

Did Archaeal and Bacterial Cells Arise Independently from Noncellular Precursors? A Hypothesis Stating That the Advent of Membrane Phospholipid with Enantiomeric Glycerophosphate Backbones Caused the Separation of the Two Lines of Descent

Y. Koga,¹ T. Kyuragi,² M. Nishihara,¹ N. Sone²

¹ Department of Chemistry, University of Occupational and Environmental Health, Yahatanishi-Ku, Kitakyushu 807, Japan

² Department of Biochemical Engineering and Science, Faculty of Computer Science and System Engineering, Kyushu Institute of Technology, Iizuka, Fukuoka 820, Japan

Received: 24 March 1997 / Accepted: 21 May 1997

Abstract. One of the most remarkable biochemical differences between the members of two domains Archaea and Bacteria is the stereochemistry of the glycerophosphate backbone of phospholipids, which are exclusively opposite. The enzyme responsible to the formation of Archaea-specific glycerophosphate was found to be NAD(P)-linked *sn*-glycerol-1-phosphate (G-1-P) dehydrogenase and it was first purified from *Methanobacterium thermoautotrophicum* cells and its gene was cloned. This structure gene named *egsA* (enantiomeric glycerophosphate synthase) consisted of 1,041 bp and coded the enzyme with 347 amino acid residues. The amino acid sequence deduced from the base sequence of the cloned gene (*egsA*) did not share any sequence similarity except for NAD-binding region with that of NAD(P)-linked *sn*-glycerol-3-phosphate (G-3-P) dehydrogenase of *Escherichia coli* which catalyzes the formation of G-3-P backbone of bacterial phospholipids, while the deduced protein sequence of the enzyme revealed some similarity with bacterial glycerol dehydrogenases. Because G-1-P dehydrogenase and G-3-P dehydrogenase would originate from different ancestor enzymes and it would be almost impossible to interchange stereospecificity of the enzymes, it seems likely that the stereostructure of membrane phospholipids of a cell must be maintained from

the time of birth of the first cell. We propose here the hypothesis that Archaea and Bacteria were differentiated by the occurrence of cells enclosed by membranes of phospholipids with G-1-P and G-3-P as a backbone, respectively.

Key words: *sn*-Glycerol-1-phosphate dehydrogenase — *sn*-Glycerol-3-phosphate dehydrogenase — Archaea — Bacteria — Ether-linked lipid — Differentiation — Cloning — Primary structure — *egsA*

Introduction

According to the recent theories (Zillig et al. 1989; Zillig 1991; Yamagishi and Oshima 1995), Archaea (archaeabacteria) and Bacteria (eubacteria) are the most fundamental groups of life. The third domain of life proposed by Woese et al. (1990), Eucarya (eukaryotes), is now considered to be of chimeric nature in the cellular components. The question of how the first two groups were differentiated from a common ancestor arises. One of the most remarkable biochemical differences between the two domains is in the structures of polar glycerolipids. Polar lipids from Archaea consist exclusively of di- and tetraethers of glycerol and isoprenoid alcohols. On the other hand, fatty acids are linked by ester linkages in glycerolipids of Bacteria. Ether-linked lipids are, however, found in some kinds of Bacteria—for example, plasmalogens (1-alkyl-1'-enyl-2-acyl-*sn*-glycerol lipids) and 1-alkyl-2-acyl-*sn*-glycerol lipids in anaerobic bacte-

ria (Goldfine and Hagen 1972). Moreover, diether lipids have been found in some bacteria (*Thermodesulfobacterium commune* [Langworthy et al. 1983] and *Aquifex pyrophilus* [Huber et al. 1992]). Note that ether bonds are at the *sn*-1 position or *sn*-1 and -2 positions of glycerol moiety in these bacterial lipids. That is, ether lipids are not confined to Archaea but are present also in Bacteria, while the stereostructure of even these ether lipids is shared with all other bacterial glycerolipids. Isoprenoid chains are not restricted to Archaea but occurred ubiquitously in all organisms. Consequently among these characteristic structural differences between archaeal lipids and bacterial lipids, stereochemistry of the glycerophosphate backbone of phospholipids is the most crucial difference. The stereostructure of glycerophosphate backbone of archaeal phospholipids is *sn*-glycerol-1-phosphate (G-1-P) and its counterpart of bacterial lipids is enantiomeric *sn*-glycerol-3-phosphate (G-3-P) without exception so far.

The enzyme responsible to G-3-P formation (biosynthetic G-3-P dehydrogenase, EC.1.1.1.94) for polar lipid biosynthesis in Bacteria has been found by Kito and Pizer (1969) in *Escherichia coli*, purified (Edgar and Bell 1978a), characterized (Edgar and Bell 1978b; Edgar and Bell 1980), and sequenced (Sofia et al. 1994). On the other hand, the enzyme for archaeal G-1-P formation was found very recently in the methanogenic archaeon *Methanobacterium thermoautotrophicum* as a new enzyme, NAD (P)-linked G-1-P dehydrogenase (Nishihara and Koga 1995).

The metabolic pathway and enzymes catalyzing CO₂ assimilation and conversion to dihydroxyacetonephosphate (DHAP) have been detected in an autotrophic methanogen *M. thermoautotrophicum* (Jansen et al. 1982; Nishihara and Koga 1995). On the other hand, G-1-P was reported to be the best acceptor of geranylgeranylpyrophosphate to form ether lipid with the Archaea-specific stereoconfiguration in *M. thermoautotrophicum* (Zhang et al. 1990; Chen et al. 1993; Zhang and Poulter 1993). G-1-P dehydrogenase must be the key enzyme in biosynthesis of archaeal enantiomeric phospholipids, considering the two processes (G-1-P formation from CO₂ and ether lipid formation from G-1-P).

The present communication reports cloning of the structure gene for *M. thermoautotrophicum* G-1-P dehydrogenase (designated *egsA* gene (*egs* represents enantiomeric glycerophosphate synthase) and its DNA and deduced amino acid sequences, which shows no similarity with those of bacterial G-3-P dehydrogenase serving phospholipid backbone for Bacteria. Significance of this fact is also discussed in the context of differentiation of Archaea and Bacteria.

Materials and Methods

Materials. *Methanobacterium thermoautotrophicum* strain ΔH was grown as previously described (Morii and Koga 1993). T4 DNA ligase,

DNA polymerase of *Thermus aquaticus*, restriction enzymes, and plasmid vector pUC118 were purchased from Takara Shuzou Co. (Kyoto, Japan). DIG DNA labeling mixture was obtained from Boehringer Mannheim (Germany). *Staphylococcus aureus* V8 protease was the product of Sigma Chemical Company.

