The *ARS309* chromosomal replicator of *Saccharomyces cerevisiae* depends on an exceptional ARS consensus sequence

James F. Theis and Carol S. Newlon*

Department of Microbiology and Molecular Genetics, University of Medicine and Dentistry of New Jersey—New Jersey Medical School, 185 South Orange Avenue, Newark, NJ 07103

ABSTRACT Autonomously replicating sequence (ARS) elements, which function as the cis-acting chromosomal replicators in the yeast *Saccharomyces cerevisiae*, depend upon an essential copy of the 11-bp ARS consensus sequence (ACS) for activity. Analysis of the chromosome III replicator *ARS309* unexpectedly revealed that its essential ACS differs from the canonical ACS at two positions. One of the changes observed in *ARS309* inactivates other ARS elements. This atypical ACS binds the origin recognition complex efficiently and is required for chromosomal replication origin activity. Comparison of the essential ACS of *ARS309* with the essential regions of other ARS elements revealed an expanded 17-bp conserved sequence that efficiently predicts the essential core of ARS elements.

The replication of yeast chromosomes is accomplished by the activation of multiple cis-acting replicators (for review, see ref. 1). These replicators were first identified by their ability to promote high-frequency transformation and stable extrachromosomal maintenance of plasmids in yeast (2, 3). On the basis of these properties, they were named autonomously replicating sequence (ARS) elements. Two-dimensional gel electrophoresis techniques for identifying replication intermediates were used to demonstrate that ARS elements colocalize with plasmid replication origins (4, 5). The analysis of chromosomal replication initiation by these techniques revealed that many ARS elements are cis-acting replicators required for chromosomal replication origin activity (6–9). However, only a subset of ARS elements are active as origins in their native chromosomal contexts (10, 11).

ARS elements are modular (Fig. 1). All contain an essential match or near match to the 11-bp ARS consensus sequence (ACS) WTTTAYRTTTW (where W is A or T, Y is T or C, and R is A or G; refs. 17 and 18). Mutations in this sequence abolish ARS function (for review, see ref. 1). ARS elements also contain additional near matches to the ACS that are dispensable in the ARS elements that have been examined (12–16). The mutation of several bases on either side of the essential match to the ACS also impairs ARS function; this larger element has been called domain A (12–16, 19, 20). The essential match to the ACS is recognized and bound by a six-subunit complex, the origin recognition complex (ORC), that functions as a DNA-dependent ATPase and is thought to be the replication initiator protein (21–26).

In addition to domain A, a region 3′ to the T-rich strand of the ACS is also required; it is referred to as domain B (19). There appears to be little or no sequence conservation among ARS elements in the domain B region. Instead, domain B is composed of several small nonredundant sequence elements (B1 to B4) that can differ from one ARS to another (12–15). Despite a lack of sequence homology, some of these elements act as functional analogues. Thus, the B1 and B2 elements of

ARS307 can substitute for their counterparts in *ARS1*, and vice versa (13). Some mutations in the B1 elements of *ARS1* and *ARS307* reduce the efficiency of ORC binding, suggesting that B1 functions along with domain A as part of the ORC binding site (27, 28). However, the observation that some B1 mutations reduce ARS efficiency without affecting ORC binding *in vitro* suggests that the B1 element plays an additional role in replicator activity (28). The B3 element of *ARS1* is a binding site for ARS binding factor 1, a transcriptional activator and repressor. ARS binding factor 1 binding sites are found 5′ of the T-rich strand of the essential match to the ACS of *ARS121* (Fig. 1*D*); these sites stimulate ARS activity from this position or when moved into domain B (16).

Despite our ability to broadly describe ARS structure, many specific issues are not resolved. In this manuscript we describe the analysis of *ARS309*, an active chromosomal replicator on the right arm of chromosome III. Detailed mutational analysis of the matches to the ACS within *ARS309* led to the unexpected discovery that the essential ACS of *ARS309* contains a mismatch in a group of three thymines that are invariant in the other 19 essential ACS elements defined to date. We show that this variant ACS is essential for chromosomal replicator activity and that it binds ORC *in vitro*. Comparison of this newly defined essential region of *ARS309* with other ARS elements allowed the definition of an expanded ACS that substantially improves predictions of the essential ACS of other ARS-containing DNA fragments.

MATERIALS AND METHODS

Plasmids and Strains. Plasmids ARS309–306 and ARS309–326 were made by cloning the *ARS309*-containing 0.35-kb *Eco*RI–*Bam*HI fragment of pVH53 (29) in pRS306 (30) and pRS326 (14), respectively. The *ARS309* fragment was derived from a genomic 0.35-kb *Eco*RI–*Hpa*I fragment (chromosome III coordinates 131, 585 to 131, 933; MIPS-Yeast Genome Project, Sept. 2, 1996, release) by adding a *Bam*HI linker to the *Hpa*I site. The plasmid 309-flk was made by ligating the 1.4-kb *Hin*dIII–*Eco*RI fragment of J11D-5a (31), the 0.9-kb *Eco*RI–*Bgl*II fragment of J11D-2 (31), and pRS306 digested with *Hin*dIII plus *Bam*HI. 309-flk was digested with *Hpa*I, and a *Bam*HI linker was inserted to create 309-flk+Bam.

Escherichia coli DH5α (Life Technologies, Grand Island, NY) was used for the propagation of plasmids. The diploid yeast strain 1C6 (14) was used for testing ARS activity, and the haploid yeast strain YP45 (30) was used for the analysis of chromosomal DNA replication intermediates.

