FEBS Letters 453 (1999) 318–326 FEBS 22112

The structural flexibility of the preferredoxin transit peptide

Hans L.J. Wienk[a,*], Michael Czisch[b], Ben de Kruijff[a]

[a] *Department Biochemistry of Membranes, Centre for Biomembranes and Lipid Enzymology, Institute of Biomembranes, Utrecht University, Padualaan 8, 3584 CH Utrecht, The Netherlands*

[b] *Department of NMR Spectroscopy, Bijvoet Center for Biomolecular Research, Utrecht University, Padualaan 8, 3584 CH Utrecht, The Netherlands*

Received 8 March 1999; received in revised form 20 April 1999

Abstract In order to obtain insight into the structural flexibility of chloroplast targeting sequences, the *Silene pratensis* preferredoxin transit peptide was studied by circular dichroism and nuclear magnetic resonance spectroscopy. In water, the peptide is unstructured, with a minor propensity towards helix formation from Val-9 to Ser-12 and from Gly-30 to Ser-40. In 50% (v/v) trifluoroethanol, structurally independent N- and C-terminal helices are stabilized. The N-terminal helix appears to be amphipathic, with hydrophobic and hydroxylated amino acids on opposite sides. The C-terminal helix comprises amino acids Met-29–Gly-50 and is destabilized at Gly-39. No ordered tertiary structure was observed. The results are discussed in terms of protein import into chloroplasts, in which the possible interactions between the transit peptide and lipids are emphasized.

Key words: Chloroplast protein import; Transit peptide; Conformational flexibility; Nuclear magnetic resonance; Trifluoroethanol

1. Introduction

Many proteins that are present in the chloroplast stroma are encoded on nuclear DNA. As they are synthesized on cytosolic ribosomes, those proteins must contain information to find the chloroplast and to translocate specifically across its envelope membranes. For this purpose, these proteins are produced as precursors with N-terminal extensions referred to as transit sequences [1]. It appears that these sequences contain the information that is necessary [2,3] and sufficient [4] to provide organelle-specific translocation. As different chloroplast-destined precursor proteins use the same protein import machinery, it can be expected that their transit sequences share common motifs. However, like most organelle-specific targeting sequences, chloroplast transit sequences show no amino acid homology [5]. Therefore, their information for organelle-specific recognition must be stored on a different level.

For different organellar targeting signals, general motifs can easily be identified. For instance, mitochondrial presequences all share an amphipathic helix, on one side positively charged and on the other side hydrophobic. This feature is readily recognized from the amino acid sequence because generally positively charged and hydrophobic residues are present at every third or fourth amino acid along the complete length of the mitochondrial presequence. The only features that characterize chloroplast transit sequences are an enrichment in hydroxylated amino acids and a positively charged C-terminus [5]. A general motif, however, is not easily pinpointed. This suggests that a common chloroplast recognition characteristic is formed in the precursor protein only after interaction with components of the chloroplast membrane. Many studies have been performed with the *Silene pratensis* precursor protein preferredoxin (preFd, for a review see [6]). Some of the findings support the hypothesis that structural motifs that are involved in the import process can be induced in the preFd transit peptide by chloroplast lipids.

The aim of this study was to get insight into the structure of the preFd transit sequence in different solvents via multidimensional nuclear magnetic resonance (NMR) spectroscopy. For this purpose, we purified the uniformly ^{15}N-labelled 47 residue transit peptide of preFd, followed by the amino acids Ala-48-Ser-49-Gly-50-Leu-51-Pro-52 (trFd), from *Escherichia coli* (Fig. 1). The results show that the peptide in aqueous solution is in a random coil conformation. Trifluoroethanol (TFE) is commonly used as solvent to investigate which parts of a peptide are most likely to undergo structural changes in a membraneous (hydrophobic) environment (e.g. [7]). It is shown that in the presence of 50% (v/v) TFE, the transit peptide forms helices from residues 2 to 14 and from 27 to 50.

2. Materials and methods

2.1. Materials

TrFd was purified to homogeneity from *E. coli*. (H. Wienk, manuscript in preparation). In short, a plasmid was constructed using the pMal-c2 vector (New England Biolabs), which, after overexpression in *E. coli*, gave rise to a fusion protein consisting (from N- to C-terminus) of maltose binding protein, a factor Xa cleavage sequence, the preFd transit peptide and the residues Ala-48-Ser-49-Gly-50-Leu-51-Pro-52. After harvesting, lysis and centrifugation, the fusion protein was purified from the supernatant on an amylose column and eluted with 10 mM maltose, according to the manufacturers' instructions. The fusion protein was cleaved with 0.5% (w/w) factor Xa for 1 h at 4°C. After ammoniumsulfate precipitation, trFd was purified by HPLC, essentially as described [8]. It was characterized by electrospray mass spectrometry (molecular weight = 5188 Da) and N-terminal sequencing. Also, it was shown to be able to interact with the chloroplast import machinery. ^{15}N-labelled trFd was obtained by growing the bacteria in SV medium containing 50 μg/ml ampicillin, 0.5% (w/w) glucose, 0.5 mg/ml thiamine and 0.5 g/l $^{15}NH_4Cl$. Chemically synthesized transit peptide was obtained as described [8]. TFA, TFE, DCl and NaOD were from Merck and ^2H_2O from Isotec. TFE-d_3 and $^{15}NH_4Cl$ were obtained from Cambridge Isotope Laboratories.

*Corresponding author. Fax: (31) (30) 2522478.
E-mail: h.l.j.wienk@chem.uu.nl

Abbreviations: CD, circular dichroism; HSQC, heteronuclear single quantum correlation spectroscopy; MGDG, monogalactosyl diacylglycerol; NOESY, nuclear Overhauser enhancement spectroscopy; preFd, precursor of the *Silene pratensis* protein ferredoxin; preSSU, precursor of the small subunit of ribulose-1,5-bisphosphate carboxylase/oxygenase from pea; TFE-d_3, [$^2H_3$]trifluoroethanol; TOCSY, total correlation spectroscopy; trFd, the 47 amino acid transit peptide of preFd followed by the amino acids Ala-48-Ser-49-Gly-50-Leu-51-Pro-52

PII: S0014-5793(99)00653-5

2.2. Circular dichroism (CD)

CD spectra were recorded as described by Horniak et al. [9] at room temperature on a Jasco-600 spectropolarimeter. Data were corrected by subtraction of spectra from peptide-free samples. Measurements were performed with samples containing 125 μM trFd in 10 mM KPi and 0.02% sodium azide at pH 3.5 with different TFE concentrations. Deconvolution was performed as described [10]. The S.D. was estimated to be 5–10%.

