# Amino Acid Sequences of Ferredoxins from *Scopolia japonica* and *Lycium chinense*: Their Similarities to That of *Datura arborea*[1)]

Yoshiki MINO

*Department of Environmental Analysis, Osaka University of Pharmaceutical Sciences; 4–20–1 Nasahara, Takatsuki, Osaka 569–1094, Japan.* Received March 18, 2002; accepted June 26, 2002

**The complete amino acid sequences of [2Fe–2S] ferredoxins from *Scopolia japonica* and *Lycium chinense* have been determined by automated Edman degradation of the entire Cm-proteins and of the peptides obtained by enzymatic digestions. These two ferredoxins exhibited only 2—7 differences in the amino acid sequence when compared to the *Datura*-ferredoxins (*D. stramonium*, *D. metel*, and *D. arborea*), and especially only 2 or 3 differences compared to *D. arborea*. On the contrary, 8—19 differences were observed among the other solanaceous ferredoxins. This suggests that *S. japonica* and *L. chinense* are closely related taxonomically to *Datura* plants, especially to *D. arborea*.**

**Key words** *Scopolia japonica*; *Lycium chinense*; Solanaceae; amino acid sequence; protein chemotaxonomy; ferredoxin

There are many examples of the use of protein characteristics in chemotaxonomy since Zuckerkandl and Pauling's seminal paper in 1965.[2)] The present study is one of a series designed to provide sequence information on solanaceous ferredoxins (Fds) with the aim of elucidating the effectiveness of 'protein chemotaxonomy,' molecular taxonomy based on the primary structures of common plant proteins. We have selected Fd, one of the iron-sulfur proteins which act as electron carriers in photosynthetic electron transport, for the study because Fds are widely distributed in plants, and because their small molecular size facilitates studies on their structure. I earlier reported the primary structures of Fds from seven *Datura* plants,[3—6)] *Physalis alkekengi* var. *francheti*,[7)] *Nicotiana tabacum*,[8)] and *Capsicum annuum*.[1)] In this report, I have isolated Fds from *Scopolia japonica* and *Lycium chinense*, important medicinal drugs in China and Japan, and determined their amino acid sequences. They were also compared with primary structures of Fds of higher plants.

## MATERIALS AND METHODS

**Materials** *Scopolia japonica* and *Lycium chinense* were cultivated in the herb garden at the author's university.

**Isolation of Ferredoxins** Each protein (1.5 or 6.5 mg) was purified from the fresh leaves (0.6 or 1.0 kg) of *S. japonica* or *L. chinense* as described previously.[3,7)]

**Sequence Determination** The amino acid sequences of the Fds were determined using a gas-phase protein sequencer by automated Edman degradation of Cm-Fds, and the peptides obtained by lysyl endopeptidase, trypsin, or endoproteinase Asp-N digestion. The C-terminal analysis was done with carboxypeptidase Y.

The detailed procedure and the other methods were described in previous reports.[3,7)]

**Construction of Phylogenetic Tree** The phylogenetic tree was constructed from the amino acid sequences (97 residues) of the higher plant Fds (24 species) using the unweighed pair-group method with arithmetical averages (UPGMA) method of Nei (GENETYX software, Software Development, Japan).[9)]

## RESULTS AND DISCUSSION

The absorption maxima in the UV–Vis spectrum of *L. chinense* (*Lc*)-Fd were at 275, 285 (sh), 330, 420, 465 nm with $A_{max}/A_{275\,nm}$ ratios of 0.87, 0.65, and 0.59, respectively, for the maxima. The *S. japonica* (*Sj*)-Fd also exhibited virtually the same spectrum as *Lc*-Fd. These spectra were characteristic of [2Fe–2S] Fds from other higher plants.[10,11)]

