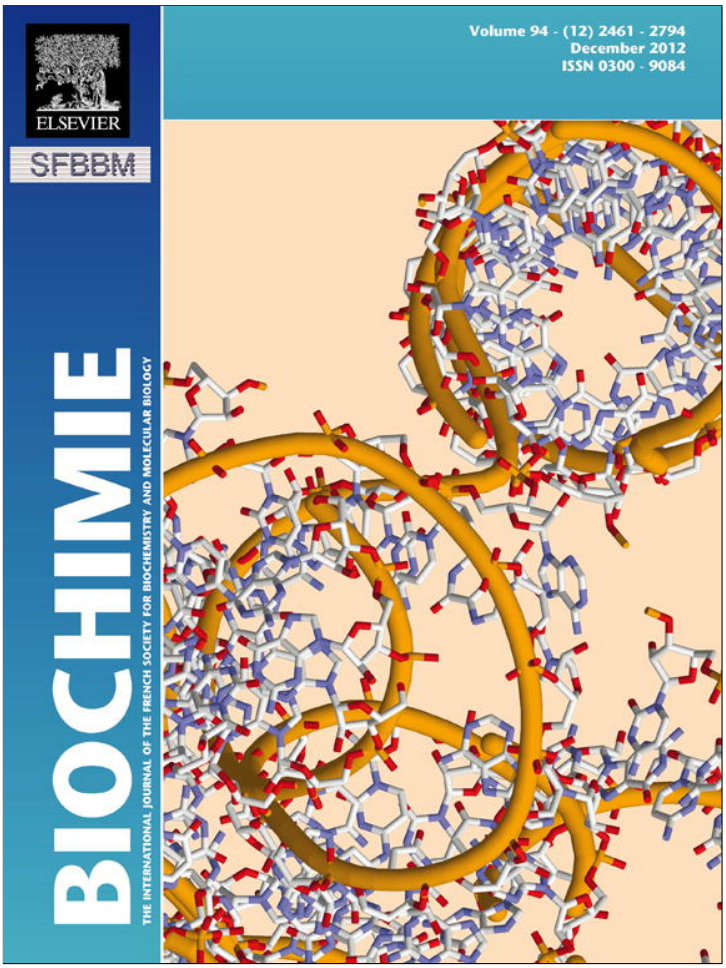

Biochimie 94 (2012) 2608–2619

Contents lists available at SciVerse ScienceDirect

Biochimie

journal homepage: www.elsevier.com/locate/biochi

Research paper

In silico and *in vitro* characterization of phospholipase A_2 isoforms from soybean (*Glycine max*)

María Elisa Mariani [a,d], Marcos Ariel Villarreal [b], Foo Cheung [c], Ezequiel Pedro Marcos Leiva [b], Ricardo Román Madoery [d], Gerardo Daniel Fidelio [a,*]

[a] *Centro de Investigaciones en Química Biológica de Córdoba, (CIQUIBIC, UNC–CONICET), Departamento de Química Biológica, Fac. de Cs. Químicas, Universidad Nacional de Córdoba, Haya de la Torre y Medina Allende, Ciudad Universitaria, X5000HUA Córdoba, Argentina*
[b] *Departamento de Matemática y Física, Facultad de Ciencias Químicas, Universidad Nacional de Córdoba, Haya de la Torre y Medina Allende, Ciudad Universitaria, X5000HUA Córdoba, Argentina*
[c] *Center for Human Immunology, Autoimmunity and Inflammation, National Institute of Health, 9000 Rockville Pike, Bethesda, MD 20892, USA*
[d] *Departamento de Fundamentación Biológica, Facultad de Ciencias Agropecuarias, Universidad Nacional de Córdoba, Córdoba, Argentina*

ARTICLE INFO

Article history:
Received 27 January 2012
Accepted 26 July 2012
Available online 7 August 2012

Keywords:
Glycine max phospholipase A_2
PLA_2 phylogenetic tree
Molecular dynamics simulations
i-face
Soybean PLA_2
Plant secreted phospholipase

ABSTRACT

At the present, no secreted phospholipase A_2 ($sPLA_2$) from soybean (*Glycine max*) was investigated in detail. In this work we identified five sequences of putative secreted $sPLA_2$ from soybean after a BLAST search in *G. max* database. Sequence analysis showed a conserved PA2c domain bearing the Ca^{2+} binding loop and the active site motif. All the five mature proteins contain 12 cysteine residues, which are commonly conserved in plant $sPLA_2$s. We propose a phylogenetic tree based on sequence alignment of reported plant $sPLA_2$s including the novel enzymes from *G. max*. According to PLA_2 superfamily, two of *G. max* $sPLA_2$s are grouped as XIA and the rest of sequences as XIB, on the basis of differences found in their molecular weights and deviating sequences especially in the N- and C-terminal regions of the isoenzymes. Furthermore, we report the cloning, expression and purification of one of the putative isoenzyme denoted as $GmsPLA_2$-XIA-1. We demonstrate that this mature $sPLA_2$ of 114 residues had PLA_2 activity on Triton:phospholipid mixed micelles and determine the kinetic parameters for this system. We generate a model based on the known crystal structure of $sPLA_2$ from rice (isoform II), giving first insights into the three-dimensional structure of folded $GmsPLA_2$-XIA-1. Besides describing the spatial arrangement of highly conserved pair HIS-49/ASP-50 and the Ca^{+2} loop domains, we propose the putative amino acids involved in the interfacial recognition surface. Additionally, molecular dynamics simulations indicate that calcium ion, besides its key function in the catalytic cycle, plays an important role in the overall stability of $GmsPLA_2$-XIA-1 structure.

1. Introduction

Soybean (*Glycine max*) is one of the most widely oil crop grain used in the world. In addition to its valuable contributions to health due to its high nutritional level, soybean is an important global source of high quality nutritional lipids, proteins and other valuable bioactive components such as: phospholipids (known as lecithin), hormones and antioxidants [1,2].

The industrial use of $sPLA_2$s from pancreatic juices and microbes, especially in food production, has a long tradition [3,4]. The aim is the utilization of $sPLA_2$ from plants for enzymatic processing to obtain lysoderivatives. This alternative has recently recognized to be of great interest since it would satisfy food regulation requirements such as in Kosher and Halal. However, no $sPLA_2$s from plants have now been available for industrial application [5].

The phospholipase A_2 (PLA_2, EC 3.1.1.4) superfamily is a broad and growing group of enzymes that stereo specifically catalyzes the cleavage at the *sn-2* acyl ester bond from diacyl-phospholipids liberating lysophospholipid and free fatty acid. In plants, secreted PLA_2 ($sPLA_2$) represents one type of phospholipase A_2 whose lipid products mediate a variety of cellular processes, including growth, development, defense and stress responses [5–11]. Although numerous $sPLA_2$s genes have been identified in plants, little is known about these enzymes in opposition to their insect, animal or human counterparts. This type of $sPLA_2$ is best known from mammals where several $sPLA_2$s have been identified within the last 20 years [12]. A lot of sPLA2s were obtained from microorganisms

* Corresponding author. Tel.: +54 351 4334168x3465.
E-mail address: gfidelio@dqb.fcq.unc.edu.ar (G.D. Fidelio).

0300-9084/$ – see front matter © 2012 Elsevier Masson SAS. All rights reserved.
http://dx.doi.org/10.1016/j.biochi.2012.07.021

as bacteria and yeasts, snake venom, bee venom and pancreatic juices, where it occurs abundantly and has a digestive role [12–15].

In the past years, several putative $sPLA_2$s identified in plants have been cloned and investigated but their functional roles are poorly understood. One enzyme was purified from elm (*Ulmus glabra*) [8] and PLA_2 activity was investigated in tobacco (*Nicotiana tabacum*) [16]. In rice (*Oryza sativa*), two cDNAs encoding $sPLA_2$ were isolated [6], four isoforms were identified in Arabidopsis (*Arabidopsis thaliana*) [9,17,18] and one $sPLA_2$ cDNA was isolated and cloned from carnation (*Dianthus caryophyllus*) [7].

Secreted PLA_2s are low molecular weight enzymes (12–18 kDa) [13]. After removal of the N-terminal signal peptide in the endoplasmic reticulum, they are secreted into the extracellular space in a mature form. In animals, $sPLA_2$s contain between 10 and 16 cysteines that have the potential to form 5–8 intramolecular disulfide bridges [13]. It is known that cysteine residues are essential for secreted $sPLA_2$ and it has been shown to play an important role in the structural stability of the mature enzyme [19]. In contrast, all $sPLA_2$s reported from plants have 12 cysteines that can form 6 disulfide bridges. Other important common features for these enzymes are the presence of: *i*) one histidine residue at the catalytic domain for nucleophilic attack at the *sn-2* acyl ester bond of the glycerol backbone, *ii*) micromolar to millimolar requirement of calcium for full activity, and *iii*) extremely heat-stable enzymes. Plant $sPLA_2$s also contain a PA2c domain with the highly conserved Ca^{2+} binding loop (YGKYCGxxxxGC) and the active site motif (DACCxxHDxC), where the HIS/ASP pair was found to be highly well conserved. At least, two characteristics are of great interest in the structure of all $sPLA_2$: the catalytic site and the interfacial recognition surface (IRS). All the enzymes of secretory type from animals ($sPLA_2$s) have the same architecture at the catalytic site level (HIS-ASP) but differ on the amino acid residues that conform the IRS [20]. It is known that plant $sPLA_2$s share little amino acid sequence similarity (only 15% of identity) compared with animal $sPLA_2$s, except in the catalytically critical Ca^{2+}-binding loop and in the active site motif (about 55% of identity) [21]. The PLA_2 superfamily was classified into 15 groups and subgroups, based on both criteria: the primary structure and functional properties [13,19]. Plant $sPLA_2$s belong to the group XI, in turn subdivided into XIA and XIB.

The structure of secreted $sPLA_2$ isoform α from *A. thaliana* was elucidated by homology modeling and molecular dynamics [22]. To date, only three structures were reported in the protein Data Bank (PDB) and all correspond to *O. sativa* (rice) isoform II, which belongs to the group XIB, and its tertiary structure was recently determined by X-ray crystallography to 2.0 Å resolution [23]. The structure of rice $sPLA_2$ shows that the half of the N-terminal chain contains mainly structured loops, including the conserved calcium binding loop together with two short anti-parallel beta-strands. The C-terminal half is folded into three anti-parallel α-helix, in which two of them are highly conserved among other $sPLA_2$s, containing the crucial catalytic histidine residue and the calcium binding/coordinating aspartate residues.

Phospholipids are constituents of biological membranes, so a very important prerequisite step to perform the lipolytic action of $sPLA_2$ is the interaction with the amphipathic nature of these interfaces; and this, in turn, determine the catalytic properties of the organized substrate [24]. The interfacial binding step is crucial for enzymatic action of $sPLA_2$, and it is mediated by a region of the protein often referred to as *i-face*. The proper intimate contact of the *i-face* of $sPLA_2$s with the interface is essential to provide the substrate access to the active site. Interfacial activation arises from an adequate coupling between the active site and the *i-face* modulating the catalytic activity [24–27]. The binding and kinetic characteristics of interfacial catalysis by $sPLA_2$ depend upon the organization and dynamics of the interface. The overall rate of catalytic turnover is not only determined by the kinetics at the interface, but also by the binding/desorption equilibrium kinetics of the enzyme with the interface. Hence, the hydrolysis of the organized substrate can occur in two distinct modes: *i*) in the *scooting* mode of catalysis, the bound enzyme remains at the interface between the catalytic turnover cycles and, *ii*) in the pure *hopping* mode, the binding and the desorption of the bound enzyme occur during each catalytic turnover cycle [28].