2.3. NMR measurements

NMR experiments were performed at the SON NMR Large-Scale facility in Utrecht on a 500 MHz Varian Inova spectrometer. A triple resonance probehead with a z-gradient accessory was used. 0.7 mM [^{15}N]-trFd either in 280 μl 18 mM KPi (pH 3.7, adjusted from pH 6 with DCl and NaOD, uncorrected readings), 0.02% (w/v) azide, 10% (v/v) $^{2}H_2O$ or in the same solution containing 50% (v/v) TFE-d_3 was measured in a 5 mm Shigemi tube. NMR experiments were performed at 20 or 5°C. Water suppression was achieved by gradient coherence selection of the ^{15}N pathway [11]. The ^{1}H B1 field strength was 35 kHz and the ^{15}N B1 was 6 kHz. Proton spectral widths were 4.4 kHz (indirect) or 7 kHz (direct), the ^{15}N spectral width was 1.125 kHz. Pulse field gradients consisting of 1 ms sine-shaped gradient pulses were applied, with a maximum power of 30 G/cm. In aqueous solution, the protons were calibrated on internal sodium 3-(trimethylsilyl)propionate. In water/TFE mixtures, the water signal was used for calibration. ^{15}N chemical shifts were referenced indirectly from the ^{1}H standards [12]. Protons were assigned according to Wüthrich [13].

Two dimensional sensitivity-enhanced gradient coherence-selected ^{15}N-^{1}H heteronuclear single quantum correlation spectroscopy (HSQC) experiments [14,15] were recorded with 2K×110 complex points. Three dimensional sensitivity-enhanced nuclear Overhauser enhancement spectroscopy (NOESY)-HSQC experiments [16] were recorded with 176×44×1K complex time points in t_1(H), t_2(N) and t_3(H^N), respectively, with STATES-TPPI [17] in t_1 and gradient coherence selection in t_2. A NOESY mixing time of 100 ms was applied. Three dimensional sensitivity-enhanced total correlation spectroscopy (TOCSY)-HSQC experiments [16] were recorded with 150×44×1K complex time points. A DIPSI [18] mixing sequence with a 50 ms length was performed with an 11 kHz field strength.

Data were processed using NMRPipe [19] and analyzed with our in house written program Regine [20]. In the ^{15}N dimension, a linear prediction was applied. The data were multiplied by a π/3-shifted sine square function in all dimensions prior to zero filling, Fourier transformation and baseline correction. Final two dimensional HSQC data matrices were 512×1K, and three dimensional (3D) data matrices were 512×128×1K points.

3. Results

3.1. CD measurements

It has been demonstrated for the chemically synthesized preFd transit peptide that it is mainly in a random coil conformation in aqueous solution at pH 8.0 [9,21]. To check the conformation of the peptide isolated from *E. coli* and to test the conformational behavior of the peptide at lower pH values, commonly used in NMR studies, CD experiments were performed with trFd in water at pH 3.5. Fig. 2 shows that at this pH, trFd is in a random coil conformation, characterized by the minimum at 200 nm. Upon titration of TFE, increasing amounts of helix are induced (minima at 206 and 222 nm), which level off at 40–50% (v/v) TFE. The helix induction by TFE is in accordance with the data obtained for the chemically synthesized transit peptide at pH 8.0 [21]. Deconvolution revealed that for trFd in aqueous solution at pH 3.5, the secondary structure was 6% helix, 17% β-strand, 3% turn and 74% random coil. These numbers indicate that in aqueous solution, the peptide is virtually structureless. In 50% (v/v) TFE at pH 3.5, the secondary structure was changed to 24% helix, 22% β-strand, 4% turn and 50% random coil.

processing site

10 20 30 40 ↓ 50

ASTLSTLSVSASLLPKQQPMVASSLPTNMGQALFGLKAGSRGRVTAMASGLP

Fig. 1. Amino acid sequence of the preFd transit peptide from *S. pratensis* followed by the amino acids Ala-48-Ser-49-Gly-50-Leu-51-Pro-52.

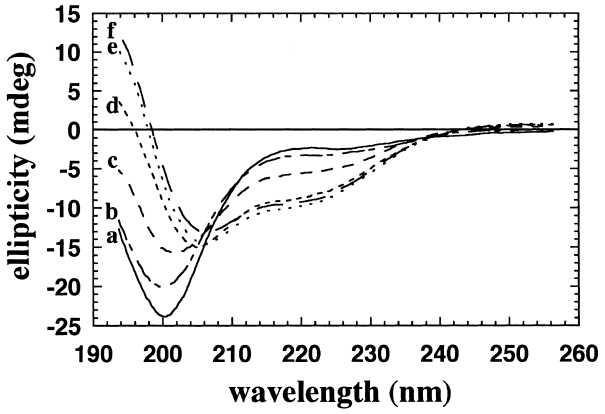

Fig. 2. CD spectra of 125 μM isolated trFd in 10 mM KPi, 0.02% (w/v) sodium azide at pH 3.5 with different percentages (v/v) TFE, recorded at 20°C. (a) 0% TFE, (b) 10% TFE, (c) 20% TFE, (d) 30% TFE, (e) 40% TFE, (f) 50% TFE.

3.2. NMR measurements

Initial NMR experiments were performed with ^{15}N-labelled trFd in aqueous solution at pH 3.5 and 20°C. A ^{15}N-^{1}H HSQC experiment pointed out that the backbone H^N signals are within a range of approximately 0.5 ppm, indicative of a random coil structure (data not shown). In NOESY experiments, no magnetization transfer between sequential H^Ns was detected which confirmed that in aqueous solution at 20°C, trFd is completely in a random coil conformation (data not shown).

In an attempt to stabilize secondary structure elements, experiments with [^{15}N]-trFd were performed at 5°C. For the structural assignment, 3D NOESY-HSQC and TOCSY-HSQC experiments were performed. In Fig. 3A, the ^{15}N-^{1}H HSQC projection of the 3D HSQC-NOESY experiment is shown. Again, the low H^N chemical shift dispersion suggests a random coil structure. Nevertheless, most of the H^N peaks are resolved. For each of the backbone H^N peaks in the ^{15}N-^{1}H HSQC projections of the 3D spectra, TOCSY and NOESY strips were isolated. Information about the spin systems could be acquired from the TOCSY strips. In combination with $d\alpha$N or $d\beta$N contacts and/or dNN cross-peaks from the NOESY strips, a complete sequential assignment could be made. The results are depicted in Fig. 3A (Gln and Asn side chain N-H^N contacts are not labelled). No Ala-1 H^N peak was found and the ^{15}N-H^N cross-peaks of Thr-3 and Ser-5 and of Val-9 and Val-21 are overlapping. Also, partial overlap is present for Ser-23 and Ser-49, for Lys-16 and Gln-17 and for Ala-32 and Ala-38. Because a rather short mixing time was used during the NOESY experiment, not all side chain protons could be assigned.

