The phosphorylation of glucose by different sugar kinases plays an essential role in Archaea because of the absence of a phosphoenolpyruvate-dependent transferase system characteristic for Bacteria. In the genome of the hyperthermophilic Archaeon Thermoproteus tenax a gene was identified with sequence similarity to glucokinases of the so-called ROK family (repressor protein, open reading frame, sugar kinase). The T. tenax enzyme, like the recently described ATP-dependent "glucokinase" from Aeropyrum pernix, shows the typical broad substrate specificity of hexokinases catalyzing not only phosphorylation of glucose but also of other hexoses such as fructose, mannose, or 2-deoxyglucose, and thus both enzymes represent true hexokinases. The T. tenax hexokinase shows strikingly low if at all any regulatory properties and thus fulfills no important control function at the beginning of the variant of the Embden-Meyerhof-Parnas pathway in T. tenax. Transcript analyses reveal that the \textit{hxk} gene of \textit{T. tenax} is cotranscribed with an upstream located \textit{orfX}, which codes for an 11-kDa protein of unknown function. Growth-dependent studies and promoter analyses suggest that post-transcriptional RNA processing might be involved in the generation of the monocistronic \textit{hxk} message, which is observed only under heterotrophic growth conditions. Data base searches revealed \textit{T. tenax} hexokinase homologs in some archaea, few eukaryal, and many bacterial genomes. Phylogenetic analyses confirm that the archaeal hexokinase is a member of the so-called ROK family, which, however, should be referred to as ROK group because it represents a group within the bacterial glucokinase fructokinase subfamily II of the hexokinase family. Thus, archaeal hexokinases represent a second major group of glucose-phosphorylating enzymes in Archaea beside the recently described archaeal ADP-dependent glucokinases, which were recognized as members of the ribokinase family. The distribution of the two types of sugar kinases, differing in their cosubstrate as well as substrate specificity, within Archaea is discussed on the basis of physiological constraints of the respective organisms.

Phosphorylation of sugars is of high significance for the carbohydrate metabolism of the cell, particularly for preparing carbohydrates for degradation and transfer reactions but also, as known for several bacteria, as a means of uptake. Among various enzymes engaged in these reactions, especially nucleotide-dependent sugar kinases with various specificities, are important for sugar phosphorylation. In the domain of Archaea, these enzymes play even a more central role for the carbohydrate metabolism than in Bacteria because Archaea in general do not seem to possess a phosphoenolpyruvate-dependent transferase system, which delivers the cell with sugar phosphates.

To the best of our knowledge the only characterized nucleotide-dependent sugar kinases in Archaea described so far are enzymes that phosphorylate glucose (1–4), fructose 6-phosphate (3, 5–11), fructose 1-phosphate (12), fructose 1-phosphate (12), fructose (13), and galactose (14). Surprisingly, a striking variety of glucose-phosphorylating enzymes occurs in Archaea: (i) the ADP-dependent glucokinases of \textit{Pyrococcus furiosus} (1, 2, 15) and \textit{Thermococcus litoralis} (2); (ii) the bifunctional ADP-dependent glucokinase/phosphofructokinase of \textit{Methanococcus jannaschii} (3, 11); and (iii) the ATP-dependent "glucokinase" of \textit{Aeropyrum pernix} (4), which is, however, because of its rather broad substrate specificity, in reality a hexokinase (ATP: hexose-6-phosphotransferase, EC 2.7.1.1). True ATP-dependent glucokinases (ATP:d-glucose 6-phosphotransferase, EC 2.7.1.2), which differ from hexokinases by their high substrate specificity for glucose (as shown for several members of Bacteria and Eukarya) have not been found yet in Archaea.

The distribution of glucokinases and hexokinases during evolution is complex and still puzzling. ADP-dependent sugar kinases (glucokinases and phosphofructokinases) are specific for a few Eurarchaeal (11, 16), and gene homologs encoding proteins of unknown function were identified in some Eukarya (11). Whereas highly specific ATP-dependent glucokinases, often together with other specific sugar kinases (e.g. fructokinases, mannokinases), are found in Bacteria and unicellular Eukarya, the nonspecific hexokinases seem to be characteristic of higher Eukarya, which, however, possess several hexokinase isoenzymes (for review, see Ref. 17).

Most eukaryal ATP-dependent hexokinases represent monomers with subunit molecular masses of 50 (fungi) or 100 kDa (vertebrates, plants). The major allosteric regulator of mammalian hexokinase isoenzymes I, II, and III is glucose 6-phosphate, which inhibits the enzyme. In addition, hexokinase I is activated by citrate and phosphate. Mammalian hexokinase isoenzyme IV is controlled mainly by a regulatory protein. Yeast hexokinases (P1, P3) are active as either monomer or dimer. In the presence of ATP and glucose the formation of the more active dimer is favored, and activation is observed by various metabolites (17, 18).

The bacterial ATP-dependent glucokinases are generally homodimers composed of 24–34-kDa subunits, and for Zymomo-
Archaeal Hexokinase

nus mobilis glucokinase an inhibition by glucose 6-phosphate and ADP is described (17, 18). The characterized archaean ADP-dependent glucokinases represent homodimeric (P. furiosus, 47-kDa subunit (1, 2)) or monomeric (T. litoralis, 52 kDa (2); M. jannaschii, 53 kDa (3)) proteins and seem to exhibit no obvious regulatory potential, although inhibition by the reaction product AMP (KIC = 0.06 mM (14)) is described for the P. furiosus enzyme.

