Structural study of the C-terminal domain of non-structural protein 1 and capsid protein from Japanese encephalitis virus

Thesis submitted in accordance with the

requirements of the University of Liverpool for

the degree of Doctor in Philosophy

By

Thanalai Poonsiri

January 2018

Abstract

Structural study of the C-terminal domain of non-structural protein 1

and capsid protein from Japanese encephalitis virus

Thanalai Poonsiri

Japanese encephalitis virus (JEV) is a mosquito-transmitted *Flavivirus* that is closely related to other emerging viral pathogens including dengue (DENV), West Nile (WNV) and Zika viruses (ZIKV). JEV infection can result in meningitis and encephalitis, which in severe cases cause permanent brain damage and death. JEV occurs predominantly in rural areas throughout South East Asia, the Pacific islands, and the Far East, causing around 68,000 cases worldwide each year. There is no specific treatment for JEV. This study aims to determine the molecular structure of new potential drug targets for JEV.

In this study, the JEV non-structural protein 1 C-terminal β-ladder domain (C-NS1) is presented at 2.1 Å resolution. The crystal structure of C-*JEV*NS1 shares a conserved fold with flavivirus C-NS1 domains. The surface charge distribution of C-*JEV*NS1 is similar to WNV and ZIKV but is significantly different from DENV. Analysis of the C-*JEV*NS1 structure, *in silico* molecular dynamics simulations and experimental solution small angle X-ray scattering, indicate extensive loop flexibility on the exterior of the protein. It is proposed that this together with charge distribution on the exterior of the protein influence NS1-host protein interaction specificity which may impact on pathogenicity. These factors may also affect the interaction with the monoclonal antibody, 22NS1, which is protective against WNV infection. Liposome and heparin binding assays indicate that only the N-terminal region of NS1 participates in the interaction with lipidic membranes and that sulphate binding sites are not the glycosaminoglycans binding interfaces.

For the first time, the crystal structure of the JEV capsid protein at 1.98 Å is also reported and compared to the existing flavivirus capsid protein. JEV capsid shows helical secondary structure (α helixes 1-4) and protein folding similar to DENV and WNV capsid proteins. It forms a homodimer by antiparallel pairing with another subunit (‘), α helix 1-1’, 2-2’, and 4-4’. The capsid dimer is believed to be the building block of the nucleocapsid. The flexibility of the N-terminal α helix 1 of the capsid could be important for its function. This dimer model agrees with a previous suggestion that the capsid protein interacts with RNA via the basic rich C-terminal, α4-α4’, and associates with lipid bilayers at the opposite hydrophobic, α2-α2’.

Acknowledgements

The Ph.D. is fully supported by Mahidol-Liverpool Stang Mongkolsuk Ph.D. scholarship. I appreciate this opportunity that allows me to have a wonderful experience in Liverpool. Thank you to my supervisors Professor Tom Solomon for his guidance and his powerful networking to make things happen, and Dr. Svetlana Antonyuk for supporting, leading, and working closely. Thank you to Professor Samar Hasnain, who runs the scholarship program and also the molecular biophysics group leader, for his support and interest in the project throughout and for an extensive discussion of the results.

This project is highly collaborative with many people. Thank for the helps from the University of Liverpool: Dr. Gareth Wright for SAXS data collection, guidance for SAXS data analysis, performing molecular dynamic simulation experiment, and mentoring throughout my PhD degree, Dr. Mark Wilkinson for conducting mass spectrometry analysis, Dr. Lance Turtle for an extensive discussion of the results and manuscript preparation, Dr. Richard Strange for assistance with the electrostatic surface map, Sujitra Keadsanti for assistance with Western bloting analysis, Kangsa Amporndanai for provinding bovine bc1 complex protein, Varunya Chantadul for providing SOD3 protein, Dr. Edwin Yates and Dr. Dominic Bryne for protein-GAGs interaction discussion, Dr. Dunhao Su for sharing bacterial cells, and Pawin Ngamlert for assistance with heparin binding assay. Thank you to friends from the Liverpool School of Tropical Medicine, Dr. Cristina Yuntayanes and Piyapon Jirawatcharadech, for

sharing the chaperone plasmid set. Thank you to all molecular biophysics group members and Liverpool brain infection group for they support.

Many thanks to the Oxford protein production facilities UK (OPPF-UK), Dr. Joanne Nattleship and Dr. Nahid Raman-huq, for performing protein expression in mammalian and insect cell systems.

Thank you to Dr. Mike Diamond, Washington University, for providing the WNV antibody and huge contribution in manuscript preparation.

Thanks to Synchrotron Soleil for the provision of Proxima 1 beamline and SAXS facilities. Use of SOLEIL was funded by the European Community's Seventh Framework Programme (FP7/2007-2013) under BioStruct-X (grant agreement number 283570 and proposal number 6714) and HHSN272201400018C (to M.S.D). Thanks to Diamond Synchrotron for the provision of I02 and I04 beamline.

Finally, big thanks to friends and family for sharing and support that have made my four years Ph.D. not too lonely.

Declaration

I declare that work in this thesis was carried out in accordance with the requirements of the University's regulations and code of practice for research degree programmes and that it has not been submitted for any other academic award. Except where indicated by specific reference in the text, the work is the candidate's own work. Work done in collaboration with or with the assistance of, others, is indicated as such. Any views expressed in the dissertation are those of the author.

Table of Contents

Publication

‘Structural study of the C-terminal domain of non-structural protein 1 from Japanese encephalitis virus’

POONSIRI, T., WRIGHT, G. S. A., DIAMOND, M. S., TURTLE, L., SOLOMON, T. & ANTONYUK, S. V. 2018. Structural study of the C-terminal domain of non-structural protein 1 from Japanese encephalitis virus. J Virol. 2018;92(7). doi: 10.1128/JVI.01868-17

See the full text copy in appendix 1.

List of Abbreviations

Abbreviations	Full term
22NS1	mAb against WNV NS1
β_2ME	β-mercaptoethanol
Ac64	autographa californica evelope protein gp64
AIMTB	autoinduction terrific broth media without trace elements
APS	ammonium persulphate
BBB	blood brain barrier
BCA	bicinchoninic acid
BLAST	local alignment search tool
C protein	capsid protein
C1S	complement component 1, s subcomponent
C4	complement component 4
C4b	complement component 4 isotype b
CAPS	3-(Cyclohexylamino)-1-propanesulfonic acid
CD8+	cytotoxic T lymphocyte
CDC	The Centers for Disease Control and Prevention
C-*JEV*NS1	JEV nonstructural protein 1 C-terminal domain
C-*JEV*NS1'	JEV nonstructural protein 1 prime C-terminal domain
CMOS	complementary metal-oxide-semiconductor
C-NS1	nonstructural protein 1 C-terminal domain
CNS	central nervous system

CNS	crystallography and NMR system
CSF	cerebrospinal fluid
C-*WNV*NS1	WNV nonstructural protein 1 C-terminal domain
Cx.	*Culex*
C-*ZIKV*NS1	ZIKV nonstructural protein 1 C-terminal domain
Da	dalton
DAXX	death-associated protein 6
DCs	dendritic cells
DDM	n-Dodecyl-β-D-Maltoside (detergent)
DHF	haemorrhagic fever
D_{max}	maximum intra-particle dimension
dsRNA	double stranded RNA
DSS	dengue shock syndrome
E protein	envelop protein
E. coli	*Escherichia coli*
ECL	enhanced chemiluminescence
EDTA	ethylenediaminetetraacetic acid
EM	electron microscopy
Eq.	equation
ER	endoplasmic reticulum
Fab	fragment antigen-binding
FDA	Food and Drug Administration
FGIN-1-27	N, N-di-n-hexyl 2-(4-fluorophenyl)indole-3-acetamide

fs	femtosecond
GAGs	glycosaminoglycans
GFP	green fluorescent protein
Gnd·HCl	guanidine hydrochloride
GPI	glycosylphosphatidylinositol
GST	glutathione S-transferase
HA	hemagglutinin
HBM	honeybee melittin
His6, HIS	hexahistidine tag
HMG-CoA	hydroxymethylglutaryl-coenzyme A
HPLC	high performance liquid chromatography
IFNs	interferons
Ig	immunoglobulin
IL	interleukin
IMAC	immobilized metal-ion affinity chromatography
IPTG	Isopropyl β-D-1-thiogalactopyranoside
IVIG	intravenous immunoglobulin
JAK	Janus kinases
JE	Japanese encephalitis
JEV	*Japanese encephalitis virus*
LB	Luria-Bertani
LDs	lipid droplets
M protein	membrane protein

mAb	monoclonal antibody
MALDI-Tof	matrix assisted laser desorption ionization-time of flight
MBP	maltose binding protein
MD	molecular dynamics
MES	2-(N-morpholino) ethanesulfonic acid
MHC-1	major histocompatibility complex class I
MPD	(+/-)-2-Methyl-2,4-pentanediol
MRI	magnetic resonance imaging
MW	molecular weight
MWCO	molecular weight cut-off
NCBI	the national center for biotechnology Information
NMR	nuclear magnetic resonance
NP 40	4-Nonylphenyl-polyethylene glycol,
NPCs	neuron progenitor stem cells
ns	nanosecond
NS	nonstructural protein 1
NTA	nitrilotriacetic acid
OD_{600}	optical density at 600 nm
OPPF-UK	Oxford protein production facility United Kingdom
PDB	protein data bank
PEG	polyethylene glycol
PKR	dsRNA-activated protein kinase
prM protein	pre-membrane protein

PRRs	pattern recognition receptors
ps	picosecond
R_g	radius of gyration or root mean square (rms) radius
RMSD	root-mean-square deviation
RPTPmu	receptor-like protein-tyrosine phosphatase
RT-PCR	reverse transcription polymerase chain reaction
S tag	pancreatic ribonueclease A
SAXS	Small angle X-ray scattering
SCH16	N-methyl isatin b--thiosemicarbazone derivative
SDS-PAGE	sodium dodecyl sulphate-polyacrylamide gel electrophoresis
SEC	size exclusion chromatography
siRNA	small-interfering RNA
spp.	two or more species of the genus indicated
STAT	signal transducer and activator of transcription
SUMO	small ubiquitin-like modifier
TBEV	*tick borne encephalitis virus*
TBS	Tris buffer saline (20 mM Tris-HCl pH 7.4, 150 mM NaCl)
TEMED	tetramethylethylenediamine
TEV	tobacco etch virus protease
TLR-3	toll-like receptor 3
TLS	Translation/Libration/Screw
TNF	tumour necrosis factor
TYK2	tyrosine Kinase 2

uPase	uridine phosphorylase
UTR	untranslated region
WHO	World Health Organization
WNV	*West Nile virus*
ZIKV	*Zika virus*
β_2ME	β-mercaptoethanol
χ	chi value
GSSH	oxidized glutathione
GSH	reduced glutathione
PMSF	phenylmethanesulfonyl fluoride
HEPES	4-(2-hydroxyethyl)-1-piperazineethanesulfonic acid
HEK 239T	human embryonic kidney cell line
DMEM	Dulbecco's modified Eagle medium
sf9	*Spodoptera frugiperda*
DTT	dithiothreitol
SOD3	superoxide dismutase 3
PC	dipalmitoyl-sn-glycero-3-phosphocholine
CHOL	cholesterol
FT	flow-through fraction
CCR4	CC chemokine receptor 4
NOT1	negative regulator of transcription 1
MS	mass-spectrometry
TFA	trifluoroacetic acid

RHB	ammonium bicarbonate
SNVs	single-nucleotide variants (SNVs)
SPG	succinic acid, sodium phosphate monobasic monohydrate
PCTP	sodium propionate, sodium cacodylate trihydrate, Bis-Tris propane

List of Figures

List of Tables

Chapter 1 Introduction to Japanese Encephalitis Virus

Japanese encephalitis virus (JEV) is a mosquito-transmitted member of the genus *Flavivirus*, family *Flaviviridae* that is closely related to other emerging viral pathogens including DENV, WNV, and ZIKV (**Figure 1.1**). JEV infection can result in meningitis and encephalitis, which in severe cases causes permanent brain damage or death. JEV occurs predominantly in rural areas throughout Southeast Asia, the Pacific islands and the Far East causing around 68,000 cases worldwide each year, despite extensive vaccination. The true number is probably much higher because many cases are still unreported. Moreover, only supportive treatment is available to cure flavivirus infection. To better manage flavivirus diseases, apart from vaccination, vector control, and personal protection, development of an anti-viral agent is needed.

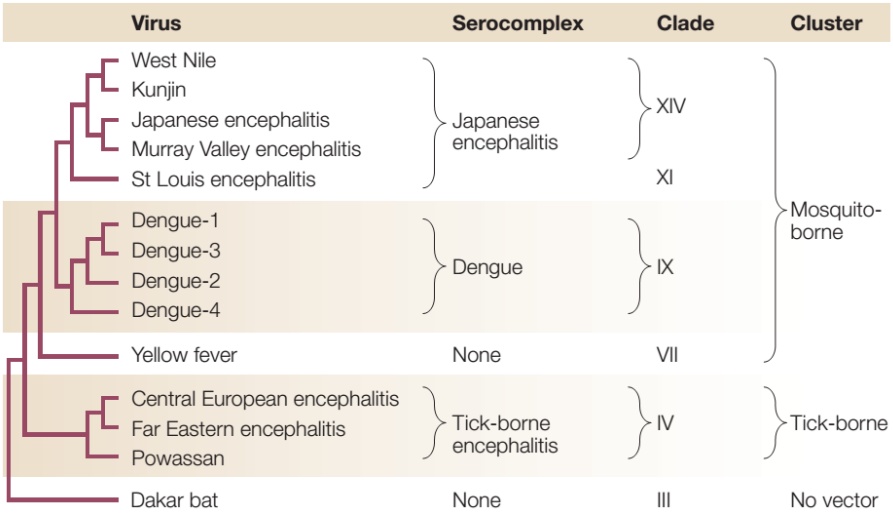


Figure 1.1 Flavivirus classification.

The phylogenetic tree was generated from NS5 genes. Evolutionary distance is not shown. (taken from (Mukhopadhyay et al., 2005). Licence permission no. 4236680904226).

1.1. Japanese encephalitis virus

1.1.1 Distribution and transmission cycle

JEV is believed to evolve from a flavivirus ancestor in the Indonesia-Malaysia area (Schuh et al., 2013, Solomon et al., 2003b). Based on the envelope protein gene (some studies use membrane-envelope or membrane-capsid protein genes), the virus is divided into 5 genotypes (Solomon et al., 2003b). Genotype IV and V are the oldest lineages and are retained in the Indonesia-Malaysia region, whereas the rest have spread wider. Genotype III and Ib (newly emerged) circulate in temperate Asia, while genotype II and Ia circulate in tropical Asia (Schuh et al., 2013). In the temperate zone, large JEV epidemics occur in the summer months but it is endemic year-round in the tropical zone. Japanese encephalitis (JE) was first described in Japan in the 1870s, but the first isolation was from the human brain in 1935 and from mosquitoes, *Culex tritaeniorhynchus,* in 1938 (Ishikawa and Konishi, 2015). In the first reported outbreaks from Japan, both adult and children were affected indicating no pre-existing immunity in the population and suggesting the virus had recently arrived in Japan (Solomon et al., 2003b). It seems that the distribution of JEV comes from Indonesia-Malaysia region to the north, west, and east. JEV was first detected in Saipan in 1990, Pakistan in 1992, Torres Strait Islands in 1995, and Australian mainland (Cape York) in 1998 (van den Hurk et al., 2006, Schuh et al., 2013, Knope et al., 2014). Now the risk area of JEV is throughout Asia (**Figure 1.2**).

In the past few decades, displacement of genotype III, which was dominant since the 1970s, with genotype I has been observed in China, India, Japan, South Korea, Taiwan, Thailand, and Vietnam (Han et al., 2014). However, the majority of human

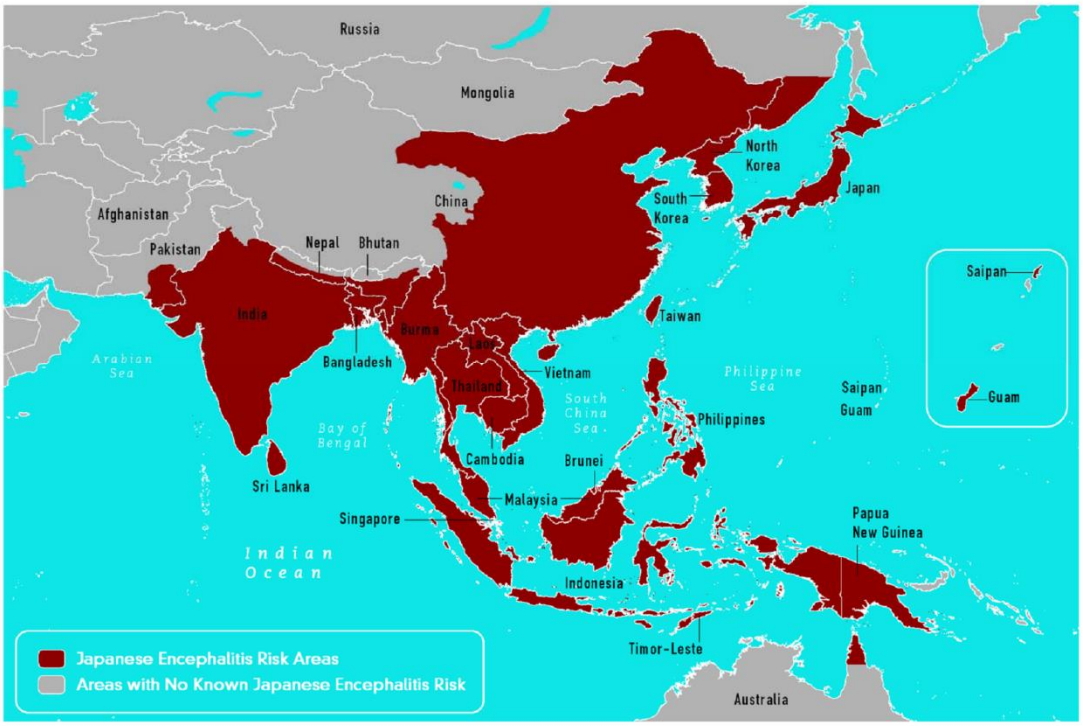


Figure 1.2 Geographic distribution of Japanese encephalitis virus
(The Centers for Disease Control and Prevention (CDC), 2015)

cases in this region are still genotype III. Replacement of genotype II with genotype I has been observed in Australia (van den Hurk et al., 2006). Moreover, genotype V, which had never been found beyond the Indonesia-Malaysia region, was more recently reported in China in 2009 and South Korea in 2010, both from mosquito samples (Li et al., 2011, Takhampunya et al., 2011). JEV circulation is dynamic and will possibly continue to spread to new areas, even though there has not been any reported JEV detection outside these areas since the arrival of JEV on the Australian mainland (CDC, 2015). The last human case in the Australian mainland was in 1998 (Knope et al., 2014) and a human case of genotype V in re-emerging areas have not been observed yet. The reasons why JEV cannot establish are unclear. It may be because the virus has not entered the zoonotic cycle in these areas. JEV probably has to compete with other flaviviruses that have already filled the ecological niche.

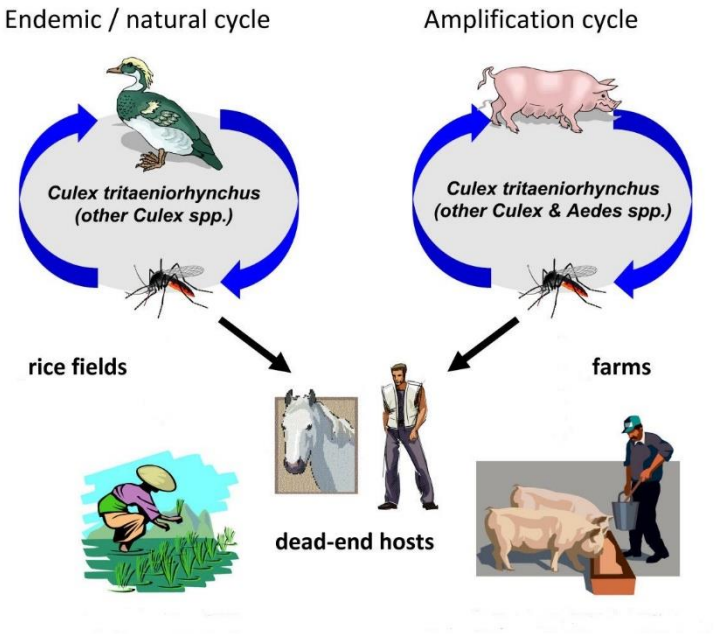


Figure 1.3 Japanese encephalitis virus transmission cycle
Modified from(Pfeffer and Dobler, 2010).

The natural transmission cycle of JEV involves the mosquito vector and wading birds (**Figure 1.3**). JEV was isolated from several species of mosquito including in genus *Culex, Aedes,* and *Anopheles*, but *Cx. tritaeniorhynchus* is the principal vector (Sucharit et al., 1989, van den Hurk et al., 2009). Detection of JEV or antibodies against it were demonstrated in more than 90 species of different avian families including chicken, pigeon, duck, sparrow and especially egret and heron, which have high viremia and recognized as the primary hosts. JEV was also detected in many mammal species such as dog, rat, goat, and cattle. Pig in particular, which has high and prolonged viremia, serve as an amplifying host. Disease symptoms in the animals are rare, but it can cause abortion in pigs (van den Hurk et al., 2009). Human and horse infections develop serious encephalitis, but both are accidental and dead-end hosts. Unlike in pigs or wading birds, JEV usually does not develop the high level of

viremia in human and horse which is therefore unlikely to result in transfer to mosquitoes.

Rice and irrigated crop fields are associated and contribute to the JEV natural transmission cycle. It is the place where wading birds meet *Culex* spp. mosquitoes that initiate transmission. Pig farming plays an important role in transmission to humans and animals nearby. Detection of JEV antibodies in pigs was found to be synchronous with a significantly increased number of human cases (Konno et al., 1966). The rapid spread of JEV may due to several reasons including a rising of the human population, increased farming, changing agricultural practices, migrating of reservoir hosts (birds), globalization, climate change, and wind patterns might explain dispersal to islands.

1.1.2 Virion structure and viral protein functions

JEV is a positive-sense single strand RNA virus. It has a genome of around 11 kb, which is translated into a polyprotein consisting of three structural proteins: capsid (C), membrane (prM/M), and envelope protein (E), and seven non-structural proteins (NS): NS1, NS2A, NS2B, NS3, NS4, NS4B, and NS5 (**Figure 1.4**). The E and M proteins are the main components inserting into the lipid bilayer to form the external shell of the virus (**Figure 1.5**) (Mukhopadhyay et al., 2005, Kuhn et al., 2002, Perera and Kuhn, 2008, Wang et al., 2017). Binding of viral E protein to cell's receptor plays a role in the host cell attachment and entry step, and fusion to the endosomal membrane to release nucleocapsid into the cytoplasm (**Figure 1.6**). JEV attachment factors and attachment receptors were identified in different cell types, for example, heparan sulphate proteoglycans, heat-shock protein 70, and laminin receptor. Flaviviruses use

different endocytic pathways to enter host cells depending on the cell type. Few candidates of JEV entry receptors have been suggested, however, the entry receptors still remain ambiguous. (Liu et al., 2017, Nain et al., 2016, Yamauchi and Helenius, 2013) The nucleocapsid is composed of multiple copies of the capsid protein enclosing the single-stranded RNA genome. Dissociation of the capsid releases the viral RNA, which serves as mRNA for viral protein translation using the host cell machinery. The signal sequence in the genome translocates E, prM and NS1 proteins to the endoplasmic reticulum (ER) lumen while C, NS3, and NS5 remain in the cytoplasm (**Figure 1.4**a). The NS2A/B and NS4A/B are transmembrane proteins. The polyprotein is cleaved into individual proteins by both host and virus enzymes (NS2B/NS3 harbours serine protease activity). The newly synthesized NS proteins form the replication complex in the vesicle packets generated from ER (Welsch et al., 2009). RNA synthesis involves the helicase domain of NS3 and RNA-dependent RNA polymerase domain of NS5, which also bears the methyltransferase domain at N-terminus. Once the viral RNA is synthesized, it is enclosed in the capsid and buds into the ER to obtain the host cell lipid bilayer together with the E and prM proteins (**Figure 1.6**). Cleaving of prM to M protein in Golgi body by the enzyme furin produces the mature virion ready to be released from the cell and start a new infection cycle. However, the molecular details of each step such as dissociation and assembly of capsid and RNA replication are still unclear.

DENV is the most well-studied flavivirus. All of the proteins coded by its genome have been structurally characterized except for NS2A and NS4A/B. Although NS2B is a small transmembrane protein, the structure was determined by co-crystallization with its co-factor, NS3.

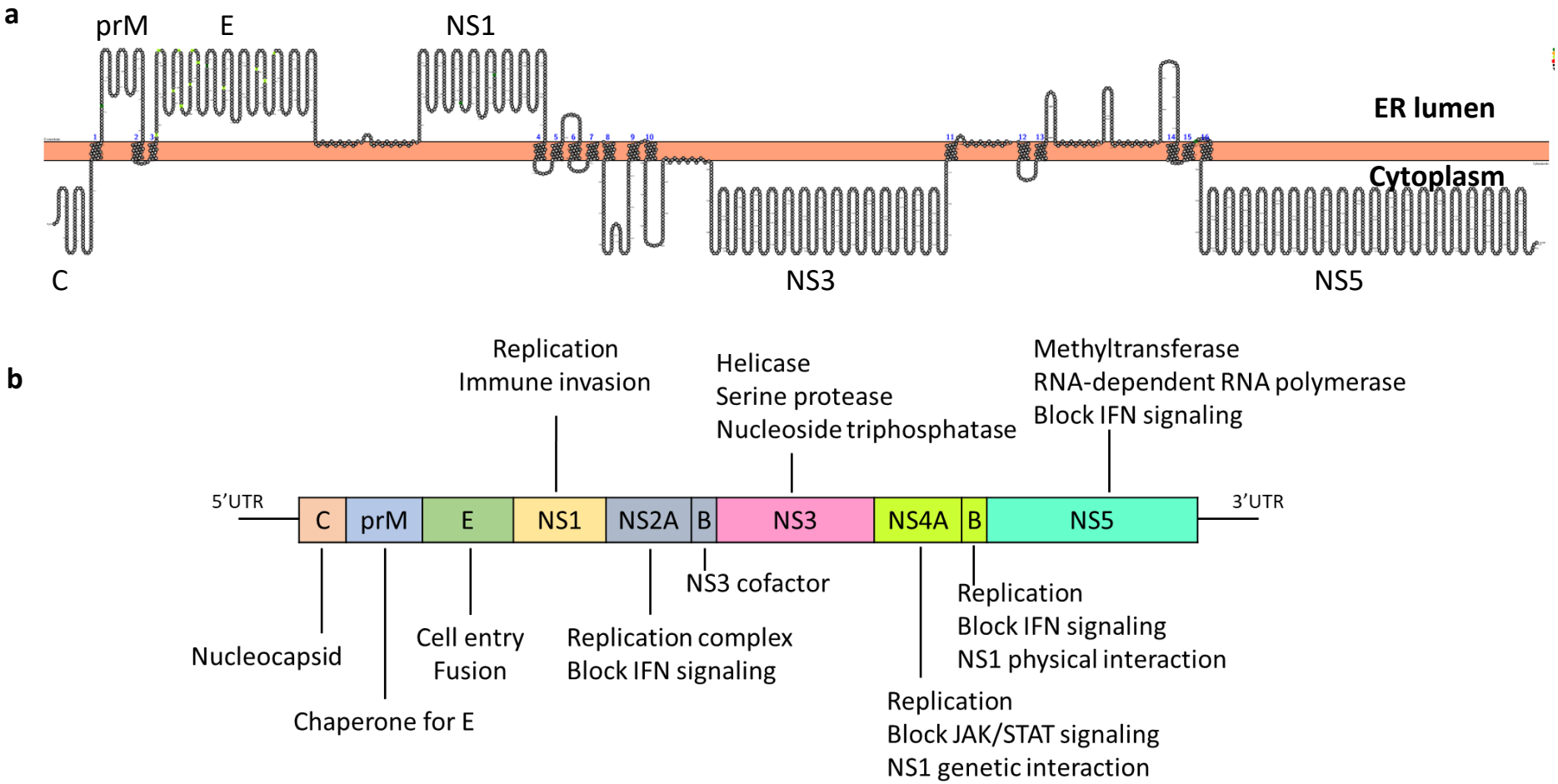


Figure 1.4 JEV polyprotein.

(a) Protein topology diagram created by PROTTER (OMASITS ET AL., 2014) **(b)** Flavivirus genome organization and protein functions.

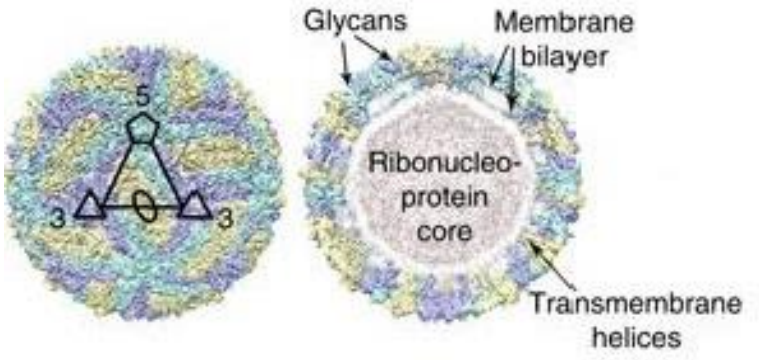


Figure 1.5 JEV cryo-EM structure reconstruction

Taken from (Wang et al., 2017)

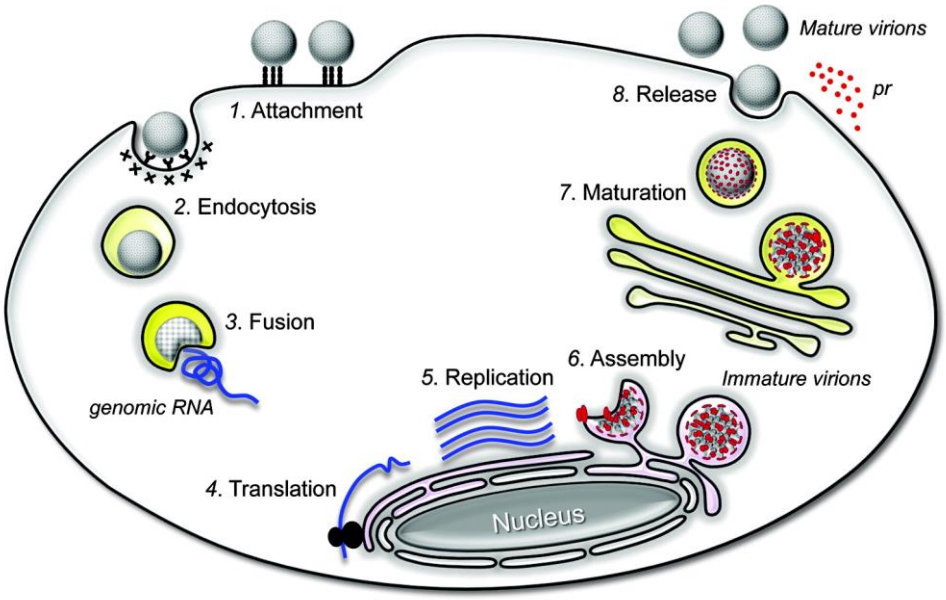


Figure 1.6 JEV replication cycle.

Taken from (Yun and Lee, 2014)

(1) Binding of virus particles to the attachment and cell entry receptors locates viruses on the cell membrane and initiates endocytic pathways. **(2-3)** Low pH in endosome causes the E protein to change its conformation and consequently triggers the membrane fusion step to release the viral genomic RNA, **(4)** which is translated and processed into mature viral proteins. **(5)** Replication of the viral RNA occurs inside the ER-derived vesicle. **(6)** The viral genomic RNA enclosed in neuclocapsid buds into the ER lumen to obtain the lipid bilayers together with the prM and E proteins. These immature virions are transported to the Golgi secretory pathway. **(7)** Cleaving of prM to M protein in Golgi body by the enzyme furin produces the mature virions. **(8)** Finally, the mature virions are released by exocytosis from the cell.

Capsid protein

Capsid protein is not only involved with nucleocapsid dissociation and assembly. Many studies have revealed the crucial functions of the capsid protein in the virus life cycle (Oliveira et al., 2017, Byk and Gamarnik, 2016). RNA secondary structure of 5' DENV capsid coding region has been found to function as the RNA cis-regulatory element involved in both RNA synthesis and cyclization (Clyde et al., 2008, Liu et al., 2013, Villordo and Gamarnik, 2009). Mutation of N-terminal residues also demonstrated impairment in virus particle formation (Samsa et al., 2012). The capsid proteins of DENV and WNV were shown to act as RNA chaperones (Pong et al., 2011, Ivanyi-Nagy et al., 2008), and the RNA binding sites were mapped to positively charged residues at the WNV C protein N and C terminus (Khromykh and Westaway, 1996). The capsid–prM junction is processed by NS2B/NS3 protease. Failure to cleave and produce mature C protein leads to a defective in viral particle assembly and release (Amberg and Rice, 1999). The capsid protein is found localized in the cytoplasm of the infected cell in ER and lipid droplets (LDs) and also in nucleoli in the nucleus (Wang et al., 2002, Bulich and Aaskov, 1992). The rationale of these subcellular distributions is unclear. Interaction with LDs was suggested to be a viral replication regulation process (Samsa et al., 2009) and it may affect cell lipid metabolism (Byk and Gamarnik, 2016). Mutagenesis studies identified a hydrophobic patch in the helix 2 region to play an important role in viral replication (Schlick et al., 2009). The function was recovered by reverting the mutation to restore the hydrophobic property of the protein indicating the importance of the residues properties rather than the residue position (Byk and Gamarnik, 2016, Schlick et al., 2009). While small mutation based deletions affect the protein function greatly, the

protein can tolerate the deletion of up to 16 residues in tick-borne encephalitis virus (TBEV) (Kofler et al., 2002) and 36 residues in WNV (Schlick et al., 2010). However, attenuation was observed in those large deletion viruses (Kofler et al., 2002, Schlick et al., 2010). Moreover, the capsid protein was reported to interact with several host proteins, for example, Death-associated protein 6 (DAXX) which may induce apoptosis and core histones that may mediate transcription through nucleosome disruption (Byk and Gamarnik, 2016).

The 3D structure of DENV (Ma et al., 2004) and WNV (Dokland et al., 2004) have been solved and show similar protein folds with the sequence identity of 36 %. (**Figure 1.7** and **Figure 1.8**) The monomer is composed of 4 α helices. Both studies demonstrated the dimeric state of the capsid. Helix 2 and 4 are an antiparallel pair together with positively charged residues at the helix 4 C-terminus. The basic rich region was suggested to interact with the viral RNA, while the opposite side of the dimer contains an apolar cleft created by the α2-α2' region, which may bind to lipid bilayer or LDs. Helix 1 shows a distinct conformation between DENV and WNV indicating its flexibility. The tetrameric form was observed only in WNV capsid structure and the capsid protein assembly to form nucleocapsid is proposed to be poorly ordered, different from the icosahedral external shell (Kuhn et al., 2002).

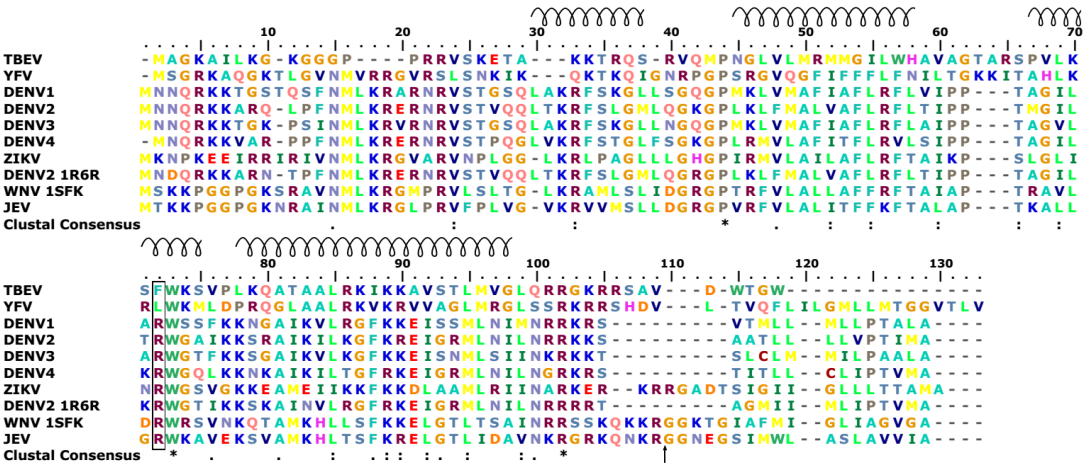


Figure 1.7 Multiple sequence alignment of flavivirus capsid protein

The alignment was produced from Clustral Omega (Sievers et al., 2011). An asterisk indicates fully conserved residue. A colon indicates conservation between groups of strongly similar properties. A period indicates conservation between groups of weakly similar properties. Spiral above sequence indicates the α helix positions.

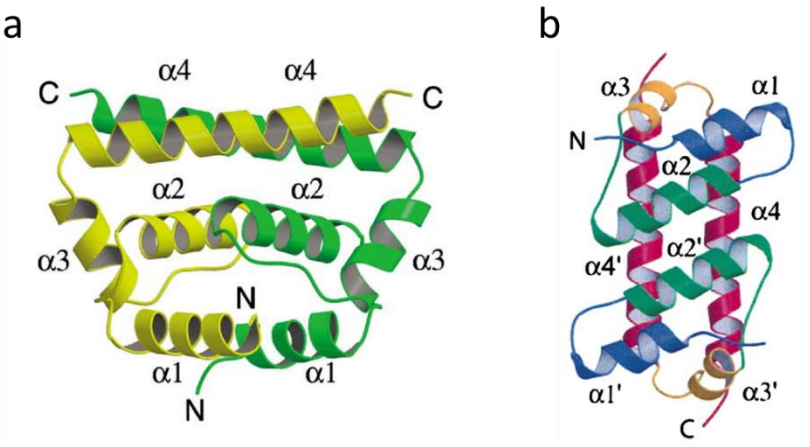


Figure 1.8 Capsid protein ribbon diagrams

(a) WNV capsid diagram protein taken from (Dokland et al., 2004) **(b)** DENV capsid protein diagram taken from (Ma et al., 2004)

NS1 and NS1' proteins

NS1 is a multifunctional glycoprotein that has drawn significant attention because of its importance in viral replication, immune modulation, and immune evasion. Mutagenesis and trans-complementation assays have established that NS1 is essential for RNA replication (Lindenbach and Rice, 1997, Mackenzie et al., 1996, Muylaert et al., 1996, Youn et al., 2013, Fan et al., 2014). It co-localizes with the replication complex (Mackenzie et al., 1996), and interacts genetically and physically with NS4A (Lindenbach and Rice, 1999) and NS4B (Youn et al., 2012), respectively. DENV NS1 was previously described as a complement fixing antigen found in patient serum and binds to complement pathway components: C1s, C4, C4b (Avirutnan et al., 2011, Avirutnan et al., 2010). WNV NS1 can also interact with factor H, complement control protein, which protects cells from complement-dependent clearance (Chung et al., 2006a). In addition, NS1 has been suggested to interfere with the dsRNA sensors. NS1 inhibits Toll-like receptor 3 (TLR-3) (Morrison and Scholle, 2014, Wilson et al., 2008) to escape host detection mechanisms.

Secreted NS1 is used as a diagnostic marker for infection as it is found in the blood at an early stage, typically before antibodies appear (Li et al., 2012, Amorim et al., 2014). Alternatively, detection of anti-NS1 antibodies, IgM and IgG, can also be used (Amorim et al., 2014, Solomon et al., 1998, Chao et al., 2015). NS-antibody complexes can trigger several inflammatory processes involving both innate and adaptive immune responses (Muller and Young, 2013, Amorim et al., 2014). Immunization of NS1 in mice or passively acquiring anti-NS1 antibodies can confer protective effects against challenge with yellow fever virus or WNV (Chung et al., 2006b, Schlesinger et al., 1986, Schlesinger et al., 1993). However, in DENV, anti-NS1 antibodies may have

auto-reactivity and reportedly bind to host extracellular matrix components, platelets, and endothelial cells (Muller and Young, 2013, Krishna et al., 2009). This cross-reactivity can induce damage to the host cells that enhance disease impact.

A signal sequence at the C-terminus of E protein translocates NS1 to the endoplasmic reticulum (ER) where it undergoes cleavage and posttranslational modification (Muller and Young, 2013). Structures of WNV (Edeling et al., 2014, Akey et al., 2014), DENV (Akey et al., 2014), and ZIKV NS1 (Xu et al., 2016, Brown et al., 2016) full-length proteins have been reported. NS1 proteins contain 6 conserved disulphide bonds and conserved N-linked glycosylation sites at Asn130 and Asn207. The JE serogroup (**Figure 1.1**) NS1 proteins have an additional glycosylation site at Asn175 linked to high-mannose carbohydrate, but this is not present in JEV NS1 itself (Muller and Young, 2013, Blitvich et al., 2001, Mandl et al., 1989). The NS1 monomer contains 3 domains: β-roll (amino acid residues 1-29), wing (38-151), and β-ladder domains (181-352) (Akey et al., 2014, Xu et al., 2016) (**Figure 1.9**). There are two characterized forms of NS1: a membrane-associated dimer (~49 kDa per monomer), found on the cell and ER surface, and a secreted hexamer (52- 55 kDa per monomer) (Muller and Young, 2013). The mass of the two NS1 forms is different due to differential glycosylation. NS1 forms a homodimer by extending the β-ladder domain and forms a cross shape (**Figure 1.9**). The hydrophobic surface of the β-roll and wing domains may mediate the interaction with the cell membrane via a number of amino acid residues including 28, 115, 118, 123, and 160-163 (Xu et al., 2016, Brown et al., 2016) (**Figure 1.9**). The opposite side to the β-barrel is composed of loops linking the surface β-strands of the ladder domain. This region is a potential protein binding surface to host immune components due to its hydrophilicity and glycosylation. NS1 dimer

attachment on the infected cell surface may occur via a glycosylphosphatidylinositol (GPI) anchor, for which a hydrophobic carboxy-terminal GPI-addition signal peptide at the N-terminus of NS2A is required (Jacobs et al., 2000, Noisakran et al., 2008, Noisakran et al., 2007). NS1 also binds to uninfected cell membranes via glycosaminoglycans (GAGs) binding, primarily heparan sulphate and chondroitin sulphate E (Avirutnan et al., 2007). Three NS1 dimers can assemble to form a hexameric pore, which lipid is found inside (PDB ID 4O6B for DENV2 NS1 and 4O6C for WNV NS1) (Akey et al., 2014).

Most knowledge of JEV NS1 functions has been inferred from studies of DENV and WNV NS1, but how it accomplishes those functions is not exactly known. Although the protein sequences are highly conserved (**Figure 1.10**) and the DENV, WNV, and ZIKV NS1 structures display the same protein fold, there are structural and non-structural evidence indicate the important differences. For example, structures of flavivirus NS1 demonstrate distinct surface charge especially DENV (Song et al., 2016). Polyclonal antibodies raised against DENV NS1 in mice were shown to cross-react with protein on the epithelial cell. The cross-reactivity epitope was mapped to amino acid residues 311-330 on NS1 (Cheng et al., 2009) (**Figure 1.10**). Although JEV NS1 shares these conserved epitopes, antibodies against JEV NS1 showed no reactivity to any of these host cell targets. As another example, WNV NS1 bound the alternative complement pathway regulator, factor H, whereas JEV NS1 did not (Krishna et al., 2009).

NS1' is specific only to JE serogroup viruses. NS1' is longer than NS1 by 52 extra amino acids at the NS2A N-terminus due to the formation of a pseudoknot by the conserved

sequences at the NS2A 5’ end. It pauses translation and shifts the reading frame back 1 base, also called a classical -1 ribosomal frameshift motif (Melian et al., 2010). The frame-shift motif creates an alternative stop codon in NS2A sequence and terminates protein translation. NS1’ was found in dimeric form (monomer molecular mass around 58 kDa) and detected in both cell lysate and culture media (Mason, 1989, Young et al., 2013). NS1’ was suggested to play a role in neurovirulence and neuroinvasiveness as lack of function reduces mortality in mice (Melian et al., 2010, Ye et al., 2012). NS1’ was found co-localized with NS1 and RNA replication complex, and could substitute for NS1 in virus replication (Young et al., 2013). However, there is a discrepancy between the results of *in vitro* and *in vivo* studies. WNV NS1’ provides an advantage only in *in vivo* studies (Melian et al., 2014). There is also variation of NS1’ involvement in replication among different viruses. Whereas WNV NS1’ does not contribute to viral replication *in vitro*, JEV NS1’ mutants have less infectivity in a cell model (Melian et al., 2014, Takamatsu et al., 2014). Therefore, the role of NS1’ in the JEV life cycle and pathogenesis remains unclear.

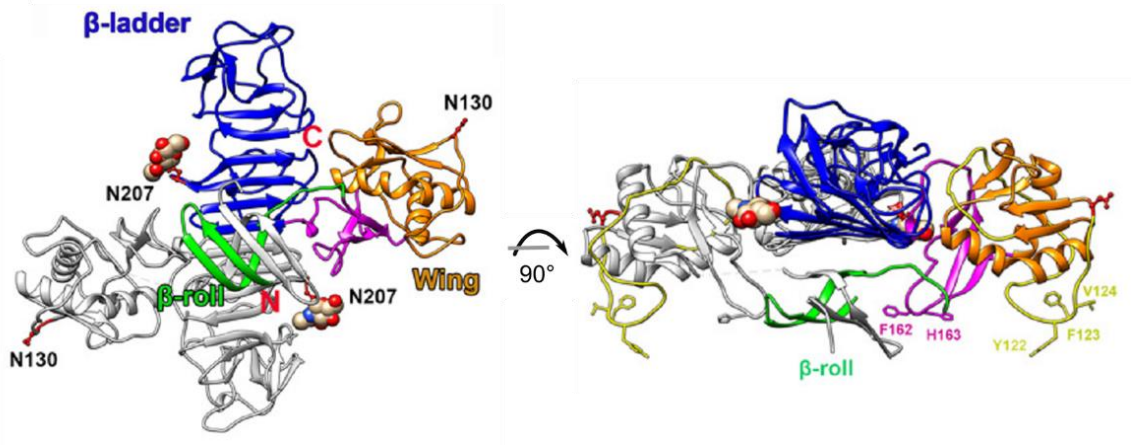


Figure 1.9 ZIKV NS1 full length

Taken from (Xu et al., 2016) (**a**) Cross-shaped homodimer NS1. One subunit is coloured in grey and another is coloured by domain. β-roll domain is coloured in green, wing domain in orange, and β-ladder domains in blue. (**b**) Side view of NS1 dimer. Connecter subdomain (close to β-roll) is indicated in magenta. The intertwine loop of wing domain shown in yellow is disordered and not visible in DENV (PDB ID 4O6B) and WNV (4O6C), but they are visible in ZIKV (PDB IDs 5GS6 and 5K6K).

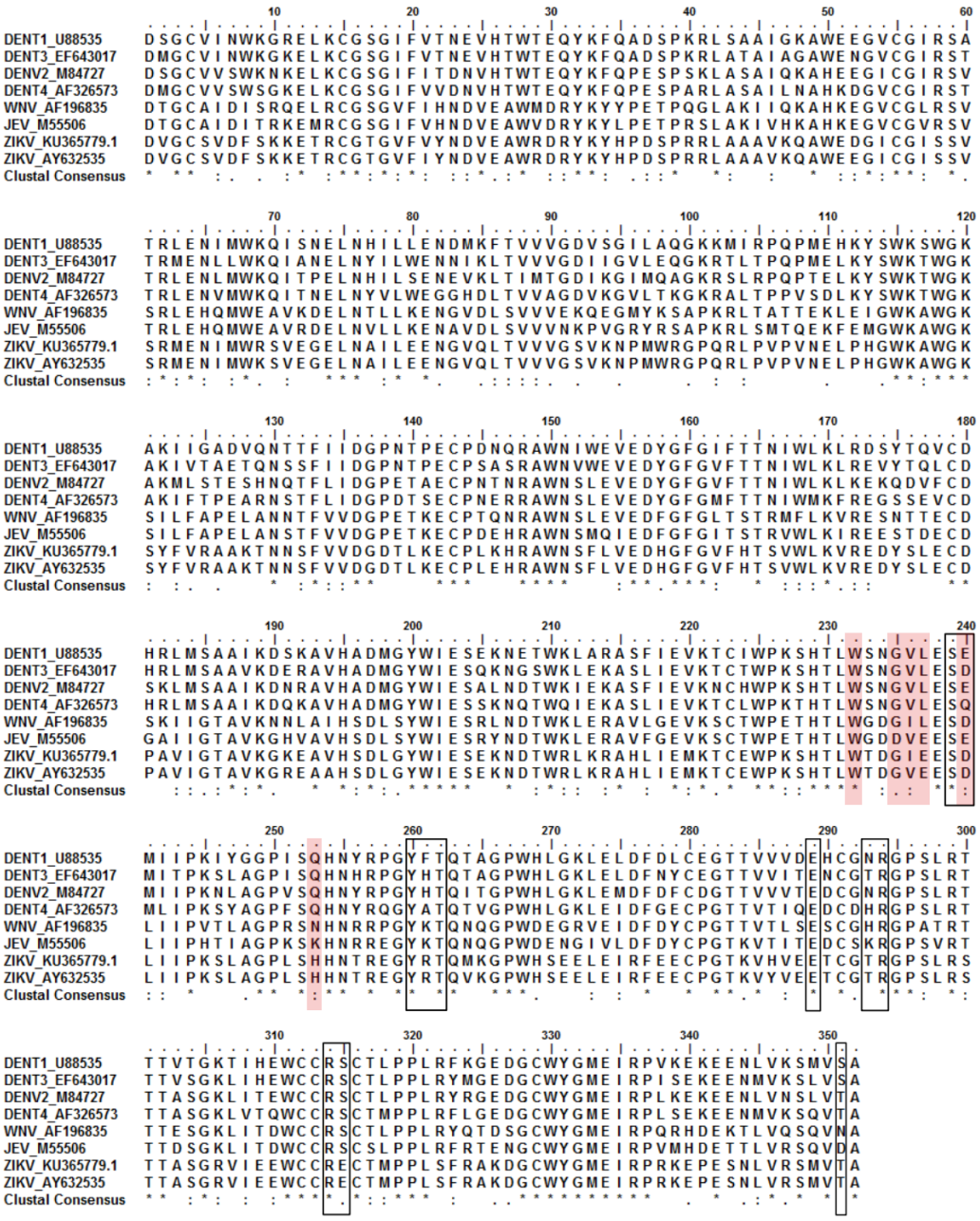


Figure 1.10 Multiple sequence alignment of flavivirus NS1 full length proteins

The alignment was produced by ClustalW (Larkin et al., 2007). An asterisk indicates fully conserved residue. A colon indicates conservation between groups of strongly similar properties. A period indicates conservation between groups of weakly similar properties. The 22NS1 light chain epitopes are highlighted in red and heavy chain epitopes are in black squares The amino acid sequences used are related to the X-ray crystal structure studies: DENV1 U88535 for PDB ID 4OIG, DENV2 M84727 for 4O6B, WNV 196835 for 4O6C and 4OIE, ZIKV KU365779 for 5IY3, and ZIKV AY632535 for 5K6K and 5GS6.

1.1.3 Disease symptoms and pathogenesis

Human JEV infection is usually asymptomatic. The ratio of unapparent and apparent JEV infections was less than 1% in endemic area (Scheld et al., 2008). The ratio in non-endemic areas was reported higher (Ishikawa and Konishi, 2015). JE occurs primarily in children <15 years old and adult usually has pre-existing immunity. However, when the virus arrives in new areas it can affect those both adults and children (Scheld et al., 2008, Solomon et al., 2003b). In addition, in childhood JE vaccinated areas, the number of cases could be similar between the two age groups. The death rate is 20-30% and 30-50% of the cases survive with permanent brain damage. After the 5-15 days of the incubation period, JEV infection begins with common cold-like febrile illness and may include shivering and diarrhoea for a few days. These could be the only manifestations or in severe cases, acute encephalitis syndrome can develop. Headache, vomiting, a reduction of consciousness, and seizure are observed. Other manifestations resulted from different parts of central nervous system (CNS) damage are observed, for example, flaccid paralysis, meningitis, Parkinsonian syndrome with mask-like face, and movement disorders. Enlargement of spleen and liver have been reported. Heart and lung complications have also been observed (Unni et al., 2011).

JEV is introduced to humans by a mosquito bite. First inoculation may occur in blood circulation or skin. However, the virus load in blood circulation is usually low and the early transmission events within the skin are not well described. DENV infected human skin explant tissue experiments demonstrated detection of DENV in the epidermis layer particularly in the basal epidermis cells, but not in the dermis, together with NS1 protein detection in the epidermis indicating viral replication

activity (Limon-Flores et al., 2005). This may be applicable to JEV, but no experiment has confirmed it. In addition, human dendritic cells and Langerhans cells were very permissive for DENV infection (Wu et al., 2000). In mouse models, JEV was detected in peritoneal macrophages early after JEV peritoneal inoculation (Mathur et al., 1988). Later, the virus was detected in splenic macrophages, splenic T-lymphocytes (but not B-lymphocyte). It is probable that the replication might occur in the epidermis and infected tissue macrophages migrate to the local lymph nodes. Virus clearance at this stage by an immune response is very important. The first line defence is committed by the pattern recognition receptors (PRRs). Detecting viral RNA and proteins by PRRs are the upstream regulation of inflammatory respond genes transcription, for example, gene codes for type I interferons (IFNs). IFNs are signalling molecules secreted by host cells to initiate the defence mechanism, which is a principal defence against viral infection. Many of JEV immune invasion strategies have been demonstrated. JEV interferes with the pattern recognition viral dsRNA-sensing that activates IFNs production by delaying the exposure of dsRNA in porcine kidney cells (Espada-Murao and Morita, 2011) and JEV NS1 was shown to inhibit TLR-3 (Wilson et al., 2008). In the downstream cascade of dsRNA IFNs activation, short fragmented RNAs of the 3'UTR blocks the phosphorylation of IFN regulatory factor-3 (Chang et al., 2013) and the JEV NS2A was observed to antagonize the IFN-induced dsRNA-activated protein kinase (PKR) (Tu et al., 2012). IFNs signalling regulates the expression of several antiviral genes via JAK/STAT signal transduction pathway. Inhibition of STAT1 and STAT2 phosphorylation by JEV NS4A was reported (Lin et al., 2008) and the JEV NS5 also blocks the phosphorylation of TYK2 and STAT1 (Lin et al., 2006). Moreover, a few different events were observed in JEV infected cells, including

depletion of cell maturation surface markers in dendritic cells (DCs), impaired DCs maturation (Cao et al., 2011) and reduction in the ability to present antigen through major histocompatibility complex class I (MHC-1) pathway to suppress the activation of T cells (Aleyas et al., 2010). T lymphocytes and neutralizing antibodies (IgM and IgG) play a major role in the adaptive immune response. Virus infected cells are eliminated by $CD8^+$ cytotoxic T cells, which may occur via recognition of NS3 protein. The cell free viruses are neutralized by antiviral antibodies, which E protein is the main target (Unni et al., 2011). If the virus titre is high enough, viruses can enter the bloodstream and progress to the CNS infection. JEV reaches the brain via the blood but how the virus crosses the blood brain barrier (BBB) is not clear. The JEV antigen was detected in the vascular endothelium, but this has not proved the permissiveness of the cells. Astrocyte is a component of the BBB and was shown to be infected by JEV (Suri and Banerjee, 1995). Infection of astrocyte and/or vascular endothelial may bring the virus into the CNS. Rupture of the BBB was also suggested, but this tends to occur after encephalitis. Thus, it might not be the main reason for CNS introduction. Infiltration of infected white blood cells may also possible (Trojan horse) (Sapkal et al., 2007). Once the virus enters the brain, neurons are the principal target and ER stress in the cell leads to apoptosis (Su et al., 2002). Neuronal death then activates astrocytes and microglia which release proinflammatory cytokines which contribute to infiltration of monocytes and leukocytes. Moreover, the release of reactive oxygen species, nitric oxide, and other cytokines like tumour necrosis factor α (TNFα) causes neuron apoptosis (Ghosh and Basu, 2009). Mouse microglial cells were highly permissive to JEV along with a long persistence of the virus suggesting that microglia may serve as a viral reservoir (Thongtan et al., 2010). In conclusion, neuronal death

is directly caused by JEV infection itself and indirect massive immune responses. The later may be the result of gradually reducing anti-inflammatory cytokine IL-10 after JEV infection (Swarup et al., 2007). JEV infection of neuron progenitor stem cells (NPCs) was reported, which could prevent NPC proliferation and consequently impair the neuron recovery. This may relate to permanent brain damage in JE survivor (Das and Basu, 2008).

1.1.4 Diagnosis and treatments

The early stages of JE are difficult to differentiate from other febrile illness and even with neurologic symptoms, JE is still not easily defined from other encephalitis diseases (Impoinvil et al., 2013). Neutrophilia is commonly found in peripheral blood tests. Laboratory testing is typically performed by using serum or cerebrospinal fluid (CSF) to detect JEV IgM specific antibodies. The test might need to be repeated as the initial test may be negative. IgM positive samples can be confirmed with neutralizing antibody testing if needed (CDC, 2015). Detection of the virus nucleic acid is sometimes also attempted. JEV nucleic acid is usually rarely detected by reverse transcription polymerase chain reaction (RT-PCR) (Sapkal et al., 2007) and also not detected in urine (Zhao et al., 2013). However, prolong detection of JEV in urine and whole blood was reported in a recent JE fatal case (Huang et al., 2017). Isolation of JEV from blood clots (white blood cells) was developed to improve the diagnosis, but a low genomic copy is still a limiting factor (Sapkal et al., 2007). Magnetic resonance imaging (MRI) of brain lesion can also help to describe flavivirus encephalitis (Kalita et al., 2016) but it should be considered with other evidence.

There is no anti-viral treatment for JE available. JE can cause several deadly complications, which only be treated by supportive care. The increase of intracranial pressure is related to seizures. Fluids or steroid may be given to maintaining body fluid balance, and sedation could help relieve the symptom. In unconscious cases, basic life support such as respiratory system and blood circulation support is required. Patients who recovered but with permanent brain damage still need an appropriate nursing care from their family.

1.1.5 Vaccines

Vaccination is a very effective way to prevent JE in humans. However, it is difficult to completely eliminate JEV, because humans are dead-end hosts and the virus is maintained in its natural cycle. Vaccines for domestic pigs are also available, but the high turnover in pig populations is a major drawback (Solomon, 2006).

Following the discovery of JEV, the first mouse-brain-derived inactivated vaccine was developed and it was used worldwide (**Table 1.1**). Later, it was replaced by Vero-cell-derived inactivated vaccines as recommended by World Health Organization (WHO) in 2006 and 2015 due to the risk of adverse effect from mouse brain derived substances and animal ethical issues. The primary hamster kidney-cell-derived inactivated vaccine was previously used, but later was also replaced by the Vero-cell-derived inactivated vaccine and a live-attenuated vaccine. The live-attenuated vaccine, also called the second-generation vaccine, has been used previously in China (from 1988). Now it is approved in many Asian countries and prequalified to use in children by WHO. Another current vaccine is a live-attenuated chimeric vaccine. The vaccine is based on yellow fever virus strain 17D inserted with JEV prM and E genes

(Ishikawa and Konishi, 2015). The vaccines are administered by intramuscular injection.

Table 1.1 List of vaccines against JEV

No.	Vaccine	JEV strain
1	mouse-brain-derived inactivated vaccine	Nagayama
2	primary hamster kidney-cell-derived inactivated vaccine	Beijing
3	Vero-cell-derived inactivated vaccines	SA14-14-2/ Beijing
4	live-attenuated vaccine	SA14-14-2
5	live-attenuated chimeric vaccine	SA14-14-2

1.1.6 Drug developments

The antiviral drug is aimed to inhibit any virus activities that facilitate replication. Every single step in the virus life cycle could be a target including the viral immune invasion mechanism. Understanding of disease pathogenesis and molecular mechanism of virus life cycle could inform on how to prevent virus replication. As mentioned above, part of the disease outcome is thought to be due to the immune response, so many therapeutic candidates are antagonistic molecules such as antioxidants that reduce reactive oxygen species production (**Table 1.2**). Some of them are immune modulators, for example, IFNs, which is a common antiviral response that is blocked by the virus. Direct RNA intervention also shows anti-viral effect. E, NS3, and NS5 are well described in structure and function. They are the main targets for drug development. Several NS3 protease inhibitors for hepatitis C virus have been recently approved by the Food and Drug Administration (FDA) for clinical use (Salam and Akimitsu, 2013). Straightforwardly, targeting E protein aims to

block virus entry, while NS3 and NS5 interfere with RNA replication. Moreover, the virus hijacks several parts of the host machinery to survive. Targeting those necessary host proteins, for example, glucosidase and hydroxymethylglutaryl-coenzyme A (HMG-CoA) reductase (cholesterol synthesis), demonstrates protection *in vitro* and *in vivo*, respectively (**Table 1.2**). However, none of the anti-viral agent candidates have succeeded at the clinical level. In many of the *in vitro* and *in vivo* studies, candidates were given before, simultaneously, or immediately after inoculation such as Lovastatin that offered success only prior to infection not after, which are not practical as patients are treated after diagnosed (usually after disease onset). In theory, products to apply on the skin that contains the candidates might interfere with replication cycle in the skin cells and helps to block the virus to enter lymph and blood circulation. The brain is a particular target for JEV, so measurement that delivers drugs into CNS is very important. Moreover, a combination of the candidates may create a synergistic effect that is better than given on its own.

.

Table 1.2 Anti-Japanese encephalitis drugs reviewed

Adapted from (Ishikawa and Konishi, 2015)

Category	Antiviral drugs	Target/mechanism	Experiment	Drug administration
Nonspecific broad spectrum	IFN	IFN-stimulated gene	Human clinical trial	After onset/ diagnosis confirm for 7 days (Solomon et al., 2003a)
	Ribavirin	Inosine monophosphate dehydrogenase, etc.	Human clinical trial	After onset/ diagnosis confirm for 7 days (Kumar et al., 2009)
	Minocycline	Antioxidant	*In vivo*	1 dpi/ after onset
	Arctigenin	Antioxidant	*In vivo*	1 dpi
	Fenofibrate	Antioxidant	*In vivo*	4 dbi
	Curcumin	Antioxidant	*In vitro*	Immediately after infection
	Pentoxifylline	Assembly or Release	*In vivo*	Immediately after infection
	Nitazoxanide	Early/mid-replication cycle	*In vivo*	1 dpi
Nucleic acid-based	siRNA	C, M, E, NS1, NS3, NS4B, NS5	*In vivo*	Simultaneous or before infection directly to the viral replicating cells.
	Peptide nucleic acid	UTR	*In vitro*	
	Morpholino oligomer	UTR	*In vivo*	Before or immediately after infection
Virus replication cycle-based	Heparan sulphate	Receptor binding	*In vivo*	Before infection

Category	Antiviral drugs	Target/mechanism	Experiment	Drug administration
	E-binding peptide	Receptor binding	*In vitro*	Incubated with virus before infection
	Nanoscale silicate platelet	Attachment	*In vivo*	Immediately after infection
	Indirubin	Attachment	*In vivo*	Immediately after infection intra cranial
	Bovine lactoferrin	Receptor binding	*In vitro*	Before infection
	Griffithsin	Receptor binding	*In vivo*	Before infection
	Recombinant E	Receptor binding	*In vivo*	Before infection
	MCPIP1	RNA replication	*In vitro*	Before infection
	Kaempferol	RNA replication	*In vitro*	Before and after infection
	FGIN-1-27	RNA replication	*In vitro*	Before infection
	Pokeweed	RNA replication	*In vivo*	Before and after infection
	SCH16	Translation	*In vivo*	Before and after infection
In silico modelling-based	Ivermectin	NS3	*In vitro*	(helicase inhibition assay)
	4-hydroxypanduratin A	NS2B/NS3	-	-
Intravenous immunoglobulin (IVIG)	intravenous immunoglobulin (IVIG)	Neutralizing antibodies, anti-inflammatory	Human clinical trial	Day 6 of hospitalization
Other potential drug targets (other flaviviruses)	NITD-451	Translation	*In vivo*	Immediately after challenging

Category	Antiviral drugs	Target/mechanism	Experiment	Drug administration
	Lovastatin	Cholesterol synthesis, Isoprenoid synthesis, Dolichol synthesis	*In vivo*	Before infection
	Iminosugar derivative	Inhibit glucosidase (glycosylation)	*In vitro*	Before infection
	Celgosivir	α-glucosidase inhibitor, interfere N-glycosylation	Human clinical trial	>48 hr after diagnosis
	ST-148	Inhibit capsid assembly and disassembly	*In vivo*	Immediately after challenging

Note: Dpi = days post infection, MCPIP1 = monocyte chemoattractant protein 1-induced protein 1, siRNA = small-interfering RNA, UTR = untranslated region, FGIN-1-27 = N, N-di-n-hexyl 2-(4-fluorophenyl)indole-3-acetamide, SCH16 = N-methylisatin β-thiosemicarbazone derivative

1.2. Dengue virus

DENV infection presents in the tropical and subtropical region including more than 100 countries (**Figure 1.11**) (Guzman and Harris, 2015) and affects infant, children, and adult. The estimated number is around 390 million infections, of which 96 million develop disease symptoms (Bhatt et al., 2013). The clinical features of dengue fever are flu-like symptoms but may include nausea, rash, and swollen lymph glands. In some cases, the disease may progress to dengue haemorrhagic fever (DHF) and dengue shock syndrome (DSS) which are deadly. Severe dengue infection causes by plasma leakage leading to severe bleeding, fluid imbalance, circulatory failure, and organ failure. Dengue viruses are transmitted by *Aedes* mosquitoes, primarily *Aedes aegypti*. DENV is originally a monkey virus circulated in sylvatic (primate) transmission cycle (Holmes and Twiddy, 2003). Dengue in human causes by four serotypes of dengue viruses (DENV1-4). The E protein substitution rate suggested that the four serotypes emerged (split) around 1000 years ago and the cross-species to human transmission happened just in a few hundred years (Twiddy et al., 2003). Severe dengue is suspected to associate with sequential infection of different dengue serotype. Lifelong antibodies produced from the primary infection are able to cross-interact with the other three serotype viruses, but cannot neutralize the virus. The virus-antibody complexes are endocytosed into cells such as macrophages, monocytes, and dendritic cells, but are not destroyed. Instead, it increases the infected cell population and so the presence of viruses in the blood. There is no specific treatment for dengue. Dengue vaccine (CYD-TDV, a live attenuated tetravalent yellow fever 17D backbone) was recently available in 2015 and has been licensed in some endemic countries(WHO, 2016).

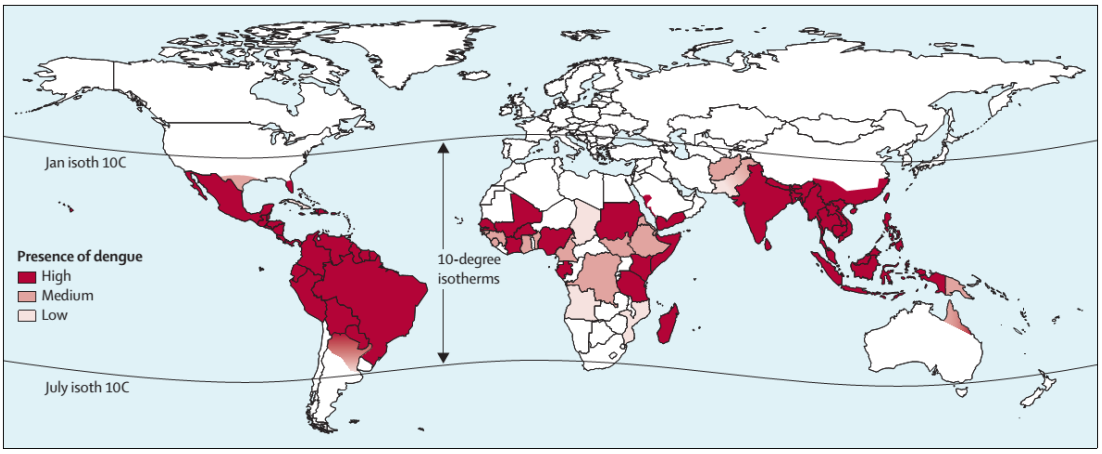


Figure 1.11 Global presence of dengue
Taken from (Guzman and Harris, 2015). License permission no. 4235380130183.

1.3. West Nile virus

WNV is closely related to JEV (**Figure 1.1**) and has a similar life cycle. WNV is transmitted by *Culex* mosquitoes. The main vector is *Culex pipiens*. The natural transmission cycle is between mosquitoes and birds such as domestic birds. In North America, more than 50 species of mosquitoes and 280 species of birds were found infected with WNV. Some mammals (e.g. squirrels) and reptiles (e.g. alligators) may act as the amplifying host. Human and horse are the dead-end hosts as they usually develop low viraemia. WNV is present in Western hemisphere throughout Canada to Venezuela, Africa, parts of Europe, Middle East, West Asia, and Australia. It was originally isolated from patient in West Nile district of Uganda, 1937. WNV was introduced to New York, the United States of America in 1999 and caused a large outbreak which later spread to all over the continent. There are 5 lineages of WNV, which lineage 1 found in all distribution area. Lineage 2-5 are geographically distributed. Lineage 2 present in Africa and Europe. Lineage 3 was isolated in Europe

(Austria), but only from mosquitoes. Lineage 4 is circulating in Russia and lineage 5 is restricted in India. Majority of infection around 80% is asymptomatic. Approximately 20% of infected people develop symptoms including flu-like symptoms, diarrhea, rash (on the torso), and swollen lymph glands. In severe cases, around 1 in 150 infections, patients may develop encephalitis or meningitis which symptoms include neck stiffness, disorientation, tremors, seizure, and paralysis. There is no specific treatment for West Nile virus infection. Even horse vaccine is available, currently, there is no vaccine available for human treatment.

1.4. Zika virus

Zika virus (ZIKV) is drawing the world's attention by its ability to spread very quickly and has affected millions of people (only in Brazil, 2015, between 440,000 and 1.3 million cases were estimated to occurred)(Slavov et al., 2016). ZIKV was first isolated from a monkey during research in the Zika Forest of Uganda in 1947 (**Figure 1.12**). The first human cases were reported in Uganda and the United Republic of Tanzania in 1952 (Smithburn, 1952) and Indonesia in 1977 (Olson et al., 1981), which was 11 years after ZIKV was isolated from mosquitoes in Malaysia in 1966. Despite the intermittent human infection, other nonspecific methods studied on human serum suggested that ZIKV circulates in both Africa and Asia (Weaver et al., 2016). Five positive human sera tests were reported in Gabon in 2007 and in the same year, an outbreak occurred in Yap Island with 49 confirmed cases (Duffy et al., 2009). In 2013, a large outbreak in French Polynesia causing approximately 28,000 cases was reported (Musso et al., 2014). In late 2014, large numbers of skin rash cases were reported in North-eastern Brazil. Almost 7,000 cases were reported but ZIKV was not suspected to be the cause. Until around May, 2015, laboratory confirmed the

circulation of ZIKV in Brazil and ZIKV was announced as associated with the outbreak (Campos et al., 2015) (WHO, 2015). ZIKV rapidly spread to South and Central American countries.

