Supporting information

For submission to: Journal of Agricultural and Food Chemistry

Functional Characterization of two Carboxylesterase Genes Involved in Pyrethroid Detoxification in *Helicoverpa armigera*

Yong-Qiang Li,[†] Li-Sha Bai,[†] Cai-Xia Zhao,[†] Jing-Jing Xu,[†] Zhong-Juan Sun,[†] Yan-Ling Dong,[‡] De-Xian Li,[‡] Xi-Li Liu,[†] Zhi-Qing Ma*[,†, §]

[†] State Key Laboratory of Crop Stress Biology for Arid Areas, College of Plant Protection, Northwest A&F University, Yangling 712100, Shaanxi, China

[‡] College of Plant Protection, Northwest A&F University, Yangling 712100, Shaanxi, China

[§]Engineering and Research Center of Biological Pesticide of Shaanxi Province, Yangling 712100, Shaanxi, China

Corresponding Author: Zhi-Qing Ma

***Tel.:** 86-13709266095**, e-mail:** zhiqing-ma@hotmail.com.

Table S1. Primers for cloning of CarEs cDNA and genomic DNA from *H. armigera*

Genes	Primer	Sequence (5′−−3′)
CarE001A	aF1	ATGAAGTGGTGGACGTGT
(cDNA)	aR1	TAGACCATCGGTTGGCTAG
Car001H	hF1	GCCTCGATGACCGATTTG
(cDNA)	hR1	CGGACAATGGTGTAACCTCGT
CarE001A	aF2	CCGTAAGTCAGTTTGCAACCGTAGCGAT
(genomic DNA)	aR2	ACTATTCACACTCGACT
Car001H	hF2	CTGGAGCTCGATGACCGATTTGAAACCA
(genomic DNA)	hR2	CGGACAATGGTGTAACCTCGT

Table S2. Primers for quantitative RT-PCR analysis of *H. armigera*

Genes	Primer	Sequence (5′−−3′)	GenBank No.
CarE001A	Forward	CCCTGATGATGTTACTATTGC	KT345936.1
	Reverse	AGATTTCCGCCACTTTCG	
CarE001H	Forward	GAAAGGGATGAGATTGGT	KT345937.1
	Reverse	AAAGTAAGATCGCCGAAG	
EF-1α	Forward	GACAAACGTACCATCGAGAAG	U20129
	Reverse	GATACCAGCCTCGAACTCAC	
ACTA3a	Forward	CTCGACTTCGAGCAGGAGAT	X97614
	Reverse	TTCCATACCCAGGAATGAGG	
RPS15	Forward	CTGAGATGATCGGCCACTAC	AY818611
	Reverse	TGTTGGTCAGCGCACTACTT	

Table S3. **Specific primers for amplification of CarEs without the Signal Peptide sequence.**

Gene	Primer	Sequence (5′--3′) [a]	Endonuclease site
CarE001A	aF3	CCG**GATATC***CCATGG*ATGGACGACGAGTGGCGCGAGGTGA[b]	***EcoR* V** / *Nco* I
	aR3	GCG**AAGCTT***CTCGAG*CTACAACTCGTTGCGTGGTCTG	***Hind* III** / *Xho* I
Car001H	hF3	CCG**GAATTC***GATATC*ATGGATGATAATGAATGGCGGC	***EcoR* I** / *EcoR* V
	hR3	GCG**AAGCTT***CTCGAG*GTTCATAGTTCTGTGTGTTTCTCCG	***Hind* III** / *Xho* I

[a] The nucleotide sequences in bold and italic indicate the two restriction endonuclease recognition sites, respectively.

[b] The specific primers corresponding to the CarE ORFs were underlined with removal of signal peptide sequence.

Table S4. Exon/intron size and junctional sequences of CarE genes in *H. armigera*.

