## The Relationship Between Structure and Function

## --take Bacillus thuringiensis Cry toxin as an example

CAASI0S2G (邓超 王金龙 庄绪静 张彦蕊 姜腾飞) Reporter: Deng Chao June 13<sup>th</sup> ,2010



## Outline

- I. Introduction
- 2. Materials
- 3.Analysis and prediction of the general properties
- 4.Analysis of the structure and function



## I. Introduction

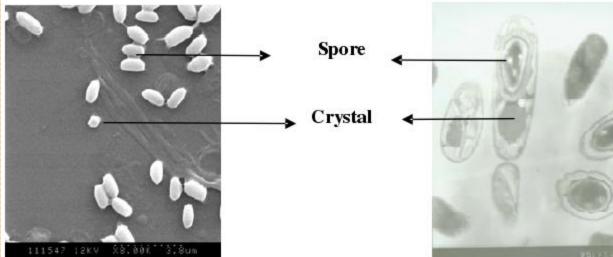
• Bacillus thuringiensis (BT)

G<sup>+</sup>

Spore-forming

Insecticidal crystal proteins

(Cry toxin, delta-endotoxins)





Kong-Ming Wu et al. *Science* 2008 (321): 1676-1678

#### Yan Guixin.2009 CAAS [D]

## BT toxins

## ~503 Bt toxins(Cry, Cyt and VIP)

( http://www.lifesci.sussex.ac.uk/home/Neil\_Crickmore/ Bt/ Updated Feb 24<sup>th</sup> ,2010)

Cry proteins

~ 386 (UniProt, with term "cry and Bacillus thuringiensis")

1 - 25 of 386 results for name:cry ≥ AND organism:"bacillus thuringiensis" ≥ in UniProtKB sorted by score descending ≥ Show only reviewed (85) ≠ (UniProtKB/Swiss-Prot) or unreviewed (301) ≠

## • Cry proteins

#### Toxic to many different species of insects larvae.



CK

BtSU4

CK

SU4-DSTK

CK

CK

HD8H

3A - SU4

CK



HD8I

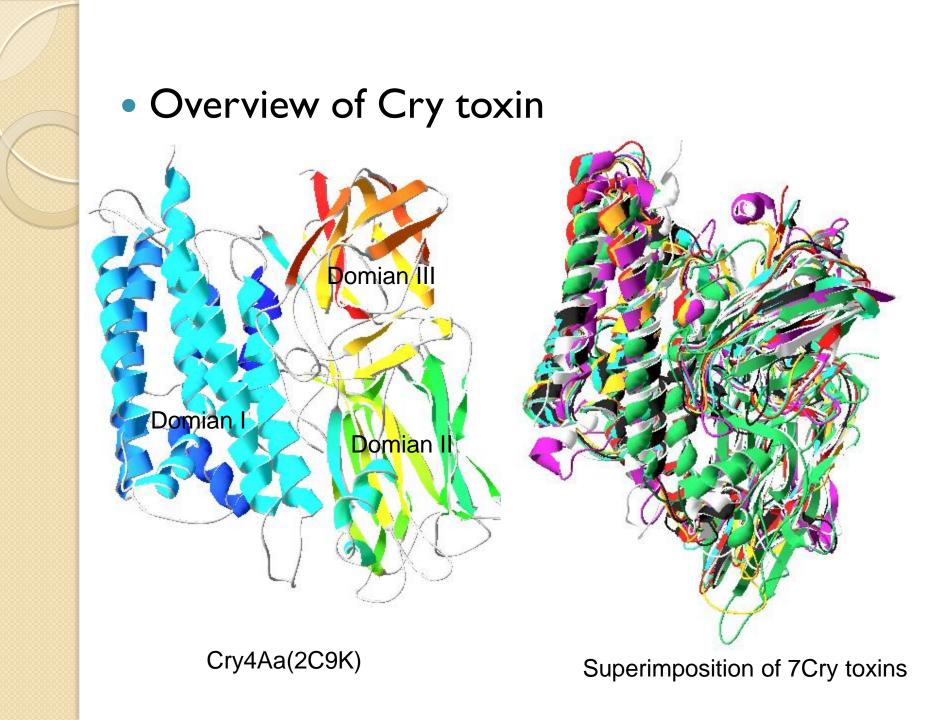


## • The nomenclature of Cry proteins

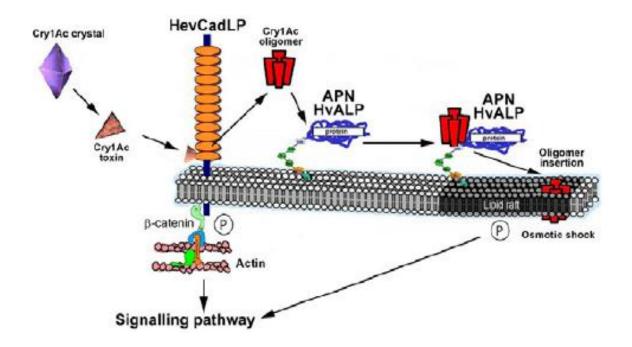
#### Based on the similarity of the Cry protein sequences

Similarity	Grade /symbol	Example(s)
<45%	I/ I,2,3	CryI,Cry2
45%-78%	II/ A,B,C	CryIA,CryIB
78%-95%	III/ a,b,c	CryIAa,CryIAb
>95%	IV/ 1,2,3	CryIAaI, CryIAa2

Crickmore N. et al. *Microbiol Mol Biol Rev.* 1998,62:807~813.



## • The mechanism of the toxic action to insects larvae



Jurat-Fuentes J.L. ,Adang M.J . *Journal of Invertebrate Pathology* 2006,92(3):166~171.



