

Computational Study of Ligand Binding in Lipid Transfer Proteins: Structures, Interfaces, and Free Energies of Protein-Lipid Complexes

Luis F. Pacios,^{*,[a]} Cristina Gómez-Casado,^[b] Leticia Tordesillas,^[b] Arantxa Palacín,^[b] Rosa Sánchez-Monge,^[b] and Araceli Díaz-Perales^[b]

Plant nonspecific lipid transfer proteins (nsLTPs) bind a wide variety of lipids, which allows them to perform disparate functions. Recent reports on their multifunctionality in plant growth processes have posed new questions on the versatile binding abilities of these proteins. The lack of binding specificity has been customarily explained in qualitative terms on the basis of a supposed structural flexibility and nonspecificity of hydrophobic protein-ligand interactions. We present here a computational study of protein-ligand complexes formed between five nsLTPs and seven lipids bound in two different ways in every receptor protein. After optimizing geometries in

molecular dynamics calculations, we computed Poisson-Boltzmann electrostatic potentials, solvation energies, properties of the protein-ligand interfaces, and estimates of binding free energies of the resulting complexes. Our results provide the first quantitative information on the ligand abilities of nsLTPs, shed new light into protein-lipid interactions, and reveal new features which supplement commonly held assumptions on their lack of binding specificity. © 2012 Wiley Periodicals, Inc.

DOI: 10.1002/jcc.23012

Introduction

Since their discovery 35 years ago,^[1] lipid transfer proteins (LTPs) have been isolated from a variety of organisms from bacteria and yeast to higher plants and animals.^[2] Some LTPs are specific, whereas others exhibit an affinity for a broad range of lipids. Nonspecific LTPs constitute a family of proteins present in up to 50 plant species^[3] and are divided into two subfamilies: nsLTP1 (~ 9 kDa) and nsLTP2 (~ 7 kDa).^[4,5] nsLTP1 proteins from various sources are well characterized whereas the structure and functions of nsLTP2 are still poorly understood.^[4] For this reason, 'nsLTP' is commonly used as synonymous with 'nsLTP1,' convention also followed here. Although there is a large literature on these proteins and their possible functions have been largely debated over the years (for a review see Salcedo et al.^[6]), new roles for nsLTPs are still being proposed. Recent reports on the functions of nsLTPs in cuticular wax deposition^[7] or the suggestion that they should be multifunctional in plant growth and organ development^[8] have posed new questions on their roles in plants.

A number of nsLTP structures have been experimentally determined by either X-ray crystallography or NMR spectroscopy. These proteins share an α -helical compact fold composed of four α -helices connected by short loops and a nonstructured C-terminal tail. This fold is stabilized by four disulfide bridges between cysteines conserved in all plant nsLTPs. The inner space left within the helices forms a characteristic tunnel-like cavity that follows the long axis of the protein. Hydrophobic side chains line the surface of this cavity which is large enough to accommodate ligands with long chains. It has been customarily assumed that big size and high plasticity of the cavity should be major factors to explain the nonspecificity of these proteins to bind ligands.

NsLTPs are known to bind a broad range of lipids including fatty acids, fatty acyl-CoA, phospholipids, glycolipids, hydroxylated fatty acids, and prostaglandin B₂.^[4,9–18] Crystal structures of maize nsLTP complexed with palmitate^[10] and with an array of fatty acids ranging from C10 to C18,^[11] wheat nsLTP in complex with lyso-myristoyl-phosphatidyl-choline (LMPC),^[12] rice nsLTP complexed with stearate,^[13] and peach nsLTP with bound laurate^[14] have been determined by X-ray crystallography. Solution NMR structures of wheat nsLTP complexed with 1,2-dimyristoyl-phosphatidyl-glycerol^[15] and with prostaglandin B₂,^[16] and barley nsLTP in complex with palmitoyl coenzyme A^[17] and palmitate^[18] have also been reported. In all these complexes, the hydrocarbon tail of the bound lipids is inserted into the hydrophobic cavity. However, a somewhat surprising finding that has gone largely unnoticed is that the solvent-exposed charged head protruding from the cavity can display two different orientations inside the tunnel. In fact, while all X-ray structures have the lipid charged heads located at the C-terminal end of the tunnel, some NMR structures exhibit the

[a] L. F. Pacios

Unidad de Química y Bioquímica, Departamento de Biotecnología, E.T.S.I. Montes, Universidad Politécnica de Madrid, Madrid 28040, Spain
E-mail: luis.fpacios@upm.es

[b] C. Gómez-Casado, L. Tordesillas, A. Palacín, R. Sánchez-Monge, A. Díaz-Perales

Unidad de Bioquímica, Departamento de Biotecnología, E.T.S.I. Agrónomos and Centro de Biotecnología y Genómica de Plantas, CBGP, Universidad Politécnica de Madrid, Madrid 28040, Spain

Contract/grant sponsors: Subdirección General de Proyectos de Investigación, Spanish Government Training Grants (MEC-FPU and FPI programs). Contract/grant sponsor: Ministerio de Ciencia e Innovación; contract grant number: BIO2009-07050.

© 2012 Wiley Periodicals, Inc.

opposite orientation. Furthermore, in one case (X-ray wheat nsLTP-LMPC complex) two lipid molecules are bound together showing both orientations simultaneously^[12] and in another case (NMR wheat nsLTP-prostaglandin B₂ complex) the ligand is fully buried in the cavity.^[16]

Although it has been argued that the flexibility of the lipid-binding cavity as well as the nonspecific nature of van der Waals protein-ligand interactions might account for the lack of binding specificity of nsLTPs,^[11] the nature of the protein-ligand interface and the lipid orientation issue have not been explored. Most of our understanding on structural effects in protein interactions comes from X-ray crystallography although several problems arise when trying to derive biologically meaningful information on protein interfaces from crystals.^[19,20] Besides effects associated to a particular biological environment, a crystal of a protein-ligand complex is a system actually different from the same complex in solution. In this regard, while the space groups of nsLTP-ligand crystal complexes available are different (*P*₂₁₂₁ for maize and wheat complexes, *P*₆₅₂₂ for the peach complex and *C*₂ for the rice complex), all of them agree in displaying the same orientation for the bound lipid. However, it is noted that the two NMR solution structures available for a same nsLTP (barley) complexed with either palmitoyl CoA^[17] or palmitate^[18] happen to show opposite orientations. This observation, which has been until now mostly ignored, is the major issue addressed in this work.

Recent reports on new possible roles of LTPs^[7,8] have highlighted the presence of these small proteins in extracellular matrix (ECM) regions where disparate molecules with negatively charged moieties are abundant.^[8] NsLTPs have net positive charge and generate electrostatic potentials predominantly positive (see below), which must be related with their distribution in the ECM. The possible multifunctional roles of nsLTPs on plant growth processes should be related with their ability to bind a broad variety of lipids. Furthermore, the versatility of these small proteins in binding ligands of considerably different size and distinct chemical nature is interesting on its own to analyze protein-ligand interactions. We report here a computational study of complexes formed between nsLTPs and lipids selected to give a representative sample of ligands. In each case, the complexes corresponding to two opposite orientations of the lipid inserted into the protein cavity are investigated. Given that all geometries were optimized in molecular dynamics (MD) calculations in the presence of water and salt ions, our results assess the differences between solvent and crystal structures available for nsLTP-lipid complexes. We have also computed polar solvation free energies, properties of protein-lipid interfaces, and estimates of binding free energies providing thus the first quantitative data on the nature of nsLTP-lipid interactions.

Methods

Initial structures of nsLTP-lipid complexes

The initial geometries for the five nsLTP proteins used in this study were taken from the following experimental structures. (i) peach nsLTP with laurate, PDB entry 2ALG,^[14] (ii) maize

nsLTP with palmitate, PDB entry 1MZM,^[10] (iii) wheat nsLTP with LMPC, PDB entry 1BWO,^[12] (iv) rice nsLTP with stearate, PDB entry 1UVC,^[13] and (v) barley nsLTP with palmitate, PDB entry 1BE2.^[18] Structures (i)-(iv) are X-ray crystal geometries whereas (v) is a solution NMR geometry.

We selected seven lipids with initial geometries taken from the following experimental structures. (i) Caprate (CAP) from its complex with maize nsLTP, PDB entry 1FK0,^[11] (ii) laurate (LAU) from peach nsLTP structure, PDB entry 2ALG,^[14] (iii) palmitate (PLM) from its complex with maize nsLTP, PDB entry 1MZM,^[10] (iv) 9,12-linoleate (EIC) from its complex with bovine β -lactoglobulin, PDB entry 3QZK,^[21] (v) α -linolenate (LNL) from its complex with maize nsLTP, PDB entry 1FK6,^[11] (vi) prostaglandin B₂ (PG2) from its complex with wheat nsLTP, PDB entry 1CZ2,^[16] and (vii) lysomyristoylphosphatidyl-choline (LPC) from its complex with wheat nsLTP, PDB entry 1BWO.^[12] Except (vi), which is a solution NMR structure, the geometries correspond to X-ray crystal structures.

Initial geometries for the 35 nsLTP-lipid complexes in orientation A (charged heads of lipids located at the cavity entry near the protein C-terminal) were prepared by (1) structural superpositions of each complex from the preceding list (except for EIC) and the remaining nsLTPs and (2) interchanging the proteins. Because of the similarity between EIC (for which no experimental structures of complexes with nsLTPs are available) and LNL lipids, initial geometries for the nsLTP-EIC complexes were obtained from the LNL complexes by just interchanging the ligands. Initial geometries for the 35 complexes in orientation B (charged head of lipids located at the cavity entry opposite to orientation A) were constructed taking as a "scaffold" the solvent NMR structure of barley nsLTP-palmitate complex which has this orientation (PDB entry 1BE2^[18]). Using this scaffold structure, every nsLTP was superposed with the barley protein and then every lipid was prepositioned to match the orientation of the charged head of palmitate and subjected to a first round of steepest-descent minimizations *in vacuo* to remove steric clashes. Protein superpositions were done with Swiss-PdbViewer 4.0.4^[22] and ligand prepositioning as well as initial optimizations were carried out with UCSF Chimera 1.5.3.^[23]

Sequence alignments, amino acid similarities, and RMS distances

Sequence identities were based upon Clustal 2.1 pairwise alignments using the ClustalW2 server (www.ebi.ac.uk/Tools/msa/clustalw2). However, since we focus on protein-ligand interactions, we used a similarity criterion based upon physico-chemical properties of amino acid side chains instead of the similarity relationships employed in alignment methods. For instance, both BLOSUM and Gonnet weight matrices mark H/Q, H/N, or A/S pairs as similar (even scoring > 0.5 in weight matrices). By grouping amino acids into the four categories charged positive, charged negative, polar, and nonpolar, we considered then two residues "similar" only if they belong to one of the following pairs: A/V, A/G, I/V, I/L, L/V, R/K, T/S, N/Q, H/K, H/R, and D/E. Root mean square (RMS) distances computed with C α atoms were obtained for structural superpositions with the CEALIGN method^[24] implemented in PyMOL 1.4 (pymol.org).^[25] Assignment of secondary structure was determined with DSSP 2.0.^[26]

