Extremophiles (1998) 2:101–108

ORIGINAL PAPER

Christian Baumann · Martin Judex · Harald Huber
Reinhard Wirth

Estimation of genome sizes of hyperthermophiles

Received: August 10, 1997 / Accepted: January 1, 1998

Abstract Genomes of various hyperthermophilic and extremely thermophilic prokaryotes were analyzed with respect to size, physical organization, and 16S rDNA copy number. Our results show that all the genomes are circular, and they are in the size range of 1.6–1.8 Mb for *Pyrodictium abyssi, Methanococcus igneus, Pyrobaculum aerophilum, Archaeoglobus fulgidus, Archaeoglobus lithotrophicus*, and *Archaeoglobus profundus* (the two bacteria *Fervidobacterium islandicum* and *Thermosipho africanus* possess genomes of 1.5-Mb size). A systematic study of all validly described species of the order *Sulfolobales* revealed the existence of two classes of genome size for these archaea, correlating with phylogenetic analyses. The *Metallosphaera–Acidianus* group, plus *Sulfolobus metallicus*, have genomes of ca. 1.9 Mb; the other members of the order *Sulfolobales* group possess genomes >2.7 Mb. The special case of *Stygiolobus azoricus* is discussed.

Key words Hyperthermophilic archaea · Genome sizes · *Sulfolobales* · Pulsed field gel electrophoresis · Thermophilic bacteria · Phylogenetic analysis

Introduction

Hyperthermophilic prokaryotes comprise the shortest branches in the phylogenetic tree of life (Stetter 1996). Therefore, these slowest evolving forms of life are thought to represent most closely the common ancestor of all present day forms of life (Woese and Fox 1977). By studying hyperthermophilic prokaryotes with respect to their biochemistry, their structural and genetic organization, their phylogenetic relationships, etc., one might obtain at least hints as to the respective features of more primitive forms of life which are no longer present on earth. In particular, the very recent explosion in knowledge of genomics [see for example Hieter and Boguski (1997) for expected developments], which is expected to continue for the next few years, could add to our understanding of primitive live.

The year 1997 has been called the year of the genome (Ash 1997), emphasizing the fact that by January 1997, the complete genome sequences of six microorganisms had been reported and at least another eight were expected to be published. Indeed, in September 1997, a total of 11 prokaryotic genome sequences were reported to be available, including the long-awaited *Escherichia coli* genome (Pennisi 1997); three of these 11 sequences are from members of the archaea, namely *Archaeoglobus fulgidus, Methanobacterium thermoautotrophicum*, and *Methanococcus jannaschii*. According to various authors, some 50–100 genome sequences are in the pipeline. Thus, an enormous wealth of information will become available over the next few years; nevertheless, the number of unassigned genes and open reading frames (ORFs) will very probably increase exponentially, too. Therefore, decisions on which genomes of archaea (but of course also of all other organisms) should be chosen for further sequencing projects should be taken very carefully; it is expected that the data reported here will provide a helpful basis for such decisions.

We have analyzed genomes of hyperthermophilic and extremely thermophilic archaea and bacteria with respect to their size, their structural organization, and their content of 16S rDNA genes. One rationale behind these experiments is that archaeal genomes are quite often referred to as being small, but that actual data for many of these organisms simply are not available. Though we were able to use standard methods to isolate DNA for Southern hybridizations with 16S rDNA probes, special protocols had to be developed for some species to obtain DNA restrictable in agarose plugs for pulsed field gel electrophoresis (PFGE).

Communicated by K. Horikoshi

C. Baumann · M. Judex · H. Huber · R. Wirth (✉)
Lehrstuhl für Mikrobiologie und Archaeenzentrum, Universität Regensburg, Universitätsstraße 31, D-93053 Regensburg, Germany
Tel. +49-941-943-1825; Fax +49-941-943-1824
e-mail: Reinhard.Wirth@biologie.uni-regensburg.de

Table 1. Hyperthermophilic microorganisms used in this study and their growth conditions

Species	Strain	DSM collection[#]	Reference	Growth medium[a]	Gas phase	Cultivation temperature
Acidianus ambivalens	Lei10	DSM 3772	Fuchs et al. 1996	Allen + 1% S^0	Air	85°C
Acidianus brierleyi	SP3a/1	DSM 1651	Segerer et al. 1986	Allen + 1% S^0	Air	65°C
Acidianus infernus	So4a	DSM 3191	Segerer et al. 1986	Allen + 1% S^0	Air	85°C
Archaeoglobus fulgidus	VC 16	DSM 4304	Stetter 1988	MGL	H_2/CO_2	85°C
Archaeoglobus lithotrophicus	TF 2	–	Unpublished	BSM	H_2/CO_2	85°C
Archaeoglobus profundus	AV 18	DSM 5631	Burggraf et al. 1990a	MGG	H_2/CO_2	85°C
Fervidobacterium islandicum	H 21 B	DSM 5733	Huber R. et al. 1990	ATCC 35602	N_2	75°C
Metallosphaera prunae	Ron 12/II	DSM 10039	Fuchs et al. 1995	Allen + 0.1% YE	Air	65°C
Metallosphaera sedula	TH2	DSM 5348	Huber G. et al. 1989	Allen + 0.1% YE	Air	65°C
Methanococcus igneus	KOL 5	DSM 5666	Burggraf et al. 1990b	MGG	H_2/CO_2	85°C
Methanopyrus kandleri	AV 19	DSM 6324	Kurr et al. 1991	BSM	H_2/CO_2	100°C
Methanothermus fervidus	V 24 S	DSM 2088	Stetter et al. 1981	MS	H_2/CO_2	85°C
Pyrobaculum aerophilum	IM 2	DSM 7523	Völkl et al. 1993	BS	H_2/CO_2	95°C
Pyrodictium abyssi	AV 2	DSM 6158	Pley et al. 1991	$^1/_2$ SME*	H_2/CO_2	100°C
Sulfolobus acidocaldarius	98-3	DSM 639	Brock et al. 1972	Allen + 0.1% YE	Air	75°C
Sulfolobus metallicus	KRA23	DSM 6482	Huber and Stetter 1991	Allen + 1% S^0	Air	65°C
Sulfolobus shibatae	B12	DSM 5389	Grogan et al. 1990	Allen + 0.1% YE	Air	70°C
Sulfolobus solfataricus	P1	DSM 1616	Zillig et al. 1980	Allen + 0.1% YE	Air	65°C
Sulfolobus solfataricus	Ron 12/III	–	Fuchs 1994	Allen + 0.1% YE	Air	85°C
Stygiolobus azoricus	FC6	DSM 6296	Segerer et al. 1991	Allen + 1% S^0 + 0.02% YE	H_2/CO_2	85°C
Thermosipho africanus	OB 7	DSM 5309	Huber R. et al. 1989	MG	N_2	75°C