The sequencing strategy is summarized in Fig. 1. The analytical results for the amino acid compositions of both Cm-Fds and the peptides obtained by enzymic digestion were consistent with the final derived sequences. Automated Edman degradation of the *Lc*-Cm-Fd yielded the amino-terminal sequence up to the 53rd cycle, except for several slightly doubtful amino acid residues. The lysyl endopeptidase digestion gave three short peptides (L-1 (1—4), L-2 (5—6), and L-5 (92—97)) and two long peptides (L-3 (7—50) and L-4 (51—91)). These peptides were isolated by reversed-phase HPLC; their $t_R$s were 14.5 for L-1, 20.0 for L-5, 44.8 for L-3, and 46.4 min for L-4, while L-2 was missing. The Edman degradation of L-3-T-2, obtained by tryptic digestion of L-3 (7—50), confirmed the sequence of 41—50. L-4 and L-5 covered the sequences of 50—91 and 92—97 (C-terminal), respectively. Sequence analysis of L-4-D-6 (84—91), obtained by endoproteinase Asp-N digestion of L-4, confirmed the end part of L-4 (50—91). In solanaceous Fds tested so far, Lys-91 was insensitive to trypsin or lysyl endopeptidase cleavage probably due to the adjacent sequence, –Lys–Glu–Glu–Glu–.[8)] Only for *C. annuum* Fd was Lys-91 sensitive to both enzymes because of the adjacent sequence, –Lys–Glu–Ala–Glu–.[1)] In the case of *Lc*-Fd, Lys-91 in the sequence, –Lys–Glu–Glu–Ala–, was sensitive to lysyl endopeptidase, but insensitive to trypsin cleavage. These results suggest that the resistances of –Lys–Glu–Glu–Glu–, –Lys–Glu–Glu–Ala–, and –Lys–Glu–Ala–Glu– against both enzymes decrease in that order. The N-terminal sequence was confirmed by the isolation of L-1 (Ala–Thr–Tyr–Lys). Carboxypeptidase Y digestion of the Cm-Fd for different periods of time suggested the C-terminal sequence to be –Leu–Thr–Gly–COOH. This result was in good agreement with the C-terminal sequence by Edman degradation of the peptide, L-5 (92—97). The sequence analysis for *Sj*-Fd was con-

e-mail: mino@gly.oups.ac.jp

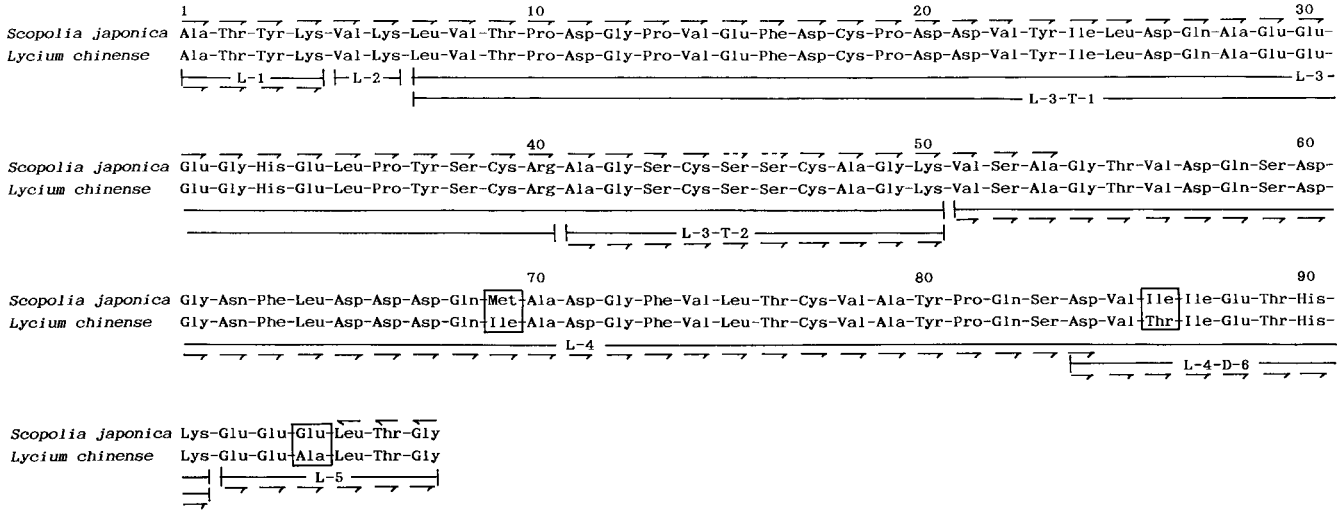


Fig. 1. Amino Acid Sequences of *Scopolia japonica* and *Lycium chinense* Ferredoxins

Arrows (⟶) and (⟵) represent residues determined by automated Edman degradation and carboxypeptidase Y digestion, respectively. Dashed arrows (⇢) indicate that the residue could not be unambiguously identified. L(1—5), T(1—2), and D(1—6) represent peptides obtained from lysyl endopeptidase, trypsin, and endoproteinase Asp–N digestions, respectively.

