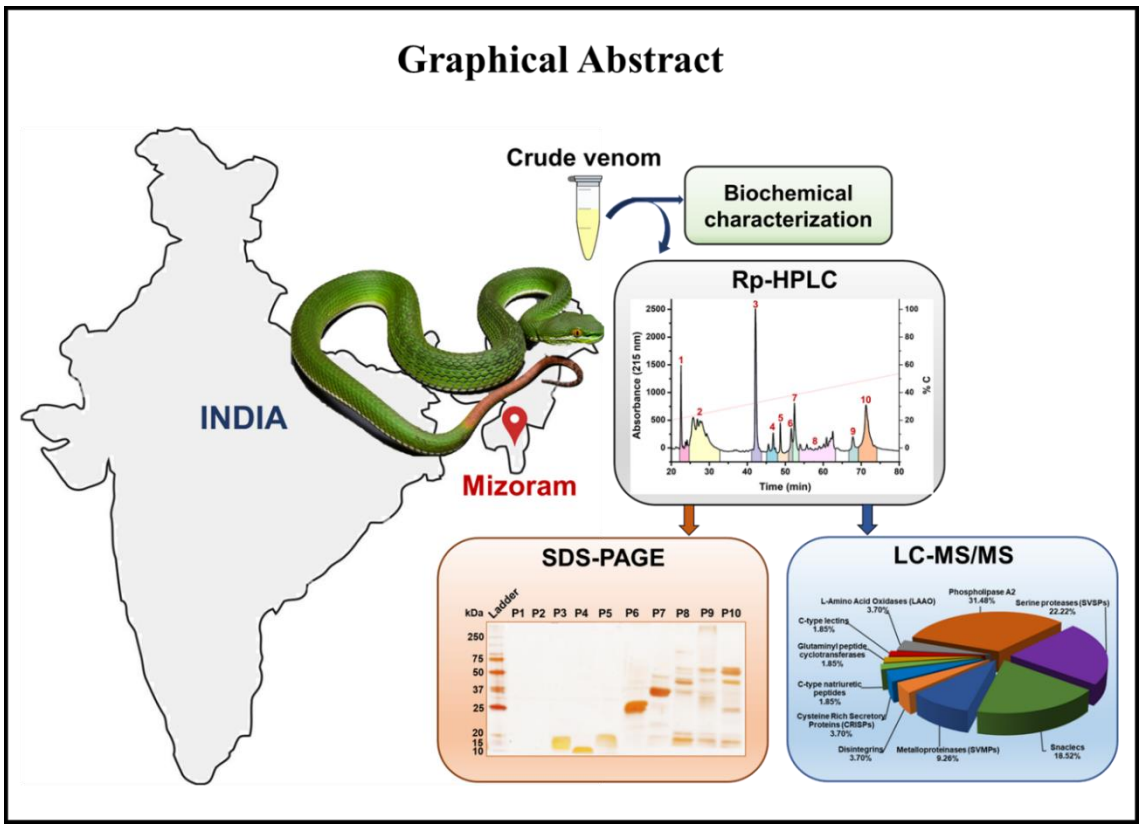
CHAPTER 5

Proteomics of *Trimeresurus erythrurus* venom from Mizoram, India

Chapter 5

Proteomics of *Trimeresurus erythrurus* venom from Mizoram, India



5.1. Introduction

Studies elucidating the venom composition of various snakes and understanding their mechanism of action causing pathophysiological alterations has remained a fruitful subject of investigation for toxinologists all around the world. The advancement in "omics" technology has made the understanding of venom dynamics more feasible. Consequently, venom composition of various green pit vipers has been studied with the help of proteomic approach, e.g., *Viridovipera stejnegeri*, *Trimeresurus albolabris*, *Popeia nebularis*, *T. insularis* and *T. purpureomaculatus* [121, 123, 125, 126]. Further, several studies have established a correlation between venom composition and the clinical toxicity presented in snakebite victims, thereby adding to the knowledge of venom-induced pathophysiology. For instance, predominance of snake venom

metalloproteases (SVMPs) in the venom of *Popeia nebularis* has been found to be responsible for acute edema and haemorrhage in the envenomated victims [121]. Therefore, understanding the snake venom complexity provides a vast range of information which helps to comprehend the venom variation among closely related species, its pharmacological profile and toxins responsible for the pathophysiological abnormalities in victims.

Trimeresurus erythrurus, commonly known as the spot-tailed pit viper, is a member of the widespread *Trimeresurus* complex of Asiatic pit vipers. They are arboreal pit vipers that are green in overall coloration, possessing a prehensile tail adorned with reddish or brownish spots. The envenomation reports of *T. erythrurus* are scanty, however, a few studies have reported the medical importance and bite-associated pathophysiology of green snakes assumed to be *T. erythrurus* based on their inhabitance range [114, 133]. The clinical manifestation is characterized by painful swelling at the bite site, along with prolonged coagulopathy leading to thrombocytopenia, hypotension and shock in a few patients [133]. Some of these clinical complications might be associated with morphological misidentification of green pit vipers, thereby leading to incorrect treatment causing further complications in patients. Moreover, there is a lack of information comprehending detailed venom composition of *T. erythrurus* and the toxins responsible for pathological disorders. Owing to their medical relevance, the current study was undertaken to perform the complete venom decomplexation of *T. erythrurus* from Mizoram, India to understand detailed venom composition uisng shot-gun proteomics.

5.2. Results

5.2.1. Fractionation of crude venom by Rp-HPLC

Crude venom was fractionated by reverse-phase HPLC and the chromatogram revealed a total of ten peaks with varying intensities (Figure 5.1). The relative proportions of proteins in each peak (in percent) are shown in table 5.1. It was observed that the lowest amount of protein was eluted in peak 5 i.e., 2.22%, corresponding to 22.2 µg and peak 2 constituted majority of venom proteins owing to 26.83% of total chromatogram.

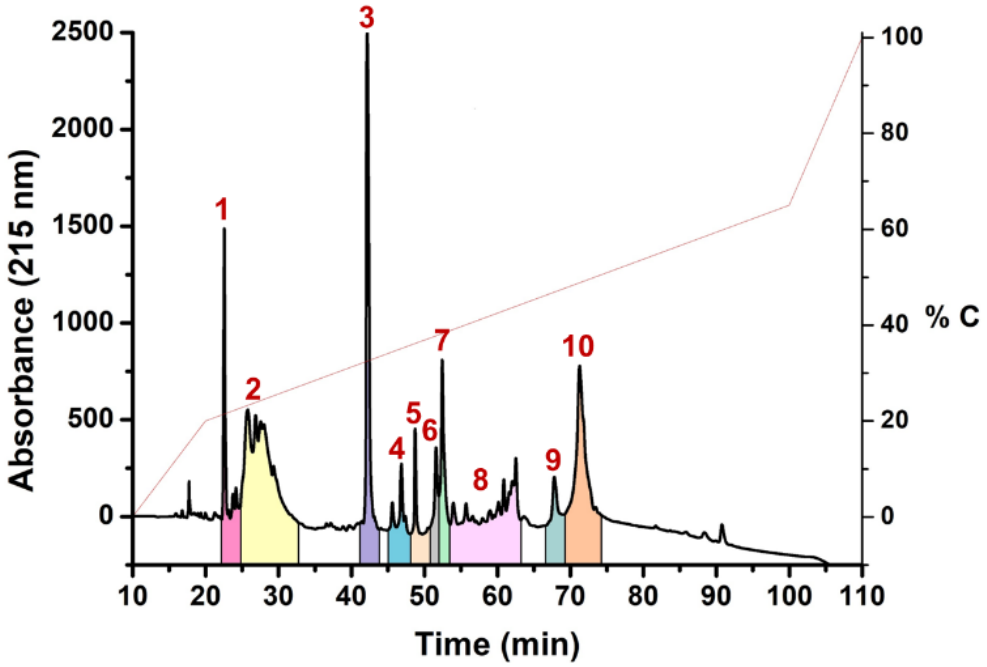


Figure 5.1. Reverse phase chromatogram of crude venom of *Trimeresurus erythrurus*. Crude venom (1 mg) was loaded on Symmetry C18 column and fractionation was carried with a linear gradient of 20-65% acetonitrile containing 0.1% TFA.

Table 5.1. Amount of protein obtained in each Rp-HPLC fraction of *T. erythrurus*.

Rp-HPLC fractions	Peak area (%)	Amount of protein (in µg)
Peak 1	6.89	68.9
Peak 2	26.82	268.2
Peak 3	13.69	136.9
Peak 4	3.86	38.6
Peak 5	2.22	22.2
Peak 6	3.02	30.2
Peak 7	10.18	101.8
Peak 8	10.37	103.7
Peak 9	4.94	49.4
Peak 10	18.01	180.1

5.2.2. SDS-PAGE analysis of Rp-HPLC fractions

The electrophoretic profile of Rp-HPLC fractions revealed presence of bands ranging from ~10-150 kDa corresponding to presence of various snake venom enzymes and toxins (Figure 5.2). Peak 3, 4 and 5 showed prominent band at 10 kDa and 15 kDa owing to the presence of low molecular weight proteins, however, peak 1 and peak 2 did

not show any band. Likewise, Peak 6 and 7 display presence of prominent bands at 25 kDa and 37 kDa which are the typical range of snake venom serine proteases. Moreover, Peak 8, 9 and 10 shows various protein bands in the range of 10 kDa to 150 kDa suggesting presence of a mixture of proteins belonging to different protein families.

