Supplementary Materials: Structural and Functional Insights into the C-terminal Fragment of Insecticidal Vip3A Toxin of *Bacillus thuringiensis*

Kun Jiang, Yan Zhang, Zhe Chen, Dalei Wu, Jun Cai and Xiang Gao



VIP3 A a11	β 6a	β 6b	β7	β8	β 9a	→		β9 b		
WTD25-11	410		0 430 	440			460			
VIP3Aa1 VIP3Aa1	SEQIYYT	NNIVFPNE	YVITKIDFTKKM YVITKIDFTKKM	TLRYEVTANFY	DSSTGEIDLNKKKVI	2 S		AEYRTLS		
Vip3Ad2	SEQIYYT	KNIAFPNE KNIAFPNE	V I TKIDFTKKM VVI TKIAFTKKM	NSLRYEATANFY	DSSTGDIDLNKTKVI	85 8 <mark>5</mark>		AEYSMLK		
Vip3Ae1 Vip3Af1	SEQIYYT SEQIYYT	NNIAFPRE NNIV <mark>FP</mark> NE	Y V I T K L T F T K K M I Y V I T <mark>K I D F T K</mark> K M I	SLRYEATANFY (T <mark>LRYEVTANF</mark> Y	DSSTGDMDLNKTKVI DSST <mark>G</mark> EI <mark>DLNK</mark> KKVI	E S	SE.	AEYSRLS AEYRTLS		
Vip3Ag2 Vip3Ah1	SEQIYYTI SQQMYYLI	NNIAFPNE HNITFPNE	YVITKITFTKKM YVITEIIFTKKK	NSLRYEATANFY NSLRYEVIANYY	DSSTGDIDLNKTKVI EFSSGDIDLNKKLVI	ES	SE.	AEYSTLS AEYSTLS		
Vip3Ai1 Vip3Ai1	SEKIYY SEOIYYT	NNIV <mark>FP</mark> NE KNIAFPNE	YVITKID <mark>F</mark> TKKMI YVITKIAFTKKM	KTLRYEVTANSY NSLRYEATANFY	D <mark>S ST</mark> GE I D L N K K K V I D S STGD I D L N K T K V I	8 <mark>5</mark>	SKI	AEYRTLS AEXSMLK		
Vip3Ba1 Vip3Bb2	REQKYYII	K D M T F P E G	YVITKIVFEKKLI YVITKIVFEKKL	NLLGYEVTANLY	DPFTGSIDLNKTIL DPFTGSIDLNKTIL	ESWKEDCCEE		ELYKIIE		
Vip3Bc Vip3Cal	LEOKYYII	KDIEFPEG KNLTEPDG	YVITKIVFEKRLM YVITKITEEKKL	NOLGYEVTANFY MILTYEADANEY	DPSTGSIDLNKVKVI	ESWKEKSCEE	DSCEDE	YSIIK		
Vipscai	SEQUELL					1 F	<u>FQ</u>			
	Domain III									
VIP3Aa11		β10	β11	β12	α6 → 00000000					
WTD23-11		480 480			510 520			550		
VIP3Aa1	ANDD.GV		ETFLTPINGFGL	ADENSRLITLT	CKSYLRELLLATDL CKSYLRELLLATDL	SNKETKLIVE	PSGFISNIVENGSI PSGFISNIVENGSI	EEDNLEP		
Vip3Ad2	ASDD.EV	Y MPLGVIS Y MPLGLIS	ETFLNPINGFGL	VDENSRLVTLT VDENSRLVTLT	CRSYLRETLLATDL CRSYLRETLLATDL	NNKETKLIVF	PISEISNIVEN GNL PNVFISNIVEN GNI	EMDTLEP		
Vip3Ae1 Vip3Af1	ASND.GV ANDD.GV	Y M P L G L I S Y M P L G V I S	ETFLTPINGFGL ETFL <mark>T</mark> PI <mark>NG</mark> FGL	ADENSRLVTLT ADENSRLITLT	CKSYLREILLATDL CK <mark>SYLRE</mark> LLLATDL	SNKETKLIVE SNKE <mark>TKLIV</mark> E	PPNGFISNIVENGNL. PPS <mark>GFIS</mark> NIVENGSI	EGENLEP		
