Structure of Aedes aegypti procarboxypeptidase B1 and its binding with Dengue virus for controlling infection

Edem Gavor1, Yeu Khai Choong1, Nikhil Kumar Tulsian1,2, Digant Nayak1, Fakhriedzwan Idris3,4, Hariharan Sivaraman1, Donald Heng Rong Ting3,4, Alonso Sylvie3,4, Yu Keung Mok1, R Manjunatha Kini1,5, J Sivaraman1

Metallocarboxypeptidases play critical roles in the development of mosquitoes and influence pathogen/parasite infection of the mosquito midgut. Here, we report the crystal structure of Aedes aegypti procarboxypeptidase B1 (PCPBAe1), characterized its substrate specificity and mechanism of binding to and inhibiting Dengue virus (DENV). We show that the activated PCPBAe1 (CPBAe1) hydrolyzes both Arg- and Lys-substrates, which is modulated by residues Asp251 and Ser239. Notably, these residues are conserved in CPBs across mosquito species, possibly required for efficient digestion of basic dietary residues that are necessary for mosquito reproduction and development. Importantly, we characterized the interaction between PCPBAe1 and DENV envelope (E) protein, virus-like particles, and infectious virions. We identified residues Asp18A, Glu19A, Glu85, Arg87, and Arg89 of PCPBAe1 as essential for interaction with DENV.

In this regard, the midgut metallocarboxypeptidases (MCPs) have been considered as transmission-blocking vaccine candidates against mosquito and other insect vectors (4, 7, 8, 9, 10, 11). MCPs are exopeptidases involved in numerous processes, including digestion, blood coagulation/fibrinolysis, inflammation, and pro-hormone and neuropeptide processing (12). In mosquitoes, the A/B subfamily of MCPs (including carboxypeptidases A and B; CPB and CPA, respectively) is primarily responsible for dietary blood digestion (3, 13, 14) for the release of amino acid–rich nutrients required for egg development (4, 5, 16). As such, MCPs are important targets for mosquito population control. MCPs also play a significant role in parasite transmission in mosquitoes. For example, the Anopheles midgut CPB1 has been shown to enhance malaria parasite transmission in various Anopheles mosquitoes (4, 5, 16). Active immunization with recombinant Anopheles gambiae procarboxypeptidase B1 (PCPBAg1) protected mice against Plasmodium berghei challenge (4). In addition, mosquitoes fed with anti-CPBAg1 antibodies exhibited reduced reproductive capacity (5), a secondary effect of a CPB-based transmission-blocking vaccine that could likely suppress mosquito populations. This effect was also observed using inhibitors against mosquito dipeptidases. Similarly, the infection rate of Plasmodium falciparum gametocytes in mosquitoes was significantly reduced after treatment with anti-CPBAg1 antibodies (5, 16). Besides, inhibitors against peptidyl dipeptidase A administered to both An. gambiae and Ae. aegypti mosquito larvae led to stunted growth, reduced fecundity, and mortality (17).

In the midgut of Ae. aegypti mosquito, procarboxypeptidase B1 (CPB1 or PCPBAe1) takes on a different role in pathogen infection, interacting with the DENV-2 envelope (E) protein/virion and impeding viral packaging, maturation, and release from the midgut (6, 7). The interaction between PCPBAe1 and the E protein/virion can also inhibit viral morphogenesis and affect viral glycoprotein processing (7). The overall impact is the release of low titers of mature viruses. Moreover, another procarboxypeptidase A homolog in Ae. aegypti interacts with the DENV capsid protein (18); albeit, the

Introduction

Mosquito-borne viruses such as dengue virus, Zika virus, and chikungunya virus, as well as the malaria parasite constitute a major threat to public health. The mosquito midgut serves as the first entry point for pathogen infection and replication (1, 2). The successful establishment of infection in the midgut cells is a prerequisite for pathogen survival and transmission. Studies have shown that some specific proteins in the midgut promote pathogenicity (3, 4, 5), whereas certain other proteins elicit antiviral/antiparasitic activities (6, 7).
role of this interaction has yet to be established. Furthermore, the Ae. aegypti carboxypeptidase gene promoters drive the expression of genes that hinder parasite development in the guts of transgenic mosquitoes (19).

Despite the crucial roles of MCPs in mosquitoes, the structural and molecular mechanisms associated with these proteases are unclear. Here, we characterized the enzymatic activity of the mature peptidase domain (CPBAe1) and determined the structure of the full-length (FL)-PCPBAe1 (proenzyme) at 2.08-Å resolution. The results provide insights into the substrate specificity of this enzyme and show how this enzyme differs from other non-mosquito insects carboxypeptidases (11). We characterized the interactions between the recombinant PCPBAe1 and DENV-2 envelope (E)-protein, virus-like particles (VLPs), and infectious virions of different serotypes with residues from the substrate-binding pocket of the peptidase domain alone (CPBAe1) and determined the structure of the substrate-binding pocket and displays a core of eight twisted β-sheets surrounded by eight α-helices (Fig 2). The globular N-terminal pro-region occupies the wide active site cleft of the mature peptidase domain; mediated via interactions with residues from the substrate-binding pocket of the peptidase.

### Results

**Sequence analysis of the PCPBAe1 gene**

Tham et al reported a putative procarboxypeptidase B1 (annotated CPB1), from Ae. aegypti mosquito midgut (6, 7). Here, we have adopted the conventional naming system to differentiate the full-length procarboxypeptidase (FL-PCPBAe1) from the mature peptidase domain alone (CPBAe1) with "Ae" denoting the Ae. aegypti's origin of the gene. Besides, the PCPBAe1 amino acid residue numbering used here adopts the conventional numbering system of carboxypeptidases (10, 11, 20). Based on sequence analyses (Fig 1A), PCPBAe1 is expressed as a preproenzyme, with a signal peptide (Met*–Ala*), an N-terminal pro-region (Ala1A–Asp75A, “A” to indicate the activation unit or pro-region), and a C-terminal mature peptidase domain (Asp72–Phe305) (CPBAe1). There is an activation loop between Asp72A and Asp7, with a primary trypsin cleavage site located between the scissile peptide bond, Arg4A–Asp7 (Fig 1A). A unique feature of PCPBAe1 is the missing 310-helix insertion, which is typical to CPBs (20, 21). From our analysis, we observe that this feature may only be unique to mammalian B-type MCPs and not insect B-type MCPs like CPBAe1 and CPBHz (Fig 1B).