Gene	Exon number	Exon size	3' splice site	Exon squences	5' splice site	Intron size
CarE001A	1	177	ACCGTAGCG	ATG AAG TGG … AAG TTC AAG	gtatttata	1051
	2	1283	ttgttttag	GCA CCT CTT … GAA AAC TGG	gtacgtttt	197
	3	208	acgtttcag	AGT GCC AGT … GAG TTG TAG	AAATTGTAA	
CarE001H	1	180	TGGAGCTCG	ATGAGGCGA… AGG TTT AAG	gtaagtctt	235
	2	1277	aattttcag	GCA CCT TAC … TAC TAA AGG	gtaagaaac	267
	3	211	tatttttcag	AAA ACC TGT … GAA CTA TGA	ACAAATTAA	

Figure Legends

Figure S1. An unrooted distance neighbor-joining tree revealing the phylogenetic relationships of insect carboxylesterases. MEGA 7 was used to construct the phylogenetic tree. Nodes with >50% bootstrap support (1000 pseudo replicates) are indicated. Clades are designated according to the nomenclature adopted by Oakeshott et al. The studied CarEs in this work are highlighted with solid squares and bold fonts. Solid rhombuses indicate the sequences from *B. mandarina* and *B. mori*, hollow rhombuses indicate the sequences from *L. migratoria*, solid triangles indicate the sequences from *Anopheles gambiae*, solid cirles indicate the sequences from *D. melanogaster*, hollow cirles indicate the sequences from *C. quinquefasciatus*, hollow squares indicate the sequences from *Aedes aegypti*, *B. dorsalis*, *M. domestica* and *L. cuprina*, respectively.

Figure S2. Comparison of amino acid differences of CarEs from different *H. armigera* strains (WH, GR, YGF), as well as HassAE2 from *H. assulta*. (A, C) CarE001A. (B) CarE001H. Amino acids the same as that of the GR strain are indicated with a dot. Numbering indicates the alignment number of different residue amino acids.

Figure S3. Sequence analysis of genomic organization and transcript mRNA of *CarE001A* (A) and *CarE001H* (B) from *H. armigera*. Exons are shown as colorful rectangles and introns as black and horizontal arrowed lines. Number indicate the length (bp) of each exon or intron.

Figure S4. Enzymatic reaction of a serial diluted recombinant carboxylesterases with the model substrate α-naphthyl acetate (α-NA) at 200 μM. Formation of *α*-naphthol was monitored by recording the change in the absorbance at 450 nm for 5 min using a microplate reader (M200 PRO, Switzerland). Sample of recombinant CarEs were diluted at 2-fold, 4-fold, 8-fold, 16-fold, 32-fold, 64-fold in the reaction mix. Sample of pET32a vector expressed in *E. coli* and the *E. coli* only were used as double negative controls.

Figure S5. Kinetic assay of recombinant carboxylesterases and substrate α-NA at a serials of concentrations (16 - 200 μM). Formation of *α*-naphthol was monitored by recording the change in the absorbance at 450 nm for 5 min using a microplate reader (M200 PRO, Switzerland). Sample of

recombinant CarEs were at 8×diluted concentration. The kinetic parameters were estimated according the initial velocity of enzymatic reaction at different substrate concentrations.

Figure S6. Liquid chromatograms at different times in the assay of metabolic activity with three pyrethroids. (A) β-cypermethrin. (B) λ-cyhalothrin. (C) Fenvalerate. Reactions were conducted at 30 ℃ for 0 - 120 min and stopped by the additional of 200 μL of absolute acetonitrile. The reactions of 0 min were set up by adding 200 uL of acetonitrile to the microfuge tube before the addition of active enzyme. Sample of vector pET32a without insertion of target gene expressed in *E. coli* were used as a negative control.

Figure S7. Three dimensional (3D) structures of CarE001A and CarE001H. (A) CarE001A. (B) CarE001H. 3D structure were predicted with the Modeller 9.20 software using the structure of *Homo sapiens* butyrylcholinesterase (PDB ID: 4TPK, Resolution = 2.7 Å) as template.

