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′ 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′, 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; Rowlsey 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).
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) HM RE 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; Backett 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 chromosome 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 DH5α (Life Technologies, Grand Rapids, NY). E. coli strain C236 (Kunkel et al. 1991) was used to prepare single-stranded DNA for oligonucleotide-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 strain was 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 mutagenesis were made in this plasmid. The pRS333 plasmids carrying ARS305 and its derivatives were created by ligating the 2.2-kb BsaHI-NgoM1 fragment of the appropriate pRS326 derivatives to the 4.1-kb BsaHI-NgoM1 fragment of pRS333. The ARS307 plasmids were constructed similarly from p307K/2′/2; the ARS308 plasmid was constructed from p308R/H0.3; the ARS310 plasmids were constructed from p310B/g0.56 (Yang et al. 1999). The pRS326 plasmid carrying the ARS306 XhoI-BglIII 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 BamH1. pCS12 was constructed by cloning the ~450-bp SpeI-NotI fragment, with BamH1 linkers, of D108 (Newlon et al. 1991) into BamH1-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
figures, and the sequences of the mutagenic oligonucleotides are available upon request. Mutations were confirmed by DNA sequencing.

Yeast transformations and plasmid stability assays: Yeast strains were transformed by electroporation as described by Becker and Guarente (1991). Conditions used for the electroporation of S. monacensis were described by Yang et al. (1999). Plasmid stability assays measured the fraction of plasmid-bearing cells in a culture growing under selection for the plasmid and were described previously (Palzkill and Newlon 1988).

DNA sequence analysis: Sequence comparisons were performed using software of the Wisconsin package (Genetics Computer Group) and the BLAST server link of the Saccharomyces Genome Database. The sequences of the S. carlsbergensis ARS elements have been deposited in GenBank: ARS304$^{carl}$ (AF087949), ARS305$^{carl}$ (AF087950), ARS306$^{carl}$ (AF087951), ARS307$^{carl}$ (AF087952), ARS308$^{carl}$ (AF087953), ARS309$^{carl}$ (AF087954), and ARS310$^{carl}$ (AF087955).

RESULTS

Overview of sequence conservation: We determined the DNA sequences of the smallest ARS-containing subclones identified in our analysis of seven ARS elements on S. carlsbergensis chromosome III (Yang et al. 1999). Of the seven ARS elements sequenced, three, ARS305$^{carl}$, ARS307$^{carl}$, and ARS310$^{carl}$, contain parts of four ORFs. The predicted amino acid sequences of these ORFs and their S. cerevisiae homologues are presented in Figure 2. Homology is quite high in the ORFs, 60 to 80% identity.

As expected, the overall level of sequence identity was lower in the intergenic regions, and there were more gaps in the alignments. In some regions, e.g., ARS305 and ARS308, the amount of overall sequence identity was not above that seen with random sequences of the same base composition. In other regions, the sequence conservation was quite high. Surprisingly, one such region contains ARS304, an ARS element that is not detectably active as a chromosomal replication origin (Dubey et al. 1991; Newlon et al. 1993). The alignments of intergenic regions also differed from alignments of ORF regions in their sensitivity to the software used to generate the alignment and the parameters used to weight matches, mismatches, and gaps. Typically, certain blocks of intergenic sequence were aligned the same way irrespective of the parameters used, while others varied greatly, especially in the location of gaps.

