Amino Acid Sequence of Cytochrome c-552 from a Thermophilic Hydrogen-Oxidizing Bacterium, *Hydrogenobacter thermophilus*

YOSHIHIRO SANBONGI, MASAHARU ISHII, YASUO IGARASHI, AND TOHRU KODAMA*

Department of Agricultural Chemistry, University of Tokyo, Bunkyo-ku, Tokyo 113, Japan

Received 11 July 1988/Accepted 6 October 1988

The complete amino acid sequence of cytochrome c-552 from an extremely thermophilic hydrogen bacterium, *Hydrogenobacter thermophilus* TK-6 (IAM 12695), was determined. It is a single polypeptide chain of 80 residues, and its molecular weight, including heme, was calculated to be 7,599. The sequence of cytochrome c-552 from *H. thermophilus* TK-6 closely resembles that of cytochromes c-551 from *Pseudomonas* species. Moreover, the tertiary structure of *Hydrogenobacter* cytochrome c-552 is suggested to be similar to that of cytochrome c-551 from *Pseudomonas aeruginosa*. The sequence similarity between *Hydrogenobacter* cytochrome c-552 and that of other bacteria physiologically related to *H. thermophilus* is not high.

*Hydrogenobacter thermophilus* TK-6 (IAM 12695) is an aerobic, extremely thermophilic hydrogen-oxidizing bacterium (13). The bacterium showed many distinct characteristics. The bacterium is an obligate autotroph (13), which was the first case among hydrogen-oxidizing bacteria. It is also characterized by a unique cellular fatty acid pattern (13) and a new sulfur-containing quinone (12). A kind of reductive tricarboxylic acid cycle was shown to operate as a CO₂ fixation pathway in the cell (15), which was the first confirmed observation that a non-Calvin type CO₂ assimilation pathway works in an aerobic organism. We concluded that this bacterium occupies a unique taxonomic position. Recently, *H. thermophilus* was proved to be able to utilize thiosulfate as a sole energy source under microaerophilic conditions. This fact suggests the high physiological similarity between *H. thermophilus* and thermophilic sulfur-oxidizing archaebacteria (3).

In *H. thermophilus*, cytochrome c-552 is a primary electron acceptor for molecular hydrogen activated by hydrogenase (11). The hydrogenase reaction is the first reaction in electron flow in hydrogen-oxidizing microorganisms (10, 11). Cytochrome c-552 isolated from *H. thermophilus* was believed to have the lowest molecular weight among bacterial cytochromes c ever reported (10).

Amino acid sequences of cytochromes c from various organisms have been determined and classified. Analyses of the primary structure of cytochromes c have contributed to the study of the evolution of the respiratory chain (7). The determination of the amino acid sequence of cytochrome c-552 from *H. thermophilus* is therefore considered to provide essential information on the taxonomic and phylogenetic position of the bacterium and will also help us to understand the structure-function relationships in energy transduction. In the present work we have determined the amino acid sequence of cytochrome c-552 from *H. thermophilus* TK-6 and have compared it with those of other organisms, especially microorganisms physiologically related to this bacterium.

**MATERIALS AND METHODS**

Preparation of cytochrome c-552. The bacterial strain used was *H. thermophilus* TK-6 (registered number, IAM 12695, at the Institute of Applied Microbiology, University of Tokyo). Growth conditions and the methods of cytochrome c-552 extraction and purification were as described previously (10) with a minor modification. The modification was that harvested cells were washed with 50 mM phosphate buffer, pH 7.0, and then cytochrome c-552 was separated from the cell washing by CM-Toyopearl 650S (Toso Corp., Tokyo) and Dye Matrex Green A (Amicon Far East, Ltd.) column chromatographies as previously described (10).

**Enzymes and reagents for sequence study.** Lysyl endopeptidase was purchased from Wako Pure Chemical Industries Ltd., Osaka, Japan. Carboxypeptidase Y was purchased from Sigma Chemical Co., St. Louis, Mo.; cyanogen bromide was purchased from Kanto Chemical Co., Inc., Tokyo, Japan.