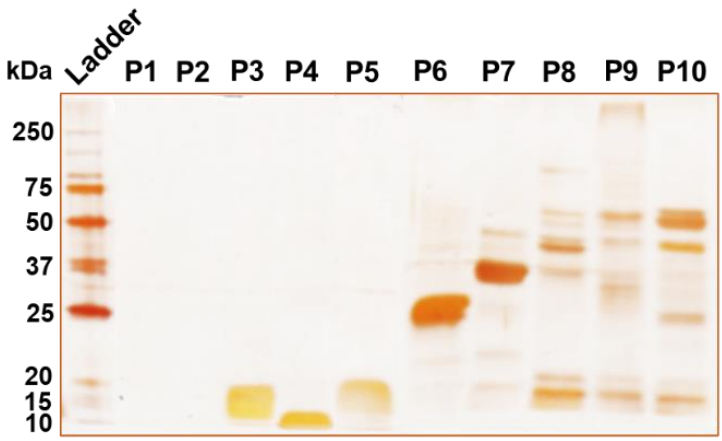


Figure 5.2. Electrophoretic profile of Rp-HPLC fractions of *Trimeresurus erythrurus*. Lane 1 contains Precision plus protein standard (Bio-Rad, USA). Lane 2-11 each contains 2.5 µg of Rp-HPLC peak 1-10 in reduced condition.

5.2.3. Protein decomplexation by ESI-LC-MS/MS

Proteome decomplexation of *Trimeresurus erythrurus* was performed using a combined approach of chromatography and mass spectrometry. The identity of proteins in each Rp-HPLC fraction along with their corresponding snake venom family, was revealed by in-solution tryptic digestion followed by ESI-LC-MS/MS of digested fragments. Analysis of the peptide fragments generated by MS/MS provided a comprehensive outline of different peptides/proteins existing in crude venom (Figure 5.3). A detailed summary of various peptide fragments detected in each Rp-HPLC peak assigned to various snake venom protein families is shown in table 5.2. Based on sequence homology, a total of 53 putative proteins/peptides belonging to 10 different venom protein families were identified viz. phospholipase A_2 (PLA_2s), snake venom serine proteases (SVSPs), snake venom metalloproteases (SVMPs), snaclecs, l-amino acid oxidases (LAAO), disintegrins, cysteine rich secretory proteins (CRiSPs), c-type lectins, c-type natriuretic peptides and glutaminyl-peptide cyclotransferases. The relative distribution of these venom protein families is shown in Figure 5.4. Pairwise sequence alignment of various peptide fragments obtained from ESI-LC-MS/MS with their homologous protein have been shown in Appendix I.

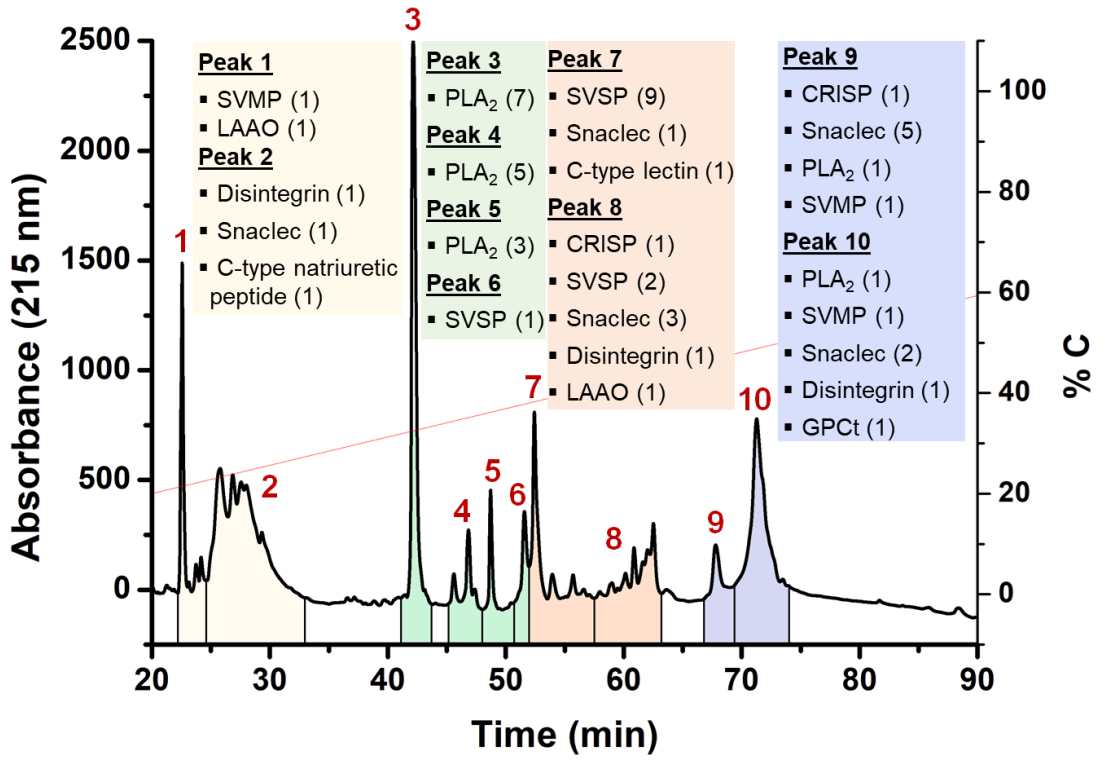


Figure 5.3. Schematic representation of various snake venom protein families identified by ESI-LC-MS/MS in each Rp-HPLC fraction of *Trimeresurus erythrurus* venom. SVMPs: snake venom metalloproteases, SVSPs: snake venom serine proteases, LAAO: L-amino acid oxidases, PLA_2: phospholipase A_2, CRISPs: cysteine rich secretory proteins, GPCts: glutaminyl-peptide cyclotransferases.

Phospholipase A_2 (PLA_2) is the most abundant protein family constituting 32% of the total proteome of *Trimeresurus erythrurus* (Figure 5.4). Rp-HPLC peaks 3, 4 and 5 were exclusively comprised of PLA_2s with 7, 5 and 3 isoforms respectively (Table 5.2) which is noticeable as 10-20 kDa bands in the electrophoretic profile of each of these peaks (Figure 5.2). One PLA_2 was also found in each of peaks 9 and 10. The PLA_2 isoform obtained in the proteome show sequence similarity with various acidic and basic PLA_2s of other Crotalinae snakes. Apart from presence of group I PLA_2s containing Asp49 functional group, two basic PLA_2 isoforms (AAP48893.1 and AAP48895.1) with Asn49 mutation in the Ca^{2+} binding site were also found. PLA_2 isoforms with modification in disulfide bonds (AAR14165.1 and P0DJP4.1) and additional substitution at functional sites have also been observed. Snake venom serine proteases (SVSPs) are the second most abundant group of proteins in *T. erythrurus* venom, constituting ~22% of the total venom protein (Figure 5.4). A total of twelve SVSPs were observed in the crude venom of *T. erythrurus* (Table 5.2) which were eluted in peak 6, 7 and 8. Many of them show similarity with proteins belonging to thrombin-like snake venom serine

proteases (TL-SVSPs), a sub-family of SVSPs. (A7LAC6.1, A7LAC7.1, POCJ41.1, POCJ41.1) which are well known for their role in consumptive coagulopathy in patients. A few of the peptides identified in the proteome displays substitution (AAQ02894.1) mutations with respect to their assigned proteins. The most abundant non-enzymatic protein family obtained in *T. erythrurus* venom was snaclecs. A total of 10 snaclecs were identified, constituting about 19% of the total proteome (Figure 5.4) and distributed in peaks 2, 7, 8, 9 and 10 (Figure 5.3, Table 5.2). Peptide fragments showing sequence similarity with α and β chain of alboaggregin A and B, stejaggregin A and purpureotin reported from *T. albolabris*, *Viridovipera stejnegeri* and *T. purpureomaculatus* respectively were observed in the proteome. The PLA_2, serine proteases and snaclecs together constitute about 72% of total proteome of *T. erythrurus*. Apart from them, snake venom metalloprotease (SVMPs) constitutes about 9% of total venom proteins identified which were eluted in peak 1, 9 and 10 (Figure 5.3). Other minor protein families detected in the venom belong to disintegrins (3.70%), L-amino acid oxidases (3.7%), Cysteine rich secretory proteins (3.70%) (Figure 5.4). Only one protein each belonging to c-type lectins, c-type natriuretic peptides and glutamyl cyclotransferases were identified in the venom which were eluted in peak 7, peak 2 and peak 10 respectively.

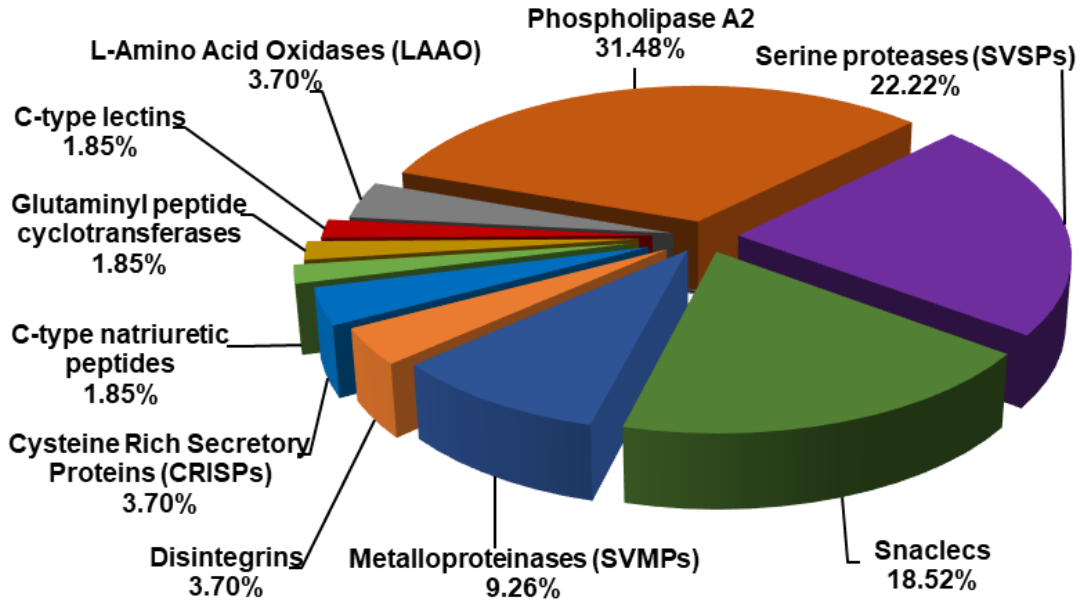


Figure 5.4. **Relative abundance of different venom protein families in the proteome of *Trimeresurus erythrurus*.** SVMPs: snake venom metalloproteases, SVSPs: snake venom serine proteases, LAAO: L-amino acid oxidases, PLA_2: phospholipase A_2, CRISPs: cysteine rich secretory proteins, GPCts: glutaminyl-peptide cyclotransferases

Table 5.2. Summary of peptide fragments identified by ESI-LC-MS/MS analysis of *Trimeresurus erythrurus* venom fractions

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.

