Refined solution structure of a liganded type 2 wheat nonspecific lipid transfer protein

Jean-Luc Pons, Frédéric de Lamotte, Marie Francoise M. F. Gautier, Marc-André Delsuc

To cite this version:
Jean-Luc Pons, Frédéric de Lamotte, Marie Francoise M. F. Gautier, Marc-André Delsuc. Refined solution structure of a liganded type 2 wheat nonspecific lipid transfer protein. Journal of Biological Chemistry, American Society for Biochemistry and Molecular Biology, 2003, 278 (16), pp.14249-14256. 10.1074/jbc.M211683200 . hal-02675602

HAL Id: hal-02675602
https://hal.inrae.fr/hal-02675602
Submitted on 31 May 2020

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L’archive ouverte pluridisciplinaire HAL, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d’enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.
The refined structure of a wheat type 2 nonspecific lipid transfer protein (ns-LTP2) liganded with \( \text{L-}\alpha\text{-palmitoylphosphatidylglycerol} \) has been determined by NMR. The \( ^{15}\text{N} \)-labeled protein was produced in \( \text{Pichia pastoris} \). Physicochemical conditions and ligandation were intensively screened to obtain the best NMR spectra quality. This ns-LTP2 is a 67-residue globular protein with a diameter of about 30 Å. The structure is composed of five helices forming a right superhelix. The protein presents an inner cavity, which has been measured at 341 Å\(^3\). All of the 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 ns-LTP2 with a diameter of about 30 Å. The structure is composed of five helices forming a right superhelix. The protein presents an inner cavity, which has been measured at 341 Å\(^3\). All of the 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 \( \text{L-}\alpha\text{-palmitoylphosphatidylglycerol} \) glycerol moiety being located on a positively charged pocket on the surface of the protein. The superhelix structure of the protein is coiled around the fatty acid chain. The overall structure shows similarities with ns-LTP1. Nevertheless, large three-dimensional structural discrepancies are observed for the H3 and H4 \( \alpha\text{-helices} \), 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 the fact that ns-LTP2 is shorter by 24 residues.

Plant nonspecific lipid transfer proteins (ns-LTPs) were first isolated from spinach leaves and named based on their ability to mediate \textit{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 Refs. 2 and 3 for review). Both families are multi-geneic, and more than 150 sequences of plant ns-LTPs are listed in data bases. Only a limited number of proteins have been isolated from plant, and \textit{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 the \textit{Zinnia elegans} cell differentiation process (9, 10), in barley and rice developing seeds (11, 12), under abiotic stress conditions in barley roots (12), or during nodulation in \textit{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), whereas ns-LTP2 has been reported only as a potent allergen of the pollen of \textit{Brassica rapa} (18).

The three-dimensional structure of four cereal ns-LTP1s 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 of 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-\( \alpha\text{-helix} \) 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 that 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-LTP1s have been studied more extensively than ns-LTP2s that are distinct in terms of primary sequence with less than 30% homology (Fig. 1), size (7 kDa \textit{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-cysteine 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-LTP2s have a fold that is similar to or different from that of ns-LTP1
**Primary sequence alignment of various ns-LTP1 and ns-LTP2.** The proteins are detailed by the Swiss–Prot accession number or the Protein Data Bank accession number when available. The disulfide bridges are shown with black lines containing 0.1% trifluoroacetic acid. A 0.5–H9251 was mixed with 0.5–H9262 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 15N HSQC NMR (34). A nonlabeled sample was also produced as described earlier (35).

