## — Article —

# Amino Acid Sequences of Ferredoxins from Several Species of Genus *Ephedra* $\stackrel{ au}{\sim}$

Yoshiki MINO\*, Takashi AZUMA, and Takaji SATO

Laboratory of Analytical Chemistry, Osaka University of Pharmaceutical Sciences 4-20-1 Nasahara, Takatsuki, Osaka 569-1094, Japan. (Received December 17, 2014; Accepted January 15, 2015)

Abstract The complete amino acid sequences of [2Fe-2S] ferredoxins (Fds) from seven species of genus Ephedra (E. sinica, E. distachya, E. equisetina, E. viridis, E. intermedia, E. foliate, and E. americana) were determined by automated Edman degradation of the entire S-carboxymethylcysteinyl proteins and of the peptides obtained by enzymatic digestion. E. sinica-Fds (I and II), which differ from one another in the amino acid residue at position 95 (Ile for I and Leu for II), have unique amino acid sequences, which include Asn-14, Asp-30, Met-51, Cys-85, Gln-88, and Gln-91, and a deletion of one amino acid residue at the carboxyl terminus. E. distachya and E. equisetina have the same Fds (I and II) as E. sinica-Fds. E. viridis and E. intermedia only have a Fd that is the same as E. sinica-Fd I. E. foliate-Fd shows a difference of only one amino acid residue (Val at position 95) compared to other Ephedra Fds. In contrast, E. americana shows differences in five and six amino acid residues from the other Ephedra-Fds, which suggests that E. americana is somewhat distantly related to the others. These Ephedra Fds have 21-34 differences in their amino acid sequences compared to those of Angiospermous plants except for Pueraria lobata. In contrast, 38-40 differences were observed when they were compared to Equisetum telmateia and *E. arvense* (horsetail plants). This suggests that *Ephedra* plants are remotely related taxonomically to horsetail plants, although they seem to be morphologically similar. In practice, *Ephedra* plants and horsetail plants belong to different phyla: Spermatophyta and Pterophyta, respectively.

Key words — ferredoxin; *Ephedra sinica*; genus *Ephedra*; Ephedraceae; amino acid sequence; protein chemotaxonomy

# INTRODUCTION

Although classical taxonomy, which is based largely on morphological and anatomical characteristics, is still a dominant concept in plant classification, chemical taxonomy has also been used to help clarify the relationships among genera and species when there is a need to confirm or revise an existing taxonomy. We have proposed the term 'protein chemotaxonomy' to describe molecular taxonomy based on the primary structures of common plant proteins, instead of so-called secondary metabolites. To evaluate the effectiveness of this concept, we carried out a series of studies on the family Solanaceae, using ferredoxin (Fd), an iron-sulfur electron-transfer protein.<sup>1)</sup> This protein was chosen because it is easy to isolate and has an appropriate molecular weight for determining the primary structure. Previously, we reported the primary structures of Fds from 14 solanaceous plants,<sup>2-11)</sup> one leguminous plant,<sup>12)</sup> and one alariaceous plant.<sup>13)</sup> Our recent results suggested that their amino acid sequences were related to their taxonomic position among plants that belong to the same genus or family, but not among plants in different families, although there may not be enough sequence data to reach any definite conclusions. It may be worthwhile to determine the amino acid sequences that belong to different families. These considerations

 $<sup>\</sup>stackrel{\wedge}{\succ}$  Part 14 in the series "Protein Chemotaxonomy".

<sup>\*</sup> e-mail: mino@gly.oups.ac.jp

led us to elucidate the amino acid sequence of Fd from *Ephedra sinica* (Ephedraceae, Ephedrales, Gnetopsida, Gymnospermae, Spermatophyta), the dried aerial part of which is one of the most commonly used traditional medicines in China, Korea, and other Asian countries for the treatment of asthma, allergic rhinitis, upper respiratory infection, and cold.

In this study, we determined the primary structures of Fds from *E. sinica* and several species of genus *Ephedra* and compared them with those of Fds from other higher plants.

## MATERIALS AND METHODS

**Materials** *E. sinica* was cultivated in the herb garden at Osaka University of Pharmaceutical Sciences. The fresh leaves of *E. distachya*, *E. equisetina*, *E. viridis*, *E. foliate*, and *E. americana* were kind gifts from the Nippon Shinyaku Institute for Botanical Research (Kyoto, Japan). The fresh leaves of *E. intermedia* were obtained from the Research Center for Medicinal Plant Resources, National Institute of Biomedical Innovation (Tsukuba, Ibaragi, Japan).

**Isolation of ferredoxin** Each Fd (*ca.* 4 mg) was purified from the fresh aerial parts (*ca.* 500 g) of each *Ephedra* plant as described previously,<sup>2. 6)</sup> except that 0.02 M Tris-HCl buffer, pH7.5, containing 0.5% Tween 80 was used for the extraction of Fd from the plant sample instead of a buffer without a surface-active agent.

Sequence determination The amino acid sequences of Fds were determined using a gas-phase protein sequencer with automated Edman degradation of S-carboxymethylcysteinyl (Cm) Fd and the peptides obtained by lysyl endopeptidase, trypsin, or endoproteinase Asp-N digestion. The peptides were purified by reversed-phase HPLC using a  $\mu$ -Bondasphere C<sub>18</sub>-100Å column (0.39×15cm, Waters) with a solvent system consisting of TFA- MeCN-H<sub>2</sub>O (A=0.1% TFA, B=MeCN containing 0.1% TFA) with a gradient program of 0-40% B in 50 min, flow rate 1 ml min<sup>-1</sup>. C-terminal analysis was carried out with carboxypeptidase Y.

