Molecular Characterization of the α-Glucosidase Gene (malA) from the Hyperthermophilic Archaeon Sulfolobus solfataricus

MICHAEL ROLFSMEIER, CYNTHIA HASELTINE, ELISABETTA BINI, AMY CLARK, AND PAUL BLUM*

George Beadle Center for Genetics, School of Biological Sciences, University of Nebraska, Lincoln, Nebraska 68588-0666

Received 18 September 1997/Accepted 12 December 1997

Acidic hot springs are colonized by a diversity of hyperthermophilic organisms requiring extremes of temperature and pH for growth. To clarify how carbohydrates are consumed in such locations, the structural gene (malA) encoding the major soluble α-glucosidase (maltase) and flanking sequences from Sulfolobus solfataricus were cloned and characterized. This is the first report of an α-glucosidase gene from the archaeal domain. malA is 2,083 bp and encodes a protein of 693 amino acids with a calculated mass of 80.5 kDa. It is flanked on the 5′ side by an unusual 1-kb intergenic region. Northern blot analysis of the malA region identified transcripts for malA and an upstream open reading frame located 5′ to the 1-kb intergenic region. The malA transcription start site was located by primer extension analysis to a guanine residue 8 bp 5′ of the malA start codon. Gel mobility shift analysis of the malA promoter region suggests that sequences 3′ to position −33, including a consensus archaeal TATA box, play an essential role in malA expression. malA homologs were detected by Southern blot analysis in other Sulfolobus strains and in Sulfolobus shibatae, while no homologs were evident in Sulfolobus acidocaldarius, lending further support to the proposed revision of the genus Sulfolobus. Phylogenetic analyses indicate that the closest Sulfolobus α-glucosidase homologs are of mammalian origin. Characterization of the recombinant enzyme purified from Escherichia coli revealed differences from the natural enzyme in thermostability and electrophoretic behavior. Glycogen is a substrate for the recombinant enzyme. Unlike maltose hydrolysis, glycogen hydrolysis is optimal at the intracellular pH of the organism. These results indicate a unique role for the Sulfolobus α-glucosidase in carbohydrate metabolism.

Microbes which are native to boiling acid hot springs include five genera assigned by 16S rRNA sequencing, G+C mole percent composition, metabolic characteristics, and protein sequence conservation studies to the order Sulfolobales (8). The order Sulfolobales is placed within the crenarchaeotal subdivision of the archaea (49). Sulfolobus is the largest genus in this order, comprising at least six species. These include S. acidocaldarius (5), S. solfataricus (13), S. shibatae (15, 52), S. metallicus (23), S. icelandicus (54), and S. hakonensis (44), all of which are obligate aerobes and either facultative chemoheterotrophs or strict lithoautotrophs. Most physiological and biochemical studies, however, focus on only three species, S. solfataricus, S. acidocaldarius, and S. shibatae.

S. solfataricus exhibits diverse modes of metabolism in batch culture at temperatures ranging between 70 and 90°C. It grows lithoautotrophically by oxidizing sulfur (5, 25, 30) and chemoheterotrophically on reduced-carbon compounds (13, 16, 19). Despite this metabolic flexibility, the utilization of reduced-carbon nutrients such as plant-derived starch or cellulose is poorly understood. Input of such carbon is typically rare in acidic hot springs and depends upon external factors such as fire or wind. Since polysaccharide hydrolysis and sugar caramlization are active processes in hot acid environments, successful competition for carbohydrates should necessitate mechanisms for rapid assimilation. Endogenous reserves of starch in plants and exogenous starch utilization in microbes often depend upon an α-amylase which generates linear maltodextrins as well as an α-glucosidase (maltase) which converts maltose and maltodextrins to glucose (26). In animals, however, α-glucosidases are also critical for utilization of intracellular stores of glycogen. For example, glycogen storage disease (Pompe’s disease) in humans is a direct consequence of α-glucosidase deficiency (21). Many α-glucosidase genes from eukaryotic and eubacterial organisms have been cloned and characterized (20). Although such enzymes also occur among the archaea (10, 40), none of their corresponding genes have yet been characterized, precluding an analysis of their intracellular functions or evolutionary origins.