Peak No.	Peptide Sequence	No. of peptides	m/z	Z	MH+[Da]	Sequest score	Coverage %	Unique peptides	Protein families	Homology match
1	NMPQCILK	1	502.26	2	1003.51	14.91	12	7	SVMPs	*Viridovipera stejnegeri* Q2LD49.1
	QGAQCAEGLCCDQCR	2	906.85	2	1812.69					
	SAECTDRFQR	1	424.19	3	1270.55					
	LYCIDSSPANKNPCNIVYLPNDEEK	1	985.13	3	2953.37					
	LYCIDSSPANK	1	634.30	2	1267.60					
	IIVQSVPAVTLK	1	634.40	2	1267.80					
	LRQGAQCAEGLCCDQCR	2	694.63	3	2081.87					
1	DPGVLKYPVKPSEEGK	1	581.65	3	1742.93	5.79	5	3	LAAO	*V. stejnegeri* AAQ16182.1
	KDPGVLKYPVKPSEEGK	1	468.51	4	1871.03					
	IFLTCTK	1	441.74	2	882.48					
2	LLPGAQCGEGLCCDQCSFMK	3	1165.99	2	2330.97	138.42	93	8	Disintegrin	*T. purpureomaculatus* QJA41976.1
	LLPGAQCGEGLCCDQCSFMKK	4	820.36	3	2459.06					
	RARGDDLDDYCNGISAGCPR	1	567.75	4	2267.99					
	ARGDDLDDYCNGISAGCPR	2	1056.45	2	2111.89					
	ARGDDLDDYCNGISAGCPRNPLHA	2	882.06	3	2644.16					
	EAGEDCDCGSPANPCCNAATCK	2	815.29	3	2443.87					
	GDDLDDYCNGISAGCPR	2	942.88	2	1884.75					
	GDDLDDYCNGISAGCPRNPLHA	2	806.35	3	2417.02					

Peak No.	Peptide Sequence	No. of peptides	m/z	Z	MH+[Da]	Sequest score	Coverage %	Unique peptides	Protein families	Homology match
2	EELSLGPEAASGPAAPQR	1	1780.87	3	594.29	2.14	21.9	2	Natriuretic peptide	*Craspidocephalus gramineus* P0C7P6.1
	GCFGLPLDR	1	1034.51	2	517.76					
3	AAAICFR	1	404.71	2	808.41	83.77	54	2	PLA_2	*V. stejnegeri* 4RFP
	YSSNNGDIVCEANNPCTK	1	682.29	3	2044.82					
	CCFVHDCCYGKVNGCDPK	2	569.97	4	2276.87					
	CCFVHDCCYGK	1	1505.54	1	1505.54					
	YWNIPMESCQESEPC	1	965.37	2	1929.75					
	EICECDKAAAICFR	2	872.38	2	1743.75					
3	MIFQETGK	3	969.47	1	969.47	149.12	30	2	PLA_2	*V. stejnegeri* AAP48893.1
	KLTDCDPIKDR	1	454.23	3	1360.69					
	AVAICFRENLDTYDK	1	605.63	3	1814.87					
	AVAICFR	2	419.22	2	837.43					
	ENLDTYDKK	1	563.27	2	1125.54					
	ENLDTYDK	2	998.43	1	998.43					
	EMCECDKAVAICFRENLDTYDK	3	696.55	4	2783.18					
	EMCECDKAVAICFR	4	602.58	3	1805.73					
3	KLTDCDPIKDR	1	454.23	3	1360.69	29.33	35	2	PLA_2	*V. stejnegeri* AAP48895.1
	MFVQEMGKNALTSYSLYGCNCGPGGR	1	977.43	3	2930.26					
	FRENLDTYDK	1	650.81	2	1300.62					
	ENLDTYDKK	1	563.27	2	1125.54					
	ENLDTYDK	2	998.43	1	998.43					

Peak No.	Peptide Sequence	No. of peptides	m/z	Z	MH+[Da]	Sequest score	Coverage %	Unique peptides	Protein families	Homology match
3	YSYSWENK	2	539.23	2	1077.45	125.78	17	2	PLA_2	*C. puniceus* AAR14165.1
	AVAICFR	1	419.22	2	837.43					
	DRYSYSWENK	1	449.87	3	1347.60					
	EMCECDKAVAICFR	4	895.37	2	1789.74					
3	AAAICFR	2	404.71	2	808.41	87.36	43	4	PLA_2	*Popeia popeiorum* AHJ09513.1
	MIVQEMGKNALTSYSLYGCNCGPGGR	7	961.09	3	2881.26					
	MIVQEMGKNALTSYSLYGCNCGPGGRR	2	756.10	4	3021.37					
	EVCECDKAAAICFR	1	576.92	3	1728.75					
	CCFVHDCCYGKVNGCNPK	3	569.97	4	2276.87					
	CCFVHDCCYGK	1	1505.54	1	1505.54					
3	MIVQEMGKNALTSYSLYGCNCGVGGR	2	956.75	3	2868.27	30.54	33	2	PLA_2	*P. Sabahi* AHJ09541.1
	NPHLKELCECDKAVAICFR	1	590.79	4	2360.13					
	AVAICFR	2	419.22	2	837.43					
3	ENLDTYDKK	1	563.27	2	1125.54	23.96	35	2	PLA_2	*T. albolabris* AHJ09518.1
	ENLDTYDK	2	998.43	1	998.43					
	EMCECDKAMAICFR	1	607.58	3	1820.73					
	MIVQEMGKNALTSYSLYGCNCGPGR	1	936.76	3	2808.25					
4	AAAICFR	2	404.71	2	808.41	142.59	35	2	PLA_2	*T. venustus* AHJ09519.1
	MDFYRYSEENGGIVCEANNPCTKEICECDK	1	926.87	4	3704.51					
	YSEENGGIVCEANNPCTK	1	1022.91	2	2044.81					
	CCFVHDCCYGR	2	512.18	3	1534.53					
	EICECDKAAAICFR	2	581.92	3	1743.75					

Peak No.	Peptide Sequence	No. of peptides	m/z	Z	MH+[Da]	Sequest score	Coverage %	Unique peptides	Protein families	Homology match
4	AAAICFR	2	404.71	2	808.41	77.7	40	2	PLA_2	*Parias malcolmi* AHJ09543.1
	YSSENEDIVCEANNPCTK	1	1066.42	2	2131.84					
	VNGCDPKMDFYKYSSENEDIVCEANNPCTK	1	1201.82	3	3603.45					
	CCFVHDCCYGK	1	502.52	3	1505.54					
	EICECDKAAAICFR	2	581.92	3	1743.75					
	CCFVHDCCYGKVNGCDPK	1	759.30	3	2275.88					
4	AAAICFR	2	404.71	2	808.41	154.77	47	2	PLA_2	*T. cardamomensis* AHJ09586.1
	YSEENGDIVCEANNPCTK	2	1051.91	2	2102.82					
	VNGCGPKMDFYRYSEENGDIVCEANNPCTK	2	886.12	4	3541.48					
	CCFVHDCCYGR	2	512.18	3	1534.53					
	EICECDKAAAICFR	2	581.92	3	1743.75					
	DNINTYDNK	1	548.75	2	1096.49					
4	SGIWWYGSYGCYCGK	1	922.38	2	1843.76	194.91	78	6	PLA_2	*T. cardamomensis* AHJ09590.1
	YCKEESEPC	1	601.23	2	1201.45					
	VNGCDPKDDFYTYREENGNIVCEEDNPCTK	2	1214.49	3	3641.46					
	GGQDRPQDASDRCCFVHDCCYGR	1	564.03	5	2816.11					
	CCFVHDCCYGRVNGCDPK	1	576.97	4	2304.87					
	CCFVHDCCYGR	2	512.18	3	1534.53					
	DDFYTYREENGNIVCEEDNPCTK	1	957.38	3	2870.14					
	EICECDKDAAICFR	1	893.89	2	1786.76					
	YWFYPAK	1	487.74	2	974.48					

Peak No.	Peptide Sequence	No. of peptides	m/z	Z	MH+[Da]	Sequest score	Coverage %	Unique peptides	Protein families	Homology match
	EENGNIVCEEDNPCTK	1	954.88	2	1908.76					
	DNINTYDNKYWFYPAK	1	1026.47	2	2051.95					
	DNINTYDNK	1	548.75	2	1096.49					
	DAAICFRDNINTYDNKYWFYPAK	1	722.10	4	2885.34					
4	YSSDNNDIVCGGNNPCLKEICECDRDAAICFR	1	956.65	4	3823.59					
	CCFVHDCCYGK	1	502.52	3	1505.54	5.12	35	2	PLA_2	*C. puniceus* P0DJP4.1
	YSSDNNDIVCGGNNPCLK	1	676.94	3	2028.83					
5	AAAICFR	2	404.71	2	808.41					
	CCFVHDCCYGK	1	502.52	3	1505.54	21.51	25	2	PLA_2	*Pa. hageni* AHJ09535.1
	AAAICFRDNVGTYDRK	1	619.64	3	1856.91					
	EVCECDKAAAICFR	1	576.92	3	1728.75					
5	AAAICFR	2	404.71	2	808.41					
	VNGCDPKDDFYK	2	486.88	3	1458.62					
	VNGCDPKDDFYKYSEENGDIVCEEDNPCTK	4	900.12	4	3597.47					
	YSEENGDIVCEEDNPCTKEICECDKAAAICFR	2	971.65	4	3883.59					
	YSEENGDIVCEEDNPCTKEICECDK	2	1032.07	3	3094.19					
	YSEENGDIVCEEDNPCTK	3	1080.92	2	2160.82	487.66	64	2	PLA_2	*T. erythrurus* AHJ09546.1
	YCKEESEPC	1	601.23	2	1201.45					
	DNIETYQNKYWSYPAK	1	674.32	3	2020.93					
	DNIETYQNK	2	563.26	2	1125.51					
	DDFYKYSEENGDIVCEEDNPCTKEICECDK	1	941.13	4	3761.49					
	DDFYKYSEENGDIVCEEDNPCTK	2	943.38	3	2828.12					

