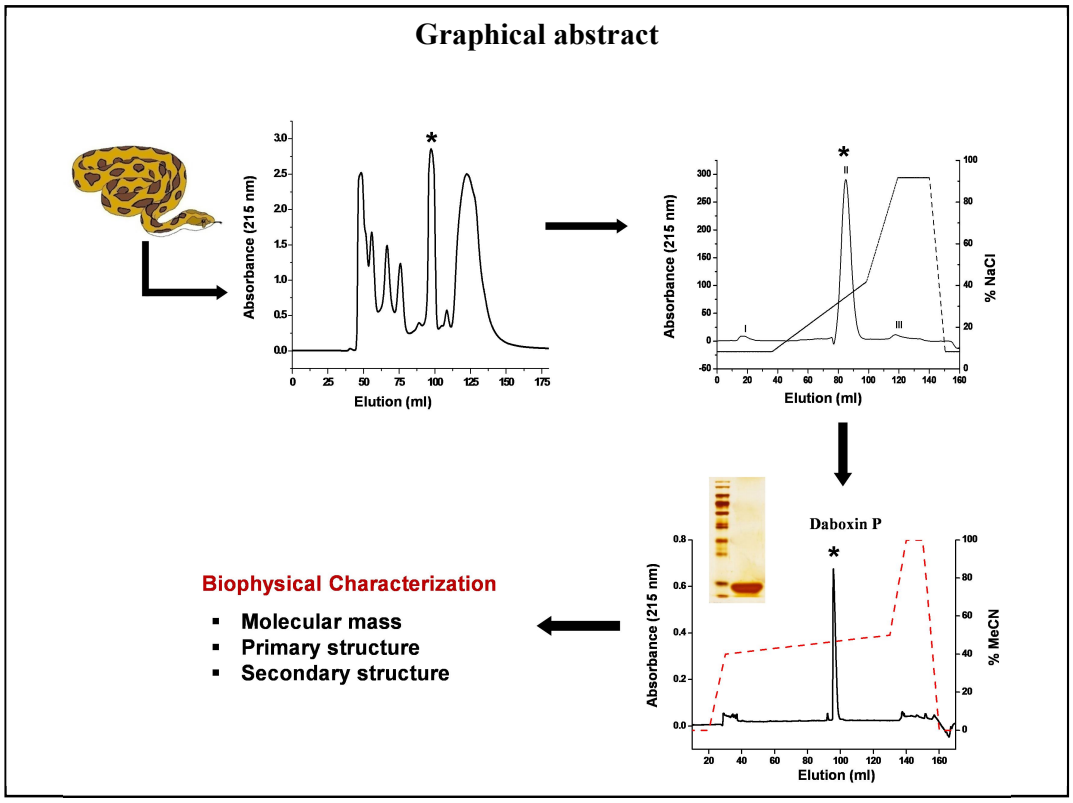
Chapter 4

Purification of a major protein from the venom of *Daboia russelii* (Irula) and its biophysical characterization



4.1 INTRODUCTION

The proteome of *Daboia russelii* venom is a rich source of both enzymatic and non-enzymatic protein families. PLA_2 enzymes are one of the most abundant enzymatic components in snake venom irrespective of their geographical origin. The members of this protein family are reported to be present in various isoforms within single snake venom (301,302). These isoforms exhibit plethora of pharmacological activities like myotoxicity, neurotoxicity, anticoagulation, edema inducing, haemolytic, etc. (46,239,455). Some of these proteins exist in the venom as monomers, dimers or multimers and enhance the toxicity of the venom (423).

In the previous chapter, the venom proteome of Indian *Daboia russelii* has been unfolded which advocates the presence of rich repertoire of various protein families with several isoforms. The proteome revealed the presence of three types of PLA_2 enzymes viz; S, H and N-type of based on the type of N-terminal amino acid residues. Moreover, acidic and alkaline forms of this enzyme were also observed in it. In the Myanmar viper venom, Daboiatoxin is the major protein which constitutes 12% of the total protein of the crude venom and exhibits neurotoxicity, edema, myonecrosis and haemolysis (426). However, proteomic analysis of the Indian viper did not reveal the presence of such peptide fragments (discussed in details in Chapter 3). Nonetheless, the gel filtration peak P6, contain a major protein in terms of its protein content which exhibits PLA_2 and potent anticoagulant activity. This protein might be one of the major proteins present in this venom proteome. Therefore, purification and understanding the biochemical, biophysical and biological activities of this protein might contribute to better understanding of the pathophysiological effects of this crude venom. The present chapter describes the purification and biophysical characterization of this major protein.

4.2 MATERIALS

4.2.1 Chemicals and Reagents

Protein sequencing: 25% trifluoroacetic acid, trifluoroacetic acid, 37% acetonitrile, chlorobutane, ethyl acetate, 5% phenyl isothiocyanate and 12% trimethylamine solution were purchased from Applied Biosystems chemicals (California, USA). BNPS-skatole [2-(2-Nitrophenylsulfenyl)-3-methyl-3-bromoindolenine)] was purchased from Bioworld (Ohio, USA). Hydroxylamine hydrochloride and 4-vinyl pyridine were bought from Sigma-Aldrich (Missouri, USA).

Other chemicals and reagents used are described in material section of previous chapters.

4.2.2 Column

The gel filtration column, Hiload^{TM} 16/600 Superdex 75 prep grade (1x120 ml) and ion exchange column, Hiprep CM FF 16/10 column (1.6x10 cm) were purchased from GE Healthcare life Sciences (Bucks, UK). The reversed phase HPLC column, Jupiter

C_{18} (3 µ, 4.6 x 250 mm, 300 Å) was bought from Phenomenex (California, USA) while Hypersil Gold C_{18} column (50 x 2.1 mm, 1.9 µm) was obtained from Thermo Scientific (Massachusetts, USA).

4.3 METHODS

4.3.1 Purification of the major anticoagulant PLA_2 enzyme

4.3.1.1 Gel filtration chromatography

For the isolation and purification of the major protein, lyophilized crude venom of *Daboia russelii* was subjected to gel filtration chromatography on Hiload^{TM} 16/600 Superdex 75 prep grade column (1x120 ml) as described in section 3.3.2 of Chapter 3.

4.3.1.2 Ion exchange chromatography

The gel filtration peak P6 with highest PLA_2 and anticoagulant activity compared to rest of the gel filtration peaks was subjected to Hiprep CM FF 16/10 (1.6x10 cm), a weak cation exchanger. Briefly, 500 µl of P6 (1 mg/ml in 50 mM of Tris-Cl, pH 7.4) was loaded onto the column pre-equilibrated with 50 mM of Tris-Cl, pH 7.4 using Äkta Purifier HPLC system. Fractionation was carried out at a flow rate of 2.25 ml/min with a linear gradient of 50 mM Tris-Cl, pH 7.4 containing 0.8 M NaCl. The elution profile was monitored at 215 nm and 1 ml of the fractions was collected.

4.3.1.3 Reversed phase high performance liquid chromatography

The ion exchange fraction (CM-II) with highest PLA_2 and anticoagulant activity was subjected to Rp-HPLC using Jupiter C_{18} column (3 µ, 4.6 x 250 mm, 300 Å). Briefly, 200 µl (0.5 mg/ml) of CM-II was loaded onto the column pre-equilibrated with 0.1% of TFA. Fractionation was carried out with a linear gradient of 80% MeCN containing 0.1% TFA using Äkta Purifier HPLC system. Elution was monitored at 215 nm and 1 ml of the fractions was collected.

To check the homogeneity of the Rp-HPLC purified protein, 5 µg of the protein (treated with 0.45 µl of βME) was loaded onto 12.5% of glycine SDS-PAGE. Electrophoretic run was carried out as described in section 2.3.2 of Chapter 2.

Subsequently, the gel was stained using Pierce™ silver staining kit, Life technologies (Thermo Fischer Scientific, Massachusetts, USA) according to the instructions of the manufacturer and documented using ChemiDoc™XRS+, Biorad (California, United States). This purified protein was named as daboxin P and was characterized in the succeeding experiments.

4.3.2 Biophysical characterization

4.3.2.1 Molecular mass determination

The molecular mass of the purified protein, daboxin P was determined by electrospray ionization mass spectrometry (ESI-MS) using Accela LCQ fleet Ion Trap, Thermo Scientific (Massachusetts, USA) (Figure 4.1). The system was nebulized with nitrogen gas and ion spray voltage of 4.4 kV was set for the program. MeCN (50%) containing 0.1% formic acid was used as the solvent and delivered into the system at a flow rate of 50 µl/min by Accela 600 pump. The ionization spectra of daboxin P were recorded in positive mode with mass to charge (m/z) ranging from 800 to 2000. The spectra was analysed and deconvoluted by Promas for Xcaliber, Thermo Scientific (Massachusetts, USA).

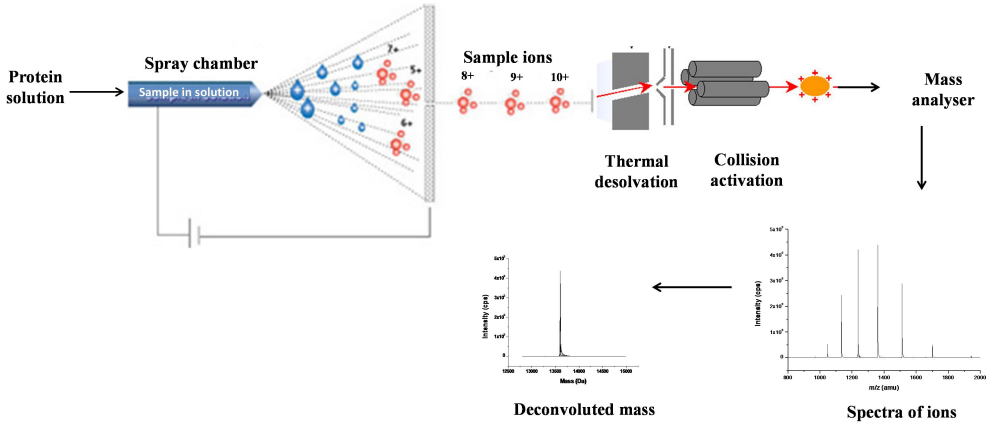


Figure 4.1: Schematic illustration of the various steps involved in the ionization of a protein solution by ESI-MS and molecular mass determination. (Adapted partly from QC analysis by mass spectrometry, Sigma-Aldrich and http://www.umass.edu/).