[a] A detailed description of the media used is given in the following references: Allen, Allen (1959), as modified by Brock et al. (1972); BS, Völkl et al. (1993); BSM, Kurr et al. (1991); MG, Balch et al. (1979) (medium 3); MGG, Balch et al. (1979), as modified by Huber et al. (1982); MGL, Balch et al. (1979); as modified by Stetter (1988); MS, Balch et al. (1979) (medium 1), as modified by Stetter et al. (1981); $^1/_2$ SME*, Stetter et al. (1983); YE, Yeast extract; S^0, colloidal sulfur; ATCC, American Type Culture Collection (Rockville, MD, USA).

Materials and methods

Strains and growth conditions

The microorganisms used in this study, together with growth conditions and references, are listed in Table 1. Cells were collected from growth media by centrifugation, washed once in ice-cold sterile growth medium, and suspended to ca. 4×10^8 cells/ml for further analysis by PFGE. In cases where growth media contained sulfur, it had to be removed by filtration (3 μm pore size) prior to cell collection, because sulfur interfered with the subsequent steps.

PFGE

A commercial PFGE apparatus (Bio-Rad Chef DR III system, with the supplier's algorithm for calculating optimal run conditions; Bio-Rad, Hercules, CA, USA) was used. PFGE was carried out in 1% FastLane agarose (FMC, Rockland, ME, USA) using the "optimized buffer system" (i.e., 0.25 × TBE[1]), allowing a 20% reduction in calculated run and switch times. Commercially available products (yeast chromosomes, lambda ladder, and 5-kb ladder; from New England Biolabs, Beverly, MA, USA) were used as size markers.

Lysis of cells in agarose plugs to obtain genomic DNA susceptible to restriction enzymes was carried out by five different methods. In each case, ca. 2×10^8 cells were embedded per 100 μl plug, which consisted of 0.8% Incert agarose (FMC) in 1 M NaCl + 10 mM Tris/Cl buffer (pH 7.5, except for species of *Sulfolobales*, for which pH 8 had to be used). Lysis method 1 followed the protocol by Smith and Cantor (1987), which in short consists of the following steps: treatment with EC lysis buffer [containing ethylenediaminetetraacetic acid (EDTA), Brij 58, deoxycholate and *N*-lauroylsarcosine], followed by extensive proteinase K treatment in ESP buffer (0.5 M EDTA, *N*-lauroylsarcosine) at 50°C. Lysozyme, however, was not included in the EC lysis buffer and the final wash steps were modified according to Noll (1989) [i.e., including 0.1% phenylmethylsulfonyl fluoride (PMSF) in step 2]. Lysis method 2 consisted of a lysis step with detergents, followed by treatment with the thermostable protease PreTaq, according to the supplier's (Boehringer Mannheim, Mannheim, Germany) instructions. Lysis method 3 was as method 1; preceding the restriction digestion, however, the following washing steps were included: overnight in 3 M KCl, 3 h each in 3 M NaCl, 1 M NaCl, and 0.3 M NaCl (all at 50°C), followed by a 30-min incubation at 37°C in 10 mM Tris/Cl, pH 7.5, 1 mM EDTA (TE-buffer), an overnight wash step at 50°C in ESP buffer, and finally three washings in TE (with wash step 2 containing PMSF). Lysis method 4 was as method 3; cell growth, however, was in media containing subinhibitory amounts of glycine [0.15% glycine for *Methanopyrus kandleri* or 0.3% glycine for

[1] Tris-borate-EDTH; Maniatis et al. 1982

Methanothermus fervidus; see König (1985) for rationale]. In addition, cells collected after growth were subjected to up to 10 freeze/thawing cycles (−20°C/37°C) before being embedded in Incert agarose. The addition of nisin at 1 mg/ml to the EC lysis buffer did not result in improved cell lysis. Lysis method 5 was as method 1, but omitting the EC lysis buffer treatment. Other procedures sometimes used for DNA preparation (such as phenol extractions or solubilization of proteins in chaotropic reagents) could not be used, since they dissolved the agarose plugs.

Southern hybridizations

Genomic DNA of all species listed in Table 1 was prepared by conventional methods (involving phenol extractions); single or double digests obtained with restriction enzymes *Bam*HI, *Cla*I, and *Pst*I were separated on 0.7% agarose gels and blotted to Hybond Nfp membranes (Amersham-Buchler, Braunschweig, Germany). Southern hybridizations to determine the number of rDNA operons were performed by standard methods (Maniatis et al. 1982), using DNA probes derived from genomic DNA of *Pyrodictium abyssi* and of *Sulfolobus solfataricus*. A 16S-rDNA-specific fragment was amplified by polymerase chain reaction (PCR) with the aid of primers CUACUACUACUATCYGGTTGATCCTGCC (forward primer; position 23 *E. coli* nomenclature) and CAUCAUCAUCAUACGGHTACCTTGTTACGACTT (reverse primer; position 1390 *E. coli* nomenclature) and cloned by use of the CloneAmp system (Gibco-BRL, Eggenstein, Germany). Detection of 16S-rDNA-hybridizing genomic DNA fragments was by use of the gel-purified insert which was labelled either by nick-translation (Maniatis et al. 1982) or by use of the ECL direct nucleic acid labelling and detection kit from Amersham-Buchler.