We have previously identified phospholipase activity in seed extracts from soybean (*G. max*), and its preliminary biochemical and enzymatic characterization were performed [29]. In the present work we intend to carry out a detailed structural and functional study of $sPLA_2$ from *G. max* focused on to further characterize these novel enzymes. We report the existence of, at least, five isoforms of putative $sPLA_2$ sequences from soybean (*G. max*), and we also tentative grouped them in the PLA_2 superfamily by sequence and phylogenetic analysis. Furthermore, one of them named as $GmsPLA_2$-XIA-1 was cloned, expressed and purified from inclusion bodies and the PLA_2 activity was analyzed using phospholipid:triton mixed micelle as organized substrate. Analysis of $GmsPLA_2$-XIA-1 sequence allowed us to characterize the essential functional domains for catalytic activity. We also generated the tree-dimensional structure of the mature $GmsPLA_2$-XIA-1 protein elucidating the topology of catalytic and calcium binding loop sites and the putative hydrophobic domain responsible for the interfacial recognition at the membrane surface level. Additionally, we describe the hydrophobic surface of this interacting membrane enzyme by using the reported approach to search druggable hot spots in proteins.

2. Materials and methods

2.1. Materials

DLPC (dilauroylphosphatidylcholine) was purchased from Avanti Polar Lipids, Inc. (Birmingham, AL, USA). Pancreatic PLA_2 and Triton X-100 were purchased from Sigma Chem. Co. (St. Louis, MO, USA), Isopropyl-L-thio-β-D-galactopyranoside (IPTG) and ampicillin from Promega (Madison, WI, USA). Ni-NTA chelating sepharose was purchased from GE Healthcare, USA. All other reagents were of the highest analytical grade.

2.2. Amplification and cloning of $GmsPLA_2$-XIA-1

The strategy used to produce a recombinant $GmsPLA_2$-XIA-1, consists in which the protein was fusioned to ubiquitin (Ub) moiety to be subcloned in the expression pHUE vector constructed by Ann-Maree Catanzariti et al. [30]. The fusion of the protein of interest to Ub can be generated using the SacII site, which has been engineered into the 3′ end of the Ub. The ligated DNA fragment must encode Gly75-Gly76 residues of Ub, which are essential for cleavage. PCR amplification of the mature sequence of $GmsPLA_2$-XIA-1 gene was performed using specific primers designed as it was recommended in Ref. [30], with the 5′extention dCTC-CGC-GGT-GGT, encoding Leu73, Arg74, Gly75, Gly76 and containing the SacII site, and using the full-length cDNA clone "JCVI-FLGm-10D19" (Plant Genomics, J. Craig Venter Institute Genbank BT092274) as template. For this purpose, PCR was carried out using 50 ng of template and GoTaq DNA polymerase (Promega, WI, USA), for 33 cycles of 10 min at 95 °C for initial denaturation, 94 °C 40 s for denaturation, 40 s with gradient annealing temperatures, 10 cycles at 45 °C, 8 cycles at 48 °C and 15 cycles at 53 °C, based on Tm value of the primers and 60 s at 72 °C for polymerization, using a forward primer 5′-TCCCCGCGGTGGATCCGATCAGGCAAACTGCAGCACC-3′ and a reverse primer 5′-TTAAAGCTTAATGAGGAACGGAGTC-3′. The amplified

product was analyzed and separated on 1% agarose gel, excised and eluted using Promega Wizard SV Gel and PCR Clean-UpSystem (Promega, WI, USA), cloned into pGEM-Teasy vector (Promega, WI, USA), transformed into XL1Blue *Escherichia coli*. Positive clone was selected according to the manufacturer's protocol. Plasmid DNA was purified from the positive clone with The Wizard® SV 96 Plasmid DNA Purification System (Promega, WI, USA). The constructed plasmid was digested using SacII (Promega, WI, USA) and HindIII (Promega, WI, USA) restriction enzymes, and the fragment coding for the Ub-GmsPLA$_2$-XIA-1 was subsequently purified from 1% agarose gel as mentioned above and subcloned into the expression vector pHUE. This system allows the protein to be expressed as histidine-tagged-ubiquitin fusions, and it has been previously found to significantly increase the yield of unstable expressed proteins [30]. The correctness of the nucleotide sequence of the construct was verified by DNA sequencing by 3730XL automated DNA Sequencer (Applied Biosystems, University of Chicago Cancer Research Center-DNA sequencing facility). The full sequence of the Ub-GmsPLA$_2$-XIA-1 construct is shown in Fig. S1B of supplementary data.

2.3. Expression and purification of recombinant GmsPLA$_2$-XIA-1

The assembled gene (Ub-GmsPLA$_2$-XIA-1) subclone into His-tagged Ub the expression vector pHUE was transformed into *E. coli* BL21 strain (DE3). Bacterial cultures were grown in fresh LB medium containing ampicillin (100 μg/ml) at 37 °C. When the optical density at 600 nm reached a value between 0.6 and 0.8, the protein expression was induced by the addition of IPTG (Isopropyl-L-thio-β-D-galactopyranoside) to a final concentration of 0.4 mM and incubated for 4 h at 37 °C. The cells were harvested by centrifugation at 5000 rpm at 4 °C. The fusioned protein was present in inclusion bodies and the purification was performed as it was reported [31] with some modifications. Poly-His-tagged recombinant protein (Poly-His-Ub-GmsPLA$_2$-XIA-1) was purified by nickel-affinity chromatography (chelating sepharose fast flow, GE Healthcare, USA) in batch under denaturing conditions. The construct was eluted from the nickel resin in 1 ml fractions with buffer Tris–HCl 50 mM pH 8 containing 6 M Gdn·HCl and 50–250 mM of imidazole. Chosen fractions were pooled and diluted 1:10 to obtain a 0.6 M of Gdn·HCl and digested with USP2cc deubiquitylating enzyme at 37 °C overnight at a 1:100 peptidase to substrate mole ratio. The cleavage occurs after the final glycine residue at the carboxyl terminal of Ub irrespective of the amino acid immediately followed in the sequence [30]. The tagged protease allows the cleavage of the Ub from the desired protein as well as its selective removal from the cleavage reaction, along with the cleavage Ub, any uncleaved fusion protein, and any co-purified contaminants, leaving the desire protein as the only soluble product. A final purification step was also performed to isolate the cleaved product from the cleavage reaction mixture by using nickel resin in the same conditions as mentioned above. In our hands, folding of the mature GmsPLA$_2$-XIA-1 with enzymatic activity was achieved by pooled the chosen fractions and dialyzed for 4 days at 4 °C against a buffer containing 50 mM Tris–HCl pH 8, 10 mM cysteine and 1 mM $CaCl_2$. The final purity of mature GmsPLA$_2$-XIA-1 was about 95% as determined by SDS-PAGE electrophoresis.

The dehubiquitilating enzyme, USP2cc was engineered by Ann-Maree Catanzariti from a mouse Usp2-45 cDNA, to provide a minimal catalytically active deubiquitylating domain and expressed and purified this as poly-histidine-tagged protein. The deubiquitylating peptidase 6H-USP2-cc was expressed in overnight cultures of *E. coli* strain BL21 (DE3) transformed with the pHUsp2-cc construct kindly donated by Ann-Maree Catanzariti and further purified by using nickel affinity chromatography (see Ref. [30]).

2.4. SDS-PAGE and protein determination

Electrophoresis on 15% polyacrylamide slab gels was performed following the Laemmli's method [32], to estimate the relative molecular mass of the protein, under reducing conditions. Runs were of 50 min with 160 mA. Proteins were visualized with Coomassie brilliant blue R-250 staining. SDS-PAGE Gels were scanned with a flat-bed scanner, and digital images were imported and quantified using Photoshop software (Adobe Photoshop CS). Then, the intensities of bands were compared according to their grayscale using bovine serum albumin as standard according to (http://lukemiller.org/index.php/2010/11/analyzing-gels-and-western-blots-with-image-j/). Protein concentration was further corroborated by absorbance at 280 nm in GenQuant Spectrophotometer (Amersham Biosciences, USA).

2.5. PLA$_2$ activity measurements

2.5.1. Indirect determination of PLA$_2$ activity by direct screening of hemolytic activity

As general screening, in the initial stages of the renaturation reaction, activity was indirectly determined by the assay of agar:-egg PC:blood in Petri dish as described [33]. 1.2 ml of packed human erythrocytes (from National University of Córdoba Blood Bank, Córdoba, Argentina) were washed four times with saline solution, 1.2 ml of 1:3 egg PC solution in saline solution, and 1 ml of 0.01 M $CaCl_2$ solution were added to 100 ml of 1% (w/v) of agar (at 50 °C), dissolved in PBS buffer at pH 8. The mixture was applied to plastic Petri dish plates and allowed to gel. Then, 3 mm-diameter wells were filled with 15 μl of samples. After 20 h of incubation at 37 °C, the diameters of hemolytic halos were visually observed. The halo formed is indicative of hydrolysis of lipids from yolk (and PLA$_2$ activity). Renaturation buffer was used as a negative control and pancreatic PLA$_2$ was used as positive control.

2.5.2. Colorimetric assay for PLA$_2$ activity in mixed micelles

PLA$_2$ activity was determined using DLPC/Triton X-100 mixed micelles (1:4, mole ratio) following the liberated fatty acids during the reaction by using the adapted NEFA-HR (2) kit from WACO (WACO Pure Chemical Industries, Japan). The activity was quantified by determining the amount of fatty acids released from DLPC according to Ref. [34]. The assay mix was composed of 5 mM DLPC, 20 mM Triton X-100, 10 mM $CaCl_2$, and 10 mM Tris/HCl pH 7 in a total volume of 30 μl. All assays were performed at 37 °C. The reaction was initiated by the addition of purified enzyme from a stock solution (the final protein concentration was 0.92 μg/μl). Substrate preparation was done by dissolving appropriate amounts of DLPC in 1 ml of chloroform/methanol (2:1, v/v) and dried with N_2 and subsequent overnight vacuum. The lipid film was dissolved in 10 mM Tris/HCl pH 7 containing 20 mM Triton X-100 and 10 mM $CaCl_2$ by vortexing for 15 min at 40 °C. In order to define the range of initial velocity (v_o), the enzyme was incubated for 1–40 min at 37 °C with DLPC (5 mM final concentration). Aliquots of 30 μl were removed after the desired incubation time, the reaction was stopped by addition of 6 μl of 0.5 M EDTA. Initial rates were determined from the released fatty acids as described above, according to the manufacture instructions, with an altered sample–reagent-ratio (36 μl of sample, 125 μl of reagent A plus 62 μl of reagent B). The absorbance of the complex was measured at 550 nm using oleic acid as standard. Enzyme activities were linearly dependent in the range of the enzyme concentration currently used in the assays. All experiments were run at least in triplicate. For the determination of the kinetic constants, the initial reaction rates were determined as described above at substrate concentrations range in the range from 1 to 6 mM of DLPC in mixed micelles. The

activity was expressed as μmol of fatty acid liberated per minute per mg of protein. The apparent kinetic parameters K_M and V_{max} were obtained by plotting initial rates versus substrate concentration and subsequent non-linear regression using the Michaelis–Menten formalism (by using Sigma Plot 10.0 as fitting program). All experiments were done in triplicate.