Molecular dynamics optimizations of nsLTP-lipid complexes

Geometry optimization of nsLTP-lipid complexes in both orientations (70 complexes in all) were performed with VMD 1.9^[27] and NAMD^[28] using the CHARMM force field.^[29] Because non-standard force field parameters were needed for lipids, atomic charges from AM1-BCC calculations^[30] were first added with the AnteCHAMBER module^[31] (used also to add hydrogens) of UCSF Chimera 1.5.3.^[23] Then, van der Waals CHARMM parameters were generated using the SwissParam (swissparam.ch) server.^[32] Standard CHARMM parameters were used for proteins. The initial structures were then immersed in a periodic rectangular solvation box containing 4762 water molecules (added according to the TIP3P water model^[33]), 23 Na⁺ ions and a number of Cl⁻ ions in excess to counter the total positive charge of the complexes. The particle mesh Ewald summation method^[34] was used for long-range electrostatics and a 10 Å cutoff was selected for short-range nonbonded interactions. A first optimization of every system during 5000 minimization steps of the conjugate gradient method implemented in NAMD was done. Steric clashes between ligands and proteins arising from prepositioning stages were removed during the first steps. With all the atoms in the protein-ligand complex fixed, water was then equilibrated during 100 ps using 2 fs time steps at 300 K and 1 atm. Langevin dynamics for temperature control and Nosé-Hoover Langevin piston method for pressure control were employed. Finally, a reoptimization of the local geometry within 8 Å of the lipid was performed monitoring the process with the AutoMD tool implemented in VMD 1.9, stopping the minimization when the two first decimal places of the total energy were stabilized, which led to gradient values about 0.1 kcal mol⁻¹ Å⁻¹.

Test molecular dynamics and docking calculations of two nsLTP-lipid complexes

With the aim to analyze the sensitivity of the methods used to compute binding energies and interface properties (explained below) to input structures, two systems were selected for obtaining new geometries: peach nsLTP-EIC and wheat nsLTP-PG2 complexes in both orientations A and B. Starting at the corresponding MD-optimized structures, test MD calculations consisting of (1) 5000 minimization steps, (2) equilibration of water during 100 ps, and (3) simulation runs during 4 ns were carried out. Same details on solvation box, added ions, treatment of long range electrostatics, temperature and pressure control methods, and time step indicated in the preceding subsection were used in these calculations also performed with NAMD. Final coordinates after these simulation runs were a first set of new structures dubbed "MD-sim" below. A second set of structures (labeled "MD-frame" below) were then selected by just picking the coordinates of the middle frame in these simulations.

To obtain new different input structures, we used a docking procedure to prepare starting geometries for a new set of MD calculations. To this end, we docked the EIC lipid to peach nsLTP and the PG2 ligand to wheat nsLTP using AutoDock Tools 1.5.6^[35] to prepare protein and ligand molecules as well as space grid definitions and then AutoDock Vina^[36] to obtain the docked conformations. As discussed below, this procedure resulted in ligand geometries that split in two sets of similar conformations corresponding to orientations A and B with

nearly identical affinity energies in both peach and wheat complexes. Upon selecting the docked conformation with best score among those in A and B groups as starting structures, a new set of MD (1)-(2)-(3) calculations were then carried out. Final coordinates after these new simulations were then the third set of input structures dubbed "MD-dock" below. Results of these simulations were analyzed with the vmdICE^[37] tool implemented in VMD 1.9.

Poisson-Boltzmann polar solvation free energies and electrostatic potentials

Electrostatic polar solvation free energies G_{polsv} and electrostatic potentials were computed by solving the Poisson-Boltzmann (PB) equation with APBS 1.3^[38] using AMBER99^[39] charges and radii assigned with PDB2PQR 1.7.^[40] The nonlinear PB equation was solved in sequential focusing multigrid calculations in 3D meshes of 192³ points (spatial grids with step size about 0.3 Å) at 298.15 K and 0.150 M ionic concentration. Dielectric constants 4 for proteins and 78.54 for water were used. Calculations for complexes and monomers were performed with identical grid spacings, lengths, and centers to ensure appropriate matching of self-energy terms. Changes of electrostatic polar solvation free energies upon complexation, ΔG_{polsv} were then computed as $\Delta G_{\text{polsv}} = G_{\text{polsv}}^{\text{complex}} - (G_{\text{polsv}}^{\text{LTP}} + G_{\text{polsv}}^{\text{lipid}})$.

PB electrostatic potentials were calculated with APBS 1.3 using above options except that 3D meshes composed of 161³ points were now selected to obtain the numerical output in scalar OpenDX format. These meshes were then mapped onto protein molecular surfaces and rendered with PyMOL 1.4.^[25] PB electrostatic potential values are given in units of kT per unit charge (k , Boltzmann's constant; T , absolute temperature)

Properties of nsLTP-lipid interfaces

Solvation free energy gain upon formation of the interface ΔG_{solv} p -value of the observed solvation free energy gain of the interface P_{int} and interface area A_{int} for optimized geometries of the complexes were computed with Protein Interfaces, Surfaces and Assemblies (PISA) software^[19,20] using the EBI-PISA server (www.ebi.ac.uk/msd-srv/prot_int/pistart.html). The meaning of PISA parameters used in this work is explained in Results and Discussion section. PISA is widely used to analyze macromolecular interfaces in both protein-protein and protein-ligand complexes and to extract biologically relevant information on protein interactions from complex structures. Although PISA has been recently introduced, it has become an essential tool in proteomics databases and servers for the analysis of protein interactions (see reviews in Refs. [41] and [42]) due in great part to the success of PISA predictions in recent rounds of CAPRI (Critical Assessment of PRediction of Interactions) experiments (www.ebi.ac.uk/msd-srv/capri/). For our purposes, given that all nsLTP-lipid structures correspond to computed geometries and hence no crystallographic information (cell parameters and space symmetry group) exists, PISA calculations were restricted to protein-ligand interface analysis.

Estimates of binding free energies

Binding free energies ΔG_{bind} were estimated with the all-atom energy based empirical scoring function^[43,44] implemented in

the Binding Affinity Prediction of Protein-Ligand, BAPPL server (www.scfbio-iitd.res.in/software/drugdesign/bappl.jsp) using "Method 2" mode calculation, i.e., the net charge on the ligand is specified in input and the server assigns the parameters needed. For the ligand, BAPPL obtains partial atomic charges from AM1-BCC calculations^[30] and assigns van der Waals parameters from GAFF force field.^[45] For proteins, both atomic charges and van der Waals parameters are assigned using a protein-adapted AMBER force field.^[46] After investigating the sensitivity to force field choice of the results on a heterogeneous dataset of 161 complexes, an empirical scoring function consisting of 25 independent variables (electrostatic, van der Waals, loss of conformational entropy and atom types for hydrophobicity) was developed.^[43,44] nsLTP-LPC complexes are omitted because BAPPL failed to give binding energies for all of them (note that this lipid is the only positively charged ligand and all nsLTPs have net positive charge).

Results and Discussion

Plant nsLTPs and lipids studied

Five plant nsLTPs were selected because experimental structures of complexes with a variety of lipids are available: peach, maize, wheat, rice, and barley. Amino acid sequence identities are between 44 and 79% whereas similarities based upon side chain chemical features vary between 9 and 15% (Table 1). Adding both identity and similarity percentages to quantity sequence conservation, maize and rice show a maximum overall 88% conservation whereas peach and rice have a minimum 59% conservation. 32 identical and 14 similar amino acids are found in the five nsLTPs (Fig. 1A), which amounts to 50% sequence conservation. It should be noted that these percentages are far more restrictive than those obtained with the rules used in conventional alignment procedures (see Methods). For instance, the 15.4% similarity between peach and barley proteins in Table 1 would be 24.2% if one had used the information provided by Gonnet and BLOSUM matrices upon aligning their sequences.

However, structures are much more similar. These proteins are characterized by a typical fold composed of four α -helices and a nonstructured C-terminal tail. While the overall fold is well conserved and the helices are nearly coincident, interhelical loops and particularly the long C-terminal tail display greater differences (Fig. 1B). Small RMS distances between 1.0 and 1.4 Å are observed except for barley that displays systematically greater values between 1.8 and 2.2 Å (Table 1). The superposition of the five structures together shows that the

tunnel-like cavity extends over a nearly identical space which runs parallel to both the longest helix H3 and the C-terminal tail (Fig. 1B). The two orientations for lipids inserted into the cavity correspond to their charged head pointing towards either the C-terminal end (orientation A) or the loop joining helices H3 and H4 (orientation B, Fig. 1B). If one considers (a) the small size of these proteins, (b) the structural coincidence of conserved helices in these superpositions, and (c) that the C-terminal long tail forms the greatest part of the cavity, it seems reasonable to expect that this unstructured segment might be a major factor in adapting the cavity to accommodate ligands of variable size, irrespective of their orientation.

To represent the diversity of lipids considered possible ligands to nsLTPs, we selected the following seven molecules (3-letter keys used hereafter to denote them are indicated after their names): (i) caprate, CAP; (ii) laurate, LAU; (iii) palmitate, PLM; (iv) 9,12-linoleate, EIC; (v) α -linolenate, LNL; (vi) prostaglandin B₂, PG2; and (vii) LMPC, LPC (Fig. 2). We used their charged forms, anionic charge -1 for (i)-(vi) and cationic charge +1 for (vii). Lipids (i)-(iii) were chosen to sample fatty acids of increasing chain length, (iv) and (v) are two unsaturated fatty acids with equal number of carbons that differ in their double bonds, and (vi) and (vii) are lipids with rather different geometries that include polar bonds and other oxygen atoms besides the carboxylate group. Only one cationic lipid was included because most lipids known to be potential ligands to nsLTPs are anionic.

Optimized structures of complexes in two possible orientations of the lipid

Structural differences between complexes of a given nsLTP arise nearly exclusively from interhelical loops and C-terminal tails (Fig. 3). In orientation A, peach, maize, and rice agree in having fatty acids in a rather similar geometry inside the cavity whereas wheat and barley exhibit a greater variety which is also noticed in the position of charged heads outside the tunnel. In orientation B, the ligands show a tighter spatial location than in orientation A. It seems that placing lipids inside the tunnel with their polar heads pointing towards the H3-H4 loop forces occupation of a more restricted inner space than in the opposite orientation. As for the LPC lipid, a large part of its long chain is necessarily outside the protein structure but the trimethyl ammonium charged group occupies similar outer regions at all nsLTPs (Fig. 3).

