JOURNAL OF BACTERIOLOGY, Aug. 1996, p. 4942–4947 Vol. 178, No. 8
0021-9193/96/$04.00+0

Fermentative Arginine Degradation in *Halobacterium salinarium* (Formerly *Halobacterium halobium*): Genes, Gene Products, and Transcripts of the *arcRACB* Gene Cluster

ANDREAS RUEPP AND JÖRG SOPPA*

Max-Planck-Institut für Biochemie, 82 152 Martinsried, Germany

Received 27 November 1995/Accepted 17 June 1996

Fermentative growth via the arginine deiminase pathway is mediated by the enzymes arginine deiminase, carbamate kinase, and catabolic ornithine transcarbamylase and by a membrane-bound arginine-ornithine antiporter. Recently we reported the characterization of catabolic ornithine transcarbamylase and the corresponding gene, *arcB*, from *Halobacterium salinarium* (formerly *Halobacterium halobium*). Upstream of the *arcB* gene, three additional open reading frames with halobacterial codon usage were found. They were identified as the *arcC* gene coding for carbamate kinase, the *arcA* gene coding for arginine deiminase, and a gene, tentatively termed *arcR*, coding for a putative regulatory protein. The identification of the *arcC* and *arcA* genes was verified, respectively, by heterologous expression of the enzyme in *Haloferax volcanii* and by protein isolation and N-terminal sequence determination of three peptides. The gene order *arcRACB* differs from the gene order *arcDABC* in *Pseudomonas aeruginosa*, the only other organism for which sequence information is available. Transcripts from *H. salinarium* cultures grown fermentatively or aerobically were characterized by Northern (RNA) blot and primer extension analyses. It was determined (i) that monocistronic transcripts corresponding to the four open reading frames exist and that there are three polycistronic transcripts, (ii) that the level of induction during fermentative growth differs for the various transcripts, and (iii) that upstream of the putative transcriptional start sites for the three structural genes there are sequences with similarities to the halobacterial consensus promoter. The data indicate that expression of the *arc* gene cluster and its regulation differ in *H. salinarium* and *P. aeruginosa*.

The arginine deiminase pathway of fermentative arginine utilization consists of the three enzymes arginine deiminase (ADI, EC 3.5.3.6), catabolic ornithine transcarbamylase (cOTCase, EC 2.1.3.3), and carbamate kinase (CK, EC 2.7.2.2) and a membrane-bound arginine-ornithine antiporter. The degradation of arginine is coupled to the equimolar generation of ATP by substrate level phosphorylation. This pathway is found in a variety of phylogenetic groups within the domain *Bacteria*, e.g., *Pseudomonas* spp., *Mycoplasma* spp., *Bacillus* spp., and lactic bacteria (for a review, see reference 8). The pathway has been most extensively studied in *Pseudomonas aeruginosa* (21), the only species for which the sequences of the four genes have been determined (1, 2, 29). The genes are organized in an operon, and a polycistronic mRNA is transcribed from a single promoter and subsequently processed into a variety of smaller transcripts (17) in an RNase E-dependent reaction (18). The transcription of the *arcDABC* operon is under the positive control of the ANR protein (15, 22), a member of the FNR protein family.

Within the domain *Archaea*, arginine-dependent anaerobic growth has only been reported for some halobacterial species (25, 37). The consumption of arginine was coupled to the equimolar production of ornithine (25), and therefore it was concluded that the ADI pathway was used. One of the enzymes, cOTCase, has been isolated from *Halobacterium salinarium* (formerly *Halobacterium halobium*). The protein was characterized and the corresponding gene, *arcB*, has been cloned and sequenced (37). Upstream of the *arcB* gene three further open reading frames were found. In the present study the sequences of these genes, coding for CK, ADI, and a putative regulatory protein, are reported, and the gene products of the two structural genes are verified. Total RNA was isolated from fermentatively and aerobically grown cultures, and the sizes of the *arc* transcripts, their level of induction, and their 5′ ends have been determined.

* Corresponding author. Present address: Institut für Mikrobiologie der J. W. Goethe-Universität, Biozentrum Niederursel, Marie-Curie-Str. 9, D-60439 Frankfurt, Germany. Phone and Fax: 49-69-798 29 564. Electronic mail address: Soppa@em.uni-frankfurt.de.

MATERIALS AND METHODS

Materials. Restriction endonucleases, T4 DNA ligase, DIG-11-dUTP, and the digoxigenin luminiscence detection kit were obtained from Boehringer GmbH (Mannheim, Germany). The Moloney murine leukemia virus reverse transcriptase kit was from GIBCO BRL (Eggenheim, Germany). The equipment for Northern (RNA) transfer was purchased from Schleicher & Schuell (Dassel, Germany), and nylon membranes were from Qiagen (Hilden, Germany). Bluescript vectors were obtained from Stratagene (Heidelberg, Germany), and the Prep-A-Gene kit was from Bio-Rad (Munich, Germany).

Microorganisms. *H. salinarium* L33 (43) was obtained from Dieter Oesterhelt (Max-Planck-Institut für Biochemie, Martinsried, Germany), the type strain *H. salinarium* was from the German culture collection (Deutsche Stammsammlung, Göttingen, Germany, strain no. DSM 670), *Haloferax volcanii* WR340 was from Moshe Mevarech (Tel Aviv University, Tel Aviv, Israel) and *Escherichia coli* XL1 Blue MRF′ was from Stratagene.

Cloning and sequence determination. General molecular biological techniques were performed according to Sambrook et al. (38). The methods used for cloning and sequencing of a 3.3-kbp fragment of genomic DNA encoding the *arcB* gene, an additional open reading frame (ORF) (*arcC*), and a part of a third ORF (*arcA*) were described earlier (37). An overlapping fragment of genomic DNA containing the remaining part of the *arcA* gene and the *arcR* gene was cloned and sequenced by using essentially the same techniques. For probe construction, the primers 3I-start, 5′-GAC ATG AGC CAG GTT CGG G-3′, and Up6, 5′-GTC CAT CTC GAC TTC CCC-3′, were used.

