Genes of De Novo Pyrimidine Biosynthesis from the Hyperthermoacidophilic Crenarchaeote
*Sulfolobus acidocaldarius*: Novel Organization in a Bipolar Operon

Thia-Lin Thia-Toong,1 Martine Roovers,1 Virginie Durbecq,2 Daniel Gigot,2 Nicolas Glansdorff,1,2 and Daniel Charlier1*

Erfelijkheidseleer en Microbiologie, Vrije Universiteit Brussel, 1 and Laboratoire de Microbiologie, Université Libre de Bruxelles and Institut de Recherches Microbiologiques J.-M. Wiame, 2 B-1070 Brussels, Belgium

Received 8 March 2002/Accepted 31 May 2002

Sequencing a 8,519-bp segment of the *Sulfolobus acidocaldarius* genome revealed the existence of a tightly packed bipolar pyrimidine gene cluster encoding the enzymes of de novo UMP synthesis. The G+C content of 35.3% is comparable to that of the entire genome, but intergenic regions exhibit a considerably lower percentage of strong base pairs. Coding regions harbor the classical excess of purines on the coding strand, whereas intergenic regions do not show this bias. Reverse transcription-PCR and primer extension experiments demonstrated the existence of two polycistronic messengers, *pyrEF-orf8* and *pyrB1-orf1-pyrCD-orf2-orf3-orf4*, initiated from a pair of divergent and partially overlapping promoters. The gene order and the grouping in two wings of a bipolar operon constitute a novel organization of *pyr* genes that also occurs in the recently determined genome sequences of *Sulfolobus solfataricus* P2 and *Sulfolobus tokodaii* strain 7; the configuration appears therefore characteristic of *Sulfolobus*. The quasi-leaderless *pyrE* and *pyrB* genes do not bear a Shine-Dalgarno sequence, whereas the initiation codon of promoter-distal genes is preceded at an appropriate distance by a sequence complementary to the 3′ end of 16S rRNA. The polycistronic nature of the *pyr* messengers and the existence of numerous overlaps between contiguous open reading frames suggests the existence of translational coupling. *pyrB* transcription was shown to be approximately twofold repressed in the presence of uracil. The mechanism underlying this modulation is as yet unknown, but it appears to be of a type different from the various attenuation-like mechanisms that regulate *pyrB* transcription in bacteria. In contrast, the *pyrE-pyrB* promoter/control region harbors direct repeats and imperfect palindromes reminiscent of target sites for the binding of a hypothetical regulatory protein(s).

De novo synthesis of UMP is universally performed in six steps catalyzed by conserved enzymes in all three domains of life (Fig. 1). In lower eukaryotes, such as the yeast *Saccharomyces cerevisiae*, the carbamoylphosphate synthetase (CPSase) (EC 6.3.5.5) and aspartate carbamoyltransferase (ATCase) (EC 2.1.3.2) activities are fused in a single polypeptide chain, also containing an inactive version of dihydroorotase (DHOase) (EC 3.5.2.3). In animals the pyrimidine biosynthetic activities are highly organized; thus, the trifunctional CAD protein shows CPSase, ATCase, and DHOase activities assembled in a single polypeptide chain, and similarly, orotate phosphoribosyltransferase (OPRTase) (EC 2.4.2.10) and orotidine-5′-monophosphate decarboxylase (OMPdecase) (EC 4.1.1.23) are fused in a bifunctional protein. In bacteria and archaea the different reactions are performed by monofunctional enzymes, but some of them do interact, at least transiently. Thus, in *Pseudomonas putida* and *Pseudomonas aeruginosa*, *pyrB* overlaps a *pyrC*′ gene encoding a nonfunctional DHOase, which is required for the assembly of the functional dodecameric ATCase (49), and feedback inhibition and thermostability of ATCase from the extreme thermophilic bacterium *Thermus ZO5* are conferred by coexpression of *pyrB* and the adjacent *bbc* (for “between *b* and *c*”) and *pyrC* genes, respectively (58, 59). In hyperthermophilic bacteria and archaea, CPSase or carbamate kinase-like CPSase appears to interact physically with ATCase and its paralogue, ornithine carbamoyltransferase, from the arginine biosynthetic pathway, thereby allowing substrate channeling and protection of the thermolabile and potentially toxic carbamoylphosphate from the hot aqueous environment (35, 57, 43, 44, 40). Indeed, since the pathway includes several thermolabile and energy-rich substrates and precursors, some of which decompose into toxic degradation products, the very possibility of pyrimidine biosynthesis is not obvious in hot environments. Yet all investigated hyperthermophilic bacteria and archaea use the same, classical route; evidently, this situation necessitates adapted strategies.

The *pyr* genes are organized in a vast number of different ways; most microorganisms exhibit an intermediate scattering, with monocistronic and polycistronic messengers. In *Esche-
richia coli and Salmonella enterica serovar Typhimurium, the
pyr genes are all scattered throughout the chromosome; only carA and carB, encoding the glutamine amidotransferase and catalytic subunits of CPSase, and pyrB and pyrI, encoding the catalytic and regulatory subunits of ATCase, are grouped into two small operons (for a review, see reference 42). The other extreme occurs in Bacillus subtilis, where 10 cistrons, comprising the pyrimidine biosynthetic genes, the gene for the regulator (pyrR), and the gene for a uracil permease (pyrP), are grouped in a compact, unipolar operon (45). A similar situation is found with the extreme thermophilic gram-positive bacterium Bacillus caldolyticus (17, 18). In the completely sequenced genomes of several euryarchaea, including methanogens, halophiles, and hyperthermophilic Thermococcales, the pyrimidine genes are highly dispersed, with only pyrB and pyrI showing more systematic clustering. An intermediate situation prevails with Aeropyrum pernix (three small clusters, pyrBI, pyrDC, and pyrFE) and Pyrobaculum aerophilum (pyrFB and pyrDEI clusters), the only crenarchaeotes not belonging to the Sulfolobales genus for which the entire genome sequences have been established (31, 15).

