Phytochemistry, 1978, Vol. 17, pp. 1863–1867. © Pergamon Press Ltd. Printed in England. 0031-9442/78/1101-1863 $02.00/0

THE AMINO ACID SEQUENCE OF FERREDOXIN FROM THE ALGA *MASTIGOCLADUS LAMINOSUS*

TOSHIHARU HASE†, SADAO WAKABAYASHI†, HIROSHI MATSUBARA†, K. KRISHNA RAO*‡ DAVID O. HALL‡
HERBERT WIDMER§, JURG GYSI§ and HERBERT ZUBER§

†Department of Biology, Faculty of Science, Osaka University, Osaka 560, Japan; ‡Kings College, 68 Half Moon Lane, London SE 24, U.K.; §Institut für molekular biologie und Biophysik, ETH-Hönggerberg, CH-8049 Zurich, Switzerland

(*Revised received* 22 *May* 1978)

Key Word Index—*Mastigocladus laminosus*; Cyanophyceae; blue-green alga; ferredoxin; thermophilic proteins; evolution.

Abstract—Ferredoxin was purified from the thermophilic blue-green alga, *Mastigocladus laminosus*. The physicochemical properties of this ferredoxin are similar to those of other [2Fe–2S] plant ferredoxins except for its unusual thermal stability. The primary structure of the protein was determined and consists of 98 amino acid residues, 5 of which are cysteines. The positions of 4 cysteines which bind the iron atoms of the active centre are identical to those in other ferredoxins. The primary structure of the ferredoxin does not reveal any special features to account for its high thermal stability.

INTRODUCTION

Ferredoxins are a class of iron–sulphur proteins which participate in many diverse electron transfer reactions in bacteria, algae and higher plants [1, 2]. In the past few years the primary structures of the [2Fe–2S] ferredoxins from one green alga and some blue-green algae and higher plants have been determined and the sequences suggest that these ferredoxins have evolved from a common ancestor [3]. *Mastigocladus laminosus* is a nitrogen-fixing, thermophilic, blue-green alga which is one of the major constituents of the algal flora of hot springs [4]. The primary structure of ferredoxin (and other proteins) from this organism is of interest since it could give us some valuable insight into the evolutionary and also structure–function relationships between thermophilic and mesophilic organisms. With this objective in mind we have purified the ferredoxin from *M. laminosus* and determined its amino acid sequence.

RESULTS AND DISCUSSION

From 1 kg of wet *M. laminosus* cells, on average 10 mg of pure ferredoxin was obtained. The optical and electron paramagnetic spectra of ferredoxin were similar to those of other plant ferredoxins [5]; the A_{420} nm/A_{280} nm ratio was 0.53. The redox potential ($E'0$) of the ferredoxin at pH 8 was −325 mV [6] which is one of the highest potentials determined for a plant type ferredoxin —only *Spirulina maxima* ferredoxin II has a more positive potential (−310 mV) [6].

The amino acid composition of carboxymethyl(Cm)-ferredoxin is given in Table 1. Amino (N)-terminal sequence of Cm-ferredoxin was determined by manual Edman degradations up to 21 residues as follows, Ala-Thr-Tyr-Lys-Val-Thr-Leu-Ile-Asn-Glu-Ala-Glu-Gly-Leu-Asn-Lys-Thr-Ile-Glu-Val-Pro. Carboxypetidase A released leucine (0.97) and tyrosine (0.99) from Cm-ferredoxin for 30 min and no other residue was released even after digestion for 24 hr. The order of leucine and tyrosine could not be decided but as shown later the C-terminal sequence was found to be Leu-Tyr.