Determination of Amino Acid Sequences. G-1-P dehydrogenase from *M. thermoautotrophicum* was purified by ammonium sulfate fractionation and seven steps of chromatography, which will be published elsewhere. N-terminal sequence of the purified enzyme (0.5 μg) and sequence of the peptide fragments prepared from the enzyme as described below were determined by automated Edman degradation with a gas-phase peptide sequencer (Applied Biosystems 477A protein/peptide sequencer) and the protocol given by the manufacturer. For determination of internal sequences, the purified enzyme (14 μg) was digested with 0.25 μg *S. aureus* V8 protease in 50 mM NaHCO₃ (total volume of 100 μl) for 10 h at 30°C. Purified G-1-P dehydrogenase (40 μg) was also degraded by CNBr treatment in the dark for 12 h. The peptides thus generated were separated by reverse-phase HPLC on a column (4.6 × 150 mm) of Cosmogel 5C18-AR (Nacalai Tesque, Japan) eluting with a gradient of acetonitrile at a flow rate of 1 ml/min. Solvent A was 0.05% trifluoroacetic acid in 100% water and solvent B was 0.05% trifluoroacetic acid in 100% acetonitrile. The sample was applied to the column equilibrated with 90% solvent A and 10% solvent B and peptides were eluted by increasing solvent B to 70% within 40 min.

Polymerase Chain Reaction, Cloning, and Sequencing. Genomic DNA from *M. thermoautotrophicum* was isolated as described (Saito and Miura 1963). Cloning of DNA and Southern hybridization was done as described by Sambrook et al. (1989). Hybridization was performed at 50°C in 3 M NaCl/0.3 M sodium citrate pH 7.5. Two primers targeting N-terminus of MDPRKIQ (sense primer) and a peptide of GIDPEY (antisense primer) obtained by CNBr treatment were designed as follows: sense primer: 5'ATGGA(T/C)CC(A/C/T)(A/C)GAAA(A/G)AT(A/C/T)CA-3'; antisense primer: 5'TA(C/T)TC(A/G/T)GG(A/G)TC(A/T/G)AT(A/C/G/T)CC-3. Polymerase chain reaction (PCR) using these primers (20 pmol each) with 100 ng *M. thermoautotrophicum* genomic DNA as template, repeated 30 cycles (94°C for 60 s, 48°C for 60 s, and 70°C for 60 s) produced three bands which were detected on gel electrophoresis. The resulting 900-bp PCR product was isolated from the gel using glass milk and ligated into *Sma*I cut pUC118. One of several transformants of the 900-bp PCR product sequenced really contained DNA flanking the two oligonucleotides used as primers. Using this plasmid as template 900 bp DNA was amplified by PCR. The product was labeled with digoxigenin by random primer method (Boehringer) and used for hybridization analyses as the probe. Southern hybridization analysis of the genomic DNA digested with *Eco*RI gave 4.1-kbp fragments, 5.3 kbp with *Hind* III, 6.0 kbp with *Bam*HI, and 7.5 kbp with *Sall*. The DNA fragments (about 3.5–4.5-kbp zone) were produced by *Eco*RI cut, and the recovered DNA (200 ng) from the agarose gel was ligated into *Eco*RI-treated pUC118 (80 ng). Colony hybridization was carried out with the same DIG-labeled probe. Deletion series were prepared by treating the *Bam*HI- and *Sph*I-digested plasmid with exonuclease III following the manufacturer's protocol (Takara). Sequencing was carried out by the dye (fluorescent isothiocyanate) primer method in a DNA sequencer (Shimadzu model DSQ-1000) using dideoxy NTP (Sanger et al. 1977).

Results and Discussion

Purification and Characterization of G-1-P Dehydrogenase

NAD (NADP)-linked G-1-P dehydrogenase in *M. thermoautotrophicum* was purified from the methanogen to

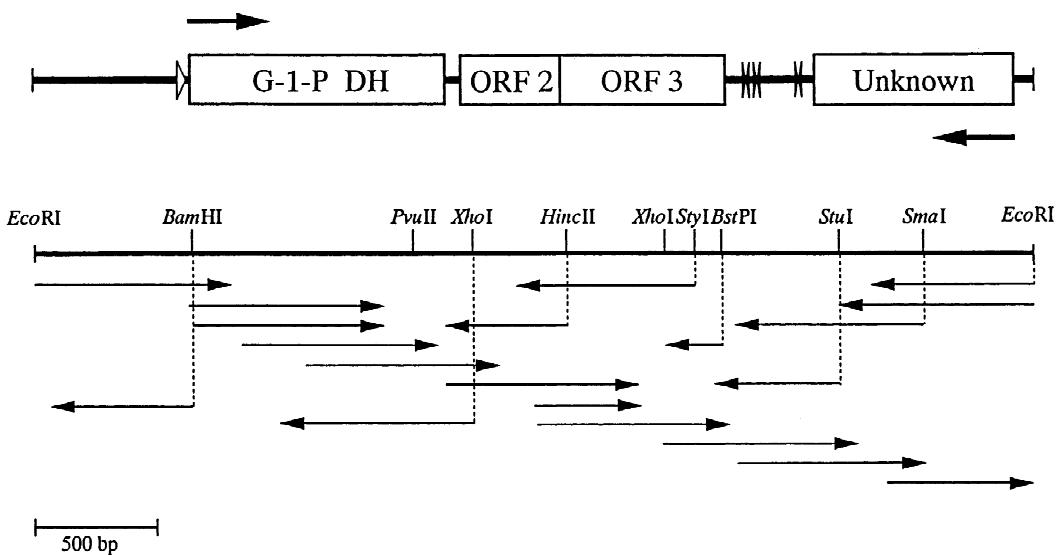


Fig. 1. A map around *M. thermoautotrophicum* *egsA* gene encoding G-1-P dehydrogenase, and sequence strategy. The *Eco*RI fragment with 4,077 bp was cloned and sequenced. Deletion series for sequencing were prepared toward right. Sequencing of the opposite side was carried out by preparing subclones using restriction sites found.

homogeneity after 7,600-fold purification by seven steps of chromatography. The molecular mass of the native form of the enzyme was 302 kDa and of the protomer shown by SDS-PAGE was approximately 38 kDa, indicating that the native enzyme consisted of an octameric form. The Michaelis constant of the enzyme for DHAP (2.17 mM in the presence of NADH) was 7.5 times smaller than that for G-1-P (16.3 mM in the presence of NAD). This fact suggests that the reduction of DHAP, that is, G-1-P formation, is the natural direction of the reaction, and this enzyme is actually responsible to the formation of G-1-P. Purification and properties of the enzyme will be published in detail elsewhere.

