Identification, description and structural analysis of beta phospholipase A<sub>2</sub> inhibitors (sbβPLIs) from Latin American pit vipers indicate a binding site region for basic snake venom phospholipases A<sub>2</sub>

Consuelo Latorre Fortes-Dias<sup>a</sup><sup>*,</sup> Carlos Alexandre H. Fernandes<sup>b,c</sup>, Paula Ladeira Ortolani<sup>a</sup>, Patrícia Cota Campos<sup>a</sup>, L.A. Melo<sup>b</sup>, Liza Figueiredo Felicori<sup>c</sup>, Marcos Roberto M. Fontes<sup>b</sup>

<sup>a</sup>Serviço de Enzimologia, Diretoria de Pesquisa e Desenvolvimento, Fundação Esquilo Dias (FUNED), Belo Horizonte, MG, Brazil
<sup>b</sup>Departamento de Física e Biofísica, Instituto de Biociências, Universidade Estadual Paulista (UNESP), Botucatu, SP, Brazil
<sup>c</sup>Departamento de Genética, Instituto de Biociências, Universidade Estadual Paulista (UNESP), Botucatu, SP, Brazil

**A R T I C L E   I N F O**

**Keywords:**
- Phospholipase A<sub>2</sub> inhibitor
- Beta inhibitors
- Phospholipase A<sub>2</sub>
- Viperidae
- Snake venom gland
- Snake liver
- Pit viper

**A B S T R A C T**

Several snake species possess, in their circulating blood, endogenous PLA<sub>2</sub> inhibitors (sbPLIs) with the primary function of natural protection against toxic enzymes from homologous and heterologous venoms. Among the three structural classes of sbPLIs – named α, β, and γ – the β class (sbβPLIs) is the least known with only four identified sequences, so far. The last class of inhibitors encompass molecules with leucine rich repeats (LRRs) motifs containing repeating amino acid segments. In the present study, we identified and characterized putative sbβPLIs from the liver and venom glands of six Latin American pit vipers belonging to Bothrops and Crotalus genera. The inhibitor from Crotalus durissus terrificus snakes (CdtββPLI) was chosen as a reference for the construction of the first in silico structural model for this class of inhibitors, using molecular modeling and molecular dynamics simulations. Detailed analyses of the electrostatic surface of the CdtββPLI model and protein-protein docking with crotinin B from homologous venoms predict the interacting surface between these proteins.

1. Introduction

Endogenous phospholipase A<sub>2</sub> inhibitors are oligomeric glycoproteins secreted by the liver into the circulating blood of several snake species. The primary purpose of these inhibitors, known as sbPLIs (an acronym for snake blood phospholipase A<sub>2</sub> inhibitors), is self-protection against the eventual presence of toxic snake venom phospholipases A<sub>2</sub> (svPLA<sub>2</sub>s) in the bloodstream.

The identification of sbPLIs from different snake species led to their classification into three distinct classes – named α, β, or γ – according to the presence of structural domains previously described for mammalian proteins (Ohkura et al., 1997). Up to now, the great majority of sbPLIs belongs to α or γ classes (Campos et al., 2016 and references therein). SβPLIs hold a characteristic C-type lectin-like domain, while sbβPLIs have a particular cysteine pattern that forms three-finger motifs. On the other hand, sbβPLIs display tandem leucine-rich repeats (LRRs) that delineate a horseshoe shaped molecular structure. Such inhibitors have been solely purified from three Asian snake species, being the first of which from the viperid Gloydius brevicaudus (formerly Agkistrodon blomhoffii sinicus). Under its native form, Gloydius brevicaudus sbβPLI (GbββPLI) occurs as a trimer composed of heavily glycosylated subunits of 50 kDa each. Each monomer binds to one PLA<sub>2</sub> molecule and GbββPLI exclusively inhibits basic PLA<sub>2</sub> from viperid snake venoms (svPLA<sub>2</sub>s), showing no inhibitory effects over acidic or neutral svPLA<sub>2</sub>s (Okumura et al., 1998, 2002). Two other sbβPLIs have been described in the nonvenomous snakes Elaphe climacophora and Elaphe quadrivirgata: EcsββPLI and EqsbββPLI, respectively. Although these species lack the secretion of toxic svPLA<sub>2</sub>s, both inhibitors maintain the narrow specificity of GbββPLI against basic svPLA<sub>2</sub> (Okumura et al., 2002; Shirai et al., 2009). The sole evidence of sbβPLIs in Latin American snakes so far was the identification of transcripts in the venom glands of Lachesis muta muta: EcββPLI (Lima et al., 2011).

Considering the limited number of sbβPLIs described in literature, this study aims at identifying and describing novel sbβPLIs from liver and venom glands of Latin American pit vipers belonging to Bothrops and Crotalus genera. The sbβPLI from C. durissus terrificus (CdtββPLI)
was used to construct the first in silico structural model of a sbβPLI. Through the molecular docking of CdsbβPLI and crotoxin B, which is the major basic neurotoxin in the homologous venom, we predicted the amino acids that interact at the interface of the inhibitor-toxin complex.

2. Materials and methods

2.1. Snake tissue collection

Bothrops alternatus, B. jararaca, B. jararacussu, B. moojeni, B. neuwiedii and C. durissus terrificus were obtained from the Serpentarium of Fundação Ezequiel Dias (Belo Horizonte, MG, Brazil). The snakes were anesthetized according to the protocol approved by the Ethics Committee on Animal Use (CEUA/FUNED 022/2012). Liver and venom glands were collected in DEPC-treated tubes, quickly frozen in liquid nitrogen and stored at −80 °C until use.