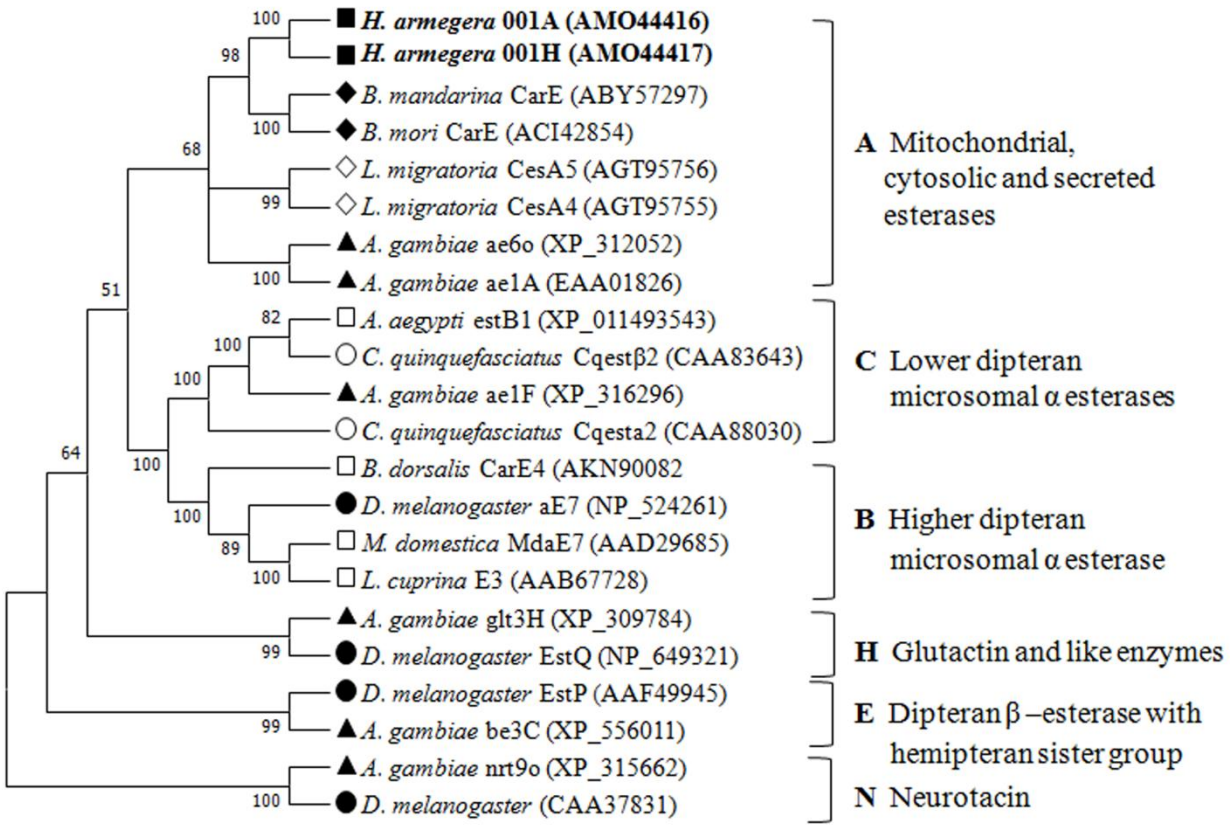

H. armegera 001A (AMO44416)
H. armegera 001H (AMO44417)
B. mandarina CarE (ABY57297)
B. mori CarE (ACI42854)
L. migratoria CesA5 (AGT95756)
L. migratoria CesA4 (AGT95755)
A. gambiae ae6o (XP_312052)
A. gambiae ae1A (EAA01826)
A. aegypti estB1 (XP_011493543)
C. quinquefasciatus Cqestβ2 (CAA83643)
A. gambiae ae1F (XP_316296)
C. quinquefasciatus Cqesta2 (CAA88030)
B. dorsalis CarE4 (AKN90082
D. melanogaster aE7 (NP_524261)
M. domestica MdaE7 (AAD29685)
L. cuprina E3 (AAB67728)
A. gambiae glt3H (XP_309784)
D. melanogaster EstQ (NP_649321)
D. melanogaster EstP (AAF49945)
A. gambiae be3C (XP_556011)
A. gambiae nrt9o (XP_315662)
D. melanogaster (CAA37831)
A Mitochondrial, cytosolic and secreted esterases
C Lower dipteran microsomal α esterases
B Higher dipteran microsomal α esterase
H Glutactin and like enzymes
E Dipteran β –esterase with hemipteran sister group
N Neurotacin

Figure S1

A

		10	19	20	30	38	42	67	92	165	183	488	524	546
GR	ADF43455	G	K	G	G	E	T	G	K	E	V	T	A	K
WH	AMO44416	A	E	W	V	A	A	W	·	D	·	V	P	·
YGF	ADJ96631	M	E	W	V	·	·	W	T	·	·	V	·	T
-	XP_021195358	V	E	W	V	·	·	W	·	·	V	·	·	

