The Dissertation Committee for Rodrigo A. Maillard Certifies that this is the approved version of the following dissertation:

THE MOLECULAR BASIS FOR EVASION OF ANTIBODY-MEDIATED NEUTRALIZATION IN FLAVIVIRUSES

Committee:

James C. Lee, PhD, Supervisor

Vince J. Hilser, PhD

Andres F. Oberhauser, PhD

Alan D. T. Barrett, PhD

Wlodzimierz Bujalowski, PhD

Hiram F. Gilbert, PhD

_____

Dean, Graduate School

THE MOLECULAR BASIS FOR EVASION OF ANTIBODY-MEDIATED NEUTRALIZATION IN FLAVIVIRUSES

by

Rodrigo A. Maillard

Dissertation

Presented to the Faculty of the Graduate School of

The University of Texas Medical Branch

in Partial Fulfillment

of the Requirements

for the Degree of

DOCTOR OF PHILOSOPHY

The University of Texas Medical Branch

March, 2007

Dedication

To the memory of my mother, Irene Andrea Reyes de Maillard

To my wife, Maria Fe Lanfranco, for her constant support, encouragement and love

Acknowledgements

I would like to express my gratitude to my mentor and friend, Dr. James C. Lee. Dr. Lee has been an excellent tutor, research adviser and role model in this initial stage of my scientific training, and he will always be. Also, I want to extend my immense appreciation to present and former members of the Lee lab; especially to Matthew Jordan, Keerthi Gottipati, James Knapp, Alex Gribenko, Lucy Lee, Andrea Garces and Shaoning Yu.

For their constant guidance and scientific challenges I want to thank Dr. Vincent Hilser, Dr. Alan Barrett and Dr. Wlodek Bujalowski; also to Dr. Andres Oberhauser and Dr. Hiram Gilbert. All of them were part of my dissertation committee. For their support and encouragement I want to thank Dr. Wayne Bolen and Dr. Lillian Chan.

At many instances during my dissertation research Dr. David Beasley, Dr. Joerg Roesgen, Greg Gromowski, Dr. Steve Whitten, Maria Jezewska, Aaron Lucius were of crucial importance. Thank you to them too.

Research work cannot be done without the help of the administrative staff in the department, laboratory and graduate program. Thanks to Shirley Broz, Debora Botting, Karen Jones and Angelina Johnson. In particular I want to thank Debora Botting, who is very supportive and a great friend.

Thanks to my friends and colleagues Jason Vertrees, Stephanie Vertrees, Sergio Santa Maria, Marty Gill, Diana Ferrari and Diego Vargas.

To all the fantastic group of students in the Biophysical, Structural and Computational Biology track I want to express my gratefulness as well. Also I extend my

gratitude to all the members in the Biological Chemistry Student Organization who were very supportive and highly involved all these years.

To my family who were always involved during my dissertation research; especially to my brothers and best friends Carlos Gonzalo and Jorge Andres; and to my dad, Carlos Alejandro. I want to thank my aunt Vicky Wilson and my uncle and Ricardo Wilson. And also to my aunt Zannie Sandoval, who gave me continuous support, compassion and encouragement. She was a very special driving force into my graduate studies.

My deepest thanks to the most important person in my life: my wife Maria Fe Lanfranco. She is my inspiration to become a better person every day.

THE MOLECULAR BASIS FOR EVASION OF ANTIBODY-MEDIATED NEUTRALIZATION IN FLAVIVIRUSES

Publication No._____

Rodrigo A. Maillard, B.Sc.

The University of Texas Medical Branch, 2007

Supervisor: James C. Lee, Ph.D.

In order to establish infection, flaviviruses induce mutations in antigenic proteins as a strategy to evade neutralizing antibodies; a phenomenon known as antibody-mediated neutralization resistance. Most of the critical mutations associated with antibody resistance are located in the envelope protein domain 3 (ED3). In West Nile virus (WNV), and in other flaviviruses, mutations in ED3 are found throughout its structure, including regions outside of the viral epitope and/or not accessible to antibodies. Besides the trivial effects of local perturbations due to mutations in the epitope, these observations are consistent with our hypothesis that there is long-range communication connecting distant residues linked to the viral epitope. Mutations at sites distant but coupled to the epitope would lead to an alteration of affinity to antibody; e.g., resistance to antibody-mediated neutralization. In support of our hypothesis, we demonstrated by multiple biophysical approaches the communications among distant sites and a network of communications of energetically coupled residues. Within this network, mutations in WNV ED3 caused perturbations only in the loop connecting strands B-C (BC loop) by changing the magnitude of energetic coupling between these distant sites. The magnitude of perturbation conveyed by the mutations is represented by a Boltzmann distribution. This suggests that neutralization resistance is the manifestation of an equilibrium process governing the distribution between ED3 conformations that are

responsible for antibody neutralization-resistance and nonresistance. Indeed we observed a linear correlation between affinity for antibody and magnitude of energetic coupling on the BC loop. To test the generality of these results, we investigated the ED3 from dengue virus type 2 (DENV2), a related flavivirus. We found that only the FG loop was susceptible to mutational perturbations. Remarkably, the BC and FG loops have been shown to be the dominant epitopes in ED3 for WNV and DENV2, respectively. Evidently these distant sites are energetically coupled to their respective viral epitope. This study reveals the strategy by which flavivirus employed to evade antibody, namely, establishment of long-range communications in viral proteins to expand the mutational repertoire to perturb the epitopes and lower the affinity for antibodies resulting in evasion of antibody-mediated neutralization.

TABLE OF CONTENTS

PAGE

LIST OF TABLES

LIST OF FIGURES

LIST OF ABBREVIATIONS

α	fractional tyrosine exposure to solvent
Å	angstrom
ALA (A)	alanine
apol	apolar
ASA	accessible surface area
ASAcal	calculated accessible surface area
<ASAcal>	averaged calculated accessible surface area
ASAexp	experimental accessible surface area
ARG (R)	arginine
ASP (D)	aspartic acid
bb	backbone
BC loop	loop connecting beta-strands B and C in ED3
bu-ex	buried that becomes exposed
C½	midpoint transition in chemical unfolding
cal	calorie
CD	circular dichroism
conf	conformational
ΔCp	heat capacity change
[D]	concentration of denaturant
Da	Dalton
DE loop	loop connecting beta-strands D and E in ED3
DENV2	dengue virus type 2
$d_{N\rightarrow U}$	peak distance between unfolded and folded state
ε	extinction coefficient
e	exponential
ED1	envelope protein domain 1
ED2	envelope protein domain 2
ED3	envelope protein domain 3
EDTA	di-sodium ethylenediamine tetra-acetic acid
E protein	envelope protein
ex-un	exposed that becomes unfolded
FG loop	loop connecting beta-strands F and G in ED3
$\gamma 1/\gamma 2$	quenching efficiency
ΔG	Gibbs free energy change
$\Delta\Delta G_{j,k}$	energetic coupling between residues j and k
$\Delta g^{pert,k}$	perturbation energy on residue k
GdnHCl	guanidinium hydrochloride
[GdnHCl]	concentration of guanidinium hydrochloride
GLU (Q)	glutamine

GLY (G)	glycine
ΔH	enthalpy change
HIS (H)	histidine
ILE (I)	isoleucine
IPTG	isopropyl beta-D-thiogalactoside
JE	Japanese encephalitis virus
$\kappa_{f,j}$	residue stability constant
k	kilo
K	Kelvin
Kd_{app}	apparent dissociation constant
K_i	statistical weight
Ksv	Stern-Volmer quenching constant
λ	excitation wavelength
LEU (L)	leucine
LYS (K)	lysine
M	molar
MAb	monoclonal antibody
[MAb]	concentration of monoclonal antibody
MBP	maltose binding protein
mM	milimolar
mol	moles
MR	mutational response
mut	mutant
m-value	slope of free unfolding energy on denaturant concentration
NaCl	Sodium chloride
NATA	N-Acetyl-L-tryptophanamide
ng	nanogram
NGC	new guinea C
N.I.	neutralization index
nm	nanomolar
NMR	nuclear magnetic resonance
N-ter	N-terminus
ω	angular velocity
OD	optical density
ρ_g	density of guanidinium
ρ	density
PBS	phosphate saline
PDB	protein data bank
P_i	Probability of conformer i
PHE (F)	phenylalanine
PMSF	Phenylmethanesulfonyl fluoride
pol	polar

PRO (P)	proline
psi	pounds per square inch
φ	quantum yield
$\phi^{pert,k}$	energetic perturbation on residue k
Q	quencher or partition function
[Q]	concentration of quencher
R	ideal gas constant
rED3	recombinant envelope protein domain 3
r.m.s.d.	root mean square deviation
RNA	ribonucliec acid
rpm	revolution per minute
Σ	summation
$s_{20,W}$	sedimentation velocity coefficient
ΔS	entropy change
Sc	new guinea C
SER (S)	serine
SDS-PAGE	sodium dodecyl sulfate polyacrylamide gel electrophoresis
solv	solvent
T	temperature
TBE	tick-borne encephalitis
TCE	trichloroethanol
THR (T)	threonine
Tm	midpoint transition in temperature unfolding
TN200	TRIS buffer with 200 mM NaCl
TN700	TRIS buffer with 700 mM NaCl
TYR (Y)	tyrosine
μg	microgram
μm	micrometer
μM	micromolar
UV	ultraviolet
ν	partial specific volume
VAL (V)	valine
W (Trp)	tryptophan
WNV	west nile virus
WT (wt)	wild type
YF	yellow fever

CHAPTER 1

INTRODUCTION

1.1 EMERGING FLAVIVIRUSES

Most members of the *Flavivirus* genus are arboviruses, the completion of whose transmission cycle requires a blood-sucking arthropod. However, a few members do not have a known vector[1] (Figure 1.1). The genus contains ~70 species, of which most are transmitted by mosquito or ticks[2] (mosquito-borne or tick-borne, respectively). Mosquito-borne members of flaviviruses represent important emerging and resurging diseases of global significance; for example, the resurgence of dengue in tropical and subtropical areas of the world, and the establishment of West Nile virus (WNV) in new habitats and environments such as the Americas[1,3]. Currently there are no human vaccines against WNV or dengue virus[1].

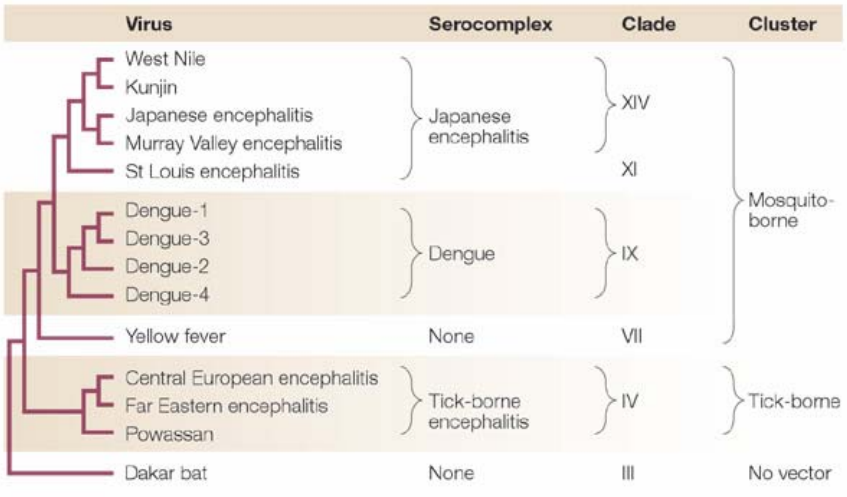

Figure 1.1. Phylogenetic classification of Flavivirus (clade and cluster) and serocomplex classification[4,5].

1.2 MOSQUITO-BORNE FLAVIVIRUSES

One of the most significant emerging pathogens is WNV in the western hemisphere. After first isolated in New York in 1999[6,7], WNV spread very rapidly throughout continental USA (Figure 1.3A) and neighboring regions such as Canada, Mexico, Central America and recently, South America[8-11]. It is thought that the rapid colonization of the Americas by WNV can be attributed to the extraordinarily broad range of bird hosts and mosquito vectors that are part of the transmission cycle[2] (Figure 1.2). More than 280 species of birds (including migratory birds) and close to 60 species of mosquito (mainly of the genus *Culex* spp.) have been documented as amplifying hosts or vectors, respectively (available at the Center of Disease Control and Prevention website: http://www.cdc.gov/ncidod/dvbid/westnile/index.htm). Humans and other vertebrates, such as horses, are incidental hosts and are thought to have a minor role in the transmission cycle because they are not capable of supporting high viremias to efficiently transmit WNV to feeding mosquitoes (Figure 1.2).

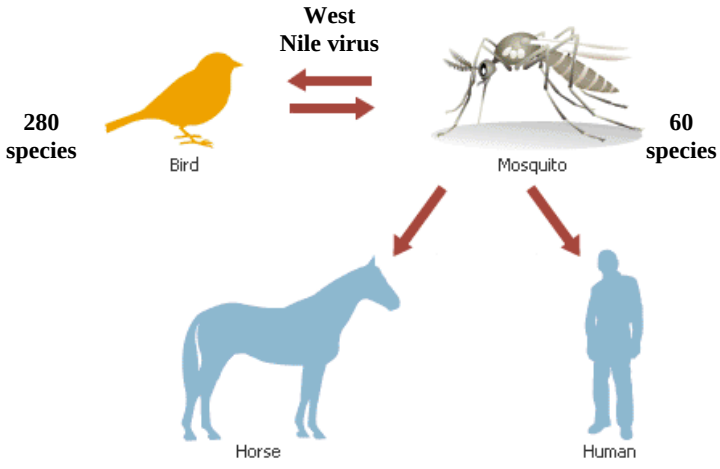

Figure 1.2. Transmission cycle of WNV between birds and the mosquito vector. Humans and horses are incidental hosts. ~280 species of birds and ~60 species of mosquito are part of the WNV transmission cycle.

Despite that humans are only incidental hosts of WNV, the incidence of WNV human infections in USA, since 2002, has increased dramatically. During the years 2002-06, more than 22,000 cases of human infections were reported to the Center of Disease Control and Prevention (CDC). Of those, 8,500 progressed into neurological conditions such as encephalitis or meningitis resulting in more than 940 deaths (Figure 1.3B) (case count update available via CDC-ArboNet).

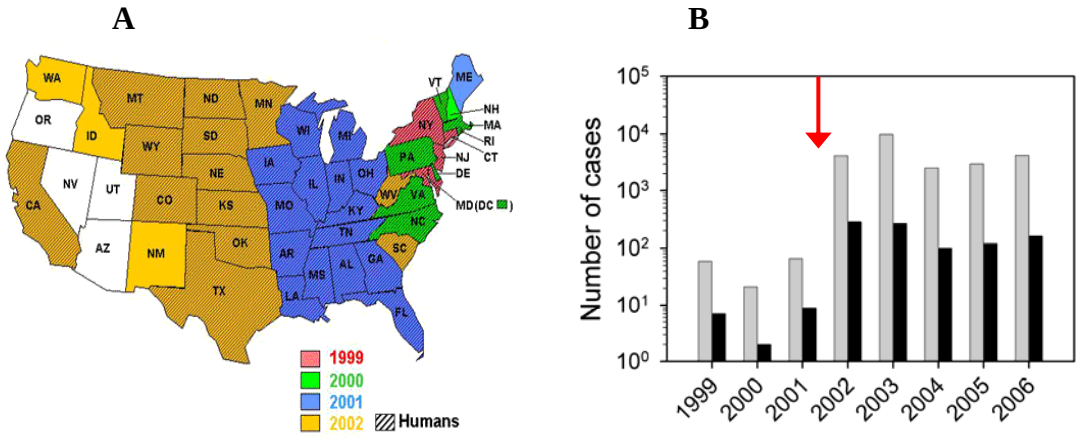

Figure 1.3. **A**, Yearly progression of WNV distribution from 1999 to 2002. **B,** Reported cases of WNV human infection (grey bars) and resulted fatalities (black bars). Information obtained from the CDC website at http://www.cdc.gov/ncidod/dvbid/westnile/index.htm. Red arrow indicates the significant increase in human infections, which correlates with the rapid geographical distribution.

The establishment of WNV human infections can be attributed to the development of mechanisms of evading the host humoral immune system, a phenomenon known as antibody neutralization resistance. Humoral immunity is an essential component of the immune response against flaviviruses as neutralizing monoclonal antibodies (MAbs) limit dissemination of infection[12-14]. In WNV, as well as in other flaviviruses, the induction of mutations in viral antigenic proteins, a phenomenon described as antigenic variation, is suggested as the principal strategy to evade potent neutralizing MAbs[12]. Most neutralizing MAbs that recognize WNV, and induce a

protective immune response, are targeted to the envelope (E) protein[14-16], which sits tangentially to the surface and fully covers the viral particle (Figure 1.4A). The E protein forms a head-to-tail homodimer (Figure 1.4B), whose exterior soluble segment (ectodomain), is composed of three distinct structural domains (ED1-3) mainly based on beta-strand structures[17] (Figure 1.4C). ED1 and ED2 are the central and dimerization domains, respectively, and ED3 is the putative receptor-binding domain[4,5].

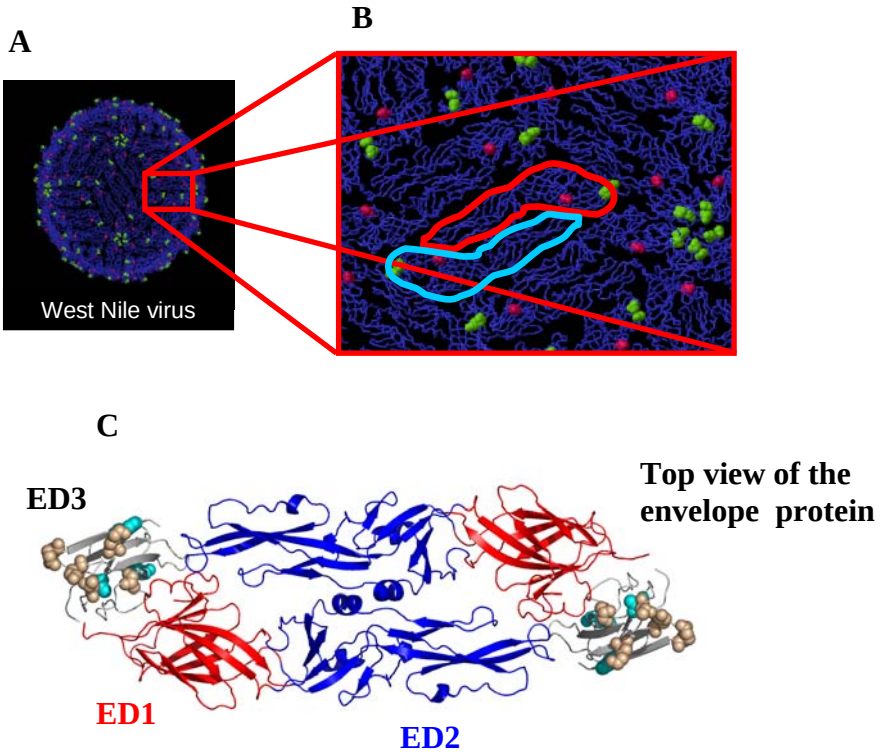

Figure 1.4. **A and B**, Cryo-electron microscopy reconstruction of the WNV viral particle. The E protein is shown in blue in which green and pink dots corresponds to solvent exposed mutation sites in ED3. Figure kindly provided by Richard Kuhn (Purdue University). **C,** Top view of the structure of the DENV2 E protein homodimer[18]. The ED3 of WNV[19] was modeled in the E protein from DENV2. Beige and cyan spheres correspond to naturally occurring mutation sites that are solvent and non-solvent exposed.

Although all three domains in the E protein contain antigenic sites, suggesting that there is not a single defined one, most of the experimental evidence strongly suggests

that the envelope protein domain 3 (ED3) is the major antigenic domain in WNV as well as in most flaviviruses. In fact, many studies have shown that potent neutralizing MAbs binding to ED3 are the most efficient at blocking virus attachment to cells[14-16,20-23], although some anti-ED3 MAbs may impair virus membrane fusion[24]. Consistent with the functional role of the ED3 is its localization in the context of the intact virion. The ED3 clearly projects above the rest of the E protein facilitating the interactions with other macromolecules[25-27] (Figure 1.5).

Figure 1.5. Side view of the structure of the DENV2 E protein homodimer[18] in which the ED3 of WNV was modeled[19]. Beige and cyan spheres correspond to naturally occurring mutation sites that are solvent and non-solvent exposed. Domains 1 and 2 are colored in red and blue, respectively. In addition, the Fab domain of the neutralizing antibody E16 was included[24]. Red arrow indicates position E390 that is part of the binding site but does decrease MAb binding.

1.3 RATIONALE FOR THE STUDY

1.3.1 Background problem, central hypothesis and specific aims

It is generally assumed that mutations that decrease MAb binding affinity, and potentially generate a MAb-neutralization resistant variant, are located in the viral epitope. However, it is interesting to note that several mutations found in different WNV variant strains[16] mapped to residues that were both solvent and non-solvent exposed in ED3 (Figure 1.5, beige and cyan spheres, respectively). And even in the case of solvent-

exposed mutations, many were distantly positioned from the epitope. Mutations outside the viral epitope that confer MAb-neutralization resistance have been documented in many flaviviruses[16,28-30] as well as in other non-related viruses[31,32]. In addition, in WNV ED3, most mutations at residues that comprise the viral epitope do not impair MAb binding[14,24] (e.g. the mutation site is highlighted by the red arrow in Figure 1.5), indicating that the mechanisms involved in MAb-neutralization resistance are not only limited to local perturbations of the antigen-antibody interaction surface. Thus, it is inappropriate to presume that the mechanism for the broad spectrum of phenomena in MAb-neutralization resistance is automatically linked to mutations in the viral epitope.

Based on those observations, and to fulfill the knowledge gap in the mechanism of MAb-neutralization resistance, ***we hypothesize that there is long-range communication connecting distant sites which are in turn linked to the viral epitope***. As a consequence to this network of communication, mutations in these distant sites would lead to a change in the epitope and decrease the affinity of MAb to the viral protein, i.e. resistance to antibody-mediated neutralization.

I tested this central hypothesis via the following specific aims:

1. *The identification of long-range communications between distant sites in ED3 from WNV* **(CHAPTER 2)**.

By investigating the solution biophysical properties in ED3 from WNV and three single-site naturally occurring mutations we will address the following question: Are there long-range interactions between distant regions in ED3? This aim will provide the experimental evidence for long-range interactions between mutation sites in the WNV ED3.

2. *Identify the role of mutations outside the viral epitope in WNV ED3 in the mechanism of MAb-neutralization resistance* **(CHAPTER 3)**.

We validated and applied the COREX algorithm to determine the correlated fluctuations of ED3 and establish a network of energetically coupled residues in ED3 from WNV. In this aim we addressed the following questions: Are there long-range interactions coupled to the epitope? Is there a residue network between distant mutation

sites and the dominant epitope in WNV ED3? If this is the case, what are the underlying principle governing these long-range communications?

3. *Validation of the generality of the mechanism for MAb-neutralization resistance in closely related flaviviruses* **(CHAPTERS 4 and 5)**.

Dengue virus type 2 (DENV2) also represents a major resurging mosquito-borne flavivirus of worldwide importance (Figure 1.6). The DENV2 ED3 is very similar structure-wise and sequence-wise to the WNV ED3 but with very different antigenic properties (Figure 1.1). We used the COREX algorithm to investigate the networks of energetically coupled residues in DENV2 ED3. This aim will address the following questions: Are the networks of communication in ED3 different between closely related viruses (e.g. WNV and DENV2)? What are similarities and differences in the residue networks between WNV and DENV2 ED3s? Is there a biological significance of establishing residue networks? Is there an additional viral advantage of having mutations outside the viral epitope to evade neutralizing MAbs that it is not known?

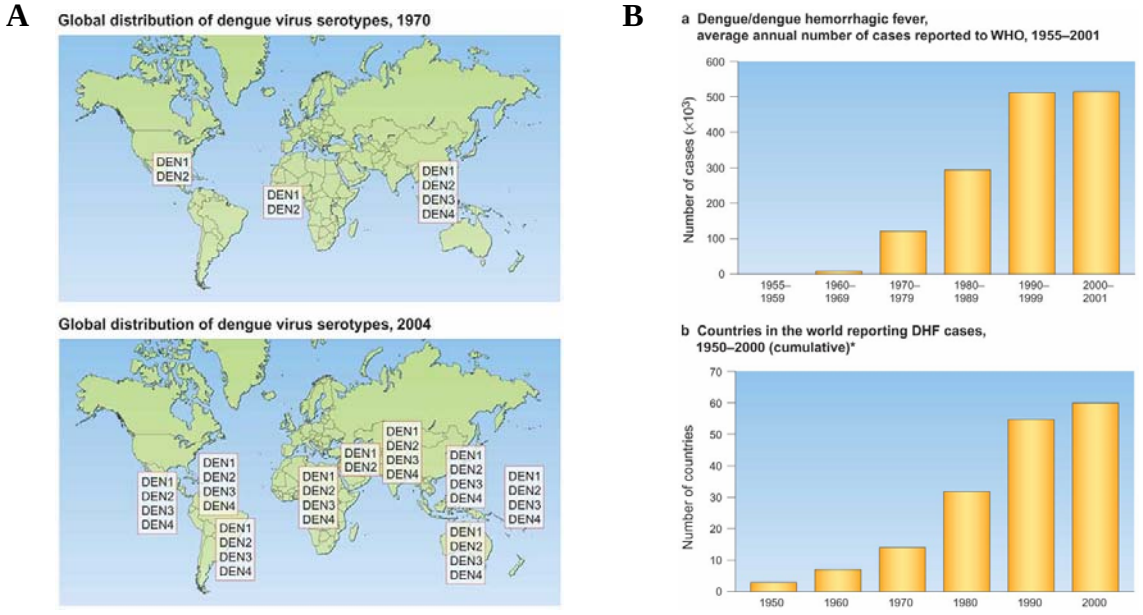

Figure 1.6. **A**, The change in geographical distribution of dengue serotypes from 1970 to 2004. **B,** The global resurgence of dengue and its associated disease (dengue hemorrhagic fever) over the past century, by number of cases (upper panel) and by country (lower panel). Adapted from Ref. 1.

CHAPTER 2

COMMUNICATION BETWEEN DISTANT SITES IN THE ENVELOPE PROTEIN DOMAIN 3 OF NEUTRALIZATION-RESISTANT MUTANTS OF WEST NILE VIRUS

2.1 ABSTRACT

Monoclonal antibody resistance is induced by mutations in viral proteins, which lead to changes in the interaction of the antibody with the virus epitope. Single mutations on domain 3 (ED3) from WNV can confer resistance to the virion against wild type-targeted neutralizing antibodies (MAbs). However, studies have indicated that structural elements other than the epitope may play a role in neutralization resistance. In order to dissect the molecular mechanism of neutralization resistance, we have studied the biophysical properties of ED3 from WNV wild type (strain USA99b) and single mutations found in WNV variant strains. These mutations are K310T, T332A and T332K. From these studies we have determined a quantitative correlation where the higher the resistance to neutralization, the less stable and more dynamic is the rED3. Furthermore, by using a variety of biophysical approaches we observed that the mutations affect the dynamics and surface chemistry of ED3 at regions outside the mutation site, implying long-range communication within ED3. Based on our results, we hypothesize that the mechanisms involved in neutralization resistance is dependent not only on regions of interaction with the antibody (epitope) but also on distant structural elements of the protein via long-range communication.

2.2 INTRODUCTION

The genus *Flavivirus* of the family Flaviviridae consists of a number of pathogens of major public health importance, which cause diseases including yellow fever (YF), Japanese encephalitis (JE), tick-borne encephalitis (TBE), West Nile (WN) and dengue[33]. All studies to date indicate that neutralizing antibodies are the major mechanism of protective immunity against flaviviruses[34], as demonstrated by the effectiveness of vaccines to prevent other flavivirus diseases (YF, JE and TBE). The primary target of virus-neutralizing antibodies is the envelope (E) protein[35]. The E protein is the major structural protein on the surface of flaviviruses and consists of three distinct structural domains (ED1-3). ED1 and ED2 are the central and dimerization domains, respectively, and ED3 (amino acids 295-395 in the E protein) is the putative receptor-binding domain[17,36]. Consistent with the functional role of ED3 is its orientation in the context of the intact virion. The ED3 of WNV clearly projects above the rest of the E protein facilitating the potential interaction with other macromolecules, such as antibodies or cell receptors[25]. Although all three domains in the E protein contain immunogenic sites, suggesting that there is not a single defined antigenic site, most of the experimental evidence strongly indicates that ED3 is the major immunogenic domain for virus type-specific neutralizing monoclonal antibodies[35] (MAbs). Many studies have shown that MAbs binding to ED3 are the most efficient at blocking virus attachment to cells[20-23], although it has recently been suggested that some anti-ED3 MAbs may impair virus membrane fusion[24,37]. Furthermore, mutations on ED3 are associated with changes in its immunogenic properties resulting in neutralization resistant variants and, frequently, in attenuation of neuroinvasiveness or neurovirulence phenotypes[35]. I has been previously determined that residue T332 in ED3 is a critical amino acid in the major neutralization site of WNV (Figure 2.1). Mutations at this position decrease MAb binding affinity by > ~100-fold and generates a fully resistant variant in neutralization assays *in vitro*[15,16].

In general, amino acid substitutions that are spatially located in the antibody binding site (epitope) interrupt the antibody-antigen interaction that leads to virus neutralization. However, several studies with flaviviruses have shown that antibody

binding (or neutralization) can be perturbed by regions other than the epitope. For example, flaviviruses that escape from neutralization due to a mutation without loss in MAb binding imply additional structural elements outside the epitope that are required for virus neutralization[16,28-30]. Furthermore, there is evidence of positive cooperativity in antibody binding (avidity) for flaviviruses indicating that virus neutralization can be the consequence of cross-talk among distant amino acids in the E protein[38-40].

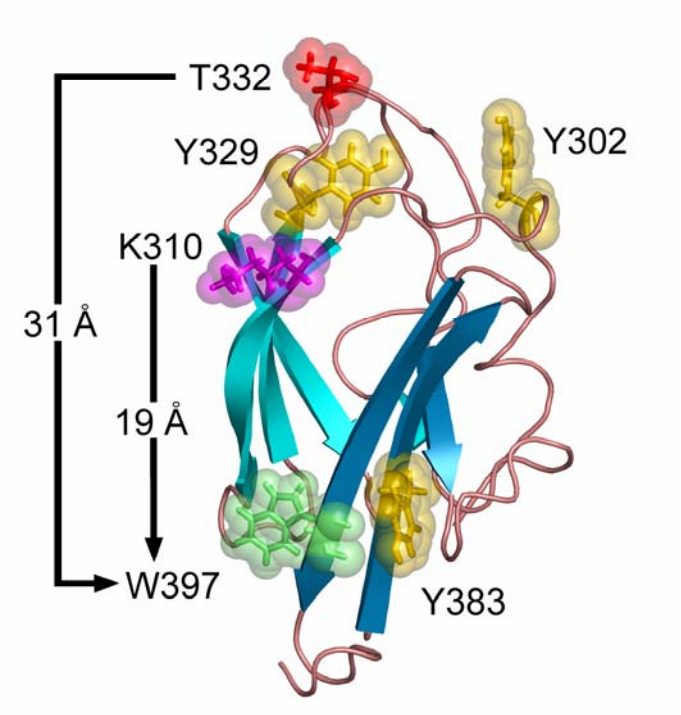

Figure 2.1. NMR structure of rED3-WT[19] (PDB 1S6N) showing the distance (in Å) between residues T332 (red), K310 (magenta) and W397 (green). Residue numbering corresponds to amino acids 296-406 of the entire E protein of WNV. Residues in yellow correspond to tyrosine residues Y302, Y329 and Y383. The ribbon diagram was rendered using PyMOL v. 0.97 (Delano Scientific LLC, San Carlos, CA).

We have studied the effect of single mutations on the biophysical properties of ED3 from the North American WNV prototype strain, USA99b. The three single

mutations rED3-K310T, rED3-T332A and rED3-T332K were chosen because they cover a broad spectrum of degree of resistance to neutralization by a panel of WNV subtype-specific MAbs[15,16]. The substitutions rED3-T332K and rED3-T332A confer full and partial resistance to the virus, respectively; while rED3-K310T has no significant effects. By using a variety of biophysical methods, we have observed changes in protein dynamics and stability which correlate with the degree of neutralization in such a way that the most resistant mutant (i.e., rED3-T332K) is the least stable and has the most dynamic ED3. Furthermore, we have observed that the mutations affected the dynamics and surface electrostatics at regions outside the mutation site, implying long-range communication. Based on these results we hypothesize that the mechanism of neutralization is dependent not only on regions of interaction with the antibody (epitope) but also on distant structural elements of ED3 via long-range communication.

2.3 MATERIALS AND METHODS

2.3.1 Chemicals and buffers

All chemicals were reagent grade. Acrylamide was from Boehringer Mannheim. 2,2,2-Trichloroethanol, Potassium iodide, Thallium acetate, N-Acetyl-L-tryptophanamide, N_{α}-Acetyl-L-tyrosinamide and Phenylmethanesulfonyl fluoride were from Sigma. Quinine Bisulfate was from Kodak. Ultrapure Guanidine hydrochloride was from ICN Biochemicals. All solutions were made with distilled and deionized >18-megaohm water. The standard buffer, TN200, was 20 mM Tris pH 8.0 (adjusted with HCl at 20 °C), 200 mM NaCl, 1 mM EDTA and 5% Glycerol v/v. Phosphate buffer was 100 mM $NaHPO_4$ pH 7.4 at 20 °C, 200 mM NaCl, 1 mM EDTA and 5% Glycerol v/v.

2.3.2 Guanidine hydrochloride preparation

A stock solution of GdnHCl (7.5 M) was prepared in TN200 buffer or phosphate buffer, filtered (0.45 μm pore size membrane), degassed and corrected for pH. Density

measurements were done with a Mettler/Paar Precise Density Meter model DMA O2D with a fix temperature of 20.0 ± 0.1 °C controlled by a circulating water bath. The concentration of GdnHCl was determined by using:

$$[GdnHCl] = \frac{W \cdot \rho_G \cdot 10^3}{95.53} \tag{2.1}$$

where W is the weight fraction of GdnHCl in solution; ρ_G is the density of the GdnHCl stock solution and 95.53 is the molecular weight of GdnHCl. W can be determined by solving the following empirical equation[41]:

$$\frac{\rho_G}{\rho_B} = 1 + 0.271W + 0.033W^2 \tag{2.2}$$

where ρ_B is the density of TN200 buffer or phosphate buffer.

2.3.3 Protein preparation

The recombinant E protein structural domain 3 (rED3) incorporates amino acids 296-406 of the entire E protein from WN virus strain 385-99. The rED3 gene was harbored in a pMAL-c2x vector (New England Biolabs) and expressed as a maltose-binding protein (MBP) fusion protein[42]. Wild type (rED3-WT) and single mutant rED3s were prepared identically by sequential chromatography on Bio-Rex 70 (Bio-Rad), amylose affinity resin (New England Biolabs) and Bio-Gel P-30 (Bio-Rad). All chromatographic steps were carried out at 4 °C unless stated otherwise. The purification scheme was as follows: Transformed cells (*Escherichia coli* strain BL21) were grown at 37 °C in rich media with glucose (2 g/L) until the $OD_{600\ nm}$ reached 0.7 and then induced with IPTG (0.2 mM final concentration). The cells were pelleted and suspended in MES buffer (20 mM MES pH 5.5, 200 mM NaCl, 1 mM EDTA and 5 % Glycerol v/v) with 1

mM PMSF and lysed with high pressure (10,000 psi) for 1 h. After centrifugation, the resulting supernatant was passed over Bio-Rex 70 pre-equilibrated with MES buffer. The matrix was washed with the same buffer until the eluant $OD_{260\ nm}$ was < 0.05. The fusion protein was eluted out from Bio-Rex 70 at room temperature with a low-to-high salt gradient (0.2 M → 1.0 M $NaHPO_4$ pH 8.3, 5 % Glycerol v/v) and dialyzed against TN200 buffer until the salt concentration was near 200 mM NaCl (the salt concentration was determined by conductivity). The dialyzate was passed over amylose resin pre-equilibrated with TN200 buffer and washed with 3X column volumes of the same buffer. The fusion protein was eluted out with a gradient of 0 M → 0.01 M maltose in TN200 buffer, concentrated to ≥ 10 mg/ml and cleaved (0.1 % w/w) with Factor Xa (Novagen-EMD Biosciences) at room temperature for 14 h. The rED3 was separated from the MBP by size exclusion chromatography using Bio Gel P-30 pre-equilibrated with degassed TN700 buffer (same as TN200 buffer but with 700 mM NaCl). The fractions containing rED3 were dialyzed against TN200 buffer and concentrated. All rED3s were > 95 % homogeneous as judged by Coomassie Blue stained SDS-PAGE gels with a loading 20-30 μg/lane of protein (Figure 2.3, Results section). The rED3s concentration was determined using an extinction coefficient of $\varepsilon_{280\ nm} = 10875\ M^{-1}cm^{-1}$ based on the protein aromatic amino acid composition[43]. Before using, the rED3s were dialyzed against fresh buffer (filtered with a 0.45 μm pore size membrane followed by degassing) and centrifuged at 14,000 rpm, at 4 °C.