2.2. RNA extraction and cDNA synthesis

Total RNA and cDNA synthesis was performed as previously described for α and γ sbPLIs from Latin American snake species (Estevão-Costa et al., 2008, 2016). Briefly, total RNA was extracted from approximately 120 mg of snake tissue (liver, venom gland or both) with Trizol® (Invitrogen, USA). RNA integrity was tested by electrophoresis on 1% agarose gel for agarose gel. For cDNA synthesis we used the SuperScript® III First-Strand Synthesis System (ThermoFisher Scientific). In the first step, cDNA was synthesized with oligo (dT)12-18. In the second step, PCR was performed with sense signal peptide (3′-ATGAAGTCTTCGGTG TCCATCTC5′) and antisense carboxyl terminus (3′-TTAGGAGGAGCAA ATTTGGGAT5′) oligonucleotides, based on the published nucleotide sequence of E. coli (Okumura et al., 2002). As an additional control of RNA integrity and PCR performance, the specific sbβPLI primers were replaced by sense and antisense beta-actin primers (Cat. G7590, Promega Co. USA), in the second step of the reaction. All the amplifications were performed in a TC-412 thermocycler (Techne, UK) under the following conditions: 5 min at 94 °C, 35 cycles of 1 min at 94 °C, 30 s at 55 °C and 30 s at 72 °C, followed by an extension period of 7 min at 72 °C. Fresh PCR products were cloned into pGEM-T vector (Promega, USA) according to the manufacturer’s instructions. Inserting-contained clones were isolated after confirmation by conventional PCR screening of transformed NM522 E. coli in the presence of T7/SP6 promoter primers. Aliquots of the amplification reactions were analyzed by electrophoresis on 1.0% agarose gel in TBE buffer, in the presence of ethidium bromide.

2.3. Primary and secondary structure analyses of sbβPLIs

DNA sequences were determined using the Big Dye Terminator Cycle Sequencing Kit on an automated ABI 3130 Genetic Analyzer (Thermo Fisher Scientific) and consensus sequences were obtained from a minimum of three complete reads in both directions. Primary sequence deductions, chemical protein properties calculations, secondary structure prediction and multiple sequence alignments using the ClustalW algorithm were performed with the MacVector 15.1.1 software (Mac Vector Inc., USA). N-glycosylation sites were predicted using NetNGlyc 1.0 software (http://www.cbs.dtu.dk) with default threshold (> 0.50).

2.4. Molecular modeling and dynamics simulations

The in silico molecular model of C. durissus terrificus sbβPLI (CdsbβPLI) was generated by threading and molecular dynamics simulation techniques (Saxena et al., 2009). Multiple alignments were performed with signal and align 2 scripts based on 4QXE_A (Score 130–139; E-value: 1e-18; Identity: 24%), 4BV4_R (Score: 128; E-value 1e-16; Identity: 20–23%), 1OZN_A (Score: 123–131; E-value: 1e-17; Identity: 24%) and 1P9A_G (Score: 120; E-value: 1e-16; Identity: 28%) templates. The alignments were manually adjusted, after incorporating the restriction of putative disulfide bonds. Fifty initial models of the selected sbβPLI primary sequence were generated by the Modeller v.9.14 software (Martí-Renom et al., 2000). The best initial model was selected based on QMEAN and Ramachandran plot values at QMEAN and RAMPAGE servers, respectively (Benkert et al., 2009; Lovell et al., 2003). Next, the initial model was submitted to molecular dynamic (MD) simulations using the GROMACS (Groming Machine for Chemical Simulation) v.4.5.3 software (Pronk et al., 2013) in the presence of explicit water molecules. The protonation states of charged groups were set according to pH 7.0. Counter ions were added to neutralize the system and GROMOS 96 53a6 force field (Oostenbrink et al., 2005) was chosen to perform MD simulations. The minimum distance between any atom of the protein and the box wall was 1.0 nm. Energy minimization (EM) using a steepest descent algorithm was performed to generate the starting configuration of the system. After this step, 200 ps of MD simulations with position restraints applied to the protein (PRMD) were performed in order to gently relax the system. Then, 100 ns of un restrained MD simulations were calculated to evaluate the stability of the structure. All MD simulations were carried out in a periodic truncated dodecahedron box under constant temperature (298 K) and pressure (1.0 bar), which were maintained by coupling to an isotropic pressure and external heat bath. Search for crystallographic data of sbβPLIs-homologous proteins was performed with the HHpred server (Söding, 2005). The final model was analyzed using Ramachandran plot (Lovell et al., 2003) and its overall quality was evaluated by Z- and QMEAN4 scores (Benkert et al., 2009; Wiederstein and Sippl, 2007). The solvent-accessible surface area (SASA) of the in silico model was calculated by Arealmol software (Saff and Kuijlaars, 1997). Electrostatic potential surfaces were generated by APBS (Adaptive Poisson-Boltzmann Solver) electrostatic calculations (Baker et al., 2001), available at Chimera v.1.10 (Peetersen et al., 2004), after the transformation from PDB to PQR file using the online server PDB2PQR (Dolinsky et al., 2004).

