

Solution structure of a lipid transfer protein extracted from rice seeds Comparison with homologous proteins

Jarek Poznanski¹, Patrick Sodano², Se Won Suh³, Jae Young Lee³, Marius Ptak² and Françoise Vovelle²

¹Institute of Biochemistry and Biophysics, Polish Academy of Sciences, Warsaw, Poland; ²Centre de Biophysique Moléculaire, Orléans, France;

³Département de Chimie, Collège of Natural Sciences, Seoul National University, Korea

Nuclear magnetic resonance (NMR) spectroscopy was used to determine the three dimensional structure of rice nonspecific lipid transfer protein (ns-LTP), a 91 amino acid residue protein belonging to the broad family of plant ns-LTP. Sequence specific assignment was obtained for all but three HN backbone ¹H resonances and for more than 95% of the ¹H side-chain resonances using a combination of ¹H 2D NOESY; TOCSY and COSY experiments at 293 K. The structure was calculated on the basis of four disulfide bridge restraints, 1259 distance constraints derived from ¹H-¹H Overhauser effects, 72 ϕ angle restraints and 32 hydrogen-bond restraints. The final solution structure involves four helices (H1: Cys3-Arg18, H2: Ala25-Ala37, H3: Thr41-Ala54 and H4: Ala66-Cys73) followed by a long C-terminal tail (T) with no observable regular structure. N-capping residues (Thr2, Ser24, Thr40), whose side-chain oxygen atoms are involved in hydrogen bonds with *i* + 3 amide proton additionally stabilize the N termini of the first three helices. The fourth helix involving Pro residues display a mixture of α and 3_{10} conformation. The rms deviation of 14 final structures with respect to the average structure is 1.14 ± 0.16 Å for all heavy atoms (C, N, O and S) and 0.72 ± 0.01 Å for the backbone atoms. The global fold of rice ns-LTP is close to the previously published structures of wheat, barley and maize ns-LTPs exhibiting nearly identical pattern of the numerous sequence specific interactions. As reported previously for different four-helix topology proteins, hydrophobic, hydrogen bonding and electrostatic mechanisms of fold stabilization were found for the rice ns-LTP. The sequential alignment of 36 ns-LTP primary structures strongly suggests that there is a uniform pattern of specific long-range interactions (in terms of sequence), which stabilize the fold of all plant ns-LTPs.

Keywords: lipid transfer protein; NMR; rice; structure.

Lipid transfer protein (LTP) extracted from rice seeds is a 91 amino acid residue protein. It belongs to the broad family of plant LTPs characterized by their ability to catalyze the exchange of a large number of lipid molecules between natural and artificial membranes *in vitro*. Due to the absence of substrate specificity [1,2], they are currently named nonspecific lipid transfer proteins (ns-LTPs). Three-dimensional solution structures have been recently determined for ns-LTPs extracted from wheat [3], maize [4] and barley [5] seeds, which exhibit a remarkable homogeneity of structure and define a new class of folds formed by four helices and a C-terminal fragment wound around an elongated cavity. Global folds of the maize ns-LTP in solution and in crystal [6] are very similar except for local differences related either to contacts with crystal lattice or to the internal mobility of the protein [7]. Recently the 1.6 Å resolution crystallographic structure of rice ns-LTP was resolved [8]. Crystallographic structure of maize ns-LTP complexed with palmitate [6] as well as solution structures of 1 : 1 complexes of barley ns-LTP with palmitoyl CoA [9], maize ns-LTP with 1-palmitoyl-2-lyso-phosphatidylcholine [4] and wheat ns-LTP

with 1,2-dimyristoylphosphatidyl-glycerol [10] show that the hydrophobic tails of fatty acids are inserted in the protein internal cavity, suggesting that such ns-LTPs could work as carriers.

Numerous other proteins capable of transferring lipids between membranes have been isolated from different animals, fungi and bacteria. Compared with plant ns-LTPs, these proteins have higher molecular masses ranging approximately from 15 kDa to more than 30 kDa. They exhibit little sequence homology and, according to preliminary results published on a sterol binding protein [11,12], on the NMR solution structure of human sterol carrier protein 2 [13] and on circular dichroism investigations of rabbit [14] and human [15] plasma lipid transfer protein, their three-dimensional structures could be different.

Several cDNAs encoding three different LTPs have been isolated from rice [16,17]. The corresponding genes belong to a multigene family and are subjected to a complex regulation as other cereal LTP genes [16]. However, the biological function of this protein family remains in debate. Their common property to catalyze transfer of lipids *in vitro* has not yet been clearly related to a similar activity *in vivo*. All known plant LTPs are synthesized with N-terminal extensions which facilitate their translocation through membranes [18,19]. They are located mainly outside the cells and thus cannot participate significantly to the cytoplasmic lipid traffic. Several compounds including abscisic acid modulate LTP gene expression [16]. According to Hollenbach [20], the primary leaves of cadmium-exposed seedlings have high contents of abscisic acid and a thicker wax layer of

Correspondence to F. Vovelle, Centre de Biophysique Moléculaire, CNRS-UPR 4301 affiliated to the University of Orléans, rue Charles Sadron, 45071 Orléans cedex 02, France. Tel.: + 33 138255574, Fax: +33 138631517, E-mail: vovelle@cnrs-orleans.fr

Abbreviations: ns-LTP, nonspecific lipid transfer protein; 1D, one-dimensional; 2D, two-dimensional.

(Received 7 September 1998, revised 2 November 1998, accepted 3 November 1998)

the cuticle. A role for plant LTPs was proposed by Sterk [21] in the transport of cutin monomers through the extracellular matrix to the sites of synthesis of the cutin layer. Several genes have been cloned and characterized in cotton fibres which are expressed in a temporal manner. The genetic diversity of plant LTPs could indicate that these proteins have several functions, depending on their location, on the age of tissues and on stress effects. Several LTP-like proteins were shown to have antifungal activities [22,23] and therefore, were proposed to be involved in the defence of plants against various pathogens. Recently, we have resolved the solution structure of *Ace*-AMP1 [24], an antifungal protein extracted from onion seeds. This structure is very similar to those of ns-LTPs except that the internal cavity is obstructed by several aromatic side-chains. It is probably why this protein is unable to transfer lipids.

Although plant ns-LTPs exhibit very similar global folds, their lipid binding capacity and more generally their interactions with lipid structures depend on their origin. Depending on the nature of amino acid side-chains pointing in the internal cavity and on tenuous variations in the organization of secondary structure elements, the shape and the volume of the internal cavity vary considerably. The capacity of ns-LTPs to extract lipids from micelles or bilayers can also vary from one protein to another [25]. That is the reason why we have determined the solution structure of the rice ns-LTP reported in the present paper. A systematic comparison of three-dimensional structures of various ns-LTPs should contribute to a better understanding of the variability of their properties which could reflect multiple functions as strongly suggested by the existence of multiple genes regulated in various manners.

MATERIALS AND METHODS

NMR experiments

The sample of rice ns-LTP was purified as described previously [26,27]. NMR experiments were performed on a 5-mm sample of rice ns-LTP in a mixture of 90% H₂O/10% ²H₂O, 50 mM phosphate buffer at pH 6.4. All spectra were recorded on a Bruker AMX-500 spectrometer operating at 500 MHz for ¹H nucleus equipped with three-axis field gradient unit. Alternatively States [28] or TPPI [29] method of the quadrature detection of indirect axes were used. All spectra were accumulated with the spectral width 6024 Hz and 1.5 s relaxation delay and then processed with the use of UXNMR program on a X32 workstation. The water signal was suppressed either using the WATERGATE [30] pulse sequence where the selective inversion was generated by 3-9-19 composite pulse [31] or by weak presaturation during relaxation delay [32]. Basic two-dimensional (2D) experiments were recorded at 293 K and 308 K. In addition to P-COSY [33], TOCSY [34,35] and NOESY [36] spectra a double quantum COSY spectrum (2Q-COSY) with 30 ms double quantum excitation time [37] recorded at 298 K was used to help glycine spin system identification. Amide proton exchange rates were estimated at 283 K as follows. The protein sample was lyophilized and redissolved in ²H₂O. Approximately 5 min after dissolution, five one-dimensional (1D) spectra, 4 min each, were recorded. Then three 60 ms mixing time TOCSY spectra with acquisition times of 2 h, 4 h, 12 h, respectively, were accumulated. At the end of acquisition of each 2D spectrum, a 1D spectrum was acquired. Finally, in order to help amino acid assignment, a 200-ms mixing time NOESY spectrum was recorded. In all 2D ²H₂O experiments a weak field presaturation was accompanied by a SCUBA sequence [38]. Additionally, for comparison,

60 ms TOCSY spectrum of the protein in H₂O solution at 283 K was recorded.

Spectrum analysis, assignment and cross-peak volume calculations were performed with XEASY [39] program. Quantitative upper distance limits were obtained from 60 ms mixing time zero-quantum filtered NOESY [40] recorded at 298 K with 650 *t*₁ values 4096 *t*₂ points each and zero-filled to the matrix 4 k × 1 k data points. Qualitative data were derived from 200 ms NOESY at 298 K. Scalar ³*J*_{H_{NH}α} spin coupling constants were determined from zero-quantum filtered NOESY spectrum by the method of inverse Fourier transformation of the INFIT program [41].

The cross-peaks obtained from low mixing time NOESY experiment were converted, after internal calibration, into upper distance limits by the CALIBA program [42]. The limits were restricted to a 2.4–5 Å range. The strongest sequential H_N*i*/H_N*i* + 1 connectivity in helical region was taken as the reference for the 2.8 Å distance. In order to take into account possible spin diffusion effects all 200 ms NOESY cross-peaks, irrespectively of the volume, were converted into an upper limit of 6 Å. Finally, after DIANA [43] filtering, 1259 upper limit distance constraints were conserved –216 intraresidual, 348 sequential, 387 short range, 308 long range. Estimated scalar coupling constant values, accompanied with upper limit constraints were used by HABAS [44] to generate 72 angular restraints for ϕ dihedral angles. Additionally, during the rounds of calculations, three pairs of upper/lower limits for each identified disulfide bridge: 1.9/2.1 Å for Sγ(*i*)-Sγ(*j*) and 2.5/3.2 Å for Cβ(*i*)-Sγ(*j*), respectively, as well as 1.8/2.5 Å restraint for O HN distances for each determined hydrogen bond were progressively introduced.

Structure determination

The three-dimensional solution structure of rice ns-LTP was determined by the standard distance geometry method followed by simulated annealing process accompanied by a final minimization in the presence of NMR constraints. Starting from 200 random structures, the assignment process was self-corrected by repeated rounds of DIANA calculations. Additional restraints were incorporated in successive stages according to the concept of structure aided assignment of ambiguous NOE cross-peaks [44] with the ASNO program [45]. Stereospecific assignment as well as some angular constraints were generated by GLOMSA [42] and HABAS programs on the basis of the set of protein structures, experimental ³*J*_{H_{NH}α} coupling constants and the interproton distance upper limits. Pseudo-atom correction for unassigned stereo partners and magnetically equivalent protons were applied, as proposed by Wüthrich [46]. Afterwards, 57 structures with target function [47] below 1.0 Å² were refined by X-PLOR 3.1 [48] simulated annealing protocol [49] followed by energy minimization in the CHARMM [50] force field. In order to mimic electrostatic screening of charges by the bulk water in the absence of the explicit water molecules, distance dependent dielectric constant ε = 4*r* was used. Finally among over 40 structures exhibiting all peptide bonds in the *trans* configuration, excluding Pro23 assigned as *cis*, 14 lowest energy structures were selected for further analysis. Quality tests as well as determination of secondary structure were performed using PROMOTIF [51], PROCHECK [52] and PROCHECK_NMR [53] packages. Cavity determinations were done by a local program CAVITE [7], chemical shift calculations were performed with TOTAL [54,55], TRIPOS SYBYL package [TRIPOS, St Louis, Missouri] and MOLMOL [56] were used for detailed analysis and graphical presentation. Solvent

accessibility calculations were performed by stochastic program ACCESS [57] parametrized with Kollman united atom radius set [58]. The calculated solvent exposed surfaces of amino acid side-chains were normalized under the assumption that maximal possible values had been obtained by analysis of MD trajectory of the corresponding Ala-X-Ala tripeptides [59]. A number of structure calculation statistics were performed by self-made X-PLOR and SYBYL procedures. Fasta [60] and CLUSTAL W [61] on the www2.ebi.ac.uk web server were used for sequential analysis and multisequence alignment.