4.3.2.2 Primary structure determination

4.3.2.2.1 N-terminal sequencing

The N-terminal sequencing of daboxin P was carried out by automated Edman degradation process using Protein sequencer PPSQ-31A Shimadzu Asia Pacific, (Singapore) (Figure 4.2). Briefly, 100 µl of daboxin P (0.89 µg/µl equivalent to 6545.56 pmol) was air dried on a PVDF membrane and loaded onto the reaction chamber of the sequencer (Figure 4.3). Sequencing was performed till the chromatogram was readable.

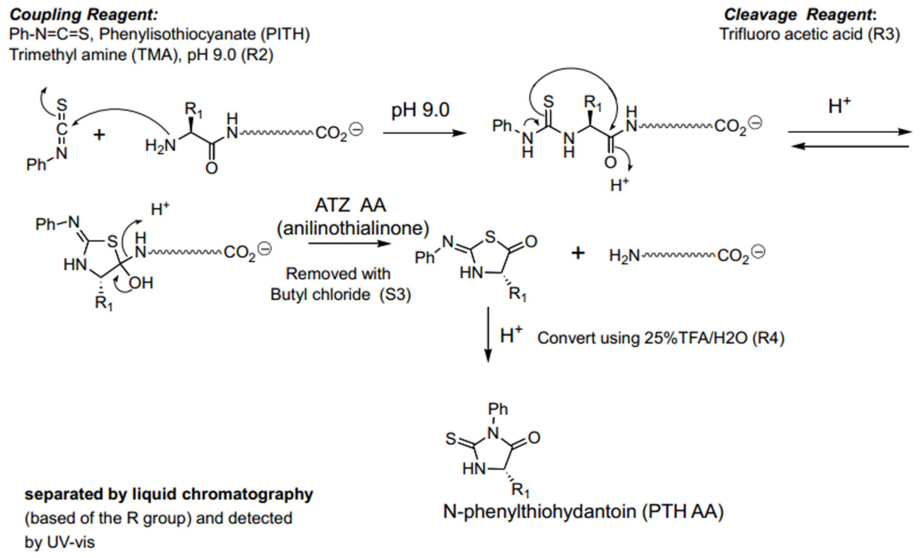


Figure 4.2: Pictorial representation of the reaction mechanism of Edman degradation (Adapted from "Introduction of Protein Sequencer PPSQ-31A", Shimadzu Asia Pacific Singapore). Under mild alkaline conditions in the presence of TMA (pH 9.0), phenyl isothiocyanate (PITC) couples to uncharged amino group of the N-terminal amino acid residue to form phenylthiocarbamoyl derivative. This is then cleaved by TFA under acidic conditions to form thiazolinone derivative. Subsequently, the derivative is then converted to phenylthiohydantoin (PTH)-amino acid (AA) derivative in the presence of 25% TFA. The PTH-AA formed is separated using liquid chromatography column (PTH column) connected to the sequencer. The PTH-AA derivative formed for each of the amino acid residues in the protein sequence has a specific elution time on PTH-LC column which helps in identification of these residues.

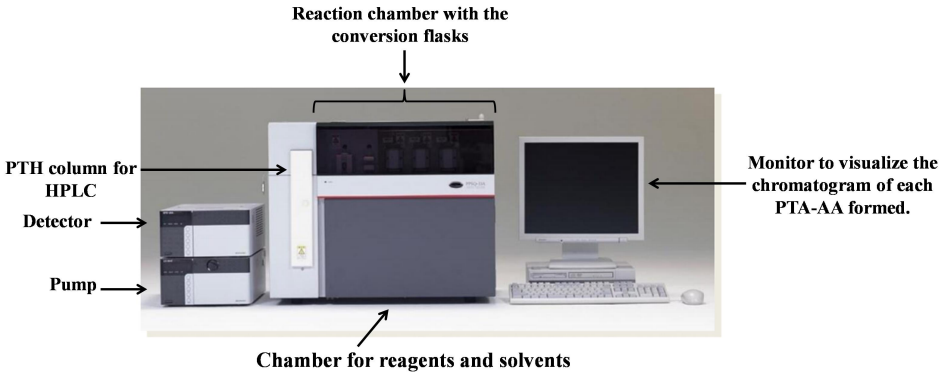


Figure 4.3: Schematic representation of the various parts of PPSQ-31A protein sequencer Shimadzu, Asia Pacific, Singapore (Adapted partly from "Introduction of Protein Sequencer PPSQ-31A", Shimadzu Asia Pacific Singapore**).**

4.3.2.2.2 Pyridylethylation

Preceding chemical or enzymatic treatment of daboxin P, it was denatured and reduced (456,457). Briefly, for denaturation of daboxin P (1 mg), it was treated with 920 µl of denaturing buffer (6 M guanidine hydrochloride in 0.1 M Tris-Cl pH 8.5). Subsequently, the cysteine groups were reduced with 30 µl βME. To expel dissolved oxygen, N_2 gas was purged over the reaction mixture and incubated for 150 min at 37^{o}C. Following this, 4-vinyl pyridine (50 µl) containing 100 ppm hydroquinone was added and incubated for 150 min at 37^{o}C. After incubation, the pyridylethylated daboxin P was desalted by Rp-HPLC using Jupiter C_{18} column (4.6 x 250 mm, 3 µm, 300 Å). Fractionation was carried out with a linear gradient of 38-52% of 80% MeCN containing 0.1% TFA for 104 min. Elution of the fractions was carried out at a flow rate of 0.8 ml/min and monitored at 215 nm.

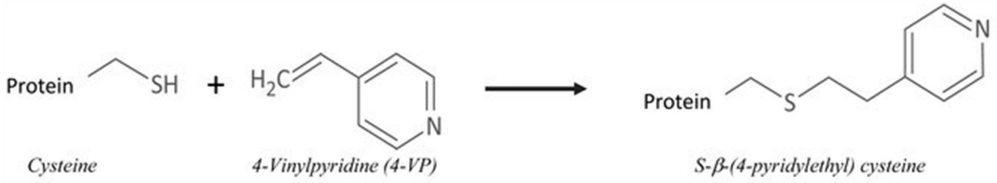


Figure 4.4: Schematic representation of the chemical reaction occurring in pyridylethylation of cysteine residues (Adapted from Rombouts., et.al (416)). Alkylation of cysteine groups helps to unfold a protein during denaturation step prior to protein cleavage. 4-vinyl pyridine, an alkylating agent converts the cysteine groups of an amino acid residue to S-pyridylethyl derivative which can be detected at 254 nm.

The molecular mass of the pyridylethylated daboxin P was analysed on ESI-MS, Accela LCQ fleet ion trap, and the desired fraction was lyophilized using pre-conditioned lyophilizer (112N-G, Hahntech, South Korea) at -80^{o}C.

4.3.2.2.3 BNPS-skatole cleavage

For chemical cleavage of pyridylethylated protein, 900 µg of the lyophilized daboxin P was treated with 0.1 % of TFA (500 µl) and 6 M guanidine hydrochloride (500 µl), pH 5.0 (458). To this reaction mixture, 1.5 mg/ml of BNPS skatole dissolved in 100 % acetic acid was added**.** The reaction blend was mixed properly by gentle shaking and kept in dark for 24 h at 37^{o}C to avoid reaction of BNPS-skatole with light. Following incubation, the reaction was stopped by addition of equal volume of Milli Q water and centrifuged at 12,000 rpm for 20 min. The clear supernatant was desalted on Jupiter C_{18} column with a linear gradient of 5-50% of 80% MeCN containing 0.1% TFA for 104 min. Fractionation was carried out at a flow rate of 0.8 ml/min and monitored at 215 nm.

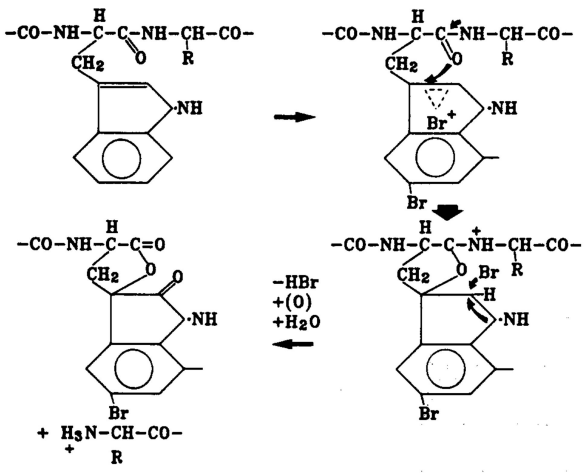


Figure 4.5: Reaction mechanism for cleavage of a protein sequence by BNPS-Skatole. It is an oxidising and brominating agent which cleaves at the C-terminus of the tryptophan (Trp-X) residue of a protein sequence.

The mass of the eluted fractions were analysed on ESI-MS (Accela LCQ fleet ion trap) for identification of the cleaved peptide fragments. The desired peptide fragments of daboxin P were lyophilized in the pre-conditioned lyophilizer at –80^{o}C. Following lyophilization, 89 µg of each of the lyophilized peptide fragment was

dissolved in 100 µl of Milli Q water and air dried on PVDF membrane. The dried membrane was loaded onto the reaction chamber of Automated Protein sequencer for performing Edman degradation.

4.3.2.2.4 Hydroxylamine hydrochloride cleavage

For cleavage with hydroxylamine hydrochloride, 1 mg of pyridylethylated daboxin P was treated with denaturing buffer (6 M guanidine hydrochloride, 0.1 M Tris-Cl, pH 9.0), 2 M hydroxylamine hydrochloride and 0.2 M dipotassium carbonate) (459). The reaction mixture was incubated for 4 h at 45^{o}C with gentle shaking. Following incubation, the mixture was desalted on Jupiter C_{18} column with a linear gradient of 5-50% of 80% MeCN containing 0.1% TFA for 104 min. Fractionation was carried out at a flow rate of 0.8 ml/min and monitored at 215 nm.