In Bacteria, pyrimidine-specific control of gene expression is exerted by a multitude of distinct mechanisms. Remarkably, even within a single organism like E. coli, the scattered genes and operons are noncoordinately regulated by multiple mechanisms, including UTP-sensitive transcription attenuation (pyrBI and pyrE), UTP-sensitive reiterative transcription (pyrBI and carAB), nucleotide-pool-dependent translational control (pyrC and pyrD), mRNA stability (pyrF) or still, as demonstrated for the carAB operon, complex transcription initiation control imposed by the concerted action of several multifunctional proteins, some of which combine catalytic and regulatory activities (5, 6, 20, 26, 33, 42). The pyrimidine operon of B. subtilis is submitted to transcriptional attenuation exerted by PyrR, an RNA binding protein that also has residual uracil phosphoribosyltransferase activity (54, 56). UTP-dependent binding of PyrR to three untranslated regions of the polycistronic mRNA disrupts the formation of the antiterminator stem-loop structures (3). An analogous mechanism appears to be operative in the gram-positive bacteria B. caldolyticus (17, 18), Lactococcus lactis (38, 39), and Enterococcus faecalis (36). An alternative PyrR-dependent attenuation mechanism was proposed in the extreme thermophilic gram-negative bacterium Thermus strain Z05 (58). In this organism, transcriptional attenuation would result from the coupling of transcription and translation of a short open reading frame (ORF), whose ribosome binding site is occluded by pyrimidine-dependent binding of PyrR to the cognate mRNA. PyrR binding would prevent translation of the leader peptide, thus promoting the formation of the terminator structure that leads to reduced expression of the downstream genes. Data on modulation of pyrimidine biosynthetic enzyme specific activities in archaea are extremely scarce and to the best of our knowledge have been reported only for Sulfolobus acidocaldarius (23; also this work); no information is yet available on the mechanism(s) that imposes pyrimidine-specific control of gene transcription in archaea.

In this study we present the isolation, organization, and characterization of the pyrimidine gene cluster of S. acidocaldarius, an aerobic thermoacidophilic sulfur-oxidizing cren-
chaetothat grows optimally around 80°C and pH 3.0 (4). In several aspects, *S. acidocaldarius* is the best-studied member of the *Sulfolobales*; its genome sequence has not yet been established, whereas those of the relatives *Sulfolobus solfataricus* P2 and *Sulfolobus tokodaii* strain 7 have recently been determined (52, 32). We demonstrate that the pyrimidine gene cluster of *S. acidocaldarius* constitutes a bipolar operon transcribed from two major promoters, an unprecedented organization for pyrimidine genes that turns out to be characteristic of *Sulfolobus*. We determined transcription initiation sites and polycistronic messenger compositions and demonstrate pyrimidine-specific modulation of transcription initiation. We performed amino acid sequence comparisons and analyzed in silico the nucleotide sequence in terms of transcription and translation initiation and stop signals.

**MATERIALS AND METHODS**

**Strains, media, and growth conditions.** Genotypes and descriptions of strains and plasmids are given in Table 1. *S. acidocaldarius* (type strain DSM 639) was grown aerobically at 75°C on a rotary shaker platform in either complex medium or minimal medium as described previously (11, 13). Growth conditions for *E. coli* were described previously (19). Ampicillin was used at a concentration of 25 μg/ml, kanamycin at 30 μg/ml, and uracil at 50 μg/ml.

**DNA preparations and manipulations, sequencing strategy, and amino acid sequence comparisons.** Plasmid DNA extraction was based on the alkaline sodiumpdodecyl sulfate lysis method and performed with the commercial Qiagen Spin Miniprep Kit (Qiagen). Oligonucleotides were purchased from Gibco BRL (Table 2). Nucleic acid digestion, ligation, phosphorylation, and dephosphorylation of DNA fragments and oligonucleotides were performed with commercial enzymes and buffers according to the manufacturer’s instructions (Roche). Competent cells were prepared by the CaCl2 method (10). DNA sequencing was performed with the commercial Expand reverse transcriptase (Roche) in the commercial buffer and in the presence of four deoxynucleoside triphosphates, 10 mM dithiothreitol, and 50 U of Expand DNA polymerase (Promega) and in the commercial buffer. Initial denaturation of 100 mM Tris-HCl, pH 8.8, and sonicated for 6 min. Cell debris was removed designed on the basis of the established sequences of the subclones. Amino acid sequences of *S. acidocaldarius* enzymes were used as queries to retrieve similar sequences from databases using the BLASTP program. Multiple and pairwise alignments of amino acid sequences were generated using the Clustal W program (55).

**Reverse transcriptase primer extension.** Total RNA was prepared with the Life Technologies procedure using the Trizol reagent (8) from 200-ml cultures of *S. acidocaldarius* cells grown in complex medium, minimal medium, and minimal medium supplemented with uracil (50 μg/ml) and uracil at the exponential phase. Total RNA (100 μg) was mixed with about 40,000 cpm of [5–32P]-end-labeled oligonucleotide primer, and after overnight hybridization at 42°C this was elongated with 10 U of avian myeloblastosis virus reverse transcriptase (Roche) at 40°C for 1 h as described previously (7). Chain-terminating DNA sequencing reactions of the noncoding strand obtained with pSPYR3 plasmid DNA as the template and the same 5‘-end-labeled nucleotides used as the primer were applied as reference ladders.

**Reverse transcription-PCR (RT-PCR).** Total RNA was extracted with the RNeasy Multi kit (Qiagen) according to the manufacturer’s instructions for *S. acidocaldarius* cells grown in minimal medium and harvested in the exponential phase. To eliminate all possible traces of contaminant DNA, a supplementary DNase I treatment was performed by incubating 1 U of enzyme (RNase free; Roche) per μg of RNA at 37°C for 30 min. DNase I was inactivated by incubation at 75°C for 5 min. CDNA synthesis was performed in a total volume of 20 μl, with 1.0 μg of RNA, 30 pmol of oligonucleotide, 1.0 mM concentration (each) of the four deoxynucleoside triphosphates, 10 mM dithiothreitol, and 50 U of Expand Reverse Transcriptase (Roche) in the commercial buffer and in the presence of 20 U RNAse inhibitor (Roche). RNA and primer were heated at 65°C for 10 min and then mixed with the other components and incubated at 42°C for 90 min. The reaction was stopped by cooling on ice. Controls were performed in the absence of Expand reverse transcriptase. Three-microliter cDNA aliquots were used as a template in the PCR amplification step with different combinations of oligonucleotide pairs (0.6 μM each) in a total volume of 50 μl with a 0.2 μM concentration (each) of the four deoxynucleoside triphosphates and 1.5 μl of Phusion DNA polymerase (Promega) in the commercial buffer. Initial denaturation was for 5 min at 94°C, followed by 30 cycles of synthesis comprising 1 min of denaturation at 94°C, 30 s of annealing at 50°C, and elongation at 72°C for 2 min per kilobase. The amplification was ended with a 10-min elongation at 72°C. Seventeen-microliter aliquots of the various PCRs were analyzed by agarose gel electrophoresis.

