Amino Acid Sequence Around the Catalytic Site in Glyceraldehyde-3-Phosphate Dehydrogenase from *Bacillus stearothermophilus*

J. BRIDGEN, J. I. HARRIS, P. W. McDONALD, R. E. AMELUNXEN, AND J. R. KIMMEL

*Medical Research Council Laboratory of Molecular Biology, Hills Road, Cambridge, England, Departments of Microbiology and Biochemistry, and the McIlvain Laboratories, University of Kansas Medical Center, Kansas City, Kansas 66103*

Received for publication 24 April 1972

The tryptic peptide containing the active-site cysteine in glyceraldehyde-3-phosphate dehydrogenase from *Bacillus stearothermophilus* 1503 was isolated after inhibition of the enzyme with $^{14}$C-iodoacetate. The amino acid sequence of the 20-residue peptide was determined by 19 successive cycles of dansyl-Edman degradation. The sequence shows considerable homology with its counterparts from mesophilic sources but differs by the addition of Ala-His-His at the N-terminus and by the substitution of phenylalanine for leucine in the prototype sequence.

The mechanism of thermostability for enzymes from obligate thermophilic bacteria is of considerable contemporary interest. Several theories have been proposed in an attempt to explain thermophily, including (i) the existence of specific protective factors within the intact cell (13), (ii) a rapid resynthesis of thermally inactivated cellular components (1, 6), and (iii) the possession of unusual structural features that confer thermal stability on the individual enzyme protein (13).

Studies of a thermostable intracellular enzyme were made possible by the crystallization of homogeneous glyceraldehyde-3-phosphate dehydrogenase [EC 1.2.1.12] from the obligate thermophile *Bacillus stearothermophilus* 1503 (2, 16). Characterization of this enzyme by physicochemical methods (3, 4, 15) has shown that it has several features in common with its counterparts from mesophilic sources. Studies of the primary structure of glyceraldehyde-3-phosphate dehydrogenases from a variety of different mesophilic organisms have shown that the amino acid sequence of the protein subunit of this enzyme is very highly conserved during evolution (5, 9, 11). Moreover the amino acid sequence of a 17-residue peptide containing the catalytically active cysteine residue is essentially identical in all of the 14 species that have been studied (e.g., reference 14, and the findings of W. S. Allison and J. I. Harris, Fed. Eur. Biochem. Soc. Proc. Meet., p. 140, 1965).

As a step towards the identification of the structural elements responsible for the marked differences in stability between mesophilic and thermophilic glyceraldehyde-3-phosphate dehydrogenases, we have investigated the amino acid sequence around the catalytically active cysteine in the enzyme from *B. stearothermophilus*.

**MATERIALS AND METHODS**

**Chemicals.** Iodoacetic-2-$^{14}$C acid was obtained from the Radiochemical Centre, Amersham, England; stock solutions contained 2 mCi/mmole. Iodoacetic-$^{14}$C acid was purchased from New England Nuclear Corp., Boston, Mass.; stock solutions contained 12 mCi/mmole.

**Enzyme preparations.** Glyceraldehyde-3-phosphate dehydrogenase was prepared in crystalline form from *B. stearothermophilus* 1503 as previously described (2, 16).

**Preparation of S-carboxymethylated enzyme and tryptic digestion.** Preparation 1: 60 mg (1.7 μmoles of subunit) of recrystallized enzyme prepared as described by Amelunxen (2) was allowed to react with 2.3 μmoles of iodoacetic-$^{14}$C acid in 0.1 M tris(hydroxymethyl)aminomethane (Tris)-hydrochloride, pH 7.4, for several hours at 4 C. After complete inactivation had been achieved, the solution was made 5 M with respect to guanidine-hydrochloride and left at room temperature for 1 hr. A 50-fold molar excess of $\beta$-mercaptoethanol was then added, followed after 3 hr by nonradioactive iodoacetate in an amount that was sufficient to react with remaining protein thiol groups, as well as excess $\beta$-
mercaptoethanol; reaction was for 1 hr at pH 7.6. The solution was then dialyzed extensively against distilled water, and the precipitated carboxymethylated protein was collected by centrifugation and redissolved in 1 ml of 8 M urea (twice recrystallized from ethanol). Trypsin (2%) was then added and the solution was immediately diluted to a final urea concentration of 2 M (12). Digestion was carried out at pH 9.0 at 25°C (Radiometer pH-stat) and was complete after 2 hr. The digest was then acidified to pH 2.5 and freeze-dried.

