Biochem. J. (1976) **159**, 55–63
Printed in Great Britain

Amino Acid Sequence of a Four-Iron–Four-Sulphur Ferredoxin Isolated from *Bacillus stearothermophilus*

By TOSHIHARU HASE, NORIKAZU OHMIYA and HIROSHI MATSUBARA
Department of Biology, Faculty of Science, University of Osaka, Toyonaka, Osaka 560, Japan

and ROGER N. MULLINGER, K. KRISHNA RAO and DAVID O. HALL
Department of Plant Sciences, King's College, 68 Half Moon Lane, London SE24 9JF, U.K.

(*Received* 27 *April* 1976)

1. The primary structure of a 4Fe–4S ferredoxin from *Bacillus stearothermophilus* was determined and shown to consist of a single polypeptide chain of 81 amino acid residues. The molecular weight of the holoprotein is about 9120. 2. There are only four cysteine residues in the molecule; three of these are located near the *N*-terminus as a Cys-X-X-Cys-X-X-Cys segment, and the fourth cysteine residue is followed by a proline and located in the *C*-terminal half. 3. The Fe–S chromophore in *B. stearothermophilus* ferredoxin was previously well characterized and was shown to consist of a single 4Fe–4S cluster. This ferredoxin sequence establishes for the first time the relative location of the four cysteine residues necessary to bind the 4Fe–4S cluster of a 4Fe ferredoxin, and is in agreement with the criteria for the relative positions of the cysteines proposed from X-ray-crystallographic studies on an 8Fe (two 4Fe–4S clusters) ferredoxin. 4. The sequence of *B. stearothermophilus* ferredoxin is homologous in many segments to that of other bacterial ferredoxins, the degree of homology being greater towards ferredoxins from *Desulfovibrio gigas* and photosynthetic bacteria than to Clostridial ferredoxins. 5. The presence of a relatively higher number of glutamic acid and lower number of cysteine residues in the molecule may explain the greater thermal stability and oxygen-insensitivity of this ferredoxin.

Ferredoxins are proteins of low molecular weight which contain particular types of iron–sulphur clusters covalently linked to the cysteine residues of the protein chain (Hall *et al.*, 1975). The physicochemical properties of *Bacillus stearothermophilus* ferredoxin were studied in detail and the protein was shown to contain a single 4Fe–4S cluster and four cysteine residues per molecule (Mullinger *et al.*, 1975). Bacterial ferredoxins with a single 4Fe–4S cluster per molecule have also been isolated from *Desulfovibrio gigas* (LeGall & Dragoni, 1966), *Desulfovibrio desulfuricans* (Zubieta *et al.*, 1973), *Bacillus polymyxa* (Yoch, 1973), *Spirochaeta aurantia* (Johnson & Canale-Parola, 1973) and *Rhodospirillum rubrum* (Yoch *et al.*, 1975). Of these, the amino acid sequence of only *D. gigas* ferredoxin is known; it consists of 56 amino acids, including six cysteine residues (Travis *et al.*, 1971). However, the physicochemical properties of this *D. gigas* ferredoxin have not been investigated in detail, and it is not certain which four of the six cysteine residues are involved in the binding of the iron–sulphur cluster. A stable well-characterized ferredoxin such as that of *B. stearothermophilus* is an ideal choice for sequence determination as a prelude to the understanding of the structure–function relationship among this important class of proteins.

This paper describes the determination of the amino acid sequence of *B. stearothermophilus* ferredoxin, which is the first sequence of a ferredoxin with a single 4Fe–4S cluster and containing only four cysteines, and which is quite different from the sequence of *D. gigas* ferredoxin.

Experimental

Materials

B. stearothermophilus ferredoxin was prepared as described in a previous paper (Mullinger *et al.*, 1975). All reagents and enzymes used in this experiment were those listed previously (Hase *et al.*, 1976) and were purchased from Wako Pure Chemical Industries (Osaka, Japan) and Nakarai Chemical Co. (Kyoto, Japan). A proteinase isolated from *Staphylococcus aureus* was a kind gift of Dr. R. P. Ambler, Department of Molecular Biology, University of Edinburgh, Edinburgh.

Sequence analyses

The essential procedures for sequence determination were as previously described (Hase *et al.*, 1976). The amino acid composition was determined as usual after acid hydrolysis of Cm*-ferredoxin prepared by the method of Crestfield *et al.* (1963), and

* Abbreviation: Cm, *S*-carboxymethyl.

the peptides produced after enzymic hydrolysis were separated as described by Matsubara & Sasaki (1968).

A manual Edman-degradation procedure (Blombäck *et al.*, 1966) was applied to the Cm-ferredoxin (approx. 0.15 μmol) and peptides (usually approx. 0.1 μmol) to determine their *N*-terminal sequences. Phenylthiohydantoin derivatives were identified by t.l.c. on Merck silica-gel-coated glass plates in two different solvent systems (Cherbuliez *et al.*, 1964; Jeppson & Sjöquist, 1967). Occasionally a subtractive method was used for the identification of the terminal residue (Konigsberg, 1972).

The hydrazinolysis method (Tsugita *et al.*, 1960) and digestion with carboxypeptidases (Matsubara *et al.*, 1968) were used to determine the *C*-terminal residue and sequence.