**Enzyme assays.** A TCase activity was measured using the colorimetric method of Prescott and Jones on dialyzed cell extracts as described previously (11). For OPRTase assays, bacterial pellets from 100-ml cultures were resuspended in 3 ml of 100 mM Tris-HCl, pH 8.8, and sonicated for 6 min. Cell debris was removed

---

**TABLE 1. Strains and plasmids used in this work**

<table>
<thead>
<tr>
<th>Strain or plasmid</th>
<th>Genotype</th>
<th>Source or reference</th>
</tr>
</thead>
<tbody>
<tr>
<td><em>S. acidocaldarius</em> DSM639</td>
<td>Wild-type</td>
<td>Deutsche Sammlung von Microorganismen</td>
</tr>
<tr>
<td><strong>E. coli</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>XLI-Blue</td>
<td>F‘::Tn10 proA”B” lacZΔ(lacZ)M15/recA1 endA1 gyrA96 (Nal’) thi hsdR17( r m ) supE44 relA1 lac endA hpyR4</td>
<td>Stratagene</td>
</tr>
<tr>
<td>JM101</td>
<td>F‘ triX36 lacPΔ(lacZ) M15 proA”B” supE thi Δ(lac-proA-B) strA slcB15</td>
<td>41</td>
</tr>
<tr>
<td>HMS174 (DE3)/pLysS</td>
<td>F‘ recA1 hsdR ( r m ) Rif’ (DE3)/pLysS (Cam’)</td>
<td>Novagen</td>
</tr>
<tr>
<td><strong>Plasmids</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>pUC18</td>
<td>Ap’, lacZ cloning vector</td>
<td>60</td>
</tr>
<tr>
<td>pGEM-7 Zf(+)</td>
<td>Ap’, cloning vector</td>
<td>Promega</td>
</tr>
<tr>
<td>pKK223-3</td>
<td>Ap’, IPTG-inducible and T7 RNA-polymerase dependent expression vector</td>
<td>Novagen</td>
</tr>
<tr>
<td>pETa3</td>
<td></td>
<td></td>
</tr>
<tr>
<td>pSPYR3</td>
<td>96-bp Prep genomic fragment of <em>S. acidocaldarius</em> in pKK223-3</td>
<td>11</td>
</tr>
<tr>
<td>pSPYRE</td>
<td>1,493-bp SacI genomic fragment of <em>S. acidocaldarius</em> cloned as BamHI</td>
<td>This study</td>
</tr>
<tr>
<td>pSPYRF</td>
<td>1,316-bp EcoRI genomic fragment of <em>S. acidocaldarius</em> cloned as BamHI</td>
<td>This study</td>
</tr>
<tr>
<td>pET3PYRE</td>
<td>pyrE coding region of <em>S. acidocaldarius</em> in pET3a</td>
<td>This study</td>
</tr>
</tbody>
</table>
by centrifugation. The assay mixture contained 300 μmol of Tris-HCl (pH 8.8), 0.25 μmol of sodium orotate, 25 μmol of MgCl₂, and 0.4 μmol of 5'-phospho-D-ribosyl-pyrophosphate (PRPP) and bacterial extract in a final volume of 3 ml. The reaction was monitored by measuring the decrease in absorbance measured at 295 nm. Protein was measured by the method of Lowry et al. (37).

**RESULTS**

Cloning and nucleotide sequence determination of the pyrimidine gene cluster of *S. acidocaldarius*. We previously reported the isolation of the *S. acidocaldarius* pyrB and pyrI genes on a 6.9-kb genomic PstI fragment (Fig. 2) by complementation of an *E. coli* mutant deficient for ATCase (11). A Southern blot analysis performed under stringent conditions indicated that this fragment indeed originates from *S. acidocaldarius* (data not shown). The complete 6,845-nucleotide (nt)-long sequence was established, and subsequently extended with another 1,674 bp (Fig. 2, left-hand border EcoRI-PstI segment) by sequencing two clones obtained by inverse PCR with SacI- and EcoRI-digested genomic DNA, respectively. Sequence analysis of the 8,519-bp-long contig and comparisons with databases revealed the presence of 13 complete ORFs and two truncated ones, some of which are homologous to pyrimidine genes of other archaea, bacteria, and eukaryotes (see below).

**TABLE 2. Oligonucleotide used in this work**

<table>
<thead>
<tr>
<th>Oligonucleotide</th>
<th>Sequence</th>
<th>Use</th>
</tr>
</thead>
<tbody>
<tr>
<td>E3Sac</td>
<td>5'-CGGGATCCCTGCTGATACACTAACAAGG-3'</td>
<td>Inverse PCR, <strong>pyrE</strong></td>
</tr>
<tr>
<td>E5Bam</td>
<td>5'-GGAGATCCGCTATGATATTGTGTCTTAC-3'</td>
<td>Inverse PCR, <strong>pyrE</strong></td>
</tr>
<tr>
<td>PYRF1</td>
<td>5'-GGGATCCCTTCTCAGCGAGATTTAG-3'</td>
<td>Inverse PCR, <strong>pyrF</strong></td>
</tr>
<tr>
<td>PYRF2</td>
<td>5'-GGGAGATCCGCTATGATATTGTTTG-3'</td>
<td>Inverse PCR, <strong>pyrF</strong></td>
</tr>
<tr>
<td>PYRENde</td>
<td>5'-GGAATTCCTAGTTGTTGGAAGCTCTTGT-3'</td>
<td>Amplification coding region, <strong>pyrE</strong></td>
</tr>
<tr>
<td>PYREBam</td>
<td>5'-CGGATCCCTATGAACTTCTAGTTTCCCTAC-3'</td>
<td>Amplification coding region, <strong>pyrE</strong></td>
</tr>
<tr>
<td>PYREB</td>
<td>5'-CTGTAAAGGGCAAATATATGTCCTTAC-3'</td>
<td>Primer extension</td>
</tr>
<tr>
<td>PYRC</td>
<td>5'-GAGATCAACTGATGCTTGCAGGAGCAG-3'</td>
<td>Primer extension</td>
</tr>
<tr>
<td>PYRD</td>
<td>5'-GGTTCAATTTACCTGTTGTAAGTG-3'</td>
<td>Primer extension</td>
</tr>
<tr>
<td>ORF1</td>
<td>5'-GCCCAAGAGCTTCTTTTTCTCCT-3'</td>
<td>Primer extension</td>
</tr>
<tr>
<td>ORF4</td>
<td>5'-CACAGAAAAAGGATCTATAATTCG-3'</td>
<td>Primer extension</td>
</tr>
<tr>
<td>1a</td>
<td>5'-GTTGGCAAGACTAAGGACG-3'</td>
<td>RT-PCR</td>
</tr>
<tr>
<td>1b</td>
<td>5'-CCTAAGACATTTAGTACTGCGG-3'</td>
<td>RT-PCR</td>
</tr>
<tr>
<td>2a</td>
<td>5'-GAGATGTAAGTATTGTTGAGAGACG-3'</td>
<td>RT-PCR</td>
</tr>
<tr>
<td>2b</td>
<td>5'-GTTATACCTGTGTGGAAGCC-3'</td>
<td>RT-PCR</td>
</tr>
<tr>
<td>3a</td>
<td>5'-GAGATCAACTGATGACGGAGCAG-3'</td>
<td>RT-PCR</td>
</tr>
<tr>
<td>3b</td>
<td>5'-GATAGTTGAAAGTAAATTGACG-3'</td>
<td>RT-PCR</td>
</tr>
<tr>
<td>4a</td>
<td>5'-GTTCAATTTACCTGTTGTAAGTG-3'</td>
<td>RT-PCR</td>
</tr>
<tr>
<td>4b</td>
<td>5'-GAGATCAACTGATGCGAGCAG-3'</td>
<td>RT-PCR</td>
</tr>
<tr>
<td>5a</td>
<td>5'-GTTCAATTTACCTGTTGTAAGTG-3'</td>
<td>RT-PCR</td>
</tr>
<tr>
<td>5b</td>
<td>5'-GCAAGACTCCCTTCATGAACTG-3'</td>
<td>RT-PCR</td>
</tr>
<tr>
<td>10b</td>
<td>5'-CACAGAAAAAGGATCTATAATTGC-3'</td>
<td>RT-PCR</td>
</tr>
</tbody>
</table>

