Purification and Sequence Analysis of a Novel NADP(H)-Dependent Type III Alcohol Dehydrogenase from *Thermococcus* Strain AN1

DONGHUI LI AND KENNETH J. STEVENSON*

Department of Biological Sciences, The University of Calgary, Calgary, Alberta, Canada T2N 1N4

Received 8 November 1996/Accepted 1 May 1997

An NADP(H)-dependent alcohol dehydrogenase was isolated from the hyperthermophilic archaeon *Thermococcus* strain AN1. This enzyme is a homotetramer with a subunit molecular weight of 46,700. The enzyme oxidizes a series of primary linear alcohols but not methanol. The pH and temperature optima with ethanol as the substrate are 6.8 to 7.0 and 85°C, respectively. The enzyme readily reduced acetaldehyde with NADPH as the cofactor. The gene encoding this enzyme has been cloned and sequenced. An open reading frame of 1,218 bp, starting with ATG and ending with TGA, was identified and corresponded to 406 amino acids. Sequence comparisons show that this *Thermococcus* strain AN1 enzyme has significant homologies with enzymes from the newly defined type III alcohol dehydrogenase family. *Thermococcus* strain AN1 alcohol dehydrogenase is the first archaeal enzyme belonging to this family.

Alcohol dehydrogenases (ADHs [EC 1.1.1.1]) are universal enzymes, widely distributed in all three life domains (3). There are three major distinct types of alcohol dehydrogenases. Type I ADHs (13), exemplified by horse liver ADH (10), were initially termed long chain but now are referred to as medium chain because of the discovery of ADHs of even longer chains (12). They are the best-studied ADHs. With subunits of 350 to 375 residues in size, they are either dimeric or tetrameric and frequently contain zinc at the active sites. Type II short-chain ADHs (18), are nonmetalloenzymes with subunits of 250 residues, as exemplified by *Drosophila* ADH (26). Type III ADHs, exemplified by *Zymomonas* ADH2, are the most recently discovered ADH family and include *Zymomonas mobilis* ADH2 (7, 17), *Saccharomyces cerevisiae* ADH4 (28), *Escherichia coli* t-1,2-propanediol oxido-reductase (POR) (6), *Clostridium acetobutylicum* ADH1 (29), *C. acetobutylicum* butanol dehydrogenases (BDH I and BDHII) (27), and *Bacillus methanolicus* methanol dehydrogenase (MDH) (9). This family of enzymes show a high degree of sequence identity among themselves but are not homologous to either medium-chain zinc-containing or short-chain, non-metal alcohol dehydrogenases.

Among the *Archaea*, a thermostable NAD⁺-dependent alcohol dehydrogenase from *Sulfolobus solfataricus*, a zinc-containing type I alcohol dehydrogenase, has been well studied (2, 21). ADHs have also been purified and characterized from two *Thermococcus* strains: an NAD⁺-specific iron alcohol dehydrogenase from *T. litoralis* (16) and a sulfur-regulated, non-heme iron alcohol dehydrogenase from *Thermococcus* strain ES-1 (15).

In this study, an NADP(H)-dependent alcohol dehydrogenase has been isolated from the hyperthermophilic, sulfur-metabolizing archaeon *Thermococcus* strain AN1 (DSM 2770) (14). The gene encoding this enzyme has been cloned and sequenced, and the amino acid sequence has been elucidated. Sequence comparison revealed that this enzyme belongs to the newly defined type III alcohol dehydrogenase family.

