Biochem. J. (1989) 263, 981–984 (Printed in Great Britain)

The amino acid sequence of a flavodoxin from the eukaryotic red alga *Chondrus crispus*

Sadao WAKABAYASHI,* Tohru KIMURA,* Keiichi FUKUYAMA,* Hiroshi MATSUBARA* and Lyndon J. ROGERS†

*Department of Biology, Faculty of Science, Osaka University, Toyonaka, Osaka 560, Japan, and
†Department of Biochemistry, University College of Wales, Aberystwyth, Dyfed SY23 3DD, Wales, U.K.

The amino acid sequence of the constitutive flavodoxin from the red alga *Chondrus crispus* was determined from the analyses of peptide fragments derived by enzymic digestions of the carboxymethylated protein. This is the first sequence reported for a flavodoxin from a eukaryote. The protein is composed of 173 amino acid residues and is a member of the longer-chain group of flavodoxins. The extent of sequence homology to the three other flavodoxins in the group for which sequences are available is in the range 36–39 %, with the most strongly conserved regions being those implicated in binding of the FMN, the redox-active prosthetic group. Nevertheless, *Chondrus crispus* flavodoxin stands apart in a number of respects, in particular the possession of an unusually high content of proline, with these residues distributed more or less regularly along the peptide chain.

INTRODUCTION

Flavodoxins are electron-transfer proteins containing one FMN molecule as the redox-active prosthetic group [1]. They function in various electron-transport systems in a range of micro-organisms, usually replacing ferredoxin when the organism is grown under iron-limiting conditions. They are functionally interchangeable with the latter protein in most ferredoxin-dependent oxido-reduction reactions, including photosynthetic electron transport [2–4]. So far, flavodoxins have been isolated from several prokaryotes and two eukaryotic algae, *Chlorella fusca* [5] and *Chondrus crispus*. *Chond. crispus*, a red macroalga, is unusual because it synthesizes flavodoxin constitutively and possesses only negligible amounts of ferredoxin [6].

Structure–function relationships in these proteins are of considerable interest. The ferredoxins have been studied extensively in this respect (see ref. [7]), and the flavodoxins are similarly of interest, being among the smallest flavoproteins, and therefore the simplest models for studying the chemistry and biochemistry of this group of proteins [1].

To date the primary structures of the flavodoxins from *Clostridium* strain MP [8], *Megasphaera elsdenii* [9], *Desulfovibrio vulgaris* [10] and *Azotobacter vinelandii* [11] have been determined by protein sequencing. The structures of some others have been deduced from nucleotide sequencing of a flavodoxin gene, in *Anacystis nidulans* (*Synechococcus* 6301) [12], and of the *nifF* genes (coding for a flavodoxin functional in the nitrogen-fixing system) from *Az. vinelandii* [13] and *Klebsiella pneumoniae* [14]. The three-dimensional structures of the flavodoxins from *D. vulgaris* [15], *Clostridium* MP [16] and *Ana. nidulans* [17] have been determined, the last-mentioned with some uncertainties since the complete amino acid sequence was not then available. *Chond. crispus* flavodoxin has been crystallized and preliminary diffraction studies have been reported [18]. We now report the complete amino acid sequence of this flavodoxin, and compare it with the other flavodoxin sequences.

METHODS

Flavodoxin was purified from *Chond. crispus* as described previously [6] and carboxymethylated [19]. The carboxymethylated flavodoxin was separately digested with lysyl-endopeptidase (Wako Pure Chemical Industries, Osaka, Japan) or staphylococcal V8 proteinase (Miles Laboratories) in 0.1 M-Tris/HCl buffer, pH 8.0, at 40 °C for 2 h. The digests were applied to a reverse-phase column (Shodex ODSpak F-511A; Showa Denko, Tokyo, Japan) and separated with a linear concentration gradient of acetonitrile in 0.1 % (v/v) trifluoroacetic acid with an Irica (Kyoto, Japan) model Σ h.p.l.c. system. The A_{220} of the eluate was monitored, and the peak fractions were collected and their amino acid compositions determined. After peptide hydrolysis with 6 M-HCl for 24 h at 110 °C in an evacuated and sealed tube amino acid analyses were performed [20] with an Irica model A-5500 amino acid analyser. Amino acid sequences of the protein and peptides were determined by manual Edman degradation [21], and the phenylthiohydantoin derivatives were identified by h.p.l.c. [22]. Some lysyl-endopeptidase-derived peptides were covalently attached to aminopropyl-glass by *p*-phenylene di-isothiocyanate [23] and analysed with an LKB 4030 solid-phase peptide sequencer. The *C*-terminal sequence was determined by digestion of the protein with carboxypeptidase A (Worthington Corp.) in 0.1 M-Tris/HCl buffer, pH 8.0, at 40 °C with direct analysis of the released amino acids by using the amino acid analyser.