Prior to determination of secondary structure elements, peak volumes were estimated and classified into five different intensity categories. Sequential and medium range NOEs are displayed in Fig. 4A (in the $d\alpha$N row, the dNδ con-

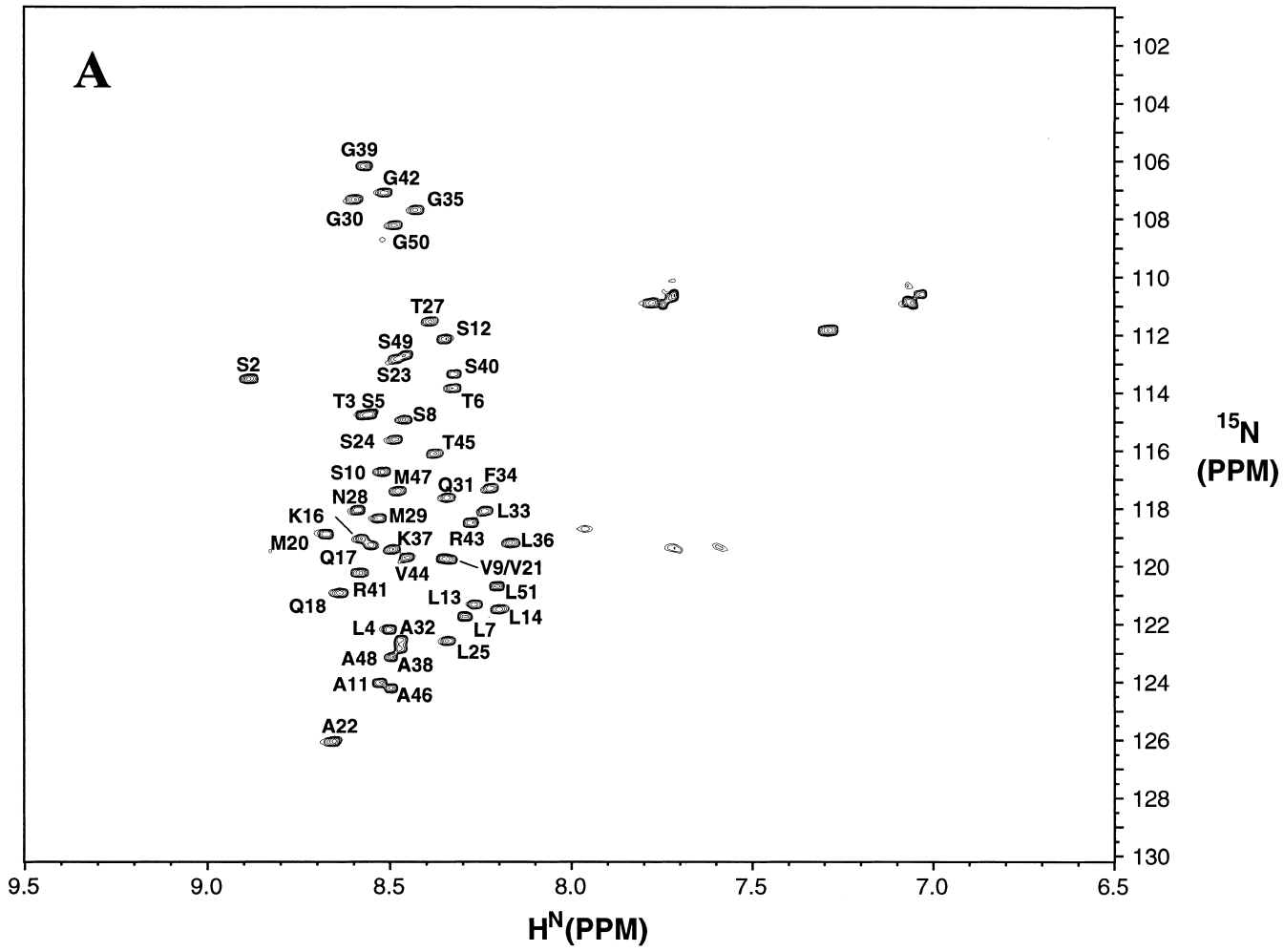

Fig. 3. HSQC projections of the 3D HSQC-NOESY spectra of ^{15}N-labelled trFd (A) in aqueous solution and (B) in 50% (v/v) TFE.

tacts for the prolines are included). The amount of connections in the $d\alpha$N, $d\beta$N and dNN rows gives an indication about the quality of the sequential assignment. The overall relatively strong $d\alpha$N and weak dNN cross-peaks indicate that the complete peptide is unstructured in aqueous solution [13]. The presence of $d\alpha$N(i, i+2) contacts between Lys-16 and Gln-18 and between Ser-23 and Leu-25 are suggesting the presence of initial structures in this region of the transit peptide [13]. The dNN(i, i+2) cross-peaks between residues 9 and 12 and residues 37 and 40 are indicative of a propensity towards a helical structure. The complete absence of $d\alpha$N(i, i+3) and $d\alpha$N(i, i+4) cross-peaks indicates that regular secondary structures present, if any, are not very stable. No long distance NOEs were observed, from which it can be concluded that no tertiary structure is present in trFd in aqueous solution.

The difference between the H^{α} proton chemical shifts in a protein and those obtained for individual amino acids in small random coil peptides correlates with secondary structures present in the protein [22]. A helical structure is characterized by an upfield change of the chemical shift of more than 0.1 ppm with respect to the random coil values for three successive amino acids, whereas a downfield shift is in agreement with a β-conformation. The results for trFd in water are depicted in Fig. 5A. The large deviations observed for residues 14, 18 and 25 are resulting from the fact that they precede proline residues. It has been shown that this causes deviations of about 0.29 ppm in this kind of data analysis [22]. It can be concluded that in aqueous solution, no stable secondary structure elements are present in trFd. There might be a propensity towards a helical structure from residue 10 to 12 and between residues 30 and 39.

From the combined results, it can be concluded that in aqueous solution at 5°C, trFd does not show regions with a stable secondary structure, but has a minor tendency towards helix formation from residues 9 to 12 and from 30 to 40.