RESULTS

Assignment

Almost full assignment of the 2D NMR spectra was performed according to the standard procedure proposed by Wüthrich [62] with the use of COSY, TOCSY and NOESY experiments. Stereospecific assignments covered 19 out of 53 H β , 8(10) glycine H α , 3(4) proline H δ , 3(4) asparagine H δ , 1(1) glutamine H ϵ and 4(5) valine C γ H $_3$, respectively. In order to solve problems of signal overlapping, the sequential assignment pathway was followed on spectra recorded either at 298 K and 308 K. In the fingerprint region of the COSY spectrum a total number of 70 cross-peaks out of 86 expected were identified. Additional cross-peaks, including Leu14, Arg32, Ala42, Ala55, Gly60, Ala63, Ala66 and Thr80 as well as Thr41, Arg44, Arg45, Thr46 and Cys48 were found in the TOCSY spectra. Finally, probably due to the relatively high value of the pH (6.6), the resonances of solvent exposed Cys3 and Gly4 amide protons were not detected

as well as the HN proton of Thr15. All nonexchangeable protons of Ile1, Cys3, Gly4 and Thr15 were assigned with the help of NOESY spectra. According to NOESY spectra, three out of the four proline residues exhibited Pro $_i$ (H δ)–(HN/H α) $_{i-1}$ sequential cross-peaks characteristic of the *trans* isomer of the X-Pro peptide bond whereas, due to the observed Pro23(H α)–Gly22(HN/H α) cross-peaks, Pro23 was assigned as a *cis* Pro. Generally every assignment made at 298 K was verified in details on the spectra recorded at 308 K. The assignments of eight out of the ten glycine resonances were additionally confirmed by a 2Q-COSY experiment, since the glycine remote connectivities resonated in a typical free spectral region. HN resonances of Gly4 and Gly11 were not observed whereas Gly20 and Gly60(δ_{NH} , $\delta_{\text{H}\alpha 1}$ + $\delta_{\text{H}\alpha 2}$) cross-peak intensities occurred at the noise level. The dispersion of the amide ^1H resonances allowed an almost continuous path of sequential NOEs over the whole molecule.

Inspection of cross-peaks of the peptide segment Thr40–Cys48 revealed line splitting caused by the presence of an isoform which made it impossible to measure $^3J_{\text{HNH}\alpha}$ coupling constants for those residues. Sequential assignment pathways of both isoforms showed that the differences arise from a single mutation where Ser40 replaces Thr40. NOESY and TOCSY connectivities suggested that both isoforms were equally populated. These observations were confirmed by mass spectroscopy showing two derived peaks with molecular weight of 8886 Da and 8900 Da. These molecular masses correspond exactly to the expected values calculated from both NMR sequences. After correction for four disulfide bonds molecular mass were estimated to 8887 and 8901 Da, respectively.

Table 1. Structural statistics and restraint violations of the 14 best structures representative of the solution structure of rice ns-LTP.

NOE statistics				
Intraresidual	216			
Sequential	348			
Medium range	387			
Long range	308			
Energy terms (kcal·mol $^{-1}$)	Average		Range	
Total	– 235.9 $\pm$ 14.9		(– 267.1; – 217.4)	
Electrostatic	– 275.9 $\pm$ 8.1		(– 297.1; – 264.6)	
VdW	– 425.2 $\pm$ 8.5		(– 437.0; – 412.4)	
H-bonds	– 143.3 $\pm$ 8.9		(– 163.3; – 130.0)	
Bond	26.4 $\pm$ 0.6		(25.5; 27.5)	
Angular	179.0 $\pm$ 3.9		(174.5; 184.9)	
Dihedral	365.1 $\pm$ 10.9		(345.8; 388.3)	
Improper	3.3 $\pm$ 0.2		(2.9; 3.7)	
NOE constraints	35.3 $\pm$ 9.1		(21.8; 53.9)	
Dihedral constraints	3.4 $\pm$ 0.9		(2.3; 4.9)	
	Number		Value	
	Average	Range	Average	Range
Violations of experimental restraints				
Distance restraints > 0.2 Å	4.1	(1, 7)	0.024	(0.019, 0.029)
Angle restraints > 5 deg.	2.3	(0, 5)	1.72	(1.43, 2.10)
Deviation from idealized geometry				
Bond distance > 0.05 Å	1.4	(0, 3)	0.0123	(0.0119, 0.0126)
Angles > 10°	0.4	(0, 2)	2.88	(2.82, 2.97)
Ramachandran	Average		Range	
Most favorable regions	76 $\pm$ 3%		(73, 81)	
Additional regions	22 $\pm$ 3%		(17, 25)	
General regions	1 $\pm$ 1%		(0, 4)	
Disallowed regions	1 $\pm$ 1%		(0; 1)	

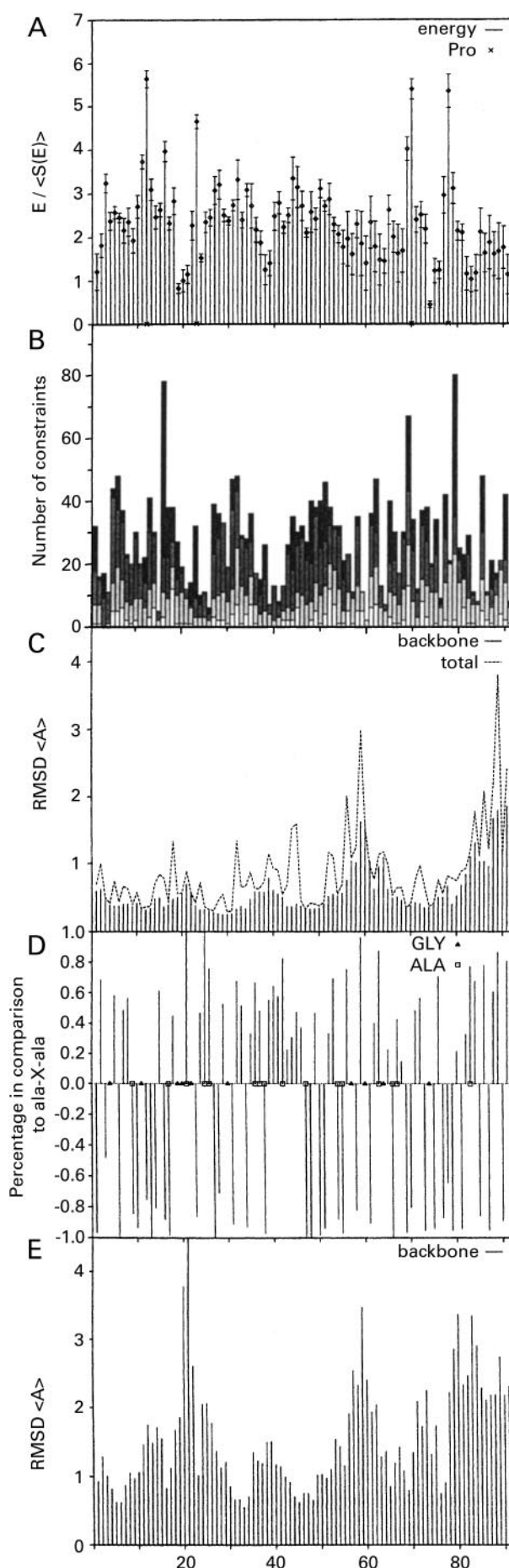
Structure quality analysis

Standard estimation of the structure quality was obtained by analyzing the ϕ - ψ distribution in a Ramachandran plot [63]. 76% of amino acid residues are located in the most favored regions and 22% more cover additional favored regions (Table 1). Details of quality and precision estimations of the solution structures are also summarized in this table. The mean deviation from idealized geometry also remains negligible. Among the analyzed structures no distance violation exceeding 0.3 Å is observed, while no more than five violations per structure are larger than 0.2 Å. The calculated local rms deviation correlates well with the number of experimental NOE constraints per residue (Fig. 1B and C). In both cases, the curve falls down in the central part of the expected helical regions and becomes larger for loop fragments. Additionally, the first part of the C-terminal loop – Ta (tail a – from Gly74 up to Tyr79) is as well defined as helical regions whereas Tb (Thr80–Ser91) exhibits extremely large structural dispersion.

Global fold

The short and medium NOE connectivities as well as amide proton exchange data and the values of $^3J_{\text{HNH}\alpha}$ vicinal coupling constants lead to determine four helical regions (Fig. 2). They are characterized by small coupling constants, slow amide proton exchange rate, strong $\text{HN}_i\text{--HN}_{i+1}$ sequential connectivities and $(\text{HN}, \text{H}\alpha, \text{H}\beta)_i\text{--HN}_{i+3}$, $\text{H}\alpha_i\text{--HN}_{i+4}$ short range NOE. The global fold of rice ns-LTP is a four-helix fold (Fig. 3) stabilized by three interhelical disulfide bonds and a long C-terminal tail (T) connected by the fourth disulfide bridge to H3. The first three helices are nearly parallel, while the fourth is significantly shorter and almost perpendicular to the first one. The helices involve residues Cys3–Arg18 (H1), Ala25–Ala37 (H2), Thr41–Ala54 (H3) and Ala66–Cys73 (H4), respectively. Although H1 formally starts from Cys3, in more than 60% of the final structures, Thr2 is involved in the α helical hydrogen-bonding pattern. Moreover, in 80% of structures Thr2 (O γ) becomes an acceptor of an additional hydrogen bond with the amide proton of Gln5, stabilizing the helical structure. This $i, i+3$ hydrogen-bond interaction was recently reported as a common mechanism of N-capping of helical termini observed in protein structures [64,65]. Analogous hydrogen bonds were found in H2 involving Cys27(HN)–Ser24(O γ) and in H3 between Asp43(HN) and Thr40(O γ 1) (Fig. 4A). Depending on structure, H1 ends up at residue Arg18 or Gly19 and H3 end is located between residues Ala54 and Arg56. Two intrahelical proline residues: Pro12 and Pro70 break the regular $i, i+4$ hydrogen-bonding pattern of the ideal α helical structure. Like in all previously published ns-LTP structures Pro12 produces a 30° bent of H1 at Val10, while Pro70 only disturbs the structure of H4 without detectable change of the helix axis direction. Among

Fig. 1. Amino acid residue contribution to the: penalty function. (A) Proline residues marked on the abscissa by *; (B) number of NOE restraints per residue (intraresidual, sequential, short and long range, respectively); (C) local rms deviation from mean structure calculated for backbone atoms (impulses) and all heavy atoms (broken line); (D) normalized accessibility of the polar side-chains (upper) and normalized protection of the nonpolar side-chains (lower), Gly residues are marked by triangles, whereas Ala residues are marked by squares and assigned as polar or non-polar according to their solvent accessibility; (E) local rms deviation from the crystallographic structure of maize ns-LTP calculated for the backbone atoms of the 14 solution structures of rice ns-LTP.