Cloning of the Gene Encoding G-1-P Dehydrogenase

In order to compare the amino acid sequences of archaeal G-1-P dehydrogenase to bacterial G-3-P dehydrogenase, we cloned the structure gene encoding G-1-P dehydrogenase. The purified enzyme was treated with *S. aureus* V8 protease or CNBr and partial peptide sequences were used for preparation of oligonucleotide primers for polymerase chain reaction (PCR). The N-terminal protein sequence of G-1-P dehydrogenase was determined to be MDPRKIQLPREIYTGPGVIEDTG. The amino acid sequences of very N-terminal part and one of the internal peptides separated by reverse-phase chromatography (GIDPEY) were used for preparation of primers for PCR reaction. Although three distinct PCR products (1.2, 0.9, and 0.7 kbp) were obtained, one of several transformants with pUC118 ligated with 0.9-kbp PCR product contained DNA flanking the two oligonucleotides used as primers and partial amino acid sequences determined by peptide sequencing. By use of the PCR product as a

probe a DNA fragment containing the gene for G-1-P dehydrogenase of *M. thermoautotrophicum* was obtained. A partial library was constructed in pUC118 with genomic *M. thermoautotrophicum* DNA cut with *Eco*RI and sized by agarose-gel electrophoresis. Colony hybridization was carried out with about 2,000 white colonies in eight petri dishes. Eight colonies turned purple within 30 min when detected with the DIG-DNA probe. They were propagated in 2 × TY medium, and plasmids obtained were treated with *Eco*RI. One of them with a 4-kbp insert hybridizing to the probe was chosen for sequencing. Fig. 1 shows a map around the *M. thermoautotrophicum* *egsA* gene encoding G-1-P dehydrogenase and sequence strategy.

Nucleotide Sequence of the *egsA* Gene and Derived Amino Acid Sequence

In the DNA region analyzed to obtain the *egsA* gene sequence an open reading frame encoding a 37-kDa protein was detected. (Figs. 1 and 2). The open reading frame was identified as the *egsA* gene by comparison of its deduced amino acid sequence with N-terminal amino acid sequence and with the sequences of seven internal peptides (shown by underlining in Fig. 2) of the 38-kDa subunit of G-1-P dehydrogenase. Figure 2 shows the DNA and deduced amino acid sequences of the cloned 4,077-bp *Eco*RI fragment. The open reading frame for G-1-P dehydrogenase consisted of 347 amino acid residues followed by two other open reading frames. These three open reading frames initiated with ATG and were preceded with putative Shine-Dalgarno-like sequences (Dennis 1986). A sequence GAGGTGATC was claimed to be complementary to the 3'-end of the *M. thermoau-*

Fig. 2. DNA and deduced amino acid sequences of *M. thermoautotrophicum* G-1-P dehydrogenase (*egsA*), ORF2, and ORF3. The sequences of internal peptides obtained by *S. aureus* V8 protease or CNBr treatment and determined by a peptide sequencer are shown by

underlining. Putative Shine-Dalgarno sequences are *boxed*. The nucleotides that may constitute the putative promoter are shown by *dotted underlines*. The arrows indicate nucleotide regions where putative stem-loop structure might form the terminator of the gene.

3130 3140 3150 3160 3170 3180 3190 3200 3210 3220 3230 3240
 CACCATCTTAAATGTTATTGGCTCATCCCTTTAAGGCCAACAGTCAGTCATCAATGTAATCAAACCGCTTTGTGAACTGAACAGCACCAAGAGAGCTGTGAGGACCGACTA
 3250 3260 3270 3280 3290 3300 3310 3320 3330 3340 3350 3360
 TAACTGCACCTGAAGATACGTTGAATATAATGAGATCCAGGCTGCGAGGGCGTATTATACTATAGCCGATGCCCTAACCATATTATCAGCCATATCTGAACTGAAATCCTTG
 3370 3380 3390 3400 3410 3420 3430 3440 3450 3460 3470 3480
 CAAGTGTGAGGGATTGTTAGGAGGGCATCACAGATTATTCGGCAGCTTATGAGAACACACATTATGCCAACAGGAGGTACAGTAAAGCCGTGACAGGA
 3490 3500 3510 3520 3530 3540 3550 3560 3570 3580 3590 3600
 CACCCAGAACCTCGGGCTACTCTCATCACAGATGGCAACGAACCTCCCTGAAAGGGGCACTTGATATTATGAGGTCAAGCACCAGCATGGTCCAGGGTCCGGTGT
 3610 3620 3630 3640 3650 3660 3670 3680 3690 3700 3710 3720
 CTACCATGAGTATGTTACCGAAGAGGTAGCTGAAGAGGTTGGTGCATACCCGGCTGAGGTTATGAAAGATATAACCCAGGCCATCCCTGCGGACAGAGCACCCCTATTGAGTGT
 3730 3740 3750 3760 3770 3780 3790 3800 3810 3820 3830 3840
 CCTCACTTACGGGCTCCGTGTCACTGCACCCATCGTAAGGGCTGCGAGGCCACCGTGAGGGTATGGCAGATGTGATCGGGTACCCCCGAGGAAGTAGCCGAGGCCACGCCAACAA
 3850 3860 3870 3880 3890 3900 3910 3920 3930 3940 3950 3960
 AGGCTGCATGGGATATGCCCACTTATTGACACTATCCTTGTATCACCGTACCTCCAACTGCAACGCTATACTTACAAGATGGTGTGCGACAGCCCCCTTCCATGAACT
 3970 3980 3990 4000 4010 4020 4030 4040 4050 4060 4070 4080
 GGTAAGCTGAGCAACTCCAGCAAGGCCCATCACTCCCTGGATCGTCACGTATAAGGATTGAGTATGTTGATTCAGCTGATGCTGTATCTTCATTGACTCGGATACCGG

Fig. 2. Continued.

totrophicum strain Marburg 16S rRNA (Hedderich et al. 1994; Winker and Woese 1991) and is considered to be the ribosomal-binding site (Reeve 1992). A/T-rich sequences, TATTAT and AATATATT, found at 38 and 76 bp upstream of the initiation codon of *egsA*, respectively, may be the promoter elements of methanogens, which was reported to be TTTA(A/T)ATA (Brown et al. 1989; Reeve 1992). The sequence TTTACTGTTTT located just after the stop codon of the open reading frame 3 may correspond to the transcription termination in methanogens (Reeve 1992). Two hair-pin structures were also found further downstream as shown by arrows. It appears that these genes, *egsA* for G-1-P dehydrogenase, ORF2, and ORF3, form an *egs* operon. The DNA sequence of the 3'-end contained one unidentified ORF (not shown) in the opposite strand. The DNA sequence of the 4077-bp EcoRI fragment clone containing *egsA* gene is available in the DDBJ data bank under accession number D88555. G-1-P dehydrogenase consisted of 347 amino acid residues. The molecular mass calculated from the deduced amino acid sequence (36,963) corresponded to the molecular mass of the protomer (about 38 kDa) shown by SDS-PAGE.