**Removal of heme from cytochrome c-552 and carboxymethylation of the apoprotein.** Heme was removed from the protein by the method of Ambler et al. (2). The apoprotein obtained was reduced and carboxymethylated by the method of Crestfield et al. (6) to obtain carboxymethylated cytochrome c-552.

**Proteolytic digestions and purification of the peptides.** The carboxymethylated protein (17 nmol) was digested with lysyl endopeptidase (2.55 μg) in 100 μl of 10 mM Tris hydrochloride buffer (pH 9.0) at 30°C for 7 h. The proteolytic digest was fractionated by reversed-phase high-performance liquid chromatography (HPLC). HPLC was performed on an LC-6A liquid chromatograph (Shimadzu Corp., Kyoto, Japan) on a Senshu Pak C18 column (Senshu Scientific Co., Ltd., Tokyo, Japan). The mobile phase was a mixture of acetonitrile and 0.1% trifluoroacetic acid at a ratio varying from 0.10 to 50:50; the flow rate was 1 ml/min.

**Chemical cleavage and purification of the peptides.** Native cytochrome c-552 (0.25 μmol) was cleaved with cyanogen bromide (25 μmol, 2.65 mg) in 2 ml of 70% formate at 38°C for 24 h. Peptides from cyanogen bromide cleavage of the native protein were separated on the column of Senshu Pak C18 by reversed-phase HPLC. The mobile phase was the same as in the case of the purification of lysyl endopeptidase digestion.

**Sequence study.** Amino acid compositions were determined with a Hitachi 835 amino acid analyzer after 24 to 74 h of acid hydrolysis. The amino acid sequence was mainly determined by an automatic sequencer (Applied BioSystems 470A). The amino-terminal amino acid assignment was also confirmed by manual Edman sequencing. The phenylthiohy-
dantoin derivatives were identified on a column of Nucleosil SC-18 (Senshu Scientific Co., Ltd.) by reversed-phase HPLC. The mobile phase was a mixture of acetonitrile and 37 mM ammonium acetate, pH 4.6, at a ratio varying from 30:70 to 86:14 and was running at 50°C; the flow rate was 1 ml/min. For a determination of the carboxyl-terminal amino acid sequence, carboxypeptidase Y digestion was performed in 0.1 M pyridine acetate, pH 6.5. The substrate and enzyme ratio was 300:1 (mol/mol). Released amino acids after digestion were determined as described above.

RESULTS

Amino acid composition and terminal sequence analysis of cytochrome c-552. The amino acid composition of carboxymethylated cytochrome c-552 is shown in Table 1. Two cysteine residues are present, which are presumably linked to the heme group in the native cytochrome. The amino-terminal sequence was determined by an automatic sequencer up to cycle 53 (Fig. 1). Figure 2 shows the results of carboxypeptidase Y digestion of carboxymethylated cytochrome c-552. Lysine was rapidly released, and subsequently isoleucine, serine, leucine, and again isoleucine were released. Thus, the carboxyl-terminal sequence was deduced as -Ile-Leu-Ser-Ile-Lys.

Isolation and sequence analysis of lysyl endopeptidase peptides. Carboxymethylated cytochrome c-552 was digested with lysyl endopeptidase and applied to the reversed-phase HPLC column; the elution profile is shown in Fig. 3. Each fraction was separately pooled and lyophilized. The amino acid compositions of peptides obtained are shown in Table 1. The sequences of the peptides digested with lysyl endopeptidase were determined by the automatic sequencer. The identified residues are indicated by arrows under the residues in Fig. 1.

