

Structure-based Classification of the Plant Non-specific Lipid Transfer Protein Superfamily Towards its Functional Characterization

Cécile FLEURY¹, Marie-Françoise GAUTIER¹, Jean-François DUFAYARD¹,
Franck MOLINA², Manuel RUIZ¹ and Frédéric DE LAMOTTE¹

cecile.fleury@cirad.fr

CONTEXT

non-specific Lipid Transfer Proteins (nsLTPs) :

- small proteins : 50 to 150 amino acid residues (including signal peptide)
- 8 cysteine motif backbone :
C-Xn-C-Xn-CC-Xn-CXC-Xn-C-Xn-C
- no glycine/proline rich N-terminal region (< 40%)
- > 70 ligands (lipids, hydrophobic compounds)
- roles in plant defense mechanism: resistance to biotic stresses and abiotic, plant germination, etc.
- α -helical folding pattern (4-5 helices)
- 4 disulfide bonds
- hydrophobic cavity (35-350 Å³)
- 32 experimental structures (10 seq.)

➔ **800 nsLTPs** from 100 plant species

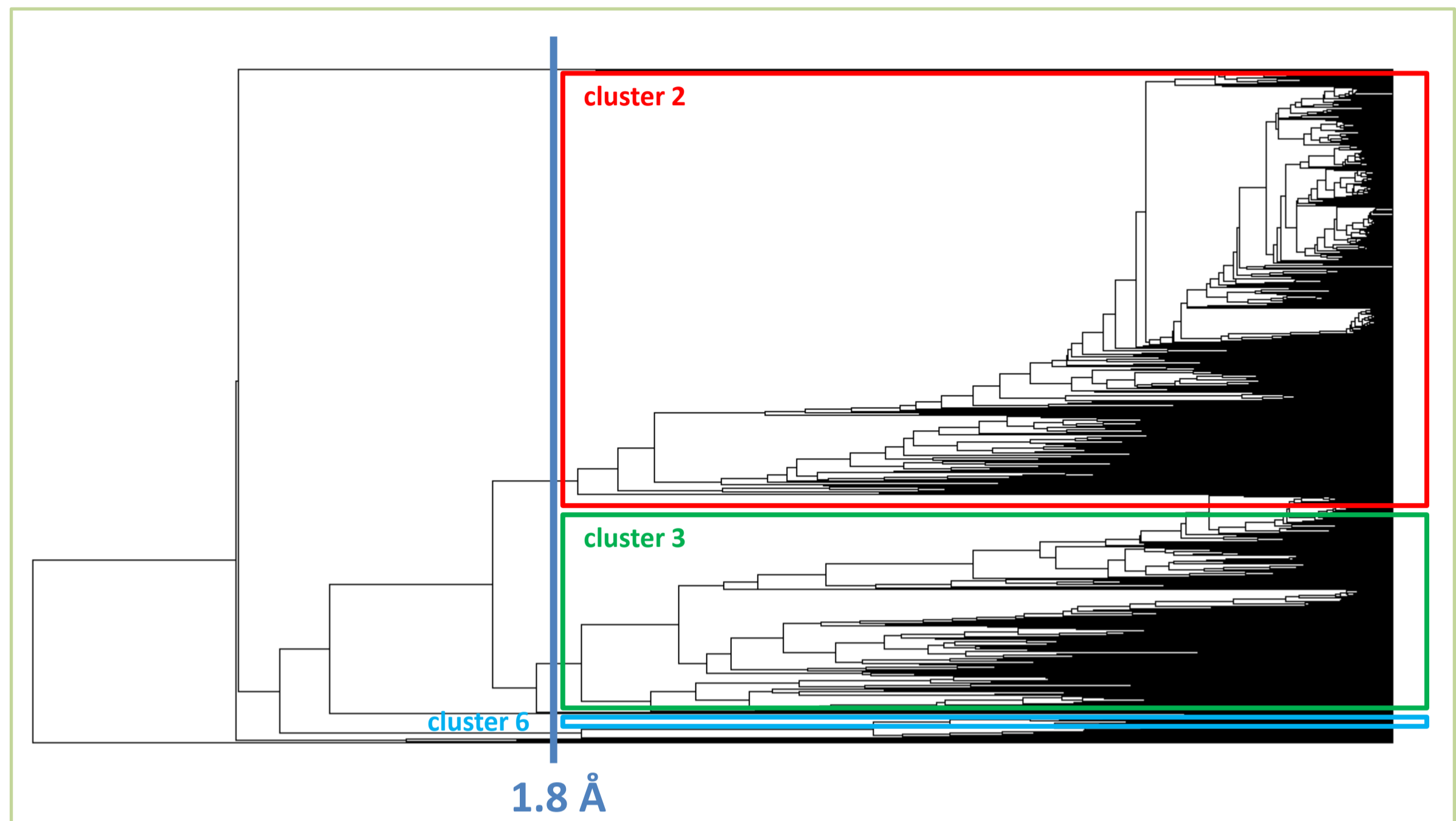


Figure 2: Dendrogram generated on the basis of structural comparison of all the nsLTPs of the new dataset. As the tree is ultrametric, one can group the structures according to a maximal RMS distance of 1.8 Å. By doing so, the superfamily is splitted into 10 groups, among which 3 main clusters arise that contain 509, 261 and 29 proteins, respectively.

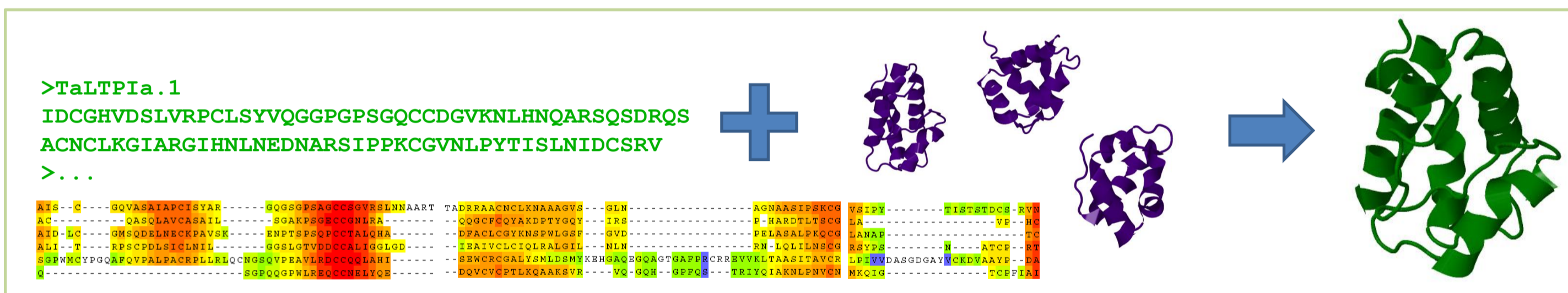


Figure 1: Homology modeling of the 800 nsLTPs with no known structure.

	1	10	20	30	40	50	60	70
Cluster 2	-----	-----	-----	-----	-----	-----	-----A	-----I
Cluster 3	-----	-----	-----	-----	-----	-----	-----	-----
Cluster 6	-----	-----	-----	-----	-----Q	-----G	-----G	-----G
	80	90	100	110	120	130	140	150