Results and discussion

Genome organization and 16S rDNA genes

The genomes of all (hyper)thermophilic microorganisms analyzed by us turned out to be circular, since in no case could a DNA band entering the gel during PFGE be detected if the genomic DNA was not treated with restriction enzymes (data not shown). In all cases, we tested for successful cell-lysis, by staining DNA inside the agarose plugs with ethidium bromide. Indeed, the degree of cell-lysis could be checked via the intensity of ethidium bromide fluorescence of stained agarose plugs. This fluorescence staining also was helpful for the development of lysis procedures for (hyper)thermophilic archaeal cells in agarose plugs.

Nearly all species tested possess only one 16S rDNA gene. The two exceptions to this turned out to be the thermophilic bacterium *Fervidobacterium islandicum*, and the hyperthermophilic archaeum *Methanococcus igneus*, which both have two 16S rDNA genes (located on DNA fragments also containing 23S rDNA genes; data not shown). Our initial attempts to analyze archaeal genomes for rDNA copy number by use of restriction enzyme I-*CeuI* – recognizing a 19-bp sequence within the 23S rDNA gene (Marshall and Lemieux 1992) – were not successful (data not shown); this recognition sequence is not contained within the archaeal 23S rDNA genes available to us. Our results are corroborated by other data. The presence of only one gene for 16S and 23S rRNA had been reported already for *Acidianus brierleyi, Acidianus infernus, Sulfolobus acidocaldarius, S. solfataricus, Sulfolobus shibatae*, and *Stygiolobus azoricus* (Trevisanato et al. 1996). For *Methanococcus jannaschii*, being a very close relative of *M. igneus*, the presence of two rDNA operons and a third unlinked 5S rRNA gene could be deduced from the complete genome sequence (Bult et al. 1996).

It has been argued that the presence of multiple rDNA operons in prokaryotic genomes enables their fast growth. This is very probably not correct, since, e.g., *Escherichia coli* possesses seven rDNA operons, at least three of which could be deleted without any effect on growth rate. Rather, the presence of multiple rDNA operons seems to be at least a prerequisite for *E. coli* to adjust quickly to altered growth conditions (Condon et al. 1993). It is not surprising, then, that archaea can grow with doubling times of less than one hour and possess only one (or two) rDNA operons.

Genome sizes

Table 2 lists the genome sizes determined by PFGE, and the lysis method used. In all cases experiments were done at least in duplicate; the fragment sizes listed in Table 2 are the mean values of all experiments. One example of the result of a PFGE experiment is given in Fig. 1 (lane 2, *Archaeoglobus profundus* × *Bss*HII; lane 4, *Sulfolobus metallicus* × *Bss*HII); in many cases an extensive smearing (resulting in pictures not acceptable for reproduction) was observed in the lanes. This smearing was very probably due to the presence of proteins complexed with genomic DNA, because it could be reduced to some extent by excessive protease treatments. For some species, only lysis method 3, including intensive high-salt washing steps, resulted in genomes susceptible to restriction enzymes or enabled us to reduce the smearing to an acceptable degree; in those cases, data were collected up to eight times to obtain reliable results. In particular, repeated experiments had to be performed for cases in which a DNA band indicative of a very big fragment was observed. By using agarose plugs containing up to 5×10^8 cells and excessive restriction digestion (100 U per plug in three successive digestions), those cases could be clarified. One example is shown in lane 4 of Fig. 1: the uppermost band for *S. metallicus* × *Bss*HII, visible in lane 4, disappeared in such an experiment; on the other hand, no indication of the presence of two bands with sizes of ca. 600 kb was observed after prolonged restriction. In cases in which two or more bands of similar size were observed, special run conditions – resulting in enhanced resolution in the corresponding size range – were used to resolve these ambiguities.

Table 2. Genome sizes, lysis methods, and DNA fragment sizes for the (hyper)thermophilic microorganisms used in this study