Vip3Ag2 Vip3Ah1	VSND.AI	Y TPLGIIS Y MPLGVIS	ETFLTPINGFGI ETFL <mark>T</mark> PI <mark>KG</mark> FGL	VDENSKLVNLT V <mark>D</mark> ESSRLVT <mark>L</mark> T	CKSYLRE <mark>ILLATDL</mark> CK <mark>SYLRE</mark> ILLATDL	SNKETKLIVE SNKATKLIVE	PIGFISNIVENGNL. PNGFISNLVENGDI	EGENLEP		
Vip3Ail Vip3Ajl	ANND.GV ASDD.EV	YMPLGVIS Y <mark>MPLGLI</mark> S	ETFLTPINGFGL ETFL <mark>NPI</mark> NG <mark>F</mark> RL	ADENSRLITLT V <mark>D</mark> ENSRLVT <mark>L</mark> T	CKSYLRELLLATDL CR <mark>SYLRE</mark> TLLATDL	SNKETKLIVP NNKE <mark>TKLIV</mark> P	PPNSFISNIVENGSI PPNVFISNIVENGNI	EEGHLEP EMDTLEP		
Vip3Bal Vip3Bb2	ADTN.GV ADTN.GV	YMPLGVIS Y <mark>MPLGVI</mark> S	ETFL <mark>TPIYSF</mark> KL] ETFL <mark>TPIYSF</mark> KL]	IIDEKTKKISLA IDERTKRISLA	GKSYLRESLLATDL GK <mark>SYLRE</mark> SLL <mark>AT</mark> DL	VNKETNLIPS VNKDTNLIPS	SPNGFISSIVQNWHI SPNGFISSIVENWNI	TSDNIEP TSDNIEP		
Vip3Bc Vip3Cal	AETD.GI IGDDDGI	YMPLGVVS Y <mark>MPLGVI</mark> S	ETFLTPIYGFGL ETFLTPINSFGL	TVDEKNQKITLT SVDAKSKTLTLK	GKSYLRE <mark>SLLET</mark> DL CK <mark>SYLRE</mark> YLL <mark>ES</mark> DL	LNNETYLIAS KNKET <mark>GLI</mark> AP	SPDGYISSIVENWNI PPNVFISNVVKNWDI	TSDNTGS EEDSLEP		
	_	A	A	A		A		·		
			B13	614 m2	615	B16	B17	ain IV 618		
VIP3Aall	5	60	570 58		600 ····	610	620 63	30		
VIP3Aa11	WKANNKN	AYVDHTGG	VNGTKALYVHKDO	GISOFIGDKLK	P <mark>KTEYVIQY</mark> TVKGK	PSIHLKDENI	GYIHYEDTNNNLED	YOTINKR		
VIP3Ab1	WIANNKN	AYVDHTGG AYVDHTGG	INGTKVLYVHKDO	EFSQFVGGKLK	SKTEYVIQYIVKGK	ASIY <mark>lk</mark> dkkn	ENSIYEEINNDLEG	FQTVTKR		
Vip3Ad2 Vip3Ae1	WKANNEN. WKANNKN	ANVDYSGG AYIDHTGG	VNGTRALYVHKDO VKGTKVLYVHKDO	GEFSQFIGDKLK	SKTEYLIRYIVKGK SKTEYVIQYIVKGK	ASIFLKDEKN AVIY <mark>LKD</mark> EKN	IGDYIYEDTNNNLED IGDYIYEEINNELED	Y QTITKR F <mark>QT</mark> VTKR		
Vip3Af1 Vip3Ag2	WKANNKN WKANNKN	AYVDHTGG AYV <mark>D</mark> HTGG	VNGTKALYVHKDO VNGTKA <mark>L</mark> YVHK <mark>D</mark> O	GFSQFIGDKLK EFSQ <mark>FIGDKL</mark> K	PKTEYVIQYTVKGK SKTEYVIQYIVKGK	PSIHLKDEN I ASIL <mark>LK</mark> DEKN	GYIHYEDTNNNLKD GDCIYEDTNNGLED	YQTITKR F <mark>QT</mark> ITKS		