Typically, symptoms of Zika virus infection are mild including fever, skin rash, muscle and joint pain, and headache. There is, however, a proposed relationship between Zika infection and Guillain-Barré syndrome and microcephaly, which later was proved by detection of ZIKV in blood and tissue samples of babies with microcephaly (Marrs et al., 2016). Moreover, studies in a mouse experimental model confirmed the disruption of neural progenitor development by ZIKV (Cugola et al., 2016, Li et al., 2016). ZIKV is transmitted by several *Aedes* spp. mosquitoes including *Aedes aegypti*, which is widespread globally. Even though case reports are declining, ZIKV has high potential to spread to new areas where *Aedes* spp. mosquitoes are present. Similar to other flaviviruses, there is no specific treatment. In addition, there is currently no vaccine available. Base on the NS5 gene, ZIKV is very closely related to DENV, so it is believed to share many characteristics with other flaviviruses. However, it is obvious that ZIKV has a distinct disease mechanism. It is interesting therefore to see the protein structures and compare to other flaviviruses with the aim to determine any differences that may cause these conserved proteins to function differently.

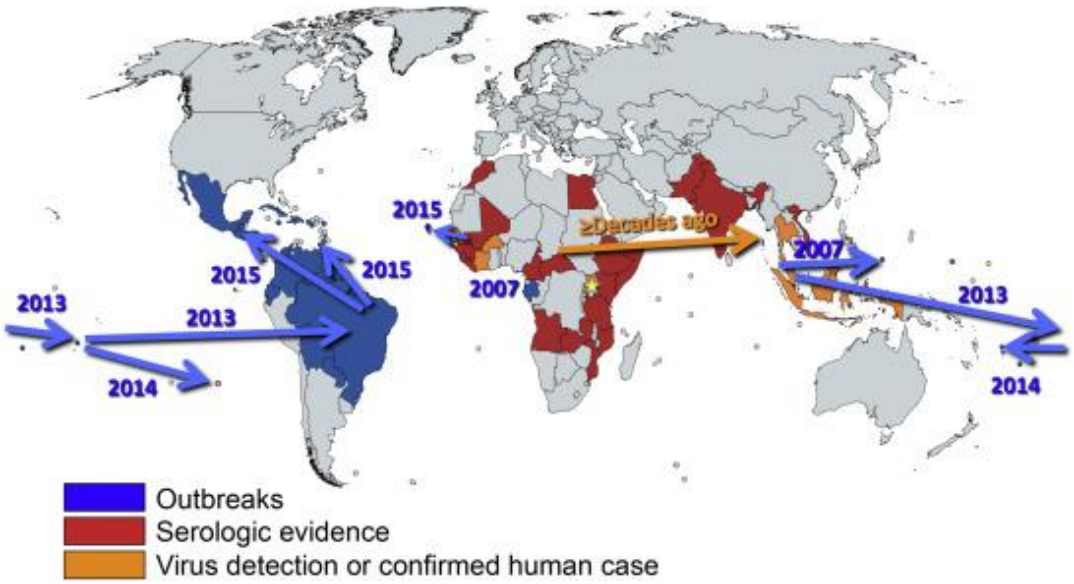


Figure 1.12 ZIKV distribution map

Uganda is indicated with yellow star. Taken from (Weaver et al., 2016). License permission no. 4171011417242.

1.5. Scope of this thesis

JEV is the major cause of viral encephalitis in Asia with no specific treatment. Flavivirus diseases management can be improved by a combination of vaccination, vector control, personal protection, and development of an anti-viral agent. Understanding virus biology is a fundamental key for us to manage the virus. Similarly, defining the molecular structure of related proteins will help us to understand how they function. With mounting evidence suggesting significant roles for C and NS1 protein in the virus life cycle, these multifunctional proteins are targets for new antiviral development as none of the current anti-viral candidates has succeeded in clinical testing. Moreover, all the protein structures in the genome of DENV, the most well-studied flavivirus, have been solved except for the transmembrane proteins: NS2A and NS4A/B. Structural characterisation of JEV is still lacking the information on both C and NS1 proteins. Therefore, the thesis sets out to determine the following:

1. Molecular structures of the two important multifunctional drug target potential proteins, C and NS1 (and its frameshift NS1'), using X-ray crystallography.
2. The molecular structure of NS1 protein of the related virus, Zika virus, to compare to JEV NS1 by using X-ray crystallography.
3. JEV NS1 protein function via protein interaction studies.

Chapter 2 Introduction to protein structural study

Protein structure determination is usually performed by the experimental methods including X-ray crystallography, electron microscopy, nuclear magnetic resonance (NMR), and small-angle X-ray (SAX) and neutron scattering. There are advantages and disadvantages in each of these techniques. For example, X-ray crystallography can give a very high-resolution structure but not all proteins can be crystallised and the protein dynamics is restricted. NMR technique can show protein dynamics and applied to proteins in solution, but it gives lower resolution than X-ray crystallography and protein molecular weight is limited to below 50,000 Da. These techniques can be used in combination to obtain data on different aspects of the research questions. Homology modeling is also useful and more accurate if the template has high degree sequence identity with the target protein. Many successful homology models have been reported especially for small proteins (Zhang and Skolnick, 2004). In addition, partial structural information, for example, molecular mass and post-translational modification can be obtained from various spectrometry methods. In this study, X-ray crystallography and SAXS were used.

2.1 Protein expression

Similar to other protein experiments, pure and functional proteins are required for structural determination, but in larger quantity. Crystallization experiment may require a different amount of protein depending on the success of crystallization screening. Usually, at least several milligrams of protein are required. Unless a large

amount of protein is isolated from the natural source, the protein sample is usually produced using bacterial, insect cell or mammalian protein expression systems (Gomes et al., 2016).

2.1.1 Bacterial expression

With its simplicity, good value for money, high yield, well known genetics, and various molecular implements available, *Escherichia coli* is a typical host for recombinant protein expression (Baneyx, 1999). Usually, sufficient amount of active soluble protein could be achieved from *E.coli* expression (functional tests are required to confirm). Occasionally protein overexpression in *E. coli* might cause partially folded or misfolded protein to accumulate together in inclusion bodies, which could be caused by insufficiency of folding machinery. Importantly, there are differences in cell environment and posttranslational modifications compared to the eukaryotic cell. *E. coli* cytoplasm has reduced environment, so disulphide bond formation is not allowed (Singh and Panda, 2005, Sorensen and Mortensen, 2005). N-linked and O-linked glycosylation have been observed in several bacteria, and the properties have succeeded to transfer to *E. coli.* However, glycoengineered *E. coli* is still in development (Nothaft and Szymanski, 2010). Lack of posttranslational modifications may result in misfolding and non-functional protein. Thus, expression of a eukaryotic glycoprotein may not suitable for the bacterial expression system. However, there are many strategies both with and without engineering of the target protein to improve the soluble protein production (**Figure 2.1**) (Sorensen and Mortensen, 2005). If none of the strategies are able to produce soluble protein, refolding from inclusion bodies may be the method of choice.

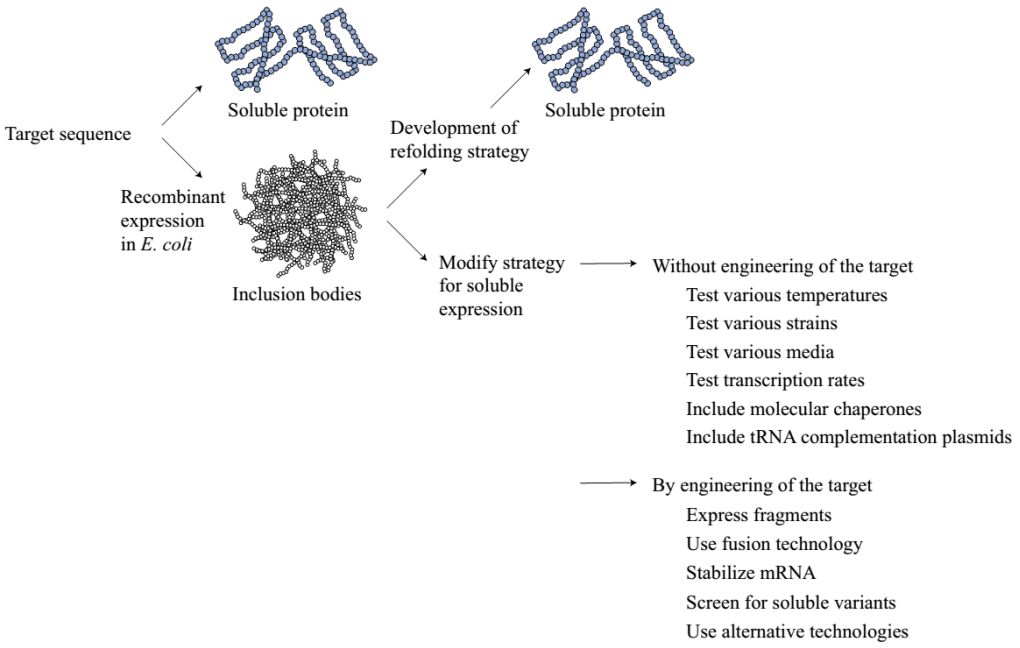


Figure 2.1 Diagram showing soluble protein production strategies used in recombinant *E. coli* expression.

taken from (Sorensen and Mortensen, 2005)

2.1.2 Eukaryotic cell expression

In many cases, glycosylation and disulphide bond formation are crucial for protein folding and functioning. In these cases the use of eukaryotic cell expression system is needed. Three eukaryotic expression systems including yeast, insect cell, and mammalian cells are mentioned in this chapter.

Yeasts are similar to *E. coli* for its ease of handling, low cost and scalability. However, yeasts have different glycosylation than mammalian cells. Mammalian protein expressed in yeast may induce immune response in human use (Ahmad et al., 2014). This is not a concern for structural study, but it may matter for protein-protein complex structure.

Insect and mammalian cells are suitable to produce functional protein due to cell environment and machinery. Mammalian cells are ideal for mammalian protein production, but the complex oligosaccharide chains generated in mammalian cells are heterogeneous and could hinder crystallization. Insect cells have similar protein glycosylation to mammalian cells, but it is more homogeneous. However, glycosylation mutant cell lines, glycosylation inhibitors, and endoglycosidase enzymes are able to manipulate and trim to produce uniform glycan chain. Glycan chain is usually flexible, so electron density of the whole chain may not be visible. Typically, only a first few residues are observed in X-ray structures (Nattleship, 2012). In addition, protein expression in insect and mammalian cells is demanding, expensive and time-consuming. Preferential choice of cells expression system depends on posttranslational modification needed and specific applications of the target protein.

2.2 Protein purification

The level of purity depends upon the purpose of use. For the structural study, the sample must not contain other protein species in the amount that interfere the structure investigation.

Usually, in order to keep the biological activity throughout the process, protein purification is conducted at low temperature (on ice or 4°C containment), using of strong buffer such as very acidic or basic buffer should be avoided. Additionally, purification steps have to be minimized due to the loss of protein yield in every extra step. Purification process should start with low-resolution methods like affinity chromatography followed by refining methods with higher resolution like gel

filtration. One advantage of the recombinant protein expression is a large amount of the expressed target protein compared to the native host expression, which makes it easier to isolate from the mixture.

Proteins differ from one another by their properties like solubility, charge, size, specific binding, and other special properties which can be exploited for protein purification. Required purification steps depend on the protein properties. What has been done previously with the close related protein might be a good starting point (Cutler, 2004).

Histidine is known to bind divalent metal ions. To utilize it for protein purification, polyhistidine (6xHis or His6) is added to the target protein and the tagged protein will selectively bind to Ni^{2+} or Co^{2+} ions. A column packed with agarose beads media which conjugated with a chelating agent, e.g. nitrilotriacetic acid (NTA), and charged with metal ions, e.g. nickel ion (Ni^{2+}) is commonly used for purification. This is called immobilized metal-ion affinity chromatography (IMAC). The target protein is eluted with buffer containing imidazole, which binds strongly to metal ions and competes with histidine-tagged protein. Protein can be directly purified with IMAC from cell lysate suspension.

Size exclusion chromatography (SEC) separates proteins by their sizes using small porous resins with a variety of pore sizes, which different pore size material gives different separation range. Difference sizes of protein are separated by their elution time from the column. Very large proteins, out of the pore size separation range, will be eluted first in the void volume so as aggregated protein. Large molecules are not able to enter the pores and pass through the column faster. Small molecules get into

the pores and take a longer time to travel along the column. Different proteins can be similar in size causing gel filtration to be a weak resolving tool. However, it very benefits for homogeneity of the sample to apply size separation chromatography as the late purification step. SEC is also used for interactional studies when the size of the protein changes due to complex formation.

In the case when protein is expressed as inclusion bodies, protein needs to undergo refolding as the first step of purification. Inclusion bodies, which have a higher density than other cell elements, are easy to separate by centrifugation after cell lysis. Traditionally, protein inclusion bodies are denatured with a high concentration (6-8 M) of chaotropic reagents such as guanidine hydrochloride and urea. Then the denaturant is removed and exchanged with refolding buffer which contains protein folding additives such as L-arginine, detergent, and glycerol. Moreover, oxidizing and reducing agents are required in refolding buffer in order to create redox shuffle for a protein containing disulphide bonds to properly form disulphide bonds. Several refolding strategies can be used, for example, dialysis, dilution, and chromatographic refolding. There is no universal recipe that works for every protein (Singh and Panda, 2005).

2.3 X-ray crystallography

To observe an object, the experimented object has to diffract light and its size has to be larger than the wavelength (λ) of the light. To distinguish between atoms in a molecule which are 0.1 -0.3 nm apart, X-ray with shorter wavelength (~0.1 nm) than visible light (400-700 nm) is required. Diffracted X-rays from one molecule are not

strong enough to detect. It is necessary to get crystals, which have millions of orderly oriented molecules, to obtain strong X-ray diffraction.

2.3.1 Obtaining crystal and crystal quality

Crystallization of inorganic molecules such as salts is well-known. A substance will solidify to form crystals in its supersaturation state, so sample purity and concentration are very important. For protein crystallization, the precipitant is added to separate protein from solution and reach its saturation. Hanging drop (**Figure 2.2a**) and sitting drop methods are commonly used. In each well is filled with reservoir solution which contains precipitant. In the crystallization drop, purified protein is mixed with reservoir solution. At the starting point, the concentration of precipitant in the drop is lower than in the reservoir. Evaporation of water from the drop will increase both protein and precipitant concentration until the concentration of the precipitant in the drop equilibrates with the precipitant concentration in the reservoir.

When the concentration of protein increases and reaches the nucleation zone (**Figure 2.2b**, grey area), nucleation can occur. Once the nucleation occurs, protein crystal growing will decrease the protein concentration to the growth zone (**Figure 2.2b,** orange area). At very high supersaturation beyond the nucleation region, protein will precipitate. However, crystals might also form after precipitation as the protein concentration decrease to the nucleation zone. During the process, besides forming crystals and precipitation, the protein might form a very dense liquid phase like gel or oil which is also in a supersaturated state that crystal may grow. Usually, various conditions are screened for crystallization. The conditions that give crystals or

promising are further optimized. Different pH, precipitant concentration, protein concentration, and additives could improve the crystal quality. The same protein may crystallise in more than one crystal form. Crystallization might take minutes, days, or months.

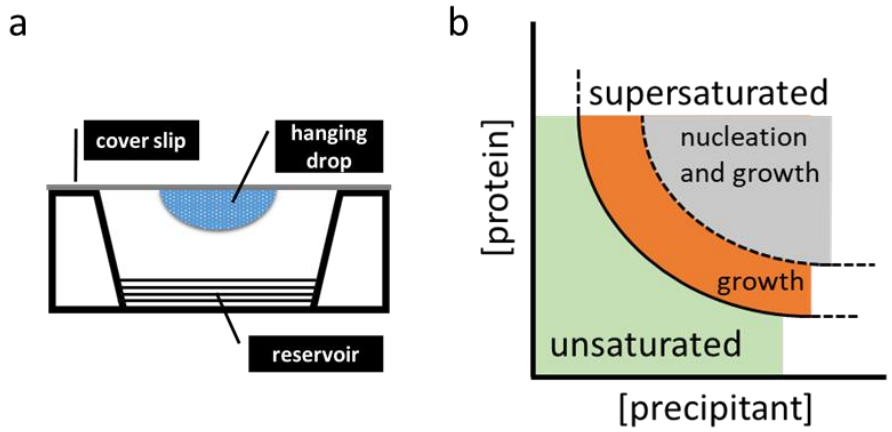


Figure 2.2 Hanging drop method crystallization and crystallization phase diagram

The crystal quality is not judged by the perfection of the crystal (beautiful crystal might not diffract at all) but by its X-ray diffraction properties. Salts (precipitants) might crystallise instead of protein. Several protein crystal properties can be tested to distinguish between protein and salt crystal, for example, protein crystal is easy to crush (see 2.3.2) and will shrink when dehydrated. Protein crystal can absorb small dye molecule through it large solvent channel, while salt cannot, and protein can be detected when washed and dissolved protein crystal is run on SDS-PAGE and stained with Coomassie dye. Besides that, experimental control can be set up alongside by mixing sample buffer and reservoir solution at the same ratio as the experiment in the new drop. (Rhodes, 2006)

2.3.2 Crystal Packing

In the crystal, millions of motif with one identical orientation (or sometimes a few) are packed together in a periodical order. This motif can be an atom, molecule, protein, or protein complex. A unit cell is a term described the smallest volume in the crystal that can translate to fill in the volume of the entire crystal. One may think of a crystal as a lot of unit cells that stack beside and over each other (**Figure 2.3**). The

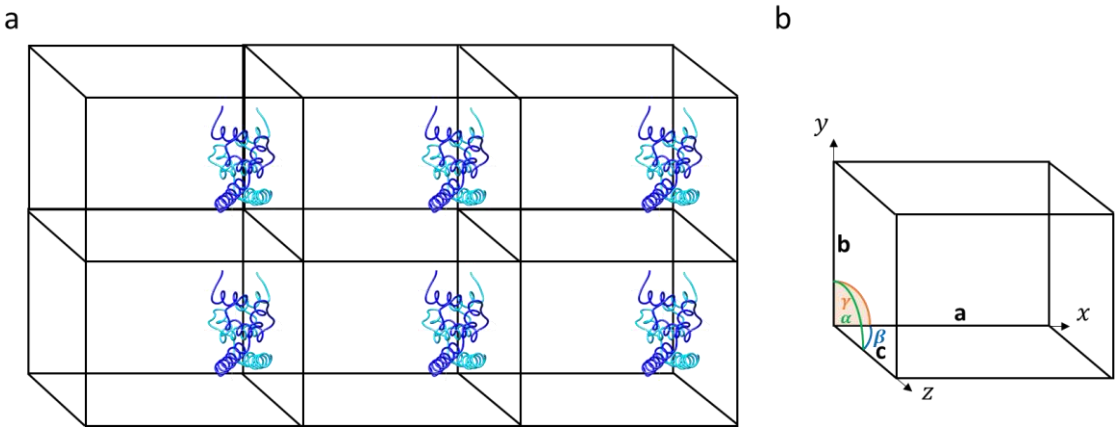


Figure 2.3 Unit cells in a crystal.
(a) Unit cell duplicates itself by stacking on top and alongside each other in every direction through the crystal. **(b)** Unit cell axes, edge lengths, and angles specification.

unit cell 3-dimensional axes are specified by x, y, and z. The lengths of 3 edges are a, b, and c and 3 angles are α, β, and γ (**Figure 2.3b**). Difference edge lengths and angles create 7 possible crystal systems (different shapes of unit cells): cubic, tetragonal, orthorhombic, rhombohedral, hexagonal, monoclinic, and triclinic. An asymmetric unit is a term described the smallest volume that could multiply in the unit cell (generating different alignment copies filling in the unit cell) by symmetry operations called point group symmetry operations: rotation axes, inversion axes, and mirror planes.

The repeated order arrangement of motif forms a pattern in the crystal, which the outline of the patterns is called lattices. Points, where the lines intersect, are called lattice points, which are the sites where atoms can be placed but not necessary. Four types of lattices: primitive, face centered, body centered, and base centered, together with 7 crystal systems give 14 Bravais lattices, which can translate and represent the crystal. Combination of point group symmetry operations, translational symmetry operations (screw axes and glide planes), and Bravais lattices allows 230 combinations in total, the 230 space groups, which describe the unit cell and atoms arrangement within the unit cell. However, inversion axes, mirror plane, and glide planes are not applicable to proteins, which are chiral molecules, making the possible point group symmetry operations and space groups reduce to 11 and 65, respectively.

Each protein molecule in a crystal is held together by the network of hydrogen bonds with water molecules within and between the protein molecules, so protein crystal is fragile and considered to be in an aqueous state. Many studies have shown that crystal structure is the same as its solution state, and protein still retains its function. However, crystal packing could affect the conformation of the structure when molecule lies very close to a neighbouring molecule. Flexible regions such as loop and terminal ends might not visible in the crystal structure.

2.3.3 X-ray diffraction

X-rays are scattered at specific angles (**θ**) by the electron cloud of the atoms in the crystal. These angles depend on the crystal symmetry, protein structure, and X-ray wavelength (**λ**).

When incoming X-ray interacts to electrons, electrons absorb energy, vibrate at the same frequency, and then emit wave at the same frequency in random directions. According to Bragg's law, the diffraction takes place when the path length difference of the two scattered beams is an integer of the wavelength (**nλ**, where n is an integral number), so they are in-phase with each other and interact constructively, while out-of-phase waves interact destructively (**Figure 2.4** and **Figure 2.5**). The X-rays are scattered from the set of lattice planes with the interplanar distance **d** with specific angle **θ**. The distance d also specifies the resolution of the structure. The total path difference between the two X-ray beams (**Figure 2.5,** orange line) is $2d \sin\theta$, which in the condition of constructive interference, it equals to $n\lambda$ (**Figure 2.5,** Bragg's equation). Each diffracted X-ray is the contribution of all atoms in the unit cell, and summation of the scattered waves reach and creates one reflection on the detector.

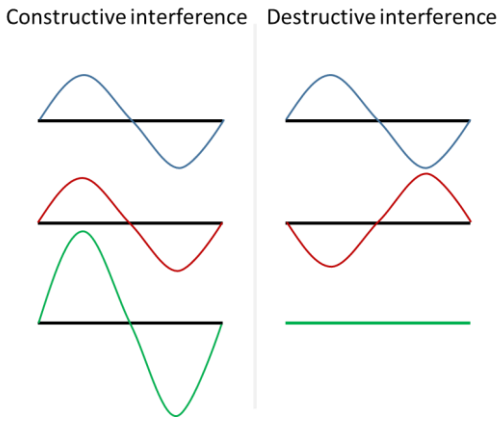


Figure 2.4 Constructive and destructive combining of waves

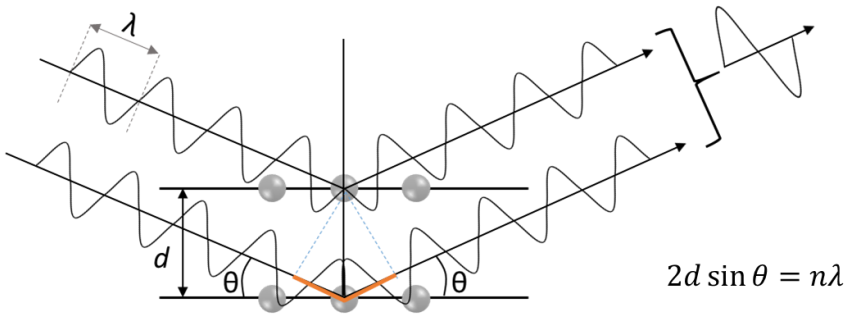


Figure 2.5 Bragg's law.

The 2 scatter beams will be in-phase if the beam distance different is equal to $n\lambda$, where n is an integral number, λ is X-ray wavelength, d is crystal plane spacing, and θ is scattering angle.

Suppose that molecules in the crystal are the arrays of spheres, the diffraction pattern of it also shows as a regular array (**Figure 2.6**). However, there is a proportional inverse relationship between the interval of real spheres and interval of reflections in the diffraction pattern. Because of this relationship, the dimensions of the unit cell in the crystalline lattice could be calculated from the diffraction pattern. For example, if a real unit cell has one of the edge lengths equal to **a**, a reciprocal unit cell has co-linear axis length of **a***:$a^* = {}^1/_a$. The 3-D imaginary space where the reflections inhabit is called *reciprocal space*.

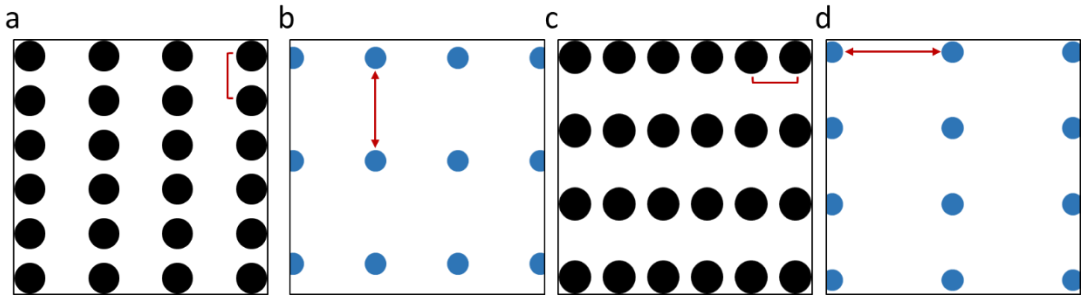


Figure 2.6 Crystalline lattice and reciprocal lattice inverse relationship.

(a) When spheres are close together only vertically, **(b)** the reciprocal lattice (diffraction pattern) expands only vertically. **(c)** When spheres are close together only horizontally, **(d)** the reciprocal lattice expands only horizontally.

2.3.4 Data collection

X-ray diffraction data are usually collected at cryogenic temperatures to improve diffraction by increasing molecular order and reducing radiation damage from X-ray beam. Crystals are held in a small loop and soaked in reservoir solution supplemented with cryoprotectant prior to the data collection. Crystal is picked up, flash frozen in liquid nitrogen, and stored in liquid nitrogen until the data collection (Rhodes, 2006). Crystal is rotated over a small range angle (0.1-0.2 degrees at the synchrotron source and 1 degree in-house) to collect reflections from different lattice points. Total rotation angle depends on crystal symmetry. Higher crystal symmetry needs the smaller overall rotation to collect all unique reflections.

Modern X-ray detectors convert X-ray photons into an electrical signal by photoelectric effect of silicon in complementary metal-oxide-semiconductor (CMOS). Information is read out from chips in each pixel in the detector. Intensity and position of each reflection are recorded (**Figure 2.7**). The phases of x-rays are not directly observed, so the phase information is lost during the measurement.

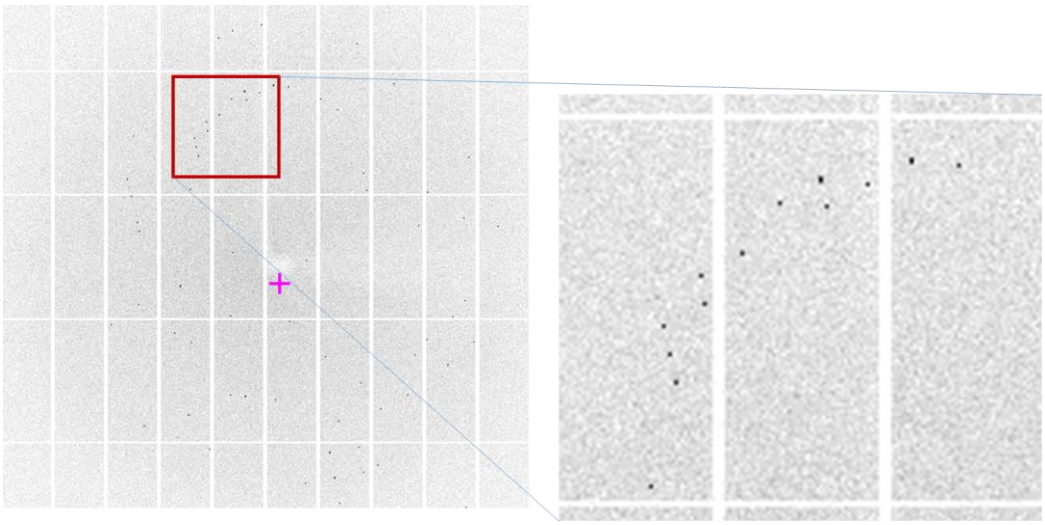


Figure 2.7 One frame of diffraction data detected from C-*JEV*NS1 protein crystal.

2.3.5 From diffraction data to electron density

There is no lens to collect X-ray beams to accurately reconstruct an image. The refractive index of all mediums is very close to the refractive index of a vacuum for X-rays. With a weak refraction, X-ray lenses are impracticable. Thus, all the processes to reconstruct the image are done by calculation.

From data frames, unit cell dimension, symmetry of the crystal, detector parameters, reflection indices (h, k, l) and their intensities (I_{hkl}) are extracted (h, k, l is a 3-dimensional coordinate used to specify a position in reciprocal space). Set of data is integrated and a single output file is produced. This process is called *data reduction.* As the crystal is rotated, reflections might be partially recorded in one frame and fully recorded in another frame. The *scaling* process is performed to improve the consistency of raw intensity data.

According to Fourier theory, a complex wave can be estimated by combining simple sine and cosine waves. Basic waveform is written as a complex number $[cos2\pi(hx) + i\ sin2\pi(hx)]$ (equation 1, Eq.1) and can also be expressed as an exponential $e^{i\ 2\pi\ (hx)}$(Eq. 2). F specifies the amplitude, h specifies frequency, and α specifies phase, which is not directly stated. The sums of n terms are as follow:

$$f(x) = \sum_{h=0}^{n} F_h\ [cos2\pi(hx) + i\ sin2\pi(hx)] \qquad (1)$$

or

$$f(x) = \sum_{h=0}^{n} F_h\ e^{2\pi i(hx)} \qquad (2)$$

or as a 3-dimensional sum

$$f(x,y,z) = \sum_h \sum_k \sum_l F_{hkl} e^{2\pi i(hx+ky+lz)} \quad (3)$$

Diffracted X-ray, which is a summation of in-phase scattered waves, is a complex wave. Therefore, it can be written as a wave equation. The summation of the reflection hkl is called the structure factor F_{hkl}. If a unit cell contains n atoms, the structure factor F_{hkl} is written as follow:

$$F_{hkl} = \sum_{j=1}^{n} f_j \, e^{2\pi i(hx_j+ky_j+lz_j)}, \quad (4)$$

where f_j is the scattering factor of atom j. Coordinate of atom j is indicated by x_j, y_j, and z_j. The frequencies in 3 directions, x, y, and z, are the reflection indices, h, k, and l, respectively.

The structure factor F_{hkl} can also be written as the summation of small volumes of electron density in the unit cell. The 3-dimensional electron density is represented by $\rho(x,y,z)$. To precisely get the average electron density, very small volume elements are integrated together where V is the unit cell volume (Eq.5).

$$F_{hkl} = \int_V \rho(x,y,z) e^{2\pi i(hx+ky+lz)} dV \quad (5)$$

Fourier transform is a mathematical operation that changes the information in one domain to another, which still represent the same data, and the 2 domains are reciprocal to each other. Fourier transform is widely used to decompose a function of time into the frequencies (Eq.6). $f(x)$ is a function of time and $F(h)$ is a function

of frequency ($1/(time)$ is frequency). Fourier transform is reversible, so it can be written as in equation 7.

$$F(h) = \int_{-\infty}^{+\infty} f(x) e^{2\pi i(hx)} dx \tag{6}$$

$$f(x) = \int_{-\infty}^{+\infty} F(h) e^{-2\pi i(hx)}\, dh \tag{7}$$

When compare equation 5 and equation 7, F_{hkl} is, in fact, the Fourier transform of $\rho(x, y, z)$. However, the electron density is the aim of the calculation revealing the structure of the target molecule. Thus, the inverse Fourier transform is as follow:

$$\rho(x, y, z) = \frac{1}{V} \sum_{h} \sum_{k} \sum_{l} F_{hkl}\, e^{-2\pi i(hx+ky+lz)} \tag{8}$$

Integral is not used and instead, the summation is used because the structure factor F_{hkl} is a set of the discrete function of reflections. Amplitude and phase are required to compute the structure factor F_{hkl}. The amplitude of F_{hkl} is proportional to the square root of the reflection intensity I_{hkl}, so it can be obtained from measured reflection intensities. Only the phase information cannot be measured by a current diffraction experiment, because the instrument only detects the scattered energy in different directions.

2.3.6 Phase problem

There are techniques available to reconstruct the phases, unmeasurable information from diffraction experiment. Three main techniques for obtaining phases of protein structure are multiple isomorphous replacement, multiwavelength anomalous

dispersion, and molecular replacement. Only details of molecular replacement used in this study are presented in this chapter.

Molecular replacement

Molecular replacement is a phasing technique of obtaining the initial phases using the homological protein model. The template model is called a *starting model*. Nowadays, there are ~100,000 protein structures solved by X-ray crystallography in the protein data bank (PDB). Among these, more than 70% were solved by molecular replacement. This method is very useful if homologue structure is available. Homologue structure usually belongs to the same protein family. The starting model typically needs at least 35% sequence identity, cover at least 50% of the whole molecule, and has $C_{\propto}$ r.m.s.d. when superimposing with the final model less than 2 Å (Abergel, 2013). Below these criteria, the success rate is decreased.

The structure of the starting model is rotated and translated to match the unknown target structure. The search of all possible orientations and locations is computed by 2 distinct steps. First, the search for the best orientation. Second, the search for the best position. Orientation search is conducted by comparing the Patterson function (Eq.9) of the starting model and the unknown target structure, which does not need phase and could be calculated directly from experimental data.

$$P(u,v.w) = \frac{1}{V}\sum_{h}\sum_{k}\sum_{l}|F_{hkl}|^2\, e^{-2\pi i(hu+kv+lw)} \qquad (9)$$

Patterson function reports Patterson map which represents the positional vectors between each pair of atoms in the structure (**Figure 2.8**).

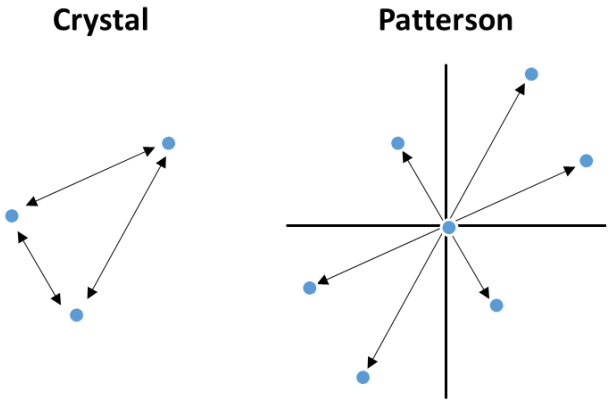


Figure 2.8 Patterson map of a three atoms structure.

The interatomic vectors between atoms in the structure including the symmetry related molecules are plotted and moved to the origin of the unit cell of the Patterson space (identical to the crystal unit cell).

Patterson map of the starting model is superimposed and rotated against the Patterson map of the unknown target structure. The top solutions of the rotation search that give a good match to the model are then used for location search. The location search is performed by translating the selected interatomic Patterson map over that of the unknown structure to find the best match. The structure factors of each new model are calculated. Agreement of the 2 models is accessed by *R-factor,* the comparing the amplitude (Eq.10) where $|F_{obs}|$ is amplitude derived from diffraction experiment and $|F_{calc}|$ is amplitude calculated from the structure factor of the new orientated and position model.

$$R = \frac{\sum\big||F_{obs}| - |F_{calc}|\big|}{\sum|F_{obs}|} \qquad (10)$$

The model with the highest correlation coefficient, small *R* value, to the unknown target structure would give the best estimation of the target phase. Perfect match *R* is zero and the worse match is one. Phases obtained from molecular replacement are only the initial estimated phases, which require an improvement before interpreting into a final electron density map. *R*-value >0.5 indicates poor agreement and may not possible to improve (near 0.59 is a random model) (Karplus and Diederichs, 2012) and *R*-value <0.4 is more amenable to improve by refinement.

2.3.7 Molecular modelling and validation

Electron density map is a contour map of $\rho(x, y, z)$ (**Figure 2.9**). The first map is calculated from the observed amplitude and estimated phases. However, when the electron density map is purely calculated by phases from the model and amplitudes from experimentally measured data, it is possible to introduce bias from the model to the electron density map. To reduce the bias, the amplitude of each term of the electron density map is calculated by subtraction of the calculated amplitude from some multiple (n) of the observed amplitude ($\big|n|F_{obs}| - |F_{calc}|\big|$) (Eq.11). A map from this calculation is a *different map* called $nF_o - F_c$ map where $F_o \; and \; F_c$ represent $F_{obs} \; and \; F_{calc}$, respectively. The difference map with *n* = 2 with weighting terms is immensely used ($2mF_o - DF_c$ where *m* is the figure of merit and *D* is the Sigma-A weighting factor).

$$\rho(x, y, z) = \frac{1}{V} \sum_h \sum_k \sum_l (n|F_{obs}| - |F_{calc}|)\, e^{-2\pi i(hx+ky+lz-\alpha'_{calc})} \qquad (11)$$

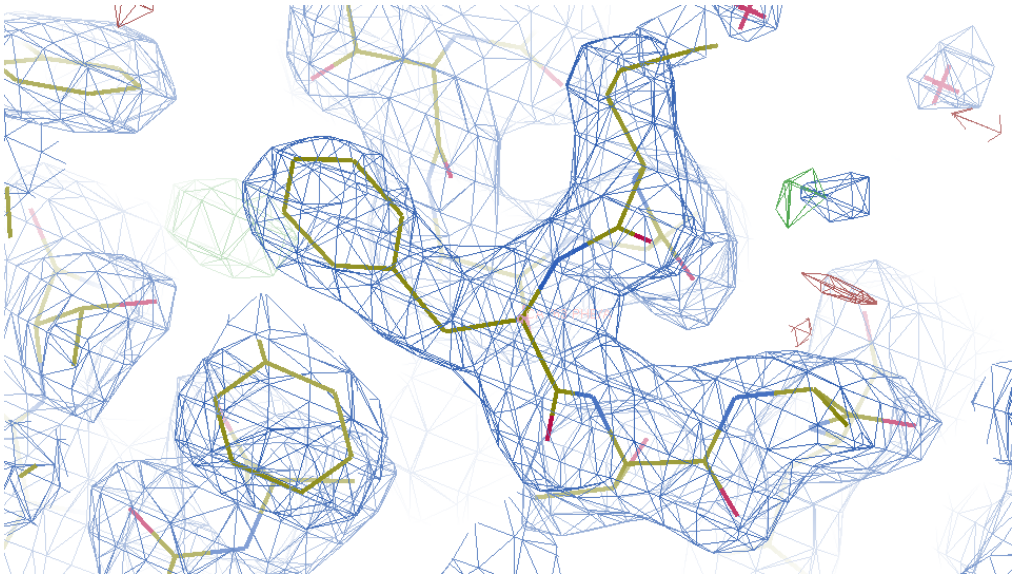


Figure 2.9 Electron density map and molecular model built into it.
The JEV capsid protein displayed by program COOT (Emsley et al., 2010).

Target protein sequence is usually known, so the molecular model is built into the electron density map according to the known sequence by partially automate computer program. However, there are still areas on the map that need user intervention to identify specific molecular features and additional molecules from crystallization condition or bound cofactors and ions.

In the different maps, electron density can be either positive or negative. Positive density, usually shows in green colour, indicates that F_{obs} are larger than F_{calc} (Eq.11). Thus, this region in the real unit cell has more electron density than indicated by the model. The positive map is suggesting that there should be atoms in this region but data are not accounted in the model. The positive density is corrected by moving or filling atoms into the region. Negative density, on the other hand, indicates that the model shows more electron density than in the real unit cell and usually shows in red colour. Adding atoms in negative density should be avoided. After optimization, new structure factors are calculated from this new model and a new phase is obtained.

New electron density map is calculated from the observed amplitudes and the new phase. These sequence of events, called *refinement*, is repeated multiple times. Stereochemical parameters and *R*-factor are monitored in each round of refinement. Bond length, bond angle, and side chain rotation should be chemically realistic. *R*-factor should keep low, for example, a model at 2.5 Å resolution is expected to have *R-factor* of 0.2. Another parameter is introduced, R_{free}, to make sure that the refined model is not over manipulated. The R_{free} calculation is the same as *R-factor* but the F_{obs} is the ~5% of the reflection data that set aside from refinement. *R* and R_{free} values should be close to each other. R_{free} is usually higher but ideally not more than 5%. The refinement is considered finished when both *R* and R_{free} reach stable values, and there is no interpretable density present.

Structure Validation

The final step is to confirm that the model is complete and thoroughly refined. Submission of the model to the online server validation, Molprobity (Chen et al., 2010), gives details on each parameter quality for each amino acid residue (**Figure 2.10**). Geometric analysis and all-atom contact analysis are conducted for structure validation. Bond length, bond angle, backbone conformation ($\varphi, Phi\ and\ \psi, Psi$ torsional angles accessed by Ramachandran diagram) should be in chemical realistic range. The residue with poor parameter should be remodelled. However, some molecular features, which categorizes as an outlier but have strong electron density map, may be the model uniqueness.

All-Atom Contacts	Clashscore, all atoms:	8.83		82nd percentile* (N=865, 1.85Å ± 0.25Å)
	Clashscore is the number of serious steric overlaps (> 0.4 Å) per 1000 atoms.			
Protein Geometry	Poor rotamers	5	3.11%	Goal: <0.3%
	Favored rotamers	147	91.30%	Goal: >98%
	Ramachandran outliers	2	1.09%	Goal: <0.05%
	Ramachandran favored	179	97.81%	Goal: >98%
	MolProbity score^	1.89		74th percentile* (N=12654, 1.85Å ± 0.25Å)
	Cβ deviations >0.25Å	1	0.58%	Goal: 0
	Bad bonds:	0 / 1501	0.00%	Goal: 0%
	Bad angles:	7 / 2053	0.34%	Goal: <0.1%
Peptide Omegas	Cis Prolines:	2 / 12	16.67%	Expected: ≤1 per chain, or ≤5%
	Twisted Peptides:	2 / 186	1.08%	Goal: 0

#	Alt	Res	High B	Clash > 0.4Å	Ramachandran	Rotamer	Cβ deviation	Bond lengths	Bond angles	Cis Peptides
			Avg: 35.52	Clashscore: 2.92	Outliers: 0 of 175	Poor rotamers: 1 of 157	Outliers: 0 of 163	Outliers: 0 of 178	Outliers: 4 of 178	Non-Trans: 1 of 176
A 176		THR	88.62	-	-	Allowed (1.9%) *p* chi angles: 81.8	0.02Å	-	-	-
A 177		THR	65.25	-	Favored (23.16%) General / -161.1,169.0	Favored (9.8%) *t* chi angles: 185.8	0.02Å	-	-	-
A 178		GLU	47.44	-	Favored (45.31%) General / -109.4,138.7	Favored (23.7%) *tt0* chi angles: 184.2,166.9,288.3	0.02Å	-	-	-
A 179		CYS	39.5	-	Favored (47.87%) General / -64.5,149.8	Favored (53.2%) *m* chi angles: 302.8	0.01Å	-	OUTLIER(S) worst is CB-SG-SG: 4.2 σ	-
A 180		ASP	56.18	-	Favored (36.11%) General / -67.0,128.4	Favored (8.4%) *t0* chi angles: 204,330.6	0.03Å	-	-	-
A 181		SER	39.21	-	Favored (42.01%) General / -101.7,7.1	Favored (63.1%) *p* chi angles: 58.1	0.01Å	-	-	-

Figure 2.10 Molprobity statistics summary.

Following by uploading the protein coordinates file to Molprobity server, the program calculate several statistical parameters. Colour coding red, yellow, and green demonstrate bad, moderate, and good values, respectively. Percentile is compared across protein structures deposited in PDB at similar resolution.

2.4 Small angle X-ray scattering (SAXS)

SAXS shares the principles of elastic X-ray scattering with X-ray crystallography. Data obtained from the experiments are similar in terms of intensity and diffraction angle measurements, but the data are sampled differently. While X-ray crystallography focuses on the data at high resolution, SAXS analyses sample at low resolution, e.g. 20 Å. SAXS can be used to probe the structure of molecules in solid, liquid, or gaseous states. The liquid sample is focused in this study as it is useful for determination of the overall shape and conformation of the protein in solution. Particles in solution do not exhibit periodicity but move freely. Particles in solution, therefore, give isotropic scattering patterns spherically averaged around the beam center. While this data cannot provide atomic detail of molecule under investigation, it does provide information on average shape, size, and molecular weight (MW). A necessity of SAXS measurements is the acquisition of pure and functional protein.

2.4.1 SAXS measurements and data analysis

Elastic scattering refers to a stage of no net energy loss before and after the interaction with matter. Incoming and scattering X-rays can be written as momentum vectors k_{in} and k_{out}, respectively **(Figure 2.11)**. Vector q, *momentum transfer*, specifies the difference momentum of X-ray before and after the collision. Vector q is:

$$q = \frac{4\pi \sin\theta}{\lambda}, \tag{12}$$

where scattering angle is 2θ and a momentum vector k is $2\pi/\lambda$. q is in reciprocal space and the unit is the inverse of length, nm^{-1} or $Å^{-1}$.

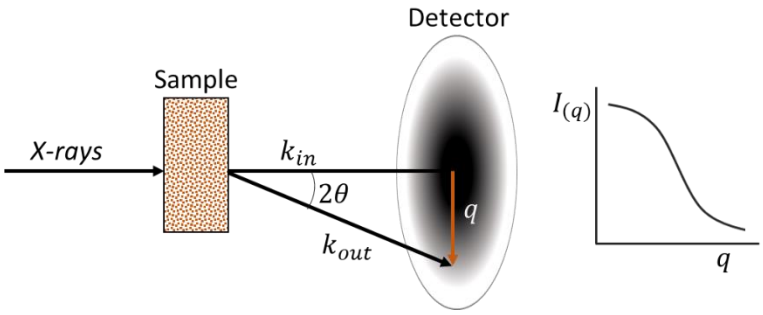


Figure 2.11 SAXS experiment diagram.

X-ray beams are written as momentum vector k. Scattered wave intensities are recorded.

Target molecules are surrounded by solvent. Incident X-rays are scattered by both target molecules and solvent. The useful information is the difference of electron density between the target molecules and solvent ($\Delta\rho$) (**Figure 2.12a**). Signal contrast is obtained by measuring the buffer signal alone then subtract it from the total signal (**Figure 2.12b).**

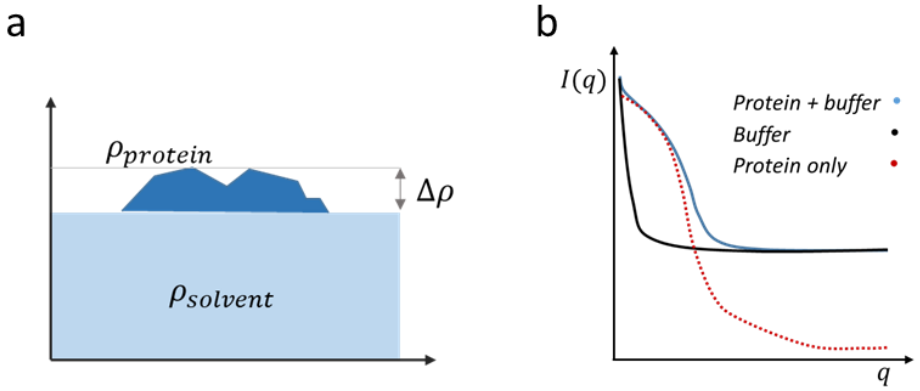


Figure 2.12 Signal contrast.

Electron density difference of protein is around 10% above the background **(a)** and obtained by subtraction of the solvent (background) signal **(b)**.

Summation of scattering waves is specified by the *form factor* of the particle $F(q)$, which contain intra-particle interaction information such as size and shape. The term *structure factor* $S(q)$ gives additional information from neighboring particles when they are packed (concentrated) or has interparticle interaction such as attraction and repulsion. The scattering intensity equation is:

$$I(q) = F(q) * S(q) \tag{13}$$

Scattering intensity $I(q)$ is experimentally measured. In order to get size and shape information of the sample, which is obtained from $F(q)$, $S(q)$ is expected to equal to 1, meaning no interparticle interaction in the solution. The ideal sample for SAXS is monodisperse (identical size and shape) with molecules far away from each other Distance between particles is larger than the wavelength. If the sample is polydisperse (varied sizes), the form factor gives the average size from all particles. An inter-particle interaction contributes to the form factor and cause inaccurate measurement of parameters. Attractive interaction usually gives larger parameters than the actual, for example, particle size, volume, and molecular weight, while repulsive interaction will give too small parameters (**Figure 2.13**). Thus, good sample preparation is very important. Multiple techniques may be used such as size exclusion chromatography to characterize the sample and avoid polydispersity or aggregation.

Another method to check data quality is Guinier method, which developed by Andre Guinier, 1939. At very low q value (close to zero), the intensity is approximated as follow:

$$I(q) = I_0 e^{-q^2 R_g^2/3} \tag{14}$$

$$\ln I(q) = \ln I_0 - \frac{R_g^2}{3} q^2 \qquad (15)$$

Equation 15 is a linear relationship plotted of $\ln I(q)$ against q^2. High-quality and reliable data should fit perfectly to the slope line (**Figure 2.13**). Attractive or aggregation interaction shows upward curve line, and repulsive interaction shows downward curve line. Poor fit data should not be further processed (**Figure 2.13**).

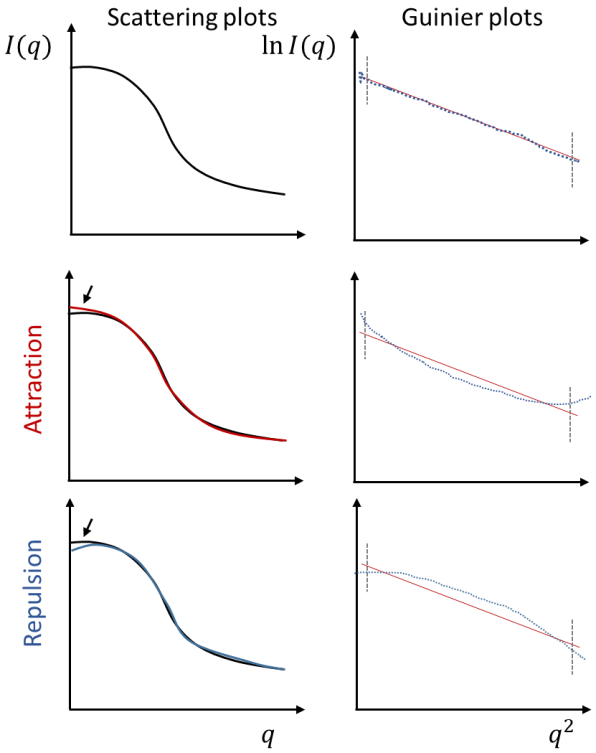


Figure 2.13 Scattering and Guinier plot examples of inter-particle interaction samples.

Attractive interaction give increase intensity at low q values of scattering plot, while repulsive interaction gives the opposite result indicated by black arrows. In addition, inter-particle interaction sample is poorly fit to the slope line in Guinier plot.

The radius of gyration defined by SAXS measurements is the root mean square (rms) radius of electron scatterers R_g, which is a measure of mass distribution from the center of mass. It is determined from the slope of the Guinier plot.

The forward scattering intensity I_0 is determined from the $I(q)$ axis intercept at the intensity at $q = 0$ and it is proportional to the number of electrons in particle. I_0 is useful for molecular weight determination. However, the SAXS intensity is not measured on an absolute scale. The indirect method to obtain the absolute intensity is calibrating to the secondary standards such as lysozyme, glucose isomerase, or bovine serum albumin. Known concentration of both sample and standard protein, and partial specific volume of the protein are required for I_0 based molecular weight estimation. Normalized to the concentrations, the partial specific volume values of the standard and protein sample are assumed identical, so the ratio of the molecular weights of the two proteins is equal to the ratio of the I_0 (Eq.16).

$$\frac{MW_{exp}}{I_{0\,exp}} = \frac{MW_{standard}}{I_{0\,standard}} \tag{16}$$

Molecular weight may be estimated from the particle volume as well (where the particle is assumed to have uniform density) called Porod volume (V_p) method (Petoukhov et al., 2012)(Eq.17).

$$V_p \simeq 2\pi^2 I_0 \,/ \int_0^\infty q^2 I(q) dq \tag{17}$$

The entire range of momentum vector q is used. However, at high q angles, the scattering intensity falls off and extrapolated by $I(q) \propto q^{-4}$ relationship by Porod's law.

The particle volume V_p is multiply by the ratio between molecular weight and Porod volume, which is averaged to 0.625, to get the estimate molecular weight (Eq.18).

$$MW = V_p * 0.625 \tag{18}$$

To extract more information from the data, scattering profile undergoes a Fourier transform to get the pair distribution function $P(r)$ in real space. The $P(r)$ distribution function describes the distances between atoms within the protein molecule (Eq. 19, **Figure 2.14**).

$$I(q) = \int P(r) \frac{\sin(qr)}{qr} dr, \tag{19}$$

where r specifies the spatial distance between scattering points and I specifies intensity.

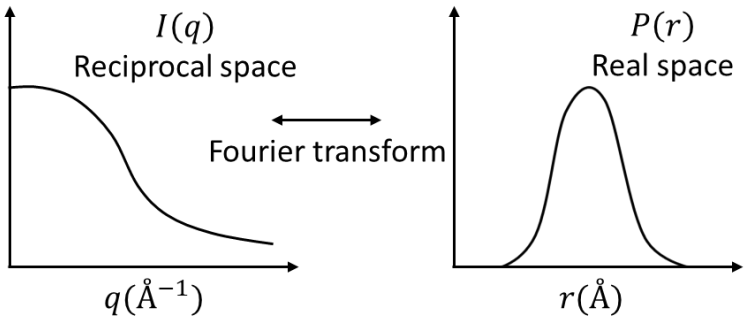


Figure 2.14 Pair distribution function.
It is calculated from X-ray scattering data via Fourier transformation.

The $P(r)$ represents the distribution of distances between pairs of points contained within a particle. The maximum distance or the maximum linear dimension D_{max} can be obtained. $P(r)$ plot can demonstrate the rough particle shape, for example, globular shape particle shows normal distribution profile (**Figure 2.14**), while elongate particle shows left skew and hollow sphere shows right skew profile. R_g can also be calculated from $P(r)$ as follows:

$$R_g^2 = \int_0^{D_{max}} r^2 P(r) dr \Big/ \int_0^{D_{max}} P(r) dr \tag{20}$$

Different from Guinier method which uses the information only at low q value, R_g from $P(r)$ uses the entire information, so it can be more accurate. R_g values from both methods are reported for data quality check and they should agree with each other.

Ab initio structural modeling is one of the methods to reconstruct a low-resolution model of the proteins, which is monodisperse and assumed to have uniform density. The method proceeds by filling up the protein volume with beads. Beads in the model are connected to resemble an amino acid chain. The scattering profile of the bead model is calculated to compare with experimental SAXS profile. The discrepancy is evaluated by χ^2 value. Searching for a model which fits the data well is an iterative process and does not yield a unique solution. Thus, even though the shape restoration program is run with the same data, differences in the reconstructed models will arise. In order to overcome this problem, many models given by *ab initio* modeling are averaged together to give a representative model of the protein shape.

Chapter 3 Materials and Methods

The general materials and methods used in this study are described in this chapter. The specific materials and methods are in its own chapter.

3.1 Materials

3.1.1 Bacterial strains

For routine cloning, *Escherichia coli* strain TOP10 was used. For *E. coli* protein expression, three strains, which were BL21(DE3), BL21(DE3)pLysS, and SHuffle®T7, were tested for soluble protein production. The SHuffle®T7 cells were kindly provided by Dr. Dunhao Su. Bacterial strains and genotypes are indicated in Table 3.1. Cells were treated with calcium chloride to generate chemically competent cells (see 3.2.3).

Table 3.1 Bacterial strains and its genotype

Name	Strain	Genotype
TOP10	K12	F- *mcrA* Δ(*mrr-hsd*RMS-*mcr*BC) Φ80*lac*ZΔM15 Δ *lac*X74 *rec*A1 *ara*D139 Δ(*araleu*)7697 *gal*U *gal*K *rps*L (StrR) *end*A1 *nup*G
BL21(DE3)	B	*fhuA2 [lon] ompT gal (λ DE3) [dcm] ΔhsdS*
BL21(DE3)pLysS	B	F-*, omp*T*, hsdS$_B$* (rB–, mB–)*, dcm, gal, λ*(DE3), pLysS, Cm^r.
SHuffle® T7	K12	F´*lac*, *pro*, *lacI^Q* / *Δ(ara-leu)7697 araD139 fhuA2 lacZ::T7 gene1 Δ(phoA)PvuII phoR ahpC* galE (or U) galK λatt*::pNEB3-r1-*cDsbC* (Spec^R, *lacI^q*) *ΔtrxB rpsL150*(Str^R) *Δgor Δ(malF)3*

3.1.2 Culture medium

For routine non-inducing bacterial growth, for example, plasmid amplification, transformation, starter culture, colonies selection, glycerol stock and competent cell preparation, Miller's formulation Luria-Bertani (LB) medium was used (Novagen). Protein expression in *E. coli* was conducted using Isopropyl β-D-1-thiogalactopyranoside (IPTG) induction in LB broth or autoinduction terrific broth media without trace elements (AIMTB) (Formedium). The medium was prepared as in the product instructions. Homemade autoinduction media (ZYM5052) was prepared as following: 1% tryptone, 0.5% yeast extract, 25 mM Na_2HPO_4, 25 mM KH_2PO_4, 50 mM NH_4Cl, 5 mM Na_2SO_4, 2 mM $MgSO_4$, 0.2× trace elements, 0.5% glycerol, 0.05% glucose, 0.2% α-lactose) (Studier, 2014). The media was autoclaved at 121°C for 20 minutes before use. The appropriate amount of required antibiotics was added to the media for a selectable marker (see **Table 3.2** and **Table 3.3**).

Table 3.2 Antibiotic concentrations for bacterial selection.

Antibiotic	Working concentration
Ampicillin	100 µg/ml (50 µg/ml, if 2 antibiotics were used in the same culture)
Chloramphenicol	25 µg/ml (20 µg/ml, if 2 antibiotics were used in the same culture)
Kanamycin	50 µg/ml

3.1.3 DNAs

Japanese encephalitis virus strain SA14 (accession: M55506) was used as a template for recombinant protein construction. Synthetic *JEV*NS1 DNA (nucleotide residues

2478-3533) optimized for *E. coli* was purchased from Life technologies and synthetic native *JEV*NS1 DNA was provided by OPPF-UK. NS1 gene fragments: N-*JEV*NS1 (nucleotide residues 2478-3017), DII-III-*JEV*NS1 (2538-3533), and C-*JEV*NS1 (2991-3533) (**Figure 3.1**), were generated by polymerase chain reaction (PCR) method using primer pairs shown in **Table 3.4**. Synthetic *JEV*NS1' and capsid DNA (1-315) were purchased from GenScript (Piscataway, NJ, USA). NS1' was the special case which the full length NS1' was the NS1 sequence with 156 additional nucleotides. The frameshift sequence was manually added by insertion of thymine at position 3561 as a result of to -1 ribosomal frameshifting which creates new stop codon at 52 amino acids far from NS1 C-terminal (**Figure 3.2**). DII-III-*JEV*NS'1 (from residue 2538) and C-*JEV*NS1' (from 2991) (**Figure 3.1**) were generated from the synthetic *JEV*NS1' by using PCR method.

Zika virus (accession: KU365779) was the template for recombinant protein construction. Synthetic *ZIKV*NS1 DNA (residues 2476-3524) and C-*ZIKV*NS1 (residues 2989-3524) optimized for *E. coli* were purchased from GenScript (Piscataway, NJ, USA).

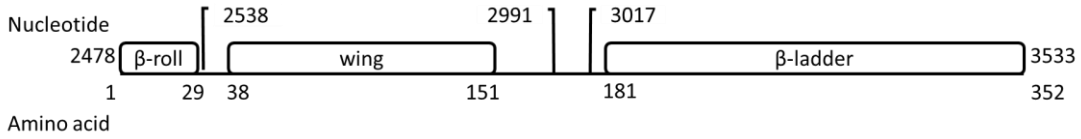


Figure 3.1 JEV NS1 gene diagram.
The diagram shows β-roll, wing and β-ladder domains together with the amino acid position of each domain and nucleotide position for each truncation.

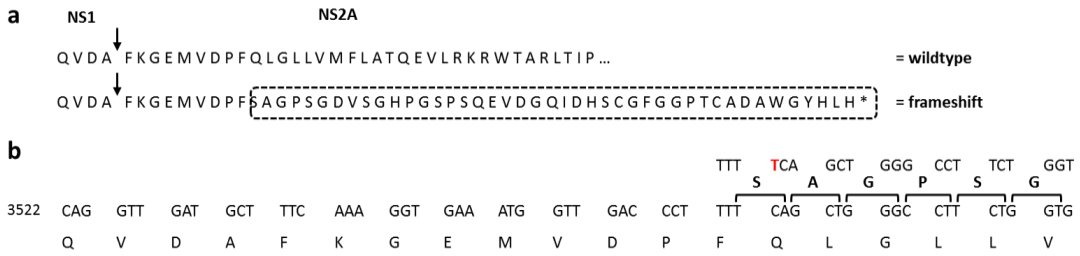


Figure 3.2 The -1 ribosomal frameshifting.

(a) Amino acid sequence of *JEV*NS1 C-terminal and *JEV*NS2A N-terminal are shown. Wildtype sequence is compared to frameshift sequence. Cleavage sites are indicated with arrows. Frameshifted sequences are in the dash box. **(b)** Repeated translation of amino acid thymine is indicated in red causes the translation frame to change and generates new amino acids.

3.1.4 Vectors

Synthetic sequences were cloned into expression vectors by the provider service (**Table 3.3**, shaded in gray). After that, if appropriate, the sequences were subcloned into new vectors to generate desired constructs. Different secretion signal sequences, fusion tags, and gene truncations were constructed as indicated in **Table 3.3**. The pOPIN HBM, G, H, HBM-M, HA-M, Ac64-M, and G-M vectors, which used for mammalian and insect cell expression provided by OPPF-UK, were cloned with the 2 version of target genes, none *E. coli* optimized and optimized *JEV*NS1. Chaperone plasmid set (Takara) for chaperone protein co-expression, which composes of pG-KJE8, pGro7, pKJE7, pG-Tf2, and pTf16, was kindly provided by Dr. Cristina Yuntayanes and Piyapon Jirawatcharadech, Liverpool School of Tropical Medicine.

3.1.5 Primers

The oligonucleotides used with pOPIN vectors were designed for In-Fusion method, while with pET303 vector were designed for ligation method. Primer sequences are given in **Table 3.4**. Cloning site information is shown in **Table 3.3**. All of the oligonucleotides were purchased from Thermo Fisher Scientific except for the oligonucleotides used with pOPIN HBM-M, HA-M, Ac64-M, G-M, H, HBM, and G were provided by OPPF-UK. See Table 3.5 for cycling conditions used for each primer pair.

3.1.6 Antibodies

Peroxidase conjugated mouse anti-histidine monoclonal antibody (Roche, AB_840259) was used to detect recombinant histidine-tagged proteins. The experiments conducted at OPPF were routinely detected by the unconjugated mouse anti-histidine monoclonal antibody (R&D, AB_357353). Mouse monoclonal antibody against WNV NS1 (22NS1) was provided by Dr. Michael Diamond. Peroxidase conjugated goat anti-mouse IgG (Santa Cruz Biotechnology, AB_631738) provided by Sujitra Keadsanti was used to detect the 22NS1 antibody.

Table 3.3 Vector construction of the interested genes used in this study

No.	Vector	Inserted gene	Cloning sites	SS/Tags	Antibiotic resistance
1	pRSET_A_A185	*JEV*NS1	BamHI/KpnI	N-HIS	amp
2	pOPIN M	*JEV*NS1	kpnI/HindIII	N-HIS-MBP-3C	amp
3	pOPIN S	*JEV*NS1	kpnI/HindIII	N-HIS-SUMO-3C	amp
4	pOPIN J	*JEV*NS1	kpnI/HindIII	N-HIS-GST-3C	amp
5	pOPIN HBM	*JEV*NS1[a]	kpnI/HindIII	HBM/C-HIS	amp
6	pOPIN G	*JEV*NS1[a]	kpnI/HindIII	uPase/C-HIS	amp
7	pOPIN H	*JEV*NS1[a]	kpnI/HindIII	uPase/N-HIS-3c	amp
8	pOPIN HBM-M	*JEV*NS1[a]	kpnI/HindIII	HBM/N-HIS-MBP-3C	amp
9	pOPIN HA-M	*JEV*NS1[a]	kpnI/HindIII	HA/N-HIS-MBP-3C	amp
10	pOPIN Ac64-M	*JEV*NS1[a]	kpnI/HindIII	Ac64/N-HIS-MBP-3C	amp
11	pOPIN G-M	*JEV*NS1[a]	kpnI/HindIII	uPase/N-HIS-MBP-3C	amp
12	pOPINTTGneo	*JEV*NS1[a]		RPTPmu/C-HIS	amp
13	pET303	C-*JEV*NS1	XbaI/XhoI	-	amp
14	pOPIN F	C-*JEV*NS1	kpnI/HindIII	N-HIS	amp
15	pET303	N-*JEV*NS1	XbaI/XhoI	-	amp
16	pET303	DII-III-*JEV*NS1	XbaI/XhoI	-	amp
17	pET-30a(+)	*JEV*NS1'	NdeI/XhoI	-	kan
18	pET303	C-*JEV*NS1'	XbaI/XhoI	-	amp
19	pET303	DII-III-*JEV*NS1'	XbaI/XhoI	-	amp
20	pET-30a(+)	*JEV*capsid	BamHI/XhoI	N-HIS-S tag-E	kan
21	pET-15b	*ZIKV*NS1	NdeI/BlpI	N-HIS-TEV	amp
22	pET-15b	C-*ZIKV*NS1	NdeI/BlpI	N-HIS-TEV	amp

Notes: [a] = both of none *E. coli* optimized and *E. coli* optimized gene were tested with the plasmid.

SS = signal sequence, N = N-terminal tagC = C-terminal tag, amp = ampicillin, kan = kanamycin, HIS = histidine tag, MBP = Maltose binding protein, GST = glutathione S-transferase, HA = hemagglutinin, HBM = honeybee melittin, SUMO = small ubiquitin-like modifier, uPase = uridine phosphorylase, S tag = pancreatic ribonueclease A, Ac64 = Autographa californica evelope protein gp64, Cleavage sites: 3C = human rhinovirus (HRV) 3C Protease, TEV = tobacco etch virus protease, E = enterokinase, RPTPmu = receptor-like protein-tyrosine phosphatase

Table 3.4 Primers used for subcloning of *JEV*NS1 N-terminal, C-terminal and domain II-III.

PCR cycle	Plasmid/ gene	Sequence
1	pOPIN M, J, S	
	optimized *JEV*NS1	Fw 5'aagttctgtttcagggcccgGATACCGGTTGTGCCATTGATATT Rv 5'atggtctagaaagctttaTGCATCAACCTGGCTACGAACCAG
2	pOPIN HBM-M, HA-M, Ac64-M, G-M, H	
	optimized *JEV*NS1	Fw 5'aagttctgtttcagggcccgGATACCGGTTGTGCCATTGATATTACCCG Rv 5'atggtctagaaagctttaTGCATCAACCTGGCTACGAACCAGGG
2	pOPIN HBM	
	optimized *JEV*NS1	Fw 5'tcttacatctatgcgGATACCGGTTGTGCCATTGATATTACCCG Rv 5'gtgatggtgatgtttTGCATCAACCTGGCTACGAACCAGGG
	Native *JEV*NS1	Fw 5'tcttacatctatgcgGACACTGGATGTGCCATTGACATCACAAG Rv 5' gtgatggtgatgtttAGCATCAACCTGTGATCTGACGAGTGT TG
2	pOPIN G	
	optimized *JEV*NS1	Fw 5'gcgtagctgaaaccggcGATACCGGTTGTGCCATTGATATTACCCG Rv 5'gtgatggtgatgtttTGCATCAACCTGGCTACGAACCAGGG
	Native *JEV*NS1	Fw 5'gcgtagctgaaaccggcGACACTGGATGTGCCATTGACATCACAAG Rv 5'gtgatggtgatgtttAGCATCAACCTGTGATCTGACGAGTGTTG
3	pET303	
	optimized C-*JEV*NS1	FW5'gctctagaatgCGTGAAGAAAGCACCGATGAATGTGAT RW5'ccgctcgagTTATGCATCAACCTGGCTACGAACCAG
	optimized N-*JEV*NS1	Fw5'gctctagaatgGATACCGGTTGTGCCATTGATATTACC Rw5'ccgctcgagttaATCACATTCATCGGTGCTTTCTTCACG
	optimized DII-III *JEV*NS1	Fw5'gctctagaatgGTGCATAATGATGTTGAAGCATGGGTG Rw5'ccgctcgagTTATGCATCAACCTGGCTACGAACCAG
	optimized C-*JEV*NS1'	FW5'gctctagaatgCGTGAAGAAAGCACCGATGAATGTGAT Rw5'ccgctcgagTTAATGCAGATGATAACCCCATGCATctg
	optimized DII-III *JEV*NS1'	Fw5'gctctagaatgGTGCATAATGATGTTGAAGCATGGGTG Rw5'ccgctcgagTTAATGCAGATGATAACCCCATGCATctg

Note: Lowercase characters indicate vector sequences and uppercase characters indicate gene sequences.

3.2 Methods

3.2.1 Polymerase chain reaction (PCR)

Replication of DNA template was performed by using KOD Hot Start DNA polymerase (Merck Millipore) or CloneAmp HiFi PCR Premix (Clontech). The reaction was set up following the product instructions with the final concentration of template DNA ~10 ng. Cycling conditions were set up as in **Table 3.5** using Techne TC-PLUS thermal cyclers. See **Table 3.5** to match the cycling conditions with the primers in **Table 3.4**. If appropriate, PCR products were purified using Wizard® SV Gel and PCR Clean-up System (Promega).

Table 3.5 PCR reaction setup.