Preparation 2: In a separate experiment (cf. reference 8), 100 mg (2.8 moles of subunit) of crystalline enzyme prepared as described by Suzuki and Harris (16) was dissolved in 10 ml of 20 mM Tris-hydrochloride containing 1 mM ethylenediaminetetraacetate, 1 mM β-mercaptoethanol, pH 7.4, and allowed to react with 15 moles of iodoacetic-2,14C until total inhibition of enzyme activity occurred (1 hr at 25°C). Solid guanidine-hydrochloride and 1 M Tris-hydrochloride (pH 8.2) were then added to a final concentration of 6 M guanidine-hydrochloride and 70 mM Tris-hydrochloride, and the carboxymethylation reaction was allowed to continue for an additional 1.5 hr at 25°C. After extensive dialysis against 50 mM NH₄HCO₃ (pH 8.0), the fully S-carboxymethylated protein that contained 2 moles of S-[2,14C] carboxymethylcysteine (cf. reference 11) was completely soluble, and was digested with trypsin (1%) for 4 hr at 37°C and freeze-dried.

Purification of radioactive peptides. Preparation 1 (peptide 1): The freeze-dried digest (containing 0.5 g of urea) was resuspended in 1.5 ml of pyridine-acetate, pH 2.5, transferred to a cation-exchange resin column (Beckman PA-28 resin; 0.9 by 30 cm column) equilibrated at 50°C, and eluted with a linear gradient of pyridine-acetate (250 ml each of 0.2 M pyridine-acetate, pH 3.1, and 2.0 M pyridine-acetate, pH 5.1). Further purification of radioactive components was achieved by anion exchange chromatography (AG-1-X2 resin; 0.9 by 55 cm column). The column was equilibrated with 2 M pyridine at 50°C and developed by using a three-chamber linear gradient, the chambers containing 125 ml of 2 N pyridine, 125 ml of 0.1 N acetic acid, and 125 ml of 3 N acetic acid, respectively. Final purification was achieved by gel filtration on Sephadex G-25 (fine grade; 1 by 100 cm column) in 0.1 N acetic acid.

Preparation 2 (peptide 2): The tryptic digest of S-[2,14C] carboxymethylated enzyme (100 mg) was first fractionated by gel filtration on a column of Sephadex G-50 (fine grade; 2.2 by 140 cm column) in 0.05 M NH₄HCO₃. All of the 14C was eluted as a single symmetrical radioactive peak. The radioactive fraction was freeze-dried, and additional purification was achieved by ion exchange chromatography on a column of Bio-Gel P-100M (Bio-Rad QAE) (25 cm column) and eluted with a linear pyridine-acetate gradient (100 ml each of 0.05 M pyridine-acetate, pH 2.5; 0.5 M pyridine-acetate, pH 3.75; 2 M pyridine-acetate, pH 5.0) at 50°C. Final purification was by anion exchange chromatography (AG-1-X2 resin; 0.9 by 25 cm column) eluted at 50°C with a three-chamber linear gradient (100 ml each of pyridine, 3%, v/v; 0.38 M pyridine-acetate, pH 5.0; acetic acid, 3%, v/v).

Thin-layer electrophoresis-chromatography. Peptide mapping was performed on cellulose (Avicel microcrystalline) plates (20-cm square) by thin-layer procedures. Electrophoresis was carried out at 200 v for 90 min in pyridine-acetic acid-water (1:10:70, v/v/v), pH 3.6, followed by chromatography in butanol-pyridine-acetic acid-water (9:6:1.8:7.2, v/v/v/v).

Amino acid analysis. Radioactive peptides were hydrolyzed in sealed evacuated tubes with glass-distilled, constant boiling HCl at 110°C for 4 hr (peptide 1) or 36 hr (peptide 2) and were analyzed with a Spinco automatic amino acid analyzer (cf. reference 8).