**Purification and properties of *Thermococcus* strain AN1 ADH.** *Thermococcus* strain AN1 was a generous gift from Hugh W. Morgan, Thermophile and Microbial Biochemistry and Biotechnology Unit of the University of Waikato, Hamilton, New Zealand. *Thermococcus* strain AN1 was routinely grown under anaerobic conditions at 80°C in four rubber-stopper-sealed 1,000-ml flasks equipped with a relief valve at the end of the branch side tube with the medium described by Klages and Morgan (14). To harvest the cells, the cell cultures were filtered once through Whatman no. 42 filter paper to remove sulfur particles, collected by centrifugation at 6,000 × g for 15 min at room temperature, and washed twice with standard buffer (50 mM potassium phosphate and 0.1 mM dithiothreitol [DTT] [pH 7.4]). The cells (2.5 g) were suspended in standard buffer at a ratio of 1 g of cells (wet weight) to 25 ml of standard buffer. Cell lysis was achieved by one passage through a French press at a pressure of 1,500 lb/in² and confirmed by phase-contrast microscopy. A cell extract was obtained by centrifugation of cell lysate at 11,000 × g for 15 min at 4°C. ADH activity was routinely monitored in a 1-ml assay mixture containing 0.87 M ethanol and 6 mM NADP⁺ at 80°C. One unit of enzyme activity was defined as the reduction of 1 μmol of NADP⁺ per minute. A 70-fold purification of ADH was performed at room temperature as shown in Table 1. (NH₄)₂SO₄ (6.7 g) was slowly added to give a final (NH₄)₂SO₄ concentration of 1 M. The precipitate was centrifuged at 11,000 × g for 20 min at 4°C, the supernatant was applied to a phenyl-Sepharose CL-4B matrix (1.6 by 10 cm; Pharmacia Fine Chemicals) equilibrated with buffer A [50 mM potassium phosphate, 0.1 mM DTT, 1 M (NH₄)₂SO₄ (pH 7.4)]. The matrix was washed extensively with the standard buffer until the A₂₈₀ of the eluent was 0.05 or less. ADH activity was eluted with Milli Q water. This enzyme fraction was re-established to the concentration of the standard buffer by adding a 1/9 volume of 0.5 M potassium phosphate buffer (pH 7.4) containing 1 mM DTT. This solution was applied directly onto a hydroxyapatite matrix (1.6 by 5 cm; Micro-prep ceramic;
Bio-Rad) equilibrated with the standard buffer. The breakthrough of this column was collected and directly applied to a DEAE-Sepharose CL-6B matrix (1.6 by 10 cm; Pharmacia Biotech) equilibrated with the standard buffer. The matrix was washed extensively with buffer B [50 mM potassium phosphate, 0.1 mM DTT, 0.25 M (NH₄)₂SO₄ (pH 7.4)] until the eluent was 0.05 or less. ADH remained adsorbed to the DEAE matrix and was eluted with 50 ml of buffer C [50 mM potassium phosphate, 0.4 M (NH₄)₂SO₄ (pH 7.4)]. Fractions with ADH activities were pooled, desalted, and concentrated by ultrafiltration with Amicon Centriprep-30 membranes and stored at 4°C. The purified enzyme was a single band on sodium dodecyl sulfate (SDS)-polyacrylamide gel electrophoresis (PAGE) which corresponded to a molecular weight of 46,000. Gel filtration on a fast-performance liquid chromatography Superose 12 column with the native protein revealed a molecular weight of 158,000, which suggests the enzyme has a homotetrameric structure. The enzyme exhibited a broad substrate specificity for primary alcohols. When a 50 mM alcohol concentration was used, the relative enzyme activities (based on an activity towards ethanol of 1) towards ethanol, propanol, and butanol were 1.00, 1.92, 2.53, and 0.92, respectively. The enzyme did not have activity towards methanol, isopropanol, or glycerol. The $K_m$ values for ethanol, propanol, and butanol were 10 mM, 12 mM, and 1.4 mM, respectively. For alcohol oxidation, the enzyme had a pH optimum of 6.8 to 7.0 at 80°C and a temperature optimum of 85°C. The enzyme had only 20% of its maximum activity at 50°C. Compared with other thermophilic enzymes, Thermococcus strain AN1 ADH is less thermostable. At 80°C, it had a half-life of only 16 min in standard buffer with a protein concentration of 50 μg/ml. The enzyme was not oxygen sensitive. All the purification steps were carried out under aerobic conditions. The half-life of the enzyme stored aerobically in standard buffer at a concentration of 50 μg/ml at 4°C, room temperature, or 80°C was 6 to 7 days, 5 to 7 days, or 16 min, respectively. These values were only slightly greater when the enzyme was stored anaerobically. The N terminus of the purified Thermococcus strain AN1 ADH was naturally blocked. Several internal peptide sequences were obtained following the cleavage of the enzyme by clostripain, one of which was TPELNLLLL MAPVEAT.