Comparison of the Amino Acid Sequence of G-1-P Dehydrogenase

The deduced amino acid sequence of G-1-P dehydrogenase was used for similarity search in protein databases and was compared with *E. coli* biosynthetic G-3-P dehydrogenase (encoded by *gpsA* gene). The results were summarized in Table 1. The highest identity score (52.6% identity in 272-amino-acid overlap residues) was obtained from putative glycerol dehydrogenase (open reading frame MJ0712) of *Methanococcus jannaschii* found in the complete genome sequence (Bult et al. 1996; The TIGR Microbial Database, <http://www.tigr.org/tdb/mdb/mdb.html>). The gene product of MJ0712 has been tentatively identified as glycerol dehydrogenase because it shows 23.3% amino acid identity in 236 overlap residues with *E. coli* glycerol dehydro-

nase. It is not surprising that putative glycerol dehydrogenase of *M. jannaschii* showed the highest score of similarity with G-1-P dehydrogenase of *M. thermoautotrophicum* because G-1-P dehydrogenase of *M. thermoautotrophicum* showed low but significant similarity with glycerol dehydrogenases of *E. coli* (*gldA*) (Blattner et al. 1993) and *Bacillus stearothermophilus* (Mallinder et al. 1992) (Table 1). Therefore MJ0712 should more probably be a G-1-P dehydrogenase gene. Figure 3 shows the alignment of *M. thermoautotrophicum* G-1-P dehydrogenase with MJ0712 and two glycerol dehydrogenases. Identical or similar amino acids are seen in almost the entire region. Hydropathy plot of *M. thermoautotrophicum* G-1-P dehydrogenase was also similar to that of glycerol dehydrogenases, especially to the enzyme from *B. stearothermophilus* (data not shown).

Table 1. Comparison matrices of amino acid sequence of *M. thermoautotrophicum* G-1-P dehydrogenase to amino acid sequences of some archaeal and bacterial dehydrogenases^a

Enzyme (gene) amino acid residues	1	2	3	4	5	6
1 Mth G-1-P DH (<i>egsA</i>) 347						
2 Mja putative GDH (MJ0712)	52.6					
3 Bst GDH (<i>gld</i>) 339	272	23.5	23.3			
4 Eco GDH (<i>gld</i>) 380	260	245				
5 Eco G-3-P DH (<i>gpsA</i>) 339	25.2	23.3	50.3			
6 Hin G-3-P DH (<i>gpsA</i>) 335	318	236	362			
7 Eco G-3-P DH (<i>glpD</i>) 501	125	125	107	89		
	17.5	27.3	19.7	21.0	66.6	
	103	33	61	62	335	
	13.0	18.3	20.5	11.9	16.3	20.0
	54	191	78	67	98	75

^a Upper figure, amino acid identity (%); lower figure, number of amino acid residues compared. Mth, *Methanobacterium thermoautotrophicum*; Mja, *Methanococcus jannaschii*; Bst, *Bacillus stearothermophilus*; Eco, *E. coli*; Hin, *Haemophilus influenzae*. G-1-P DH, sn-glycerol-1-phosphate dehydrogenase; G-3-P DH, sn-glycerol-3-phosphate dehydrogenase; GDH, glycerol dehydrogenase

Mth-G1PDH	1:M-D-PRKI----OLPREIYTGPVIEDTGR---CRDL-RFEGRAMVTGPRTLQIAGEA	50
Mja-G1PDH	1:-----MI.----VT..YTIIED.A.NKIEE.---LKK.-NLK-NPL.I..KN.KKYC-R-	44
Eco-GDH	1:.PHLALL.SKGAIMD.I.QSPGKY.QGADV.NRLGEY.KPLXE.WL..GDKFV.GF.QST	60
Bst-GDH	1:-----MAAE.VFISPACYVQGKNV.TKIANY.EGIGNKTV.IADEIVWK...HT	49
	*	*
Mth-G1PDH	51:--AIESLQAEGFEVDQTVDDATMASVRNVQDGGLDG-VSVVLGVGGGKVIDVAK-MSATL	106
Mja-G1PDH	45:---F-FYDI-VY-Y.EI-LNNLEI-EJKKY-TAY.---C.I.-I...RS..TG.YLAYK.	91
Eco-GDH	61:VEKSFKDAGLVV.IAPFGGECSQNEID.LRGIAETAQCGAI..I....TL.T..ALAHFM	120
Bst-GDH	50:IVNELKKGNIAA.EVVFSGEASRNEVE.IANIARKAEAAI.I....TL.T..AVADE.	109
	* ***	* *
Mth-G1PDH	107:EGLHFISVPTAASHGIASPRASI-RN-G-EGTASLEASSPIGVIADTEITISRAPFRLLA	163
Mja-G1PDH	92:--IP.....T.N.....IV..---Q--P.FMVA..AT.....KKS.R...S	143
Eco-GDH	121:GVPVA.-A..I..APC.ALSV.YTDE.EFDRLY.LPNN.NM..V..K.VAG..A....	179
Bst-GDH	110:DAYTV.-.....T.APT..ALSV.YSDD.VFESYRFYKKN.DL.LV..K..AN..P....	168
	* *** * *	* *** * *
Mth-G1PDH	164:SGCADIISNYTAIMDWKLAHRLLNERYSESAAAALSLMTAKMIKSADAIKEGLEE-SARL	222
Mja-G1PDH	144:A.MG..V..I..VL.....YEKG.K....S.IF.KTI..EL...Y-VLNSD.S.YHNK.	201
Eco-GDH	180:A.IG.ALATWFEARACSRSGATTM-A-GGKCTQAA.AL.ELCYNTLLEEG.KAMLAAEQH	237
Bst-GDH	169:..I..ALATWEARSVIKSGGKTM-A-GGIPTIAAEAI.EKCEQLFKYKLAY.SVKAK	226
	* *	*
Mth-G1PDH	223:AVKSLISSGIAISIAGSSRPASGSEHKFSHALD-M-IA-P-K--PALHGEQCGVGTTIMMM	276
Mja-G1PDH	202:..A.VG....A.N.....L.....KLKEEYNLNI-NS.....I.....S	259
Eco-GDH	238:V.TPALERV.EANTYL.GVGFESGLAAA..VHNGLT.I-PDAHHYY..KVAF..LTQL	296
Bst-GDH	227:V.TPALEAVVEANTLL.GLFESGLAAA..IHNGFT.LEGEIHHLT..KVAF..LVQL	286
	* *	**
Mth-G1PDH	277:HLH-G-GDWQF--IRD---ALARTNAPTTA-AELGIDPEYI-IEALTMNAHNRERRYTI	326
Mja-G1PDH	260:Y..-EKENKKLDSLHEKIKMS.KKVD.....-K..F.EDI.-.....K.-N.W..	315
Eco-GDH	297:V.ENAPVEEITVAALSHAVG.PITL.QLDIKEDVPAKMRIVAEA.CAEGET.HNMPGGA	356
Bst-GDH	287:A.EEHSSQQETERY.ELYLSLD.PVTLEDITKLKDASRE.ILKVAKA.TAEGET.HN-AFNV	345
	* *	* *
Mth-G1PDH	327:LGDR-GLTREAERLAKITEVI---	347
Mja-G1PDH	316:..R.---S..E.RK..EE.G.---	335
Eco-GDH	357:TP.QVYAAALLV.DOYGQRFLQ-EWE	380
Bst-GDH	346:TA.DVADAIIF..DQY..AYKEKHRK	370
	* *	*

