Identification of a novel carboxylesterase dominantly expressed in the booklouse *Liposcelis bostrychophila*

Osamu Ishibashi*

Laboratory of Biological Macromolecules, Graduate School of Agriculture, Osaka Metropolitan University, Sakai 599-8531, Japan

*To whom correspondence should be addressed. E-mail: ishibos1@omu.ac.jp

Keywords: *Liposcelis bostrychophila*, booklouse, carboxylesterase, insecticide, synergist

Summary

Booklice are tiny insect pests commonly found in an indoor environment and characterized by extremely high proliferation ability. Booklouse contamination in stored foods causes serious food loss; therefore, they are recognized as so-called stored-food pests. Furthermore, booklice in house dust have been reported to be a potent allergen by which approximately 20% of asthma patients are sensitized. Therefore, the safe extermination of booklice would contribute to the solution of both food loss and allergic concerns. Organophosphates (OPs) are compounds generally used as harmless insecticides against various pests; nonetheless, cases of OP poisoning are sometimes reported worldwide. Therefore, considering usage for protecting stored food products, OP doses should be reduced as far as possible. It is generally known that carboxylesterase-mediated inactivation of OPs decreases the efficacy of OPs, suggesting that an inhibitor of carboxylesterase could be a synergist to reduce the effective dose of OP. Here, I report the molecular cloning and characterization of a novel carboxylesterase highly expressed in a representative species of indoor booklice, *Liposcelis bostrychophila*.

Introduction

Losses in the food chain from production to harvest, storage, distribution, processing, sales, and consumption have recently been recognized as a problem because they not only cause possible food shortages but also indirectly contribute to global warming. It was reported that global food loss and waste during 2010–2016 were estimated to equal 8–10% of the total anthropogenic greenhouse gas emissions and cost about 1 trillion USD per year [1]. Further, it is estimated that more than 40% of global food loss is so-called "post-harvest loss" [1]; thus, countermeasures to address this issue should be urgently made. Biotic post-harvest food losses include the infestation of microorganisms, pests, and rodents. Pests, which

include some insects and mites, occasionally infest food products, *e.g.*, grains and cereals, during storage and distribution. Particularly, insect pests, including cockroaches, ants, molds, warehouse beetles, and booklice (psocids), are considered a nuisance and cause huge losses in the grain [2]. The amounts of stored grain lost due to insect infestation are estimated to be 5% to 10% in developed countries and as much as 35% in developing countries [3]. Nonetheless, the insect pests, particularly booklice, had not attracted much attention for a long period because solid evidence as to quantitative and qualitative food losses caused by booklice was missing. However, the status of booklice in terms of food pests began to change in the late 1980s as booklouse infestation in stored grains in diverse locations worldwide was reported [4]. Furthermore, Kučerová reported quantitative data that the average weight of grain samples (broken wheat kernels) infested with booklice decreased by 9.7% after 3 months of infestation [5]. These reports indicate that food loss caused by booklouse infection should be seriously considered.

Booklice are tiny insects measuring 1-4 mm in length. The proliferation of booklice is extremely rapid. Therefore, once the outbreak of booklice occurs, it is extremely difficult to eliminate them. Booklice include numerous species, and some of them are often found in food factories and warehouses. Booklice generally prefer to feed on molds; thus, they can infest stored foods not only directly but also indirectly through food-infesting molds. Among the booklouse species, *Liposcelis* (*L.*) *bostrycophila*, *L. decolor*, *L. entomophila*, and *L. paeta* are common food-infesting booklice. Of these, *L. bostrychophila* is a representative species of indoor booklice, which are commonly found in food facilities, as well as ordinary houses; therefore, it has been studied more commonly than the other species [5].

Additionally, it has recently been reported that *L. bostrychophila* in house dust is an allergen that causes allergic asthma [6,7]. Considering that the cases of severe allergic symptoms developed immediately after eating mite-contaminated foods have been reported [8], accidental ingestion of foods infested with booklice may induce allergic symptoms. Although it is possible to suppress the proliferation of booklice by maintaining proper temperature and humidity, providing an environment suitable for food storage is sometimes difficult for various reasons such as cost issues. Thus, the safe extermination of booklice using less harmful insecticides would contribute to the solution of food loss and health hazard problems, which could lead to the achievement of SDGs. More effective and less harmful insecticides have been developed so far. However, their effectiveness is not sufficient to extirpate booklice, which is attributed to their extremely high proliferative capacity and the emergence of insecticide-resistant booklice [9]. Of insecticides, organophosphates (OPs), which are known as acetylcholinesterase (AChE) inhibitors, exert a relatively high anti-booklouse effect [10]. Therefore, high-dose OPs have the potential to extirpate booklice. However, OPs should be carefully used for food control, as cases of OP poisoning often occur even though OPs are considered to have low toxicity to humans. Additionally, as with many other insects, the acquisition of resistance to OPs has been reported for booklice [11]. Thus far, several

mechanisms underlying the acquisition of resistance to OPs have been documented [12,13], which includes carboxylesterase (CE)-mediated OP inactivation in the insect body. Notably, Correy *et al*. showed that a specific inhibitor of α-carboxylesterase (αE7), which was designed based on its protein structure, significantly suppressed OP degradation and increased the efficacy of OPs, indicating that the combination of OPs and CE inhibitors possibly enables to extirpate at lower insecticide concentrations [14].

Here, we report the molecular cloning of a novel CE expressed at a high level in *L. bostrycophila*, designated LBCE1, and the generation of recombinant LBCE1 protein that shows activity to hydrolyze *p*-nitrophenol acetic acid (PNPA), a general substrate for sensitive esterase assay. Interestingly, LBCE1 has also been demonstrated to show a weak AChE activity.

Methods

Analysis of RNA-sequencing dataset

The RNA-sequencing dataset from *L. bostrychophila* (DRX080823), which was previously obtained and registered by the author in the Sequence Reads Archive database, was analyzed for *de novo* contig construction and annotation. Contigs were generated by assembling qualified sequence reads, as described previously [7]. Annotation of the contig sequences was carried out at both nucleotide and amino acid sequence levels as follows. First, the contig sequences were subjected to a blastn search against the nr/nt database of NCBI. Second, open reading frames (ORFs) in the contig sequences were predicted using the TransDecoder tool, and amino acid sequences deduced from the ORF sequences were subjected to blastp search against the UniProtKB database (https://www.uniprot.org/). Gene ontology terms were then assigned to the individual contigs based on the retrieved UniProt IDs using the Blast2Go tool.

Fragments per kilobase of exon per million reads mapped (FPKM) values were calculated as an index for gene expression level by the Expectation-Maximization (RSEM) tool. For this process, the *de novo*-constructed contigs were used as a reference.

