Thermostable NAD$^+$-Dependent Alcohol Dehydrogenase from *Sulfolobus solfataricus*: Gene and Protein Sequence Determination and Relationship to Other Alcohol Dehydrogenases†

Sergio Ammendola,‡ Carlo A. Raia,§ Carla Caruso,§ Laura Camardella,§ Sabato D'Auria,§ Mario De Rosa,∥ and Mosè Rossi*,§

Istituto di Biochimica delle Proteine ed Enzimologia, CNR, Napoli, Italy, Menarini Ricerche Sud, Dipartimento di Biotecnologie, Pomezia, Italy, and Dipartimento di Chimica Organica e Biologica, Università di Napoli, Napoli, Italy

Received June 8, 1992; Revised Manuscript Received September 21, 1992

ABSTRACT: The NAD$^+$-dependent alcohol dehydrogenase (EC 1.1.1.1) from the thermoacidophilic archaebacterium *Sulfolobus solfataricus*, DSM1617 strain (SSADH), has been purified and characterized. Its gene has been isolated by screening two *S. Solfataricus* genomic libraries using oligonucleotide probes. The encoding sequence consists of 1041 base pairs, and it shows a high preference for codons ending in T or A. The primary structure, determined by peptide and gene analysis, consists of 347 amino acid residues, yielding a molecular weight of 37 588. A level of identity of 24–25% was found with the amino acid sequences of horse liver, yeast, and *Thermoanaerobium brockii* alcohol dehydrogenases. The coenzyme-binding and catalytic and structural zinc-binding residues typical of eukaryotic alcohol dehydrogenases were found in SSADH with the difference that one out of the four structural zinc-binding Cys residues is substituted by Glu. The protein contains four zinc atoms per dimer, two of which are removed by chelating agents with a concomitant loss of structural stability.

Alcohol dehydrogenases are widely distributed enzymes and have been found in many microorganisms and in fungal, plant, and animal cells (Sund & Theorell, 1963). The most thoroughly characterized alcohol dehydrogenases are the yeast and horse liver enzymes (Brändén et al., 1975) to which all the others subsequently studied have been related.

Alcohol dehydrogenases have been subdivided into different families according to structure: the short-chain family, which groups non-metalloenzymes with subunits of about 250 residues [reviewed in Persson et al. (1991)], the medium-chain family with subunits of 350–375 residues and, often, containing zinc [reviewed in Jörnvall et al. (1987)], and the long-chain family with over 700 residues (Inoue et al., 1989; Goodlove et al., 1989). Alcohol dehydrogenases lacking zinc have been described (Jeffery et al., 1981; Jörnvall et al., 1981, 1984; Villaroya et al., 1989) as well as alcohol dehydrogenases requiring iron for activation (Scopes, 1983; Neale et al., 1986; Williamson & Paquin, 1987). Zinc-dependent ADHs[1] have a dimeric or tetrameric structure. Dimeric enzymes are the alcohol dehydrogenases from mammals [reviewed in Jörnvall et al. (1987)] and from higher plants (Llewellyn et al., 1987). Tetrameric enzymes are the alcohol dehydrogenases from yeasts (Jörnvall, 1977; McKnight et al., 1985; Young & Pilgrim, 1985) and from bacteria (Jendrossek et al., 1988; Inoue et al., 1989; Peretz & Burstein, 1989). Mammalian ADHs have been subdivided into classes according to the electrophoretic and kinetic properties and amino acid sequences of their isoforms (Kaiser et al., 1989; Estonius et al., 1990).

Among the thermophilic bacteria, ADHs from *Bacillus stearothermophilus* (Sheehan et al., 1988), *Thermoanaerobacter ethanolicus* (Bryant et al., 1988), and *Thermoanaerobium brockii* (Peretz & Burstein, 1989) have been well characterized. The latter enzyme is the first prokaryotic and NADP-linked thermophilic ADH of which the complete amino acid sequence has been determined. Furthermore, a novel ADH from the thermoacidophilic archaebacterium *Sulfolobus solfataricus* strain MT4 was purified and characterized (Rella et al., 1987a). The enzyme is a dimeric, NAD$^+$-dependent alcohol/aldehyde oxidoreductase, with a stereoselective preference to produce (*S*)-alcohol from prochiral ketones (Rella et al., 1987b; Trincone et al., 1990); from this point of view, it is more similar to horse liver ADH than to yeast and *T. brockii* ADHs. So far, no other alcohol dehydrogenases have been described from Archaebacteria, one of the three primary kingdoms which, together with Eukaryotes and Eubacteria, include all living organisms (Woese et al., 1990). The analysis of a large number of ADH encoding genes from different sources has shown that eukaryotic ADH genes have intron/exon structures which, together with gene polymorphism, reflect the evolutionary history of dehydrogenases (Brändén et al., 1984; Gaut & Clegg, 1991). On the other hand, prokaryotic and archaebacterial genes coding for proteins cloned so far have an uninterrupted structure.

This paper describes the general properties, protein primary structure, and gene sequencing of *S. solfataricus* strain DSM1617 alcohol dehydrogenase (SSADH) as well as its relationship to well-known eukaryotic and prokaryotic alcohol dehydrogenases.

† This work has been carried on under a research contract with Tecnofarmaci Spa., Pomezia, Italy, within the Advanced Biotechnology National Research Plan of the Italian Minister of University and Scientific and Technological Research.

* Address correspondence to this author at IBPE, CNR, Via Marconi 10, 80125 Napoli, Italy.

‡ Menarini Ricerche Sud.

§ IBPE, CNR.

∥ DCOB.

[1] Abbreviations: ADH, alcohol dehydrogenase; FAB-MS, fast atom bombardment mass spectrometry; DSM1617, German collection of microorganisms and cell culture 1617; DTT, dithiothreitol; SDS–PAGE, sodium dodecyl sulfate–polyacrylamide gel electrophoresis; SSADH, *Sulfolobus solfataricus* ADH from DSM1617 strain; SSC, sodium citrate solution; TBADH, *Thermoanaerobium brockii* ADH.

0006-2960/92/0431-12514$03.00/0

MATERIALS AND METHODS

Protein Purification. SSADH was purified using a procedure similar to that previously described by Rella et al. (1987a). *S. solfataricus* strain DSM1617 was grown on glucose and harvested during the stationary phase. In a typical preparation, 105 g of bacteria was suspended in 20 mM Tris-HCl, 5 mM 2-mercaptoethanol, 10% glycerol, 1 mM $MgCl_2$, and 0.3 mM EDTA, pH 8.4 (buffer A), plus 1 M NaCl and homogenized as described (Rella et al., 1987a). After dialysis against buffer A, the sample was loaded on a DEAE-Sepharose Fast Flow column (Pharmacia) equilibrated in the same buffer. The enzyme was eluted with a NaCl gradient (0–0.3 M). The active pool was dialyzed against 50 mM Tris-HCl, pH 7.4, and 5 mM 2-mercaptoethanol, and then loaded onto a Matrex Gel Red A column (Amicon), equilibrated in the same buffer. The enzyme was eluted with a gradient of 0–1.25 M NaCl. Blue A (Amicon) chromatography was performed in 20 mM Tris-HCl, pH 7.4, on the Red A active pool dialyzed against the same buffer. After washing, the enzyme was eluted with 2 mM NAD^+ in 20 mM Tris-HCl, pH 8.4. The active pool from the Blue A column was adjusted to 1 M ammonium sulfate and applied directly onto a phenyl-Superose column equilibrated with 50 mM Tris-HCl and 1 M ammonium sulfate, pH 7.4. The column was extensively washed with the same buffer, and then the enzyme was eluted with a decreasing ammonium sulfate salt gradient (1.0–0 M). The active pool was dialyzed against 10 mM Tris-HCl, pH 8.4, and concentrated to 1–5 mg of protein/mL by ultrafiltration on a PM30 Amicon membrane. The enzyme thus prepared was stored at 4 °C without loss of activity for several months. SDS–PAGE (Laemmli, 1970) was performed using a Bio-Rad Miniprotean apparatus. Protein concentration was determined using the Bio-Rad Protein Assay kit, following the manufacturer's instructions and using bovine serum albumine as standard.