Species	Genome Size	Lysis method[a]	Sizes of DNA fragments[b]
Acidianus ambivalens	1855	5	*Bss*HII: 505; 495; 290; 230; 200; 135; (1855 kb)
Acidianus brierleyi	1880	2	*Eag*I: 430; 330; 190; 170; 160; 135; 130; 110; 105; 75; 45; (1880 kb)
Acidianus infernus	1829	5	*Bss*HII: 500; 375; 235; 210; 195; 140; 94; 80; (1829 kb)
Archaeoglobus fulgidus	1784	1	*Bss*HII: 240; 230; 180; 165; 145; 140; 135; 127; 117; 82; 65; 62; 59; 37; (1784 kb) *Sfi*I: 400; 270; 217; 180; 165; 127; 117; 110; 95; 75; 28; (1784 kb)
Archaeoglobus lithotrophicus	1891	1	*Sfi*I: 360; 250; 200; 170; 157; 115; 110; 107; 90; 76; 68; 45; 42; 37; 35; 29; (1891 kb)
Archaeoglobus profundus	1813	3	*Bss*HII: 470; 335; 295; 220; 175; 90; 88; 72; 68; (1813 kb)
Fervidobacterium islandicum	1535	3	*Not*I: 540; 320; 255; 245; 180; (1540 kb) *Sfi*I: 630; 610; 150; 90; 50; (1530 kb)
Metallosphaera prunae	1879	1	*Bss*HII: 260; 255; 245; 190; 180; 175; 165; 150; 65; 57; 40; 35; 22; 18; 12; 10; (1879 kb)
Metallosphaera sedula	1890	1	*Bss*HII: 260; 255; 245; 185; 183; 180; 165; 145; 78; 57; 40; 35; 22; 18; 12; 10; (1890 kb)
Methanococcus igneus	1658	2	*Eag*I: 350; 335; 250; 245; 200; 190; 80; (1650 kb) *Sac*II: 400; 360; 300; 200; 160; 90; 60; 50; 45; (1665 kb)
Pyrobaculum aerophilum	1709	3	*Asc*I: 270; 230; 210; 180; 155; 130; 105; 85; 65; 60; 55; 45; 40; 32; 25; 22; (1709 kb)
Pyrodictium abyssi	1627	3	*Swa*I: 375; 365; 310; 185; 150; 145; 50; 47; (1627 kb)
Sulfolobus acidocaldarius	2760[c]	n.a.[c]	n.a.[c]
Sulfolobus metallicus	1932	2	*Bss*HII: 600; 590; 250; 140; 135; 95; 45; 32; 24; 21; (1932 kb)
Sulfolobus shibatae	3010	5	*Bss*HII: 600; 550; 460; 450; 370; 180; 140; 135; 125; (3010 kb)
Sulfolobus solfataricus	2795	1	*Eag*I: 445; 280; 270; 260; 225; 200; 187; 175; 150; 140; 110; 90; 82; 60; 54; 45; 39; (2812 kb) *Ksp*I: 450; 320; 310; 290; 200; 190; 155; 140; 130; 120; 105; 95; 90; 60; 55; 40; 35; 30; 10; (2825 kb) *Bss*HII: 400; 330; 300; 230; 205; 185; 165; 145; 125; 115; 105; 100; 72; 62; 55; 45; 44; 38; 25; (2746 kb)
Methanopyrus kandleri	n.d.	4	n.d.
Methanothermus fervidus	n.d.	4	n.d.
Sulfolobus solfataricus Ron 12/III	2705	1	*Eag*I: 460; 435; 260; 240; 210; 190; 170; 168; 157; 107; 85; 80; 54; 50; 35; (2701 kb) *Ksp*I: 330; 305; 300; 215; 200; 195; 160; 155; 150; 135; 125; 100; 95; 80; 65; 60; 40; (2710 kb)
Stygiolobus azoricus	1543	2	*Eag*I: 340; 180; 160; 155; 150; 130; 125; 80; 70; 56; 55; 45; (1546 kb) *Ksp*I: 210; 190; 150; 120; 110; 100; 95; 75; 72; 62; 60; 58; 57; 55; 53; 35; 21; 18; (1541 kb)
Thermosipho africanus	1550	1	*Bss*HII: 290; 260; 200; 170; 115; 105; 95; 80; 60; 50; 45; 40; 30; 15; 10; (1565 kb) *Eag*I: 410; 390; 335; 235; 160; (1530 kb)

n.d., No data available since cells were lysed, but no fragments obtained (see text); n.a., not applicable.
[a] Lysis methods were as described in Materials and Methods.
[b] Data are given in the following way: restriction enzyme used: resolved fragments in kb; (sum of resolved fragments).
[c] For *S. acidocaldarius*, data from Kondo et al. (1993) are cited.

PFGE methodology and accuracy of genome size determinations

Separation of genomic DNA fragments could be performed by conventional PFGE techniques. However, we had to use specially developed methods to obtain these fragments. This very probably is due to two factors:

1. Archaea possess cell walls of very rigid structure, which in many cases cannot be lysed by conventional techniques (see, e.g., Kandler and König 1993). In some cases their crystalline surface layers will not even be denatured by prolonged treatment at 100°C in 1% sodium dodecyl sulfate (SDS). Therefore, only extensive treatment with proteinase K at 50°C in 1% *N*-lauroylsarcosine resulted in cell-lysis. In the case of *Methanopyrus kandleri* and *Methanothermus fervidus* we were able to lyse cells (indicated by ethidium bromide fluorescence of agarose plugs treated with lysis procedure 4); their genomic DNA, however, was not susceptible to any restriction enzyme tested (including some active at 65°C). It therefore remains open whether for these two archaea – which have pseudomurein in their cell walls – the lysis procedure was still not efficient enough to generate holes big enough for restriction enzymes to diffuse inside the lysed cells or for DNA fragments to move out of the cells during electrophoresis.
2. Archaea contain proteins with much stronger DNA-binding capacity than do bacteria. In the case of *Pyrococcus furiosus*, such proteins protect genomic DNA 20-fold better against thermal damage compared with *E. coli* (Peak et al. 1995). For the archaeal isolate AN1 – a member of the *Thermococcales* – the histone-like protein HAN1 has been shown to compact DNA, thereby protecting it from thermal denaturation and cleavage by the restriction enzyme *TaqI* at 65°C (Ronimus and Musgrave 1996). Such DNA-binding proteins are major components of