Table 1. Amino acid composition of Cm-ferredoxin of *Mastigocladus laminosus*

Amino acid	From acid hydrolysis*	From the sequence
Lys	3.87 (4)	4
His	0.96 (1)	1
Arg	1.00 (1)	1
Cm-Cys	5.18 (5)	5
Asp	12.30 (12)	12†
Thr	7.90 (8)	8
Ser	6.18 (6)	6
Glu	14.10 (14)	14‡
Pro	2.92 (3)	3
Gly	6.33 (6)	6
Ala	9.76 (10)	10
Val	5.70 (6)	6
Ile	7.40 (7)	7
Leu	7.56 (8)	8
Tyr	6.34 (6)	6
Phe	1.10 (1)	1
Total residues	98	98

*Results from 6 N HCl hydrolysis for 24 and 72 hr. The values of threonine, serine and tyrosine were obtained by extrapolating to zero time and those of valine and isoleucine were of 72 hr hydrolysis.

†Sum of 10 aspartic acid and 2 asparagine.

‡Sum of 10 glutamic acid and 4 glutamine.

*To whom correspondence should be addressed.

Table 2. Amino acid compositions of tryptic and staphylococcal protease peptides of Cm-ferredoxin

	T-1	T-2	T-3	T-4	T-5	T-6	T-7	T-8	T-9	S-1
Lys	0.96(1)	0.96(1)		0.83(1)			1.01(1)	1.00(1)		0.86(1)
His							0.98(1)	0.94(1)		
Arg			0.96(1)							0.98(1)
Cm-cys			0.74(1)	2.30(2)		1.88(2)				2.93(3)
Asp		1.03(1)	4.19(4)		4.94(5)	1.30(1)				5.93(6)
Thr	0.76(1)		1.02(1)	1.01(1)	1.17(1)	1.89(2)	1.07(1)	0.92(1)		2.13(2)
Ser			0.84(1)	0.97(1)	2.42(3)*	0.98(1)		0.58		4.29(5)*
Glu		1.00(1)	4.20(4)		3.91(4)	1.37(1)	3.24(3)	3.20(3)		4.19(4)
Pro			2.15(2)			1.08(1)				1.10(1)
Gly		0.98(1)	1.32(1)	2.10(2)	0.88(1)	0.90(1)		0.35		3.84(4)
Ala	0.94(1)	1.00(1)	3.31(3)	2.98(3)	0.40	1.78(2)				4.00(4)
Val			1.08(1)		1.32(1)	2.63(3)				1.19(1)
Ile			2.93(3)		1.92(2)	0.98(1)				3.19(3)
Leu		1.02(1)	1.95(2)		2.08(2)	1.02(1)	0.97(1)		0.91(1)	3.04(3)
Tyr	1.09(1)		1.60(2)			1.50(2)	0.93(1)		1.08(1)	0.89(1)
Phe					0.85(1)					0.94(1)
Total	4	6	26	10	20	18	8	6	2	40
Yield(%)	26	72	17	60	14	32	27	9	26	86
Color reaction†	P		P			P	P	P	P	P
Purification‡	PE_1	—	DE_{52}, PE_2	PE_2	DE_{52}, PE_2	DE_{52}, PE_2	PE_1	PE_1	—	—

*From the sequence studies.
†P, Pauli-positive (20).
‡PE_1 and PE_2, paper electrophoresis at pH 3.6 and 6.5, respectively.
DE_{52}, chromatography on a DE_{52} column.

Nine peptides, T-1 to T-9, were isolated from typtic digest of Cm-ferredoxin. The amino acid compositions and characteristics of the isolated peptides are given in Table 2 and the sequence studies of these peptides are summarized in Fig. 1. Peptides T-1, T-2, T-4, T-6, T-7 and T-9 were completely sequenced by Edman degradations. Peptide T-7 must be the C-terminal peptide of the original protein; its C-terminal sequence, -Leu-Tyr, was consistent with the result obtained from carboxypeptidase A digestion of Cm-ferredoxin as described above. Partial sequences of peptides T-3 and T-5 were determined by Edman degradations and no sequence study of peptide T-8 was performed. As shown in Fig. 1, unexpected cleavages with trypsin were observed; Glu-Ala (10-11, 72-73), Glu-Thr (90-91) and Glu-Leu (96-97) bonds were cleaved. Judging from the cleavage site, these must be due to contamination of staphylococcal protease in the trypsin solution.

