Wobbe, C. R., Dean, F. B., Weissbach, L., & Hurwitz, J. (1985) *Proc. Natl. Acad. Sci. U.S.A. 82*, 5710-5714.
Wold, M. S., Li, J. J., & Kelly, T. J. (1987) *Proc. Natl. Acad. Sci. U.S.A. 84*, 3643-3647.
Wood, E. R., & Matson, S. W. (1987) *J. Biol. Chem. 262*, 15269-15276.
Yarranton, G. T., & Gefter, M. L. (1979) *Proc. Natl. Acad. Sci. U.S.A. 76*, 1658-1662.

Alanine Dehydrogenases from Two *Bacillus* Species with Distinct Thermostabilities: Molecular Cloning, DNA and Protein Sequence Determination, and Structural Comparison with Other NAD(P)$^+$-Dependent Dehydrogenases†

Shun'ichi Kuroda,‡ Katsuyuki Tanizawa,§ Yonekazu Sakamoto, Hidehiko Tanaka,‖ and Kenji Soda*

Laboratory of Microbial Biochemistry, Institute for Chemical Research, Kyoto University, Uji, Kyoto 611, Japan

Received April 24, 1989; Revised Manuscript Received September 8, 1989

ABSTRACT: The gene encoding alanine dehydrogenase (EC 1.4.1.1) from a mesophile, *Bacillus sphaericus*, was cloned, and its complete DNA sequence was determined. In addition, the same gene from a moderate thermophile, *B. stearothermophilus*, was analyzed in a similar manner. Large parts of the two translated amino acid sequences were confirmed by automated Edman degradation of tryptic peptide fragments. Each alanine dehydrogenase gene consists of a 1116-bp open reading frame and encodes 372 amino acid residues corresponding to the subunit (M_r = 39 500–40 000) of the hexameric enzyme. The similarity of amino acid sequence between the two alanine dehydrogenases with distinct thermostabilities is very high (>70%). The nonidentical residues are clustered in a few regions with relatively short length, which may correlate with the difference in thermal stability of the enzymes. Homology search of the primary structures of both alanine dehydrogenases with those of other pyridine nucleotide-dependent oxidoreductases revealed significant sequence similarity in the regions containing the coenzyme binding domain. Interestingly, several catalytically important residues in lactate and malate dehydrogenases are conserved in the primary structure of alanine dehydrogenases at matched positions with similar mutual distances.

Alanine dehydrogenase (L-alanine:NAD$^+$ oxidoreductase, deaminating, EC 1.4.1.1) catalyzes the reversible deamination of L-alanine to pyruvate and is found in vegetative cells (Hong et al., 1959; Zink & Sanwal, 1962; McCowen & Phibbs, 1974) and spores (O'Conner & Halvorson, 1961; Nitta et al., 1974) of various bacilli and in some other bacteria (Germano & Anderson, 1968; Holmes et al., 1965; Ohshima & Soda, 1979). This enzyme is a key factor in the assimilation of L-alanine as an energy source through the tricarboxylic acid cycle during sporulation (McCowen & Phibbs, 1974). Alanine dehydrogenase has been purified to homogeneity from *Bacillus subtilis*, *B. sphaericus*, and *B. cereus*, and its enzymological properties have been elucidated (Yoshida & Freese, 1965; Ohshima & Soda, 1979; Porumb et al., 1987). The kinetic and chemical mechanisms of the enzyme reaction have also been extensively studied (Grimshaw & Cleland, 1981; Grimshaw et al., 1981). It has been shown that the enzyme has different mechanisms for coenzyme binding (A stereospecificity) and substrate binding from those of glutamate dehydrogenase.

To date, a large number of NAD(P)$^+$-dependent dehydrogenases have been sequenced, and several of these have been structurally analyzed in atomic detail. For the dehydrogenases that act on amino acids, however, structural data are available only for primary structures of B-stereospecific dehydrogenases, such as glutamate (Smith et al., 1975), leucine (Nagata et al., 1988), and phenylalanine (Okazaki et al., 1988) dehydrogenases. In this work, we report the cloning and sequencing of the alanine dehydrogenase gene from a mesophile, *B. sphaericus*, and also the complete sequence of the thermostable alanine dehydrogenase from a moderate thermophile, *B. stearothermophilus*, the gene for which has been cloned recently in this laboratory (Nagata et al., unpublished results). The difference in thermostability between the two alanine dehydrogenases is discussed on the basis of their primary structures. Sequence comparison with other NAD(P)$^+$-dependent dehydrogenases has suggested possible locations of the coenzyme binding region and several catalytically important residues in the primary structure of alanine dehydrogenases.

Experimental Procedures

Strains and Media. *Escherichia coli* C600 $r_k^-m_k^-$ *thi thr leu* was used as a host strain for cloning and plasmid construction. A mesophile, *B. sphaericus* IFO3525, was used as the source for chromosomal DNA. Transformants were grown

† This study was supported in part by Grants-in-Aid for Scientific Research from the Ministry of Education, Science and Culture and by special coordination funds from the Ministry of Agriculture, Forestry and Fisheries of Japan.

* Author to whom correspondence should be addressed.
‡ Present address: Biotechnology Laboratories, Central Research Division, Takeda Chemical Industries, Ltd., Yodogawa-ku, Osaka 532, Japan.
§ Present address: Laboratory of Food Chemistry, Institute of Scientific and Industrial Research, Osaka University, Ibaraki, Osaka 567, Japan.
‖ Present address: Laboratory of Applied Gene Chemistry, Faculty of Agriculture, Okayama University, Okayama 700, Japan.

0006-2960/90/0429-1009$02.50/0

in YT medium (1% polypeptone, 0.5% yeast extract, and 0.5% NaCl; pH 7.2), with or without 2% agar, containing appropriate antibiotics (50 μg/mL ampicillin and/or 15 μg/mL tetracycline).

