Amino acid sequence of a ferredoxin from thermoacidophilic archaebacteria, *Thermoplasma acidophilum*

S. Wakabayashi, N. Fujimoto[†], K. Wada, H. Matsubara, L. Kerscher[+,*] and D. Oesterhelt*

Department of Biology, Faculty of Science, Osaka University, Toyonaka, Osaka 560, Japan and **Max-Planck-Institut für Biochemie, 8033 Martinsried bei München, FRG*

Received 11 August 1983

The amino acid sequence of *Thermoplasma acidophilum* was determined by various conventional techniques. It is composed of 142 amino acid residues including 9 cysteines. One of the two cysteine clusters shows exactly the same cysteine distribution as clostridial ferredoxins, but the other shows a unique distribution. Comparison of various archaebacterial ferredoxins suggests multiple origins.

Thermoplasma acidophilum *Ferredoxin* *Amino acid sequence* *Archaebacteria* *Phylogeny* *Cysteine distribution*

1. INTRODUCTION

Archaebacteria include three major groups, the methanogens, the extreme halophiles, and the thermoacidophiles [1]. Ferredoxins were isolated from organisms of each group and characterized. Ferredoxin from the methanogen, *Methanosarcina barkeri*, has a very similar primary structure to those of clostridial ferredoxins but contains only one trinuclear Fe–S-cluster [2] or two Fe–S-clusters [3]. *Halobacteria*, however, contain 2 Fe–2 S ferredoxins which show amino acid sequences homologous with those of higher plants and cyanobacteria [4,5], and function as an electron acceptor in 2-oxoacid oxidation system [6]. Ferredoxins of unique characteristics, such as M_r-values and amino acid compositions, were isolated from the thermoacidophiles [7]. These are electron acceptors of 2-oxoacid:ferredoxin oxidoreductases as well. The elucidation of the amino acid sequences of these ferredoxins will provide important information on their structure–function relationship and will have implications, moreover, for the phylogeny of archaebacteria [8].

We report here the complete amino acid sequence of a ferredoxin from the thermoacidophile, *Thermoplasma acidophilum*, and discuss the unique arrangement of cysteine residues.

2. MATERIALS AND METHODS

Thermoplasma acidophilum ferredoxin was prepared as in [7]. The ferredoxin was treated with 10% trichloroacetic acid to remove Fe and S, followed by reduction with 2-mercaptoethanol and carboxymethylation [9]. Carboxymethylated (Cm)-ferredoxin was cleaved either by trypsin or by cyanogen bromide. The resulting peptide mixture was separated by gel filtration on a Toyopearl HW-40 column (2 × 180 cm), by ion-exchange chromatography on an SP-Sephadex column (1.5 × 37 cm), by paper electrophoresis at pH 3.6 and 6.5, and by high-performance liquid chromatography (HPLC). The large peptides were further digested with staphylococcal V8 protease, chymotrypsin, and thermolysin. The amino acid compositions of Cm-ferredoxin and the purified peptides were determined with an amino acid analyzer (Irica Instruments Inc., model A-3300, Kyoto) HY

[†] Present address: Otsuka Pharmaceutical Co., Ltd., Tokushima 771-01, Japan

[+] Present address: Boehringer Mannheim GmbH, Biochemica Werk Tutzing, D-8132 Tutzing, FRG

Published by Elsevier Science Publishers B.V.
00145793/83/$3.00

after acid hydrolysis at 110°C for 24 h. The amino(N)-terminal sequences of Cm-ferredoxin and the purified peptides were determined by manual and automated solid-phase Edman degradations. The phenylthiohydantoin derivatives were identified by HPLC and TLC. The carboxyl (C)-terminal sequences were studied by carboxypeptidase digestion. The detailed procedures were essentially as in [4,10]. To confirm the Asn–Gly sequence located at residues 31 and 32, cleavage of Cm-ferredoxin with hydroxylamine was done as in [11].

3. RESULTS AND DISCUSSION

3.1. *Terminal sequences of* Thermoplasma acidophilum *ferredoxin*

The N-terminal sequence of Cm-ferredoxin was determined by manual Edman degradation as:

Val–Lys–Leu–Glu–Glu–Leu–Asp–Phe–Lys–
Pro–Lys–Pro–Ile–Asp–Glu–

Carboxidase Y released after 30 min digestion proline (91%) and threonine (96%) together with low amounts of several amino acids from Cm-ferredoxin.

