J. Microbiol. Biotechnol. (1999), 9(4), 404–413

JOURNAL OF
MICROBIOLOGY
AND
BIOTECHNOLOGY

© The Korean Society for Applied Microbiology

Random Sequence Analysis of the Genomic DNA of *Methanopyrus kandleri* and Molecular Cloning of the Gene Encoding a Homologue of the Catalytic Subunit of Carbon Monoxide Dehydrogenase

SHIN, HYUN-SEOCK[1,2], JAE-RYEON RYU[1], YE-SUN HAN[1], YONG-JIN CHOI[2], AND YEON GYU YU[1]*

[1]*Structural Biology Center, Korea Institute of Science and Technology, P.O.Box 131, Cheongryang, Seoul 136-791, Korea*
[2]*Graduate School of Biotechnology, Korea University, Seoul 136-701, Korea*

Received: February 20, 1999

Abstract *Methanopyrus kandleri* is a hyperthermophilic methanogen that represents one of the most heat-resistant organisms: the maximum growth temperature of *M. kandleri* is 110°C. A random sequence analysis of the genomic DNA of *M. kandleri* has been performed to obtain genomic information. More than 200 unique sequence tags were obtained and compared with the sequences in the GenBank and PIR databases. About 30% of the analyzed tags showed strong sequence similarity to previously identified genes involved in various cellular processes such as biosynthesis, transport, methanogenesis, or metabolism. When statistics relating to the frequency of codons were examined, the sequenced open reading frames showed highly biased codon usage and a high content of charged amino acids. Among the identified genes, a homologue of the catalytic subunit of carbon monoxide dehydrogenase (CODH) that reduces CO_2 to CO was cloned and sequenced in order to examine its detailed gene structure. The cloned gene includes consensus promoters. The amino acid sequence of the cloned gene shows a strong homology with the CODH genes from methanogenic Archaea, especially in the presumed binding sites for Fe-S centers.

Key words: Carbon monoxide dehydrogenase (CODH), Hyperthermophile, *Methanopyrus kandleri*

Hyperthermophiles are microorganisms that are able to grow at 80°C or higher temperatures. Proteins from hyperthermophiles show strong resistance to heat, and this characteristic has evoked interest for their potential application in biotechnology. Most of hyperthermophiles belong to the Archaea, and are classified as methanogens, chemolithotrophs, or chemoorganotrophs depending on the nature of their energy metabolism [2]. *M. kandleri*, a methanogenic Archaeon that was first found at an abyssal hot vent near the Gulf of California [13, 19], is one of the most heat resistant organisms ever known: its optimum and highest growth temperatures are recorded to be 98°C and 110°C, respectively. An analysis of the sequence of 16S rRNA locates this organism in a separate lineage origination near the root of the phylogenetic tree [5, 32]. The genomic structure of *M. kandleri* has not been systematically studied, although a few genes have been cloned and their gene products expressed in *Escherichia coli* [16, 18, 28].

M. kandleri is an obligate chemolithoautotroph, and it is assumed to assimilate CO_2 into acetyl-CoA using the modified Ljungdahl-Wood pathway like other autotrophic methanogens, such as *Methanococcus jannaschii* [4]. In this pathway, the methyl carbon of acetyl-CoA is derived from CO_2 via reduction in the tetrahydrofolate pathway and the carbonyl group is obtained by the reduction of CO_2 catalyzed by carbon monoxide dehydrogenase (CODH) [24]. The same enzyme dissimilates acetate into CO and the methyl group in acetogenic methanogens, such as *Methanosarcina* species [10]. Hence, CODH is involved in either the synthesis or the cleavage of acetyl-CoA depending on the organism.

Most CODH enzymes contain Ni ion as a metal cofactor except for the enzyme from carboxydotrophic microorganisms that uses molybdenum as a cofactor [10]. The CODH from *Clostridium thermoaceticum* has been well characterized. It is a dimer ($\alpha_2\beta_2$) of a heterodimer ($\alpha\beta$) [33], in which the α subunit is a 82-kDa protein composed of 729 amino acids, and the β subunit is a 73 kDa protein composed of 674 amino acids [21]. The α subunit contains the A center where the acetyl-CoA synthesis occurs and the β subunit contains two metal binding centers where the CO oxidation takes place [33]. Since the β subunit is responsible for the oxidation or reduction of CO or CO_2, respectively, this subunit is called

*Corresponding author
Phone: 82-2-958-5936; Fax: 82-2-958-5939;
E-mail: ygy@kistmail.kist.re.kr

the catalytic subunit. The organization of metal clusters and their involvement in the assimilation of CO_2 in the catalytic subunit of CODH is of current interest.

To obtain genome information on the organisms that are difficult to culture and be analyzed by conventional genetic methods, the sequence determination of the randomly selected fragments of genomic DNA and comparison to known DNA sequences was successfully performed on *Thermotoga maritima* [15] and *Aquifex pyrophilus* [6]. In this study, a plasmid library that contained the genomic DNA of *M. kandleri* was sequenced, and the sequence information was then analyzed. More than 70 of the *M. kandleri* sequences showed significant sequence similarity with previously identified genes including the catalytic subunit of CODH. The whole gene of a homologue of the catalytic subunit of CODH was cloned and its structure examined.

MATERIALS AND METHODS

Bacterial Strains and Growth Conditions

Methanopyrus kandleri cells (DSM #6324) were obtained from Deutsch Sammlung von Mikroorganismen (DSM, Braunschweig, Germany). The cultivation of *M. kandleri* cells was performed according to the method developed by Kurr *et al.* [19]. A BSM medium was prepared under anaerobic conditions in a H_2:CO_2 gas mixture. The *M. kandleri* cells were inoculated (5% inoculum) in a 125-ml serum bottle (Wheaton, Scientific Co., Milliville, NJ, U.S.A.) containing 40 ml of the BSM medium in 3 atm of a H_2:CO_2 gas mixture (80:20, v/v). After 24 h of cultivation at 98°C while shaking at 200 rpm, the cells were harvested by centrifugation at 12,000 ×g for 20 min at 4°C and stored at -80°C. The *E. coli* host strain for plasmid DNA amplification was DH5α. The *E. coli* cells were grown at 37°C while shaking at 250 rpm in a Luria-Bertani (LB) medium [26] supplemented with appropriate antibiotics.

Enzymes and Reagents

All restriction and DNA modification enzymes used for this study were of molecular biology grade. The sequencing reaction was performed using the Sequenase version 2.0 reaction kit (USB, Cleveland, U.S.A.), Sequitherm™ Long-Read™ Cycle sequencing kit (Epicentre, Madison, U.S.A.), and *Taq* Dye Deoxy Terminator Cycle sequencing kit (PE Applied Biosystems, Warrington, U.K.).

Isolation of Genomic DNA from *Methanopyrus kandleri*

The genomic DNA of *M. kandleri* was isolated using the method developed by Shima *et al.* [28] with slight modification. About 0.4 g of *M. kandleri* cell paste was ground with a pestle in a cold mortar under liquid nitrogen, and the powder was mixed with 10 ml of a TES buffer [20 mM Tris-HCl (pH 8.0), 10 mM EDTA (pH 8.0), 0.25 M Sucrose] containing proteinase K (0.5 mg/ml). Sodium dodecyl sulfate (SDS) was added to the suspension at a final concentration of 1%, and incubated at 60°C overnight. After incubation, a 0.1 volume of 5 M NaCl was added to the mixture, and the insoluble materials in the sample were removed by centrifugation at 18,000×g for 30 min at 4°C. An equal volume of isopropanol was added to the supernatant, and the precipitated DNA was recovered by centrifugation at 14,000 ×g for 10 min at 4°C. The DNA was dissolved in 1 ml of a TE buffer (10 mM Tris-HCl, 1 mM EDTA, pH 8.0), extracted with phenol (a 50:50 phenol-chloroform mixture) and finally with chloroform.