Materials. The plasmid pICD301 containing the gene for thermostable alanine dehydrogenase from a moderate thermophile, *B. stearothermophilus* IFO12550, has been constructed by Nagata and Sakamoto of this laboratory (unpublished results). The thermostable enzyme was purified to homogeneity from cell extracts of *E. coli* C600/pICD301. All restriction endonucleases and T4 DNA ligase were obtained from Takara Shuzo (Kyoto, Japan). The other chemicals were of highest purity available.

DNA Cloning, Mapping, and Sequencing. The chromosomal DNA was isolated from *B. sphaericus* according to the method of Saito and Miura (1963). Plasmids were prepared by the method of Birnboim and Doly (1979). Chromosomal DNA (2 μg) from *B. sphaericus* was partially digested with *Hin*dIII, and the resultant fragments were ligated into the *Hin*dIII site of pBR322 (0.5 μg) and then used to transform *E. coli* C600 cells (Mandel & Higa, 1970). To identify the clones expressing alanine dehydrogenase, the replica-printing assay (Nagata et al., 1988) was employed, by incubation of the filter papers at 37 °C for 10 min in 1.5 mL of the reaction mixture containing 50 mM L-alanine, 50 mM glycine–KOH buffer (pH 10.5), 0.625 mM NAD^+, 0.064 mM phenazine methosulfate, and 0.24 mM nitroblue tetrazolium. The positive colonies were detected as dark blue spots on the replica disk.

DNA sequence was determined by the chain termination method using an M13 sequencing kit (Takara Shuzo) after cloning into bacteriophage M13 mp18 and mp19 sequencing vectors (Messing, 1983). Several regions of the sequence were also determined by selective priming with the oligonucleotides synthesized on an Applied Biosystems 381A DNA synthesizer.

Enzyme and Protein Assays. Alanine dehydrogenase was assayed at 25 °C as previously described (Ohshima & Soda, 1979). One unit of the enzyme is defined as the amount of enzyme that catalyzes the formation of 1 μmol of NADH/min in oxidative deamination. Specific activity is expressed as units per milligram of protein. Protein concentration was measured by the method of Lowry et al. (1951), with crystalline bovine serum albumin as standard.

Purification of Alanine Dehydrogenase from E. coli C600/pICD322 Cells. E. coli C600 carrying pICD322 (see below) was cultivated at 37 °C for 16 h in 180 L of YT medium containing 50 mg/L ampicillin. After preparation of cell-free extract and the first ion-exchange chromatography, which were carried out as previously described for purification of the enzyme from *B. sphaericus* (Ohshima & Soda, 1979), ammonium sulfate was added to the concentrated enzyme solution (200 mL) to give 40% saturation. The enzyme solution was applied to a column (3 × 30 cm) of butyl-Toyopearl 650M equilibrated with 10 mM potassium phosphate buffer (pH 7.2) containing 40% saturated ammonium sulfate and was eluted with a linear gradient of 40–0% saturated ammonium sulfate. The pooled active fractions were dialyzed against 500 volumes of 20 mM Tris-HCl buffer (pH 8.0) containing 0.01% (v/v) 2-mercaptoethanol, concentrated by ultrafiltration to about 15 mL, and applied to a Mono Q ion-exchange column in a fast protein liquid chromatography system (FPLC; Pharmacia, Piscataway, NJ). The enzyme was eluted with a linear gradient of NaCl (0.3–0.4 M) in 20 mM Tris-HCl buffer. The active fractions were combined, concentrated to 2 mL, and further purified by gel filtration with two tandemly connected Superose 12 columns in FPLC, using 0.15 M Tris-HCl buffer (pH 8.0) containing 0.15 M NaCl and 0.01% 2-mercaptoethanol as an eluting buffer. The enzyme obtained was homogeneous by polyacrylamide gel electrophoresis.

Amino Acid Analysis. The purified enzyme (1 nmol) dissolved in distilled water was lyophilized and hydrolyzed at 110 °C in 1 mL of 6 N HCl for 24, 48, and 72 h in sealed and evacuated tubes. After the hydrolysates were evaporated to dryness, the residue was dissolved in 0.2 M sodium citrate buffer (pH 2.2) and analyzed with a Beckman 7300 high-performance amino acid analyzer. Half-cystine was determined as cysteic acid after performic acid oxidation and hydrolysis. Tryptophan and tyrosine were determined spectrophotometrically by the method of Edelhoch (1967).

N-Terminal Sequence Analysis. The purified enzyme [about 2 nmol in 1 mM potassium phosphate buffer (pH 7.2)] was used directly for the N-terminal sequence analysis by automated Edman degradation with an Applied Biosystems 470A gas–liquid-phase protein sequencer. The phenylthiohydantoin (PTH) amino acids were separated and identified by an on-line PTH analyzer Model 120A (Applied Biosystems) with a PTH-C_{18} column.

Trypsin Digestion, Isolation of Peptide Fragments, and Peptide Sequencing. Digestion of the purified alanine dehydrogenases with bovine pancreas trypsin, isolation of tryptic peptides by reverse-phase high-performance liquid chromatography, and amino acid sequencing of the tryptic peptides were carried out as described previously for leucine dehydrogenase (Nagata et al., 1988).

Computer Search for Sequence Similarities. A search of the National Biomedical Research Foundation (NBRF) protein sequence data bank (George et al., 1986) for sequence similarities was carried out on an IBM PC-AT personal computer with a Beckman MICROGENIE program.