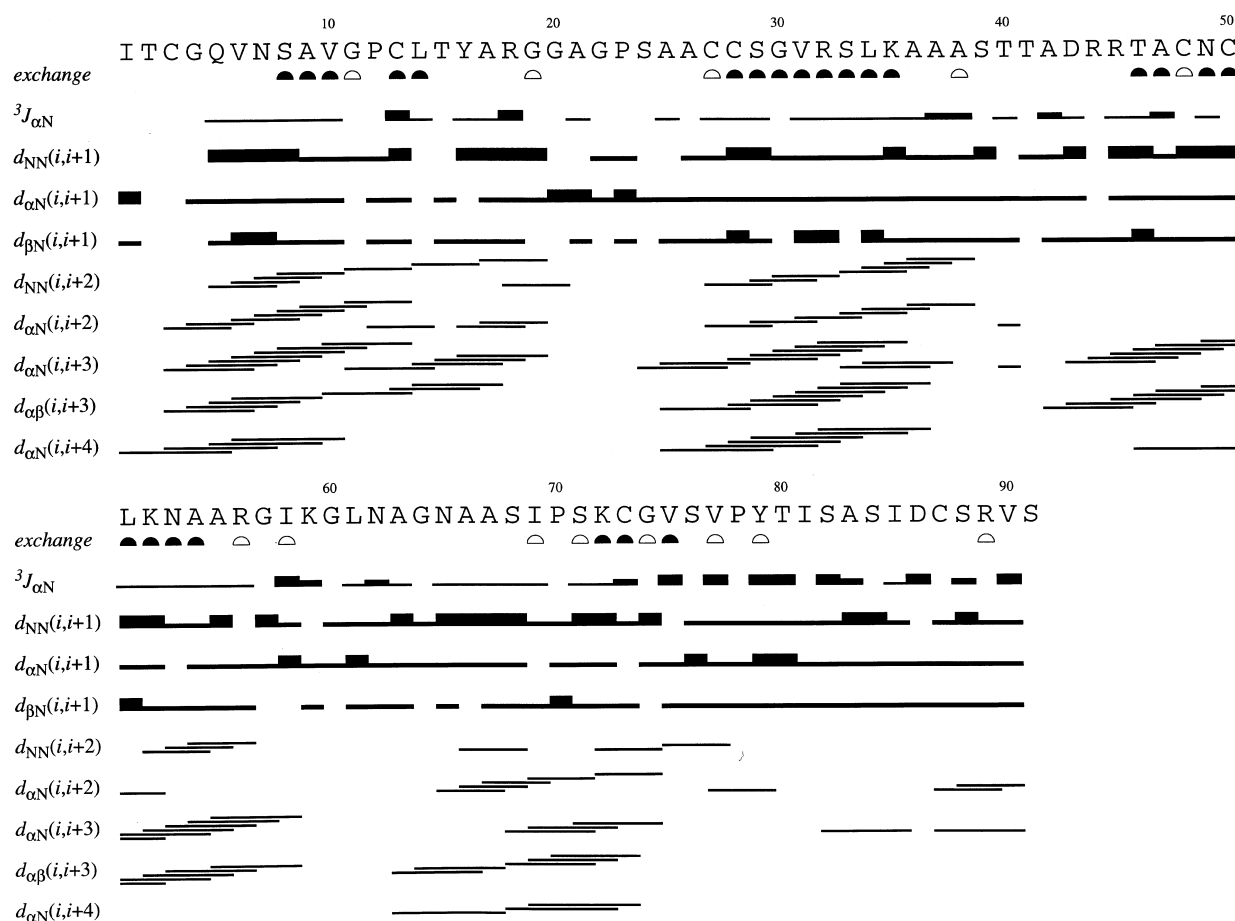


Fig. 2. Overview of the NMR data showing amide proton exchange rates, $^3J_{\text{HNH}\alpha}$ coupling constants, sequential (HN_i , $\text{H}\alpha_i$, $\text{H}\beta_i$)/ HN_{i+1} and medium range connectivities in NOESY spectrum recorded at 298 K. The exchange rates are reported as slow (filled shapes) and medium (empty shapes); the coupling constants are reported as weak (< 7 Hz), medium (7–9 Hz) and large (> 9 Hz) as well as the sequential $i, i+1$ dipolar interactions are reported as weak, medium or strong according to the thickness of line.

the interhelical loops only L3 between H3 and H4 exhibits a recurrent classical type I β -turn conformation between residues Asn62 and Asn65. Moreover, this turn is stabilized by inter-residual Asn62(O δ)-Gly64(HN) hydrogen bond. Another hydrogen bond involving residue side-chain is formed between Tyr16(O) and Gly20(HN). This bond constrains the structure of the first part of L1 (Gly19, Gly20) in the nearly α helical conformation. The long C-terminal tail is organized by three successive β -turns with type: IV, VIII for Val77-Thr80, IV for Ser82-Val85 and Cys87-Val90, respectively. The first turn is stabilized by hydrophobic interaction between the aliphatic side-chains of Val77, Pro78 and Tyr79, whereas the second one is stabilized by the Ser82(O γ)-Ser84(HN) hydrogen bond. In addition to interhelical disulfide bridges, the spatial organization of the helices is constrained by two well-defined long-range hydrogen bonds: Asn62(HN)-Ala17(O) and Asn65(H δ 21)-Tyr16(O). Both maintain the C terminus of H1 close to loop L3 (Fig. 4B).

Disulfide bridges

In the absence of chemical determination, the pairing of the four disulfide bridges was determined by six intercysteine NOEs: SS1: Cys3(H α)-Cys50(H β 1, H β 2), SS2: Cys13(H α , HN)-Cys27(H β 1), SS3: Cys28(H β 2)-Cys73(C β H $_2$) and SS4: Cys48(H β 2)-Cys87(C β H $_2$), respectively. Five additional NOEs: Val6(H α , C γ 1H $_3$, C γ 2H $_3$)-Cys50(H β 1, H β 2) and two

Asn49(H δ 21)-Cys87(H β 1, H β 2) confirm two out of the four pairings. Although the classical disulfide distance restraints [66] deduced from intercysteine NOEs were not introduced into the initial distance constraint set, sulfur atoms of each cysteine pair were found close to each other just after the first round of structural calculations. The obtained pairing entirely agrees with the disulfide pattern described in all known ns-LTP structures as well as with chemically determined cysteine pairing made for castor bean ns-LTP [67]. The disulfide distance restraints were confidently added to the constraint list for all following structure calculations. An analysis of the cysteine side-chain conformations demonstrates that SS1 and SS4 display a left-handed spiral conformation, SS2 a right-handed hairpin, whereas SS3 exhibits a 50 : 50 mixture of a left-handed spiral and a right-handed hairpin conformation.

Protein surface – solvent exposed polar residues

Almost all polar side-chains are solvent exposed and exhibit large conformational dispersion, which is reflected by large values of χ_1 and χ_2 circular variances. However, some of them are involved in local interactions resulting in a reduction of both their conformational flexibility as well as side-chain solvent accessibility (Fig. 1D). For instance the Thr46 side-chain is very well defined due to steric hindrances between its C γ H $_3$ methyl group and the side-chains of residues Ile1 and Asp43. This

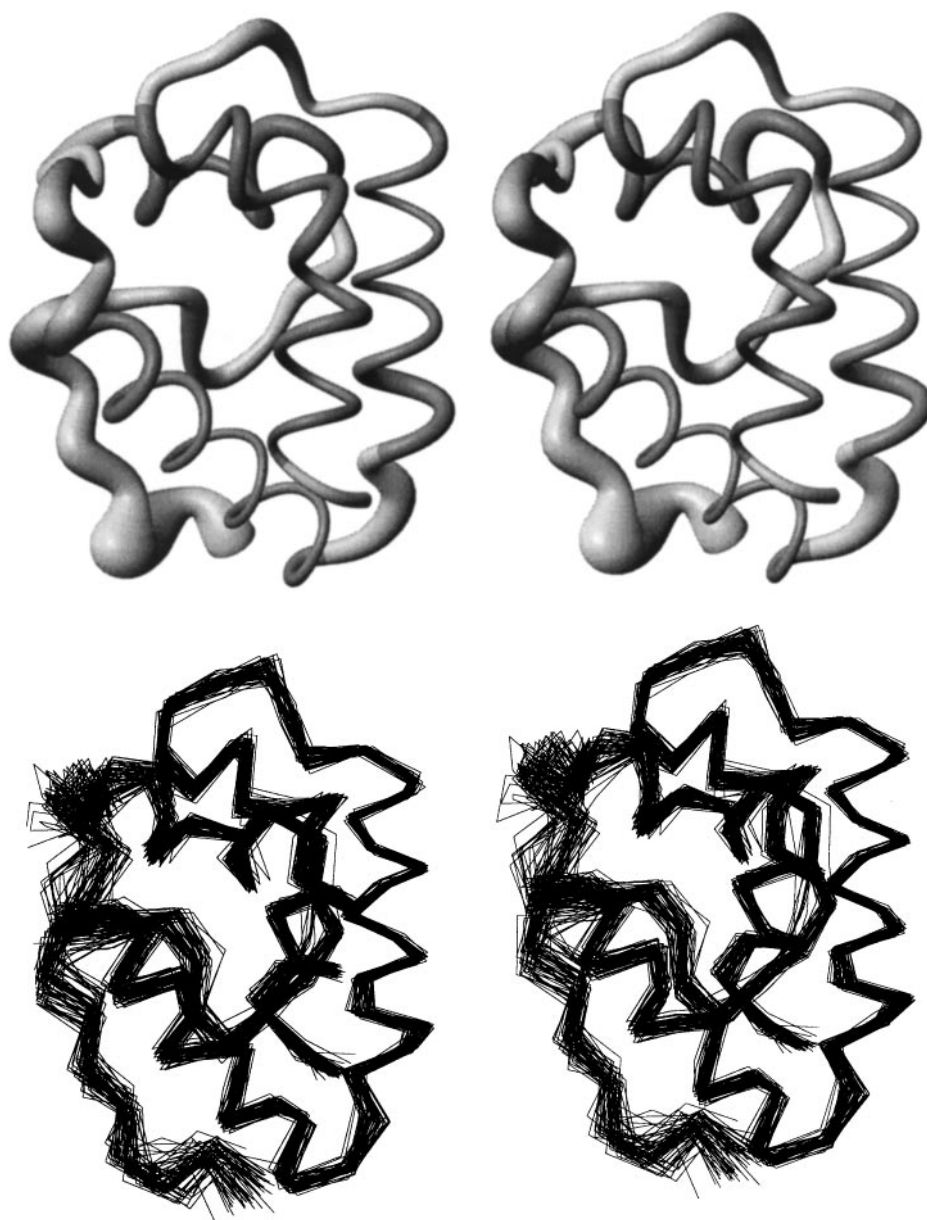


Fig. 3. Stereoscopic view of the 14 superposed solution structures of rice ns-LTP represented with the MolMol sausage model. The thickness of the tube is proportional to the local rms deviation value/upper/and by the C α trace backbone model (lower). the structures were best fitted on the backbone atoms of residues 3–79.

interaction reduces Thr46 side-chain exposition to the solvent. The motion of β and γ methylene groups of Arg32 and Arg44 is reduced due to the packing with other side-chains: Val75 for Arg32 and Pro78, Tyr79, Val90 in the case of Arg44, respectively, when their guanidinium groups are solvent exposed. Moreover, in a majority of structures Arg44 and Arg45 guanidinium groups interact with the carbonyl group of Val90 and/or with C-terminal carboxylate (Fig. 4C). Some other amino acid side-chains display also a reduced mobility because they are involved in long-range hydrogen bonds, Asn49(H δ 21) interacts with the carbonyl group of Arg45, Asn62(H δ 21) with Arg18(O) and Asn65(H δ 21) with Tyr16 (O) (Fig. 4B). Asn62 side-chain is involved in another hydrogen bond between its O δ and Gly64(HN). Asp43 side-chain carboxylate group interacts both with the charged N-terminal ammonium group of Ile1 and either with the HN of Ser39 or Thr40 (Fig. 4C). The side-chain

conformation of Gln5 is stabilized by a hydrogen bond between its O ϵ and the amide proton of Thr2. As a consequence, Thr2(HN) is the only amide proton of the N-terminal part of the protein observed in the spectra. Finally, a hydrogen bond is observed between the H δ 21 of Asn65 and the C=O oxygen of Tyr16, in agreement with experimental proton exchange data showing that both Asn65 side-chain amide protons are well protected from the solvent exhibiting a half-time of exchange exceeding 10 h.