Cloning and sequence analysis of the gene encoding ADH. A 20-base oligonucleotide probe (AAC ATG GCI CCI GTI GAA/G GC) was designed on the basis of the C-terminal sequence of the above internal peptide (NMAPVEA). Thermococcus strain AN1 genomic DNA was separately digested to completion with a series of restriction endonucleases. The DNA fragments of each digest were subsequently resolved by agarose gel electrophoresis and transferred onto a nylon membrane by a modification of the method described by Sambrook et al. (22) with a vacuum blotter. The oligonucleotide probe was labelled at its 5'-terminus with [γ-32P]ATP and T4 polynucleotide kinase. Hybridization was carried out for 10 h at 30°C according to the procedure described by Sambrook et al. (22). The membrane was washed with solution I (0.3 M NaCl, 0.03 M sodium citrate, 0.5% [wt/vol] SDS [pH 7.0]) at room temperature for 15 min followed by washing with solution II (15 mM NaCl, 1.5 mM sodium citrate, 0.5% [wt/vol] SDS [pH 7.0]) at 48°C for 15 min. Autoradiography was carried out with Kodak X-Omat diagnostic film. A 3-kb EcoRI fragment from the digest of Thermococcus strain AN1 genomic DNA was identified by Southern hybridization and was inserted into the EcoRI site of the plasmid vector pBluescript KS (Stratagene). The ligation mixture was used to transform competent

<table>
<thead>
<tr>
<th>Step</th>
<th>Protein (mg)</th>
<th>Activity (U)</th>
<th>Sp act (U/mg)</th>
<th>Yield (%)</th>
<th>Purification (fold)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Crude extract</td>
<td>475.0</td>
<td>23</td>
<td>0.05</td>
<td>100</td>
<td>1</td>
</tr>
<tr>
<td>Ammonium sulfate</td>
<td>275.0</td>
<td>20</td>
<td>0.07</td>
<td>90</td>
<td>1.5</td>
</tr>
<tr>
<td>Phenyl-Sepharose</td>
<td>20.0</td>
<td>14</td>
<td>0.70</td>
<td>60</td>
<td>15</td>
</tr>
<tr>
<td>Hydroxyapatite</td>
<td>9.0</td>
<td>11</td>
<td>1.22</td>
<td>50</td>
<td>25</td>
</tr>
<tr>
<td>DEAE-Sepharose</td>
<td>2.5</td>
<td>8</td>
<td>3.20</td>
<td>35</td>
<td>70</td>
</tr>
</tbody>
</table>

μM. For ethanol oxidation, the enzyme had a pH optimum of 6.8 to 7.0 at 80°C and a temperature optimum of 85°C. The enzyme had only 20% of its maximum activity at 50°C.