Results and Discussion

Cloning of the Alanine Dehydrogenase Gene of B. sphaericus. We recently cloned the gene for the thermostable alanine dehydrogenase of a thermophile, *B. stearothermophilus*, into *E. coli* with pBR322 as the vector (Nagata et al., unpublished results). In the present studies, the gene for the alanine dehydrogenase with lower thermostability (Ohshima & Soda, 1979) was similarly cloned by using *Hin*dIII digests of the chromosomal DNA from a mesophilic bacterium, *B. sphaericus*. Of approximately 11 000 Amp^rTc^s transformants obtained, three clones showed strong dark blue color on the replica disks in activity staining. The plasmids isolated from each clone generated the same physical map with several restriction endonucleases and were designated pICD321 [10.9 kilobase pairs (kb)]. *Hin*dIII digestion of pICD321 produced four fragments of 4.36, 3.81, 2.56, and 0.17 kb, and *Eco*RI digestion yielded five fragments of 5.05, 3.23, 1.24, 1.20, and 0.17 kb. Southern blot analysis indicated that three *Hin*dIII fragments (3.81, 2.56, and 0.17 kb) are derived from the *B. sphaericus* chromosomal DNA (data not shown). To locate the alanine dehydrogenase gene in pICD321, it was subcloned as follows: about 5 μg of pICD321 was digested completely with *Hin*dIII and *Eco*RI, and the resulting DNA fragments were separated and religated between the *Eco*RI and *Hin*dIII sites of pBR322. As a result, a 3.10-kp *Eco*RI–*Hin*dIII fragment was found to contain the entire gene for the expression of alanine dehydrogenase, and the plasmid containing this fragment was designated pICD322 (7.50 kb; Figure 1, left).

Purification of Alanine Dehydrogenase from E. coli C600/pICD322. Since the expression level of alanine de-

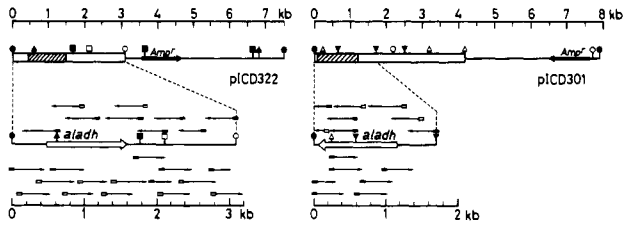

FIGURE 1: Cleavage maps and DNA sequencing schemes for the 3.10-kb *Eco*RI–*Hin*dIII fragment of pICD322 (*B. sphaericus*, left) and the 1.70-kb *Hin*dIII–*Sac*I fragment of pICD301 (*B. stearothermophilus*, right). The single line represents pBR322 DNA. The open bar indicates the genomic insert, and the hatched region indicates the coding region for the alanine dehydrogenase genes. Restriction endonucleases are *Hin*dIII (●), *Eco*RI (○), *Hin*cII (■), *Kpn*I (□), *Sph*I (▲), *Sal*I (△), and *Sac*I (▼). The restriction fragments were subcloned into M13 vectors, and the resulting single-stranded templates were sequenced by the dideoxy chain termination method with the universal sequencing primers (solid bars). The synthesized site-specific primers (open bars) were used for sequencing of double-stranded templates.

hydrogenase in *E. coli* C600 carrying pICD322 was unexpectedly lower (about 0.4% of the total soluble protein) than that in *E. coli* C600 carrying pICD301 (about 4%), which contains the thermostable alanine dehydrogenase gene of *B. stearothermophilus*, we cultivated the *E. coli* C600/pICD322 cells on a large scale (180 L of YT medium) for enzyme purification. The enzyme was purified to homogeneity in five steps (see Experimental Procedures), with a final yield of 6% of the total enzyme activity (52 mg of protein).

Properties of Alanine Dehydrogenase from E. coli C600/pICD322. The molecular weight of the purified alanine dehydrogenase was estimated to be 235 000 by the gel filtration method, using tandemly connected Superose 12 columns in FPLC. In sodium lauryl sulfate/polyacrylamide gel electrophoresis, the enzyme migrated as a single polypeptide, the size of which was estimated to be 39 500 on the basis of its mobility in the gel. On the basis of these results, alanine dehydrogenase is composed of six subunits with identical molecular weights, which coincides with the subunit structure of the thermostable enzyme from *B. stearothermophilus* (Nagata et al., unpublished results). In addition to these structural properties, its catalytic properties, including kinetic parameters for substrates and the coenzyme, were identical with those of the enzyme purified from the original bacterium, *B. sphaericus* (Ohshima & Soda, 1979) (data not shown). The amino acid composition of the enzyme is summarized in Table I. The predominant residues of the enzyme protein were alanine, glutamic acid, valine, and glycine, whereas only a few half-cystine and tryptophan residues were found in the enzyme. The analytic data are in good agreement with the amino acid composition predicted from the DNA sequence of the enzyme gene.

DNA Sequence of the Alanine Dehydrogenase Gene of B. sphaericus. Sanger analysis of the M13 recombinants containing the *Taq*I and *Sau*3AI digests provided an initial set of five continuous, nonoverlapping sequences corresponding to nearly 50% of the total length of the 3.10-kb *Eco*RI–*Hin*dIII fragment of pICD322. After sequence determination with 8 directional recombinants and 11 specific primers for double-stranded pICD322, sequencing of the entire region was completed in both directions (Figure 1, left). The open reading frame of the alanine dehydrogenase gene comprises 1116 bp encoding 372 amino acids, whose ATG codon locates 491 bp downstream of the unique *Hin*dIII site (Figure 2, left). We identified 23 residues from the N terminus of the enzyme protein by automated Edman degradation of the purified gene product: Met-Lys-Ile-Gly-Ile-Pro-Lys-Glu-Ile-Lys-Asn-Asn-Glu-Asn-Arg-Val-Ala-Met-Thr-Pro-Ala-Gly-Val. The N-terminal sequence exactly matched that deduced from the DNA sequence. About 50% of the entire amino acid sequence, including the C-terminal region, was also confirmed by amino acid sequence analyses of 12 peptides obtained by trypsin digestion and purification by HPLC. These peptides are shown at their positions in the amino acid sequence (Figure 2, left). Thus, we identified the 1116-bp open reading frame as the translational unit for alanine dehydrogenase. On the basis of the deduced amino acid sequence (372 residues), the molecular weight of the enzyme subunit is calculated as 39 465, which corresponds to the value (39 500) determined by sodium lauryl sulfate/polyacrylamide gel electrophoresis of the protein (see above).