On the other hand, search showed that virtually no similarity was observed between the sequences of G-1-P dehydrogenase and G-3-P dehydrogenase although both enzymes are NAD(P)-linked alcohol dehydrogenases with similar polypeptide lengths and the functions of both enzyme are completely the same except for the stereospecificity. A catabolic enzyme G-3-P dehydrogenase (EC.1.1.1.8, encoded by *glpD* gene) of *E. coli* showed almost no similarity to not only methanogens' G-1-P dehydrogenases but also glycerol dehydrogenases and biosynthetic G-3-P dehydrogenase of *E. coli*.

Harr plots of methanogen G-1-P dehydrogenase vs bacterial glycerol dehydrogenase or bacterial G-3-P dehydrogenase (Fig. 4) confirmed the above observation. The plot for G-1-P dehydrogenase from two methanogens gave a clear and almost straight diagonal line throughout the sequences (Fig. 4A). The plot for G-1-P dehydrogenase vs glycerol dehydrogenase showed short but significant diagonal lines (Fig. 4B) but only scattered dots were observed in the similar plot for G-1-P dehydrogenase and *E. coli* biosynthetic G-3-P dehydrogenase (Fig. 4C). Biosynthetic G-3-P dehydrogenase and glycerol dehydrogenase from the same bacterium, *E. coli*, did not any significant diagonal line in the similar plot (not shown).

M. thermoautotrophicum G-1-P dehydrogenase also showed significant similarity with several other NAD-

linked dehydrogenases, such as butanol dehydrogenase B (26.0% in 104 overlap) (Walter et al. 1992) and yeast alcohol dehydrogenase type IV (25.5% in 106 overlap) (Williamson and Paquin 1987). If the comparison is restricted to around the NAD-binding region, most of NAD-dependent dehydrogenases such as lobster glyceraldehyde-3-phosphate dehydrogenase (Davidson et al. 1967) and bovine glutamate dehydrogenase (Moon and Smith 1973) show similarity. Figure 5 shows the alignment of the sequences around the putative NAD-binding site, which is known as a common structure for enzymes of the NAD-dependent dehydrogenase superfamily, which contains the motif composed of several glycine residues and aspartic acid sited on top of a $\beta\alpha\beta$ barrel structure (Wierenga et al. 1986). In conclusion, it is clear that *M. thermoautotrophicum* G-1-P dehydrogenase has no close relationship with bacterial G-3-P dehydrogenase for ester phospholipids biosynthesis but *M. thermoautotrophicum* G-1-P dehydrogenase is a member of NAD-dependent alcohol dehydrogenase family and has a rather close relationship with glycerol dehydrogenase.

The second open reading flame was short (133 amino acid residues) and showed no clear similarity with any protein in the data bases including the complete genome sequence of *M. jannaschii*, while the third open reading flame had 44.6% identity in a 222 residue overlap with ribose-5-phosphate isomerase A (*rpiA*) from *E. coli*

Fig. 3. Alignment of G-1-P dehydrogenase (G1PDH) from *M. thermoautotrophicum* (Mth) and *M. jannaschii* (Mja) with glycerol dehydrogenase (GDH) from *E. coli* and *B. stearothermophilus*. —, gap; dot (full stop), identical amino acid residue to that of Mth-G1PDH.* shows identical amino acid for all four sequences.

