

Iron-Sulfur Flavoprotein (Isf) from *Methanosarcina thermophila* Is the Prototype of a Widely Distributed Family

TONG ZHAO, FRANCISCO CRUZ, AND JAMES G. FERRY*

Department of Biochemistry and Molecular Biology, Eberly College of Science, The Pennsylvania State University, University Park, Pennsylvania 16802-4500

Received 11 April 2001/Accepted 7 August 2001

A total of 35 homologs of the iron-sulfur flavoprotein (Isf) from *Methanosarcina thermophila* were identified in databases. All three domains were represented, and multiple homologs were present in several species. An unusually compact cysteine motif ligating the 4Fe-4S cluster in Isf is conserved in all of the homologs except two, in which either an aspartate or a histidine has replaced the second cysteine in the motif. A phylogenetic analysis of Isf homologs identified four subgroups, two of which were supported by bootstrap data. Three homologs from metabolically and phylogenetically diverse species in the *Bacteria* and *Archaea* domains (Af3 from *Archaeoglobus fulgidus*, Cd1 from *Clostridium difficile*, and Mj2 from *Methanococcus jannaschii*) were overproduced in *Escherichia coli*. Each homolog purified as a homodimer, and the UV-visible absorption spectra were nearly identical to that of Isf. After reconstitution with iron, sulfide, and flavin mononucleotide (FMN) the homologs contained six to eight nonheme iron atoms and 1.6 to 1.7 FMN molecules per dimer, suggesting that two 4Fe-4S or 3Fe-4S clusters and two FMN cofactors were bound to each dimer, which is consistent with Isf data. Homologs Af3 and Mj2 were reduced by CO in reactions catalyzed by cell extract of acetate-grown *M. thermophila*, but Cd1 was not. Homologs Af3 and Mj2 were reduced by CO in reactions catalyzed by *A. fulgidus* and *M. jannaschii* cell extracts. Cell extract of *Clostridium thermoaceticum* catalyzed CO reduction of Cd1. Our database sequence analyses and biochemical characterizations indicate that Isf is the prototype of a family of iron-sulfur flavoproteins that occur in members of all three domains.

Methane produced by the acetate fermentation pathway accounts for two-thirds of all biologically produced methane. In *Methanosarcina thermophila*, acetate is cleaved into carbonyl and methyl groups by the CO dehydrogenase (CODH)–acetyl coenzyme A (acetyl-CoA) synthase (ACS) enzyme complex. The methyl group is subsequently reduced to methane by electrons derived from oxidation of the carbonyl group to CO₂ by the CODH-ACS complex (7). An iron-sulfur flavoprotein (Isf) from *M. thermophila* has been overproduced in *Escherichia coli*, purified, and characterized. Isf is purified as a homodimer which binds two 4Fe-4S clusters and two flavin mononucleotide (FMN) cofactors (2, 14). Ferredoxin, the electron acceptor of the CODH-ACS complex, is the physiological electron donor for Isf. Additional results support the hypothesis that Isf plays a role in electron transport during fermentation of acetate to methane (14). The deduced amino acid sequence of Isf contains an unusually compact cysteine motif (CX₂CX₂CX₅C) that has been shown, on the basis of site-directed mutagenesis and electron paramagnetic resonance studies, to ligate the 4Fe-4S cluster (15). The midpoint potential values are -394 mV for the 4Fe-4S cluster and -277 mV for the FMN in Isf, suggesting that electrons derived from oxidation of the carbonyl group of acetate flow from ferredoxin to the 4Fe-4S cluster of Isf and then to the FMN (2). The downstream electron acceptor of Isf has not been identified yet.

When Isf was initially discovered, no sequences which ex-

hibited significant overall identity were found in the databases (14). Subsequently, homologs were identified in the genomes of methanogenic and nonmethanogenic members of the *Archaea* domain (2) and several organisms belonging to the *Bacteria* domain (15). Here we describe an analysis of these homologs and 24 additional homologs and also describe biochemical characterization of three homologs from diverse species. Our findings indicate that Isf is the prototype of a family of iron-sulfur flavoproteins that are more widely distributed in nature than previously recognized.

MATERIALS AND METHODS

Materials. *Archaeoglobus fulgidus* genomic DNA was a gift from Michael Adams, University of Georgia. *A. fulgidus* cells grown on lactate were a gift from Robert Kelly, North Carolina State University. *Clostridium difficile* genomic DNA was a gift from TechLab, Inc., Blacksburg, Va. *Clostridium thermoaceticum* cells grown on glucose and CO₂ as previously described (1) was a gift from Steve Ragsdale, University of Nebraska. *Methanococcus jannaschii* clone AMJAI79 was purchased from the American Type Culture Collection. Autotrophically grown cells of *M. jannaschii* JAL-1 were a gift from Biswarup Mukhopadhyay, University of Illinois. Acetate-grown *M. thermophila* cells were a gift from Birthe Borup. pSJS1240 was a gift from Rosalind Kim, University of California, Berkeley. Restriction endonucleases, T4 DNA ligase, and Vent DNA polymerase were purchased from New England Biolabs, Inc. Oligonucleotide primers were synthesized at Integrated DNA Technologies, Inc. All *E. coli* strains were obtained from Novagen, Inc. Phenyl Sepharose high-performance matrix, HiTrapQ, HiTrap SP, HiTrap Desalting, PD-10, Superose-12, and Mono-Q columns were obtained from Pharmacia. Amicon stirred cell 8200 and YM-10 membranes were obtained from Millipore. Tryptone and yeast extract were obtained from Difco. All other chemicals were obtained from Sigma.

Sequence analysis. BLAST searches were performed at <http://www.ncbi.nlm.nih.gov> and http://www.jgi.doe.gov/JGI_microbial/html/index.html. BLAST searches were also performed with preliminary genome data for *Methanosarcina mazei* from the Göttingen Genomic Laboratory at <http://www.g21.bio.uni-goettingen.de/>. Sequences were aligned by using Clustal X, version 1.64b. A phylogenetic analysis of protein sequences was performed with the MEGA

* Corresponding author. Mailing address: Department of Biochemistry and Molecular Biology, Eberly College of Science, The Pennsylvania State University, 205 South Frear Laboratory, University Park, PA 16802-4500. Phone: (814) 863-5721. Fax: (814) 863-5721. E-mail: JGF3@psu.edu.

* * *

Ca1	-----MKVIGINGSARKDGNISILIKTVFKELEKE--GIETELIELSGKTIKGMGCFCKEK--KQCIIRDDFFNECF	70
Msb3	-----MKVIGINGSRKRKDGNTAILIQTVFEELTKE--GIETELIQLSGKNIEGGKACWACHKNRNKQCVITDDFFNECF	72
Ca2	-----MKVVAFNKSPKKNKNGTNEAIKVAEAELEKE--NIEVEIIHVGNKVRIGCLDCGGCSRNMRNRCVMKNDEVNDWI	72
Gs	-----MKVVAFNKSPKADGNTWHALKMTAELESE--GIETEVHVGNKVVVGGIACGQCAKNKNERCVLPGDEVNEWI	72
Ch1	-----MKVVAFNKSPRANGNTYQALKLVLAELEKE--GIEGEEIIQVGDKVRIGCLACGKCVRNKNERCAID--DEVNDWI	71
De1	-----MKVVAFNKSPRAQGNTRRGLEMVLAELEKE--GIQTEIVQLGGKRVFGCLACGQCKINKNLRCARNDDMNFI	72
Msb4	-----MKVVAFNKSPRKGNTAALIKHVLAELEKE--GIETEVVQIGGKKIHGCTACGKCYEKMDKKCVI DKKI VNECI	72
Dv	-----MYALATNGSPRKGNTIIMLNKALEPLAAA--GWETELVQVGGRNIRGCIACYKCFENKDRCAVTKDKFNDVM	72
Eh	-----MSKPIKVFMINSGSPREKGNFTMLSWVKEGLEKE--GIEVEIYQLGKKEIKPCVACVACQKT--GKCWKDSDVMDELL	74
Ca3	-----MKVLLINGSHPHAKGCTYTALSEVAKALEEE--NIETEIFQIGTKPVRGCIQGGTCSKTGSGKCVFDDDTLNLAG	72
Cd2	-----MKVLLINGSBNQYCTYTALNETKVLKSKH--DIKTEILYLKGETIPGCISCASCFTG--KCI RNDK-VNELI	69
De2	-----MKVLAVSGSPRPQGNATFLLEDEVLSELEKA--GIETGRIDIAQANIKPCLGHDNCAEYK--KCLQHDD-MDAIF	69
Mst	-----MKITGISGSPRKGQCEKIGALEVAKER--GFETDVTVFISNEEVAPCKACGACR-DQ-DFCVLDDDM-DEIY	69
Mm1	-----MKILGISGSPRKGQCEKMIASALEVAKER--GFETDVTFLSCSEVAPCKACGACR-EK-DCCVLDDSM-EEII	69
Msb1	-----MKILGISGSPRKGQCEKMAIAALEIAKDR--GFETETVFLSNEEIAAPCKVCGACI-EK-SSCVIEDSM-EKVI	69
Mm2	---MQETEAKPIRILGISGSPRN-MATDYLVOEALKIAKEY--GAETEYFSAKGGKLNFCIHDFCIRKK-EGCINHDDIAAELY	79
Msb2	---MQETEAKPIKILGISGSPRN-MATDFLVREALKIAEKEY--GAETDYFSAKGGKLNFCIHDFCLKTK-KGCIHDDISAEIY	79
Msb5	-----MKVIGIVGSPRKNGNINILVQVLEGAEEA--GAETRTIFILNEMNYKGCQGCYDCKGHD--KCKLEDD-LTEVF	69
Cd4	-----MIITVINGSPRKNGATSKVLTYLYKDIERLIPDVKINYFDLSEVNPSYICGLNLCYKMG--KCIQNQDKVEYIH	72
Dh	-----MKILGINGSHRKGNKNTAALLKIVLEETERL--GCSTELLELADLDIKLGIACNVLCRS--RCSVAGDDMAMVE	70
Ch3	-----MAVANSKYLILSTSPRIGNSDTAARYVETLLKAK--GDGVTFLNTRRKKINPCLGNLCSRDG--RCVQKDD-MAEII	74
Cd3	MDSNDKKGKLSPEIIIIINGSPRINKNCSFTANEVKAMLEEC--GLSCKIFNIYNDIEYCTACGFCCKTG--YCKFNDD-MKPLY	80
Ch2	-----MKVVFALMGSPREGSNTDILIEALKPAREA--GWEVEKFNIDQLYINPCRGCMCECRGTG--DCSLEEDDILVVA	70
Xf	-----MKICILVGSNRRNGISRQVCGLIKQYINCR--AIVSFI FLAEKRIELQDADNNS-NK--KCSLNDD-MEYII	77