Protein interior – hydrophobic core and the cavity

All hydrophobic residues including some alanines are significantly protected from the solvent (Fig. 1D) and point inwards the core of the protein (Fig. 4D). Their side-chains are relatively well packed, resulting in small χ 1 and χ 2 circular variances, in

most cases lower than 0.005. Only Leu14, Leu34 and Val90 side-chains display significant variability with circular variances $\chi_1 = 0.74$, $\chi_2 = 0.41$ (Leu14), $\chi_1 = 0.25$, $\chi_2 = 0.60$ (Leu34) and $\chi_1 = 0.35$ (Val90), respectively. In contrast, the conformation of Ile1, probably due to a large number of medium and long range distance constraints is very well resolved. The proximity of residues Tyr16, Pro23 and Ile69 side-chains builds a small hydrophobic cluster in which Tyr16 phenol ring is nearly parallel to Pro23 pyrrolidine ring (Fig. 4B). Such hydrophobic association decreases the solvent accessible surface of both amino acids. Furthermore, the Tyr16(O)-Gly20(HN) hydrogen bond stabilizes the spatial organization of L1. In consequence, despite the potential flexibility of the Gly19-Gly20-Ala21-Gly22 sequence, the ϕ and ψ circular variances as well as local the rms deviation of L1 are relatively low (Fig. 1C). The interaction of Tyr16 with Pro23 is experimentally confirmed by the large upfield shift of Pro23(H α , H δ) protons. Tyr16 is also involved in another surface interaction with Asn65 side-chain amide group resulting in a large upfield shift of asparagine H δ 22 proton. Analogous interactions with Tyr79 resulting in upfield shifts of Ala47(H α), Cys48(H α), Leu51(C δ H $_3$) are confirm by chemical shift calculations.

Besides the Tyr16-Pro23 cluster, two aggregates of hydrophobic side-chains are found in the core of the protein (Fig. 4D). The first one created by hydrophobic interactions of H1 with H2 and H3, includes the side-chains of Ile1, Val6, Ala9, Val10, Val31, Leu34, Ala38, Ala47, Leu51. The second one, generated by the interaction of H1 and L3, involves the side-chains of Leu14, Ala17, Arg18, Ile58, Leu61, Ala66, the Arg18 guanidinium group being exposed to the solvent. These three hydrophobic regions are present in all analyzed structures of rice ns-LTP. As a matter of fact all these hydrophobic residues are conserved in the sequences of all other ns-LTPs (Table 3) demonstrating that hydrophobic forces may play an important role in the stabilization of the protein scaffold as well as in the ligand binding. The use of CAVITE program [7] allows the determination of protein internal cavity for each solution structure. It clearly appears that its shape and dimension strongly depend on the protein conformation and vary from 140 Å³ to 350 Å³, the estimate average value being 270 ± 70 Å³. The largest, nearly spherical cavity common to all structures has a volume of 100 Å³ and is delimited by the side-chains of Val31, Leu34, Leu51, Ile69, Pro70, Val77, Tyr79 and the methyl group of Thr80 (Fig. 5A). There is only one structure which exhibits one large cavity (Fig. 5B). In three others the relatively large cavity is split into two smaller ones by the aliphatic side-chains of Val10, Leu51 and Ile81 (Fig. 5C). Both dimension and shape of the cavity are very sensitive to the global fold. Indeed small reorganizations of bordering amino acid orientations induce a large change in the cavity shape. The cavity can expand towards the peripheral regions of the molecule. Two main directions of expansion are in the neighbourhood of L2 delimited by Ala47, Ala38 and Val6 (Fig. 5D) and of H1a-L3 interaction.

DISCUSSION

Comparison to the crystallographic structure of rice ns-LTP

The statistics of rms deviations calculated for the ensemble of rice LTP solution structures in relation to the crystallographic one (deposited as 1r2l in the Protein Data Bank [68]) are

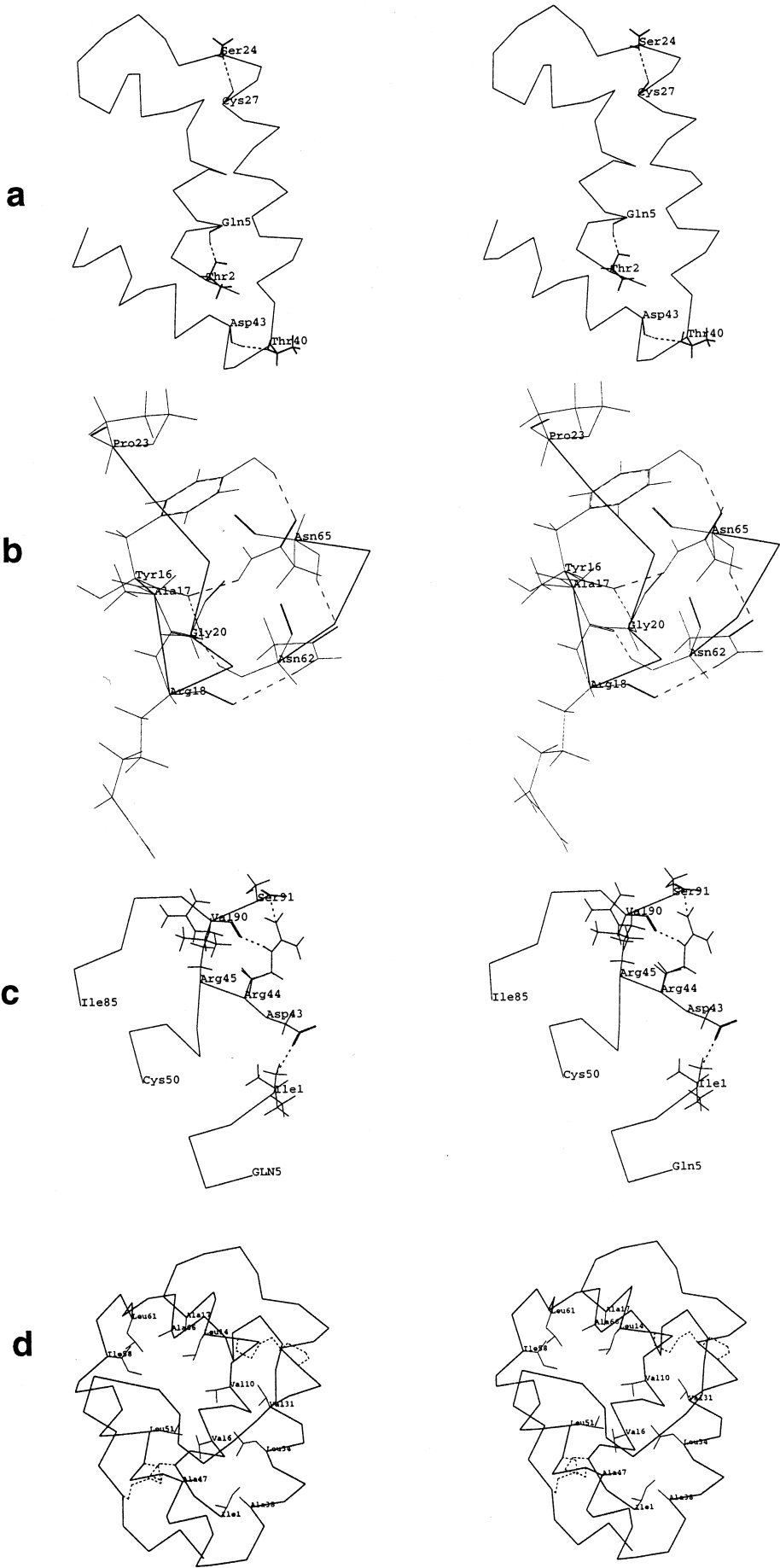
presented in Table 2. Despite local differences, both structures are globally similar with a total rms deviation 1.52 Å (Fig. 6). Generally the helical cores are closer to each other than the rest of the molecules. For individual helices, the rms deviations vary in a range of 0.14–0.35 Å for the well defined H1a, H2 and H3 helices while H1b and H4 show a larger value of 0.5–0.6 Å. The structure of the first part of the C-terminal tail Ta is close to the crystallographic one (rms deviation 0.53 Å) whereas the second part Tb differs significantly (rms deviation 1.12 Å).

Both structures exhibit nearly identical hydrogen bonding pattern. Beside the intrahelical network, the first turns of the H1, H2 and H3 helices are additionally stabilized by *i,i*+3 hydrogen bonds between the O γ side-chain oxygen of the N-terminal amino acid of the helices: Thr2(O γ 1)-Gln5(HN), Ser24(O γ)-Cys27(HN) and Thr40(O γ 1)-Asp43(HN), respectively. Furthermore, the hydrogen bond Ser82(O γ)-Ser84(HN) found to stabilize the (82–85) β -turn in solution, exists also in the crystal structure. The stabilization of protein termini by Asp43-Arg44-Arg45 fragment described in the solution structure is also observed in the crystallographic one. Moreover a hydrophobic interaction between residues Val77-Pro78-Tyr79 located in a turn, as well as the pseudo-stacking conformation of Pro23 and Tyr16 is confirmed by crystallographic structure analysis. In addition to specific local interactions there is a number of long range ones (in the sequence terms) which also stabilize the protein structure. Among them, there are the long range hydrogen bonds between Ala17(O)-Asn62(HN) and Tyr16(O)-Asn65(H δ 21) which were described in the previous section and were found in the X-ray structure. Two hydrophobic clusters involving H1a, H2, H3 and H1b, L3, respectively, are also conserved and both the dimension and shape of the cavity are similar.

Comparison to other structures of ns-LTPs

To date, in addition to rice ns-LTP, another crystallographic structure: maize (1m2l) [7]; and three solution structures of ns-LTPs: wheat (1ltp) [3], maize (1afh) [4] and barley (1lip) [5] are available in the PDB. They all exhibit a similar fold and differ only in details of the secondary structure (Fig. 6). The statistics of rms deviations calculated for the ensemble of rice LTP solution structures in relation to the crystallographic structure of maize ns-LTP and average NMR structures are presented in Table 2. In all of them, as in rice ns-LTP, intrahelical proline causes $\approx 30^\circ$ bend of helix H1 axis. The largest differences are located in the H4 helix. They all contain a proline residue corresponding to rice Pro70 that, as aforementioned, causes a distortion of the helical structure differing from the canonical α structure. The N-terminal amino acid of this helix changes from protein to protein, which is consistent with the structural description exhibiting large distortion of the N-terminal part of H4. In contrast H1, H2 and H3 occupy almost the same sequence range. The most striking feature is that the same interaction pattern stabilizing the global fold is conserved for all structures. For instance the Asn65(H δ)-Tyr16(O) hydrogen bond is found in all analyzed structures. The second important hydrogen bond: Asn62(HN)-Ala17(O) can also be observed in a significant number of structures. The salt bridges connecting the guanidinium groups of Arg44 and Arg45 with the C-terminal carboxylate group were already determined in the case of maize ns-LTP solution and crystal structures [7]. A

Fig. 4. Stereoscopic views of the conserved mechanisms of the fold stabilization involving N-capping residues. (A) H1b-L3 interaction; (B) Asp43-Arg44-Arg45 couple interacting with both protein termini; and (C) hydrophobic interactions of H1, H2, H3 and L3 (D).