Molecular surfaces reveal dramatical changes of tunnel entries in a same nsLTP depending on the bound lipid (Fig. 4).

Furthermore, it is found that the size of this entry can decrease drastically leaving the ligand fully embedded in the cavity as it happens with PG2 at both orientations A and B. Complexes in the two lipid orientations may differ considerably in the precise location of tunnel ends at the protein surface. Whereas

Table 1. Percentage of sequence identity (*Ide*) and chemical similarity (*Sim*) and RMS distances (Å) between corresponding C α atoms of experimental structures of nsLTPs (PDB codes in parentheses).

	Peach (2ALG)			Maize (1MZM)			Wheat (1BWO)			Rice (1UVC)		
	<i>Ide</i>	<i>Sim</i>	RMS	<i>Ide</i>	<i>Sim</i>	RMS	<i>Ide</i>	<i>Sim</i>	RMS	<i>Ide</i>	<i>Sim</i>	RMS
Maize (1MZM)	62.6	8.8	1.33									
Wheat (1BWO)	45.6	13.3	0.95	58.9	8.9	1.35						
Rice (1UVC)	52.7	13.2	1.30	79.1	8.8	1.12	61.1	10.0	1.24			
Barley (1BE2)	44.0	15.4	2.12	56.0	13.2	2.22	72.2	8.9	2.10	62.6	11.0	1.81

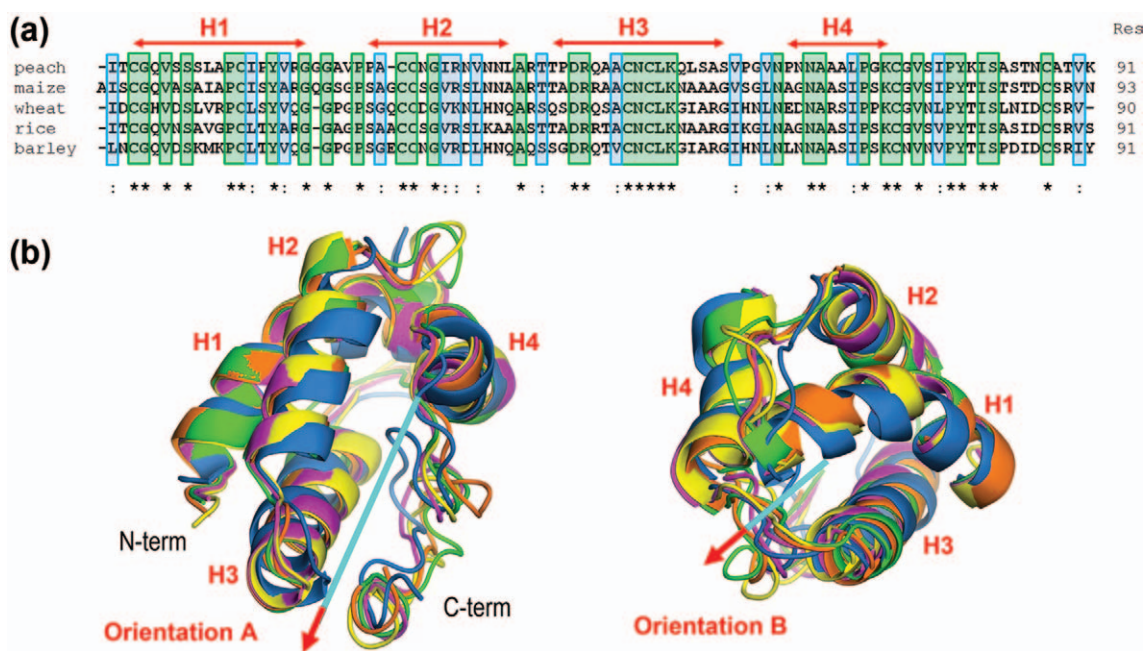


Figure 1. a) Multiple alignment of nsLTPs. Green boxes enclose fully conserved residues. Blue boxes enclose chemically similar residues in conserved positions. Red arrows indicate sequence segments forming the four α -helices (peach numbering): H1 (3-19), H2 (25-37), H3 (41-57), and H4 (63-72). b) Ribbon diagram of structural superposition of peach (green), maize (yellow), wheat (orange), rice (magenta), and barley (blue) nsLTPs. Cyan lines and red arrows sketch hydrophobic chains and charged heads of bound lipids, respectively, at orientations A and B. Image of orientation B is generated from that of orientation A upon a 180° rotation around a vertical axis followed by a 45° rotation around a horizontal axis.

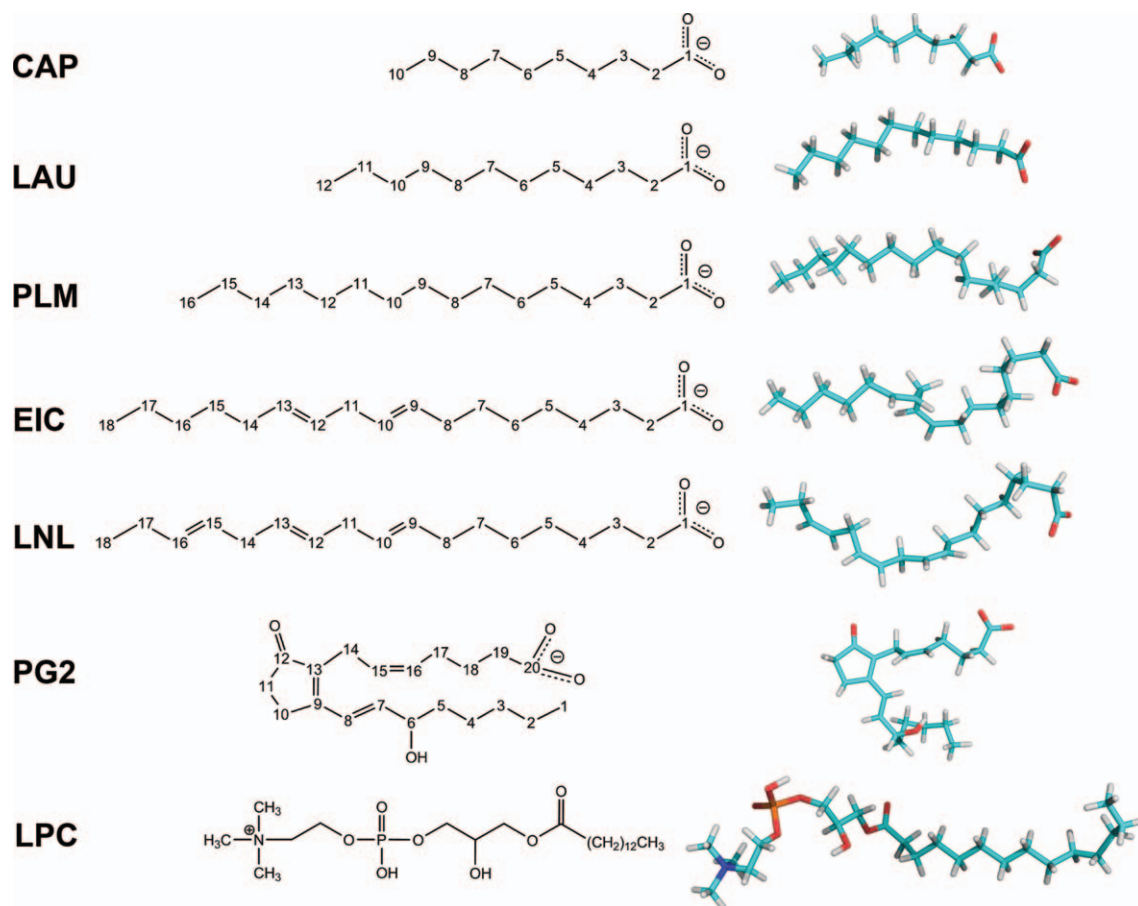


Figure 2. Lipids selected as nsLTP ligands. CAP: caprate, LAU: laurate, PLM: palmitate, EIC: 9,12-linoleate, LNL: α -linolenate, PG2: prostaglandin B₂, LPC: lysomiristoyl-phosphatidylcholine. Sticks diagrams depict initial geometries taken from experimental structures for complexes with different plant nsLTPs except EIC which pertains to a complex with bovine β -lactoglobulin.

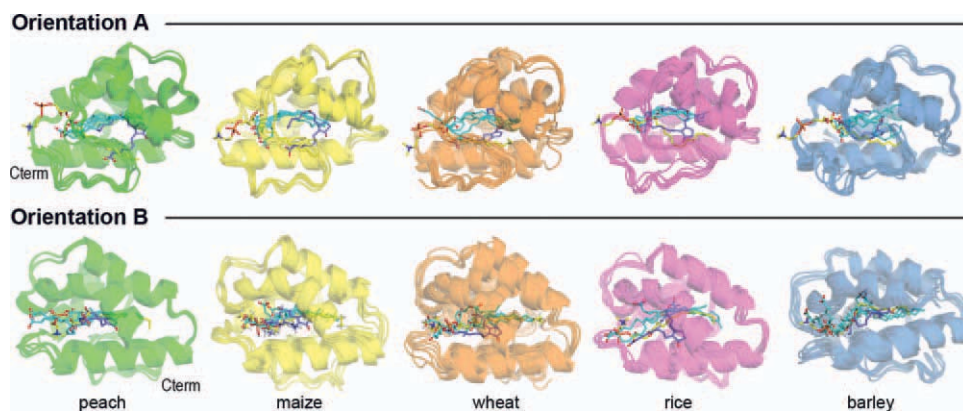


Figure 3. Optimized MD structures of nsLTP-lipid complexes. Superposition of ribbon diagrams of complexes with bound lipids drawn as sticks in orientation A (upper row) and B (lower row). Carbons are colored cyan in CAP, LAU, PLM, EIC, and LNL, violet in PG2 and yellow in LPC. Oxygens are colored red. Nitrogen and phosphorus in LPC are colored blue and orange, respectively. [Color figure can be viewed in the online issue, which is available at wileyonlinelibrary.com.]

the entry is situated at nearly the same surface region in orientation A (note the location of the opening with respect to helix H3), it occurs at markedly different regions in orientation B (Fig. 4). Although only peach nsLTPs complexes are depicted, similar results were found for the remaining nsLTPs.