Ca1	NKMVNADGIIIGSPVYSADISSKMKAFLEBRAG-----VIVATNSG-MLKHKVGCVVVA-----VRRGGMTAVDALNHFLLNK	142
Msb3	AKMLVSDGIIIGSPVYSAGVTSQMKALIDRAS-----MVLAGNRG-LFKHKV GASVVA-----ARRGGATSAFGLTNNFLHSK	144
Ca2	QKMKEADGIIIGSPVHYSATAGTMKSFDRAF-----LVTSVNNG-MLRHKV GASVVA-----VRRSGGIPVFNQLNNYINYS	144
Gs	QKMKQADGIIISPPVYSATAGTMKSFDRAF-----YVTGVNDG-MLRHKV GASVVA-----VRRSGGIPVFNQLNNFLTYA	144
Ch1	QKMKEADGIIIGSPVHYSATATMKAFLEBRAF-----YVASANGG-LFRHKV GASVVA-----VRRSGGVAAFDQLNHVLYHA	143
De1	AKMFEADGIIIGSPYFNSVNAEVKALIXRAG-----MVARANDH-ALKGKVGAPLVA-----ARRAGATHVYSAINFFFGIN	144
Msb4	EKMLEADGIIIGSPYFADLTPELKAIDRAG-----FVAKANNE-MFRHKV GASVVA-----VRRAGSIHVFDISINHEFTIS	144
Dv	EKMLRADALIGSPYFTDVTAEKAVLDRSG-----LVAIANGT-AFRGKIGAAVVV-----ARRGGATHVYDTINHMFLMS	144
Eh	QKIEADGIIIGSPYFADVPVGVKTFIDRSQ-----PIIYYS-KK-TLFRKVGAAVVM-----ARRGGATHVYDTINHWFGIN	145
Ca3	EKAKTADGFIIGSPVHYAAASGALTSFLDR-F-----FYAGK-----DFAYKPGAIVS-----CRRGAAAADFQLNKYFTIA	140
Cd2	EDLDNIDGIIIGSPVYFSSATGQLTSFLDRLF-----FIAGS-----RMAATKLGASVVS-----CRRGGASATFQDLNKYFSTIS	138
De2	KDMAEADGLIIVATPVYYNNTAQLKAFIDR-H-----YFLYKHHI-SLEFFKAVGVVIL-----AEEQGLDLDLNTLVQYLDSD	140
Mst	EKMRAADGIIIVAAPVYMGNYPAQLKALFDRS-----VLLRRKNF-ALKNKVGAAALSVG-----GSRNGGQEKTIQSIHDMWHI	142
Mm1	EKMRAADGIIIVAAPVYMGNYPAQLKALFDRS-----VLLRRKDF-ALKNKVGAAALSVG-----GSRNGGQEKTIQSIHDMWHI	142
Msb1	EKMKAADGIIIVASPVYMGNYSQLKALFDRS-----VLLRRNNF-ALKNKVGAAALSVG-----GSRNGGQEKTIQSIHDMWHI	142
Mm2	DKMIWADAWIIGTPVYQGTLSAQTKTIMDRCR-----AVVARDPK-VFLNKVGMGIADG-----GDRIGGQEPAIQTIHTFYIN	153
Msb2	DKMIWADAWIIGTPVYQGTLSAQTKTIMDRCR-----AVVARNPK-VFLNKVGMGIADG-----GDRIGGQEPAIQTIHNFYIN	153
Msb5	EELTSADGVVFGSPYIYFGQFTGQMLFLDRCY-----SLINAD-----FSPRLPAGKAV-----IVGTQGAPDPAAFKGVYEEFK	139
Cd4	DIITKSDGVIIGSPYIYSSVTGLFKVFTDRAH-----MMLER-----LLYRKPCTIAVTT-----YENARGSKAISFIKSMVLDS	141
Dh	HKLEADGIIIGSPVYWSNVTALMKNFMDRSR-----YLHITRN--LLAGKAGAAVTNA-----GLLHGGQEATLKIMEGELQAH	143
Ch3	ELLPQINYLIVASPTIYAMGLCAQAKAIIIDRLQPYWAQYLLKQSEM-QQRRKGLYLLTAGTNFNFTVGTAEITLKYLNNEVE	158
Cd3	KMFDDSKGTIVISPVHFDISAKLKTVDRTQAIYASKYVLNKPISDRSKKRIGMYIAGGS--KSYDSQFKGGQIVMDFFFKSI	163
Ch2	RAMAEDADYIIGAPVYGNHLPQGVKLVFLDRLSG--LIHKVSYDGASLRSQSLRPHKKRRVFI FAVAAGRPESCDGVLRYLKFV	153
Xf	QKMEADGIIYMPVMHAYGTNSCFQAFLEERSG-----YGYFRPMGRPLKDKVASVIVIG-----BRYGHTAVYSQIVLNMILN	140