Isolation of thermolysin fragments. The radioactive tryptic peptide from preparation 2 was digested with thermolysin (250 nmoles of enzyme incubated with 0.1 mg of enzyme at 37°C for 4 hr), and resulting fragments were purified by combinations of gel filtration on Sephadex G-25 and chromatography on cation exchange resins. Amide assignments were based on the mobilities of peptides during electrophoresis in pyridine-acetic acid buffer at pH 6.5.

Amino acid sequence analysis. N-terminal residues were identified by dansylation, and sequences were established by the dansyl-Edman method (7) with identification of dansyl-amino acids by thin-layer chromatography on polyamide sheets (17).

RESULTS

Isolation of active-site peptide. The tryptic digest of preparation 1 gave two radioactive fractions on the cation exchanger, one of which emerged unretarded from the column and contained 40% of the total 14C. This fraction, which was also unretarded on the anion exchanger, contained peptides that stained poorly with ninhydrin and were only sparingly soluble in the absence of urea. The radioactive peptide from this fraction did not appear to react with dansyl chloride, and it is probable that carbamylation of its amino groups had occurred when it was exposed to urea during tryptic digestion or ion exchange chromatography, or both. After anion exchange chromatography of peak 2 (which contained the remaining 60% of the 14C), several components were observed by thin-layer electrophoresis and chromatography, only one of which was radioactive. Final purification was achieved by gel filtration on Sephadex G-25; the radioactive peptide showed only one component by thin-layer analyses. Because of the many procedures tried in developing a method for final purification of the peptide, the final yield (about 5% based on the recovery of radioactivity) was sufficient only to determine its amino acid composition.
acid composition. This showed (peptide 1, Table 1) that the active-site tryptic peptide of thermophilic glyceraldehyde-3-phosphate dehydrogenase consists of 20 amino acids.

Material for sequence analysis was obtained from preparation 2. All of the radioactivity chromatographed as a single peak on Sephadex G-50 as well as on the cation and anion exchangers, showing that all of the S-[2-14C] carboxymethylcysteine in the protein occurred in a single tryptic peptide. The pure peptide was obtained in a yield of 450 to 500 nmoles, and its amino acid composition (peptide 2, Table 1) showed that it was essentially identical to peptide 1 except that the two residues of S-carboxymethylcysteine were both radioactive (cf. reference 9). These combined results showed the active-site peptide of glyceraldehyde-3-phosphate dehydrogenase from B. stearothermophilus to be similar in overall composition to its counterparts from mesophilic sources, except that it contains two residues of histidine and one of phenylalanine.

**Sequence analysis.** Alanine was shown to be the N-terminal residue by the dansyl-Edman degradation led to the elucidation of the complete sequence of peptide 2 as given in Table 2. The sequence was confirmed from the results of thermolysin digestion. Four major peptides (one radioactive and therefore containing the two residues of S-carboxymethylcysteine) were obtained, and these were again sequenced by the dansyl-Edman method. Amide assignments were based on the mobilities of peptides at pH 6.5, and the presence of two asparagine residues in peptide 2 (Table 1) was confirmed by amino acid analysis after digestion with aminopeptidase-M.

**DISCUSSION**

Active-site tryptic peptides from glyceraldehyde-3-phosphate dehydrogenases in most cases consisted of 17 residues, comprising positions 4 to 20 in Table 2, and positions 143 to 159 in the sequence of the entire protein chain (9, 11). The sequence of this 17-residue segment has been very highly conserved in all the glyceraldehyde-3-phosphate dehydrogenases that have been studied (11, 14), and also as reported in the findings of W. S. Allison and J. I. Harris (Fed. Eur. Biochem. Soc. Proc. Meet., p. 140, 1965). The peptide from the thermophilic enzyme differs from the prototype sequence only at position 18, where a leucine residue (valine in the lobster muscle enzyme) is