The Cm-ferredoxin (approx. 2.6 μmol) was digested with either trypsin (0.4 mg) for 2 h or staphylococcal proteinase (0.25 mg) for 2 h at pH 8 and 40°C. The tryptic digest was chromatographed on a Bio-Gel P-4 column (1.6 cm × 90 cm) with 0.1 M-NH_3. Each fraction was monitored by the E_{280} and E_{220} (Koide *et al.*, 1973) as well as by the fluorescence method (Vanderkerckhove & Van Montagu, 1974). Fractions containing peptides were further purified by paper electrophoresis and chromatography (Matsubara & Sasaki, 1968). A large peptide containing methionine was treated with CNBr (Steers *et al.*, 1965) and the fragments were further separated by paper electrophoresis at pH 3.6. Further digestion was performed on a large CNBr peptide by chymotrypsin and peptides were purified on paper.

The staphylococcal-proteinase digest was purified by chromatography on a Bio-Gel P-2 column (1.6 cm × 90 cm) with 0.2 M-NH_4HCO_3, pH 9.0, and peptides were monitored as mentioned above. Further purification was carried out by paper electrophoresis.

Nomenclature

T-, S-, C- and CN- refer to tryptic, staphylococcal-proteolytic, chymotryptic and CNBr fragments derived from Cm-ferredoxin and parent peptides from Cm-ferredoxin respectively. All values for amino acid composition are expressed in mol/mol of peptide or protein.

Results and Discussion

Amino acid composition and N- and C-terminal sequences

The amino acid composition of *B. stearothermophilus* Cm-ferredoxin is shown in Table 1. The total number of residues calculated was 83, two residues more than that deduced from the sequence study. The calculations from the amino acid hydrolysis showed one glutamic acid and one threonine residue in excess.

Only four Cm-cysteine residues were present, and arginine, histidine and tryptophan were absent from this ferredoxin molecule.

Manual Edman degradation revealed the *N*-terminal sequence up to 29 residues without any ambiguity in identification as shown below:

```
1                                   10
Pro-Lys-Tyr-Thr-Ile-Val-Asp-Lys-Glu-Thr-Cys-Ile-
                                              |
                                             Cm
                                20
Ala-Cys-Gly-Ala-Cys-Gly-Ala-Ala-Ala-Pro-Asp-
    |           |
   Cm          Cm
                        29
Ile-Tyr-Asp-Tyr-Asp-Glu.
```

Treatment of the protein (0.5 mg) with carboxypeptidase A (0.01 mg) released only aspartic acid and phenylalanine, Asp (0.32) and Phe (0.25) at 20 min, and Asp (0.67) and Phe (0.60) at 3 h, suggesting the *C*-terminal sequence to be Phe-Asp. This agrees with the specificity of the enzyme, and this sequence was confirmed by hydrazinolysis experiment on a *C*-terminal peptide, as shown below.

Tryptic peptides

Three peptides, T-1 to T-3, were recovered in pure

Table 1. *Amino acid composition of B. stearothermophilus Cm-ferredoxin*

Values are expressed as molar ratio for the minimum mol.wt. 8700 and are mean values of 24 h and 72 h hydrolysates. Values for threonine and serine are calculated after extrapolation to zero hydrolysis time.

Amino acid	Content (mol/mol of protein)	Content Nearest integer
Lysine	3.45	4
Cm-cysteine	4.27	4
Aspartic acid	16.8	17
Threonine	4.53	5
Serine	1.39	1
Glutamic acid	7.88	8
Proline	6.23	6
Glycine	6.29	6
Alanine	8.12	8
Valine	5.02	5
Methionine	2.00	2
Isoleucine	7.85	8
Leucine	2.18	2
Tyrosine	4.05	4
Phenylalanine	3.08	3
Total residues		83

Table 2. *Amino acid compositions of tryptic peptides of Cm-ferredoxin*

Values are the means for single 24h and 72h hydrolysates; values for threonine and serine are calculated after extrapolation to zero time. Numbers in parentheses are integral values; the value for valine is corrected for incomplete hydrolysis. Abbreviations: P, Pauli-positive (Ames & Mitchell, 1952); PIN, paper chromatography in pyridine/3-methyl-1-butanol/0.1M-NH_3 (6:3:5, by vol.); PE, paper electrophoresis at pH3.6 (Matsubara & Sasaki, 1968); *R*, relative mobility of peptide with respect to leucine; *m*, mobility of the peptide, in cm, from the origin towards anode (+) or cathode (−) after 1h of electrophoresis at 43V/cm.

		Content (mol/mol of peptide)		
Amino acid	Peptide ...	T-1	T-2	T-3
Lysine		1.03 (1)	1.05 (1)	0.86 (1)
Cm-cysteine			4.00 (4)	
Aspartic acid		1.51 (1)	11.34 (11)	5.00 (5)
Threonine		0.93 (1)	2.67 (3)	
Serine			0.83 (1)	
Glutamic acid		0.29	6.32 (6)	1.14 (1)
Proline			3.36 (3)	2.01 (2)
Glycine		0.24	4.67 (5)	1.01 (1)
Alanine		0.30	6.63 (6)	1.04 (1)
Valine		0.82 (1)	2.62 (3)	0.99 (1)
Methionine			1.72 (2)	
Isoleucine		0.89 (1)	6.60 (7)	
Leucine			2.06 (2)	
Tyrosine		0.80 (1)	2.88 (3)	
Phenylalanine			0.99 (1)	1.95 (2)
Total residues		6	59	14
Yield (%)		30	54	21
Colour reaction		P	P	
Purification			PIN (*R* 0.21)	PIN (*R* 0.53) PE (*m* −6)

form. The *N*-terminal peptide, Pro-Lys, was accidentally lost, and the Lys-Phe bond at the *C*-terminal region was not digested. The compositions of these peptides are in Table 2. Fig. 1 should be consulted for the following experimental details.