PCR-based molecular cloning of *L.bostrychophila* carboxylesterase 1 (LBCE1)

Total RNA extracted from *L. bostrychophila* bodies was subjected to reverse transcription using a PrimeScript RT Master Mix (Perfect Real Time) (Takara-bio, Japan) and the LBCE1-reverse primer. This was followed by PCR using KOD-plus2 DNA polymerase (Takara-bio) and the primers for LBCE1. The sequences of the forward and reverse primers used are 5'-CTGAATTCAATGCAGTTCGGCTCCGACCT-3' and 5'-TTGTCGACCTTGTGTCTTGTGCAAACGGA-3', respectively (the underlined nucleotides indicate the

103 restriction sites used for cloning). The PCR products were purified, inserted in pGEM-Teasy (Promega),
104 and subjected to Sanger sequencing to verify their nucleotide sequences. The determined nucleotide
105 sequence was registered in the DNA Data Bank of Japan under the submission identifier LC742390. This
106 was then followed by the subcloning of the PCR product into the expression vector pET22b to construct
107 pET22b-LBCE1.
108
109 **Expression and purification of LBCE1 using an *E. coli* expression system**
110 *E. coli* cells of the strain BL21 (DE3) were transformed with pET22b-LBCE1, and the expression of
111 LBCE1 tagged with 6x His at the C-terminus (LBCE1-His) was induced by culturing the cells in a 500-ml
112 flask containing 100 ml of an autoinduction medium, OvernightExpress TB medium (Sigma)
113 supplemented with 1% glycerol, for 8 h. Since LBCE1-His was shown to be secreted and/or leaked into
114 the culture medium at a detectable level, it was purified from the culture medium as follows. First, the
115 culture medium was concentrated and buffer-changed to the equivalation buffer (20 mM phosphate, 500
116 mM NaCl, pH7.4) using an Amicon Ultra-15 filtration unit (10-kDa cutoff) (Merck-Millipore). The
117 concentrate was then applied to an affinity column filled with TALEN Metal Affinity Resin (Takara-bio)
118 and washed with the equivalation buffer. Resin-bound proteins were then eluted with the same buffer
119 containing 10 mM imidazole. The concentration of the purified LBCE1 protein was determined according
120 to the Bradford protein assay [15].
121
122 **LBCE1 activity assay**
123 The recombinant LBCE1 protein was assayed for esterase activity at pH4-8 using *p*-nitrophenol acetic
124 acid (pNPA) as a substrate as described earlier [16]. The hydrolysis of the substrate was monitored by
125 measuring OD_{405}. Since the substrate can be spontaneously hydrolyzed, data in the enzyme-free condition
126 were also taken as background and used for subtraction.
127
128 **Western blotting**
129 The purified LBCE1 was subjected to SDS-PAGE, transferred to a nylon membrane, and probed with the
130 anti-6xHis tag monoclonal antibody (Clone 6C4) (MBL, Tokyo, Japan) diluted at 1:1,000 in TBST
131 containing 5% skim milk. The blot was then subjected to immunodetection using a WesternBreeze
132 Chromogenic Kit, anti-mouse (Thermo) according to the manufacturer's instructions.
133
134 **Protein structure modeling and docking simulation**
135 The 3D-structural model of LBCE1 was constructed using the protein structure prediction tool
136 Alphafold2.1 Notebook (https://github.com/deepmind/alphafold) [17] with HHsearch through PDB70.

The analysis returned five model structures, which were ranked in the order of accuracy. Docking simulation was performed using Webina (https://durrantlab.pitt.edu/webina/), a web application that runs AutoDock Vina [18]. *In silico* detection of surface cavities was performed using the web-based tool CavityPlus (http://www.pkumdl.cn:8000/cavityplus/index.php#/) [19].

Results and Discussion

Identification of novel carboxylesterases in *L. bostrychophila* using the RNA-sequencing dataset DRX080823

I first searched for contigs that encompass complete or >1,000-nt partial ORFs coding for possible CEs from the RNA sequencing dataset of *L. bostrychophila*, which was previously obtained by my own [7]. Of the contigs with FPKM values of more than 30, 13 contigs were identified to be associated with the gene ontology term "carboxylic ester hydrolase activity" based on their homology with CEs previously identified from various organisms (Table S1). Of these contigs, Comp61175, Comp62403, and Comp61080 are highly homologous to the previously identified *L. bostrychophila* esterases, esterase-1 (95% identity), -2 (96% identity), and -4 (97% identity), respectively (Table S1); therefore, they are conceivably derived from these known esterase genes. Notably, the top-6 contigs in the order of the FPKM value are not highly homologous to the four CEs previously identified from *L. bostrychophila* (LB esterases 1-4) (Table S1), indicating that there remain unidentified CEs that are abundantly expressed in *L. bostrychophila*.

A contig with the highest FPKM value at a gene level is Comp59220, which is predicted to encompass a complete ORF encoding a putative novel CE consisting of 554 amino acids (Figure S1). The amino acid sequence is moderately, but not highly, similar to the previously identified *L. bostrychophila* esterases (44.3% identity at the most with *L. bostrychophila* esterase-1). Then, I designated the novel CE *L. bostrychophila* carboxylesterase 1 (LBCE1). Previous studies showed that esterase E4/FE4 activity elevated by gene amplification in insecticide-resistant insects was involved in the detoxication of insecticides such as OPs [20]. Therefore, considering the high FPKM value of Comp59220, it is likely that LBCE1 is involved in the desensitization to OPs. Then, I was motivated to focus on LBCE1 as it might contribute to a decrease in the efficacy of OPs against the booklouse *L. bostrychophila*.

Molecular cloning of LBCE1 cDNA by polymerase chain reaction

A cDNA encompassing the putative coding sequence (CDS) of LBCE1 was successfully cloned by RT-PCR and sanger-sequenced (Figure 1). Consistently, the RNA- and genomic DNA-sequencing datasets of *L. Bostrychophila* (registered as ERR073018, SRR17191995, and SRR17191998 in the NCBI Sequence

171 Read Archive) contained sequence reads homologous to the LBCE1 cDNA (Figure S2). Furthermore, the
172 nucleotide sequence registered as GAYV02033066.1 in the NCBI Transcriptome Shotgun Assembly
173 database shows 99% identity with the CDS of LBCE1 (Figure S3). Considering that these datasets were
174 obtained from *L. bostrychophila* collected independently in different countries, I consider that the LBCE1
175 gene is present and transcribed in *L. bostrychophila* inhabiting many areas worldwide.

176 The amino acid sequence deduced from the LBCE1 cDNA was shown to match that deduced from the
177 contig Comp59220, except for the 3 amino acid residues (Figures 1 and S1). Concretely, Leu37, Pro280,
178 and Glu289 in the amino acid sequence deduced from the Comp59220 sequence were replaced by Ser37,
179 Gln280, and Asp289 in LBCE1, respectively (Figure 1). I cloned several cDNAs for LBCE1 and showed
180 that they all have the same CDS; therefore, the difference in sequence is possibly due to the inaccuracy of
181 next-generation sequencing. Using the bioinformatics tool Expasy (https://web.expasy.org/compute_pi/),
182 the theological molecular weight and isometric point and of LBCE1 were calculated to be 62528.48 and
183 7.16, respectively. The *in silico* analysis using SignalP [21] revealed that LBCE1 lacks a possible signal
184 peptide (Figure S4), indicating that it is possibly localized in the cytoplasm. This result is consistent with
185 the fact that 9 of 11 CEs (including CE-related esterases) that have been previously associated with
186 insecticide resistance lack possible signal peptides (Table S2). In contrast, 3 of the four previously
187 identified *L. bostrychophila* esterases, *i.e.*, esterases 2, 3, and 4, have putative signal peptides, suggesting
188 that they are secreted proteins. Multiple alignments indicated that LBCE1 exhibited no more than 46%,
189 38%, 28%, and 37% identities at an amino acid level with *L. bostrychophila* esterases 1, 2, 3, and 4,
190 respectively (Figure 2). BLASTP homology search and phylogenetic tree analyses showed that LBCE1
191 protein showed moderate similarities with CEs of various insects (Table 1 and Figure S5). Of these CEs,
192 the putative esterase FE4 protein of *Pediculus humanus* (human louse) showed the highest identity (47%)
193 with LBCE1, consistent with a previous report that booklice and sucking lice (Anoplura) are evolutionally
194 close [22]. Importantly, amino acid residues forming the esterase catalytic triad, *i.e.*, serine, glutamic acid,
195 and histidine, are conserved in LBCE1 (Figures 1 and 2). Furthermore, common motifs typical for CEs,
196 such as the Gly-X-Ser-X-Gly motif and the site of an oxygen anion hole (His-Gly-Gly), are also
197 conserved in LBCE1 (Figures 1 and 2). Notably, when the homology search was performed against the
198 Protein Data Bank (PDB) database, a dedicated repository for proteins with 3D-structural data, the top hit
199 was *Lucilia cuprina* αE7 (LCαE7), which shared 36% amino acid identity with LBCE1 (Table S3).
200 LCαE7 was previously reported to be involved in the increased OP resistance of *Lucilia cuprina* as it
201 detoxicated OPs [23,24]. Consistently, it was previously reported that the use of *in silico*-designed
202 inhibitors of LCαE7 as synergists significantly increased the insecticide efficacy of OPs. Amino acid
203 sequence identities between LCαE7 and the putative CEs encoded by the contigs that have complete
204 ORFs (Table S1), *i.e.*, Comp59220 (corresponding to LBCE1), Comp62262, Comp61411, Comp62642,