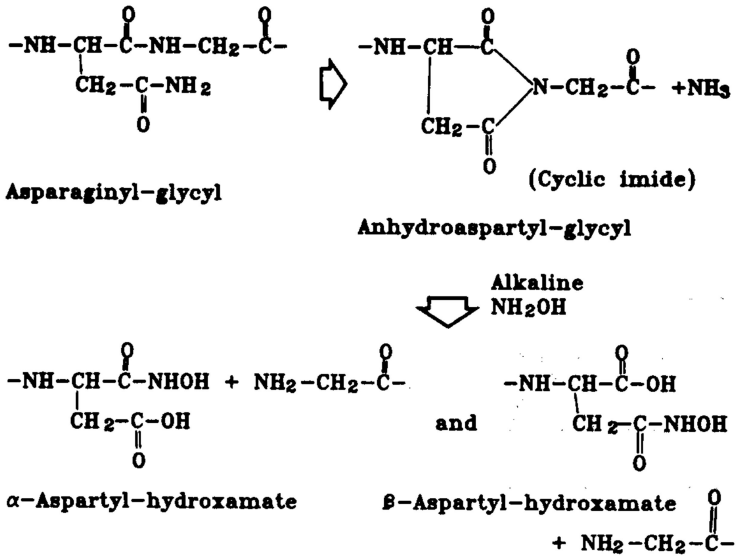


Figure 4.6: Reaction mechanism for cleavage of a protein sequence by hydroxylamine hydrochloride. Under mild alkaline conditions, hydroxylamine hydrochloride cleaves at the asparaginyl-glycine (N-G) bond of the protein sequence and generates peptide fragments.

The mass of the eluted fractions were analysed on ESI-MS (Accela LCQ fleet ion trap) for identification of the cleaved peptide fragments. The desired peptide fragments were lyophilized in pre-conditioned lyophilizer at –80^{o}C. Following this,

89 µg of the lyophilized sample was reconstituted in 100 µl of Milli Q water and air dried on PVDF membrane. The dried membrane was loaded into the reaction chamber of the Automated Protein Sequencer for performing Edman degradation.

4.3.2.3 Tandem mass spectrometry

For tandem mass spectrometry, 50 µg of daboxin P was denatured and alkylated in the presence of Protease max and trypsinized with 1 µg/µl of Trypsin. The detailed experimentation of mass spectrometry is described in section 3.3.4 of Chapter 3.

4.3.2.4 Multiple sequence alignment

For multiple sequence alignment, the PLA_2 enzymes from snake venom with maximum sequence similarity with daboxin P were retrieved from NCBI database by blastp (blast.ncbi.nlm.nih.gov) search. These sequences were aligned using DNAMAN 4.1.5.1 (Lynnon BioSoft) which is highly versatile and accurate software for carrying out various molecular biology applications. DNAMAN employs Clustal W algorithm for multiple sequence alignment (460).

4.3.2.5 Phylogenetic analysis

The phylogenetic relationship of daboxin P with the PLA_2 enzymes of snake venom exhibiting anticoagulant activity were constructed using the Mega 5.05 software (461). Anticoagulant PLA_2 enzymes having complete primary structure were retrieved from Pubmed Protein Database (www.ncbi.nlm.nih.gov/pubmed) and aligned using ClustalW with default parameters. To construct the phylogenetic tree of the sequences, Neighbour joining (NJ) algorithm was used which takes into account the distance matrix to calculate the distance between a pair of sequence (462). A bootstrap value of 1000 replicates was considered for the analysis. The pairwise distance among the sequences was calculated using the software Mega 5.0.

4.3.2.6 Secondary structure determination

The secondary structure of daboxin P was evaluated using Jasco Spectropolarimeter J-810 (Tokyo, Japan). In brief, 240 µg of daboxin P was dissolved in 600 µl of Milli Q water to make a final concentration of 0.4 mg/ml. Subsequently, 200 µl of the dissolved sample was loaded into a quartz cuvette having a path length of 0.1 cm and

subjected to the spectropolarimeter. The spectra of the differential absorbance of the plane polarized light by daboxin P was recorded at far UV scan from 260-190 nm at 25°C. The spectra were recorded at room temperature with a speed of 50 nm/min to determine the secondary structure. For the effect of temperature on the structural integrity of daboxin P, the scans were recorded at different temperature (from 25°C to 100°C). Spectra of Milli Q water without daboxin P, under the same experimental conditions was considered as the blank. To check the effect of pH on the structural conformation of daboxin P, it was dissolved in acidic (pH 3.0), basic (pH 7.4) and alkaline (pH 12) buffer. PBS buffer with the respective pH without daboxin P was considered as the blank for each experimental setup.

The scans were performed in triplicates and the average spectra of the sample were subtracted from the blank to obtain the final spectra of daboxin P.

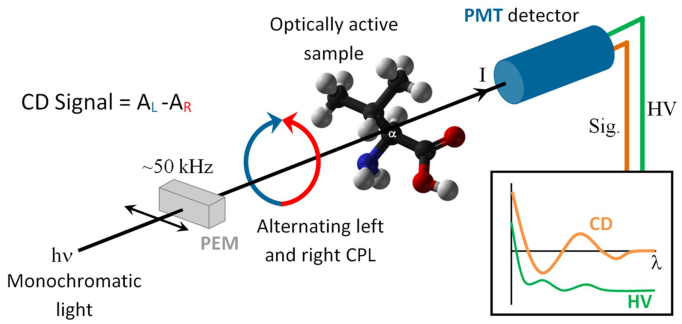


Figure 4.7: Pictorial representation of the underlying principle of circular dichroism spectroscopy (Adapted from www.isa.au.dk). The linearly polarised monochromatic light converts into alternating left and right handed polarised light when passed through a Photo Elastic Modulator (PEM). Upon passing through a sample, these differently polarized light gets absorbed variably which can be detected with a Photo Multiplier Tube (PMT).

4.4 RESULTS

4.4.1 Purification of daboxin P

4.4.1.1 Gel filtration chromatography

The gel filtration chromatography of the crude venom of *Daboia russelii* has resolved it into 8 peaks (P1 to P8) as shown in figure 4.8. Out of these 8 peaks, P6 accounted

for ~32% of the total protein of the crude venom with strong anticoagulant activity (clotting time higher than 600 s) and highest PLA_2 (93.69 µmol/min) (Figure 3.22 & 3.23 of Chapter 3). Based on maximum protein content and prominent tested activities, P6 was selected for purification of the major protein.

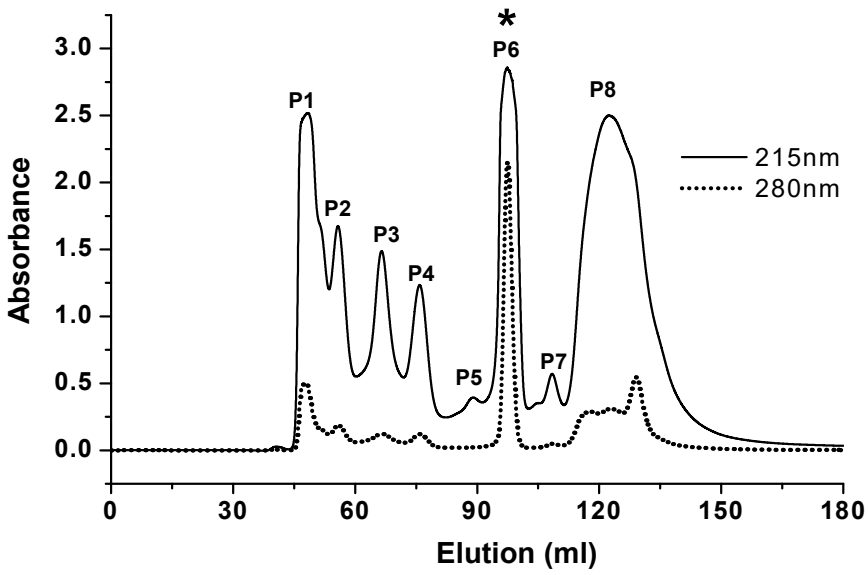


Figure 4.8: Fractionation of the crude venom of Indian *Daboia russelii* by gel filtration chromatography. Elution of the protein was monitored at 215 and 280 nm and the protein peaks were collected manually. * indicates the peak of interest.

4.4.1.2 Ion exchange chromatography

Fractionation of gel filtration peak P6 on weak cation exchange chromatography resolved it into two minor (CM-I and CM-III) and one major peak (CM-II) (Figure 4.9). It was observed that 3.33 µg/ml of all the three peaks delayed the recalcification time of plasma. However, the delay was more significant (p-value < 0.01) for CM-II with a 5.2 fold increase in the clotting time compared to NCT then CM-I (1.8 fold) and CM-III (1.1 fold) (Figure 4.10 A). Although 0.04 µg/ml of all the three peaks exhibited PLA_2 activity but the effect was more significant (p-value < 0.01) for CM-II (77.82 ± 2.79 u) compared to CM-I (9.59 ± 1.53 u) or CM-III (11.58 ± 2.01 u) (Figure 4.10 B).

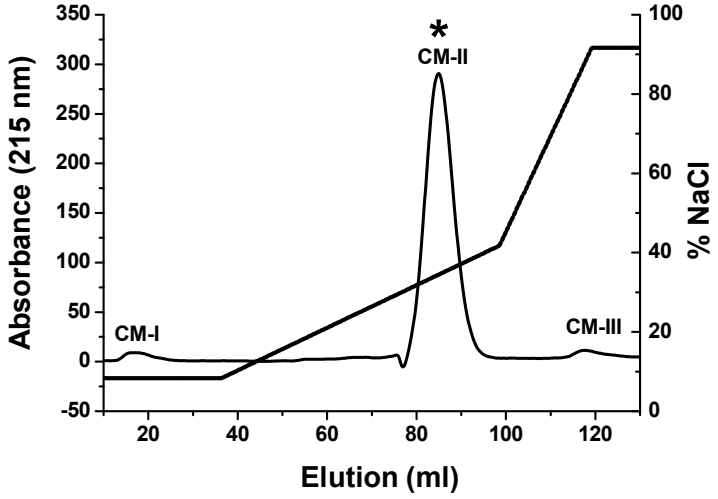


Figure 4.9: Ion exchange chromatography profile of P6: Fractionation was carried out on Hiprep CM FF 16/10 at a flow rate of 2.25 ml/min with a linear gradient of 50 mM Tris-Cl, pH 7.4 containing 0.8 M NaCl. Elution was monitored at 215 nm and the protein peaks were collected manually. * indicates the peak of interest.