Table I: Amino Acid Composition of Alanine Dehydrogenases from *B. sphaericus* and *B. stearothermophilus*

	B. sphaericus		*B. stearothermophilus*	
amino acid	predicted	observed	predicted	observed
Cys	2	2	3	3
Asn	15		13	
Asx		34		30
Asp	15		17	
Thr	22	23	28	27
Ser	16	17	13	15
Gln	15		6	
Glx		38		30
Glu	22		23	
Pro	17	16	16	19
Gly	38	37	40	37
Ala	44	44	40	40
Val	36	37	40	44
Met	11	12	10	9
Ile	28	26	21	20
Leu	28	31	32	33
Tyr	15	12	12	10
Phe	5	6	8	8
His	10	9	12	9
Lys	22	22	18	17
Arg	10	10	18	19
Trp	1	1	2	3

DNA Sequence of the Alanine Dehydrogenase Gene of B. stearothermophilus. Alanine dehydrogenase of *B. stearothermophilus* shows higher thermostability than the enzyme of *B. sphaericus* (Ohshima & Soda, 1979). Thus, the *B. stearothermophilus* enzyme retained about 50% of its initial activity when heated at 85 °C for 5 min in 50 mM potassium phosphate buffer (pH 7.2), whereas the *B. sphaericus* enzyme lost the same activity when heated at only 65 °C for 5 min. To correlate the difference in thermostability between the two alanine dehydrogenases with the difference in their primary structures, we sequenced the gene of *B. stearothermophilus* in the plasmid pICD301 (7.90 kb). This plasmid consists of a 4.20-kb *Hin*dIII–*Sal*I fragment of *B. stearothermophilus* chromosomal DNA and the 3.70-kb *Hin*dIII–*Sal*I fragment of pBR322 (Figure 1, right). Because subcloning analysis showed that deletion of the 0.29-kb *Hin*dIII–*Sal*I fragment at the end of the inserted fragment from pICD301 led to loss of the enzyme activity in the recombinant *E. coli* cells, we started the sequence determination with two directional clones for the 0.29-kb fragment. After sequence analysis of the M13 recombinants from *Taq*I, *Sau*3AI and *Hpa*II digests allowed two overlapping sequences corresponding to about 40% of the total length of the 1.74-kb *Hin*dIII–*Sac*I fragment of pICD301. Further analysis with seven directional clones and four specific primers provided the complete sequence in both directions (Figure 1, right). The open reading frame of the alanine dehydrogenase gene includes 1116 bp encoding 372