Construction of Plasmid and Lambda Libraries of *M. kandleri*

For the construction of a plasmid library, the genomic DNA of *M. kandleri* was digested with *Kpn*I to generate DNA fragments ranging in size from 0.1 to 20 kb. The plasmid vector pBluescript II KS (+) was digested with *Kpn*I and dephosphorylated with calf intestinal alkaline phosphatase (CIP) (New England Biolab, U.K.). The DNA fragments of *M. kandleri* were then ligated into the pBluescript II KS (+) by T4 DNA ligase (Promega, Madison, U.S.A.), and the ligated DNAs were transformed into the *E. coli* strain DH5α by electroporation using an electroporator (Invitrogen electroporator II, Netherland). The transformed cells were plated on LB agar supplemented with ampicillin (100 μg/ml), isopropylthio-β-D-galactoside (IPTG), and 5-bromo-4-chloro-3-indolyl-β-D-galactoside (X-Gal) as described by Sambrook *et al.* [26]. The plasmid DNAs from white colonies were isolated and used for further study.

To generate a lambda library, the genomic DNA (2 μg) of *M. kandleri* was partially digested with *Bam*HI until the average size of the DNA fragments were within the range of 10-20 kb. The DNA fragments were ligated into 1 μg of lambda Dash II (Qiagen, Germany) digested with *Bam*HI. The ligated DNAs were mixed with a Gigapack II XL packaging extract (Stratagene, U.S.A.) as described in the instruction manual. *E. coli* strain [XL1-Blue MRA (P2)] was transfected with the recombinant bacteriophages and plated on an LB agar plate. The phages from the developed plaques were collected and used for further screening.

DNA Sequencing and Sequence Analysis

Sequence determination was performed using the dideoxy chain termination method [24]. The sequencing reaction products were separated on a 6% acrylamide gel and the sequences were read using either an ABI 373A automatic sequencer (Applied Biosystems, U.S.A.) or an LI-COR automatic sequencer (LI-COR, Inc., U.S.A.). The nucleotide

sequences obtained were compared with the known proteins in the database using the sequence comparison program, advanced BLAST (Basic Local Alignment Search Tool) [1] program. The CLUSTAL W program [30] was used for the conserved regions in the protein sequences between similar proteins.

Preparation of RNA from *M. kandleri*

The preparation of the total RNA of *M. kandleri* was performed using Weil's method [31] with the following modification. About 0.4 g of cells (wet weight) was ground in a cold mortar under liquid nitrogen, and then the powder of the cell lysate was resuspended in 10 ml of a TES buffer containing 5 mg of proteinase K and 1% SDS. After incubation at 60°C for 2 h, a 0.1 volume of 5 M NaCl was added to the mixture, and any particulate materials in the reaction mixture were removed by centrifugation at 18,000×g for 30 min at 4°C. Nucleic acids were precipitated by the addition of an equal volume of isopropanol, and the resulting pellet was dissolved in a TE buffer. The nucleic acids were treated with RNase-free DNase (Promega, U.S.A.) and further purified by phenol extraction [7, 22]. After ethanol precipitation, the RNA pellet was dissolved in diethyl pyrocarbonate (DEPC)-treated water and stored at - 80°C in a deep-freezer until use.

Primer Extension

Primer extension was performed according to Sambrook *et al.* [26]. About 20 pmol of oligonucleotide was labeled with [γ-^{32}P] ATP (5,000 Ci/mmol, Amersham, U.K.), and the labeled oligonucleotide was hybridized with 10 µg of the total RNA of *M. kandleri* in a hybridization buffer (100 mM $MgCl_2$, 500 mM NaCl) overnight at 34°C. The nucleic acid was recovered by ethanol precipitation and a reverse transcription reaction was performed in 50 mM Tris-HCl (pH 7.6), 60 mM KCl, 1 mM dithiothreitol (DTT), and 1 unit/µl of a placental RNase inhibitor using 27 units of an avian myeloblastosis virus (AMV) reverse transcriptase (Promega, U.S.A.). The reaction product was separated on a 6% (w/v) polyacrylamide gel along with the sequencing reaction product using the same oligomer as the sequencing primer.

Cloning of a Homologue of the Catalytic Subunit of CODH from *M. kandleri*

The general procedures for λ library screening and DNA cloning were performed as described by Sambrook *et al.* [26]. The λ library was screened using the sequence tag lmk-ef as a probe. The plaques that hybridized with the probe were amplified and their DNA was isolated. One of the phage DNA had a 7 kb-*Eco*RI fragment that hybridized with the probe. The 7 kb-*Eco*RI fragment was cloned into pBluescript II KS (+) (Stratagene, Germany) to produce pMkCD1. The 7-kb *Eco*RI fragment was further digested with *Hin*dIII, and a 2.8 kb-*Hin*dIII fragment that contained the whole coding region of the catalytic subunit of the CODH was further cloned into pBluescript II KS (+) to produce pHIN. The DNA sequence of the 2.8 kb-*Hin*dIII/*Eco*RI fragment was determined by the dideoxy chain termination method [24] using T3, T7, and synthetic primers (synthesized by Bio-synthesis, Co., U.S.A.).

RESULTS AND DISCUSSION

Cell Culture and Library Construction

The growth of the *M. kandleri* cell was confirmed by detecting the methane gas in the culture bottle after 12 h of culture. The cylindrical shape (0.5 µm diameter and 1.5-2 µm of length) of the cultured cell observed by electron microscopy (data not shown) was consistent with the previously observed shape of *M. kandleri* [13, 19]. About 0.4 g of cells (wet weight) was obtained from 1 l of culture and approximately 200 µg of genomic DNA was isolated from 0.4 g of cell paste.

After the construction of a plasmid library, the plasmid DNAs from 290 white colonies were purified. When the *Bam*HI fragments of the genomic DNA of *M. kandleri* were ligated into λ DASH II DNA, about 3,000 plaques were generated. The average size of the insert DNA in the plaques was about 12-20 kb. The phages from the 3,000 plaques were pooled and amplified. The amplified lambda library was used for the screening of the CODH gene.

Sequence Determination and Analysis of the Plasmid Library of *M. kandleri*

To obtain the genetic information of *M. kandleri*, the constructed plasmid libraries were sequenced. From the 290 plasmid DNAs in the plasmid libraries, 202 plasmid DNAs were successfully sequenced. The failure to sequence was probably due to an inadequate purity of the DNA or a failure of the sequencing reaction procedure. The 202 sequence tags obtained were analyzed using the advanced BLAST program. Amino acid sequences of six possible reading frames deduced from the obtained sequence tags were compared with protein sequences in the database. Among the 202 tags, 68 clones were selected with sequences longer than 300 bp and a matched score of over 80 from a BLAST search of previously identified genes. There were 51 clones that matched genes with assigned functions. The remaining 17 clones were matched with genes with hypothetical functions.