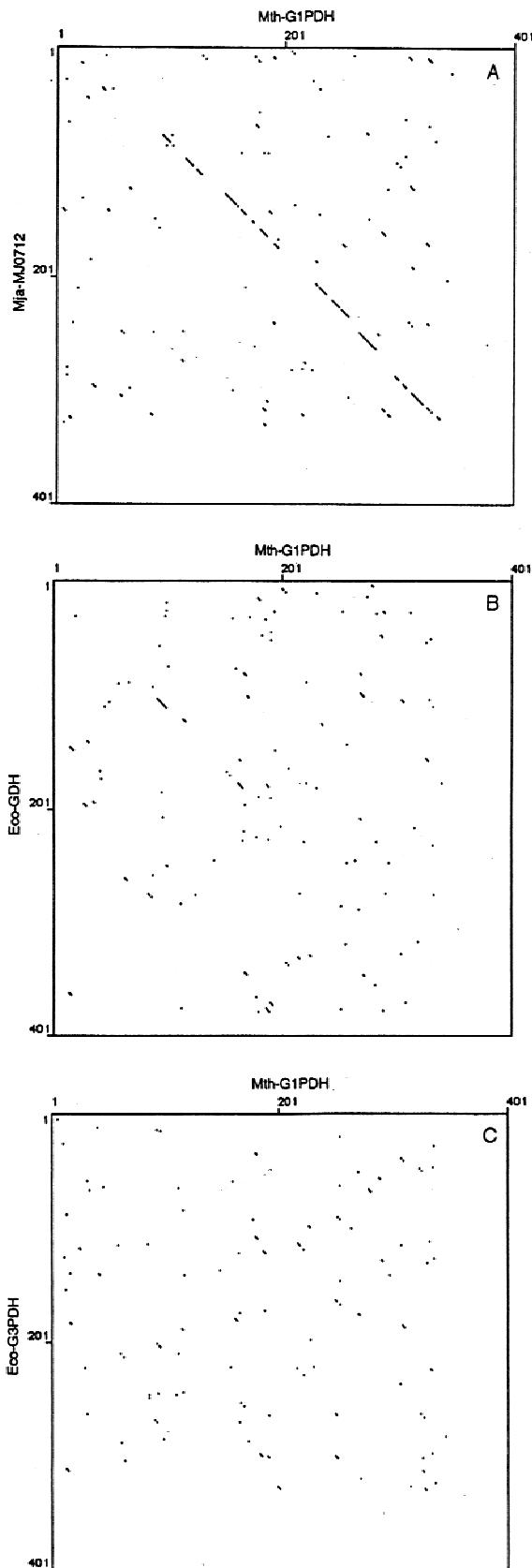


Fig. 4. Comparison of G-1-P dehydrogenase (*G1PDH*) from *M. thermoadaptrophicum* (*Mth*) and *Methanococcus jannaschii* (*Mja*), glycerol dehydrogenase (*GDH*) and G-3-P dehydrogenase (*G3PDH*) from *E. coli* (*Eco*) in the dot matrix plots. The points where three similar residues were appeared in a window of five residues are dotted.

```

1 VSVLIGVGG-GKVIDVAKMSATLEGLH--EISVPTA-ASHDG
2 AAIIVIGVGG-GKTLDTAKAVADELDAY--IVIVPTA-ASTDA
3 CGATTLGIGG-GVVGVDLAGFVAAATPMRGIDEIQMPTTLLAHD
4 VGVVLAIGG-GSAIDCAKVIIAACHEYD---GNPWD-TVLDG
5 SEIVVSIGG-GSAHDNAKAIAITLATNG-----GEIGDYEG
6 TCAVFLGGVGCLSVIMGCKAAGAA-----RIIGVDI
7 KIGIDGFGRIGRVLRAALSCGAQ-----VVAVNDP
8 TFAVQGFGNVGLHSMRYLHRGAK-----CVAVGES
9 NASMTVIGAGSYGTALAITLARNGHEV-----VVLWGHD

```

Fig. 5. Multiple alignment of *M. thermoautotrophicum* G-1-P dehydrogenase with several dehydrogenase around the NAD-binding site. The amino acid residues very similar (identical, L = I = V and E = D) to those of *M. thermoautotrophicum* G-1-P dehydrogenase are shadowed. The stars (*) on the top show the residues which are analyzed to be important for NAD-binding (Wierenga et al. 1986). 1, *M. thermoautotrophicum* G-1-P dehydrogenase; 2, *B. stearothermophilus* glycerol dehydrogenase (Mallinder et al. 1992); 3, *E. coli* glycerol dehydrogenase (Blattner et al. 1993); 4, *Clostridium acetobutylicum* butanol dehydrogenase (Walter et al. 1992); 5, yeast alcohol dehydrogenase type IV (Williamson and Paquin 1987); 6, horse liver alcohol dehydrogenase (Jörnvall 1970); 7, lobster glyceraldehyde-3-phosphate dehydrogenase (Davidson et al. 1967); 8, bovine glutamate dehydrogenase (Moon and Smith 1973); 9, biosynthetic *E. coli* G-3-P dehydrogenase (gpsA).

(Hove-Jensen and Marigaard 1993). The amino acid sequence of this open reading flame showed 54.3% similarity with the amino acid sequence deduced from the nucleotide sequence of the gene MJ1603 found in the *M. jannaschii* complete genome, which has been identified also as ribose-5-phosphate isomerase A (Bult et al. 1996, The TIGR Microbial Database). Figure 6 shows the alignment of putative ribose-5-phosphate isomerase from three sources. MJ1603 is located distantly from MJ0712 on the *M. jannaschii* chromosome and those genes do not seem to constitute an operon. It is not known at present whether the third open reading flame of *M. thermoauto-trophicum* or MJ1603 is related to G-1-P metabolism or lipid metabolism.

Hypothesis for the Differentiation of Archaea and Bacteria by the Different Enantiomers of Glycerophosphate in Their Membrane Phospholipids

To interpret the significance and the reason for the existence of the different enantiomers of glycerophosphate in the cells of Archaea and Bacteria we propose that ancestors of Archaea and Bacteria happened to differentiate by adoption of different enantiomers of glycerophosphate in their membrane phospholipids when first cells were enclosed by the membranes (Fig. 7). Because the enzymes with an opposite stereospecificity have quite different sequences, as shown above in the case of G-1-P dehydrogenase and G-3-P dehydrogenase, it seems quite unlikely that the stereospecificity of the glycerophosphate-forming enzyme of Archaea or Bacteria could be reversed. Such an example can be also seen

Mth-ORF3	1:MEVFMNLKKMAALRAVDEIDDGDVVGLGTGSTTHYFIEELGRRVREEGLEVMGVPTSYQS	60
Mja-MJ1603	1:VSN-ED..LKV.KE..KLK..M.I.....AAL..R...N.I...E.T.F.I...FEA	59
Eco-RPIA	1:.TQ-DE..AVGWA.LQYVQP.TI..V....AAH..DA..T-MKGQ-I.GAVSSSDAST	57
	*** * * * * * * * * * *	
Mth-ORF3	61:MFLAAESGIKVTSLAHD-VDVAVDGADEV-DPDLNLTKGGAAHTLEKIVDSSAASFIV	118
Mja-MJ1603	60:KM..MQYE.PLVT.D.Y...I.F.....EETT.F.....GC..Q.....YN.NE.V.	118
Eco-RPIA	58:EK..-KSL..H.FD.N.V.SLGIV.....IN-GHMQM.....L.R...IA.V.EK..C	115
	* * * * * * * * * * *	
Mth-ORF3	119:IVDESKLVERLGA-FPLPVVEIPAAACRPVKLKLESMGAVSNTRSSSEGKDGPVTDNGNFV	177
Mja-MJ1603	119:L.....KK..E.K..I.....S.Y.V.IRA.SE..GEAV..LGDR.R...I.....MI	178
Eco-RPIA	116:.A.A..Q.DI..-K.....M.RSA.ARQ.VKL---G..-P.YRQ.-.....VI	168
	* * * * * * * * * * *	
Mth-ORF3	178:LDAAFGVIDDPGAMESRLNNIPGVENGIF-AGIADMVIAGTSEGLKILR-	226
Mja-MJ1603	179:I..-V.MN...AIEL.KEI.....-TKV..K.LV..KK.V.T.KK	226
Eco-RPIA	169:..VHGME.L..I...NAI.A.....TV.L.ANRG..VALI..PD.V.TIVK	219
	* * * * * * * * * * *	

Fig. 6. Alignment of amino acid sequence of putative ribose-5-phosphate isomerase from *M. thermoaotrophicum*, *M. jannaschii*, and *E. coli*. RPIA, ribose-5-phosphate isomerase A. Other abbreviations and symbols are the same shown in Fig. 3.