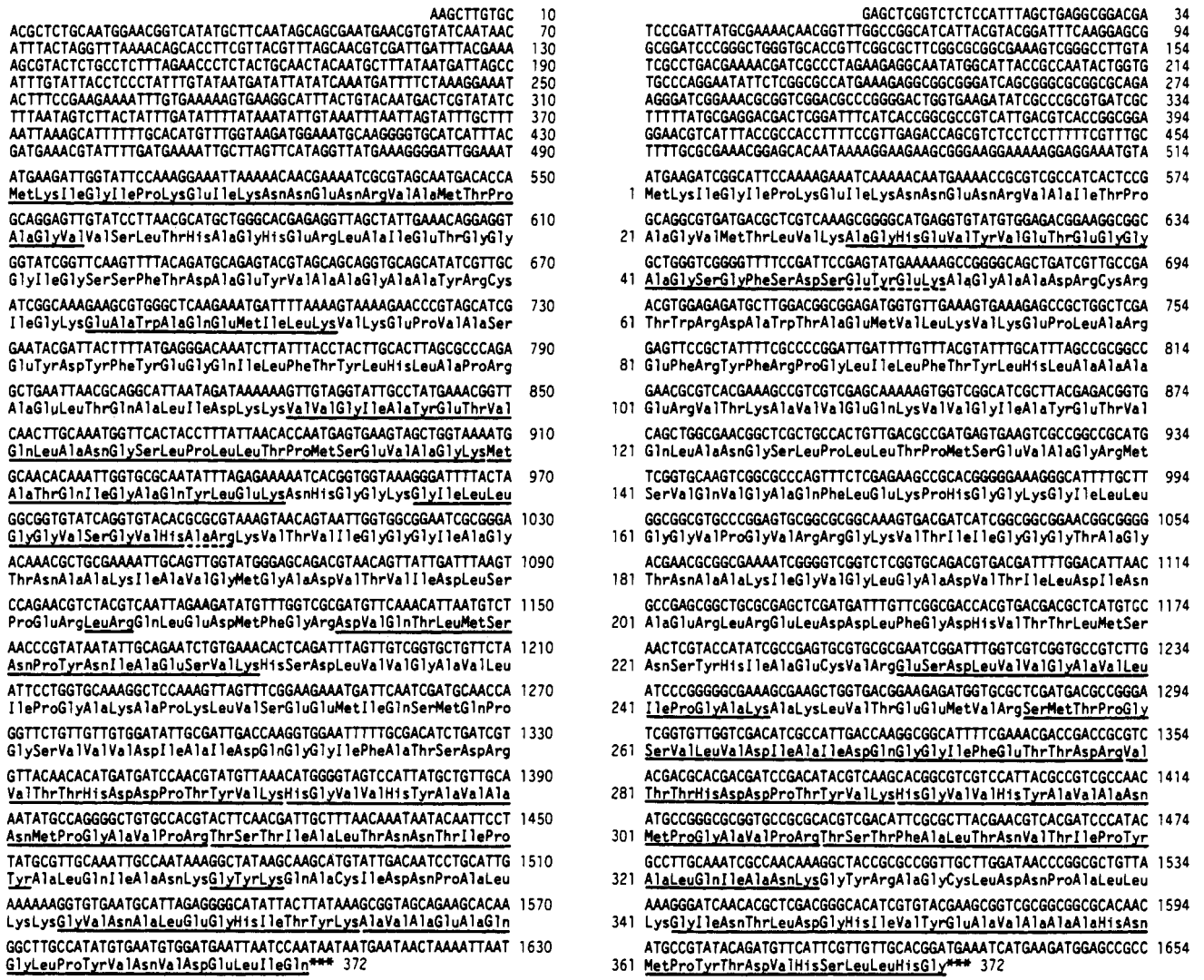

FIGURE 2: DNA sequences of the gene and the deduced amino acid sequences for alanine dehydrogenases from *B. sphaericus* (left) and *B. stearothermophilus* (right). Tryptic peptides, as well as the N-terminal region of the *B. sphaericus* enzyme, for which sequences were determined from the purified gene product by automated Edman degradation, are underlined.

amino acids, whose ATG codon locates 515 bp downstream of the *Sac*I site (Figure 2, right); the total number of nucleotides in the coding region is identical with that of the gene for the *B. sphaericus* enzyme. The high similarity of the amino acid sequence to that of the *B. sphaericus* enzyme (see below) also enabled us to identify the open reading frame accurately, and amino acid sequencing of seven tryptic peptides (Figure 2, right) confirmed about 40% of the entire sequence. On the basis of the deduced amino acid sequence of the thermostable alanine dehydrogenase (372 amino acids), the molecular weight of the enzyme subunit is calculated as 39 699, which agrees well with the value (40 000) determined experimentally for the enzyme purified from the *E. coli* clone cell extract. In addition, the amino acid composition estimated from the deduced sequence is in good agreement with the analytic value of the enzyme purified from the transformant (Table I).

Comparison of the Alanine Dehydrogenase Genes of B. sphaericus and B. stearothermophilus. The open reading frame of the alanine dehydrogenase gene of *B. stearothermophilus* shows a higher GC content (59.6 mol %) than that of *B. sphaericus* (40.8 mol %), both of which correspond to the genomic GC contents of *B. stearothermophilus* (52–58 mol %) and *B. sphaericus* (36.5–37 mol %). However, G or C is preferentially used at the third base of the codons in the *B. stearothermophilus* gene: 75.0 mol % in the *B. stearothermophilus* codons and 25.5 mol % in the *B. sphaericus* codons. Changes from A/T to G/C in DNA sequences, particularly in the third position of each codon, are believed to be one of the mechanisms of gene stabilization at high temperatures (Kagawa et al., 1984).

Amino Acid Sequence Comparison of the Alanine Dehydrogenases of B. sphaericus and B. stearothermophilus. In addition to the size parity of polypeptides and the identical hexameric subunit structure, the hydropathy profiles (Kyte & Doolittle, 1982) and the predicted secondary structures (Chou & Fasman, 1978) for each alanine dehydrogenase sequence showed no significant differences between the two enzymes (data not shown). Linear alignment of the two sequences on the basis of a mutation data scoring matrix (Dayhoff et al., 1983) revealed marked similarity between the two enzymes (Figure 3); 272 residues are identical of 375 possible residue matches. The overall sequence homology is calculated as 73%, which is considerably high for the degree of interspecies sequence homology between thermostable and thermolabile enzymes (cf. 54% homology between alanine racemases from *B. stearothermophilus* and *B. subtilis*; Tanizawa et al., 1988), strongly suggesting that the two alanine dehydrogenases have very similar three-dimensional structures.

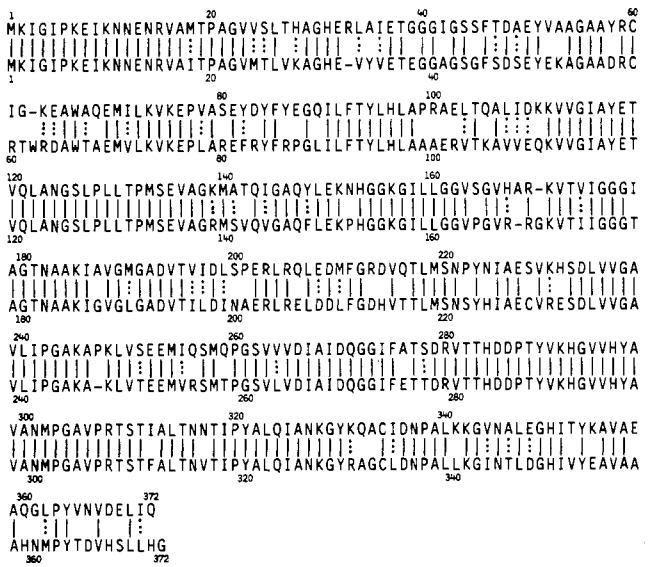

FIGURE 3: Linear alignment of the protein sequences of alanine dehydrogenases from *B. sphaericus* (upper) and *B. stearothermophilus* (lower). The two sequences were aligned by introducing gaps (hyphens) to maximize identities according to the method of Dayhoff et al. (1983). Identical residues and conservative substitutions between the two sequences are shown by verticle lines and dotted lines, respectively.