2.3.4 Chemical unfolding

Unfolding of rED3s was evaluated by measuring changes in the maximum emission fluorescence intensity at 340 nm (normalized to the lamp intensity) as a function of GdnHCl concentration. The excitation wavelengths were 280 nm and 295 nm. Fluorescence recordings were monitored with a SLM 8000C spectrofluorometer using a 5.0 mm-path length quartz cuvette and maintaining the temperature constant at 20.0 ± 0.1 °C. Chemical denaturation measurements were performed by mixing appropriate

proportions of stock denaturant solution (i.e., 7.5 M GdnHCl in TN200 buffer), TN200 buffer and protein solution (i.e., 80 μM dialyzed exhaustively against TN200 buffer), followed by incubation at 4 °C for ≥ 14 h. The protein concentration was 18 μM. Baseline corrections were made for buffer and denaturant contributions to the fluorescence intensity. GdnHCl-induced unfolding data were analyzed by nonlinear least-square fitting using the following equation, which describes a two-state unfolding process:

$$\Delta F_D = \frac{(a_N + m_N[D]) + (a_U + m_U[D]) \cdot e^{-\left[\frac{-m_D \cdot ([D]-C_{1/2})}{RT}\right]}}{1 + e^{-\left[\frac{-m_D \cdot ([D]-C_{1/2})}{RT}\right]}} \qquad (2.3)$$

where ΔF_D is the change in maximum emission intensity at a given concentration of GdnHCl ([D]); a_N, m_N, a_U and m_U are the intercepts (a) and slopes (m) of the base lines corresponding to the native (N) and unfolded (U) states, respectively. $C_{1/2}$ is the concentration of GdnHCl at which the population of folded and unfolded states is equal. m_D is the slope of the linear dependence of ΔG on denaturant concentration as described by the Linear extrapolation method or LEM[44] ($\Delta G = \Delta G^{H_2O}_{N \to U} + m_D[D]$). The reported $\Delta G^{H_2O}_{N \to U}$ is the free energy change for conversion of native to unfolded protein in the absence of denaturant at 20 °C.

2.3.5 Thermal unfolding

Thermal denaturation measurements were evaluated by changes in the maximum emission fluorescence intensity at 340 nm using a Perkin-Elmer LS50B spectrofluorometer. The excitation wavelengths were 280 nm and 295 nm. The temperature in the cuvette was increased using temperature-controlled circulating water with a precision of ± 0.1 °C. The scan rate was 60 °C/h. To avoid temperature differences between the circulating water and the cuvette, the temperature of the protein solution was

measured directly using a thermocouple. The temperature scans for the rED3s were performed as a function of increasing concentrations of GdnHCl (0.0 M → 1.0 M). Samples were prepared in the following order: First, appropriate proportions of stock denaturant solution (7.5 M GdnHCl in phosphate buffer) and phosphate buffer were mixed, degassed and checked for correct pH. Then, the protein solution was added to a final concentration of $4 \cdot 10^{-1}$ µM. Samples with GdnHCl were incubated at 4 °C overnight for ≥ 14 h. Baseline corrections were made for buffer and denaturant contributions to the fluorescence intensity. Temperature-induced unfolding data for all GdnHCl concentrations were simultaneously analyzed by nonlinear least-squares fitting using the following equation, which describes a two-state unfolding process:

$$\Delta F_{T,D} = \frac{(a_N + m_{N,T}T + m_{N,D}[D]) + (a_U + m_{U,T}T + m_{U,D}[D]) \cdot e^{-\left(\frac{\Delta G_{N\to U}(T,D)}{RT}\right)}}{1 + e^{-\left(\frac{\Delta G_{N\to U}(T,D)}{RT}\right)}} \quad (2.4)$$

where $\Delta F_{T,D}$ is the change in maximum emission intensity at a given temperature *(T)* and concentration of GdnHCl ([*D*]); a_N, $m_{N,T}$, $m_{N,D}$, a_U, $m_{U,T}$ and $m_{U,D}$ are the intercepts (*a*) and slopes (*m*) for the temperature (*T*) and denaturant (*D*) dependences of the base lines corresponding to the native (*N*) and unfolded (*U*) states, respectively. $\Delta G_{N\to U}(T,D)$ includes the temperature and denaturant dependence on the change of Gibb's free energy of unfolding and it is defined by:

$$\Delta G_{N\to U}(T,D) = [\Delta H^0 + \Delta Cp^0(T - T_m)] - T \cdot \left[\frac{\Delta H^0}{T_m} + \Delta Cp^0 \cdot \ln\left(\frac{T}{T_m}\right)\right] - m_D[D] \quad (2.5)$$

where ΔH^0 and ΔCp^0 are the enthalpy and heat capacity changes for conversion of native to unfolded protein; T_m is the temperature where the population of native and unfolded states is equal and m_D is the slope of the linear dependence of ΔG on denaturant concentration (see above).

2.3.6 Fluorescent measurements

All steady-state fluorescence intensity measurements were carried out in 10 mm-path length quartz cuvettes using a Perkin-Elmer LS50B spectrofluorometer. The temperature of the cuvette holder was regulated by circulating water at 20.0 ± 0.1 °C. The excitation and emission polarizers were set at 90° and 55° (magic angle) to avoid potential artifacts due to fluorescence anisotropy of the sample. All measurements were corrected for inner filter effects using the formula[45]:

$$F_{cor} = F_{exp} \cdot 10^{0.5b(A_{\lambda ex})} \tag{2.6}$$

where F_{cor} is the corrected value of fluorescence intensity, F_{exp} is the experimentally measured fluorescence intensity, b is the length in the optical path in the cuvette (1 cm) and $A_{\lambda ex}$ is the absorbance of the sample at the excitation wavelength.

2.3.7 Tryptophan fluorescence quenching

Quenching experiments were monitored at 340 nm with excitation wavelength at 295 nm. The protein concentration was 2 μM in TN200 buffer. Four different quenching molecules (quenchers) were use for these experiments: Acrylamide (neutral quencher), iodide (negatively charged quencher), thallium (positively charged quencher) and trichloroethanol or TCE (hydrophobic quencher). Due to the low solubility of thallium in the presence of chloride, the standard TN200 buffer was substituted by buffer Tris-Acetate (20 mM Tris-Acetate pH 8.0, 200 mM Sodium Acetate, 1 mM EDTA and 5% glycerol v/v). The Job method[46] for continuous variation was applied to keep the solutions at constant ionic strength during the titrations with iodide and thallium. For each titration, four cuvettes were used and measured independently: one for baseline (A), one for dilution (B), one for the titration itself with a given quencher (C) and one to measure photobleaching (D). Cuvettes B, C and D contained the same protein concentration. Cuvette A contained only buffer. During the experiment, aliquots of

quencher (in TN200 buffer or Tris-Acetate buffer) were added to cuvettes A and C; and to cuvette B, the same volume of buffer was added. To cuvette D neither acrylamide nor buffer was added. Under this experimental set up, the actual effect of the acrylamide on the fluorescent properties of the tryptophan residue in the rED3 can be extracted by the following relation:

$$F_Q = \left(\frac{F_C - F_A}{F_B - F_A} \right) / \left(F_D \right) \tag{2.7}$$

where F_Q is the corrected steady-state fluorescent intensity and; F_A, F_B, F_C and F_D are the steady-state fluorescent intensity of each cuvette as described above. If only the dynamic quenching process is present, then, the process is described by the classic Stern-Volmer equation[47]:

$$\frac{F_{Q=0}}{F_Q} = 1 + K_{SV}[Q] \tag{2.8}$$

where $F_{Q=0}$ and F_Q are the corrected steady-state fluorescent intensity in the absence or presence of quencher at concentration [Q], respectively. K_{SV} is the Stern-Volmer quenching constant in M^{-1} units. The presence of both static and dynamic quenching processes may generate deviations from the ideal Stern-Volmer behavior and can be accurately described by[48]:

$$\frac{F_{Q=0}}{F_Q} = (1 + K_{SV}[Q]) \cdot e^{(V[Q])} \tag{2.9}$$

where V is the static quenching constant in M^{-1} units.

2.3.8 Tyrosine exposure determination

The fractional exposure of tyrosine residues, α, in the native state of the rED3s was calculated from the 2nd-derivative absorbance spectra as described previously[49] using the following equation:

$$\alpha = \frac{(r_N - r_a)}{(r_U - r_a)} \tag{2.10}$$

where $r_{i=N,U,a}$ is the ratio between two peak-to-peak distances in the 2nd-derivative spectra. The peaks correspond to one between the maximum at 287 nm and the minimum at 283 nm, and one between the maximum at 295 nm and the minimum at 290.5 nm. This ratio ($r_{i=N,U,a}$) is determine mainly by the polarity of the tyrosine environment: $r_{i=N}$ and $r_{i=U}$ correspond to the ratio of the folded and fully unfolded state of the protein; and, $r_{i=a}$ corresponds to the ratio of the molar mixture of aromatic amino acids (tyrosine and trytophan), as in the studied protein, dissolved in ethylene glycol, which represents an environment similar to the complete burial of the aromatic amino acids in the interior of the protein. Absorbance spectra were recorded using a Hitachi U-3010 spectrophotometer equipped with a Peltier temperature controlled unit, with all measurements conducted at 20.0 ± 0.1 °C. The protein concentration was 28 μM. Spectra were recorded in triplicate every 0.2 nm with 2 nm bandwidth over the range of 340-240 nm at a speed of 30 nm/min. Savitzky-Golay smoothing was performed on the spectral data using a 30-point moving window size and fourth degree polynomial approximations to acquire the 2nd-derivative spectra[50].

2.3.9 Quantum yield determination

The quantum yield of the rED3s was determined by the comparative method of Parker and Rees[51] using the following equation:

$$\phi_P = \phi_R \cdot \frac{\int F_P \partial\lambda}{\int F_R \partial\lambda} \cdot \frac{A_{R-\lambda ex}}{A_{P-\lambda ex}} \quad (2.11)$$

where the subscripts P and R refer the protein and reference fluorophore. ϕ is the quantum yield; $\int F \partial\lambda$ is the area under the fluorescence emission spectra corrected for baseline and inner filter effect. $A_{P\text{-}\lambda ex}/A_{R\text{-}\lambda ex}$ is the absorbance of the protein/reference fluorophore at the excitation wavelength, respectively. Quinine bisulfate in 0.1 N H_2SO_4 was used as reference fluorophore with an absolute quantum yield (ϕ_R) of 0.55[52].

2.3.10 Circular dichroism (CD) spectroscopy

CD measurements were performed on an AVIV model 60 DS spectropolarimeter. Far-UV CD spectra of the rED3s in the native state (in TN200 buffer) were measured over the ranges of 260-200 nm using a 1.0 mm-path length cell and 220-190 nm using a 0.1 mm-path length cell. Each spectrum was recorded in 0.5 wavelength increments. For each rED3, 3 repetitive scans were obtained, averaged and baseline-corrected.

2.3.11 Analytical ultracentrifugation

Sedimentation velocity experiments were carried out in a Beckman Optima XL-A analytical ultracentrifuge equipped with absorbance optics and An60Ti rotor at 60,000 rpm, 20 °C. Velocity data were collected at the absorption band of 280 nm and were analyzed using the program SedFit[53]. The reported weight-average sedimentation coefficient values ($\overline{S}_{20,w}$) were corrected for the solvent viscosity and temperature to standard conditions. Sedimentation equilibrium experiments were performed at 40,000 rpm, 20 °C. The apparent weight-average molecular weights of the rED3s were obtained by non-linear square fitting of the sedimentation equilibrium data to the following equation:

$$C = E + C_{1,r_0} \cdot e^{\left[\frac{(1-\bar{\nu}\rho)\omega^2}{2RT} M (r^2 - r_0^2)\right]} \qquad (2.12)$$

where C is the observed protein concentration in absorbance at radial position r, E is the baseline offset, and C_{1,r_0} is the concentration of monomeric protein. At the meniscus, r_0, $\bar{\nu}$ is the partial specific volume, ρ is the solvent density, ω is the angular velocity, M is the apparent weight-average molecular weight; and R and T are the gas constant and temperature in degrees Kelvin, respectively. The value of $\bar{\nu}$ was 0.742 ml/g for the studied proteins in TN200 buffer, which was derived from the amino acid composition using the method of Cohn and Edsall[54]. A density value (ρ) of 1.019 g/ml for TN200 buffer was determined using an Mettler/Paar Precise Density Meter model DMA O2D with a fix temperature of 20.0 ± 0.1 °C. For all experiments the concentration of the rED3s was 30 μM.

2.3.12 Monoclonal antibody binding

Binding of WNV type-specific MAbs 3A3 and 5H10 (Bioreliance Corporation, Rockville, MD) to purified MBP-rED3 fusion proteins was evaluated by indirect ELISA at 20 °C, as described previously[16]. Briefly, ELISA plates were coated overnight at 4 °C with 30 ng of pure fusion protein in borate saline buffer pH 9.0. The coated plates were blocked with a solution of PBS containing 3 % bovine serum albumin for 1 h. The MAbs were diluted in PBS with 0.5 % Tween 20 to a concentration of 10 nM, which corresponds to ~100-fold greater than the previoulsy reported Kd_{app} value for rED3-WT[19]. As control an anti-MBP serum was included. All binding reactions were performed in triplicate. After a 45 min incubation time between the MAb and the fusion protein, the plates were washed in triplicated with PBS with Tween. Then, a peroxidase-labeled anti-mouse IgG antibody was added to each well for 45 min. The plates were

again washed in triplicate and the colorimetric reaction was started by adding 3,3',5,5'-tetramethylbenzidene substrate (Sigma) for 10 min. The reaction is stopped with 3.0 M HCl and the absorbance values read on a model 3550-UV plate reader (Bio-Rad) at 450 nm with reference of 595 nm. The apparent dissociation constant (Kd_{app}) was obtained by fitting the data to a single-site binding isotherm using a standard package from Sigma Plot version 8.0.

2.4 RESULTS

2.4.1 Optimization of the purification protocol of the rED3s

For our solution biophysical studies, it was necessary to develop an efficient purification scheme of the rED3s to obtain large quantities in high purity. Because the rED3s is expressed as a fusion protein (Maltose Binding Protein (MBP) fusion protein)[42], the affinity chromatography step using and amylose resin, in which MBP binds, was preceded by a cation exchange chromatographic step to elute out principally nucleic acid contaminations. This step significantly improved the purification efficiency of the amylose affinity chromatographic step. The flow diagram of the purification scheme, applied to all rED3s (wild type and the mutants rED3-K310T, rED3-T332A and rED3-T332K is illustrated in Figure 2.2.

All the rED3s studied here were in the soluble fraction after over expression with IPTG. In Figure 2.3 it is shown a SDS-PAGE 12 % summarizing the purification protocol illustrated in Figure 2.2.

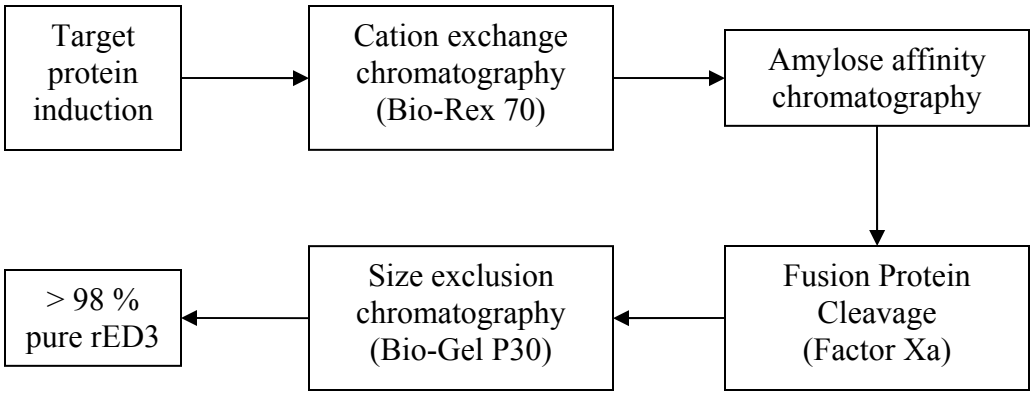

Figure 2.2. Flow diagram of purification scheme of rED3 from WNV from a Maltose Binding Protein (MBP) fusion protein. The purification scheme was applied to rED3-WT and single mutations.

2.4.2 Correlation between MAb neutralization *in vitro* and MAb binding affinity

We first assessed the biological relevance of using the recombinant ED3s (rED3s) from wild type (WNV prototype strain USA99b) and the mutants rED3-K310T, ED3-T332A, rED3-T332K to represent the phenomenon of virus neutralization resistance to MAbs in cell assays (*in vitro*). The reduction of the number of virus plaques (or virus titer) in cell-infected plates in the presence of MAb, compared to a medium-only control, is the *neutralization index* (N.I.)[16]. The neutralization index qualitatively characterizes the degree of resistance conferred to the virus due to a mutation relative to wild type virus.

It was previously reported that WNV strains harboring the mutations T332A and T332K generated a partially and fully resistant virus, respectively[15,16]. In agreement with this behavior *in vitro*, the binding affinity to MAbs for rED3-T332A and rED3-T332K was reduced by ~5 to > 100-fold relative to rED3-WT, respectively (Table 2.1). Figure 2.4 shows the binding isotherms of rED3s to two WNV type-specific MAbs (3A3 and 5H10). The data was fitted to a single-site binding isotherm to obtain the apparent dissociation constant (Kd_{app}). The mutation K310T had no effect in MAb binding affinity.

These observations are consistent for both MAbs 3H3 and 5H10, and the results are summarized in Table 2.1.

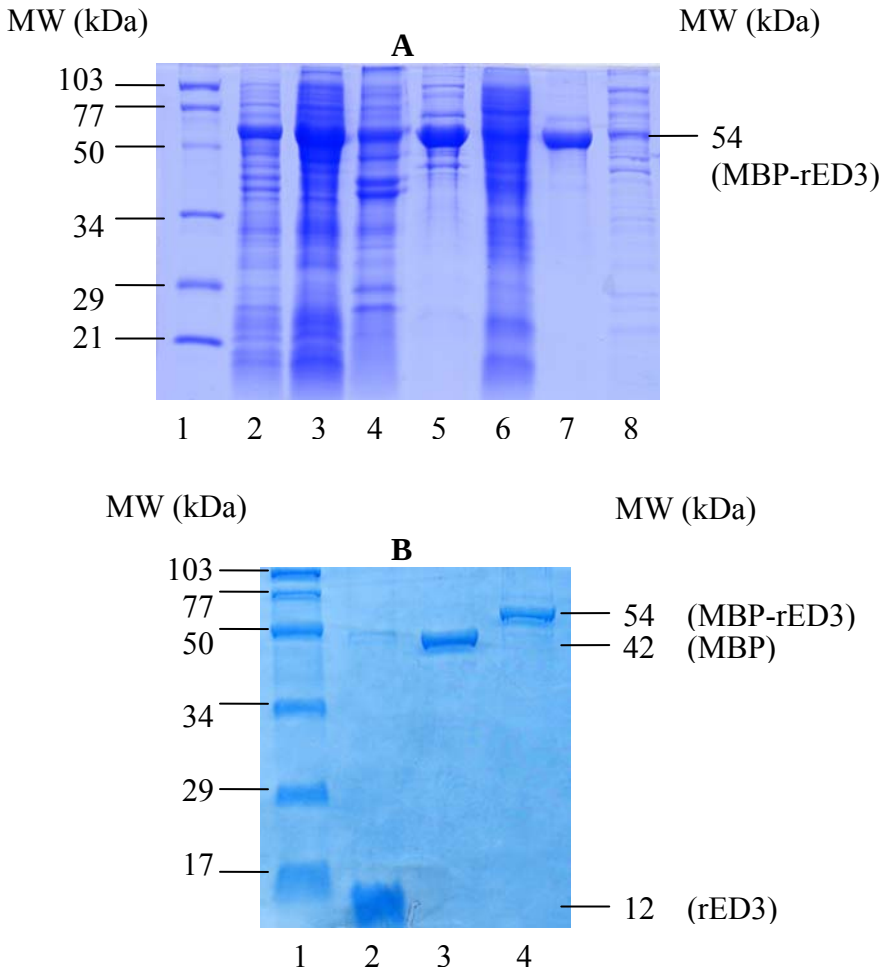

Figure 2.3. Induction and purification of rED3-WT from an MBP fusion protein. **A**, (1) Low molecular weight ladder; Induction with IPTG for 1 h (2) and 4 h (3); (4) pellet after breaking cells; (5) fusion protein after Bio-Rex 70; (6) remnants from Bio-Rex 70; (7) fusion protein from Amylose affinity chromatography and (8) remnants from Amylose resin. The SDS-PAGE 12 % was run for 60 min at 140 volts. **B**, SDS-PAGE 15 % after digestion of the MBP-rED3 fusion protein with Factor Xa and separation of the rED3 using size exclusion chromatography. Lane (1) is the low molecular weight ladders; lane (2) is the purified rED3 from Bio-Gel P30; lane (3) is MBP after cleavage and MBP-rED3 is the fusion protein before cleavage.

Table 2.1. Dissociation constants and Neutralization index (N.I.) for rED3s

	WNV strain	MAb-3A3 (Kd_{app}, nM)	MAb-5H10 (Kd_{app}, nM)	N.I.[a] MAb-5H10
rED3-WT	USA99b	0.17 ± 0.05	0.13 ± 0.03	2.3
rED3-K310T	AUS60	0.39 ± 0.01	0.16 ± 0.02	1.1
rED3-T332A	ISR53	1.00 ± 0.16	0.30 ± 0.03	0.9
rED3-T332K	MAD88 / SA58a	> 10.0	> 10.0	0.2

[a] N.I. is the $\log_{10}$ reduction in virus titer in the presence of MAb-5H10 compared to a culture medium only as control. Values were obtained from references 15 and 16.

By directly comparing the Kd_{app} values from the binding isotherm and neutralization index for MAb-5H10, for rED3-WT and the single mutants, we found an inverse correlation of systematic increase in Kd_{app} for decreasing neutralization index values as shown in Figure 2.5. For example, the mutation T332K has the lowest neutralization index value, indicating a high degree of resistance *in vitro*, and the highest Kd_{app}, indicating the weakest binding affinity. The opposite holds for rED3-WT. Therefore, the rED3s preserve their immunological properties as in its native environment in the intact viral particle.

These results validate the functional and biological significance of using the purified rED3s to assess the effects of single mutations on the solution biophysical properties of the rED3-WT in order to understand the basic mechanism of neutralization resistance induced by mutations.

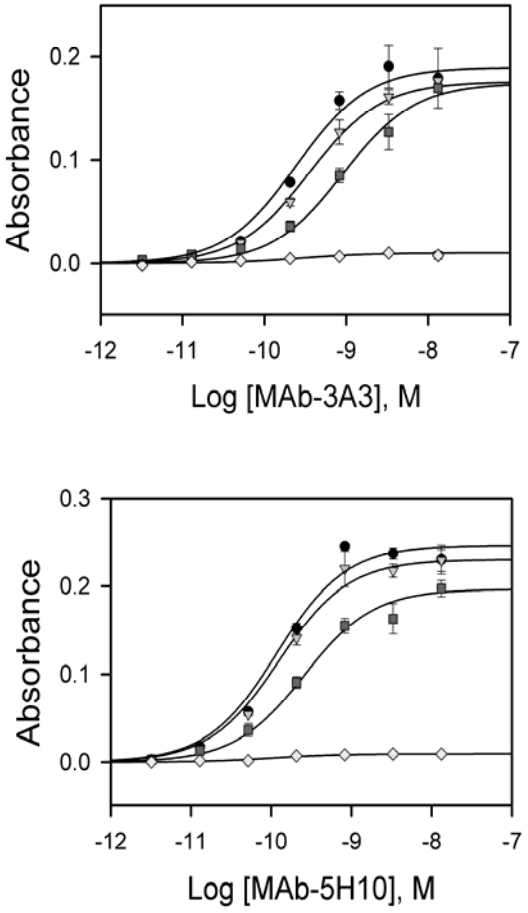

Figure 2.4. Binding isotherms of MBP-rED3 fusion protein to MAbs 3A3 and 5H10. The symbols correspond to rED3-WT (●), rED3-K310T (▼), rED3-T332A (■) and rED3-T332K (◆). The solid line correspond to the best fit to a single-site binding polynomial: Obs(x) = (Max·[MAb])/(Kd_{app} + [MAb]), where Max is the maximum signal observed, [MAb] is the concentration of MAb and Kd_{app} is the apparent dissociation constant.

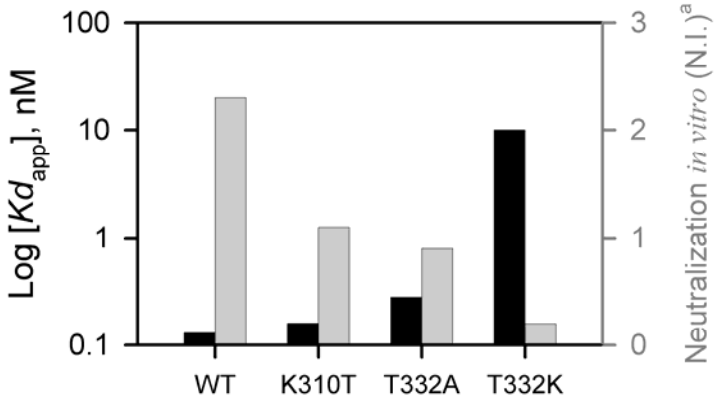

Figure 2.5. Correlation between *in vitro* MAb neutralization (grey bars) and MAb binding affinity (black bars). Neutralization index (N.I.) is the $\log_{10}$ reduction in virus titer in the presence of MAb-5H10 compared to a culture medium only as control. The values of N.I. were obtained with permission from Li, L., et al. (2005)[16]. Binding constants were obtained from the fitting of data in Figure 2.4 with MAb-5H10.

2.4.3 Protein structural stability

a. Chemical denaturation

Chemical unfolding of the rED3s with GdnHCl was monitored by changes in the fluorescence emission maximum intensity of the single tryptophan residue (at λ_{ex} = 295 nm) and in coordination with the fluorescence properties of the three tyrosine residues (at λ_{ex} = 280 nm). The emission wavelength was 340 nm. The three tyrosine residues in the rED3s are spatially located in different regions of the protein (see Figure 2.1). Therefore, monitoring the chemical unfolding at λ_{ex} = 280 nm provides an observable for global unfolding. Whereas monitoring the chemical unfolding at λ_{ex} = 295 nm provides an observable only on the local environment around the single tryptophan residue, at position 397.

All four proteins showed reversibility of GdnHCl-induced unfolding as judged by the similarity in maximum emission intensity and maximum emission wavelength

number between the native and refolded states. The refolded protein was obtained by dialyzing the GdnHCl extensively (5 times using 100-fold volumes of the sample). Within experimental error, the thermodynamic parameters obtained for rED3-WT and the mutants rED3-K310T, rED3-T332A and rED3-T332K were independent of the excitation wavelength. Figure 2.6 shows the chemical unfolding of rED3-WT and single mutants with λ_{ex} = 280 nm.

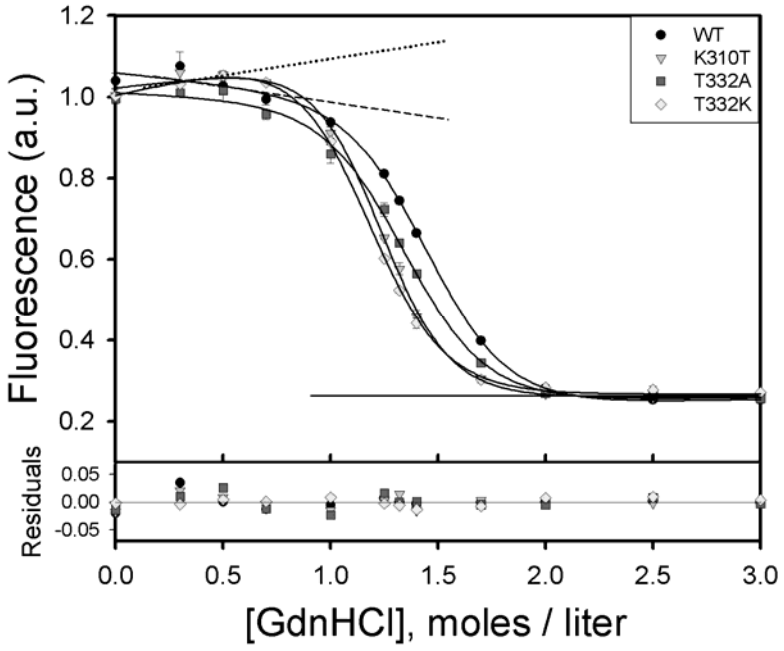

Figure 2.6. GdnHCl-induced chemical denaturation of rED3-WT and mutants rED3-K310T, rED3-T332A and rED3-T332K monitored by changes in the fluorescence emission maximum (λ_{ex} = 280 nm, λ_{em} = 340 nm). The unfolding measurements were performed in TN200 buffer at 20.0 °C. The solid lines are the nonlinear best fit to equation 2.3. The dashed curve (----) and dotted curve (·····) correspond to the pre-transition base lines for rED3-WT and rED3-K310T, respectively. The solid line (—) correspond to the post-transition base line.

From Figure 2.6 a shift is observed in the unfolding curves of the rED3s single mutants towards lower GdnHCl concentrations, indicating lower values of $C_{1/2}$. The $C_{1/2}$

for rED3-WT, rED3-K310T, rED3-T332A and rED3-T332K were 1.45 M, 1.23 M, 1.34 M and 1.16 M, respectively. The $C_{1/2}$ for rED3-T332K was the lowest and also shows the lowest stability ($\Delta G_{N\rightarrow U}^{H_2O}$) compared to rED3-WT. The $\Delta G_{N\rightarrow U}^{H_2O}$ of unfolding for rED3-T332K was 3.9 kcal·mol^{-1} whereas for rED3-WT was 4.6 kcal·mol^{-1}. Despite that the single mutant rED3-K310T has also a significantly lower $C_{1/2}$ compare to rED3-WT ($C_{1/2}$ = 1.23 M), there was no change in the stability. This similarity in stability (or $\Delta G_{N\rightarrow U}^{H_2O}$) can be attributed to a higher m_D-value for rED3-K310T (observed at λ_{ex} = 280 nm and at λ_{ex} = 295 nm). The higher m_D-value can be explained by the differences in the slope of the pre-transition or folded state base line. Figure 2.6 includes also the pre-transition base lines for rED3-WT (dashed line) and rED3-K310T (dotted line), which shows a considerable difference in their slopes. No differences were observed in the post-transition or unfolded state base line for all four proteins (solid line). Hence, variations in the m_D-value are predominantly caused by the changes in the slope of the pre-transition base lines. There were no significant differences between rED3-T332A and rED3-WT in $\Delta G_{N\rightarrow U}^{H_2O}$ and m_D-value. The thermodynamic parameters from the GdnHCl-induced unfolding are listed in Table 2.2.

To examine in more detail the effects of the single mutations on the thermodynamic properties of rED3-WT, we conducted thermal denaturation experiments as a function of GdnHCl concentration. This approach allowed us to extract all the thermodynamic parameters corresponding to the thermal unfolding (ΔH^0, ΔS^0, ΔCp^0 and T_m) and in addition, the m_D-value.

b. Thermal denaturation

Temperature-induced unfolding of the rED3s was monitored by changes in the fluorescence emission maximum intensity (λ_{em} = 340 nm) at two excitation wavelengths: 295 nm and 280 nm. Likewise to GdnHCl-induced unfolding, there were no observable

differences between the two excitation wavelengths in the thermodynamic parameters obtained from the thermal unfolding data analysis.

Table 2.2. Thermodynamic parameters for the GdnHCl-induced unfolding of rED3s[a]

	-m_D-value (kcal·mol^{-1}·M^{-1})	$C_{1/2}$ (M)[b]	$\Delta G_{N\rightarrow U}^{H_2O}$ (kcal·mol^{-1})[c]
rED3-WT	3.19 ± 0.19	1.45	4.64 ± 0.19
rED3-K310T	3.70 ± 0.17	1.23	4.52 ± 0.14
rED3-T332A	3.24 ± 0.20	1.34	4.35 ± 0.18
rED3-T332K	3.37 ± 0.08	1.16	3.91 ± 0.06

[a] Obtained by global analysis of the chemical unfolding data at λ_{ex} = 280 nm and λ_{ex}= 295 nm. [b] The error correspond to ± 0.01 M. [c] $\Delta G_{N\rightarrow U}^{H_2O}$ (at 20 °C) was calculated using the corresponding m_D-value and $C_{1/2}$ according to the formula $\Delta G = \Delta G_{N\rightarrow U}^{H_2O} + m_D[D]$; where $\Delta G = 0$ and $[D]$ is equal to $C_{1/2}$.

Figure 2.7 shows typical fluorescence emission scans (λ_{ex} = 280 nm) at increasing temperatures (black curves). Upon denaturation of the rED3s, it is observed –in addition to a decrease in the fluorescence intensity– a red shift in the maximum fluorescence intensity from 340 nm to approximately 360 nm, indicating that the single tryptophan residue in the rED3 is fully exposed to the solvent. All rED3s were thermally reversible as judged by the same maximum emission intensity and wavelength number (located at 340 nm) during the refolding process (Figure 2.7, grey curves).

We observed very good agreement between the temperature-induced unfolding and GdnHCl-induced unfolding as reflected in the similarity of the rank orders of $C_{1/2}$ and T_m for the rED3s. The rED3-WT had the highest T_m value (60.9 °C) concomitantly with the highest $C_{1/2}$ (1.45 M), followed by rED3-T332A (T_m = 59.0 °C and $C_{1/2}$ = 1.34

M) and finally rED3-T332K and rED3-K310T with the lowest T_m and $C_{1/2}$ values (~57 °C and ~1.2 M). For instance, Figure 2.8 shows the temperature-induced unfolding curves of rED3-WT and rED3-T332K in the absence of GdnHCl. The unfolding curves with closed symbols were obtained with an excitation wavelength of 280 nm; whereas open symbols correspond to 295 nm. There is a clear shift of the curves corresponding to rED3-T332K –at both wavelengths– towards lower temperatures, due to the lower T_m compared to rED3-WT. The solid lines are the nonlinear best fit of equation 2.4 but neglecting the terms related to denaturant-induced effects on the stability.

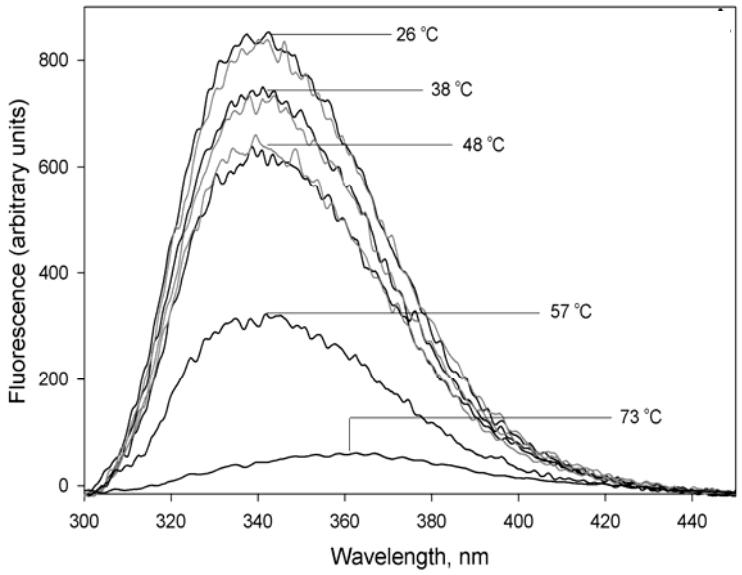

Figure 2.7. Representative fluorescence emission spectra of rED3-T332K as a function of increasing temperatures (black lines). Emission spectra during the refolding process of rED3-T332K (grey lines).

To extract all the thermodynamic parameters that characterize the rED3s –namely ΔH^0, ΔS^0, ΔCp^0 and T_m– we conducted the temperature-induced unfolding experiments with increasing concentrations of GdnHCl (0.0 M → 1.0 M). Figure 2.9 shows representative unfolding curves at increasing concentration of GdnHCl. Increasing the

concentration of GdnHCl had two major effects in the temperature-induced unfolding curves: 1) shifted the T_m towards lower temperatures; and 2) broadened the folded-unfolded transition region (concomitantly with a decrease in ΔH^0). These two effects allowed us to determine ΔCp^0 and m_D-value by nonlinear global fitting of the data to equation 2.4. The results obtained are listed in Table 2.3.