Structural changes upon accommodating different ligands were quantitatively evaluated by computing RMS distances for the superposition of each complex with its corresponding uncomplexed nsLTP (Fig. 5). With the only exception of PG2 (see below), RMS deviations computed with backbone atoms are between 0.8 and 1.5 Å for the two orientations. Structural differences given by RMS computed with all atoms are obviously greater because the effect of conformational changes of amino acid side chains are now included. However, note that all-atom values are still below 1.7 Å (except again for PG2). It must be emphasized that these calculations were carried out with all the protein atoms, not only with those selected by the structural alignments achieved upon superimposing the backbone chains. Since RMS values computed this way magnify structural deviations, our results indicate that these proteins are able to bind rather different lipids in two opposite orientations inside the cavity without significant structural distortion even upon including conformational changes in amino acid side chains. Structural differences are consistently smaller for peach and maize complexes and consistently greater for wheat and barley complexes. In only one case (wheat nsLTP-PG2 complex), ligand binding produces a conformational change of side chains large enough to yield RMS deviations of 3.0 Å in both orientations.

Binding of prostaglandin B₂ has marked differences with other lipids.^[16] Our optimized structures for PG2 complexes show qualitatively distinct features for the bound lipid in peach, maize, and rice nsLTPs on one hand and in wheat and barley nsLTPs on the other hand. Surfaces of peach and wheat complexes as representative examples of these two sets reveal that although PG2 molecule is completely embedded in the cavity, the former set exhibits large openings that leave ring oxygen (orientation A) or carboxylate oxygens (orientation B)

openly accessible to the solvent whereas the latter set shows very small surface openings that barely connect hydroxyl (orientation A) or carboxylate (orientation B) groups with the solvent (Fig. 6). The burial of PG2 occurs in spite of the presence of polar oxygens in intermediate positions (Fig. 2), a finding at odds with the conventionally assumed hydrophobic nature of the inner walls of the tunnel.^[16] It is noted that the optimized structure of wheat nsLTP-PG2 complex in orientation A reproduces accurately the features shown by the experimental structure^[16]: hydroxyl group connected with the sol-

vent through a unique opening in the cavity and buried carboxylate orientated towards the protein C-terminal end (Fig. 6). In fact, when experimental and optimized structures of this wheat nsLTP-PG2 complex are superimposed, lipid molecules are nearly coincident in orientation A whereas they show an opposite spatial arrangement in orientation B (Fig. 6B).

Solvation free energies upon formation of nsLTPs-lipid complexes

The recently developed PISA (Protein, Interfaces, Surfaces, and Assemblies) procedure was primarily intended to identify macromolecular assemblies in crystals using empirical estimates for the binding free energy of complexes.^[19,20] Focusing on interface properties, PISA has proven successful to distinguish between natural interactions and artifacts of crystal packing, a problem that the identification of protein contacts must face. All the analyses and algorithmic approaches employed by PISA may directly be applied to ligands by considering them as any other monomeric unit after introducing some corrections to cope with difficulties arising from the large variety of interface types in protein-ligand interactions.^[19] Interaction with the solvent is a dominant factor in protein-ligand complex formation, hence the leading term in PISA binding free energy is solvation energy gain upon formation of the interface, ΔG_{solv} . Three model energy terms accounting for disulphide bonds, hydrogen bonds, and ionic pairs (salt bridges) formed across the interface are then added to ΔG_{solv} . This solvation energy is assumed to depend on the interface surface area and is calculated at a semiempirical level as a sum of terms that runs over all atom types of the subunits in the interface.^[19,20]

However, given that PISA calculations make no explicit consideration of electrostatic solvation effects, we computed separate polar solvation energies, $\Delta G_{\text{polsolv}}$ using the nonlinear Poisson-Boltzmann (PB) equation, a continuum model for describing electrostatic interactions between macromolecular solutes and salty, aqueous media^[38,47,48] sensitive enough to provide detailed local changes arisen from conformational

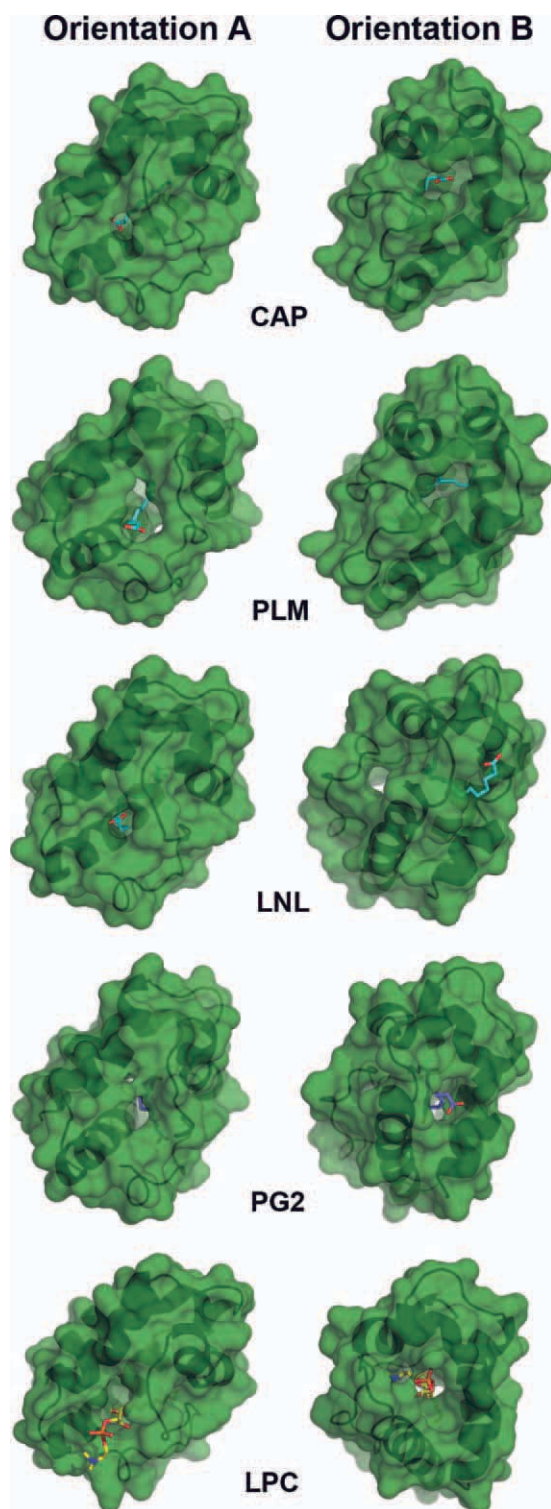


Figure 4. Molecular surfaces of optimized structures of complexes of peach nsLTP with CAP, PLM, LNL, PG2, and LPC lipids in orientations A and B. Atom colors for ligands as in Figure 3. [Color figure can be viewed in the online issue, which is available at wileyonlinelibrary.com.]

variations in proteins.^[49] PB equation relates the electrostatic potential to the dielectric properties of solute and solvent, the ionic strength of the solution and the distribution of atomic partial charges in the solute.^[47,48] Although an implicit solvent

model of nonpolar solvation that provides accurate results for small alkane solutes has been recently introduced as an add-on to PB equation solvers,^[50] uncertainties on the parameters needed have hampered its widespread application to protein-ligand solvation. Hence, we restricted PB calculations on nsLTP-lipid interactions to solvation polar effects. Before proceeding further, it should be recalled that protein-protein complex formation in aqueous media between mostly hydrophilic partners (such as antigen-antibody complexes: see for instance the studies on the electrostatic contributions to antibody interactions in Refs. [51] and [52]) sequesters surfaces from water, solvation is not a favorable effect and polar ΔG_{polsv} energies should be positive. In protein-ligand complex formation, hydrophobicity and small size of the ligand could make solvation a favorable effect and polar ΔG_{polsv} energies can be negative. However, in nsLTP-lipid complexes the burial of the ligand inside the cavity introduces a new feature regarding polar solvation effects.

Both solvation free energies, ΔG_{polsv} and ΔG_{solvr} are gathered in Table 2. The main feature shown by ΔG_{polsv} energies is that the majority of nsLTP-lipid complexes in the two orientations have positive values, i.e., polar effects do not play a favorable role in their formation. Other specific trends are

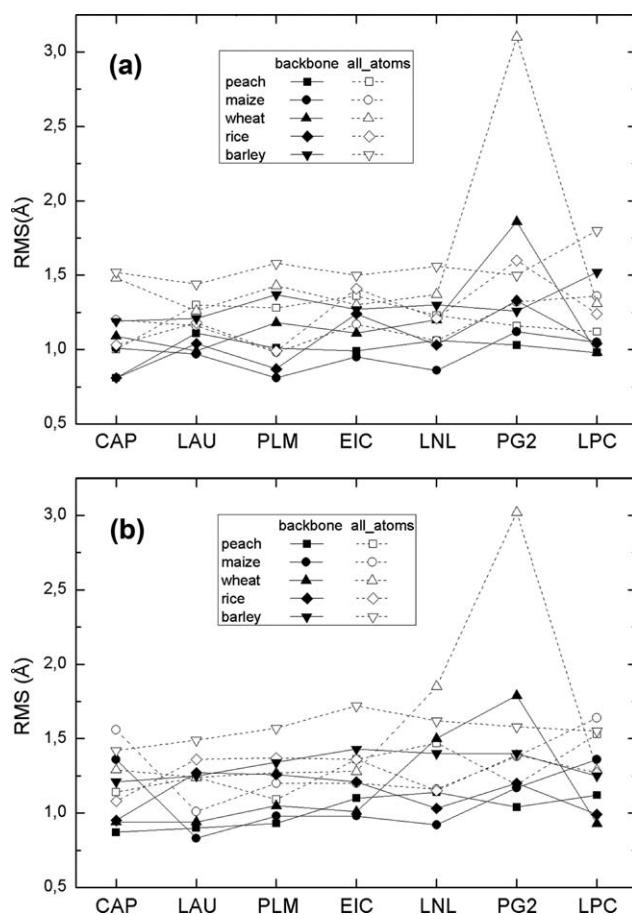


Figure 5. RMS distances computed in backbone structural alignments of nsLTP-lipid complexes with their corresponding uncomplexed nsLTP. a) Lipid inserted in the cavity in orientation A. b) Lipid inserted in the cavity in orientation B.