205 Comp60632, and Comp51226, are 36%, 31%, 26%, 31%, 26%, and 26%, respectively. Additionally,
206 LCαE7 share 36%, 35%, 27%, and 31% amino acid identities with the previously identified *L.*
207 *bostrychophila* esterases-1, -2, -3, and -4, respectively. Thus, LBCE1 is conceivably one of the most
208 LCαE7-homologous CEs in *L. bostrychophila*. Taken together with the high expression level of LBCE1,
209 the comparatively high similarity to LCαE7 led me to the idea that LBCE1 can be a potential target for
210 developing OP synergists used for booklouse extermination.
211
212 **Purification and enzyme activity assay of LBCE1**
213 LBCE1-His was expressed using an *E. coli* expression system. Although the expression plasmid of
214 LBCE1-His was designed to drive the periplasmic localization of LBCE1-His by the action of the *pelB*
215 signal sequence, LBCE1-His was detected in the culture medium but not in the periplasm fraction (Figure
216 3A); therefore, most of the LBCE1-His protein was conceivably secreted or leaked into the culture
217 medium. Then, LBCE1-His was purified from the concentrated culture medium by nickel affinity
218 chromatography (Figure 3B). The yield of LBCE1-His after purification was 30 μg from 100 ml of the *E.*
219 *coli* cell culture. The low yield may be attributed to the instability of LBCE1 by the analogy of LCαE7, of
220 which the recombinant enzyme was shown to be unstable without inhibitors [15]. Due to the poor yield, it
221 was difficult to examine the OP-hydrolyzing activity of the recombinant LBCE1. Then, in the present
222 study, it was tested whether the recombinant protein has an esterase activity using pNPA as a sensitive
223 substrate. Since pNPA is hydrolyzed enzyme-independently at alkaline conditions (pH > 8), the esterase
224 activity of LBCE1-His was evaluated between pH5 and pH8. As shown in Figure 3C, the extent of pNPA
225 hydrolysis was low at pH less than 6.4 and dramatically increased at pH higher than 7. Collectively, the
226 recombinant LBCE1 produced using an *E.coli* expression system was demonstrated to be enzymatically
227 active, although efforts to overcome the poor yield is necessary.
228
229 **Structural Modeling of LBCE1**
230 Since it was also difficult to perform 3D-structural analysis, *e.g.*, X-ray crystallography, of LBCE1 due to
231 the poor yield, 3D-structural modeling of LBCE1 was alternatively performed. Five 3D-structural models
232 of LBCE1 were built by Alphafold2. All the predicted models have pIDDT values and pTM scores greater
233 than 90 and 0.9, respectively, suggesting the high reliability of these models. Of note, the overall 3D
234 structures of these models are nearly overlapping except for their N-terminal regions, of which prediction
235 reliability was poor (Figure 4A). However, I do not think that the N-terminal region critically affects the
236 substrate/inhibitor specificity of LBCE1 as it is distant from the catalytic center. In this study, the model
237 3, which was top-ranked based on the pIDDT value of 91.1 and the ptmscore of 0.926, was selected for
238 the following analyses (Figure 4B).

239 Structural superposition revealed that the 3D model of LBCE1 had an overall structure similar to the
240 X-ray crystal structure of LCαE7 (PDB ID: 4FNG) except for its N-terminal α-helix region (Figures 4C-
241 E). The root-mean-square deviation (RMSD) value of the two structures was calculated to be no more
242 than 0.948Å. Importantly, the residues forming the catalytic triad in the two structures are almost
243 overlapping (Figure 4F), suggesting the possibility that they show similar substrate specificities. *In silico*
244 surface cavity detection revealed that LBCE1 had several druggable cavities, but strongly druggable
245 cavities were localized near its catalytic center (Figure S6).
246
247 **Docking simulation of LBCE1 with CE inhibitors**
248 Next, it was evaluated whether LBCE1 potently interacted with OPs at its active center by *in silico*
249 molecular docking. For this purpose, malathion and diazinon were chosen as representative OPs. When
250 grid parameters were set to restrict possible binding sites to the neighbor of the catalytic centers, nine
251 docking poses were predicted for both malathion- and diazinon-binding (Tables S4). The binding sites of
252 malathion and diazinon in LBCE1 were shown to be similar, and the common residues (Met90, Ile142,
253 Ile347, Arg351, Leu447, and Ala462) were predicted to interact with both OPs. Molecular docking
254 analysis also revealed that these OPs are potently bound to sites near the catalytic center of LCαE7.
255 Unexpectedly, the predicted docking poses and ligand-interacting residues of LCαE7 were different from
256 those of LBCE1 (Table S5). This difference can be attributed to the size of cavities leading to the catalytic
257 center. Two strongly druggable cavities leading to the catalytic center were detected on the surface of
258 LBCE1 (Figures 5A and 5B), while only one was on the surface of LCαE7. Although one of the cavities
259 in LBCE1 and the cavity in LCαE7 are located in the corresponding regions in the superposed structures,
260 the former appears narrower than the latter and the other cavity in LBCE1 (Figures 5C-F). Therefore, the
261 OP is considered to dominantly interact with the residues of the wider cavity of LBCE1.
262
263 **Conclusions**
264 In the present study, a novel CE expressed dominantly in *L. bostrychophila*, designated LBCE1, was
265 successfully identified. By structure-based analogy with LCαE7, LBCE1 was speculated to be involved in
266 the detoxication of OPs as well, and the results of my *in silico* analyses including docking simulation
267 indeed supported this possibility. However, the results also raised the possibility that the spectrum of OP
268 docking can be different between LBCE1 and LCαE7.
269 Methods for the genetic manipulation of *L. bostrychophila* have not been established so far; therefore,
270 selective LBCE1 inhibitors are necessary to assess the validity of LBCE1 as a target of OP synergists.
271 Thus, further enzymological studies will be necessary to characterize LBCE1 in detail.
272

Availability of data and materials

The datasets and materials used and/or analyzed for the present study are available from the corresponding author upon reasonable request.

Funding

This work was supported by the Tojuro Iijima Foundation for Food Science and Technology.

Author contributions

OI is responsible for all matters regarding the present study, which includes receiving a fund, conducting both the wet-lab experiments and dry analyses, and writing the manuscript.

Declaration of Competing Interest

The authors declare no conflicts of interest.

Acknowledgment

The production and purification of recombinant LBCE1 were performed with the support of BEX Co. Ltd (Tokyo, Japan).