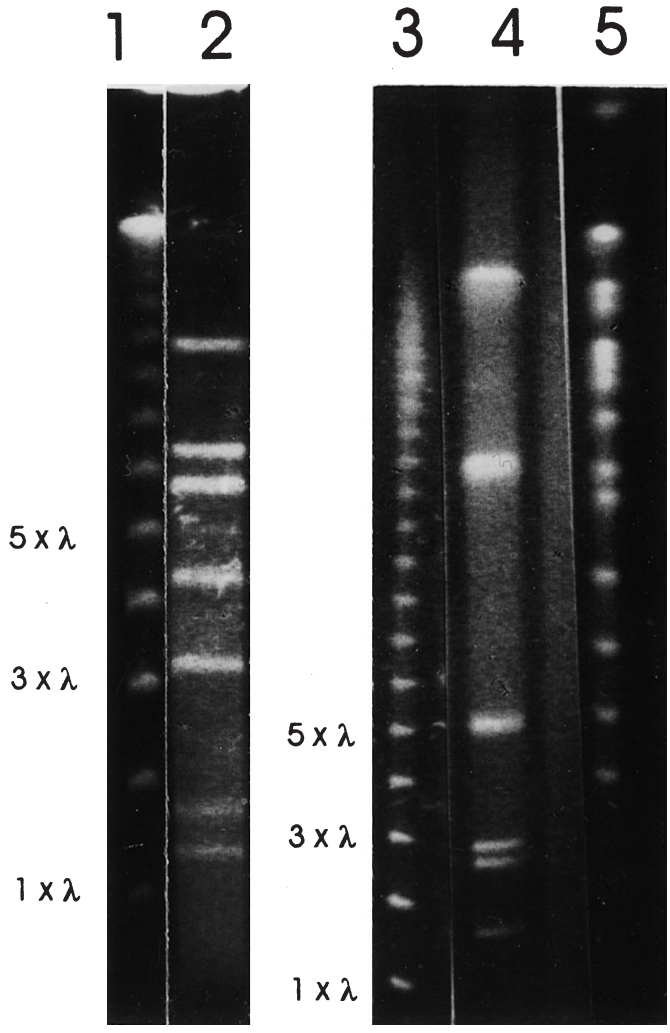

Fig. 1. Analysis of genomes of (hyper)thermophilic archaea by pulsed-field gel electrophoresis (PFGE). Growth and lysis conditions were as given in Table 1; the following samples were analyzed: *lane 1*, size marker λ-ladder (multiples of 48.5 kb); *lane 2*, *Archaeoglobus profundus*, restricted with *Bss*HII *lane 3*, size marker λ-ladder (multiples of 48.5 kb) *lane 4*, *Sulfolobus metallicus*, restricted with *Bss*HII *lane 5*, size marker yeast chromosomes (225–1900 kb)

archaeal cells; for example, in the case of *M. jannaschii*, not only one, but five genes for histones have been identified (three of which are encoded on the 1665-kb genome; Bult et al. 1996). It is therefore not surprising that susceptibility to restriction enzymes is not trivial for DNA prepared from cells in agarose plugs under our conditions. Indeed, in many cases, only restriction enzyme *Bss*HII – active at 50°C – gave satisfactory results.

We have confidence in the genome size values reported here for the following reasons: (1) In the seven cases in which different restriction enzymes could be used for size determination, the values obtained agree very well (see Table 2). (2) In a few cases, data from genome-sequencing projects will become available very soon; these correlate well with our data. For *Pyrobaculum aerophilum* we determined a genome size of ca. 1700 kb, while the value obtained by sequencing will be ca. 1800 kb. For *S. solfataricus* genome sequencing will give a value of ca. 3000 kb, compared to our 2800 kb. (3) The first genome sequence of an archaeum, namely that of *M. jannaschii*, revealed a genome size of 1665 kb (Bult et al. 1996). Our estimate for its closest relative, *M. igneus*, is 1658 kb, very close to that value.

We also want to emphasize that the size estimates given here are minimal values, for two reasons. Firstly, in some cases we could not resolve similar sized fragments from each other via PFGE; only clearly resolved DNA bands were used for our calculations. Secondly, comparison with size standards with a different G + C mol% content very often results in underestimates. In the case of *Enterococcus faecalis* (38 G + C mol%) comparisons with size markers derived from *E. coli* phage λ consistently resulted in values ca. 10% too low (unpublished results). G + C mol% contents of members of the *Sulfolobales* are in the same range.

Archaeoglobus fulgidus is special case. Here, our size estimate of 1784 kb is clearly lower than the value obtained by genome sequencing (2.18 Mb: H-P Klenk, personal communication). In this case, however, a very special situation occurs, in that the genome of *A. fulgidus* contains 13 repeats with a length of >300 bp; also the occurrence of multiple gene families (up to five copies of structural genes for certain anabolic enzymes) indicates that in this case some redundancy in the genomic organization has to be expected. Therefore, it is possible that we missed a few very similar sized DNA fragments of *A. fulgidus*.

The order *Sulfolobales* and the case of *Stygiolobus azoricus*

According to phylogenetic analyses (see inset in Fig. 2) using 16S rDNA and 23S rDNA data (Fuchs et al. 1996; Trevisanato et al. 1996), members of *Sulfolobales* belong to four major groups. The first one comprises *Metallosphaera* and *Acidianus* species. The second contains the sole species *Sulfolobus metallicus. Sulfolobus acidocaldarius* and *Stygiolobus azoricus* are members of the third group, while *Sulfolobus solfataricus*, *Sulfolobus shibatae* and isolate Ron12/III are in the fourth branch. Due to the lack of data available on physiological and biochemical differences, a reclassification of some genera in this order (otherwise fully justified) has been postponed (see Fuchs et al. 1996 for further discussion). Acoording to the data presented here, members of the first two groups have genome sizes of ca. 1.9 Mb, and those of the other two groups, >2.7 Mb, with *S. azoricus* being a distinct exception, possessing the smallest genome of all species within the order *Sulfolobales*. Indded, *Stygiolobus* is a unique case in that *S. azoricus* is the only strictly anaerobic organism within the order *Sulfolobales*; the data presented here again highlight this distinguishing feature. It remains open whether the greater metabolic potential of all other members of the order *Sulfolobales* correlates with their bigger genome size.