3.2. *Tryptic peptides*

Trypsin produced 15 peptides which were all isolated in pure form and sequenced. The amino acid sequences of peptides, T-1, T-4, T-7, T-10, T-11, T-12, T-13, T-14 and T-15, were completely determined by manual Edman degradation (fig.1). The sequence study of peptide T-2 was not performed since it was clear from its amino acid composition that this peptide was derived from the N-terminal region of the protein. Peptide T-3 was rather long and was digested with thermolysin after the N-terminal sequence analysis. Three sub-

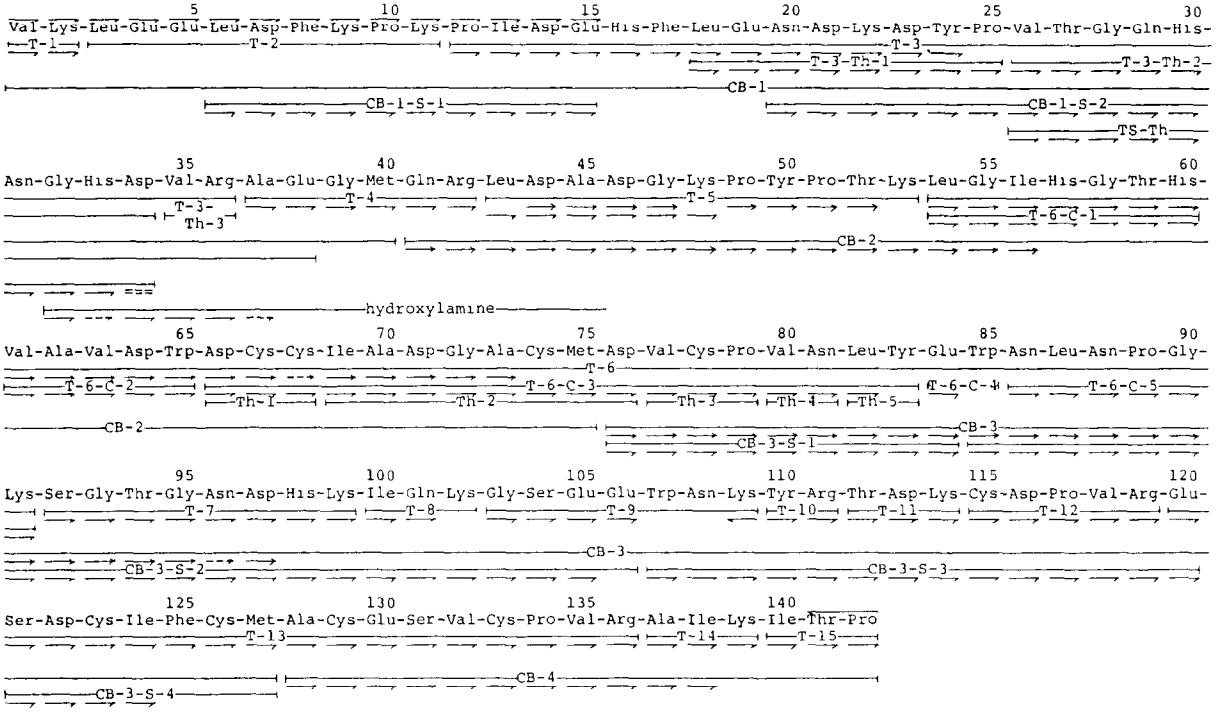

Fig.1. Summary of the sequence studies of *Thermoplasma* ferredoxin. T- and CB- refer to tryptic and cyanogen bromide peptides of the Cm-ferredoxin, respectively. Th-, C- and S- refer to the subfragments obtained by second digestion of peptides with thermolysin, chymotrypsin and staphylococcal protease, respectively; (⟶), (⟶) and (⟵) indicate a manual and automated Edman degradations and carboxypeptidase digestion, respectively. Dotted arrows indicate ambiguous identifications. The aspartic acid at residue 34 was directly identified on the analyzer after 8 steps of Edman degradation of peptide TS–Th as indicated by (= = =).

fragments were obtained in almost pure form. Sequence determination of peptide T-3–Th-2 was unsuccessful due to the His–Asn–Gly–His sequence (see below). Peptide T-5 was analyzed by manual and automated solid-phase Edman degradation except for the C-terminal lysine which followed from the specificity of trypsin. The largest peptide T-6 was first analyzed by the solid-phase method in order to determine the 20 N-terminal amino acid residues. Next, the peptide was digested with chymotrypsin. Five subfragments were purified and their amino acid sequences almost completely determined by manual Edman degradation. Peptide T-6–C-3 was further digested with thermolysin. To provide additional evidence for the distribution, unusual among ferredoxins, 5 thermolysin fragments were purified. Their amino acid compositions confirmed the proposed sequence. The amino acid sequences of peptides T-8 and T-9 were only partially determined. Carboxypeptidase B released lysine (76%) from peptide T-9 after 1.5 h digestion.