## 2. Materials

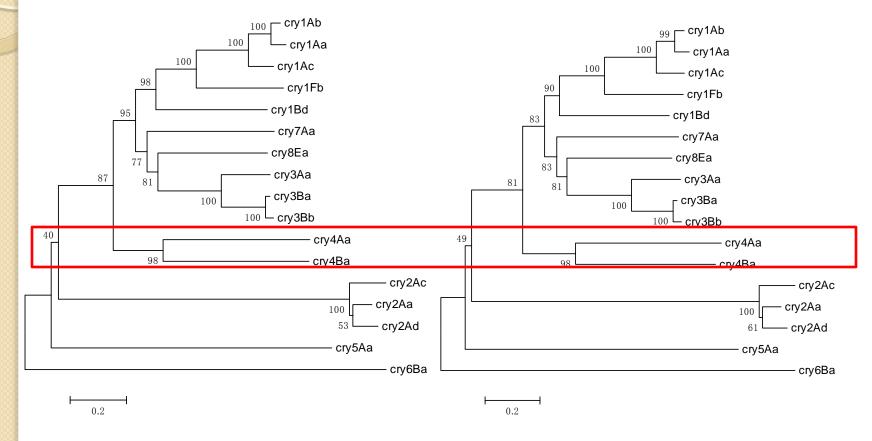
Get the total 7 Cry proteins structures in PDB.

[Cry1Aa(1CIY), Cry2Aa(115P), Cry3Aa(1DLC), Cry3Bb1(1J16), Cry4Aa(2C9K), Cry4Ba (1W99), Cry8Ea1(3EB7)]

 Get full sequences of 17 Cry proteins in UniProt (including the full sequences of the 7 PDB Cry proteins and their toxic sequences).

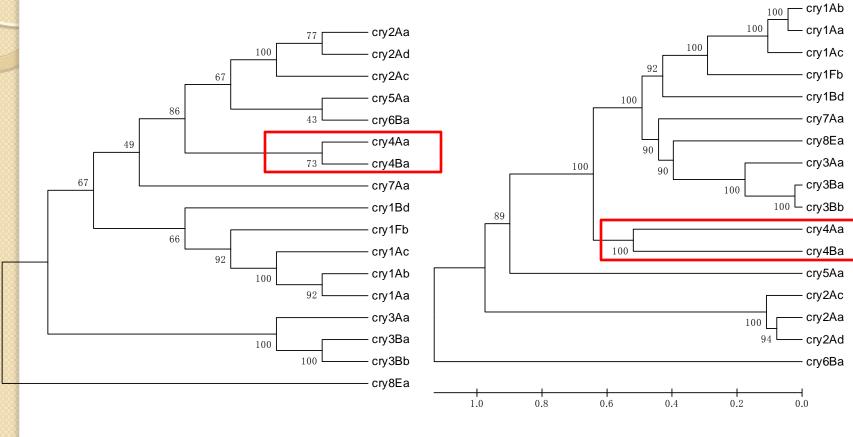
(CryIAa, CryIAb, CryIAc, CryIBd, CryIFb, Cry2Aa, Cry2Ac, Cry2Ad, Cry3Aa, Cry3Ba, Cry3Bb, Cry4Aa, Cry4Ba, Cry5Aa, Cry6Ba, Cry7Aa, Cry8EaI)

## Phylogenetic tree of 17 Cry proteins. (by MEGA 4.0)



Neighbor-Joining method

Minimum Evolution method



Maximum Parsimony method

**UPGMA** method

## Choose Cry4Aa(PI6480) and Cry4Ba(P05519) for further analysis

-both with PDB structures [Cry4Aa(2C9K), Cry4Ba (1W99)] -both have toxic to Aedes (伊蚊), Culex (库 蚊) and Anopheles (按蚊) larvae but significantly differ in toxicity level

 Table 1. Larvicidal activity of Bacillus thuringiensis svar. israelensis powders containing individual spore/crystals mixtures, against Aedes aegypti early 4<sup>th</sup> instar larvae

Powder	No. assays	No. larvae	LC <sub>50</sub> <sup>a</sup> (95% fiducial limits)
Crv11Aa	3	1140	1.35 (1.01-1.82)
Cry4Aa	3	1020	13.01 (8.82-20.04)
Cry4Ba	3	1020	0.12 (0.08-0.54)
Bti <sup>b</sup>	3	1080	0.013 (0.011–0.016)

Henrique de Barros Moreira Beltrao, Maria Helena Neves Lobo Silva-Filha. *FEMS Microbiol Lett* 2007, (266) 163–169

## 3. General property(full length)

	Cry4Aa	Cry4Ba	Program/method
Length (aa)	1180	1136	ProtParam
M(Da)	134538.7	127764.2	ProtParam
PI(theoretical)	5.04	4.87	ProtParam
Hydrophobicity	-0.424152	-0.346919	SOSUI
SecP score	0.901246	0.822789	SecretomeP
Identity	56.5%		Jemboss
Similarity	68.1%		Jemboss

**Conclusion 1:** The acidity of the Cry ptoteins make it easier to dissolve in the midgut where is alkaline.

## • Transmembrane prediction (SOSUI)

Cry4Aa

[Hydropathy profile]

#### Cry4Ba

This amino acid sequence is of a MEMBRANE PROTEIN which have 1 transmembrane helix.

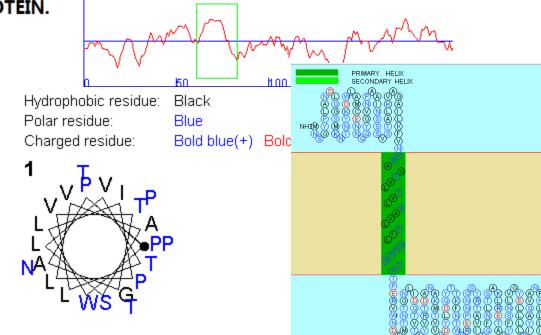


 No.
 N terminal
 transmembrane region
 C terminal
 type
 length

 1
 61
 PPAGTVLTVLSAVLPILWPTNTP
 83
 PRIMARY
 23

 [Hydropathy profile]

This amino acid sequence is of a SOLUBLE PROTEIN.



