

Contents lists available at [ScienceDirect](http://ScienceDirect)

## Virology

journal homepage: [www.elsevier.com/locate/yviro](http://www.elsevier.com/locate/yviro)

## Rapid Communication

## Structure of yellow fever virus envelope protein domain III

David E. Volk<sup>a,1,2</sup>, Fiona J. May<sup>b,1</sup>, Sai H.A. Gandham<sup>a</sup>, Anjenique Anderson<sup>b,3</sup>, Jana J. Von Lindern<sup>b</sup>, David W.C. Beasley<sup>c,e,f,g</sup>, Alan D.T. Barrett<sup>b,e,f,g</sup>, David G. Gorenstein<sup>a,d,\*</sup><sup>a</sup> Departments of Biochemistry and Molecular Biology, University of Texas Medical Branch, Galveston, TX 77555-1157, USA<sup>b</sup> Department of Pathology, University of Texas Medical Branch, Galveston, TX 77555-1157, USA<sup>c</sup> Microbiology and Immunology, University of Texas Medical Branch, Galveston, TX 77555-1157, USA<sup>d</sup> Sealy Center for Structural Biology and Molecular Biophysics, University of Texas Medical Branch, Galveston, TX 77555-1157, USA<sup>e</sup> Center for Biodefense and Emerging Infectious Diseases, University of Texas Medical Branch, Galveston, TX 77555-1157, USA<sup>f</sup> Institute for Human Infections and Immunity, University of Texas Medical Branch, Galveston, TX 77555-1157, USA<sup>g</sup> Sealy Center for Vaccine Development, University of Texas Medical Branch, Galveston, TX 77555-1157, USA

## ARTICLE INFO

## Article history:

Received 3 August 2009

Returned to author for revision

20 August 2009

Accepted 2 September 2009

Available online 8 October 2009

## Keywords:

Yellow fever virus

Envelope protein domain III

Neutralization site

Flavivirus

## ABSTRACT

The structure of recombinant domain III of the envelope protein (rED3) of yellow fever virus (YFV), containing the major neutralization site, was determined using NMR spectroscopy. The amino acid sequence and structure of the YFV-rED3 shows differences from ED3s of other mosquito-borne flaviviruses; in particular, the partially surface-exposed BC loop where methionine-304 and valine-324 were identified as being critical for the structure of the loop. Variations in the structure and surface chemistry of ED3 between flaviviruses affect neutralization sites and may affect host cell receptor interactions and play a role in the observed variations in viral pathogenesis and tissue tropism.

© 2009 Elsevier Inc. All rights reserved.

## Introduction

Yellow fever virus (YFV) is an arthropod-borne virus belonging to the family *Flaviviridae*, genus *Flavivirus*. Most flaviviruses are typically transmitted by either mosquitoes or ticks, and include major human pathogens such as YFV, dengue virus (DENV types 1–4), West Nile virus (WNV), Japanese encephalitis virus (JEV), and tick-borne encephalitis virus (TBEV). Yellow fever is an acute viral disease that causes hemorrhagic fever and jaundice. The virus is transmitted between humans by the *Aedes aegypti* mosquito and about 200,000 cases are reported annually, including 30,000 deaths. Because no treatment or cure exists for yellow fever, there is great interest in developing strategies to control the disease. Unlike other mosquito-borne flaviviruses, YFV has a tropism for the liver and causes a viscerotropic disease whereas many other mosquito-borne flaviviruses have a tropism for the brain, or in the case of the DEN viruses they target cells of reticuloendothelial origin.

The YFV genome is an 11 kb single-stranded positive-sense RNA genome coding for a polyprotein, which is post- and co-translationally processed into three structural proteins and seven non-structural proteins. The largest of the structural proteins, the envelope (E) protein, is the major component of the virion surface. It is the primary immunogen and plays a central role in receptor binding and membrane fusion (Heinz and Allison, 2003). The structure of the ectodomain (the soluble N-terminal portion, consisting of 395 residues) of the E protein of TBEV was determined by x-ray crystallography (Rey et al., 1995). Based on this structure, three distinct structural domains, domains I, II, and III, have been identified in the ectodomain. This structure has been confirmed by x-ray crystallographic studies of other flaviviruses, including DENV1 (Nayak et al., 2009), DENV2 (Modis et al., 2003), DENV3 (Modis et al., 2005), and WNV (Kanai et al., 2006; Nybakken et al., 2006). Domains I and II lie parallel to the virion surface in the mature, pre-fusion form. They contain the fusion peptide and the hinge region, both involved in the low-pH induced conformational change observed upon fusion and entry into the cell, and the N-linked glycosylation site(s) (Rey et al., 1995). Domain III (ED3) is involved in receptor binding and contains epitopes critical for type-specific neutralization of the virus (i.e., those neutralization epitopes that distinguish each flavivirus, e.g. YFV from DENV2) (Chu et al., 2005; Crill and Roehrig, 2001). The major neutralization epitopes of WNV (Beasley and Barrett, 2002; Nybakken et al., 2005; Sánchez et al., 2005), YFV (Ryman et al., 1998), DENV2

\* Corresponding author. The Brown Foundation Institute of Molecular Medicine, The University of Texas Health Science Center at Houston, 1825 Pressler St., Suite 205, Houston, TX 77030, USA. Fax: +1 713 500 2420.

E-mail address: [David.G.Gorenstein@uth.tmc.edu](mailto:David.G.Gorenstein@uth.tmc.edu) (D.G. Gorenstein).

<sup>1</sup> These authors contributed equally to this work.