| | 1 2 3 4 5 6 7 8 9<br>0 0 0 0 0 0 0 0 0 |
|---|---|
| (1) *Scopolia japonica* | ATYKVKLVTPDGPVEFDCPDDVYILDQAEEEGHELPYSCRAGSCSSCAGKVSAGTVDQSDGNFLDDDQMADGFVLTCVAYPQSDVIIETHKEEELTG |
| (2) *Lycium chinense* | ATYKVKLVTPDGPVEFDCPDDVYILDQAEEEGHELPYSCRAGSCSSCAGKVSAGTVDQSDGNFLDDDQIADGFVLTCVAYPQSDVTIETHKEEALTG |
| (3) *Capsicum annuum* | ASYKVKLITPDGPIEFDCPDDVYILDQAEEAGHDLPYSCRAGSCSSCAGKIAGGAVDQTDGNFLDDDQLEEGWVLTCVAYPQSDVTIETHKEAELVG |
| (4) *Nicotiana tabacum* | ASYKVKLITPEGAVEFDCPDDVYILDQAEEMGHDLPYSCRAGSCSSCAGKVTAGNVDQSDGNFLDDDQMADGFVLTCVAYPQSDVTIETHKEEELTA |
| (5) *Physalis alkekengi** | ATYKVKLITPDGPVVFDCPDNEYILDAAEEQGHDLPYSCRAGSCSSCAGKVTAGTVDQSDGNFLDDDQVADGFVLTCVAYPQSDVTIETHKEEELTA |
| (6) *Datura stramonium*† | ATYKVKLVTPDGPVEFNCPDDVYILDQAEEEGHDLPYSCRAGSCSSCAGKVTAGTVDQSDGNYLDDDQMADGFVLTCVAYPQSDVTIETHKEEELTG |
| (7) *Datura metel*‡ | ATYKVKLVTPDGPVEFDCPDDVYILDRAEEEGHDLPYSCRAGSCSSCAGKVTAGTVDQSDGNYLDDDQMAEGFVLTCVAYPQSDVTIETHKEEELTG |
| (8) *Datura arborea* | ATYKVKLVTPDGPVEFDCPDDVYILDQAEEEGHELPYSCRAGSCSSCAGKVTAGTVDQSDGNFLDDDQMADGFVLTCVAYPQSDVTIETHKEEELTG |

Fig. 2. Comparison of Sequences of [2Fe–2S] Ferredoxins from Solanaceous Plants

Amino acids are represented by a one letter abbreviation. ∗: *Physalis alkekengi* var. *francheti*, †: var. *stramonium* and var. *tatula*, and *D. quercifolia*, ‡: *D. metel*, *D. innoxia*, and *D. fastuosa*. References for the sequences are: (3) listed in 1), (4) in 8), (5) in 7), (6) in 3), (7) in 4), (8) in 5).

ducted in the same manner except that both peptides, L-5 (92—97) and L-4,5 (51—97), were obtained because of the considerable resistance of –Lys–Glu–Glu–Glu– to lysyl endopeptidase.

These results led to the complete amino acid sequences shown in Fig. 1 for *Sj*- and *Lc*-Fds. A comparison of their amino acid sequences with those of solanaceous Fds is shown in Fig. 2. As compared with the other solanaceous Fds (*Datura*-, *Pa*-, *Nt*- and *Ca*-Fds), differences were observed only at Ser-52 (in both Fds), Ile-69 (in *Lc*-Fd), and Ala-94 (in *Lc*-Fd), in addition to Thr-2, Val-8, Asp-11, Pro-13, Val-14, Glu-15, Asp-17, Asp-21, Val-22, Gln-27, Glu-31, Glu-34, Val-51, Ala-53, Thr-55, Ser-59, Phe-63, Ala-70, Asp-71, Phe-73, Glu-93, Thr-96 and Gly-97 in which the amino acid differs among the other solanaceous Fds. The residues Ser-52 and Ala-94, in the primary structure of *Sj*- and/or *Lc*-Fd, are rare in sequences of the other higher plants.[1)] These residues are probably characteristic of *Sj*- and/or *Lc*-Fd. Also, the residues His-33, Asn-62, and Gln-82 may be characteristic of the solanaceous Fds. In Fds, the sequence 35—50, including the sequences –C39–C44–C47– which participate in chelating to the iron atoms, the sequence 74—77, containing the last cysteine ligand (–C77–) for the iron atom, and the later region, 83—93, are almost perfectly conserved. This was also true in the case of *Sj*- and *Lc*-Fd.