Ca1	EVLVVGSTYWNM-VYG-----KDVGD-----VILDEEGIKMNRNLGENMAWLMKKIRNV-----	190
Msb3	EMFLVGSYWNM-AYG-----NAIGE-----VEQDQEGIDNMKNLGNMAWLLKKTIHNS-----	192
Ca2	EMMPTSNYWNV-IHG-----TKPGE-----ATQDEEGQIMRVLGKNMAWLLKTI DSSKNN-VKENIEEIKIFTNFIR	211
Gs	EMLVPTSNYWNV-IHG-----TKPGE-----ATQDEEGQIMRVLGKNMAWLLKLVHEHGKA-VTPPERETKTYMSFIR	211
Ch1	EMFIPVSNYWNV-IHG-----TKPGE-----ALQDEEGQIMRVLGKNMAYLLKLEAKGKGE-IEAPEKESKYVTNFIR	210
De1	QMIIVPGSIYWNM-TLG-----GAPGD-----IEKDEEGKNTFIQLGKNMADLLKLN-----	190
Msb4	QMIIVPGSIYWNM-GMG-----LAEGE-----VEKDEEGIQTMQTLGQNLAWLLKLN-----	191
Dv	QMLVPGSIYWNL-GYG-----RDKGE-----VESDAEAMAMNDHLGKVIDWLGKATAPHRDA-YPLGVRPPAE-----	205
Eh	QMITIGSSYWNM-GYNPN-----VTNQHE-----VEKDEEAKNTMKNLAENMAFVLRKRVRTNWN-SIYYIKF-----	206
Ca3	NMPIVSSQYWNM-VHG-----NTPEE-----VKQDEEGQIMRVLGKNMAWLLKSIQAGKEAGVKLPEKEQIIRTNFIR	208
Cd2	NMPIVSSQYWNQ-VHG-----FTPED-----VMKDEEGLQTMRTLGENMAWLLKCSISAGKAGIKEPQYEETIMTNFIK	206
De2	FKVKAHELVLVT-GYA-----AREGD-----AKDNAELISLARDMGQTLAATLKGAS-----	186
Mst	GMIIVVGDN--S-HFGGI-----TWNP-----AEEDTVGMQTVSETAKKLCVLELQKNRDK-----	191
Mm1	GMIIVVGDN--A-HFGGI-----TWNP-----SEEDSVGMQTVFETAKKLCVLELVHKKK-----	189
Msb1	GMIIVVGDN--S-HFGGI-----AWNPN-----VEEDSIGMQTVSETAKKLCVLELQKNNKR-----	190
Mm2	EMIPVGGGSFGA-NLGGT-----FWSKDKGAEGVSESEGMRSRLRRLTKLIIQTAQLVCRADASEEPQTPAPEISGEAHKQ	228
Msb2	EMIPVGGGSFGA-NLGGT-----FWSKDKGEEGVSESEGMRSRLRRLTKLIIQTTQLVKNASGSETSKPQSETLKEBEGK-	227
Msb5	GQVS-----G-----FMQ-----AEVKDVL-----	155
Cd4	GGYVCGSLSIKT-GFNQNPITEKVESKIQKVSKKFYICYIEKKNPVFLSQIYNFIAINAVLKPMAPKDIIEQYKGIIDRWEEQGI	225
Dh	GLHIVDARNPNTGIFTIGVVGAMMSGFKEDKAVWRRSGADELVEQATSRLQGRNMVALIQQLSGASKMT-----	212
Ch3	EFRIVG-----FKNLEEKGEIKKDPNNFNRIIEAINQMGVS-----	194
Cd3	NTKLLHNIYLN-----NTDD-----VSYIDNQEFKDIYKLNKIYDSI-----	202
Ch2	RPELNGGEVFLAVTG-----VGAKGQIG-FGPEELEKMLIKYNYDNPPEEKIKEYLDLHQYLNLAAREFGQKILTL---	223
Xf	KMILAGSG-FPVLFMG-----MLGQ-----ASEDIEALEALEETLNRMISISYRLRVGLFMEAPAYA-----	196

program (13), using a neighbor-joining algorithm with gamma distance estimation ($\alpha = 2$). A phylogenetic tree was constructed based on pairwise estimates of the expected number of amino acid replacements per site. A total of 1,500 bootstrap cycles were performed, and the values reported below are percentages of times that proteins grouped together on a branch.

Amplification, cloning, and expression of genes encoding Af3, Cd1, and Mj2.

The genes encoding Af3 and Cd1 were amplified by performing PCR with *A. fulgidus* and *C. difficile* genomic DNAs, respectively. The gene encoding Mj2 was amplified by performing PCR with *M. jannaschii* clone AMJAI79 DNA. PCR primers were designed to introduce an *NdeI* site at the start codon and a *BamHI* site (or an *EcoRI* site for the gene encoding Cd1) downstream of the stop codon of each gene. The sequences of the primers used are as follows: (restriction sites are underlined): 5'-CGATCGCATTCCATATGAACTGCTGGCAATA-3' (5' primer) and 5'-GGTGAAGGATCCGGTGGTTTAGGACTCTC-3' (3' primer) for the gene encoding Af3, 5'-CATCGACCATATGGCAAGATATTAGGT-3' (5' primer) and 5'-CCATCGAGAATTCGCAATAATCTATTACATCC-3' (3' primer) for the gene encoding Cd1, and 5'-CCCCCGGGTCCCATATGAAAGTTATAGGGATAAGTGG-3' (5' primer) and 5'-CCGTAGGATCCAAAGAGAAGGTAATAA-3' (3' primer) for the gene encoding Mj2. The PCR products were purified from agarose gels and digested with *NdeI* and *BamHI* (or *NdeI* and *EcoRI* for the gene encoding Cd1). The resulting DNA fragments were cloned into the *NdeI* and *BamHI* sites (or *NdeI* and *EcoRI* sites for the gene encoding Cd1) of pTZ7-7 to obtain recombinant plasmids pTZ993 (the gene encoding Af3), pFC991 (the gene encoding Cd1), and pTZ983 (the gene encoding Mj2). To confirm the gene sequences in the plasmids, DNA sequencing was done by the automated dideoxy method at the nucleic acid facility of the Pennsylvania State University Biotechnology Institute. To express the genes encoding Cd1 and Mj2, pFC991 and pTZ982, respectively, were used to transform *E. coli* BL21(DE3), and the transformed *E. coli* cells were grown at 37°C in Luria-Bertani broth containing 100 mg of ampicillin per liter. To express the gene encoding Af3, both pTZ993 and pSJS1240 (a plasmid expressing some rare *E. coli* tRNAs with spectinomycin resistance) were transformed into *E. coli* BL21(DE3), and the transformed *E. coli* cells were grown at 37°C in Luria-Bertani broth containing 100 mg of ampicillin per liter and 50 mg of spectinomycin per liter. The transformed *E. coli* cells were grown to an absorbance at 600 nm (A_{600}) of 0.6 to 0.8; then the cultures were cooled (for approximately 10 min) to approximately 25°C and induced by adding 1% (wt/vol) (final concentration) lactose. Ferric ammonium citrate (1 g/liter) was also added at the same time to supply iron to the cultures. After 6 h of growth at 25°C, the cells were harvested by centrifugation at $5,322 \times g$ for 10 min at 4°C. Cell pellets were stored at -80°C.

Purification of Af3, Cd1, and Mj2 overproduced in *E. coli*. All steps in the purification procedure were performed anaerobically. Where appropriate, a COY anaerobic chamber (Coy Laboratory Products, Ann Arbor, Mich.) was employed. All steps were performed at 21°C except as indicated below. Cells (approximately 15 g [wet weight]) were resuspended in 90 ml of buffer (50 mM HEPES [pH 7.0] for Af3, 50 mM morpholineethanesulfonic acid [MES] [pH 6.2] for Cd1, and 50 mM Tris [pH 7.6] for Mj2) and lysed by two passages through a French press at 20,000 lb/in² (138 MPa). The lysates were centrifuged at $74,400 \times g$ for 20 min at 4°C. For Af3 and Mj2, the supernatants were heated at 65°C for 30 min anaerobically. For Cd1, streptomycin sulfate (final concentration, 1% [wt/vol]) was added to the supernatant. The treated supernatants were centrifuged as described above. The supernatants resulting from the second centrifugation were filtered (pore size, 0.45 μ m) and loaded on 5-ml HiTrap columns (HiTrap SP for Af3 and Cd1 and HiTrap Q for Mj2). A 50-ml linear 0 to 1.0 M NaCl gradient was applied at a rate of 5 ml/min. Fractions containing the Isf homologs (as determined by a yellow-brown color and sodium dodecyl

sulfate-polyacrylamide gel electrophoresis [SDS-PAGE] [see Fig. 4]) were pooled. The pooled fractions were then adjusted so that the $(\text{NH}_4)_2\text{SO}_4$ concentration in the resuspension buffer was 0.6 M by addition of 3.0 M $(\text{NH}_4)_2\text{SO}_4$. The solutions were then passed over a custom-made 9-ml phenyl Sepharose high-performance column that had been preequilibrated with 0.6 M $(\text{NH}_4)_2\text{SO}_4$ in the resuspension buffer. The flowthrough fractions containing Isf homologs were diluted 1:10 with 50 mM Tris (pH 7.6 for Mj2 and pH 8.6 for Af3 and Cd1), filtered (pore size, 0.2 μ m), and loaded onto a Mono-Q 10/10 column that had been preequilibrated with the dilution buffer. A 50-ml 0 to 1.0 M NaCl gradient was applied at a rate of 2 ml/min. The fractions containing the desired proteins were pooled, and their purity was analyzed by SDS-PAGE and Coomassie blue staining.

Reconstitution of Af3, Cd1, and Mj2. The reconstitution method used was a modified version of the method described previously (16). To 100 ml of anaerobic 50 mM Tris buffer (pH 7.6 for Mj2 and pH 8.6 for Af3 and Cd1), 800 μ l of β -mercaptoethanol was added dropwise while the preparation was gently stirred. After 10 min, 2.5 to 10 mg of pure protein (eluted from the Mono-Q column) was added. Then 300 μ l of 60 mM ferric chloride, 300 μ l of 60 mM sodium sulfide, and 300 μ l of 15 mM FMN (all in 50 mM Tris) were added stepwise at 10-min intervals. The mixtures were incubated anaerobically at 4°C overnight and then concentrated by using an ultrafiltration apparatus (Amicon stirred cell 8200 and YM-10 membrane) under an N₂ atmosphere at 4°C. The excess Fe, S, and FMN were removed by several ultrafiltration cycles and/or by passage through a desalting column.