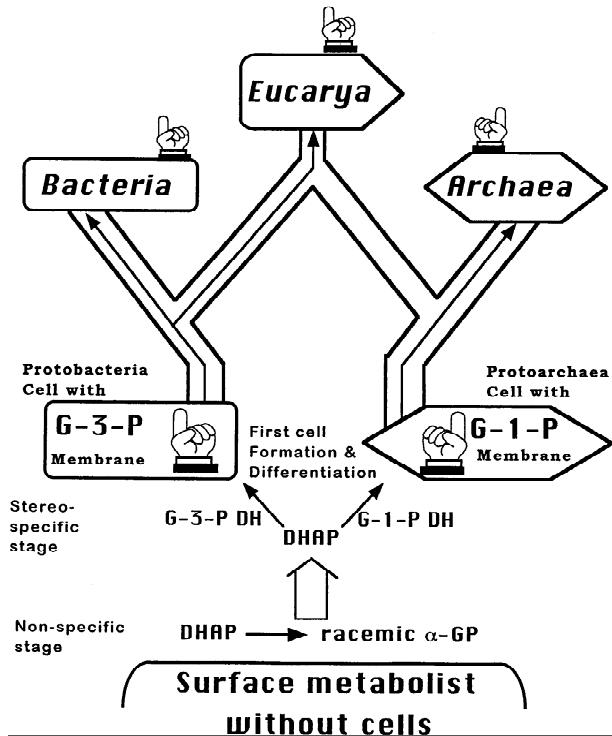


Fig. 7. A diagrammatic representation of hypothetical differentiation of Archaea and Bacteria by membrane phospholipids with glycerophosphate enantiomers as a backbone. DHAP, dihydroxyacetonephosphate; G-1-P, sn-glycerol-1-phosphate; G-3-P, sn-glycerol-3-phosphate; G-1-P DH, G-1-P dehydrogenase; G-3-P DH, G-3-P dehydrogenase; α -GP, α -glycerophosphate. One-hand symbol represents chirality of a glycerophosphate enantiomer.

in D- and L-lactate dehydrogenases, which are claimed to have different ancestral enzymes (Taguchi and Ohta 1991) because of their lack of similarity. It is likely that the enantiomer of glycerophosphate of membrane phospholipids was hardly able to be replaced by another enantiomer in one organism once the stereospecificity of the enzyme had been established in an organism. This means that the origin of the stereochemical structures of glycerophosphate in Archaea and Bacteria could be ret-

rospected to the time of the differentiation of the two groups.

Because glycerophosphate is the first stereospecific precursor of phospholipid biosynthesis in organisms of both domains (Ailhaud and Vagelos 1966; Zhang et al. 1993), the stereostructure of glycerophosphate prescribes the stereostructure of the derived lipids, and the difficulty of changeover from one stereoisomer of glycerophosphate to the other in an organism is multiplied by the number of stereospecific enzymes involved in the synthesis of polar lipids. Thus it seems likely that the established stereospecificity of the enzymes on the pathway is kept almost permanently by heredity. Apart from the case of D- and L-lactate, which are the end-products of glycolysis of lactic acid bacteria, it is noteworthy that glycerophosphate constitutes the backbone of membrane phospholipids, even though it is also a water-soluble metabolite in a cell. A cell membrane is of special significance in differentiation of cells of Archaea and Bacteria because membrane structure is absolutely essential for a cell. A cell is a small room separated from the outside environment by a partition (membrane). A cell membrane defines cell; no membrane—no cell. A cell was first born when soluble metabolites, genetic machineries, and biological catalysts, etc., were enclosed by a membrane (probably made of phospholipids). A membrane made of phospholipids with either enantiomer of glycerophosphate might separately insulate intracellular process (metabolism) and it should have established a cell, which would be an ancestor of either domain of life (Archaea and Bacteria). Substantial metabolism without cells has been suggested to be present as a form such as surface metabolism on pyrite, proposed by Wächtershäuser (1988, 1992) before cells were enclosed by lipid membranes, since most basic biochemical features were shared by Archaea and Bacteria (Bult et al. 1996) (i.e., coenzymes, central metabolic pathways, and genetic machineries). At the chemical evolution stage, both enantiomers of glycerophosphate would be synthesized as phospholipid precursors when stereospecificity of a biosynthetic catalyst would not be established. At a later time

two types of catalysts or enzymes specific for either stereoisomer became evolved, and the membrane would be composed of either enantiomer of glycerophosphate. Thus separation of Archaea and Bacteria, two domains of life, might be caused by cellularization by membranes with two enantiomeric lipids synthesized with G-1-P dehydrogenase and G-3-P dehydrogenase evolved from different enzymes, respectively. Although no ether lipid biosynthetic enzymes but digeranylgeranylglyceryl phosphate synthase is known at present, elucidation and comparison of the properties, stereospecificities, and primary structures of them will help to give evidence for or against this hypothesis.