Many primary structures have been reported for chloroplast [2Fe–2S] Fds.[1,3—8,12—14)] The numbers of amino acid differences are 14—40 for different family and zero to four for the same genus.[1,3—8,12—14)] The sequences of several solanaceous Fds have recently been determined.[1,3—8)] Table 1 shows the matrix of amino acid differences for the higher plant Fds examined so far. Eight to nineteen amino acid differences were observed among *C. annuum*, *N. tabacum*, *P. alkekengi* var. *francheti*, and *Datura* plants. Of interest is the fact that, in the present study, only two to seven amino acid differences were observed among *S. japonica*, *L. chinense*, and *Datura* plants; *Sj*-Fd showed 5, 6, 2, 11, 10, and 17 differences, and *Lc*-Fd showed 6, 7, 3, 11, 11, and 17 differences respectively, when compared to the Fds of *D. stramonium*, *D. metel*, *D. arborea*, *P. alkekengi* var. *francheti*, *N. tabacum*, and *C. annuum*. Two to seven differences among *Sj*-, *Lc*-, and *Datura*-Fds were significantly smaller than those (8—19) among the solanaceous Fds except for *Sj*- and *Lc*-Fds. Figure 3 shows the phylogenetic tree based on Fd sequences of the higher plants.[9)] Eight solanaceous plants form one cluster distinctly separated from the other angiospermous plants, fern, and horsetails by appreciable long branch lengths. The solanaceous cluster could be divided into three groups: ① *Datura* group (genetic distance=0.02—0.06) including three *Datura* plants, *S. japonica*, and *L. chinense*, ② *P. alkekengi* var. *francheti* and *N. tabacum* group (0.11), and ③ *C. annuum* group. These results suggest that *S. japonica* and *L. chinense* are placed in a close taxonomic position to the *Datura* plants, especially *D. arborea*, although genera

Table 1. Amino Acid Difference Matrix of [2Fe–2S] Ferredoxin Sequences for Higher Plants[a)]

| | (1) | (2) | (3) | (4) | (5) | (6) | (7) | (8) |
|---|---|---|---|---|---|---|---|---|
| (1) *Scopolia japonica* | 0 | 3 | 17 | 10 | 11 | 5 | 6 | 2 |
| (2) *Lycium chinense* | 3 | 0 | 17 | 11 | 11 | 6 | 7 | 3 |
| (3) *Capsicum annuum* | 17 | 17 | 0 | 16 | 19 | 17 | 16 | 16 |
| (4) *Nicotiana tabacum* | 10 | 11 | 16 | 0 | 10 | 9 | 10 | 8 |
| (5) *Physalis alkekengi*[a)] | 11 | 11 | 19 | 10 | 0 | 10 | 10 | 9 |
| (6) *Datura stramonium* | 5 | 6 | 17 | 9 | 10 | 0 | 3 | 3 |
| (7) *Datura metel* | 6 | 7 | 16 | 10 | 10 | 3 | 0 | 4 |
| (8) *Datura arborea* | 2 | 3 | 16 | 8 | 9 | 3 | 4 | 0 |
| (9) *Brassica napus* | 25 | 24 | 25 | 22 | 24 | 24 | 23 | 24 |
| (10) *Leucaena glauca* | 24 | 24 | 23 | 26 | 28 | 24 | 24 | 24 |
| (11) *Medicago sativa* | 20 | 19 | 21 | 18 | 23 | 20 | 19 | 19 |
| (12) *Petroselinum sativum* | 25 | 27 | 29 | 26 | 29 | 25 | 26 | 25 |
| (13) *Phytolacca americana* | 28 | 26 | 30 | 28 | 26 | 27 | 26 | 26 |
| (14) *Phytolacca esculenta* | 27 | 25 | 30 | 27 | 25 | 26 | 25 | 25 |
| (15) *Spinacia oleracea* | 25 | 24 | 26 | 24 | 26 | 24 | 23 | 24 |
| (16) *Sambucus nigra* | 23 | 22 | 20 | 20 | 22 | 22 | 21 | 22 |
| (17) *Arctium lappa* | 21 | 22 | 24 | 17 | 21 | 19 | 19 | 19 |
| (18) *Colocasia esculenta* | 20 | 19 | 22 | 19 | 22 | 19 | 19 | 19 |
| (19) *Triticum aestivum* | 22 | 24 | 27 | 22 | 27 | 22 | 23 | 22 |
| (20) *Hordeum vulgare* | 21 | 23 | 25 | 21 | 26 | 21 | 21 | 21 |
| (21) *Oryza sativa* | 24 | 24 | 27 | 25 | 26 | 24 | 25 | 24 |
| (22) *Equisetum telmateia* | 37 | 36 | 38 | 36 | 36 | 38 | 36 | 37 |
| (23) *Equisetum arvense* | 38 | 40 | 39 | 37 | 37 | 39 | 37 | 38 |
| (24) *Gleichenia japonica* | 35 | 35 | 36 | 34 | 34 | 34 | 33 | 34 |

*a*) Data among the higher plants except for solanaceous plants are available in ref. 1 and so are omitted here. See legend to Fig. 2. (1)—(8) belong to Solanaceae; (9) to Cruciferae; (10) and (11) to Leguminosae; (12) to Umbelliferae; (13) and (14) to Phytolaccaceae; (15) to Chenopodiaceae; (16) to Caprifoliaceae; (17) to Compositae; (18) to Araceae; (19)—(21) to Gramineae; (22) and (23) to Equisetales; (24) to Filicales, respectively.