References

1. Mbow C, Rosenzweig C, Barioni LG, Benton TG, Herrero M, Krishnapillai M, Liwenga E, Pradhan P, Rivera-Ferre MG, Sapkota T, Tubiello FN, Xu Y. 2019: Food Security. In: Climate Change and Land: an IPCC special report on climate change, desertification, land degradation, sustainable land management, food security, and greenhouse gas fluxes in terrestrial ecosystems. doi.org/10.1017/9781009157988.007.
2. Kumar D, Kalita P. Reducing Postharvest Losses during Storage of Grain Crops to Strengthen Food Security in Developing Countries. *Foods* 2017;6:8. doi: 10.3390/foods6010008.
3. Campbell JF, Arthur FH, Mullen MA. Insect management in food processing facilities. *Adv Food Nutr Res*. 2004;48:239-95. doi: 10.1016/S1043-4526(04)48005-X.
4. Nayak MK, Collins PJ, Throne JE, Wang JJ. Biology and management of psocids infesting stored products. *Annu Rev Entomol*. 2014;59:279-97. doi: 10.1146/annurev-ento-011613-161947.
5. Kučerová Z. Weight losses of wheat grain caused by psocid infestation (Liposcelis bostrychophila: Liposcelididae: Psocoptera). *Plant Protect. Sci*. 2002;38:103-107. doi:10.17221/4858-PPS.
6. Fukutomi Y, Kawakami Y, Taniguchi M, Saito A, Fukuda A, Yasueda H, Nakazawa T, Hasegawa M,

Nakamura H, Akiyama K. Allergenicity and cross-reactivity of booklice (*Liposcelis bostrichophila*): a common household insect pest in Japan. *Int Arch Allergy Immunol*. 2012;157:339-348. doi: 10.1159/000329853.

7. Ishibashi O, Sakuragi K, Fukutomi Y, Kawakami Y, Kamata Y, Sakurai M, Nakayama S, Uchiyama H, Kobayashi H, Kojima H, Inui T. Lip b 1 is a novel allergenic protein isolated from the booklouse, *Liposcelis bostrychophila*. *Allergy* 2017;72:918-926. doi: 10.1111/all.13091.
8. Sánchez-Borges M, Capriles-Hulett A, Fernandez-Caldas E. Oral mite anaphylaxis: who, when, and how? *Curr Opin Allergy Clin Immunol*. 2020;20:242-247. doi: 10.1097/ACI.0000000000000624.
9. Clemmons EA, Taylor DK. Booklice (Liposcelis spp.), Grain Mites (Acarus siro), and Flour Beetles (Tribolium spp.): 'Other Pests' Occasionally Found in Laboratory Animal Facilities. *J Am Assoc Lab Anim Sci.* 2016;55:737-743. No DOI available.
10. Collins PJ, Nayak MK, Kopittke R. Residual efficacy of four organophosphate insecticides on concrete and galvanized steel surfaces against three liposcelid psocid species (Psocoptera: Liposcelidae) infesting stored products. *J Econ Entomol*. 2000;93:1357-1363. doi: 10.1603/0022-0493-93.4.1357.
11. Ding W, Zhao ZM, Wang JJ, Tao HY, Zhang, YQ. The relationship between resistance to controlled atmosphere and insecticides of Liposcelis bostrychophila Badonnel (Psocoptera:Liposcelididae). *Scientia Agricultura Sinica*. 2004;37:1308–1315. No DOI available.
12. Khan S, Uddin MN, Rizwan M, Khan S, Uddin MN, Rizwan M, Khan W, Farooq M, Sattar Shah A, Subhan F, Aziz F, Rahman KU, Khan A, Ali S, Muhammad M. Mechanism of Insecticide Resistance in Insects/Pests. *Polish Journal of Environmental Studies*. 2020;29:2023-2030. doi: 10.15244/pjoes/108513.
13. Boyer S, Zhang H, Lempérière G. A review of control methods and resistance mechanisms in stored-product insects. *Bull Entomol Res*. 2012;102:213-229. doi: 10.1017/S0007485311000654.
14. Correy GJ, Zaidman D, Harmelin A, Carvalho S, Mabbitt PD, Calaora V, James PJ, Kotze AC, Jackson CJ, London N. Overcoming insecticide resistance through computational inhibitor design. *Proc Natl Acad Sci USA*. 2019;116:21012-21021. doi: 10.1073/pnas.1909130116.
15. Bradford MM. A rapid and sensitive method for the quantitation of microgram quantities of protein utilizing the principle of protein-dye binding. *Anal Biochem*. 1976;72:248-54. doi: 10.1006/abio.1976.9999.
16. Ashour MB, Gee SJ, Hammock BD. Use of a 96-well microplate reader for measuring routine enzyme activities. *Anal Biochem*. 1987;166:353-360. doi: 10.1016/0003-2697(87)90585-9.
17. Senior AW, Evans R, Jumper J, Kirkpatrick J, Sifre L, Green T, Qin C, Žídek A, Nelson AWR, Bridgland A, Penedones H, Petersen S, Simonyan K, Crossan S, Kohli P, Jones DT, Silver D, Kavukcuoglu K, Hassabis D. Improved protein structure prediction using potentials from deep learning. Nature 2020;577:706–710. doi: 10.1038/s41586-019-1923-7.

341 18. Kochnev Y, Hellemann E, Cassidy KC, Durrant JD. Webina: an open-source library and web app that
342 runs AutoDock Vina entirely in the web browser. Bioinformatics. 2020;36:4513-4515. doi:
343 10.1093/bioinformatics/btaa579.
344 19. Xu Y, Wang S, Hu Q, Gao S, Ma X, Zhang W, Shen Y, Chen F, Lai L, Pei J. CavityPlus: a web server for
345 protein cavity detection with pharmacophore modelling, allosteric site identification and covalent ligand
346 binding ability prediction. Nucleic Acids Res. 2018;46(W1):W374-W379. doi: 10.1093/nar/gky380.
347 20. Hemingway J, Karunaratne SH. Mosquito carboxylesterases: a review of the molecular biology and
348 biochemistry of a major insecticide resistance mechanism. *Med Vet Entomol*. 1998;12:1-12. doi:
349 10.1046/j.1365-2915.1998.00082.x.
350 21. Almagro Armenteros JJ, Tsirigos KD, Sønderby CK, Petersen TN, Winther O, Brunak S, von Heijne G,
351 Nielsen H. SignalP 5.0 improves signal peptide predictions using deep neural networks. *Nat Biotechnol*.
352 2019;37:420-423. doi: 10.1038/s41587-019-0036-z.
353 22. Yoshizawa K, Johnson KP. Phylogenetic position of Phthiraptera (Insecta: Paraneoptera) and elevated rate
354 of evolution in mitochondrial 12S and 16S rDNA. *Mol. Phylogen. Evol*. 2003;29:102-114. doi:
355 10.1016/s1055-7903(03)00073-3.
356 23. Newcomb RD, Campbell PM, Russell RJ, Oakeshott JG. cDNA cloning, baculovirus-expression and
357 kinetic properties of the esterase, E3, involved in organophosphorus resistance in *Lucilia cuprina*. *Pestic*
358 *Biochem Physiol*. 1996;55:85-99. doi: 10.1016/s0965-1748(96)00065-3.
359 24. Jackson CJ, Liu JW, Carr PD, Younus F, Coppin C, Meirelles T, Lethier M, Pandey G, Ollis DL, Russell
360 RJ, Weik M, Oakeshott JG. Structure and function of an insect α-carboxylesterase (αEsterase7) associated
361 with insecticide resistance. *Proc Natl Acad Sci U S A*. 2013;110:10177-82. doi:
362 10.1073/pnas.1304097110.
363
364

365 **Figure legends and table headings**

366

367 **Figure 1** The nucleotide and deduced amino acid sequences of LBCE1 cDNA. Residues forming the
368 catalytic triads conserved in carboxylesterase are indicated by boxes with a solid line. Residues for the
369 oxyanion hole conserved in carboxylesterase are indicated by a box with a dotted line. The GXSXG
370 motif, which is conserved in esterase, is indicated by a shadow. Asterisks indicate residues that differ
371 from those coded by the contig Comp59220.

