Refined solution structure of a liganded type 2 wheat non-specific Lipid Transfer Protein.

Jean-Luc Pons¹, Frédéric de Lamotte², Marie-Françoise Gautier², Marc-André Delsuc¹*.

¹ Centre de Biochimie Structurale, INSERM, CNRS, U. Montpellier I, 15 Ave C. Flahault, 34 060 Montpellier - France.

² UMR Génomique Appliquée aux Caractères Agronomique, INRA , 2 Place Viala, 34 060 Montpellier cedex 01- France.

* address for correspondence
tel : +33 467 04 34 36
fax : +33 467 52 96 23
e-mail : MA.Delsuc@cbs.cnrs.fr

Running title :
Refined solution structure of a liganded wheat ns-LTP2.

Keywords :
3D NMR spectroscopy; lipid transfer protein; L⁰-palmitoylphosphatidylglycerol ligand; nitrogen labeling; protein assignment; protein structure; wheat protein.
Abbreviations and symbols

Summary

The refined structure of a wheat type 2 non-specific lipid transfer protein (ns-LTP2) liganded with \( \text{La-palmitoylphosphatidylglycerol (LPG)} \) has been determined by NMR.

The \(^{15}\text{N}\) labeled protein was produced in \( \text{P. pastoris} \). Physicochemical conditions and ligandation were intensively screened in order to obtain the best NMR spectra quality.

This ns-LTP2 is a 67 residues globular protein with a diameter of about 30 Å. The structure is composed of five helices forming a right super helix.

The protein presents an inner cavity, which has been measured at 341 Å\(^3\). All helices display hydrophobic side chains oriented toward the cavity. The phospholipid is found in this cavity. Its fatty acid chain is completely inserted in the protein, the LPG glycerol moiety being located on a positively charged pocket on the surface of the protein. The super-helix structure of the protein is coiled around the fatty acid chain.

The overall structure shows similarities with ns-LTP1. Nevertheless, large 3D structural discrepancies are observed for the H3 and H4 alpha helices, the C-terminal region, and the last turn of the H2 helix. The lipid is orthogonal to the orientation observed in ns-LTP1. The volume of the hydrophobic cavity appears to be in the same range as the one of ns-LTP1, despite ns-LTP2 is shorter by 24 residues.
Introduction

Plant non specific lipid transfer proteins (ns-LTPs) were first isolated from spinach leaves and named on their ability to mediate in vitro the transfer of phospholipids between membranes (1). ns-LTPs are widely distributed and form a superfamily of related proteins subdivided into two families: the type 1 ns-LTPs (ns-LTP1) and the type 2 ns-LTPs (ns-LTP2) (see ref (2,3) for review). Both families are multigenic and more than 150 sequences of plant ns-LTPs are listed in databases. Only a limited number of proteins have been isolated from plant and in vitro lipid transfer or binding has been demonstrated for an even more limited number of proteins.

The biological functions of ns-LTP1 have not yet been clearly determined, the most favored hypothesis being a role in the transport of cutin monomers (4,5) or in plant defense mechanisms (6-8) for ns-LTP1. ns-LTP2 gene expression has been reported in Zinnia elegans cell differentiation process (9,10), barley and rice developing seeds (11,12), under abiotic stress conditions in barley roots (12) or during nodulation in Vigna unguiculata root hairs (13). However there is no biological evidence of their function in these different contexts. The recent discovery that some ns-LTPs are pan-allergens of plant-derived foods has brought new interest for their study. Most of the ns-LTP allergens identified so far belong to the ns-LTP1 family (14-17) while ns-LTP2 have been reported only as a potent allergen of the pollen of Brassica rapa (18).

The three-dimensional structure of four cereal ns-LTP1 has been determined, i.e. wheat (19), barley (20), maize (21,22) and rice (23,24). In addition, seven structures of plant ns-LTP1 in complex with ligands have been determined, including those of maize ns-LTP1 with palmitate (21) or palmitoyl-lyso-phosphatidylcholine (22), barley ns-LTP1 with palmitoyl coA (25) or palmitate (26) and wheat ns-LTP1 with di-myristoyl-phosphatidylglycerol (27), lyso-myristoyl-phosphatidylcholine (28) or prostaglandin B2 (29). All these data showed that ns-LTP1 are compact single domain proteins whose fold is stabilized by four disulfide bonds. They are characterized by a four [-]-helix bundle and a C-terminal region with no regular secondary structure. The most interesting feature of ns-LTP1 structure is the tunnel-like hydrophobic cavity which runs through the molecule and appears as a potential site for lipid binding. Although, plant ns-LTP1s exhibit very similar global folds, the shape and size of this hydrophobic cavity vary considerably depending on the protein and/or on the ligand. This clearly indicates a high plasticity of the cavity that is able to accommodate a variety of hydrophobic molecules. In contrast, an antifungal protein extracted from onion seeds that showed a structure similar to those
of ns-LTP1 except that the internal cavity, obstructed by several aromatic side chains, is unable to transfer lipids (30).

ns-LTP1 have been studied more extensively than ns-LTP2 that are distinct in terms of primary sequence with less than 30% homology (Figure 1), size (7 kDa versus 9 kDa), and lipid transfer efficiency. In wheat, both ns-LTP1 and ns-LTP2 have been biochemically characterized and ns-LTP2 exhibits a higher lipid transfer activity than ns-LTP1 (31). They share the same 8 cysteines skeleton but their disulfide bond assignment has been shown to be different (32). Because the wheat ns-LTP2 exhibits a lipid transfer activity, one can hypothesize the presence of an hydrophobic cavity. However, whether ns-LTP2 have a similar or different fold than ns-LTP1 remained to be elucidated at the initiation of the present work. This paper presents the refined solution structure of a recombinant wheat ns-LTP2 liganded with a C16 lyso-phospholipid. In the final stages of preparation of this manuscript, the structure of a rice ns-LTP2 was reported (33) (PDB entry not available as of date of submission).
Experimental Procedures