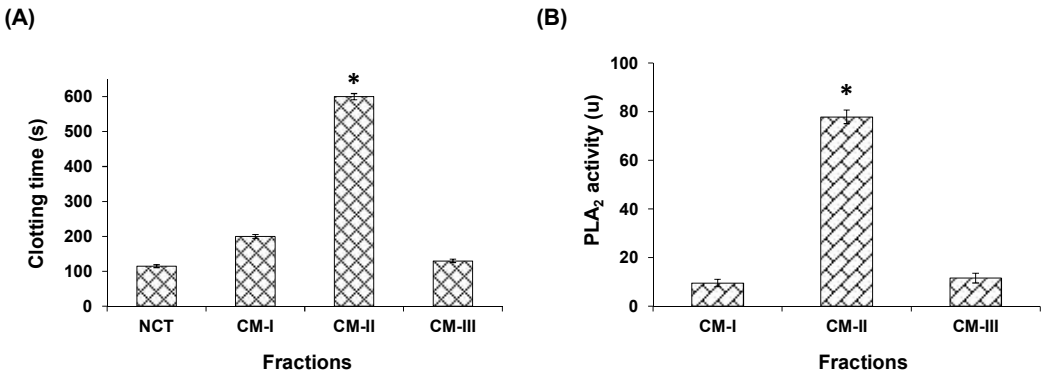


Figure 4.10: Recalcification time and PLA_2 activity of ion exchange fractions of P6. (A): Ion exchange fractions (3.33 µg/ml) were pre-incubated with PP plasma at 37°C for 2 min. Clot formation was initiated with 50 mM of $CaCl_2$. **(B):** The PLA_2 activity of ion exchange fractions (0.04 µg/ml) were determined using $sPLA_2$ assay kit. Amount of substrate hydrolysis was quantified at 414 nm. * indicates the major peak of interest. Results are mean ± SD of three independent experiments.

4.4.1.3 Rp-HPLC

Fractionation of ion exchange fraction, CM-II by Rp-HPLC further separated it into 2 prominent peaks, Rp-1 and Rp-2 (Figure 4.11 A). Although 3.33 µg/ml of both the

peaks delayed recalcification time of plasma, but the delay in clotting time was more significant (p-value < 0.01) for Rp-2 (beyond 600 ± 4.98 s) compared to Rp-1 (142.6 ± 3.43 s) (Figure 4.12 A). Moreover, 0.04 µg/ml of Rp-2 displayed significantly higher (p-value < 0.01) PLA_2 activity compared to Rp-1 (Figure 4.12 B).

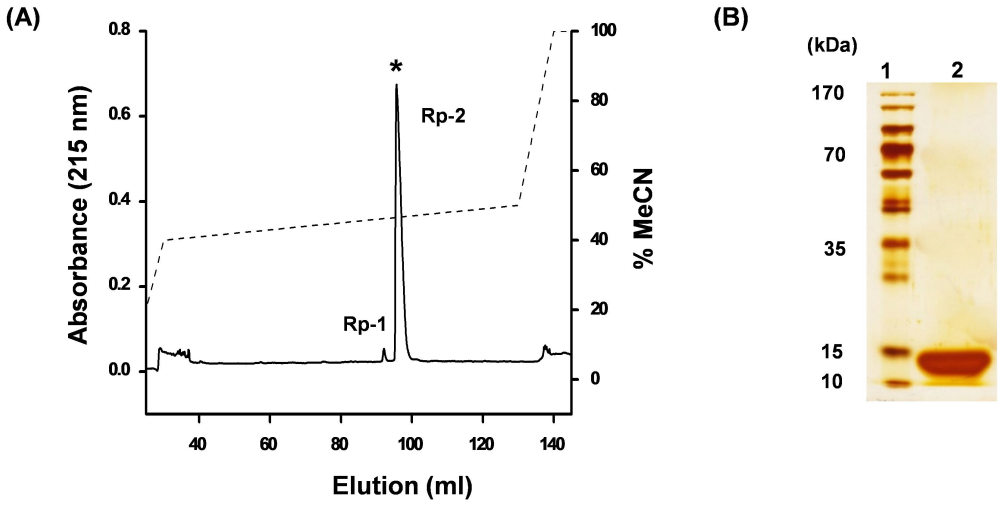


Figure 4.11: Rp-HPLC profile of CM-II and homogeneity of Rp-2. (A): RP-HPLC was carried out using Jupiter C_{18} column. Fractionation was monitored at 215 nm and the protein peaks were collected manually. * indicates the major peak of interest. **(B):** 12.5% Glycine SDS-PAGE profile of Rp-2 after silver staining. **Lane 1:** PageRuler^{TM} pre-stained protein marker ranging from 170-10 kDa. **Lane 2:** Rp-2 after treatment with β-mercaptoethanol.

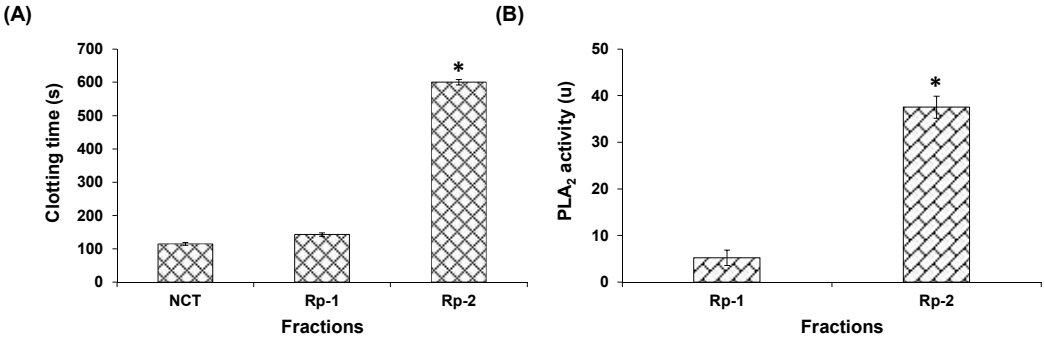


Figure 4.12 Recalcification time and PLA_2 activity of Rp-HPLC fractions of CM-II. (A): 3.33 µg/ml of the Rp-HPLC fractions were pre-incubated with PP plasma at 37°C for 2 min. Clot formation was initiated with 50 mM of $CaCl_2$. **(B):** The PLA_2 activity of 0.04 µg/ml of Rp-HPLC fractions were determined using $sPLA_2$ assay kit. Amount of substrate hydrolysis was quantified at 414 nm. * indicates the major peak of interest. Results are mean ± SD of three independent experiments.

The electrophoretic profile of Rp-2 displayed a single protein band between 15-10 kDa under reduced state (Figure 4.11 B). This confirms the homogeneity of Rp-2 and this purified major protein is named as daboxin P.

4.4.2 Biophysical characterization

4.4.2.1 Molecular mass of daboxin P

ESI-MS spectra of daboxin P revealed the presence of 6 mass to charge (m/z) ratio peaks with charges ranging from +8 to +13. The deconvoluted mass of the spectra displayed a molecular mass of 13597.62 ± 1.28 Da, which is in the range of PLA_2 enzymes (Figure 4.13).

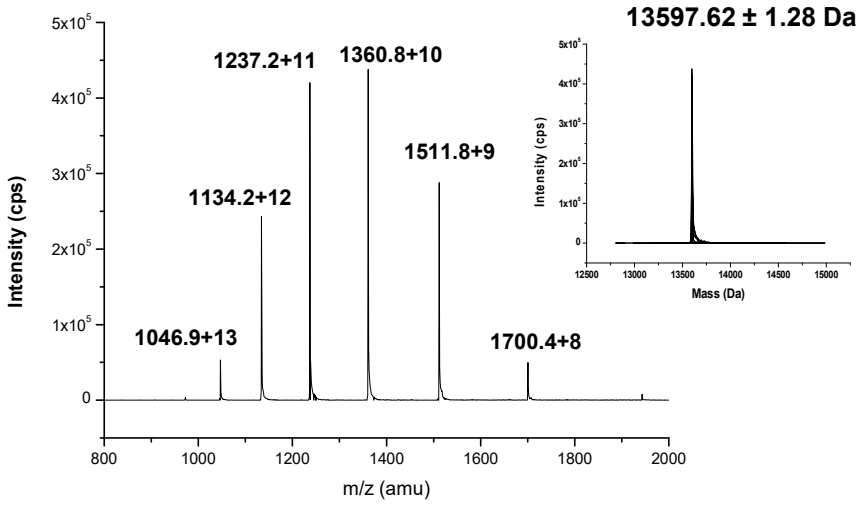


Figure 4.13: Molecular mass of daboxin P. The ESI-MS spectra show a series of multiple charged ions corresponding to a homogenous peptide. **Inset:** Reconstructed mass of daboxin P (cps: counts per second, amu: atomic mass unit).

4.4.2.2 Primary sequence of daboxin P

4.4.2.2.1 N-terminal sequencing

The first 30 residues of daboxin P were deduced by N-terminal sequencing using Edman degradation method (table 4.1). Sequencing reveals it to be a S-type PLA_2 enzyme as the first amino acid residue is found to be serine (S).

Table 4.1: Amino acid residues of daboxin P deduced by N-terminal sequencing. X represents Cys residue.

Method	Amino acid sequence
N-terminal	(1) SLLEFGKMILEETGKLAIPSYSSYGxYxGW (30)

4.4.2.2.2 Pyridylethylation

The desalted peak of pyridylethylated daboxin P was eluted at 44% of 80% MeCN (Figure 4.14 A). The mass of the protein was observed to be 15083.5 Da (Figure 4.14 B). The increase in mass of the protein after pyridylethylated was due to alkylation of cysteine residues by 4-vinyl pyridine which has a mass of 105.14 Da. This increase in molecular mass of 1486 Da corresponds to the presence of 14 cysteine residues in daboxin P.

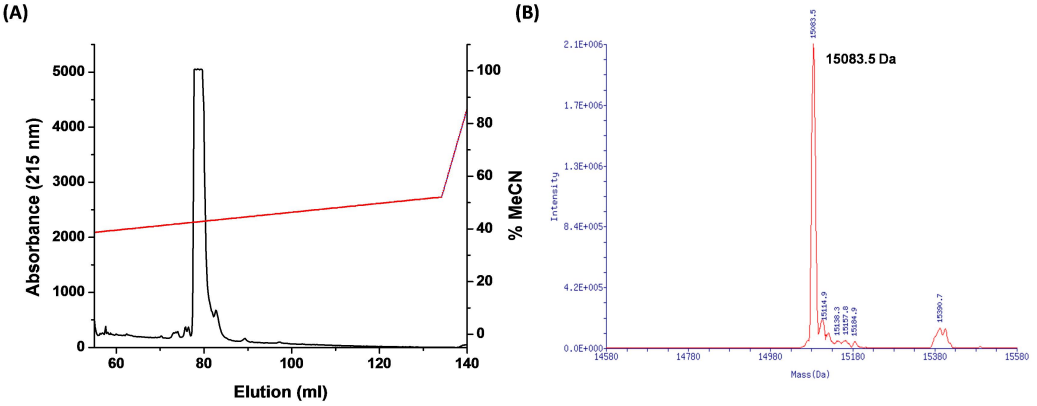


Figure 4.14: Rp-HPLC profile and molecular mass of pyridylethylated daboxin P. (A): Rp-HPLC profile of the desalted daboxin P at 215 nm on Jupiter C_{18} column after pyridylethylation. (B): Deconvoluted mass of pyridylethylated daboxin P by Promass for Xcaliber.