Genes from *M. kandleri* and Codon Usage

The lists of the 68 sequence tags are shown in Table 1 along with genes exhibiting a strong similarity. In general, the sequence tags from *M. kandleri* showed strong similarity to the genes from methanogens or other Archaea

Table 1. Lists of identified genes from the *M. kandleri* genomic library.

Pir No.	Gene description	Score	Clone # (High)
	Biosynthesis of amino acid, cofactors, prosthetic group, and carriers		
pir||C64459	dihydroxy-acid dehydrogenase [*Methanococcus jannaschii*]	429	smk-153
pir||D64413	cobalamin biosynthesis protein N [*Mc. jannaschii*]	332	smk-116
pir||JC4352	aspartate-tRNA ligase [*Pyrococcus* sp.]	277	pmk-135
sp|P06195|PABA_SERMA	*p*-aminobenzoate synthase glutamine amidotransferase component II	214	smk-118
sp|P21157|COBQ_METVO	cobyric acid synthase [*Mc. voltae*]	198	smk-128
gnl|PID|e299306	TDP-D-Glucose synthase [*Streptomyces argillaceus*]	146	smk-114
pir||C64375	glutamate-1-semialdehyde 2,1-aminomutase [*Mc. jannaschii*]	130	pmk-21
gnl|PID|d1011268	cobyric acid synthase [*Synechocystis* sp.]	112	pmk-161
gn1|PID|d1030554	173aa long hypothetical molybdenum cofactor biosynthesis protein [*Pyrococcus horikoshii*]	106	pmk-97, 101
	Cell envelope and cellular processes related protein		
pir||E64318	centromere/microtubule-binding protein [*Mc. jannaschii*]	241	pmk-70
gi|2258418	dolichol monophosphate mannose synthase [*Homo sapiens*]	174	pmk-123
	Central intermediary and energy metabolism		
pir||H64390	carbon-monoxide dehydrogenase beta subunit [*Mc. jannaschii*]	124	lmk-ef
gi|2622619	succinate dehydrogenase, flavoprotein subunit [*Mb. thermoautotrophicum*]	490	pmk-68
pir||A64304	succinate dehydrogenase, flavoprotein [*Mc. jannaschii*]	374	pmk-106, 18
pir||B64361	L-lactate dehydrogenase [*Mc. jannaschii*]	306	pmk-46
sp|P53191|SUCD_COXBU	succinyl-CoA Synthetase alpha chain [*Coxiella burnetii*]	195	pmk-25
gi|2982926	nitrate reductase (NAD(P)H) large subunit [*Aquifex aeolicus*]	126	pmk-47
pir||F64327	H^+-transporting ATP synthase subunit K [*Mc. jannaschii*]	120	smk-102
sp|P08640|AMYH_YEAST	glucoamylase S1/S2 precursor	102	smk-53
pbd|1FCA|	ferredoxin	83	pmk-102
	Methanogenesis		
pir||H64406	*N*5-methyl-H4MPT-coenzyme M methyltransferase homolog [*Mc. jannaschii*]	255	pmk-118
gi|1354836	methyl coenzyme M reductase beta-chain [*Methanopyrus kandleri*]	212	pmk-30
pir||A64455	methyl coenzyme M reductase system, component A2 [*Mc. jannaschii*]	190	pmk-16
pir||S57465	formylmethanofuran DHase subunit A [*Mb. thermoautotrophicum*]	150	pmk-76
pir||H64405	*N*5-methyl-H4MPT-coenzyme M methyltransferase homolog [*Mc. jannaschii*]	149	pmk-59
pir||S57453	formylmethanofuran DHase (tungsten) subunit F [*Mc. jannaschii*]	139	pmk-164
gnl|PID|e334007	*N*5-methyl-H4MPT-coenzyme M methyltransferase, subunit MtrD [*Methanopyrus kandleri*]	138	pmk-105, 143, smk-103
pir||B37777	methyl viologen-reducing hydrogenase [*Methanothermus fervidus*]	85	lmk-hb
	Purines, pyrimidines, nucleosides, and nucleotides		
pir||D64437	glucose-1-phosphate thymidylyltransferase [*Mc. jannaschii*]	250	pmk-117
sp|P50845|KDGK_BACSU	2-dehydro-3-deoxygluconokinase [*Bacillus subtilis*]	126	smk-101
sp|P80305|TYSY_METTH	thymidylate synthase (TS) [*Methanobacterium thermoautotrophicum*]	100	pmk-69
	Replication, transcription, and translation		
gnl|PID|d1026092	DNA polymerase II subunit 2 [*Pyrococcus furiosus*]	317	smk-170
sp|P54109|RS2_METJA	30S ribosomal protein S2P [*Mc. jannaschii*]	393	smk-112
sp|P54067|RS6E_METJA	30S ribosomal protein S6E [*Mc. jannaschii*]	368	pmk-24
gi|2621738	ATP-dependent RNA helicase related protein [*Mb. thermoautotrophicum*]	285	pmk-89
gi|2984318	initiation factor eIF-28 alpha subunit [*Aquifex aeolicus*]	282	pmk-09
gnl|PID|d1031905	967aa long hypothetical leucyl-tRNA synthetase [*Pyrococcus horikoshii*]	248	smk-133
gn|PID|d1031046	480aa long hypothetical prolyl-tRNA synthetase [*Py. horikoshii*]	221	smk-26
gi|2650146	mRNA 3'-end processing factor, putative [*A. fulgidus*]	199	pmk-25
sp|P54064|R24E_METJA	50S ribosomal protein L24E [*Mc. jannaschii*]	152	smk-24
gi|2826316	putative transcription initiation factor (TF II E, subunit alpha)	102	smk-32

Table 1. Continued.

Pir No.	Gene description	Score	Clone # (high)
	Transport and binding proteins		
gi|2649523	ATP transporter, ATP-binding protein, putative [*Archaeoglobus fulgidus*]	268	pmk-14
gi|2104726	V-ATPase A subunit [*Desulfurococcus* sp. SY]	213	pmk-104
gnl|PID|d1030679	355aa long hypothetical GTP-binding protein [*Py. horikoshii*]	207	pmk-110
gnl|PID|e1294480	putative cationic amino acid transporter [*Streptomyces coelicolor*]	133	pmk-108
	Other categories		
sp|031662|YKRS_BACSU	hypothetical 38.9-kD protein in dat-spooe intergenic region [*Bacillus subtilis*]	206	pmk-93
pir||I40767	catalase [*Campylobacter jejuni*]	158	lmk-pb2
gi|927403	F46G10.3 [*Caenorhabditis elegans*]	134	pmk-107
pir||S72844	β-aspartate methyltransferase PimT [*Mycobacterium leprae*]	115	smk-55
pir||S50120	activating factor-common tabacco [*Nicotiana tabacum*]	96	pmk-166
sp|Q09625|Y58A_CAEEL	hypothetical 84.3-kD protein ZK945.10 in chromosome II	85	pmk-77
	Unknown and hypothetical protein		
sp|Q60362|Y057_METJA	hypothetical protein MJ0057 [*Mc. jannaschii*]	226	pmk-137
sp|Q58482|YA82_METJA	hypothetical protein MJ1082 [*Mc. jannaschii*]	209	pmk-08
sp|Q57897|Y455_METJA	hypothetical protein MJ0455 [*Mc. jannaschii*]	180	pmk-156
sp|Q58952|YF57_METJA	hypothetical protein MJ1557 [*Mc. jannaschii*]	155	pmk-22
sp|Q58095|Y682_METJA	*malT-glpR* intergenic region hypothetical intein-containing protein homolog MJ0682 [*Mc. jannaschii*]	90	pmk-167
sp|Q58778|YD83_METJA	hypothetical protein MJ1383 [*Mc. jannaschii*]	89	lmk-kf
gi|2622305	conserved protein [*Mb. thermoautotrophicum*]	266	pmk-129
gi|2622180	conserved protein [*Mb. thermoautotrophicum*]	206	pmk-141
gi|2621853	conserved protein [*Mb. thermoautotrophicum*]	145	pmk-48
gi|2622692	conserved protein [*Mb. thermoautotrophicum*]	137	smk-9
gi|2621447	conserved protein [*Mb. thermoautotrophicum*]	95	pmk-60
gi|2622971	unknown [*Mb. thermoautotrophicum*]	97	pmk-28
sp|026762|Y666_METTH	hypothetical protein MTH666 [*Mb. thermoautotrophicum*]	95	pmk-19
gi|264940	conserved hypothetical protein [*A. fulgidus*]	280	smk-129
gi|2650030	conserved hypothetical protein [*A. fulgidus*]	100	smk-154
gnl|PID|e291024	*orf 2* [*Mp. kandleri*]	377	smk-93
gnl|PID|e291026	*orf 4* [*Mp. kandleri*]	582	pmk-103