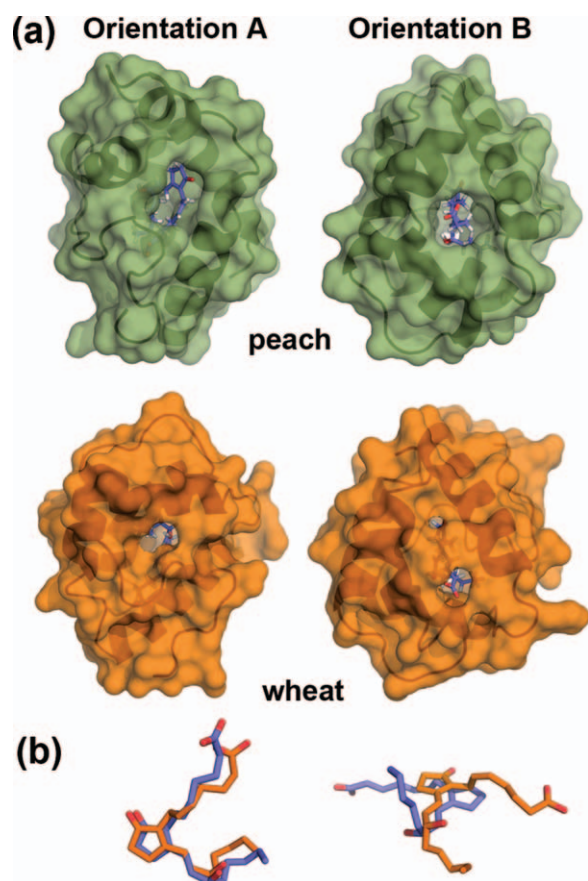


Figure 6. a) Molecular surfaces of optimized structures of complexes of peach and wheat nsLTPs with prostaglandin PG2 in orientations A and B. Atom colors for PG2 as in Figure 3. b) Superposition of PG2 molecules drawn as sticks in the optimized geometry (violet carbons) and the experimental structure^[16] (orange carbons) of wheat nsLTP-PG2 complexes.

noticed. (a) Only peach complexes (except LPC) in orientation A display systematic negative ΔG_{polsv} values. (b) LPC shows in most cases the greatest positive values, an understandable result if one considers that its protrusion outside protein (Fig. 3) renders polar solvation a minor effect in complex formation. (c) Among fatty acids, only the smallest one (CAP) shows slightly favorable effects. (d) Polar solvation effects noticeably favor orientation A in peach complexes and orientation B in barley complexes whereas they have no significant effect on the remaining cases.

Solvation free energy gain of the interface ΔG_{solv} is calculated in PISA as difference in solvation energies of all residues between dissociated and associated (interfacing) structures. Hence, a positive value is interpreted as corresponding to a hydrophobic interface.^[19] Because ΔG_{solv} is computed in terms of surface areas buried upon complexation, it includes nonpolar solvation effects providing thus a complementary information to that given by PB free energies. ΔG_{solv} energies in Table 2 exhibit more systematic trends than ΔG_{polsv} energies. In fact, all nsLTP-fatty acid complexes show small positive values in the range 0.1–3.8 kcal mol⁻¹ in agreement with their hydrophobic carbon chains. In contrast, PG2 complexes have negative energies between -0.1 and -2.2 kcal mol⁻¹ while LPC

complexes have much greater energies about -9 kcal mol⁻¹ (orientation A) and -11 kcal mol⁻¹ (orientation B). The different sign with respect to fatty acids is due to the presence of intermediate polar groups in PG2 and LPC (Fig. 2) while the different magnitude of their ΔG_{solv} energies should be explained recalling that whereas PG2 is almost completely buried, LPC protrudes into the solvent. Finally, it is interesting to note that peach, maize, and rice PG2 complexes in orientation A show ΔG_{solv} energies more negative than the remaining PG2 complexes. As discussed earlier, these three nsLTPs exhibit large openings of the cavity that leave the ring oxygen of the ligand openly accessible to water.

Electrostatic potentials of nsLTPs

Electrostatic potential maps provide an invaluable tool to analyze protein properties of interest in protein-ligand interactions, both at long range and at short distances.^[53] Examination of PB electrostatic potentials mapped onto the protein surfaces of nsLTPs reveals differences that may straightforwardly be attributed to the number of basic and acidic charged amino acids in them (Fig. 7). All these proteins have in common a relatively large number of positively charged amino acids but they differ in the negatively charged amino acids. The resulting net charges split nsLTPs into two groups: peach, maize, and rice with net charges between +6 and +8 and wheat and barley both sharing a net charge +2.

Protein surface sides corresponding to the C-terminal end of the tunnel (orientation A) display a dominant positive electrostatic potential in peach, maize, and rice nsLTPs. On the contrary, large regions of neutral and negative potential are seen in this surface side in wheat and barley nsLTPs. Opposite surface sides (orientation B) show a dominance of positive potential that decreases from rice to peach to maize nsLTPs. In contrast, wheat and barley proteins have a different electrostatic

Table 2. Change of PB electrostatic polar solvation free energy upon complexation (ΔG_{polsv}) and PISA solvation free energy gain upon formation of the interface (ΔG_{solv}) for nsLTP-lipid complexes (all values in kcal mol⁻¹).

Lipid	Peach	Maize	Wheat	Rice	Barley
	ΔG_{polsv} ΔG_{solv}	ΔG_{polsv} ΔG_{solv}	ΔG_{polsv} ΔG_{solv}	ΔG_{polsv} ΔG_{solv}	ΔG_{polsv} ΔG_{solv}
Orientation A					
CAP	-6.1, +1.9	+1.2, +2.1	+0.2, +1.4	+0.7, +1.2	+1.3, +2.4
LAU	-1.6, +0.4	+3.5, +0.3	+1.9, +0.1	+1.7, +0.8	+0.6, +0.8
PLM	-0.1, +0.1	+6.5, +0.1	+2.0, +0.3	+1.5, +0.3	+1.3, +1.2
EIC	-4.0, +3.5	+7.8, +1.6	+3.6, +3.2	-1.5, +2.5	+3.4, +3.8
LNL	-3.5, +2.2	+1.1, +1.7	+8.2, +1.8	-1.8, +3.2	+0.9, +3.1
PG2	-2.4, -1.5	+7.1, -2.2	+8.6, -0.1	+3.8, -1.6	+4.0, -0.7
LPC	+6.7, -8.9	+4.2, -9.5	+8.3, -9.0	+7.9, -9.4	+8.6, -8.7
Orientation B					
CAP	+2.0, +1.5	+1.7, +0.4	+6.0, +0.4	+4.6, +2.2	-5.9, +2.6
LAU	+6.0, +0.1	+4.1, +0.2	+10., +0.4	-1.8, +0.6	-7.9, +1.3
PLM	+5.8, +0.6	+2.1, +0.7	+8.1, +0.2	-0.5, +0.2	-5.9, +1.4
EIC	+2.7, +2.1	-3.7, +1.2	+0.3, +2.2	-1.2, +0.7	+3.4, +0.3
LNL	+0.7, +0.5	-4.9, +0.4	+1.4, +0.1	-0.3, +0.1	+2.1, +1.5
PG2	+4.2, -0.4	+0.2, -0.3	+2.7, -0.5	+3.5, -0.1	-0.4, -0.3
LPC	+8.3, -12.	+11., -11.	+5.3, -11.	+7.0, -12.	+6.1, -11.

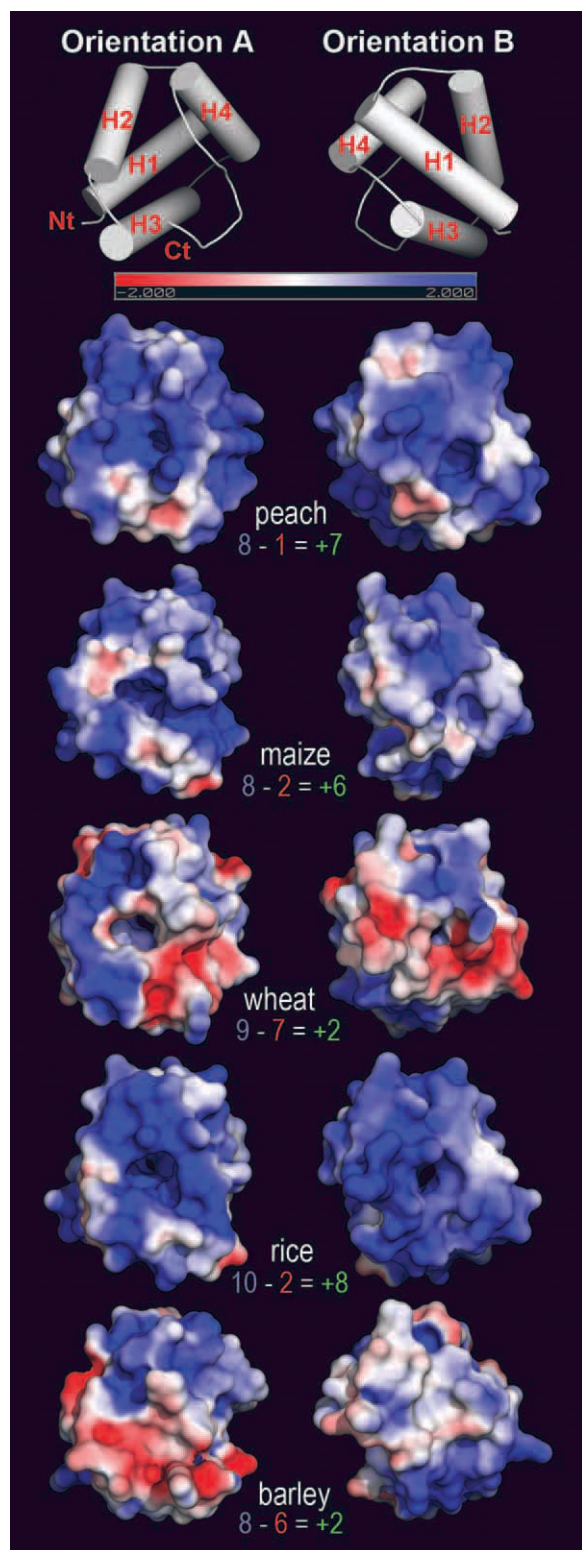


Figure 7. PB electrostatic potentials mapped onto protein surfaces of uncomplexed plant nsLTPs in orientations A and B. Secondary structure cartoons at the upper row indicate the perspective at which surface views are rendered underneath. Below every nsLTP name, the number of positive (blue) and negative (red) charged amino acids as well as the net protein charge (green) are indicated. [Color figure can be viewed in the online issue, which is available at wileyonlinelibrary.com.]

nature, negative in the former and neutral in the latter. Wheat nsLTP has one acidic residue (Asp2) which confers a negative character to this surface side and is absent in barley. These electrostatic features suggest more favorable attractions toward negatively charged ligands in rice and peach, then maize and to a lesser extent, barley and wheat. In either case, with the exception of wheat (particularly in orientation B), both ends of the tunnel are located at surface region with positive or neutral electrostatic potential (Fig. 7).