To get insight into possible changes in the trFd structure upon entering a hydrophobic environment, as promoted for instance by the membrane lipids, we studied the conformation of the transit peptide in the presence of TFE which is a commonly used solvent for these purposes. In Fig. 3B, the ^{15}N-^{1}H HSQC spectrum of ^{15}N-labelled trFd in 50% (v/v) TFE is shown. As was shown for the transit peptide in aqueous solution, also in this spectrum, most of the N-H^N backbone atoms are resolved. 3D NOESY-HSQC and 3D TOCSY-HSQC spectra were recorded and analyzed. A complete sequential assignment was made, making use of dNN and $d\alpha$N or $d\beta$N contacts. The results are depicted in the HSQC in Fig. 3B. Again, the Ala-1 H^N was not observed. The strips for Ser-23 and Thr-27 show partial overlap.

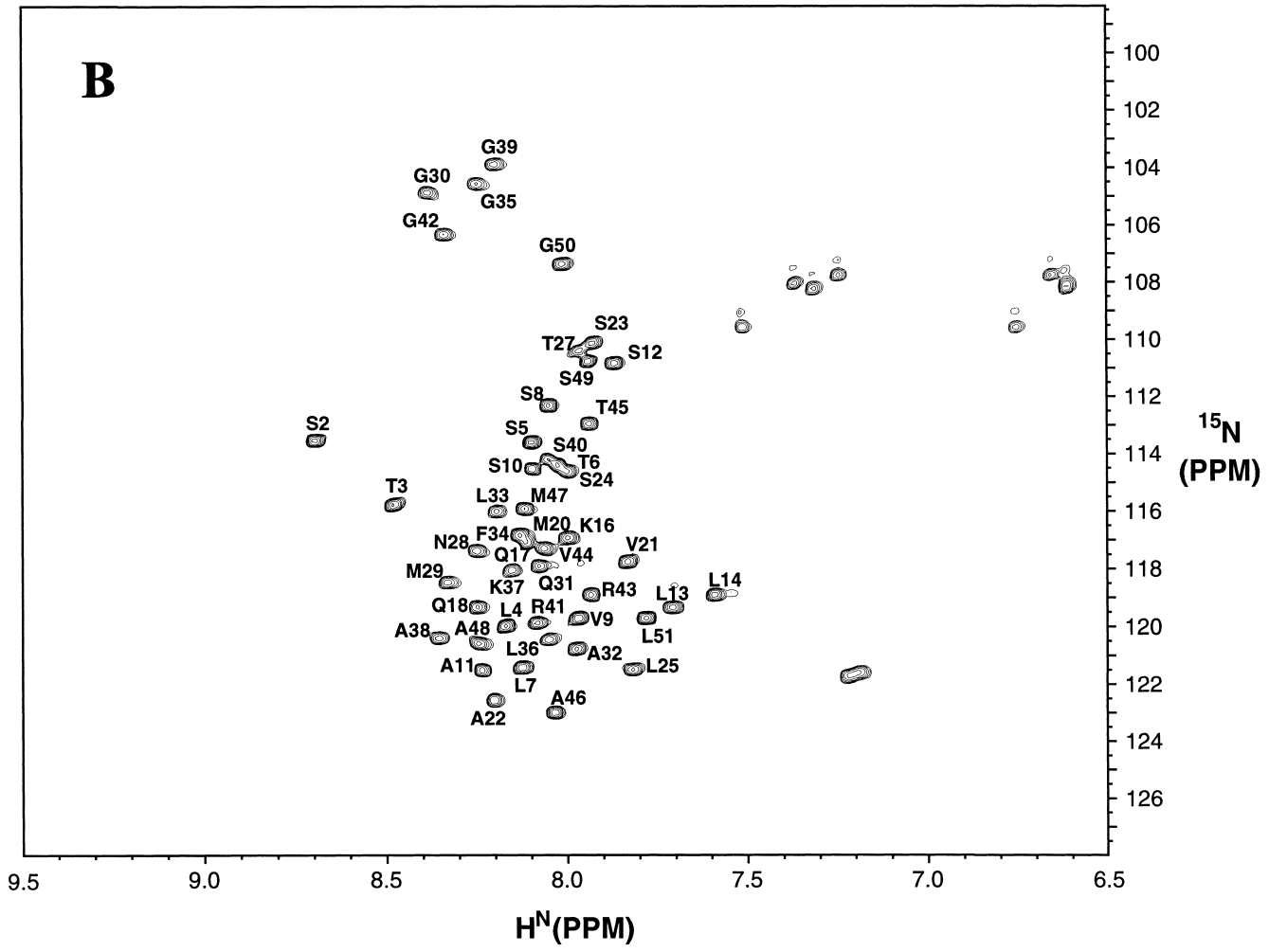

Fig. 3 (*continued*)

All short and medium range NOE results are summarized in Fig. 4B. The intensity ratio between the *d*NN and *d*αN cross-peaks is indicative of a helical propensity in the complete peptide [13,23]. Several medium range *d*NN(*i*, *i*+2) and *d*αN(*i*, *i*+3) NOEs could be assigned unambiguously. These contacts indicate that the regions Ala-1–Pro-15 and Met-29–Ser-49 form the most stable helical structures. The position of the assigned *d*αN(*i*, *i*+4) cross-peak between Phe-34 and Ala-38 suggests that the least flexible part of the induced helices is present in this region.

The deviations of the H^{α} chemical shifts from their random coil values are presented in Fig. 5B. The results imply that an N-terminal helix seems to extend from residues Thr-3 to Ser-12, whereas a helical structure is also present from Met-29 to Met-47. These observations confirm the data obtained from the short and medium range NOE connectivities. The C-terminal helix seems to be destabilized at Gly-39 as its H^{α} chemical shift is similar to the random coil value for glycines [22].

The fact that no long range NOEs were detected indicates that in TFE, no tertiary structure is present in trFD. The combined results suggest that trFd obtains two structurally independent helical regions in TFE. The first extending from Ser-2 to Leu-14, the second from Met-29 to Gly-50. The latter helical structure might be disturbed around Gly-39.

4. Discussion

In this study, for the first time, information about the position of structural elements in a higher plant chloroplast targeting sequence has been obtained. The NMR and CD data show that in water, trFd is predominantly in a random coil conformation, which was predicted to be an essential feature of chloroplast transit peptides [24]. Some medium range connectivities could be assigned (Fig. 4A), indicative of initial structures between Lys-16 and Gln-18 and between Ser-23 and Leu-25. Also, two nascent helical structures were found [23], between Val-9 and Ser-12 and between Lys-37 and Ser-40. As there were no large deviations from random coil H^{α} chemical shifts, and also no reliable medium range *d*αN(*i*, *i*+3) and *d*αN(*i*, *i*+4) contacts, it can be concluded that potential structured regions are not well-defined. The NMR data suggest that in 50% (v/v) TFE, N- and the C-terminal helices are present. This indicates that the nascent helices present in trFd in water are stabilized by TFE. In Fig. 6, the observed helices

are presented as α-helical wheels. For clarity, the C-terminal helical region is divided into two parts separated by the helix destabilizing residue Gly-39 (Fig. 5B).

