

# In Silico Analysis of Short Linear Motifs Present in Snake Venom Phospholipases

Caterina Peggion <sup>1</sup>, and Fiorella Tonello <sup>2,\*</sup>

<sup>1</sup> Department of Biomedical Sciences, University of Padua, Padua, Italy; caterina.peggion@unipd.it

<sup>2</sup> Institute of Neuroscience of the National Research Council, Padua, Italy

\* Correspondence: fiorella.tonello@cnr.it

**Abstract:** Phospholipases A2 (PLA2s) are important constituents of snake venom that, depending on their amino acid composition, possess several toxic properties whose main ones are neurotoxicity, myotoxicity and impairing of haemostasis. They are proteins of about 120 amino acids, having a structure conserved since basal metazoa, and very similar to that of mammalian secretory PLA2s. Some snake venom PLA2s are heteromultimers, others are monomers or homodimers. In this work we analyze the sequence alignment of monomeric or homodimeric snake venom PLA2s, grouped according to their myotoxic and neurotoxic properties, and we compare this alignment with that of the most similar mammalian secretory PLA2s. We found short linear motifs, present in three regions of secretory PLA2s that can have a role in their toxic and physiological functions. This work suggests important molecular interactions of secretory PLA2s that can focus and shorten the experimental work of characterization of the mechanism of action of these proteins.

**Keywords:** snake venom phospholipases A2; neurotoxins; myotoxins

## 1. Introduction

Secretory PLA2s (sPLA2s) (EC 3.1.1.4) are evolutionary well-conserved lipolytic enzymes involved in a variety of processes from signal transduction, lipid mediator production and dietary lipids digestion in mammals to snakebite envenomation, since sPLA2s represent the major component of snake venom [1–3].

More than 400 different snake venom PLA2s toxins sequences are reported in the UniProtKB database, with an overall conserved tertiary structure [4]. Given this highly conservation, all of these sPLA2s toxins represent optimal models for the study of sPLA2s function and mechanism of action. Based to the homology to mammalian sPLA2s, snake venom toxins are subdivided in two groups. Toxins derived from Elapidae and Hydrophiidae snakes belong to group I and are more similar to the mammalian pancreatic phospholipases (PLA2G1B), whereas toxins derived from Viperidae and Crotalidae belong to group II, given their homology with the mammalian phospholipase enriched in synovial fluids and tears (PLA2G2A) [5]. sPLA2s of both groups share a highly conserved amino acid sequence and a well-conserved tertiary structure, and although the majority of snake venom PLA2s exist as monomer, some of them acquired a quaternary structure by a non-covalent interaction between two or more PLA2s [1,5–7]. PLA2s snake toxins are also subdivided in three different categories based on their toxic effects: myotoxins, targeting skeletal muscles, neurotoxins, that act on pre- or post- synaptic element of neuromuscular junction and hemostasis-impairing toxins [5,8]. Unfortunately, a correlation between toxic effects of different snake venom PLA2s and their primary structure is still unknown. Moreover, a lot of snake venom PLA2s with high myo/neuro activity are catalytically inactive due to a missense mutation of D49 amino acid (more frequently substituted with a Lysine) [5,9]. Therefore, the emerging idea is that the sPLA2s toxic activities depend not only by their enzymatic functions, but also by their interaction with target

**Citation:** Peggion, C.; Lastname, F.; Tonello, F. In Silico Analysis of Short Linear Motifs Present in Snake Venom Phospholipases. *Proceedings* **2021**, *65*, x. <https://doi.org/10.3390/xxxxx>

Received: 30 October 2020

Accepted: 4 January 2021

Published: 14 January 2021

**Publisher's Note:** MDPI stays neutral with regard to jurisdictional claims in published maps and institutional affiliations.



**Copyright:** © 2021 by the authors. Submitted for possible open access publication under the terms and conditions of the Creative Commons Attribution (CC BY) license (<http://creativecommons.org/licenses/by/4.0/>).

