

Studies on structure and function of
the rRNA introns in the hyperthermophilic
archaeon *Aeropyrum pernix*

Norimichi Nomura

1998

To my family and to my dear friends

Contents

Chapter 1

Introduction

Chapter 2

Isolation and Characterization of a Marine Aerobic Hyperthermophilic Archaeon *Aeropyrum pernix* : Discovery of the Novel Genus in the *Archaea* Domain

Chapter 3

Intraspecies Polymorphism in the Single rRNA Operon of *A. pernix*, Implying the Presence of Hotspots for Intron-insertion

Chapter 4

Post-splicing Dynamism of the Excised Intronic RNAs in *A. pernix* K1 Cells

Chapter 5

Functional Analyses of the Latent Intron-encoded Protein I- *Ape* I: Potential Role in the Horizontal Transfer of the Archaeal Introns

Summary

Acknowledgment

References

List of Publications

Chapter 1

Introduction

Marine hydrothermal environments, e. g. submarine hot vents and coastal hot springs, are considered to be extreme by human beings in terms of temperature; however, it is currently well recognised that many microorganisms are specifically adapted to these ecological niches (Stetter, 1996). These organisms, designated as hyperthermophiles, not only survive but actively grow at temperatures above 90 °C. Over past fifteen years, extensive studies of the ecology, physiology, taxonomy and molecular biology of hyperthermophiles have been undertaken (reviewed by Cowan, 1992; Olsen and Woese, 1997; Belfort and Weiner, 1997; Dennis, 1997). These have resulted in a complete reassessment of our concept of microbial evolution. In particular, the identification of the *Archaea* (originally called archaeobacteria)(Woese and Fox, 1977; Woese *et al.*, 1990) as the third domain of life has given considerable impetus to hyperthermophile research. Since then, on the basis of the comparative sequence analyses of rRNAs (Olsen *et al.*, 1994) and a few other proteins such as elongation factors EF-1 α /Tu, EF-2/G (Iwabe *et al.*, 1989), RNA polymerase β , β' subunits (Puhler *et al.*, 1989), and V- and F- type ATPases (Gogarten *et al.*, 1989), all extant organisms were classified into three domains of the *Bacteria* (eubacteria), the *Eucarya* (eukaryote), and the *Archaea*, although the phylogenetic relationship among these three domains is still a matter of controversy. The last is comprised of two separate kingdoms; the Euryarchaeota which includes the methanogens, the extreme halophiles and certain related extreme thermophiles, and the Crenarchaeota which is entirely comprised of thermophilic members of orders '*Igneococcales*', *Sulfolobales*, and *Thermoproteales*. Since hyperthermophiles occur on evolutionarily deep branches both within the kingdoms Euryarchaeota and Crenarchaeota, thermophily is presumed to be the ancestral phenotype. Conversely, hyperthermophilic archaea, therefore, might provide some valuable hints for the exploration of early evolution of life and the nature of the most recent universal ancestor.

Hyperthermophilic archaea are currently presenting many new challenges in basic biological sciences, because there is so much to be learned ---about this group of exotic microorganisms themselves, thermophily, their relationship to the eukaryotic cells; certain complex eukaryotic functions can be effectively studied in simpler hyperthermophilic archaeal systems, molecular structures can be inferred from thermostable archaeal proteins, and the functional essence of an enzyme or system can be revealed by a broader comparative analysis. From this point of view, the latest whole genome sequencing of hyperthermophilic archaea (Smith *et al.*, 1997; Klenk *et al.*, 1997; Bult *et al.*, 1996) seems remarkably insightful and productive.

Whilst considered to be mere 'scientific curiosities', it is also generally accepted that hyperthermophilic archaea have considerable biotechnological and commercial significance. Hyperthermophiles provide a valuable resource for exploitation in novel biotechnological processes and in developing our understanding of how biomolecules are stabilized when subjected to extreme high temperatures. They have provided thermostable enzymes for application in industrial processes, which are used as diagnostic enzymes (reviewed by Cowan, 1992) and have applications in genetic engineering (reviewed by Bergquist and Morgan, 1992). They also provide models for protein engineers attempting to determine the basis of protein stability (Bohm and Jaenicke, 1994; Britton *et al.*, 1995; Schultes and Jaenicke, 1991; Cannio *et al.*, 1994; Arias and Argos, 1989; Tamakoshi *et al.*, 1995; Kotsuka *et al.*, 1996; Chan *et al.*, 1995). In the field of bioremediation, they have considerable application in the removal of heavy metals and the microbial desulfurization of coal to reduce sulfur emissions (reviewed in Gadd, 1992). In the search for alternative energy resources to replace fossil fuels, they offer potential for the large scale production of ethanol, organic solvents, methane and hydrogen. Some hyperthermophiles have also been exploited for many years in the leaching of metals from low grade ores and more recently in precious metal recovery.

Hyperthermophiles known to date have been restricted to the strictly anaerobes (Stetter *et al.*, 1990; Stetter, 1996), which include the methanogenic archaea, the archaeal sulfate reducers, the archaeal S⁰-metabolizers, and the genera *Thermotoga* and *Aquifex* within the *Bacteria*. This might stem partly from the preconception that oxygen availability at the elevated temperatures is

low due to its poor solubility. However, because of their lower growth rates and cell yields and highly complex procedures for cultivation, detailed research on biochemistry and molecular biology of hyperthermophilic archaea has been precluded. Therefore, I mounted an exploration of the aerobic organism which can grow optimally at temperatures above 90 °C.

After months of trial and error, I succeeded in isolation and cultivation of a novel type of the strictly aerobic hyperthermophilic archaeon *Aeropyrum pernix*, and quite significant increases in cell yield could be achieved with minor modifications to culture conditions (Chapter 2; Sako *et al.*, 1996). Subsequently, in the process of the study on phylogenetic characterization of this organism based on rRNA sequence, I unexpectedly found the intervention of multiple introns within the rRNA gene locus (Chapter 3). Furthermore, the results implied the presence of hotspots for intron-insertion. These unusual observations promoted me to conduct an intensive investigation of the structure and regulation of the rRNA operon (*arnSL*) of this attractive organism.

The overall aim of this study is to clarify the structure and function of the rRNA introns in the hyperthermophilic archaeon *A. pernix*, and to understand how the intraspecies polymorphism in intron-insertion was generated. In Chapter 3, I reported the intraspecies polymorphism in the rRNA operon of *A. pernix* and also discuss the implication of the presence of hotspots for intron-insertion within this locus. In Chapter 4, to examine whether the protein-coding intronic RNAs, $I\alpha$ and $I\gamma$, can function as mRNAs for corresponding ORFs after splicing events, post-splicing dynamics of the excised intronic RNAs in *A. pernix* K1 cells was investigated. In Chapter 5, I describe the functional characterization of the intron-encoded protein and discuss the role that it may have played in the horizontal transfer of *A. pernix* rRNA introns

Chapter 2

Isolation and Characterization of a Marine Aerobic Hyperthermophilic Archaeon *Aeropyrum pernix* : Discovery of the Novel Genus in the *Archaea* Domain

INTRODUCTION

The proposal of the *Archaea* (originally called archaeobacteria) as a discrete domain (Woese and Fox, 1977; Woese *et al.*, 1990) shed new light on the central problems of both early evolution of life and prokaryotic systematics. Although this concept is now generally accepted because of the several biochemical features peculiar to the *Archaea*, it is still a matter of controversy how the *Archaea* domain is phylogenetically related to the other two domains, the *Eucarya* (eukaryote) and the *Bacteria* (eubacteria) (Cavalier-Smith, 1992).

This problem stems partially from an essential lack of information on the deepest (earliest) branches within the universal phylogenetic trees. For instance, a "deep missing branch" might decrease the reliability of the deep-branching topologies of the phylogenetic trees. In order to obtain additional information pertinent to this problem, I aimed to isolate organisms that might be representative of the deep missing branches.

Interestingly, it was pointed out that the deepest and shortest branches within the universal phylogenetic trees are dominated by hyperthermophiles, which grow optimally at temperatures above 80 °C (Achenbach-Richter *et al.*, 1987a; Achenbach-Richter *et al.*, 1988; Achenbach-Richter *et al.*, 1987b; Woese, 1987). During the past decade, many new hyperthermophiles were isolated from solfataric fields and submarine volcanic vents (Stetter *et al.*, 1990). Those include the hyperthermophilic methanogenic archaea, the archaeal sulfate reducers, the hyperthermophilic S⁰-metabolizers, and the genera *Thermotogales* and *Aquifex* within the *Bacteria*. However, since oxygen availability in the hydrothermal environments is low because of

poor solubility, most studies of life in these ecosystems have been restricted to the anaerobic organisms (Hoaki *et al.*, 1994; Huber *et al.*, 1990) .

Recently, I succeeded in isolation and cultivation of a novel type of the strictly aerobic hyperthermophilic archaeon, growing optimally at temperatures above 90 °C. In this chapter, I characterize the new isolate and propose that it should be classified in a new genus, *Aeropyrum* (Ae.ro.py'rum. Gr. n. *aer* air; Gr. neutr. n. *pyr* fire; M.L. masc. n. *Aeropyrum* air fire, referring to the hyperthermophilic respirative character of the organism).

MATERIALS AND METHODS

Collections of samples. Samples were collected from several sites in Japan where volcanic activities are associated.

(i) Tachibana Bay and Obama Hot Spring, Nagasaki Prefecture. Hot effluent water samples from submarine hydrothermal vents were collected from Tachibana Bay (32° 45' N, 130° 10' E) at a depth of 22 m, and hot water, soils and muds were collected from coastal hot springs at Obama (32° 43' N, 130° 12' E) in 1992. Aerobic extremely thermophiles were isolated from this fields (Sako *et al.*, 1997; Nomura *et al.*, unpublished observation).

(ii) Kodakara Island, Kagoshima Prefecture. Hot sedimentary materials and venting water were collected at a coastal solfataric thermal vent in Kodakara Island (29° 13' N, 129° 20' E) during a cruise of the research vessel *Sogen-maru* in 1993. Anaerobic hyperthermophiles were isolated from this field (Hoaki *et al.*, 1994; Hoaki *et al.*, 1995; Morikawa *et al.*, 1994).

In situ temperature was determined with a mercury filled Celsius thermometer (0 to 200 °C), and the pH was estimated by a pH meter (Toa Electronics, Tokyo, Japan). After collection, the samples were brought back to the laboratory, cooled on ice, and stored at 4 °C until enrichment.

Culture conditions. The base for all media described in this dissertation is the synthetic sea water JS, which was prepared from Jamarine S synthetic sea salts (Jamarine Laboratory, Osaka, Japan) according to the manufacturer's instructions. Unless specified otherwise, the new isolate was

TABLE 2.1 Source of the isolates.

Strain	Source	
	Locality	Temp. (°C)
K1	Coastal sulfataric vent in Kodakara Island (Kagoshima)	98-104
OH2	Coastal hot spring at Obama (Nagasaki)	85- 90
TB1	Shallow hydrothermal vent at Tachibana Bay (Nagasaki)	120-125
TB7	Shallow hydrothermal vent at Tachibana Bay (Nagasaki)	120-125

grown in the standard medium referred to as JXT, containing (per liter of JS): yeast extract (Difco Laboratories, Detroit, Mich.), 1g; trypticase peptone (BBL Microbiology Systems, Cockeysville, Md.), 1g; $\text{Na}_2\text{S}_2\text{O}_3 \cdot 5\text{H}_2\text{O}$, 1g; pH 7.0 (adjusted with HCl at room temperature). For enrichment cultures and isolation, JX medium was used, which prepared by omitting $\text{Na}_2\text{S}_2\text{O}_3 \cdot 5\text{H}_2\text{O}$ from JXT medium.

Cultures were routinely grown in screw-capped test tubes (180 by 18mm) containing 10 ml of medium and incubated at 90 °C in a forced convection oven (FC-610; Advantec, Tokyo, Japan) without shaking. In order to maintain aerobic culture conditions, the gas phase of the closed test tubes was not evacuated and the atmospheric air (100 kPa) was confined. Resazurin (1 μg / ml) was used as a redox indicator.

Batch cultures were grown in cotton-plugged 2 liter-Erlenmeyer flasks containing 500ml of medium and incubated at 90 °C in an air bath rotary shaker (RGS-32.TT; Sanki Seiki, Osaka, Japan) under vigorously shaking (180 rpm).

Substrate utilization. In an attempt to find carbon substrates which would sustain growth and to test the optimum concentration for growth, each of the following organic substances (obtained from Nacalai Tesque, Kyoto, Japan, except where noted otherwise) was added to 10ml of the JS base, supplemented with 0.1% (wt/vol) $\text{Na}_2\text{S}_2\text{O}_3 \cdot 5\text{H}_2\text{O}$ in concentrations (wt/vol) of 0.05, 0.1, 0.2, and 0.4%: D-(-)-ribose, L-(+)-arabinose, D-(+)-xylose, D-(+)-glucose, L-(-)-glucose, D-(+)-galactose, D-(-)-fructose, L-(-)-fucose, L-(+)-rhamnose, D-(+)-mannose, maltose, lactose, sucrose, D-(+)-cellobiose,

D-(+)-melibiose, D-(+)-raffinose, cellulose, starch, D-sorbitol, D-mannitol, glycerol, acetate, alanine plus glycine, 20 different L-amino acids, yeast extract (Difco), tryptone (Difco), nutrient broth (Difco), Casamino acid (vitamin free; Difco), trypticase peptone (BBL). To test the various sulfur compounds, the $\text{Na}_2\text{S}_2\text{O}_3$ in JXT medium was replaced with 10mM (each) Na_2SO_4 , Na_2SO_3 , S^0 , cysteine, or methionine. These cultures were incubated at 90 °C without shaking.

Sulfide analysis. The sulfide production during growth was detected qualitatively as described by Huber *et al.* (1990).

Oxygen requirement for growth. In order to ensure strict anaerobiosis, the anaerobic techniques described by Balch and Wolfe (1976) were employed. Oxygen was reduced by adding 0.05%(wt/vol) $\text{Na}_2\text{S} \cdot 9\text{H}_2\text{O}$. Microaerobic culture conditions were achieved as described by Huber *et al.* (1992) with minor modifications. Prior to autoclaving, 10ml of medium was dispensed into a 120ml-glass bottle (Schott Glaswerke, Mainz, Germany) that was tightly sealed with butyl rubber stoppers, and the gas phase was exchanged with either $\text{H}_2\text{-CO}_2$ (80:20, 300 kPa) or N_2 (100%, 300 kPa). In both experiments, the state of oxygenation was monitored by resazurin (1 μ g/ ml), and cultures were incubated at 90 °C without shaking.

Light and electron microscopy. Cells were inspected with a Nikon Optiphot XF-NT differential interference light microscope equipped with an oil immersion objective 100/1.25 . Transmission electron microscopy of negatively stained cells was carried out as described by Zillig *et al.* (1990). The cells were negatively stained with 1 to 2% (wt/vol) uranyl acetate and observed in a Hitachi H-7000 transmission electron microscope at an accelerating voltage of 75kV. For ultrathin sectioning, cells were fixed in JXT medium with 2% (vol/vol) glutaraldehyde and postfixed with 1% OsO_4 . The fixed cells were then dehydrated through a series of ethanol and embedded in Quetol 812 (Nisshin EM, Tokyo, Japan) epoxy resin. Thin sections were double contrasted with uranyl acetate and lead citrate. The specimens were observed with a JEOL 1200EXII electron microscope operated at 80 kV.

Determination of growth. Growth was determined either by measurement of the optical density at 660nm(OD_{660}) or by direct cell counting using a Thoma chamber (depth: 0.02 mm). A good correlation was found

between these two methods in exponential and stationary phase cultures. In the standard culture conditions, an OD₆₆₀ of 1 corresponded to a cell density of about 1.4×10^9 cells per ml.

To study the growth characteristics, the batch culture technique described above was applied. At temperatures above 100 °C, screw-capped 1-liter glass bottles (Schott Glaswerke) were used to prevent the evaporation of the medium. The pH dependency of growth was studied in JXT medium with various buffer systems at a concentration of 20mM, e.g., sodium succinate (below pH 5.0), MES [2-(*N*-morpholino) ethanesulfonic acid] (pH 5.5 to 6.5), HEPES [*N*-(2-hydroxyethyl) piperazine-*N'*-2-ethanesulfonic acid] (pH 7.0 to 8.0), and BTP {1,3-bis [tris (hydroxymethyl) methylamino] propane} (pH 8.5 or more). The pH of the medium was adjusted by the addition of HCl or NaOH at room temperature. Growth at different salinities was determined using a 2× JS diluted appropriately in the medium. Doubling times were calculated using linear regression analysis from 3 to 5 points along the logarithmic part of the resulting growth curves.

Antibiotics sensitivity. Sensitivity to the antibiotics, chloramphenicol, ampicillin, vancomycin, cycloserine (all from Sigma Chemical, St. Louis, Mo.) was tested at a concentration of 100 μg/ml. The temperature for this test was 70 °C. Preliminary experiments with the bacterium *Thermus aquaticus* (ATCC 25104) had indicated that the antibiotics were effective at 70 °C.

Lipid analysis. Total lipid was extracted from the cells as described by Bligh and Dyer (1959) and analyzed by two-dimensional thin layer chromatography (2D-TLC) performed on a silica gel 60 plates (Art. 5721; Merk, Darmstadt, Germany). Solvents included chloroform-methanol-7M aqueous ammonia (65:35:8, vol/vol/vol) in the first dimension and chloroform-methanol-acetic acid-water (85:30:15:5, vol/vol/vol/vol) in the second dimension. Polar lipid composition was determined by the measurement of phosphorus (Bartlett, 1959) of individual lipid spots on the 2D-TLC chromatogram.

Core lipids were prepared by splitting off the polar head groups by acid methanolysis (5% HCl-methanol at 100 °C for 2h) of the total lipid extracts and separated by TLC with the solvent, light petroleum-diethyl ether-acetic acid (50:50:1, vol/vol/vol). Hydrocarbon chains were prepared

from the core lipids by hydroiodic acid degradation followed by LiAlH_4 reduction as previously reported (Nishihara *et al.*, 1989) and identified by gas chromatography-mass spectrometry and electron impact mass spectrometry employing instrumentation and conditions as previously described (Nishihara and Koga, 1991). The polar head groups obtained in the aqueous fraction of the acid methanolysis was further acid-hydrolyzed (1M HCl at 100 °C for 3h, or 4M HCl at 100 °C for 16h) and acetylated. The acetylated hydrolysate was analyzed by gas liquid chromatography as described previously (Nishihara *et al.*, 1989). Glycerophosphoric esters prepared from the total lipids extracts by dealkylation with BCl_3 (Nishihara and Koga, 1988) were analyzed by cellulose TLC (Art. 5716, Merck)(Nishihara and Koga, 1988) with the solvent phenol-water (100:38, vol/vol). Authentic samples of archaeol and caldarchaeol (Nishihara *et al.*, 1987) were prepared from total lipids of *Methanobacterium thermoautotrophicum* or *Sulfolobus solfataricus*.

DNA preparation. Cell pellets (about 1g) were suspended in 10ml NET buffer (50mM Tris-HCl, 150mM NaCl, 100mM $\text{Na}_2\text{-EDTA}$ [pH 8.0]) containing 2% (wt/vol) sodium dodecyl sulfate and 0.1mg of proteinase K (Nacalai Tesque) per ml and incubated at 65 °C for 30 min. The crude genomic DNA was prepared by the method of Sambrook *et al.* (1989) and was purified by a protocol which employed cetyltrimethyl ammonium bromide (Murray and Thompson, 1980) and by CsCl density-gradient centrifugation .

DNA base composition. The G+C content of the genomic DNA was determined by high-performance liquid chromatography (Mesbah *et al.*, 1989) with a DNA-GC kit (Yamasa Shoyu, Tokyo, Japan). Calf thymus DNA (42 mol% G+C; Sigma)(Marmur and Doty, 1962) and nonmethylated DNA from bacteriophage λ (49.8 mol% G+C; Takara Shuzo, Kyoto, Japan)(Sanger *et al.*, 1982) were used as references.

16S and 23S rRNA preparation. 16S and 23S rRNA of isolate K1 were prepared essentially according to Traub *et al.* (1971) with modifications as follows. The exponentially growing cells (about 1g) were suspended in 20ml of buffer I [10 mM Tris-HCl, 10 mM $\text{Mg}(\text{CH}_3\text{COO})_2$, 60mM NH_4Cl , 6mM 2-mercaptoethanol, pH 7.8] and 20 ml glass beads (type I; Sigma) was added. The suspension was vortexed for 15 min at 4 °C and centrifuged

twice at $30,000 \times g$ for 30 min. The supernatant was removed and recentrifuged at $105,000 \times g$ for 6h to precipitate the ribosomes. The pellet was resuspended in buffer I and treated with RNase-free DNase I (Boehringer, Mannheim, Germany) at a concentration of $2 \mu\text{g/ml}$. rRNAs were obtained after three extractions with phenol-chloroform followed by two ethanol precipitations and separated by a 5 to 20% sucrose density gradient centrifugation at $100,000 \times g$ for 17 h (TST 28.38 rotor; Kontron Instruments, Milano, Italy). Fractions containing 16S or 23S rRNA were pooled. Purity of the rRNAs was checked electrophoretically.

16S and 23S rRNA analyses. cDNA clones of the 16S and 23S rRNA of isolate K1 were obtained after reverse transcription and PCR amplification. The oligonucleotide (1521R, 5'-AGGTGATTCAGCCGCA GGTT-3'), complementary to 3'-tail conserved region of the archaeal 16S rRNA were employed to prime first-strand synthesis with SuperscriptII RNaseH⁻ (GIBCO Laboratories, Grand Island, N.Y.). The reverse transcription reaction was performed according to the manufacture's instructions, and the product was amplified by PCR with Arch21F (5'-TTCCGGTTGATCCYGCCGGA-3') (DeLong, 1992) and 1521R as primers. Thirty five amplification cycles of 90 s at 96°C , 1 min at 62°C , 2 min at 72°C were performed. The double-stranded PCR product was cloned directly into the pCRII vector (Invitrogen, San Diego, CA.) to yield plasmid pNA4. On the other hand, the oligonucleotide (DmLR3070R, 5'-CGGGCTCTTG GGAGCGGCG-3'), complementary to 3'-tail region of the 23S rRNA of *Desulfurococcus mobilis* (Leffers *et al.*, 1987) was used to prime first-strand synthesis, and the product was amplified by PCR with 0227aF (5'-GAGGAAAAGAAATCAA-3') (Achenbach and Woese, 1995) and DmLR3070R as primers. The resulting double-stranded PCR product was cloned directly into the pCRII vector (Invitrogen, San Diego, CA.) to yield plasmid pNB23. Nested deletion mutants were produced by progressive exonuclease III digestion (Yanisch-Perron *et al.*, 1985), and the nucleotide sequence was determined by the dideoxy sequencing method (Sanger *et al.*, 1977). The sequence was aligned to a collection of archaeal rRNA sequences (DNA Data Bank of Japan [DDBJ], National Institute of Genetics, Shizuoka, Japan) using the Clustal W (Thompson *et al.*, 1994). Numbers of base substitutions

per site were estimated by the method of Kimura's 2-parameter model (Kimura, 1980), and the phylogenetic tree was inferred based on neighbor joining method (Saitou and Nei, 1987). The resulting tree was tested using bootstrap analysis (Felsenstein, 1985).

Amplification of 16S-23S internal transcribed spacer (ITS) and sequence determination. The ITS region was amplified by PCR. The primers used for amplification were 1538F (5'-CGGTTGGATCACCTC-3') and 0213aR (5'-GTTGGTTTCTTTTCCT-3')(Achenbach and Woese, 1995). The PCR products (ca. 360 bp) were cloned directly into the pCRII vector, and the nucleotide sequences were analyzed as described above.

Nucleotide sequence accession number. The 16S and 23S rRNA sequences of isolate K1 have been deposited in the DDBJ, EMBL and GenBank nucleotide sequence databases with the accession number D83259 and AB004787, respectively.

RESULTS

Enrichment and isolation of the aerobic hyperthermophile. In order to enrich aerobic hyperthermophiles, 10 ml of JX medium were inoculated with approximately 1g of sample in the laboratory. All the samples were mixtures of sandy sediment and clear sea water. Microbial mat material was not visible in the samples. The original temperature were shown in Table 2.1. The enrichment was performed at 90 °C in screw-capped test tubes with air as the gas phase (100 kPa) without shaking. Within 2 days, turbidity caused by cell growth was observed. This growth consisted of a mixed population of various sizes of rods and cocci. During this enrichment culture, resazurin remained red, and sulfide was not produced. All positive enrichment cultures could be successfully transferred in the same medium.

In order to obtain pure cultures, a dilution-to-extinction technique was employed (Baross, 1995). After the cell density of a enrichment culture reached to around 10^7 cells per ml, 5 separate dilution series were conducted in which 5 1:10 dilutions of the culture were followed by 20 1:3 dilutions. Each dilution in the dilution-to-extinction series was carried out in triplicate

and incubated for at least 7 days. The cultures in the tube showing growth at the highest dilution were designated isolates K1, OH2, TB1, and TB7, which usually grew up to about 10^7 cells per ml in 4 days. The purity of these cultures were routinely confirmed by microscopic examinations and by the restriction fragment length polymorphism (RFLP) analysis of the PCR-amplified 16S rRNA gene fragment (data not shown). For further investigations, JX medium was replaced by JXT medium, in which a higher final cell density was obtained as described below. In this medium, the cell yield in batch culture was about 1 g/liter (wet weight) in stationary phase. Packed cell masses exhibited a brown color.

Morphology. Cells of four isolates were irregular cocci with some sharp edges. The cell size was 0.8 to 1.0 μm in diameter. The cells stained Gram negative. They often appeared singly, but pairs were also observed (Fig. 2.1). Motility was observed under the light microscope in logarithmic and stationary phase cultures. Motility was strongly enhanced by heating the microscopic slide glass to 90 °C. The cells were frequently surrounded by pili-like appendages (Fig. 2.2a). Flagella were not observed.

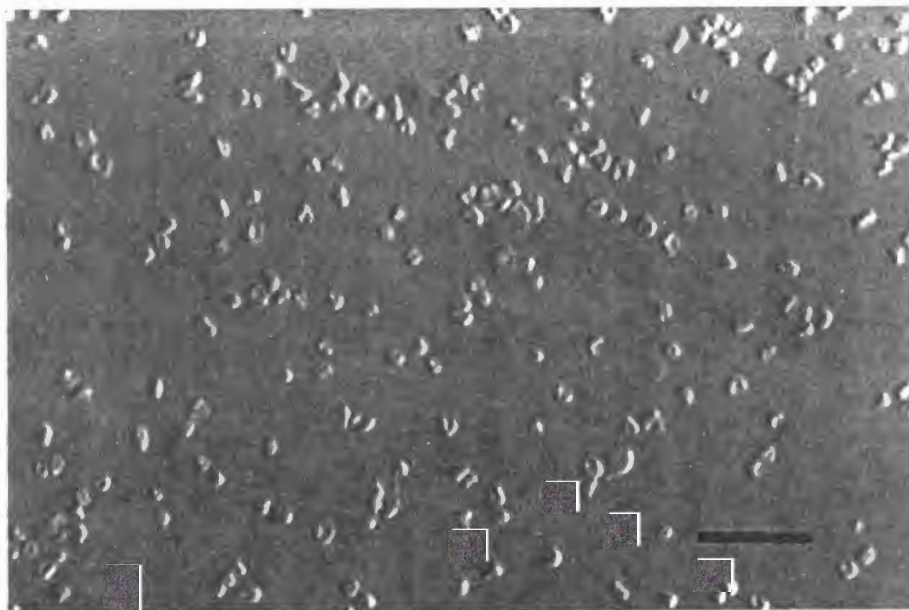


FIG. 2.1 Differential interference micrograph of *Aeropyrum pernix* K1 cells. Bar, 5 μm .