Production / purification

The Pichia pastoris transformant GS115-Tdtp18-tr5.2 (34) expressing a wheat ns-LTP2 was used for the production of the $^{15}$N-labeled protein. Production was carried out in an Applikon fermentor (400 ml culture) with 99.4% $^{15}$N-labeled ammonium sulfate (Eurisotop) as nitrogen source. Labeled ($^{15}$NH$_4$)$_2$SO$_4$ (0.9% w/v) was added from the very beginning of the biomass production phase and during the induction phase at 0, 24 h and 48 h (0.27, 0.27 and 0.36 w/v respectively). Production of ns-LTP2 was induced by methanol after 22 h of culture and lasted for 77 h. The amount of secreted protein was estimated by densitometer analysis of SDS gels, using purified ns-LTP2 as a standard. Folding and accumulation of the recombinant protein was also directly monitored by $^{15}$N HSQC NMR (34). A non labeled sample was also produced as described earlier (35).

The protein was purified from culture supernatant by a single step procedure (36) using expanded-bed chromatography (Streamline SP-XL, Amersham Biosciences), 650 mg of protein were obtained. The column was equilibrated in sodium acetate 30 mM, pH 5.5 and proteins were eluted with 1M NaCl. The purified protein was dialyzed against H$_2$O, lyophilized and analyzed by SDS-PAGE, MALDI-TOF and HSQC. MALDI-TOF-MS experiments were conducted on BiFlex III mass spectrometer (Bruker Daltonics, Bremen, Germany). 100 pmoles of the protein were solubilized in 10 [l] of water containing 0.1% T.F.A. A 0.5 [l] sample of this solution was mixed with 0.5 [l] of matrix (a saturated solution of alpha-cyano-4-hydroxycinnamic acid in water/acetonitrile 50/50 v/v) and deposited on the target (dried droplet preparation). Spectra were recorded on the linear positive mode using the Pulsed Ion Extraction. For the tryptic digestion analysis, the protein was reduced by DTT and alkylated by iodoacetamide prior to digestion. The protein was digested for 4 h by trypsin (modified, sequencing grade from Boehringer Mannheim, Mannheim, Germany) in 25 mM ammonium carbonate buffer (pH 8) at an enzyme/protein ratio of 1/20 (w/w). Proteolysis was stopped by adding diluted TFA. The digest mixture was analyzed by MALDI-TOF-MS in the positive reflector mode.

MS analysis revealed the heterogeneity of the purified protein, 15% being glycosylated. The sugar moiety, assumed to be mannose, were located in the first 29 residue fragment. The protein was thus further purified using reverse phase chromatography (Source 15 RPC / Akta Purifier 10, Amersham Biosciences), for a final amount of 300 mg of protein.
NMR and fluorescence ns-LTP2/Lipid screening

Preliminary experiments revealed that NMR spectra could be significantly improved when ns-LTP2 interacts with a lipid. In order to find the best conditions for the study, several ligands and physico-chemical conditions were screened by fluorimetry as well as by NMR. The interaction with L\[\text{-}\]palmitoylphosphatidylglycerol (LPG) L\[\text{-}\]palmitoylphosphatidylcholine (LPC), dimyristoylphosphatidylglycerol (DMPG), dimyristoylphosphatidylcholine (DMPC) and cholesterol was tested by fluorimetry and NMR. Lipids from Avanti polar were prepared in solution in ethanol at a concentration of 10 mM or in small unilamellar vesicle obtained by sonication in H\textsubscript{2}O.

Fluorescence experiments were performed on a ISS instrument, equipped with Hamatsu detector, at 294 K in a 150 \textmu l volume, excitation was set at 275 nm (300 W), and spectra were recorded between 285 nm and 400 nm. Increasing amounts of ligand (0.1 to 100 equivalent) were added to a solution of ns-LTP2 (30 \textmu M).

NOESY spectra were recorded for 12 h on a Bruker 600 MHz spectrometer at 300 K on a 2.7 mM ns-LTP2 solution in 60 mM phosphate buffer containing 1 equivalent lipid. HSQC spectra were recorded for 30 min on a Bruker 400 MHz on a 1.1 mM ns-LTP2 solution, 10\% D\textsubscript{2}O. The impact of several parameters was investigated : temperature (300 to 315 K) ; pH (3.45 to 6.5) ; ionic strength (0 to 100 mM NaCl and 0 to 60 mM phosphate buffer) and lipid concentration (0 to 1.5 equivalent).

Sequence alignment and Model Prediction

Primary sequence alignment were performed with the following software : CLUSTAL W (37) and Psi-Blast (38). Early 3D models construction and 3D comparison were performed using the web metaserver: http://bioserv.cbs.cnr.fr (39) with the following threading methods: 3DPSSM(40) and TITO(41).