References

- Ailhaud GP, Vagelos PR (1966) Palmityl-acyl carrier protein as acyl donor for complex lipid biosynthesis in *E. coli*. J Biol Chem 241: 3866–3869
- Blattner FR, Burland V, Plunkett III G, Sofia HJ, Daniels DL (1993) Analysis of the *Escherichia coli* genome. IV. DNA sequence of the region 89.2 to 92.8 minutes. Nucleic Acid Res 21:5408–5417
- Brown JW, Daniels CJ, Reeve JN (1989) Gene structure, organization and expression in archaeabacteria. Crit Rev Microbiol 16:287–338
- Bult CJ, White O, Olsen GJ, Zhou LX, Fleischmann RD, Sutton GG, Blake JA, Fitzgerald LM, Clayton RA, Gocayne JD, Kerlavage AR, Dougherty BA, Tomb JF, Adams MD, Reich CI, Overbeek R, Kirkness EF, Weinstock KG, Merrick JM, Glodek A, Scott JL, Geoghegan NSM, Weidman JF, Fuhrmann JL, Nguyen D, Utterback TR, Kelley JM, Peterson JD, Sadow PW, Hanna MC, Cotton MD, Roberts KM, Hurst MA, Kaine BP, Borodovsky M, Klenk HP, Fraser CM, Smith HO, Woese CR, Venter JC (1996) Complete genome sequence of the methanogenic archaeon, *Methanococcus jannaschii*. Science 273:1058–1073
- Chen AJ, Zhang DL, Poulter CD (1993) (*S*)-Geranylgeranylglyceryl phosphate synthase—purification and characterization of the first pathway-specific enzyme in archaeabacterial membrane lipid biosynthesis. J Biol Chem 268:21701–21705
- Davidson BE, Sajgo M, Noller HF, Harris JI (1967) Amino-acid sequence of glyceraldehyde 3-phosphate dehydrogenase from lobster muscle. Nature 216:1181–1185
- Dennis PP (1986) Molecular biology of archaeabacteria. J Bacteriol 168:471–478
- Edgar JR, Bell RM (1978a) Biosynthesis in *Escherichia coli* of *sn*-glycerol 3-phosphate, a precursor of phospholipid. Purification and physical characterization of wild type and feedback-resistant forms of the biosynthetic *sn*-glycerol-3-phosphate dehydrogenase. J Biol Chem 253:6348–6353
- Edgar JR, Bell RM (1978b) Biosynthesis in *Escherichia coli* of *sn*-glycerol 3-phosphate, a precursor of phospholipid. Kinetic characterization of wild type and feedback-resistant forms of the biosynthetic *sn*-glycerol-3-phosphate dehydrogenase. J Biol Chem 253: 6354–6363
- Edgar JR, Bell RM (1980) Biosynthesis in *Escherichia coli* of *sn*-glycerol 3-phosphate, a precursor of phospholipid. Further kinetic characterization of wild type and feedback-resistant forms of the biosynthetic *sn*-glycerol-3-phosphate dehydrogenase. J Biol Chem 255:3492–3497
- Goldfine H, Hagen P-O (1972) Bacterial plasmalogens. In: Snyder F (ed) Ether lipids. Chemistry and biology. Academic Press, New York, pp 329–350
- Hedderich R, Koch J, Linder D, Thauer RK (1994) The heterodisulfide reductase from *Methanobacterium thermoautotrophicum* contains sequence motifs characteristic of pyridine-nucleotide-dependent thioredoxin reductases. Eur J Biochem 225:253–261
- Hove-Jensen B, Marigaard M (1993) *Escherichia coli rpiA* gene encoding ribose phosphate isomerase A. J Bacteriol 175:5628–5635
- Huber R, Wilharm T, Huber T, Trincone A, Burggraf S, König H, Rachel R, Rockinger I, Fricke H, Stetter KO (1992) *Aquifex pyrophilus* gen nov sp nov represents a novel group of marine hyperthermophilic hydrogen-oxidizing bacteria. Syst Appl Microbiol 15: 340–351
- Jansen K, Stupperich E, Fuchs G (1982) Carbohydrate synthesis from acetyl Co A in the autotroph *Methanobacterium thermoautotrophicum*. Arch Microbiol 132:355–364
- Jörnvall H (1970) Horse liver alcohol dehydrogenase. On the primary structures of the isoenzymes. Eur J Biochem 16:41–49
- Kito M, Pizer LI (1969) Purification and regulatory properties of the biosynthetic L-glycerol-3-phosphate dehydrogenase from *Escherichia coli*. J Biol Chem 244:3316–3323
- Langworthy TA, Holzer G, Zeikus JG, Tornabene TG (1983) Iso- and anteiso-branched glycerol diethers of the thermophilic anaerobe *Thermodesulfobacterium commune*. Syst Appl Microbiol 4:1–17
- Mallinder PR, Pritchard A, Moir A (1992) Cloning and characterization of a gene from *Bacillus stearothermophilus* var. non-diastaticus encoding a glycerol dehydrogenase. Gene 110:9–16
- Moon K, Smith EL (1973) Sequence of bovine liver glutamate dehydrogenase VIII. Peptides produced by specific chemical cleavages; the complete sequence of the protein. J Biol Chem 248:3082–3088
- Morii H, Koga Y (1993) Tetraether type polar lipids increase after logarithmic growth phase of *Methanobacterium thermoautotrophicum* in compensation for the decrease of diether lipids. FEMS Microbiol Lett 109:283–288
- Nishihara M, Koga Y (1995) *sn*-glycerol-1-phosphate dehydrogenase in *Methanobacterium thermoautotrophicum*: Key enzyme in biosynthesis of the enantiomeric glycerophosphate backbone of ether phospholipids of archaeabacteria. J Biochem 117:933–935
- Reeve JN (1992) Molecular biology of methanogens. Annu Rev Microbiol 46:165–191
- Saito H, Miura K (1963) Preparation of transforming deoxyribonucleic acid by phenol treatment. Biochim Biophys Acta 72:619–629
- Sambrook J, Fritsch EF, Maniatis T (1989) Molecular cloning: a laboratory manual. 2nd ed. Cold Spring Harbor Laboratory Press, New York
- Sanger F, Nicklen S, Coulson AR (1977) DNA sequencing with chain-terminating inhibitors. Proc Natl Acad Sci USA 74:5463–5467
- Sofia HJ, Burland V, Daniels DL, Plunkett III G, Blattner FR (1994) Analysis of the *Escherichia coli* genome. V. DNA sequence of the region from 76.0 to 81.5 minutes. Nucleic Acid Res 22:2576–2586
- Taguchi H, Ohta T (1991) D-Lactate dehydrogenase is a member of the D-isomer-specific 2-hydroxyacid dehydrogenase family. Cloning, sequencing, and expression in *Escherichia coli* of the D-lactate dehydrogenase gene of *Lactobacillus plantarum*. J Biol Chem 266: 12588–12594
- Wächtershäuser G (1988) Before enzymes and templates: theory of surface metabolism. Microbiol Rev 52:452–484
- Wächtershäuser G (1992) Groundworks for an evolutionary biochemistry. The iron-sulphur world. Prog Biophys Mol Biol 58:85–201
- Walter KA, Bennett GN, Papoutsakis ET (1992) Molecular characterization of two *Clostridium acetobutylicum* ATCC824 butanol dehydrogenase isozyme genes. J Bacteriol 174:7149–7158
- Wierenga RK, Terpstra P, Hol WGJ (1986) Prediction of the occurrence of the ADP-binding $\beta\alpha\beta$ -fold in proteins, using an amino acid sequence fingerprint. J Mol Biol 187:101–107
- Williamson VM, Paquin CE (1987) Homology of *Saccharomyces cerevisiae* ADH4 to an iron-activated alcohol dehydrogenase from *Zymomonas mobilis*. Mol Gen Genet 209:374–381
- Winker S, Woese CR (1991) A definition of the domains *Archaea*,

- Bacteria and Eucarya* in terms of small subunit ribosomal RNA characteristics. *Syst Appl Microbiol* 14:305–310
- Woese CR, Kandler O, Wheelis ML (1990) Towards a natural system of organisms; proposal for the domains Archaea, Bacteria and Eucarya. *Proc Natl Acad Sci USA* 87:4576–4579
- Yamagishi A, Oshima T (1995). Return to dichotomy: Bacteria and Archaea. In: Chera-Fores J, Chadha M, Negron-Mendoza A, Oshima T (eds) Chemical evolution: self-organization of the macromolecules of life. A. Deepak, Hampton, VA, pp 155–158
- Zhang D, Poulter CD (1993) Biosynthesis of archaeabacterial ether lipids. Formation of ether lipkages by prenyltransferases. *J Am Chem Soc* 115:1270–1277
- Zhang D-L, Daniels L, Poulter CD (1990) Biosynthesis of archaeabacterial membranes. Formation of isoprene ethers by a prenyl transfer reaction. *J Am Chem Soc* 112:1264–1265
- Zillig W (1991) Comparative biochemistry of Archaea and Bacteria. *Curr Opin Genet Dev* 1:544–551
- Zillig W, Klenk H-P, Palm P, Leffers H, Pühler G, Gropp F, Garrett RA (1989) Did eukaryotes originate by a fusion event? *Endocytobiosis Cell Res* 6:1–25