to *ksgA* gene was found near the *dnaA* gene in *M. capricolum*, but this gene is located about 60 kb and 800 kb away from the *dnaA* gene in *B. subtilis* (Ogasawara, personal communication) and in *E. coli* (36), respectively. This result suggests that regions outside the *dnaN*-*dnaA*-*rpmH*-*rnpA* region are not conserved in *M. capricolum*.

An assay for *ars* (autonomously replicating sequence) has been useful for identifying the origin region of bacterial genomes (8,9,10,34). However, in *M. capricolum*, for which no transformation system has yet been constructed, the *ars* assay is not feasible. The technique of neutral/alkaline 2D-gel analysis (16) can provide crucial information about the movement of forks. Although this method is theoretically applicable to any replicon that has already been cloned, it had not been applied to the analysis of bacterial genomes. We applied 2D-gel analysis to the mapping of the initiation site of *M. capricolum* and were able to show clearly the directions of the fork movements.

This successful analysis might be due to the small size of the genome of *M. capricolum*, namely, 1156 kb (15). The intensity of a hybridization signal generally depends on the ratio of the length of the region to the size of the genome, and the probability that the region in question includes the replication forks also depends on this ratio.

The initiation site was located in a 1569-bp region that consisted of the *dnaA* gene and its downstream non-coding region (Fig. 2). We tried to map the initiation site to a smaller region using probes of about 300 bp, but no significant changes in the length of the nascent curves on the hybridization image were found (data not shown).

In each of the hybridization images, a vertical line from the double-mass point was observed, as shown schematically by line (e) in Fig. 6. This line has not been reported previously and has the following features. [i] The line (Fig. 6,e) extended for the same distance as the line that corresponded to the nascent strand (f) in the second dimension. This feature suggests that the vertical line (e) is composed of nascent strands of various sizes; [ii] The line extended from the double-mass position (a), in other words, all the molecules that formed the vertical line migrated most slowly in the first dimension. Therefore, the molecules must have very large mass in spite of the various sizes of the nascent strands. From these features, this line can be supposed to be composed

of the nascent strands from intermediates with multiple forks. Our previously reported data (14) also suggested the formation of multiple forks during the replication of the genome in *M. capricolum*.

The separate vertical line (g) from the single-mass position (c) also changed in length, depending on the probe used. The molecules forming this line have the same mass as the nonreplicating molecule, as reflected by the migration rate in the first dimension. The line (g) is interpreted as being composed of the nascent and nicked parental strands of intermediates, with breaks at the joints of the fork.

Some signals are observed on the upper side of the line (b) and on the left side of the point (a) in the images of Figs. 4 and 5. They are possibly owing to durations of the attending molecules in the migration in the both dimensions, which were mainly consisted of comigration of the molecules with molecules of higher mass.

The results obtained by the 2D-gel analysis do not provide any direct information about the initiation-controlling *cis*-factors, but information about the location of the initiation site *in vivo* provides an important clue to the nature of the initiation mechanism.

The DnaA proteins bind specifically to the 9-bp sequences (6), namely, the DnaA-boxes, that are found in the origin regions of many bacterial replicons (11). The consensus sequence in *E. coli*, TTATCCACA, is generally conserved in many replicons, including other bacterial genomes (11,33). However, Fujita, *et al.* found no consensus sequences in the regions adjacent to the *dnaA* gene in *M. capricolum* (13). In the genome of *B. subtilis*, which is phylogenetically related to *M. capricolum* (37,38), a cluster of DnaA-boxes is located in the upstream region about 2.6 kb away from the *dnaA* gene. Therefore, we searched for the consensus sequence of the DnaA-box in the extended sequences but we did not find it. In a previous report, Fujita *et al.* (13) reported three homologs of DnaA-boxes with single base changes from the consensus sequence in the upstream non-coding region of the *dnaA* gene and suggested that these homologs might be DnaA-boxes, judging from the similarity between the organization of the genes adjacent to *dnaA* and that in *B. subtilis*. We also searched for DnaA-boxes with one-base and two-base changes in the entire sequence of 9540 bp and found one new DnaA-box-like sequence with a single base change near the N-terminus of the *dnaA* gene. This result allows us to exclude the possibility that the one-base-changed, putative DnaA-boxes were just coincidences, and it supports the assumption that they function as the initiation-controlling *cis*-factors. Homologs with two-base changes were found nonspecifically all over the 9540-bp sequence (data not shown).