DNA isolation and transformation. Small-scale plasmid isolation was done by the LiCl boiling method (44), and large-scale isolations were done with columns from Qiagen. Halobacterial genomic DNA was isolated as described by Rosenshine et al. (36). *E. coli* was transformed by electroporation (10), and *Haloferax volcanii* was described by Cline et al. (7).

Heterologous expression of *arcC* in *Haloferax volcanii*. To prove that the *arcC* gene encodes a CK, it was cloned into pSD1-R1/6. pSD1 is a shuttle vector with

origins of replication and marker genes for *E. coli* and for *Haloferax volcanii*. The selection of plasmids with strong promoters from a plasmid library containing a randomized promoter region upstream of a marker gene led to the identification of the plasmid pSD1-R1/6 (9). The promoter of pSD1-R1/6 has a higher activity than natural halobacterial promoters, and thus by replacing the marker gene with the gene of interest this plasmid can be used as an expression vector. The *arcC* gene was amplified by PCR with the primers CK-NcoI, 5′-GGC ACC CAT GGC TTA CAC CGT TGT AGC TCT G-3′, and CK-Asp718, 5′-CCA ACT GGT ACC TCA CTC GTC GGC TGG CAC GAC-3′, providing a *Nco*I restriction site at the N terminus of the gene and an *Asp*718 restriction site at the C terminus, respectively. The PCR fragment was ligated into pSD1-R1/6, which had been cut with the same enzymes. The resulting plasmid, pR1/6-CK, was transformed into *Haloferax volcanii* WR340. CK activity was measured by an indirect assay. In a first step, crude extracts of a transformant and an untransformed control culture were incubated for different times at 37°C in the presence of 40 mM ornithine, 5 mM carbamyl phosphate, 5 mM ADP, 60 mM Tris-HCl (pH 7.2), and 3 M KCl. The carbamyl phosphate, which was not degraded by CK, was subsequently converted into citrulline by the addition of ornithine and purified cOTCase of *H. salinarium* (37). Citrulline was quantitated with a colorimetric assay essentially as described by Oginsky (31). The amount of citrulline, which corresponds to the amount of unreacted carbamyl phosphate, was plotted versus the time of incubation.

Northern blot analysis. *H. salinarium* was grown aerobically in the absence of arginine as well as fermentatively in the presence of 1% arginine to the late exponential growth phase in complex medium as described (37). RNA was isolated as described by Chomczynski and Sacchi (6). Multiple pairs of aliquots were separated in formaldehyde gels (38) together with a molecular weight marker (USB, Bad Homburg, Germany). The RNA was transferred to nylon membranes (Qiagen) by downward capillary Northern blots (5) overnight by using 10× SSC (1× SSC is 0.15 M NaCl and 0.015 M sodium citrate) as transfer buffer. The RNA was stained with methylene blue and the membranes were cut into strips, which were hybridized with probes specific for the individual genes. Single-stranded DNA probes labelled with DIG-11-dUTP (Boehringer GmbH) were obtained using a PCR approach as described earlier (37). The following oligonucleotides were used for probe construction: Ups, 5′-CCG CGA GTT GGG AAC AGT C-3′, and REV-2, 5′-AGG CGG TCT TTC TCC TGC-3′, for the *argB*-specific probe; Sk3, 5′-GGT GGG ACG TTC ACT GTG-3′, and Up4, 5′-TTC GCG GAG CTG CTG CTG-3′, for the *arcC*-specific probe; GO1, 5′-GAT CTC AGT GAA CTC GGA G-3′, and Up7, 5′-GAA CCT GGC TCA TGT CGG CG-3′, for the *argA*-specific probe; and GO5, 5′-TCG GAT TGG TCT TGG GCG-3′, and Up10, 5′-GGT AAC GCT CAC TGC GCC C-3′ for the *argR*-specific probe. Hybridization and washing were performed in the presence of 50% formamide at 50°C. Digoxigenin detection was performed according to the instructions of the manufacturer (Boehringer GmbH).

Primer extension experiments. The primer extension experiments were performed with the reverse transcriptase kit from GIBCO BRL and modified according to Reiter et al. (34). The primers were 5′-labelled with [γ-^{32}P]ATP. For each reaction, 20 μg of RNA was used. The primers were OTCext, 5′-ACT TGT TGT CCA TCA CGC CG-3′, for *arcB*; Up5a, 5′-TTC GCC ACC TCT GAG TAG-3′, for *arcC*; Primerex1, 5′-AGA TGC TCC TGT TGG GCG TCG-3′, for *arcA*; and Up7, 5′-GAA CCT GGC TCA TGT CGG CG-3′, for the 3.1-kb RNA. As a sizing standard, single-stranded DNA of M13 (Sequenase kit; USB) was used in a sequencing reaction with the −40 primer.

Computer programs. The software packages of the Genetics Computer Group (20) and the Protein Identification Resource (33) were used on a microvax (Digital Equipment). The multiple sequence alignments were done with the program Pile up (20) and edited manually with the program Line up (20).

Nucleotide sequence accession number. The sequence of the *arc* gene cluster has been deposited in the EMBL nucleotide sequence database and has the accession number X80931.

RESULTS AND DISCUSSION

The *arc* gene cluster. Characterization of the cOTCase of *H. salinarium* and the corresponding gene, *arcB*, has been described previously (37). By using protein sequence information, a 3.3-kbp *Eco*RI fragment of genomic DNA had been cloned, and the *arcB* gene was sequenced by primer walking. Sequence determination of the remaining part of the cloned fragment led to the detection of a second ORF 56 bp upstream of the *arcB* gene, and 241 bp upstream of this the 3′-part of a third ORF was found. The deduced protein sequences were similar to the CK and ADI sequences, respectively, showing that the genes for the enzymes involved in fermentative arginine degradation are clustered in the genome of *H. salinarium*. To get access to the remaining part of the gene cluster, a 6.2-kbp *Bam*HI fragment of genomic DNA was cloned, which overlaps the *Eco*RI fragment by 279 bp. The missing part of the *arcA* gene was sequenced, and 81 bp upstream of it a fourth ORF was found. All four ORFs exhibit the halobacterial codon usage (40), and thus presumably are expressed. The four genes have high G+C contents of 62 and 63%, respectively, which is typical for halobacteria with a genomic G+C content of about 68% (30). Downstream of the *arcB* gene a stretch of T residues is located which is thought to be involved in archaeal transcriptional termination. In the regions 420 bp downstream and 1.3 kbp upstream of the four genes no further ORFs with the halobacterial codon usage could be found. Therefore, it can be concluded that no further genes belong to this gene cluster. Figure 1 shows an overview of the genomic organization of the *arc* gene clusters from *H. salinarium* and *P. aeruginosa*. The individual genes and their products are discussed below.