NMR Spectroscopy

All NMR samples were prepared by dissolving the \textsuperscript{15}N labeled and unlabeled recombinant wheat ns-LTP2 in H\textsubscript{2}O (10\% D\textsubscript{2}O), at a concentration of 2.8 mM in presence of 70 mM phosphate at pH 3.5 and 1 mM NaN\textsubscript{3}. 1.5 equivalent of L\[\text{-}\]palmitoylphosphatidylglycerol (1-Palmitoyl-2-Hydroxy-sn-Glycero-3-[Phospho-rac-(1-Glycero)] sodium salt named LPG) was added to all the protein preparations. Optimum phospholipid concentration was determined by a step by step titration, monitored by \textsuperscript{15}N HSQC. An improvement of the spectrum was observed with increasing lipid concentration from 0 to 1.2 equivalent, further lipid addition having no effect on the protein spectrum. NMR spectra were recorded on a Bruker AMX 600 spectrometer,
operating at 599.94 MHz, equipped with a triple-resonance inverse probe with a field gradient unit on the z-axis. All data processing was performed with the version 4 of the Gifa software (42).

On the $^{15}$N labeled sample, a 3D HSQC-TOCSY (45 ms mixing time, 80 hours acquisition 305.2 K), a 3D HSQC-NOESY (200 ms mixing time, 64 hours acquisition 305.2 K) and a 3D HSQC NOESY (200 ms mixing time, 82 hours acquisition, 295.2 K) were performed. 3D experiments were processed by linear prediction along the proton and $^{15}$N indirect axes in order to obtain at least 256 complex points along each axis. Two 2D NOESY (200 ms mixing time 18.5 hours acquisition at 295.2 K and 305.2 K) were also performed on this sample.

On the unlabeled sample, COSY, NOESY (200 ms mixing time), and TOCSY experiments were performed at the temperatures of 295.2 K, 310 K and 323 K. Additional experiments were also performed in pure D$_2$O. Three NOESY ([$\tau$],=50 ms, 100 ms, 200 ms) and a COSY experiments were performed on the unlabeled protein liganded with a perdeuterated dodecylphosphocholine lipid (DPC) as ligand and confirmed the assignment. Deuterium exchange experiments were performed by following the evolution of the TOCSY spectrum on a 3.3 mM sample in D$_2$O (mixing time 60 ms, pH 3.3, 1 equivalent LPG. In order to follow the most labile amide protons, the lyophilized protein was dissolved in D$_2$O at a temperature of 273 K. The exchange was then monitored by several experiments, raising slowly the temperature. A first spectrum was acquired at 273 K, starting directly after mixture and a second one at the same temperature starting 3.5 hours after the mixture. Then several spectra were acquired while raising the temperature in order to follow less labile protons: at 295.2 K 7 hours after initial mixture, at 310 K 21 hours after mixture. The sample was then left at room temperature during 4 days, and a final experiment was acquired at 295.2 K. A natural abundance $^{13}$C HSQC experiment was performed in D$_2$O to confirm methylene assignment. ROESY experiments were performed at 295.2 K (50 ms and 100 ms) on the unlabeled protein liganded with LPG and with DPC. A translational diffusion measure (DOSY (43) was performed at 295.2 K (diffusion duration 200 ms - gradient duration 1.5 ms, with varying intensities from 0.5 to 47 G/cm, one hour acquisition). DOSY experiments were performed on the unlabeled sample at 290 K, the STE-LED sequence (44,45) was used with a WATERGATE filter applied for water suppression. A set of small monomeric globular proteins was measured under the same conditions to determine the molecular mass calibration (46). Several $^{31}$P NMR experiments were performed at 295.2 K on a AMX 400 Bruker spectrometer, on a 2 mM sample (pH 3.5, 80 mM NaH$_2$PO$_4$ with 2.5 mM of LPG. Two $^{31}$P EXSY 2D experiments (mixing time of 0.2 s and 2 s respectively) did not present any evidence of a bound / free equilibrium.
All the NMR data-sets have been deposited to the NMRb databank (http://nmrb.cbs.cnrs.fr) (NMRb # ltpt2a).

**NMR spectra Assignment**

The assignment of the wheat ns-LTP2 complexed with the phospholipid was performed from the set of 3D (HSQC-NOESY & HSQC-TOCSY) and 2D experiments (COSY, TOCSY, NOESY) using the sequential assignment strategy (47). with the help of the Rescue software for the amino-acid typing step (48). The assignment module of the Gifa program was used for this purpose (49).

In the HSQC experiment, all amide peaks could be found. All the side chain labile hydrogens from Lys, Asn, Arg and Gln residues were also assigned. Additional peaks belonging to the glycosylated proteins were sorted out in the HSQC spectrum from their reduced intensity. On the basis of sequential dNN(i,i+1) and dN(i,i+1) all amino-acids have been found, and 98% of all the non-labile 1H chemical shifts have been assigned. The natural abundance13C HSQC spectrum was used to help the assignment of side-chain methylenes. Stereo-specific assignment of protons of methylene groups was performed by examination of the NOE intensity and coupling constant patterns.

Under the experimental conditions, and because of the blurring of the NOE spectra due to the exchange cross peaks, the phospholipid signal could not be unambiguously assigned, except for the C16-terminal methyl protons.

The complete assignment of the 1H and 15N chemical shifts of the wheat ns-LTP2 complexed with LPG has been deposited in the BioMagResBank (BMRB # 4977).