4.4.2.2.3 Cleavage with BNPS-skatole

Pyridylethylated daboxin P was cleaved into two peptide fragments designated as "a" and "b" (Figure 4.15 A). The peptide fragment "a" corresponds to 3.6 kDa while the peptide fragment "b" corresponds to 11.7 kDa (Figure 4.15 B & C). Edman sequencing of fragment "a" unveiled the first 30 residues of daboxin P while

sequencing of "b" unfolded next 38 amino acid residues from 31st to 68th position (Table 4.2).

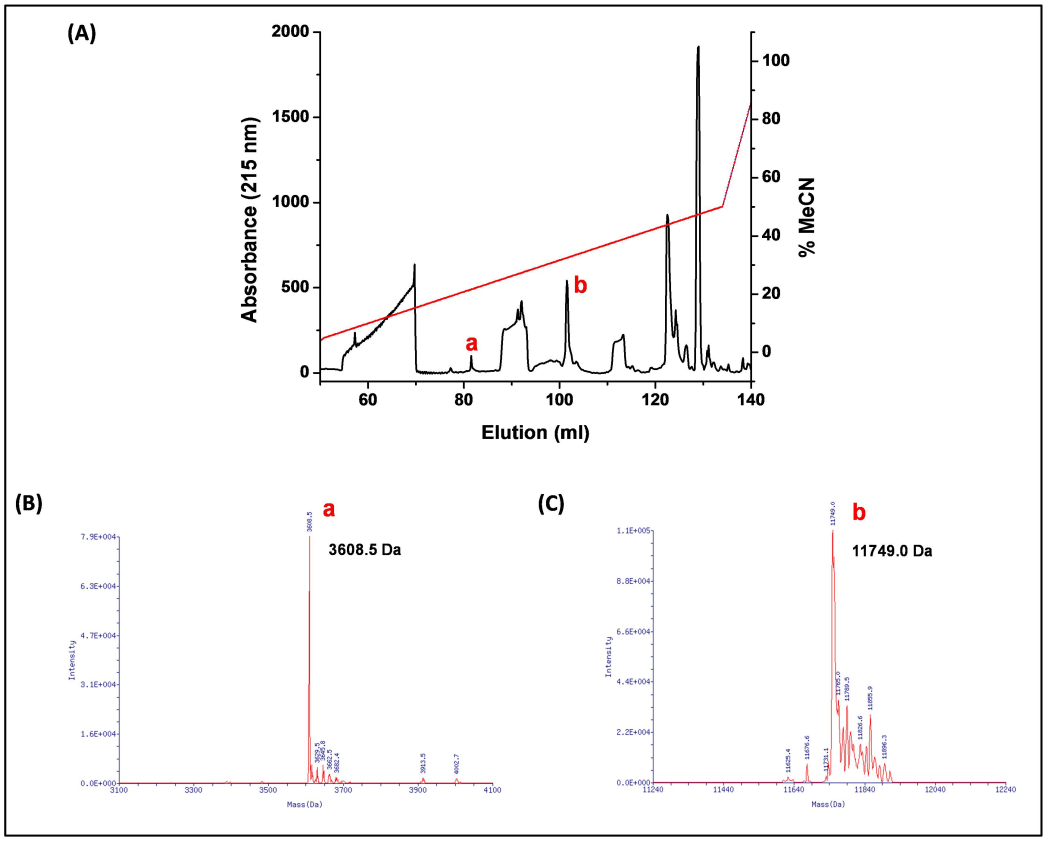


Figure 4.15: Rp-HPLC profile and molecular mass of the peptide fragments of daboxin P after cleavage with BNPS-skatole. (A): Rp-HPLC profile of the desalted fragments of daboxin P at 215 nm on Jupiter C_{18} column. Two peptide fragments were generated viz., "a" and "b". The mass of the cleaved and uncleaved peptides of daboxin P were analysed by ESI-MS. **(B) & (C):** Deconvoluted mass of the two peptide fragments of daboxin P were generated by Promass for Xcaliber.

Table 4.2: Sequence of amino acid residues deduced by Edman degradation of the peptide fragments of daboxin P after cleavage with BNPS-skatole.

Fragment no.	Mass (kDa)	Amino acid sequence
a	3.6	(1) SLLEFGKMILEETGKLAIPSYSSYGCYCGW (30)
b	11.7	(31) GGKGTPKDATDRCCFVHDCCYGNLPDCNNK SKRYRYKK (68) …..

4.4.2.2.4 Cleavage with hydroxylamine hydrochloride

Hydroxylamine hydrochloride fragmented daboxin P into two peptide fragments named as "a" and "b" (Figure 4.16). The peptide fragment "a" and "b" correspond to a molecular mass of 8.5 kDa and 6.5 kDa respectively (Figure 4.16 B & C). Sequencing of the peptide fragment "a" revealed the first 36 residues of daboxin P while sequencing of fragment "b" resolved the last 50 amino acid residues starting from 71st to 121st residue (Table 4.3).

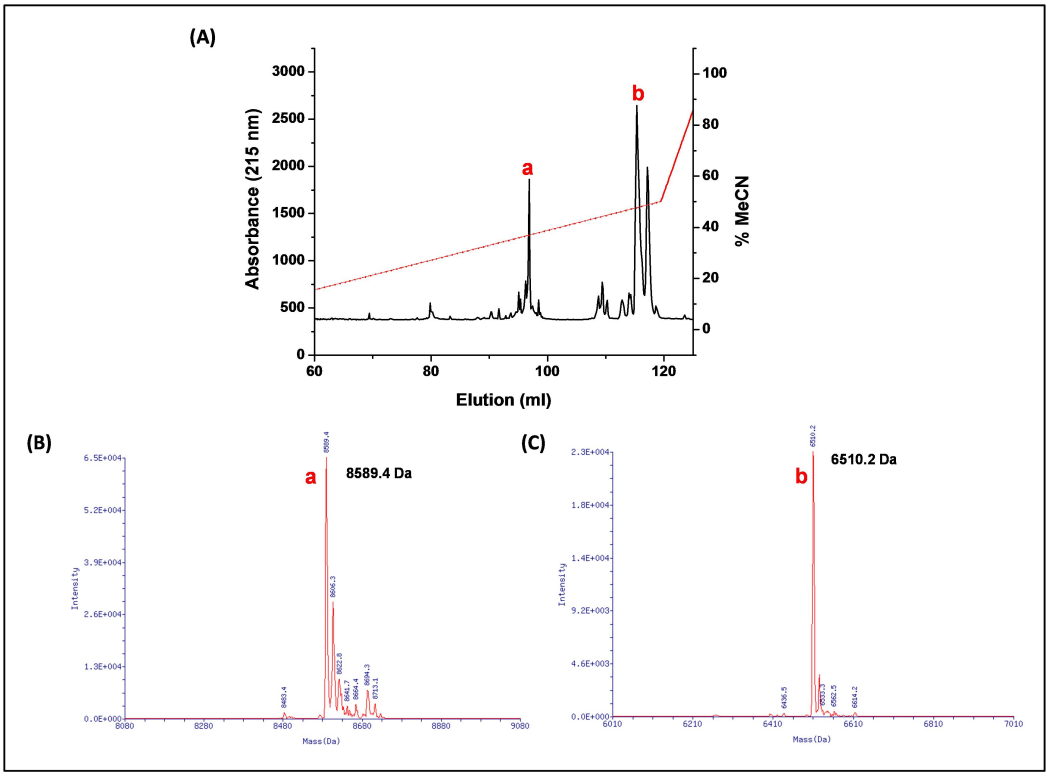


Figure 4.16 Rp-HPLC profile and molecular mass of the peptide fragments of daboxin P after cleavage with hydroxylamine hydrochloride. (A): Rp-HPLC profile of the desalted fragments of daboxin P at 215 nm on Jupiter C_{18} column. Two peptide fragments were generated viz., "a" and "b". The mass of the cleaved and uncleaved mass of daboxin P were analysed by ESI-MS. **(B) & (C):** Deconvoluted mass of the two peptides of daboxin P were generated by Promass for Xcaliber.

Table 4.3: Sequence of amino acid residues deduced by Edman degradation of the peptide fragments of daboxin P after cleavage with hydroxylamine hydrochloride.

Fragment no.	Mass (kDa)	Amino acid sequence
a	6.5	(1) SLLEFGKMILEETGKLAIPSYSSYGCYCGWGGKGTP (36)……
b	8.5	(71) GAIVCEKGTSCENRICECDKAAAICFRQNLNTYSKKYMLY PDFLCKGELVC (121)

4.4.2.3 Tandem mass spectrometry

Tandem mass spectrometry of trypsinized daboxin P resolved it into eleven peptide fragments. The amino acid sequence of the various peptide fragments obtained has been shown in Table 4.4.

The assembled protein sequence of daboxin P (Figure 4.17) submitted to UniProt (http://www.uniprot.org/) is available in NCBI database (https://www.ncbi.nlm.nih.gov/protein/C0HK16.1) with an accession number of C0HK16.1.

Table 4.4: Peptide fragments generated after tandem mass spectrometry of trypsinized daboxin P.