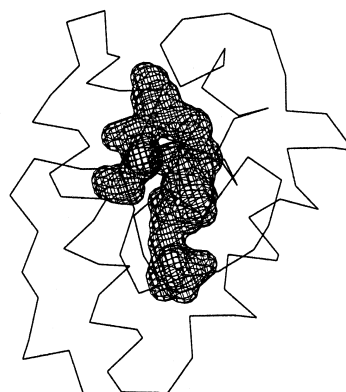
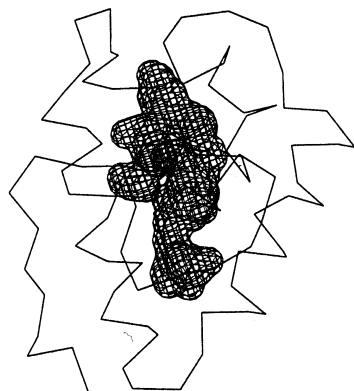
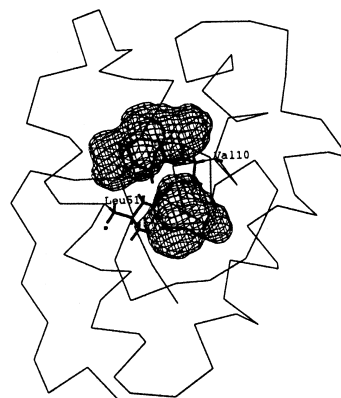
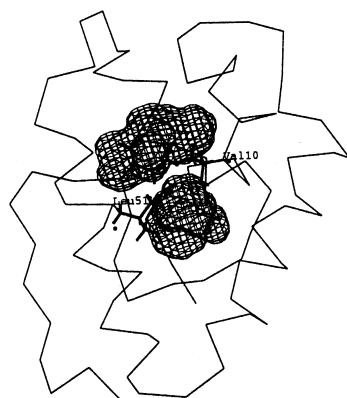
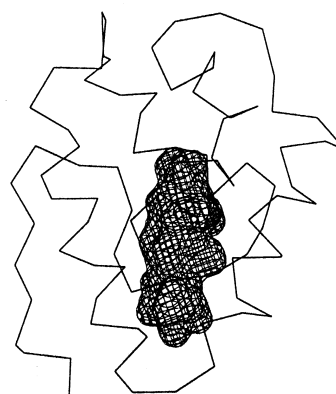
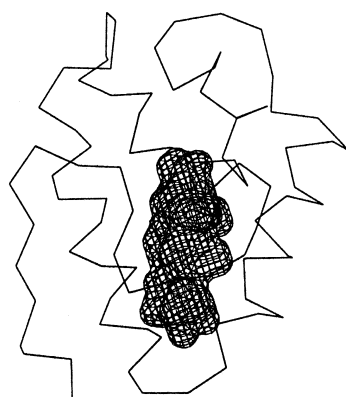
a**b****c****d**

Table 2. Average root-mean-square deviations (RMSDs) of rice ns-LTP solution structure calculated in relation to different reference structures.

		Backbone		All heavy atoms	
RMSD (Å)		Average	Range	Average	Range
From average structure					
All	(2-90)	0.72 ± 0.01	(0.56,0.88)	1.14 ± 0.16	(0.85,1.41)
Helices		0.46 ± 0.06	(0.37,0.56)	0.78 ± 0.07	(0.68,0.92)
C-terminal region	(74-90)	1.13 ± 0.25	(0.85,1.63)	1.77 ± 0.42	(1.15,2.49)
	(74-79)	0.58 ± 0.15	(0.37,0.85)	0.72 ± 0.14	(0.53,1.08)
	(80-90)	1.33 ± 0.31	(0.93,1.94)	2.13 ± 0.52	(1.33,3.05)
Pairwise					
All	(2-90)	1.06 ± 0.15	(0.72,1.41)	1.68 ± 0.25	(1.17,2.26)
Helices		0.67 ± 0.10	(0.46,0.96)	1.15 ± 0.11	(0.91,1.43)
C-terminal region	(74-91)	1.66 ± 0.37	(0.87,2.81)	2.58 ± 0.70	(1.23,4.25)
	(74-79)	0.85 ± 0.24	(0.36,1.41)	1.05 ± 0.23	(0.51,1.75)
	(80-90)	1.95 ± 0.46	(0.98,3.34)	3.10 ± 0.88	(1.39,5.20)
From PDB data of rice ns-LTP					
All		1.52 ± 0.07	(1.39,1.63)		
Helices		0.97 ± 0.11	(0.77,1.16)		
C-terminal region	(74-91)	1.77 ± 0.13	(1.56,2.02)		
From PDB data of maize ns-LTP					
All	(2-18 + 25-90)	1.64 ± 0.14	(1.38,1.94)		
Helices		1.22 ± 0.07	(1.08,1.39)		
C-terminal region	(74-91)	2.32 ± 0.31	(1.95,3.11)		
Mean solution structure of rice ns-LTP from other mean NMR structures.					
	Amino acid		Helical regions		
	residue 2 : 90				
Maize	2.05		1.89		
Wheat	1.96		1.94		
Barley	2.77		2.32		

detailed analysis of the ns-LTP solution structures demonstrates that for these proteins analogous interactions can be also observed as the interaction of the carboxyl group of Asp43 with the N-terminal amino group. As in rice ns-LTP two large hydrophobic clusters were found. Analogous interaction involving aliphatic residue 77, Pro78 and Tyr79 side-chains were found to stabilize the 77–81 turn structure. In contrast, the side-chain pseudo-stacking interaction between Tyr16 and Pro23 is poorly resolved. Analysis of the sets of structures shows that for maize and wheat ns-LTPs a perpendicular organization of the phenol and pyrrolidine rings is preferred, whereas in barley ns-LTP a pseudo-stacking form is dominating. But, despite structural differences, the chemical shift values obtained for Pro23, especially for H α and H δ are almost identical for rice, maize, wheat and barley ns-LTPs in solution [69–70].

Cavity capacity to the ligand binding

The analysis of the crystallographic structure of maize ns-LTP complexed with the palmitate [6] (1mzm in PDB) and solution structures of barley ns-LTP complexed with the palmitoyl CoA [9] (1jTb), maize ns-LTP with 1-palmitoyl-2-lysophosphatidylcholine [9] and wheat ns-LTP with 1,2-dimyristoylphosphatidylglycerol [10] demonstrates that hydrophobic fatty acid chains are inserted into the protein cavity. Despite similarities in the shape and dimension of the cavities the position of the ligand entrance of the cavity could differ. In maize and wheat ns-LTPs the cavity opens in the region of L2 and Val90 while in barley ns-LTP the palmitoyl CoA entrance is located in neighbourhood

of L3. To date, there is no consensus about protein or lipid specificity of the ligand entrance localization. Analysis of the uncomplexed rice ns-LTP cavities show that most of them are appropriate to the ligand binding with the entrance located close to L2, as displayed in the crystallographic structure of maize ns-LTP complexed with the palmitate and in the solution structure of wheat LTP complexed with 1,2-dimyristoylphosphatidylglycerol [10]. However, the biggest U-shaped cavity seems to have a possibility to adapt its structure to both entrance locations (Fig. 7A and B). This strongly suggests, that, in the case of rice ns-LTPs, as well as for other ns-LTPs, both types of ligand binding may be possible.

Possible role of the evolutionary conserved amino acids

The sequential alignment of 36 ns-LTPs sequences including all proteins and isoproteins found in the SWISS-PROT sequence data base [71,72] is presented in Table 3. The structure of ns-LTP, in the absence of a uniform but tight hydrophobic core, is stabilized by numerous specific interactions which were previously described. All these amino acid residues are conserved among the sequences. Besides these hydrophobic interactions, the protein fold is stabilized by four strictly conserved disulfide bonds (Cys3-Cys50, Cys13-Cys27, Cys28-Cys73 and Cys48-Cys87 in the case of rice ns-LTP) and electrostatic interactions as well. The most important interactions, constraining the terminal amino acids of the proteins involve Asp43, whose carboxylate group makes a salt bridge with the N-terminal ammonium group, as well as Arg44 and Arg45 guanidinium

Fig. 5. Stereoscopic views of the protein and the cavity obtained for four solution structures of rice ns-LTP exhibiting different types of the cavity organization. (A) Organization of the amino acids bordering the basic cavity; (B) the largest cavity; (C) cavity split into two regions; and (D) cavity expanded to the neighbourhood of L2, almost open to the exterior.

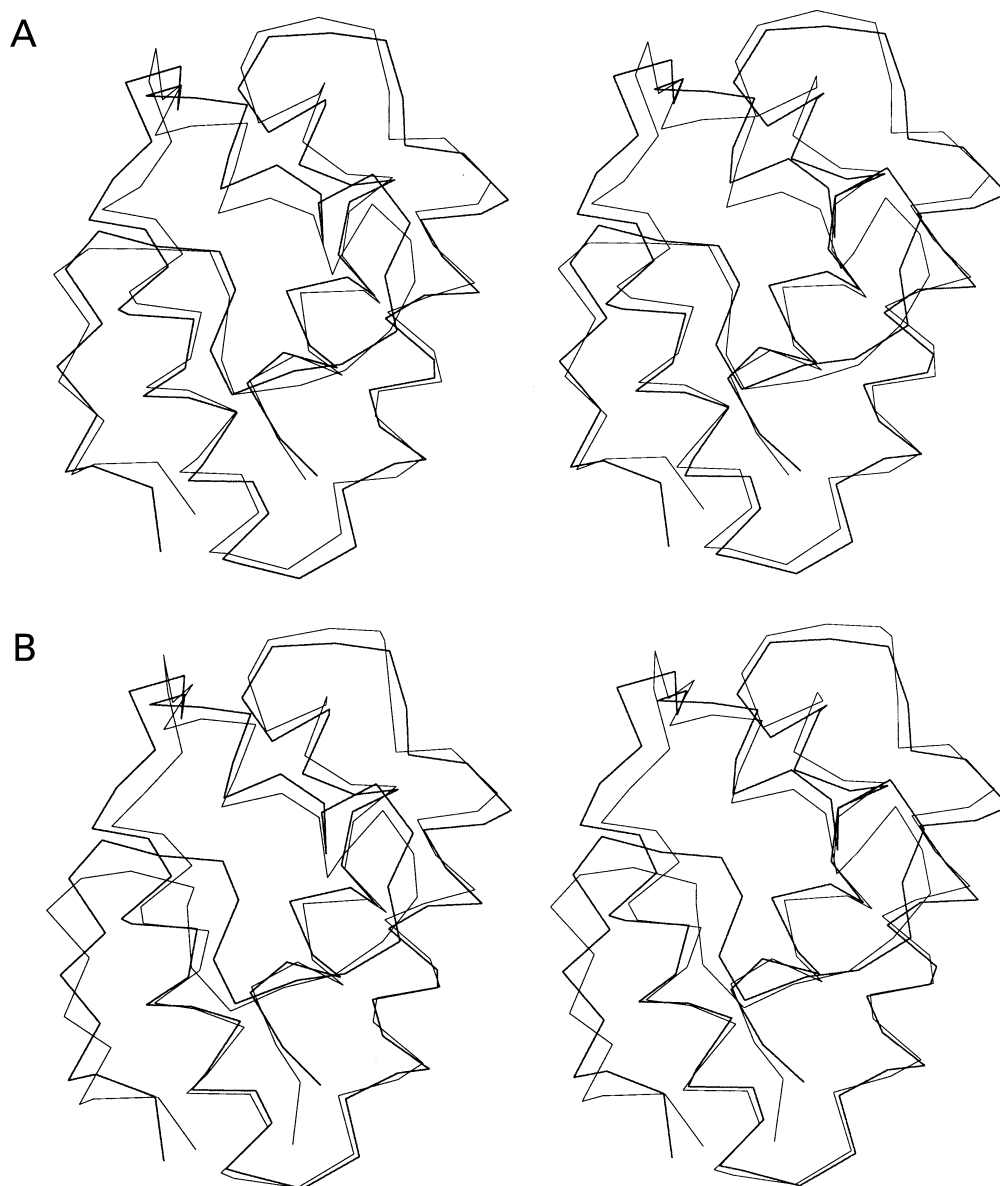


Fig. 6. Comparison of the mean solution structure of rice ns-LTP (thick line) and the crystallographic structure of rice (A) and maize (B) ns-LTP. The best fit was obtained for the backbone heavy atoms corresponding to rice residues 3–18 and 26–79.

groups, which are hydrogen bonded to the C-terminal carboxylate group and to the carbonyl group of the last but one amino acid residue, respectively. The triad Asp43-Arg44-Arg45 is strictly conserved. The only observed Asp43 mutation which breaks the interaction appears in cotton ns-LTPs (into Val), whereas Arg44 undergoes a conservative mutation into Lys in barley, carrot and some castor bean ns-LTPs. Arg45 is slightly less conserved and could be replaced either by a lysine or by a glutamine. In all cases at this sequence location the amino acid is a donor whose side-chain is long enough to generate the hydrogen bond with the 90th residue. Asp86 carboxylate group could be involved into a salt bridge with Lys52 amino group stabilizing the conformation of the C-terminal tail. This interaction may drive in the folding process to generate the Cys48-Cys87 disulfide bridge [7], both residues being highly conserved, Asp 86 could be mutated only into Asn. Asn62 and Asn65 are involved in hydrogen-bonding interactions between L3 and the end of H1. Asn62 is additionally involved in L3

stabilization (Fig. 4B). Asn62 is conserved in 70% of the primary structures, whereas Asn65 is less conserved and could be replaced either by Lys or Arg. Helix N-capping residues, Thr2, Ser24 and Thr40, whose side-chains may be involved in hydrogen bonding, are also conserved. At these N-capping positions both Thr and Ser residues stabilize preferably the N terminus of α helix [65]. In H1 they could be replaced by other good N-capping residues: Asn or Asp (both stabilize either α or 3_{10} helices) whereas in H2, a proline residue, exhibiting analogous N-capping properties as Asn and Asp [65], is frequently observed. Ser82, which stabilizes the turn of the C-terminal loop, through hydrogen bonding to its side-chain oxygen, is strictly conserved. There is only one occurrence with a mutation into threonine. All these amino acids are making either salt bridges or hydrogen-bonding patterns constraining the different parts of the protein, as it is presented in Fig. 4A–C.