Step	PCR cycle 1 KOD Hot Start	PCR cycle 2 KOD Hot Start	PCR cycle 3 CloneAmp HiFi
1.Polymerase activation	95°C 2 min	95°C 2 min	
2.Denature	95°C 20 sec	95°C 20 sec	98°C 10 sec
3.Annealing	65°C 10 sec	70°C 10 sec	55°C 15 sec
4.Extension	70°C 40 sec	70°C 40 sec	72°C 5 sec
5.Final extension	70°C 10 min	70°C 10 min	
Repeat steps 2-4	35 cycles	30 cycles	35 cycles

3.2.2 Agarose gel electrophoresis

DNA fragments were separated and analysed by agarose gel electrophoresis. Low melting point agarose was dissolved in 1x Tris-acetate-EDTA (TAE) buffer (40 mM Tris, 20 mM acetic acid, 1 mM ethylenediaminetetraacetic acid (EDTA)), which prepared from 10x stock, to achieve 1% agarose gel with Midori green DNA stain (Bulldog Bio)

at the concentration of 6 µl/100 ml. DNA samples or 0.5 µg of 1 kb DNA ladder (New England Biolabs, N3232S) were mixed with 6x purple loading dye (New England Biolabs, B7025S) at 5:1 ratio DNA to dye. Electrophoresis was run in 1x TAE buffer at 120 volts constant until the dye reached the bottom of the gel.

3.2.3 Chemically competent cells

To make chemically competent cells, each step was done with aseptic technique. Desire *E. coli* strain was recovered from glycerol stock by streaking on LB agar plate which contains appropriate antibiotics, if necessary, and incubating overnight at 37°C (12-16 hours). A colony was inoculated into 5 ml of LB broth with appropriate antibiotics and incubated at 37°C overnight to make an overnight culture. Cells were transferred to 500 ml LB broth with appropriate antibiotics and allowed to grow at 37°C until OD_{600} reached 0.4. Cells were placed on ice for 20 minutes and kept cold throughout the procedure. Cells were pelleted at 3000xg for 10 min, resuspended in 30 ml of cold 0.1 M $CaCl_2$, and incubated on ice for 30 min. Cells were pelleted again and resuspended in 8 ml of cold 0.1 M $CaCl_2$ and 15% glycerol. Cells were aliquot into 1.5 ml tubes, freeze in liquid nitrogen and stored at -80°C.

3.2.4 Plasmid extraction

Five ml of overnight culture of *E. coli* that contain plasmid of interest was prepared. Plasmids were extracted by using Wizard® Plus SV Minipreps DNA Purification System (Promega) or Qiagen miniprep following the product instructions.

3.2.5 Cloning

In-fusion cloning

The primer used for infusion cloning was designed by the 5′ end was the 15 bases of one end of the vector sequences at the restriction site that the DNA fragment will be joined. Second part at 3′ end was the sequence which specific to target DNA, size about 18-25 bases, and contained 40-60% GC content. pOPIN vector was double digested using HindIII and KpnI restriction enzymes (New England Biolabs) at 37°C overnight. Cloning steps were followed the In-fusion HD Enzyme Premix protocol (Clontech Laboratories). Briefly, in 10 µl reaction, In-fusion enzyme (1x), linearized vector (~50 ng), DNA fragment PCR product (~50 ng), and deionized water were mixed together and incubated at 50°C for 15 min. Then the reaction was placed on ice and continued to the transformation experiment.

Ligation cloning

The primer used for ligation cloning was designed to contain restriction site at their 5’ end. pET303 vector and DNA fragment PCR product were double-digested with XbaI and XhoI restriction enzymes (New England Biolabs) to create sticky ends at 37°C overnight and heat inactivated at 65°C for 20 min. Phosphorylated end of the cut vector was removed by incubated at 37°C for 15 minutes with Antarctic phosphatase (New England Biolabs) at 1 u/µg DNA and heat inactivated at 70°C for 5 min. The cut vector was purified with Wizard® SV Gel and PCR Clean-up System (Promega). The ligation reaction was set up by mixing of 1:3 vector to insert (usually 0.02:0.06 pmol) with T4 DNA ligase at the concentration of 20 u/ul (New England Biolabs). The

reaction was incubated overnight at 4°C and heat inactivated at 65°C for 10 min. Then the reaction was placed on ice and continued to the transformation experiment.

3.2.6 Bacterial transformation

Competent cells were thawed on ice. As soon as the cells were completely thawed, 2 µl (~5 ng) of interested plasmid was added to each 50 µl reaction and mixed well by gently flicking the bottom of the tube a few times with a finger. After that, the mixture was incubated on ice for 20 minutes and heat shocked in 42°C water bath for 42 sec. Then the mixture was placed on ice for 5 minutes and 200 µl of LB media was added. Next, the cells were allowed to grow at 37°C for 45 min. All of the transformation reaction was plated onto 10 cm LB agar plate which contains appropriate antibiotics. Competent cells alone without plasmid were treated with the same procedure and used as negative control. Agar plates were incubated at 37°C for 12-16 hours. Plates containing colonies were kept at 4°C until use, but no longer than a month.

3.2.7 DNA sequencing

DNA sequencing experiments were performed by supreme Sanger sequencing service, GATC Biotech. The concentration of isolated plasmids from miniprep kit was measured using NanoDrop® ND-1000 (Thermo Scientific). The samples were prepared at concentration and volume according to supreme Sanger sequencing service requirements (30-100 ng/µl, 20 µl). Universal primers, T7 and pET RP, were used as sense and antisense primer, respectively.

3.2.8 Protein expression

Bacterial expression

Isopropyl β-D-1-thiogalactopyranoside (IPTG) induction

Five ml of *E. coli* overnight culture that contains interested plasmid was prepared and transferred to 500 ml of LB broth with appropriate antibiotics. After incubation at 37°C for 2-3 hours or until OD_{600} reach ~0.5, IPTG was added at the concentration of 0.5 M or 1 M and the temperature was changed to the appropriate temperature for protein expression. Cells were allowed to grow for 12-16 hour overnight and pelleted by centrifugation at 4500xg, 8°C for 20 min.

Autoinduction

Cells were prepared similarly to IPTG induction method. Instead of LB broth, cells were transferred to AIMTB or ZYM5052 with appropriate antibiotics. Cells were allowed to grow for 2-3 hours at 37°C and the temperature was changed to the appropriate temperature to express the protein for 12-16 hour overnight. Expression at 16°C and 18°C were conducted for 24 hr. No IPTG was added.

3.2.9 Protein purification

Lysis buffer screening

After protein was expressed, 100 µl of the culture was aliquoted into 1.5 ml tube and centrifuged at 5000xg, 8°C for 10 min. Cell pellets were resuspended in 4 different initial lysis buffers as indicated in **Table 3.6** (Perry, 2016) and incubated at room temperature for 30 min. Cells were briefly vertex before quickly frozen in liquid nitrogen and thawed in a water bath at 42°C. The freeze-thaw cycle was repeated for 5 times. Cells were centrifuged at 16000xg, 8°C for 10 minutes to separate

supernatant, which contains soluble protein, and pellet, which contain insoluble protein. Supernatant and pellet were analysed separately by sodium dodecyl sulphate-polyacrylamide gel electrophoresis (SDS-PAGE). The result was followed by another set of lysis buffer screen which could be pH, salt, urea, or detergents solubility screen depend on the result from the initial screen. In this study, only salt solubility screen was used (see **Table 3.7**). After the suitable buffer was found, the experiment was scaled up proportionally to the amount of pellet for routine protein purification (at least 1:4 pellet to buffer).

Table 3.6 Lysis buffer screen.

Name	Buffer
7.5N	50 mM Tris pH 7.5, 50 mM NaCl, 5 mM EDTA, 1 mg/ml lysozyme
2S	50 mM Tris pH 7.5, 2 M NaCl, 5 mM EDTA, 1 mg/ml lysozyme
0.5U	50 mM Tris pH 7.5, 50 mM NaCl, 5 mM EDTA, 0.5 M urea, 1 mg/ml lysozyme
D	20 mM Tris pH 7.5, 50 mM NaCl, 0.2% NP 40, 1 mg/ml lysozyme

Table 3.7 Salt solubility screen

Name	Buffer
0.1S	50 mM Tris pH 7.5, 0.1 M NaCl, 5 mM EDTA, 1 mg/ml lysozyme
0.5S	50 mM Tris pH 7.5, 0.5 M NaCl, 5 mM EDTA, 1 mg/ml lysozyme
1S	50 mM Tris pH 7.5, 1 M NaCl, 5 mM EDTA, 1 mg/ml lysozyme
0.1K	50 mM Tris pH 7.5, 0.1 M KCl, 5 mM EDTA, 1 mg/ml lysozyme
1K	50 mM Tris pH 7.5, 1 M KCl, 5 mM EDTA, 1 mg/ml lysozyme

Protein isolation

For routine protein purification, cells from 1 L culture were pelleted in appropriate lysis buffer (20mM Tris·HCl pH 7.4, 150 mM NaCl, and 1 mg/mL lysozyme). Cells were incubated at room temperature for 30 minutes before sonicated on ice with 9.5 mm probe at amplitude 10 for 30 seconds with 30 seconds interval on ice for 10 times. Then, cells were spun down at 17000xg, 4°C for 20 min. The supernatant and pellet fractions were analysed by SDS-PAGE and Coomassie staining. If target protein is insoluble, the pellet will further use for protein refolding or stored at-80°C until use. If protein is soluble, suspension fraction is further purified.

JEV NS1 insoluble protein refolding, and purification

Modified Edelling's method

Protein was expressed by auto-induction and insoluble protein pellet was isolated and resuspended in resolubilization buffer (7 M guanidine hydrochloride (Gnd·HCl), 30 mM β-mercaptoethanol (β_2ME)) for 1 hour at 37°C at 1:10 pellet to buffer. Then centrifuged at 16000xg for 10 minutes at 8°C to remove undissolved protein, and diluted with 50 mM sodium acetate pH 5.2 to reduce Gnd·HCl concentration to 2 M. The sample was filtered with 0.45 micron filter and refolded by adding 10 ml of denatured protein at 1 ml/hr to 1 L of 400 mM L-arginine, 100 mM Tris, pH 8.3, 2 mM EDTA, 0.5 mM oxidized glutathione (GSSH), 5 mM reduced glutathione (GSH), 0.2 mM phenylmethanesulfonyl fluoride (PMSF) at 4°C. After overnight incubation, the refolded protein was filtered with 0.2 µm filter and concentrated with Amicon stirred cell using PL-10 membrane and 10000 Da molecular weight cut-off (MWCO) centrifuge concentrator. NS1 protein was subjected to Superpose 6 10x300 mm or Superdex 200 10x300 mm size-exclusion chromatography column (GE Healthcare Life

Science) equilibrated with 20 mM 4-(2-hydroxyethyl)-1-piperazineethanesulfonic acid (HEPES) pH 7.4, 150 mM NaCl.

Protein quantification

Protein concentration was determined by the bicinchoninic acid assay (BCA assay) following the product protocol (Pierce™ BCA Protein Assay Kit) or UV absorbance at 280 nm by using NanoDrop® ND-1000 (Thermo Scientific). After obtained the absorbance value, protein concentration was calculated from Beer's law:

$$A = \varepsilon * b * c,$$

where *A* is the absorbance value, ε is extinction coefficient, *b* is container path length, and *c* is protein concentration in molarity.

3.2.10 Sodium dodecyl sulphate-polyacrylamide gel electrophoresis (SDS-PAGE)

Protein yield and purity were examined by SDS-PAGE in Laemmli buffer system. Protein was separated by 4% (stacking gel) and 12% polyacrylamide gel (resolving gel) at 0.75 or 1.0 mm thickness. See gel recipes in **Table 3.8**. Samples were mixed with 4x sample buffer (277.8 mM Tris-HCl pH 6.5, 44.4% (v/v) glycerol, 4.4 % SDS, 0.02% (w/v) bromophenol blue) supplemented with 355 mM β_2ME (100 µl in 900 µl 4x sample buffer) at 3:1 sample to buffer ratio and heated at 95°C for 5 min. Samples with a high concentration of Gnd·HCl were precipitated with ethanol prior mixing with sample buffer. Briefly, samples were mixed with chill absolute ethanol at 1:40 ratio sample to ethanol and incubated at -80°C for 20 minutes to 1 hour. Then the samples were centrifuged at 16000xg for 5 minutes at 8°C to separate the

precipitated protein. Ethanol was discarded and protein pellet was air dried in the tube with the lid open for 10 minutes. The samples were then mixed with 50 µl of 1x sample buffer and heated normally. Protein standard (New England Biolabs, P7712 or PageRuler, 26616) was used to determine the protein size. Electrophoresis was conducted at 200 volts constant in 1x electrode buffer (25 mM Tris-base, 200 mM glycine, 1% (w/v) SDS) until the dye reaches the bottom of the gel by using Mini-PROTEAN®3 Cell electrophoresis system (BIO-RAD). The gel was subjected to Western blot or Coomassie staining. For Coomassie staining, the gel was stained in 0.1% (w/v) Coomassie Brilliant Blue R-250, 40 % (v/v) methanol and 10 % (v/v) glacial acetic acid for 20 minutes or until the gel was stained consistency blue and sample dye cannot be seen. Then gel was destained in 40 % (v/v) methanol 10 % (v/v) glacial acid solution until the background was clear.

Table 3.8 Gel formulations (10 ml)

12% Resolving gel		**4% Stacking gel**	
10% (w/v) SDS	100 µl	10% (w/v) SDS	100 µl
Gel buffer (1.5 M Tris, pH 8.8)	2.5 ml	Gel buffer (0.5 M Tris, pH 6.8)	2.5 ml
Water	3.3 ml	Water	6 ml
30% Acrylamide/Bis	4 ml	30% Acrylamide/Bis	1.3 ml
10% APS	100 µl	10% APS	100 µl
TEMED	10 µl	TEMED	10 µl

3.2.11 Western blot

After protein samples were resolved on polyacrylamide gels, the protein was transferred to Immobilon®-P membrane (Merck Millipore) by electro-tank transfer (Mini Trans-Blot® cell wet electroblotting systems, BIO-RAD) or dry transfer method

by using iBlot® 2 Gel Transfer Device (Thermo Fisher Scientific). Protein was transferred according to transfer apparatus manufacturer's instructions. For a tank transfer, gels were transferred in transfer buffer (25 mM Tris, pH 8.3, 192 mM glycine, 20% methanol) at constant 400 mA for 1 hour with the cooling unit. For dry transfer, iBlot® 2 Transfer Stacks PVDF was used with P_0 default method (7 min). After that, the membrane was incubated in blocking buffer (5% (w/v) non-fat dried milk, 20 mM Tris pH 7.6, 150 mM NaCl, 0.1% Tween 20) for 30 min. Next, the membrane was incubated with primary antibody (in blocking buffer) for 1.5 hours and washed with 20 mM Tris pH 7.6, 150 mM NaCl, 0.1% Tween 20 for 5 minutes 3 times. After that, the membrane was repeated with the incubation and washing steps with secondary antibody (these steps were omitted if using peroxidase conjugated anti-histidine antibody). Amersham ECL start (GE Healthcare Life Science), a chemiluminescent detection reagent, was used and the result was analysed by ImageQuant LAS 4000 (GE Healthcare Life Science). See antibodies used in this study in section 3.1.6.

3.2.12 Crystallization and data collection

The proteins were concentrated and screened using commercial crystallization screens. The SaltRx, PEGRx, and Natrix screens from Hampton research, and Structure, PACT premier, and JCSG screens from Molecular Dimension were used on crystallization robot (SCREENMAKER 96+8, Innovadyne/Gilson) by adding 200 nl of protein and 200 nl of reservoir solution to drop well of Intelli-Plate 96 sitting drop crystallization. All positive conditions are listed in **Appendix 3**. Successful conditions were optimized by hanging drop method in 24 (500 µl reservoir) or 48 wells plate (200 µl reservoir).

X-ray data were collected at a cryogenic temperature at beamline I24 or I02 at Diamond Light Source, UK and PROXIMA 1 at Soleil synchrotron, France. Data reduction was carried out by iMosflm (Battye et al., 2011) (C-*JEV*NS1), XDS (Kabsch, 2010) (C-*JEV*NS1') or Xia2 programs (Winter, 2010) (JEV capsid). The C-*JEV*NS1 and C-*JEV*NS1' protein structures were determined by molecular replacement using the structure of WNV NS1 C-terminal domain (PDB: 4OIE, >70% sequence identity) as a starting model by MOLREP (Vagin and Teplyakov, 2010) in the CCP4 program suite as well as capsid protein which WNV capsid protein (PDB: 1SFK, 68% sequence identity) was a starting model. For capsid protein, the automated model building was performed by Buccaneer (Cowtan, 2006). The structures were refined by REFMAC5 (Murshudov et al., 2011) and built in COOT (Emsley et al., 2010) in CCP4.

3.2.13 Small-angle X-ray scattering (SAXS)

The purified C-*JEV*NS1 1st peak (2.4 mg/ml), C-*JEV*NS1 2st peak (3.4 mg/ml), C-*JEV*NS1' 1st peak (5 mg/ml), C-*JEV*NS1' 2nd peak (5 mg/ml) (see **Chapter 5**), and C-*JEV*NS1-22NS1 complex (3 mg/ml) (see **chapter 6**) in TBS buffer (20 mM Tris-HCl pH 7.4, 150 mM NaCl) were analysed with SEC-SAXS on beamline SWING at Soleil synchrotron, France. Samples were load onto Agilent BioSEC-3 4.6x300 mm column at flow rate of 0.25 ml/min, 15°C. Data were collected at a distance of 1.8 m and X-ray wavelength of 1 Å. Data were collected by Dr. Gareth Wright. Data processing was conducted in PRIMUS (Konarev et al., 2003). Comparison of scattering profile was done in FoXS (Schneidman-Duhovny et al., 2013). *Ab initio* model was the average from 10 (C-*JEV*NS1 and C-*JEV*NS1') or 20 (protein complex) independently model calculations with (protein complex) or without symmetry (C-*JEV*NS1 and C-*JEV*NS1') using

DAMMIF (Franke and Svergun, 2009). The model was average with DAMAVER (Volkov and Svergun, 2003) and refine with DAMMIN (Svergun, 1999). The low resolution model surface representation was created in CHIMERA (Pettersen et al., 2004) using 'molmap' command. The molecular mass was calculated from Porod volume (Petoukhov et al., 2012).

Chapter 4 JEV NS1 and NS1' protein expression and purification trials

Flavivirus NS1 protein has been studied for more than 30 years, but recombinant soluble protein has never been successfully produced from a bacterial expression system. The expression has been reported from insect and mammalian cells expression systems in the 1990s (Flamand et al., 1992, Falgout et al., 1989). Later with different vector constructions, expression of flavivirus NS1 protein in baculovirus and mammalian cell expression systems have been widely used in many studies to produce recombinant NS1 for protein characterization and functional study (Flamand et al., 1995, Brown et al., 2011, Falgout et al., 1989, Flamand et al., 1992, Noisakran et al., 2008, Avirutnan et al., 2010, Avirutnan et al., 2011, Somnuke et al., 2011). In 2014, WNV and DENV full length NS1 protein X-ray structures were presented for the first time (Akey et al., 2014, Edeling et al., 2014). The proteins were produced by insect cells according to the previously described study of Brown, 2011 (Brown et al., 2011).

In this study, JEV NS1 expression was attempted in bacterial cells using both engineering and non-engineering of the gene strategies. Gene engineering strategies were including *E. coli* codon optimization, fusion tags, and expression of protein fragments. Non-engineering strategies were the testing of various culture temperature, medium, *E. coli* strains, and chaperone protein co-expression. Moreover, refolding from inclusion body was also tested.

Parallel to *E. coli* expression, JEV NS1 was expressed in mammalian and insect cell systems since the protein is glycosylated and both cells are the natural host of the virus. The protein produced from these systems should resemble natural NS1 and reveal its native characteristics and functions. The protein production trial was conducted at OPPF-UK. The vector construction for insect cell expression was designed followed Brown, 2011 study.

4.1 Materials and methods

4.1.1 Comparison of percentage identity of NS1 sequence

Virus strain was chosen to use in this project by comparing of NS1 amino acid sequences of JEV genotype 1 to 5 and other flaviviruses taken from NCBI database. Five of genotype III, which are common and widely spreads strain, and 1 of each genotype II-V were used to represent each genotype. Percentage identity was obtained by using the local alignment search tool (BLAST), the National Center for Biotechnology Information (NCBI) with two strains aligned at a time. Moreover, JEV NS1 sequences were compared to WNV lineage 1, WNV lineage 2, Murray, Kunjin, St Louis, Usutu, Alfuy, DengueT1, DengueT2, DengueT3, DengueT4, TBE, YFV, and ZIKV. Percent identity was depicted in (**Table 4.1**).

Table 4.1 Percentage identity of NS1 amino acid sequences compared within *Flavivirus* genus

Flavivirus	NCBI accession number	JEV strain								
		Genotype III					Genotype IV	Genotype I	Genotype V	Genotype II
		JEV SA 14-14-2	JEV SA 14	JEV P3	JEV Nakayama	JEV NC_001437	JEV JKT6468	JEV K94P05	JEV MUAR	JEV FU
JEV SA 14-14-2	AF315119									
JEV SA 14	M55506	99%								
JEV Beijing	L48961	98%	98%							
JEV P3	U47032	99%	99%							
JEV Nakayama	EF571853	99%	99%	99%						
JEV NC_001437	NC_001437	99%	99%	100%	99%					
JEV JKT6468	AY184212	94%	95%	96%	95%	95%				
JEV K94P05	AF045551	95%	95%	95%	96%	96%	93%			
JEV MUAR	HM596272	91%	91%	91%	92%	91%	92%	89%		
JEV FU	AF217620	95%	96%	97%	96%	97%	92%	95%	91%	
WNV lineage 1	NC_009942	77%	77%	78%	78%	78%	77%	77%	78%	78%
WNV lineage 2	NC_001563	76%	76%	77%	77%	77%	76%	76%	77%	77%
Murray	NC_000943	78%	78%	78%	78%	78%	79%	78%	81%	79%
Kunjin	AY274504	76%	76%	77%	76%	77%	75%	76%	76%	77%
St Louis	JF460774	64%	64%	65%	65%	65%	65%	64%	66%	64%
Usutu	NC_006551	76%	76%	76%	77%	76%	77%	75%	78%	77%
Alfuy	AY898809	78%	77%	78%	78%	78%	78%	76%	81%	78%
DengueT1	U88535	51%	51%	51%	51%	51%	51%	50%	52%	51%

Flavivirus	NCBI accession number	JEV strain								
		Genotype III					Genotype IV	Genotype I	Genotype V	Genotype II
		JEV SA 14-14-2	JEV SA 14	JEV P3	JEV Nakayama	JEV NC_001437	JEV JKT6468	JEV K94P05	JEV MUAR	JEV FU
DengueT2	M29095	53%	53%	53%	53%	53%	54%	53%	54%	53%
DengueT3	EF643017	51%	51%	51%	51%	51%	50%	51%	52%	52%
DengueT4	AF326573	52%	52%	53%	53%	53%	51%	51%	53%	53%
TBE	NC_001672	44%	44%	44%	44%	44%	44%	44%	44%	44%
YFV	NC_002031	46%	46%	46%	46%	46%	46%	47%	47%	46%
ZIKV	KU365779	56%	56%	57%	57%	57%	56%	57%	57%	56%

4.1.2 Plasmid construction

Synthetic *JEV*NS1 DNA optimized for *E. coli* in the pRSET_A_A185 vector was used for *JEV*NS1 full length expression. The *JEV*NS1 full length was subcloned into pOPIN M, J and S vectors to produce *JEV*NS1 with MBP, GST, and SUMO fusion tags, respectively. NS1 gene fragments, N-*JEV*NS1, DII-III-*JEV*NS1, and C-*JEV*NS1, were generated from the full length template by PCR method.

Synthetic *JEV*NS1' in pET-30a(+) was used for *JEV*NS1' full length expression. DII-III-*JEV*NS1', and C-*JEV*NS1' were generated from *JEV*NS1' full length. All the truncated DNA fragments were cloned into a pET303 vector.

For mammalian and insect cell expression, the experiments were divided into 2 sets by the inserted DNA. The first was the synthetic *JEV*NS1 DNA optimized for *E. coli* and the second was the synthetic native *JEV*NS1 DNA, which was provided by OPPF-UK. Both were cloned into pOPIN HBM, G, H, HBM-M, HA-M, Ac64-M, and G-M vectors. pOPINTTGneo was used later instead of pOPIN G for native DNA expression in mammalian cells expression for a better yield. The cloning step (except for pOPINTTGneo) was done in Liverpool with the guidance of Dr. Louise Bird, OPPF-UK.

4.1.3 Protein expression and purification trials

All of the trial conditions were summarised in **Table 4.2**.

_JEV_NS1 full length with histidine tag

_JEV_NS1 was expressed in _E. coli_ BL21(DE3)PLysS by IPTG induction in LB media at 30°C and 18°C. The protein was also expressed in _E. coli_ BL21(DE3) at 30°C and SHuffle® T7 by autoinduction in AIMTB at 25°C, 30°C, and 37°C. Cells were pelleted and resuspended in lysis buffer. Protein was isolated. The supernatant and pellet fractions were analysed by SDS-PAGE and Coomassie staining.

_JEV_NS1 full length with fusion tags

The _JEV_NS1 with MBP, GST, and SUMO tags were expressed in _E. coli_ BL21(DE3)PLysS at 30°C and 18°C by IPTG induction in LB media and SHuffle® T7 at 37°C by autoinduction in AIMTB. Protein was isolated and analysed by SDS-PAGE and Coomassie staining.

Truncated _JEV_NS1, full length _JEV_NS1' and its truncations

C-_JEV_NS1 was expressed by autoinduction in AIMTB in _E. coli_ BL21(DE3) at 16°C, 25°C, 30°C, and 37°C and expressed in SHuffle® T7 at 25°C, 30°C, and 37°C. N-_JEV_NS1, DII-III-_JEV_NS1, _JEV_NS1', DII-III-_JEV_NS1', and C-_JEV_NS1' were expressed in _E. coli_ BL21(DE3) by AIMTB at 18°C, 30°C, and 37°C. Protein was analysed by initial lysis buffer screening, SDS-PAGE, and Coomassie staining.

Chaperone proteins co-expression

Full length NS1 with different fusion proteins, MBP, SUMO, GST, and HIS, and NS1 C-terminus were co-expressed with bacteria chaperone protein vectors, pG-KJE8,

pGro7, pKJE7, pTf16, and pG-Tf2 (TaKaRa chaperone plasmid set, cat.#3340). First, each chaperone protein vector that contains different chaperone proteins (see **Table 4.3**) was transformed into *E.coli* BL21(DE3), and/or K12 Shuffle T7. The selection medium was supplemented with 20 µg/ml chloramphenicol as all of the vectors resistance to. Then the competent cells were prepared as explained in section 3.2.3. After that, the competent cells were retransform with the interested plasmids. Cells were grown in ZYM5052 media (Studier, 2014) containing appropriate induction compounds (see Table 4.3) at 16°C, 25°C, 30°C, and 37°C for protein expression. Protein was isolated and analysed by SDS-PAGE and Coomassie staining.

Mammalian cell expression

The experiment was conducted at OPPF-UK and followed OPPF-UK mammalian expression standard protocols. Human embryonic kidney cell line (HEK 239T) was used and cultured in Dulbecco's modified Eagle medium (DMEM) supplemented with non-essential amino acids, 2 mM L-glutamine, and fetal calf serum at required concentration. In brief, HEK 239T was transiently transfected by using GeneJuice® transfection reagent (Novagen) with 1 µl vector DNA (vector construct 5-11, **Table 3.3**), which prepared by using PureLink® HiPure Plasmid Megaprep Kit (Invitrogen), at 2.66:1 ratio of GeneJuice to DNA (w/w) per 1 ml culture media. Transfection reaction was incubated for 3 days at 37°C. Kifunensine at concentration 1 mg/L was presented during protein expression. Cell supernatant (media) was collected, and cells were lysed by freeze-thaw method and were centrifuged to remove cell debris. Then cell supernatant and cells were analysed separately to indicate secreted products and intracellular soluble products, respectively, on SDS-PAGE. The expression was analysed by the anti-histidine Western blot. For large scale

purification, crude protein was purified by Ni-NTA column (His Trap FF, GE Healthcare) and followed with HiLoad Superdex 200 16x600 mm column (GE Healthcare). The experiments were conducted by Dr. Joanne Nattleship and Dr. Nahid Raman-huq.

Insect cell expression

The experiment was conducted at OPPF-UK and followed OPPF-UK Baculovirus expression standard protocols. *Spodoptera frugiperda* (Sf9) cells were cultured in Sf-900™ II serum free medium (Invitrogen). Bacmid was prepared by using BacMax kit (Cambio) and linearized by Bsu36I restriction enzyme (New England Biolabs). Bacmid with target DNA (vector construct 5-11, **Table 3.3**) was transfected into Sf9 cells by using FuGENE® HD transfection reagent (Roche). Transfection reaction was incubated for 5-6 days at 27°C and then P0 viruses were harvested. Sf9 cells were infected with the P0 viruses to obtain P1 viruses. Protein expression screening was conducted by infection of the P1 viruses to Sf9 cells and incubated for 3 days at 27°C. Intracellular soluble protein was analysed by Ni-NTA affinity beads and followed by Western blot against histidine. The experiments were conducted by Dr. Joanne Nattleship and Dr. Nahid Raman-huq.

Table 4.2 Protein expression trial conditions.

Protein construct	Strain	Induction	Temperature (°C)
N-HIS-*JEV*NS1	BL21(DE3)PLysS	IPTG	30 18
	BL21(DE3)	IPTG	25, 30, 37
	SHuffle® T7	auto	25, 30, 37
	BL21(DE3) with chaperone plasmids	auto	16, 25, 30, 37
N-HIS-MBP-*JEV*NS1 N-HIS-GST-*JEV*NS1 N-HIS-SUMO-*JEV*NS1	BL21(DE3)PLysS	IPTG	18, 30
	SHuffle® T7	auto	37
	BL21(DE3) with chaperone plasmids	auto	16, 37
C-*JEV*NS1	BL21(DE3)	auto	16, 25, 30, 37
	SHuffle® T7	auto	25, 30, 37
	BL21(DE3) with chaperone plasmids	auto	16, 25, 30, 37
N-*JEV*NS1 DII-III-*JEV*NS1 *JEV*NS1' DII-III-*JEV*NS1' C-*JEV*NS1'	BL21(DE3)	auto	18, 30, 37
HBM/C-HIS-*JEV*NS1 uPase/C-HIS-*JEV*NS1 uPase/N-HIS-*JEV*NS1 HBM/N-HIS-MBP-*JEV*NS1 HA/N-HIS-MBP-*JEV*NS1	HEK239T	-	37
Ac64/N-HIS-MBP-*JEV*NS1 uPase/N-HIS-MBP-*JEV*NS1	Sf9	-	27

Table 4.3 Chaperone protein vectors components.

Protein vector	Chaperone/MW (kDa)	Antibiotic resistance	Induction compound
pG-KJE8	dnaK/70-dnaJ/40-grpE/22-groES/10-groEL/60	20 µg/ml of chloramphenicol	0.5 mg/ml L-arabinose, 5 ng/ml tetracycline
pGro7	groES/10-groEL/60	20 µg/ml of chloramphenicol	0.5 mg/ml L-arabinose
pKJE7	dnaK/70-dnaJ/40-grpE/22		
pTF16	groES/10-groEL/60-tig/56		
pG-Tf2	Tig/56	20 µg/ml of chloramphenicol	5 ng/ml tetracycline

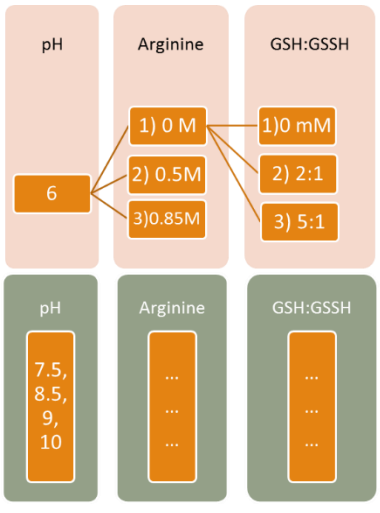
***JEV*NS1 protein refolding buffer screening**

The inclusion bodies from IPTG induction were resuspended in resolubilization buffer, TBS-Gnd·HCl containing 20 mM Imidazole (20 mM Tris-HCl pH 7.5, 150 mM NaCl, 6 M Guanidine hydrochloride (Gnd·HCl), 5 mM Dithiothreitol (DTT)), at wet weight concentration of 0.1 g/ml and incubated on roller at room temperature overnight. The protein was purified using 1 ml Ni-NTA column chromatography (GE Healthcare Life Science). The column was equilibrated with 5-10 ml re-solubilization buffer before the protein sample was loaded onto the column. The column was then washed with 30 ml of re-solubilization buffer and eluted by gradient steps of TBS-Gnd·HCl with 50 increments of imidazole concentration from 50-500 mM (5 ml for each step) or a single step of 5-10 ml of TBS-Gnd·HCl contained 500 mM Imidazole. Protein concentration in the eluted fraction was quantified by using BCA assay (Pierce™ BCA Protein Assay Kit), and protein was refolded by dilution method to dilute out Gnd·HCl. The eluted NS1 was diluted 10 times in refolding buffer for 36-48

hours at room temperature. Totally 45 conditions of refolding buffer were screened to find the right buffer condition in 1 ml reaction by varying 3 factors, including pH (0.1 M of 2-(N-morpholino) ethanesulfonic acid (MES) pH 6, Tris-HCl pH 7.5, 8.5, 9, and 3-(Cyclohexylamino)-1-propanesulfonic acid (CAPS) pH 10), arginine concentration (0 M, 0.5 M and 1 M), and glutathione concentration (0 mM, 3 mM at 2:1, and 5:1 reduced to oxidized) (**Figure 4.1**). The reactions were centrifuged at 16000xg for 10 minutes or filtered with 0.45 μm filter to separate the aggregated protein. The supernatant was analysed by SDS-PAGE. The positive condition was chosen and expanded to large scale experiment. The selected refolding buffer from the screening result was further tested with different refolding methods. First, slow dilution refolding. After the denatured NS1 was purified from NiNTA column, the protein was 10 times diluted with either slowly added drop-by-drop of protein to refolding buffer (0.1 M MES pH 6, 0.85 M arginine, 3 mM at 2:1 reduced to oxidized glutathione) or the refolding buffer to protein. Second, refolding on the column. Denatured full length NS1 was prepared as mention above and applied to 5 ml Ni-NTA column. 400 ml of refolding buffer (0.1 M MES pH 6, 0.85 M arginine, 3 mM at 2:1 reduced to oxidized glutathione) was washed through the column and the protein was eluted with 30 ml of refolding buffer contained 500 mM imidazole. NS1 protein was subjected to Superdex 200 10x300 mm size-exclusion chromatography column equilibrated with TBS buffer.

Refolded *JEV*NS1 protein purification

Refolded protein (10 times dilution method) was purified with different methods. First, the refolded protein was dialyzed against TBS without Gnd·HCl to remove the remaining Gnd·HCl. Second, the refolded protein was diluted to reduce the imidazole concentration to 20 mM and subjected to Ni-NTA column again and eluted with TBS buffer supplemented with 500 mM Imidazole. The protein was concentrated with 3000 Da MWCO centrifugal concentrator and protein concentration was quantified by the scanning wavelength from 250 to 600 nm. The concentration was calculated using the Beer-Lambert's law. Third, refolded NS1 protein was subjected to Ni-NTA column chromatography similar to the second method, but the protein was eluted from the column in 1 ml fractions for 50 ml to help concentrate the protein. Forth, concentration in refolding buffer. The NS1 full length was refolded in 2 different refolding buffers chosen from previous screening experiment (0.1 M MES pH 6 or 0.1



611	711	811	911	X11
612	712	812	912	X12
613	713	813	913	**X13**
621	721	821	921	X21
622	722	822	922	X22
623	723	**823**	923	X23
631	731	831	931	X31
632	**732**	**832**	**932**	X32
633	733	833	933	X33

Figure 4.1 Refolding buffer screen.

Forty five buffer conditions were screened by varying of pH, arginine, and GSH: GSSH. The pH 6, 7.5, 8.5, 9, and 10 are labelled 6, 7, 8 ,9, and X, respectively. Arginine concentration of 0 M, 0.5 M, and 0.85 M are labelled 1, 2, and 3, respectively. No glutathione is labelled 1. Glutathione concentration of 3 mM at 2:1 molar ratio GSH to GSSH is labelled 2 and at 5:1 is labelled 3. Positive results was coloured in blue.

M Tris pH 8 added with 0.85 M arginine, 3 mM at 2:1 reduced to oxidized) and concentrated by 10000 Da MWCO centrifuge concentrator. The Ni-NTA column chromatography after the refolding was omitted. NS1 protein was subjected to Superdex 200 10x300 mm size exclusion chromatography column equilibrated with TBS buffer.

4.2 Results

4.2.1 Comparison of NS1 sequences

Among the JEV strains from different genotypes, NS1 sequences are almost identical and the percent identity ranges from 89% to 100% (**Table 4.1**). When compared to viruses in the same serocomplex, which contain *West Nile, Kunjin, Murray Valley encephalitis, and St Louis encephalitis virus*, the protein sequences average about 75% identity. The value reduces to 50% when aligned with other viruses in the same genus such as *dengue, tick-borne encephalitis, Yellow fever, and Zika virus*. With the high percent identity (89-100%), only one JEV strain can represent all genotypes, and comparable to the remaining in the genus (50-70%). JEV SA-14 was used as the template for this study. The advantage of the conserved sequence makes the findings applicable to all genotypes.

4.2.2 Expression of *JEV*NS1 full length protein with histidine tag

The soluble full length protein expressed in *E. coli* BL21(DE3)PLysS could only be detected at 30°C expression. No soluble protein expression was observed at 18°C analysed by SDS-PAGE (**Figure 4.2a**). However, the protein was isolated in pellet fraction indicating the insoluble protein expression. Expression of *JEV*NS1 in *E. coli* BL21(DE3) yield non-soluble protein from all the temperatures tested, 25°C, 30°C, and 37°C (**Figure 4.2b**). The best yield was obtained from expression at 37°C, which greater than expression in BL21(DE3)PLysS. A small amount of soluble protein was found in this *E. coli* strain expression (**Figure 4.2b**). *E. coli* SHuffle®T7 expression was also succeeded at 25°C, 30°C, and 37°C, and the best yield was at 30°C (**Figure 4.2c**). However, none of these conditions could produce soluble protein.

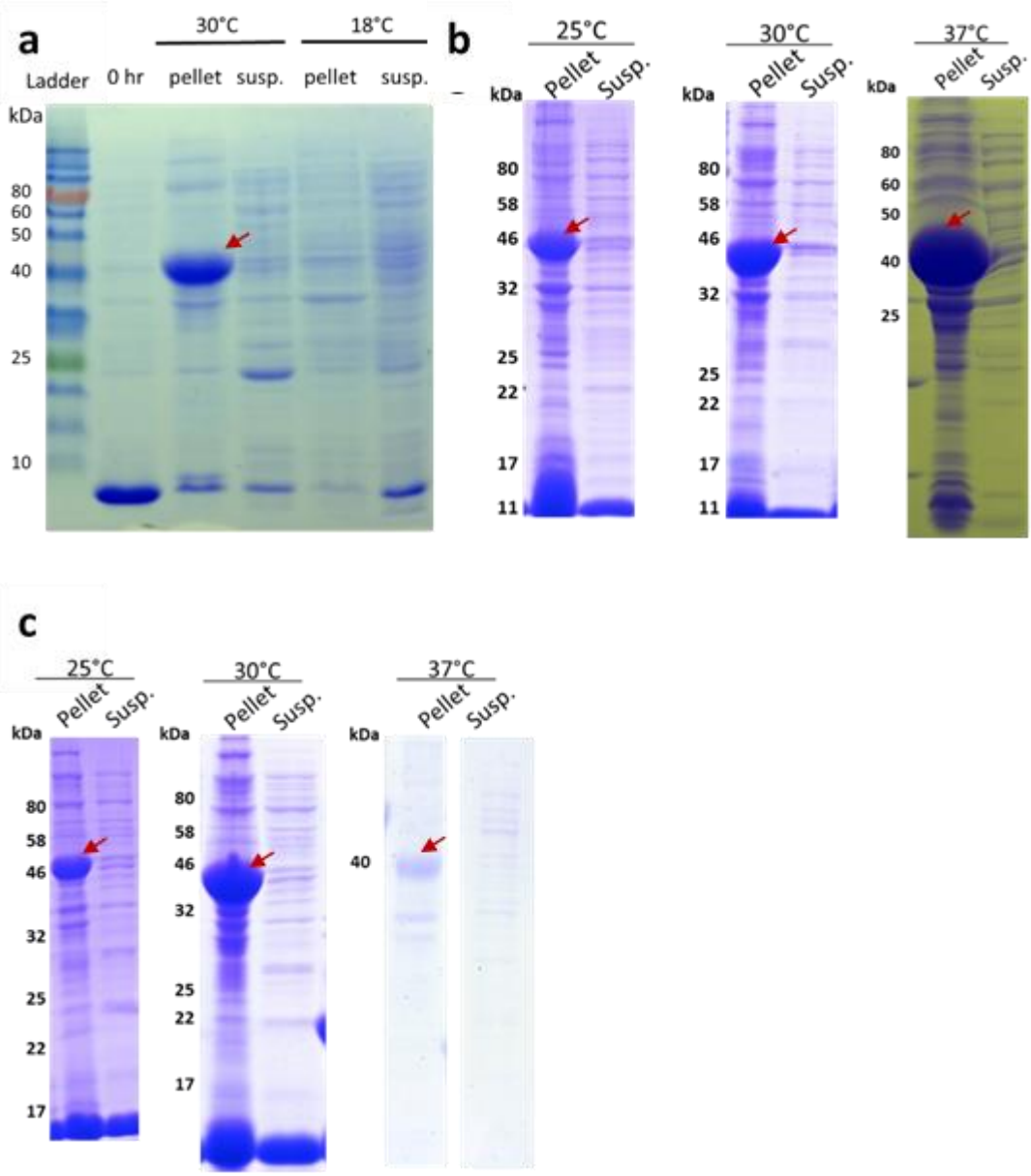


Figure 4.2 SDS-PAGE analysis of *JEV*NS1 full length with histidine tag protein expression.

(a) *JEV*NS1 expressed in BL21(DE3)PLysS at 18 and 30°C. Sample before induction is shown in 0 hr lane. **(b)** *JEV*NS1 expressed in BL21(DE3) at 25, 30, and 37°C. **(c)** JEVNS1 expressed in SHuffle®T7 at 25, 30, and 37°C. Suspension fraction is labelled susp. NS1 protein is indicated with red arrows.

4.2.3 Expression of *JEV*NS1 full length protein with fusion tags

With fusion tags, the theoretical size of the target protein was increased to 81.9 kDa for MBP, 67 kDa for GST, and 52.9 kDa for SUMO. Bands at expected sized were observed only in the pellet fraction of BL21(DE3)PLysS at 30°C and SHuffle®T7 at 37°C. Protein not effectively expressed in BL21(DE3)PLysS at 18°C. No expected bands was observed analysed by SDS-PAGE (**Figure 4.3**).

4.2.4 Expression of truncated *JEV*NS1, full length *JEV*NS1' protein and its truncations

The theoretical size of each protein constructs is 20.57 kDa for N-*JEV*NS1, 38 kDa for DII-III-*JEV*NS1, 20 kDa for C-*JEV*NS1, 45 kDa for *JEV*NS1', 43 kDa for DII-III-*JEV*NS1', and 26 kDa for C-*JEV*NS1'. Majority of C-*JEV*NS1 expressed in BL21(DE3) was in pellet fractions at all temperatures tested. Very small amounts of protein was found in supernatant fractions at 16°C and 25°C (**Figure 4.4a**). A similar result was found in the expression of C-*JEV*NS1 in SHuffle® T7 at 25°C and 30°C (**Figure 4.4b**).

The growth of DII-III-*JEV*NS1 and C*JEV*NS1' at 18°C was very slow even after 24 hours expression. Thus, they were excluded from the lysis buffer screening experiment (**Figure 4.4c**). However, they grew normally at 30°C and 37°C. Analyse by SDS-PAGE indicates proteins were expressed as inclusion bodies. Overall, overexpression of the truncated protein obtained quite a low yield except for the C-*JEV*NS1. Tiny amounts of protein was found in suspension fractions of DII-III-*JEV*NS1' expressed at 18°C in 7.5N lysis buffer (**Figure 4.4c**), N-*JEV*NS1 and C-*JEV*NS1' expressed at 30°C in 7.5N lysis buffer (**Figure 4.5**), C-*JEV*NS1' expressed at 30°C in 0.5U lysis buffer, and *JEV*NS1'

expressed at 37°C in 7.5N lysis buffer (**Figure 4.6**). Protein solubilized in a buffer contained urea indicating that it may need to be refolded.

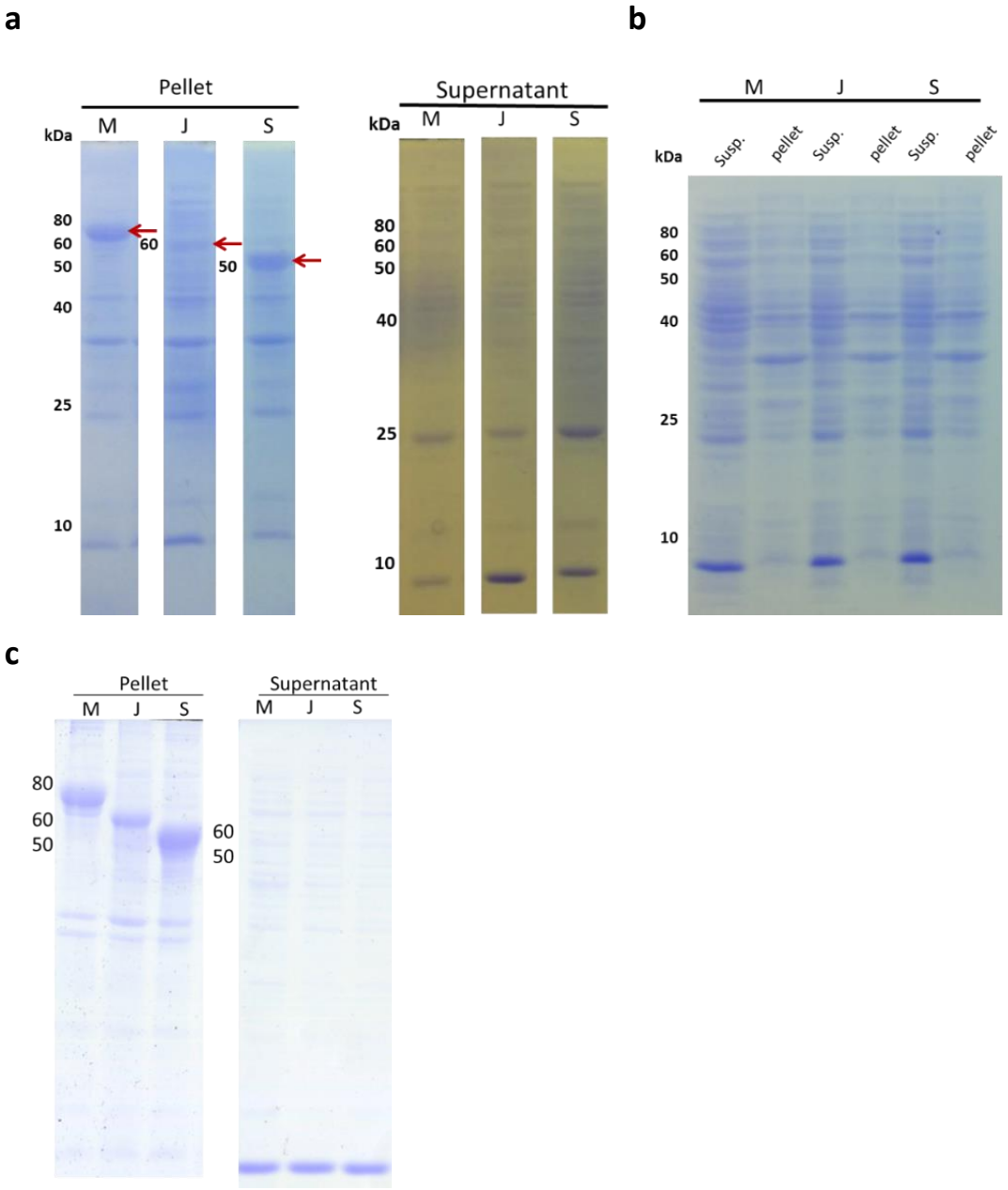


Figure 4.3 SDS-PAGE analysis of *JEV*NS1 full length with fusion tags protein expression.

(a) *JEV*NS1 expressed in BL21(DE3)PLysS at 30°C. **(b)** *JEV*NS1 expressed in BL21(DE3)PLysS at 18°C. **(c)** *JEV*NS1 expressed in SHuffle®T7 at 37°C. Fusion tag MBP, GST, and SUMO were labelled M, J, and S, respectively, according to the vector name. NS1 protein is indicated with red arrows.

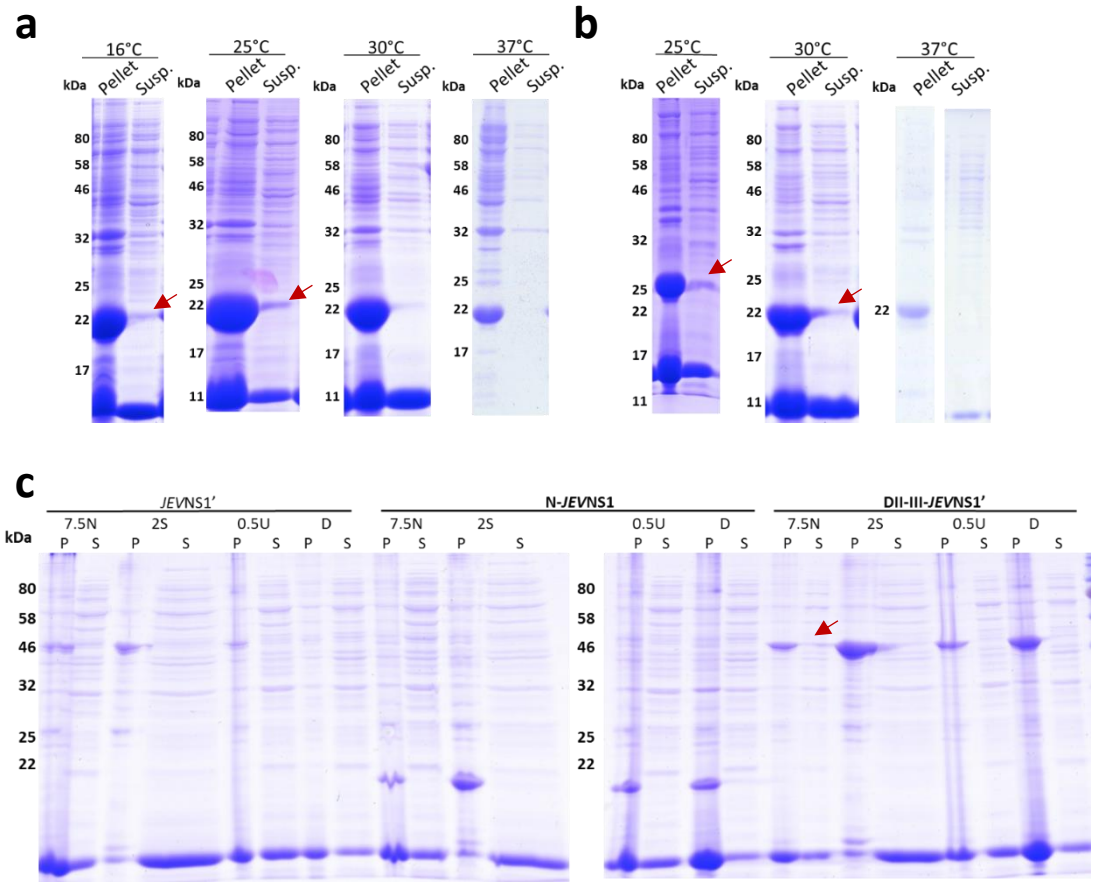


Figure 4.4 Truncated *JEV*NS1 expression trial.

(a) C-*JEV*NS1 expression in *E. coli* BL21(DE3) at 16°C, 25°C, 30°C, and 37°C. **(b)** C-*JEV*NS1 expression in *E. coli* SHuffle® T7 at 25°C, 30°C, and 37°C. **(c)** N-*JEV*NS1, *JEV*NS1', and DII-III-*JEV*NS1' expression in *E. coli* BL21(DE3) at 18°C. Protein solubility was analysed in 4 different buffer conditions. The buffer added with the following additives: no additive (**7.5N**), 2 M NaCl (**2S**), 0.5 M urea (**0.5U**), and 0.2% Triton X 100 (**D**). Soluble protein is indicated with red arrow.

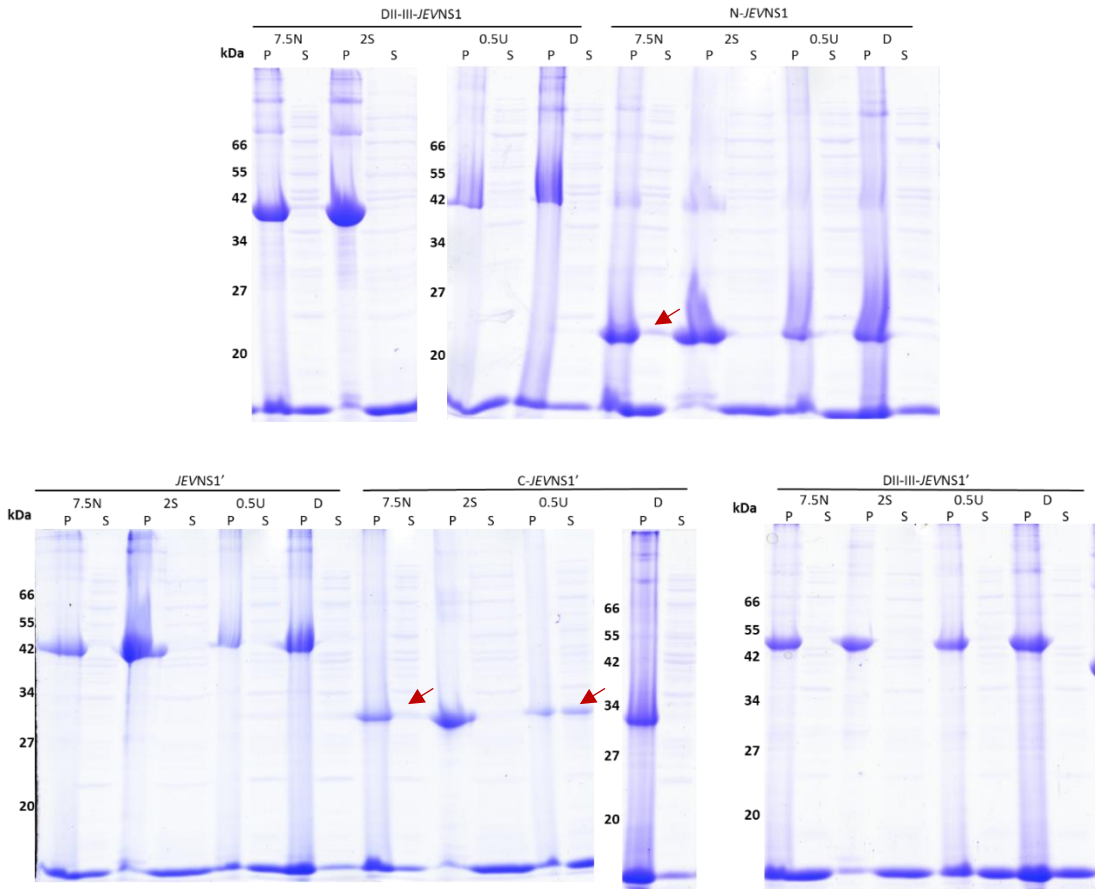


Figure 4.5 Truncated *JEV*NS1 expression trial at 30°C.

N-*JEV*NS1, DII-III-*JEV*NS1, *JEV*NS1', C-*JEV*NS1', and DII-III-*JEV*NS1' expression in *E. coli* BL21(DE3). Protein solubility was analysed in 4 different buffer conditions, which added with the following additives: no additive (**7.5N**), 2 M NaCl (**2S**), 0.5 M urea (**0.5U**), and 0.2% Triton X 100 (**D**). Soluble protein is indicated with red arrow.

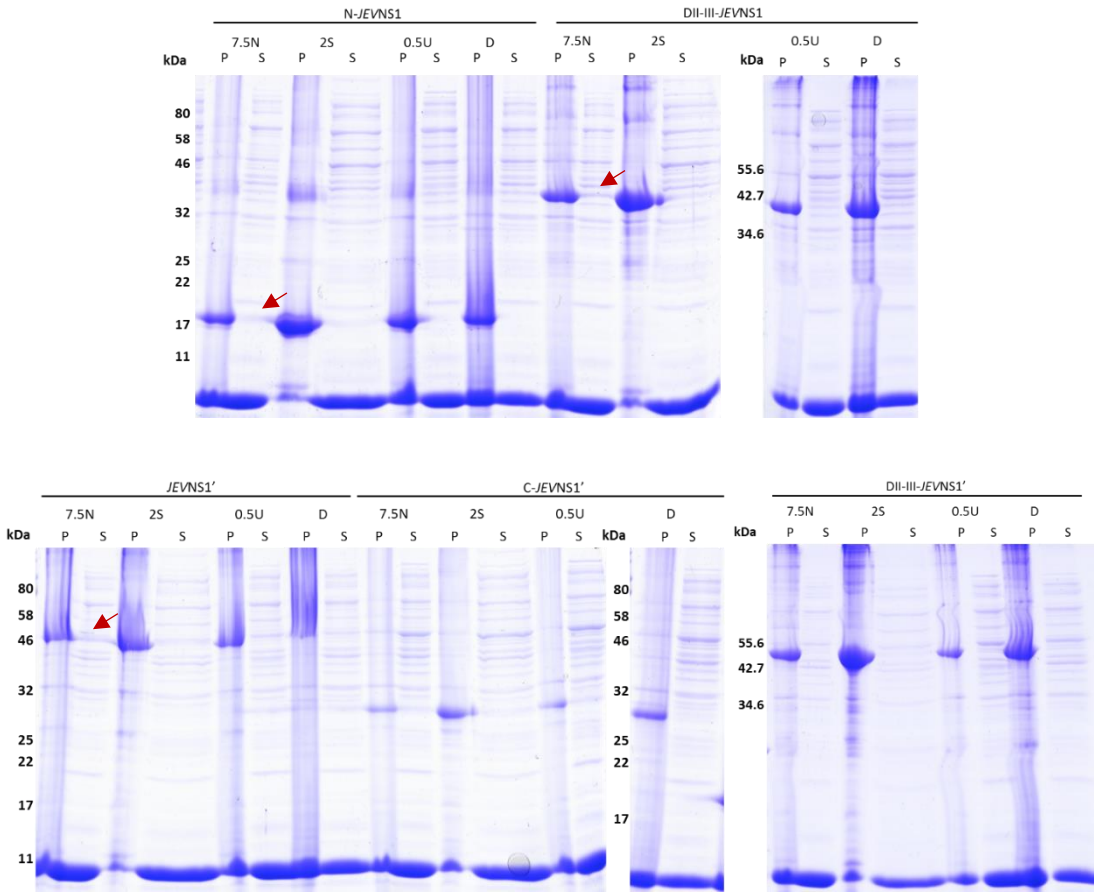


Figure 4.6 Truncated *JEV*NS1 expression trial at 37°C.

N-*JEV*NS1, DII-III-*JEV*NS1, *JEV*NS1', C-*JEV*NS1', and DII-III-*JEV*NS1' expression in *E. coli* BL21(DE3). Protein solubility was analysed in 4 different buffer conditions. Four lysis buffers were the buffer added with the following additives: no additive (**7.5N**), 2 M NaCl (**2S**), 0.5 M urea (**0.5U**), and 0.2% Triton X 100 (**D**). Soluble protein is indicated with red arrow.

4.2.5 Expression of *JEV*NS1 and C-*JEV*NS1 protein with chaperone proteins

Multiple bands were observed because of the expression of chaperone proteins (see **Table 4.3** for the molecular weight of each chaperone protein). *JEV*NS1 co-expression with chaperone plasmid pTF16 and pG-Tf2 grew slowly at 16°C (**Figure 4.7a**) and were eliminated from SDS-PAGE analysis. All cells containing chaperone plasmids grew normally at 25°C (**Figure 4.7b**), 30°C (**Figure 4.7c**), and 37°C (**Figure 4.7d**). Protein was isolated mainly in the pellet fraction and small bands of the same size were found in suspension fractions at 16°C, 25°C and 30°C expression (**Figure 4.7c**).

The growth was again slow at 16°C expression of *JEV*NS1 with fusion tags and some of them were excluded out including MBP-*JEV*NS1 co-expressed with chaperone plasmid pGro7 and pTF16, GST-*JEV*NS1 with pTF16, and SUMO-*JEV*NS1 with pKJE7, pTF16, and pG-Tf2. Proteins were expressed as inclusion bodies (**Figure 4.8a, b**) except for SUMO-*JEV*NS1 with pG-KJE8 and pGro7 that shown large amounts of protein in suspension fractions. However, Western blot analysis demonstrated that the bands in suspension fractions were not the target protein (**Figure 4.8a**).

C-*JEV*NS1 co-expressed well with chaperone proteins at 16°C (**Figure 4.9a**), 25°C (**Figure 4.9b**), 30°C (**Figure 4.9c**), and 37°C (**Figure 4.9d**). However, almost all of C-*JEV*NS1 was in pellet fractions, and only small amounts were in suspension fractions especially at 16°C, 25°C and 30°C.

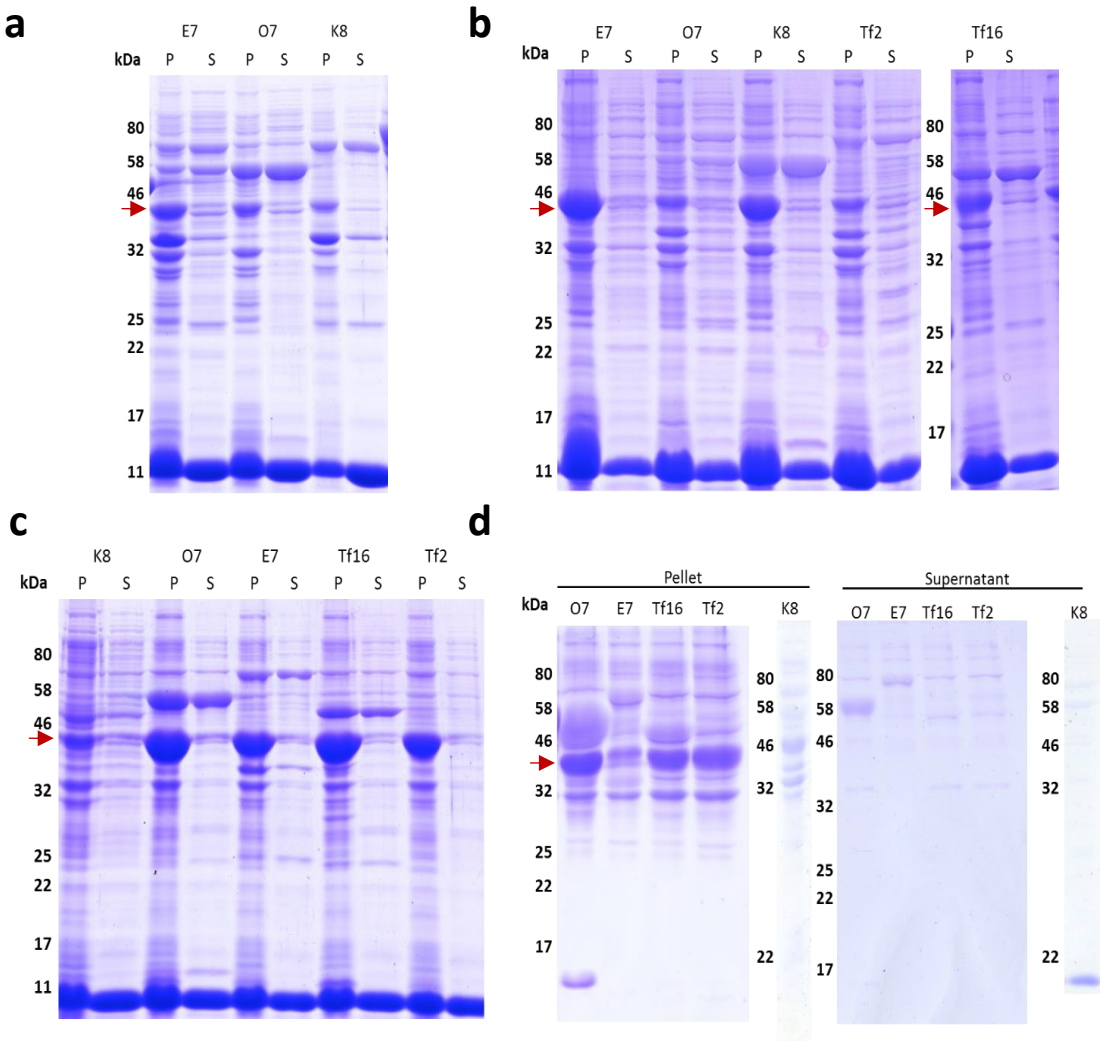


Figure 4.7 *JEV*NS1 chaperone proteins co-expression

at **(a)** 16°C, **(b)** 25°C, **(c)** 30°C, and **(d)** 37°C. Chaperone plasmid pG-KJE8, pGro7, pKJE7, pTf16, and pG-Tf2 were labelled K8, O7, E7, Tf16, and Tf12, respectively. NS1 full length size is indicate with red arrow.

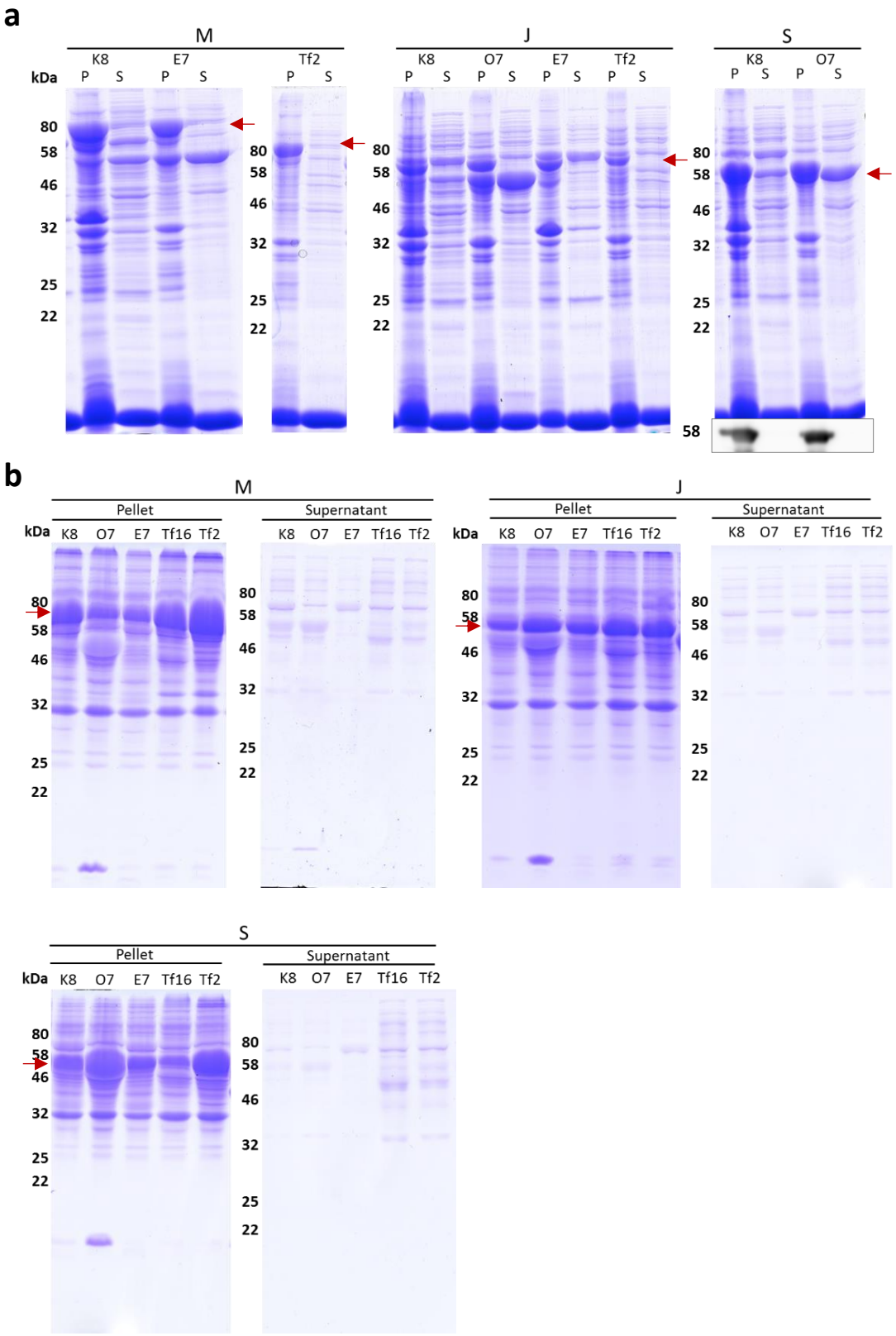


Figure 4.8 *JEV*NS1 with fusion tags chaperone proteins co-expression

at 16°C **(a)** and 37°C **(b).** Chaperone plasmid pG-KJE8, pGro7, pKJE7, pTF16, and pG-Tf2 were labelled K8, O7, E7, Tf16, and Tf12, respectively. Expected band sizes are indicated with red arrows.