When treated by a low osmotic shock, ghost cells of lower electron density were observed in some projections. They exhibited larger diameters than cells with high electron density, and revealed a structure of the envelope. But its lattice regularity in ultrastructure could not be discerned (Fig. 2.2b). The cells were covered by a cell envelope (S-layer-like structure) about 25 nm wide outside the cytoplasmic membrane (Fig. 2.3). A typical subunit structure of the S-layer was not clearly discerned in this cell envelope. However, small areas of the cell envelope were sectioned at a shallow angle (margin of the cell appearing in the lower left corner of Fig. 2.3) and a faint repeating pattern was visible. Extracellular materials in sheets and small fragments shown in Fig. 2.3 might be the cell envelope and was present in stationary phase cell pellet. Isolates also had a dense layer just outside the cytoplasmic membrane, with a separate cell envelope spaced much further out from the cell. These structural features are different from the archaeal cell surfaces described to date (Baumeister and Lembcke, 1992).

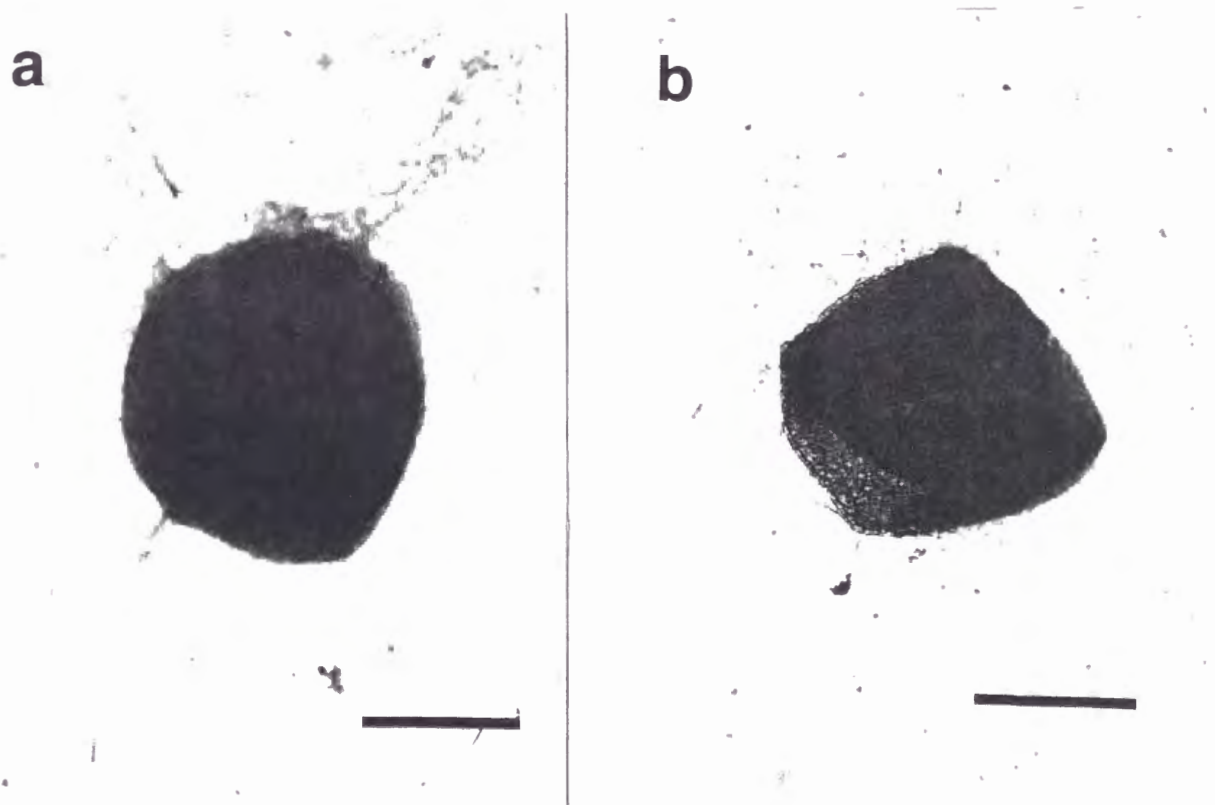


FIG. 2.2 Electron micrographs of *Aeropyrum pernix* K1. (a) Single cell with pili. (b) Ghost exhibiting a cell envelope. All cells were negatively stained with uranyl acetate. Bars, 0.5 μ m.

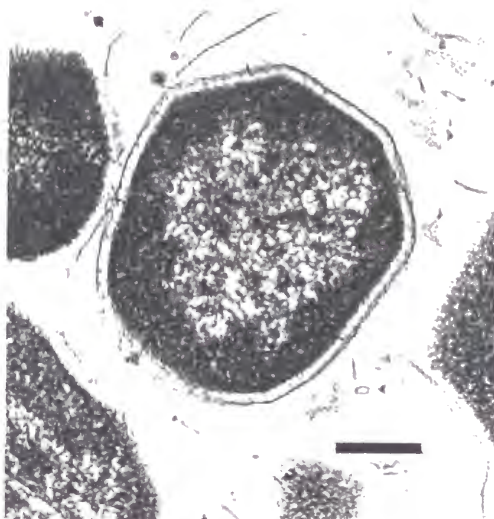


FIG. 2.3 Ultrathin section of *Aeropyrum pernix* K1. Bar, 0.2 μ m.

Physiological characterization of growth. Isolates were heterotroph and grew only under aerobic culture conditions, indicating that isolates grew by aerobic respiration. Both under microaerobic and strictly anaerobic conditions, growth was completely inhibited. The cells exposed to strictly anaerobic condition but not microaerobic conditions were rapidly inactivated and lysed. Growth did not resume even upon return to aerobic conditions. The microaerobic incubation merely prevented the growth and did not kill cells. Maximum cell densities (up to 1.5×10^9 cells per ml) were obtained by cultures incubated with shaking.

The new isolates were able to grow on proteinaceous complex substrates such as yeast extracts, trypticase peptone, tryptone, or nutrient broth but not on Casamino acids. Yeast extract, trypticase peptone, and tryptone allowed the growth at concentrations of 0.05 to 0.4%. Yeast extract and trypticase peptone (0.1%) were the most efficacious in supporting growth. A slight growth inhibition was observed with yeast extract, trypticase peptone, and tryptone at a 0.4% concentration. Nutrient broth concentrations of over 0.1% inhibited growth. Any single carbohydrate, organic acid, or amino acid tested within the concentration range of 0.05 to 0.4% did not suffice as a sole carbon and energy source. The growth yield was stimulated about eightfold by thiosulfate. However, thiosulfate was not required for growth. H_2S was not formed during growth in either the presence or absence of thiosulfate. The growth yield was not stimulated by other sulfur-containing compounds such as sulfate, sulfite, S^0 , sulfide, cysteine or methionine.

Isolates grew well between 70 and 97 °C, and cell division still occurred at 100 °C (Fig. 2.4). No growth was observed at 68 °C and 102 °C. The optimum growth temperature was 90 to 95 °C, with a doubling time of 200 min (Fig. 2.5a). The optimum pH for growth was around 7.0. Growth was obtained in a pH range from 5.0 to 9.0 (Fig. 2.5b). Metabolism of the medium components did not cause the initial medium pH to change significantly during these experiments (data not shown). The optimum salinity in the medium was found to be approximately 3.5 %, which was the same level of ionic strength as JXT medium (Fig. 2.5c). Growth was observed at a salinity from 1.8 to 7.0%. At a salinity of 1.5% in the medium, rapid cell lysis occurred, as indicated by the presence of ghost cells upon electron microscopic examination and by the decreased turbidity of the culture.

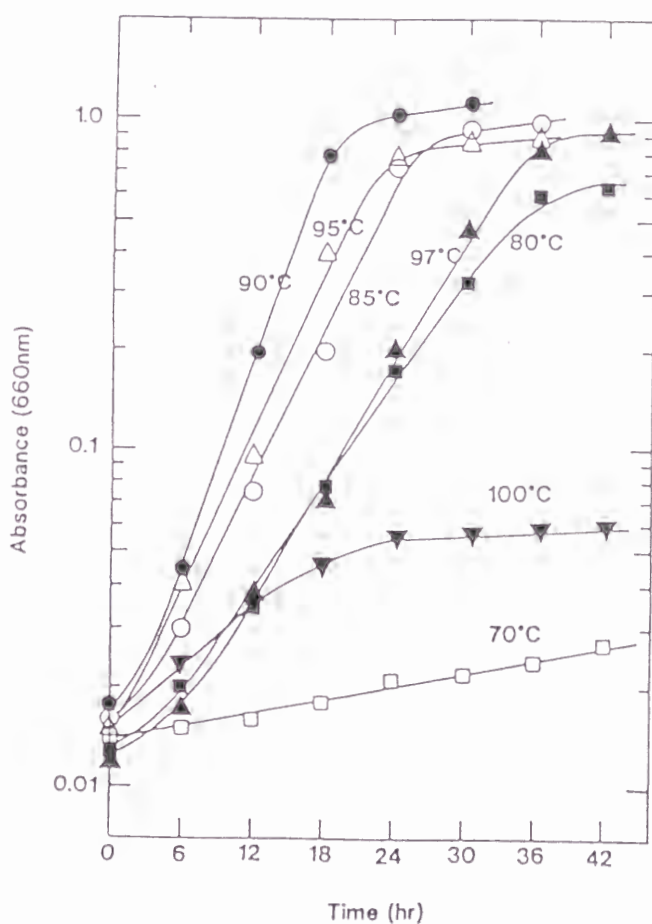


FIG. 2.4 Temperature dependence of growth of *Aeropyrum pernix* K1.

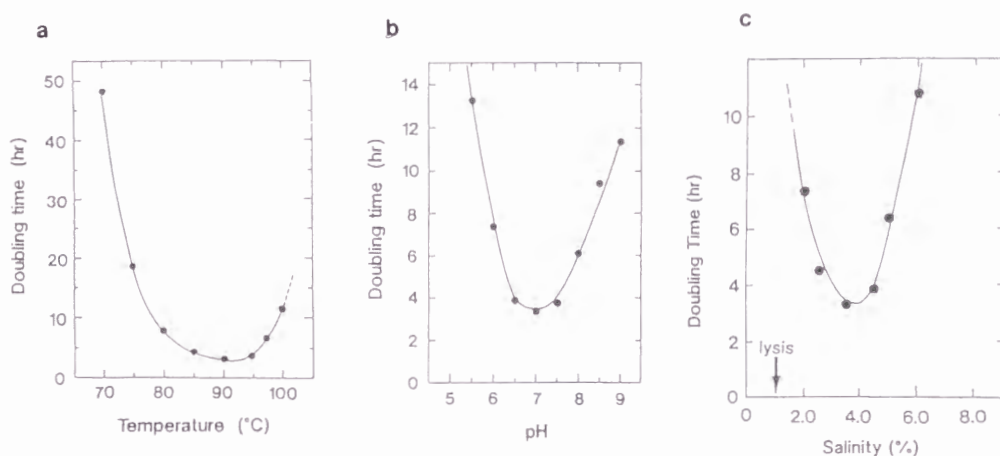


FIG. 2.5 Optimum growth conditions of *Aeropyrum pernix* K1. Doubling times were calculated from the slopes of the growth curves. (a) Effect of temperature on growth. (b) Effect of pH on growth. (c) Effect of temperature on growth.

Sensitivity to antibiotics. Isolate K1 was insensitive to 100 μ g/ml of ampicillin, vancomycin, and cycloserine but was completely inhibited in the presence of the same concentration of chloramphenicol.

Lipids. The 2D-TLC pattern of the total lipid of isolate K1 appeared to be different from those of archaeal total lipids investigated to date. At least five polar lipids were detected. Among these five polar lipids, one phosphoglycolipid (designated as AGI) and one phospholipid (designated as AI) were predominant. Assuming that each lipid contains one phosphate moiety, the content of AGI and AI in the total lipid was 91 mol% and 9 mol%, respectively. The other three polar lipids (one phosphoglycolipid and two glycolipids) were at trace levels.

The hydrocarbon chain prepared from the core lipid was identified by gas chromatography-mass spectrometry as C_{25} -isoprenoid, and the structure of the glycerol diether core lipid was identified disesterterpanyl (C_{25} , C_{25}) glycerol. AGI contained glucose and inositol and AI contained inositol as polar head group components. Glycerophosphoinositol was detected as a sole phosphorus-containing product of BCl_3 treatment of AGI or AI. The complete structures of the polar lipids and their core portions have been recently determined (Fig. 2.6) (Morii *et al.*, unpublished observation).

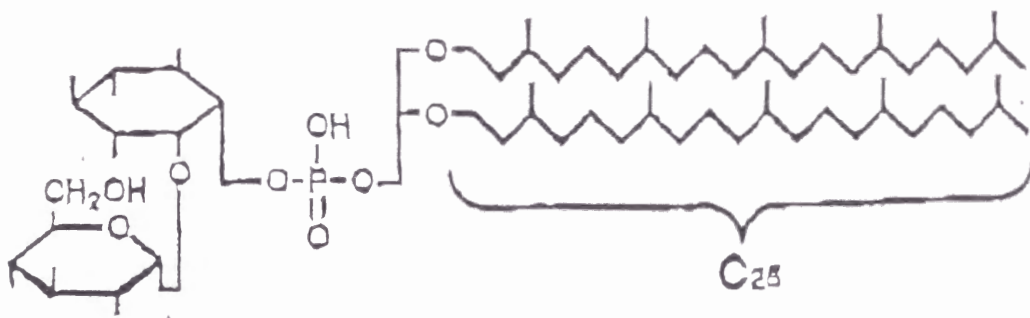


FIG. 2.6 Structure of the major diether phosphoglycolipid (AGI) from *Aeropyrum pernix* K1.

DNA base composition. The genomic DNA of isolate K1 had a G+C content of 67 mol% as calculated by direct analysis of the nucleosides.

16S-23S ITS sequences. Sequence alignment of 16S-23S ITS region between the four isolates, K1, OH2, TB1, and TB7, are shown in Fig. 2.7. The four isolates showed extremely high levels of similarity values (>99%) each other.

Phylogenetic analysis. The G+C contents of the 16S rRNA sequences employed in this analysis ranged from 62.7 to 68.7 mol%, and it was therefore assumed that they were relatively free from the variation of biased base compositions. Evolutionary distances were estimated by the comparison of representative archaeal 16S rRNA sequences. Isolate K1 represented considerably close relationships to *Pyrodictium occultum* and *Desulfurococcus mobilis* as indicated by estimated exchange of 7.9 and 11.0 per 100 positions, respectively. A phylogenetic tree was inferred on the basis of these evolutionary distances by neighbor-joining method (Fig. 2.8a). Isolate K1 represented a deep lineage within the Crenarchaeota. The cluster consisted of isolate K1, *Pyrodictium* spp., and *Desulfurococcus* spp. was specifically associated with the genus *Sulfolobus*, with a 95% confidence in a bootstrap analysis of this tree on the basis of 100 resamplings. However, the branching of isolate K1 more deeply than *Pyrodictium* spp. or *Desulfurococcus* spp. is given with a low level of confidence (24%). The level might be increased by isolating and sequencing the 16S rRNA of more species within the Crenarchaeota, a work which is currently carried out in our laboratories. Furthermore, this tree topology was contrasted to that inferred from

transversion distances, since the effect of compositional bias can be mitigated by confining analysis to transversion substitution only (Woese *et al.*, 1991). All sequences employed in this analysis have purine contents in the range of 56.1 to 57.2%. The resulting tree derived by transversion distance analysis also gave the same topology as that in Fig. 1.7a (data not shown).

On the other hand, the 23S rRNA sequences employed in the comparative analysis had G+C contents in the range of 51.8 to 67.6 mol%, while purine contents ranged from 55.1 to 56.6 mol%. Therefore, to avoid the artifact caused by compositional disparities, transversion evolutionary distances were estimated, and a phylogenetic tree was constructed with the distance matrix (Fig. 2.8b). Overall topology of the tree was similar to Fig. 1.8a and the reported ones based on 16S rRNA (Olsen *et al.*, 1994) with the exception of the convergence between the *Thermococcales* and the *Methanococcales-Methanobacteriales* cluster. The tree showed that isolate K1 was a deeply branching representative, which was located above the order *Sulfolobales* and below branching the family Pyrodictiaceae from the family Desulfurococcaceae within the order 'Igneococcales'. The bootstrap value of 100% among 1000 resampling supported the grouping of isolate K1, the family Pyrodictiaceae, and the family Desulfurococcaceae. Besides, the bootstrap value for the nodes branching the three species *Pyrodictium occultum*, *Desulfurococcus mobilis*, and *Staphylothermus marinus* were relatively high (85% and 87%), indicating the probability of their branching order. The particularly noteworthy was isolate K1 represented the second shortest, therefore primitive, lineage in this tree.

		10	20	30	40	50	
str.K1	1	CCGAAGCTAG	GCGGCAGCCG	GGGGCGGGCT	CGTGAAAGGC	CCGTGAGGGG	50
str.OH2	1	CCGAAGCTAG	GCGGCAGCCG	GGGGCGGGCT	CGTGAAAGGC	CCGTGAGGGG	50
str.TB1	1	CCGAAGCTAG	GCGGCAGTCG	GGGGCGGGCT	CGTGAAAGGC	CCGTGAGGGG	50
str.TB7	1	CCGAAGCTAG	GCGGCAGTCG	GGGGCGGGCT	CGTGAAAGGC	CCGTGAGGGG	50
		60	70	80	90	100	
str.K1	51	GTGTGGGCGC	CGTCCTAGAT	CCCATTATGC	TTATGTTTAG	CCATTGGGAT	100
str.OH2	51	GTGTGGGCGC	CGTCCTAGAT	CCCATTATGC	TTATGTTTAG	CCATTGGGAT	100
str.TB1	51	GTGTGGGCGC	CGTCCTAGAG	CCCATTATGC	TTATGTTTAG	CCATTGGGAT	100
str.TB7	51	GTGTGGGCGC	CGTCCTAGAT	CCCATTATGC	TTATGTTTAG	CCATTGGGAT	100
		110	120	130	140	150	
str.K1	101	GCAGCTCAGG	AGCCGTGGGC	TCCTGGGCGA	AGGGCGCGAG	GGGGCTAGCC	150
str.OH2	101	GCAGCTCAGG	AGCCGTGGGC	TCCTGGGCGA	AGGGCGCGAG	GGGGCTAGCC	150
str.TB1	101	GCAGCTCAGG	AGCCGTGGGC	TCCTGGGCGA	AGGGCGCGAG	GGGGCTAGCC	150
str.TB7	101	GCAGCTCAGG	AGCCGTGGGC	TCCTGGGCGA	AGGGCGCGAG	GGGGCTAGCC	150
		160	170	180	190	200	
str.K1	151	CGGGTGCTTA	CTGCCCGTTC	CGCAGTCCTG	CCCCGGGGGC	CCTGCTGTGA	200
str.OH2	151	CGGGTGCTTA	CTGCCCGTTC	CGCAGTCCTG	CCCCGGGGGC	CCTGCTGTGA	200
str.TB1	151	CGGGTGCTTA	CTGCCCGTTC	CGCAGTCCTG	CCCCGGGGGC	CCTGCTGTGA	200
str.TB7	151	CGGGTGCTTA	CTGCCCGTTC	CGCAGTCCTG	CCCCGGGGGC	CCTGCTGTGA	200
		210	220	230	240	250	
str.K1	201	GCGCGCCCTC	CATCCACCAG	CGGGGCGAGC	CGAGGGTCCG	GGGTGCCCGT	250
str.OH2	201	GCGCGCCCTC	CATCCACCAG	CGGGGCGAGC	CGAGGGTCCG	GGGTGCCCGT	250
str.TB1	201	GCGCGCCCTC	CATCCACCAG	CGGGGCGAGC	CGAGGGTCCG	GGGTGCCCGT	250
str.TB7	201	GCGCGCCCTC	CATCCACCAG	CGGGGCGAGC	CGAGGGTCCG	GGGTGCCCGT	250
		260	270	280	290	300	
str.K1	251	GCGGGCACAG	AGCCGCCCTA	GACAGCGCTC	CCTGCCCTTGG	GGGGCTGCCG	300
str.OH2	251	GCGGGCACAG	AGCCGCCCTA	GACAGCGCTC	CCTGCCCTTGG	GGGGCTGCCG	300
str.TB1	251	GCGGGCACAG	AGCCGCCCTA	GACAGCGCTC	CCTGCCCTTGG	GGGGCTGCCG	300
str.TB7	251	GCGGGCACAG	AGCCGCCCTA	GACAGCGCTC	CCTGCCCTTGG	GGGGCTGCCG	300
		310	320	330	340	350	
str.K1	301	GGGCGCCTG	TACCCCTGCAG	GCCGCCGGGC	CCTAGGGGGC	CCGCCCAGGC	350
str.OH2	301	GGGCGCCTG	TACCCCTGCAG	GCCGCCGGGC	CCTAGGGGGC	CCGCCCAGGC	350
str.TB1	301	GGGCGCCTG	TACCCCTGCAG	GCCGCCGGGC	CCTAGGGGGC	CCGCCCAGGC	350
str.TB7	301	GGGCGCCTG	TACCCCTGCAG	GCCGCCGGGC	CCTAGGGGGC	CCGCCCAGGC	350
		360	370	380	390	400	
str.K1	351	GGGCCTACG					400
str.OH2	351	GGGCCTACG					400
str.TB1	351	GGGCCTACG					400
str.TB7	351	GGGCCTACG					400

FIG. 2.7 Alignment of nucleotide sequences of 16S-23S ITS. Identical nucleotides are shaded.

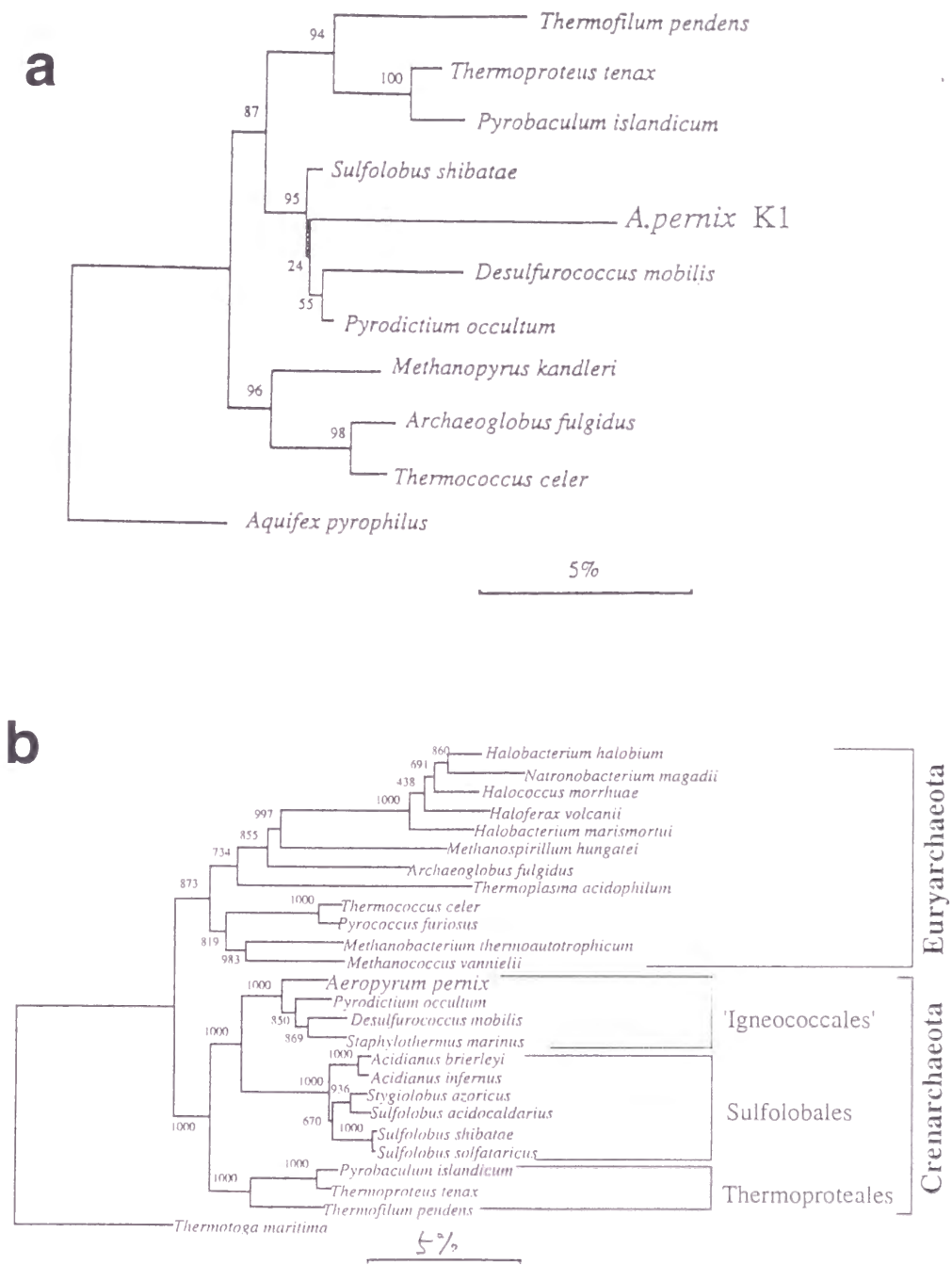


FIG. 2.8 Phylogenetic position of *Aeropyrum pernix* K1 within the representative archaea. (a) The tree inferred from the 16S rRNA evolutionary distances. Scale bar indicates 5 substitutions per hundred nucleotides. Numbers represent bootstrap values. Nucleotide sequence accession numbers employed in the phylogenetic analysis are as follows: *Aquifex pyrophilus*, M83548; *Archaeoglobus fulgidus*, Y00275; *Desulfurococcus mobilis*, M36474; *Methanopyrus kandleri*, M59932; *Pyrobaculum islandicum*, L07511; *Pyrodictium occultum*, M21087; *Sulfolobus shibatae*, M32504; *Thermococcus celer*, M21529; *Thermophilum pendens*, X14835; *Thermoproteus tenax*, M35966. (b) The tree inferred from the 23S rRNA transversion evolutionary distances. Scale bar represents 5 transversional

nucleotide substitutions per 100 sites. The sequence of *Thermotoga maritima* was used as an outgroup. Numbers at nodes represents bootstrap values for that node based on 1000 bootstrap resamplings. The sequences used in this calculations were obtained from DDBJ/EMBL/GenBank nucleotide sequence databases with the exception of the sequence of *Haloferax volcanii* and *Thermoproteus tenax*, which were extracted from LSU rRNA database (De Rijk *et al.*, 1997). Nucleotide sequence accession numbers are as follows: *Pyrodictium occultum*, M86626; *Desulfurococcus mobilis*, X05480; *Staphylothermus marinus*, M86623; *Acidianus brierleyi*, U32317; *Acidianus infernus*, U32318; *Sulfolobus acidocaldarius*, U05018; *Sulfolobus shibatae*, U32321; *Sulfolobus solfataricus*, U32322; *Stygiolobus azoricus*, U32319; *Thermofilum pendens*, X14835; *Pyrobaculum islandicum*, M86622; *Thermococcus celer*, M67497; *Pyrococcus furiosus*, M86627; *Methanococcus vanneilii*, X02729; *Methanobacterium thermoautotrophicum*, X15364 and X05482; *Thermoplasma acidophilum*, M32298 and M20822; *Archaeoglobus fulgidus*, M64487; *Methanospirillum hungatei*, M81323 and M61738; *Halobacterium marismortui*, X13738; *Halococcus morrhuae*, X05481; *Natronobacterium magadii*, X72495; *Halobacterium halobium*, X03407, X00872, and X01699; *Thermotoga maritima*, M67498. A total of 1242 unambiguous nucleotides were used in the analysis.