B

		16	39	43	57	255	271	282	293	305	307	361	389	440	466	518	536	550
GR	ADF43464.1	G	G	A	E	F	T	S	P	D	S	K	N	G	K	G	N	D
WH	AMO44417	C	A	T	E	I	A	L	·	·	·	N	K	·	E	N	K	E
-	XP_021181761	·	·	A	D	·	·	·	A	E	Y	·	·	R	E	N	K	E

C

		10	38	42	55	84	92	165	275	341	462	495	521	546
H. assulta	ATJ44546	V	E	T	R	K	T	E	Q	F	E	R	V	T
WH	AMO44416	A	A	A	Q	M	K	D	K	L	D	W	A	K

Figure S2

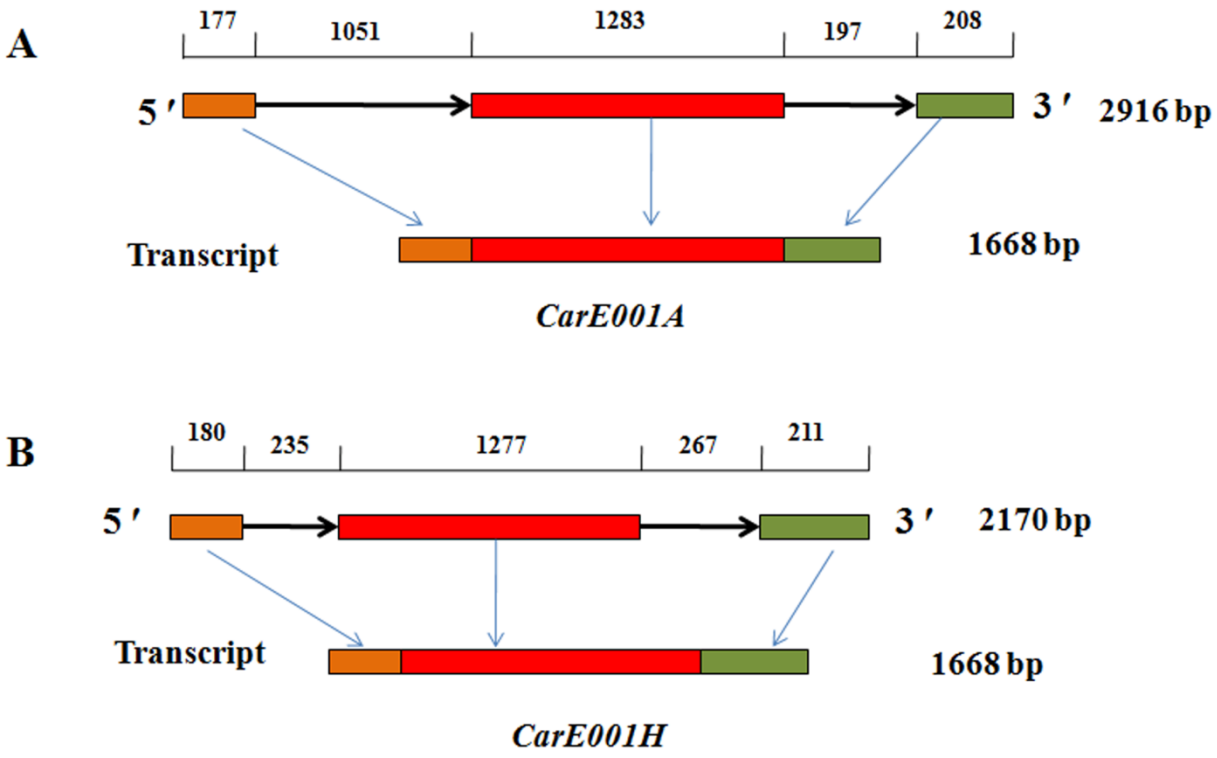

A
177
1051
1283
197
208
5′
3′
2916 bp
Transcript
1668 bp
CarE001A
B
180
235
1277
267
211
5′
3′
2170 bp
Transcript
1668 bp
CarE001H