The details of the procedure and other methods have been described previously.<sup>2, 6)</sup>

**Construction of a phylogenetic tree** A phylogenetic tree was constructed from the amino acid sequences (97 residues) of higher-plant Fds (39 species) using the unweighed pair-group method with arithmetical averages (UPGMA) as described by Nei (1994) (GENETYX software, Software Development, Japan).<sup>14)</sup>

### **RESULTS AND DISCUSSION**

**Properties of ferredoxins** The absorption maxima in the UV-Vis spectrum of *E. sinica* (*Es*)-Fd were at 275, 285(sh), 330, 420, and 465 nm, and showed  $A_{max}/A_{275nm}$  ratios of 0.65, 0.44, and 0.38, respectively. This spectrum was characteristic of [2Fe-2S] Fds from other higher plants.<sup>15)</sup> The molar absorption coefficient at 420 nm, based on the spectrum and protein determination, was 11000 M<sup>-1</sup>cm<sup>-1</sup>, which was similar to those of other higher-plant Fds.<sup>1,15)</sup> The biological activities and other physico-chemical properties of *Es*-Fd will be published elsewhere. The other *Ephedra* Fds exhibited properties similar to those of *Es*-Fd.

**Sequence determination** The sequencing strategy for *Es*-Fd is summarized in Fig.1. The analytical results regarding the amino acid compositions of Cm-Fd and the peptides obtained by enzymatic digestion were consistent with the derived sequences. Automated Edman degradation of *Es*-Cm-Fd yielded the amino-terminal sequence up to the 42nd cycle. Lysyl endopeptidase digestion gave two short peptides [L-1 (1-4) and L-2 (5-6)] and four long peptides [L-3 (7-50), L-4 (51-82), L-5 (83-96), and L-5' (83-96)]. Although Lys-91 was conserved in almost all of the Fds, except for Gleichenia japonica (fern)-Fd, in this case the residue was changed to Gln-91. These peptides were isolated by reversed-phase HPLC; their  $t_{\rm R}$  values were 14.4 for L-1(1-4), 29.6 for L-5(83-96), 30.4 for L-5' (83-96), 46.4 for L-3(7-50), and 46.8 min for L-4(51-82), while L-2 was missing. The isolation of L-5 (83-96) and L-5' (83-96) in almost the same amounts clearly indicates the existence of isoforms Fd-I and -II as Es-Fd. Sequence analyses of L-5 and L-5' clarified the sequences of 83-96 and a difference in the amino acid residue at position 95 between Fd-I and -II (Ile for I and Leu for II). Edman degradation of L-3-T-2, obtained by tryptic digestion of L-3 (7-50), confirmed the sequence of 41-50. Since there was not enough of the peptides, L-4 (51-82), to determine the sequence near the carboxyl terminus of the peptide, a proper short peptide containing the carboxyl terminus was needed. Endoproteinase Asp-N digestion of L-4 should give several short peptides [L-4-D-1 (51-56), L-4-D-2 (57-59), L-4-D-3 (60-64), and L-4-D-4 (65-82)]. These peptides were also isolated by HPLC; their  $t_{\rm R}$  values were 20.8 for L-4-D-1,2 (51-59), 31.6 for L-4-D-3, and 41.6 min for L-4-D-4, while L-4-D-1 and L-4-D2 were missing because of their small yields. Sequence analysis of L-4-D-4 confirmed the end part of 65-82. The N-terminal sequence was confirmed by the isolation of L-1 (Ala-Thr-Tyr-Lys). In addition, carboxypeptidase Y digestion of Cm-Fd for different periods of time suggested that the C-terminal sequence was reasonably consistent with the C-terminal sequence obtained by Edman degradation of the peptide, L-5 (83-96). These results led to the complete amino acid sequences for *Es*-Fds, as shown in Fig. 1.

In the case of *E. americana* (*Ea*)-Fd, due to the lack of Lys-82, no peak appeared near 30 min ( $t_R$ ) in the chromatogram of the peptides obtained by lysyl endopeptidase digestion. Instead, the long peptide 51-96 appeared at 46.8 min. Sequence analyses of this peptide clarified the sequences of 51-96 of *Ea*-Fd. This result was confirmed by sequence analyses of the short peptides obtained by Endoproteinase Asp-N digestion



Fig. 1. Amino Acid Sequences of Ephedra sinica Ferredoxins

Arrows  $(\rightarrow)$  and  $(\leftarrow)$  represent residues determined by automated Edman degradation and carboxypeptidase Y digestion, respectively. L (1-5), T -2, and D-4 represent peptides obtained from lysyl endopeptidase, trypsin, and endoproteinase Asp-N digestion, respectively. Only the amino acid sequence of ferredoxin I is shown; for ferredoxin II, the difference in the amino acid residue at position 95 (Leu instead of IIe) is shown in parentheses.

(1)

(2)

(3)

(4)

(5)

of the long peptide. The other *Ephedra* Fds could be analyzed in almost the same manner as for *Es*-Fd.

Figure 2 shows a comparison of the amino acid sequences among Ephedra plant-Fds. E. sinica has two isoforms of Es-Fd. These isoforms differed from one another in the amino acid residues at position 95; Ile for Es-Fd I and Leu for Es-Fd II. E. distachya and E. equisetina also have two isoforms, which have the same amino acid sequences as Es-Fds I and II. In contrast, E. viridis and E. intermedia have only one kind of Fd, which has the same amino acid sequence as Es-Fd I. A minor (ca. 20%) E. intermedia (Ei)-Fd showed differences in two or three amino acids compared to Es-Fd I or II. The Fd of E. foliate differs from Es-Fd I in only one amino acid residue at position 95, which gives Val instead of Ile or Leu. Interestingly, the Fd from E. americana had five or six differences in the amino acid sequence compared to the other Ephedra Fds, which suggests that E. americana is somewhat remotely related to the other Ephedra plants, although the other Ephedra plants are very closely related to each other. It is also very interesting that only E. americana does not contain the alkaloid ephedrine.