### Table 1. Summary of the purification of Thermococcus strain AN1 ADH

<table>
<thead>
<tr>
<th>Step</th>
<th>Protein (mg)</th>
<th>Activity (U)</th>
<th>Sp act (U/mg)</th>
<th>Yield (%)</th>
<th>Purification (fold)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Crude extract</td>
<td>475.0</td>
<td>23</td>
<td>0.05</td>
<td>100</td>
<td>1</td>
</tr>
<tr>
<td>Ammonium sulfate</td>
<td>275.0</td>
<td>20</td>
<td>0.07</td>
<td>90</td>
<td>1.5</td>
</tr>
<tr>
<td>Phenyl-Sepharose</td>
<td>20.0</td>
<td>14</td>
<td>0.70</td>
<td>60</td>
<td>15</td>
</tr>
<tr>
<td>Hydroxyapatite</td>
<td>9.0</td>
<td>11</td>
<td>1.22</td>
<td>50</td>
<td>25</td>
</tr>
<tr>
<td>DEAE-Sepharose</td>
<td>2.5</td>
<td>8</td>
<td>3.20</td>
<td>35</td>
<td>70</td>
</tr>
</tbody>
</table>
FIG. 3. Alignment of primary structures for type III alcohol dehydrogenases. Residues that are conserved in at least five of eight sequences are shaded. Twenty-five strictly conserved residues are shown in "consensus" and are boxed and shaded. The 15-amino-acid stretch is highlighted in an extended box. Abbreviations: adh-therm, Thermococcus strain AN1 alcohol dehydrogenase; adh1-cloab, C. acetobutylicum alcohol dehydrogenase 1; adh2-zymmo, Z. mobilis alcohol dehydrogenase 2; medh-bacmt, B. methanolicus methanol dehydrogenase; adh4-yeast, S. cerevisiae alcohol dehydrogenase 4; adh-b-cloab, C. acetobutylicum butanol dehydrogenase II; adha-cloab, C. acetobutylicum butanol dehydrogenase I.
E. coli DH5α [supE44 ΔlacU169 (Δ80lacZΔM15) hsdR17 recA1 endA1 gyrA96 thi-1 relA1] cells, which were screened for recombinant plasmids by plating on Luria Bertani agar plates containing ampicillin (60 μg/ml), 5-bromo-4-chloro-3-indolyl-β-D-galactopyranoside (800 μg/plate), and isopropyl-β-D-thiogalactopyranoside (800 μg/plate). Plasmids from white, ampicillin-resistant colonies were analyzed by digesting with EcoRI and performing a Southern hybridization with the oligonucleotide probe. The clone containing an insertion of the 3-kb EcoRI fragment was named pADH-3kb. A subclone of a 600-bp fragment, named pADH-600bp, was constructed by the same approach from a fragment of the HindIII digest of pADH-3kb which hybridized to the oligonucleotide probe. Sequence analysis commenced with pADH-600bp and was carried out by the dideoxy chain termination method (23) with the universal forward (T7) and reverse (T3) primers. The DNA sequence encoding the amino acid sequence used to design the probe was found in pADH-600bp. A stop codon was also found downstream from this sequence. Specific primers were in turn synthesized on the basis of the nucleotide sequence so obtained and were subsequently utilized to determine the remaining ADH sequence in pADH-3kb. The sequence data were analyzed by using the Staden program (25). The entire sequence of the Thermococcus strain AN1 ADH gene was determined for both strands and is presented in Fig. 1. An open reading frame of 1,218 bp was identified, corresponding to 406 amino acids with an aggregate molecular weight of 46,672. The N-terminal sequence of Thermococcus strain AN1 ADH was identified by the homology observed with two other ADHs, isolated from T. litoralis (16) and Thermococcus strain ES-1 (15) (Fig. 2). The coding region of adh is preceded by the sequence GGAGGG, a putative archaean Shine-Dalgarno ribosome-binding site (4, 24) located six bases upstream of the starting codon, ATG. Two consensus sequences were recognized: the ATGA motif and the TTATAA motif, 20 and 44 bases upstream of the starting ATG codon, respectively, corresponding to box B and box A as described by Reiter et al. (19). Three putative terminator sequences, TTCTTT, TTCTT CT, and TTTTTTATT, were found beginning at 8, 25, and 72 bases, respectively, downstream of the stop codon TGA (20). The amino acid sequence used to design the oligonucleotide probe was found near the C terminus of the protein sequence, with two discrepancies at positions 3 and 9 of the peptide. Positions 3 and 9 were glutamine and asparagine, respectively, determined by peptide sequencing but were serine residues derived from the DNA sequence. The deduced amino acid sequence of Thermococcus strain AN1 ADH was used for searching for homologous sequences in the GenBank and protein databases with the BLAST program (1). Significant homologies