Enzyme Assay and Kinetic Parameters. Enzymatic activity was measured spectrophotometrically as described (Rella et al., 1987a) with some modifications. The standard reaction mixture contained 5 mM benzyl alcohol and 2 mM NAD^+ in 25 mM barbital hydrochloride, pH 8.5. The enzyme solution was added to preheated cuvettes, and initial rates were followed with a Varian DMS200 spectrophotometer. Enzyme activity units are expressed as micromoles of NADH produced per minute at 65 °C under the conditions described. Enzymatic parameters were determined at 65 °C in 25 mM barbital hydrochloride, pH 8.5, or in 25 mM Tris-HCl, pH 7.2, for the oxidation and reduction reactions, respectively. Kinetic results were analyzed using the ENZFITTER program (Elsevier-BIOSOFT).

Amino Acid Analysis. Hydrolysis was carried out under vacuum for 24 h, at 110 °C in 6 M hydrochloric acid containing 0.5% (w/v) phenol and 0.04% 2-mercaptoethanol (Moore & Stein, 1963). Amino acid analysis was performed using either a Carlo Erba Model 3A29 automatic analyzer, with a single-column protein hydrolysate program, or an ABI Model 420A derivatizer, equipped with an automatic hydrolysis station. Amino-terminal sequences were performed on an ABI Model 477A protein sequencer equipped with an on-line phenylthiohydantoin analyzer, Model 120A, following manufacturer's instructions.

Zinc Analysis. Contamination of protein samples by adventitious zinc was minimized by employing buffers prepared with ultrapure water obtained from a Milli-Q apparatus and by the use of plastic- and glassware cleaned exaustively with 20% HNO_3 and then rinsed with the same water. VISKING dialysis tubing (Serva) was washed in ultrapure water at 50–60 °C several times before use. Purified SSADH, at a concentration of 16 μM or over, was dialyzed at room temperature for 6–8 h against at least five changes of 200 volumes of 25 mM Tris-HCl, pH 7.5, containing 1 mM 1,-10-phenanthroline (Sigma) or 3 mM EDTA. Residual chelating agent was removed by dialysis against 4 or 5 changes of 200 volumes of the same buffer over a 16-h period. Another enzyme sample was dialyzed under the same conditions in the absence of chelating agent. Both samples were then stored in metal-free plastic test tube and used for activity assay and for protein and zinc analysis. The protein concentration was determined by amino acid analysis, while the zinc content was estimated with a Perkin-Elmer 5100 PC atomic absorption spectrophotometer equipped with an AS 60 autosampler and a Zeeman 5100 furnace.

Protein Reduction and Alkylation. The purified protein was reduced with DTT (molar ratio over protein SH groups 5:1) in the presence of 6 M guanidine hydrochloride, 0.3 M Tris-HCl, pH 7.8, and 1 mM EDTA, at 37 °C for 4 h under nitrogen; 4-vinylpyridine (Sigma) was added (molar ratio over protein SH groups 150:1), and the reaction proceeded in the dark at room temperature for 90 min under nitrogen. The S-pyridylethylated protein was then purified from excess reagents using reverse-phase chromatography on a Waters μBondapak C18 column (0.39 × 30 cm), eluted with a gradient from 5% to 60% acetonitrile in 0.1% TFA. The S-pyridylethylated protein showed a slightly lower retention time with respect to the unmodified protein.

Enzymatic Digestion and Chemical Cleavage. Trypsin digestion was performed at 37 °C for 4 h in 1% ammonium bicarbonate (trypsin/protein ratio 1:50 w/w). Peptides were first separated by gel filtration on a Sephadex G-50 SF column (1.4 × 190 cm), and then some of the collected pools were separated by reverse-phase HPLC. Peptide separation was performed on a Waters μBondapack C18 column, using a Beckman liquid chromatographer Model 324, equipped with a variable-wavelength monitor; elution was accomplished by mixing 0.1% aqueous TFA with 0.08% TFA in acetonitrile at a flow rate of 1 mL/min; detection was carried out at 220 nm. Digestions of S-pyridylethylated protein with carboxypeptidases A and B were performed as described by Ambler (1967). Cleavage at methionyl residues was obtained by incubating the S-pyridylethylated protein in 70% formic acid with CNBr (molar ratio 100:1 over methionyl residues) for 24 h in the dark at room temperature under nitrogen (Fontana & Gross, 1986). The peptide mixture obtained was lyophilized and separated by reverse-phase HPLC. In situ cleavage of Asp–Pro bonds was performed by adsorbing the protein dissolved in 70% formic acid on the sequencer glass fiber filter and incubating it at 40 °C for 20 h (Schininà et al., 1988). The sample was then applied to the sequencer: a first cycle with *o*-phthalaldehyde abolished all sequences starting with a primary amino group, which meant that only peptides starting with proline were sequenced (Brauer et al., 1984).

Identification of Genomic Clones. Two *S. solfataricus* genomic libraries of different sizes were constructed as described by Cubellis et al. (1990). The first was obtained after digestion of *S. solfataricus* DNA by DNase I; the fragments, about 1 kb, were cloned in phage λGT11. The second was obtained by partial *Nde*II digestion of *S. solfataricus* DNA; the DNA fragments, ranging between 10 and 15 kb, were inserted into phage λEMBL3. On the basis of the N-terminal amino acid sequence of the protein and homology comparison among known ADHs, two mixed

Table I: Purification of SSADH

step	total protein (mg)	total activity (units)	specific activity (units/mg)	yield (%)	purification (x-fold)
crude extract	4100	38[a]	0.01	100	
DEAE-Sepharose Fast Flow	745	90	0.12	230	12
Matrex gel Red A	97	110	1.13	290	113
Matrex gel Blue A	40	91	2.32	240	232
phenyl-Superose HR 10/10	12	60	5.31	157	531

[a] Apparent units.

oligonucleotides were synthesized on an Applied Biosystems Synthesizer, Model A381, and used as molecular probes to screen the first genomic library. Probe A (corresponding to residues 61–74 of the SSADH amino acid sequence) was 5′ GGT GTT AA(A/G) (C/T)TA CCA GT(T/A) AC(T/G) (T/C)TA GGC CAT GAA AT(T/A) GC(T/A) GGT AA-(A/G) AT; probe B (corresponding to residues 42–49) was 5′ GA(T/C) GT(T/A) CA(T/C) ATG (C/A)G(A/G) GG. The probes were 5′-end-labeled with [γ-^{32}P]ATP (Amersham; 3000 Ci/mmol) by use of T4 polynucleotide kinase (Boehringer). Lysis plaques from the first library were screened for hybridization with probe A; the positive clones were tested for hybridization with probe B. The hybridization conditions were as described by Maniatis et al. (1983), according to the different sizes of the probes. The filters were washed in stringent conditions and finally autoradiographed. The isolated phage DNAs were extracted, digested, and analyzed using agarose gel electrophoresis (Bio-Rad).

The second genomic library was screened with the largest positive fragment isolated from the first library (probe C). This DNA insert was random labeled with [α-^{32}P]ATP by means of DNA polymerase I, Klenow fragment (Boehringer). Hybridization conditions were as described by Maniatis et al. (1983). Phage DNA was extracted from the positive lysis plaque and digested with several restriction enzymes, and the fragments were again tested with probe C.

DNA Sequence Analysis. The restriction enzyme fragment obtained from the λEMBL3 library was ligated into the pEMBL 130 vector (Dente et al., 1988). Sequence analysis was carried out using the Sanger method, modified for double-stranded DNA, with the universal forward and reverse primers (17-mer, Pharmacia) and specific primers synthesized on the basis of the nucleotide sequence obtained. The labeled nucleotide was [α-^{35}S]dATP (Amersham; 600 Ci/mmol), and the sequence reaction was performed using the Sequenase kit (U.S.B.). The specific sequencing primers used were as follows: SAM1 = 5′TTTTATGAAGTAGCATG; SAM2 = 5′AAGTTTTTCCTTTTTAT; SP1 = 5′GGGTGTAAAAT-TGCCCG; SP2 = 5′TTGCCAAAGCTTTAGGG; Sp3 = 5′GACTTAACGCTGTAGAA; Sp4 = 5′CATCAATGCAA-GACCCC; Sp6 = 5′ATTTAGAGAATTTTAAG; Sp9 = 3′AGTGCATAGCCAAAG; Sp10 = 3′GGCAGGGAGAAG-GAAAT; Sp12 = 3′ATGAAGTAGCATGAGAG. Analysis of the gene structure was performed using the computer program DNASIS (Pharmacia).