In Fig. 3, these amino acid sequences are compared to those of higher-plant Fds.<sup>2-13, 16-18)</sup> In comparison to other Fds, a noticeable feature of the present representative sequence, *Es*-Fd, is the isoform with Ile or Leu at position 95 from the amino terminus and a deletion of one amino acid residue at the carboxyl terminus. In comparison to other higher-plant Fds, differences were observed at Phe-7, as with *Brassica napus* (*Bn*) (Cruciferae)-and *Gleichenia japonica* (*Gj*)(Filicales)-

Fds, at Ile-16, as with Phytolacca americana (Pa)-, P. esculenta (Pe)(Phytolaccaceae)-and Gi-Fds, at Val-24, as with Bn-, Petroselinum sativum (Ps) (Umbelliferae), Pa- and Pe-Fds, at Leu-33, as with Pa-, Pe- and Gj-Fds, at Asn-55, as with Solanum lyratum (Solanaceae)-, Nicotiana tabacum (Solanaceae)-, and Trifolium pratense (Leguminosae)-Fds, at Ala-94, as with Solanum indium (Solanaceae)- and Lycium chinense (Solanaceae)-Fds, at Ile-95, as with Ps-, Pa-, and Pe-Fds, and at Ala-96, as with Pe-Fd. The residues Asn-14, Asp-30, Met-51, Ser-70, Cys-85, Gln-88, and Gln-91 were only observed in the primary structure of this Es-Fd among these higher-plant Fds. These residues are characteristic of Es-Fd. The residues Met-2 and Leu-8 are also characteristic of Ea-Fd. In Fds, the sequence 35-50, including the sequence -C39-C44-C47-, which participates in chelation to iron atoms, the sequence 74-77, which contains the last cysteine ligand (-C77-) for the iron atom, and the region 83-93 are almost perfectly conserved. This was also true in the case of Ephedra plant-Fds, except for four differences observed in the region 83-93.

**Taxonomic Considerations** Ephedrales consist of a single family (Ephedraceae) containing a single genus (*Ephedra*), and are known as the jointfirs because they have long slender branches which bear tiny scale-like leaves at their nodes. The aerial parts of some *Ephedra* plants have been traditionally used as a stimulant, but are controlled substances today in many jurisdictions because of the risk of harmful or even fatal overdosing. The genus *Ephedra*, which contains about 35 species,

7

8

9

0 0 0 0 ٥ 0 0 0 ATYKVKFVTPDGENE I ECPDDVYVLDAAEDAGLDLPYSCRAGSCSSCAGKMVAGNVDQSDGNFLDDDQ I SEGFVLTCVAYPKSDCV I QTHQEEA I A-Ephedra sinica I ¶ ATYKVKFVTPDGENEIECPDDVYVLDAAEDAGLDLPYSCRAGSCSSCAGKMVAGNVDQSDGNFLDDDQISEGFVLTCVAYPKSDCVIQTHQEEAUA-Ephedra sinica II \$ ATYKVKFVTPDGENEIECPDDVYVLDAAEDAGLDLPYSCRAGSCSSCAGKMVAGNVDQSDGNFLDDDQISEGFVLTCVAYPKSDCVIQTHQEEAWA-Ephedra foliata ATYKVKFVTPDGENEIECPDDVYVLDAAEDAGLDLPYSCRAGSCSSCAGKMVSGNVDQSDGNFLDDDQISEGFVLTCVAYPOSDCVIQTHQEEAIA-AMYKVKFUTPGGENEIECPDDVYVLDAAEDAGLDLPYSCRAGSCSSCAGKMVSGNVDQSDGNFLDDDQISEGFVLTCVAYPOSDCVIQTHQEEAIA-Ephedra intermedia £ Ephedra americana

5

6

2

Amino acids are represented by one-letter abbreviations.  $\P$ , *E. distachya* I, *E. equisetina* I, *E. viridis*, and *E. intermedia*; \$, *E. distachya* II and *E. equisetina* II;  $\pounds$ , minor Fd from *E. intermedia* 