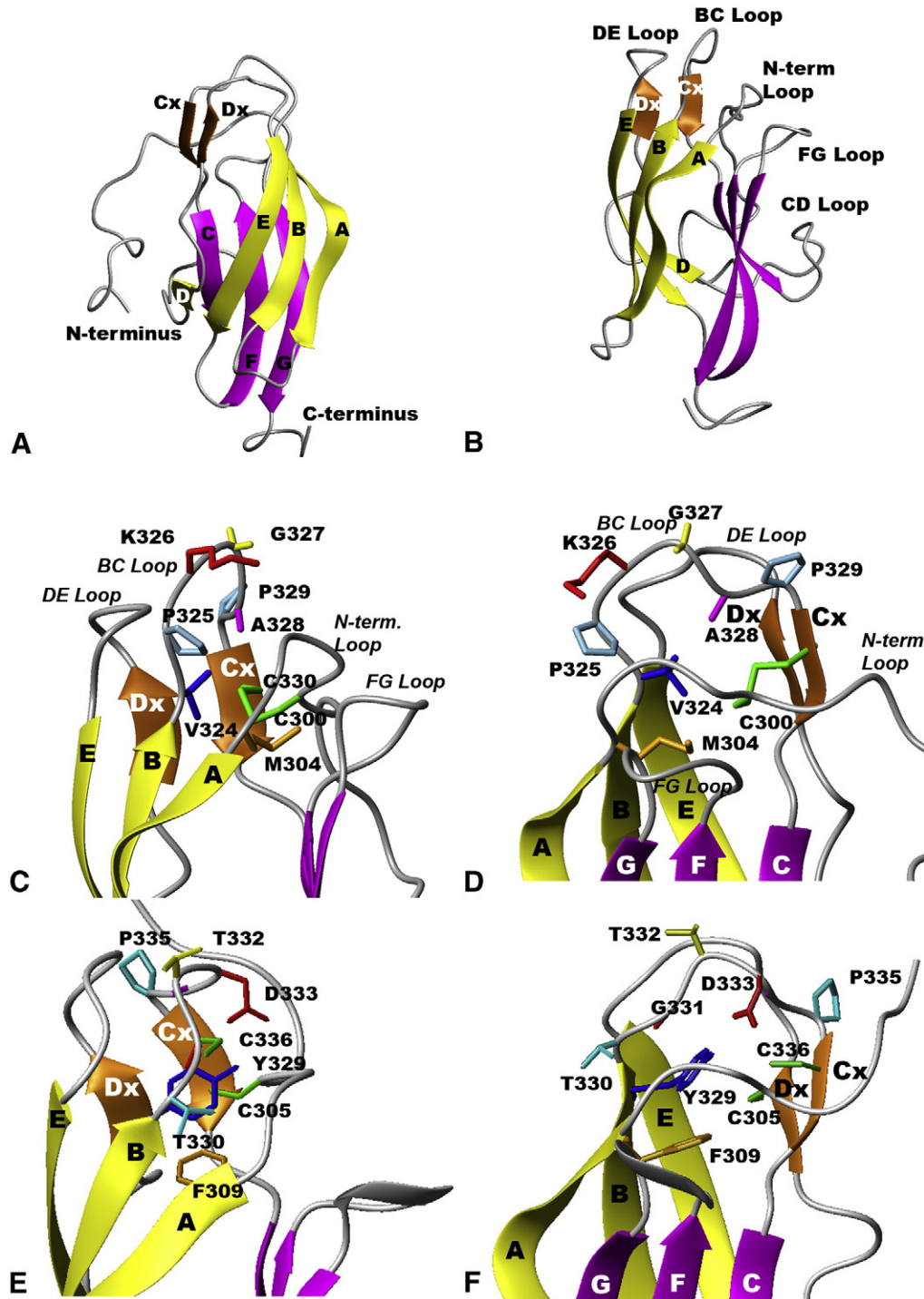
<sup>2</sup> Present address: The Brown Foundation Institute of Molecular Medicine, University of Texas Health Science Center at Houston, TX 77030, USA.

<sup>3</sup> Present address: Rollins School of Public Health, Emory University, Atlanta, GA 30322, USA.

(Hiramatsu et al., 1996; Roehrig et al., 1988; Gromowski and Barrett, 2007; Sukupolvi-Petty et al., 2007), TBEV (Mandl et al., 1989; Holzmann et al., 1997), and JEV (Cecilia and Gould, 1991; Wu and Lin, 2001; Lin and Wu, 2003; Wu et al., 1997, 2003, 2004; Goncalvez et al., 2008) have all been mapped to ED3.

Cryoelectron microscopic reconstructions of several flaviviruses indicate that the E protein is arranged as dimers parallel to the virion surface, such that ED3 projects slightly above the viral surface (Kuhn et al., 2002; Mukhopadhyay et al., 2003). Interactions between five ED3 subunits at the virion 5-fold axes of symmetry form pores on the

virion surface where cell receptors may bind. NMR-derived solution structures of the JEV (Wu et al., 2003), WNV (Volk et al., 2004), Omsk hemorrhagic fever virus (OHFV (Volk et al., 2006)), Langat virus (LGTV (Mukherjee et al., 2006)), and DENV4 (Volk et al., 2007b) rED3 illustrate an overall similar structural fold for this domain of these flaviviruses, with specific differences between those viruses transmitted by mosquito or tick vectors. In this study, we have solved the solution structure of rED3 of wild-type strain Asibi of YFV and demonstrate that it is markedly different from ED3 of other mosquito-borne flaviviruses that have been solved.



**Fig. 1.** Ribbon diagrams of rED3 of YFV and WNV.  $\beta$ -Sheets 1–3 are colored yellow, orange, and magenta, respectively, and the disulfide bridge between C300 and C330 is colored green. (A and B) Two orthogonal views of the NMR-derived YF rED3 backbone atom structures. Surface loop structure of (C and D) YF rED3 and (E and F) WNV rED3 (Volk et al., 2004).

**Table 1**  
Summary of NMR structure constraints and statistics.

Total restraints	1833
NOE restraints	1278
Intra-residue	609
Sequential	319
Medium range	59
Long range	291
TALOS phi/psi dihedral restraints	170
Omega dihedral restraints	111
Chirality restraints	274
Structural statistics	
NOE violations > 0.5 Å	3 ± 1
NOE violations > 0.3 Å	13 ± 2
Dihedral angle violation > 20°	0
Dihedral angle violation > 10°	2 ± 1
RMSD from ideal geometry	
Bond lengths (Å)	0.013
Bond angles (°)	2.1
Restraint error RMSD	
Distance restraints (Å)	0.019 ± 0.071
Dihedral restraints (°)	0.54 ± 1.51
Atomic pairwise RMSD	
Backbone atoms	1.38 ± 0.44
All heavy atoms	1.59 ± 0.41
Ramachandran statistics	
Most favored regions (%)	81.8
Additionally allowed regions (%)	16.9
Generously allowed regions (%)	1.1
Disallowed regions (%)	0.2