372

373 **Figure 2** Multiple alignments of the amino acid sequences of LBCE1 and the previously identified
374 esterases of L. bostrychophila. A box with a solid red line indicates the GXSXG motif. A box with a
375 dashed red line indicates a sequence of a typical anion hole. Percent identities to LBCE1 are shown at
376 the bottom line.

377

378 **Figure 3** Recombinant protein expression and purification of 6xHis-tagged LBCE1. (A) SDS-PAGE
379 and western blot analyses of the concentrated culture medium from His-tagged LBCE1-expressing *E.*
380 *coli* cell culture. LBCE1. (B) SDS-PAGE and western blot analyses of the purified His-tagged LBCE1.
381 Arrowheads indicate the bands corresponding to His-tagged LBCE1. M: Size marker. CM:
382 Concentrated culture medium. E: Elute from the Ni-affinity column. (C) pH-dependent hydrolysis of
383 *p*-nitrophenyl acetate by LBCE1.

384

385 **Figure 4** Structure modeling of LBCE1 using Alphafold2. (A) Predicted IDDT values per position for the five
386 predicted models (rank 1-5) of LBCE1. (B) Color-based visualization of predicted IDDT value per position for
387 the rank 3 model. (C) The predicted 3D structure of LBCE1 (D) The crystal structure of LCαE7 (registered as
388 4FNG in PDB). (E) Superpose of the LBCE1 and LCαE7 structures shown in panels C and D, respectively. (F)
389 The magnified view of the catalytic triangle region, which is shown by a yellow box in panel E.

390

391 **Figure 5** Docking simulation for LBCE1- and LCaE7-malathion binding. The superposed views of docking
392 poses for LBCE1-malathion binding (A) and LCaE7-malathion binding (B) (also see Tables S4 and S5). The
393 possible ligand-interacting residues of these enzymes are also indicated. Surface view of LBCE1 (C) and
394 LCaE7 (D). The positions of the surface cavities leading to their catalytic centers are indicated by arrows. The
395 numbering of the two cavities in panel C corresponds to that in Figure S7. (E, F) The magnified views of the
396 cavities of LBCE1 (shown in panel C) and LCaE7 (shown in panel D) together with the superposed docking
397 poses of malathion.

398

399 **Table 1** BLAST search hits identified as LBCE1-homologous protein.