Analyses. UV-visible absorption spectra and absorbance were determined with a Beckman DU 600 spectrophotometer. The molecular masses of Af3, Cd1, and Mj2 subunits were estimated by SDS-15% PAGE, using a low-molecular-weight protein standard obtained from Sigma (14). The molecular masses of the native proteins were based on elution from a Superose-12 gel filtration column as described previously (14). The following proteins with known molecular masses were used as standards: bovine serum albumin (66 kDa), ovalbumin (45 kDa), carbonic anhydrase (29 kDa), and chymotrypsinogen (25 kDa) or horse heart cytochrome *c* (12.4 kDa). The flavins were extracted from the purified Af3, Cd1, and Mj2 (without reconstitution) by adding trichloroacetic acid (final concentration, 5% [vol/vol]) to the protein solutions. The precipitated proteins were removed by centrifugation at 4°C and $13,500 \times g$ for 15 min. The supernatants containing the flavins were adjusted to pH 6.0 with ammonium acetate, and the flavins were identified as FMN by high-performance liquid chromatography as described previously (14). The quantity of extracted FMN was determined by measuring the A_{450} and using a standard curve generated from pure FMN. The iron content was determined as described previously (14). The biuret method was used to determine the protein concentration; bovine serum albumin was used as the standard. N-terminal sequences were determined with a model 477A protein sequencer (PE Biosystems, Inc., Foster City, Calif.) at the Macro Core Facility of the Pennsylvania State University Hershey Medical Center.

Reduction of the homologs and enzyme activity assays. Cell extracts were made anaerobically as follows. Cells were resuspended in 50 mM Tris (pH 7.6) and then lysed by one passage through a French press at 20,000 lb/in². The lysates were centrifuged at $30,000$ to $40,000 \times g$ for 30 min at 4°C. The supernatants (cell extracts) were divided into aliquots and stored in liquid nitrogen. The reduction experiments were performed at 21°C. In each reduction experiment, 180 μ g of cell extract protein (or 250 μ g of protein in the case of the *M. jannaschii* cell extract) in 50 mM Tris (pH 7.6) was equilibrated at 21°C for 10 min in a stoppered 1.0-ml cuvette. Each reduction reaction was initiated by adding 0.2 mg of Isf or an Isf homolog protein (or 0.125 mg of protein for reduction of Mj2 by *M. jannaschii* cell extract). The total volume of the reduction mixture was 0.5 ml. The UV-visible spectra during reduction of the proteins were recorded every 20 s

FIG. 1. Amino acid sequence alignment for 24 Isf homologs from diverse microorganisms not previously reported (15). The sequence of *M. thermophila* Isf is included for reference. The sequences included exhibit more than 20% identity with the *M. thermophila* Isf sequence (191 residues) and bear the conserved cysteine motif. Abbreviations: Ca1, Ca2, and Ca3, ORFs from a *C. acetobutylicum* unfinished fragment of the complete genome; Cd2, Cd3, and Cd4, ORFs from *C. difficile* unfinished fragments of the complete genome; Ch1, Ch2, and Ch3, ORFs from *Carboxydotherrnus hydrogeniformans* unfinished fragments of the complete genome; De1 and De2, ORFs from two *Dehalococcoides ethenogenes* unfinished fragments of the complete genome; Dh, ORF from preliminary sequence data for the genome of *Desulfotobacterium hafnense* (http://www.jgi.doe.gov/JGI_microbial/html/index.html); Dv, ORF from a *Desulfovibrio vulgaris* unfinished fragment of the complete genome; Eh, ORF from *E. histolytica* unfinished fragment of the complete genome; Gs, ORF from a *Geobacter sulfurreducens* unfinished fragment of the complete genome; Mm1 and Mm2, ORFs from preliminary sequence data for the genome of *M. maezi* provided by the Göttingen Genomic Laboratory (<http://www.g21.bio.uni-goettingen.de/>); Msb1, Msb2, Msb3, Msb4, and Msb5, ORFs from preliminary sequence data for the genome of *Methanosarcina barkeri*; Mst, *M. thermophila* iron-sulfur flavoprotein (GenBank accession no. U50189); and Xf, *X. fastidiosa* ORF XF1919 (GenBank accession no. AE004011). Residues that are conserved in at least 7 of 10 sequences are shaded. The cysteines in the cysteine motif ligating the 4Fe-4S cluster are indicated by asterisks.

with a Hewlett-Packard 8452A diode array spectrophotometer. The A_{480} was used to monitor the amount of protein reduced after the absorbance was corrected for turbidity. Assays to determine the CODH activities of the cell extracts were conducted at 21°C and pH 7.6 by using a Hewlett-Packard 8452A diode array spectrophotometer as described previously (22). Hydrogenase assays were conducted in the same way as the CODH assays, except that CO was replaced by H_2 .

RESULTS

Analysis of sequences from diverse prokaryotes with identity to Isf. Previous database searches identified 11 homologs of Isf (2, 15). A recent search (15 July 2001) of all available databases with BLASTP and TBLASTN, which included finished and unfinished genomes, was conducted with the entire 191-amino-acid sequence of Isf as the query. Less biased BLAST searches with the Isf sequence minus the conserved region containing the cysteine motif (residues 41 to 61) yielded the same results. The searches identified 24 additional homologs (Fig. 1) that exhibited at least 20% overall identity to Isf and also contained the unique cysteine motif (CX₂CX₂CX₄₋₇C) previously shown to ligate the 4Fe-4S cluster in Isf (15). This motif is conserved in all of the homologs except De3 and Xf, in which either an aspartate or a histidine has replaced the second cysteine in the motif. For all of the homologs the number of residues (155 to 228 residues) and the predicted subunit molecular masses (16.9 to 25.4 kDa) are within narrow ranges. Inspection of the predicted protein sequence immediately upstream of the annotated start of Mbt3 (21) revealed 16 residues with high levels of identity to the conserved N-terminal residues in all other homologs shown in Fig. 1, suggesting that the annotated start of Mbt3 was not correct. In general, the C-terminal halves of the homologs are less similar than the N-terminal halves. The homologs shown in Fig. 1 and those described previously are homologs from physiologically and phylogenetically distinct microbes representing all three domains. The species represented include methane-producing and sulfate-reducing members of the *Archaea* domain. The *Bacteria* domain is represented by gram-positive, green nonsulfur, and purple nonsulfur species. All of the species belonging to the *Bacteria* and *Archaea* domains are anaerobes except for *Xylella fastidiosa*. The only Isf homolog identified in a member of the *Eucarya* domain is a homolog from the intestinal anaerobic pathogen *Entamoeba histolytica*. The previously described homologs Mj1-2 and Af1-3 are from the hyperthermophiles *M. jannaschii* and *A. fulgidus*, which have optimum growth temperatures of 85 and 83°C, respectively (10). Multiple Isf homologs were found in many species (Fig. 1 and 2).

A neighbor-joining phylogenetic tree for the previously described homologs and the homologs in Fig. 1 is shown in Fig. 2. Although the number of homologs is relatively small, there appear to be four subgroups. Bootstrap data support the grouping of subgroups I and II. Closely related sequences are present in physiologically and phylogenetically diverse species, and several strains have homologs belonging to different subgroups, suggesting that they have different functions in the cell. Subgroup I contains homologs that are mostly from the *Archaea* domain, and subgroups III and IV contain mainly homologs from the *Bacteria* domain.