In the aligned sequences, two portions (1–32 and 90–340 of the *B. sphaericus* sequence) that comprise about 75% of the entire sequence show particularly high similarity (>80%). Several residues with catalytic importance and those identified to be involved in binding of the coenzyme (see below) are all conserved in this highly similar portion (90–340). The remaining two portions (33–89 and 341–372 of the *B. sphaericus* sequence) are less similar (<60%), indicating that most of the nonidentical residues are clustered in these short segments. In view of the fact that the only significant difference between the two alanine dehydrogenases is thermostability, we speculate that either or both of the two segments with low similarity might contribute a role in the thermostability of alanine dehydrogenases. This is consistent with the idea that a considerable increase in thermal resistance of proteins can be acquired by the addition of only a few intramolecular bonds, such as hydrogen, ionic, and hydrophobic bonds (Hocking & Harris, 1976). Studies of the temperature dependence of enzyme activity and the heat inactivation of alanine dehydrogenases from *B. subtilis* and *Thermus thermophilus* (Väli et al., 1980) have also indicated that the factors ensuring the thermoresistance of the *T. thermophilus* enzyme do not affect enzyme activity, and therefore the structural differences between the two enzymes are subtle.

We previously showed the presence of a free thiol group(s) in the *B. sphaericus* alanine dehydrogenase, modification of which by thiol blocking agents [e.g., *p*-(chloromercuri)benzoate and $HgCl_2$] led to reversible inactivation of the enzyme (Ohshima et al., 1979). The *B. sphaericus* enzyme contains only two cysteinyl residues at positions 60 and 334 (Figure 2, left). Therefore, there is probably no intramolecular disulfide bond in the *B. sphaericus* enzyme. It is interesting to note that the more thermostable enzyme of *B. stearothermophilus* contains an additional cysteinyl residue (at position 228) besides the two residues located at positions (59 and 333) equivalent to those in the *B. sphaericus* enzyme. This contrasts with the lack of cysteine in the *T. thermophilus* enzyme with very high thermostability (Väli et al., 1980). Because free cysteines occurring in the exterior of proteins are a potential source of thermal instability (Hocking & Harris, 1976), Cys-228 of the *B. stearothermophilus* enzyme may occur in the interior of the protein and/or form a disulfide bond with either Cys-59 or Cys-333, exerting a positive effect on the thermostability.

Prediction of the Coenzyme Binding Domain by Sequence Comparison with Other NAD(P)$^+$-Dependent Dehydrogenases. Although a computer-aided search of the National Biomedical Research Foundation (NBRF) protein sequence data bank (George et al., 1986) revealed rather low overall sequence similarities (11–20%) between alanine dehydrogenases and other pyridine nucleotide-dependent dehydrogenases, a partial sequence of 29 residues (from Lys-170 to Asp-198) in each alanine dehydrogenase was found to contain many residues identical with those in the nicotinamide coenzyme binding region of various dehydrogenases (Figure 4). The coenzyme binding domain of dehydrogenases has been shown to display a high degree of conservation of tertiary structure (Birktoft & Banaszak, 1984): it consists of a four-stranded parallel β-sheet and one α-helix with virtually identical arrangement. In the dehydrogenases acting on amino acids (i.e., glutamate and leucine dehydrogenases), however, only two parallel β-sheets and one α-helix have been identified as the binding domain for the coenzyme, in particular, for the adenine nucleotide moiety (Birktoft & Banaszak, 1984; Nagata et al., 1988). The similarity of the 29-residue sequence with the coenzyme binding domain of other dehydrogenases is very high, giving scores of 11 and 10 for the sequences of the *B. sphaericus* and *B. stearothermophilus* enzymes, respectively, in prediction of the occurrence of the ADP-binding $\beta\alpha\beta$-fold in proteins by the amino acid sequence fingerprint method

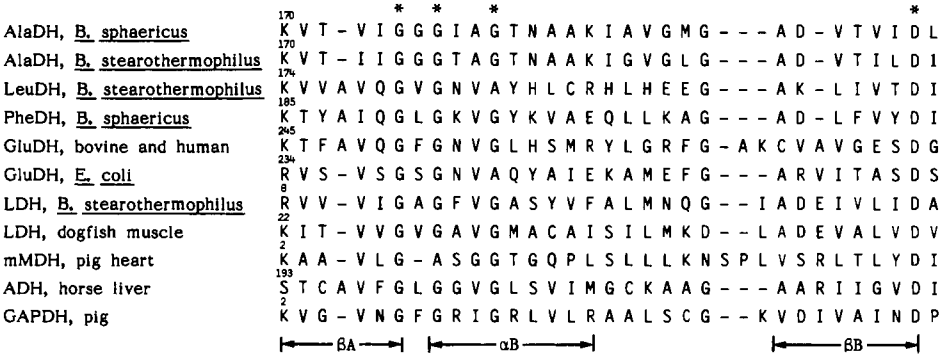

FIGURE 4: Sequence comparison of the βA–αB–βB regions of several NAD(P)$^+$-dependent dehydrogenases. Conserved residues that are believed to be important for coenzyme binding are indicated by asterisks. The sequence numbers of the first and last residues involved in this region are shown above each sequence. Abbreviations: AlaDH, alanine dehydrogenase; LeuDH, leucine dehydrogenase; PheDH, phenylalanine dehydrogenase; GluDH, glutamate dehydrogenase; LDH, lactate dehydrogenase; mMDH, mitochondrial malate dehydrogenase; ADH, alcohol dehydrogenase; GAPDH, glyceraldehyde-3-phosphate dehydrogenase.