2.6. Fluorescence spectroscopy studies

Steady state intrinsic fluorescence measurements of purified recombinant *GmsPLA$_2$-XIA-1* were done on a Varian Eclipse Cary Spectrofluorimeter. The excitation bandlength was set at 290 nm. The excitation and emission monochromators were set at 10 nm slit width. Emission was collected in the 300–400 nm range. For each condition, ten emission spectra were averaged and smoothed according to inner routine of the software provided by the manufacturer. The folded condition was achieve with active enzyme as indicated in Section 2.3, dissolved in final buffer containing 50 mM Tris–HCl pH 8, 10 mM cysteine and 1 mM $CaCl_2$. Calcium-depleted condition was achieved by adding concentrated EDTA solution to a final concentration of 10 mM. Denaturing condition was achieved with 10 mM EDTA and 4 M Gdn·HCl pH 8.5 final concentration in the cuvette. Before each measurement the samples were equilibrated by 10 min with a gently stirring. The volume in the 10 mm pass cuvette was 50 μl.

2.7. In silico studies

2.7.1. Sequence analysis

A local database was created with FORMATdb [35] to perform a TBLASTN search in the *G. max* mRNA database (http://www.plantgdb.org/GmGDB/) for fishing out the probable genes and the putative CDS of sPLA$_2$s. We have used sPLA$_2$s sequences reported for *A. thaliana* as templates that were downloaded from UNIPROTKB database. The putative nucleotide sequences of sPLA$_2$s from *G. max* (soybean) were downloaded from Glycine max available database (http://www.plantgdb.org/GmGDB/). Protein sequences were subjected to protein functional analysis using PFAM [36] and PROSITE [37] to find out conserved motifs. For posttranslational modifications the data were analyzed by using signalIP for topology prediction [38], PROTPARAM [37] to determine the isoelectric point. The entire DNA sequences and information of the genes were downloaded from http://www.phytozome.net. The sequences of the mature proteins were aligned using Clustal X [39] and phylogenetic tree was constructed by means of the parsimony method using the PHYLIP package software V 3.63a [40]. Moreover, these sequences were also further analyzed using the facility Sequence LOGO online server [41]. Similarities and identities were analyzed using the Matgat 2.0 software [42].

2.7.2. Model building

An initial model for the protein was constructed using the I-TASSER server, which combines threading and *ab-initio* modeling for the prediction [15,43]. This server was ranked as the number one server, in the recent CASP structure prediction competitions [44]. The principal template used in the modeling was the structure reported for *O. sativa* sPLA$_2$ isoform II (PDB ID: 2WG7). This template has an identity of 40% in sequence and a coverage of 0.97 (defined as the number of aligned residues divided by the length of query protein). The model created had a *C*-score of 1.29. The *C*-score assesses the quality of the model, and typically is in the range [5,2], wherein a higher score reflects a model of better quality. In general, models with a *C*-score >-1.5 have a correct fold [15]. The *C*-score obtained indicates that the model is of high quality with an estimated deviation (rmsd) of 0.18 ± 0.15 nm from the experimental structure. The superposition of the generated model with the crystal structure of *O. sativa* sPLA$_2$ gives a rmsd of 0.07 nm. As can be observed in Fig. S2 of supplementary data, the superposition of the tertiary structures of sPLA$_2$s is very well conserved in the regions that are essential for the catalytic activity.

2.7.3. Molecular dynamics

The GROMOS56a3 [45] and the AMBER99SB [46] force fields were used for the protein. The water model used was SPCE [47] for the simulations with the GROMOS force field and TIP3 [48] for the simulations with AMBER. The parameters for calcium ion in the GROMOS force field were the proposed by Project et al. [49] while the default parameters for Ca^{2+} were used in the AMBER force field. The amino acid side chains were charged according to the pKa of each free amino acid in aqueous solution. Considering a pH of 7, the total charge on the protein was 0. For all cases we considered one Ca^{+2} ion and two Cl^- to obtain an electrically neutral system. The electrostatic interactions were handled with the SPME version of the Ewald sums [50,51]. The settings for the SPME method were as follows: a real space cut off of 0.9 nm, a grid spacing of 0.12 nm, and a cubic interpolation. In all the simulations, a dielectric permittivity of 1 was assumed and the van der Waals interactions were cut off at 1.2 nm. The simulations were carried out in the NPT ensemble using the velocity rescale algorithm for the thermostat [52] and the Berendsen barostat [53]. The protein and the solvent were coupled separately to a temperature bath with a reference temperature of 310 K and a relaxation constant of 0.3 ps. The pressure was maintained constant by coupling to a reference pressure of 1 bar with a relaxation constant of 2.0 ps. The bonds in the protein were constrained using the LINCS algorithm [54]; for the bonds and bond-angle of the water molecules we used the SETTLE algorithm [55]. The time step for the integration of the equation of motion was 2 fs with the AMBER, and 5 fs with the GROMOS force field due to the use of virtual sites in the polar hydrogen atoms of the protein [56]. In total we performed seven different simulations with an aggregated time of over 700 ns. The non-bonded list was updated every 20 fs. To release steric clashes we performed 1000 steepest descent cycles before the simulation. Prior to every production run, a series of 4 equilibration steps of 500 ps each were performed applying position restraints to all the atoms in the protein. The restraint forces were 1000, 600, 200, and 75 kJ/nm^2 in each equilibration step, respectively. Every run, whether of equilibration or production, was started with a different set of initial velocities in order to produce different trajectories. The secondary structure content of the protein was analyzed using the DSSP program reported by Kabsch and Sander [57]. A contact between two atoms was defined if they were separated by 0.4 nm or less. With this definition, no water molecules could be placed between the interacting pair of atoms. The simulations and part of the analysis of the trajectories were performed using the GROMACS 4.0.7 software package [58,59].

3. Results and discussion

3.1. Identification and sequence analysis of sPLA$_2$s from G. max

From a TBLASTN search in the *G. max* mRNA database we obtained the sequences of putative secreted PLA$_2$ (sPLA$_2$s), by using the sequences previously published for *A. thaliana* AtsPLA$_2$-α and AtsPLA$_2$-β (Uniprot ID: Q8S8N6 and Q8GZB4, respectively) as templates. The search enabled us to identify five putative sPLA$_2$s isoenzymes in soybean, denoted as GmsPLA$_2$-XIA-1, GmsPLA$_2$-XIA-2, GmsPLA$_2$-XIB-1, GmsPLA$_2$-XIB-2 and GmsPLA$_2$-XIB-3 (GenBank ID: BT092274, BT094641, BT095220, BT091171 and BT099163, respectively).

The full-length cDNA sequences indicated that each sPLA$_2$ putative gene in soybean was found to be of 789 (GmsPLA$_2$-XIA-1),

875 (GmsPLA$_2$-XIA-2), 826 (GmsPLA$_2$-XIB-1), 762 (GmsPLA$_2$-XIB-2) and 821 (GmsPLA$_2$-XIB-3) nucleotides, with open reading frames (ORF) consisting of 417, 417, 474, 471, and 477 nucleotides that codify proteins of 138, 138, 157, 156, and 158 amino acids for GmsPLA$_2$-XIA-1, -2 and GmsPLA$_2$-XIB-1, -2, and -3, respectively (including the signal sequence peptide). It was observed that genes encoding for GmsPLA$_2$-XIA-1 and GmsPLA$_2$-XIB-1 are located in chromosome I, GmsPLA$_2$-XIA-2 y GmsPLA$_2$-XIB-2 are located in chromosome 7 and the gene of GmsPLA$_2$-XIB-3 in chromosome 8. While GmsPLA$_2$-XIB-1, GmsPLA$_2$-XIB-2 and GmsPLA$_2$-XIB-3 have 3 introns and 4 exons, the genes of GmsPLA$_2$-XIA-1 and GmsPLA$_2$-XIA-2 have 2 introns and 3 exons, respectively. These facts are indicative that during the course of evolution events of divergence and duplication might have occurred.

The comparison between the amino acid sequences of these novel proteins and other known sPLA$_2$s from plants indicates the presence of conserved clusters of amino acids among them (Fig. 1A). All sPLA$_2$s sequences found in soybean hold a PA2c domain that contains the highly conserved Ca^{2+}-binding loop and the active site motif that hold the highly conserved HIS/ASP pair [60] that corresponds to the position 49 and 50 in GmsPLA$_2$-XIA-1, respectively.

However, as it was observed previously, there is a difference that remains unclear in the HIS/ASP of the catalytic dyad in sPLA$_2$s from plants compared with those found in animals [22]. While water molecules were proposed to assist in the Ca^{+2} coordination at the active site composed of HIS48 and ASP49 in bovine pancreatic sPLA$_2$ [61] the role of ASP99 in this sPLA$_2$ [62,63], and ASP64 in bee venom sPLA$_2$ were also postulated to participate in the hydroxyl–imidazole–carboxylate motif [64] but this seems not to be conserved in plant sPLA$_2$ enzymes. This important catalytic residue is replaced by a histidine or an asparagine residue in enzymes from group XIA and a serine or asparagine in enzymes belonging to group XIB [22] as it is shown in the alignment (Fig. 1A). Mansfeld et al. demonstrated that SER, ASN or HIS in plant sPLA$_2$s may adopt the catalytic role assigned to ASP in animals sPLA$_2$s [22]. Even though the comparison showed a low homology among them within the overall amino acid sequences, both the catalytic site and the Ca^{2+} binding loop are highly conserved (Fig. 1A). Other relevant conserved residues within the Ca^{2+} binding loop are two tyrosines and two glycines which participate in the hydrogen bonding network reported in either animal or plant sPLA$_2$s [21]. The mature proteins contain 12 cysteine residues (Fig. 1A) known to form six structural disulfide bonds that also exist in the corresponding positions of other known PLA$_2$s from plants [22]. The histidine (at position 49 in GmsPLA$_2$-XIA-1, Fig. 1A) was suggested to play a crucial role in the nucleophilic attack at the *sn-2* bond in the glycerol backbone of phospholipids for all sPLA$_2$s [14,19,20].

The existence of the N-terminal signal peptide in these proteins reinforces the secretory nature of these enzymes, as it was suggested for other plant sPLA$_2$s [16,21,22]. The potential cleavage positions of the N-terminal signal peptide in the putative sPLA$_2$s present *in G. max* was predicted by using the available SignalIP tool. SignalIP computer analysis suggests that the signal sequence preceding the N-terminal amino acid of mature enzyme contains 24, 25, 29, 28 and 30 residues long, yielding a mature protein of 114, 115, 128, 128, and 128 amino acids (deduced molecular weights are: 12.3 kDa, 12.,6 kDa, 13.9 kDa, 13.9 kDa, 14 kDa) for GmsPLA$_2$-XIA-1, GmsPLA$_2$-XIA-2, GmsPLA$_2$-XIB-1, GmsPLA$_2$-XIB-2 and GmsPLA$_2$-XIB-3, respectively.