DISCUSSION

The four novel marine isolates represent the first strictly aerobic hyperthermoneutrophile growing optimally above 90 °C. On account of the presence of isopranyl ether lipids and its 16S and 23S rRNA sequences, isolate K1 belongs to the phylogenetic domain of the *Archaea* (De Rosa and Gambacorta, 1988; Koga *et al.*, 1993; Winker and Woese, 1991). In addition, its insensitivity to inhibitors of peptidoglycan synthesis such as ampicillin, vancomycin, and cycloserine provides further support that isolate K1 is a member of the *Archaea* (Hilpert *et al.*, 1981; Konig and Stetter, 1989). The other three isolates belong to one species because of the phenotypic similarities described above and high levels of similarity values in 16S-23S ITS sequence.

Isolate K1 does not fit into any of the described taxa within the *Archaea* domain. In particular, on the basis of its hyperthermophily and strict aerophily, isolate K1 is distinct from other archaea. Several members within the orders *Sulfolobales* and *Thermoplasmatales* are able to grow aerobically, depending on the culture conditions. *Sulfolobus* spp. have an irregular coccoid shape and capable of growing under aerobic conditions, either heterotrophically, or by the oxidation of S⁰ (Seegerer and Stetter, 1989a), chemolithoautotrophically. Among them, *Sulfolobus solfataricus* DSM1616^T grows optimally at 87 °C (Zillig *et al.*, 1980). *Acidianus* spp. and

Desulfurolobus spp. are facultatively aerobic S^0 oxidizers (Sererer and Stetter, 1989b; Zillig *et al.*, 1986). *Metallosphaera sedula* is strictly an aerobe, either by heterotrophy or chemolithoautotrophy (Huber *et al.*, 1989). *Thermoplasma* spp. are facultative aerobes (Langworthy and Smith, 1989). However, all of them have distinctly lower G+C contents, which range from 31 to 46 mol%, and grow optimally below pH 4.0, whereas isolate K1 has a G+C content of 67 mol% and grows within a pH range from 5.0 to 9.0. Furthermore, the other genera within the *Archaea* domain consist of obligate anaerobes.

Besides its unusual phenotypes described above, isolate K1 possesses unique physiological properties. Cultures of isolate K1 grow very rapidly, with a doubling time of about 200 min. This feature of rapid growth should make it an attractive organism for basic studies of aerobic hyperthermophile, and gram quantities of cells can be obtained easily in 2 days. In this connection, it was observed that thiosulfate greatly increased cell yields without H_2S production. I have also demonstrated thiosulfate-oxidizing activity in cell-free extracts of isolate K1 (Nomura and Sako, unpublished observation). These results suggest that isolate K1 has a highly developed thiosulfate-oxidizing system for energy generation. Similar to this, thiosulfate-oxidizing heterotrophs have been found in fresh water (Tuttle, 1980; Veinstein, 1976) and in a variety of marine habitats (Tuttle and Jannasch, 1972; Tuttle and Jannasch, 1973; Tuttle and Jannasch, 1977).

Isolate K1 contains novel lipids which for hyperthermophilic archaea are novel. Generally, most sulfur-dependent, thermophilic members of the *Archaea* such as *Sulfolobus* spp., *Thermoplasma* spp., and several species within the order *Thermoproteales* contain a caldarchaeol core along with a trace amounts of archaeol, except *Thermococcus celer*, which has only an archaeol core (De Rosa *et al.*, 1983). In contrast, the core lipid of isolate K1 consists only of archaeol. A similar core lipid is so far only known from the alkaliphilic halophile *Natronobacterium* SP8, but its content is only 1 to 6% (De Rosa *et al.*, 1987). Besides its core lipid structure, the lipid composition of isolate K1 is also unique. Only one phosphoglycolipid (AGI) is predominant. AGI seems to be a novel diether phosphoglycolipid, which contains both glucose and inositol.

The phylogenetic analysis based on 16S rRNA sequence indicates that isolate K1 is related to the genera *Pyrodictium* and *Desulfurococcus* within the Crenarchaeota. However, isolate K1 may be assigned a separate genus, since (i) the evolutionary distances between isolate K1 and each of these two archaeal genera (7.9 and 11.0 per 100 nucleotides, respectively) is significantly higher than that between *Pyrobaculum islandicum* and *Thermoproteus tenax* (2.0 per 100 nucleotides), each of which belongs to different archaeal genera, and (ii) *P. occultum* and *D. mobilis* are unable to grow in aerobic conditions (Stetter *et al.*, 1983; Zillig *et al.*, 1982).

The phylogenetic depth of isolate K1 within the *Archaea* domain raises problems about the origin and evolution of oxygen respiration on the Earth. In the traditional view, the Earth's archaean atmosphere has been presumed to be highly reducing (Kasting, 1987), and the aerobic organisms are presumed to have evolved after the oxygenic photosynthetic organisms became active. In contrast, Towe (1990) suggested that the archaean atmosphere contained a low (0.2 to 0.4%) but stable proportion of oxygen as a component of the global carbon cycle. Furthermore, based on sulfur isotope analyses of microscopic-sized grains of pyrite, it is likely that sulfur-reducing organisms converted seawater sulfate into pyrites at least 3.4 billion years ago (Ohmoto *et al.*, 1993). This observation implies that by about 3.4 billion years ago the oceans were rich in sulfate and the atmosphere contained appreciable amounts of free oxygen. Interestingly, *Aquifex pyrophilus* and *Pyrobaculum aerophilum*, which represent the deep and short lineages within the *Bacteria* and *Archaea* domains, respectively, are able to grow in microaerophilic conditions (Huber *et al.*, 1992; Volkel *et al.*, 1993). It is still unclear when the oxygen respiring metabolism was acquired and how it was developed, but the origin of aerobic organisms might be more ancient than is commonly believed. Further study of the aerobic respiration of isolate K1 could elucidate these issues.

The ecological distribution of this new organism is at present unknown. Isolate K1 was obtained from a geothermally heated marine solfataric biotope on a seashore of Kodakara Island, Kagoshima, Japan. The inability of isolate K1 to grow appreciably at the expense of sulfide is puzzling in light of the abundance of sulfide at this site, but it could actually be a reflection of

the unusual growth conditions this organism experiences in situ. Close relatives of isolate K1 were isolated from a shallow submarine hydrothermal vent in Tachibana Bay, Nagasaki, Japan (32° 45' N, 130° 10' E) and coastal hot springs at Obama, Nagasaki, Japan (32° 43' N, 130° 12' E). There would seem to be several ways that geothermally heated water could contain, or become contaminated with, oxygen in these hydrothermal environments. Therefore, isolate K1 and its relatives may thrive in widespread oxidative zones within marine hydrothermal environments.

On the basis of the strictly aerobic mode of life, a distinct G+C content of 67 mol%, the idiosyncratic lipid composition, and the low level of similarity of the 16S and 23S rRNA sequences with any member of the *Archaea*, I proposed a name *Aeropyrum pernix* gen. nov., sp. nov. for the four isolates. The type strain of *A. pernix* is K1, which has been deposited in the Japan Collection of Microorganisms, The Institute of Physical and Chemical Research (RIKEN), Saitama, Japan, with the accession number of JCM 9820.

Chapter 3

Intraspecies Polymorphism in the Single rRNA Operon of *A. pernix*, Implying the Presence of Hotspots for Intron-insertion

INTRODUCTION

Reflecting the scope of their early paradigm, introns and RNA splicing were considered as characteristics that distinguish eukaryotes from prokaryotes (bacteria and archaea). However, studies in the last decades elucidated that several prokaryotic genes encoding stable RNAs, such as rRNAs and tRNAs, are scattered with introns despite with the exceedingly low frequency. The first evidence for an intron in prokaryotic chromosome came from the study of the 23S rRNA gene in the anaerobic hyperthermophilic archaeon *Desulfurococcus mobilis* (Kjems and Garrett, 1985). It circularizes after splicing and forms a normal 5', 3'- phosphodiester bond at the ligation junction. This 622-nt circular intron contains a putative ORF possessing LAGLI-DADG motifs (Kjems and Garrett, 1988) that are common to endonucleases and maturases encoded by group I introns and inteins found in mitochondria, chloroplasts, and nuclei of lower eukaryotes (Mueller *et al.*, 1993). Subsequently, archaeal ORF-containing introns were also detected in the 23S rRNA gene of *Pyrobaculum organotrophum* (Dalgaard and Garrett, 1992), and the 16S rRNA gene of *Pyrobaculum aerophilum* (Burggraf *et al.*, 1993). The other characterized archaeal introns from tRNA-encoding genes (Datta *et al.*, 1989; Kaine, 1987; Kaine *et al.*, 1983; Wich *et al.*, 1987) and other rRNA-encoding genes (Kjems and Garrett, 1991) are small, 15 to 110 nt. On the other hand, introns of bacterial origin have also been found in tRNA genes in cyanobacterial and proteobacterial chromosomes (Ferat and Michel, 1993; Kuhsel *et al.*, 1990; Reinhold-Hurek and Shub, 1992; Xu *et al.*, 1990), in *td*, *nrdB*, or *sunY* genes of the T-even phages (Belfort, 1990; Chu *et al.*, 1984), and in DNA polymerase gene (*g31*) of the *Bacillus subtilis* phage SPO1 (Goldrich-Blair *et al.*, 1990). Considering that the large number of nucleotide sequences of prokaryotic

genes are currently available, such cases may represent a rather idiosyncratic phenomenon. Nevertheless, eukaryotic genomes can no longer be considered to make a monopoly of introns.

In the previous chapter, a marine aerobic hyperthermophilic archaeon *Aeropyrum pernix* was isolated from several marine hydrothermal environments in Japan. In the process of the previous study on phylogenetic characterization of *A. pernix* K1, we obtained a nearly complete nucleotide sequence of the 16S rRNA gene retrieved from the chromosomal DNA by PCR and unexpectedly found an intervening sequence. This observation promoted us to conduct an extensive investigation of the organization and nucleotide sequence of the rRNA operon (*arnSL*) of this organism. In this chapter, I report the occurrence of three introns in the single copy of the rRNA operon from *A. pernix* K1, their structural features, and intraspecies polymorphism in rRNA operon of *A. pernix*. I also discuss the implication of the presence of hotspots for intron-insertion within these region.

MATERIALS AND METHODS

Strains, vectors, and culture conditions. *A. pernix* K1(JCM9820), OH2, TB1, and TB7 were cultivated at 90°C with vigorous shaking as described in Chapter 2. *E. coli* INV α F' (Invitrogen) was used for plasmid construction. This strain was routinely grown at 37°C in Luria-Bertani (LB) broth and agar. Ampicillin (50mg/liter) was included in the medium to select for cells harboring ampicillin-resistant plasmids. 5-Bromo-4-chloro-3-indolyl- β -D-galactopyranoside (20mg/liter) was used to identify recombinant plasmids with DNA insertions that inactivated β -galactosidase activity in *E. coli* INV α F'. The vectors pGEM-3Zf(+) (Promega), or pCR2.1 (Invitrogen) was used for subcloning.

Recombinant DNA techniques. The procedure used for isolation of chromosomal DNA of *A. pernix* was carried out as described in Chapter 2. Plasmid DNA was isolated from *E. coli* with an alkaline-sodium dodecyl sulfate (SDS) cell lysis miniprep protocol(Sambrook *et al.*, 1989). DNA restriction digests, ligations, transformations, and other DNA manipulations

were carried out by using standard methods (Sambrook *et al.*, 1989), as specified by the manufacturers. PCR was performed in a Perkin-Elmer Cetus DNA thermal cycler. Approximately 50 ng of *A. pernix* chromosomal DNA was amplified with AmpliTaq DNA polymerase (Perkin-Elmer) and 1 μ M of each primer. Reactions were subjected to multiple round (30-35) of denaturation for 90 s at 96°C, annealing for 1 min at the optimal temperature for each set of primers, and extension for 2 min at 72°C. A final extension step of 72°C for 15 min was completed at the end of the multiple cycles. The plasmids constructed by PCR cloning were shown in Table 2.1.

Probe construction and labeling. Digoxigenin-labeled antisense RNA probes were produced by run-off transcription using T7 RNA polymerase. Probe S was derived from plasmid pNA4 which is the cDNA clone of the mature 16S rRNA (Chapter 2; Sako *et al.*, 1996). Probes 1, 2, 3, and 4 were prepared from plasmids pNC2, pNB5, pIP231, and pIP232, respectively. The plasmids were linearized with appropriate restriction enzymes prior to in vitro transcription and then used for synthesis of digoxigenin-dUTP-labeled RNAs with T7 RNA polymerase according to the recommendations of the supplier (Boehringer, Mannheim, Germany). The amount of digoxigenin-labeled probe was estimated by direct detection of the labeled RNA probe with anti-digoxigenin-alkaline phosphatase.

The additional three 5'-digoxigenin labeled oligonucleotide probes were prepared for Southern blot analysis of introns I δ , I ϵ , I ζ . The sequences were as follows: Probe 5 (I δ specific), 5'-CAATTCTACTTAACCACGTC-3'; Probe 6 (I ϵ specific), 5'-AAGAGATCTTAAGGTCTTCT-3'; Probe 7 (I ζ specific), 5'-CCGCCTTATATGCTATGACC-3'.

Southern blot analysis. For the determination of the rRNA operon copy number, chromosomal DNA of *A. pernix* K1 was digested with *Eco*RI, *Hind*III, *Bam*HI, or *Pst* I, and then electrophoresed on a 0.8% agarose gel. DNA was depurinated, denatured, and transferred to a Biodyne Plus nylon membrane (Pall) by using 20 $\times$ SSC (3M NaCl, 0.3M sodium citrate, pH 7.0), and then hybridized with the digoxigenin-labeled RNA probes overnight at 60°C in buffer, containing 50% formamide, 5 $\times$ SSC, 0.1% *N*-lauroylsarcosine, 0.02% SDS, and 2% blocking reagent (Boehringer). Detection was performed according to the protocol obtained from Boehringer with CSPD (Boehringer)

TABLE 3.1. Strains and plasmids used in this chapter

Strain or plasmid	Description	Source or reference
<i>E. coli</i>		
INV α F'	<i>recA1 lacZ</i> Δ M15	Invitrogen
<i>Aeropyrum pernix</i>		
K1	Natural isolate of <i>A. pernix</i> JCM9820	Chapter 2 ; Sako <i>et al.</i> , 1996
OH2	Natural isolate	Chapter 2
TB1	Natural isolate	Chapter 2
TB7	Natural isolate	Chapter 2
Plasmids		
pGEM-3Zf(+)	Cloning vector; Amp ^r	Promega
pCR2.1	Cloning vector; Amp ^r Km ^r	Invitrogen
pRU363	4.5-kb <i>Bam</i> HI- <i>Eco</i> RI chromosomal DNA fragment containing 5'-portion of <i>arnS</i> and its upstream region (nt ^a 1 to 4477 ^b) inserted into pGEM-3Zf(+)	This chapter
pRD542	6.2-kb <i>Eco</i> RI- <i>Sal</i> I chromosomal DNA fragment containing 3'-portion of <i>arnS</i> and entire <i>arnL</i> (nt 4472 to 10717) inserted into pGEM-3Zf(+)	This chapter
pNB5	348-bp PCR fragment containing the partial internal region of I α (nt 4515 to 4862) inserted into pCR2.1	This chapter
pNC2	663-bp PCR fragment containing the partial internal region of ES1 (nt 3380 to 4042) inserted into pCR2.1	This chapter
pIP231	45-bp PCR fragment containing the partial internal region of I β (nt 7278 to 7322) inserted into pCR2.1	This chapter
pIP232	228-bp PCR fragment containing the partial internal region of I γ (nt 8439 to 8666) inserted into pCR2.1	This chapter
pNA4	1444-bp reverse transcriptase-mediated PCR fragment containing entire cDNA of mature 16S rRNA inserted into pCR2.1	Chapter 2

^ant, nucleotides.

^bThe nucleotide positions refer to the numbering for the sequence submitted to DDBJ, EMBL, GenBank under accession number AB008745.

as the substrate. The membrane was exposed to X-ray film for appropriate periods at room temperature.

For the investigation of the distribution and copy number of the introns, chromosomal DNA samples from the four strains of *A. pernix* (K1, OH2, TB1, and TB7) were digested with *Pst* I and *Sal* I. The blotting and hybridization were carried out as described above.

Construction and screening of the *A. pernix* K1 genomic libraries.

A 4.5-kb of *Bam*HI-*Eco*RI fragment, including the 5'-portion of the *arnS* gene and its upstream region, was cloned as follows. Chromosomal DNA of *A. pernix* K1 was digested with *Bam*HI and *Eco*RI, and the fragments were ligated to the pGEM-3Zf(+) vector treated with *Bam*HI and *Eco*RI. The recombinant clones containing target locus were selected by in situ hybridization of colonies with both probe S and probe 1 according to standard procedures (Sambrook *et al.*, 1989).

A 6.2-kb of *Eco*RI-*Sal* I fragment, containing the 3'-portion of the *arnS* gene and the *arnL* gene, was cloned in the following manner. Chromosomal DNA of *A. pernix* K1 was digested with *Eco*RI and *Sal* I, and the restriction fragments were ligated to the pGEM-3Zf(+) vector treated with *Eco*RI and *Sal*I. The recombinant clones containing target locus were selected by in situ hybridization of colonies with probe S as well as probe 2.

PCR cloning. The PCR fragment primed by appropriate oligonucleotides (Achenwach and Woese, 1995) with chromosomal DNA of *A. pernix* OH2, TB1, or TB7 as a template was cloned directly into the pCR2.1 vector. The resulting plasmids, which contains the *arnSL* locus, were sequenced.

DNA sequencing. DNA fragment cloned into plasmids were sequenced in both directions by the dideoxy chain termination method described by Sanger *et al.* (1977), using a DyeDeoxy terminator cycle sequencing FS Ready Reaction kit (Perkin-Elmer) and a ABI 373A automated DNA sequencer (Applied Biosystems). Each of the plasmid DNA inserts was treated with exonuclease III and mung bean nuclease to create nested deletion series for sequencing (Henikoff, 1984). For confirmation of intron sequences within the *arnS* and *arnL* genes (Fig. 3.2), the sequencing ladder was generated with DIG Taq DNA sequencing kit (Boehringer) and 5'-digoxigenin labeled oligonucleotide primers. The sequences of oligonucleotides used as primers were as follows:

P1, 5'-AAACTATCAAGAGTTTG TAA-3', complementary to the I α sequence from bases 53 to 72 downstream of the 5' splice site of the I α ; P2, 5'-GGTTCCCGGCGTTGACTCCA-3', complementary to the ES2 sequence from bases 54 to 73 downstream of the 3' splice site of the I α ; P3, 5'-TC TACCCTATCCTGGCATGA-3', complementary to the I β sequence from bases 65 to 84 downstream of the 5' splice site of the I β ; P4, 5'-CTCGGCGGCC GG CTTAAGCC-3', complementary to the EL2 sequence from bases 53 to 72 downstream of the 3' splice site of the I β ; P5, 5'-CTCCAGTAGACACT GATCCA-3', complementary to the I γ sequence from bases 78 to 97 downstream of the 5' splice site of the I γ ; P6, 5'-AGTGGGGACCT CGTTGACCC-3', complementary to the EL3 sequence from bases 45 to 64 downstream of the 3' splice site of the I γ . 5'-end of the each oligonucleotide was labeled with digoxigenin. In order to eliminate ambiguities due to band compression, dGTP was occasionally replaced by 7-deaza-dGTP in the sequencing reactions.

Direct RNA sequencing. The mature 16S and 23S rRNA were extracted and purified as described in Chapter 2. RNA sequencing across intron insertion sites was performed as described by Qu *et al.* (1983) with minor modifications. Superscript II RNase H⁻ reverse transcriptase (Gibco BRL) and appropriate 5'-digoxigenin labeled oligonucleotide primers were used to initiate cDNA synthesis on either mature 16S or 23S rRNA template. Reaction products were resolved on sequencing gels alongside M13 sequencing reactions of the corresponding plasmid DNA templates.

Computer analysis. Alignment of primary sequence data and calculation of degree of sequence identity were based on the results of FASTA (Pearson and Lipman, 1988) and Clustal W (Thompson *et al.*, 1994) analyses. RNA secondary structures were inferred according to Zuker and Stieger(1982).

Nucleotide sequence accession number. The sequence of the rRNA operon (*arnSL*) of *A. pernix* K1 will appear in the DDBJ, EMBL, GenBank nucleotide sequence databases with the accession number AB008745.

RESULTS

Organization and sequence of the single rRNA operon (*arnS-arnL*) in *A. pernix* K1. In this section, the nucleotide positions refer to the numbering for the sequence accession No. AB008745. Recombinant clones pRU363 and pRD542 served as the primary source of DNA for sequencing of the rRNA operon of *A. pernix* K1. A 10.7 kb *Bam*HI-*Sal*I segment of the genome of *A. pernix* K1 containing the 16S and 23S rRNA genes (designated as *arnS* and *arnL*, respectively) was sequenced. The determined sequence also included nearly 3.4 kb of 5' flanking region of *arnS*, *arnS-arnL* internal transcribed spacer region(ITS) containing no tRNA gene, and nearly 1.0 kb of 3' flanking region of *arnL*. The relevant portion is depicted in Fig. 3.1A. Extrapolation of transcription studies in archaea (Garrett *et al.*, 1991) makes it highly likely that the *A. pernix* K1 rRNA gene cluster is in fact an operon, although Northern blot analyses failed to reveal the presence of a primary transcript. A similar gene organization of the rRNA operon has been reported for those of crenarchaea *Sulfolobus acidocaldarius* (Durovic and Dennis, 1994) and *Desulfurococcus mobilis* (Kjems and Garrett, 1987; Leffers *et al.*, 1987).

The *arnS* and *arnL* spanned DNA segments of 2143 bp and 4158 bp, respectively. Both genes were in size of extraordinarily large, considering that the most prokaryotic 16S and 23S rRNA genes ranges from 1.4 to 1.5 kbp, and from 2.9 to 3.0 kbp, respectively. Sequence comparison with so far known counterparts from archaea and bacteria revealed that an intervening sequence of 699 bp, designated here as $I\alpha$ (positions 4254 to 4952) was inserted into *arnS*, and two distinct intervening sequences resided within *arnL* and contained 202 bp and 575 bp, designated $I\beta$ (positions 7118 to 7319) and $I\gamma$ (positions 8439 to 8456), respectively. The five discontinuous segments encoding rRNAs were designated as follows: ES1 (positions 3378 to 4253) and ES2 (positions 4953 to 5520) in *arnS*, and EL1 (positions 5878 to 7117), EL2 (positions 7320 to 8174), and EL3 (positions 8750 to 9735) in *arnL*.

In the upstream region flanking the *arnS* gene, two putative promoter signals were detected, which correspond to the AT-rich TATA-like sequence of the archaeal box A element (Hain *et al.*, 1992; Palmer and Daniels, 1995; Reiter *et al.*, 1988; Thomm and Wich, 1988)(CTTATA; positions 3260 to

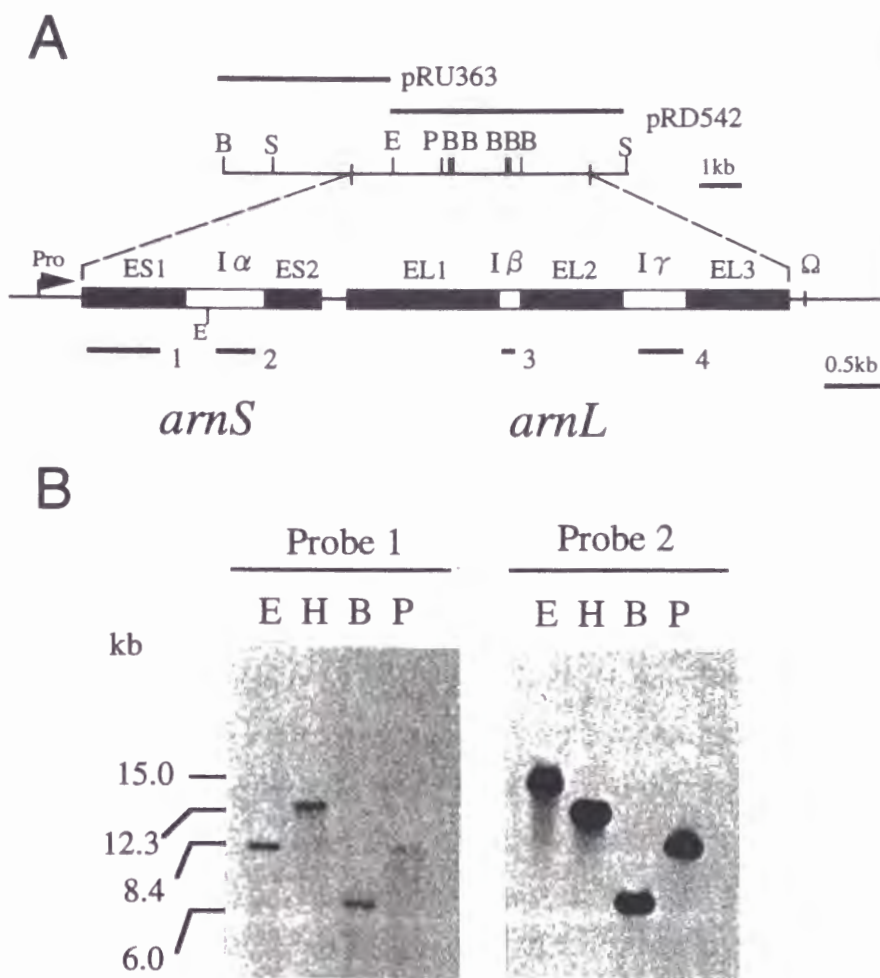


FIG. 3.1. (A) Schematic representation of the organization of the 16S and 23S rRNA gene (*arnS* and *arnL*, respectively) locus on the chromosome of *Aeropyrum pernix* K1. Upper, restriction map of the sequenced 10.7 kbp *Bam*HI-*Sal*I region harbouring the rRNA operon. Thick lines indicate DNA fragments cloned into the two plasmids pRU363 and pRD542. The nucleotide sequence accession number of this region is AB008745. Lower, gene map of the subregion encompassing the rRNA operon as deduced from sequence data. Filled areas denote exons of *arnS* (ES1 and ES2) and *arnL* (EL1, EL2, and EL3), while open boxes denote introns ($I\alpha$, $I\beta$, and $I\gamma$). Antisense RNA fragments numbered 1 to 4 were used for Southern and/or Northern blot analyses as probes. Pro denotes the putative promoter sequence, and the possible termination site of *arnS-arnL* transcript is also denoted with Ω . Size scales, in terms of base pairs (bp), are given at the right. (B) Southern blot analyses of the rRNA operon of *A. pernix* K1. Probe 1 or 2 was hybridized to restriction digests of the chromosomal DNA. Molecular size (in kb) are indicated on the left. Restriction endonuclease abbreviations: B, *Bam*HI; E, *Eco*RI; H, *Hind*III; P, *Pst*I; S, *Sal*I.

3265) and a box B element (CAGGA; positions 3290 to 3294) located 24 bp downstream of the box A. In contrast, GC-rich inverted repeat (positions 9838 to 9843, and 9851 to 9856) was located 103 bp downstream of *arnL*, suggestive of a transcriptional terminator.

To determine the copy number of the rRNA operon harbored on the chromosome of *A. pernix* K1, Southern blot analyses were performed (Fig. 3.1B). Four restriction enzymes, *EcoRI*, *BamHI*, *HindIII*, and *PstI* were employed in the analyses, one of which, *EcoRI*, cleaved within I α . Hybridization was carried out by using either probe 1 or 2. Single bands ranging from 6.0 to 15.0 kbp were observed in all lanes, suggesting that the genome of *A. pernix* K1 exhibits a single copy of the rRNA operon. The *EcoRI* site within I α , directly upstream of probe 2, justify the different size of signals in the *EcoRI*-treated lanes. In this context, it might be concluded that the single, therefore transcriptionally active, rRNA operon always contains the aforementioned three intervening sequences, I α , I β , and I γ .