Peak No.	Peptide Sequence	No. of peptides	m/z	Z	MH+[Da]	Sequest score	Coverage %	Unique peptides	Protein families	Homology match
	CCFVHDCCYGRVNGCDPKDDFYK	1	744.04	4	2973.15					
	CCFVHDCCYGRVNGCDPK	1	576.97	4	2304.87					
	CCFVHDCCYGR	2	767.77	2	1534.53					
	AAAICFRDNIETYQNK	1	957.46	2	1913.92					
	EICECDKAAAICFRDNIETYQNK	1	950.09	3	2848.27					
	EICECDKAAAICFR	2	872.38	2	1743.75					
	GRPQDASDRCCFVHDCCYGR	1	629.76	4	2516.01					
5	AAAICFR	2	404.71	2	808.41	653.22	78	6	PLA_2	*T. albolabris* AHJ09577.1
	VNGCDPKDDFYK	2	486.88	3	1458.62					
	SGIWWYGSYGCYCGKGGQGRPQDASDR	1	767.83	4	3068.32					
	SGIWWYGSYGCYCGK	1	922.39	2	1843.76					
	VNGCDPKDDFYKYSEENGDIVCEEDNPCTK	4	900.12	4	3597.47					
	YWFYPAK	1	974.48	1	974.48					
	YSEENGDIVCEEDNPCTKEICECDKAAAICFR	2	971.65	4	3883.59					
	YSEENGDIVCEEDNPCTKEICECDK	2	1032.07	3	3094.19					
	YSEENGDIVCEEDNPCTK	3	1080.92	2	2160.82					
	YCKEESEPC	1	601.23	2	1201.45					
	DNIETYQNKYWFYPAKYCK	1	633.55	4	2531.17					
	DNIETYQNKYWFYPAK	2	1040.99	2	2080.97					
	DNIETYQNK	2	563.26	2	1125.51					
	DDFYKYSEENGDIVCEEDNPCTKEICECDK	1	941.13	4	3761.49					
	DDFYKYSEENGDIVCEEDNPCTK	2	943.38	3	2828.12					

Peak No.	Peptide Sequence	No. of peptides	m/z	Z	MH+[Da]	Sequest score	Coverage %	Unique peptides	Protein families	Homology match
	CCFVHDCCYGRVNGCDPKDDFYK	1	744.04	4	2973.15					
	CCFVHDCCYGRVNGCDPK	1	576.97	4	2304.87					
	CCFVHDCCYGR	2	767.77	2	1534.53					
	AAAICFRDNIETYQNKYWFYPAK	2	957.47	3	2870.36					
	AAAICFRDNIETYQNK	1	957.46	2	1913.92					
	GGQGRPQDASDRCCFVHDCCYGR	2	690.53	4	2759.09					
	GGQGRPQDASDR	1	415.20	3	1243.58					
	EICECDKAAAICFRDNIETYQNK	1	950.09	3	2848.27					
	EICECDKAAAICFR	2	872.38	2	1743.75					
	YWFYPAKYCKEESEPC	1	1078.96	2	2156.91					
6	TAYSWR	1	783.38	1	783.38	111.51	30	12	SVSPs	*T. albolabris* P0DJF5.1
	TLNEDEQTR	1	553.26	2	1105.51					
	TLNEDEQTRDPK	2	723.34	2	1445.69					
	TIPTKDIYPDVPHCANINILDHAVCR	2	758.88	4	3032.51					
	EKFFCPNR	1	549.27	2	1097.52					
	EKFFCPNRK	1	409.21	3	1225.61					
	DATCPP	1	660.26	1	660.27					
	DDEVDKDIMLIK	1	717.36	2	1433.72					
	DIMLIK	1	732.43	1	732.43					
	DIYPDVPHCANINILDHAVCR	1	831.40	3	2492.18					
	IMGWGK	2	691.36	1	691.36					
	FFCPNR	2	420.70	2	840.38					

Peak No.	Peptide Sequence	No. of peptides	m/z	Z	MH+[Da]	Sequest score	Coverage %	Unique peptides	Protein families	Homology match
	FFCPNRK	1	484.74	2	968.48					
	KDDEVDKDIMLIK	1	521.28	3	1561.81					
7	RDSGGPLICNGQLQGVVFWGPKPCAQPR	1	774.64	4	3095.53					
	SRTLCAGVLQGGIDTCK	1	612.64	3	1835.91					
	SRTLCAGVLQGGIDTCKR	2	498.76	4	1992.01					
	NVQFDDEQRRYPK	2	565.94	3	1695.81					
	NVQFDDEQRR	1	436.21	3	1306.61					
	NMYIYLGMHNK	4	692.82	2	1384.64					
	NSEHIAPLSLPSSPPSVGSVCR	2	1146.08	2	2291.14					
	NVQFDDEQR	2	1150.51	1	1150.51					
	VFNHLDWIQSIIAGNTTVTCPP	1	828.75	3	2484.22					
	WDKDIMLIR	2	595.32	2	1189.64	356.82	64	6	SVSPs	*T. albolabris* A7LAC6.1
	TLCAGVLQGGIDTCK	1	796.89	2	1592.78					
	TLCAGVLQGGIDTCKR	2	583.63	3	1748.88					
	DSGGPLICNGQLQGVVFWGPKPCAQPR	3	980.15	3	2938.45					
	EYVLTAAHCETR	2	483.90	3	1449.68					
	CANINLLNYTVCR	2	805.89	2	1610.78					
	DIMLIR	1	760.44	1	760.44					
	LNRPVRNSEHIAPLSLPSSPPSVGSVCR	1	1010.51	3	3029.55					
	KKYFFR	1	444.76	2	888.51					
	CSNNFTRWDKDIMLIR	2	518.01	4	2069.01					

Peak No.	Peptide Sequence	No. of peptides	m/z	Z	MH+[Da]	Sequest score	Coverage %	Unique peptides	Protein families	Homology match
7	RDSGGPLICNGKLQGVVFWGPKPCAQPR	1	775.14	4	3097.53					
	SRTLCAGVLQGGIDTCK	1	612.64	3	1835.91					
	SRTLCAGVLQGGIDTCKR	2	498.76	4	1992.01					
	NMYIYLGMHNK	4	1383.65	1	1383.65					
	NSEHIAPLSLPSSPPSVGSVCR	2	1146.08	2	2291.14					
	WDKDIMLIR	2	595.32	2	1189.64					
	VFDHLDWIQSIIAGNTTVTCPP	1	828.75	3	2484.22					
	TLCAGVLQGGIDTCK	1	796.89	2	1592.78					
	TLCAGVLQGGIDTCKR	2	583.63	3	1748.88	248.81	59	4	SVSPs	*T. albolabris* A7LAC7.1
	DSGGPLICNGKLQGVVFWGPKPCAQPR	2	980.82	3	2940.45					
	EYVLTAAHCETR	2	483.90	3	1449.68					
	LQGVVFWGPKPCAQPR	1	613.99	3	1839.97					
	CANINLLNYTVCR	2	805.89	2	1610.78					
	DIMLIR	1	760.44	1	760.44					
	LNRPVRNSEHIAPLSLPSSPPSVGSVCR	1	1010.51	3	3029.55					
	KKYFFR	1	444.76	2	888.51					
	CSNNFTRWDKDIMLIR	2	518.01	4	2069.01					
7	SRTLCAGILEGGK	1	681.36	2	1361.72					
	NSAHIEPLSLPSSPPSVGSVCR	2	764.38	3	2291.14					
	TLCAGILEGGK	1	559.80	2	1118.59	20.39	32	3	SVSPs	*V. stejnegeri* AAQ02894.1
	WNKDIMLIK	1	580.83	2	1160.65					
	LQFGLHSK	1	465.26	2	929.52					

Peak No.	Peptide Sequence	No. of peptides	m/z	Z	MH+[Da]	Sequest score	Coverage %	Unique peptides	Protein families	Homology match
	DIMLIK	2	732.43	1	732.43					
	FFCESNK	1	466.20	2	931.40					
	GSGTLINPEWVLTAAHCETEEMK	2	858.41	3	2573.20					
7	SRTLCAGILEGGK	1	681.36	2	1361.72	34.83	32	2	SVSPs	*V. stejnegeri* AAQ02895.1
	NSAHIEPLSLPSSPPSVGSVCR	2	764.38	3	2291.14					
	TLCAGILEGGK	1	559.80	2	1118.59					
	WNKDIMLIK	1	580.83	2	1160.65					
	LQFGLHSK	1	465.26	2	929.52					
	DIMLIK	2	732.43	1	732.43					
	FFCESNK	1	466.20	2	931.40					
	GSGTLINPEWVLTAAHCETEEMK	2	858.41	3	2573.20					
7	NPFVCKFPPQC	1	697.32	2	1393.64	17.96	16	4	Snaclec	*V. stejnegeri* AAQ15166.1
	WDNTDCQAK	1	569.24	2	1137.46					
	WTNFLK	1	808.43	1	808.44					
	WTNFLKWDNTDCQAK	1	963.94	2	1926.88					
7	TLCAGIVQGGK	2	1103.59	1	1103.59	36.57	29	3	SVSPs	*C. gramineus* BAA19979.1
	EKFICPNK	1	518.27	2	1035.53					
	DIMLIK	2	732.43	1	732.43					
	IMGWGSITPTK	2	595.81	2	1190.62					
	IMGWGSITPTKVTYPDVPYCANINLLDDAECKPGYPELLPEYR	2	1240.09	4	4957.38					
	KNNEVLDKDIMLIK	1	563.98	3	1689.91					