Analyses combining sequence, structure, and functional information of 60 different sugar kinases (20 different sugar binding activities) indicate the presence of three distinct sugar kinase families (the hexokinase, ribokinase, and galactokinase families) which exhibit different three-dimensional structures and show no significant similarity (19). In addition, the authors provide evidence of convergent evolution of similar specificity in different structural families as well as in different branches of the same structural family. Recently, the crystal structure of the ADP-dependent glucokinase from T. litoralis was solved characterizing the enzyme as a member of the ATP-dependent ribokinase family (19, 20). Despite the striking differences observed between ATP-dependent hexokinases and glucokinases regarding their enzymatic properties and oligomeric state, structural comparisons reveal that they share a common three-dimensional fold and conserved sequence signatures characterizing them as members of the hexokinase family (19). The hexokinase family is subdivided into three subfamilies: (i) subfamily I, eukaryal hexokinases and glucokinases (yeast); (ii) subfamily II, bacterial glucokinases and fructokinases; and (iii) subfamily III, bacterial sugar kinases with diverse specificities (e.g. fucokinase, ribulokinase, glucokinase). More recent phylogenetic studies of the two glucokinases from the eukaryal parasasalid Trichomonas vaginalis indicate that they, as well as the enzyme of the diplomonad Giardia intestinalis, are members of the bacterial glucokinase, fructokinase subfamily (subfamily II) and only distantly related to eukaryal hexokinases (21). Thus, this former exclusively bacterial subfamily was shown to comprise a mixture of eukaryal as well as bacterial glucokinases (group A), bacterial glucokinases (group B), and fructokinases (group C) (21).

A close phylogenetic relationship of sugar kinases to proteins with different function was shown by Tittgemeyer and co-workers (22). Based on sequence similarity they proposed a novel family of proteins, the “ROK family” (bacterial repressor protein, open reading frames of unknown function, and sugar kinases). So far, several glucokinases of the so-called ROK family have been characterized, and surprisingly the bacterial homologs seem to be involved in carbon catabolite (glucose) repression in addition to their metabolic activity as reported for the enzymes of Streptomyces coelicolor (23), Bacillus megaterium (24), Bacillus subtilis (25), Staphylococcus xylosus (26) and Corynebacterium glutamicum (27). Phylogenetic analyses identified also the recently described archaean hexokinase of A. pernix, which was, however, designated by the authors as glucokinase, and other archaean homologs as members of the ROK family (4). Unfortunately, available phylogenetic studies do not address the relationship between the so-called ROK family and the hexokinase family, thus leaving some confusion about the evolutionary relationship between these phylogenetic entities.

In Archaea ATP-dependent hexokinase activity has been reported recently for the A. pernix enzyme (glucokinase according to Ref. 4), and respective activities have been proposed for gene homologs identified in the genomes of Halobacterium sp. strain NRC-1, Thermoplasma acidophilum, T. volcanium, and Pyrolobus fumarii (4, 11). In addition the ATP-dependent phosphorylation of glucose has been shown in crude extracts of Thermoproteus tenax (28) and Desulfovibrio desulfuricans (29).

Thus the scarce knowledge about archaean sugar kinases, especially hexokinases and glucokinases, motivates intense studies to get more insight into the evolution of these enzymes, their diversification with respect to substrate specificity, the physiological background of cosubstrate specificity (ATP versus ADP), and their regulatory potential for directing the carbon flux through the various pathways.

To address the questions about the dominant phenotype(s) of nucleotide-dependent glucose-phosphorylating enzymes and their metabolic function in Archaea we focused on the ATP-dependent hexokinase of the hyperthermophilic Archaean T. tenax. T. tenax is a facultative chemoorganotroph (30, 31) that uses a modified nonphosphorylating Entner-Doudoroff (ED)1 pathway and a variant of the Embden-Meyerhof-Parnas (EMP) pathway for carbohydrate catabolism (28, 32). The EMP pathway is characterized by several unique features: (i) a reversible, nonallosteric PF-dependent phosphofructokinase (5); (ii) two different glycolaldehyde-3-phosphate dehydrogenases (33–35); and (iii) a pyruvate kinase with reduced allosteric potential (36). Thus, deviating from the classical version of the pathway with control points at the beginning (ATP-dependent hexokinase and phosphofructokinase) and at the end of the pathway (pyruvate kinase) regulation was shown to take place at the level of glycolaldehyde-3-phosphate (33, 34). Although ATP-dependent phosphorylation of glucose was demonstrated in T. tenax (28), so far no information was available about the enzyme, which catalyses the first committed step of the pathway and represents an important control point in many organisms.

**Experimental Procedures**

Chemicals and Plasmids—All chemicals and enzymes were purchased from Sigma, Merck, or Roche Diagnostics in analytical grade. For heterologous expression of hexokinase the vector pET-11c (Novagen) and for generating antisense mRNA the vector pSPT 19 (Roche Diagnostics) were used.

Strains and Growth Conditions—Mass cultures of T. tenax Kna1 (DSM 2078 (30, 31)) were grown under autotrophic and heterotrophic conditions as described previously (34). Escherichia coli strains DH5α (Invitrogen), BL21(DE3), and BL21-CodonPlus(DE3)-RII (Stratagene) for cloning and expression studies were cultured under standard conditions (37) following the instructions of the manufacturer.

Enzyme Assay—The hexokinase activity was determined at 50 °C using two different coupled assays. The phosphorylation of glucose was measured by coupling the reaction to the reduction of NADP+ using glucose-6-phosphate dehydrogenase (bakers yeast, EC 1.1.1.49) as auxiliary enzyme. The standard assay (1 ml total volume) was performed in 100 mM Tris/HCl (pH 7.5, 50 °C) in the presence of hexokinase (10–20 μg of protein), 4 mM ATP, 4 mM MgCl2, 1 mM NADP+, and 3 units of glucose-6-phosphate dehydrogenase. The reaction was started by the addition of 10 μl glucose.