† To whom correspondence should be addressed.
These sequence data have been submitted to the EMBL/GenBank Data Libraries.

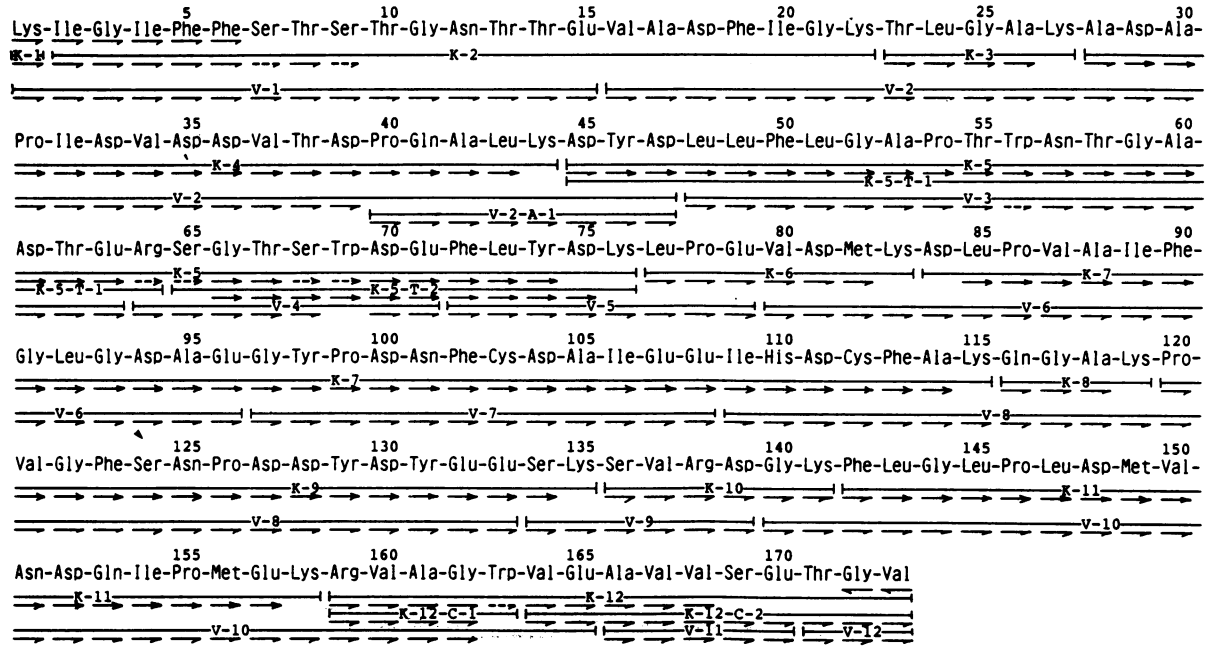

Fig. 1. Summary of the sequence studies of *Chond. crispus* flavodoxin

K- and V- refer to the peptides obtained by lysyl-endopeptidase and staphylococcal V8 proteinase digestion respectively. T-, C- and A- refer to the subfragments obtained by the second cleavage of the peptides with trypsin, chymotrypsin and acid treatment respectively. Arrows →, ← and ↞ indicate that the residues were identified by manual and automated Edman degradation and carboxypeptidase digestion respectively. A broken arrow indicates an ambiguous identification.