such as *Methanococcus jannaschii, Methanobacterium thermoautotrophicum*, and *Archaeoglobus fulgidus*. Two sequence tags, pmk-30 and pmk-105, had the same DNA sequences as the previously identified methyl coenzyme M reductase beta subunit [23] and N_5-methyl-H_4MPT-coenzyme M methyltransferase [12], respectively. There were no mismatched bases between the two sequence tags and reported sequences. This result indicates that the accuracy of the sequencing was more than 99%. The identified genes were grouped according to their cellular function such as the biosynthesis of amino acids, cell envelope formation, energy metabolism, nucleotide synthesis, replication/transcription/translation, transport, other categories, and hypothetical genes. To examine the detailed structure of the *M. kandleri* genes, the complete gene encoding the catalytic subunit of carbon monoxide dehydrogenase (CODH) was cloned and its DNA sequence was determined.

The number of codons for each amino acid was determined to examine the codon frequency in *M. kandleri.* All the codons from the most probable open reading frames of the 68 sequence tags were collected, and the codon percentages for each amino acid are listed in Table 2 along with those for *M. jannaschii* and *E. coli.* The most distinct characteristic of codon usage in *M. kandleri* is the dominant occurrence of G or C in the third position of the codon. For Phe and Tyr, more than 90% of codons ended with C. More than 70% of the remaining codons for most amino acids, except Arg and Gly, used G and C at the third position of the codon. The 6 and 4 codons for Arg and Gly, respectively, were used evenly in *M. kandleri.* This trend of codon usage in *M. kandleri* was quite different to that in *E. coli* and *M. jannaschii*, a hyperthermophilic methanogen.

Molecular Cloning of a Homologue of the Catalytic Subunit of the CODH Gene

To examine the detailed structure of a constantly expressed gene from *M. kandleri*, a homologue of the catalytic subunit of CODH was cloned using the sequence tag lmk-

Table 2. Codon usage in *M. kandleri* and comparison with other organisms.

Amino acid	Codon	number	*M. kandleri* %	*M. jannaschii* %	*E. coli* %
Phe	UUU	6	6	71	51
	UUC	90	94	29	49
Leu	UUA	12	4	32	11
	UUG	15	5	26	11
	CUU	27	9	9	10
	CUC	110	38	19	10
	CUA	27	9	7	3
	CUG	98	34	7	55
Ile	AUU	20	11	41	47
	AUC	139	76	13	46
	AUA	24	13	47	7
Met	AUG	51	100	100	100
Val	GUU	30	9	64	29
	GUC	133	38	6	20
	GUA	49	14	21	17
	GUG	138	39	9	34
Ser	UCU	12	5	18	19
	UCC	80	32	6	17
	UCA	23	9	30	12
	UCG	67	26	4	13
	AGU	25	10	29	13
	AGC	46	18	13	27
Pro	CCU	29	10	19	16
	CCC	98	33	5	10
	CCA	33	11	68	20
	CCG	134	46	8	55
Thr	ACU	28	12	35	21
	ACC	94	41	11	43
	ACA	15	6	52	13
	ACG	94	41	3	23
Ala	GCU	55	16	48	19
	GCC	130	39	11	25
	GCA	29	9	37	22
	GCG	120	36	4	34
Tyr	UAU	8	10	76	53
	UAC	70	90	24	47
His	CAU	26	24	76	52
	CAC	82	76	24	48
Gln	CAA	28	34	66	31
	CAG	55	66	34	69
Asn	AAU	20	25	71	39
	AAC	61	75	29	61
Lys	AAA	34	23	64	76
	AAG	116	77	36	24
Asp	GAU	72	31	82	59
	GAC	163	69	18	41
Glu	GAA	84	24	43	70
	GAG	272	76	57	30
Cys	UGU	16	24	65	43
	UGC	51	76	35	57
Trp	UGG	39	100	100	100
Arg	CGU	73	18	<1	42
	CGC	76	19	<1	37
	CGA	74	19	<1	5
	CGG	73	18	<1	8
	AGA	39	10	74	4
	AGG	65	16	24	3
Gly	GGU	95	26	20	38
	GGC	98	27	7	40
	GGA	94	26	59	9
	GGG	72	20	15	13

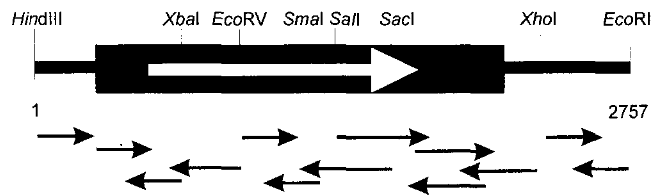

Fig. 1. Restriction map of the 2.8-kb *M. kandleri* DNA fragment containing a homologue of the catalytic subunit of the CODH gene.