The three introns in the rRNA operon of *A. pernix* K1 are excised during rRNA maturation. In order to examine whether the intervening sequences were excised post-transcriptionally and to confirm that the flanking RNA segments encoding rRNAs were ligated to yield mature 16S and 23S rRNAs, the appropriate mature rRNA species were directly sequenced by reverse transcriptase across the putative insertion sites of the intervening sequences. Figure 3.2 shows the results alongside the *arnS* and *arnL* sequences correspond to the boundaries at the DNA level. All of the three intervening sequences were absent from the mature 16S and 23S rRNAs. Furthermore, the lack of termination in reverse transcription at each junction implies that the split rRNA segments (ES1/ES2, and EL1/EL2/EL3) were post-transcriptionally ligated in normal 5', 3'-phosphodiester bonds to produce two mature rRNAs, 16S and 23S rRNA of 1444 nt and 3081 nt, respectively. Therefore, it is concluded that all the three intervening sequences are introns.

Intraspecies polymorphism in rRNA intron-insertion of *A. pernix*. The additional sequence analyses of the single rRNA operon in the other three isolates (OH2, TB1, and TB7) of *A. pernix* showed that a total of 7 introns is present within the loci. Based on the sequence differences observed in the introns, we have assigned these introns into six types (I α ; I β ; I γ ; I δ , 62nt ;

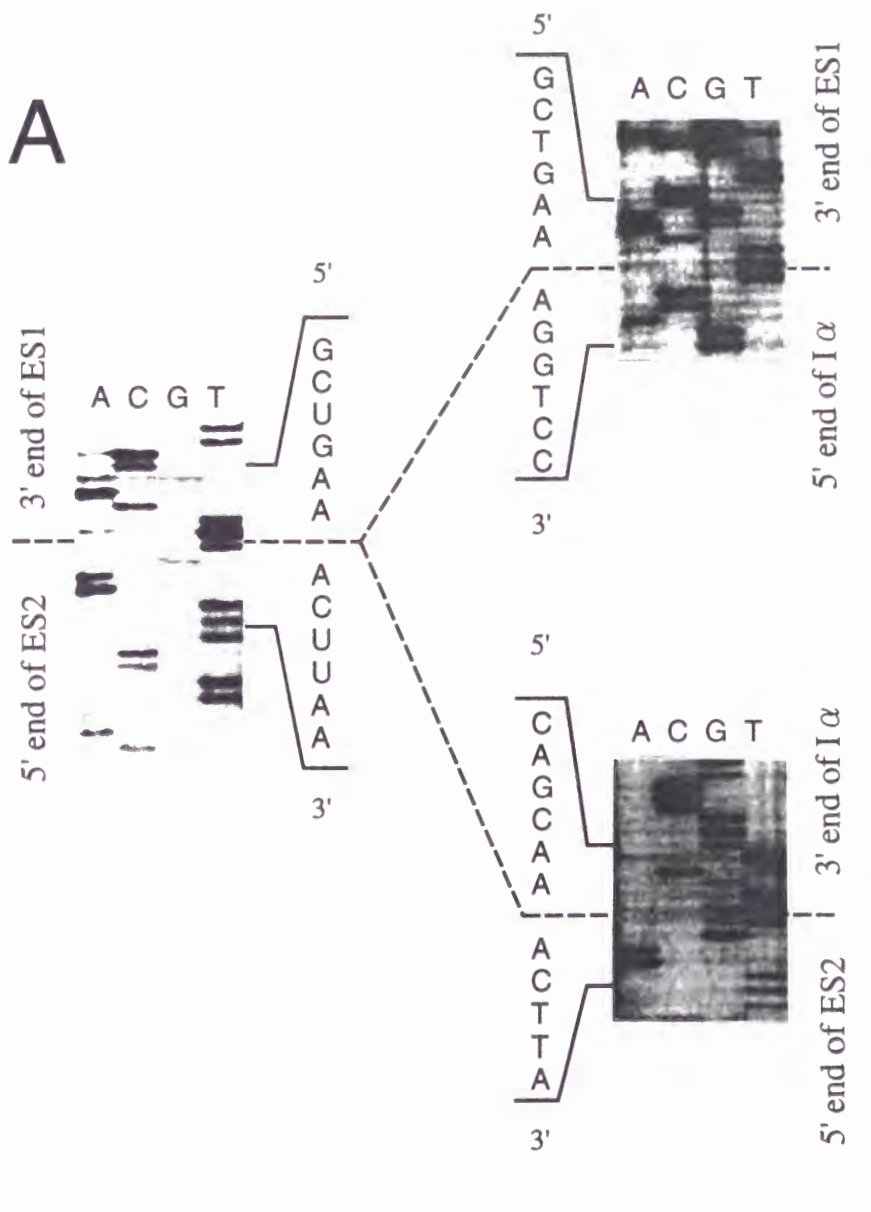


FIG. 3.2. rRNA/rDNA sequencing gels demonstrating that intron excision and exon ligation occur for introns I α (A), I β (B), and I γ (C). Left panels are of mature rRNA sequencing gels generated by reverse transcriptase dideoxy sequencing primed from oligonucleotide specific for the sequence spanning insertion site of introns, whereas right panels are of rDNA sequencing gels spanning the positions of 5' flanking exon/intron junction (top) and intron/3' flanking exon (bottom). The actual rRNA/rDNA nucleotide sequences are shown to facilitate comparison with Fig. 3.5.

B

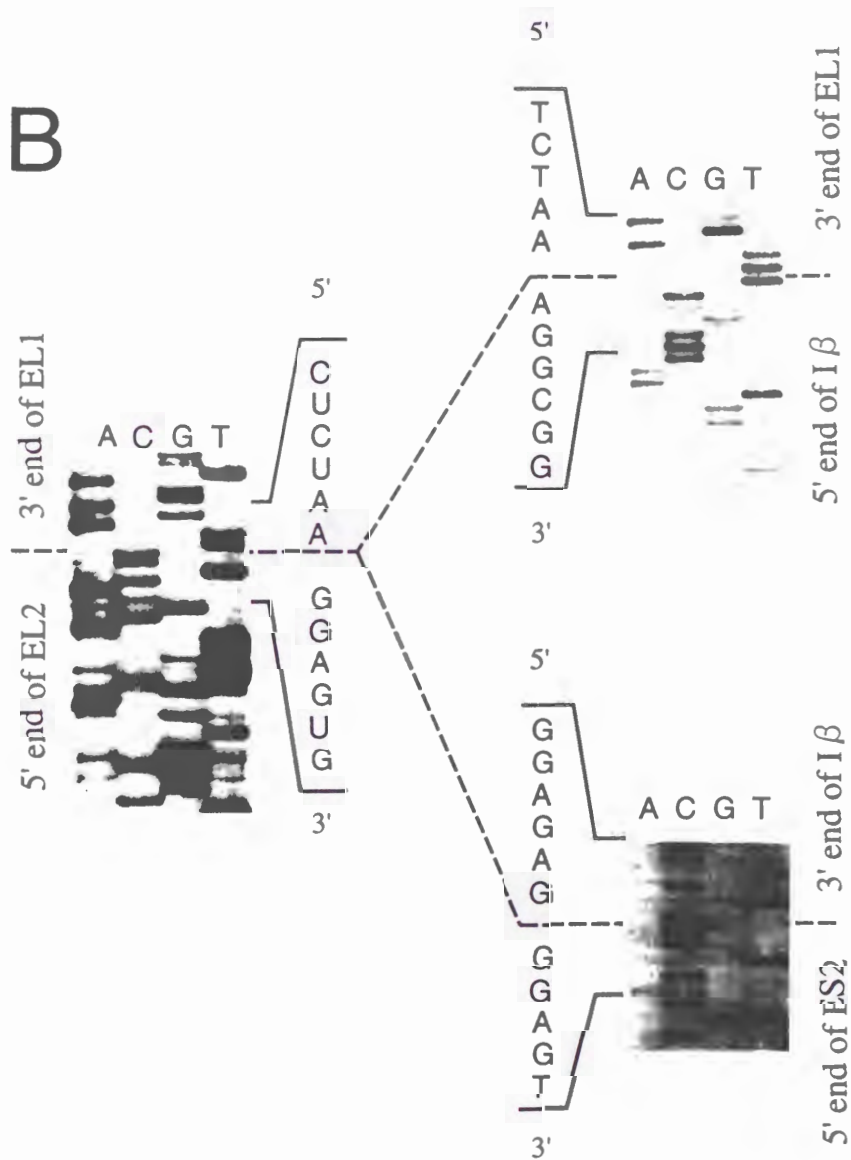


FIG. 3.2. (continued.)

I ϵ , 124nt ; and I ζ , 57nt)(Fig. 3.3). Interestingly, I α and I δ precisely occupy the same insertion sites in the *arnS* genes, although they differ in length and primary structure. As is the case observed between I γ and I ζ in the *arnL* genes. These results lead us to suppose the presence of at least four hotspots (I α / I δ , I β , I γ / I ζ , and I ϵ) for intron-insertion within the *arnSL* locus of *A. pernix*.

FIG. 3.3. Schematic representation for intraspecies polymorphism of the rRNA operon in *A. pernix*.

Structural features of the introns. The six rRNA introns in *A. pernix* are highly diversified in length and potential secondary structure. There are no obvious primary sequence similarities either within the introns themselves or at the splice sites in the comparison with other archaeal counterparts so far characterized. However, they all are remarkably AT-rich (42.4, 43.1, 41.5, 48.3, 46.3, and 53.5 mol% G+C for I α , I β , I γ , I δ , I ϵ , and I ζ , respectively) compared with the exons (67.7 and 69.2 mol% G+C for mature 16S and 23S rRNA, respectively) and contain 16 to 20 bp of GC-rich terminal inverted repeat sequences which are predicted to form helical structures in the nascent transcripts. Their insertion site is of particular interest, because the three

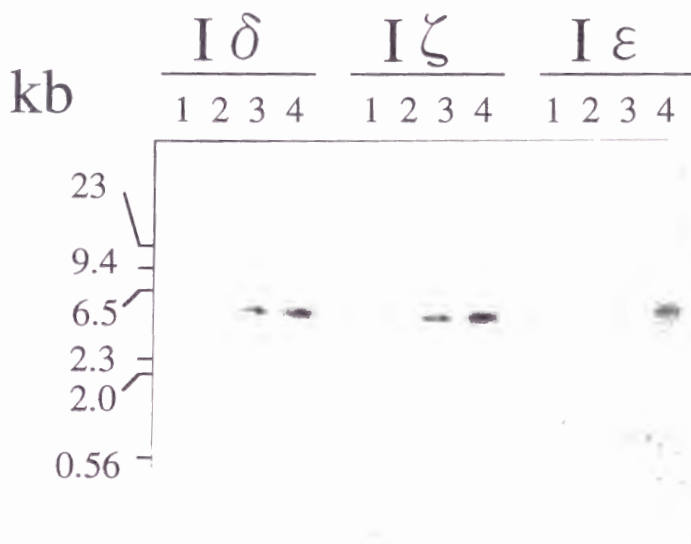
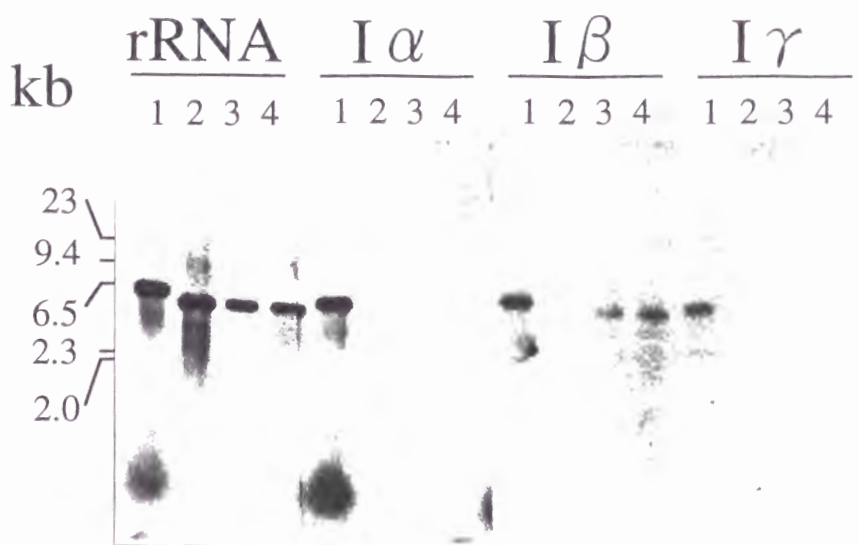


FIG. 3.4. Southern blotting analyses demonstrating the copy number of the *arnSL* operon and distribution of the introns. Oligonucleotideprobe which is specific to the 5' portion of the *arnS* or each intron was hybridized to *Pst* I and *Sal* I- digested of the chromosomal DNA from *A. pernix* strains. Molecular size (in kb) are given on the left. Lanes 1, 2, 3, and 4 are the *Pst* I and *Sal* I restricted chromosomal DNA from *A. pernix* K1, OH2, TB1, and TB7, respectively.

introns reside close to the sites for tRNA-binding or interaction of elongation factor, implying that both 30S and 50S ribosomal subunit could neither assemble completely nor function until they are spliced.

(i) $I\alpha$. The intron $I\alpha$ consists of 699 nt and lies within region of highly conserved primary and secondary structure in central domain of 16S rRNA (Noller, 1984), detached from the insertion site of *Pb. aerophilum* 16S rRNA intron (Burggraf *et al.*, 1993). The insertion site is located near an site that has been implicated in A-site binding of tRNA (Moazed and Noller, 1986; Moazed and Noller, 1990). The putative secondary structure of $I\alpha$ and surrounding rRNA agree with the rigidly defined bulge-helix-bulge motif, 3-base loops on opposite strands separated by 4-base pairs helix (Chant and Dennis, 1986; Dennis, 1991; Garrett *et al.*, 1991; Kjems and Garrett, 1987)(Fig. 3.5A). As shown in Fig. 3.6A, $I\alpha$ contained ORF, designated here as $I\alpha$ ORF, which starts at position 4285 (AB008745), continues through the 3' junction and extends 52 nt into the ES2 (position 5004; AB008745). Given that intronic RNA $I\alpha$ circulizes after splicing without nucleotide insertions and deletions as observed in the circular archaeal introns (Dalgaard and Garrett, 1992; Kjems and Garrett, 1988), it provides no stop codon in the same reading frame upstream of the start codon. Furthermore, Northern blot analyses (Chapter 4; Figs. 4.2 and 4.3) failed to reveal the presence of precursor 16S rRNA containing $I\alpha$, while free intronic RNA species $I\alpha$ was shown to retain stably in the cell. These observations raised the possibility that the linear intron L- $I\alpha$ is used as a ORF-encoding template for translation inspite of the absence of stop codon. Assuming it to be true, the deduced product is a 25-kDa, 221-amino acid polypeptide. Putative ribosome-binding site (GGGAGGG) precede $I\alpha$ ORF, however, no obvious transcriptional promoter and terminator sequences is found within $I\alpha$.

(ii) $I\beta$. The intron $I\beta$ is 202 nt long (Fig. 3.6B) and map within domain II of 23S rRNA. The insertion site has not been documented so far in other systems and coincides closely with nucleotide which has been to be the site of action of the EF-G-dependent GTPase inhibitor thiostrepton in *E. coli* (Thompson *et al.*, 1982). At intron/exon junction, $I\beta$ and surrounding rRNA could fold into an aberrant bulge-helix-bulge structure shown in Fig. 3.5B. No possible open reading frame was found in $I\beta$.

(iii) **I γ** . The intron I γ (575nt) interrupts the domain IV of 23S rRNA at a position close to that of large subunit rRNA introns of an archaeon *D. mobilis* (Kjems and Garrett, 1988) and an eukaryote *Physarum polycepharum* (Otsuka *et al.*, 1983). The insertion site is also close to a 23S rRNA site that has been implicated in P-site binding of tRNA (Moazed and Noller, 1989). The putative secondary structure of I γ at the intron/exon junction does not fit into the typical bulge-helix bulge motif as described in I α (Fig. 3.5C). The free standing ORF (411nt; designated as I γ ORF), starting with a GTG start codon at position 8253 (AB008745) and ending with a TAA stop codon at position 8666 (AB008745), encodes a deduced 15-kDa, 137-amino acid polypeptide (Fig. 3.6C). A potential ribosome-binding site, GAGGA, is located within the upstream of I γ ORF and is spaced 9 nt from the start codon. A computer search of the sequence of I γ failed to reveal any canonical transcriptional promoter and terminator-like sequences.

(iv) **I δ** . The intron I δ is 62 nt long (Fig. 3.6D) and, the insertion site is same as that of I α . At intron/exon junction, I δ and surrounding rRNA could fold into an typical bulge-helix-bulge structure shown in Fig. 3.5D. No possible open reading frame was found in I δ .

(v) **I ϵ** . The intron I ϵ is 124 nt long (Fig. 3.6E) and exist in 16S rRNA of *A. pernix* TB1 and TB7. At intron/exon junction, I ϵ and surrounding rRNA could fold into an typical bulge-helix-bulge structure shown in Fig. 3.5E. No possible open reading frame was found in I ϵ .

(vi) **I ζ** . The intron I ζ is 57 nt long (Fig. 3.6F) and, the insertion site is same as that of I γ . At intron/exon junction, I ζ and surrounding rRNA could fold into an typical bulge-helix-bulge structure shown in Fig. 3.5F. No possible open reading frame was found in I ζ .

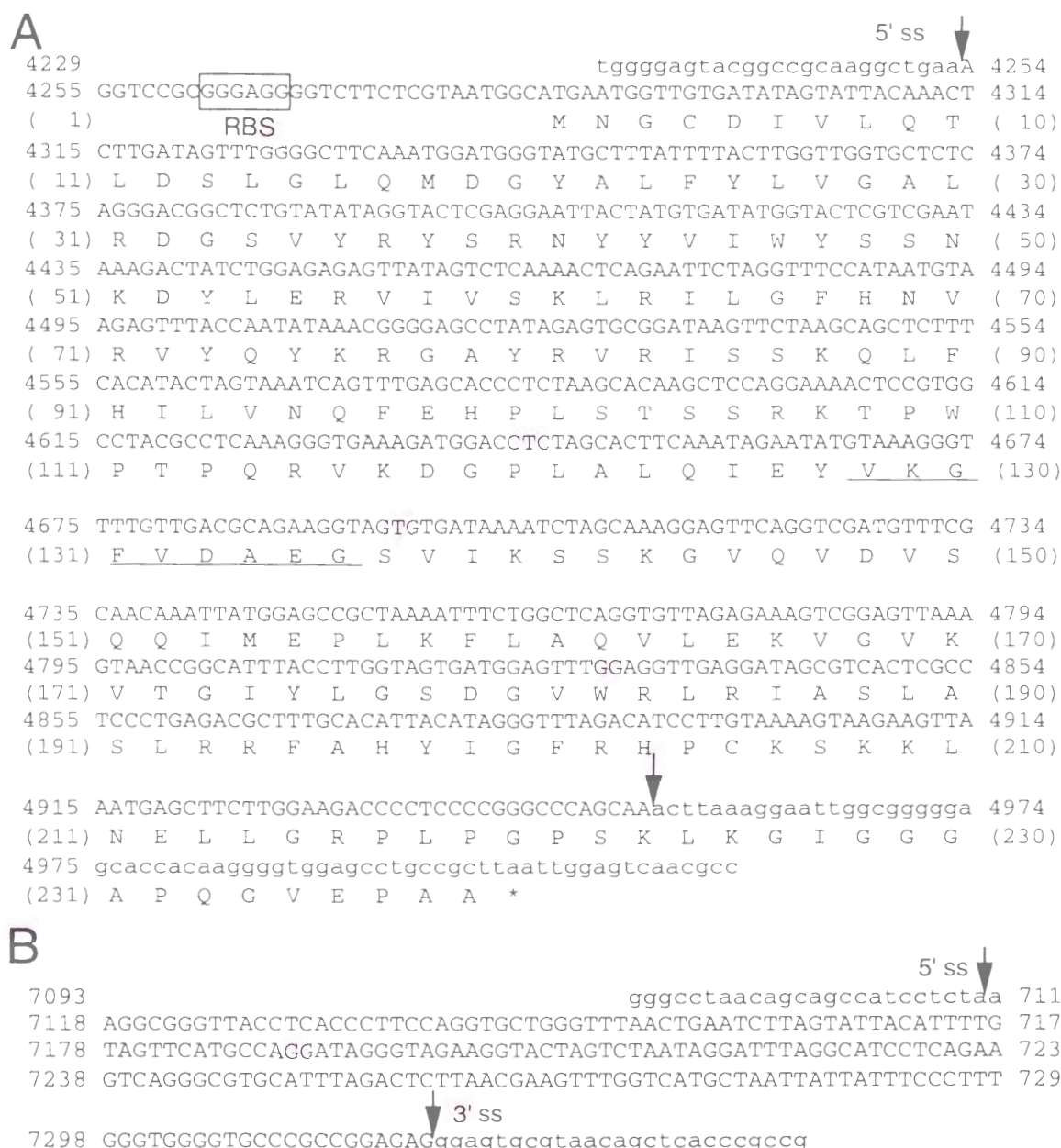


FIG. 3.6. Nucleotide sequences corresponding to the three rRNA introns and flanking exons. (A) I α ; (B) I β ; (C) I γ ; (D) I δ ; (E) I ϵ ; (F) I ζ . The nucleotide positions refer to the numbering for the sequence AB008745. 5'ss and 3'ss denote the 5' and 3' splice sites, respectively. The amino acid sequences of intronic ORF are included. Putative ribosome binding sites (RBS) are boxed. Underlined amino acid sequences indicate LAGLIDADG motif (Mueller *et al.*, 1993) that is shared by intron/intein-encoded homing endonucleases.

C

7150 gcccgggagtaaccctgactctctta 8174 5' ss ↓
8175 ACGGGGGTTATCCACCGCCAGTCACAGTATATATTAGGGTTTTGGTTTCTTGGGCTTCTT 8234
8235 TATTGAGGATAGATTAGAGTGGATCAGTGTCTACTGGAGGGATATGTGGCGGAATAATA 8294
(1) RBS V D Q C L L E G Y V A G I I (14)

8295 GATGCCGAAGCTAGCCTCTCAGTAAGCATAAAAAATTCAAGAGGATCTAAGATGTCGTGTG 8354
(15) D A E A S L S V S I K I Q E D L R C R V (34)

8355 CGCGTCGATCCTGTGTTCAGTATAACACAAGACAGCAAAGACGTGTAAACGTGGTGAAA 8414
(35) R V D P V F S I T Q D S K D V L N V V K (54)
8415 AACTACTTTGGCTGTGGCAGGTTGATGCCGAAACCAGGTCAAGAACATCTAACATTATAT 8474
(55) N Y F G C G R L M P K P G Q E H L T L Y (74)
8475 GTGGTTGATAGGCTAGAGGCTCTTGCTGATTGTCTAATCCCTAAATTGGATCGTCTACCA 8534
(75) V V D R L E A L A D C L I P K L D R L P (94)
8535 TTAATCGTCAAGAAAAGAGGATTCGAGATGTTCCGTGAGATTGTATTAACGTAAACCCGA 8594
(95) L I V K K R G F E M F R E I V L T L T R (114)
8595 ATGAAATATAGACGAGTCGAGTGCTGCGTCATAAGAGACTTGGTATTAAATCGTACTCCC 8654
(115) M K Y R R V E C C V I R D L V L N R T P (134)
8655 TATCTAGCCTAAATAAAAAAGTCTAAGCGAAAACGCTCGCTAGAAGAAATTTGAAGATTA 8714
(135) Y L A * 3' ss ↓
8715 TTCTTGTGATAAGGCGGTGGAGCCCCCGGGAGAAaggtagccaaatgccttgccgggta

D

-25 tggggagtagcgccgcaaggctgaa -1 5' ss ↓
1 AGGTCCGCGGGAGGGTTCCTGACGTGGTTAAGTAGAATTGTTACCCTCCCCGGGCCCAGC 60
61 AAacttaaaggaattggcggggga
↑ 3' ss

E

-10 ccgtcgccag -1 5' ss ↓
1 ACCAGCUGUAGGCGCCCGCACGGUACGCUGAAAUAUAGUGUAAAUAAGAAGUACCUUAGG 60
61 AGAAGACCUUAAGAUCUCUUCACUUUAAUAAACGCGCCGUGCGGGCGCCUGUGCUGGGGC 120
121 GAACTcgtgcccgtgaggtgtcctgtta
↑ 3' ss

F

-25 gcccgggagtaaccctgactctctta -1 5' ss ↓
1 ACGGGGGTTATCCACCGCCGGTCATAGCATATAAGGCGGGTGGAGCCCCCGGGAGAAagg 60
61 tagccaaatgccttgccgggta 3' ss ↑

FIG. 3.6. (continued.)

4.4

DISCUSSION

The finding that the single rRNA operon of the hyperthermophilic archaeon *Aeropyrum pernix* K1 contains three introns is not totally unexpected, considering that the intervention was known to coincide with sequences of 16S or 23S rRNA genes from a few species in the Crenarchaeota kingdom, such as *Desulfurococcus mobilis*, *Staphylothermus marinus*, *Pyrobaculum organotrophum*, and *Pyrobaculum aerophilum* (Burggraf *et al.*, 1993; Dalgaard and Garrett, 1992; Kjems and Garrett, 1985; Kjems and Garrett, 1991). However, this is the first description of the presence of introns within both 16S and 23S rRNA genes in prokaryotic system. Detailed RNA analyses confirmed that three introns are removed during the post-transcriptional rRNA processing events and that the flanking rRNA segments (exons) are ligated to generate the mature 16S and 23S rRNAs. In this respect, they are distinguished from a number of intervening sequences observed within the 23S rRNA genes from certain genera of bacteria, such as *Campylobacter* (Konkel *et al.*, 1994; Lipton *et al.*, 1994; Trust *et al.*, 1994), *Leptospira* (Hsu *et al.*, 1990; Ralph and McClelland, 1994), *Rhodobacter* (Kordes *et al.*, 1994; Lessie, 1965), *Salmonella* (Bulgin *et al.*, 1990; Hsu *et al.*, 1994; Hsu *et al.*, 1992, Skurnik and Toivanen, 1991), and *Yersinia* (Skurnik and Toivanen, 1991), and eukarya, such as *Crithidia* (Spencer *et al.*, 1987), *Euglena* (Schnare and Gray, 1990), and *Chlamydomonas* (Boer and Gray, 1988; Turmel *et al.*, 1991).

The sequence analyses of the single rRNA operon in several isolates of *Aeropyrum pernix* from various marine hydrothermal sources demonstrate that four distinct genotypes (K1, OH2, TB1, and TB7) based on the introns present within this region harbor a total of 10 introns. On the basis of the sequence differences observed in the introns, we have assigned these introns into six types ($I\alpha$, $I\beta$, $I\gamma$, $I\delta$, $I\epsilon$, and $I\zeta$). Interestingly, $I\alpha$ and $I\delta$ precisely occupy the same insertion sites in the 16S rRNA genes, although they differ in length and primary structure. As is the case observed between $I\gamma$ and $I\zeta$ in the 23S rRNA genes. These observations lead us to suppose the presence of at least four hotspots ($I\alpha/I\delta$, $I\beta$, $I\gamma/I\zeta$, and $I\epsilon$) for intron-insertion within the rRNA coding region of *Aeropyrum pernix*.

One challenge is to understand the evolution and phylogenetic distribution

of these introns. The pattern of introns suggests that they were independently inserted into the hotspots within the rRNA genes of *A. pernix*. Introns at same positions, e. g. I α and I δ , are not recognizably similar in sequence or are significantly different in size and are difficult to compare. If a prototypical insertion pattern of introns was present before the divergence of *A. pernix* subspecies, then a staggering number of mode for deletion events is required to account for the present distribution. Therefore, it is plausible that intron-insertion promoted the subspecies genetic divergence of this archaeon.