**EXPERIMENTAL PROCEDURES**

Production/Purification—The Pichia pastoris transformant GS115–Purification transformant GS115-Pichia pastoris was prepared in solution in ethanol at a concentration of 10 mM or in 300 mg of protein. A 0.5–H9251 was solubilized in 10

---

[Fig. 1. Primary sequence alignment of various ns-LTP1 and ns-LTP2. This figure shows the alignment of primary sequences of various ns-LTP1 and ns-LTP2 proteins. The sequences are aligned using the Clustal W algorithm. The alignment includes the primary sequence of wheat ns-LTP2 with reference to its Swiss-Prot accession number (P38433) and Protein Data Bank accession number (1L6H). The alignment highlights the conservation of disulfide bridges and other structural elements among the sequences.]
NMR Spectra Assignment—The assignment of the wheat ns-LTP2 complexed with the phospholipid was performed from the set of three-dimensional (HSQC-NOSY and HSQC-TOCSY) and two-dimensional experiments (COSY, TOCSY, and 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 of the amide peaks could be found. All of the side chain labile hydrogens from Lys, Asn, Arg, and Gln residues were also assigned. Additional peaks belonging to the glycansylated protein were sorted out in the HSQC spectrum from their reduced intensity. On the basis of sequential dNN(i,j+1) and doNN(i,j+1) all of the amino acids have been found, and 98% of all the nonlabile 1H chemical shifts have been assigned. The natural abundance 13C HSQC spectrum was used to help the assignment of side chain methylenes. Stereospecific assignment of β 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 caused by 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 13N chemical shifts of the wheat ns-LTP2 complexed with LPG has been deposited in the BioMagResBank (number 4977).

NMR Structure Calculation and Analysis—The two-dimensional and three-dimensional NOESY experiments, performed at three different temperatures (295.2, 310, and 323 K), have been used to extract all of the nonlabile 1H distances; pseudo-atom corrections were used. Additional constraints have been found, and 98% of all the nonlabile 1H NOE spectra caused by 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 13N chemical shifts of the wheat ns-LTP2 complexed with LPG has been deposited in the BioMagResBank (number 4977).

The dynamical annealing protocol (anneal.inp) of the crystallography NMR (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, because all of the sulfur atoms were located in compatible distances. The observed disulfide bridges are: Cys28/Cys69, Cys30/Cys49, Cys28/Cys54, and Cys36/Cys67. Hβ-Hβ NOE contacts were further observed in the NOEY 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 Hα and Hβ of Tyr-44 and the middle of the lipidic chain (C8) with H3 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 (code 1N89).

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

RESULTS AND DISCUSSION

Production and Purification—The recombinant protein preparations were obtained as previously described (34–36). The N-terminal sequence (ACQASLQLVC) 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 15N-labeled protein. A 7055-Da average molecular mass was measured, confirming a 15N isotopic labeling over 95%. The mass spectrum also revealed that approximately 15% of the protein has been glycosylated with one to four C6 sugar moieties, assumed to be mannose (162 Da). This is consistent with the fact that P. pastoris adds O-glycosyltransferases composed solely of mannose residues. Peptide mass fingerprinting 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-Xaa-Ser/Thr consensus sequence, whereas O-glycosylation requires the presence of a Ser or Thr residue. The wheat ns-LTP2 does not contain the Asn-Xaa-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 that 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 of these results indicated that a small fraction of the wheat ns-LTP2 is expressed as an O-glycosylated protein. To our knowledge, no glycosylation has been reported for ns-LTP purified from plant.

Physicochemical Context Screening and Ligand Choice—Intensive screening of the physicochemical solution conditions was performed 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 negatively charged phospholipids. We finally selected LPG with an average chain length (C16). The titration of the 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 two-dimensional HSQC spectrum. This evolu-

Assessments and Secondary Structure Elements—The assignment of the protein resonances was performed from the set of \(^{15}\text{N}\)-edited NOESY and TOCSY spectra. The sequential strategy was used, aided by the \(^{15}\text{N}\) and natural abundance \(^{13}\text{C}\) HSQC spectra (Fig. 3).