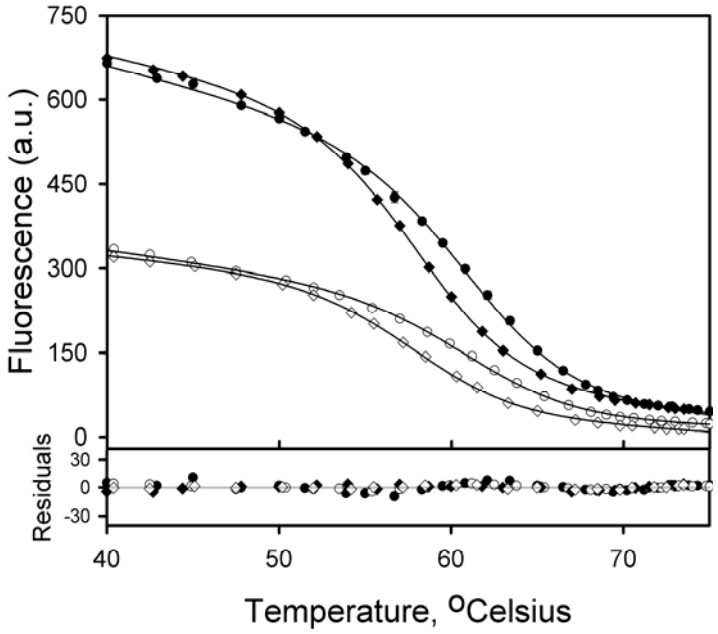

Figure 2.8. Temperature-induced unfolding of rED3-WT (○ or ●) and rED3-T332K (◇ or ◆) monitored by changes in the fluorescence emission maximum (λ_{em} = 340 nm) with λ_{ex} = 280 nm (filled symbols) and λ_{ex} = 295 nm (open symbols). The solid lines are the best fit to equation 2.4**.**

Differences in the thermodynamic parameters between rED3-WT and single mutants were observed. All rED3s had similar ΔH^0 of unfolding (~73.0 kcal·mol^{-1}) except for rED3-T332A (~67.0 kcal·mol^{-1}). No significant differences in ΔS^0 were observed between the rED3s. Removal or addition of a surface charged residue as in the case of the mutations rED3-K310T and rED3-T332K generated an increase in ΔCp^0 in comparison to rED3-WT and rED3-T332A. An increase of ΔCp^0 by changes of the

protein surface chemistry (i.e. addition of electrostatic interactions) has been widely observed[55-57]. In addition, the single mutations rED3-K310T and rED3-T332K had higher m_D-values, as observed in the GdnHCl-induced unfolding (see Tables 2 and 3). Therefore, the differences in the m_D-values between rED3-WT and, rED3-K310T and rED3-T332K observed in the GdnHCl-induced unfolding are qualitatively validated by the m_D-values determined in the temperature-induced unfolding.

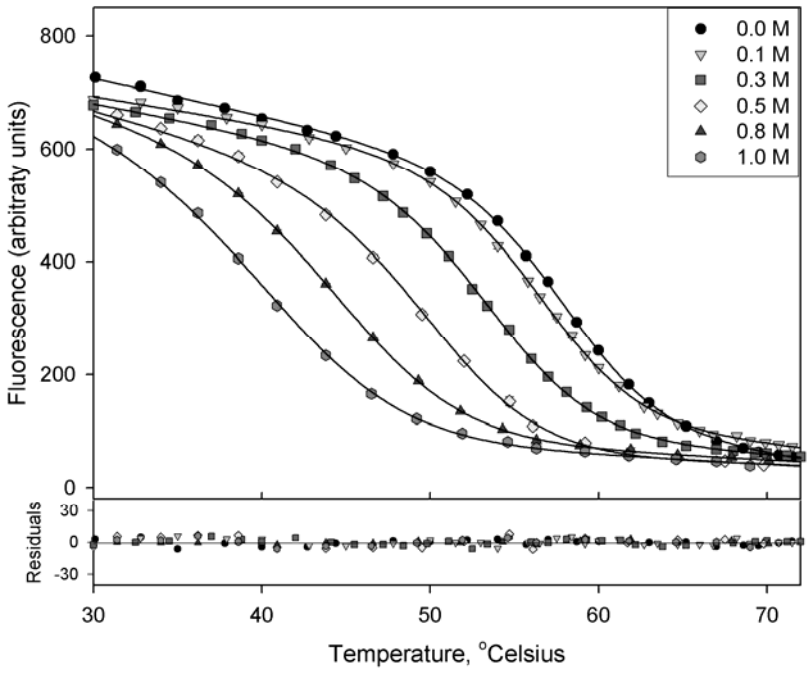

Figure 2.9. Temperature-induced unfolding of rED3-WT (1 μM) as a function of [GdnHCl] (λ_{em} = 340 nm, λ_{ex} = 280 nm). The [GdnHCl] were 0.0 M (●), 0.1 M (▼), 0.3 M (■) 0.5 M (◆), 0.8 M (▲) and 1.0 M (●). The solid lines are the best fit to equation 2.4**.**

Table 2.3. Thermodynamic parameters for the Temperature-induced unfolding of rED3s[a]

	ΔH^0 (kcal·mol^{-1})	ΔS^0 (kcal·mol^{-1}·K^{-1})[b]	ΔCp^0 (kcal·mol^{-1}·K^{-1})	T_m (C)[c]	-m_D-value (kcal·mol^{-1}·M^{-1})	$\Delta G_{N\rightarrow U}^{20\,^{\circ}C}$ (kcal·mol^{-1})[d]
rED3-WT	72.1 ± 2.5	0.22 ± 0.01	1.66 ± 0.21	60.9	3.00 ± 0.09	4.50
rED3-K310T	75.2 ± 1.0	0.23 ± 0.01	1.99 ± 0.10	57.3	3.30 ± 0.10	4.13
rED3-T332A	67.1 ± 1.1	0.20 ± 0.01	1.55 ± 0.17	59.0	2.92 ± 0.06	4.18
rED3-T332K	73.8 ± 1.0	0.23 ± 0.01	2.06 ± 0.10	57.7	3.19 ± 0.10	3.81

[a] Thermodynamic parameters obtained with λ_{ex} = 280 nm. [b] Calculated according to the following relation: $\Delta S^0 = \Delta H^0/(T_m + 273.15)$. [c] The error correspond to ± 0.1 °C. [d] Calculated using equation 2.5, with T = 20 °C, and [D] = 0.

2.4.4 Dynamic motions and communication between distant sites in ED3

a. Tryptophan fluorescence quenching with acrylamide: a neutral quencher

The rED3-WT and single mutants rED3-K310T, rED3-T332A and rED3-T332K have a single tryptophan residue at position 397. Tryptophan quenching experiments with acrylamide (neutral quencher) were performed by titrating aliquots of 5.0 M stock solution to a 2 μM protein solution, taking into consideration the effects of buffer absorption, dilution, and photobleaching (see material and methods). In addition, we performed fluorescence emission scans every 0.025 M of acrylamide up until 0.3 M. We observed a constant shape of the fluorescence emission spectra and a constant maximum intensity wavelength meaning that the presence of the quencher (acrylamide) at concentrations up to 0.3 M does not perturb the fold of the proteins studied.

Figure 2.10 shows the raw data from a typical tryptophan quenching experiment with acrylamide: Curve *1* shows a stable signal over the course of the titration indicating that there are not photobleaching effects. Curve *2* corrects for the dilution of the protein concentration when aliquots of acrylamide were added. Curve *3* is the signal from the titration before any correction and curve *4* corrects for baseline of buffer absorption. Each point in curves *1* thru *4* corresponds to an average of 5 single reads. Under this experimental set up, the actual effect of acrylamide on the single tryptophan fluorescence intensity can be determined by using equation 2.7.

The corrected data was plotted in the form of the Stern-Volmer equation, as shown in Figure 2.11. Acrylamide fluorescence quenching of free NATA and the single tryptophan residue in rED3-WT showed a significant curvature, indicating the presence of both dynamic (collisional) and static processes. The values of K_{SV} (collisional quenching) and V (static quenching) for rED3-WT were 5.6 ± 0.2 M^{-1} and 1.5 ± 0.1 M^{-1}, respectively (closed circles). Division of $F_{Q=0}/F_Q$ by $e^{(V[Q])}$ corrects the steady-state intensities for the static component and reveals the dynamic (collisional) portion of the observed quenching[58], which is included in Figure 2.11 (open circles). The K_{SV} and V for

NATA, under the same conditions, were 17.0 ± 0.1 M^{-1} and 2.2 ± 0.1 M^{-1} respectively, in agreement with previously reported values[59,60] (Figure 2.11, closed and open triangles). The 3-fold difference between the K_{SV} for NATA and rED3-WT indicates that the single tryptophan residue in rED3-WT is located in a hydrophobic environment. This conclusion will be further supported by the blue-shifted emission fluorescence spectrum of rED3-WT compared to both free NATA and unfolded rED3-WT (Figure 2.18).

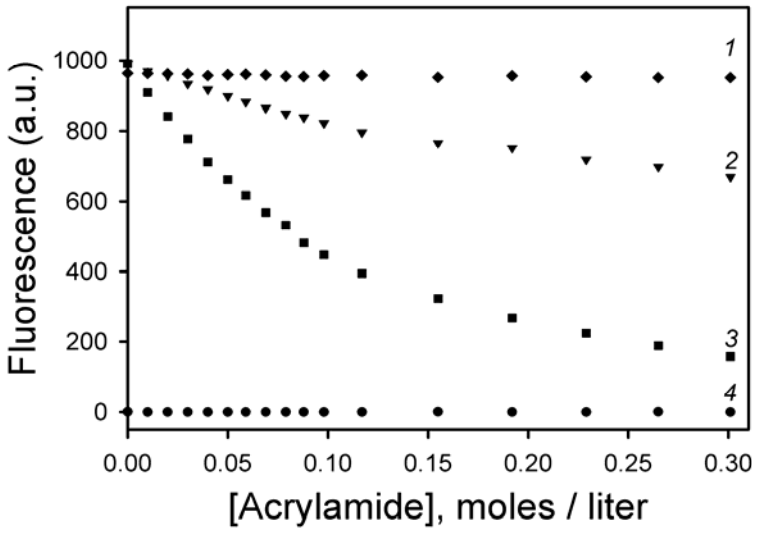

Figure 2.10. Raw data of rED3-WT tryptophan fluorescence quenching experiment with acrylamide, representing a typical quenching experiment. Curve *1* (◆) shows the stability of the fluorescent signal over the course of the experiment. Curve *2* (▼) shows the effect of dilution on the fluorescent signal by each titrating point. Curve *3* (■) shows the effect of acrylamide in the fluorescent signal of the tryptophan before correction for baseline, dilution or photobleaching. Curve *4* (●) is the buffer fluorescent signal. The protein concentration was 1 μM.

Considering NATA as a model for a fully solvent-exposed tryptophan residue with 100 % accessible surface area (ASA); then, the K_{SV} for NATA would represent the Stern-Volmer quenching constant of a tryptophan residue fully accessible on the surface of the protein. The value of K_{SV} determined for rED3-WT (K_{SV} = 5.6 ± 0.2 M^{-1}) is ~30 %

of the corresponding value for NATA, which indicates that in the native state of rED3-WT, the single tryptophan residue has 30 % ASA to the solvent. In agreement with the experimental ASA (ASA_{exp}), the calculated ASA[46] (ASA_{cal}) of the single tryptophan residue, based the rED3-WT NMR structure[38] (PDB 1S6N), is 27 ± 3% (see Table 2.4).

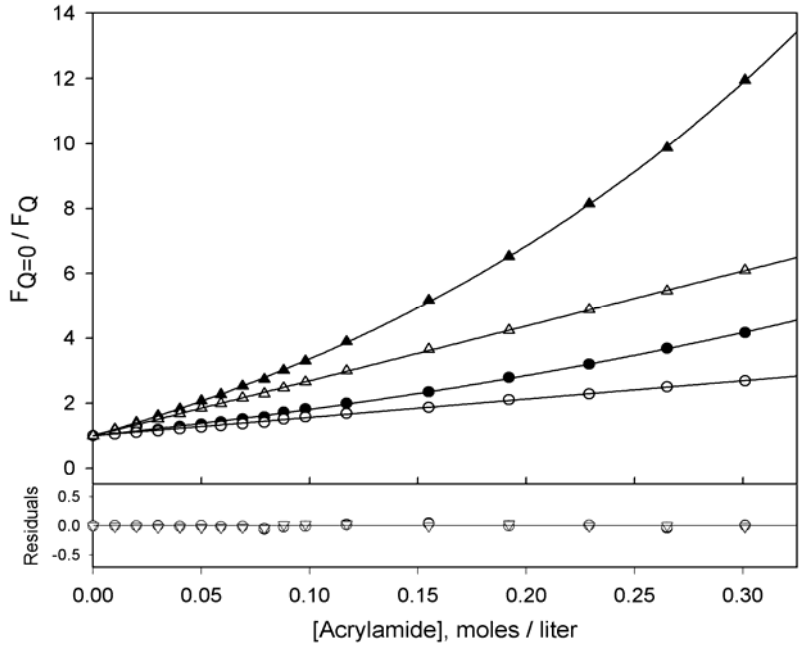

Figure 2.11. Stern-Volmer plots for acrylamide quenching of free NATA fluorescence and rED3-WT single tryptophan fluorescence in TN200 buffer at 20 °C (λ_{ex} = 295 nm, λ_{em} = 340 nm). Closed triangles and circles (▲ and ●) show the deviation of the classic Stern-Volmer behavior for NATA and rED3-WT (respectively). The solid lines are the nonlinear best fit to equation 2.9. Open triangles and circles (△ and ○) show the classic (collisional) Stern-Volmer behavior for NATA and rED3-WT (respectively) and were analyzed using equation 2.8.

The classic Stern-Volmer plots for rED3-WT and mutants rED3-K310T, rED3-T332A and rED3-T332K are shown in Figure 2.12. The values of K_{SV} for rED3-K310T (K_{SV} = 5.6 ± 0.2 M^{-1}) and rED3-WT were the same within experimental error; however,

the K_{SV} for rED3-T332A and rED3-T332K were 6.2 ± 0.3 M^{-1} and 6.8 ± 0.4 M^{-1}, respectively. These results indicate that the ASA_{exp} of tryptophan to solvent is higher for the single mutants at position T332 (rED3-T332A/K) and that the mutant rED3-K310T has the same tryptophan ASA_{exp} for rED3-WT (see Table 2.4).

Table 2.4. Acrylamide Stern-Volmer quenching constants (K_{SV})

	K_{SV} (M^{-1})	V (M^{-1})	ASA_{exp} (%)[a]	Max. (nm)[b]
NATA	17.0 ± 0.1	2.2 ± 0.1	100	358
rED3-WT	5.6 ± 0.2	1.5 ± 0.1	32 ± 2 (27 ± 3)[c]	341
rED3-K310T	5.6 ± 0.2	1.2 ± 0.1	32 ± 2	341
rED3-T332A	6.2 ± 0.3	1.2 ± 0.1	36 ± 2	344
rED3-T332K	6.8 ± 0.4	1.1 ± 0.2	40 ± 2	344

[a] The ASA_{exp} to solvent was determined using NATA as reference. [b] Maximum emission intensity (peak wavelength number) with λ_{ex} = 295 nm. [c] The value in parenthesis correspond to ASA_{cal} using the NMR structure of rED3-WT[19].

To validate the differences in the K_{SV} (and therefore in the ASA_{exp}) between the mutants rED3-T332A/K and rED3-WT, we determined the maximum fluorescence intensity (peak value in nm) of several (≥ 7) emission scans for all four proteins at λ_{ex} = 295 nm. We observed a statistically significant increase [$F(3, 37) = 18.2$, $p < 0.05$] in the maximum fluorescence intensity, or red shift, for the single mutants rED3-T332A/K compared to rED3-K310T and rED3-WT. There were no significant differences between rED3-WT and rED3-K310T. Figure 2.13 shows the wavelength number of the average peak value (± 2-fold the standard deviation or ± 2 S.D.) of the fluorescence emission spectra for rED3-WT, rED3-K310T, rED3-T332A and rED3-T332K; including the number of scans (N) that were subjected to the statistical analysis. The increase in peak

wavelength number for rED3-T332A and rED3-T332K indicates that the environment surrounding the single tryptophan residue for these mutants is more polar[62], which is in accordance with the observation that the single tryptophan residue is more accessible (higher K_{SV}).

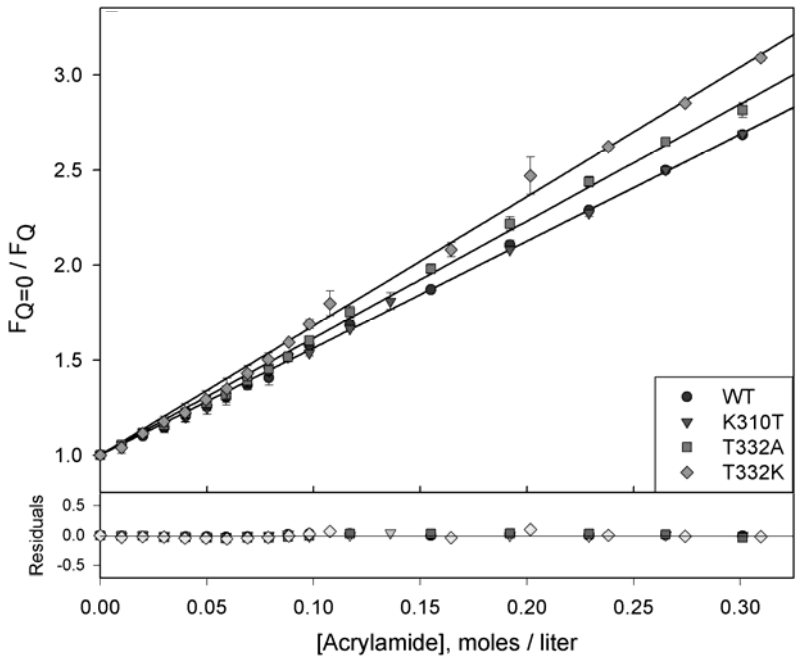

Figure 2.12. Stern-Volmer plots for acrylamide quenching of rED3-WT (●), rED3-K310T (▼), rED3-T332A (■) and rED3-T332K (◆). The solid lines are the best fit to $F_{Q=0}/F_Q = (1 + K_{SV}[Q])$. The sample concentrations were 2 μM.

b. Tryptophan fluorescence quenching with iodide, thallium and trichloroethanol (TCE): ionic and hydrophobic quenchers

To further examine the effect of the single mutations in the solution structure of rED3-WT, we investigated the physicochemical properties of the surrounding environment of the single tryptophan residue by performing fluorescence quenching experiments using a hydrophobic and two ionic (charged) quenchers. The experimental

procedure when using hydrophobic and ionic quenchers was identical to the procedure during the tryptophan quenching experiments with acrylamide. That means that controls for the deviation to the fluorescence intensity due to buffer emission, sample dilution and photobleaching were considered (see Figure 2.10).

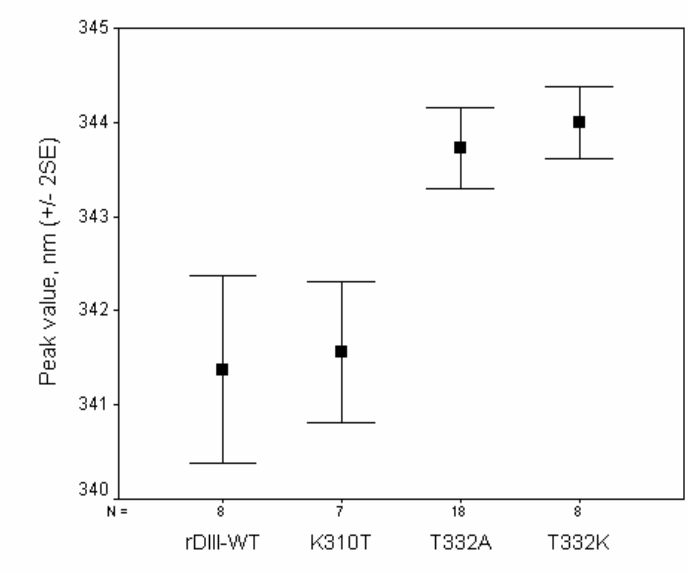

Figure 2.13. Average wavelength number of the maximum fluorescence emission spectra of the single tryptophan residue for rED3-WT, rED3-K310T, rED3-T332A and rED3-T332K. "N" is the number of spectra used in the statistical analysis. The bars represent a 2-fold standard deviation. All the samples were prepared in TN200 at 20 °C (λ_{ex} = 295 nm).

In proteins, the efficiency of a quencher will depend on both the physical accessibility of the fluorophore in the macromolecular matrix and/or the physicochemical properties of the environment surrounding the fluorophore (i.e. charge density, hydrophobicity, etc.). Therefore, when exploring the physicochemical properties of the environment surrounding a fluorophore by using hydrophobic or ionic quenchers, it is imperative to take into account the degree of exposure (ASA_{exp}) or physical accessibility of the fluorophore to the quencher. Acrylamide is a neutral quencher; thus, it provides principally information about the ASA_{exp} to the solvent, and the effects of the

physicochemical properties of the environment surrounding a fluorophore are minimized. By considering the K_{SV} of acrylamide as a correction factor for the degree of quenching of a hydrophobic or ionic quencher, it is possible to determine the actual effect of the physicochemical properties of the environment on the quenching process, and the effect of ASA_{exp} is minimized. The corrected degree of quenching for a hydrophobic or ionic quencher is called quenching efficiency. The efficiency of a hydrophobic or ionic quencher (Q_1) relative to the accessibility of the tryptophan residue, determined by acrylamide quenching (Q_2), is defined by equation 2.13:

$$\frac{\gamma_1}{\gamma_2} = \frac{Ksv_1^b}{Ksv_1^f}\frac{Ksv_2^f}{Ksv_2^b} \qquad (2.13)$$

The subscripts *1* and *2* refer to the quencher Q_1 (TCE or ionic quenchers) and the reference quencher Q_2 (acrylamide), respectively. The superscripts *b* and *f* refer to fluorophore in the bound (i.e. in the native state of the rED3s) and free states (free NATA in solution). The quantity γ_1/γ_2 reflects the accessibility of the fluorophore to Q_1 as compared with Q_2 considering steric hindrances and physicochemical properties of the surroundings of the fluorophore[63]. A value of $\gamma_1/\gamma_2 = 1$ means that the physicochemical properties of the environment are not affecting the efficiency of the quencher (ionic or hydrophobic). Values of $\gamma_1/\gamma_2 > 1$ or < 1 indicate that the physicochemical nature of the environment of the fluorophore is increasing or decreasing the efficiency of the quencher, respectively.

The tryptophan fluorescence quenching with trichloroethanol (TCE) explores the hydrophobicity of the tryptophan residue environment[64]. The K_{SV} values for rED3-WT, rED3-K310T, rED3-T332A and rED3-T332K were 6.3 ± 0.2 M^{-1}, 6.0 ± 0.1 M^{-1}, 7.1 ± 0.3 M^{-1} and 5.7 ± 0.1 M^{-1}, respectively. Using equation 2.13, we determined the quenching efficiency of TCE relative to acrylamide ($\gamma_{TCE}/\gamma_{Acrylamide}$). The $\gamma_{TCE}/\gamma_{Acrylamide}$ for rED3-WT was 1.2, indicating that TCE is a more efficient quencher that acrylamide, in agreement with previous results that show that the single tryptophan residue is located in a

predominantly hydrophobic environment. The $\gamma_{TCE}/\gamma_{Acrylamide}$ for rED3-K310T and rED3-T332A was identical to rED3-WT ($\gamma_{TCE}/\gamma_{Acrylamide}$ = 1.2). The $\gamma_{TCE}/\gamma_{Acrylamide}$ for rED3-T332K was 0.9. As expected, the lower value of $\gamma_{TCE}/\gamma_{Acrylamide}$ is in agreement with the observation that the single tryptophan residue in rED3-T332K is significantly more accessible to the solvent and therefore, the surrounding environment is more hydrophilic.

To investigate the surface electrostatics around the single tryptophan residues in the rED3s we used iodide and thallium as negatively and positively charged quenchers, respectively. Both quencher molecules have comparable diffusion coefficients, and their indole fluorescence quenching efficiencies are identical[65]. Therefore, any differences in the quenching efficiencies between iodide and thallium can be attributed to differences in the physicochemical properties (i.e. surface electrostatics) surrounding the tryptophan residue. There was a significant difference between the K_{SV} of iodide and thallium for rED3-WT with 3.1 ± 0.3 M^{-1} and 20.9 ± 0.7 M^{-1}, respectively. Relative to acrylamide, thallium is a quencher ~3-fold more efficient than iodide and TCE (see Table 2.5). This result indicates that the indole group of the single tryptophan residue in rED3-WT is surrounded principally by negatively charged residues. Despite that the tryptophan residue is located near basic residues in the primary sequence, inspection of the rED3-WT NMR structure shows that the tryptophan residue is spatially located very close to 3 negatively charged residues: ASP317, ASP281 and GLU376.

The single mutations affect distinctly the K_{SV} and quenching efficiencies of iodide and thallium compared to rED3-WT. The single mutant rED3-K310T had a lower K_{SV} for iodide (1.8 ± 0.3 M^{-1}) concomitantly with a higher K_{SV} for thallium (24.5 ± 0.3 M^{-1}), which corresponds to a lower iodide quenching efficiency ($\gamma_{iodide}/\gamma_{Acrylamide}$ = 0.5) and higher thallium quenching efficiency ($\gamma_{thallium}/\gamma_{Acrylamide}$ = 3.6). On the contrary, the single mutant rED3-T332A generated a significant increase in iodide quenching (K_{SV} = 5.0 ± 0.2 M^{-1}) corresponding to a higher iodide quenching efficiency of 1.3. However, there was no difference in thallium K_{SV} and quenching efficiency (20.9 ± 0.5 M^{-1} and 2.8, respectively). In Figure 2.14 is shown the different behavior of rED3-K310T and rED3-T332A. The mutant rED3-T332K had a lower K_{SV} for iodide and higher K_{SV} for thallium

(1.5 M^{-1} and 23.3 M^{-1}, respectively) that corresponds to a significant lower quenching efficiency for iodide ($\gamma_{iodide}/\gamma_{Acrylamide}$ = 0.4) and no significant difference for thallium ($\gamma_{thallium}/\gamma_{Acrylamide}$ = 2.8). In Table 2.5 is summarized the K_{SV} and quenching efficiencies for tryptophan fluorescence using iodide, thallium and TCE.

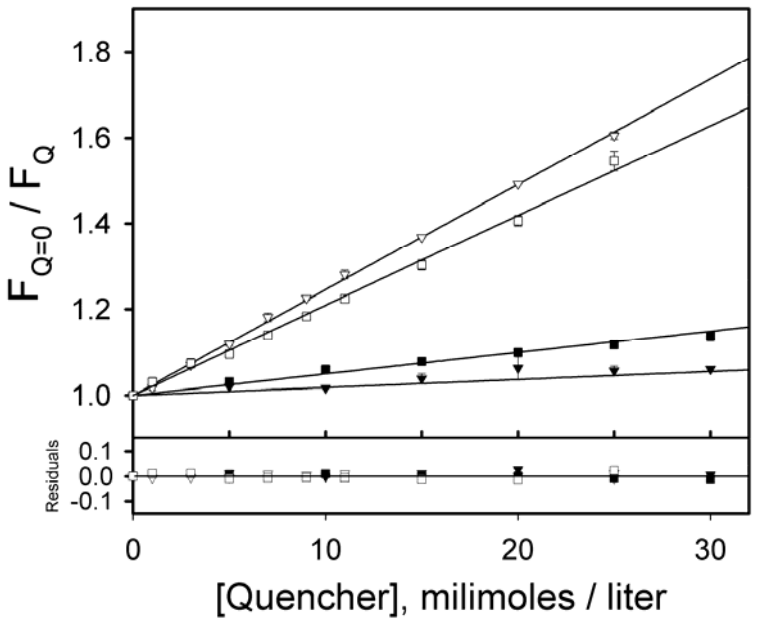

Figure 2.14. Stern-Volmer plots of the tryptophan fluorescence quenching of rED3-K310T (triangles) and rED3-T332A (squares) with iodide and thallium in TN200 buffer (or Tris-Acetate buffer) at 20 °C (λ_{ex} = 295 nm, λ_{em} = 340 nm). Closed symbols (▼ and ■) correspond to iodide quenching whereas open symbols (▽ and □) correspond to thallium quenching. The solid lines are the nonlinear best fit to equation 2.8. The protein concentration was 1 μM.

Table 2.5. Stern-Volmer quenching (K_{SV}) and quenching efficiencies ($\gamma_{Quencher}/\gamma_{Acrylamide}$) by iodide, thallium and TCE

	K_{SV} (M^{-1})			$\gamma_{Quencher}/\gamma_{Acrylamide}$[a]		
	iodide	thallium	TCE	iodide	thallium	TCE
NATA	10.6 ± 0.1	20.7 ± 0.7	15.6 ± 0.3	1.0	1.0	1.0
rED3-WT	3.1 ± 0.3	20.9 ± 0.7	6.3 ± 0.2	0.9	3.0	1.2
rED3-K310T	1.8 ± 0.3	24.5 ± 0.3	6.0 ± 0.1	0.5	3.6	1.2
rED3-T332A	5.0 ± 0.2	20.9 ± 0.5	7.1 ± 0.3	1.3	2.8	1.2
rED3-T332K	1.5 ± 0.1	23.3 ± 0.3	5.7 ± 0.1	0.5	2.8	0.9

c. Tyrosine accessible surface area to the solvent

In addition to study the effect of single mutations in the physicochemical properties of the single tryptophan residue and its local surrounding environment, we studied the effects of mutations on the degree of exposure to the solvent of the three tyrosine residues by UV absorption spectroscopy. The three tyrosine residues are located in spatially distinct regions of the rED3s (see Figure 2.1). Therefore, they provide an observable for the effect of single mutations at a global surface level compared to the local environment of the single tryptophan residue. In Figure 2.15 is shown a typical UV absorbance spectrum of the rED3s in the native state (solid line) and fully unfolded state (dotted line), which correspond to the single mutant rED3-T332K. It is evident that there is a spectral shift from the native to the unfolded state along with a shift in the maximum absorption band (native state = 278.8 nm and unfolded state = 276.2 nm). The peak distance in the UV absorption spectra for rED3-WT, rED3-K310T and rED3-T332K between the native and unfolded states were 2.80 ± 0.65 nm, 2.10 ± 0.37 nm and 2.60 ± 0.16 nm. In the case of the single mutant rED3-T332A, the peak distance was shorter with a value of 1.30 ± 0.61 nm.

The UV spectral shift of a protein (i.e. native vs. unfolded, Figure 2.15) is due mainly to changes in the spectral properties of tyrosine residues because of its greater sensitivity to the solvent, compared to tryptophan residues[66,67]. Thus, the longer the peak distance between the native and unfolded states of the rED3s, the larger is the change in the environment polarity of the tyrosine residues. Therefore, the shortest peak distance observed for rED3-T332A implies that the change in the solvent exposure of the three tyrosine residues upon denaturation is smaller or simply that the fractional exposure of the three tyrosine residues is larger. This last observation is further supported by the tyrosine fractional exposure determined by using the 2nd-derivative UV absorbance spectrum as described in the materials and methods section.

The 2nd derivative UV absorbance spectra of rED3-T332K in the native (solid line) and unfolded (dotted line) states are shown in Figure 2.16. The two peak-to-peak distances, necessary to calculate the fractional exposure (α) of the three tyrosine residues,

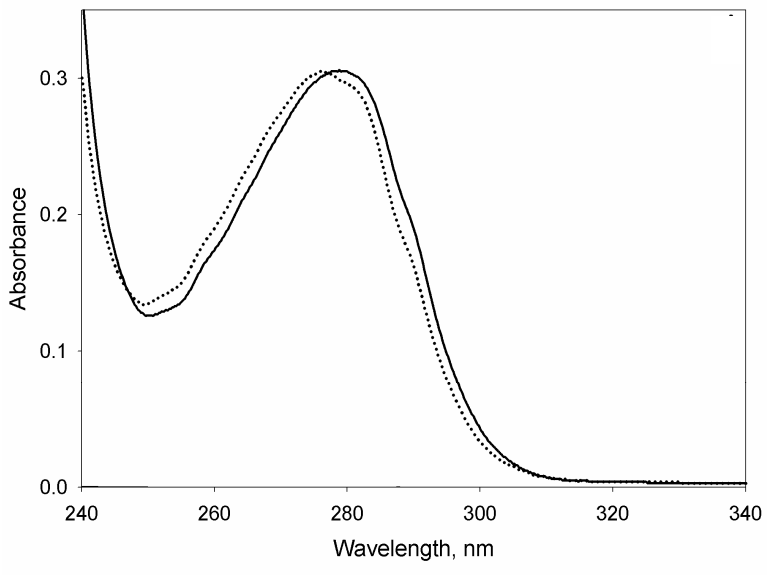

Figure 2.15. Absorbance spectra of rED3-T332K in TN200 buffer (solid line) and in 4.0 M GdnHCl (dotted line) at 20 °C.

are labeled '*a*' (287-283 nm) and '*b*' (295-290.5 nm). A close inspection of the peak '*a*' in Figure 2.16 shows a significant change in the absorption amplitude as compared to the peak '*b*' when the protein is in the unfolded vs. folded state (dotted vs. solid line, respectively). It was shown previously that the changes in the amplitude of peak '*a*' are due only to changes in the polarity of the tyrosine residues and not to tryptophan residues[49]. Therefore, this method allows the selective investigation of the tyrosine residues of a protein without the need of chemical modification of the protein that can alter the native conformation.

Within experimental error, there is not a significant difference in the fractional exposure of the three tyrosine residues between rED3-WT and the mutants rED3-K310T and rED3-T332K. The value of α for rED3-WT was 0.22 ± 0.04, which is in agreement with the calculated average ASA of the three tyrosine residues using the NMR structure ($<ASA_{cal}>$ = 23 ± 7 %, equivalent to α = 0.23 ± 0.2). The value of α for rED3-K310T and rED3-T332K was 0.17 ± 0.03. In the case of the single mutant rED3-T332A, α was

higher with a value of 0.31 ± 0.04, meaning that the average degree of exposure to the solvent of the three tyrosine residues is higher. This last result is in agreement with the observed shortest peak distance between the native and unfolded states UV absorbance spectra for rED3-T332A. The results are listed in Table 2.6.

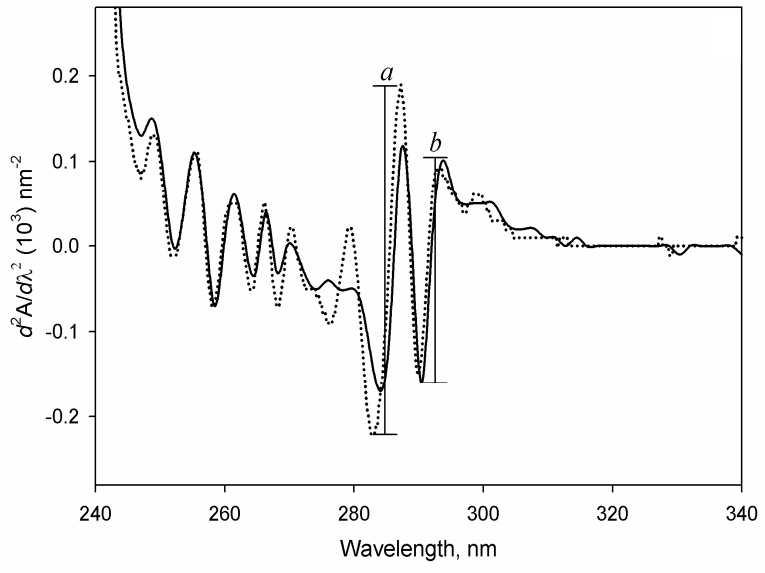

Figure 2.16. 2nd-derivative absorbance spectra obtained from Figure 2.15. The solid line and dotted lines correspond to the native and unfolded states, respectively. Peaks labeled '*a*' and '*b*' correspond to the peak distances between 287-283 nm and 295-290.5 nm, respectively.