**NMR Structure calculation and analysis**

The 2D and 3D NOESY experiments, performed at three different temperatures (295.2 K, 310 K and 323 K), have been used to extract all the distance constraints used for the structure reconstruction. Peak intensities were calibrated against a set of reference peaks, using the standard tools provided with the Gifa program (49). Intensity were analyzed with a 1/r6 law, and distance constraints were obtained from the intensities by classifying in long, medium and short range distances, pseudo-atoms corrections were used. Additional constraints were obtained on the angle by measuring 1J NH-H coupling constants using the Ludvigsen method (50) and on the l angle by considering the NOESY peak intensity between HN, Hl, spins and Hl, Hl, spins, when a stereospecific assignment was available.
The dynamical annealing protocol (anneal.inp) of the CNS software (51) was used to generate the protein structure from the set of constraints. A first set of structures was obtained from the constraint list, without any constraint on the cysteine linkage and with no phospholipid in the topology. From this set of structures, the disulfide linkage could be unambiguously assigned, as all the sulfur atoms were located in compatible distances. The observed disulfide bridges are: Cis-25/Cis-60, Cis-10/Cis-24, Cis-2/Cis-34, and Cis-36/Cis-67. H-H NOE contacts were further observed in the NOESY map which confirmed this assignment.

The model of the liganded protein was constructed from the same set of constraints. Additional constraints were used, corresponding to the observed connectivity of the C16 terminal CH3 of the fatty acid chain with H of Tyr-44 and Tyr-47 and the middle of the lipidic chain (C8) with H of Phe-35 with a long distance constraint. No additional constraint was used to force the exit of the phospholipid from the protein core. The distance constraints corresponding to the S-S bounds were then added to the constraint list. The dynamical annealing protocol was used for the calculation of 250 structures, from which the 10 structures with the lowest global energy were conserved. The obtained set of structures was deposited to the Protein Data Bank (PDB # 1N89).

The Procheck (52) program was used to check the quality of the obtained structures, as well as to compute the Ramachandran maps. The Voidoo (53-55) was used to compute the volume of the inner cavity.

**Results & Discussion**

*Production & purification*

Recombinant proteins preparations were obtained as previously described (34-36). The N-terminal sequence (ACQASQLAVC) of the recombinant ns-LTP2, determined by mass spectroscopy, is identical to those of the wheat-purified ns-LTP2 (56) indicating that the recombinant protein was correctly processed by the *P. pastoris* KEX2 protease.

The mass spectroscopy was performed on the ^15^N labeled protein. A 7055 Da average molecular mass was measured, confirming a ^15^N isotopic labeling over 95%. The mass spectrum also revealed that ca. 15% of the protein has been glycosylated with 1 to 4 C6 sugar moieties, assumed to be mannose (162 Da). This is consistent with the fact that *P. pastoris* add O-oligosaccharides composed solely of mannose residues. Peptide mass fingerprint of the digested protein revealed that only the first tryptic fragment at the N-terminus of the protein (fragment 1-29) is glycosylated.
N-glycosylation requires the Asn-X-Ser/Thr consensus sequence while O-glycosylation requires the presence of a Ser or Thr residue. The wheat ns-LTP2 does not contain the Asn-X-Ser/Thr sequence in its first 29 residues; however it does contain four serine residues (at positions -5, 12, 16 and 21) and one threonine residue (at position 27), residues which are accessible for possible glycosylation. The presence of a proline residue in the vicinity of a serine or threonine residue can enhance O-mannosylation (57). Of the four serine residues available for O-linkage, residue 21 is the only one close to a proline residue. All these results indicated that a small fraction of the wheat ns-LTP2 is expressed as a O-glycosylated protein. To our knowledge, no glycosylation has been reported for ns-LTP purified from plant.

**Physicochemical context screening & ligand choice**

Intensive screening of the physicochemical solution conditions was performed in order to find a set of conditions that provides NMR spectra of good quality. Ligand nature and ionic strength conditions were found to be critical.

Fluorescence screening experiments were performed on a series of phospholipids and fatty acids, highlighting a higher affinity for negative charged phospholipids. We finally selected LPG with an average chain length (C16). The titration of ns-LTP2 by LPG, monitored by NMR, shows that the ns-LTP2 becomes more structurally constrained in the presence of this ligand. A narrowing and a spreading out of the peaks are observed on the 2D HSQC spectrum. This evolution is stabilized around 1 equivalent of LPG. Addition of 70 mM phosphate buffer was observed to also improve spectral quality (Figure 2). On the other hand, we found that the protein could adapt a large range of pH (3.45 to 6.5) and temperatures (273 K to 323 K) without important modification of the HSQC spectra.

Under these conditions, DOSY experiment was used to determine the translational diffusion coefficient of the protein in the experimental conditions, found at 117 m2/s. Based on a calibration performed on a series of small globular proteins measured in the same conditions, this value confirms that the sample is monomeric. This result is in good agreement with the relaxation study, which has determined a rotational correlation time compatible with a monomeric form (58).

Under these optimized conditions, the 1H-15N HSQC spectrum of the 15N labeled ns-LTP2 protein liganded with LPG, displays more peaks than expected, considering the number of residues. A second step of purification, designed to remove the glycosylated proteins was performed on the sample, and a 1H-15N HSQC recorded. This
spectrum displays a reduced number of peaks, but the peaks are still too numerous (labeled with a question mark in Figure 2).