For peptide L1, the sequence analysis revealed 22 residues. Amino acid 22 was lysine, which may be the carboxyl terminal of this peptide. For peptide L2, the sequence analysis revealed 10 residues. The sequence of residues 6 through 10 of this peptide was -Ile-Leu-Ser-Ile-Lys. This sequence is in accordance with the results of CNBr cleavage and lysyl endopeptidase digestion.

\[\text{Residue no.} \quad \text{Composition}^a \]

<table>
<thead>
<tr>
<th>Residue no.</th>
<th>Intact CM protein$^b$</th>
<th>L1'</th>
<th>L2'</th>
<th>C1'</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>49-70</td>
<td>71-80</td>
<td>60-80</td>
</tr>
<tr>
<td>Asp</td>
<td>7.6 (5)</td>
<td>2.0 (1)</td>
<td>2.3 (1)</td>
<td></td>
</tr>
<tr>
<td>Glu</td>
<td>7.3 (2)</td>
<td>1.9 (1)</td>
<td>2.1 (1)</td>
<td></td>
</tr>
<tr>
<td>Gly</td>
<td>8.5 (4)</td>
<td>3.0 (3)</td>
<td>2.3 (2)</td>
<td></td>
</tr>
<tr>
<td>Ala</td>
<td>13.3 (12)</td>
<td>2.2 (2)</td>
<td>1.2 (1)</td>
<td></td>
</tr>
<tr>
<td>Val</td>
<td>5.4 (6)</td>
<td>2.4 (3)</td>
<td>1.3 (1)</td>
<td></td>
</tr>
<tr>
<td>Met</td>
<td>1.1 (2)</td>
<td>0.9 (1)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Ile</td>
<td>2.4 (3)</td>
<td>1.3 (2)</td>
<td>1.0 (2)</td>
<td></td>
</tr>
<tr>
<td>Leu</td>
<td>4.6 (5)</td>
<td>1.9 (2)</td>
<td>1.8 (2)</td>
<td></td>
</tr>
<tr>
<td>Tyr</td>
<td>2.7 (3)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Phe</td>
<td>0.9 (1)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>His</td>
<td>11.7 (13)</td>
<td>1.1 (1)</td>
<td>0.9 (1)</td>
<td></td>
</tr>
<tr>
<td>Arg</td>
<td>1.2 (1)</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

$^a$ Numbers in parentheses are residues found in the sequence.

$^b$ Average value obtained from 24, 53, and 74-h hydrolysis with 6 M HCl at 110°C. CM, Carboxymethylated.

$^c$ Taken from 24-h hydrolysis.

$^d$ Taken from 74-h hydrolysis.

$^e$ Trp was not determined in the present study.

FIG. 1. Amino acid sequence of Hydrogenobacter cytochrome c-552. Symbols: →, result of Edman degradation by an automatic sequencer of the carboxymethylated apoprotein (1), of lysyl endopeptidase digests (3, 5), and of the peptide after CNBr cleavage (4) (alignment [2] is the result of manual Edman degradation of the native protein); ←, result of carboxypeptidase Y digestion.
with that of the result of carboxypeptidase Y digestion, which showed that this peptide is placed at the carboxyl terminal of cytochrome c-552.

**Isolation and sequence analysis of cyanogen bromide peptides.** Native cytochrome c-552 was cleaved with cyanogen bromide, and the products were applied to the same column for the lysyl endopeptidase digests (Fig. 4). Peptide C1 was assumed to be derived from the carboxyl-terminal portion of cytochrome c-552 because of its lack of homoserine (Table 1). The sequence analysis revealed 17 amino acid residues.

**Peptide alignment and proof of the sequence.** The sequence of the whole carboxymethylated cytochrome c-552 overlapped with peptide L1. The sequence of peptide C1 established the overlap among peptides L1 and L2. The sequence of the carboxyl-terminal portion of peptide L2 was coincident with the result of the carboxypeptidase Y digestion of carboxymethylated cytochrome c-552. Thus, we conclude that the total sequence of cytochrome c-552 from *H. thermophilus* TK-6 is constructed as shown in Fig. 1.

**DISCUSSION**

The total number of residues of cytochrome c-552 was 80, and the molecular weight, including heme, was calculated to be 7,599, which was in good agreement with the value of 7,600 estimated by the gel filtration method (10).