One large peptide S-1, which was useful for overlapping tryptic peptides, was obtained from staphylococcal protease digest of Cm-ferredoxin. The amino acid composition of this peptide is included in Table 2. The sequence of the unknown region of peptide T-3 was established after 24 steps of Edman degradation.

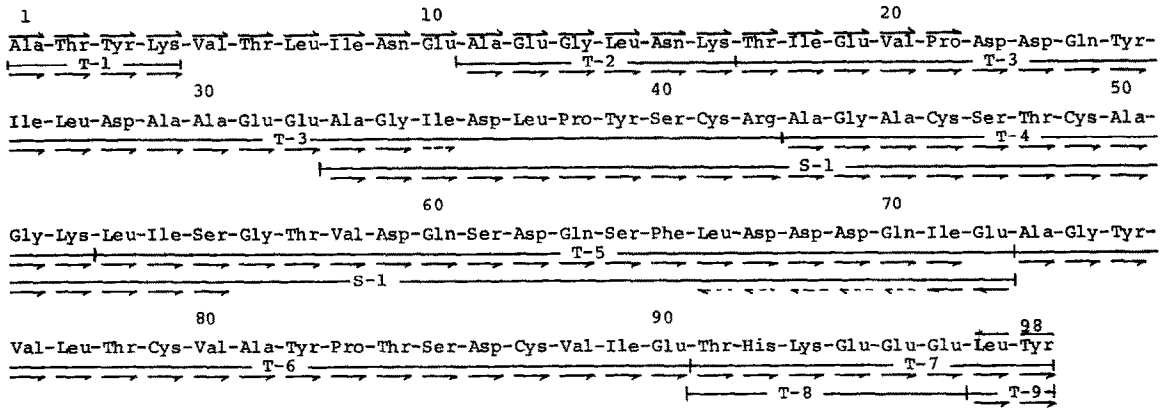

Fig. 1. Summary of sequence studies of *Mastigocladus* ferredoxin. T and S refer to tryptic and staphylococcal protease peptides, and Cys refers to Cm-cysteine. Arrows, (→) and (←), above the sequence represent Edman degradation and carboxypeptidase A digestion on Cm-ferredoxin, respectively. Arrows, (→) and (←), below the sequence represent Edman degradations and carboxypeptidase digestion on peptide fragments. Dotted arrows indicate amino acids released by carboxypeptidase without decision of sequence.

Carboxypeptidase Y digestion of peptide S-1 for 10 min released aspartic acid (0.53), amide (0.29), glutamic acid (1.06), isoleucine (0.77) and leucine (0.23). These results and staphylococcal protease specificity indicated that the C-terminal sequence of peptide S-1 was -Ile-Glu and other residues were near the C-terminus.

The complete amino acid sequence of Cm-ferredoxin of *Mastigocladus laminosus* is shown in Fig. 1. Edman degradation of the original Cm-protein allowed the alignment of peptides T-1, T-2, and T-3. The peptide corresponding to residues 5–10 could not be isolated. Peptide S-1 overlapped peptides T-3, T-4, and T-5. Peptide T-7 was the C-terminal peptide of Cm-ferredoxin. Therefore, peptide T-6 must be placed between peptides T-5 and T-7. The amino acid composition deduced from the sequence studies agreed with that of the original protein (Table 1), supporting the proposed sequence. Thus the complete amino acid sequence of *Mastigocladus laminosus* ferredoxin was established. The total number of residues was 98 with a MW of 10618 excluding iron and sulphur atoms. Residue 20 was valine (and not cysteine) thus reconfirming our earlier suggestion that the 4 invariant cysteines involved in the co-ordination of the [Fe–S] centre in plant type ferredoxins are at positions 41, 46, 49 and 80 [7].

