Isolation and Molecular Analysis of the Gene Cluster for the Arginine Deiminase System from *Streptococcus gordonii* DL1

Yiqian Dong, Yi-Ywan M. Chen, Jennifer A. Snyder,† and R. A. Burne*

Department of Oral Biology, University of Florida, Gainesville, Florida 32610

Received 3 June 2002/Accepted 14 August 2002

The arginine deiminase (AD) system (ADS) is one of two major ammonia-generating pathways in the oral cavity that play important roles in oral biofilm pH homeostasis and oral biofilm ecology. To initiate a study of the *Streptococcus gordonii* ADS, the ADS gene cluster was isolated from subgenomic DNA libraries of *S. gordonii* DL1 by using an *arcB*-specific probe. Nucleotide sequence analysis revealed six open reading frames (ORFs) that were arranged contiguously; the first five ORFs were transcribed in the same direction, as an apparent operon, and the sixth was transcribed in the opposite direction. The ORFs were found to share significant homologies and to correspond closely in molecular mass to previously characterized *arc* genes; thus, they were designated *arcA* (AD), *arcB* (ornithine carbamyltransferase), *arcC* (carbamate kinase), *arcD* (arginine-ornithine antipporter), *arcT* (dipeptidase), and *arcR* (regulator). A putative σ^70^ promoter (*ParcA [TTGTGT-N\_19-T AGAAT]*) was mapped 5′ to *arcA* by primer extension, and the expression of *ParcA* was shown to be inducible by arginine and repressible by glucose, in agreement with AD-specific activities measured in the wild-type strain. To investigate the function of ArcR in the differential expression of the *arc* operon, *arcR* was insertionally inactivated by a Km resistance marker flanked by T4 transcription/translation termination signals, and the expression of *ParcA* was monitored by primer extension in the wild-type and ArcR-deficient strains. Lower levels of *arcR* expression, as well as lower levels of AD activity, were consistently observed in the ArcR-deficient strain compared to wild-type cells, regardless of the growth conditions. Thus, ArcR is a transcriptional activator that is required for induction and optimal expression of the *S. gordonii* ADS gene cluster.

The arginine deiminase (AD) system (ADS) is a three-enzyme pathway that catalyzes the conversion of arginine to ornithine, ammonia, and CO₂ with the concomitant production of ATP (16). Arginine is first hydrolyzed by AD, encoded by *arcA*, to generate citrulline and ammonia. Citrulline is then converted to ornithine and carbamylphosphate via a catabolic ornithine carbamyltransferase (cOTC) encoded by *arcB*. Finally, carbamate kinase (CK), encoded by *arcC*, transfers phosphate from carbamylphosphate to ADP to produce ATP, CO₂, and ammonia. In some cases, an arginine-ornithine antipporter (ArcD), which catalyzes the uptake of arginine and concomitant export of ornithine, and a putative transaminase or peptidase (ArcT) have been found to be part of *arc* gene clusters (37).

Arginine metabolism via the ADS is widely distributed in both bacteria and archaeabacteria, and the primary structures of the enzymes involved in the pathway have been reasonably conserved throughout evolution. In contrast, the physiologic role and genetic regulation of expression of the ADS vary among microorganisms. For instance, both *Bacillus licheniformis* (32) and *Pseudomonas aeruginosa* (26) utilize the ADS exclusively under anaerobic conditions and the presence of arginine is required for full induction of the pathway. Expression of the *P. aeruginosa* arc operon requires a transcriptional activator, ANR, that binds to a conserved ANR box (5′-TTG ACN₅ATCAG-3′) located 5′ to the transcriptional initiation site. Binding of ANR occurs only under conditions of low oxygen tension, and induction by ANR can be further enhanced by the addition of arginine in the presence of a transcriptional regulator, ArgR. A conserved ArgR box (5′-TTGTC GN₅AAN₅-3′) comprising a pair of half sites in a palindromic arrangement is required for ArgR binding. Similarly, expression of the *Rhizobium etli* arc operon is induced under anaerobic conditions (18) and it is believed that a functional ADS is essential for efficient nitrogen fixation by this organism. In some lactic acid bacteria, such as *Streptococcus sanguis* (19) and *Lactobacillus sakei* (37), the expression of the operon is under the control of catabolite repression and is inducible by arginine. It has been suggested that ATP generation via the ADS in *L. sakei* promotes the survival of the organisms in the stationary phase (11). *S. sanguis* is protected against lethal acidification by catabolism of arginine via the ADS (10, 29), presumably through the production of ammonia and subsequent increases in local pH, as well as by ATP generation, which provides energy for growth and contributes to the maintenance of ΔpH.