The phosphorylation of different substrates was followed by coupling the formation of ADP to the oxidation of NADH via pyruvate kinase (rabbit muscle (EC 2.7.1.40) and β-lactate dehydrogenase (rabbit muscle, EC 1.1.1.27). The assay (1 ml total volume) was performed in 100 mM Tris/HCl (pH 7.5, 50 °C), 2 mM ATP, 2 mM MgCl2, 5 mM phosphoenolpyruvate, 20 mM of the tested sugar, 0.5 mM NADH, 25 units of pyruvate kinase, and 90 units of β-lactate dehydrogenase. The reaction was started by the addition of hexokinase (10–20 μg of protein).

Enzymatic activities were measured by monitoring the increase in absorption at 366 nm (ε2ADH 50°C = 3.36 mM cm−1; ε2ADP 50°C = 3.43 mM cm−1).

The substrate specificity for different sugars was examined following the formation of ADP and replacing glucose by other substrates such as fructose or mannose. The nucleotide and cation specificity, as well as effector studies, were determined via the standard enzyme assay with glucose as substrate. Effector studies were performed in the presence of

1 The abbreviations used are: ED, Entner-Doudoroff; EMP, Embden-Meyerhof-Parnas.

Downloaded from http://www.jbc.org/ on August 16, 2017
1 mM MgCl₂ and half-saturating concentrations of glucose (60 μM) and ATP (300 μM). The different substances were added at a concentration of 0.1, 1, 5, or 10 mM. To test the metal ion requirement, 2.5 mM EDTA or metal ions (Mg²⁺ and Mn²⁺) were added to the mixture. The protein concentration was determined according to the method of Bradford (38) using the Bio-Rad protein assay with bovine standard albumin as standard.

Cloning of the hxx Gene and Expression in E. coli—In the course of the ongoing T. tenax genome sequencing project an open reading frame with significant sequence similarity to glucokinases was identified (accession number AY510140, EMBL nucleotide sequence data base). For expression the hxx gene was cloned into pET-11c using two new restriction sites (Ndel and BamHI) introduced by PCR mutagenesis with the following primer set: GK-2-Ndel-F (TCCGAGTCTGAGATACTTGCTGGCGATCG, sense) and GK-2-BamHI-rev (GGCCATCG, antisense). The mutations are shown in boldface, and the resulting Ndel and BamHI restriction sites are underlined. PCR mutagenesis was performed using Pfu polymerase and genomic T. tenax DNA as template. The sequence of the expression clones was controlled by sequencing. Expression of the T. tenax enzyme in E. coli BL21(DE3) and BL21-CodonPlus(DE3)-RIL was performed following the instructions of the manufacturer (Stratagene).

Purification of the Recombinant Hexokinase—Recombinant E. coli cells (5 g, wet weight) were suspended in 15 ml of 100 mM HEPES/KOH (pH 7.5) containing 7.5 mM diithiothreitol (buffer A) and passed three times through a French pressure cell at 150 MPa. Cell debris and unbroken cells were removed by centrifugation (85,000 × g for 30 min at 4 °C), and the resulting crude extract was diluted 1:1 with buffer A, heat-pretreated (90 °C, 30 min), centrifuged again (20,000 × g for 30 min at 4 °C), and dialyzed overnight against 50 mM HEPES/KOH (pH 7.5), 7.5 mM diithiothreitol (2-liter volume, 4 °C). The dialyzed fraction was applied to a Q-Sepharose fast flow (Amersham Biosciences) column (volume 30 ml, C16/20) equilibrated in the same buffer at a flow rate of 0.25 ml/min. Hexokinase activity was detected only in flow-through fractions, and those containing the homogeneous enzyme solution were pooled. Gel filtration experiments were performed as described previously (5).

Northern Blot Analyses of the hxx Transcript—Digoxigenin-labeled antisense mRNA of hexokinase was obtained by cloning a part of the hxx gene (278 bp) into pSPT19 (Roche Diagnostics) by PCR mutagenesis using the primer set GK-pSPT19-EcoRI-f (TGGCCCAACAGGATCTTTCTGGCCGATCG, sense) and GK-pSPT19-BamHI-rev (GGCCGCTGGATCCGAACCTGACG, antisense). The introduced EcoRI and BamHI restriction sites are underlined and the mutations are shown in boldface. In vitro transcription from the T7 promoter of pSPT19 was performed according to the instructions of the manufacturer. Preparation of sense and antisense RNA and partial and full sequences were obtained from the analyses leaving a total of 53 sequences and 212 amino acid positions. The phylogenetic tree shown in Fig. 4 corresponds to the consensus tree obtained by the MrBayes analysis, the option majority rule consensus tree was used (41). MrBayes version 2.01 was used for the Bayesian inference with 100,000 generations using the JTT model and gamma distributed rates, the trees were sampled every 10 generations (41). The number of generations needed until convergence around a stable likelihood value was in the range of 5–10%. The posterior probabilities given in the MrBayes consensus tree are indicated as percent values in the phylogenetic tree. A maximum likelihood based estimate of the gamma parameter as well as of the statistical support for internal nodes (quartet puzzling support values) was performed using the program TREE-PUZZLE v.5 (42). Distance analyses including 1000 bootstrap replicates were performed with the MEGA2 package using gamma correction and the Minimal evolution approach (43). A Maximum parsimony bootstrap analysis was performed using PAUP* with 1000 bootstrap replicates and 10 times random addition (44).