Sl. No.	MS/MS sequence	No. of peptides	MH+ (Da)	Z	Modifications	Coverage & score
1.	YMLYPDFLcKGELK	7	1776.14	3	C9*	80.17 & 31.65
2.	VNGAIVcEK	8	989.36	1	C7*	
3.	SLLEFGKMILEETGK	7	1695.85	3	-	
4.	GTScENTRIcEcDK	7	1457.00	2	C4*, 10*, 12*	
5.	QNLNTYSK	7	967.47	1	-	
6.	MILEETGK	10	920.47	1	-	
7.	LAIPSYSSYGcYcGWGGK	7	1911.51	3	C11*, 13*	
8.	KYmLYPDFLcK	7	1494.36	2	M3*, C10*	
9.	IcEcDK	11	824.27	1	C2*, C4*	
10.	AAAIcFR	23	808.40	1	C5*	
11.	IcEcDKAAAIcFRQNLNTYSK	8	2504.97	3	C2*,C4*, C11*	

MH+ stands for mass/charge (m/z) of the protonated molecular ions (peptide), **Z** stands for the number of charges a peptide carries after ionization, **Score** implies the sum of all peptide cross-correlation (Xcorr) values. **C*** stands for carbamidomethylation of cysteine residue; **M*** stands for oxidation of methionine residue.

Table 4.5: Various peptide fragments of daboxin P obtained after Edman degradation sequencing and tandem mass spectrometry.

Method	Peptide sequence
N-terminal sequencing	(1)SLLEFGKMILEETGKLAIPSYSSYGCYCGW(30)
BNPS-skatole cleavage	(1)SLLEFGKMILEETGKLAIPSYSSYGCYCGW(30)
	(31)GGKGTPKDATDRCCFVHDCCYGNLPDCNNKSKRY RYKK(68)
Hydroxylamine hydrochloride	(1)SLLEFGKMILEETGKLAIPSYSSYGCYCGWGGKGTP(36)
	(71)GAIVCEKGTSCENRICECDKAAAICFRQNLNTYSKK YMLYPDFLCKGELVC(121)
Tandem mass spectrometry	YMLYPDFLCKGELK
	VNGAIVCEK
	SLLEFGKMILEETGK
	GTSCENTRICECDK
	QNLNTYSK
	MILEETGK
	LAIPSYSSYGCYCGWGGK
	KYMLYPDFLCK
	ICECDK
	AAAICFR
	ICECDKAAAICFRQNLNTYSK

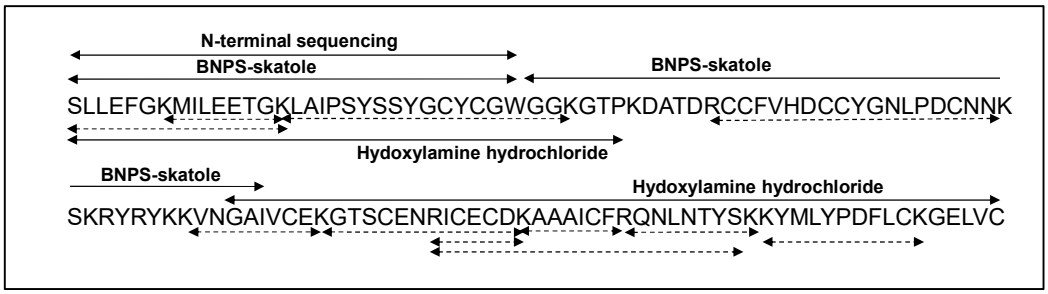


Figure 4.17: Assembled peptide fragments of daboxin P. Amino acid sequence was deciphered by Edman degradation and tandem mass spectrometry. The peptide sequences obtained from N-terminal sequencing and chemical cleavage by BNPS-skatole and hydroxylamine hydrochloride are shown with two headed solid arrows whereas peptide fragments obtained from tandem mass spectrometry are shown with two headed doted arrows.

4.4.2.4 Sequence analysis of daboxin P

The various peptide fragments of daboxin P obtained from N-terminal sequencing, chemical cleavage and tandem mass spectrometry have been shown in table 4.5. These peptide fragments were overlapped and assembled together to generate the complete sequence of daboxin P (Figure 4.17). It is composed of 121 amino acids with 14 cysteine residues. Based on the sequence, the calculated molecular mass of daboxin P is found to be 13596.70 Da which is in close agreement with the molecular mass of 13597.62 ± 1.28 Da observed in ESI-MS. Histidine residue is observed at the 48th position which is the active site residue, followed by an aspartate residue at the 49th position. The predicted anticoagulant region (53NLPDCNNKSKRYRYKK68) is found to be rich in positively charged amino acid residues like lysine and arginine. The homology search of daboxin P using blastp algorithm revealed its maximum sequence similarity to PLA_2 enzymes from various sub-species of *Daboia russelii* which are basic in nature. Amino acid substitutions at 5 different positions revealed its uniqueness from the already reported PLA_2 enzymes (Figure 4.18).

```
This study    SLLEFGKMILEETGKLAIPSYSSYGCYCGWGGKGTPKDATDRCCFV**H**DCCYGNLPDCN**N**KS
24638087      SLLEFGKMILEETGKLAIPSYSSYGCYCGWGGKGTPKDATDRCCFVHDCCYGNLPDCNPKS
408407675     SLLEFGKMILEETGKLAIPSYSSYGCYCGWGGKGTPKDATDRCCFVHDCCYGNLPDCNPKS
31615955      SLLEFGKMILEETGRLAIPSYSSYGCYCGWGGKGTPKDATDRCGFVHDCCYGNLPDCNPKS
49259309      SLLEFGKMILEETGKLAIPSYSSYGCYCGWGGSGTPKDATDRCGFVHDCCYGNLPDCNPKS
31615954      SLLEFGRMILEETGKLAIPSYSSYGCYCGWGGKGTPKDATDRCGFVHDCCYGNLPDCNPQS
109157490     SLLEFGKMILEETGKLAIPSYSSYGCYCGGCGSGTPKDATDRCCFVHCCCYGNLPDCNPKS
48425253      SLLEFGKMILEETGKLAIPSYSSYGCYCGGCGSGTPKDATDRCCFVHCCCYGNLPDCNPKS
298351762     SLLEFGMMILEETGKLAVPFYSSYGCYCGWGGKATPKDATDRCGFVHDCCYGNLPDCNPKS
81174981      SLLEFGMMILEETGKLAVPFYSSYGCYCGWGGKGTPKDATDRCGFVHDCCYGNLPDCTPKP

This study   **K**RY**R**YK**K**VNGAIVCEKGTSCENRICECDKAAAICFRQNLNTYSKKYMLYPDFLCKGEL**V**C
24638087     DRYKYKRVNGAIVCEKGTSCENRICECDKAAAICFRQNLNTYSKKYMLYPDFLCKGELKC
408407675    DRYKYKRVNGAIVCEKGTSCENRICECDKAAAICFRQNLNTYSKKYMLYPDFLCKGELRC
31615955     DRYKYKRVNGAIVCEKGTSCENRICECDKAAAICFRQNLNTYSKKYMLYPDFLCKGELKC
49259309     DRYKYKRVNGAIVCEKGTSCENRICECDKAAAICFRQNLNTYSKKYMLYPDFLCKGELKC
31615954     DRYKYKRVNGAIVCEKGTSCENRICECDKAAAICFRQNLNTYSKKYMLYPDFLCKGELKC
109157490    DRYKYKRVNGAIVCEKGTSCENRICECDKAAAICFRQNLNTYSKKYMLYPDFLCKGELKC
48425253     DRYKYKRVNGAIVCEKGTSCENRICECDKAAAICFRQNLNTYSGKYMLYPDFLCKGELKC
298351762    DRYKYKRVNGAIVCEQGTSCENRICECDKAAAICFRRNLNTYSKIYMLYPDFLCKGELKC
81174981     DRYKYKRVNGAIVCEQGTSCENRICECDKAAAICFTKNLNTYSKIYMLYPDFLCKGELKC
```

Figure 4.18: Multiple sequence alignment of daboxin P (C0HK16.1) with snake venom PLA_2 enzymes based on homology search using blastp algorithm. (**24638087:** *D. r. russelii*, **408407675:** *D. siamensis*, **31615955:** *D. r. pulchella*, **49259309:** *D. r. russelii*, **31615954:** *D r.pulchella*, **109157490:** *D. r. pulchella*, **48425253:** *D. r. pulchella*, **298351762:** *D. r. russelii*, **81174981:** *D. r. russelii*). Conserved cysteine residues are highlighted in grey and the amino acid substitutions in daboxin P are highlighted in bold and underlined. The His48 is highlighted in bold and the predicted anticoagulant region is marked with a solid black line at the top.