Mutagenesis of *ARS309*. Exonuclease III deletions of *ARS309* were constructed as described (14). For deletions from the *Eco*RI end, ARS309–306 was digested with *Eco*RI plus *Kpn*I; for deletions from the *Hpa*I end, ARS309–326 was

Genetics: Theis and Newlon
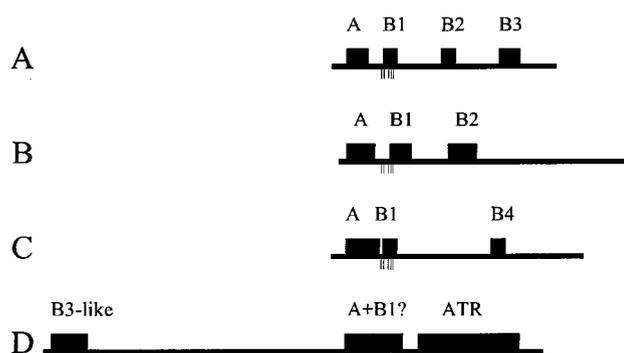
Proc. Natl. Acad. Sci. USA 94 (1997)   10787

Fig. 1. Modular nature of ARS elements. The functional elements defined by mutational analysis of four ARS elements are indicated by boxes on the line diagrams. All four are aligned by their essential matches to the ACS and drawn to the same scale. The tick marks below the lines indicate the positions +8, +10, +15, +17, and +18 relative to the end of the essential match to the ACS (see text). (*A*) *ARS1* (12). (*B*) *ARS307* (13, 14). (*C*) *ARS305* (15). (*D*) *ARS121* (16). The region labeled A+B1? corresponds to the "core" element of Walker *et al.* (16). ATR is the A+T-rich domain. B3-like corresponds to the pair of ARS binding factor 1 binding sites.

digested with *Bam*HI plus *Sac*I. The deletion fragment "57 to 135" is the genomic *Rsa*I fragment cloned into pRS326 at the *Eco*RI site, which was blunted by filling in with T4 polymerase. The site-specific mutations in the matches to the ACS were generated by the method of Kunkel *et al.* (32), as described by Theis and Newlon (14). The oligonucleotides used were 22 nucleotides long and were designed to change two of the three thymidines at positions 8, 9, and 10 of the matches to the ACS to other bases. The alterations made are shown in Table 1. The plasmid ARS309–326 was used as the template for all the single mutations. The double and triple mutants were constructed by using the appropriate singly or doubly mutant plasmid as template with the desired oligonucleotide. Oligonucleotides were synthesized by the New Jersey Medical School Molecular Resource Facility.

To integrate mutant *ARS309* derivatives into the chromosome, the 0.35-kb *Eco*RI–*Bam*HI fragment from the desired mutant plasmid was swapped for the wild-type 0.35-kb fragment in 309-flk+Bam. A two-step gene replacement (33) was performed in YP45 (30), using the unique *Eco*RI or *Blp*I sites to direct integration. Mutant popouts were identified by the presence of a new *Bam*HI site detected by Southern blotting and/or PCR amplification of genomic DNA. PCR products amplified from the mutant strains were sequenced to confirm the presence of the mutations.

Examination of DNA Replication Intermediates. Preparation of genomic DNA and two-dimensional gel electrophoresis were described (14).

ORC Footprinting. Probes were made by PCR as described (34) by using T7 primer (Life Technologies) end-labeled by

Table 1. ACS matches and mutations

ACS match	Location, bp	Mutant version	Positions altered
ATCTATATTTT	31–41	ATCTATATaaT	39, 40 T → A
TTATATGTTTA	169–159	TTATATGaaTA	161, 162 T → A
TTTTATATGTT	171–161	TTTTATATGaa	161, 162 T → A
TTTTTTATTTT	241–231	TTTTTTAaaTT	233, 234 A → T
TTTGAAGTTTT	75–85	TTTGAAGTaaT	83, 84 T → A
TTAAATGTTTT	104–94	TTAAATGTggT	95, 96 A → C
GTTTATATCTT	62–72	GTTTATAgCgT	69, 71 T → G

Matches to the ACS in the 350-bp *Hpa*I–*Eco*RI fragment. Positions that differ from the ACS are underlined. Alterations in the mutant versions are indicated as lowercase type. All matches are shown as the T-rich strand.

incubation with polynucleotide kinase and [γ-^{32}P]ATP and unlabeled M13 reverse primer (New England Biolabs). The plasmid templates for PCR were ARS309–326 and the appropriate mutant derivatives. Footprinting was performed as described by Klemm *et al.* (22). Recombinant ORC was the generous gift of Stephen Bell (Massachusetts Institute of Technology). Chemical sequencing reactions (A+G and T) were performed as described by Richterich *et al.* (35).

RESULTS AND DISCUSSION

ARS309 is associated with a chromosomal replication origin on the right arm of *Saccharomyces cerevisiae* chromosome III (36). The DNA sequences required for maximum plasmid stability are contained within a 0.35-kb *Eco*RI–*Hpa*I genomic DNA fragment (29, 31). In contrast to *ARS1* and *ARS307* (12–14), *ARS309* contains no perfect match to the ACS but rather contains four 10 of 11 matches (Table 1).