## Results

### Quality of the NMR structure

The 20 final structures of YFV-rED3 in the ensemble (Figs. 1A and B) had low molecular and restraint energy penalties. The structure presented here is well defined, as shown by the RMSD values and restraint violations listed in Table 1. The final structures, determined in an automated fashion, had  $13 \pm 2$  distance violations over 0.3 Å,  $3 \pm 1$  violation over 0.5 Å and  $2 \pm 1$  dihedral angle violations over 10°, and no dihedral angle violations over 20° (Table 1). Thus, 99.9% of the NMR-derived restraints fit the structures determined. Most of the violations occur in four or fewer of the twenty structures, although six are nearly always violated. The violations occur because the NOE interactions and cutoff distances were set in an automated fashion into distance spins based on cross-peak volumes, disregarding confounding effects such as amide proton exchange rates, equivalent geminal methyl groups, ambiguous NOE assignments, and differing spin-spin relaxation rates. The RMSD on the distance restraint error was  $0.019 \pm 0.071$  Å, and the RMSD on dihedral angle error was  $0.54 \pm 1.51$ °. The structural ensemble has an average pairwise backbone atom RMSD of  $1.38 \pm 0.44$  Å and an average pairwise heavy atom RMSD of  $1.59 \pm 0.41$  Å. The program PROCHECK was used to analyze the quality of the final ensemble. Analysis of the non-glycine, non-proline residues indicated that 98.7% of these residues are in the two most favored regions of a Ramachandran plot. Specifically, 81.8% of the residues are in the most favored regions, 16.9% are in the additionally allowed regions, 1.1% are in the generously allowed regions, and 0.2% are in the disallowed regions.

### Structural details of the NMR ensemble

The overall structure ensemble of YFV ED3 determined by NMR (Figs. 1A and B; Table 2, all amino acid numbers in the text refer to amino acid number in the specific viral E protein being discussed, unless otherwise defined) is similar to that reported for the ED3 of other flaviviruses, including DENV1 (Nayak et al., 2009), DENV2 (Modis et al., 2003), DENV3 (Modis et al., 2005), DENV4 (Volk et al., 2007b), JEV (Wu et al., 2003), LGTV (Mukherjee et al., 2006), OHFV (Volk et al., 2006), TBEV (Rey et al., 1995), and WNV (Volk et al., 2004). The YFV-rED3 structure has nine  $\beta$ -strands in three  $\beta$ -sheets arranged in an IgG-like  $\beta$ -barrel configuration. The first  $\beta$ -sheet (yellow) contains  $\beta$ -strands A from Ser305 to Asp312,  $\beta$ -strand B from Val318 to Lys323,  $\beta$ -strand D from Ile348 to Leu349, and  $\beta$ -strand E from Glu362 to Asn368. The second  $\beta$ -sheet (orange) is formed by only two short  $\beta$ -strands, Cx and Dx, encompassing residues Cys330-Lys331 and Ile355-Ala356, respectively. The last  $\beta$ -sheet (magenta) is comprised of  $\beta$ -strand C from Val334-Ala337,  $\beta$ -strand F from Gly372-Val378 and  $\beta$ -strand G from Leu385-Lys391. Both the overall global fold and the secondary structures of YFV-rED3 are grossly similar to the structures reported for mosquito-borne DENV1, DENV2, DENV3, DENV4, JEV, and WNV rED3, although small differences in the lengths of  $\beta$ -sheets do exist. However, the major difference between the YFV ED3 structure and other flavivirus structures is found at the surface-exposed loops; particularly in the BC loop. This difference is directly related to the addition of Pro325 in YFV (found in no other flavivirus; see Figs. 1C and D), and the presence of relatively small, non-aromatic residues at positions Met304 and Val324 of YFV ED3 (Figs. 1C and D and Table 2) compared to other mosquito-borne flaviviruses (see WNV ED3 in Figs. 1E and F).

The residues comprising the flavivirus BC loops differ significantly in mosquito and tick vectors and between flavivirus complexes (Table 2). All of the mosquito-borne flaviviruses, excluding the YFV complex, contain a conserved tyrosine immediately before the BC loop (amino acid position 329 for WNV in Table 2 and Figs. 1E and F), which has been shown to be essential for viability of WNV (Zhang, S. and Beasley, D.W.C., unpublished data) and presumably plays a role in stabilizing the ED3 protein fold while some of the tick-borne viruses have a phenylalanine substitution in place of the tyrosine. The phenylalanine at amino acid position 305 in the alignment (equivalent to F309 in WNV and M304 in YFV), which is packed closely with the tyrosine at position 329 in the WNV ED3 structure, is also conserved in these viruses. In contrast, YFV-rED3, as well as other YFV complex viruses (Wesslesbron [WSLV], Sepik [SEPV], Saboya [SABV], Jugra [JUGV], Edge Hill [EHV], Yokose [YOKV], and Entebbe Bat viruses [ENTV]), contain a methionine at position 304 and a valine at position 324. Immediately following the valine at 324, the BC loop of YFV both starts and ends with a proline (amino acids 325 and 329), whereas all other flaviviruses have BC loops ending with a proline only. The proline present at position 325 in the YFV E protein removes the need for a tyrosine or phenylalanine at position 324 by forcing the BC loop to start turning towards the next beta strand. The smaller sizes of Met304, relative to a phenylalanine, and Val324, relative to either a tyrosine or a phenylalanine, allow the length of the BC loop to be smaller in YFV and related viruses compared to the other mosquito-borne flaviviruses.

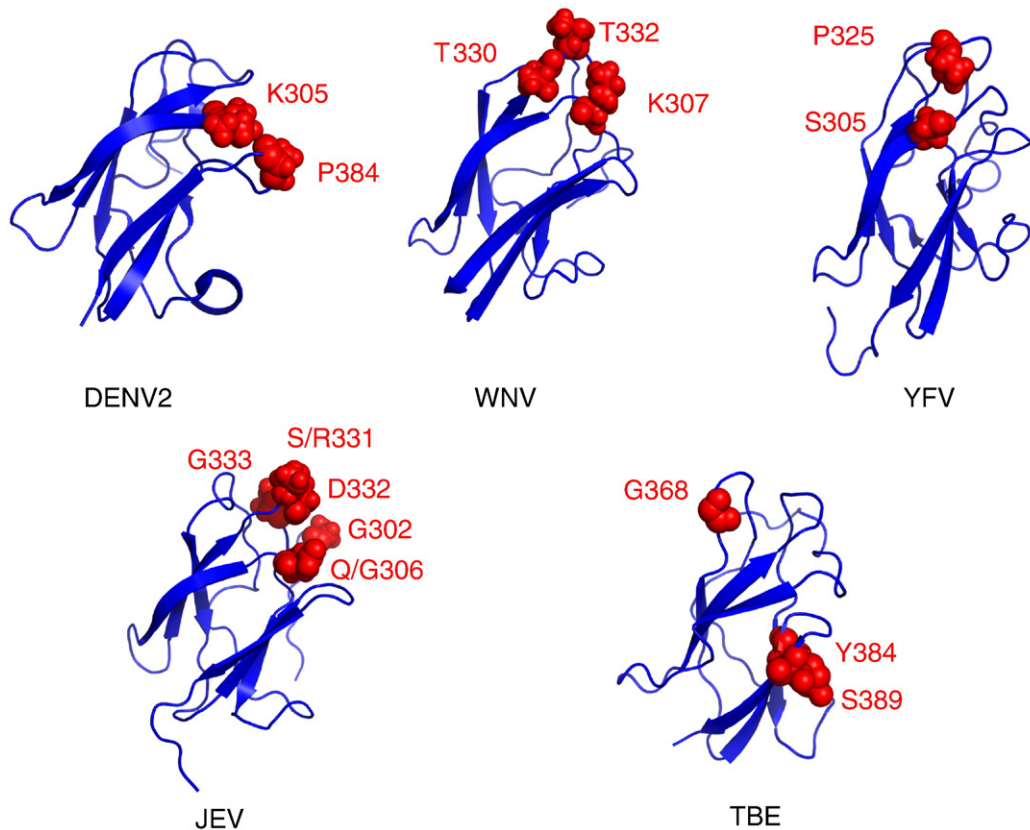
#### Notes to Table 2:

Asterisks (\*) indicate highly conserved residues and hashes (#) indicate residues with group-specific differences.  $\beta$ -Strands are indicated with underlined residues and are labeled underneath using TBEV (Rey et al., 1995) nomenclature. The number of the first amino acid of ED3 for each flavivirus is shown in superscript to the right of the name of each flavivirus. Those without available full length E genes sequences are left blank. Biochemically similar amino acids are the same color to allow easier understanding of the alignment.

<sup>1</sup>ICTV abbreviations used are as follows: yellow fever virus = YFV, Edge Hill virus = EHV, Jugra virus = JUGV, Saboya virus = SABV, Wesslesbron virus = WSLV, Sepik virus = SEPV, Entebbe bat virus = ENTV, Yokose virus = YOKV, West Nile virus = WNV, St Louis encephalitis virus = SLEV, Ntaya virus = NATV, Rocio virus = ROCV, Kokobera virus = KOKV, Bussuquara virus = BSQV, Iguape virus = IGUV, Zika virus = ZIKAV, Spondweni virus = SPOV, dengue virus = DENV, Kedougou virus = KEDV, tick-borne encephalitis virus = TBEV, Royal Farm virus = RFV, Kadam virus = KADV, Meaban virus = MEAV, Modoc virus = MODV, Apoi virus = APOIV, Montana myotis leukoencephalitis virus = MMLV, Rio Bravo virus = RBV.

