Genome sequence of *Picrophilus torridus* and its implications for life around pH 0

O. Fütterer*, A. Angelov*, H. Liesegang*, G. Gottschalk*, C. Schleper†, B. Schepers‡, C. Dock‡, G. Antranikian‡, and W. Liebl*§

*Institut of Microbiology and Genetics, University of Goettingen, Grisebachstrasse 8, D-37075 Goettingen, Germany; †Institut of Microbiology and Genetics, Technical University Darmstadt, Schnittspahnstrasse 10, 64287 D-Darmstadt, Germany; and ‡Technical Microbiology, Technical University Hamburg–Harburg, Kasernenstrasse 12, 21073 D-Hamburg, Germany

Edited by Dieter Söll, Yale University, New Haven, CT, and approved April 20, 2004 (received for review February 26, 2004)

The euryarchaea *Picrophilus torridus* and *Picrophilus oshimae* are able to grow around pH 0 at up to 65°C, thus they represent the most thermoacidophilic organisms known. Several features that may contribute to the thermoacidophilic survival strategy of *P. torridus* were deduced from analysis of its 1.55-megabase genome. *P. torridus* has the smallest genome among nonparasitic aerobic microorganisms growing on organic substrates and simultaneously the highest coding density among thermoacidophiles. An exceptionally high ratio of secondary over ATP-consuming primary transport systems demonstrates that the high proton concentration in the surrounding medium is extensively used for transport processes. Certain genes that may be particularly supportive for the extreme lifestyle of *P. torridus* appear to have been internalized into the genome of the *Picrophilus* lineage by horizontal gene transfer from crenarchaea and bacteria. Finally, it is noteworthy that the thermoacidophiles from phylogenetically distant branches of the *Archaea* apparently share an unexpectedly large pool of genes.

Picrophilus torridus is a thermoacidophilic euryarchaeon thriving optimally at 60°C and pH 0.7. Strains of this species were first isolated from a dry solfataric field in northern Japan. In addition to being moderately thermophilic, the *Picrophilaceae* are the most acidophilic organisms known and are also able to grow at negative pH values. It was reported for *P. torridus* that even adaptation to conditions such as those in 1.2 M sulfuric acid is possible (1). Another unusual trait of *Picrophilus* is a very low intracellular pH of 4.6, in contrast to other thermoacidophilic organisms that maintain internal pH values close to neutral (2, 3). The high specialization of *Picrophilus* strains for growth in extremely acidic habitats is evident from their inability to grow at pH values above 4.0, and makes them model organisms to study thermoacidophilic adaptation. Points of major interest include questions about properties and mechanisms that ensure viability under these conditions. Specific adaptation mechanisms can be expected at the levels of (*i*) structure and function of biomolecules and (sub-)cellular structures, (*ii*) physiology and metabolic features, and (*iii*) regulation of gene expression. For example, *P. torridus* cells need a specifically adapted membrane. Indeed, the membrane of *P. torridus* displays a very low proton permeability, is highly acid stable, and loses its integrity when incubated at pH 7 (2). Furthermore, an acid-stable cell envelope and acid-resistant extracellular enzymes for the degradation of polymeric or oligomeric carbon sources are required. Finally, the organism needs to generate metabolic energy in substantial amounts to maintain the intracellular pH at an acceptable value.

The phylogenetic position of *P. torridus* lies within the order of Thermoplasmales, which comprises the thermoacidophilic group of the euryarchaeal branch of the domain *Archaea*, consisting of the *Thermoplasmaceae*, *Ferroplasmaceae*, and *Picrophilaceae* (Fig. 1). A second group of thermoacidophiles can be found inside the crenarchaeal branch, the *Sulfolobaceae*. With only a few exceptions, members of both groups are aerobic or microaerophilic, heterotrophic organisms that are often found to share the same habitat (4–6). After analysis of a number of archaeal and bacterial genomes, it has been argued that microorganisms that live together swap genes at a higher frequency (7, 8). With the genome sequence of *P. torridus*, five complete genomes of thermoacidophilic organisms are available, which allows a more complex investigation of the evolution of organisms sharing the extreme growth conditions of a unique niche in the light of horizontal gene transfer.

Methods

Sequencing Strategy. Genomic DNA of *P. torridus* strain DSM 9790 was extracted and sheared randomly. A shotgun library was constructed by using fractions with sizes ranging from 2 to 3 kb. All generated sequences were assembled into contigs with the PHRAP assembling tool (9) and have been edited with GAP4 of the STADEN software package (10). A total of 25,694 sequences were aligned leading to a database with a 9.4-fold coverage and a statistical error rate of below 1 in 2,000,000 bp. Gap closure was accomplished by primer walking on plasmids originating from the plasmid library and from PCRs with chromosomal DNA as template. Gene and gene order comparison with already sequenced genomes served as verification for the assembly of the contigs to the closed chromosome.

ORF Prediction and Annotation. ORFs likely to encode proteins were predicted by the YACOP software package (11) (www.g2l.bio.uni-goettingen.de/software), based on the algorithms CRITICA (12), ORPHEUS (13), and GLIMMER (14). Automatic and manual annotation was accomplished with the ERGO annotation tool (Integrated Genomics), which was refined by searches against PFAM, PROSITE, PRODOM, and COGS databases, in addition to the BLASTP (15) versus SWISSPROT, NR, and TCDB databases. Putative signal peptides were predicted by using the SIGNALP program (16).

Homology Prediction and Horizontal Gene Transfer. For gene comparison, homology was specified as 30% amino acid sequence identity. Ortho- and paralogous sequences were counted only once. The threshold for specifying genes into the categories archaeal, bacterial, eukaryotic, or thermoacidophilic was set at a BLAST *e*-value of 1e-05 at the amino acid sequence level.

