International Journal of Mosquito Research

Author's details are given below the reference section

Proteomic analysis of AeD7L1 and AeD7L2 in *Aedes aegypti* to the biogenic amines in host hemostasis

Darren Nichole Ocampo, Hans Ezekiel Olorosisimo, Erica Julienne Palting, Lyka Angel Pascual, Edrole Siegfred Ramos, Julienne Aubrey Ruela, Sebastiane Salonga, Andrea Bianca Francesca Sanchez and Mary Rose Lirio

Abstract
AeD7L1 and AeD7L2 are components of the salivary proteins of *Aedes aegypti*, contributing to the successful blood-feeding of the arthropod. This event is caused by inhibiting the hemostatic response of the host by acting as scavengers for biogenic amines such as serotonin, inducing vasodilation and increased vascular permeability that leads to excessive bleeding. In silico methods have been utilized to establish the similarities and distinct differences of these proteins, as well as the determination of their structures in correlation with their binding mechanisms through the use of homology modeling. AeD7L1 has a higher antigenicity score, but AeD7L2 gained the ability to bind to a thromboxane analog due to a difference in the amino acid composition. Their structures also showed areas responsible for their binding mechanisms and protein interactions. Recognizing these functional properties may aid in drug developments, protein modifications, and others that may alleviate the burden of arthropod-borne viruses.

Keywords: *Aedes aegypti,* D7 proteins, hemostasis, *in silico*, microbiology, salivary proteins

1. Introduction

The Dengue Virus (DENV) has been endemic to the Philippines since 1966 and continues to infect the country in the present at a steady rate [1]. The primary vector for arboviruses such as DENV is a hematophagous infected female *Aedes aegypti* which employs a complex saliva composition during blood feeding, with D7 proteins emerging as crucial components [2]. The abundance and defensive mechanisms of D7 proteins were the significant effects of the habitual and environmental adaptability of the vector. Its continual evolution gave way to a diverse array of salivary proteins that progressed to have anti-hemostatic, anti-inflammatory, and binding characteristics that inhibit the immune response of their targeted host—making them almost invulnerable to their immediate opponents and have an increased chance of successful hematophagy [3]. The intricate interplay between these salivary proteins and the host's biogenic amines is central to the mosquito's successful blood meal acquisition.

Consequently, this blood-feeding mechanism critically affects the normal process of hemostasis of the host as the D7 protein family scavenges for biogenic amines such as serotonin, histamine, and epinephrine. Hemostasis is the primary response of the host against breakage in the capillary or blood vessels that reduces the risk of blood loss, and when binding occurs among the D7 proteins and these biogenic amines, phases in the hemostatic process such as vasoconstriction, platelet aggregation, and blood coagulation are inhibited [4].

These mechanisms are inherently evident in the manifestation of symptoms in DENV. Although the virus can span from a self-limiting disease such as Dengue Fever (DF), it can progress into its more serious forms such as Dengue Hemorrhagic Fever (DHF) and Dengue Shock Syndrome (DSS) [5]. Significant hallmarks of these diseases are severe bleeding associated with impaired coagulation process in the body or coagulopathy, organ-associated

Corresponding Author:
Darren Nichole Ocampo
Far Eastern University, Institute of Health Sciences and Nursing, Department of Medical Technology, Sampaloc, Manila, 1015, Philippines

illnesses or failures, thrombocytopenia, petechiae, unusual plasma leakage leading to hypovolemic shock [6], and sepsis due to increased vascular permeability [7].

Despite its high prevalence rate in places with warm climates such as Asia and its implication for the well-being of people of all ages - including its severe viral symptoms and fatal complications regarding the hematological as well as organ processes, there is still a scarcity of effective medications and antiviral drugs against it [8]. Although there are a few antiviral compounds identified for the alleviation of the virus, many of them are still in the early process of drug development [9]. Today in the local setting in the Philippines, the Dengue Virus is being alleviated by controlling the spread of the vectors which is seen in the Department of Health's National Dengue Prevention and Control Program consisting of surveillance, case management in hospitals, outbreak response, health promotion, and further research about the virus which is condensed in its S4 strategy [10]. Though the country has already permitted the administration of 'Dengvaxia' in the past as a vaccine for people who already encountered the virus, tons of controversies have sparked up around it, causing people to be uneasy and distrust the vaccine, consequently plummeting the confidence level of people in all types of vaccines [11].

Due to the imbalance between its viral implications and the number of medications available, this study aims to bridge the gap between the two factors. Characterization of the proteins affecting the host and the mechanisms circulating around the process at a molecular level may aid in the identification of possible characteristics of therapeutic drugs for the virus. This goal can be feasible by utilizing in silico methods that make use of online resources and software that analyzes the physicochemical properties as well as the protein structures of the examined protein using their foundational amino acid sequence. This method makes it an advantageous component in the early stage of the drug development process [12].

2. Materials and Methods
2.1 Retrieval of Protein Sequences and Proteomics
Amino acid sequences for AeD7L1, AeD7L2, and serotonin were retrieved from the database of National Centre for Biotechnology Information (NCBI) using BLAST and UniProt. These sequences are then subjected to ClustalOmega for alignment, showing parallels in the results. The AAEL006424 (ProtParam: Q0IF93_AEDAE) for *Aedes aegypti* D7L1 and AAEL006417 (ProtParam: A0A1S4FDI7_AEDAE) for D7L2 with their C-terminal and N-terminal, and AGB76027.1 for the partial sequence of serotonin were selected sequences to be utilized for this study.

2.2 Determination of Physicochemical Properties
The molecular characterization of the D7 proteins and the biogenic amine serotonin was determined by submitting the identified protein sequences to ProtParam in the Swiss Bioinformatics Resource Portal, mining for the functional characteristics of the proteins through the determination of its physicochemical characteristics. Moreover, the antigenicity of AeD7L1 and AeD7L2 were established by using the Predicted Antigenic Peptides that showed antigenic determinants based on their amino acid sequence and average antigenic scores. These sequences also went through SOPMA (Self Optimized Prediction Method from Alignment) and PredictProtein for the determination of the secondary structures contributing to the protein mechanisms and interactions. The alignment of the sequences using ClustalOmega also showed significance in the process. Predictions from these software were utilized to analyze in accordance with the proteins' characteristics contributing to its mechanisms in host hemostasis.
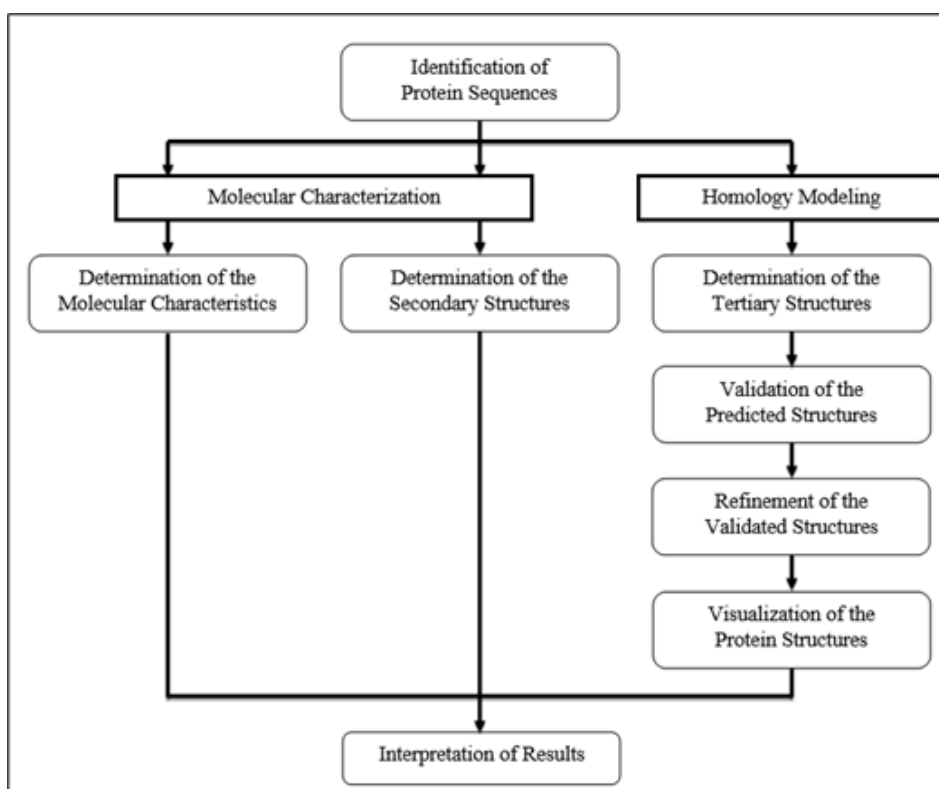