Peptide T-1. Edman degradation (six steps) determined the sequence.

Peptide T-2. A mixture of carboxypeptidases A and B (0.01mg each) released lysine and isoleucine successively, suggesting the *C*-terminal sequence to be -Ile-Lys. CNBr treatment produced three peptides, T-2-CN-1, T-2-CN-2 and T-2-CN-3. Edman degradation (20 steps) of peptide T-2-CN-1, was successfully carried out. The *C*-terminal residue was homoserine, which was released by carboxypeptidase A. Edman degradations of peptides T-2-CN-2 and T-2-CN-3 (ten and nine steps respectively) were also performed. The fact that methionine was at the *N*-terminus of peptide T-2-CN-2 suggested the presence of a Met-Met bond. Carboxypeptidase B digestion of peptide T-2-CN-3, followed by treatment with carboxypeptidase A, determined the *C*-terminal sequence as Ile-Lys. Chymotryptic digestion of peptide T-2-CN-1 produced two more peptides, T-2-CN-1-C-1 and -2. Edman degradations of peptide T-2-CN-1-C-1 and peptide T-2-CN-1-C-2 (16 and 7 steps respectively) confirmed and extended the partial sequence of peptide T-2-CN-1. Carboxypeptidase A released tyrosine, followed by isoleucine, from peptide T-2-CN-1-C-1, and tyrosine was followed by alanine and isoleucine at a low rate, from peptide T-2-CN-1-C-2.

Peptide T-3. Edman degradation (11 steps) and carboxypeptidase A digestion determined the sequence of the *C*-terminal peptide.

Staphylococcal proteinase peptides

Five peptides were obtained after chromatography on Bio-Gel P-2 and paper electrophoresis (Table 3).

Peptide S-1. Edman degradation (four steps) was carried out. Glutamic acid was detected by hydrazinolysis as the *C*-terminus.

Peptide S-2. Only two steps of Edman degradation were carried out. Glutamic acid was again detected by hydrazinolysis.

Peptide S-3. Edman degradation (19 steps) was performed. This peptide was further digested by the same enzyme, staphylococcal proteinase. Three main peptides were obtained: S-3-1, S-3-2 and S-3-3. Edman degradation (15 steps) of peptide S-3-1 completed the sequence. After the 15th step, free glutamic