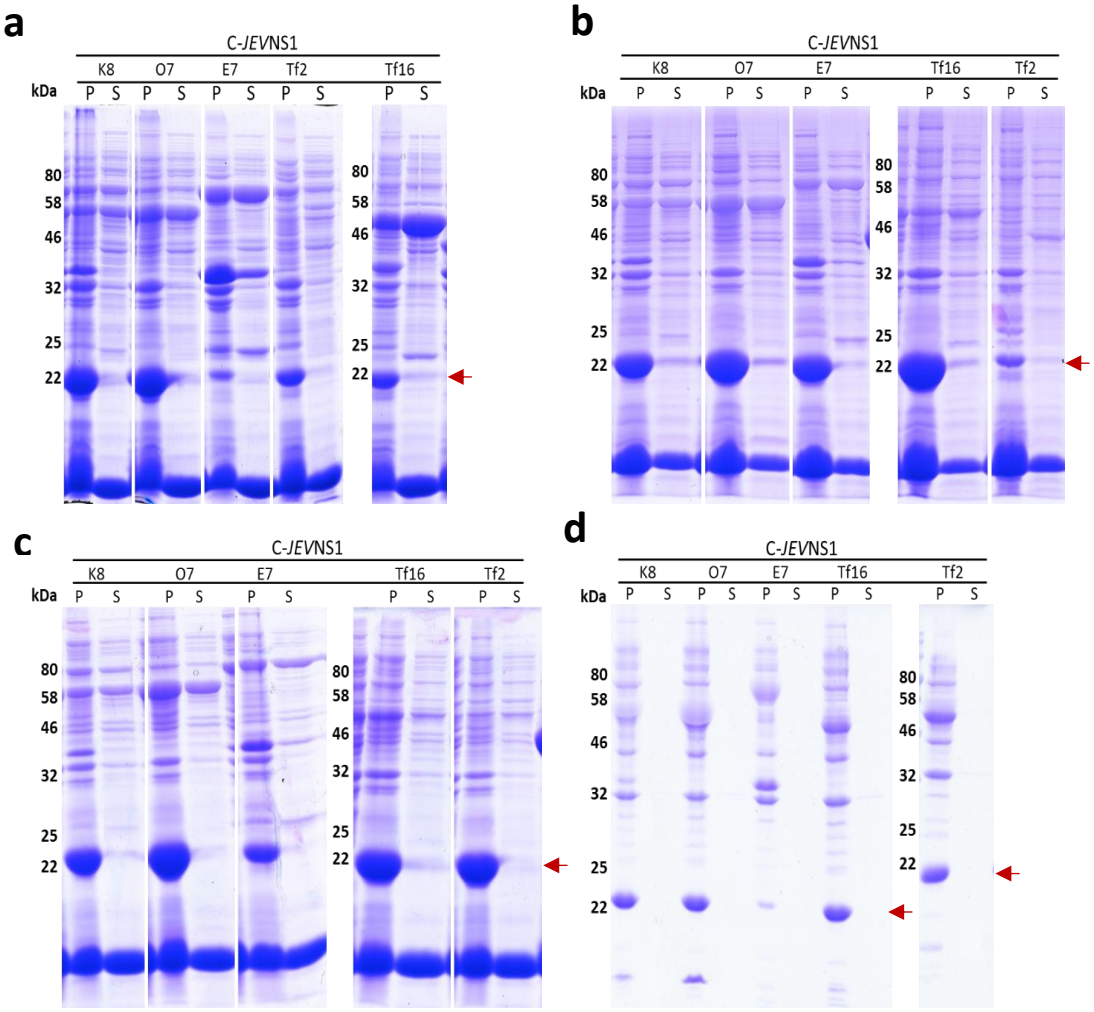


Figure 4.9 C-*JEV*NS1 chaperone proteins co-expression

at **(a)** 16°C, **(b)** 25°C, **(c)** 30°C, and **(d)** 37°C. Chaperone plasmid pG-KJE8, pGro7, pKJE7, pTf16, and pG-Tf2 were labelled K8, O7, E7, Tf16, and Tf12, respectively. NS1 C-terminal domain size is indicated with red arrow.

4.2.6 Expression of *JEV*NS1 protein in mammalian cells

E. coli codon optimized *JEV*NS1 with HBM/C-His, uPase/N-His and Ac64/N-His-MBP tags for mammalian cell expression were excluded because of the low sequencing quality check after cloning step. From Western blot analysis, native *JEV*NS1 with uPase/C-HIS was strongly expressed in both secreted and in-cell soluble protein (**Table 4.4 and Figure 4.10**, lane 2). There was a moderate expression of native *JEV*NS1 with HA/N-HIS-MBP, uPase/N-His-MBP (**Figure 4.10**, lane 3 and 15), and *E. coli* codon optimized *JEV*NS1 with uPase-N-His-MBP in media and cell lysis fractions (**Figure 4.10,** lane 19). Native *JEV*NS1 with HBM/C-HIS showed faint bands in both fractions (**Figure 4.10**, lane 8). Native *JEV*NS1 with HBM/N-HIS-MBP shown faint band only in secretion fraction (**Figure 4.10**, lane 20). *JEV*NS1 was subcloned into pOPINTTGneo vector. The pOPINTTGneo has a similar construct to pOPIN G (uPase/C-HIS) and contains RPTPmu secretion signal (**Table 3.3**). This new vector was believed to give a better protein expression. The new protein construct was used for large scale protein expression (2 liters). After Ni-NTA purification and size exclusion chromatography, *JEV*NS1 displayed multiple bands and some background in SDS-PAGE analysis (**Figure 4.11**, upper). Possibly due to too little or no protein expression that led to non-specific binding to the Ni-NTA column. Elution fraction B3-B10 was pooled together and concentrated to~1 ml. The concentration was ~5 mg/ml. Western blot analysis showed only faint bands in lane 9 (**Figure 4.11**, lower).

Table 4.4 Secretion screen in HEK293-T Cells.

Gene	Lane	SS/Tags	Secreted	Soluble in the cell
E. coli codon optimized	1	uPase/N-HIS-3c		
Native	2	uPase/C-HIS		
Native	3	HA/N-HIS-MBP-3C		
E. coli codon optimized	7	HBM/N-HIS-MBP-3C		
Native	8	HBM/C-HIS		
Native	9	Ac64/N-HIS-MBP-3C		
E. coli codon optimized	13	HA/N-HIS-MBP-3C		
Native	14	uPase/N-HIS-3c		
Native	15	uPase/N-HIS-MBP-3C		
E. coli codon optimized	19	uPase/N-HIS-MBP-3C		
Native	20	HBM/N-HIS-MBP-3C		Very faint

Note: Dark green indicates strong expression. Light green indicates low expression.

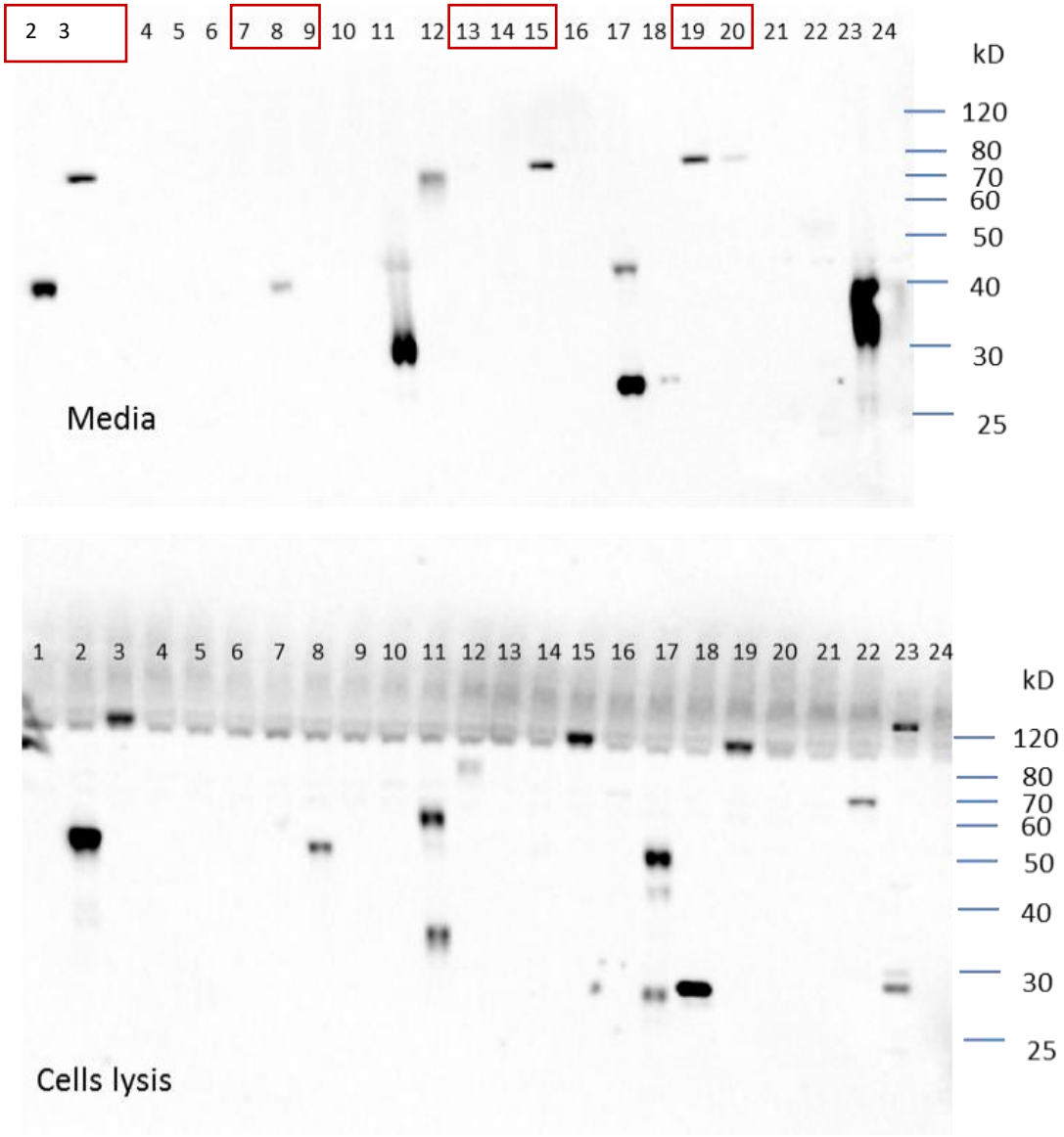


Figure 4.10 Western blot analysis of HEK239-T transient expression.

Secreted protein is labelled media (upper) and soluble protein in the cell is labelled cells lysis (lower). Protein in each lane is indicated in table 4.4.

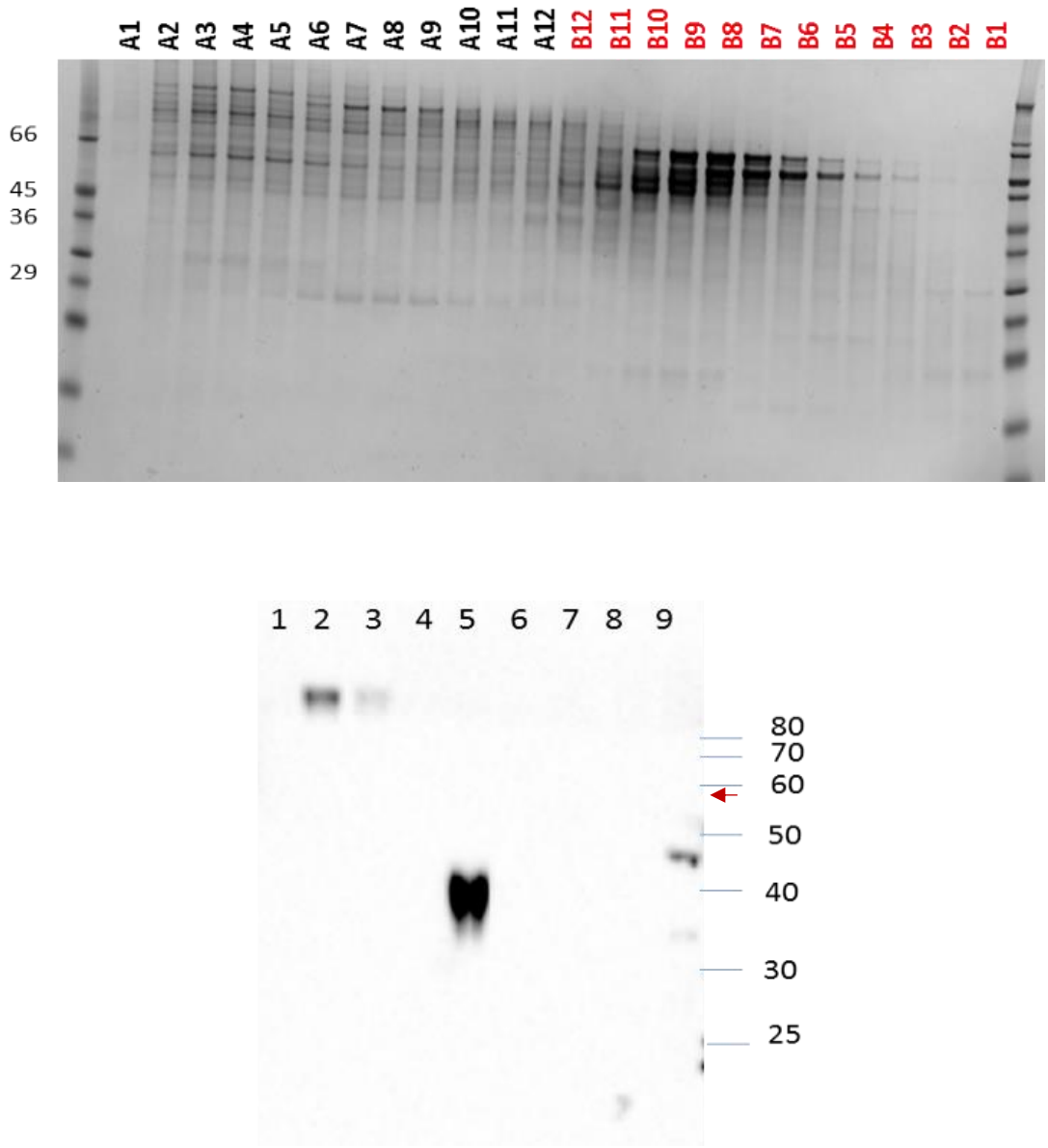


Figure 4.11 HEK239-T scaled up transient transfection.

HiLoad Superdex 200 16x600 mm column purification after Ni-NTA (upper). Western blot analysis. Native *JEV*NS1 was loaded in lane 9 indicated with red arrow. Positive control was loaded in lane 5 (lower).

4.2.7 Expression of *JEV*NS1 protein in insect cells

E. coli codon optimized *JEV*NS1 with Ac64/N-His-MBP was excluded from the expression screening because of the low sequencing quality. Western blot analysis showed that *E. coli* codon optimized *JEV*NS1 and native *JEV*NS1 with uPase/C-HIS and native *JEV*NS1 with HBM/N-HIS-MBG demonstrated low protein expression (**Table 4.5** and **Figure 4.12**, lane B1, H1, and C2). Higher expression was observed in *E. coli* codon optimized and native *JEV*NS1 with uPase/N-His-MBP. However, the expression was not strong enough to continue to scale up compared to green fluorescent protein (GFP) positive control (**Figure 4.12**, lane H2).

Table 4.5 Secretion screen in *Sf9* cells.

Gene	Lane	Vector	Secreted
E. coli codon optimized	A1	HBM/C-HIS	
E. coli codon optimized	B1	uPase/C-HIS	Very faint
E. coli codon optimized	C1	uPase/N-HIS-3c	
E. coli codon optimized	D1	HBM/N-HIS-MBP-3C	
E. coli codon optimized	E1	HA/N-HIS-MBP-3C	
E. coli codon optimized	F1	Ac64/N-HIS-MBP-3C	
E. coli codon optimized	G1	uPase/N-HIS-MBP-3C	
Native	H1	uPase/C-HIS	Very faint
Native	A2	HBM/C-HIS	
Native	B2	uPase/N-HIS-3c	
Native	C2	HBM/N-HIS-MBP-3C	Very faint
Native	D2	HA/N-HIS-MBP-3C	
Native	E2	Ac64/N-HIS-MBP-3C	
Native	F2	uPase/N-HIS-MBP-3C	
Blank	G2		
GFP	H2		

Note: Dark green indicates strong expression. Light green indicates low expression.

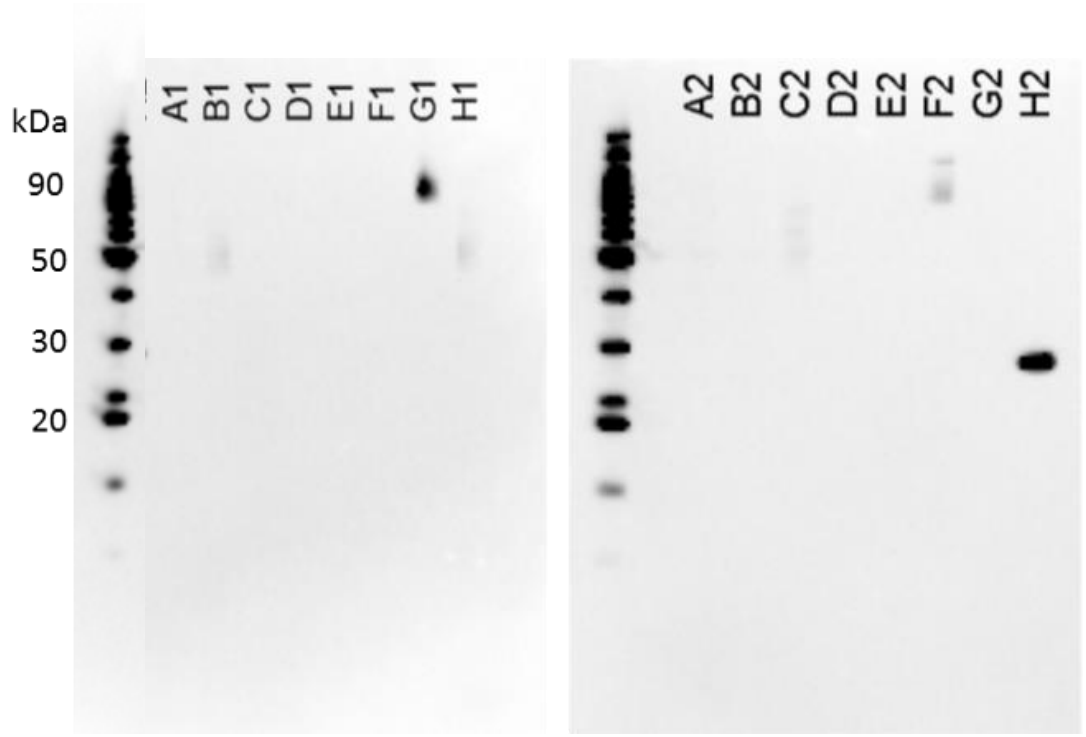


Figure 4.12 Western blot analysis of *Sf9* expression.

Protein in each lane is indicated in Table 4.5. The strongest expression is His-MBP-tagged proteins in lane F2 and G1. H2 is the GFP positive control.

4.2.8 *JEV*NS1 protein refolding buffer screening

From 45 conditions, 5 were succeed (**Figure 4.1 and Figure 4.13**) which were condition X13, 823, 632, 732, 832, and 932. The strongest band was acquired from condition 632 (0.1 M MES pH 6, 0.85 M arginine, and 3 mM glutathione at a 2:1 GSH to GSSH) and it was used for the large scale protein refolding and other refolding methods. Slow dilution (drop-by-drop) refolding did not give significant yield increase compared to one step dilution (data not shown). Refolding on column yielded a large amount of protein in elution fraction (**Figure 4.14a)**.

4.2.9 Refolded *JEV*NS1 protein purification

Refolded *JEV*NS1 sample with buffer condition 632 appeared a weak band at the size of ~46 kDa (**Figure 4.15a**, lane 1). Followed by filtration and dialysis, the band disappeared (**Figure 4.15a**, lane 2). It is possible that the protein already aggregated and was eliminated via filtration.

Ni-NTA purification after refolding step and elution with TBS buffer gave a clear band at 46 kDa (**Figure 4.15b**). However, it started to have some precipitation when stored at 4°C overnight and more debris was observed when the protein was concentrated to the volume about 1 ml.

Small fraction elution. The elution fractions precipitated immediately especially in the fraction 8 which had a high concentration of protein (**Figure 4.15c and d**).

Refolding on Ni-NTA column. A large amount of protein refolded on the column was eluted, but when elution profile of size exclusion chromatography was further

analysed, it become clear that protein eluted in a void volume of the column indicating protein aggregation (**Figure 4.14b**).

Concentration in refolding buffer. Protein precipitation happened during the concentration step, but not in massive level. In both refolding conditions, pH 6 and pH 8, the majority of the protein was eluted from SEC in the void volume indicated protein aggregation (**Figure 4.16a, b**). Even though NS1 protein was eluted in a few fractions confirmed by Western blot, each peak was not separated well (**Figure 4.16a, b**) and multiple bands were observed make it difficult to obtain a pure and homogenous sample for crystallization.

Even though the refolding buffer condition was found and several purification methods have been tried, none of these could give protein in quantity and purity for crystallization trials.

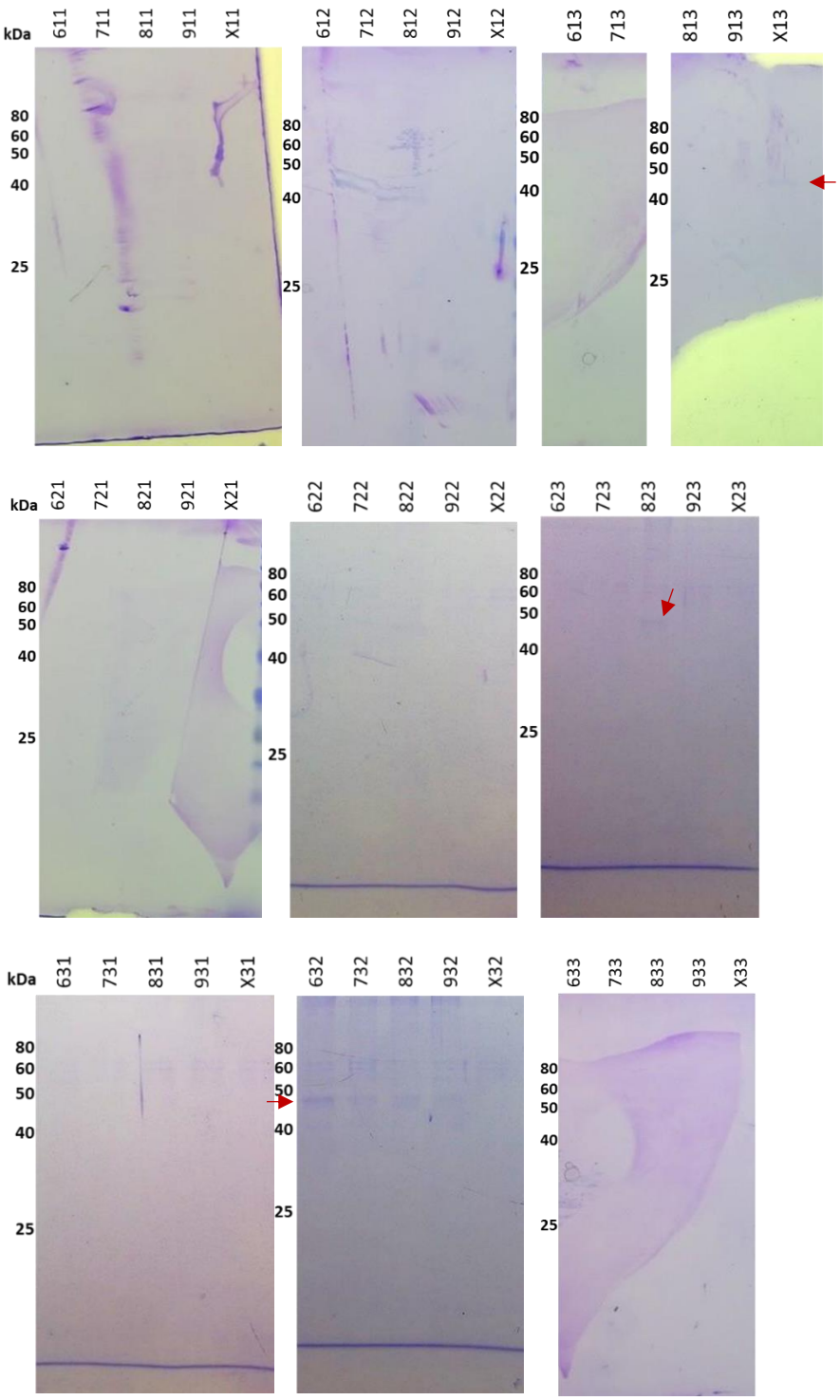


Figure 4.13 Protein refolding screen.

Three digits code indicate pH, arginine concentration, and glutathione ratio, respectively. pH 6, 7.5, 8.5, 9, and 10 are labelled 6, 7, 8 ,9, and X, respectively. Arginine concentration of 0 M, 0.5 M, and 0.85 M are labelled 1, 2, and 3, respectively. No glutathione is labelled 1. Glutathione concentration of 3 mM at 2:1 molar ratio GSH to GSSH is labelled 2 and at 5:1 is labelled 3.

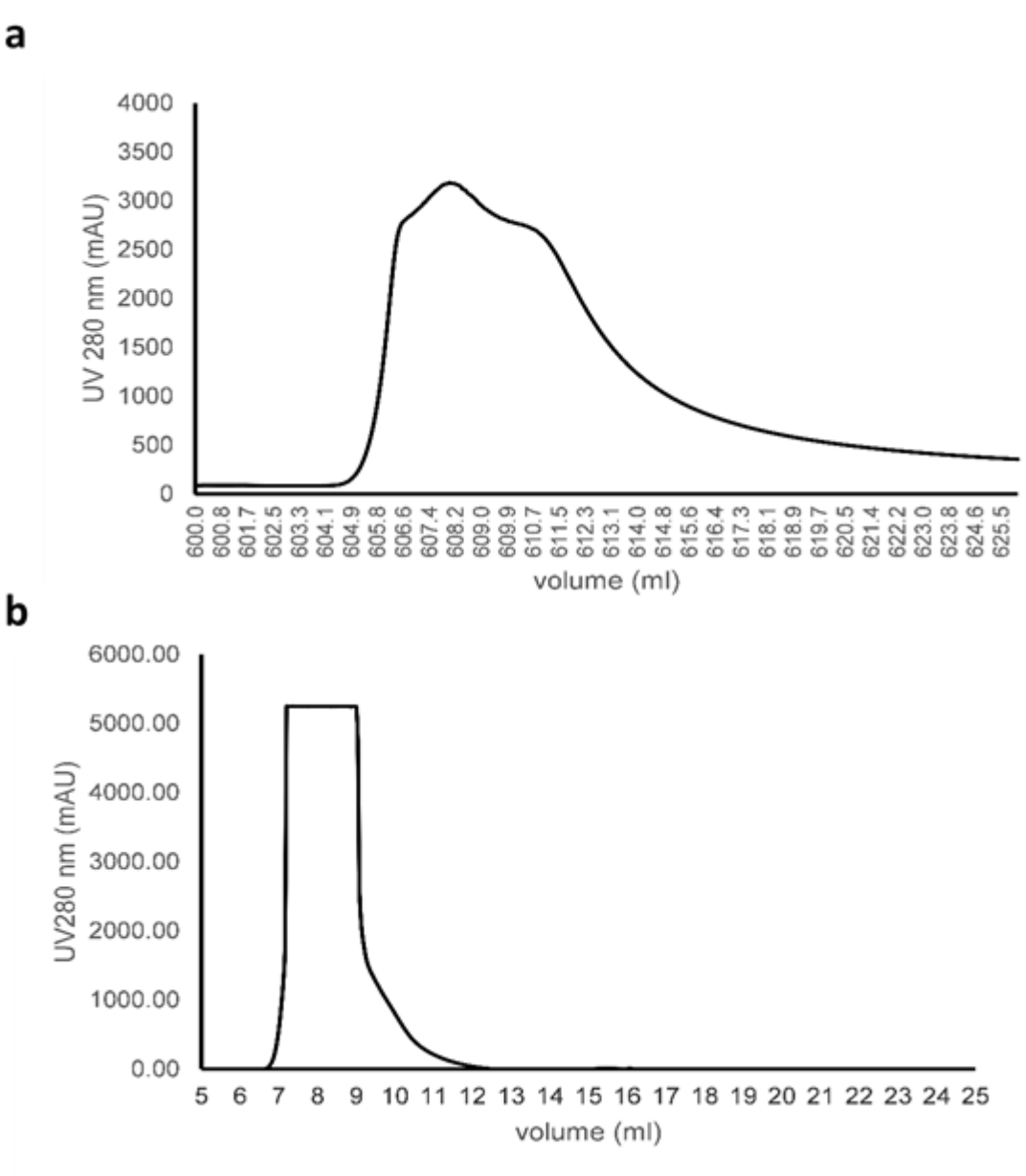


Figure 4.14 *JEV*NS1 refolding on column.

(a) Ni-NTA elution. Large amount of the protein was eluted by refolding buffer supplemented with 500 mM imidazole. **(b)** Refolding on column gel filtration profile. *JEV*NS1 protein was eluted in void volume of Superdex 200 10x300 mm column.

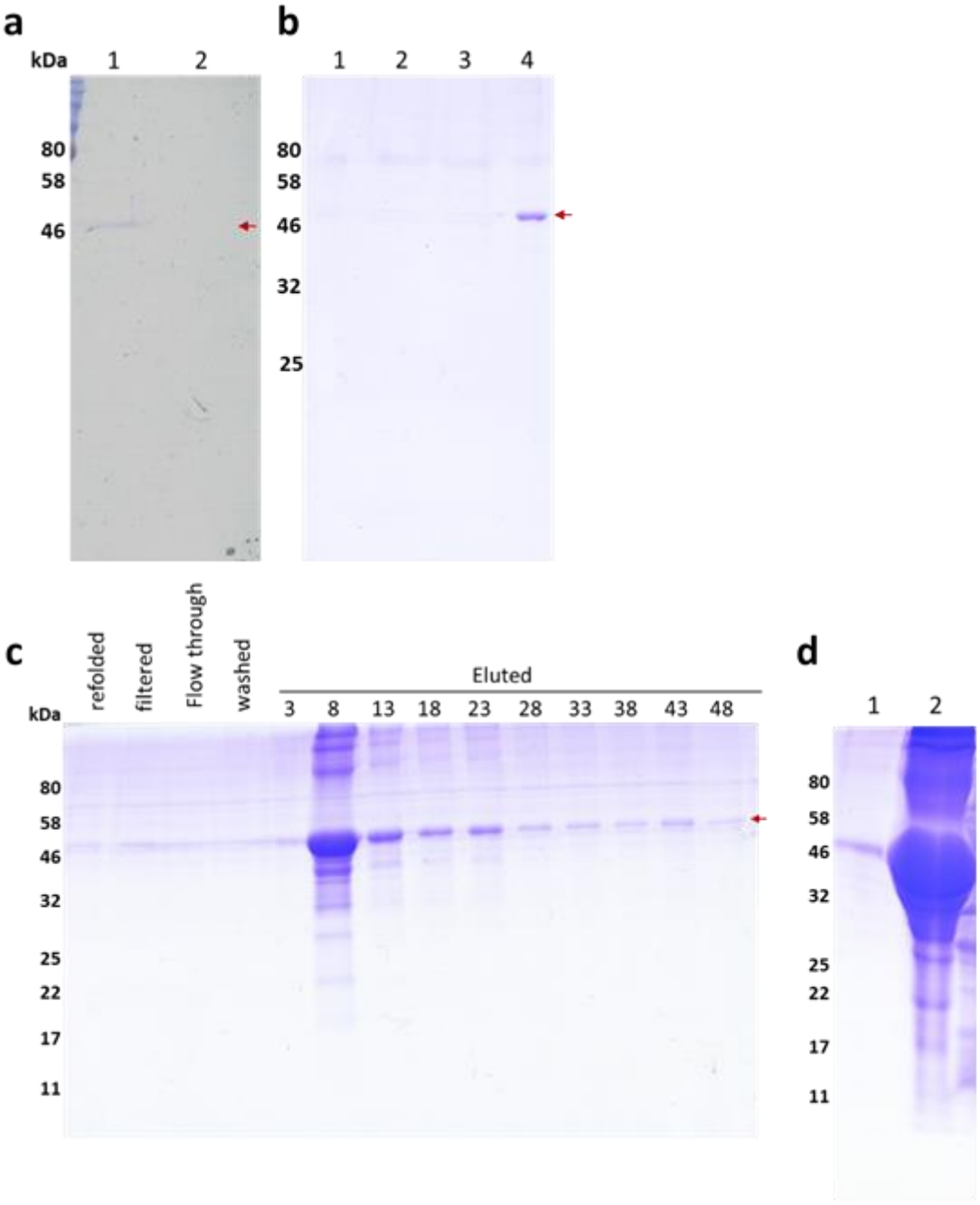


Figure 4.15 SDS-PAGE analysis of protein purification trials.

(a) Dialysis after protein refolding. Refolded protein before dialysis shows in lane 1 and after dialysis shows in lane 2. **(b)** Second Ni-NTA purification. Lane 1 is refolded protein. Lane 2 is flow-through fraction, lane 3 is washed fraction, and lane 4 is eluted fraction. **(c)** Small elution fraction. Protein was eluted with 500 mM Imidazole in TBS buffer for 50 ml (1 ml for each fraction). Elution fraction 3, 8, 13, 18, 23, 28, 33, 38, 43, and 48 were analysed. **(d)** Fraction number 8. Protein precipitation was removed by centrifugation. Lane 1 is supernatant fraction and lane 2 is precipitant.

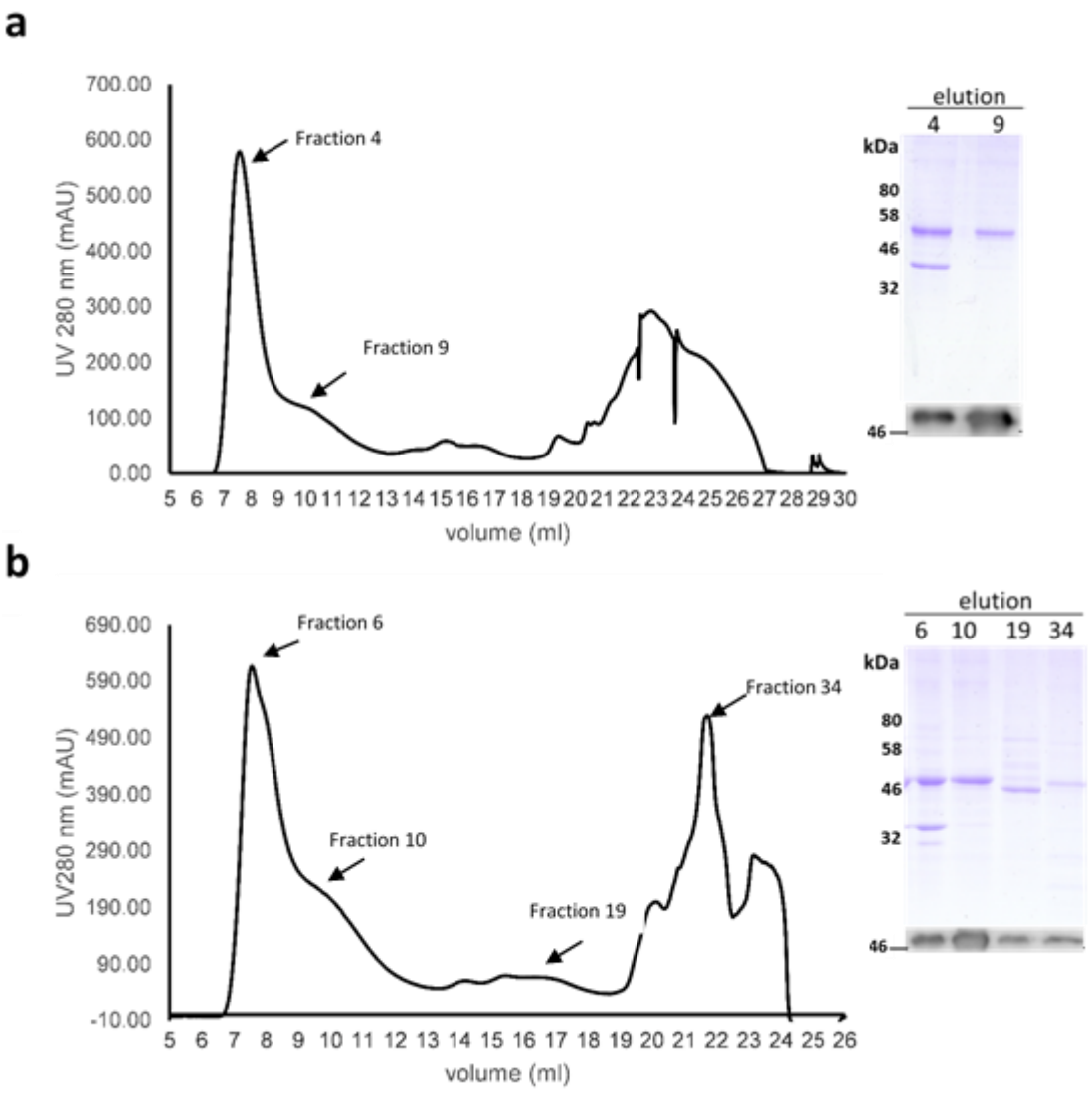


Figure 4.16 Size exclusion chromatography profiles of *JEV*NS1 after refolding

After refolding in buffer pH 6 **(a)** and 8 **(b)** show poor resolution of separation. Elution fractions were analysed by SDS-PAGE and Western blot (a, b).

4.3 Discussion

Protein expression in *E. coli* has advantages in that it is inexpensive and can give a good yield. However, the major drawback is the lack of post-translational modification found in more complex organisms. When heterologous proteins are expressed this can result in misfolding and failure to express the protein in the soluble fraction. Nowadays, there are many strategies to improve expression of the foreign protein in *E. coli* both engineering and non-engineering of the target protein. Some techniques are experimentally determined and there is no recipe that works for every protein such as fusion tag proteins and culture temperature. High throughput experiments would be more appropriate to test every possible method available. Thus, in this study, the methods were chosen based on the available materials and equipment. All the strategies used to modify protein expression in this study could not produce a satisfactory yield. This is not surprising as *JEV*NS1 protein carries 6 disulphide bonds and also 2 glycosylation sites which are not supported in *E. coli*.

Similar results were obtained from mammalian cell and insect cell expression systems. While the protein was expressed, the yield was too low and not worth scaling up. The conditions used in this study were developed from what has been reported to succeed (Brown et al., 2011). However, differences in small details could lead to different results such as signal sequence, vector construct, and culture techniques. Further modification of conditions to get a better protein yield might help and it is very promising.

The strategy was, therefore, changed to refold protein from *E. coli* inclusion bodies. Even though the refolding buffer condition was found, *JEV*NS1 aggregated during the

purification process. It is possible that the protein was not properly refolded. Typically, dilution refolding is done in 100 fold dilution, but in this study, it was only 10 fold dilution. Low dilution resulted in a moderate concentration of Gnd·HCl that may be too high for protein to refold and cause partial unfolding and keep the protein in the intermediate state which is prone to aggregation. Once the protein buffer was immediately changed, the aggregation might occur.

In the future, people might develop new strategies to produce protein effectively in bacteria and also in higher expression systems. New technologies may prove more efficient in large-scale protein expression and purification. Besides that, if we understand more about protein folding, we may able to better design a method to express and purify the protein.

4.4 Conclusion

In this study, soluble *JEV*NS1, *JEV*NS1', and the truncations of *JEV*NS1 and *JEV*NS1' were attempted to produce in *E. coli.* Many strategies were examined to improve foreign protein expression. A great deal of protein was expressed as inclusion bodies. Only some conditions could express a soluble form of the proteins but the yield was too low. The protein was refolded from inclusion bodies instead. Our own refolding protocol was created, but a lot of protein aggregated during protein purification processes. In parallel, mammalian and insect cell expression screen were performed at OPPF-UK. Some vector constructs could produce *JEV*NS1. However, the yield was low hence not pursued further.

Chapter 5 Structural study of C-*JEV*NS1 and C-*JEV*NS1'

Since the expression trials conducted in chapter 4 could not give the protein yield and purity required for crystallization, the strategy was changed to follow Edeling's work published in 2014 (Edeling et al., 2014). Refolding of WNV NS1 truncations had previously been successful (Chung et al., 2006b), but it was used for structural study for the first time in Edeling's publication. With highly conserved sequence, it is likely that the refolding method could work for *JEV*NS1 as well. The protein inclusion bodies were refolded by using the protocol modified from WNV NS1 and enterokinase refolding protocols (Edeling et al., 2014, Skala et al., 2013). Parts of the work reported in this chapter is now published (Poonsiri et al., 2018).

5.1 Methods

5.1.1 Plasmid construction

Synthetic *JEV*NS1 DNA optimized for *E. coli* in pRSET_A_A185 vector and *JEV*NS1' in pET-30a(+) were subcloned into pET303 to create C-*JEV*NS1 and C-*JEV*NS1' without histidine tag by using the primers indicated in **Table 3.4**.

5.1.2 Protein expression, refolding, and purification

Protein was expressed in *E. coli* by autoinduction and refolded using a modified method previously described in Edeling et al, 2014. Briefly, inclusion bodies were denatured in 7 M Gnd·HCl, 30 mM β_2ME and refolded by 100 times dilution in 400 mM L-arginine, 100 mM Tris pH 8.3, 2 mM EDTA, 0.5 mM GSSH, 5 mM GSH, 0.2 mM PMSF. After refolding, the protein was purified once by size exclusion

chromatography column, Superose 6 10x300 mm or superdex 200 10x300 mm equilibrated with 20 mM HEPES pH 7.4, 150 mM NaCl.

5.1.3 Protein oligomeric state analysis

Agilent Bio SEC-3 4.6x300 mm HPLC column (300 Å pore structure) was equilibrated with 100 ml (20 column volumes) of a filtered-degassed buffer, 20 HEPES pH 7.4 and 150 mM NaCl. The column was operated at a flow rate 0.25 ml/min, 35 °C. The purified C-*JEV*NS1 first peak, second peak, and total protein (80 µl each) at day 0, 3, 6, 10, and 13 after the purification were loaded onto the column. Signals at the wavelength of 220, 260, and 280 nm were observed.

5.1.4 Protein crystallization and data collection

The protein was concentrated to ~6 mg/ml (C-*JEV*NS1) or ~7 mg/ml (C-*JEV*NS1') and screened by using SaltRx, PEGRx, and Natrix screens from Hampton research, and Structure, PACT premier, and JCSG screens from Molecular Dimension. All the positive condition are given in **Appendix 3**. Needle crystals of C-*JEV*NS1 were produced from a range of salt concentrations, 0.7-1 M Ammonium sulphate or 1 M Li_2SO_4 and buffer pH, 0.1 M MES pH 5.5-6.7, while NS1' C-terminus which also crystallised in needle form crystallised in 2 M Ammonium sulphate and 5% propanol. The crystals were flash frozen in reservoir solution added with 20-25% ethylene glycol or glycerol (**Figure 5.1**).

X-ray data were collected at cryogenic temperature at 0.98 Å wavelength, at beamline PROXIMA 1 at Soleil synchrotron, France for C-NS1 protein and at beamline I02 at Diamond Light Source, UK for C-NS1' protein. Data reduction was carried out

by XDS programs for C-NS1 or iMosflm for C-NS1'. Scaling was performed by Scala or Aimless (Evans, 2011). Both protein structures were determined by molecular replacement using the structure of WNV NS1 C-terminal domain (PDB: 4OIE, sequence identity 73%) as a starting model by MOLREP in the CCP4 program package. The structure was refined by REFMAC5 and built in COOT. The C-NS1 data was collected twice from a single crystal at a time. The C-NS1 first model was refined with weight term of 0.01. The second data was refined against the first model with weight term of 0.2. The C-NS1' models were built separately from C-NS1. The reflection data was refined against the C-NS1' model with weight term of 0.15. Data collection and refinement statistics are shown in **Table 5.1**. Both of C-*JEV*NS1 and NS1' have good protein geometry and those statistical parameters are in high quality compared to protein structures at similar resolution. The C-*JEV*NS1 refinement statistic of Ramachandran plot is 95.98% favoured and 0% outliers. The MolProbity score is 1.6. The JEV NS1'-C refinement statistic of Ramachandran plot is 94.89% favoured and 0% outliers. The MolProbity score is 1.84.

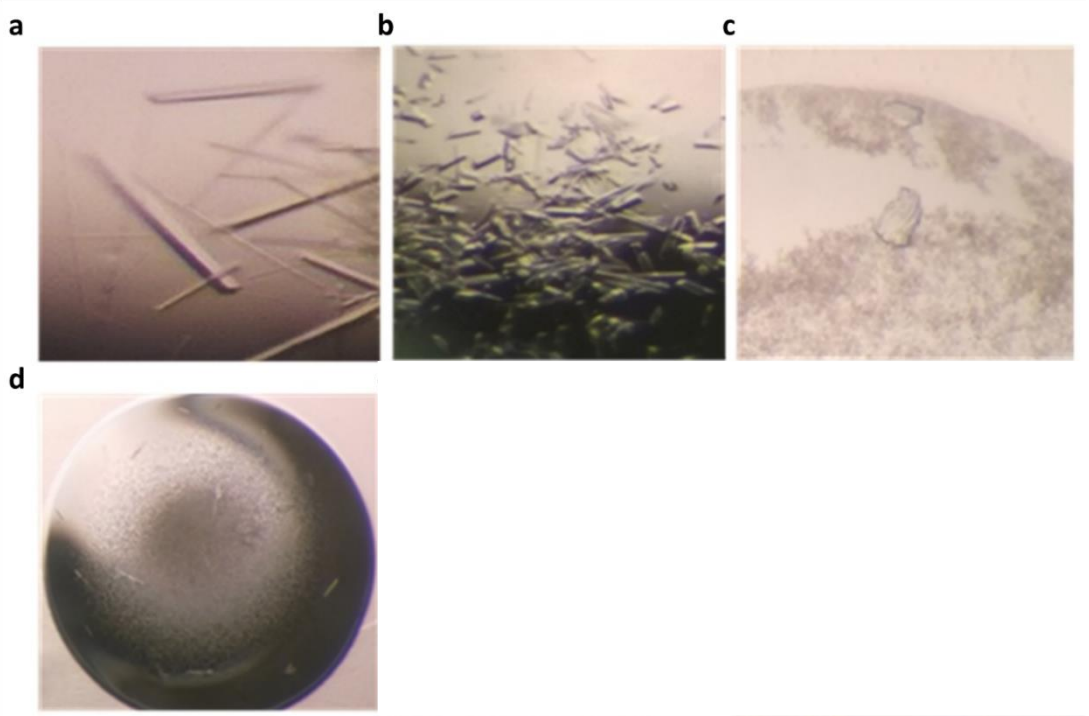


Figure 5.1 Protein crystallisation.

C-*JEV*NS1 protein was crystallised under different crystallisation conditions and produced different shape crystals, for example, **(a)** needle shape in 0.1 M MES pH 6.2, 0.8 M $(NH_4)_2SO_4$, **(b)** short rod shape in 0.1 M MES pH 5.5, 1 M Li_2SO_4, and **(c)** sheet-like in 18% (v/v) 2-propanol, 0.1 M Sodium citrate tribasic dihydrate pH 5.5, 20% (w/v) polyethylene glycol (PEG) 4000. **(d)** C-*JEV*NS1' in 1.2 M $(NH_4)_2SO_4$, 5% (v/v) Propanol.

5.1.5 Structure analysis

Assembly analysis was performed by program PISA (Krissinel and Henrick, 2007). Conservation scores of residues on protein structures were given by Consurf (Landau et al., 2005) using 21 homologous sequences (**Appendix 4**). Electrostatic surface maps were generated by using PDB2PQR (Dolinsky et al., 2007) and APBS (Baker et al., 2001) without pKa prediction.

5.1.6 Small angle X-ray scattering (SAXS)

The purified C-*JEV*NS1 1st peak (2.4 mg/ml), C-*JEV*NS1 2st peak (3.4 mg/ml), C-*JEV*NS1' 1st peak (5 mg/ml), and C-*JEV*NS1' 2nd peak (5 mg/ml) were analysed with SEC-SAXS (SEC-3 4.6x300 mm HPLC column) on beamline SWING at Soleil synchrotron, France. The low resolution model surface representation was created and docked with C-*JEV*NS1 or C-*JEV*NS1' dimer.

5.1.7 Molecular dynamics (MD) simulations

Molecular dynamics (MD) simulations were performed by Dr. Gareth Wright using GROMACS 4.6.5 (Berendsen et al., 1995) and GROMOS96 54A7 (Scott et al., 1999) force field in a cubic box solvated with single point charge-E water molecules on C-*JEV*NS1 dimers. A neutral charge was introduced at 150 mM NaCl. The distance between C-*JEV*NS1 dimers and the box edge was set to 10 Å. Long range interactions were defined using the particle mesh Ewald algorithm and other non-bonded interactions were restricted to 10 Å. An energy minimization was performed using the steepest descent algorithm followed by a 100 picosecond (ps) NVT ensemble (the number of particles N, the volume V, and the temperature T of the system are kept constant) at 310 K and a 200 ps NPT ensemble (the number of particles N, the pressure V, and the temperature T of the system are kept constant) at 310 K and 1 bar. Production MD was performed at 310 K and 1 bar for 40 nanoseconds (ns). Cα displacement was calculated with the GROMACS RMSF function. Torsion angle MD was performed with crystallography and NMR System (CNS) at 100,000 K for 37.5 picoseconds (ps) with sampling every 7.5 femtoseconds (fs) in eight separate simulations. The best structure was found with FoXS (Schneidman-Duhovny et al.,

2013) using experimental data over data range 0.017<q<0.25 $Å^{-1}$ and was the refined with another eight separate 7.5 ps simulations and energy minimization in GROMACS using the procedure described above. Models were again compared with FoXS. Freeing loop 214-243 gave a fit with experimental data of χ 1.66, however, expanding the flexible region to 218-272 allowed improvement of the fit to χ 1.48.

Table 5.1 X-ray data collection and refinement statistics

	C-*JEV*NS1 (1st)	C-*JEV*NS1 (2nd)	C-*JEV*NS1'
Data collection			
Space group	$I2_12_12_1$	$I2_12_12_1$	$I2_12_12_1$
Cell dimensions			
a, *b*, *c* (Å)	50.98, 77.48, 164.28	49.42, 78.24, 163.18	50.32, 77.94, 163.49
α, β, γ (°)	90, 90, 90	90, 90, 90	90, 90, 90
Resolution (Å)	50-2.15	47.3-2.10 (2.21-2.10)	81.75-2.6 (2.72-2.6)
R_{merge}	-	0.103 (0.907)	0.2 (1.413)
R_{meas}	0.108(1.048)	0.113(0.986)	0.246(1.758)
R_{pim}	-	0.045 (0.383)	0.141 (1.030)
I / σI	14.27 (1.76)	11.5 (2.3)	7.1 (2.1)
CC half	0.999 (0.883)	0.998 (0.923)	0.981 (0.226)
Completeness (%)	99.3 (96)	99.8 (99.6)	99.9 (99.9)
Redundancy	-	6.3 (6.5)	5.3 (5.2)
Refinement			
Resolution (Å)	44.43-2.13	47.3-2.10	81.75-2.6
No. reflections	17335	17944	9719
R_{work} / R_{free}	0.24/0.27	0.19/0.23	0.17/0.23
No. atoms	1472	1574	1573
Protein	1411	1398	1418
Sulphate ion	41	60	60
Ligand	-	24 (MES)	4 (POL)
Water	20	92	91
B-factors(Å)			
Protein	38.15	41.52	38.64
Sulphate ions	38.36	90.12	85.15
Ligands	-	86.74 (MES)	61.91 (POL)
Water	73.17	50.19	48.67
R.m.s. deviations			
Bond lengths (Å)	0.007	0.016	0.016
Bond angles (°)	1.144	1.785	1.741

*Values in parentheses are for highest-resolution shell.

5.2 Results

5.2.1 Protein expression, refolding, and purification

The theoretical molecular weight of C-*JEV*NS1 and C-*JEV*NS1' are 20.54 kDa and 25.98 kDa, but the SDS-PAGE appearance sizes are 23 kDa (**Figure 5.2a-c**) and 30 kDa (**Figure 5.2d**), respectively. Conformational difference of the target proteins from the protein standards may cause the difference in size when analysed on polyacrylamide gel. Samples from each purification step were analysed by SDS-PAGE (**Figure 5.2a, b**). The protein pellets were successfully refolded and purified indicated by detection of protein in the soluble fraction. C-*JEV*NS1 was loaded onto a Superpose 6 10x300 mm column and the elution showed a single peak at a retention volume of 16.4 ml (**Figure 5.3a**). C-*JEV*NS1' also show a similar profile, but the peak was slightly separated at 14.5 and 15.4 ml (**Figure 5.3b**). The C- *JEV*NS1 and C- *JEV*NS1'peak fractions were collected and used for crystallization.

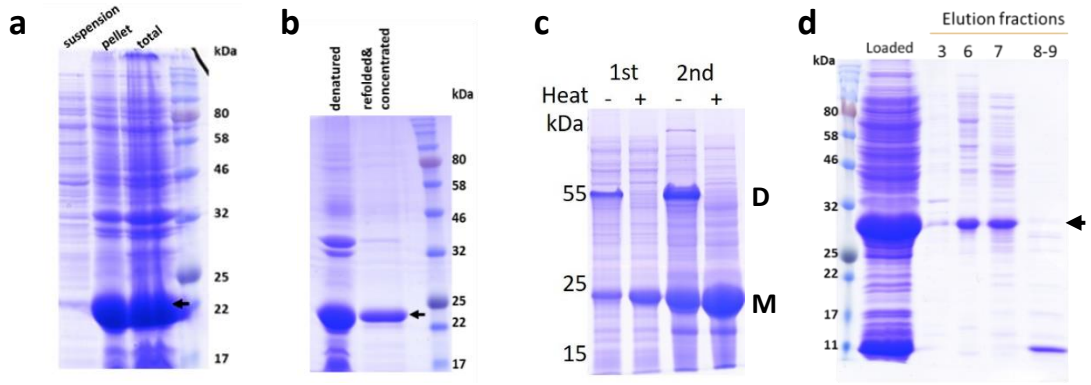


Figure 5.2. JEV NS1 and NS1' C-terminus expression and purification.
(a) C- *JEV*NS1 is mainly expressed as inclusion body. **(b)** After protein denaturation and refolding, sample suspensions were analysed and soluble protein was observed. **(c)** C- *JEV*NS1 peak 1 and peak 2 showed heat sensitivity when analysed by non-denaturing SDS-PAGE. D indicate dimer. M indicate monomer. **(d)** C- *JEV*NS1' purification by Superose 6 10x300 mm column. Peak 1 and peak 2 are in lane 6 and 7, respectively. Target protein size is indicated with black arrow.

However, when both C- *JEV*NS1 and C- *JEV*NS1' were purified with superdex 200 10x300 mm column, the elution profiles showed 2 peaks (**Figure 5.3b, c**) that were not clearly separated. To further investigate the 2 populations of protein, C-*JEV*NS1 was analysed by analytical SEC over a time-course which showed a time-dependent oligomerization between peak 1 and peak 2. Total protein which contains both peaks showed a small amount of peak 1 at day 0 (**Figure 5.4a**). The amount of peak 1 protein was gradually raised on day 3, 6, 10 and 13 samples, while the amount of peak 2 protein was gradually reduced. Protein aggregation also increase which showed at the retention volume of ~2 ml. As the protein was not well separated, C- *JEV*NS1 peak 1 fraction was contaminated with peak 2 protein. Both peaks were observed on day 0 (**Figure 5.4b**), but peak 2 reduced on day 3 and could not detect on day 6. However, it was detected again on day 10 and 13. C- *JEV*NS1 peak 2 fraction alone did not shift to form peak 1 within 13 days (**Figure 5.4c**).

Heat sensitivity is a characteristic of flavivirus NS1. The dimeric state will break down into monomer when heated, but is tolerant of reducing agents (Winkler et al., 1988) and detergent (Muller and Young, 2013). Both C- *JEV*NS1 peaks demonstrated heat sensitive property and showed the same oligomeric size at ~55 kDa (**Figure 5.2c**).

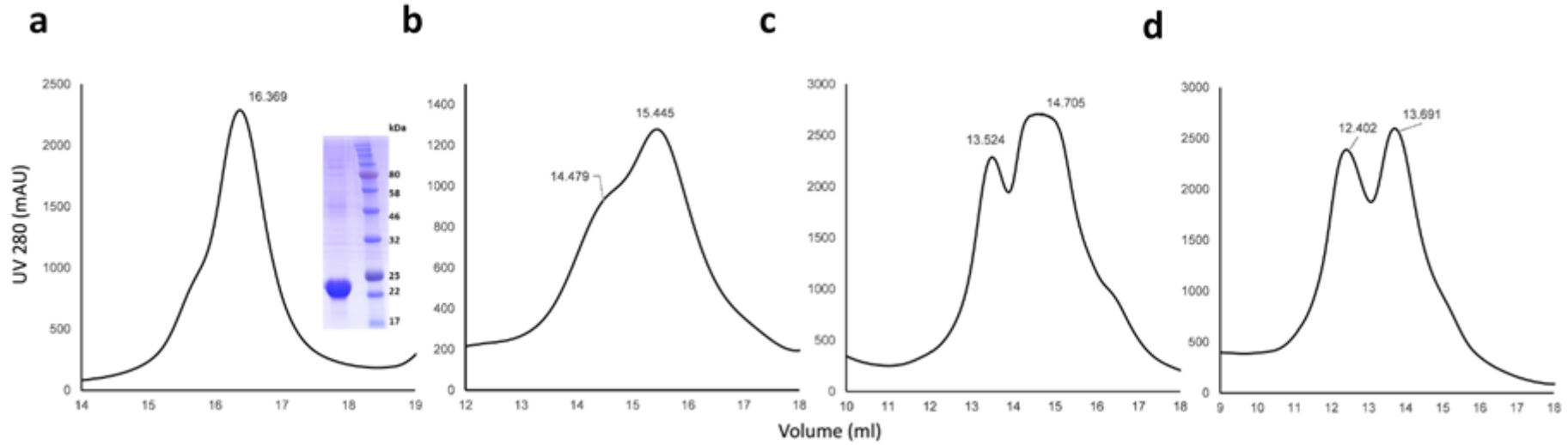


Figure 5.3 Size exclusion chromatography purification.

C-*JEV*NS1 **(a)** and C-*JEV*NS1' **(b)** elution profile purified from Superose 6 10x300 mm column. C-*JEV*NS1 protein purity is shown (right inset). **(c)** C-*JEV*NS1 and **(d)** C-*JEV*NS1' elution profile purified from Superdex 200 10x300 mm column.

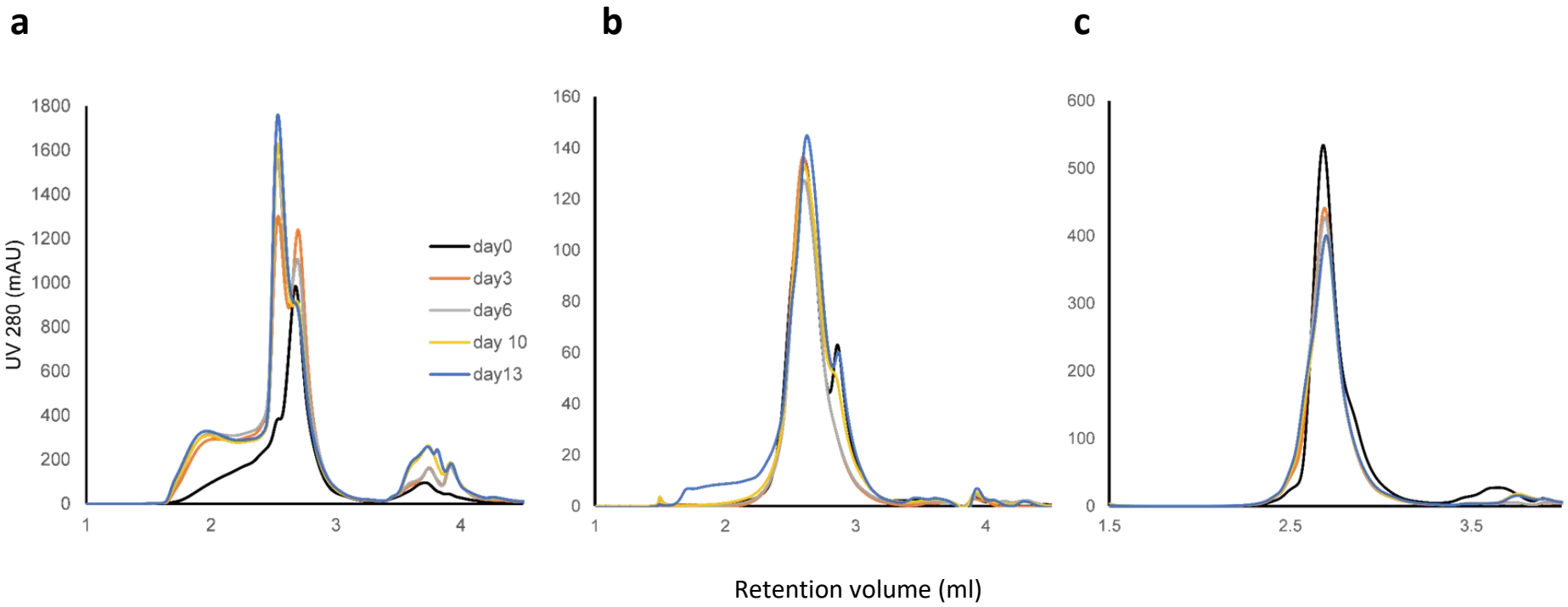


Figure 5.4 Protein dynamics analysis.

(a) C-*JEV*NS1 total protein. **(b)** C-*JEV*NS1 peak 1. **(c)** C-*JEV*NS1 peak 2. Day 0, 3, 6, 10, and 13 sample are shown in black, orange, grey, yellow, and blue, respectively.

5.2.2 C-*JEV*NS1 and C-*JEV*NS1' structure

The crystal structure of the C-terminal part of *JEV*NS1 at 2.1 Å resolution is similar to all previously described NS1 structures. Measured from its widest point, monomeric C-*JEV*NS1 is 53 x 23 Å (**Figure 5.5**). Electron density is clearly seen from residues 177-352, while the first 5 residues at the N-terminus are absent. One side of the monomer consists of 10 β-strands and another side are 4 helices and unstructured loops. β-strands are connected by β-turns and short loops except for the β4 and β5 which are separated by a long unstructured loop (residues 218-273) (Figure 5.5a, b). The protein is held together by four conserved disulphide bonds (C179-C229, C280-C329, C291-C312, and C313-C316) and hydrogen bonds between β-strands and loops. C-*JEV*NS1 forms 20 β-strands head-to-head in the dimer similar to ZIKV, WNV, and DNEV NS1 C-terminus with the dimer length of 96.5 Å at its widest point (Figure 5.5c). Twenty one residues from each monomer form the dimer interface with an average distance of 2.9 Å (**Table 5.2**). Among the interface residues, 8 of these are conserved compared to other flavivirus NS1 (**Table 5.3**, score 7-9). Twelve of hydrogen bonds are found at the dimer interface and there are 6 common residues which habitually have the same bond arrangement: Thr (JEV, ZIKV, WNV) / Ala (DENV) **186 –** Val (JEV, ZIKV, WNV) / Ile (DENV) **188**, Thr (JEV, WNV) / Ser (ZIKV, DENV) **228 -** His**254**, and Thr**230-**Trp**232** (**Table 5.4** and **Figure 5.6**) compared to ZIKV, WNV, and DENV NS1 C-terminus.

The C-*JEV*NS1' structure was determined at 2.6 Å resolution. The similar protein fold and identical disulphide bond orientations were revealed. The interface residues and hydrogen bond formation residues were identical (**Table 5.5-5.6**). C-*JEV*NS1' dimer had the same orientation as NS1. However, the electron density for the C-terminus

created by the -1 frame-shift was not observed. Only 2 extra amino acids longer than the C-terminal domain of JEV NS1 were present (0.337 Å Cα root-mean-square deviation (RMSD)) (**Figure 5.5d**) indicating C-terminal is disordered.

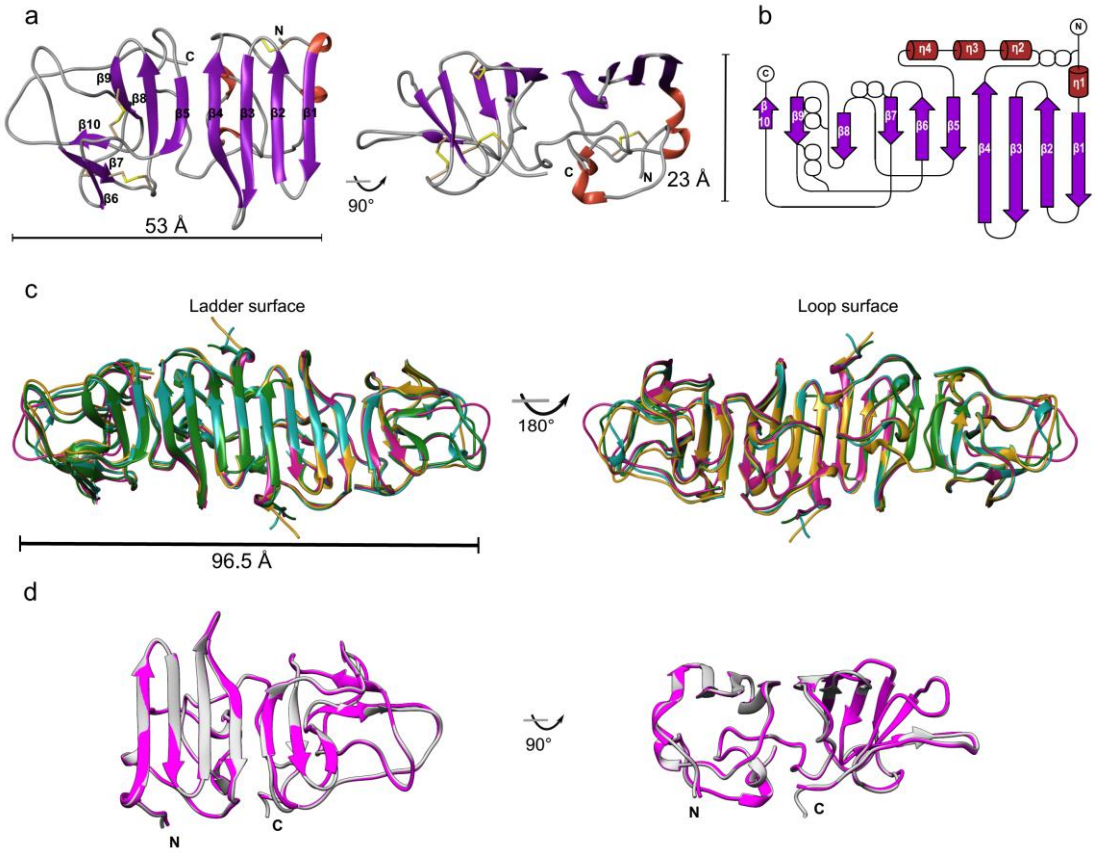


Figure 5.5 Structure of C-JEVNS1.

(a) Ribbon model of C-*JEV*NS1 monomer. One side is built of 10 β-strands and the opposite is the non-structured loops. Disulphide bonds are shown in yellow. **(b)** Topology diagram of C-*JEV*NS1. Four disulphide bonds are indicated as white spheres. β represent the β-sheet and η represent 310 helix. **(c)** Superimposed ribbon diagram of C-NS1 of JEV (magenta), ZIKV (PDB: 5IY3, blue), WNV (PDB: 4OIE, green), and DENV1 (PDB: 4OIG, gold). (**d**) Superimposition of C-*JEV*NS1 (magenta) and C-*JEV*NS1' (grey).

Table 5.2 Hydrogen bonds between C-*JEV*NS1 dimer interfacing residues and the distance

Number	Structure 1	Distance (Å)[1]	Structure 2
1	Gly190 [N]	2.93	Ile184 [O]
2	Val188 [N]	2.86	Thr186 [O]
3	Thr186 [N]	2.89	Val188 [O]
4	His229 [NE2][2]	2.83	Gly190 [O]
5	His254 [NE2]	2.94	Thr228 [O]
6	Trp232 [N]	2.96	Thr230 [O]
7	Ile184 [O]	2.93	Gly190 [N]
8	Thr186 [O]	2.86	Val188 [N]
9	Val188 [O]	2.89	Thr186 [N]
10	Gly190 [O]	2.83	His229 [NE2]
11	Thr228 [O]	2.94	His254 [NE2]
12	Thr230 [O]	2.96	Trp232 [N]

[1] Assembly analysis in the program PISA.

[2]See histidine atoms nomenclature in Appendix 5.

Table 5.3 C-*JEV*NS1 dimer interfacing residues reported with accessible (ASA) and buried surface area (BSA), solvation energy effect (ΔG) and conservation score.

Number	Residue	ASA (Å^2)[1]	BSA (Å^2)[2]	ΔG (kcal/mol)[3]	Conservation[4]
1	Gly181	23.56	6.81	0.03	3
2	Ala182	91.29	45.68	0.30	1
3	Ile184	22.47	16.36	-0.18	5
4	Gly185	40.33	16.05	0.26	7
5	Thr186	37.46	36.73	-0.22	7
6	Ala187	63.58	21.49	0.34	9
7	Val188	64.98	63.40	0.34	6
8	Lys189	181.59	9.98	0.16	8
9	Gly190	63.17	54.83	0.30	5
10	His191	110.70	33.66	0.74	1
11	Trp210	60.01	29.42	0.08	5
12	Glu227	104.51	54.79	0.51	5
13	Thr228	120.12	94.74	0.66	6
14	His229	54.08	52.04	0.90	9
15	Thr230	24.89	21.26	-0.20	9
16	Leu231	48.09	48.09	0.77	8
17	Trp232	95.68	59.18	0.38	5
18	Gly233	39.95	30.60	-0.02	4
19	Asp234	91.67	54.93	0.26	6
20	Asp235	128.72	0.58	-0.01	1
21	His254	13.56	10.75	0.73	8

[1] ASA= Accessible Surface Area

[2] BSA= Buried Surface Area

[3] ΔG= Solvation energy effect

[1,2,3] Assembly analysis in the program PISA. The values are taken from one monomer in a dimer.