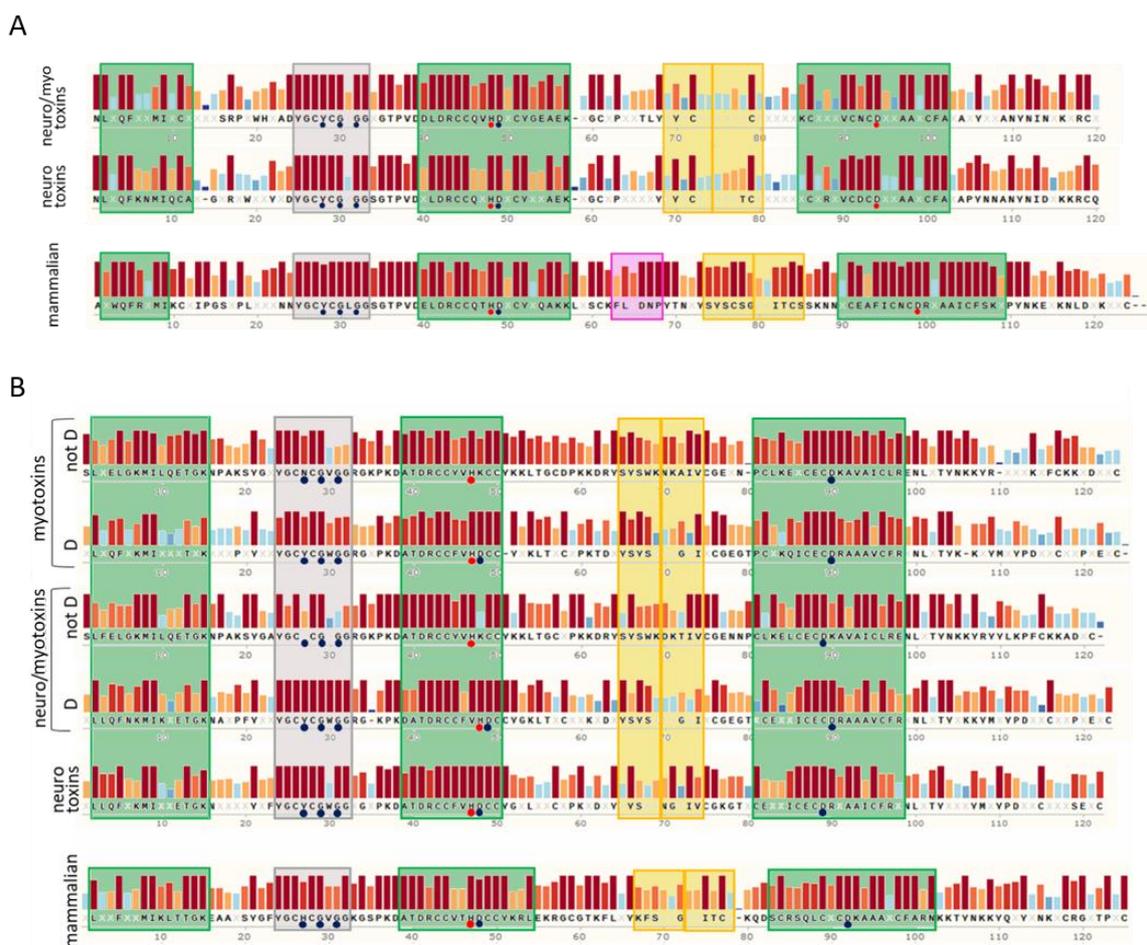
proteins that may mediate their entry into the cells, activating different signaling cascades, and their intracellular activity.

Here we present an *in silico* analysis of the primary and 3D structures of snake toxins PLA2s, comparing with their mammalian counterparts, aimed to unveil exposed sites that may represent potential candidates for modification or interaction with other proteins. This analysis was made possible thanks to the “animal toxin annotation project” of the Swiss-Prot database, in which venom protein sequences are systematically curated to the standards of UniProtKB/SwissProt [10]. Thanks also to the Eukaryotic Linear Motif (ELM) resource [11,12] it was possible to predict not yet discovered post-translational modifiable sites (i.e, by phosphorylation), or of interaction with other proteins in sPLA2s, helping in the elucidation of the mechanism of action of snake toxins and mammalian sPLA2s.

## 2. Results

### 2.1. Sequence Alignment Comparison

To unveil similarities or critical differences between snake PLA2s toxins and mammalian sPLA2s belonging to group I and II, we have aligned their sequences with Clustal Omega, represented the alignments using the software Snapgene, and compared the alignments in figure 1 A and B.