3.3. *Cyanogen bromide peptides and construction of the complete sequence*

Thermoplasma ferredoxin contains 3 methionine residues. It was treated with cyanogen bromide and the resulting 4 peptides were purified. To obtain overlaps of tryptic peptides, peptides CB-2 and CB-3 were analyzed with the solid-phase sequencer, and peptide CB-4 by manual Edman degradation. Peptides CB-1 and CB-3 were further digested with staphylococcal protease and several subfragments were isolated and analyzed.

The amino acid sequence from residue 31 to 34 could not be determined successfully by the analysis of tryptic and cyanogen bromide peptides; therefore, the peptide corresponding to residues 26–34 was obtained through mixed digestion by trypsin and staphylococcal protease followed by thermolysin treatment. Careful studies on this TS–Th-peptide revealed the sequence proposed. It was confirmed by hydroxylamine cleavage and successive manual Edman degradation of the cleaved mixture (fig.1).

From these studies, the complete amino acid sequence of *Thermoplasma* ferredoxin could be established. Summarizing, tryptic peptides T-1 to T-4 could be arranged by the N-terminal sequence analysis of the uncleaved protein and the amino acid composition of CB-1. Sequence studies of CB-2, CB-3 and their subfragments made it possible to place tryptic peptides T-4–T-13 in order. Peptide CB-4 overlapped peptides T-13–T-15.

Thermoplasma ferredoxin is composed of 142 amino acid residues with 15 Asp, 7 Asn, 6 Thr, 4 Ser, 10 Glu, 3 Gln, 10 Pro, 11 Gly, 7 Ala, 9 Cys, 10 Val, 3 Met, 7 Ile, 7 Leu, 4 Tyr, 3 Phe, 12 Lys, 6 His, 5 Arg, 3 Trp, confirming the previous analysis [7]. The exact M_r-value, excluding Fe and S, is 15963.

3.4. *Characteristic features of the primary structure of* Thermoplasma *ferredoxin*

Thermoplasma ferredoxin contains about 8 Fe, 8 S and 9 cysteine residues per molecule which might constitute two 4 Fe–4 S clusters [7]. Cysteines complexing 4 Fe–4 S 2+(2+,1+) clusters typically arrange in a sequence of the type:

–C–X–X–C–X–X–C–X–X–X–C–P– [12]

and this sequence was thought to be essential for the constitution of Fe–S chromophore. Indeed, one of the clusters (residues 123–134) of ferredoxin shows this arrangement but the other cysteine cluster present at residues 67–79 is unique. That is, the second cysteine residue of the usual arrangement is substituted by aspartic acid 71 in *Thermoplasma* ferredoxin and instead, another cysteine residue (67) directly precedes the first one. It seems likely that these successive two cysteine residues cannot contribute to the chelation of Fe in the same Fe–S center. Therefore, *Thermoplasma* ferredoxin is expected to have a unique chelate structure, which might also involve the cysteine at position 115, unique to this protein. The elucidation of the chelate structure therefore must await for the determination of the three-dimensional structure of this ferredoxin.

As far as intramolecular homology is concerned *Thermoplasma* ferredoxin shows similarities of the regions between residues 68–80 and residues 123–135 (8 identical residues out of 13), indicating that gene duplication occurred in the evolutionary past.

On comparing the primary structure of *Thermoplasma* ferredoxin to those of other bacterial ferredoxins [12] the presence of a long stretch at the N-terminal region is obvious. Halobacterial ferredoxins have also long N-terminal extensions