Figure S3

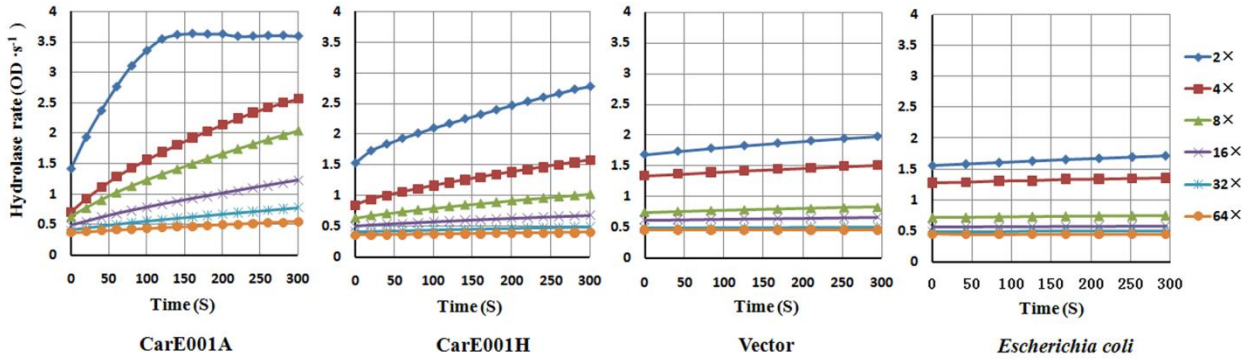

Hydrolase rate (OD ·s-1)
Time (S)
CarE001A
CarE001H
Vector
Escherichia coli
2×
4×
8×
16×
32×
64×