**Structure of PCPBAe1 and the pro-region auto-inhibition mechanism**

Crystal structure of PCPBAe1 was determined to 2.08-Å resolution (Fig 2) 7EQX. PCPBAe1 elutes as a monomer in solution as revealed by size-exclusion chromatography and dynamic light scattering (Fig S1A–C). However, the structure consists of two molecules in the asymmetric unit of the crystal. Each molecule consists of 394 amino acid residues (Glu5A–Phe305) (Fig 2). The activation loop (Asp72A–Asp7) that connects the pro-region to the mature peptidase domain is disordered and could not be modeled in our structure. The pro-region adopts an open-sandwich antiparallel-α/antiparallel-β-fold, consisting of two α-helices and four β-strands. The mature peptidase domain contains a zinc-coordinated active site triad and the substrate-binding pocket and displays a core of eight twisted β-sheets surrounded by eight α-helices (Fig 2).

The globular N-terminal pro-region occupies the wide active site cleft of the mature peptidase domain, mediated via interactions with residues from the substrate-binding pocket of the peptidase.
Figure 2. Crystal structure of PCPBAe1 and the autoinhibition mechanism by the pro-region. (A) Overall structure of PCPBAe1 7EQX. There are two molecules in the asymmetric unit of the crystal. Pro-region (magenta) and mature peptidase domain (green) are indicated for each molecule in cartoon and surface representations. Close-up view of the active site pocket (inset) highlights the active site triad His68, Glu71, and Tyr194 (C) and affects the substrate-binding efficiency of the enzyme (Fig 3A and C). Of particular interest, Asp251 at the bottom of the substrate specificity pocket in PCPBAe1 (Fig 2A) is also critical for Arg-substrate hydrolysis (Table 1). Thus, our kinetic studies showed that the substrate specificity is dictated by amino acid residue combinations in positions 251 and 239 (Table 1).

So far, the only insect digestive CPB structure available is that of CPBHz from Culex pipiens (11). CPBHz does not hydrolyze Arg-substrates, and its activity is crucial for blocking small-peptide substrates from contacting the zinc atom, and thus maintain PCPBAe1 in its inactive form. In particular, the double hydrogen bond between Asp251 and Arg120 (Fig 2C) helps maintain PCPBAe1 in its inactive form. The double hydrogen bond between Asp251 and Arg120 is also conserved in both pancreatic and insect CPBs and CPAs, except for human pancreatic CPBh and porcine CPBp, where the residue in the equivalent position is occupied by Ser (Fig S2). The presence of Ser239 in CAE1, which is conserved in Helicoverpa zea CPBHz and other mosquito CPBs, narrows the S1′ specificity pocket and affects the substrate-binding efficiency of the enzyme (Fig 3A and C).

So far, the only insect digestive CPB structure available is that of CPBHz from H. zea (11). CPBHz does not hydrolyze Arg-substrates, and has highly specific interaction towards Lys-substrates; this property was speculated to be dictated by the presence of Glu255 (Asp255 in PCPBAe1) in place of Asp in CPBHz (11) (Fig S2). From our site-directed mutagenesis experiments, Asp255-Glu mutation abrogated Arg-substrate hydrolysis completely in PCPBAe1, whereas Lys-substrate hydrolysis remained unaffected (Table 1). Interestingly, the double mutant Asp255Glu and Ser239Gly restored Arg-substrate hydrolysis by PCPBAe1, which was comparable to WT CPBPAe1. Besides, Asp255 and Ser239Gly further enhance Arg-substrate hydrolysis >30-fold as compared with WT CPBPAe1 (Table 1). To further understand these substrate specificities, we compared Arg-hydrolyzing and non-Arg-hydrolyzing carboxypeptidases (Fig 3A). A detailed analysis of these results suggests that the substrate specificity is dictated by amino acid residue combinations in positions 251 and 239 (Table 1).
of the substrate-binding pocket of CPBAe1 and CPBHz superimposed onto the duck CPD-GEMSA (arginine analogue carboxypeptidase inhibitor) complex (PDB: 1H8L) (27) revealed an apparent difference in the substrate-binding pockets of these enzymes (Fig 3A). Moreover, Ala246, a key residue in the substrate-binding pocket of CPBAe1, is conserved among MCPs except in CPBHz, which has an Ile at this position (Fig S2).

Comparison of PCPBAe1 structure with other procarboxypeptidase structures

A DALI structural homology search identified structural similarities among MCP proenzymes of the A/B subfamily of both insect and mammalian origins (Fig S3A–C and Table S1). Most of the secondary structures aligned well with the PCPBAe1 structure (Fig S3A–C). The nearest structural homolog of PCPBAe1 is the insect proCPA from Helicoverpa armigera (PCPAHa) (Fig S3A). PCPAHa aligns well with PCPBAe1 (PDB: 1JQG, 39% sequence identity, RMSD 1.4 Å for 394 Cα atoms) over the mammalian counterpart (PDB: 1KWM, 36% sequence identity, RMSD 1.6 Å for 394 Cα atoms) (Table S1).

Overall, most of the secondary structures of PCPBAe1 are well-conserved. However, an apparent rotation of ~8–10° is observed in the pro-region, perhaps because of the less-conserved nature of the residues in the pro-region (~20% sequence identity) compared with that of the mature peptidase domain (~40% sequence identity) of the homologs (PDBs: 1JQG, 1KWM, and 1NSA) (Figs S2 and S3A–C and Table S1).