The black box represents an open reading frame. The white arrow in the black box indicates the direction of the translation. The black arrows under the restriction map indicate the direction and length of the sequenced DNAs.

ef as a probe. Since CODH is essential for carbon fixation, the structure of the gene encoding the catalytic subunit of CODH may include the canonical features of the *M. kandleri* gene. The DNA sequence of the 2.8 kb-*Hin*dIII/*Eco*RI fragment that hybridized to the probe was determined. There was only one open reading frame (ORF) larger than 1 kb (Fig. 1). The sequence of 2,758

```
1    AAGCTTTGTCGAGCAAATACACCGCACGGAACAAGGAGCGCTTATCTACGCCCGGTCGGAACGCGACGATGTACTTCATCAGCGGCCGCA
91   AGTTAGCGCCCGTCGGTTTGACCGTGTGACACCTGCACCGGGCGTGCAACCTGGGTCCCCCTTGTATCAGTATGTTCGAAAGTGTCCAAA
181  CGTTTGATGAAAATCGAAACTAATATAAGTCAAGTCACCCCTCCAATGGTCGGATAGTATCGAACGTTCCCATAGTACCGAACAAGGGGA
                          boxA                                boxB                     *     RBSs
271  TCCCCGTGGAGGAAAAGAGATCCTCGTGCCCGTACGCCGATGAAGCCGTGTGTGAGCTAGTCGAGCACGCGAAGGAGCTGAATGAGGAGA
        M  E  E  K  R  S  S  C  P  Y  A  D  E  A  V  C  E  L  V  E  H  A  K  E  L  N  E  E  I   29
361  TTCCGGAGATAGAGACCCCGCACATCCGATGGCCCGTTCAGTTCCCGAAGTGCCCGTACGGCAAGCAAGGCGTCTGGTGTAACATCTGCT
30    P  E  I  E  T  P  H  I  R  W  P  V  Q  F  P  K  C  P  Y  G  K  Q  G  V  W  C  N  I  C  S   59
451  CGAACGGCCCGTGCCGAATTACGGAGAAGACACCACGCGGTGTCTGCGGAGCGACGGCGGACGTGATCGTGGCACGCAACTTCTTGCGTC
60    N  G  P  C  R  I  T  E  K  T  P  R  G  V  C  G  A  T  A  D  V  I  V  A  R  N  F  L  R  H   89
541  ACGTCGCCGCCGGAGCCGCGTGCTACGTGCACTGCCTGGAGAACGCCGCCCGGGCCCTCAAGTCCGTGGCGGACGAGGAATCTCCTTACG
90    V  A  A  G  A  A  C  Y  V  H  C  L  E  N  A  A  R  A  L  K  S  V  A  D  E  E  S  P  Y  E  119
631  AGATCGCAGACGAGAAGGCGCTGAGGCACGCTGCCGAAGTATACGGTCTAGATACCTCCGGTAAACCGGAAGATGTGGCCGAGGAGATCG
120   I  A  D  E  K  A  L  R  H  A  A  E  V  Y  G  L  D  T  S  G  K  P  E  D  V  A  E  E  I  A  149
721  CAGAGTTCATTCTCGAGGATATCTACAGGCCCAGATATGAGGAGTCCGAGGTGTTCAAGGCTGTTGTACCCGATTGGCGCATCGAGATGT
150   E  F  I  L  E  D  I  Y  R  P  R  Y  E  E  S  E  V  F  K  A  V  V  P  D  W  R  I  E  M  Y  179
811  ACGAGGAAATGGGCCTAATTCCGGGCGGAGCTAAGTCGGAGATACACGACGCCCTTGTCAAGACGAGTACGAACCTCAACTCAGACCCCG
180   E  E  M  G  L  I  P  G  G  A  K  S  E  I  H  D  A  L  V  K  T  S  T  N  L  N  S  D  P  V  209
901  TTGACATGCTCCTCCACGTGCTGCGACTGGGACTGATCACCGGCCCTGTCGCGCTCTTCGGTGTGGAGACGATCAACGATATCCTCTTCG
210   D  M  L  L  H  V  L  R  L  G  L  I  T  G  P  V  A  L  F  G  V  E  T  I  N  D  I  L  F  G  239
991  GCAGCCCGAAGATCACGCAAACGGAGGGTGGTCCCGGAATACTGGACCCGGACTACGTCAACATCATGACGACCGGCCACCAGATGGCTC
240   S  P  K  I  T  Q  T  E  G  G  P  G  I  L  D  P  D  Y  V  N  I  M  T  T  G  H  Q  M  A  L  269
1081 TCATGAAGTACCTGACGGACGCCGCCGAGAAGCTGGAAGAAGAGGCCAAGGCGGCGGGAGCTAAAGGTATCCGGATCATCGGTGCGACGT
270   M  K  Y  L  T  D  A  A  E  K  L  E  E  E  A  K  A  A  G  A  K  G  I  R  I  I  G  A  T  C  299
1171 GCGTCGGCGACGATTTCGAGGCGCGCGCCGAGCACCTACCAGACACGTACGCGGGCTTCGCGGGTAACAATTTCGCGACGGAGGCGCTCG
300   V  G  D  D  F  E  A  R  A  E  H  L  P  D  T  Y  A  G  F  A  G  N  N  F  A  T  E  A  L  A  329
1261 CAGCAACGGGTCTCGTGGACGCGATCGTGAGCGAGTTCAACTGCACGTTCCCGGGTCTGAAGTTCTACAAGGAGAAGCTGGGACGTCGAG
330   A  T  G  L  V  D  A  I  V  S  E  F  N  C  T  F  P  G  L  K  F  Y  K  E  K  L  G  R  R  A  359
1351 CTGGTCGCCGTCGACGACGTGGCCAAGGTGTGGGGCGCGGACTGATCCTCTGGGACCCGGAGCGTGCCGAGGAGGTCGCGGAGGAGGCCG
360   G  R  R  R  R  R  G  Q  G  V  G  R  G  L  I  L  W  D  P  E  R  A  E  E  V  A  E  E  A  V  389
1441 TCCAGAGGGCCATTGAGGCGTTCAAGGAGCGCCGGTCGAAGCACGAGGACAAGATCATGGAGCCGAATGCACAGGCACGAGAACGTGGTG
390   Q  R  A  I  E  A  F  K  E  R  R  S  K  W  E  D  K  I  M  E  P  N  A  Q  A  R  E  R  G  G  419
1531 GGGTTCGGATACTTCTCGATCGAGGAGGCCGTCGGCTGGGAGAACGTGCTCAAGCTGATCGAAGAAGCACCATCCGCGGCGTGTGCGCCA
420   V  R  I  L  L  D  R  G  G  R  R  L  G  E  R  A  Q  A  D  R  R  S  T  I  R  G  V  C  A  I  449
1621 TCATGGGCTGCACGAATCTCTCCAGCGGCGGCCACAACGTGCCGGCGGTCGAGCTCGCCAAGGAGATGATCAAGCGCGACGTGCTCGTGC
450   M  G  C  T  N  L  S  S  G  G  H  N  V  P  A  V  E  L  A  K  E  M  I  K  R  D  V  L  V  L  479
1711 TCGGCGCCGGCTGTGTGAACGGGGCGTTCGCCAACGCTGGACTGTTCAACCCTGAAGCAGCCGAGCTGGCCGGTGACAACCTCCGACAGG
480   G  A  G  C  V  N  G  A  F  A  N  A  G  L  F  N  P  E  A  A  E  L  A  G  D  N  L  R  Q  V  509
1801 TGTGCGAGGAGTTAGGTATTCCACCGGTGCTGCATTACGGCCCATGTCTGGCGATCGGTAAGATCGAGCATCTCGTCTTCGAGATCGCCG
510   C  E  E  L  G  I  P  P  V  L  H  Y  G  P  C  L  A  I  G  K  I  E  H  L  V  F  E  I  A  E  539
1891 AGATACTTCGAGAGAAAACCGGCGAGGAAATCGACATCCCGGACGTACCCGCCGTCGCGAGCGCGCCGCAGTGGCTCGAGGAGCAAGCCC
540   I  L  R  E  K  T  G  E  E  I  D  I  P  D  V  P  A  V  A  S  A  P  Q  W  L  E  E  Q  A  L  569
1981 TAGCGGACGCGTCCTCGGCCTTAGCTCTAGGTATCACGCTGCACGTGAGTCCCGTACCGCCGGTGACCGGTAGCGAGCTGGTTACGAAGA
570   A  D  A  S  S  A  L  A  L  G  I  T  L  H  V  S  P  V  P  P  V  T  G  S  E  L  V  T  K  T  599
2071 CGCTGCTGGAGGACCTACCGGATCTCACCGGCGGCGAGCTGATCGTGGAGACGGACATGAAGCGTGCCGGAGAGATTCTCGCCGAGAAGA
600   L  L  E  D  L  P  D  L  T  G  G  E  L  I  V  E  T  D  M  K  R  A  G  E  I  L  A  E  K  I  629
2161 TCGAGGAGAAGCGGAAACGGCTTGGTATCTAATTCTCACCCCCACTTCCGGGATGGTGGTGACCATGGAGATCGAGGCGATCCGTGTAGC
630   E  E  K  R  K  R  L  G  I  #                                                             638
2251 AACGGGAGCTCCTGGCGAACTGTATATGGTAGCCGAGATCGCGGAGGACAAGGTCGTGAACGCCGAAGTGTGCTCGGGCACGACACCGAT
2341 CTCCTGGGAGACGCTCGTCCTCGAGAAGCCCGTGGAGTTCGCCGTAGTGGCCGCCGAACGAGTATGCGCCCAGGTGACGCGAGCCTAGCG
2431 CGTCGTGGTCGAGGCGGCCGAGGCAGCTCTCAACGTCGAGATCCCGGACGAGGCGGACGAGCGCGCAGATCCTGCATGGCTAACGTCGTC
2521 AGGGCGCACCACCGTCTGGCAGGACTGACTTCCTCGACGTGAAGAAGCCCGCGTACGACCTAGTGAGAGCTGCGAAGGACATCATGCACA
2611 CGCCGTCGGAGGGAAGCCCGACCACCCACCGGCCGTGACCGTAGGCGGTTTGGACCTCGAAAAGCTGCCCGTAGACAGGGTTGAATCCGT
2701 AGCTAAGGACGCCGCCGAGGTGGCAGACAGGCTTGAGGACGAGGTCGGATCCGAATTC
```