Figure S4

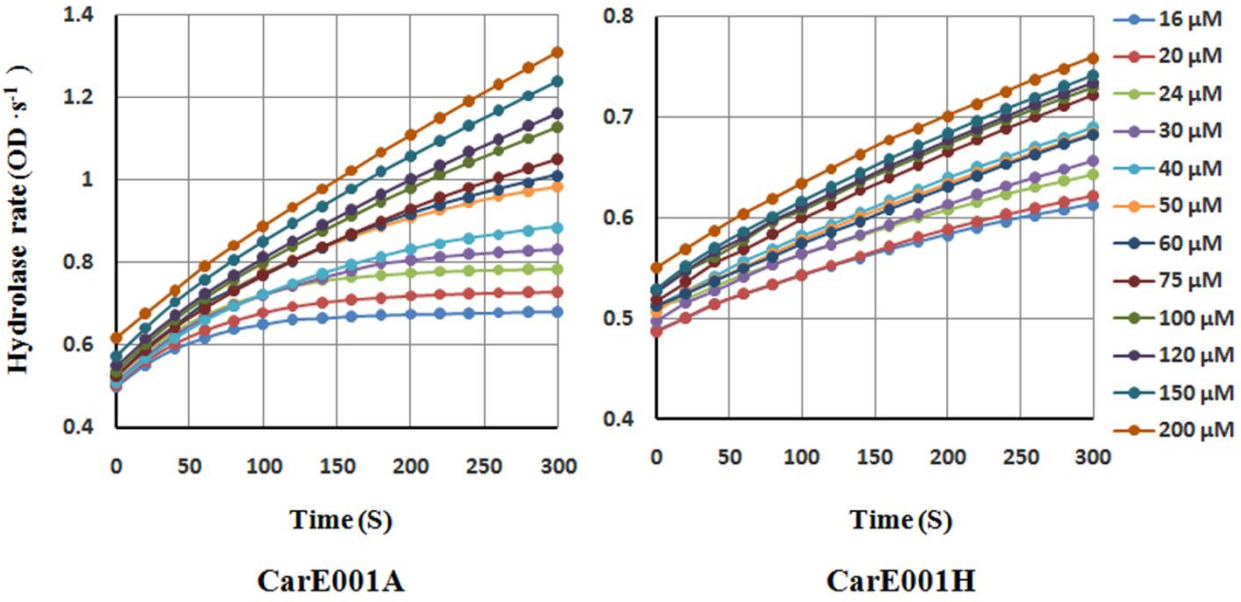