PCPBAe1 directly interacts with DENV-2 infectious virion, VLP, and E protein

Previous work suggested that PCPBAe1 inhibited DENV-2 release from the mosquito midgut cells, by binding to and hijacking the viral E protein/virion, thereby preventing the encapsulation of newly formed nucleocapsids or preventing virion maturation (6, 7). Interestingly, PCPBAe1 overexpression affects mature DENV-2 release from the mosquito C6/36 cell line, but not from mammalian Vero cells (7). Here, we conducted ELISA assays using recombinant PCPBAe1 purified from Escherichia coli (ECP) and insect baculoviral expression systems (ICP) (Fig S1A and C). ECP and ICP were

Table 1. Substrate-hydrolysis of WT and mutant CPBAe1.

| Enzymes/Substrates | SEA/Units/mg | Hip-L-Arg | Hip-L-Lys | Hip-L-Phe | PCPBAe1 |
|---------------------|--------------|-----------|-----------|-----------|-----------|
| Asp Ser             | 17,000.34    | 11,600.41 | 0.00      | 0.00      |
| Glu Ser             | 0.00         | 11,443.22 | 0.00      | 0.00      |
| Asp Gly             | 25,012.10    | 11,216.91 | 0.00      | 0.00      |
| Glu Gly             | 15,971.50    | 11,930.00 | 0.00      | 0.00      |

SEA, specific enzyme activity; 251 and 239 residues positions of WT and mutant CPBAe1.
incubated with DENV-2 E protein obtained from insect and mammalian expression systems (Fig S1F). VLPs purified from mammalian cells, and infectious virions of different DENV serotypes from the mosquito C6/36 cell line (Fig 4A). Different sources of proteins/VLPs/virions were used as a way to possibly uncover the mechanism behind the different effects of PCPBAe1 on DENV-2 release from different cell lines (7).

ECP and ICP interacted with insect- and mammalian cell-derived viral coat antigens/virions (Fig 4A). ICP showed marginally higher interaction with all the viral antigens/virions (~10–13-fold higher) as compared with ECP (Fig 4A). The reason for this higher interaction is unclear. Gel electrophoresis analysis showed that both ICP and ECP have similar molecular weights (Fig S1A and C). Hence, the difference in interaction is unlikely due to any post-translational modifications on ICP. We also noted that PCPBAe1 had a stronger interaction (~25–30-fold higher) with the insect-purified E protein than the mammalian-purified E protein/VLP (Fig 4A). This interaction disparity might be due to the source of the viral antigens, which likely have variations in post-translational modifications on the E glycoproteins. There are two N-glycosylation sites (Asn67 and Asn153) on the DENV E protein (28), and the differences in the compositions of glycans between insect- and mammalian-purified glycoproteins (29) might play a role in the interaction disparities. The interaction with the peptidase domain (CPBAe1) or the pro-region (ProCP) alone (Fig 4B) (purified from E. coli; Fig S1D and E), led to a reduction in binding with DENV (>50-fold lower) suggesting that the full-length-PCPBAe1 is required for interaction. There was no interaction between PCPBAe1 and VLPs of chikungunya (CHIKV) and Mayaro (MAYV) viruses (two alphaviruses that are also transmitted by Aedes mosquitoes) (data not shown).

### Mapping the interaction regions of PCPBAe1 with the E protein using HDXMS

We used HDXMS and high-resolution structures to map the interacting regions of PCPBAe1 and the DENV-2 E protein (Fig 5A–C). We observed that the N-terminal pro-region of PCPBAe1 Ala10A–Asp20A (notably region Tyr18A–Ala23A) showed a single highest deuterium uptake than the peptidase domain Asp71A–Phe295, which exhibited multiple increased deuterium uptake (peptide clusters 171-182, 206-218 and 226-252) upon binding to E protein (Fig 5A and C). Only one locus (the linker between the pro-region and the peptidase domain [Pro23A–Leu62A]) of PCPBAe1 showed significant protection against deuterium exchange in the PCPBAe1E protein complex. Consequently, these identified sites were mutated to alanine or deleted from loop regions in each peptide and showed varying levels of reduction in the interaction as compared with wild-type PCPBAe1 (Fig 6). The extreme N-terminal exposed Asp18A and Glu19A deletion caused the most dramatic reduction in the interaction (~50-fold lower), followed by the charged residues mutations to alanine in the buried linker region ([Pro23A–Leu62A; Glu19A, Arg20A, and Arg22A–AlaAlaAla]) which showed moderate reduction in interaction (~20-fold) (Fig 6A). Deletion of the entire linker loop resulted in inclusion body formation and could not be purified and tested. The extreme C-terminal residues mutations Glu308, Glu170, Arg172–AlaAlaAla (peptide Gly162–Arg172), and Val100–Phe305 deletion did not impact the interaction (Fig 6A).

PCPBAe1 primarily targets the exposed flexible loop regions of the E protein domains I/II; although, the interaction does not include the fusion loop (Asp20A–Gly279) (Fig 5B and C). The peptide regions of E protein domains I/II (Lys35A–Glu41A, Val31A–Val35A, Leu278–Leu283) showed decreased deuterium exchange, suggesting that these regions are buried upon binding to PCPBAe1. Consequently, we used independent domains (domains I/II, and domain III; purified from E. coli) (Fig S1G and H) as truncated versions of the E protein to localize the importance of these regions in interacting with PCPBAe1. The interaction of the individual domains I/II or III alone with PCPBAe1 is significantly less (~70–85-fold lower) than that of the full-length E protein, suggesting that the full-length E protein is necessary for maximum binding (Fig 6B).