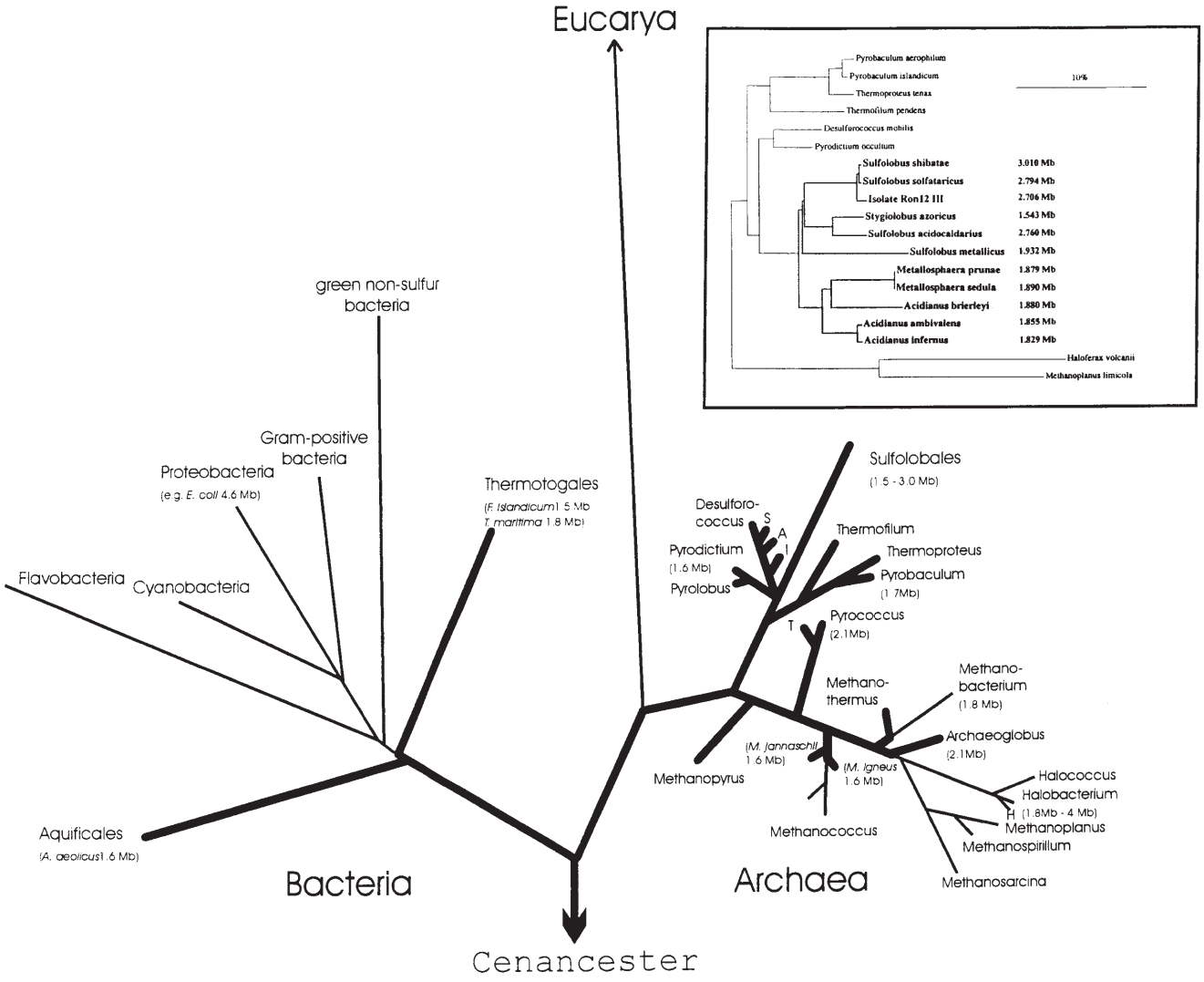

Fig. 2. Genome sizes of (hyper)thermophiles and their position within the phylogenetic tree. This scheme is redrawn from Stetter (1992), modified to show only some groups of bacteria, but representing most archaeal groups. *A, H, I, S*, and *T* represent *Aeropyrum, Haloferax, Igneococcus, Staphylothermus*, and *Thermococcus*. *Thick lines* indicate hyperthermophiles. Genome sizes not determined here were taken from Fox (1997) or provided by H-P. Klenk (personal communication). The genome sizes of haloarchaea were reported by Lopez-Garcia et al. (1996). The *inset* gives data for all validly described *Sulfolobales*. The evolutionary distance method was used for construction of the inset tree (see Fuchs et al. 1996); the *scale bar* indicates 10 substitutions per 100 nucleotides

Concluding remarks

We believe that the data given here can be helpful for defining further genome-sequencing projects. Of course, analyses of the deepest branches in the tree of life are very interesting, especially from the viewpoint of evolutionary studies/theories. For comparisons of metabolic functions and their evolution, one also should consider other points, as exemplified for the order *Sulfolobales*. The genome sequence of *S. solfataricus* should be completed within a few months; any other member of the *Sulfolobales* might have been chosen instead for genome sequencing if potential advantages with respect to genetic systems had not spoken for *S. solfataricus* [see Sensen et al. (1996)]. The ability of *Acidianus* species to oxidize sulfur under aerobic conditions, and reduce sulfur (with the aid of H_2) under anaerobic conditions, together with the fact that it has a much smaller genome than *S. solfataricus*, indicate that species of this genus are very interesting candidates for genome-sequencing projects. Obviously, *Stygiolobus azoricus* will also be a very interesting case.

In Fig. 2 an overview is given for the position of hyperthermophiles within the evolutionary tree and their genome sizes (for comparison a few other genome sizes are shown). This emphasizes that the genome sizes of (hyper)thermophiles are in the range of ca. 1.5–2 Mb; on the other hand, the genomes of *Mycoplasma genitalium* and *Mycoplasma pneumoniae* are around 0.6 and 0.8 Mb, respectively (Fraser et al. 1995; Himmelreich et al. 1996).