Since the *M. laminosus* was grown at 58° and since ferredoxin is essential for photosynthesis by the alga, one can expect this ferredoxin to be highly heat-stable within the cellular environment. When we compared the thermal stabilities of ferredoxins isolated from *M. laminosus*, *Zea mays*, *Spirulina maxima* and *Spinacia oleracea*, we found that *M. laminosus* ferredoxin is remarkably stable at 65° even in the isolated state. It takes more than 2 hr incubation at 65° and 1 hr incubation at 75° to destroy the active [Fe–S] centre of *M. laminosus* ferredoxin. Among the other ferredoxins tested only *Zea mays* ferredoxin was able to survive heating to 65° and that only for 10 min. When discussing the higher thermal stability and oxygen-insensitivity of (4Fe–4S) *Bacillus stearothermophilus* ferredoxin compared to other bacterial ferredoxins, we suggested that the presence of a relatively higher number of glutamic acid residues and lower content of cysteine residues may contribute to the stability of the molecule [8]. However, comparison of the amino acid sequence of *M. laminosus* ferredoxin with those of other ferredoxins (Fig. 2) reveals no special features in the *M. laminosus* ferredoxin sequence to account for its extreme thermal stability. We have to conclude that only the tertiary structures of these ferredoxins will provide an answer to the question of relative heat stability.

There are two schools of thought regarding the evolutionary status of thermophiles. One is that the early waters of the world were not saline as at present and were very much warmer and that it is in such places that the most primitive forms of life may be found [9]. According to Brock [10], hot springs provide stable environments for ecosystems which have remained un-