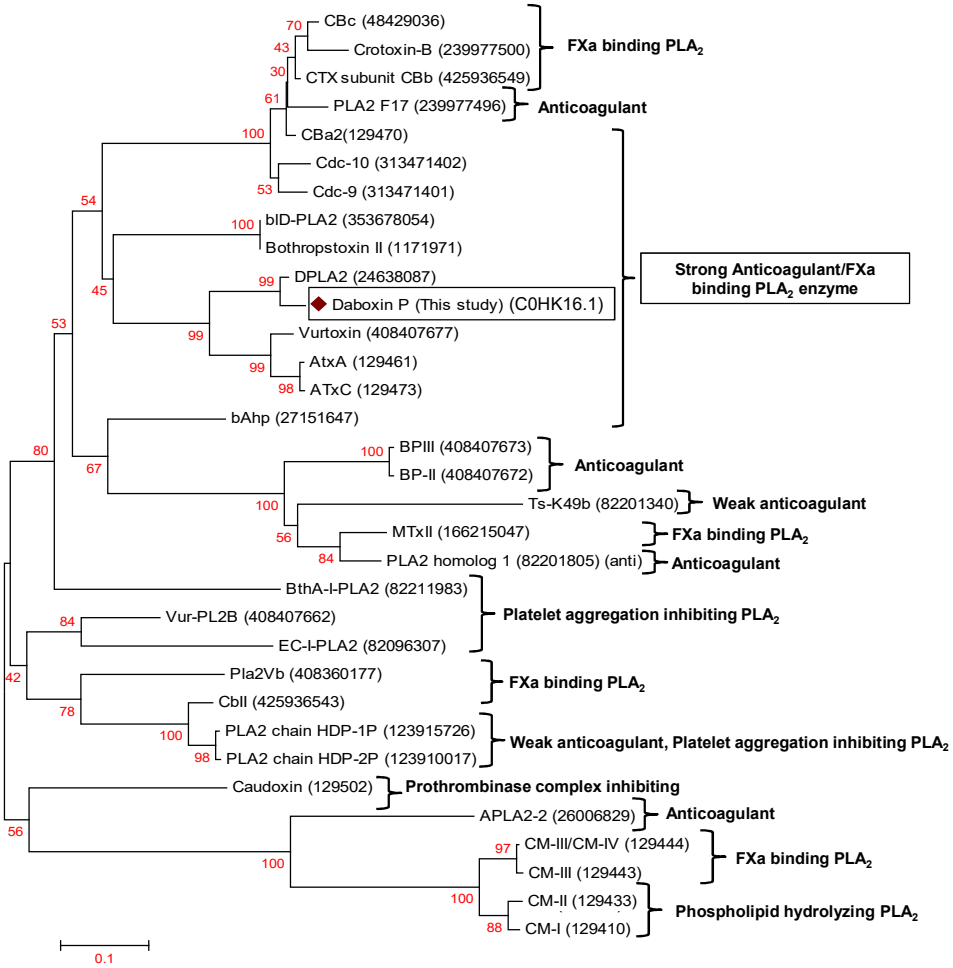


Figure 4.19: Phylogenetic tree of daboxin P with anticoagulant PLA_2 enzymes from snake venom. The sequences used to construct the tree were CBc (*Crotalus durissus terrificus*), CB crotoxin (*Crotalus durissus collilineatus*), CBb (*Crotalus durissus terrificus*), PLA2 F17 (*Crotalus durissus terrificus*), CBa2 (*Crotalus durissus terrificus*), Cdc-9 (*Crotalus durissus cumanensis*), Cdc-10 (*Crotalus durissus cumanensis*), bID-PLA2 (*Bothrops leucurus*), Bothropstoxin (*Bothrops jararacussu*), DPLA2 (*Daboia russelii russelii*), vurtoxin (*Vipera renardi*), AtxA (*Vipera ammodytes ammodytes*), AtxC (*Vipera ammodytes ammodytes*), bAhp (*Gloydius halys*), BP III (*Protobothrops flavoviridis*), II-BP (*Protobothrops flavoviridis*), Ts-K49b (*Trimeresurus stejnegeri*), MTxII (*Bothrops asper*), PLA_2 homology (*Bothrops atrox*), BthA-I-PLA2 (*Bothrops jararacussu*), Vur PL2B (*Vipera renardi*), EC-I-PLA2 (*Echis carinatus*), PLa2Vb (*Vipera berus berus*), CbII (*Pseudocerastes fieldi*), HDP-1P (*Vipera nikolskii*), HD-2P (*Vipera nikolskii*), Caudoxin (*Bitis caudalis*), APLA2-2 (*Ophiophagus hannah*); CM-III/CM-IV (*Naja nigricollis*); CM-III (*Naja mossambica*); CM-II (*Naja mossambica*) and CM-I (*Naja mossambica*).

4.4.2.5 Phylogenetic relationship of daboxin P

The phylogenetic relationship of daboxin P with the reported anticoagulant PLA_2 enzymes, showed its close relation to strong anticoagulant PLA_2 enzymes which are basic in nature (Figure 4.19). It has shown close evolutionary relation with vurtoxin from *Vipera renardi*, DPLA2 from *Daboia russelii pulchella*, AtxA and AtxC from *V. a. ammodytes* with a bootstrap value of 99 (304,354,387). Daboxin P is clustered along with PLA_2 enzymes which are reported either as strong anticoagulant or as FXa binding proteins. The p-distance of the tree is 0.1 suggesting a genetic change of 0.1%.

4.4.2.6 Secondary structure of daboxin P

The far UV CD spectrum of daboxin P, at 25^{o}C in Milli Q water has shown negative minima at 222 nm and 208 nm and a positive maximum at 190 nm which indicated that the secondary structure is mainly composed of α-helix and β-sheet (Figure 4.20 A). The percentage of secondary structure was estimated to be 42.73% of α-helix and 12.36% of β-sheet by the online software K2D3 (Figure 4.20 B) (463,464).

The structural conformation of daboxin P was found to be stable at acidic pH (3.0) and neutral pH (7.4) maintaining the percentage of α-helix and β-sheet (Figure 4.21 A). However, its structural integrity was completely distorted to 5.51% of α-helix and 23.75% of β-sheet at alkaline pH (12.0) (Figure 4.21 B).

Nonetheless, the CD spectra of daboxin P displayed a typical α-helical pattern up to 70^{o}C beyond which there was a gradual loss of secondary structure (Figure 4.22 A). Its melting temperature (Tm) was calculated to be 71.59 ± 0.4^{o}C (Figure 4.22 B). At 100^{o}C, its secondary structure contained 2.29% of α-helix and 42.16% of β-helix (Figure 4.22 C).

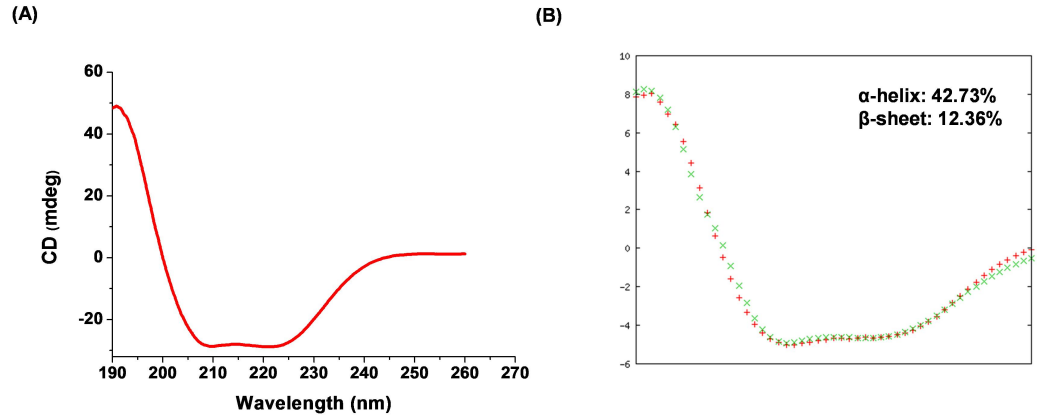


Figure 4.20: Secondary structure of daboxin P (**A):** Far UV CD spectrum of daboxin P dissolved in Milli Q water at 25°C, recorded using Jasco J-18 spectropolarimeter. **(B):** Predicted percentage of secondary structure of daboxin P using K2D3 software. The red points correspond to the input spectrum and the green points correspond to the predicted spectrum.

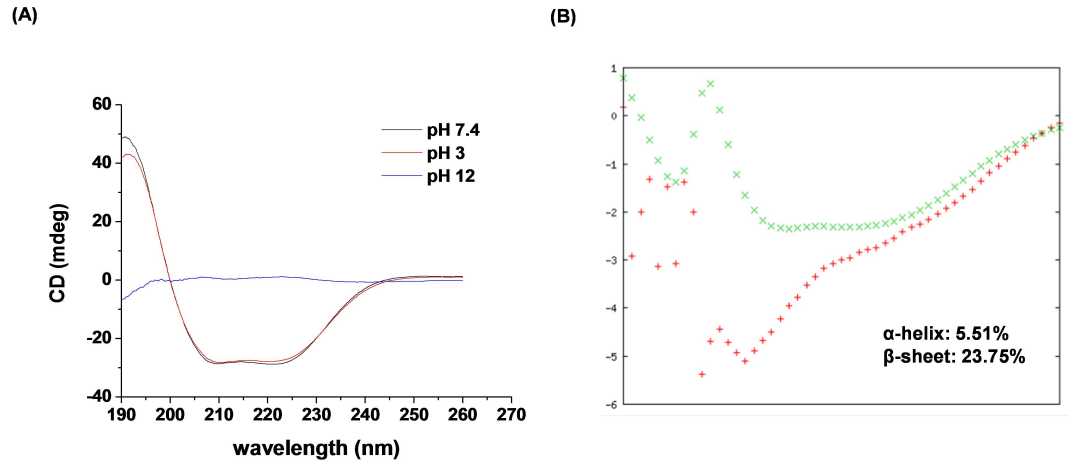


Figure 4.21: Secondary structure of daboxin P at different pH. (**A):** Far UV CD spectrum of daboxin P was recorded at acidic (3.0), neutral (7.4) & alkaline pH (12), recorded using Jasco J-18 spectropolarimeter. **(B):** A snapshot of the secondary structure of daboxin P at pH 12.0 using K2D3 software. The red points correspond to the input spectrum and the green points correspond to the predicted spectrum.

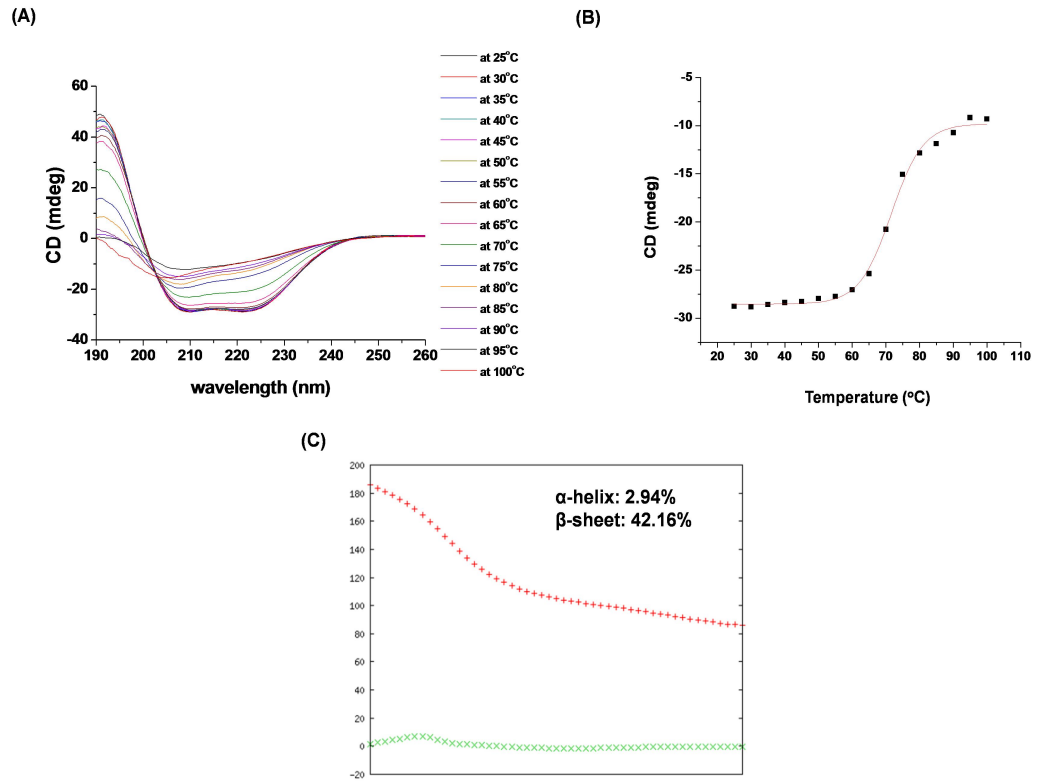


Figure 4.22: Secondary structure of daboxin P at different temperature. (**A):** Far UV CD spectra of daboxin P from 25°C to 100°C at pH 7.4, recorded using Jasco J-18 spectropolarimeter. **(B)** Melting curve of daboxin P plotted using sigmoidal curve fit at 222 nm, taking temperature as a function. Tm value is calculated by Boltzman equation using Origin (OriginLab). **(C):** Graphical representation of the predicted percentage of secondary structure of daboxin P at 100°C using K2D3 software. The red points correspond to the input spectrum and the green points correspond to the predicted spectrum.