RESULTS

Enzyme Purification, Structure, and Properties. S. solfataricus ADH was obtained in homogeneous form by means of a four-step purification summarized in Table I. This procedure is similar to that used for the previous *S. solfataricus* strain MT4 (Rella et al., 1987a) and is extremely useful as an enzyme with a higher specific activity is obtained. Recovery is over 100% because the activity is barely detectable in the homogenate and increases during purification.

Using gel filtration and sucrose gradient centrifugation, a molecular weight of about 74000 was obtained for the native enzyme, while a single band corresponding to a molecular weight of about 37 000 was observed when SDS–PAGE was used (data not shown). The enzymatic activity was strictly NAD^+/NADH dependent. The K_m values measured at 65 °C, pH 8.5, toward benzyl alcohol and NAD^+ were 0.22 and 0.32 mM, respectively; at the same temperature and pH 7.2 the K_m values toward *p*-anisaldehyde and NADH were 0.09 and 0.036 mM, respectively. At pH 8.5 and with benzyl alcohol as the substrate, the enzyme exhibited an activity which increased with temperature up to an instrumental limit of about 95 °C. At this temperature, the specific activity was 3 times higher than the value measured at 65 °C. Moreover, no loss of activity was detected after preincubation for 16 h at 65 °C, pH 8.0, at a protein concentration of 2.5 mg/mL. At a lower concentration (0.5 mg/mL), 50% of the activity was lost after 8 h at 65 °C and after 3 h at 85 °C.

Triplicate analysis of enzyme dialyzed against buffer at pH 7.5 in the absence and in the presence of chelating agents revealed an average of 4.2 and 1.9 mol of zinc per mol of dimer, respectively, while no loss of specific activity was detected in both samples after dialysis procedures.

Thermostability studies showed that the enzyme dialyzed against buffer retained 95% of activity after incubation for 3 h at 65 °C, at a protein concentration of 0.5 mg/mL. The enzyme dialyzed against chelating agents lost the whole activity after 1 h under the same experimental conditions. Similar results were obtained by incubation in 6 M urea or 0.1% SDS at 30 °C.

Protein Sequence Analysis. Data on the primary structure of the protein were obtained by direct N-terminal sequencing of the S-pyridylethylated protein and by peptide analysis, utilizing chemical (cyanogen bromide and mild acid treatment) and enzymatic (trypsin) cleavages (Figure 1). N-Terminal sequencing was carried out for 50 cycles, yielding identifiable residues. Methionine was present at the first cycle of Edman degradation on the protein obtained from different preparations. Peptides derived from cyanogen bromide cleavage and separated by reverse-phase HPLC (Figure 2) were sequenced until identifiable residues were produced. Peptide CB1, lacking the first methionyl residue, was not extensively sequenced, since a few cycles revealed that it corresponded to the previously determined amino-terminal sequence. Peptide CB2 was a large fragment (residues 45–138), and automatic sequencing proceeded up to residue 79; its C-terminal portion (residues 120–138) was worked out from the sequence of peptide CB2a, which did not arise from a methionyl cleavage and was occasionally found in a CNBr digest having suffered a specific acid cleavage.

Peptides from trypsin digestion were first separated by gel filtration on a Sephadex G-50 column (Figure 3) and then by reverse-phase HPLC (Figure 4). The relevant peptides were sequenced; the others were aligned according to amino acid composition.

Mild acid treatment of the S-pyridylethylated protein produced a cleavage at the two Asp–Pro bonds; direct sequencing, performed after blocking the primary amino groups by *o*-phthalaldehyde, yielded a double sequence. The sequence starting at Pro 228 was otherwise elucidated by sequencing the peptides derived from cyanogen bromide and trypsin cleavages (CB5 and T15). The sequence starting at

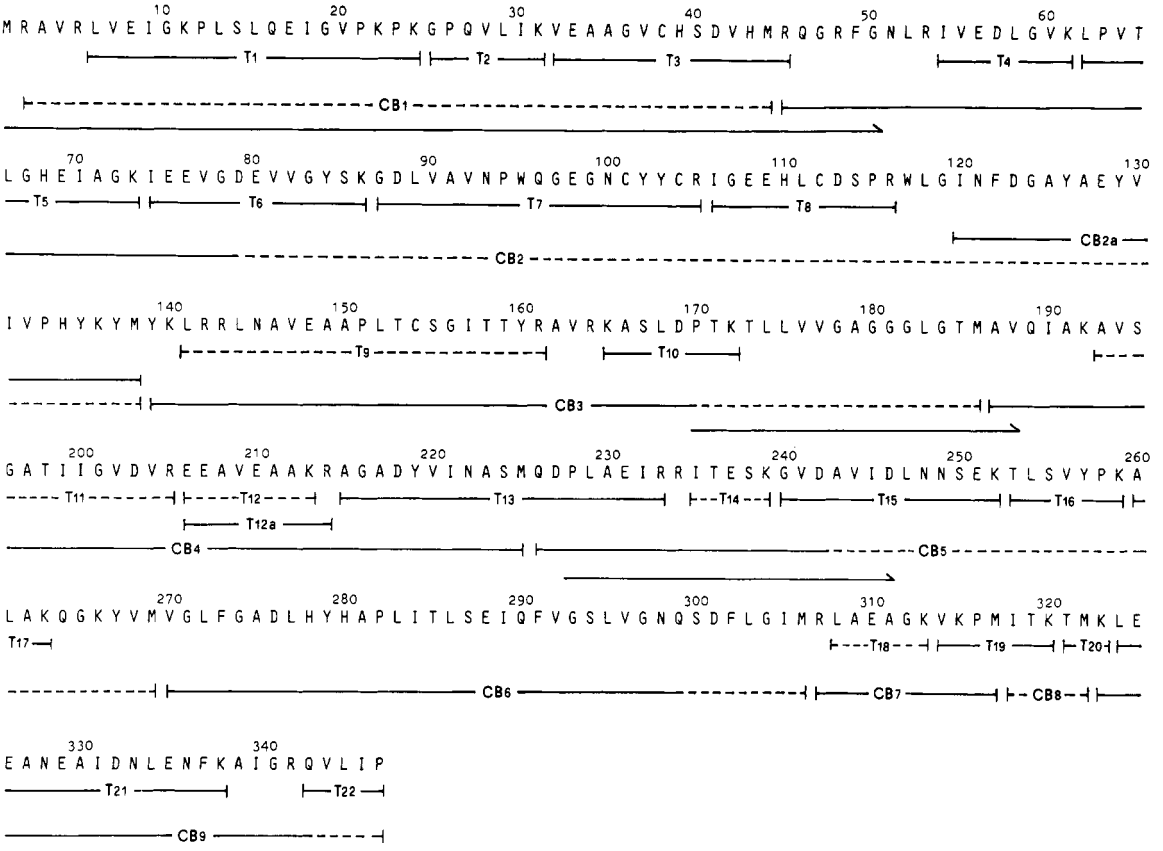

FIGURE 1: Amino acid sequence of alcohol dehydrogenase from *S. solfataricus* strain DSM1617 obtained by direct analysis of the protein sequence. Arrows indicate sequencing from the amino terminus and after Asp–Pro bond cleavage. Peptides are indicated by solid lines when determined by Edman degradation and dashed lines when aligned by amino acid composition. Peptides are denoted by letters (T for peptides derived from trypsin cleavage, CB for peptides derived from CNBr cleavage) and are numbered according to their order in the final structure.