400

Table 1 BLAST search hits identified as LBCE1-homologous proteins

Description	Organism	Query Cover	E value	Identity	Accession
Esterase FE4 precursor, putative	*Pediculus humanus corporis*	96%	1.00E-170	47.19%	XP_002424560.1
esterase E4 isoform X1	*Cryptotermes secundus*	94%	3.00E-153	46.13%	XP_023716379.1
esterase 1	*Liposcelis bostrychophila*	95%	3.00E-152	45.96%	ACI16653.1
esterase E4 isoform X2	*Cryptotermes secundus*	94%	5.00E-153	45.94%	XP_023716388.1
Esterase FE4	*Cryptotermes secundus*	94%	6.00E-154	45.89%	PNF42130.1
Esterase FE4	*Cryptotermes secundus*	94%	2.00E-153	45.89%	PNF42133.1
Esterase FE4	*Cryptotermes secundus*	94%	7.00E-154	45.80%	PNF42132.1
Esterase FE4	*Cryptotermes secundus*	94%	2.00E-153	45.80%	PNF42131.1
uncharacterized protein LOC111869379	*Cryptotermes secundus*	94%	1.00E-146	45.80%	XP_023716645.1
carboxylesterase	*Locusta migratoria*	97%	7.00E-153	45.78%	AHJ81320.1
Esterase FE4	*Cryptotermes secundus*	94%	9.00E-140	45.62%	PNF42134.1
esterase E4-like isoform X2	*Zootermopsis nevadensis*	95%	4.00E-152	45.50%	XP_021912901.1
carboxylesterase	*Dioryctria abietella*	89%	3.00E-135	45.30%	QZC92223.1
juvenile hormone esterase-like isoform X2	*Schistocerca gregaria*	94%	5.00E-147	45.25%	XP_049853109.1
esterase FE4-like isoform X1	*Schistocerca gregaria*	94%	9.00E-147	45.25%	XP_049853108.1
Esterase FE4	*Cryptotermes secundus*	96%	3.00E-149	45.10%	PNF42088.1
carboxylesterase	*Oxya chinensis*	94%	1.00E-147	45.07%	AJP62543.1

```
ATG CAG TTC GGC TCC GAC CTT CGA ACG AGT TAC AAT CGC GAA AAA AAA AAA ATG ACT GAA
 M   Q   F   G   S   D   L   R   T   S   Y   N   R   E   K   K   K   M   T   E    20
AGC CAA CCG ATT ATC CGC ATC GCG GAT GGC TCG ATC CGA GGG GAA AAA TCG GAT TCA ATT
 S   Q   P   I   I   R   I   A   D   G   S   I   R   G   E   K   S*  D   S   I    40
CGC GGT GGT TCT TAT TAC AGC TTT AAG GGG ATC CCT TAT GCC AAA CCT CCT GTT GGG GAT
 R   G   G   S   Y   Y   S   F   K   G   I   P   Y   A   K   P   P   V   G   D    60
TTG AGG TTT AAG GCC CCG GTA CCG GTG GAA CCT TGG ACA GGT GTA AGA GAT GCC TTA AAA
 L   R   F   K   A   P   V   P   V   E   P   W   T   G   V   R   D   A   L   K    80
CAT GGA AGC GAA GCC CCG GCA AAG GAC ATG TTG AAA CAT GAA TAT ATG GAA AAT ACG AGT
 H   G   S   E   A   P   A   K   D   M   L   K   H   E   Y   M   E   N   T   S   100
GAG GAT TGC CTG TTC ATC AAC GTC TAC ACG CCA GAA CTT CCG AAA AGC AAA AAT GAC AAA
 E   D   C   L   F   I   N   V   Y   T   P   E   L   P   K   S   K   N   D   K   120
TTG AAA TCA GTC CTC GTC TGG GTG CAC GGA GGA GGA TTC TCC ATG GGA TCT GGA AAC TCT
 L   K   S   V   L   V   W   V   H   G   G   G   F   S   M   G   S   G   N   S   140
GAA ATC TAC GGC CCC GAC TAC CTC ATC ACG GAA GAC GTC GTC CTG GTC ACT TTC AAC TAT
 E   I   Y   G   P   D   Y   L   I   T   E   D   V   V   L   V   T   F   N   Y   160
CGG TTG GGA GTT TTG GGA TTC CTC AGT CTC GGA ACA GTC GAA TGC CCC GGG AAC TTC GGT
 R   L   G   V   L   G   F   L   S   L   G   T   V   E   C   P   G   N   F   G   180
TTG AAG GAT ATG GTC CTT GCC TTA AAA TGG GTT CAA AAG AAC ATT GCC GCT TTC GGC GGA
 L   K   D   M   V   L   A   L   K   W   V   Q   K   N   I   A   A   F   G   G   200
GAT CCG AAC AAC GTC ACG ATT TTC GGT GAA AGC GCG GGA GGA GCC GCC GTT CAG TAC CTT
 D   P   N   N   V   T   I   F   G   E   S   A   G   G   A   A   V   Q   Y   L   220
TTG ATT TCG AAA GCG ACC AGA GGA TTG TTC CAT AAG GCC ATT TCC CAA TCG GGA ACC ACT
 L   I   S   K   A   T   R   G   L   F   H   K   A   I   S   Q   S   G   T   T   240
TTG GAC CCG TGG GCG CAT AGA CTG AAT CCC AGA GAT TTC GCG TTT GCT TTG GGG GAA GAG
 L   D   P   W   A   H   R   L   N   P   R   D   F   A   F   A   L   G   E   E   260
TTG GGA TGC AAA ACA ACC GAC GAC AAA GTG CTT CTC GAC TTC TTG AAA AAA GCA TCG CAA
 L   G   C   K   T   T   D   D   K   V   L   L   D   F   L   K   K   A   S   Q*  280
AAA GAT TTC GTA GAA AAA GAA GGG GAC TTG CCG AAG AAG CTG TAC CCC GAC AGG ATT TTT
 K   D   F   V   E   K   E   G   D*  L   P   K   K   L   Y   P   D   R   I   F   300
CTC CAG TTA TCG TTT GTT CCC GTA GTC GAA CCC GAA CAC GAA GGG GCC TTT TTA ACC AAA
 L   Q   L   S   F   V   P   V   V   E   P   E   H   E   G   A   F   L   T   K   320
AGC CCA AGG GAA ATT ATT CAA AGC GGG GAT TTC AAT GAT GTC CCG TAT ATT ATC GGA GGA
 S   P   R   E   I   I   Q   S   G   D   F   N   D   V   P   Y   I   I   G   G   340
GTT AGC TTG GAA GGC CTT ATT ATT ATC TAC AGA AAT TTC GAA TAT AAA GAA TCG ACG GCG
 V   S   L   E   G   L   I   I   I   Y   R   N   F   E   Y   K   E   S   T   A   360
GAT GAG GAT TTG GAA CAA GTC CTC CCT CTG GGA ACA TTA AAC ATT CAA AAG GGA TCG AAA
 D   E   D   L   E   Q   V   L   P   L   G   T   L   N   I   Q   K   G   S   K   380
GAA TCC AAG GAA ATA ACG AAG AAA ATT CGG GAC TTT TAC TTC CCC AAC GGA TAT GAG AAG
 E   S   K   E   I   T   K   K   I   R   D   F   Y   F   P   N   G   Y   E   K   400
GAG AAA CTA GTA GCT GTT CTC TCC GCC ATT TAT TTT CTG AAC GGA ATC GGC AAA ACC TGC
 E   K   L   V   A   V   L   S   A   I   Y   F   L   N   G   I   G   K   T   C   420
GAT TGG ATC GGC AGA TTA AAG AAC AGA AAT TCT CCC ACT TAT CTG TAC CAT TTC CTG TTC
 D   W   I   G   R   L   K   N   R   N   S   P   T   Y   L   Y   H   F   L   F   440
GAC GGA ACC AAG GCC TTC CTT AAG CAT CTT ATA GGC TAC GGG GAT TGG AAA GGA ACT TGC
 D   G   T   K   A   F   L   K   H   L   I   G   Y   G   D   W   K   G   T   C   460
CAT GCT GAC GAG CTC GGC TAT CTC TTC CAC ATG CCC ATG CTC CAA GCT AAA CTC GAG CCG
 H   A   D   E   L   G   Y   L   F   H   M   P   M   L   Q   A   K   L   E   P   480
AAC ACC CCT GAA TAT ACG ACA GTT CAA CGC ATG ACC AAA TTA TGG ACC GAT TTT GCG AAA
 N   T   P   E   Y   T   T   V   Q   R   M   T   K   L   W   T   D   F   A   K   500
ACC GGA AAC CCG ACG CCG AAG GAT AAC TCC TGG AAA CCG ATA TCT GAG AAT GAC AAC ACG
 T   G   N   P   T   P   K   D   N   S   W   K   P   I   S   E   N   D   N   T   520
TAT CTG GAA ATC GAA AAA GAA TTA ACT CTC AAG AAG AAT TTC AAC GAG AAA GAG GCG AAA
 Y   L   E   I   E   K   E   L   T   L   K   K   N   F   N   E   K   E   A   K   540
TTG TGG AAT GAA ATT TAC AAA TCC GTT TGC ACA AGA CAC AAG TAA
 L   W   N   E   I   Y   K   S   V   C   T   R   H   K   *                       564
```

Figure 1 The nucleotide and deduced amino acid sequences of LBCE1 cDNA. Residues forming the catalytic triads conserved in carboxylesterase are indicated by boxes with a solid line. Residues for the oxyanion hole conserved in carboxylesterase are indicated by a box with a dotted line. The GXSXG motif, which is conserved in esterase, is indicated by a shadow. Asterisks indicate residues that differ from those coded by the contig Comp59220.