RESULTS AND DISCUSSION

Nucleotide Sequence—The hxx gene was identified in the T. tenax genome sequencing project by similarity to bacterial and archaeal ATP-dependent glucokinases of the so-called ROK family. The open reading frame of the hxx gene (906 bp) encodes a polypeptide of 301 amino acid residues with a calculated molecular mass of 32,543 kDa. The coding region of the hxx gene overlaps by 1 bp with an upstream located open reading frame (orfX) of 333 bp (110 amino acid residues), which encodes a protein of unknown function. The adenosine (A) of the start codon (ATG) is the last nucleotide of the triplet encoding the final amino acid lysine (AAA) of orfX (Fig. 1). The same overlap was found in the fba-pfp operon, coding for fructose-1,6-bisphosphate aldolase and PP₄-dependent phosphofructokinase (45), and in the pfp-gap operon, coding for 3-phosphoglycerate kinase and NADP⁺-dependent glyceraldehyde-3-phosphate dehydrogenase (44), suggesting that also the hxx gene is organized in an operon. Similar to the fba-pfp operon, AT-rich regions, which correspond well to the crenarchaeal consensus promoter sequences, were recognized 22–35 bp upstream of orfX (Fig. 1) but were absent in front of the succeeding hxx gene (46–48). The putative TATA box (CTTTAAGT, consensus sequence (C/T)TTTTAAA) is centered on position (−25)–(−26), and directly upstream the transcription factor B recognition element (AGTGAA; consensus sequence (A/G)(A/T)AA(A/T)) could be identified. As already reported for the fba-pfp operon, also in the case of the orfX-hxx operon, a putative ribosome binding site (RBS, GGAGG) seems to be absent (45).

The recently described so-called ATP-dependent glucokinase (gk) gene of A. pernix (APE2091 (1,131 bp) (4)) was shown to contain an additional noncoding extension of 171 bp (5’-end), as demonstrated by N-terminal sequencing of the native enzyme. Thus, the coding gk gene of A. pernix comprises 963 bp and codes for a protein of 320 amino acids, which is quite similar to the T. tenax enzyme.

Transcript Analysis—Northern blot experiments were per-
formed to determine whether the hxx gene and the upstream located orfX are cotranscribed, as suggested by the juxtaposition of both genes, and whether the gene expression is under differential transcriptional control. The effect of carbon source on hxx transcription was analyzed with total RNA from autotrophically and heterotrophically grown cells and digoxigenin-labeled, hxx-specific antisense mRNAs. The RNA molecular size standard (left) and the derived transcript size (arrows, right) are shown.

The role and function of the upstream located orfX are still unclear. In BLASTX as well as PSI BLAST searches only one homolog was identified which encodes a hypothetical protein (AE099933) in *P. aerophilum* (44.6% identity). However, the position of this homolog in *P. aerophilum* (NP_560752.1, location 2080434–2080739) differs from the one of the orfX in *T. tenax*. It is located upstream of a molybdenum cofactor biosynthesis protein (moaA, NP_560753.1, location 2080744–2082663) and not adjacent to the putative glucokinase of *P. aerophilum* (NP_560732.1, location 2060281–2061171). A similar coexpression of the glucokinase gene with an upstream located open reading frame (ORF2, 573 bp) of unknown function was described for the bacterial glucokinase of *S. coelicolor* A3 (2) (49), but the bacterial open reading frame and the two open reading frames do not exhibit any similarity. Thus, from these comparisons no hints for a functional connection of hexokinase and the orfX gene product can be deduced.

However, the absence of putative archaea promoter structures upstream of the hxx gene (Fig. 1) suggests that posttranscriptional RNA processing rather than alternative transcription termination and initiation might be involved in the generation of the two different transcripts. Particularly the presence of the hxx transcript only under heterotrophic growth conditions, where the active enzyme is required, hints at a regulatory effect. The involvement of posttranscriptional processes in the generation of mono- and bicistronic messages is currently under investigation.

Expression of the hxx Gene in *E. coli* and Purification of the Recombinant Protein—The hxx gene coding for hexokinase of *T. tenax* was expressed in *E. coli* using the pET-11c vector system. Analysis of the hxx gene sequence revealed differences in codon usage of *T. tenax* and the bacterial expression host especially for arginine (seven times AGA, nine times AGG). Thus the expression was performed in BL21(DE3) in the absence and presence of the RIL plasmid, which codes for rare tRNAs of arginine, isoleucine, and leucine (Stratagene) in *E. coli*. Expression was only observed by coexpression of the RIL plasmid. The enzyme was purified by heat precipitation and Q-Sepharose for biochemical characterization. From 5 g of recombinant *E. coli* cells 70 mg of homogeneous protein with a specific activity of 15.2 units/mg protein (90 °C) was recovered.

**Biochemical Properties of the Hexokinase of *T. tenax*—**The hexokinase of *T. tenax* revealed a molecular mass of 33 kDa in SDS-PAGE, thus being in good agreement with the calculated molecular mass of 32.543 kDa. The native molecular mass determined by gel filtration experiments was ~40 kDa, suggesting a monomeric structure. A similar monomeric structure with a subunit size of about 33 kDa has been reported for the *A. pernix* hexokinase (4). Thus, regarding subunit size (Table I), the two archaeal enzymes resemble bacterial glucokinases ranging from about 24 kDa (e.g. *Streptococcus mutans, E. coli*) to 33 kDa (e.g. *Z. mobilis, Bacillus stearothermophilus*) and numerous other bacterial sugar kinases that exhibit subunit sizes of about 32–37 kDa (e.g. fructokinase *Z. mobilis*, ribokinase *E. coli* (for review, see Ref. 17)). However, most of these sugar kinases represent dimers, whereas the two archaeal enzymes seem to be monomer like most eukaryal hexokinases, even though they have much higher subunit molecular masses of 50 kDa (fungi, invertibrate hexokinases) and 100 kDa (vertebrate hexokinases). In contrast the nonhomologous, archaeal ADP-dependent glucokinases (1–3) resemble eukaryal hexokinases in both subunit size and oligomerization state.