**Table 2**  
Alignment of flavivirus ED3 proteins from mosquito-, tick-, and non-vector-borne flaviviruses.

	293	300	310	320	330	340								
<b>YFV complex</b>	YFV1 293 K G T S Y K M C T - D K M S F V K N P T D T G H G T V V M Q V K V P K G - - A P C K I P V I V A D D L T A	EHV K G S T Y T M C K - G G F S F V K T P T D T G H G T A V M Q V K V S K G - - T P C R I P V Q A V D S S N G	JUGV K G T T Y Q N C R - G G L S F T K T P A D T G H G T V V M Q V K V T K N - - T P C R L T A I A S D D A S G	SABV K G T T Y Q N C R - G G L S F T K T P A D T G H G T V V M Q V K V T K N - - A P C R L T A I A A D D A S G	WSLV 290 K G S T Y S M C K - R G M S F A K Q P V E T D H G T A V M Q I K V T T G - - A P C R I P V I A A D S M A G	SEPV 290 K G S T Y P M C K - K G M S F V K Q P V E T D H G T A V M Q V K V T N G - - A P C R I P V I A S D S M A G	ENTV 290 K G K T Y A M C R - G G Y S F S K T P V T S G H Q T V L M K V K V S K G - - T P C R I P V T M S D S L T V	YOKV 290 K G S T Y T M C K - G G Y S F S K T P V D S G H Q T V I M K V K V S K A - - T P C R I P V A V I D S M Q S						
<b>Mosquito-borne</b>	WNV 298 K G T T Y G V C S - K A F K F L G T P A D T G H G T V V L E L Q Y - T G T D G P C K V P I S S V A S L N D	SLEV 298 K G T T Y G M C D - S A F T F S K N P T D T G H G T V I V E L Q Y - T G S N G P C R V P I S V T A N L M D	NATV K G M T Y P M C S - N K F S L A R N P T D T G H G T V V V K L S Y - A G S D G P C R I P I S M T A N L Q D	ROCV 298 K G S T Y L M C K - D K F A F A K N P V D T G H G T I V T E V Q Y - A G S D G P C R I P I T M T E N L H D	KOKV 298 K G T T Y H M C K - G S F A F T K T P S D T G H G T V V L E L T Y - S G S D G P C R V P I S M S V S L S N	BSQV 300 K G I T Y G Q C S - G T F A K M E K H P A D T G H G T V V L D V S Y - Q G D D A P C K I P I V I T S N L A E	IGUV 293 K G T T Y H M C A - K A F T M K K D P T D T G H G T V V M E L T Y - K G I D V P C R V P I T I A R S P N D	ZIKAV 297 K G V S Y S L C T - A A F T F T K V P A E T L H G T V T V E V Q Y - A G T D G P C K I P V Q M A V D M Q T	SPOV K G M S Y A L C T - G A F T F A R T P S E T I H G T A T V E L Q Y - A G E D G P C K V P I V I T T D T N S	DENV1 295 K G M S Y V M C T - G S F K L E K E V A E T Q H G T V L V Q V K Y - E G T D A P C K I P F S T Q D D K G A	DENV2 295 K G M S Y S M C T - G K F K V V K E I A E T Q H G T I V I R V Q Y - E G D G S P C K I P F E I M D L E K R	DENV3 293 K G M S Y A M C T - N T F V L K K E V S E T Q H G T I L I K V E Y - K G E D A P C K I P F S T E D G Q G K	DENV4 295 K G M S Y T M C S - G K F S I D K E M A E T Q H G T T V V K V K Y - E G A G A P C K V P I E I R D V N K E	KEDV 301 R G V S Y A M C G - G K F S F H R N P A P T Q H G T V T V D I G Y - S G - D A P C K V P I S V S S E A N S
<b>Tick-borne</b>	TBEV 300 K G L T Y T M C D K T K F T W K R A P T D S G H D T V V M E V T F - S G T - K P C R I P V R A V A H G S P	RFV 300 K G I T Y S M C E S G K F S W K R P P T D S G H D T V V M E L S Y - S G A T K P C R I P V M A T A H G E E	KADV 294 V G M T Y S A C E S S K F T W K Q T P R D S A H D T V V M K L A Y - T G T - K P C R A L V R A Y R P G A E	MEAV 293 R G L T Y G M C A V G D F S W K R V P T D S Q H D T V V M E V T Y - T G S S T P C R I P V R A Y H P G T P										
<b>Non-vector-borne</b>	MODV 292 K G M T Y V V C G - G K F A W A K K P I A T N H D T V A M E V T Y - T G N D T P C R V T V K N V K E N S D	APOIV 293 V G A T Y S Q C T - K P F E W I K K P V L T Q H G T V V M E V K Y - T G E G A P C R I P F R V E R V D K P	MLLV 292 K G T T Y P Y C G - D S F V W K R R P V A T H H G T V A M E V T Y - Q G T D V P C K V S V I V E K D G Q N	RBV 292 K G L T Y Q M C S - S S F V W H K R P V A T Q H G T V A M E V K Y - K G S D A P C R I P V S V E K E G Y N										
	* * * * *	* # # # #	* # * * *	* # #	* # * * *	* # #								
	A <sub>x</sub>	A	B	C <sub>x</sub>	C									
<b>YFV complex</b>	YFV A I N K G I L V T V N P I A S T N D D - - E V L I E V N - P P F G D S Y I I V G T G D S R L T Y Q W H K	EHV G T N R A T L I T A N P I A A T T E D - - K V M I E L S - P P Y G E S Y I M I G T G D D K L T Y H W L K	JUGV R V N R G T L V T S N P V A N S A N D - - E V L I E I N - P P Y G E S Y L I A G V G D D K L V Y Q W F Q	SABV K V N R G T L V T S N P I A N A A N D - - E V L I E I N - P P F G E S Y L I V G T G D D K L V Y Q W K K	WSLV T E N R G S V I T T N P I A A S N N D - - E V L V E I S - P P F G E S Y I I V G N G D D K L T Y H W Q R	SEPV T E N R G S V I T T N P I A A L N N D - - E V L V E I S - P P F G E S Y I I V G S G D D K L T Y H W Q R	ENTV T K N Q G V I V T T N P I A F D A N E - - V L I E V I - P P F G D S H I I I G N G E D R L T H R W H Q	YOKV N I N R G V V V T T N P V A F E A A T - - E V M I E V V - P P F G E S V I T I G N G E D R L T Y Q W H Q						
<b>Mosquito-borne</b>	WNV L T P V G R L V T V N P F V S V A T A N A K V L I E L E - P P F G D S Y I V V G R G E Q Q I N H H W H K	SLEV L T P V G R L V T V N P F I S T G G A N N K V M I E V E - P P F G D S Y I V V G R G T T Q I N Y H W H K	NATV L T P I G R M I T V N P Y V S T S S T G T K V I E L E - P P F G D S F I L V G S G E N Q I K Y Q W H K	ROCV L T P I G R L V T V N P F V P S S E T A Q K I L I E L E - P P F G T S F I L V G T G P N Q V K Y Q W H K	KOKV I E P V G R M V T V N P I V L S S S P Q K T I M I E V E - P P F G D S F I I A G T G E P R A H Y H W R K	BSQV V E P V G R L V S A H P V I T A K N V - - R T M L E V E - P P Y G D S Y I I V I G G D R L K Q H W F K	IGUV G E M V G R M V S V N P L A M T T S S - - V F M V E V E - P P Y G D S N I I V G S Y D N V L K H H W F K	ZIKAV L T P V G R L I T A N P V I T E S T E N S K M M L E L D - P P F G D S Y I V I G V G D K K I T H H W H R	SPOV M A S T G R L I T A N P V V T E S G A N S K M M V E I D - P P F G D S Y I I V G T G T K I T H H W H R	DENV1 T Q - N G R L I T A N P I V T D K E K - - P V N I E A E - P P F G E S Y I V V G A G E K A L K L S W F K	DENV2 H V - L G R L I T V N P I V T E K D S - - P V N I E A E - P P F G D S Y I I I G V E P G Q L K L N W F K	DENV3 A H - N G R L I T A N P V V T K K E E - - P V N I E A E - P P F G E S N I V I G I G D K A L K I N W Y R	DENV4 K V - V G R I I S S T P L A E N T N S - - V T N I E L E - P P F G D S Y I V I G V G N S A L T L H W F R	KEDV H K N V G R L V T A N P I V M K N G D - - S V L V E V E - P P F G D S Y I V V G T G P T K I N Y H W Y K
<b>Tick-borne</b>	TBEV D V N V A M L I T P N P T I E N N G G - - - G F I E M Q L P P - G D N I I Y V G E - - - L S H Q W F Q	RFV S N - V A M L I T S N P T I E T D K G - - - G F I E M Q V P P - G D I T I K I G D - - - L K Q Q W F Q	KADV T L D V A K L I T S N P I C T N D M T - - D L F V E M Q V P P - G D T I I A V G D - - - L R F Q W F Q	MEAV E K D V A S V I T A N P V V E S T H V K D - I F I E M Q L P P - G D N V I A V G S - - - L R Y Q W F Q										
<b>Non-vector-borne</b>	MODV D Q - G T L I T T N P F V E S N G A - - T I F L E L E - P V Y G L S T I K V G D - - - I T Y Q W N Q	APOIV M E N V G N L V T G N P Y A S Q K D A - - V V F L E A E V P P - G I S I K I G D - - - I D V Q W N Q	MLLV G G N A G S L I T S N P I T A Q G S - - S V F L E L E V P L - G F S T I K V G A - - - A K Q Q W R Q	RBV G K N F G N L I T A N P F A A N N E A - - V V F L E L E A P L - G V S T I K V G G - - - A V F Q W K Q										
	# * * * *	# * * * *	# * * # * *	* * * * *	* # # # #	*								
	D	D <sub>x</sub>	E	F	G									



**Fig. 2.** Neutralizing epitopes on ED3 of yellow fever (YF), dengue-2 (DENV2), Japanese encephalitis (JE), tick-borne encephalitis (TBE), and West Nile (WN) viruses. Red dots identify amino acids recognized by type-specific monoclonal antibodies.