RESULTS AND DISCUSSION

The *N*-terminal sequence of *Chond. crispus* flavodoxin up to the 34th residue was reported previously [24], and the present sequence studies of carboxymethylated flavodoxin and of peptides K-1, K-2, K-3, K-4, V-1 and V-2 confirmed the earlier analysis, except for serine at residues 29 and 33, the latter tentative, both of which were now firmly identified as aspartic acid. Carboxypeptidase A treatment of the carboxymethylated flavodoxin released only valine (62%) at 10 min and glycine (52%) and valine (103%) after 2 h. Therefore the *C*-terminal sequence is -Gly-Val. Lysyl-endopeptidase produced 12 peptides, which were successfully separated by reverse-phase h.p.l.c. Peptides K-2, K-3, K-6, K-8, K-10 and K-12 were analysed by manual Edman degradation and peptides K-4, K-5, K-7, K-9 and K-11 by use of the solid-phase sequencer. Peptides K-5 and K-12 were further digested with trypsin and chymotrypsin respectively, and the derived subfragments were purified by reverse-phase h.p.l.c. These subfragments were also sequenced by automated or manual Edman degradation. The largest peptide, V-2, was treated with 70% (v/v) formic acid to cleave the acid-labile Asp–Pro bond [25]. The resulting subfragments were purified by h.p.l.c. and sequenced. These two sets of peptides overlapped each other, and in conjunction with previous analyses the complete amino acid sequence could be firmly established (Fig. 1). The amino acid composition obtained by the direct analysis of the protein agreed well with that deduced from the sequence (in parentheses), i.e. Cys (as carboxymethylcysteine) 1.7 (2), Asp with Asn 27.5 (23+5), Thr 10.8 (11), Ser 8.5 (8), Glu with Gln 16.2 (12+3), Pro 9.7 (10), Gly 15.7 (16), Ala 14.3 (14), Val 13.0 (13), Met 3.9 (3), Ile 7.7 (8), Leu 12.6 (12), Tyr 5.1 (5), Phe 9.5 (10), Lys 9.3 (11), His 1.2 (1), Arg 3.2 (3) and Trp (3), and was in good agreement with an earlier report [6]. The M_r of the apoprotein was calculated to be 18871; the value derived from meniscus-depletion sedimentation equilibrium was 20800 [26].

The sequence was compared with those from other sources (Fig. 2); to obtain maximum extent of homology insertions or deletions are introduced as necessary. This adjustment is particularly notable at residues 123–143, which reflects the division of the flavodoxins into two groups. The flavodoxins from the non-photosynthetic bacteria have an M_r of about 15000, with the exception of the flavodoxin from *Az. vinelandii*, which like those from the cyanobacteria and algae has an M_r of about 20000. The placement of the long deletion in the shorter-chain flavodoxins is based on comparisons of the three-dimensional structures of *Ana. nidulans*, *D. vulgaris* and *Clostridium* MP flavodoxins.

Chond. crispus flavodoxin is the only representative from eukaryotes, but clearly shows some degree of homology to the other flavodoxins; within the longer-chain group the extent of homology is 36–39%, rather greater than that seen (25–35%) among members of the small-M_r group [10]. In a comparison of sequences the recognition of residues that are invariant throughout suggests that these may be important for structure and/or function. For the [2Fe-2S] ferredoxins a comparison of 20 or so sequences [27] showed that the invariant residues, apart from the four cysteine residues that bind the iron–sulphur centre, were only glycine residues at positions 46, 53, 58 (one exception) and 77 (two exceptions). For the flavodoxins as aligned in Fig. 2 the invariant residues are Thr-10, Gly-11, Thr-13, Ala-17, Ile-20 and