1D and EXSY 2D $^{31}$P spectra were recorded on this sample. Two phosphorous lines can be observed in the 1D spectrum, corresponding to the bound phospholipid and to the slight excess of free phospholipid in solution. No additional peaks, nor any exchange peak could be observed in the EXSY spectra, even at very long mixing time. In order to observe eventual chemical exchange, a 2D $^1$H ROESY experiment was performed on the LPG-ligated protein. In this spectrum a strong exchange peak is visible in the H[β] region, as well as several weaker peaks close to the diagonal in the aliphatic and amide regions. The same peaks are observed when the LPG ligand is replaced with a fully deuterated DPC phospholipid. An exchange between the holo and the apo states of the protein being excluded by the $^{31}$P experiments, this is an indication that a conformational equilibrium between a major and a minor form of the protein is taking place. This equilibrium is the source of the additional peaks observed in the HSQC spectra. It was taken into account during the assignment phase, and several peaks were assigned to the minor form. However no effort was done to fully assign this minor form.

**Assignments and secondary structure elements**

The assignment of the protein resonances was performed from the set of $^{15}$N-edited NOESY and TOCSY spectra. The sequential strategy was used, aided by the $^{15}$N and natural abundance $^{13}$C HSQC spectra (Figure 3).

The solvent protection experiments indicate that the secondary structure is quite strongly established, as about 40% of the amide protons remain unexchanged during the first hours of the D$_2$O exchange experiment at 1°C, and 8 remain unexchanged after 5 days at room temperature. The solvent protection patterns, as well as J-coupling and NOESY patterns are indicative of a structure mostly helical. On the other hand, the secondary structure prediction program Jpred (59) anticipates 2 helical zones, ranging from residues 22 to 29, and from residues 33 to 39.

Chemical assays as well as mass spectroscopy have shown that all the 8 cysteines of the protein are engaged in disulfide bridges. We have not been able to unambiguously assign the H[α]-H[β] NOE contacts characteristic of this structure due to the crowding of this spectral region. However, after the first run of structure generation, all the obtained structures exhibited side chain proximities permitting the disulfide skeleton based on these pre structures to be assigned. The disulfide bridges thus found are: Cys-25/Cys-60, Cys-10/Cys-24, Cys-2/Cys-34, and Cys-36/Cys-67. This is in agreement with the chemically determined assignment (32).
The $^1$H assignments of the LPG in the complexed form were obtained from 2D COSY homonuclear experiments on the free LPG, and by comparing the homonuclear spectra obtained from the LPG-ns-LTP2 complex and from the DPC-ns-LTP2 complex. A few chemical shifts of the fatty acid chain were clearly identified: the terminal methyl group (C16), its vicinal methylene (C15) as well as the proximal methylene (C2), the glycerol moiety attached to the fatty acid chains was also assigned. Some intermolecular NOE contacts were observed.

Structure determination

From the complete set of geometric constraints extracted from the NMR spectra, a set of 10 structures has been obtained. They do not display any important constraint violation, and all the residues are localized in the allowed regions of the Ramachandran plot. The ensemble of the 10 best structures present an rmsd computed on the residues 2 to 67 of 0.9 Å for all the heavy atoms, and a rmsd of 0.67 Å for the backbone atoms. All the statistics of the geometrical constraints and the structure reconstruction are given in table 1. The set of solution structures of ns-LTP2 liganded with LPG as obtained from this experimental work is presented in figure 4.

The protein is observed as a globular protein with a diameter of about 30 Å. The structure is composed of five helices arranged in a super-helix tertiary structure. Helix 1 is a 3-10 helix, encompassing residues from 7 to 15. All other helices are $\alpha$ helices. Helix 2 is located from residue 22 to residue 31; helix 3 from residue 34 to residue 40, helix 4 from residue 44 to residue 49 and helix 5 from residue 51 to residue 60. The overall fold is a right super helix. The localization of the helices is in good agreement with the proton exchange experiment, which has shown that there are five main zones in which the amide protons are protected against solvent exchange: (14-15, 26-31, 38-41, 47-50, 56-64). The LPG molecule is found partly embedded in the structure of the protein, with the super-helix structure of the protein coiled around the fatty acid chain, and with the phosphate group and the external glycerol moiety unstructured and located outside of the core of the protein.

Helices 1 and 2 are organized in a near anti-parallel conformation. The contact between helices 1 and 2 is tightened by the Cys-10/Cys-24 disulfide bridge. They are linked by an extended strand from Ser-16 to Gly-22. Lys-19 appears to be oriented toward the solvent, and does not display any NOE contact with other residue. It should be noted that in a previous dynamic study, it has been observed that the H-N vector of Lys-19 is highly mobile in the ns range (58). Helices 3, 4 and 5 form a square configuration. Helix 3 contains a characteristic Cys-34 Phe-35 Cys-36 pattern, with Phe-35 buried into the structure, and contributing to the hydrophobic core of the structure. Cys-34 and 36 are respectively engaged in disulfide bridges with Cys-2 and Cys-67, forming
two diametrically opposed bonds relative to the helix axis. All prolines are observed in trans conformation, as
confirmed by the H[4](i) - H[4]-Pro(i+1) contacts observed in the NOESY spectra, for all the proline residues
(Pro-20, Pro-42, Pro-51, and Pro-65).