Peak No.	Peptide Sequence	No. of peptides	m/z	Z	MH+[Da]	Sequest score	Coverage %	Unique peptides	Protein families	Homology match
7	SIIAGNKDATCPP	1	672.33	2	1343.66	21.3	22	3	SVSPs	*C. gramineus* O13060.1
	TLCAGILEGGK	1	559.80	2	1118.59					
	WNKDIMLIR	4	594.83	2	1188.66					
	FFCLSSK	1	444.72	2	888.43					
	DIMLIR	1	760.44	1	760.44					
	FHCGGTLINQEWVLSAAR	1	687.01	3	2059.02					
7	SNEILDK	1	409.72	2	818.43	294.83	69	15	SVSPs	*T. albolabris* P0CJ41.1
	SNEILDKDIMLIK	2	766.42	2	1531.84					
	SNEILDKDIMLIKLDSPVSNSAHIAPLSLPSSPPSVGSVCR	2	869.66	5	4344.26					
	TLCAGIVQGGK	2	1103.59	1	1103.59					
	TLCAGIVQGGKDTCGGDSGGPLICNEK	2	1382.64	2	2764.27					
	DTCGGDSGGPLICNEK	2	840.36	2	1679.70					
	EKFICPNK	1	518.27	2	1035.53					
	DIMLIK	2	732.43	1	732.43					
	DIMLIKLDSPVSNSAHIAPLSLPSSPPSVGSVCR	2	886.97	4	3544.85					
	LDSPVSNSAHIAPLSLPSSPPSVGSVCR	2	944.49	3	2831.44					
	LHGIVSYGGHPCGQSHKPGIYTNVFDYNDWIQSIIAGNTDATCLS	1	993.47	5	4963.32					
	LLNEDEQIR	2	565.30	2	1129.58					
	LLNEDEQIRNPK	2	490.26	3	1468.78					
	LLNEDEQIRNPKEK	1	432.23	4	1725.91					

Peak No.	Peptide Sequence	No. of peptides	m/z	Z	MH+[Da]	Sequest score	Coverage %	Unique peptides	Protein families	Homology match
	IMGWGSTTPIEVTYPDVPYCANINLLDDAECKPGYPELLPEYR	3	1240.09	4	4957.34					
	KSNEILDKDIMLIK	2	419.74	4	1675.93					
	KLLNEDEQIRNPK	2	533.29	3	1597.85					
	KLLNEDEQIR	1	419.90	3	1257.68					
7	WDKDIMLIR	2	595.32	2	1189.64	29.67	27	6	SVSPs	*T. albolabris* P0DJF6.1
	YFCLSSNNDTKWDKDIMLIR	1	630.56	4	2519.21					
	TVTCPP	1	674.32	1	674.32					
	IVCAGILR	1	451.27	2	901.53					
	DIMLIR	1	760.44	1	760.44					
	GDSGGPLICNAQLQGIVSAGGDPCAQPR	2	932.45	3	2795.32					
	GGKGSCKGDSGGPLICNAQLQGIVSAGGDPCAQPR	1	868.17	4	3469.64					
	GSCKGDSGGPLICNAQLQGIVSAGGDPCAQPR	2	1076.51	3	3227.50					
7	VPAIYTK	1	791.46	1	791.47	30.34	39	3	SVSPs	*V. stejnegeri* Q71QJ0.1
	TLCAGILQGGK	1	559.80	2	1118.59					
	LQFGLHSK	1	465.26	2	929.52					
	EIYPDVPHCANINILDHAVCR	1	627.31	4	2506.20					
	FFCLSSK	1	444.72	2	888.43					
	AFYPGLLEK	1	519.29	2	1037.57					
	DICQGDSGGPLICNGQIQGIVSVGGDPCAEPR	1	1110.50	3	3329.47					
	DIMLIR	1	760.44	1	760.44					

Peak No.	Peptide Sequence	No. of peptides	m/z	Z	MH+[Da]	Sequest score	Coverage %	Unique peptides	Protein families	Homology match
7	SCTDYLNWDK	1	651.28	2	1301.55	62.71	52	11	C- lectins	*V. stejnegeri* Q9YGP1.1
	SCTDYLNWDKNQPDHYKDK	1	810.35	3	2429.04					
	WNDQVCESK	2	583.25	2	1165.49					
	YKPGCHLASFHR	1	491.58	3	1472.72					
	EFCVELVSLTGYHR	1	570.62	3	1709.83					
	DFSWEWTDR	1	1241.52	1	1241.52					
	KYKPGCHLASFHR	1	400.96	4	1600.82					
	IFDEPKTWEDAEMFCR	2	691.98	3	2073.90					
	IFDEPKTWEDAEMFCRK	2	551.26	4	2202.00					
	KDFSWEWTDR	1	685.31	2	1369.62					
	KKDFSWEWTDR	1	749.36	2	1497.71					
8	SVNPTASNMLR	1	595.30	2	1189.60	4.3	12	4	CRISPs	*V. stejnegeri* ACE73572.1
	RSVNPTASNMLR	1	449.24	3	1345.70					
	MEWYPEAADNAER	1	791.34	2	1581.66					
	MEWYPEAADNAERWAYR	1	719.99	3	2157.94					
8	TLCAGVLQGGTDTCNR	2	861.90	2	1722.79	10.3	18	3	SVSPs	*C. gramineus* BAA19981.1
	WDKDIMLIK	1	581.32	2	1161.63					
	VFDHLDWIQSIIAGNTDAACPP	1	814.06	3	2440.16					
	DIMLIK	1	732.43	1	732.43					
8	TLCAGILQGGK	1	559.80	2	1118.59	7.88	8	2	SVSPs	*C. gramineus* O13062.1
	AFYPGLLEK	1	519.29	2	1037.57					

Peak No.	Peptide Sequence	No. of peptides	m/z	Z	MH+[Da]	Sequest score	Coverage %	Unique peptides	Protein families	Homology match
8	QLKTWEDAER	2	425.88	3	1275.63	60.45	80	13	Snaclec	*T. purpureomaculatus* P0DJL2.1
	QLKTWEDAERFCLDQMK	1	550.51	4	2199.02					
	QYCYQIIK	1	558.28	2	1115.56					
	TWEDAER	1	453.70	2	906.40					
	TWEDAERFCLDQMK	2	914.91	2	1828.80					
	TWEDAERFCLDQMKGAHLVSIESYR	1	612.29	5	3057.42					
	TWSNVYCEQK	2	657.79	2	1314.58					
	TWSNVYCEQKHIFMCK	1	533.50	4	2130.96					
	EAVFVAELLSENVK	1	774.42	2	1547.83					
	ESEFRTWSNVYCEQK	1	654.96	3	1962.87					
	FCLDQMK	1	941.42	1	941.42					
	FCLDQMKGAHLVSIESYR	1	543.27	4	2170.04					
	DCPSDWSSFK	1	614.75	2	1228.49					
	DCPSDWSSFKQYCYQIIK	1	775.68	3	2325.03					
	GAHLVSIESYR	1	411.22	3	1231.64					
	GQQCSSEWSDGSTVSYENLVKPNPK	2	933.10	3	2797.27					
	GQQCSSEWSDGSTVSYENLVKPNPKK	1	732.10	4	2925.37					
	HIFMCK	1	418.20	2	835.40					
8	SVGIVR	1	630.39	1	630.39	61.79	15	5	SVMPs	*T. albolabris* P0C6B6.1
	RCIELVMVADHR	3	500.26	3	1498.76					
	CIELVMVADHR	3	671.83	2	1342.66					
	LTPGSQCAEGLCCAQCK	2	970.41	2	1939.81					

Peak No.	Peptide Sequence	No. of peptides	m/z	Z	MH+[Da]	Sequest score	Coverage %	Unique peptides	Protein families	Homology match
	MAHGLGHNLGIHHDGDSCSCGANSCIMSATVSNEPSSR	1	1015.18	4	4057.66					
8	STTDLPSR	1	438.72	2	876.44	36.18	10	3	LAAO	*T. purpureomaculatus* P0DPS2.1
	NPLEECFR	1	532.75	2	1064.48					
	VTVTYQTPAK	2	554.31	2	1107.60					
	ETDYEEFLEIAR	1	757.85	2	1514.70					
	FWEDDGIHGGK	1	630.79	2	1260.56					
	KFWEDDGIHGGK	1	463.56	3	1388.66					
8	NWEDAER	2	919.39	1	919.39	100.98	65	13	Snaclec	*T. albolabris* P81112.1
	NWEDAERFCAK	2	713.32	2	1425.62					
	WINLGCIQLNPFVCK	2	931.48	2	1861.95					
	VQNKEQQCSSEWSDGSSVTYENLIK	2	973.11	3	2917.32					
	VFNEPK	1	733.39	1	733.39					
	VFNEPKNWEDAER	1	545.26	3	1633.76					
	VFNEPKNWEDAERFCAK	1	535.75	4	2139.99					
	EQQCSSEWSDGSSVTYENLIK	1	1224.04	2	2447.07					
	DFHCLPGWSAYDQYCYR	2	1120.45	2	2239.88					
	CGALEQESGFR	1	627.28	2	1253.56					
	CGALEQESGFRK	2	461.22	3	1381.65					
	KCGALEQESGFR	1	461.22	3	1381.65					
	KCGALEQESGFRK	2	504.25	3	1510.73					