Fig 1: The process and flow of the study. The protein sequences of female *Aedes aegypti* D7L1 (AAEL006424) and D7L2 (AAEL006417) will be the elected sequences in the study

2.3 Modeling, Validation, and Refinement of Structures

In homology modeling, the structures went through a complex process for the reliability of predicted structures. First, the amino acid sequences retrieved from the database are uploaded to tertiary structure predicting programs *id est* SWISS-MODEL, Phyre 2, and i-TASSER. Models created are then screened using the validation tool Ramachandran plots via the SAVES server. Subsequently, passing model structures are further refined in the GalaxyWEB. Lastly, refined structures are visualized using the Mol* viewer. The product models of structures are now used to interpret results along with the data gathered from molecular characterization.

3. Results and Discussion

3.1 Analysis of the Sequences of AeD7L1, AeD7L1, and Serotonin:
Based on the results shown in Table 1, AeD7L1

has a total of 321 amino acids with a molecular weight of 36, 895.10 kDa, while AeD7L2 is composed of 332 amino acids having a weight of 38,627.77 kDa. Serotonin has a lower number of amino acids with 118 in total, weighing 12,998.72 kDa.

The molecular weight of the two salivary proteins in *Aedes aegypti* correlates with their ability to cause an immune response from the host. As stated by Miller and Stevens, the body can detect an immunogen with a molecular weight of 10,000 Da [13]. One of the most important factors in an immunogen to elicit an immune response is its composition, wherein proteins are the most immunogenic. With the molecular weight and nature of both AeD7L1 and AeD7L2, they are the perfect vessels to elicit an immune response in their chosen host.

Table 1: Physicochemical properties of AeD7L1, AeD7L1, and Serotonin

Properties	AeD7L1	AeD7L2	Serotonin
Number of amino acids	321	332	118
Molecular weight	36895.10 kDa	38627.77 kDa	12998.72 kDa
Formula	$C_{1646}H_{2563}N_{435}O_{494}S_{17}$	$C_{1736}H_{2717}N_{467}O_{489}S_{21}$	$C_{562}H_{904}N_{148}O_{188}S_{8}$
Total number of atoms	5155	5430	1810
Theoretical pI	7.91	9.32	4.35
Extinction coefficients	39475	48985	14230
Estimated Half-life	30 hours (mammalian reticulocytes, *in vitro*) >20 hours (yeast, *in vivo*) >10 hours (*Escherichia coli*, *in vivo*)	30 hours (mammalian reticulocytes, *in vitro*) >20 hours (yeast, *in vivo*) >10 hours (*Escherichia coli*, *in vivo*)	30 hours (mammalian reticulocytes, *in vitro*) >20 hours (yeast, *in vivo*) >10 hours (*Escherichia coli*, *in vivo*)
Instability index	40.45 (unstable)	33.04 (stable)	38.07 (stable)
Aliphatic index	67.73	65.81	101.69
N-terminal	M (Met)	M (Met)	M (Met)
Grand Average of Hydropathicity (GRAVY)	-0.659 (hydrophilic)	-0.649 (hydrophilic)	0.43 (hydrophobic)
Secondary Structures Alpha Helix	53.52%	56.33%	30.51%
Extended Strands	9.03%	8.73%	30.51%
Random coils	32.09%	32.23%	28.81%

3.1.1 Aliphatic Index
The aliphatic index of AeD7L1 and AeD7L2 are 67.73 and 65.81, respectively. Meanwhile, serotonin has a relatively higher aliphatic index of 101.69 (see Table 1). This parameter looks for aliphatic side chains (alanine, glycine, isoleucine, leucine, proline, and valine) that contribute to their aliphatic index associated with the thermostability of the protein when subjected to a range of temperatures. High aliphatic indices mean that the protein can withstand heat without denaturation.

3.1.2 Instability Index
The instability index of AeD7L1 and AeD7L2 is 40.45 and 33.04 respectively. Meanwhile, serotonin has a score of 38.07. In proteins, an instability index of below 40 is observed to be stable, while proteins that have an index of above 40 are more likely to be unstable. A study by Gamage *et al*. stated that proteins with a high instability index are more likely to be unstable because of the heightened risk of misfolding and aggregation, ultimately contributing to the pathogenicity of the disease they are causing [14].

3.1.3 N-Terminal
Both the D7 proteins and serotonin have methionine (Met or

M) as the starting amino acid in their N-terminal. This amino acid is synthesized from the codon 'AUG' which is also known as the start codon. The presence of Met at the beginning of the N-terminal denotes its presence as an initiator of protein synthesis. Moreover, it also has a crucial role in maintaining the stability of the protein and in protein-to-protein interactions [15, 16].

3.1.4 Theoretical pI
The theoretical pI is the generated pH at which the net charge of a protein molecule is zero. Amino acids are the main driving force that dictates the pI of a protein by combining their pKa values. Generally, proteins are positively charged at a pH below their pI and negatively charged at a pH above their pI. The folded states of proteins are also often most stable at the pH near their pI which also relates to their optimal pH. A study by Tokmakov, *et al*. showed that the theoretical pI of proteins is correlated in terms of their solubility and subcellular localization which also corresponds to the environment that the proteins are usually subjected to [17]. In terms of solubility at a given pH, proteins are least soluble in aqueous solutions with pH close to their pI, and the lowest rate of soluble expression is observed in proteins with

a pI of 7.0-7.5. Relating to the results, both salivary proteins express a theoretical pI of above 7.5, which makes them relatively more soluble than those within the range of 7.0-7.5. The solubility of proteins has been considered one of the factors that greatly affect the manufacture of therapeutic advancements making it a key attribute in choosing druggable targets that can be utilized for experimental characterizations [18, 19].

The subcellular localization of these proteins can also be determined by their theoretical pI because of the varying pH of each compartment in a cell. In correspondence with specific pI patterns observed by Tokmakov, *et al*. wherein proteins in the cytoplasm, Golgi apparatus, and vacuoles are highly biased towards pI that is acidic [17]. Meanwhile, proteomes inside the mitochondria have a higher avidity with a basic pI. The acidity and basicity of the proteins also contribute to the degree of their interactions inside the cell.

According to Chasapis and Konstantinoudis in eukaryotes, basic proteins have the lowest average of interactions and acidic proteins have the highest degree of interactions [20].