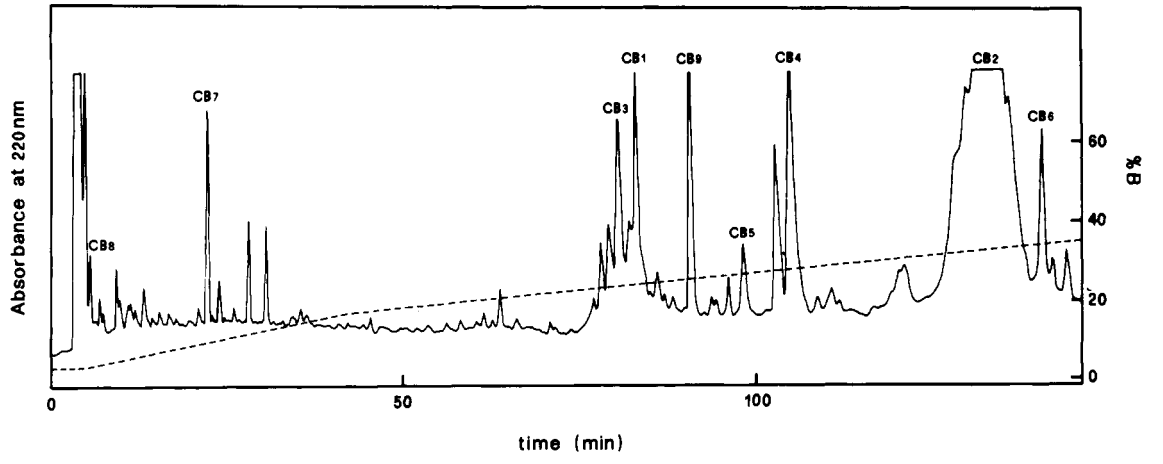

FIGURE 2: HPLC profile of peptides generated by CNBr digestion. The column was a C18 μBondapak. Solvent A was 0.1% TFA; solvent B was 0.08% TFA in acetonitrile. Flow rate: 1 mL/min.

Pro 170 was deduced from a comparison with the previous one up to Val 188, providing the overlap between CB3 and CB4.

Carboxypeptidase A and B digestions failed to release any amino acid residues from the protein, explained by the presence of a proline residue at the carboxyl-terminal end of the polypeptide chain.

Peptides were aligned either by overlapping fragments or by means of the nucleotide sequence. Although not complete, direct analysis of the protein sequence covered almost the whole sequence deduced from DNA analysis (95%). Furthermore, using protein sequence analysis, it was possible to recognize at least two sites of lysine methylation (position 11 and 213). N-Terminal sequencing of the intact S-pyridylethylated protein revealed the presence at position 11 of both monomethylysine (18–28%) and unmodified lysine (72–82%). The presence of the methylated residue at position 11 was also ascertained by FAB-MS analysis of the fraction containing peptide T1: two mass signals at 2057 and 2043, corresponding to peptides containing methyllysine and unmodified lysine were observed, respectively. Both peptides T12 and T12a were isolated from the tryptic digest: T12 contained unmodified lysine, and T12a contained methylated lysine (residue 213) which was not cleaved by trypsin. The relative amount of each peptide was about 50%.

Isolation and Sequencing of the ADH Gene. The screening of about 45000 λgt11 phages with the two oligonucleotide probes, described in the Materials and Methods section, yielded

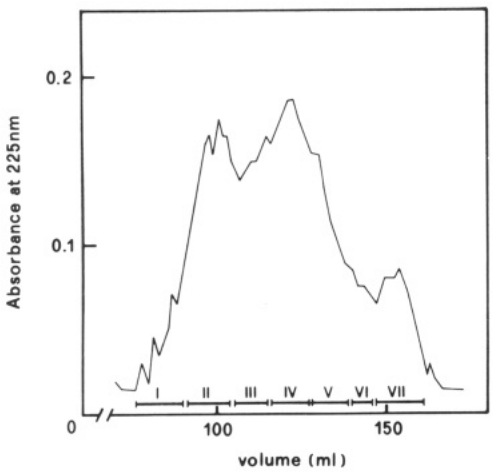

FIGURE 3: Separation of peptides generated by trypsin digestion using gel filtration on a Sephadex G-50 column (1.4 × 190 cm) equilibrated in 0.1% ammonium bicarbonate. Flow rate: 6 mL/h. Solid bars indicate pooled fractions.

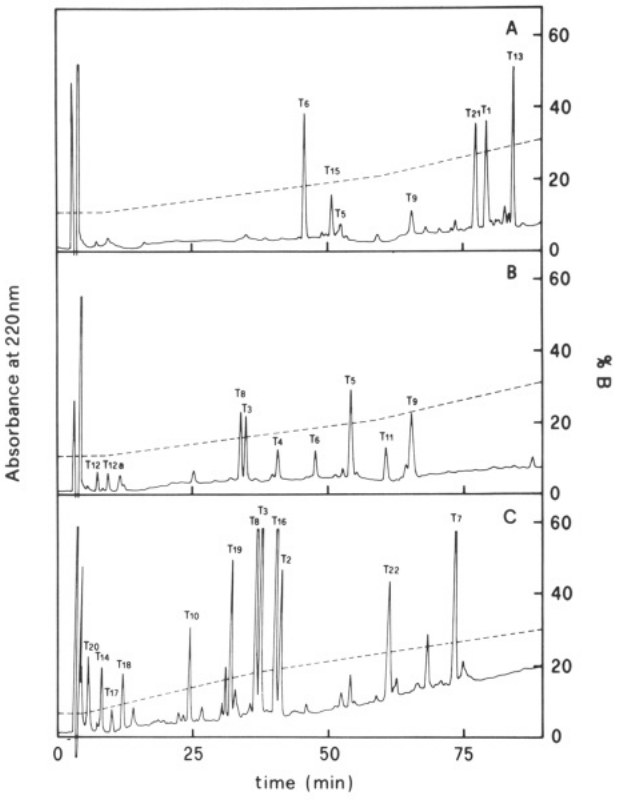

FIGURE 4: Separation by HPLC of pooled fractions from the Sephadex G-50 column. Panels A, B, and C correspond to pools II, III, and IV, respectively. Conditions as in Figure 2.

six clones giving positive hybridization signals with probe A; three of them, also giving positive signals with probe B, were selected for further analysis. The inserts were completely sequenced and overlapped for a total length of 1300 bp, including the entire sequence coding for alcohol dehydrogenase. In order to identify the smallest genomic region containing the whole gene coding for SSADH, about 15 000 λEMBL3 phages of the second genomic library were screened with probe C. Two hybridization positive phages, containing inserts of about 12 kb were isolated. A *Pst*I fragment (about 5 kb) was subcloned from one of these clones for sequence analysis.

The partial DNA sequence of fragment *Pst*I was determined according to the strategy shown in Figure 5. The specific primers were designed to proceed with the DNA sequence on

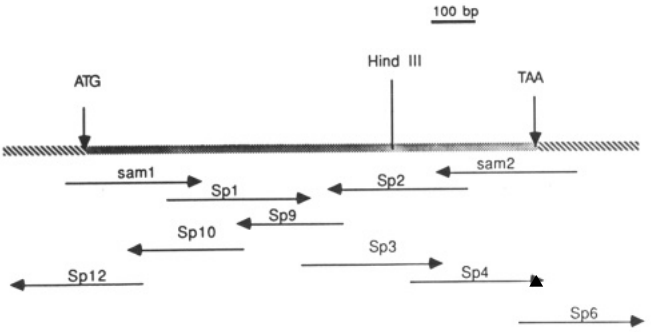

FIGURE 5: Sequence strategy of the SSADH gene. The specific primers used in the sequence, as described under Materials and Methods, are indicated.

both strands as indicated. An open reading frame from nucleotide 346 to 1389 was identified (Figure 6); an ATG codon is the first translation site, in agreement with data from protein amino acid sequence determination, indicating that the N-terminal methionine is not posttranslationally removed in the cell. The first stop codon, TAA, is encountered after 1041 bp and is preceded by CCA, coding for proline. The gene codifies for a polypeptide of 347 amino acids corresponding to a molecular weight of 37 588.

Two conserved sequences were recognized: the ATGA motif and the TTTAAT motif, 10 p and 42 bp before the starting ATG codon, respectively, corresponding to box B and box A as described by Reiter et al. (1988a). The first putative terminator signal TTTTTCTTT was encountered at position 82 bp downstream from the stop codon, similar to that described by Reiter et al. (1988b).