The N-terminal helix from Ser-2 to Leu-14 seems to be amphipathic (Fig. 6A). Along one side of the helix, the hydrophobic amino acids Leu-4, Leu-7, Ala-11 and Leu-14 are present and on the opposite side, mainly the hydroxylated amino acids serine and threonine occur. It has been shown that the presence of the N-terminal 14 amino acids is indispensable for preFd import in chloroplasts, both in vitro [25] and in vivo [26]. This suggests that formation of an N-terminal amphipathic helix is necessary for transit sequence functionality. Fig. 6B shows the C-terminal helix from Met-29 to Gly-39. Pilon et al. [25] suggested that the semi-conserved FGLK motif can be involved in the efficiency rather than the occurrence of the import process. It is shown that Phe-34-Gly-35-Leu-36-Lys-37 is one of the most stable structural elements in trFd. In water, this region has a propensity towards secondary structure formation, because of the observed $d\alpha N(i, i+2)$ contacts (Fig. 4A). In TFE, it forms the most stable helical structure, as its H^{α}s show large upfield deviations from the random coil values (Fig. 5B). In addition, in TFE, this tetrapeptide shows the only $d\alpha N(i, i+4)$ connectivity. The findings suggest that in water, this region forms a nascent helix [23], which is possibly involved in the chloroplast import process by mediating the efficiency of induction

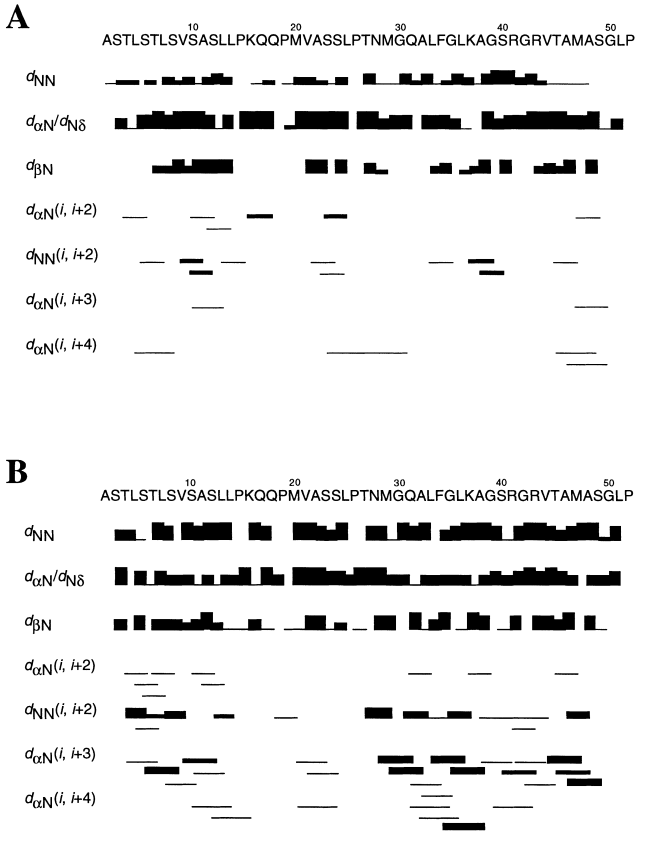

Fig. 4. Sequential and medium range NOE contacts observed in the 3D HSQC-NOESY spectra of trFd (A) for trFd in aqueous solution and (B) in 50% (v/v) TFE. The height of the bars is a indication of the relative intensity of the observed NOEs and is presented in five levels from low to maximum intensity. For prolines, the dNδ contacts are included in the $d\alpha$N row. Lines indicate NOEs whose presence or absence cannot be assessed due to overlap.

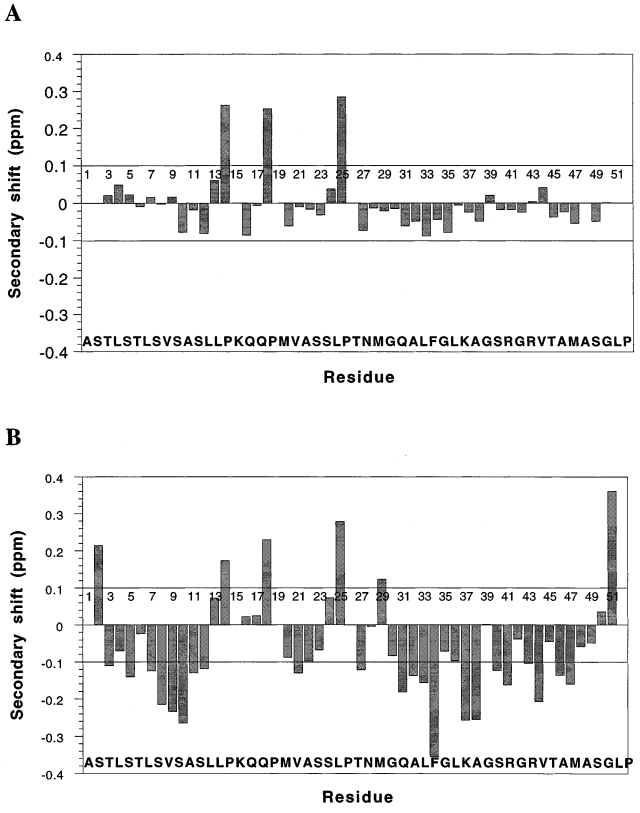

Fig. 5. Difference between the H^{α} chemical shifts for random coil peptides and those observed are displayed as a function of the residue position for trFd in aqueous solution (A) and in 50% (v/v) TFE (B).

of C-terminal helical structures in trFd. These helices might be involved more directly in the import process. In Fig. 6C, the second part of the C-terminal helix, formed by amino acids Gly-39–Gly-50, is shown. One side of the helix is enriched in hydrophobic residues (Ala-48, Val-44 and Met-47), whereas the opposite side contains mainly residues that are hydroxylated or have small side chains. These two regions are separated from each other by the positively charged amino acids Arg-41 and Arg-43.