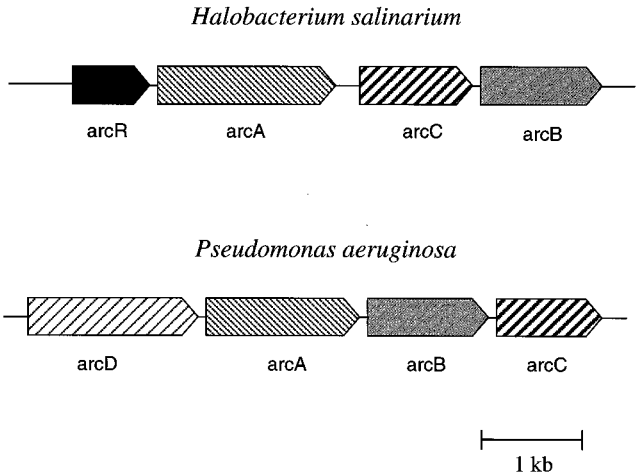

FIG. 1. Genomic organization of the *arc* genes in *H. salinarium* and *P. aeruginosa*. The four genes of the *arcRACB* gene cluster of *H. salinarium* and the *arcDABC* operon of *P. aeruginosa* are shown schematically. The same graphical pattern was used to indicate corresponding genes. A scale bar is included.

The *arcB* gene and cOTCase. Isolation of cOTCase and characterization of the *arcB* gene have been described in a previous paper (37). It should be noted that 30% of the deduced protein sequence was verified by determining the N-terminal sequences of four peptides and that the halobacterial enzyme has an amino acid identity of between 30 and 45% with cOTCase of *P. aeruginosa* and various anabolic OTCases.

The *arcC* gene and CK. The ORF which was identified as the *arcC* gene has a length of 921 bp and thus codes for a polypeptide of 307 amino acids. The calculated molecular mass of the deduced protein is 32.5 kDa. The excess of acidic amino acids (difference between the fractions of acidic and basic residues) is 14.7%, which is much higher than the value of 1.9% for proteins of enteric bacteria. This is a common feature of soluble halobacterial proteins (28, 41) and is thought to be important for solubility at the high ionic strength found in the cytoplasm (3, 12, 13). A database search showed that the CK of *P. aeruginosa* is the only protein with significant similarity to the halobacterial enzyme. The percent identity between the two proteins is 42%, a strong indication that they are related in evolution. To give further evidence that the halobacterial gene codes for a CK, the ORF was cloned into the halobacterial expression vector pSD1-R1/6 (9) downstream of a strong constitutive promoter. The resulting plasmid, pR1/6-CK, was transformed into *Haloferax volcanii* WR340, which does not possess the ADI pathway. In crude extracts of transformants, CK activity, which was absent in untransformed control cultures, could be measured (Fig. 2). Thus, we could prove ex-

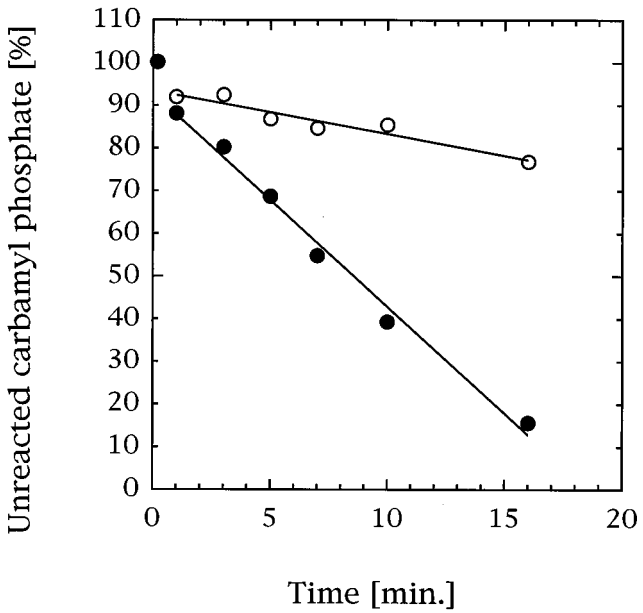

FIG. 2. Heterologous expression of CK in *Haloferax volcanii*. A two-step procedure was used to measure CK activities with lysates of either *Haloferax volcanii* cultures containing the expression vector pR1/6-CK (●) or untransformed control cultures (○). First, the lysates were incubated for the indicated times in the presence of the substrate carbamyl phosphate. In a second step, the unreacted carbamyl phosphate was converted into citrulline by the addition of ornithine and cOTCase, and the citrulline was detected by a colorimetric assay. The mean values of duplicate measurements as well as regression curves calculated with the omission of the first datum point are shown.

perimentally that the ORF upstream of *arcB* is the *arcC* gene encoding a halobacterial CK. This is the second sequence of a CK known until now.

The *arcA* gene and ADI. The ORF upstream of *arcC* consists of 1,448 bp corresponding to a polypeptide of 486 amino acids. The calculated molecular mass of the deduced protein is 54.7 kDa, and the excess of acidic residues is 12.8%. In the protein sequence databases, three ADIs were found. The similarities of the ADIs of *P. aeruginosa* (1), *Mycobacterium arginini* (27, 32), and *Mycoplasma hominis* (24) to the deduced halobacterial protein are 14.1, 14.0, and 13.4%, respectively. The sequence identities of the putative halobacterial ADI and the bacterial enzymes are much lower than those that have been found for cOTCase and CK (see above). The values of about 15% do not support an evolutionary relationship between the halobacterial enzyme and these bacterial enzymes; however, the bacterial enzymes themselves are not highly conserved, e.g., the ADIs of *P. aeruginosa* and *M. arginini* exhibit only 25.7% sequence identity.