The solvent protection experiments indicate that the secondary structure is quite strongly established, because about 40% of the amide protons remain unexchanged during the first hours of the \(\text{D}_2\text{O}\) exchange experiment at 1 °C, and eight 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 two 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 eight cysteines of the protein are engaged in disulfide bridges. We have not been able to unambiguously assign the HJ-HJ NOE contacts characteristic of this structure because of the crowding of this spectral region. However, after the first run of structure generation, all of the obtained structures exhibited side chain proximities permitting the disulfide skeleton based on these prestrutures 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 \(^{1}\text{H}\) assignments of the LPG in the complexed form were obtained from two-dimensional 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 of the residues are localized in the allowed regions of the Ramachandran plot. The ensemble of the best structures present a root mean square deviation computed on residues 2–67 of 0.9 Å for all of the heavy atoms and a root mean square deviation of 0.67 Å for the backbone atoms. All of the statistics of the geometrical constraints and the structure reconstruction are given in Table I. The set of solution structures of ns-LTP2 liganded with LPG as obtained from this experimental work is presented in Fig. 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 superhelix tertiary structure. Helix 1 is a 3–10 helix, encompassing residues 7–15. All of the other helices are \(\alpha\)-helices. Helix 2 includes residues 22–31, helix 3 includes residues 34–40, helix 4 includes residues 44–49, and helix 5 includes residues 51–60. The overall fold is a right superhelix. 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: residues 14 and 15, residues 26–31,
residues 38–41, residues 47–50, and residues 56–64. The LPG molecule is found partly embedded in the structure of the protein, with the superhelix 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 3–5 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 Cys-36 are respectively engaged in disulfide bridges with Cys-2 and Cys-67, forming two diametrically opposed bonds relative to the helix axis. All of the prolines are observed in trans conformation, as confirmed by the Hα(i)–Hα(i + 1) contacts observed in the NOESY spectra, for all of 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-terminus 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 γ turn is observed between helices 2 and 3, corresponding to residues Gln-31, Gln-32, and Gly-33. A type-1 β turn is observed between helices 3 and 4, corresponding to residues Asp-41, Pro-42, Thr-43, and Tyr-44, with the acidic head of Asp-41 engaged in a hydrogen bond with the HN of Thr-43.

Three classic helix cappings can also be observed: the N-capping of helix 2 with the OH of Ser-21 hydrogen-bonded to the HN of Glu-23; the C-capping of helix 2 with the side chain of Gln-32 bonded to the CO of Arg-29; and the N-capping of helix 5, with the OH of Ser-50 bonded to the HN of His-52. Finally, a transient slat bridge between the amide moiety of Lys-40 and the C 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 of the 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 10 retained structures. The fatty acid chain is completely embedded in the protein structure. Its axis is aligned with the axis of the tertiary superhelix and is orthogonal to the α-helix axes (Fig. 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 8 and exits the cavity in a cleft between helices H4 and 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 (Fig. 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 intermolecular 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 Hβ of Phe-35. All of 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 shift 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 S^2 ranging from 0.8 to 0.9, except for Lys-19, exposed to the solvent.

ns-LTP2 versus 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 four α-helices organized in a superhelix structure and connected by four disulfide bridges. We compared the structure we present here with wheat ns-LTP1 (Protein Data Bank code 1gh1) (19) using the Visual Molecular Dynamics software (46) (Fig. 6).

The best superposition is obtained when the H1, H2, and H5 (residues 3–16, 22–32, and 50–63) helices in ns-LTP2 are superimposed with the H1, H2, and H4 helices (residues 5–8, 10–
The orientation of the phospholipid main chain, observed in ns-LTP2, is roughly orthogonal to the α-helix axes, and the chain runs from helices H1 to H5. This is in contrast with most described ns-LTP1 cavities found with a main axis parallel to the α-helix axes.

**Conclusion**—The present work presents the refined structure of the wheat ns-LTP2 protein, liganded with 1-α-palmitooylphosphatidylglycerol, as determined by NMR spectroscopy. The protein was observed as being composed of five helices, structured as a right superhelix, surrounding the phospholipid. This structure presents some homologies with other lipid transfer proteins 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 also presents homology with the one soy bean hydrophobic protein that does not exhibit any lipid transfer activity. This permits us to devise a protein family encompassing ns-LTP1 and ns-LTP2 but also soy bean hydrophobic protein and other related vegetal proteins, based on structural homologies, rather than on the function of primary sequence homologies.

Extensive ligand screening had to be undertaken 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 be explained 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 us to draw stronger links between structural features and biological functions.

**Acknowledgments**—We acknowledge N. Sommerer for mass spectrometry 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**