![FIG. 2. Schematic drawing of the pyrimidine gene cluster of *S. acidocaldarius* and surrounding genes. Grey boxes represent pyrimidine biosynthetic genes. Cross-hatched boxes indicate potential ORFs with unknown function (orf2 to orf8) and the transcriptional regulator Sa-Lrp. Arrows represent mRNA molecules and indicate the start, approximate stop, and orientation of transcription. Positions of the recognition sites for the restriction enzymes EcoRI, SacI, and PstI are indicated (in nucleotides).](http://jb.asm.org/)

The 8,519-bp-long sequence has been deposited in the EMBL data bank under accession number AJ459777.
pyrE, and pyrF encoding DHOase, dihydroorotate dehydrogenase (DHOdehase), OPRTase, and OMPdecase, respectively (Fig. 2). pyrI and pyrC are separated by the 177-amino-acid (aa)-long orf1, involved in the electron transport associated with the formation of orotate (see below). pyrD is followed by six relatively short ORFs (see below), one of which corresponds to Sa-lrp, encoding an Lrp-like archaeal transcription regulator (7, 13). Thus, the entire 8,519-bp-long EcoRI-PvuI segment of the S. acidocaldarius genome contains the information for the pyrE, pyrF, and truncated orf8 gene products on one strand, and for the pyrB, pyrI, orf1, pyrC, pyrD, orf2, orf3, orf4, Sa-lrp, orf5, orf6, and truncated orf7 gene products on the complementary strand (Fig. 2). The G+C content of the cluster is 35.3%, comparable to that of the entire genome (36 to 38%) (61, 24).

This divergent pyr gene cluster encodes the functional enzymes responsible for de novo UTP synthesis, as demonstrated by sequence analysis of mutant alleles and enzyme assays. The pyrB allele from mutant DG64, affected in ATCase activity (23), proved to bear a substitution of phenylalanine (TTT) for His than at 37°C, whereas no activity was measurable in S. acidocaldarius pyrE gene products on the complementary strand (Fig. 2). The G+C content of the cluster is 35.3%, comparable to that of the entire genome (36 to 38%) (61, 24).

Pyrococcus abyssi Type I enzymes (14), and Type II enzymes are predominantly found in bacteria, gram-negative archaea, and cyanobacteria (active enzymes), but also the inactive archaeal DHOases (monofunctional) identified till now. Type II enzymes are predominantly found in bacteria, gram-negatives and cyanobacteria (active enzymes), but also the inactive DHOase sequences of several fluorescent Pseudomonas species, the multifunctional active enzyme of mesophilic and thermophilic gram-positive bacteria, the active monofunctional Ura2 gene product of S. cervesiae, active multifunctional hamster CAD, and all the archaean DHOases (monofunctional) identified till now. Type II enzymes are predominantly found in bacteria, gram-negatives and cyanobacteria (active enzymes), but also the inactive DHOase-like sequence embedded in the multifunctional Ura4 gene product from S. cervesiae belongs to the Type II. The best scores for S. acidocaldarius DHOase are with the crenarchaeotes: S. tokodaii (69% identity), S. solfataricus (69 and 59%) followed by P. aerophilum (53 and 51%), A. pernix (50 and 35%), and Pyrococcus horikoshii (50 and 47%).

Sequence comparisons of active DHOases and silent DHOase domains of multifunctional enzymes indicate that DHOases fall into two main phylogenetically distinct groups (14). Type I includes the silent DHOase-like sequences of several fluorescent Pseudomonas species, the monofunctional active enzyme of mesophilic and thermophilic gram-positive bacteria, the active monofunctional Ura2 gene product of S. cervesiae, active multifunctional hamster CAD, and all the archaean DHOases (monofunctional) identified till now. Type II enzymes are predominantly found in bacteria, gram-negatives and cyanobacteria (active enzymes), but also the inactive DHOase-like sequence embedded in the multifunctional Ura4 gene product from S. cervesiae belongs to the Type II. The best scores for S. acidocaldarius DHOase are with the crenarchaeotes: S. tokodaii (69% identity), S. solfataricus (57%), A. pernix (34%), and P. aerophilum (31%). These enzymes are clearly Type I enzymes (14), and S. acidocaldarius DHOase also shows the conserved stretch (consensus: PGLV) of 4 aa (PASV, positions 45 to 48), typical for Type I enzymes, that immediately precedes the aspartate residue of the first DHOase signature, DLHVHRVGA (positions 49 to 57 in S. acidocaldarius DHOase). Sequence analyses strongly suggest that S. acidocaldarius pyrC translation initiates at a TTG codon and overlaps

FIG. 3. Nucleotide sequences of the border regions of adjacent genes in the divergent pyrimidine gene cluster. Potential start codons are indicated in bold capital letters and are overlined; stop codons are indicated in bold capitals and underlined. Potential SD sequences are indicated in bold capitals.
the C terminus of orf1 by 10 aa (Fig. 3), rather than at the in-phase ATG codon 99 (nt) downstream. Several arguments support this proposal: the TTG codon is preceded by a good Shine-Dalgarno (SD) sequence (Table 3), whereas the ATG codon is not; downstream initiation would result in a truncated DHOase missing 34 aa residues that show good sequence conservation with other DHOase sequences; moreover, this 102-nt-long stretch separating the TTG and ATG codons shows all the characteristics of a coding region (Pu:Py ratio, 2.90; G+C content, 34.3%) rather than of an intergenic part (see also below, gene organization in the divergent gene cluster) and finally, a similar overlap of PyrC with the C terminus of Orf1 appears to exist in S. solfataricus, though not in S. tokodaii.