## • Smart: different results!

#### Domains within *Bacillus thuringiensis* protein CR4AA\_BACTI (P16480)

Pesticidal crystal protein cry4Aa (Insecticidal delta-endotoxin CryIVA(a))

100 200

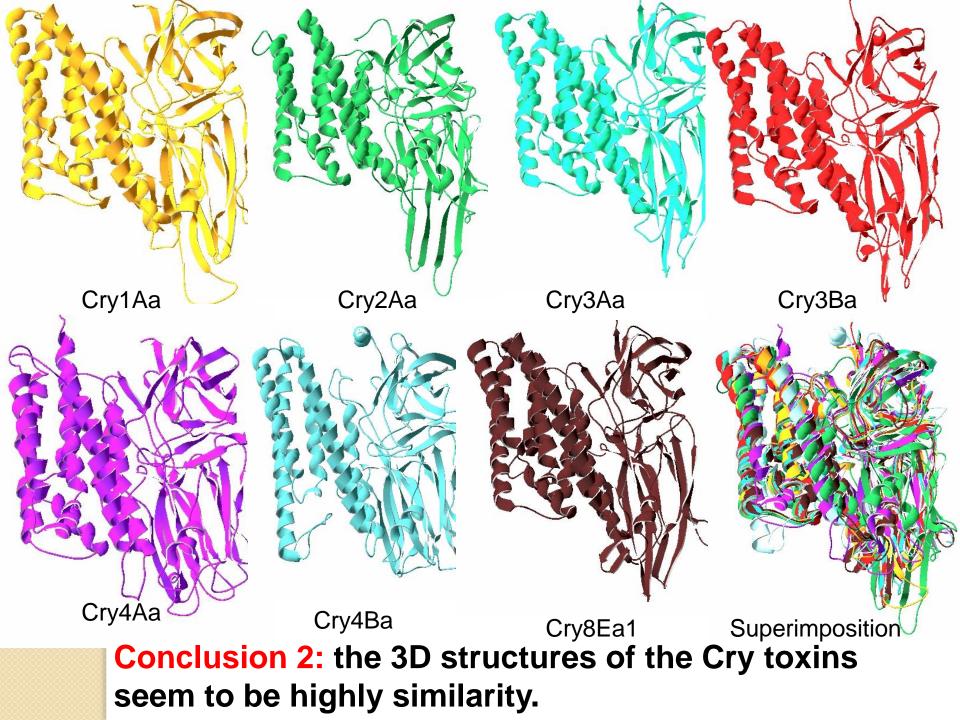
#### Domains within *Bacillus thuringiensis* protein CR4BA\_BACTI (P05519)

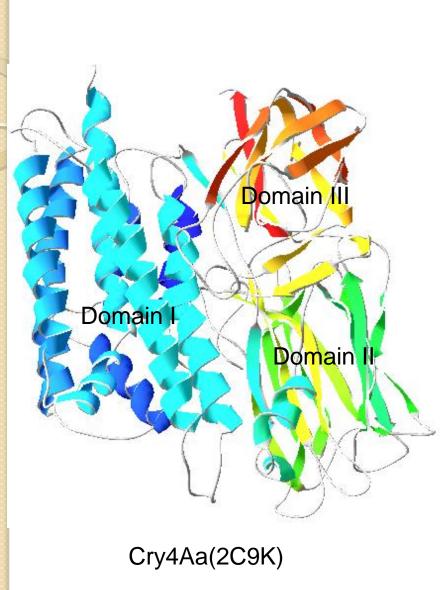
Pesticidal crystal protein cry4Ba (Insecticidal delta-endotoxin CryIVB(a))

1 100 200

# 4. Structure and function (toxin sections)

- The full length protein (protoxin) dissolved and cleaved in the midgut of larvae, produce the active toxin.
- Cry4Aa toxin: 68-679
- Cry4Ba toxin: 84-641





## • Domain I

Helixes rich (usually 7), responsibility for the formation of ion channel in the midgut cell membrane (the toxic domain).

## Domain II

Beta-sheet rich, involved in the specific binding to the receptor.

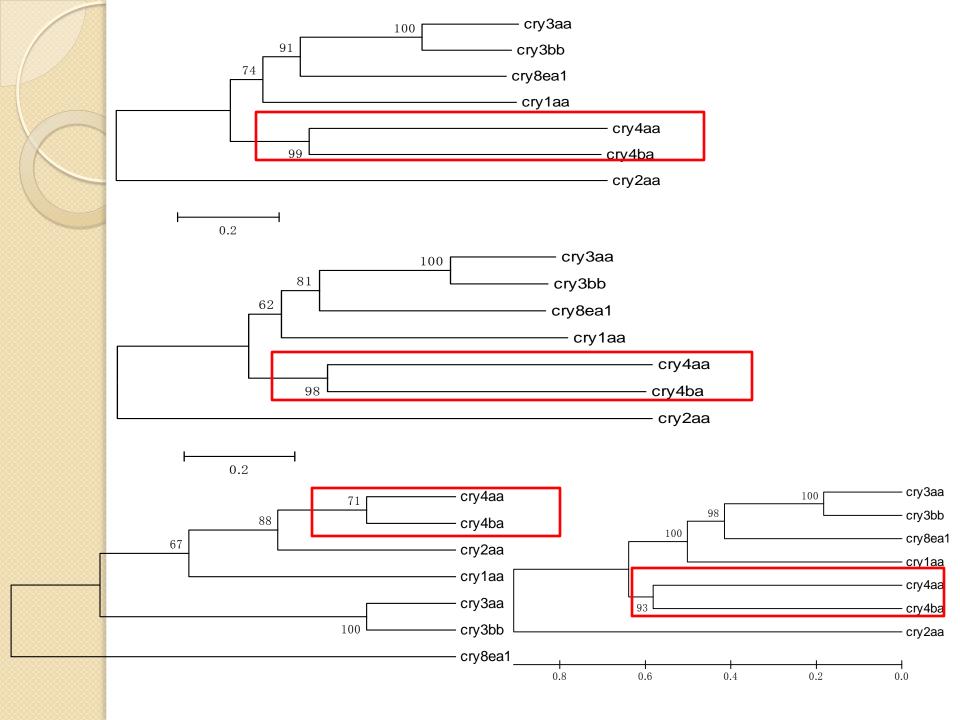
• Domain III

Not very clear, maybe with both function.