Some eukaryotic group I introns are considered autonomous genetic elements (Dujon, 1989; Lambowitz, 1989) that are mobile by intron insertion (Jaquier and Dujon, 1985; Woodson and Cech, 1989; Mohr and Lambowitz, 1991) or , possibly, horizontal transfer (Dover and Coen, 1981; Lambowitz and Perlman, 1990). Some introns are known, genetically and experimentally, to be inserted or 'home' into specific target positions that are recognized by an endonuclease encoded by the same intron (Jaquier and Dujon, 1985; Belfort, 1991). With intron homing, an intron could be inserted into an intron-lacking rDNA repet as a result of interaction with an intron-containing rDNA repeat from another nucleus, perhaps during sexual reproduction. On the analogy of these knowledges, the presence of hotspot for rRNA introns in the archaeon *A. pernix* might be rationalize by the further investigation of the homing-site specific endonuclease (homing endonuclease) and the observation of mating phenomenon.

Chapter 4

Post-splicing Dynamism of the Excised Intronic RNAs in *A. pernix* K1 Cells

INTRODUCTION

Molecular approaches to mechanistic aspects of prokaryotic RNA splicing might lead us to better understanding on regulation of prokaryotic gene expression systems. The primary transcript of rRNA genes in both prokaryotes and eukaryotes must undergo a number of post-transcriptional processing reactions to produce the mature, functional form of the molecule before assembly and activation of ribosomes. According to a number of investigation for eukaryotic ribosome synthesis (reviewed in Beserga and Steitz, 1993; Hernandez-Veerduin, 1991), the 37-45S rRNA transcript is packaged in a large ribonucleoprotein particles imported into nucleus. While this particle remain in the nucleolus, selected RNAs and proteins involved in rRNA processing and splicing are discarded as it is processed into immature large and small ribosomal subunits. These two subunits are thought to attain their final functional form only as each is individually transported through the nuclear pores into the cytoplasm. In contrast, little is known about rRNA precursor processing and ribosome-assembly pathway in prokaryotic cells. Particularly, many important questions remain to be answered pertinent to prokaryotic rRNA splicing. In what way the temporal order of the splicing events are controled without nucleolus and nuclear envelope? How many rRNA splicing enzymes exist per cell? rRNA splicing is influenced by cell growth or rRNA transcription? What *trans*-acting factors regulate rRNA splicing and how do they function? What is the mechanism of excised intronic RNA decay, i. e. what are the degradation intermediates? Archaeal rRNA introns could serve as an attractive experimental object suitable for the inquiry.

In the previous chapter, I present that (i) three introns ($I\alpha$, $I\beta$, and $I\gamma$) reside within the rRNA operon of *A. pernix* K1, (ii) two of them, $I\alpha$ and $I\gamma$, contain long ORFs consist of 221 and 137 amino acids, respectively, (iii) putative Shine-Dalgarno sequences occur 10-15 nucleotides upstream of the

start codons. To examine whether the free intronic RNAs, I α and I γ , can function as mRNAs for corresponding ORFs, post-splicing dynamism of the excised intronic RNAs in *A. pernix* K1 cells was investigated.

MATERIALS AND METHODS

RNA preparation. Total cellular RNA was isolated from cells grown to either mid-logarithmic growth phase (optical density at 660nm of 0.4) or stationary phase (optical density at 660nm of 0.9) by the acid guanidinium-phenol-chloroform method (Chomczynski and Sacchi, 1987). The resulting RNA pellet was dissolved in H₂O and stored at -80°C.

Primer extension analysis. The 5' termini of the spliced introns were determined by primer extension analysis, using oligonucleotides P1, P3, and P5 described above as primers for I α , I β , and I γ , respectively. A five-microgram amount of total RNA from cells grown to the stationary phase and 50 pmol of 5'-digoxigenin labeled oligonucleotide primer were heated at 70 °C for 10 min and then hybridized at 50 °C for 2 min in 20 μ l of 50 mM Tris-HCl (pH 8.3), 75 mM KCl, 3 mM MgCl₂, 10 mM dithiothreitol, 0.5 mM each deoxyribonucleotide. Reverse transcription with 200 U of SuperScript II RNase H⁻ reverse transcriptase was carried out at 50 °C for 1 h. RNA was then degraded with DNase-free RNase A (100 μ g/ml, 1h, 37 °C). A sequencing reaction with the same primer and either pRU363 (for I α) or pRD542 (for I β and I γ) as the template was run in parallel as a reference for determining the endpoint of the extension product. The sequencing ladder was generated with DIG Taq DNA sequencing kit (Boehringer) and 5'-digoxigenin labeled oligonucleotide primers.

Northern blot analysis. Total cellular RNA samples (2.5 μ g) were electrophoresed on a 1.5% agarose-formaldehyde denaturing gel in MOPS buffer (20mM 3-[N-morpholine] propanesulfonic acid, 5mM sodium acetate, 2mM EDTA [pH 7.0]), blotted onto a Biotrans Plus membrane by using 20 $\times$ SSC, and then the blot was probed according to the same protocol employed for Southern blot analysis. The RNA size standard was obtained from Gibco BRL.

Two-dimensional denaturing gel electrophoresis. In order to examine the topological properties of the spliced intronic RNAs, two-dimensional denaturing gel electrophoresis was carried out on horizontal gels according to a modified version of the method of Ford and Ares (1994). The first dimension was 1.5% agarose-formaldehyde denaturing gel and was run for 3 h at 5 V/cm in MOPS buffer. A lane from the first dimension was excised and laid on the horizontal surface, and the second dimension was poured directly around it. The second dimension was 2.2% agarose-formaldehyde denaturing gel and was run for 2 h at 10 V/cm in MOPS buffer with 0.3 μ g/ml ethidium bromide. Electrophoresis in the second dimension was performed at 4 °C. The total cellular RNA (2.5 μ g) from stationary phase cells was used in this analysis. The spots correspond to the $I\alpha$, $I\beta$, $I\gamma$ were detected by the northern blot analysis with mixed probe of probes 2, 3, and 4.

RESULTS

Determination of the splice sites. Primer extension analysis identified 5' termini of the three introns (Fig. 4.1). The observed products indicated the 5' splice sites are located at positions 4254, 7118, and 8175 (AB008745) of the 5' ends within $I\alpha$, $I\beta$, and $I\gamma$, respectively. The generation of termination signals implies that all the intronic RNA molecules are not circulized after splicing events *in vivo*. This is in good agreement with the results of Northern blot analyses described above, in which the intronic RNA species in linear form ($L-I\alpha$, $L-I\beta$, and $L-I\gamma$) were detected much more than those in circular form ($C-I\alpha$, $C-I\beta$, or $C-I\gamma$). In Fig. 3.1A, additional extension products with strong signal intensity were observed at 11 to 15 nt downstream of 5' splice site; these could represent positions for post-splicing processing at 5' end of the free intronic RNA species $I\alpha$.

Spliced introns are stable and accumulated in the cells grown to the stationary phase. Post-splicing fate of the intronic RNAs *in vivo* was investigated. The results of Northern blot analyses for total RNA prepared from cells before and after entry into stationary phase are presented in Fig. 4.2. Signals detected by $I\alpha$ -specific probe 2 revealed that (i) free intronic RNA

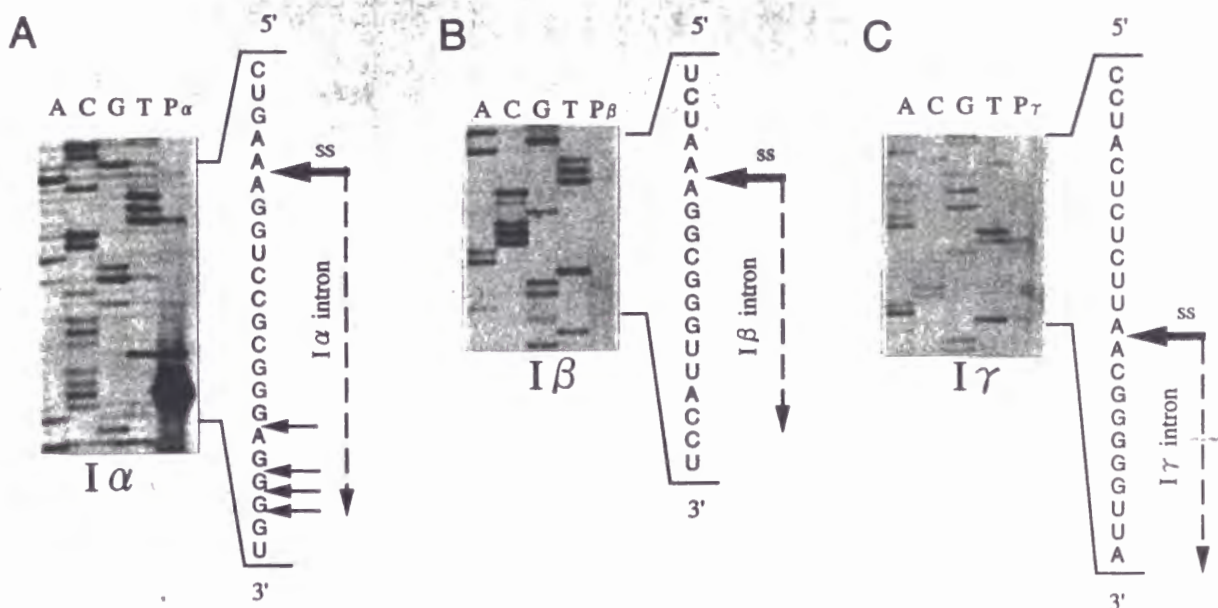


FIG. 4.1. Primer extension analyses determining the 5' termini of the rRNA introns. Synthetic oligonucleotides specific for $I\alpha$ (A), $I\beta$ (B), and $I\gamma$ (C) were used as primers for reverse transcriptase extension with total cellular RNA as template (lanes P_α , P_β , and P_γ), and as primers for dideoxy sequencing of the rRNA genes (lanes A, C, G, T). The deduced RNA sequence is shown to the right of each panel, with the arrowheads indicating the 5' termini of the abundant spliced intronic RNA species. 5' spliced sites are indicated by thick horizontal arrows with ss. Probable 5' end processing sites are denoted by thin horizontal arrows.

species $I\alpha$ was present in both logarithmic-phase cells and stationary-phase cells, (ii) the molecular ratio of free intronic RNA species $I\alpha$ to total cellular RNA increased considerably after entry into the stationary growth phase, and (iii) major signal with the expected migration behavior ($L-I\alpha$, 699 nt) was detected as well as minor amount of retarded band ($C-I\alpha$). Similar results were obtained for species $I\beta$ and $I\gamma$ (Fig. 4.2).

To inquire further into the topological properties of the free intronic RNA molecules in vivo, two-dimensional denaturing gel electrophoresis was carried out by using total cellular RNA from stationary-phase cells. Figure 4.3 shows the signals hybridized to the mixed probes 2, 3 ($I\beta$ specific), and 4 (I

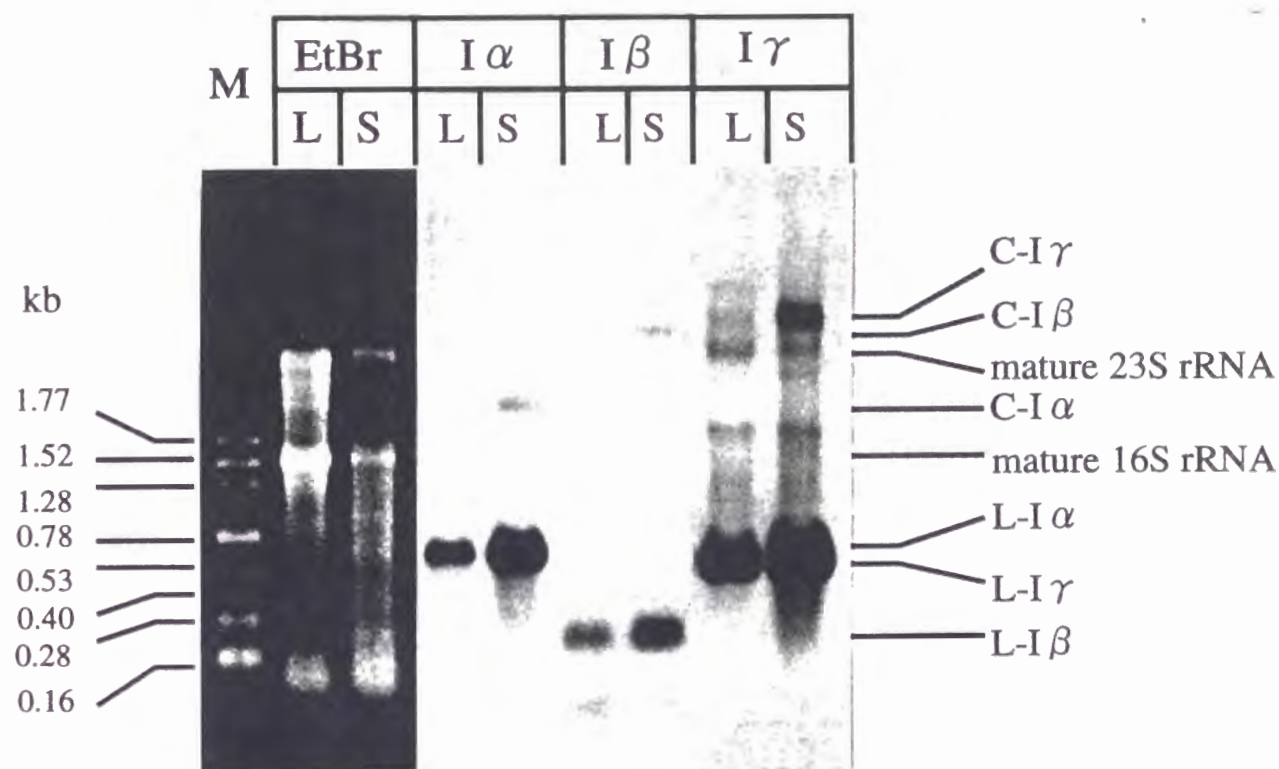


FIG. 4.2. Northern blot analyses of the spliced intronic RNAs accumulated in the cell. Total cellular RNA samples ($2.5 \mu\text{g}$) from cells grown on different growth phases were loaded as follows: L, mid-logarithmic growth phase; S, stationary phase. The left three lanes show an ethidium bromide-stained gel. Intronic RNAs I α , I β , and I γ were detected by using probes 2, 3, and 4, respectively. Size markers (lane M) in kb are provided to the left side of the panel. RNA species are identified on the right. The prefix L denotes intronic RNAs with expected mobility on agarose-formaldehyde denaturing gels, while the prefix C denote intronic RNA species with a significantly slower mobility.

γ specific). Each spot was identified by individual probe (data not shown). A smooth diagonal arc containing L-I α , L-I β , and L-I γ was generated, while C-I α , C-I β , and C-I γ appeared above the arc, suggesting that the latter isomeric RNA species possess intramolecular circular section and therefore retard in higher agarose concentration at the second dimension. Minor signals within the diagonal arc possibly reflect tandem oligomerized intronic RNAs in linear form, although further detailed identification is needed. These observation indicate that the three spliced intronic RNAs are stable and exhibit polymorphism in vivo and that the post-splicing intramolecular circulization of these intronic

RNAs are less efficient under the physiological conditions we investigated here.

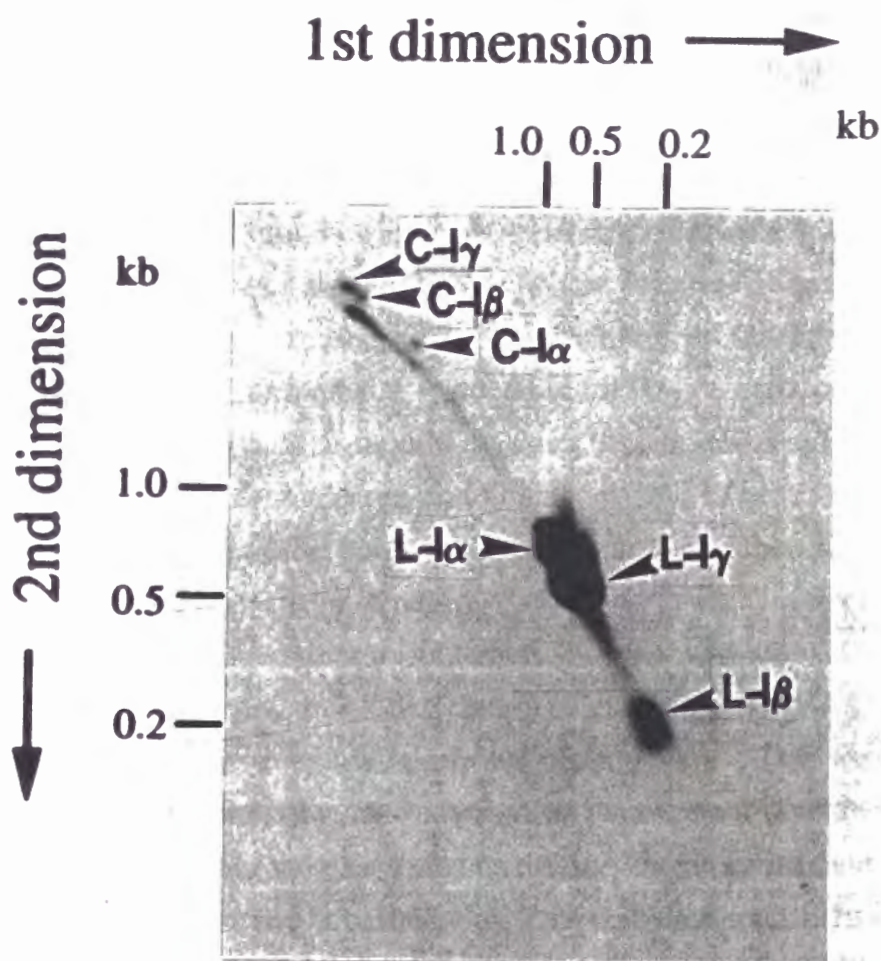


FIG. 4.3. Two-dimensional Northern blot analysis of the topological properties of the spliced intronic RNA species. Total cellular RNA (2.5 μ g) derived from stationary phase cells were subjected to electrophoresis as described in the text. Spots of the intronic RNA species were visualized by mixed probes 2, 3, and 4. The nomenclature of the RNA species is the same as shown in Fig. 4.2. Sizes (in kb) and expected migration of linear RNA species are shown on the top and left sides.

DISCUSSION

Mechanistic aspects for splicing of the three rRNA introns presented here remain to be investigated. The splicing of most RNAs with introns depends on structures or specific sequences that lie primarily within the introns themselves (Belfort *et al.*, 1995). However, due to the lack of significant levels of identity among primary structures of so far reported archaeal introns, it is considered that substrate recognition of the splicing endoribonuclease occurs mainly at a high order structure level (Kjems and Garrett, 1991). There are strong evidences that *Haloferax volcanii* tRNA intron endoribonuclease (EndA) requires the bulge-helix-bulge motif, which consists of two three-nucleotide bulge loops on opposite strands which are separated by precisely four helical base pairs near the center of a much longer helical structure (Kleman-Leyer *et al.*, 1997; Thompson and Daniels, 1988). Furthermore, so far reported rRNA introns in archaea were experimentally probed to fold bulge-helix-bulge helical structure at intron/exon junctions in vitro by using nucleotide-specific chemicals and ribonucleases (Lykke-Andersen and Garrett, 1994). These observations indicate the importance of the bulge-helix-bulge motif for their splicing. In contrast, unlike this consensus structure, I β and I γ were inferred to fold aberrant bulge-helix-bulge structure (Chapter 2; Figs. 2.5B and C), although both of them can form GC-rich (72.5 and 73.5 mol% G+C for I β and I γ , respectively) long stem structures at intron/exon junctions. This exceptional structures can be explained by two alternative hypotheses: (i) rRNA intron endoribonuclease has a broad substrate specificity, therefore it could tolerate the structural variant of the substrate, (ii) There are separate pathways for splicing of disparately structured substrates. A distinction between the two models is how stringently the rRNA intron endoribonuclease(s) recognize the junction structure.

Less information is currently available on rRNA intron endoribonuclease in *A. pernix* K1. In the processing of the primary rRNA transcript of halophilic archaeon *Halobacterium cutirubrum*, helical structures formed by the long inverted repeats surrounding both 16S and 23S rRNA-coding regions are recognized and cleaved by proteinaceous bulge-helix-bulge endoribonuclease

(Chant and Dennis, 1986; Dennis, 1991). Besides, the aforementioned *H. volcanii* EndA are shown to recognize the bulge-helix bulge motif (Kleman-Leyer *et al.*, 1997; Thompson and Daniels, 1988). Based on the analogy with these observations, the idea was proposed that rRNA introns capable of folding into bulge-helix-bulge structures could be cleaved by the same helix endoribonuclease, despite a few experimental evidences (Garrett *et al.*, 1991). In contrast, the processing helix surrounding the pre-16S rRNA in *Sulfolobus acidocaldarius* contains only a 2-base bulge on the 5' side. Analysis of in vitro rRNA processing intermediates indicated that little or no cleavage occurs within the region of the helical stem. Instead, cleavage occurs at two site further upstream (Durovic and Dennis, 1994). Potter *et al.* (1995) developed an in vitro processing system to study the two cleavages in the 5' external transcribed spacer (ETS) and the cleavage at the 5' ETS-16S rRNA junction and concluded that the cleavage activity is sensitive to digestion by micrococcal nuclease, suggesting that the responsible endoribonuclease contains one or more essential RNA component. Hence, it remains matter how archaeal rRNA intron splicing is initiated, especially the intron could fold into aberrant bulge-helix-bulge structures. Pursuing the splicing endoribonuclease of the three rRNA introns in *A. pernix* K1 might offer the clue to this problem.

Post-splicing, three rRNA introns found in this study was shown to retain in the cell without efficient circulization (Figs. 4.2 and 4.3), unlike the case of 23S rRNA introns in *D. mobilis* (Kjems and Garrett, 1988) and *Pb. organotrophum* (Dalgaard and Garrett, 1992), or 16S rRNA intron in *Pb. aerophilum* (Burggraf *et al.*, 1993). Under the conditions we investigated here, the relative amount of the excised intronic RNAs to total cellular RNA increased in the cells grown to stationary phase. DiRuggiero *et al.* (1993) reported that the regulation of rRNA transcription in hyperthermophilic archaeon *Pyrococcus furiosus* depends on alterations in growth rate, as occurs in *E. coli* (Gourse *et al.*, 1996). Thus, supposing that the rRNA operon in *A. pernix* K1 is no longer newly transcribed after entry into the stationary growth phase, our results leads us to suggest a mechanism in which the excised intronic RNAs are specifically exempted from degradation, presumably due to the secondary structure unique to these introns such as GC-rich long stem at their terminals or some unknown RNA modifications. Taking it into account that free intronic

RNAs, $I\alpha$ and $I\gamma$, could work as mRNA for intron-encoded ORFs, further investigations on post-splicing fate of the three introns might provide a molecular rationale of regulation of intronic ORF expression. Site specific 5' end processing of the excised intronic RNA $I\alpha$ (Fig. 4.1A) could also be involved in the translational control of intronic ORF, since the resulting trimmed oligoribonucleotide portions contain ribosome-binding site upstream of the intronic ORF.

Chapter 5

Functional Analyses of the Latent Intron-encoded Protein I- *Ape* I; Potential Role in the Horizontal Transfer of the Archaeal Introns

INTRODUCTION

Mobile introns have been identified in genetic crosses involving *Saccharomyces cerevisiae* (Colleaux *et al.*, 1986; Delahodde *et al.*, 1989; Wenzlau *et al.*, 1989), *Physarum polycephalum* (Muscarella *et al.*, 1989), bacteriophage T4 (Quirk *et al.*, 1989), and *Chlamydomonas* species (Colleaux *et al.*, 1990; Lemieux and Lee, 1987; Remacle *et al.*, 1990). Such introns belong to the group I, and some of them contain internal open reading frames (ORFs) encoding LAGLI-DADG type homing endonucleases (Mueller *et al.*, 1993) that are responsible for the 'homing' (see Dujon *et al.*, 1989, for definition of the term) of these introns into cognate intronless sites. These homing endonucleases generate double-strand breaks at or near the site of intron insertion and initiate efficient unidirectional gene conversion events during which the intron is most probably inserted by the double-strand break repair mechanism proposed by Szostack *et al.* (1983), resulting in their idiosyncratic distribution (Lambowitz and Belfort, 1993; Dujon, 1989; Lambowitz, 1989; Belfort, 1990).

In Chapter 3, I suggested that there are at least four hotspots for intron-insertion within the *arnSL* locus of *A. pernix* and described that two of the rRNA introns of *A. pernix* K1 have ORFs encoding putative proteins that contain variations of the LAGLI-DADG motif (see Fig. 5.1). In Chapter 4, I showed that the free intronic RNA molecules are stably accumulated in the cell after splicing events and can function as mRNAs. These observations led us to wonder whether the intron-encoded proteins are involved in the intergenomic, i. e. intercellular mobility and homing of the archaeal introns. A better understanding of the mechanism underlying rRNA intron homing in the genome of *A. pernix*, therefore, requires functional characterization of the intron-encoded proteins and the molecular interactions between the proteins

and their substrates.

In this context, I have overexpressed the I α ORF in *E. coli* and purified the recombinant I-*Ape* I protein to homogeneity to analyze its function. We also showed the evidence for the expression of I-*Ape* I in *A. pernix* K1 cells. Furthermore, the role that I-*Ape* I may have played in the horizontal transfer of *A. pernix* rRNA introns is discussed.



FIG. 5.1. Diagrammatic representation of the primary structure of the I-*Ape* I protein. The single copy of the conserved LAGLI-DADG motif found in I-*Ape* I and other mobile intron endonucleases and maturases is located at residues 128 to 136. This motif was aligned with the counterpart in I-*Ceu* I endonuclease, which is encoded by the fifth group I intron (CeLSU • 5) in the green algae *Chlamydomonas eugametos* chloroplast large subunit rRNA gene.

MATERIALS AND METHODS

Bacterial strains and plasmids. *Escherichia coli* BL21(DE3) carrying plasmids pLysS and pAH352 (see below) was used to produce recombinant I-*Ape* I protein; this strain harbors a λ lysogen that contains the T7 RNA polymerase gene under control of the *lac* UV5 promoter (Studier *et al.*, 1990). pET-11b (Stratagene) was used for the construction of the expression plasmid pAH352 (see below). The plasmid pO2A2, which contains the homing site of

the intron I α within the *arnS* gene of *A. pernix* OH2, were constructed by cloning directly the PCR fragment (1.26 kb) into the pCR2.1 vector (Invitrogen). For cleavage assay, the pO2A2 plasmid was linearized with *Eco* RI and then used as a substrate.

Recombinant DNA techniques. Standard procedures were used as described in Chapter 3.

Construction of expression plasmid for the intron-encoded protein

I-*Ape* I. To express the presumptive I-*Ape* I coding sequence independently of the rest of the *arnS* gene of *A. pernix* K1, codons 1 to 222 (refer to the numbering in Fig. 3.6A) were inserted into a bacterial expression vector, as follows. To place an initiator methionine codon (ATG) and an *Nde* I site at the 5' end of the I-*Ape* I coding segment, two oligonucleotide primers (5'-ACATATGAATGGTTGTGATATA -3' and 5'-GGATCCTTAGCTGGGCCCGGGGAG -3') and *A. pernix* K1 chromosomal DNA as a template, a corresponding double-stranded fragment was generated by PCR. One end of this PCR product has the sequence corresponding to the carboxy-terminal end of the presumptive I-*Ape* I polypeptide followed immediately thereafter by a translational stop codon and an *Bam* HI site. Therefore, a 677-bp *Nde* I-*Bam* HI fragment of the PCR product that contains this region was subcloned into pET-11b to yield pAH352. Nucleotide sequence analysis confirmed that the inserted segment did not contain any unwanted PCR-generated mutations.