The three putative DnaA-boxes are located in the region adjacent to the 1569-bp region that includes the initiation site. Assuming that the initiation mechanism of *M. capricolum* is not very different from that of *E. coli* whose opening site of initiation is located less than 30 bases from the nearest DnaA-box (39), we propose that the initiation site may be located in the 5'-end region of the *dnaA* gene in the genome of *M. capricolum*.

Recently, T at position 4, C at position 6, and A at position 7 were shown to be the most invariant bases in the sequences of DnaA-boxes of various replicons (10,40). The sequences of the putative mycoplasmal DnaA-boxes are as follows: [i] TTATaCACA (nt 4207–4215); [ii] TTtTCCACA (nt 4334–4326); [iii] TTATCaACA (nt 4424–4432); and [iv] TTATCtACA (nt 4466–4474). Sequences [iii] and [iv] contain changes at position 6. The changes may indicate that the DnaA-

box sequences of *M. capricolum* are significantly different from those of other bacteria. Alternatively, a small number of DnaA-boxes may be sufficient for the initiation of replication of the *M. capricolum* genome.

ACKNOWLEDGEMENTS

We are grateful to Drs H. Yoshikawa and N. Ogasawara of Osaka University Medical School and to Dr T. Shinomiya of the Mitsubishi Kasei Institute of Life Science for helpful discussions. We thank the Department of Biology, Faculty of Science, Osaka University, for allowing us to use the Bio Image-Analyzer, BAS 2000. This work was supported in part by a Grant-in-Aid for Scientific Research on Priority Areas from the Ministry of Education, Science and Culture of Japan.