Since these two species are the free-living microorganisms with the smallest genome known today, one has to assume that genomes of hyperthermophiles are at least twice the size of a minimal functional genome (indeed, other calculations indicate an even smaller minimal genome; Mushegian and Koonin 1996), and therefore have undergone extensive evolution.

Acknowledgments We wish to thank K. Stetter for continuous interest, very helpful discussions, and generous support; H-P. Klenk for communicating unpublished results; and A. Muscholl-Silberhorn for very valuable technical advice. The expert technical help of T. Fuchs, K. Roth, and E. Samberger is gratefully acknowledged. This work was supported in part by a grant to RW from the Bayerische Forschungsstiftung.

References

Allen MB (1959) Studies with *Cyanidium caldarium*, an anomalously pigmented chlorophyte. Arch Microbiol 32:270–277

Ash C (1997) Year of the genome. Trends Microbiol 5:135–139

Balch W, Fox GE, Magrum LJ, Woese CR, Wolfe RS (1979) Methanogens: Reevaluation of a unique biological group. Microbiol Rev 45:260–296

Brock TD, Brock KM, Belly RT, Weiss RL (1972) *Sulfolobus*: A new genus of sulfur-oxidizing bacteria living at low pH and high temperature. Arch Microbiol 84:54–68

Bult CJ, Whiteo, Olsen G, Zhou L, Fleischmann RD, Sutton GG, Blake JA, FitzGerald LM, Clayton RA, Gocayne JD, Kerlavage AR, Dougherty BA, Tomb J-F, Adams MD, Reich CI, Overbeek R, Kirkness EF, Weinstock KG, Merrick JM, Glodek A, Scott JL, Geoghagen NSM, Weidman JF, Fuhrmann JT, Nguyen D, Utterback TR, Kelley JM, Peterson JD, Sadow PW, Hanna MC, Cotton MD, Roberts KM, Hurst MA, Kaine BP, Borodovsky M, Klenk H-P, Fraser CM, Smith HO, Woese CR, Venter JC (1996) Complete genome sequence of the methanogenic archaeon, *Methanococcus jannaschii*. Science 273:1058–1073

Burggraf S, Jannasch HW, Nicolaus B, Stetter KO (1990a) *Archaeoglobus profundus* sp. nov., represents a new species within the sulfate-reducing archaebacteria. Syst Appl Microbiol 13:24–28

Burggraf S, Fricke H, Neuner A, Kristjansson J, Rouvier P, Mandelco L, Woese CR, Stetter KO (1990b) *Methanococcus igneus* sp. nov., a novel hyperthermophilic methanogen from a shallow submarine hydrothermal system. Syst Appl Microbiol 13:263–269

Condon C, French S, Squires C, Squires CL (1993) Depletion of functional ribosomal RNA operons in *Escherichia coli* causes increased expression of the remaining intact copies. EMBO J 12:4305–4315

Fox JL (1997) Whole *E. coli!* Microbial sequencing in log-phase growth. ASM News 63:187–192

Fraser CM, Gocayne JD, White O, Adams MD, Claytor RA, Fleischmann RD, Bult CJ, Kerlavage AR, Sutton G, Kelley JM, et al. (1995) The minimal gene complement of *Mycoplasma genitalium*. Science 270:397–403

Fuchs T (1994) Physiologische und molekularbiologische Untersuchungen an neuisolierten thermoacidophilen Archaeen. Diplomarbeit, Universität Regensburg

Fuchs T, Huber H, Teiner K, Burggraf S, Stetter KO (1995) *Metallosphaera prunae*, sp. nov., a novel, metal-mobilizing, thermoacidophilic archaeum, isolated from an uranium mine in Germany. Syst Appl Microbiol 18:560–566

Fuchs T, Huber H, Burggraf S, Stetter, KO (1996) 16S rDNA-based phylogeny of the archaeal order *Sulfolobales* and reclassification of *Desulfurolobus ambivalens as Acidianus ambivalens comb.* nov. Syst Appl Microbiol 19:56–60

Grogan D, Palm P, Zillig W (1990) Isolate B12, which harbours a virus-like element, represents a new species of the archaebacterial genus *Sulfolobus*, *Sulfolobus shibatae*, sp. nov. Arch Microbiol 154:594–599

Hieter P, Boguski M (1997) Functional genomics: it's all how you read it. Science 278:601–602

Himmelreich R, Hilbert H, Plagens H, Pirkl E, Li B-C, Herrmann R (1996) Complete sequence analysis of the genome of the bacterium *Mycoplasma pneumoniae*. Nucleic Acids Res 24:4420–4449

Huber G, Stetter KO (1991) *Sulfolobus metallicus*, sp. nov., a novel strictly chemolithoautotrophic thermophilic archaeal species of metal-mobilizers. Syst Appl Microbiol 14:372–378

Huber G, Spinnler C, Gambacorta A, Stetter KO (1989) *Metallosphaera sedula* gen. and sp. nov. represents a new genus of aerobic, metal-mobilizing, thermoacidophilic archaebacteria. Syst Appl Microbiol 12:38–47

Huber H, Thomm M, König H, Thies G, Stetter KO (1982) *Methanococcus thermolithotrophicus*, a novel thermophilic lithotrophic methanogen. Arch Microbiol 132:47–50

Huber R, Woese CR, Langworthy TA, Fricke H, Stetter KO (1989) *Thermosipho africanus* gen. nov., represents a new genus of thermophilic eubacteria within the "*Thermotogales*." Syst Appl Microbiol 12:32–37

Huber R, Woese CR, Langworthy TA, Kristjansson JK, Stetter KO (1990) *Fervidobacterium islandicum* sp. nov., a new extremely thermophilic eubacterium belonging to the "*Thermotogales*". Arch Microbiol 154:105–111