DISCUSSION

Gene Structure. Archaebacteria have a structural genomic organization resembling that of Eubacteria, although gene products and regulative elements are often more homologous with those of their eukaryotic counterparts. The analysis of the gene structure of SSADH (Figure 6) shows typical archaebacterial gene features (Zillig et al., 1988): location and composition of the promoter motifs, including the TATA-like box, which, as suggested by Reiter et al. (1990), is recognized by eukaryotic RNA polymerase types II and III. In addition, there are several stem and loop structures in the gene. At the 3′-untranslated region, three pyrimidine-rich terminator signals are present; a long stem structure is located between the first and second signals which may have the functional importance suggested by Cheng et al. (1991) for eubacterial systems (see Figure 6).

It has been suggested that the evolution of genes encoding for dehydrogenases proceeded through the rearrangement of exons corresponding to functional domains, i.e., exon-shuffling mediated by the presence of introns (Duester et al., 1986). However, no introns have been found in the SSADH gene, analogously to other protein encoding genes in *S. solfataricus* (Cubellis et al., 1989, 1990). Interestingly, sequence 481–487 of the SSADH gene is identical to the sequence of the exon/intron II junction in the maize ADH gene (Dennis et al., 1984; Brändén et al., 1984); an insertion of nine amino acid residues is observed at this position (48–56) in SSADH in the alignment of its amino acid sequence with that of other ADHs (Figure 7). This observation seems to indicate that introns are the result of an insertion of transposable elements into previously uninterrupted genes.

In Table II, the analysis of the SSADH codon usage, confirmed by direct protein sequence analysis, shows an

```
                                                                  30                                                                  60
TGC TAG TGA TAT TGT TGT TGC AAT GAG CCT CCT CAC AGA AGT TGA GAG CAA TTG GAA AAA

                                                                  90                                                                 120
ATT AAA TGT TGA TAG AGT TCT ATT CTA ACA ATA AAA AAT TGG TTA AGA GTA GAT AAG GAG

                                                                 150                                                                 180
AAG ATA AGG AAG GCG GTT ATG GAA TTA GGC GTT GAT GGC GTT GGG TAA TTC CGC AAA AGA

                                                                 210                                                                 240
GAA AAT GAA GGA ATT CAT GAA CAG AAA ACG TAC ATA ATG AGG TGC CTT TGT CTA TAT AAA

                                                                 270                                                                 300
TCT TAA TAT ATC ATA ACT CTA GAC ATT GTC AGT AAT GCT ATT ACG TTA TAT AAC ATT ATA

                                                                 330                                                                 360
ATA TTT AAT AAT ATC CAG AAA CCT ATT TAT ATT TTA TGA AGT AGC ATG AGA GCA GTT AGA
                                                                                                     M   R   A   V   R

                                                                 390                                                                 420
TTA GTA GAA ATA GGA AAA CCC CTT AGC TTA CAA GAG ATA GGT GTG CCT AAA CCC AAA GGA
 L   V   E   I   G   K   P   L   S   L   Q   E   I   G   V   P   K   P   K   G

                                                                 450                                                                 480
CCT CAA GTC TTA ATA AAA GTA GAG GCA GCG GGA GTT TGT CAT TCT GAT GTG CAC ATG AGA
 P   Q   V   L   I   K   V   E   A   A   G   V   C   H   S   D   V   H   M   R

                                                                 510                                                                 540
CAA GGA AGG TTT GGG AAT TTA AGA ATA GTG GAG GAT TTG GGT GTA AAA TTG CCC GTA ACT
 Q   G   R   F   G   N   L   R   I   V   E   D   L   G   V   K   L   P   V   T

                                                                 570                                                                 600
TTA GGT CAT GAG ATT GCA GGA AAA ATA GAG GAA GTT GGA GAT GAA GTA GTT GGA TAT TCT
 L   G   H   E   I   A   G   K   I   E   E   V   G   D   E   V   V   G   Y   S

                                                                 630                                                                 660
AAG GGG GAT CTA GTT GCA GTT AAT CCT TGG CAG GGA GAA GGA AAT TGC TAC TAT TGT AGA
 K   G   D   L   V   A   V   N   P   W   Q   G   E   G   N   C   Y   Y   C   R

                                                                 690                                                                 720
ATA GGA GAG GAG CAT CTG TGT GAC TCT CCT AGA TGG TTG GGG ATT AAT TTT GAT GGA GCT
 I   G   E   E   H   L   C   D   S   P   R   W   L   G   I   N   F   D   G   A

                                                                 750                                                                 780
TAT GCA GAG TAT GTA ATA GTT CCT CAT TAT AAA TAT ATG TAT AAG CTG AGA AGA CTT AAC
 Y   A   E   Y   V   I   V   P   H   Y   K   Y   M   Y   K   L   R   R   L   N

                                                                 810                                                                 840
GCT GTA GAA GCA GCA CCG TTA ACT TGT TCT GGA ATA ACT ACA TAT AGG GCA GTA AGA AAG
 A   V   E   A   A   P   L   T   C   S   G   I   T   T   Y   R   A   V   R   K

                                                                 870                                                                 900
GCA TCT TTA GAT CCA ACT AAG ACA CTA CTA GTA GTG GGT GCT GGA GGA GGC TTA GGC ACA
 A   S   L   D   P   T   K   T   L   L   V   V   G   A   G   G   G   L   G   T
```

FIGURE 6: Partial nucleotide sequence of the genomic *Pst*I fragment containing the full-length SSADH gene and the translated amino acid sequence. Numbers refer to the nucleotide sequence.

```
                                                                 930                                                                 960
ATG GCA GTG CAG ATA GCC AAA GCT GTT AGT GGT GCA ACG ATA ATA GGT GTG GAT GTG AGG
 M   A   V   Q   I   A   K   A   V   S   G   A   T   I   I   G   V   D   V   R

                                                                 990                                                                1020
GAA GAG GCA GTG GAG GCT GCT AAG AGA GCT GGA GCT GAC TAC GTA ATA AAC GCA TCA ATG
 E   E   A   V   E   A   A   K   R   A   G   A   D   Y   V   I   N   A   S   M

                                                                1050                                                                1080
CAA GAC CCC TTA GCG GAG ATA AGA AGG ATA ACT GAA TCG AAA GGA GTG GAT GCC GTA ATA
 Q   D   P   L   A   E   I   R   R   I   T   E   S   K   G   V   D   A   V   I

                                                                1110                                                                1140
GAC CTG AAT AAC TCT GAG AAA ACG CTT TCA GTT TAC CCT AAA GCT TTG GCA AAA CAA GGT
 D   L   N   N   S   E   K   T   L   S   V   Y   P   K   A   L   A   K   Q   G

                                                                1170                                                                1200
AAA TAC GTA ATG GTG GGA TTG TTT GGT GCA GAT TTG CAT TAT CAT GCA CCT CTA ATA ACT
 K   Y   V   M   V   G   L   F   G   A   D   L   H   Y   H   A   P   L   I   T

                                                                1230                                                                1260
TTA TCA GAA ATA CAA TTT GTA GGC AGT TTA GTT GGG AAT CAA TCA GAC TTC TTG GGG ATA
 L   S   E   I   Q   F   V   G   S   L   V   G   N   Q   S   D   F   L   G   I

                                                                1290                                                                1320
ATG AGA TTA GCA GAG GCT GGT AAA GTC AAA CCT ATG ATA ACT AAA ACT ATG AAA CTA GAA
 M   R   L   A   E   A   G   K   V   K   P   M   I   T   K   T   M   K   L   E

                                                                1350                                                                1380
GAG GCA AAT GAG GCT ATT GAT AAT TTA GAG AAT TTT AAG GCT ATT GGA AGA CAA GTA CTC
 E   A   N   E   A   I   D   N   L   E   N   F   K   A   I   G   R   Q   V   L

                                                                1410                                                                1440
ATA CCA TAA AAA GGA AAA ACT TAA GAA GGT TTG AGA TAA AGT ATA AAT ACT TTT TCT AAT
 I   P

                                                                1470                                                                1500
CTA AAT GGC GAT ACT GAA GTC ACC GTA TTT TTC TTT GAT GTA ATC GAC AAT TCT CTC TTT

                                                                1530                                                                1560
TTA GCA TCT TCA ACT CGT TGA TTA CTT CAG ATG GTA TAT CTT CAA ATT TAA GGT AAA TTT

                                                                1590                                                                1620
CAA TTA TAC GTT CGG TTT TCT CTA AGT ATT TCC TCA ATT TTA TTG TGA ATT CGT TTT CAT

                                                                1650                                                                1680
ACT CTT CCT AGT TGT ATT CCT TTC CTA AAA CTC GTT TGA AAA GGG CTT TAA GAT CTT CAT

                                                                1710                                                                1740
AAA ATG GGA TAA GAC CAA TAG GTG ATA TAG TCG ACT TAG ACG TTT CCT CCA CTC TTT TTA

                                                                1770
CTG CCC AGC TAA CCC ATA CTT TCT TAT CTT CCT TT
```

Box A and box B are located at positions 304–309 and 336–339, respectively. The terminator signals are present at positions 1432–1437, 1469–1474, and 1576–1581.