Peak No.	Peptide Sequence	No. of peptides	m/z	Z	MH+[Da]	Sequest score	Coverage %	Unique peptides	Protein families	Homology match
8	SYCLVSK	1	428.72	2	856.42	47.26	30	5	Snaclec	*T. albolabris* P81113.1
	TRYPVCK	1	924.47	1	924.46					
	TTNNEWLSMDCSR	4	807.34	2	1613.67					
	ECTKEWSDGAR	1	669.79	2	1338.57					
	EWSDGAR	1	410.68	2	820.36					
9	RSVNPTASNMLR	1	449.24	3	1345.70	5.81	16	4	CRISPs	*V. stejnegeri* 1RC9
	SVNPTASNMLR	1	595.30	2	1189.60					
	MEWYPEAADNAER	1	791.34	2	1581.66					
	NVDFDSESPR	1	583.26	2	1165.51					
9	YKAWAEESYCVYFK	1	615.28	3	1843.84	19.5	10	2	Snaclec	*T. albolabris* AAB26045.1
	AWAEESYCVYFK	1	776.84	2	1552.68					
9	AAAICFR	1	404.71	2	808.41	11.24	18	2	PLA_2	*C. puniceus* AAR14167.1
	QICECDKAAAICFR	1	581.59	3	1742.77					
	CCFVHDCCYGK	1	502.52	3	1505.54					
9	MNWEDAEK	2	1022.42	1	1022.42	77.72	93	3	Snaclec	*T. purpureomaculatus* P0DJL3.1
	MNWEDAEKFCR	3	495.88	3	1485.62					
	TLPILK	1	684.47	1	684.47					
	YNAWTAESECIASK	1	815.37	2	1629.72					
	TTDNQWWTR	2	604.28	2	1207.55					
	TYPFVCK	1	457.73	2	914.44					
	TYPFVCKLEV	1	628.32	2	1255.64					
	ADFVWIGLTDVWSACR	1	948.46	2	1895.91					

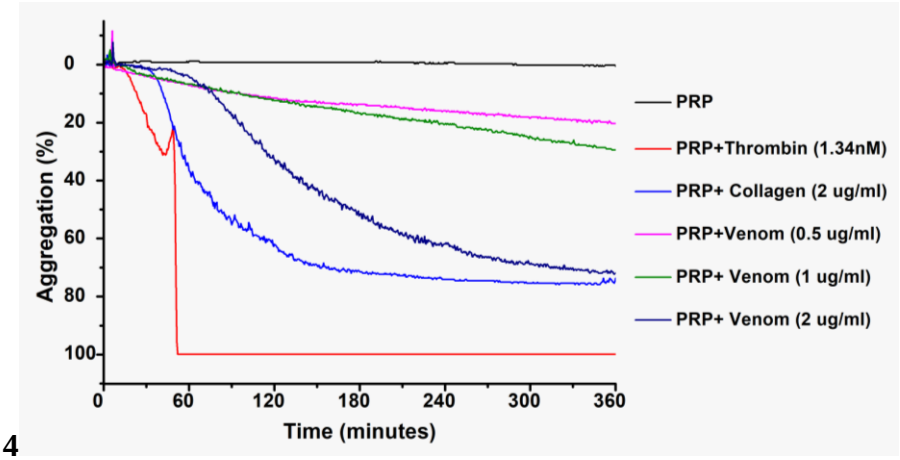
Peak No.	Peptide Sequence	No. of peptides	m/z	Z	MH+[Da]	Sequest score	Coverage %	Unique peptides	Protein families	Homology match
	DCPSDWSSYDLYCYK	1	979.88	2	1958.76					
	LQWSDGTELK	1	588.80	2	1176.59					
	LQWSDGTELKYNAWTAESECIASK	1	929.77	3	2787.29					
	QQHTGSHLLSFHSSEEVDFVVSK	1	520.46	5	2598.26					
9	LHSWVECESGECCEQCR	2	742.95	3	2226.83	29.05	6	2	SVMPs	*V. stejnegeri* ABA40759.1
	HDCAQLLTAIDFDGPTIGR	2	700.68	3	2100.02					
9	TWSNVYCGHEYPFVCK	2	1023.95	2	2046.88	21.36	28	3	Snaclec	*T. albolabris* P81111.1
	CYGLEK	1	769.35	1	769.35					
	DCPSDWSSYDQYCYR	1	1001.37	2	2001.74					
9	QYCYQIVK	2	551.27	2	1101.54	44.39	46	4	Snaclec	*T. albolabris* P81115.2
	TWEDAEK	1	439.70	2	878.39					
	TWSNVYCEQK	2	657.79	2	1314.58					
	TWSNVYCEQKHIFMCK	1	533.49	4	2130.96					
	EAVFVAELLSENVK	2	774.42	2	1547.83					
	GQQCSSEWSDGSSVSYENLVKPNPK	2	928.42	3	2783.26					
	GQQCSSEWSDGSSVSYENLVKPNPKK	1	728.59	4	2911.35					
	HIFMCK	1	418.20	2	835.40					
9	YKAWAEESYCVYFK	1	615.28	3	1843.84	16.01	10	2	Snaclec	*V. stejnegeri* Q71RR1.1
	AWAEESYCVYFK	1	776.84	2	1552.68					
10	TDIITPPVCGNELLEEGEECDCGSPENCQYQCCDAASCK	1	1142.21	4	4565.79	66.73	16	6	SVMPs	*V. stejnegeri* ABA40760.1
	YNYSDIGIVDHGTK	2	527.92	3	1581.75					

Peak No.	Peptide Sequence	No. of peptides	m/z	Z	MH+[Da]	Sequest score	Coverage %	Unique peptides	Protein families	Homology match
	AHIASMCNSK	3	559.76	2	1118.51					
	CESGECCDQCR	1	731.23	2	1461.45					
	HCVDVTTVY	1	547.25	2	1093.50					
10	AAAICFR	1	404.71	2	808.41	6.22	39	2	PLA_2	*P. popeiorum* AHJ09512.1
	SGIWWYGSYGCYCGK	1	922.39	2	1843.76					
	YCKEESEPC	1	601.23	2	1201.45					
	DNIETYQNKYWFYPAK	1	694.00	3	2079.98					
	EICECDKAAAICFR	1	581.59	3	1742.77					
10	TSHDHAQLLTATIFNGNVIGR	3	567.05	4	2265.17	72.67	6	2	Disintegrin	*C. gramineus* CAA35910.1
	CNCNTCIMSK	3	644.25	2	1287.49					
10	RHDNAQLLTGMIFNEK	1	629.66	3	1886.95	2.29	8	3	SVMPs	*V. stejnegeri* P0DM87.1
	RHDNAQLLTGMIFNEKIEGR	1	586.31	4	2342.20					
	LTPGSQCAEGLCCEQCR	1	1013.42	2	2025.82					
10	VFNEPQNWADAEK	2	775.34	2	1549.68	61.01	36	4	Snaclec	*T. albolabris* P81114.1
	VFNEPQNWADAEKFCTQQHK	3	826.38	3	2477.13					
	EQFECLVSR	2	584.28	2	1167.55					
	DCPSDWSSYEGHCYR	2	959.86	2	1918.71					
10	RHPVEDDHIPFLR	1	408.47	4	1630.84	5.54	51	15		*T. gracilis* AFE84763.1
	TFSNIISTLNPLAK	1	759.93	2	1518.85					
	TFSNIISTLNPLAKR	1	558.99	3	1674.95					
	MWQNDLHPILIER	1	555.62	3	1664.86					
	NPVFPVYFLNTAR	1	769.41	2	1537.82					

Peak No.	Peptide Sequence	No. of peptides	m/z	Z	MH+[Da]	Sequest score	Coverage %	Unique peptides	Protein families	Homology match
	VWHTMEDNEENLDKPTIDNLSK	1	657.81	4	2628.22					
	WSPSDSLYGSR	1	627.79	2	1254.57					
	YFPPQLDGK	1	532.77	2	1064.54					
	YPGSPGSYAVR	1	577.29	2	1153.56					
	LIFFDGEEAFVR	1	721.87	2	1442.73					
	LQGLQAGWLVEEDTFQSHTPYGYR	1	932.45	3	2795.34					
	GVPILHLIPSPFPR	1	514.98	3	1542.92					
	HLVIACHYDSK	1	448.22	3	1342.66					
	HPVEDDHIPFLRR	1	408.47	4	1630.84					
	VFVGATDSAVPCAMMLELAR	1	713.35	3	2138.04					

5.2.4. Platelet aggregation effect of crude venom

The crude venom of *T. erythrurus* exhibited platelet aggregation effect in a dose-dependent manner in absence of any agonist (Figure 5.5). At 2 µg/ml dose, crude venom showed 81% aggregation which was almost similar to 1 µg collagen in reaction (87%) when used as the agonist; however, it was less than 1.34 nM thrombin, which showed 100% aggregation (Figure 5.5).



4

Figure 5.5. Platelet aggregation activity of crude venom of *Trimeresurus erythrurus.* Platelet rich plasma (PRP) without any treatment was taken as negative control. PRP with 1.34 nM Thrombin and PRP with 1µg Collagen was taken as control agonist for aggregation.

5.3. Discussion

Snake venom, being a heterogenous mixture of several proteins and peptides, presents an enigma of diversified functions which were primarily evolved for efficient predation and/or defense. These proteins/peptides exhibit various pathophysiological manifestations in the envenomated victims, either independently or in association with others. However, the continuous investigation of these intricate mixture by researchers over the years has led to understanding of detailed pharmacological profile of these components. Recent advancements in proteomics technology have provided tools for more feasible and in-depth venomics studies. Understanding the detailed venom composition provides information on unique and major toxins responsible for clinical symptoms observed post envenomation in victims. Further studies, in correlation with experimental evaluation can give an insight on mechanism of action of these toxins and thereby provide useful strategies to combat clinical challenges of snake envenomation.