Hydrophobic interactions are supposed to play an important role in the stabilization of the four-helix fold. In the ensemble of

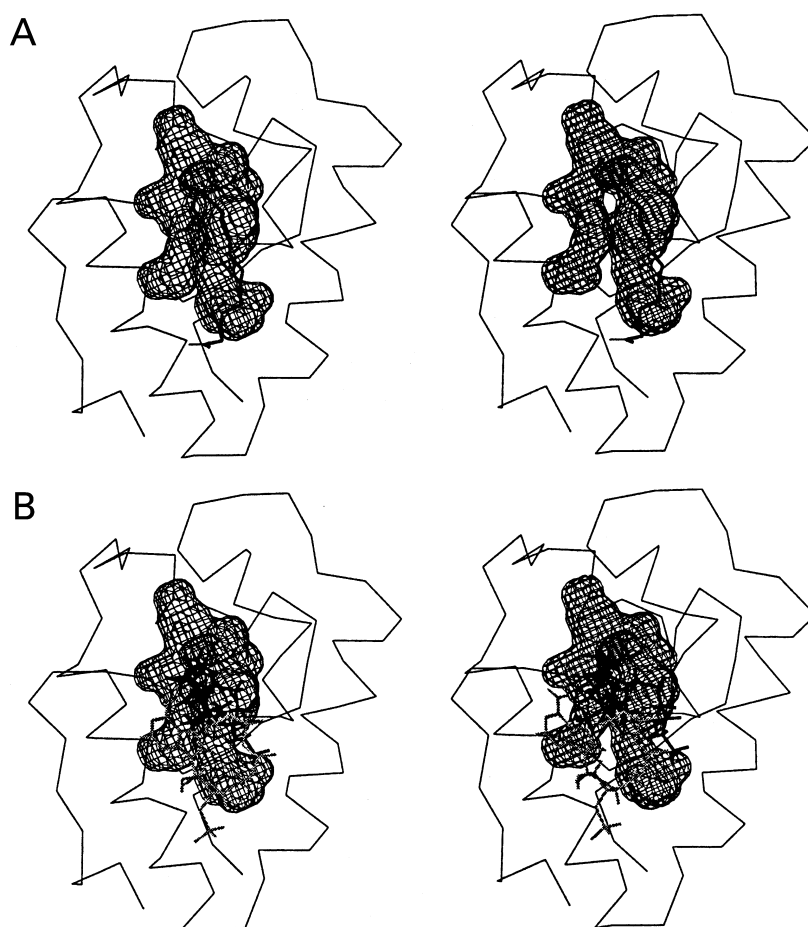


Fig. 7. Comparison of the largest cavity obtained for the 14 solution structures of rice ns-LTP. With the hypothetical ligand positions adapted from the crystallographic structure of maize ns-LTP complexed with the palmitate (A) and solution structure of barley ns-LTP complexed with the palmitoyl CoA (B).

the analyzed structures two hydrophobic clusters involving helix residues were distinguished (Fig. 4D). The first one, originating from the interaction of H1, H2, H3 occurs between residues Ile1, Val6, Val10, Val31, Leu34, Ala38, Ala47, Leu51. The second cluster made of residues Leu14, Ala17, Ile58, Leu61 and Ala66 packs the C-terminal part of H1 against L3. All described amino acids are conserved among the sequences. In addition, Val75, Val77, Ile81 and Val90 are the centres of hydrophobic interactions of the C-terminal tail with the helical regions and are present in all primary structures. Additionally the conservative Val77, Pro78 and Tyr79 are supposed to stabilize the first turn of the C-terminal tail by hydrophobic clustering of their side-chains. Three other prolines, Pro12, Pro23 and Pro70, are strictly conserved in all plant ns-LTP sequences. Two of them, Pro12 and Pro70 are found in the helical structures. Pro12 produces the bend of H1 whereas Pro70 only distorts the structure of H4. Pro23 is supposed to be involved in pseudo-stacking with Tyr16, stabilizing the conformation of the L1 loop and additionally exhibiting hydrophobic contact with Ile69 from H4. Both Tyr16 and Ile69 are conserved, whereas in a number of sequences the position of a proline corresponding to rice Pro23 only changes by only one amino acid towards either C or N terminus.

Despite the presented pattern of well-defined interactions, a number of solvent exposed, as well as solvent protected amino acids, are conserved in terms of their polarity. The latter makes a pattern of hydrophobic/hydrophilic amino acid residues along the sequences common to all plant ns-LTPs.

CONCLUSIONS

Clearly, considering the three-dimensional structures of wheat, maize and barley ns-LTPs, rice ns-LTP belongs to the previously defined structural family of lipid carriers. Even if, all these proteins bind a variety of hydrophobic molecules, their interaction with organized structures such as micelles, bilayers from which they extract these molecules as well as the form and stability of the protein–lipid complexes will not be strictly identical. The volume and the organization of the internal cavity which is the site of lipid binding exhibit subtle differences from one protein to another. It would be especially interesting now to establish the three-dimensional structures of several complexes involving various lipids with different bulkiness and different polar heads in order to test the plasticity of ns-LTPs. Classically, the hydrophobic interior of a number of proteins, is believed to be tightly packed and was shown to have very low compressibility [72]. Nevertheless, a sound analysis of an extended basis of structures has revealed the existence of empty spaces especially in the regions where hydrophobic faces of secondary structures are interacting [73]. Plant ns-LTPs seem to be extreme cases since their internal core is loose enough to allow binding of very bulky molecules. A simulation of expansion of their internal cavity as well as an analysis of the four helix arrangement (unpublished results) reveals that both its shape and volume can considerably vary without major changes in backbone fold. Local reorganization of side-chains and very minor changes in the relative positions of secondary structure elements

Table 3. Sequence alignment of 36 plant ns-LTPs. Helical regions marked by grey background. The N-terminal signal extensions were removed.

	1										2										3										4																		
	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24	25	26	27	28	29	30	31	32	33	34	35	36	37	38	39	40	41								
NLT1_PRUDU almond	-	A	A	V	L	N	C	G	Q	Q	V	S	S	N	L	A	P	C	I	P	Y	V	R	G	G	-	-	G	A	V	P	P	A	C	C	C	T	G	V	R	S	T	T	T					
NLT3_PRUDU almond	A	A	V	S	C	G	Q	Q	V	D	S	K	M	K	L	A	P	C	I	N	Y	V	Q	G	-	G	P	G	-	P	S	G	E	C	C	N	G	V	R	S	T	T	T						
NLT1_HORVU barley 1LIP	V	A	I	S	C	G	Q	Q	V	S	A	L	S	P	C	L	A	P	C	I	S	Y	A	R	G	N	G	A	K	-	P	P	V	A	C	C	S	G	V	K	R	L	A	Q	Q	S	T		
NLT4_HORVU barley	A	A	I	S	C	G	Q	Q	V	S	G	A	L	A	P	C	L	G	Y	A	R	S	Q	V	N	V	K	-	P	P	V	L	T	C	C	N	V	R	G	L	N	A	A	A	R	T	T		
NLTP_DAUCA carrot	A	V	L	T	C	G	Q	Q	V	N	S	S	L	A	S	C	I	P	F	L	T	G	-	G	V	A	S	P	S	A	S	C	C	A	G	V	Q	M	E	L	K	R	L	A	P	T	T	V	
NLTA_RICCO castor bean riccinu	-	-	V	D	C	G	Q	Q	V	N	K	A	L	S	S	C	V	P	F	L	T	G	-	F	D	T	T	P	S	L	T	C	C	A	G	V	Q	M	E	L	K	R	L	A	P	T	T	V	
NLTB_RICCO castor bean riccinu	-	-	V	N	C	G	Q	Q	V	N	K	A	L	S	S	C	V	P	F	L	T	G	-	F	D	T	T	P	S	L	T	C	C	A	G	V	Q	M	E	L	K	R	L	A	P	T	T	V	
NLTC_RICCO castor bean riccinu	A	A	V	P	C	S	T	V	D	M	K	A	A	A	C	V	G	F	F	A	T	G	K	D	S	K	-	P	S	Q	A	C	C	T	G	L	Q	Q	L	A	Q	T	V	K	S	V	V		
NLTD_RICCO castor bean riccinu	A	A	V	P	C	S	T	V	D	M	K	A	A	A	C	V	G	F	F	A	T	G	K	D	S	K	-	P	S	Q	A	C	C	T	G	L	Q	Q	L	A	Q	T	V	K	S	V	V		
NLTA_BRAOL cauliflower	R	A	L	T	C	G	T	V	N	S	N	V	A	A	P	C	I	G	Y	I	T	Q	G	-	T	-	L	P	G	A	C	C	T	G	V	T	S	L	N	N	M	A	R	T	T	T			
NLTB_BRAOL cauliflower	A	A	L	S	C	G	T	V	S	G	N	L	A	A	P	C	I	G	Y	L	A	Q	N	A	-	P	-	L	P	R	T	G	C	C	T	G	V	T	S	L	N	N	M	A	R	T	T	T	
NLT1_GOSHI cotton	A	A	L	S	C	G	T	V	S	G	N	L	A	A	P	C	I	G	Y	L	A	Q	N	A	-	P	-	L	P	R	T	G	C	C	T	G	V	T	S	L	N	N	M	A	R	T	T	T	
NLT2_GOSHI cotton	G	A	V	T	S	G	Q	V	T	N	S	L	A	P	C	I	N	Y	L	R	G	S	G	A	G	A	-	P	-	V	P	P	A	C	C	T	G	I	K	S	L	N	S	A	A	Q	T	T	T
NLTP_GERHY gerbera	G	A	V	T	S	G	Q	V	T	N	S	L	A	P	C	I	N	Y	L	R	G	S	G	A	G	A	-	P	-	V	P	P	A	C	C	T	G	I	K	S	L	N	S	A	A	Q	T	T	T
NLTP_AMACA inca-wheat	A	-	I	S	C	G	Q	V	T	S	G	L	V	P	C	F	G	Y	L	A	A	G	-	P	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-		
NLTP_PINTA lobliolly pine	G	A	I	S	C	N	Q	V	V	S	A	M	T	P	C	A	T	Y	L	K	G	T	G	A	T	P	P	P	A	N	C	C	A	G	V	R	S	L	N	S	A	A	Q	T	T	T	T		
NLTP_MAIZE maize 1M2L	A	A	I	S	C	G	Q	V	A	S	A	I	A	P	C	I	S	Y	A	R	Q	G	S	-	G	-	P	S	A	A	G	C	C	S	G	V	R	S	L	N	S	A	A	Q	T	T	T		
NLT1_ARATH mouse-ear cress	A	A	L	S	C	G	S	V	N	S	N	L	A	A	P	C	I	G	Y	V	L	Q	G	-	V	-	I	P	P	A	A	C	C	Q	S	G	V	R	S	L	N	S	A	A	Q	T	T	T	
NLTP_ELECO ragi	-	A	I	S	C	G	Q	V	S	S	A	I	G	P	C	L	A	Y	A	R	G	A	A	-	P	-	P	S	A	S	A	C	C	Q	S	G	V	R	S	L	N	S	A	A	Q	T	T	T	
NLT1_BRANA rape	A	A	L	S	C	G	T	V	S	G	N	L	A	A	P	C	I	G	Y	L	T	Q	N	G	-	P	-	L	P	R	G	C	C	T	G	V	T	N	L	N	N	M	A	R	T	T	T		
NLT2_BRANA rape	A	A	L	S	C	G	T	V	S	G	N	L	A	A	P	C	I	G	Y	L	T	Q	N	G	-	P	-	L	P	R	G	C	C	T	G	V	T	N	L	N	N	M	A	R	T	T	T		
NLT3_BRANA rape	A	A	L	S	C	G	T	V	S	G	N	L	A	A	P	C	I	G	Y	L	T	Q	N	G	-	P	-	L	P	R	G	C	C	T	G	V	T	N	L	N	N	M	A	R	T	T	T		
NLTP_ORYSA rice	-	-	I	T	C	G	Q	V	N	S	A	V	G	P	C	L	T	Y	A	R	G	-	G	A	G	-	P	S	A	A	C	C	S	G	V	R	S	L	K	A	A	A	S	T	T	T	T		
NLT2_ORYSA rice	A	A	I	S	C	G	Q	V	N	S	A	V	S	P	C	L	S	Y	A	R	G	-	G	S	G	-	P	S	A	A	C	C	S	G	V	R	S	L	N	S	A	A	A	S	T	T	T		
NLT3_ORYSA rice	A	A	I	S	C	G	Q	V	N	S	A	V	S	P	C	L	S	Y	A	R	G	-	G	S	G	-	P	S	A	A	C	C	S	G	V	R	S	L	N	S	A	A	A	S	T	T	T		
NLT4_ORYSA rice	A	A	V	S	C	G	D	V	T	S	S	I	A	P	C	L	S	Y	V	M	G	-	R	E	S	S	-	P	S	S	S	C	C	S	G	V	R	S	L	N	S	A	A	A	S	T	T		
NLT1_SORVU sorghum	A	A	V	T	C	G	Q	V	S	S	A	I	A	L	P	C	L	S	Y	A	R	G	Q	G	S	-	P	S	A	A	G	C	C	S	G	V	R	S	L	N	S	A	A	A	S	T	T		
NLTP_SORVU sorghum	A	A	V	T	C	G	Q	V	S	S	K	L	A	P	C	I	G	Y	L	K	G	-	G	-	P	-	L	G	G	G	C	C	G	G	I	K	A	L	N	A	A	A	A	A	S	T	T		
NLTP_SPIOL spinach	A	G	I	T	C	G	M	V	S	S	K	L	A	P	C	I	G	Y	L	Q	N	T	G	-	A	-	P	G	P	S	A	A	C	C	G	G	I	K	A	L	N	S	A	A	A	S	T	T	
NLTP_BETVU sugar beet	A	-	I	T	C	G	L	V	A	S	K	L	A	P	C	I	G	Y	L	R	N	T	G	-	P	-	-	L	G	S	C	C	G	G	V	K	A	L	L	V	N	S	A	A	A	S	T	T	
NLT1_TOBAC tobacco	A	-	I	T	C	G	Q	V	T	S	N	L	A	P	C	L	A	Y	L	Q	N	T	G	-	P	-	-	L	G	S	C	C	G	G	V	K	A	L	L	V	N	S	A	A	A	S	T	T	
NLTP_TOBAC tobacco	A	-	L	S	C	G	Q	V	T	S	G	L	A	P	C	L	A	Y	L	Q	N	T	G	-	P	-	-	L	G	S	C	C	G	G	V	K	A	L	L	V	N	S	A	A	A	S	T	T	
NLT1_LYCES tomato	S	-	L	S	C	G	E	V	T	A	G	L	A	P	C	L	P	Y	L	E	G	R	G	-	P	-	-	L	G	S	C	C	G	G	V	K	A	L	L	V	N	S	A	A	A	S	T	T	
NLTP_LYCES tomato	A	-	L	T	C	G	Q	V	T	S	G	L	A	P	C	L	P	Y	L	Q	N	T	G	-	P	-	-	L	G	S	C	C	G	G	V	K	A	L	L	V	N	S	A	A	A	S	T	T	
NLTA_WHEAT wheat 1LTP	V	A	I	D	C	G	H	V	D	S	L	V	R	P	C	L	S	Y	V	Q	G	-	G	P	G	-	P	S	G	G	C	C	C	D	G	V	K	A	L	L	H	N	Q	A	R	S	Q		
consensus	IL	n	C	V	h	P	C	IL	Y	h	Q	G	-	G	P	G	-	P	S	G	G	-	G	P	G	-	P	S	G	G	C	C	C	D	G	V	K	A	L	L	H	N	Q	A	R	S	Q		