The formation of orotate constitutes the sole redox reaction in de novo UMP biosynthesis. On the basis of phylogenetic studies and enzyme characterizations, DHOdehase sequences have been subdivided into two major groups (29, 53). Class I enzymes (predominantly present in gram-positive bacteria and archaea) are cytosolic homodimers (pyrD-encoded) that use NAD as the electron acceptor (type 1B), whereas class II enzymes (gram-positive bacteria and higher eukaryotes) are monomeric enzymes constituted by a catalytic subunit and a new type of electron transfer subunit (53).

The best fits of S. acidocaldarius ORPTase are with the crenarchaeotes S. tokodaii (57% identity), S. solfataricus (55%), A. pernix (39%), and P. aerophilum (36%). The PRPP binding and active site is well conserved (VIVVDD-VAATTGGS, residues 108 to 120 in the S. acidocaldarius enzyme).

The best fits of S. acidocaldarius OMPdecase are with archaean homologues, the best being with S. tokodaii (60% identity), S. solfataricus (48%), A. pernix (43%), M. jannashii (32%), and P. abyssi (31%).

orf2 shows only very weak amino acid sequence similarity to the pyrK (pyrDII)-encoded electron transfer subunits and is much shorter, but it bears an iron-sulfur cluster. Since orf1 from S. acidocaldarius (177 aa) and that from S. solfataricus (208 aa) show 38.9% sequence identity (41% to S. tokodaii 7 [197 aa]), Sulfolobus DHOdehases appear to be heteromorphic enzymes constituted by a catalytic subunit and a new type of electron transfer subunit (53).

### TABLE 3. Function and characteristics of ORFs in the 8,519-bp genomic segment of S. acidocaldarius

<table>
<thead>
<tr>
<th>Gene</th>
<th>Function</th>
<th>Length of ORF (in aa)</th>
<th>Potential initiation codon</th>
<th>Stop codon</th>
<th>Potential SD sequence</th>
<th>Pu:Py ratio</th>
<th>G+C (mol %)</th>
<th>5'-overlap (in nt)</th>
<th>3'-overlap (in nt)</th>
</tr>
</thead>
<tbody>
<tr>
<td>orf8</td>
<td>Hypothetical protein</td>
<td>202 (truncated)</td>
<td>ATG</td>
<td>TAG</td>
<td>GAAaGGAATAA</td>
<td>1.164</td>
<td>36.2</td>
<td>26</td>
<td>?</td>
</tr>
<tr>
<td>pyrF</td>
<td>OMPdecase</td>
<td>219</td>
<td>TTG</td>
<td>TGA</td>
<td>TAA</td>
<td>GGAaGTTGA</td>
<td>1.286</td>
<td>32.6</td>
<td>14</td>
</tr>
<tr>
<td>pyrE</td>
<td>ORPTase</td>
<td>197</td>
<td>ATG</td>
<td>TAG</td>
<td>None</td>
<td>1.309</td>
<td>36.2</td>
<td>14</td>
<td></td>
</tr>
<tr>
<td>pyrB</td>
<td>ATCase catalytic</td>
<td>299</td>
<td>TTA</td>
<td>TGA</td>
<td>None</td>
<td>1.342</td>
<td>35.0</td>
<td></td>
<td></td>
</tr>
<tr>
<td>pyrA</td>
<td>ATCase regulatory</td>
<td>164</td>
<td>ATG</td>
<td>TAA</td>
<td>TGA</td>
<td>GGTTGA</td>
<td>1.800</td>
<td>34.8</td>
<td></td>
</tr>
<tr>
<td>orf1</td>
<td>DHOdehase electron transfer</td>
<td>177</td>
<td>TTG</td>
<td>TGA</td>
<td>None</td>
<td>1.414</td>
<td>36.4</td>
<td>29</td>
<td></td>
</tr>
<tr>
<td>pyrC</td>
<td>DHOase</td>
<td>389</td>
<td>TTG</td>
<td>TAA</td>
<td>None</td>
<td>1.432</td>
<td>37.0</td>
<td>29</td>
<td>23</td>
</tr>
<tr>
<td>pyrD</td>
<td>DHOdehase catalytic</td>
<td>291</td>
<td>TTG</td>
<td>TGA</td>
<td>None</td>
<td>1.473</td>
<td>37.6</td>
<td>23</td>
<td>2</td>
</tr>
<tr>
<td>orf2</td>
<td>Hypothetical protein</td>
<td>103</td>
<td>ATG</td>
<td>TAA</td>
<td>None</td>
<td>1.836</td>
<td>38.2</td>
<td>2</td>
<td>29</td>
</tr>
<tr>
<td>orf3</td>
<td>Hypothetical protein</td>
<td>102</td>
<td>ATG</td>
<td>TAA</td>
<td>None</td>
<td>1.597</td>
<td>34.0</td>
<td>29</td>
<td>2</td>
</tr>
<tr>
<td>orf4</td>
<td>Conserved hypothetical</td>
<td>203</td>
<td>ATG</td>
<td>TGA</td>
<td>None</td>
<td>1.061</td>
<td>37.6</td>
<td>2</td>
<td></td>
</tr>
<tr>
<td>Sa-lrp</td>
<td>Transcriptional regulator</td>
<td>145</td>
<td>ATG</td>
<td>TAA</td>
<td>None</td>
<td>1.450</td>
<td>31.8</td>
<td></td>
<td></td>
</tr>
<tr>
<td>orf5</td>
<td>Hypothetical Zn finger</td>
<td>99</td>
<td>ATG</td>
<td>TAA</td>
<td>None</td>
<td>2.000</td>
<td>40.0</td>
<td></td>
<td></td>
</tr>
<tr>
<td>orf6</td>
<td>Hypothetical protein</td>
<td>80</td>
<td>ATG</td>
<td>TAA</td>
<td>None</td>
<td>2.076</td>
<td>28.8</td>
<td></td>
<td></td>
</tr>
<tr>
<td>orf7</td>
<td>Hypothetical protein</td>
<td>18 (truncated)</td>
<td>ATG</td>
<td></td>
<td></td>
<td>1.040</td>
<td>30.9</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

*?, unknown.
(NP342138) and S. tokodaii (NP377445), but these are not contiguous to pyrF.