Several additional secondary structure elements can be observed in the structure. The C-terminal of helix 1
presents an unusual hydrogen bond pattern, with the carboxyl moiety of Ser-12 being engaged with the HN of
Ala-18. This structure is found in the 10 structures generated, and the HN of residue Ala-18 is found to be
protected against solvent exchange, thus confirming this result. A type-1 gamma turn is observed between
helices 2 and 3, corresponding to residues Gln-31 Gln-32 and Gly-33. A type-1 beta turn is observed between
helices 3 and 4, corresponding to residues Asp-41 Pro-42 Thr-43 and Tyr-44, with acidic head of Asp-41
engaged in a H-bound with the HN of Thr-43. Three classic helix capping can also be observed : the N-capping
of helix 2 with the OH of Ser-21 H-bounded to the HN of Glu-23; the C-capping of helix 2 with side chain of
Gln-32 bounded to the CO of Arg-29; the N-capping of helix 5, with the OH of Ser-50 bounded to the HN of
His-52. Finally, a transient slat bridge between the amide moiety of Lys-40 and the carboxyl terminus of the
backbone can be observed in several structures of the NMR ensemble.

The protein presents an inner cavity, which has been measured at 341 Å³. All helices present hydrophobic side
chains directed toward the cavity. The phospholipid is found in this cavity. Only one unique phospholipid
position is observed in the cavity, for all the 10 retained structures. The fatty acid chain is completely embedded
in the protein structure. Its axis is aligned with the axis of the tertiary super-helix, and is orthogonal to the [4]-
helix axes (See figure 4c). The terminal methyl group is positioned between the H1 and H4 helices. The fatty
acid chain is inserted in the cavity constituted by the hydrophobic residues (Leu-7, Leu-28, Phe-35, Tyr-38,
Tyr-44, Tyr-47, Ile-48, Ala-53, Leu-57, Val-64, and Pro-65.). The chain presents a turn near carbon C8 and
exits the cavity in a cleft between helices H4, H5 and the C-terminal residue. The inner glycerol moiety is
located on a pocket on the surface of the protein. This pocket presents a basic environment constituted by the
Arg-49, Arg-54 and His-66 side chains. These basic residues are observed in close proximity with the phosphate
moiety, equilibrating the phosphate charge (see figure 5). The cavity of ns-LTP2 is asymmetric. The proximal
entrance of the cavity, where the phosphate group is found, presents several hydrophilic and basic groups : Arg-
49, Arg-54, Thr-58 and His 66. The distal opening of the cavity is characterized with hydrophobic residues,
such as Leu-7, Tyr-38, Tyr-44 and Tyr-47.
This location is in good agreement with the following spectroscopic observations: In the fatty acid chain only three unambiguous Noe inter-molecular contacts have been identified. Four Noe contacts connect the LPG terminal methyl groups (C16) with the Tyr-44 and Tyr-47 aromatic part. The other Noe connects the one of the methylenes of the fatty acid chain with the H[4] of Phe-35. All these residues are found in the hydrophobic cavity. No Noe contacts were found between the protein and LPG glycerol moiety. The chemical shifts of the phospholipid glycerol moiety do not present much shift upon complexation with the protein. On the other hand, the fatty acyl chain chemical shifts differences between isolated and liganded LPG are important.

The structure of the minor form present in solution has not been studied, even though it appears to be structured. About one third of the amino acids seem to be involved in the conformation equilibrium; no attempt was made to assign the residues involved in this equilibrium. Previous dynamic study (58) has shown that the major form is predominantly rigid, with S2 ranging from 0.8 to 0.9, except for Lys-19, exposed to the solvent.

**ns-LTP2 vs ns-LTP1 Comparison**

The three-dimensional structure is known for several ns-LTP1 from different species (19-29). They are very similar among plant and consist in 4 [i]-helices organized in a super helix structure and connected by 4 disulfide bridges. We compared the structure we present here with wheat ns-LTP1 (PDB # 1gh1) (19) using the VMD software (60) (figure 6).

The best superposition is obtained when the H1, H2 and H5 (residues 3-16; 22-32; 50-63) helices in ns-LTP2 are superimposed with the H1, H2 and H4 helices (residues 5-8,10-19; 23-35; 63-76) in ns-LTP1. This gives a root-mean-square deviation of 1.88 Å (backbone only). The global structure is similar and consists in an hydrophobic cavity structure adapted to the lipid transport. Nevertheless, large 3D structural differences between ns-LTP1 and ns-LTP2 are observed in two regions: the H3 and H4 (respectively H3) [i]-helices and the C-terminal region. The last turn of the H2 helix is also different in both structures.

The two short H3 and H4 [i]-helices orthogonal to each other in ns-LTP2 are replaced by the longer H3 helix in ns-LTP1. ns-LTP1 H3 helix is oriented differently and flanked by two large loops presenting no regular secondary structure. This structural difference opens up the cleft corresponding to the entrance of the cavity in which the ligand is located. The C-terminus of ns-LTP1 presents a large loop with no secondary structure. The corresponding zone is deleted in ns-LTP2, thus releasing access to the hydrophobic cavity. This deletion has
globally no impact on the position of the disulfide bridge involving the last cysteine (C67) which is conserved. As a consequence of these structural changes, the wheat ns-LTP2 is smaller and more globular than ns-LTP1.

All cysteine residues are involved in disulfide bridges, but the wheat ns-LTP2 differs from ns-LTP1 in the way the connections are made between the cysteines. ns-LTP1 exhibit a conserved CNC motif located in helix 3, with N being an hydrophilic residue. The two cysteines flanking this hydrophilic residue are linked to cysteines at distal positions in a crossed scheme (as showed in figure 1). At the opposite, in the wheat ns-LTP2, an hydrophobic residue (Phe-35) is present and the two flanking cysteines are involved in disulfide bridges in a non-crossed scheme. The hydrophobic side chain of Phe-35 is displayed on the cavity surface of the wheat ns-LTP2 whereas, in ns-LTP1 the corresponding hydrophilic residue is exposed to the solvent, due to a 180° rotation of the entire helix 3 along its main axis.

