

# DNA Sequence and Functional Analysis of Homologous ARS Elements of *Saccharomyces cerevisiae* and *S. carlsbergensis*

James F. Theis, Chen Yang, Christopher B. Schaefer and Carol S. Newlon

Department of Microbiology and Molecular Genetics, New Jersey Medical School and Graduate School of Biomedical Sciences, University of Medicine and Dentistry of New Jersey, Newark, New Jersey 07103

Manuscript received September 3, 1998

Accepted for publication April 1, 1999

## ABSTRACT

ARS elements of *Saccharomyces cerevisiae* are the *cis*-acting sequences required for the initiation of chromosomal DNA replication. Comparisons of the DNA sequences of unrelated ARS elements from different regions of the genome have revealed no significant DNA sequence conservation. We have compared the sequences of seven pairs of homologous ARS elements from two *Saccharomyces* species, *S. cerevisiae* and *S. carlsbergensis*. In all but one case, the *ARS308-ARS308<sup>cat</sup>* pair, significant blocks of homology were detected. In the cases of *ARS305*, *ARS307*, and *ARS309*, previously identified functional elements were found to be conserved in their *S. carlsbergensis* homologs. Mutation of the conserved sequences in the *S. carlsbergensis* ARS elements revealed that the homologous sequences are required for function. These observations suggested that the sequences important for ARS function would be conserved in other ARS elements. Sequence comparisons aided in the identification of the essential matches to the ARS consensus sequence (ACS) of *ARS304*, *ARS306*, and *ARS310<sup>cat</sup>*, though not of *ARS310*.

ARS elements of *Saccharomyces cerevisiae* are the best understood eukaryotic origins of DNA replication. Much of our knowledge of ARS structure is derived from the detailed analysis of a few ARS elements (Figure 1). ARS activity depends on the ARS consensus sequence (ACS; included in domain A) and on additional sequences 3' to the T-rich strand of the ACS, referred to as domain B (reviewed by Newlon 1996). Some ARS elements (e.g., *ARS121*, *ARS1501*, and *HMR-E*) have stimulatory sequences on the opposite side of the ACS, domain C (Figure 1 and references therein). The ARS activity of *HMR-E*, which also functions as the silencer of mating-type information at the *HMR* locus, has been studied most thoroughly in the context of a construct, the synthetic silencer, that maintains the spacing and DNA sequence of known protein-binding sites but scrambles DNA sequences surrounding these sites (McNally and Rine 1991).

Mutations in the ACS decrease or abolish ARS function (reviewed by Newlon 1996) by decreasing binding of the origin recognition complex (ORC), the yeast initiator protein (reviewed by Dutta and Bell 1997) that recognizes the ACS and several bases to either side (Bell and Stillman 1992; Lee and Bell 1997). ORC also recognizes bases within the B1 element (Rao and Stillman 1995; Rowley *et al.* 1995; Lee and Bell 1997), though there are mutations in the B1 element of *ARS1* whose associated phenotypes cannot be explained by a

reduction in ORC binding (Rao and Stillman 1995). Although ORC usually recognizes a single ACS within an ARS element, it is often not possible to identify the essential ACS by DNA sequence inspection because most ARS elements contain multiple matches to the ACS, and ORC does not necessarily recognize the best match to the ACS (Theis and Newlon 1997).

The precise function of the B2 elements of *ARS1* and *ARS307* is unknown, though these elements can substitute for one another (Rao *et al.* 1994). Consistent with the proposal that the B2 element is contained within a DNA unwinding element (Huang and Kowalski 1996), Bielinsky and Gerbi (1998) have mapped the 5' ends of nascent strands to this region of *ARS1*, indicating that the first primers are laid down in this interval. The function of the B4 element of *ARS305* is also unknown, and it cannot be substituted by any of the domain B elements of *ARS1* (Huang and Kowalski 1996).

Many of the ARS elements examined contain binding sites for transcription factors Abf1p and/or Rap1p, which act to enhance ARS function by an unknown mechanism. The B3 element of *ARS1* is a binding site for Abf1p (Marahrens and Stillman 1992), and two Abf1p binding sites constitute the replication enhancer of *ARS121* (Walker *et al.* 1990). *ARS1501* also contains a pair of Abf1p binding sites, but they appear not to contribute to ARS activity; instead, a different enhancer (REN1501) activates this ARS (Raychaudhuri *et al.* 1997). The Abf1p binding site in the synthetic *HMR-E* is dispensable for origin function, while mutation of the Rap1p binding site abolishes origin function (Fox *et al.* 1993).

Corresponding author: Carol S. Newlon, Department of Microbiology and Molecular Genetics, UMDNJ-New Jersey Medical School, 185 S. Orange Ave., Newark, NJ 07103. E-mail: newlon@umdnj.edu