[4]Amino acid conservation scores are given by Consurf. (9 = conserved and 1 = variable)

Table 5.4 Residues forming hydrogen bond at dimer interface compared among existing flavivirus NS1 structures

JEV	ZIKV			WNV			DENV	
	5k6k	5gs6	5iy3	4o6d	4o6c	4oie	4o6b	4oig
	Asp1	His1		Asp1	Asp1			
	Val2	Val2		Thr2	Thr2		Ser2	
	Cys4	Cys4		Cys4	Cys4		Cys4	
	Ser5	Ser5						
	Val6	Val6		Ile6	Val6		Ile6	
	Phe8							
	Ser9							
				Arg10	Arg10			
	Lys11							
	Glu12			Glu12	Glu12			
				Leu13				
	Arg14	Arg14		Arg14	Arg14		Lys14	
	Thr17	Thr17		Ser17	Ser17		Ser17	
	Val19	Val19		Val19	Val19		Ile19	
	Phe20	Phe20		Phe20	Phe20			
	Ile21	Val21		Ile21	Ile21		Ile21	
	Tyr22	Tyr22						
	Asn23	Asn23		Asn23	Asn23		Asp23	
	Asp24	Asp24		Asp24	Asp24			
	Arg31			Arg31	Arg31			
	Tyr32			Tyr32	Tyr32			
	Asp157	Asp157						
							Tyr158	

JEV	ZIKV			WNV			DENV	
	5k6k	5gs6	5iy3	4o6d	4o6c	4oie	4o6b	4oig
				Phe160				
	Thr165			Thr165	Thr165			
				Ser181	Ser181			
				Lys182	Lys182			Arg182
Ile184	Ile184	Ile184	Ile184					
								Ser185
Thr186	Thr186	Thr186	Thr186	Thr186	Thr186	Thr186	Ala186	Ala186
Val188	Val188	Val188	Val188	Val188	Val188	Val188	Ile188	Ile188
	Lys189			Lys189	Lys189			Lys189
Gly190	Gly190	Gly190	Gly190					Asp190
		Lys191		Asn191	Asn191			
	192Glu	Glu192						
	193Ala							
	Glu203			Glu203				
	Lys227	Lys227	Lys227					
Thr228	Ser228	Ser228	Ser228	Thr228	Thr228	Thr228	Ser228	Ser228
His229	His229	His229	His229					
Thr230	Thr230	Thr230	Thr230	Thr230	Thr230	Thr230	Thr230	Thr230
Trp232	Trp232	Trp232	Trp232	Trp232	Trp232	Trp232	Trp232	Trp232
	Thr233	Thr233	Thr233				Ser233	Ser233
	Asp234	Asp234	Asp234				Asn234	Asn234
His254	His254	His254	His254	His254	His254	His254	His254	His254

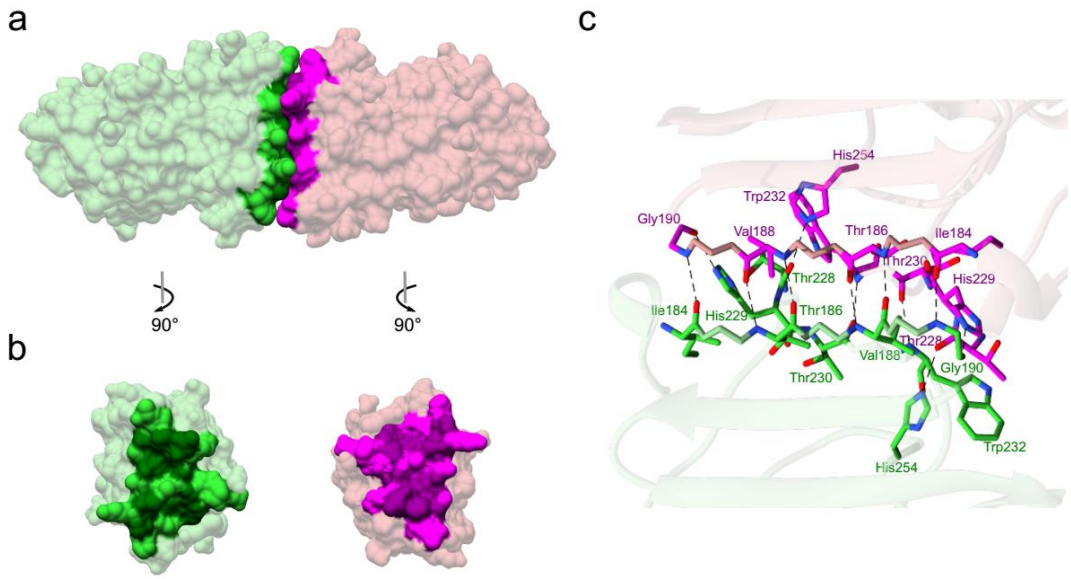


Figure 5.6 Dimer interface of C-*JEV*NS1.

(a, b) The surface of 21 residues from each monomer involved in dimer interface is coloured in lime green and the surface that form hydrogen bonds are coloured in dark green. Similarly, another monomer interfacing surface is in magenta and surface forming hydrogen bonds are in dark magenta. **(c)** Residues involved in hydrogen formation at the dimer interface are highlighted in lime green and magenta, respectively. Hydrogen bonds are indicated by dashed lines.

Table 5.5 Hydrogen bonds between C-*JEV*NS1' dimer interfacing residues and the distance

Number	Structure 1	Distance (Å)[1]	Structure 2
1	Gly190 [N]	3.10	Ile184 [O]
2	Val188 [N]	2.90	Thr186 [O]
3	Thr186 [N]	2.96	Val188 [O]
4	His229 [NE2][2]	3.11	Gly190 [O]
5	His254 [NE2]	3.00	Thr228 [O]
6	Trp232 [N]	3.07	Thr230 [O]
7	Ile184 [O]	3.10	Gly190 [N]
8	Thr186 [O]	2.90	Val188 [N]
9	Val188 [O]	2.96	Thr186 [N]
10	Gly190 [O]	3.11	His229 [NE2]
11	Thr228 [O]	3.00	His254 [NE2]
12	Thr230 [O]	3.07	Trp232 [N]

[1] Assembly analysis in the program PISA.

[2]See histidine atoms nomenclature in Appendix 5.

Table 5.6 C-*JEV*NS1' dimer interfacing residues reported with accessible (ASA) and buried surface area (BSA), solvation energy effect (ΔG) and conservation score.

Number	Residue	ASA (Å²)[1]	BSA (Å²)[2]	ΔG (kcal/mol)[3]	Conservation[4]
1	Gly181	23.27	3,57	-0.02	3
2	Ala182	91.85	39.09	0.09	1
3	Ile184	19.82	15.44	-0.16	5
4	Gly185	39.39	16.40	0.26	7
5	Thr186	37.67	36.77	-0.26	7
6	Ala187	66.81	24.00	0.38	9
7	Val188	64.00	62.27	0.38	6
8	Lys189	183.40	10.95	0.18	8
9	Gly190	68.07	56.39	0.31	5
10	His191	114.57	9.46	0.15	1
11	Trp210	59.38	28.61	0.08	5
12	Glu227	85.91	55.86	0.30	5
13	Thr228	120.24	98.04	0.75	6
14	His229	53.99	44.78	0.76	9
15	Thr230	20.31	20.28	-0.19	9
16	Leu231	47.77	47.69	0.76	8
17	Trp232	93.95	56.42	0.36	5
18	Gly233	29.10	24.02	-0.04	4
19	Asp234	96.17	56.52	0.33	6
20	Asp235	128.64	1.02	-0.01	1
21	His254	12.61	9.62	0.78	8

[1] ASA= Accessible Surface Area

[2] BSA= Buried Surface Area

[3] ΔG= Solvation energy effect

[1,2,3] Assembly analysis in the program PISA. The values are taken from one monomer in a dimer.

[4]Amino acid conservation scores are given by Consurf. (9 = conserved and 1 = variable)

5.2.3 SAXS analysis of C-*JEV*NS1 and C-*JEV*NS1'

The dimeric state of the proteins was confirmed by SAXS. The first and second peak experimental profiles of C-*JEV*NS1 and C-*JEV*NS1' were compared to its calculated monomer and dimer SAXS profiles (**Figure 5.7**, *panel 1*). For C-*JEV*NS1 first peak, C-*JEV*NS1' first and second peaks, the calculated profiles fit the experimental profiles with poor χ value (**Table 5.7**). Only the C-*JEV*NS1 second peak demonstrates a good match with χ of 4.02. R_g of all samples obtained from Guinier analysis is consistent with the value extracted from the pair distribution function (**Table 5.7**). The pair distribution function of all samples shows characteristics of a prolate ellipsoid particle. The maximum intra-particle distance (D_{max}) of C-*JEV*NS1 first, C-*JEV*NS1 second, C-*JEV*NS1' first, and C-*JEV*NS1 second peaks are 122.36, 94.13, 106.20, and 119.11 Å, respectively. Only C-*JEV*NS1 D_{max} is similar to the widest point of C-*JEV*NS1 dimer crystal structure (96.5 Å) (**Figure 5.5c and Figure 5.7b.3**). Molecular mass calculated from Porod volume is shown in **Table 5.7**, where C-*JEV*NS1 has the closest weight to the protein dimer, 45.5 kDa. The averaged *ab initio* models of all samples were generated at 30 Å resolution with good similarity agreement (**Table 5.7**) and were compared with C-*JEV*NS1 or C-*JEV*NS1' dimer crystal structures (**Figure 5.7b.4**). The structures of the C-*JEV*NS1 second peak are well matched overall but there is an extra region of mass near the dimer interface in the SAXS model (labelled M in **Figure 5.7b.4**). This feature is also seen in the SAXS model of WNV (Edeling et al., 2014) indicating that this is not an artefact. The rest of the samples showed different low resolution models. C-*JEV*NS1 first peak shows a symmetric SAXS model while C-*JEV*NS1' first and second peak models are not (**Figure 5.7**, *panel 4*). Asymmetric models are unexpected as the protein is a symmetric dimer. However, SAXS results

indicated that NS1 crystal structures may not fully represent the structure of the protein in solution.

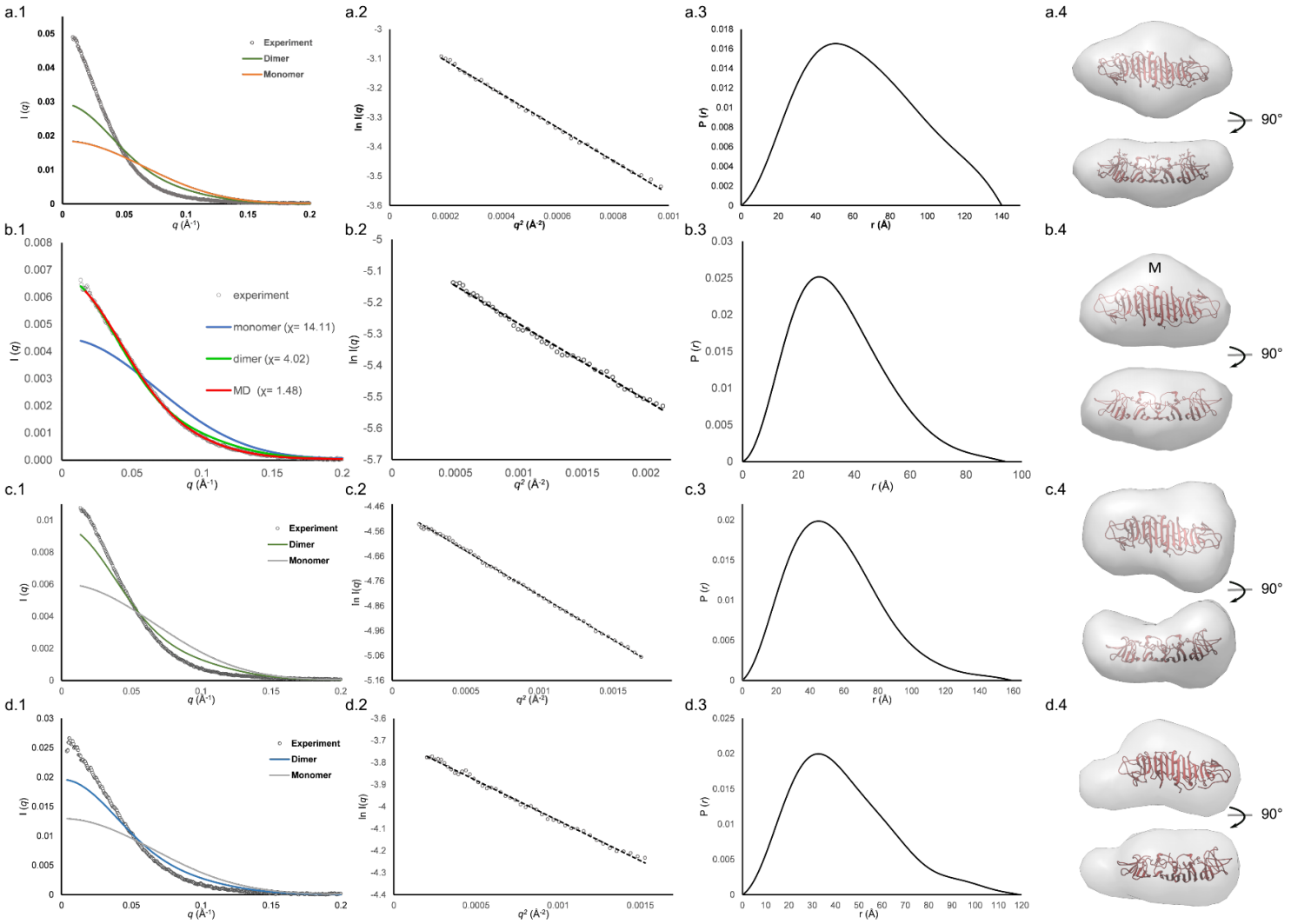

a.1
Experiment
Dimer
Monomer
I (q)
q (Å-1)
a.2
ln I(q)
q2 (Å-2)
a.3
P (r)
r (Å)
a.4
90°
b.1
experiment
monomer (χ= 14.11)
dimer (χ= 4.02)
MD (χ= 1.48)
I (q)
q (Å-1)
b.2
ln I(q)
q2 (Å-2)
b.3
P (r)
r (Å)
b.4
M
90°
c.1
Experiment
Dimer
Monomer
I (q)
q (Å-1)
c.2
ln I(q)
q2 (Å-2)
c.3
P (r)
r (Å)
c.4
90°
d.1
Experiment
Dimer
Monomer
I (q)
q (Å-1)
d.2
ln I(q)
q2 (Å-2)
d.3
P (r)
r (Å)
d.4
90°

Figure 5.7 SAXS analysis of C-*JEV*NS1 and NS1'-C dimer.

(a) SAXS analysis of C-*JEV*NS1 first peak, **(b)** C-*JEV*NS1 second peak, **(c)** C-*JEV*NS1' first peak, and **(d)** C-*JEV*NS1' second peak. Panel 1 is SAXS scattering curves. Experimental scattering curve is shown in black scattering. Scattering profile of monomer and dimer calculated with FoXS are shown. Panel 2 is Guinier plots. Panel 3 is pair distribution functions. Panel 4 Low-resolution model calculated from SAXS profiles docked with the crystal structure. An extra region of mass is labelled with M.

Table 5.7 SAXS analysis of C-*JEV*NS1 and C-*JEV*NS1' first peak and second peak.

	Chi value (χ)		Radius of gyration (R_g)		D_{max}	Molecular mass (kDa)	normal spatial discrepancy (NSD)
	monomer	dimer	Guinier analysis	Pair distribution function			
C-NS1 1st	27.48	12.16	33.5	33.58	122.36	73.6	0.575±0.02
C-NS1 2nd	14.11	4.02	27.02	27.08	94.13	45.46	0.513±0.016
C-NS1' 1st	38.62	24.53	38.06	38.06	106.20	105.73	0.698±0.026
C-NS1' 2nd	15.23	7.32	33.91	33.97	119.11	70.40	0.632±0.035

5.2.4 MD simulations of C-*JEV*NS1

Examination of B-factors in the C-*JEV*NS1 structure shows that surface areas of loop 218-272, especially sub-loop 235-237, have high conformational freedom within the crystal lattice (**Figure 5.8a)**. A 40 ns all-atom molecular dynamics simulation (MD) of the C-*JEV*NS1 dimer at 37°C confirmed that movement of this loop is relatively unrestrained in both monomers (**Figure 5.8**). The obvious extra region of mass observed in the C-*JEV*NS1 and WNV SAXS envelope structures could be the result of loop 218-272 movement causing expansion of the particle's volume. To more accurately model C-*JEV*NS1 behaviour in solution, a pool of possible structures with varying loop 218-272 conformations was created and compared with the SAXS data. Using this approach the model fitting to the experimental SAXS data was improved from χ of 4.02 to 1.48 (**Figure 5.7 b.1**).

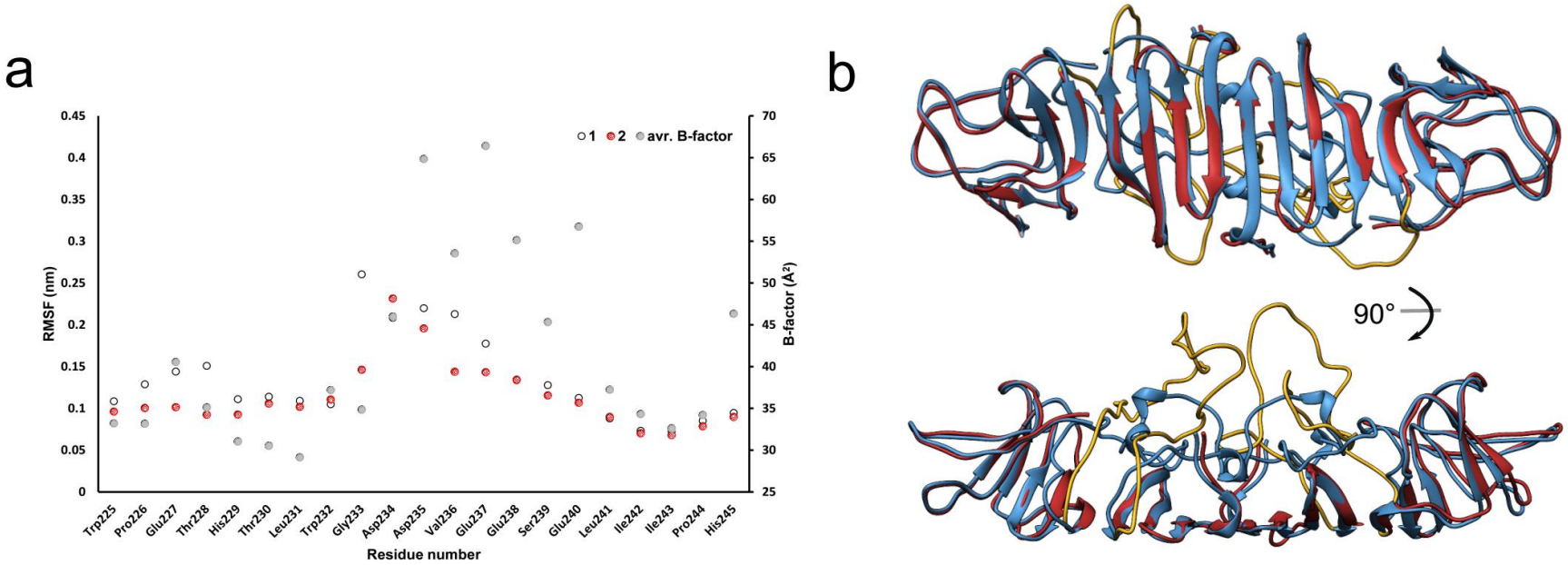


Figure 5.8 Flexible loop structure of C-*JEV*NS1

(a) The root mean square fluctuation (RMSF) plot of the molecular dynamic simulation dimer structure at the flexible loop. RMSF values of each monomer were plot in grey and red. Average β-factor of each residue was plot in solid grey**. (b)** The best molecular dynamic simulation structure (red) was superimposed with the C-*JEV*NS1 crystal structure (blue). The flexible loop 218-272 was shown in yellow.

5.2.5 C-*JEV*NS1 structure compared to other flavivirus C-NS1.

C-*JEV*NS1 has the same fold as ZIKV, WNV, DENV, and superimposition gives Cα RMSD closest to WNV NS1 (1.16 Å for ZIKV, 0.96 Å for WNV, and 1.3 Å for DENV) (**Figure 5.5c**). The N-terminus, C-terminus, and beta-turns show low positional conservation. The electrostatic surface potential map of known C-NS1: ZIKV, WNV, and DENV, were compared and show a symmetric pattern. On the β-ladder surface, all display neutral charge in the central regions flanked by negatively charged regions (**Figure 5.9**). This negatively charged region is small in DENV, larger in WNV, and is expanded diagonally from the top left to the bottom right pattern in JEV and ZIKV. Next to it toward the end are small positively charged pockets which are clearly seen only in JEV and ZIKV, and the tips of all C-NS1 have mixed charge. Loop surface is more variable than the ladder surface. DENV has a distinct positively charged central region, whereas JEV and WNV have negative charge in their central area. ZIKV is different, as the middle region displays both positive and negative charge. The adjacent area has positively charged pockets in all NS1 structures (**Figure 5.9** and **Table 5.8**). Three pockets are found in WNV and DENV, whereas ZIKV has only pocket 1 and 2, and JEV has pocket 1 and 3 (**Figure 5.9**). The residues building the positively charged pockets are conserved in pocket 1 and partially conserved in pocket 2, pocket 3, and front pocket on ladder surface (**Figure 5.9** and **Table 5.8**).

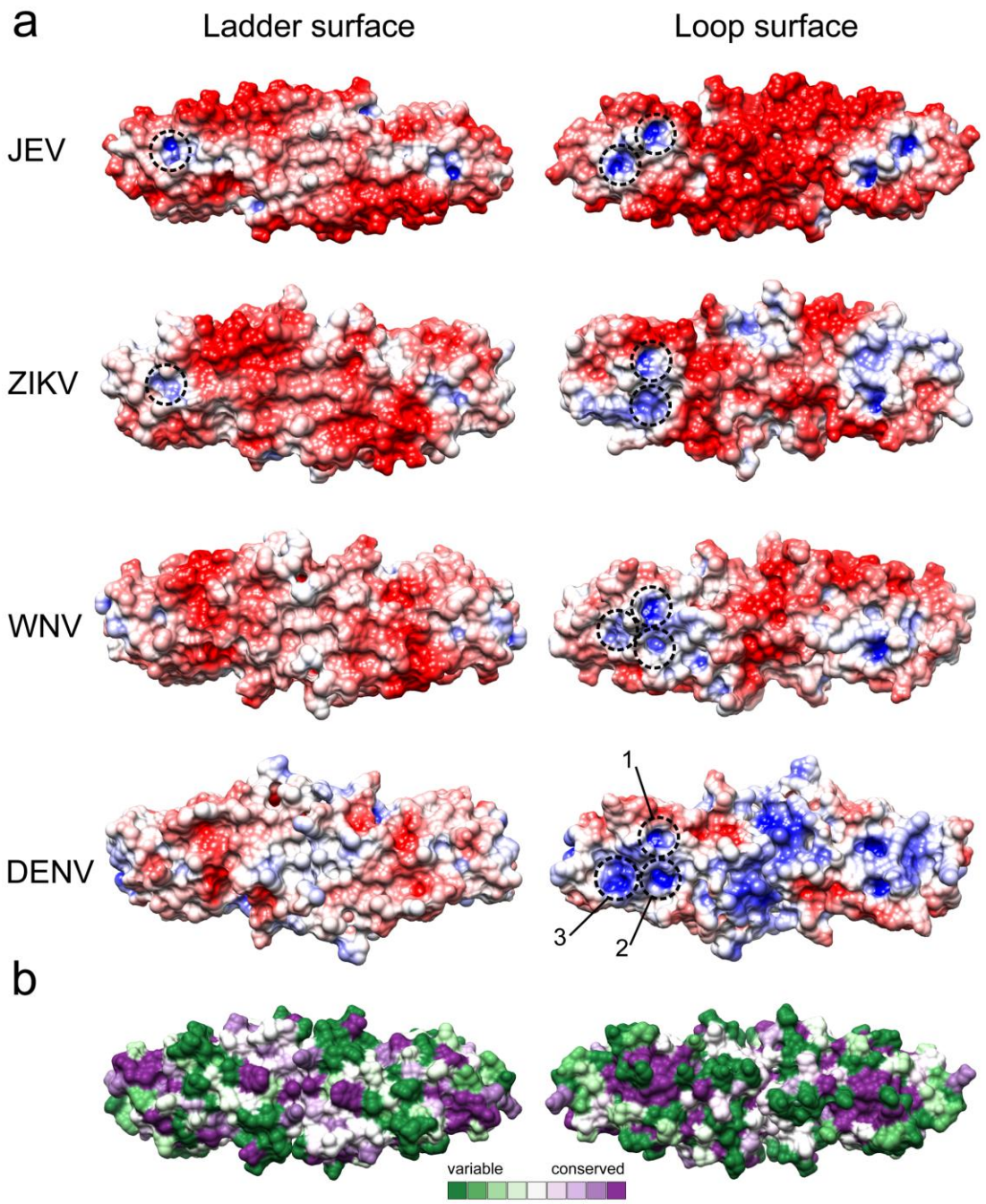


Figure 5.9 Comparison of C-*JEV*NS1 with other flavivirus C-NS1

(a) Electrostatic surface map of C-NS1 from JEV, ZIKV, WNV, and DENV. Surface is coloured by electrostatic potential from -5 kT/e (red) to 5 kT/e (blue). Positive potential pockets are depicted in dash circles. **(b)** Surface model colour-coded by conservation. The most conserved residues are represented in dark magenta and the most variable residues are represented in dark green.

Table 5.8 Residue forming positively charge pockets compared to existing C-NS1 structures

JEV	ZIKV 5iy3	WNV 4oie	DENV 4oig	Conservation[1]
Pocket 1				
Gly259	Gly259	Gly259	Gly259	9
Tyr260	Tyr260	Tyr260	Tyr260	9
Lys261	Arg261	Lys261	Phe261	1
			Ala265	1
Ser292	Gly292	Gly292	Gly292	1
Lys293	Thr293	His293	Asn293	1
Arg294	Arg294	Arg294	Arg294	9
Cys313				9
Arg314	Arg314	Arg314	Arg314	9
Ser315	Glu315	Ser315	Ser315	5
Cys316	Cys316	Cys316	Cys316	9
Glu334	Glu334	Glu334	Glu334	9
Pocket 2				
	Thr262	Thr262	Thr262	6
	Met264	Asn264	Thr264	1
	Lys265			1
	Gly295	Gly295	Gly295	9
	Pro296	Pro296	Pro296	4
		Gly332		6
	Met333	Met333	Met333	9
	Thr351	Asn351	Ser351	3
Pocket 3				
Gly295				9
Pro296				4
Ser297		Ala297	Ser297	9

JEV	ZIKV 5iy3	WNV 4oie	DENV 4oig	Conservation[1]
Val298		Thr298	Leu298	1
Arg336		Arg336	Arg336	9
Pro337		Pro337	Pro337	9
Met339				3
			Glu340	2
		Glu342	Glu342	8
Leu345		Leu345	Leu345	6
Arg347		Gln347	Lys347	3

5.3 Discussion

Flavivirus NS1 proteins have generated much interest and have been studied for more than 30 years. Since 2014, the structures of nine of the NS1 proteins have been solved (Brown et al., 2016, Akey et al., 2014, Xu et al., 2016, Song et al., 2016, Edeling et al., 2014). These proteins were expressed in bacterial or insect cell expression systems. In this study, we describe the structure of *E. coli* expressed JEV NS1 C-terminus, which when compared to other NS1 structures, shows a high degree of structural conservation. Four conserved disulphide bonds are found in the same arrangement in crystal structures. Mass spectrometry analysis of reduced and non-reduced DENV2 NS1 assigned different pair arrangement, C291-C313 and C312-C316 (Edeling et al., 2014, Wallis et al., 2004), while the crystal structures show C291-C312 and C313-C316 pairing. Conservation of the protein fold indicates that a bacterial expression system can produce stable NS1 comparable to that produced by an insect cell expression system, which is the virus natural vector. Moreover, C-*JEV*NS1 in this study was refolded from inclusion bodies which supports the view that NS1 protein has a very stable protein fold. The protein produced may be useful in categorizing NS1 protein-protein interactions. The key success of Edeling's method over the in-house method (Chapter 4) could be the ratio of the protein to the refolding buffer which was 1: 10 in in-house method and 1:100 in Edeling's method. Moreover, differences in small details such as dilution with buffer at pH 5.2 prior the refolding step, addition of protease inhibitor, or ratio of oxidizing and reducing agents may also play a role in the success.

Two populations of JEV NS1 were observed in both C-NS1 and C-NS1', but it has never been observed before in WNV or DENV NS1 recombinant protein purification (Edeling

et al., 2014, Avirutnan et al., 2007). The NS1 first elution peak is suspected to be the higher order of C-*JEV*NS1 that the protein dimer is the building block. The oligomeric stage of the dimers might be easily broken, for example by reducing agent, because it did not show higher molecular weight than the dimeric form when analysed with non-heat treatment SDS-PAGE. The second peak protein only shift to from the first peak when the first peak exists in the environment. However, it is also possible that changing from peak 2 on its own to peak 1 takes a much longer time. The crystallization of C-*JEV*NS1 and C-*JEV*NS1' were made from total protein which contains both peak 1 and 2. However, the higher order of NS1 has never been observed in crystal structures. This may due to the dimeric form protein, which is the majority of the purified protein, is enough to form crystals. However, SAXS analysis did not explain the higher order NS1 supposition because the extra mass shown in the models was not large enough to fit another NS1 monomer or dimer. Conformation shift may describe this dynamic observation, but more evidence is required. Otherwise, it could be the sample processing that generates 2 species of NS1 protein.

All NS1 proteins are dimeric *in crystallo*, even though the recombinant protein contains only the C-terminal domain (Edeling et al., 2014, Xu et al., 2016). The molecular mass and low resolution model generated from SAXS data confirm the dimeric nature of the isolated C-terminal domain in solution. The availability of WNV, DENV, and ZIKV NS1 structures allowed us to assess both general and different characteristics of the protein which may be connected to their functional role. In contrast to previous work, which suggested that the β-roll domain is responsible for dimerization (Smith et al., 2015), this study proposes that 6 common residues which

form hydrogen bonds at the dimer interface of all NS1 structures mediate dimer formation (Smith et al., 2015). In principle, inhibition of dimer formation by interposing a ligand at this site could facilitate anti Flavivirus drug development.

Both faces of the C-NS1 dimer display electrostatic surface charge diversity (Figure 5.9). However, when considering the full length Flavivirus NS1 protein structure (Brown et al., 2016, Akey et al., 2014, Xu et al., 2016), the ladder face of the C-terminal domain is positioned underneath the β-roll domain (**Figure 1.9**). The N-terminus protects the central region of the ladder face from the environment. Besides that, the β-roll domain is contained by a hydrophobic region that is suspected to interact with the cell membrane making it harder for the ladder face to make an interaction. This conflicts with a hypothesis that the β-ladder may bind to the complement control protein domain (sushi domain) of complement proteins (Akey et al., 2015). On the other hand, the loop face is fully exposed with its diverse surface charge. Particularly, DENV has the most distinct positive central area while the rest are negatively charged. This may influence the protein-protein interactions. Positively charged pockets found on the loop face of the NS1 crystal structure could be responsible for an important anionic ligand binding because it is composed of conserved sequences and exists in all known NS1 structures especially pocket 1. The presence or absence of each pocket may relate to virus evolution and its ability to interact with specific target proteins. B-factor and MD analysis provide evidence suggesting that loop 218-272 is flexible. Although the B-factors are high in this region, the X-ray structure does not show disorder. This may be due to bias from cryocooling which accesses a low-energy conformation. Loop 218-272 links strands β4 and β5 and is the longest C-NS1 loop. It was suspected that the dynamic 218-272 loop may

harbour distinct protein-protein interaction functions. This phenomenon has been independently found in WNV (Edeling et al., 2014) thus all NS1 are likely to share this characteristic. Taken together, it is very probable that the NS1 protein orientates with the N-terminus facing the membrane and the loop facing outward (Brown et al., 2016, Akey et al., 2014, Edeling et al., 2014) making an interacting interface.

5.4 Conclusion

This study presents a 2.1 Å resolution crystal structure of C-*JEV*NS1 and 2.6 Å resolution of C-*JEV*NS1'. C-*JEV*NS1 share a conserved fold of flavivirus NS1-C domains. The surface charge distribution of C-*JEV*NS1 is similar to WNV and ZIKV but is significantly different from DENV. This is likely to be important for their specificity and ability to interact with other proteins in the cell. Analysis of the JEV NS1 structure, *in silico* molecular dynamics simulations, and experimental solution small angle X-ray scattering indicate extensive loop flexibility on the exterior of the protein. Taken together with charge distribution on the exterior of the protein, the loop may govern protein-protein interaction function.

Chapter 6 Antibody, complement proteins and cell membrane interaction with C-*JEV*NS1

Even though NS1 is known to be involved in viral replication and immune invasion, the molecular detail of its function is a mystery. From the structure, surface diversity gave a clue of the unique character of flavivirus NS1. It would be interesting therefore to investigate its functions further; however, NS1 is not known to have any enzymatic function. It is a challenge to develope the experiment to test the function of the protein.

Antibodies against NS1 are known to have a protective effect in mice. In a WNV anti-NS1 study, the 22NS1 antibody was notable as it had a strong protective effect specific to WNV (Chung et al., 2006b). The epitope was identified (Edeling et al., 2014) and quite conserved compared to JEV (Trp232, Ser239, Tyr260, Lys261, Thr262, Glu289, Arg294, Arg314, and Ser315). Cross reactivity of the antibody is investigated in this study.

Unfortunately, JEV is not as widely studied as DENV, WNV, or ZIKV. Most information is inferred from other *Flaviviruses*. Immune molecules and cell membrane interaction were reported in DENV and WNV NS1. To prove that for JEV, complement proteins, liposome, and heparin interactions are examined in this study. Parts of the work reported in this chapter is now published (Poonsiri et al., 2018).

6.1 Materials and methods

6.1.1 Recombinant proteins

C-*JEV*NS1 with the histidine tag was generated from the full length template by PCR method for pulldown assay. Primer sequences are given in **Table 3.4** in chapter 3. The truncated DNA fragment was cloned into pOPIN F vector at kpnI/HindIII cloning site to get the histidine tag at N-terminus. C-*JEV*NS1 and C-*JEV*NS1' were produced using method described in chapter 4. All of the recombinant proteins were refolded and purified as mentioned in chapter 5.

6.1.2 Protein complex formation

Complex formation of the recombinant C-*JEV*NS1 and 22NS1 antibody was confirmed by Western blot. To detect C-*JEV*NS1, WNV 22NS1 monoclonal antibody was used as primary antibody and goat anti-mouse IgG was the secondary antibody. Purified C-*JEV*NS1 and 22NS1 fragment antigen-binding (Fab) (prepared by Pierce™ Fab Preparation Kit, Thermo Fisher Scientific) were mixed at 2:1, 1:1, and 1:2 molar ratio of protein to antibody overnight at 4°C and purified by Bio-SEC3 4.6x300 mm HPLC column (Agilent Technologies) or Superdex 200 10x300 mm size exclusion chromatography column.

6.1.3 Small angle X-ray scattering (SAXS)

C-*JEV*NS1-22NS1 Fab complex at concentration of 3 mg/ml in TBS buffer (20 mM Tris-HCl pH 7.4, 150 mM NaCl) were analysed with SEC-SAXS (SEC-3 4.6x300 mm HPLC column) at beamline SWING, Soleil synchrotron, France. The low resolution model

surface representation was created and docked with C-*WNV*NS1-22NS1 complex (PDB ID 4OII).

6.1.4 Pulldown assay for complement proteins interaction study

C-*JEV*NS1 with the histidine tag (bait) was produced for pull-down assay. The protein was expressed and purified by the same methods as C-*JEV*NS1. Small scale Ni-NTA column was set up by adding 50 µl of Super nickel-NTA agarose affinity resin (Generon) slurry into 1 ml pipette tip which was plugged by 20 µl tip's filter. The column was operated with standard Ni-NTA protocol, except that 1000 µl pipette was used to dispense the buffer through the column. The binding buffer was 20 mM HEPES pH 7.4, 150 mM NaCl, 20 mM Imidazole and elution buffer was 20 mM HEPES pH 7.4, 150 mM NaCl, 500 mM Imidazole. Equilibration was done by adding 400 µl of binding buffer to the column. The steps were repeated 3 times. C-*JEV*NS1 (100-150 µg) was added to the equilibrated column and incubated on a roller (both ends of the tip were wrapped with parafilm) for 30 minutes at 4°C. The bait flow-through was dispensed and the column was washed with 400 µl binding buffer 3 times and 400 µl of normal human serum (pray) (Sigma, H4522), which was 10 times diluted with binding buffer, was added. The column was incubated on a roller for 1 hour at 4°C and washed again with 400 µl binding buffer 3 times before eluted twice with 100 µl of elution buffer. Samples from each step: load, flow-through, wash, and elute, were analysed by SDS-PAGE and Coomassie staining. Bait protein alone and pray protein alone were applied to the column as the positive and negative control, respectively.

6.1.5 Lipid binding assay

Liposomes were prepared from cholesterol (CHOL) (Sigma, C8667) and 1,2-Dipalmitoyl-sn-glycero-3-phosphocholine (PC) (Sigma, P4329) at 1:9 CHOL to PC (Smith et al., 2015). CHOL and PC powder were dissolved in chloroform. To achieve total 400 nmol, 40 nmol of CHOL and 360 nmol of PC were mixed together in 2 ml tube and the lipid mixture was dried under nitrogen gas stream. The lipid films were kept in -20°C until use. To hydrate the lipid sheets, 50 µl of buffer (50 mM Bis-Tris pH 5.5, 50 mM $(NH_4)_2S0_4$, 10 % glycerol or 150 mM KCl, 25 mM Tris-HCl pH 7.5, 1 mM DTT, 0.5 mM EDTA.) was added and incubated at room temperature on a shaker for 30 min. Then the lipid was sonicated with an exponential probe at amplitude 4 for 30 seconds with 30 seconds interval for 5 times in a warmed water bath to avoid overheat. The liposomes were kept at 4°C for 1-2 weeks or -20°C until use. Liposome binding reaction (50 µl) was set up at 400 nmol, 125 nmol and 25 nmol of total lipid and mixed with 5000 ng of C-*JEV*NS1 (5 µl of 1 mg/ml protein). The reactions were incubated at 37°C for 2 hours. After that, the reactions were centrifuged at 16000xg for 30 minutes at 22°C and the supernatant was harvested to a new tube. The lipid pellet was resuspended in 200 µl buffer and transferred to a new tube. Liposomes were pelleted again and the supernatant was discarded. Liposome pellet was resuspended in 30 µl of 1x SDS-PAGE sample buffer. Bovine cytochrome bc1 complex, membrane proteins provided by Kangsa Amporndanai, was used as positive control in 25mM phosphate buffer pH 7.5, 100mM NaCl, 3mM NaN3, 0.015% DDM buffer. The supernatant and pellet fractions were analysed by SDS-PAGE and Coomassie staining.

6.1.6 Heparin binding

C-*JEV*NS1 was used in this assay. Small scale column was set up and operated as described in pull down assay by using heparin agarose beads (Affi-Gel heparin gel, BIO-RAD). The binding buffer was 20 mM HEPES pH 7.4, 150 mM NaCl and the elution buffer was 20 mM HEPES pH 7.4 supplemented with different concentrations of salt (1.5 M and 2 M NaCl). The column was equilibrated with 400 µl of binding buffer for 3 times. C-*JEV*NS1 5000 ng was applied to the column and incubated on a roller for 30 minutes at 4°C. The column was washed with 400 µl of binding buffer 3 times before eluted twice with 100 µl of 1.5 M and 2 M NaCl elution buffer, respectively. Samples from each step: load, flow-through, wash, and elute, were analysed by SDS-PAGE and Coomassie staining. Superoxide dismutase 3 (SOD3) provided by Varunya Chantadul was used as positive control.

6.1.7 Differential scanning fluorimetry (DSF).

Polymers of heparan sulphate (average molecular weight 30,000), chondroitin sulphate (62% chondroitin 4-sulphate and 33% chondroitin 6-sulphate, average molecular weight 45,400), and dermatan sulphate (average molecular weight 41,000) from Iduron at final concentration of 100, 50, 25, 10, 5, 1, 0.5 µM were mixed with C-*JEV*NS1 and Sypro Orange 5000X (Invitrogen) at final concentration of 10 µM and 10X, respectively. The reaction volume was 10 µM. The experiments were set in 96 well-plates and perform using StepOnePlus™ Real-Time PCR Systems (software version 2.3) (Applied Biosystems). The experiment was set to follow the Protein Thermal Shift™ Studies user guide. The heating cycle was a 2 minutes prewarming step at 25°C

and followed by a ramp to 95°C with a ramp rate of 1°C. The experiment type was set as Melt Curve, the reporter was ROX, and the quencher was None.

6.2 Results

6.2.1 NS1 protein and 22NS1 Fab interaction

The interaction of WNV antibody with C-*JEV*NS1 and C-*JEV*NS1' were confirmed by Western blot analysis (**Figure 6.1**, lower left inset) and size exclusion chromatography (**Figure 6.1**). C-JEVNS1 and 22NS1 Fab alone eluted at a retention volume of 3.16 and 3.32 ml, respectively. C-JEVNS1 incubated with 22NS1 Fab eluted faster at a retention volume of 2.74 ml corresponding to complex formation with a small amount of free 22NS1 Fab fragments left. The eluted fraction was analysed by SDS-PAGE and 2 peaks representing C-JEVNS1 and 22NS1 (~25 kDa) were identified. The incubation also generated a small peak at a retention volume of 2.44 ml which is believed to be a higher order oligomeric stage of C-JEVNS1 (front peak of C-JEVNS1 at a retention volume of ~3 ml) complexed with 22NS1. This confirmed that NS1 and 22NS1 monoclonal antibody (mAb) interact in solution (**Figure 6.1**, lower right inset). The protein interaction was also analysed by SAXS. As the 22NS1 antibody, which binds to C-WNVNS1, was suspected to bind by the same epitope to JEVNS1 protein, it should have a similar structure for the complex. The JEV complex experimental profile was compared to calculated scattering profile of C-WNVNS1-22NS1 complex (PDB ID 4OII) (**Figure 6.2a, c**). The complex (4OII) gave a poor fit to the experimental data with χ of 6.82. Guinier analysis gave the radius of gyration of 52.89 ± 0.34 Å which coincides with 52.50 Å extracted from the pair distribution function. The pair distribution function of the complex has multiple peaks which signify the multi-domain geometric shape with Dmax of 154.9 Å (**Figure 6.2b**), similar to the C-WNVNS1-22NS1 complex. The calculated molecular mass was 149.96 kDa. An

averaged ab initio model was generated at 30 Å resolution. The Fab part of WNV complex (4OII) structure did not fit into the SAXS envelope. The Fab part in SAXS structure moves away from the position in 4OII model indicating the flexibility of Fab epitope in solution, while the C-WNVNS1 dimer fits well (**Figure 6.2c-e**) into the envelope shape. I generated a pseudo-atomic model of the C-JEVNS1 antibody complex by replacing the C-WNVNS1 with C-JEVNS1 and optimizing the position of the Fab molecules. This model had better fit to the SAXS data which improved the fit parameter from χ of 6.82 to 3.09 (**Figure 6.2**a, c-e).

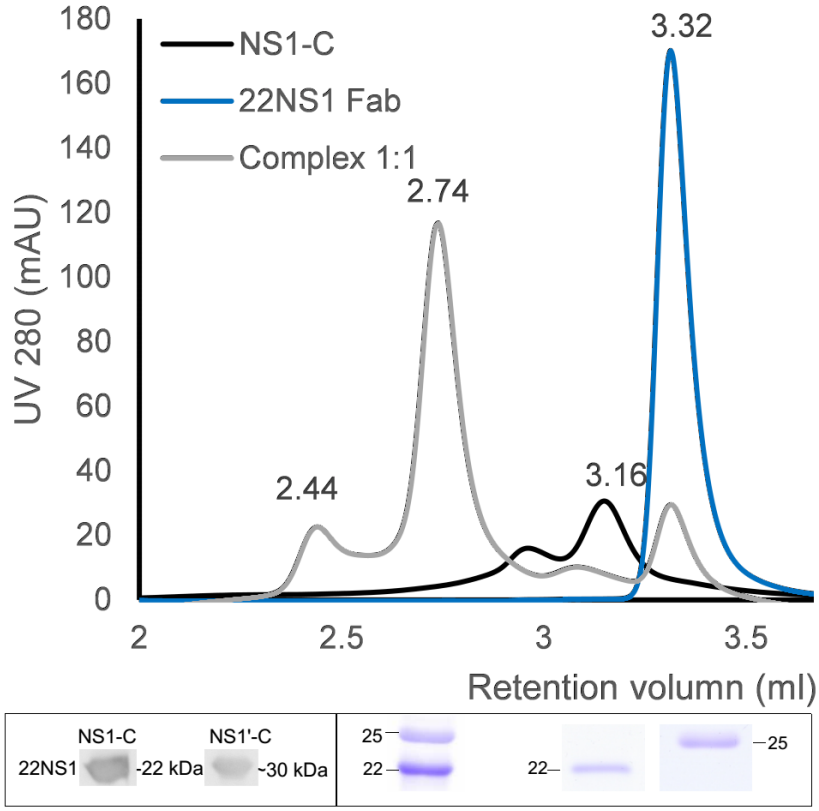


Figure 6.1 C-*JEV*NS1 in complex with 22NS1 Fab.

C-*JEV*NS1 was detected by 22NS1 mAb (lower left inset). C-*JEV*NS1 was incubated with 22NS1 Fab at 1:1 molar ratio protein to Fab fragment and the complex formation was analysed on an Agilent BioSEC-3 4.6/300. The lower right panel show SDS-PAGE analysis of each elution fraction.

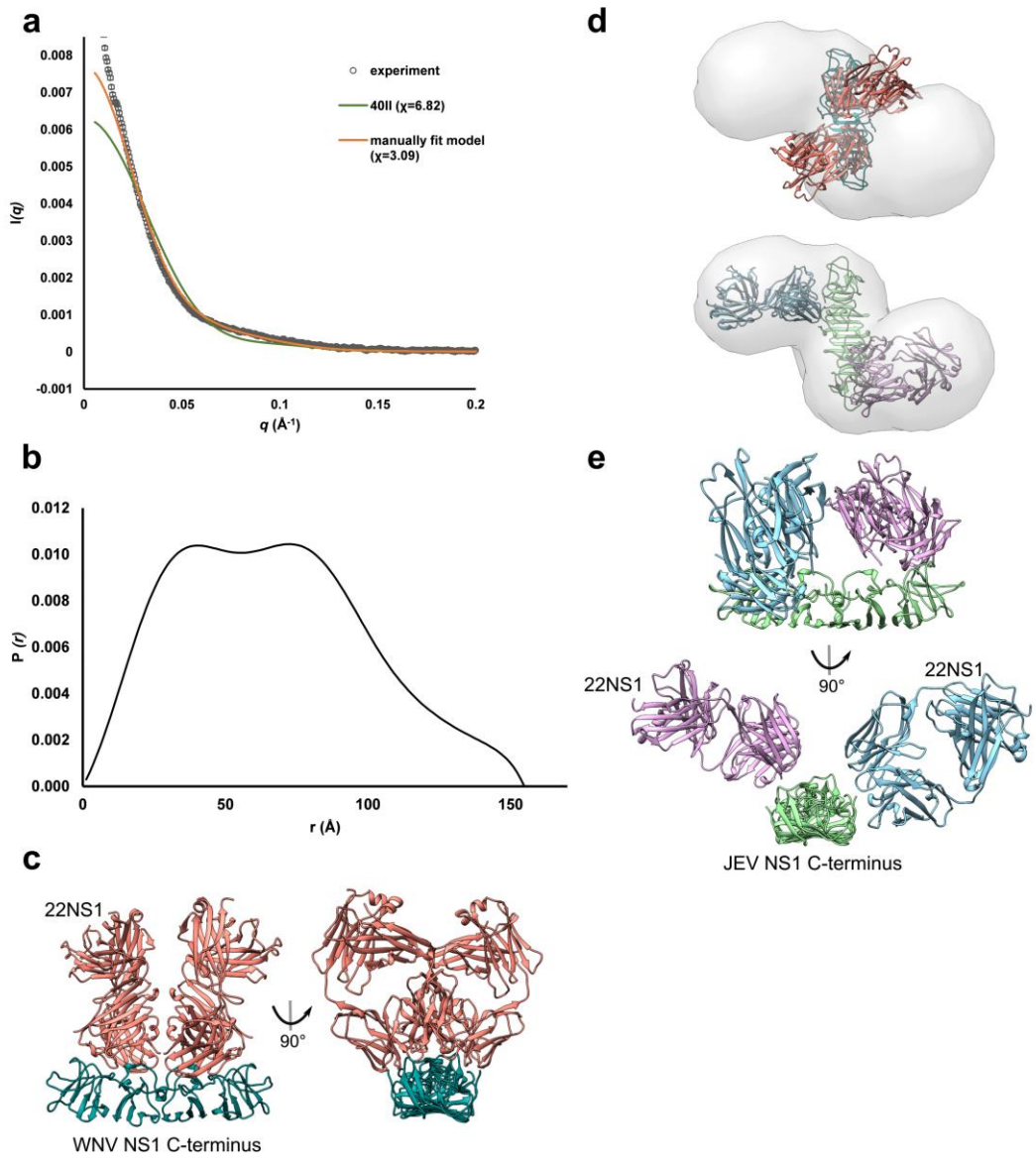


Figure 6.2 SAXS analysis of C-*JEV*NS1-22NS1 Fab complex.

(a) SAXS scattering curve. Experimental scattering curve for C-*JEV*NS1-22NS1 Fab complex is shown in black scattering. Calculated scattering profile of C-*WNV*NS1-22NS1 complex (4OII) is displayed in green and C-*JEV*NS1-22NS1 Fab complex manually fit model is shown in orange. **(b)** Pair distribution function shows 2 peaks signify the two-domain geometric shape. **(c)** C-*WNV*NS1-22NS1 complex (4OII). C-*WNV*NS1 is coloured in blue. 22NS1 Fabs are coloured in pink. **(d)** C-*WNV*NS1-22NS1 complex (4OII) fit the C-*JEV*NS1-22NS1 Fab complex ab initio model (upper). A pseudo-atomic model C-*JEV*NS1-22NS1 Fab complex are manually fit the ab initio model (lower). **(e)** C-*JEV*NS1-22NS1 Fab complex pseudo-atomic model.

6.2.2 C-*JEV*NS1 and complement protein interaction

The C-*JEV*NS1 with the histidine tag at N-terminus (bait) incubated with 10 times diluted normal human serum (pray) was isolated by Ni-NTA purification. Analysed by SDS-PAGE, no pray protein was eluted together with C-*JEV*NS1 (**Figure 6.3**) indicating the absence of interaction with complement protein.

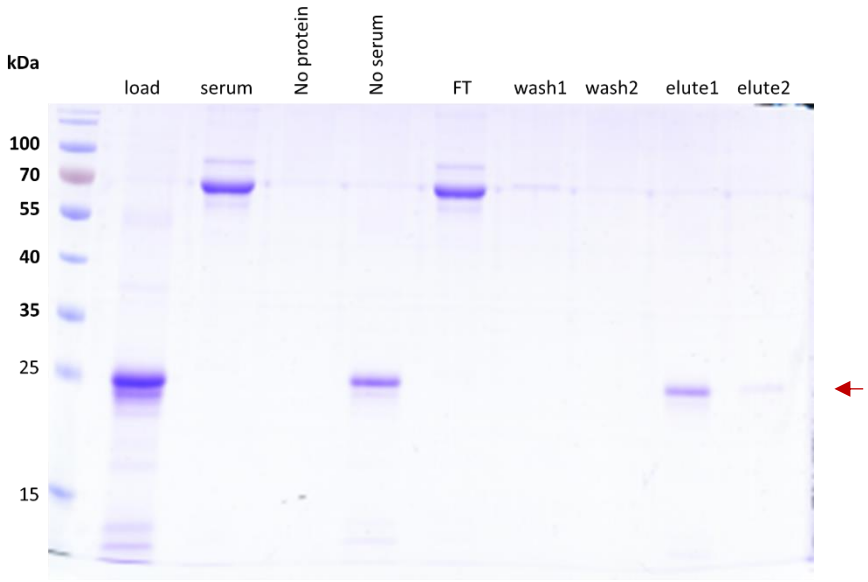


Figure 6.3 C-*JEV*NS1 pull down assay.

The protein complex was isolated by Ni-NTA. First lane show protein loaded onto Ni-NTA column. Second is the 10 times diluted human serum. Lane 3 and 4 are the control. The column was loaded with human serum alone and the elution fraction is shown in lane 3. Lane 4 is the elution of the column loaded only with C-*JEV*NS1. C-*JEV*NS1 was loaded onto the column and mixed with human serum. Flow-through fraction (FT) is shown in lane 5. Wash fractions are in lane 6 and 7. Elution fractions are in lane 8 and 9.

6.2.3 C-*JEV*NS1 and cell membrane interaction via GAGs

Sulphate molecules were found on the surface of C-*JEV*NS1 X-ray structure similar to ZIKV (PBD ID 5K6K), WNV (4O6C), and DENV (4OIG) (**Figure 6.4** and **Table 6.1**). Moreover, they are distributed near the positively charged pockets. Hence, it is possible that this positively charged area might be the binding site of negatively charged ligands. Specifically, natural sulphate molecules like GAGs that are involved in membrane attachment may interact here (Avirutnan et al., 2007).

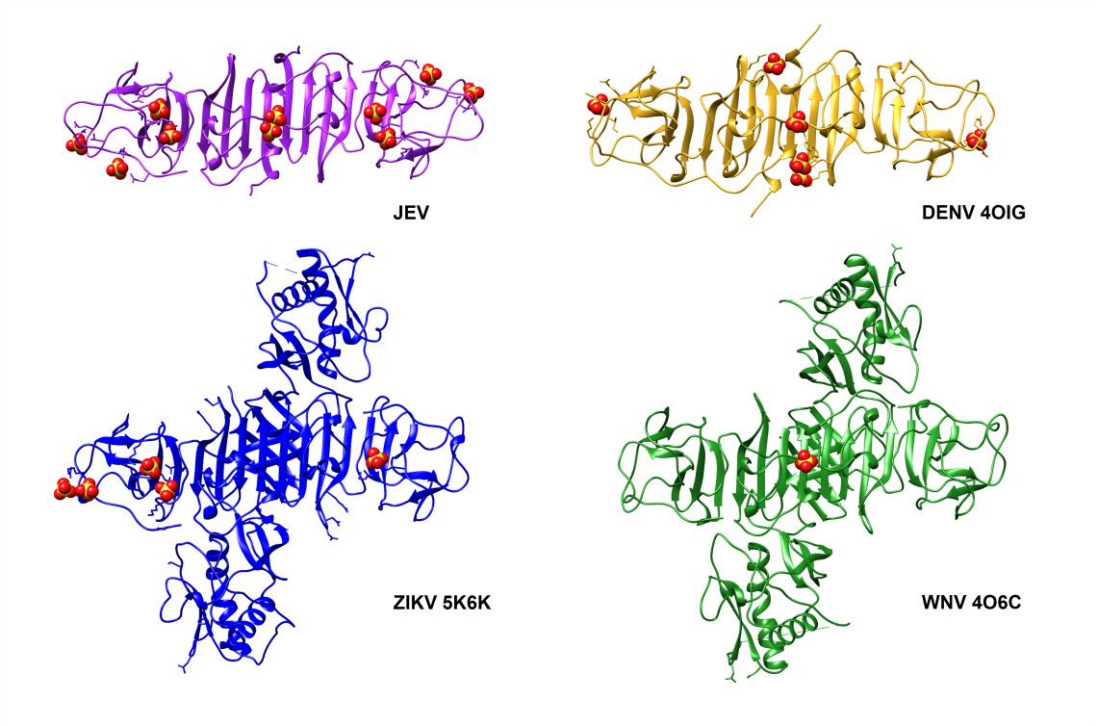


Figure 6.4 Sulphate molecules bound to the loop surface of C-JEVNS1

Sulphate molecules were found not only for JEVNS1, but also in DENV 4OIG, ZIKV 5K6K, and WNV 4O6C. Thus, it is suspected to be an importance sulphate binding interface.

Table 6.1 Sulphate contact residues from assembly analysis in the program PISA

Area	C-JEV	ZIKV 5K6K	WNV 4O6C	C-DENV 4OIG
Tip	Arg347 Gln349 Thr302 Ser304 Lys306 Thr343 Thr344	Ser342 Glu343 Thr302 Ser304 Arg306		His309 Glu310 Lys339
Positively charge pockets	Arg294 Arg314	Arg294 Arg261		
Central	Asp235		Gly235	His181 Lys206 Thr210 Ser228 Trp232 Asn234 Gly235

To prove that interaction with GAGs happens via sulphate binding sites at C-terminus, C-*JEV*NS1 binding to heparin agarose beads was analysed. The 20 kDa C-*JEV*NS1 was found only in flow-through and wash fractions (**Figure 6.5a**) indicating that C-*JEV*NS1 could not interact with heparin. The interaction of heparan sulphate, chondroitin sulphate, and dermatan sulphate polymers with C-*JEV*NS1 was further investigated by protein thermal shift assay. No apparent of C-*JEV*NS1 stabilizing effect was observed for any of GAG polymers tested even at high concentration (100 µM) (**Table 6.2** and **Appendix 6**). No change of the unfolding temperature (Δ melting temperature) more than approximately 1°C was observed. No GAG binding was consistent with the pull-down experiments. These suggest that the interaction of NS1 with GAGs does not happen at the C-terminal sulphate binding sites.

Interaction of C-*JEV*NS1 with the cell membrane was tested by liposome binding assay. Liposomes could not bind to C-*JEV*NS1 at either pH 7.5 or 5.5 (**Figure 6.5b**). While full length NS1 does bind liposomes (Akey et al., 2014, Smith et al., 2015), this study shows that the NS1 C-terminus is not responsible for membrane interaction. In contrast to this view, the hydrophobic residues of β-roll and wing domains were previously suggested to play a role in membrane binding (Xu et al., 2016, Brown et al., 2016, Akey et al., 2014). The C-*JEV*NS1 does not responsible for cell membrane interaction via GAGs.

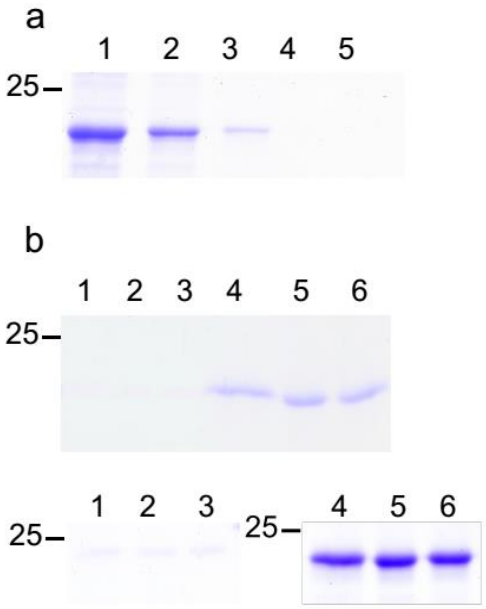


Figure 6.5 Cell membrane interaction determination.

(a) Heparin binding determination. C-*JEV*NS1 was incubated with heparin agarose beads. Total C-*JEV*NS1 loaded to the column is shown in lane 1. Lane 2 is flow-through fraction. Lane 3-4 are wash fraction. The column was eluted with buffer supplemented with 1.5 M NaCl shown in lane 5. **(b)** Liposome binding assay. The experiments were conducted at pH 7.5 (upper) and pH 5.5 (lower). Supernatant and pellet fractions separated by centrifugation were analysed by SDS-PAGE. Lane 1, 2, and 3, were pellet of 400 nmol, 100 nmol, and 25 nmol reactions, respectively. Lane 4, 5, and 6, were supernatant of 400 nmol, 100 nmol, and 25 nmol reactions, respectively.

Table 6.2 Mean values of melting temperatures calculated from experiments performed in duplicate and its ΔTm.

HS (μM)	Mean Tm	±SD	ΔTm
100	54.2	1.1	0.3
50	54.6	1.2	-0.1
25	54.8	0.9	-0.3
10	54.8	0.9	-0.3
5	54.8	0.9	-0.3
1	54.7	1.1	-0.2
0.5	54.6	1.3	-0.1
0	54.3	1.3	0.2
DS (μM)	**Mean Tm**	**±SD**	**ΔTm**
100	54.5	1.2	0.0
50	54.7	0.7	-0.2
25	55.2	1.1	-0.7
10	55.0	0.8	-0.5
5	54.7	0.4	-0.2
1	54.8	0.6	-0.3
0.5	55.2	1.4	-0.7
0	54.8	1.2	-0.3
CS (μM)	**Mean Tm**	**±SD**	**ΔTm**
100	53.2	1.3	1.3
50	53.9	0.9	0.6
25	54.2	0.8	0.3
10	54.2	1.5	0.3
5	54.1	1.9	0.4
1	54.6	1.2	-0.1
0.5	54.4	1.4	0.1
0	54.5	1.4	0.0

Note: The concentration of JEVNS1-C was kept constant at 10 μM. The average Tm of JEVNS1-C alone was 54.5°C. The experiments were performed in total 6 replicates (3 replicates for each batch of protein purification).

6.3 Discussion

A protective anti-WNV NS1 mAb, 22NS1, did not cross react with DENV-2 (Chung et al., 2006b). This study demonstrated that this mAb can cross-react with the more closely related C-*JEV*NS1 at the same epitope, but with some conformational flexibility. This finding agrees with our MD result (**Chapter 5**) showing elasticity in the epitope loop, which may affect the antibody-NS1 structure in solution. As 22NS1 may be able to interact with multiple viruses in the JE serocomplex, it could be useful in diagnostic testing and for passive antibody therapy development against related viruses (e.g., St. Louis encephalitis and Murray Valley encephalitis viruses). Moreover, the antibody may inhibit NS1 protein-protein interaction and limit viral activity. Even though the interaction partners of JEV NS1 are not fully defined, it is interesting to investigate further the inhibition of protein-protein interactions by the antibody. Even though the C-*JEV*NS1' has the extra amino acids, the C-terminal tail does not obstruct the binding surface of the WNV 22NS1 mAb. JEV NS1'-C can interact with WNV 22NS1 mAb. The C-tail may then locate at the side flanking the dimer. The presence of NS1' is a shared characteristic of JE serocomplex viruses. The protein may have specific protein interactions that consequently target the brain cells and lead to encephalitis in humans. Targeting NS1' may prove useful for cross-protective vaccination against the JE serocomplex.

Even though the flavivirus NS1 proteins have a conserved protein fold, these related proteins differ in their charge distribution so are capable of having unique interactions with host proteins. The fact that WNV 22NS1 mAb interacts positively with JEV NS1 is consistent with a close similarity of charge distribution of WNV and

JEV NS1. This similarity extends to ZIKV NS1 and may be a productive avenue for developing a common diagnostic and therapeutic strategy, which provide insight into the possible route for developing a strategy for diagnostic test and therapeutics for this group of *Flavivirus*.

A pull-down assay failed to detect protein association between C-*JEV*NS1 and complement proteins in human serum. However, as interaction to complement proteins were reported in DENV and WNV (Avirutnan et al., 2010, Avirutnan et al., 2011), JEV NS1 is still believed to bare an immune invasion function by interacting with complement proteins or other immune modulators. The interaction may occur via N-terminus or require full length protein. Moreover, it is possible that the interaction is weak and transient, so it is not accessible by pull down assay. Also, SDS-PAGE may not be sensitive enough to detect a small amount of interaction (low percentage) complement proteins. Higher sensitivity method is suggested, for example, mass spectrometry.

Sulphate molecules distributed on the NS1 surface agree with previous findings for DENV and ZIKV indicating its potential anionic ligand interaction, such as GAGs (Avirutnan et al., 2007). Although NS1 was thought to interact with uninfected cell membranes via these sulphate binding sites, our study shows that C-*JEV*NS1 cannot bind heparin, GAGs polymers, or liposomes. Thus, the sulphate binding sites are not GAG binding interfaces and could represent a crystallographic artifact. Our results suggest that cell membrane interactions via GAGs may occur at the β-roll and wing domains, as was suggested previously (Akey et al., 2014, Xu et al., 2016, Brown et al., 2016).

6.4 Conclusion

WNV 22NS1 antibody able to cross react to C-*JEV*NS1 by interacting at the flexible loop, the same epitope as WNV. No binding of C-*JEV*NS1 and complement proteins observed by Pull-down assay. C-*JEV*NS1 could not interact with either liposome, heparin, or GAGs polymers indicated no cell membrane association via GAGs. The study suggests a significant roles of the N-terminal domain for cell attachment and probably the complement proteins association.

Chapter 7 ZIKV NS1 and C terminal domain NS1 protein expression and purification

While the existence of ZIKV has been known for several decades it has not been at the forefront of public healthcare discussions until the recent outbreak in Brazil in 2015. ZIKV quickly spread to Central and South American countries and has affected millions of people. Even though the majority of patients have mild symptoms, ZIKV has been proven to relate to Guillain-Barré syndrome and babies born with microcephaly in pregnant women. As little was known about this virus, many projects and funding on ZIKV have emerged to respond to the situation. Every aspect of ZIKV study such as protein structure, immunology, diagnosis, vaccine development, and therapeutic are growing fast. Even though ZIKV is new to flavivirus research field and it has distinct disease phenotype, it is closely related to DENV based on NS5 gene (Weaver et al., 2016) and it may share typical characters to other flaviviruses. Laboratory protocols develop for other flaviviruses may applicable for ZIKV. A structural study of NS1 protein could develop the knowledge of this virus especially the disease distinction. ZIKV NS1, which has about 50% sequence identity compared to other *Flaviviruses* and is suspected to have the same protein fold, was expressed and purified with the same protocol as C-*JEV*NS1 in this study.

7.1 Materials and methods

7.1.1 Plasmid construction

Synthetic *ZIKV*NS1 and C-*ZIKV*NS1 were constructed in the pET15b vector at NdeI/BlpI cloning site. The protein constructs contain N-terminus histidine tag, TEV cleavage site, and followed by ZIKV NS1 or NS1 C-terminus.

7.1.2 Protein expression and purification

*ZIKV*NS1 and C-*ZIKV*NS1 were produced and purified by using modified Edelling's method previously described in chapter 3.

7.1.3 Protein crystallization

*ZIKV*NS1 was concentrated to ~4.4-4.6 mg/ml and screened by using SaltRx 1-2, PEGRx 1-2, and Natrix screens from Hampton research, and Structure, PACT premier, and JCSG screens from Molecular Dimension.

Apart from robot screen, crystallization of *ZIKV*NS1 was seeded with C-*JEV*NS1 needle crystals (**Figure 7.1**). C-*JEV*NS1 needle crystals from 0.7 M Li_2SO_4, 0.05 M MES pH 6 buffer condition were crashed into small pieces under a stereo microscope by fine tip glass rod made from heated Pasteur pipette and then vortex with some glass beads

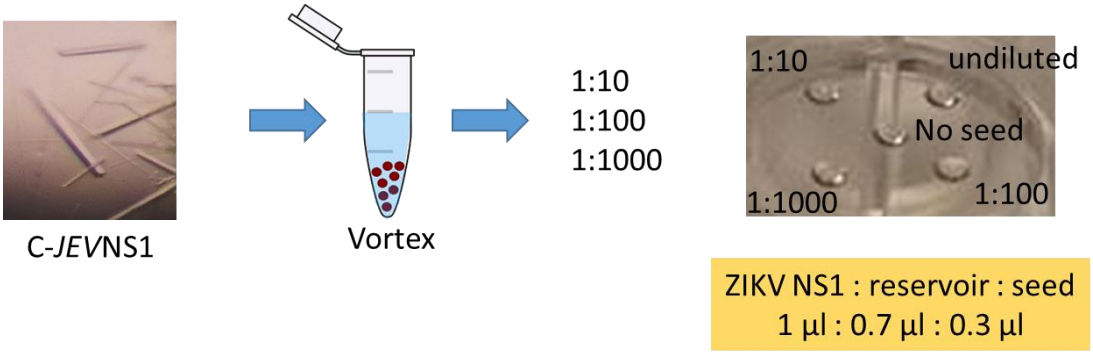


Figure 7.1 Crystal seeding for ZIKVNS1 crystallization

in a micro centrifuge tube for 10 cycles of 3 minutes vortex and 1 minute on ice interval. The seed stock was diluted to make 1:10, 1:100, and 1:1000 seed stocks. Crystal trays were set in hanging drop manner in 24-wells plates by using JCSG-plus Set 1 (formulation 1-24) and SaltRX Set 2 (formulation 1-24) commercial screens. On a cover slip of each well, 5 seed stock conditions were tested: no seed control, undiluted, 1:10, 1:100, and 1:1000, by mixing 1 µl of *ZIKV*NS1 protein, 0.7 µl of reservoir solution, and 0.3 µl of seed stock. Each well was filled with 500 µl of reservoir solution.

7.2 Results

7.2.1 Protein expression and purification

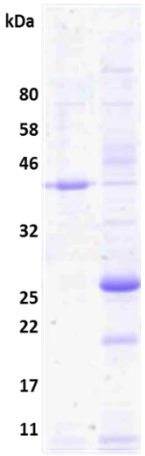


Figure 7.2 *ZIKV*NS1 and C-*ZIKV*NS1 after refolding
in lane 1 and 2, respectively

The protein inclusions were effectively refolded and purified as demonstrated by detection of protein in soluble fraction analysed by SDS-PAGE (**Figure 7.2**). From the protein sequences, the molecular weight of *ZIKV*NS1 and C-*ZIKV*NS1 is 40.1 kDa and 20.8 kDa, respectively. The SDS-PAGE appearance size of *ZIKV*NS1 was ~40 kDa, while C-*ZIKV*NS1 showed 2 bands at >25 and <22 kDa, which Western blot analysis against histidine tag confirmed that they all were the target proteins (**Figure 7.3**). By using SEC, I expected to see C-*ZIKV*NS1 eluted at the retention volume similar to C-*JEV*NS1 based on its similar MW and earlier in the case of NS1 full length. However, they all were eluted at a retention volume of ~16 ml, about the same retention volume as C-*JEV*NS1. *ZIKV*NS1 aggregated a lot during the refolding step and gave very low yield after SEC purification which might cause by buffer and pH that not appropriate to the protein (**Figure 7.3a**). Chromatogram of *ZIKV*NS1 showed similar profile to C-*JEV*NS1 with 3 elution peaks: void peak, middle peak, and end peak. SDS-PAGE analysis showed high protein purity of all 3 peak fractions (**Figure7.3**). Even the end peak had a high absorbance at UV 280 nm, a small amount of protein was shown in the SDS-PAGE gel. Moreover, the retention volume of peak 3 did not relate to protein size. Thus, only the middle peak was used for protein crystallization. C-*ZIKV*NS1 had less aggregation and also had 3 peak elution

profile with a much better yield of the second peak (**Figure 7.3b**). Two bands observed in refolding step were purified together in the middle peak and were not affected by heat treatment, but appeared as a single band in non-reduced condition. In addition, dimer protein at the size approximately double of the monomer, 40-50 kDa, was not observed (**Figure 7.3b**).

