

# Enhanced hydrolysis of $\beta$ -cypermethrin caused by deletions in the glycine-rich region of carboxylesterase 001G from *Helicoverpa armigera*

Li-Sha Bai,<sup>a,b†</sup> Jing-Jing Xu,<sup>a,b†</sup> Cai-Xia Zhao,<sup>a,b</sup> Ya-Li Chang,<sup>a,b</sup>  
 Yan-Ling Dong,<sup>a,b</sup> Kai-Ge Zhang,<sup>a,b</sup> Yong-Qiang Li,<sup>a,b\*</sup> Yi-Ping Li,<sup>a,b</sup>  
 Zhi-Qing Ma<sup>a,b</sup> and Xi-Li Liu<sup>a,b</sup>



## Abstract

**BACKGROUND:** Carboxylesterase (CarE) is a major class of enzyme involved in the detoxification of toxic xenobiotics in various insect species. Previous work has shown that the carboxylesterase gene *CarE001G* found in *Helicoverpa armigera* is more active and can metabolize synthesized pyrethroids, such as  $\beta$ -cypermethrin, one of the commonly used commercial insecticides for lepidopteran pest control. In addition, *CarE001G* is very special as it has a very specific glycine-rich region located adjacent to its C-terminal. But whether mutations in this unique sequence can change the biochemistry and function of *CarE001G* are unknown.

**RESULTS:** In this study, four variants of *CarE001G* with different deletions in the glycine-rich region were obtained and functionally expressed in *Escherichia coli*. The recombinant proteins were purified and confirmed by Western blot and mass spectrometry analyses. These mutant enzymes showed high catalytic efficiency toward the model substrate  $\alpha$ -naphthyl acetate. Inhibition study showed that  $\beta$ -cypermethrin had relatively strong inhibition on CarE activities. *In vitro* metabolism assay showed that the mutant enzymes significantly enhanced their metabolic activities toward  $\beta$ -cypermethrin with specific activities between 4.0 and 5.6 nmol L<sup>-1</sup> min<sup>-1</sup> mg<sup>-1</sup> protein. Molecular docking analyses consistently demonstrated that deletion mutations in the glycine-rich region may facilitate the anchoring of the  $\beta$ -cypermethrin molecule in the active binding pocket of the mutant enzymes.

**CONCLUSION:** The data show that deletion mutations can cause qualitative change in the capacity of CarEs in the detoxification of  $\beta$ -cypermethrin. This indicates that deletion mutations in the glycine-rich region may have the potential to cause synthesized pyrethroid (SP) resistance in *H. armigera* in the future.

© 2020 Society of Chemical Industry

Supporting information may be found in the online version of this article.

**Keywords:** *Helicoverpa armigera*; carboxylesterase; deletion mutation; metabolic activity; insecticide resistance

## 1 INTRODUCTION

The cotton bollworm, *Helicoverpa armigera* (Hübner), is a devastating pest feeding on various kinds of economical crops such as cotton, corn, and soybean, and causes severe yield losses on them globally.<sup>1,2</sup> This is largely because cotton bollworms have the following important inherent characteristics: they are polyphagous, undergo facultative diapause and high fecundity and mobility.<sup>2</sup>

Organophosphates (OPs) and synthesized pyrethroids (SPs) are two major commercial chemical insecticides which have been frequently used for cotton bollworm control over the past 30–40 years.<sup>3</sup> Intensive spraying of insecticides in farmlands has led to the evolution of insecticide resistance in cotton bollworms since the 1980s.<sup>1,4,5</sup> Genetically-engineered crops producing *Bacillus thuringiensis* (BT) insecticidal proteins have been planted

widely for controlling cotton bollworm since 1996.<sup>6</sup> However from 2005, the evolution of resistance to transgenic BT cotton in the cotton bollworms in China has been reported.<sup>7–9</sup> Although the extensive planting of transgenic crops suppressed pests and reduced chemical insecticides use over the last two decades,<sup>10,11</sup>

\* Correspondence to: Y-Q Li, College of Plant Protection, Northwest A&F University, Yangling 712100, Shaanxi, China. E-mail: yongqiangli@nwfau.edu.cn

† These two authors contributed equally to this work.

a College of Plant Protection, Northwest A&F University, Yangling, China

b State Key Laboratory of Crop Stress Biology for Arid Areas, Northwest A&F University, Yangling, China

it has been widely reported that cotton bollworms have gained metabolic resistance to SPs and OPs worldwide.<sup>11–14</sup> For example, recent studies have shown that field strains of *H. armigera* collected from northern and north-eastern China have developed high-level fenvalerate resistance and moderate-level  $\lambda$ -cyhalothrin resistance, respectively.<sup>12,15</sup> The metabolic resistance is mainly caused by cytochrome P450s (CYPs) and carboxylesterases (CarEs), which are two multigene superfamilies playing key roles in the detoxification of xenobiotic chemical compounds among insect pests including the cotton bollworm.<sup>5,16–19</sup>

CarEs (EC3.1.1.1) are esterases under the main enzyme group known as hydrolases. They hydrolyze carboxylic esters.<sup>20</sup> CarEs have the catalytic triad (Ser-His-Glu) in their active sites, which is essential for their hydrolytic capability and are highly conserved in insect CarEs.<sup>21</sup> The other residues in the oxyanion hole and the nucleophilic elbow (Gly-X-Ser-X-Gly) termed as pentapeptide are also necessary for their catalytic hydrolysis.<sup>21</sup> CarEs are ubiquitous in various organisms including mammals and insects, and play crucial roles in the detoxification of xenobiotic chemicals including various drugs and insecticides.<sup>22</sup>

The involvement of CarEs in the detoxification of different agrochemicals has been reported in insect species, such as in *Myzus persicae* (Sulzer),<sup>23</sup> *Culex* mosquitoes,<sup>24</sup> *H. armigera*,<sup>25</sup> *Lucilia cuprina*,<sup>26</sup> *Cydia pomonella*,<sup>27</sup> *Bactrocera dorsalis*,<sup>28</sup> and *Locusta migratoria*.<sup>29</sup> Meanwhile, other studies have implicated CarEs in insects resistance to insecticides.<sup>2,17,23</sup> For examples, in *Myzus persicae* (Sulzer), amplifications of *EF4* and *E4* were found to be responsible for insecticide resistance in different field populations;<sup>23</sup> and in *Nilaparvata lugens*, more copies of the *N1-EST1* gene were detected in resistant strains.<sup>30</sup> Additionally, in *Aphis gossypii* and *B. dorsalis*, high messenger RNA (mRNA) expression levels were also found to be associated with deltamethrin and malathion resistance,<sup>31,32</sup> respectively. Besides these, in *Lucilia cuprina* and *Musca domestica*, single point mutations of CarE genes (i.e. Gly137 to Asp or Trp251 to Leu of *E3* and *E7*) were reported to also cause OP resistance, respectively.<sup>33–35</sup> Therefore, two general molecular mechanisms for CarE-mediated metabolic resistance are currently well known in insects.<sup>20,21</sup> The first involves overproduction of CarEs via (i) amplification of individual CarE genes at the DNA level or (ii) increased transcripts of a specific CarE gene at the mRNA level.<sup>20</sup> It is generally assumed that the overproduction of the high amount of enzymes enhance insecticide detoxification by effective sequestration. Over production of CarEs has been reported in peach aphids,<sup>23,36</sup> mosquitoes,<sup>24,37</sup> brown planthopper,<sup>30</sup> cotton aphid,<sup>31</sup> and the oriental fruit fly.<sup>38</sup> The other involves single point mutations in specific CarEs which causes a significant increase in insecticide detoxification by elevating hydrolytic activity. Point mutation has been well characterized in a few insect species, including *Lucilia cuprina* and *Musca domestica*.<sup>33–35</sup>

CarE-mediated insecticide resistance in cotton bollworm, has been widely reported since the 1990s in Australia,<sup>3</sup> India,<sup>4</sup> Africa,<sup>39</sup> and China.<sup>13</sup> However, few CarEs involved in this resistance have been fully elucidated at the molecular level to date. A recent study found that four CarE genes were constitutively overexpressed in a fenvalerate-resistant cotton bollworm strain.<sup>40</sup> So far, there have been no reports about point mutations in CarEs responsible for insecticide resistance in the cotton bollworm. In our previous work, Asp and Leu mutations occurring in natural populations of *Lucilia cuprina*<sup>33</sup> and *Musca domestica*,<sup>34</sup> were *in vitro* introduced into a few CarEs of *H. armigera*. The results showed that Asp substitutions in CarE001C increased OP hydrolyase activity about 14-fold, and Leu substitutions caused about

four- to six-fold increase in SP hydrolyase activity of three other CarEs, CarE001B, CarE001D and CarE001F.<sup>41</sup> However, these point mutations have never been found in either laboratory-selected strains or field populations of *H. armigera*.

There are limited reports on the metabolic activities of CarEs in the cotton bollworm. An earlier study suggested that the homogenate of the third instar larvae of *H. armigera* exhibited hydrolytic activity toward esfenvalerate.<sup>42</sup> Another study provided forceful evidence that the crude enzyme sample from larval midguts could metabolize *alpha*-cyano-3-phenoxybenzyl SPs into smaller metabolites.<sup>25</sup> Recently, another study reported that an enzyme preparation from the midgut samples of a laboratory-selected resistant *H. armigera* strain showed hydrolyase activity toward fenvalerate.<sup>40</sup> Although genomic analysis has shown that about 40 CarEs are found in the cotton bollworm,<sup>43</sup> current knowledge on the metabolic capacities of individual CarEs toward insecticides is rather limited. In 2013, a study which involved the heterologous expression of CarEs from *H. armigera* using a baculovirus system, firstly reported that individual CarEs had relatively low hydrolyase activity toward either OPs or SPs.<sup>44</sup> Similarly in our previous studies, heterologous expression in the *Escherichia coli* demonstrated that a few CarEs, including CarE001D, CarE001G, CarE001A and CarE001H from *H. armigera* strain (WH), had metabolic activities toward SPs.<sup>45–47</sup> Particularly CarE001G exhibited highly catalytic efficiency toward  $\alpha$ -naphthyl acetate ( $\alpha$ -NA), strong binding affinity to OPs, and relatively low hydrolyase activity toward three types of SPs. Furthermore, CarE001G has a longer open reading frame (ORF) encoding 747 amino acid (aa) residues than most of other CarEs (550–560 aa), and contains a glycine-rich region (GRR) (approximately 60% glycine residues) of 146 aa located adjacent to the C-terminal. We discovered that this GRR sequence consisted of both 12 copies of 8 aa motif (Glu-Gly-Asp-Gly-Gly-Asp-Gly-Gly) and ten copies of 5 aa motif (Glu-Gly-Asp-Gly-Gly).<sup>46</sup> To the best of our knowledge, no similar GRR sequences have yet been documented in either the cotton bollworm or other insect species to date, and the functions of this unique ‘glycine-rich region’ in insecticide metabolism and insecticide resistance are unknown.

In this study, we identified and obtained four variants of CarE001G with different deletions in the GRR, and heterologously expressed them in an *E. coli* system. We characterized the biochemical properties of the purified mutant enzymes (including kinetic parameters and insecticide inhibition properties), measured their hydrolyase activity toward  $\beta$ -cypermethrin, and performed molecular docking analysis and molecular dynamic (MD) simulations. We found that the mutant CarEs can significantly enhance the detoxification of  $\beta$ -cypermethrin. This was confirmed by a binding mode analysis. Our results indicate that deletion mutations in GRR may play a role in the development of SP resistance in *H. armigera* in the future.