Members of the genus Sulfolobus utilize starch as the sole carbon and energy source (16) and have both α-amylase and α-glucosidase activities (4). It has been reported, however, in surveys of Sulfolobus species that carbohydrate utilization profiles are distinct (16). For example, S. solfataricus contains a β-glucosidase (17, 34) which appears to be largely absent in S. acidocaldarius (17). We recently reported the purification and characterization of the major soluble α-glucosidase (maltase) and the secreted α-amylase from S. solfataricus (19, 40). To further explore archaeal mechanisms for carbohydrate utilization and to examine their relationship with those of eukaryotes and eubacteria, the α-glucosidase gene from S. solfataricus was cloned and characterized, and its distribution and associated activity in common Sulfolobus cultivars were examined.

MATERIALS AND METHODS

Strains and cultivation. The identity of S. solfataricus 98/2 (22, 40) was confirmed by DNA sequence analysis of a cloned PCR fragment spanning residues 99 to 626 of the 16S rRNA gene. The GenBank accession numbers for this sequence and that determined for S. solfataricus P2 (DSM 1617) are L30990 and L30991, respectively. Comparison of the resulting sequences with previously published citations for S. shibatae (GenBank accession no. M32504) and S. acidocaldarius (30, 35, 53) (GenBank accession no. X03235) confirmed the identity

* Corresponding author. Mailing address: School of Biological Sciences, E234, Beadle Center for Genetics, University of Nebraska, Lincoln, NE 68588-0666. Phone: (402) 472-2769. Fax: (402) 472-8722. E-mail: pblum@crcvms.unl.edu.
of strain 98/2 as *S. solfataricus*. Cells were cultured at 80°C in a minimal salts medium (1), modified as described previously (5), at pH 3.0 with various carbon hydrates at 0.2% (wt/vol) as sole carbon and energy sources. All manipulations of *Escherichia coli* were as described previously (38).

**Methods.** Restriction digestion and ligation of DNA were performed as described previously (3). Plasmid transformation was performed with DH5α cells as described previously (18). Isolation of plasmid DNA was performed by the alkali lysis procedure (2). DNA sequence analysis was as described previously (39), and DNA alignment and analysis were performed with the fragment assembly programs of the Wisconsin Genetics Computer Group software package, version 5.1.

**RESULTS**

### Generation times and α-glucosidase activities of *Sulfobolus* species during growth on different sole carbon and energy sources

<table>
<thead>
<tr>
<th>Organism</th>
<th>Glucose</th>
<th>Maltose</th>
<th>Starch</th>
<th>Glucose</th>
<th>Maltose</th>
<th>Starch</th>
</tr>
</thead>
<tbody>
<tr>
<td><em>S. solfataricus</em> 98/2</td>
<td>8</td>
<td>8</td>
<td>10</td>
<td>130 ± 3</td>
<td>221 ± 9</td>
<td>230 ± 26</td>
</tr>
<tr>
<td><em>S. shibatae</em></td>
<td>12</td>
<td>9</td>
<td>14</td>
<td>29 ± 2</td>
<td>181 ± 21</td>
<td>228 ± 35</td>
</tr>
<tr>
<td><em>S. acidocaldarius</em></td>
<td>23</td>
<td>ND</td>
<td>ND</td>
<td>&lt;0.1 ND</td>
<td>12 ± 2</td>
<td></td>
</tr>
</tbody>
</table>

* Variations between replicate samples are indicated.
* NG, no growth.
* ND, not determined.

### Genome/Derivatives

The *malA* coding region has been deposited in GenBank (accession AF042494).