2.5. Protein-protein docking predictions

The interface interaction between the CdsbβPLI in silico model and the crystal structure of crotoxin B (CB) isoform CBc (PDB ID 2QOG) was computationally predicted using docking algorithms available at HADDOCK 2.2 (Van Zundert et al., 2016). The docking protocol consisted in three stages (Domínguez et al., 2003): (i) Randomization of orientations around its mass center and rigid body energy minimization (EM), in which each protein is allowed to rotate to minimize the intermolecular energy function. Then, both translation and rotation are allowed, and the proteins are docked by rigid body EM. Typically, 1000 complex conformations are calculated at this stage and the best 200 solutions in terms of intermolecular energies are subsequently refined. (ii) Three simulated annealing refinements: first, both proteins are considered as rigid bodies (1000 steps from 2000 to 50 K with 8 fs time steps); second, the side chains at the interface are allowed to move (4000 steps from 2000 to 50 K with 4 fs time steps); and third, the side chains and the backbone at the interface are allowed to move, allowing conformational rearrangements (1000 steps from 500 to 50 K with 2 fs time steps). (iii) A series of molecular dynamics simulations with explicit solvent, for a final refinement. The final docking solutions are clustered through the pairwise backbone root mean square deviation (RMSD) at the interface, and the cluster is defined as an ensemble that displays I-RMSD smaller than 1.0 Å. The resulting clusters ranked on HADDOCK score summarize the average interaction energies (electrostatic interaction energy; van der Waals interactions and restraints violation energy) and their average buried surface area. The interactions of protein-protein predicted surfaces were plotted using the DIMPLOT software (Laskowski and Swindells, 2011).
3. Results

3.1. Detection and characterization of putative sbβPLIs in Latin American pit vipers

The integrity of the starting RNAs was confirmed by the unique presence of two bands that corresponds to 18S and 28S bands of ribosomal RNA, by electrophoresis analysis (data not shown). RT-PCR products with the expected sizes for sbβPLIs (about 1000bp) confirmed the presence of sbβPLIs transcripts in snake tissues (liver, venom glands or both) of B. alternatus (Bal), B. jararaca (Bja), B. jararacussu (Bju), B. moojeni (Bmo), B. neuwiedii (Bne) and C. durissus terrificus (Cd). The species abbreviation is followed by an asterisk when the source tissue was the venom gland.

Multiple nucleotide (nt) alignments of the cDNA sequences were performed and the primary structures of putative sbβPLIs were deduced from the consensus nt sequences, according to snake species and tissue source (provided as Supplementary data). Although the RNA and cDNA profiles on gel indicated the presence of transcripts in B. moojeni liver and B. jararacussu venom glands, the nucleotide sequences showed incomplete reads and were not included in our study.

All sbβPLI precursors from Latin American pit vipers displayed a signal peptide comprising 23 amino-acid residues. The mature proteins comprised 309 or 308 amino-acid residues. This difference was due to a proline deletion observed at the 221th position in seven out of the eleven sequences available for comparison (Fig. 2). The average molecular mass was 34667.40 ± 359.90 kDa (mean ± S.D.). The iso-electric points (pIs) varied from 6.4 to 7.6, therefore ensuring slightly acidic to basic character to these molecules. Leucine and proline residues accounted for 18.6 ± 1.0% and 7.3 ± 0.5% of the total amino acid content, respectively. Nine cysteines were present in most sequences, except for Bne and Lmm*, in which Ser was replaced by a tenth cysteine.

3.2. Comparison between Latin American and Asian sbβPLIs

The deduced primary structures of sbβPLIs from Latin American pit vipers were compared with known sbβPLIs from Asian species (G. brevicaudus; Gb; E. clima cophora; Ec; E. quadrivirgata, subunits A and B: EqA and EqB, respectively) by calculating the identity (IS) and similarity (SS) scores. Both scores were higher for the sbβPLIs from colubrid species when compared to G. brevicaudus vipers. For the whole set of inhibitors, ISs varied from 61.9% (Lmm* vs Cdt) to 98.4% (EqB vs Ec), while SSs ranged from 75.8% (Bmo vs Cdt/Cdt*) to 99.4% (EqB vs Ec). A pair to pair comparison between ISs suggests three clusters: cluster 1 is formed by colubrid sbβPLIs, while cluster 2 encompasses sbβPLIs from vipers, except for the sbβPLIs from Lmm and Bne, which form a third cluster. A bar-coded representation of the primary structures, outlining cysteine, proline and leucine residues, expresses the consensus sequences in each cluster (Fig. 3).

3.3. Structural characterization of the sbβPLI from C. durissus terrificus (CdsbβPLI)

CdsbβPLI is composed of 308 amino acids, containing 54 leucines (17.53%), 22 prolines (7.14%) and nine cysteines (2.92%). The calculated pl (7.2) is close to neutrality. The in silico prediction of the secondary structure of CdsbβPLI points 27% of turns (T), 43% of sheets (S) and 17% (H) of helices, besides 13% of H/T, H/S and S/T segments. N102 (N102ASS105) and N209 (N209SSL212) were predicted as N-linked glycosylation sites, with respective scores of 0.55 and 0.59. These segments are highly conserved in clusters 1 and 2 of Latin American sbβPLIs (Fig. 3).