## 2 MATERIALS AND METHODS

### 2.1 Chemicals

The chemicals, including  $\beta$ -cypermethrin, triphenyl phosphate (TPP),  $\alpha$ -NA,  $\alpha$ -naphthol, and Fast Blue RR salt, were of technical grade and were obtained from Aladdin Regent (Shanghai, China). Other chemicals and solvents used in this work were all reagent grade.

### 2.2 Construction and bacterial expression of mutant enzymes

The DNA of variants of CarE001G were amplified with gene-specific primers containing restriction endonuclease recognition

sites (Supporting Information, Table S1) by using the PrimeSTAR HS DNA Polymerase (TaKaRa, Shiga, Japan) following the method of Li *et al.*<sup>47</sup> with slight modification. Briefly, the amplification program was 98 °C for 30 s, followed by 30 cycles of 98 °C for 12 s, 68 °C for 25 s, 72 °C for 2.5 min, and with a final extension of 72 °C for 5 min. The *CarE001G* gene has been cloned and sequenced in our previous work.<sup>46</sup> The plasmid having the ORF of *CarE001G* was used as template in the amplification. The polymerase chain reaction (PCR) products were applied on 1% agarose gel. The two PCR product bands (approximately 2.0 kb and 2.3 kb) were separately excised from the 1% agarose gel (Fig. 1 (A)). After purification and digestion, the two products were inserted into pET30a vector and transformed into *E. coli* BL21 (DE3) competent cells, respectively. Then the recombinant plasmids were sequenced by Invitrogen Company (Beijing, China). To further verify these four mutant CarEs, they were sequenced again by another company (Sangon Company, Shanghai, China).

To characterize the function of these mutant CarEs, the sequencing-verified plasmids were used to yield mutant CarEs according to the protocol of our previous report.<sup>46,47</sup> The protein expression was induced by isopropylthio-galactoside (IPTG) at a final concentration of 0.2 mmol L<sup>-1</sup> at 18 °C for 48 h with shaking at 200 rpm.

### 2.3 Purification and verification of recombinant mutant enzymes

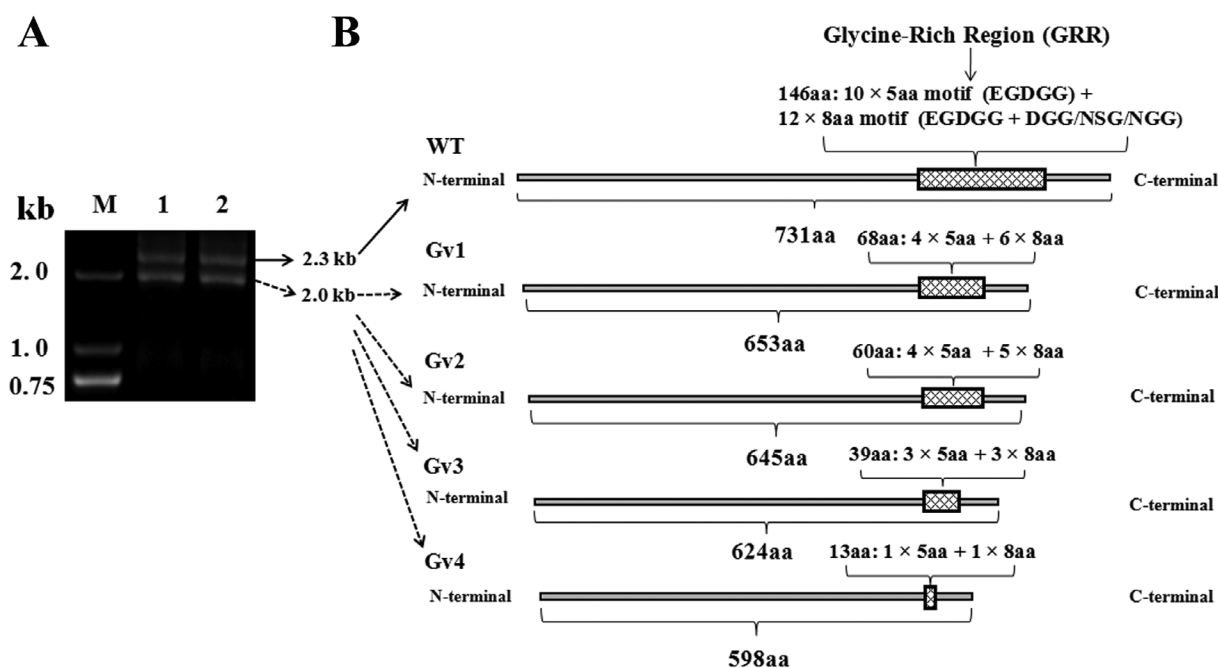
The recombinant proteins were purified following the protocol of Bai *et al.*<sup>46</sup> Briefly, cells were harvested by centrifugation after incubation for 48 h. The cell pellets were then re-suspended with 25 mmol L<sup>-1</sup> Tris-HCl (pH 8.0) and lysed by sonication. The soluble recombinant proteins in the supernatant were collected after centrifuging at 15 000 rpm at 4 °C for 0.5 h (CF16RX II, Hitachi Ltd, Tokyo, Japan), and filtered by passing through a sterile 0.22 µm filter. The recombinant proteins were

finally purified by a Ni<sup>2+</sup>-NTA resin (TransGen Biotech, Beijing, China) with a linear gradient of 50 to 200 mmol L<sup>-1</sup> imidazole solutions following the manufacturer's instructions. The purified mutant proteins were subsequently dialyzed with 25 mmol L<sup>-1</sup> Tris-HCl (pH 8.0) three times, followed by the verification with 12% sodium dodecyl sulfate polyacrylamide gel electrophoresis (SDS-PAGE) gel and Western blot analysis. The concentrations of the purified mutant proteins were measured using the Bradford method.<sup>48</sup>

To investigate the activity of the recombinant mutant CarEs, the purified proteins were submitted to a native polyacrylamide gel electrophoresis (native-PAGE) following the method of Teese *et al.*<sup>43</sup> with slight modification. Briefly, after running for about 60 min at 200 V in a 4 °C fridge, gels were incubated in freshly prepared 50 mL of staining solution (containing 1 mmol L<sup>-1</sup> α-NA and 3 mmol L<sup>-1</sup> Fast Blue RR salt) at 50 rpm for 5 to 10 min at room temperature. Next, 30 mL of 0.05% (v/v) glacial acetic acid was used to replace the staining solution and to stop the staining reaction. Finally, gels were washed by deionized water three times for about 15 min.

### 2.4 Mass spectrometry analysis of purified mutant enzymes

To further identify the purified recombinant proteins, a mass spectrometry (MS) assay was performed following the method of Bai *et al.*<sup>46</sup> Briefly, the excised fractions corresponding to the two protein bands on the SDS-PAGE gel were separately digested by trypsin (sequence grade, Promega, Madison, WI, USA) and incubated at 37 °C for 20 h. After the resultant peptides were separated and desalted, the peptide analysis was conducted by a 5800 MALDI-TOF/TOF (AB SCIEX) by Shanghai Applied Protein Technology Company, China. The amino acid residues of the protein bands identified by MS were aligned with the deduced sequences of each mutant CarEs.



**Figure 1.** Agarose gel analysis of PCR products and the frames of mutant CarEs and wild-type CarE001G (WT). (A) The results of PCR amplification of *CarE001G* by 1% agarose gel assay. (B) The length of the deduced amino acid sequences of each mutant CarEs and the organizations of short simple tandem repeats in each GRR sequences in their C-terminals. M indicates the DNA marker, lanes 1 and 2 are the PCR products of *CarE001G*.

## 2.5 Sequence analysis

The molecular weight of the mutant CarEs was predicted by the Compute pl/Mw tool of ExPASy web ([https://web.expasy.org/compute\\_pi/](https://web.expasy.org/compute_pi/)). Protein glycosylation sites were predicted with the NetNGlyc1.0 server (<http://www.cbs.dtu.dk/services/NetNGlyc/>). The glycosylphosphatidylinositol (GPI) anchor sites were predicted with GPI-anchor Predictor (<http://gpcr2.biocomp.unibo.it/gpipe/pred.htm>). Comparison of amino acid sequences was performed using the GENEDOC software (<https://genedoc.software.informer.com/2.7/>).

## 2.6 Kinetic assay of mutant CarEs with $\alpha$ -naphthyl acetate

The kinetics of purified mutant CarEs were determined using continuous assay on a microplate reader (M200 PRO, Tecan Trading AG, Männedorf, Switzerland) with 96-well microplate following the protocol of Teese *et al.*<sup>44</sup> and Li *et al.*<sup>47</sup> Briefly, a stock solution of  $\alpha$ -NA in absolute methanol was used to prepare a 400  $\mu\text{mol L}^{-1}$  working solution (containing 3  $\text{mmol L}^{-1}$  Fast Blue RR salt) with 0.1  $\text{mol L}^{-1}$  sodium phosphate buffer (PBS) pH 7.5 (0.1  $\text{mol L}^{-1}$   $\text{Na}_2\text{HPO}_4$  and 0.1  $\text{mol L}^{-1}$   $\text{NaH}_2\text{PO}_4$ ). Then, the first round assay was carried out to determine the suitable quantity of enzyme to be added to the reaction. This reaction system consisted of 100  $\mu\text{L}$  of 200  $\mu\text{mol L}^{-1}$   $\alpha$ -NA and equal volume of serial dilutions of the enzymes (two-fold, four-fold, eight-fold, 12-fold, 32-fold, 64-fold) in 0.1  $\text{mol L}^{-1}$  PBS buffer (pH 7.5). The second round assay was subsequently carried out using the optimal quantity of enzyme to measure the initial velocity of each reaction at serial concentrations of  $\alpha$ -NA (16–200  $\mu\text{mol L}^{-1}$ ). Finally, the initial rates of all the 12 reactions were used to accurately estimate the exact kinetic parameters [including maximum velocity ( $V_{\text{max}}$ ) and Michaelis–Menten constant ( $K_m$ )] by fitting the data to the Michaelis–Menten equation, and the intrinsic clearance values ( $\text{CL}_{\text{int}} = V_{\text{max}}/K_m$ ) were calculated following the method of Han *et al.*<sup>49</sup>

## 2.7 Inhibition assay of mutant CarEs activity by $\beta$ -cypermethrin

The inhibition of mutant CarEs by insecticides were examined following the method of Gunning *et al.*<sup>50</sup> with modification. Briefly, serial stock solutions of  $\beta$ -cypermethrin and TPP were dissolved in acetone, and 1.0–2.0  $\mu\text{L}$  of stock solution was added to the enzyme solution (1  $\mu\text{g}$  in 100  $\mu\text{L}$ ) to give final concentrations ranging from 1 to 1000  $\mu\text{mol L}^{-1}$ . All treatments were incubated at 30 °C for 10 min, then the residual CarE activity was screened at 450 nm by adding 100  $\mu\text{L}$  of 400  $\mu\text{mol L}^{-1}$   $\alpha$ -NA containing 3  $\text{mmol L}^{-1}$  Fast Blue RR salt with a microplate reader (M200 PRO). TPP was used as a positive control inhibitor. The half inhibitory concentration ( $\text{IC}_{50}$ ) values were estimated by the Trimmed Spearman–Karber method.<sup>51</sup>

## 2.8 HPLC assay of CarEs hydrolase activity toward $\beta$ -cypermethrin

The hydrolase activity of mutant CarEs was determined by measuring the difference between the residual concentration and the initial concentration of  $\beta$ -cypermethrin using high-performance liquid chromatography (HPLC) assay following the protocol of Li *et al.*<sup>47</sup> First, a stock solution of 20  $\text{mmol L}^{-1}$   $\beta$ -cypermethrin in absolute acetone was diluted to 200  $\mu\text{mol L}^{-1}$  in 0.1  $\text{mol L}^{-1}$  sodium phosphate buffer (PBS) pH 7.5 (0.1  $\text{mol L}^{-1}$   $\text{Na}_2\text{HPO}_4$  and 0.1  $\text{mol L}^{-1}$   $\text{NaH}_2\text{PO}_4$ ). Next, 100  $\mu\text{L}$  substrate solution was added into 100  $\mu\text{L}$  of appropriately diluted enzyme sample, and reactions were

incubated at 30 °C for 2 h and then stopped by adding 200  $\mu\text{L}$  acetonitrile. The residual  $\beta$ -cypermethrin was measured using an Agilent HPLC 1260 Series equipped with a Symmetry C18 column (250 mm  $\times$  4.6 mm, 5  $\mu\text{m}$ , Waters, Milford, MA, USA) by using acetonitrile and water as the solvents (80:20, v/v) with an 1.0  $\text{mL min}^{-1}$  flow rate. The column temperature was maintained at 30 °C, and the detection wavelength was 210 nm. Heat-inactivated mutant CarE was used as a negative control. All reactions were tested in three replicates. The hydrolase activity toward  $\beta$ -cypermethrin was presented as the specific activity, which is defined as the nanomoles of substrate loss per minute per milligram enzyme.