### Gel mobility shift assays

A Gel mobility shift assay was performed generally as described previously (7). The labeled probe was generated by the incorporation of [γ-32P]ATP into the 5′ overhanging end of a purified restriction fragment. Cell extracts used as a source of DNA binding proteins were prepared as described previously (24). Briefly, mid-exponential-phase cultures of *S. solfataricus* 98/2 were harvested by centrifugation, resuspended, and lysed by addition of Triton X-100. The resulting lysate was clarified by centrifugation, and 1 μl of the clarified preparation was added to a final concentration of 20% (vol/vol) and the extract was rapidly frozen and stored at −80°C. Labeled probe with or without added unlabeled competitor DNAs was incubated with cell extract at 70°C for 30 min, and the samples were then analyzed by nondenaturing PAGE. The resulting gel was dried and used to prepare autoradiographs.

### Nucleotide sequence accession number

The sequence resulting from analysis of regions lying 3′ to the *malA* coding region has been deposited in GenBank under accession no. AF042494.
α-Glucosidase activities in crude cell extracts of S. solfataricus and S. shibatae were typically 10-fold greater than those seen with S. acidocaldarius during growth on all carbon sources (Table 1). Activities varied nearly twofold for S. solfataricus (comparing all carbon sources), while in S. shibatae, α-glucosidase activity was six- to sevenfold higher during growth on maltose or starch than during growth on glucose. α-Glucosidase in S. acidocaldarius was undetectable during growth on glucose and was 28-fold lower during growth on starch relative to that of either of the other two species.

Cloning and characterization of the S. solfataricus malA gene. To further investigate the apparent differences in polysaccharide utilization of maltose by these Sulfolobus species, the gene encoding the α-glucosidase from S. solfataricus was cloned and characterized. Cloning of the gene was accomplished by using gene-specific PCR primers derived from the S. solfataricus α-glucosidase protein sequence. Amino acid sequencing indicated that the mature N-terminal maltase sequence was MQTIKIYENLGVYLWIGEP. Since the purified S. solfataricus α-glucosidase is generally resistant to proteolytic degradation (40), protein fragments for internal N-terminal sequence analysis were generated by chemical cleavage with cyanogen bromide. A fragment of 19 kDa was selected and yielded the N-terminal sequence VGKYLLYAPI. The resulting amino acid sequence information was used to design degenerate oligonucleotides which were then used to amplify a DNA fragment of 1.6 kb by PCR. A 731-bp HindIII-EcoRV fragment derived from the resulting 1.6-kb PCR product was used to generate a radiolabeled probe for Southern hybridization to verify the origin of the amplification product. This probe cross-hybridized with single DNA fragments of 1.2, 1.7, and 1.4 kb in HindIII, XbaI, and HindIII genomic digests, respectively, of S. solfataricus DNA. The probe was used to screen a genomic S. solfataricus phage l library consisting of 672 individually propagated recombinant phages by Southern blot analysis. A single isolate was identified (l-7F7), which contained a 15.1-kb insert of S. solfataricus DNA. Southern blot analysis of restriction digests of the l-7F7 phage was performed with the PCR-derived probe. Cross-hybridizing restriction fragments which were identical in size to those observed previously with genomic DNA were observed and indicated that the α-glucosidase-coding region is contained in a 4.3-kb BamHI fragment of the l-7F7 insert. This 4.3-kb BamHI fragment was subcloned and sequenced. The sequence located immediately 5' to the malA coding region was subcloned from α-7F7 as a 2.3-kb SacI-HindIII fragment. A 344-bp SacI-BamHI fragment from the extreme 5' end of this fragment was used as a Southern blot probe to identify the next upstream overlapping l clone from the genomic S. solfataricus 98/2 library. This isolate was named l-1H4. A 4.1-kb HindIII-SacI

FIG. 1. DNA sequence analysis of malA. Glycosyl hydrolase and ATP/GTP binding motifs are indicated by double and single underlining, respectively. Putative promoter and termination sequences are in boldface. The start of transcription is indicated as +1.
fragment which cross-hybridized to the same SacI-BamHI 344-bp fragment was subcloned from \( \lambda \)-1H4. A 2-kb segment of this 4.1-kb HindIII-SacI fragment was sequenced to complete the analysis of putative genes lying \( 5' \) to \( malA \). Analysis of regions lying \( 3' \) to the \( malA \) coding region was done with a 1.8-kb \( 3' \) overlapping HindIII-HindIII fragment derived from \( \lambda \)-7F7. The resulting sequence, comprising a nearly 7-kb DNA contig, has been deposited in GenBank (see Materials and Methods).