ARS307$^{carl}$: ARS307$^{carl}$ is 64% identical to ARS307 over a 308-bp region, allowing nine gaps (data not shown). Linker-substitution mutagenesis of ARS307 identified three functional elements, an essential domain A, which includes the ACS, and two accessory elements in domain B, B1, and B2 (Figure 1B; Rao et al. 1994; Theis and Newlon 1994). As can be seen from the alignment in Figure 3A, the essential match to the ACS is within a 41-bp region of homology that also includes most of the B1 element. The B1 element is conserved, except for 5 bp in the middle, where linker substitutions have little or no effect on ARS activity (Rao et al. 1994; Theis and Newlon 1994). To assess whether these conserved sequences are important for the function of ARS307$^{carl}$, we made the base substitution mutations indicated in Figure 3A. Base changes in the ACS and the adjacent 3 bp, which are part of domain A, abolished ARS activity. Similarly, mutation of residues in the ACS-proximal half of the B1 element also abolished ARS activity, resulting in a more severe phenotype than was seen for a compara-
Figure 3—Analysis of ARS elements. In the alignments, bars indicate positions of identity and periods mark the positions of gaps. Plasmid stabilities are expressed as the percentage of plasmid-bearing cells in a log-phase culture under selection. Ars denotes that the mutant failed to yield transformants. (A) ARS307\textsuperscript{carl}: The sequences presented are the reverse-complement of those in the genome database, to show the T-rich strand of the ACS. BASES below the alignment indicate changes made in the ARS307\textsuperscript{carl} sequence. The black boxes with white letters indicate the sequences of ARS307 domain A and the B1 elements (Rao et al. 1994; Theis and Newlon 1994). The more severe phenotype likely reflects the fact that the construct analyzed lacked an intact B2 element. Mutation of the distal part of the B1 element resulted in an unusual phenotype: transformant colonies exhibiting two different plasmid stabilities. Some colonies showed plasmid stabilities indistinguishable from wild type (25 ± 2%), while other colonies, even from the same transformant, gave a lower plasmid stability (14 ± 1%). This behavior is reminiscent of the epigenetic inheritance of the silenced state seen in sir1 mutants, which reflects a compromised ability to establish the repressive chromatin structure associated with silenced mating-type genes (Pillus and Rine 1989). Metastable silencing defects are also caused by some mutations in the cis-acting silencers (Laurenson and Rine 1992), a situation akin to the mutations in ARS307\textsuperscript{carl}. Similarly, some deletion derivatives of Schizosaccharomyces pombe centromeres are delayed in assuming a functional conformation when they are introduced via DNA transformation. Minichromosomes carrying these centromeres are initially very unstable and are stabilized by the establishment of the required higher-order chromatin structure mediated by an enhancer of centromere function (Steiner and Clarke 1994; Ngan and Clarke 1997). Perhaps this mutation in the B1 element compromises the ability of ARS307\textsuperscript{carl} to recruit a protein that participates in replication initiation or allows the formation of a chromatin structure that compromises ARS function. (B) ARS305\textsuperscript{carl}: The essential match to the ACS, Box 3, and B1 of ARS305 are shown as white letters in black boxes (Huang and Kowalski 1996). The sequences presented are the reverse-complement of those in the genome database, to show the T-rich strand of the ACS. (C) ARS309\textsuperscript{carl} sequence alignment.
mutation in ARS305 is that the construct used to analyze ARS305carl was truncated at position 245 and therefore lacked the B4 element.

**ARS309carl**: Comparison of the DNA sequences of ARS309 and ARS309carl 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 11 bp 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 ARS309carl. 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 ARS305carl 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.

**ARS306carl**: 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 ARS306carl 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, RTCRYYNNNACG (Della Set al. 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.

**ARS310carl**: 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 ARS310carl 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 contains 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 ARS304carl. 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 ARS304carl has not been tested.

ARS308 contains 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. CEN3carl is contained in a 1.1-kb HindIII fragment, and ARS308carl lies in the adjacent 0.3-kb HindIII-EcoRI fragment (Yang et al. 1999). Therefore, ARS308carl 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 ARS308carl 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 ARS308carl 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 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. ARS305carl, ARS307carl, ARS310carl, 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, ARS305carl, ARS307carl, and ARS310carl, 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-

### TABLE 1

<table>
<thead>
<tr>
<th>ARS element</th>
<th>Mutation</th>
<th>S. cerevisiae phenotype</th>
<th>S. monacensis phenotype</th>
</tr>
</thead>
<tbody>
<tr>
<td>ARS307carl</td>
<td>None</td>
<td>Ars+</td>
<td>Ars+</td>
</tr>
<tr>
<td>ARS307carl</td>
<td>ACS</td>
<td>Ars+</td>
<td>Ars+</td>
</tr>
<tr>
<td>ARS310carl</td>
<td>None</td>
<td>Ars+</td>
<td>Ars+</td>
</tr>
<tr>
<td>ARS310carl</td>
<td>ACS</td>
<td>Ars+</td>
<td>Ars+</td>
</tr>
</tbody>
</table>

*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. Ars+ indicates that transformants were obtained at high frequency, while Ars- indicates that transformants were obtained rarely or not at all.*
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 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 ARS308carl and the intergenic region of ARS305carl 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 ARS308carl, 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 ARS305carl and ARS307carl, whose S. cerevisiae homologues had previously been analyzed by linker-substitution mutagenesis (Rao et al. 1994; Theis and Newlon 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 Newlon 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 ARS307carl and ARS305carl 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 Newlon 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 Newlon 1990; Bell and Stillman 1992) and a comparison of the essential matches of known ARS elements (Theis and Newlon 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 Newlon 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 ARS307carl are expected to be neutral for ARS activity. Similarly, the conserved substitution at position 3 of the ARS305 and ARS305carl essential matches is found in ARS602 (Shirahige et al. 1993). However, the change at position 6 of the ARS305carl ACS is a change that reduced but did not abolish ARS307 activity (Van Houten and Newlon 1990). The most surprising finding was the conservation of the nonconsensus C at position 9 of the ARS309 essential match in ARS309carl. This change abolishes the activity of ARS307 (Van Houten and Newlon 1990) and is found in no other known essential match (Theis and Newlon 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 Kluveromyces 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 extrachromosomal 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, its 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 Matμ2-Mcm1 operator site within the minimal RE is required for the activity of the minimal fragment (Wu et al. 1998). A second Matμ2-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 GM53679 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


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.


Communicating editor: M. Johnston