**Enzymatic Properties of the Recombinant Hexokinase of *T. tenax*—**Although the hxx gene product of *T. tenax* shows high sequence similarity to bacterial glucokinases the enzyme exhibits, like the so-called *A. pernix* glucokinase, broad substrate specificity. The enzyme phosphorylates glucose, fructose, mannose as well as 2-deoxyglucose (*K_m* values of 8179, 1389, 4276, and 4416 (mm⁻¹ min⁻¹), respectively) and thus represents a hexokinase rather than a glucokinase, however, with preferred specificity for glucose. The *T. tenax* hexokinase follows classical Michaelis-Menten kinetics for glucose, fructose, mannose, 2-deoxyglucose, and ATP. The respective *K_m* and *V_max* values are given in Table II. A similar preference for glucose is reported for the *A. pernix* enzyme, although the efficiency for mannose and 2-deoxyglucose (44 and 31% efficiency of the *T. tenax* enzyme, respectively) is reduced (4). The broad substrate specificity of both archaeal enzymes clearly classifies them as hexokinases.

Galactose, ribose, xylose, and phosphorylated hexoses (glucose 6-phosphate, glucose 1-phosphate, fructose 6-phosphate, fructose 1-phosphate) are not phosphorylated by the *T. tenax* enzyme. In addition to ATP (100% relative activity), ITP (40% relative activity), GTP (18% relative activity), and cGMP (17% relative activity) also serve as phosphoryl group donors, whereas CTP, ADP, UTP, and polyphosphates (P₅₋, P₆₋, P₇₋, P₂₀ₒ) were not active. The *A. pernix* enzyme exhibits a similar cosubstrate specificity (100, 40, 9, and 3% relative activity for ATP, ITP, GTP, and UTP, respectively), although a rudimentary activity was observed with PP, (1%). This feature of archaeal enzymes is rather unusual for most eukaryal hexokinases, which prefer ATP as nucleotide substrate and resemble more bacterial glucokinases. For example, the glucokinase of *B. stearothermophilus* was reported to use ATP (100%) as well as ITP (75%), UTP (32%), GTP (30%), and CTP (10%) as phosphoryl donors (50).

The recently described archaeal ADP-dependent glucoki-
nases of *P. furiosus* and *T. litoralis* display high specificity for glucose, no activity with fructose, and only very limited activity with mannose, galactose, glucosamine, 1,5-anhydroglucitol, and 2-deoxyglucose ((1, 2), Table II) was detected. Both enzymes showed comparable activity with ADP and CDP, whereas nucleotide triphosphates, GDP, IDP, and UDP and aszymes showed comparable activity with ADP and CDP, and 2-deoxyglucose ((1, 2), Table II) was detected. Both enzymes showed comparable activity with ADP and CDP, whereas nucleotide triphosphates, GDP, IDP, and UDP and as demonstrated for the *P. furiosus* enzyme phosphoenolpyruvate, PP, and polyphosphates were not used as phosphoryl group donors. The bifunctional ADP-dependent glucokinase/phosphofructokinase of *M. jannaschii* catalyzes the phosphorylation of glucose (glucokinase activity) and fructose 6-phosphate (phosphofructokinase activity) and is only weakly active with fructose and 2-deoxyglucose using ADP and to some extent GDP as phosphoryl group donors (3).

Thus, regarding the broad sugar specificity, the archaeal hexokinases of *T. tenax* and *A. pernix* resemble classical eu- karyal hexokinases, whereas their cosubstrate specificity is more similar to bacterial glucokinases.

The *T. tenax* enzyme, like the *A. pernix* hexokinase and all sugar kinases described so far, showed strict requirement for divalent metal ions and was completely inhibited in the presence of 2.5 mM EDTA. Maximal activity (8.9 units/mg protein) was observed in the presence of 1 mM Mg2+ which could be replaced partially by Mn2+ (5.1 units/mg of protein).

### Influence of Metabolites on Hexokinase Activity and Physiological Role of the Enzyme—The regulation of hexokinase activity by various metabolites was tested at half-saturating concentrations of 60 μM glucose and 300 μM ATP and saturating concentrations of 1 mM MgCl2. As shown in Table III inhibition was observed in the presence of glucose 6-phosphate, ADP, UDP, AMP, Pi, and PP; however, significant effects were only observed at high concentration of the ligands, especially of the nucleotides and glucose 6-phosphate, leaving doubts as to whether these effects are of physiological relevance.

The regulation by glucose 6-phosphate seems to be a general feature of some eukaryal hexokinases and bacterial glucoki- nases. Glucose 6-phosphate is the major inhibitor of the 100-kDa vertebrate hexokinases (hexokinase A, B, and C), and inhibition by glucose 6-phosphate and ADP has been reported for the bacterial *Z. mobilis* glucokinase (18, 51) and by glucose 6-phosphate and ADP as well as AMP for the *S. mutans* enzyme (52). At first glance, the inhibition by PP, (1 mM PP, 30% residual activity) seems to be significant, but considering the low intracellular PP concentration in *T. tenax*, as deduced from micromolar *Km* values of PP-utilizing enzymes, such as the PP-dependent phosphofructokinase (23 μM (5)), this effect might be of minor importance. Other species-specific inhibitors or activators of eukaryal hexokinases which are not regulated by glucose 6-phosphate (e.g., ATP, glucose, citrate, trehalose 6-phosphate (17, 18)) and intermediates of carbohydrate metabolism, such as maltose, glucose 1-phosphate, fructose 6-phosphate, fructose 1,6-bisphosphate, 6-phosphogluconate,gluonate, galactose, ribose, dihydroxyacetonephosphate, glyceralddehyde 3-phosphate, 2-phosphoglycerate, phosphoenolpyruvate, pyruvate, glyceraldehyde, acetaldheyde, and α-keto- glutarate showed no effect on the *T. tenax* hexokinase.

None of the identified effectors of the *T. tenax* hexokinase affects the *A. pernix* enzyme (tested: glucose 6-phosphate, ADP, AMP, P); (4). Thus archaeal hexokinases seem to possess low if at all any regulatory properties. Strikingly, also archaeal ADP-dependent glucokinases do not exhibit remarkable regulatory potential (1–3), although inhibition by AMP was reported for the *P. furiosus* enzyme (14).