3.2 Hydropathicity of AeD7L1, AeD7L2, and Serotonin
The grand average of hydropathicity (GRAVY) is based on the hydrophobicity value of each peptide in a protein divided by the length of the protein sequence. Positive GRAVY values indicate hydrophobic; negative values mean hydrophilic [21]. Correlating with the given result, the D7 proteins AeD7L1 and AeD7L2 obtained scores of –0.659 and -0.649, respectively (Table 1), and resulted in being hydrophilic, while serotonin obtained a positive score of 0.43 is said to be hydrophobic (Table 1). Below are the hydropathicity plots made by Gagsteiger *et al*. based on the study of Kyte and Dolittle [22].
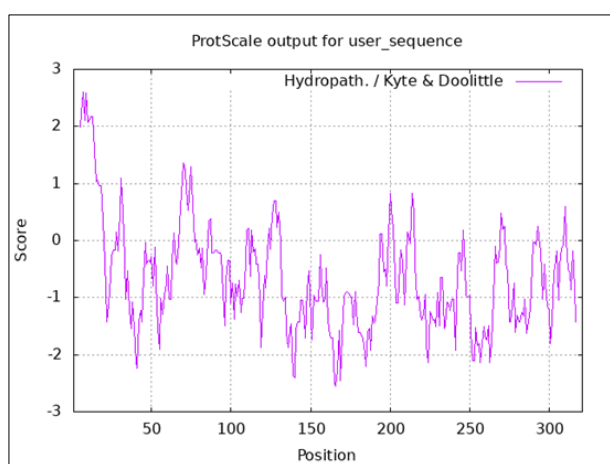

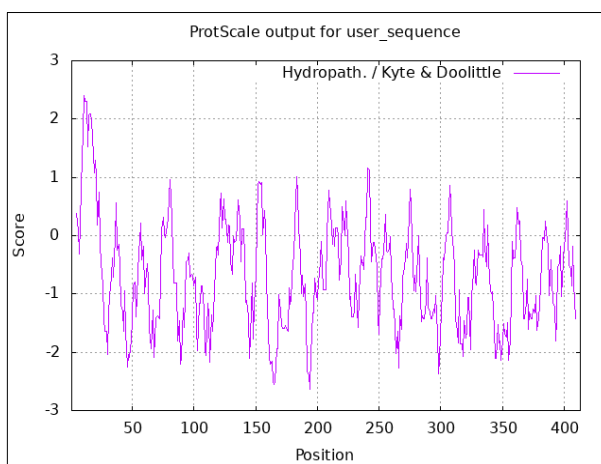**Fig 2:** Hydropathicity plot for AeD7L1
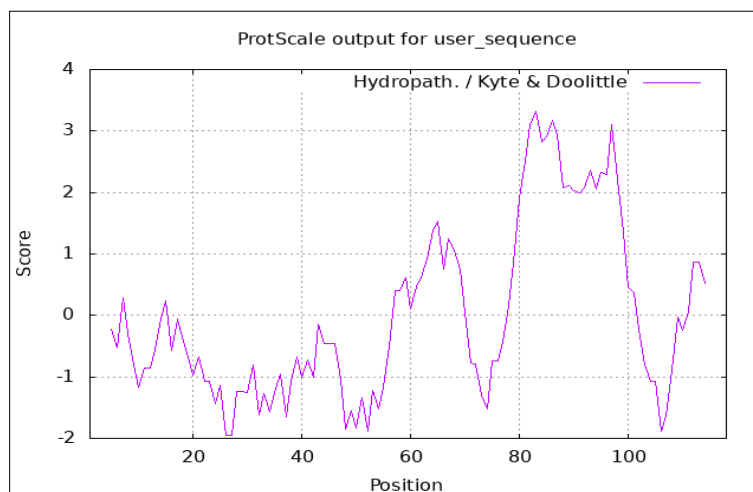

Fig 3: Hydropathicity plot for AeD7L2

Fig 4: Hydropathicity plot for Serotonin

Aside from the theoretical pI of the proteins, the GRAVY index is also useful in determining the subcellular localization of the proteins. Regions of the graph that have hydrophobic regions are more likely to be found in non-aqueous environments such as the interior of the plasma membrane. Meanwhile, the hydrophilic regions of the protein have a high chance of being found in aqueous regions such as the cytosol or the nucleus. As both the D7 proteins are hydrophilic in nature, it might be attributed to their distant relation to

odorant-binding proteins (OBPs).
OBPs usually reside in the sensory organs of arthropods, aiding in the binding and transport of hydrophobic molecules in aqueous environments [23].

3.3 Antigenicity of AeD7L1 and AeD7L2
The researchers were able to identify the antigenicity of the two D7 proteins through the usage of Predicted Antigenic Peptides based on the Kolaskar and Tongaonkar method.

There are 11 antigenic determinants for AeD7L1 with a total average antigenic propensity score of 1.0295 (Figure 5; Table 2). AeD7L2 has a higher number of antigenic determinants but a lower average antigenic propensity of 1.0216 (Figure 6; Table 3) which may be attributed to its greater number of amino acids and total number of atoms than AeD7L1.
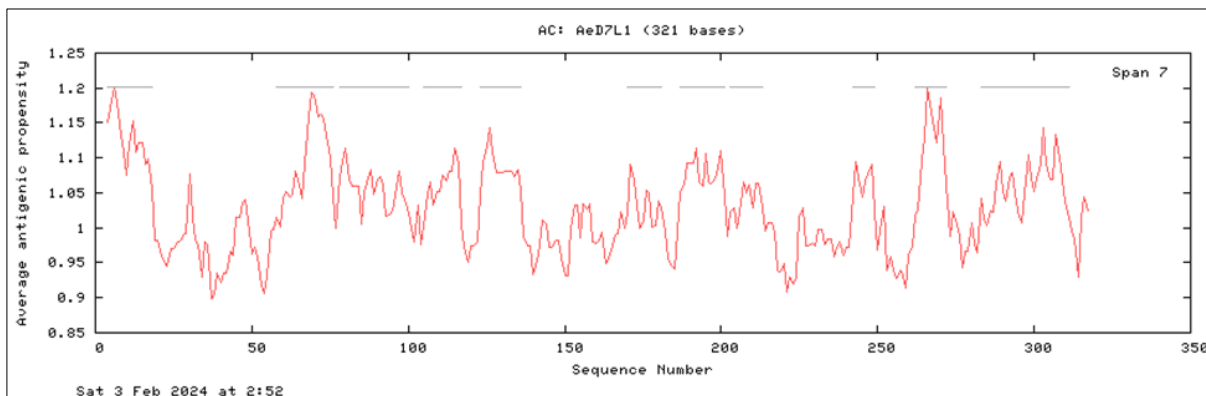

Fig 5: Antigenic Plot for AeD7L1. The average antigenic propensity is 1.0295, and the 11 antigenic determinants are indicated by gray lines
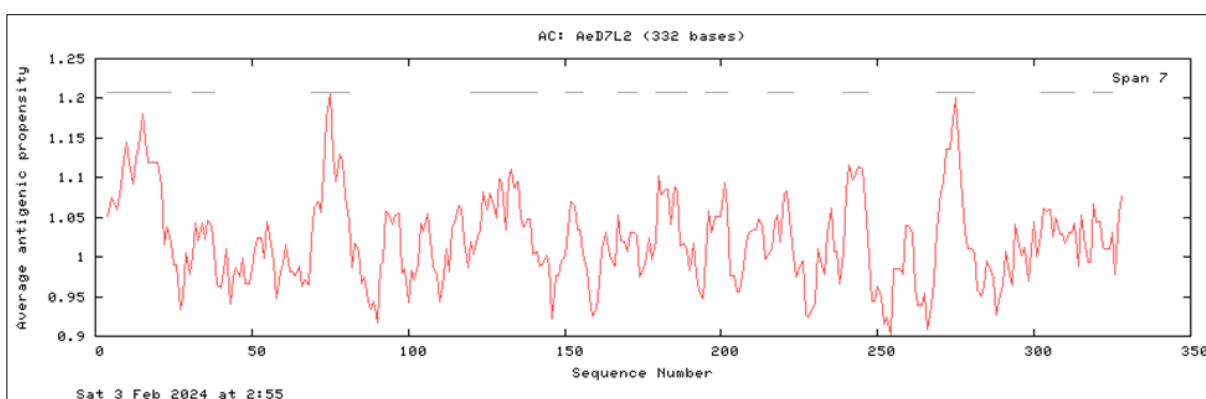

Fig 6: Antigenic Plot for AeD7L2. The average antigenic propensity is 1.0216, and the 13 antigenic determinants are indicated by gray lines.