Legend to the consensus: n, helix N-cap; h, hydrophobic; hh, large hydrophobic; p, polar; capital letters, amino acids.

are sufficient to considerably expand the cavity. The intrinsic stability to high pressure (6 kbar) and temperature (60 °C) (S. Tassin, personal communication) of these proteins which does not present a closed packed hydrophobic core but are rather stabilized by disulfide bridges and a series of subtle long range electrostatic and hydrophobic interactions, is puzzling. As a matter of fact, ns-LTPs are probably members of a larger structural family of proteins, including *Ace*-AMP1 an antifungal protein from onion seeds [24] and the hydrophobic protein from soybean (HPS) [74]. Although unable to transfer or to bind lipids *in vitro*, *Ace*-AMP1 and HPS display a similar global fold to ns-LTPs, however, they do not present a hydrophobic cavity, which may explain why they are not lipid carriers. This outlines the importance of the cavity in plant ns-LTP structure which is presumably related to their true *in-vivo* role.

The existence of several LTP genes in rice, which are subjected to complex regulation strongly suggests that these proteins be involved in some important physiological mechanisms. Unfortunately, due to this complexity the actual role of these proteins has not been yet established. Their capacity to bind lipids supports the hypothesis of a role in the protection of seeds, leaves and different parts of plants by transporting cutin monomers [21]. Actually, ns-LTPs bind such molecules (D. Marion, personal communication), but such a mechanism *in situ* has yet to be confirmed. According to the *in-vitro* antifungal properties of some ns-LTPs [22] they may also be active as plant defence proteins. Such a role would be in agreement with their spatial and temporal distribution.

A first step in the understanding of a potential multifunctionality of ns-LTPs should be to analyze the three-dimensional structures and the lipid binding capacity as well as the antifungal properties of the proteins corresponding to the different genes which are expressed and regulated along the plant life. Plant ns-LTPs appears ideally designed to bind and carry not only lipids but also a variety of organic molecules which makes it attractive for biotechnological applications.

ACKNOWLEDGEMENTS

The authors want to thank Anita Caille for her kind experimental support. We thank Eric Forest for running mass spectrometry experiments on rice ns-LTP. This work was supported in part by the Korea-French collaborative research program of Korea Science and Engineering foundation.

REFERENCES

- Rueckert, D.G. & Schmidt, K. (1990) Lipid transfer proteins. *Chem. Phys. Lipids* **56**, 1–20.
- Kader, J.C. (1990) Intracellular transfer of phospholipids, galactolipids and fatty acids in plant cells. *Subcell. Biochem.* **16**, 69–110.
- Gincel, E., Simorre, J.P., Caille, A., Marion, D., Ptak, M. & Vovelle, F. (1994) Three-dimensional structure in solution of a wheat lipid-transfer protein from multidimensional ¹H-NMR data. A new folding for lipid carriers. *Eur. J. Biochem.* **226**, 413–422.
- Gomar, J., Petit, M.C., Sodano, P., Sy, D., Marion, D., Kader, J.C., Vovelle, F. & Ptak, M. (1996) Solution structure and lipid binding of a nonspecific lipid transfer protein extracted from maize seeds. *Protein Sci.* **5**, 565–577.
- Heinemann, B., Andersen, K.V., Nielsen, P.R., Bech, L.M. & Poulsen, F.M. (1996) Structure in solution of a four-helix lipid binding protein. *Protein Sci.* **5**, 13–23.
- Shin, D.H., Lee, J.Y., Hwang, K.Y., Kim, K.K. & Suh, S.W. (1995) High-resolution crystal structure of the non-specific lipid-transfer protein from maize seedlings. *Structure* **3**, 189–199.
- Gomar, J., Sodano, P., Sy, D., Shin, D.H., Lee, J.Y., Suh, S.W., Marion, D., Vovelle, F. & Ptak, M. (1998) Comparison of solution and crystal structures of maize non-specific lipid transfer protein. A model for a potential *in vivo* lipid carrier protein. *Proteins: Struct. Funct. Genet.* **31**, 160–171.
- Lee, J.Y., Min, K., Cha, H., Shin, D.H., Hwang, K.Y. & Suh, S.W. (1998) Rice Nos-specific lipid transfer protein: the 1.6 Å crystal structure in the unliganded state reveals a small hydrophobic cavity. *J. Mol. Biol.* **276**, 437–448.
- Lerche, M.H., Kragelund, B.B., Bech, L.M. & Poulsen, F.M. (1997) Barley lipid-transfer protein complexed with palmitoyl CoA: the structure reveals a hydrophobic binding site that can expand to fit both large and small lipid-like ligands. *Structure* **5**, 291–306.
- Sodano, P., Caille, A., Sy, D., de Person, G., Marion, D. & Ptak, M. (1997) ¹H NMR and fluorescence studies of the complexation of DMPG by wheat non-specific lipid transfer protein. Global fold of the complex. *FEBS Lett.* **416**, 130–134.
- Wirtz, K.W. & Gadella, T.W. Jr (1990) Properties and modes of action of specific and non-specific phospholipid transfer proteins. *Experientia* **46**, 592–599.
- Sha, B., Phillips, S.E., Bankaitis, V.A. & Luo, M. (1998) Crystal structure of the *Saccharomyces cerevisiae* phosphatidylinositol-transfer protein. *Nature* **391**, 506–510.
- Szyperki, T., Scheek, S., Johansson, J., Assmann, G., Seedorf, U. & Wüthrich, K. (1993) NMR determination of the secondary structure and the three-dimensional polypeptide backbone fold of the human sterol carrier protein 2. *FEBS Lett.* **335**, 18–26.
- Ko, K.W., Oikawa, K., Ohnishi, T., Kay, C.M. & Yokoyama, S. (1993) Purification, characterization, and conformational analysis of rabbit plasma lipid transfer protein. *Biochemistry* **32**, 6729–6736.
- Ohnishi, T., Hicks, L.D., Oikawa, K., Kay, C.M. & Yokoyama, S. (1994) Properties of human plasma lipid transfer protein in aqueous solution and at interfaces. *Biochemistry* **33**, 6093–6099.
- Vignols, F., Wigger, M., Garcia-Garrido, J.M., Grellet, F., Kader, J.C. & Delseny, M. (1997) Rice lipid transfer protein (LTP) genes belong to a complex multigene family and are differentially regulated. *Gene* **195**, 177–186.
- Vignols, F., Lund, G., Pammi, S., Tremousaygue, D., Grellet, F., Kader, J.C., Puigdomenech, P. & Delseny, M. (1994) Characterization of a rice gene coding for a lipid transfer protein. *Gene* **142**, 265–270.
- Bernhard, W.R. & Somerville, C.R. (1989) Coidentity of putative amylase inhibitors from barley and finger millet with phospholipid transfer proteins inferred from amino acid sequence homology. *Arch. Biochem. Biophys.* **269**, 695–697.
- Vergnolle, C., Arondel, V., Tchang, F., Grosbois, M., Guerbet, F., Jolliot, A. & Kader, J.C. (1988) Synthesis of phospholipid transfer proteins from maize seedlings. *Biochem. Biophys. Res. Commun.* **157**, 37–41.
- Hollenbach, B., Schreiber, L., Hartung, W. & Dietz, K.J. (1997) Cadmium leads to stimulated expression of the lipid transfer protein genes in barley: implications for the involvement of lipid transfer proteins in wax assembly. *Planta* **203**, 9–19.
- Sterk, P., Booij, H., Schellekens, G.A., Van Kammen, A. & De Vries, S.C. (1991) Cell-specific expression of the carrot EP2 lipid transfer protein gene. *Plant Cell* **3**, 907–921.
- Molina, A., Segura, A. & Garcia-Olmedo, F. (1993) Lipid transfer proteins (nsLTPs) from barley and maize leaves are potent inhibitors of bacterial and fungal plant pathogens. *FEBS Lett.* **316**, 119–122.
- Cammue, B.P., Thevissen, K., Hendriks, M., Eggermont, K., Goderis, I.J., Proost, P., Van Damme, J., Osborn, R.W., Guerbet, F., Kader, J.C., *et al.* (1995) A potent antimicrobial protein from onion seeds showing sequence homology to plant lipid transfer proteins. *Plant Physiol.* **109**, 445–455.
- Tassin, S., Broekaert, W.F., Marion, D., Acland, D.A., Ptak, M., Vovelle, F. & Sodano, P. (1998) Solution structure of *Ace*-AMP1, a potent antimicrobial protein extracted from onion seeds. Structural analogies with plant non-specific lipid transfer proteins. *Biochemistry* **37**, 3623–3637.
- Subirade, M., Salesse, C., Marion, D. & Pezolet, M. (1995) Interaction of a nonspecific wheat lipid transfer protein with phospholipid