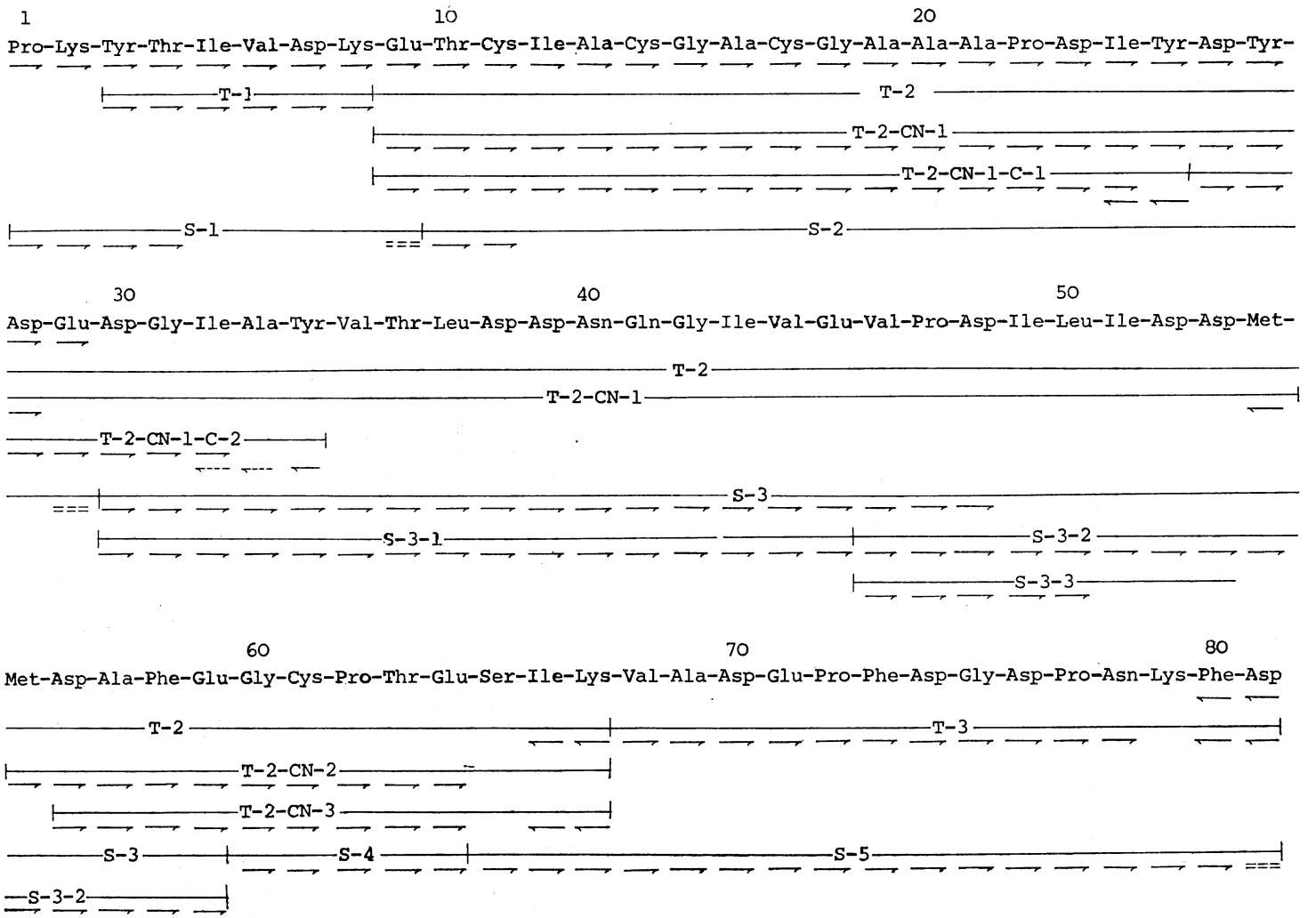

Fig. 1. *Summary of the sequence study of B. stearothermophilus ferredoxin*

→ and ← just below the sequence represent Edman degradation and carboxypeptidase digestion respectively. === represents hydrazinolysis and ⇠ uncertain identification by carboxypeptidase. T, S and CN represent peptides formed after treatment with trypsin, staphylococcal proteinase and CNBr respectively.

Table 3. *Amino acid compositions of staphylococcal proteolytic peptides of Cm-ferredoxin*

Conditions of hydrolysis and meaning of abbreviations are as in Table 2.

Amino acid	Peptide ...	Content (mol/mol of peptide)				
		S-1	S-2	S-3	S-4	S-5
Lysine		2.00 (2)				1.90 (2)
Cm-cysteine			3.18 (3)		1.00 (1)	
Aspartic acid		1.01 (1)	3.10 (3)	7.94 (8)		4.79 (5)
Threonine		0.87 (1)	0.88 (1)	0.82 (1)	0.99 (1)	
Serine						0.90 (1)
Glutamic acid		1.09 (1)	1.11 (1)	3.11 (3)	1.06 (1)	1.13 (1)
Proline		0.94 (1)	1.00 (1)	1.34 (1)	1.11 (1)	2.06 (2)
Glycine			2.02 (2)	1.84 (2)	0.84 (1)	1.02 (1)
Alanine			4.96 (5)	2.31 (2)		1.03 (1)
Valine		0.89 (1)		2.31 (3)		1.04 (1)
Methionine				2.18 (2)		
Isoleucine		0.87 (1)	1.90 (2)	3.64 (4)	0.10	1.02 (1)
Leucine				1.81 (2)		
Tyrosine		0.99 (1)	1.94 (2)	0.91 (1)		
Phenylalanine				1.17 (1)		1.96 (2)
Total residues		9	20	30	5	17
Yield (%)		52	23	37	44	48
Colour reaction		P	P	P		
Purification		PE (m −10)	PE (m +2)	PE (m 0)	PE (m −0.5)	PE (m −5.3)

acid was detected on the amino acid analyser. Edman degradation (13 steps) of peptide S-3-2 was carried out. The last step gave free glutamic acid. Edman degradation (five steps) of peptide S-3-3 showed that this peptide originated from the *N*-terminal region of peptide S-3-2.

Peptide S-4. Edman degradation (four steps) and the identification of free glutamic acid after completion of the fourth step completed the sequence.

Peptide S-5. Edman degradation (16 steps) and hydrazinolysis, which gave aspartic acid, completed the sequence of this peptide.

Complete amino acid sequence

Edman degradation on the original Cm-ferredoxin placed peptide T-1 before T-2. Peptide S-1 confirmed the *N*-terminal sequence and also the positions of peptide T-1 and the *N*-terminal peptide which was lost. Peptide T-2-CN-1-C-2 connected peptide S-3 and a part of peptide T-2-CN-1, extending the sequence up to residue 48. Peptide S-3-2 extended the sequence of peptide S-3 up to its *C*-terminal end, i.e. residue 59. Peptides T-2-CN-2 and T-2-CN-3 overlapped with the *C*-terminal peptide, S-5. Together with other supplement peptides the complete amino acid sequence was established as shown in Fig. 1. The total number of residues was 81, giving a mol.wt. of about 8770 for the apoprotein, 9120 for the holoprotein. This value is in fair agreement with the mol.wt. of 8500 deduced from gel-electrophoretic analysis (Mullinger *et al.*, 1975).

The staphylococcal proteinase was very useful in this study and in other sequence determinations (R. F. Ambler, personal communication), and as suggested by the original authors (Houmard & Drapean, 1972; Ryden *et al.*, 1974) the enzyme cleaved mainly the peptide bonds at the *C*-terminal side of glutamic acid (9–10, 29–30, 45–46, 59–60 and 64–65). Only one bond at the *C*-terminal side of aspartic acid (53–54) was weakly cleaved, as shown in Fig. 1.