Previously described Isf homologs accounted for the first nine pairwise alignments with the lowest *E* values produced by the BLASTP search. These homologs were immediately fol-

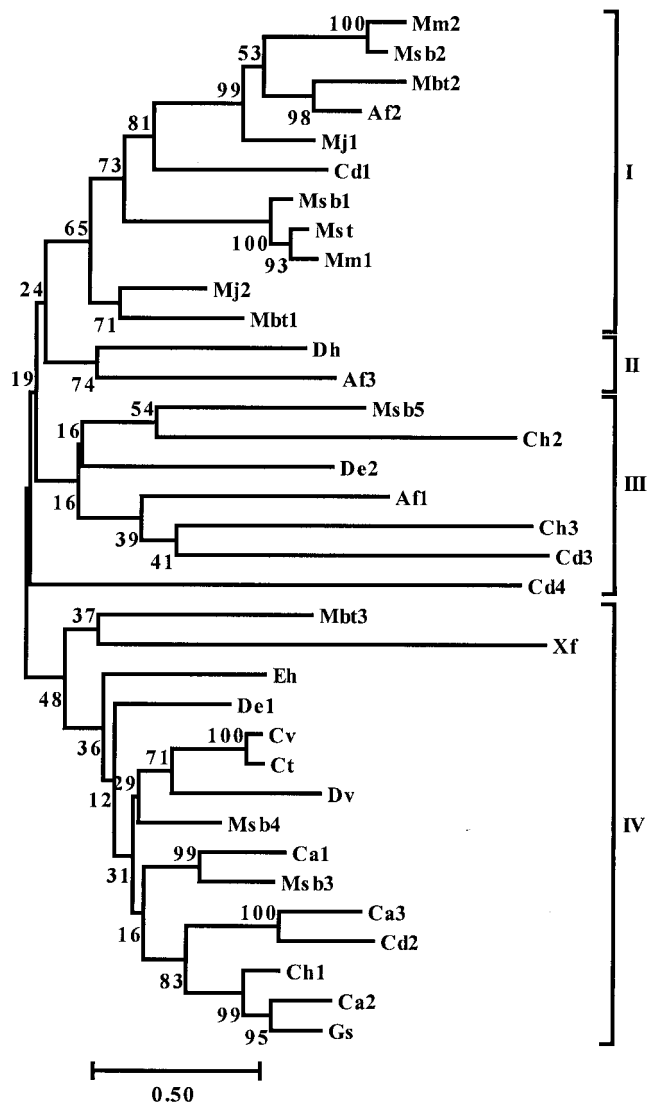


FIG. 2. Phylogenetic tree for 35 Isf homologs. The following homologs were described previously (15) and are not included in Fig. 1: Af1, Af2, and Af3, *A. fulgidus* ORFs AF1436, AF1519, and AF1896 (GenBank accession no. AE001004, AE000997, and AE000972), respectively; Cd1, ORF from a *C. difficile* unfinished fragment of the complete genome; Ct, ORF from a *Chlorobium tepidum* unfinished fragment of the complete genome; Cv, ORF (EMBL accession no. Z83933) in a 15-kb genomic DNA fragment of *Chlorobium vibrioforme*; Mbt1, Mbt2, and Mbt3, *M. thermoautotrophicum* ORFs MTH135, MTH1473, and MTH1595 (GenBank accession no. AE000908 and AE000919), respectively; and Mj1 and Mj2, *M. jannaschii* ORFs MJ0731 and MJ1083 (GenBank accession no. C64391 and B64435), respectively. The tree was constructed based on pairwise distance estimates of the expected number of amino acid replacements per site. Bar = 0.50 amino acid replacement per site. The bootstrap values are indicated to the left of the branches. The four subgroups (subgroups I to IV) are indicated on the right.

lowed by three FMN-binding tryptophan repressor binding proteins, the Isf homolog Xf, two additional tryptophan repressor binding proteins, and then a flavodoxin from *Clostridium acetobutylicum* exhibiting 28% identity with Isf residues 42 to 154. The identity of Isf and the flavodoxin was reported

previously (14). No other alignments with flavodoxins were obtained up to an E value of 8.2. The search also resulted in pairwise alignments with the FMN-containing enzymes 1,4-benzoquinone reductase ($E = 0.002$) and NADH dehydrogenase ($E = 0.70$).

Biochemical characterization of Af3, Cd1, and Mj2. Isf homologs from physiologically and phylogenetically diverse species in the *Archaea* and *Bacteria* domains were selected for production in *E. coli* to determine if they have biochemical properties characteristic of Isf. The homologs chosen were Af3 from the sulfate-reducing archaeon *A. fulgidus*, Mj2 from the CO₂-reducing hyperthermophilic methanoarchaeon *M. jannaschii*, and Cd1 from *C. difficile*, which is an intestinal pathogen classified in the *Bacteria* domain. Cd1 is also one of the sequences in Fig. 1 with the largest number of residues. Sequencing of the expression plasmids containing the genes encoding Af3, Cd1, and Mj2 confirmed that the correct sequences were present. Af3 was produced at low levels in *E. coli*; however, when plasmid pSJS1240 was cotransformed with the expression plasmid, the level of production of Af3 increased more than 10-fold (data not shown). Plasmid pSJS1240 harbors and expresses *argU* and *ileX*, genes that encode the rare *E. coli* tRNA^{arg} (AGA/AGG) and tRNA^{ile} (AUA) codons. All 11 arginines and 5 of 16 isoleucines in Af3 are encoded by the rare codons. Heating *E. coli* cell extracts containing Af3 or Mj2 precipitated most *E. coli* proteins, whereas the recombinant proteins from the hyperthermophiles remained soluble, which facilitated purification of Af3 and Mj2. The proteins produced in *E. coli* were purified to apparent homogeneity, as determined by SDS-PAGE (Fig. 3). Two bands, at 46 and 20 kDa, were observed for Mj2. N-terminal analysis revealed that the first five residues of both bands were the same as the residues deduced from the DNA sequence, suggesting that the two bands correspond to the dimer and monomer forms of Mj2. Without boiling, the amount of the 46-kDa band increased significantly, whereas the amount of the 20-kDa band decreased (data not shown), findings which further support the identity of the bands as dimer and monomer forms of Mj2. The dimer band persisted even when Mj2 was boiled for 20 min in the presence of dithiothreitol, which indicated that Mj2 dimerizes strongly at high temperatures, as would be expected for the hyperthermophilic organism *M. jannaschii*. N-terminal sequencing confirmed the identities of Af3 and Cd1 produced in *E. coli*. The experimental subunit molecular masses agreed with the predicted values deduced from the gene sequences (Table 1). The experimental native molecular masses suggest that each Isf homolog was purified as an α_2 dimer, although each value was slightly higher than the calculated value (Table 1). Each of the proteins purified from *E. coli* contained FMN, as shown by high-performance liquid chromatography analysis of the extracted flavins (data not shown). The UV-visible spectra of the proteins purified from *E. coli* were similar to those of Isf (data not shown), suggesting that iron-sulfur centers were present in addition to FMN. However, the intensity of absorbance centered at 430 nm varied among preparations, indicating that there was incomplete incorporation of the iron-sulfur center. Prior to quantification of iron and FMN, each purified protein was reconstituted with iron, sulfur, and FMN by a procedure similar to that previously described for apo-Isf (15). The reconstituted proteins had UV-

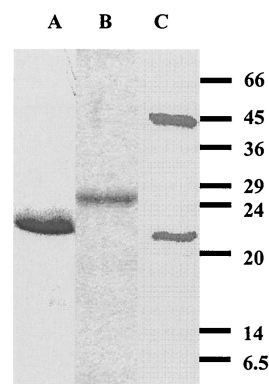


FIG. 3. SDS-PAGE of purified homologs Af3 (lane A), Cd1 (lane B), and Mj2 (lane C). The positions of molecular mass markers (in kilodaltons) are indicated on the right.

visible spectra nearly identical to that of Isf, with absorbance maxima at approximately 484, 452, 430, 378, and 280 nm (Fig. 4). Exposure to air did not alter the spectra, indicating that the reconstituted proteins were in their oxidized states (data not shown). Mild acidification of the reconstituted proteins to destroy the iron-sulfur clusters produced visible spectra with absorbance maxima at approximately 445, 374, and 269 nm typical of flavins (Fig. 4, insets). Analyses of the reconstituted proteins showed that 1.6 to 1.7 mol of FMN and 5.6 to 7.5 mol of nonheme iron were bound to 1 mol of dimer (Table 1), indicating that two FMN molecules and two 4Fe-4S or 3Fe-4S clusters were bound to each dimer. Addition of dithionite decreased the absorbance at 350 to 550 nm (Fig. 4), whereas exposure to air resulted in recovery of the initial spectra, which indicated that Af3, Cd1, and Mj2 are redox-active proteins. In summary, Af3, Cd1, and Mj2 have the distinguishing biochemical characteristics of Isf.

Reconstituted Af3 and Mj2 were reduced by CO in reactions catalyzed by cell extract of acetate-grown *M. thermophila*, but Cd1 was not (Table 2). Reconstituted Af3 and Mj2 were reduced by CO in reactions catalyzed by *A. fulgidus* and *M. jannaschii* cell extracts, respectively (Table 3). *C. difficile* cell extract was not available; however, *C. thermoacetum* cell extract catalyzed CO reduction of Cd1 (Table 3). When H₂ replaced CO as the electron donor, only Isf and Mj2 showed significant reduction (Table 3). Most cell extracts had significant CODH and hydrogenase activities (Table 4); the only exception was *A. fulgidus* cell extract, which had no detectable hydrogenase activity, which is consistent with previous reports (25).