The expected isoelectric point (p*I*) for each putative enzyme has been calculated with PROTPARAM to be 6.93, 7.42, 5.76, 5.76 and 6.78 for GmsPLA$_2$-XIA-1, GmsPLA$_2$-XIA-2, GmsPLA$_2$-XIB-1, GmsPLA$_2$-XIB-2 and GmsPLA$_2$-XIB-3, respectively. As it can be observed, four of the putative sPLA$_2$s are rather acidic or neutral (GmsPLA$_2$-XIA-1, GmsPLA$_2$-XIB-1, -2 and -3) as in the case reported for sPLA$_2$s isolated from *Bothrops diporus* (*ex-neuwiedii*) venom [65,66]. Acidic sPLA$_2$s were also reported for some enzymes found in the *Crotalinae* subfamily [67] and those found in rice (isoform I and III) [21]. On the other hand, the expected p*I* of GmsPLA$_2$-XIA-2 is slightly alkaline similar to all the sPLA$_2$s found in *Arabidopsis* [21]. Other known sPLA$_2$s from plants were shown to have a p*I* almost neutral as it is the case of those found for carnation and tomato [21]. The functional role of the diverse p*I*s found in different sPLA$_2$s has not clearly been elucidated yet.

3.2. Cloning, expression and purification of GmsPLA$_2$-XIA-1

In order to obtain functional and structural information on these novel secreted phospholipases, we selected GmsPLA$_2$-XIA-1 as

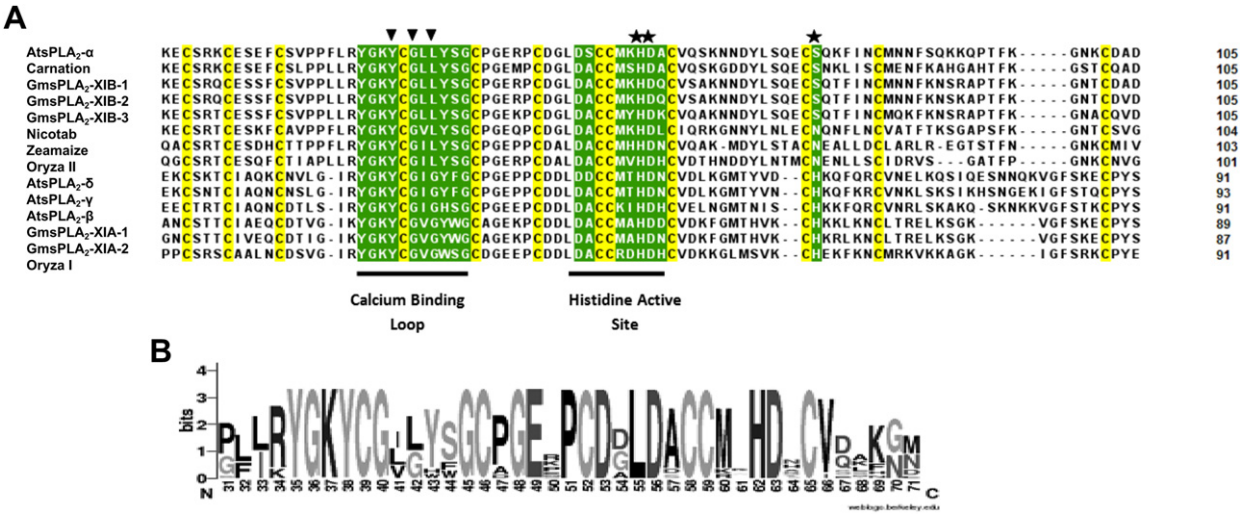

Fig. 1. Alignment of the amino acid sequences of reported sequence of plant sPLA$_2$s including the sequence of *Glycine max*. (A) The alignment was performed by using the program Clustal X (version 1.64b) and edited with the Jalview program. The triangles denote the amino acid residues involved in binding Ca^{2+} and the stars denote amino acid residues putatively involved in catalysis. In yellow are marked cysteine residues. (B) Sequence logo (http://weblogo.brekeley.edu) of the conserved motif of plants sPLA$_2$s. The height of each letter is proportional to its frequency. The height of the entire stack (*y* axis) is adjusted to signify the information content of sequences at that position (measured in bits). The *x*-axis represents the 41 positions (corresponding to 31 to 71 positions from the alignment) of the most conservative domains (histidine active site and Ca^{2+} binding site). (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

representative of the identified isoforms. It has been reported that an extra residue at the N-terminus in some animals sPLA$_2$s can substantially reduce or completely inhibit the activity [68–70]. The counterpart information for plant sPLA$_2$ is unknown. To overcome this potential problem, we have fusioned the GmsPLA$_2$-XIA-1 enzyme to the ubiquitin moiety in the PHUE system. The system is designed to obtain a recombinant protein without an extra residue at the N-terminus and also enable to yield high expression in inclusion bodies, as it was previously used in other proteins [30]. Thus, the gene sequence that encodes for the mature GmsPLA$_2$-XIA-1 was fused with a sequence encoding ubiquitin and cloned into the SacII and HindIII sites of the PHUE vector. To obtain the recombinant GmsPLA$_2$-XIA-1 in bacteria, the recombinant vector was used to transform *E. coli* BL21 (DE3). When the expression of the recombinant protein was induced by IPTG, the enzyme accumulated in inclusion bodies was isolated by solubilization with 6 M Gdn·HCl. In a first step of purification, this fusioned protein was purified by using a nickel affinity chromatography. As shown in Fig. 2A (line 1), a band of about 24 kDa was detected in the mixture and assigned to the fusioned protein. The mature protein was obtained almost pure after a second purification step, and further digested with USP2cc deubiquitylating peptidase (Fig. 2B). Since renaturation was critically dependent on the composition of the redox shuffling system, optimal refolding was achieved by dialyzing the enzyme against the refolding Tris–HCl buffer containing cysteine and $CaCl_2$. The presence of Ca^{+2} throughout all the purification steps was completely necessary for folding of the protein with enzymatic activity. The purification steps was followed by using the indirect screening assay of GmsPLA$_2$-XIA-1 activity by the appearance of halos in plates of agar-egg PC-blood as described above (data not shown). Activity toward of DLPC/Triton X-100 mixed micelles (1:4, mole ratio, see Ref. [71]) at pH 7 shows an apparent K_M of 0.64 mM and a V_{max} of 376 μmol min^{-1} mg^{-1} (Fig. 2C). The specific activity found for recombinant GmsPLA$_2$-XIA-1 was about 14 times lower than that reported for AtsPLA$_2$-α from *Arabidopsis*, although a different substrate system was used instead [22]. Uncleaved fused protein show no activity in all systems assayed.

3.3. Phylogenetic analysis and classification of the GmsPLA$_2$ isoforms

All the sPLA$_2$s sequences reported for plants were aligned using Clustal X in order to determine the extent of similarity and identity among them. When compared, all the data showed a close evolutionary relationship in the different plants. Fig. 3 shows the most important parts of the alignment of the selected sPLA$_2$ sequences reported for plants. To determine if sPLA$_2$s from *G. max* might fit into the existing XIA or XIB groups, the alignment of the sequences based on amino acid sequence similarities of all recently identified proteins recognized as plant sPLA$_2$s were analyzed in order to propose a phylogenetic tree. The analysis reveals that *G. max* sPLA$_2$ enzymes can be sorted in two distinct groups according to the classification done by Six and Dennis [19]. According to the PLA$_2$ superfamily numbering system [13,14,19] these categories would correspond to the subgroups XIA and XIB. GmsPLA$_2$-XIB-1, -2 and -3 belong to group XIB, with the existing AtsPLA$_2$-α, *O. sativa* isoform II, *carnation*, *N. tabacum* and *Zea maize*. On the other hand, GmsPLA$_2$-XIA-1 and -2 are taking part of group XIA, which includes AtsPLA$_2$-γ, AtsPLA$_2$-β, AtsPLA$_2$-δ, *O. sativa* isorform I (see Fig. 3). The comparison of the level of overall amino acid sequence identities within each group is about 47.6% and 57.3% for group XIA and XIB, respectively; while the identity between the two subgroups is only about 29%. The highest similarity of amino acid sequences was observed between GmsPLA$_2$-XIA-1 and GmsPLA$_2$-XIA-2, being of

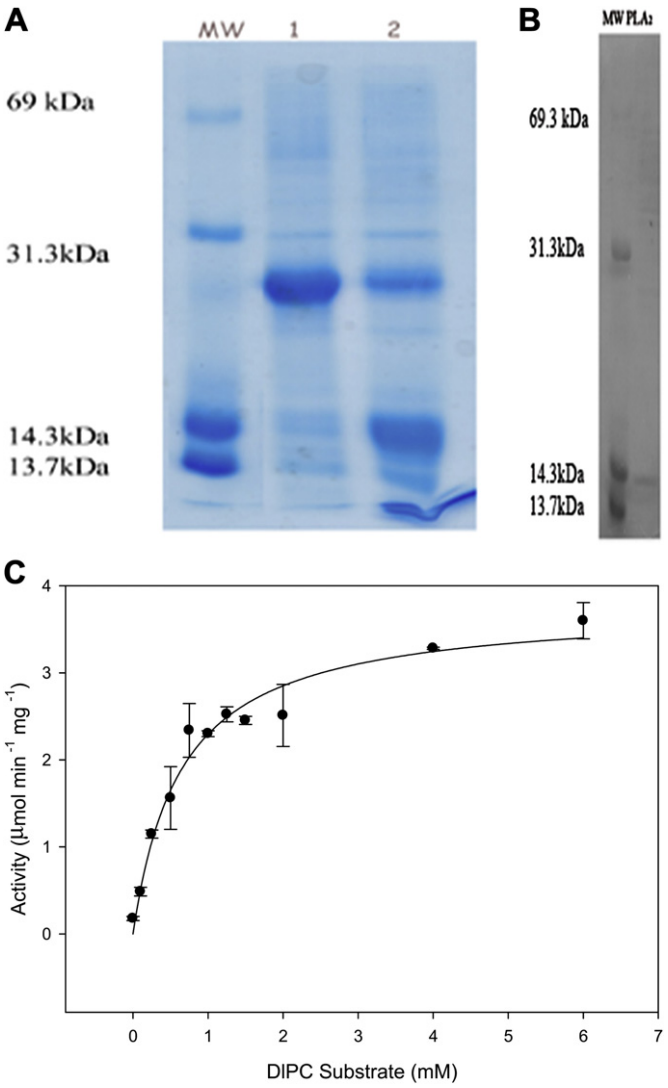

Fig. 2. Expression, purification and activity of GmsPLA$_2$-XIA-1. (A) SDS-PAGE analysis of the GmsPLA$_2$-XIA-1 enzyme purification. The first line MW on the left shows the protein markers; line 1 shows the recombinant protein Poly-Histag-ubiquitin-GmsPLA$_2$-XIA-1 eluted from inclusion bodies without treatment with USP2cc deubiquitylating peptidase; line 2 shows the digested proteins with USP2cc peptidase recombinant fusioned proteins (MW 24 kDa), GmsPLA$_2$-XIA-1 (MW 11.8 kDa) and the cleaved ubiquitin (MW 10 kDa). (B) SDS-PAGE showing the purification fraction of the recombinant protein eluted from nickel affinity chromatography. (C) The activity of purified GmsPLA$_2$-XIA-1 as function of DLPC substrate concentration in mixed micelles (see the text for details).