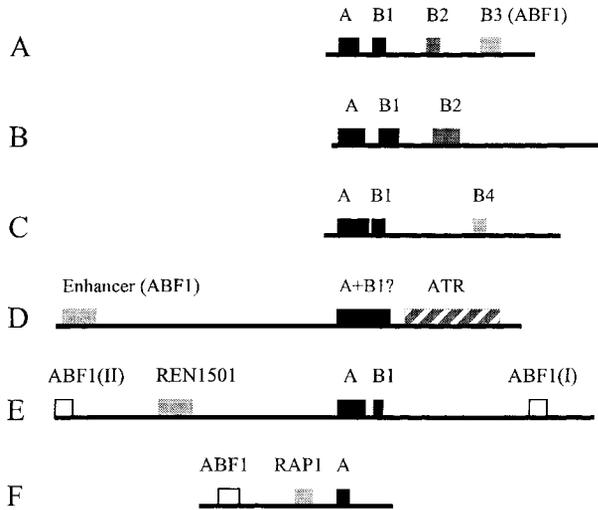


Figure 1.—Comparison of ARS elements. Line diagrams of the ARS elements are drawn to the same scale. (A) *ARS1* with functional elements as described by Marahrens and Stillman (1992); (B) *ARS307* (Rao *et al.* 1994; Theis and Newlon 1994); (C) *ARS305* (Huang and Kowalski 1996); and (D) *ARS121* (Walker *et al.* 1991). The region labeled A + B1 corresponds to the “core” element. ATR is the “A + T rich domain.” Enhancer is the replication enhancer, which consists of a pair of binding sites for Abf1p. (E) *ARS1501* (Raychaudhuri *et al.* 1997). The region labeled A corresponds to the essential match to the ACS plus the d1 linker substitution. The element labeled B1 corresponds to the d3 linker substitution; this assignment is based on the similar spacing from the ACS. The Abf1p binding sites and the replication enhancer (REN1501) are also indicated. (F) *HMR-E* synthetic silencer (McNally and Rine 1991; Fox *et al.* 1993). The region labeled A corresponds to the ACS. The binding sites for Rap1p and Abf1p are also indicated.

Our view of these ARS elements is based on extensive mutational analysis of their sequences. It should be possible to identify important functional elements by comparing the DNA sequence of a single ARS element with its homologs in other species. Comparisons of evolutionarily-related intergenic regions in *Drosophila* have been useful in the identification of promoter elements (Pelham and Bienz 1982; Blackman and Meselson 1986), and a similar strategy has been used to identify DNA sequences important for the function of the *E. coli* replication origin, *oriC* (Zyskind *et al.* 1983), as well as DNA elements important for the function of a recombination enhancer on *S. cerevisiae* chromosome III that controls the directionality of mating-type switching (Wu *et al.* 1998). Here we report on the use of this approach to identify DNA sequences important for ARS function.

For these studies we have made use of seven ARS elements that we identified on a chromosome III isolated from Carlsberg Brewery production strain 244, hereafter referred to as the *S. carlsbergensis* chromosome (Yang *et al.* 1999). This chromosome has diverged in nucleotide sequence while maintaining homologous genes in the same relative order as *S. cerevisiae* chromo-

some III (referenced in Yang *et al.* 1999). Our finding that each of these seven ARS elements is in an intergenic region flanked by open reading frames (ORFs) homologous to the ORFs that flank *S. cerevisiae* chromosome III ARS elements provides strong evidence that these pairs of ARS elements are related by descent (Newlon *et al.* 1991; Yang *et al.* 1999). In contrast to previous sequence analyses of unrelated *S. cerevisiae* ARS elements, which revealed little or no sequence homology beyond matches to the ACS, our analysis of homologous ARS elements revealed significant sequence conservation. In most cases, DNA sequence conservation allowed the essential match to the ACS to be distinguished from nonessential matches and highlighted additional sequences in domain B. However, the conservation of domain B sequences was not always apparent, suggesting that there may be considerable flexibility in the requirements for accessory sequences or that they are short sequences not readily identified by such comparisons.

## MATERIALS AND METHODS

**Strains and plasmids:** Routine plasmid manipulations were performed using *Escherichia coli* strain DH5 $\alpha$  (Life Technologies, Grand Rapids, NY). *E. coli* strain CJ236 (Kunkel *et al.* 1991) was used to prepare single-stranded DNA for oligo-directed mutagenesis. ARS activity was assessed by transforming strain *S. cerevisiae* strain 1C6 (Theis and Newlon 1994; ATCC no. 201543). The *S. monacensis* type strain was obtained from the American Type Culture Collection (no. 76670; an isolate of CBS1503).

The plasmids constructed for assays of ARS activity and plasmid stability made use of the shuttle vectors pRS326 (Theis and Newlon 1994) and pRS333 (Yang *et al.* 1999). pRS326-based plasmids were used for *S. cerevisiae* transformations, selecting for expression of *URA3*. pRS333-based plasmids were used for transformation of *S. monacensis*, selecting for resistance to G418. Both plasmids carry centromeres to direct plasmid segregation. The *ARS305<sup>carl</sup>* constructs were derived from a plasmid created by ligating the 2.0-kb *FspI* fragment of p305K/D0.55 (Yang *et al.* 1999) to the 4.7-kb *FspI* fragment of pRS326. Oligo-directed mutations were made in this plasmid. The pRS333 plasmids carrying *ARS305<sup>carl</sup>* and its derivatives were created by ligating the 2.2-kb *BsaHI-NgoMI* fragments of the appropriate pRS326 derivatives to the 4.1-kb *BsaHI-NgoMI* fragment of pRS333. The *ARS307<sup>carl</sup>* plasmids were constructed similarly from p307k/s2'/2; the *ARS308<sup>carl</sup>* plasmid was constructed from p308R/H0.3; the *ARS310<sup>carl</sup>* plasmids were constructed from p310Bg/B0.56 (Yang *et al.* 1999). The pRS326 plasmid carrying the *ARS306 XhoI-BglII* fragment was constructed by swapping the *FspI* fragment of the pRS306 plasmid carrying this fragment, pAD16 (Deshpande 1995). The pRS326 plasmid carrying the *ARS306 HindIII-BglII* fragment was constructed by ligating the 220-bp *HindIII-BglII* fragment of pAD15 (Deshpande 1995) to pRS326 digested with *HindIII* plus *BamHI*. pCS12 was constructed by cloning the ~450-bp *SpeI-DraI* fragment, with *BamHI* linkers, of D10B (Newlon *et al.* 1991) into *BamHI*-cut pRS326.