Table 2: Antigenicity plot for AeD7L1

N	Start Position	Sequences	End Position
1.	4	PLLLAIVTTFSVVAS	18
2.	58	EPVDSPATQCGFKCVLVRT	76
3.	78	LYDPVAQKFDASDVIQEQFKAYPS	100
4.	105	SKVEAYANAVQQL	117
5.	123	DCAAVFKAYDPVHK	136
6.	170	SYFEFCENKYYP	181
7.	187	RQQLCKIRQYTVLDD	201
8.	203	LFKEHTDCVMK	213
9.	242	ALEKVLND	249
10.	262	SWHYYKCLVES	272
11.	283	DYREVRSQIYAFNLPKKQVYSKPAVQSQV	311

Table 3: Antigenicity plot for AeD7L2

N	Start Position	Sequences	End Position
1.	4	PRKFLLSSFILAALHVTAAPL	24
2.	31	EQLRFITS	38
3.	69	QATQCYTKCVLEK	81
4.	120	STFDFIPPLKSSSCSEVFEAFK	141
5.	150	TIRAILF	156
6.	167	QEKGVKI	173
7.	179	SLFMHCEALNY	189
8.	195	QRKDLCGI	202
9.	215	RHMECIFKG	223
10.	239	ARDFIVVKK	247
11.	269	GKIAVHYYKCLLMN	281
12.	302	DYFAALTGKLKP	313
13.	319	VRKQVDD	325

The antigenicity of these D7 proteins is important in enhancing the replication and pathogenicity of viruses carried by the female *Aedes aegypti*, specifically in the Dengue Virus (DENV). Being the most abundant in the array of salivary proteins, they have a definite role in inhibiting inflammatory and hemostatic responses of the host, leading to the upregulated effects of the disease once the event of a blood vessel injury occurs [24]. AeD7L1 has a higher antigenicity score which may be attributed to its higher affinity for serotonin than AeD7L2, effectively inhibiting the functions of the biogenic amine. This leads to the hindrance of sensing itch and pain as well as vasoconstriction and increased vascular permeability of the blood vessels, leading to a higher risk of blood loss that may progress into severe cases of bleeding if left untreated [25, 26].

3.4 Alignment of the Amino Acid Sequences of AeD7L1 and AeD7L2

Using the Clustal Omega, the researchers have aligned the sequences of both AeD7L1 and AeD7L2 to observe the differences and similarities between the two salivary proteins in terms of their amino acids. Their components vary in positions that greatly affect their structure, ultimately causing a degree of variation in their mechanism and role in the inhibition of primary hemostasis of the host— primarily focusing on the binding of biogenic amines such as Serotonin. These effects will be further explained in the presentation of the protein structures.


```
sp|P18153|ALL2_AEDAE        ------MKLPLLLAIVTTFSVVASTGPFDPEEMLFTFTRCMEDNLEDGPNRLPMLAKWKE        54
tr|A0A1S4FDI7|A0A1S4FDI7_AEDAE  MFPPRKFLLSSFILAALHVTAAPLWDAKDPEQLRFITSRCMEDWYPKAKNPKAALQNWLG        60
                            : *  :: .  .:..  .  ***:: *  :***** ..  *    * :*

sp|P18153|ALL2_AEDAE        WINEPVDSPATQCFGKCVLVRTGLYDPVAQKFDASVIQEQFKAYPSLGE--KSKVEAYAN        112
tr|A0A1S4FDI7|A0A1S4FDI7_AEDAE  WKLEPSDDQATQCYTKCVLEKIGFYEPGEKRFKGVRVMQQWETFHKYLNADREKVHDLTS        120
                            *  ** *. ****: ****  : *:*:* ::*..  : :*:::: .  :  :.**. :.

sp|P18153|ALL2_AEDAE        AVQ-QLPSTNNDCAAVFKAYDPVHKAHKDTSKNLFHGNKELTKGLYEKLGKDIRQKKQSY        171
tr|A0A1S4FDI7|A0A1S4FDI7_AEDAE  TFDFIPPLKSSSCSEVFEAFKKVNGKHSETIRAILFGKGESSKKYYQEKGVKIKQKEQSL        180
                            ::: *  ....*: **:*:. *:  *.:* : ::.*: *:*  *:: *.*:**:**

sp|P18153|ALL2_AEDAE        FEFCENKYYPAGSDKRQQLCKIRQYTVLDDALFKEHTDCVMKGIRYITKNNELDAEEVKR        231
tr|A0A1S4FDI7|A0A1S4FDI7_AEDAE  FMHCEALNYPKGSPQRKDLCGIRKYQMGSGIVFERHMECIFKGLRYMTSKNELDVDEIAR        240
                            * .** . ** ** :*::** **:* : .. :*:.* :*::**:**:*.:****.:*: *

sp|P18153|ALL2_AEDAE        DFMQVNKDTKALEKVLNDCKSKE-PSNAGEKSWHYYKCLVE-SSVKDDFKEAFDYREVRS        289
tr|A0A1S4FDI7|A0A1S4FDI7_AEDAE  DFIVVKKKPDAMKAMMKTCKANLKEKNPGKIAVHYYKCLMNDSKVTNDFKEAFDYREVRS        300
                            **:. *:.*. .*:: ::: **::    .* *: * ******:: *.*.:*************

sp|P18153|ALL2_AEDAE        QIYAFNLP-KKQVYSKPAVQSQVMEIDGKQCPQ        321
tr|A0A1S4FDI7|A0A1S4FDI7_AEDAE  KDYFAALTGKLKPYSRSDVRKQVDDIDKIQCS-        332
                            : *   *  * : **: *:.** :** **
```

Fig 7: Sequence Alignment of AeD7L1 and AeD7L2 using Clustal Omega. AeD7L1 (P18153) has 321 amino acids and AeD7L2 (A0A1S4FDI7) has 332 amino acids

3.5 Analysis of the Secondary Structure of AeD7L1, AeD7L2, and Serotonin

The secondary structures of the AeD7L1, AeD7L2, and Serotonin were processed through the SOPMA software. The results concluded that AeD7L1 has 54.52% alpha helix compared to AeD7L2 which has 56.33% alpha helix and 32.33% random coil, which has the highest percentage among the three proteins. Serotonin has the lowest percentage of both alpha helix and random coil with 30.51% and 28.81% results respectively, it also had the highest percentage in the extended strand with 30.51% compared to AeD7L1 at 9.03% and AeD7L2 at 8.73% (see Table 1).

Alpha helices in the protein structure contribute to the stability of the protein. It may also be related to the hemolytic activities of a protein as studied by Chen *et al*., suggesting that strong hemolytic activities of the peptides correlate with high hydrophobicity, high amphipathic residues, and increased helical structures in the protein [27]. Meanwhile, extended strands that are usually exposed to the protein's surroundings have a role in establishing hydrogen bonds with polar side chains such as water molecules. Eswar *et al*. stated that these exposed structures help in connecting distant parts of a protein using polar contacts, thus minimizing the process of folding into a compact beta-sheet [28].