Molecular and structural properties and comparison with other 4Fe-ferredoxins

The complete amino acid sequence of *B. stearothermophilus* ferredoxin contained 81 amino acid residues. In molecular size and in the number of cysteine residues per molecule, this ferredoxin resembles ferredoxins I and II of *B. polymyxa* (Yoch, 1973). In contrast, the four-iron ferredoxins of *S. aurantia*, *D. gigas* and *D. desulfuricans* (see Zubieta *et al.*, 1973) are smaller in size (mol.wt. 6000) and contain six cysteine residues per molecule. The four-iron ferredoxin from *R. rubrum* is a comparatively larger molecule of mol.wt. 14500 (Yoch *et al.*, 1975), and the cysteine content is uncertain. Thus ferredoxins containing single four-iron clusters, with varying molecular sizes and amino acid compositions, occur in many diverse species of bacteria.

Position of the cysteine residues. The most valuable information derived from the present sequence study is the relationship between the location of the cysteine residues in the protein and the formation of the

```
      1
(a)  Pro-Lys-Tyr-Thr-Ile-Val-Asp-Lys-Glu-Thr-Cys-Ile-Ala-Cys-Gly-Ala-Cys-Gly-Ala-Ala-Ala-Pro- -
(b)  Pro-Ile- - Gln-Val- - - - Asp-Asn-Cys-Met-Ala-Cys-Gln-Ala-Cys-Ile-Asn-Glu-Cys-Pro-Val-
(c)  Ala-Leu-Tyr- - Ile- - - Thr-Glu-Glu-Cys-Thr-Tyr-Cys-Gly-Ala-Cys-Glu-Pro-Glu-Cys-Pro-Val-
(d)  Ala-His- - Arg-Ile- - - Thr-Glu-Glu-Cys-Thr-Tyr-Cys-Ala-Ala-Cys-Glu-Pro-Glu-Cys-Pro-Val-
(e)  Ala-Leu-Met- - Ile- - - Thr-Asp-Gln-Cys-Ile-Asn-Cys-Asn-Val-Cys-Gln-Pro-Glu-Cys-Pro-Asn-
(f)  Ala- - Tyr-Lys-Ile- - - Ala-Asp-Ser-Cys-Val-Ser-Cys-Gly-Ala-Cys-Ala-Ser-Glu-Cys-Pro-Val-
(g)  Ala-His-Ile- - Ile- - - Thr-Asp-Glu-Cys-Ile-Ser-Cys-Gly-Ala-Cys-Ala-Ala-Glu-Cys-Pro-Val-
(h)  Met-His-Val- - Ile- - - Ser-Asp-Glu-Cys-Val-Lys-Cys-Gly-Ala-Cys-Ala-Ser-Thr-Cys-Pro-Thr-

      23
(a)  Asp-Ile-Tyr- - Asp-Tyr-Asp-Glu-Asp-Gly-Ile-Ala-Tyr-Val-Thr-Leu-Asp-Asp-Asn-Gln-Gly-Ile-Val-
(b)  Asp-Val-Phe- - Gln-Met-Asp-Glu-Gln-Gly-Asp-Lys-Ala-Val-Asn-Ile- - - - - - - - -
(c)  Thr-Ala-Ile-Ser-Ala-Gly-Asp-Asp-Ile-Tyr-Val-Ile- - - - Asp-Ala-Asn-Thr-Cys-Asn- -
(d)  Asn-Ala-Ile-Ser-Ala-Gly-Asp-Glu-Ile-Tyr-Ile-Val- - - - Asp-Glu-Ser-Val-Cys-Thr- -
(e)  Gly-Ala-Ile-Ser-Gln-Gly-Asp-Glu-Thr-Tyr-Val-Ile- - - - Glu-Pro-Ser-Leu-Cys-Thr- -
(f)  Asn-Ala-Ile-Ser-Gln-Gly-Asp-Ser-Ile-Phe-Va- Ile- - - - Asp-Ala-Asp-Thr-Cys-Ile- -
(g)  Glu-Ala-Ile-His-Glu-Gly-Thr-Gly-Lys-Tyr-Gln-Val- - - - Asp-Ala-Asp-Thr-Cys-Ile- -
(h)  Gly-Ala-Ile-Glu-Glu-Gly-Glu-Thr-Lys-Tyr-Val-Val- - - - - Thr-Asp-Ser-Cys-Ile- -

      45
(a)  Glu-Val-Pro-Asp- - Ile-Leu-Ile-Asp-Asp-Met-Met-Asp-Ala-Phe-Glu-Gly-Cys-Pro-Thr-Glu-Ser-Ile-
(b)  - - Pro-Asn-Ser-Asn-Leu-Asp-Asp-Gln-Cys-Val-Glu-Ala-Ile-Gln-Ser-Cys-Pro-Ala- - Ala-Ile-
(c)  Glu-Cys- - - Ala-Gly-Leu-Asp-Glu-Gln- - - Ala-Cys-Val-Ala-Val-Cys-Pro-Ala-Glu-Cys-Ile-
(d)  Asp-Cys- - - Glu-Gly-Tyr-Tyr-Asp-Glu-Pro- - Ala-Cys-Val-Ala-Val-Cys-Pro-Val-Asp-Cys-Ile-
(e)  Glu-Cys- - - Val-Gly-His-Tyr-Glu-Thr-Ser- - Gln-Cys-Val-Glu-Val-Cys-Pro-Val-Asp-Cys-Ile-
(f)  Asp-Cys- - - - - - - - - - Gly-Asn-Cys-Ala-Asn-Val-Cys-Pro-Val-Gly-Ala-Pro-
(g)  Asp-Cys- - - - - - - - - - Gly-Ala-Cys-Gln-Ala-Val-Cys-Pro-Thr-Gly-Ala-Val-
(h)  Asp-Cys- - - - - - - - - - Gly-Ala-Cys-Glu-Ala-Val-Cys-Pro-Thr-Gly-Ala-Ile-

      67                                                    81
(a)  Lys-Val-Ala- - - Asp-Glu-Pro-Phe-Asp-Gly-Asp-Pro-Asn-Lys-Phe-Asp
(b)  Arg-Ser
(c)  Val-Gln-Gly
(d)  Ile-Lys-Val
(e)  Lys-Asp-Pro-Ser-His-Glu-Glu-Thr-Glu-Asp-Glu-Leu-Arg-Ala-Lys-Tyr-Glu-Arg-Ile-Thr-Gly-Glu-Gly
(f)  Val-Gln-Glu
(g)  Lys-Ala-Glu
(h)  Ser-Ala-Glu
```