**Site directed mutagenesis:** Oligonucleotide-directed mutagenesis was performed by the method of Kunkel *et al.* (1991). The oligonucleotides used for the mutagenesis were synthesized by the Molecular Resources Facility of UMDNJ–New Jersey Medical School. The mutations are indicated in the





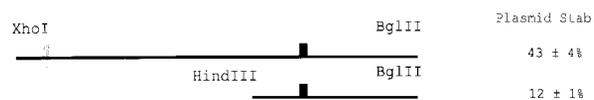
mutation in *ARS305* is that the construct used to analyze *ARS305<sup>carl</sup>* was truncated at position 245 and therefore lacked the B4 element.

***ARS309<sup>carl</sup>*:** Comparison of the DNA sequences of *ARS309* and *ARS309<sup>carl</sup>* revealed 69% identity over 177 bp (data not shown). Within this region is a 33-bp block containing the essential ACS of *ARS309* (Theis and Newlon 1997; Figure 3C). This match is unusual in two respects. First, this 9 of 11 match to the ACS was essential, despite the presence of four 10 of 11 matches in the same fragment. Second, this match contains a C at position 9. All previously identified essential matches have a T at this position (Newlon and Theis 1993); a T-to-C change at this position abolished the activity of *ARS307* (Van Houten and Newlon 1990), and base modification at this position eliminated ORC binding to *ARS1* (Lee and Bell 1997). Therefore, it was surprising to find that the C at this position is conserved in *ARS309<sup>carl</sup>*. Two other positions within the aligned ACS match are changed, but neither substitution is expected to have a deleterious effect on ARS activity (Van Houten and Newlon 1990). We have not yet determined whether this match to the ACS is essential for *ARS309<sup>carl</sup>* nor have we determined functional elements other than the ACS for *ARS309*. However, the region of homology containing the essential ACS extends into the B domain far enough to encompass the expected B1 element. In addition, a 67-bp region of 87% identity is present further into the B domain.

***ARS306<sup>carl</sup>*:** Previous work on *ARS306* had localized ARS activity to a 220-bp *HindIII-BglII* fragment (Newlon *et al.* 1991). During the subcloning of this ARS element, it was noted that plasmids carrying the 530-bp *XhoI-BglII* fragment appeared more stable than those carrying the larger *HindIII-BglII* fragment, and direct assays of plasmid stability confirmed this observation (Figure 4A). The previous analysis had not identified the essential match to the ACS for *ARS306*. We used homology to *ARS306<sup>carl</sup>* to guide our analysis of *ARS306*. Overall, the two ARS elements are 62% identical over 418 bp, allowing 13 gaps (data not shown). Most of this homology lies outside of the *HindIII-BglII* fragment, but we found a perfect match to the ACS within a block of homology in this fragment. Introduction of a 2-bp change within this match abolished the activity of *ARS306* (Figure 4B), demonstrating that this conserved match is the essential match to the ACS.

The domain C region of *ARS306* includes the *XhoI-HindIII* fragment that stimulates ARS activity (Figure 4A). We found no significant alignment with the *S. carlsbergensis* sequence of the 75-bp region of domain C adjacent to the ACS. Interestingly, however, there are a number of blocks of homology farther away that fall within the stimulatory *XhoI-HindIII* fragment. One of these blocks caught our attention because it contains a pair of conserved 7-of-8 matches to the Abf1p binding site consensus, RTCRYNNNNNACG (Della Seta *et al.*

#### A Stability of *ARS306* plasmids



#### B *ARS306* mutagenesis summary

Position	Sequence	Plasmid Stab.
Conserved ACS		
375	AATTGTATTATGTTTACTCAGTATTATTGAATTAGAAATAT 421	43 ± 4%
74487	AATTGTATTATATTATTTAGTACTTCTTGAGTTTACATAT 74527	
Conserved ABF1 sites		
24	.TTTACGTAATGTTTCGGCTAAGACGCCACAGCTGGGATTTGCTCGGATAAT 73	39 ± 3%
74132	CTTTACGTAATTTTCAGCTAAGACGCCACAGCTCGGATTTTTCGAAATAT 74182	
Non-conserved ABF1 site		
74269	CTGTTTCGTTATTGATGTAGGAGAA	44 ± 4%
	CTGTT...CAGCT...GGAGAA	
Double ABF1 site mutant		
		34 ± 3%

Figure 4.—Analysis of *ARS306*. (A) Plasmid stabilities of two *ARS306* subclones. Gray boxes mark the positions of the matches to the Abf1p binding site consensus, and the black boxes mark the perfect matches to the ACS. Plasmid stability was measured as described in the legend to Figure 3. (B) In this summary of the phenotypes of mutant derivatives of *ARS306*, the top line is the *ARS306<sup>carl</sup>* sequence. Matches to the ACS and the Abf1p binding site consensus are underlined.

1990). The Abf1p binding sites in *ARS121*, which lie in domain C, can function as distance- and position-independent enhancers of *ARS121* function (Walker *et al.* 1989, 1990). Therefore, it seemed possible that these putative Abf1p binding sites were responsible for the stimulatory activity of the *XhoI-BglII* fragment on *ARS306*. To test this hypothesis, the sites were deleted. This deletion mutation had no effect on *ARS306* activity (Figure 4B). The *ARS306* sequence contains an additional 7 of 8 match to the Abf1p binding site consensus, which was also deleted. This mutation also had no significant effect on *ARS306* activity, either alone or in combination with the previously described deletion (Figure 4B). We conclude that these putative Abf1p binding sites play no role in *ARS306* activity, and other sequences within the *XhoI-HindIII* fragment are likely to be responsible for the enhanced plasmid stability seen for the *XhoI-BglII* fragment.