As a first step in localizing *ARS309* within the 0.35-kb genomic DNA fragment, we constructed exonuclease III deletion series from either the *Eco*RI or *Hpa*I sides. As shown in Fig. 2, deletions from the *Eco*RI side cause a gradual reduction in plasmid stability (constructs A, B, and C and data not shown). The small fragment containing nucleotides 1–147 (Fig. 2C) was still Ars$^+$ despite deletion of three of the four 10 of 11 matches to the ACS. Deletion of 53 bp from the *Hpa*I side had no effect on ARS function (Fig. 2D), despite removing the 10 of 11 match to the ACS at positions 31–41. The data from the two sets of unidirectional deletions suggest that all of the 10 of 11 matches are dispensable. This was confirmed in two ways. (*i*) Two-base changes were made in each 10 of 11 match, both singly and in combination; similar alterations of the ACS have been shown to abolish activity of other ARS elements (15,
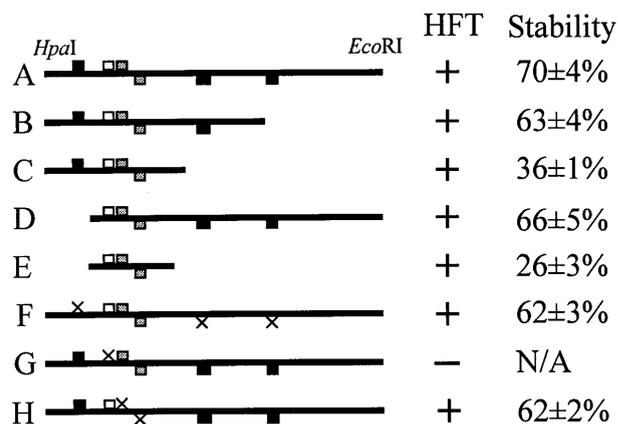

Fig. 2. Effect of deletions and point mutations on *ARS309* function. Lines represent the sequences present in plasmids. Position 1 is the first adenine of the *Hpa*I site; position 349 is the cytidine of the *Eco*RI site. Solid boxes show the 10 of 11 matches to the ACS at positions 31–41, 159–169, 161–171, and 231–241. The open box marks the position of the 9 of 11 match at positions 62–72. The stippled boxes show the positions of the 9 of 11 matches at positions 75–85 and 94–104. × indicates point mutations altering the ACS matches (see Table 1). Boxes above the line indicate that the T-rich strand of the ACS is in the top strand, and boxes below the line indicate the T-rich strand of the ACS is in the lower strand. HFT: +, high-frequency transformation; −, no high-frequency transformation. Stability: Numbers show percentage of plasmid bearing cells in a logarithmic-phase culture under selection (mean ± SEM). (*A*) Full-length wild-type *ARS309* (base pairs 1–349). (*B*) Deletion mutant containing base pairs 1–227. (*C*) Deletion mutant containing base pairs 1–147. (*D*) Deletion mutant containing base pairs 54–349. (*E*) *Rsa*I subclone (base pairs 57–135). (*F*) Triple point mutation altering all four 10 of 11 matches (see Table 1). (*G*) Mutation altering the 9 of 11 match at positions 62–72 (see Table 1). (*H*) Mutation altering the two 9 of 11 matches at positions 75–85 and 94–104 (see Table 1).
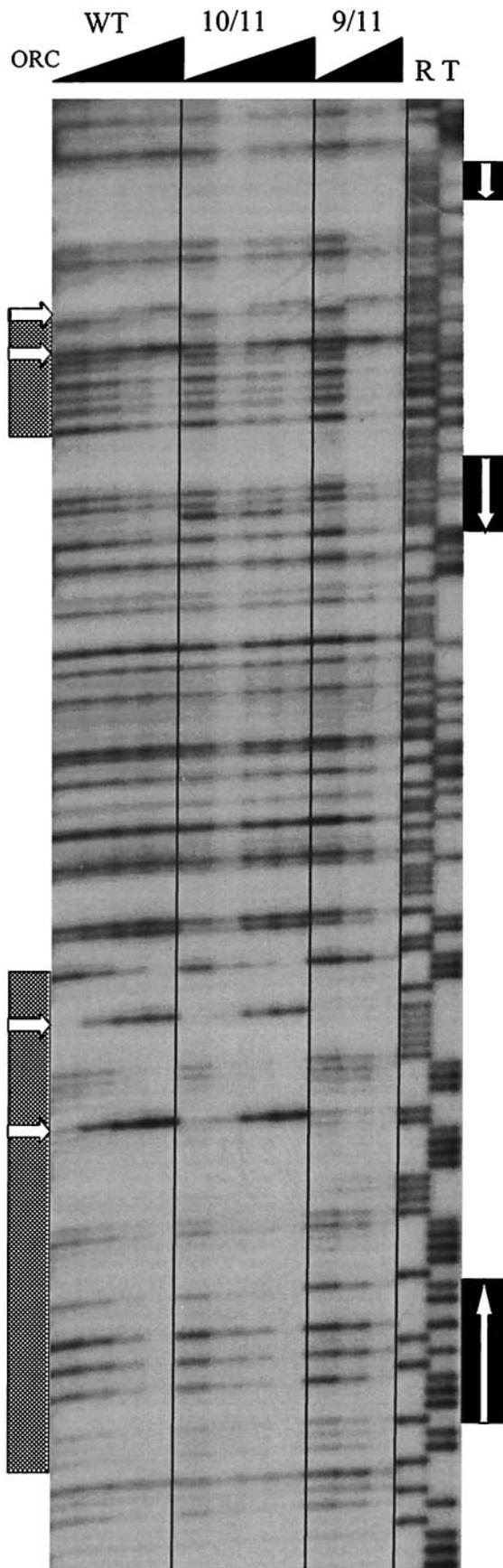
FIG. 3.   ORC footprinting of *ARS309* and mutant derivatives. End-labeled fragments were footprinted with ORC. WT, wild-type *ARS309* (Fig. 2*A*); 10/11, *ARS309* with all four 10 of 11 matches mutated (Fig. 2*F*); 9/11, *ARS309* with the essential 9 of 11 match

18, 23, 37, 38). None of the single mutations had a significant effect on ARS function, nor did the triple mutation that altered all four 10 of 11 matches (Fig. 2*F*, Table 1, and data not shown). It was possible to alter all four 10 of 11 matches with three mutations because two of them overlap. (*ii*) An *Rsa*I fragment retained ARS activity (Fig. 2*E*).