In order to gain insights about the tertiary structure of sbβPLIs and their interaction with basic phospholipases A2, we generated an in silico model for CdsbβPLI. We searched for crystal structures of CdsbβPLI-homologous proteins due to the absence of crystallographic data for sbβPLIs. The output structures were the LRR-containing G-protein coupled receptor 4 from Epitreatus burgert (PDB ID 4QXE), the Toll ecotodomain from Dro sophila melanogaster (PDB ID 48V4) and the reticulon 4 receptor from Homo sapiens (PDB ID 1OZM). All these crystal structures presented e-values of −20 and IS of about 29% with CdsbβPLI and were used for the generation of our template by molecular modeling and molecular dynamics simulations (Fig. 4). According to the Ramachandran plot, 93.8% of the residues are distributed in favored and allowed regions. Z- and QMEAN4 scores of −4.10 and −8.28, respectively, indicate an overall good quality for the proposed model.

The search for consensus sequences of LRR motifs in the primary structure of CdsbβPLI revealed eight typical motifs with full conservation of the ionic character of the Highly Conserved Segment (HCS) and high conservation of the Variable Segment (VS) (Table 2; Fig. 4). In five of the eight conserved HCS, the last leucine is replaced by amino acids with the same chemical character. The in silico model of CdsbβPLI is in accordance to a homeshoe-shaped molecule, with the HCS regions located at its concave face, displaying the typical three-residue parallel β-strands of LRR motifs. A ninth LRR motif, flanked by a type I C-terminal cysteine-rich cap, starts at the 248th residue and displays a four-residue cysteine cluster with Cys260/Cys286 and Cys262/Cys306, forming disulfide bridges, besides a free thiol group Cys308 (Fig. 4). Two other cysteines (Cys147 and Cys196) are located at the convex face of the protein and may not be involved in disulfide bridges, since this would provide an unfavorable geometry for the LRR conformation. The LRR motifs and their flanking regions are conserved in every sbβPLI (Figs. 2 and 5). Remarkably, the N-terminal portion of CdsbβPLI does not present the main characteristics of the consensus sequence of an N-terminal cysteine-rich flanking domain, despite the existence of a two-residue cysteine cluster with a Cys1/Cys18 disulfide bridge. An enrichment of eight proline residues in the N-terminal region is observed for all sbβPLIs (Figs. 2, 4 and 5). A cluster of acidic amino acids can be noted in the concave surface of the inhibitor, therefore ascribing a negative net charge to it. This charged surface is compatible with the previously reported specificity of sbβPLIs from Asian species to basic svPLA2 (Fig. 4).
3.4. Mapping the interface region between CdtβPLIs and basic svPLA2s

A large negatively-charged area mostly constituted by aspartic and glutamic acid residues was observed at the concave surface of the CdtβPLI in silico model in the N-terminal region, as well as in the six initial LRRs motifs (Fig. 6). This area also contains several serine residues, which are usually involved in hydrogen bonds as proton donors or acceptors. The residues that form this negatively-charged area in the in silico model display a large solvent-accessible surface area (SASA), exposing an acidic area of $\sim 730 \text{ Å}^2$. This region is in the neighborhood of the above-mentioned proline-rich region at the N-terminal portion of CdtβPLI, probably offering structural stability for the binding of basic svPLA2. Remarkably, 10–50 N-terminal region presented the lowest root mean square fluctuation (RMSF) values of the overall structure during the molecular dynamic simulations (Fig. 4).

Comparative amino acid sequence analyses between CdtβPLI and other putative sbβPLIs from Latin American pit vipers showed that these acidic (aspartic and glutamic acids) and serine residues have great level of conservation in all sequences, except for a few residues (Fig. 2). Ser38, Glu40 and Ser64 are substituted by Tyr38, Gln40 and Thr64 residues in B. neuwiedi and L. muta muta sequences, respectively. However, these substitutions maintain the acidic character of the residue position. Other substitutions are Glu60/Gly60 and Glu184/Lys184 in B. neuwiedi and L. muta muta sequences, Ser112/Ile112 in B. alternatus venom gland sequence, and D136/Asn136 in B. moojeni venom gland sequence. Amino acid sequence comparisons between CdtβPLI and sbβPLIs from Asian snakes also revealed a great level of conservation of the residues present in this negatively-charged area, with only two substitutions: Ser110/Pro110 in GbβPLI and Ser154/Pro154 in non-venomous snake sequences (Fig. 5). Therefore, this negatively-charged area is present in all sbβPLIs that have been identified so far and possibly comprises the binding site region for positively-charged basic svPLA2s.

Unfortunately, no native CdtβPLI was available for experimental...
tests. Nevertheless, based on the expected selectivity of sbβPLIs for basic svPLA2, it is possible to suggest that CdsbβPLI may target crotoxin B (CB), the major basic PLA2 in C. d. terrificus venom. Analyses of the electrostatic surface in the crystal structure of the monomeric CB (PDB ID 2QOG) showed that positively charged areas, composed of lysine and arginine residues, mainly concentrate in the β-wing (residues 75–84), N- and C-terminal regions (residues 1–20 and 115–133, respectively), besides Arg36, Arg38 and Arg43 residues (Fig. 6).