### PCPBAe1 binds to all DENV serotypes and 67NTT–67QTV and N153Q mutants

Sequence alignment of all four DENV serotype E proteins (Fig 7A) revealed some residue conservations among the buried peptides.

![Diagram showing interaction between PCPBAe1 and DENV-2 E protein](image)

**Figure 4.** PCPBAe1, CPBAe1, and ProCP interaction with DENV-2 E protein/virus-like particles or infectious virion. (A, B) All the interaction studies (panel A, B) were performed by ELISA in triplicates. (A) Insect cell-purified PCPBAe1 (ICP) and E. coli purified CPBAe1 (CP) interaction with insect cell purified E protein and mammalian HEK purified E protein/virus-like particles. (B) PCPBAe1 (ECP/ICP), mature peptidase domain (CPBAe1) and pro-region (ProCP) interaction with DENV-2 infectious virion. CPBAe1 and ProCP were produced from the E. coli expression system. Data were analyzed by independent unpaired t-test between two sets of data (control and treated samples), and by two-way ANOVA for multiple groups of data using GraphPad Prism version 7.00. Results are shown as mean ± S.D. *P < 0.05, **P < 0.01, ***P < 0.001, and ****P < 0.0001.
Figure 5. Effect of E protein-PCPBAe1 binding on each protein and proposed binding mode.
(A) Plot depicting the differences in deuterium exchange between E protein bound and free PCPBAe1 for various pepsin digested fragments at 1 min (orange) and 10 min (black) labeling timescales. Residue numbers spanning the N- to C termini of PCPBAe1 are indicated, as per the pro-region and peptidase domains. (B) Difference plot of differences in deuterium exchange (y-axis) between PCPBAe1-bound and free E protein for various pepsin digest fragments of E protein (x-axis). Residue numbers of various peptides with domain (I, II, and III) organization is shown. Peptide Ala35-Asp43 spanning domains I showed large-scale increased deuterium exchange in the complex. (A, B) Positive differences indicate increased deuterium exchange and negative values indicate decreased deuterium exchange in E protein bound to PCPBAe1 as compared with (A) PCPBAe1 alone or (B) E protein alone. Blue boxes highlight peptides of E protein or PCPBAe1 showing significant “protection” against deuterium uptake in the peptides in PCPBAe1:E protein complex. Each value is an average of three independent labeling measurements and the standard deviations are in grey. A threshold of ± 0.3 D is considered as significance cut-off and indicated by red-dashed lines. (C) Differences in deuterium exchange at 10 min labeling time are mapped on to the crystal structures of PCPBAe1 (PDB: 7EQX) (solved in this study) (left panel, top) and E protein (left panel, bottom) dimer (PDB: 1OAN), shown in surface representation. Difference maps are colored according to the key with regions showing decreased deuterium exchange in shades of blue and increased deuterium exchange in shades of red. A surface representation model of PCPBAe1-E protein complex (right panel) showing the interaction interface is shown, with the predicted regions of interaction indicated by double-headed arrows.
Here, we determined the properties of the mature peptidase domain (CPBAe1). We identified critical differences between CPBAe1 substrate hydrolysis and that of mammalian and other insects MCPs. Furthermore, through a combination of HDXMS, ELISA, and mutational studies, we described and proposed how PCPBAe1 binds to and suppresses DENV infection.

PCPBAe1 hydrolyzes Arg- and Lys-substrates, and this substrate specificity is dictated by residues Asp$^{251}$ and Ser$^{239}$. From the comparison of substrate-binding pocket, the presence of Gly$^{239}$ compensates for Glu/Gln$^{251}$, whereas Asp$^{251}$ compensates for Ser$^{239}$ to accommodate Arg-substrate hydrolysis. Although Asp and Glu have similar physicochemical properties, their side chains affect Arg- and Lys-substrates binding and hydrolysis differently. The Arg side chain of Arg-substrates is bulkier and does not insert easily into the substrate-binding pocket compared with the less bulky Lys side chain. Thus, Gly$^{251}$/Ser$^{239}$ severely restricts the substrate-binding pocket and does not allow accommodation of the bulky guanidyl group of peptides containing C-terminal Arg. On the contrary, Asp$^{251}$/Gly$^{239}$ allows the side chain of a C-terminal Arg residue of a bound substrate to move away from Asp$^{251}$, which increases Arg-substrate hydrolysis.

Sequence alignment shows that Asp$^{251}$ and Ser$^{239}$ are conserved across mosquito CPBs, and these amino acid residues might be critical for blood digestion and in facilitating egg development (11, 19). This might explain the poor fecundity observed in female mosquitoes when CPB activity was blocked (11, 19). Thus, blocking the enzymatic activity of CPBs in mosquitoes could contribute as a mosquito population control strategy.

Solving the structure of PCPBAe1 allowed us to investigate how the full-length PCPBAe1 binds to the DENV-2 E protein to suppress infection. In the absence of PCPBAe1, ingested DENV undergoes normal replication in the midgut (Fig 8A). During blood ingestion, PCPBAe1 becomes activated (CPBAe1) to digest blood proteins. It is possible that the mature CPBAe1, when released into DENV-infected blood in the midgut lumen, binds to mature virions and partially inhibits binding and entry into midgut cells. We surmise that both the full-length-PCPBAe1 and the mature form (CPBAe1) could potentially bind to the virus inside the mosquito (Fig 8B); albeit the full-length-PCPBAe1 might exhibit a stronger inhibition potency against virus propagation due to its ability to bind more strongly (~50-fold stronger) to the virus. The weaker interaction of the mature peptidase domain alone in our in vitro assay might be due to the disruption of the putative interaction linker region (Tyr$^{46}$–Glu$^{78}$) to produce the active enzyme. The N-terminal peptide region
Figure 7. Sequence alignment of E protein from the four DENV serotypes.