## 2.9 Homology modeling of mutant and wild-type CarEs

After querying the Protein Data Bank (PDB database) with the amino acid sequence of CarE001G, the crystal structure of *Lucilia cuprina alpha-esterase-7* (LcaE7) (PDB ID: 5TYM, resolution = 1.84 Å) was selected as the suitable template for homology modeling (about 31% homology).<sup>52</sup> The three-dimensional (3D) structures of wild-type CarE (CarE-wt) and mutant CarE (CarE-mut) (Gv4) were built by modeller 9.21 package (<https://salilab.org/modeller/>). Subsequently, the Amber 14 and AmberTools15 programs were used for MD simulations of the selected pose.<sup>53</sup> The 3D structure of the CarE was placed in a rectangular box with explicit TIP3P water extending at least 10 Å in each direction using the ‘SolvateOct’ command with the minimum distance between any solute atoms. Equilibration of the solvated complex was performed by carrying out a short minimization (the steepest descent for the first 2000 steps and the conjugate gradient method for the subsequent 2000 steps), 1000 ps of heating, and 500 ps of density equilibration with weak restraints using the GPU accelerated PMEMD (Particle Mesh Ewald Molecular Dynamics) module. Finally, 200 ns of MD simulations were carried out. All the MDs were performed on a Dell Precision T5500 workstation.

## 2.10 Binding mode analysis of CarE- $\beta$ -cypermethrin complexes

Molecular docking analyses were performed to investigate the binding mode between the wild type and the mutant with the  $\beta$ -cypermethrin compound using Autodockvina 1.1.2.<sup>54</sup> The AutoDock Tools 1.5.6 package was employed to generate the docking input files.<sup>55</sup> The search grid of the CarE-wt and CarE-mut were both identified as center\_x: 31.792, center\_y: 49.617, and center\_z: 71.429 with dimensions size\_x: 15, size\_y: 15, and size\_z: 15. In order to increase the docking accuracy, the value of exhaustiveness was set to 20. For Vina docking, the default parameters were used if it was not mentioned. MD simulations of both target complexes were carried out following the earlier mentioned method. At least, 100 ns of MD simulations were carried out.

The binding free energies ( $\Delta G_{\text{bind-kal}}$ ) of the generated CarE- $\beta$ -cypermethrin complexes were calculated using the Molecular Mechanics/Generalized Born Surface Area (MM/GBSA) method in AmberTools 15. For each complex, the binding free energy of MM/GBSA was estimated as follows:

$$\Delta G_{\text{bind}} = G_{\text{complex}} - G_{\text{protein}} - G_{\text{ligand}}$$

where  $\Delta G_{\text{bind}}$  is the binding free energy,  $G_{\text{complex}}$ ,  $G_{\text{protein}}$  and  $G_{\text{ligand}}$  are the free energies of complex, protein, and ligand, respectively.

### 3 RESULTS

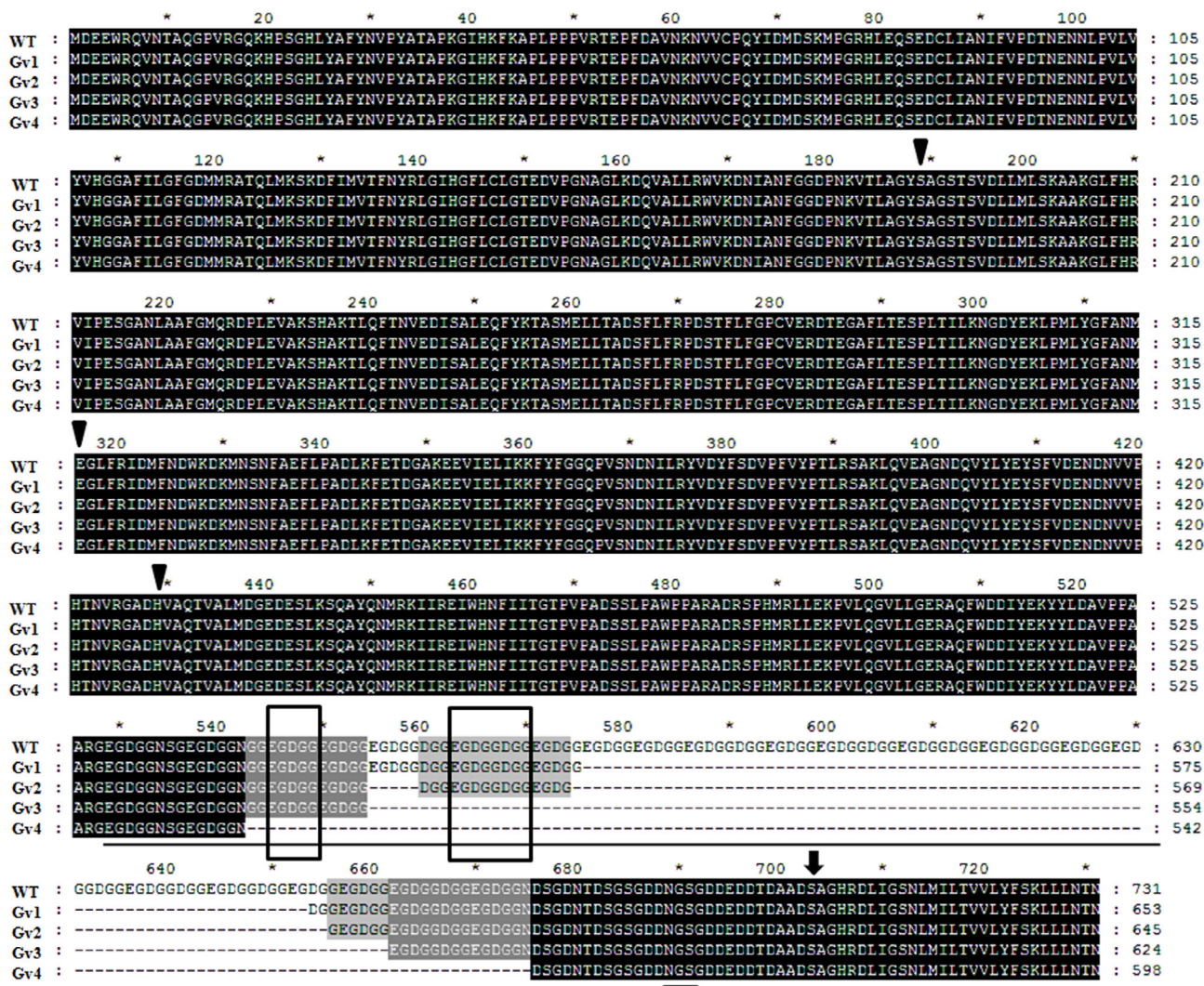
#### 3.1 Identification and sequence analysis of mutant CarEs

With the exception of the 2.3 kb PCR product band, a smaller PCR product band (about 2.0 kb) was amplified in PCR amplification (Fig. 1(A)). The two product bands were then purified and ligated into the pET30a vector. Four mutant CarEs, Gv1, Gv2, Gv3 and Gv4, were identified from the cloning of smaller PCR products by sequencing analysis using Vector NT II software (Figs 1 and 2). The results showed that Gv1, Gv2, Gv3, and Gv4 had complete ORFs encoding 653, 645, 624 and 598 aa residues, respectively. Comparison of the deduced amino acid sequences of these mutant CarEs with the wild-type CarE001G, revealed that their protein sequences were well matched with the wild type except the variation of the length of the GRR sequences (Fig. 2). Specifically all mutant CarEs had a shorter GRR sequences in their ORFs compared with the 146 aa of GRR in the ORF of wild type. As shown in Fig. 1 and Table 1, Gv1 had 653 aa residues in length with a short GRR sequence (68 aa), which comprised six copies of an 8 aa motif and four copies of a 5 aa motif. The ORF of Gv2 encoded 645 aa residues containing a 60 aa of GRR sequence,

which consisted of five copies of an 8 aa motif and four copies of a 5aa motif. By contrast, the ORF of Gv3 encodes a shorter protein of 624 aa residues in length with a GRR sequence of 39 aa, which has both three copies of an 8 aa motif and a 5 aa motif. Particularly, the fourth mutant CarE (Gv4) had the shortest ORF encoding 598 aa residues with the shortest GRR sequence of 13 aa residues in length and consisted of each copy of an 8 aa motif and a 5aa motif (Fig. 1, Table 1).

#### 3.2 Purification and verification of recombinant mutant CarEs

To characterize the function of the mutant CarEs, the recombinant plasmids harboring mutant CarE genes were transferred into competent *E. coli* BL21 (DE3) cells to yield recombinant proteins. The purified mutant proteins were identified by SDS-PAGE analysis (Fig. 3(A)). As shown, a relatively heavy staining band termed Band 1 (70–75 kDa) for each purified mutant proteins was clearly detected on the gel. These bands (Band 1) were the target proteins as they are agreed with the predicted molecular weight of Gv1, Gv2, Gv3 and Gv4, which are approximately 76.0 kDa, 75.4

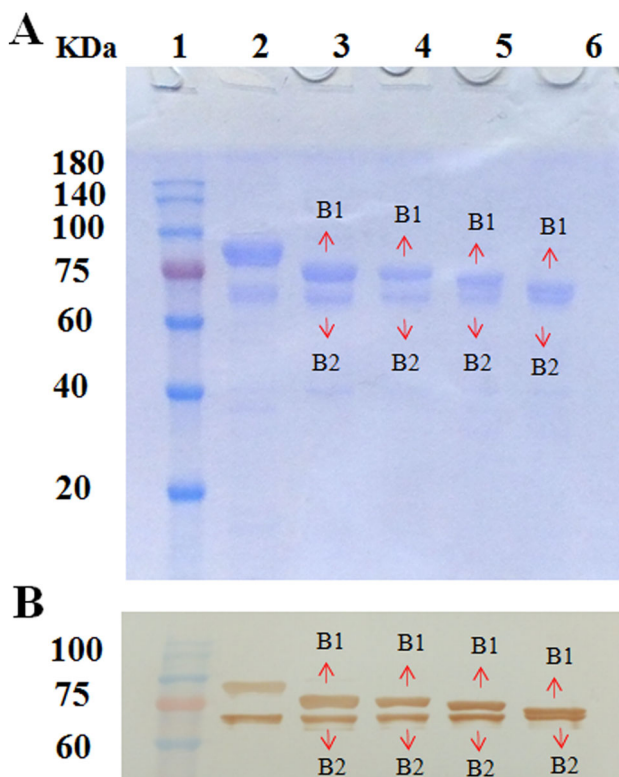


**Figure 2.** Comparison of the deduced amino acid sequence of four mutant CarEs with the wild-type CarE001G (WT). The glycine-rich region (GRR) sequences are underlined, and each 5 aa and 8 aa simple tandem repeats are boxed, respectively. The conserved catalytic triad residues are labeled with triangles, the potential N-glycosylation site is marked with double underlines, and the predicted GPI-anchor site is shown with a dark arrow.