Open reading frames (ORFs) near the genes encoding the Isf homologs in all of the completed genomes (*A. fulgidus*, *M. jannaschii*, *Methanobacterium thermoautotrophicum*, and *X. fastidiosa*) were examined to determine if the ORFs encode redox proteins that could partner with the homologs. An ORF annotated as the shikimate 5-dehydrogenase ORF (*aroE*) is directly adjacent to and downstream of the gene encoding the *M. jannaschii* Mj2 homolog and is transcribed in the same direction (4). Also, an ORF annotated as the ORF encoding a heterodisulfide reductase-related protein (MTH139) is located four ORFs downstream of the gene encoding the *M. thermoautotrophicum* homolog Mbt1, and the two regions are transcribed in the same direction (21). The genes encoding the

TABLE 1. Properties of Isf homologs

Homolog	Predicted isoelectric point	Predicted molecular mass (kDa)	Subunit molecular mass (kDa)	Native molecular mass (kDa)	Amt of FMN (mol/mol of dimers) ^a	Amt of iron (mol/mol of dimers) ^a
Af3	8.9	22	21	51	1.6 ± 0.1	6.6 ± 0.2
Cd1	6.9	24	26	58	1.7 ± 0.1	7.5 ± 0.1
Mj2	6.3	21	20	51	1.6 ± 0.1	5.6 ± 0.2

^a Means ± standard deviations based on three measurements.

A. fulgidus Af2 and *M. jannaschii* Mj1 homologs are flanked by two ORFs annotated as ORFs encoding flavoproteins (AF1518, AF1520, MJ0730, and MJ0732) (Fig. 5) (12, 21). The deduced sequences of the putative proteins encoded by AF1518 and MJ0730 are very similar (50% identity and 73% similarity), as are the sequences of the flavoproteins encoded by AF1520 and MJ0732 (46% identity and 65% similarity). The deduced sequences of Af2 and Mj1 are 56% identical and 74% similar. Thus, it appears that the *A. fulgidus* and *M. jannaschii* genomes have the same flavoprotein-Isf homolog-flavoprotein gene arrangement. An ORF (MJ0728) encoding the CODH catalytic subunit (CooS) in *M. jannaschii* and an unidentified ORF (MJ0729) are located upstream of MJ0730 (Fig. 5); however, the two ORFs immediately upstream of *A. fulgidus* AF1518 exhibit little similarity with MJ0729 and MJ0728, and the homologous CooS protein in *A. fulgidus* is encoded by a remote ORF (AF1849) (12). No ORFs encoding redox proteins were found adjacent to the ORF encoding homolog Xf in *X. fastidiosa*.

DISCUSSION

The results presented here indicate that physiologically and phylogenetically diverse organisms belonging to all three domains contain homologs of Isf from *M. thermophila*, suggesting that this protein is distributed more broadly than previously recognized. Although the databases contained data for aerobes and facultative anaerobes, the BLAST searches primarily identified Isf homologs in organisms capable of anaerobic growth and metabolism. The only exception was the aerobe *X. fastidiosa*. However, these results should not be interpreted to indicate that Isf homologs are found in all anaerobes or are not also widespread in aerobes.

Our results also indicate that the homologs belongs to a new family of flavoproteins. In addition to the overall sequence identity that separates Isf from other flavoproteins, a distinguishing characteristic is the unique cysteine motif which ligates the 4Fe-4S center (15). This motif is conserved in all of the Isf homologs, suggesting that all of these homologs have an iron-sulfur center. The iron contents and UV-visible spectra of homologs Af3, Cd1, and Mj1 support this interpretation, although it cannot be determined from the results presented here if the clusters are of the 4Fe-4S or 3Fe-4S type. Histidine and aspartate replace the second cysteine in the motif of De3 and Xf, suggesting that these residues are ligands for the iron-sulfur cluster in these homologs. A role for histidine and aspartate in ligation of iron-sulfur centers is uncommon but not without precedent (3, 5, 8, 9, 24, 26). Following the previously described Isf homologs, the BLASTP alignments with the next lowest *E* values were the alignments with several FMN-binding tryptophan repressor binding proteins that do not contain iron-

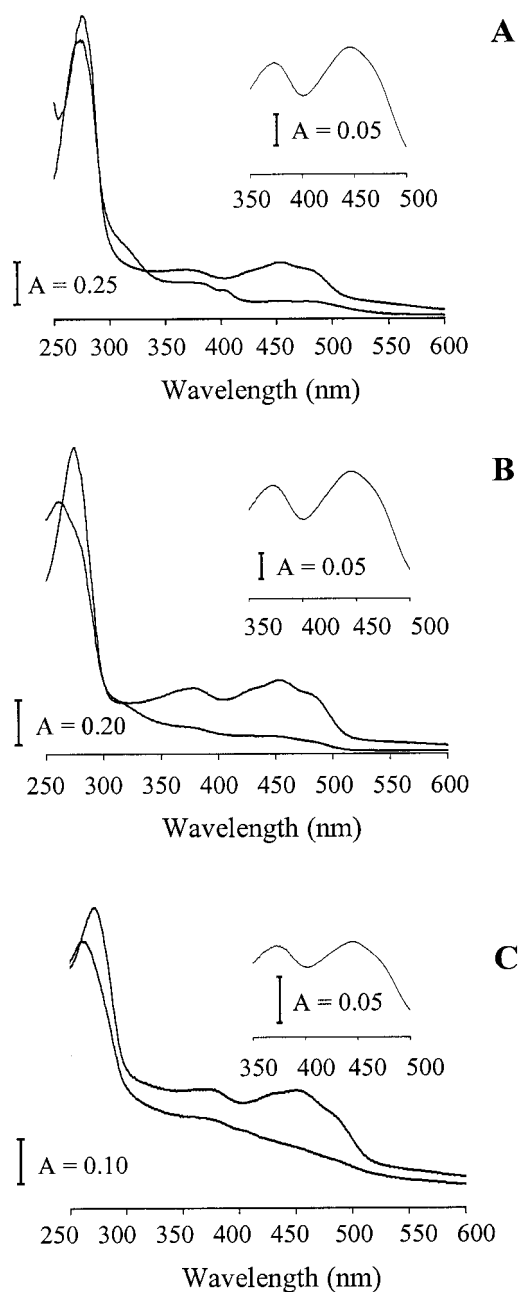


FIG. 4. UV-visible spectra of reconstituted homologs Af3 (A), Cd1 (B), and Mj2 (C). In each panel, the upper and lower traces are the traces obtained before and after reduction by dithionite. The insets show the spectra obtained after mild acidification. The proteins were in 50 mM Tris-Cl (pH 7.6 for Mj2 and pH 8.6 for Af3 and Cd1). The spectra were recorded at room temperature.

TABLE 2. Rates of CO-dependent reduction of Isf and Isf homologs catalyzed by *M. thermophila* cell extract

Protein	Rate of reduction (nmol/min/mg of cell extract protein) ^a	Relative rate (%) ^b
Isf	20 ± 2 (0.37 ± 0.03)	100
Af3	7 ± 1 (0.1 ± 0.02)	35
Cd1	ND ^c	NA ^d
Mj2	9.7 ± 0.8 (0.12 ± 0.01)	49

^a Each rate was determined from the slope of the linear portion of the time course. The values in parentheses are the changes in A_{480} per minute per milligram of cell extract protein. The values are means ± standard deviations based on three determinations.

^b Relative rates were normalized to the rate of reduction of Isf.

^c ND, not detected (below the level of detection, 0.1 nmol/min/mg of cell extract protein).

^d NA, not applicable.

sulfur clusters. Of the two other flavoproteins whose alignments had E values less than 1.0, only NADH dehydrogenase contains iron-sulfur clusters; however, this enzyme has low overall identity with Isf ($E = 0.70$) and does not contain the unique cysteine motif characteristic of Isf homologs. The BLAST searches also produced a pairwise alignment with a flavodoxin, as previously reported (14), which exhibited significant identity with selected regions of Isf; however, no other flavodoxin alignments were obtained, and members of the flavodoxin family do not contain iron-sulfur clusters. Indeed, one of the defining features of the flavodoxin family is the ability to replace ferredoxin under iron-limiting growth conditions when the synthesis of iron-sulfur proteins is compromised. Nonetheless, the pairwise alignments of Isf with flavoproteins belonging to diverse families identify residues in Isf with the potential to bind FMN. Finally, less biased BLAST searches with only the Isf sequences that predominantly aligned with the sequences of flavoproteins (complete Isf sequence minus residues spanning the cysteine motif) produced the same alignments of flavoproteins as searches with the complete Isf sequence produced. These results support the notion that Isf is the prototype of a new family of flavoproteins.