Fig. 2. *Comparison of bacterial ferredoxin sequences*

(*a*) *B. stearothermophilus* 4Fe–4S, facultative thermophile (present study); (*b*) *D. gigas* 4Fe–4S, sulphate reducer (Travis *et al.*, 1971); (*c*) *Chlorobium limicola* I, 8Fe–8S, green photosynthetic (Tanaka *et al.*, 1974); (*d*) *Chlorobium limicola* II (Tanaka *et al.*, 1975); (*e*) *Chromatium vinosum* 8Fe–8S, purple photosynthetic (Matsubara *et al.*, 1970), modified as suggested by Tanaka *et al.* (1974); (*f*) *C. pasteurianum* 8Fe–8S, obligate anaerobe, mesophile (Yasunobu & Tanaka, 1973); (*g*) *C. tartarivorum* 8Fe–8S, obligate anaerobe, thermophile (Yasunobu & Tanaka, 1973); (*h*) *Peptostreptococcus elsdenii*, obligate anaerobe, mesophile (Yasunobu & Tanaka, 1973).

ferredoxin type of 4Fe–4S chromophore. X-ray-crystallographic studies on the eight-iron ferredoxin from *Peptococcus aerogenes* have shown that the iron–sulphur chromophore in this molecule exists as two almost identical cubical clusters, each cluster containing four iron atoms linked to four cysteine sulphur atoms of the protein chain and four atoms of inorganic sulphur (Adman *et al.*, 1973). The four iron atoms of one of the clusters are linked to cysteines-8, -11, -14 and -45, and the four iron atoms of the other cluster are linked to cysteines-35, -38, -41 and -18 of the ferredoxin molecule. Thus it has been suggested that the formation of a bacterial-ferredoxin-type 4Fe–4S cluster requires the presence of three cysteine residues as a -Cys-X-X-Cys-X-X-Cys segment and a fourth isolated -Cys- distal to this segment in the molecule of the apoprotein (Wada *et al.*, 1974). This observation is unambiguously proved by the present study, which shows that the only four cysteine residues of *B. stearothermophilus* ferredoxin are located at positions 11, 14, 17 and 61 of the protein chain. Further, the isolated cysteine residue is followed by a proline, a sequence found in other bacterial ferredoxins also. We may therefore conclude that the requirement for formation of a 4Fe–4S cluster as found in ferredoxins is the presence of a -Cys-X-X-Cys-X-X-Cys- segment and a -Cys-Pro-sequence at some distance from the segment in the molecule of the apoprotein. From a calculation of the conformational energies associated with a 4Fe–4S centre and the position of the cysteines in a peptide chain, Van Rooten & De Coen (1974) have deduced that the first three cysteine residues can be close together, but the fourth cysteine needs to be at a distance for the formation of a stable chromophore.

Comparison of the sequence with sequences of other ferredoxins. Fig. 2 compares the sequences of various bacterial ferredoxins containing one or two 4Fe–4S centres. The sequence of *D. gigas* ferredoxin is aligned in such a way that the -Cys-Pro- segment in the *C*-terminal half of this molecule matches with the -Cys-Pro- segment in the other ferredoxin sequences. From the present study it seems more likely that the isolated cysteine residue at position 50 of *D. gigas* ferredoxin is the fourth cysteine that is involved in the iron–sulphur cluster formation than the cysteine residue at position 43, as was suggested by Yasunobu & Tanaka (1973). The published sequence of *Chromatium* ferredoxin (Matsubara *et al.*, 1970), is tentatively modified to incorporate the suggestions of Tanaka *et al.* (1974). Although it is not highly meaningful to compare protein sequences after placing a number of gaps, the homologous nature of the ferredoxins is quite obvious, especially if one compares the first 30 residues in all the ferredoxins.

B. stearothermophilus ferredoxin seems more closely related to the four-iron ferredoxin from *D. gigas* and to the eight-iron ferredoxins of photosynthetic bacteria than to the eight-iron ferredoxins of the anaerobic fermenters, such as the clostridia. A matrix indicating the similarities among various bacterial ferredoxin sequences is presented in Table 4. Clostridial-type ferredoxins are quite similar to each other and show a fair degree of similarity to photosynthetic bacterial ferredoxins, but they are less related to *D. gigas* and *B. stearothermophilus* ferredoxins. It is possible that facultative aerobes, such as *B. stearothermophilus*, and sulphate reducers, such as *D. gigas*, evolved, with respect to the photosynthetic bacteria, at a very early stage and in a separate line, from the clostridial type of anaerobic fermenters.