To prove that the ORF codes for a halobacterial ADI, the ADI was isolated from a fermentatively grown culture of *H. salinarium*. In denaturing sodium dodecyl sulfate-polyacrylamide gel electrophoresis, the protein appeared as a single band with an apparent molecular weight of about 70,000. The high abundance of acidic residues in halobacterial proteins leads to a systematic overestimation of the apparent molecular weight by about 25%; therefore, this value agrees well with the calculated molecular mass of 54.7 kDa. The protein was cleaved with Lys-C. Three peptides were isolated and their N-terminal sequences were determined to be TLERPDVNFG, FEIVDV YDYVDP, and QFIDGIQEDYDVTVIPDGE (28a). The sequences are identical to the positions 376 to 385, 393 to 404, and 442 to 460, respectively, of the polypeptide sequence deduced from the *arcA* gene, confirming that *arcA* codes for an arginine deiminase.

In conclusion, for all three structural genes of the halobacterial *arc* gene cluster, *arcA*, *arcC*, and *arcB*, experimental evidence was given that they code for the three enzymes involved in fermentative arginine degradation.

The putative regulatory protein ArcR. At a location 81 bp upstream of *arcA*, an ORF with the same orientation as the three structural genes was found. It consists of 582 bp coding for a polypeptide of 194 amino acids. In a databank search five proteins possessing significant similarities to the deduced protein were found. The identities of ArcR with the five proteins are 28% for KdgR (from *Erwinia chrysanthemi*), 27% for GylR from *Streptomyces griseus*, 25% for GylR from *Streptomyces coelicolor*, 21% for IclR from *Salmonella typhimurium*, and 20% for IclR from *E. coli*. All five proteins are bacterial proteins which function as negative regulators of transcription of genes of catabolic pathways. They are involved in the catabolism of pectin (KdgR [35]), glycerol (GylR [4, 26, 39]), and acetate (IclR [42, 16]).

A multiple sequence alignment of the six proteins is shown in Fig. 3. At 24 positions (12%), the residues of at least four of the five regulators are identical to the halobacterial sequence. The similarity of ArcR to these regulators, the position of *arcR* immediately upstream of *arcACB*, and its mode of transcription (see below) make it likely that ArcR is involved in regulation of *arcACB* transcription. However, compared with the bacterial regulators, ArcR is devoid of more than 60 N-terminal amino acids. The N-terminal parts of IclR, GylR, and KdgR contain a helix-turn-helix motif proposed to be involved in DNA binding. Therefore, either ArcR has a DNA-binding motif in another part of the sequence or it performs its function via interaction with a DNA-binding protein still to be discovered. This could be regarded as a eucaryotic-like feature of ArcR, as many eucaryotic transcription factors perform their function via protein-protein interaction and not by direct DNA binding. On the other hand, there are also a few bacterial regulators which act via protein-protein interaction, e.g. the prrA gene product of *Rhodobacter sphaeroides* (14) and nifL of *Klebsiella pneumoniae* (11). Clearly, further experiments are needed to prove that the putative ArcR is involved in regulation of *arc* gene transcription and to elucidate its mechanism.

Transcript analysis of the *arc* gene cluster. Total RNA was isolated from cultures grown either fermentatively or aerobically to the late exponential growth phase. Northern blot analyses were performed by using specific probes for *arcR*, *arcA*, *arcC*, and *arcB*. The results are summarized in Fig. 4. In fermentatively grown cells, monocistronic transcripts were found for each of the four genes. The sizes of the mRNAs of the three structural genes correlate with the sizes of the ORFs, while the *arcR* message is about 260 nucleotides (nt) longer than the ORF. Three additional transcripts were observed, i.e., a transcript of 1,930 nt which included *arcC* and *arcB* and two transcripts (2,600 and 3,350 nt) which included *arcB*, *arcC*, and part of *arcA*. The bicistronic *arcCB* transcript is about twice as abundant as the monocistronic *arcC* transcript but less abundant than the monocistronic *arcB* transcript (60%). The other two transcripts are too short to include a full-length transcription of *arcA* in addition to *arcCB*, which would require a transcript size of at least 3,564 nt. Furthermore, they cannot be detected with the *arcA*-specific probe derived from the 5′-part of the gene. Both transcripts are of low abundance and together they constitute less than 10% of the RNAs detected with the *arcB*- and *arcC*-specific probes, respectively.