**Gene organization in the divergent pyrimidine cluster.**

The genetic information comprised within the 8,519-bp stretch containing the divergent pyr cluster is extremely tightly packed. ORFs cover more than 94% of the sequence, and the gene density of 1.643 per kb largely exceeds those determined for the complete S. solfataricus P2 and S. tokodaii 7 genomes (1.013 and 1.049, respectively). Considering only those genes clearly involved in pyrimidine biosynthesis, the density becomes 1.333 ORFs per kb, which is still high with respect to gene densities calculated for archaeal genomes. In seven instances, adjacent ORFs overlap over a short distance, varying between 2 and 29 nt (Table 3; Fig. 3). Of these, five show an overlap at both the N and C termini, resulting in 1.5% of sequence information being translated into two different proteins. In three other instances, the stop and start codons of contiguous ORFs are separated by a short nucleotide stretch of variable length comprised of between 1 and 133 nt (Table 4).

Of the 15 ORFs, 10 (66.7%) likely start with an ATG codon and five (33.3%) start with TTG, an observation in good agreement with the statistics on the utilization of ATG (65%), TTG (28%), and GTG (14%) as initiation codons in a 156-kb segment with the statistics on the utilization of ATG (65%), TTG (28%), and GTG (14%) as initiation codons in a 156-kb sequence of the related S. solfataricus (51). The most frequent stop codon utilized in the divergent gene cluster is TGA (53.8%), followed by TAA (38.5%) and TAG (7.7%) (Table 3); pyrF shows a pair of adjacent TGA and TAA stop codons. This occurrence of stop codons is unlike the situation reported for the 156-kb sequence of S. solfataricus (51) where TAA is predominant (50%). The coding parts show a purine bias (Pu/Py ratios varying between 1.164 and 2.076 for the individual genes; mean value of 1.416 for the 15 ORFs) (Table 3) close to the typical 55% A+G (ratio of 1.222) found in all sequenced genomes, including archaea, bacteria, and eukaryotes, except for orf4 (Pu/Py ratio of 1.061). In contrast, the intergenic regions with Pu/Py ratios of between 0.78 and 1.133 do not show this bias (Table 4).

Experimental identification of the transcription starts indicate that the 5′ mRNA leader sequences of the two polycistronic pyr messengers (pyrF-orf8 and pyrB-orf1-pyrC-orf2-orf3-orf4; see below) and of monocistronic Sa-lrp (13) are very short and do not harbor a ribosome binding site upstream of the initiation codon. A similar situation might prevail for orf6. Within the polycistronic mRNAs (see below), however, the situation is different and the initiation codon is nearly always preceded at an appropriate distance by a sequence that shows extensive complementarity to the 3′ end of the 16S rRNA (Table 3; Fig. 3). This is not the case for orf4, but as the ATG initiation codon overlaps the stop codon of orf3, translation might possibly proceed without the need for reassociation of ribosomes on the polycistronic messenger.

**Polycistronic messenger analyses and potential transcription stop signals.**

The existence of polycistronic pyr messengers was demonstrated by RT-PCR experiments (Fig. 4). To demonstrate that two or more genes are transcribed as a single mRNA molecule, we used total RNA as template, reverse transcriptase, and pairs of oligonucleotides hybridizing to complementary strands of different genes, in PCR amplification reactions. In the first step, the oligonucleotide complementary to the mRNA of the downstream gene was used to synthesize cDNA. In the second amplification reaction, the oligonucleotide complementary to the cDNA and located in a different gene was added (for details, see Materials and Methods). In the sole presence of RNA as a template, the amplification reaction will be possible only if the mRNA molecule indeed spans the different genes. Thus, we have demonstrated that pyrF and pyrE are part of the same mRNA molecule, initiated at the pyrE promoter (see below for the initiation site). Indeed, using the pair of oligonucleotides 5a (hybridizing to the N-terminal part of the pyrF messenger, for cDNA synthesis) and 5b (hybridizing to the C terminus of pyrE, for further amplification), we detected the expected 226-bp fragment (Fig. 4a and b, lane 9), whereas no amplification at all was detected in the absence of reverse transcription in the assay (lane 10). Similarly, we have demonstrated cotranscription of pyrI and pyrB (oligonucleotides 1b and 1a; 204-bp fragment, lanes 1 and 2), of orf1 and pyrF (oligonucleotides 2b and 2a; 231-bp fragment, lanes 3 and 4), of pyrC and orf1 (oligonucleotides 3b and 3a; 193-bp fragment, lanes 5 and 6), and of pyrD and pyrC (oligonucleotides 4b and 4a; 217-bp fragment, lanes 7 and 8). In further combinations we demonstrated cotranscription of pyrI, orf1, and pyrC (oligonucleotides 3b and 2a, 816-bp fragment; Fig. 4c, lanes 3 and 4), of pyrB, pyrI, orf1, and pyrC (oligonucleotides 3b and 1a; 1,329-bp fragment, Fig. 4c, lanes 1 and 2) and of pyrC, pyrD, orf2, orf3, and orf4 (oligonucleotides 10b and 4a; 1,600-bp fragment, Fig. 4d, lanes 1 and 2). These various combinations already strongly suggested that transcription initiated at the pyrE promoter (see below) might proceed till the end of orf4. This was further proven by an RT-PCR experiment using the pair of oligonucleotides 10b and 1a, which produced the expected 3,861-bp fragment (Fig. 4e). Previously, members of our group had shown that the Sa-lrp gene is transcribed as a monocistronic messenger that most likely stops at the type I transcriptional stop (5′-TTTTTATT) located 1 nt downstream of the TAA stop codon and that there is only very little readthrough from orf4 into Sa-lrp (13). Thus, the bipolar pyrimidine operon from S. acidocaldarius is transcribed as two polycistronic messengers initiated from the divergent pair of pyrE and pyrB promoters (see below). Transcription initiated at the pyrB promoter does not proceed into Sa-lrp and most likely stops at the type I transcriptional stop signal 5′-TTTTTATT, located 22 nt downstream of the TGA stop codon of orf4 (13). Similar thymine stretches (5′-TTTTTATT and 5′-TTTTTATT) are found 70 and 68 nt downstream of the TAA stop codons of orf5 and orf6, respectively. Such a run of T residues is not present in the region 3′ of the pyrF, pyrE, pyrB, pyrI, orf1, pyrC, pyrD, orf2, and orf3 genes and occurs in four instances only in the 7,636-nt-