***ARS310<sup>carl</sup>*:** Previous work had localized *ARS310* to an 850-bp *EcoRV* fragment (positions 166214–167063; Palzkill *et al.* 1986; Newlon *et al.* 1991). A total of 181 bp of this sequence could be aligned significantly with the *S. carlsbergensis* sequence (59% identity allowing 5 gaps; data not shown). This region of *ARS310<sup>carl</sup>* contains a single 10-of-11 match to the ACS, and mutation of this match abolished ARS activity (Figure 3D). The ho-

mologous match in *ARS310* contains an A at position 10 of the ACS, an alteration that inactivates *ARS307* (Van Houten and Newlon 1990). No other ARS element has a base other than T at position 10 of its essential ACS (Theis and Newlon 1997). A change in this match caused a very slight reduction in plasmid stability, indicating that this match is not essential for *ARS310* activity. Our further analysis of *ARS310* is presented elsewhere (J. F. Theis and C. S. Newlon, unpublished results).

***ARS304<sup>carl</sup>*:** *ARS304* has been localized to a 460-bp *SpeI-DraI* fragment (30351–30641; Newlon *et al.* 1991). It lies in the 2.5-kb intergenic region between *KAR4* and YCL054W. Plasmids carrying *ARS304* are unstable, suggesting that it is a weak ARS (I. Collins and C. S. Newlon, unpublished observation), and it is not detectably active as a chromosomal replication origin (Dubey *et al.* 1991; Newlon *et al.* 1993). Remarkably, this region is highly conserved in *S. carlsbergensis*. A block of 68% identity, allowing six gaps, extends over 501 bp and contains half of the *SpeI-DraI* fragment (data not shown). This is the most impressive region of sequence homology that we detected in an intergenic region.

The sequence of *ARS304* contains one 10-of-11 and six 9-of-11 matches to the ACS. Only one 9-of-11 match is conserved in *ARS304<sup>carl</sup>*. Alteration of this match inactivated *ARS304*, establishing this 9-of-11 match as the essential ACS in *ARS304* (Figure 3E). The role of this match to the ACS in the function of *ARS304<sup>carl</sup>* has not been tested.

***ARS308<sup>carl</sup>*:** The weak, centromere-associated origin, *ARS308*, has been localized to the *Sau3A-BamHI* fragment that contains *CEN3*. Its location in this fragment places *ARS308* within 0.5 kb of *CEN3*. *CEN3<sup>carl</sup>* is contained in a 1.1-kb *HindIII* fragment, and *ARS308<sup>carl</sup>* lies in the adjacent 0.3-kb *HindIII-EcoRI* fragment (Yang *et al.* 1999). Therefore, *ARS308<sup>carl</sup>* is 0.7 to 1.0 kb farther away from its centromere than is *ARS308*. In marked contrast to the other ARS elements examined, no significant homology was found between *ARS308<sup>carl</sup>* and sequences within 5 kb of *CEN3*. We were, therefore, unable to use sequence homology to identify functional elements in these ARS elements. The sequence of *ARS308<sup>carl</sup>* contained a single 14 of 17 match to the EACS (Theis and Newlon 1997), and alteration of this match abolished ARS activity (Figure 3F). There is an exact match to the ACS, which is also the best match to the EACS, in the *Sau3A-BamHI* fragment containing *ARS308*. This match lies within the CDEII element of *CEN3*. Alteration of this match did not inactivate *ARS308* (Figure 3F). Given its location within the centromere and the unusual chromatin structure associated with the centromere itself (Basrai and Hieter 1995), it is not surprising that the essential ACS lies elsewhere in the fragment.

**Analysis of ARS function in *S. monacensis*:** The preceding analysis was performed in *S. cerevisiae* due to the

**TABLE 1**  
**Comparison of ARS activity in *S. cerevisiae***  
**and *S. monacensis***

ARS element	Mutation	<i>S. cerevisiae</i> phenotype <sup>a</sup>	<i>S. monacensis</i> phenotype <sup>a</sup>
<i>ARS307<sup>carl</sup></i>	None	Ars <sup>+</sup>	Ars <sup>+</sup>
	ACS	Ars <sup>-</sup>	Ars <sup>-</sup>
<i>ARS305<sup>carl</sup></i>	None	Ars <sup>+</sup>	Ars <sup>+</sup>
	ACS	Ars <sup>-</sup>	Ars <sup>-</sup>
	Box 3'	Ars <sup>-</sup>	Ars <sup>-</sup>
	B1	Ars <sup>-</sup>	Ars <sup>-</sup>
<i>ARS310<sup>carl</sup></i>	None	Ars <sup>+</sup>	Ars <sup>+</sup>
	ACS	Ars <sup>-</sup>	Ars <sup>-</sup>

Mutations refer to the mutations described in Figures 3 and 4. pRS333 plasmids carrying the ARS construct indicated were used to transform *S. cerevisiae* and *S. monacensis* to G418 resistance.

<sup>a</sup> Ars<sup>+</sup> indicates that transformants were obtained at high frequency, while Ars<sup>-</sup> indicates that transformants were obtained rarely or not at all.

ease of experimental design. We wished to assess the function of the *S. carlsbergensis* ARS elements in their natural background. For the reasons described by Yang *et al.* (1999), the brewing strain itself was not acceptable, and we chose *S. monacensis* for our analyses. *ARS305<sup>carl</sup>*, *ARS307<sup>carl</sup>*, *ARS310<sup>carl</sup>*, and some of their mutant derivatives were transferred into the vector, pRS333; ARS-containing fragments in this vector transform yeast to G418 resistance (Yang *et al.* 1999). All three wild-type ARS elements were functional in *S. monacensis* (Table 1). Furthermore, mutations that inactivated these ARS elements, as tested in *S. cerevisiae*, also inactivated them in *S. monacensis*. It appears that the same sequences are required for ARS function in the two species.