Results

General Features of the *P. torridus* Genome. The genome of *P. torridus* consists of a 1,545,900-bp large single circular chromosome and contains 1,535 ORFs (Table 1). A total of 92% of the sequence is coding, which represents the highest coding density in the genomes of the thermoacidophile group (89%, 87%, 85%, and 85% for *Thermoplasma acidophilum*, *Thermoplasma volcanii*, *Sulfolobus*

This paper was submitted directly (Track II) to the PNAS office.

Abbreviation: ABC, ATP-binding cassette.

Data deposition: The sequence reported in this paper has been deposited in the GenBank database (accession no. AE017261).

§To whom correspondence should be addressed. E-mail: wliebl@gwdg.de.

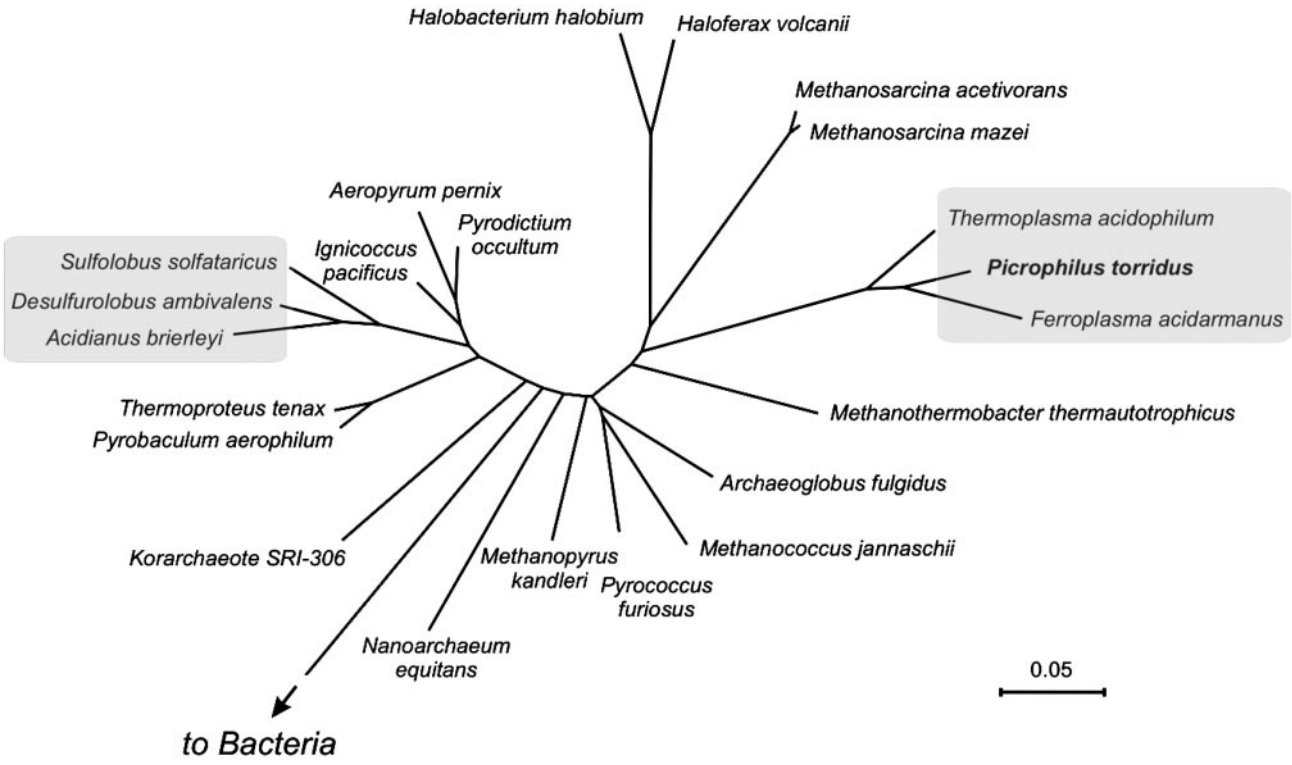

Fig. 1. 16S rRNA phylogenetic tree. Highlighted are the two thermoacidophilic groups of the archaea. Sequences were aligned with the CLUSTALW algorithm. The tree was build by neighbor joining by using the Kimura 2-parameter for distance calculation.

solfataricus, and *Sulfolobus tokodaii*, respectively). For 74% of all ORFs found in the genome, it was possible to assign a function. Of the 397 hypothetical ORFs, 79 are unique to *P. torridus*, whereas 318 showed similarities to hypothetical ORFs of other organisms. Interestingly, 174 of the latter ORFs have orthologs only in the genomes of other thermoacidophilic organisms, indicating that the thermoacidic environment forms an old and genetically distinct niche of life (see supporting information, which is published on the PNAS web site). This is supported by whole genome comparison of the amino acid sequences of the complete genomes of three prominent members of the thermoacidophilic group, *P. torridus*, *T. acidophilum*, and the crenarchaeon *S. solfataricus*, for homology (Fig. 2). *P. torridus* and *T. acidophilum* display significant homology in 66% of all genes, and these two euryarchaea share 58% and 62% genes, respectively, with the crenarchaeon *S. solfataricus*, but only ≈35% with the phylogenetically more closely related euryarchaeon *Pyrococcus furiosus*, meaning that in this case ecological closeness overrides phylogenetic relatedness. The statement that *P. torridus* shares nearly the same number of homologs with *T. acidophilum* and *S. solfataricus* but significantly fewer homologs with *Pyrococcus furiosus* remains true even when lowering the threshold for homology prediction from 30% to 25% identity.