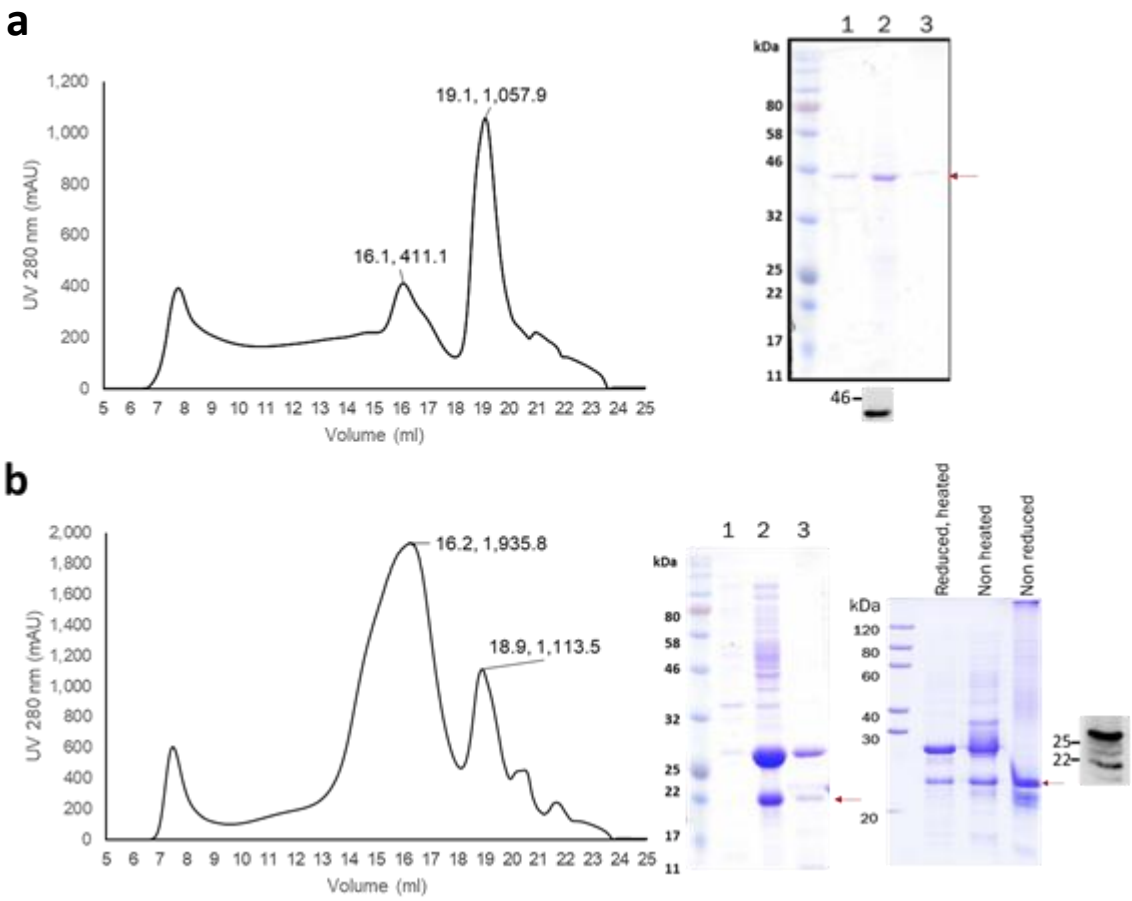


Figure 7.3 ZIKV NS1 and NS1 C terminal domain purification.

Elution profile of *ZIKV*NS1 **(a)** and C-*ZIKV*NS1 **(b).** Three peaks observed in chromatogram for each protein were analysed on SDS-PAGE in lane 1, 2, and 3, respectively. Bands at estimated size is indicated with red arrows. For C-*ZIKV*NS1 **(b)**, band size at ~22 kDa was shown with another higher band at >25 kDa. The 2 bands were combined when run under non-reducing conditions. Western blot analysis against the histidine tag is shown in greyscale beneath or on the right on the gel.

7.2.2 Protein crystallization

None of the *ZIKV*NS1 crystal tray set by the robot gave crystal, so seeding method was used. By seeding with C-*JEV*NS1, positive results were obtained from 4 formulations of SaltRX 2 and 2 formulations of JCSG-plus 1 (**Figure 7.4**, and **Table 7.1**). Conditions in figure a, b, d, g, and h looked like the solid phase of the protein. Crystals in figure c, e, and f were probed by a needle under a microscope which suggested that they were salt as the crystals were hard and not easy to break. The crystals were also tested at the in-house X-ray source, Barkla, University of Liverpool. In agreement with the previous tests, the crystals were salt as shown by sporadic mid high resolution. In the meantime, three crystal structures of ZIKA virus NS1 protein have appeared (Brown et al., 2016, Xu et al., 2016, Song et al., 2016). The solid phase conditions and C-*ZIKV*NS1 were not further investigated.

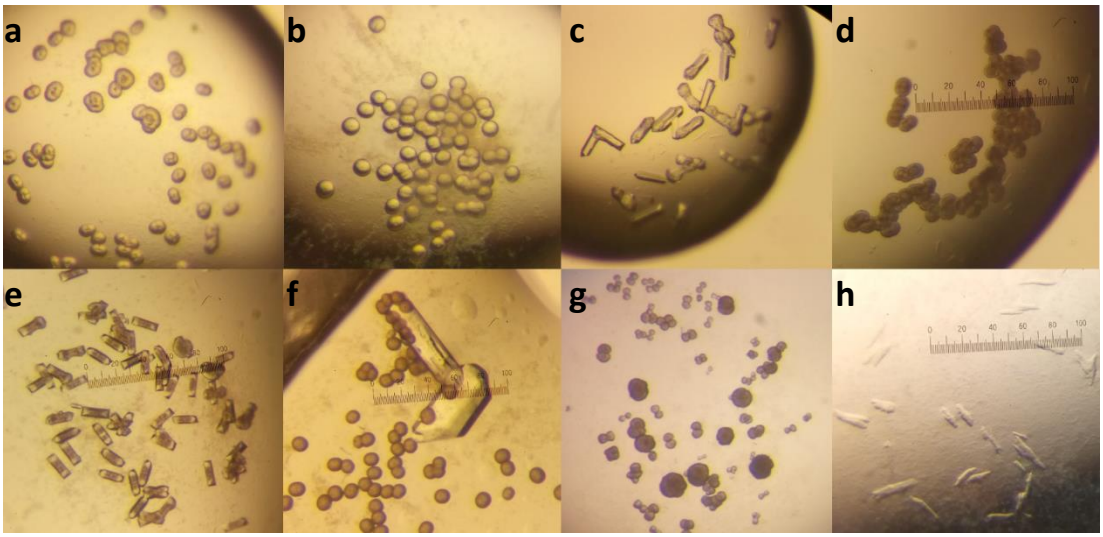


Figure 7.4 ZIKVNS1 seeding.

Positive results from SaltRX 2 screen **(a-f)** corresponding to buffer formulation 3, 4, 6, 7, 9, and 10, respectively, and from JCSG-plus 1 screen **(g, h)** corresponding to buffer formulation 11 and 17, respectively.

Table 7.1 Buffer formulations and seed stock solutions that gave crystals

Formulation	Seed stock				
	No seed	undiluted	1:10	1:100	1:1000
SaltRX 2: 0.1 M Tris pH 8.5, 1.5 M ammonium phosphate dibasic			✓		
SaltRX 2: 0.1 M Tris pH 8.5, 2.4 M ammonium phosphate dibasic			✓		
SaltRX 2: 1.0 M sodium phosphate monobasic monohydrate, potassium phosphate dibasic pH 6.9			✓		
SaltRX 2: 1.0 M sodium phosphate monobasic monohydrate, potassium phosphate dibasic pH 8.2		✓	✓	✓	✓
SaltRX 2: 1.8 M sodium phosphate monobasic monohydrate, potassium phosphate dibasic pH 6.9			✓		✓
SaltRX 2: 1.8 M sodium phosphate monobasic monohydrate, potassium phosphate dibasic pH 8.2		✓	✓		✓
JCSG-plus 1: 0.2 M ammonium hydrogen phosphate, 0.1 M Tris pH 8.5, 50% v/v MPD		✓	✓		✓
JCSG-plus : 0.1 M Sodium cacodylate pH 6.5, 40% v/v (+/-)-2-Methyl-2,4-pentanediol (MPD)	✓			✓	

7.3 Discussion

ZIKV NS1 both full length and C-terminal domain were successfully refolded by using the same protocol as JEV NS1, which had been adapted from the WNV NS1 protocol, indicating the similarity between the protein's properties. However, there were some significant differences in details of the protein purification and refolding. The *ZIKV*NS1 full length was more success to refold than JEVNS1. These may reflect that NS1 proteins of flaviviruses share common features and also have distinct features at the same time, which the diversity of the protein surface charge and hydrophobicity may

be the cause of the differences. From NS1 X-ray structures, NS1 is dimerized by hydrogen bonds (Akey et al., 2014, Edeling et al., 2014, Song et al., 2016). The dimer is usually dissociated when heated as reported in DENV, TBEV, and JEV (Winkler et al., 1988, Flamand et al., 1992, Crooks et al., 1994). Two populations of C-*ZIKV*NS1 at the size of ~25 and ~22 kDa were observed which were not detected in C-*JEV*NS1 expression or another C-*ZIKV*NS1 study (Song et al., 2016). The ~25 kDa band was not a protein dimer as it was detected even after the sample was boiled and the size was much smaller than the dimer (40-50 kDa on a gel). The upper band protein can form a complex with inter-molecular disulphide bonds as demonstrated by reduced size when analysed in non-reduced SDS-PAGE, while NS1 dimer normally intact when treated with a reducing agent (Winkler et al., 1988, Flamand et al., 1992). The ~22 kDa band protein is suspected to be the C-*ZIKV*NS1 monomer because it is the similar size when the protein was denatured (any interaction was broke) and it was unchanged in non-reducing condition. Moreover, C-*ZIKV*NS1 was unable to form a dimer without non-heat treatment. In Song, *et al*. ZIKV NS1 structural study, refolded ZIKV NS1 C-terminus was proved to be able to dimerize and has the same protein fold as DENV (Akey et al., 2014), WNV (Edeling et al., 2014), and JEV NS1 (from this study). Thus, the protocol used in this study might not suit C-*ZIKV*NS1 and need some optimization. The results indicate that C-*ZIKV*NS1 produced in this study has unique protein assembly and might have wrong protein fold. This finding could be an experimental error during the purification or it might hint the uniqueness of ZIKV NS1 protein.

Indistinguishable elution time of *ZIKV*NS1 and C-*ZIKV*NS1 probably because size determination by SEC is good for globular protein. C-terminus NS1 has a rod shape as

determined from other flavivirus NS1, so SEC may not reveal its true molecular weight on the chromatogram.

7.4 Conclusion

*ZIKV*NS1 was successfully refolded but still needed some optimization in crystallization. C-*ZIKV*NS1 in this study formed a unique complex interacted by disulphide bonds and unable to form a dimer, in contrast to the feature of NS1 proved to be correct in another C-*ZIKV*NS1 study (Song et al., 2016). C-*ZIKV*NS1 in this study is believed to have the unusual protein fold. Then, I focused on *ZIKV*NS1 structure determination in order to compare to other flavivirus NS1. In the competitive field of ZIKV study, other research groups had accomplished the goal before us (Brown et al., 2016, Xu et al., 2016, Song et al., 2016). *ZIKV*NS1 and C-*ZIKV*NS1 were not further investigated in this study.

Chapter 8 Structural study of JEV capsid protein

The capsid protein forms the inner shell to enclose virus genome which happens in all viruses. The protein plays an important role in viral encapsidation (genome packaging) and dissociation to release the genome. Consequently, it involves viral particle assembly and propagation. Raising of many evidence shows that flavivirus capsid protein may relate to different functions in viral life cycle such as RNA replication, and interaction with different host proteins interfering cell functions (Byk and Gamarnik, 2016). The available X-ray structure of WNV capsid (Dokland et al., 2004) and NMR structure of DENV capsid (Ma et al., 2004) have been solved but with poor resolution and statistics. In this study, the crystal structure of JEV capsid was obtained, which has never been visualized. Structure-function studies have the potential to increase our knowledge of a critical stage in viral transmission and provide information useful in its inhibition. Please note that the mass spectrometry analysis was done by Dr. Mark Wilkinson.

8.1 Methods

8.1.1 Plasmid construction

JEV capsid DNA (nucleotide residue 1-315) lacking the hydrophobic C-terminal (**Figure 1.** and **Figure 8.1**), which end at natural NS3 protease cleavage site (^{102}QNKR↓GGNE^{109}) (Shiryaev et al., 2007), was constructed in pET30a(+) vector at BamHI/XhoI cloning site. The protein construct is N-terminal histidine, S-tag, enterokinase cleavage site, and capsid protein (N-HIS-S tag-E-capsid) (**Figure 8.2a**).

Another capsid protein construct without S-tag was introduced into the same vector at NdeI/XhoI cloning site.

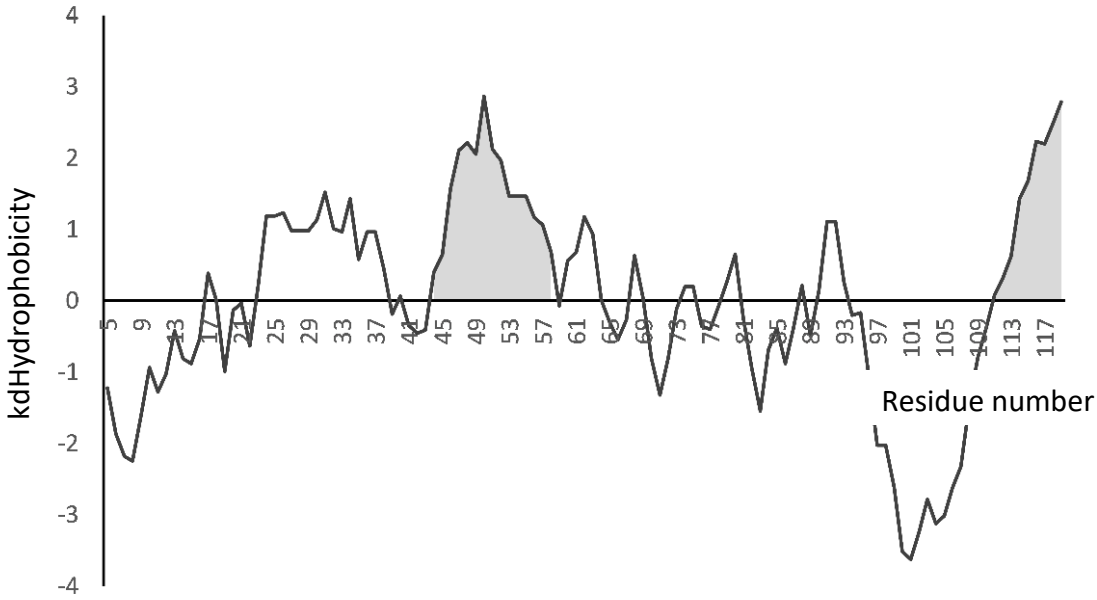


Figure 8.1 JEV capsid sequence hydrophobicity plot

The plot was produced by ProtScale (McWilliam et al., 2013). The amino acid Kyte&Doolittle scale was used. The middle hydrophobic region (Arg44- Ala58) corresponding to the α helix 2 and C-terminal hydrophobic are demonstrated in grey.

8.1.2 Protein expression and purification

Capsid protein was expressed in AIMTB at 30°C overnight in *E. coli* BL21(DE3). Lysis buffer screening was performed to identify the lysis buffer. Both first screen and salt solubility screen were tested. Finally, cells were lysed in high salt buffer (50 mM Tris pH 7.5, 1 M NaCl, 5 mM EDTA, 1 mg/ml lysozyme). Crude sample was purified by Ni-NTA column and dialyzed against 25 mM Tris pH 7.6, 50 mM NaCl, 2 mM $CaCl_2$ buffer overnight. The sample was then centrifuged at 16000xg for 20 minutes at 4°C to separate precipitation. S-tag was cleaved by enterokinase (New England Biolabs® inc., P8070) at 1:25, 1:50, and 1:100 ratio enzyme to capsid protein at 4°C overnight. The sample was purified by using Ni-NTA again to remove fusion tags. Flow through

fraction was subjected to superdex 75 10x300 mm column size-exclusion chromatography column (GE Healthcare Life Science). The peak at a retention volume of 13.5 ml (fraction A2-A4) was concentrated and used for crystallization.

8.1.3 Tricine-SDS-PAGE

Protein yield and purity were examined by Tricine-SDS-PAGE follow Schägger's protocol (Schagger, 2006). Protein was separated by 4% (stacking gel) and 16% polyacrylamide gel (resolving gel) at 0.75 or 1.0 mm thickness. See gel recipes in **Table 8.1**. Samples were mixed with 4x sample buffer (12% (w/v) SDS, 6% (v/v) mercaptoethanol, 30% (w/v) glycerol, 0.05% Coomassie blue G-250, 150 mM Tris HCl pH 7.0) at 3:1 sample to buffer ratio and heated at 95°C for 5 min. Electrophoresis was conducted at 200 volts constant in 1x anode buffer (100 mM Tris HCl pH 8.9) and cathode buffer (100 mM Tris pH~8.25, 100 mM Tricine, 0.1% SDS) by using Mini-PROTEAN®3 Cell electrophoresis system (BIO-RAD). At the end of the run voltage was increased to 300 V. Gel was fixed in 50% methanol, 10% acetic acid, 100 mM ammonium acetate for 1 hour before stained in 0.025% Coomassie blue G-250 in 10% acetic acid for 20 minutes or until the gel was stained consistency blue and sample dye cannot be seen. Then gel was destained in 10% acetic acid solution until the background was clear.

Table 8.1 Gel formulation (10 ml)

16% Resolving gel		4% Stacking gel	
3x Gel buffer (3 M Tris HCl, pH 8.45, 0.3% SDS)	3.3 ml	3x Gel buffer (3 M Tris HCl, pH 8.45, 0.3% SDS)	3.3 ml
Glycerol	1 ml	-	-
Water	-	Water	4.85 ml
30% Acrylamide/Bis	5.3 ml	30% Acrylamide/Bis	1.3 ml
10% APS	50 µl	10% APS	50 µl

16% Resolving gel		4% Stacking gel	
TEMED	5 µl	TEMED	5 µl

8.1.4 Mass-spectrometry (MS) analysis

Gel pieces of digested capsid and purified capsid (fraction A2-A4) from SDS-PAGE which was preserved in 20% ethanol were washed for 30 minutes twice with 50% acetonitrile, 0.2 M ammonium bicarbonate pH 8.9 and then dried in a rotary evaporator. The gel pieces were rehydrated in 2 M urea, 0.2 M ammonium bicarbonate pH 7.8 (RHB) containing 0.1 ug trypsin and incubated at 37°C overnight. Excess RHB was then removed to a new 1.5 ml microfuge tube and peptides were extracted from the gel pieces with 60% acetonitrile and 0.1% trifluoroacetic acid (TFA). The total peptide extract was then concentrated to 10 ul in a rotary evaporator and then desalted using C18 (200 Å pore size silica resin) ZipTips (Milipore) according to the manufacturer's instructions. MS analysis was performed using a MALDI-Tof instrument (Waters-Micromass) using a saturated solution of alpha-cyano-4 hydroxycinnaminic acid (CHCA) in 50% acetonitrile/0.1% trifluoroacetic acid. Samples were selected in the mass range of 850 – 2500 Da.

8.1.5 Protein crystallization

Capsid protein at the concentration of ~6mg/ml was screened by using commercial crystallization screens: the PEGRx screen from Hampton Research, the PACT premier and JCSG screens from Molecular Dimension. All the conditions that produce crystals are given in appendix 3. The 2 conditions: 10% v/v 2-propanol, 0.1 M BICINE pH 8.5, 30% w/v polyethylene glycol 1,500 and 18% v/v 2-propanol, 0.1 M sodium citrate tribasic dihydrate pH 5.5, 30% w/v polyethylene glycol 4,000 were further optimized.

Needle crystals produced from 18% v/v 2-propanol, 0.1 M sodium citrate tribasic dihydrate pH 4.6, 5.6, and 6.2, 16 and 18% w/v polyethylene glycol 4,000 were used for diffraction experiment. The crystals were flash frozen in reservoir solution pH 5.6 added with 25% ethylene glycol.

8.1.6 Diffraction experiment, data processing, and model building

X-ray data were collected at a cryogenic temperature at beamline PROXIMA 1 at Soleil synchrotron, France, I04 and I24 at Diamond Light Source, UK. Data were processed by HKL2000 (Otwinowski and Minor, 1997). Autoprocessing by Xia2 was used for processing of data collected at Diamond Light Source. The protein structure was determined by molecular replacement using the structure of WNV NS1 capsid protein (PDB ID 1SFK, 63% sequence identity) as a starting model. An automated model building was performed by Buccaneer. The structure was refined by REFMAC5 and built in COOT in CCP4. The data were collected 3 times each from a single crystal. The first model was refined with weight term of 0.115 and the difference between R factor and R free was greater than 0.05. The second and third data were refined against the previously built model with weight term of 0.15. TLS (Translation/Libration/Screw) refinement was used once and tight NCS restraint was set on both chain A and B. Data collection and refinement statistics are shown in **Table 8.2**. The JEV capsid refinement statistics of Ramachandran plot are 100% favoured and 0% outliers. The MolProbity score is 1.03.

Table 8.2 X-ray data collection and refinement statistics

	1st (I24 Diamond)	2nd (Soleil)	3rd (I04 Diamond)
Data collection			
Space group	$P2_12_12_1$	$P2_12_12_1$	$P2_12_12_1$
Cell dimensions			
a, *b*, *c* (Å)	46.1, 49.14, 67.75	45.6, 49.32, 67.74	46.31, 49.78, 68.25
α, β, γ (°)	90, 90, 90	90, 90, 90	90, 90, 90
Resolution (Å)	38.11-2.43(2.47-2.43)	39.82-2.3(2.38-2.3)	38.32-1.98(2.03-1.98)
R_{merge}	0.156(1.429)		0.058(0.862)
R_{pim}	0.07(0.616)	0.041(0.127)	0.044(0.686)
I / σI	7.8(1.6)	16.06(4.49)	12.1(1.4)
CC half	0.995(0.656)	0.999(0.918)	0.999(0.551)
Completeness (%)	100(99)	91.1(66.5)	99.6(99)
Redundancy	6.1(6.4)	5(3.3)	4.5(4.3)
Refinement			
Resolution (Å)	38.11-2.43	39.82-2.3	38.32-1.98
No. reflections	5839	5936	10871
R_{work} / R_{free}	0.194/0.255	0.185/0.266	0.188/0.237
No. atoms	1178	1203	1229
Protein	1142	1147	1144
Water	36	32	56
Ethylene glycol	-	16	16
2-propanol	-	8	-
Citrate ion	-	-	13
B-factors (Å)			
Protein	54.981	32.53	42.77
Water	57.276	35.25	53.26
Ethylene glycol	-	40.58	52.91
2-propanol	-	50.94	-
Citrate ion	-	-	66.22
R.m.s. deviations			
Bond lengths (Å)	0.02	0.01	0.01
Bond angles (°)	2.28	1.63	1.43

*Values in parentheses are for highest-resolution shell.

8.2 Results

8.2.1 Protein expression and purification of JEV capsid protein

JEV capsid without S-tag failed to express (data not shown). JEV capsid with S-tag was soluble in high salt lysis buffer (2 M NaCl) (**Figure 8.2b**). Later, the lower concentrations of salt were tested and JEV capsid protein was still soluble in 1 M NaCl buffer. The 20 kDa JEV capsid protein was successfully purified by Ni-NTA column with high protein yield and purity. Multiple bands were observed (~17 kDa and 15 kDa) after the protein was stored at 4°C (**Figure 8.2c**, first lane). The S-tag was cleaved by enterokinase at the 1:25, 1:50, and 1:100 ratio enzyme unit to protein. There was no difference among the 3 ratios, so 1:100 ratio was used for large scale protein preparation. After cleavage, the protein's apparent size reduced from 20 kDa to 15 kDa (**Figure 8.2c**) and smaller bands size <15 kDa, 10 kDa, and ~8 kDa were observed in some batch of protein after digestion (**Figure 8.2c**). As the capsid protein had no tag after S-tag was cleaved, the protein band was identified by size. The strong band that showed a reduction in size was suspected to be the capsid protein. The band sometimes appears as hand shape (**Figure 8.2d**, first lane). However, after digestion, all 3 extra bands at <15 kDa, 10 kDa, and ~8 kDa were shown in Ni-NTA flow-through, wash, and elute fraction (**Figure 8.2d**), which makes it difficult to distinguish the target protein. The bands corresponding to proteins of 10 kDa and ~8 kDa size were not observed in the first digestion experiment (**Figure 8.2c**). Multiple bands observed during the process that usually occurs after storage most probably due to the instability and inclining to degrade of the capsid protein. The size exclusion chromatography elution demonstrated 2 peaks at UV 220 nm (**Figure 8.3**). The first

peak contained 15 kDa and <15 kDa size proteins, while the second peak had only ~8 kDa size protein (**Figure 8.2d**). The first peak elution fractions A2 to A4 were concentrated together and appeared as the non-homogeneous band with the new size ~13 kDa (**Figure 8.2d**). Mass spectrometry confirmed that the 2 different size bands of the digested capsid and purified capsid (A2-A4) were the same protein, but N-terminus was lost after purification (**Figure 8.5**). The protein from fractions A2-A4 was used for crystallization and it gave crystals in several conditions (**Figure 8.4**) (see **Appendix 3**).

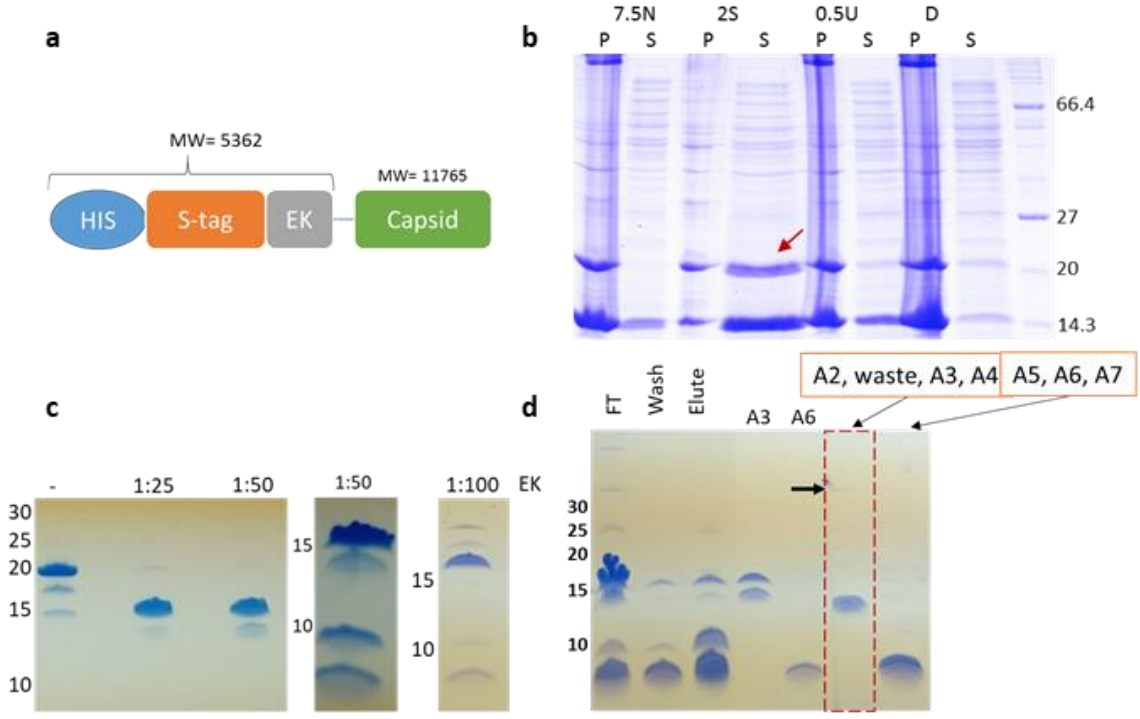


Figure 8.2 Capsid protein purification.

(a) JEV capsid protein construct. EK indicates enterokinase cleavage site. **(b)** Lysis buffer screening demonstrated that capsid protein soluble in 2 M salt buffer. Four lysis buffers were the buffer added with the following additives: no additive (**7.5N**), 2 M NaCl (**2S**), 0.5 M urea (**0.5U**), and 0.2% Triton X 100 (**D**). P indicates pellet fraction and S indicates suspension fraction. **(c)** Enterokinase digestion at 1:25, 1:50, and 1:100 enzyme unit to protein. Different batch 1:50 enterokinase digestion is shown. **(d)** Ni-NTA purification after enterokinase digestion. Lane 1 is flow-through from the Ni-NTA column. Lane 2 and 3 are the wash and elute fraction, respectively. Size exclusion chromatography elution fractions are in lane 5-7. Enterokinase enzyme is shown as a faint band at ~30 kDa (indicated with arrow).

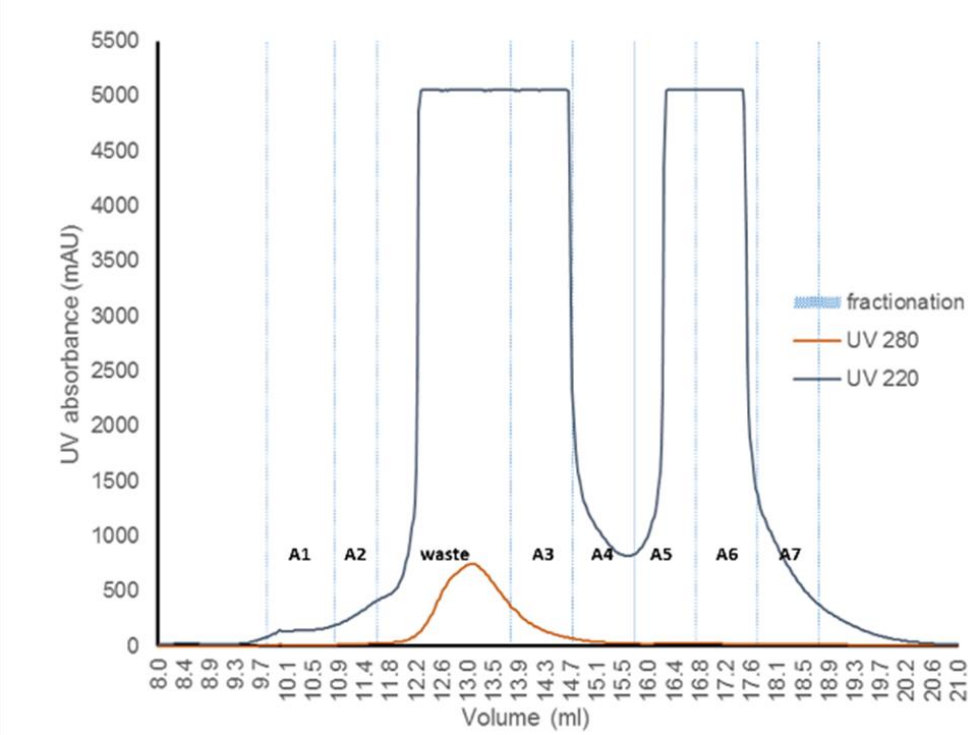


Figure 8.3 Capsid protein size exclusion chromatography profile.

At UV 280, only one peak is observed at retention volume ~13 ml. At UV 220, there are 2 very high peaks at retention volume ~13 ml (fraction A2-A4) and ~17 ml (A5-A7).

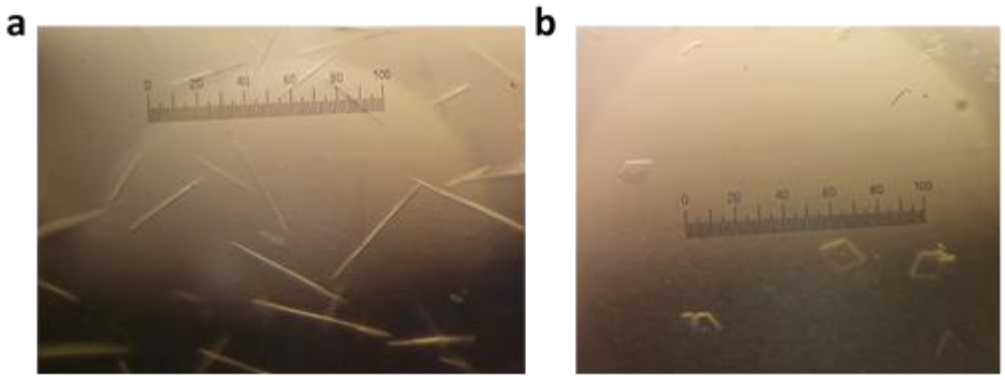


Figure 8.4 Capsid protein crystallization.

(a) Needle crystals were obtained from 10% v/v 2-propanol, 0.1 M BICINE pH 8.5, 30% w/v polyethylene glycol 1,500. **(b)** Thin rhomboid crystals were obtained from 18% v/v 2-propanol, 0.1 M sodium citrate tribasic dihydrate pH 5.5, 30% w/v polyethylene glycol 4,000. This condition can also give needle crystals.

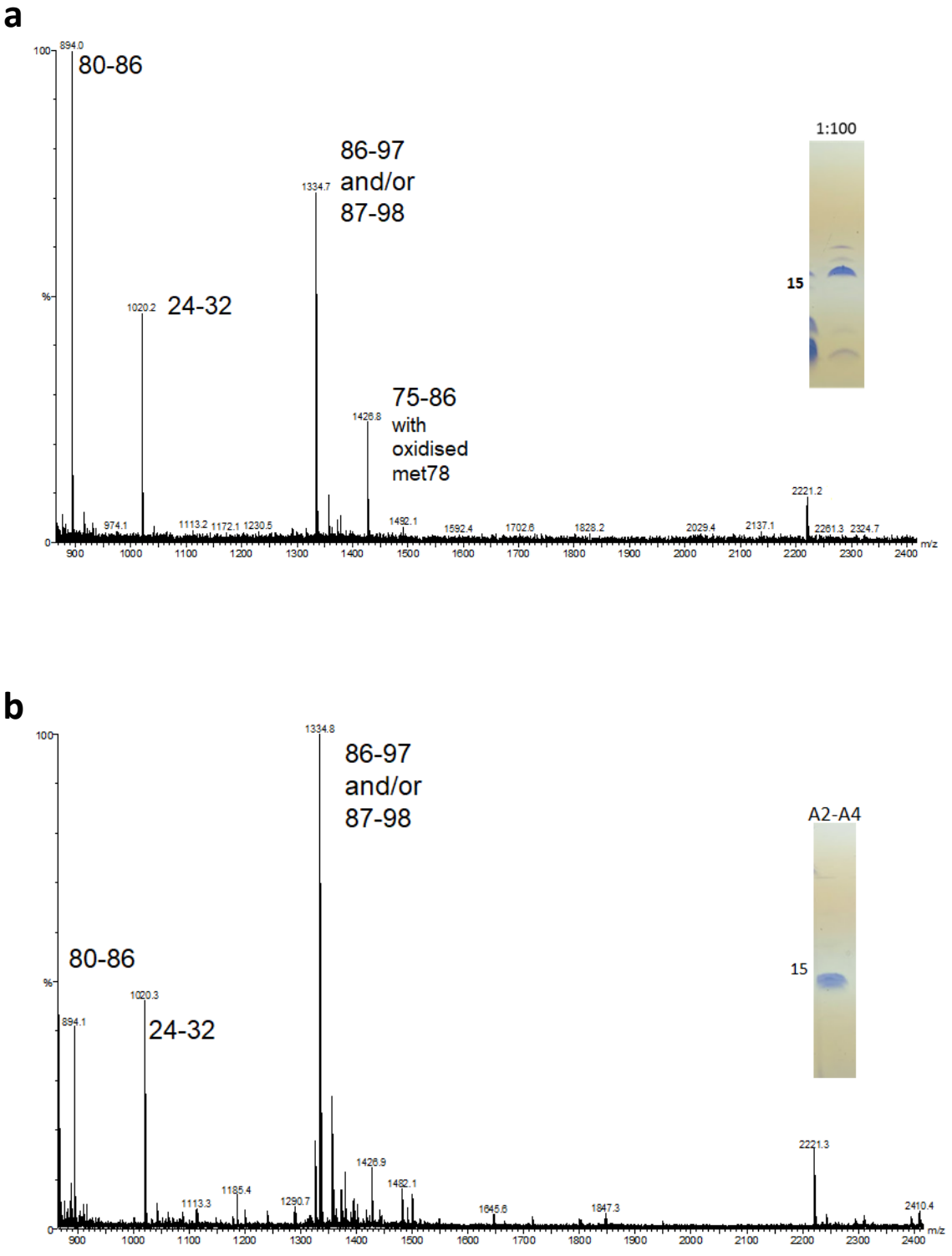


Figure 8.5 Mass-spectrometry analysis of JEV capsid protein

(a) MS analysis of digested and **(b)** purified capsid (fraction A2-A4). SDS-PAGE gels of the samples used for MS analysis are shown. Residue numbers are indicated next to the peaks.

8.2.2 Structure of JEV capsid protein

Crystals of full length JEV capsid protein diffracted to 1.98 Å resolution. Data collection and refinement statistics are shown in **Table 8.2**. The first 25 residues at N-terminus are not visible. Same as DENV and WNV, the crystal structure reveals a dimer in the asymmetric unit. Each monomer is composed of 4 helices: α1 (amino acid 30-38), α2 (44-57), α3 (63-71), and the longest α4 (74-94), connected by short loops (**Figure 8.6a-b**). The dimer is created by the anti-parallel pairing of the α1-α1', α2-α2', and α4-α4' connected with hydrogen bonds and salt bridges (**Table 8.3, Table 8.4**, and **Figure 8.7**). The hydrogen bonds also hold the α2 and α 4 together. From the side view, there are 3 layers of protein and flanking by the α3 and α3' (**Figure 8.7c-d**). Almost half of the protein are dimer interfacing residues, 43 residues out of 105 (**Table 8.5**) and the surface is quite hydrophobic (**Figure 8.8b, c**). However, it is concealed after protein dimerization and the rest of the surface is hydrophilic (**Figure 8.8a**). At neutral pH, the dimer net charge is +19 and the electrostatic surface map indicates total positive surface charge (**Figure 8.9**). The higher-order structure is not observed. Structural homolog searching was performed by using Dalilite v.3 (Holm and Rosenstrom, 2010). Except for the flavivirus itself, JEV capsid resemblance to the transcription factor II B (TFIIB) subunit of yeast polymerase II transcription initiation complex (PDB code 5FYW; Z= 4.5, RMSD = 2.9) and the human CCR4-NOT transcription complex subunit 1 (CNOT1) (4CQO; Z = 4.4, RMSD = 3.0). They are both gene regulatory proteins that involved in nucleic acid interactions. Only a monomer of JEV capsid (α1- α4) could align to Asn124-Lys217 of the TFIIB which is the DNA interacting subunit with the α4 closest to the DNA fragment (**Figure 8.10a**). CNOT1

directly interacts with mRNA. However, the RNA binding site of CNOT1 has not revealed yet. The monomer of JEV capsid (α2- α4) aligns to N-terminal of CNOT1 at Tyr1842-Pro1921 (**Figure 8.10b**).

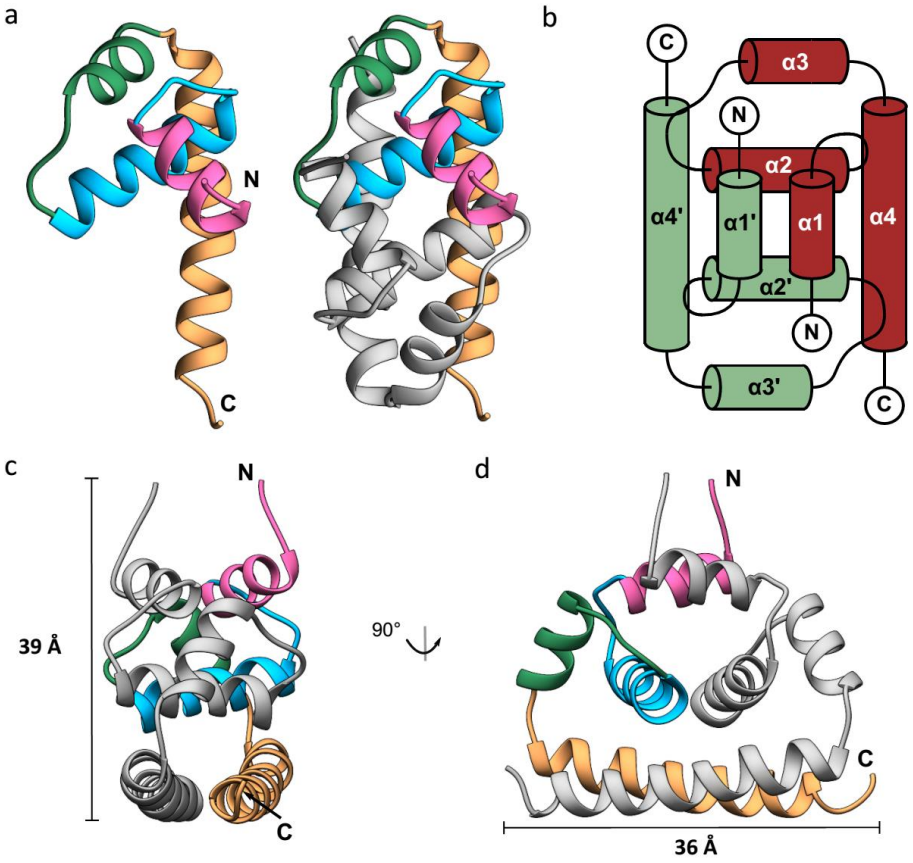


Figure 8.6 Structure of JEV capsid.

(a) Ribbon model of JEV capsid monomer coloured by the 1-4 α helices in pink, blue, green, and sandy brown, respectively. Another subunit is in grey. **(b)** Topology diagram of JEV capsid dimer. The α helices of one subunit are indicated with apostrophe symbol. **(c, d)** Side view of JEV capsid dimer with the dimension of 39x36 Å.

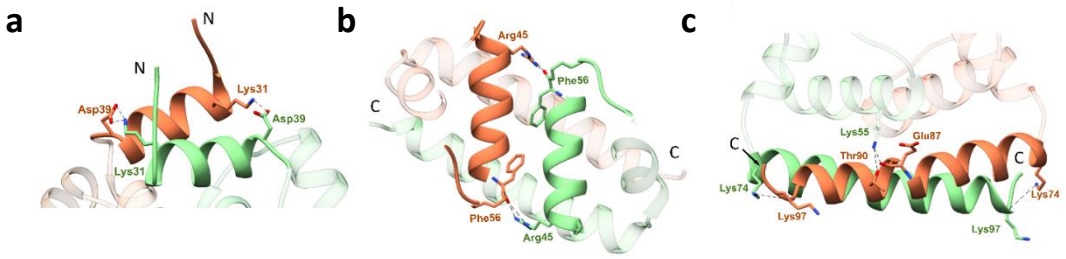


Figure 8.7 Hydrogen bonds at JEV capsid dimer interface.

Hydrogen bonds at the α1-α1', α2-α2', and α3-α3' are shown with dashed lines in figure **(a)**, **(b)**, and **(c)**, respectively. One monomer is in orange. Another is in green.

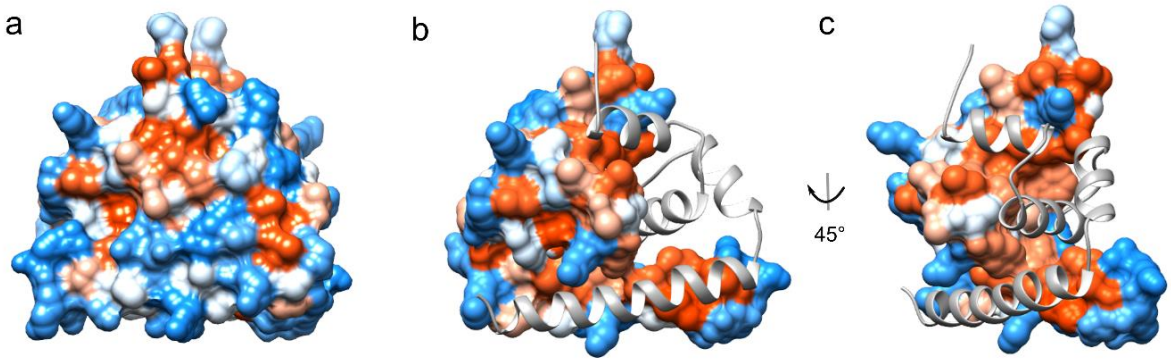


Figure 8.8 Hydrophobic surface coloured by Kyte&Doolittle scale.

(a) Hydrophobic surface of the dimer. **(b, c)** Hydrophobic surface of the dimer with one subunit in ribbon model. Surface colour ranging from dodger blue for the most hydrophilic to white for neutral and orange red for the most hydrophobic.

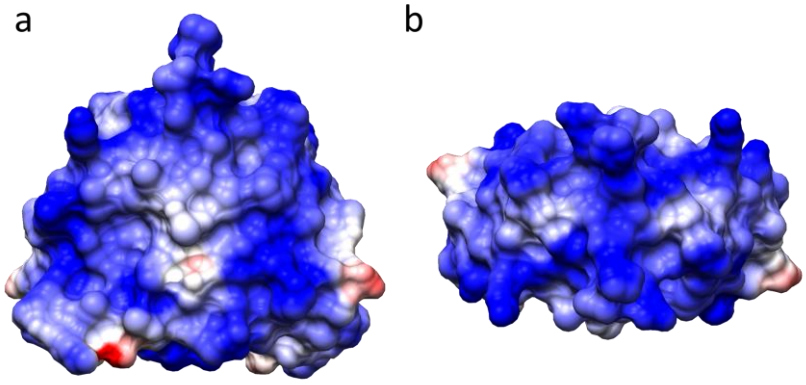


Figure 8.9 Coulombic surface colouring of the capsid dimer.

Electrostatics potential was calculated according to Coulomb's law with thresholds ± 10 kcal/mol*e.

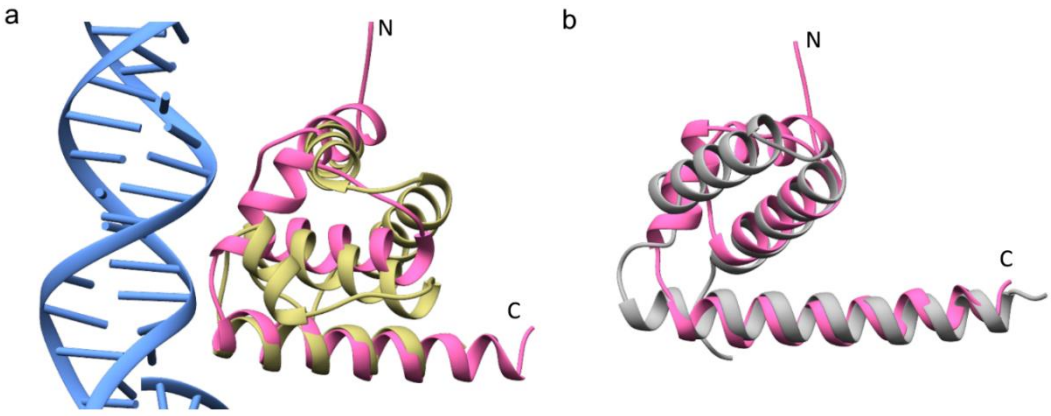


Figure 8.10 Superimposition of structural homologs

(a) JEV capsid in pink colour superimposed to TFIIB subunit of transcription initiation complex (PDB 5FYW) and **(b)** human CNOT1 (PDB 4CQO).

Table 8.3 Hydrogen bonds between JEV capsid dimer interfacing residues and the distance

Number	Structure 1	Distance (Å)[1]	Structure 2
1	Lys 31 [NZ][2]	2.64	Asp 39 [OD1][2]
2	Arg 45 [NE][2]	3.11	Phe 56 [O]
3	Arg 45 [NH2]	3.13	Phe 56 [O]
4	Arg 98 [NH2]	3.36	Trp 69 [O]
5	Arg 98 [NH1}	2.94	Val 72 [O]
6	Lys 85 [NZ]	3.87	Lys 85 [O]
7	Lys 55 [NZ}	3.03	Glu 87 [O}
8	Lys 74 [NZ]	2.77	Arg 98 [O]
9	Asp 39 [OD1]	2.69	Lys 31 [NZ]
10	Phe 56 [O]	3.16	Arg 45 [NH2]
11	Phe 56 [O]	3.07	Arg 45 [NE]
12	Glu 87 [O]	3.00	Lys 55 [NZ]
13	Arg 98 [O]	2.98	Lys 74 [NZ]

[1] Assembly analysis in the program PISA.

[2]See lysine, aspartic acid, and arginine atoms nomenclature in Appendix 5.

Table 8.4 Salt bridges between JEV capsid dimer interfacing residues and the distance

Num15ber	Structure 1	Distance (Å)[1]	Structure 2
1	Lys 31 [NZ]	2.64	Asp 39 [OD1]
2	Lys 31 [NZ]	3.92	Asp 39 [OD2]
3	Lys 74 [NZ]	2.77	Arg 98 [O]
3	Asp 39 [OD1]	2.69	Lys 31 [NZ]
4	Asp 39 [OD2]	4.00	Lys 31 [NZ]
5	Arg 98 [O]	2.98	Lys 74 [NZ]

[1] Assembly analysis in the program PISA.

[2]See lysine, aspartic acid, and arginine atoms nomenclature in Appendix 5.

Table 8.5 JEV capsid dimer interfacing residues reported with accessible (ASA) and buried surface area (BSA), solvation energy effect (ΔG) and conservation score.

(The values are reported from one monomer)

Number	Residue	ASA (Å^2)	BSA (Å^2)	ΔG (kcal/mol)	Conservation
1	Leu27	84.03	20.42	0.33	1
2	Val30	118.46	92.59	1.48	2*
3	Lys31	141.43	88.01	0.13	8
4	Val33	68.52	17.91	0.29	3*
5	Val34	63.66	54.72	0.93	6*
6	Met35	103.38	55.26	1.33	1
7	Leu37	11.11	10.11	0.16	5*
8	Leu38	49.81	39.11	0.61	3*
9	Asp39	71.47	20.51	-0.37	2*
10	Arg45	83.50	17.61	-1.16	6*
11	Phe46	94.02	43.47	0.68	2*
12	Ala49	41.64	41.64	0.57	7
13	Leu50	24.48	24.48	0.39	2*
14	Ile51	18.69	7.03	0.11	5*

Number	Residue	ASA (Å^2)	BSA (Å^2)	ΔG (kcal/mol)	Conservation
15	Thr52	28.35	28.11	0.17	5*
16	Phe53	94.72	94.08	1.51	7*
17	Phe54	22.93	16.25	0.26	5*
18	Lys55	164.08	90.08	-1.12	7*
19	Phe56	144.14	141.80	1.46	6*
20	Thr57	62.62	45.02	0.44	6*
21	Leu59	102.61	51.41	0.82	7*
22	Ala60	89.91	20.19	-0.12	5*
23	Thr62	52.55	26.77	0.43	8
24	Trp69	89.96	32.30	0.52	8
25	Lys74	157.5	81.09	-0.41	6*
26	Ala77	17.07	11.38	0.18	9
27	Met78	106.84	41.91	1.07	7*
28	Leu81	58.55	58.55	0.81	8
29	Thr82	69.27	21.75	0.35	2*
30	Phe84	36.05	36.05	0.51	7*
31	Lys85	157.40	87.05	-0.32	7
32	Glu87	97.04	48.21	0.17	6*
33	Leu88	79.65	79.65	1.24	7
34	Gly89	30.26	14.10	0.20	5*
35	Thr90	92.33	10.55	0.17	4*
36	Leu91	109.62	92.39	1.46	7
37	Ile92	91.20	74.11	1.18	6*
38	Val95	70.00	67.32	0.92	4*
39	Asn96	116.67	40.08	-0.45	8
40	Lys97	167.41	0.31	0.00	3*
41	Arg98	241.04	124.28	-0.98	9

*Below the confidence cut-off - The calculations for this site were performed on less than 6 non-gaped homologue sequences, or the confidence interval for the estimated score is equal to- or larger than- 4 colour grades.

8.2.3 Structure of JEV capsid protein compared to WNV and DENV

Compare to DENV capsid NMR structure and WNV X-ray structure, the flavivirus capsid protein fold is conserved with Cα RMSD of 2.1 Å and 1.4 Å, respectively, except for the helices 1 which is oriented in the different position **(Figure 8.11**). The α1- α1' of JEV and WNV are on top of the α2- α2' covering the hydrophobic surface of the α2- α2', whereas in DENV the α1 is not paired with α1' and move aside perpendicularly to the helices 4 when looking from the top view (**Figure 8.11b, c**). This movement allows the hydrophobic patch at the α2- α2' surface to expose **Figure 8.11d, e**).

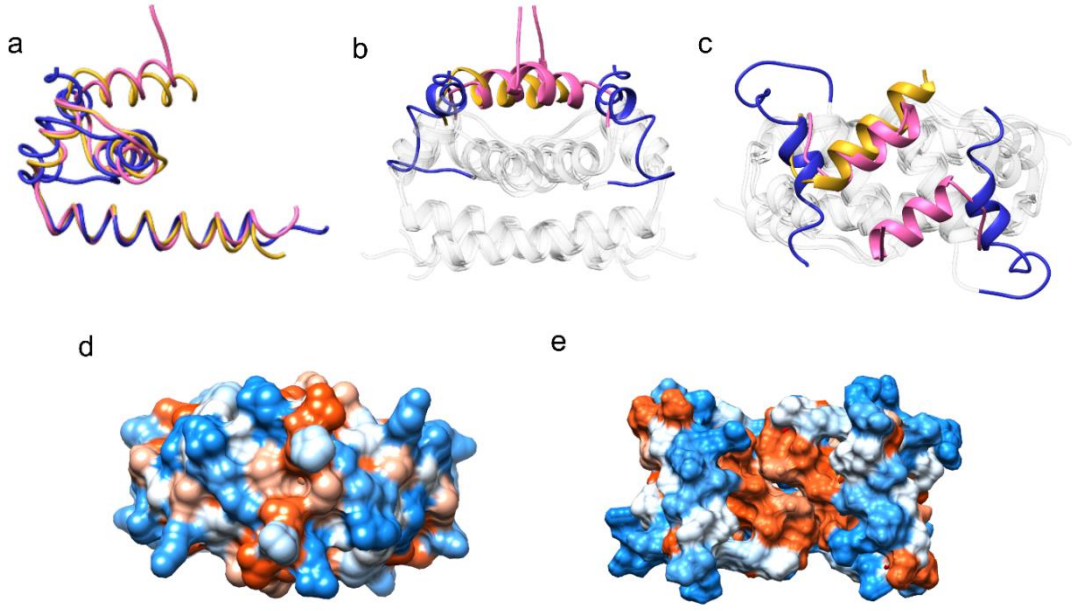


Figure 8.11 JEV capsid superimposing to WNV and DENV capsid proteins.

(a) Superimposition of of WNV capsid (PDB 1SFK) in gold and DENV capsid (1R6R) in blue to JEV capsid in pink. **(b)** superposition of the α1-α1' of WNV, DENV, and JEV from the side view and **(c)** top view in the same color scheme. Please note that the helix 1 in chain B of WNV capsid is missing. **(d)** Hydrophobic surface of JEV capsid from the top view showing closed position and **(e)** DENV capsid from the top view showing opened position. Surface colour ranging from dodger blue for the most hydrophilic to white for neutral and orange red for the most hydrophobic.

8.3 Discussion

Other flavivirus capsid structures have been solved by different research groups using both NMR (Ma et al., 2004) and X-ray crystallography (Dokland et al., 2004). One common feature of the protein is the unstable N-terminus. The first 20 residues of DENV capsid (PDB 1R6R) are conformational labile and WNV capsid (1SFK) is stable from residue 23. The helix 1 of JEV capsid may well be longer than we observed as the first 25 residues from N-terminal are not visible. It is consistent with the expression profile and mass spectra that the protein is degraded from N-terminus. Moreover, it is possible that the protein degradation happened during the JEV capsid protein purification creating a non-uniform N-terminus which gave poor density in this region leading to a non-structure of the first 4 residues. Compared to WNV capsid, the helix 1 of WNV capsid from residue 24 is well defined (Dokland et al., 2004). It maybe because the WNV capsid was trypsinized to generate stable fragments before crystallization.

With a large interfacing surface, there are only 9 hydrogen bonds between the dimer. However, the dimer is packed by inclusive hydrophobic interaction. The fact that capsid monomer bears a large hydrophobic patch which later not expose after dimerization leads to the conclusion that the capsid protein may preferable in dimeric form because the monomer protein might not be able to solubilize. Furthermore, in our study, JEV capsid was isolated from bacteria cells only in high salt buffer (1-2 M NaCl) similar to what have found in TBEV capsid (Kiermayr et al., 2004). High concentration of salt may facilitate the capsids to interact with each other and form a stable dimer. Even after buffer exchange to low salt buffer (50 mM NaCl), the dimer still intact. Importantly, dimeric stage of flavivirus capsid is confirmed by structural

studies of DENV and WNV (Dokland et al., 2004, Ma et al., 2004). This supports the notion that capsid dimer is probably the building block of the nucleocapsid (Mukhopadhyay et al., 2005, Kiermayr et al., 2004).

RNA binding site of flavivirus capsid was mapped to N-terminal and C-terminal of WNV (Khromykh and Westaway, 1996) but was mapped to the middle hydrophobic region of JEV (Tseng et al., 2007). In our study, the homolog search matched the C terminus of the JEV capsid monomer to the TFIIB subunit of transcription complex, in partial agreement with the WNV study. Besides this, the 3D structure revealed that the middle hydrophobic region or the helix 2 is obscured under the α1-α1' making it more difficult to access. In addition, Cryo-EM structures of flavivirus virions suggest the capsid protein is poorly ordered, which may reflect random interactions between capsid and RNA (Zhang et al., 2003, Mukhopadhyay et al., 2005). These observations might explain the difference in RNA binding site mapping (Khromykh and Westaway, 1996). Other positive-stranded RNA viruses typically exhibit RNA binding as a function of the N-terminus of the protein, which binds in a non-icosahedral manner (Zhang et al., 2003). The flavivirus capsid is very small and the whole protein might function as RNA-binding similar to N-terminal of other positive stranded RNA viruses (Zhang et al., 2003, Mukhopadhyay et al., 2005).

Two conformations of the α helix 1 were detected; close conformation in JEV and WNV capsid crystal structures, and open conformation in DENV NMR structure. The open conformation of capsid protein allows more access to the α2-α2' hydrophobic patch, but it is likely to make hydrophobic interaction to, for example, cell membrane rather than interaction with a negatively charged molecule such as RNA. NMR

structure might allow more flexibility than crystal structure, but it does not explain the protein disorder in WNV capsid structure (helix 1 of chain B is missing). Thus, flexibility is an inherent characteristic of the N-terminus itself. The buffer pH did not affect the conformation as 3 of capsid crystals used in this study were in buffer pH 4.6, 5.6 and 6.2 gave the same fold of the protein. DENV capsid was in pH 6 buffer, while WNV capsid was in pH 10.5 buffer. The factor that triggers conformational changing of capsid protein is unclear, but the flexibility of helix 1 indicates that it may bear functionally importance roles such as lipid bilayer and LDs association and higher order structure assembly.

Capsid protein is known to form a spherical core enclosing viral genome. However, cryo-EM structural studies showed that the maximum height of nucleocapsid density was low (~25-50%) compared with the envelope (Zhang et al., 2003, Kuhn et al., 2002). This, together with the missing of an oligomeric stage of DENV NMR structure and in our study, suggests that flavivirus nucleocapsid is disordered (Kuhn et al., 2002, Zhang et al., 2003, Wang et al., 2017, Zhang et al., 2013). It is also possible that lacking higher order structure is due to the conformation of the protein. Protein assembly probably occurs after conformational changing obtain by certain stimulation. Moreover, as virus assembly is a complex process and occurs with coordinating factors, an *in vitro* experiment that contains only capsid protein might not generate the core protein assembly process. Addition of interacting molecules is possibly required; for example, capsid-like particles were successfully produced from dimeric capsid isolated from TBEV virions incubated with viral RNA (Kiermayr et al., 2004).

8.4 Conclusion

JEV capsid protein has an unstable N-terminus and also has conserved protein folds compared with DENV and WNV. The N-terminus corresponding to α helix 1 in the 3D structure is flexible which might be functionally important. However, the factors governing the conformational change are unclear. A change in conformation of α helix 1, might in turn expose the α2-α2' hydrophobic in order to allow interactions with lipid bilayers or LDs. JEV capsid protein C-terminus, which corresponds to helix 4, may interact with RNA as suggested by the RNA binding study of WNV capsid and homolog protein search in this study. Capsid oligomerization to form nucleocapsid was not observed but may consist of capsid dimers as a building block.

Chapter 9 General Discussion

Proteins are the machinery of all organisms including viruses. Understanding of protein is the key to understand the building block of life. In most of the studies, for example, this work, even though the target proteins were selected for the potential to be drug targets, this fundamental knowledge is also applicable to several topics covering disease aetiology, vaccine, treatment and diagnosis development.

The target proteins in this work, NS1 and capsid protein, were chosen for their multifunctional character involved in crucial roles in the viral life cycle and also to complete JEV structural proteomics study. The two are potential targets for drug development.

Flavivirus capsid proteins are not very conserved in sequence, but structurally conserved. This study presented the first JEV and the third flavivirus capsid protein structure available. Despite the lack of functional study, evidence provided here demonstrated important capsid protein properties such as the dimer building block of the nucleocapsid and α-helix 1 conformation.

Nine structures of flavivirus NS1 proteins have been described before this study started. The first structures were DENV and WNV NS1 full length proteins published in 2014. Even only NS1 and NS1' C-terminal domain are presented in this study, it is useful as no JEV NS1 protein structure has been revealed before. This protein truncation helps to link protein region to its functions, for example, cell membrane and GAG binding may locate to its N-terminal. In addition, even C-*JEV*NS1 protein

structure is almost identical to WNV, DENV, and ZIKV which described previously, several differences are pointed out in this study including protein surface charge, loop flexibility, and WNV antibody cross-interaction.

What we have learned here may be just the tip of the iceberg. The NS1 protein may not be amenable as a drug target as the DENV whole-genome sequencing research suggested. As the result of error-prone NS5 (viral RNA polymerase), the viral genome is diverse from the single-nucleotide variants (SNVs) which accumulated on prM, E and NS1 gene on viral genome derived from human cases (Sim et al., 2015). The study suggested that drug-resistance mutations are more probable in these areas. However, this field of study is emergent and there is still some controversy over the data (Wash and Soria, 2015).

With the fast development of instruments and understanding of structural studies (both X-ray crystallography and cryo-EM) and protein production, the breakthrough of difficult to obtain protein structures like membrane proteins or large protein complex structures has become true. Virus replication complex or assembled capsid structures are very probable in the near future. These will provide us the information of protein-protein interaction from the actual working part.

Nowadays, there are still some boundaries of flavivirus endemic areas. For example, JEV has not established yet in the Australian mainland or America. JEV genotype 4 and 5 are restricted in the Malaysia-Indonesia area (no human case outside this region). WNV is majorly present in the western hemisphere. However, migration of birds (reservoir hosts), globalization, changing agricultural practices, and climate change may influence the introduction of viruses to the naïve regions. Although the

establishment of exotic viruses depends on many factors such as vectors, reservoir hosts, and climate, we have seen ZIKV that successfully adapted and caused outbreaks in many countries in a short time. While JEV is preventable by vaccination and a DENV vaccine has been recently introduced, there is no vaccine for WNV and ZIKV. Moreover, immune cross-reactivity between endemic virus infection, which people in the endemic region are usually pre-exposed or vaccinated, and the viruses that have not been circulated in the area before, is not well understood. It could be a cross-protection or adverse effect between the two as has been observed in different serotypes of dengue virus. Besides that, if the viruses co-circulate in the same region, the specific and sensitive diagnosis will be required due to the high serological cross-reactivity of flaviviruses. Currently, a combination of more than one enzyme-linked immunosorbent assays is needed to be able to specifically distinguish JEV from WNV (Latif et al., 2015) or ZIKV from previous DENV infection (Tsai et al., 2017).

Therefore, not only JEV but flaviviruses are a critical concern to global public health. Although the findings of this study cannot clearly describe the full protein function, we are building up knowledge, which will lead us to a greater understanding of virus biology, especially at the molecular level which will facilitate the next-generation of diagnostic, vaccine, and treatment development.

References

ABERGEL, C. 2013. Molecular replacement: tricks and treats. *Acta Crystallogr D Biol Crystallogr,* 69**,** 2167-73.

AHMAD, M., HIRZ, M., PICHLER, H. & SCHWAB, H. 2014. Protein expression in Pichia pastoris: recent achievements and perspectives for heterologous protein production. *Appl Microbiol Biotechnol,* 98**,** 5301-17.

AKEY, D. L., BROWN, W. C., DUTTA, S., KONWERSKI, J., JOSE, J., JURKIW, T. J., DELPROPOSTO, J., OGATA, C. M., SKINIOTIS, G., KUHN, R. J. & SMITH, J. L. 2014. Flavivirus NS1 structures reveal surfaces for associations with membranes and the immune system. *Science,* 343**,** 881-5.

AKEY, D. L., BROWN, W. C., JOSE, J., KUHN, R. J. & SMITH, J. L. 2015. Structure-guided insights on the role of NS1 in flavivirus infection. *Bioessays,* 37**,** 489-94.

ALEYAS, A. G., HAN, Y. W., GEORGE, J. A., KIM, B., KIM, K., LEE, C. K. & EO, S. K. 2010. Multifront assault on antigen presentation by Japanese encephalitis virus subverts CD8+ T cell responses. *J Immunol,* 185**,** 1429-41.

AMBERG, S. M. & RICE, C. M. 1999. Mutagenesis of the NS2B-NS3-mediated cleavage site in the flavivirus capsid protein demonstrates a requirement for coordinated processing. *J Virol,* 73**,** 8083-94.

AMORIM, J. H., ALVES, R. P., BOSCARDIN, S. B. & FERREIRA, L. C. 2014. The dengue virus non-structural 1 protein: risks and benefits. *Virus Res,* 181**,** 53-60.

AVIRUTNAN, P., FUCHS, A., HAUHART, R. E., SOMNUKE, P., YOUN, S., DIAMOND, M. S. & ATKINSON, J. P. 2010. Antagonism of the complement component C4 by flavivirus nonstructural protein NS1. *J Exp Med,* 207**,** 793-806.

AVIRUTNAN, P., HAUHART, R. E., SOMNUKE, P., BLOM, A. M., DIAMOND, M. S. & ATKINSON, J. P. 2011. Binding of flavivirus nonstructural protein NS1 to C4b binding protein modulates complement activation. *J Immunol,* 187**,** 424-33.

AVIRUTNAN, P., ZHANG, L., PUNYADEE, N., MANUYAKORN, A., PUTTIKHUNT, C., KASINRERK, W., MALASIT, P., ATKINSON, J. P. & DIAMOND, M. S. 2007. Secreted NS1 of dengue virus attaches to the surface of cells via interactions with heparan sulfate and chondroitin sulfate E. *PLoS Pathog,* 3**,** e183.

BAKER, N. A., SEPT, D., JOSEPH, S., HOLST, M. J. & MCCAMMON, J. A. 2001. Electrostatics of nanosystems: application to microtubules and the ribosome. *Proc Natl Acad Sci U S A,* 98**,** 10037-41.

BANEYX, F. 1999. Recombinant protein expression in Escherichia coli. *Curr Opin Biotechnol,* 10**,** 411-21.

BATTYE, T. G., KONTOGIANNIS, L., JOHNSON, O., POWELL, H. R. & LESLIE, A. G. 2011. iMOSFLM: a new graphical interface for diffraction-image processing with MOSFLM. *Acta Crystallogr D Biol Crystallogr,* 67**,** 271-81.

BERENDSEN, H. J. C., VAN DER SPOEL, D. & VAN DRUNEN, R. 1995. GROMACS: A message-passing parallel molecular dynamics implementation. *Computer Physics Communications,* 91**,** 43-56.

BHATT, S., GETHING, P. W., BRADY, O. J., MESSINA, J. P., FARLOW, A. W., MOYES, C. L., DRAKE, J. M., BROWNSTEIN, J. S., HOEN, A. G., SANKOH, O., MYERS, M. F., GEORGE, D. B., JAENISCH, T., WINT, G. R., SIMMONS, C. P., SCOTT, T. W.,

FARRAR, J. J. & HAY, S. I. 2013. The global distribution and burden of dengue. *Nature,* 496**,** 504-7.

BLITVICH, B. J., SCANLON, D., SHIELL, B. J., MACKENZIE, J. S., PHAM, K. & HALL, R. A. 2001. Determination of the intramolecular disulfide bond arrangement and biochemical identification of the glycosylation sites of the nonstructural protein NS1 of Murray Valley encephalitis virus. *J Gen Virol,* 82**,** 2251-6.

BROWN, W. C., AKEY, D. L., KONWERSKI, J. R., TARRASCH, J. T., SKINIOTIS, G., KUHN, R. J. & SMITH, J. L. 2016. Extended surface for membrane association in Zika virus NS1 structure. *Nat Struct Mol Biol,* 23**,** 865-7.

BROWN, W. C., DELPROPOSTO, J., RUBIN, J. R., LAMIMAN, K., CARLESS, J. & SMITH, J. L. 2011. New ligation-independent cloning vectors compatible with a high-throughput platform for parallel construct expression evaluation using baculovirus-infected insect cells. *Protein Expr Purif,* 77**,** 34-45.

BULICH, R. & AASKOV, J. G. 1992. Nuclear localization of dengue 2 virus core protein detected with monoclonal antibodies. *J Gen Virol,* 73**,** 2999-3003.

BYK, L. A. & GAMARNIK, A. V. 2016. Properties and Functions of the Dengue Virus Capsid Protein. *Annu Rev Virol,* 3**,** 263-281.

CAMPOS, G. S., BANDEIRA, A. C. & SARDI, S. I. 2015. Zika Virus Outbreak, Bahia, Brazil. *Emerg Infect Dis,* 21**,** 1885-6.

CAO, S., LI, Y., YE, J., YANG, X., CHEN, L., LIU, X. & CHEN, H. 2011. Japanese encephalitis Virus wild strain infection suppresses dendritic cells maturation and function, and causes the expansion of regulatory T cells. *Virol J,* 8: 39.

CHANG, R. Y., HSU, T. W., CHEN, Y. L., LIU, S. F., TSAI, Y. J., LIN, Y. T., CHEN, Y. S. & FAN, Y. H. 2013. Japanese encephalitis virus non-coding RNA inhibits activation of interferon by blocking nuclear translocation of interferon regulatory factor 3. *Vet Microbiol,* 166**,** 11-21.

CHAO, D. Y., GALULA, J. U., SHEN, W. F., DAVIS, B. S. & CHANG, G. J. 2015. Nonstructural protein 1-specific immunoglobulin m and g antibody capture enzyme-linked immunosorbent assays in diagnosis of flaviviral infections in humans. *J Clin Microbiol,* 53**,** 557-66.

CHEN, V. B., ARENDALL, W. B., 3RD, HEADD, J. J., KEEDY, D. A., IMMORMINO, R. M., KAPRAL, G. J., MURRAY, L. W., RICHARDSON, J. S. & RICHARDSON, D. C. 2010. MolProbity: all-atom structure validation for macromolecular crystallography. *Acta Crystallogr D Biol Crystallogr,* 66**,** 12-21.