Vip3Ah1 Vip3Ai1	WKGNNKN WKANNKN	A Y V D H T G G A Y V <mark>D</mark> H T G G	VNGTKALYTQDD VNGTKALYVHEDO	∃EFSQFIGDKLK GVSQ <mark>F</mark> M <mark>GDKL</mark> K	SKTEYIIQYTVKGN PKTEYVIQYTVKGKI	TSIYLKDKKN PSIH <mark>LK</mark> DENI	ENVIYEDKNNNLEA GYILYEDTNNDLED	F QT ITKR F <mark>QT</mark> ITKR		
Vip3Aj1 Vip3Bal	WKANNEN WKANNKN	ANVDYSGG AYV <mark>D</mark> KTDA	VNGTRALYVHKDO MVGFSSLYTHKDO	FFSHFIGDKLK FFLQFI <mark>G</mark> AKLK	SKTEYLIRYIVKGK AKTEYIIQYTVKGNI	A S I F <mark>L K D</mark> E K N P E V Y <mark>L K</mark> N N K D	ENYIYEDTNNNLED ICYEDKTNNFDT	Y <mark>QT</mark> ITKR F QT ITKK		
Vip3Bb2 Vip3Bc	WKANNKN WRANN <mark>N</mark> N	A Y V <mark>D</mark> K T D D A F V <mark>D</mark> K A D T	MVGFNSLYTHKDO IKGSSSLYTHKDO	EFLQFIGAKLK EFSQFIGNKLK	AKTEYIIQYTVKGS PKTNYVIQYVIKGR	P E V Y <mark>L K</mark> N N K G P A I Y <mark>L K</mark> N N K C	IFYEDTTNKFDT TLFEDTKNNFSD	F <mark>QT</mark> ITKK F QT VTKK		
Vip3Ca1	WVANNKN	AYV <mark>D</mark> NTGG	IE <u>RSKAL</u> FTQGD(EFSQFIGDKLK	PNTDYIIQYTVKGK	PAIYLKNKSI	GYIT <mark>Y B</mark> DTNGN <u>SEE</u>	F <mark>QT</mark> IAVK		
				D	omain IV					
VTP3Aa11	→	β19	β20	_	η3 β21	β22 β	β23 β24	_		
VTP2Ac11	640	650	660 600 CDEDWGDWGD			690 	700 710 710			
VIP3Aali VIP3Aal	FTTGTDL	KGVYLILK	S Q N G D E A W G D N F I S Q N G D E A W G D N F I	ILEISPSEKLL	SPELINTNNWTSTG: SPELINTNNWTSTG:	S.TNISGNTI	TLYQGGRGILKONL	OLDSFST		
Vip3Ad2	FTTGTDS	TGVYLIFN	SQNGEGAFGGNF SQNGDEAWGDNF	I LEISPCEKLL	SPELIMSDAWVGSQC SPELIKTDKWNSTG:	S.TYISGNSI	TLYRGGRGILKONL	QLDGFST		
Vip3Ae1 Vip3Af1	FITGTDS: FTTGTDL	SGVHLIFT KGVYLILK	S Q N G E E A F G G N F 1 S Q N G D E <mark>A W</mark> G D K F 1	ISEIRPSELL ILEIKPAEDLL	SPELIKSDAWVGTQ(SPELINPNSWITTP(G.AWNSGNSL G.AS <mark>ISG</mark> NKL	FINLGTNGTFRONL	SLNSYST		
Vip3Ag2 Vip3Ah1	FITGTDS: FTTELDS:	SGVHLIFN SDVYLVFK	S Q N G D E A F G E N F 1 C K N G Y K <mark>A W</mark> G D N F I	ISEIRLSEDLL ITEIRPKE.VV	SPELINSDAWVGSQ(SPELIKVENW <mark>IG</mark> MG(G.TWISGNSL GSNH <mark>VNP</mark> DSL	LINSNVNGTFRONL LFTGGRSILK <mark>O</mark> NL	SLESYST QLDSY <mark>S</mark> T		
Vip3Ail Vip3Aj1	FTTGTDLI FTTGTDS	MRVYLILK TGVYLIFN	S Q S G H E AWG D N F 1 S Q N G D E <mark>A</mark> F <mark>G</mark> E N F 1	TILEIKPAEALV TISEIRLSEDLL	SPELINPNSWITTQ(SPELINSDAWVGSQ(G.AS <mark>ISG</mark> DKI G.TW <mark>ISG</mark> NSI	FISLGTNGTFRONL TINSNVNGTFRONL	SLNSYST SLESYST		