Protein decomplexation of *Trimeresurus erythrurus* venom was performed using proteomics approach, combining together reverse phase chromatography, electrophoresis and mass spectrometric techniques. Reverse phase chromatography of crude venom fractionated the venom into 10 peaks of varying intensities. SDS-PAGE profile of Rp-HPLC fractions showed presence of multiple protein bands in the range of 10-150 kDa revealing a mixture of various high and low molecular weight proteins. Presence of prominent band at 10-15 kDa in peak 3, 4 and 5 suggested the existence of PLA_2 isoforms. However, peak 1 and peak 2 did not show any band, probably due to presence of mixture of various low molecular weight proteins in relatively lesser amount which might not be easily detectable in gel. Likewise, peak 6 and 7 showed prominent bands in the range of 25-37kDa, owing to the presence of proteases in the venom. On the other hand, peak 8, 9 and 10 consisted of various bands ranging from 10 to 150 kDa proposing the presence of PLA_2s, snaclecs, proteases, LAAO etc., however, the identity of proteins present in each band needs further evaluation by proteomics approach.

ESI LC-MS/MS of trypsin digested fragments of Rp-HPLC peaks of crude venom provided a comprehensive outline of different proteins/peptides belonging to various snake venom protein families. Analysis of the peptide fragments generated by MS/MS based on sequence homology revealed a total of 53 proteins/peptides belonging to 10 different venom protein families viz. phospholipase A_2 (PLA_2s), snake venom serine proteases (SVSPs), snake venom metalloproteases (SVMPs), snaclecs, l-amino acid oxidases (LAAO), disintegrins, cysteine rich secretory proteins (CRiSPs), c-type lectins, c-type natriuretic peptides and glutaminyl-peptide cyclotransferases.

Phospholipase A_2 were identified to be most abundant enzymatic family in the proteome of *T. erythrurus* venom constituting approximately 32% of total venom. A total of 15 isoforms were identified which were distributed in Rp-HPLC peaks 3, 4, 5, 9 and 10. PLA_2s are the major enzymatic protein family of Viperidae snakes which are primarily responsible for immobilization and digestion of prey [64, 195]. The abundance of PLA_2s in the proteome suggests its significant role in various pathological disorders (viz. coagulopathy, inflammation, platelet aggregation inhibition, haemolysis, haemorrhage, tissue/organ damage etc.) in envenomed patients [195]. Thus, the presence of large number of PLA_2 isoforms have been attributed to the overall toxicity of crude venom. The PLA_2 isoforms present in the proteome of *T. erythrurus* show homology

with PLA_2s of other pit viper species from south-east Asian countries such as *T. albolabris* (Accession AHJ09518.1 and AHJ09577.1), *T. venustus* (Accession AHJ09519.1), *Parias hageni* (Accession AHJ09535.1), *T. erythrurus* (Accession AHJ09546.1), *Popeia popeiorum* (Accession AHJ09512.1 and AHJ09513.1), *Po. sabahi* (Accession AHJ09541.1), *Pa. malcolmi* (Accession AHJ09543.1) and *T. cardamomensis* (Accession AHJ09586.1 and AHJ09590.1). These PLA_2s were identified by Malhotra and team during the studies on the evolution of multigene toxin families [196, 197]. PA2-Vb, a dimeric acidic PLA_2 isolated from *V. stejnegeri* venom, acts as a protease-activated receptor (PAR-I) agonist to facilitate muscular contraction [212]. PAR-I receptors are members of seven transmembrane G protein-coupled receptor family which are generally activated by N-terminal proteolytic cleavage by thrombin and ultimately leads to thrombin-mediated platelet aggregation [213, 214]. Apart from their role in platelet aggregation, PAR-I mediates the contraction, cell-migration, cell-proliferation, hypertrophy and production of the extracellular matrix in the smooth muscle cells [215]. Six peptides showing similarity with chain A of this dimeric complex was identified in this proteome (Accession 4RFP). Peptide fragments similar to two basic PLA_2 isoforms Ts-K49a (Accession AAP48893.1) and Ts-K49b (Accession AAP48895.1), reported from the venom gland cDNA library of *V. stejnegeri*, were identified in the proteome [216]. These K49 isoenzymes have Asn at the 49^{th} residue instead of Asp in the Ca^{2+} binding site, hence are catalytically inactive. However, two PLA_2 isoforms similar to Tpu-K49a (Accession AAR14165.1) and Tpu-E6c (Accession P0DJP4.1) that have been reported from *Craspedocephalus puniceus* that lack the C61-C91 bond were found in this proteome [217]. The K49-PLA_2s are known to display several Ca^{2+}-independent activities such as myotoxicity, bactericidal and edema-inducing effects [218]. Three peptides similar to a basic D49-PLA_2s with a unique Gly6 substitution (Accession AAR14167.1), isolated from *C. puniceus* and showing edema-inducing and anticoagulating activities [217], were also identified in this proteome. The wide range of pathophysiological abnormalities observed in patients bitten by green pit vipers can be largely attributed to occurrence of multiple isoforms of PLA_2. Abundance of enzymatically active PLA_2s in *T. erythrurus* venom has been observed by assessment of *in vitro* catalytic activity of PLA_2s (discussed in Chapter 3). However, along with the pathophysiological activities, the non-catalytic nature of PLA_2 homologues present in

venom resulting from mutations at the calcium-binding site can be explored for diagnostic marker for arboreal pit vipers.

Snake venom serine proteases (SVSPs) were found to be the second most abundant group of proteins in *T. erythrurus* venom. A total of 12 SVSPs, constituting ~22% of the total venom protein were identified which were eluted in peak 6, 7 and 8. They are one of the major snake venom enzymatic family of viperid venom which catalyze a broad range of reactions affecting the hemostasis system of prey and are responsible for proteolytic and fibrinogenolytic activity [64]. The thrombin-like enzymes (TL-SVSPs), a subgroup of SVSPs, cleave fibrinogen to fibrin monomers leading to formation of unstable clots (initial procoagulant effects) which readily gets dissolved [150], thereby giving an overall anticoagulant effect (consumptive coagulopathy) caused by defibrination [151, 162]. This pathological effect has been frequently observed in victims of *Trimeresurus* bites [102, 107]. Interestingly, four proteins showing sequence homology with serine proteases cloned from a venom gland cDNA library of *T. albolabris* [219] were observed in the *T. erythrurus* proteome. Two of them were thrombin-like enzymes 1 and 2, TA1 (Accession A7LAC6.1) and TA2 (Accession A7LAC7.1), an alpha-fibrinogenase albofibrase (Accession POCJ41.1), which cleaves the α chain of fibrinogen, and a venom plasminogen activator GPV-PA (Accession P0DJF5.1), which activates plasminogen and is responsible for fibrinolysis of thrombin, suggesting their collective role in hypofibrinogenemia in patients. Markedly, an additional asparagine residue was observed in one of the peptide fragments assigned to thrombin-like enzyme 1 (Appendix I). A homolog of thrombin-like enzyme, chitibrisin (Accession P0DJF6.1) was also observed in the proteome which is reported from *T. albolabris,* and known to activates fibrinogen by releasing fibrinopeptides A and B [166]. The *T. erythrurus* venom profile also possesses peptide sequences showing similarity with four serine proteases (Accession BAA19979.1, O13060.1, BAA19981.1 and O13062.1) identified in the venom gland cDNA library of *Craspedocephalus gramineus*. Apart from these, homologs of serine proteases KN1 (AAQ02894.1), KN2 (Q71QJ0.1) and KN6 (AAQ02895.1) precursors reported from venom gland cDNA library of *V. stejnegeri* were observed in the proteome, however, 6 out of 8 peptide fragments assigned to KN1 precursor exhibited some dissimilarities in their amino acid composition when compared with their actual sequence (Appendix I). They are often synthesized as inactive precursor zymogens that are cleaved during proteolysis to

generate their active forms. The abundance of SVSPs, specifically TL-SVSPs in the venom, might be responsible for remarkable *in vitro* thrombin-like activity and fibrinogenolytic activity shown by *T. erythrurus*, leading to overall procoagulant nature of venom (discussed in Chapter 3).

Snaclecs constitute the most abundant family of non-enzymatic proteins with 10 isoforms constituting about 19% of the total proteome. Alboaggregin A and B are snaclecs isolated from *T. albolabris*, which in association with each other stimulate platelet agglutination and induce platelet aggregation [220]. Peptide sequences showing homology with 2α chains (Accession P81111.1 and P81112.1) and 2β chains (Accession P81113.1 and P81114.1) of alboaggregin A as well as α and β chains of alboaggregin B (Accession P81115.2 and AAB26045.1) were observed. Peptides similar to α chain of stejaggregin A (Accession AAQ15166.1) and β chain of snaclec coagulation factor IX/X binding protein (Q71RR1.1) of *V. stejnegeri,* were also identified. Peptides showing similarity with snaclec purpureotin, a heterodimer of α and β subunit (Accession P0DJL2.1 and P0DJL3.1) isolated from *T. purpureomaculatus*, was also observed in the proteome with a loss of one phenylalanine residue in one of the peptides assigned to α subunit (Appendix I). Purpureotin is a GPIb binding protein which induces platelet aggregation without any cofactor [221]. Along with snaclecs, the venom profile also contains homolog of Ca^{2+} dependent galactose binding C-type lectin, TsL (Accession Q9YGP.1), reported in the venom gland cDNA library of *V. stejnegeri* [222]. Snaclecs, being a non-enzymatic group, function predominantly by binding to platelet surface receptors and contributing to alterations in platelet functions, leading to a clinical condition of low platelet count known as thrombocytopenia in green pit viper envenomated patients [109].