2.4.5 Global protein solution structure

a. CD spectroscopy

The secondary structure of rED3 wild type (rED3-WT) and single mutants was monitored by CD spectroscopy, as shown in Figure 2.17. All rED3s spectra had strong negative ellipticity at 216-217 nm and positive ellipticity at 198 nm (Figure 2.17),

Table 2.6. Fractional exposure (α) of the tyrosine residues of rED3s

	r_N	r_U	r_a	α[a]	<ASA_{cal}> (%)	$d_{N\rightarrow U}$ (nm)[b]
rED3-WT	1.05	2.08	0.75	0.22 ± 0.04 (0.7)[c]	23 ± 7 (0.69)[d]	2.80 ± 0.65
rED3-K310T	1.00	2.22	0.75	0.17 ± 0.03 (0.5)[c]	---	2.10 ± 0.37
rED3-T332A	1.25	2.34	0.75	0.31 ± 0.04 (1.0)[c]	---	1.30 ± 0.61
rED3-T332K	0.98	2.11	0.75	0.17 ± 0.04 (0.5)[c]	---	2.60 ± 0.16

[a] Determined using equation 2.10 as described in Materials and Methods. [b] $d_{N\rightarrow U}$ is the peak distance between the native and unfolded states of the rED3s in the UV absoprtion spectra. [c] Based of α, the number is parenthesis corresponds to the average number of tyrosine residues exposed to the solvent in the rED3. [d] The number in parenthesis is the summation of the three tyrosine residues ASA_{cal} in the rED3-WT: TYR302 (0.48), TYR329 (0.0) and TYR383 (0.21).

which are characteristic of β-strand structure[68]. The negative absorption bands at about 210 nm and 222 nm, distinctive of α-helical structure, were absent in all rED3s CD spectra[69]. This result indicates that in native conditions the secondary structure of the rED3s in solution is largely composed of β-strand structure, in agreement with the NMR structure of WN virus rED3-WT[19].

Despite the similarity of the CD spectra between rED3-WT and all single mutants, there are small but significant differences in the absorption bands that corresponds to β-turn structures around 205-207 nm[68,69]. Close inspection of Figure 2.17 shows that rED3-WT and rED3-K310T had very similar absorptions in that region,

whereas rED3-T332A and rED3-T332K had strong-negative and weak-positive absorptions, respectively.

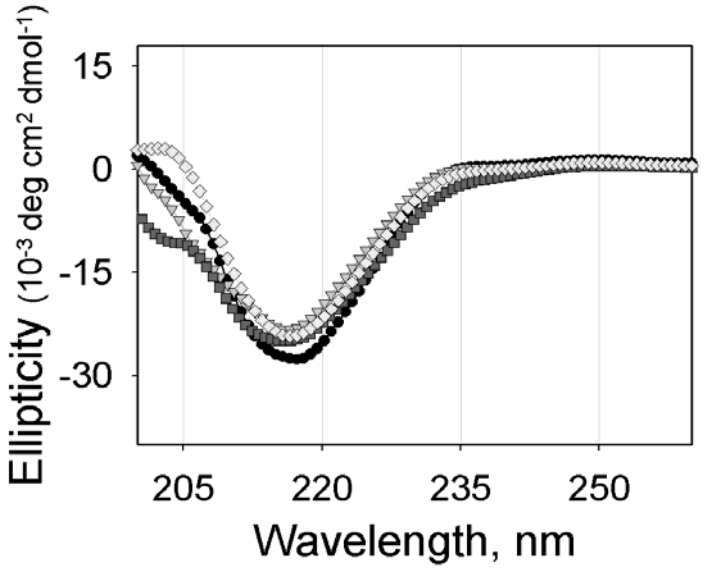

Figure 2.17. Circular dichroism spectra of rED3s in TN200 buffer at 20 °C corresponding to rED3-WT (●), rED3-K310T (▼), rED3-T332A (■) and rED3-T332K (◆). The protein concentration was 92 μM.

Quantitative analysis of the CD spectra by curve fitting confirms that these proteins are composed mostly of β-strand and of β-turn structures. The secondary structure fractional content was determined by using two different curve fitting procedures and both methods gave very similar results[70,71]. The summary of the secondary structure fractional content is listed in Table 2.7. None of the rED3s have significant α-helical structure component and there were no differences in the amount of random structures. The secondary structure content between rED3-WT and rED3-K310T were very similar. However, the mutants rED3-T332A and rED3-T332K showed significant differences in their β-strand and β-turn content. The mutant rED3-T332A had lower β-strand (0.35 ± 0.04) and higher β-turn (0.30 ± 0.07) content; whereas rED3-T332K had higher β-strand (0.50 ± 0.02) and lower β-turn (0.15 ± 0.04) content.

Table 2.7. Secondary structure fractional content of rED3s[a]

	α-helix	β-strand	β-turn	random coil
rED3-WT	0	0.47 ± 0.06	0.17 ± 0.03	0.35 ± 0.09
rED3-K310T	0.04 ± 0.02	0.40 ± 0.07	0.21 ± 0.08	0.35 ± 0.11
rED3-T332A	0.08 ± 0.07	0.35 ± 0.04	0.30 ± 0.07	0.27 ± 0.06
rED3-T332K	0	0.50 ± 0.02	0.15 ± 0.04	0.35 ± 0.03

[a] Determined using the programs CDSSTR[70] and CONTIN/LL[71]. The error corresponds to the standard deviation of the fractional contents obtained from each program using two different reference set of proteins.

Although the content of β-strand secondary structure was very similar between all rED3s (except for rED3-T332A with a slightly lower value compare to rED3-WT), the differences in the absorption bands at about 205-207 nm indicates that the local environments of the β-turn structures of the single mutants were different. This last statement is supported by the fact that the mutations at position T332 (rED3-T332A/K) correspond to a residue spatially located in the loop that connects strand number 2 and 3 of the protein. Furthermore, the mutation at position K310 is located on the side of strand number 1 and probably do not perturb the β-turn environment as shown in the CD spectra (at 205 nm) and in the secondary structure fractional content.

b. Hydrodynamic properties

The hydrodynamic properties of the rED3s were monitored by sedimentation velocity and equilibrium. The weight-average sedimentation velocity coefficient, ($\overline{S}_{20,w}$), were 1.15, 1.09, 1.00 and 1.02 for rED3-WT, rED3-K310T, rED3-T332A and rED3-T332K, respectively. All rED3s sedimentation velocity data were fitted to a single species using the program SedFit[53]. The apparent molecular weights, determined by sedimentation equilibrium, were 11.3 kDa, 10.5 kDa, 11.3 kDa and 11.1 kDa for rED3-WT, rED3-K310T, rED3-T332A and rED3-T332K, respectively. These results indicate

that there are not significant differences between rED3-WT and single mutants in their hydrodynamic size or shape.

2.4.6 Local protein solution structure and surface chemistry

a. Protein fluorescence spectrum analysis

The corrected emission spectrum of rED3-WT in TN200 buffer at 20 °C is shown in Figure 2.18 (curve A) with a maximum intensity located at 341 nm (at λ_{ex} = 295 nm). The maximum emission intensity of rED3-WT in 4.0 M GdnHCl was significantly decreased and it is located at 357 nm (curve B). The red shift of 16nm between curves A and B indicates that the tryptophan residue in the native state of rED3-WT is largely located in a hydrophobic environment and that upon denaturation it is in full contact with the polar solvent. This statement is supported by the similarity of emission maxima between unfolded rED3-WT and free NATA, which is located at 358 nm (dashed curve C). The maximum emission intensity for the rED3 mutants rED3-K310T, rED3-T332A and rED3-T332K were located at 341 nm, 344 nm and 344 nm, respectively (spectra not shown), indicative that the tryptophan in all rED3 mutants is predominantly situated in a hydrophobic environment.

In order to analyze in detail the total protein fluorescence spectrum of the rED3s, it is necessary to examine the contribution of tryptophan and tyrosine residues to the total emission intensity. The rED3-WT and single mutants have 1 tryptophan and 3 tyrosine residues. The contribution of the 3 tyrosine residues to the total protein fluorescence spectrum can be estimated by subtraction of the emission spectrum of the tryptophan residue (at λ_{ex} = 295 nm) from the total protein emission spectrum (at λ_{ex} = 280 nm). Before subtraction, both spectra (λ_{ex} = 295 nm and λ_{ex} = 280 nm) need to be normalized at 370 nm, where only tryptophan residues emit[62]. Figure 2.19 (black lines) shows the emission spectrum of the tryptophan residue of rED3-WT with λ_{ex} = 295 nm (solid curve), the protein emission spectrum with λ_{ex} = 280 nm (dashed curve), and the

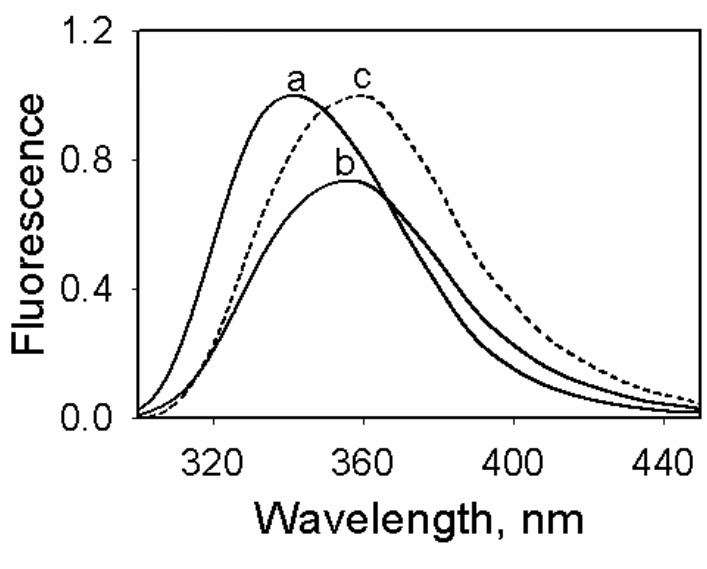

Figure 2.18. Corrected fluorescence emission spectra (at λ_{ex} = 295 nm) of rED3-WT in TN200 buffer (curve A); rED3-WT in 4.0 M GdnHCl dissolved in the same buffer (curve B) and free NATA in TN200 buffer (dashed curve, C). The concentration of rED3 and NATA was 1 μM. All measurements were performed at 20 °C. Curves A and C were normalized to their maximum. Curve B was normalized to the maximum of curve A.

difference spectrum [(λ_{ex} = 280 nm) — (λ_{ex} = 295 nm)], which corresponds to the emission spectrum of the 3 tyrosine residues (dotted curve). The tyrosine emission spectrum has a maximum located around 307 nm, characteristic of tyrosine residues in proteins[72]. Analogously, the emission spectrum of rED3-WT in 4.0 M GdnHCl (at λ_{ex} = 295 nm, λ_{ex} = 280 nm and the difference spectrum) is shown in Figure 2.19 (grey lines). It is clear that the emission intensity for the 3 tyrosine residues was greater in the unfolded state than in the folded state (dotted line, grey color vs. black color) with a maximum located at 303 nm, indicating that the tyrosine residues are in contact with the solvent[62]. Therefore, this result indicates that the protein is unfolded. Moreover, the increase in tyrosine emission intensity indicates that there is a fluorescence quenching process for the tyrosine residues in the native structure of rED3-WT that is disrupted upon denaturation of the protein.

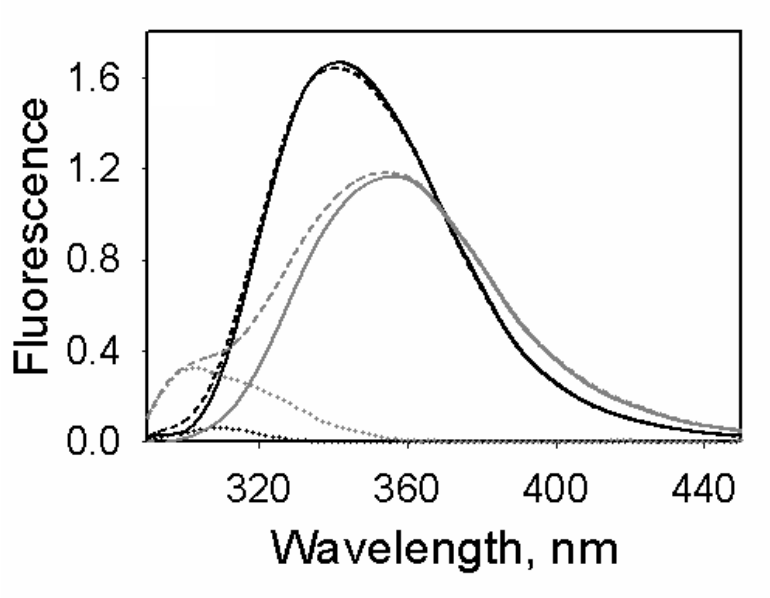

Figure 2.19. Corrected emission spectra of rED3-WT in TN200 buffer (black lines) and in 4.0 M GdnHCl (grey lines) dissolved in the same buffer. The spectra were recorded at different excitation wavelengths: λ_{ex} = 295 nm (solid curve) and λ_{ex} = 280 nm (dashed curve). The spectra were normalized at 370 nm. The difference spectra [(λ_{ex} = (280 nm) — (λ_{ex} = 295 nm)] correspond to the fluorescence emission spectrum of the 3 tyrosine residues (dotted curve).

Figure 2.20 shows the tyrosine fluorescence emission spectra of rED3-WT and the mutants rED3-K310T, rED3-T332A and rED3-T332K in the native and fully unfolded states (in 4.0 M GdnHCl). It is clear that for all the mutants there is also a dominant quenching process in the fluorescence emission intensity of the 3 tyrosine residues that were disrupted upon unfolding of the protein, as revealed by the significant increase in the fluorescence emission intensity. The effect of the mutations on the tyrosine fluorescence emission was negligible in the native state with a maximum intensity located around 307 nm for all three mutants. The maximum intensity for rED3-K310T and rED3-T332K was slightly higher than for rED3-WT and rED3-T332A in the unfolded state. However, the peak for all four proteins in the unfolded state is located at 303 nm, indicating that all tyrosine residues are in full contact with the solvent[6].

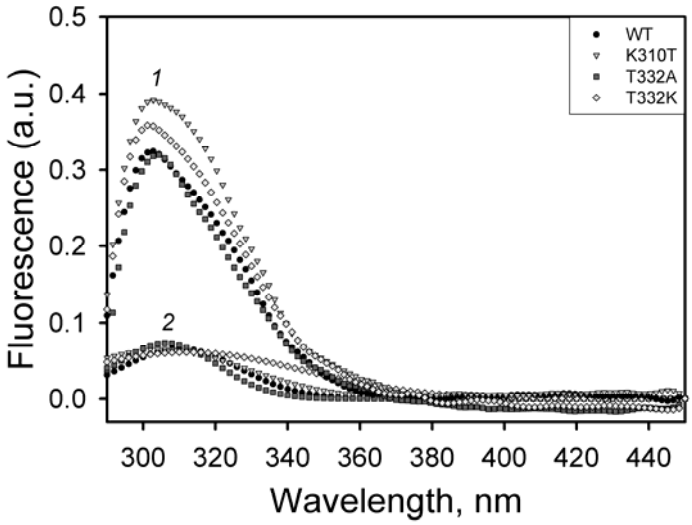

Figure 2.20 Difference spectra [(λ_{ex} = 280 nm) — (λ_{ex} = 295 nm)] corresponding to the fluorescence emission spectrum of the 3 tyrosine residues in rED3-WT, rED3-K310T, rED3-T332A and rED3-T332K in TN200 buffer (curves *2*) and in 4.0 M GdnHCl prepared in the same buffer (curves *1*). The protein concentration was 1 μM.

b. Quantum yield determination

To quantitatively characterize the contribution of the single tryptophan and three tyrosine residues to the protein fluorescence emission spectrum the quantum yield (ϕ) in the native and unfolded states of the protein was determined. The excitation wavelengths were 280 nm ($\phi_{P,280}$) where both tyrosine and tryptophan residues absorb, and 295 nm ($\phi_{P,295}$) where tryptophan residues are selectively excited. The values for rED3-WT were $\phi_{P,280}$ = 0.116 and $\phi_{P,295}$ = 0.226 in the native state and $\phi_{P,280}$ = 0.039 and $\phi_{P,295}$ = 0.060 in the unfolded state. The quantum yield of the protein at λ_{ex} = 280 nm and λ_{ex} = 295 nm is higher in the native state of rED3-WT compare to the unfolded state, indicating that the tryptophan and tyrosine residues are located in an ordered/structured environment. This result is in agreement with the observation that the tryptophan residue in the native state of the protein is located in a hydrophobic environment. The summary of results for the

mutants rED3-K310T, rED3-T332A and rED3-T332K in the native and unfolded states is shown in Table 2.8. Only the mutant rED3-T332A showed a significant decrease in the quantum yield of the protein at both excitation wavelengths (280 nm and 295 nm) compared to rED3-WT. The other two mutants, rED3-K310T and rED3-T332K, have essentially identical values of quantum yield.

Using the spectra in Figure 2.20 we determined the quantum yield of the single tryptophan residue (ϕ_{Trp}) and the quantum yield of the 3 tyrosine residues (ϕ_{Tyr}) of rED3-WT and single mutants, based on a previously described calculation that incorporates the fractional absorbance of tyrosine and tryptophan residues according to the molar ratio of the aromatic amino acids as in the studied protein[59]. The calculation was done for both, native and unfolded states of the rED3s. The values of ϕ_{Trp} for rED3-WT were 0.194 and 0.056 in the native and unfolded states, respectively; and are, within experimental error, the same values as determined using the protein fluorescence emission spectrum with λ_{ex} = 295 nm (see above). This agreement is expected because at λ_{ex} = 295 nm only tryptophan residues are selectively excited. The values for ϕ_{Tyr} in the native and unfolded states of rED3-WT were 0.008 and 0.013. The higher ϕ_{Tyr} in the unfolded state of rED3-WT indicates that the quenching process is eliminated upon destruction of the native architecture around the tyrosine residues. The results for all the single mutants are shown in Table 2.8. Similarly, only the mutant rED3-T332A showed a significant decrease in quantum yield for tryptophan (ϕ_{Trp}) and tyrosine (ϕ_{Tyr}) residues. There are no significant differences between rED3-WT and the mutants rED3-K310T and rED3-T332K neither in the native nor in the unfolded state.

Table 2.8. Quantum yield (ϕ) in native and unfolded states of rED3s

	Native state[a]		Unfolded state[a]		Native state[b]		Unfolded state[b]	
	$\phi_{P,280}$	$\phi_{P,295}$	$\phi_{P,280}$	$\phi_{P,295}$	ϕ_{Trp}	ϕ_{Tyr}	ϕ_{Trp}	ϕ_{Tyr}
rED3-WT	0.116	0.226	0.039	0.060	0.194	0.008	0.056	0.013
rED3-K310T	0.114	0.217	0.039	0.055	0.192	0.007	0.053	0.017
rED3-T332A	**0.077**	**0.147**	**0.025**	**0.042**	**0.129**	**0.003**	**0.037**	**0.007**
rED3-T332K	0.115	0.222	0.041	0.067	0.191	0.010	0.058	0.015

[a] Determined with the method of Parker and Rees[51]. [b] Determined as described in Bujalowski and Klonowska (1994)[59].

2.5 DISCUSSION

Single mutations in the envelope protein structural domain 3 (rED3) that generate virus neutralization resistance and decrease in monoclonal antibody binding affinity have been extensively described in flaviviruses[35]. However, and most intriguing, how these single amino acid substitutions impair the interaction with a monoclonal antibody remains unknown. Or in other words, what is the mechanism by which single mutations on the domain 3 lead to neutralization resistance? In order to address this question, we have studied the effects of three single mutations on the solution properties of rED3 from WN virus wild type (strain USA99b). These single mutations are rED3-K310T, rED3-T332A and rED3-T332K, which are encoded by three different WN virus variant strains[16]. The mutations at position T332 had the most significant effects in monoclonal antibody binding and degree of neutralization, being rED3-T332A partially and rED3-T332K fully resistant; whereas the mutation at position K310 had very small effects.

2.5.1 Linear correlation between neutralization-resistance, stability and dynamics of rED3s

The most significant conclusion of our study is the establishment of a quantitative, linear relation between the stability and the local conformational heterogeneity, represented by changes in solvent accessibility of W397, of rED3 and binding affinity to Mab expressed as $-\text{Log}[Kd_{app}]$, shown in Figure 2.21. There is an inverse correlation between stability and dynamics, namely, the less stable the protein, the more dynamic it is; a consistent set of observations. The rank order in decreasing stability or increasing dynamics is rED3-WT > rED3-K310T > rED3-T332A > rED3-T332K. A decrease in affinity for Mab means an increase in the degree of resistance. The Mab binding data indicate that WT rED3 has the highest affinity. The affinity decreases in the order of rED3-T332K, rED3-T332A and rED3-K310T. This is exactly the rank order of effectiveness in eliciting MAb neutralization resistance to intact viral particles[15,16]. Thus, for the first time and for any virus, we have identified a physical

property of a viral protein that is intimately linked to the biological phenomenon of MAb-neutralization resistance.

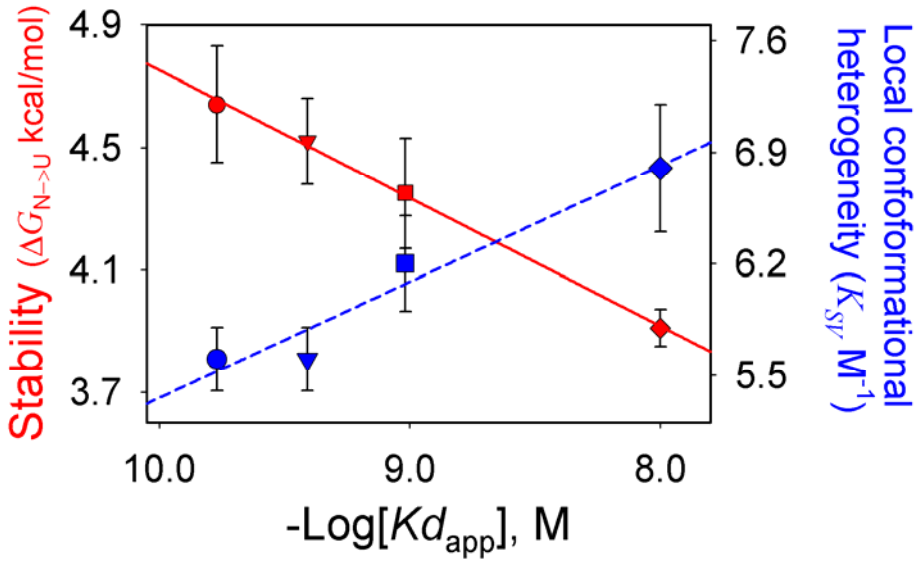

Figure 2.21. Correlation between protein stability, dynamics and binding to MAb-3A3. Protein stability and dynamics data was obtained from Tables 1 and 3, respectively. The values of Kd_{app} for rED3-WT (●), rED3-K310T (▼) and rED3-T332A (■) were 0.17 ± 0.03 nM, 0.39 ± 0.01 nM and 0.96 ± 0.16 nM, respectively. For rED3-T332K (◆) the Kd_{app} is > 10 nm because there was no detectable binding at [Mab]=10 nm (corresponding to ~100-fold greater [Mab] than the Kd_{app} for rED3-WT).

2.5.2 Escape mutations preserve the global solution structure of rED3-WT

By comparison of the global solution properties of rED3-WT and single mutants, we observed that none of the single mutations affected dramatically the global solution properties relative to rED3-WT. The fold of the rED3s was investigated by CD spectroscopy. The CD spectra for all four proteins show the absorption bands characteristic of β-sheet structures; and the calculated composition of secondary structure showed no significant difference in β-sheet content between rED3-WT and all single mutants (see Table 2.7). The tertiary solution structure around the tryptophan residue was

assessed by fluorescence spectroscopy. The maximum fluorescence emission intensity for all rED3s was located at about 341-344 nm (λ_{ex} = 295 nm). Despite the small differences in the emission maxima between rED3-WT and the single mutants; the blue shift of 14-17 nm relative to the emission maximum of free NATA (358 nm) or unfolded rED3 indicates that the local environment of the tryptophan residue in all rED3s is highly structured (Figure 2.18). Furthermore, the quantum yield of the tryptophan residue in all rED3s are equal or higher than the quantum yield of free NATA (ϕ_{NATA} = 0.13)[73]. The higher quantum yield also can be attributed to a more structured environment around the tryptophan residue[62]. The hydrodynamic properties of these proteins (namely, sedimentation velocity coefficient and apparent molecular weight), which are related to the size and shape of the protein, were essentially identical. Altogether, the measurements of the global solution properties of these proteins suggest that the rED3 harboring the single mutations are not escaping from the monoclonal antibodies by generating large structural or conformational changes. Moreover, these observations are consistent with the putative role of ED3 as the primary receptor binding domain, which presumably places significant constraints on the range of mutations that can be tolerated in this domain without adversely affecting virus viability.

2.5.3 Escape mutations affect the structural stability of rED3s

a. Agreement between chemical and thermal denaturation

The observation that there were no large conformational or structural changes due to single mutations can be anticipated because these mutations are spatially located in the surface of rED3 (see Figure 2.1), and most likely are not affecting the core of the protein. However, while the global solution structures between all rED3s are very similar, still there were differences in their structural stability, as judged by differences in the thermodynamic parameters between rED3-WT and mutants (summarized in Tables 1 and 2). The changes in stability (or $\Delta G_{N\rightarrow U}^{H_2O}$) due to a single mutation correlate with the

degree of resistance to neutralization in such a way that the most resistant mutant (i.e. rED3-T332K) is the least stable compared to rED3-WT (Figure 2.21). The rank order in stability was consistent among chemical and thermal denaturation. For instance, rED3-WT was the most stable protein with the highest $\Delta G_{N\rightarrow U}^{H_2O}$ followed by the mutants rED3-K310T and rED3-T332A, and the least stable was rED3-T332K. The agreement between chemical and thermal unfolding was observed also in the rank orders of $C_{1/2}$ and T_m, where rED3-WT has the highest $C_{1/2}$ and T_m, followed by rED3-T332A and finally rED3-K310T and rED3-T332K. Both parameters represent the midpoint of the unfolding process or the concentration/temperature at which the population of folded and unfolded states is equal.

b. Monitoring intrinsic thermodynamic parameters of unfolding

The agreement in the thermodynamic parameters between chemical and thermal denaturation suggests that the $\Delta G_{N\rightarrow U}^{H_2O}$ obtained by GdnHCl- and temperature-induced unfolding is an intrinsic property of the rED3. Furthermore, the fact that the unfolding of the rED3s is accompanied by the disruption of the fluorescence quenching process of the tyrosine residues (as shown in Figures 2.15A and 2.15B), indicates that monitoring the chemical or thermal unfolding process by changes in the fluorescence emission intensity with λ_{ex} = 280 nm reflects the changes in the fluorescent properties of both tryptophan and tyrosine residues simultaneously. Since, the three tyrosine residues are located at different regions of the protein (Figure 2.1), then, the unfolding process monitored with λ_{ex} = 280 nm represents a global unfolding process. Whereas monitoring the unfolding process with λ_{ex} = 295 nm (where tryptophan residues are excited) represents an observable of the local tryptophan residue environment or local unfolding process. Despite the differences between monitoring the unfolding process at λ_{ex} = 280 nm and λ_{ex} = 295 nm, the thermodynamic parameters obtained by global analysis are independent of the excitation wavelength (295 nm or 280 nm), which further supports that the

$\Delta G_{N\to U}^{H_2O}$ obtained by GdnHCl- and temperature-induced unfolding is indeed an intrinsic property of the rED3.

c. Effect of single mutations on the intrinsic thermodynamic parameters of unfolding of rED3-WT

Although there are differences in the stability between the rED3s, the structural information extracted only from changes in the Gibbs free energy of unfolding is limited because $\Delta G_{N\to U}(T,D)$ is a composite of various thermodynamic parameters; namely, change in enthalpy (ΔH^0), entropy (ΔS^0), heat capacity (ΔCp^0), T_m and m_D-value (the latter if the unfolding process is performed in the presence of denaturant). By performing temperature-induced unfolding experiments at increasing concentrations of GdnHCl (as shown in Figure 2.9), we were able to deconvolute with high precision all the thermodynamic parameters characterizing the unfolding process of the rED3s (Table 2.3).

c.1 Enthalpy changes (ΔH^0)

There were no differences in the change of enthalpy upon unfolding (ΔH^0) between rED3-WT, rED3-K310T and rED3-T332K. The average ΔH^0 for these proteins was 73.7 ± 1.5 kcal·mol^{-1}. However, the single mutant rED3-T332A had a lower value of approximately 67.1 ± 1.1 kcal·mol^{-1}. Dissecting with high accuracy the source of the 6.6 ± 2.6 kcal·mol^{-1} difference in ΔH^0 between rED3-T332A and the other rED3s is out of the scope of this investigation. However, by analyzing the different sources that lead to changes in ΔH^0 due to a mutation, it is possible to extract qualitative differences about the solution structure of rED3-T332A compared to the other rED3s. Differences in ΔH^0 due to a mutation can be dissected primarily into two components: differences in the degree of hydration and in the packing interactions of the protein (i.e. changes in van der Waals interactions or hydrogen bonding)[74].

Since the residue T332 is located on the surface of the protein having 98 % ASA to the solvent, the changes in the packing interactions of the core of rED3 due to a

mutation at this position are probably small. In fact, previous studies have shown that the changes in packing interactions have very significant effects on ΔH^0 compare to changes in hydration of non-polar residues[74,75]. Therefore, the small difference of 6.6 ± 2.6 kcal·mol^{-1} between rED3-T332A and the other rED3s can be attributed preferably to differences in the hydration of the protein due to the hydrophobic nature of the substitution on the surface of the protein (threonine→alanine represents a polar→hydrophobic substitution) and to a lesser extend to changes in the packing interaction of the core of rED3-WT.

Because rED3-T332K and rED3-K310T are substitutions of residues that are solvent exposed with a conservation of the hydrophilic nature of the wild type residues, no significant changes in hydration and no changes in the packing interaction are expected for these two mutants. Thus, it is predicted that there should be no significant changes in ΔH^0, an expectation in total agreement with the results observed in the experimental determination of ΔH^0.

c.2 Entropy changes (ΔS^0)

No difference in the change of entropy of unfolding (ΔS^0) was observed for any protein, indicating that the single mutations are not generating large global conformational changes. This agrees with the experimental results, obtained by CD, fluorescence and ultracentrifugation, that show that the mutations do not affect the global solution structure of rED3-WT.

c.3 Heat capacity changes (ΔCp^0)

The most significant variation among the thermodynamic parameters between rED3-WT and the single site mutants was observed in the heat capacity change of unfolding (ΔCp^0). For instance, the mutants rED3-K310T and rED3-T332K had higher ΔCp^0 values than rED3-WT and rED3-T332A (Table 2.3). More precisely understanding the origin of the increment in ΔCp^0 for rED3-K310T and rED3-T332K can provide additional structural information regarding the effects of the single mutations on the

solution structure of these proteins. Changes in ΔCp^0 are associated to: 1) changes in the protein surface chemistry (i.e. changes in the protein surface electrostatics due to a mutation)[55], or 2) changes in the solvent-accessible hydrophobic surface area that is exposed upon unfolding or $\Delta ASA_{N\rightarrow U}$ (i.e. due to larger surface area exposed during the unfolding process); or both[76]. Furthermore, changes in $\Delta ASA_{N\rightarrow U}$ can be attributed to: a) changes in the size of the protein, for example by addition or deletion of residues or, b) changes in the folded state structure or conformation due to a mutation or, c) changes in the unfolded state configuration or conformation (i.e. residual structure caused by a mutation), or any combination of a, b or c[77,78].

It is conceivable that the single mutations rED3-K310T and rED3-T332K perturb the surface electrostatics of rED3-WT because the mutation sites are located on the surface of the protein, and the physicochemical nature of the substitution involves a removal and addition of a positively charged residue, respectively. Therefore, it is plausible that changes in the surface electrostatics cause an increment in ΔCp^0. However, it is important to determine if structural changes other than surface electrostatics play a role in the increment of ΔCp^0 for these two mutants.

Based on the linear correlation between number of residues of a protein, $\Delta ASA_{N\rightarrow U}$, ΔCp^0 and m_D-values for globular proteins observed by Myers et al., (1996)[78] –which means, higher number of residues → higher $\Delta ASA_{N\rightarrow U}$ → higher ΔCp^0 → higher m_D-values and vice versa– it can be anticipated that the higher ΔCp^0 for rED3-K310T and rED3-T332K due to an increment in protein size (that will concomitantly increase $\Delta ASA_{N\rightarrow U}$) is most unlikely. On one hand our experimental results show no significant differences in the global solution properties like fold, size and shape between all the rED3s. On the other hand, an increase in $\Delta ASA_{N\rightarrow U}$ and/or increase in protein size can be easily illustrated by applying the linear correlation determined by Myers et al.[78] to rED3-WT. The predicted values based on the number of residues of rED3-WT (= 115) were the following: 9788 $Å^2$ for $\Delta ASA_{N\rightarrow U}$, 1.61 $kcal\cdot mol^{-1}\cdot K^{-1}$ for ΔCp^0 and 3.01 $kcal\cdot mol^{-1}\cdot M^{-1}$ for m_D-value. The predicted values are in excellent agreement with the experimental ΔCp^0 and m_D-value (Table 2.3), suggesting that the rED3-WT behaves like an ideal

globular protein; and, the agreement validates the application of this correlation to our system. Using ΔCp^0 of rED3-K310T and rED3-T332K (1.99 kcal·mol^{-1}·K^{-1} and 2.06 kcal·mol^{-1}·K^{-1}, respectively) to calculate back the corresponding $\Delta ASA_{N\rightarrow U}$ for these proteins we obtained 11795 Å^2 and 12163 Å^2, respectively. The difference in $\Delta ASA_{N\rightarrow U}$ between both mutants and rED3-WT would be at least 2000 Å^2, which is equivalent to have on average 20 extra residues in the primary sequence of the rED3s. By performing the same calculation but using, instead of ΔCp^0, the m_D-values of rED3-K310T and rED3-T332K we obtained similar differences in $\Delta ASA_{N\rightarrow U}$ compared to rED3-WT. Therefore, these calculations and our experimental results show that for our system it is very improbable to have an increase in ΔCp^0 of ~0.4 kcal·mol^{-1}·K^{-1} caused only by large changes in the protein structure or conformation.

As mentioned above, structural/conformational differences in the folded and unfolded states due to a mutation are two additional potential contributors that lead to changes in $\Delta ASA_{N\rightarrow U}$ and therefore potentially increase ΔCp^{o76} .These two contributors can potentially be responsible of the increment in ΔCp^0 observed for rED3-K310T and rED3-T332K. Differences in the folded state of the rED3s can be indirectly related to differences in the slope of the pre-transition or folded state base line between rED3-K310T/rED3-T332K and rED3-ET/rED3-T332A during the GdnHCl-induced unfolding process (See Figure 2.6). The slopes for rED3-K310T and rED3-T332K were positive while the slopes for rED3-WT and rED3-T332A were negative. Considering GdnHCl at low concentration (before the folded↔unfolded transition starts) as a perturbation that probes the different conformations of the protein in the folded state; then, the variations in the slope would indicate changes or differences in the folded state conformation between rED3-K310T/rED3-T332K and rED3-WT/rED3-T332A. The direct effect of having differences in the slope of the folded state base lines is that there are differences in the m_D-values. The mutants rED3-K310T and rED3-T332K had higher m_D-values compared to rED3-WT and rED3-T332A. These differences were consistent among the two excitation wavelengths (λ_{ex} = 280 nm and λ_{ex} = 295 nm) in both chemical and thermal unfolding experiments. Provided that it has been proven that the single mutations

rED3-K310T and rED3-T332K do not generate large changes in $\Delta ASA_{N\rightarrow U}$ due to gross conformational changes (i.e. changes in protein size), then, higher m_D-values (which implies higher $\Delta ASA_{N\rightarrow U}$ and higher ΔCp^0) indicates that these mutations in addition to generating changes in the surface chemistry, generate subtle changes in the folded conformation of the protein.

On the other hand, alteration of the charge distribution of a protein can affect its conformation in the unfolded state due to repulsion or attraction of charges[79]. If this is the case for rED3-K310T and rED3-T332K, then the unfolded conformation of the protein could have an impact in the $\Delta ASA_{N\rightarrow U}$ and ΔCp^0. However, it was observed that rED3-WT, rED3-K310T and rED3-T332K had the same unfolded state base line during the GdnHCl- and temperature-induced unfolding process, indicating that there are no major differences in the unfolded conformation between these proteins. This last observation is further supported by the fact that these three proteins had essentially the same quantum yield for the tryptophan residue alone (ϕ_{Trp}), the three tyrosine residues (ϕ_{Tyr}) and both natural fluorophores ($\phi_{P,280}$) in the unfolded state (in 4.0 M GdnHCl) as shown in Table 2.8. Differences in the quantum yield can be attributed to differences in the conformation of the environment surrounding the fluorophores in the unfolded state (i.e. residual structure). Therefore, a change in the $\Delta ASA_{N\rightarrow U}$ or ΔCp^0 due to changes in the unfolded conformation of rED3-K310T and rED3-T332K is most unlikely.