3.5.1 Solvent Accessibility

These secondary structures may be attributed to the degree of accessibility of certain solvents interacting with the protein. Solvent-accessible surface area refers to the relative area where a solvent, such as water, interacts with the amino acid residue of the given protein. This will predict the functional stability and folding of the secondary structure of the protein. Moreover, the degree reflects the exposure of the residue that allows hydrophobicity interaction between two molecules. For instance, buried amino acid residues tend to be more inaccessible to the surrounding molecules compared with exposed which interact with the solvent molecules. As such, the buried class elicits hydrophobic behavior that makes up the protein core essential for protein folding, consequently decreasing the access of the solvent to the surface of the protein [29]. Each type of secondary structure such as alpha-helix, beta-sheets, and random-coil/turns corresponds to specific accessibility classes. Studies mentioned that random coils had more accessible exposed residues while beta-pleaded sheets were the most inaccessible because they were buried in the protein core [30].

AeD7L1 and AeD7L2 have secondary structures exhibiting highly marked alpha-helix structures. Lins *et al*. stated that most alpha-helices were amphipathic which follows the behavior of hydropathicity when the protein folds; knowing that both AeD7 long forms are hydrophilic, they are water-accessible at the surface of the protein structure. Furthermore, both proteins also displayed more random coils, meaning that they have more accessible residues, especially hydrophilic amino acid residues of the protein [30]. In contrast, serotonin is hydrophobic in which extended beta-sheets are predominantly present. It was suggested that beta-sheets are more likely inaccessible to water since their hydrophobic surface residues

decrease, thus, they were mostly buried at the protein core.

3.5.2 Salt Bridge Formation

The possibility of a salt bridge formation between AeD7 protein and the biogenic amine serotonin can be evaluated based on their structural and chemical properties. While such interactions have not been explicitly characterized in studies, salt bridge formation is possible. In the study of Mans *et al*., they described the structure of D7r4, which shows a similar sequence to that of the AeD7 protein including the "long forms" of D7 being comparable with that of the whole D7r protein, this is further supported by the study Calvo *et al*. where the binding pocket structures of the ligand-bound AeD7 C-terminal domain and the D7r4-serotonin complex are nearly similar with one another [31, 32]. It is shown in the study that the central binding pocket accommodates biogenic amines like serotonin via aromatic residues and hydrogen bonding interactions and that an apparent salt bridge was formed between Arg and Glu to establish a hydrogen bond with the amino group of the serotonin [31]. The presence of glutamate indicates that the protein possesses ionizable amino acid chains that can carry negative charges at physiological pH.
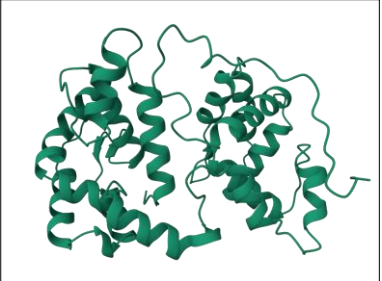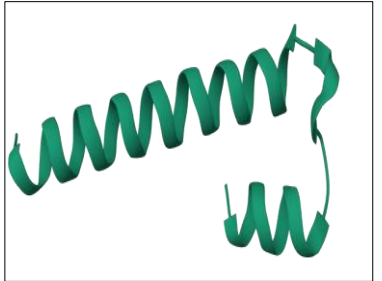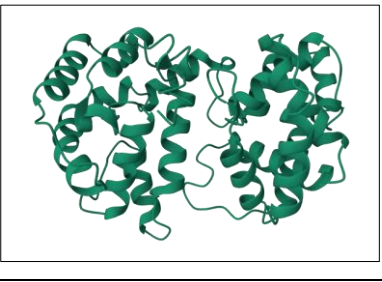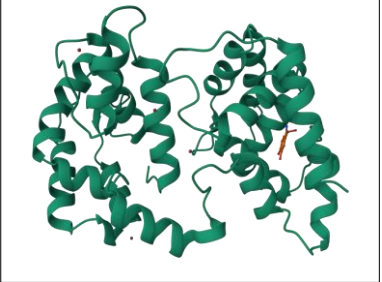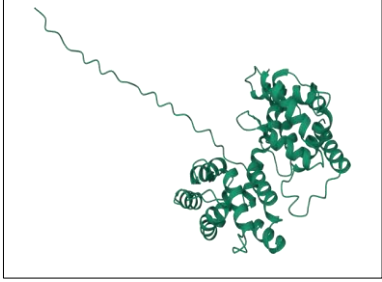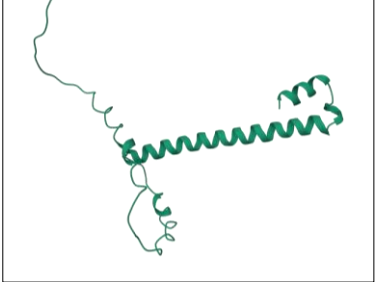
On the other hand, serotonin is said to have an ionizable amino group that carries a positive charge under physiological pH [33]. As such, the negatively charged carboxylate groups of Asp or Glu residues in AeD7 and the positively charged amino group of serotonins could potentially interact with one another and form salt bridges which are ionic interactions between oppositely charged groups. However, for a salt bridge to be formed, both charged groups should be in proximity and properly oriented such that the binding pocket of AeD7 where the serotonin could bind positions its amino group closely to the negatively charged residues of AeD7 to enable the formation of such salt bridge [34]. The formation of such bonding contributes greatly to the binding affinity and specificity of the protein-ligand interactions and the overall binding of serotonin to the AeD7 protein.

3.6 Homology Modeling

Amino-acid sequence-based predicted structures are shown in Table 4, which illustrates the raw data from homology modeling programs (i.e., SWISS-MODEL, Phyre 2, and i-TASSER), visualized using Mol* viewer. These structures, however, are prone to residual errors. Thus, these are subjected to validation under the SAVES server. The models with the highest values in the most favored core regions (Co) in the Ramachandran plots pass for the validation screening. With 94.5% Co, SWISS-MODEL is selected for AeD7L1. For AeD7L2, the SWISS-MODEL (model 2) with 90.6% Co passes the validation. Lastly, SWISS-MODEL (model 2) has the highest Co with 97.8%.

Table 4: Raw Protein Structures from SWISS-MODEL, Phyre 2, and i-TASSER (Mol* viewer)

Software	AeD7L1 (AAEL006424)	AeD7L2 (AAEL006417)	Serotonin (AGB76027.1)
Phyre 2			
I-TASSER			
SWISS-MODEL			
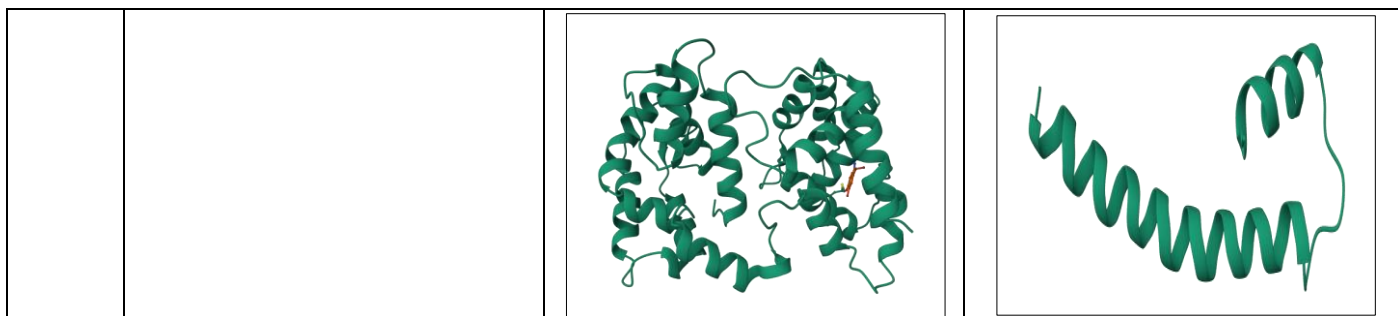
These models are further improved in the refinement process, wherein the qualified structures from validation are submitted to GalaxyWEB to detect unreliable regions and will be refined to improve the models [35]. For the selection among the refined structures, the criteria used are: (1) highest Rama (Ramachandran) favored regions; and (2) lowest MolProbity, whereas a lower value score implies that the protein model has more favorable and realistic geometry, with fewer steric clashes [36].
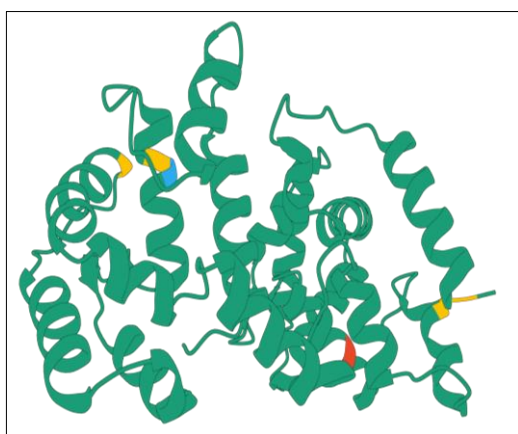