## DISCUSSION

ARS elements play a key role in chromosome maintenance, functioning as chromosomal origins of DNA replication in the budding yeast *S. cerevisiae*. A small number of ARS elements from different regions of the yeast genome have been analyzed in detail, yielding a partial picture of the DNA sequence elements required for ARS activity, but virtually nothing is known of their evolution. In this article we presented our analysis of the sequences of seven ARS elements derived from chromosome III of *S. carlsbergensis* and their comparison to homologous ARS elements from chromosome III of *S. cerevisiae*. As shown in Figure 2, three of the sequences, *ARS305<sup>carl</sup>*, *ARS307<sup>carl</sup>*, and *ARS310<sup>carl</sup>*, contain ORFs homologous (60 to 80% identity) to ORFs present in the *S. cerevisiae* chromosome III sequence. The presence of these homologous ORFs localizes these *S. carlsbergensis* ARS elements to the same intergenic regions as their *S. cerevisiae* counterparts. The remaining *S. carlsbergensis* ARS ele-

ments were anchored to the *S. cerevisiae* sequence by sequencing the ends of larger clones to identify homologous ORFs (Yang *et al.* 1999). Comparison of the sequences of the *S. carlsbergensis* ARS elements to the complete *S. cerevisiae* genome sequence using the BLAST algorithm (Altschul *et al.* 1990) provided additional strong evidence that these pairs of ARS elements are evolutionarily related. Significant homologues were found for five of the seven *S. carlsbergensis* ARS elements, and, in each case, the homology was detected in a region corresponding to the appropriate *S. cerevisiae* ARS element. Only *ARS308<sup>carl</sup>* and the intergenic region of *ARS305<sup>carl</sup>* failed to give significant BLAST scores (data not shown).

Comparisons of the limited coding regions of these clones with their *S. cerevisiae* homologues revealed patterns of DNA sequence conservation consistent with evolutionary constraints on maintaining ORFs. The small number of gaps in the alignment were in multiples of 3 bp, and approximately one-third of the single base changes were in the third position of codons that did not alter the predicted amino acid sequence of the protein product.

In contrast to the coding regions, alignments of the DNA sequences of the intergenic regions containing the ARS elements showed much lower levels of DNA sequence identity and the presence of many more gaps. Nevertheless, except for *ARS308<sup>carl</sup>*, it was possible to find regions of high identity that were constant, regardless of the software used to do the alignment, and that were insensitive to variation in parameters.

**DNA sequence conservation of domain A and element B1:** As the core of the binding site for ORC, the ACS is the most easily recognized conserved feature of ARS elements. However, most ARS elements contain multiple near matches to the ACS, making it difficult to identify the essential match by DNA sequence inspection alone. Because several ARS elements have an essential ACS that matches the consensus at only 9 of 11 positions and because the best match to the ACS may not be the essential one, any sequence matching at least 9 positions has the potential to be the essential match for the ARS element under study. Each of the ARS elements described in this study has five or more matches to the ACS that meet this criterion. A striking feature of the DNA sequence conservation is that only the essential match to the ACS was found in a region of significant homology. This pattern first became apparent in the analysis of *ARS305<sup>carl</sup>* and *ARS307<sup>carl</sup>*, whose *S. cerevisiae* homologues had previously been analyzed by linker-substitution mutagenesis (Rao *et al.* 1994; Theis and Newl on 1994; Huang and Kowalski 1996; Figure 3, A and B). This pattern of DNA sequence conservation extended to the third ARS element whose essential ACS had been identified, *ARS309* (Theis and Newl on 1997; Figure 3C). Within these regions of homology, positions

that had been shown by mutagenesis to be important for ARS activity were generally conserved.

The functional significance of the DNA sequence conservation was tested by constructing a series of base substitutions in *ARS307<sup>carl</sup>* and *ARS305<sup>carl</sup>* and determining their effect on ARS activity. In every case the mutations resulted in phenotypes similar to their *S. cerevisiae* counterparts, demonstrating directly that sequences important for ARS function in the *S. cerevisiae* homologue are conserved in *S. carlsbergensis*. The observation that other near matches to the ACS were not aligned in these pairs of homologous ARS elements suggests that they are not required for ARS activity, which is consistent with the mutagenesis of *ARS307* (Rao *et al.* 1994; Theis and Newl on 1997), *ARS305* (Huang and Kowalski 1996), and *ARS1* (Marahrens and Stillman 1992). Their occurrence may reflect simply the relatively high A + T content of the sequences.

The regions of homology containing the essential match to the ACS extended into the B domain far enough to include the B1 element, which is also known to contribute to ORC binding (Rao and Stillman 1995; Lee and Bell 1997). ORC binding to *ARS1* is reduced or abolished by modifications of nucleotides within the ACS, at positions +3 to +5 (between the ACS and element B1) and at +16 to +18, within the B1 element (Lee and Bell 1997). While the ACS is well conserved, it is clear from both mutational analysis of the ACS (Van Houten and Newl on 1990; Bell and Stillman 1992) and a comparison of the essential matches of known ARS elements (Theis and Newl on 1997) that some variation is tolerated. Moreover, the *ARS1* nucleotides that ORC contacts at positions +3 to +5 and +16 to +18 are not highly conserved, with the exception of the T:A base pair at position +18 (Theis and Newl on 1997). The sequence variations observed in the conserved regions of *ARS305* and *ARS307* are consistent with these generalizations. The two substitutions observed in the essential match of *ARS307<sup>carl</sup>* are expected to be neutral for ARS activity. Similarly, the conserved substitution at position 3 of the *ARS305* and *ARS305<sup>carl</sup>* essential matches is found in *ARS602* (Shirahige *et al.* 1993). However, the change at position 6 of the *ARS305<sup>carl</sup>* ACS is a change that reduced but did not abolish *ARS307* activity (Van Houten and Newl on 1990). The most surprising finding was the conservation of the nonconsensus C at position 9 of the *ARS309* essential match in *ARS309<sup>carl</sup>*. This change abolishes the activity of *ARS307* (Van Houten and Newl on 1990) and is found in no other known essential match (Theis and Newl on 1997).