* T. tenax uses two different pathways for glucose catabolism, the variant of the EMP pathway and the modified ED pathway (28, 32). Thus, glucose is processed via the nonphosphorylated ED pathway, whereas glucose 6-phosphate is metabolized via glycolysis or forms the substrate for trehalose and glycogen synthesis.

### Phylogenetic Analysis—Data base searches with the *hxk* gene of *T. tenax* revealed a high number of sequences with apparent similarity to glucokinases and other members of the so-called ROK family in all three domains of life. No similarity was observed to archaeal ADP-dependent glucokinases of *P. furiosus* and *T. litoralis* or the bifunctional glucokinase/phosphofructokinase of *M. jannaschii*, which were shown previously to belong to the ribokinase family (19, 20). Archaeal *hxk* homologs were identified in the completely sequenced genomes of *P. aerophilum*, *A. pernix*, *Halobacterium* sp. strain NRC-1, *T. acidophilum*, *T. volcanium*, *Ferroplasma acidarmanus*, and *Archaeoglobus fulgidus*. So far the only characterized enzymes are the hexokinases of *T. tenax* and *A. pernix* (4), and nothing is known about the enzymatic function of the other archaeal gene products. BLAST searches with the *hxk* gene of *T. tenax* reveal significant similarity to members of the hexokinase family, which exhibit a common three-dimensional fold and con-

![](https://bioclient.biochemistry.gatech.edu/figs/150083.png)

**TABLE I**

<table>
<thead>
<tr>
<th>Domain</th>
<th>Species</th>
<th>Enzyme type</th>
<th>Growth optimum</th>
<th>Molecular mass</th>
<th>Oligomeric state</th>
<th>Ref.</th>
</tr>
</thead>
<tbody>
<tr>
<td>Archaea</td>
<td><em>T. tenax</em></td>
<td>Hexokinase</td>
<td>86</td>
<td>41</td>
<td>33</td>
<td>1</td>
</tr>
<tr>
<td><em>A. pernix</em></td>
<td>Hexokinase</td>
<td>90–95</td>
<td>30–33</td>
<td>36</td>
<td>1</td>
<td>(4)</td>
</tr>
<tr>
<td><em>P. furiosus</em></td>
<td>ADP-glucokinase</td>
<td>100</td>
<td>90–100</td>
<td>47/47</td>
<td>2</td>
<td>(1, 2)</td>
</tr>
<tr>
<td><em>T. litoralis</em></td>
<td>ADP-glucokinase</td>
<td>88</td>
<td>54</td>
<td>52</td>
<td>1</td>
<td>(2)</td>
</tr>
<tr>
<td>Bacteria</td>
<td><em>E. coli</em></td>
<td>Glucokinase</td>
<td>37</td>
<td>49</td>
<td>24.5</td>
<td>2</td>
</tr>
<tr>
<td><em>B. stea-thermotophilus</em></td>
<td>Glucokinase</td>
<td>50</td>
<td>67</td>
<td>34.5</td>
<td>2</td>
<td>(50)</td>
</tr>
<tr>
<td>Eukarya</td>
<td><em>S. cerevisiae</em></td>
<td>Hexokinase</td>
<td>25</td>
<td>102</td>
<td>51</td>
<td>2</td>
</tr>
<tr>
<td>Rat</td>
<td>Hexokinase A</td>
<td>37</td>
<td>98</td>
<td>98</td>
<td>1</td>
<td>(17)</td>
</tr>
</tbody>
</table>

* Described as glucokinase with broad substrate specificity (4).
served sequence signatures (19, 21). Closer inspection of members of the hexokinase family, reported previously by Bork and co-workers (19), and the so-called ROK family showed that the latter sequence group comprises enzymes of the prokaryotic glucokinase fructokinase group B and C (21) of subfamily II of the hexokinase family (19, 21, 22). Further on, the glucokinase of *Z. mobilis*, which is referred to as member of the so-called non-ROK protein family (23, 27) because it shows no obvious homology to glucokinases of the so-called ROK family and is not involved in catabolite repression, was shown earlier to be a member of the same subfamily II, group A. Thus the hexokinase subfamily II comprises both members of the so-called ROK family (group B and C) as well as homologs of the non-ROK protein family (group A), and because of their position within the hexokinase family they should be referred to as ROK group and non-ROK protein group.

Sequence identities in the range of 20–25% between the basal lineages (based on the 212 amino acid residues used for the phylogenetic analysis) indicate that the ROK group represents a divergent set of proteins, although conserved sequence motifs confirm their origin from a common ancestor. As documented in the sequence alignment (Fig. 3) the transcriptional repressor protein (*e.g.* N-acetylglucosamine repressor (NagC/R) *E. coli*) shows an N-terminal extension of about 80 amino acid residues, which harbors the N-terminal DNA binding domain with the helix-turn-helix motif (22). Sugar kinases contain the typical N-terminal ATP binding site (Fig. 3, residues 4–10 of the *T. tenax hxk* sequence, consensus motif AID-LGGT (54)/DIGGT (27, 55)), and all members exhibit the two proposed ROK consensus motifs (Fig. 3, residues 119–146 and 157–170 of the *T. tenax hxk* sequence (27, 55)).