(A) The E protein peptides that are buried upon binding to PCPBAe1 are indicated in blue dash boxes spanning protein domains I/II (a.a 64–84, 238–252, and 278–287), whereas peptide regions that underwent increased deuterium exchange are indicated in black boxes spanning E protein domains I/II (a.a 16–22, 35–43, and 214–231). The Fusion loop is also indicated. The two N-glycosylation sites (N67 and N153) are shown with blue arrows. White fonts with red background show conserved residues, whereas red fonts with white background are areas of high similarities. The alignment was done with Clustal Omega and ESPript 3.0.

(B) Zoomed-in sequence alignment of E protein regions targeted by PCPBAe1. Important amino acid residue changes are indicated with bold font and yellow highlights. Conserved proline residues are indicated with bold red font.
(Tyr4A–Glu21A) of PCPBAe1 seems to be targeted primarily by the E protein, as shown by the significant deuterium exchange at this site, and further corroborated by the alanine site-directed mutagenesis of residues Asp18A and Glu19A. However, using the pro-region alone (ProCP) shows an extreme reduction in interaction (~60-fold reduction), suggesting that the conformational dynamics required for interaction of full-length PCPBAe1 might be different from the individual pro-region. The identified DENV E protein-interacting peptide regions of PCPBAe1 in other Ae. aegypti PCPs (PCPAe1 and PCPBAe111) are poorly conserved, suggesting that PCP homologs may possess different levels of inhibitory activities, if any, against DENV.

Furthermore, the HDXMS results suggest that PCPBAe1 binds across the E protein’s dimeric interface by targeting flexible loop regions but not the conserved fusion loop (Asp98–Lys110). The 4G2 anti-DENV E protein antibody used in our ELISA experiment targets an extreme reduction in interaction (~60-fold), suggesting that the conformational dynamics required for interaction of full-length PCPBAe1 might be different from the individual pro-region. The identified DENV E protein-interacting peptide regions of PCPBAe1 in other Ae. aegypti PCPs (PCPAe1 and PCPBAe111) are poorly conserved, suggesting that PCP homologs may possess different levels of inhibitory activities, if any, against DENV.

The interaction between PCPBAe1 and all DENV serotypes suggests that PCPBAe1 could potentially suppress all serotypes in the mosquito. There are three discontinuous interaction sites on the DENV E protein that bind to PCPBAe1. We systematically compared the sequences of these sites to understand the differential interactions (DENV-3 > DENV-2 > DENV-4 > DENV-1) with DENV serotypes. At first, we compared the sites of DENV-3 and DENV-1, which are at the opposite ends of affinity. Site-3 is completely conserved between DENV-3 and DENV-1, whereas site-2 shows a single residue change (Thr242Asn) (Fig 7B). In contrast, site-1 shows several key differences. Specifically, the two hydrophilic Thr residues to hydrophobic residues in DENV-3 (Thr68Ile and Thr81Val). These hydrophobic residues may enhance the interaction due to the...
tendency to bury themselves in the interaction surface. In addition, a Pro$^{83}$ residue replaces Val$^{83}$ in DENV-3, creating a proline bracket with Pro$^{78}$ residue (32). Such proline brackets are known to enhance protein–protein interactions (33). Thus, it appears that the Site-1 (62–82 residues) mostly defines the affinity differences observed between DENV-3 and DENV-1. There are several amino acid residue differences observed in all three sites between DENV-2 and DENV-4. Interestingly, sites-1 of both DENV-2 and DENV-4 have slightly shorter regions within the proline brackets than DENV-3. They also have a conserved Pro$^{243}$ residue unlike DENV-1 and DENV-3. It is unclear which of these changes are responsible for the observed differences in affinity in these two serotypes.

Only the glycan motif at Asn$^{67}$ on the E protein seems to influence interaction. The different glycan compositions at Asn$^{67}$ between the insect and mammalian cell-purified viral antigens might partially account for the marginal differences observed in the interaction between PCPBAe1 and DENV (29). However, because this difference is only marginal (and not significant), these findings cannot wholly explain why PCPBAe1 only suppresses the release of DENV-2 in mosquito C6/36 cell line but not in a mammalian Vero cell line (7). Thus, apart from PCPBAe1 interaction with the E protein, other conditions in the mosquito cell line microenvironment such as cellular proteins/receptors or attachment factors (34), that are different from the mammalian cell line microenvironment likely play a role in virus inhibition (Fig 8). For example, virus internalization might occur through distinct entry pathways, including clathrin-mediated or non-classical clathrin-independent endocytosis, depending on whether mosquito or human cell or specific host cells (34). Furthermore, C6/36 is a better cell line for DENV infection than Vero cell line (34).

In summary, we show that the residues Asp$^{251}$ and Ser$^{299}$ control substrate-specificity, and such residues may be targeted for developing anti-CBP molecules to arrest digestion and control mosquito reproduction. Because of its ability to bind all DENV serotypes, PCPBAe1 may also inhibit other DENV serotypes by targeting the E protein/virion, which could hamper viral packaging or morphogenesis, leading to an overall reduced viral maturation and release from the mosquito midgut. Alternatively, activated CPBPAe1 may also inhibit mature DENV when released into DENV-infected blood in the midgut lumen, although this inhibition may be less than that of the full-length PCPBAe1. Besides, because the mature region usually degrades the pro-region of procarboxypeptidases upon tryptic cleavage, the independent pro-region may not be available in the midgut when cleaved from the proenzyme (35). However, PCPBAe1 might be specific to DENV as we did not observe PCPBAe1 binding to CHIKV or MAYV VLPs.

Thus, our structural and functional studies of PCPBAe1 provide a blueprint for developing two distinct strategies: (i) carboxypeptidase inhibitors or antibodies as a strategy for mosquito population control and (ii) how PCPBAe1 elicits antiviral properties against DENV. However, it would be essential to evaluate the strategies aimed at blocking carboxypeptidase activities in DENV endemic areas to avoid compromising CPB antiviral activities in the mosquito. Moreover, even though PCPBAe1 serves as an anti-dengue viral molecule in the mosquito, it may not represent an effective dengue or other mosquito-borne viral diseases control strategy as the inhibition of viral propagation seems to only occur in the vector and may be specific to DENV.