Fig. 2. Nucleotide sequences of the cloned gene from *M. kandleri* and its deduced amino acid sequence denoted by a single letter code.

The underlined part is the conserved archaeal promoter region [box A (TATA box), box B] and putative ribosome binding site (RBSs). The bold **C** at the 257 nucleotide, marked by an asterisk is the start site of transcription. The ochre stop codon is marked by the # symbol.

nucleotides including the unique ORF and adjacent sequences are shown in Fig. 2. The amino acid sequence translated from the ORF is denoted by a single letter code under the first codon base. Around the 5' end of the ORF, there was no ATG codon. When the sequence homology of the ORF was compared with the catalytic subunit of a CODH gene (Fig. 2), the translational initiation site of the ORF was assumed to be located at nucleotides 260~290. There was one GTG codon at the 276th base and this GTG codon was assigned as a putative translational start codon region, since the GTG codon is occasionally used as a translation initiation codon in methanogens [11, 25]. When the presumed protein coding sequence of the cloned gene was expressed in *E. coli* using a pET expression system, it produce a soluble and heat stable protein (data not shown). These results indicated that the cloned ORF represented a homologue of CODH catalytic subunit, and produced a well-folded protein.

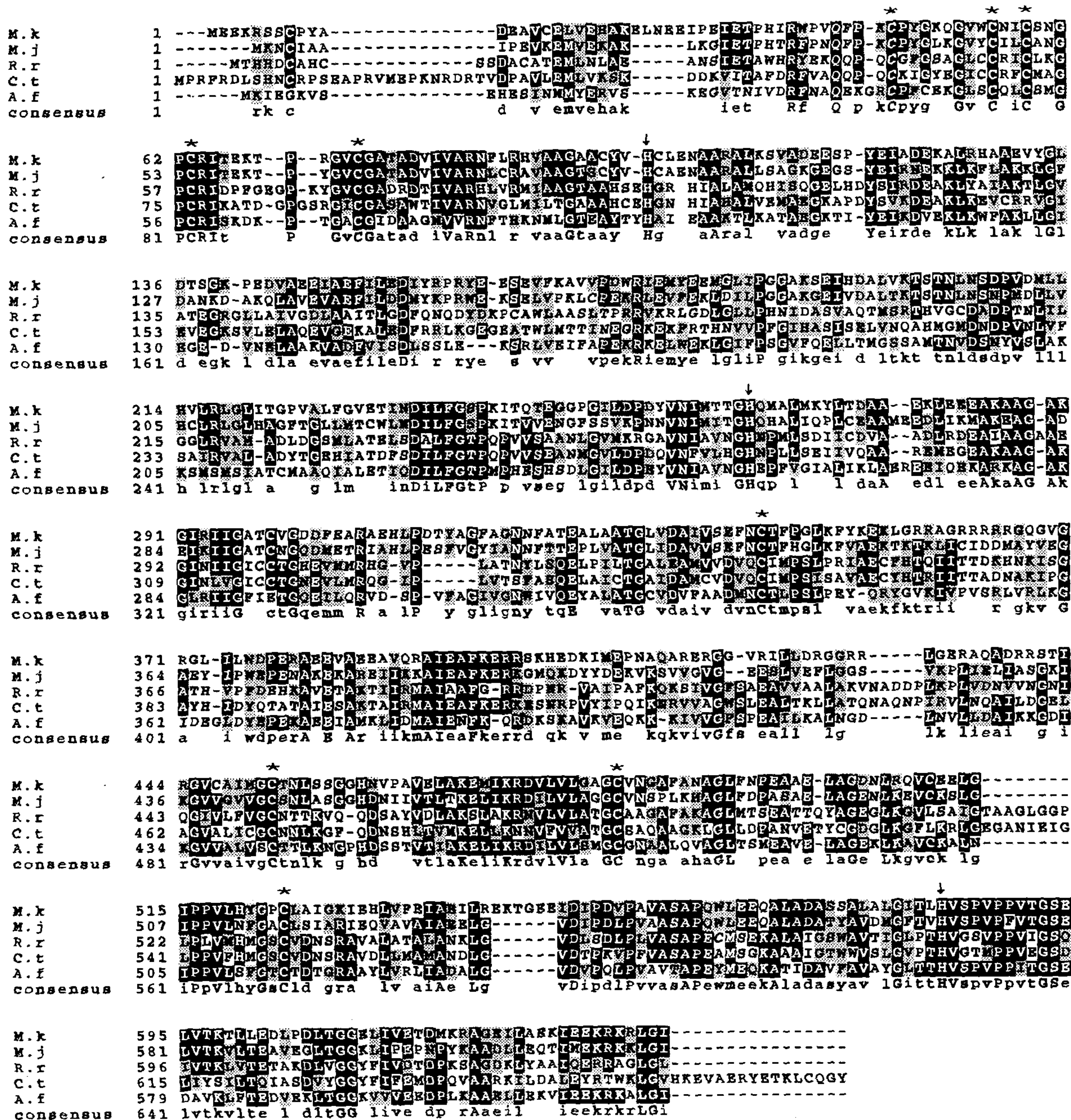

Fig. 3. Comparison of the amino acid sequence of the cloned gene with the CODH catalytic subunits from *Methanopyrus kandleri* (M. k.), *Methanococcus jannaschii*, *Rhodospirillum rubrum* (R. r.), *Clostridium thermoaceticum* (C. t.), and *Archaeoglobus fulgidus* (A. f.). The conserved amino acids are represented as black boxes, and similar amino acids are represented as gray boxes. The conserved Cys and His residues are marked by * and ↓.