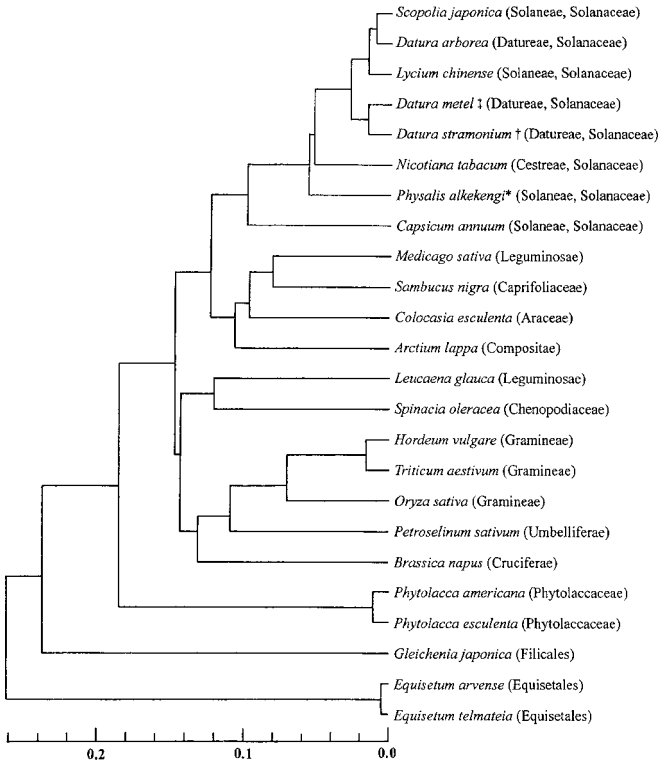


Fig. 3. Phylogenetic Tree Based on the Amino Acid Sequences of Ferredoxins from Higher Plants

The phylogenetic tree was constructed using the UPGMA method of Nei[9)] (GENETYX software). Genetic distances were represented by the proportion of amino acid differences between each *taxon* (1.0=100%).

*Scopolia* and *Lycium*, and genus *Datura* belong to different tribes, *Solaneae* and *Datureae*, respectively.

**Acknowledgements** The author is grateful to Emeritus Prof. Nagayo Ota, Prof. Kiyoshi Ikeda, and Dr. Seiji Inoue for helpful discussions. Thanks are also due to Mr. Shunji Kita for supplying the fresh leaves of *S. japonica* and *L. chinense*.

## REFERENCES AND NOTES

1) This paper is Part VIII of "Protein Chemotaxonomy of The Solanaceae," Part VII: Mino Y., Iwao M., *Nat. Med.*, **53**, 37—41 (1999).
2) Zuckerkandl E., Pauling L., *J. Theor. Biol.*, **8**, 357—366 (1965).
3) Mino Y., Usami H., Inoue S., Ikeda K., Ota N., *Phytochemistry*, **33**, 601—605 (1993).
4) Mino Y., *Phytochemistry*, **35**, 385—387 (1994).
5) Mino Y., *Phytochemistry*, **37**, 429—431 (1994).
6) Mino Y., *Chem. Pharm. Bull.*, **43**, 1186—1189 (1995).
7) Mino Y., Yasuda K., *Phytochemistry*, **49**, 1631—1636 (1998).
8) Mino Y., Iwao M., *Biol. Pharm. Bull.*, **22**, 96—99 (1998).
9) Nei M., "Molecular Evolutionary Genetics," Columbia University Press, New York, 1944, pp. 293—298.
10) Palmer G., "Iron-Sulfur Proteins II," ed. by Lovenberg W., Academic Press, New York, 1973, pp. 285—325.
11) Buchanan B. B., Arnon D. I., *Methods Enzymol.*, **23**, 413—440 (1971).
12) Matsubara H., Hase T., "Proteins and Nucleic Acid in Plant Systematics," ed. by Jansen U., Fairbrothers D. E., Springer, Berlin, 1983, pp. 168—181.
13) Takruri I. A. H., *Phytochemistry*, **30**, 415—418 (1991).
14) Kamo M., Kotani N., Tsugita A., He Y. K., Nozu Y., *Protein Sequences Data Anal.*, **2**, 289—293 (1989).