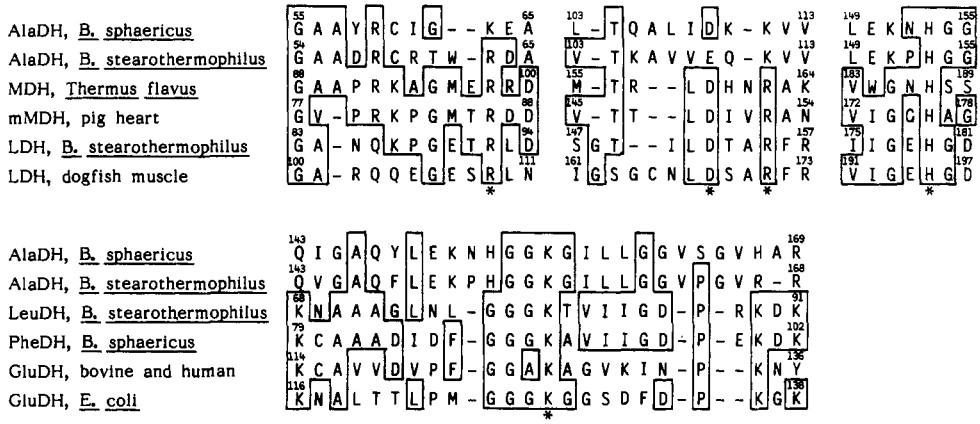

FIGURE 5: Comparison of partial sequences of alanine dehydrogenases with the catalytic regions of other NAD(P)$^+$-dependent dehydrogenases. The sequences were aligned by introducing gaps (hyphens) to maximize identities. Numbers of the first and last residues are referred to the sequences registered in the data bank and are shown above each sequence. Catalytically important residues are shown by asterisks, and conserved residues are blocked. Abbreviations are the same as used in Figure 4.

(Wierenga et al., 1986). These findings support the suggestion that this region of alanine dehydrogenases participates in the binding of NAD$^+$ in a manner analogous to other NAD(P)$^+$-dependent dehydrogenases.

Further Decipherment of the Alanine Dehydrogenase Sequence. By a computer search for local sequence homology, three segments (55–65, 103–113, and 143–169 of the *B. sphaericus* sequence) were found to be significantly similar to partial sequences from other dehydrogenases (Figure 5). The enzymes chosen from the data bank include lactate and malate dehydrogenases, which share the reaction mechanism and catalytically important residues in the highly similar primary structures (Wilks et al., 1988). Interestingly enough, His-195 (an acid–base catalyst; Parker & Holbrook, 1977), Asp-168 (an anionic group supporting the protonated form of His-195 in the ternary complex; Clarke et al., 1988), Arg-109 (a cationic group stabilizing the negative charge on the substrate C2 oxygen during a transition state; Clarke et al., 1986), and Arg-171 (a cationic group providing a tight and bifurcated binding with the C1 carboxylate of the substrate; Hart et al., 1987) of lactate dehydrogenase are found in these segments from both alanine dehydrogenase sequences at matched positions with similar mutual distances (except that Lys is substituted for Arg at the position corresponding to Arg-171 of lactate dehydrogenase). According to Grimshaw et al. (1981), a cationic acid group (probably histidine) plays an essential role as an acid–base catalyst in alanine dehydrogenase, unlike glutamate dehydrogenase which appears to employ a lysyl residue (Rife & Cleland, 1980a,b). Thus, it seems probable that the essential histidine corresponds to His-153 of the *B. sphaericus* and *B. stearothermophilus* alanine dehydrogenases (which is matched to His-195 of lactate dehydrogenase; see Figure 5).

Furthermore, the longest segment (143–169 of the *B. sphaericus* alanine dehydrogenase sequence) compared in Figure 5 contains a tetrapeptide sequence (from Gly-154 to Gly-157) analogous to the consensus sequence, G-G-(G or A or S)-K, uniquely conserved in glutamate (Piszkiewicz et al., 1970; Blumenthal et al., 1975) and leucine dehydrogenases (Nagata et al., 1988); such a motif does not occur in the sequences of lactate and malate dehydrogenases. The lysyl residue in the Gly-rich region of glutamate dehydrogenases probably functions in catalysis by forming a Schiff base adduct with α-ketoglutarate (Smith et al., 1975). Alanine dehydrogenase of *B. sphaericus* contains no reactive lysyl residues (Ohshima & Soda, 1979) whereas leucine dehydrogenase, the sequence of which is similar to glutamate dehydrogenase sequences (Nagata et al., 1988), suffers from chemical modification of a reactive lysyl residue by pyridoxal 5′-phosphate (Ohshima & Soda, 1984). These facts suggest that Lys-156 of alanine dehydrogenase, which is matched to the essential Lys-126 of mammalian glutamate dehydrogenase, is not involved in catalysis. However, Grimshaw et al. (1981) suggested the presence of a cationic acid group (probably lysine) in alanine dehydrogenase which must be protonated for activity and good binding of inhibitors and which appears to be important for maintaining the proper conformation of the enzyme.

In conclusion, we have shown here by complete sequencing and detailed sequence decipherment that alanine dehydrogenases from *B. sphaericus* and *B. stearothermophilus* comprise many residues which are of interest as targets for site-directed mutagenesis in structural studies of amino acid dehydrogenases.

ACKNOWLEDGMENTS

We thank Dr. S. Kuramitsu and Prof. H. Kagamiyama, Osaka Medical College, for amino acid sequence analysis and Drs. N. Esaki, T. Ohshima, and S. Nagata of this laboratory and Dr. S. A. Benner, ETH Zurich, for their helpful discussions.

Registry No. DNA (*Bacillus sphaericus* alanine dehydrogenase gene), 124153-11-3; alanine dehydrogenase (*Bacillus sphaericus* protein moiety reduced), 124153-10-2; DNA (*Bacillus stearothermophilus* alanine dehydrogenase gene), 124153-12-4; alanine dehydrogenase (*Bacillus stearothermophilus* protein moiety reduced), 124153-09-9; alanine dehydrogenase, 9029-06-5.

REFERENCES

Birktoft, J. J., & Banaszak, L. J. (1984) *Pept. Protein Rev. 4*, 1–46.