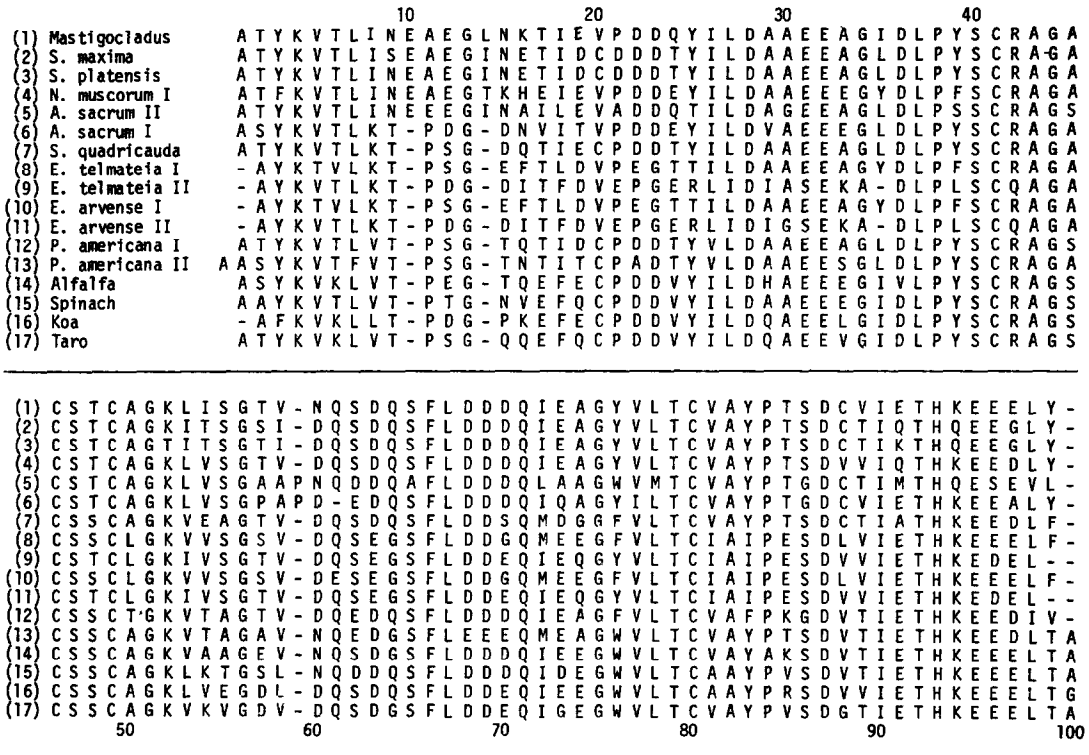

Fig. 2. Comparison of chloroplast-type ferredoxin sequences. 1. *M. laminosus* (present work); 2. *Spirulina maxima* [27]; 3. *S. platensis* [28a, b and c]; 4. *Nostoc muscorum* I [29]; 5. *Aphanothece sacrum* II [30]; 6. *Asacrum* I [31]; 7. *Scenedesmus quadricauda* [32]; 8. *Equisetum telmateia* I [33]; 9. *E. telmateia* II [33]; 10. *E. arvense* I [34]; 11. *E. arvense* II [34]; 12. *Phytolacca americana* I [35]; 13. *P. americana* II [35]; 14. Alfalfa (*Medicago sativa*) [36]; 15. Spinach (*Spinacia oleracea*) [20]; 16. Koa (*Leucaena glauca*) [37]; 17. Taro (*Colocasia esculenta* [38].

Abbreviations for amino acids: A = Ala, C = Cys, D = Asp, E = Glu, F = Phe, G = Gly, H = His, I = Ile, K = Lys, L = Leu, M = Met, N = Asn, P = Pro, Q = Gln, R = Arg, S = Ser, T = Thr, V = Val, Y = Tyr.

changed for thousands of years. *M. laminosus* can develop a wide range of morphological forms, a characteristic attributed to primitive algae [9]. It is perhaps the most abundant alga in natural and artificial hot springs and it has a nitrogenase activity even at 60°, the highest temperature at which biological nitrogen fixation was recorded on Earth [11]. All these properties suggest that *M. laminosus* may be one of the primitive forms of extant algae. The second school of thought assumes that the thermophilic flora has become specially adapted to their environment by mutations which increased the thermal stability of their proteins (see [12]). What can we learn about the evolutionary origin of *M. laminosus* from the amino acid sequence of its ferredoxin? A comparison of the primary structure of *M. laminosus* ferredoxin with those of all other ferredoxins so far sequenced (Fig. 2) show that this ferredoxin is closely homologous to the ferredoxins from the *Spirulina* species and from nitrogen-fixing blue-green alga, *Nostoc muscorum*. There is more than 75% homology in the sequence of these 4 ferredoxins whereas only 27% homology is observed between the amino acid sequences of all the 17 ferredoxins in Fig. 2. Fossil evidence [13] suggests that *Spirulina* is one of the earliest forms of life that has been present since the Pre Cambrian era. It is difficult to judge from the sequences of the ferredoxins alone as to whether *Spirulina* or *Mastigocladus* appeared first. But it is tempting to speculate that *M. laminosus* could be one of the earliest nitrogen-fixing algae to have evolved on Earth.

EXPERIMENTAL

Mastigocladus laminosus cells were cultured at 55–60° in a hot spring near Reykjavik, Iceland [14]. A suspension of cells in 20 mM Pi buffer, pH 8, was disrupted by homogenization in an MSE Atomix followed by compression through a Manton Gaulin homogenizer. The debris was removed from the cell homogenate by centrifugation and the supernatant was brought up to 50% satn with $(NH_4)_2SO_4$ to precipitate the pigments and biliproteins [15]. The ppt. was separated by centrifugation and the supernatant dialysed against 20 mM Tris–HCl buffer, pH 8. The dialysate was passed through a column of DEAE–cellulose (Whatman DE23) equilibrated with 0.2 M Cl^-. The ferredoxin adsorbed at the top of the DE23 column was eluted as a red band by passing 0.8 M Cl^- through the column. The crude ferredoxin was purified by standard techniques [16], viz. chromatography on DEAE–cellulose, gel filtration on Sephadex G-75 and chromatography on hydroxyl-apatite. All these operations starting from the disruption of the cells were done at 4°. Other plant and algal ferredoxins were purified by published procedures [17]. For determination of the thermal stability of ferredoxins the proteins were dialysed against a soln of 0.1 M NaCl in 50 mM HEPES buffer, pH 7.5. They were placed in 1 ml quartz cuvettes and in a thermostated cavity (65° or 75°) of a Unicam SP 800A UV spectrophotometer. The *A* of the protein at 420 nm was recorded automatically against time.