The Isf homologs are relatively small (155 to 228 residues) compared to members of other families of flavoproteins (17, 23). The sizes of most of the Isf homologs are close to the sizes of flavodoxins (ca. 175 residues) that contain no iron-sulfur clusters and one FMN per subunit. Thus, Isf homologs are unusually compact considering that two iron-sulfur clusters and two FMN are bound per dimer. Only the N-terminal regions of the homologs Eh, Mbt1, Msb2, Mm2, and Cd3 deviate from the highly conserved N-terminal region by having

TABLE 4. Enzyme activities in cell extracts

Cell extract	Enzyme activities (μmol of methyl viologen reduced/min/mg of cell extract protein) ^a	
	CODH	Hydrogenase
<i>M. thermophila</i>	2.29 ± 0.07	0.46 ± 0.02
<i>A. fulgidus</i>	0.136 ± 0.003	ND ^b
<i>C. thermoacetikum</i>	1.31 ± 0.04	0.96 ± 0.02
<i>M. jannaschii</i>	0.072 ± 0.005	1.06 ± 0.03

^a Each rate was determined from the slope of the linear portion of the time course. The values are means ± standard deviations based on three measurements. The assays were performed at 21°C.

^b ND, not detected (below the limit of detection, 0.01 μmol of methyl viologen reduced/min/mg of cell extract protein).

an extension consisting of 4 to 10 residues, the function of which is unknown.

The CO-dependent reduction of homologs Af3 and Cd1 by extracts of *A. fulgidus* and *C. thermoacetikum*, respectively, indicates that Isf homologs could potentially have a role in energy-yielding pathways in these nonmethanogenic anaerobes analogous to the role proposed for Isf during acetotrophic growth of *M. thermophila* (2, 14). When the sulfate-reducing organism *A. fulgidus* grows on lactate, a CODH-ACS complex cleaves acetyl-CoA, oxidizes the carbonyl group, and donates electrons to ferredoxin (6); thus, it is hypothesized that Af3 or another Isf homolog in *A. fulgidus* accepts electrons from the ferredoxin, thereby participating in the electron transport chain for reduction of sulfate to sulfide. The strictly anaerobic homoacetogen *C. thermoacetikum* belonging to the *Bacteria* domain utilizes the energy-yielding Wood-Ljungdahl pathway. Central to this pathway is CODH-ACS, which synthesizes acetyl-CoA from a methyl group, CoA, and CO obtained by reduction of CO₂ catalyzed by CODH-ACS. Analogous to the situation in *M. thermophila* and *A. fulgidus*, ferredoxin is the proposed electron donor for CODH-ACS in *C. thermoacetikum* (18, 19). Thus, it is hypothesized that an Isf homolog similar to Cd1 in *C. difficile* is present in *C. thermoacetikum* and that this homolog donates electrons to the ferredoxin. Contrary to this hypothesis, extracts of *C. thermoacetikum* were unable to catalyze H₂-dependent reduction of the Isf homolog Cd1, although ferredoxins are electron acceptors for the hydrogenase of *C. thermoacetikum*. Clearly, additional experiments are necessary to test the hypotheses described above.

BLAST searches of the unfinished *C. difficile* genome revealed two ORFs which encode proteins that exhibit 45 and 37% identity to the α subunit (contig 939, nucleotides 49244 to

TABLE 3. Rates of CO- or H₂-dependent reduction of Isf and Isf homologs catalyzed by cell extracts

Protein	Cell extract	Rates of reduction with the following electron donors (nmol/min/mg of cell extract protein) ^a	
		CO	H ₂
Isf	<i>M. thermophila</i>	20 ± 2 (0.37 ± 0.03)	2.7 ± 0.5 (0.05 ± 0.01)
Af3	<i>A. fulgidus</i>	12 ± 2 (0.18 ± 0.03)	ND ^b
Cd1	<i>C. thermoacetikum</i>	6.2 ± 0.6 (0.11 ± 0.01)	ND
Mj2	<i>M. jannaschii</i>	1.2 ± 0.2 (0.024 ± 0.003)	0.4 ± 0.1 (0.007 ± 0.001)

^a Each value was determined from the slope of the linear portion of the time course. The values in parentheses are the changes in A_{480} per minute per milligram of cell extract protein. The values are means ± standard deviations based on three determinations.

^b ND, not detected (below the limit of detection, 0.1 nmol/min/mg of cell extract protein).

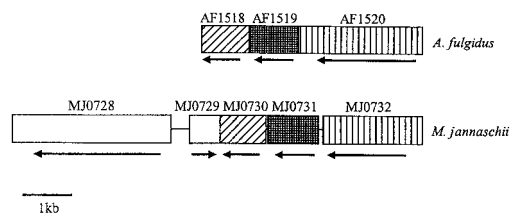


FIG. 5. Flavoprotein-Isf homolog-flavoprotein gene arrangements in the genomes of *A. fulgidus* and *M. jannaschii*. ORFs that encode homologous proteins are aligned and are indicated by the same pattern. The transcription direction for each ORF is indicated by the arrow below the ORF. AF1519 and MJ0731 are the ORFs that encode Af2 and Mj1, respectively. AF1518, AF1520, MJ0730, and MJ0732 encode flavoproteins. MJ0728 encodes the catalytic subunit (CooS) of CODH. MJ0729 is an undefined ORF.

51367) and β subunit (contig 939, nucleotides 35452 to 37368), respectively, of the CODH-ACS complex of *C. thermoacetum*, in addition to two ORFs that encode proteins homologous to *C. thermoacetum* ferredoxin I (data not shown). These results are consistent with a role for Cd1 in electron transport coupled to CODH-ACS in *C. difficile*, as shown previously for *M. thermophila*.

The CO-dependent reduction of Mj2 catalyzed by extracts of *M. jannaschii* suggests that the Isf homologs are components of electron transport chains coupled to CODH in this methanarchaeon. Indeed, the annotated genome of *M. jannaschii* contains the CODH homologs CdhA from *M. thermophila* and CooS from *Rhodospirillum rubrum* (4). CdhA from *M. thermophila* catalyzes reversible oxidation of CO to CO₂ and is a subunit of the CODH-ACS complex. *M. jannaschii* oxidizes H₂ and reduces CO₂ to methane for growth and, therefore, does not utilize the CODH-ACS complex for energy generation; however, the CODH-ACS complex is required by this autotroph for synthesis of acetyl-CoA from 2CO₂ for cell carbon (20). The ability of *M. jannaschii* extracts to link CO oxidation and H₂ oxidation to reduction of Mj2 is consistent with a role for Isf homologs in an electron transport chain required for H₂-dependent reduction of CO₂ to CO for incorporation into acetyl-CoA. Thus, it is hypothesized that Mj2 or another Isf homolog in *M. jannaschii* reduces a ferredoxin that then donates electrons to the CdhA homolog for reduction of CO₂ to CO, analogous to the reverse role of Isf in *M. thermophila* during growth with acetate. The Isf homologs identified in the autotrophic CO₂-reducing methanarchaeon *M. thermoautotrophicum* may also have a role analogous to the role postulated for acetyl-CoA synthesis in *M. jannaschii*. During growth of *R. rubrum* in the dark, CooS oxidizes CO, which is coupled to the reduction of protons to H₂ and generation of ATP (11). The role of the CooS homolog in *M. jannaschii* is unknown; however, the low level of overall identity between CooS and CdhA in *M. jannaschii* (less than 7%) suggests that these proteins have different functions. The close proximity of the genes encoding Mj1 and CooS in *M. jannaschii* is consistent with a role for Mj1 in an electron transport chain coupled to CooS that may involve the flavoproteins flanking Mj1. Further research is needed to test the hypotheses described above.

Although the results presented here lead to hypothesized roles in electron transport for the Isf homologs, the electron

donors and acceptors are unknown, except that ferredoxin is the electron donor for Isf in *M. thermophila*. The genes encoding putative flavoproteins flank the genes encoding Mj1 and Af2, which is consistent with the hypothesis that these Isf homologs interact with the flavoproteins. The same flavoprotein-Isf homolog-flavoprotein gene arrangement, together with the high level of identity between the proteins encoded by these genes in *A. fulgidus* and *M. jannaschii*, is consistent with this hypothesis. The Isf homologs in Fig. 1 exhibit less identity in their C-terminal halves, which may reflect adaptations for interactions with different redox partners or cofactors. The results obtained for Cd1 of *C. difficile* are consistent with this interpretation. Although *C. thermoacetum* extract catalyzed robust CO-dependent reduction of Cd1, no CO-dependent reduction of Cd1 by *M. thermophila* cell extract was detected. Cd1 is one of the largest Isf homologs identified and has an extended C-terminal region compared to the C-terminal regions of Isf, Af3, and Mj1. The most straightforward interpretation of these results is that Cd1 is unable to interact with the ferredoxin that donates electrons to Isf in *M. thermophila* (14). Most species in which Isf homologs were found contain more than one homolog, which not only reflects the importance of this iron-sulfur flavoprotein family in physiology but also underscores the potential for diverse functions. Finally, the levels of identity for some Isf homologs from the same organism are lower than the levels of identity for homologs from different organisms; for example, Af2 exhibits 76% identity with Mbt2 but only 29 and 30% identity with Af3 and Af1, respectively, which indicates that homologs in the same organism may have different functions.