The genome sequence data indicate that, in contrast to a previous report (4), the DNA-dependent RNA polymerase (RNAP) of *P.*

Table 1. General features of the *P. torridus* genome sequence

General features	Number
Size, bp	1,545,900
Coding region, %	91.7
G + C content, %	36
Total number of ORFs	1,535
ORFs with assigned function	983
ORFs without function	553
ORFs without function or similarity	73
ORFs without function, with similarity	480
ORFs with putative signal peptides	121
ORFs involved in transport	170
Stable RNAs	
rRNAs	1 5S; 1 16S; 1 23S
tRNAs	47

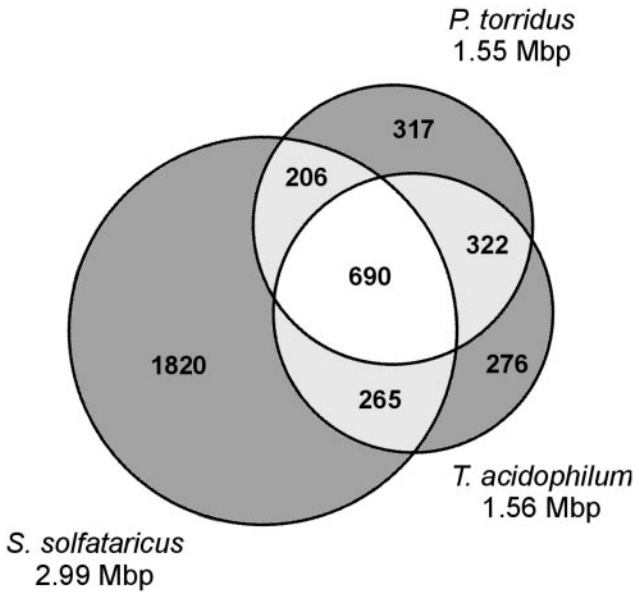

Fig. 2. Occurrence of homologous ORFs in *P. torridus*, *T. acidophilum*, and *S. solfataricus*. The size of the circles is proportional to the size of the genomes.

torridus is identical in subunit composition and highly similar in amino acid sequence to the RNAPs of *Ferroplasma acidarmanus* and *T. acidophilum*. The subunit composition of the *S. solfataricus* RNAP is different from the one of *P. torridus*, underlining the phylogenetic distance of these two organisms observed in the 16S–rRNA tree (Fig. 1).

The intracellular pH of 4.6 of *Picrophilus* cells is unusually low even compared to other thermoacidophiles, which maintain intracellular pH values close to neutral (2, 17). It was therefore assumed that not only the extracellular but also the intracellular proteins display acid stability, offering the unique opportunity for a genome-wide survey of the encoded polypeptides for the possible reasons of acid stability. A comparative analysis of the isoelectric point distribution and amino acid composition of the genome-derived proteome of *P. torridus* and other prokaryotes was carried out (see supporting information). In both cases, distributions strikingly different from the average could only be found in *Halobacterium sp*. However, *P. torridus* showed a slight increase of the average isoleucine content of its proteins compared to the reference organisms. In accordance, it was recently argued that an increase in hydrophobic amino acid residues on the protein surface may contribute to acid stability (18).

Clearly, one of the major keys to the adaptation of *P. torridus* to the acidic environment is the nature of its cell wall and membrane. The membranes of *Picrophilus* mainly consist of polar ether lipids like caldarchaeol and modified derivatives thereof. S-layer proteins apparently linked to polysaccharide chains form the cell wall (4). We could detect several ORFs necessary for diether and tetraether lipid biosynthesis and a putative S-layer protein gene (data not shown). However, it was not possible to deduce reasons for the acid resistance of the membrane or cell wall by the genome sequence alone.

Replication, Repair, and Restriction. The replication apparatus of *P. torridus* is of the classical archaeal type. An ORF encoding an Orc1/Cdc6 homolog that recognizes the replication origin was found near an ORF for a DNA–helicase that is probably involved in the unwinding of the parental duplex DNA in cooperation with single-strand DNA binding proteins and a topoisomerase. Coding sequences for a two-subunit gyrase were found in the genome, whereas no reverse gyrase gene could be detected. Synthesis of the RNA/DNA primer can be accomplished by a two-subunit primase, and genes coding for all of the DNA–polymerase complex proteins necessary for strand elongation were identified: "clamp loader" and "sliding clamp" polypeptides as well as three DNA polymerases of the DNA polymerase families X, B, and D. RNA primers attached to the 5′ end of the Okazaki fragments can be removed by an endonuclease and a ribonuclease before gap-filling and joining by a ligase.

To ensure DNA integrity, *P. torridus* contains the coding capacity for a large number of repair and recombination proteins. Two repair endonucleases of types III and IV, one repair endonuclease of type V, three repair DNA helicases, two proteins with MutT-like domains, and the repair proteins RadA, RadB, MRE11, and Rad50 are exclusively involved in DNA repair or, in part, play a role in recombination processes, together with a RecJ exonuclease homolog and the topoisomerase and ligase already mentioned. Uncommon in archaea is the deoxyribodipyrimidine photolyase gene, which shows high similarity to bacterial orthologs. Besides a type II restriction/modification system usually found in the genomes of other thermoacidophiles, *P. torridus* also possesses a type I system.