CHENG, H. J., LIN, C. F., LEI, H. Y., LIU, H. S., YEH, T. M., LUO, Y. H. & LIN, Y. S. 2009. Proteomic analysis of endothelial cell autoantigens recognized by anti-dengue virus nonstructural protein 1 antibodies. *Exp Biol Med (Maywood),* 234**,** 63-73.

CHUNG, K. M., LISZEWSKI, M. K., NYBAKKEN, G., DAVIS, A. E., TOWNSEND, R. R., FREMONT, D. H., ATKINSON, J. P. & DIAMOND, M. S. 2006a. West Nile virus nonstructural protein NS1 inhibits complement activation by binding the regulatory protein factor H. *Proc Natl Acad Sci U S A,* 103**,** 19111-6.

CHUNG, K. M., NYBAKKEN, G. E., THOMPSON, B. S., ENGLE, M. J., MARRI, A., FREMONT, D. H. & DIAMOND, M. S. 2006b. Antibodies against West Nile Virus nonstructural protein NS1 prevent lethal infection through Fc gamma receptor-dependent and -independent mechanisms. *J Virol,* 80**,** 1340-51.

CLYDE, K., BARRERA, J. & HARRIS, E. 2008. The capsid-coding region hairpin element (cHP) is a critical determinant of dengue virus and West Nile virus RNA synthesis. *Virology,* 379**,** 314-23.

COWTAN, K. 2006. The Buccaneer software for automated model building. 1. Tracing protein chains. *Acta Crystallogr D Biol Crystallogr,* 62**,** 1002-11.

CROOKS, A. J., LEE, J. M., EASTERBROOK, L. M., TIMOFEEV, A. V. & STEPHENSON, J. R. 1994. The NS1 protein of tick-borne encephalitis virus forms multimeric species upon secretion from the host cell. *J Gen Virol,* 75 (Pt 12)**,** 3453-60.

CUGOLA, F. R., FERNANDES, I. R., RUSSO, F. B., FREITAS, B. C., DIAS, J. L., GUIMARAES, K. P., BENAZZATO, C., ALMEIDA, N., PIGNATARI, G. C., ROMERO, S., POLONIO, C. M., CUNHA, I., FREITAS, C. L., BRANDAO, W. N., ROSSATO, C., ANDRADE, D. G., FARIA DDE, P., GARCEZ, A. T., BUCHPIGEL, C. A., BRACONI, C. T., MENDES, E., SALL, A. A., ZANOTTO, P. M., PERON, J. P., MUOTRI, A. R. & BELTRAO-BRAGA, P. C. 2016. The Brazilian Zika virus strain causes birth defects in experimental models. *Nature,* 534**,** 267-71.

CUTLER, P. 2004. *Protein Purification Protocols. [electronic book]*, [s.l.] : Humana Press, 2004.

DAS, S. & BASU, A. 2008. Japanese encephalitis virus infects neural progenitor cells and decreases their proliferation. *J Neurochem,* 106**,** 1624-36.

DOKLAND, T., WALSH, M., MACKENZIE, J. M., KHROMYKH, A. A., EE, K. H. & WANG, S. 2004. West Nile virus core protein; tetramer structure and ribbon formation. *Structure,* 12**,** 1157-63.

DOLINSKY, T. J., CZODROWSKI, P., LI, H., NIELSEN, J. E., JENSEN, J. H., KLEBE, G. & BAKER, N. A. 2007. PDB2PQR: expanding and upgrading automated preparation of biomolecular structures for molecular simulations. *Nucleic Acids Research,* 35**,** W522-W525.

DUFFY, M. R., CHEN, T. H., HANCOCK, W. T., POWERS, A. M., KOOL, J. L., LANCIOTTI, R. S., PRETRICK, M., MARFEL, M., HOLZBAUER, S., DUBRAY, C., GUILLAUMOT, L., GRIGGS, A., BEL, M., LAMBERT, A. J., LAVEN, J., KOSOY, O., PANELLA, A., BIGGERSTAFF, B. J., FISCHER, M. & HAYES, E. B. 2009. Zika virus outbreak on Yap Island, Federated States of Micronesia. *N Engl J Med,* 360**,** 2536-43.

EDELING, M. A., DIAMOND, M. S. & FREMONT, D. H. 2014. Structural basis of Flavivirus NS1 assembly and antibody recognition. *Proc Natl Acad Sci U S A,* 111**,** 4285-90.

ELODIE FOULQUIER, C. G., ZOHRA OUARAY AND MARIE-PAULE LEFRANC 2017. Amino acids. *IMGT Education*.

EMSLEY, P., LOHKAMP, B., SCOTT, W. G. & COWTAN, K. 2010. Features and development of Coot. *Acta Crystallogr D Biol Crystallogr,* 66**,** 486-501.

ESPADA-MURAO, L. A. & MORITA, K. 2011. Delayed cytosolic exposure of Japanese encephalitis virus double-stranded RNA impedes interferon activation and enhances viral dissemination in porcine cells. *J Virol,* 85**,** 6736-49.

EVANS, P. R. 2011. An introduction to data reduction: space-group determination, scaling and intensity statistics. *Acta Crystallogr D Biol Crystallogr,* 67**,** 282-92.

FALGOUT, B., CHANOCK, R. & LAI, C. J. 1989. Proper processing of dengue virus nonstructural glycoprotein NS1 requires the N-terminal hydrophobic signal sequence and the downstream nonstructural protein NS2a. *J Virol,* 63**,** 1852-60.

FAN, J., LIU, Y. & YUAN, Z. 2014. Critical role of Dengue Virus NS1 protein in viral replication. *Virol Sin,* 29**,** 162-9.

FLAMAND, M., CHEVALIER, M., HENCHAL, E., GIRARD, M. & DEUBEL, V. 1995. Purification and renaturation of Japanese encephalitis virus nonstructural glycoprotein NS1 overproduced by insect cells. *Protein Expr Purif,* 6**,** 519-27.

FLAMAND, M., DEUBEL, V. & GIRARD, M. 1992. Expression and secretion of Japanese encephalitis virus nonstructural protein NS1 by insect cells using a recombinant baculovirus. *Virology,* 191**,** 826-36.

FRANKE, D. & SVERGUN, D. I. 2009. DAMMIF, a program for rapid ab-initio shape determination in small-angle scattering. *J Appl Crystallogr,* 42**,** 342-346.

GHOSH, D. & BASU, A. 2009. Japanese encephalitis-a pathological and clinical perspective. *PLoS Negl Trop Dis,* 3**,** e437.

GOMES, A. R., BYREGOWDA, S. M., VEEREGOWDA, B. M. & VINAYAGAMURTHY, B. 2016. An Overview of Heterologous Expression Host Systems for the Production of Recombinant Proteins. *Advances in Animal and Veterinary Sciences* 4**,** 346-356.

GUZMAN, M. G. & HARRIS, E. 2015. Dengue. *Lancet,* 385**,** 453-65.

HAN, N., ADAMS, J., CHEN, P., GUO, Z. Y., ZHONG, X. F., FANG, W., LI, N., WEN, L., TAO, X. Y., YUAN, Z. M. & RAYNER, S. 2014. Comparison of genotypes I and III in Japanese encephalitis virus reveals distinct differences in their genetic and host diversity. *J Virol,* 88**,** 11469-79.

HOLM, L. & ROSENSTROM, P. 2010. Dali server: conservation mapping in 3D. *Nucleic Acids Res,* 38**,** W545-9.

HOLMES, E. C. & TWIDDY, S. S. 2003. The origin, emergence and evolutionary genetics of dengue virus. *Infect Genet Evol,* 3**,** 19-28.

HUANG, G. K. L., TIO, S. Y., CALY, L., NICHOLSON, S., THEVARAJAN, I., PAPADAKIS, G., CATTON, M., TONG, S. Y. C. & DRUCE, J. 2017. Prolonged Detection of Japanese Encephalitis Virus in Urine and Whole Blood in a Returned Short-term Traveler. *Open Forum Infectious Diseases,* 4**,** ofx203-ofx203.

IMPOINVIL, D. E., BAYLIS, M. & SOLOMON, T. 2013. Japanese encephalitis: on the One Health agenda. *Curr Top Microbiol Immunol,* 365**,** 205-47.

ISHIKAWA, T. & KONISHI, E. 2015. Potential chemotherapeutic targets for Japanese encephalitis: current status of antiviral drug development and future challenges. *Expert Opin Ther Targets,* 19**,** 1379-95.

IVANYI-NAGY, R., LAVERGNE, J. P., GABUS, C., FICHEUX, D. & DARLIX, J. L. 2008. RNA chaperoning and intrinsic disorder in the core proteins of Flaviviridae. *Nucleic Acids Res,* 36**,** 712-25.

JACOBS, M. G., ROBINSON, P. J., BLETCHLY, C., MACKENZIE, J. M. & YOUNG, P. R. 2000. Dengue virus nonstructural protein 1 is expressed in a glycosyl-phosphatidylinositol-linked form that is capable of signal transduction. *FASEB J,* 14**,** 1603-10.

KABSCH, W. 2010. Xds. *Acta Crystallogr D Biol Crystallogr,* 66**,** 125-32.

KALITA, J., MISRA, U. K., MANI, V. E. & BHOI, S. K. 2016. Can we differentiate between herpes simplex encephalitis and Japanese encephalitis? *J Neurol Sci,* 366**,** 110-115.

KARPLUS, P. A. & DIEDERICHS, K. 2012. Linking crystallographic model and data quality. *Science,* 336**,** 1030-3.

KHROMYKH, A. A. & WESTAWAY, E. G. 1996. RNA binding properties of core protein of the flavivirus Kunjin. *Arch Virol,* 141**,** 685-99.

KIERMAYR, S., KOFLER, R. M., MANDL, C. W., MESSNER, P. & HEINZ, F. X. 2004. Isolation of capsid protein dimers from the tick-borne encephalitis flavivirus and in vitro assembly of capsid-like particles. *J Virol,* 78**,** 8078-84.

KNOPE, K. E., DOGGETT, S. L., KURUCZ, N., JOHANSEN, C. A., NICHOLSON, J., FELDMAN, R., SLY, A., HOBBY, M., EL SAADI, D., MULLER, M., JANSEN, C. C., MUZARI, O. M., NATIONAL, A. & MALARIA ADVISORY, C. 2014. Arboviral diseases and malaria in Australia, 2011-12: annual report of the National Arbovirus and Malaria Advisory Committee. *Commun Dis Intell Q Rep,* 38**,** E122-42.

KOFLER, R. M., HEINZ, F. X. & MANDL, C. W. 2002. Capsid protein C of tick-borne encephalitis virus tolerates large internal deletions and is a favorable target for attenuation of virulence. *J Virol,* 76**,** 3534-43.

KONAREV, P. V., VOLKOV, V. V., SOKOLOVA, A. V., KOCH, M. H. J. & SVERGUN, D. I. 2003. PRIMUS: a Windows PC-based system for small-angle scattering data analysis. *Journal of Applied Crystallography,* 36**,** 1277-1282.

KONNO, J., ENDO, K., AGATSUMA, H. & ISHIDA, N. 1966. Cyclic outbreaks of Japanese encephalitis among pigs and humans. *Am J Epidemiol,* 84**,** 292-300.

KRISHNA, V. D., RANGAPPA, M. & SATCHIDANANDAM, V. 2009. Virus-specific cytolytic antibodies to nonstructural protein 1 of Japanese encephalitis virus effect reduction of virus output from infected cells. *J Virol,* 83**,** 4766-77.

KRISSINEL, E. & HENRICK, K. 2007. Inference of macromolecular assemblies from crystalline state. *J Mol Biol,* 372**,** 774-97.

KUHN, R. J., ZHANG, W., ROSSMANN, M. G., PLETNEV, S. V., CORVER, J., LENCHES, E., JONES, C. T., MUKHOPADHYAY, S., CHIPMAN, P. R., STRAUSS, E. G., BAKER, T. S. & STRAUSS, J. H. 2002. Structure of dengue virus: implications for flavivirus organization, maturation, and fusion. *Cell,* 108**,** 717-25.

KUMAR, R., TRIPATHI, P., BARANWAL, M., SINGH, S., TRIPATHI, S. & BANERJEE, G. 2009. Randomized, controlled trial of oral ribavirin for Japanese encephalitis in children in Uttar Pradesh, India. *Clin Infect Dis,* 48**,** 400-6.

LANDAU, M., MAYROSE, I., ROSENBERG, Y., GLASER, F., MARTZ, E., PUPKO, T. & BEN-TAL, N. 2005. ConSurf 2005: the projection of evolutionary conservation scores of residues on protein structures. *Nucleic Acids Res,* 33**,** W299-302.

LARKIN, M. A., BLACKSHIELDS, G., BROWN, N. P., CHENNA, R., MCGETTIGAN, P. A., MCWILLIAM, H., VALENTIN, F., WALLACE, I. M., WILM, A., LOPEZ, R., THOMPSON, J. D., GIBSON, T. J. & HIGGINS, D. G. 2007. Clustal W and Clustal X version 2.0. *Bioinformatics,* 23**,** 2947-8.

LATIF, B., KANNAN KUTTY, M., MUSLIM, A., HUSSAINI, J., OMAR, E., HEO, C. C., ROSSLE, N. F., ABDULLAH, S., KAMARUDIN, M. A. & ZULKARNAIN, M. A. 2015. Light microscopy and molecular identification of Sarcocystis spp. in meat producing animals in Selangor, Malaysia. *Trop Biomed,* 32**,** 444-52.

LI, C., XU, D., YE, Q., HONG, S., JIANG, Y., LIU, X., ZHANG, N., SHI, L., QIN, C. F. & XU, Z. 2016. Zika Virus Disrupts Neural Progenitor Development and Leads to Microcephaly in Mice. *Cell Stem Cell,* 19**,** 120-126.

LI, M. H., FU, S. H., CHEN, W. X., WANG, H. Y., GUO, Y. H., LIU, Q. Y., LI, Y. X., LUO, H. M., DA, W., DUO JI, D. Z., YE, X. M. & LIANG, G. D. 2011. Genotype v Japanese encephalitis virus is emerging. *PLoS Negl Trop Dis,* 5**,** e1231.

LI, Y. Z., COUNOR, D., LU, P., LIANG, G. D., VU, T. Q., PHAN, T. N., HUYNH, T. K., SUN, G., GRANDADAM, M., BUTRAPET, S., LAVERGNE, J. P., FLAMAND, M., YU, Y. X., SOLOMON, T., BUCHY, P. & DEUBEL, V. 2012. A specific and sensitive antigen capture assay for NS1 protein quantitation in Japanese encephalitis virus infection. *J Virol Methods,* 179**,** 8-16.

LIMON-FLORES, A. Y., PEREZ-TAPIA, M., ESTRADA-GARCIA, I., VAUGHAN, G., ESCOBAR-GUTIERREZ, A., CALDERON-AMADOR, J., HERRERA-RODRIGUEZ, S. E., BRIZUELA-GARCIA, A., HERAS-CHAVARRIA, M., FLORES-LANGARICA, A., CEDILLO-BARRON, L. & FLORES-ROMO, L. 2005. Dengue virus inoculation to human skin explants: an effective approach to assess in situ the early infection and the effects on cutaneous dendritic cells. *Int J Exp Pathol,* 86**,** 323-34.

LIN, C. W., CHENG, C. W., YANG, T. C., LI, S. W., CHENG, M. H., WAN, L., LIN, Y. J., LAI, C. H., LIN, W. Y. & KAO, M. C. 2008. Interferon antagonist function of Japanese encephalitis virus NS4A and its interaction with DEAD-box RNA helicase DDX42. *Virus Res,* 137**,** 49-55.

LIN, R. J., CHANG, B. L., YU, H. P., LIAO, C. L. & LIN, Y. L. 2006. Blocking of interferon-induced Jak-Stat signaling by Japanese encephalitis virus NS5 through a protein tyrosine phosphatase-mediated mechanism. *J Virol,* 80**,** 5908-18.

LINDENBACH, B. D. & RICE, C. M. 1997. trans-Complementation of yellow fever virus NS1 reveals a role in early RNA replication. *J Virol,* 71**,** 9608-17.

LINDENBACH, B. D. & RICE, C. M. 1999. Genetic interaction of flavivirus nonstructural proteins NS1 and NS4A as a determinant of replicase function. *J Virol,* 73**,** 4611-21.

LIU, C. C., ZHANG, Y. N., LI, Z. Y., HOU, J. X., ZHOU, J., KAN, L., ZHOU, B. & CHEN, P. Y. 2017. Rab5 and Rab11 Are Required for Clathrin-Dependent Endocytosis of Japanese Encephalitis Virus in BHK-21 Cells. *J Virol,* 91.

LIU, Z. Y., LI, X. F., JIANG, T., DENG, Y. Q., ZHAO, H., WANG, H. J., YE, Q., ZHU, S. Y., QIU, Y., ZHOU, X., QIN, E. D. & QIN, C. F. 2013. Novel cis-acting element within the capsid-coding region enhances flavivirus viral-RNA replication by regulating genome cyclization. *J Virol,* 87**,** 6804-18.

MA, L., JONES, C. T., GROESCH, T. D., KUHN, R. J. & POST, C. B. 2004. Solution structure of dengue virus capsid protein reveals another fold. *Proc Natl Acad Sci U S A,* 101**,** 3414-9.

MACKENZIE, J. M., JONES, M. K. & YOUNG, P. R. 1996. Immunolocalization of the dengue virus nonstructural glycoprotein NS1 suggests a role in viral RNA replication. *Virology,* 220**,** 232-40.

MANDL, C. W., HEINZ, F. X., STOCKL, E. & KUNZ, C. 1989. Genome sequence of tick-borne encephalitis virus (Western subtype) and comparative analysis of nonstructural proteins with other flaviviruses. *Virology,* 173**,** 291-301.

MARRS, C., OLSON, G., SAADE, G., HANKINS, G., WEN, T., PATEL, J. & WEAVER, S. 2016. Zika Virus and Pregnancy: A Review of the Literature and Clinical Considerations. *Am J Perinatol,* 33**,** 625-39.

MASON, P. W. 1989. Maturation of Japanese encephalitis virus glycoproteins produced by infected mammalian and mosquito cells. *Virology,* 169**,** 354-64.

MATHUR, A., BHARADWAJ, M., KULSHRESHTHA, R., RAWAT, S., JAIN, A. & CHATURVEDI, U. C. 1988. Immunopathological study of spleen during Japanese encephalitis virus infection in mice. *Br J Exp Pathol,* 69**,** 423-32.

MCWILLIAM, H., LI, W., ULUDAG, M., SQUIZZATO, S., PARK, Y. M., BUSO, N., COWLEY, A. P. & LOPEZ, R. 2013. Analysis Tool Web Services from the EMBL-EBI. *Nucleic Acids Res,* 41**,** W597-600.

MELIAN, E. B., HALL-MENDELIN, S., DU, F., OWENS, N., BOSCO-LAUTH, A. M., NAGASAKI, T., RUDD, S., BRAULT, A. C., BOWEN, R. A., HALL, R. A., VAN DEN HURK, A. F. & KHROMYKH, A. A. 2014. Programmed ribosomal frameshift alters expression of west nile virus genes and facilitates virus replication in birds and mosquitoes. *PLoS Pathog,* 10**,** e1004447.

MELIAN, E. B., HINZMAN, E., NAGASAKI, T., FIRTH, A. E., WILLS, N. M., NOUWENS, A. S., BLITVICH, B. J., LEUNG, J., FUNK, A., ATKINS, J. F., HALL, R. & KHROMYKH, A. A. 2010. NS1' of flaviviruses in the Japanese encephalitis virus serogroup is a product of ribosomal frameshifting and plays a role in viral neuroinvasiveness. *J Virol,* 84**,** 1641-7.

MORRISON, C. R. & SCHOLLE, F. 2014. Abrogation of TLR3 inhibition by discrete amino acid changes in the C-terminal half of the West Nile virus NS1 protein. *Virology,* 456-457**,** 96-107.

MUKHOPADHYAY, S., KUHN, R. J. & ROSSMANN, M. G. 2005. A structural perspective of the flavivirus life cycle. *Nat Rev Microbiol,* 3**,** 13-22.

MULLER, D. A. & YOUNG, P. R. 2013. The flavivirus NS1 protein: molecular and structural biology, immunology, role in pathogenesis and application as a diagnostic biomarker. *Antiviral Res,* 98**,** 192-208.

MURSHUDOV, G. N., SKUBAK, P., LEBEDEV, A. A., PANNU, N. S., STEINER, R. A., NICHOLLS, R. A., WINN, M. D., LONG, F. & VAGIN, A. A. 2011. REFMAC5 for the refinement of macromolecular crystal structures. *Acta Crystallogr D Biol Crystallogr,* 67**,** 355-67.

MUSSO, D., NILLES, E. J. & CAO-LORMEAU, V. M. 2014. Rapid spread of emerging Zika virus in the Pacific area. *Clin Microbiol Infect,* 20**,** 595-6.

MUYLAERT, I. R., CHAMBERS, T. J., GALLER, R. & RICE, C. M. 1996. Mutagenesis of the N-linked glycosylation sites of the yellow fever virus NS1 protein: effects on virus replication and mouse neurovirulence. *Virology,* 222**,** 159-68.

NAIN, M., ABDIN, M. Z., KALIA, M. & VRATI, S. 2016. Japanese encephalitis virus invasion of cell: allies and alleys. *Rev Med Virol,* 26**,** 129-41.

NATTLESHIP, J. E. 2012. Structural Biology of Glycoproteins.

NOISAKRAN, S., DECHTAWEWAT, T., AVIRUTNAN, P., KINOSHITA, T., SIRIPANYAPHINYO, U., PUTTIKHUNT, C., KASINRERK, W., MALASIT, P. & SITTISOMBUT, N. 2008. Association of dengue virus NS1 protein with lipid rafts. *J Gen Virol,* 89**,** 2492-500.

NOISAKRAN, S., DECHTAWEWAT, T., RINKAEWKAN, P., PUTTIKHUNT, C., KANJANAHALUETHAI, A., KASINRERK, W., SITTISOMBUT, N. & MALASIT, P. 2007. Characterization of dengue virus NS1 stably expressed in 293T cell lines. *J Virol Methods,* 142**,** 67-80.

NOTHAFT, H. & SZYMANSKI, C. M. 2010. Protein glycosylation in bacteria: sweeter than ever. *Nat Rev Microbiol,* 8**,** 765-78.

OLIVEIRA, E. R., MOHANA-BORGES, R., DE ALENCASTRO, R. B. & HORTA, B. A. 2017. The flavivirus capsid protein: Structure, function and perspectives towards drug design. *Virus Res,* 227**,** 115-123.

OLSON, J. G., KSIAZEK, T. G., SUHANDIMAN & TRIWIBOWO 1981. Zika virus, a cause of fever in Central Java, Indonesia. *Trans R Soc Trop Med Hyg,* 75**,** 389-93.

OMASITS, U., AHRENS, C. H., MULLER, S. & WOLLSCHEID, B. 2014. Protter: interactive protein feature visualization and integration with experimental proteomic data. *Bioinformatics,* 30**,** 884-6.

OTWINOWSKI, Z. & MINOR, W. 1997. Processing of X-ray diffraction data collected in oscillation mode. *Methods Enzymol,* 276**,** 307-26.

PERERA, R. & KUHN, R. J. 2008. Structural proteomics of dengue virus. *Curr Opin Microbiol,* 11**,** 369-77.

PERRY, J. 2016. Preparation of soluble/insoluble protein from cells. *solubilitiy studies* [Online]. Available: https://www.embl.de/pepcore/pepcore_services/protein_purification/extraction_clarification/solubility_studies/index.html [Accessed 2016-12-12].

PETOUKHOV, M. V., FRANKE, D., SHKUMATOV, A. V., TRIA, G., KIKHNEY, A. G., GAJDA, M., GORBA, C., MERTENS, H. D. T., KONAREV, P. V. & SVERGUN, D. I. 2012. New developments in the ATSAS program package for small-angle scattering data analysis. *Journal of Applied Crystallography,* 45**,** 342-350.

PETTERSEN, E. F., GODDARD, T. D., HUANG, C. C., COUCH, G. S., GREENBLATT, D. M., MENG, E. C. & FERRIN, T. E. 2004. UCSF Chimera--a visualization system for exploratory research and analysis. *J Comput Chem,* 25**,** 1605-12.

PFEFFER, M. & DOBLER, G. 2010. Emergence of zoonotic arboviruses by animal trade and migration. *Parasites & Vectors,* 3: 35.

PONG, W. L., HUANG, Z. S., TEOH, P. G., WANG, C. C. & WU, H. N. 2011. RNA binding property and RNA chaperone activity of dengue virus core protein and other viral RNA-interacting proteins. *FEBS Lett,* 585**,** 2575-81.

POONSIRI, T., WRIGHT, G. S. A., DIAMOND, M. S., TURTLE, L., SOLOMON, T. & ANTONYUK, S. V. 2018. Structural study of the C-terminal domain of non-structural protein 1 from Japanese encephalitis virus. *J Virol*.

RHODES, G. 2006. *Crystallography made crystal clear. [electronic book] : a guide for users of macromolecular models*, Amsterdam : Elsevier/Academic Press, 2006. 3rd ed.

SALAM, K. A. & AKIMITSU, N. 2013. Hepatitis C virus NS3 inhibitors: current and future perspectives. *Biomed Res Int,* 2013.

SAMSA, M. M., MONDOTTE, J. A., CARAMELO, J. J. & GAMARNIK, A. V. 2012. Uncoupling cis-Acting RNA elements from coding sequences revealed a requirement of the N-terminal region of dengue virus capsid protein in virus particle formation. *J Virol,* 86**,** 1046-58.

SAMSA, M. M., MONDOTTE, J. A., IGLESIAS, N. G., ASSUNCAO-MIRANDA, I., BARBOSA-LIMA, G., DA POIAN, A. T., BOZZA, P. T. & GAMARNIK, A. V. 2009. Dengue virus capsid protein usurps lipid droplets for viral particle formation. *PLoS Pathog,* 5**,** e1000632.

SAPKAL, G. N., WAIRAGKAR, N. S., AYACHIT, V. M., BONDRE, V. P. & GORE, M. M. 2007. Detection and isolation of Japanese encephalitis virus from blood clots

collected during the acute phase of infection. *Am J Trop Med Hyg,* 77**,** 1139-45.
SCHAGGER, H. 2006. Tricine-SDS-PAGE. *Nat. Protocols,* 1**,** 16-22.
SCHELD, W. M., HAMMER, S. M. & HUGHES, J. M. 2008. *Emerging Infections, 8,* Washington, UNITED STATES, ASM Press.
SCHLESINGER, J. J., BRANDRISS, M. W., CROPP, C. B. & MONATH, T. P. 1986. Protection against yellow fever in monkeys by immunization with yellow fever virus nonstructural protein NS1. *J Virol,* 60**,** 1153-5.
SCHLESINGER, J. J., FOLTZER, M. & CHAPMAN, S. 1993. The Fc portion of antibody to yellow fever virus NS1 is a determinant of protection against YF encephalitis in mice. *Virology,* 192**,** 132-41.
SCHLICK, P., KOFLER, R. M., SCHITTL, B., TAUCHER, C., NAGY, E., MEINKE, A. & MANDL, C. W. 2010. Characterization of West Nile virus live vaccine candidates attenuated by capsid deletion mutations. *Vaccine,* 28**,** 5903-9.
SCHLICK, P., TAUCHER, C., SCHITTL, B., TRAN, J. L., KOFLER, R. M., SCHUELER, W., VON GABAIN, A., MEINKE, A. & MANDL, C. W. 2009. Helices alpha2 and alpha3 of West Nile virus capsid protein are dispensable for assembly of infectious virions. *J Virol,* 83**,** 5581-91.
SCHNEIDMAN-DUHOVNY, D., HAMMEL, M., TAINER, J. A. & SALI, A. 2013. Accurate SAXS profile computation and its assessment by contrast variation experiments. *Biophys J,* 105**,** 962-74.
SCHUH, A. J., WARD, M. J., LEIGH BROWN, A. J. & BARRETT, A. D. T. 2013. Phylogeography of Japanese Encephalitis Virus: Genotype Is Associated with Climate. *PLoS Negl Trop Dis,* 7**,** e2411.
SCOTT, W. R. P., HÜNENBERGER, P. H., TIRONI, I. G., MARK, A. E., BILLETER, S. R., FENNEN, J., TORDA, A. E., HUBER, T., KRÜGER, P. & VAN GUNSTEREN, W. F. 1999. The GROMOS Biomolecular Simulation Program Package. *The Journal of Physical Chemistry A,* 103**,** 3596-3607.
SHIRYAEV, SERGEY A., KOZLOV, IGOR A., RATNIKOV, BORIS I., SMITH, JEFFREY W., LEBL, M. & STRONGIN, ALEX Y. 2007. Cleavage preference distinguishes the two-component NS2B–NS3 serine proteinases of Dengue and West Nile viruses. *Biochemical Journal,* 401**,** 743-752.
SIEVERS, F., WILM, A., DINEEN, D., GIBSON, T. J., KARPLUS, K., LI, W., LOPEZ, R., MCWILLIAM, H., REMMERT, M., SODING, J., THOMPSON, J. D. & HIGGINS, D. G. 2011. Fast, scalable generation of high-quality protein multiple sequence alignments using Clustal Omega. *Mol Syst Biol,* 7: 539.
SIM, S., AW, P. P., WILM, A., TEOH, G., HUE, K. D., NGUYEN, N. M., NAGARAJAN, N., SIMMONS, C. P. & HIBBERD, M. L. 2015. Tracking Dengue Virus Intra-host Genetic Diversity during Human-to-Mosquito Transmission. *PLoS Negl Trop Dis,* 9**,** e0004052.
SINGH, S. M. & PANDA, A. K. 2005. Solubilization and refolding of bacterial inclusion body proteins. *J Biosci Bioeng,* 99**,** 303-10.
SKALA, W., GOETTIG, P. & BRANDSTETTER, H. 2013. Do-it-yourself histidine-tagged bovine enterokinase: a handy member of the protein engineer's toolbox. *J Biotechnol,* 168**,** 421-5.

SLAVOV, S. N., OTAGUIRI, K. K., KASHIMA, S. & COVAS, D. T. 2016. Overview of Zika virus (ZIKV) infection in regards to the Brazilian epidemic. *Braz J Med Biol Res,* 49**,** e5420.

SMITH, J. L., AKEY, D. L., BROWN, W. C. & KUHN, R. J. 2015. Vaccine compositions and uses thereof WO 2015095735 A2. Google Patents.

SMITHBURN, K. C. 1952. Neutralizing antibodies against certain recently isolated viruses in the sera of human beings residing in East Africa. *J Immunol,* 69**,** 223-34.

SOLOMON, T. 2006. Control of Japanese encephalitis--within our grasp? *N Engl J Med,* 355**,** 869-71.

SOLOMON, T., DUNG, N. M., WILLS, B., KNEEN, R., GAINSBOROUGH, M., DIET, T. V., THUY, T. T., LOAN, H. T., KHANH, V. C., VAUGHN, D. W., WHITE, N. J. & FARRAR, J. J. 2003a. Interferon alfa-2a in Japanese encephalitis: a randomised double-blind placebo-controlled trial. *Lancet,* 361**,** 821-6.

SOLOMON, T., NI, H., BEASLEY, D. W. C., EKKELENKAMP, M., CARDOSA, M. J. & BARRETT, A. D. T. 2003b. Origin and Evolution of Japanese Encephalitis Virus in Southeast Asia. *Journal of Virology,* 77**,** 3091-3098.

SOLOMON, T., THAO, L. T., DUNG, N. M., KNEEN, R., HUNG, N. T., NISALAK, A., VAUGHN, D. W., FARRAR, J., HIEN, T. T., WHITE, N. J. & CARDOSA, M. J. 1998. Rapid diagnosis of Japanese encephalitis by using an immunoglobulin M dot enzyme immunoassay. *J Clin Microbiol,* 36**,** 2030-4.

SOMNUKE, P., HAUHART, R. E., ATKINSON, J. P., DIAMOND, M. S. & AVIRUTNAN, P. 2011. N-linked glycosylation of dengue virus NS1 protein modulates secretion, cell-surface expression, hexamer stability, and interactions with human complement. *Virology,* 413**,** 253-64.

SONG, H., QI, J., HAYWOOD, J., SHI, Y. & GAO, G. F. 2016. Zika virus NS1 structure reveals diversity of electrostatic surfaces among flaviviruses. *Nat Struct Mol Biol,* 23**,** 456-8.

SORENSEN, H. P. & MORTENSEN, K. K. 2005. Soluble expression of recombinant proteins in the cytoplasm of Escherichia coli. *Microb Cell Fact,* 4: 1.

STUDIER, F. W. 2014. Stable expression clones and auto-induction for protein production in E. coli. *Methods Mol Biol,* 1091**,** 17-32.

SU, H. L., LIAO, C. L. & LIN, Y. L. 2002. Japanese encephalitis virus infection initiates endoplasmic reticulum stress and an unfolded protein response. *J Virol,* 76**,** 4162-71.

SUCHARIT, S., SURATHIN, K. & SHRESTHA, S. R. 1989. Vectors of Japanese encephalitis virus (JEV): species complexes of the vectors. *Southeast Asian J Trop Med Public Health,* 20**,** 611-21.

SURI, N. K. & BANERJEE, K. 1995. Growth and cytopathic effect of Japanese encephalitis virus in astrocyte-enriched cell cultures from neonatal mouse brains. *Acta Virol,* 39**,** 143-8.

SVERGUN, D. I. 1999. Restoring low resolution structure of biological macromolecules from solution scattering using simulated annealing. *Biophysical Journal,* 76**,** 2879-2886.

SWARUP, V., GHOSH, J., DUSEJA, R., GHOSH, S. & BASU, A. 2007. Japanese encephalitis virus infection decrease endogenous IL-10 production:

correlation with microglial activation and neuronal death. *Neurosci Lett,* 420**,** 144-9.

TAKAMATSU, Y., OKAMOTO, K., DINH, D. T., YU, F., HAYASAKA, D., UCHIDA, L., NABESHIMA, T., BUERANO, C. C. & MORITA, K. 2014. NS1' protein expression facilitates production of Japanese encephalitis virus in avian cells and embryonated chicken eggs. *J Gen Virol,* 95**,** 373-83.

TAKHAMPUNYA, R., KIM, H. C., TIPPAYACHAI, B., KENGLUECHA, A., KLEIN, T. A., LEE, W. J., GRIECO, J. & EVANS, B. P. 2011. Emergence of Japanese encephalitis virus genotype V in the Republic of Korea. *Virol J,* 8: 449.

THONGTAN, T., CHEEPSUNTHORN, P., CHAIWORAKUL, V., RATTANARUNGSAN, C., WIKAN, N. & SMITH, D. R. 2010. Highly permissive infection of microglial cells by Japanese encephalitis virus: a possible role as a viral reservoir. *Microbes Infect,* 12**,** 37-45.

TSAI, W. Y., YOUN, H. H., BRITES, C., TSAI, J. J., TYSON, J., PEDROSO, C., DREXLER, J. F., STONE, M., SIMMONS, G., BUSCH, M. P., LANTERI, M., STRAMER, S. L., BALMASEDA, A., HARRIS, E. & WANG, W. K. 2017. Distinguishing Secondary Dengue Virus Infection From Zika Virus Infection With Previous Dengue by a Combination of 3 Simple Serological Tests. *Clin Infect Dis,* 65**,** 1829-1836.

TSENG, H. N., LEE, C. C., WONG, M. L., CHEN, S. O. & LIU, J. J. 2007. DNA-binding property of recombinant capsid protein of Japanese encephalitis virus. *Virus Genes,* 35**,** 483-8.

TU, Y. C., YU, C. Y., LIANG, J. J., LIN, E., LIAO, C. L. & LIN, Y. L. 2012. Blocking double-stranded RNA-activated protein kinase PKR by Japanese encephalitis virus nonstructural protein 2A. *J Virol,* 86**,** 10347-58.

TWIDDY, S. S., HOLMES, E. C. & RAMBAUT, A. 2003. Inferring the rate and time-scale of dengue virus evolution. *Mol Biol Evol,* 20**,** 122-9.

UNNI, S. K., RUZEK, D., CHHATBAR, C., MISHRA, R., JOHRI, M. K. & SINGH, S. K. 2011. Japanese encephalitis virus: from genome to infectome. *Microbes Infect,* 13**,** 312-21.

VAGIN, A. & TEPLYAKOV, A. 2010. Molecular replacement with MOLREP. *Acta Crystallogr D Biol Crystallogr,* 66**,** 22-5.

VAN DEN HURK, A. F., MONTGOMERY, B. L., NORTHILL, J. A., SMITH, I. L., ZBOROWSKI, P., RITCHIE, S. A., MACKENZIE, J. S. & SMITH, G. A. 2006. Short report: the first isolation of Japanese encephalitis virus from mosquitoes collected from mainland Australia. *Am J Trop Med Hyg,* 75**,** 21-5.

VAN DEN HURK, A. F., RITCHIE, S. A. & MACKENZIE, J. S. 2009. Ecology and geographical expansion of Japanese encephalitis virus. *Annu Rev Entomol,* 54**,** 17-35.

VILLORDO, S. M. & GAMARNIK, A. V. 2009. Genome cyclization as strategy for flavivirus RNA replication. *Virus Res,* 139**,** 230-9.

VOLKOV, V. V. & SVERGUN, D. I. 2003. Uniqueness of ab initio shape determination in small-angle scattering. *Journal of Applied Crystallography,* 36**,** 860-864.

WALLIS, T. P., HUANG, C. Y., NIMKAR, S. B., YOUNG, P. R. & GORMAN, J. J. 2004. Determination of the disulfide bond arrangement of dengue virus NS1 protein. *J Biol Chem,* 279**,** 20729-41.

WANG, S. H., SYU, W. J., HUANG, K. J., LEI, H. Y., YAO, C. W., KING, C. C. & HU, S. T. 2002. Intracellular localization and determination of a nuclear localization signal of the core protein of dengue virus. *J Gen Virol,* 83**,** 3093-102.

WANG, X., LI, S. H., ZHU, L., NIAN, Q. G., YUAN, S., GAO, Q., HU, Z., YE, Q., LI, X. F., XIE, D. Y., SHAW, N., WANG, J., WALTER, T. S., HUISKONEN, J. T., FRY, E. E., QIN, C. F., STUART, D. I. & RAO, Z. 2017. Near-atomic structure of Japanese encephalitis virus reveals critical determinants of virulence and stability. *Nat Commun,* 8**,** 14.

WASH, R. & SORIA, C. D. 2015. True Blood: dengue virus evolution. *Nat Rev Microbiol,* 13**,** 662.

WEAVER, S. C., COSTA, F., GARCIA-BLANCO, M. A., KO, A. I., RIBEIRO, G. S., SAADE, G., SHI, P. Y. & VASILAKIS, N. 2016. Zika virus: History, emergence, biology, and prospects for control. *Antiviral Res,* 130**,** 69-80.

WELSCH, S., MILLER, S., ROMERO-BREY, I., MERZ, A., BLECK, C. K., WALTHER, P., FULLER, S. D., ANTONY, C., KRIJNSE-LOCKER, J. & BARTENSCHLAGER, R. 2009. Composition and three-dimensional architecture of the dengue virus replication and assembly sites. *Cell Host Microbe,* 5**,** 365-75.

WHO 2016. Dengue vaccine: WHO position paper - July 2016. *Wkly Epidemiol Rec,* 91**,** 349-64.

WILSON, J. R., DE SESSIONS, P. F., LEON, M. A. & SCHOLLE, F. 2008. West Nile virus nonstructural protein 1 inhibits TLR3 signal transduction. *J Virol,* 82**,** 8262-71.

WINKLER, G., RANDOLPH, V. B., CLEAVES, G. R., RYAN, T. E. & STOLLAR, V. 1988. Evidence that the mature form of the flavivirus nonstructural protein NS1 is a dimer. *Virology,* 162**,** 187-96.

WINTER, G. 2010. xia2: an expert system for macromolecular crystallography data reduction. *Journal of Applied Crystallography,* 43**,** 186-190.

WU, S. J., GROUARD-VOGEL, G., SUN, W., MASCOLA, J. R., BRACHTEL, E., PUTVATANA, R., LOUDER, M. K., FILGUEIRA, L., MAROVICH, M. A., WONG, H. K., BLAUVELT, A., MURPHY, G. S., ROBB, M. L., INNES, B. L., BIRX, D. L., HAYES, C. G. & FRANKEL, S. S. 2000. Human skin Langerhans cells are targets of dengue virus infection. *Nat Med,* 6**,** 816-20.

XU, X., SONG, H., QI, J., LIU, Y., WANG, H., SU, C., SHI, Y. & GAO, G. F. 2016. Contribution of intertwined loop to membrane association revealed by Zika virus full-length NS1 structure. *EMBO J,* 35**,** 2170–2178.

YAMAUCHI, Y. & HELENIUS, A. 2013. Virus entry at a glance. *Journal of Cell Science,* 126**,** 1289-1295.

YE, Q., LI, X. F., ZHAO, H., LI, S. H., DENG, Y. Q., CAO, R. Y., SONG, K. Y., WANG, H. J., HUA, R. H., YU, Y. X., ZHOU, X., QIN, E. D. & QIN, C. F. 2012. A single nucleotide mutation in NS2A of Japanese encephalitis-live vaccine virus (SA14-14-2) ablates NS1' formation and contributes to attenuation. *J Gen Virol,* 93**,** 1959-64.

YOUN, S., AMBROSE, R. L., MACKENZIE, J. M. & DIAMOND, M. S. 2013. Non-structural protein-1 is required for West Nile virus replication complex formation and viral RNA synthesis. *Virol J,* 10**,** 339.

YOUN, S., LI, T., MCCUNE, B. T., EDELING, M. A., FREMONT, D. H., CRISTEA, I. M. & DIAMOND, M. S. 2012. Evidence for a genetic and physical interaction

between nonstructural proteins NS1 and NS4B that modulates replication of West Nile virus. *J Virol,* 86**,** 7360-71.

YOUNG, L. B., MELIAN, E. B. & KHROMYKH, A. A. 2013. NS1' colocalizes with NS1 and can substitute for NS1 in West Nile virus replication. *J Virol,* 87**,** 9384-90.

YUN, S. I. & LEE, Y. M. 2014. Japanese encephalitis: the virus and vaccines. *Hum Vaccin Immunother,* 10**,** 263-79.

ZHANG, X., GE, P., YU, X., BRANNAN, J. M., BI, G., ZHANG, Q., SCHEIN, S. & ZHOU, Z. H. 2013. Cryo-EM structure of the mature dengue virus at 3.5-A resolution. *Nat Struct Mol Biol,* 20**,** 105-10.

ZHANG, Y., CORVER, J., CHIPMAN, P. R., ZHANG, W., PLETNEV, S. V., SEDLAK, D., BAKER, T. S., STRAUSS, J. H., KUHN, R. J. & ROSSMANN, M. G. 2003. Structures of immature flavivirus particles. *EMBO J,* 22**,** 2604-13.

ZHANG, Y. & SKOLNICK, J. 2004. Tertiary structure predictions on a comprehensive benchmark of medium to large size proteins. *Biophys J,* 87**,** 2647-55.

ZHAO, H., WANG, Y. G., DENG, Y. Q., SONG, K. Y., LI, X. F., WANG, H. J., ZHU, C. M., QIN, E. D. & QIN, C. F. 2013. Japanese encephalitis virus RNA not detected in urine. *Clin Infect Dis,* 57**,** 157-8.

Appendices

Appendix 1 Publication

AMERICAN SOCIETY FOR MICROBIOLOGY | Journal of Virology®

STRUCTURE AND ASSEMBLY

Check for updates

Structural Study of the C-Terminal Domain of Nonstructural Protein 1 from Japanese Encephalitis Virus

Thanalai Poonsiri,[a,c] Gareth S. A. Wright,[a] Michael S. Diamond,[d,e,f] Lance Turtle,[b,c] Tom Solomon,[c] Svetlana V. Antonyuk[a]

[a]Molecular Biophysics Group, Institute of Integrative Biology, Faculty of Health and Life Sciences, University of Liverpool, Liverpool, United Kingdom

[b]Centre for Global Vaccine Research, Institute of Infection and Global Health, University of Liverpool, Liverpool, United Kingdom

[c]Health Protection Research Unit on Emerging and Zoonotic Infections, Institute of Infection and Global Health, University of Liverpool, Liverpool, United Kingdom

[d]Department of Medicine, Washington University School of Medicine in St. Louis, St. Louis, Missouri, USA

[e]Department of Molecular Microbiology, Washington University School of Medicine in St. Louis, St. Louis, Missouri, USA

[f]Department of Pathology & Immunology, Washington University School of Medicine in St. Louis, St. Louis, Missouri, USA

ABSTRACT Japanese encephalitis virus (JEV) is a mosquito-transmitted flavivirus that is closely related to other emerging viral pathogens, including dengue virus (DENV), West Nile virus (WNV), and Zika virus (ZIKV). JEV infection can result in meningitis and encephalitis, which in severe cases cause permanent brain damage and death. JEV occurs predominantly in rural areas throughout Southeast Asia, the Pacific Islands, and the Far East, causing around 68,000 cases of infection worldwide each year. In this report, we present a 2.1-Å-resolution crystal structure of the C-terminal β-ladder domain of JEV nonstructural protein 1 (NS1-C). The surface charge distribution of JEV NS1-C is similar to those of WNV and ZIKV but differs from that of DENV. Analysis of the JEV NS1-C structure, with *in silico* molecular dynamics simulation and experimental solution small-angle X-ray scattering, indicates extensive loop flexibility on the exterior of the protein. This, together with the surface charge distribution, indicates that flexibility influences the protein-protein interactions that govern pathogenicity. These factors also affect the interaction of NS1 with the 22NS1 monoclonal antibody, which is protective against West Nile virus infection. Liposome and heparin binding assays indicate that only the N-terminal region of NS1 mediates interaction with membranes and that sulfate binding sites common to NS1 structures are not glycosaminoglycan binding interfaces. This report highlights several differences between flavivirus NS1 proteins and contributes to our understanding of their structure-pathogenic function relationships.

IMPORTANCE JEV is a major cause of viral encephalitis in Asia. Despite extensive vaccination, epidemics still occur. Nonstructural protein 1 (NS1) plays a role in viral replication, and, because it is secreted, it can exhibit a wide range of interactions with host proteins. NS1 sequence and protein folds are conserved within the *Flavivirus* genus, but variations in NS1 protein-protein interactions among viruses likely contribute to differences in pathogenesis. Here, we compared characteristics of the C-terminal β-ladder domain of NS1 between flaviviruses, including surface charge, loop flexibility, epitope cross-reactivity, membrane adherence, and glycosaminoglycan binding. These structural features are central to NS1 functionality and may provide insight into the development of diagnostic tests and therapeutics.

Received 26 October 2017 **Accepted** 12 January 2018

Accepted manuscript posted online 17 January 2018

Citation Poonsiri T, Wright GSA, Diamond MS, Turtle L, Solomon T, Antonyuk SV. 2018. Structural study of the C-terminal domain of nonstructural protein 1 from Japanese encephalitis virus. J Virol 92:e01868-17. https://doi.org/10.1128/JVI.01868-17.

Editor Julie K. Pfeiffer, University of Texas Southwestern Medical Center

Address correspondence to Svetlana V. Antonyuk, S.Antonyuk@liverpool.ac.uk.

KEYWORDS Japanese encephalitis virus, protein crystallography, neutralizing antibodies, nonstructural protein 1, protein structure-function

Japanese encephalitis virus (JEV) is a positive-sense single-stranded RNA virus with a 10.9-kb genome which is translated into a polyprotein consisting of three structural proteins (capsid, membrane, and envelope protein [E]) and seven nonstructural proteins, including nonstructural protein 1 (NS1), NS2A, NS2B, NS3, NS4, NS4B, and NS5. Flavivirus NS1 is a multifunctional glycoprotein that has drawn attention because of its importance in viral replication, immune modulation, and immune evasion. Mutagenesis and transcomplementation assays have established that flavivirus NS1 is essential for RNA replication (1–5) and colocalizes with the replication complex (2). Transcomplementation suppressor mutagenesis studies indicate that yellow fever virus (YFV) NS1 interacts with NS4A (6) and that West Nile virus (WNV) NS1 interacts with NS4B (7). WNV NS1 forms a physical complex with NS4B based on coimmunoprecipitation experiments (7). NS1 has been described as a complement-fixing antigen (8–11), and dengue virus (DENV) NS1 binds to complement pathway components C1s, C4, and C4b (12, 13), whereas WNV NS1 can also interact with factor H (14) for protection of infected cells from complement-dependent clearance. NS1 may also interfere with the double-stranded RNA (dsRNA) sensor Toll-like receptor 3 (TLR-3) (15) to escape host pathogen recognition receptor detection. DENV NS1 can induce inflammatory cytokine production, endothelial cell permeability, and changes to the glycocalyx (16), possibly through interactions with TLR-4; all of those activities appear to contribute to the development of severe dengue (17, 18). Although direct interactions between JEV NS1 and TLR-4 have not been evaluated, it may play a role in JEV pathogenesis, because deletion of the TLR-4 gene enhances resistance to JEV (19).

A signal sequence at the C terminus of E protein translocates NS1 to the endoplasmic reticulum (ER), where it undergoes cleavage and posttranslational modification (20). There are two characterized forms of NS1: a membrane-associated dimer (~49 kDa per monomer), found on ER surface and the plasma membrane, and a secreted hexamer (52 to 55 kDa per monomer) (20). The masses of the two NS1 forms are different due to differential glycosylation. Structures of full-length NS1 proteins of WNV (21, 22), DENV (22), and ZIKV (23, 24) have been reported. Most NS1 proteins contain six conserved disulfide bonds. The NS1 forms share a conserved N-linked glycosylation site at Asn 207. YFV, DENV, WNV, and JEV share a second glycosylation site at Asn 130, and most of the JEV serogroup NS1 proteins have a third glycosylation site at Asn175 linked to high-mannose carbohydrate, but it is not present in JEV NS1 itself (20, 25–27). The NS1 monomer of WNV, DENV, and ZIKV contains 3 domains: the β-roll (amino acid residues 1 to 29), wing (38 to 151), and β-ladder (181 to 352) domains (22–24). NS1 forms a homodimer by extending the β-ladder domain and connecting at the β-roll domain, forming a cross-shaped protein. One face of the dimer comprises the protruding β-roll domain and part of the wing domain. The hydrophobic surface of the β-roll and wing domains may mediate the interaction with the cell membrane (22) via a number of amino acid residues identified in ZIKV, including amino acid residues 28, 115, 118, 123, and 160 to 163 (23, 24). The opposite side is composed of loops linking the surface β-strands of the ladder domain. This region is a potential host protein interacting surface due to its hydrophilicity. Three NS1 dimers can assemble to form a hexameric pore, which can act as a lipid depot (22, 28). DENV NS1 expression on the infected cell surface may occur via a glycosylphosphatidylinositol (GPI) anchor, for which a hydrophobic carboxy-terminal GPI addition signal peptide at the N terminus of NS2A is required (29–31). Soluble NS1 also binds to uninfected cell membranes via glycosaminoglycans (GAGs), primarily heparan sulfate and chondroitin sulfate E (32).

Secreted NS1 is used as a diagnostic marker for flavivirus infection, as it is found in the blood at early stages (33, 34). Alternatively, detection of anti-NS1 IgM and IgG can be used (34, 35). Immunization of NS1 in mice or passive transfer of anti-NS1 antibodies (Abs) can confer protective effects against flavivirus challenge (34, 36–39). However,

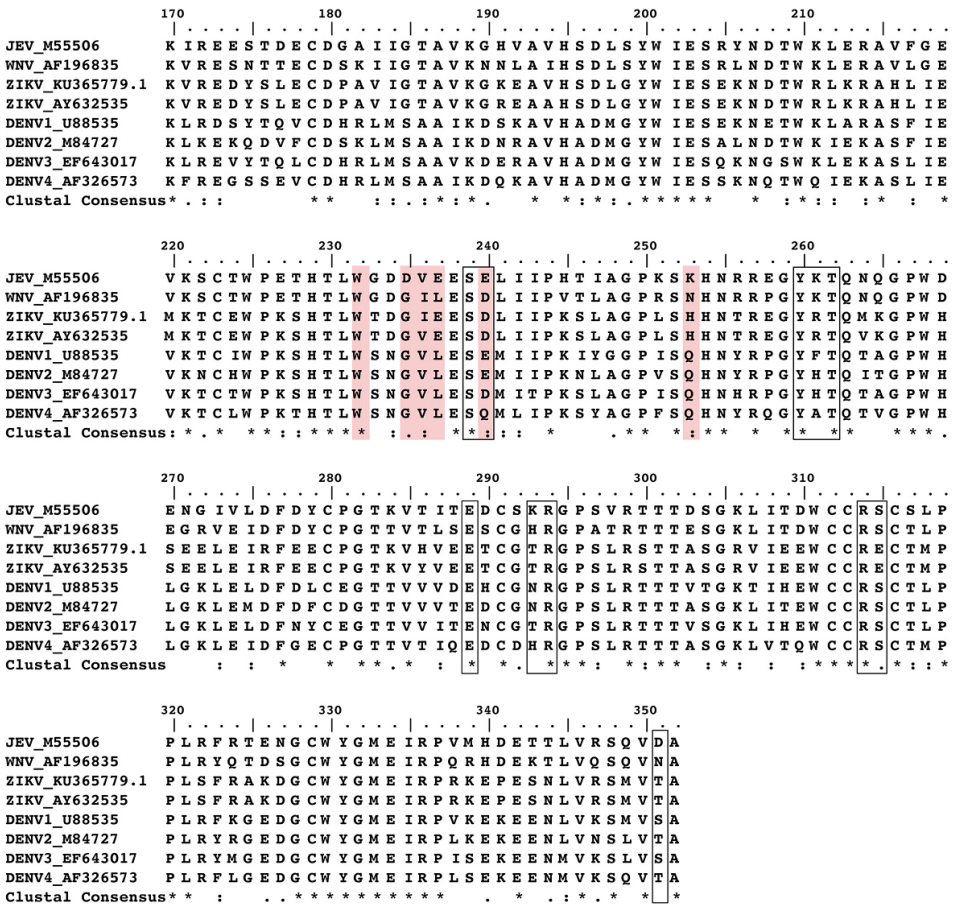


FIG 1 Sequence alignment of full-length flavivirus NS1 produced from Clustal W (1). An asterisk indicates a fully conserved residue. A colon indicates conservation between groups of strongly similar properties. A period indicates conservation between groups of weakly similar properties. The following amino acid sequences were used for X-ray structure studies: DENV1 U88535 for PDB ID 4OIG; DENV2 M84727 for PDB ID 4O6B; WNV 196835 for PDB ID 4O6C and PDB ID 4OIE; ZIKV KU365779 for PDB ID 5IY3; and ZIKV AY632535 for PDB ID 5K6K and PDB ID 5GS6. The 22NS1 light-chain epitopes are highlighted in red, and heavy-chain epitopes are in black rectangles.

some anti-DENV NS1 antibodies reportedly are autoreactive and bind to host extracellular matrix components, platelets, and endothelial cells (8, 20), which may have pathogenic consequences. Flavivirus NS1 transferred with blood meal was found to enhance viral infection in mosquitoes by downregulating mosquito midgut immune genes (40).

Most of our knowledge of JEV NS1 has been inferred from studies of DENV and WNV NS1. Although the protein sequences are highly conserved (Fig. 1) and the DENV, WNV, and ZIKV NS1 structures display the same protein fold, there are important differences. For example, polyclonal antibodies raised against DENV NS1 in mice were shown to

cross-react with proteins on epithelial cells: ATPase, protein disulfide isomerase, vimentin, and heat shock protein 60. The cross-reactive epitope was mapped to amino acid residues 311 to 330 on DENV NS1 (41) (Fig. 1). Although JEV NS1 shares these conserved epitopes, antibodies against JEV NS1 did not react to any of these host cell targets (41). NS1 alone was shown to cause endothelial leakage in DENV, but this was not detected in WNV, consistent with the non-vascular-leakage phenotype of WNV disease (42). Similarly to WNV, other encephalitic flaviviruses, including JEV, may vary in their NS1-endothelium interactions. As another example, WNV NS1 binds the alternative complement pathway regulator factor H, whereas JEV NS1 does not (8).

NS1′ is an extended form of NS1 with 52 extra amino acids from the NS2A N terminus, generated by a −1 ribosomal frameshift (43). It is specific to the JE serogroup of flaviviruses. NS1′ was found in dimeric form (monomer molecular mass of around 58 kDa), has been detected in both cell lysate and culture media (44, 45), and has been suggested to play a role in neuroinvasiveness; selectively abolishing NS1′ production reduces WNV mortality in mice (43, 46). NS1′ colocalized with viral RNA replication complex and can substitute for NS1 in cells (45). However, there is a discrepancy between the results of *in vitro* and those of *in vivo* studies. WNV NS1′ provides an advantage only in *in vivo* studies (47). There is also variation with respect to NS1′ involvement in replication among different viruses. Whereas WNV NS1′ does not contribute to viral replication *in vitro*, JEV NS1′ mutants have less infectivity in a cell model (47, 48). Therefore, the role of NS1′ in JEV life cycle and pathogenesis remains unclear.

Here, we report the crystal structure of the C-terminal domain (amino acids 172 to 352) of JEV NS1 and compare it with published DENV, WNV, and ZIKV NS1 structures. Our findings reveal diversity in protein surface charges. Furthermore, the solution conformation of the protein was examined by small-angle X-ray scattering (SAXS) and molecular dynamics (MD) simulations along with analysis of cell membrane association. Importantly, we define a cross-reactive epitope on NS1 using an antibody that shows protective activity against WNV infection. Our study results show the common and contrasting features of flavivirus NS1 structures and contribute to our knowledge of the molecular basis of multiple NS1 functions.

RESULTS

Structure of C-terminal domain of JEV NS1 and NS1′. The crystal structure of the C-terminal region of JEV NS1 determined at 2.1-Å resolution is similar to all previously solved flavivirus NS1 structures with respect to fold characteristics (Fig. 2a). The electron density is visible for residues 177 to 352, whereas that of the first 5 residues at the N terminus is not visible. The monomer consists of 10 β-strands on one side and 4 helices and unstructured loops on the other side. Between the β-strands in each pair are β-turns and short loops, apart from β4 and β5, which are separated by a long unstructured loop (residues 218 to 273) (Fig. 2a and b). The protein contains four conserved disulfide bonds (C179—C229, C280—C329, C291—C312, and C313—C316) and hydrogen bonds between β-strands and loops. JEV NS1-C forms 20 β-strands oriented in a head-to-head arrangement in the dimer, as do ZIKV, WNV, and DENV NS1, with a dimer length of 9.65 nm at its widest point (Fig. 2c). The dimer interface is created by 21 residues from each monomer, with an average distance of 2.9 Å (Tables 1 and 2). Eight of these interface residues are conserved among flavivirus NS1 proteins (Table 1, scores 7 to 9). The dimer is connected by 12 hydrogen bonds. Comparing the hydrogen bonding networks at the dimer interface of the C-terminal domains of ZIKV (PDB identifier [ID] 5IY3), WNV (PDB ID 4OIE), and DENV NS1 (PDB ID 4OIG), there are 6 common residues with the same bond arrangement: Thr (JEV, ZIKV, WNV)/Ala (DENV) 186, Val (JEV, ZIKV, WNV)/Ile (DENV) 188, Thr (JEV, WNV)/Ser (ZIKV, DENV) 228, His 254, and Thr 230 to Trp 232 (Tables 1, 2, and 3) (Fig. 3). In addition, we solved the structure of JEV NS1′-C, which is distinguished from NS1 by the presence of an extra 52 amino acids at the C terminus of the protein, to 2.6-Å resolution. The structure revealed the same protein fold and dimer orientation. However, it showed only 2 extra amino acids

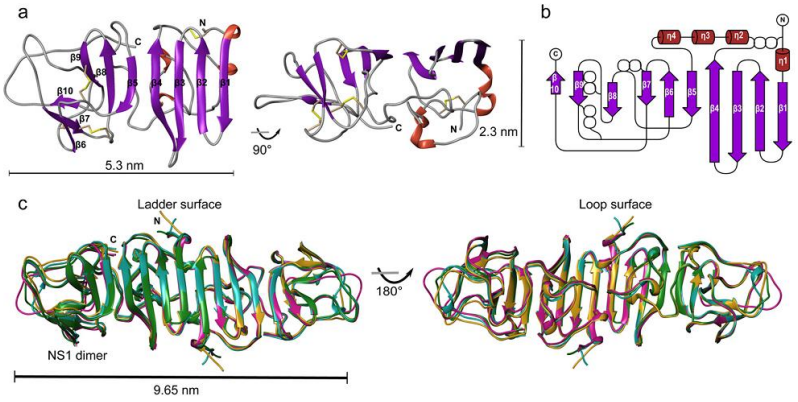


FIG 2 The C-terminal domain structure of JEV NS1. (a) Ribbon model of JEV NS1-C monomer. One side is built of 10 β-strands, and the opposite side consists of the nonstructured loops. Disulfide bonds are shown in yellow. (b) Topology diagram of JEV NS1-C. Four disulfide bonds are indicated as white spheres. "β" represents the β-sheet, and "η" represents the 310 helix. (c) Superimposed ribbon diagram of NS1-C of JEV (magenta), ZIKV (PDB ID 5IY3, blue), WNV (PDB ID 4OIE, green), and DENV1 (PDB ID 4OIG, gold).

in comparison with the C-terminal domain of JEV NS1 (0.348-Å Cα root mean square deviation [RMSD]) (data not shown). The C terminus is disordered, so the electron density is not visible.

Solution model of JEV NS1-C dimer. The dimeric nature of JEV NS1-C was confirmed by SAXS studies performed on the protein in solution. The SAXS profiles calculated from the monomer and dimer of JEV NS1-C crystal structure were compared with the JEV NS1-C experimental SAXS data (Fig. 4a). A monomer of JEV NS1-C yielded a poor fit to the experimental data with χ of 14.11, whereas a dimer provided an improved fit with χ of 4.02. A radius of gyration of 27.02 Å was obtained from Guinier

TABLE 1 JEV NS1 C-terminal dimer interfacing residues

No.	Residue	ASA (Å^2)[a]	BSA (Å^2)[b]	ΔG (kcal/mol)[c]	Conservation[d]
1	Gly181	23.56	6.81	0.03	3
2	Ala182	91.29	45.68	0.30	1
3	Ile184	22.47	16.36	−0.18	5
4	Gly185	40.33	16.05	0.26	7
5	Thr186	37.46	36.73	−0.22	7
6	Ala187	63.58	21.49	0.34	9
7	Val188	64.98	63.40	0.34	6
8	Lys189	181.59	9.98	0.16	8
9	Gly190	63.17	54.83	0.30	5
10	His191	110.70	33.66	0.74	1
11	Trp210	60.01	29.42	0.08	5
12	Glu227	104.51	54.79	0.51	5
13	Thr228	120.12	94.74	0.66	6
14	His229	54.08	52.04	0.90	9
15	Thr230	24.89	21.26	−0.20	9
16	Leu231	48.09	48.09	0.77	8
17	Trp232	95.68	59.18	0.38	5
18	Gly233	39.95	30.60	−0.02	4
19	Asp234	91.67	54.93	0.26	6
20	Asp235	128.72	0.58	−0.01	1
21	His254	13.56	10.75	0.73	8

[a]ASA, accessible surface area (determined by assembly analysis in the program PISA).
[b]BSA, buried surface area (determined by assembly analysis in the program PISA).
[c]ΔG, solvation energy effect (determined by assembly analysis in the program PISA).
[d]Amino acid conservation values represent Consurf scores (9, conserved; 1, variable).

TABLE 2 Hydrogen bonds between JEV NS1 C-terminal dimer interfacing residues

No.	Structure 1	Distance (Å)[a]	Structure 2
1	Gly190 [N]	2.93	Ile184 [O]
2	Val188 [N]	2.86	Thr186 [O]
3	Thr186 [N]	2.89	Val188 [O]
4	His229 [NE2]	2.83	Gly190 [O]
5	His254 [NE2]	2.94	Thr228 [O]
6	Trp232 [N]	2.96	Thr230 [O]
7	Ile184 [O]	2.93	Gly190 [N]
8	Thr186 [O]	2.86	Val188 [N]
9	Val188 [O]	2.89	Thr186 [N]
10	Gly190 [O]	2.83	His229 [NE2]
11	Thr228 [O]	2.94	His254 [NE2]
12	Thr230 [O]	2.96	Trp232 [N]

[a]Data represent results of assembly analysis in the program PISA.

analysis, and that value is consistent with the value extracted from the pair distribution function, 27.08 Å. The pair distribution function of JEV NS1-C shows characteristics of a lengthy ovoid particle with a maximum intraparticle distance (D_{max}) of 94.1 Å, similar to the widest point of JEV NS1-C dimer crystal structure (96.5 Å) (Fig. 2c and 4b). The calculated molecular mass was 45.5 kDa, corresponding to the dimeric form of C terminus NS1. An averaged *ab initio* model was generated at 30-Å resolution with good similarity agreement (normal spatial discrepancy [NSD] = 0.513 ± 0.016) and was compared with the JEV NS1-C dimer crystal structure (Fig. 4c). The structures were well matched, although there was an extra region of mass near the dimer interface in the SAXS model (labeled M; Fig. 4c). This feature also is seen in the SAXS model of WNV, suggesting that the NS1 crystal structures of JEV and WNV may not fully represent the structure of the protein in solution (21). Analysis of the crystallographic atomic mean square displacements or B-factors in our JEV NS1-C crystal structure indicates that surface regions of loop 218 to 272, particularly subloop 235 to 237, had high conformational freedom within the crystal lattice (Fig. 4d and e). A 40-ns all-atom molecular dynamics (MD) simulation of the JEV NS1-C dimer at 37°C confirmed that the movement of this loop was unrestrained in both monomers (Fig. 4d and e). We hypothesized that the apparent extra region of mass observed in the JEV NS1-C and WNV SAXS structures could be accounted for by the dynamic nature of loop 218 to 272 and the resulting expansion of volume in the solution structures. To model JEV NS1-C behavior in solution more accurately, we created a pool of possible structures with various conformations of loop 218 to 272 and compared them with our SAXS data. Using this approach, we improved the fit to the experimental SAXS data from χ of 4.02 to χ of 1.48 (Fig. 4a).

Comparison of JEV NS1-C with other flavivirus NS1-C structures. JEV NS1-C has the same fold characteristics as ZIKV (2.2-Å resolution), WNV (2.6-Å resolution), and DENV (3-Å resolution) NS1-C, and superposition gives Cα RMSD values closest to that determined for WNV NS1 (1.162 Å for ZIKV, 0.959 Å for WNV, and 1.333 Å for DENV) (Fig. 2c). The structural superimposition showed low positional conservation only at the N terminus, C terminus, and β-turns. The electrostatic surface potential maps of known NS1-C domains (ZIKV, WNV, and DENV) showed symmetric patterns consistent with homodimers. On the β-ladder surface, all displayed a neutral charge in the central regions flanked by negatively charged regions (Fig. 5). This negatively charged region is small in DENV and larger in WNV and can be seen to expand diagonally from the top left to bottom right in JEV and ZIKV in the figure. Adjacent to it, toward the ends, are small positively charged pockets that are seen clearly only in JEV and ZIKV, and the tips of all NS1-C proteins have mixed charges. The loop surface is more variable than the ladder surface. DENV has a distinct positively charged central region, whereas JEV and WNV have a negative charge in their central area. ZIKV is different, as the middle region displays both positive and negative charges. The adjacent area has positively charged pockets in all NS1 structures (Fig. 5 and Table 4). Three pockets are found in WNV and

TABLE 3 Residues forming a hydrogen bond at a dimer interface compared with existing flavivirus NS1[a]

	ZIKV residue			WNV residue			DENV residue	
JEV residue	5K6K	5GS6	5IY3	4O6D	4O6C	4OIE	4O6B	4OIG
	Asp1	His1		Asp1	Asp1			
	Val2	**Val2**		**Thr2**	**Thr2**		**Ser2**	
	Cys4	**Cys4**		**Cys4**	**Cys4**		**Cys4**	
	Ser5	**Ser5**						
	Val6	**Val6**		**Ile6**	**Val6**		**Ile6**	
	Phe8							
	Ser9							
				Arg10	Arg10			
	Lys11							
	Glu12			Glu12	Glu12			
				Leu13				
	Arg14	**Arg14**		**Arg14**	**Arg14**		**Lys14**	
	Thr17	**Thr17**		**Ser17**	**Ser17**		**Ser17**	
	Val19	**Val19**		**Val19**	**Val19**		**Ile19**	
	Phe20	Phe20		Phe20	Phe20			
	Ile21	**Val21**		**Ile21**	**Ile21**		**Ile21**	
	Tyr22	Tyr22						
	Asn23	**Asn23**		**Asn23**	**Asn23**		**Asp23**	
	Asp24	Asp24		Asp24	Asp24			
	Arg31			Arg31	Arg31			
	Tyr32			Tyr32	Tyr32			
	Asp157	Asp157						
							Tyr158	
				Phe160				
	Thr165			Thr165	Thr165			
				Ser181	Ser181			
				Lys182	Lys182			Arg182
Ile184	Ile184	Ile184	Ile184					
								Ser185
Thr186	**Thr186**	**Thr186**	**Thr186**	**Thr186**	**Thr186**	**Thr186**	**Ala186**	**Ala186**
Val188	**Val188**	**Val188**	**Val188**	**Val188**	**Val188**	**Val188**	**Ile188**	**Ile188**
	Lys189			Lys189	Lys189			Lys189
Gly190	Gly190	Gly190	Gly190					Asp190
		Lys191		Asn191	Asn191			
	192Glu	Glu192						
	193Ala							
	Glu203			Glu203				
	Lys227	Lys227	Lys227					
Thr228	**Ser228**	**Ser228**	**Ser228**	**Thr228**	**Thr228**	**Thr228**	**Ser228**	**Ser228**
His229	His229	His229	His229					
Thr230	**Thr230**	**Thr230**	**Thr230**	**Thr230**	**Thr230**	**Thr230**	**Thr230**	**Thr230**
Trp232	**Trp232**	**Trp232**	**Trp232**	**Trp232**	**Trp232**	**Trp232**	**Trp232**	**Trp232**
	Thr233	Thr233	Thr233				Ser233	Ser233
	Asp234	Asp234	Asp234				Asn234	Asn234
His254	**His254**	**His254**	**His254**	**His254**	**His254**	**His254**	**His254**	**His254**

[a]Shared residues are indicated in boldface.

DENV, whereas ZIKV has only pockets 1 and 2 and JEV has pockets 1 and 3. The residues building the positively charged pockets are conserved in pocket 1 and partially conserved in pocket 2, pocket 3, and the front pocket on the ladder surface (Fig. 5b and Table 4).