This 79-bp *Rsa*I fragment contains three 9 of 11 matches to the ACS (Table 1). To determine whether one of these matches functions as the essential match to the ACS, 2-bp changes that alter each of these 9 of 11 matches were made in the 0.35-kb *Hpa*I–*Eco*RI fragment. Mutation of the matches at positions 75–85 and 94–104 had no effect on ARS function, either singly or in combination (Fig. 2*H*). In contrast, ARS activity was abolished when the match at positions 62–72 was altered, identifying this as the essential match (Fig. 2*G*).

It was surprising to find that this 9 of 11 match, gTT-TATATcTT, was the essential ACS. In particular, the TcT instead of TTT was surprising because all of the previously identified essential matches to the ACS conserved the three thymines at positions 8, 9, and 10 (39). In addition, any base-pair substitution at any one of the TTT sites inactivates *ARS307* (18). Some changes at these positions have also been shown to inactivate *ARS1*, and these changes reduced ORC binding to *ARS1* (23). *ARS309* differs from all other ARS elements examined to date in that its essential match has a cytidine at position 9.

Because all essential matches to the ACS tested to date bind ORC, it was important to determine whether ORC binds to the exceptional ACS in *ARS309*. In other ARS elements, ORC protects a region of ~50 bp that includes the essential ACS. ORC binding also induces DNase I-hypersensitive sites within the B domain, the most prominent hypersensitive sites occurring 12–16 and 22–26 bp away from the ACS. *In vitro* footprinting with purified ORC was performed on the 0.35-kb *ARS309* fragment. As can be seen in Fig. 3, ORC protects the region from positions 60 to 102, which contains the essential ACS. Moreover, ORC binding induces two hypersensitive sites at positions 87 and 97, which are 15 and 25 bp away, respectively, from the ACS. The hypersensitive sites are detected in the wild-type fragment with as little as 5 ng of ORC, although 25 ng are required for full protection of this region. When the mutated ACS that abolishes origin function was tested for ORC binding, protection of the 60 to 102 region was abolished, even at 100 ng of ORC, and no hypersensitive sites were seen. The absence of hypersensitive sites at 100 ng of ORC indicates that the 2-bp mutation in the essential ACS reduced ORC binding at least 20-fold.

ORC protected neither the 10 of 11 match at positions 31–41 (data not shown) nor the two overlapping sites at positions 159–171. Although there was weak protection of the DNase I cleavage sites at positions 159, 165, and 167, there was no other protection in the region from positions ~120 to 171, as would be expected for ORC binding at positions 159–171. In addition, there are no hypersensitive sites induced within this region. Because there are no DNase I cleavage sites in naked

---

mutated (Fig. 2*G*). For WT and 10/11, the four footprinting reactions contained 0, 5, 12.5, and 25 ng of ORC; for 9/11, the three reactions contained 0, 25, and 100 ng. Lanes R and T are chemical sequencing reactions for A+G, and T, respectively. The hatched boxes to the left indicate the protected regions (positions 60–102, lower box; 180–204, upper box), and the arrows point to hypersensitive sites (positions 87, 97, 196/197, 205/206 from bottom to top). Solid boxes to the right of the sequencing reactions mark the positions of the 10 of 11 matches to the ACS at positions 231–241 (upper box), 159–169 and 161–171 (middle box), and the essential 9 of 11 match at positions 62–72 (lower box). Arrows within these boxes indicate the orientation of the T-rich strand of the ACS. Note that neither the protected regions nor the hypersensitive sites are altered in the 10 of 11 match mutant, but the 9 of 11 match mutant abolishes the protection and hypersensitive sites in the region from positions 60 to 102.
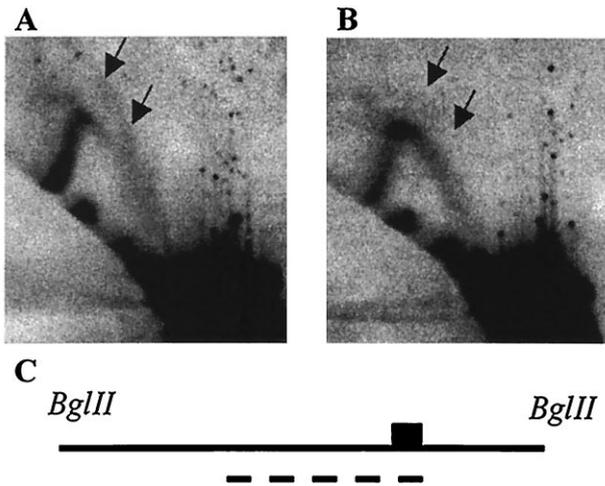
Genetics: Theis and Newlon

Proc. Natl. Acad. Sci. USA 94 (1997)   10789

FIG. 4. Analysis of *ARS309* replication intermediates. Genomic DNA was digested with *Bgl*II, electrophoresed, and blotted as described by Brewer and Fangman (5). The 3.6-kb *Bgl*II fragment containing *ARS309* was detected by probing with the 1.4-kb *Hin*dIII–*Eco*RI fragment indicated by the dashed lines in *C*. The 0.35-kb *Hpa*I–*Eco*RI fragment containing *ARS309* is denoted by the box. (*A*) Wild-type *ARS309*. The arrows point to bubble-shaped replication intermediates, indicating that *ARS309* is active as an origin on the chromosome. (*B*) *ARS309* with the essential 9 of 11 match mutated (Fig. 2*G*). There are no bubble-shaped replication intermediates (arrows), demonstrating that origin function is abolished in the mutant.