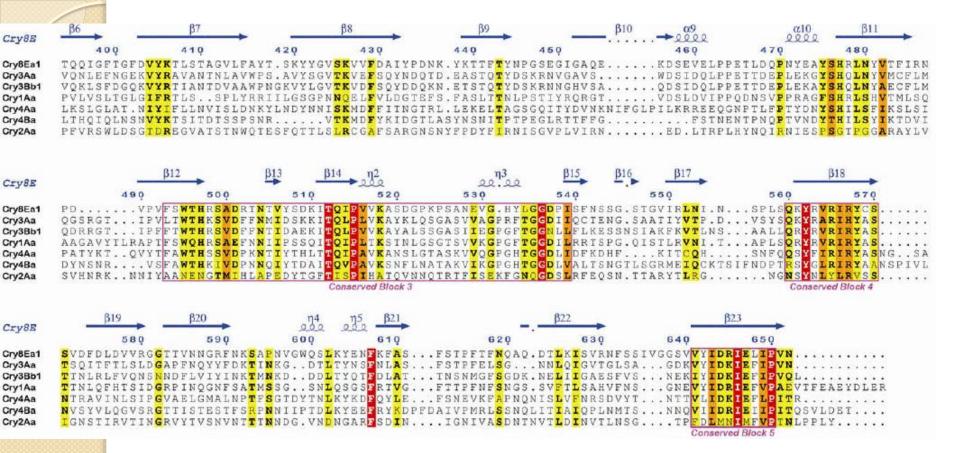
## Multiple alignment of 7 Cry toxins (by MEGA 4.0)

	F	т	т	Р	L	G	L	Α	L	1	-	-	G	F	G	Т	Ľ	.	I F	V	L	F	Р	4															
✓ 1. cry4aa											-	-																											
2. cry4ba	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	•		-	-	т		E															
3. cry1aa	-	۷	Ρ	G	А		F	۷		G	-	-	L	V	D	1	1	٧	V G	1	F	G		\$															
4. cry2aa	L	L	к	κ	v		s	L	Т	G	к	R	1	L	S	E		٧	V G	1	Т			\$															
✓ 5. cry3aa	G	F	Ρ	F	G		Α	L	٧	S	-	-	F	Y	Т	N	F	: 1	LN	Т	Т	w		\$															
<ol> <li>6. cry3bb</li> </ol>	G	٧	Р	F	А		Α	L	т	s	-	-	F	Y	Q	s	F	: 1	LN	Т	Т	w		\$															
7. cry8ea1	G	v	Р	L	v		Р	Т	٧	s	-	-	L	Y	S		۰.		. 🖸			w		¢															
_	_																					l																	
	I.	L Y	v i	. :	6 9	5 N	( )	A C	2	A	A	N	L	н	L	т	v	L	-	_		- 1	т	к	Α	I	E C	Y	Т	Ν	Y	С	v	т	т	Y	K	к	G
1. cry4aa																			<b>v</b> 1	. cry	4aa	_																	
2. cry4ba	L	. 1	L	. 1	•	ī .				v			F	N		L	L	Т	2	. cry	4ba		٧	Q	Y	т	K E		Т	Α	н	s	Т		w		Ν		
3. cry1aa	V	P	L		. 1	1.	1	v .								s			_	. cry				R	L	. (	3 N			D		Α		R	w		Ν	т	
4. cry2aa	L	. 1	L		• I	. F	•						м			s	F	Т	_	. cry			R	N	Y	т	₹.		s				1	N			٩	т	Α
5. cry3aa	v	.	F	. 1	r 1	r .							т			F	L		_	. cry			L		L	т	2 E			D	н			к	w		N	v	
6. cry3bb	v	.	F		• 1	r .							т			L	L		_	. cry			L		L	т	2 0			D	н			N	w		N	v	
7. cry8ea1			L	-	_	1.	-									L	L		_	. cry			м	s	L		A C	_	s	D	н			Q	w		R	т	
	-															-	_							_	-														
	N	-	Q	-	-	-	<u>u</u>	S	Y	-	-	R	-	R	1 T		A :	5	N C	; -				L	I I	) k	- I	Е	F	L	Р	Т	т	R	s				
✓ 1. cry4aa		·	•	-	-	-	•	•	•		· ·	•	•	· ·	·		·   :		_	-																			
2. cry4ba		•	N	D	Ρ	Т	R	•	•	G	L	•	•	·	·		. /	٩.	_	S P				L		. F			Т	Т				Q		١			
				- 1	- 1	-		R	÷.	R	V	-	•	_		_	_		_	• •			-	Y		. F				۷		Α	Е	۷	т	1			
	_ P	-	S	-																								-											
✔ 4. cry2aa	G	N	s	-	-	-	-	-	·	N		Y	L	_	V	/ 5	5 .	· _	•	• •				D	L   I	A N		М		۷		Т	Ν	L	Ρ	1			
✓ 4. cry2aa ✓ 5. cry3aa	G	N Y				•	-	к	•	R	A	-		· H	_		-	-	_	-			-	D Y	-	и N	_	M	•	V I	•	T V	N N	L -	P -				
<ul> <li>✓ 3. cry1aa</li> <li>✓ 4. cry2aa</li> <li>✓ 5. cry3aa</li> <li>✓ 6. cry3bb</li> <li>✓ 7. cry8ea1</li> </ul>	G	N Y	s	-	-				•	-		-		н					•							-	•	-	-		•	-				-			