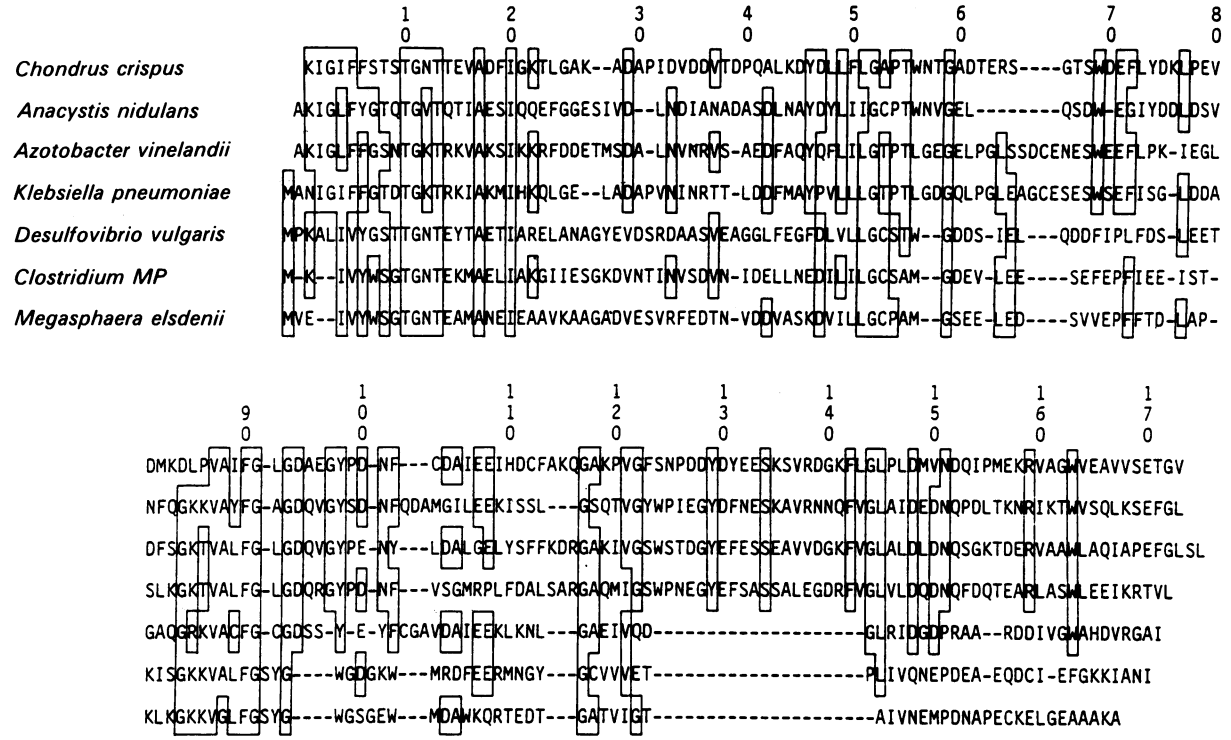

Fig. 2. Comparison of flavodoxin sequences

The sequences are of the flavodoxins from *Chond. crispus* (the present work), *Ana. nidulans* [12], *Az. vinelandii* [11,13], *K. pneumoniae* [14], *D. vulgaris* [15], *Clostridium* MP [8] and *M. elsdenii* [9]. For consistency the numbering of residues is that for *Chond. crispus* flavodoxin in Fig. 1. Residues common in at least four flavodoxins are boxed.

Gly at positions 52, 59, 84 (one exception), 91, 94 and 117. The conservation of glycine residues suggests the importance of these in stabilizing the structure of the flavodoxin molecule. Some of the other invariant residues are sited at the FMN–protein interface, which involves particularly the sequences 8–13, 54–60 and 93–104; inspection of Fig. 2 shows the high conservation of sequence in these regions, where also most of the invariant glycine residues are to be found. In *D. vulgaris* flavodoxin (following the numbering in Fig. 2 for *Chond. crispus* flavodoxin), Ser-8, Thr-10, Thr-13 and Ser-54 hydrogen-bond to the FMN phosphate, and Asn-12, Thr-55 and Asp-150 to the 4′-hydroxy group of the ribityl side chain. In *Chond. crispus* flavodoxin the differences are Thr-8, Pro-54 and Val-150 and in the protein from *Ana. nidulans* Thr-8, Val-12 and Pro-54. These, in conjunction with other changes in the sequences, presumably produce subtle differences in flavin environment that account for the observed differences in redox properties of these flavodoxins [28].

The flavodoxin from *Chond. crispus* is unusual in other respects compared with *Ana. nidulans* flavodoxin in undergoing a marked conformational change on dissociation of the prosthetic group [26]. A clue to this behaviour may lie in the relative amounts of proline, ten residues in *Chond. crispus* flavodoxin, three in *Ana. nidulans* flavodoxin, this being the most notable difference in amino acid compositions. Moreover, in the red-algal flavodoxin these residues are distributed more or less regularly along the peptide chain (positions 31, 40, 54, 78, 86, 99, 120, 126, 146 and 155). The possible importance of these feature and also the details of FMN binding will become clearer when elucidation of the tertiary structure of *Chond. crispus* flavodoxin is completed.

Note added in proof (received 4 September 1989)

After submission we noted publication of the sequence of the flavodoxin gene from the cyanobacterium *Anabaena variabilis* by Leonhardt & Straus [29]. This flavodoxin shows 67% homology with that from *Ana. nidulans* and 35% homology with *Chond. crispus* flavodoxin. With respect to detailed comment in our paper *Anabaena variabilis* flavodoxin agrees with that for *Ana. nidulans* flavodoxin (except for Lys-12 in place of Val-12 and possession of only a single proline residue), and is diverse from *Chond. crispus* flavodoxin.