The same flavoprotein-Isf homolog-flavoprotein gene arrangement and the high level of deduced sequence identity of the flavoproteins of *A. fulgidus* and *M. jannaschii* are consistent with lateral gene transfer. On the other hand, the G+C content of each gene in an arrangement is similar to the total G+C content of the genome to which it belongs (49 to 50% G+C for the three genes from *A. fulgidus* and 33 to 38% G+C for the genes from *M. jannaschii*), suggesting that if lateral gene transfer occurred, the event was not recent.

ACKNOWLEDGMENTS

We thank Nancy L. Scott and Juliette T. Lecomte of the Chemistry Department of the Pennsylvania State University for their kindness and help with the high-performance liquid chromatography analysis of the flavins extracted from Af3, Cd1, and Mj2. We also especially thank Rebecca D. Miles for critically reading the manuscript. We thank the Göttingen Genomic Laboratory for providing the preliminary genome data for *M. maei*.

REFERENCES

- Andresen, J. R., A. Schaupp, C. Neurauter, A. Brown, and L. G. Ljungdahl. 1973. Fermentation of glucose, fructose, and xylose by *Clostridium thermoacetum*: effect of metals on growth yield, enzyme, and the synthesis of acetate from CO₂. *J. Bacteriol.* **114**:743-751.
- Becker, D. F., U. Leartsakulpanich, K. K. Surerus, J. G. Ferry, and S. W. Ragsdale. 1998. Electrochemical and spectroscopic properties of the iron-sulfur flavoprotein from *Methanosarcina thermophila*. *J. Biol. Chem.* **273**:26462-26469.
- Britt, R. D., K. Sauer, M. P. Klein, D. B. Knaff, A. Kriauciunas, C. A. Yu, L. Yu, and R. Malkin. 1991. Electron spin echo envelope modulation spectroscopy supports the suggested coordination of two histidine ligands to the Rieske Fe-S centers of the cytochrome *b₆f* complex of spinach and the cytochrome *bc₁* complex of *Rhodospirillum rubrum*, *Rhodobacter sphaeroides* R-26, and bovine heart mitochondria. *Biochemistry* **30**:1891-1901.
- 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, and J. C. Venter. 1996. Complete genome sequence of the methanogenic archaeon *Methanococcus jannaschii*. *Science* **273**:1058–1073.
5. Busch, J. L., J. L. Breton, B. M. Bartlett, F. A. Armstrong, R. James, and A. J. Thomson. 1997. [3Fe-4S]↔[4Fe-4S] cluster interconversion in *Desulfovibrio africanus* ferredoxin III: properties of an Asp14→Cys mutant. *Biochem. J.* **323**:95–102.
 6. Dai, Y.-R., D. W. Reed, J. H. Millstein, P. L. Hartzell, D. A. Grahame, and E. DeMoll. 1998. Acetyl-CoA decarbonylase/synthase complex from *Archaeoglobus fulgidus*. *Arch. Microbiol.* **169**:525–529.
 7. Ferry, J. G. 1997. Enzymology of the fermentation of acetate to methane by *Methanosarcina thermophila*. *BioFactors* **6**:25–35.
 8. Gurbiel, R. J., C. J. Batie, M. Sivaraja, A. E. True, J. A. Fee, B. M. Hoffman, and D. P. Ballou. 1989. Electron-nuclear double resonance spectroscopy of ¹⁵N-enriched phthalate dioxygenase from *Pseudomonas cepacia* proves that two histidines are coordinated to the [2Fe-2S] Rieske-type clusters. *Biochemistry* **28**:4861–4871.
 9. Gurbiel, R. J., T. Ohnishi, D. E. Robertson, F. Daldal, and B. M. Hoffman. 1991. Q-band ENDOR spectra of the Rieske protein from *Rhodobacter capsulatus* ubiquinol-cytochrome c oxidoreductase show two histidines coordinated to the [2Fe-2S] cluster. *Biochemistry* **30**:11579–11584.
 10. Horikoshi, K., and W. D. Grant (ed.). 1998. *Extremophiles: microbial life in extreme environments*. John Wiley & Sons, Inc., New York, N.Y.
 11. Kerby, R. L., P. W. Ludden, and G. P. Roberts. 1995. Carbon monoxide-dependent growth of *Rhodospirillum rubrum*. *J. Bacteriol.* **177**:2241–2244.
 12. 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, J. C. Venter, et al. 1997. The complete genome sequence of the hyperthermophilic, sulphate-reducing archaeon *Archaeoglobus fulgidus*. *Nature* **390**:364–370.
 13. Kumar, S., K. Tamura, and M. Nei. 1994. MEGA: molecular evolutionary genetic analysis. *Comput. Appl. Biosci.* **10**:189–191.
 14. Latimer, M. T., M. H. Painter, and J. G. Ferry. 1996. Characterization of an iron-sulfur flavoprotein from *Methanosarcina thermophila*. *J. Biol. Chem.* **271**:24023–24028.
 15. Leartsakulpanich, U., M. L. Antonkine, and J. G. Ferry. 2000. Site-specific mutational analysis of a novel cysteine motif proposed to ligate the 4Fe-4S cluster in the iron-sulfur flavoprotein of the thermophilic methanoarchaeon *Methanosarcina thermophila*. *J. Bacteriol.* **182**:5309–5316.
 16. Li, N., J. D. Zhao, P. V. Warren, J. T. Warden, D. A. Bryant, and J. H. Golbeck. 1991. PsaD is required for the stable binding of PsaC to the photosystem I core protein of *Synechococcus* sp. PCC 6301. *Biochemistry* **30**:7863–7872.
 17. Lim, L. W., N. Shamala, F. S. Mathews, D. J. Steenkamp, R. Hamlin, and N. H. Xuong. 1986. Three dimensional structure of the iron-sulfur flavoprotein trimethylamine dehydrogenase at 2.4-Å resolution. *J. Biol. Chem.* **261**:15140–15146.
 18. Ragsdale, S. W., J. E. Clark, L. G. Ljungdahl, L. L. Lundie, and H. L. Drake. 1983. Properties of purified carbon monoxide dehydrogenase from *Clostridium thermoaceticum*, a nickel, iron-sulfur protein. *J. Biol. Chem.* **258**:2364–2369.
 19. Ragsdale, S. W., L. G. Ljungdahl, and D. V. DerVartanian. 1983. ¹³C and ⁶¹Ni isotope substitutions confirm the presence of a nickel (III)-carbon species in acetogenic CO dehydrogenase. *Biochem. Biophys. Res. Commun.* **115**:658–665.
 20. Selkov, E., N. Maltsev, G. J. Olsen, R. Overbeek, and W. B. Whitman. 1997. A reconstruction of the metabolism of *Methanococcus jannaschii* from sequence data. *Gene* **197**:11–26.
 21. 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, and J. N. Reeve. 1997. Complete genome sequence of *Methanobacterium thermoautotrophicum* deltaH: functional analysis and comparative genomics. *J. Bacteriol.* **179**:7135–7155.
 22. Terlesky, K. C., M. J. K. Nelson, and J. G. Ferry. 1986. Isolation of an enzyme complex with carbon monoxide dehydrogenase activity containing corrinoid and nickel from acetate-grown *Methanosarcina thermophila*. *J. Bacteriol.* **168**:1053–1058.
 23. Vanoni, M. A., and B. Curti. 1999. Glutamate synthase: a complex iron-sulfur flavoprotein. *Cell. Mol. Life Sci.* **55**:617–638.
 24. Volbeda, A., M. H. Charon, C. Piras, E. C. Hatchikian, M. Frey, and J. C. Fontecilla-Camps. 1995. Crystal structure of the nickel-iron hydrogenase from *Desulfovibrio gigas*. *Nature* **373**:580–587.
 25. Vorholt, J., J. Kunow, K. O. Stetter, and R. K. Thauer. 1995. Enzymes and coenzymes of the carbon monoxide dehydrogenase pathway for autotrophic CO₂ fixation in *Archaeoglobus lithotrophicus* and the lack of carbon monoxide dehydrogenase in the heterotrophic *A. profundus*. *Arch. Microbiol.* **163**:112–118.
 26. Zhou, Z. H., and M. W. Adams. 1997. Site-directed mutations of the 4Fe-ferredoxin from the hyperthermophilic archaeon *Pyrococcus furiosus*: role of the cluster-coordinating aspartate in physiological electron transfer reactions. *Biochemistry* **36**:10892–10900.