**Materials and Methods**

**Cloning, recombinant protein expression and purification**

**PCPBAe1 and E protein domain I/II and III expression and purification from the E. coli system**

The gene coding for proenzyme *Aedes aegypti* procarboxypeptidase B1 (PCPBAe1) (Genbank accession number: Q6J661) was codon optimized and synthesized by Genscript for recombinant protein expression in E. coli (BL-21(DE3)) (shuffle) cells. The gene was subcloned into pGEX-6p-1 vector (Novagen) using the BamHI/XhoI restriction sites to form the pGEX-precision protease site-PCPBAe1 fusion construct. Confirmed fusion plasmid was transformed into the E. coli shuffle (BL-21(DE3)) cells. Transformed cells were inoculated into 10 ml LB medium supplemented with 100 μg/ml ampicillin and grown at 37°C overnight as a primary culture inoculum. The overnight culture was used to inoculate 1L flask secondary LB medium and grown until the optical density (OD) reached 0.8. The culture was cooled at room temperature and induced with 0.3 mM IPTG, supplemented with 60 μM zinc sulphate. The induced culture was shaken at 18°C, 150 rpm overnight. Overnight bacteria culture was pelleted down at 4,000 rpm (centrifuge rotor ID: JLA-8.1000) for 30 min. Bacteria cell pellets were resuspended in 35 ml/l pellet in lysis buffer containing 50 mM citrate buffer, pH 6.0, 0.3 M NaCl, 5% glycerol 1 mM DTT, and a protease cocktail inhibitor tablet (Roche). Pellets were vortexed and sonicated for three cycles (5 min, 1 s on, 2 s off). Sample was then spun down at 18,000 rpm (centrifuge rotor ID: JA-20) for 30 min at 4°C and bound to GST beads (GE) previously equilibrated with the same lysis buffer at 4°C for 4 h for affinity purification. Beads were washed three times with 90 ml of lysis buffer each time and incubated overnight with elution buffer supplemented with PreScission Protease (GE) to cleave the GST tag. First step purification of anion exchange chromatography was done on QHP column (GE), applying a gradient from 100% buffer A (20 mM Tris, pH 8.0, 1 mM DTT) to 100% buffer B (20 mM Tris/NaCl 1 M, pH 8.0, 1 mM DTT) in 30 min. A final size exclusion chromatography on a Superdex S-200 column (GE) (20 mM sodium citrate pH 6.0, 0.1 M NaCl, 5% glycerol, and 1 mM DTT) was performed to achieve maximal purity. The peak fractions were collected and concentrated to 10 mg/ml. All the other PCPBAe1 mutants were purified using the E. coli expression and purified as above.

DENV-2 E protein domains I/II (aa 1–296) and domain III (aa 296–394) were cloned separately into a modified pet23b vector carrying sumo-tag and modified pet32a vector carrying only 6X-Histidine tag. Protein expression conditions were similar to the above described. Purification was carried out with nickel beads affinity chromatography, followed by ion exchange chromatography (buffer similar to above described) and gel filtration in the buffer condition (Tris, pH 8.0, 150 mM NaCl, 5% glycerol, and 1 mM DTT). The tags were not cleaved before further assays.

**PCPBAe1 expression and purification from the insect baculoviral system**

The PCPBAe1 construct was further cloned into pFastBac1 vector with 6X-Histidine purification tag for protein expression in insect-cell baculovirus expression system using *Spodoptera frugiperda*
(SF9) cells by following previous protocols (36, 37). Protein was secreted into the culture medium and purified using nickel-NTA beads and gel filtration chromatography. Dynamic light scattering studies were carried out on a DynaPro Light Scattering instrument (Protein Solutions) to assess protein quality.

**Enzymatic assay of CPBAe1**

The proenzyme PCPBAe1 was activated using trypsin limited proteolysis into the active form CPBAe1 and further purified using gel filtration and ion exchange chromatography. CPBAe1 activity was determined by the continuous spectrophotometric rate determination method of Folk et al (1960) where the reaction velocity is determined by an increase in absorbance at 254 nm resulting from the hydrolysis of Hippuryl-L-Arginine (Hip-L-Arg), Hip-L-Lys, and Hip-L-Phe. One unit causes the hydrolysis of 1 μmol of Hippuryl-L-arginine per minute at 25°C and pH 7.65 under the specified conditions. Briefly, substrates at different concentrations (0.05–10 mM) were diluted in buffer (0.025 M Tris–HCl, pH 7.65, containing 0.1 M sodium chloride). Different concentrations of stock trypsin activated CPBAe1 solution were diluted in assay buffer (0.025 M Tris–HCl buffer, pH 7.65, containing 0.1 M NaCl). The spectrophotometer was set at 254 nm and 25°C. Next, 2.9 ml of substrate solution was pipetted into two separate quartz cuvettes and incubated in spectrophotometer at 25°C for 3–4 min to reach temperature equilibration and blank rate was established. To start activity measurement, 0.1 ml of diluted enzyme was added to the substrate in the cuvette and the increase in absorbance (A254) was recorded for 3–4 min. The fastest linear rate (ΔA254nm/minute) over a 1-min interval for the test and the blank reactions was then determined from the initial linear portion of the curve and used to calculate the specific enzyme activity values. All the OD readings were done in triplicates. The Km was also calculated using a kinetics plot.