In the oral cavity, bacteria experience repeated cycles of acidification resulting from periodic ingestion of carbohydrates by the host. The acidification phases are followed by alkalization of the environment, with the pH returning to near neutrality during fasting periods (31). Alkali production by plaque bacteria is believed to contribute to alkalization of oral biofilms, which, during resting periods, are generally at a higher pH than saliva. Most of the alkali in the oral cavity is generated by hydrolysis of urea and arginine by bacterial ureases and the ADS, respectively. Ammonia can neutralize acids generated by bacterial glycolysis, thereby increasing plaque...
pH. The elevated pH induced by urea and arginine catabolism is thought to be important in inhibiting the development of tooth decay and in modulating the composition of plaque ecology (6, 8). A correlation between higher concentrations of arginine in saliva and a lower incidence of dental caries has been demonstrated (36), further strengthening the idea that base production is one of the key factors in oral health and oral biofilm pH homeostasis. The concomitant production of ATP via arginine metabolism may also provide a strong selective advantage for ADS-positive organisms during carbohydrate limitation. Among the comparatively few ADS-positive organisms that colonize the mouth, Streptococcus gordonii is one of the most abundant organisms in tooth biofilms. S. gordonii belongs to the viridans group of streptococci, microorganisms that are early colonizers of the teeth and make up a major portion of healthy human supragingival dental plaque. It has been suggested that establishment of healthy biofilms may inhibit subsequent colonization by the highly acidogenic, aciduric species that are responsible for dental caries (8). Thus, the ability of S. gordonii to produce base may play an important role in plaque ecology and the development of oral diseases. To begin to understand the importance of the ADS in oral health, we have initiated a molecular analysis of the S. gordonii ADS in this study.

MATERIALS AND METHODS

Bacterial strains, growth conditions, and reagents. S. gordonii DL1 was maintained in brain heart infusion (Difco Laboratories, Detroit, Mich.) broth at 37°C in 5% CO₂ and 95% air. An S. gordonii ArcR-deficient strain was selected and maintained on brain heart infusion agar supplemented with 250 µg of kanamycin (KM) ml⁻¹. Recombinant Escherichia coli strains were routinely maintained on L agar supplemented, when indicated, with 25 µg of chloramphenicol ml⁻¹ or 50 µg of KM ml⁻¹. All chemical reagents were obtained from Sigma.

DNA manipulations and construction of a subgenomic DNA library of S. gordonii DL1. Genomic DNA from S. gordonii DL1 was isolated as previously described (12). Plasmid DNA was isolated from E. coli by the method of Birnboim and Doly (5). Plasmid DNA to be used in sequencing reactions was prepared from E. coli DH10B by using the QIAprep Spin Plasmid Kit (Qiagen Inc., Valencia, Calif.). Cloning and electrophoretic analysis of DNA fragments were carried out by using established protocols (2). Southern hybridization and washes at high stringency were carried out as described previously (25). Restriction and DNA-modifying enzymes were purchased from Life Technologies Inc. (Rockville, Md.) or New England BioLabs (Beverly, Mass.).

For preparation of subgenomic DNA libraries, total chromosomal DNA of S. gordonii DL1 was digested to completion with XbaI or HincII. The digested DNA fragments were separated on agarose gels, and the XbaI fragments of 6 to 8 kbp or HincII fragments of 3 to 5 kbp were enriched by gel purification with the Elu-Quik DNA Purification Kit (Schleicher & Schuell, Keene, N.H.). The isolated DNA fragments were ligated onto XbaI- or HincII-digested, phosphate-treated pSU20 (3), respectively. The ligation mixtures were used to transform E. coli DH10B, and chloramphenicol-resistant transformants were selected.

Development of an arcB-specific probe. An internal fragment of S. gordonii DL1 arcB, encoding COTC, was generated by PCR using degenerate primers that were designed on the basis of alignments of known ornithine carbamyltransferases and COTCs. Primer 5′-arcB [5′-GGNGATCCTCGAC/AC GNAAT/C/TGATCA/TGAT-3′] contained the sequence encoding approximately amino acid positions 160 to 170 of ornithine carbamyltransferase (GDARNNM). Primer 3′-arcB [5′-TTGCATG/GCTGG/TGTATCC/GTCCG/CTGG-3′] contained the antisense sequence of that encoding approximately amino acid positions 315 to 325 (HMRINEAQ). Each reaction consisted of 5 cycles at a less stringent annealing temperature (50°C), followed by 20 cycles at a stringent annealing temperature (55°C). The product with the correct predicted size was gel purified prior to cloning into pCRII to generate pZ223. Southern blot analysis confirmed that the PCR product hybridized to the S. gordonii DL1 chromosome at high stringency (requiring greater than 90% homology). Sequencing analysis and BLAST searches were performed to confirm that the product shared high degrees of homology with other COTCs.