Amino Acid Metabolism. In contrast to *T. acidophilum* and *T. volcanii*, biosynthetic pathways for all 20 amino acids were detected in the *P. torridus* genome. With respect to the utilization of amino acids, a major source of carbon and energy for *Picrophilus*, it has been reported that *Picrophilus oshimae* cells rapidly take up the amino acids histidine, proline, glutamate, and serine, although only glutamate, proline, and leucine were able to drive respiration (2). Genome analysis of *P. torridus* revealed that this organism possesses particular genes and pathways for the degradation of aspartate, glutamate, serine, arginine, histidine, glycine, threonine, and the aromatic amino acids phenylalanine and tyrosine. The breakdown of serine, glycine, and histidine in *P. torridus* requires an operating folate or modified folate C1 metabolism. We found some of the genes for later steps in tetrahydrofolate (THF) biosynthesis as well as all genes needed for the backbone of the one-carbon transfer reactions, from THF and formate to 5.10-methylene THF. The availability of a one-carbon folate pool can greatly enhance the metabolic capacity of an organism because it can be used to gain reducing equivalents by various catabolic reactions and to provide C1 compounds for nucleotide, methionine, and panthotenate biosynthesis (19). In *P. torridus*, most of the genes at the periphery of the one-carbon folate pool have catabolic functions apart from some involved in purine and methionine biosynthesis. No genes for THF-dependent reactions have been found for the synthesis of panthotenate, formylmethionine–tRNA, or thymidine. In most archaea, C1 compounds are carried by modified pterin-containing compounds that are structurally related to folate (20). This also seems to be the case in *P. torridus* because we found ORFs for tetrahydrofolate rather than tetrahydromethanopterin biosynthesis in its genome sequence. Bearing the high number of homologous ORFs between *P. torridus* and *S. solfataricus* in mind (see above), it is noteworthy in this context that a modified folate with structural features common to both methanopterin and folate was identified in *S. solfataricus* ATCC 35091 (21).

Protein and Peptide Degradation. Because *P. torridus* is believed to live as a scavenger (4), peptides and proteins are important growth substrates. Proteins can be degraded by several predicted extracellular acid proteases, including two thermopsin-like proteins and two serine proteases. Most of these proteins possess a putative transmembrane helix at their C-terminal end, which is thought to serve as a membrane anchor. Similar hydrophobic C termini are also observed with exported proteins from other archaea (22) and also seems to be true for other extracellular proteins of *P. torridus*. We found four ATP-binding cassette (ABC) transporters for the uptake of di- and oligopeptides, which can be further degraded to free amino acids by a tricorn peptidase, two tricorn cofactors (F2 and F3), an acylaminoacyl peptidase, a proline dipeptidase, and a metallo-carboxypeptidase.

Protein Synthesis and Export. A large number of ORFs predicted to encode chaperones were found: the Hsp60 system (two thermosome subunits and two prefoldin/Gim subunits), a VAT-protein, a Lon-2-related ATPase, two Hsp20 proteins, and the complete Hsp70 system DnaK, DnaJ, and GrpE, which is found in the euryarchaeal branch of the thermoacidophiles and in methanogens, but is absent from most other archaea (8, 23). Genes required for the twin arginine and signal recognition particle protein export systems can be detected, and a total of 121 putatively exported proteins are predicted by the SIGNALP algorithm. Most of them have been annotated as transporters, exported binding proteins, proteases, components of the respiratory chain, or hypothetical proteins. A total of 38% of the predicted exported proteins possess a C-terminal transmembrane helix that could serve as membrane anchor. Interestingly, in five genes annotated as ABC-transport binding proteins, we were able to detect signal peptides but no means for anchoring the proteins to the cell wall or membrane.

Transporters. The genome of *P. torridus* contains a large number of genes coding for transporters (Fig. 2). A total of 170 ORFs, or 12% of all genes, play a role in transport. A total of 21 transporters are predicted to be involved in drug export. We assume that most of these are required in detoxification of the cell, because we could not detect any genes for secondary metabolite biosynthesis. Uptake