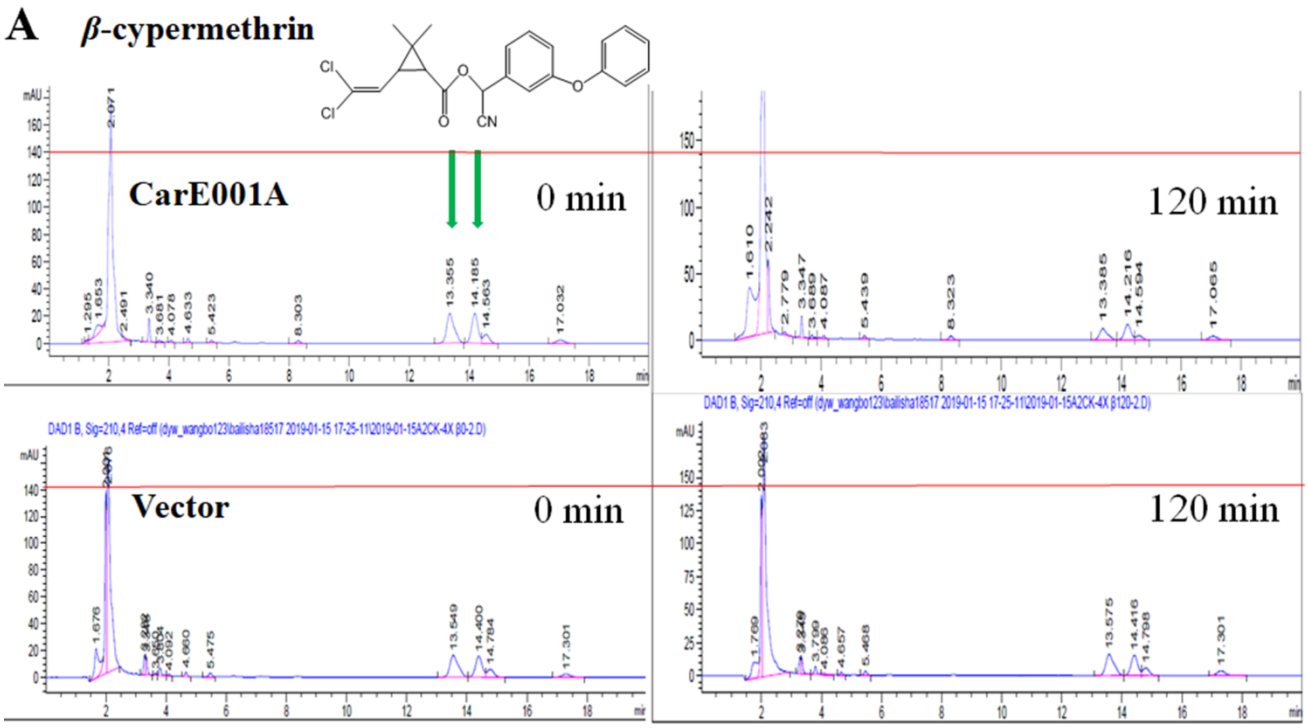
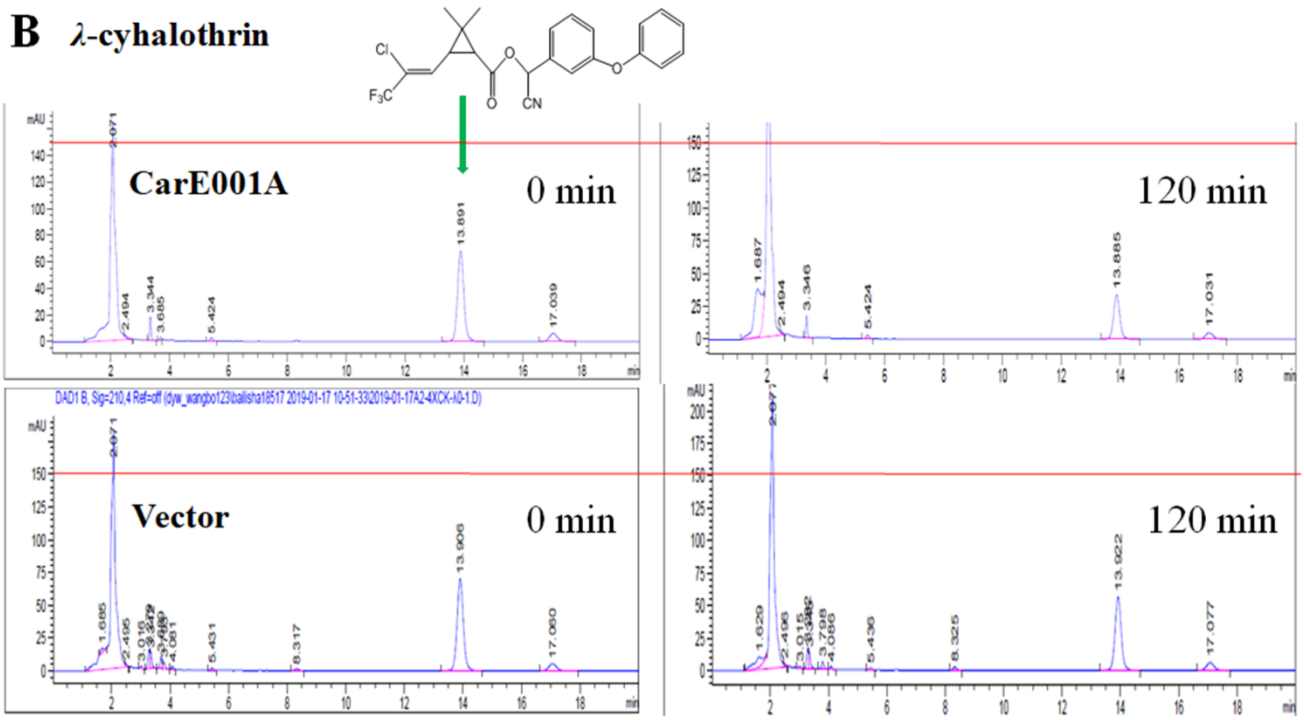