REFERENCES

1. Razin, S. (1985) *Microbiol. Rev.*, **49**, 419–455.
2. Pyle, L. E., Corcoran, L. N., Cocks, A. D., Bergemann, A. D., Whitly, J. C. and Finch, L. R. (1988) *Nucl. Acids Res.*, **16**, 6015–6025.
3. Herrmann, R. (1992) In Maniloff, J., McElhaney, R. N., Finch, L. R. and Baseman J. B. (eds.), *Mycoplasmas—Molecular Biology and Pathogenesis*. American Society for Microbiology, Washington, pp. 157–168.
4. Muto A., Andachi, Y., Yamao F., Tanaka R. and Osawa S. (1992) In Maniloff, J., McElhaney, R. N., Finch, L. R. and Baseman J. B. (eds.), *Mycoplasmas—Molecular Biology and Pathogenesis*. American Society for Microbiology, Washington, pp. 331–347.
5. Labaree, J. (1992) In Maniloff, J., McElhaney, R. N., Finch, L. R. and Baseman J. B. (eds.), *Mycoplasmas—Molecular Biology and Pathogenesis*. American Society for Microbiology, Washington, pp. 309–323.
6. Kornberg, A. and Baker, K. (1991) *DNA replication*, 2nd ed., W. H. Freeman and Company, N.Y. pp. 471–552.
7. Ogasawara, N., Moriya S. and Yoshikawa, H. (1991) In Ishikawa, A. and Yoshikawa, H. (eds.), *Control of Cell Growth and Division*. Japan Sci. Soc. Press, Tokyo/Springer-Verlag Berlin, pp. 3–21.
8. Yee, T. W. and Smith, D. W. (1990) *Proc. Natl. Acad. Sci. USA*, **87**, 1278–1282.
9. Marczyński, G. T. and Shapiro L. (1992) *J. Mol. Biol.*, **226**, 959–977.
10. Zakrzewska-Czerwinska, J. and Shrempf, H. (1992) *J. Bacteriol.*, **174**, 2688–2693.
11. Yoshikawa, H. and Ogasawara, N. (1991) *Mol. Microbiol.*, **5**, 2589–2597.
12. Pyle, L. E. and Finch, L. R. (1988) *Nucl. Acids Res.*, **16**, 6027–6039.
13. Fujita, M. Q., Yoshikawa, H. and Ogasawara, N. (1992) *Gene*, **93**, 73–78.
14. Miyata, M., Wang, L. and Fukumura, T. (1993) *J. Bacteriol.*, **175**, 655–660.
15. Miyata, M., Wang, L. and Fukumura, T. (1991) *FEMS Microbiol. Lett.*, **79**, 329–334.
16. Huberman, J. A., Spotilia, L. D., Nawotka, K. A., El-Assouli, S. M. and Davis, L. R. (1987) *Cell*, **51**, 473–481.
17. Razin, S. and S. Rottem. (1976) In Maddy, A. H. (ed.) *Biochemical Analysis of Membranes*, Chapman and Hall, London, pp. 3–25.
18. Sawada, M., Muto, A., Iwami, M., Yamao, F. and Osawa, S. (1984) *Mol. Gen. Genet.*, **196**, 311–316.
19. Saiki, R. A. (1990) In Innis, M. A., Gelfand, D. H., Sninsky, J. J., and White, T. J. (eds.), *PCR protocols—a Guide to Methods and Applications*. Academic Press, Inc., San Diego, pp. 13–20.
20. Sambrook, J., Fritsch, E. F. and Maniatis, T. (1989) *Molecular Cloning: A Laboratory Manual*, 2nd ed. Cold Spring Harbor Laboratory, Cold Spring Harbor NY.
21. Southern, E. (1975) *J. Mol. Biol.*, **98**, 503–517.
22. Medveczky, P., Chang, C. O. and Mulder, C. (1987) *Biotechniques*, **5**, 242–244.
23. Feiberg, A. P. and Volstein, B. (1983) *Anal. Biochem.*, **132**, 6–13.
24. Church, G. M. and Gilbert, W. (1984) *Proc. Natl. Acad. Sci. USA*, **81**, 1991–1995.
25. Perbal, B. (1988) *A Practical Guide to Molecular Cloning*, 2nd ed. Wiley-Interscience, New York.
26. Sanger, F., Nicklen, S. and Coulson, A. R. (1977) *Proc. Natl. Acad. Sci. USA*, **74**, 5463–5467.
27. Shinomiya, T. and Ina S. (1991) *Nucl. Acids Res.*, **19**, 3935–3941.
28. Dijkwel, P. A., Vaughn, J. P. and Hamlin, J. L. (1991) *Mol. Cell. Biol.*, **11**, 3850–3859.

29. Lai, C.Y. and Baumann, P. (1992) *Gene* **113**, 175–181.
30. Ogasawara, N. and Yoshikawa, H. (unpublished). X62539.
31. Roa, B. B., Connolly, D. M. and Winkler, M. E. (1989) *J. Bacteriol.* **171**, 4767–4777.
32. Weiser, J. N., Love, J. M. and Moxon, E. R. (1989) *J.N., Cell*, **59**, 657–665.
33. Fuller, R. S. Funnell, B. E. and Kornberg, A. (1984) *Cell*, **38**, 889–900.
34. Moriya, S., Atlung, T., Hansen, F. G., Yoshikawa, H. and Ogasawara, N. (1992) *Mol. Microbiol.*, **6**, 309–315.
35. Razin, S. (1992) In Maniloff, J., McElhaney, R. N., Finch, L. R. and Baseman J. B. (eds.), *Mycoplasmas—Molecular Biology and Pathogenesis*. American Society for Microbiology, Washington, pp. 3–22.
36. Bachmann, B. J. (1990) *Microbiol. Rev.*, **54**, 130–197.
37. Hori, H., Sawada, M., Osawa, S., Murao, K. and Ishikura, H. (1981) *Nucl. Acids Res.*, **9**, 540–541.
38. Rogers, M. J., Simmons, J., Walker, R. T., Weisburg, W. G., Woese, C. R., Tanner, R. S., Robinson, I. M., Stahl, D. A., Olsen, G., Leach, R. H. and Maniloff, J. (1985) *Proc. Natl. Acad. Sci. USA*, **82**, 1160–1164.
39. Bramhill, D. and Kornberg, A. (1988) *Cell* **52**, 743–755.
40. Schaefer, C. and Messer, W. (1991) *Mol. Gen. Genet.*, **226**, 34–40.