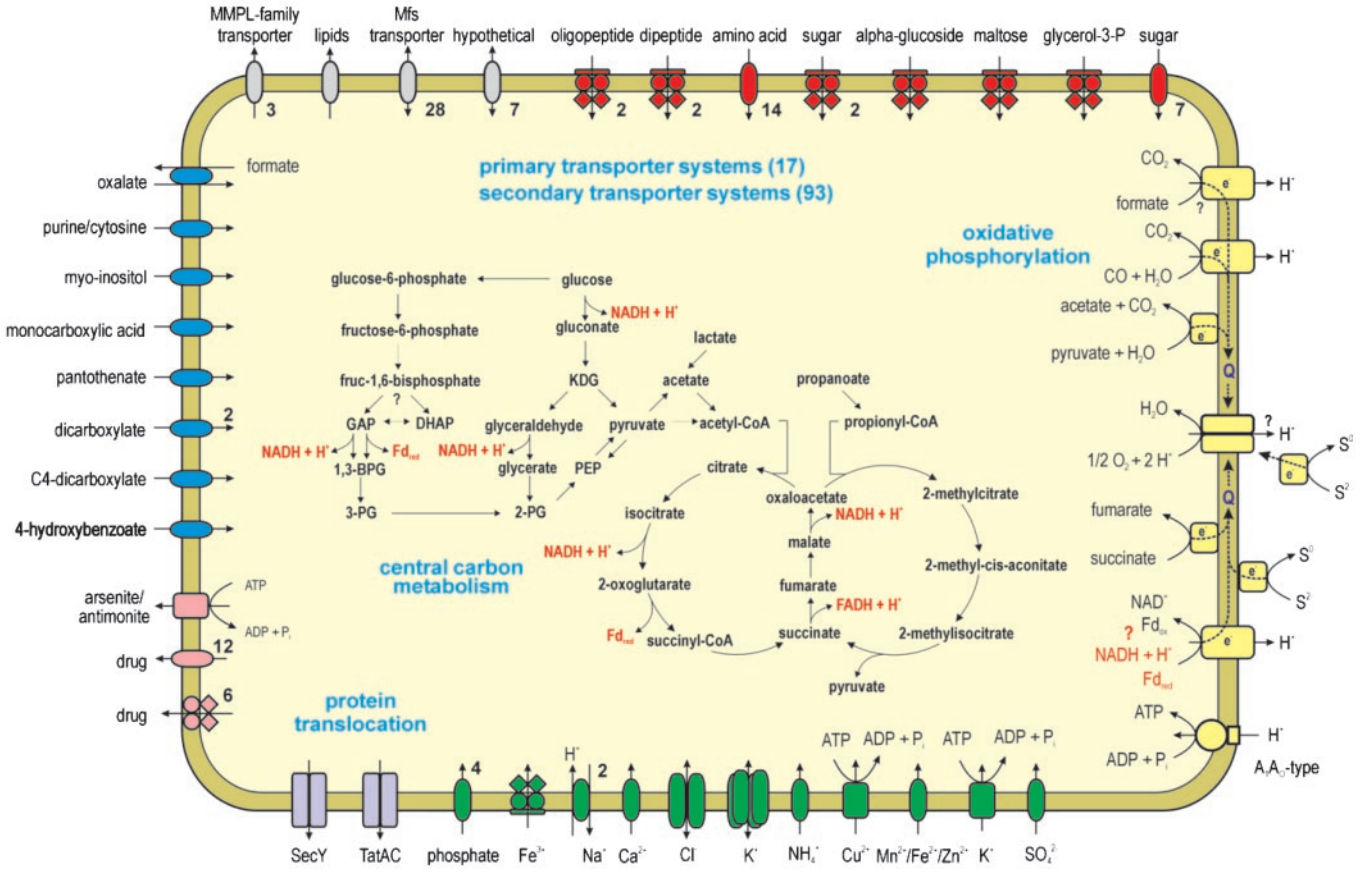

Fig. 3. Overview of the transport, central metabolism, and energy production in *P. torridus*. Sugar, peptide, and amino acid uptake systems are shown in red, drug exporters are shown in pink, trace elements transport systems are shown in green, other and hypothetical transporters are shown in gray. Bold numbers mark the number of each transporter. Protein translocation systems are shown in violet, and the components of the respiratory chain are shown in yellow. A total of 93 secondary and 17 primary transporters were found in the genome sequence, resulting in an unusual ratio of 5.6:1. So far, no aldolase gene is found. Enzyme activity essays indicate a functional nonphosphorylated Entner Doudoroff pathway for glucolysis (unpublished data). Pathways for the respiration of the organic acids acetate, lactate, and propanoate were identified. $NADH_2$ and reduced ferredoxin is produced in the *P. torridus* central carbon metabolism but the final reducing compound of the NADH-oxidoreductase is still unknown as no electron-input module for it was detected.

systems for Fe^{3+}, NH_4^+, Cu^{2+}, Mn^{2+}, Zn^{2+}, SO_4^{2-}, and phosphate were found, as well as two proton/sodium exchangers and transport channels for Cl^- and K^+ (Fig. 3). Besides the K^+ channel, the organism possesses a K^+-transporting ATPase whose functional role is most probably potassium uptake to invert the $\Delta\psi$ to positive inside, which is necessary to cope with the high ΔpH, by counteracting the proton influx and bringing the proton motive force to a range found also in neutrophiles and other acidophiles (3). Also present are several transporters for nucleotides and a number of organic acids. A large number of ORFs seem to be necessary for the uptake of peptides, amino acids (34 ORFs), and sugars (32 ORFs). Nearly half of them code for primary ABC-transporter subunits. Primary transporters of the ABC-transporter type usually comprise five subunits and use the free energy of ATP hydrolysis as energy source. In contrast, active secondary transporters are single subunit proteins and use the transmembrane electrochemical gradient of protons or sodium ions to drive the transport.

We determined the overall ratio of secondary to primary transporters for thermoacidophilic archaea. Interestingly, the ratio for the representatives of the euryarchaeal branch, the thermoplasmales, was unusually high (10:1 and 5.6:1 for *T. acidophilum* and *P. torridus*, respectively) compared to other microorganisms such as *S. solfataricus* (2.7:1), *E. coli* (2.6:1), *Pyrococcus horikoshii* (1.5:1), or *Thermotoga maritima* (0.5:1). Because we could not find any candidate genes for secondary transporters that use Na^+ to drive the transport, we believe that *P. torridus* relies mainly on the high proton motive force to drive its metabolite transport. The large number of ABC-transporter genes for peptide and sugar uptake on the other hand indicates the importance of such compounds as nutrient source, and points to the need of high-affinity transporter systems for the efficient uptake of these substrates.

Importantly, the exceptionally high ratio of secondary to primary solute transport systems found in *Picrophilus* indicates that the predominant use of proton-driven secondary transport represents a highly relevant strategy for the adaptation of this organism to its extremely acidic environment. In contrast, in most hyperthermophilic bacteria and archaea primary uptake systems are preferred (22). This strategy of acidophilic adaptation can also be observed in the genomes of other thermoacidophilic euryarchaea like *T. acidophilum* and *F. acidarmanus*, but not in thermoacidophilic organisms of the crenarchaeota, and thus seems to be a trait only common to the former branch.