In the case of the single mutant rED3-T332A, which does not involve a direct alteration of the surface electrostatics on the protein, there was not a significant difference in ΔCp^0 and m_D-value compared to rED3-WT. Nevertheless, this substitution subtlety affects the solution structure or conformation of the protein. The first experimental evidence of changes in the solution structure of rED3-T332A is noted in the higher calculated β-turn content compare the other rED3s. Based on the quantitative analysis of the CD spectra (Table 2.7), the calculated β-turn content of rED3-T332A was 30 %, while for the other rED3s is ~18 %. It is conceivable that this mutation will affect the β-turn structures because the substitution is located on a surface loop of rED3 (see

Figure 2.1) and also involves a drastic change in the hydrophilic nature of the residue (polar→hydrophobic).

It was stated before that a change in ΔH^0 due to a mutation can be attributed to changes in the degree of hydration of non-polar residues or changes in the packing interactions of the core of the protein[74]; both are related to the solution structure of the protein. Therefore, the observed decrease in ΔH^0 (of ~6.6 ± 2.6 kcal·mol^{-1}) for rED3-T332A correlates with the observed change in β-turn structure content. In addition, changes in the solution structure are observed indirectly by the large difference in quantum yield between rED3-T332A and the other rED3s (see Table 2.8). The quantum yield of rED3-T332A for the tryptophan residue alone (ϕ_{Trp}), the three tyrosine residues (ϕ_{Tyr}) and both natural fluorophores ($\phi_{P,280}$) in the native and unfolded state was reduced by ~30 %, suggesting that the solution structure surrounding the natural fluorophores of rED3-T332A were being perturbed compared to rED3-WT, rED3-K310T and rED3-T332K. For instance, direct evidence of change on the solution structure around the tyrosine residues comes from the fractional exposure of the three tyrosine residues calculated from the second derivative absorbance spectra (Figure 2.16). The fractional tyrosine exposure for rED3-T332A was around 30 % while for the other rED3s was on average ~19 % (Table 2.6). Then, the statement that the lower quantum yield for the tyrosine residues is indicative of variations in the solution structure (tertiary structure) is supported by the differences in fractional tyrosine exposure.

Overall, the analysis of the thermodynamic parameters indicates that, despite of the small differences in the stability ($\Delta G_{N\to U}^{H_2O}$) among all rED3s, there were significant effects on the different thermodynamic components that include $\Delta G_{N\to U}(T,D)$. All single mutations affected the folded conformation of the protein but were reflected in different thermodynamic parameters. For example, the single mutant rED3-T332A decreased ΔH^0, whereas ΔCp^0 and m_D-value remain essentially identical to those corresponding to rED3-WT. On the other hand, for mutants rED3-K310T and rED3-T332K, an increase was observed in both ΔCp^0 and m_D-value, while ΔH^0 remained similar to the value for rED3-

WT. Since all three mutants show subtle perturbation in the solution structure/conformation of the protein (not in the same manner), then it is conceivable that the direct change in the electrostatic surface or net charge of rED3, as in the cases of rED3-K310T and rED3-T332K, is the major cause for the increase in ΔCp^0 and m_D-value because rED3-T332A does not affect these thermodynamic parameters.

2.5.4 Escape mutations affect the surface chemistry of rED3s at distant regions

Direct evidence of changes on the surface electrostatics of rED3-WT due to a single mutation is observed in the tryptophan quenching experiments using charged molecules: iodide and thallium (see Table 2.5). The tryptophan quenching efficiency of rED3-K310T and rED3-T332K by iodide was decreased by ~50 % compared to rED3-WT. Noteworthy, these two substitutions have the same effect in quenching efficiency and they involved removal (rED3-K310T) or addition (rED3-T332K) of a positively charged residue that will affect the net charge of the protein. In the case of rED3-K310T, the net charge of the protein will have one less positive charge compared to rED3-WT, which is consistent with the increased thallium quenching efficiency, since thallium will be attracted to negatively charged regions of the protein or, in other words, to less positively charge environments. Despite that in the case of rED3-T332K the net charge will have one extra positive charge compared to rED3-WT, there was not a significant effect on thallium quenching efficiency. The mutation rED3-T332A had the opposite effect compared to rED3-K310T and rED3-T332K. It was observed that the iodide quenching was ~50 % higher for both the degree of quenching (K_{SV}) and quenching efficiency compared to rED3-WT.

It is important to note that the substitutions rED3-K310T and rED3-T332K not only involve a direct change in the net charge of the protein, but the differences in K_{SV} and quenching efficiency of iodide and thallium show that there is also a redistribution of charges on the surface of rED3 because the mutation sites are 19 Å and 31 Å away from the tryptophan residue (see Figure 2.1). The substitution rED3-T332A does not involve a

direct change in the net charge of the protein, but the differences in K_{SV} and quenching efficiencies of iodide compared to rED3-WT indicate that there is also charge redistribution on the surface of the protein. Furthermore, since the mutation sites and the tryptophan residues are at distant regions of rED3, the observed effects on K_{SV} and quenching efficiencies due to the single mutations are a first evidence of long-range communication between distant sites of rED3.

2.5.5 Long-range communication between distant sites of rED3s

The data from the tryptophan fluorescence quenching with acrylamide can provide additional information about the accessibility of the tryptophan residue to the solvent, and thus indirectly, the dynamic nature of the environment surrounding the tryptophan residue. Significant differences in the degree of tryptophan quenching (K_{SV}) with acrylamide between rED3-WT and the mutants were observed. In the case of rED3-WT, the K_{SV} for acrylamide was 5.6 ± 0.2 M^{-1} while for the mutants rED3-T332A and rED3-T332K were 6.2 ± 0.3 M^{-1} and 6.8 ± 0.4 M^{-1}, respectively. No changes were observed for the mutant rED3-K310T (K_{SV} = 5.6 ± 0.2 M^{-1}). These differences imply that for rED3-T332A and rED3-T332K the tryptophan residue is more accessible to the solvent and thus, more hydrophilic. This last conclusion is supported by the observed lower quenching efficiency of the rED3-T332K tryptophan residue by TCE, which is a hydrophobic quencher. This experimental evidence supports that the mutations are not only affecting the surface electrostatics around the tryptophan residue but also the structural dynamics of that region in the folded state, being rED3-T332K the most dynamic. Again, changes in tryptophan accessibility and dynamics due to distal mutations indicate that there is long-range communication between distant sites of rED3.

2.6 CONCLUSIONS

The results and conclusions derived from these studies most probably reflect the intrinsic solution properties of the native viral E protein ED3, although we have investigated rED3 as an isolated domain and not in the context of the whole ectodomain of the E protein or in the context of the intact virion. However, it has been observed that the Ig-like fold of the rED3 from WNV[19] is preserved as in the whole E protein ectodomain[17,36]. For example, the *r.m.s.d.* between backbone atoms of the WNV NMR structure and dengue virus type 2[36] rED3s is 1.4 Å, despite the fact that the latter structure was determined as part of the complete ectodomain of the E protein and by x-ray crystallography. Furthermore, the agreement between the degree of virus neutralization (using the intact viral particle harboring the mutations studied in this report)[15,16] and our results on Mab binding using the rED3s (Figures 2.5 and 2.21) indicates that the rED3, as an isolated domain, has very similar antigenic properties as compared to the ED3 in the context of the intact virion.

The molecular mechanism of antibody neutralization resistance of flaviviruses, including WNV, has been elusive in spite of the intense investigation. However, for the first time, as a result of this study a quantitative relation between the important biological phenomenon of antibody mediated neutralization and the biophysical properties of ED3 is established:

1. The substitution that conferred the highest degree of neutralization resistance to the virus possesses the least stable and most dynamic ED3 (Figure 2.21).
2. These mutations elicit effects at distant sites that are spatially located many Å away i.e. long-range communications between distant sites in ED3 (Figure 2.1).

These experimental results are strong evidence in support of our hypothesis that the mechanism of neutralization involves not only regions of interaction with the antibody (epitope) but also distant structural elements of the protein via long-range communication. Furthermore, this hypothesis provides an explanation for other studies of flaviviruses where it was observed that an escape from neutralization due to a mutation can happen without a loss in MAb binding[16,28-30]. Our hypothesis is also consistent with

other studies in which a positive cooperativity in antibody binding (avidity) was observed[38-40].

CHAPTER 3

RESISTANCE TO ANTIBODY-MEDIATED NEUTRALIZATION IN WEST NILE VIRUS IS MODULATED BY THE REDISTRIBUTION OF CONFORMATIONAL STATES OF THE ENVELOPE PROTEIN DOMAIN 3

3.1 ABSTRACT

Monoclonal antibody resistance is induced by mutations in viral proteins, which lead to changes in the interaction of the antibody with the virus epitope. The envelope protein domain 3 (ED3) of West Nile virus (WNV) harbors most of the critical mutations that have been shown to induce resistance against antibody neutralization. Because these mutations are distributed throughout the ED3 structure, including regions that are outside the epitope, or not accessible to antibodies, we hypothesized that there must be long-range communication between distant sites of ED3 that affect the interaction between the antibody and the epitope. Using a previously validated ensemble-based description of the equilibrium, we investigated the network of communications among distant residues in ED3 and demonstrated new insight into the physical basis of resistance to antibody neutralization. First, the basic mechanism of communications in wild type ED3 was not perturbed by any mutation, implying that the fold of ED3 is thermodynamically robust. Second, only mutations that generate resistance to antibody neutralization were able to redistribute the ensemble of conformational states of ED3. Furthermore, the region affected by these mutations corresponds to the major neutralizing epitope on WNV ED3, or the BC loop, indicating that resistance to antibody neutralization is function of the energetic coupling among distant sites.

3.2 INTRODUCTION

To understand the molecular and structural basis of antibody-mediated neutralization resistance in West Nile virus (WNV), the central dogma has been the study of the effect of mutations on antibody binding to the envelope protein domain 3 (ED3), the receptor-binding domain of flaviviruses[17]. By using a variety of cellular, molecular and biophysical approaches, it has been shown that ED3 of flaviviruses harbors most of the critical epitopes recognized by virus type-specific monoclonal antibodies[14-16,19,24] (MAbs).

ED3 in WNV is an IgG-like protein of approximately 100 residues that consists of a classic beta-sandwich fold (Figure 3.1A). The N-terminus and loops connecting beta-strands B-C, D-E and F-G conforms a solvent-exposed surface patch in the intact viral particle[19,25] (Figure 3.1B). This surface patch has been shown to be the epitope of high titer neutralizing MAbs with very high degree of surface complementarity[24,80] (S_c = 0.763). The high surface complementarity and the absence of global conformational changes between the free[19] and bound conformations of ED3[24] (*r.m.s.d.* of c-alpha atoms = 1.8 Å), suggests that the MAb-ED3 complex formation follows a *lock-and-key* mechanism[81-83].

If WNV antibody neutralization is due to a *lock-and-key* mechanism, then any perturbation (i.e. a mutation) on the interaction surface between the complex MAb-ED3 should be able to impair, to a certain extent, MAb binding and decrease neutralization. However, mutations of most of the residues that are part of this interaction surface have no observable effect, and only mutations in a few sites are able to decrease MAb binding, especially depending on the amino acid side chain substitution[14-16]. For example, the mutation T332K decreases MAb binding affinity > 100-fold relative to wild type ED3 and generates a fully resistant variant in *in vitro* neutralization assays against a panel of virus type-specific MAbs[14-16,84]. However, the mutation T332A reduces MAb binding affinity only by ~10-fold and generates a partially resistant virus while the mutation T332V shows no significant antibody neutralization resistance[14]. In a closely-related flavivirus, Japanese encephalitis virus (JEV), mutations at residues S331 and D332,

located in the equivalent epitope of WNV, decrease MAb binding affinity also dependent on the type of amino acid side chain substitution as well[85]. Noteworthy, for both WNV and JE there is no a correlation between the changes in charge, size or hydrophobicity of the mutated aminoacid side chain and the magnitude of the effect in MAb binding[85], which would be expected for a *lock-and-key* binding mechanism.

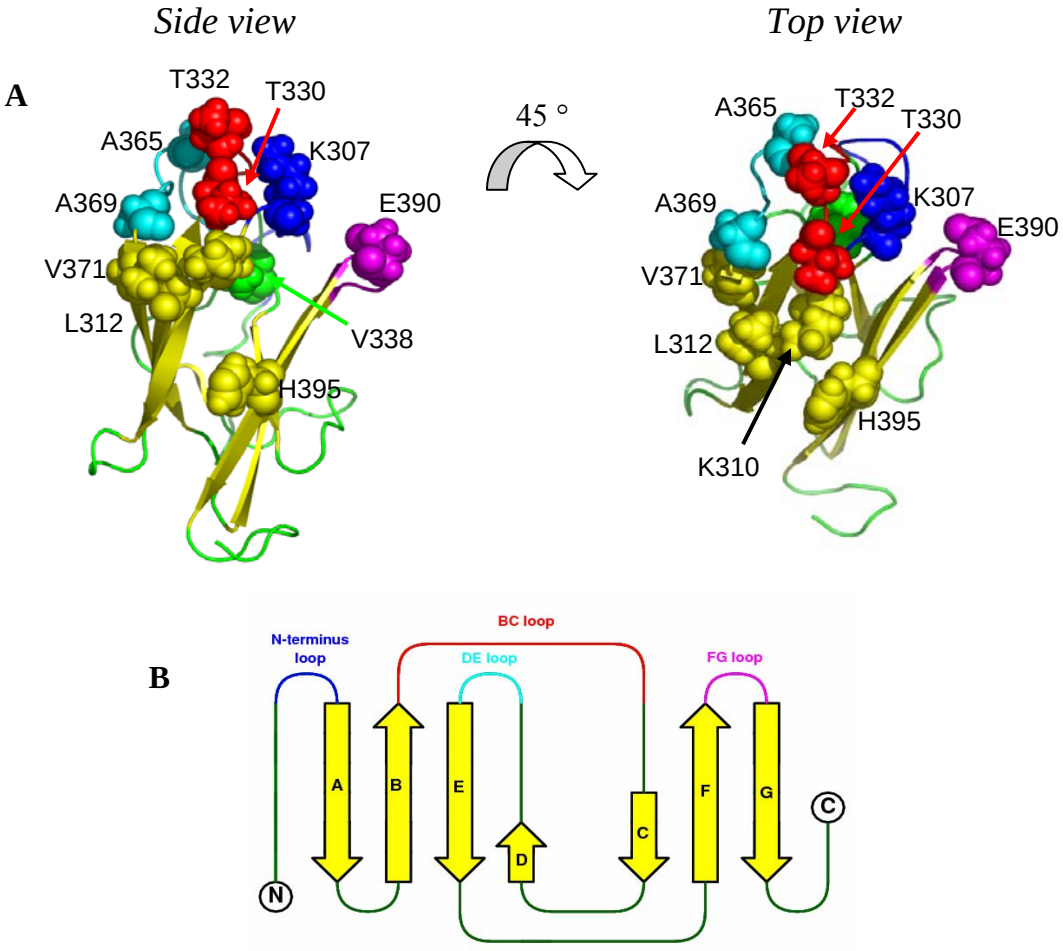

Figure 3.1. Naturally occurring mutation sites in ED3 from WNV. **A**, NMR structure of wild type ED3[19] (PDB 1S6N). Residue numbering corresponds to amino acids 296-406 of the entire E protein of WNV. Residues in yellow correspond to mutation sites in beta-strand structures while residues in green correspond to mutation sites in the N-terminus and loops B-C, D-E and F-G. The ribbon diagram was rendered using PyMOL v. 0.97 (Delano Scientific LLC, San Carlos, CA). **B**, Topological map of ED3 showing the two beta-sheets forming the beta-sandwich fold. The N-terminus and loops B-C, D-E and F-G conforms a solvent-exposed surface patch in ED3 that is recognized by MAb E16[24].

These observations indicate that MAb-neutralization resistance is not a consequence of the mutational perturbation of a *lock-and-key* mechanism between the MAb and protein-antigen complex. Therefore, the interaction of MAbs with ED3 must include a dynamical component as well, to accommodate distortions in the high surface complementary between the MAb and ED3. In fact, there are significant local displacements in the N-terminus and loops B-C, D-E and F-G in ED3 after binding to the Fab domain of a potent neutralizing MAb[24] (*r.m.s.d.* of c-alpha atoms = 3.3 Å vs. 1.8 Å when all c-alpha atoms of the structure in ED3 in the free and bound forms are considered), indicating that neither ED3 nor MAbs bind as rigid bodies, and in agreement with previous observations showing that epitopes tend to localize in mobile regions of protein antigens[86,87]. More recent studies have shown; however, that too high segmental mobility can impair MAb binding as well[88-90]. Therefore, there must be a range of conformational states in the protein antigen that optimize MAb binding, and have to be accounted in the process of MAb neutralization.

If the conformational states of the protein antigen and epitope are important for MAb-binding and neutralization; then, what are the MAb-resistant mutations doing in addition to perturbing locally the interaction surface between the MAb and the epitope? It is clear that most mutations in the epitope do not have any effect in binding affinity[14-16] and there are mutations in ED3 that are not in accessible regions to MAbs[16]. Moreover, antibody neutralization resistance due to mutations outside the epitope has been reported in other flaviviruses[28-30,38-40], and our preceding studies showed that MAb-resistant mutations in WNV have long-range effects over residues distantly positioned more than 20 Å away[84] from where the MAb interacts with the epitope. Thus, we can hypothesize that MAb-resistant mutations, independent of their location, can shift the conformational states of ED3 towards a *MAb-resistant-like conformation* in addition to the expected local perturbation in the epitope.

To test this hypothesis and to gain insight into the role of conformational states of ED3 in antibody-mediated neutralization in WNV we investigated the coupling of

fluctuations (or energetic coupling) between distant residues in ED3 by using a previously validated ensemble-based description of the equilibrium[91-95]. The results indicate that despite of the occurrence of mutations in multiple residues throughout ED3 that may affect many different regions, only the BC loop is poised to respond to perturbation of mutations located anywhere in ED3. Our findings underscore the crucial role of conformational dynamics in antibody-mediated neutralization by perturbations of sparse but connected residues via long-range communication.

3.3 METHODS

Most biological processes facilitated by proteins involve inter- and intra-molecular communication between distant groups of residues that allows the transmission of signals exerted by protein effectors such as small molecules, nucleic acids, lipids or proteins. In this study, a statistical mechanical approach was used to investigate the underlying mechanism that governs the role of mutations distantly positioned from the epitope in ED3 from WNV by treating the protein molecule as a fluctuating ensemble of conformational states. In Figure 3.2 it is shown the graphic representation of the fully folded state of ED3 from WNV and as an ensemble of conformational states with fluctuating local regions in yellow color.

3.3.1 Generation of the ensemble of conformational states of ED3 from West Nile virus

An ensemble of conformational states for the envelope protein domain 3 (ED3) from West Nile virus (PDB: 1S6N) was generated using the COREX algorithm[91]. The COREX algorithm models the native state of a protein by generating an ensemble of conformational states that involves local fluctuations, each of which is derived from the high resolution structure. The local fluctuations are modeled as local unfolding reactions or changes in solvent accessible surface area (ΔASA) of contiguous group of residues

defined as *folding units*. The scheme of the method employed by COREX for generating partially unfolded states in ED3 from WNV is shown in Figure 3.3. Briefly, the folding units are treated or as native-like or as being unfolded with the same thermodynamic properties as unstructured peptides; and by systematically varying the boundaries of the folding units, an exhaustive enumeration of different conformational states is achieved.

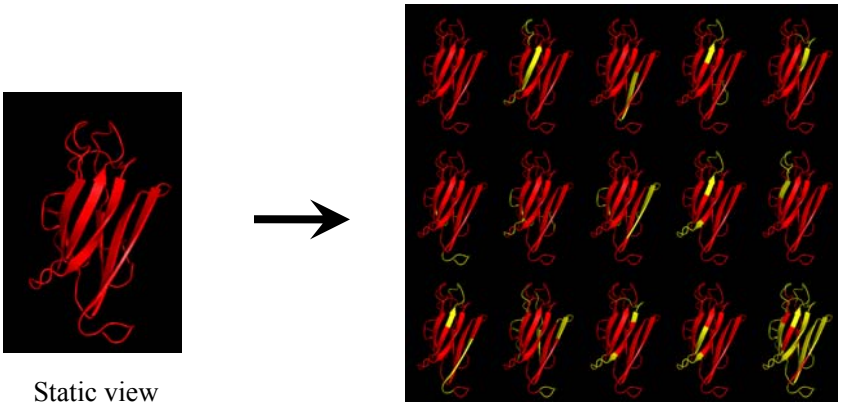

Figure 3.2. Graphic representation of the ensemble of conformational states of ED3 from WVN obtained from the fully folded state or high resolution structure (PDB 1S6N)[19]. Regions in red represent folded-like residues while regions in yellow represent the fluctuating residues contained in the native state ensemble. The ribbon diagram was rendered using PyMOL v. 0.97 (Delano Scientific LLC, San Carlos, CA).

3.3.2 Calculation of the probability of each ED3 conformer[a]

The probability of any given conformation in the native state ensemble of ED3 can be obtained by recasting the Gibbs free energy change between the conformer *i* and the fully folded state (ΔG_i) into the form of an equilibrium constant ($K_i = Exp[-\Delta G_i / RT]$):

[a] This section represents a summary from the review article Modeling the Native State Ensemble in *Methods in Molecular Biology* **168**, 93-116[116].

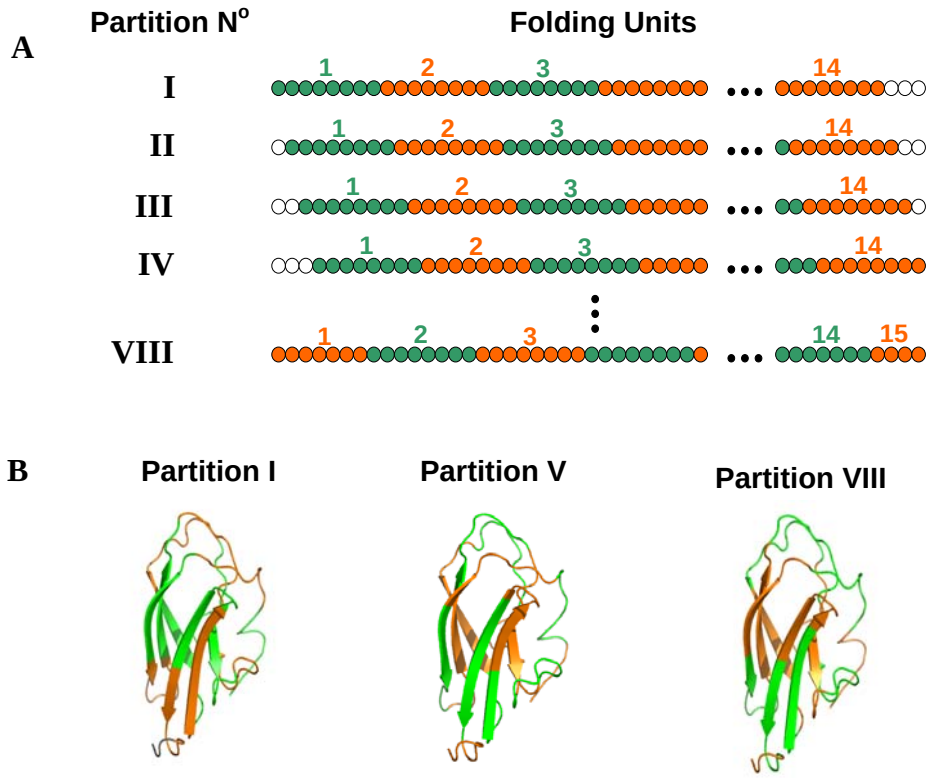

Figure 3.3. **A**, Illustration of the method used by COREX[91]. For ED3, a 115-residue protein, the folding units incorporated 8 residues, except for the first and last residues after moving the boundary of the folding unit by one. A minimum of 4 residues constitutes a folding unit. Each partition is overlaid in the high resolution structure and the folding units are unfolded in all possible combinations resulting in 2^N-2 (=16,382 or 32,766) partially unfolded states per partition. **B**, Overlay of partitions I, V and VIII into the high resolution structure of ED3 from WNV. Figures rendered in PyMOL v. 0.97 (Delano Scientific LLC, San Carlos, CA).

$$P_i = \frac{K_i}{\sum_i K_i} = \frac{K_i}{Q} \tag{3.1}$$

where Q is the partition function defined as $Q = \sum_{i=0}^{N} \exp(-\Delta G_i / RT)$ and N is the total number of states accessible to the system. T and R are the temperature in Kelvin and ideal gas constant in cal·K^{-1}·mol^{-1}, respectively.

The free energy of each conformational state can be accurately described by using a surface-area-based parameterization that has been validated experimentally[91-95]. This parameterization allows to calculate the heat capacity (ΔCp_i), enthalpy (ΔH_i) and entropy (ΔS_i) change between the conformer *i* and the fully folded state (or reference state). The change in all these thermodynamic functions have been parameterized in terms of changes in of ΔASA of polar (ΔASA_{pol}) and apolar (ΔASA_{ap}) residues. The change in heat capacity is described by equations 3.2 and 3.3.

$$\Delta Cp_i = \Delta Cp_{ap} + \Delta Cp_{pol} \tag{3.2a}$$

$$\Delta Cp_i = a_c \cdot \Delta ASA_{ap} + b_c \cdot \Delta ASA_{pol} + (c_c - b_c) \cdot \Delta ASA_{OH} \tag{3.2b}$$

where a_c, b_c and c_c are equal to 0.45, 0.26 and 0.17 in cal·K^{-1}·mol^{-1}·Å^{-2}, respectively. The third term accounts for the contribution of hydroxyl groups to ΔCp[91]. Because most of the calculations performed in this research were at 20 °C, the dependence of ΔCp by temperature is insignificant[91].

The enthalpy change also scales with changes in ΔASA and can be written as:

$$\Delta H_i(60\ ^\circ C) = a_H \cdot \Delta ASA_{ap} + b_H \cdot \Delta ASA_{pol} \tag{3.3}$$

where a_H and b_H are equal to -8.44 and 31.4 in cal·mol^{-1}·Å^{-2}, respectively. The enthalpy change parameterization is calculated at 60 °C. This turned out to be a very good approximation since ED3 wild type from WNV has a experimental T_m of 61 °C. The entropy change of unfolding has a contribution from changes in solvation ($\Delta S_{solv,i}$) and changes in conformational entropy ($\Delta S_{conf,i}$) describe in equation 3.5:

$$\Delta S_i = \Delta S_{solv,i} + \Delta S_{conf,i} \tag{3.4a}$$

and,

$$\Delta S_{solv,i}(T) = \Delta Cp_{ap} \cdot \ln(T / T^*_{s,ap}) + \Delta Cp_{pol} \cdot \ln(T / T^*_{s,pol}) \tag{3.4b}$$

where $T^*_{s,ap}$ and $T^*_{s,pol}$ have been shown to be ~385 K and ~335 K, respectively[91]. The conformational-entropy change is calculated taking into account the contributions of:

1. The summation of entropy change for all side chains that are buried in the folded state and become solvent-exposed in a partially unfolded state, or $\sum \Delta S_{bu-ex}$;
2. the summation of entropy change of solvent-exposed side chains upon unfolding of the peptide backbone, or $\sum \Delta S_{ex-un}$; and,
3. the summation of entropy change of the backbone for residues that become unfolded in partially unfolded states, or $\sum \Delta S_{bb}$. Therefore:

$$\Delta S_{conf,i} = \sum \Delta S_{bu-ex} + \sum \Delta S_{ex-un} + \sum \Delta S_{bb} \tag{3.4c}$$

The three terms contributing to the conformational entropy, $\Delta S_{conf,i}$, constitute 95 % of the entropy of stabilization of most proteins. The entropy difference, corresponding to ≈ 5 % is used to define a conformational entropy-scaling factor.

3.3.3 Thermodynamic descriptors to interrogate the ensemble of ED3

After obtaining a value for ΔG_i we can resolve it into the contributions from each residue in the conformational state i ($\Delta G_i = \sum_{j}^{N=residues} \Delta H_{i,j} - T\Delta S_{i,j}$), which allows the

investigation of thermodynamic properties at the residue-specific level contained in the native state ensemble[91].

a. Residue stability constant ($\kappa_{f,j}$)

An important thermodynamic descriptor that follows from the native state ensemble of ED3 is the residue-specific free energy ($\sum_{i}^{N=states}\Delta G_{f,j}$ or simply $\Delta G_{f,j}$), which provides a measure of the relative stability of the different regions in the protein and places the statistical mechanical foundation to investigate the long-range communications in ED3 (see the following section). The residue-specific free energy is defined as

$$\Delta G_{f,j} = -RT \cdot \ln \kappa_{f,j} = -RT \cdot \ln \left[\frac{\sum_{i}^{N=states} P_{f,j}}{\sum_{i}^{N=states} P_{u,j}} \right] \tag{3.5}$$

where $\kappa_{f,j}$ corresponds to the residue-specific equilibrium constant and, $\sum_{i}^{N=states} P_{f,j}$ and $\sum_{i}^{N=states} P_{u,j}$ are the summations of the probabilities of states '*i*' where residue '*j*' is folded and unfolded, respectively.

b. Energetic coupling between each residue pair in ED3 ($\Delta\Delta G_{j,k}$)

The long-range communications in ED3 are described as the mutual perturbation between each residue pair (*j*, *k*) in the protein. The first step to obtain a quantitative measurement of long-range communication is to evaluate the effect of an energetic perturbation, of 2 kcal·mol^{-1}, on residue '*k*' ($\Delta g^{pert,k}$) over residue '*j*', when residue '*k*' is in the folded state. Mathematically this is accomplished by rewriting the expression for

$\kappa_{f,j}$ in order to consider the independent folding status of the second residue, '*k*', and evaluate the effect of $\Delta g^{pert,k}$ over $\kappa_{f,j}$,

$$\Delta G_{f,j}^{pert,k} = -RT \cdot \ln \kappa_{f,j}^{pert,k} = -RT \cdot \ln \left[\frac{\sum_{i}^{N=states} \phi^{pert,k} \cdot P_{f,j|f,k} + \sum_{i}^{N=states} P_{f,j|u,k}}{\sum_{i}^{N=states} \phi^{pert,k} \cdot P_{u,j|f,k} + \sum_{i}^{N=states} P_{u,j|u,k}} \right] \qquad (3.6)$$

where $\sum_{i}^{N=states} P_{f,k}$ and $\sum_{i}^{N=states} P_{u,k}$ are the summations of the probabilities of states '*i*' where residue '*k*' is folded and unfolded, respectively; $\phi^{pert,k} = Exp[-\Delta g^{pert,k} / RT]$ corresponds to the effect of perturbing all states in which residue '*k*' is folded, and $\kappa_{f,j}^{pert,k}$ corresponds to the residue-specific equilibrium constant of residue '*j*' after perturbing residue '*k*' by $\Delta g^{pert,k}$.

In view that energetic perturbation is not a bidirectional phenomenon *per se*[92,94,95]; we consider here the mutual perturbation ($\Delta\Delta G_{j,k}$) as a bidirectional effect defined as the sum of the influence of a perturbation on residue '*j*' over residue '*k*' ($\Delta\Delta G_{k}^{pert,j} = \Delta G_{f,k} - \Delta G_{f,k}^{pert,j}$) and the reciprocal effect; namely, the influence of a perturbation on residue '*k*' over residue '*j*' ($\Delta\Delta G_{j}^{pert,k} = \Delta G_{f,j} - \Delta G_{f,j}^{pert,k}$). In conclusion, we obtain a quantitative descriptor for long-range communications in ED3 defined as the mutual perturbation between each residue pair (*j*, *k*)[92,94,95],

$$\Delta\Delta G_{j,k} = \Delta\Delta G_{j}^{pert,k} + \Delta\Delta G_{k}^{pert,j} = -RT \cdot \ln \left[\frac{\kappa_{f,j}^{WT}}{\kappa_{f,j}^{pert,k}} \cdot \frac{\kappa_{f,k}^{WT}}{\kappa_{f,k}^{pert,j}} \right] \qquad (3.7)$$

c. Nature of the energetic coupling: Positive, Negative and Neutral[95]

The energetic coupling between two residue pairs in a protein can be manifested in positive, negative and neutral coupling. The ensemble-based description of ED3 provides a pathway-independent basis for coupling because only the conformational states in the ensemble are interrogated. To illustrate the nature of the energetic coupling between a residue pairs, equation 3.5 is expanded to account the folding state of a second residue, '*k*'; the residue stability constant for residue '*j*' becomes,

$$\Delta G_{f,j} = -RT \cdot \ln \kappa_{f,j} = -RT \cdot \ln \left[\frac{\sum_{i}^{N=states} P_{f,j|f,k} + \sum_{i}^{N=states} P_{f,j|nf,k}}{\sum_{i}^{N=states} P_{nf,j|f,k} + \sum_{i}^{N=states} P_{nf,j|nf,k}} \right] \qquad (3.8)$$

where $\sum_{i}^{N=states} P_{f,j|f,k}$ and $\sum_{i}^{N=states} P_{f,j|nf,k}$ are the summations of the probabilities of states '*i*' where residue '*j*' is always folded and residue '*k*' is either folded or unfolded, respectively. And, $\sum_{i}^{N=states} P_{nf,j|f,k}$ and $\sum_{i}^{N=states} P_{nf,j|nf,k}$ are the summations of the probabilities of states '*i*' where residue '*j*' is always unfolded and residue '*k*' is either folded or unfolded, respectively. These four subpopulations in the ensemble can illustrate the observed positive, negative and neutral coupling.

c.1 Positive coupling

Positive coupling is manifested when the states in which one residue (e.g. '*j*') in the folded state are stabilized as a result from the increase in stability of the states in which a second residue in the folded state (e.g. '*k*'); and vice versa. Therefore, under this constrain, $\sum_{i}^{N=states} P_{f,j|f,k} >> \sum_{i}^{N=states} P_{f,j|nf,k}$ and $\sum_{i}^{N=states} P_{nf,j|nf,k} >> \sum_{i}^{N=states} P_{nf,j|f,k}$. Thus, the residue stability constant from equation 3.8 becomes,

$$\kappa_{f,j} \cong \frac{\sum P_{f,j|f,k}}{\sum P_{nf,j|nf,k}} \tag{3.9}$$

When an energetic perturbation, $\phi^{pert,k} (= Exp[-\Delta g^{pert,k} / RT])$ on residue '*k*' in the folded state is evaluated, the stability constant of residue '*j*' becomes,

$$\kappa_{f,j}^{pert,k} \cong \frac{\phi_{f,k} \cdot \sum P_{f,j|f,k}}{\sum P_{nf,j|nf,k}} \tag{3.10a}$$

Analogously, in the case of the introduction of an energetic perturbation on residue '*j*' in the folded state, the stability constant for residue '*k*' becomes,

$$\kappa_{f,k}^{pert,j} \cong \frac{\phi_{f,j} \cdot \sum P_{f,j|f,k}}{\sum P_{nf,j|nf,k}} \tag{3.10b}$$

In this studies, the energetic perturbation, $\phi^{pert,k}$, corresponds to an stabilizing energy equal to -2.0 kcal·mol^{-1}. Therefore, $\phi^{pert,k} = 30$. By introducing equations 3.10a and 3.10b into equation 3.7, it is obtained an expression for the case of positive coupling between residues (*j*, *k*),

$$\Delta\Delta G_{j,k} \cong -RT \cdot \ln\left[\frac{\dfrac{\sum P_{f,j|f,k}}{\sum P_{nf,j|nf,k}}}{\dfrac{\phi_{f,k} \cdot \sum P_{f,j|f,k}}{\sum P_{nf,j|nf,k}}} \cdot \frac{\dfrac{\sum P_{f,j|f,k}}{\sum P_{nf,j|nf,k}}}{\dfrac{\phi_{f,j} \cdot \sum P_{f,j|f,k}}{\sum P_{nf,j|nf,k}}} \right] \tag{3.11a}$$

or

$$\Delta\Delta G_{j,k} \cong -RT \cdot \ln\left[\frac{1}{\phi_{f,k}} \cdot \frac{1}{\phi_{f,j}}\right] \tag{3.11b}$$

Since $\phi^{pert,k}$ is a positive number, then the value of $\Delta\Delta G_{j,k}$ will be positive as well. Importantly, positive coupling will hold when only one constrain is considered in the analysis and not the other one.

c.2 Negative coupling

Analogously, in the case of negative coupling the stability of one residue will result in the destabilization of the second one, and vice versa. Therefore, $\sum_{i}^{N=states} P_{f,j|nf,k} >> \sum_{i}^{N=states} P_{f,j|f,k}$ and $\sum_{i}^{N=states} P_{nf,j|f,k} >> \sum_{i}^{N=states} P_{nf,j|nf,k}$; and equation 3.7 becomes:

$$\Delta\Delta G_{j,k} \approx -RT \cdot \ln\left[\frac{\frac{\sum P_{f,j|nf,k}}{\sum P_{nf,j|f,k}}}{\frac{\sum P_{f,j|nf,k}}{\phi_{f,k} \cdot \sum P_{nf,j|f,k}}} \cdot \frac{\frac{\sum P_{nf,j|f,k}}{\sum P_{f,j|nf,k}}}{\frac{\sum P_{nf,j|f,k}}{\phi_{f,j} \cdot \sum P_{f,j|nf,k}}}\right] \tag{3.12a}$$

or

$$\Delta\Delta G_{j,k} \approx -RT \cdot \ln\left[\phi_{f,k} \cdot \phi_{f,j}\right] \tag{3.12b}$$

In this case, an stabilizing energetic perturbation will result in a negative value of $\Delta\Delta G_{j,k}$.

c.3 Neutral coupling

Neutral coupling arises from the absence of any perturbation on residue '*j*' due to the stabilization of residue '*k*'. In this case the denominator and numerator of equation 3.7 becomes unity, which results in $\Delta\Delta G_{j,k} = 0$.

d. Mutational response on the energetic coupling of ED3 ($MR_{j,k}^{mut-wt}$)

To investigate the potential long-range effect of naturally occurring mutation in ED3 from West Nile virus we can characterize the mutual perturbation between each residue pair in the presence of the mutation ($\Delta\Delta G_{j,k}^{mut}$) as described in the previous section. Then, equation 3.7 becomes:

$$\Delta\Delta G_{j,k}^{mut} = [\Delta\Delta G_{j}^{pert,k}]^{mut} + [\Delta\Delta G_{k}^{pert,j}]^{mut} \qquad (3.13)$$

To dissect the effect of a naturally occurring mutation on the long-range communications (or energetic coupling) of ED3 wild type ($\Delta\Delta G_{j,k}^{wt}$), we defined the mutational response (MR) as:

$$MR_{j,k}^{mut-wt} = \Delta\Delta G_{j,k}^{mut} - \Delta\Delta G_{j,k}^{wt} \qquad (3.14)$$

3.4 RESULTS

The computational studies on ED3 from WNV were performed with the 10th conformer deposited in the NMR structure (PDB 1S6N). The 10th conformer (out of fifteen) was chosen because it represents the structure closest to the centroid of the largest residue-cluster in ED3. This cluster represents more that 90 % of the total residues in ED3. This analysis was obtained by the application programs NMRCORE and NMRCLUST (http://pqs.ebi.ac.uk/pqs-nmr.html).