3

Fig. 2. Comparison of the Sequences of [2Fe-2S] Ferredoxins from Ephedra Plants

|      |                          | 1  | 2  | 3   | 4                          | 5                                       | 6                         | 7                                       | 8   | 9  |
|------|--------------------------|--|--|---|----------------------------|---|---------------------------|---|---|--|
|      |                          | 0  | 0  | 0   | 0                          | 0                                       | 0                         | 0                                       | 0   | 0  |
| (1)  | <i>Ephedra sinica</i> I¶ | ATYKVK <b>F</b> VTPDG <mark>E</mark> I       | NE I ECPDDVYVL                               | .DAAEDAGLDL                               | PYSCRAGSCS                 | SCAGK <mark>M</mark> VA(                | NVDQSDG <mark>N</mark>    | FLDDDQ I <mark>s</mark> e               | EGFVLTCVAYP <mark>K</mark> SI               | DCVIOTHQEEAIA-   |
| (2)  | Ephedra sinica 🏾 \$      | ATYKVK <mark>F</mark> VTPDG <mark>E</mark> I | N <mark>e i e</mark> cpddvy <mark>v</mark> l | .D <mark>AAE</mark> DA <mark>GL</mark> DL | PYSCRAGSCS                 | SCAGK <mark>M</mark> VA(                | NVDQSDG <mark>N</mark> I  | FLDDDQ I <mark>s</mark> e               | EGFVLTCVAYP <mark>K</mark> SI               | DCVIQTHQEEALA-   |
| (3)  | Ephedra foliata          | ATYKVK <mark>F</mark> VTPDG <mark>E</mark> N | N <mark>e i e</mark> cpddvy <mark>v</mark> l | .D <mark>AAE</mark> DA <mark>GL</mark> DL | PYSCRAGSCS                 | SCAGK <mark>M</mark> VA(                | NVDQSDG <mark>N</mark> I  | FLDDDQ I <mark>s</mark> e               | EGFVLTCVAYP <mark>K</mark> SI               | DCVIQTHQEEAVA-   |
| (4)  | Ephedra intermedia £     | ATYKVK <mark>F</mark> VTPDG <mark>E</mark> I | N <mark>e i e</mark> cpddvy <mark>v</mark> l | .D <mark>AAEDA</mark> GLDL                | PYSCRAGSCS                 | SCAGK <mark>MVS</mark> O                | NVDQSDG <mark>N</mark> F  | FLDDDQI <mark>s</mark> e                | EGFVLTCVAYPQSI                              | DCVIOTHQEEAIA-   |
| (5)  | Ephedra americana        | AMYKVKFLTPEGEN                               | N <mark>e i e</mark> cpddvy <mark>v</mark> l | _D <mark>AAEDA</mark> GLDL                | PYSCRAGSCS                 | SCAGK <mark>M</mark> V <mark>S</mark> ( | NVDQSDG <mark>N</mark> I  | FLDDDQI <mark>s</mark> e                | EGFVLTCVAYPQS                               | DCVIQTHQEEAIA-   |
| (6)  | Atropa belladonna        | ATYKVKLVTPDGP                                | /EFDCPDDVYIL                                 | DQAEEEGH                                  | PYSCRAGSCS                 | SCAGKVSAG                               | TVDQSDG <mark>N</mark>    |   | GFVLTCVAYPQS                                | DVTIETHKEEELTG   |
| (7)  | Hyoscyamus niger         | ATYKVKLVTPDGP                                | /EF <mark>I</mark> CPDDVYIL                  | .DQAEEEGH <mark>E</mark> L                | PYSCRAGSCS                 | SCAGKV <mark>S</mark> AG                | atvdqsdg <mark>n</mark> i |   | DGFVLTCVAYPQSI                              | DVTIETHKEE <mark>D</mark> LTG                              |
| (8)  | Solanum nigrum           | ATYKVKLVTPDGP                                | IEFDCPDDVYIL                                 | .DQAEEEGH <mark>E</mark> L                | PYSCRAGSCS                 | SCAGKV <mark>T</mark> AG                | itvdqsdg <mark>n</mark> i |   | OGFVLTCVAYP <mark>K</mark> SI               | OVTIETHKEE <mark>D</mark> LTG                              |
| (9)  | Solanum lyratum          | ATYKVKL                                      | FICPDDVYIL                                   | .D <mark>S</mark> AEE <mark>N</mark> GHDL | .PYSCRAG <mark>A</mark> CS | SCAGK <mark>IT</mark> A(                | NVDQSDNSI                 | -LDDDQ <mark>V</mark> AE                | GFVLTCVAYP <mark>K</mark> S                 | VTIETHKE <mark>DD</mark> LVG                               |
| (10) | Solanum indicum          | ASYKVKLITPDGP                                | IEF <mark>N</mark> CPDDVYIL                  | .DRAEEEGHDL                               | .PYSCRAG <mark>A</mark> CS | SCAGKIVDO                               | SVDQSD <mark>N</mark> SI  | LDDDQIGO                                | GFVLTCVAYP <mark>K</mark> S                 | VTIETHKEEALVG  |
| (11) | Solanum abutiloides      | ATYKVKLVTPDGP                                | FECPDDYIL                                    | .D <mark>R</mark> AEEEGHDL                | PYSCRAGSCS                 | SCAGK <mark>IA</mark> A(                | SVDQSDGN                  |   | GFVLTCVAYPOS                                | <b>DVTIETHKEEELT</b> A                                     |
| (12) | Scopolia japonica        | ATYKVKLVTPDGPV                               | /EFDCPDDVYIL                                 | .DQAEEEGHEL                               | PYSCRAGSCS                 | SCAGKVSAG                               | TVDQSDG <mark>N</mark>    |   | GFVLTCVAYPQSI                               | DVIIETHKEEELTG   |