Cis-Elements for Gene Expression

Two elements for transcriptional control, (A/T)(A/T)TATA and (A/T)TG(C/G), were observed at the 5' end of the house-keeping genes from the methanogens [3, 11, 25]. Sequences similar to these promoters were identified at 74 bp (AATATA) and 50 bp (ATGG) upstream from the translational initiation codon (Fig. 2). Between these promoter-like sequences and the putative translation codon, an AGGGGA sequence that was similar to the consensus ribosomal binding sites was found. To confirm the role of these sequences, the transcription start site was determined. When a reverse transcription reaction was performed using the total RNA of *M. kandleri*, two bands were observed as the major products. The primer-extension reaction was ended at the C at the 18th base upstream from the putative translation start site, and the minor product ended at the 16th base upstream (Fig. 4). These results indicate that the cloned gene is expressed in *M. kandleri,* and the location of the transcription start site suggests that the putative assigned sequences are the promoters and ribosomal binding sites of the cloned gene from *M. kandleri.* No canonical transcriptional termination sequence such as an inverted-repeat or purine-rich sequence [3, 25] was found downstream of the stop codon.

Sequence Comparison

When the amino acid sequence of the cloned gene and the CODH catalytic subunits from *Methanococcus jannaschii*

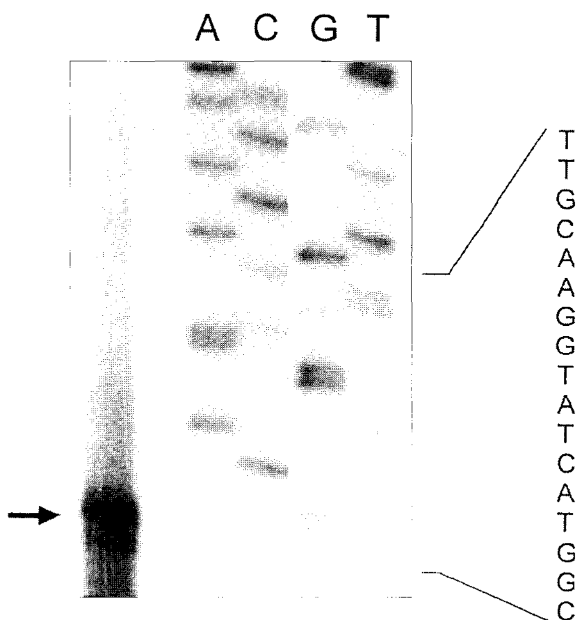

Fig. 4. Determination of the initiation site of transcription.
The primer extension reaction was performed on the total RNAs of *M. kandleri* using the labeled primer EXT. An asterisk marks the start site of transcription. The recombinant plasmid pHIN was sequenced using EXT as a primer to determine the length of the reaction product of the primer extension reaction.

[4], *Clostridium thermoaceticum* [22], *Rhodospirillum rubrum* [14], and *Archaeoglobus furgidus* [17] were compared using the CLUSTAL W program, they showed a strong homology to each other. The sequence identities between the cloned gene and the CODH of *M. jannaschii, C. thermoaceticum, R. rubrum,* and *A. furgidus* were 54%, 31%, 33%, and 37%, respectively. Metal clusters (Fe-S clusters) and Ni ions are essential to the activity of CODH, and cysteine and histidine residues are involved in the coordination of these metal ions [9, 10, 14]. There were 9 and 3 conserved cysteins and histidines, respectively, observed in all the tested sequences (Fig. 3). The conserved 9 cysteines have been reported as coordinating with an Fe/S cluster and the 3 histidines are known for chelation in a Ni-containing enzyme. In particular, the cysteine cluster that occurred at the N-terminus of the protein (Cys-X_2-Cys-X_4-Cys-X_{10-14}-Cys) was assumed to be involved in the coordination of Fe_4S_4 centers [8].

The structural analysis of proteins from hyperthermophiles has shown that they have additional ion-pairs or salt bridges

Table 3. Amino acid composition of CODH catalytic subunits.

Amino acids	*M. kandleri*	*M. jannaschii*	*A. fulgidus*	*C. thermoaceticum*	*R. rubrum*
Ala	73	63	62	75	90
Cys	16	20	13	21	20
Asp	**31**	**25**	**31**	**33**	**38**
Glu	**73**	**56**	**57**	**46**	**25**
Phe	16	20	18	20	12
Gly	57	51	46	56	59
His	14	11	8	17	21
Ile	42	43	46	44	41
Lys	**31**	**55**	**54**	**35**	**27**
Leu	57	63	58	53	62
Met	11	12	18	22	18
Asn	19	26	18	29	21
Pro	34	32	25	31	28
Gln	11	10	15	21	19
Arg	**42**	**15**	**19**	**31**	**26**
Ser	21	24	33	30	32
Thr	28	30	31	33	37
Val	44	51	55	57	50
Trp	5	4	3	6	3
Tyr	13	13	12	14	10
Total	638	624	622	674	639
Number[a]	177	151	161	145	116
Ratio[b]	27.7	24.1	25.8	21.5	18.1
T_{opt} (°C)[c]	98	85	85	55	30

[a]The number of charged amino acids.
[b]The percentage of charged amino acids (%).
[c]Optimum growth temperature of the organism.
The charged amino acids are indicated in bold.

compared to mesophilic proteins [20], and the percentage of charged amino acids in hyperthermophilic proteins is generally higher than that in mesophilic or thermophilic proteins [6, 20]. To examine whether the cloned gene had such characteristics, its amino acid composition was obtained and compared with those of the CODHs from different organisms (Table 3). In general, the number of charged amino acids was well correlated to the growth temperature of the organism. The ratio of charged amino acid in the cloned gene was more than 50% higher than in mesophilic CODH. This result suggests that the cloned gene from *M. kandleri* includes more charged amino acids, and that it exhibits an increased ion-pair interaction that can help maintain heat stability. The cloned gene contained a significantly higher quantity of glutamate. It would be interesting to examine the role of this increased glutamate in the stability of the protein.

Acknowledgments

This research was supported by grants from the Korea Institute of Science and Technology (KIST) and from the Ministry of Science and Technology, Korea (Biotech2000 program).