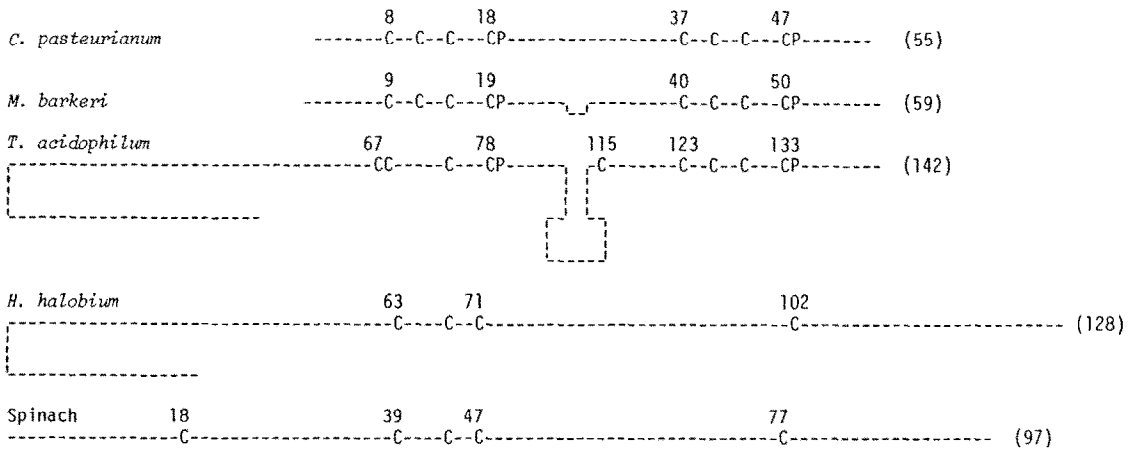

Fig.2. Comparison of the distribution of cysteine residues (C). P refers to proline residue emphasized at certain positions.

[4,5], but the amino acid sequences of the iron clusters are quite different from that of *Thermoplasma* ferredoxin (fig.2). *Thermoplasma* ferredoxin has another extra sequence between the two cysteine clusters. These long extra sequences might contribute to protect the Fe–S-cluster from extremely hazardous environments in which the bacteria live. The 3 bacterial groups above are called archaebacteria on the nucleotide sequence homology of their 16 S rRNA [1,8]. However, the amino acid sequences of their ferredoxins (fig.1) suggest multiple origins in evolution. Halobacteria have been suggested to be close to blue-green algae and green plants [4,5,12,13]. A more detailed comparison of the archaebacterial ferredoxins with each other, and with eubacterial type ferredoxins, will include the *Sulfolobus* ferredoxin, the sequence of which will be completed soon.

REFERENCES

[1] Fox, G.E., Stackebrandt, E., Hespell, R.B., Gibson, J., Maniloff, J., Dyer, T.A., Wolfe, R.S., Balch, W.E., Tanner, R.S., Magrum, L.J., Zablen, L.B., Blakemore, R., Gupta, R., Bonen, L., Lewis, B.J., Stahl, D.A., Luehrsen, K.R., Chen, K.N. and Woese, C.R. (1980) Science 209, 457–463.

[2] Hausinger, R.P., Moura, I., Moura, J.J.G., Xavier, A.V., Santos, M., LeGall, J. and Howard, J.B. (1982) J. Biol. Chem. 257, 14192–14197.

[3] Hatchikian, E.C., Bruschi, M., Forget, N. and Scandellari, M. (1982) Biochem. Biophys. Res. Commun. 109, 1316–1323.

[4] Hase, T., Wakabayashi, S., Matsubara, H., Kerscher, L., Oesterhelt, D., Rao, K.K. and Hall, D.O. (1978) J. Biochem. (Tokyo) 83, 1657–1670.

[5] Hase, T., Wakabayashi, S., Matsubara, H., Mevarech, M. and Werber, M.M. (1980) Biochim. Biophys. Acta 623, 139–145.

[6] Cammack, R., Kerscher, L. and Oesterhelt, D. (1980) FEBS Lett. 118, 271–273.

[7] Kerscher, L., Nowitzki, S. and Oesterhelt, D. (1982) Eur. J. Biochem. 128, 223–230.

[8] Woese, C.R. and Fox, G.E. (1977) Proc. Natl. Acad. Sci. USA 74, 5088–5090.

[9] Crestfield, A.M., Moore, S. and Stein, W.H. (1963) J. Biol. Chem. 238, 622–627.

[10] Wakabayashi, S., Takeda, H., Matsubara, H., Kim, C.H. and King, T.E. (1982) J. Biochem. (Tokyo) 91, 2077–2085.

[11] Steinman, H.M., Naik, V.R., Abernethy, J.L. and Hill, R.L. (1974) J. Biol. Chem. 249, 7326–7338.

[12] Matsubara, H., Hase, T., Wakabayashi, S. and Wada, K. (1980) in: The Evolution of Protein Structure and Function (Sigman, D.S. and Brazier, M.A.B. eds) pp.245–266, Academic Press, New York.

[13] Matsubara, H., Hase, T., Wakabayashi, S. and Wada, K. (1978) in: Evolution of Protein Molecules (Matsubara, H. and Yamanaka, T. eds) pp.209–220, Jap. Sci. Soc., Tokyo.