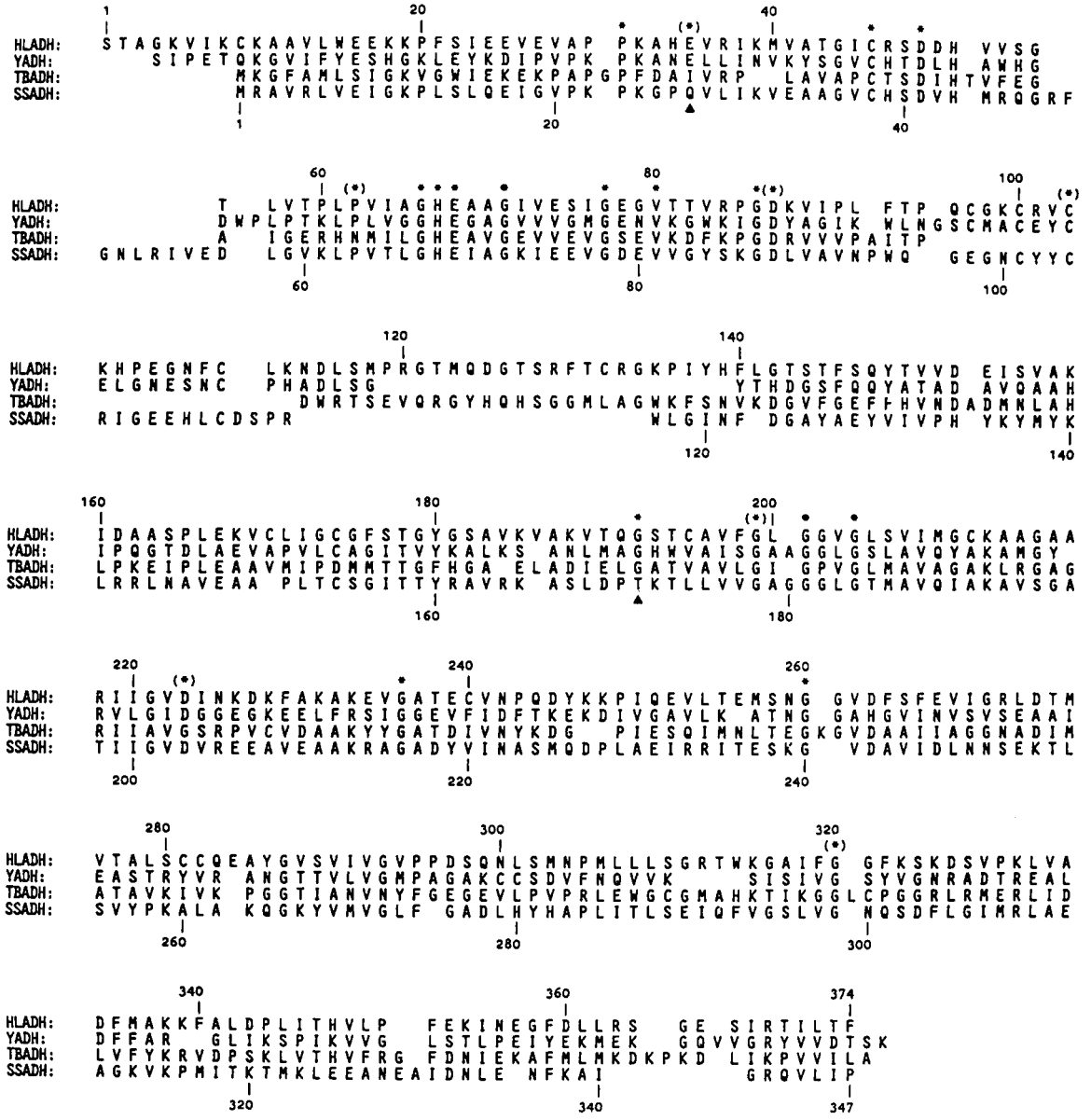

FIGURE 7: Comparison of the primary structure of alcohol dehydrogenases from horse liver (HLADH, EE enzyme), from the yeast *S. cerevisiae* (YADH, isoenzyme 1), from the bacterium *T. brockii* (TBADH) and from the archaebacterium *S. solfataricus* strain DSM1617 (SSADH). Gaps were introduced to optimize identity between sequences. Residue numbering of HLADH is above the sequence; that of SSADH is below the sequence. Asterisks indicate the 22 strictly conserved residues in 16 different alcohol/polyol dehydrogenases (Jornvall et al., 1987); asterisks within parentheses indicate conserved residues which have already been replaced in *T. brockii* and/or *Alcaligenes eutrophus* alcohol dehydrogenase (Peretz & Burnstein, 1989); black triangles represent conserved residues which are replaced in SSADH. From this alignment, the identity percentages have been calculated by dividing the number of identical residues at the corresponding positions by the total number of amino acid residues of SSADH, not considering gaps and insertion(s) in the calculation.

interesting feature: the six time-degenerated amino acids are preferentially codified by triplets with T or A at the first and/or the third position; the four time-degenerated residues, by triplets with T or A at the third position. This preference is still observed in the isoleucine codons and in the twice degenerated amino acid codons; the only exception is represented by the glutamic acid codons (GAA/GAG = 0.56). This feature, which has also been observed in other *S. solfataricus* genes, can be related to the level of protein expression in the cell (Cubellis et al., 1989, 1990).

Protein Structure. The alcohol dehydrogenase purified from *S. solfataricus*, strain DSM1617, is similar in its general properties to that obtained from the MT4 strain (Rella et al., 1987a). The enzyme is a NAD^+-dependent alcohol dehydrogenase which is remarkably thermophilic and thermostable. It was purified to homogeneity with high yields, from the crude extract using an improved procedure. Enzyme activity increased during purification, probably due to the presence of factors inhibiting enzyme activity in the early steps.

The primary structure of SSADH was determined by combining protein and gene analysis; it consists of 347 amino acid residues with free N-terminal methionine and C-terminal proline. These analyses support the subunit size and quaternary structure deduced from electrophoretic and chromatographic behavior and establish subunit homogeneity, as no heterogeneity was found during protein sequence analysis. Therefore, this archaebacterial enzyme can be included in the family of dimeric medium-chain alcohol dehydrogenases.

In order to analyze the functional and structural features of SSADH, its primary structure was compared to that of three evolutionarily distinct alcohol dehydrogenases: horse liver alcohol dehydrogenase (E-type subunit) from protein data (Brändén et al., 1975) verified by X-ray crystallography (Eklund et al., 1976); yeast alcohol dehydrogenase from

Table II: Codon usage of SSADH-Encoding ORF (nt 346–1386 in Figure 6)

a	b	c	a	b	c	a	b	c
TTT	Phe	83.3	CCT	Pro	53.3	AAA	Lys	73.9
TTC	Phe	16.7	CCC	Pro	26.7	AAG	Lys	26.1
			CCA	Pro	13.3			
TTA	Leu	40.6	CCG	Pro	6.7	GAT	Asp	66.6
TTG	Leu	22.0				GAC	Asp	33.4
CTT	Leu	9.3	ACT	Thr	61.5			
CTC	Leu	3.1	ACC	Thr	0.0	GAA	Glu	36.0
CTA	Leu	15.6	ACA	Thr	23.1	GAG	Glu	64.0
CTG	Leu	9.3	ACG	Thr	15.4			
						TGT	Cys	80.0
ATT	Ile	16.7	GCT	Ala	35.3	TGC	Cys	20.0
ATC	Ile	0.0	GCC	Ala	5.9			
ATA	Ile	83.3	GCA	Ala	52.9	TGG	Trp	100
			GCG	Ala	5.9			
ATG	Met	100				CGT	Arg	0.0
			TAT	Tyr	69.3	CGC	Arg	0.0
GTT	Val	27.8	TAC	Tyr	30.7	CGA	Arg	0.0
GTC	Val	5.5				CGG	Arg	0.0
GTA	Val	38.9	CAT	His	85.7	AGA	Arg	76.5
GTG	Val	27.8	CAC	His	14.3	AGG	Arg	23.5
TCT	Ser	42.8	CAA	Gln	80.0	GGT	Gly	25.8
TCC	Ser	0.0	CAG	Gln	20.0	GGC	Gly	8.6
TCA	Ser	28.6				GGA	Gly	51.4
AGT	Ser	14.3	AAT	Asn	75.0	GGG	Gly	14.2
AGC	Ser	7.1	AAC	Asn	25.0			
TCG	Ser	7.2						

[a] Codon. [b] Amino acid. [c] Percent (%) codon utilization in SSADH.