| (13) | Lycium chinense          | ATYKVKLVTPDGPV                               | /EFDCPDDVYIL                                 | .DQAEEEGH <mark>E</mark> L                | PYSCRAGSCS                 | SCAGKV <mark>S</mark> AG                | ATVDQSDG <mark>N</mark>   |   | GFVLTCVAYPQS                                | DVTIETHKEE <mark>A</mark> LTG                              |
| (14) | Capsicum annuum          | ASYKVKLITPDGP                                | IEFDCPDDVYIL                                 | DQAEEAGHDL                                | PYSCRAGSCS                 | SCAGK <mark>IAG</mark> (                | AVDQTDGN                  | -LDDDQ <mark>LE</mark> E                | GWVLTCVAYPQS                                | DVTIETHKE <mark>A</mark> ELVG                              |
| (15) | Nicotiana tabacum        | ASYKVKLITPEGA                                | FDCPDDVYIL                                   | .DQAEE <mark>M</mark> GHDL                | PYSCRAGSCS                 | SCAGKVTAG                               | NVDQSDGN                  | FLDDDQMA                                | GFVLTCVAYPQS                                | OVTIETHKEEELTA   |
| (16) | Physalis alkekengi*      | ATYKVKLITPDGP                                | / <mark>v</mark> fdcpd <mark>ne</mark> yil   | .D <mark>A</mark> AEE <mark>Q</mark> GHDL | PYSCRAGSCS                 | SCAGKV <mark>T</mark> AG                | TVDQSDG <mark>N</mark> I  | FLDDDQ <mark>V</mark> a <mark>d</mark>  | GFVLTCVAYPQSI                               | DVTIETHKEEELT <mark>A</mark>                               |
| (17) | Datura stramonium†       | ATYKVKLVTPDGP                                | /EFICPDDVYIL                                 | .DQAEEEGHDL                               | PYSCRAGSCS                 | SCAGKV <mark>T</mark> AG                | TVDQSDG <mark>N</mark>    | /LDDDQMA                                | GFVLTCVAYPQS                                | OVTIETHKEEELTG   |
| (18) | Datura mete/‡            | ATYKVKLVTPDGP                                | /EFDCPDDVY1L                                 | DRAEEEGHDL                                | PYSCRAGSCS                 | SCAGKV <mark>T</mark> AG                | atvdqsdg <mark>n</mark> ` | /LDDDQ <mark>M</mark> AE                | EGFVLTCVAYPQSI                              | OVTIETHKEEELTG   |
| (19) | Datura arborea           | ATYKVKLVTPDGPV                               | /EFDCPDDVYIL                                 | .DQAEEEGHEL                               | PYSCRAGSCS                 | SCAGKV <mark>T</mark> AG                | TVDQSDG <mark>N</mark>    |   | GFVLTCVAYPQS                                | OVTIETHKEEELTG   |
| (20) | Pueraria lobata          | ASYKVKLTPEGE                                 | /EFECPDDVYIL                                 | . QAEEEG <mark>VE</mark> L                | .PFSCKAGSCS                | VCTGKVVS                                | STVD <mark>NS</mark> EAA  | LEDEQIE/                                | GFTLTCVALPTS                                | VVIETHKENDLTG  |
| (21) | Leucaena glauca          | -AFKVKLLTPDGP                                | KEF <mark>E</mark> CPDDVYIL                  | DQAEELGIDL                                | PYSCRAGSCS                 | SCAGKEVE                                | DLDQSDQSI                 | ELDDEQIE                                | GWVLTCAAYPRSI                               | DV <mark>V</mark> IETHKEEELTG                              |
| (22) | Medicago sativa          | <b>ASYKVKLVTPEGT</b>                         | EF <mark>E</mark> CPDDVYIL                   | DAEEEGIVL                                 | PYSCRAGSCS                 | SCAGKVAA                                | EVNOSDGSI                 | FLDDDQIEE                               | EGWVLTCVAYAKSI                              | DVTIETHKEEELTA   |
| (23) | Pisum sativum            | ASYKVKLVTPDGT                                | DEFECPSDVYIL                                 | .D <mark>H</mark> AEE <mark>V</mark> GIDL | PYSCRAGSCS                 | SCAGKVVG                                | EVDQSDGSI                 |   | GFVLTCVAYPTSI                               | DV <mark>V</mark> IETHKEE <mark>D</mark> LT <mark>A</mark> |
| (24) | Trifolium pratense       | ATYKVKL TPEGP                                | EFDCPDDVYIL                                  | .D <mark>H</mark> AEE <mark>V</mark> GIEL | PYSCRAGSCS                 | SCAGKVV <mark>N</mark> (                | NVNQEDGSI                 | ELDD <mark>E</mark> QI <mark>E</mark> @ | GWVLTCVAFPTSI                               | OVTIETHKEEELTA   |
| (25) | Brassica napus           | ATYKVKFITPEGEG                               | DEVECDDDVYVL                                 | .D <mark>AAEEA</mark> GIDL                | PYSCRAGSCS                 | SCAGKVV <mark>S</mark> (                | FVDQSDESI                 | LDDDQIAE                                | GFVLTCAAYPTSI                               | OVTIETHKEEEL <mark>V</mark> -                              |
| (26) | Panax ginseng            | ATYKVKLLTPEGE                                | EFECPDDVYIL                                  | DQAEE <mark>IG</mark> IDL                 | PYSCRAGSCS                 | SCAGKVV <mark>S</mark> (                | K <mark>vdqsdq</mark> si  | EL <mark>qddq</mark> md <i>i</i>        | AGYVLTCHAYAQSI                              | DV <mark>V</mark> IETHKEEE <mark>V</mark> −−               |