Comparison of the Northern blot analyses performed with total RNAs isolated from cultures grown fermentatively or

```
           1                                                                                   80
S. typh.   .......... .......... .........M VAPVPAKRGR KPAATTAPVT GQVQSLTRGL KLLEWIAESN GT..VALTEL
E. coli    .......... .......... .........M VAPIPAKRGR KPAVATAPAT GQVQSLTRGL KLLEWIAESN GS..VALTEL
S. gris.   .......... .......... .......... .......... ........MA RNIQSLERAA AMLRLLA..G GERRLGLSDI
S. coel.   .......... .......... .......... .......... ........MA RNIQSLERAA AMLRLLA..G GERRLGLSDI
E. chrys.  MDDINRSVTY SNARLPYSKR SLYTKTRVLF FLKQKILSRV TTKMAIADLD KQPDSVSSVL KVFGILQALG EEREIGITEL
H. salin.  .......... .......... .......... .......... .......... .......... .......... ..........

           81                                                                                 160
S. typh.   AQQAGLPNSS THRLLTTMQQ QGFVRQVGEL GHWAVGPHAF IVGSSFLQSR NLLAIVHPIL RKLMEDSGET VNLAVLDQSD
E. coli    AQQAGLPNST THRLLTTMQQ QGFVRQVGEL GHWAIGAHAF MVGSSFLQSR NLLAIVHPIL RNLMEESGET VNMAVLDQSD
S. gris.   ASSLGLAKGT AHGILRTLQQ EGFVEQDDAS GRYQLGAELL RLGTTYLDVH ELRARALVWT DDLARSSGES VHLGVLHQ..
S. coel.   ASTLGLAKGT AHGILRSLQA EGFVEQEPAS GRYQLGAELL ALGNSYLDVH ELRARALVWT DDLARSSGEA AYLGVLHQ..
E. chrys.  SQRVMMSKST VYRFLQTMKS LGYVAQEGES EKYSLTLKLF ELGAKALQNV DLIRSADIQM RELSALTRET IHLGALDE..
H. salin.  .......... .......... .......... ..MQLGDWFV PVSNYARNST DLYRLGREEI DKLATQTRHT AHLVTESDGR

           161                                                                                240
S. typh.   HQAII....I DQVQCTQLMR MSAPIGASCP MPASGAGKAF LSQLSEEQVT SLLHRKGLHA YTHATLVSPL HLKDDLAQTA
E. coli    HEAII....I DQVQCTHLMR MSAPIGGKLP MHASGAGKAF LAQLSEEQVT KLLHRKGLHA YTHATLVSPV HLKEDLAQTR
S. gris.   QGVLI....V HHVFRPDDSR QVLEIGAMQP LHSTALGKVL SAY.DPVAHS EAL.EADRKA FTDRTVCEPD SFEHVLDITR
S. coel.   QGVLI....V HHVFRPDDSR QVLEVGAMHP LHSTAHGKVI SAF.DPVAHS EVL.EGDRAT LTGRTVTEAA AFEEVLDLTR
E. chrys.  DSIVY....I HKIDSMYNLR MYSRIGRRNP LHSTAIGKVL LAWRDREEVK EILSQVEFKR TTVHTIGSTE ELLPQLDLVR
H. salin.  QIVLYESMGE ESVTTEYHLR MR...ETPRK LHTSAAGKSI LAFLPEGRRE TLISEIEFGT ESSTPIGSPD ALREQLATVR

           241                                                                      314
S. typh.   KRGYSFDDEE HALGLRCVAS CIYDEHREPF AALSISGPRS RITDD..RVT ELGAMVIKAA KEVTLAYGGT R...
E. coli    KRGYSFDDEE HALGLRCLAA CIFDEHREPF AAISISGPIS RITDD..RVT EFGAMVIKAA KEVTLAYGGM R...
S. gris.   ARGYAADVEE TWEGIASIAA PIHDRRRMPV GAVGITGAVE RLCREGELRP ELVAAVRDCA RAVSRDLGAG RF..
S. coel.   ARGWALDLEE TWEGVASLAA PVHDRRRMAV GAVGVTGPVE RLCPDGAPAT ELVTAVRDCA AAVSRDLGAG RF..
E. chrys.  QQGYGEDNEE QEEGLRCIAV PVFDRFGVVI AGLSISFPTI RFSEDN..KH EYVAMLHTAA RNISDQMGYH DYPF
H. salin.  DQGYAINDEE EIHGIRSIGA PIRGRTEDVA GAVSVTAPKT RL.QNAEFAG EVPALVMEAA NIIEVRLETA ....
```

FIG. 3. Alignment of the deduced ArcR sequence and bacterial regulatory proteins. The alignment includes the following sequences: IclR from *Salmonella typhimurium* (S. typh.) and *E. coli* (E. coli), GylR from *S. griseus* (S. gris.) and *S. coelicolor* (S. coel.), KdgR from *Erwinia chrysanthemi* (E. chrys.), and ArcR from *H. salinarium*. The alignment was constructed using the program Pileup (20). Amino acids which are identical in the sequence of *H. salinarium* and one or more of the other sequences are shaded.

aerobically gave a first insight into the regulation of *arc* gene expression (Fig. 5, compare lanes a and f). The following points were revealed by using the four gene-specific probes.

(i) *arcB*-containing transcripts. All four *arcB*-containing messages were found in aerobically grown cultures at late exponential growth phase, but the levels were highly elevated in fermentatively grown cultures. The levels of induction were different, being about 5-fold for the monocistronic *arcB* transcript and about 20-fold for the bicistronic *arcCB* transcript. The levels of induction of the two minor transcripts were about three- to eightfold.

(ii) *arcC*-containing transcripts. Both of the major *arcC*-containing messages are highly regulated. The 20-fold induction of the *arcCB* transcript was verified with the *arcC*-specific probe, while the induction of the monocistronic transcript could not be quantitated because of its low abundance at aerobic conditions.

(iii) *arcA*-containing transcripts. Surprisingly, the same amount of the *arcA* transcript was found in fermentatively grown cells as was found in aerobically grown cells. Clearly, the *arcA* transcript level is regulated differently from those of *arcC*, *arcB*, and *arcCB*. To elucidate whether *arcA* transcription is regulated at all, a culture was grown aerobically to the early exponential growth phase and shifted to fermentative growth conditions. Northern blot analysis using aliquots taken before and after the shift revealed that the level of the *arcA* message is also regulated (data not shown).

(iv) *arcR*-containing transcripts. Using an *arcR*-specific probe, only very faint signals could be detected. The same amount of the monocistronic *arcR* message is present in aerobically and in fermentatively grown cells, and induction experiments confirmed that this message is not regulated. Both the low abundance of this message and the lack of regulation agree well with the proposed role of *arcR* as a regulator.

To obtain indications as to whether several promoters exist

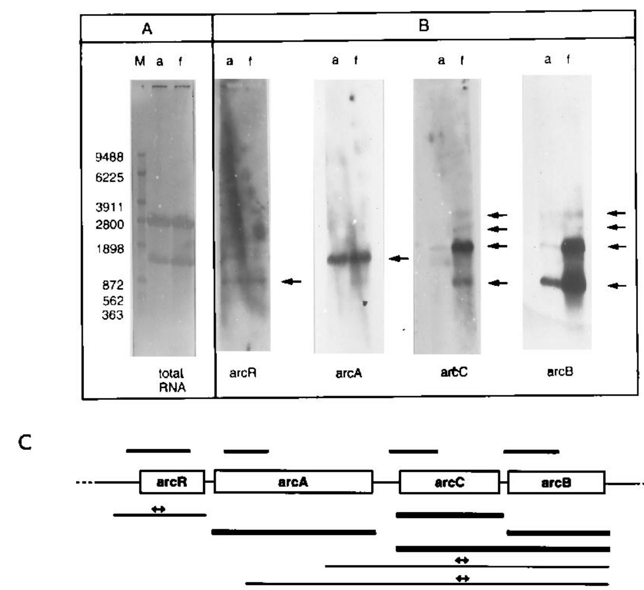

FIG. 4. Transcription analysis of the *arc* gene cluster. (A and B) Total RNA was isolated from *H. salinarium* cultures grown aerobically (a) or fermentatively (f) to the late exponential growth phase. Five replicates of the two RNA samples were separated electrophoretically and transferred by Northern blotting onto a nylon membrane, which was then cut into strips. Panel A shows one of the strips stained with methylene blue to visualize the rRNA and an RNA molecular weight marker (M). Panel B shows the results after hybridization with probes specific for the four genes, *arcR*, *arcA*, *arcC*, and *arcB*. The major and minor transcripts which were detected are indicated by arrows. (C) The genomic organization of the *arc* gene cluster is shown schematically. The probes used for Northern blot analyses are indicated above the genes. Below the genes, the major transcripts are presented as thick lines and the minor transcripts are presented as thin lines. The double arrows above the thin lines indicate that for these transcripts neither 5′-ends nor 3′-ends of the transcripts have been determined, and thus the exact location of the transcripts is uncertain to some extent.