For a better resolution of the functional diversification of the ROK group in the course of evolution we constructed phylogenetic trees with 53 protein sequences (212 unambiguously aligned amino acid positions), representing members with different specificities of all important taxa, using maximum likelihood, maximum parsimony, and distance-based methods (Fig. 4). The vast majority of the sequences is of bacterial (a subset of 44 sequences was incorporated into the data set), only eight are of archaeal, and one is of eukaryal (mammalian) origin. This unequal distribution can only partially be explained by the fact that more bacterial than archaeal and eukaryal genomes (166, 18, and 18 genomes, respectively) are released. It is more likely because Archaea and Eukarya possess only one...
homolog per genome, whereas some bacteria contain several
distantly related genes (e.g. E. coli and B. subtilis; Fig. 4),
which code for proteins with different enzymatic activities. In
addition, the occurrence in Eukarya is restricted to mammals
(55), and several Archaea (e.g. methanogens, Sulfolobales, Py-
rococcales, Thermococcales etc.) do not have any homologous
copy at all.

The phylogenetic analyses showed a complex tree topology
with a high number of deeply branching lineages and members,
displaying different enzymatic activities. The presence of sev-
eral paralogous genes from the same organism in different
basal lineages (e.g. enzymes of E. coli and Bacillus halodurans)
suggests that early gene duplications confer largely to the
characteristic topology of the tree. Examples for more recent
gene duplications are found in the glucokinases of the high
GC Gram-positive bacteria (Thermobifida fusca, S. coelicolor)
and for the transcriptional regulators (E. coli). The obvious
limited archaeal distribution (eight members in three groups)
raises questions about their evolutionary history. We assume
that the present day picture presents the results of early gene
duplications (at least some of them older than the separation
of the bacterial and archaeal lineages, and some probably
happened early in bacterial evolution) combined with subse-
quent differential losses and probably some late lateral gene
transfers.

Interestingly, the biochemically characterized members
found in different branches indicate that the early gene
duplications are the basis for the creation of different substrate
specificities, e.g. enzyme activities. Unfortunately there are,
at least, only few biochemically characterized members in the
different lineages, thus allowing only very preliminary assign-
ments of the enzymatic phenotypes. In contrast to previous
work (4), hexokinase activity is assumed for the archaeal clus-
ter, which comprises both crenarchaeal and euryarchaeal ho-
mologs including the characterized T. tenax and A. pernix
enzymes. For the bacterial sistergroup (Proteobacteria, low GC
Gram-positive bacteria) N-acetylmannosamine kinase activity
is suggested because of the presence of the yhcI gene product of
E. coli for which the respective kinase activity is proposed
because of its organization in the nan operon, although no
biochemical information is available (56). No enzymatic infor-
mation is available for the archaeal Thermoplasmales group.
Because at least T. acidophilum is supposed to use only the
nonphosphorylated ED pathway for carbohydrate catabolism
(57), the presence of hexokinase would be rather surprising,
although an involvement in biosynthesis cannot be excluded.
Fructokinase activity is suggested for the low GC Gram-posi-
tive and proteobacterial sequences with the characterized en-
zyme of Z. mobilis (58). For several branches with one or two
representatives no suggestion for enzyme activity is possible.
The only eukaryal member of the ROK group is the
N-acetylmannosamine kinase domain of the bifunctional UDP-
N-acetylglucosamine 2-epimerase/N-acetylmannosamine kinase
of mammals (55, 59) indicating that either losses in other
Eukarya or lateral gene transfer might be involved. The eu-
karyal sequence is deeply separated from the bacterial se-

Fig. 3. Multiple sequence alignment of nine selected members of the ROK group. The enzyme activity is given at the end: HK, hexokinase; NMAK, N-acetylmannosamine kinase; GK, glucokinase; R, repressor protein; FK, fructokinase; ALSK, allose kinase. Boldface letters indicate amino acid residues used in the phylogenetic analyses. Conserved sequence motifs, such as the helix-turn-helix (HTH) motif (NagCR, E. coli) (22), the postulated ATP binding site and the two ROK consensus motifs (27, 54, 55) are shown above the sequences as black and white boxes, respectively.
quence of Listeria innocua and the prokaryotic sistergroup including the archaeal A. fulgidus sequence, thus allowing no prediction for the three prokaryotic sequences. Strikingly, eukaryal homologs are only identified in mammals. The presence of characterized glucokinases in the bacterial glucokinase group (S. xylosus (26), B. megaterium (24), B. subtilis (54), not included in Fig. 4 (P54495), S. coelicolor (49) and C. glutamicum (27)) let us assume that at least these two branches bear clearly glucokinases. In addition, glucokinase activity might be assumed for the other related bacterial group (Cyanobacteria and Thermotoga). The monophyletic group of the transcriptional repressors is in addition defined by their characteristic N-terminal extension and the characterized NagC/R (E. coli) and XylR (B. subtilis) repressor proteins. Putative allose kinase activity is suggested for the E. coli enzyme because of the organization of the gene in the allose operon (63).

In summary the phylogenetic tree (Fig. 4) presents proteins of the ROK group, which is with respect to the phylogenetic distribution as well as regarding the metabolic function of its members much more complex than reported earlier (22). In addition to many new bacterial homologs also archaeal members were identified with the hexokinases of T. tenax and A. pernix being the first and only characterized archaeal enzymes (4). The functional assignment of ROK group members, which needs further experimental confirmation, suggests the presence of hexokinases, fructokinases, N-acetylmannosamine kinases, glucokinases, repressor proteins and allose kinases in Bacteria and of hexokinases in Archaea. The identification of hexokinases as members of the ROK group strengthens the hypothesis of convergent evolution by Bork and co-workers (19), who proposed independent evolution of similar specificities not only in different structural families (e.g. fructokinases in the hexokinase and ribokinase family) but also in different branches of the same structural family (e.g. glucokinases and hexokinases in the eukaryotic subfamily I and the prokaryotic glucokinase and fructokinase subfamily II (ROK group)).