**Figure 1.** Consensus sequences analyses of snake venom PLA2s and their mammalian homologs. Panel A- Consensus sequences of Group I snake venom neuro-myotoxic and neurotoxic PLA2s compared with that of mammalian PLA2G1B. Panel B- Consensus sequences of Group II snake venom myotoxic, neuro-myotoxic and neurotoxic PLA2s compared with that of mammalian PLA2G2A. Myotoxins and neuro-myotoxins are divided according to the amino acid present in the position known as ‘49’ of the active site. The sequences were collected from the Swiss-Prot database and aligned with the align tool of UniProt (Clustal O, <https://www.uniprot.org/align/>). The alignments were then visualized with SnapGene

Viewer version 4.3.11. Conserved residues are reported in the consensus sequence, whereas X code indicates any amino acid. The colored bars above the protein sequence show the degree of conservation where a red bar indicates the maximum level of conservation (100%). Green, yellow and grey squares represent  $\alpha$ -helix,  $\beta$ -sheet and calcium binding loop, respectively. The pancreatic loop (pink square, panel A) is present only in mammalian PLA2G1B. Blue circles indicate amino acids involved in the  $Ca^{2+}$  binding, whereas red circles represent the catalytic sites.

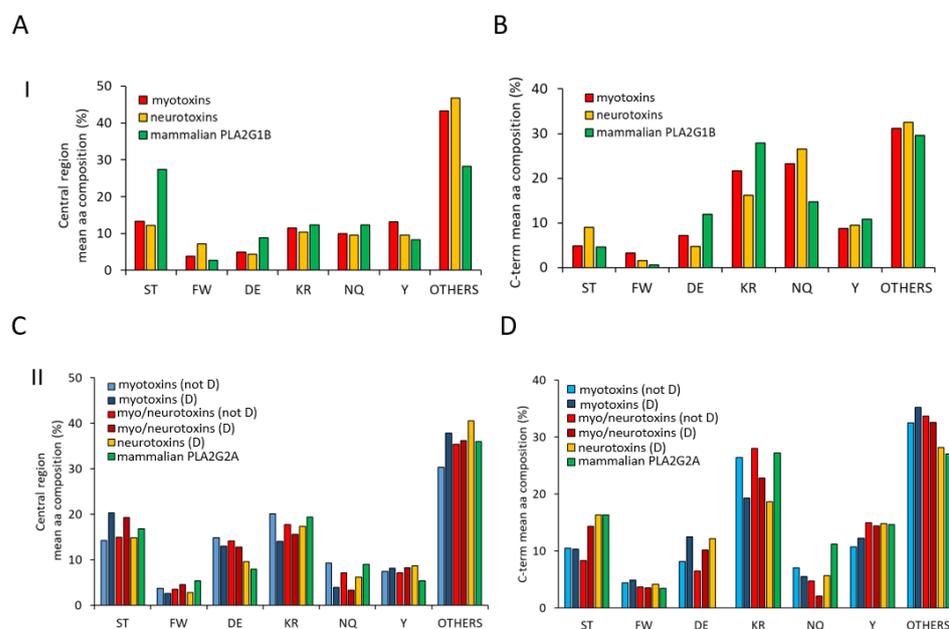
Secretory PLA2 appears to have three regions with higher variability: the region between the first  $\alpha$ -helix and the calcium binding loop, the region between the second and the third  $\alpha$ -helices and that comprises the  $\beta$ -sheets, and the C-terminal region.

### 2.2. Amino-acid Composition Analysis of the Central and C-terminal Regions

To unveil amino acids more abundant in the central and the C-terminal regions that may be important for the biological roles of sPLA2s, we have evaluated the mean abundance of different categories of amino acids, taking in consideration particular amino acidic properties.

As shown in figure 2A, the central region of sPLA2s (Group I) neuro-myotoxins and neurotoxins (red and yellow bars) is characterized by a lower mean abundance of phosphorylatable Ser/Thr amino acids and of negative charged amino acids (D/E), with respect to mammalian PLA2G1B (green bars). As regards the C-terminal region, it is possible to observe an increased amount of N/Q and a lower amount of charged amino acid (D/E and K/R) in toxins with respect to mammalian PLA2G1 (figure 2 B).