| (27) | Petroselinum sativum     | ATYNVKLITPDGE                                | /EF <mark>KCD</mark> DDVYVL                  | DQAEEEGID1                                | PYSCRAGSCS                 | SCAGKVV <mark>S</mark> (                | STDQSD <mark>q</mark> SI  | ELDDEQMD/                               | GYVLTCHAYPTSI                               | DV <mark>V</mark> IETHKEEE <mark>IV-</mark>                |
| (28) | Phytolacca americana     | ATYKVILVTPSGT                                | AT I DCPDDTYVL                               | DAAEEAGLDL                                | PYSCRAGSCS                 | SCTGKVTAG                               | ATVDQEDQSI                | LDDDQIE/                                | GFVLTCVAFPKG                                | OVTIETHKEEDIV-   |
| (29) | Phytolacca esculenta     | ATYKV <mark>T</mark> LVTP <mark>S</mark> GT0 | AT I DCPDDTYVL                               | .D <mark>AAEEAGL</mark> DL                | PYSCRAGSCS                 | SCTGKVTAG                               | ATVDQ <mark>e</mark> dqsi | FLDDDQ1 <mark>E</mark> /                | GFVLTCVAYPKG                                | DVTIETHKEE <mark>DIA-</mark>                               |
| (30) | Spinacia oleracea        | AAYKVTLVTPTGN                                | /EFCCPDDVYIL                                 | .D <mark>AAEEEG</mark> IDL                | PYSCRAGSCS                 | SCAGKLKT                                | SLNQDDQSI                 | ELDDDQI <mark>D</mark> E                | GWVLTCAAYP <mark>V</mark> SI                | OVTIETHKEEELTA   |
| (31) | Sambucus nigra           | ASYKVKLITPDGP                                | EFECPDDVYIL                                  | <b>EHAEELGIDI</b>                         | PYSCRAGSCS                 | SCAGKLVA                                | SVDQSDQSI                 | FLDDEQIE                                | EG <mark>WVLTCVAYP</mark> KSI               | DVTIETHKEEELT <mark>A</mark>                               |
| (32) | Arctium lappa            | ATYKVILITPEGKO                               | D <mark>EFEV</mark> PDDVYIL                  | DHAAEEVGDL                                | PYSCRAGSCS                 | SCAGKVTAG                               | S <mark>S</mark> VDQSDGSI |   | G <mark>WVLTCVAYP</mark> TSI                | DVTIETHKEEELT <mark>A</mark>                               |
| (33) | Colocasia esculenta      | ATYKVKLVTP <mark>S</mark> GQ0                | DEFQCPDDVYIL                                 | DQAEE <mark>V</mark> GIDL                 | PYSCRAGSCS                 | SCAGKV <mark>KV</mark> (                | DVDQSDGSI                 | ELDDEQIGE                               | G <mark>WVLTCVAYP</mark> VSI                | D <mark>G</mark> TIETHKEEELT <mark>A</mark>                |
| (34) | Triticum aestivum        | ATYKVKLVTP <mark>E</mark> G <mark>E</mark> V | /E <mark>LEV</mark> PDDVYIL                  | .DQAEEEG <mark>I</mark> DL                | PYSCRAGSCS                 | SCAGK <mark>L</mark> VS                 | EIDQSDQSI                 | LDDDQME/                                | G <mark>WVLTCH</mark> AYP <mark>K</mark> SI | D <mark>IVIETHKEEELT</mark> A                              |
| (35) | Hordeum vulgare          | ATYKVKLVTP <mark>E</mark> G <mark>E</mark> V | /E <mark>LEV</mark> PDDVYIL                  | .DQAEEEG <mark>I</mark> DL                | PYSCRAGSCS                 | SCAGK <mark>LV</mark> S                 | EIDQSDQSI                 | -LDDDQ <mark>me</mark> e                | G <mark>WVLTC</mark> AAYP <mark>K</mark> SI | DVVIETHKEEELT <mark>A</mark>                               |
| (36) | Oryza sativa             | ATYNVKL TPDG                                 | /ELQVPDDVYIL                                 | DQAEEEGIDL                                | PYSCRAGSCS                 | SCAGKVVS                                | EIDQSDQSI                 | LDDDQVA                                 | GWVLTC <mark>H</mark> AYPKSI                | DVVIETHKEDDLI-   |
| (37) | Equisetum telmateia      | -AYKTVLKTPSGEF                               | TLDVPEGTTIL                                  | DAAEEAGYDL                                | PFSCRAGACS                 | SCLGKVVS                                | SVDQSEGSI                 |   | <b>GFVLTCIAIPES</b>                         | DLVIETHKEEELF-   |
| (38) | Equisetum arvense        | -AYKTVLKTPSGEFTI                             | DVPEGTTILD/                                  | AEEAGYDLPF                                | SCRAGACSSC                 | GKVVSGS                                 | DESEGSFLI                 | DDCQMEEGF                               | VLTCIAIPESD                                 | IETHKEEELF-  |
| (39) | Gleichenia japonica      | AIFKVK FLTPDGERTI                            | IEVPDDKFILD/                                 | GEEAGLDLPY                                | SCRAGACSSC                 | TGKLLDGR                                | /DOS <mark>EC</mark> SFLI | DDDQMAEGF                               | VLTCVAYPAGDI                                | TIETHAEEKL   |
|      |                          |  |  |   |                            |   |                           |   |   |  |