Some conserved blocks



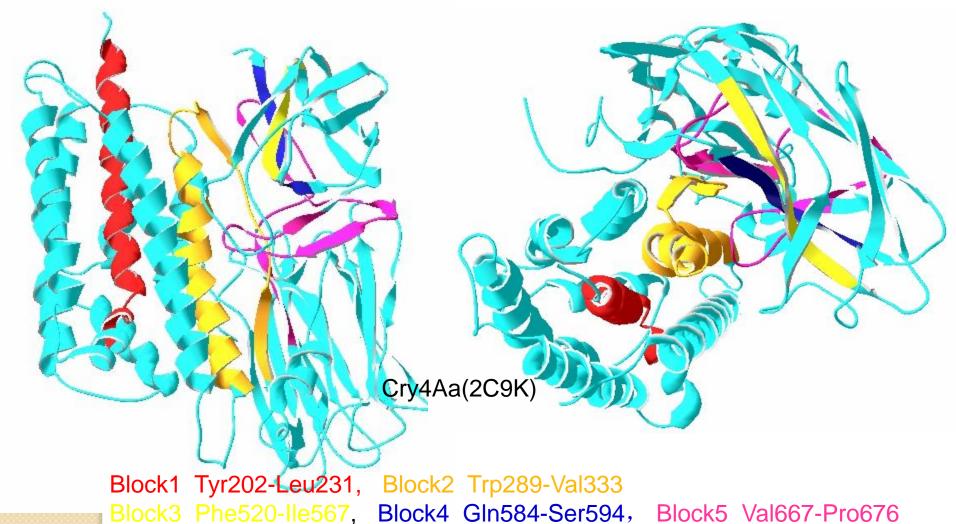
CryBE						000	2000000	1 00000000	0	000000	α2a	2	0000	a2b	00
Cry8Ea1 Cry3Aa Cry3Bb1 Cry1Aa Cry4Aa Cry4Ba Cry2Aa	MNNVLNSGRI	TICDA		PESFEHKSL	· · · · · · · · · · · · · · · · · · ·	TTKI 	OVIQKGI DAVGTGI TPIDISL SELSAYT	SVVGQILG SLTQFLLS IVVGTVLT	VVG.FPH VVG.VPH EFVPC GFGFTTH	FGGALVSF FAGALTSF GAGFVLGI PLGLALIC	VDIIWGII FGTLIPVI	GPSQ. FPAQD TPER.	DP <mark>W</mark> KA DPWKA WDA QSNT <mark>W</mark> SD	FLVQIEQI FITQTKN FMTNTGNI	LM LI LI LI
Cry8Ea1 Cry8Ea1 Cry3Aa Cry3Bb1 Cry1Aa Cry4Aa Cry4Ba Cry2Aa	QQQQQ 130 NQKIAEYARA DQKIADYAKN DKKIEEYAKS NQRIEEFARN KKEIASTYIS DQTVTAYVRT NQRINTDTLA	KALAE KALAE KALAE QAISR NANKI	EGLGNNY QGLQNNY QGLQNNY QGLQNNY EGLSNLY NRSFNYJ ATVVKDYI	VÊD <mark>YVS</mark> ALS FEDYVNALN VQIYAESFR ISTYHNHLK LDQ <mark>Y</mark> T <b>T</b> KFN	160 EWQENPS SWOKNPV SWKKTPL EWEADPTI TWENNPN TWKREPNI	SRNPH: SLRSKR: NPALRE PONTOD NOSYRT	179 DVRN SOGRIRE SQDRIRE EMRI VRI AVIT	1 RFEILDSL LFSQAESH LFSQAESH QFNDMNSA QIQLVHYH QFNLTSAK	FRNSMPS FRNSMPS LTTAIPI FONVIPE LRETAV	SFAI <mark>S</mark> G SFAV <b>S</b> K LLAVQN ELVN <b>S</b> CPF YFSNLVG.	NPSDCDY	EVPLL EVLFL EVLFL QVPLL (NILVL (ELLL	PTYAQAA SVYVQAA SSYAQAA PIYAQVA PLFAQAA	210 NLHLLLL NTHLFLL NTHLLLL NLHLSVL NLHLTVL NFNLLLI	KD KD RD RD RD RD
Cry8E Cry8Ea1 Cry3Aa Cry3Bb1 Cry1Aa Cry4Aa Cry4Aa Cry4Aa	220 ASIFGEEWGE AQIYGEEWGY AQVFGEEWGY VSVFGQRWGE AVKFEAYLKN GLINAQEWSI VILNADEWGI	EK SS DA INRQFD AR	LEPLPT	NSRYNDLT	240 SLIAQYS KLTQEYT KLTQQYT RLIGNYT KAIEDYT QYTKEYI	2 DHCVQW DHCVKW DHCVNW DHCVNW DYAVRW NYCVTT AHSITW	50 (RT <mark>GL</mark> DR (NVGLDK (NVGLNG (NTGLER (KKGLNL (NKGLDV	260 LKGSNAK. LRGSSYE. LRGSTYD. VWGPDSR. IKTTPDSN LRNKSNG.	Q 	VNFNRYF VKFNRFF VVRYNQFF NNTYNTYF NITFNDYF	28 REMTLSVI REMTLTVI REMTLTVI RELTLTVI RELTLTVI RTKMTTAVI	DIMTL DLIAL DLIVL DIVAL DLVAL DLVAL	FPEYDVR: FPFYDIR FSNYDSRI FPNYDVGI FASYDPRI MVSSGAN	LYPKE LYSKG RYPIR KYPIG RYPADKII LYASG	• • • • • •
Cry8E Cry8Ea1 Cry3Aa Cry3Bb1 Cry1Aa Cry4Aa Cry4Ba Cry4Ba	220 ASIFGEEWGE AQIYGEEWGY AQVFGEEWGY VSVFGQRWGE AVKFEAYLKN GLINADEWGI VILNADEWGI	(EK (SS DA INRQED JAR	TA ED ED AT YLEPLPT	LAEFYKRQL VAEFYHRQL INSRYNDLI AIDY <mark>Y</mark> PVLI	240 SLIAOYS KLTQEYT KLTQQYT RLIGNYT KAIEDYT QYTKEYI	2 DHCVQW DHCVKW DHCVNW DYAVRW NYCVTT AHSITW	50 YRIGLDH YNVGLDH YNVGLN YNIGLEH YKKGLNI YNKGLDV	LRGSSYE. LRGSTYD. VWGPDSR. IKTTPDSN LRNKSNG.	Q 	270 WVEYNRF WVNFNRY WVKFNRF WVRYNQF WNTYNTY WITFNDY	2 RREMTLSV RREMTLTV RREMTLTV RRELTLTV RTKMTTAV KREMTIQV	LDIMTI LDLIAI LDLIVI LDLVAI LDLVAI LDLVAI	290 FPMYDMH FPLYDVH FPFYDIH FSNYDSH FPNYDVO FASYDPH	RTYPME RLYPKE RLYSKG RRYPIR SKYPIG RRYPADKI	   IDN
Cry8E Cry8Ea1 Cry3Aa Cry3Bb1 Cry1Aa Cry4Aa	β] 300 TKAQLTE VKTELTE VKTELTE TVSQLTE VQSELTE	β2 SEVYTD DVLTD DIFTD EIYTN REIYQV	IQ PIGAIGA( PIVGVN. PIFSLN. PVLENFD( LNFEES.	nl 320 QGSWYDSAP .NLRGYGT .TLQEYGP GSFRG PYKYY	QS QQQQQQ 339 SFNTLES TFSNIEN TFLSIEN MAQRIEQ DFQYQED	QQ IFIRGK YIRKP NIRCP SLTRRP	340 HLFDFI HLFDYL HLFDYL HLFDYL HLFTWL	β3 350 IRLSIYTGE IRLQFHTRE IGIEFHTRI ISITIYTD OSLNFYEKA	SS.FSA OPGYYG OPGYFG VHR	360 SNYLKKW NDSFNYW KDSFNYW GFNYW TPNNF	β4 370 IGHQISSQ SGNYVSTR SGNYVETR SGHQITAS FTSHYNMF	PIGG <mark>S</mark> PSIGS PSIGS PVGFS HYTLD	β5 38 Q LQTQTYGT NDIITSPI SKTITSPI SPEFAFPI NISQKSST	390 TSGS.SV TSGS.SV TSGNK.SS TYGDK.SS TYGDK.GN FGNHNVT	SEP TEP AAP TDK
Cry4Ba Cry2Aa	TKLSKTEF			SKSI SLFQVNSNY											