---

**TABLE 4. Characteristics of intergenic regions**

<table>
<thead>
<tr>
<th>Location</th>
<th>Length (nt)</th>
<th>Pu/Py ratio</th>
<th>G+C (mol %)</th>
</tr>
</thead>
<tbody>
<tr>
<td>pyrE-pyrF</td>
<td>71</td>
<td>1.03</td>
<td>26.8</td>
</tr>
<tr>
<td>pyrB-pyrF</td>
<td>1</td>
<td>0.93</td>
<td>28.2</td>
</tr>
<tr>
<td>pyrF-orf1</td>
<td>85</td>
<td>1.073</td>
<td>34.1</td>
</tr>
<tr>
<td>orf4-Sa-lrp</td>
<td>60</td>
<td>1.14</td>
<td>18.3</td>
</tr>
<tr>
<td>Sa-lrp-orf5</td>
<td>66</td>
<td>0.784</td>
<td>24.2</td>
</tr>
<tr>
<td>orf5-orf6</td>
<td>133</td>
<td>0.847</td>
<td>27.8</td>
</tr>
<tr>
<td>orf6-orf7</td>
<td>96</td>
<td>1.133</td>
<td>31.4</td>
</tr>
</tbody>
</table>

---

Downloaded from http://jb.asm.org/ on July 9, 2017 by guest
long coding part of the cluster (one in pyrF, two in pyrB, one in orf4; they are far more numerous on the noncoding strand). The presence of these thymine runs at the 3’ end of mono- and polycistronic messengers appears therefore statistically significant. No potential transcriptional stop signal (type I or type II) could be identified at the 3’ end of pyrF. This observation, and the physical overlap of the coding parts of pyrF and truncated orf8, suggest that transcription initiated at pyrE proceeds into...
orf8. In contrast, orf5 and orf6 might produce monocistronic messengers.

Transcription initiation and promoter sequences in the bipolar pyrimidine operon. Start points of transcription were determined by primer extension using total RNA extracted from S. acidocaldarius cells grown in minimal medium supplemented with uracil (50 μg/ml) and minimal medium, respectively, and arrested in the exponential phase. Lanes 7 to 9, primer extension reactions with 100 μg of total RNA from independent cultures grown on complex medium, minimal medium, and minimal medium supplemented with uracil, as indicated. G, A, T, and C, chain terminating DNA sequencing reactions of the non-coding strand obtained with the same 32P-labeled oligonucleotide used to perform the extension reactions. (b) Mapping of the potential transcription initiation site of pyrE by primer extension. The extension reaction was performed with 100 μg of total RNA extracted from cells grown in minimal medium. (c) Mapping of the potential transcription initiation site located within the pyrC gene. The extension reaction was performed with 100 μg of total RNA extracted from cells grown in minimal medium. (d) Nucleotide sequences of the pyrE-pyrB intergenic region and of the potential internal promoter located in the pyrC coding region. Arrows indicate the start site and orientation of transcription. Potential TATA promoter elements are boxed. (d) Comparison of the pyrE-pyrB intergenic regions of S. acidocaldarius, S. solfataricus, and S. tokodaii. Sequences have been aligned with respect to the strictly conserved TATA box elements. The most important direct repeats (black bars) and imperfect palindromic sequences (grey bars) present in the S. acidocaldarius sequence are indicated. Translation initiation codons are in bold capitals (complementary strand for pyrE), and arrows indicate the start sites and orientations of transcription.
ation codon (Fig. 5d). Similarly, pyrE transcription, proceeding in the opposite direction, was found to initiate with an A residue 9 nt upstream of the ATG initiation codon (Fig. 5b). Consequently, the start sites for divergent transcription are separated by a 52-nt-long stretch only (Fig. 5d). Its G+C content of 21.2% (26.8% for the extended 71-nt-long noncoding region) is considerably lower than that of the complete cluster (35.3%) but comparable to the mean value of 25% for intergenic regions in S. solfataricus (50). No transcription start could be identified in front of orf1, pyrD, and orf4 (not shown), but a cDNA product corresponding to a potential initiation at an A residue within the pyrC coding part was detected (Fig. 5c and d). The start site of the monocistronic Sa-lrp gene was determined previously (7, 13); transcription initiation of orf5, orf6, and of the truncated orf7 and orf8 has not been investigated. Three of the identified transcripts (pyrE, pyrB, and Sa-lrp) initiate with a purine residue that is preceded by a thymine, thus creating the pyrimidine-purine dinucleotide required for correct start site selection of archaeal promoters (25). A good match to the consensus TATA box of archaeal promoters is the 5′-TAAAATA sequence, ideally centered around −26.5 nt upstream of the pyrB start site (Fig. 5d). This sequence is preceded (positions −36 to −31) by a sequence stretch, 5′-AAGAC, that shows good sequence conservation with the transcription factor B responsive element sequence of archaeal promoters, a purine-rich stretch with a consensus RNWAAW (R = purine; W = A or T) (2). The potential TATA box element of the pyrE promoter (5′-TATTAAA), comprising between −23 and −30 upstream of the pyrE transcription start, largely overlaps the TATA box element of the pyrB promoter (Fig. 5d). Therefore, this region is apparently shared by the divergent pair of pyrE and pyrB promoters. The hypothetical TATA box element (5′-TAAAGGAA) comprising between −23 and −30 nt upstream of the potential pyrC transcription start showed only weak sequence identity with the consensus (Fig. 5d). Therefore, the detected cDNA product might correspond to a discrete degradation product of the longer mRNA molecule initiated at pyrB, or it could result from a premature arrest of the reverse transcriptase. The lack of a stable hairpin structure in the region makes the latter unlikely, however.

**Regulation of pyrB transcription initiation.** To determine the effect of nutrient composition on the abundance of pyrB transcription, we performed primer extension experiments with total RNA extracted from S. acidocaldarius cells grown on complex medium or on minimal medium either devoid of or supplemented with uracil. The densitometric analysis of the autoradiograph (Fig. 5a) indicated that the pyrB transcript is about 2.0-fold more abundant in the absence than in the presence of uracil and about 2.5-fold more abundant on minimal medium than on complex medium (mean for two assays performed with independent RNA preparations). Similarly, an approximately twofold difference was also observed in the ATCase specific activities measured in cell extracts of S. acidocaldarius cells grown in the presence or absence of uracil. Therefore, expression of the S. acidocaldarius pyrimidine biosynthetic genes is regulated, at least in part, at the transcriptional level.