Vip3Bal Vip3Bb2	FNSGVDP FNSGVDP	SEIYLVFK SEIYLVFK	NQIGYEAWGNNF] NQIGYEAW <mark>G</mark> NKF]	ILEIKSLET	LPQILKPENWIPLGI LPQILKPENWMPFGI	NAEIKEDGKI NAEIKEDGKI	EISGNGSMNOYI EISGNGTMTONI	QLEQNSK QLEQNSK		
Vip3Bc Vip3Cal	FNSGVNP: FTSETDL	SEIYFLFK SQTHLVFK	NQSEYE <mark>AWG</mark> NNF SQNGYE <mark>AWG</mark> DNF	ILEIKSLEF ILEAKLFETPE	LPQMLKPEDWIPSGI SPELIKFNDWERFG	NVQMKD <mark>G</mark> GRI F.TY <mark>ITG</mark> NEI	EILGDGYFKOFI RIDHSRGGYFROSL	KLENDST NIDSYST		
	-	Dom	nin IV	_		De	omain V			
	625	BO R2	6 B2	7	628	629	n4 630a 630)b		
VIP3Aall	720	→ <u>730</u>	→ 7 <u>40</u>	→ ► 750	7 6 0	770	780 poor poor	•		
VIP3Aa11	YRVYFS.	. VSGDAN <mark>V</mark>	RIRNSR EVLFE	R.YMSGAK	DVSEMFTTKFEKDN	YIELSQGNN	LYGGPIVHFYDVSI	K		
VIP3Aal VIP3Abl	YRVYFS. YSMNFT.	. VSGDANV . VNGFGKV	KIRNSR.EVLFEH TVRNSR.EVLFEH	K.YMSGAK	DVSEMETTKFEKDNI DISEKETTAANNTGI	FILELSQGNN LYVELSRST.	.SGGAIN.FRDFSI	K K		
Vip3Ad2 Vip3Ae1	YRVNFS. YSMNFN.	. VDGDANV . ITGFGKV	KIRNSR.EVLLEH TIRNSR.EVLFEH	KYLNRK	GVSEMFTTKFDKDN DYSEKFTTAANNTGI	FYVELSQGD. FYVELSRGT.	.NLGTIVHEYDFSI .QGGNIT.FRDFSI	K K		
Vip3Af1 Vip3Ag2	YSISFT. YSMNFN.	. ASGPFNV . VNGFAKV	IVRNSR.EVLFE IVRNPR.EVLFE	SNLMSSTS NYPQLSPK	H I SGTFKTESNNTG D I SEKFTTAANNTGI	LYVELSRRS. LYVE <mark>L</mark> SRFT.	.GGGGHISFENVSI .SGGAIN.FRDFSI	K K		
Vip3Ah1 Vip3Ai1	YRVRFSLI YSISFT.	MVIGKAK <mark>V</mark> .ASGPFNV	IIRNSS.EVLFEN TVRNSR.EVLYEN	SYVND <mark>S</mark> EGVLE NNLMS <mark>STS</mark>	G <mark>VS</mark> ET <mark>FTT</mark> KSIQDNI HISGEFKTESNNTGI	F <mark>YVEL</mark> SNEGI L <mark>YVEL</mark> SRRS.	FGSKDVAYFYNF <mark>S</mark> I GGAGHISFENISI	K K		
Vip3Aj1 Vip3Bal	YSMNFN. YHLRFS	. VNGFAK <mark>V</mark> . VKGKGRV	TVRNSR.EVLFER	NYPQLSPK	DIFEKFTTAANNTGI EVSIMIETTRLYGE	L <mark>YVEL</mark> SRFT. GIISLLNDE	.SGGAIT.FRDFSI	К <mark></mark> VKЕ		
Vip3Bb2 Vip3Bc	YHLRFS.	VKGKGRV	AIQTOSSHI	N.VPATNE	EVSTMITTRNLYGE DLTRVIKNTSSKGE	GMIYLFNDD. CFIALEGTY	.VENSKVIFSDVSL VENSSTIFSNVST	VKE		
Vip3Ca1	YDLSFS.	FSGLWAKV	IVKNSRGVVLFE	VKNNG <mark>S</mark> SYE	DISESFTTASNKDG	F <mark>FIELT</mark> AER.	.TSSTFHSFRDISI	KEKIE		
				Domai	n V					