In order to gain insights on the interface region between sbβPLIs and basic svPLA2s, we performed docking predictions between CdsbβPLI in silico model and the crystal structure of crotoxin B (CB) CBc isoform (PDB ID 2QOG) to obtain an in silico structural model for the CdsbβPLI/CB complex (Fig. 7). For the in silico model of CdsbβPLI, the previously identified negatively charged area formed by Glu15, Ser38, Glu40, Glu60, Ser64, Asp86, Thr84, Ser88, Ser110, Ser112, Glu113, Asp136, Ser156, Asp158, Ser160 and Glu184 residues was chosen as an active region to drive the docking (directly involved in the interaction) (Fig. 6). Regarding CB crystal structure, two different positively

Table 1

| Family      | Species or subspecies | Source tissue | GenBank             | Reference                                |
|-------------|-----------------------|---------------|---------------------|------------------------------------------|
| Colubridae  | Elaphe climacophora  | Liver         | AB462511            | Shirai et al. (2009)                     |
|             | Elaphe quadrivirgata | Liver         | AB060637 AB060638   | Okumura et al. (2002)                    |
| Viperidae   | Bothrops alternatus   | Liver         | MH479016            | Present study                            |
|             | Bothrops jararaca    | Venom gland   | MH479017            | Present study                            |
|             | Bothrops jararacussu | Venom gland   | MH479018            | Present study                            |
|             | Bothrops moojeni     | Venom gland   | MH479021            | Present study                            |
|             | Bothrops neuwiedi    | Venom gland   | MH479022            | Present study                            |
|             | Crotalus durissus terrificus | Venom gland | MH479024            | Present study                            |
|             | Gloydius brevicaudus | Liver         | AB007198            | Ohkura et al. (1997); Okumura et al. (1998) |
|             | Lachesis muta muta   | Venom gland   | –                   | Lima et al. (2011)                       |
charged areas were chosen as active regions for dock driving comparison. Region 1 is formed by Lys7, Lys10, Arg14, Lys16, Arg74, Arg78 and Lys87 residues, located at the N-terminal and β-wing regions from CB; while region 2 is formed by Arg36, Lys38, Arg43 and Lys114, Lys125 and Arg127, located at the C-terminal portion of CB (Fig. 6). The passive residues (surrounding surface residues) are automatically defined around the active residues. Comparative analyses between the statistical values for the best cluster among the solutions found for docking predictions of regions 1 and 2 of CB with CdtβPLI in silico model showed that region 1 has the best HADDOCK score and the best free energy values for binding (Table 3). These data suggest that the N-terminal and β-wing regions from CB may constitute the region for interaction with CdtβPLI. A protein/protein interface analysis of the structural model of CdtβPLI/CB complex (Fig. 7) shows that this interface contains fourteen hydrogen bonds. Five residues from the β-wing, a neighboring residue (Lys87) and four residues from the N-terminal portion of CB establish nine, one and four hydrogen bonds, respectively, with serine and aspartic/glutamic acid residues located at the negatively charged area from CdtβPLI in LRR motif 1 (one residue/one hydrogen bond), LRR3 (one residue/one hydrogen bond), LRR4 (one residue/one hydrogen bond), LRR5 (two residues/three hydrogen bonds), LRR6 (two of these residues and the exception Arg180/six hydrogen bonds) and even in LRR8 (one residue/one hydrogen bond), which is located outside the negatively charged area. Besides hydrogen bonding, hydrophobic contacts are also established in the CdtβPLI/CB interface by three residues from the β-wing, two from its vicinities, six from the N-terminal region and three (Phe24, Trp31 and Trp70) from other regions of CB crystal structure; and two residues from LRR1, four from LRR2, two from LRR3, two from LRR4, three from LRR5, one from LRR6 and two from LRR7 of CdtβPLI in silico model.

Table 2

| LRR motif | Sequence | No. of matching residues of HCS region with consensus sequence | No. of matching residues of VS region with consensus sequence |
|-----------|----------|---------------------------------------------------------------|-------------------------------------------------------------|
| Consensus sequence | LxxLxLxLxLnxLxxLxLPxxoFxzLxx | – | – |
| 1 | 58LQELHLSNNRL68 | 5/5 | 5/5 |
| 2 | 82LHTLDLSRNFL92 | 5/5 | 4/5 |
| 3 | 106LTHLSLSENQL116 | 5/5 | 4/5 |
| 4 | 130LRILGLDHNQV140 | 5/5 | 2/5 |
| 5 | 154LTSLDLSFNLI164 | 5/5 | 2/5 |
| 6 | 178LERLVLESNPI188 | 5/5 | 2/5 |
| 7 | 202LSVLSLNSNS125 | 5/5 | 12/15* |
| 8 | 225ELLLDLDNFE225 | 2/5 | 12/15* |
| 9 | 248NFLDLDSGNWACDRCLEELRLW(13)FVC(19)CPC308 | 2| |