Construction of an ArcR-deficient strain of S. gordonii DL1. An EcoRI-XbaI fragment of 1.5 kbp containing the S. gordonii arcR gene was amplified by recombinant PCRs (22) to introduce a unique BamHI site 152 bp 3′ to the start codon of arcR. A BamHI fragment containing a KM resistance cassette flanked by transcription/translation termination signals (Ttkan) (30) was subsequently cloned into the BamHI site within the PCR product. The resulting plasmid was used to transform S. gordonii (24) to generate an ArcR-deficient mutant via double-crossover recombination. The correct configuration in the integration event in KM-resistant transformants of S. gordonii was confirmed by Southern blot analysis with flkan- and arcR-specific probes.

RNA isolation and primer extension analysis. Wild-type S. gordonii DL1 and its ArcR-deficient derivative were grown to mid-exponential phase (optical density at 600 nm, ~0.5 to 0.6) in TV medium (9) containing 10 mM glucose or galactose, with or without 1% arginine. Total RNA was isolated from each culture as previously described (13). The arcR transcription initiation site was determined by primer extension analysis (2) with primer ASarcA (5′-CAGGTC TATGTAACTACACTTTCATAC3′), which contained the antisense sequence of arcR located 42 bases 3′ to the translational start site. For each reaction, radiolabeled primers were incubated with 50 µg of total cellular RNA at 42°C for 90 min to allow complete annealing. Products were analyzed alongside a DNA sequencing reaction using the same primer.

AD assay. AD activity was measured by monitoring citrulline production from arginine as previously described (1). Briefly, S. gordonii DL1 and the ArcR-deficient strain were grown in TV medium containing 10 mM glucose or galactose, with or without 1% arginine, to an optical density at 600 nm of ~0.55. Cells were harvested by centrifugation, washed once with 10 mM hexanoic acid, and resuspended in 1/10 of the original culture volume in the same buffer. The cell suspension was disrupted in a bead beater (Biospec Products, Inc., Bartlesville, Okla.) for a total of 40 s at 4°C. The total cell lysates were centrifuged at 18,000 × g, the soluble fraction was recovered, and the concentration of the protein was determined by using a protein assay (Bio-Rad, Hercules, Calif.) based on the method of Bradford. Bovine serum albumin was used as the standard.

Nucleotide sequence accession number. The complete sequences of the arc operon and arcR reported in this article have been deposited in the GenBank database, and the accession number is AF534569.

RESULTS AND DISCUSSION

Isolation of the arc genes of S. gordonii DL1. To prepare a subgenomic DNA library, total chromosomal DNA isolated from S. gordonii DL1 was digested to completion with various restriction enzymes. The digested DNA fragments were separated on 0.8% agarose gel and screened for the presence of arcB by hybridization under stringent hybridization conditions with a 0.5-kbp PCR product internal to the S. gordonii arcB gene, which was generated as described in Materials and Methods. The results indicated that the arcB gene was contained on an XbaI fragment of approximately 7 kbp. To clone this fragment, a subgenomic DNA library of XbaI fragments was constructed in moderate-copy-number plasmid pSU20 (3). The library was screened by colony hybridization under stringent conditions with an arcB-specific probe. A positive clone, containing a 7.0-kbp DNA insert (pZ240) was identified. Southern blot analysis confirmed that the origin of this 7.0-kbp XbaI DNA fragment was S. gordonii and demonstrated that this fragment was continuous on the chromosome (data not shown).

The results of sequence analysis of the 7.0-kbp XbaI fragment (see below) indicated that this fragment contained the 3′ portion of a partial open reading frame (ORF) that shared homology with other known arcA genes, followed by five complete ORFs. To obtain genomic DNA fragments that contained the complete ORF and potentially identify other genes tightly linked to the arginine catabolism cluster, a subgenomic DNA library of HincII fragments was constructed as detailed in Materials and Methods and the library was screened with a 1-kbp DNA fragment containing the 3′ portion of arcA. A
3.2-kbp HinCII fragment was subsequently isolated, and the origin of this fragment was confirmed by Southern blot analysis of *S. gordonii* chromosomal DNA under stringent hybridization conditions.