Energy Metabolism. Polymeric sugar compounds outside of the cell are subject to enzymatic attack by an extremely acid-stable glucoamylase that has recently been investigated (24). Genes for five ABC transport systems and seven secondary transporters were identified for the uptake of oligomeric and monomeric sugars. The fate of glucose in the cell is less clear. *P. torridus* most likely catabolizes glucose via a nonphosphorylated variant of the Entner–Doudoroff (ED) pathway (Fig. 3), which is usually used by thermoacidophilic archaea (23, 25, 26). Genes for all steps have been

assigned, including a gluconate dehydratase gene, which to date has not been identified in other genomes. The predicted gluconate dehydratase shares high similarity with the predicted galactonate dehydratase genes of enterobacteria and orthologs are clustered with the KDG aldolase gene in *S. solfataricus* and with the glucose dehydrogenase gene in *T. acidophilum*. This finding indicates that a completely promiscious nonphosphorylated ED pathway (27) operates in *P. torridus*.

Except for the fructose-1,6-bisphosphate aldolase, all genes required for the Embden–Meyerhof–Parnas (EMP) pathway are present. In contrast to *T. acidophilum* and *S. solfataricus*, *P. torridus* possesses a phosphofructokinase gene that would be unnecessary unless its reaction product is further cleaved in an aldolase reaction or vice versa in the gluconeogenic orientation of the pathway. We therefore assume that a nonclassical fructose-1,6-bisphosphate aldolase may be present in *P. torridus* and that the EMP pathway is used, at least, for gluconeogenesis.

Pyruvate as the final product of glucose breakdown can be converted to acetyl-CoA by either a NAD^+-dependent pyruvate dehydrogenase or a ferredoxin-dependent pyruvate oxidoreductase. It is unclear at present whether both enzymes operate *in vivo*. It is possible that the ferredoxin-dependent enzyme is used in the reverse direction for anabolic purposes, i.e., for growth on C2-compounds such as acetate, because *P. torridus* has no glyoxylate pathway.

P. torridus appears to contain a complete set of genes for the oxidative tricarboxylic acid (TCA) cycle. In parallel, the organism maintains the coding capacity for the 2-methylcitrate pathway for the oxidation of propionyl–CoA, employing enzymes of the TCA cycle responsible for the conversion of succinate to oxaloacetate. A gene coding for a propionyl–CoA synthase has been detected, enabling *P. torridus* to grow on propionate. Interestingly, it was reported previously that the addition of propionate, lactate, acetate, or formate to *P. oshimae* cells inhibited respiration (3). Because we also found a lactate-2-monooxygenase, which converts lactate and 0.5 O_2 to pyruvate, two acetyl-CoA synthetases, parts of a formate hydrogen lyase operon, and a formyl-tetrahydrofolate synthetase, it is possible that the tested compounds are not metabolized in substantial amounts, but that the enzymes and pathways serve mainly as a means of protection against uncoupling of the respiratory chain by organic acids. It must be noted that the formyl–tetrahydrofolate ligase ORF shows high similarity to bacterial ORFs, and a lactate monooxygenase homolog has so far not been detected in any other archaeal genome. We therefore conclude that at least some of the organic acid metabolic pathways have been acquired by horizontal gene transfer. Because compounds such as acetic or propionic acid function as uncouplers at low pH because of the diffusion of the undissociated form into the cell followed by dissociation, they may be very harmful for extreme thermoacidophiles. Therefore, the acquisition of efficient pathways for the degradation of organic acids may have been important for the evolution of the extreme thermoacidophily found in the *Picrophilus* lineage.

To prevent further acidification of its already low intracellular pH, active respiration is required. In contrast to its close relative, the microaerophilic *T. acidophilum*, *P. torridus* is an obligate aerobe and uses a more complex electron transport chain to generate a membrane potential (Fig. 3). Although no complete set of quinone synthesis genes was identified, many genes for protein complexes that feed electrons into the transport chain were found, including several quinone oxidoreductases. Among the latter are a pyruvate oxidase, a CO dehydrogenase, the formate lyase complex mentioned above, and a sulfide–quinone oxidoreductase. In addition, a gene for a putative sulfide dehydrogenase, which transfers the electrons directly to cytochromes, was found.

All type I NADH oxidoreductase-homologous genes of *Paracoccus denitrificans* are found in *P. torridus* except the ones coding for the electron input module, *nuoEFG*. The same is true for the complex I of *T. acidophilum* and *F. acidarmanus*, but not for members of the *Sulfolobales* that lack some of the integral membrane and electron transfer subunits. We therefore assume that the complex I in thermoacidophiles of the euryarchaeota is able to transfer protons over the cytoplasmic membrane. However, most archaea do not contain genes for homologs of the subunits of the electron input module, which oxidizes NADH and subsequently channels the electrons to the membrane-associated quinone reductase module. It is still unknown how electrons are fed into the transport chain in organisms without the NADH input module and whether NAD^+ or ferredoxin is the electron mediator between metabolism and the electron transport chain (28, 29)

Quinol oxidation in *P. torridus* is accomplished by a complex similar to the SoxM complex, which has been described in *Sulfolobus acidocaldarius* (30). This complex consists of the quinone oxidase and the terminal oxidase with a blue copper protein (sulfocyanin) as electron shuttle between the two. Although we could find homologs of the quinone oxidase in the *T. acidophilum* and *S. solfataricus* genomes, significant similarities of the terminal oxidase part are only found with bacterial cytochrome *c* oxidases. Genes similar to the blue copper protein gene of *P. torridus* were only found in the *Sulfolobus* and *Ferroplasma* genomes. Thus, based on the detection of genes for a putative proton pumping NADH dehydrogenase (complex I) as well as those for quinol and cytochrome oxidation and those for an A_OA_1-type ATPase, it can be concluded that *P. torridus* obviously has an efficient respiratory system that accomplishes the expulsion of protons with a rate sufficiently high to prevent overacidification of the cytoplasm. In addition, the proton gradient upheld in this way can be extensively exploited for solute transport, as mentioned above. Because several components of the respiratory chain of *P. torridus* are by far most similar to genes from organisms of the distant crenarchaea or bacteria but are not found in members of the *Thermoplasmaceae*, we assume that these were obtained relatively late in evolution by horizontal gene transfer.