### Fig. 3. Comparison of the Sequences of [2Fe-2S] Ferredoxins from Higher Plants

Amino acids are represented by one-letter abbreviations.  $\P$ , *E. distachya* I, *E. equisetina* I, *E. viridis*, and *E. intermedia*; \$, *E. distachya* II and *E. equisetina* II;  $\pounds$ , minor Fd from *E. intermedia*; *\**, *Physalis alkekengi* var. *francheti*; *†*, var. *stramonium* and var. *tatula*, and *D. quercifolia*, *‡*, *D. metel*, *D. innoxia*, and *D. fastuosa*. References for the sequences are: (6) and (7) in 11), (8) -(11) in 10), (12) and (13) in 9), (14) in 8), (15) in 7), (16) in 6), (17) in 2), (18) in 3), (19) in 4), (20) in 12), (21), (22), (25), (27)-(34), and (37)-(39) in 16), (26) in 13), (35) in 18), (36) in 17), and (23) and (24) listed in accession numbers M31713 and AY340639, respectively.

Table 1. Amino Acid Differences Between Ephedra Ferredoxins and Other Higher-plant Ferredoxins

|      |                      | (1) | (2) | (3) | (4) | (5) |      |                      | (1) | (2) | (3) | (4) | (5) |
|------|----------------------|-----|-----|-----|-----|-----|------|----------------------|-----|-----|-----|-----|-----|
| (1)  | Ephedra sinica I ¶   | 0   | 1   | 1   | 2   | 5   | (21) | Leucaena glauca      | 31  | 30  | 31  | 31  | 31  |
| (2)  | Ephedra sinica 🏽 \$  | 1   | 0   | 1   | 3   | 6   | (22) | Medicago sativa      | 28  | 27  | 28  | 30  | 30  |
| (3)  | Ephedra foliata      | 1   | 1   | 0   | 3   | 6   | (23) | Pisum sativum        | 26  | 25  | 26  | 26  | 28  |
| (4)  | Ephedra intermedia £ | 2   | 3   | 3   | 0   | 3   | (24) | Trifolium pratense   | 32  | 31  | 32  | 32  | 32  |
| (5)  | Ephedra americana    | 5   | 6   | 6   | 3   | 0   | (25) | Brassica napus       | 22  | 21  | 22  | 21  | 21  |
| (6)  | Atropa belladonna    | 26  | 25  | 26  | 26  | 29  | (26) | Panax ginseng        | 29  | 29  | 28  | 27  | 26  |
| (7)  | Hyoscyamus niger     | 25  | 24  | 25  | 25  | 28  | (27) | Petroselinum sativum | 30  | 31  | 31  | 29  | 31  |
| (8)  | Solanum nigrum       | 25  | 24  | 25  | 27  | 30  | (28) | Phytolacca americana | 26  | 27  | 27  | 28  | 30  |
| (9)  | Solanum lyratum      | 29  | 28  | 29  | 31  | 31  | (29) | Phytolacca esculenta | 24  | 25  | 25  | 26  | 28  |
| (10) | Solanum indicum      | 28  | 27  | 28  | 29  | 30  | (30) | Spinacia oleracea    | 33  | 32  | 33  | 33  | 34  |
| (11) | Solanum abutiloides  | 24  | 23  | 24  | 24  | 27  | (31) | Sambucus nigra       | 28  | 27  | 28  | 30  | 31  |
| (12) | Scopolia japonica    | 26  | 25  | 26  | 26  | 29  | (32) | Arctium lappa        | 32  | 31  | 32  | 33  | 33  |
| (13) | Lycium chinense      | 24  | 23  | 24  | 24  | 27  | (33) | Colocasia esculenta  | 28  | 27  | 28  | 28  | 30  |
| (14) | Capsicum annuum      | 29  | 28  | 29  | 28  | 29  | (34) | Triticum aestivum    | 28  | 27  | 28  | 28  | 29  |
| (15) | Nicotiana tabacum    | 27  | 26  | 27  | 27  | 26  | (35) | Hordeum vulgare      | 27  | 26  | 27  | 27  | 28  |
| (16) | Physalis alkekengi*  | 28  | 27  | 28  | 28  | 30  | (36) | Oryza sativa         | 30  | 29  | 30  | 30  | 32  |
| (17) | Datura stramonium†   | 26  | 25  | 26  | 26  | 29  | (37) | Equisetum telmateia  | 39  | 38  | 39  | 38  | 38  |
| (18) | Datura mete/‡        | 25  | 24  | 25  | 25  | 28  | (38) | Equisetum arvense    | 40  | 39  | 40  | 39  | 39  |
| (19) | Datura arborea       | 26  | 25  | 26  | 26  | 29  | (39) | Gleichenia japonica  | 31  | 30  | 31  | 31  | 31  |
| (20) | Pueraria lobata      | 40  | 39  | 40  | 39  | 38  |      |                      |     |     |     |     |     |

See legend to Fig. 3. (1) belongs to Ephedraceae (order: Ephedrales, class: Gnetopsida, subphylum: Gymnospermae, phylum: Spermatophyta) (6)-(19) to Solanaceae; (20)-(24) to Leguminosae; (25) to Cruciferae; (26) to Araliaceae; (27) to Umbelliferae; (28) and (29) to Phytolaccaceae; (30) to Chenopodiaceae; (31) to Caprifoliaceae; (32) to Compositae; (33) to Araceae; (34)~(36) to Gramineae; (37) and (38) to Equisetaceae (order: Equisetales, class: Articulatae, phylum: Pteridophyta); and (39) to Gleicheniaceae (order: Filicales, class: Filicinae, phylum: Pteridophyta).

is represented by E. sinica.<sup>19)</sup>

Many primary structures have been reported for chloroplast [2Fe-2S] Fds.<sup>2-13, 16-18)</sup> The number of amino acid differences is 14-40 for different families and 0 to 4 for the same genus, except for the genus Solanum.<sup>10)</sup> In our recent study, 2 to 19 amino acid differences were observed among different genera of Solanaceae; Datura, Physalis, Nicotiana, Capsicum, Scopolia, and Lycium. Table 1 shows amino acid differences in Ephedra plant-Fds compared to other higher-plant Fds that have been determined so far. These Ephedra plant-Fds exhibited 21-34 differences in their amino acid sequences compared to those of Angiosperms, except for P. lobata (38-40 differences). In contrast, 38-40 differences were observed compared to E. telmateia and arvense (horsetails), respectively. This suggests that *Ephedra* plants are remotely related taxonomically to horsetails. Note that only 21-24 differences were observed between *Ephedra* plant-Fds and those of several dicotyledonous plants, *B. napus* (Cruciferae) and some solanaceous plants. This does not necessarily indicate a close taxonomic relation between *Ephedra* plants and these dicotyledonous plants. As described by Matsubara and Hase,<sup>16)</sup> it may be difficult to deduce the relation at the family or order level based only on Fds. Nevertheless, it is interesting that *Ephedra* plant-Fds showed the lowest similarity to *Equisetum*-Fd (Equisetales) among those of higher plants, despite their morphological similarity. In practice, the genus *Ephedra* (phylum Spermatophyta) is thought to be remotely related to the genus *Equisetum* (phylum Pteridophyta).