Two Types of Glucose-phosphorylating Enzymes in Archaea: ATP-dependent hexokinases and ADP-dependent Glucokinases—At our present knowledge, Archaea use two types of kinases for phosphorylation of glucose differing in cosubstrate specificity: ATP-dependent hexokinases (T. tenax, A. pernix, Halobacterium sp. strain NRC1) and ADP-dependent glucokinases (Pyrococcus, Methanococcus, Methanosarcina, and Thermococcus species). In some archaeal genomes no homologs of known phosphorylating enzymes have been found, leaving the question open as to whether glucose-phosphorylating enzymes are generally absent or if hitherto unknown enzyme types are present (e.g. Sulfolobus species). ATP-dependent hexokinases

FIG. 4. Phylogenetic tree of the ROK group (groups B and C) of the hexokinase family. Bootstrap proportions according to Bayesian analysis (Ba), Quartet Puzzling (QP), minimal evolution (ME), and maximum parsimony (MP) are given as numbers at the corresponding nodes. Only values greater 30% are shown. Archaeal members are indicated by boldface letters, the biochemical characterized enzymes are underlined, and the respective enzyme activity is given: HK, hexokinase; GK, glucokinase; R, repressor protein; FK, fructokinase; ALSK, allose kinase; NMAK, N-acetylmannosamine kinase. Additional taxonomic information is given in parentheses in front of the species name (A, Archaea; Ca, Crenarchaea; Ea, Euryarchaea; Es, Eosarchaeota; y, group of Proteobacteria; low GC, low GC Gram-positive Bacteria; high GC, high GC Gram-positive bacteria; cyano, Cyanobacteria; Toga, Togales; green sulfur, green sulfur bacteria). The accession numbers are given after the @ sign.
and ADP-dependent sugar kinases are members of two different enzyme families, the hexokinase and ribokinase family, respectively (19, 20). Whereas ADP-dependent sugar kinases are found only in Euryarchaeum (14), hexokinases were identified in Euryarchaeum as well as Crenarchaeum.

The unusual presence of ADP-dependent enzymes (glucokinases and phosphofructokinases) in Archaea is generally discussed in terms of metabolic adaptation to high temperatures and to conditions of low energy (1–3, 6, 11, 14, 20). Several factors argue against the assumption that the use of ADP as phosphoryl group donor is the result of metabolic redirection. (i) Although ATP is less thermostable than ADP (half-lives of 115 and 750 min at 90 °C, respectively (14)) the half-life of ATP is much higher than for several intermediates of the EMP pathway such as phosphoenolpyruvate, glyceraldehyde 3-phosphate, or 1,3-bisphosphoglycerate (with half-lives of 6 min at 90 °C (36), 3.4 min at 80 °C (64), and 1.6 min at 60 °C, respectively). Therefore, the thermostability of ATP should not represent the bottleneck for thermoadaptation. (ii) ADP-dependent sugar kinases (hexokinases/phosphofructokinases) are present in other hyperthermophiles (T. tenax optimum temperature = 86 °C; A. pernix optimum temperature = 90–95 °C) with a optimum growth temperatures similar to that of P. furiosus (optimum temperature = 100 °C), and vice versa ADP-dependent sugar kinases are also found in mesophiles (Methanosarcina mazei), which certainly do not need to avoid ATP for thermoadaptive reasons (4, 7–11, 16). (iii) ATP- and ADP-dependent kinases coexist in the same organism as shown for the hyperthermophile P. furiosus, which possesses in addition to its ADP-dependent sugar kinases an ADP-dependent galactokinase (14). All of these facts indicate that extreme growth temperature is not the selection factor for the occurrence of ADP-dependent sugar kinases in the central metabolism of Archaea.

However, the presence of ADP-dependent sugar kinases might be governed with generally or transiently limiting ATP concentrations. Strikingly, ADP-dependent sugar kinases seem to be restricted to obligate heterotrophs such as members of the genera Pyrococcus and Thermococcus and to the chemolithoautotrophic, glycerol-consuming methanogenic genera Methanococcus, Methanosaeta, and Methanosaeta (14). The heterotrophs depend exclusively on energy modified by the central version of the EMP pathway for energy conservation and from the rather low energy gain of four ATP/glucose, including the succeeding acetyl-CoA synthetase reaction (ADP-forming) (65); one would expect a rather low energy charge and with that a low intracellular ATP content in general. On the other hand, chemolithoautotrophs methanogens presumably need their sugar kinases in an energy emergency, i.e. when the energy charge falls below a certain level, to use stored energy sources such as glycerol (66). In both cases, sugar kinases of the central metabolism must work generally or temporarily in the presence of low ATP concentrations, which cannot guarantee a trouble-free function, and therefore ADP-dependent kinases might be preferred. Contrary to that, the more efficient and more versatile energy metabolism of A. pernix (aerobic respiration), T. tenax (sulfur respiration; SV/H2 autotrophy), and Halobacterium sp. (fermentation, aerobic respiration, photosynthesis) probably guarantees a higher and continuous level of ATP, which allows the usage of ATP-dependent kinases for a reliable supply of glucose 6-phosphate in their central carbohydrate metabolism.

Acknowledgment—We thank R. Hensel for critically reading the manuscript and for continuous support.

3 D. Hess and R. Hensel, unpublished data.
The Hexokinase of the Hyperthermophile Thermoproteus tenax: ATP-DEPENDENT HEXOKINASES AND ADP-DEPENDENT GLUCOKINASES, TWO ALTERNATIVES FOR GLUCOSE PHOSPHORYLATION IN ARCHAEA
Christine Dörr, Melanie Zaparty, Britta Tjaden, Henner Brinkmann and Bettina Siebers

doi: 10.1074/jbc.M301914200 originally published online March 7, 2003

Access the most updated version of this article at doi: 10.1074/jbc.M301914200

Alerts:
• When this article is cited
• When a correction for this article is posted

Click here to choose from all of JBC's e-mail alerts

This article cites 64 references, 20 of which can be accessed free at http://www.jbc.org/content/278/21/18744.full.html#ref-list-1