**Table 1.** Organization of glycine-rich region (GRR) sequences in wild-type and mutant carboxylesterases (CarEs)

Enzymes	ORF (aa)	Sequences and copies of 5 aa/8 aa motifs in GRR (5'–3')											GRR (aa)	
WT <sup>a</sup>	731	2 × 8aa	2 × 5aa	2 × 8aa	3 × 5aa	1 × 8aa	1 × 5aa	3 × 8aa	1 × 5aa	3 × 8aa	2 × 5aa	1 × 8aa	1 × 5aa	146
Gv1	653	2 × 8aa	2 × 5aa	3 × 8aa	1 × 5aa	1 × 8aa	1 × 5aa							68
Gv2	645	2 × 8aa	1 × 5aa	2 × 8aa	2 × 5aa	1 × 8aa	1 × 5aa							60
Gv3	624	2 × 8aa	2 × 5aa	1 × 8aa	1 × 5aa									39
Gv4	598	1 × 8aa	1 × 5aa											13

<sup>a</sup> WT indicates wild-type CarE001G.

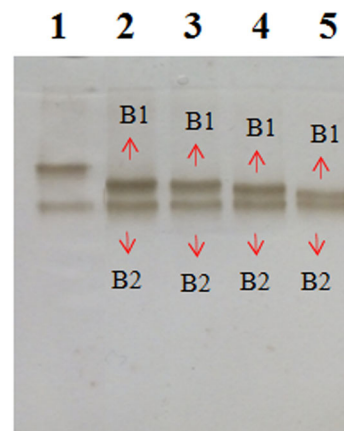


**Figure 3.** SDS-PAGE and Western blot analysis of purified recombinant proteins of mutant CarEs. (A) The samples were separated on 12% gels. (B) Western blot analysis of purified protein using anti-His tag antibody. Lane 1: pre-stained protein ladder; Lane 2: purified protein of wild-type CarE001G (WT) as a control; Lanes 3–6: purified recombinant proteins of Gv1, Gv2, Gv3, and Gv4, respectively. B1 and B2 are abbreviations for Band 1 and Band 2, respectively.

kDa, 73.7 kDa, and 71.6 kDa, respectively (Fig. 3(A)). Besides Band 1 on the SDS-PAGE gel, another smaller protein band termed Band 2 with relatively light staining intensities was also observed from each purified mutant CarEs (Fig. 3(A)). The Western blot analysis confirmed that these two protein bands were in accordance with the SDS-PAGE analysis (Fig. 3(B)). Furthermore, both heavy staining bands were detected by native-PAGE analysis (Fig. 4), thus indicating these two protein bands are active and could hydrolyze the substrate ( $\alpha$ -NA) into metabolites.

### 3.3 Mass spectrometry analysis of the expressed mutant CarEs

The MS assays were performed to further investigate the relations between the two bands of each mutant CarEs. About 22–25

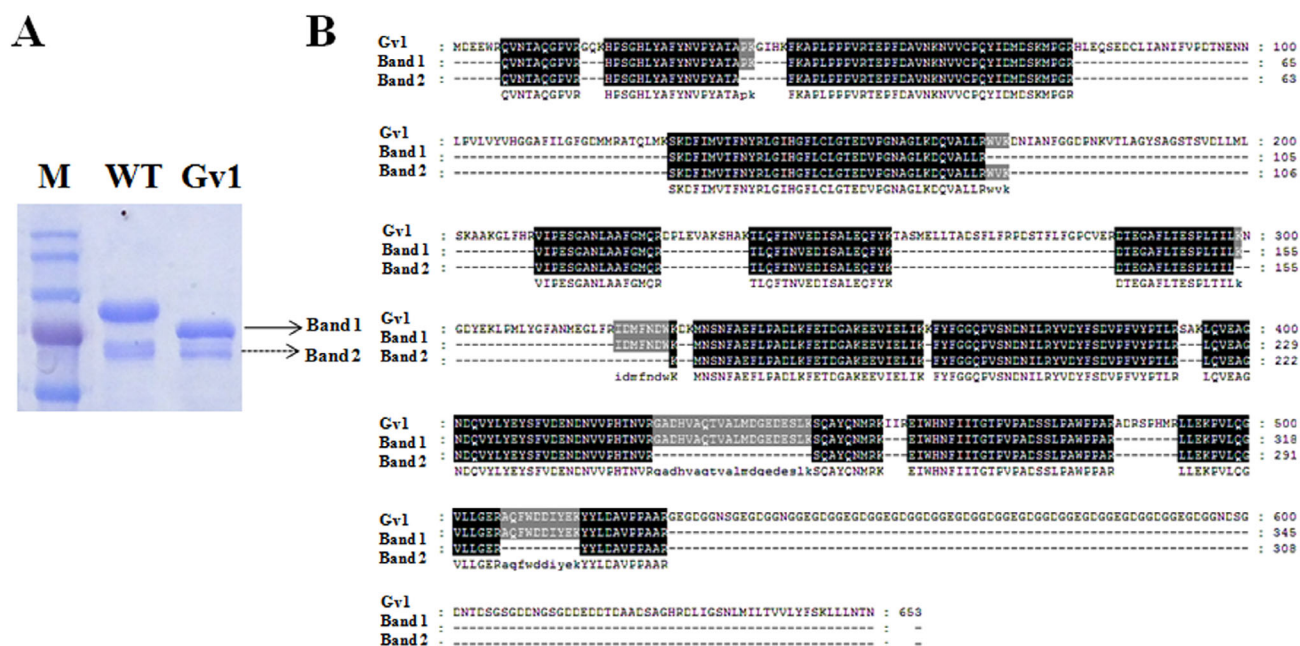


**Figure 4.** Verification of the activity of purified mutant enzymes by native-PAGE analysis. The gel was stained with the freshly prepared solution contained 1 mmol L<sup>-1</sup>  $\alpha$ -NA and Fast Blue RR salt. Lane 1: purified recombinant proteins of wild-type CarE001G (WT) as a control; Lanes 2–4: purified recombinant proteins of Gv1, Gv2, Gv3, and Gv4, respectively.

protein-peptides were identified from the target band and the smaller band (Band 1 and Band 2) (Tables S2 and S3). Comparison of these protein-peptide data with the deduced amino acid sequences of each mutant CarEs, indicated that the peptide sequences of Band 1 and Band 2 well matched with the amino acid sequence of the individual corresponding mutant CarEs (Fig. 5 and Supporting Information, Fig. S2). Thus these results in combination with the results from the Western blot and native PAGE analyses, suggested that Band 1 was the target protein of each mutant CarEs, but Band 2 may be the fragment from the intact protein of each mutant CarEs.

### 3.4 CarE activity of mutant CarEs toward $\alpha$ -NA

The enzymatic activity toward the model substrate  $\alpha$ -NA was determined by spectrophotometric microplate assays (Figs S3 and S4). As seen in Table 2, although the bind affinities of  $\alpha$ -NA to both the mutant esterases and wild type were very similar and no significant differences were observed between their  $K_m$  values,  $V_{max}$  for the four mutant esterases appeared to be significantly higher than that of wild type, thus showing very high rate constant ( $k_{cat}$ ) values (12.3–20.3 s<sup>-1</sup>) toward  $\alpha$ -NA. Particularly, three of four mutant CarEs (Gv1, Gv2 and Gv3) exhibited an maximal specific activity of approximately 16 nmol s<sup>-1</sup> mg<sup>-1</sup> protein, which was about two-fold higher than that of the wild type. For only two of the mutant CarEs, Gv1 and Gv2, the  $CL_{int}$  value and specificity constant ( $k_{cat}/K_m$ ) values were significantly larger than those of wild type. These results indicated that deletion mutations



**Figure 5.** Comparison of the amino acid residues of the two protein bands isolated from the SDS-PAGE and identified by MS with the deduced sequences of Gv1. (A) SDS-PAGE analysis of the purified proteins of wild-type CarE001G (WT) and mutant Gv1. (B) Comparison of the amino acid sequences of protein Band 1 and Band 2 with the known amino acid residues of Gv1 by GENDOC software. The amino acid residues shaded in black indicate the identical peptide residues between the two protein bands (Bands 1 and 2) and the deduced amino acid sequence of Gv1. Wild-type CarE001G (WT) was used as a control in the SDS-PAGE analysis.

**Table 2.** Kinetic parameters of purified mutant carboxylesterases (CarEs) expressed in *Escherichia coli* toward  $\alpha$ -naphthyl acetate ( $\alpha$ -NA)

Enzymes	$V_{max}$ (nmol s <sup>-1</sup> mg <sup>-1</sup> )	$K_m$ ( $\mu$ mol L <sup>-1</sup> )	$CL_{int}$ (mL s <sup>-1</sup> mg <sup>-1</sup> ) <sup>†</sup>	$k_{cat}$ (s <sup>-1</sup> )	$k_{cat}/K_m$ (L s <sup>-1</sup> $\mu$ mol <sup>-1</sup> )	Specific activity ( $\mu$ mol s <sup>-1</sup> $\mu$ g <sup>-1</sup> ) <sup>‡</sup>
WT <sup>§</sup>	106.4 ± 8.9 c	9.6 ± 2.8	11.7 ± 3.4 b	8.8 ± 0.8 c	1.0 ± 0.3 b	8.4 ± 0.9 b
Gv1	240.1 ± 4.1 a	9.1 ± 0.9	26.7 ± 2.1 a	18.2 ± 0.3 a	2.0 ± 0.2 a	16.6 ± 0.6 a
Gv2	268.8 ± 12.4 a	8.3 ± 2.6	34.9 ± 11.0 a	20.3 ± 0.9 a	2.6 ± 0.8 a	16.2 ± 1.3 a
Gv3	253.7 ± 20.8 a	12.7 ± 1.4	20.2 ± 4.1ab	18.7 ± 1.5 a	1.5 ± 0.3 ab	16.3 ± 3.1 a
Gv4	171.1 ± 20.7 b	9.5 ± 0.1	18.0 ± 1.9 b	12.3 ± 1.5 b	1.3 ± 0.1 b	10.1 ± 0.6 b

Different letters following the means for each enzyme indicate the significant differences (Tukey's test,  $P \leq 0.05$ ).  
 $V_{max}$ , maximum velocity;  $K_m$ , Michaelis–Menten constant;  $k_{cat}$ , rate constant;  $k_{cat}/K_m$ , specificity constant.  
<sup>†</sup>  $CL_{int}$  indicates intrinsic clearance, values were calculated based on  $V_{max}/K_m$ .  
<sup>‡</sup> Specific activity was determined at 200  $\mu$ mol L<sup>-1</sup> substrate.  
<sup>§</sup> WT indicates wild-type CarE001G, data is cited from Bai *et al.*<sup>46</sup>

in the GRR sequence could alter the enzymatic activity of CarE001G.