Porphyrin Metabolism. *P. torridus* seems to be able to synthesize porphyrins like cytochromes and adenosylcobalamin. We found 28 genes that are required to synthesize adenosyl-cobalamin from glutamate, which represents nearly 2% of the whole genome. Synthesis starts with L-glutamate and proceeds via glutamate-1-semialdehyde and aminolevulinate to uroporphyrinogen III, the last common precursor of the porphyrins. For cobalamin biosynthesis, uroporphyrinogen III is converted to precorrin 2 in which either Co^{2+} is inserted in organisms employing the anaerobic route yielding cobalt-precorrin 2, or which undergoes a methylation step followed by an oxygen-dependent ring contraction yielding precorrin 3 in the aerobic pathway. Organisms using the anaerobic pathway are known to possess type II, or "early" metal ion chelatases for cobalt insertion (31). *P. torridus* contains no type II metal chelatase-homologous genes and, in contrast to *Thermoplasma* and *Sulfolobus*, seems to employ the aerobic pathway, where the insertion of the cobalt ion takes place after the main modifications of the porphyrin ring system, and is catalyzed by a type I or "late" metal ion chelatase protein complex similar to CobNST of *Pseudomonas denitrificans*. However, most enzymes of the porphyrin biosynthesis of *P. torridus* revealed their highest similarity scores with sequences deduced from the porphyrin biosynthesis genes of other thermoacidophilic archaea that employ the anaerobic pathway, thus indicating that the late metal ion chelatase complex was acquired later through horizontal gene transfer.

We did not find *cob*S or *cob*T homologs in *P. torridus*, but we did find two genes homologous to magnesium chelatase subunits *chl*I and *chl*D flanking the *cob*N gene. Comparative studies with other prokaryotes (data not shown) showed that this is not uncommon, and it has recently been suggested that ChlI and ChlD may take over the function of CobS and CobT (32).

Coping with Oxygen Stress. As a strict aerobic organism, *P. torridus* possesses several mechanisms to protect the cell against oxidative damage. We found genes coding for a superoxide dismutase, three putative peroxiredoxin-like proteins, and an alkyl hydroperoxide reductase that is present only in thermoacidophilic archaea and *Pyrococcus furiosus*. Flanking the predicted OriR region (supporting information), a β-carotene biosynthetic operon strongly resembling genes from marine ε-proteobacteria and corynebacteria was detected. β-Carotene formation in other archaea has only been predicted for *S. solfataricus* and biochemically studied in *Halobacteria* (33, 34).

Conclusions

P. torridus lives in an extremely hostile environment and grows at the lowest pH values known among all organisms. Even adaptation to conditions such as in 1 M sulfuric acid is possible (4). Points of major interest include questions about properties and mechanisms which ensure viability under these conditions. Specific adaptation mechanisms can be expected at the levels of (*i*) structure and function of biomolecules and (sub)cellular structures, (*ii*) physiology and metabolic features, and (*iii*) regulation of gene expression. For example, *P. torridus* cells need a specifically adapted membrane. Indeed, the membrane of *P. torridus* displays a very low proton permeability, is highly acid stable, and loses its integrity when incubated at pH 7 (2). Furthermore, an acid-stable cell envelope and acid-resistant extracellular enzymes for the degradation of polymeric or oligomeric carbon sources are required. Finally, the organism needs to generate metabolic energy in substantial amounts to maintain the intracellular pH at an acceptable value.

To prevent further acidification of its already low intracellular pH, active respiration is required. Genes for a putative proton pumping NADH degydrogenase (complex I) have been detected, as have those for quinol and cytochrome oxidation. Genes for an A_OA_1-type ATPase were also present. This organism obviously accomplishes the expulsion of protons with the necessary rate, and the proton gradient upheld in this way can be extensively exploited for solute transport, as is evident from the large number of secondary transport systems.

Organic acids, such as acetic or propionic acid, function as uncouplers at low pH because of the diffusion of the undissociated form into the cell followed by dissociation. Genes encoding enzymes for the degradation of these compounds as soon as they enter the cell are present.

It is important to note that many genes that enhanced the abilities of *P. torridus* to cope with its extremely acidic environment have been obtained by horizontal gene transfer. This includes some of the organic acid degradation pathways, the main components of the electron transport chain, and mechanisms to deal with oxygen stress.

P. torridus has the smallest genome of a nonparasitic free living organism. The genomes of thermophilic methanogenic archaea and hyperthermophilic bacteria are not much larger ($\approx$1.6–1.8 megabases). These organisms, which all grow at higher temperatures than *P. torridus*, have to cope with a temperature-dependent increased error rate in their nucleic acids due to cytosine deamination (35), which amongst other factors may have led to selective pressure favoring small genomes. *P. torridus*, albeit being only moderately thermophilic (T_{opt} 60°C), in addition must cope with pH values around 0 in the medium and an extremely low intracellular pH. Although the consequences for cell physiology and DNA integrity are not yet understood, we believe that the combination of two extremophilic conditions, low pH and high temperature, may have led to the small size and one of the highest coding densities for the genome of *P. torridus*.

We thank Mechthild Bömeke for excellent technical support and Hans-Peter Klenk for helpful discussions. This work was supported by the Bundesministerium für Bildung, Wissenschaft, Forschung und Technologie within the framework Genomforschung an Bakterien für die Analyse der Biodiversität und die Nutzung zur Entwicklung neuer Produktionsverfahren.