Type 1 Cysteine-rich flanking consensus sequence

| LxxLxLxLxLnxLxxLxLPxxoFxzLxx |

Fig. 5. Amino acid sequences alignment of CdtβPLI and known βPLIs from Asian species. Abbreviations: G. brevicaudus (GbβPLI), E. climacophora (EcβPLI), and E. quadrivirgata (EqβPLIA and B) species. The highly conserved segments (HCS) of the eight leucine-rich repeat (LRR) motifs from the sequences are highlighted in red boxes. The highly conserved core LxxLxLxLnxLxxLxLPxxoFxzLxx of HCS from typical LRR motifs (1–8 in yellow) is indicated below the yellow boxes, where L is Leu, Ile, Val, or Phe; N is Asn, Thr, Ser, or Cys; C is Cys, Ser, or Asn, and x is any residue. The C-terminal LRR motif flanked by a cysteine-rich cap is also highlighted in a yellow box. The cysteines involved in disulfide bridges in the C-terminal (Cys260/Cys286 and Cys262/Cys306) and N-terminal (Cys4/Cys18) are highlighted in green boxes and by C symbols. Proline residues in the N-terminal region are highlighted in purple boxes and by P symbols, evidencing the proline-enrichment of this region. The acidic amino acids (Glu and Asp) and serine residues that form the negatively charged area located at the N-terminal region and LRR motifs 1–6 (see Fig. 6) in the concave face of the protein are highlighted in red boxes and by E, D and S symbols.
GbßPLI was the first prototype for the structural model of blood plasma β class inhibitors from snakes, in spite of the absence of any structural model of its tertiary or quaternary structures. The native trimer (160 kDa) is composed of homogenous subunits of 50 kDa, as estimated by SDS-PAGE (Ohkura et al., 1997). It was reported to be heavily glycosylated with N-linked, since the estimated molecular mass of the monomer dropped from 50 to 39 kDa after enzymatic deglycosylation (Okumura et al., 1998). Two (N102 and N209) of the four potential N-linked glycosylations proposed for GbßPLI are present in CdtßPLI, therefore confirming the glycosylated nature of sbßPLIs. The average molecular mass of 34.7 kDa calculated for the sbßPLIs monomers described herein is compatible with that of deglycosylated GbßPLI monomer (34.6 kDa). Analysis of GbßPLI primary structure showed that this protein had LRR motifs homologous to leucine-rich α2-glycoprotein (Okumura et al., 1998).

The in silico model of CdtßPLI reveals eight typical LRR motifs in a horseshoe-shaped structure, and a ninth LRR motif flanked by a type I C-terminal cysteine-rich cap (Figs. 4 and 5). Two cysteines (Cys147 and Cys190) located at the convex face of the protein probably occur as free thiols, in agreement with the GbßPLI description, since an intra-chain disulfide bond would cause an unfavorable geometry for the LRR conformation. Basic PLA2 molecules would supposedly bind to the concave face of the GbßPLI due to negatively charged residues at LRR1 (Okumura et al., 1998). Moreover, the LRR1 region might be responsible for the specific binding to G. brevicaudus basic PLA2s, since it is fully conserved among three sbßPLIs from Asian species (Shirai et al., 2009). Curiously, GbßPLI gene expression in the liver was shown to be upregulated by acidic svPLA2 and not by basic svPLA2, as could be expected based on the selectivity of the inhibitor to the ionic character of svPLA2 (Kinkawa et al., 2010).

The in silico model of CdtßPLI highlights the relevance of a negatively charged area in its concave surface to the specificity of sbßPLIs to basic svPLA2s. Furthermore, the structural model shows that this negatively charged area is not restricted to LRR1 as previously suggested, but it spreads to the N-terminal region and LRR motifs 1 to 6. The clustering of negative amino acid residues results in the exposition to solvent of an acidic-charged area of ∼730 Å². The availability of the concave surface of the inhibitor to svPLA2 binding is in accordance with the most typical profile of LRR-containing proteins. This negatively charged area in the CdtßPLI is in the neighborhood of the C-terminal cysteine-rich cap (Cys147 and Cys190) located at the convex face of the protein probably occur as free thiols, in agreement with the GbßPLI description, since an intra-chain disulfide bond would cause an unfavorable geometry for the LRR conformation. Basic PLA2 molecules would supposedly bind to the concave face of the GbßPLI due to negatively charged residues at LRR1 (Okumura et al., 1998). Moreover, the LRR1 region might be responsible for the specific binding to G. brevicaudus basic PLA2s, since it is fully conserved among three sbßPLIs from Asian species (Shirai et al., 2009). Curiously, GbßPLI gene expression in the liver was shown to be upregulated by acidic svPLA2 and not by basic svPLA2, as could be expected based on the selectivity of the inhibitor to the ionic character of svPLA2 (Kinkawa et al., 2010).
a proline-rich region at the N-terminal portion of the molecule, which can confer structural stability to the binding of basic svPLA₂ (Fig. 3). Remarkably, this area in our model is larger than the positively charged area pointed as the acidic svPLA₂ binding region at the monomer of Bothrops alternatus venom. CB has an isoelectric point of 8.4 and displays widely spread positively charged areas in its structure (Fernandes et al., 2017; Marchi-Salvador et al., 2008) mainly concentrated in the N-terminal region and LRR motifs 1 to 6 (Fig. 6). Both N- and C-terminal regions have been pointed out to be involved in the neurotoxic effect induced by CB (Carin-Serbec et al., 1994; Fernandes et al., 2017; Fortes-Dias et al., 2009). Docking predictions between CdsbβPLI in silico model and CB crystal structure showed that, differently from the previous studies with GbsbβPLI that suggested LRR1 as the central point for the binding of basic svPLA₂, the negatively charged area located at the N-terminal region and LRR motifs 1 to 6 of CdsbβPLI establishes hydrogen bonds and hydrophobic contacts with CB, especially LRR motifs 2 (4 hydrogen bonds), 5 (3 hydrogen bonds and 3 residues involved in hydrophobic interactions) and 6 (1 hydrogen bond and 6 residues in hydrophobic interactions) (Fig. 7). It is important to point out that this binding region is similar to the binding site of MD-2 at the concave face of the ectodomain of Toll-like Receptor 4 (TLR4), at the junction of the N-terminal and central LRR motifs (Kim et al., 2007). Regarding CB, N-terminal and β-wing regions may constitute the region of interaction with CdsbβPLI, displaying 14 hydrogen bonds with serine and asparagine/glutamic acid residues located at the negatively charged area of CdsbβPLI, and 14 residues involved in hydrophobic interactions (Fig. 7).