It has been suggested that recognizable motifs are induced in transit peptides only after initial interactions with components of the chloroplast outer membrane. These interactions might occur with either membrane proteins or lipids. In the following section, we would like to focus on the putative interactions between the preFd transit peptide and outer membrane lipids. Comparison of the CD data for the chemically synthesized preferredoxin transit peptide obtained by Horniak et al. [9] and Pilon et al. [21] suggests that (part

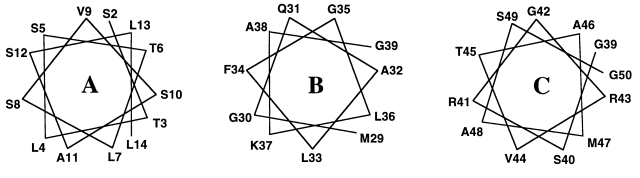

Fig. 6. Helical wheel presentations for the helices present in trFd in 50% (v/v) TFE: the N-terminal helix comprising amino acids 1–14 (A), the N-terminal part of the C-terminal helix, residues 29–39 (B), the C-terminal part of the C-terminal helix, residues 39–50 (C).

of) the structural changes that are induced in the peptide by TFE may also be induced by chloroplast lipids. Upon interaction with negatively charged lipids, the amount of helix of the transit peptide increased. This effect was more pronounced if also monogalactosyl diacylglycerol (MGDG) was present [9]. This implies that the N- and C-terminal helices presented for trFd in Fig. 6 can be formed upon interaction between precursor protein and lipids. It was shown that deletion of preFd residues 6–14 disturbs the interaction of preFd with MGDG [25]. Because the chloroplast is the only organelle exposing MGDG to the cytosol, it has been proposed that this interaction could be involved in precursor protein targeting. The present results lead to the suggestion that the nascent N-terminal helix in water formed by residues 9–12 may be stabilized by interactions with MGDG. The helix from Ser-2 to Leu-14 (Fig. 6A) can then be formed in the MGDG lipid interface by means of hydrophobic attraction between the side chains of the hydrophobic residues and the lipid acyl chains and hydrogen bonding between the headgroup hydroxyl groups and the hydroxyl groups of the serine and threonine residues. This helix induction might be chloroplast-specific and could, directly or indirectly, form a general motif to be recognized by the general receptor of the import machinery.

Deletions in the preFd transit sequence C-terminus, mainly from Gly-35 to Met-47, reduced but did not cause complete absence of import. Also, these C-terminal deletions caused a decreased insertion into phosphatidyl glycerol monolayers [25]. This suggests that the import efficiency can be influenced by interaction between the transit sequence C-terminus and negatively charged phospholipids. The part of the C-terminal helix presented in Fig. 6C might be involved in this interaction. The positive charges of the residues Arg-41 and Arg-43 may show electrostatic interactions with negatively charged lipids. The hydrophobic side chains of Val-44, Met-47 and Ala-48 are able to insert into the lipid acyl chains. It might even be possible that the OH-groups of the hydroxylated amino acids Thr-45 and Ser-49 are interacting with the lipid headgroup hydroxyl groups, which are present in all chloroplast outer membrane negatively charged lipids. In conclusion, the helical wheel shows that if these amino acids form an α-helix, this can easily be anchored into a membrane containing negatively charged lipids. This insertion might be necessary for trFd to anchor itself firmly in the chloroplast outer membrane and present a more N-terminal general motif to the protein import receptor.

The preFd transit sequence region 15–25 was shown not to be important for recognition of the chloroplast, but appears to be indispensable for import. Also, the region 15–25 was predicted to have a low probability for secondary structure [21], which is confirmed by the described NMR data. It was suggested that this region has a role as a flexible connector allowing optimal spatial arrangement of other parts of the transit peptide during later stages of translocation, which might account for the fact that this region shows often prolines and also large differences in length [25]. Another possibility, however, is that the N- and C-terminal structurally independent helices bring about that, somehow, this part of trFd is recognized by the proteinaceous receptor of the import machinery.

To obtain information about general motifs in chloroplast import signals, a comparison must be made between trFd and other targeting peptides. A few years ago, Lancelin et al. [27] published the NMR structure of the chemically synthesized chloroplastic ferredoxin transit peptide from the unicellular green alga *Chlamydamonas reinhardtii. Chlamydamonas* transit sequences generally differ from higher plant transit sequences in that they are shorter and, furthermore, they resemble both mitochondrial and chloroplast targeting sequences. The authors [27] showed that in water, the peptide is unstructured, whereas upon addition of TFE, a N-terminal helix was formed, followed by a more flexible region. They suggested that the N-terminal helix can anchor the *Chlamydamonas* transit peptide in the membrane by means of a hydrophobic ridge. The following unstructured region is exposed to the cytosol and available for a putative interaction with proteins of the translocation machinery. Our results show that also the *S. pratensis* preFd transit peptide in TFE forms an N-terminal helix with a hydrophobic ridge (i.e. Leu-4, Leu-7, Ala-11 and Leu-14). In addition, the subsequent region, Pro-15–Leu-25, remains unstructured in TFE. Moreover, the N-terminus was shown to be necessary for in vivo preFd import [26] and the N-terminus as well as the region 15–25 were found to be indispensable for in vitro transit sequence functionality [25]. This suggests that an N-terminal helix followed by an unstructured amino acid stretch is a vital functional unit for chloroplast protein import.

A closer look at the amino acid sequence of the transit peptide of the pea chloroplast precursor of the small subunit of ribulose-1,5-bisphosphate carboxylase/oxygenase (preSSU) [28] suggests that in theory, the characteristics of the helices formed in trFd might also be induced in this transit sequence. A stretch containing hydrophobic and hydroxylated amino acids (potentially forming an amphipathic helix) followed by a proline- and glycine-rich region is present at the N-terminus. Regions containing the FGLK motif can also be detected. Furthermore, electrostatic anchoring might be performed by positively charged residues located towards the C-terminus. In a study in which large parts of the transit peptide of preSSU were deleted [29], it was shown that, in analogy with the trFd N-terminus, the preSSU N-terminus is indispensable for import. However, not the N-terminus but the C-terminus was shown to interact specifically with MGDG [28,30]. Furthermore, the C-terminus also contains information which cannot be missed, for if it is deleted, import is completely absent.

The combined data suggest that for chloroplast-specific import, an N-terminal amphipathic helix is absolutely necessary, whereas a C-terminal helix is also favorable. Furthermore, a specific interaction with the chloroplast lipid MGDG appears to be required, possibly as an initial recognition motif that stabilizes the interaction between the incoming precursor protein transit peptide and the chloroplast.