In Figure 3.3 it is shown the schematic representation of the enumeration process of microstates performed by the COREX algorithm. In the case of ED3 from WNV, windows of 8 residues were chosen as folding units with a minimum of 4. Under this condition, close to 200,000 microstates were obtained in 8 partitions.

3.4.1 Effect of the global Gibbs free energy of unfolding on the thermodynamic descriptor of ED3

To characterize each conformer generated by COREX, the probability of each conformer was calculated by using equation 3.1 in the Methods section. For this calculation, it is necessary to determine the Gibbs free energy for every microstate using as reference the high resolution structure or fully folded state. The enthalpic and entropic contribution to the Gibbs free energy for each conformer is calculated using a surface-area-based-parameterization. However, and as explained in the Methods section, 5% of the conformational entropy is not determined by COREX, which can be normalized by using a conformational entropy-scaling factor based on the experimental global free energy of unfolding. Because we have previously determined the global Gibbs free energy of unfolding (Chapter 1) for wild type ED3 and three single mutants (K310T, T332A and T332K), we were able to determine the conformational entropy-scaling factor without assuming a hypothetical global unfolding free energy. The single mutants K310T, T332A and T332K had a destabilizing effect of ~0.5 kcal·mol^{-1} relative to ED3 wild type, which had a value of 4.5 kcal·mol^{-1}. Therefore, for additional mutations, we utilized a Gibbs free energy of unfolding of 4.0 kcal·mol^{-1}

a. Dependence of the residue stability constant on the free energy of unfolding

The relevance of the accuracy of the conformational entropy-scaling factor is illustrated in Figure 3.4, which shows the changes in the residue stability constant, determined using equation 3.5, for wild type ED3 as a function of the free energy of

unfolding. Noteworthy, the residues with highest stability constant were the most affected by the changes in the global free energy of unfolding; while the residues with lowest stability constant remained unchanged. The changes in free energy of unfolding will also have a large impact in the energetic coupling between in each residue pair in ED3 as shown in the next section.

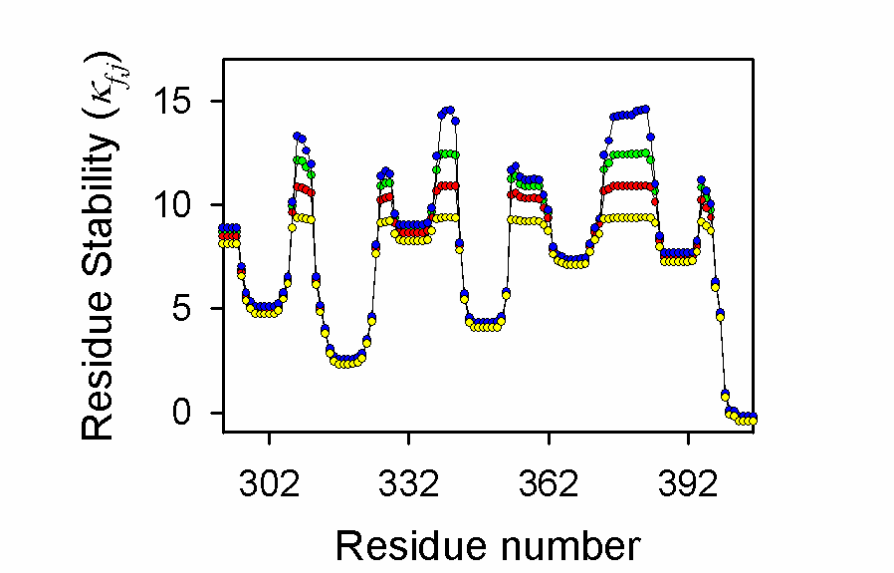

Figure 3.4. Effect of increasing free energy of unfolding in the residue stability constant in wild type ED3 from WNV. The color coding corresponds to different global free energies of unfolding as follows: yellow (4.5 kcal·mol^{-1}), red (5.5 kcal·mol^{-1}), green (6.5 kcal·mol^{-1}) and blue 7.8 (kcal·mol^{-1}).

b. Dependence of the energetic coupling on the free energy of unfolding

Changes in the global free energy of unfolding also had an impact in the energetic coupling calculations. The energetic coupling between pairs of residues is rendered as color maps (or heat maps) in which positive, negative and neutral coupling are represented in red, blue and green colors, respectively. When the stability of ED3 is increased from 4.5 kcal·mol^{-1} to 7.8 kcal·mol^{-1}, the degree of cooperativity in decreased significantly. Figure 3.5A shows residues in red color that share a high degree of positive coupling forming a cooperative core in the protein. By mapping the residues sharing high

positive coupling onto the NMR structure of wild type ED3[19], it was observed that as the global stability increases in ED3, the cooperative core reduces significantly not only in magnitude but in the number of residues comprising the cooperative core (Figure 3.5D).

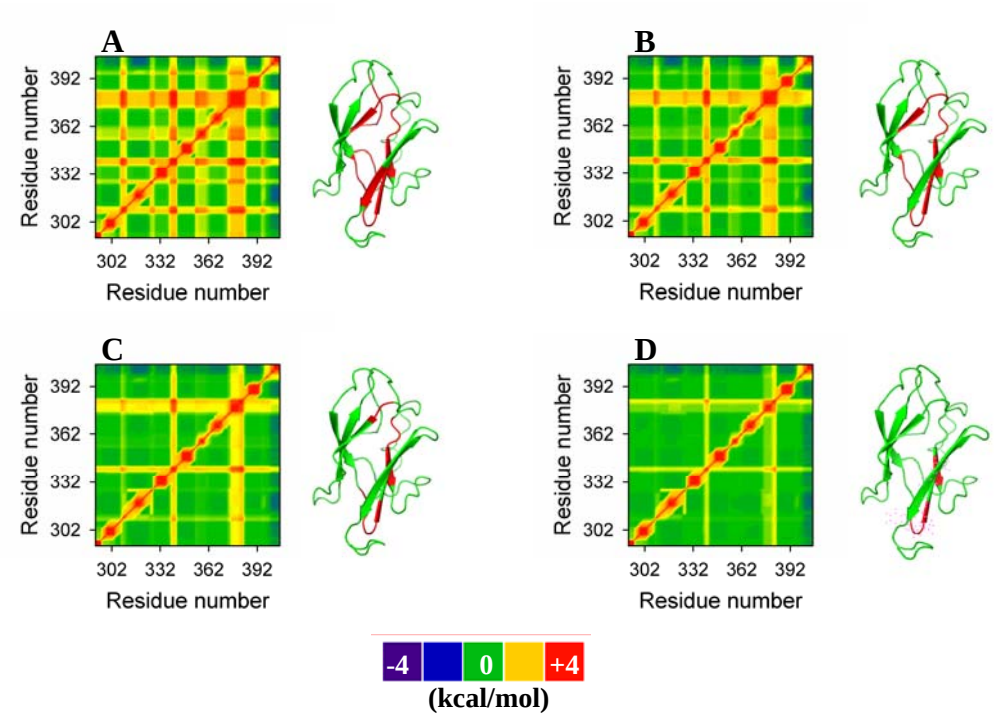

Figure 3.5. Effect of increasing free energy of unfolding in energetic coupling ($\Delta\Delta G_{j,k}$) in ED3 wild type from WNV[19]. The letter coding corresponds to different free energy of unfolding: **A** (4.5 kcal·mol^{-1}), **B** (5.5 kcal·mol^{-1}), **C** (6.5 kcal·mol^{-1}) and **D** (7.8 kcal·mol^{-1}). The conformational entropy-scaling factors were 0.955, 0.950, 0.945 and 0.938 kcal·mol^{-1}·K^{-1}, respectively. The ribbon diagram was rendered using PyMOL v. 0.97 (Delano Scientific LLC, San Carlos, CA).

3.4.2 Experimental validation of the ensemble-based calculation of energetic coupling in ED3

a. Calculation of the energetic coupling in single mutants of ED3

Because only the high-resolution structure of wild type ED3 is available, all the calculations for the single mutants required modification *in silico* of the wild type protein.

To avoid any perturbation of the aminoacid backbone due to the mutation in ED3, the calculations with COREX were performed without any further energy minimization of the new mutant structure. However, to prevent atomic clashes due to the substitution, the COREX algorithm will determine the best rotamer for the new aminoacid side chain. Noteworthy, all the calculations using COREX performed on single mutants of WNV ED3 have been reported in naturally occurring variants of the virus (Table 3.1). Thus, it is not expected that this mutations will significantly impair the global fold of ED3 wild type, because the integrity of this domain is required for virus viability during infection, multiplication and viral-particle assembly processes. For example, when introduced a non-naturally occurring mutation into WNV ED3 (e.g. Y329F/K or D333A) the virus was not viable[96]. These observations are in agreement with out previous studies that have shown that the naturally occurring single mutations K310T, T332A and T332K in ED3 do not have any significant effects in the secondary structure and global solution properties (Figure 3.15, Chapter 2).

b. Ensemble-based description captures the long-range communications in ED3

The arguments presented in the preceding section support the modification *in silico* of wild type ED3 to calculate the energetic coupling pattern of single mutants. However, we have, in addition, experimentally validated the application of the COREX algorithm. Because the COREX algorithm retrieves the thermodynamic information of the target protein (i.e. ED3) by using a surface-area based parameterization or changes in solvent accessible surface area[91]; then the effect over residue '*j*' due to an energetic perturbation on residue '*k*', described by $\Delta\Delta G_j^{pert,k}$ in equation 3.6, was correlated with experimental values of changes in solvent accessible surface area of residue '*j*' due to a mutation at residue '*k*' ($\Delta ASA_j^{mut,k}$). In this particular case, we use as experimental values the changes in solvent accessible surface area of the single tryptophan residue, W397, in the presence of the mutations K310T, T332A and T332K relative to wild type ED3. The distance between the mutation sites and W397 are ~20 Å and 30 Å for K310 and T332

(respectively). Therefore, the perturbation indicates long-range effects. These data was obtained from Table 4 in Chapter 1.

There was a strong correlation (R=0.98) between the calculated and experimental long-range effect over W397 exerted by the mutations described above as shown in Figure 3.6. The excellent correlation factor demonstrates that the energetic coupling determined by the COREX algorithm accurately describes the long-range communications in wild type ED3 and the effect of single mutants at the residue level.

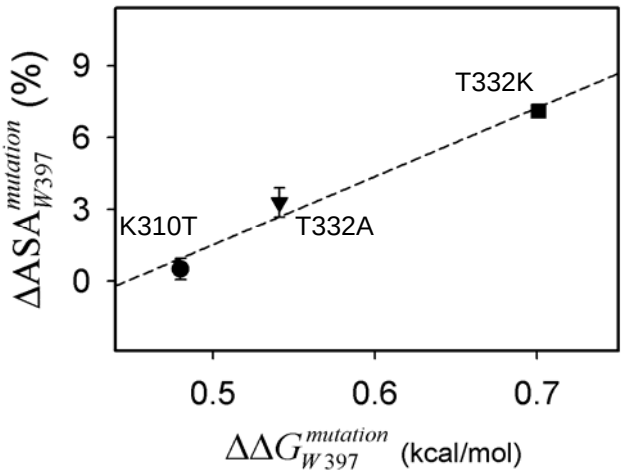

Figure 3.6. Correlation between calculated and experimental long-range effects. The linear correlation (R = 0.98) between the calculated effect of a mutation on the energetic coupling of residue W397 due to a distant mutation ($\Delta\Delta G_{W397}^{mutation}$), and the experimental values of changes in solvent accessible surface area of the same residue ($\Delta ASA_{W397}^{mutation}$) validates the application of the COREX algorithm. The symbols corresponds to mutations K310T (●), T332A (▼) and T332K (■). The distances between the residues K310-W397 and T332-W397 are ~20 Å and 30 Å.

3.5 DISCUSSION

3.5.1 ED3 in the viral particle is an ensemble of conformational states

Having thus validated the ensemble-based approach, we further investigated the long-range communications in wild type ED3 by calculating the mutual perturbation

between each residue pair (j, k) in the protein defined as $\Delta\Delta G_{j,k}$ and shown in Figure 3.7A.In addition to the expected local perturbation to nearby residues (delineated by the diagonal running from the lower left to the upper right), there is both positive and negative energetic coupling between groups of residues that are not close in space, which can be as much as 28 Å away from each other (i.e., residues 307-311 and 375-384 shown in Figure 3.7B). Furthermore, many other groups of distant residues share high positive energetic coupling indicating that there is a complex long-range communication network immersed in the native state ensemble of ED3 (Figure 3.7C). Therefore, ED3 in the viral particle does not behave as a rigid structure, rather as a conformational ensemble that is capable to transmit long-range effects.

3.5.2 The energetic coupling pattern of ED3 wild type is thermodynamically robust

Upon investigation of the long-range communications of ED3 harboring naturally occurring single mutations of WNV, which are distributed throughout its structure (Figure 3.1), we found very similar overall energetic coupling patterns compared to wild type ED3. In Figure 3.8 it is shown the energetic coupling patterns of ED3 harboring the mutations K310T, T332A and T332K. It is evident that the energetic coupling patterns between residue pairs of groups of residues are very similar in all cases, whether or not the mutation confers MAb-neutralization resistance to the virus, suggesting that the basic mechanism of interacting residue-networks is conserved as in wild type ED3. This observation is consistent for all mutations listed in Table 3.1. Therefore, the fold of ED3 is thermodynamically robust despite of the presence of single mutations. This conclusion is in agreement with our preceding studies that showed that single mutations maintained the same overall fold, size and shape of wild type ED3 as judged by essentially identical spectral and transport properties[84].

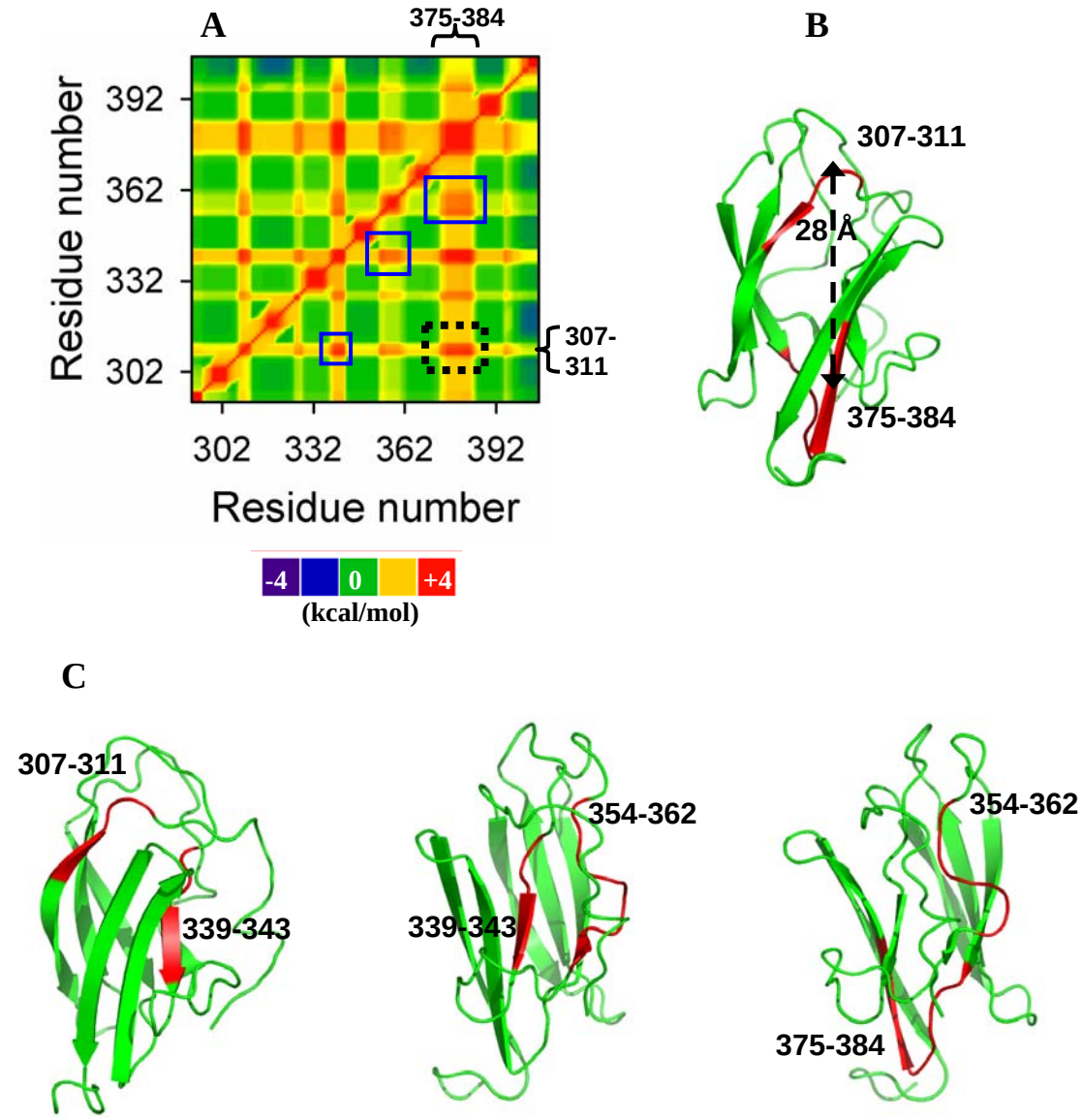

Figure 3.7. Residue networks in wild type ED3. **A**, Bidirectional energetic coupling or mutual perturbation between pair of residues in ED3 wild type ($\Delta\Delta G_{j,k}$). Groups of residues in boxes represent a few examples of distant residues that are highly energetically coupled (≥ 2 kcal·mol^{-1}), conforming a residue interaction network. **B and C**, 3D representation of resides energetically coupled (≥ 2 kcal·mol^{-1}), even when they are located more than 28 Å away from each other. The ribbon diagram was rendered using PyMOL v. 0.97 (Delano Scientific LLC, San Carlos, CA).

Table 3.1. Mutations found in ED3 from West Nile virus variant strains

West Nile virus strain	Harbored mutations
Lineage 1[a]	
USA99b	Wild type ED3
ISR52	H398Y
ISR53	***T332A***
AUS60	K310T, V338I, A365S
IND80	L312V, V338I, E390D
Lineage 2[a]	
MAD78	L312V, A369S, V371I, V375I
MAD88	L312A, ***T332K***, A369S
SA89	L312A, A369S
SA58a	L312A, ***T332K***, A369S
Selected *in vitro*	
385-99 MAb^R-5C5[b]	***K307R***
385-99 MAb^R-5H10[b]	***T330I***
rED3-K307E[c]	***K307E***

[a] Li et al. (2005)[16]. [b] Beasley and Barrett (2002)[15]. [c] Oliphant et al. (2006)[14]. Mutations in bolded and italics correspond to mutations that decrease MAb binding relative to wild type ED3.

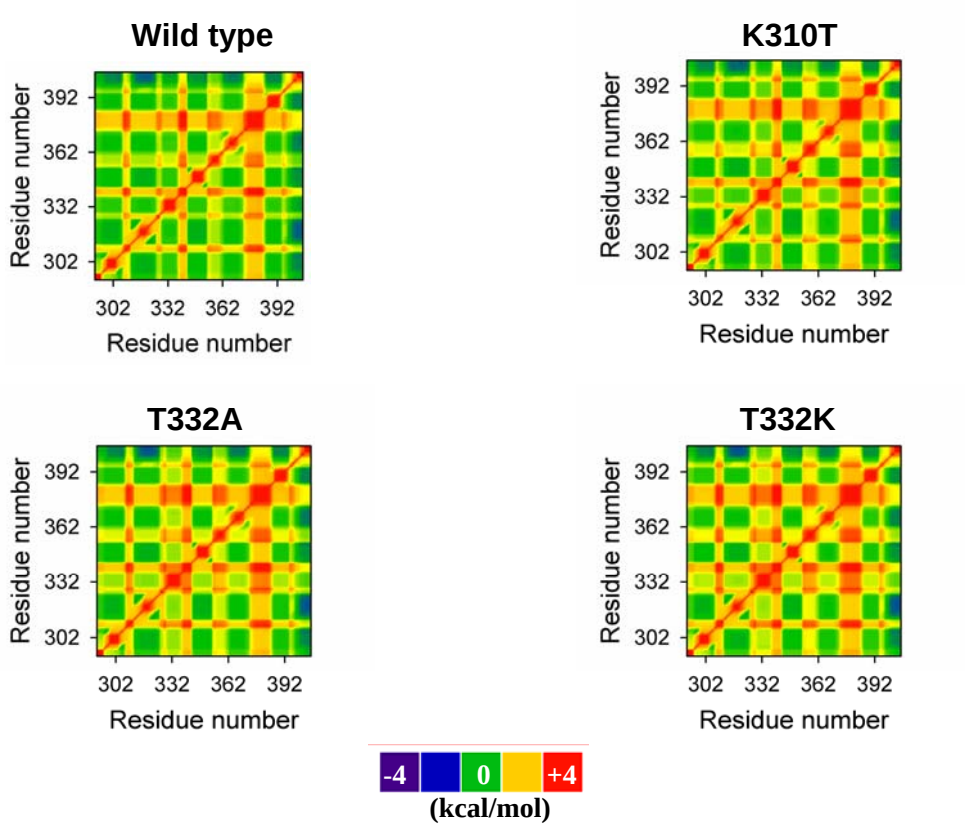

Figure 3.8. The energetic coupling pattern ($\Delta\Delta G_{j,k}$) or mutual perturbation between pair of residues in ED3 wild type from WNV is thermodynamically robust. The mutation K310T does not generate a resistant variant of the virus, while the mutations T332A and T332K generate a partially and fully resistant virus.

3.5.3 Single mutations in ED3 modulate the ensemble of conformational states

To amplify potential differences in the energetic coupling patterns between a single mutant and wild type ED3, we calculate the difference in energetic coupling between each residue pair and defined it as the mutational response of the energetic coupling: $MR_{j,k}^{mut-wt}$ (described by equation 3.14 in the Methods section).

We found that despite the thermodynamic robustness of the wild type ED3 fold manifested by the same overall energetic coupling pattern compared to the single mutants (Figure 3.8), at the residue level the MAb-resistant mutations generated very significant

effects localized on residues 328 to 338 (Figure 3.9). For example, the single mutant T332K, which has been shown to decrease MAb binding by 100-fold relative to wild type ED3 and generates a MAb-neutralization resistant strain in WNV[15,16,84], had a high positive energetic coupling over residues 328 to 338 (Figure 3.9C) compared to ED3 wild type. The average differential energetic coupling value in this region, defined as $< MR^{mut-wt}_{328-338} >$, was ~0.85 kcal·mol^{-1}. Similarly, the mutation T332A, which generates a partially resistant strain in WNV[16], had a significant but smaller effect over the same group of residues (0.68 kcal·mol^{-1}, Table 3.2). This result shows that not only the position of the mutation site can modulate the energetic coupling distribution in ED3 but the type of amino acid substitution as well.

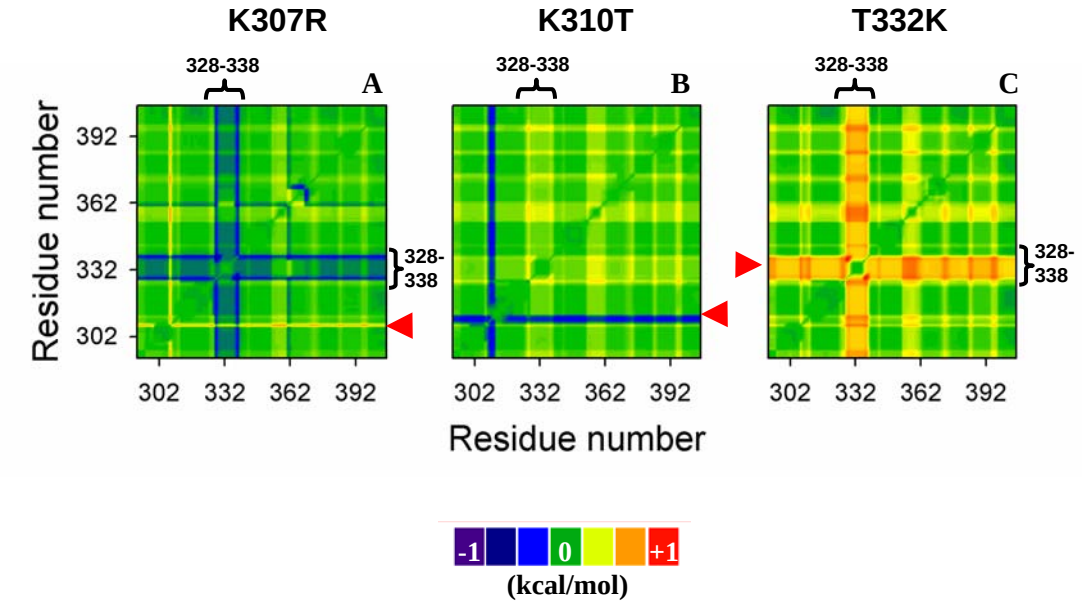

Figure 3.9. Continuum behavior in the modulation of MAb-neutralization resistance. Different mutations perturbed the same group of residues (328-338) by different magnitudes ranging from negative (**A**, K307R), intermediate (**B**, K310T) and positive (**C**, T332K) values of energetic coupling relative to the wild type ED3. The red arrow heads show the location of the mutation. Note that the mutations also have local effects in the protein. Each plot corresponds to the effect of the mutation on the long-range communication of ED3 wild type defined as $< MR^{mut-wt}_{j,k} >$.

Surprisingly, the MAb-resistant mutations at position K307 (by an R or E) also perturbed almost exclusively residues 328 to 338, even though the mutation site is outside the affected region. However, in the case of K307R, the effect is opposite compared to T332K; namely, it generated negative energetic coupling of –0.2 kcal·mol^{-1} (Figure 3.9A). Analysis of mutations that have been shown to not affect MAb binding in WNV (i.e., K310T)[15,16,84], revealed minimal effects over residues 328 to 338, or other regions in ED3 (Figure 3.9B).

The biological importance of residues 328 to 338 is that they map to a solvent-exposed loop in the viral particle connecting beta strands B and C in ED3 (named BC loop, Figure 3.1B) that has been identified as the major MAb neutralizing epitope in WNV by our group and others[15,16,84]. Thus, this ensemble-based description of the native state precisely captures the biologically relevant regions in ED3; the perturbation of which WNV uses to acquire MAb-neutralization resistance.

3.5.4 MAb-neutralization resistance behaves as a continuum phenomenon

Both the MAb-resistant and MAb-nonresistant mutations were energetically coupled to the same region, the BC loop in ED3, but with different magnitudes ranging from most negative (K307R/E), to most positive (T332A/K), via intermediate values (K310T) (Figure 3.9). Thus, it would be predicted that mutations at any position in ED3, which are energetically coupled to the BC loop to a higher magnitude relative to an intermediate value (corresponding to MAb-nonresistant mutants) will result in a MAb-neutralization resistant phenotype. Therefore we investigated the effect of several mutations over the BC loop of ED3, which are located throughout the ED3 structure (Figure 3.1A). These mutations were selected from different strains of WNV and cover a broad spectrum of MAb-neutralization resistance (Table 3.1) [15,16].

The effects of 15 naturally occurring mutations in WNV ED3 over the BC loop was plotted in rank order –from most negative to most positive values of $< MR_{BC\ loop}^{mut-wt} >$– to avoid any bias from previous knowledge of whether the mutation comes from a MAb-

neutralization resistant variant or not (Figure 3.10). All mutants that generated high positive or high negative values of $< MR_{BC\ loop}^{mut-wt} >$ corresponded to mutations that reduce MAb binding affinity by ~10-100-fold compared to wild type ED3[16,84] (Table 3.2). These values deviate from the mean by ± 0.3-0.6 kcal·mol^{-1}. The remaining mutations with intermediate values ($< MR_{BC\ loop}^{mut-wt} > \cong 0.34$ kcal·mol^{-1}) had essentially identical binding affinities to MAbs relative to wild type ED3 and corresponded to strains of WNV that were not resistant to MAb-neutralization[15,16].

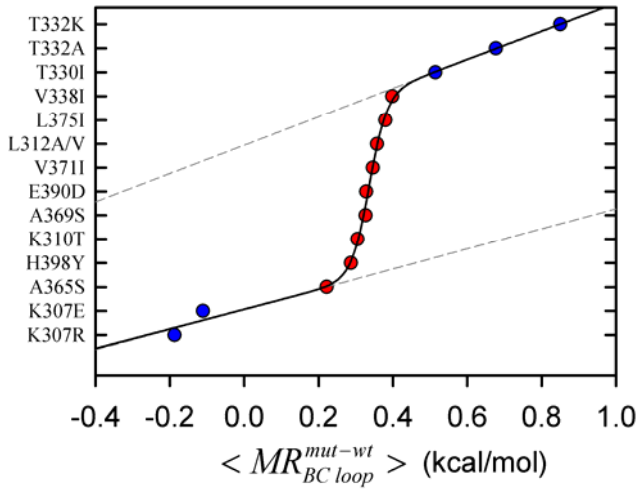

Figure 3.10. Boltzmann distribution of the energetic coupling perturbation in the BC loop ($< MR_{BC\ loop}^{mut-wt} >$). Rank order plot of the averaged difference of energetic coupling in the BC loop (residues 328-338) for single mutations found in WNV variants. The solid line represents the Boltzmann fit (equation 3.15) and was obtained by assuming identical distances in the rank order between each mutant. The dotted lines correspond to the pre- and post-transition baselines. The midpoint value was 0.34 kcal·mol^{-1}, and the maximum deviances correspond to ± 0.3-0.6 kcal·mol^{-1}. Blue spheres represent MAb-resistant mutations, while red spheres MAb-nonresistant mutations.

We observed that the data in Figure 3.10 had a sigmoidal shape and was fitted to a Boltzmann distribution (solid line) using the following equation,

$$F(x) = pre(x) + [post(x) - pre(x)] \cdot f_{post}(x) \quad (3.15)$$

where 'x' is the mutational response over the energetic coupling of wild type ED3 or $< MR_{BC\ loop}^{mut-wt} >$; *pre(x)* and *post(x)* represent the linear baselines before and after the transition of the distribution, respectively; and $f_{post}(x)$ is defined as,

$$f_{post}(x) = \frac{K(x)}{1 + K(x)} \quad (3.16)$$

where $K(x) = Exp[-C \cdot (x - x_0)/R \cdot T]$, and represents an equilibrium constant in the distribution of conformational states between populations characterized by negative energetic coupling ($MR_{BC\ loop}^{mut-wt}$ < 0 kcal·mol^{-1}) and high positive energetic coupling ($MR_{BC\ loop}^{mut-wt}$ > 0.34 kcal·mol^{-1}). The value 'x_0' is the mean average mutational response corresponding to the midpoint of the transition (= 0.34 kcal·mol^{-1}). The constant 'C' is a dimensionless parameter that has been described as an equilibrium cooperative factor (C = 24.6) in other systems[97,98]. R and T are the gas constant (1,987 cal·mol^{-1}·K^{-1}) and absolute temperature (293 K).

The Boltzmann distribution of the rank order plot of the energetic coupling indicates that the effect of mutations located in different regions in ED3 behaves like an equilibrium process between different conformational states in the ensemble that must include both the MAb-non neutralization resistant and MAb-resistant-like conformations. Thus, the function *F(x)* of equation 3.15 represents an phenomenological description of the probability of a mutation of generating MAb neutralization resistance in the WNV, in which the modulation resistance will depend on the ability of a mutation to shift the equilibrium of the ensemble towards a conformation characterized by high positive or negative energetic coupling; i.e., significant deviations from a mean value of $< MR_{BC\ loop}^{mut-wt} >$.

Table 3.2. Correlation between Mutational Response in the BC loop ($< MR_{BC\ loop}^{mut-wt} >$) and MAb binding affinity (Kd_{app}, nM) in WNV ED3

	Mutational Response in BC loop ($< MR_{BC\ loop}^{mut-wt} >$)	MAb binding affinity[a] (Kd_{app}, nM)
Wild type	0.00	0.17 ± 0.03
K307R	-0.20	> 10.0[b]
K310T	0.31	0.39 ± 0.05
T330I	0.52	> 7.50[b]
T332A	0.68	1.00 ± 0.16
T332K	0.85	> 10.0[b]

[a] Binding affinity values were obtained from Refs. 15, 16 and Figure 3.21, Chapter 3. [b] For T330I, T332K and K307R the Kd_{app} is > 7.5 nM because there was no detectable binding at [Mab] ≥ 7 nM (corresponding to ~50-fold greater [Mab] than the Kd_{app} wild type ED3).

Therefore, these results suggest that MAb neutralization resistance behaves as a continuum phenomenon where mutations, independent of their location, decrease MAb neutralization by modulating the ensemble of conformational states, in which the BC loop adopts a MAb-resistant-like conformation. We do not wish to imply any specific mechanism or model from this phenomenological observation, as other schemes will be able to fit the data; however, our description is the most parsimonious explanation for the data presented.

We did not find any significant effects other than in the BC loop, including other regions of interaction with MAbs such as the DE and FG loops[24]. Both, MAb neutralization resistant and MAb-nonresistant mutations had mean values of $< MR_{DE\ /\ FG\ loops}^{mut-wt} >$ around ~0.2 kcal·mol^{-1}, even if the mutation is located in one of the

loops (i.e. A365S is located in the DE loop). In agreement with these observations, the mutations in the loops DE and FG have been shown to not decrease MAb binding affinity. This behavior is also observed in the energetic coupling map in Figure 3.9, where most of the significant effects are located only in the BC loop. Thus, there was no evidence for effects in other loops or in any other region in ED3 caused by the MAb-resistant or nonresistant mutations studied here, despite the fact that the DE and FG loops are also part of the binding site of a previously described neutralizing MAb[24].

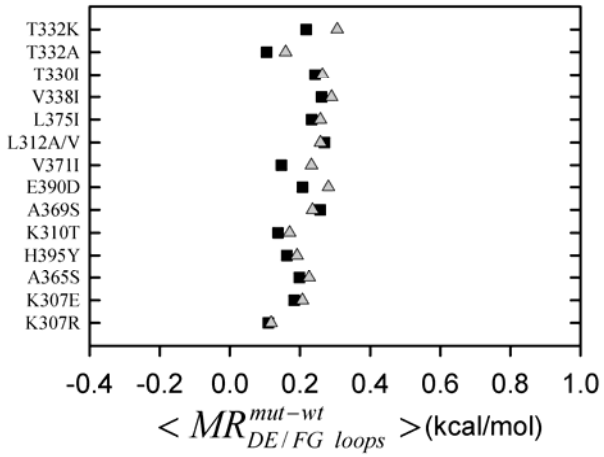

Figure 3.11. Plot of the averaged difference of energetic coupling in the DE and FG loops ($< MR^{mut-wt}_{DE\ /\ FG\ loops} >$). The rank order was kept as in Figure 3.10. All the values for the mutants oscillate near the mean value of 0.2 kcal·mol^{-1}. Squared and triangular symbols correspond to the DE and FG loops, respectively.