**Nucleotide sequence analysis of the arc genes.** The complete nucleotide sequences of both strands of the 7.0-kbp *Xba*I fragment and the 3.2-kbp HinCII fragment were determined from a series of nested deletions generated by exonuclease III (21). Translation of the nucleotide sequences revealed six ORFs that were arranged contiguously on the chromosome. The first five ORFs were transcribed in the same direction as an operon, and the sixth was transcribed in the opposite direction. All six ORFs were preceded by a putative Shine-Dalgarno sequence and began with an ATG codon. A putative rho-independent terminator was located between ORFs 5 and 6 (5'-AAAAGG AGTGGAAAATCTCAACTCCTTTT-3' [the bases forming the stem are underlined]; \(\Delta G = -12.2\) kcal). The ORFs had significant homologies and corresponded closely in molecular mass to previously characterized arc genes; thus, they were designated *arcA, -B, -C, -D, -T, and -R* (Fig. 1). In addition, the ORFs immediately flanking the *arc* gene cluster did not appear to be involved in arginine catabolism (data not shown). We also compared the sequence of the *arc* gene cluster from this study with that of the incomplete *S. gordonii* genome from The Institute for Genomic Research Microbial Genome Database. However, only one contig in the database could be identified that contains the 3' portion of the *arc* gene cluster. The existing partial sequence of the *arc* gene cluster in The Institute for Genomic Research database (*arcT* and *arcR*) is identical to the sequence completed in this study.

The similarities between the deduced amino acid sequences of *arcABCDCTR* of *S. gordonii* and *arc* genes of other bacteria were evaluated. Briefly, *S. gordonii* AD, encoded by *arcA*, was more than 92 and 84% identical to *Streptococcus pneumoniae* and *Streptococcus pyogenes* AD, respectively. Lesser but still significant similarities were also observed with the AD enzymes of *Enterococcus faecalis* (65% identity), *L. sakei* (51% identity), and *B. licheniformis* (52% identity). *S. gordonii* arcB, encoding cOTC enzymes, was more than 94 and 89% identical to the *S. pneumoniae* and *S. pyogenes* cOTCs, respectively. Lower but still substantial levels of homology were observed with cOTCs from gram-negative bacteria. The motifs for carbamylphosphate binding and catalysis (4, 23), which are highly conserved in both anabolic ornithine carbamyltransferase and cOTC, STRTR and HPTQ, were identified at protein residues 58 to 62 and 136 to 139, respectively. Other conserved amino acid residues between these two regions that are essential for substrate binding and catalysis were also present in the *S. gordonii* cOTC (K88, K91, R101, and G110). The ornithine binding motif LHCLP was located at amino acid residues 271 to 275. *S. gordonii* CK was 89 and 65% identical to the *S. pneumoniae* and *S. pyogenes* CKs, respectively. Similar to what was noted for the *E. faecalis* CK (28), the conserved motif for carbamylphosphate binding, STRTR, could not be identified in the *S. gordonii* CK nor could the P loop that is known to be involved in mononucleotide binding. *S. gordonii* ArcD, the arginine-ornithine antiporter, was 74 and 64% identical to ArcD of *S. pneumoniae* and *S. pyogenes*, respectively. Eleven transmembrane helices were predicted on the basis of the deduced amino acid sequence of *S. gordonii* arcD by using the dense alignment surface method accessed from the “DAS”-Transmembrane Prediction server at http://www.sbc.usc.edu/~miklos/DAS (15). Finally, ArcT was found to have 70% homology with an *S. pyogenes* Xaa-His dipeptidase.

Significant levels of homology were detected between the ArcR sequences of *S. gordonii* and *S. pyogenes* (65% identity). A substantial degree of similarity was also observed when *S. gordonii* ArcR was compared to other known arginine metabolism regulatory proteins, such as *Bacillus subtilis* AhrC (64%) and *B. licheniformis* ArgR (64%). Similar to *B. subtilis* AhrC and *E. coli* ArgR (20), the pI of the first 100 amino acids of *S. gordonii* ArcR was basic (9.71) while the pI of the last 56 residues was acidic (3.75). Furthermore, the serine-arginine motif (34) that has been shown to be essential for DNA binding was found at the N terminus of ArcR (amino acids 43 and 44). Conserved amino acid residues for arginine binding that were defined by studies of the *E. coli* arginine repressor (7, 34, 35) were found in *S. gordonii* ArcR at positions 103 (alanine), 125 (aspartic acid), and 126 (aspartic acid), and the conserved glycine residue for oligomerization (35) was also present at position 124.