were found between Thermococcus strain AN1 ADH and the enzymes of the type III ADH family including Z. mobilis ADH2 (7), S. cerevisiae ADH4 (28), E. coli POR (6), C. acetobutylicum ADH1 (29), C. acetobutylicum BDHI and BDHII (27), and B. methanolicus MDH (9); thus, Thermococcus strain AN1 ADH became the eighth member of this newly defined ADH family. A summary of these homologies is shown in Table 2. This family of ADHs are not homologous to either the medium-chain, zinc-containing type I ADHs or the short-chain, non-metal type II ADHs. Alignment of the primary sequence of all the type III ADHs performed with the PredictProtein program of EMBL internet computational services (22a) as presented in Fig. 3. Twenty-five residues are strictly conserved in all eight proteins, including seven glycine residues, four histidine residues, three proline residues, three aspartic acid residues, two glutamic acid residues, two lysine residues, two alanine residues, one serine, and one isoleucine. Comparison with a previous protein sequence alignment for the other seven members of the type III family (27) shows that seven strictly conserved residues were replaced in Thermococcus strain AN1 ADH as follows: P→A, V→I, G→A, T→A, P→S, Y→A, and P→T in the order of N terminus to C terminus. The most significant changes were the replacement of three proline residues with smaller residues (A, S, and T), which could impact the conformation of the protein. Thermostable proteins have been found to have a more compact structure contributed by loop shortening, improved internal packing, and increased subunit interactions (8). Since Thermococcus strain AN1 ADH is the only thermophile enzyme in the type III family, the variations in highly conserved residues present in Thermococcus strain AN1 ADH may play a role in increasing the thermostability of this enzyme by contributing to the above structural aspects. Future expression and mutagenesis studies are needed. In Thermococcus strain AN1 ADH, three of the four strictly conserved histidine residues are in a 16-amino-acid segment (Fig. 3) which is highly homologous to the 15-amino-acid stretch described by Cabisco et al. (5) as being exclusive only to the type III ADHs. These histidine residues were thought to serve as metal binding ligands and be a portion of the active center (5). In addition, substantial homology was found between Thermococcus strain AN1ADH and the C-terminal region of the long-chain multifunctional alcohol dehydrogenases, such as ADHE (alcohol dehydrogenase/acetaldehyde dehydrogenase) from C. acetobutylicum, and ADHE (alcohol dehydrogenase/acetaldehyde dehydrogenase) from E. coli (Swissprot protein sequence database). These proteins were thought to be the evolutionary fusion product of the two original proteins, alcohol dehydrogenase and acetaldehyde dehydrogenase (11). No significant homologies were found when Thermococcus strain AN1 ADH was compared with the primary structures of procaryotic or eucaryotic type I or type II alcohol dehydrogenases. A search in GenBank with the sequence HXMXHXLGAXXXPHG (letters in boldface represent the three highly conserved histidine residues) revealed only the type III and long-chain multifunctional ADHs and related enzymes; thus, this sequence may serve as a consensus signature sequence to probe type III ADH homologs in other organisms with available complete genome data.

**Nucleotide sequence accession number.** The nucleotide sequence of the Thermococcus strain AN1 adh gene has been deposited with GenBank under accession number U72646.

This work was supported by the Natural Sciences and Engineering Research Council of Canada by a grant (OGP0005859) to K.J.S.

We thank Lynn Parker, Thermophile and Microbial Biochemistry and Biotechnology Unit of the University of Waikato, New Zealand, for her generous assistance with the initial culture conditions for Thermococcus strain AN1 in our laboratory and Ron Ronimus of this unit for the codon usage data of Thermococcus strain AN1. The amino acid sequencing was performed by Don McKay, Protein Sequencing Facility, Faculty of Medicine, the University of Calgary, Canada. The sequencing primers were synthesized by University Core DNA Services of the University of Calgary. We also thank George Din for his helpful discussions.

**REFERENCES**