The central region of Group II sPLA2s show a pattern less distinctive for each family, with a mild increment of D/E and of phosphorylatable Y in toxins with respect to mammalian PLA2G2A (green bars, figure 2 C). Conversely, more differences could be notice in the C-terminal region of sPLA2s (Group II). In particular, it is evident the total absence of D/E in mammalian PLA2G2A and lower amount of N/Q in toxins and of S/T in myotoxins and neuro-myotoxins (not-D49) (figure 2 D).

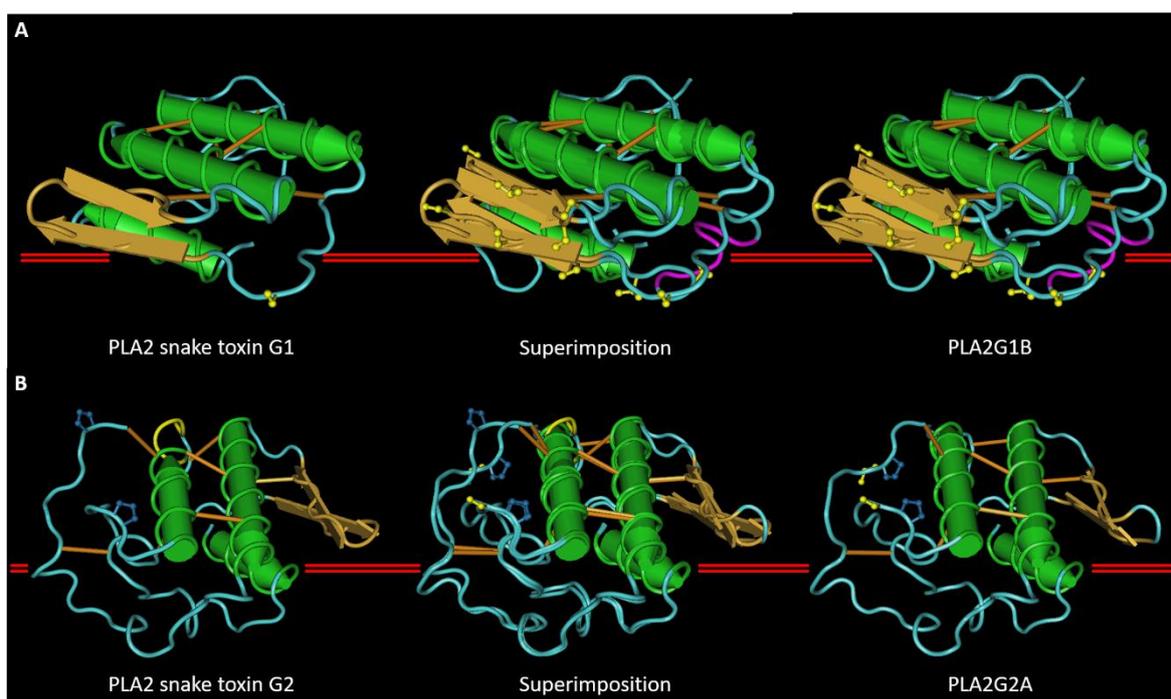


**Figure 2.** Mean amino acidic composition of the two low complexity regions of sPLA2s. Analyses of the mean amino acidic composition of the central regions (55–85) (panel A and C) and of the C-terminal stretches (from 103) (panel B and D) belonging to toxins of group I and mammalian PLA2G1B (panel A and B) and to toxins of group II and mammalian PLA2G2B (panel C and D), respectively. Evaluation of amino acidic composition was performed for each sequence using the ProtParam tool (<https://web.expasy.org/protparam/>) and then the mean value for each amino acid was calculated and reported in the bar diagram. Amino acids reported in the x axis were separated based on their chemical properties (S,T =

phosphorylable sites of ser/thr kinases; F,W: aromatic; D,E: negative charged at neutral pH; K,R: positive charged at neutral pH; N,Q: amine added amino acids; Y: aromatic phosphorylable sites of Tyr kinases; Others: A,G, V, L, I, H, P,C,M).

### 2.3. Slims Conserved in Toxins and not in Mammalian PLA2s or Vice Versa

We searched the ELM database [12] for SLiMs present and conserved in the different groups of sPLA2s. We included in the search the motifs present in any cellular organelles because PLA2s have been shown to be internalized in cells and localized in vesicles, in the cytoplasm, as well as in mitochondria and the nucleus [13–17]. Of all the motifs identified, we selected those conserved in toxins but not in mammalian PLA2s, or vice versa. Furthermore, as these motifs can be false positives, to decrease this possibility, we have only considered those present in more flexible and exposed areas of the proteins. The main motifs we identified are reported in figure 3.