94.9%. Between GmsPLA$_2$-XIB-1 and GmsPLA$_2$-XIB-2 the extent of similarity is about 93.6%. In Table S1 of the supplementary data, the identity level among the different sPLA2s isoforms present in *G. max* can be observed. All calculations of identities are based on the sequence of the mature proteins.

As it was described by Mansfeld et al. [22] enzymes from these subgroups differ in the third Ca^{2+} coordinating amino acid, being a Gly residue in subgroup XIA or Leu residue in subgroup XIB, see Fig. 1A. As it was discussed previously, the common HIS/ASP catalytic dyad found in animal sPLA$_2$s does not correlate in the counterpart found in plants. Instead of the additional aspartate residue of the dyad located upstream of the conserved HIS-ASP pair, the plant enzymes that belong to group XIA contain a histidine residue, and the enzymes of group XIB contain a serine or an asparagine residue [22]. The biological/functional role of these differences upon catalytic properties concerning to lipid specificity (either for

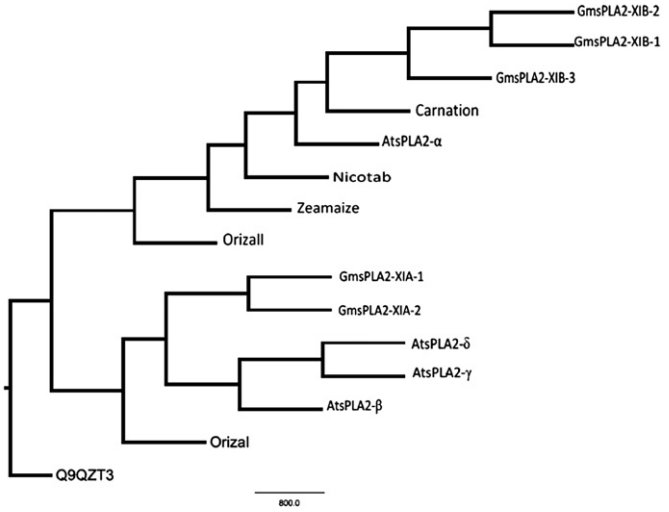

Fig. 3. Phylogenetic Cladogram shows the distribution of sPLA$_2$ *Glycine max* within the sPLA$_2$ family of plants. The program used for the alignment of the sequences was Clustal X (version 1.64b) and phylogenetic analysis by parsimony was performed with the PHYLIP software package. The numbers at nodes indicate the number of replicas obtained by 1000 bootstrap replicates. The root was determined using the sequence of the rat sPLA$_2$ (Q9QZT3) as outgroup. The Genbank ID are for *Carnation*, *Zea mays*, *O. sativa* isoform-I and -II, *Arabidopsis thaliana* (isoforms alpha, beta, gamma and delta) and *Nicotiana tabacum*, AF064732, EU968759, AJ238116, AJ238117, AY136317, AF541915, AY148346, AY148447 and AB190177 respectively, while for the isoforms of sPLA$_2$ *Glycine max* (GmsPLA$_2$-XIA-1, -2 and GmsPLA$_2$-XIB-1, -2, and -3) are ACU16523, ACU18951, ACU19490, ACU15282 and ACU24341, respectively.

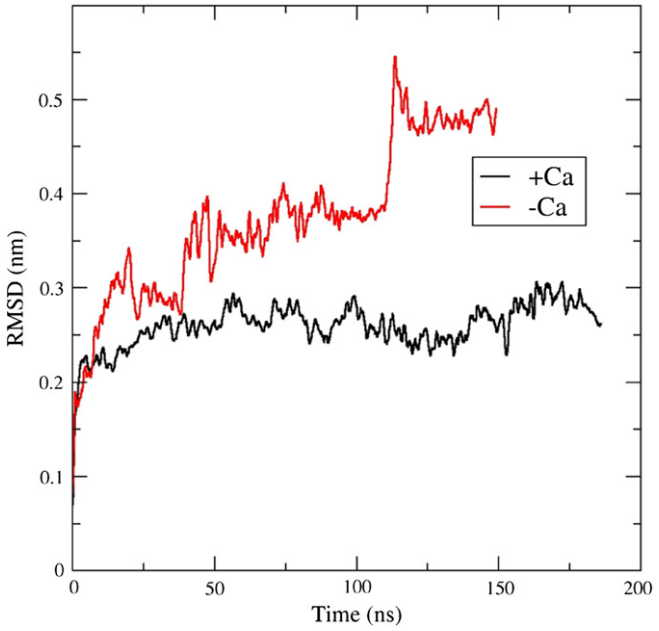

Fig. 4. Root mean square deviation (rmsd) of the alpha carbon from the initial structure of GmsPLA$_2$-XIA-1. The curve in black is the average of the four runs of the protein with calcium, while the red curve correspond to the average of the two runs of the protein without the calcium ion. The individual rmsd curves are shown in Fig. S3 of supplementary data. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

subtle differences in the affinity for a particular lipid polar head group or in the hydrophobic chain at the *sn-2* position) or special empathy for a particular physical state on the organized lipid substrate, compared with animals' sPLA$_2$s, has not been completely elucidated yet.

3.4. Molecular dynamics

3.4.1. Molecular dynamics in water of GmsPLA$_2$-XIA-1

In order to refine the model and to give insight into the structural dynamics of the protein we performed molecular dynamics simulations of the complete mature sequence of GmsPLA$_2$-XIA-1 protein. The strategy we employed was to perform simulations with two different force field approaches. Each force field runs initialized with a different set of initial velocities. This strategy is not commonly used, and we give a brief explanation for its application in this work. As there is no current gold standard in force fields, we decided to use two different options and obtain the consensus picture from the simulations. The repetition with different seeds helps us to explore the conformational space around the initial structure more efficiently than a single long run with equal added time [72].

The average for the four simulations in the root mean square deviation (rmsd) of the alpha carbon from the initial structure is shown in Fig. 4. The curves for the individual runs are shown in Fig. S3 of supplementary data. In three of the four simulations, the protein remains within 0.22–0.29 nm from the initial structure. These values are typically obtained when simulating proteins derived from X-ray or NMR experimental data. Taken together these results make us confident that we have obtained a good model of the protein.

In Fig. S3 of supplementary data, it is also observed that one of the simulations with Amber force field approach starts to increase the rmsd near 50 ns and reaches a value of almost 0.44 nm. A closer observation of this trajectory shows that at this point the calcium ion is lost from the protein. This indicates that the calcium atom plays a pivotal role in maintaining the protein folded in its native structure. For this reason we decided to study the effect of the ion calcium. The results are shown in a section below. Given the lost of the calcium ion, in the following calculations we only take into account the first 50 ns of this simulation.

Fig. 5 shows the root mean square fluctuation (rmsf) of the alpha carbons, which is an indicator of the flexibility of the backbone. The curve shows the weighted average from the four simulations of the holo protein, and indicates that the most mobile regions are the N- and C-terminal, followed by the loops in residues 74–85, 53–62, 34–37 of GmsPLA$_2$-XIA-1. These loops connect, respectively, the last two helices, the first with the second helix, and the last beta-sheet with the first helix. In Fig. S4 of supplementary data the individual rmsf curves are shown. These data show that there are no major differences between runs and force fields, which indicate that the flexibility inferred is robust and very likely.

Table 1 shows the weighted average of the secondary structure of the GmsPLA$_2$-XIA-1 protein. As other sPLA$_2$s in the family, the principal secondary structure is the alpha helix, with only a small portion of beta sheet. Also there are abundant regions in turns and bends. In this structural aspect there are differences between the force fields. The AMBER force field gives consistently higher helix content, while for GROMOS there are fewer residues with secondary structures and the number of bends is increased. The difference in the number of residues in alpha helix between both force fields is concentrated mainly in the terminal helix, where the AMBER force field retains more the initial structure which is fully helical. Given the low number of contacts with the rest of the protein and the high water exposure of this terminal helix, it is possible that this helix is not as stable as the others and that it suffers more transitions to coil structures. Therefore, the results obtained with the GROMOS force field are more likely. Nevertheless, consistently, these differences between force fields are small since they amount to only 5%.

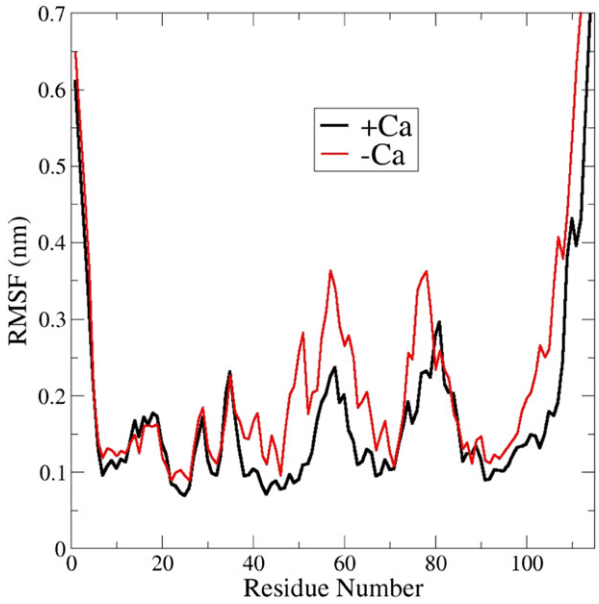

Fig. 5. Root mean square fluctuations of the alpha carbons (rmsf). The curve in black is the weighted average of the four runs of the GmsPLA$_2$-XIA-1 in presence of calcium, while the red curve correspond to the weighted average of the two runs of the protein without calcium ion. The individual rmsf curves are shown in Fig. S4 of supplementary data. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