Figure S1. Sequence alignment of selected Vip3 family members. Each domain is indicated by the lines below the sequences, coloured as in Figure 1A. Secondary structural elements of Vip3Aa11 are shown above the sequences. The conserved hydrophobic amino acid residues discussed in domain II and domain III are marked with green and magenta triangles, respectively. The potential cleavage

site between domain III and domain IV is highlighted with blue triangle. ClustalX2 was used for the sequence alignment. ESPript-3.0 was used to generate the figure.



Figure S2. Structure of MBP-Vip3Aa11_{200-end} in the *P*₂₁ space group. Two views of MBP-Vip3Aa11_{200-end} and structure in one asymmetric unit. There are four molecules of MBP-Vip3Aa11_{200-end} in one asymmetric unit and they are arranged into two copies of dimer in the different orientations. The molecule A, B, C and D are shown in green, cyan, magenta and yellow, respectively. The MBP (Maltose Bind Protein) tags are shown in silver color in all four molecules. The interaction area between molecule B and C is less than 500 Å², as calculated by PISA server.



Figure S3. Structural alignment between Molecule A and B from the Vip3Aa11_{200-end} dimer. Structure superimposition for the Vip3Aa11_{200-end} and each domain between molecule A and B from the Vip3Aa11_{200-end} dimer structure. Molecule A is coloured as Figure 1A, and Molecule B is shown in cyan color. The root mean square deviation (r.m.s.d) of each alignment is listed.



Figure S4. Two hydrophobic helices from domain II. The hydrophobic amino acid residues are shown as stick and labelled with residue numbers. The amino acid residues involved in the hydrophobic (red) and polar (yellow) interactions between α 1 and α 4 helices are shown as sticks, and the polar interactions are shown in black dashes.



Figure S5. Images of Sf9 cells treated with RFP. Fluorescence microscope images of Sf9 cells treated with RFP protein only for 6 h as control. The images are representative of three independent experiments. Nuclei are stained with DAPI (blue), Scale bar: $10 \mu m$.



Figure S6. Overall structure of Vip3B2160. Domains I, II, III, IV and V are coloured in light purple, blue, light brown, magenta and green, respectively.