**Expression of AD in S. gordonii.** The ADS in lactic acid bacteria is highly regulated by growth conditions (17, 37). Specifically, in *S. sanguis* NCTC 10904, the system requires induction by arginine and expression is subject to carbohydrate catabolite repression (CCR), with higher levels of activity in cells grown on nonrepressing sugars, such as galactose (17). To investigate whether the expression of the *S. gordonii* DL1 *arc* operon is regulated in the same manner, AD specific activities were measured in cells of the wild-type strain grown on various energy sources (Fig. 2). When cells were grown in galactose, higher levels of AD activity were observed than in cells grown in glucose. Addition of 1% arginine to the growth medium resulted in a further increase in AD activity, although the presence of glucose had a repressive effect on AD expression. Thus, *S. gordonii* arcA was inducible by arginine and sensitive to CCR.

**Localization of ParcA.** To investigate the control of *arc* operon expression, we first localized the transcriptional initiation site of the most 5' gene in the *arc* cluster, *arcA*, by primer extension with total cellular RNA isolated from cells grown under inducing and derepressing conditions (1% arginine and 10 mM galactose). A single product, 50 bases 5’ to the ATG start codon at a G residue, was observed, and a putative σ^70^ promoter (ParcA [TTGTTG-N_{15}-TGAAT]) was identified 8
bases 5′ to the signal (Fig. 3). To identify the possible cis elements that are involved in the regulation of arc operon expression, we first searched for the presence of a highly conserved catabolite-responsive element (cre) near ParcA. CCR in low-GC gram-positive bacteria has been studied extensively (33). The two major regulatory elements involved are the trans-acting factor catabolite control protein A (CcpA) and a cis-acting palindromic sequence designated cre. In the presence of glucose, CcpA binds to cre and represses the expression of catabolic genes and operons. Two cre sites were identified 5′ to the transcription initiation site. One, AGAAAACGCTTCAA, spanning positions −107 to −94, differed by three bases from the consensus sequence, and the other, TGTAAGTGTTTTCA, spanning positions −35 to −22, differed by only one base from the consensus sequence, suggesting that catabolite repression of ParcA may be controlled by CcpA. On the other hand, an Arg box, the cis element essential for the activation of arginine-dependent, anaerobic growth of B. licheniformis and P. aeruginosa (26, 27), was not found 5′ to the arcA translation start site.

Functional analysis of ArcR. To determine whether ArcR is involved in the regulation of arc operon expression, an ArcR-deficient derivative was constructed by allelic exchange with the insertion of a KM resistance marker. The levels of arcA message in both the wild-type and ArcR-deficient strains were evaluated by primer extension analysis under different growth conditions (Fig. 4). In agreement with the AD activities observed in wild-type cells under different growth conditions, the primer extension signal was lower in cells grown with glucose and higher in cells grown with arginine. In the absence of ArcR, lower levels of expression were consistently observed than in wild-type cells growing on the same substrates, indicating that ArcR is required for optimal arc expression. Similar to wild-type cells, expression from ParcA in the ArcR-deficient strain was still sensitive to CCR, with consistently lower levels of expression being observed when glucose was the sole carbohydrate source. The aberrant expression of arc genes was further examined at the level of AD activity in the ArcR-deficient strain (Fig. 2), in which lower levels of activity were consistently observed than in the wild-type strain, regardless of the growth conditions. Thus, it is clear that ArcR is essential for optimal expression of the arc operon.

In conclusion, the genes encoding enzymes for arginine catabolism in S. gordonii DL1 were identified in this study. The expression of the arc operon of S. gordonii was sensitive to CCR and inducible by arginine. ArcR was shown to be a positive regulator of arc gene expression. The use of genetically engineered, ammonia-producing oral streptococci as potential agents for the control of dental caries has been demonstrated by expressing the urease genes of Streptococcus salivarius in Streptococcus mutans (14). The information obtained in this study provides the foundation for testing of the potential of
exploiting the use of arginine catabolism to moderate plaque acidification and to control the emergence of a cariogenic flora.

ACKNOWLEDGMENT

This work was supported by Public Health Service grant DE10362 from the National Institute of Dental and Craniofacial Research.

REFERENCES