Multiple alignment by DSSP: 5 conserved blocks

Shuyuan Guo, et al. Journal of Structural Biology. 168 (2009) 259–266

#### The location of the conserved blocks



**Conclusion 3:** All conserved blocks located in the *center* of molecular or the *interfaces* between 2 domains. Maybe important for the structure and stability.

## Comparison of domain I/ II between Cry4Aa and Cry4Ba

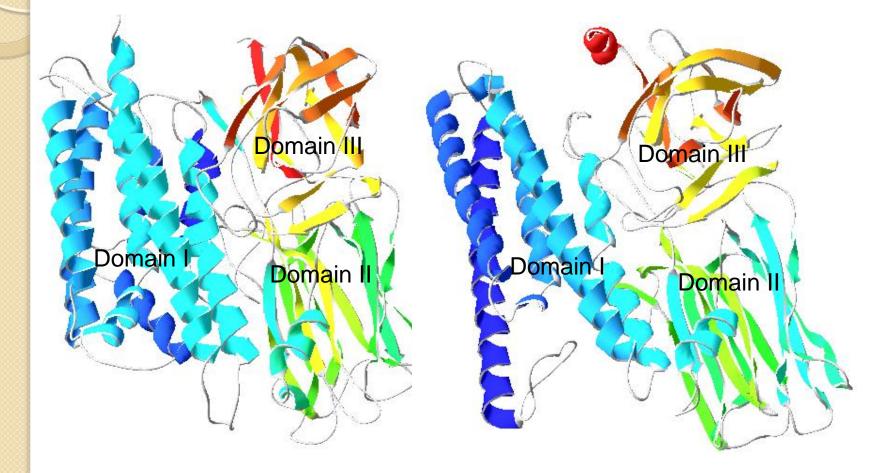
 

 Table 1. Larvicidal activity of Bacillus thuringiensis svar. israelensis powders containing individual spore/crystals mixtures, against Aedes aegypti early 4<sup>th</sup> instar larvae

Powder	No. assays	No. Iarvae	LC <sub>50</sub> ª (95% fiducial limits)
Crv11Aa	3	1140	1.35 (1.01-1.82)
Cry4Aa	3	1020	13.01 (8.82-20.04)
Cry4Ba	3	1020	0.12 (0.08-0.54)
Bti <sup>b</sup>	3	1080	0.013 (0.011–0.016)

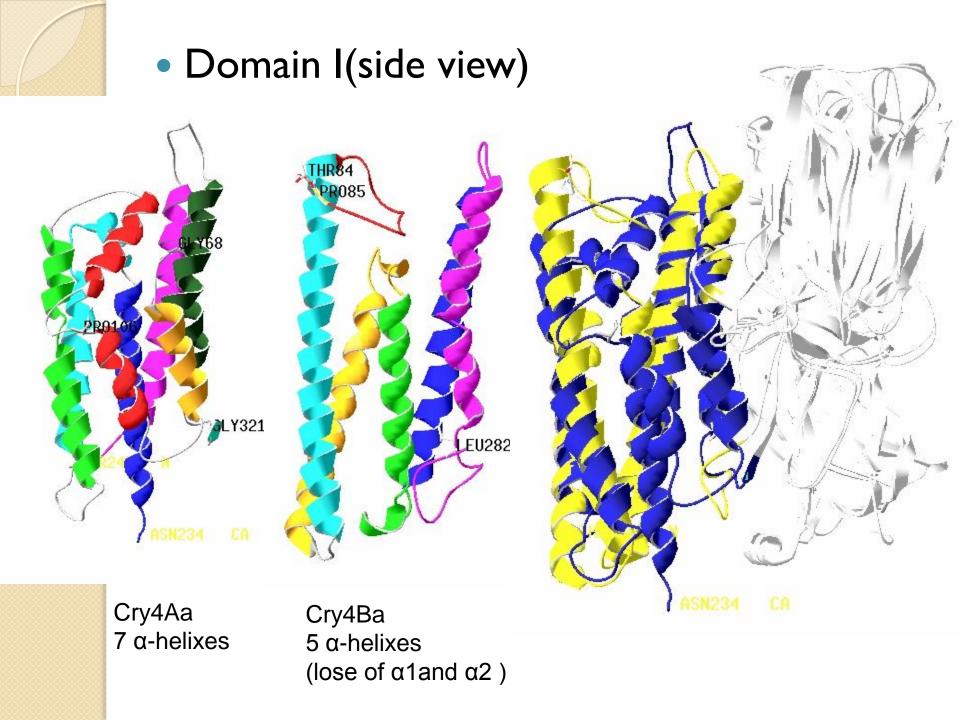
Henrique de Barros Moreira Beltrao, Maria Helena Neves Lobo Silva-Filha. *FEMS Microbiol Lett* 2007 (266) 163–169