1. Schleper, C., Pühler, G., Kühlmorgen, B. & Zillig, W. (1995) *Nature* **375,** 741–742.
2. van de Vossenberg, Driessen, A. J., Zillig, W. & Konings, W. N. (1998) *Extremophiles* **2,** 67–74.
3. She, Q., Singh, R. K., Confalonieri, F., Zivanovic, Y., Allard, G., Awayez, M. J., Chan-Weiher, C. C., Clausen, I. G., Curtis, B. A., De Moors, A., *et al.* (2001) *Proc. Natl. Acad. Sci. USA* **14,** 7835–7840.
4. Schleper, C., Puehler, G., Holz, I., Gambacorta, A., Janekovic, D., Santarius, U., Klenk, H.-P. & Zillig, W. (1995) *J. Bacteriol.* **177,** 7050–7059.
5. Gonzalez-Toril, E., Llobet-Brossa, E., Casamayor, E. O., Amann, R. & Amils, R. (2003) *Appl. Environ. Microbiol.* **69,** 4853–4865.
6. Okibe, N., Gericke, M., Hallberg, K. B. & Johnson, D. B. (2003) *Appl. Environ. Microbiol.* **69,** 1936–1943.
7. DeLong, E. F. (2000) *Genome Biol*. **1,** 1029.1–1029.3.
8. Deppenmeier, U., Johann, A., Hartsch, T., Merkl, R., Schmitz, R. A., Martinez-Arias, R., Henne, A., Wiezer, A., Baumer, S., Jacobi, C., *et al.* (2002) *J. Mol. Microbiol. Biotechnol.* **4,** 453–461.
9. Ewing, B., Hillier, L., Wendl, M. C. & Green, P. (1998) *Genome Res.* **8,** 175–185.
10. Staden, R., Beal, K. F. & Bonfield, J. K. (2000) *Methods Mol. Biol.* **132,** 115–130.
11. Tech, M. & Merkl, R. (2003) *In Silico Biol.* **3,** 441–451.
12. Badger, J. H. & Olsen, G. J. (1999) *Mol. Biol. Evol.* **16,** 512–524.
13. Frishman, D., Mironov, A., Mewes, H.-W. & Gelfand, M. (1998) *Nucleic Acids Res.* **26,** 2941–2947.
14. Delcher, A., Harmon, D., Kasif, S., White, O. & Salzberg, S. (1999) *Nucleic Acids Res.* **27,** 4636–4641.
15. Altschul, S. F., Gish, W., Miller, W., Myers, E. W. & Lipman, D. J. (1990) *Mol. Biol.* **215,** 403–410.
16. Nielsen, H., Engelbrecht, J., Brunak, S. & von Heijne, G. (1997) *Protein Eng.* **10,** 1–6.
17. Moll, R. & Schafer, G. (1988) *FEBS Lett.* **232,** 359–363.
18. Schafer, K., Magnusson, U., Scheffel, F., Schiefner, A., Sandgren, M. O., Diederichs, K., Welte, W., Hulsmann, A., Schneider, E. & Mowbray, S. L. (2004) *J. Mol. Biol.* **335,** 261–274.
19. Edward, B. & Maden, H. (2000) *Biochem. J.* **350,** 609–629.
20. Angelaccio, S., Chiaraluce, R., Consalvi, V., Buchenau, B., Giangiacomo, L., Bossa, F. & Contestabile, R. (2003) *J. Biol. Chem.* **278,** 41789–41797.
21. Zhou, D. & White, R. H. (1992) *J. Bacteriol.* **174,** 4576–4582.
22. Albers, S. V., Van de Vossenberg, J. L., Driessen, A. J. & Konings W. N. (2001) *Extremophiles* **5,** 285–294.
23. Ruepp. A., Graml, W., Santos-Martinez, M.-L., Koretke, K. K., Volker, C., Mewes, H. W., Frishman, D., Stocker, S., Lupas, A. N. & Baumeister, W. (2000) *Nature* **407,** 508–513.
24. Serour, E. & Antranikian, G. (2002) *Antonie Leeuwenhoek* **81,** 73–83.
25. Selig, M., Xavier, K. B., Santos, H. & Schonheit, P. (1997) *Arch. Microbiol.* **167,** 217–232.
26. Budgen, N. & Danson, M. J. (1986) *FEBS Lett.* **196,** 207–210.
27. Lamble, H. J., Heyer, N. I., Bull, S. D., Hough, D. W. & Danson, M. J. (2003) *J. Biol. Chem.* **278,** 34066–34072.
28. Schäfer, G., Engelhard, M. & Müller, V. (1999) *Microbiol. Mol. Biol. Rev.* **63,** 570–620.
29. Deppenmeier, U. (2002) *Cell. Mol. Life. Sci.* **59,** 1513–1533.
30. Lubben, M., Arnaud, S., Castresana, J., Warne, A., Albracht, S. P. & Saraste, M. (1994) *Eur. J. Biochem.* **224,** 151–159.
31. Scott, A. I. & Roessner, C. A. (2002) *Biochem. Soc. Trans.* **30,** 613–620.
32. Rodionov, D. A., Vitreschak, A. G., Mironov, A. A. & Gelfand, M. S. (2003) *J. Biol. Chem.* **278,** 41148–41159.
33. Hemmi, H., Ikejiri, S., Nakayama, T. & Nishino, T. (2003) *Biochem. Biophys. Res. Commun.* **305,** 586–591.
34. Spudich, J. L., Yang, C. S., Jung, K. H. & Spudich, E. N. (2000) *Annu. Rev. Cell. Dev. Biol.* **16,** 365–392.
35. Wang, R. Y., Kuo, K. C., Gehrke, C. W., Huang, L. H. & Ehrlich, M. (1982) *Biochim. Biophys. Acta.* **697,** 371–377.