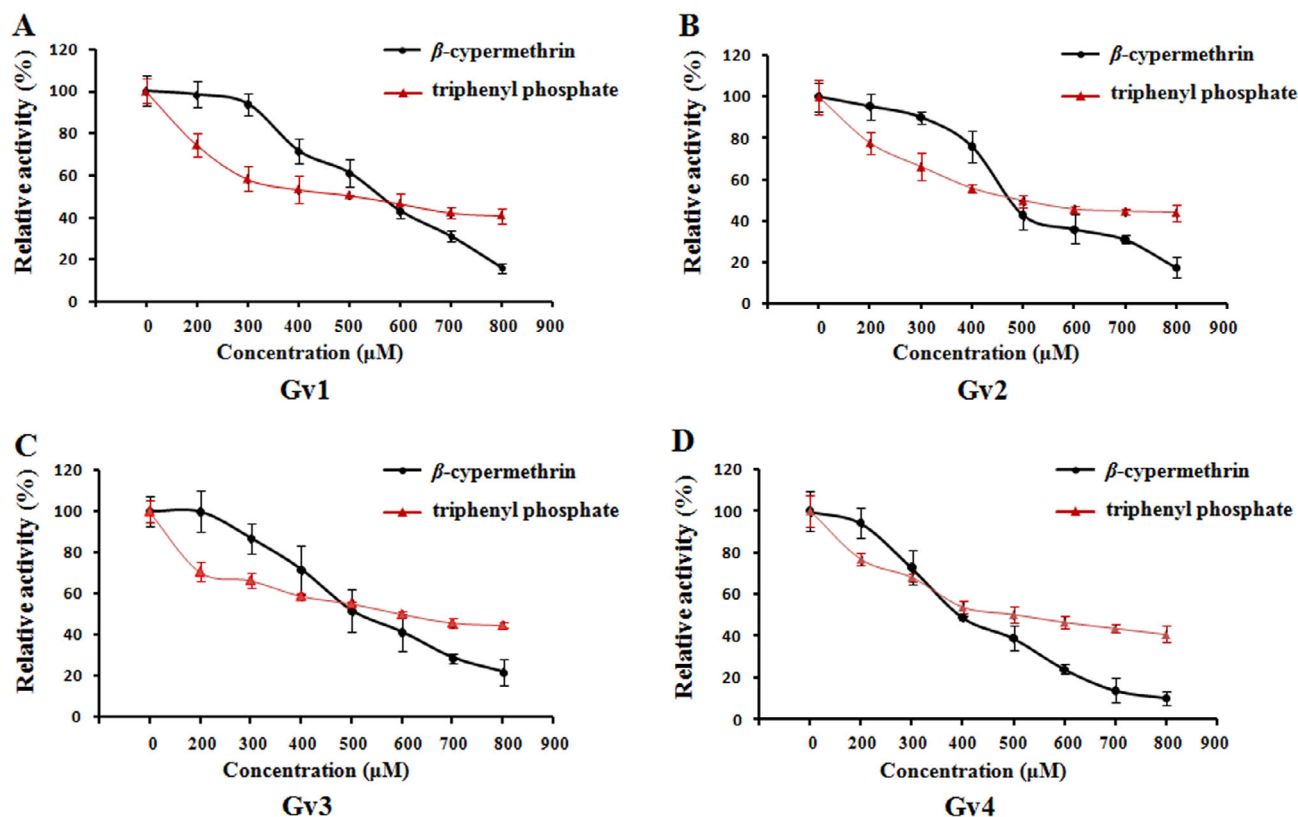
### 3.5 Inhibition properties of $\beta$ -cypermethrin against mutant esterase activity

Inhibition ( $IC_{50}$  values) of the four mutant CarEs was determined (Fig. 6). The results showed that  $\beta$ -cypermethrin exhibited relatively strong inhibition on the mutant CarEs with  $IC_{50}$  values varying from 400 to 550  $\mu$ mol L<sup>-1</sup> compared to wild type (Table 3). The inhibition of  $\beta$ -cypermethrin on Gv1, Gv2 and Gv3 was as strong as their positive inhibitors. Particularly the inhibition of  $\beta$ -cypermethrin was significantly stronger for Gv4 ( $IC_{50} = 408.5 \pm 21.5 \mu$ mol L<sup>-1</sup>) than the positive inhibitor ( $IC_{50} = 539.5 \pm 82.3 \mu$ mol L<sup>-1</sup>). However, in contrast to these mutants,  $\beta$ -cypermethrin exhibited significantly weaker inhibition ( $IC_{50} = 501.2 \pm 22.7 \mu$ mol L<sup>-1</sup>) on the wild-type CarE001G than its positive inhibitor control ( $IC_{50} = 360.3 \pm 15.9 \mu$ mol L<sup>-1</sup>). The

results earlier suggested these mutants had stronger binding affinity with  $\beta$ -cypermethrin than the wild type.

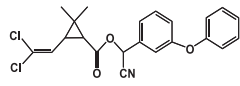
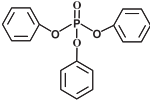
### 3.6 Metabolic activity of mutant CarEs toward $\beta$ -cypermethrin

The measurement of the metabolic activity of mutant CarEs was performed by HPLC assay (Fig. S5). The *in vitro* metabolism assay showed that four mutant CarEs had significantly higher metabolic activity toward  $\beta$ -cypermethrin than wild type (Fig. 7). Specifically, Gv4 had the highest metabolic activity with a specific activity of  $5.6 \pm 1.1$  nmol min<sup>-1</sup> mg<sup>-1</sup> protein, followed by Gv2 and Gv3 with specific activities of  $5.3 \pm 0.5$  and  $4.6 \pm 0.9$  nmol min<sup>-1</sup> mg<sup>-1</sup> protein, respectively. Comparatively, Gv1 had the lowest metabolic activity among these mutants. These results indicated that the deletion mutation in the GRR of CarE001G significantly enhanced



**Figure 6.** Inhibition tests of  $\beta$ -cypermethrin and triphenyl phosphate (TPP) on activities of mutant CarEs. (A) Gv1. (B) Gv2. (C) Gv3. (D) Gv4. The  $\alpha$ -NA was used as a substrate, with TPP as a positive control inhibitor. Residual activity was expressed as the percentage of the initial activity without insecticide. Data are presented as mean ( $\pm$  standard error) for three replicates.

**Table 3.** Half inhibitory concentration ( $IC_{50}$ ) values for insecticide and inhibitor

Insecticide	Structure	$IC_{50}$ ( $\mu\text{mol L}^{-1}$ ) <sup>a</sup>				
		Gv1	Gv2	Gv3	Gv4	WT <sup>b</sup>
$\beta$ -Cypermethrin		547.1 $\pm$ 26.6	514.9 $\pm$ 49.8	533.4 $\pm$ 27.6	408.5 $\pm$ 21.5 *	501.2 $\pm$ 22.7 **
Triphenyl phosphate <sup>c</sup>		504.2 $\pm$ 88.5	548.7 $\pm$ 88.9	611.5 $\pm$ 104.0	539.9 $\pm$ 82.3	360.3 $\pm$ 15.9

The statistic *t*-test was done for the pair comparison between  $\beta$ -cypermethrin and triphenyl phosphate. Means for each inhibitor followed by the asterisks indicate the significant differences by student's *t*-test ( $P \leq 0.05$ ;  $**P \leq 0.01$ ).  
<sup>a</sup> Concentration of inhibitor that inhibits 50% of enzyme activity, values are means ( $\pm$  standard error) based on an average of three replicates.  
<sup>b</sup> WT indicates wild-type CarE001G, data is cited from Bai et al.<sup>46</sup>  
<sup>c</sup> Triphenyl phosphate was used as a positive inhibitor control in this work.

the metabolic ability toward  $\beta$ -cypermethrin, which was at least two-fold higher than that of the wild type.

### 3.7 Binding mode analyses of CarE-mut- $\beta$ -cypermethrin and CarE-wt- $\beta$ -cypermethrin complexes

To better explore the different underlying mechanism of CarE-wt and CarE-mut in metabolizing insecticide, the 3D structures of the CarE-wt and the CarE-mut were constructed using Modeler 9.21 package and optimized by MD simulations (Figs S6 and S7). After

a 200 ns MD simulation, the 3D structures of CarE-wt and CarE-mut achieved equilibrium, suggesting stability of the protein structures. The CarE-mut- $\beta$ -cypermethrin and CarE-wt- $\beta$ -cypermethrin complexes were also obtained by molecular docking and analyzed by MD simulations (Figs 8 and S7). As shown by the 100 ns of MD simulations (Fig. S8(A)), the CarE-mut- $\beta$ -cypermethrin complex reached equilibrium much earlier at the 15 ns point with an averaged root-mean-square deviation (RMSD) of  $2.0 \pm 0.2$  Å, compared to the CarE-wt- $\beta$ -cypermethrin



complex. The root-mean-square fluctuation (RMSF) analysis showed that the residues of CarE-mut appeared to be more rigid as a result of binding to  $\beta$ -cypermethrin. Its RMSF value was less than 2 Å. Comparatively, the CarE-wt exhibited a big degree of flexibility with an RMSF value almost 6 Å (Fig. S8(B)). These results probably imply that differences exist in the binding mode between the CarE-mut- $\beta$ -cypermethrin complex and the CarE-wt- $\beta$ -cypermethrin complex.

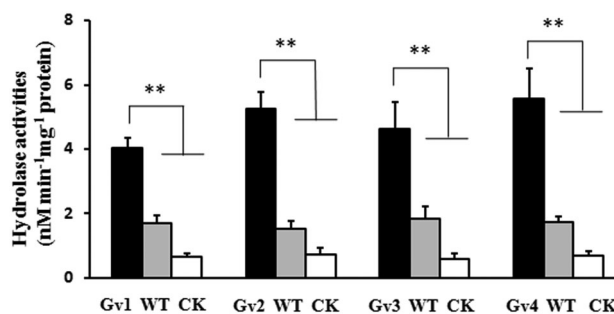
The theoretical binding modes of CarE\_wt and CarE-mut with  $\beta$ -cypermethrin are illustrated in Fig. 8. The  $\beta$ -cypermethrin adopted a compact conformation to bind in the binding pocket of the CarE-wt and CarE-mut, respectively (Fig. 8(A, B)). The canonical catalytic triad of the  $\alpha/\beta$ -hydrolase fold is conserved in both CarE-wt and CarE-mut with Ser205-His445-Glu332. The structure also reveals the presence of an aryl binding pocket consisting of Phe238, Phe291, and Phe401. Thus, the catalytic groups in both CarE-wt and CarE-mut are essentially identical to those groups in most other structural homolog, such as LcaE7 from *Lucilia cuprina*.<sup>56</sup> Detailed analysis further showed the cyhalothric acid part of the  $\beta$ -cypermethrin molecule was well located in the hydrophobic pocket which consisted of the amino acid residue Phe238, Phe291, Phe401, Leu292, and Val446, forming a strong binding, while the oxydibenzene scaffold of  $\beta$ -cypermethrin molecule located in another hydrophobic pocket consisted of Ala127, Phe335 and Thr449 (Fig. 8(A, B)). The interactions between the two binding complexes appeared to be nearly similar. However, comparison of their binding modes (Fig. 8(C)), indicated that a primary difference existed between the two complexes, which was that insecticide compound in the CarE\_mut- $\beta$ -cypermethrin complex was well oriented to the residues Ser205, Glu332 and His445 of the catalytic triad (Ser-Glu-His) of CarE-mut, which probably facilitated the compound to bind at the active sites of the enzyme. Particularly, the cation- $\pi$  interaction was observed between residue His445 and  $\beta$ -cypermethrin in the CarE-mut- $\beta$ -cypermethrin complex. Furthermore, the distance between the carbonyl carbon atom of the ester group of cypermethrin and the key residue Ser205 of CarE-mut was 3.3 Å, which was slightly shorter than that of Ser205 of CarE-wt (3.4 Å) (Fig. 8(C)). These points may make the CarE-mut more active in insecticide detoxification. Additionally, the calculations of the total binding free energy showed that the estimated  $\Delta G_{\text{bind}}$  for CarE-mut- $\beta$ -cypermethrin ( $-47.8 \pm 2.6 \text{ kcal mol}^{-1}$ ), was slightly higher than that for CarE-wt- $\beta$ -cypermethrin ( $-44.8 \pm 2.5 \text{ kcal mol}^{-1}$ ). This was consistent with the order of  $\beta$ -cypermethrin hydrolase activity determined *in vitro*. All together, these results provide rational explanation of the different interactions between CarE-mut and CarE-wt with  $\beta$ -cypermethrin molecules.