Expression and purification of the recombinant I-*Ape* I protein.

Overexpression of I-*Ape* I was obtained by transforming pAH352 into *E. coli* strain BL21 (DE3) pLysS. A culture (6 liter) of transformed BL21 (DE3) pLysS cells was grown at 37 °C in LB broth (1% Bacto-tryptone, 0.5% Bact-yeast extract, 0.5% NaCl, 1 mM NaOH) supplemented with 100 μ g/ml ampicillin to an OD₆₀₀=1.0. Induction of the T7 RNA polymerase, and subsequently the I-*Ape* I gene under control of the T7 RNA polymerase-dependent phage promoter, was achieved by the addition of isopropyl-1-thio- β -D-galactopyranoside (IPTG) to a concentration of 1.0 mM followed by further growth for 3 h at 37 °C. The culture was then chilled on ice, and the cells (15 g, wet weight) were harvested by centrifugation at 4 °C. Unless otherwise stated, all subsequent operations were performed at 4 °C. The cell pellet was washed and resuspended in 1.5 volumes of phosphate-buffered saline [PBS; 8.1 mM

Na_2HPO_4 , 1.47 mM KH_2PO_4 , 137 mM NaCl, 2.7 mM KCl (pH 7.4)]. The cell suspension was sonically irradiated at maximum output using a model UR-200P sonicator (Tomy Seiko, Tokyo, Japan) using five 20-s pulses with intermittent 20-s periods of cooling on ice. The crude lysate was diluted with 10 volumes of PBS and subjected to centrifugation for 20 min at $20,000 \times g$ to remove unbroken cells and cellular debris. The resulting supernatant fraction was heated at 60°C for 30 min, and the thermal denatured proteins were removed by the sedimentation at $20,000 \times g$ for 20 min. The supernatant solution was adjusted slowly to a final concentration of ammonium sulfate at 25% of saturation by the addition of a sufficient amount of the solid salt with stirring over 30 min on ice before centrifugation for 20 min at $20,000 \times g$. Ammonium sulfate was then added to the supernatant to a final concentration of 50%. After incubation on ice and centrifugation, the pellet was dissolved in 20 ml of PBS, and dialyzed extensively against the same buffer. The dialysate was filtrated through a $0.2\text{-}\mu\text{m}$ cellulose acetate membrane filter (Advantec, Tokyo) and loaded onto a Mono Q HR 5/5 column (Pharmacia Biotech Inc.) at 1.0 ml/min on a FPLC apparatus. After washing the column with the same buffer, bound protein was eluted using a linear gradient from 0 to 1 M NaCl in PBS buffer. Fractions (1 ml each) were collected, assayed for endonuclease activity, and those containing the bulk of the I-*Ape* I activity were pooled, and aliquots were stored at -20°C .

Partial purification of the native I-*Ape* I protein from *A. pernix* K1 cells. *A. pernix* K1 cells were grown to stationary phase ($\text{OD}_{660} = 0.9$) by batch culture technique as described in Chapter 1. Native I-*Ape* I protein was extracted and purified with the same procedure as those of the recombinant I-*Ape* I except that heat treatment step was skipped.

Molecular weight determination. Purified recombinant I-*Ape* I protein was applied to a Superdex 200 HR column (Pharmacia) previously equilibrated with PBS buffer and eluted using a FPLC apparatus at a flow rate of 0.2 ml/min. The elution position (V_e) of the I-*Ape* I protein was determined by its absorbance at $A_{280\text{nm}}$. The column was calibrated by determining V_e for bovine serum albumin (66 kDa), ovalbumin (44 kDa), chymotrypsinogen (25 kDa) run under identical conditions and by determining V_0 (9.2 ml) from the elution position of blue dextran (Pharmacia) and V_i (24 ml).

Mapping of the cleavage site. The I- *Ape* I cleavage site was delimited by a primer extension method using plasmid pO2A2 which contain the I- *Ape* I cleavage site in each orientation. Single -stranded form of this plasmid was prepared by standard procedures (Sambrook *et al.*, 1989) and used as template for synthetic primers, using a modification of the DIG sequencing protocol (Boehringer) to generate partially double-stranded DNA molecules. Each single-stranded template (1 pmol) was annealed to 5 pmol of an appropriate 5'-end digoxigenin labeled oligonucleotide primer (Acs-1: 5'-ACGATGCGGGCTAG GTGTTG- 3' and Acs-2: 5'- GGCCGTCATCCTGCTGTCGC- 3') in 20 μ l of standard reaction buffer. Extension was then carried out by using Klenow fragment (Takara) in each of the four standard dideoxynucleotide sequencing reactions at 37°C for 2 h. The extension products were extracted with phenol: chloroform: isoamyl alcohol, precipitated with ethanol, and then resuspended in loading buffer (100% formamide, 0.5 mM EDTA, 0.5 mg/ml bromophenol blue, 0.5 mg/ml xylene cyanol FF). The resulting samples were subjected to electrophoresis in a 6% polyacrylamide gel containing 7.5 M urea in 89 mM Tris-borate (pH 8.0) and 2 mM EDTA. Standard sequencing reactions were performed by the dideoxynucleotide termination method (Chapter 3) and analyzed on the same gel.

Cleavage assays. Target DNA (*Eco* RI-linearized pO2A2 plasmid; 1 μ g) was incubated with 0.5 μ g of the purified I- *Ape* I protein in a total volume of 15 μ l containing 10 mM Tris-HCl (pH 7.5), 10 mM MgCl₂, 1mM dithiothreitol, and 50 mM NaCl at 85 °C for 30 min. During the incubation, the mixture was overlaid with 25 μ l of mineral oil. The reactions were stopped on ice with 2 μ l of dye mixture [0.25%(wt/vol) bromphenol blue, 0.25%(wt/vol) xylene cyanol FF, 1mM EDTA, 30%(vol/vol)glycerol]. Cleavage products were separated by electrophoresis in 0.8% agarose gels and visualized by staining with ethidium bromide.

RESULTS

Expression of the I- *Ape* I protein. As the I- *Ape* I protein might be highly toxic to *E. coli*, the I α ORF was inserted into the pET-11b vector.

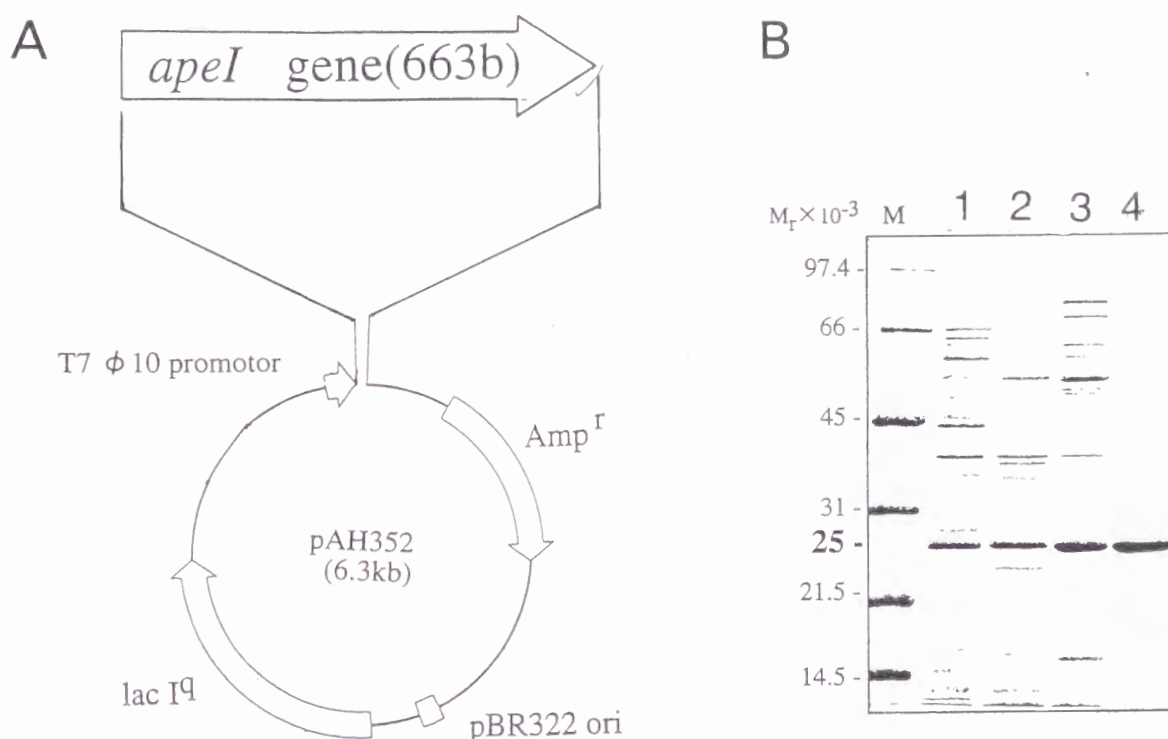


FIG. 5.2. Overexpression and purification of I-*Ape I*. (A) Physical map of the construct pAH352 for overexpression. (B) Analysis of the recombinant I-*Ape I* purification by SDS-PAGE. Samples (5 μ g each) were from the indicated stages of the purification of recombinant I-*Ape I* from *E. coli*. Lane 1, clarified lysate; lane 2, supernatant fraction after heat treatment; lane 3, redissolved 25 to 50% ammonium sulfate precipitate; lane 4, pooled Mono Q eluate; lane M, molecular mass standards. All samples were suspended in SDS-PAGE sample buffer, subjected to electrophoresis in 14% SDS-polyacrylamide gels, and stained with Coomassie Brilliant Blue.

Tight regulation of this gene in the resulting pAH352 construct is ensured by the operator sequence of the *E. coli* lactose operon which is situated downstream of the T7 promoter and which binds the *lacI^q* repressor encoded by the plasmid vector, thus preventing transcription by basal levels of T7 RNA polymerase. Upon addition of IPTG (1 mM) to the *E. coli* culture carrying the pAH352 construct, expression of the T7 RNA polymerase gene is first induced, followed by the expression of the I-*Ape I* gene. It was confirmed by a strong band of 25 kDa in SDS-PAGE analysis (Fig. 5.2B).

I-*Ape I* has site-specific endonucleolytic activity. The recombinant I-*Ape I* overexpressed in *E. coli* cells were tested in cleavage assays with the intron-minus *arnS* allele form *A. pernix* OH2 as the DNA substrate. It showed

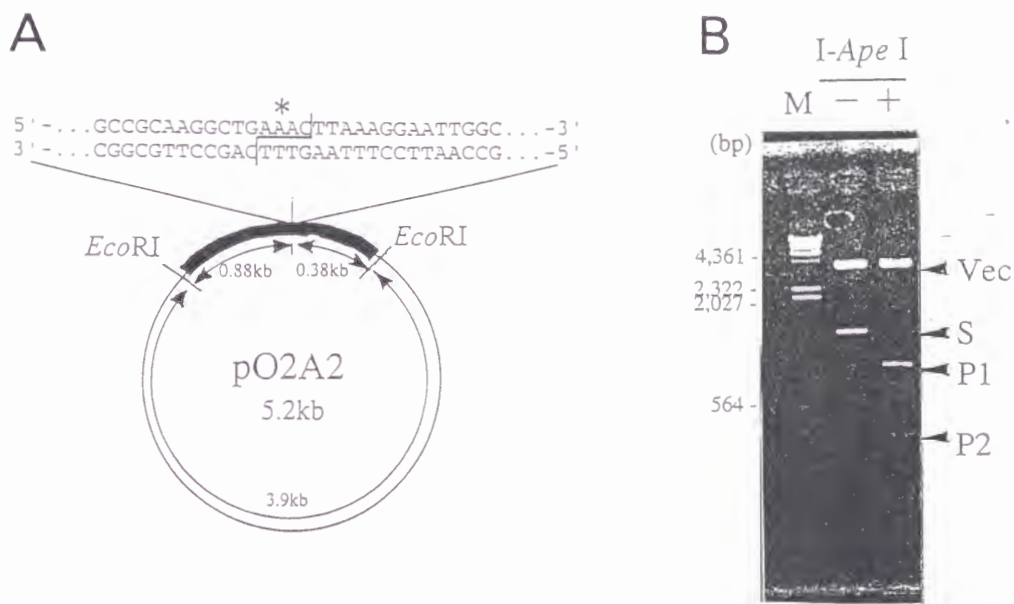


FIG. 5.3. DNA substrate and assay for recombinant I-*Ape I* activity. (A) Schematic representation of plasmid pO2A2, constructed as described under 'Materials and Methods'. The unique *EcoRI* site in the vector (solid circular line) and the DNA fragment containing intron-minus *armS* allele that was inserted (closed arc) are indicated. The nucleotide sequence of the I-*Ape I* recognition/ cleavage site, the staggered double-stranded break generated by I-*Ape I* action, and the predicted sizes of the two DNA fragments produced by I-*Ape I* cleavage from plasmid that was linearized by digestion with *EcoRI* are also shown. (B) Substrate DNA (pO2A2 linearized with *EcoRI*; 0.5 μ g) was incubated for 30 min at 85 °C under the otherwise standard reaction conditions specified under 'Materials and Methods' in the absence (-) or in the presence of 0.5 μ g of purified recombinant I-*Ape I* from *E. coli*. The procedures were analyzed by electrophoresis in a 0.8% agarose gel using a Tris acetate/EDTA buffer system.

high levels of cleavage activity (Fig. 5.3B)

Properties of I-*Ape I* endonuclease. Consistent with the hyperthermophilic nature of *A. pernix*, which grow optimally at 90 to 95 °C, I-*Ape I* is thermostable. The enzyme requires Mg^{2+} , is virtually inactive at temperatures up to 37 °C, reaches peak activity in the 85 to 90 °C range, and loses activity at higher temperatures (data not shown). It is unclear whether loss of activity above 90 °C reflects denaturation of the enzyme and/ or the substrate DNA.

The single cleavage site, which was observed exclusively for the intron-minus and not the intron-plus allele (data not shown), was mapped by bidirectional primer extension analysis (Fig. 5.4A). The cleavages on the two strands generate 4-nt 3' overhangs, which overlap the site at which the intron is inserted in the intron-plus allele. Like intron endonucleases in the two other

domains of life, the bacteria and eukarya, I- *Ape* I appears to have an asymmetric recognition sequence (Fig. 5.4A) and generates 3'-hydroxyl and 5'-phosphate termini that can be ligated with T4 DNA ligase (data not shown).

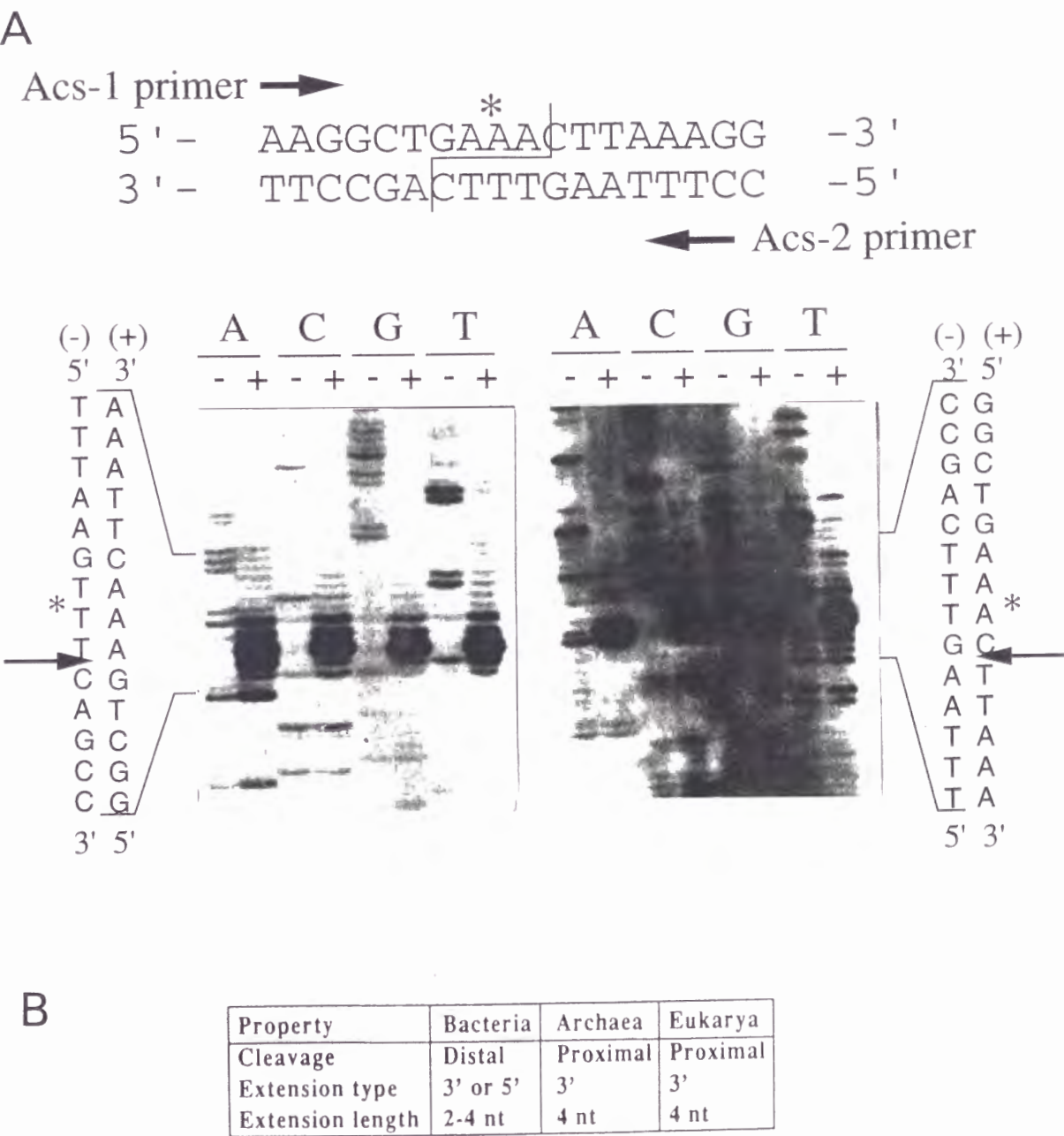


FIG. 5.4. Cleavage characteristics. (A) Cleavage-site mapping. Primer extension products treated with (+) or without (-) I- *Ape* I; dideoxy-sequencing reaction alone (A, C, G, T), and dideoxy-sequencing reactions with I- *Ape* I included (+ I- *Ape* I). All extensions were from primers annealed upstream (left) or downstream (right) from the putative insertion site of intron I α . The cleavage sites are indicated by horizontal arrows and the intron insertion site is marked by asterisks. The experimental procedures are described in 'Materials and Method's. (B) Properties of intron-encoded endonucleases

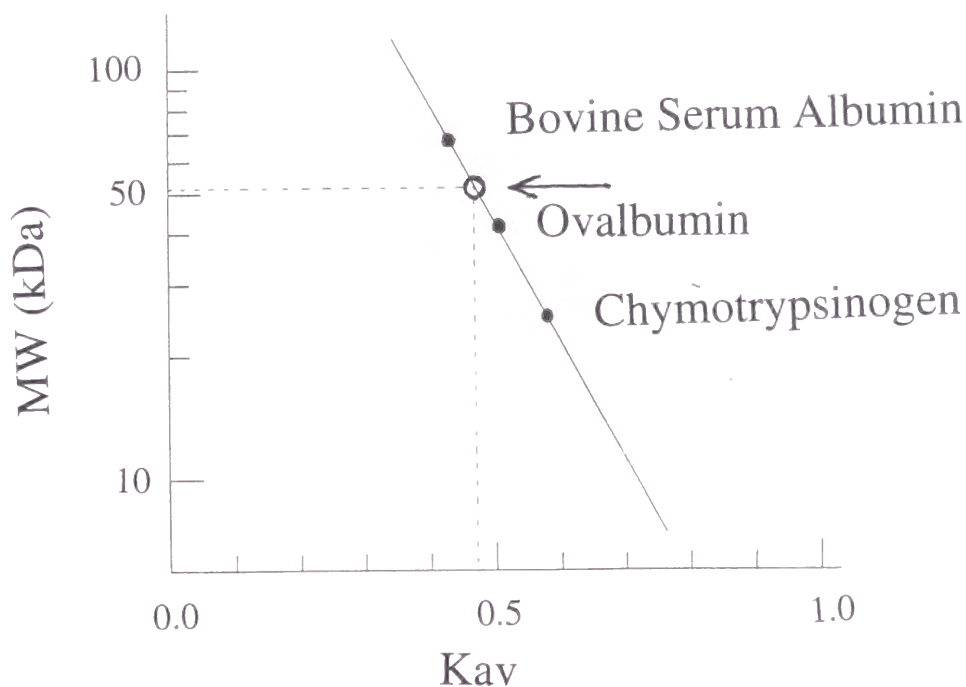


FIG. 5.5. Molecular weight determination of I-*Ape* I by gel filtration. Purified recombinant I-*Ape* I was subjected to gel exclusion chromatography on a bed of Superdex 200 under nondenaturing conditions, as described under 'Materials and Methods'. Molecular mass standards (and corresponding observed Kav values, where $K_{av} = (V_e - V_0) / (V_t - V_0)$) were : bovine serum albumin, 66 kDa (0.43); ovalbumin, 43 kDa (0.50); chymotrypsinogen, 25 kDa (0.57). The arrow indicates the elution position of I-*Ape* I ($K_{av} = 0.47$), which yielded an estimated apparent molecular mass of 50 kDa.

Detection of the native I-*Ape* I activity in *A. pernix* K1 cells. Active I-*Ape* I fractions were eluted at 0.6 to 0.7 M NaCl in the Mono Q column chromatography (data not shown). The recovered activity possessed the identical properties as that of the recombinant I-*Ape* I activity described above.

Determination of native molecular mass of I-*Ape* I. Electrophoretic analysis indicated that the recombinant I-*Ape* I, when denatured with SDS, had an apparent molecular mass of about 25 kDa (Fig. 5.2B). Gel permeation chromatography was used to estimate the molecular mass in solution of undenatured I-*Ape* I. Purified recombinant enzyme was subjected to chromatography on a column of Superdex 200 which had been calibrated as described under 'Materials and Methods', immediately before application of the I-*Ape* I sample. The V_e observed for I-*Ape* I yielded a Kav of 0.47, which, when compared with the values for the three protein standards used for calibration, gave an apparent molecular mass of 50,000 (Fig. 5.5). The molecular weight of I-*Ape* I calculated from the amino acid sequence deduced from the

nucleotide sequence of the corresponding segment of the intron I α is 25 kDa. The close correspondence between the apparent molecular weight of I- *Ape* I determined by three independent methods (gel exclusion chromatography, SDS-PAGE, and direct calculation from amino acid sequence) indicates that I- *Ape* I is a homodimeric enzyme in solution. The same result was obtained from the native I- *Ape* I.

DISCUSSION

Both introns I α and I γ contain ORFs, and the following circumstantial evidences suggest that they are expressed *in vivo* : (i) only one reading frame is possible for each intron, (ii) putative ribosome-binding sites, GGGAGGG or GAGGA, complementary to the 3' end of mature 16S rRNA occur 10 to 15 nt upstream of their start codons, (iii) free intronic RNAs, I α and I γ , which could function as mRNA are stably accumulated in the cells, although their relative quantity could fluctuate in dependence on the physiological conditions. Comparative sequence analyses against DDBJ, EMBL, GenBank nucleotide sequence databases and SWISS-PROT, PIR protein sequence databases reveal no genes or proteins with any overall similarity. However, a search made with each sequence, ¹²⁸VKGFVDAEG¹³⁶ in I α ORF and ¹⁰VAGIIDAEA¹⁸ in I γ ORF, revealed a similar motif (LAGLI-DADG) characteristic of proteins encoded by eukaryotic group I introns, archaeal introns and inteins (Mueller *et al.*, 1993)(Fig. 5.1). In *Saccharomyces cerevisiae*, the LAGLI-DADG protein encoded by the intronic ORF of 21S rRNA possesses endonuclease activity and causes a site-specific double strand break of intron minus (ω^-) variants of the gene, promoting efficient conversion to the intron plus (ω^+) form (Colleaux *et al.*, 1986; Collaux *et al.*, 1988). In this context, I have developed the overexpression system of the I α ORF (I- *Ape* I) in *E. coli* and verified the intron-homing endonucleolytic activity. Moreover, I detected the I- *Ape* I endonuclease activity in *A. pernix* K1 cells.

Active I- *Ape* I is expressed from excised forms of the intron I α , suggesting that splicing is a prerequisite for synthesis of the active enzyme. Like the phage and eukaryotic enzymes, I- *Ape*I cleaves an intronless allele by

making a staggered double-strand cut at an asymmetric target site. In general, I- *Ape* I is strikingly similar to the eukaryotic endonucleases (Fig. 5.4) but contrasts with that of the phage intron endonucleases, which cleave at some distance from the intron insertion site (Lanbowitz and Belfort, 1993; Belfort, 1990). Furthermore, all known eukaryotic homing endonucleases generate 4-nt 3' extensions, as opposed to the phage enzymes, which yield 3' or 5' extensions that are shorter than 4-nt (Lanbowitz and Belfort, 1993; Belfort, 1990). Additionally, expression of several eukaryotic intron endonucleases is dependent upon splicing, whereas this is not the case for the phage enzymes (reviewed in Belfort, 1990). Finally, the common LAGLI-DADG motif suggests that the archeal and eukaryotic intron endonucleases are members of the same family (see below).

Although homing of the rRNA introns cannot be demonstrated with the genetic tools currently available for *A. permix*, mobility is implied by the existence of intron-plus and intron-minus variants (Chapter 3) and by cleavage of the intron-minus allele by I- *Ape* I encoded in the intron-plus variant. Proof of mobility in an archaeon would be of interest for several reasons. Homing introns are most prevalent in multicopy genomes, such as those of bacteriophages, mitochondria, and chloroplasts, and in repetitive nuclear rRNA genes (Lanbowitz and Belfort, 1993; Dujon, 1989; Lambowitz, 1989; Belfort, 1990). *A. permix* rRNA genes are present in only one copy (Chapter 3); demonstration of homing would therefore provide a rare example of a mobile intron in a single-copy gene. Furthermore, coexistence of intron-plus and intron-minus alleles is implicit in homing, and intron mobility would therefore suggest that intercellular genetic exchange takes place in the hyperthermophilic archaeon *A. permix*.

It has been argued that ORFs were acquired by preexisting introns, based on closely related introns bearing different ORFs (Lanbowitz and Belfort, 1993; Dujon, 1989; Lambowitz, 1989; Belfort, 1990; Perlman and Butow, 1989). For example, the homologous *td*, *nrdB*, and *sunY* introns of phage T4 contain three heterologous ORFs (Belfort, 1990) and the *Neurospora* ND1 intron contains different ORFs at different locations depending on species (Mota and Collins, 1988). The discovery of an archaeal endonuclease in an intron that differs in structure and splicing pathway from the group I introns that frequent the other two domains of life supports the argument that the

endonuclease ORFs and the introns that house them arose separately (Bell-Pedersen *et al.*, 1990).

The independent origin of the endonuclease ORF from that of the intron core is further supported by the existence of LAGLI-DADG -containing endonucleases in both freestanding form between genes (Nakagawa *et al.*, 1991) and as protein fusions in an archaeon (Perler *et al.*, 1992), a yeast (Hirata *et al.*, 1990; Gimble and Thorner, 1992; Kane *et al.*, 1990), and a mycobacterium (Shub and Goldrich-Blair, 1992; Davis *et al.*, 1992). Another family of proteins, containing the GIY-YIG motif, is encoded by fungal mitochondrial introns (Cummings *et al.*, 1989) and also occurs in the form of both freestanding (Sharma *et al.*, 1992) and intron-encoded (Michel and Dujon, 1986) endonucleases in phage T4. These observations underscore the phylogenetically widespread distribution of different families of endonucleases and also support their intron-independent ancestry, since the endonuclease genes do not occur in exclusive association with introns but also at other genetic loci.

