SUPPLEMENTARY DATA

RESEARCH REPORT

Toxin transcripts in *Crotalus atrox* venom and *in silico* structures of toxins

Ying Jia*, Ivan Lopez and Paulina Kowalski

Biology Department, The University of Texas Rio Grande Valley, Brownsville, Texas 78520, USA

***Correspondence to:** Ying Jia, Email: Ying.jia@utrgv.edu

J Venom Res (2020), Vol 10

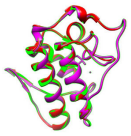
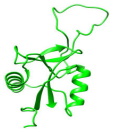
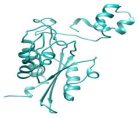
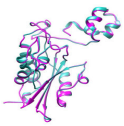
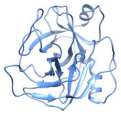
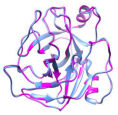
Received: 16 May 2020 | **Revised:** 15 June 2020 | **Accepted:** 16 June 2020 | **Published:** 17 June 2020

Supplementary Table 1. Deduced mature amino acid sequences of venom toxin transcripts.

Transcript	Deduced mature amino acid sequence
Lys49	SLVELGKMILQETGKNPITSYGIYGCNCGVGSRHKPKDGTDRCCFVHKCCYKKLTDCDPKMDGYTYSFKDKTIICDVNNPCLKEM CECDKAVAICLRENLDTYNKKYKIYPKFLCKKPDTC 100% identity with Lys49 (Q8UVZ7) from *Crotalus atrox*
Asp49	NLLQFNKMIKIMTKKNAFPFYTSYGCYCGWGGRGRPKDATDRCCFVHDCCYEKLTNCSPKTDIYSYSWKRGVIICGKGTPCEKQI CECDRAAAVCFRENLPTYKKRYMFYLDFLCTDPSEKC 100% identity with Asp49 (APD70896) from *Crotalus atrox*
SVMP I	VNDYEVVYPRKVTALPKGAVQPKYEDSMQYELKVNGEPVVLHLEKNIGLFSKDYSETHYSPDGRKITTNPSVEDHCYYHGRIEND ADSTASISACNGLKGHFKLQGEMYLIEPLELSDSEAHAVFKLENVEKEDEAPKMCGVTQNWESYEPIKKASQLNLNPEHQRYVEL FIVVDHGMFTKYNGDSDKIRQRVHQMVNIMKESYSYMYIDILLAGIEIWSNGDLINVQPAAPDTLNSFGEWRESDLLKRKSHD NAQLLTSIDFNGPTIGIANIGGICDPQCSTGVVQDHSKINLRVAVTMTHELGHNLGIRHDTGSCSCGGYSCIMSPVISDEPSKYFS DCSYIQCWNFIMNQKPQCILKKPLRTDTVSTPVSGNELLEA 73% identity with SVMP I (P15167) from *Crotalus atrox*
SVMP II	VNDYEVVYPRKVTALPKGAVQPKYEDSMQYELKVNGEPVVLHLEKNIGLFSKDYSETHYSPDGRKITTNPSVEDHCYYHGRIEND ADSTASISACNGLKGHFKLQGEMYLIEPLELSDSEAHAVFKLENVEKEDEAPKMCGVTQNWESYEPIKKASQLNLNPEHQRYVEL FIVVDHGMFTKYNGDSDKIRQRVHQMVNIMKESYSYMYIDILLAGIEIWSNGDLINVQPAAPDTLNSFGEWRESDLLKRKSHD NAQLLTAIDLDRVIGLAYVGSMCHPKRSTGIIQDYSEINLVVAVIMAHEMGHNLGINHDSGYCSCGDYACIMRPEISPEPSTFFSN CSYFECWDFIMNHNPECILNEPLRTDTVSTPVSGNELLEAGEECDCGSPANPCCDAATCKLRPGAQCADGLCCDQCRFIKKGTV CRPARGDWNDDTCTGQSADCPRNGLYG 86% identity with SVMP II (P34182) from *Crotalus atrox*
SVMP II	VNDYEVVYPRKVTALPKGAVQPKYEDSMQYELKVNGEPVVLHLEKNIGLFSKDYSETHYSPDGRKITTNPSVEDHCYYHGRIEND ADSTASISACNGLKGHFKLQGEMYLIEPLELSDSEAHAVFKLENVEKEDEAPKMCGVTQNWESYEPIKKASQLNLNPEHQRYVEL FIVVDHGMFTKYNGDSDKIRQRVHQMVNIMKESYSYMYIDILLAGIEIWSNGDLINVQPAAPDTLNSFGEWRESDLLKRKSHD NAQLLTAIDLDRVIGLAYVGSMCHPKRSTGIIQDYSEINLVVAVIMAHEMGHNLGINHDSGYCSCGDYACIMRPEISPEPSTFFSN CSYFECWDFIMNHNPECILNEPLRTDTVSTPVSGNELLEAGEECDCGSPANPCCDAATCKLRPGAQCADGLCCDQCRFIKKGTV CRPARGDWNDDTCTGQSADCPRNGLYG 86% identity with SVMP II (P34182) from *Crotalus atrox*
SVMP III	VNDYEIVYPRKVTALPKGAVQPKYEDAMQYELKVNGEPVVLHLEKNKQLFSKDYSETHYSPDGREITTYPLVEDHCYYHGRIEND ADSTASISACNGLKGHFKLQGEMYLIEPLKLSDSEAHAVYKYENVEKEDEALKMCGVTQNWESYEPIKKASQLVVTAEHQKYNP FRYVELVLVVDKAMVTKNNGDLDKIKTRMYEIVNTVNEIYRYMYIHVALVGLEIWSNEDKITVKPEAGYTLNAFGEWRKTDLLTR KKHDNAQLLTAIDLDRVIGLAYVGSMCHPKRSTGIIQDYSEINLVVAVIMAHEMGHNLGINHDSGYCSCGDYACIMRPEISPEPS TFFSNCSYFECWDFIMNHNPECILNEPLGTDIISPPVCGNELLEVGEECDCGTPENCQNECCDAATCKLKSGSQCGHGDCCEQC KFSKSGTECRASMSECDPAEHCTGQSSECPADVFHKNGQPCLDNYGYCYNGNCPIMYHQCYDLFGADVYEAEDSCFERNQKG NYYGYCRKENGNKIPCAPEDVKCGRLYCKDNSPGQNNPCKMFYSNEDEHKGMVLPGTKCADGKVCSNGHCVDVATAY 99% identity with SVMP III (Q90282) from *Crotalus atrox*
C-type lectin	DCPSGWSSYEGHCYQIFHVYKTWDDAERFCSEQAKGGHLVSIESSEEADFVAQLVANNVRKGISYIWIGLRVQGEEKQCSTKW SDGSSVSYENWIEAESKTCLGLEEDTNHKWVNIYCGQRNPFVCKA 96% identity with C-type lectin-like protein (AEJ31967) from *Crotalus adamanteus*
3FTx	LECEACNQPNCDFLPSIRCPKGFNQCYKKWNKIGLSVRTFERGCTANCTPNAQTKCCKTNLCNA 86% identity with 3FTx (ABZ89717) from *Sistrurus catenatus edwardsi*
CRISP	SVDFDSESPRKPEIQNKIVDLHNFLRRSVNPTASNMLKMEWYPEAAANAERWAYRCIESHSPRDSRVLGGIKCGENIYMSPVPI KWTEIIHAWHGENKNFKYGIGAVPPNAVTGHFSQVVWYKSYRIGCAAAYCPSSKYSYFYVCQYCPAGNIIGKTATPYKSGPPCG DCPSACDNGLCTNPCTKEDKYTNCKSLVQQAGCQDKQMQSDCPAICFCQNKII 100% identity with CRISP (Q7ZT99) from *Crotalus atrox*
Vespryn	ILLFTLCFFADQENGGKALASPPGIWKRADVTFDSNTAFSSLVVSANKKTVKNVGVPQVVPDNPERFNSNPCVLGSPGFRSGKH YFEVKYGTQREWAVGIAGKSVKRKGNLMLVPEERIWQMGLWWLRHLETDPGRVHSTSGKITVFVDYNGGNVIFDLNHINTTL