Spatial isosurfaces of PB potentials cast light on far reaching electrostatic effects arising from these proteins (Fig. 8). Similarly to protein surfaces, the zero PB potential isosurface splits these proteins into two groups: peach, maize, and rice on one side, and wheat and barley on the other side. The former group displays a zero isocontour that encircles tightly most of the protein structure except a protuberance that corresponds to an exposed Lys80 residue in peach and to two neighbor Arg residues (46 and 47 in maize, 44 and 45 in rice). In sharp contrast, the latter group shows zero PB potential isosurfaces that extend far from the protein structure setting two well differentiated spatial domains that leave a large belt of negative potential between (Fig. 8). Since the zero electrostatic potential divides the space around a molecule into positive and negative domains, the zero isosurface that surrounds tightly peach, maize, and rice proteins indicates that the negative space is constrained underneath. In other words, the positive electrostatic potential dominates the whole space around these three nsLTPs whereas it is restricted to just a part of space around wheat and barley nsLTP. Hence, according to these PB isosurfaces, negatively charged ligands should be spatially better directed towards peach, maize, or rice proteins than towards wheat or barley proteins.

Properties of protein-ligand interface in nsLTPs-lipid complexes

The PISA analysis of protein-ligand interfaces also provides P_{int} , a hydrophobic p -value obtained from the observed ΔG_{solv} . P_{int} measures the probability of getting a ΔG_{solv} lower than obtained if interface atoms were picked randomly from protein surface such as to amount to the observed interface area A_{int} . The parameter may be interpreted as the probability to find a same-area patch on the protein surface that would be more hydrophobic than the interface.^[20] Since P_{int} provides thus a measure of interface specificity, $P_{\text{int}} = 0.5$ means that the chances to find surface patches that are more or less hydrophobic than the protein-ligand interface are equal, i.e., the hydrophobic properties of the interface are standard. $P_{\text{int}} > 0.5$ indicates nonspecific binding, i.e., there is no strong preference to any particular geometry among those permitted by structures, which indicates interfaces with low hydrophobicity. On the contrary, $P_{\text{int}} < 0.5$ means interaction-specific interfaces with specific hydrophobic spots likely to play a preferential role in protein-ligand interactions. Limit case $P_{\text{int}} = 0$ indicates that no other interface of the observed area has a lower ΔG_{solv} and the interface is a truly unique spot.^[20]

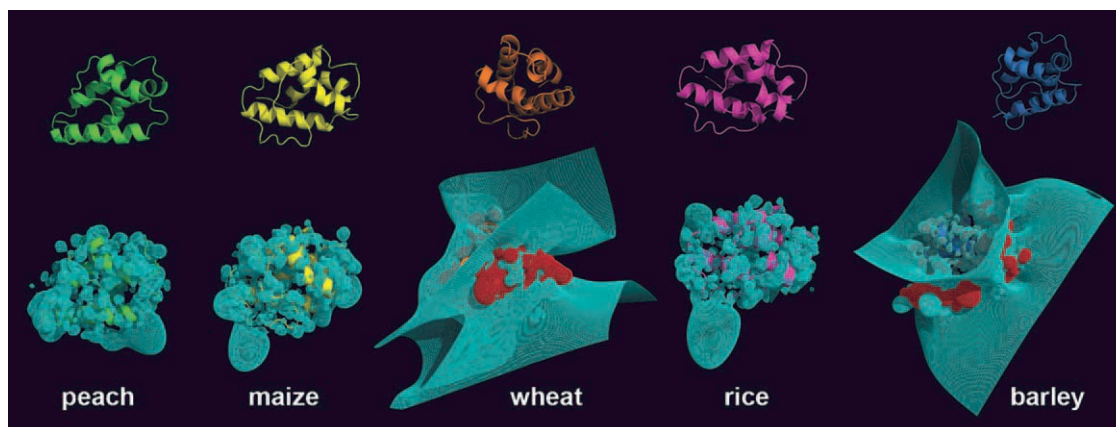


Figure 8. Spatial isosurfaces of PB electrostatic potential zero (cyan grids) and -1 (red grids) values for uncomplexed plant nsLTPs at perspectives drawn in the upper row. Negative -1 isosurfaces lie inside zero isosurfaces for peach, maize, and rice nsLTPs. [Color figure can be viewed in the online issue, which is available at wileyonlinelibrary.com.]

A_{int} interface areas and P_{int} parameters for nsLTP-lipid interfaces are collected in Table 3. The first point to note is that the great majority of interfaces exhibit specific hydrophobic features, i.e., $P_{\text{int}} < 0.5$ values. Only the smaller CAP and LAU lipids present interfaces with low hydrophobicity properties ($P_{\text{int}} > 0.5$), particularly in peach complexes. The longer carbon chain in PLM has interfaces with greater hydrophobic specificity, although large variations are noticed: compare for instance P_{int} values in peach and wheat PLM complexes. Interface areas for the three saturated fatty acids obviously increase with their size and are slightly smaller in orientation B, revealing a minor shrink of carbon chains inside the tunnel in this orientation. Interfaces for unsaturated EIC and LNL lipids are markedly different. Both show the lowest P_{int} values, particularly in orientation B where they have values about 0.1 or lower that indicate highly specific hydrophobic interactions whereas in orientation A, only peach complexes display less specific interfaces (higher P_{int}). Despite their identical chain length, EIC and LNL show

rather distinct features which reveals that the additional double bond at C_{15} in LNL introduces a noticeable effect on the hydrophobicity properties of the interaction with protein. As for the interface areas and in contrast with saturated fatty acids, EIC and LNL complexes show large variations with much more compact interfaces in orientation B (in six out of ten complexes, A_{int} values are 50 \AA^2 smaller than in orientation A). Together with the associated low P_{int} values, these results suggest that nsLTPs bind unsaturated fatty acids forming more unusual interfaces than with other lipids.

Interfaces for PG2 complexes exhibit relatively constant features regarding both interface areas and P_{int} parameter. The nearly complete burial inside the cavity associated to the more compact geometry of PG2 as compared with the remaining lipids yields a uniformly high hydrophobic specificity and almost constant areas for the interfaces of its complexes. Finally, LPC has interfaces with low hydrophobic specificity ($P_{\text{int}} \sim 0.4 - 0.5$) similar to those of CAP-PLM lipids. This is an expected result if one considers that the hydrophobic saturated carbon chain of the phospholipid is inserted into the cavity while its polar and charged moieties are exposed to the solvent. However, A_{int} areas for the LPC interface show clear differences between orientations. In orientation B, LPC occupies a great extent of the protein core as compared with its protrusion into the solvent in orientation A (Fig. 3). Consequently, interface areas are far greater in orientation B.

Estimates of binding energies for nsLTPs-lipid complexes

We list in Table 4 estimates of binding free energies ΔG_{bind} . It must be mentioned that neither experimental binding affinity data exist in the BindingDB (www.bindingdb.org) database^[54] nor, to the best of our knowledge, any other binding affinities had been reported before for nsLTP-lipid complexes. We obtained estimates of ΔG_{bind} with the BAPPL (Binding Affinity Prediction of Protein-Ligand) approach,^[43,44] a computationally fast procedure for predicting binding affinities of nonmetal protein ligand complexes. This method is based upon an all-

Table 3. PISA interface areas (A_{int} in \AA^2) and p -values of observed solvation free energy gain upon formation of the interface (P_{int}) for nsLTP-lipid complexes.

Lipid	Peach		Maize		Wheat		Rice		Barley	
	A_{int}	P_{int}	A_{int}	P_{int}	A_{int}	P_{int}	A_{int}	P_{int}	A_{int}	P_{int}
Orientation A										
CAP	312	0.649	318	0.551	297	0.475	328	0.516	302	0.464
LAU	361	0.708	393	0.490	384	0.303	377	0.564	352	0.517
PLM	414	0.419	435	0.358	416	0.210	413	0.352	386	0.351
EIC	489	0.305	513	0.149	481	0.132	513	0.199	446	0.265
LNL	460	0.327	490	0.252	471	0.141	478	0.321	429	0.262
PG2	474	0.266	466	0.173	519	0.199	488	0.254	456	0.216
LPC	517	0.506	467	0.365	508	0.359	559	0.488	567	0.438
Orientation B										
CAP	324	0.592	321	0.468	305	0.318	309	0.560	306	0.485
LAU	390	0.618	375	0.484	374	0.296	378	0.444	355	0.509
PLM	397	0.498	395	0.342	403	0.238	411	0.307	396	0.351
EIC	435	0.087	462	0.098	463	0.065	447	0.061	424	0.025
LNL	433	0.108	426	0.096	396	0.030	401	0.072	429	0.117
PG2	444	0.274	454	0.202	475	0.146	457	0.224	470	0.203
LPC	607	0.478	596	0.434	644	0.351	608	0.403	587	0.313

Table 4. BAPPL binding free energies ΔG_{bind} for nsLTP-lipid complexes (kcal mol⁻¹).

Lipid	Peach	Maize	Wheat	Rice	Barley
Orientation A					
CAP	-4.3	-3.5	-3.2	-3.5	-4.7
LAU	-3.6	-4.6	-4.5	-4.9	-4.6
PLM	-4.7	-5.4	-5.8	-5.7	-5.6
EIC	-5.7	-4.0	-4.4	-4.8	-5.6
LNL	-4.9	-4.6	-4.1	-5.4	-4.2
PG2	-6.0	-4.7	-5.9	-5.0	-4.7
Orientation B					
CAP	-3.7	-3.5	-2.9	-3.9	-4.6
LAU	-3.7	-4.6	-3.8	-4.9	-4.8
PLM	-6.0	-5.5	-5.9	-6.2	-6.1
EIC	-4.5	-0.7	-4.3	-5.2	-3.5
LNL	-3.6	-0.2	-3.1	-4.2	-3.9
PG2	-5.8	-5.6	-6.4	-4.5	-6.5

atom energy empirical scoring function that explicitly comprises four terms: electrostatics and van der Waals interactions, hydrophobicity effects, and an entropy term. The electrostatic contribution is computed via Coulomb's law with a sigmoidal dielectric function. Van der Waals interactions are determined using a 6–12 potential between all the atoms of the protein and ligand. Hydrophobicity interactions are modeled in terms of desolvation effects via Gurney parameter approach.^[55] The entropy term gives the loss in conformational entropy of protein side chains upon binding^[56] and is computed using an empirical scale of side chain conformational entropy.^[57] BAPPL has been validated on a heterogeneous dataset of 161 complexes yielding a $r = 0.92$ correlation for the predicted binding free energies against experimental binding affinities.^[43] Despite its computational simplicity, the method is able to provide reasonable estimates for binding free energies of protein-ligand complexes and its implementation as web server offers a fast protocol to explore protein-ligand interactions.