Birnboim, H. C., & Doly, J. (1979) *Nucleic Acids Res. 7*, 2251–2266.

Blumenthal, K. M., Moon, K., & Smith, E. L. (1975) *J. Biol. Chem. 250*, 3644–3654.

Chou, P. Y., & Fasman, G. D. (1978) *Adv. Enzymol. Relat. Areas Mol. Biol. 47*, 45–148.

Clarke, A. R., Wigley, D. B., Chia, W. N., Barstow, D. A., Atkinson, T., & Holbrook, J. J. (1986) *Nature (London) 324*, 699–702.

Clarke, A. R., Wilks, H. M., Barstow, D. A., Atkinson, T., Chia, W. N., & Holbrook, J. J. (1988) *Biochemistry 27*, 1617–1622.

Dayhoff, M. O., Barker, W. C., & Hunt, L. T. (1983) *Methods Enzymol. 91*, 524–545.
Edelhoch, H. (1967) *Biochemistry 6*, 1948–1954.
George, D. G., Barker, W. C., & Hunt, L. T. (1986) *Nucleic Acids Res. 14*, 11–15.
Germano, G. J., & Anderson, K. E. (1968) *J. Bacteriol. 96*, 55–60.
Grimshaw, C. E., & Cleland, W. W. (1981) *Biochemistry 20*, 5650–5655.
Grimshaw, C. E., Cook, P. F., & Cleland, W. W. (1981) *Biochemistry 20*, 5655–5661.
Hart, K. W., Clarke, A. R., Chia, W. N., Barstow, D. A., Atkinson, T., & Holbrook, J. J. (1987) *Biochem. Biophys. Res. Commun. 146*, 346–353.
Hocking, J. D., & Harris, J. I. (1976) in *Enzymes and Proteins from Thermophilic Microorganisms* (Zuber, H., Ed.) Experimentia Supplementum 26, pp 121–133, Birckhauser Verlag, Basel.
Holmes, P. K., Dundas, I. E., & Halvorson, H. O. (1965) *J. Bacteriol. 90*, 1159–1160.
Hong, M. M., Shen, S. C., & Braunstein, A. E. (1959) *Biochim. Biophys. Acta 36*, 288–289.
Kagawa, Y., Nojima, H., Nukiwa, N., Ishizuka, M., Nakajima, T., Yasuhara, T., Tanaka, T., & Oshima, T. (1984) *J. Biol. Chem. 259*, 2956–2960.
Kyte, J., & Doolittle, R. F. (1982) *J. Mol. Biol. 157*, 105–132.
Lowry, O. H., Rosebrough, N. J., Farr, A. L., & Randall, R. J. (1951) *J. Biol. Chem. 193*, 265–275.
Mandel, M., & Higa, A. (1970) *J. Mol. Biol. 53*, 159–162.
McCowen, S. M., & Phibbs, P. Y., Jr. (1974) *J. Bacteriol. 118*, 590–597.
Messing, J. (1983) *Methods Enzymol. 101*, 20–78.
Nagata, S., Tanizawa, K., Esaki, N., Sakamoto, Y., Ohshima, T., Tanaka, H., & Soda, K. (1988) *Biochemistry 27*, 9056–9062.
Nitta, Y., Yasuda, Y., Tochikubo, K., & Hachisuka, Y. (1974) *J. Bacteriol. 117*, 588–592.
O'Conner, R. J., & Halvorson, H. (1961) *Arch. Biochem. Biophys. 91*, 290–299.
Ohshima, T., & Soda, K. (1979) *Eur. J. Biochem. 100*, 29–39.
Ohshima, T., & Soda, K. (1984) *Agric. Biol. Chem. 48*, 349–354.
Okazaki, N., Hibino, Y., Asano, Y., Ohmori, M., Numano, N., & Kondo, K. (1988) *Gene 63*, 337–341.
Parker, D. M., & Holbrook, J. J. (1977) in *Pyridine Nucleotide-Dependent Dehydrogenases* (Sund, H., Ed.) pp 485–502, de Gruyter, New York.
Piszkiewicz, D., Landon, M., & Smith, E. L. (1970) *J. Biol. Chem. 245*, 2622–2626.
Porumb, H., Vancea, D., Muresan, L., Presecan, E., Lascu, I., Petrescu, I., Porumb, T., Pop, R., & Barzu, O. (1987) *J. Biol. Chem. 262*, 4610–4615.
Rife, J. E., & Cleland, W. W. (1980a) *Biochemistry 19*, 2321–2328.
Rife, J. E., & Cleland, W. W. (1980b) *Biochemistry 19*, 2328–2333.
Saito, H., & Miura, K. (1963) *Biochim. Biophys. Acta 72*, 619–629.
Smith, E. L., Austen, B. M., & Nyc, J. F. (1975) *Enzymes (3rd Ed.) 11*, 293–367.
Tanizawa, K., Ohshima, A., Scheidegger, A., Inagaki, K., Tanaka, H., & Soda, K. (1988) *Biochemistry 27*, 1311–1316.
Váli, Z., Kilár, F., Lakatos, S., Venyaminov, S. A., & Závodszky, P. (1980) *Biochim. Biophys. Acta 615*, 34–47.
Wierenga, R. K., Terpstra, P., & Hol, W. G. J. (1986) *J. Mol. Biol. 187*, 101–107.
Wilks, H. M., Hart, K. W., Feeney, R., Dunn, C. R., Muirhead, H., Chia, W. N., Barstow, D. A., Atkinson, T., Clarke, A. N., & Holbrook, J. J. (1988) *Science 242*, 1541–1544.
Yoshida, A., & Freese, E. (1965) *Biochim. Biophys. Acta 96*, 248–262.
Zink, M. W., & Sanwal, B. D. (1962) *Arch. Biochem. Biophys. 99*, 72–77.