A new genus and species, *H. thermophilus*, was created by our previous studies (13); but the taxonomic position of this bacterium is uncertain because our previous studies showed only the uniqueness, but not the taxonomic position and phylogenetic situation, of this bacterium. *Hydrogenobacter* cytochrome c-552 belongs to the S (small) class consisting of *Pseudomonas* and *Azotobacter* cytochromes c-551 in a phylogenetic tree for the cytochrome c superfamily proposed by Dickerson (7). Cytochrome c-552 from *H. thermophilus* is most similar to *Pseudomonas* cytochrome c-557 (Fig. 5); for example, 56% of the residues are identical to those of *Pseudomonas aeruginosa*. Cytochrome c-552 differs from the cytochrome c-551 family in being two residues shorter at the amino terminus (Fig. 5). Cytochrome c-552 contains two cysteine residues (Cys-10 and Cys-13) and a histidine residue (His-14), as shown in Fig. 1, and the methionine residue (Met-59) is in a location where it is likely to be the sixth iron ligand. The facts that the methionine residue at position 59 is surrounded by a cluster of proline residues and that lysine is the carboxyl-terminal residue are also typical features of *Pseudomonas* cytochromes c-551.

Cytochrome c-552 has, however, a preponderance of basic amino acid residues (15 residues) to acidic ones (7 residues), as reflected by its basic isoelectric point of above 10 (data not shown), while cytochrome c-551 is an acidic protein (pI, 4.70 [9]). The high similarity of the amino acid sequences between cytochromes c-552 and c-551 attracts much attention because we can list many hydrogen oxidizers among *Pseudomonas* species (4), but unfortunately no cytochrome c has been sequenced from hydrogen-oxidizing pseudomonads.

Sequences of cytochromes c from some bacteria which are physiologically related to *H. thermophilus* are available in the literature. Those are cytochrome c-555 (18) from an anaerobic bacterium, *Chlorobium thiosulfotrophicum*, which has the reductive tricarboxylic acid cycle as the CO2 fixation pathway (8); cytochrome c-552 (17) from a thermophilic bacterium, *Thermus thermophilus*; cytochrome c-554 (2) from a thiosulfate-oxidizing bacterium, *Thiobacillus neapolitanus*; cytochrome c-552 (14) from an autotrophic ammonia-oxidizing bacterium, *Nitrosomonas europaea*; and cytochrome c-550 (16) from an autotrophic nitrate-oxidizing bacterium, *Nitrobacter winogradskyi*. Among those cyto-
chromes, only cytochrome c-552 of N. europaea belongs to the S class, and the amino acid sequence is similar to that of Hydrogenobacter cytchrome c-552. But cytochromes c from other bacteria do not belong to the S class, and the amino acid sequences from those bacteria are not similar to that of Hydrogenobacter cytchrome c-552, which shows that the position of Hydrogenobacter cytchrome c-552 in the phylogenetic tree is distant from these physiologically related bacteria. For further phylogenetic discussions, amino acid sequences of cytochromes c from hydrogen-oxidizing pseudomonads and thermophilic sulfur oxidizers are eagerly desired. Studies on rRNA sequences of H. thermophilus will also be helpful in determining the taxonomic position of the strain.

From the study of the tertiary structure of cytochrome c-551 from P. aeruginosa, cytchrome c-551 is proved to be composed of several alpha-helical regions (Fig. 5) (1). Most of the regions in Hydrogenobacter cytrome c-552, which correspond to the alpha-helical regions in P. aeruginosa cytrome c-551, contain clusters of alpha-helical formers as designated by Chou and Fasman (5). The respective sequences of these parts show high similarity, and the observed amino acid substitutions are those which are frequently found in homologous protein. The heme-binding regions of these cytochromes are also similar. Therefore, the amino acid sequences, as well as the tertiary structure of Hydrogenobacter cytrome c-552, are believed to be similar to those of cytrome c-551 from P. aeruginosa. Nevertheless, cytrome c-552 from H. thermophilus is extremely stable to heat compared with cytrome c-551 from P. aeruginosa (data not shown). The heat stability of Hydrogenobacter cytrome c-552 will be published elsewhere soon.

ACKNOWLEDGMENTS

We are grateful to H. Nagasawa of the Department of Agricultural Chemistry, Tokyo University, for valuable technical aid and discussions.

This research was supported in part by Grant-in-Aid for Scientific Research no. 63760085 from the Ministry of Education, Science and Culture of Japan.

LITERATURE CITED