In our analysis of the remaining ARS elements, we exploited the homology to *S. carlsbergensis* to identify functional elements in their *S. cerevisiae* counterparts. Sequence conservation correctly identified the essential ACS of *ARS304* and *ARS306*. In the case of *ARS310*, we identified a 10-of-11 match to the ACS conserved in

*ARS310<sup>carl</sup>*. As expected, a mutation in the conserved match of *ARS310<sup>carl</sup>* abolished activity (Figure 3D). However, in contrast to our previous observations, a mutation in this sequence did not have a significant effect on *ARS310* function, perhaps because this match contains an A at position 10 of the ACS, an alteration known to abolish the function of *ARS307* (Van Houten and Newlon 1990). It is intriguing that an ARS element is found within 500 bp of the start of YCR026C, in both *S. cerevisiae* (J. F. Theis and C. S. Newlon, unpublished results) and *S. carlsbergensis*, yet the sequences responsible for activity appear to be different in the two species.

In contrast to the other six ARS elements studied, no significant sequence identity was detected between *ARS308* and *ARS308<sup>carl</sup>*, *i.e.*, no alignments scored better than the alignments generated with random sequences of the same base composition. These two ARS elements were considered homologues because both were isolated in association with their centromere (Yang *et al.* 1999). We identified the essential match to the ACS of *ARS308<sup>carl</sup>*, but the mutation made in *ARS308*, the ACS match within the CDEII element of *CEN3*, did not abolish ARS activity (Figure 3F). This situation is similar to that seen with *ARS310* and *ARS310<sup>carl</sup>*, in which the position of the ARS element is conserved but the sequences essential for ARS function are not.

**Conservation of ARS structure in *S. monacensis*:** In studies of ARS elements of other yeasts, for example *Kluyveromyces lactis* (Fabiani *et al.* 1996) and *S. pombe* (Maundrell *et al.* 1985), DNA fragments have been identified that have ARS activity in *S. cerevisiae* as well as the species from which they were derived. However, further analysis revealed that the DNA sequences that promote ARS activity in the two species are different. Therefore, it was important to test the *S. carlsbergensis* ARS elements in their native host. *ARS305<sup>carl</sup>*, *ARS307<sup>carl</sup>*, and *ARS310<sup>carl</sup>* are all active, as judged by their ability to promote high frequency transformation and extra-chromosomal maintenance of plasmids in *S. monacensis*. Importantly, mutations that inactivated these ARS elements in *S. cerevisiae* also inactivated them in *S. monacensis*, indicating that the same functional elements are used. Thus it appears that replication origins are highly conserved in these closely related yeasts.

**Additional DNA sequence conservation in domains B and C:** One of the motivations for undertaking this study was the hope that comparisons of homologous ARS elements would aid in the identification of DNA sequences important for ARS activity. Aside from the clear conservation of the ORC-binding region discussed above, the homologous pairs of ARS elements varied widely in the presence of additional significant blocks of sequence homology. No significant sequence homology was detected outside of the region containing domain A and the B1 element of *ARS305*, which is consistent with the results of the linker substitution mutagenesis of Huang and Kowalski (1996), which identified only

one linker substitution that affected *ARS305* activity outside of domains A and B1. The domain B regions of *ARS307* and *ARS309*, and the domain C region of *ARS306* were found to contain blocks of significant homology 70 to 100 bp in length. These regions have not yet been analyzed in detail to assess their contribution to ARS activity.

It is unlikely that the sequence alignment algorithms used for these studies would have been useful in identifying conserved sequences the size of typical protein binding sites if they were not embedded in longer regions of homology. Although we recognized conservation of part of the *ARS307* B2 element and the Abf1p binding sites near *ARS306*, it was their location in larger regions of homology and existing experimental data that brought them to our attention. Having additional homologues of these ARS elements would aid in the identification of short conserved elements.

***ARS304*:** The extensive region of conservation was unexpected because *ARS304* is inactive as a chromosomal replication origin (Dubey *et al.* 1991; Newlon *et al.* 1991). Moreover, the 501-bp conserved region is substantially larger than the sequences typically required for ARS function (~100 bp; reviewed by Newlon 1996). *ARS304* lies adjacent to YCL054W within an unusually large intergenic region flanked on the other end by *KAR4* (Yang *et al.* 1999). Although the homology could reflect a complex regulatory region controlling the expression of one of the flanking genes, there is no precedent for regulatory regions of this size in yeast.

A more likely explanation for the high degree of sequence conservation is the presence of the recombination enhancer (RE) in the same intergenic region (Wu and Haber 1995; Szeto *et al.* 1997; Wu *et al.* 1998). The RE, responsible for the preferential use of *HML* by **a** cells during mating-type switching, activates recombination on the entire left arm of chromosome III. A 700-bp fragment from the *KAR4* end of the intergenic region that excludes *ARS304* was initially defined as being necessary and sufficient for RE activity (Wu and Haber 1995), and a 244-bp region of this fragment, which can be interchanged with *S. carlsbergensis* DNA, contains most of the RE activity (Wu *et al.* 1998). However, while not essential for RE function, deletions within or extending into the conserved block containing *ARS304* diminish the preference for *HML* in **a** cells (Szeto *et al.* 1997). A conserved Mat $\alpha$ 2-Mcm1 operator site within the minimal RE is required for the activity of the minimal fragment (Wu *et al.* 1998). A second Mat $\alpha$ 2-Mcm1 operator site is conserved (data not shown) in the *S. carlsbergensis* sequence within the *ARS304* homology. The two operator sites in *S. cerevisiae* appear to be redundant, as deletion of both sites is required to affect donor preference in **a** cells (Szeto *et al.* 1997). Thus it seems likely that the marked sequence conservation in this region results from the presence of accessory elements for the RE. Whether *ARS304* itself plays a role in RE

function has not been examined, but the mutation in its essential match to the ACS provides a means of testing its role.