REFERENCES

1. Mayhew, S. G. & Ludwig, M. L. (1975) Enzymes 3rd Ed. **12**, 57–118
2. Yoch, D. C. & Valentine, R. C. (1972) Annu. Rev. Microbiol. **26**, 139–162
3. Hutber, G. N., Hutson, K. G. & Rogers, L. J. (1977) FEMS Microbiol. Lett. **1**, 193–196
4. Hutber, G. N. & Rogers, L. J. (1981) Photosynth. Res. **2**, 269–280
5. Zumft, W. G. & Spiller, H. (1971) Biochem. Biophys. Res. Commun. **45**, 112–118
6. Fitzgerald, M. P., Husain, A. & Rogers, L. J. (1978) Biochem. Biophys. Res. Commun. **81**, 630–635
7. Rogers, L. J. (1987) in The Cyanobacteria (Fay, P. & Van Baalen, C., eds.), pp. 35–67, Elsevier Science Publishers, Amsterdam
8. Tanaka, M., Haniu, M., Yasunobu, K. T. & Mayhew, S. G. (1974) J. Biol. Chem. **249**, 4393–4396

9. Tanaka, M., Haniu, M., Yasunobu, K. T., Mayhew, S. G. & Massey, V. (1973) J. Biol. Chem. **248**, 4354–4366; revisions **249**, 4397
10. Dubourdieu, M. & Fox, J. L. (1977) J. Biol. Chem. **252**, 1453–1463
11. Tanaka, M., Haniu, M., Yasunobu, K. T. & Yoch, D. C. (1977) Biochemistry **16**, 3525–3537
12. Laudenbach, D. E., Reith, M. E. & Straus, N. A. (1988) J. Bacteriol. **170**, 258–265
13. Bennett, L., Jacobson, M. & Dean, D. R. (1988) J. Biol. Chem. **263**, 1364–1369
14. Drummond, M. (1985) Biochem. J. **232**, 891–896
15. Watenpaugh, K. D., Sieker, L. C. & Jensen, L. H. (1973) Proc. Natl. Acad. Sci. U.S.A. **70**, 3857–3860
16. Burnett, R. M., Darling, G. E., Kendall, D. S., Lequesne, M. E., Mayhew, S. G., Smith, W. W. & Ludwig, M. L. (1974) J. Biol. Chem. **249**, 4383–4392
17. Smith, W. W., Pattridge, K. A., Ludwig, M. L., Petsko, G. A., Tsernoglou, D., Tanaka, M. & Yasunobu, K. T. (1983) J. Mol. Biol. **163**, 737–755
18. Fukuyama, K., Matsubara, H., Katsube, Y. & Rogers, L. J. (1989) J. Biochem. (Tokyo) **105**, 348–350
19. Crestfield, A. M., Moore, S. & Stein, W. H. (1963) J. Biol. Chem. **238**, 622–627
20. Spackman, D. H., Moore, S. & Stein, W. H. (1958) Anal. Chem. **30**, 1190–1206
21. Blombäck, B., Blombäck, M., Edman, P. & Hessel, B. (1966) Biochim. Biophys. Acta **115**, 371–396
22. Zimmerman, C. L., Appella, E. & Pisano, J. J. (1977) Anal. Biochem. **77**, 569–573
23. Laursen, R. A. (1977) Methods Enzymol. **47**, 277–288
24. Takruri, I., Boulter, D., Fitzgerald, M. P., Hutber, G. N. & Rogers, L. J. (1986) Phytochemistry **25**, 2113–2115
25. Landon, M. (1977) Methods Enzymol. **47**, 145–149
26. Fukuyama, K., Matsubara, H., Katsube, Y., Sykes, G. A. & Rogers, L. J. (1989) Biochem. Soc. Trans., in the press
27. Masui, R., Wada, K., Matsubara, H., Williams, M. M. & Rogers, L. J. (1988) Phytochemistry **27**, 2817–2820
28. Sykes, G. A. & Rogers, L. J. (1984) Biochem. J. **217**, 845–850
29. Leonhardt, K. G. & Straus, N. A. (1989) Nucleic Acids Res. **17**, 4384

Received 17 August 1989; accepted 6 September 1989