NMR studies have also been performed with mitochondrial presequences (e.g. [31–33]). Similar to chloroplast transit peptides, presequences appear to be able to obtain non-continuous helical structures. It was suggested that for mitochondrial import, an amphipathical helix determines the import functionality as well. A major difference, however, between mitochondria and chloroplasts is that the latter lacks N-terminal positively charged amino acids. In general, it seems that functional targeting peptides are characterized by their ability to adopt, possibly upon interaction with organellar lipids, N-terminal amphipathic helices. These may be recognized by protein components of the translocation complex.

Acknowledgements: A.V.E. George is gratefully acknowledged for assisting the ^{15}N-labelling of the transit peptide. The NMR spectra were acquired at the SON NMR Large Scale Facility. Support from the European Community for MC (contract ERBFM GECT 950032) is acknowledged. This work is part of the program 'Chloroplast Protein Import' and was carried out under auspices of the Foundation of Life Sciences (SLW) with financial aid from the Netherlands Foundation for Scientific Research (NWO).

Appendix. Table of chemical shifts

Chemical shifts of trFd in aqueous solution (ppm)

Residue	^{15}N	H^N	H^α		H^β		H^γ		H^δ
Ala-1									
Ser-2	107.89	8.62	4.44		3.73				
Thr-3	109.29	8.35	4.21		4.08		1.05		
Leu-4	116.56	8.27	4.21		1.48				
Ser-5	109.11	8.33	4.34		3.73				
Thr-6	108.23	8.10	4.18		4.12		1.05		
Leu-7	116.20	8.08	4.20		1.47				0.76
Ser-8	109.29	8.23	4.31		3.71				
Val-9	114.12	8.13	3.99		1.96		0.79		
Ser-10	111.24	8.30	4.23		3.71				
Ala-11	118.51	8.30	4.16		1.26				
Ser-12	106.62	8.12	4.23		3.71				
Leu-13	115.85	8.08	4.22		1.44	1.50			
Leu-14	115.85	7.98	4.43		1.40	1.46			
Pro-15			4.25		1.78				
Lys-16	113.54	8.35	4.09		1.62		1.30		
Gln-17	113.72	8.33	4.16		1.81	1.91	2.20		
Gln-18	115.32	8.41	4.43		1.77	1.92	2.23		
Pro-19			4.35		1.71	2.14			
Met-20	113.37	8.45	4.26		1.89				
Val-21	114.25	8.11	3.95		1.89		0.77		
Ala-22	120.45	8.43	4.16		1.24				
Ser-23	107.34	8.25	4.28		3.69				
Ser-24	110.00	8.26	4.35		3.71				
Leu-25	117.09	8.12	4.45		1.45				
Pro-26			4.35		1.77				
Thr-27	105.93	8.16	4.12		4.05		1.06		
Asn-28	112.48	8.36	4.57		2.64	2.69			
Met-29	112.84	8.31	4.30		1.84	1.98	2.35		
Gly-30	101.85	8.38	3.79						
Gln-31	112.13	8.12	4.11		1.82	1.92	2.19		
Ala-32	117.27	8.24	4.13		1.22				
Leu-33	112.48	8.02	4.08		1.20	1.36			
Phe-34	111.77	8.01	4.43		3.01	2.87			
Gly-35	102.20	8.20	3.69	3.76					
Leu-36	113.54	7.94	4.16		1.42	1.50			
Lys-37	113.90	8.27	4.15		1.61	1.69			
Ala-38	117.27	8.25	4.09		1.22				
Gly-39	100.61	8.34	3.79	3.85					
Ser-40	107.70	8.10	4.29		3.75				
Arg-41	114.61	8.36	4.17		1.64	1.75	1.47	1.51	
Gly-42	101.50	8.29	3.77						
Arg-43	112.84	8.05	4.20		1.60	1.67	1.45		
Val-44	114.08	8.23	4.00		1.93		0.79		
Thr-45	110.53	8.16	4.15		4.07		1.05		
Ala-46	118.68	8.27	4.14		1.23				
Met-47	111.77	8.25	4.26		1.84				
Ala-48	117.62	8.27	4.17		1.26				
Ser-49	107.17	8.24	4.26		3.74				
Gly-50	102.74	8.26	3.80						
Leu-51	115.14	7.98			1.42				
Pro-52									

Chemical shifts of trFd in 50% (v/v) TFE (ppm)

Residue	^{15}N	H^N	H^α		H^β		H^γ		H^δ	
Ala-1			4.12							
Ser-2	109.80	8.64	4.63		4.03	3.91				
Thr-3	112.26	8.45	4.20				1.23			
Leu-4	115.96	8.09	4.20		1.61				0.85	
Ser-5	109.85	8.03	4.29		3.91	3.99				
Thr-6	110.86	7.96	4.28		4.04		1.22			
Leu-7	117.78	8.11	4.14		1.74		1.62		0.84	
Ser-8	108.45	8.00	4.21		3.99	3.90				
Val-9	116.12	7.94	3.83		2.13		1.01	0.91		
Ser-10	110.72	8.05	4.15		4.00	3.89				
Ala-11	117.65	8.19	4.15		1.45					
Ser-12	107.17	7.81	4.30		4.02	3.94				
Leu-13	115.58	7.65	4.33		1.75		1.56		0.80	
Leu-14	115.12	7.53	4.44		1.75		1.51		0.89	
Pro-15			4.37		2.27				3.54	3.78
Lys-16	113.15	7.92	4.29		1.87		1.49		1.77	
Gln-17	113.28	8.06	4.30		1.98		2.33			
Gln-18	115.51	8.19	4.51		1.97	2.11	2.35			
Pro-19			4.27		1.88	2.26	1.98		3.65	3.72
Met-20	112.95	8.05	4.34		2.07		2.51	2.57		
Val-21	113.95	7.76	3.93		2.10		0.92			
Ala-22	118.53	8.13	4.18		1.40					
Ser-23	106.19	7.85	4.36		3.88	3.96				
Ser-24	110.88	7.93	4.49		3.94	3.90				
Leu-25	117.78	7.74	4.56		1.71		1.53		0.89	
Pro-26			4.44		1.97	2.27	2.08		3.64	3.84
Thr-27	106.36	7.86	4.18				1.21			
Asn-28	113.74	8.18	4.69		2.85					
Met-29	114.72	8.29	4.31		2.07					
Gly-30	101.10	8.33	3.78	3.88						
Gln-31	114.17	8.02	4.10		2.13		2.36	2.46		
Ala-32	117.10	7.93	4.13		1.48					
Leu-33	112.92	8.15	4.11		1.73		1.47		0.84	
Phe-34	113.19	8.09	4.22		3.18					
Gly-35	100.78	8.21	3.78	3.90						
Leu-36	116.89	8.02	4.18		1.82		1.62		0.86	
Lys-37	114.46	8.12	4.02		1.81		1.37			
Ala-38	116.75	8.34	4.03		1.22					
Gly-39	100.17	8.18	3.87	3.96						
Ser-40	110.71	8.00	4.30		3.97	3.93				
Arg-41	116.15	8.01	4.13		1.88					
Gly-42	102.69	8.31	3.84	3.89						
Arg-43	115.53	7.87	4.19		1.89		1.63			
Val-44	113.41	8.02	3.86		2.11		0.93	1.01		
Thr-45	108.98	7.88	4.25		4.04		1.20			
Ala-46	119.25	7.98	4.15		1.45					
Met-47	112.18	8.08	4.27		2.13		2.53	2.68		
Ala-48	116.63	8.24	4.24		1.44					
Ser-49	106.87	7.87	4.37		3.94					
Gly-50	103.55	7.93	3.95							
Leu-51	116.06	7.71	4.63		1.62		1.47		0.74	
Pro-52										