**Crystallization and structure determination of proenzyme PCPBAe1**

Crystallization screens were conducted using the sitting drop vapor diffusion method at room temperature (22°C) by mixing protein with the crystallization reservoir solutions at a 1:1 volume ratio. Crystals appeared after about 2 wk in the condition Crystal Screen 1 (0.2 M ammonium sulphate, 0.1 M sodium codbadate trihydrate, pH 6.5, 30% wt/vol PEG 8000) (Hampton Research). The diffraction quality crystals were immersed in mother liquor solution supplemented with 25% ethylene glycol as a cryoprotectant for data collection. A complete native 2.0 Å resolution diffraction data set was collected at 100K using Rigaku MicroMax-007 HF equipped with Saturn 944+ CCD detector, National University of Singapore. Data were processed with the HKL-2000 program (38). The structure was determined using molecular replacement (39) using the coordinates of CPAHa (PDB: 1JQG) (10) and the final model was refined at 2.08 Å resolution. The structure has good stereochemical parameters evaluated with PROCHECK (Table S2). PyMOL and CHIMERA (40) were used to prepare all structure-related figures.

**ELISA assay to characterize PCPBAe1-DENV virion, E protein and VLP interaction**

High binding 96-well plates (Thermo Fisher Scientific) were coated overnight at 4°C with a 5 μg of all PCPBAe1 constructs and mutants or 5 μg BSA or sumo tag protein alone in coating buffer. All samples were coated in triplicates. The next day, plates were washed twice with PBS and blocked with blocking buffer (1% BSA in 1XPBST) for 1 h at 28°C. Plates were incubated with 100 μl of DENV-2 VLPs/ infectious virion serotypes (DENV-1, DENV-2, DENV-3, and DENV-4) (2 × 10⁵ p.f.u) or recombinant DENV-2 E protein (1 μg) (full-length, sumo-tagged domain I/II, or domain III) at 28°C for 1 h, considering the temperature of the mosquito vector. DENV-2 E protein (DENV-2-ENV-100) and VLPs (DENV-2-VLP-100, CHIKV-VLP-100, and MAYV-VLP: REC31616-100) were purchased from The Native Antigens Company (purified from HEK293 cell cultures), whereas infectious virion was obtained from our collaborators (A/P Sylvie Alonso, National University of Singapore). Unbound proteins/virion were washed three times with buffer (1× PBST+ 0.01% Tween 20) and incubated with 100 μl of 1 μg/ml mouse anti-DENV-2 primary antibody (MAB10216, Millipore or anti-E protein domain III [2073], absolute antibody) for 2 h at RT to detect VLP/E protein/virion. Unbound primary antibody was then washed three times with buffer (1× PBST+ 0.01% Tween 20) and plates incubated with 100 μl of 1.5,000 dilution rabbit anti-mouse HRP conjugated secondary antibody and washed six times after 2 h of incubation at RT. Reaction was visualized after incubating with 80 μl of tetramethylbenzidine (TMB). After 5 min, the reaction was stopped with 1 M HCl and plates were read at 450 nm in Tecan Infinite-200 pro (Life Sciences and Diagnostics).

**Hydrogen-deuterium exchange (HDXMS) of PCPBAe1-E protein interaction**

Interaction of PCPBAe1 and DENV-2 E protein were mapped by amide hydrogen-deuterium exchange mass spectrometry (HDXMS) (41). HDX reaction was initiated by diluting the protein in 90 μl deuterated buffer (50 mM PBS pH 7.4) to a final 90% D₂O (Cambridge Isotopes) concentration. For HDXMS of free proteins, 3 μl of 1 mg/ml of DENV-2 E protein and 5 μl of 1 mg/ml of PCPBAe1 were used. For HDXMS of complex, the proteins were mixed in 1:2 E protein:PCPBAe1 stoichiometric ratio and incubated for 20 min before deuteration labeling reaction. Deuterium labeling was carried out for 1- and 10-min time points at 28°C temperature, similar to mosquito vector. The exchange reaction was stopped by lowering the pH to ~2.6 by addition of 100 μl quench solution (0.8 M guanidinium hydrochloride, 0.25 M TCEP) and temperature to 0°C to minimize back-exchange. Non-deuterated control experiments of PCPBAe1 and E protein alone were also carried out by diluting the samples in aqueous buffer, followed by quench solution.

The quenched samples were incubated on ice (0°C) for 30 s followed by 3 min proteolytic digestion by immobilized pepsin and the digested pepsin-cleaved fragments (“peptides”) were resolved by reverse-phase chromatography using C18 column. The peptides were eluted using a 10 min 8–40% gradient of 0.1% formic acid in acetonitrile, pumped at 40 μl/min by nanoACQUITY binary solvent manager (42). The peptides eluted were then subjected to mass
For the activity assays, three mutants were generated: Asp251Glu, Site-directed mutagenesis from N- to C terminus. to another condition is represented as a difference plot listed from the deuterium exchange reactions and are not corrected for back-exchange. All values reported are an average of three deuteration exchange reactions and are not corrected for back-exchange. The absolute difference in deuteron exchange by all peptides relative to another condition is represented as a difference plot listed from the N– to C terminus.

Site-directed mutagenesis

For the activity assays, three mutants were generated: Asp251Glu, Ser239Gly, and Asp251Gly+Ser239Gly. Based on our HDXMS results, four PCPBAe1 mutants were generated including Asp18A and Glu19A deletion, linker region (Glu85Ala Arg87AlaArg89Ala) and (Glu168Ala, Glu170Ala, and Arg172Ala) alanine mutagenesis, and Val300-Phe305 deletion. All the primer/oligonucleotide information are provided in Table S3. All the PCR reactions were carried out using KAPA-BIOSYSTEMS HiFi PCR reaction Kits (Roche).