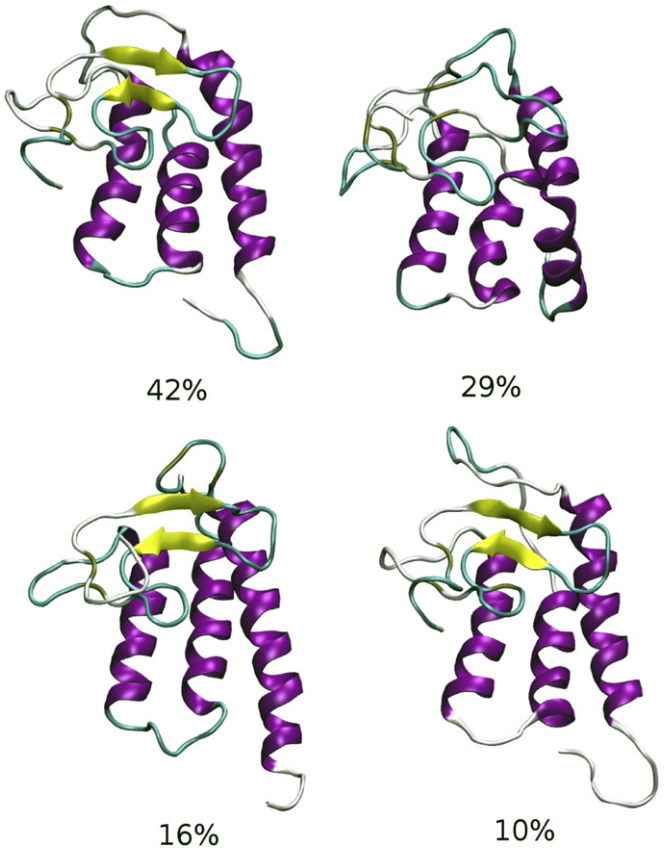

Fig. 6. Centers of the clusters found using a rmsd cutoff of 0.2 nm. The helices are shown in magenta, the beta sheets in yellow, the turns in cyan, and the residues in coil are white. The figure was generated using the program VMD. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

In order to analyze the tertiary structure of the GmsPLA$_2$-XIA-1 protein we performed a clustering of the structures sampled in all the trajectories [73]. In this analysis, the structures are clustered by similarity measured by the rmsd, which was set to 0.2 nm in this case. The center of the first four clusters is shown in Fig. 6. Together these four clusters represent the 97% of the structures sampled. The main differences observed between the cluster centers are principally located in residues 99–114 (the last residues of helix 3) and residues 74–85 (the loop connecting the last two helices). The principal cluster which amounts to a 42% of the population has the last 7 residues in coil. These residues are in full alpha helix in the initial structure as well as in the crystal structure of *O. sativa* sPLA$_2$. The second cluster, with a 29% of the population, has also lost the helix in the same segment of the last helix, and shows a kink in this helix occurring at residue 99. This kink may be caused by the presence of a GLY residue at position 98, which is known for its low intrinsic heliticity. In the third cluster (population 16%) the terminal helix is fully formed. The fourth cluster has the terminal residues in a similar conformation to the first cluster, but has a different conformation comprising residues 74–85. This observation indicates that the terminal helix is rather a dynamic region and has three principal conformations, one fully helical, other with the last seven residues in coil, and the third one with a kink plus coil. As noted before, this behavior can be attributed to the low contact number of this region, its high solvent exposed area, and the presence of the flexible residue GLY98.

Table 1
Secondary structure[a] of the GmsPLA$_2$-XIA-1 protein in presence ($+Ca^{+2}$) and absence ($-Ca^{+2}$) of calcium in the active site.

	$+Ca^{+2}$	$-Ca^{+2}$
Helix	47	44
Turn	13	16
Bend	20	18
Beta bridge	3	2
Beta sheet	4	6
Coil	27	28

[a] The secondary structure was calculated with the program DSSP.

The active site of the protein contains a calcium ion commonly present in other sPLA$_2$s from plants [22,23]. The calcium ion is a crucial cofactor in the catalysis mechanism and is a prerequisite for the enzyme activity. In GmsPLA$_2$-XIA-1, the calcium ion is coordinated by the backbone carbonyls of TYR25, GLY27, and TYR30 (Fig. 7). The carboxylic side chain of ASP50 and three water molecules complete the coordination of the calcium ion. One of the water molecules makes a bridge between the calcium ion and HIS49. The HIS49 residue is known to be part of the catalytic site of all known sPLA$_2$s regardless whether they are from animals or plants, and is localized in a cavity that is accessible from the *i-face*, thus mediating the hydrolysis of the substrate. The HIS-ASP pair constitutes the active center and the calcium binding loop is essential for the proper function of the enzyme [25]. All secreted PLA$_2$s catalyze the hydrolysis through the same mechanism of abstraction of a proton from a water molecule followed by a nucleophilic attack on the *sn-2* bond of diacylglycerophospholipids. One water molecule is activated by the presence of a HIS/ASP dyad with the absolute dependence of the coordinated Ca^{2+} ion [74,75]. This correlates with the catalytic mechanism proposed called the Ca^{+2}-coordinated oxyanion mechanism [76], where two water molecules are involved. The two water molecules present in the structure surrounding the calcium binding site suggest that it is essential that one water molecule be bound to the histidine active site through the reorientation and activation in the catalytic mechanism. The second water molecule seems to be

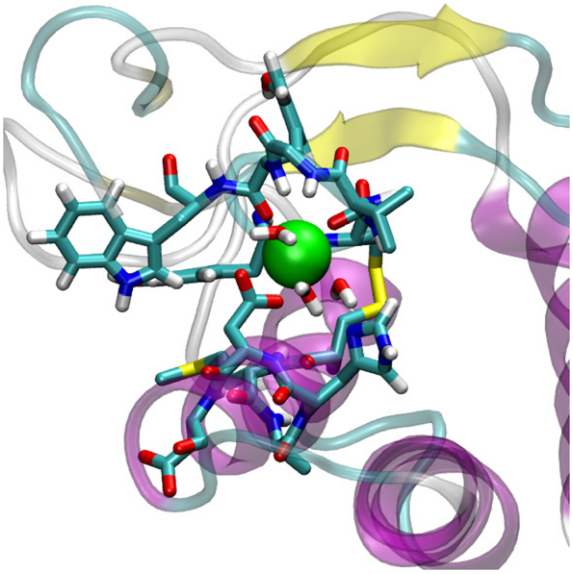

Fig. 7. Active site of the GmsPLA$_2$-XIA-1. The protein backbone is shown in cartoon (with the same color code as Fig. 6, the calcium ion is in green, and residues 25–31 and residues 46–51 are in sticks). There are also three water molecules coordinating the calcium ion. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

involved in the catalytic mechanism, connecting the calcium coordinated water with the histidine of the active site. In Fig. S5 of supplementary data, the number of contacts between the calcium ion and the side chain of the protein is shown for the individual simulations.

3.4.2. Molecular dynamics in the absence of calcium ion

In order to evaluate the effect of the calcium ion in the structure of the protein, we also performed simulations of the GmsPLA$_2$-XIA-1 in absence of this cation in the active site. When the calcium ion is removed from the active site, the structure deviates more with respect to the initial model, as is clearly shown in Fig. 4. Also, the fluctuations of the alpha carbons are more pronounced (Fig. 5). This increase in flexibility is more evident in the terminal helix and the loops connecting the helices. The protein is tightly connected by a net of disulfide bridges, since it contains six disulfide bridges which connect cysteines 6–33, 10–39, 15–86, 26–46, 45–70 and 52–63. The position and the number of cysteine residues are well conserved among all the sPLA$_2$s from plants (Fig. 1A). Thus, the cysteines are important for the structural stability of the enzyme and for the activity as well as for animal sPLA$_2$s. As the majority of the residues involved in this disulfide bridges are in the first 50 amino acid of the protein sequence, this region is less affected by the absence of calcium (Fig. 5). The secondary structure is slightly affected by the removal of calcium (Table 1). The retention of a major portion of the secondary structure after removal of the calcium ion may be due to the already mentioned high number of disulfide bridges which stabilize the native structure, or it may simply imply that the simulated time is not long enough to observe a complete unfolding. Taken together all the data, the results emerging from simulations indicate that the calcium ion plays an important role not only in the catalytic cycle but also in the overall flexibility of native structure. This finding is keeping with the necessity of calcium ion for a correct folding (indicated by its enzymatic activity) in the recombinant expressed protein (see above). In order to further demonstrate the role of calcium ion on protein structure, we have investigated the effect of this cation on the intrinsic protein fluorescence as an indirect measurement of the compactness of protein tertiary structure. GmsPLA$_2$-XIA-1 has an only one TRP residue in the calcium loop domain at the 31st position of the sequence (Fig. 1A). The steady-state intrinsic fluorescence emission has already been used to determine more exposed domains of lipolytic enzymes where TRP is involved [77]. Both the maximum emission of TRP and its fluorescence intensity are very sensitive to environmental conditions such as exposure to the solvent, dynamic quenching of soluble quenchers or neighboring groups [78]. In an enzymatic active conformation in presence of 1 mM of Ca^{+2}, the maximum emission is centered at 335 nm indicative of a rather less exposed environment of the domain in which the TRP is located. Depletion of calcium by adding EDTA in the medium shifted the maximum emission to 337 nm and reduced the fluorescence intensity. In denaturing condition of high Gdn·HCl, the fluorescence profile has the typical behavior of TRP in water with a maximum emission at 352 nm (Fig. S7 of supplementary data). Fluorescence data is also indicative that Ca^{+2} is inducing an increase in protein compactness in agreement with model obtained from molecular dynamics.

3.4.3. Molecular dynamics in mixed solvent

In order to gain insight into the putative membrane face of the protein, we performed molecular dynamics simulations of the GmsPLA$_2$-XIA-1 protein in a solvent which is a mixture of water and 30% isopropyl alcohol using the GROMOS force field. The central idea behind this approach is that the regions of the protein that are going to contact the membrane need to lose, totally or partially, their solvating water molecules in order to interact with a more hydrophobic surface such as the membrane interface. The regions with a high tendency to desolvate may be spotted out by the increase in the local concentration of isopropyl alcohol molecules compared to the bulk solvent. This reasoning is similar to the one used in the detection of hot spots for the binding of drugs in the surfaces of proteins [79]. With this in mind, we performed one simulation of 60 ns of the protein in mixed solvent. Before calculating the isopropyl density, the stability of the protein was monitored by measuring the rmsd, the rmsf, and secondary structure content. All these parameters showed that the structure simulated in mixed solvent behaves similarly to the structure in water (Fig. S6 of supplementary data).

Fig. 8 shows an isosurface of density equal to 10 times the average isopropyl alcohol density in bulk. In this figure, two different patches can be distinguished, one thin patch that is mainly on top of the first helix and runs below the loop in residues 74–85, and one thick patch surrounding the C-terminal helix. This last patch directly contacts the catalytic HIS49 and therefore we propose it as the membrane *i-face* of the enzyme. The residues that are in this face are VAL18, GLY19, VAL28, HIS49, HIS64, LEU101, ALA102, ILE103, LEU104, LEU105 and LEU108. Note that the majority of these residues are hydrophobic as expected for a membrane binding surface, so this constitutes a hydrophobic environment surrounding the active site. The hydrophobic side chains of the residues that putative are forming the *i-face* are able to partition to the interface, allowing the interfacial binding, and excluding the water molecules from the region that surrounds the active site that could permit the diffusion of the substrate into the pocket of the active site to be hydrolyzed. As it is known, sPLA$_2$s must penetrate the phospholipid interface to exert their action. Thus, the successful binding to the phospholipid surface is a prerequisite step in the catalytic cycle, and this property may also determine some specificities of enzymatic activity. Nevertheless, there are a limited number of charged residues alongside the planar topography of the *i-face* that may further modulate the interaction with the substrate interface in a way in which has not been fully elucidated yet [24].