ΔG_{bind} values in Table 4 indicate small binding energies between 3 and 6 kcal mol⁻¹ for most complexes, that is, dissociation constants K_D between 6.4×10^{-3} and 4.0×10^{-5} M. The only exception are the maize complexes with EIC and LNL lipids in orientation B which show binding energies smaller than 1 kcal mol⁻¹. Both palmitate and prostaglandin B₂ have slightly greater binding affinities about 6 kcal mol⁻¹ in all complexes in both orientations. It must be stressed that as far as these estimates are concerned, no significant difference is noticed between orientation A and orientation B for the lipid inserted into the tunnel-like cavity of plant nsLTps.

Test MD simulations and docking of two nsLTP-lipid complexes

Two complexes were used to perform a computational test with a twofold purpose: (i) to analyze the sensitivity of computed energies and interface properties to input structures and (ii) to explore the existence of alternate ligand binding modes. To this end, MD simulations were performed as explained in Methods starting at (1) the preceding optimized

structures and (2) new conformations prepared with docking calculations. Besides the final geometries obtained after completion of MD simulations (1) and (2), labeled hereafter "MD-sim" and "MD-dock," respectively, structures picked from frames corresponding to the middle time of MD runs (1) were also taken (labeled "MD-frame"). Given the computational cost of these calculations, we restricted this test to peach-EIC and wheat-PG2 complexes in orientations A and B as representative examples of the 70 nsLTP-lipid systems investigated. Moreover, MD simulation times were restricted to only 4 ns (which at 2 fs time-steps involve two million steps). While this is too short a time for an in-depth MD study of a protein-ligand complex, it may be regarded as enough to achieve our purposes. Before proceeding further with data computed for these new structures, we start addressing the docking results.

In both test complexes, AutoDock Vina^[36] found nine conformations of the ligand with affinity energies between -6.5 and -6.3 kcal mol⁻¹ for the EIC lipid docked to peach nsLTP and between -7.3 and -7.0 kcal mol⁻¹ for PG2 docked to wheat nsLTP. All docked conformations happen to show two groups inside the tunnel cavity with regard to the orientation of their charged head (Supporting Information Fig. S1). The proportion of orientation A to orientation B docked conformations is 3 to 6 in the peach complex and 6 to 3 in the wheat complex. While there are some expected differences on the hydrophobic tails, particularly in EIC lipid, the concordance in the positions of charged heads is satisfactory. The docked conformation showing the highest score in each orientation group was then selected to prepare the starting structure for MD runs (2).

The superposition of final structures obtained after MD simulations on the four systems (Supporting Information Fig. S2) indicate a satisfactory structural agreement with the previous geometries. Protein secondary structure is tightly conserved and, as expected, only the C-terminal tail exhibits noticeable differences. As for ligands, the EIC lipid shows clearly greater differences than PG2, particularly with regard to the position of the charged head outside the cavity (Supporting Information Fig. S2). This notwithstanding, the structural comparisons suggest that these MD simulations rule out alternate binding modes. The differences noticed in ligand conformations should be regarded as likely to occur if one considers the nature of the computational procedures followed to obtain them. Even choosing the system with the worst agreement between previous and new geometries as measured by the RMS deviations between ligands (peach nsLTP-EIC complex in orientation B; see Table 5), the output from MD simulations (Supporting Information Fig. S3) indicate no hints on alternate binding modes. It should be mentioned that an in-depth MD study of lauric acid (LAU) bound to peach nsLTP with a 20 ns simulation revealed that the time-evolving structure in solution remained in a state very similar to that of the starting model.^[14] These calculations explored also thermal effects increasing the temperature by 30 K intervals at 8 and 14 ns without finding significant changes in the structure of both protein and ligand. While this analysis was primarily intended to check if the geometry observed in the crystal structure of

Table 5. Free energies ΔG_{polsv} , ΔG_{solv} , and ΔG_{bind} (kcal mol⁻¹), areas A_{int} (Å²) and p -values P_{int} of protein-ligand interfaces, and RMS deviations $RMSD$ (Å) on nonhydrogen atoms of the lipids with respect of reference (Ref) geometries of MD structures of peach nsLTP-EIC and wheat nsLTP-PG2 complexes in orientations A and B.

	MD-sim	MD-dock	MD-frame	Ref
Peach nsLTP – EIC complex, orientation A				
ΔG_{polsv}	-4.7	-5.1	-1.5	-4.0
ΔG_{solv}	+3.2	+4.4	+3.7	+3.5
ΔG_{bind}	-4.9	-6.4	-4.7	-5.7
A_{int}	471	502	493	489
P_{int}	0.174	0.225	0.186	0.305
$RMSD$	1.54	2.07	1.76	-
Peach nsLTP – EIC complex, orientation B				
ΔG_{polsv}	+2.7	+0.1	+2.4	+2.7
ΔG_{solv}	+2.9	+2.7	+1.7	+2.1
ΔG_{bind}	-4.9	-5.8	-3.6	-4.5
A_{int}	461	473	446	435
P_{int}	0.131	0.179	0.124	0.087
$RMSD$	2.31	2.92	2.46	-
Wheat nsLTP – PG2 complex, orientation A				
ΔG_{polsv}	+3.5	+4.5	+3.3	+8.6
ΔG_{solv}	-0.5	-0.1	-0.4	-0.1
ΔG_{bind}	-4.8	-5.9	-5.8	-5.9
A_{int}	484	519	488	519
P_{int}	0.136	0.141	0.169	0.199
$RMSD$	1.71	1.83	1.92	-
Wheat nsLTP – PG2 complex, orientation B				
ΔG_{polsv}	+1.4	+2.5	+0.7	+2.7
ΔG_{solv}	-0.2	-0.8	-0.2	-0.5
ΔG_{bind}	-5.5	-5.3	-6.7	-6.4
A_{int}	443	476	467	475
P_{int}	0.132	0.112	0.143	0.146
$RMSD$	1.58	1.67	1.73	-

the peach nsLTP-LAU complex represented a stable or transient state in solution, its conclusions demonstrated a stable, unique binding mode of the ligand.^[14]

We finally show in Table 5 energies and protein-ligand interface data computed with the new MD-sim, MD-dock, and MD-frame structures of peach nsLTP-EIC and wheat nsLTP-PG2 complexes. Reference results for these complexes presented in Tables 2–4 are included here for ease of comparison (column labeled “Ref”). RMS deviations in Table 5 are computed with respect to Ref structures upon superposing the proteins without optimizing the structural fit around ligand binding sites. Data in this table reveal that the employed methods to estimate nsLTP-lipid binding features are able to calculate quantities robust enough to provide useful information regarding the nature of these complexes. This is particularly apparent for PISA data and BAPPL estimates of binding free energies, whereas PB polar solvation energies are more sensitive to input structures, although sign and magnitude of ΔG_{polsv} values are kept in all cases. As for our second reason to perform this computational test, data in Table 5 and output from MD simulations discussed in the preceding paragraph suggest no indication of alternate ligand binding modes.

Conclusions

Plant nsLTPs are able to bind a broad variety of lipids into a tunnel-like cavity which spans a similar space of the protein

core in the available experimental structures of nsLTPs. Our study revealed that the cavity accommodates lipids in two opposite orientations without significant differences in the resulting structures, solvation effects or binding energies. Despite the long held belief that the flexibility of the cavity might account for the lack of binding specificity of these proteins, optimized structures of nsLTP-lipid complexes showed very small deviations from uncomplexed structures. Except for one single lipid (prostaglandin B₂), the structural differences between complexes were not noticeably greater than those between uncomplexed nsLTPs, irrespective of the lipid orientation. Insofar as our study included lipids of rather distinct size, this result suggests that the insertion of a variety of ligands can occur without significant structural adaptations. The exception noted concerns a special case, i.e., the complete burial of prostaglandin B₂ inside the cavity.

Molecular surfaces and electrostatic potentials were properties able to distinguish between nsLTPs and orientations. Surfaces unveiled differences between orientations in size and location of cavity openings, even for buried prostaglandin B₂ for which these openings connect polar bonds of the lipid with the solvent. Electrostatic potentials split plant nsLTPs into two sets: one composed of peach, maize, and rice (net positive charges between +6 and +8), and another composed of wheat and barley (both net positive charge +2). Electrostatic potentials on protein surfaces as well as three-dimensional potential isosurfaces agree in predicting for the former set a more favorable binding of negatively charged ligands, not only with regard to their location in the tunnel but also to their possible directioning through the space surrounding the protein. Zero potential isosurfaces showed that nsLTPs in the former set can be located at any negative electrostatic environment without spatial orientation constraints which are observed in nsLTPs in the latter set.

Solvation polar effects played no favorable energetic role in complex formation. Solvation free energy gains upon interface formation showed dominance of hydrophobicity effects except for prostaglandin B₂ (which has polar groups at intermediate positions) and positive LPC phospholipid (which protrudes largely into the solvent). Binding free energies between 3 and 6 kcal mol⁻¹ were found for most complexes without significant differences between orientations. Taken together, these results showed that nsLTP-lipid affinity is small but non-negligible and that complexation occurs without help of favorable solvation effects.


Irrespective of the lipid orientation inside the cavity, the great majority of protein-ligand interfaces showed features indicative of hydrophobic specificity. Although saturated carbon chains in both saturated fatty acids and LPC phospholipid displayed a lesser specificity, plant nsLTPs were found to bind unsaturated fatty acids and prostaglandin B₂ forming unusual interfaces that indicate specific hydrophobic spots. This result is at odds with the long assumed belief that the nonspecific nature of van der Waals protein-ligand interactions might be a major factor in the lack of binding specificity of nsLTPs.

Among the proteins investigated in this work, peach nsLTP had a higher ability to bind negatively charged lipids than the remaining plant nsLTPs. This conclusion is supported by the following features: smaller structural differences upon complexation with all lipids, larger ability to adapt the cavity openings to different lipids, more extended positive electrostatic potential and tighter spatial zero potential isosurfaces around the uncomplexed protein, more favorable polar solvation energies to lipid binding, and lesser hydrophobic specificity upon formation of protein-ligand interfaces.

Recently reported evidence on the possible multifunctionality of nsLTPs in plant growth processes^[6–8] and open questions on the antigenic interactions of plant nsLTPs in alimentary allergies (particularly peach nsLTP),^[58–61] pose the need to deepen our understanding on their lipid binding properties. In addition, we aimed to gain insight into general protein-ligand interactions by providing quantitative information on the lipid binding abilities of these small proteins. While it is evident that a better estimate of binding free energies should call for more sophisticated procedures (e.g. thermodynamic integration methods), especially for addressing how the flexibility of nsLTPs affects the binding affinities, their computational cost precludes a systematic study for a large number of complexes such as those studied here. The detailed analysis of protein-lipid interactions at the atomic level and thermodynamic integration calculations via MD simulations for a restricted sample of complexes between peach nsLTP and some lipids playing a central role in plant growth (currently under research in our laboratory) will be addressed in the near future.