The \( \alpha \)-glucosidase open reading frame (ORF) (\( malA \)) was identified by comparison of peptide sequences derived from the N-terminal and internal N-terminal sequencing of the natural protein to the deduced amino acid sequence (Fig. 1). The \( malA \) sequence comprises 2,083 bp encoding a protein of 693 amino acids with a predicted mass of 80.5 kDa. This closely corresponds to the apparent mass of the previously purified enzyme subunit (40). Sequence analysis of the deduced \( malA \) product identified a glycosyl hydrolase motif at residues 316 to 323 and an ATP/GTP binding site motif (P loop) at residues 583 to 590 (single underline). The glycosyl hydrolase motif contains the putative active-site asparagine previously identified for the human \( \alpha \)-glucosidase gene (20). Only two cysteine residues are evident, consistent with the low cysteine content seen previously in thermophilic proteins. There are 14 methionine residues, and the predicted mass of the largest sequence uninterrupted by methionines is 19 kDa, as suggested by the cyanogen bromide cleavage pattern of the \( \alpha \)-glucosidase. The codon composition favors adenosine or thymidine in the wobble position, as expected for the 38 mol\% G+C \( malA \) coding sequence. Significant bias was evident for arginine; this amino acid is coded for twice by CG(A/T/C/G) codons and 31 times by AG(G/A) codons. However, contrary to the apparent low G+C content of the genome of this organism, four amino acids (asparagine, tyrosine, phenylalanine, and histidine) which can be coded for by codons with either a C or a T in the third position show no bias towards T. Additional examination of the 7.05-kb contig identified several ORFs (Fig. 2A) with G+C contents of 37 to 38 mol\%, as expected from previous analysis (13, 15). No sequence homologs of these ORFs were evident in searches of sequence databases. An unusual intergenic region of nearly 1 kb located immediately \( 5' \) to \( malA \) was identified in this contig. It exhibits a G+C content of 30.8 mol\%, a value considerably lower than that for the flanking coding regions.

**Northern blot analysis of the \( malA \) region.** Northern blot analysis was performed to evaluate the expression of \( malA \) and its surrounding regions during growth of \( S. solfataricus \) on maltose as the sole carbon and energy source. Probes P1 to P4 were dsDNA probes, and probes P5 and P6 were RNA probes (riboprobes). A probe derived from a 682-bp \( ClaI-HincII \) fragment from the \( 3' \) end of the \( malA \) coding region (Fig. 2A, probe P3) cross-hybridized to a single transcript of approximately 2.4 kb, indicating that \( malA \) is expressed during growth.
on maltose (Fig. 2C, lane 2). Probes derived from regions either 5' or 3' to the malA coding region were used to assess gene expression of immediately flanking sequences. These included an EcoRV-EcoRI fragment of 499 bp located 150 bp 5' to malA (Fig. 2A, probe P2) and an EcoRI-BamHI fragment of 231 bp located 192 bp 3' to malA (Fig. 2A, probe P4). No cross-hybridization was evident with either of these probes (Fig. 2C, lanes 1 and 3). An additional probe (Fig. 2A, probe P1) was used to examine expression of ORF1, located approximately 1 kb 5' of malA. This probe was derived from a SacI-BamHI fragment located in the center of ORF1. An approximately 2.4-kb transcript was detected with this probe (Fig. 2B, lane 1), while again no transcript was evident with probe P2 (Fig. 2B, lane 2). Riboprobes (Fig. 2A, probes P5 and P6) were used to determine the direction of transcription of malA. The 2.4-kb transcript was evident with the antisense riboprobe P6 (Fig. 2D, lane 2), while no transcript was detected with the sense riboprobe P5 (Fig. 2D, lane 1). These results indicate that the malA gene is transcribed from a site immediately adjacent to the gene and away from the large noncoding intergenic region.