Fig 10: Serotonin

Fig 8: AeD7L1. Key cysteines (Cys-67, Cys-24, Cys-191, and Cys-319) important for disulfide bonding with serotonin are marked yellow, Histidine-207 known to bind the 5-hydroxyl group of Serotonin via hydrogen bond is marked red, and the blue region indicates Phenylalanine-68.
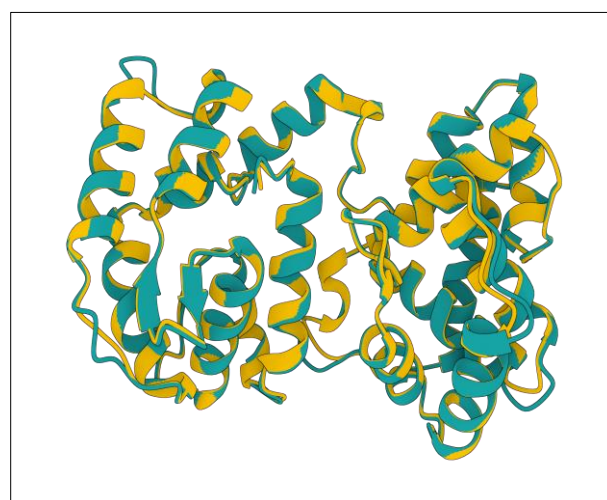

Fig 11: AeD7L1 (green) and AeD7L2 (yellow) in superimposition showing regions of similarities and differences

Based on molecular characterization and superimposition of AeD7L1 and AeD7L2 quaternary structures, it can be derived that these two proteins share the same functionality in that they inhibit host hemostasis by acting as scavengers [26]. This pathological event is attributed to the binding of biogenic amines, specifically, serotonin in the C-terminal of the proteins. After the alignment using the clustal omega and viewing of positions of amino acids via Jalview, the researchers were able to identify the key amino acids responsible for binding. The disulfide bonds formed from key cysteines of the AeD7L1 (see Figure 8) such as Cys-67, Cys-24, Cys-191, and Cys-319, and the binding of the 5-hydroxyl group of serotonin to the histidine-207 firmly attaches the biogenic amine to the AeD7L1. Similarly, AeD7L2 attaches to serotonin with disulfide bonds on Cys-73, Cys-133, Cys-200, and Cys-331, and hydrogen bonds to histidine-216. These cysteine molecules are generally responsible for the attachment of biogenic amines such as norepinephrine, epinephrine, and serotonin by disulfide bonds. Stabilization of this binding mechanism is achieved by the disulfide bonds by increasing the number of amino residues between cysteine molecules. It was also suggested by Feige *et al*., that these

Fig 9: AeD7L2. Key cysteines (Cys-73, Cys-133, Cys-200, and Cys-331) important for disulfide bonding with serotonin are marked yellow, Histidine-216 known to bind the 5-hydroxyl group of Serotonin via hydrogen bond is marked red, and the yellow-green region indicates Tyrosine-74.

molecular bonds decrease the solvent-binding accessibility of the amino acid residues, thus burying the hydrophobic surface residue and hydrogen-bond donors and acceptors [37].

A study by Calvo *et al*. also stated that this induced-fit type of binding process carries out a conformational change in the protein, leading to the increase in a variety of ligands to be accommodated as well as the closing of the entry path of the biogenic amine and trapping the bound ligand [32]; this mechanism is also seen in other salivary proteins of arthropods that can bind biogenic amines. In cases of the absence of these key cysteines, this binding process will not occur, and the salivary protein is observed to lose the function of binding to biogenic amines as seen in other arthropods. On the other hand, histidine molecules of both D7 proteins specifically bind and stabilize the scavenged serotonin which may contribute to its high affinity.

Additionally, the negatively charged carboxylate groups of glutamate and aspartate of AeD7L1 and AeD7L2 can participate in favorable or unfavorable electrostatic interactions with positively charged groups, such as the amine group of biogenic amines, present within the adjacent biomolecules [38, 39]. These interactions contribute to the stabilization of the binding of these proteins, which may also explain their affinity to the ligand.

Despite the similarities in the binding of biogenic amines in their C-terminal, there are various key differences in their amino acid sequence, consequently altering their function and mechanism in the event of blood feeding. One distinct difference between the two salivary proteins is the substitution of Tyrosine-74 in AeD7L2 from the Phenylalanine-68 of AeD7L1 (see Figure 9). This key position in the alignment of amino acids changes the function of the N-terminal of the protein. As for AeD7L2, this caused an additional change that let it possibly attach to Thromboxane A_2, as this protein shows binding to TXA_2 analog, U-46619 [4]. TXA_2 is a known procoagulant eicosanoid that plays a vital role in triggering physiological host hemostasis, through vasoconstriction and promoting platelet aggregation by decreasing adenylate cyclase and cAMP reactions, inducing the release of ionized calcium in the dense tubular system of the platelets [40]. If this process is inhibited by the AeD7L2, it may lead to the vasodilation of blood vessels, and the failure to produce a stable platelet plug, leading to the successful blood-feeding of the female *Aedes aegypti* due to excess bleeding.

Meanwhile, TXA_2 is considered immeasurable *in vivo* and *in vitro* due to its short half-life of only 30 seconds. This explains why efforts to produce a clinical assay in measuring it have proved to be unsuccessful. On the other hand, 11-dehydrothromboxane B_2, a resulting product of liver enzymes after acting upon the TXA_2, can be considered as it is more stable and measurable. Assays measuring this substance have been useful in monitoring aspirin therapy and detecting aspirin therapy failure [40]. Thus, this shows the potentiality of urinary 11-dehydrothromboxane B_2 in becoming a parameter or variable of a future study related to the pathological phenomenon of AeD7L2.

3.7 Protein Modifications of AeD7L1 and AeD7L2

In the pharmaceutical aspect, various methods concerning druggable targets in proteins remain one of the pillars for drug discoveries. With regards to the established physicochemical properties, binding mechanisms, and protein structures stated in this study, it can pave the way for a significant progression into alleviating the severity of arthropod-borne viruses.