The identification of sbPLIs from three different structural classes in non-venomous snakes raises questions about other physiological roles of these proteins beyond a natural resistance against the venom. In the particular case of sbβPLIs, Okumura et al. (1998) raised the hypothesis that GsbβPLI could correspond to snake LRG and that mammalian LRG might function as an svPLA₂ inhibitor. Although the last assumption has not been proved, GsbβPLI was shown to bind cytochrome c (Cyt c), an endogenous ligand of LRG, with higher affinity ($K_d = 2.37 \times 10^{-12} \text{M}$) than to basic PLA₂ from G. brevicaudus venom ($K_d = 1.21 \times 10^{-9} \text{M}$) (Shirai et al., 2010). In fact, the tight binding of GsbβPLI to Cyt c led the authors to suggest that the original function of sbβPLIs might be the clearance of autologous Cyt c from the circulatory system of snakes. They also proposed that the sbβPLIs may have acquired a new function as PLA₂ inhibitory proteins in addition to the ancestral Cyt c binding, during snake evolution. This possibility indicates a great biotechnological potential of sbβPLIs, beyond the action of protecting against toxic svPLA₂, since Cyt c is involved in apoptosis induction and proinflammatory mediation (Hiraoaka et al., 2004; Pullerits et al., 2005). Cyt c has a positively-charged lysine-rich region on its surface involved in electrostatic interactions with subunit II of cytochrome c oxidase (Döpner et al., 1999). Thus, Cyt c may bind to sbβPLIs in the same way as it binds to basic svPLA₂, by electrostatic interactions with the negatively-charged area. The structure identified in the CdsbβPLI in silico model has high level of conservation among known sbβPLIs (Figs. 2, 5 and 6).

CdsbβPLI, as well as other previously described sbβPLIs, belong to the ‘typical’ subfamily of the structural class III of repetitive proteins, which contain an invariant eleven-residue core LxxLxLxxCxxL plus a variable segment (Matsumiwa et al., 2007). The whole repetitive segment is formed by the 24-25-residue sequence LxxLxLxxNxLxxLxxOFxZxLxx, where, in the case of type C-terminal cysteine region identified in CdsbβPLI, L is Leu, Ile, Val, or Phe; N is Asn, Thr, Ser; C is Cys, P is Pro but can be substituted by Trp or Phe residues in some sequences (Kobe and Kajava, 2001; Ng et al., 2011). SbβPLIs belong to type α/β solenoid structures, which have been associated to the PFAM clan CI022 of LRRs (Paladin and Tosatto, 2015). Since LRRs occur in a large number of proteins in plants, invertebrates and vertebrates, and participate in several biologically important processes (Ng et al., 2011), we believe that the present study contributes, not only to the study of sbβPLIs, but to the general knowledge on the widely distributed LRR-containing proteins.
Table 3

| Protein and parameters        | Cdtβ PB   | Cdtβ PB in silico model | Cdtβ PB crystal structure |
|-------------------------------|-----------|-------------------------|---------------------------|
| RMSD from the overall lowest-energy structure (Å) | 0.9 ± 0.6 | 0.9 ± 0.6 | 0.9 ± 0.6 |
| Van der Waals energy (kcal/mol) | 49.1 ± 7.7 | 49.1 ± 7.7 | 49.1 ± 7.7 |
| Electrostatic energy (kcal/mol) | -48.7 ± 6.0 | -48.7 ± 6.0 | -48.7 ± 6.0 |
| Desolvation energy (kcal/mol) | 60.1 ± 8.0 | 60.1 ± 8.0 | 60.1 ± 8.0 |
| Restraints violation energy (kcal/mol) | 38.2 ± 4.5 | 38.2 ± 4.5 | 38.2 ± 4.5 |
| Z-score (a.u.) | -10.9 ± 7.9 | -10.9 ± 7.9 | -10.9 ± 7.9 |
| HADDOCK score | 2387.5 ± 107.7 | 2387.5 ± 107.7 | 2387.5 ± 107.7 |
| Cluster size | 18 | 18 | 18 |
| RMSD to the overall lowest-energy structure (Å) | 2.3 ± 0.4 | 2.3 ± 0.4 | 2.3 ± 0.4 |
| HAMDOCK score (a.u.) | -215.7 ± 7.7 | -215.7 ± 7.7 | -215.7 ± 7.7 |
| Buried Surface Area (Å²) | 2857.3 ± 65.9 | 2857.3 ± 65.9 | 2857.3 ± 65.9 |
| Restraints violation energy (kcal/mol) | 57.7 ± 34.21 | 57.7 ± 34.21 | 57.7 ± 34.21 |
| Z-score (a.u.) | -1.6 | -1.6 | -1.6 |
| HADDOCK score | 2637.3 ± 60.1 | 2637.3 ± 60.1 | 2637.3 ± 60.1 |
| Cluster size | 21 | 21 | 21 |
| RMSD to the overall lowest-energy structure (Å) | 6.2 ± 0.3 | 6.2 ± 0.3 | 6.2 ± 0.3 |
| HAMDOCK score (a.u.) | -188.4 ± 4.0 | -188.4 ± 4.0 | -188.4 ± 4.0 |
| Buried Surface Area (Å²) | 521.1 ± 52.8 | 521.1 ± 52.8 | 521.1 ± 52.8 |
| Restraints violation energy (kcal/mol) | -38.0 ± 6.2 | -38.0 ± 6.2 | -38.0 ± 6.2 |
| Z-score (a.u.) | -1.4 | -1.4 | -1.4 |
| HADDOCK score | 2067.5 ± 107.7 | 2067.5 ± 107.7 | 2067.5 ± 107.7 |
| Cluster size | 26 | 26 | 26 |
| RMSD to the overall lowest-energy structure (Å) | 7.9 ± 0.3 | 7.9 ± 0.3 | 7.9 ± 0.3 |
| HAMDOCK score (a.u.) | -79.2 ± 4.1 | -79.2 ± 4.1 | -79.2 ± 4.1 |
| Buried Surface Area (Å²) | 38.0 ± 6.2 | 38.0 ± 6.2 | 38.0 ± 6.2 |
| Restraints violation energy (kcal/mol) | 57.7 ± 38.21 | 57.7 ± 38.21 | 57.7 ± 38.21 |
| Z-score (a.u.) | -1.8 | -1.8 | -1.8 |
| HADDOCK score | 2387.5 ± 107.7 | 2387.5 ± 107.7 | 2387.5 ± 107.7 |
| Cluster size | 20 | 20 | 20 |
| RMSD to the overall lowest-energy structure (Å) | 7.9 ± 0.3 | 7.9 ± 0.3 | 7.9 ± 0.3 |