Characterization of the malA regulatory region. There is a potential archaean promoter sequence located 32 bp 5' to the start codon of the malA gene (Fig. 1). The putative promoter (TTTATA) closely matches the consensus promoter sequence for Sulfolobus (37). A box B motif (TGA) (37) is also evident 7 bp 5' to the start codon. Primer extension analysis indicated that malA transcription initiates on the guanine of the putative box B motif (Fig. 3). The mapped start site is 8 bp 5' to the start codon. Although there is a potential ribosome binding site spanning positions −3 to +3, which are complementary to the six 3'-terminal bases of the 16S rRNA of S. solfataricus (30, 35, 53), this sequence overlaps in part the site of malA transcription initiation. The utilization of this sequence for the initiation of translation is therefore unclear. The malA mRNA is only slightly larger than the coding region of the gene (Fig. 2). Since transcription initiates very close to the start of the coding region, termination of transcription of the gene must occur close to the end of malA. The near-consensus terminator sequence (TTTTTCA) (11) located immediately 3' to the stop codon of malA may play a role in this process.

The interaction between purified archaean TATA binding proteins and archaean promoters can be characterized by gel shift analysis (36). Crude cell extracts prepared as described previously for use in an in vitro transcription system (24) were used as sources of DNA binding proteins. The probe was a 233-bp EcoRI fragment which starts 151 bp 5' to the malA transcription start site and extends 80 bp into the malA trans-
in the analysis, strain 98/2 from Yellowstone National Park and strain P2 (DSM 1617) from Italy. Genomic digests prepared with EcoRV (Fig. 5A) or HindIII (Fig. 5B) were then probed under stringent hybridization conditions with a 731-bp EcoRV-HindIII malA gene fragment encompassing nucleotides 714 to 1445 of the malA coding region (Fig. 1). Both strains of *S. solfataricus* exhibited strongly hybridizing bands of 2.9 kb following EcoRV digestion and 1.2 kb following HindIII digestion (Fig. 5, lanes 3 and 4), in agreement with the Southern blot results obtained previously with the 731-bp HindIII-EcoRV probe fragment derived from the initial malA PCR product. For *S. shibatae*, single weakly hybridizing bands of 0.65 kb following EcoRV digestion (Fig. 5A, lane 1) and 3.7 kb following HindIII digestion (Fig. 5B, lane 1) were also observed. No cross-hybridization was observed, however, between the *S. solfataricus* malA gene and *S. acidocaldarius* genomic DNA digests (Fig. 5, lanes 2).

**Phylogenetic analysis of α-glucosidase sequences.** Amino acid sequences of α-glucosidases and the related sucrase isomaltases were retrieved from the Swiss-Prot and EMBL/GenBank/DDBJ databases. A multiple sequence alignment of 6 bacterial and 11 eukaryotic sequences in addition to the *S. solfataricus* sequence was made. The region of the *S. solfataricus* α-glucosidase used for the alignment included 569 amino acid residues spanning positions 50 to 618. The *S. solfataricus* α-glucosidase is the only representative of the archaea, since no other archaeal α-glucosidases were found in the databases. A conserved stretch of amino acids located in the middle of the three fungal sequences was deleted to minimize sequence gaps in the alignment. The alignment of sequences then was analyzed by distance, parsimony, and maximum-likelihood methods. The *E. coli* malZ gene product was used as the outgroup. The sequences clustered into two groups typically of either eubacterial or eukaryotic affiliation by all three methods of analysis. Nearest-neighbor distance analysis and parsimony analysis indicate that the *S. solfataricus* α-glucosidase is most closely related to mammalian enzyme homologs (Fig. 6). Maximum-likelihood analysis gave similar results (data not shown).