Generation of de-glycosylated DENV-2 mutant

D2Y98P-PP1 (GenBank: JF327392.1) RNA genome was extracted using QIAamp Viral RNA Kits (QIAGEN). cDNA synthesis was performed using GoScript Reverse Transcriptase (Promega) as per manufacturer’s instructions. Four PCR fragments of around 2,700 nucleotides long were generated from cDNA using primer pairs with NEB Q5 Hot-Start high-fidelity 2× Master Mix (New England Biolabs). Fragments were gel-purified with MinElute gel extraction Kit (QIAGEN) after agarose gel electrophoresis and blunt end cloning was performed into pCR-Blunt II-TOPO Vector (Invitrogen). Site-directed mutagenesis was performed on plasmid that carried E gene to introduce the desired mutation using primer pairs. The mutated gene was amplified with other genes fragments via PCR and seamlessly assembled with vector contains CMV promoter sequence, hepatitis delta virus ribozyme, and simian virus 40 (SV40) poly-A sequence using NEBuilder HiFi DNA Assembly Master Mix (New England Biolabs) at 50°C for 60 min. This CMV-DENV genome-HDVr-SV40pA assembled product was transfected into BHK-21 cells using Lipofectamine 2000 (Invitrogen). Site-directed mutagenesis was performed at 5°C for 60 min. This CMV-DENV genome-HDVr-SV40pA assembled product was transfected into BHK-21 cells using Lipofectamine 2000 (Invitrogen). Four to 6 d later, the viral supernatant was collected and the viral supernatant was collected and sequenced for validation using Sanger sequencing. All the primer/oligonucleotide information are provided in Table S4.

Data Deposition

Three dimensional atomic coordinates, and structure factors of PCPBAe1 have been deposited in the Protein Data Bank, www.pdb.org (PDB ID codes 7EQX).

Supplementary Information

Supplementary information is available at https://doi.org/10.26508/lsa.202101211.

Acknowledgements

This work was supported by Ministry of Education, Singapore (MoE Tier-3) grant (R154–000–697–112) and R154–000–C07–114 (AcRF Tier 1 grant), respectively. E Gavor is a graduate scholar in receipt of the Singapore International Graduate Award (SINGA) research scholarship.

Author Contributions

E Gavor: conceptualization, data curation, software, formal analysis, validation, visualization, methodology, and writing—original draft, review, and editing.
YK Choong: data curation, software, formal analysis, validation, visualization, methodology, and writing—review and editing.
NK Tulsian: data curation, software, formal analysis, validation, visualization, methodology, and writing—review and editing.
D Nayak: software, formal analysis, validation, and visualization.
F Idris: formal analysis, validation, methodology, and visualization.
H Sivaraman: methodology.
DHR Ting: methodology.
A Sylvie: resources, formal analysis, and writing—review and editing.
YK Mok: writing—review and editing.
RM Kini: formal analysis, validation, visualization, and writing—review and editing.
J Sivaraman: conceptualization, resources, software, formal analysis, supervision, funding acquisition, validation, project administration, and writing—original draft, review, and editing.

Conflict of Interest Statement

The authors declare that they have no conflict of interest.

References

1. Franz AW, Kantor AM, Passarelli AL, Clem RJ (2015) Tissue barriers to arbovirus infection in mosquitoes. Viruses 7: 3741–3767. doi:10.3390/v7072795
2. Kato N, Mueller CR, Fuchs JF, McElroy K, Wessely V, Higgs S, Christensen BM (2008) Evaluation of the function of a type I peritrophic matrix as a physical barrier for midgut epithelium Invasion by mosquito-borne pathogens in Aedes aegypti. Vector Borne Zoonotic Dis 8: 701–712. doi:10.1089/vbz.2007.0270
3. Edwards MJ, Lemos FJ, Donnelly-Doman M, Jacobs-Lorena M (1997) Rapid induction by a blood meal of a carboxypeptidase gene in the gut of the mosquito Anopheles gambiae. Insect Biochem Mol Biol 27: 1063–1072. doi:10.1016/S0965-1748(97)00093-3
4. Lavazec C, Boudin C, Lacroix R, Bonnet S, Diop A, Thiberge S, Boisson B, Tahar R, Bourgoin C (2007) Carboxypeptidases B of Anopheles gambiae as targets for a Plasmodium falciparum transmission-blocking vaccine. Infect Immun 75: 1635–1642. doi:10.1128/IAI.00864-06
PCPBAe1 interaction with Dengue virus  Gavor et al.
37. Scholz J, Suppmann S (2017) A new single-step protocol for rapid baculovirus-driven protein production in insect cells. BMC Biotechnol 17: 83–89. doi:10.1186/s12896-017-0400-3
38. Powell HR (2017) X-ray data processing. Biosci Rep 37: 1–14. doi:10.1042/BSR20170227
39. McCoy AJ (2006) Solving structures of protein complexes by molecular replacement with Phaser. Acta Crystallogr D Biol Crystallogr 63: 32–41. doi:10.1107/S0907444906045975
40. Mooers BH (2016) Simplifying and enhancing the use of PyMOL with horizontal scripts. Protein Sci 25: 1873–1882. doi:10.1002/pro.2996
41. Raghuvamsi PV, Tulsian NK, Samsudin F, Qian X, Purushotorman K, Yue G, Kozma MM, Hwa WY, Lescar J, Bond PJ, et al (2021) Sars-cov-2 s protein: Ace2 interaction reveals novel allosteric targets. Elife 10: 1–47. doi:10.7554/elife.63646
42. Wales1 TE, Fadgen KE, Gerhardt GC, Engen JR (2011) High-speed and high-resolution UPLC separation at zero degrees celsius. Bone 23: 1–7. doi:10.1021/ac8008862.High-speed
43. Li GZ, Vissers JP, Silva JC, Golick D, Gorenstein MV, Geromanos SJ (2009) Database searching and accounting of multiplexed precursor and product ion spectra from the data independent analysis of simple and complex peptide mixtures. Proteomics 9: 1696–1719. doi:10.1002/pmnic.200800564

License: This article is available under a Creative Commons License (Attribution 4.0 International, as described at https://creativecommons.org/licenses/by/4.0/).