These differences in loops are unique in the YFV complex viruses and would be predicted to contribute to differences in antigenicity and differences in the individual amino acids that constitute the major type-specific neutralization epitopes on different flaviviruses. In particular, the major neutralization epitope in YFV involves the serine at residue 305 and proline at residue 325 (Ryman et al., 1998), while it is the lysine at residue 307, the threonine at residue 330, and threonine at residue 332 for WNV (Beasley and Barrett, 2002; Nybakken et al., 2005); the lysine at residue 305 and proline at residue 384 for DENV2 (Hiramatsu et al., 1996; Gromowski and Barrett, 2007; Sukopolvi-Petty et al., 2007); the glycine at residue 302, glutamine or glycine at residue 306, serine or arginine at residue 331, aspartic acid at residue 332, and glycine at residue 333 for JEV (Cecilia and Gould, 1991; Wu and Lin, 2001; Lin and Wu, 2003; Goncalvez et al., 2008), and the glycine at residue 368, tyrosine at residue 384, and serine at residue 389 for TBEV (Mandl et al., 1989; Holzmann et al., 1990). Based on this information, Fig. 2 shows that the location of the type-specific epitopes associated with neutralization for WNV, JEV, TBEV, DENV2 and YFV viruses are not in identical locations on ED3. (i.e., YFV: 305 and 325 [B-C loop], DENV2: 305 and 284 [F-G loop], WNV: 310 and 332 [B-C loop], JEV: 302, 306, 331, 332, and 333 [B-C loop], TBEV: 384 and 389). Thus, different surface-exposed loops on ED3 of different flaviviruses are important for neutralizing epitopes.

## Discussion

Identification of structural and/or amino acid differences in ED3 has revealed differences in the critical type-specific neutralization epitopes that is leading to a greater understanding of how each flavivirus is distinguished immunologically. The newly determined structure of the ED3 of YFV, representing a major flavivirus

serocomplex not previously subjected to detailed structural analysis, was compared with the structure of the ED3s of other flaviviruses.

The structure of the ED3 of YFV differs from the structures of other mosquito-borne flaviviruses; in particular, the surface-exposed loops, especially the BC loop, are different (see Figs. 1C–F). In YFV, the BC loop is one amino acid shorter than in the mosquito-borne and non-vector-borne viruses, but the same length as most tick-borne viruses. Although the function of the BC loop is unknown, it contains the major neutralization determinant for YFV (residue 325), WNV (residue 332; Beasley and Barrett, 2002; Nybakken et al., 2005), and JEV (residue 333; Wu and Lin, 2001). The structures in Figs. 1C–F suggest that not all flavivirus type-specific neutralization epitopes are in analogous positions, which supports the hypothesis that the function of the BC loop may be different for at least YFV, WNV, and JEV. In addition, a nearby loop, the FG loop, located on the same surface of ED3, has been shown to be the major neutralization determinant for TBEV and DENV2 (see Fig. 2) and is involved in vector-specific receptor binding of DENV2 (Hung et al., 2004). Like the BC loop, the FG loop is longer in most mosquito-borne viruses than the tick-borne viruses. The differences in this loop, in combination with other variations in surface chemistry, most likely contribute to the diversity in antigenicity, and possibly receptor binding and host specificity and tissue tropism.

The overall structure of the ED3s of most mosquito-borne flaviviruses, including DENV1 (Nayak et al., 2009), DENV2 (Modis et al., 2003), DENV4 (Volk et al., 2007b), WNV (Volk et al., 2004), and JEV (Wu et al., 2003) are very similar, and comparison of the amino acid sequences reveals several motifs unique to these virus complexes. The same is true of the tick-borne viruses such as TBEV (Rey et al., 1995), LGTV (Mukherjee et al., 2006), and OHFV (Volk et al., 2006) (see Table 2). In contrast, the structure of the YFV ED3 has several

unique differences when compared with other mosquito-borne flaviviruses; in particular, the surface-exposed BC loop is shorter in YFV than any other mosquito-borne virus. These differences are reflected in the amino acid sequence of this region, and due to a high level of similarity of the amino acid sequence of ED3 between members of the YFV complex (see Table 2), these structural differences can be expected to occur in other members of the YFV complex.

## Materials and methods

### Protein expression and purification

Uniformly  $^{15}\text{N}$ ,  $^{13}\text{C}$ -labeled human YFV-rED3 protein (Asibi strain) encompassing residues (Ser288-Lys398) was expressed using the pET-15b vector (Novagen), with an added methionine residue on the N-terminus but lacking the N-terminal His-tag sequence encoded in that plasmid. The cells were lysed using the native lysis buffer and centrifuged to obtain the YFV-rED3 protein along with the crude cell debris in the pellet. The pellet is then dissolved in denaturing lysis buffer containing 6 M guanidine hydrochloride to solubilize the protein. The insoluble cell debris is removed by centrifugation. The guanidine HCl in the supernatant was then removed by dialysis ( $6\times 1:2000$  dilution). The expressed protein was filtered through an Amicon centrifugal filter concentrator with a 50 kDa molecular weight cutoff to remove proteins with higher molecular weight. Size exclusion chromatography was performed using Sephadex G-75 beads to further purify the protein. Centricon concentrators with a 3 kDa cutoff membrane were used for the final concentration step and to remove low-molecular weight impurities, as well as to exchange the material into the final NMR buffer.

### NMR spectroscopy and the generation of NMR restraints

The NMR samples contained 0.1–0.4 mM protein in 50 mM deuterated Tris (pH 5.8), 50 mM NaCl, 1 mM  $\text{NaN}_3$  in 90%  $\text{H}_2\text{O}$ , and 10%  $\text{D}_2\text{O}$ . NMR experiments were performed at 25 °C on Varian Inova 750 MHz (UTMB) or 600 MHz (with cold probe, Rice University) spectrometers with triple resonance probes. The  $^{13}\text{C}$  and  $^{15}\text{N}$  dimensions were referenced indirectly using frequency ratios. Sequence-specific backbone assignments were obtained using the 2D  $^1\text{H}$ ,  $^{15}\text{N}$ -HSQC, 3D HNCACB, 3D CBCA(CO)NH and 3D HNCO experiments as described previously (Volk et al., 2007a). Non-aromatic side chain assignments were obtained using the HCCH-TOCSY, TOCSY- $^1\text{H}$ ,  $^{15}\text{N}$ -HSQC, H(C)CH-TOCSY, H(CCO)NH, and C(CO)NH experiments as described previously (Volk et al., 2007a). Aromatic proton assignments were obtained from the (HB)CB(CGCD)HD and (HB)CB(CGCDCE)HE experiments. A NOESY- $^1\text{H}$ ,  $^{15}\text{N}$ -HSQC experiment provided several missing side chain assignments as described previously (Volk et al., 2007a). Stereo-specific assignments for some of the side chain protons were obtained after initial rounds of structure calculations using ambiguous restraints. The NMR spectra were processed in VNMRJ (Varian, Inc.) or Felix2000 (Felix, Inc.) software. SANE was used to facilitate the assignment of the  $^{15}\text{N}$ -edited or  $^{13}\text{C}$ -edited NOE cross-peaks and for the generation of restraints as described previously (Volk et al., 2007a). Chemical shifts, distance cutoffs, and contribution cutoffs were used within the program. The NMR restraints were separated into four bins, based on the NOESY cross-peak volumes from which they were derived, with upper distance limits of 2.5, 3.5, 4.5, and 6.0 Å for all NOE data. The 1833 NOE-based restraints (see Table 1) consist of 609 intra-residue, 319 sequential, 59 medium-range, and 291 long-range distance restraints. TALOS was used to derive 170 phi/psi dihedral angle restraints based on the chemical shifts of the amino acids. Additional angular restraints for the omega angles and correct chiralities were generated within AMBER6.0.