**Recombinant *S. solfataricus* α-glucosidase activity.** To prove that malA encodes a hyperthermophilic α-glucosidase, the
The recombinant α-glucosidase exhibited greater thermostability than the natural enzyme, with a half-life of 39 h at 85°C at a pH of 6.0.

α-Glucosidases of mammalian origin can be generally distinguished from those of higher plants and eubacteria by their affinities for glycogen as a substrate. Glycogen was hydrolyzed efficiently by the S. solfataricus enzyme. It exhibited a pH optimum for glycogen hydrolysis of 5.5 (Fig. 7B), a $K_m$ of 64.9 mg/ml, and a $V_{max}$ of 1.0 μmol of glucose/min at 85°C.

**DISCUSSION**

We report here the identification and characterization of the gene (malA) encoding the major soluble α-glucosidase (maltase) of S. solfataricus. This is the first report of an α-glucosidase sequence from the archaeal domain. The presence of an S. solfataricus malA homolog and corresponding α-glucosidase activity in S. shibatae suggests that these Sulfolobus species have similar pathways for the utilization of maltose and maltodextrins. Lack of a malA homolog or significant α-glucosidase activity in S. acidocaldarius may explain the inability of S. acidocaldarius to utilize maltose as a sole carbon and energy source. An α-glucosidase thus may be essential for utilization of maltose among certain members of the genus Sulfolobus and represents a distinguishing physiological feature for Sulfolobus species identification. Such metabolic divergence lends further support to the suggestion that the Sulfolobus genus be revised (8).

Maltose utilization by these Sulfolobus species necessitates mechanisms for assimilation of maltose or maltodextrins, and specific transport systems have been identified recently in S. shibatae (51). However, the purified S. solfataricus enzyme also uses glycogen as a substrate. S. solfataricus accumulates glycogen as the major intracellular storage polysaccharide (28); thus, glycogen utilization may require the S. solfataricus α-glucosidase. This is further supported by the observation that unlike maltose hydrolysis (40), glycogen hydrolysis by the S. solfataricus α-glucosidase exhibits a more neutral pH optimum approximating that of the intracellular environment of this organism (31).

Perhaps a dual role for the α-glucosidase in the utilization of endogenous and exogenous polysaccharides can explain the apparent α-glucosidase activity observed in both S. solfataricus and S. shibatae during growth on glucose. Constitutive expression of malA may be necessary to balance catabolic and anabolic metabolic needs. Since eubacterial α-glucosidases lack glycogen-hydrolytic activity (33, 47, 48), the results presented here further distinguish archaeal α-glucosidases from those of eubacteria.

The large intergenic sequence located 5′ to malA is a distinguishing feature of the malA region. Northern blot analysis indicates that there is a lack of apparent transcripts encoded on either strand in the region covered by the dsDNA probe P2 (bp −651 to −152) produced during chemoheterotrophic growth on maltose. The entire intergenic region is also largely devoid of sequences encoding proteins; there is only one deduced sequence containing putative coding regions for malA. The entire intergenic region is also largely devoid of sequences encoding proteins; there is only one deduced sequence containing putative coding regions for malA. The results presented here further distinguish archaeal α-glucosidases from those of eubacteria.

The large intergenic sequence located 5′ to malA is a distinguishing feature of the malA region. Northern blot analysis indicates that there is a lack of apparent transcripts encoded on either strand in the region covered by the dsDNA probe P2 (bp −651 to −152) produced during chemoheterotrophic growth on maltose. The entire intergenic region is also largely devoid of sequences encoding proteins; there is only one deduced sequence containing putative coding regions for malA. The entire intergenic region is also largely devoid of sequences encoding proteins; there is only one deduced sequence containing putative coding regions for malA. The results presented here further distinguish archaeal α-glucosidases from those of eubacteria.