Hydrolase rate (OD ·s-1)
1.4
1.2
1
0.8
0.6
0.4
0
50
100
150
200
250
300
Time (S)
CarE001A
0.8
0.7
0.6
0.5
0.4
0
50
100
150
200
250
300
Time (S)
CarE001H
16 μM
20 μM
24 μM
30 μM
40 μM
50 μM
60 μM
75 μM
100 μM
120 μM
150 μM
200 μM

Figure S5


A
β-cypermethrin
Cl
Cl
O
O
O
CN
CarE001A
0 min
120 min
Vector
0 min
120 min

B *λ*-cyhalothrin



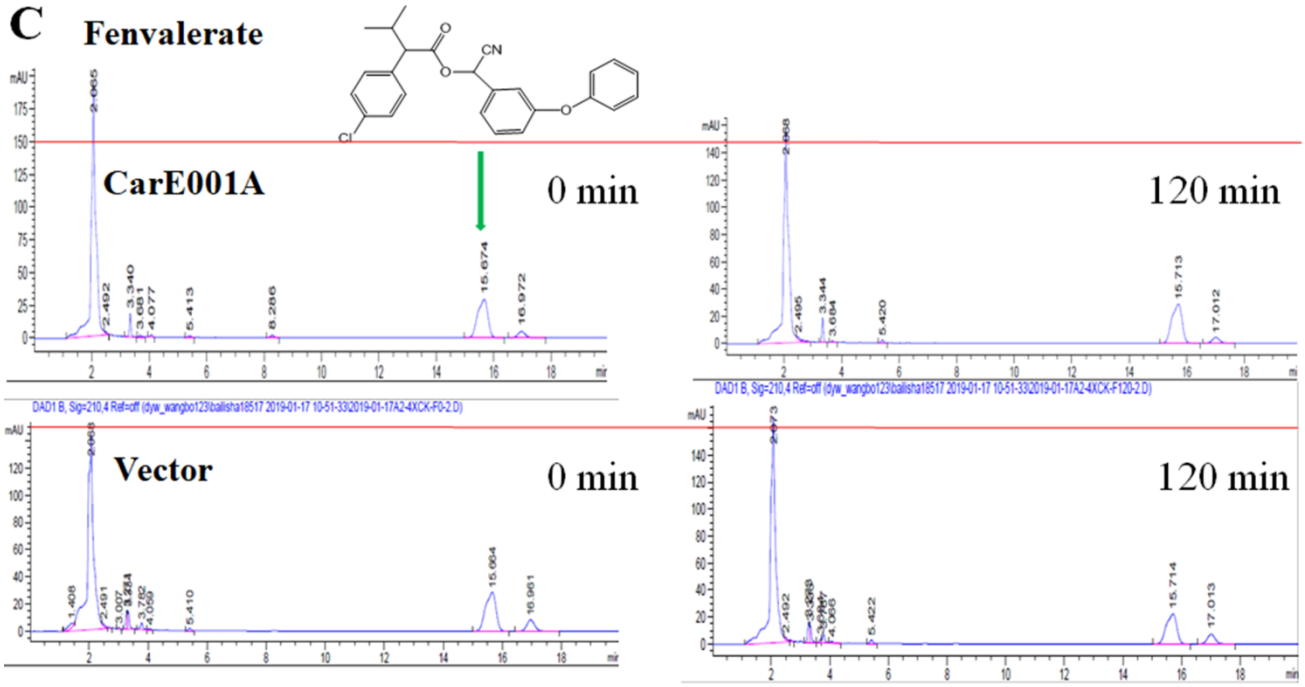


Figure S6

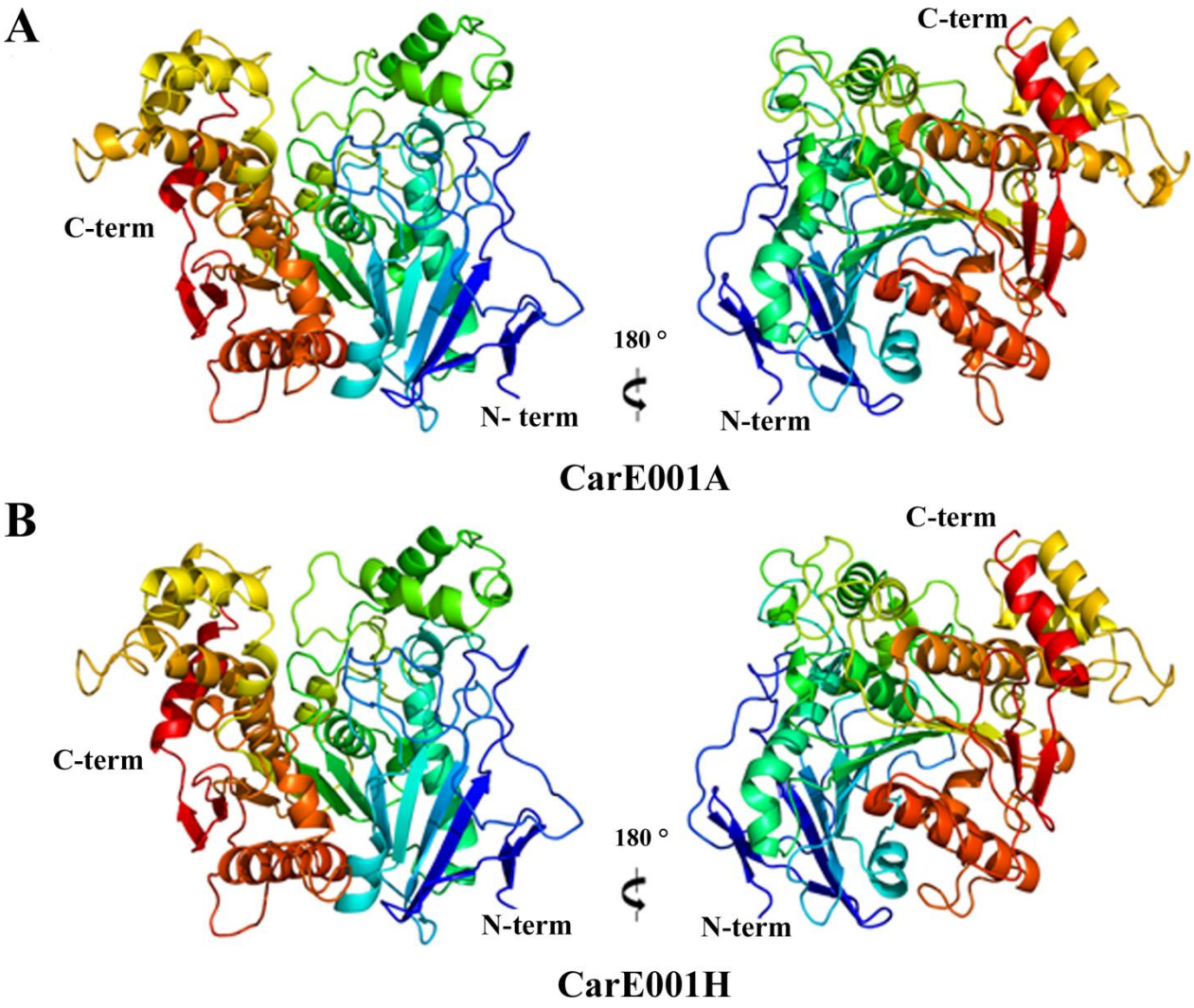

A
C-term
C-term
N- term
180 °
N-term
CarE001A
B
C-term
C-term
N-term
180 °
N-term
CarE001H

Figure S7