3.5.5 Physical meaning of modulation of energetic coupling by MAb-resistant mutations

We have observed that the modulation of the energetic coupling in the BC loop, $< MR^{mut-wt}_{BC\ loop} >$, follows a Boltzmann distribution. While this behavior introduces the

concept of an equilibrium process between multiple conformations in ED3 involved in MAb neutralization resistance, it does not provide any information about the effect of mutations at the molecular or structural level. However, analysis of the effect of the MAb-resistant mutations on the *residue stability constant* defined as $\kappa_{f,j}$ (see Methods section), provides insight into the effect on the stability of each residue in ED3 due to a mutation (Figure 3.12). The mutants shown are K310T, a non-resistant mutation; and, T332K and K307R, which are MAb-resistant. The mutant K310T has a very similar value of $< MR_{BC\ loop}^{mut-wt} >$, while T332K and K307R have high positive and negative values in the same region, respectively.

The main effect over $\kappa_{f,j}$ is again located in the region comprising residues 328 to 338 or the BC loop, in which each mutant had a different behavior. In the case of K310T, which does not have any differential effect compared to ED3-WT, the residues corresponding to the BC loop had higher values of $\kappa_{f,j}$ compared to the calculated global stability (defined as the average residue stability or $<\kappa_{f,j}>=\frac{1}{N}\sum_{j}^{N=total\ residues}\kappa_{f,j}$). However, relative to the maximum residue stability, residues in the BC loop were less stable (Table 3.3). In Figure 3.12B, it is shown the stability corresponding to the residues in the BC loop. The ratio of stability between the BC loop and the maximum stability for K310T is 0.94. In the case of the mutant T332K, this ratio was 0.98, and for K307 it was 0.75. The latter indicates that for this mutant there was a 20% decrease in stability in the BC loop relative to the maximum stability of wild type ED3 (or K310T). For T332K there was an increase of stability by 11%, and K310T had comparable values with wild type ED3. While Figure 3.12B shows differences in local regions of the protein (BC loop), the global stability for K310T, T332K and K307R were essentially identical.

The importance of the analysis of the magnitude of this ratio, which represent the probability of local unfolding to the probability of global unfolding or

$\left(\frac{1}{n} \sum_{j}^{\substack{n=residues \\ in\ BC\ loop}} \kappa_{f,j} \right) \Big/ < \kappa_{f,j} >$, is that it reflects the cooperative response of a given residue, or group of residues, relative to the rest of the protein[26,31].

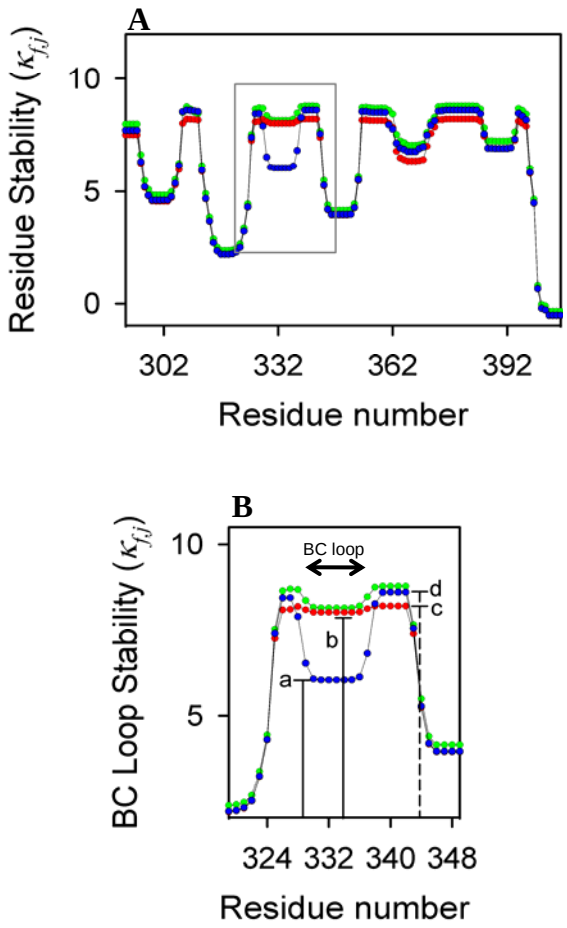

Figure 3.12. Physical interpretation of modulation of energetic coupling or mutation response. **A**, Residue stability plot for the mutants K307R (blue), K310T (green) and T332K (red). Box encloses the residues mostly affected by the presence of the mutations and corresponds to residues 328-338. **B**, Relative stability of residues 328-338 of BC loop in ED3. Labels 'a', 'b' and 'c' correspond to the stability of residues 328-338 for K307R, T332K and K310T, respectively. At the same time, label 'c' matches the maximum residue stability of T332K. Label 'd' corresponds to the maximum stability of K307R and K310T.

Therefore, residues with the lowest ratios will have a low cooperative response as in the case of the mutant K307R that shows negative coupling relative to wild type. This can be interpreted as if the BC loop in the presence of the mutations K307R/E is out-of-phase by being unfolded with very high probability without jeopardizing the integrity of the fold of ED3. The opposite was observed for the mutant T332K with a high ratio and high cooperative response. Thus, the BC loop in the presence of the T332K mutation is in-phase with the most stable residues in ED3, which are frequently located in the core of the protein. Reported partially MAb-resistant mutations such as K307E, T332A and T330I[15,16] had intermediate values of these ratios and, accordingly, have intermediate cooperative responses, either in-phase or out-of-phase, which supports the continuum-like behavior in the mechanism of MAb neutralization resistance. The mutation A374S, a reported MAb-nonresistant mutation[16], behaves as wild type ED3 or K310T.

The calculations of the cooperativity responses of the BC loop suggest that in the case of the mutations K307R/E, this region might be in a highly dynamic or an unfolded-like conformation without risking the fold integrity of ED3 required for proper assembly of the virus. Whereas the mutation T332K generates a very stable and rigid BC loop that probably does not have the flexibility required to be engaged by MAbs. Interestingly, the BC loop in the presence of the mutations K310T, and other MAb-nonresistant mutations, is as stable as the wild type ED3; probably within the range of conformational dynamics where MAbs are able to interact efficiently with protein. This interpretation will require a direct experimental validation; however, it is in agreement with previous studies showing that epitopes that have too high dynamic behavior are not recognized by MAbs[88-90] and correlates very well with our binding data.

Table 3.3. Effect of mutations in the Stability of the BC loop in WNV ED3[a]

	Global Stability[b]	BC loop (Local) Stability[c]	*Ratio Local/Global Stability*	Maximum Residue Stability[d]	*Ratio Local/Max Stability*	Effect of mutations on BC loop stability relative to wild type ED3
Wild type	6.66	8.36	1.26	9.29	0.90	
K307R	6.13	6.20	1.01	8.44	0.73	<< *stability*
K307E	6.10	6.27	1.03	8.27	0.76	<< *stability*
A365S	6.27	7.72	1.23	8.45	0.91	*Behaves like wild type*
K310T	6.47	8.22	1.27	8.74	0.93	*Behaves like wild type*
T330I	6.23	8.01	1.29	8.34	0.96	> *stability*
T332A	6.33	8.48	1.34	8.65	0.98	>> *stability*
T332K	6.11	8.05	1.32	8.18	0.98	>> *stability*

[a] All stability units are in kcal·mol^{-1}. [b] The Global Stability was calculated by averaging the Residue Stability Constants of each residue in ED3. [c] BC loop stability was calculated by averaging the Residue Stability Constants of residues 328-338. [d] The Maximum Residue Stability found in the ED3 for wild type and single mutants.

3.6 CONCLUSIONS

MAb mediated-neutralization is a central component of the immune response against WNV and other flaviviruses as MAbs restrict and clear dissemination of infection[12]. Passive transfer of neutralizing MAb before and after WNV infection has been proven to increase survival in challenged animals[14]. Currently humanized MAbs represent possible therapeutic measures for humans. In the present study we have investigated the role of conformational dynamics in the mechanism of MAb neutralization resistance. For that purpose we have investigated the ensemble of conformational states of the ED3 from WNV by using the COREX algorithm. In examining the effect of naturally occurring mutations on the conformational fluctuations or energetic coupling between distant residues in ED3 we were able to determine a residue interaction network between the mutation sites (found in WNV variant strains) and functional regions of the protein. The determined network of residues showed strong correlation with experimental data and the mapping of the location of the major neutralizing epitope in WNV ED3 ; namely, the BC loop.

The energetic coupling describe here is founded on the thermodynamic basis of fluctuations of proteins in solution. Therefore, the physical interaction between distantly positioned residues suggests that there is not a direct mechanical pathway that allows propagation of perturbations. Instead, it shows that the propagation of an effect throughout the structure is originated by the redistribution of conformational states within the native state ensemble of the protein. We found that the redistribution of conformational states generated by naturally occurring mutations in WNV ED3 follows a Boltzmann distribution, indicating that there exists equilibrium between multiples conformations that is shifted not only depending on the position of the mutation site, but the nature of the amino acid substitution plays an important role as well. The Boltzmann distribution suggests that both MAb-resistant and MAb-nonresistant mutations are part of the same phenomenon that behaves as a continuum. However, only mutations that are able to generate MAb neutralization resistance were also able to perturb best this equilibrium towards a *MAb-resistant like* conformation, characterized either by very low

or very high stability of the BC loop relative to the global stability and maximum residue stability. The quantitative description of this continuum behavior suggests that there is an optimal range of conformational states required for MAb binding and neutralization.

A central result from this study is that multiple mutation sites are part of a residue network that affects principally one region of the protein via long-range interactions, the BC loop in WNV ED3. This result questions the mapping of multiple natural mutations as indicators of multiple epitopes in a protein antigen, unless direct structural evidence is presented. But what will the advantage of having one dominant epitope connected to a network of distantly positioned residues be? We can hypothesize that long-range interactions can expand the "mutational-space" of the virus, serving as an additional viral strategy to increase antigenic variation to evade the host immune system.

CHAPTER 4

IDENTIFICATION OF ANTIGENIC DETERMINANTS OF NEUTRALIZATION IN FLAVIVRUSES BY MUTATIONAL RESPONSE OF THE ENSEMBLE OF THE ENVELOPE PROTEIN DOMAIN 3

4.1 ABSTRACT

We have previously described quantitatively the phenomenon of MAb-neutralization resistance in West Nile virus (WNV) by using a continuous function of the energetic coupling among mutation sites in ED3 that followed a Boltzmann distribution, in which the BC loop was the region most susceptible to perturbations due to mutations in ED3. To test the generality of this results, we analyzed the mutational response on the energetic coupling of ED3 from Dengue virus type 2 (DENV2), a related flavivirus, that has a highly conserved fold compared to WNV ED3, but its antigenic properties are very different. By evaluating the ability of a panel of single mutations to modulate the ensemble of conformational states in DENV2 ED3, we determined that mutations in DENV2 ED3 perturbed residues located in the loop connecting beta-strand F-G, a solvent-exposed loop in the virus that corresponds to an immunodominant epitope in DENV2 ED3. Remarkably, we found that the perturbation of the energetic coupling of both the antibody-neutralization resistant and non-resistant mutations was consistent with a Boltzmann distribution, wherein the antibody-resistant mutants were best able to perturb this region (FG loop). These results are analogous to those found in WNV ED3, supporting the generality of the continuous-like behavior in antibody-neutralization resistance. However, the residues perturbed by mutations in WNV and DENV2 corresponded to virus-specific regions in ED3.

4.2 INTRODUCTION

Mosquito-borne flaviviruses include important emerging and resurging diseases of global significance. Two main examples are the resurgence of dengue in tropical and subtropical areas of the world, and WNV in new habitats and environments (i.e. the Americas)[1,99]. To date, there is no effective human vaccine for DEV2 or WNV[1,34]. Because of their high neutralizing activity *in vitro* and *in vivo* of DENV2 and WNV, monoclonal antibodies (MAbs) represent an attractive avenue for the development of an effective human vaccine[15,100,101]. The primary target of potent neutralizing MAbs in DENV2 and WNV is the envelope protein domain 3 (ED3)[14-16], which assumes an Ig-G-like fold conserved among all flaviviruses (Figure 4.1).

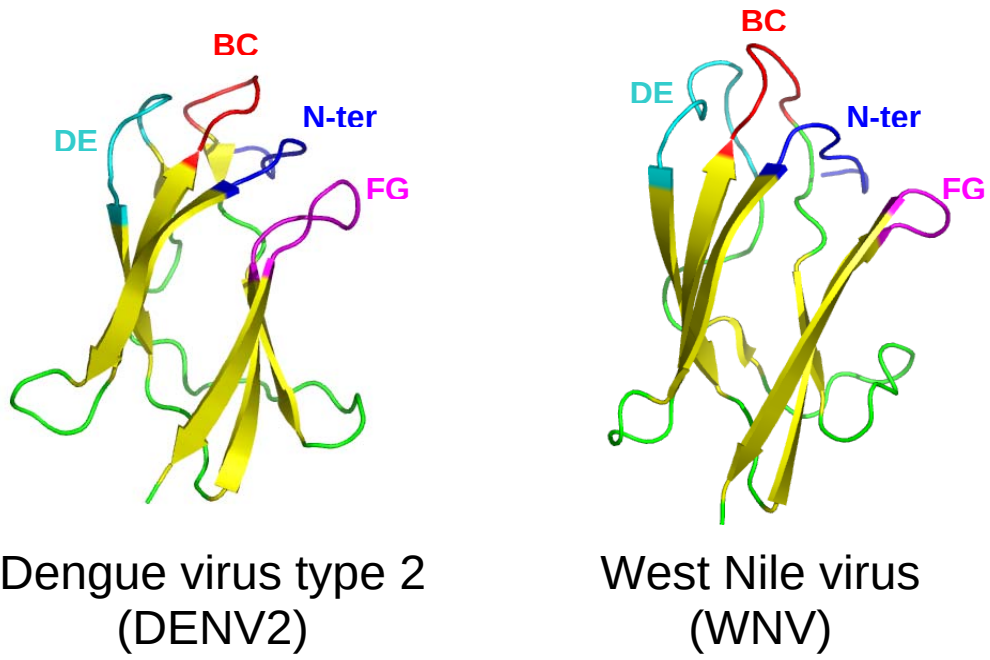

Figure 4.1. Structural comparison of the ED3 from Dengue virus type 2 (left; PDB 1TG8)[18] and West Nile virus (right; PDB 1S6N)[19]. Both ED3s have ~100 residues and are IgG-like proteins conforming two beta-sheets (yellow color in the cartoon). The r.m.s.d. between DENV2 and WNV ED3s is 2.5 Å (between all atoms). The loops that are solvent-exposed in the viral particle are the DE (cyan), BC (red), N-terminus (blue) and FG (magenta). The ribbon diagram was rendered using PyMOL v. 0.97 (Delano Scientific LLC, San Carlos, CA).

Despite that the structures of the ED3 in DENV2 and WNV are highly conserved, their antigenic properties are very dissimilar (e.g. a mutation in the same structurally homologous site can have two completely different outcomes in their effect on MAb binding)[33]. These differences in antigenic properties are not obviously derived either from their structure or from their amino acid sequence. In particular, the ED3 from DENV2 (strain NGC)[18] and WNV (strain USA99b)[19] have high structural similarly (*r.m.s.d.* = 2.5 Å between all atoms), and very high sequence identity and similarity (44 % and 60 %, respectively, Figure 4.2). Thus, the differences in antigenic properties, described as the phenomenon of virus-specific antigenic determinants, can not be explained solely by sequence or structural analyses. Because the identification of virus-specific residues that are important for binding of neutralizing MAbs is critical for the development of human vaccines[3], we investigated the underlying principles by which closely related flaviviruses have different but specific MAb-neutralization antigenic determinants.

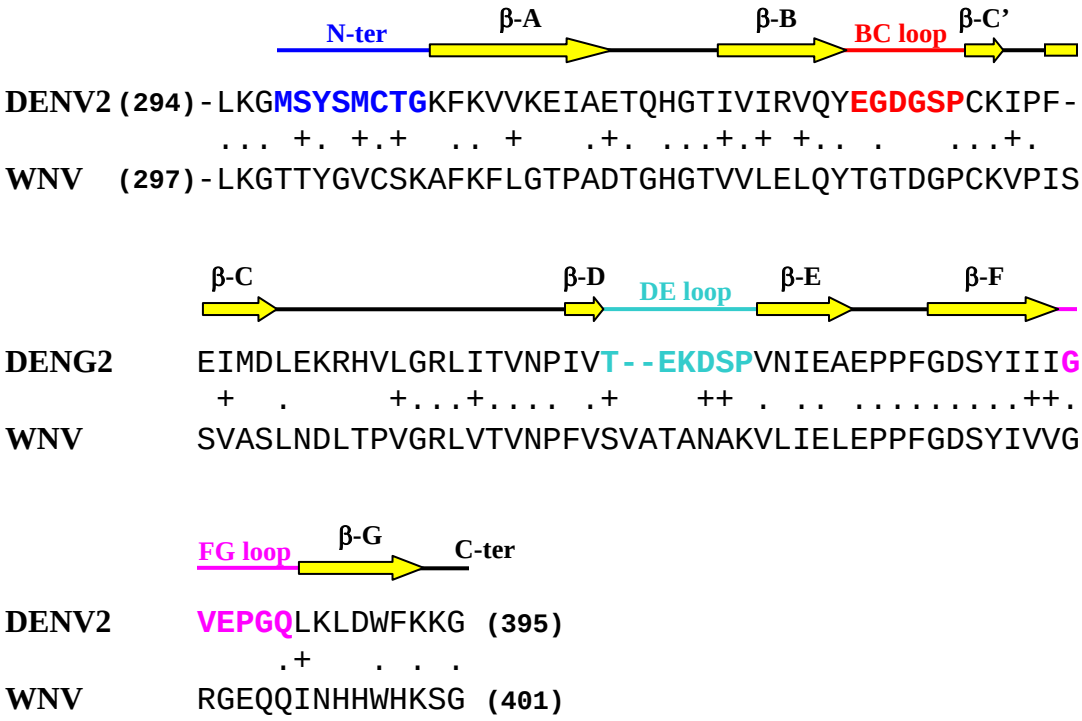

Figure 4.2. Sequence alignment of ED3 from Dengue virus type 2 (DENV2) and West Nile virus (WNV). Cartoon displaying the secondary structures were mapped using the tertiary structure from DEV2 ED3 in Figure 4.1. Dots and crosses indicate residue identity and similarity, respectively.

By using an ensemble-based description of the equilibrium[91], we have shown in the previous chapter (Chapter 3) that mutations in ED3 from WNV perturbed a residue interaction network of energetically coupled residues that followed a Boltzmann distribution; and the mutations that were best able to perturb the energetic coupling of the BC loop (Figure 4.1) corresponded to MAb-neutralization resistance mutations. Remarkably, the BC loop has been shown to be the immunodominant epitope in WNV ED3 by our group and others[14-16]. Based on this study, we investigated the effect of single site mutations on the residue interaction networks in ED3 from DENV2, and identified that the MAb-neutralization resistant mutations perturbed the energetic coupling of a region different compared to WNV ED3; namely, the FG loop (Figure 4.1). The FG loop has been proposed as the immunodominant epitope in DENV2 ED3[102-104]. These results for DENV2 ED3 are in parallel to the findings in WNV ED3, but not identical. Therefore, the virus-specific antigenic determinants depend on the localization of the response of MAb-neutralization resistant mutations on the energetic coupling of the viral protein; in this particular case, the FG and BC loops in ED3 for DENV2 and WNV, respectively.

4.3 METHODS

The methods used in this chapter correspond to the application of the COREX algorithm[91] to the envelope protein domain 3 (ED3) from Dengue virus type 2 (DENV2). Please refer to the method section in Chapter 3, where the applications of the COREX algorithm were described and explained.

4.4 RESULTS

4.4.1 Mutations located at structurally homologous sites have different effects in ED3 from DENV2 and WNV

Despite the high structural and sequence similarity (*r.m.s.d.* = 2.5 Å and 44 % sequence identity) between the ED3 from DENV2 and WNV (Figures 4.1 and 4.2), the

effect of single mutations at structurally homologous sites have significantly different effects in MAb binding affinity. For example, mutations at residue T332 generate a resistant phenotype in WNV, whereas mutations at the homologous position in DENV2 ED3 (S331) have no effect. The opposite is observed by mutations in position K310 for WNV ED3, and position K307 for DENV2 ED3. This observation was quantified by determining the MAb binding affinity for single mutations, at the sites described above, relative to the wild type ED3. In Figure 4.3A is shown the location of the mutations in the structures of DENV and WNV ED3s, which correspond to structurally homologous sites. In Figure 4.3B is plotted the dissociation constant, Kd_{app}, of the binding to type-specific MAbs for each mutant and wild type ED3. The single mutation S331A in DENV2 ED3 did not have any effect compared to its wild type protein. However, a mutation T332A in WNV ED3 decreased the Kd_{app} ~5-fold. In the case of a mutation a position K307G in DENV2, the Kd_{app} decreased by ~4-fold and in the case of WNV, the mutation K310T had no evident effect. The phenomenon described above reflects the virus-specific antigenic determinant for each virus. However, the underlying principles by which closely related flaviviruses have different but specific antigenic determinants remain elusive. We therefore investigated the effect of mutations on the residue networks in ED3 from DENV2 by determining the mutational response on the energetic coupling defined as $MR_{j,k}^{mut-wt}$ (equation 3.14, Chapter 3).

4.4.2 Generation of the conformational ensemble of ED3 from Dengue virus type 2

The computational studies on ED3 from DENV2 were analogously performed as in the case of ED3 from WNV (Chapter 3). In the case of DENV2, the structure of ED3 was obtained by modifying the x-ray structure of the entire ectodomain of the envelope protein (PDB 1TG8)[18]. Namely, heteroatoms and water molecules, and domains 1 and 2 were removed from the original structure, leaving only residues 294 to 395 corresponding to ED3 (Figure 4.1). As in the calculations for the ED3 from WNV, in the case of

DENV2 ED3 we used windows of 8 residues as *folding units* with a minimum of 4. Under this condition, the ensemble comprised ~50,000 microstates from 8 partitions.

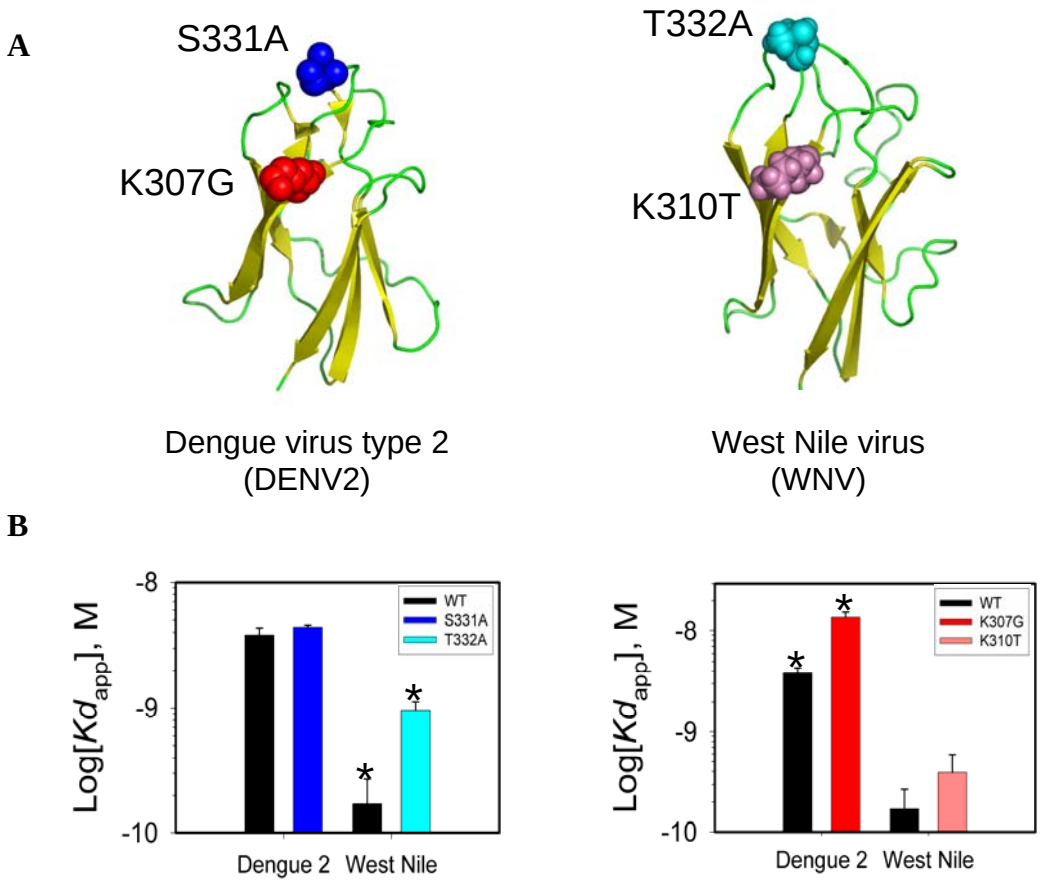

Figure 4.3. Mutations at structurally homologous sites have different effect in MAb binding affinity. **A**, The structurally homologous mutations in ED3 are S331A (blue spheres) and T332A (cyan spheres), and K307G (red spheres) and K310T (pink spheres) from DENV2 (left) and WNV (right), respectively. The ribbon diagram was rendered using PyMOL v. 0.97 (Delano Scientific LLC, San Carlos, CA). **B** (left panel), Effect of mutations in the dissociation constant for mutations S331A (DENV2) and T332A (WNV). **B** (right panel), same effect for mutations K307G (DENV2) and K310T (WNV). The MAbs used for binding isotherms were 9F16 and 3H3 for DENV2 and WNV, respectively. '*' represents a statistically significant decrease in MAb binding affinity relative to the wild type protein.

4.4.2 Gibbs free energy of global unfolding affects the thermodynamic descriptors of ED3 from DENV2

a. Dependence of the residue stability constant on the free energy of unfolding

In Figure 4.4 is shown the changes in the residue stability constant of wild type ED3 as a function of the free energy of unfolding from 4.5 kcal·mol^{-1} to 7.8 kcal·mol^{-1}. Previous studies have shown that the unfolding energy of wild type ED3 from DEN2 is qualitatively higher than that for wild type ED3 from WNV[42]. Therefore we used a value of 5.5 kcal·mol^{-1} for wild type ED3 and 5.0 kcal·mol^{-1} for single site mutations.

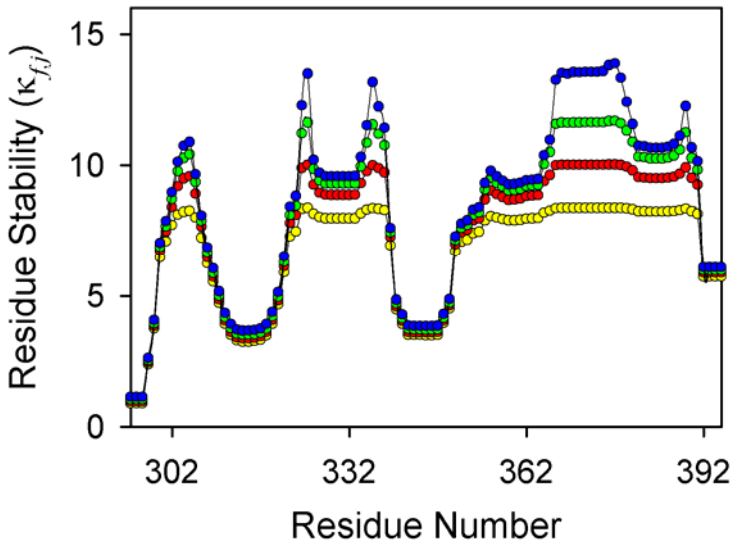

Figure 4.4. Effect of increasing free energy of unfolding in the residue stability constant in wild type ED3 from DENV2. The colors correspond to different free energies of unfolding as follows: yellow (4.5 kcal·mol^{-1}), red (5.5 kcal·mol^{-1}), green (6.5 kcal·mol^{-1}) and blue 7.8 (kcal·mol^{-1}).

b. Dependence of the energetic coupling on the free energy of unfolding

Figure 4.5 shows the dependence of energetic coupling between pairs of residues (or $\Delta\Delta G_{j,k}$) on the stability of unfolding of wild type ED3, in which residues in red color

share a high degree of positive energetic coupling and formed a cooperative core in the protein. When the stability of wild type ED3 is increased from 4.5 kcal·mol^{-1} to 7.8 kcal·mol^{-1}, the degree of cooperativity between residues decreased significantly. We observed that as the global stability increases, the cooperative core reduces significantly not only in magnitude but in the number of residues comprising the cooperative core as well (Figure 4.5D).

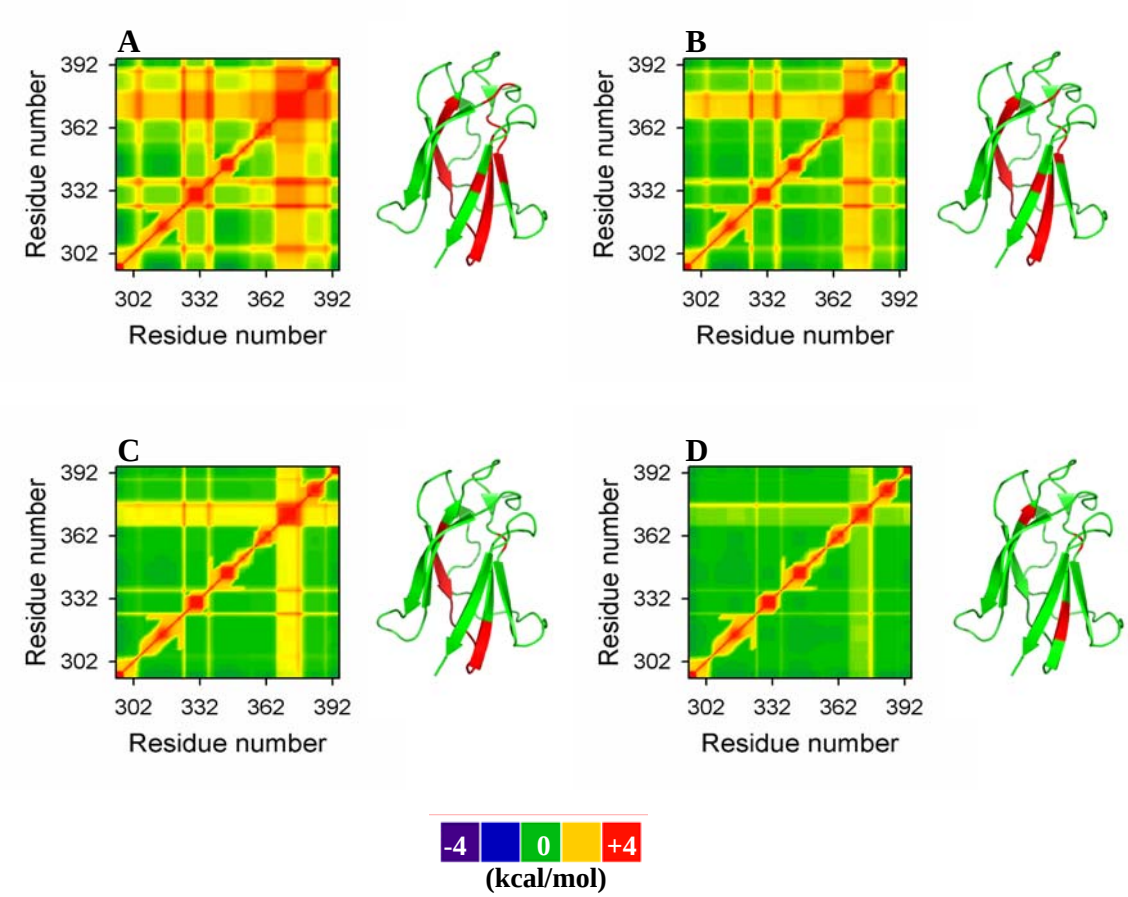

Figure 4.5. Effect of increasing free energy of unfolding in energetic coupling ($\Delta\Delta G_{j,k}$) in wild type ED3 from DENV2. The letter coding corresponds to different global free energies of unfolding: **A** (4.5 kcal·mol^{-1}), **B** (5.5 kcal·mol^{-1}), **C** (6.5 kcal·mol^{-1}) and **D** (7.8 kcal·mol^{-1}). The conformational entropy-scaling factors were 0.892, 0.886, 0.880 and 0.872 kcal·mol^{-1}·K^{-1}, respectively. The ribbon diagram was rendered using PyMOL v. 0.97 (Delano Scientific LLC, San Carlos, CA).

4.4.3 Wild type and mutants of ED3 in DENV2 share similar energetic coupling patterns

We investigated the effect of a panel of mutations on the energetic coupling of wild type ED3. These mutations are located throughout the ED3 structure (Figure 4.11) and correspond to both naturally and non-naturally occurring mutations (Table 4.1). The former corresponded to amino acid sequence differences between Asian genotype viruses of DENV2 (involved in human epidemics) and the sylvatic genotypes (found in primate-mosquito-primate cycles in the jungles of Africa and Asia)[105-107]. The non-naturally occurring substitutions were made based on the amino acid sequence diversity of ED3 derived from 50 DENV2 strains relative to DENV2 prototype strain NGC[104].

Table 4.1. Mutations in ED3 from Dengue virus type 2

Naturally occurring mutations	Non natural mutations
R345K	K305G
I379V	E393G
K393R	D329G
S331A	K388G
K334Q	K307G
P384D/N	E327G

In wild type ED3 as well as in the single mutants, residues sharing high positive coupling were conserved. As shown in Figure 4.6, the energetic coupling map of ED3 from DENV2 for wild type and mutants K388G, I379V and P384D is rendered as color maps (or heat maps), wherein positive, negative and neutral coupling are represented in red, blue and green colors, respectively. The single mutations K388G and P384D have been described as MAb-neutralization resistant mutations, while the mutation I379V behaves like wild type ED3[104]. It is important to note that many of these energetically coupled residues are not contiguous in the primary sequence, suggesting that, as in the

case of ED3 from WNV, there is also a network of interacting residues that is conserved among wild type ED3 and single mutants. Therefore, the fold of ED3 from DENV2 is thermodynamically robust despite of the presence of single mutations.

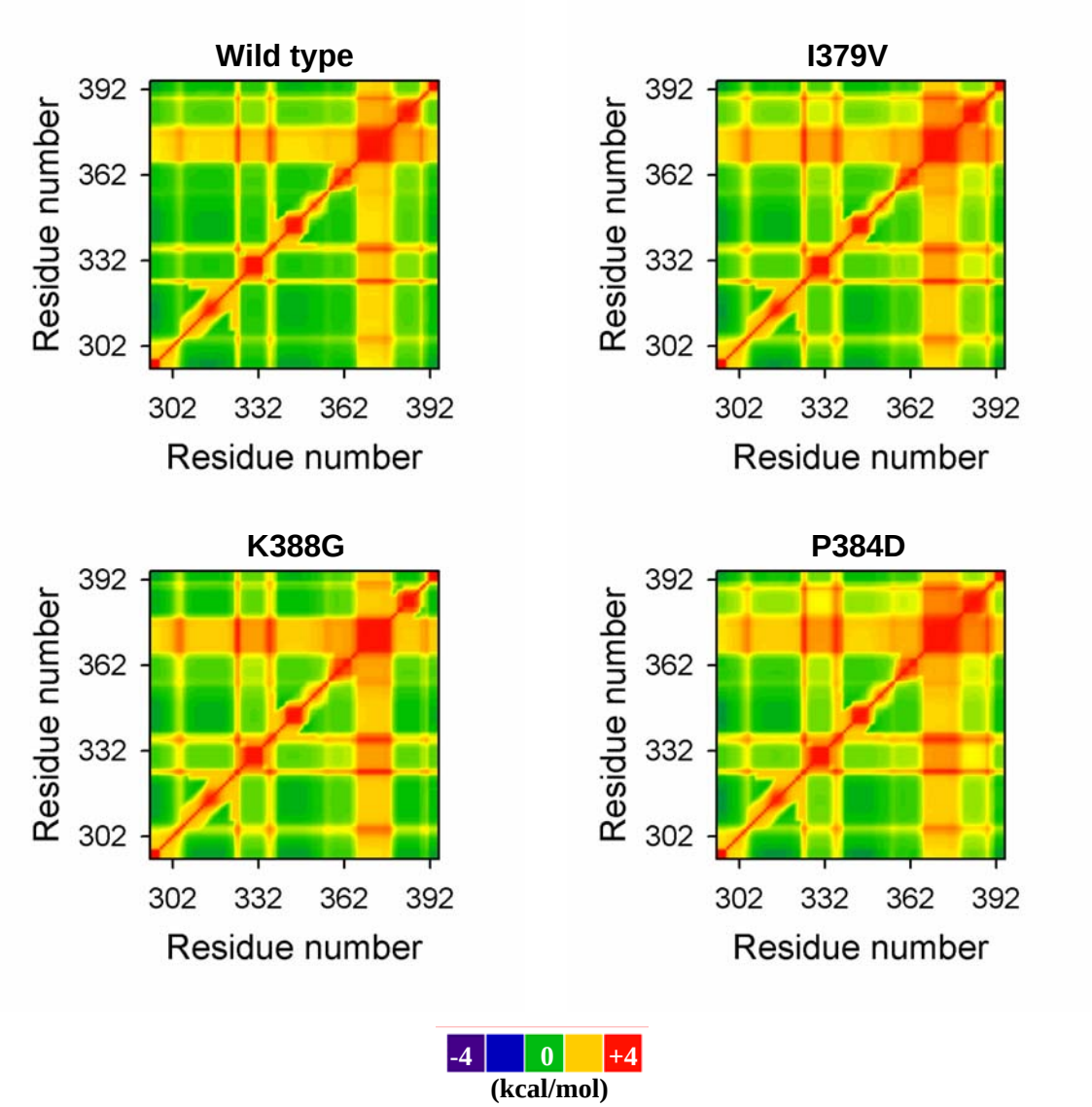

Figure 4.6. The energetic coupling pattern ($\Delta\Delta G_{j,k}$) or mutual perturbation response between pair of residues in ED3 wild type from DEN2 is thermodynamically robust. Relative to wild type ED3, the mutation I379V does not decrease antibody binding, while the mutations K388G and P384D significantly decrease antibody binding affinity by ~4-fold and 25-fold (Table 4.1).