Parameter name	Se-SAD	Native I	Native II
Data collection			
Wavelength (Å)	0.9792	0.9793	0.9793
Space group	P212121	P212121	P21
Cell dimensions			
<i>a, b, c</i> (Å)	125.71,139.35,163.20	126.08,140.20,167.64	136.09,127.88,149.18
α, β, γ (°)	90, 90, 90	90, 90, 90	90,91.378,90
R _{meas}	0.153 (0.829)*	0.207(0.915)	0.145 (0.753)
Ι/σΙ	9.30(2.80)	7.47(1.32)	6.11 (1.00)
CC1/2	0.981(0.674)	0.983(0.565)	0.993(0.541)
Completeness (%)	91.25 (79.60)	93.28(91.40)	97.33 (95.63)
Redundancy	3.5 (3.7)	3.9(3.8)	2.3 (2.3)
Wilson B-factor	48.4	99.2	71.4
MR-SAD Phasing			
Selenium sites (ShelX/D)	11		
PATFOM (/D)	24.47		
Overall CC (ShelX/E) (%)	6.57		
Pseudo-Free CC (/E) (%)	60.29		
Final Map CC (/E)	0.995(0.897)		
Figure of Merit (FOM)	0.564		
Refinement			
$\mathbf{B}_{\text{assolution}}(\mathbf{\hat{A}})$	49.93-3.91	42.42-3.62	28.93-3.195
Resolution (A)	(4.04–3.90)	(3.75–3.62)	(3.309–3.195)
No. reflections			82266 (8040)
$R_{ m work}/R_{ m free}$			0.1980/0.2389
No. atoms			30179
Protein			29960
Ligand/ion			42
Water			177
Average <i>B</i> -factors			74.14
Protein			74.32
Ligand			86.70
Water			40.48
R.m.s. deviations			
Bond lengths (Å)			0.002
Bond angles (°)			0.58
Ramachandran			
Favored (%)			95.01
Allowed (%)			4.80
Outliers (%)			0.18

Table S1. X-ray and refinement statistics.

*Values in parentheses are for highest-resolution shell.\

Primer Name	Sequence(5'→3')	Function
Vip200-F	GAGCTTTCGCTGCAGCGTCC <u>GGCTCTCCTGCAGATATTC</u>	Vip200-end cloning
Vip200-R	GTTAGCAGCCGGATCTCAGTGTTACTTAATAGAGACATCGTAAAAATGTAC	Vip200-end cloning
DmI-III-F	AAGAAGGAGATATACCATGGGC <u>ATGAACAAGAATAATACTAAATTAAGC</u>	DmI-III cloning
DmI-III-R	TCGACTGCAGAGGCCTGCAT <u>AGAAAGTGTAGGGAGGATGTTTAC</u>	DmI-III cloning
DmVI-V-F	AAGAAGGAGATATACCATGGGC <u>GGTTTTATTAGCAATATTGTAGAG</u>	DmVI-V cloning
DmVI-V-R	TCGACTGCAGAGGCCTGCAT <u>CTTAATAGAGACATCGTAAAAATG</u>	DmVI-V cloning
DmI-II-F	AAGAAGGAGATATACCATGGGC <u>ATGAACAAGAATAATACTAAATTAAGC</u>	DmI-II cloning
DmI-II-R	TCGACTGCAGAGGCCTGCAT <u>AGAAAGTGTAGGGAGGATGTTTAC</u>	DmI-II cloning
DmII-III-F	AAGAAGGAGATATACCATGGGC <u>GATGGCTCTCCTGCAGATATTCTTG</u>	DmII-III cloning
DmII-III-R	TCGACTGCAGAGGCCTGCAT <u>AGAAAGTGTAGGGAGGATGTTTAC</u>	DmII-III cloning
DmIII-F	AAGAAGGAGATATACCATGGGC <u>ACACTTTCTAATACTTTTTCTAATC</u>	DmIII cloning
DmIII-R	TCGACTGCAGAGGCCTGCAT <u>TTCTTTATTGCTTAAGTCTGTTGC</u>	DmIII cloning
pET-MBP-F	CACTGAGATCCGGCTGCTAAC	pET28-MBP cloning
pET-MBP-R	<u>TTCTTTATTGCTTAAGTCTG</u>	pET28-MBP cloning
pET-RFP-F	ATGCAGGCCTCTGCAGTCGACGGG	pET28-RFP cloning
pET-RFP-R	<u>GCCCATGGTATATCTCCTTCTT</u>	pET28-RFP cloning

The nucleic acid bases corresponding to each gene are labeled with underscores.