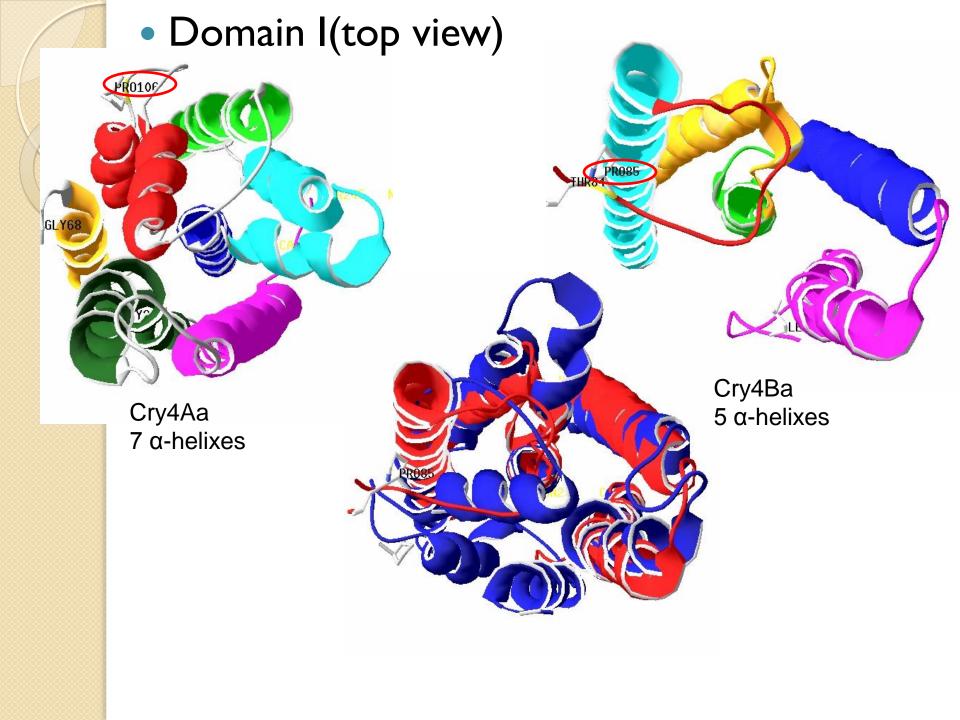
	Domain I	Domain II	Domain III
Cry4Aa	68-321	322-524	525-679
Cry4Ba	84-282	283-466	467-641



#### Cry4Aa(2C9K)

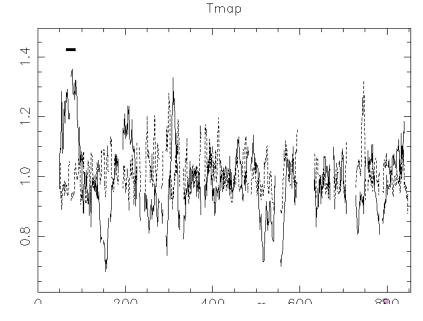
Cry4Ba (1W99)





 α-4 and α-5 are the keys to pore-forming in the midgut membrane, maybe because of their high hydrophobicity

## • T-map of Cry4Aa toxin(by Weblab)



This amino acid sequence is of a MEMBRANE PROTEIN which have 1 transmembrane helix.

No.	N	terminal	transmembrane region	С	terminal	type	length
1		14	VGTVLTGFGFTTPLGLALIGFGT		36	PRIMARY	23

The hydrophobic section located in the  $\alpha$ -1 and  $\alpha$ -2

(Î)

V

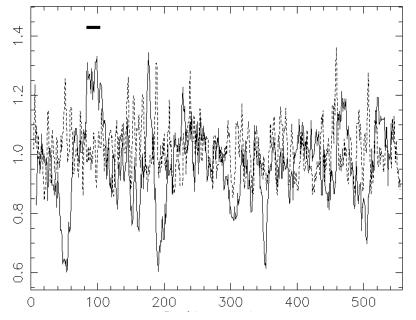
L

 $\langle \hat{T} \rangle$ 

Π

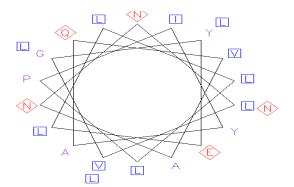
The helical wheel of Cry4Aa α-5 shows its hydrophobicity. 208-SSYAQAANLHLTVLNQAVKFEA

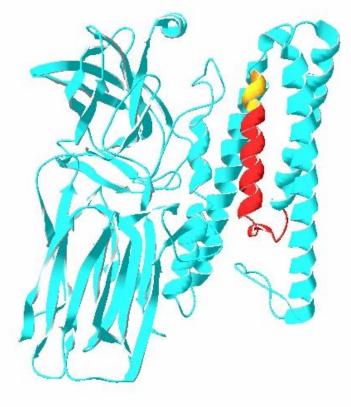
## • T-map of Cry4Ba toxin(by Weblab)



This amino acid sequence is of a MEMBRANE PROTEIN which have 1 transmembrane helix.

No. N	terminal	transmembrane region	C terminal	type	length
1	4	NLVGYELLLLPIYAQVANFNLLL	26	PRIMARY	23