We thank Dr. Lynn S. Ripley for helpful discussions and Dr. Steve Oliver for comments on the manuscript. We also thank George Arhin, Brad Hornbeck, Karen Fleshman, and Shanaz Ghandhi for their work on this project as summer or rotation students. This work was supported by National Institutes of Health research grant GM35679 to C.S.N. Partial support for C.Y. and C.B.S. was provided by fellowships from UMDNJ-Graduate School of Biomedical Sciences.

#### LITERATURE CITED

- Altschul, S. F., W. Gish, W. Miller, E. W. Myers and D. J. Lipman, 1990 Basic local alignment search tool. *J. Mol. Biol.* **215**: 403–410.
- Basrai, M. A., and P. Hieter, 1995 Is there a unique form of chromatin at *Saccharomyces cerevisiae* centromeres? *Bioessays* **17**: 669–672.
- Becker, D. M., and L. Guarente, 1991 High-efficiency transformation of yeast by electroporation. *Methods Enzymol.* **194**: 182–187.
- Bell, S. P., and B. Stillman, 1992 ATP-dependent recognition of eukaryotic origins of DNA replication by a multiprotein complex. *Nature* **357**: 128–134.
- Bielinsky, A. K., and S. A. Gerbi, 1998 Discrete start sites for DNA synthesis in the yeast *ARS1* origin. *Science* **279**: 95–98.
- Blackman, R. K., and M. Meselson, 1986 Interspecific nucleotide sequence comparisons used to identify regulatory and structural features of the *Drosophila hsp82* gene. *J. Mol. Biol.* **188**: 499–515.
- Della Seta, F., I. Treich, J. M. Buhler and A. Sentenac, 1990 ABF1 binding sites in yeast RNA polymerase genes. *J. Biol. Chem.* **265**: 15168–15175.
- Deshpande, A. M., 1995 The initiation and movement of DNA replication forks in *Saccharomyces cerevisiae*. Ph.D. Thesis, UMDNJ-Graduate School of Biomedical Sciences, Newark, NJ.
- Dubey, D. D., L. R. Davis, S. A. Greenfeder, L. Y. Ong, J. G. Zhu *et al.*, 1991 Evidence suggesting that the ARS elements associated with silencers of the yeast mating-type locus *HML* do not function as chromosomal DNA replication origins. *Mol. Cell. Biol.* **11**: 5346–5355.
- Dutta, A., and S. P. Bell, 1997 Initiation of DNA replication in eukaryotic cells. *Annu. Rev. Cell Dev. Biol.* **13**: 293–332.
- Fabiani, L., L. Frontali and C. S. Newlon, 1996 Identification of an essential core element and stimulatory sequences in a *Kluyveromyces lactis* ARS element, *KARS101*. *Mol. Microbiol.* **19**: 756–766.
- Fox, C. A., S. Loo, D. H. Rivier, M. A. Foss and J. Rine, 1993 A transcriptional silencer as a specialized origin of replication that establishes functional domains of chromatin. *Cold Spring Harbor Symp. Quant. Biol.* **58**: 443–455.
- Huang, R. Y., and D. Kowalski, 1996 Multiple DNA elements in *ARS305* determine replication origin activity in a yeast chromosome. *Nucleic Acids Res.* **24**: 816–823.
- Kunkel, T. A., K. Bebenek and J. McClary, 1991 Efficient site-directed mutagenesis using uracil-containing DNA. *Methods Enzymol.* **204**: 125–139.
- Laurenson, P., and J. Rine, 1992 Silencers, silencing, and heritable transcriptional states. *Microbiol. Rev.* **56**: 543–560.
- Lee, D. G., and S. P. Bell, 1997 Architecture of the yeast origin recognition complex bound to origins of DNA replication. *Mol. Cell. Biol.* **17**: 7159–7168.
- Marahrens, Y., and B. Stillman, 1992 A yeast chromosomal origin of DNA replication defined by multiple functional elements. *Science* **255**: 817–823.
- Maudrell, K., A. P. Wright, M. Piper and S. Shall, 1985 Evaluation of heterologous ARS activity in *S. cerevisiae* using cloned DNA from *S. pombe*. *Nucleic Acids Res.* **13**: 3711–3722.
- McNally, F. J., and J. Rine, 1991 A synthetic silencer mediates SIR-dependent functions in *Saccharomyces cerevisiae*. *Mol. Cell. Biol.* **11**: 5648–5659.
- Newlon, C. S., 1996 DNA replication in yeast, pp. 873–914 in *DNA Replication in Eukaryotic Cells*, edited by M. L. DePamphilis. Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY.
- Newlon, C. S., and J. F. Theis, 1993 The structure and function of yeast ARS elements. *Curr. Opin. Genet. Dev.* **3**: 752–758.
- Newlon, C. S., L. R. Lipchitz, I. Collins, A. Deshpande, R. J. Devenish *et al.*, 1991 Analysis of a circular derivative of *Saccharomyces cerevisiae* chromosome III: a physical map and identification and location of ARS elements. *Genetics* **129**: 343–357.
- Newlon, C. S., I. Collins, A. Dershowitz, A. M. Deshpande, S. A. Greenfeder *et al.*, 1993 Analysis of replication origin function on chromosome III of *Saccharomyces cerevisiae*. *Cold Spring Harbor Symp. Quant. Biol.* **58**: 415–423.