Kandler O, König H (1993) Cell envelopes of archaea: structure and chemistry. In: Kates M (ed) The biochemistry of archaea (archaebacteria). Elsevier, Amsterdam

König H (1985) Influence of amino acids on growth and cell wall composition of methanobacteriales. J Gen Microbiol 131:3271–3275

Kondo S, Yamagishi A, Oshima T (1993) A physical map of the sulfur-dependent archaebacterium *Sulfolobus acidocaldarius* 7 chromosome. J Bacteriol 175:1532–1536

Kurr M, Huber R, König H, Jannasch HW, Fricke H, Trincone A, Kristjansson JK, Stetter KO (1991) *Methanopyrus kandleri*, gen. and sp. nov. represents a novel group of hyperthermophilic methanogens, growing at 110°C. Arch Microbiol 156:239–247

Lopez-Garcia P, Amils R, Anton J (1996) Sizing chromosomes and megaplasmids in haloarchaea. Microbiology 142:1423–1428

Maniatis T, Fritsch EF, Sambrook J (1982) Molecular cloning: a laboratory manual. Cold Spring Harbor Laboratory Press, Cold Spring Harbor, New York

Marshall P, Lemieux C (1992) The I-Ceu I endonuclease recognises a sequence of 19 base pairs and preferentially cleaves the coding strand of the *Chlamydomonas moewusii* large subunit rRNA gene. Nucleic Acids Res 29:6401–6407

Mushegian AR, Koonin EV (1996) A minimal gene set for cellular life derived by comparison of complete bacterial genomes. Proc Natl Acad Sci USA 93:10268–10273

Noll KM (1989) Chromosome map of the thermophilic archaebacterium *Thermococcus celer*. J Bacteriol 171:6720–6725

Peak MJ, Robb FT, Peak JG (1995) Extreme resistance to thermally induced DNA backbone breaks in the hyperthermophilic archaeon *Pyrococcus furiosus*. J Bacteriol 177:6316–6318

Pennisi E (1997) Laboratory workhorse decoded. Science 277:1432–1434

Pley U, Schipka J, Gambacorta A, Jannasch HW, Fricke H, Rachel R, Stetter KO (1991) *Pyrodictium abyssi* sp. nov. represents a novel heterotrophic marine archaeal hyperthermophile growing at 110°C. Syst Appl Microbiol 14:245–253

Ronimus RS, Musgrave DR (1996) Purification and characterization of a histone-like protein from the archaeal isolate AN1, a member of the *Thermococcales.* Mol Microbiol 20:77–86

Segerer A, Neuner A, Kristjansson JK, Stetter KO (1986) *Acidianus infernus* gen. nov., sp. nov., and *Acidianus brierleyi* comb. nov.: facultatively aerobic, extremely acidophilic thermophilic sulfur-metabolizing archaebacteria. Int J Syst Bacteriol 36:559–564

Segerer AH, Trincone A, Gahrtz M, Stetter KO (1991) *Stygiolobus azoricus* gen. nov., sp. nov. represents a novel genus of anaerobic, extremely thermoacidophilic archaebacteria of the order Sulfolobales. Int J Syst Bacteriol 41:495–501

Sensen CW, Klenk H-P, Singh RK, Allard G, Chan CC-Y, Liu QY, Penny SL, Young F, Schenk ME, Gaasterland T, Doolittle WF, Ragan MA, Charlebois RL (1996) Organizational characteristics and information content of an archaeal genome: 156 kb of sequence from *Sulfolobus solfataricus* P2. Mol Microbiol 22:175–191

Smith CL, Cantor CR (1987) Purification, specific fragmentation, and separation of large DNA molecules. In: Wu R (ed) Methods in enzymology, vol 155. Academic San Diego

Stetter KO (1988) *Archaeoglobus fulgidus* gen. nov., sp. nov.: a new taxon of extremely thermophilic archaebacteria. Syst Appl Microbiol 10:172–173

Stetter KO (1992) Life at the upper temperature border. In: Tran Than Van J, Tran Than Van K, Mounolou JC, Schneider J, McKay C (eds) Frontier of life. Editions Frontieres, Gif-sur-Yvette, pp 195–219

Stetter KO (1996) Hyperthermophilic procaryotes. FEMS Microbiol Rev 18:149–158

Stetter KO, Thomm M, Winter J, Wildgruber G, Huber H, Zillig W, Janecovic D, König H, Palm P, Wunderl S (1981) *Methanothermus fervidus*, sp. nov. a novel extremely thermophilic methanogen isolated from an icelandic hot spring. Z Bakteriol Hyg I. Abteilung, Originale C 2:166–178

Stetter KO, König H, Stackebrandt E (1983) *Pyrodictium* gen. nov., a new genus of submarine disk-shaped sulfur-reducing archaebacteria growing optimally at 105°C. Syst Appl Microbiol 4:535–551

Trevisanato SI, Larsen N, Segerer AH, Stetter KO, Garrett RA (1996) Phylogenetic analysis of the archaeal order of *Sulfolobales* based on sequences of 23S rRNA genes and 16S/23S rDNA spacers. Syst Appl Microbiol 19:61–65

Völkl P, Huber R, Drobner E, Rachel R, Burggraf S, Trincone A, Stetter KO (1993) *Pyrobaculum aerophilum* sp. nov., a novel nitrate-reducing hyperthermophilic archaeum. Appl Environ Microbiol 59:2918–2926

Woese CR, Fox GE (1977) Phylogenetic structure of the prokaryotic domain: The primary kingdoms. Proc Natl Acad Sci USA 74:5088–5090

Zillig W, Stetter KO, Schulz W, Priess H, Scholz I (1980) The *Sulfolobus* – "*Caldariella*" group: taxonomy on the basis of the structure of DNA-dependent RNA-polymerases. Arch Microbiol 125:259–269