3.7.1 Chemical Modifications

A study made by Naowarojna *et al*. stated that chemical modifications in proteins by the usage of a probe can determine the functional characteristics of a protein and may contribute to stabilizing the protein and enhancing its half-life, leading to the control of these macromolecules [41]. One example is the cysteine modification; AeD7L1 and AeD7L2 have key cysteine residues that form disulfide bonds with their bound biogenic amine and with the nature of cysteines having a thiol with high nucleophilicity, it can be a good candidate for single-site modifications. Moreover, sites of binding in both salivary proteins exemplify high affinities with biogenic amines. This functional characteristic may be a target in ligand-directed protein modifications, wherein the alteration is driven by the formation of bonds; allowing for the alteration at the active site of binding [42]. Moreover, the N-terminal of AeD7L1, AeD7L2, and serotonin is methionine - a hydrophobic residue (see Table 1). This position is usually solvent-exposed and may provide regions of well-defined reactive chemical pockets for its respective protein or peptide targets [43]. Aledo also stated in his study that methionine plays a role in various protein activities, and with the modification of this residue with reversible redox reactions may be manipulated with enzyme catalysts, possibly impacting the structure and properties of the protein or the substrate in the targeted interaction [44].

4. Summary of findings, conclusion, and recommendations

Molecular characterization and homology modeling through *in silico* methods have served as useful tools in determining the physicochemical characteristics of biomolecules by delineating unique molecular structures and properties. The two clinically significant long-form proteins from the salivary gland of female *Aedes aegypti* - AeD7L1 and AeD7L2 play a vital role in disrupting the host hemostasis during bloodletting by binding to biogenic amine serotonin.

The researchers were able to determine that during a successful bloodletting of *Aedes aegypti*, both D7 proteins, AeD7L1 and AeD7L2, show the ability to bind to serotonin. The researchers also determined the clinical significance of their aliphatic and instability indices, the theoretical pI and its relationship with the proteins' solubility and subcellular localization, and the relationship with their hydropathicity. Moreover, the analysis of structure with the help of homology modeling led the researchers to determine the possible salt bridge formation, and key amino acids that play vital roles in their ability to bind to biogenic amines wherein AeD7L1 shows possibly higher affinity than AeD7L2 as shown by their antigenicity scores.

While the function of binding to serotonin is common among the two long forms, AeD7L2 may be as equally as significant as AeD7L1 due to its additional function in the potential binding of TXA_2 to its N-terminal end. Future wet lab experiments and research may be conducted to prove this novel function of AeD7L2 through the inclusion of urinary 11-dehydrothromboxane B_2 as a parameter or variable of a future study, as assays to measure TXA_2 proved to be cumbersome.

The global burden of various arthropod-borne viruses such as the Dengue Virus (DENV) continues to plague various

countries without an effective method to combat it. In the determination of the salivary proteins that play crucial roles in its pathogenesis, the study was able to establish its physicochemical properties, structure, and binding mechanisms that contribute to the inhibition of host hemostasis leading to excessive bleeding. The pharmaceutical industry may utilize the information indicated in this study to develop therapeutic agents that will aid in drug discoveries, vaccine developments, and other advancements that will possibly reduce the severity of the virus by conducting chemical modifications and tuning important protein parameters such as the extent of its hydrophobicity, amino acid composition, and structure.

References
1. Dengue situation update 681. [Internet]. 15 February 2024 [cited 2024 Feb 15]. Available from: https://cdn.who.int/media/docs/default-source/wpro---documents/emergency/surveillance/dengue/dengue-20240215.pdf?sfvrsn=fc80101d_138.
2. Conway MJ, Londoño-Rentería B, Troupin A, Watson AM, Klimstra WB, Fikrig E, *et al. Aedes aegypti* D7 Saliva Protein Inhibits Dengue Virus Infection. PLoS Negl Trop Dis. 2016;10(9):e0004941.
3. Ribeiro JM. Blood-feeding arthropods: live syringes or invertebrate pharmacologists? Infect Agents Dis. 1995;4(3):143-52.
4. Martin-Martin I, Kern O, Brooks S, Smith LB, Valenzuela-Leon PC, Bonilla B, *et al*. Biochemical characterization of AeD7L2 and its physiological relevance in blood feeding in the dengue mosquito vector, *Aedes aegypti*. FEBS J. 2020;288:2014-29.
5. Sam S, Omar SFS, Teoh B, Abd-Jamil J, AbuBakar S. Review of Dengue Hemorrhagic fever Fatal cases Seen among adults: A retrospective study. PLoS Negl Trop Dis. 2013;7(5):e2194.
6. Yacoub S, Wills B. Predicting outcome from dengue. BMC Med., 2014, 12(1).
7. Rajapakse S. Dengue shock. J Emerg Trauma Shock. 2011;4(1):120.
8. Murugesan A, Mythreyee M. Dengue Virus. Elsevier eBooks; c2020. p. 281-359.
9. Troost B, Smit JM. Recent advances in antiviral drug development towards dengue virus. Curr. Opin. Virol. 2020;43:09-21.
10. Dengue Prevention and Control Program [Internet]. n.d. [Cited 2024 Apr 8]. Available from: https://doh.gov.ph/national-dengue-prevention-and-control-program.
11. Larson HJ, Hartigan-Go K, de Figueiredo A. Vaccine confidence plummets in the Philippines following dengue vaccine scare: Why it matters to pandemic preparedness. Hum Vaccin. Immunother. 2019;15(3):625-7.
12. Hodos RA, Kidd BA, Shameer K, Readhead BP, Dudley JT. In silico methods for drug repurposing and pharmacology. Wiley Interdiscip. Rev. Syst. Biol. Med. 2016;8(3):186-210.
13. Miller LE, Stevens CD. Clinical Immunology and Serology: A Laboratory Perspective. 5th ed. F.A. Davis; c2021.
14. Gamage DG, Gunaratne A, Periyannan GR, Russell TG. Applicability of instability index for *in vitro* protein stability prediction. Protein Pept. Lett. 2019 May 1;26(5):339-47.
15. Lim JM, Kim G, Levine RL. Methionine in Proteins: It's not just for protein initiation anymore. Neurochem Res. 2018;44(1):247.
16. Valley CC, Cembran A, Perlmutter JD, Lewis AK, Labello NP, Gao J, *et al*. The methionine-aromatic motif plays a unique role in stabilizing protein structure. J Biol. Chem. 2012;287(42):34979-91.
17. Tokmakov AA, Kurotani A, Sato K. Protein pI and Intracellular Localization. Front Mol. Biosci. 2021;8:775736.
18. Hon K, Marusiak M, Martínek T, Kunka A, Zendulka J, Bednář D, *et al*. SoluProt: prediction of soluble protein expression in *Escherichia coli*. Bioinformatics. 2021;37(1):23-8.
19. Vaňáček P, Šebestová E, Babková P, Bidmanová Š, Daniel L, Dvořák P, *et al*. Exploration of Enzyme Diversity by Integrating Bioinformatics with Expression Analysis and Biochemical Characterization. ACS Catal. 2018;8(3):2402-12.
20. Chasapis CT, Konstantinoudis G. Protein isoelectric point distribution in the interactomes across the domains of life. Biophys. Chem. 2020;256:106296.
21. Chang KY, Yang R. Analysis and Prediction of Highly Effective Antiviral Peptides Based on Random Forests. PLoS ONE., 2013, 8(8).
22. Kyte J, Doolittle RF. A simple method for displaying the hydropathic character of a protein. J Mol. Biol. 1982;157(1):105-32.
23. Rihani K, Ferveur J. The 40-Year Mystery of Insect Odorant-Binding Proteins. Biomolecules. 2021;11(4):509.
24. Cox JC, Mota J, Sukupolyi-Petty S, Diamond MS, Rico-Hesse R. Mosquito bite delivery of dengue virus enhances immunogenicity and pathogenesis in humanized mice. J Virol. 2012;86(14):7637-49.