Amino acid composition and sequence. M. laminosus ferredoxin was carboxymethylated with iodoacetic acid according to the method of ref. [18] after removal of Fe and S by TCA treatment. Amino acid composition of Cm-ferredoxin was determined with an amino acid analyser (Beckman Model 120 B) according to the method of ref. [19]. The Cm-protein was hydrolysed with 6 N HCl in an evacuated, sealed tube at 110° for 24 and 72 hr. The Cm-ferredoxin (1.27 μmol) was digested with 0.2 mg of trypsin treated with tosylphenylalanine chloromethyl ketone (Worthington) in 1 ml of 0.1 M Tris–HCl buffer, pH 8, at 40° for 2 hr. The digestion mixture was directly applied to a Bio-Gel p-4 column (2 × 197 cm) which was eluted with 0.2 M ammonium bicarbonate pH 9 and the flow rate was 10 ml/hr. Fractions (2.2 ml) were collected and the *A* was monitored at 220 and 280 nm. Pooled fractions were lyophilized and subjected to analyses or further purifications by ion exchange chromatography on a DE-52 (Whatman) column or by paper electrophoresis at pH 3.6 or 6.5 [20]. In order to align the tryptic peptides, the Cm-ferredoxin (0.88 μmol) was also digested with 0.15 mg of staphylococcal protease (a gift of Dr. R. Ambler, Edinburgh University) in 1 ml of Tris–HCl buffer, pH 8 at 40° 3 h. The digest was separated on a Bio–Gel P-10 column under the same conditions as described above. A manual Edman degradation procedure [21] was applied to the Cm-ferredoxin and peptides to determine their N-terminal sequences. PTH-derivatives were identified by TLC on Si gel using various solvent systems [22, 23]. PTH-histidine and PTH-arginine were identified by paper electrophoresis at pH 6.5 [24]. The C-terminal sequences of the Cm-protein and Peptide S-1 were determined by digesting them with carboxypeptidase A [25] or Y [26].

Acknowledgements—This work was supported by grants from the Ministry of Education Science and Culture of Japan, U.K. Science Research Council and the Schweizerischer National fonds Zur Forderung der wissenschaftlichen Forschung Project 3, 005-076.