- Ngan, V. K., and L. Clarke, 1997 The centromere enhancer mediates centromere activation in *Schizosaccharomyces pombe*. *Mol. Cell. Biol.* **17**: 3305–3314.
- Oliver, S. G., Q. J. van der Aart, M. L. Agostoni-Carbone, M. Aigle, L. Alberghina *et al.*, 1992 The complete DNA sequence of yeast chromosome III. *Nature* **357**: 38–46.
- Palzkill, T. G., and C. S. Newlon, 1988 A yeast replication origin consists of multiple copies of a small conserved sequence. *Cell* **53**: 441–450.
- Palzkill, T. G., S. G. Oliver and C. S. Newlon, 1986 DNA sequence analysis of ARS elements from chromosome III of *Saccharomyces cerevisiae*: identification of a new conserved sequence. *Nucleic Acids Res.* **14**: 6247–6264.
- Pelham, H. R. B., and M. Bienz, 1982 A synthetic heat-shock promoter element confers heat-inducibility of the herpes simplex thymidine kinase gene. *EMBO J.* **1**: 1473–1477.
- Pillus, L., and J. Rine, 1989 Epigenetic inheritance of transcriptional states in *S. cerevisiae*. *Cell* **59**: 637–647.
- Rao, H., and B. Stillman, 1995 The origin recognition complex interacts with a bipartite DNA binding site within yeast replicators. *Proc. Natl. Acad. Sci. USA* **92**: 2224–2228.
- Rao, H., Y. Marahrens and B. Stillman, 1994 Functional conservation of multiple elements in yeast chromosomal replicators. *Mol. Cell. Biol.* **14**: 7643–7651.
- Raychaudhuri, S., R. Byers, T. Upton and S. Eisenberg, 1997 Functional analysis of a replication origin from *Saccharomyces cerevisiae*: identification of a new replication enhancer. *Nucleic Acids Res.* **25**: 5057–5064.
- Rowley, A., J. H. Cocker, J. Harwood and J. F. Diffley, 1995 Initiation complex assembly at budding yeast replication origins begins with the recognition of a bipartite sequence by limiting amounts of the initiator, ORC. *EMBO J.* **14**: 2631–2641.
- Shirahige, K., T. Iwasaki, M. B. Rashid, N. Ogasawara and H. Yoshikawa, 1993 Location and characterization of autonomously replicating sequences from chromosome VI of *Saccharomyces cerevisiae*. *Mol. Cell. Biol.* **13**: 5043–5056.
- Steiner, N. C., and L. Clarke, 1994 A novel epigenetic effect can alter centromere function in fission yeast. *Cell* **79**: 865–874.
- Szeto, L., M. K. Fafalios, H. Zhong, A. K. Vershon and J. R. Broach, 1997 *Mata2p* controls donor preference during mating type interconversion in yeast by inactivating a recombinational enhancer of chromosome III. *Genes Dev.* **11**: 1899–1911.
- Theis, J. F., and C. S. Newlon, 1994 Domain B of *ARS307* contains two functional elements and contributes to chromosomal replication origin function. *Mol. Cell. Biol.* **14**: 7652–7659.
- Theis, J. F., and C. S. Newlon, 1997 The *ARS309* chromosomal replicator of *Saccharomyces cerevisiae* depends on an exceptional ARS consensus sequence. *Proc. Natl. Acad. Sci. USA* **94**: 10786–10791.
- Van Houten, J. V., and C. S. Newlon, 1990 Mutational analysis of the consensus sequence of a replication origin from yeast chromosome III. *Mol. Cell. Biol.* **10**: 3917–3925.
- Walker, S. S., S. C. Francesconi, B. K. Tye and S. Eisenberg, 1989 The OBF1 protein and its DNA-binding site are important for the function of an autonomously replicating sequence in *Saccharomyces cerevisiae*. *Mol. Cell. Biol.* **9**: 2914–2921.
- Walker, S. S., S. C. Francesconi and S. Eisenberg, 1990 A DNA replication enhancer in *Saccharomyces cerevisiae*. *Proc. Natl. Acad. Sci. USA* **87**: 4665–4669.
- Walker, S. S., A. K. Malik and S. Eisenberg, 1991 Analysis of the interactions of functional domains of a nuclear origin of replication from *Saccharomyces cerevisiae*. *Nucleic Acids Res.* **19**: 6255–6262.
- Wu, X., and J. E. Haber, 1995 *MATa* donor preference in yeast mating-type switching: activation of a large chromosomal region for recombination. *Genes Dev.* **9**: 1922–1932.

- Wu, C., K. Weiss, C. Yang, M. A. Harris, B. K. Tye *et al.*, 1998 *Mcm1* regulates donor preference controlled by the recombination enhancer in *Saccharomyces* mating-type switching. *Genes Dev.* **12**: 1726–1737.
- Yang, C., J. F. Theis and C. S. Newlon, 1999 Conservation of ARS elements and chromosomal DNA replication origins on chromosomes III of *Saccharomyces cerevisiae* and *S. carlsbergensis*. *Genetics* **152**: 933–941.
- Zyskind, J. W., J. M. Cleary, W. S. Brusilow, N. E. Harding and D. W. Smith, 1983 Chromosomal replication origin from the marine bacterium *Vibrio harveyi* functions in *Escherichia coli*: *oriC* consensus sequence. *Proc. Natl. Acad. Sci. USA* **870**: 1164–1168.

Communicating editor: M. Johnston