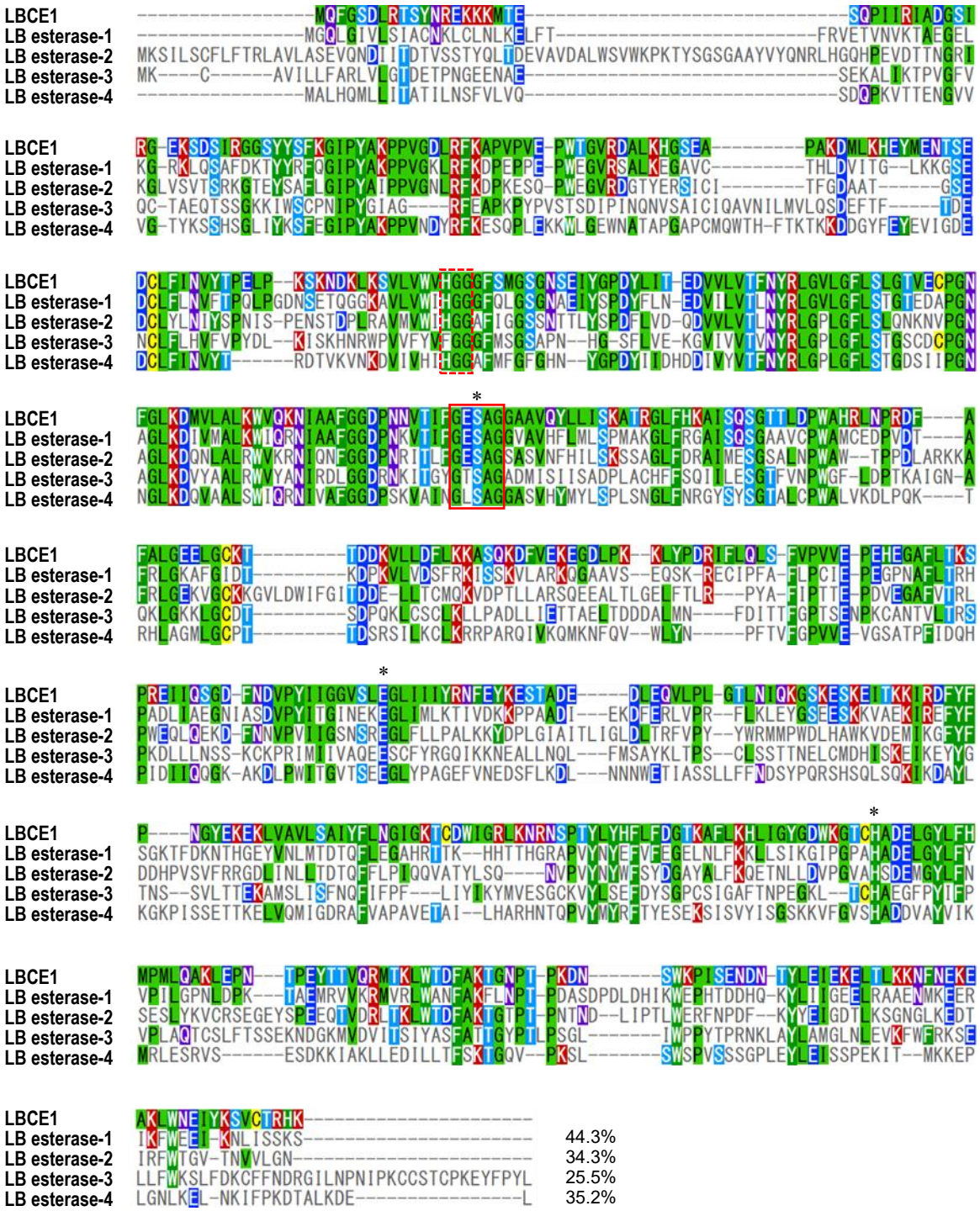

Figure 2 Multiple alignment of the amino acid sequences of LBCE1 and the previously identified esterases of *L. bostrychophila*. A box with a solid red line indicates the GXSXG motif. A box with a dashed red line indicates a sequence of a typical anion holes. Percent identities to LBCE1 are shown at the bottom line.

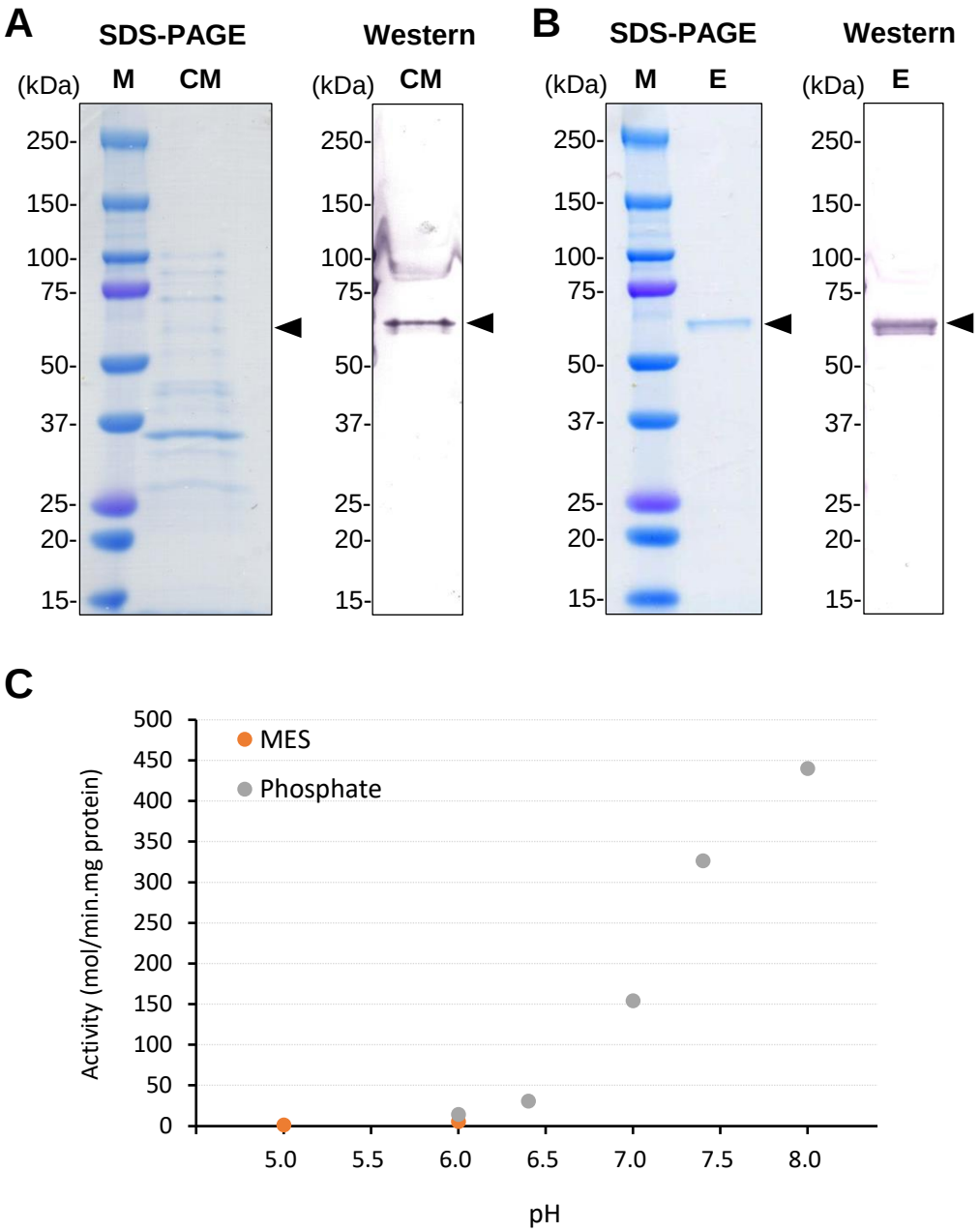

Figure 2 Recombinant protein expression and purification of 6xHis-tagged LBCE1. (A) SDS-PAGE and western blot analyses of the concentrated culture medium from His-tagged LBCE1-expressing *E. coli* cell culture. LBCE1. (B) SDS-PAGE and western blot analyses of the purified His-tagged LBCE1. Arrowheads indicate the bands corresponding to His-tagged LBCE1. M: Size marker. CM: Concentrated culture medium. E: Elute from the Ni-affinity column. (C) pH-dependent hydrolysis of *p*-nitrophenyl acetate by LBCE1.