- monolayers imaged by fluorescence microscopy and studied by infrared spectroscopy. *Biophys. J.* **69**, 974–988.
26. Yu, Y.G., Chung, C.H., Flower, A. & Suh, S.W. (1988) Amino acid sequence of a probable amylase/protease inhibitor from rice seeds. *Arch. Biochem. Biophys.* **265**, 465–475.
 27. Park, K.D., Hwang, K.Y., Lee, S.Y., Kimm, K.K., Eom, S.H., Chung, C.H., Flower, A. & Suh, S.W. (1991) Purification and characterization of probable amylase/protease inhibitor-B protein from rice seeds. *Korean Biochem. J.* **24**, 308–314.
 28. States, D.J., Haberkorn, R.A. & Ruben, D.J. (1982) A two-dimensional nuclear Overhauser experiment with pure absorption phase in four quadrants. *J. Magn. Res.* **48**, 286–292.
 29. Marion, D. & Wüthrich, K. (1983) Application of phase-sensitive two-dimensional correlated spectroscopy (COSY) for measurement of proton-proton spin-spin coupling constants. *Biochem. Biophys. Res. Commun.* **113**, 967–974.
 30. Piotto, M., Saudek, V. & Sklenar, V. (1992) Gradient-tailored excitation for single-quantum NMR spectroscopy of aqueous solutions. *J. Biomol. NMR* **2**, 661–665.
 31. Sklenar, V., Piotto, M., Lepik, R. & Saudek, V. (1992) Gradient-tailored water suppression for ^1H - ^{15}N HSQC experiments optimized to retain full sensitivity. *J. Magn. Res.* **102**, 241–245.
 32. Wüthrich, K. (1976) *NMR in Biological Research. Peptides and Proteins*. North-Holland Publishing Company, Amsterdam.
 33. Marion, D. & Bax, A. (1988) PCOSY, a sensitive alternative for double-quantum-filtered COSY. *J. Magn. Res.* **80**, 528–533.
 34. Davis, D.G. & Bax, A. (1985) Assignment of complex ^1H NMR spectra via two-dimensional homonuclear Hartmann-Hahn spectroscopy. *J. Am. Chem. Soc.* **107**, 2820–2821.
 35. Griesinger, C., Otting, G., Wüthrich, K. & Ernst, R.R. (1988) Clean TOCSY for ^1H spin system identification in macromolecules. *J. Am. Chem. Soc.* **110**, 7870–7872.
 36. Jeener, J., Meier, B.H., Bachmann, P. & Ernst, R.R. (1979) Investigation of exchange processes by two-dimensional NMR spectroscopy. *J. Chem. Phys.* **71**, 4546–4553.
 37. Mareci, T.H. & Freeman, R. (1983) Mapping proton-proton coupling via double-quantum coherence. *J. Magn. Res.* **51**, 531–535.
 38. Brown, S.C., Weber, P. & Muller, L. (1988) Toward complete ^1H NMR spectra in proteins. *J. Magn. Res.* **77**, 166–169.
 39. Bartels, C., Xia, T.E., Billeter, M., Güntert, P. & Wüthrich, K. (1995) The Program XEASY for the computer-supported NMR spectral analysis of biological macromolecules. *J. Biomol. NMR* **5**, 1–10.
 40. Mithang, L., Keeler, J., Davis, A.L. & Oschkinat, H. (1992) Removal of zero-quantum interference in NOESY spectra of proteins by utilizing the natural inhomogeneity of the radiofrequency field. *J. Biomol. NMR* **2**, 545–556.
 41. Szyperski, T., Güntert, P., Otting, G. & Wüthrich, K. (1992) Determination of scalar coupling constants by inverse Fourier transformation of in-phase multiplets. *J. Magn. Res.* **99**, 552–560.
 42. Güntert, P., Braun, W. & Wüthrich, K. (1991) Efficient computation of three-dimensional protein structures in solution from nuclear magnetic resonance data using the program DIANA and the supporting programs CALIBA, HABAS and GLOMSA. *J. Mol. Biol.* **217**, 517–530.
 43. Güntert, P. & Wüthrich, K. (1991) Improved efficiency of protein structure calculations from NMR data using the program DIANA with redundant dihedral angle constraints. *J. Biomol. NMR* **1**, 447–456.
 44. Güntert, P., Braun, W., Billeter, M. & Wüthrich, K. (1989) Automated stereospecific ^1H NMR assignments and their impact on the precision of protein structure determinations in solution. *J. Am. Chem. Soc.* **111**, 3997–4004.
 45. Güntert, P., Berndt, K. & Wüthrich, K. (1993) The program ASNO for computer-supported collection of NOE upper distance constraints as input for protein structure determination. *J. Biomol. NMR* **3**, 601–606.
 46. Wüthrich, K., Billeter, M. & Braun, W. (1983) Pseudo-structures for the 20 common amino acids for use in studies of protein conformations by measurements of intramolecular proton-proton distance constraints with nuclear magnetic resonance. *J. Mol. Biol.* **169**, 949–961.
 47. Braun, W. & Go, N. (1985) Calculation of protein conformations by proton-proton distance constraints. a new efficient algorithm. *J. Mol. Biol.* **186**, 611–626.
 48. Brünger, A.T. (1992) *The X-PLOR Software Manual, Version 3.1*. Yale University Press, New Haven, CT.
 49. Chou, K.C. & Carlacci, L. (1991) Simulated annealing approach to the study of protein structures. *Protein Eng.* **4**, 661–667.
 50. Brooks, B., Bruccoleri, R., Olafson, B.O., States, D.J., Swaminathan, S. & Karplus, M. (1983) CHARMM: a program for macromolecular energy, minimization and dynamics calculations. *J. Comp. Chem.* **4**, 187–217.
 51. Hutchinson, E.G. & Thornton, J.M. (1996) PROMOTIF: a program to identify and analyze structural motifs in proteins. *Protein Sci.* **5**, 212–220.
 52. Laskowski, R.A. (1993) PROCHECK: a program to check the stereochemical quality of protein structures. *J. Appl. Cryst.* **26**, 283–291.
 53. Laskowski, R.A., Rullmann, J.A., MacArthur, M.W., Kaptein, R. & Thornton, J.M. (1996) AQUA and PROCHECK-NMR: programs for checking the quality of protein structures solved by NMR. *J. Biomol. NMR* **8**, 477–486.
 54. Williamson, M.P. (1990) Secondary-structure dependent chemical shifts in proteins. *Biopolymers* **29**, 1423–1431.
 55. Williamson, M.P. & Asakura, T. (1993) Empirical comparisons of models for chemical-shift calculation in proteins. *J. Magn. Res. B* **101**, 63–71.
 56. Koradi, R., Billeter, M. & Wüthrich, K. (1996) MOLMOL: a program for display and analysis of macromolecular structures. *J. Mol. Graphics* **14**, 51–55.
 57. Lee, B. & Richards, F.M. (1971) The interpretation of protein structures: estimation of static accessibility. *J. Mol. Biol.* **55**, 379–400.
 58. Weiner, S.J., Kollman, P.A., Case, D.A., Singh, U.C., Ghio, C., Alagona, G., Profeta, S. & Weiner, P. (1984) A new force field for molecular mechanical simulation of nucleic acids and proteins. *J. Am. Chem. Soc.* **106**, 765–784.
 59. Zielenkiewicz, P. & Saenger, W. (1992) Residue solvent accessibilities in the unfolded polypeptide chain. *Biophys. J.* **63**, 1483–1486.
 60. Pearson, W.R. & Lipman, D.J. (1988) Improved tools for biological sequence comparison. *Proc. Natl Acad. Sci. USA* **85**, 2444–2448.
 61. Thompson, J.D., Higgins, D.G. & Gibson, T.J. (1994) CLUSTAL W: improving the sensitivity of progressive multiple sequence alignment through sequence weighting, position-specific gap penalties and weight matrix choice. *Nucl. Acids Res.* **22**, 4673–4680.
 62. Wüthrich, K., Wider, G., Wagner, G. & Braun, W. (1982) Sequential resonance assignments as a basis for determination of spatial protein structures by high resolution proton nuclear magnetic resonance. *J. Mol. Biol.* **155**, 311–319.
 63. Ramachandran, G.N. & Sassiakaran, V. (1968) Conformation of polypeptides and proteins. *Adv. Protein Chem.* **28**, 283–437.
 64. Parker, M.H. & Hefford, M.A. (1997) A consensus residue analysis of loop and helix-capping residues in four- α -helical-bundle proteins. *Protein Eng.* **10**, 487–496.
 65. Doig, A.J., MacArthur, M.W., Stapley, B.J. & Thornton, J.M. (1997) Structures of N-termini of helices in proteins. *Protein Sci.* **6**, 147–155.
 66. Williamson, M.P., Havel, T.F. & Wüthrich, K. (1985) Solution conformation of proteinase inhibitor IIA from bull seminal plasma by ^1H nuclear magnetic resonance and distance geometry. *J. Mol. Biol.* **182**, 295–315.
 67. Takishima, K., Watanabe, S., Yamada, M., Suga, T. & Mamiya, G. (1988) Amino acid sequences of two nonspecific lipid-transfer proteins from germinated castor bean. *Eur. J. Biochem.* **177**, 241–249.
 68. Bernstein, F.C., Koetzle, T.F., Williams, G.J., Meyer, E.E. Jr, Brice, M.D., Rodgers, J.R., Kennard, O., Shimanouchi, T. & Tasumi, M. (1977) The Protein Data Bank: a computer-based archival file for macromolecular structures. *J. Mol. Biol.* **112**, 535–542.
 69. Simorre, J.P., Caille, A., Marion, D., Marion, D. & Ptak, M. (1991) Two- and three-dimensional ^1H NMR studies of a wheat phospholipid transfer protein: sequential resonance assignments and secondary structure. *Biochemistry* **30**, 11600–11608.
 70. Petit, M.C., Sodano, P., Marion, D. & Ptak, M. (1994) Two-dimensional ^1H -NMR studies of maize lipid-transfer protein. Sequence-specific

- assignment and secondary structure. *Eur. J. Biochem.* **222**, 1047–1054.
71. Barioch, A. & Boeckmann, B. (1993) The SWISS-PROT protein sequence data bank, recent developments. *Nucl. Acids Res.* **21**, 3093–3096.
72. Morozov, V.N. & Morozova, T. (1986) Thermal motion of whole protein molecule in proteins solids. *J. Theor. Biol.* **121**, 73–88.
73. Bairoch, A. & Apweiler, R. (1996) The SWISS-PROT protein sequence data bank and its new (Suppl.)TREMBL. *Nucl. Acids Res.* **24**, 21–25.
74. Baud, F., Pebay-Peroula, E., Cohen Addad, C., Odani, S. & Lehmann, M.S. (1993) Crystal structure of hydrophobic protein from soybean: a member of a new cysteine rich family. *J. Mol. Biol.* **231**, 877–887.
75. Kocher, J.P., Prevost, M., Wodak, S.J. & Lee, B. (1996) Properties of the protein matrix revealed by the free energy of cavity formation. *Structure* **4**, 1517–1529..