REFERENCES

1. Hall, D. O., Rao, K. K. and Cammack, R. (1975) *Sci. Prog. Oxf.* **62**, 285.
2. Buchanan, B. B. and Arnon, D. I. (1970) *Adv. Enzymol.* **33**, 119.
3. Matsubara, H., Hase, T., Wakabayashi, S. and Wada, K. (1978) in *Evolution of Protein Molecules* (Matsubara, H. and Yamanaka, T., eds.) (in press). Centre for Academic Publications Japan, Tokyo.
4. Fogg, G. E., Stewart, W. D. P., Fay, P. and Walsby, A. E. (1973) in *The Blue-Green algae*, p. 278. Academic Press, London.
5. Hall, D. O., Cammack, R. and Rao, K. K. (1973) *Pure Appl. Chem.* **34**, 553.
6. Cammack, R., Rao, K. K., Bargeron, C. P., Hutson, K. G., Andrew, P. W. and Rogers, L. J. (1977) *Biochem. J.* **168**, 205.
7. Kagamiyama, H., Rao, K. K., Hall, D. O., Cammack, R. and Matsubara, H. (1975) *Biochem. J.* **145**, 121.
8. Hase, T., Ohmiya, N., Matsubara, H., Mullinger, R., Rao, K. K. and Hall, D. O. (1976) *Biochem. J.* **159**, 55.
9. Chapman, V. J. and Chapman, D. J. (1973) in *The Algae*, p. 17. Macmillan, London.
10. Brock, T. D. (1967) *Science* **158**, 1012.
11. Stewart, W. D. P. *et al.* Ref. [4], p. 339.
12. Stanier, R. Y., Adelberg, A. E. and Ingraham, J. L. (1976) in *General Microbiology* 4th edn, p. 308. Macmillan, London.
13. Schopf, J. W. (1970) *Biol. Rev.* **45**, 319.
14. Binder, A., Locher, P. and Zuber, H. (1972) *Arch. Hydrobiol.* **70**, 541.
15. Gysi, J. and Zuber, H. (1974) *FEBS Letters* **48**, 209.
16. Hall, D. O. and Rao, K. K. (1977) in *Encyclopedia of Plant Physiology*, Vol. 5, Photosynthesis I (Trebst, A. and Avron, M., eds.) pp. 206–216. Springer, Berlin.
17. Hall, D. O., Rao, K. K. and Cammack, R. (1972) *Biochem. Biophys. Res. Commun.* **47**, 798.
18. Crestfield, A. M., Moore, S. and Stein, W. H. (1963) *J. Biol. Chem.* **238**, 622.
19. Spackman, D. H., Stein, W. H. and Moore, S. (1958) *Analyt. Chem.* **30**, 1190.
20. Matsubara, H. and Sasaki, R. M. (1968) *J. Biol. Chem.* **243**, 1732.
21. Blombäck, B., Blombäck, M., Edman, P. and Hessel, B. (1966) *Biochim. Biophys. Acta* **115**, 371.
22. Cherbuliez, E., Baehler, Br. and Rabinowitz, J. (1964) *Helv. Chim. Acta* **47**, 1350.
23. Jeppson, J. O. and Sjoquist, J. (1967) *Analyt. Biochem.* **18**, 264.

24. Edman, P. and Henschen, A. (1975) in *Protein Sequence Determination* (Needleman, S. B., ed.) pp. 232–279. Springer, Berlin.
25. Matsubara, H., Sasaki, R. M. and Chain, R. K. (1968) *J. Biol. Chem.* **243**, 1725.
26. Hayashi, R. (1977) *Methods Enzymol.* **47**, 84.
27. Tanaka, M., Haniu, M., Yasunobu, K. T., Rao, K. K. and Hall, D. O. (1975) *Biochemistry* **14**, 5535.
28. (a) Wada, K., Hase, T., Tokunaga, H. and Matsubara, H. (1975) *FEBS Letters* **55**, 102; (b) Matsubara, H., Wada, K. and Masaki, R. (1976) in *Iron and Copper Proteins* (Yasunobu, K. T. *et al.*, eds.) pp. 1–15. Plenum, New York; (c) Tanaka, M., Haniu, M., Yasunobu, K. T., Rao, K. K. and Hall, D. O. (1976) *Biochem. Biophys. Res. Commun.* **69**, 759.
29. Hase, T., Wada, K., Ohmiya, M. and Matsubara, H. (1976) *J. Biochem* **80**, 993.
30. Hase, T., Wakabayashi, S., Wada, K. and Matsubara, H. (1978) *J. Biochem.* **83**, 761.
31. Hase, T., Wada, K. and Matsubara, H. (1976) *J. Biochem* **79**, 329.
32. Sugeno, K. and Matsubara, H. (1969) *J. Biochem.* **244**, 2979.
33. Hase, T., Wada, K. and Matsubara, H. (1977) *J. Biochem.* **82**, 267.
34. Hase, T., Wada, K. and Matsubara, H., *ibid.*, 277.
35. Wakabayashi, S., Hase, T., Wada, K., Matsubara, H., Suzuki, K. and Takaichi, S. (1978) *J. Biochem.* **83**, 1305.
36. Keresztes-Nagy, F., Perini, F. and Margoliash, E. (1969) *J. Biol. Chem.* **244**, 891.
37. Benson, A. M. and Yasunobu, K. T. (1969) *J. Biol. Chem.* **244**, 955.
38. Rao, K. K. and Matsubara, H. (1970) *Biochem. Biophys. Res. Commun.* **38**, 500.