## 4 DISCUSSION

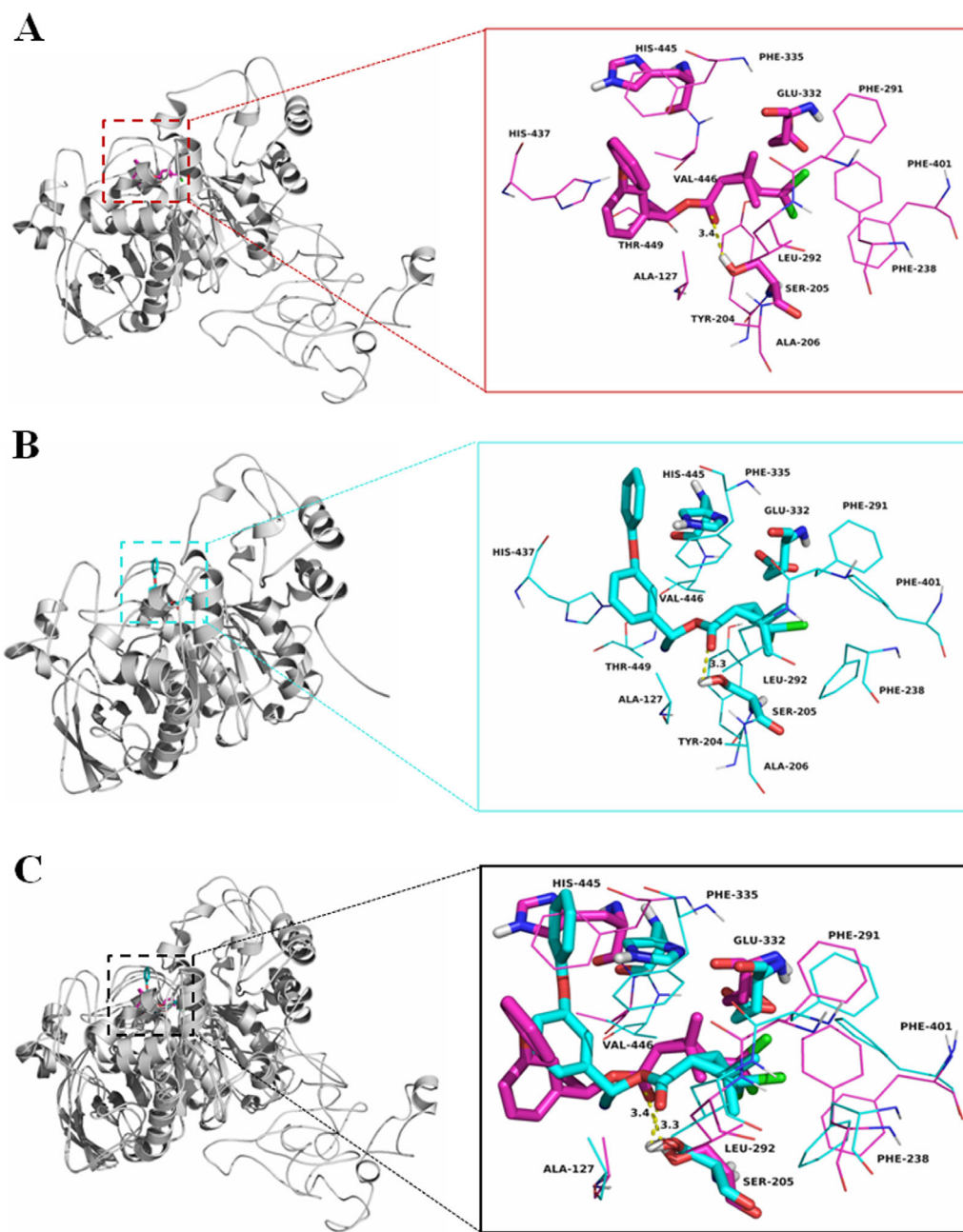
Insects have multiple CarE genes playing important roles in the detoxification of xenobiotic chemicals because they can metabolize or sequester various agrochemicals containing ester bond like SPs.<sup>20</sup> Approximately 40 CarE genes are located in the genome of *H. armigera*,<sup>43</sup> but few CarEs have been characterized in their biochemistry and functions. In a previous work, we found that CarE001G had 747 aa in length with 146 aa of unique GRR sequence, which was located at its C-terminal (the residues 545–690) and consisted of 12 copies of an 8 aa motif and ten copies of a 5 aa motif.<sup>46</sup> Furthermore, comparison of the protein sequence of CarE001G with other insect CarEs, indicated that the ORF of *CarE001G* had the longest amino acid sequence.<sup>46</sup>

These may suggest that *CarE001G* is an unique gene in the CarE super-family of *H. armigera* with unknown biological role. Interestingly, in this work we observed a smaller PCR product band (about 2.0 kb) in the 1% agarose gel analysis (Fig. 1(A)). Unexpectedly, we discovered four mutant CarEs from this smaller PCR band through cloning sequencing analysis (Figs 1 and 2). We further confirmed that no alternative mRNA splicing of *CarE001G* was observed in the *H. armigera* WH strain, and the 438 bp of GRR cDNA sequence (encoding the 146 aa residues) was located in the third exon of the genomic DNA of *CarE001G* (Fig. S1(A)). We suggest the following points as possible reasons for our observation. First, the addition of a higher amount of templates into the reaction mix led to a non-specific amplification. We confirmed that the smaller PCR product band was generally amplified by the addition of templates more than 10 ng to a 25  $\mu\text{L}$  of reaction mix (Fig. S1(B)). Second, a lower annealing temperature for PCR probably also resulted in a non-specific amplification. We clearly observed the smaller PCR product band when the annealing temperature was below 68 °C (Fig. S1(C)), suggesting that the specific amplification of *CarE001G* required a higher annealing temperature as the guanylic acids enriched in its DNA sequence. Lastly, the short simple tandem repeat in the GRR sequence may be prone to PCR errors during the amplification. These assumptions required further inquires.

Comparison of the deduced amino acid sequence of these four mutant CarEs with the wild-type CarE001G, indicated that they have shorter but complete ORFs (Figs 1 and 2). To characterize these four mutant CarEs, they were successfully expressed in an *E. coli* system and purified by  $\text{Ni}^{2+}$ -NTA resin. The target protein band (Band 1) for each sample was identified by SDS-PAGE and Western blot analysis (Fig. 3), while another smaller protein band (Band 2) was also observed. The native-PAGE assay indicated that the protein Band 2 was also active toward  $\alpha$ -NA (Fig. 4), and the MS assays further provided evidence that this band was the products of mutant CarE (Figs 5 and S2). Therefore, Band 1 and Band 2 were all active products from the bacterial expression of each mutant CarEs. The source of protein Band 2 is not clear. We hypothesized that it might be similar to the heterologous expression of wild-type CarE001G as we observed in our previous study.<sup>46</sup> We reasoned firstly that the mutant CarEs



**Figure 7.** Hydrolase activities of purified recombinant mutant CarEs toward  $\beta$ -cypermethrin. The values are shown as nanomoles of substrate hydrolyzed by per milligram of protein per minute ( $\text{nmol min}^{-1} \text{mg}^{-1}$  protein). They are means with standard errors based on an average of three replicates. Heat-inactivated CarE was used as a negative control (CK). WT is the abbreviation for wild-type CarE001G. The statistic *t*-test was done for the pair comparisons between each mutant CarEs and the wild-type CarE001G or the negative control, respectively. Asterisks above the error bars indicated the significant differences analyzed by Student's *t*-test (\*\* $P \leq 0.01$ ; \* $P \leq 0.05$ ).



**Figure 8.** Binding modes of CarE-wt- $\beta$ -cypermethrin complex and CarE-mut- $\beta$ -cypermethrin complex were shown by illustration in the surface mode and cartoon mode. (A) CarE-wt- $\beta$ -cypermethrin complex. (B) CarE-mut- $\beta$ -cypermethrin complex. (C) Comparison between binding modes of CarE-wt- $\beta$ -cypermethrin complex and CarE-mut- $\beta$ -cypermethrin complex. Enzyme is represented with cartoon mode and the involved residues are shown in lines. The catalytic triad Ser205-His445-Glu332 and the  $\beta$ -cypermethrin molecule are presented with stick mode.

are predicted to be GPI anchored proteins (Fig. 2), which usually attaches to the cell plasma membrane after translation and thus results in an incomplete release of protein from the cell. Secondly, that the protein may have been broken into a short form during lyses by sonication or purification by Ni-NTA<sup>2+</sup>. Lastly, that this may be due to irregular/error gene transcription or protein translation due to its special sequence characteristic (having part of GRR sequence located adjacent to their C-terminals). Further work is still needed to explore the accurate molecular mechanism for the protein Band 2.

Two types of generic artificial substrates,  $\alpha$ -NA and  $\beta$ -NA, are generally used as model substrates to examine activity of CarEs. In

particular, the  $\alpha$ -NA is more frequently implicated in the kinetic analysis of insect CarEs. We previously found that CarE001G was more efficient in its binding with and hydrolysis of  $\alpha$ -NA compared to  $\beta$ -NA.<sup>46</sup> In this work, the kinetic constants of mutant CarEs toward  $\alpha$ -NA were determined. In comparison with the wild type, the results indicated that all four mutant CarEs showed relatively high enzymatic activities toward  $\alpha$ -NA (Table 2). Moreover, the  $CL_{int}$  values were also measured in this work, the  $CL_{int}$  values of Gv1 and Gv2 were significantly larger than that of wild type, which were in accordance with the trend of  $k_{cat}/K_m$  values change. Additionally, the inhibition assay suggested that they were strongly inhibited by  $\beta$ -cypermethrin with  $IC_{50}$  values between 400 and 550  $\mu\text{mol L}^{-1}$

(Table 3 and Fig. 6). Notably,  $\beta$ -cypermethrin exhibited the strongest inhibition on Gv4 ( $IC_{50} = 408.5 \pm 21.5 \mu\text{mol L}^{-1}$ ) compared to either the positive control ( $IC_{50} = 539.9 \pm 82.3 \mu\text{mol L}^{-1}$ ) or the wild-type CarE001G ( $IC_{50} = 501.2 \pm 22.7 \mu\text{mol L}^{-1}$ ). This suggested strong binding between Gv4 and  $\beta$ -cypermethrin. Additionally, the inhibition of  $\beta$ -cypermethrin against the other three mutant CarEs, Gv1, Gv2 and Gv3 were also strong as their  $IC_{50}$  values were lower or close to their positive controls, while they were close to or even slightly higher than that of the wild-type CarE00G. These results indicated that  $\beta$ -cypermethrin could bind to the active sites of these four mutant CarEs.