REFERENCES

1. Altschul, S. F., L. M. Thomas, A. S. Alejandro, Z. Jinghui, Z. Zheng, M. Webb, and D. J. Lipman. 1997. Gapped BLAST and PSI-BLAST: A new generation of protein database search programs. *Nucl. Acids Res.* **25:** 3389-3402.
2. Barross, J. A. and J. F. Holden. 1996. Overview of hyperthermophiles and heat-shock proteins. *Adv. Prot. Chem.* **48:** 1-34.
3. Brown, J. W., C. J. Daniels, and J. N. Reeve. 1989. Gene structure, organization, and expression in Archaebacteria. *CRC Crit. Rev. Microbiol.* **16:** 287-338.
4. 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. L. Fuhrmann, D. Nguyen, T. R. Utterback, J. M. Kelly, J. D. Peterson, P. W. Sadow, M. C. Hanna, M. D. Cotton, K. M. Roberts, M. A. Hurst, 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.
5. Burggraf, S., K. O. Stetter, P. Pouviere, and C. R. Woese. 1991. *Methanopyrus kandleri*: An archael methanogen unrelated to all other know methanogens. *Syst. Appl. Microbiol.* **14:** 346-351.
6. Choi, I.-G., S. S. Kim, J.-R. Ryu, Y. S. Han, W.-G. Bang, S.-H. Kim, and Y. G. Yu. 1997. Random sequence analysis of genomic DNA of a hyperthermophile: *Aquifex pyrophilus. Extremophiles* **1:** 125-134.
7. Chomczynski, P. and N. Sacchi. 1987. Single-step method of RNA isolation by acid guanidium thiocyanate-phenol-chloroform extraction. *Anal. Biochem.* **162:** 156-159.
8. Colden, C. D., H. W. Hellinga, and J. P. Caradonna. 1997. The rational design and construction of a cuboidal iron-sulfur protein. *Proc. Natl. Acad. Sci. USA* **94:** 6635-6640.
9. Ensign, S. A. and P. W. Ludden. 1991. Characterization of the CO oxidation/H_2 evolution system of *Rhodospirillum rubrum. J. Biol. Chem.* **266:** 18395-18403.
10. Ferry, J. G. (1995) CO dehydrogenase. *Annu. Rev. Microbiol.* **49:** 305-333.
11. Hansner, W., G. Frey, and M. Thomm. 1991. Control regions of an archael gene: A TATA box and an inhibition element promote cell-free transcription of the $tRNA^{val}$ gene of *Methanococcus vannielii. J. Mol. Biol.* **222:** 495-508.
12. Harms, U. and R. K. Thauer. 1997. Identification of the active site histidine in the corrinoid protein MtrA of the energy conserving methyltransferase complex from *Methanobacterium thermoautotrophicum. Eur. J. Biochem.* **250:** 783-788.
13. Huber, R., M. Kurr, H. W. Jannasch, and K. O. Stetter. 1989. A novel group of abyssal methanogenic archaebacteria (*Methanopyrus*) growing at 110. *Nature* (*London*) **342:** 833-834.
14. Kerby, R. L., S. S. Hong, S. A. Ensign, L. J. Coppoc, P. W. Ludden, and G. P. Roberts. 1992. Genetic and physiological characterization of the *Rhodospirillum rubrum* carbon monoxide dehydrogenase system. *J. Bacteriol.* **174:** 5284-5294.
15. Kim, C. W., P. Markiewicz, J. J. Lee, C. F. Schierle, and J. H. Miller. 1993. Studies of the hyperthermophile *Thermotoga maritima* by random sequencing of cDNA and genomic libraries: Identification and sequencing of the *trpEF* (D) operon. *J. Mol. Biol.* **231:** 960-981.
16. Klein, A. R. and R. K. Thauer. 1997. Overexpression of the coenzyme-F_{420}-dependent N^5, N^{10}-methylenetetrahydromethanopterin dehydrogenase gene from the hyperthermophilic *Methanopyrus kandleri. Eur. J. Biochem.* **245:** 386-391.
17. 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. Peterson, 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, sulphate-reducing archaeon *Archaeoglobus fulgidus. Nature* **390:** 364-370.
18. Krah, R., M. H. O'Dea, and M. Gellerb. 1997. Reverse gyrase from *Methanopyrus kandleri. J. Biol. Chem.* **272:** 13986-13990.

19. Kurr, M., R. Huber, H. Konig, H. W. Jannasch, H. Fricke, A. Trincone, I. K. Kristjansson, and K. O. Stetter. 1991. *Methanopyrus kandleri*, gen. and sp. nov. represents a novel group of hyperthermophilic methanogens, growing at 110°C. *Arch. Microbiol.* **156:** 239-247.
20. Lim, J. W., Y. G. Yu, I. G. Choi, J. R. Ryu, B. Y. Ahn, S. H. Kim, and Y. S. Han. 1997. The crystal structure of an Fe-Superoxide dismutase from the hyperthermophile *Aquifex pyrophilus* at 1.9 resolution: Structural basis for thermostability. *J. Mol. Biol.* **270:** 259-274.
21. Morton, T. A., J. A. Runquist, S. W. Ragsdale, T. Shanmugasundaram, H. G. Wood, and L. G. Ljungdahl. 1991. The primary structure of the subunits of carbon monoxide dehydrogenase/acetyl-CoA synthase from *Clostridium thermoaceticum*. *J. Biol. Chem.* **266:** 23828.
22. Murakawa, G. J., C. Kwan, J. Yamashita, and D. P. Nierlich. 1991. Transcription and decay of the lac messenger: Role of an intergenic terminator. *J. Bacteriol.* **173:** 28-36.
23. Nolling, J., A. Elfner, J. R. Palmer, V. J. Steigerwald, T. D. Pihl, J. A. Lake, and J. N. Reeve. 1996. The phylogeny of the *Methanopyrus kandleri* and methyl coenzyme M reductase operons. *Int. J. Syst. Bacteriol.* **46:** 1170-1173.
24. Ragsdale, S. W. 1991. Enzymology of the Acetyl-CoA pathway of CO_2 fixation. *Crit. Rev. Biochem. Mol. Biol.* **26:** 261-300.
25. Reeve, J. N. 1992. Molecular biology of methanogens. *Annu. Rev. Microbiol.* **46:** 165-191.
26. Sambrook, J., E. F. Fritsch, and T. Maniatis. 1989. *Molecular Cloning: A Laboratory Manual.* 2nd Ed. Cold Spring Harbor Laboratory Press, Cold Spring Harbor, N.Y.
27. Sanger, F., S. Nicklen, and A. R. Coulson. 1977. DNA sequencing with chain-terminating inhibitors. *Proc. Natl. Acad. Sci. USA* **74:** 5463-5467.
28. Shima, S., D. S. Weiss, and R. K. Thauer. 1995. Formylmethanofuran: Tetrahydromethanopterin formyltransferase (Ftr) from the hyperthermophilic *Methanopyrus kandleri*. *Eur. J. Biochem.* **230:** 906-913.
29. Thauer, R. K., R. Hedderich, and R. Fischer. 1993. Reactions and enzymes involved in methanogenesis from CO_2 and H_2, *In* Ferry, J. G. (ed.), *Methanogenesis* pp. 207-252. Champman & Hall, Inc.
30. 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. *Nucl. Acids Res.* **22:** 4673-4680.
31. Weil, C. F., D. S. Cram, B. A. Sherlf, and J. N. Reeve. 1988. Structure and comparative analysis of the genes encoding component C of methyl Coenzyme M reductase in the extremely thermophilic Archaebacterium *Methanothermus fervidus*. *J. Bacteriol.* **170:** 4718-4726.
32. Woese, C. R., O. Kandler, and M. W. Wheelis. 1990. Towards a natural system of organisms: Proposal for the domains Archaea, Bacteria, and Eucarya. *Proc. Natl. Acad. Sci. USA* **87:** 4576-4579.
33. Xia, J., J. F. Sinclair, T. O. Baldwin, and P. A. Lindahl. 1996. Carbon monoxide dehydrogenase from *Clostridium thermoaceticum*: Quaternary structure, stoichiometry of its SDS-induced dissociation, and characterization of the faster-migrating form. *Biochemistry* **35:** 1965-1971.