Keywords: protein-ligand interactions · lipid binding · protein structure · electrostatic potential · solvation effects

How to cite this article: L. F. Pacios, C. Gómez-Casado, L. Tordesillas, A. Palacín, R. Sánchez-Monge, A. Díaz-Perales, *J. Comput. Chem.* **2012**, *33*, 1831–1844. DOI: 10.1002/jcc.23012

 Additional Supporting Information may be found in the online version of this article.

[1] J. C. Kader, *Biochim. Biophys. Acta* **1975**, *380*, 31.
 [2] D. G. Rueckert, K. Schmidt, *Chem. Phys. Lipids* **1990**, *56*, 1.
 [3] J. C. Kader, *Annu. Rev. Plant Physiol. Plant Mol. Biol.* **1996**, *47*, 627.
 [4] J. Douliez, K. Michon, K. Elmorjani, D. J. Marion, *D. J. Cereal Sci.* **2000**, *32*, 1.
 [5] F. Boutrot, A. Guirao, R. Alary, P. Joudrier, M. F. Gautier, *Biochim. Biophys. Acta* **2005**, *1730*, 114.
 [6] G. Salcedo, R. Sánchez-Monge, D. Barber, A. Díaz-Perales, *Biochim. Biophys. Acta* **2007**, *1771*, 781.
 [7] A. DeBono, T. H. Yeats, J. K. C. Rose, D. Bird, R. Jetter, L. Kunst, L. Samuelsa, *Plant Cell* **2009**, *21*, 1230.
 [8] K. Chae, B. J. Gonong, S. C. Kim, C. A. Kieslich, D. Morikis, S. Balasubramanian, E. M. Lord, *J. Exp. Bot.* **2010**, *61*, 4277.
 [9] J. Douliez, S. Jegou, C. Pato, D. Molle, V. Tran, D. Marion, *Eur. J. Biochem.* **2001**, *267*, 1117.
 [10] D. H. Shin, J. Y. Lee, K. Y. Hwang, K. K. Kim, S. W. Suh, *Structure* **1995**, *3*, 189.
 [11] G. W. Han, J. Y. Lee, H. K. Song, C. Chang, K. Min, J. Moon, D. H. Shin, M. L. Kopka, M. R. Sawaya, H. S. Yuan, T. D. Kim, J. Choe, D. Lim, H. J. Moon, S. W. Suh, *J. Mol. Biol.* **2001**, *308*, 263.

[12] D. Charvolin, J. P. Douliez, D. Marion, C. Cohen-Addad, E. Pebay-Peyroula, *Eur. J. Biochem.* **1999**, *264*, 562.
 [13] H. C. Cheng, P. T. Cheng, P. Peng, P. C. Lyu, J. Sun, *Protein Sci.* **2004**, *13*, 2304.
 [14] N. Pasquato, R. Berni, C. Folli, S. Folloni, M. Cianci, S. Pantano, J. R. Hellwell, G. Zanotti, *J. Mol. Biol.* **2006**, *356*, 684.
 [15] P. Sodano, A. Caille, D. Sy, G. De Person, D. Marion, M. Ptak, *FEBS Lett.* **1997**, *416*, 130.
 [16] S. Tassin-Moindrot, A. Caille, J. P. Douliez, F. Vovelle, *Eur. J. Biochem.* **2000**, *267*, 1117.
 [17] M. H. Lerche, B. B. Kragelund, L. M. Bech, M. Poulsen, *Structure* **1997**, *5*, 291.
 [18] M. H. Lerche, M. Poulsen, *M. Protein Sci.* **1998**, *7*, 2490.
 [19] E. Krissinel, K. Henrick, *J. Mol. Biol.* **2007**, *372*, 774.
 [20] E. Krissinel, *J. Comput. Chem.* **2010**, *31*, 133.
 [21] J. Loch, A. Polit, P. Bonarek, D. Ries, K. Kurpiewska, M. Dziedzicka-Wasylewska, K. Lewinski, (in press).
 [22] N. Guex, M. C. Peitsch, *Electrophoresis* **1997**, *18*, 2714.
 [23] E. F. Pettersen, T. D. Goddard, C. C. Huang, G. S. Couch, D. M. Greenblatt, E. C. Meng, T. E. Ferrin, *J. Comput. Chem.* **2004**, *25*, 1605.
 [24] I. N. Shindyalov, P. E. Bourne, *Protein Eng.* **1998**, *11*, 739.
 [25] The PyMOL Molecular Graphics System, Version 1.4, Schrödinger, LLC.
 [26] W. Kabsch, C. Sander, *Biopolymers* **1983**, *22*, 2577.
 [27] W. Humphrey, A. Dalke, K. Schulten, *J. Mol. Graphics* **1996**, *14*, 33.
 [28] J. C. Phillips, R. Braun, W. Wang, J. Gumbart, E. Tajkhorshid, E. Villa, C. Chipot, R. D. Skeel, L. Kalé, K. Chulten, *J. Comput. Chem.* **2005**, *26*, 1781.
 [29] A. D. MacKerell, D. Bashford, M. Bellott, R. L. Dunbrack, J. Evanseck, M. J. Field, S. Fischer, J. Gao, H. Guo, S. Ha, D. Joseph, L. Kuchnir, K. Kuczera, F. T. K. Lau, C. Mattos, S. Michnick, T. Ngo, D. T. Nguyen, B. Prodhom, I. W. E. Reiher, B. Roux, M. Schlenkrich, J. Smith, R. Stote, J. Straub, M. Watanabe, J. Wiorkiewicz-Kuczera, D. Yin, M. Karplus, *J. Phys. Chem. B* **1998**, *102*, 3586.
 [30] A. Jakalian, B. L. Bush, D. B. Jack, C. I. Bayaly, *J. Comput. Chem.* **2000**, *21*, 132.
 [31] J. Wang, W. Wang, P. A. Kollman, D. A. Case, *J. Mol. Graph. Model* **2006**, *25*, 247.
 [32] V. Zoete, M. A. Cuendet, A. Grosdidier, O. Michielin, *J. Comput. Chem.* **2011**, *32*, 2359.
 [33] W. Jorgensen, J. Chandrasekhar, J. Madura, M. Klein, *J. Chem. Phys.* **1983**, *79*, 926.
 [34] T. Darden, D. York, L. Pedersen, *J. Chem. Phys.* **1993**, *98*, 10089.
 [35] M. F. Sanner, *J. Mol. Graph.* **1999**, *17*, 57.
 [36] O. Trott, A. J. Olson, *J. Comput. Chem.* **2010**, *31*, 455.
 [37] B. Knapp, N. Lederer, U. Omasits, W. Schreiner, *J. Comput. Chem.* **2010**, *31*, 2868.
 [38] N. A. Baker, D. Sept, S. Joseph, M. J. Holst, J. A. McCammon, *Proc. Natl. Acad. Sci. USA* **2001**, *98*, 10037.
 [39] J. Wang, P. Cieplak, P. A. Kollman, *Comp. Chem.* **2000**, *21*, 1049.
 [40] T. Dolinsky, J. Nielsen, J. A. McCammon, N. A. Baker, *Nucleic Acids Res.* **2004**, *32*, 665.
 [41] R. P. Bahadur, M. Zacharias, *Cell Mol. Life Sci.* **2008**, *65*, 1059.
 [42] Q. Xu, A. A. Canutescu, G. Wang, M. Shapovalov, Z. Obradovic, R. L. Dunbrack, *J. Mol. Biol.* **2008**, *381*, 487.
 [43] T. Jain, B. Jayaram, *B. FEBS Lett.* **2005**, *579*, 6659.
 [44] N. Arora, B. Jayaram, *J. Phys. Chem. B* **1998**, *102*, 6139.
 [45] J. Wang, R. M. Wolf, J. W. Caldwell, P. A. Kollman, D. A. Case, *J. Comput. Chem.* **2004**, *25*, 1157.
 [46] W. D. Cornell, P. Cieplak, C. I. Bayly, I. R. Gould, K. M. Merz, D. Ferguson, D. C. Spellmeyer, T. Fox, J. W. Caldwell, P. A. Kollman, *J. Am. Chem. Soc.* **1995**, *117*, 5179.
 [47] M. E. Davis, J. A. McCammon, *Chem. Rev.* **1990**, *90*, 509.
 [48] B. Honig, A. Nicholls, *Science* **1995**, *268*, 1144.
 [49] D. R. Livesay, D. J. Jacobs, J. Kanjanapangka, E. Chea, H. Cortez, J. Garcia, P. Kidd, M. Pulido Marquez, S. Pande, D. Yang, *J. Chem. Theor. Comput.* **2006**, *2*, 927.
 [50] J. A. Wagoner, N. A. Baker, *Proc. Natl. Acad. Sci. USA* **2006**, *103*, 8381.
 [51] D. Livesay, S. Linthicum, S. Subramaniam, *Mol. Immunol.* **1999**, *36*, 397.
 [52] D. R. Livesay, S. Subramaniam, *Protein Eng. Des. Sel.* **2004**, *17*, 463.
 [53] D. R. Livesay, P. Jambeck, A. Rojnuckarin, S. Subramaniam, *Biochemistry* **2003**, *42*, 3464.

- [54] T. Liu, Y. Lin, X. Wen, R. N. Jorissen, M. K. Gilson, *Nucleic Acids Res.* **2007**, *35*, D198.
- [55] B. Jayaram, D. L. Beveridge, *J. Phys. Chem.* **1990**, *94*, 4666.
- [56] A. J. Doig, M. J. E. Sternberg, *Protein Sci.* **1995**, *4*, 2247.
- [57] S. D. Pickett, M. J. E. Sternberg, *J. Mol. Biol.* **1993**, *231*, 825.
- [58] G. García-Casado, L. F. Pacios, A. Díaz-Perales, R. Sánchez-Monge, M. Lombardero, F. J. García-Selles, F. Polo, D. Barber, G. Salcedo, *J. Allergy Clin. Immunol.* **2003**, *112*, 599.
- [59] A. Palacín, S. Quirce, A. Armentia, M. Fernández-Nieto, L. F. Pacios, T. Asensio, J. Sastre, A. Díaz-Perales, G. Salcedo, *J. Allergy Clin. Immunol.* **2007**, *120*, 1132.
- [60] L. F. Pacios, L. Tordesillas, J. Cuesta-Herranz, E. Compes, R. Sánchez-Monge, A. Palacín, G. Salcedo, A. Díaz-Perales, *Mol. Immunol.* **2008**, *45*, 2269.
- [61] L. Tordesillas, L. F. Pacios, A. Palacín, S. Quirce, A. Armentia, D. Barber, G. Salcedo, A. Díaz-Perales, *Mol. Immunol.* **2009**, *47*, 534.

Received: 20 December 2011

Revised: 19 April 2012

Accepted: 20 April 2012

Published online on 23 May 2012