**Figure 3.** Main SLiMs that differentiates snake venom PLA2s from their mammalian homologues. The plasma membrane is represented with a double red line. **A)** 3D structures of the snake venom neuro-myotoxin notexin, of human PLA2G1B and their superimposition, with in yellow the side chain of S/T residues in the central region. The S/T stretch present in mammalian PLA2G1B contains four superimposed motifs of phosphorylation by GSK3. **B)** 3D structures of the snake venom myotoxin bothropstoxin-I, of human PLA2G2A and their superimposition. The S/T residue (in yellow), when phosphorylated, allows isomerisation of the adjacent proline by Pin1. The loop following the second  $\alpha$ -helix in the toxin structure contains a PKA phosphorylation site and an SH2 binding site, not present in PLA2G2A.

### 3. Discussion and Conclusions

The higher differences in the amino acidic composition of the mammalian and snake venom sPLA2s concern the regions that in the 3D structure are on the sides of the proteins, perpendicular to the plasmatic membrane. These regions are enriched in amino acids typical of prion-like domains (N, Y, S, G), and can have or not charged amino acids that can influence the propensity of these proteins to form oligomers on the plasma membrane, and condensates with other proteins such as melittin, HSP70, nucleolin. Such condensates can modulate the enzymatic activity of sPLA2s or regulate the activity of other enzymes [13,18,19]

Finally, the main finding of this analysis are that sPLA2s may be subjected to post-translational modifications that could modulate their function/toxicity. In particular, PLA2G1B has several motifs, not conserved in group I toxins, that can be modified both

by GSK3, which generally has an inhibitory function [20,21], and by prolyl isomerase 1 (Pin1), an [ST] phosphorylation regulated prolyl isomerase. Structural modifications by prolyl isomerase Pin1 can remove sPLA2s from the complexes in which they participate and redirect it towards the degradation system as happen for many proteins [22,23]. Hence, the toxicity of myotoxins may be (partly) due to the lack of this control by Pin1 so that proline always remains in the *trans* conformation and this may explain why the removal of proline in bothropstoxin-I greatly reduces its myotoxic activity [24]. Moreover, Group II PLA2s snake toxins possess motifs that interact with SH2 domains and PDZ domains.

These interaction domains, and post-translational modification motifs, can have a role in the formation and breakdown of functional molecular condensates on the plasma membrane that can modulate the enzymatic activity of the sPLA2s, and carry out their non-enzymatic activity. Further studies will be necessary to unravelling all the activities triggered by sPLA2s and all its molecular partners.

## 4. Methods

### 4.1. Sequence Collection and Alignment

The sequences of the mono or homomeric snake venom phospholipases A2 were collected in the database UniProt using the string (family: "phospholipase a2 family Group I subfamily" fragment:no keyword: "Toxin [KW-0800]" keyword: "Myotoxin [KW-0959]" AND keyword: "Neurotoxin [KW-0528]" taxonomy: "Serpentes (snakes) [8570]" not (annotation:(type:subunit heterodimer) OR annotation:(type:subunit heterotrimer) OR annotation:(type:subunit heterohexamer)), substituting 'Group I' with 'Group II' to search phospholipases A2 of the second Group, and leaving or removing the keywords indicating the type of toxins (myo- or neuro-) to collect the proteins based on their site of action. Only curated entries were considered. To distinguish the catalytically active subfamily the string (family: "D49 sub-subfamily") was included with the conjunctions AND or not. The sequences of mammalian phospholipases were collected in the database UniProt by entering the gene names PLA2G1B and PLA2G2A, considering only the ten curated entry for PLA2G1B, and the five curated entry plus five other entries in the case of PLA2G2A. The sequence alignment was performed with Clustal Omega and visualized with the SnapGene viewer software (version 4.3.11). Pre- and pro-peptides, when present, were removed from the sequences.