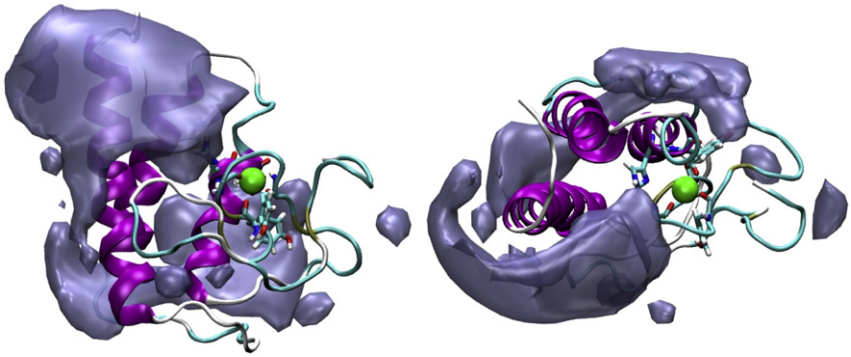

Fig. 8. Top and side view of an isosurface of density equal to ten times the value of isopropyl alcohol density in the bulk. The GmsPLA$_2$-XIA-1 protein is shown in cartoons with the same color code as in Fig. 6. The calcium ion is in green and the residues in the catalytic site are in sticks. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

4. Conclusions

The BLAST search clearly demonstrates the existence of five putative isoforms of sPLA$_2$ within the *G. max* genome. In spite of a relatively low identity among them, the sequence analysis performed shows the presence of clusters with conserved amino acid residues, when compared with those reported sequences from plant sPLA$_2$s. The catalytic motif, the Ca^{2+}-binding loop and the active site domain are well conserved. Besides, alignment studies showed the existence of conserved residues in the plant enzymes absent in sPLA$_2$s from animals. However, the relationship between the proposed protein structure and the modulation of the "*interfacial quality of the interface*" on catalytic activity still remains to be assessed.

The five sPLA$_2$s identified in *G. max* were consistently grouped into two categories by phylogenetic analysis which separates them clearly from the other sPLA$_2$ from plants into two subgroups XIA and XIB based on the differences in their molecular weights and deviating sequences especially in the N- and C- terminal regions of the proteins. Moreover, there are marked differences in the amino acids that are essential for catalysis. The role of ASP residue in the dyad HIS/ASP responsible for catalytic activity in animal sPLA$_2$s, has been replaced by HIS residue for GmsPLA$_2$-XIA-1 and -2, and by SER residue for GmsPLA$_2$-XIB-1, -2 and -3, a distinctive characteristic of phospholipases coming from plants.

We have expressed and purified the mature form of GmsPLA$_2$-XIA-1. The recombinant enzyme showed PLA$_2$ activity against phospholipid:Triton X-100 mixed micelles.

We have created a model for the GmsPLA$_2$-XIA-1 protein based on the structure template reported for rice (*O. sativa*) isoform II, in order to gain insight into its structure and dynamics. The simulations, which added a total time of over 700 ns, were performed in pure water (in the presence or absent of Ca^{+2} in the active site), and in a water–isopropyl alcohol mixture. The initial fold is globally retained, suggesting that indeed the structure of *G. max* sPLA$_2$ is very similar to *O. sativa* isoform II. We observed a conformational flexibility in the last residues of C-terminal helix. This portion, of the otherwise long helix, samples coil conformations. It is possible that upon interaction with membranes, the fully helical state is recovered.

Despite of being highly connected by a network of six disulfide bridges, the molecular dynamic data suggest that the removal of the calcium ion from the binding site destabilize the structure. The apoprotein is more flexible, but there is only a little lost of secondary structure. The intrinsic fluorescence data from the unique TRP residue localized in the Ca^{+2} binding domain in presence or in absence of this cation is in keeping with molecular dynamics simulation.

The simulations in the mixed solvent allowed us to identify a putative *i-face* which comprises residues VAL18, GLY19, VAL28, HIS49, HIS64, LEU101, ALA102, ILE103, LEU104, LEU105 and LEU108, and directly contacts the catalytic HIS49.

Acknowledgments

This work was supported, in part, by grants from MinCyT Córdoba (ex-Agencia Córdoba Ciencia), CONICET, FONCYT and SeCyT-UNC. M.E.M. is a fellow from MinCyT Córdoba-CONICET under special agreement. M.A.V., E.P.M.L. and G.D.F. are members of the Scientific Research Career (CIC) from CONICET. We thank to Ann-Maree Catanzariti for the PHUE vector system. The clone was generated under funding from the USDA (Grant Number 2006-35300-17144) to the J. Craig Venter Institute (P.I. Christopher D. Town). We also thank Rodrigo Quiroga for useful comments in bioinformatic analysis.

Appendix A. Supplementary data

Supplementary data related to this article can be found online at http://dx.doi.org/10.1016/j.biochi.2012.07.021.

References

[1] M.J. Messina, Legumes and soybeans: overview of their nutritional profiles and health effects, Am. J. Clin. Nutr. 70 (1999) 439S–450S.
[2] M.S. Choi, K.C. Rhee, Production and processing of soybeans and nutrition and safety of isoflavone and other soy products for human health, J. Med. Food 9 (2006) 1–10.
[3] Z. Guo, A.F. Vikbjerg, X. Xu, Enzymatic modification of phospholipids for functional applications and human nutrition, Biotechnol. Adv. 23 (2005) 203–259.
[4] L. De Maria, J. Vind, K.M. Oxenboll, A. Svendsen, S. Patkar, Phospholipases and their industrial applications, Appl. Microbiol. Biotechnol. 74 (2007) 290–300.
[5] J. Mansfeld, Plant phospholipases A2: perspectives on biotechnological applications, Biotechnol. Lett. 31 (2009) 1373–1380.
[6] U. Stahl, M. Lee, S. Sjodahl, D. Archer, F. Cellini, B. Ek, R. Iannacone, D. MacKenzie, L. Semeraro, E. Tramontano, S. Stymme, Plant low-molecular-weight phospholipase A2S (PLA2s) are structurally related to the animal secretory PLA2s and are present as a family of isoforms in rice (*Oryza sativa*), Plant Mol. Biol. 41 (1999) 481–490.
[7] J.Y. Kim, Y.S. Chung, S.H. Ok, S.G. Lee, W.I. Chung, I.Y. Kim, J.S. Shin, Characterization of the full-length sequences of phospholipase A2 induced during flower development, Biochim. Biophys. Acta 1489 (1999) 389–392.
[8] U. Stahl, B. Ek, S. Stymne, Purification and characterization of a low-molecular-weight phospholipase A2 from developing seeds of elm, Plant Physiol. 117 (1998) 197–205.
[9] H.Y. Lee, S.C. Bahn, Y.M. Kang, K.H. Lee, H.J. Kim, E.K. Noh, J.P. Palta, J.S. Shin, S.B. Ryu, Secretory low molecular weight phospholipase A2 plays important roles in cell elongation and shoot gravitropism in *Arabidopsis*, Plant Cell 15 (2003) 1990–2002.