Saccharomyces cerevisiae, isoenzyme 1, from protein data (Jörnvall et al., 1978); and *T. brockii* alcohol dehydrogenase, from protein data (Peretz & Burstein, 1989).

The alignment obtained according to Jörnvall et al.'s (1987) suggestion is given in Figure 7. Several gaps and insertions were introduced to optimize identity among sequences.

In the alignment of a large set of alcohol/polyol dehydrogenase sequences (Jörnvall et al., 1987), a common gap segment just before position 60 (numbering of HLADH) was introduced in correspondence to an intron in plant alcohol dehydrogenase genes. At this position, an insertion of 9 residues is found in SSADH. Moreover, a long gap, consisting of 21 residues at position 116 of SSADH, is found in the corresponding position of the yeast enzyme. This position is very close to an intron position in the β human alcohol dehydrogenase gene. Other minor gaps are present, four of which, at positions 186, 284, 321, and 364 (numbering of HLADH), are found at or very close to the intron position in plant and β human alcohol dehydrogenase genes. These observations are in agreement with the hypothesis that a change in the polypeptide chain length can arise at the level of the intron positions due to ADH gene evolution events (Jörnvall et al., 1987).

SSADH is quite different from the other three alcohol dehydrogenases compared, exhibiting 25% sequence identity to both horse liver and yeast enzyme and 24% sequence identity to the *T. brockii* enzyme.

Comparison of the aligned sequences reveals that 20 out of 22 strictly conserved residues which Jörnvall et al. (1987) have stated are typical in alcohol dehydrogenase conformation are preserved in SSADH. These residues are 10 Gly, 3 Asp, 2 Cys, 2 Pro, and 1 each of His, Glu, and Val (see Figure 7). The remaining two strictly conserved residues Glu-35 and Gly-192 (HLADH numbering) are replaced in SSADH by Gln and Thr, respectively. As stated previously, the number of invariant residues may decrease when more sequences are compared (Jörnvall et al., 1987). In fact, a reduction of the number of invariant residues from 22 to 15 has already been observed when comparing the sequences of TBADH and *Alcaligenes eutrophus* ADH (Peretz & Burstein, 1989). In SSADH, 1 of these 15 residues, Gly 192, is replaced by a Thr residue; thus the strictly conserved residues of medium-chain alcohol dehydrogenases are further reduced to 14. As mentioned above, half of the strictly conserved residues are glycine. This retention of a high number of conserved glycines has already been observed as a characteristic feature of distantly related enzymes with conserved conformations and functions [see Jörnvall et al. (1987)]. The pattern of glycine residues directly involved in the coenzyme-binding domain is located in the central part of the molecule (positions 178–184 of SSADH). The spacing of this glycine motif differs from that found in horse liver ADH (GXGXXG), but it is identical to that found in the enzymes from yeast and *Aspergillus nidulans* (GXXGXXG) (Jörnvall et al., 1987). Asp 223 (HLADH numbering), a residue essential for the binding of NAD^+, is conserved in SSADH (position 203), as expected for an NAD^+-dependent alcohol dehydrogenase (Brändén et al., 1975; Fan et al., 1991).

The horse liver enzyme contains two zinc atoms per subunit, one catalytic, bound to amino acid residues Cys 46, His 67, and Cys 174, and one structural, having as ligands four cysteine residues at positions 97, 100, 103, and 111 (Brändén et al., 1975; Vallee & Auld, 1990a). In SSADH the same three residues involved in the binding of the catalytic zinc are present at positions 38, 68, and 154, respectively (Figure 7). Consequently, it seems likely that this enzyme contains zinc in the catalytic site. Zinc analysis combined with dialysis experiments showed that the native SSADH contains 4 mol of zinc per mol of dimer and that the chelating agents used removed only two of the four zinc atoms. The enzyme with only two zinc retained its catalytic activity, but lost its resistance to heat and denaturing agents, suggesting that two zinc atoms have a catalytic role and two have a structural role. As far as the structural zinc binding site is concerned, it should be noted that only three out of four cysteine residues are conserved in SSADH; in fact, Cys 97 (numbering of HLADH) is replaced by Glu at position 98 of the archaebacterial enzyme. The substitution Cys→Glu is remarkable as it has been found that Glu is a zinc ligand in the catalytic site of the computer model of sheep liver sorbitol dehydrogenase (Eklund et al., 1985) and is a ligand of this metal ion more frequently than cysteine and less than histidine in the catalytic site of zinc enzymes (Vallee & Auld, 1990b). At present, this is the first example of a Cys→Glu replacement in the structural zinc site of an alcohol dehydrogenase. Furthermore, if this substitution is linked to the remarkable thermostability of SSADH, stabilizing both local conformation and overall dimeric structure, it remains an interesting hypothesis to be demonstrated and correlated to other general trends in protein thermostabilization.

As already observed for other thermophilic enzymes, the number of Cys residues in SSADH is significantly reduced with respect to its counterpart from mesophiles (only 5 thiols per subunit, versus 8 in YADH and 14 in HLADH). This may be related to protein thermostability in that cysteine replacement avoids the presence of side chains sensitive to oxidation or other reactions on the protein surface (Nosoh & Sekiguchi, 1988; Maras et al., 1992). As already observed (Argos et al., 1979), the ratio Arg/Lys is larger in proteins from thermophilic organisms than in those from mesophilic, and more recently, direct proof for the stabilizing influence of arginine with respect to lysine has been provided (Mrabet

et al., 1992). Accordingly, a ratio Arg/Lys about twice as high with respect to other mesophilic ADHs is found in SSADH. Furthermore, the methylation of lysine has also been claimed to contribute to protein thermostability (Maras et al., 1992). Although in SSADH two lysyl residues at positions 11 and 213 were methylated to different extents, the role of methylation and that of other peculiar amino acid substitutions as stabilizing factors in this archaebacterial enzyme need to be investigated with the aid of three-dimensional computer models or directly by X-ray crystallography and site-specific mutagenesis.

ACKNOWLEDGMENT

We are grateful to Prof. A. Gambacorta of the ICMIB, CNR, Arco Felice, Naples, who kindly provided us with some of biomass; to Prof. P. Pucci of Centro di Spettrometria di Massa, CNR, Naples, for mass spectral analyses; to Dr. E. Parisi and Dr. M. R. Scudiero for zinc analyses; to Ms. N. Vespa for her efficient collaboration; and to Mrs. C. Di Fusco for the assistance in the manuscript typing. Skillful assistance in protein sequencing by Mr. V. Carratore is gratefully acknowledged.

REFERENCES

Ambler, R. P. (1967) *Methods Enzymol. 11*, 436–445.

Argos, P., Rossmann, M. G., Grau, U. M., Zuber, H., Frank, G., & Tratschin, J. D. (1979) *Biochemistry 18*, 5698–5703.

Brändén, C.-I., Jörnvall, H., Eklund, H., & Furugren, B. (1975) in *The Enzymes* (Boyer, P. D., Ed.) 3rd ed. Vol. 11, pp 103–190, Academic Press, New York.

Brändén, C.-I., Eklund, H., Cambillau, C., & Pryor, A. J. (1984) *EMBO J. 3*, 1307–1310.

Brauer, A. W., Oman, C. L., & Margolies, M. N. (1984) *Anal. Biochem. 137*, 134–142.