### 4.2. Amino-acidic Composition Analysis of The B-Sheet Containing Region and C-terminal Stretch

The amino-acidic composition analysis of the central region (located between the second and third  $\alpha$ -helices and including the two  $\beta$ -sheet secondary structures) and of the C-terminal stretch (next to the third  $\alpha$ -helix) of PLA2s proteins was performed using the PROTParm tool of Expasy (<https://web.expasy.org/protparam/>). Amino-acids were Grouped considering their chemical and biochemical properties as follows: S,T; F,W; D,E; K,R; N,Q; Y and G,A,V,M,L,I,C,H (called as others). Thus, the frequency of each amino acidic category in the sequence analysed was obtained by dividing its abundance to the total length of the analysed sequence and expressing it as a percentage value. Finally, the mean value of the frequency of each amino acidic category was calculated for each family of PLA2A toxins and their mammalian counterparts.

### 4.3. Short Linear Motifs Identification

Short linear motifs (SLIMs) were identified using the 'The Eukaryotic Linear Motif resource for Functional Sites in Proteins' (<http://elm.eu.org/>) [12], considering all cell compartments. Only motifs conserved in toxins and not in mammalian phospholipases A2 or vice versa were considered. The syntax used to express amino acid motifs is that adopted by the ELN database and described at the page <http://elm.eu.org/infos/help.html>.

#### 4.4. Tri-dimensional Structure Representations

The tri-dimensional superimposition of the protein structures was obtained with the 'Vector Alignment Search Tool Plus' [25], pictures and animations were obtained with the 'Cn3D 4.3.1 3-D structure viewer'. The PDB files were AE7 (notexin), 3ELO (human PLA2G1B), 3I3I (bothropstoxin-I of *Bothrops jararacussu*) and 1BBC (PLA2G2A).