DNA in the region containing the fourth 10 of 11 match (positions 231–241), we cannot assess protection by ORC in this region. However, binding of ORC to this match to the ACS would be expected to protect the region from positions ~190 to 241, and it is clear that the DNase I cleavage sites at positions 217 and 221 were not protected. Furthermore, the absence of hypersensitive sites within this region suggests that ORC does not bind to this match. We also examined ORC binding to the fragment containing mutations in all four 10 of 11 matches to the ACS. As seen with the wild-type fragment, there is protection of the region from positions 60 to 102 and induction of the two hypersensitive sites at positions 87 and 97. The amount of ORC required for complete protection or the induction of hypersensitive sites was not different in the case of this mutant fragment.

In addition to the protection and induction of hypersensitive sites in the region containing the essential ACS, there was partial protection of a second region from positions 180 to 203 or 204 at 25 ng of ORC. Moreover, there was induction of two hypersensitive sites in this region at positions 196 or 197 and 205 or 206. These hypersensitive sites became visible at 12.5 ng of ORC, whereas the sites associated with the essential ACS were visible at 5 ng. This second protected region is present in both the 9 of 11 match mutant and the triple 10 of 11 match mutant. Despite the ORC binding detected *in vitro*, this region cannot provide ARS function (Fig. 2*G* and data not shown). The deletion analysis in Fig. 2 demonstrates that sequences between positions 147 and 227 contribute to ARS function, but it is not clear whether the enhanced stability of the larger fragment is due to the presence of this weak ORC binding site or some other sequence.

To determine whether the chromosomal replication origin associated with *ARS309* depends upon the essential ACS, the chromosomal copy of *ARS309* was replaced with mutant derivatives, and replication intermediates were examined by two-dimensional gel electrophoresis. Plasmids carrying mutant ARS derivatives and about 1 kb of flanking DNA were constructed and used to replace the chromosomal copy of *ARS309*. Fig. 4*A* shows the replication intermediates of a restriction fragment carrying wild-type *ARS309*. The presence

of bubble-shaped intermediates (arrows) and the relative paucity of small Y-shaped intermediates indicate that *ARS309* is active as a chromosomal replication origin. In contrast, when the Ars⁻ mutation (Fig. 2*G*) was present in the chromosome, no bubble-shaped replication intermediates were detected (Fig. 4*B*, arrows), and there was a commensurate increase in the signal from small Y-shaped intermediates, indicating that origin function is abolished by this mutation. Strains carrying the triple mutation altering all four 10 of 11 matches (Fig. 2*F*) and the double mutation changing the two nonessential 9 of 11 matches (Fig. 2*H*) gave a pattern similar to the strain carrying wild-type *ARS309* (data not shown).
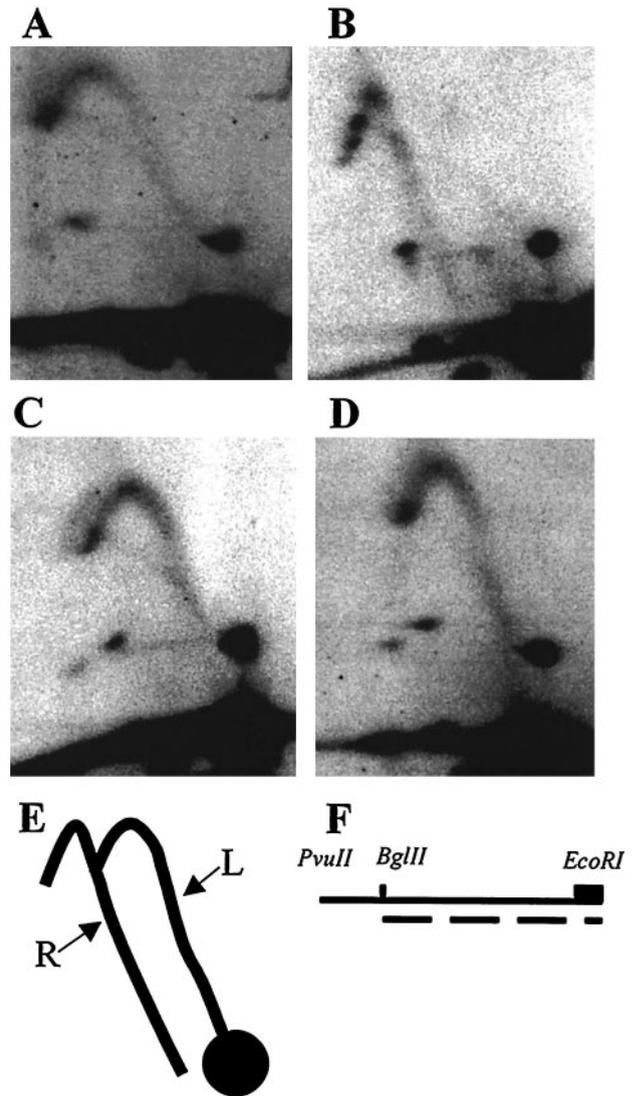

FIG. 5. Direction of replication fork movement in the left-flanking region of *ARS309* and mutant derivatives. Genomic DNA was digested with *Pvu*II and *Eco*RI before electrophoresis in the first dimension. Gel slices were excised, and the DNA was digested *in situ* with *Bgl*II prior to electrophoresis in the second dimension. (*A*) Wild-type *ARS309*. (*B*) Essential ACS mutant (Fig. 2*G*). (*C*) *ARS309* with all 10 of 11 matches mutated (Fig. 2*F*). (*D*) *ARS309* with nonessential 9 of 11 matches mutated (Fig. 2*H*). (*E*) Diagram of replication intermediates. When *ARS309* is active, leftward-moving replication forks traverse the *Eco*RI–*Pvu*II fragment (see *F*), and replication intermediates fall on the arc labeled L. When *ARS309* is inactive, rightward-moving forks from *ARS307* or *ARS308* traverse the fragment, and the replication intermediates fall on the arc labeled R. (*F*) Line drawing indicating the position of relevant restriction sites and of the 0.35-kb *Hpa*I–*Eco*RI fragment containing *ARS309* (box). The 2.7-kb *Eco*RI–*Bgl*II fragment used as a probe is represented by the dashed line.