Figure 4 shows a phylogenetic tree based on the Fd sequences of higher plants.<sup>14)</sup> Fourteen solanaceous



**Fig. 4.** Phylogenetic Tree Based on the Amino Acid Sequences of Ferredoxins from Higher Plants The phylogenetic tree was constructed using the UPGMA method of Nei (1987) (GENETYX software).<sup>14)</sup> Genetic distances are represented by the proportion of amino acid differences between each taxon (1.0=100%).

plants form a cluster that is distinctly separated from other angiospermous plants, ferns, and horsetails by appreciably long branch lengths, which increase in that order. In the solanaceous cluster, five genera, Atropa, Hyoscyamus, Scopolia, Lycium, and Datura, are separated from each other by short branch lengths, which suggests a close taxonomic relationship among them. On the other hand, Ephedra plants (subphylum Gymnospermae) form a small cluster with short branch lengths. This cluster forms a greater cluster together with B. napus and two Phytolacca plants which belong in a different subphylum (Angiospermae), with a considerably long branch length. Furthermore, this cluster forms a greater cluster with other plants of Angiospermae except for *P. lobata*, which suggests that the correlation between the Fd structures and the taxonomic position of plant taxa is not reasonable. This can be partially accounted for by the rapid evolution of Fds. The number of mutations seems to have been saturated in a relatively short period for a small protein, and differences in the numbers of amino acids in Fds of remotely related plants do not reflect real phylogenic distances.<sup>16)</sup> Nevertheless, it is interesting that differences in the numbers of amino acids in Fds reflect the most remote relation between Ephedra plants and horsetails.

In conclusion, Ephedra plant-Fds possess unique amino acid sequences that are distinct from those of other Fds based on Asn-14, Asp-30, Met-51, Cys-85, Gln-88, Gln-91, and the deletion of one amino acid residue at the C-terminus. While E. sinica, E. distachya, E. equisetina, E. viridis, and E. intermedia have identical or very similar Fds, the Fd from E. americana was somewhat different from those of the other Ephedra plants. These results suggest that E. americana, which does not contain ephedrine, is somewhat distantly related to the other ephedrinecontaining Ephedra plants, although the others are very closely related to each other. A comparison of the amino acid sequence of Ephedra plant-Fds to those of other higher plants indicated that Ephedra plants (class Gnetopsida, phylum Spermatophyta) and horsetails (class Articulatae, phylum Pteridophyta) are remotely related. For further discussion, we would need additional information regarding the amino acid sequences of Fds from these two classes.

Acknowledgements The author is grateful to Dr. Tooru Akita, Head of the Nippon Shinyaku Institute for Botanical Research (Kyoto, Japan), for supplying the fresh leaves of several *Ephedra* plants, and for participating in helpful discussions. Thanks are also due to Misses Akiko Ishii, Masumi Yamashita, and Yuko Murata for their technical assistance. This work was supported in part by a Grant-in-Aid for High Technology Research from the Ministry of Education, Culture, Sports, Science, and Technology, Japan.

## References

- Palmer G., in "Iron-Sulfur Proteins II," ed. by Lovenberg W., Academic Press, New York, 1973, pp. 285-325.
- Mino Y., Usami, H., Inoue, S., Ikeda, K., Ota, N., *Phytochemistry*, **33**, 601-605 (1993).
- 3) Mino Y., Phytochemistry, 35, 385-387 (1994).
- 4) Mino Y., Phytochemistry, 37, 429-431 (1994).
- 5) Mino Y., Chem. Pharm. Bull., **43**, 1186-1189 (1995).
- Mino Y., Yasuda K., *Phytochemistry*, **49**, 1631-1636 (1998).
- Mino Y., Iwao M., Biol. Pharm. Bull., 22, 96-99 (1999).
- 8) Mino Y., Iwao M., Natural Med., 53, 37-41 (1999).
- 9) Mino Y., Biol. Pharm. Bull., 25, 1367-1369 (2002).
- 10) Mino Y., Hazama T., Machida Y., *Phytochemistry*, 62, 657-662 (2003).
- Mino Y., Yukita M., Hiratsuka N., Wariishi H., Biol. Pharm. Bull., 28, 1535-1538 (2005).
- 12) Mino Y., Machida Y., Wariishi H., *Natural Med.*, 59, 181-185 (2005).
- 13) Mino Y., Biol. Pharm. Bull., 29, 1771-1774 (2006).
- Nei M., in "Molecular Evolutionary Genetics," ed. by Nei M., Columbia University Press, New York,

1987, pp. 287-326.

- 15) Buchanan B. B., Arnon D. I., *Methods Enzymol.*, 23, 413-440 (1971).
- Matsubara H., Hase T., in "Proteins and Nucleic Acids in Plant Systematics," ed. by Jansen U., Fairbrothers D.E., Springer-Verlag, Berlin, 1983, pp. 168-181.
- 17) Kamo M., Kotani N., Tsugita A., He Y. K., Nozu Y., *Protein Seq. Data Anal.*, 2, 289-293 (1989).
- 18) Takruri I. A. H., *Phytochemistry*, **30**, 415-418 (1991).
- 19) Tobe H., in "Asahi Encyclopedia, World of Flora,"ed. by Suzuki, M., Asahi Shimbun Company, Tokyo, 1994, pp.11-170-11-172.