#### References

1. Dennis, E.A.; Cao, J.; Hsu, Y.-H.; Magrioti, V.; Kokotos, G. Phospholipase A2 enzymes: physical structure, biological function, disease implication, chemical inhibition, and therapeutic intervention. *Chem. Rev.* **2011**, *111*, 6130–6185.
2. Singer, A.G.; Ghomashchi, F.; Le Calvez, C.; Bollinger, J.; Bezzine, S.; Rouault, M.; Sadilek, M.; Nguyen, E.; Lazdunski, M.; Lambeau, G.; et al. Interfacial kinetic and binding properties of the complete set of human and mouse groups I, II, V, X, and XII secreted phospholipases A2. *J. Biol. Chem.* **2002**, *277*, 48535–48549.
3. Murakami, M.; Sato, H.; Taketomi, Y. Updating Phospholipase A(2) Biology. *Biomolecules* **2020**, *10*.
4. Tasoulis, T.; Isbister, G.K. A Review and Database of Snake Venom Proteomes. *Toxins (Basel)*. **2017**, *9*.
5. Tonello, F.; Rigoni, M. Cellular Mechanisms of Action of Snake Phospholipase A2 Toxins. In *Snake Venoms*; Inagaki, H., Vogel, C.-W., Mukherjee, A.K., Rahmy, T.R., Eds.; Springer Netherlands: Dordrecht, 2017; pp. 49–65 ISBN 978-94-007-6410-1.
6. Murakami, M.; Taketomi, Y.; Sato, H.; Yamamoto, K. Secreted phospholipase A2 revisited. *J. Biochem.* **2011**, *150*, 233–255.
7. Murakami, M.; Sato, H.; Miki, Y.; Yamamoto, K.; Taketomi, Y. A new era of secreted phospholipase A2. *J. Lipid Res.* **2015**, *56*, 1248–1261.
8. Ferraz, C.R.; Arrahman, A.; Xie, C.; Casewell, N.R.; Lewis, R.J.; Kool, J.; Cardoso, F.C. Multifunctional Toxins in Snake Venoms and Therapeutic Implications: From Pain to Hemorrhage and Necrosis. *Front. Ecol. Evol.* **2019**, *7*, 218.
9. Lomonte, B.; Gutiérrez, J.M. Phospholipases A2 from viperidae snake venoms: how do they induce skeletal muscle damage? *Acta Chim. Slov.* **2011**, *58*, 647–659.
10. Jungo, F.; Bougueleret, L.; Xenarios, I.; Poux, S. The UniProtKB/Swiss-Prot Tox-Prot program: A central hub of integrated venom protein data. *Toxicon* **2012**, *60*, 551–557.
11. Gouw, M.; Sámano-Sánchez, H.; Van Roey, K.; Diella, F.; Gibson, T.J.; Dinkel, H. Exploring Short Linear Motifs Using the ELM Database and Tools. *Curr. Protoc. Bioinforma.* **2017**, *58*, 8.22.1–8.22.35.
12. Kumar, M.; Gouw, M.; Michael, S.; Sámano-Sánchez, H.; Pancsa, R.; Glavina, J.; Diakogianni, A.; Valverde, J.A.; Bukirova, D.; Čalyševa, J.; et al. ELM-the eukaryotic linear motif resource in 2020. *Nucleic Acids Res* **2020**, *48*, D296–D306.
13. Massimino, M.L.; Simonato, M.; Spolaore, B.; Franchin, C.; Arrigoni, G.; Marin, O.; Monturiol-Gross, L.; Fernández, J.; Lomonte, B.; Tonello, F. Cell surface nucleolin interacts with and internalizes *Bothrops asper* Lys49 phospholipase A. *Sci Rep* **2018**, *8*, 10619.
14. Spolaore, B.; Fernández, J.; Lomonte, B.; Massimino, M.L.; Tonello, F. Enzymatic labelling of snake venom phospholipase A. *Toxicon* **2019**, *170*, 99–107.
15. Šribar, J.; Kovačič, L.; Oberčkal, J.; Ivanušec, A.; Petan, T.; Fox, J.W.; Križaj, I. The neurotoxic secreted phospholipase A. *Sci Rep* **2019**, *9*, 283.
16. Murakami, M.; Kambe, T.; Shimbara, S.; Kudo, I. Functional coupling between various phospholipase A2s and cyclooxygenases in immediate and delayed prostanoid biosynthetic pathways. *J. Biol. Chem.* **1999**, *274*, 3103–3115.
17. Rigoni, M.; Paoli, M.; Milanese, E.; Caccin, P.; Rasola, A.; Bernardi, P.; Montecucco, C. Snake phospholipase A2 neurotoxins enter neurons, bind specifically to mitochondria, and open their transition pores. *J. Biol. Chem.* **2008**.
18. Mahalka, A.K.; Code, C.; Jahromi, B.R.; Reza Jahromi, B.; Kirkegaard, T.; Jäätelä, M.; Kinnunen, P.K. Activation of phospholipase A2 by Hsp70 in vitro. *Biochim Biophys Acta* **2011**, *1808*, 2569–2572.
19. Code, C.; Domanov, Y.A.; Killian, J.A.; Kinnunen, P.K. Activation of phospholipase A2 by temporin B: formation of antimicrobial peptide-enzyme amyloid-type cofibrils. *Biochim Biophys Acta* **2009**, *1788*, 1064–1072.
20. Suzuki, T.; Bridges, D.; Nakada, D.; Skiniotis, G.; Morrison, S.J.; Lin, J.D.; Saltiel, A.R.; Inoki, K. Inhibition of AMPK catabolic action by GSK3. *Mol. Cell* **2013**, *50*, 407–419.
21. Souder, D.C.; Anderson, R.M. An expanding GSK3 network: implications for aging research. *GeroScience* **2019**, *41*, 369–382.
22. Nakatsu, Y.; Yamamotoya, T.; Ueda, K.; Ono, H.; Inoue, M.K.; Matsunaga, Y.; Kushiyama, A.; Sakoda, H.; Fujishiro, M.; Matsubara, A.; et al. Prolyl isomerase Pin1 in metabolic reprogramming of cancer cells. *Cancer Lett* **2020**, *470*, 106–114.
23. Liou, Y.-C.; Zhou, X.Z.; Lu, K.P. Prolyl isomerase Pin1 as a molecular switch to determine the fate of phosphoproteins. *Trends Biochem. Sci.* **2011**, *36*, 501–514.
24. Chioato, L.; Aragão, E.A.; Lopes Ferreira, T.; Medeiros, A.I.; Faccioli, L.H.; Ward, R.J. Mapping of the structural determinants of artificial and biological membrane damaging activities of a Lys49 phospholipase A2 by scanning alanine mutagenesis. *Biochim Biophys Acta* **2007**, *1768*, 1247–1257.
25. Madej, T.; Lanczycki, C.J.; Zhang, D.; Thiessen, P.A.; Geer, R.C.; Marchler-Bauer, A.; Bryant, S.H. MMDB and VAST+: tracking structural similarities between macromolecular complexes. *Nucleic Acids Res* **2014**, *42*, D297–303.