We also noted that the inhibition constants ( $IC_{50}$  values) for  $\beta$ -cypermethrin against both mutant CarEs and wild type were significantly higher ( $300\text{--}550 \mu\text{mol L}^{-1}$ ) than those in a previous study on an Australian *H. armigera* strain,<sup>50</sup> in which the  $IC_{50}$  values were  $0.52 \mu\text{mol L}^{-1}$  and about  $5 \mu\text{mol L}^{-1}$  for  $\zeta$ -cypermethrin and  $\alpha$ -cypermethrin, respectively. The lower inhibitory potency of the CarEs in the current work is probably attributed to the following two factors. The first is likely to be relevant to excess substrate inhibition as the inhibitor  $\beta$ -cypermethrin was also a substrate for the CarEs. The other is likely to be the amount of CarE enzymes, which may be larger than the optimal amount for the inhibition study to some extent even if it has been already optimized by two-fold serial dilutions in the kinetic assay. However, the next laboratory researches including investigation of excess substrate inhibition and optimization of the inhibition assay are further needed in the future.

CarEs are enzymes in  $\alpha/\beta$  hydrolase fold and can hydrolyze carboxyl ester to its component alcohol and acid.<sup>20</sup> Therefore, CarEs may have metabolic activity toward SPs because most commercial SP insecticides belong to carboxyl esters (i.e.  $\beta$ -cypermethrin,  $\lambda$ -cyhalothrin and fenvalerate). In this study, we conducted an *in vitro* metabolism assay using HPLC to investigate the metabolic activity of the mutant CarEs toward  $\beta$ -cypermethrin (Figs 7 and S5). The results suggested that all the four mutant enzymes showed significantly higher metabolic activities toward  $\beta$ -cypermethrin, with specific activities varied from 4.0 to 5.6  $\text{nmol min}^{-1} \text{mg}^{-1} \text{protein}$  (Fig. 7). Particularly, among these mutant enzymes, Gv4 exhibited the highest metabolic activity toward  $\beta$ -cypermethrin with a specific activity of  $5.6 \pm 1.1 \text{ nmol min}^{-1} \text{mg}^{-1} \text{protein}$ , which are over three-fold higher than that of the wild-type CarE001G. These results matched the results from the insecticide inhibition analysis. However, comparison of the metabolic activity toward  $\beta$ -cypermethrin with that toward  $\alpha$ -NA, failed to provide a consistent pattern between these two different types of substrates. This may be attributed to the different chemical structures of the artificial substrate ( $\alpha$ -NA) and the toxic insecticide ( $\beta$ -cypermethrin), one of the *alpha*-cyano-3-phenoxybenzyl pyrethroids.<sup>45</sup>

As mentioned earlier, the *in vitro* metabolism assay showed that all the mutant CarEs had at least a two-fold increase of metabolic activity toward  $\beta$ -cypermethrin (Fig. 7). The results indicated that the deletion mutations of the GRR sequence caused qualitative changes in CarE001G and therefore its enhancement in insecticide detoxification. In previous work, single point mutations had been demonstrated to confer OP resistance in *Lucilia cuprina*,<sup>33,35</sup> and the *in vitro* mutagenesis and heterologous expression had confirmed that they enhanced the hydrolysis of either OPs or individual SP isomers.<sup>57,58</sup> In another study, it was assumed that Asp and Leu mutations may be a common mechanism in the development of OP resistance in insects, because these point mutations in a few CarEs from different insect orders

were shown to enhance the hydrolysis of OPs.<sup>59</sup> In this work, we found that deletion mutations in the GRR sequence of CarE001G can enhance the hydrolysis of  $\beta$ -cypermethrin. Previous proteomic studies had suggested that *CarE001G* is probably involved in OPs and SPs resistance in both Australian and Chinese *H. armigera* strains,<sup>13,43,60</sup> but to date such resistance has not been observed in any natural populations or laboratory-selected resistant strains of *H. armigera*. In a recent study, Asp and Leu mutations (Ala127 to Asp and Phe238 to Leu) were introduced in CarE001G and expressed in a baculovirus system, but no significant improvement was detected in the hydrolysis of either OPs or SPs.<sup>41</sup> These may indicate that the enhancement of the hydrolysis of insecticides by CarE001G may be caused by novel mechanisms such as the deletion mutations of the GRR sequence. This implies that the deletion mutations of the GRR sequence may have potential in the development of SP resistance in *H. armigera* in the future.

In this study, a homology modeling using the crystal structure of LcaE7 (PDB ID: 5TYM, resolution = 1.84 Å) as a suitable template was carried out to assess the effects of the deletions of part of the GRR sequence on the binding mechanism between CarE and insecticide compounds.<sup>52</sup> The intact 3D structure of the whole sequence of CarE001G (CarE-wt) was partly built according to the crystal structure of LcaE7 as template, finalized and optimized by a 200 ns MD simulation (Fig. S6(A, B)). The results of MD simulations showed that both the 3D structures of CarE-wt and CarE-mut were stable (Figs S6 and S7). The binding mode analyses further demonstrated that  $\beta$ -cypermethrin molecules were better oriented to the residues Ser205, Glu332 and His445 of the catalytic triad of CarE-mut compared with the CarE-wt (Fig. 8(C)). This may explain why the CarE-mut exhibited higher metabolic activity toward  $\beta$ -cypermethrin than that of CarE\_wt. Similarly, in *Lucilia cuprina*, the G137D mutation of *LcaE7* was found to alter the orientation of the Ser200 attacking water molecule.<sup>33</sup> The 3D crystal structure of LcaE7 was further determined as the first insect carboxylesterase.<sup>56</sup> Based on this crystal structure, the binding mode analysis revealed that the molecular mechanism underlying the enhanced hydrolase activity toward OP which was caused by a single amino acid substitution. It was discovered that in the active site of mutant LcaE7, an asymmetric and hydrophobic substrate binding cavity was found to be well-suited to the substrate.<sup>56</sup> In another investigation of the evolution of *LcaE7*, molecular docking analysis revealed that substitutions of several residues around the catalytic triad (Ser-Glu-His) enlarged the volume of the binding pocket and therefore increased the metabolic activity toward deltamethrin.<sup>26</sup> Notably, based on the crystal structure and binding mode analysis of LcaE7, development of inhibitors of CarE which can be used as novel insecticides for insect resistance control can be done through computational inhibitor design.<sup>52</sup>

## 5 CONCLUSIONS

In this work, four CarE variants were identified and successfully expressed in a bacterial system. The HPLC assay showed that the deletion mutations in GRR sequence can significantly enhance the detoxification of  $\beta$ -cypermethrin by increasing the enzyme hydrolytic activity. Binding mode analyses further confirmed the results of the metabolism assay. This indicates that the deletion mutations of the GRR sequence may play a role in the development of SP resistance in *H. armigera* in the future. Further researches on determination of the complete 3D structure by

crystallization and functional analyses through either site-directed or random mutagenesis are needed, in order to comprehensively understand the potential role of *CarE001G* and its mutants in insecticide detoxification and resistance in *H. armigera*.

## ACKNOWLEDGEMENTS

The authors thank Professor De-guang Liu in Northwest A&F University for proof reading this manuscript. This work was supported by the National Natural Science Foundation of China (grants U1903111), the Fundamental Research Funds of Northwest A&F University (grant 2452018109 and 2452015013) and the Provincial Undergraduate Innovation and Entrepreneurship Training Programs of Northwest A&F University (grant S202010712422).

## SUPPORTING INFORMATION

Supporting information may be found in the online version of this article.

## REFERENCES

- Denholm I, Pickett JA, Devonshire AL and McCaffery AR, Resistance to insecticides in Heliiothine Lepidoptera: a global view. *Philos Trans R Soc of Lond B Biol Sci* **353**:1735–1750 (1998).
- Fitt GP, The ecology of heliothis species in relation to agroecosystems. *Annu Rev Entomol* **34**:17–53 (1989).
- Forrester NW, Cahill M, Bird LJ and Layland JK, Management of pyrethroid and endosulfan resistance in *Helicoverpa-armigera* (Lepidoptera, Noctuidae) in Australia. *B Entomol Res* **1**:R1–R132 (1993).
- Armes NJ, Jadhav DR and DeSouza KR, A survey of insecticide resistance in *Helicoverpa armigera* in the Indian subcontinent. *B Entomol Res* **86**:499–514 (1996).
- Brun-Barale A, Hema O, Martin T, Suraporn S, Audant P, Sezutsu H *et al.*, Multiple P450 genes overexpressed in deltamethrin-resistant strains of *Helicoverpa armigera*. *Pest Manag Sci* **66**:900–909 (2010).
- Sanahuja G, Banakar R, Twyman RM, Capell T and Christou P, *Bacillus thuringiensis*: a century of research, development and commercial applications. *Plant Biotechnol J* **9**:283–300 (2011).
- Zhang DD, Xiao YT, Chen WB, Lu YH and Wu KM, Field monitoring of *Helicoverpa armigera* (Lepidoptera: Noctuidae) Cry1Ac insecticidal protein resistance in China (2005–2017). *Pest Manag Sci* **75**:753–759 (2019).
- Jin L, Wang J, Guan F, Zhang JP, Yu S, Liu SY *et al.*, Dominant point mutation in a tetraspanin gene associated with field-evolved resistance of cotton bollworm to transgenic Bt cotton. *Proc Natl Acad Sci USA* **115**:11760–11765 (2018).
- Zhang HN, Tian W, Zhao J, Jin L, Yang J, Liu CH *et al.*, Diverse genetic basis of field-evolved resistance to Bt cotton in cotton bollworm from China. *Proc Natl Acad Sci USA* **109**:10275–10280 (2012).
- Wu KM, Lu YH, Feng HQ, Jiang YY and Zhao JZ, Suppression of cotton bollworm in multiple crops in China in areas with Bt toxin-containing cotton. *Science* **321**:1676–1678 (2008).
- Yang YH, Li YP and Wu YD, Current status of insecticide resistance in *Helicoverpa armigera* after 15 years of Bt cotton planting in China. *J Econ Entomol* **106**:375–381 (2013).
- Xu L, Li D, Qin J, Zhao W and Qiu L, Over-expression of multiple cytochrome P450 genes in fenvalerate-resistant field strains of *Helicoverpa armigera* from north of China. *Pestic Biochem Phys* **132**:53–58 (2016).
- Han YC, Wu SW, Li YP, Liu JW, Campbell PM, Farnsworth C *et al.*, Proteomic and molecular analyses of esterases associated with monocrotophos resistance in *Helicoverpa armigera*. *Pestic Biochem Phys* **104**:243–251 (2012).
- Qayyum MA, Wakil W, Arif MJ, Sahi ST, Saeed NA and Russell DA, Multiple resistances against formulated organophosphates, pyrethroids, and newer-chemistry insecticides in populations of *Helicoverpa armigera* (Lepidoptera: Noctuidae) from Pakistan. *J Econ Entomol* **108**:286–293 (2015).
- Wang ZG, Jiang SS, Mota-Sanchez D, Wang W, Li XR, Gao YL *et al.*, Cytochrome P450-mediated  $\lambda$ -cyhalothrin-resistance in a field strain of *Helicoverpa armigera* from northeast China. *J Agr Food Chem* **67**:3546–3553 (2019).
- Jousen N, Agnolet S, Lorenz S, Schone SE, Ellinger R, Schneider B *et al.*, Resistance of Australian *Helicoverpa armigera* to fenvalerate is due to the chimeric P450 enzyme CYP337B3. *Proc Natl Acad Sci U S A* **109**:15206–15211 (2012).
- Oakeshott JG, Farnsworth CA, East PD, Scott C, Han YC, Wu YD *et al.*, How many genetic options for evolving insecticide resistance in heliothine and spodopteran pests? *Pest Manag Sci* **69**:889–896 (2013).
- Abd El-Latif AO and Subrahmanyam B, Pyrethroid resistance and esterase activity in three strains of the cotton bollworm, *Helicoverpa armigera* (Hubner). *Pestic Biochem Phys* **96**:155–159 (2010).
- Shi Y, Wang H, Liu Z, Wu S, Yang Y, Feyereisen R *et al.*, Phylogenetic and functional characterization of ten P450 genes from the CYP6AE subfamily of *Helicoverpa armigera* involved in xenobiotic metabolism. *Insect Biochem Mol* **93**:79–91 (2018).
- Whelock CE, Shan G and Ottea J, Overview of carboxylesterases and their role in the metabolism of insecticides. *J Pestic Sci* **30**:75–83 (2005).
- Oakeshott JG, Claudianos C, Campbell PM, Newcomb RD and Russell RJ, 5.10 – Biochemical genetics and genomics of insect esterases, in *Comprehensive Molecular Insect Science*, ed. by Gilbert LI. Elsevier, Amsterdam, pp. 309–381 (2005).
- Satoh T and Hosokawa M, Structure, function and regulation of carboxylesterases. *Chem Biol Interact* **162**:195–211 (2006).
- Devonshire AL and Moores GD, A carboxylesterase with broad substrate specificity causes organophosphorus, carbamate and pyrethroid resistance in peach-potato aphids (*Myzus persicae*). *Pestic Biochem Phys* **18**:235–246 (1982).
- Mouches C, Magnin M, Berge JB, de Silvestri M, Beyssat V, Pasteur N *et al.*, Overproduction of detoxifying esterases in organophosphate-resistant *Culex* mosquitoes and their presence in other insects. *Proc Natl Acad Sci USA* **84**:2113–2116 (1987).
- Ai GM, Zou DY, Shi XY, Li FG, Liang P, Song DL *et al.*, HPLC assay for characterizing alpha-cyano-3-phenoxybenzyl pyrethroids hydrolytic metabolism by *Helicoverpa armigera* (Hubner) based on the quantitative analysis of 3-phenoxybenzoic acid. *J Agr Food Chem* **58**:694–701 (2010).
- Coppin CW, Jackson CJ, Sutherland T, Hart PJ, Devonshire AL, Russell RJ *et al.*, Testing the evolvability of an insect carboxylesterase for the detoxification of synthetic pyrethroid insecticides. *Insect Biochem Mol* **42**:343–352 (2012).
- Yang XQ, Liu JY, Li XC, Chen MH and Zhang YL, Key amino acid associated with acephate detoxification by *Cydia pomonella* carboxylesterase based on molecular dynamics with alanine scanning and site-directed mutagenesis. *J Chem Inf Model* **54**:1356–1370 (2014).
- Wang LL, Lu XP, Smagghe G, Meng LW and Wang JJ, Functional characterization of BdB1, a well-conserved carboxylesterase among tephritid fruit flies associated with malathion resistance in *Bactrocera dorsalis* (Hendel). *Comp Biochem Phys C* **200**:1–8 (2017).
- Zhang JQ, Ge PT, Li DQ, Guo YP, Zhu KY, Ma EB *et al.*, Two homologous carboxylesterase genes from *Locusta migratoria* with different tissue expression patterns and roles in insecticide detoxification. *J Insect Physiol* **77**:1–8 (2015).
- Vontas JG, Small GJ and Hemingway J, Comparison of esterase gene amplification, gene expression and esterase activity in insecticide susceptible and resistant strains of the brown planthopper, *Nilaparvata lugens* (Stal). *Insect Mol Biol* **9**:655–660 (2000).
- Cao CW, Zhang J, Gao XW, Liang P and Guo HL, Differential mRNA expression levels and gene sequences of carboxylesterase in both deltamethrin resistant and susceptible strains of the cotton aphid, *Aphis gossypii*. *Insect Sci* **15**:209–216 (2008).
- Wang LL, Lu XP, Meng LW, Huang Y, Wei D, Jiang HB *et al.*, Functional characterization of an alpha-esterase gene involving malathion detoxification in *Bactrocera dorsalis* (Hendel). *Pestic Biochem Phys* **130**:44–51 (2016).
- Newcomb RD, Campbell PM, Ollis DL, Cheah E, Russell RJ and Oakeshott JG, A single amino acid substitution converts a carboxylesterase to an organophosphorus hydrolase and confers insecticide resistance on a blowfly. *Proc Natl Acad Sci USA* **94**:7464–7468 (1997).
- Claudianos C, Russell RJ and Oakeshott JG, The same amino acid substitution in orthologous esterases confers organophosphate