Cluster 2	---S---	---C---	-----G	-----	-----Q	---VA---	---SAIAPCIS	---YAR---G---
Cluster 3	-----	-----	-----	-----	-----A	---C---QA	---SQLAVCAS	---AIL---S---
Cluster 6	---G---	-----	-----E	-----C	---V---P	---QL---	---NRLACRA	---YAV---P---
	160	170	180	190	200	210	220	230
Cluster 2	-----Q	-----G	-----S	GP-SAGCCSG	VRSLNNAAR	T--T-AD--R	---RA---	ACNC LKN---AA-
Cluster 3	-----	-----	-----A	KP-SGECGN	LRAQQ	-----	---GC---	FCQY AKD---PTY-
Cluster 6	-----G	-----A	-----G	DP-SAECCSA	LSSIS	-----	---Q---	GC--ACSA IS---
	240	250	260	270	280	290	300	310
Cluster 2	--A-G---	--V---S---	-----	-----G	-----L	-----N	AGNAASIPSK	CGVSI---
Cluster 3	--G---	-----Q	-----	-----Y	-----I	-----R	SPHARDTLTS	CGLAV---
Cluster 6	-----	-----	-----	-----	-----	-----	---I---MNSLPSR	CHLSQ---
	320	330	340	350	360	370	380	390
Cluster 2	P-----	---Y---	-----T	-----I	S---T---	S---T---D---	---C---	S---RV-N---
Cluster 3	P-----	---H---	-----C	-----	-----	-----	-----	-----
Cluster 6	I-----	---N---	-----C	-----S	-----A	-----	-----	-----
	400	410	420	430				
Cluster 2	-----	-----	-----	-----				
Cluster 3	-----	-----	-----	-----				
Cluster 6	-----	-----	-----	-----				

Figure 3: Structure-based sequence alignment of the reference proteins of the 3 main structural clusters. The original alignment has been performed on the whole set of protein structures, but only the 3 representative proteins are shown here. The class-specific evolutionarily important residues are highlighted in one colour for each cluster, while the important residues roughly conserved among the whole superfamily are underlined.