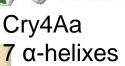




The hydrophobic section located in the  $\alpha$ -5

#### **Conclusion4** :α-5 are highly hydrophobic

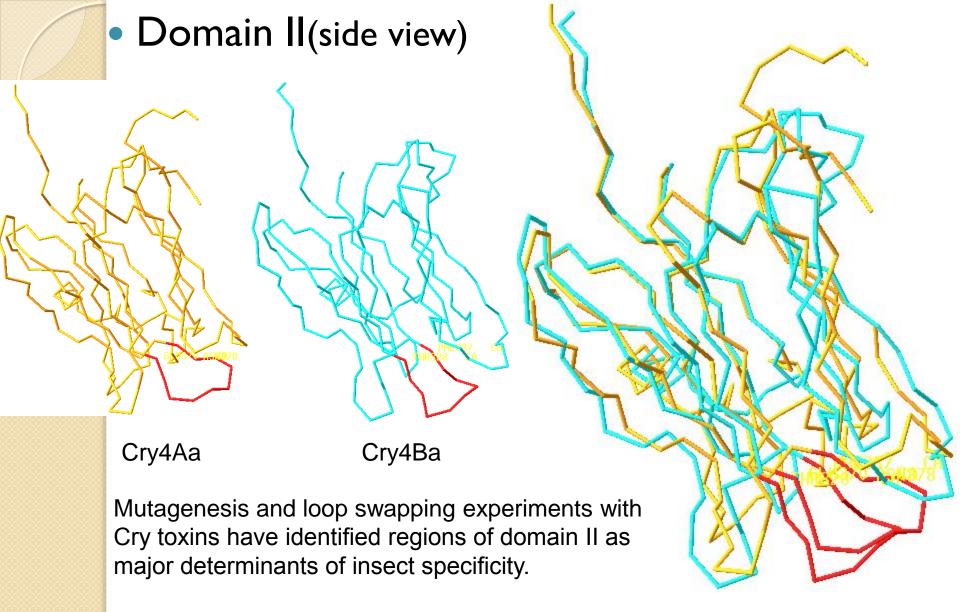
### Domain I(top view)



Cry4Ba 5 α-helixes

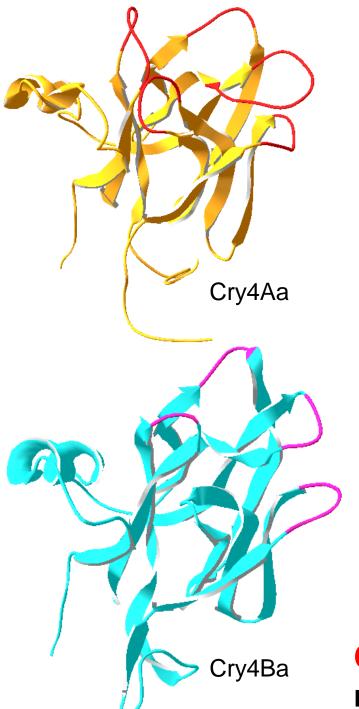
The  $\alpha$ -1 and  $\alpha$ -2 maybe not necessary for the toxicity of Cry toxin

Conclusion 5: the conserved Pro may play a important role as the link of the "lip", and maybe involved in the stability of Cry toxin. Shuyuan Guo, et al. *Journal of Structural Biology.* 168 (2009) 259–266



Poncet, S. et al. J. Invertebr. Pathol.1995,66:131-135.

Abdullah, M. A, et al. *Appl. Environ. Microbiol.* 2003,69:5343–5353.



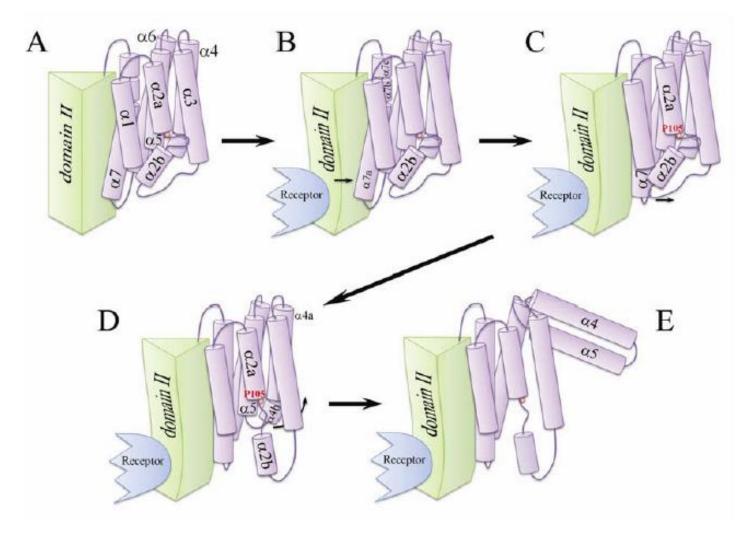


**Conclusion 6:** The loops in domain II major determine the insect specificity.

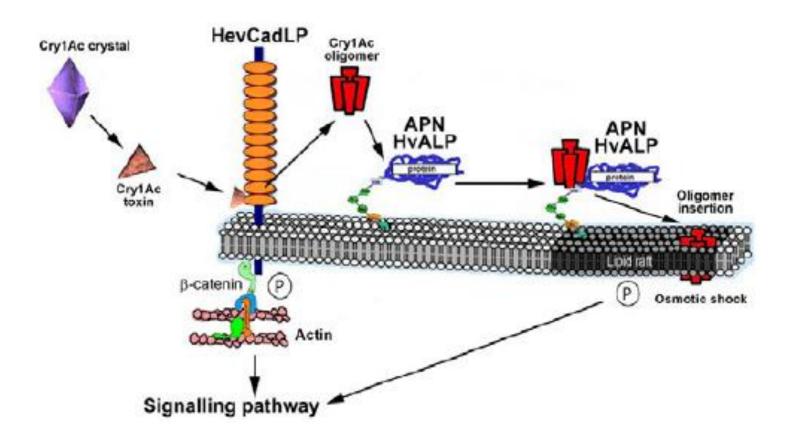
## Summary:

- The acidity of the Cry proteins make it easier to dissolve in the midgut where is alkaline.
- The 3D structures of the Cry toxins share high similarity.
- The conserved sections located in the center of molecular or the interfaces between 2 domains. Maybe important for the structure and stability.
- α-5 are highly hydrophobic thus it can easily insert into the midgut membrane
- The conserved Pro play a important role as the link of the "lip" which may protect the  $\alpha$ -5 .
- The loops in domain II major determine the insect specificity and activity.





Shuyuan Guo, et,al. Journal of Structural Biology 168 (2009) 259–266



Jurat-Fuentes J.L. ,Adang M.J . *Journal of Invertebrate Pathology* 2006,92(3):166~171.



## Acknowledgements

This work was done by the hands of my team members. Thank them for their hardworking and help.

Appreciate my classmates for their meaningful suggestions.

Pro. Jingchu Luo guided me to the gate of bioinformatics, thank him very much!

Thank you!