- resistance on the house fly and a blowfly. *Insect Biochem Mol* **29**: 675–686 (1999).
- 35 Campbell PM, Newcomb RD, Russell RJ and Oakeshott JG, Two different amino acid substitutions in the ali-esterase, E3, confer alternative types of organophosphorus insecticide resistance in the sheep blowfly, *Lucilia cuprina*. *Insect Biochem Molec* **28**:139–150 (1998).
- 36 Field LM and Devonshire AL, Evidence that the E4 and FE4 esterase genes responsible for insecticide resistance in the aphid *Myzus persicae* (Sulzer) are part of a gene family. *Biochem J* **330**:169–173 (1998).
- 37 Mouches C, Pasteur N, Berge JB, Hyrien O, Raymond M, de Saint Vincent BR *et al.*, Amplification of an esterase gene is responsible for insecticide resistance in a California *Culex* mosquito. *Science* **233**:778–780 (1986).
- 38 Wang LL, Huang Y, Lu XP, Jiang XZ, Smagghe G, Feng ZJ *et al.*, Overexpression of two-esterase genes mediates metabolic resistance to malathion in the oriental fruit fly, *Bactrocera dorsalis* (Hendel). *Insect Mol Biol* **24**:467–479 (2015).
- 39 Achaleke J, Martin T, Ghogomu RT, Vaissayre M and Brevault T, Esterase-mediated resistance to pyrethroids in field populations of *Helicoverpa armigera* (Lepidoptera: Noctuidae) from central Africa. *Pest Manag Sci* **65**:1147–1154 (2009).
- 40 Wu S, Yang Y, Yuan G, Campbell PM, Teese MG, Russell RJ *et al.*, Overexpressed esterases in a fenvalerate resistant strain of the cotton bollworm, *Helicoverpa armigera*. *Insect Biochem Mol Biol* **41**:14–21 (2011).
- 41 Li YQ, Farnsworth CA, Coppin CW, Teese MG, Liu JW, Scott C *et al.*, Organophosphate and pyrethroid hydrolase activities of mutant esterases from the cotton bollworm *Helicoverpa armigera*. *PLoS One* **8**:1–7 (2013).
- 42 Gunning RV, Devonshire AL and Moores GD, Metabolism of esfenvalerate by pyrethroid-susceptible and -resistant Australian *Helicoverpa armigera* (Lepidoptera: Noctuidae). *Pestic Biochem Phys* **51**:205–213 (1995).
- 43 Teese MG, Campbell PM, Scott C, Gordon KHJ, Southon A, Hovan D *et al.*, Gene identification and proteomic analysis of the esterases of the cotton bollworm, *Helicoverpa armigera*. *Insect Biochem Mol* **40**:1–16 (2010).
- 44 Teese MG, Farnsworth CA, Li YQ, Coppin CW, Devonshire AL, Scott C *et al.*, Heterologous expression and biochemical characterisation of fourteen esterases from *Helicoverpa armigera*. *PLoS One* **8**:1–9 (2013).
- 45 Li YQ, Liu JW, Lu M, Ma ZQ, Cai CL, Wang YH *et al.*, Bacterial expression and kinetic analysis of carboxylesterase 001D from *Helicoverpa armigera*. *Int J Mol Sci* **17**:1–13 (2016).
- 46 Bai LS, Zhao CX, Xu JJ, Feng C, Li YQ, Dong YL *et al.*, Identification and biochemical characterization of carboxylesterase 001G associated with insecticide detoxification in *Helicoverpa armigera*. *Pestic Biochem Phys* **157**:69–79 (2019).
- 47 Li YQ, Bai LS, Zhao CX, Xu JJ, Sun ZJ, Dong YL *et al.*, Functional characterization of two carboxylesterase genes involved in pyrethroid detoxification in *Helicoverpa armigera*. *J Agr Food Chem* **68**: 3390–3402 (2020).
- 48 Bradford MM, A rapid and sensitive method for the quantitation of microgram quantities of protein utilizing the principle of protein-dye binding. *Anal Biochem* **72**:248–254 (1976).
- 49 Han X, Nabb DL, Mingoia RT and Yang C-H, Determination of xenobiotic intrinsic clearance in freshly isolated hepatocytes from rainbow trout (*Oncorhynchus mykiss*) and rat and its application in bioaccumulation assessment. *Environ Sci Technol* **41**:3269–3276 (2007).
- 50 Gunning RV, Moores GD, Jewess P, Boyes AL, Devonshire AL and Khambay BPS, Use of pyrethroid analogues to identify key structural features for enhanced esterase resistance in *Helicoverpa armigera* (Huebner) (Lepidoptera: Noctuidae). *Pest Manag Sci* **63**:569–575 (2007).
- 51 Hamilton MA, Russo RC and Thurston RV, Trimmed Spearman–Kärber method for estimating median lethal concentrations in toxicity bioassays. *Environ Sci Technol* **11**:714–719 (1977).
- 52 Correy GJ, Zaidman D, Harmelin A, Carvalho S, Mabbitt PD, Calaora V *et al.*, Overcoming insecticide resistance through computational inhibitor design. *Proc Natl Acad Sci USA* **116**:21012–21021 (2019).
- 53 Götz AW, Williamson MJ, Xu D, Poole D, Le Grand S and Walker RC, Routine microsecond molecular dynamics simulations with AMBER on GPUs. 1. Generalized born. *J Chem Theory Comput* **8**:1542–1555 (2012).
- 54 Trott O and Olson AJ, AutoDock Vina: improving the speed and accuracy of docking with a new scoring function, efficient optimization, and multithreading. *J Comput Chem* **31**:455–461 (2010).
- 55 Morris GM, Huey R, Lindstrom W, Sanner MF, Belew RK, Goodsell DS *et al.*, AutoDock4 and AutoDockTools4: automated docking with selective receptor flexibility. *J Comput Chem* **30**: 2785–2791 (2009).
- 56 Jackson CJ, Liu JW, Carr PD, Younus F, Coppin C, Meirelles T *et al.*, Structure and function of an insect alpha-carboxylesterase (alpha esterase 7) associated with insecticide resistance. *Proc Natl Acad Sci USA* **110**:10177–10182 (2013).
- 57 Heidari R, Devonshire AL, Campbell BE, Bell KL, Dorrian SJ, Oakeshott JG *et al.*, Hydrolysis of organophosphorus insecticides by in vitro modified carboxylesterase E3 from *Lucilia cuprina*. *Insect Biochem Mol* **34**:353–363 (2004).
- 58 Devonshire AL, Heidari R, Huang HZ, Hammock BD, Russell RJ and Oakeshott JG, Hydrolysis of individual isomers of fluorogenic pyrethroid analogs by mutant carboxylesterases from *Lucilia cuprina*. *Insect Biochem Mol* **37**:891–902 (2007).
- 59 Cui F, Lin Z, Wang HS, Liu SL, Chang HJ, Reeck G *et al.*, Two single mutations commonly cause qualitative change of nonspecific carboxylesterases in insects. *Insect Biochem Mol* **41**:1–8 (2011).
- 60 Pearce SL, Clarke DF, East PD, Elfekih S, Gordon KHJ, Jermini LS *et al.*, Genomic innovations, transcriptional plasticity and gene loss underlying the evolution and divergence of two highly polyphagous and invasive *Helicoverpa* pest species. *BMC Biol* **15**:30 (2017).