### Molecular dynamics calculations

One hundred random structures were generated by annealing the protein at 700 K, obtaining the coordinates every 5 ps and minimizing the structures obtained. The structures were then subjected to r-MD using dihedral angle restraints (Table 1) followed by the application of all restraints at 300 K. Finally, the structures were energy minimized for 2000 steps. Twenty structures with the lowest restraint penalties were then chosen for the structural ensemble. The SANDER module within AMBER6.0 (Case et al., 1999) was used for all NMR structure calculations, and MIDAS (Ferrin et al., 1988) and MOLMOL (Koradi, Billeter, and Wuthrich, 1996) were used to visualize the structures. Coordinates for the ensemble of NMR structures of YFV-rED3 have been deposited with the Protein Data Bank (PDB ID 2JQM) and the chemical shifts have been deposited with the BMRB (Volk et al., 2007a; accession code 15034).

### Acknowledgments

This work was supported by the Pediatric Dengue Vaccine Initiative, the Centers for Disease Control (U90CCU618754), the NIAID (U01 AI054827), and the Welch Foundation (H1296). We thank Sean Moran for assistance with the Rice University Varian Inova 600 MHz NMR spectrometer whose Carbon-Enhanced HCN Cold Probe was obtained through the SPRING (Strategic Partnership for Research in Nanotechnology Grant) AFRL (AFOSR) FA9550-04-1-0328.

The atomic coordinates (2JQM) and NMR restraints (2JV6) for YFV-rED3 were deposited in the Protein Data Bank, Research Collaboratory for Structural Bioinformatics, Rutgers University, New Brunswick, NJ (<http://www.rcsb.org/>). NMR chemical shifts (15034) have been deposited at the BioMagResBank, University of Wisconsin-Madison (<http://www.bmrb.wisc.edu/>). The sequence of the WSLV has been deposited with GenBank (accession number EU075555).

### References

- Beasley, D.W., Barrett, A.D., 2002. Identification of neutralizing epitopes within structural domain III of the West Nile virus envelope protein. *J. Virol.* 76 (24), 13097–13100.
- Case, D.A., Pearlman, D.A., Caldwell, J.W., Cheatham, T.E., III, Ros, W.S., Simmerling, C.L., Darden, T.A., Merz, K.M., Standon, R.V., Cheng, A.L., Vincent, J.J., Crowley, M., Tsui, V., Radmer, R.J., Duan, Y., Pitera, J., Massova, I., Seibel, G.L., Singh, U.C., Weiner, P.K., Kollman, P.A. 1999. AMBER6, University of California, San Francisco.
- Cecilia, D., Gould, E.A., 1991. Nucleotide changes responsible for loss of neuroinvasiveness in Japanese encephalitis virus neutralization-resistant mutants. *Virology* 181, 70–77.
- Chu, J.J., Rajamanonmani, R., Li, J., Bhuvanankantham, R., Lescar, J., Ng, M.L., 2005. Inhibition of West Nile virus entry by using a recombinant domain III from the envelope glycoprotein. *J. Gen. Virol.* 86 (Pt 2), 405–412.
- Crill, W.D., Roehrig, J.T., 2001. Monoclonal antibodies that bind to domain III of dengue virus E glycoprotein are the most efficient blockers of virus adsorption to Vero cells. *J. Virol.* 75 (16), 7769–7773.
- Ferrin, T.E., Huang, C.C., Jarvis, L.E., Langridge, R., 1988. The MIDAS display system. *J. Mol. Graph.* 6 (1), 13–27.
- Goncalves, A.P., Chien, C.H., Tubthong, K., Gorshkova, I., Roll, C., Donau, O., Schuck, P., Yoksan, S., Wang, S.D., Purcell, R.H., Lai, C.J., 2008. Humanized monoclonal antibodies derived from chimpanzee Fabs protect against Japanese encephalitis virus in vitro and in vivo. *J. Virol.* 82 (14), 7009–7021.
- Gromowski, G.D., Barrett, A.D.T., 2007. Characterization of an antigenic site that contains a dominant, type-specific, neutralization determinant on the envelope protein domain III (ED3) of dengue 2 virus. *Virology* 366, 349–360.
- Heinz, F.X., Allison, S.L., 2003. Flavivirus structure and membrane fusion. *Adv. Virus Res.* 59, 63–97.
- Hiramatsu, K., Tadano, M., Men, R., Lai, C.J., 1996. Mutational analysis of a neutralization epitope on the dengue type 2 virus (DEN2) envelope protein: monoclonal antibody resistant DEN2/DEN4 chimeras exhibit reduced mouse neurovirulence. *Virology* 224 (2), 437–445.
- Holzmann, H., Heinz, F.X., Mandl, C.W., Guirakhoo, F., Kunz, C., 1990. A single amino acid substitution in envelope protein E of tick-borne encephalitis virus leads to attenuation in the mouse model. *J. Virol.* 64, 5156–5159.
- Holzmann, H., Stiasny, K., Ecker, M., Kunz, C., Heinz, F.X., 1997. Characterization of monoclonal antibody-escape mutants of tick-borne encephalitis virus with reduced neuroinvasiveness in mice. *J. Gen. Virol.* 78 (Pt 1), 31–37.