Table 4. *Matrix for the comparison among bacterial ferredoxins*

(*a*) to (*h*) correspond to ferredoxins listed in Fig. 2. The values are expressed as percentage similarity between two ferredoxins, and was calculated by dividing the number of the same amino acid residues, compared between two ferredoxins as aligned in Fig. 2, by the total number of residues compared.

	(A)	(B)	(C)	(D)	(E)	(F)	(G)	(H)
(*a*)	100	22	21	18	18	16	18	14
(*b*)		100	23	21	16	20	17	16
(*c*)			100	61	38	51	46	38
(*d*)				100	39	42	44	38
(*e*)					100	29	26	24
(*f*)						100	60	53
(*g*)							100	70
(*h*)								100

Primary structure and stability. The ferredoxin used in this study was isolated from *B. stearothermophilus* cells grown aerobically at 60°C. This ferredoxin has been found to be more resistant to heat (65°C) denaturation (under aerobic conditions) than the eight-iron ferredoxins from *Clostridium pasteurianum* and the two-iron ferredoxin from *Spirulina maxima*, which is the most stable plant-type ferredoxin analysed so far (Cammack *et al.*, 1976). Also, *B. stearothermophilus* ferredoxin is remarkably stable to cycles of reduction by sodium dithionite and reoxidation by oxygen, a process which normally leads to decomposition of many other ferredoxins. Hence the iron–sulphur chromophore should be in a very stable environment within the tertiary structure of this ferredoxin.

Devanathan *et al.* (1969) compared the relative stabilities of ferredoxins from the mesophiles *C. pasteurianum* and *Clostridium acidi urici* and from the thermophiles *Clostridium tartarivorum* and *Clostridium thermosaccharolyticum*; they found that the thermophilic ferredoxins are much more resistant to heat denaturation and less sensitive to oxygen than are the mesophilic ferredoxins, *C. thermosaccharolyticum* ferredoxin being the most stable and *C.*

pasteurianum ferredoxin the least stable. Although the three-dimensional structure of none of these ferredoxins has been determined, because of the similarities in properties and sequences of these ferredoxins to that of *P. aerogenes* ferredoxin (two 4Fe–4S clusters) it may be assumed that the crystal structures of all the clostridial eight-iron ferredoxins are similar to the known structure of *P. aerogenes* ferredoxin.

What differences in the primary structures can account for the differences in heat-stability and oxygen-sensitivity among these ferredoxins? Tanaka *et al.* (1973) have pointed out that the glutamine or glutamic acid residues in positions 31 and 44, found only in the sequences of thermophilic clostridial ferredoxins, can be hydrogen-bonded to other amino acid side chains and this possibly can account for the increased heat-stability of these ferredoxins. Perutz & Raidt (1975) suggest that the greater heat-stability of thermophilic ferredoxins arises mainly from external salt bridges linking residues near the *N*-terminus to others near the *C*-terminus. A closer look at the primary structures of these clostridial ferredoxins (Yasunobu & Tanaka, 1973) reveals that there is a parallel between the heat-stability of the ferredoxins and the number of glutamic acid residues in the interior of the molecule. If we ignore the glutamic acid at the *C*-terminus (which in the tertiary structure is outside the chromophore and may not be involved in the formation of salt bridges or hydrogen bonds) of some of these ferredoxins, then the number of glutamic acid residues in the interior of the protein molecule increases in the order *C. pasteurianum*, 1; *C. acidi urici*, 2; *C. tartarivorum*, 4; *C. thermosaccharolyticum*, 6. We cannot offer a theoretical explanation for this correspondence between the increased heat-stability of the proteins and the presence of additional numbers of glutamic acid residues, except that glutamic acid residues are the best helix makers in proteins (Robson & Pain, 1971). Since we now know the sequence of the *B. stearothermophilus* 4Fe ferredoxin, we can speculate that the six glutamic acid residues in the interior of this ferredoxin may contribute to the increased heat-stability of this protein.

Another factor which may contribute to the greater thermal stability of *B. stearothermophilus* ferredoxin is the lower number of cysteine residues in this molecule when compared with the number of cysteines in other four-iron and eight-iron ferredoxins whose sequences are known. When discussing the higher stability of alcohol dehydrogenase from *B. stearothermophilus* compared with horse liver alcohol dehydrogenase, Bridgen *et al.* (1973) have pointed out that several enzymes from thermophilic sources contain fewer thiol groups than their mesophilic counterparts. The enzyme glyceraldehyde 3-phosphate dehydrogenase prepared from *Thermus aquaticus* (grown at 70°C) is stable at temperatures up to 98°C and in denaturing agents such as 8M-urea (Hocking & Harris, 1976). The amino acid sequence of this enzyme is homologous to a great extent with that of the glyceraldehyde phosphate dehydrogenases of mesophiles, except that the *T. aquaticus* enzyme contains only one thiol group per subunit (Hocking & Harris, 1976). whereas the corresponding enzyme from mesophiles contain two or more thiol groups. Only when we know the amino acid sequences of ferredoxins from a few more thermophiles and mesophiles could we ascertain whether the higher stability of the *B. stearothermophilus* ferredoxin is due to its lower cysteine content.