Bryant, F. O., Wiegel, J., & Ljungdahl, L. G. (1988) *Appl. Environ. Microbiol.*, 460–465.

Cheng, S.-W. C., Lynch, E. C., Leason, K. R., Court, D. L., Shapiro, B. A., & Friedman, D. I. (1991) *Science 254*, 1205–1207.

Cubellis, M. V., Rozzo, C., Nitti, G., Arnone, M. I., Marino, G., & Sannia, G. (1989) *Eur. J. Biochem. 186*, 375–381.

Cubellis, M. V., Rozzo, C., Montecucchi, P., & Rossi, M. (1990) *Gene 94*, 89–94.

Dennis, E. S., Gerlach, W. L., Pryor, A. J., Bennetzen, J. L., Inglis, A., Llewellyn, D. L., Sachs, M. M., Ferl, R. J., & Peacock, W. J. (1984) *Nucleic Acids Res. 12*, 3983–4000.

Dente, L., Sollazzo, M., Baldari, C., Cesareni, G., & Cortese, R. (1988) in *DNA Cloning* (Glover, D. M., Ed.) 1st ed. Vol. I, pp 101–107, IRL Press, Oxford and Washington, DC.

Duester, G., Jörnvall, H., & Hatfield, G. W. (1986) *Nucleic Acids Res. 14*, 1931–1940.

Eklund, H., Nordström, B., Zeppezauer, E., Söderlund, G., Ohlsson, I., Boiwe, T., Söderberg, B.-O., Tapia, O., Brändén, C.-I., & Åkeson, Å. (1976) *J. Mol. Biol. 102*, 27–59.

Eklund, H., Horjales, E., Jörnvall, H. Brändén, C.-I., & Jeffery, J. (1985) *Biochemistry 24*, 8005–8012.

Estonius, M., Karlsson, C., Fox, E. A., Höög, J.-O., Holmquist, B., Vallee, B. L., Davidson, W. S., & Jörnvall, H. (1990) *Eur. J. Biochem. 194*, 593–602.

Fan, F., Lorenzen, J. A., & Plapp, B. V. (1991) *Biochemistry 30*, 6397–6401.

Fontana, A., & Gross, E. (1986) in *Practical Protein Chemistry* (Darbre, A., Ed.) pp 68–115, John Wiley and Sons, New York.

Gaut, B. S., & Clegg, M. T. (1991) *Proc. Natl. Acad. Sci. U.S.A. 88*, 2060–2064.

Goodlove, P. E., Cunningham, P. R., Parker, J., & Clark, D. P. (1989) *Gene 85*, 209–214.

Inoue, T., Sunagawa, M., Mori, A., Imai, C., Fukuda, M., Takagi, M., & Yano, K. (1989) *J. Bacteriol. 171*, 3115–3122.

Jeffery, J., Cummins, L., Carlquist, M., & Jörnvall, H. (1981) *Eur. J. Biochem. 120*, 229–234.

Jendrossek, D., Steinbuchel, A., & Schlegel, H. G. (1988) *J. Bacteriol. 170*, 5248–5256.

Jörnvall, H. (1977) *Eur. J. Biochem. 72*, 425–442.

Jörnvall, H., Eklund, H., & Brändén, C.-I. (1978) *J. Biol. Chem. 253*, 8414–8419.

Jörnvall, H., Persson, M., & Jeffery, J. (1981) *Proc. Natl. Acad. Sci. U.S.A. 78*, 4226–4230.

Jörnvall, H., von Bahr-Lindstrom, H., & Jeffery, J. (1984) *Eur. J. Biochem. 140*, 17–23.

Jörnvall, H., Persson, M., & Jeffery, J. (1987) *Eur. J. Biochem. 167*, 195–201.

Kaiser, R., Holmquist, B., Vallee, B. L., & Jörnvall, H. (1989) *Biochemistry 28*, 8432–8438.

Laemmli, U. K. (1970) *Nature 227*, 680–685.

Llewellyn, D. J., Finnegan, E. J., Ellis, J. G. Dennis, E. S., & Peacock, W. J. (1987) *J. Mol. Biol. 195*, 115–123.

Maniatis, T., Fritsch, E. F., & Sambrook, J. (1983) *Molecular Cloning, A Laboratory Manual*, Cold Spring Harbor Laboratory, Cold Spring Harbor, NY.

Maras, B., Consalvi, V., Chiaraluce, R., Politi, L., De Rosa, M., Bossa, F., Scandurra, R., & Barra, D. (1992) *Eur. J. Biochem. 203*, 81–87.

McKnight, G. L., Kato, H., Upshall, W., Parker, M. D., Saari, G., & O'Hara, P. J. (1985) *EMBO J. 4*, 2093–2099.

Moore, S., & Stein, W. H. (1963) *Methods Enzymol. 6*, 819–831.

Mrabet, N. T., Van den Broeck, A., Van den brande, I., Stanssens, P., Laroche, Y., Lambeir, A.-M., Matthijssens, G., Jenkins, J., Chiadmi, M., van Tilbeurgh, H., Rey, F., Janin, J., Quax, W. J., Lasters, I., De Maeyer, M., & Wodak, S. J. (1992) *Biochemistry 31*, 2239–2253.

Neale, A. D., Scopes, R. K., Kelley, J. M., & Wettenhall, R. E. H. (1986) *Eur. J. Biochem. 154*, 119–124.

Nosoh, Y., & Sekiguchi, T. (1988) *Biocatalysis 1*, 257–273.

Peretz, M., & Burstein, Y. (1989) *Biochemistry 28*, 6549–6555.

Persson, B., Krook, M., & Jörnvall, H. (1991) *Eur. J. Biochem. 200*, 537–543.

Reiter, W.-D., Palm, P., & Zillig, W. (1988a) *Nucleic Acids Res. 16*, 1–19.

Reiter, W.-D., Palm, P., & Zillig, W. (1988b) *Nucleic Acids Res. 16*, 2445–2459.

Reiter, W.-D., Hüdepohl, U., & Zillig, W. (1990) *Proc. Natl. Acad. Sci. U.S.A. 87*, 9509–9513.

Rella, R., Raia, C. A., Pensa, M., Pisani, F. M., Gambacorta, A., De Rosa, M., & Rossi, M. (1987a) *Eur. J. Biochem. 167*, 475–479.

Rella, R., Raia, C. A., Trincone, A., Gambacorta, A., De Rosa, M., & Rossi, M. (1987b) in *Biocatalysis in Organic Media* (Laane, C., Tramper, J., & Lilly, M. D., Eds.) Vol. 29, pp 273–278, Elsevier, Amsterdam.

Schininà, M. E., De Biase, D., Bossa, F., & Barra, D. (1988) *J. Protein Chem. 7*, 284–286.

Scopes, R. K. (1983) *FEBS Lett. 156*, 303–306.

Sheehan, M. C., Bailey, C. J., Dowds, B. C. A., & McConnell, D. J. (1988) *Biochem. J. 252*, 661–666.

Sund, H., & Theorell, H. (1963) *Enzymes, 2nd Ed. 7*, 25–83.

Trincone, A., Lama, L., Lanzotti, V., Nicolaus, B., De Rosa, M., Rossi, M., & Gambacorta, A. (1990) *Biotechnol. Bioeng. 35*, 559–564.

Vallee, B. L., & Auld, D. S. (1990a) *Biochemistry 29*, 5647–5659.

Vallee, B. L., & Auld, D. S. (1990b) *Proc. Natl. Acad. Sci. U.S.A. 87*, 220–224.
Villaroya, A., Juan, E., Egestad, B., & Jörnvall, H. (1989) *Eur. J. Biochem. 180*, 191–197.
Williamson, V. M., & Paquin, C. E. (1987) *Mol. Gen. Genet. 209*, 374–381.
Woese, C. R., Kandler, O., & Wheelis, M. L. (1990) *Proc. Natl. Acad. Sci. U.S.A. 87*, 4576–4579.
Young, E. T., & Pilgrim, D. (1985) *Mol. Cell. Biol. 5*, 3024–3034.
Zillig, W., Palm, P., Reiter, W.-D., Gropp, F., Pühler, G., & Klenk, H.-P. (1988) *Eur. J. Biochem. 173*, 473–482.