The large intergenic sequence located 5′ to malA is a distinguishing feature of the malA region. Northern blot analysis indicates that there is a lack of apparent transcripts encoded on either strand in the region covered by the dsDNA probe P2 (bp −651 to −152) produced during chemoheterotrophic growth on maltose. The entire intergenic region is also largely devoid of sequences encoding proteins; there is only one deduced sequence containing putative coding regions for malA. The entire intergenic region is also largely devoid of sequences encoding proteins; there is only one deduced sequence containing putative coding regions for malA. The results presented here further distinguish archaeal α-glucosidases from those of eubacteria.

The large intergenic sequence located 5′ to malA is a distinguishing feature of the malA region. Northern blot analysis indicates that there is a lack of apparent transcripts encoded on either strand in the region covered by the dsDNA probe P2 (bp −651 to −152) produced during chemoheterotrophic growth on maltose. The entire intergenic region is also largely devoid of sequences encoding proteins; there is only one deduced sequence containing putative coding regions for malA. The entire intergenic region is also largely devoid of sequences encoding proteins; there is only one deduced sequence containing putative coding regions for malA. The results presented here further distinguish archaeal α-glucosidases from those of eubacteria.

The large intergenic sequence located 5′ to malA is a distinguishing feature of the malA region. Northern blot analysis indicates that there is a lack of apparent transcripts encoded on either strand in the region covered by the dsDNA probe P2 (bp −651 to −152) produced during chemoheterotrophic growth on maltose. The entire intergenic region is also largely devoid of sequences encoding proteins; there is only one deduced sequence containing putative coding regions for malA. The entire intergenic region is also largely devoid of sequences encoding proteins; there is only one deduced sequence containing putative coding regions for malA. The results presented here further distinguish archaeal α-glucosidases from those of eubacteria.
mol% (13, 15). However, the approximately 1-kb intergenic region lying 5' to malA is distinctly lower in its G+C composition, with an average of 30 mol%, in contrast to the two flanking coding regions, which exhibit values of 38 mol%. It must therefore be assumed that intergenic regions such as this are rare in the _S. solfataricus_ genome, as they would otherwise reduce global measurements of base composition.

Transcription of malA appears to utilize a consensus archaenal promoter motif. This sequence has been previously described as the box A motif TTTATA (11, 12). Gel shift analysis directly supports a role for the TATA box A in the 5'-flanking region for malA. Complex formation was dependent upon the 3'-terminal 42 bp, of which the most 5'-terminal region comprises the TATA box A sequence. Additionally, primer extension analysis indicates that the point of transcription initiation is at the conserved guanine located within the so-called box B region. As this residue lies only 8 bp 5' to the malA start codon, the mechanism employed for transcription initiation of malA must operate within significant sequence constraints. Similar observations have been made for other archaenal genes, and as yet the mechanisms employed for transcription initiation of these types of genes remain obscure.

Phylogenetic analysis of the β-glucosidase and surcosase isoforms yielded similar trees with nearly identical branching topologies. These methods place the _S. solfataricus_ enzyme with those of eukaryotes, specifically mammals, rather than with euarchaenal orthologs. Recent studies on the relatedness of archaea to eubacteria and eukaryotes have suggested that archaeal central metabolic enzymes exhibit greater relatedness to those of eubacteria. The results presented here, however, indicate that at least some archaenal metabolic pathways, such as those associated with carbohydrate metabolism, may have an evolutionary origin more common in those of certain eukaryotes.

**ACKNOWLEDGMENTS**

We thank D. Grogan for cultures. This work was supported by NSF grants EPS-9255225 and MCB-9604000 and by the University of Nebraska Center for Biotechnology.

**REFERENCES**


16. S. solfataricus_ genome, as they would otherwise reduce global measurements of base composition.

17. S. solfataricus_ genome, as they would otherwise reduce global measurements of base composition.