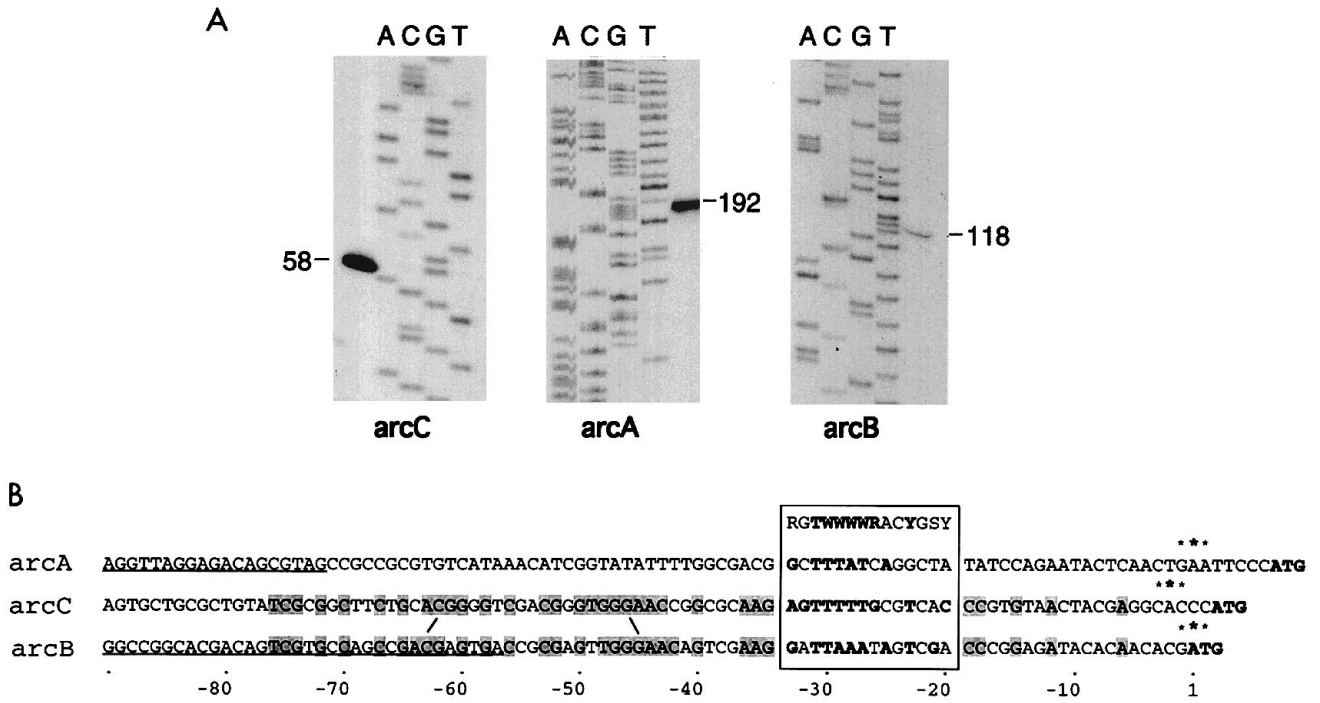

FIG. 5. Determination of 5′-ends of *arc* gene transcripts. (A) The 5′-ends of the *arc* transcripts starting near the structural genes were determined by primer extension analyses as described in the text. The M13 sequence (lanes A, C, G, T) was used for size determination. Numbers indicate fragment lengths. (B) The sequences upstream of the *arcA*, *arcC*, and *arcB* genes including the first codon (shown in bold) are presented. Presumed transcriptional start sites, as determined by the primer extension analyses, are indicated by larger asterisks. The smaller asterisks indicate that because of the usage of the M13 sizing standard determination of the 5′ end may be uncertain by one or two nucleotides. The distal promoter element is boxed. For comparison, the consensus sequence for a strong halobacterial promoter (9) is shown above the sequences and the most important nucleotides for promoter strength are shown in bold. Nucleotides upstream of the *arc* genes which match the consensus promoter are also highlighted in bold. The 3′ ends of the ORF preceding the *arcA* and *arcB* genes are underlined. Nucleotides which are identical upstream of the *arcC* and *arcB* genes are shaded.

within the *arc* gene cluster, the 5′-ends of the transcripts were determined by primer extension analyses. As is shown in Fig. 5, one distinct 5′-end could be detected for the transcripts starting near the *arcA*, *arcC*, and *arcB* genes. In the required spacing to the putative transcriptional start sites, sequences with similarities to the consensus sequence for a constitutive halobacterial promoter (9, 23) were found, indicating that the structural genes are transcribed from individual promoters. Furthermore, a region of striking sequence similarity was found directly upstream of the putative distal promoter elements for the *arcB* and *arcC* genes (Fig. 5). If two sequence motifs are included which have an offset of 1 nt, 27 of 43 nt (63%) are identical. It is tempting to speculate that this region is involved in the regulation of transcription. This region (−34 to −76) is totally different upstream of *arcA*, which is in agreement with the difference in transcript induction found for the *arcA* compared with the *arcC* and *arcB* genes (see above).