An analysis of the direction of replication fork movement (40) through the regions flanking *ARS309* was performed on these mutants to determine whether there was any small reduction in origin usage that might have been missed by the standard two-dimensional gel analysis in Fig. 4. The region immediately to the left of *ARS309* was examined as a 3.5-kb *Pvu*II–*Eco*RI fragment that was cleaved *in situ* with *Bgl*II prior electrophoresis in the second dimension (Fig. 5F). When *ARS309* was active (Fig. 5 *A* and *E*), replication forks moved leftward from *ARS309* at the extreme right end of this fragment, resulting in a fork-direction pattern in which the early replication intermediates originate from the position of the unreplicated *Bgl*II–*Eco*RI fragment. In contrast, when *ARS309* was inactive (Fig. 5 *B* and *E*), replication forks originating from *ARS307* or *ARS308* moved rightward through this region, resulting in a pattern in which the early replication intermediates arose from a position to the left of the unreplicated *Bgl*II–*Eco*RI fragment. The presence of rightward-moving replication forks in the Ars⁻ mutants confirms and extends the conclusion that the 2-bp substitution that inactivates *ARS309* also inactivates its activity as a chromosomal replicator.

We also examined a mutant in which all four 10 of 11 matches were eliminated (Fig. 2F) and a mutant in which the two nonessential 9 of 11 matches were eliminated (Fig. 2H); both gave patterns indistinguishable from wild-type, indicating that these mutations have no effect on origin function (Fig. 5 *C* and *D*).

The direction of replication fork movement through the region to the right of *ARS309* was also examined. As we reported previously for a small circular derivative of chromosome III (36), forks traverse this region from right to left regardless of the activity of *ARS309*, so this region is not informative (data not shown).

This analysis of the *ARS309* chromosomal replicator demonstrates that *ARS309* is atypical, depending for activity upon an ACS that agrees with the consensus at only 9 of 11 positions. Moreover, the T → C change at position 9 present in the essential ACS inactivates other ARS elements. How then does *ARS309* function? Three observations suggested a solution to us. (*i*) Our *in vitro* analysis of *ARS309* indicates that ORC binds to the essential ACS. (*ii*) Linker-substitution mutagenesis of *ARS307, ARS1, ARS121,* H4 ARS, and *ARS305* indicates that domain A, the region essential for ARS function, extends several base pairs to either side of the ACS (12–16, 41). (*iii*) Some mutations within the B1 element of *ARS1* and *ARS307*

decrease ORC binding (27, 28), suggesting that the ORC recognition sequence is larger than the ACS.

In an attempt to identify additional sequences important for ARS function, we aligned sequences flanking the essential match to the ACS for all those ARS elements where the essential match had been identified by mutation (Table 2). There is substantial conservation of the 3 bp that flank both sides of the essential match to the ACS. On the 5′ side of the T-rich strand of the ACS, at position −1, all of the sequences have an adenine or thymine. There is also a strong bias for adenine or thymine at positions −2 (80% of ARS elements) and −3 (85% of ARS elements). Positions −4 to −6 show less bias, with 60–65% of ARS elements having an adenine or thymine at these positions. On the 3′ side of T-rich strand of the essential match, 75% of sequences have a guanine at position +1. Single-base changes at this position inhibit, but do not abolish, the activity of *ARS307, ARS305,* and HO ARS (15, 18, 20). Positions +2 and +3 both have thymines in 65% of sequences. This analysis suggests that the ACS should be expanded to the 17-bp sequence WWWWTTTAYRTT-TWGTT, which we will refer to as the extended ARS consensus sequence (EACS).

We searched the sequence of the 350-bp *ARS309* fragment for matches to the EACS. The best match (positions 59–75, ATT*g*TTTATAT*c*TTGTT) contains the essential match to the ACS identified by mutation. Thus, the 17-bp EACS correctly identifies the essential sequence for *ARS309* function, whereas the conventional 11-bp ACS does not. A similar analysis was performed with the sequences of restriction fragments carrying the other ARS elements in Table 2. The sequences analyzed varied from 120 to 1400 bp, with an average size of 640 bp. The EACS does a better job of identifying the essential sequence than the conventional ACS does. For example, the 640-bp fragment carrying *ARS608* has three 10 of 11 matches to the ACS (37). The best match to the EACS correctly identifies the essential match. Similarly, for the rDNA ARS, there is one perfect match to the ACS and two 10 of 11 matches; one of the 10 of 11 matches is essential for function (44). The best match to the EACS corresponds to the essential match. For *ARS1,* there is one perfect match to the ACS that is essential for function (19). However, there are four 15 of 17 matches to the EACS, one of which corresponds to the essential match. Even in this case, the EACS identifies the correct match as a candidate. Overall, the EACS correctly