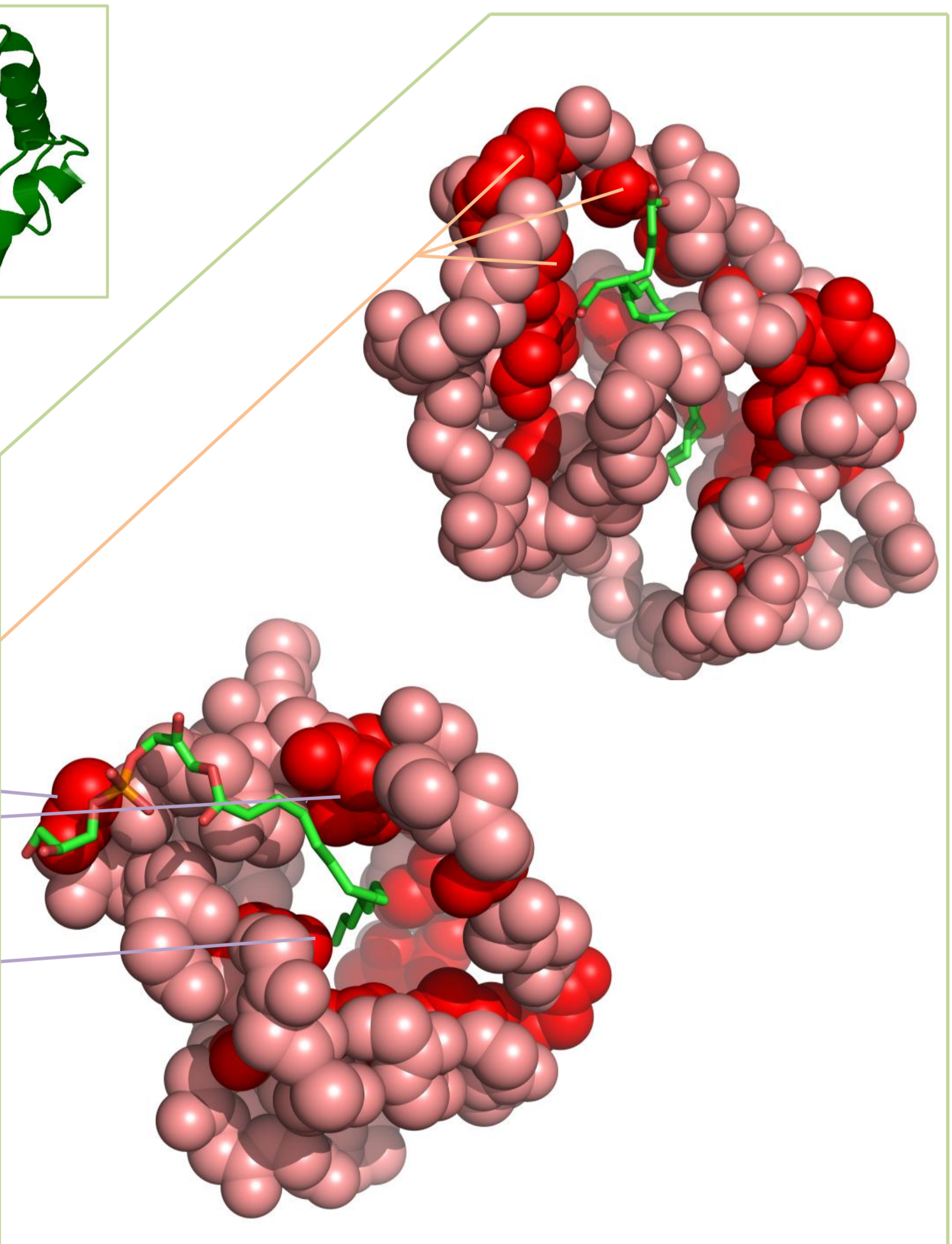


Figure 4: Experimentally determined three-dimensional structures of cluster 2 (up : maize type I nsLTP) and cluster 3 (down : wheat type II nsLTP) representative proteins. Class-specific evolutionarily important residues are highlighted in red on each structure and a natural ligand is bound inside the cavity.

¹UMR AGAP CIRAD/INRA, avenue Agropolis, 34398, Montpellier, Cedex 5, France

²SysDiag, UMR3145 CNRS/Bio-Rad, 1682 rue de La Valsière, CS 61003, Montpellier, Cedex 4, France