This cysteine linkage pattern is not unique to ns-LTP2 and is observed in the soybean hydrophobic protein. Searching structural databases (40) for a protein presenting a similar structural arrangement to ns-LTP2, the soybean hydrophobic protein (PDB # 1hyp) is found as the best match. This protein shares some structural similarity with ns-LTPs but does not transfer lipids. It presents an analogous organisation in terms of helix number and orientation, as well as disulfide bridge topology. Helix 3 is involved in a non-crossed disulfide bridges scheme and appears not to be amphiphilic when compared with ns-LTP2. This could explain why soybean hydrophobic protein does not exhibit lipid transfer capability.

The presence of an hydrophobic cavity is a characteristic of ns-LTPs, the hydrophobic ligands are bound in this cavity in a rather non-specific manner. No major differences have been observed between the structure of the free and palmitate complexed maize (21) and barley (25) ns-LTP1 and free and prostaglandin B2-ligated wheat ns-LTP1 (29) while large conformational changes have been seen for barley ns-LTP1 when it complexes with palmitoyl coA (25). Orientation of the lipid within the hydrophobic cavity was found to be opposite in maize (21) and barley (26) liganded structures whereas wheat ns-LTP1 is able to bind two monoacylated lipids insert head to tail in the hydrophobic cavity (28).

Measured volumes of this hydrophobic cavity are highly variable in ns-LTP1 (23), with typical values ranging from 150 Å³ to 580 Å³. However barley ns-LTP1 exhibits a large volume change upon palmitoyl-coA binding, with a measured cavity volume increasing from 39 to 620 Å³. The volume observed for the wheat ns-LTP2 appears to be in the same range, even though the protein is smaller by 24 amino acids.
The orientation of the phospholipid main-chain, observed in ns-LTP2, is roughly orthogonal to the $\alpha$-helix axes, and the chain runs from helix H1 to helix H5. This is in contrast with most described ns-LTP1 cavities found with a main axis parallel to the $\alpha$-helix axes.

**Conclusion**

The present work presents the refined structure of the wheat ns-LTP2 protein, liganded with L$\alpha$-palmitoylphosphatidylglycerol, as determined by NMR spectroscopy.

The protein was observed as being composed of five helices, structured as a right super-helix, surrounding the phospholipid. This structure presents some homologies with other Lipid Transfer Protein such as ns-LTP1, however the phospholipid was found in a quite different location than in most LTP1s. An hydrophobic cavity was also observed, with a volume equivalent to the one found in ns-LTP1, but with a different geometry.

The ns-LTP2 protein structure presents also homology with the one of the soybean hydrophobic protein, which does not exhibit any lipid transfer activity. This permits to devise a protein family, encompassing ns-LTP1 and ns-LTP2, but also soybean hydrophobic protein and other related vegetal proteins, based on structural homologies, rather than on function of primary sequence homologies.

Extensive ligand screening had to be undertaken in order to find a set of conditions allowing structural study. This indicates that even if the ns-LTP2 is able to adapt a large number of hydrophobic ligands in its hydrophobic pocket, it certainly presents varying affinities depending on the nature of the ligand. The better affinity for single chain phospholipid, negatively charged, can be explained *a posteriori*, by the size of the pocket, and by the patch of positively charged residues located around its entrance.

Further comparative studies on this protein and on homologous proteins will have to be undertaken to improve the understanding on the phospholipid binding affinity, and of the phospholipid transfer activity.

Finally it remains to elucidate how the structural differences observed in the plant LTP family might be related to the various *in vivo* activities putatively assigned to its members. Comparative structural and dynamical studies on several isoforms of wheat ns-LTPs are currently in progress in our group. This complementary work should permit to draw stronger links between structural features and biological functions.
Acknowledgments:

The authors acknowledge N. Sommerer for MS analysis, C. Royer, C. Blanchard, and F. Vagner for helping in the fluorescence and HSQC screening experiments, and the VMD team for efficient support.
References


Figure legends

Figure 1
Primary sequence alignment of various ns-LTP1 and ns-LTP2. Protein are detailed by the Swiss-Prot accession number, or the PDB accession number when available. Disulfide bridges are shown with black lines and numbered according to wheat ns-LTPs.

Figure 2
HSQC spectrum of $^{15}$N labeled recombinant wheat ns-LTP2. The spectrum has been recorded at 295.2K on a Bruker AMX 600, on a 2.8 mM sample complexed with 4.2 mM of L-$\alpha$-palmitoylphosphatidyglycerol in 70 mM phosphate buffer at pH 3.5. Assignment is indicated by numbering the peaks with the id of each residue. Unassigned peaks are indicated with a ? mark.

Figure 3
Synthetic plot showing the elements used for the structure determination.

Figure 4
a) Stereo view of the best 10 NMR structures of wheat ns-LTP2. The phospholipid is not shown for clarity
b) mean NMR structure, with the ensemble of the phospholipid structure as observed in the 10 best protein generated structures. The orientation is rotated by 90° along OX axis with respect with figure 4a. Basic residues around the phosphate group are marked in green.

c) Schematic view of the structure of ns-LTP2 with a CPK model of the liganded phospholipid. Orientation is the same as in Figure 4b.