- Hung, J.J., Hsieh, M.T., Young, M.J., Kao, C.L., King, C.C., Chang, W., 2004. An external loop region of domain III of dengue virus type 2 envelope protein is involved in serotype-specific binding to mosquito but not mammalian cells. *J. Virol.* 78 (1), 378–388.
- Kanai, R., Kar, K., Anthony, K., Gould, L.H., Ledizet, M., Fikrig, E., Marasco, W.A., Koski, R.A., Modis, Y., 2006. Crystal structure of West Nile virus envelope glycoprotein reveals viral surface epitopes. *J. Virol.* 80 (22), 11000–11008.
- Koradi, R., Billeter, M., and Wuthrich, K., (1996). MOLMOL: a program for display and analysis of macromolecular structures. *J. Mol. Graph.* 14(1), 51-5, 29-32.
- 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 (5), 717–725.
- Lin, C.W., Wu, S.C., 2003. A functional epitope determinant on domain III of the Japanese encephalitis virus envelope protein interacted with neutralizing-antibody combining sites. *J. Virol.* 77 (4), 2600–2606.
- Mandl, C.W., Guirakhoo, F., Holzmann, H., Heinz, F.X., Kunz, C., 1989. Antigenic structure of the flavivirus envelope protein E at the molecular level, using tick-borne encephalitis virus as a model. *J. Virol.* 63 (2), 564–571.
- Modis, Y., Ogata, S., Clements, D., Harrison, S.C., 2003. A ligand-binding pocket in the dengue virus envelope glycoprotein. *Proc. Natl. Acad. Sci. U. S. A.* 100 (12), 6986–6991.
- Modis, Y., Ogata, S., Clements, D., Harrison, S.C., 2005. Variable surface epitopes in the crystal structure of dengue virus type 3 envelope glycoprotein. *J. Virol.* 79 (2), 1223–1231.
- Mukherjee, M., Dutta, K., White, M.A., Cowburn, D., Fox, R.O., 2006. NMR solution structure and backbone dynamics of domain III of the E protein of tick-borne Langat flavivirus suggests a potential site for molecular recognition. *Protein Sci.* 15 (6), 1342–1355.
- Mukhopadhyay, S., Kim, B.S., Chipman, P.R., Rossmann, M.G., Kuhn, R.J., 2003. Structure of West Nile virus. *Science* 302 (5643), 248.
- Nayak, V., Dessau, M., Kucera, K., Anthony, K., Ledizet, M., Modis, Y., 2009. Crystal structure of dengue virus type 1 envelope protein in the postfusion conformation and its implications for membrane fusion. *J. Virol.* 83 (9), 4338–4344.
- Nybakken, G.E., Oliphant, T., Johnson, S., Burke, S., Diamond, M.S., Fremont, D.H., 2005. Structural basis of West Nile virus neutralization by a therapeutic antibody. *Nature* 437 (7059), 764–769.
- Nybakken, G.E., Nelson, C.A., Chen, B.R., Diamond, M.S., Fremont, D.H., 2006. Crystal structure of the West Nile virus envelope glycoprotein. *J. Virol.* 80 (23), 11467–11474.
- Rey, F.A., Heinz, F.X., Mandl, C., Kunz, C., Harrison, S.C., 1995. The envelope glycoprotein from tick-borne encephalitis virus at 2 Å resolution. *Nature* 375 (6529), 291–298.
- Roehrig, J.T., Bolin, R.A., Kelly, R.G., 1988. Monoclonal antibody mapping of the envelope glycoprotein of the dengue 2 virus, Jamaica. *Virology* 246 (2), 317–328.
- Ryman, K.D., Ledger, T.N., Campbell, G.A., Watowich, S.J., Barrett, A.D., 1998. Mutation in a 17D-204 vaccine substrain-specific envelope protein epitope alters the pathogenesis of yellow fever virus in mice. *Virology* 244 (1), 59–65.
- Sánchez, M.D., Pierson, T.C., McAllister, D., Hanna, S.L., Puffer, B.A., Valentine, L.E., Murtadha, M.M., Hoxie, J.A., Doms, R.W., 2005. Characterization of neutralizing antibodies to West Nile virus. *Virology* 336 (1), 70–82.
- Sukupolvi-Petty, S., Austin, S.K., Purtha, W.E., Oliphant, T., Nybakken, G.E., Schlesinger, J.J., Roehrig, J.T., Gromowski, G.D., Barrett, A.D., Fremont, D.H., Diamond, M.S., 2007. Type- and subcomplex-specific neutralizing antibodies against domain III of dengue virus type 2 envelope protein recognized adjacent epitopes. *J. Virol.* 81, 12816–12826.
- Volk, D.E., Beasley, D.W., Kallick, D.A., Holbrook, M.R., Barrett, A.D., Gorenstein, D.G., 2004. Solution structure and antibody binding studies of the envelope protein domain III from the New York strain of West Nile virus. *J. Biol. Chem.* 279 (37), 38755–38761.
- Volk, D.E., Chavez, L., Beasley, D.W., Barrett, A.D., Holbrook, M.R., Gorenstein, D.G., 2006. Structure of the envelope protein domain III of Omsk hemorrhagic fever virus. *Virology* 351 (1), 188–195.
- Volk, D.E., Gandham, S.H.A., May, F.J., Anderson, A., Barrett, A.D., Gorenstein, D.G., 2007a. NMR assignments of the yellow fever virus envelope protein domain III. *Biomol. NMR Assign.* 1, 49–50.
- Volk, D.E., Lee, Y.C., Li, X., Thiviyanathan, V., Gromowski, G.D., Li, L., Lamb, A.R., Beasley, D.W., Barrett, A.D., Gorenstein, D.G., 2007b. Solution structure of the envelope protein domain III of dengue-4 virus. *Virology* 364 (1), 147–154.
- Wu, S.C., Lin, C.W., 2001. Neutralizing peptide ligands selected from phage-displayed libraries mimic the conformational epitope on domain III of the Japanese encephalitis virus envelope protein. *Virus Res.* 76 (1), 59–69.
- Wu, S.C., Lian, W.C., Hsu, L.C., Liao, M.Y., 1997. Japanese encephalitis virus antigenic variants with characteristic differences in neutralization resistance and mouse virulence. *Virus Res.* 51 (2), 173–181.
- Wu, K.P., Wu, C.W., Tsao, Y.P., Kuo, T.W., Lou, Y.C., Lin, C.W., Wu, S.C., Cheng, J.W., 2003. Structural basis of a flavivirus recognized by its neutralizing antibody: solution structure of the domain III of the Japanese encephalitis virus envelope protein. *J. Biol. Chem.* 278 (46), 46007–46013.
- Wu, S.C., Lin, Y.J., Chou, J.W., Lin, C.W., 2004. Construction and characterization of a Fab recombinant protein for Japanese encephalitis virus neutralization. *Vaccine* 23 (2), 163–171.