Platelet aggregation takes place as a result of some injury, where firstly platelets adhere to collagen, followed by activation of platelets, induced by binding of some agonists with surface receptors. After activation, a cascade of downstream reactions takes place which leads to aggregation of platelets to form a platelet plug. Various agonists like collagen, thrombin, ADP, ristocetin, arachidonic acid etc. play a vital role in platelet aggregation via binding to their specific receptors. When tested for platelet aggregation, the crude venom of *T. erythrurus*, exhibited its affect in a dose-dependent manner without the addition of any agonist. This might be attributed to binding of some snake venom

toxins directly to platelet receptors, similar to collagen, thereby activating the aggregation cascade. Snaclecs are known to be a major group of snake venom toxins that bind to platelet receptors and alter their functions [223]. Apart from snaclecs, some PLA_2s, serine proteases and disintegrins also shows binding with platelet surface receptors to either promote or inhibit aggregation [224-226]. In thrombin-mediated platelet aggregation, thrombin binds to PAR-1 and PAR-4 receptors of platelets to initiate aggregation. PLA_2 peptides observed in venom, showing similarity with PA2-Vb, are known to bind PAR-I receptors of platelets to aid in muscular contraction, which might also secondarily lead to platelet aggregation. Thus, the effect of crude venom on *in vitro* aggregation on platelet-poor plasma is suggested to be combined effects of snake venom enzymes and toxins which bind to various platelet surface receptors to alter platelet functions. The aggregation of platelets causes consumption of functional platelets and thereby leads to low platelet count resulting in thrombocytopenia as observed in white-lipped green pit-viper envenomated patients [104, 105, 127]. Similarly, thrombocytopenia persisting up to 11 days of bite was reported in a patient from Yangon, Myanmar bitten by *T. erythrurus* [127]. Another report from Myanmar confirms low platelet count up to $150x10^9$ per litre in 10 out of 333 patients with "green snake bite" [109]. However, the frequency of thrombocytopenia is found to be low in patients bitten by green pit vipers. The reason might be attributed to the diverse venom composition of different green pit vipers as well as presence of PLA_2s, disintegins and other venom proteins which serves as platelet aggregation inhibitors, in contrast to the snaclecs.

Snake venom metalloproteases (SVMPs) constitute a large group of enzymatic proteins of Viperidae venom which cause local as well as systemic injuries in the envenomed victims. The pathophysiological abnormalities caused by SVMPs include hemorrhage, edema, myonecrosis, hypotension, blistering, dermonecrosis and coagulopathy [151, 227]. The clinical manifestations of green pit vipers are characterized by local symptoms such as extensive pain, bleeding and progressive swelling at bite site, also tissue necrosis in few [100, 102, 109]. These symptoms can be attributed to SVMPs constituting approximately 10% of *T. erythrurus* venom. SVMPs are classified into three classes (P-I, P-II and P-III) based on the structural differences in various domains. The P-I class is the most simplified and includes precursors containing a signal sequence, a pro-domain, and a metalloproteinase domain. The precursor proteins of class P-II SVMPs comprise, in addition to the domains of class I, a disintegrin domain, connected to the

metalloproteinase domain by a short spacer sequence [151, 227]. However, in many cases, the disintegrin domain gets separated by proteolytic cleavage during post-translational modification. Homolog of albolatin, a zinc dependent class P-II metalloproteinase homolog (Accession P0C6B6.1), reported from a cDNA library from *T. albolabris* venom gland, has been identified in the *T. erythrurus* proteome [228]. The disintegrin domain of albolatin can inhibit collagen-induced platelet aggregation. Sequence similarity with another class P-II protein stejnitin (Accession P0DM87.1), isolated from *V. stejnegeri*, was also observed in the *T. erythrurus* proteome. Stejnitin can degrade the Aα as well as Bβ band of fibrinogen along with inhibiting ADP-induced platelet aggregation [229]. Class P-III of SVMPs comprise an additional cysteine-rich domain along with components of class P-II in their precursor form [227]. Peptide sequences showing homology with a class P-III SVMPs, TSV-DM (Accession Q2LD49.1), a basic metalloproteinase isolated from *V. stejnegeri* was observed in the proteome. This has roles in cell proliferation inhibition and induction of cell morphology, along with fibrinogenolytic activity [230]. Also, proteins having sequence similarity with class P-III metalloprotease, stejnihagin-A (Accession no- ABA407601.1) and stejnihagin-B (ABA40759.1), reported from a cDNA library of *V. stejnegeri* venom gland, was identified. Stejnihagin A and B are known to have a cysteine residue at position 100 of metalloprotease domain (features of class IIIc) and function to inhibit L-type Ca^{2+} channels [231, 232].

Disintegrins, cysteine-rich polypeptides which are generated by the proteolytic cleavage of multidomain metalloproteinases [233], comprise 3.7% of the venom. Peptide fragments homologous to Trigramin (Accession CAA35910.1), a platelet aggregation inhibitor isolated from *Crespedocephalus gramineus*, has also been identified in the proteome. Trigramin is a cysteine rich single chain peptide (approx. 9 kDa) containing a RGD motif which is generated by post-translational processing of a metalloprotease precursor protein [226, 234]. It is a highly specific competitive inhibitor of fibrinogen binding to platelet receptors, leading to inhibition of fibrinogen-induced aggregation of platelets stimulated by ADP [226]. Purpureomaculin homolog (Accession QJA41976.1), a disintegrin isolated from *T. purpuromaculatus*, also having a RGD motif, was also observed in the venom. Purpureomaculin differs from trigramin-gamma by a single substitution of Asp to Asn at the 17^{th} position [235].

L-amino acid oxidases (LAAO) are homodimeric flavoproteins which catalyzes the oxidation of L-amino acids, releasing corresponding α-ketoacid, ammonia and H_2O_2. The released H_2O_2 causes the LAAO-associated pathological disorders such as edema induction, platelet aggregation/induction/inhibition, apoptosis induction, antibacterial, anticoagulant, and anti-parasitic effects [151]. Peptide sequences homologous to two LAAOs (Accession AAQ16182.1 and P0DPS2.1) reported from *V. stejnegeri* and *T. purpureomaculatus* respectively, has been observed in the proteome. Both of these LAAO has been reported to show anticancer properties by anti-proliferative and apoptosis activity [236, 237].

Other non-enzymatic peptides/proteins observed in the proteome shows similarity with two cysteine rich secretory proteins (CRISPs), A chain of Stecrisp (Accession 1RC9) and Ts-CRPYa reported from V. *stejnegeri* having functions of blockage of ion channels [238]. Furthermore, two peptide fragments similar to a bradykinin-potentiating c-type natriuretic peptide (Accession P0C7P6.1) reported in the venom gland cDNA library of *C. gramineus* were identified in the proteome. These peptides are reported to have a hypotensive function and vasodepressor activity [239]. Fragments with sequence homology to glutaminyl-peptide cyclotransferases (Accession AFE84763.1) isolated from *T. gracilis* was also obtained from peak 10. However, the specific pathology associated with above mentioned components might not be visible in the envenomed patients due to low abundance in the venom. Many of these proteins/peptides work in association with other groups for the expression of some pathologies and inhibition of others, thereby, working in a holistic manner to produce different pathophysiological manifestations observed in envenomated patients. Moreover, many other commonly found minor toxin families such as phosphodiesterase (PDE), aminopeptidase (AP), 5'-nucleotidase (5-NU), endonuclease (END), vascular endothelial growth factor, hyaluronidase etc. were lacking in the proteome of *T. erythrurus.*

Trimeresurus radiation, being a large group of venomous snakes consists of individuals of multiple genus showing morphological and ecological similarity among each other. Among them, the green pit vipers possess a typical colouration with diverse patterns on their body and are mostly morphologically indistinguishable, however, their venom composition varies across species and across geographical locations. A recent

study has shown comparative venom composition of *Trimeresurus albolabris*, *T. insularis*, *Craspidocephalus puniceus* and *T. purpureomaculatus* from Java and Sumatra island of Indonesia revealing SVMPs, snake c-type lectins, SVSPs and PLA_2s to be the major families [121]. Another study accounting for venom composition of *T. albolabris* from Thailand and *T. purpureomaculatus* from Malaysia shows similar pattern of composition, although with different proportions [117]. Venomics of another Malaysian pit viper *Popeia nebularis* further reveals SVMPs as the most abundant group [115]. Markedly, proteomic studies of an Indian pit viper species, *C. malabaricus*, abundant in Western ghats, shows SVMPs and LAAO as major components [78]. In contrast, the present study undertaking proteomic analysis of the *T. erythrurus* from Mizoram, India reveals presence of PLA_2 and SVSPs as major enzymatic families and snaclecs as dominant non-enzymatic family. Interestingly, a similar pattern with PLA_2s as most abundant family has also been observed in a geographically distantly located species *T. insularis* from Lesser Sundra Island of Indonesia [119]. Another study disclosing proteomics of *C. puniceus* of Indonesian origin reveals major proteins to be PLA_2, SVSPs, disintegrins and SVMPs [116]. Along with the disparities in major protein families, the proteome of *T. erythrurus* lacks many minor toxins families (discussed above), however, some of the proteins belonging to SVSPs, snaclecs and SVMPs families such as albofibrase, venom plasminogen activator, alboaggregin, purpureotin, C-type lectin TsL, class P-II albolatin, stejnitin, class P-III TSV-DM etc. were common in many green pit vipers. Thus, *Trimeresurus erythrurus*, distributed in Indo-Malayan and Himalayan region, shows a very distinctive pattern of venom composition unlike those species found in peninsular India and other geographical locations of South-east Asia, though, with some noticeable similarities with distantly located Indonesian cousins *T. insularis* and *C. puniceus*.

The discrepancies in the venom composition of such morphologically similar species might be the potential cause for the disparities in clinical pathologies shown by patients bitten by green pit vipers across the south-east Asian countries. Envenomated patients are observed to show some typical symptoms such as extensive swelling and pain in the bite site [100] with incoagulable blood, however, some cases are assisted with systematic bleeding, tissue necrosis associated with prolonged coagulopathy and thrombocytopenia [102, 104, 105, 109]. The discrepancies in venom composition leading

to unpredictable pattern of clinical toxicity can lead to the bigger problem of mistreatment of patients. As symptomatic treatments are enough in most cases, however, in extreme cases para-specific antivenom might be administered which might lead to further complications like anaphylactic reactions in the patients [110]. The proteomic studies of morphologically similar species of a particular geographical area, thus, leads to a better insight on inter-specific venom variation along with identification of major toxins, thereby, provide a stronger theoretical basis for manufacturing region-specific heterologous antivenom to mitigate the challenges of para-specific inefficacy of present-day antivenom therapy.