5. Conclusion remarks

In the last years, efforts have been made to determine the structure of endogenous phospholipase inhibitors and their binding mode with phospholipases A₂ due to the potentiality of these endogenous molecules to be used as complements of conventional serum therapy and/or as inhibitors of secretory PLA₂ in pathological processes in humans. Regarding sbβPLIs, few of these molecules have been identified so far and the reasons for their specific inhibition of basic svPLA₂ are poorly understood. In the present study, we identified precursors of novel sbβPLIs in six species of Latin American pit vipers, and an extensive structural analysis of the transcribed proteins led to the construction of the first structural model of a sbβPLI, using the inhibitor from C. d. terriceps (CdtβPB) as a prototype. We investigated the electrostatic surface of CdtβPB in silico model and identified a conserved negatively charged area located at the N-terminal region and LRR motifs 1–6 (in the neighborhood of a proline-rich region). Docking predictions between CdtβPB in silico model and CB crystal structure highlighted the role of this area in the binding interface between these two molecules, where LRR motifs 2, 5 and 6 from CdtβPB establish the major number of contacts with CB. We also emphasized the positively charged areas at the surface of the basic svPLA₂ from C. d. terriceps venom crotoxin B (CB), and pointed that the N-terminal and β-wing portions of this toxin may constitute the interacting interface with CdtβPB. Our data contribute to a better understanding of sbβPLIs as well as of LRR-containing proteins in general.

Conflict of interest

On behalf of the authors of the manuscript entitled “Identification, description and structural analysis of beta phospholipase A₂ inhibitors (sbβPLIs) from Latin American pit vipers indicate a binding site region for basic phospholipases A₂ from snake venoms”, I declare no conflicts of interest regarding the manuscript.

Financial support

We thank the Brazilian funding agencies: Fundação de Amparo à Pesquisa de Minas Gerais (FAPEMIG), Fundação de Amparo à Pesquisa do Estado de São Paulo (FAPESP), Conselho Nacional de Desenvolvimento Científico e Tecnológico (CNPq), Coordenação de Aperfeiçoamento de Pessoal de Ensino Superior (CAPES Toxicology Program), Coordenação de Aperfeiçoamento de Pessoal de Ensino Superior, Fundação de Amparo à Pesquisa do Estado de São Paulo (FAPESP), Conselho Nacional de Desenvolvimento Científico e Tecnológico (CNPq), Coordenação de Aperfeiçoamento de Pessoal de Ensino Superior (CAPES Toxicology Program), Coordenação de Aperfeiçoamento de Pessoal de Ensino Superior, and Instituto Nacional de Ciência e Tecnologia em Toxinas (INCTTox/CNPq) for financial support.

Acknowledgements

We acknowledge the contribution of MI Estevão-Costa and RM Lima in sequence data acquisition.

Appendix A. Supplementary data

Supplementary data to this article can be found online at https://doi.org/10.1016/j.toxcx.2019.100009.

References

Baker, N.A., Sept, D., Joseph, S., Holst, M.J., McCammon, J.A., 2001. Electrostatics of nanosystems: Application to microtubules and the ribosome. Proc. Natl. Acad. Sci. 98, 10037–10041. https://doi.org/10.1073/pnas.181342398.
Benkert, P., Künzli, M., Schwede, T., 2009. QMEAN server for protein model quality estimation. Nucleic Acids Res. 37. https://doi.org/10.1093/nar/gkp322.
Campos, P.C., de Melo, L.A., Dias, G.L.F., Fortes-Dias, C.L., 2016. Endogenous phospholipase A₂ inhibitors in snakes: a brief overview. J. Venom. Anim. Toxins Incl. Trop. Dis. 22, 37. https://doi.org/10.1186/s40409-016-0092-5.
Čurin-Šerbek, V., Délot, F., Faure, G., Saliou, B., Gubeneck, F., Bon, C., Choumet, V., 1994. Antipeptide antibodies directed to the C-terminal part of ammodytoxin A react with