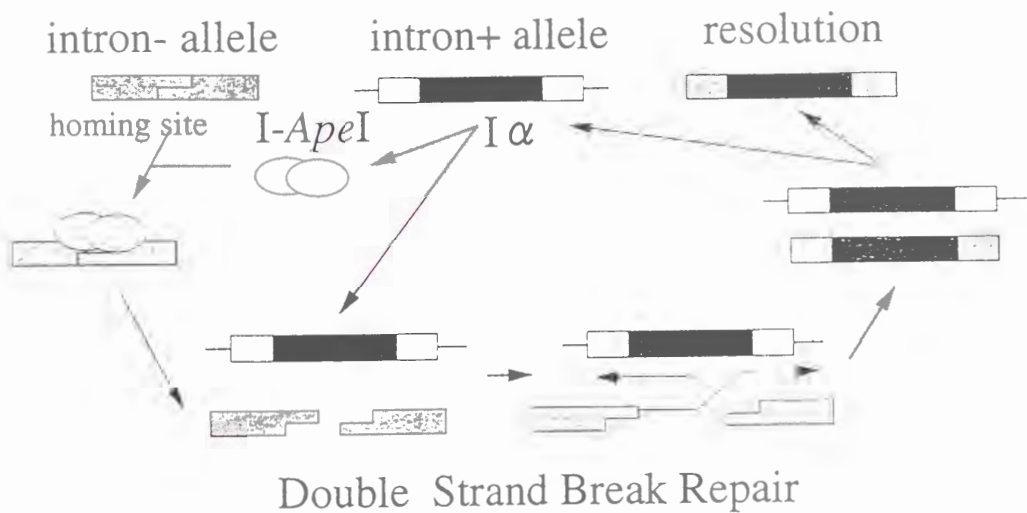


Fig. 5.6. Model for the I-Ape I intron homing cycle. Intron mobility or homing is initiated and targeted by a site-specific DNA double strand break made in an intron-less allele (homing site; gray box, left) by I- Ape I encoded by the mobile intron ORF (black box of intron-plus allele, top center). Double strand break repair transfers the mobile intron containing the endonuclease ORF to the cleaved recipient allele (double strand break repair, bottom). Resolution generates two alleles that contain the mobile intron (resolution, top right). Intron insertion disrupts the homing site, and renders intron-plus alleles resistant endonuclease cleavage.

The endonucleases have the ability to facilitate recombination by virtue of their double-strand cleavage activity (Fig. 5.6; reviewed in Lambowitz and Belfort, 1993; Belfort, 1990). This property can facilitate the entry of DNA, including that of the endonuclease gene itself, into a genome. Whether or not the foreign DNA could be tolerated within a gene would depend upon an innate splicing activity at either RNA or protein levels [the latter has been recently demonstrated for endonucleases in several different systems (Kane *et al.*, 1990; Davis *et al.*, 1992; Hodges *et al.*, 1992)] or upon the fortuitous colonization of a site that confers this ability. Introns of different kinds provide such hospitable landing sites and, if invaded by endonuclease ORFs, can be converted into mobile elements (Belfort, 1990). Mobility of the composite intron is then driven by the intron-encoded endonuclease, while the intron core splicing structure maintains gene function in subsequent transfer events (Lambowitz and Belfort, 1993; Belfort, 1990; Bell-Pedersen *et al.*, 1990). In this situation, as with endonuclease genes that are inserted directly into protein coding sequences and are spliced at the protein level, the endonuclease ORF can be viewed as the primary mobile element, with all of the invasive, propagative, and selfish properties of classical transposons.

Summary

The results obtained through a series of this study are summarized as follows:

(i) A novel aerobic hyperthermophilic archaeon was isolated from coastal solfataric vent at Kodakara Island, Japan. The new isolate, strain K1, was the first strictly aerobic organism growing up to 100 °C. It grew optimally at 90 to 95 °C, pH 7.0, and a salinity of 3.5 %. The cells were spherical shaped, 0.8 to 1.2 μ m in diameter. Various proteinaceous complex compound served as substrates during aerobic growth. Thiosulfate stimulated growth without producing H₂S. The core lipids consisted solely of C₂₅, C₂₅-isopranyl archaeol (glycerol diether). The G+C content of the genomic DNA was 67 mol%. Phylogenetic analysis based on 16S and 23S rRNA sequences indicated that strain K1 is a new member of Crenarchaeota. On the basis of our results, the name *Aeropyrus pernix* gen. nov., sp. nov. is proposed for the organism (type strain: K1; JCM 9820). This name was validated by the International Committee for Systematic Bacteriology.

The other three strains (OH2, TB1, and TB7), which were isolated from geographically distinct area, belong to the same species *A. pernix* because of the phenotypic similarities described above and high levels of similarity values in 16S-23S ITS sequence.

(ii) The single rRNA operon (*arnS-arnL*) of the hyperthermophilic archaeon *Aeropyrum pernix* K1 was sequenced. The DNA sequence data along with the results of detailed RNA analyses disclosed an unusual features of the presence of three introns within the rRNA genes, whose insertion positions have not been described previously. The 699-nucleotide (nt) intron I α lies at the position 908 (*E.coli* number [Noller, H. F., Annu. Rev. Biochem. 53: 119-162, 1984]) of the 16S rRNA, and the 202-nt intron I β and the 575-nt intron I γ at positions 1085 and 1927 (*E. coli* number) of the 23S rRNA, respectively. They are located within highly conserved sites which has been implicated to be crucial for rRNA function in *E. coli*. All the three introns are remarkably AT-rich (41.5 to 43.1 mol% G+C) compared with the mature rRNAs (67.7 and 69.2 mol% G+C for 16S and 23S rRNA, respectively). No

obvious primary sequence similarities were detected among them.

The additional sequence analyses of the single rRNA operon in the other three isolates (OH2, TB1, and TB7) of *A. pernix* showed that a total of 7 introns is present within the loci. Based on the sequence differences observed in the introns, we have assigned these introns into six types ($I\alpha$; $I\beta$; $I\gamma$; $I\delta$, 62 nt; $I\epsilon$, 124 nt; and $I\zeta$, 57 nt). Interestingly, $I\alpha$ and $I\delta$ precisely occupy the same insertion sites in the *arnS* genes, although they differ in length and primary structure. As is the case observed between $I\gamma$ and $I\zeta$ in the *arnL* genes. These results lead us to suppose the presence of at least four hot spots ($I\alpha/I\delta$, $I\beta$, $I\gamma/I\zeta$, and $I\epsilon$) for intron-insertion within the *arnSL* locus of *A. pernix*.

(iii) After splicing from rRNA transcripts, they are stably retained in the cell, although high abundant circular intronic RNA molecules were not observed unlike the case of several archaeal rRNA introns so far characterized. A secondary structural model of the $I\alpha$ -containing rRNA precursors agree with the rigidly defined bulge-helix-bulge motif, whereas those of $I\beta$ and $I\gamma$ could fold into aberrant bulge-helix-bulge structures.

(iv) The two of the introns, $I\alpha$ and $I\gamma$, contain open reading frames (ORFs) whose protein translation have no overall similarity with proteins so far reported. However, both share a LAGLI-DADG motif characteristic of homing endonucleases encoded by group I introns and inteins. The existence of the putative Shine-Dalgarno sequences occurring 10-15 nt upstream of the start codons suggest that they are expressed in *A. pernix* K1 cells.

In this context, we have overexpressed the $I\alpha$ ORF (I-*Ape* I) in *E. coli* and purified the recombinant I-*Ape* I protein to homogeneity by heat treatment and chromatography. I-*Ape* I recognizes the double stranded DNA whose sequence is 5' GGCTGAAAC*TTAAAGG 3' and cleaves as indicated by the asterisk, generating four nucleotide 3'-overhangs. The enzyme reaches peak activity at temperatures ranging from 85 to 90 °C and requires Mg^{2+} . The I-*Ape* I endonucleolytic activity was also detected in the crude lysate prepared from *A. pernix* K1 cells.

Considering that the target sequence corresponds to that lies at intron-

insertion site of an intronless *arnS* allele of *Aeropyrum pernix* OH2, these result raised the possibility that I-Ape I might also promote intron mobility, as was demonstrated for the *Saccharomyces cerevisiae* mitochondrial *al4* α intronic ORF.

Acknowledgment

I would like to express my sincere gratitude to Yoshihiko Sako, Associate Professor of Marine Microbiology, Kyoto University, for his constant kind guidance and valuable discussion throughout the course of this investigation. This dissertation owes much to his critical reading and helpful suggestion, although responsibility with any surviving errors rests entirely upon me.

I am also greatly indebted to Aritsune Uchida, Professor of Marine Microbiology, Kyoto University, who gave me gentle encouragement and a number of constructive criticisms of certain sections of this investigation. I appreciate Taeko Kogishi and Yayoi Morinaga for their excellent collaborations and valuable discussions. The advice through discussion with Yuzabura Ishida, Ikuo Yoshinaga, and Ken Takai are gratefully acknowledged. Thanks are also due to the other members of Laboratory of Marine Microbiology, Division of Applied Bioscience, Graduate School of Agriculture, Kyoto University, for support of my work.

I thank Tadashi Maruyama of Marine Biotechnology Institute, Toshihiro Hoaki of Taisei Co., Ltd., and crews of the research vessel *Sogen-maru* for their help in collecting samples at Kodakara Island. Yosuke Koga and Hiroyuki Morii of the University of Occupational and Environmental Health are acknowledged for their excellent collaborations in chemical analysis of lipid. I am grateful to Iwao Furusawa, Kyoto University and Tadaaki Yoshida, Kureha Chemical Industry Co. Ltd., for electron microscopy.

I was supported by the Research Fellowship of the Japan Society for the Promotion of Science for Young Scientists.

Kyoto, 1998

野村紀通

Norimichi Nomura

References

- Achenbach, L., and C. R. Woese.** 16S and 23S rRNA-like primers. p. 15-23. *In* F. T. Robb, and A. R. Place (ed.), *Archaea, a laboratory manual, thermophiles*. Cold Spring Harbor Laboratory, Cold Spring Harbor, N.Y.
- Achenbach-Richter, L., K. O. Stetter, and C. R. Woese.** 1987a. A possible biochemical missing link among archaebacteria. *Nature (London)* **327**: 348-349.
- Achenbach-Richter, L., R. Gupta, K. O. Stetter, and C. R. Woese.** 1987b. Were the original eubacteria thermophiles? *Syst. Appl. Microbiol.* **9**: 34-39.
- Achenbach-Richter, L., R. Gupta, W. Zillig, and C. R. Woese.** 1988. Rooting the archaebacterial tree: The pivotal role of *Thermococcus celer* in archaebacterial evolution. *Syst. Appl. Microbiol.* **10**: 231-240.
- Arias, L. M., and P. Argos.** 1989. Engineering protein thermal stability: sequence statistics point to residue substitutions in α -helices. *J. Mol. Biol.* **206**: 397-406.
- Balch, W. E., and R. S. Wolfe.** 1976. New approach to the cultivation of methanogenic bacteria: 2-mercaptoethanesulfonic acid (HS-CoM)-dependent growth of *Methanobacterium ruminantium* in a pressurized atomosphere. *Appl. Environ. Microbiol.* **32**: 781-791.
- Baross, J. A.** 1995. Isolation growth, and maintenance of hyperthermophiles, p. 15-23. *In* F. T. Robb, and A. R. Place (ed.), *Archaea, a laboratory manual, thermophiles*. Cold Spring Harbor Laboratory, Cold Spring Harbor, N.Y.
- Bartlett, G. R.** 1959. Phosphorus assay in column chromatography. *J. Biol. Chem.* **234**: 466-468.
- Baumeister, W., and Lembecke, G.** 1992. Structural features of archaebacterial cell envelopes. *J. Bioenergetics. Biomembranes.* **24**: 567-575.
- Belfort, M.** 1990. Phage T4 introns: self-splicing and mobility. *Annu. Rev. Genet.* **24**: 363-385.
- Belfort, M.** 1991. Self-splicing introns in prokaryotes: migrant fossils? *Cell* **64**: 9-11.
- Belfort, M., and A. Weiner.** 1997. Another bridge between kingdoms: tRNA splicing in Archaea and Eukaryotes. *Cell* **89**: 1003-1006.
- Belfort, M., M. E. Reaban, T. Coetzee, and J. Z. Dalggaard.** 1995. Prokaryotic

- introns and inteins: a panoply of form and function. *J. Bacteriol.* **177**: 3897-3903.
- Bell-Pedersen, D., S. Quirk, J. Clyman, and M. Belfort.** 1990. Intron mobility in phage T4 is dependent upon a distinctive class of endonucleases and independent of DNA sequences encoding the intron core: Mechanistic and evolutionary implications. *Nucleic Acids Res.* **18**: 3763-3770.
- Bergquist, P. L., and H. W. Morgan.** 1992. The molecular genetics and biotechnological application of enzymes from extremely thermophilic eubacteria. *In* R. A. Herbert, and R. J. Sharp (ed.), *Molecular biology and biotechnology of extremophiles*. Blackie, London.
- Beserga, S. J., and J. A. Steitz.** 1993. The diverse world of small ribonucleoproteins, p. 359-381. *In* R. F. Gesteland, and J. F. Atkins (ed.), *The RNA world*. Cold Spring Harbor Laboratory Press, Cold Spring Harbor Laboratory, N. Y.
- Bligh, E.G., and W.J. Dyer.** 1959. A rapid method of total lipid extraction and purification. *Can. J. Biochem. Physiol.* **37**: 911-917.
- Boer, P. H., and M. W. Gray.** 1988. Scrambled ribosomal RNA gene pieces in *Chlamydomonas reinhardtii* mitochondrial DNA. *Cell* **55**: 399-411.
- Bohm, G., and R. Jaenicke.** 1994. Relevance of sequence statistics for the properties of extremophilic proteins. *Int. J. Protein Res.* **43**: 97-106.
- Britton, K. L., P. J. Baker, K. M. M. Borges, P. C. Engel, A. Pasquo, D. W. Rice, F. T. Robb, R. Scandurra, T. J. Stillman, and K. S. P. Yip.** 1995. Insight into thermal stability from a comparison of the glutamate dehydrogenases from *Pyrococcus furiosus* and *Thermococcus litoralis*. *Eur. J. Biochem.* **229**: 688-695.
- Bulgin, A. B., K. Parodos, D. J. Lane, and N. R. Pace.** 1990. The excision of intervening sequences from *Salmonella* 23S ribosomal RNA. *Cell* **60**: 405-414.
- Bult, C. J., O. White, G. J. Olsen, L. Zhou, R. D. Fleischmann, G. G. Sutton, J. A. Blake, L. M. FitzGerald, R. A. Clayton, J. D. Gocayne, A. R. Kerlavage, B. A. Dougherty, J. -F. Tomb, M. D. Adams, C. I. Reich, R. Overbeek, E. F. Kirkness, K. G. Weinstock, J. M. Merrick, A. Glodek, J. L. Scott, N. S. M. Geoghagen, J. F. Weidman, J. F. Fuhrmann, E. A. Presley, D. Nguyen, T. R. Utterback, J. M. Kelley, J. D. Peterson, P. W. Sadow, M. C. Hanna, M. D. Cotton, M. A. Hurst, K. M. Roberts, B. P.**

- Kaine, M. Borodovsky, H. -P. Klenk, C. M. Fraser, H. O. Smith, C. R. Woese, and J. C. Venter.** 1996. Complete genome sequence of the methanogenic archaeon, *Methanococcus jannaschii*. *Science* **273**: 1058-1073.
- Burggraf, S., N. Larsen, C. R. Woese, and K. O. Stetter.** 1993. An intron within the 16S ribosomal RNA gene of the archaeon *Pyrobaculum aerophilum*. *Proc. Natl. Acad. Sci. USA* **90**: 2547-2550.
- Cannio, R., M. Rossi, and S. Bartolucci.** 1994. A few amino acid substitutions are responsible for the higher thermostability of a novel NAD⁺-dependent bacillar alcohol dehydrogenase. *Eur. J. Biochem.* **222**: 345-352.
- Cavalier-Smith, T.** 1992. Bacteria and eukaryotes. *Nature (London)* **356**: 570.
- Chan, M. K., S. Mukund, A. Kletzin, M. W. W. Adams, and D. C. Rees.** 1995. Structure of a hyperthermophilic tungstopterin enzyme, aldehyde ferredoxin oxidoreductase. *Science* **267**: 1463-1469.
- Chant, J., and P. P. Dennis.** 1986. Archaeobacteria: transcription and processing of ribosomal RNA sequences in *Halobacterium cutirubrum*. *EMBO J.* **5**: 1091-1097.
- Chomczynski, P., and N. Sacchi.** 1987. Single-step method of RNA isolation by acid guanidinium thiocyanate-PhOH-chloroform extraction. *Anal. Biochem.* **162**: 156-159.
- Chu, F. K., G. F. Maley, F. Maley, and M. Belfort.** 1984. Intervening sequence in the thymidylate synthase gene of bacteriophage T4. *Proc. Natl. Acad. Sci. USA* **81**: 3049-3053.
- Colleaux, L., L. d' Auriol, M. Betermier, G. Cottarel, A. Jacquier, F. Galibert, and B. Dujon.** 1986. Universal code equivalent of a yeast mitochondrial intron reading frame is expressed in *E. coli* as a specific double strand endonuclease. *Cell* **44**: 521-533.
- Colleaux, L., M. -R. Michel, R. F. Matagne, and B. Dujon.** 1990. The apocytochrome *b* gene of *Chlamydomonas smithii* contains a mobile intron related to both *Saccharomyces* and *Neurospora* introns. *Mol. Gen. Genet.* **223**: 288-296.
- Colleaux, L., L. d' Auriol, F. Galibert, and B. Dujon.** 1988. Recognition and cleavage site of the intron-encoded omega transposase. *Proc. Natl. Acad. Sci. USA* **85**: 6022-6026.

- Cowan, D. A.** 1992. Biotechnology and molecular biology of the extremely thermophilic archaeobacteria. In R. A. Herbert, and R. J. Sharp (ed.), Molecular biology and biotechnology of extremophiles. Blackie, London.
- Cummings, D. J., F. Michel, and K. L. McNally.** 1989. DNA sequence analysis of the apocytochrome b gene of *Podospira anserina*: A new family of intronic open reading frame. *Curr. Genet.* **16**: 407-418.
- Dalgaard, J. Z., and R. A. Garrett.** 1992. Protein-coding introns from the 23S rRNA-encoding gene form stable circles in the hyperthermophilic archaeon *Pyrobaculum organotrophum*. *Gene* **121**: 103-110.
- Datta, P. K., L. K. Hawkins, and R. Gupta.** 1989. Presence of an intron in elongator methionine-tRNA of *Halobacterium volcanii*. *Can. J. Microbiol.* **35**: 189-194.
- Davis, E. O., P. J. Jenner, P. C. Brooks, M. J. Colston, and S. G. Sedgwick.** 1992. Protein splicing in the maturation of *Mycobacterium tuberculosis* RecA protein: A mechanism for tolerating a novel class of intervening sequence. *Cell* **71**: 201-210.
- Delahodde, A., V. Goguel, A. M. Becam, F. Creusot, J. Banroques, and C. Jacq.** 1989. Site-specific DNA endonuclease and RNA maturase activities of two homologous intron encoded proteins from yeast mitochondria. *Cell* **56**: 431-442.
- DeLong, E. F.** 1992. Archaea in coastal marine environments. *Proc. Natl. Acad. Sci. USA* **89**: 5685-5689.
- Dennis, P. P.** 1991. The ribosomal RNA operons of halophilic archaeobacteria, p. 251-258. In F. Rodriguez-Valera (ed.), General and applied aspects of halophilic microorganisms. Plenum Press, N. Y.
- Dennis, P. P.** 1997. Ancient ciphers: Translation in archaea. *Cell* **89**: 1007-1010.
- De Rijk, P., Y. Van de Peer, and R. De Wachter.** 1997. Database on the structure of large ribosomal subunit RNA. *Nucleic Acids Res.* **24**: 117-123.
- DeRosa, M., and A. Gambacorta.** 1988. The lipids of archaeobacteria. *Prog. Lipid Res.* **27**: 153-157.
- De Rosa, M., A. Gambacorta, A. Trincone, A. Basso, W. Zillig, and I. Holz.** 1987. Lipids of *Thermococcus celer*, a sulfur-reducing archaeobacterium: structure and biosynthesis. *Syst. Appl. Microbiol.* **9**: 1-5.

- De Rosa, M., A. Gambacorta, B. Nicolaus, and W. D. Grant.** 1983. A C₂₅, C₂₅ diether core lipid from archaebacterial haloalkaliphiles. *J. Gen. Microbiol.* **129**: 2333-2337.
- DiRuggiero, J., L. A. Achenbach, S. H. Brown, R. M. Kelly, and F. T. Robb.** 1993. Regulation of ribosomal RNA transcription by growth rate of the hyperthermophilic archaeon, *Pyrococcus furiosus*. *FEMS Microbiol. Lett.* **111**: 159-164.
- Dover, G., and E. Coen.** 1981. Spring cleaning ribosomal DNA: a model for multigene evolution? *Nature (London)* **290**: 731-732.
- Dujon, B.** 1989. Group I introns as mobile genetic elements: facts and mechanistic speculations- a review. *Gene* **82**: 91-113.
- Dujon, B., M. Belfort, R. A. Butow, C. Jacq, C. Lemieux, P. S. Perlman, and V. M. Vogt.** Mobile introns: definitions of terms and recommended nomenclature. *Gene* **82**: 115-118.
- Durovic, P., and P. P. Dennis.** 1994. Separate pathways for excision and processing of 16S and 23S rRNA from the primary rRNA operon transcript from the hyperthermophilic archaebacterium *Sulfolobus acidocaldarius*: similarities to eukaryotic rRNA processing. *Mol. Microbiol.* **13**: 229-242.
- Felsenstein, J.** 1985. Confidence limits on phylogenies: an approach using bootstrap. *Evolution* **39**: 783-791.
- Ferat, J. -F., and F. Michel.** 1993. Group II self-splicing introns in bacteria. *Nature* **364**: 358-361.
- Ford, E., and M. Ares, Jr.** 1994. Synthesis of circular RNA in bacteria and yeast using RNA cyclase ribozymes derived from a group I intron of phage T4. *Proc. Natl. Acad. Sci. USA* **91**: 3117-3121.
- Gadd, G. M.** 1992. Molecular biology and biotechnology of microbial interactions with organic and inorganic heavy metal compounds. *In* R. A. Herbert, and R. J. Sharp (ed.), *Molecular biology and biotechnology of extremophiles*. Blackie, London.
- Garrett, R. A., J. Dalgaard, N. Larsen, J. Kjems, and A. S. Mankin.** 1991. Archaeal rRNA operon. *Trends Biochem. Sci.* **16**: 22-26.
- Gogarten, J. P., B. Kibak, P. Dittrich, I. Taiz, E. J. Bowman, B. J. Bowman, M. F. Manolson, R. J. Poole, T. Date, T. Oshima, J. Konishi, K. Denda, and M. Yoshida.** 1989. Evolution of the vacuolar H⁺-ATPases: Implications

for the origin of eukaryotes. *Proc. Natl. Acad. Sci. USA* **86**: 6661-6665.

Goldrich-Blair, H., V. Scarlato, J. M. Gott, M. -Q. Xu, and D. A. Shub. 1990. A self-splicing group I intron in the DNA polymerase gene of *Bacillus subtilis* bacteriophage SPO1. *Cell* **63**: 417-424.

Gourse, R. L., T. Gaal, M. S. Bartlett, J. A. Appleman, and W. Ross. 1996. rRNA transcription and growth rate-dependent regulation of ribosome synthesis in *Escherichia coli*. *Annu. Rev. Microbiol.* **50**: 645-677.

Hain, J., W. D. Reiter, U. Hudepohl, and W. Zillig. 1992. Elements of an archaeal promoter defined by mutational analysis. *Nucleic Acids Res.* **20**: 5423-5428.

Henikoff, S. 1984. Unidirectional digestion with exonuclease III creates targeted breakpoints for DNA sequencing. *Gene* **28**: 351-359.

Hernandez-Veerduin, D. 1991. The nucleolus today. *J. Cell. Sci.* **99**: 465-471.

Hilpert, R., J. Winter, W. Hammes, and O. Kandler. 1981. The sensitivity of archaebacteria to antibiotics. *Zentralbl. Bakteriol. Parasitenkd. Infektionskr. Hyg. Abt. 1 Orig. Reihe C* **2**: 11-20.

Hoaki, T., M. Nishijima, M. Kato, K. Adachi, S. Mizobuchi, N. Hanzawa, and T. Maruyama. 1994. Growth requirements of hyperthermophilic sulfur-dependent heterotrophicarchaea isolated from a shallow submarine geothermal system with reference to their essential amino acids. *Appl. Environ. Microbiol.* **60**: 2898-2904.

Hoaki, T., M. Nishijima, H. Miyashita, and T. Maruyama. 1995. Dense community of hyperthermophilic sulfur-dependent heterotrophs in a geothermally heated shallow submarine biotope near Kodakara-Jima Island, Kagoshima, Japan. *Appl. Environ. Microbiol.* **61**: 1931-1937.

Hodges, R. A., F. B. Perler, C. J. Noren, and W. E. Jack. 1992. Protein splicing removes intervening sequences in an archaea DNA polymerase. *Nucleic Acids Res.* **20**: 6153-6157.

Hsu, D., L. Shih, and Y. Zee. 1994. Degradation of rRNA in *Salmonella* strains: a novel mechanism to regulate the concentration of rRNA and ribosomes. *J. Bacteriol.* **176**: 4761-4765.

Hsu, D., M. Pan, Y. Zee, and R. B. LeFebvre. 1990. Unique ribosomal structure of *Leptospira interrogans* is composed of four rRNA component. *J.*

Bacteriol. **172**: 3478-3480.

Hsu, D., Y. Zee, J. Ingraham, and L. Shih. 1992. Diversity of cleavage patterns in *Slmonella* 23S rRNA. J. Gen. Microbiol. **138**: 199-203.

Huber, G., C. Spinnler, A. Gambacorta, and K. O. Stetter. 1989. *Metallospaera sedula* gen. and sp. nov. represents a new genus of aerobic, metal-mobilizing, thermoacidophilic archaeobacteria. Syst. Appl. Microbiol. **12**: 38-47.

Huber, R., P. Stoffers, J. L. Cheminee, H. H. Richnow, and K. O. Stetter. 1990. Hyperthermophilic archaeobacteria within the crater and open-sea plume of erupting Macdonald Seamount. Nature (London) **345**: 179-182.

Huber, R., T. A. Langworthy, H. Konig, M. Thomm, C. R. Woese., U. B. Sleytr, and K. O. Stetter. 1986. *Thermotoga maritima* sp. nov. represents a new genus of unique extremely thermophilic eubacteria growing up to 90 °C. Arch. Microbiol. **144**: 324-333.

Huber, R., T. Wilharm, D. Huber, A. Trincone, S. Burggraf, H. Konig, R. Rachel, I. Rockinger, H. Fricke, and K. O. Stetter. 1992. *Aquifex pyrophilus* gen. nov. represents a novel group of marine hyperthermophilic hydrogen-oxidizing bacteria. Syst. Appl. Microbiol. **15**: 340-351.

Iwabe, N., K. Kuma, M. Hasegawa, S. Osawa, and T. Miyata. 1989. Evolutionary relationship of archaeobacteriaeubacteria and eukaryotes inferred from phylogenetic trees of duplicated genes. Proc. Natl. Acad. Sci. USA **86**: 9355-9359.

Jaquir, A., and B. Dujon. 1985. An intron-encoded protein is active in a gene conversion process that spreads an intron into a mitochondrial gene. Cell **41**: 383-394.

Kaine, B. P. 1987. Intron containing tRNA genes of *Sulfolobus solfataricus*. J. Mol. Evol. **25**: 248-254.