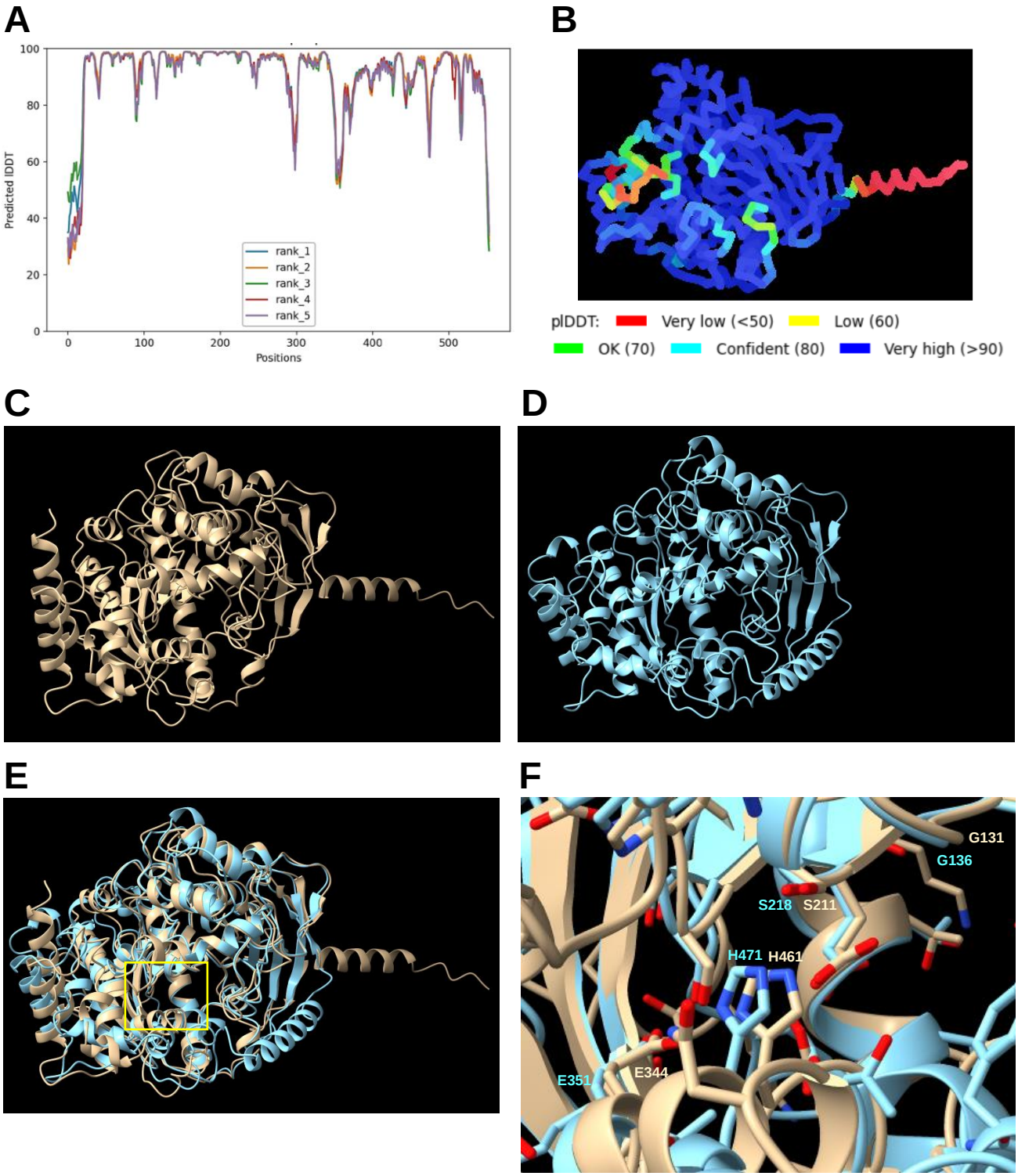

Figure 3 Structure modeling of LBCE1 using Alphafold2. (A) Predicted IDDT values per position for the five predicted models (rank 1-5) of LBCE1. (B) Color-based visualization of predicted IDDT value per position for the rank 3 model. (C) The predicted 3D structure of LBCE1 (D) The crystal structure of LCαE7 (registered as 4FNG in PDB). (E) Superpose of the LBCE1 and LCαE7 structures shown in panels C and D, respectively. (F) The magnified view of the catalytic triangle region, which is shown by a yellow box in panel E.

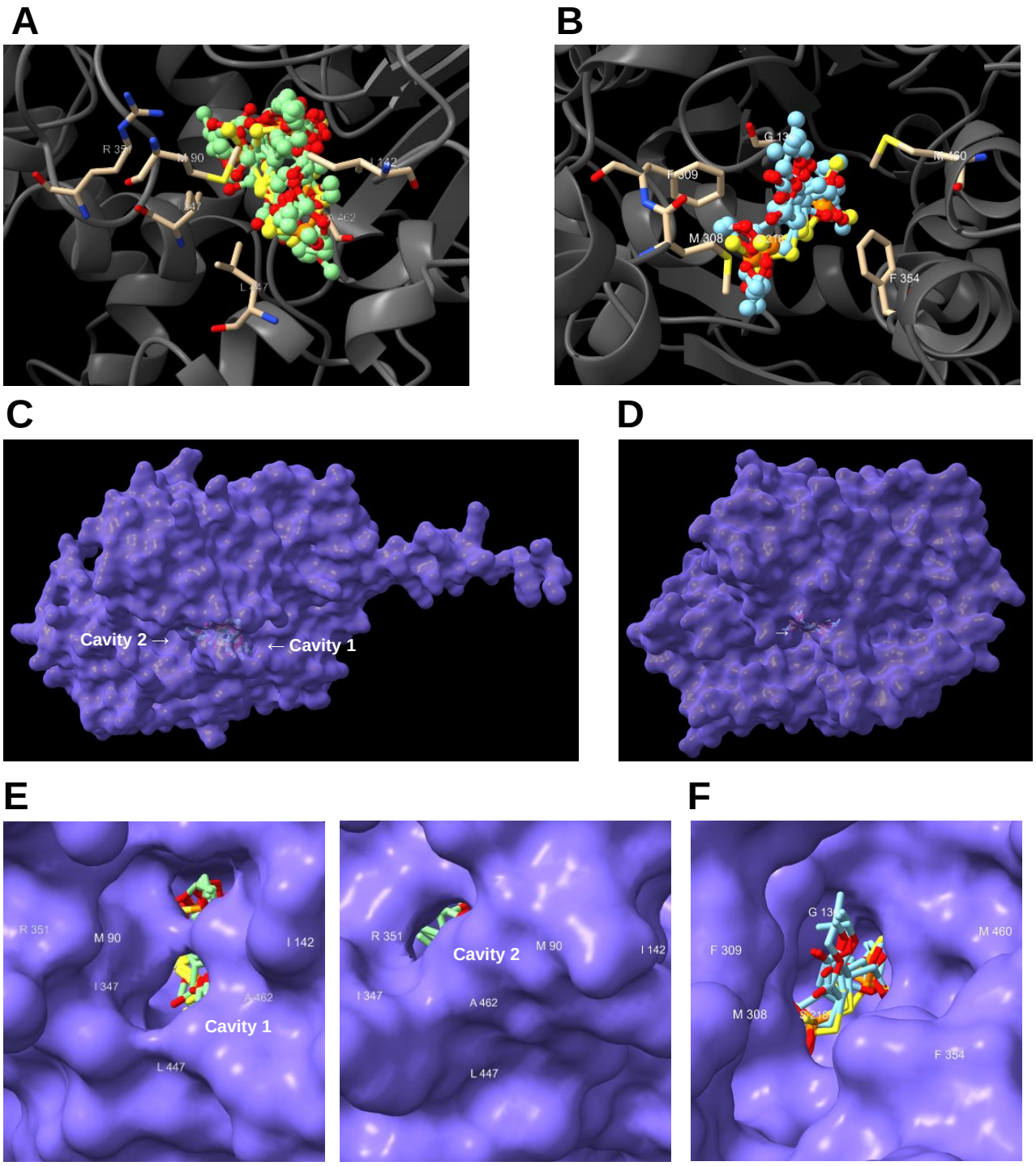

Figure 4 Docking simulation for LBCE1- and LCαE7-malathion binding. The superposed views of docking poses for LBCE1-malathion binding (A) and LCαE7-malathion binding (B) (also see Tables S4 and S5). The possible ligand-interacting residues of these enzymes are also indicated. Surface view of LBCE1 (C) and LCαE7 (D). The positions of the surface cavities leading to their catalytic centers are indicated by arrows. The numbering of the two cavities in panel C corresponds to that in Figure S7. (E, F) The magnified views of the cavities of LBCE1 (shown in panel C) and LCαE7 (shown in panel D) together with the superposed docking poses of malathion.