This work was supported by grants from the Japanese Ministry of Education and the U.K. Science Research Council. We thank Dr. R. Cammack for valuable discussions.

References

Adman, E. T., Sieker, L. C. & Jensen, L. H. (1973) *J. Biol. Chem.* **248**, 3987–3996

Ames, B. N. & Mitchell, H. K. (1952) *J. Am. Chem. Soc.* **74**, 252–253

Blombäck, B., Blombäck, M., Edman, P. & Hessel, B. (1966) *Biochim. Biophys. Acta* **115**, 371–396

Bridgen, J., Kolb, E. & Harris, J. I. (1973) *FEBS Lett.* **33**, 1–3

Cammack, R., Hall, D. O., Mullinger, R. N. & Rao, K. K. (1976) *Experientia Suppl.* **26**, 307–314

Cherbuliez, E., Baehler, Br. & Rabinowitz, J. (1964) *Helv. Chim. Acta* **47**, 1350–1353

Crestfield, A. M., Moore, S. & Stein, W. H. (1963) *J. Biol. Chem.* **238**, 622–627

Devanathan, T., Akagi, J. M., Hersch, R. T. & Himes, R. H. (1969) *J. Biol. Chem.* **244**, 2846–2853

Hall, D. O., Cammack, R. & Rao, K. K. (1975) *Sci. Prog. (Oxford)* **62**, 285–317

Hase, T., Wada, K. & Matsubara, H. (1976) *J. Biochem. (Tokyo)* **69**, 329–343

Hocking, J. D. & Harris, J. I. (1976) *Experientia Suppl.* **26**, 121–133

Houmard, J. & Drapean, G. R. (1972) *Proc. Natl. Acad. Sci. U.S.A.* **69**, 3506–3509

Jeppson, J.-O. & Sjöquist, J. (1967) *Anal. Biochem.* **18**, 264–269

Johnson, P. W. & Canale-Parola, E. (1973) *Arch. Mikrobiol.* **89**, 341–353

Koide, T., Tsunasawa, S. & Ikenaka, T. (1973) *Eur. J. Biochem.* **32**, 408–416

Konigsberg, W. (1972) *Methods Enzymol.* **25**, 326–332

LeGall, J. & Dragoni, N. (1966) *Biochem. Biophys. Res. Commun.* **23**, 145–149

Matsubara, H. & Sasaki, R. M. (1968) *J. Biol. Chem.* **243**, 1732–1757

Matsubara, H., Sasaki, R. M. & Chain, R. K. (1968) *J. Biol. Chem.* **243**, 1725–1731

Matsubara, H., Sasaki, R. M., Tsuchiya, D. K. & Evans, M. C. W. (1970) *J. Biol. Chem.* **245**, 2121–2131
Mullinger, R. N., Cammack, R., Rao, K. K., Hall, D. O., Dickson, D. P. E., Johnson, C. E., Rush, J. D. & Simopoulos, A. (1975) *Biochem. J.* **151**, 75–83
Perutz, M. F. & Raidt, R. H. (1975) *Nature (London)* **255**, 256–259
Robson, E. & Pain, R. H. (1971) *J. Mol. Biol.* **58**, 237–259
Ryden, A.-C., Ryden, L. & Philipson, L. (1974) *Eur. J. Biochem.* **44**, 105–114
Steers, E., Jr., Craven, G. R., Anfinsen, C. B. & Bethune, J. L. (1965) *J. Biol. Chem.* **240**, 2478–2484
Tanaka, M., Haniu, M., Yasunobu, K. T., Himes, R. H. & Akagi, J. M. (1973) *J. Biol. Chem.* **248**, 5215–5217
Tanaka, M., Haniu, M., Yasunobu, K. T., Evans, M. C. W. & Rao, K. K. (1974) *Biochemistry* **13**, 2953–2959
Tanaka, M., Haniu, M., Yasunobu, K. T., Evans, M. C. W. & Rao, K. K. (1975) *Biochemistry* **14**, 1938–1943
Travis, J., Newman, D. J., Le Gall, J. & Peck, H. D. (1971) *Biochem. Biophys. Res. Commun.* **45**, 452–458
Tsugita, A., Gish, D. T., Young, J., Fraenkel-Conrat, H., Knight, C. A. & Stanley, W. M. (1960) *Proc. Natl. Acad. Sci. U.S.A.* **46**, 1463–1469
Vanderkerckhove, J. & Van Montagu, M. (1974) *Eur. J. Biochem.* **44**, 279–288
Van Rooten, M. & De Coen, J. L. (1974) *Arch. Int. Physiol. Biochim.* **82**, 793
Wada, K., Kagamiyama, H., Shin, M. & Matsubara, H. (1974) *J. Biochem. (Tokyo)* **76**, 1217–1225
Yasunobu, K. T. & Tanaka, M. (1973) in *Iron–Sulfur Proteins* (W. Lovenberg, ed.), vol. 2, pp. 27–130, Academic Press, New York and London
Yoch, D. C. (1973) *Arch. Biochem. Biophys.* **158**, 633–640
Yoch, D. C., Arnon, D. I. & Sweeny, W. V. C. (1975) *J. Biol. Chem.* **250**, 8330–8336
Zubieta, J. A., Mason, R. & Postgate, J. R. (1973) *Biochem. J.* **133**, 851–854