---

**Table 1. Amino acid composition of the active-site peptide of thermophilic glyceraldehyde-3-phosphate dehydrogenase**

<table>
<thead>
<tr>
<th>Amino acid*</th>
<th>Peptide 1</th>
<th>Peptide 2</th>
<th>No. of residues</th>
</tr>
</thead>
<tbody>
<tr>
<td>CMCys</td>
<td>1.6</td>
<td>1.6</td>
<td>2</td>
</tr>
<tr>
<td>Asp</td>
<td>2.0</td>
<td>2.1</td>
<td>2</td>
</tr>
<tr>
<td>Thr</td>
<td>1.9</td>
<td>1.8</td>
<td>2</td>
</tr>
<tr>
<td>Ser</td>
<td>1.7</td>
<td>1.9</td>
<td>2</td>
</tr>
<tr>
<td>Pro</td>
<td>1.0</td>
<td>1.1</td>
<td>1</td>
</tr>
<tr>
<td>Ala</td>
<td>3.8</td>
<td>3.9</td>
<td>4</td>
</tr>
<tr>
<td>Val</td>
<td>1.0</td>
<td>1.1</td>
<td>1</td>
</tr>
<tr>
<td>Ile</td>
<td>1.1</td>
<td>1.0</td>
<td>1</td>
</tr>
<tr>
<td>Leu</td>
<td>1.2</td>
<td>0.9</td>
<td>1</td>
</tr>
<tr>
<td>Phe</td>
<td>1.1</td>
<td>0.9</td>
<td>1</td>
</tr>
<tr>
<td>Lys</td>
<td>1.0</td>
<td>1.0</td>
<td>1</td>
</tr>
<tr>
<td>His</td>
<td>2.0</td>
<td>2.1</td>
<td>2</td>
</tr>
<tr>
<td>N-terminal</td>
<td>Not determined</td>
<td>Alaine</td>
<td>20</td>
</tr>
</tbody>
</table>

*Abbreviations: CMCys, S-carboxymethylcysteine; Asp, aspartic acid; Thr, threonine; Ser, serine; Pro, proline; Ala, alanine; Val, valine; Ile, isoleucine; Leu, leucine; Phe, phenylalanine; Lys, lysine; His, histidine.

* Shows to be asparagine by amino acid analysis after digestion with aminopeptidase-M.

---

**Table 2. Amino acid sequence of the active-site peptide (T1) of thermophilic glyceraldehyde-3-phosphate dehydrogenase and its thermolysin fragments (T1: Th-1 through 4)**

<table>
<thead>
<tr>
<th>Peptide</th>
<th>Sequence*</th>
</tr>
</thead>
<tbody>
<tr>
<td>T1</td>
<td>Ala-His-His-Ile-Val-Ser-Ala-Asp-Pro-Thr-Met-CMCys-Thr-Thr-Asn-CMCys</td>
</tr>
<tr>
<td>T1:Th-1</td>
<td>Ala-His-His</td>
</tr>
<tr>
<td>Th-2</td>
<td>Ile-Val-Ser-Ala-Ser-CMCys-Thr-Asn-CMCys</td>
</tr>
<tr>
<td>Th-3</td>
<td>Leu-Ala-Pro</td>
</tr>
<tr>
<td>Th-4</td>
<td>Phe-Ala-Lys</td>
</tr>
</tbody>
</table>

* Indicates sequence analyses by dansyl-Edman method (7); a box around an amino acid denotes a difference from prototype sequence (8, 14).

* Active-site cysteine (8, 9).
replaced by phenylalanine. Moreover, in most of the glyceraldehyde-3-phosphate dehydrogenases that have been sequenced, position 3 (Table 2) is occupied by a lysine residue. The only known variation among 14 different examples of the prototype sequence is the lobster muscle enzyme (5) in which the lysine residue is replaced by threonine. In the thermophilic enzyme this position is occupied by histidine. The adjacent histidine residue (position 2) occurs in a highly variable part of the protein chain (cf. residues 138 to 143 in reference 9). The presence of two adjacent histidine residues is a novel feature of the active-site sequence, but it remains to be established whether this particular structural feature is in any way involved in the thermal stability of the enzyme. In this connection, it will be of interest to establish whether similar amino acid replacements also occur in active-site peptides derived from other thermophilic glyceraldehyde-3-phosphate dehydrogenases. Further studies to elucidate the mechanism of thermal stability for this enzyme are in progress at Cambridge and involve sequence analysis of the subunit protein chain and a study of the three-dimensional structure by X-ray diffraction analysis of the crystalline enzyme.

ACKNOWLEDGMENTS

This investigation was supported by Public Health Service grant AM-09072 from the National Institute of Arthritis and Metabolic Diseases to J.R.K., and by a grant from the American Cancer Society, Kansas Division, to R.E.A.

We thank Miss Sandra Poindexter for her excellent technical assistance.

LITERATURE CITED