	KANFNGEEVVPFFYLGGTVSLTTL 97% identity with Vespryn (F8S122) from *Crotalus adamanteus*
Crotamine	QSQCEQEGGFCRFLLCPSRTSDIGKLGCEPLWKCCKSKIGK 44% identity with Crotamine (P24334) from *Crotalus durissus terrificus*
LAAO	MSSVTVLEASERVGGRVRTYRKKDWYANLGPMRLPTKHRIVREYIKKFDLKLNEFSQENENAWYFIKNIRKRVREVKNNPGLLEY PVKPSEEGKSAAQLYVESLRKVVEELRSTNCKYILDKYDTYSTKEYLLKEGNLSPGAVDMIGDLLNEDSGYYVSFIESLKHDDIFGYE KRFDEIVGGMDQLPTSMYEAIKEKVQVHFNARVIEIQQNDREATVTYQTSANEMSSVTADYVIVCTTSRAARRIKFEPPLPPKKA HALRSVHYRSGTKIFLTCKKKFWEDDGIHGGKSTTDLPSRFIYYPNHNFTSGVGVIIAYGIGDDANFFQALDFKDCADIVINDLSLI HQLPKEDIQTFCRPSMIQRWSLDKYAMGGITTFTPYQFQHFSEALTAPFKRIYFAGEYTAQFHGWIDSTIKSGLTAARDVNRASE NPSGIHLSNDNEF 99% identity with LAAO (O93364) from *Crotalus adamanteus*
SVSP	VVGGDECNINEHRSLVVLFNSSSFLCAGTLINEEWVLTAAHCDSKNFQMQLGVHSKKVLNEDEQTRDPKEKFICPNKKKDDEN DKDIMLIRLDSPVSNSEHIAPLSLPSSPPSVGSVCRIMGWGTITPTKETYPDVPRCANINILDHAVCRAAYPWNPVASTTLCAGTQ QGGKDTCRGDSGGPLICNGQFQGIVSWGGHPCGQAHEPGVYTKVFDYTDWIQSIIAGNTDATCPP 100% identity with SVSP (AUS824890) from *Crotalus atrox*
EGF-like	ALGTSEERLRQACNTCRGIADRFTQGLTDTAKKNFGGGNTAWEEKTLSKYESSEIRLVEIIENLCDSSNFECNNMVEEHEEHIEN WWFKWKKKYPDLFKWLCIETIEVCCPAGTHGPDCVACRGGSERPCHGNGDCDGDGTRAGDGSCKCQKEYQGEFCLDCSDGY YNFYKNDTHSVCTACHDSCKTCTGATNKDCKDCKEGWLRNEEACVDVDECAVEESPCNSDQYCLNTDGSFSCKACDLSCLGCT GEGPNKCKSCVTGYEMKEETCTDVDECSQTEEVCTRENTNCINTPGGYKCICSEGFEDKDDICVPSIKAEEKTSANISSPDTHEDL 100% identity with EGF- like domain (AEJ31965) from *Crotalus adamanteus*
PLB	DIHYATVYWLEAEKSFQIKDVLDKNGDAYGYYNDAIQSTGWGILEIKAGYGNQPISNEILMYAAGFLEGYLTASHMSDHFANLF PLMIKNVIIEQKVKDFIQKQDEWTRQQIKNNKDDPFWRNAGYVIAQLDGLYMGNVEWAKRQKRTPLTDFEISFLNAIGDLLDLI PALHSELRKSDFRSMPDVSRIYQWDMGHCSALIKVLPGYENIYFAHSSWFTYAATLRIYKHLDFKITDPQTKTGRASFSSYPGFLV SLDDFYILGSGXIMLQTTNSVFNLSLLKKIVPESXFAWERVRIANMMADSGKTWAETFEKQNSGTYNNQYMILDTKKIKLQRSLE DGTLYIIEQVPKLVKYSDQTKVLRNGYWPSYNIPFDKEIYNMSGYGEYVQRHGLEFSYEMAPRAKIFRRDQGKVTDMESMKFIM RYNNYKEDPYAKRNPCNTICCRQDLDRRTPVPAGCYDSKVADISMAAKFTAYAINGPPVEKGLPVFSWVHFNKTKHQGLPESY NFDFVTMKPVL 98% identity with PLB (F8S101) from *Crotalus adamanteus*