Figure 5
CPK representation of the ns-LTP2 structure. Positively charged residues are colored in blue and negatively charged residues are colored in red, the phospholipid is represented in licorice mode. It can be seen how the terminal glycerol group sticks out of the molecular surface, and how the basic residues surround the phosphate group (in yellow). Orientation is similar to Figure 4a.
Figure 6

Structural comparison of wheat ns-LTP2 (red) with barley ns-LTP1 (PDB # 1gh1) (blue). Orientation is the same as in Figure 4a, only ns-LTP2 C-term and N-term are labeled. Superposition of helices 1 2 and 5 (ns-LPT2 numbering) with equivalent helices in ns-LTP1 can clearly been observed. Helices 3 and 4, with no equivalent in ns-LTP1 are also observed in front.
Table 1

NMR restraints and structural statistics of the ensemble calculated for LPG liganded wheat ns-LTP2 (10 structures)

Distance & angles restraints
- Noe Intraresidual: 403
- Noe Sequential \((i,i+1)\): 188
- Noe Medium range \((i-j)\leq 3\): 72
- Noe Long range \((i-j)> 3\): 95
- Noe Intermolecular: 5
- S-S bridges: 4
- Dihedral angle \(\phi\): 29
- Dihedral angle \(\psi\): 8
- Total: 800

Deviation from idealized geometry
- Bonds (Å): 0.00323 +/- 0.00013
- Angles: 0.372 +/- 0.012
- Improper angles: 0.187 +/- 0.017
- Dihedral angles: 30.16 +/- 3.04

Deviation from restraints
- Noe restraints: 0.029 +/- 0.001
- Dihedral restraints: 0.45 +/- 0.1

Ramachandran plot (see ref (52))
- Most favored regions: 74.2 %
- Allowed regions: 23.5 %
- Generously allowed regions: 2.4 %
- Disallowed regions: 0.0 %

Mean energies (kcal / mol)
- \(E_{bon}\): 10.29 +/- 0.61
- \(E_{ang}\): 37.14 +/- 2.5
- \(E_{imp}\): 2.67 +/- 0.54
- \(E_{vdw}\): 23.5 +/- 3.8
- \(E_{noe}\): 48.5 +/- 3.93
- \(E_{cdih}\): 0.94 +/- 0.39

RMSD from average structure (Å) (residues 2-67)
- Backbone: 0.67
- Heavy atom: 0.90
Figure 1

LTP type 2
Wheat (AJ297768) - ACGASQ---LVCASAILS-GAXPSEGCCCNLR---------AQQ-GCFQYAKDPYQQVIRSPARLIT/TSCG---LAVP--H---C-----
Barley (P20145) - ACEFAQ---LAVCASAIGL-GTXPSEGCCCNLR---------AQQ-GLCLQYDKPRYRYGVSPPARLITLSCG---IPV--H---C-----
Barley (O81135) - APECVQIQ---LTVCMPAIIG-TGAXPSEGCCCNLR---------AQQ-GLCFQYAKDPILGYRITSPARLITLSCG---LAVP--R---C-----
Rice (Q40631) - ASCNAQQ---LTVCANAIAG-GARPAAACCSSLR---------AQQ-GCFQFAPKDPYRYVYVNNAPRIFTVSGC---IALP--T---C-----
Mustard (O64431) - ACDFKOQ---LQFCCLAAITG-GDPFSGCCCAKL---------EQQ-FLCQFGSKNPAPAQYISSPAVITACG---IFPP--S---C-----
A.thaliana (Q9FF40) - VTCDATQ---LSCQVTAVST-GAFSTDDCCCGKL---------EH=--TQCTTYQNPILFVTSSPARKLTAACD---VAP--T---C-----

LTP type 1
Wheat (1BJW) - IDCQVDSVRPCLSYVQQ-GPGPSQGDCGKLNQANSQRRASACMLKDIARGIINH-LNEDNARSIPKCVNLPYTISLIDCSKV----
Barley (1JTB) - LINCQVDSVRPCLSYVQQ-GPGPSQGDCGKLNQANSQRRASACMLKDIARGIINH-LNEDNARLIPKCVNLPYTISLIDCSKV----
Rice (1IRL) - ITCQVDSVRPCLSYVQQ-GPGPSQGDCGKLNQANSQRRASACMLKDIARGIINH-LNEDNARSIPKCVNLPYTISLIDCSKV----
Maize (1MUM) - ISCCQVSAGITPCLYVQQ-GPGPSQGDCGKLNQANSQRRASACMLKDIARGIINH-LNEDNARSIPKCVNLPYTISLIDCSKV----
Radish (P29420) - LCQVDSVRPCLSYVQQ-GPGPSQGDCGKLNQANSQRRASACMLKDIARGIINH-LNEDNARSIPKCVNLPYTISLIDCSKV----
Sorghum (Q43194) - VTCQVDSVRPCLSYVQQ-GPGPSQGDCGKLNQANSQRRASACMLKDIARGIINH-LNEDNARSIPKCVNLPYTISLIDCSKV----
Chickpea (O23758) - VTCQVDSVRPCLSYVQQ-GPGPSQGDCGKLNQANSQRRASACMLKDIARGIINH-LNEDNARSIPKCVNLPYTISLIDCSKV----

Figure 1
Figure 4b
Refined solution structure of a liganded type 2 wheat non-specific lipid transfer protein
Jean-Luc Pons, Frédéric de Lamotte, Marie-Françoise Gautier and Marc A. Delsuc
J. Biol. Chem. published online January 13, 2003

Access the most updated version of this article at doi: 10.1074/jbc.M211683200

Alerts:
  • When this article is cited
  • When a correction for this article is posted

Click here to choose from all of JBC's e-mail alerts