References

[1] Chua, N.-H. and Schmidt, G.W. (1979) J. Cell Biol. 81, 461–483.
[2] Reiss, B., Wasmann, C.C. and Bohnert, H.J. (1987) Mol. Gen. Genet. 209, 116–121.
[3] Smeekens, S., Geerts, D., Bauerle, C. and Weisbeek, P. (1989) Mol. Gen. Genet. 216, 178–182.
[4] Van den Broeck, G., Timko, M.P., Kausch, A.P., Cashmore, A.R., Van Montagu, M. and Herrera-Estrella, L. (1985) Nature 313, 358–363.
[5] Von Heijne, G., Steppuhn, J. and Herrmann, R.G. (1989) Eur. J. Biochem. 180, 535–545.
[6] De Kruijff, B., Pilon, M., Van 't Hof, R. and Demel, R. (1996) in: Molecular Dynamics of Biomembranes (Op den Kamp, J.A.F., Ed.), Vol. H96, pp. 99–136, Springer-Verlag, Berlin.
[7] Wray, V., Mertins, D., Kiess, M., Henklein, P., Trowitzsch-Kienast, W. and Schubert, U. (1998) Biochemistry 37, 8527–8538.
[8] Van 't Hof, R., Van Klompenburg, W., Pilon, M., Kozubek, A., De Korte-Kool, G., Demel, R.A., Weisbeek, P.J. and De Kruijff, B. (1993) J. Biol. Chem. 268, 4037–4042.
[9] Horniak, L., Pilon, M., Van 't Hof, R. and De Kruijff, B. (1993) FEBS Lett. 334, 241–246.
[10] De Jongh, H.H.J. and De Kruijff, B. (1990) Biochim. Biophys. Acta 1029, 105–112.
[11] Hurd, R.E. and John, B.K. (1991) J. Magn. Reson. 91, 648–653.
[12] Wishart, D.S., Bigam, C.G., Yao, J., Abildgaard, F., Dyson, H.J., Oldfield, E., Markley, J.L. and Sykes, B. (1995) J. Biomol. NMR 6, 135–140.
[13] Wüthrich, K. (1986) NMR of Proteins and Nucleic acids, Wiley, New York.
[14] Bodenhausen, G. and Ruben, D.J. (1980) Chem. Phys. Lett. 69, 185–189.
[15] Kay, L.E., Keifer, P. and Saarinen, T. (1992) J. Am. Chem. Soc. 114, 10663–10665.
[16] Marion, D., Driscoll, P.C., Kay, L.E., Wingfield, P.T., Bax, A., Gronenborn, A.M. and Clore, G.M. (1989) Biochemistry 28, 6150–6156.
[17] Marion, D., Ikura, M., Tschudin, R. and Bax, A. (1989) J. Magn. Reson. 85, 393–399.
[18] Shaka, A.J., Lee, C.J. and Pines, A. (1988) J. Magn. Reson. 77, 274–293.
[19] Delaglio, F., Grzesiek, S., Vuister, G.W., Zhu, G., Pfeifer, J. and Bax, A. (1995) J. Biomol. NMR 6, 277–293.
[20] Kleywegt, G.J., Vuister, G.W., Padilla, A., Knegtel, R.M.A., Boelens, R. and Kaptein, R. (1993) J. Magn. Reson. B102, 166–176.
[21] Pilon, M., Rietveld, A.G., Weisbeek, P.J. and De Kruijff, B. (1992) J. Biol. Chem. 267, 19907–19913.
[22] Wishart, D.S. and Sykes, B.D. (1994) Methods Enzymol. 239, 363–392.
[23] Dyson, H.J., Rance, M., Houghten, R.A., Wright, P.E. and Lerner, R.A. (1988) J. Mol. Biol. 201, 201–217.
[24] Von Heijne, G. and Nishikawa, K. (1991) FEBS Lett. 278, 1–3.
[25] Pilon, M., Wienk, H., Sips, W., De Swaaf, M., Talboom, I., Van 't Hof, R., De Korte-Kool, G., Demel, R., Weisbeek, P. and De Kruijff, B. (1995) J. Biol. Chem. 270, 3882–3893.
[26] Rensink, W.A., Pilon, M. and Weisbeek, P. (1998) Plant Physiol. 118, 691–699.
[27] Lancelin, J.-M., Bally, I., Arlaud, G.J., Blackledge, M., Gans, P., Stein, M. and Jacquot, J.-P. (1994) FEBS Lett. 343, 261–266.
[28] Pinnaduwage, P. and Bruce, B.D. (1996) J. Biol. Chem. 271, 32907–32915.
[29] Reiss, B., Wasmann, K., Schell, J. and Bohnert, H.J. (1989) Proc. Natl. Acad. Sci. USA 86, 886–890.
[30] Van 't Hof, R., Demel, R.A., Keegstra, K. and De Kruijff, B. (1991) FEBS Lett. 291, 350–354.
[31] Chupin, V., Leenhouts, J.M., De Kroon, A.I.P.M. and De Kruijff, B. (1995) FEBS Lett. 373, 239–244.
[32] Jarvis, J.A., Ryan, M.T., Hoogenraad, N.J., Craik, D.J. and Høj, P.B. (1995) J. Biol. Chem. 270, 1323–1331.
[33] Karslake, C., Piotto, M.E., Pak, Y.K., Weiner, H. and Gorenstein, D.G. (1990) Biochemistry 29, 9872–9878.