Kaine, B. P., R. Gupta, C. R. Woese. 1983. Putative introns in tRNA genes of prokaryotes. Proc. Natl. Acad. Sci. USA **80**: 3309-3312.

Kasting, J. F. 1987. Theoretical constraints on oxygen and carbon dioxide concentrations in the Precambrian atmosphere. Precambrian Res. **34**: 205-229.

Kjems, J., and R. A. Garrett. 1985. An intron in the 23S ribosomal RNA gene of the archaeobacterium *Desulfurococcus mobilis*. Nature **318**: 675-677.

Kjems, J., and R. A. Garrett. 1987. Novel expression of the ribosomal

RNA genes in the extreme thermophile and archaeobacterium *Desulfurococcus mobilis*. EMBO J. **6**: 3521-3530.

Kjems, J., and R. A. Garrett. 1988. Novel splicing mechanism for the ribosomal RNA intron in the archaeobacterium *Desulfurococcus mobilis*. Cell **54**: 693-703.

Kjems, J., and R. A. Garrett. 1991. Ribosomal RNA introns in archaea and evidence for RNA conformational changes associated with splicing. Proc. Natl. Acad. Sci. USA **88**: 439-443.

Kleman-Leyer, K., D. W. Armbruster, and C. J. Daniels. 1997. Properties of *H. volcanii* tRNA intron endonuclease reveal a relationship between the Archaeal and Eucaryal tRNA intron processing systems. Cell **89**: 839-847.

Klenk, H. -P., R. A. Clayton, J. -F. Tomb, O. White, K. E. Nelson, K. A. Ketchum, R. J. Dodson, M. Gwinn, E. K. Hickey, J. D. Pederson, D. L. Richardson, A. R. Kerlavage, D. E. Graham, N. C. Kyrpides, R. D. Fleischmann, J. Quackenbush, N. H. Lee, G. G. Sutton, S. Gill, E. F. Kirkness, B. A. Dougherty, K. McKenney, M. D. Adams, B. Loftus, S. Peterson, C. I. Reich, L. K. McNeil, J. H. Badger, A. Glodek, L. Zhou, R. Overbeek, J. D. Gocayne, J. F. Weidman, L. McDonald, T. Utterback, M. D. Cotton, T. Spriggs, P. Artiach, B. P. Kaine, S. M. Sykes, P. W. Sadow, K. P. D'Andrea, C. Bowman, C. Fujii, S. A. Garland, T. M. Mason, G. J. Olsen, C. M. Fraser, H. O. Smith, C. R. Woese, and J. C. Venter. 1997. The complete genome sequence of the hyperthermophilic, sulfate-reducing archaeon *Archaeoglobus fulgidus*. Nature (London) **390**: 364-370.

Koga, Y., M. Nishihara, H. Morii, and M. Akagawa-Matsushita. 1993. Ether polar lipids of methanogenic bacteria: structures, comparative aspects, and biosyntheses. Microbiol. Rev. **57**: 164-182.

Konig, H., and K. O. Stetter. 1989. Section 25. Archaeobacteria, p. 2171-2173. In J. T. Staley, M. P. Bryant, N. Pfennig, and J. G. Holt (ed.), Bergey's manual of systematic bacteriology, vol. 3. The Williams & Wilkins Co., Baltimore.

Konkel, M. E., R. T. Marconi, D. J. Mead, and W. Ceiplak. 1994. Identification and characterization of an intervening sequence within the 23S ribosomal RNA genes of *Campylobacter jejuni*. Mol. Microbiol. **14**: 235-241.

Kordes, E., S. Jock, J. Fritsch, F. Bosch, and G. Klug. 1994. Cloning of a gene involved in rRNA precursor processing and 23S rRNA cleavage in

- Rhodobacter capsulatus*. J. Bacteriol. **176**: 1121-1127.
- Kotsuka, T., S. Akamura, M. Tomuro, A. Yamagishi, and T. Oshima.** 1996. Further stabilization of 3-isopropylmalate dehydrogenase of an extreme thermophile, *Thermus thermophilus*, by a suppressor mutation method. J. Bacteriol. **178**: 723-727.
- Kuhse, M. G., R. Strickland, and J. D. Palmer.** 1990. An ancient group I intron shared by eubacteria and chloroplasts. Science **250**: 1570-1573.
- Lambowitz, A. M.** 1989. Infectious introns. Cell **56**: 323-326.
- Lambowitz, A. M., and M. Belfort.** 1993. Introns as mobile genetic elements. Annu. Rev. Biochem. **62**: 587-622.
- Lambowitz, A. M., and P. S. Perlman.** 1990. Involvement of aminoacyl-transfer RNA-synthetases and other proteins in group-I and group-II intron splicing. Trends Biochem. Sci. **15**: 440-444.
- Langworthy, T. A., and P. F. Smith.** 1989. Group IV. Cell wall-less archaeobacteria, p. 2233-2236. In J. T. Staley, M. P. Bryant, N. Pfennig, and J. G. Holt (ed.), Bergey's manual of systematic bacteriology, vol. 3. The William & Wilkins Co., Baltimore.
- Leffers, H., J. Kjems, L. Ostergaard, N. Larsen, and R. A. Garrett.** 1987. Evolutionary relationship amongst archaeobacteria. A comparative study of a sulphur-dependent extreme thermophile, an extreme halophile and a thermophilic methanogen. J. Mol. Biol. **195**: 43-61.
- Lemieux, C., and R. W. Lee.** 1987. Non-reciprocal recombination between alleles of the chloroplast 23S rRNA gene in interspecific *Chlamydomonas* crosses. Proc. Natl. Acad. Sci. USA **84**: 4166-4170.
- Lessie, T. G.** 1965. The atypical ribosomal RNA complement of *Rhodospseudomonas sphaeroides*. J. Gen. Microbiol. **39**: 311-320.
- Linton, D., F. E. Dewhirst, J. P. Clewley, R. J. Owen, A. P. Burnens, and J. Stanley.** 1994. Two types of 16S rRNA gene are found in *Campylobacter helveticus*: analysis, applications and characterization of the intervening sequence found in some strains. Microbiology **140**: 847-855.
- Lykke-Andersen, J., and R. A. Garrett.** 1994. Structural characteristics of the stable RNA introns of archaeal hyperthermophiles and their splicing junctions. J. Mol. Biol. **243**: 846-855.
- Marmur, J., and P. Doty.** 1962. Determination of the base composition of

deoxyribonucleic acid from its thermal denaturation temperature. *J. Mol. Biol.* **5**: 109-118.

Mesbah, M., U. Premachandran, and W. B. Whitman. 1989. Precise measurement of the G+C content of deoxyribonucleic acid by high-performance liquid chromatography. *Int. J. Syst. Bacteriol.* **39**: 159-167.

Michel, F., and B. Dujon. 1986. Genetic exchanges between bacteriophage T4 and filamentous fungi? *Cell* **46**: 323.

Moazed, D., and H. F. Noller. 1986. Transfer RNA shields specific nucleotides in 16S ribosomal RNA from attack by chemical probes. *Cell* **47**: 985-994.

Moazed, D., and H. F. Noller. 1989. Interaction of tRNA with 23S rRNA in the ribosomal A, P, and E sites. *Cell* **57**: 585-597.

Moazed, D., and H. F. Noller. 1990. Binding of tRNA to the ribosomal A and P site protects two distinct sets of nucleotides in 16S rRNA. *J. Mol. Biol.* **211**: 135-145.

Morikawa, M., Y. Izawa, N. Rashid, T. Hoaki, and T. Imanaka. 1994. Purification and characterization of a thermostable thiol protease from a newly isolated hyperthermophilic *Pyrococcus* sp. *Appl. Environ. Microbiol.* **60**: 4559-4566.

Mueller, J. E., M. Bryk, N. Loizos, and M. Belfort. 1993. Homing endonucleases, p. 111-143. *In* S. M. Linn, R. S. Lloyd, and R. J. Roberts (ed.), *Nucleases*. Cold Spring Harbor Laboratory Press, Cold Spring Harbor, N.Y.

Murray, M.G., and W.F. Thompson. 1980. Rapid isolation of high molecular-weight plant DNA. *Nucleic Acids Res.* **8**: 4321-4325.

Muscarella, D. E., and V. M. Vogt. 1989. A mobile group I intron in the nuclear rDNA of *Physarum polycephalum*. *Cell* **56**: 443-454.

Nakagawa, K., N. Morishima, and T. Shibata. 1991. A maturase-like subunit of the sequence-specific endonuclease Endo. *Sce* I from yeast mitochondria. *J. Biol. Chem.* **266**: 1977-1984.

Nishihara, M., H. Morii, and Y. Koga. 1987. Structure determination of a quartet of novel tetraether lipids from *Methanobacterium thermoautotrophicum*. *J. Biochem.* **101**: 1007-1015.

Nishihara, M., H. Morii, and Y. Koga. 1989. Heptads of polar ether lipids of an archaebacterium, *Methanobacterium thermoautotrophicum*: structure and biosynthetic relationship. *Biochemistry* **28**: 95-102.

- Nishihara, M., and Y. Koga.** 1988. Quantitative conversion of diether or tetraether phospholipids to glycerophosphoesters by dealkylation with boron trichloride: a tool for structural analysis of archaebacterial lipids. *J. Lipid Res.* **29**: 384-388.
- Nishihara, M., and Y. Koga.** 1991. Hydroxyarchaetidylserine and hydroxyarchaetidyl-*myo*-inositol in *Methanosarcina barkeri*: polar lipids with a new ether core portion. *Biochim. Biophys. Acta* **1082**: 211-217.
- Noller, H. F.** 1984. Structure of ribosomal RNA. *Annu. Rev. Biochem.* **53**: 119-162.
- Ohmoto, H., T. Kanegawa, and D. R. Lowe.** 1993. 3.4-billion-year-old biogenic pyrites from Barberton, South Africa: Sulfur isotope evidence. *Science* **262**: 555-557.
- Olsen, G. J., and C. R. Woese.** 1997. Archaeal genomics: An overview. *Cell* **89**: 991-994.
- Olsen, G. J., C. R. Woese, and R. Overbeek.** 1994. The wind of (evolutionary) change: Breathing new life into microbiology. *J. Bacteriol.* **176**: 1-6.
- Otsuka, T., H. Nomiya, H. Yoshida, T. Kukita, S. Kuhara, and Y. Sakaki.** 1983. Complete nucleotide sequence of the 26S ribosomal RNA gene of *Physarum polycephalum*: its significance in gene evolution. *Proc. Natl. Acad. Sci. USA* **80**: 3163-3167.
- Palmer, J. R., and Daniels, C. J.** 1995. In vivo definition of an archaeal promoter. *J. Bacteriol.* **177**: 1844-1849.
- Pearson, W. R., and D. J. Lipman.** 1988. Improved tools for biological sequence comparison. *Proc. Natl. Acad. Sci. USA* **85**: 2444-2448.
- Perler, F. B., D. G. Comb, W. E. Jack, L. S. Moran, B. Quiang, R. B. Kucera, J. Benner, B. E. Slatko, D. O. Nwankwo, S. K. Hempstead, C. K. S. Carlow, and H. Jannasch.** 1992. Intervening sequences in an Archaea DNA polymerase gene. *Proc. Natl. Acad. Sci. USA* **89**: 5577-5581.
- Perlman, P. S., and R. A. Butow.** 1989. Mobile introns and intron-encoded proteins. *Science*. **246**: 1106-1109.
- Potter, S., P. Durovic, and P. P. Dennis.** 1995. Ribosomal RNA precursor processing by a eukaryotic U3 small nucleolar RNA-like molecule in an archaeon. *Science* **268**: 1056-1060.
- Puhler, G., H. Leffersy, F. Gropp, P. Palm, H.- P. Klenk, F. Lottspeich,**

- R. A. Garrett, W. Zillig.** 1989. Archaeobacterial DNA-dependent RNA polymerases testify to the evolution of the eukaryotic nuclear genome. *Proc. Natl. Acad. Sci. USA* **86**: 4569-4573.
- Qu, L. H., B. Michot, and J. -P. Bachellerie.** 1983. Improved methods for structure probing in large RNAs: A rapid 'heterologous' sequencing approach is coupled to the direct mapping of nuclease accessible sites. Application to the 5' terminal domain of eukaryotic 28S rRNA. *Nucleic Acids Res.* **11**: 5903-5920.
- Quirk, S. M., D. Bell-Pedersen, and M. Belfort.** 1989. Intron mobility in the T-even phages: High frequency inheritance of group I introns promoted by intron open reading frames. *Cell* **56**: 455-465.
- Ralph, D., and M. McClelland.** 1994. Phylogenetic evidence for horizontal transfer of an intervening sequence between species in a spirochete genus. *J. Bacteriol.* **19**: 5982-5987.
- Reinhold-Hurek, B., and D. A. Shub.** 1992. Self-splicing introns in tRNA genes of widely divergent bacteria. *Nature* **357**: 173-176.
- Reiter, W. D., P. Palm, and W. Zillig.** 1988. Analysis of transcription in the archaeobacterium *Sulfolobus* indicates that archaeobacterial promoters are homologous to eukaryotic pol II promoters. *Nucleic Acids Res.* **16**: 1-19.
- Remacle, C., C. Bovie, M. -R. Michel-Wolwertz, R. Loppes, and R. F. Matagne.** Mitochondrial genome transmission in *Chlamydomonas* diploids obtained by sexual crosses and artificial fusions: role of the mating type and of a 1 kb intron. *Mol. Gen. Genet.* **223**: 180-184.
- Mohr, G., and A. M. Lambowitz.** 1991. Integration of a group I intron into a ribosomal RNA sequence promoted by a tyrosyl-tRNA synthetase. *Nature (London)* **354**: 164-167.
- Mota, E. M., and R. A. Collins.** 1988. Independent evolution of structural and coding regions in a *Neurospora* mitochondrial intron. *Nature (London)* **332**: 654-656.
- Saitou, N., and M. Nei.** 1987. The neighbor-joining method: A new method for reconstructing phylogenetic trees. *Mol. Biol. Evol.* **4**: 406-425.
- Sako, Y., N. Nomura, A. Uchida, Y. Ishida, H. Morii, Y. Koga, T. Hoaki, and T. Maruyama.** 1996. *Aeropyrum pernix* gen. nov., sp. nov., a novel aerobic hyperthermophilic archaeon growing at temperatures up to 100°C. *Int.*

J. Syst. Bacteriol. **46**: 1070-1077.

Sambrook, J. E., E. Fritsch, and T. Maniatis. 1989. Molecular cloning: a laboratory manual, 2nd ed. Cold Spring Harbor Laboratory Press, Cold Spring Harbor, N. Y.

Sanger, F., A. R. Coulson, G. F. Hong, O. F. Hill, and G. B. Petersen. 1982. Nucleotide sequence of bacteriophage λ DNA. J. Mol. Biol. **162**: 729-773.

Sanger, F., S. Nicklen, and A. R. Coulson. 1977. DNA sequencing with chain-terminating inhibitors. Proc. Natl. Acad. Sci. USA **74**: 5463-5467.

Schultes, V., and R. Jaenicke. 1991. Folding intermediates of hyperthermophilic D-glyceraldehyde-3-phosphate dehydrogenase from *Thermotoga maritima* are trapped at low temperature. FEBS Lett. **290**: 235-238.

Schnare, M. N., and M. N. Gray. 1990. Sixteen discrete RNA components in the cytoplasmic ribosome of *Euglena gracilis*. J. Mol. Biol. **215**: 73-83.

Seegerer, A., and K. O. Stetter. 1989. Genus I. *Sulfolobus* Brock, Brock, Belly and Weiss 1972, 66^{AL}, p.2250-2251. In J. T. Staley, M. P. Bryant, N. Pfennig and J. G. Holt (ed.), Bergey's manual of systematic bacteriology, vol. 3. The Williams & Wilkins Co., Baltimore.

Seegerer, A., and K. O. Stetter. 1989. Genus II. *Acidianus* Seegerer, Neuner, Kristjansson and Stetter 1986, 562^{VP}, p.2251-2253. In J. T. Staley, M. P. Bryant, N. Pfennig, and J. G. Holt (ed.), Bergey's manual of systematic bacteriology, vol. 3. The Williams & Wilkins Co., Baltimore.

Sharma, M., R. L. Ellis, D. M. Hinton. 1992. Identification of a family of bacteriophage T4 genes encoding proteins similar to those present in group I introns of fungi and phage. Proc. Natl. Acad. Sci. USA **89**: 6658-6662.

Shub, D. A., and H. Goodrich-Blair. 1992. Protein introns: A new home for endonucleases. Cell. **71**: 183-186.

Skurnik, M., and P. Toivanen. 1991. Intervening sequences (IVSs) in the 23S ribosomal RNA genes of pathogenic *Yersinia enterocolitica*. The IVSs in *Y. enterocolitica* and *Salmonella typhimurium* have a common origin. Mol. Microbiol. **5**: 585-593.

Smith, D. R., L. A. Doucette-Stamm, C. Deloughery, H. Lee, J. Dubois, T. Aldredge, R. Bashirzadeh, D. Blakely, R. Cook, K. Gilbert, D. Harrison, L. Hoang, P. Keagle, W. Lumm, B. Pothier, D. Qiu, R. Spadafora, R.

- Vicaire, Y. Wang, J. Wierzbowski, R. Gibson, N. Jiwani, A. Caruso, D. Bush, H. Safer, D. Patwell, S. Prabhakar, S. McDougall, G. Shimer, A. Goyal, S. Pietrokovski, G. M. Church, C. J. Daniels, J. Mao, P. Rice, J. Nolling, and J. N. Reeve.** 1997. Complete genome sequence of *Methanobacterium thermoautotrophicum* Δ H: Functional analysis and comparative genomics. *J. Bacteriol.* **179**: 7135-7155.
- Spencer, D. F., J. C. Collings, M. N. Schnare, and M. W. Gray.** 1987. Multiple spacer sequences in the nuclear large subunit ribosomal RNA gene of *Crithidia fasciculata*. *EMBO J.* **6**: 1063-1071.
- Stetter, K. O.** 1996. Hyperthermophilic prokaryotes. *FEMS Microbiol. Rev.* **18**: 149-158.
- Stetter, K. O., G. Fiala, G. Huber, R. Huber, and A. Segerer.** 1990. Hyperthermophilic microorganisms. *FEMS Microbiol. Rev.* **75**: 117-124.
- Stetter, K. O., H. Konig, and E. Stackebrandt.** 1983. *Pyrodictium* gen. nov., a new genus of submarine disc-shaped sulphur reducing archaeobacteria growing optimally at 105 °C. *Syst. Appl. Microbiol.* **4**: 535-551.
- Studier, F. W., A. H. Rosenberg, J. J. Dunn, and J. W. Dubendorff.** 1990. Use of T7 RNA polymerase to direct expression to cloned genes. *Methods Enzymol.* **185**: 60-89.
- Szostak, J. W., T. L. Orr-Weaver, R. J. Rothstein, and F. W. Stahl.** 1983. The double-strand break repair model for recombination. *Cell* **33**: 25-35.
- Tamakoshi, M., A. Yamagishi, and T. Oshima.** 1995. Screening of stable proteins in an extreme thermophile, *Thermus thermophilus*. *Mol. Microbiol.* **16**: 1031-1036.
- Thomm, M., and G. Wich.** 1988. An archaeobacterial promoter element for stable RNA genes with homology to the TATA box of higher eukaryotes. *Nucleic Acids Res.* **16**: 151-163.
- Thompson, J. D., D. G. Higgins, and T. J. Gibson.** 1994. Clustal W: improving the sensitivity of progressive multiple sequence alignment through sequence weighting, position-specific gap penalties and weight matrix choice. *Nucleic Acids Res.* **22**: 4673-4680.
- Thompson, J., F. Schmidt, and E. Cundliff.** 1982. Site of action of a ribosomal RNA methylase conferring resistance to thiostrepton. *J. Biol. Chem.* **257**: 7915-7917.

- Thompson, L. D., and C. J. Daniels.** 1988. A tRNA^{1P} intron endonuclease from *Halobacterium volcanii*. Unique substrate recognition properties. J. Biol. Chem. **263**: 17951-17959.
- Towe, K. M.** 1990. Aerobic respiration in the Archaean? Nature (London) **348**: 54-56.
- Traub, S., S. Mizushima, C. V. Lowry, and M. Nomura.** 1971. Reconstitution of ribosomes from subribosomal components. Methods Enzymol. **20**: 391-407.
- Trust, T. J., S. M. Longan, C.E. Gustafson, P. J. Romaniuk, N. W. Kim, V. L. Chan, M. A. Ragan, P. Guerry, and R. R. Gutell.** 1994. Phylogenetic and molecular characterization of a 23S rRNA gene positions the genus *Campylobacter* in the epsilon subdivision of the Proteobacteria and shows that the presence of transcribed spacers is common in *Campylobacter* spp. J. Bacteriol. **176**: 4597-4609.
- Turmel, M., J. Boulanger, M. N. Schnare, M. W. Gray, C. Lemieux.** 1991. Six group I introns and three internal transcribed spacers in the chloroplast large subunit ribosomal RNA gene of the green alga *Chlamydomonas eugametos*. J. Mol. Biol. **218**: 293-311.
- Tuttle, J. H.** 1980. Thiosulfate oxidation and tetrathionate reduction by intact cells of marine pseudomonad 16B. Appl. Environ. Microbiol. **39**: 1159-1166.
- Tuttle, J. H., and H. W. Jannasch.** 1972. Occurrence and types of *Thiobacillus*-like bacteria in the sea. Limnol. Oceanogr. **17**: 532-543.
- Tuttle, J. H., and H. W. Jannasch.** 1973. Sulfide and thiosulfate-oxidizing bacteria in anoxic marine basins. Mar. Biol. **20**: 64-70.
- Tuttle, J. H., and H. W. Jannasch.** 1977. Thiosulfate stimulation of microbial dark assimilation of carbon dioxide in shallow marine waters. Microb. Ecol. **4**: 9-25.
- Veinstein, M. B.** 1976. On taxonomy of *Thiobacillus trautweinii*. Microbiology (USSR) **45**: 137-141.
- Volkel, P., R. Huber, E. Drobner, R. Rachel, S. Burggraf, A. Trincone, and K. O. Stetter.** 1993. *Pyrobaculum aerophilum* sp. nov., a novel nitrate-reducing hyperthermophilic archaeum. Appl. Environ. Microbiol. **59**: 2918-2926.
- Wenzlau, J. M., R. J. Saldanha, R. A. Butow, and P. S. Perlman.** 1989. A

latent intron-encoded maturase is also an endonuclease needed for intron mobility. *Cell* **56**: 421-430.

Wich, G., W. Leinfelder, A. Bock. 1987. Genes for stable RNA in the extreme thermophile *Thermoproteus tenax* introns and transcription signals. *EMBO J.* **6**: 523-528.

Winker, S., and C. R. Woese. 1991. A definition of the domains Archaea, Bacteria, and Eucarya in terms of small subunit ribosomal RNA characteristics. *Systm. Appl. Microbiol.* **14**: 305-310.

Woese, C. R. 1987. Bacterial evolution. *Microbiol. Rev.* **51**: 221-271.

Woese, C. R., and G. E. Fox. 1977. Phylogenetic structure of the prokaryotic domain: the primary kingdoms. *Proc. Natl. Acad. Sci. USA* **74**: 5088-5090.

Woese, C. R., L. Achenbach, P. Rouviere, and L. Mandelco. 1991. Archaeal phylogeny: Reexamination of the phylogenetic position of *Archaeoglobus fulgidus* in light of certain composition-induced artifacts. *Syst. Appl. Microbiol.* **14**: 364-371.

Woese, C. R., O. Kandler, and M. L. Wheelis. 1990. Towards a natural system of organisms: Proposal for the domains Archaea, Bacteria, and Eucarya. *Proc. Natl. Acad. Sci. USA* **87**: 4576-4579.

Woodson, S. A., and T. R. Cech. 1989. Reverse self-splicing of the Tetrahymena group I intron: Implication for the directionality of splicing and for intron transposition. *Cell* **57**: 335-345.

Xu, M.-Q., S. D. Kathe, H. Golgrich-Blair, S. A. Nierzwicki-Bauer, D. A. Shub. 1990. Bacterial origin of a chloroplast intron: Conserved self-splicing group I introns in cyanobacteria. *Science* **250**: 1566-1570.

Zillig, W., I. Holz, D. Janekovic, H.-P. Klenk, E. Imself, J. Trent, S. Wunderl, V. H. Forjaz, R. Coutinho, and T. Ferreira. 1990. *Hyperthermus butylicus*, a hyperthermophilic sulfur-reducing archaebacterium that ferments peptides. *J. Bacteriol.* **172**: 3959- 3965.

Zillig, W., K. O. Stetter, D. Prangishvilli, W. Schafer, S. Wunderl, D. Janekovic, I. Holz, and P. Palm. 1982. *Desulfurococcaceae*, the second family of the extremely thermophilic, anaerobic, sulfur-respiring *Thermoproteales*. *Zentralbl. Bakteriell. Parasitenkd. Infektionskr. Hyg. Abt. 1 Orig. Reihe C* **3**: 304-317.

Zillig, W., K. O. Stetter, S. Wunderl, W. Schulz, H. Priess, and I. Scholz.

1980. The *Sulfolobus*-*Caldariella* Group: Taxonomy on the basis of the structure of DNA-dependent RNA polymerases. Arch. Microbiol. **125**: 259-269.

Zillig, W., S. Yeats, I. Holz, A. Bock, M. Rettenberger, F. Gropp, and G. Simon. 1986. *Desulfurolobus ambivalens*, gen. nov., sp. nov., an autotrophic archaeobacterium facultatively oxidizing or reducing sulfur. Syst. Appl. Microbiol. **8**: 197-203.

Zuker, M., and P. Stieger. 1982. Optimal computer folding of large RNA sequences using thermodynamics and auxiliary information. Nucleic Acids Res. **9**: 133-148.

List of Publications

1. **Sako, Y., N. Nomura, A. Uchida, Y. Ishida, H. Morii, Y. Koga, T. Hoaki, and T. Maruyama.** 1996. *Aeropyrum pernix* gen. nov., sp. nov., a novel aerobic hyperthermophilic archaeon growing at temperatures up to 100°C. *Int. J. Syst. Bacteriol.* **46**: 1070-1077.
2. **Nomura, N., and Y. Sako.** 1997. An old-fashioned newcomer: The marine hyperthermophilic archaeon *Aeropyrum pernix*. *Bioscience and Industry* **55**: 631-632 (in Japanese)
3. **Nomura, N., Y. Sako, and A. Uchida.** Molecular characterization and post-splicing fate of the three introns within the single rRNA operon of the hyperthermophilic archaeon *Aeropyrum pernix* K1. Submitted to *J. Bacteriol.*
4. **Sako, Y., and N. Nomura.** 1998. Genus. *Aeropyrum* Sako, Nomura, Uchida, Ishida, Morii, Koga, Hoaki and Maruyama 1996, 1075^{VP}. In *Bergey's manual of systematic bacteriology* (2 nd Ed.), vol.1. In press.
5. **Nomura, N., Y. Sako, and A. Uchida.** Intraspecies polymorphism in rRNA Operon of the Hyperthermophilic Archaeon *Aeropyrum pernix*, Implying the Presence of Hotspots for Intron-insertion. Submitted to *J. Bacteriol.*
6. **Nomura, N., Y. Sako, and A. Uchida.** Functional analyses of the latent intron-encoded protein I- *Ape* I: Potential role in the horizontal transfer of the archaeal introns. Submitted to *J. Bacteriol.*
7. **Morii, H., H. Yagi, H. Akutsu, N. Nomura, Y. Sako, and Y. Koga.** The sesterterpanyl ether-linked polar lipids of the aerobic hyperthermophilic archaeon *Aeropyrum pernix*. In preparation.
8. **Van Wagoner, R., E. Bruenger, N. Nomura, Y. Sako, and J. A. McCloskey.** Posttranscriptional modifications in transfer RNA from the hyperthermophilic archaeon *Aeropyrum pernix*. In preparation.