Comparison of the ADI pathways of *P. aeruginosa* and *H. salinarium*. Previously the ADI pathway has been most thoroughly studied in *P. aeruginosa*. The present study opens the possibility of comparing ADI pathways from a bacterial and an archaeal species. While both organisms use the same pathway of fermentative arginine degradation, analysis of the halobacterial *arc* gene cluster revealed more differences from than similarities with the *arc* gene cluster for *P. aeruginosa.*

In *P. aeruginosa* the genes are organized in the *arcDABC* operon, while, in contrast, the gene order *arcRACB* was found in *H. salinarium*. The degree of similarity among the three enzymes is very different. The CKs and the cOTCases of both organisms share about 40% amino acid identity, while the similarity between the ADIs is low (15%) and only short regions are conserved. The cOTCase from *H. salinarium* is more distantly related to the cOTCase from *P. aeruginosa* than to its anabolic OTCase or to anabolic OTCases from two gram-positive bacteria (37).

In *P. aeruginosa* a single promoter was found, and the polycistronic primary transcripts are subsequently processed into smaller mRNAs (reference 18 and references therein). In contrast, the analyses of transcript abundances and transcript induction levels, the differential regulation of *arcA* compared with *arcC* and *arcB*, and the existence of putative distal promoter elements and a conserved sequence motif indicate the presence of four individual promoters governing *arc* gene transcription in *H. salinarium.*

In *P. aeruginosa* transcription is controlled by the positive regulator ANR, a member of the FNR protein family. There is no indication that an ANR-like protein exists in *H. salinarium*. ANR boxes, i.e., consensus sequences for ANR binding (19), could not be detected upstream of the halobacterial *arc* genes. In contrast, the gene for a putative regulator without any similarity to ANR was found to be included in the halobacterial *arc* gene cluster. Taken together, these data indicate that the ADI pathway was not present in the last common ancestor of the two species but evolved independently in *H. salinarium* and *P. aeruginosa.*

ACKNOWLEDGMENTS

This work was supported by the Deutsche Forschungsgemeinschaft with the grant So 264/1.

We thank Uta Schimanko for supply of the oligonucleotides and Karin Rodewald for help with sequence determination. We specifically thank Johannes Hain for helping us with the 5′-end determination of the different transcripts, the Tübinger student Gunnar Jacobs for isolation of the ADI, and Friedrich Lottspeich for cleavage of the protein, peptide isolation, and N-terminal sequence determination. We are grateful to Dieter Oesterhelt for continuous support and interest in our work over the last years.

REFERENCES

1. **Baur, H., E. Luethi, V. Stalon, A. Mercenier, and D. Haas.** 1989. Sequence analysis and expression of the arginine deiminase and carbamate kinase genes of *Pseudomonas aeruginosa*. Eur. J. Biochem. **179:**53–60.
2. **Baur, H., V. Stalon, P. Falmagne, E. Luethi, and D. Haas.** 1987. Primary and quaternary structure of the catabolic ornithine carbamoyltransferase from *Pseudomonas aeruginosa*. Eur. J. Biochem. **166:**111–117.
3. **Böhm, G., and R. Jaenicke.** 1994. A structure-based model for the halophilic adaptation of dihydrofolate reductase from *Halobacterium volcanii*. Protein Eng. **207:**213–220.
4. **Bolotin, A., and S. Biro.** 1990. Nucleotide sequence of the putative regulatory gene and major promoter region of the *Streptomyces griseus* glycerol operon. Gene **87:**151–152.
5. **Chomczynski, P.** 1992. One-hour downward alkaline capillary transfer for blotting of DNA and RNA. Anal. Biochem. **201:**134–139.
6. **Chomczynski, P., and N. Sacchi.** 1987. Single-step method of RNA isolation by acid guanidinium thiocyanate-phenol-chloroform extraction. Anal. Biochem. **162:**156–159.
7. **Cline, S. W., W. L. Lam, R. L. Charlebois, L. C. Schalkwyk, and W. F. Doolittle.** 1989. Transformation methods for halophilic archaebacteria. Can. J. Microbiol. **35:**148–152.
8. **Cunin, R., N. Glansdorff, A. Pierard, and V. Stalon.** 1986. Biosynthesis and metabolism of arginine in bacteria. Microbiol. Rev. **50:**314–352.
9. **Danner, S., and J. Soppa.** 1996. Characterization of the distal promoter element of halobacteria *in vivo* using saturation mutagenesis and selection. Mol. Microbiol. **19:**1265–1276.
10. **Dower, W., J. Miller, and C. Ragsdale.** 1988. High efficiency transformation of *E. coli* by high voltage electroporation. Nucleic Acids Res. **16:**6127–6145.
11. **Drummond, M. H., A. Contreras, and L. A. Mitchenall.** 1990. The function of isolated domains and chimaeric proteins constructed from the transcriptional activators NifA and NtrC of *Klebsiella pneumoniae*. Mol. Microbiol. **4:**29–37.
12. **Dym, O., M. Mevarech, and J. L. Sussman.** 1995. Structural features that stabilize halophilic malate dehydrogenase from an archaebacterium. Science **267:**1344–1346.
13. **Eisenberg, H., M. Mevarech, and G. Zacchai.** 1992. Biochemical, structural, and molecular genetic aspects of halophilism. Adv. Protein Chem. **43:**1–62.
14. **Eraso, J. M., and S. Kaplan.** 1994. *prrA*, a putative response regulator involved in oxygen regulation of photosynthesis gene expression in *Rhodobacter sphaeroides*. J. Bacteriol. **176:**32–43.
15. **Galimand, M., M. Gamper, A. Zimmermann, and D. Haas.** 1991. Positive FNR-like control of anaerobic arginine degradation and nitrate respiration in *Pseudomonas aeruginosa*. J. Bacteriol. **173:**1598–1606.
16. **Galinier, A., D. Negre, J. Cortay, S. Marcandier, S. Maloy, and A. Cozzone.** 1990. Sequence analysis of the iclR gene encoding the repressor of the acetate operon in *Salmonella typhimurium*. Nucleic Acids Res. **18:**3656.
17. **Gamper, M., B. Ganter, M. Polito, and D. Haas.** 1992. RNA processing modulates the expression of the arcDABC operon in *Pseudomonas aeruginosa*. J. Mol. Biol. **226:**943–957.
18. **Gamper, M., and D. Haas.** 1993. Processing of the *Pseudomonas* arcDABC mRNA requires functional RNase E in *Escherichia coli*. Gene **129:**119–122.
19. **Gamper, M., A. Zimmermann, and D. Haas.** 1991. Anaerobic regulation of transcription initiation in the arcDABC operon of *Pseudomonas aeruginosa*. J. Bacteriol. **173:**4742–4750.
20. **Genetics Computer Group.** 1991. Program manual for the GCG package, version 7. Genetics Computer Group, Madison, Wis.
21. **Haas, D., M. Galimand, M. Gamper, and A. Zimmermann.** 1990. Arginine network in *Pseudomonas aeruginosa*: specific and global controls, p. 303–316. *In* S. Silver, A. M. Chakrabarty, B. Iglewski, and S. Kaplan (ed.), Pseudomonas. American Society for Microbiology, Washington, D.C.
22. **Haas, D., M. Gamper, and A. Zimmermann.** 1992. Anaerobic control in *Pseudomonas aeruginosa*, p. 177–187. *In* E. Galli, S. Silver, and B. Witholt (ed.), Pseudomonas. American Society of Microbiology, Washington, D.C.
23. **Hain, J., W.-D. Reiter, U. Hüdepohl, and W. Zillig.** 1992. Elements of an archaeal promoter defined by mutational analysis. Nucleic Acids Res. **20:** 5423–5428.
24. **Harasawa, R., K. Koshimizu, M. Kitagawa, K. Asada, and I. Kato.** 1992. Nucleotide sequence of the arginine deiminase gene of *Mycoplasma hominis*. Microbiol. Immunol. **36:**661–665.
25. **Hartmann, R., H. Sickinger, and D. Oesterhelt.** 1980. Anaerobic growth of halobacteria. Proc. Natl. Acad. Sci. USA **77:**3821–3825.
26. **Hindle, Z., and C. Smith.** 1994. Substrate induction and catabolite repression of the *Streptomyces coelicolor* glycerol operon are mediated through the GylR protein. Mol. Microbiol. **12:**737–745.
27. **Kondo, K., H. Sone, H. Yoshida, T. Toida, K. Kanatani, Y. M. Hong, N. Nishino, and J. Tanaka.** 1990. Cloning and sequence analysis of the arginine deiminase gene from *Mycoplasma arginini*. Mol. Gen. Genet. **221:**81–86.
28. **Lanyi, J.** 1974. Salt-dependent properties of proteins from extremely halophilic bacteria. Bacteriol. Rev. **38:**272–290.