4.5 DISCUSSION

Russell's viper is one of the most venomous snakes of the Indian sub-continent responsible for most of the envenomation cases (9). Thus, having a comprehensive understanding of its venom proteome along with characterization of its major protein is crucial for better management of snakebite cases.

PLA_2 enzymes are one of the most abundant protein families in snake venom responsible for manifestation of wide range of clinical signs and symptoms in envenomated victims. Altogether 15 isoforms of PLA_2 enzymes were identified in the

proteome of Indian *Daboia russelii* venom (discussed in details in Chapter 3). A major protein was identified in the gel filtration peak P6 of the crude venom which was purified to homogeneity by ion exchange chromatography followed by Rp-HPLC. This purified protein, named daboxin P exhibited highest anticoagulant and PLA_2 activity and accounts for 24% of the total protein of the crude venom. It showed single protein band of ~14 kDa under reduced and non-reduced state on SDS-PAGE suggesting it to be monomeric in nature. The exact mass of daboxin P was found to be 13597.62 ± 1.28 Da by ESI-MS which is in the range of PLA_2 enzymes (13-15 kDa) (184). Its primary structure consists of 121 amino acids with 14 cysteine residues that correspond to 7 disulphide bond. The 7 disulphide bonds in this group of PLA_2 enzymes are reported to be formed between Cys27-Cys126; Cys29-Cys45; Cys44-Cys105; Cys51-Cys98; Cys61-Cys91; Cys84-Cys96; Cys50-Cys133 (465). The disulphide bond formed between the 50^{th} cys and the 133^{rd} cys residue is known to be the unique pair for group II A PLA_2 enzymes along with the extended C-terminal loop which differentiate them from the group I PLA_2 enzymes (240,265,266). The C-terminal loop in daboxin P comprises 114SKKYMLYPD-FL**C**KGEL-V**C**133 (sequence is numbered based on Renetseder numbering system (350). This group of enzymes are proposed to have evolved from non-toxic ancestral PLA_2 genes after species diversification and are abundantly reported in human synovial fluid and rattle snake venom apart from the viper venoms (275,280). At position 48 and 49 of the primary sequence of daboxin P, histidine and aspartate residues are observed respectively suggesting it to be a calcium dependent PLA_2 enzyme (466). The His48-Asp49 pair plays a very crucial role in hydrolysis of phospholipids in the presence of calcium ions as the co-factors (271). Apart from these residues, Tyr52, Tyr73 and Asp99 are observed in daboxin P that are also reported to play a crucial role during catalysis (267). The presence of aspartate residue at the 49^{th} position in daboxin P categorizes it as Asp49 group II A PLA_2 enzymes (271). Figure 4.23 shows a summary of the primary structure of daboxin P with conserved disulphide bonds and amino acid segments. The members of group IIA PLA_2 enzymes are also reported with other amino acid residues like serine, asparagine, lysine or arginine at the 49^{th} position, categorizing them as Ser49, Asn49, Lys49 or Arg49 respectively (268-270,272). The absence of Asp49 residue in these PLA_2 enzymes prevent their binding to the Ca^{2+} ions thus lowering their catalytic efficiency. Hence, the Asp49 substituted

PLA_2 enzymes are often reported with reduced or no enzymatic effect on the phospholipid hydrolysis (273).

Figure 4.23: Primary structure of daboxin P. The sequence is numbered based on Renetseder numbering system (350). The calcium binding loop (25YGCYCGWGG33) is highlighted with a box, the catalytically active residues (His48-Asp49) are highlighted with **. The predicted anticoagulant region (54NLP-D-C-----NNKSKRYRYKK77) is underlined. The C-terminal region is underlined with a dotted line. The disulphide bonds between the cysteine residues are connected with black dotted lines.

Kini and Evans proposed that the presence of positively charged amino acid residues in the anticoagulant region as an essential criteria for strong anticoagulant PLA_2 enzyme (292,349). On the other hand, the weak or non-anticoagulant PLA_2 enzymes contain neutral or acidic amino acid residues in this region. The presence of positively charged amino acid residues like lysine and arginine in the predicted anticoagulant region of daboxin P suggests it to be a strong anticoagulant PLA_2 enzyme.

Apart from the anticoagulant region, various other pharmacological sites have been predicted in snake venom PLA_2 enzymes. Site directed mutagenesis study revealed that the presence of exposed Phe124 residue at the C-terminal region of group II A PLA_2 enzymes play a crucial role in exhibiting neurotoxicity (314). On the other hand, the presence of hydrophobic as well as cationic sites and overall basicity of the PLA_2 enzymes are essential for the myotoxic and antibacterial activity respectively (319,327,328). Thus, the presence of phenylalanine at the 124^{th} position and the overall basicity in daboxin P suggest it to be a potential neurotoxic and myotoxic PLA_2 enzyme.

Based on the amino acid sequence, the theoretical pI of daboxin P is calculated to be 8.5 by the online software ExPASy ProtParam (http://web.expasy.org/protparam/) suggesting it to be basic in nature (467). Blastp analysis also shows its sequence similarity with basic PLA_2 enzymes like VRV-PL-VIIIa and VRV-PL-V from *Daboia r. russelii*, DsM-S1 from *Daboia siamensis,* further confirming the basic nature of daboxin P (387,429,468). Daboxin P displayed maximum sequence similarity of 96% with VRV-PL-VIIIa (accession no. 24638087) from *Daboia russelii* (387). The amino acid residues of VRV-PL-VIIIa viz., Pro59, Asp62, Lys65, Arg68 and Lys120 were found to be substituted to Asn59, Lys62, Arg65, Lys68 and Val120 in daboxin P (Figure 4.24). Although there is a substitution from Lys65 to Arg65 and Arg68 to Lys68 in daboxin P but the overall charge in these two positions is unchanged. However, at position 59 and 62, proline (polar uncharged) and aspartate (negatively charged) are replaced by asparagine (polar uncharged) and lysine (positively charged) in daboxin P. At its C-terminal loop Lys120 (positively charged) substituted to Val120 (nonpolar) in daboxin P. Hence, the overall substitutions at these positions in the two proteins are well balanced with one extra positive charge to daboxin P which makes it more basic than VRV-PL-VIIIa (theoretical pI 8.35 based on amino acid sequence).

```
Daboxin P     SLLEFGKMILEETGKLAIPSYSSYGCYCGWGGKGTPKDATDRCCFVHDCCYGNLPDCNNKS
24638087      SLLEFGKMILEETGKLAIPSYSSYGCYCGWGGKGTPKDATDRCCFVHDCCYGNLPDCNPKS

Daboxin P     KRYRYKKVNGAIVCEKGTSCENRICECDKAAAICFRQNLNTYSKKYMLYPDFLCKGELVC
24638087      DRYKYKRVNGAIVCEKGTSCENRICECDKAAAICFRQNLNTYSKKYMLYPDFLCKGELKC
```

Figure 4.24: Sequence alignment of daboxin P (C0HK16.1) with VRV-PL-VIIIa (24638087). The substituted amino acid residues in daboxin P with respect to VRV-PL-VIIIa are highlighted in bold and underlined. The conserved cysteine residues are highlighted in grey.

PLA_2 enzymes are known to share a highly conserved scaffold of secondary structure with an N-terminal α-helix A, helix B, Ca^{2+} binding loop, Helix C, β-wing of two antiparallel β-sheets, helix D followed by the C-terminal loop (180). The residues of the N-terminal helix (1-12) are reported to form an integral part of the hydrophobic channel opening and stabilize the adjacent β-sheets (267,469). On the other hand, the

calcium binding loop comprising of the conserved residues of "YGCYCGXGG", aid in facilitating the binding of the Ca^{2+} ions to the enzyme during catalysis (240). The circular dichroism spectrum of daboxin P has revealed an α-helical pattern typical for PLA_2 enzymes with 42.73% of α-helix and 12.36% of β-sheets.

The secondary structure of daboxin P is found to be stable at acidic and neutral pH conditions but not at alkaline condition. The stability of the secondary structure of daboxin P, at acidic or neutral pH and its instability at alkaline pH could be attributed to the role of intramolecular and intermolecular hydrogen (H) -bonding. The presence of excess of hydroxyl ions in the buffer at high pH might have interfered with the H-bonds involved in maintaining the secondary structure of daboxin P within itself and also with its surrounding. This interference in H-bonding in turn, might have caused its structural deformation at alkaline condition. The thermal stability of daboxin P was observed up to 70^{o}C with a Tm value of 71.59 ± 0.4^{o}C. The loss of structural conformation at higher temperature could also be accredited to the role of the H-bonds which breaks at high temperature. Although H-bonds are not very strong interactions but they play one of the most crucial roles in protein folding and maintaining rigidity in protein structure. To satisfy this interaction within the hydrophobic core of protein and its aqueous environment, α-helix and β-sheets are formed. Adverse conditions like alkaline pH or high temperature might temper the stability of these weak yet very important non-covalent interactions. Thus, daboxin P might have lost its structural integrity at higher pH and temperature due to destabilization of its H-bonds and might lose its activity.

Thus, the present study describes the purification of a major protein from the venom of Indian *Daboia russelii* by a collective approach of chromatographic techniques. This major protein, named daboxin P has a molecular mass of 13597.62 ± 1.28 Da. Its primary sequence is composed of 121 amino acids with 14 cysteine reisdues and His48-Asp49 pair. These characteristics suggest it to be a calcium dependent catalytic PLA_2 enzyme. Its CD spectrum shows α-helical pattern typical for PLA_2 enzymes which is stable at acidic and neural pH and has a Tm value of 71.59 ± 0.4^{o}C.