Supplementary Table 2. Venom toxin homology modelling and structure validation.

Model (M)	Template (T)	Superimposition	Ident (%)	RMSD(Å)	ERRAT (%)	Z-Score	ϕ/ψ-plot (%)
 Lys49	1PPA		80.2	0.19	92.9(T) 92.1 (M)	-5.4(T) -5.0(M)	89.9 (T) 90.5 (M)
 Asp49	2OQD		76.2	0.18	100(T) 91.2(M)	-4.6(T) -4.9(M)	89.2(T) 93.4(M)
 PIII	2DW0		99.8	0.58	91.8(T) 88.7(M)	-9.9(T) -9.7(M)	87.2(T) 88.2(M)
 C-type lectin	1BJ3		80.6	0.24	71.7(T) 84.1(M)	-7.1(T) -7.0(M)	80.5(T) 90.4(M)
 3FTx	2VLW		41.8	0.58	97.5(T) 80.1(M)	-4.3(T) -3.4(M)	87.5(T) 78.2(M)
 CRISP	1RC9		81.5	0.20	92.5(T) 86.6(M)	-6.3(T) -5.7(M)	87.5(T) 88.8(M)
 LAAO	5TS5		87.9	0.34	98.5(T) 88.2(M)	-9.8(T) -9.3(M	93.2(T) 92.9(M)
 SVSP	1BQY		83.8	0.24	85.7(T) 85.8(M)	-7.5(T) -7.3(M)	84.2(T) 89.8(M)
 PLB	4BWC		68.4	0.21	93.2(T) 76.7(M)	-6.7(T) -6.7(M)	86.7(T) 89.8(M)