25. Smith LJ, Fiebig KM, Schwalbe H, Dobson CM. The concept of a random coil: Residual structure in peptides and denatured proteins. Folding Des. 1996;1(5):R95-R106.
26. Alvarenga PH, Andersen JF. An Overview of D7 Protein Structure and Physiological Roles in Blood-Feeding Nematocera. Biology. 2022;12(1):39.
27. Chen Y, Mant CT, Farmer SW, Hancock REW, Vasil ML, Hodes RS. Rational design of alpha-helical antimicrobial peptides with enhanced activities and specificity/therapeutic index. J Biol. Chem.; c2005.
28. Eswar N, Ramakrishnan C, Srinivasan N. Stranded in isolation: structural role of isolated extended strands in proteins. Protein Eng Des Sel., 2003, 16(5).
29. Ausaf Ali S, Hassan I, Islam A, Ahmad F. A review of methods available to estimate solvent-accessible surface areas of soluble proteins in the folded and unfolded states. Curr Protein Pept. Sci. 2014;15(5):456-476.
30. Van Der Lee R, Buljan M, Lang B, Weatheritt RJ, Daughdrill GW, Dunker AK, et al. Classification of intrinsically disordered regions and proteins. Chem. Rev. 2014;114(13):6589-6631.
31. Mans BJ, Calvo E, Ribeiro JMC, Andersen JF. The Crystal Structure of D7r4, a Salivary Biogenic Amine-binding Protein from the Malaria Mosquito Anopheles gambiae. J Biol. Chem. 2007;282(50):36626-36633.
32. Calvo E, Mans BJ, Ribeiro JM, Andersen JF. Multi-functionality and mechanism of ligand binding in a mosquito anti-inflammatory protein. Proc. Natl. Acad. Sci. U S A. 2009;106(10):3728-3733.
33. Andrews PW, Boysj C, Brenton L, Green L, Gasser PJ, Lowry CA, Pickel VM. All the brain's a stage for serotonin: the forgotten story of serotonin diffusion across cell membranes. Proc. R Soc. B., 2022, 289(1986).
34. Bosshard HR, Marti DN, Jelesarov I. Protein stabilization by salt bridges: concepts, experimental approaches and clarification of some misunderstandings. J Mol. Recognit. 2004;17(1):1-16.
35. Ko J, Park H, Heo L, Seok C. Galaxy WEB server for protein structure prediction and refinement. Nucleic Acids Res. 2012 May 30;40(W1):W294-7.
36. Williams CK, Headd JJ, Moriarty NW, Prisant MG, Videau LL, Deis LN, et al. MolProbity: More and better reference data for improved all-atom structure validation. Protein Sci. 2018;27:293-315.
37. Feige MJ, Braakman I, Hendershot LM. CHAPTER 1.1. Disulfide bonds in protein folding and stability. Chem Biol.; c2018. p. 1-33.
38. Dill KA. Dominant forces in protein folding. Biochemistry. 1990;29(31):7133-7155.
39. Pace CN, Grimsley GR, Scholtz JM. Protein Ionizable Groups: pK Values and Their Contribution to Protein Stability and Solubility. J Biol. Chem. 2009;284(20):13285-13289.
40. Keohane E, Otto C, Walenga J. Rodak's Hematology: Clinical Principles and Applications. 6th ed. Elsevier; c2020.
41. Naowarojna N, Cheng R, Lopez J, Wong C, Qiao L, Liu P. Chemical modifications of proteins and their applications in metalloenzyme studies. Synth Syst. Biotechnol. 2021;6(1):32-49.
42. Shadish JA, DeForest CA. Site-selective protein modification: from functionalized proteins to functional biomaterials. Matter. 2020;2(1):50-77.
43. Jiang H, Chen W, Wang J, Zhang R. Selective N-terminal modification of peptides and proteins: Recent progresses and applications. Chin. Chem. Lett. 2022;33(1):80-88.
44. Aledo JC. Methionine in proteins: The Cinderella of the proteinogenic amino acids. Protein Sci. 2019;28(10):1785-1796.

Darren Nichole Ocampo
Far Eastern University, Institute of Health Sciences and Nursing, Department of Medical Technology, Sampaloc, Manila, 1015, Philippines

Hans Ezekiel Olorosisimo
Far Eastern University, Institute of Health Sciences and Nursing, Department of Medical Technology, Sampaloc, Manila, 1015, Philippines

Erica Julienne Palting
Far Eastern University, Institute of Health Sciences and Nursing, Department of Medical Technology, Sampaloc, Manila, 1015, Philippines

Lyka Angel Pascual
Far Eastern University, Institute of Health Sciences and Nursing, Department of Medical Technology, Sampaloc, Manila, 1015, Philippines

Edrole Siegfred Ramos
Far Eastern University, Institute of Health Sciences and Nursing, Department of Medical Technology, Sampaloc, Manila, 1015, Philippines

Julienne Aubrey Ruela
Far Eastern University, Institute of Health Sciences and Nursing, Department of Medical Technology, Sampaloc, Manila, 1015, Philippines

Sebastiane Salonga
Far Eastern University, Institute of Health Sciences and Nursing, Department of Medical Technology, Sampaloc, Manila, 1015, Philippines

Andrea Bianca Francesca Sanchez
Far Eastern University, Institute of Health Sciences and Nursing, Department of Medical Technology, Sampaloc, Manila, 1015, Philippines

Mary Rose Lirio
Far Eastern University, Institute of Health Sciences and Nursing, Department of Medical Technology, Sampaloc, Manila, 1015, Philippines

Appendices
Appendix A: Ramachandran plot calculations in the SAVES Server
AeD7L1

SWISS-MODEL	94.5%	5.1%		0.4%
Phyre 2	92.0%	6.6%		1.5%
i-TASSER	83.8%	13.1%	2.1%	1.0%

AeD7L2

SWISS-MODEL 1	89.7%			10.3%	
SWISS-MODEL 2	90.6%	8.3%		0.4%	0.7%
Phyre 2	89.5%	8.0%	1.8%	0.7%	
i-TASSER	76.0%	20.0%	2.3%	1.7%	

Serotonin

SWISS-MODEL 1	77.5%	15.3%	6.3%	0.9%
SWISS-MODEL 2	97.8%	2.2%		
Phyre 2	95.3%	4.7%		
i-TASSER	44.6%	35.7%	13.4%	6.2%

Appendix B: Galaxy WEB Refined Model Scores

AeD7L1	GDT-HA	RMSD	MolProbity	Clash score	Poor rotamers	Rama favored
Initial	1.0000	0.000	0.705	0.6	0.0	98.7
MODEL 1	0.9925	0.266	1.584	10.7	1.1	99.0
MODEL 2	0.9834	0.315	1.518	9.9	0.4	99.3
MODEL 3	0.9900	0.278	1.476	8.9	0.7	99.3
MODEL 4	0.9900	0.281	1.555	10.9	0.4	99.0
MODEL 5	0.9892	0.282	1.504	8.7	1.1	99.0

AeD7L2	GDT-HA	RMSD	MolProbity	Clash score	Poor rotamers	Rama favored
Initial	1.0000	0.000	1.682	5.6	0.4	94.4
MODEL 1	0.9829	0.333	1.909	14.8	1.1	96.7
MODEL 2	0.9805	0.322	1.826	14.2	0.4	97.0
MODEL 3	0.9764	0.335	1.948	12.8	1.5	96.7
MODEL 4	0.9796	0.326	1.915	16.6	1.1	97.0
MODEL 5	0.9829	0.326	1.964	16.6	1.5	97.4

SEROTONIN	GDT-HA	RMSD	MolProbity	Clash score	Poor rotamers	Rama favored
Initial	1.0000	0.000	0.500	0.0	0.0	100.0
MODEL 1	0.9745	0.373	0.841	1.2	0.0	100.0
MODEL 2	0.9796	0.352	1.030	2.5	0.0	100.0
MODEL 3	0.9745	0.338	1.316	2.5	2.4	100.0
MODEL 4	0.9847	0.342	0.500	0.0	0.0	100.0
MODEL 5	0.9949	0.269	0.841	1.2	0.0	100.0