Table 2.   Alignment of ARS sequences by essential ACS matches

ARS	5′	Essential match	3′
ARS309	TACATT	GTTTATATCTT	GTTTTGAAGTTTTAGCTTTG
ARS121 (42)	GTGAGA	TGTTTTGTTTA	ACATTAGTTTCAAATTAACA
ARS1 (43)	TACAGA	TTTTATGTTTA	GATCTTTTATGCTTGCTTTT
rDNA ARS (44)	TTGATT	GTTTATGTTTT	GTGTGATGATTTTACATTTT
HO ARS (20)	GTAAAA	TTTAATATTTT	GGATGAAAAAAACCATTTTT
H4 ARS (41)	ATTTAT	TTTTATGTTTT	GTATTTGGATTTTAGAAAGT
HMRE ARS (45)	AACATT	TTTTATATTTA	GGTATTAAATTGCGATTTAA
ARS305 (15)	TGGTTT	TTATATGTTTT	GTTATGTATTGTTTATTTTC
ARS307 (18. 46)	TTTTTT	ATTTATGTTTT	CTTCTTCACACATGGGTTAC
ARS601 (36)	GTACTA	ATTTCCATTTT	TGAGTTTAGGCGGTGCTTTC
ARS602 (36)	ACCAAT	TTATACGTTTA	GTTTTAACATCATCACAATG
ARS603 (36)	CCCTAA	TTTCATATTTT	GTATCAAAACTTTAAAATCT
ARS603 (36)	CCAGAT	TTTAAAGTTTT	GATACAAAATATGAAATTAG
ARS604 (36)	TGTTCA	TTTTACGTTTT	GTTACACAAATCCAATCGAA
ARS605 (36)	AGAATT	AATTACGTTTT	GTTGCTAAAGGAAACTTTAC
ARS606 (36)	ATATAT	ATTTATATTTT	CGTTGCAATTTCGCAATTTT
ARS607 (36)	TGTCTT	GTTTATATTTA	GTTACGTTGGGATCAAGTTT
ARS608 (36)	TTTTTT	TTTTACTTTTA	GTTTTCTTCTATGCGCAAGC
ARS609 (36)	CCATTT	TTTTATGTTTT	TTCCGGAATTGGCTAAATTC
ARS1413 (37)	GAAACA	ATTTGTATTTA	GTTGGTGTGACCTCAAATTT
Consensus	---WWW	WTTTAYRTTTW	GTT----W-W----R-WT--

W, A or T; Y, T or C; R, A or G.

Genetics: Theis and Newlon

Proc. Natl. Acad. Sci. USA 94 (1997)   10791

identifies the essential match to the ACS in 14 of the 20 sequences. In three other cases, the EACS identifies the correct match, but also one or more sequences with the same number of mismatches. In contrast, the conventional ACS only identifies 12 essential matches correctly, and in five other cases, the correct match is one of several with the same number of mismatches. We are using the 17-bp EACS to guide our analysis of other ARS elements from chromosomes III of *S. cerevisiae* and *Saccharomyces carlsbergensis* and have found it to be a better predictor of essential sequences than the original ACS, identifying the essential ACS in four of six ARS elements examined (C. Yang, J.F.T., and C.S.N., unpublished results).

The EACS also identifies a 14 of 17 match at positions 167–183 of the *ARS309* sequence that may account for the weak ORC binding noted in Fig. 3. In addition, the weak protection of positions 159, 165, and 167 may also result from ORC binding at this EACS. The hypersensitive sites at positions 195–196 and 204–205 are spaced the same distance from this sequence as the hypersensitive sites at positions 87 and 97 are from the best match to the EACS. However, there are two other 14 of 17 matches that fail to bind ORC (positions 98–114 and 205–221).

In addition to the conserved sequences that we propose are part of an expanded ACS, the sequence comparison in Table 2 revealed five other conserved positions. The base at positions +8, +10, and +17 is adenine or thymine in 85%, 80%, and 90% of the sequences, respectively. Position +15 is a purine in 85% of the sequences, and position +18 is thymine in 75% of the sequences. These positions are indicated in Fig. 1 for the three ARS elements that have been analyzed by linker-substitution mutagenesis. Positions +17 and +18 are within the B1 elements of all three of these ARS elements (12–14, 47). In fact, mutations at position +18 of *ARS1* diminish ORC binding (28). Positions +10 and +15 fall within the B1 elements of *ARS1* and *ARS305* but not within that of *ARS307*, which is further away from the ACS. Thus, the conservation of DNA sequence at these positions likely identifies base pairs of functional importance. The reason for conservation at position +8 is less clear. Position +8 lies between the ACS and the B1 elements of *ARS1*, *ARS305*, and *ARS307*, a region where mutations have no effect on ARS function. Further experiments will be required to determine whether this position is important in other ARS elements.

In summary, the identification of an atypical ACS in *ARS309* has led to a reassessment of the sequences required for ARS function. It seems likely that the detailed analysis of additional ARS elements will reveal features that are not apparent from the small sample currently available.