Cell membrane interaction via determination of GAGs. Sulfate molecules were found on the surface of JEV NS1-C (Fig. 6 and Table 5) that were similar to those seen with ZIKV (PDB ID 5K6K), WNV (PDB ID 4O6C), and DENV (PDB ID 4OIG). Moreover, they were found to be distributed near the positively charged pockets. Hence, it is possible that this positively charged area is the binding site of negatively charged ligands. Sulfate-containing molecules, such as GAGs, which are involved in NS1-dependent membrane attachment (32) could interact here. To test if the interaction with GAGs occurs via sulfate binding sites at the C terminus, JEV NS1-C binding to heparin agarose beads was analyzed. However, the 20-kDa JEV NS1-C was found only in the flowthrough

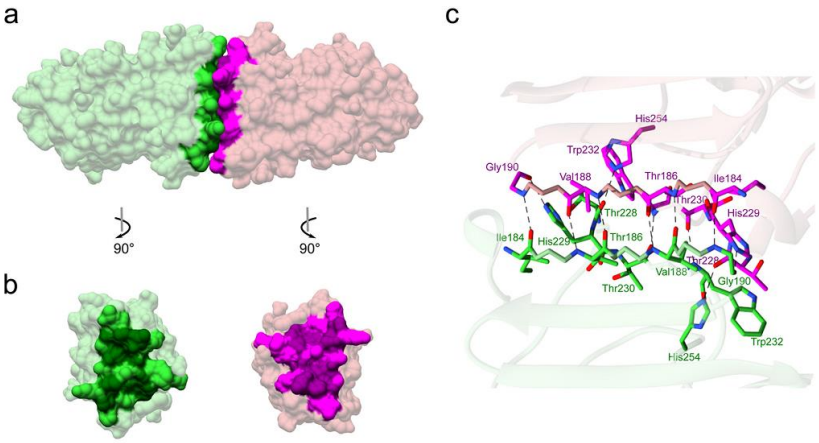


FIG 3 Dimer interface of JEV NS1-C. (a and b) The surfaces of 21 residues from one monomer involved in dimer interface are colored in lime green, and the surfaces that form hydrogen bonds are colored in dark green (a). Similarly, another monomer-interfacing surface is in magenta and surfaces forming hydrogen bonds are in dark magenta (b). (c) Residues involved in hydrogen formation at the dimer interface are highlighted in lime green and magenta as described for panels a and b. Hydrogen bonds are indicated by dashed lines.

and wash fractions (Fig. 7a), indicating that it did not interact efficiently with heparin. The interaction of heparan sulfate, chondroitin sulfate, and dermatan sulfate polymers with JEV NS1-C was investigated by protein thermal shift assay. No JEV NS1-C stabilizing effect was observed for any of GAG polymers tested even at a high concentration of GAG (100 μM). The absence of GAG binding was consistent with the results of the pulldown experiments.

Interactions of JEV NS1-C with lipids common to cell membranes were tested using a liposome binding assay. JEV NS1-C did not associate with liposomes at either pH 7.5 or pH 5.5 (Fig. 7b). As full-length NS1 can bind liposomes (22, 49), it appears that the NS1 C terminus is not responsible for membrane binding. We note that the hydrophobic residues of β-roll and wing domains have been suggested to play a role in membrane binding (22–24).

JEV NS1-C and JEV NS1′-C complexed with 22NS1 fragment antigen binding (Fab). Comparison of the 22NS1 antibody epitope of WNV-NS1-C with that of JEV NS1-C showed that 9 of 16 residues (Trp232, Ser239, Tyr260, Lys261, Thr262, Glu289, Arg294, Arg314, and Ser315) are conserved between the two viruses (Fig. 1) (36). Indeed, 22NS1 monoclonal antibody (MAb) cross-reacts with JEV NS1-C protein and NS1′-C protein, which was confirmed by Western blotting (Fig. 8, lower left inset) and size exclusion chromatography (SEC) (Fig. 8). JEV NS1-C and 22NS1 Fab alone eluted at retention times of 7.9 and 8.3 min, respectively. JEV NS1-C incubated with 22NS1 Fab eluted faster at a retention time of 6.9 min, corresponding to complex formation with a small amount of free 22NS1 Fab fragments left. The eluted fraction was analyzed by SDS-PAGE, and 2 peaks representing JEV NS1-C and 22NS1 (~25 kDa) were identified. This confirms that NS1 and 22NS1 MAb interact in solution (Fig. 8, lower right inset). The incubation also generated a small peak at a retention time of 6.1 min. This may represent a higher-order oligomer of JEV NS1-C, which recruits multiple 22NS1 monomers into a complex with a higher hydrodynamic radius than the 2:2 complex observed at 6.9 min. In support of this idea, we note the absence of the NS1-C species eluting at ~7.4 min in the complex chromatogram. The protein-Fab complex was also analyzed by SAXS. The complex experimental profile was compared to that of the WNV NS1-C–22NS1 complex (PDB ID 4OII) using the calculated SAXS profile (Fig. 9a and c). However,

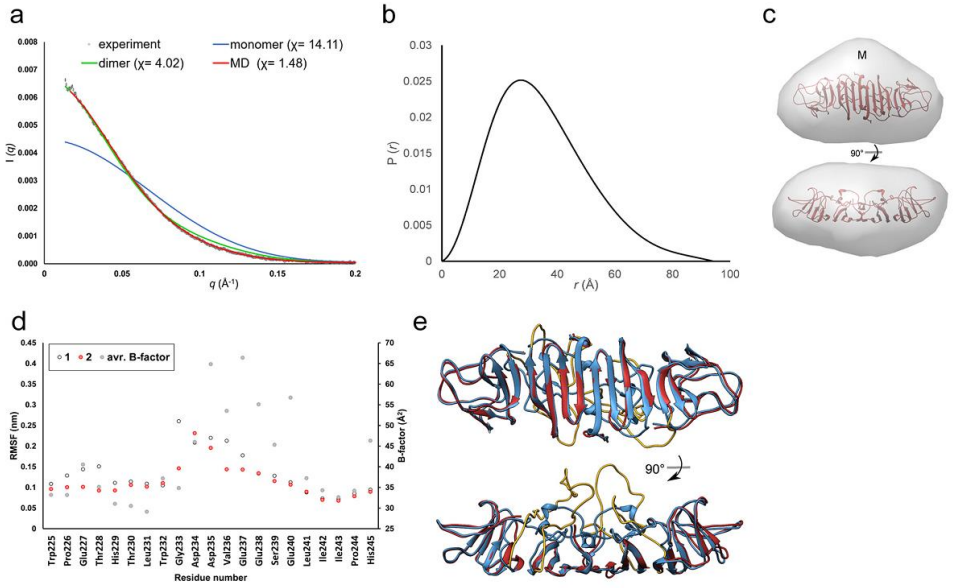


FIG 4 Solution model of JEV NS1-C dimer. (a) SAXS curve. An experimental scattering curve is shown in black scattering. Scattering profiles of JEV NS1-C monomer and dimer and the best molecular dynamic simulation structure calculated with FoXS are shown in blue, green, and red, respectively. (b) Pair distribution functions. (c) Low-resolution model of JEV NS1-C calculated from SAXS profiles docked with the crystal structure of the JEV NS1-C dimer. An extra region of mass is labeled with an "M." (d) RMSF plot of the molecular dynamic simulation at the flexible loop. RMSF values of each monomer are indicated in black and red. The average (avr.) β-factor value for each residue is indicated in gray. (e) The best molecular dynamic simulation structure (red) was superimposed with the JEV NS1-C crystal structure (blue). The flexible loop consisting of residues 218 to 272 is shown in yellow.

the complex (PDB ID 4OII) gave a poor fit to the experimental SAXS data, with a χ value of 6.82. Guinier analysis gave a radius of gyration of 52.89 $\pm$ 0.34 Å, which coincides with the 52.50-Å value extracted from the pair distribution function. The pair distribution function of the complex had multiple peaks which signify the multidomain geometric shape, with a D_{max} value of 154.9 Å (Fig. 9b). The calculated molecular mass was 149.96 kDa. An averaged *ab initio* model was generated at 30-Å resolution. The Fab part of the WNV complex (PDB ID 4OII) did not fit into the SAXS envelope and shifted from the positions in the 4OII model, whereas the WNV NS1-C dimer fit well (Fig. 9c to e), indicating flexibility of the Fab epitope in solution. We generated a pseudoatomic model of the JEV NS1-C antibody complex by replacing the WNV-NS1-C with JEV NS1-C and optimizing the position of the Fab molecules. This model, with the 2 Fab molecules shifted away from the primary location in the 4OII model, had a better fit to the SAXS data (χ of 6.82 to 3.09; Fig. 9a and c to e). Both JEV NS1-C and NS1′-C were able to cross-interact with the protective WNV 22NS1 MAb, and JEV NS1-C interacted with some flexibility.

DISCUSSION

Flavivirus NS1 proteins have generated much interest because of their multiple functions in viral replication, cell signaling, and immune evasion. Since 2014, the structures of nine NS1 proteins have been solved (21–24, 50). These proteins were expressed in bacterial or insect cell expression systems. Here, we expressed the JEV NS1 C terminus in *Escherichia coli* after the failure of several attempts to express full-length JEV NS1 in *E. coli*, insect cells, and mammalian cells. We describe the first structure of

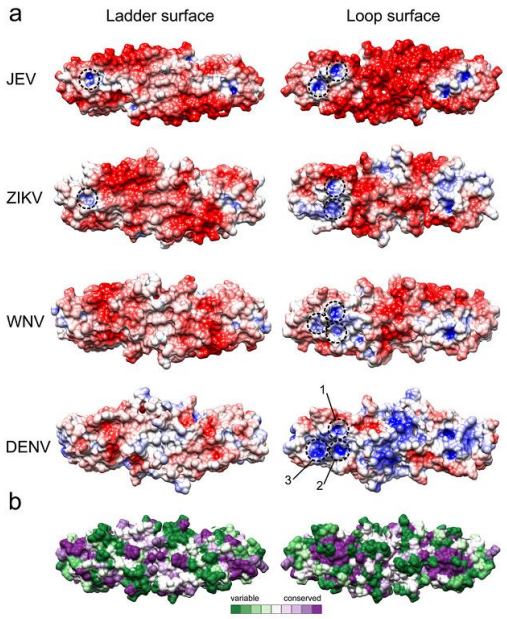


FIG 5 JEV NS1-C structure compared to other flavivirus NS1-C structures. (a) Electrostatic surface map of NS1-C structures from JEV, ZIKV, WNV, and DENV. The surface is colored according to electrostatic potential values from −5 kT/e (red) to 5 kT/e (blue). Positive potential pockets are depicted in dashed circles. (b) Surface model color coded by conservation. The most conserved residues are represented in dark magenta, and the most variable residues are represented in dark green.

the *E. coli*-expressed JEV NS1 C terminus, which, compared to other NS1 structures as well as that of JEV NS1′-C, shows a high degree of structural conservation. As expected for NS1′, the presence of the same fold could explain the similar functions of NS1 and NS1′ seen *in vitro*. However, the specific role of the extra amino acids is not yet clear, although WNV lacking the NS1′ form is less neuroinvasive (43).

The availability of WNV, DENV, and ZIKV NS1 structures has allowed us to assess similarities and differences which may be relevant to their functional behavior. All NS1 proteins are dimeric *in crystallo*, even though the recombinant protein contains only the C-terminal domain (21, 23). The molecular mass and the low-resolution model generated from SAXS data confirm the dimeric nature of the isolated C-terminal domain in solution. In contrast to previous work, which suggested that the β-roll domain is responsible for dimerization (49), we propose that 6 common residues which form hydrogen bonds at the dimer interface of all NS1 structures mediate dimer formation (49). In principle, inhibition of dimer formation by interposition of a ligand at this site could facilitate antiflavivirus drug development.

Both faces of the JEV NS1-C dimer display electrostatic surface charge diversity. However, considering the full-length flavivirus NS1 protein structure (22–24), the ladder face of the C-terminal domain is positioned under the β-roll domain (Fig. 10). The N terminus protects the central region of the ladder face from the environment. In addition, the β-roll domain is contained by a hydrophobic region that is suspected to interact with the cell membrane or to form a lipid cargo pore in NS1 hexamer, making it harder for the ladder face to interact. This model conflicts with a previous suggestion

TABLE 4 Residues forming positively charged pockets compared with existing C-NS1

JEV residue	ZIKV residue (5IY3)	WNV residue (4OIE)	DENV residue (4OIG)	Conservation[a]
Pocket 1				
Gly259	Gly259	Gly259	Gly259	9
Tyr260	Tyr260	Tyr260	Tyr260	9
Lys261	Arg261	Lys261	Phe261	1
			Ala265	1
Ser292	Gly292	Gly292	Gly292	1
Lys293	Thr293	His293	Asn293	1
Arg294	Arg294	Arg294	Arg294	9
Cys313				9
Arg314	Arg314	Arg314	Arg314	9
Ser315	Glu315	Ser315	Ser315	5
Cys316	Cys316	Cys316	Cys316	9
Glu334	Glu334	Glu334	Glu334	9
Pocket 2				
	Thr262	Thr262	Thr262	6
	Met264	Asn264	Thr264	1
	Lys265			1
	Gly295	Gly295	Gly295	9
	Pro296	Pro296	Pro296	4
		Gly332		6
	Met333	Met333	Met333	9
	Thr351	Asn351	Ser351	3
Pocket 3				
Gly295				9
Pro296				4
Ser297		Ala297	Ser297	9
Val298		Thr298	Leu298	1
Arg336		Arg336	Arg336	9
Pro337		Pro337	Pro337	9
Met339				3
			Glu340	2
		Glu342	Glu342	8
Leu345		Leu345	Leu345	6
Arg347		Gln347	Lys347	3

[a]Amino acid conservation values represent Consurf scores (9, conserved; 1, variable).

that the β-ladder might bind to the complement control protein domain (sushi domain) of complement proteins (51). In comparison, the loop face in the JEV NS1-C, with its diverse surface charges, is fully exposed compared to that of ZIKV, WNV, and DENV. In particular, DENV has the most distinct positive central area whereas the rest are negatively charged. Positively charged pockets found on the loop face of the NS1 crystal structure could mediate anionic ligand binding. Moreover, the pockets, especially pocket 1, are composed of conserved sequences and are found in all known NS1 structures. The presence or absence of each pocket in NS1 from different flaviviruses may confer upon the individual NS1 proteins the ability to interact with different target proteins or ligands in a virus-specific manner. The presence of sulfate molecules distributed on the NS1 surface agrees with previous findings for DENV and ZIKV and indicates the potential for anionic ligand interaction. We thought that NS1 might interact with uninfected cell membranes via these sulfate binding sites (32), but further experiments confirmed that JEV NS1-C cannot bind efficiently to heparin, heparan sulfate, chondroitin sulfate, or dermatan sulfate polymers. Thus, the sulfate binding sites are not GAG binding interfaces and could represent a crystallographic artifact. Moreover, JEV NS1-C cannot bind to liposomes. Our results also suggest that the NS1 C terminus is not responsible for binding to the cell membrane through GAGs. Instead, cell membrane interactions may occur at the β-roll and wing domains, as was suggested previously (22–24).

B-factor and MD analyses suggest that loop 218 to 272 is conformationally dynamic. Although the B-factor values are high for this region, the X-ray structure does not show

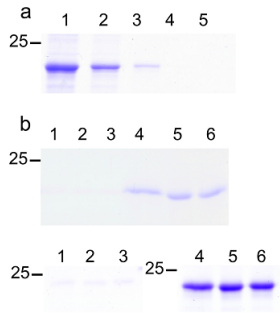


FIG 7 Cell membrane interaction via GAG determination. (a) Heparin binding determination. JEV NS1-C was incubated with heparin agarose beads. Total JEV NS1-C loaded to the column is shown in lane 1. Lane 2 shows the flowthrough fraction. Lanes 3 and 4 show wash fractions. The column was eluted with buffer supplemented with 1.5 M NaCl, and the results are shown in lane 5. Three independent experiments were conducted. (b) Liposome binding assay. The experiments were conducted at pH 7.5 (upper panel) and pH 5.5 (lower panel). Supernatant and pellet fractions separated by centrifugation were analyzed by SDS-PAGE. Lanes 1, 2, and 3 represent pellets from 400-nmol, 100-nmol, and 25-nmol reaction mixtures, respectively. Lanes 4, 5, and 6 represent supernatants from 400-nmol, 100-nmol, and 25-nmol reaction mixtures, respectively. Three independent experiments were conducted.

the more closely related JEV NS1-C at the same epitope, but with some conformational flexibility. This finding agrees with our MD result showing elasticity in the epitope loop, which may affect the antibody-NS1 structure in solution. Even though JEV NS1′-C has extra amino acids at the C terminus, JEV NS1′-C can interact with WNV 22NS1 MAb, indicating the C-terminal tail does not obstruct the binding surface of 22NS1. The C-tail may then locate at the side flanking the dimer. The presence of NS1′ is a shared characteristic of JE serocomplex viruses, and NS1′ may have specific properties that contribute to the propensity of JE serogroup viruses to cause encephalitis.

Despite flavivirus NS1 proteins having a conserved protein fold, these proteins differ

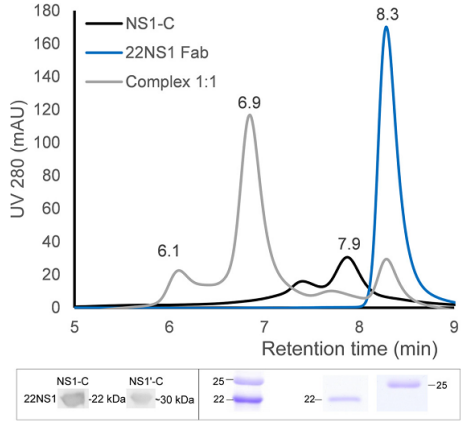


FIG 8 JEV NS1-C complexed with 22NS1 Fab. JEV NS1-C was detected by 22NS1 MAb (lower left inset). JEV NS1-C was incubated with 22NS1 Fab at 1:1 molar ratio of protein to Fab fragment, and complex formation was analyzed on an Agilent BioSEC-3 4.6/300 column. The lower right panel shows the results of SDS-PAGE analysis of each elution fraction. AU, arbitrary units.

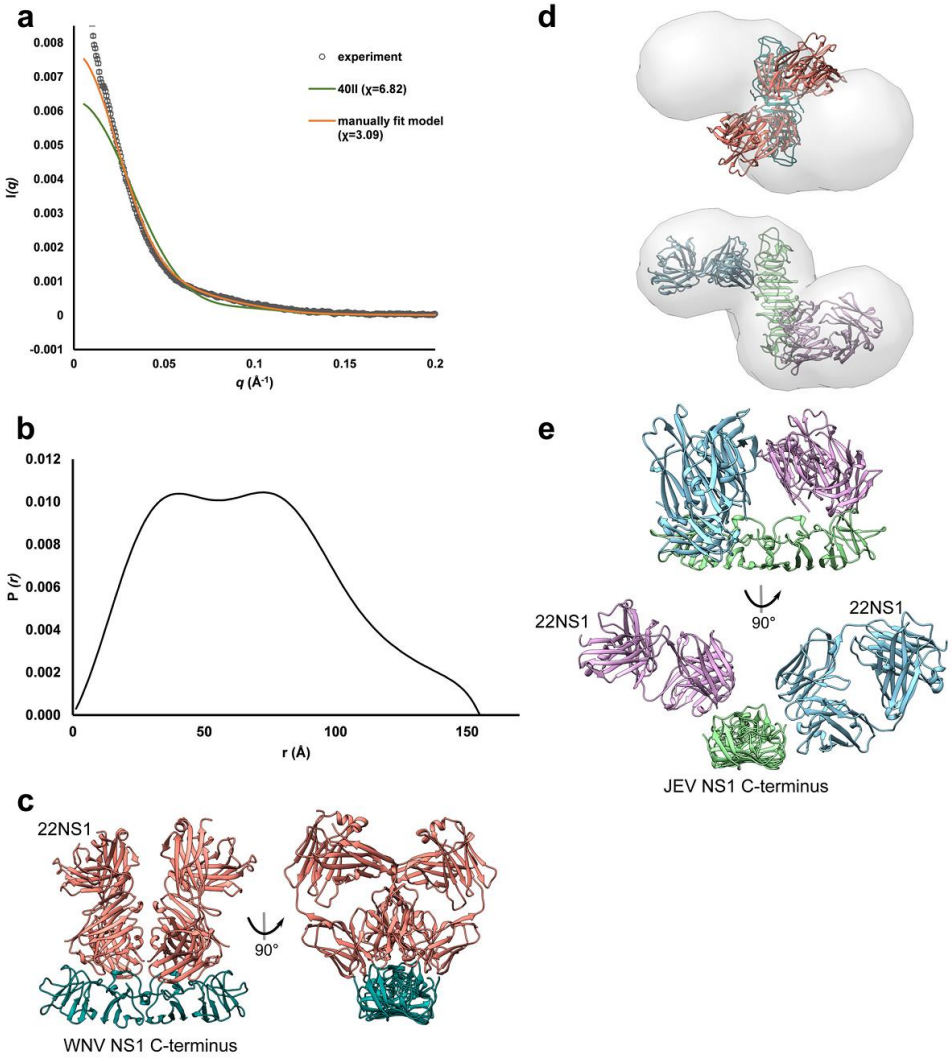


FIG 9 SAXS analysis of JEV NS1-C–22NS1 Fab complex. (a) SAXS curve. An experimental scattering curve of the JEV NS1-C–22NS1 Fab complex is shown in black scattering. The calculated scattering profile of the WNV NS1-C–22NS1 complex (PDB ID 4OII) is displayed in green, and data from the JEV NS1-C–22NS1 Fab complex manually fit model are shown in orange. (b) Pair distribution functions show multiple peaks signifying the multidomain structure. (c) WNV NS1-C–22NS1 complex (PDB ID 4OII). WNV NS1-C is colored in deep sky blue. 22NS1 Fabs are colored in salmon. (d) (Upper panel) WNV NS1-C–22NS1 complex (PDB ID 4OII) fitted to the JEV NS1-C–22NS1 Fab complex *ab initio* model. (Lower panel) A pseudoatomic model of the JEV NS1-C–22NS1 Fab complex was manually fitted to the *ab initio* model. (e) JEV NS1-C–22NS1 Fab complex pseudo-atomic model. JEV NS1-C is colored in light green. 22NS1 Fab is colored in orchid, and another Fab is colored in sky blue.

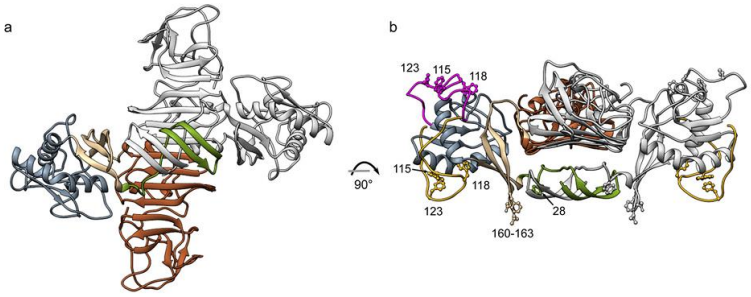


FIG 10 JEV NS1 homology model. The JEV NS1 full-length model was created by using SWISS-MODEL homology modeling. Dimerization was generated by superimposition of the JEV NS1 homology model onto ZIKV NS1 PDB ID 5GS6. (a) Cross-shaped homodimer NS1. One subunit is colored in gray, and another is colored by domain. A β-roll domain (amino acid residues 1 to 29) is colored in green, a wing domain (amino acids 38 to 151) is colored in blue, and β-ladder domains (amino acids 181 to 352) are colored in brown. (b) Side view of NS1. Residues 108 to 128 of the JEV homology model are indicated in magenta. Residues 108 to 128 are disordered and not visible in DENV (PDB ID 4O6B) or WNV (PDB ID 4O6C), but they are visible in ZIKV (PDB ID 5GS6 [shown in yellow] and 5K6K). Hydrophobic residues (namely, residues 28, 115, 118, 123, and 160 to 163) suspected to be involved with cell membrane interaction are labeled.

in their charge distributions, which may enable unique interactions with host proteins (8, 41). The fact that WNV 22NS1 MAb interacts with JEV NS1 is consistent with the high similarity of charge distributions of WNV and JEV NS1. This similarity also extends to ZIKV. Overall, these results provide structural details that aid NS1 function determination and highlight both similarities and contrasts among NS1 orthologs, which may be a productive avenue for developing common diagnostic and therapeutic strategies against diseases caused by the members of this important group of flaviviruses.

MATERIALS AND METHODS

Protein expression and refolding. JEV strain SA14 (GenBank accession no. M55506) was used as a template. Synthetic DNA optimized for expression of JEV NS1 in *E. coli* was acquired from Life Technologies. To create JEV NS1-C (amino acid residues 172 to 352), the target sequences were cloned into pET303 at XbaI/XhoI cloning sites by using forward primer 5′-gctctagaatgCGTGAAGAAAGCACCG ATGAATGTGAT-3′ and reverse primer 5′-ccgctcgagTTATGCATCAACCTGGCTACGAACCAG-3′; lowercase characters indicate the vector sequence that was required for the cloning step, and uppercase characters indicate the target sequence. Synthetic *JEV*NS1′ was purchased from GenScript (Piscataway, NJ, USA). The full-length NS1′ was the NS1 sequence with 156 additional nucleotides. The frameshift sequence was manually added by insertion of thymine at position 3561 as a result of −1 ribosomal frameshifting. JEV NS1′-C was generated from the synthetic *JEV*NS1′ by using forward primer 5′-gctctagaatg CGTGAAGAAAGCACCGATGAATGTGAT-3′ and reverse primer 5′-ccgctcgagTTAATGCAGATGATAACCC CATGCATctg-3′. Proteins were expressed in *E. coli* by autoinduction and refolded by using a method modified from that previously described by Edeling et al. (21). The theoretical molecular masses of JEV NS1-C and JEV NS1′-C are 20.54 kDa and 25.98 kDa. Protein yield and purity were analyzed by SDS-PAGE.

Protein crystallization and data collection. JEV NS1-C (~6 mg/ml) and JEV NS1′-C (5 to 7 mg/ml) were screened using commercial crystallization screens. Successful conditions were optimized by the use of the hanging drop method. Needle crystals of JEV NS1-C were produced from 1 M ammonium sulfate and 0.1 M MES [2-(*N*-morpholino)ethanesulfonic acid; pH 5.5]. The crystals were flash frozen in reservoir solution with 20% to 25% ethylene glycol added. JEV NS1′-C also crystallized in needle form in 1 M ammonium sulfate–5% propanol. JEV NS1′-C was cryo-protected in reservoir solution–20% glycerol.

X-ray data were collected at cryogenic temperature, at a wavelength of 0.98 Å, at beamline PROXIMA 1 at the Soleil synchrotron (France) and at beamline I02 at the Diamond Light Source (United Kingdom). Data reduction was carried out by the use of XDS programs (52) or iMOSFLM (53). The protein structure was determined by molecular replacement using the structure of the WNV NS1 C-terminal domain (PDB ID 4OII; sequence identity, >70%) as a starting model by the use of MOLREP (54) in the CCP4 program suite. The structure was refined by the use of REFMAC5 (55) and built in COOT (56). Data collection and refinement statistics are shown in Table 6. The JEV NS1-C refinement statistic values from a Ramachandran plot are 95.98% favored and 0% outliers. The MolProbity score is 1.6. The JEV NS1′-C refinement statistic values from a Ramachandran plot are 94.89% favored and 0% outliers. The MolProbity score is 1.84.

Protein structure analysis. Assembly analysis was performed by the use of the program PISA (57). Conservation scores of residues on protein structures were determined by the use of ConSurf (58) with

TABLE 6 Data collection and refinement statistics

Parameter	Value(s)[b]	
	JEV NS1-C	JEV NS1′-C
Data collection		
Space group	I212121	I212121
Cell dimensions		
a, *b*, *c* (Å)	49.42, 78.24, 163.18	50.32, 77.94, 163.49
a, b, g (°)	90, 90, 90	90, 90, 90
Resolution (Å)	47.3–2.10 (2.21–2.10)	81.75–2.6 (2.72–2.6)
R_{merge}	0.103 (0.907)	0.2 (1.413)
R_{pim}	0.045 (0.383)	0.141 (1.030)
$I/\sigma I$	11.5 (2.3)	7.1 (2.1)
Completeness (%)	99.8 (99.6)	99.9 (99.9)
Redundancy	6.3 (6.5)	5.3 (5.2)
Refinement		
Resolution (Å)	47.3–2.10	81.75–2.6
No. of reflections	17,944	9719
R_{work}/R_{free}	0.189/0.228	0.166/0.225
No. of atoms	1574	1573
Protein	1398	1418
Sulfate ion	60	60
Ligand	24 (MES)	4 (POL[a])
Water	92	91
B-factors		
Protein	41.516	38.612
Sulfate ion	90.121	85.148
Ligand	86.736 (MES)	61.722 (POL)
Water	50.187	48.669
RMS deviations		
Bond lengths (Å)	0.016	0.016
Bond angles (°)	1.785	1.741

[a]POL, *n*-propanol.
[b]Values in parentheses are for the highest-resolution shell.

21 homologous sequences. The input homologous sequences of NS1 C terminus were searched by the program, and existing NS1 structure sequences were added manually. Electrostatic surface maps were generated by using PDB2PQR (59) to convert PDB files into PQR files and Adaptive Poisson-Boltzmann Solver (APBS) for electrostatics calculations (60) without pK_a prediction.

JEV NS1 C terminus-22NS1 complex formation. Complex formation was confirmed by Western blotting, with 22NS1 (36) and goat anti-mouse IgG-horseradish peroxidase (IgG-HRP; Santa Cruz Biotechnology, sc-2055) used as primary and secondary antibodies, respectively. Purified JEV NS1 C terminus and 22NS1 fragment antigen binding (Fab) (prepared from 22NS1 IgG MAb using a Pierce Fab Preparation kit; catalog no. 44985) were mixed overnight at 4°C at a 1:1 ratio of JEV NS1-C to 22NS1 and purified by the use of Agilent Bio SEC-3 4.6 300 or GE Superdex 200 10 300 GL. Eluted fractions were analyzed by SDS-PAGE.

SAXS data collection and processing. JEV NS1-C at a concentration of 3.4 mg/ml and JEV NS1-C–22NS1 Fab complex at concentration of 3 mg/ml in TBS buffer (20 mM Tris-HCl [pH 7.4], 150 mM NaCl) were analyzed with SEC-SAXS on beamline SWING at the Soleil synchrotron (France). Samples were loaded onto an Agilent BioSEC-3 4.6/300 column at a flow rate of 0.25 ml/min at 15°C. Data were collected at a distance of 1.8 m and an X-ray wavelength of 1 Å. Data processing was conducted in PRIMUS (61). Comparison of scattering profiles was done in FoXS (62). The *ab initio* model included averages from 10 (JEV NS1-C) or 20 (protein complex) independent model calculations with symmetry (protein complex) or without symmetry (JEV NS1-C) determined using DAMMIF (63). The model data were averaged with DAMAVER (64) and refined with DAMMIN (65). The low-resolution-model surface representation was created in CHIMERA (66) using the 'molmap' command. The molecular mass was calculated from Porod volume data (67). Molecular dynamics (MD) simulations were performed using GROMACS 4.6.5 and GROMOS96 54A7 force fields in a cubic box solvated with single-point charge-E water molecules on JEV NS1-C dimers. A neutral charge was introduced at 150 mM NaCl. The distance between the JEV NS1-C dimers and the box edge was set to 10 Å. Long-range interactions were defined using the particle mesh Ewald algorithm, and other nonbonded interactions were restricted to 10 Å. An energy minimization step was performed using the steepest descent algorithm followed by a 100-ps NVT ensemble at 310 K and a 200-ps NPT ensemble at 310 K and 10^5 Pa. Production MD analyses were performed at 310 K and 10^5 Pa for 40 ns. Cα displacement was calculated with the GROMACS root mean square fluctuation (RMSF) function. Torsion angle MD analyses were performed with CNS software at 100,000 K for 37.5 ps with sampling performed every 7.5 fs in eight separate simulations. The best structure was found with FoXS using experimental data over a data range of $0.017 < q < 0.25$ Å^{-1}, where

$q = (4\pi \sin \theta)/\lambda$, 2θ is the angle between the incident X-ray beam and the detector measuring the scattered intensity, and λ is the wavelength of the X-rays, and was refined with another eight separate 7.5-ps simulations and energy minimization in GROMACS using the procedure described above. The models were compared again using FoXS. Freeing loop 214 to 243 gave a value of 1.66 corresponding to the fit with experimental data. Expanding the flexible region to 218 to 272 allowed us to improve the fit to 1.48.

Liposome binding assay. The liposome preparation method was modified from methods described in previous publications (49, 68). Liposomes were prepared from cholesterol (CHOL) (Sigma, C8667) and 1,2-dipalmitoyl-sn-glycero-3-phosphocholine (PC) (Sigma, P4329) at a 1:9 ratio of CHOL to PC (22, 49, 68). CHOL and PC powders were dissolved in chloroform. To achieve a total of 400 nmol, 40 nmol of CHOL and 360 nmol of PC were mixed together in a 2-ml tube, and the lipid mixture was dried under a nitrogen gas stream. To hydrate the lipid sheets, 50 μl of buffer consisting of 50 mM Bis-Tris (pH 5.5), 50 mM $(NH_4)_2SO_4$, and 10% glycerol or 150 mM KCl, 25 mM Tris-HCl (pH 7.5), 1 mM dithiothreitol (DTT), and 0.5 mM (EDTA) was added and the reaction mixture was incubated at room temperature on a shaker for 30 min. Then, the lipid was sonicated with an exponential probe at amplitude 4 for 30 s with a 30-s interval on a warmed water bath for 5 times. A liposome binding reaction (50 μl) was set up using 400 nmol, 125 nmol, and 25 nmol of total lipid, and the reaction mixture was then mixed with 5 μg of protein. The reaction mixtures were incubated at 37°C for 45 min. After that, the reactions were centrifuged at 16,000 $\times$ *g* for 30 min at 22°C and the supernatant was transferred to a new tube. The lipid pellet was resuspended in 200 μl buffer and also transferred to a new tube. Liposomes were pelleted again, and the supernatant was discarded. The liposome pellet was resuspended in 30 μl of 1$\times$ SDS-PAGE sample buffer. Bovine cytochrome bc1 complex (membrane proteins) was used as the positive control in a mixture consisting of 25 mM phosphate buffer (pH 7.5), 100 mM NaCl, 3 mM NaN_3, and 0.015% DDM (*n*-dodecyl-D-maltoside) buffer. The supernatant and pellet fractions were analyzed by SDS-PAGE.

Heparin binding assay. A small-scale 50-μl column was set up using a pipette tip and heparin agarose beads (Affi-Gel heparin gel; Bio-Rad). The binding buffer was 20 mM HEPES (pH 7.4) supplemented with 150 mM NaCl, and the elution buffer was 20 mM HEPES (pH 7.4) supplemented with 1.5 M and 2 M NaCl. The column was equilibrated with 400 μl of binding buffer. JEV NS1-C (5 μg) was applied to the column, and the column was incubated on a roller for 30 min at 4°C. The column was washed 3 times with 400 μl binding buffer before elution was performed twice with 100 μl of 1.5 M and 2 M NaCl elution buffer. Superoxide dismutase 3 (SOD3), which contains a heparin binding domain, was used as a positive control. Samples from each step (including the loading, flowthrough, washing, and elution steps) were analyzed by SDS-PAGE.

Differential scanning fluorimetry (DSF). Polymers of heparan sulfate (average molecular weight, 30,000), chondroitin sulfate (62% chondroitin 4-sulfate and 33% chondroitin 6-sulfate; average molecular weight, 45,400), and dermatan sulfate (average molecular weight, 41,000) (Iduron) at final concentrations of 100, 50, 25, 10, 5, 1, and 0.5 μM were mixed with JEV NS1-C and Sypro Orange 5000X (Invitrogen) at final concentrations of 10 μM and 10$\times$, respectively. The reaction volume was 10 μl. The experiments were set up in 96-well plates and performed using StepOnePlus Real-Time PCR Systems (software version 2.3) (Applied Biosystems). The reactions were equilibrated at 25°C for 2 min followed by an increase to 95°C at 1°C min^{-1}. The experiments were performed in three replicates.

Data availability. The atomic coordinates and structure factors have been deposited in the Protein Data Bank (www.pdb.org) (PDB ID 5O19 for JEV NS1-C and 5O36 for JEV NS1′-C).

ACKNOWLEDGMENTS

This work was supported by a Mahidol-Liverpool Stang Mongkolsuk Ph.D. scholarship. We acknowledge help from the following colleagues at the University of Liverpool: Richard Strange for assistance with the electrostatic surface map; Sujitra Keadsanti for assistance with Western blotting; Kangsa Amporndanai for providing bovine cytochrome *bc1*; Varunya Chantadul for providing human SOD3; and Pawin Ngamlert for assistance with the heparin binding assay. We also thank Samar Hasnain for support and interest in the project throughout and for extensive discussions of the results. We acknowledge Synchrotron Soleil for provision of the Proxima 1 beamline and SAXS facilities.

Use of Soleil was funded by the European Community's Seventh Framework Programme (FP7/2007-2013) under BioStruct-X (grant agreement number 283570 and proposal number 6714). We gratefully acknowledge the Diamond Synchrotron for providing support at the I02 beamline. M.S.D. acknowledges support from HHSN272201400018C, L.T. from The Wellcome Trust (grant number 205228/Z/16/Z), and T.S. and L.T. from the National Institute for Health Research (NIHR; www.nihr.ac.uk) Health Protection Research Unit in Emerging and Zoonotic Infections.

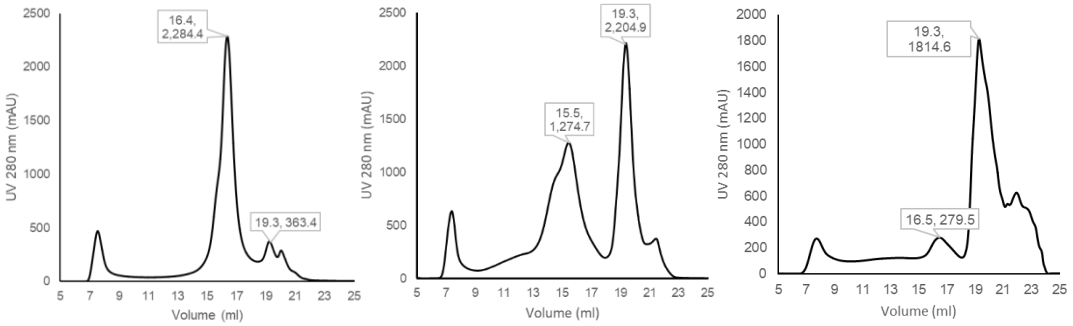
T.S. and S.V.A. originated and designed the project; T.P. expressed and purified proteins; T.P. and G.S.A.W. performed the experiments; T.P., G.S.A.W., and S.V.A. undertook data analysis; and T.P., G.S.A.W., M.S.D., T.S., L.T., and S.V.A. contributed to interpretation of data and wrote the manuscript.

REFERENCES

1. Lindenbach BD, Rice CM. 1997. trans-Complementation of yellow fever virus NS1 reveals a role in early RNA replication. J Virol 71:9608–9617.
2. Mackenzie JM, Jones MK, Young PR. 1996. Immunolocalization of the dengue virus nonstructural glycoprotein NS1 suggests a role in viral RNA replication. Virology 220:232–240. https://doi.org/10.1006/viro.1996.0307.
3. Muylaert IR, Chambers TJ, Galler R, Rice CM. 1996. Mutagenesis of the N-linked glycosylation sites of the yellow fever virus NS1 protein: effects on virus replication and mouse neurovirulence. Virology 222:159–168. https://doi.org/10.1006/viro.1996.0406.
4. Youn S, Ambrose RL, Mackenzie JM, Diamond MS. 2013. Non-structural protein-1 is required for West Nile virus replication complex formation and viral RNA synthesis. Virol J 10:339. https://doi.org/10.1186/1743-422X-10-339.
5. Fan J, Liu Y, Yuan Z. 2014. Critical role of dengue virus NS1 protein in viral replication. Virol Sin 29:162–169. https://doi.org/10.1007/s12250-014-3459-1.
6. Lindenbach BD, Rice CM. 1999. Genetic interaction of flavivirus nonstructural proteins NS1 and NS4A as a determinant of replicase function. J Virol 73:4611–4621.
7. Youn S, Li T, McCune BT, Edeling MA, Fremont DH, Cristea IM, Diamond MS. 2012. Evidence for a genetic and physical interaction between nonstructural proteins NS1 and NS4B that modulates replication of West Nile virus. J Virol 86:7360–7371. https://doi.org/10.1128/JVI.00157-12.
8. Krishna VD, Rangappa M, Satchidanandam V. 2009. Virus-specific cytolytic antibodies to nonstructural protein 1 of Japanese encephalitis virus effect reduction of virus output from infected cells. J Virol 83:4766–4777. https://doi.org/10.1128/JVI.01850-08.
9. Winkler G, Randolph VB, Cleaves GR, Ryan TE, Stollar V. 1988. Evidence that the mature form of the flavivirus nonstructural protein NS1 is a dimer. Virology 162:187–196. https://doi.org/10.1016/0042-6822(88)90408-4.
10. Schlesinger JJ, Brandriss MW, Putnak JR, Walsh EE. 1990. Cell surface expression of yellow fever virus non-structural glycoprotein NS1: consequences of interaction with antibody. J Gen Virol 71(Pt 3):593–599. https://doi.org/10.1099/0022-1317-71-3-593.
11. Lee JM, Crooks AJ, Stephenson JR. 1989. The synthesis and maturation of a non-structural extracellular antigen from tick-borne encephalitis virus and its relationship to the intracellular NS1 protein. J Gen Virol 70(Pt 2):335–343. https://doi.org/10.1099/0022-1317-70-2-335.
12. Avirutnan P, Hauhart RE, Somnuke P, Blom AM, Diamond MS, Atkinson JP. 2011. Binding of flavivirus nonstructural protein NS1 to C4b binding protein modulates complement activation. J Immunol 187:424–433. https://doi.org/10.4049/jimmunol.1100750.
13. Avirutnan P, Fuchs A, Hauhart RE, Somnuke P, Youn S, Diamond MS, Atkinson JP. 2010. Antagonism of the complement component C4 by flavivirus nonstructural protein NS1. J Exp Med 207:793–806. https://doi.org/10.1084/jem.20092545.
14. Chung KM, Liszewski MK, Nybakken G, Davis AE, Townsend RR, Fremont DH, Atkinson JP, Diamond MS. 2006. West Nile virus nonstructural protein NS1 inhibits complement activation by binding the regulatory protein factor H. Proc Natl Acad Sci U S A 103:19111–19116. https://doi.org/10.1073/pnas.0605668103.
15. Morrison CR, Scholle F. 2014. Abrogation of TLR3 inhibition by discrete amino acid changes in the C-terminal half of the West Nile virus NS1 protein. Virology 456–457:96–107.
16. Glasner DR, Ratnasiri K, Puerta-Guardo H, Espinosa DA, Beatty PR, Harris E. 2017. Dengue virus NS1 cytokine-independent vascular leak is dependent on endothelial glycocalyx components. PLoS Pathog 13:e1006673. https://doi.org/10.1371/journal.ppat.1006673.
17. Modhiran N, Watterson D, Blumenthal A, Baxter AG, Young PR, Stacey KJ. 21 February 2017. Dengue virus NS1 protein activates immune cells via TLR4 but not TLR2 or TLR6. Immunol Cell Biol https://doi.org/10.1038/icb.2017.5.
18. Modhiran N, Watterson D, Muller DA, Panetta AK, Sester DP, Liu L, Hume DA, Stacey KJ, Young PR. 2015. Dengue virus NS1 protein activates cells via Toll-like receptor 4 and disrupts endothelial cell monolayer integrity. Sci Transl Med 7:304ra142. https://doi.org/10.1126/scitranslmed.aaa3863.
19. Han YW, Choi JY, Uyangaa E, Kim SB, Kim JH, Kim BS, Kim K, Eo SK. 2014. Distinct dictation of Japanese encephalitis virus-induced neuroinflammation and lethality via triggering TLR3 and TLR4 signal pathways. PLoS Pathog 10:e1004319. https://doi.org/10.1371/journal.ppat.1004319.
20. Muller DA, Young PR. 2013. The flavivirus NS1 protein: molecular and structural biology, immunology, role in pathogenesis and application as a diagnostic biomarker. Antiviral Res 98:192–208. https://doi.org/10.1016/j.antiviral.2013.03.008.
21. Edeling MA, Diamond MS, Fremont DH. 2014. Structural basis of flavivirus NS1 assembly and antibody recognition. Proc Natl Acad Sci U S A 111:4285–4290. https://doi.org/10.1073/pnas.1322036111.
22. Akey DL, Brown WC, Dutta S, Konwerski J, Jose J, Jurkiw TJ, DelProposto J, Ogata CM, Skiniotis G, Kuhn RJ, Smith JL. 2014. Flavivirus NS1 structures reveal surfaces for associations with membranes and the immune system. Science 343:881–885. https://doi.org/10.1126/science.1247749.
23. Xu X, Song H, Qi J, Liu Y, Wang H, Su C, Shi Y, Gao GF. 30 August 2016. Contribution of intertwined loop to membrane association revealed by Zika virus full-length NS1 structure. EMBO J https://doi.org/10.15252/embj.201695290.
24. Brown WC, Akey DL, Konwerski JR, Tarrasch JT, Skiniotis G, Kuhn RJ, Smith JL. 2016. Extended surface for membrane association in Zika virus NS1 structure. Nat Struct Mol Biol 23:865–867. https://doi.org/10.1038/nsmb.3268.
25. Blitvich BJ, Scanlon D, Shiell BJ, Mackenzie JS, Pham K, Hall RA. 2001. Determination of the intramolecular disulfide bond arrangement and biochemical identification of the glycosylation sites of the nonstructural protein NS1 of Murray Valley encephalitis virus. J Gen Virol 82:2251–2256. https://doi.org/10.1099/0022-1317-82-9-2251.
26. Mandl CW, Heinz FX, Stockl E, Kunz C. 1989. Genome sequence of tick-borne encephalitis virus (Western subtype) and comparative analysis of nonstructural proteins with other flaviviruses. Virology 173:291–301. https://doi.org/10.1016/0042-6822(89)90246-8.
27. Watterson D, Modhiran N, Young PR. 2016. The many faces of the flavivirus NS1 protein offer a multitude of options for inhibitor design. Antiviral Res 130:7–18. https://doi.org/10.1016/j.antiviral.2016.02.014.
28. Gutsche I, Coulibaly F, Voss JE, Salmon J, d'Alayer J, Ermonval M, Larquet E, Charneau P, Krey T, Megret F, Guittet E, Rey FA, Flamand M. 2011. Secreted dengue virus nonstructural protein NS1 is an atypical barrel-shaped high-density lipoprotein. Proc Natl Acad Sci U S A 108:8003–8008. https://doi.org/10.1073/pnas.1017338108.
29. Jacobs MG, Robinson PJ, Bletchly C, Mackenzie JM, Young PR. 2000. Dengue virus nonstructural protein 1 is expressed in a glycosyl-phosphatidylinositol-linked form that is capable of signal transduction. FASEB J 14:1603–1610. https://doi.org/10.1096/fj.99-0829com.
30. Noisakran S, Dechtawewat T, Avirutnan P, Kinoshita T, Siripanyaphinyo U, Puttikhunt C, Kasinrerk W, Malasit P, Sittisombut N. 2008. Association of dengue virus NS1 protein with lipid rafts. J Gen Virol 89:2492–2500. https://doi.org/10.1099/vir.0.83620-0.
31. Noisakran S, Dechtawewat T, Rinkaewkan P, Puttikhunt C, Kanjanahaluethai A, Kasinrerk W, Sittisombut N, Malasit P. 2007. Characterization of dengue virus NS1 stably expressed in 293T cell lines. J Virol Methods 142:67–80. https://doi.org/10.1016/j.jviromet.2007.01.008.
32. Avirutnan P, Zhang L, Punyadee N, Manuyakorn A, Puttikhunt C, Kasinrerk W, Malasit P, Atkinson JP, Diamond MS. 2007. Secreted NS1 of dengue virus attaches to the surface of cells via interactions with heparan sulfate and chondroitin sulfate E. PLoS Pathog 3:e183. https://doi.org/10.1371/journal.ppat.0030183.
33. Li YZ, Counor D, Lu P, Liang GD, Vu TQ, Phan TN, Huynh TK, Sun G, Grandadam M, Butrapet S, Lavergne JP, Flamand M, Yu YX, Solomon T, Buchy P, Deubel V. 2012. A specific and sensitive antigen capture assay for NS1 protein quantitation in Japanese encephalitis virus infection. J Virol Methods 179:8–16. https://doi.org/10.1016/j.jviromet.2011.06.008.
34. Amorim JH, Alves RP, Boscardin SB, Ferreira LC. 2014. The dengue virus non-structural 1 protein: risks and benefits. Virus Res 181:53–60. https://doi.org/10.1016/j.virusres.2014.01.001.
35. Solomon T, Thao LT, Dung NM, Kneen R, Hung NT, Nisalak A, Vaughn DW, Farrar J, Hien TT, White NJ, Cardosa MJ. 1998. Rapid diagnosis of Japanese encephalitis by using an immunoglobulin M dot enzyme immunoassay. J Clin Microbiol 36:2030–2034.
36. Chung KM, Nybakken GE, Thompson BS, Engle MJ, Marri A, Fremont DH, Diamond MS. 2006. Antibodies against West Nile virus nonstructural protein NS1 prevent lethal infection through Fc gamma receptor-dependent and -independent mechanisms. J Virol 80:1340–1351. https://doi.org/10.1128/JVI.80.3.1340-1351.2006.
37. Schlesinger JJ, Brandriss MW, Cropp CB, Monath TP. 1986. Protection

against yellow fever in monkeys by immunization with yellow fever virus nonstructural protein NS1. J Virol 60:1153–1155.
38. Schlesinger JJ, Foltzer M, Chapman S. 1993. The Fc portion of antibody to yellow fever virus NS1 is a determinant of protection against YF encephalitis in mice. Virology 192:132–141. https://doi.org/10.1006/viro.1993.1015.
39. Li Y, Counor D, Lu P, Duong V, Yu Y, Deubel V. 2012. Protective immunity to Japanese encephalitis virus associated with anti-NS1 antibodies in a mouse model. Virol J 9:135. https://doi.org/10.1186/1743-422X-9-135.
40. Liu J, Liu Y, Nie K, Du S, Qiu J, Pang X, Wang P, Cheng G. 2016. Flavivirus NS1 protein in infected host sera enhances viral acquisition by mosquitoes. Nat Microbiol 1:16087. https://doi.org/10.1038/nmicrobiol.2016.87.
41. Cheng HJ, Lin CF, Lei HY, Liu HS, Yeh TM, Luo YH, Lin YS. 2009. Proteomic analysis of endothelial cell autoantigens recognized by anti-dengue virus nonstructural protein 1 antibodies. Exp Biol Med (Maywood) 234:63–73. https://doi.org/10.3181/0805-RM-147.
42. Beatty PR, Puerta-Guardo H, Killingbeck SS, Glasner DR, Hopkins K, Harris E. 2015. Dengue virus NS1 triggers endothelial permeability and vascular leak that is prevented by NS1 vaccination. Sci Transl Med 7:304ra141. https://doi.org/10.1126/scitranslmed.aaa3787.
43. Melian EB, Hinzman E, Nagasaki T, Firth AE, Wills NM, Nouwens AS, Blitvich BJ, Leung J, Funk A, Atkins JF, Hall R, Khromykh AA. 2010. NS1′ of flaviviruses in the Japanese encephalitis virus serogroup is a product of ribosomal frameshifting and plays a role in viral neuroinvasiveness. J Virol 84:1641–1647. https://doi.org/10.1128/JVI.01979-09.
44. Mason PW. 1989. Maturation of Japanese encephalitis virus glycoproteins produced by infected mammalian and mosquito cells. Virology 169:354–364. https://doi.org/10.1016/0042-6822(89)90161-X.
45. Young LB, Melian EB, Khromykh AA. 2013. NS1′ colocalizes with NS1 and can substitute for NS1 in West Nile virus replication. J Virol 87:9384–9390. https://doi.org/10.1128/JVI.01101-13.
46. Ye Q, Li XF, Zhao H, Li SH, Deng YQ, Cao RY, Song KY, Wang HJ, Hua RH, Yu YX, Zhou X, Qin ED, Qin CF. 2012. A single nucleotide mutation in NS2A of Japanese encephalitis-live vaccine virus (SA14-14-2) ablates NS1′ formation and contributes to attenuation. J Gen Virol 93:1959–1964. https://doi.org/10.1099/vir.0.043844-0.
47. Melian EB, Hall-Mendelin S, Du F, Owens N, Bosco-Lauth AM, Nagasaki T, Rudd S, Brault AC, Bowen RA, Hall RA, van den Hurk AF, Khromykh AA. 2014. Programmed ribosomal frameshift alters expression of West Nile virus genes and facilitates virus replication in birds and mosquitoes. PLoS Pathog 10:e1004447. https://doi.org/10.1371/journal.ppat.1004447.
48. Takamatsu Y, Okamoto K, Dinh DT, Yu F, Hayasaka D, Uchida L, Nabeshima T, Buerano CC, Morita K. 2014. NS1′ protein expression facilitates production of Japanese encephalitis virus in avian cells and embryonated chicken eggs. J Gen Virol 95:373–383. https://doi.org/10.1099/vir.0.057968-0.
49. Smith JL, Akey DL, Brown WC, Kuhn RJ. June 2015. Vaccine compositions and uses thereof.US patent WO 2015095735 A2.
50. Song H, Qi J, Haywood J, Shi Y, Gao GF. 2016. Zika virus NS1 structure reveals diversity of electrostatic surfaces among flaviviruses. Nat Struct Mol Biol 23:456–458. https://doi.org/10.1038/nsmb.3213.
51. Akey DL, Brown WC, Jose J, Kuhn RJ, Smith JL. 2015. Structure-guided insights on the role of NS1 in flavivirus infection. Bioessays 37:489–494. https://doi.org/10.1002/bies.201400182.
52. Kabsch W. 2010. XDS. Acta Crystallogr D Biol Crystallogr 66:125–132. https://doi.org/10.1107/S0907444909047337.
53. Battye TG, Kontogiannis L, Johnson O, Powell HR, Leslie AG. 2011. iMOSFLM: a new graphical interface for diffraction-image processing with MOSFLM. Acta Crystallogr D Biol Crystallogr 67:271–281. https://doi.org/10.1107/S0907444910048675.
54. Vagin A, Teplyakov A. 2010. Molecular replacement with MOLREP. Acta Crystallogr D Biol Crystallogr 66:22–25. https://doi.org/10.1107/S0907444909042589.
55. Murshudov GN, Skubak P, Lebedev AA, Pannu NS, Steiner RA, Nicholls RA, Winn MD, Long F, Vagin AA. 2011. REFMAC5 for the refinement of macromolecular crystal structures. Acta Crystallogr D Biol Crystallogr 67:355–367. https://doi.org/10.1107/S0907444911001314.
56. Emsley P, Lohkamp B, Scott WG, Cowtan K. 2010. Features and development of Coot. Acta Crystallogr D Biol Crystallogr 66:486–501. https://doi.org/10.1107/S0907444910007493.
57. Krissinel E, Henrick K. 2007. Inference of macromolecular assemblies from crystalline state. J Mol Biol 372:774–797. https://doi.org/10.1016/j.jmb.2007.05.022.
58. Landau M, Mayrose I, Rosenberg Y, Glaser F, Martz E, Pupko T, Ben-Tal N. 2005. ConSurf 2005: the projection of evolutionary conservation scores of residues on protein structures. Nucleic Acids Res 33:W299–W302. https://doi.org/10.1093/nar/gki370.
59. Dolinsky TJ, Czodrowski P, Li H, Nielsen JE, Jensen JH, Klebe G, Baker NA. 2007. PDB2PQR: expanding and upgrading automated preparation of biomolecular structures for molecular simulations. Nucleic Acids Res 35:W522–W525. https://doi.org/10.1093/nar/gkm276.
60. Baker NA, Sept D, Joseph S, Holst MJ, McCammon JA. 2001. Electrostatics of nanosystems: application to microtubules and the ribosome. Proc Natl Acad Sci U S A 98:10037–10041. https://doi.org/10.1073/pnas.181342398.
61. Konarev PV, Volkov VV, Sokolova AV, Koch MHJ, Svergun DI. 2003. PRIMUS: a Windows PC-based system for small-angle scattering data analysis. J Appl Crystallogr 36:1277–1282. https://doi.org/10.1107/S0021889803012779.
62. Schneidman-Duhovny D, Hammel M, Tainer JA, Sali A. 2013. Accurate SAXS profile computation and its assessment by contrast variation experiments. Biophys J 105:962–974. https://doi.org/10.1016/j.bpj.2013.07.020.
63. Franke D, Svergun DI. 2009. DAMMIF, a program for rapid ab-initio shape determination in small-angle scattering. J Appl Crystallogr 42:342–346. https://doi.org/10.1107/S0021889809000338.
64. Volkov VV, Svergun DI. 2003. Uniqueness of ab initio shape determination in small-angle scattering. J Appl Crystallogr 36:860–864. https://doi.org/10.1107/S0021889803000268.
65. Svergun DI. 1999. Restoring low resolution structure of biological macromolecules from solution scattering using simulated annealing. Biophys J 76:2879–2886. https://doi.org/10.1016/S0006-3495(99)77443-6.
66. Pettersen EF, Goddard TD, Huang CC, Couch GS, Greenblatt DM, Meng EC, Ferrin TE. 2004. UCSF Chimera–a visualization system for exploratory research and analysis. J Comput Chem 25:1605–1612. https://doi.org/10.1002/jcc.20084.
67. Petoukhov MV, Franke D, Shkumatov AV, Tria G, Kikhney AG, Gajda M, Gorba C, Mertens HDT, Konarev PV, Svergun DI. 2012. New developments in the ATSAS program package for small-angle scattering data analysis. J Appl Crystallogr 45:342–350. https://doi.org/10.1107/S0021889812007662.
68. Julkowska MM, Rankenberg JM, Testerink C. 2013. Liposome-binding assays to assess specificity and affinity of phospholipid-protein interactions. Methods Mol Biol 1009:261–271. https://doi.org/10.1007/978-1-62703-401-2_24.

Appendix 2 Chromatogram of C-*JEV*NS1, C-*JEV*NS1', and *JEV*NS1, respectively, purified by Superose 6 10x300 mm column.



Appendix 3 Crystallization screen formulation that gave crystals.

No.	Screens	Formulations	Protein concentration
1	SaltRx	1.5 M Ammonium sulphate, 0.1 M Sodium acetate trihydrate pH 4.6	C-*JEV*NS1 4 mg/ml
2		2.4 M Ammonium phosphate dibasic, 0.1 M Tris pH 8.5	
3	PEGRx	18% (v/v) 2-propanol, 0.1 M Sodium citrate tribasic dihydrate pH 5.5, 20% (w/v) PEG 4000	C-*JEV*NS1 6 mg/ml
4		15% (v/v) 2-propanol, 0.1 M Sodium citrate tribasic dihydrate pH 5.0, 10% (w/v) PEG 10000	
5	Natrix	0.02 M Magnesium sulphate hydrate, 0.002 M Cobalt(II) chloride hexahydrate, 0.05 M Sodium	

No.	Screens	Formulations	Protein concentration
	Natrix	cacodylate trihydrate pH 6, 25% (v/v) (+/-)-2-Methyl-2,4-pentanediol, 0.0005 M Spermine	C-*JEV*NS1 6 mg/ml
6		0.08 M Sodium chloride, 0.02 M Barium chloride dihydrate, 0.04 M Sodium cacodylate trihydrate pH 7, 40% (v/v) (+/-)-2-Methyl-2,4-pentanediol, 0.012 M Spermine tetrahydrochloride	
7		0.08 M Strontium chloride hexahydrate, 0.02 M Magnesium chloride hexahydrate, 0.04 M Sodium cacodylate trihydrate pH 7, 20% (v/v) (+/-)-2-Methyl-2,4-pentanediol, 0.012 M Spermine tetrahydrochloride	
8	Structure	1.0 M Ammonium phosphate monobasic, 0.1 M Sodium citrate pH 5.6	
9	JCSG	0.02 M Calcium chloride dihydrate, 0.1 M Sodium acetate pH 4.6, 30% (v/v) (+/-)-2-Methyl-2,4-pentanediol	
10		0.1 M Sodium cacodylate pH 6.5, 40% (v/v) (+/-)-2-Methyl-2,4-pentanediol, 5% (w/v) PEG 8000	

No.	Screens	Formulations	Protein concentration
11	Structure	0.1 M Sodium citrate pH 5.6, 20 % (w/v) PEG 4000, 20% (v/v) 2-propanol	C-*JEV*NS1' 7 mg/ml
12		0.2 M Magnesium acetate tetrahydrate, 0.1 M Sodium cacodylate pH 6.5, 20% (w/v) PEG 8000	
13		1.5 M Litium sulphate, 0.1 M Sodium HEPES pH 7.5	
14		2.0 M Ammonium sulphate, 0.1 M Sodium HEPES pH 7.5, 2% (v/v) PEG 400	
15		0.1 M Sodium HEPES pH 7.5, 20% (v/v) PEG 400, 10% 2-Propanol	
16		0.05 M Potassium phosphat monobasic	
17		30% (w/v) PEG 1500	
18		1.5 M Ammonium sulphate, 0.1 M Tris pH 8.5, 12% (v/v) Glycerol	
19	Structure	0.1 M Sodium HEPES pH 7.5, 70% (v/v) MDP	C-*JEV*NS1' 7 mg/ml
20		1.6 M Ammonium sulphate, 0.1 M MES pH 6.5, 10% (v/v) 1,4-Dioxane	

No.	Screens	Formulations	Protein concentration
21		0.1 M Sodium HEPES pH 7.5, 20% (w/v) PEG 10000	
22		2.0 M Ammonium sulphate, 0.2 M Potassium sodium tartrate tetrahydrate, 0.1 M Sodium citrate pH 5.6	
23		2.0 M Sodium chloride, 10% (w/v) PEG 6000	
24		2.0 M Ammonium sulphate, 5% (v/v) propanol	
25	PEGRx	10% (v/v) 2-Propanol, 0.1 M BICINE pH 8.5, 30% PEG 1500	JEV capsid 6 mg/ml
26		0.1% (w/v) n-Octyl-β-D-glucoside, 0.1 M Sodium citrate tribasic dehydrate pH 5.5, 22% (w/v) PEG 3350	
27	PEGRx	18% 2-Propanol, 0.1 M Sodium citrate tribasic dehydrate pH 5.5, 20% (w/v) PEG 4000	
29	PACT *premier*	0.1 M SPG buffer pH 6, 25% (w/v) PEG 1500	
30		0.1 M SPG buffer pH 7, 25% (w/v) PEG 1500	
31		0.1 M SPG buffer pH 8, 25% (w/v) PEG 1500	
32		0.1 M SPG buffer pH 9, 25% (w/v) PEG 1500	

No.	Screens	Formulations	Protein concentration
33		0.1 M PCTP buffer pH 8, 25% (w/v) PEG 1500	JEV capsid 6 mg/ml
34	JCSG-*plus*	0.2 M Ammonium citrate dibasic, 20% (w/v) PEG 3350	
35		0.1 M BICINE pH 9, 10% (w/v) PEG 20000, 2% (v/v) 1,4-Dioxane	
36		0.1 M HEPES pH 7.5, 20% (v/v) Jeffamine® M-600	
37		0.1 M Potassium thiocyanate, 30% (w/v) PEG 2000 MME	

Note: The formulation used for data collections are shaded in grey.

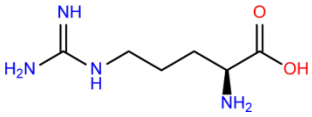
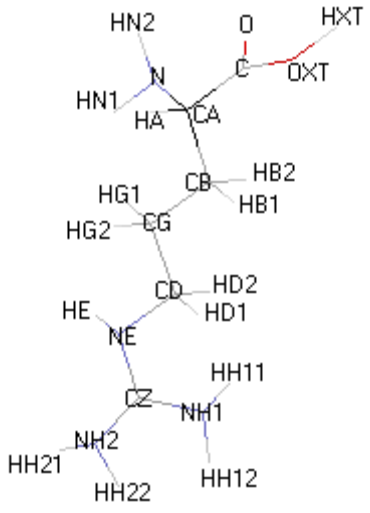
SPG = Succinic acid, Sodium phosphate monobasic monohydrate

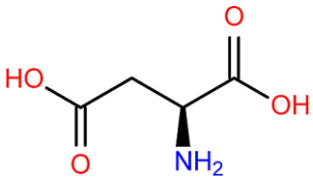
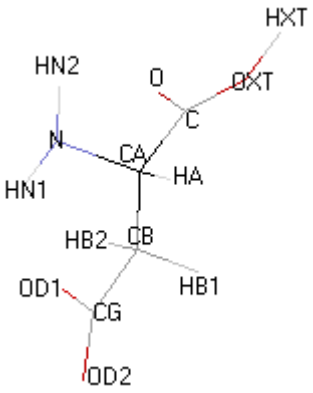
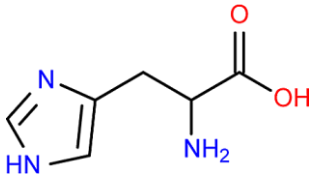
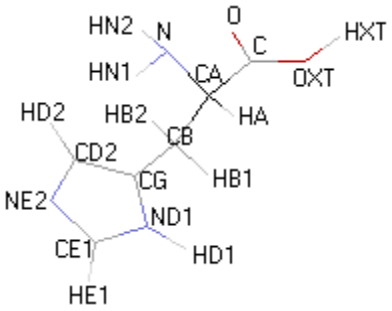
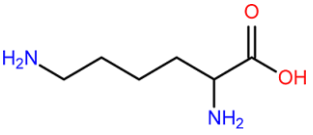
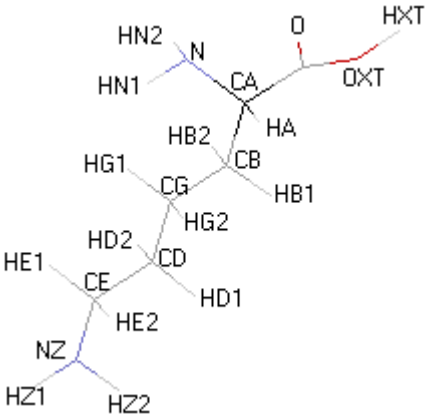
PCTP = Sodium propionate, Sodium cacodylate trihydrate, Bis-Tris propane

Appendix 4 Homologous sequences of NS1 C-terminus submitted to Consurf program

No.	Homologous sequence	No.	Homologous sequence
1	P06935 POLG_WNV_965_1139	12	P29991 POLG_DEV27_954_1127
2	Q32ZE1 POLG_ZIKV_968_1142	13	Q04538 POLG_POWVL_954_1128
3	Q89277 POLG_YEFVF_956_1130	14	P14335 POLG_KUNJM_969_1143
4	5IY3	15	Q1X880 POLG_YEFVU_956_1130
5	4OIE	16	JEV M55506
6	Q9WDA6 POLG_DEV2Q_954_1127	17	P05769 POLG_MVEV5_971_1145
7	4OIG	18	P07720 POLG_TBEVS_954_1128
8	P22338 POLG_LIV_954_1128	19	Q2YHF2 POLG_DEN4H_953_1126
9	P29837 POLG_LANVT_954_1128	20	Q6YMS4 POLG_DEN3S_952_1125
10	Q01299 POLG_TBEVH_954_1128	21	P33478 POLG_DEN1S_953_1126
11	P09732 POLG_STEVM_967_1141		

Appendix 5 Arginine, aspartic acid, histidine, and lysine atoms nomenclature

Amino acid	**Molecular formula**	**Atoms nomenclature (Elodie Foulquier, 2017)**
Arginine		

Amino acid	Molecular formula	Atoms nomenclature (Elodie Foulquier, 2017)
Aspartic acid		
Histidine		
Lysine		

Appendix 6 Mean values of melting temperatures calculated from experiments performed in duplicate and its ΔTm.

HS (µM)	**Tm1**	**Tm2**	**Tm3**	**Tm4**	**Tm5**	**Tm6**	**Mean Tm**	**±SD**	**ΔTm**
100	55.3	54.5	55.5	53.3	53.7	52.7	54.2	1.1	0.3
50	55.5	55.8	55.6	54.2	53.6	53	54.6	1.2	-0.1
25	55.4	55.6	55.8	54	54	53.8	54.8	0.9	-0.3
10	55.5	55.7	55.6	54.1	54.1	53.6	54.8	0.9	-0.3
5	55.6	55.5	55.5	54.2	54	53.8	54.8	0.9	-0.3
1	55.9	55.6	55.6	53.6	53.6	53.8	54.7	1.1	-0.2
0.5	56.2	55.5	55.6	53.3	53.7	53.3	54.6	1.3	-0.1
0	55.2	55.4	55.8	53.2	52.7	53.6	54.3	1.3	0.2
DS (µM)	**Tm1**	**Tm2**	**Tm3**	**Tm4**	**Tm5**	**Tm6**	**Mean Tm**	**±SD**	**ΔTm**
100	56.3	53.9	55.4	53.3	54.3	53.5	54.5	1.2	0.0
50	55.2	55.6	55.2	54.1	54.1	53.9	54.7	0.7	-0.2
25	56.6	55.8	56	54.2	54.5	54.1	55.2	1.1	-0.7
10	55.5	55.9	55.7	54.3	54.6	54.1	55.0	0.8	-0.5
5	54.8	55.4	54.8	54.4	54.6	54.1	54.7	0.4	-0.2

DS (µM)	Tm1	Tm2	Tm3	Tm4	Tm5	Tm6	Mean Tm	±SD	ΔTm
1	54.8	55.6	55.4	54.5	54.4	54.1	54.8	0.6	-0.3
0.5	56.7	56.6	55.9	54.5	53.6	53.6	55.2	1.4	-0.7
0	55.9	55.7	55.7	54.6	53.6	53.2	54.8	1.2	-0.3
CS (µM)	**Tm1**	**Tm2**	**Tm3**	**Tm4**	**Tm5**	**Tm6**	**Mean Tm**	**±SD**	**ΔTm**
100	53.7	53.7	55.2	52.7	51.5	52.4	53.2	1.3	1.3
50	54.5	54.5	54.9	53	52.8	53.5	53.9	0.9	0.6
25	55	54.7	55.1	53.3	53.5	53.8	54.2	0.8	0.3
10	55.2	55.6	55.7	53.5	52.1	53.2	54.2	1.5	0.3
5	55.3	55.8	56	51.8	52.3	53.2	54.1	1.9	0.4
1	55.6	55.6	55.6	53.9	53.1	53.5	54.6	1.2	-0.1
0.5	55.5	55.7	55.7	53.3	52.6	53.4	54.4	1.4	0.1
0	55.5	55.7	55.7	54	53.5	52.5	54.5	1.4	0.0

Note: The concentration of JEVNS1-C was kept constant at 10 µM. The average Tm of JEVNS1-C alone was 54.5°C.

The experiments were performed in total 6 replicates (3 replicates for each batch of protein purification).