[10] S.B. Ryu, Phospholipid-derived signaling mediated by phospholipase A in plants, Trends Plant Sci. 9 (2004) 229–235.
[11] G. Chen, C.L. Snyder, M.S. Greer, R.J. Weselake, Biology and biochemistry of plant phospholipases, Crit. Rev. Plant Sci. 30 239–258.
[12] M. Murakami, Y. Taketomi, H. Sato, K. Yamamoto, Secreted phospholipase A2 revisited, J. Biochem. 150 (2011) 233–255.
[13] R.H. Schaloske, E.A. Dennis, The phospholipase A2 superfamily and its group numbering system, Biochim. Biophys. Acta 1761 (2006) 1246–1259.
[14] J.E. Burke, E.A. Dennis, Phospholipase A2 structure/function, mechanism, and signaling, J. Lipid Res. 50 (Suppl.) (2009) S237–S242.
[15] M. Murakami, Y. Taketomi, Y. Miki, H. Sato, T. Hirabayashi, K. Yamamoto, Recent progress in phospholipase A research: from cells to animals to humans, Prog. Lipid Res. 50 (2010) 152–192.
[16] R. Fujikawa, Y. Fujikawa, N. Iijima, M. Esaka, Molecular cloning, expression, and characterization of secretory phospholipase A2 in tobacco, Lipids 40 (2005) 901–908.
[17] S.C. Bahn, H.Y. Lee, H.J. Kim, S.B. Ryu, J.S. Shin, Characterization of *Arabidopsis* secretory phospholipase A2-gamma cDNA and its enzymatic properties, FEBS Lett. 553 (2003) 113–118.
[18] J. Mansfeld, R. Ulbrich-Hofmann, Secretory phospholipase A2-alpha from *Arabidopsis thaliana*: functional parameters and substrate preference, Chem. Phys. Lipids 150 (2007) 156–166.
[19] D.A. Six, E.A. Dennis, The expanding superfamily of phospholipase A(2) enzymes: classification and characterization, Biochim. Biophys. Acta 1488 (2000) 1–19.
[20] O.G. Berg, M.H. Gelb, M.D. Tsai, M.K. Jain, Interfacial enzymology: the secreted phospholipase A(2)-paradigm, Chem. Rev. 101 (2001) 2613–2654.
[21] H.Y. Lee, S.C. Bahn, J.S. Shin, I. Hwang, K. Back, J.H. Doelling, S.B. Ryu, Multiple forms of secretory phospholipase A2 in plants, Prog. Lipid Res. 44 (2005) 52–67.
[22] J. Mansfeld, S. Gebauer, K. Dathe, R. Ulbrich-Hofmann, Secretory phospholipase A2 from *Arabidopsis thaliana*: insights into the three-dimensional structure and the amino acids involved in catalysis, Biochemistry 45 (2006) 5687–5694.
[23] J.E. Guy, U. Stahl, Y. Lindqvist, Crystal structure of a class XIB phospholipase A2 (PLA2): rice (*Oryza sativa*) isoform-2 pla2 and an octanoate complex, J. Biol. Chem. 284 (2009) 19371–19379.
[24] M.K. Jain, O.G. Berg, Coupling of the i-face and the active site of phospholipase A2 for interfacial activation, Curr. Opin. Chem. Biol. 10 (2006) 473–479.
[25] D.L. Scott, S.P. White, Z. Otwinowski, W. Yuan, M.H. Gelb, P.B. Sigler, Interfacial catalysis: the mechanism of phospholipase A2, Science 250 (1990) 1541–1546.
[26] S.A. Tatulian, S. Qin, A.H. Pande, X. He, Positioning membrane proteins by novel protein engineering and biophysical approaches, J. Mol. Biol. 351 (2005) 939–947.
[27] J.M. Winget, Y.H. Pan, B.J. Bahnson, The interfacial binding surface of phospholipase A2s, Biochim. Biophys. Acta 1761 (2006) 1260–1269.
[28] O.G. Berg, M.K. Jain, Interfacial Enzyme Kinetics, John Wiley & Sons, Ltd., 2002.
[29] M. Minchiotti, M.B. Scalambro, L. Vargas, C. Coronel, R. Madoery, Isolation of phospholipase A2 from soybean (*Glycine max*) seeds: the study of its enzymatic properties, Enzym. Microb. Technol. 42 (2008) 389–394.
[30] A.M. Catanzariti, T.A. Soboleva, D.A. Jans, P.G. Board, R.T. Baker, An efficient system for high-level expression and easy purification of authentic recombinant proteins, Protein Sci. 13 (2004) 1331–1339.
[31] M.J. Kelley, R.M. Crowl, E.A. Dennis, Renaturation of cobra venom phospholipase A2 expressed from a synthetic gene in *Escherichia coli*, Biochim. Biophys. Acta 1118 (1992) 107–115.
[32] U.K. Laemmli, Cleavage of structural proteins during the assembly of the head of bacteriophage T4, Nature 227 (1970) 680–685.
[33] J.M. Gutierrez, C. Avila, E. Rojas, L. Cerdas, An alternative in vitro method for testing the potency of the polyvalent antivenom produced in Costa Rica, Toxicon 26 (1988) 411–413.
[34] G.E. Hoffmann, D. Schmidt, B. Bastian, W.G. Guder, Photometric determination of phospholipase A, J. Clin. Chem. Clin. Biochem. 24 (1986) 871–875.
[35] A. Pearson, A blast from the past: whatever happened to the 'new nursing' and 'nursing beds'? Int. J. Nurs. Pract. 9 (2003) 67–69.
[36] R.D. Finn, J. Mistry, B. Schuster-Bockler, S. Griffiths-Jones, V. Hollich, T. Lassmann, S. Moxon, M. Marshall, A. Khanna, R. Durbin, S.R. Eddy, E.L. Sonnhammer, A. Bateman, Pfam: clans, web tools and services, Nucleic Acids Res. 34 (2006) D247–D251.
[37] E. Gasteiger, A. Gattiker, C. Hoogland, I. Ivanyi, R.D. Appel, A. Bairoch, ExPASy: the proteomics server for in-depth protein knowledge and analysis, Nucleic Acids Res. 31 (2003) 3784–3788.
[38] O. Emanuelsson, S. Brunak, G. von Heijne, H. Nielsen, Locating proteins in the cell using TargetP, SignalP and related tools, Nat. Protoc. 2 (2007) 953–971.
[39] J.D. Thompson, T.J. Gibson, D.G. Higgins, Multiple sequence alignment using ClustalW and ClustalX, Curr. Protoc. Bioinform. (2002) (Chapter 2, Unit 2 3).
[40] J.D. Retief, Phylogenetic analysis using PHYLIP, Methods Mol. Biol. 132 (2000) 243–258.
[41] G.E. Crooks, G. Hon, J.M. Chandonia, S.E. Brenner, WebLogo: a sequence logo generator, Genome Res. 14 (2004) 1188–1190.
[42] J.J. Campanella, L. Bitincka, J. Smalley, MatGAT: an application that generates similarity/identity matrices using protein or DNA sequences, BMC Bioinformatics 4 (2003) 29.
[43] Y. Zhang, I-TASSER server for protein 3D structure prediction, BMC Bioinformatics 9 (2008) 40.
[44] Y. Zhang, I-TASSER: fully automated protein structure prediction in CASP8, Proteins 77 (Suppl. 9) (2009) 100–113.
[45] C. Oostenbrink, A. Villa, A.E. Mark, W.F. van Gunsteren, A biomolecular force field based on the free enthalpy of hydration and solvation: the GROMOS force-field parameter sets 53A5 and 53A6, J. Comput. Chem. 25 (2004) 1656–1676.
[46] V. Hornak, R. Abel, A. Okur, B. Strockbine, A. Roitberg, C. Simmerling, Comparison of multiple Amber force fields and development of improved protein backbone parameters, Proteins 65 (2006) 712–725.
[47] H.J.C. Berendsen, J.R. Grigera, T.P. Straatsma, The missing term in effective pair potentials, J. Phys. Chem. 91 (1987) 6269–6271.
[48] W.L. Jorgensen, J. Chandrasekhar, J.D. Madura, R.W. Impey, M.L. Klein, Comparison of single potential functions for simulating liquid water, J. Chem. Phys. 79 (1983) 2577–2637.
[49] E. Project, E. Nachliel, M. Gutman, Parameterization of Ca^{+2}-protein interactions for molecular dynamics simulations, J. Comput. Chem. 29 (2008) 1163–1169.
[50] T. Darden, D. York, L. Pedersen, Particle mesh Ewald: an $N - \log(N)$ method for Ewald sums in large systems, J. Chem. Phys. 98 (1993) 10089–10092.
[51] U. Essmann, L. Perera, M.L. Berkowitz, T. Darden, H. Lee, L.G. Pedersen, A smooth particle mesh Ewald potential, J. Chem. Phys. 103 (1995) 8577–8592.
[52] G. Bussi, D. Donadio, M. Parrinello, Canonical sampling through velocity rescaling, J. Chem. Phys. 126 (2007) 014101.
[53] H.J.C. Berendsen, J.P.M. Postma, A. DiNola, J.R. Haak, Molecular dynamics with coupling to an external bath, J. Phys. Chem. 81 (1984) 3684–3690.
[54] B. Hess, H. Beckker, H.J.C. Berendsen, J.G.E.M. Fraaije, LINCS: a linear constraint solver for molecular simulations, J. Comput. Chem. 18 (1997) 1463–1472.
[55] S. Miyamoto, P.A. Kollman, SETTLE: an analytical version of the SHAKE and RATTLE algorithms for rigid water models, J. Comput. Chem. 13 (1992) 952–962.
[56] K.A. Feenstra, B. Hess, H.J.C. Berendsen, Improving efficiency of large time-scale molecular dynamics simulation of hydrogen-rich systems, J. Comput. Chem. 20 (1999) 786–798.
[57] W. Kabsch, C. Sander, Dictionary of protein secondary structure: pattern recognition of hydrogen-bonded and geometrical features, Biopolymers 22 (1983) 2577–2637.
[58] D. Van Der Spoel, E. Lindahl, B. Hess, G. Groenhof, A.E. Mark, H.J. Berendsen, GROMACS: fast, flexible, and free, J. Comput. Chem. 26 (2005) 1701–1718.
[59] B. Hess, C. Kutzner, D. van Der Spoel, E. Lindahl, GROMACS 4: algorithms for highly efficient, load-balanced, and scalable molecular simulations, J. Chem. Theory Comp. 4 (2008) 435–447.
[60] N. Singh, R.K. Somvanshi, S. Sharma, S. Dey, P. Kaur, T.P. Singh, Structural elements of ligand recognition site in secretory phospho-lipase A2 and structure-based design of specific inhibitors, Curr. Top. Med. Chem. 7 (2007) 757–764.
[61] B.J. Bahnson, Structure, function and interfacial allosterism in phospholipase A2: insight from the anion-assisted dimer, Arch. Biochem. Biophys. 433 (2005) 96–106.
[62] A. Kumar, C. Sekharudu, B. Ramakrishnan, C.M. Dupureur, H. Zhu, M.D. Tsai, M. Sundaralingam, Structure and function of the catalytic site mutant Asp 99 Asn of phospholipase A2: absence of the conserved structural water, Protein Sci. 3 (1994) 2082–2088.
[63] K. Sekar, B.Z. Yu, J. Rogers, J. Lutton, X. Liu, X. Chen, M.D. Tsai, M.K. Jain, M. Sundaralingam, Phospholipase A2 engineering. Structural and functional roles of the highly conserved active site residue aspartate-99, Biochemistry 36 (1997) 3104–3114.
[64] R.R. Annand, M. Kontoyianni, J.E. Penzotti, T. Dudler, T.P. Lybrand, M.H. Gelb, Active site of bee venom phospholipase A2: the role of histidine-34, aspartate-64 and tyrosine-87, Biochemistry 35 (1996) 4591–4601.
[65] J.J. Daniele, I.D. Bianco, G.D. Fidelio, Kinetic and pharmacologic characterization of phospholipases A2 from *Bothrops neuwiedii* venom, Arch. Biochem. Biophys. 318 (1995) 65–70.
[66] J.J. Daniele, I.D. Bianco, C. Delgado, D.B. Carrillo, G.D. Fidelio, A new phospholipase A2 isoform isolated from *Bothrops neuwiedii* (*Yarara chica*) venom with novel kinetic and chromatographic properties, Toxicon 35 (1997) 1205–1215.
[67] J.I. dos Santos, M. Cintra-Francischinelli, R.J. Borges, C.A. Fernandes, P. Pizzo, A.C. Cintra, A.S. Braz, A.M. Soares, M.R. Fontes, Structural, functional, and bioinformatics studies reveal a new snake venom homologue phospholipase A class, Proteins 79 61–78.
[68] S. Di Marco, F. Marki, H. Hofstetter, A. Schmitz, J. van Oostrum, M.G. Grutter, Purification, analysis, and enzymatic activity of recombinant human synovial fluid phospholipase A2 and N-terminal variants, J. Biochem. 112 (1992) 350–354.
[69] Y.L. Chiou, Y.C. Cheng, P.H. Kao, J.J. Wang, L.S. Chang, Mutations on the N-terminal region abolish differentially the enzymatic activity, membrane-damaging activity and cytotoxicity of Taiwan cobra phospholipase A2, Toxicon 51 (2008) 270–279.
[70] Y.L. Chiou, S.R. Lin, L.S. Chang, Mutations on N-terminal region of Taiwan cobra phospholipase A(2) result in structurally distorted effects, J. Pept. Sci. 14 (2008) 890–897.
[71] D. Lombardo, E.A. Dennis, Immobilized phospholipase A2 from cobra venom. Prevention of substrate interfacial and activator effects, J. Biol. Chem. 260 (1985) 16114–16121.

[72] L.S. Caves, J.D. Evanseck, M. Karplus, Locally accessible conformations of proteins: multiple molecular dynamics simulations of crambin, Protein Sci. 7 (1998) 649–666.
[73] X. Daura, K. Gademann, B. Jaun, D. Seebach, W.F. van Gunsteren, A.E. Mark, Peptide folding: when simulation meets experiment, Angew. Chem. Int. Ed. 38 (1999) 236–240.
[74] J.E. Burke, E.A. Dennis, Phospholipase A2 structure/function, mechanism, and signaling, J. Lipid Res. 50 (2009) S237–S242.
[75] L. Yu, E.A. Dennis, Critical role of a hydrogen bond in the interaction of phospholipase A2 with transition-state and substrate analogues, Proc. Natl. Acad. Sci. U. S. A. 88 (1991) 9325–9329.
[76] J. Rogers, B.Z. Yu, M.D. Tsai, O.G. Berg, M.K. Jain, Cationic residues 53 and 56 control the anion-induced interfacial k*cat activation of pancreatic phospholipase A2, Biochemistry 37 (1998) 9549–9556.
[77] A. Bourbon-Freie, R.E. Dub, X. Xiao, M.E. Lowe, Trp-107 and trp-253 account for the increased steady state fluorescence that accompanies the conformational change in human pancreatic triglyceride lipase induced by tetrahydrolipstatin and bile salt, J. Biol. Chem. 284 (2009) 14157–14164.
[78] J. Lakowicz, Principles of Fluorescence Spectroscopy, third ed., Springer, New York, 2006.
[79] J. Seco, F.J. Luque, X. Barril, Binding site detection and druggability index from first principles, J. Med. Chem. 52 (2009) 2363–2371.