28a. **Lottspeich, F.** Personal communication.

29. **Lüthi, E., H. Baur, M. Gamper, F. Brunner, D. Villeval, A. Mercenier, and D. Haas.** 1990. The arc operon for anaerobic arginine catabolism in *Pseudomonas aeruginosa* contains an additional gene, *arcD*, encoding a membrane protein. Gene **87:**37–43.
30. **Moore, R. L., and B. J. McCarthy.** 1969. Characterization of the deoxyribonucleic acid of various strains of halophilic bacteria. J. Bacteriol. **99:**291–298.
31. **Oginsky, E.** 1957. Isolation and determination of arginine and citrulline. Methods Enzymol. **3:**639–643.
32. **Ohno, T., O. Ando, K. Sugimura, M. Taniai, M. Suzuki, S. Fukuda, Y. Nagase, K. Yamamoto, and I. Azuma.** 1990. Cloning and nucleotide sequence of the gene encoding arginine deiminase of *Mycoplasma arginini*. Infect. Immun. **58:**3788–3795.
33. **Protein Identification Resource.** 1991. National Biomedical Research Foundation, Georgetown University Medical Center, Washington, D.C.
34. **Reiter, W.-D., P. Palm, and W. Zillig.** 1988. Analysis of transcription in the archaebacterium *Sulfulobus* indicates that archaebacterial promoters are homologous to eucaryotic pol II promoters. Nucleic Acids Res. **16:**1–19.
35. **Reverchon, S., W. Nasser, and J. Robert-Baudouy.** 1991. Characterization of kdgR, a gene of *Erwinia chrysanthemi* that regulates pectin degradation. Mol. Microbiol. **5:**2203–2216.
36. **Rosenshine, I., T. Zusman, R. Werczberger, and M. Mevarech.** 1987. Amplification of specific DNA sequences correlates with resistance of the archaebacterium *Halobacterium volcanii* to the dihydrofolate reductase inhibitors trimethoprim and methotrexate. Mol. Gen. Genet. **208:**518–522.
37. **Ruepp, A., H. N. Müller, F. Lottspeich, and J. Soppa.** 1995. Catabolic ornithine transcarbamylase of *Halobacterium halobium* (*salinarium*): purification, characterization, sequence determination, and evolution. J. Bacteriol. **177:**1129–1136.
38. **Sambrook, J., E. F. Fritsch, and T. Maniatis.** 1989. Molecular cloning: a laboratory manual, 2nd ed. Cold Spring Harbor Laboratory Press, Cold Spring Harbor, N.Y.
39. **Smith, C., and K. Chater.** 1988. Structure and regulation of controlling sequences for the *Streptomyces coelicolor* glycerol operon. J. Mol. Biol. **204:** 569–580.
40. **Soppa, J.** 1994. Compilation of halobacterial protein-coding genes, the halobacterial codon usage table and its use. Syst. Appl. Microbiol. **16:**725–733.
41. **Soppa, J., J. Duschl, and D. Oesterhelt.** 1993. Bacterioopsin, haloopsin and sensory opsin I of the halobacterial isolate *Halobacterium* sp. strain SG1: three new members of a growing family. J. Bacteriol. **175:**2720–2726.
42. **Sunnarborg, A., D. Klumpp, T. Chung, and D. C. LaPorte.** 1990. Regulation of the glyoxylate bypass operon: cloning and characterization of *iclR*. J. Bacteriol. **172:**2642–2649.
43. **Wagner, G., D. Oesterhelt, G. Krippahl, and J. Lanyi.** 1981. Bioenergetic role of halorhodopsin in *Halobacterium halobium* cells. FEBS Lett. **131:**341–345.
44. **Wilimzig, M.** 1985. LiCl-boiling method for plasmid mini-preps. Trends Genet. **1985:**158.