**DISCUSSION**

In this study we present the sequence determination and analysis of an 8,519-bp segment of the S. acidocaldarius genome carrying an extremely tightly packed divergent pyrimidine gene cluster that exhibits a novel organization. Primer extension and RT-PCR experiments indicate that the pyrimidine genes are transcribed as two polycistronic messengers, pyrBI-orf1-pyrCD-orf2-orf3-orf4 and pyrEF. The latter most likely also includes orf8, as suggested by the physical overlap of the pyrF and hypothetical orf8 cistrons and the lack of an identifiable transcription stop signal at the 3′ end of pyrF. A similar clustering of pyr genes occurs in the entirely sequenced genomes of S. solfataricus P2 (52) and S. tokodaii strain 7 (32) and appears therefore characteristic of Sulfolobus. Beyond the orf5 and pyrF genes, however, the similarity in genome organization vanishes. Also novel is the gene order orf1-pyrC-pyrD, in which pyrC is intercalated between the electron transfer and catalytic subunit of DHOdehase; in bacteria, pyrK (pyrDII) and pyrD are generally contiguous.

The seven cistrons involved in de novo synthesis of UMP are tightly packed, and only pyrB and pyrI, encoding the catalytic and regulatory subunits of ATCase, respectively, show no overlap at either the 5′ or 3′ end; pyrB and pyrI are separated by a single nucleotide, and pyrD and orf1 are separated by 85 nt. The start points of the divergent transcripts initiated at the pyrE and pyrB promoters are separated by 52 nt only, and the corresponding initiation codons are separated by 71 nt. This short intergenic region appears to contain a pair of overlapping divergent promoters showing a good match to the consensus TATA and transcription factor B responsive element promoter elements. This peculiar situation constitutes an interesting model system for the study of basal and regulated transcription in hyperthermophilic archaea, still a poorly documented area of investigation. Indeed, the intimate intertwining of divergent promoters creates an interesting test case for the analysis of sequence determinants imposing transcription polarity, for the study of mutual influences of nearby promoters on binding of transcription factors (TBP, TFB, and TFE) and polymerase recruitment, and for the analysis of variations in template topology generated by the moving polymerases engaged in divergent transcription (positive supercoiling in front, negative ahead) on transcription initiation frequencies. This last aspect is particularly interesting, since DNA in hyperthermophiles is known to be relaxed to slightly positively supercoiled, in contrast to the negative supercoiling of mesophilic genomes.

We have shown that pyr gene expression in S. acidocaldarius is regulated at least in part at the transcriptional level. pyrB mRNA levels were found to differ by a factor of 2 in minimal medium devoid of or supplemented with uracil. On the other hand, ATCase enzyme activities measured in cultures of various pyrimidine auxotrophs grown on limiting amounts of uracil showed a 2- to 10-fold derepression (23). Taken together, these data suggest the existence of a several-fold regulation of transcription initiation that would provide S. acidocaldarius with the flexibility required to attune pyrimidine biosynthesis to the cellular needs and fluctuations in available nutrients. The mode of pyrimidine-specific repression of pyrB transcrip-
tion initiation in *S. acidocaldarius* is not known, but the mecha-
anism appears to be different from the various types of attenuation-like mechanisms operative in mesophilic and thermophilic bacteria. The pyrE and pyrB leader sequences are too short to contain attenuator-antia attenuation sites. In contrast, the short pyrE-pyrB intergenic region exhibits features (direct repeats and imperfect palindromic sequences) charac-
teristic of potential binding sites for a hypothetical transcriptional regulator(s), which are at least in part conserved in *S. solfataricus* and *S. tokodaii* (Fig. 5e).

The quasi-leaderless pyrB and pyrE messengers bear no SD site, whereas in contrast, the ORFs embedded in the polys
tronic messengers show a rather extensive complementarity to the 3′ end of 16S rRNA (Table 3). Interestingly, it was previously
shown that the disruption of the SD sequence abolished translation, but most importantly, the effect of this disruption could be suppressed by deleting the 5′-untranslated region, thus creating a leaderless messenger (9). In all organisms, bacteria, archaea, and eucarya, leaderless messengers exist, even for abundant proteins (28). Translation initiation of the archaeal leaderless mRNAs might require the formation of a complex equivalent to the bacterial 30S-Met-tRNA-MetIF2 intermediate implicated in translation of such mRNAs in *E. coli*, as is suggested by the faithful in vitro translation of leader-
less λ cfl mRNA with an archaeal translation system (21).

The bipolar pyr operon of *S. acidocaldarius* does not harbor the genes encoding carbamoylphosphate synthetase. Studies with mutants indicate that *S. acidocaldarius* has only one car-amoylphosphate synthetase (23), and on the opposite flank of the *S. solfataricus* P2 genome the adjacent carA and carB genes encoding the sole CPSase of this organism are embedded in an arginine gene cluster (52). This peculiar situation raises interesting questions regarding the regulation of CPSase activity in relation to the cellular needs for arginine and pyrimidine biosynthesis. A com-
parable situation exists in *B. caldolyticus* and the thermophile

Charlier, D., M. Roovers, T.-L. Thia-Toong, V. Durbecq, and N. Glansdorff. 1997. Cloning and identification of the *Sulfolobus solfataricus* pyrK gene en-


Condó, I., A. Ciammarucioni, D. Bellini, D. Ruggiero, and P. Londé. 1991. Cis-acting signals controlling translation initiation in the thermophilic ar-

Dagert, M., and S. D. Ehrlich. 1979. Prolonged incubation in calcium chlo-
ride improves the competence of *E. coli* cells. Gene 6:23–28.

Durbecq, V., T.-L. Thia-Toong, D. Charlier, V. Villeret, M. Roovers, R. Wattiez, C. Le Grain, and N. Glansdorff. 1999. Aspartate carbamoyltrans-


Enoro-Glansdorff, J., D. Gigot, T.-L. Thia-Toong, N. Glansdorff, and D. Charlier. 2000. Purification and characterization of Sa-Lrp, a DNA-binding protein from the extreme thermophilic archaeon *Sulfolobus acidocaldarius* homolo-

history for a complex enzyme. Paths Pyrimidines 7:49–63.

Fitz-Gibbon, S. T., H. Ladner, U.-J. Kim, K. O. Stetter, M. I. Simons, and R. L. Switzer. 1994. The pyrimidine biosynthesis operon of *Bacillus caldolyticus* reveals a number of unique traits and the univer-