1. Theis, J. F. & Newlon, C. S. (1996) in *The Mycota*, eds. Brambl, R. & Marzluf, G. A. (Springer, Berlin), Vol. 3, pp. 1–28.
2. Hsiao, C. L. & Carbon, J. (1979) *Proc. Natl. Acad. Sci. USA* **76,** 3829–3833.
3. Stinchcomb, D. T., Struhl, K. & Davis, R. W. (1979) *Nature (London)* **282,** 39–43.
4. Huberman, J. A., Spotila, L. D., Nawotka, K. A., el-Assouli, S. M. & Davis, L. R. (1987) *Cell* **51,** 473–481.
5. Brewer, B. J. & Fangman, W. L. (1987) *Cell* **51,** 463–471.
6. Deshpande, A. M. & Newlon, C. S. (1992) *Mol. Cell. Biol.* **12,** 4305–4313.
7. Rivier, D. H. & Rine, J. (1992) *Science* **256,** 659–663.
8. Huang, R. Y. & Kowalski, D. (1993) *EMBO J.* **12,** 4521–4531.
9. Marahrens, Y. & Stillman, B. (1994) *EMBO J.* **13,** 3395–3400.
10. Dubey, D. D., Davis, L. R., Greenfeder, S. A., Ong, L. Y., Zhu, J. G., Broach, J. R., Newlon, C. S. & Huberman, J. A. (1991) *Mol. Cell. Biol.* **11,** 5346–5355.
11. Newlon, C. S., Collins, I., Dershowitz, A., Deshpande, A. M., Greenfeder, S. A., Ong, L. Y. & Theis, J. F. (1993) *Cold Spring Harbor Symp. Quant. Biol.* **58,** 415–423.
12. Marahrens, Y. & Stillman, B. (1992) *Science* **255,** 817–823.
13. Rao, H., Marahrens, Y. & Stillman, B. (1994) *Mol. Cell. Biol.* **14,** 7643–7651.
14. Theis, J. F. & Newlon, C. S. (1994) *Mol. Cell. Biol.* **14,** 7652–7659.
15. Huang, R. Y. & Kowalski, D. (1996) *Nucleic Acids Res.* **24,** 816–823.
16. Walker, S. S., Malik, A. K. & Eisenberg, S. (1991) *Nucleic Acids Res.* **19,** 6255–6262.
17. Broach, J. R., Li, Y. Y., Feldman, J., Jayaram, M., Abraham, J., Nasmyth, K. A. & Hicks, J. B. (1983) *Cold Spring Harbor Symp. Quant. Biol.* **47,** 1165–1173.
18. Van Houten, J. V. & Newlon, C. S. (1990) *Mol. Cell. Biol.* **10,** 3917–3925.
19. Celniker, S. E., Sweder, K., Srienc, F., Bailey, J. E. & Campbell, J. L. (1984) *Mol. Cell. Biol.* **4,** 2455–2466.
20. Kearsey, S. (1984) *Cell* **37,** 299–307.
21. Liang, C., Weinreich, M. & Stillman, B. (1995) *Cell* **81,** 667–676.
22. Klemm, R. D., Austin, R. J. & Bell, S. P. (1997) *Cell* **88,** 493–502.
23. Bell, S. P. & Stillman, B. (1992) *Nature (London)* **357,** 128–134.
24. Bell, S. P., Kobayashi, R. & Stillman, B. (1993) *Science* **262,** 1844–1849.
25. Fox, C. A., Loo, S., Dillin, A. & Rine, J. (1995) *Genes Dev.* **9,** 911–924.
26. Loo, S., Fox, C. A., Rine, J., Kobayashi, R., Stillman, B. & Bell, S. (1995) *Mol. Biol. Cell* **6,** 741–756.
27. Rowley, A., Cocker, J. H., Harwood, J. & Diffley, J. F. (1995) *EMBO J.* **14,** 2631–2641.
28. Rao, H. & Stillman, B. (1995) *Proc. Natl. Acad. Sci. USA* **92,** 2224–2228.
29. Van Houten, J. V. (1990) Ph.D. thesis (University of Medicine and Dentistry of New Jersey, Newark).
30. Sikorski, R. S. & Hieter, P. (1989) *Genetics* **122,** 19–27.
31. Newlon, C. S., Lipchitz, L. R., Collins, I., Deshpande, A., Devenish, R. J., Green, R. P., Klein, H. L., Palzkill, T. G., Ren, R. B., Synn, S. & Woody, S. T. (1991) *Genetics* **129,** 343–357.
32. Kunkel, T. A., Bebenek, K. & McClary, J. (1991) *Methods Enzymol.* **204,** 125–139.
33. Boeke, J. D., Trueheart, J., Natsoulis, G. & Fink, G. R. (1987) *Methods Enzymol.* **154,** 164–175.
34. Ausubel, F. M., Brent, R., Kingston, R. E., Moore, D. D., Seidman, J. G., Smith, J. A. & Struhl, K. (1994) *Current Protocols in Molecular Biology* (Wiley, New York).
35. Richterich, P., Lakey, N. D., Lee, H. M., Mao, J. I., Smith, D. & Church, G. M. (1995) *Nucleic Acids Res.* **23,** 4922–4923.
36. Greenfeder, S. A. & Newlon, C. S. (1992) *Mol. Biol. Cell* **3,** 999–1013.
37. Shirahige, K., Iwasaki, T., Rashid, M. B., Ogasawara, N. & Yoshikawa, H. (1993) *Mol. Cell. Biol.* **13,** 5043–5056.
38. Friedman, K. L., Diller, J. D., Ferguson, B. M., Nyland, S. V., Brewer, B. J. & Fangman, W. L. (1996) *Genes Dev.* **10,** 1595–1607.
39. Newlon, C. S. & Theis, J. F. (1993) *Curr. Opin. Genet. Dev.* **3,** 752–758.
40. Friedman, K. L. & Brewer, B. J. (1995) *Methods Enzymol.* **262,** 613–627.
41. Bouton, A. H. & Smith, M. M. (1986) *Mol. Cell. Biol.* **6,** 2354–2363.
42. Walker, S. S., Francesconi, S. C. & Eisenberg, S. (1990) *Proc. Natl. Acad. Sci. USA* **87,** 4665–4669.
43. Celniker, S. E. & Campbell, J. L. (1982) *Cell* **31,** 201–213.
44. Miller, C. A. & Kowalski, D. (1993) *Mol. Cell. Biol.* **13,** 5360–5369.
45. Brand, A. H., Micklem, G. & Nasmyth, K. (1987) *Cell* **51,** 709–719.
46. Palzkill, T. G. & Newlon, C. S. (1988) *Cell* **53,** 441–450.
47. Huang, T. & Campbell, J. L. (1995) *J. Biol. Chem.* **270,** 9607–9614.