4.5 DISCUSSION

4.5.1 ED3 from Dengue virus type 2 and West Nile virus have unique virus-specific thermodynamic signatures

The ED3 in all flaviviruses have a conserved IgG-like fold, with sequence identities between 40 to 80 %[21,108]. In particular, the ED3 from WNV and DENV2 share very high structural similarity (Figure 4.1) and sequence identity (40 %, Figure 4.2); yet, its antigenic determinants are significantly different, as judged by the differential effect of mutations in MAb binding (Figure 4.3). We found that a given substitution in DENV2 ED3 can impair MAb binding affinity, while a similar substitution in the structurally homologous site in WNV ED3 has no effects (Figure 4.3). In agreement with the functional differences between the ED3 from DENV2 and WNV, we have observed a differential effect in the response of the effect of mutations on the residue stability constant ($\kappa_{f,j}$) and energetic coupling ($MR_{j,k}^{mut-wt}$). These differences, as it will be seen in the next sections, are localized in residues that have been shown to be functionally relevant for the virus.

a. Differential effect of mutations on the residue stability constant

In the analysis of the calculated thermodynamic properties of DEN2 ED3 relative to WNV ED3, we found that there were similarities in the residue stability constant in ED3, an expected result since both proteins are homologous. The residue stability constants of wild type ED3 from DENV2 and WNV are shown in Figure 4.7. The black bars represent the residue stability constant for DENV2 ED3, and the grey bars represent the *negative* of the residue stability constant of WNV ED3. It is evident that the overall profile resembles a "mirror image" between the two ED3s. Namely, almost the same groups of residues share either low or high stability within the protein. The correlation factor between the two stability constant profiles was 0.6. However and not so expected, the effect of structurally homologous mutations showed very distinct patterns, suggesting that the response of the ensemble to perturbations is unique to each flaviviral ED3. This

last statement will be further supported by the response of the energetic perturbations to single mutations.

In Figure 4.8 (upper panel) it is shown the residue stability constants of the MAb-neutralization resistant mutation K305G from DENV2 ED3. Together in the same panel is shown the residue stability constants of the MAb-nonresistant mutation I379V. Remarkably, the only differential effect between the mutations K305G and I379V was the destabilization of residues located in the FG loop (residues 378-388) generated by the K305G mutation (Figures 8, red triangles). Analogously, in the lower panel of Figure 4.8 it is shown the residue stability constants of the MAb-neutralization resistant mutation K307R and the MAb-nonresistant mutation K310T in WNV ED3. The mutation site K307 is structurally homologous to the K305 site in DENV2 ED3. However, the mutation K307R in WNV destabilized the BC loop, corresponding to residues 328-338. No other difference was observed between K307R and K310T (Figure 4.8, lower panel).

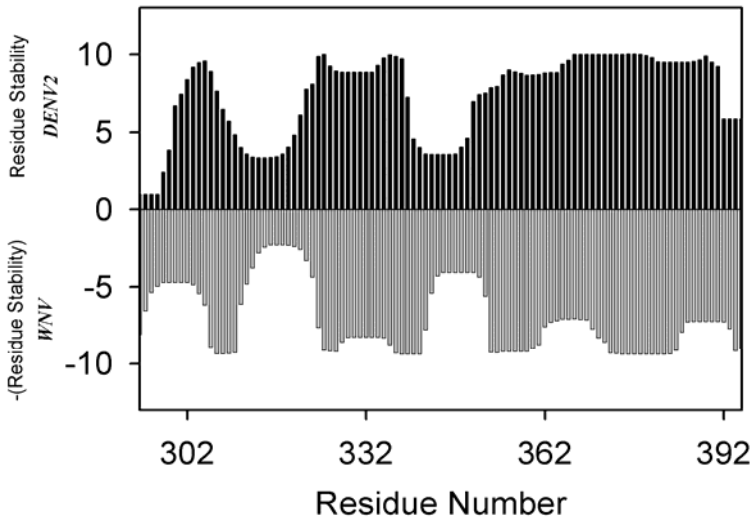

Figure 4.7. Residue stability constants of ED3 form DENV2 (black bars) and WNV (grey bars). The values corresponding to WNV are rendered as the negative of the residue stability constants. The free energy of unfolding and conformational entropy-scaling factor for the calculations for DEN2 were 5.5 kcal·mol^{-1} 0.986 kcal·mol^{-1}·K^{-1}, and for WNV, 4.5 kcal·mol^{-1} 0.955 kcal·mol^{-1}·K^{-1}.

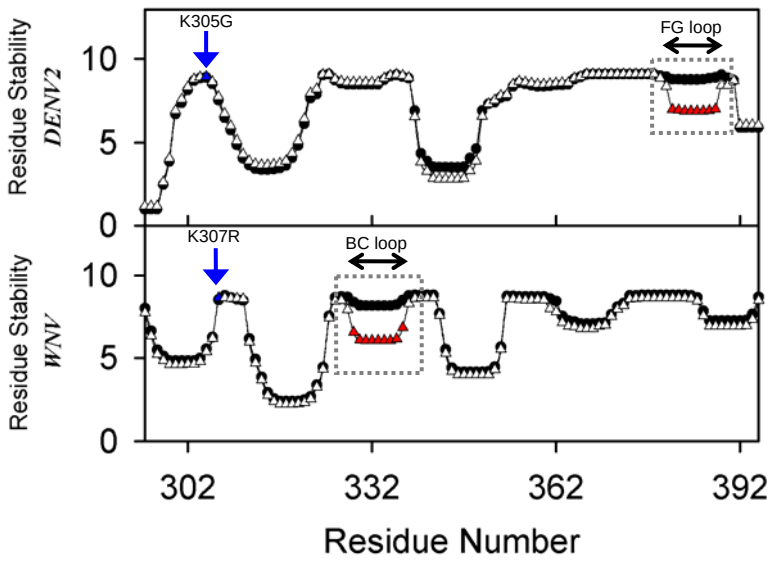

Figure 4.8. Differential effect of MAb-resistant mutations in the residue stability constant between the ED3 form DENV2 (upper panel) and WNV (lower panel). The mutation for DENV2 was K305G and for WNV, K307R (blue arrows). The residue stability constant for K305G and K307R (open triangles) overlap almost perfectly with the residue stability constant of the MAb nonresistant mutation, except in the regions corresponding to the FG loop for DENV2; and for WNV, residues in the BC loop (red triangles). Black circles represent the residue stability constant of the MAb nonresistant mutations I379V and K310T, for DENV2 and WNV, respectively

b. Differential effect of mutations on the energetic coupling

It was shown that the residue stability constant plot of the wild type ED3s from DENV2 and WNV were very similar (Figure 4.7). However, inspection of the energetic coupling patterns for these two proteins showed noticeable differences. The difference resided principally in the 'amount' of residues sharing high positive energetic coupling, being for DENV2 ED3 significantly lower (Figure 4.9A). Nevertheless, the positively coupled residues in DEV2 ED3 also mapped to residues in the core of the protein, as in the case of WNV ED3 (Figure 4.9B).

The difference in energetic coupling between DENV2 ED3 and WNV ED3 suggest that the response of mutations on the energetic coupling will be different too. And in fact, that was the case. Single mutations in ED3 from DENV2 perturbed principally the energetic coupling over residues 378 to 388, or the FG loop (Figure 4.10),

while in the case of WNV ED3, single mutations perturbed residues 328-338 or the BC loop (Figure 3.9, Chapter 3).

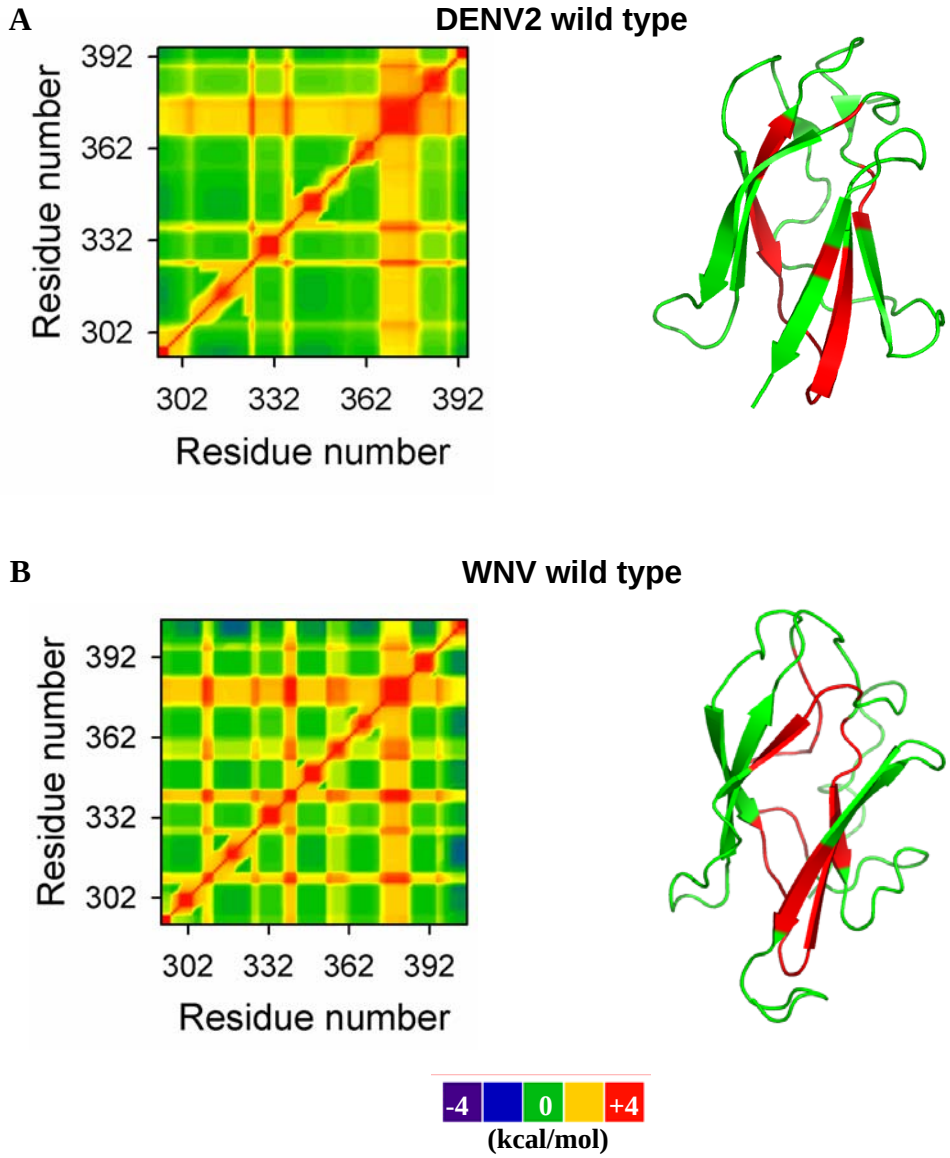

Figure 4.9. **A**, Energetic coupling ($\Delta\Delta G_{j,k}$) in ED3 wild type from DENV2 and **B**, for WNV. Residues in ED3 that shared high positive energetic coupling, forming a cooperativity core, were rendered in red, while residues having neutral coupling in green. The values of global free energy and conformational entropy-scaling factor for the calculation of the residue stability constant for DEN2 were 5.5 kcal·mol^{-1} 0.986 kcal·mol^{-1}·K^{-1} and for WNV were 4.5 kcal·mol^{-1} 0.955 kcal·mol^{-1}·K^{-1}, respectively. The ribbon diagrams were rendered using PyMOL v. 0.97 (Delano Scientific LLC, San Carlos, CA).

In Figure 4.10 it is shown the energetic coupling difference, defined as mutational response or $MR_{j,k}^{mut-wt}$ (equation 3.14, Chapter 3), between wild type ED3 from DENV2 and three single mutants: K388G, I379V and P384D. Remarkably, the effect of these mutations was conserved; namely, it perturbed the same group of residues that mapped to the FG loop, a solvent-exposed loop in the viral particle. While the perturbation by the mutation was conserved position-wise, the magnitude of the effect was very different between the single mutants. In the case of the mutation K388G (Figure 4.10A), the average mutational response in the FG loop was $< MR_{FG\ loop}^{mut-wt} >$ = –0.42 kcal·mol^{-1}. In the case of P384D (Figure 4.10C), the average mutational response in the same region was 0.65 kcal·mol^{-1}. Interestingly, both the K388G and P384D substitutions decrease MAb binding affinity relative to wild type ED3 (Table 4.2). For MAb-nonresistant mutations, there was not a significant effect over residues in the FG loop, as in the case of I379V (Figure 4.10B), with an average mutational response of 0.19 kcal·mol^{-1}.

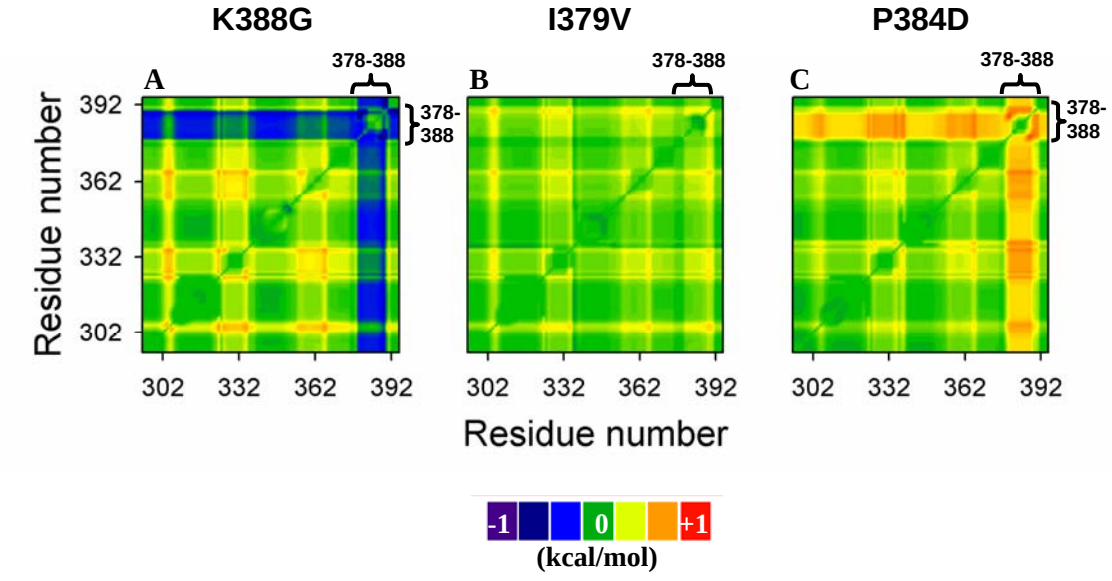

Figure 4.10. Residues perturbed by single mutations are conserved position-wise. Mutations K388G, I379V and P384D perturbed the same group of residues (378-388) by different magnitudes of energetic coupling relative to the wild type ED3 protein, ranging from negative (**A**, K388G), intermediate (**B**, I379V) and positive (**C**, P384D) values. The heat maps were obtained by subtracting the wild type ED3 energetic coupling from the mutant one and it is defined as the mutational response or $< MR_{j,k}^{mut-wt} >$.

The biological relevance of the solvent-exposed FG loop in DENV2 ED3 is that this loop has been proposed as the dominant epitope in this domain by independent groups of investigators[102-104]. Thus, as in the case of WNV ED3, the ensemble-based description of the COREX algorithm captures the biologically relevant regions in ED3 for DENV2.

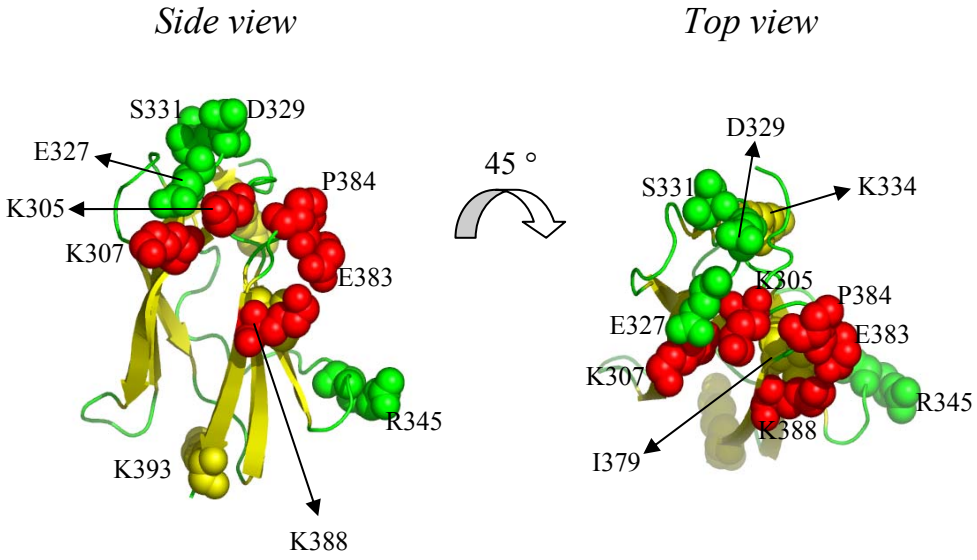

Figure 4.11. Top and side view of mutation sites in DENV2 ED3 (PDB 1TG8). Red spheres show mutation sites that decrease MAb binding affinity by > 30~fold relative to wild type ED3. Mutation sites in green and yellow do not decrease MAb binding affinity as strongly as mutation sites in red spheres. Yellow and green colors represent residues in beta-strand or loops, respectively. The ribbon diagram was rendered using PyMOL v. 0.97 (Delano Scientific LLC, San Carlos, CA).

4.5.2 Continuum-like behavior in the mechanism of MAb-neutralization resistance is a common feature in both the DENV2 and WNV

In the analysis of WNV ED3, we observed that both the MAb-resistant and MAb-nonresistant mutations were energetically coupled to the same region, the BC loop in ED3, but with different magnitudes ranging from negative to positive energetic coupling (Figure 3.9, Chapter 3). However, only those mutations that generated large deviations on the energetic coupling (negative or positive) corresponded to MAb-neutralization

resistant mutants. In the case of DENV2 ED3 we found an analogous behavior between K388G (negative energetic coupling), I379V (intermediate energetic coupling) and P384D (positive energetic coupling) but localized in the FG loop (Figure 4.10). Thus, we can hypothesize that the single mutations best able to perturb the energetic coupling of the FG in ED3 from DENV2 will result in the reduction of MAb binding affinity. Therefore we investigated the effect 13 single mutations (listed in Table 4.1) over the energetic coupling of the FG loop. The substitutions include both naturally and non-naturally occurring mutations (Table 4.1) and are located throughout the ED3 structure (Figure 4.11).

The values of the mutational response of the 13 mutations over the FG loop (defined as $< MR_{FG\ loop}^{mut-wt} >$) was plotted in rank order, from most negative to most positive (Figure 4.12A). By doing the rank-order plot of $< MR_{FG\ loop}^{mut-wt} >$ for all 13 mutations, we observed a sigmoidal curve, wherein the values at both extremes (high negative and high positive coupling) correspond to mutations that decrease MAb binding affinity by ~30-fold or more (Table 4.2). These values deviate from the mean by + 0.3 kcal·mol^{-1} to –0.8 kcal·mol^{-1}. The remaining mutations had a mean value of 0.42 kcal·mol^{-1} and corresponded to mutations that had very similar binding affinities to MAbs relative to wild type ED3 (Table 4.2).

The sigmoidal shape in Figure 4.12A was also observed in the rank-order plot of the mutational response in WNV ED3 in the BC loop or $< MR_{BC\ loop}^{mut-wt} >$ (Figure 3.10, Chapter 3), which conformed a Boltzmann distribution. Similarly, the data is Figure 4.12A was fitted to a Boltzmann distribution (solid line) using equation 3.15 in Chapter 3. The Boltzmann distribution of the data suggests that the ensemble of ED3 from DENV2 includes conformational states in which the ED3 adopts a *MAb-non neutralization resistant conformation,* characterized by intermediate values of energetic coupling in the FG loop; and a *MAb-neutralization resistant conformation* characterized by either high negative or high positive energetic coupling in the same region. These results are consistent with the observed continuum-like behavior in MAb-neutralization resistance for WNV ED3. Thus, in DEV2, MAb neutralization resistance also behaves as

a continuum phenomenon and the decrease in MAb binding depends on the relative ability of a mutation to modulate the energetic coupling in the FG loop (Table 4.2).

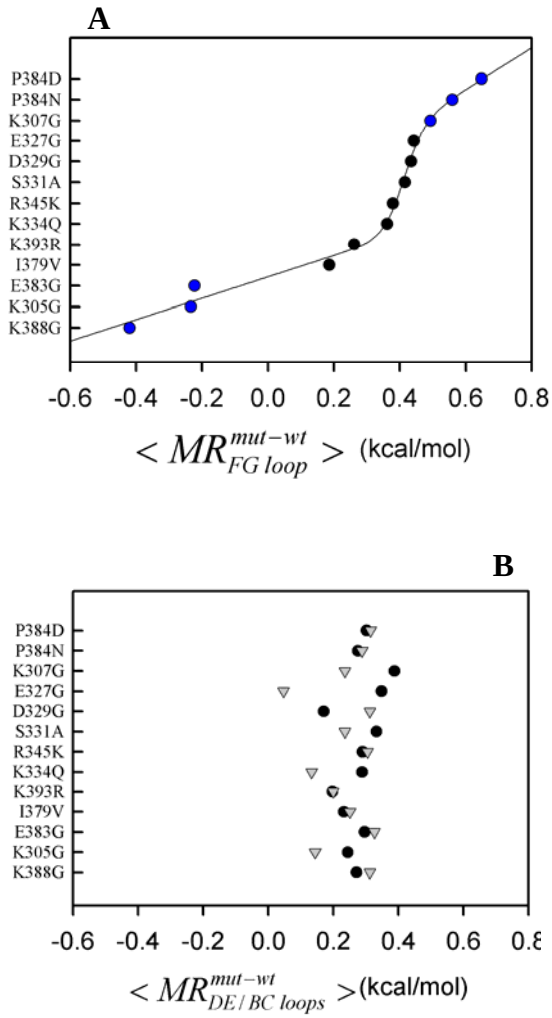

Figure 4.12. **A**, Boltzmann distribution of the rank-order plot of the energetic coupling perturbation by single mutations in the FG loop of ED3 from DENV2. The solid line represents the Boltzmann fit (equation 3.15, Chapter 3) and was obtained by assuming identical distances in the y-axis between each mutant. The midpoint value of the transition was 0.42 kcal·mol^{-1}, and the maximum deviances correspond to + 0.3 kcal·mol^{-1} and –0.8 kcal·mol^{-1}. Blue spheres represent MAb-resistant mutations, while black spheres represent MAb-nonresistant mutations. **B**, Plot of the energetic coupling perturbations in the DE (spheres) and BC (triangles) loops of ED3 from DENV2. The rank order was kept as in Figure 12A. All the values for the mutants oscillate near the mean value of 0.3 kcal·mol^{-1}.

Table 4.2. Correlation between Mutational Response in the FG loop ($< MR_{FG\ loop}^{mut-wt} >$) and MAb binding affinity (Kd_{app}, nM) in DENV2 ED3

	Mutational Response in BC loop ($< MR_{FG\ loop}^{mut-wt} >$)[a]	MAb binding affinity[b] (Kd_{app}, nM)
Wild type	0.00	**3.81 ± 0.5**
K388G	**-0.42**	**11.7 ± 0.8**
K305G	**-0.23**	**108.0 ± 24**
E383G	**-0.22**	**80.0 ± 1.7**
I379V	0.19	4.67 ± 0.7
K393R	0.26	5.42 ± 0.9
K334Q	0.36	3.68 ± 0.2
R345K	0.38	4.34 ± 0.3
S331A	0.42	4.38 ± 0.2
D329G	0.43	14.2 ± 0.8
E327G	0.44	13.3 ± 1.2
K307G	0.49	13.8 ± 1.7
P384N	**0.56**	**> 80[c]**
P384D	**0.65**	**> 100[c]**

[a] Values in kcal·mol^{-1}. [b] Binding affinity values for MAb 9F16 were kindly provided by Gregory Gromowski and Dr. Alan Barrett[104]. [c] For P384N/D the data was obtained from Hiramatsu et al., 1996[102].

While the qualitative behavior between the mutational response of the energetic coupling in ED3 from DENV2 and WNV is very similar, the values of the average mutational response on the energetic coupling are different (Table 4.3). This observation is consistent with the fact that the ED3 in DENV2 and WNV are homologous proteins, but they maintain virus-specific properties. For example, we observed differences in the

fitted parameters in the Boltzmann distribution function, which included the mean average mutational response (= 0.42 kcal·mol^{-1} for DENV2 ED3 in the FG loop, vs. 0.34 kcal·mol^{-1} for WNV ED3 in the BC loop) and the constant '*C*', which is a dimensionless parameter that describes the equilibrium cooperative factor[97,98] (= to *20* for DENV2 ED3 in the FG loop, vs. *24.6* for WNV ED3 in the BC loop). The mean mutational response corresponds to the midpoint in the transition between high negative and high positive values of energetic coupling. Table 4.3 summarizes the differences between the ED3 from DENV2 and WNV in the mutational response of the energetic coupling of loops BC, DE and FG, which delineate a solvent-exposed patch in ED3 accessible to MAbs (Figure 4.1).

Table 4.3. Differences between the ED3 from DENV2 and WNV in the mutational response on the energetic coupling in the BC, DE and FG loops

	Dengue virus type 2			West Nile virus		
Mutational Response[a]	BC loop	DE loop	FG loop[b]	BC loop[b]	DE loop	FG loop
Mean $< MR_{j,k}^{mut-wt} >$	0.24 ± 0.09	0.28 ± 0.06	**0.42 ± 0.02**	**0.34 ± 0.04**	0.19 ± 0.06	0.23 ± 0.05
Maximum $< MR_{j,k}^{mut-wt} >$	0.33 (E383G)	0.39 (K307G)	**0.65 (P384D)**	**0.85 (T332K)**	0.26 (V338I)	0.31 (T332K)
Minimum $< MR_{j,k}^{mut-wt} >$	0.05 (E327G)	0.17 (D329G)	**-0.42 (K388G)**	**-0.20 (K307R)**	0.10 (T332A)	0.12 (K307R)

[a] Values in kcal·mol^{-1}. [b] Transition midpoint in Figure 4.12A and Figure 3.10 (Chapter 3). Mutation in parenthesis that generated the min/max observed $< MR_{j,k}^{mut-wt} >$.

It is evident from Table 4.3 that in the case of DENV2 ED3, the loop that was the most susceptible to perturbations by mutations is the FG loop compared to the DE and BC loops. This mutational susceptibility was reflected not only by the largest deviations of energetic coupling from a mean value (= 0.42 ± 0.02 kcal·mol^{-1}), but also the mean value itself is ~1.5-fold higher that the mean value of for the DE and BC loops. This behavior is illustrated in the Figure 4.12B, in which the mutational response on the DE and BC loops (for MAb-resistant and non-resistant mutations) is plotted in the same order as in Figure 4.12A. However, no significant deviations were observed for any mutant. The mean mutational response for the DE and FG loops were 0.24 kcal·mol^{-1} and 0.28 kcal·mol^{-1}, respectively. In the case of WNV ED3, the BC loop had the highest mutational response (or mutational susceptibility) compared to the DE and FG loops. And similarly, the BC loop showed the largest deviations from the mean value (0.34 kcal·mol^{-1}) upon mutation; and its mean energetic coupling is ~1.5-fold higher than those for the DE and FG loops. It is interesting to note that almost all the mutations (except one, V338I in WNV ED3) that had the most significant perturbations on the energetic coupling in any loop in ED3 for both DENV and WNV corresponded to mutations that reduce MAb binding affinity by ~4-100-fold. This result suggests that the effect of the MAb-neutralization resistant mutations on the energetic coupling of ED3 is more pronounced, but not limited, to one particular region of the domain.

4.5.3 Modulation of energetic coupling depends on the stability of virus-specific residues

Analysis of the effect of MAb-neutralization resistant and non-resistant mutations on the residue stability constant of the ED3 provides a physical understanding at the molecular level of the modulation of the energetic coupling from high negative to high positive values. In Figure 4.8 we observed that the MAb-neutralization resistant mutation K305G in DENV2 ED3 destabilized residues only in the FG loop compared to a non-resistant mutation, while the rest of the residues in ED3 remained essentially unaffected.

Subsequent calculations showed that the mutation K305G generated high negative coupling in the same group of residues (Table 4.2). In Figure 4.13A, a similar pattern is observed in the residue stability constants of wild type ED3 from DENV2 (green spheres) and two MAb-neutralization resistant mutants: K388G and P384D (blue and red spheres, respectively). Both mutations had almost the same destabilizing effect, relative to wild type ED3, except in the residue stability constants of residues 378 to 388, or the FG loop.

In the case of the mutation K388G the residue stability constant in the FG loop was significantly lower compared to wild type ED3, and also showed high negative energetic coupling. The correlation *low stability* ↔ *negative-energetic-coupling* in the FG loop was also observed for the mutants K305G and E383G (Table 4.4). The same correlation but with different magnitudes was observed for single mutations that generated high positive coupling (e.g. P384D/N); namely, a correlation of *high-stability* ↔ *positive-energetic-coupling*. Though, this correlation was not observed when the residue stability constant of the FG loop of the mutant was directly compared the values of wild type ED3. Instead, the higher stability was observed in the ratio *FG-loop-stability* to *maximum-residue-stability*, which was higher for mutants that generated high positive energetic coupling compared to wild type ED3. These results indicate that the correlation *high-stability* ↔ *positive-energetic-coupling* is not the consequence of more stable residues in the FG loop, but by a decrease in the residue stability constant in the rest of the protein (Figure 4.13B). The value of the ratio *FG-loop-stability* to *maximum-residue-stability* of MAb-non resistant mutations (e.g. I379V or K393R) was very similar to the value for wild type ED3 (Table 4.4), in agreement with the minimal mutational response of the energetic coupling in the FG loop (Table 4.3).

This correlation (*low/ high-stability* ↔ *negative/ positive-energetic-coupling*) is analogous to the findings in WNV ED3, where mutations that showed high negative coupling in the BC loop, also showed the lowest stabilities in the same groups of residues; whereas mutations that generated high positive coupling showed the highest residue stability (normalized to the maximum residue stability). Therefore, the modulation of the energetic coupling, which correlates with the degree of resistance

conveyed by the mutation, depends on the effect of the mutation over the relative stability of virus-specific residues in ED3; namely, the FG loop in DENV2 the BC loop in WNV.

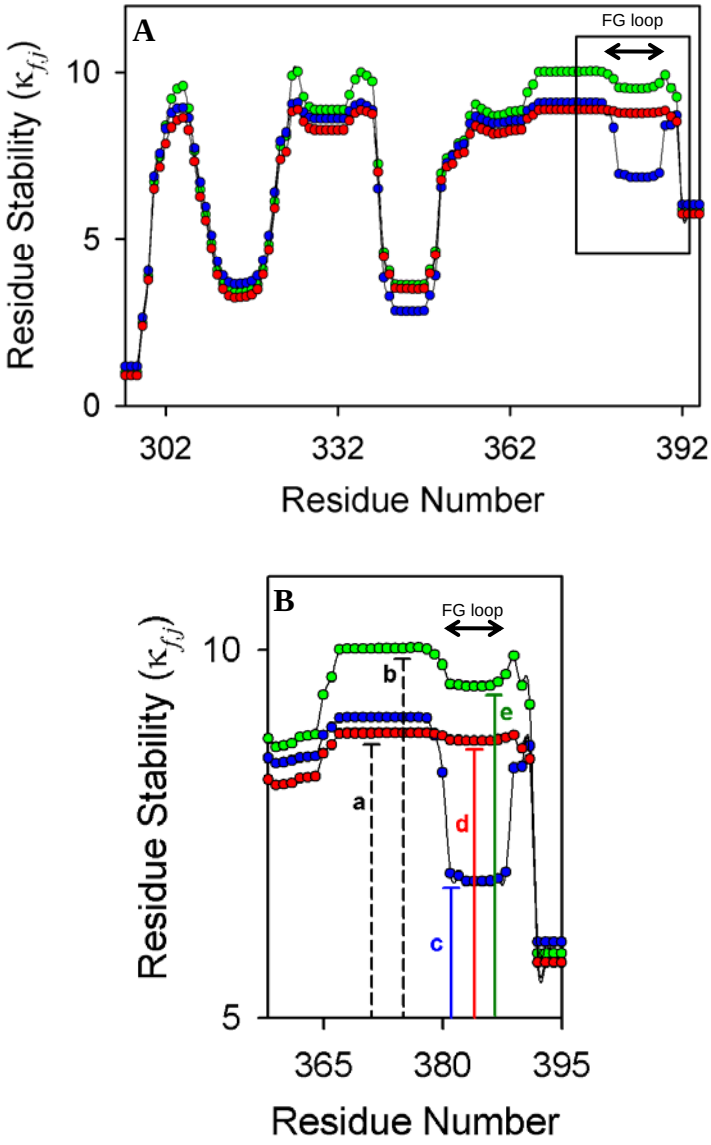

Figure 4.13. **A**, Residue stability plot for the mutants K388G (blue) and P384D (red) together with wild type ED3 from DENV2 (green). Box encloses the residues mostly affected by the presence of the mutations, relative to wild type ED3, and corresponds to residues 378-388 or FG loop. **B**, Relative residue stability of the FG loop in ED3 from DENV2. Labels 'a' and 'b' correspond to the maximum stability in the protein for the mutants (K388G and P384D) and wild type ED3, respectively. Labels 'c' (blue), 'd' (red) and 'e' (green) correspond to the stability of the residues in the FG loop for K388G, P384D and wild type ED3, respectively.

Table 4.4. Effect of mutations in the Stability of the FG loop in DENV2 ED3[a]

	Global Stability[b]	FG loop (Local) Stability[c]	*Ratio Local/Global Stability*	Maximum Residue Stability[d]	*Ratio Local/Max Stability*	Mutational Response $< MR_{BC\ loop}^{mut-wt} >$	Effect on FG loop stability relative to wild type ED3
Wild type	7.39	9.64	1.31	10.02	0.96	---	---
K388G	6.84	7.41	1.08	9.08	**0.81**	**-0.42**	*<<< stability*
K305G	6.95	8.35	1.20	9.07	**0.92**	**-0.23**	*<< stability*
E383G	6.86	8.14	1.19	8.82	**0.92**	**-0.22**	*<< stability*
I379V	6.98	8.84	1.27	9.08	0.97	0.19	*Behaves like wild type*
K393R	6.89	8.95	1.30	9.16	0.98	0.26	*Behaves like wild type*
K334Q	6.82	8.87	1.30	9.03	0.98	0.36	*> stability*
R345K	6.90	8.71	1.26	8.86	0.98	0.38	*> stability*
S331A	6.85	8.69	1.27	8.84	0.98	0.42	*> stability*
D329G	6.85	8.75	1.28	8.93	0.98	0.43	*> stability*
E327G	6.82	8.72	1.28	8.85	0.98	0.44	*> stability*
K307G	6.82	8.76	1.29	8.88	**0.99**	**0.49**	*>> stability*
P384N/D	6.84	8.80	1.29	8.87	**0.99**	**0.61**	*>> stability*

[a] All stability units are in kcal·mol^{-1}. [b] The Global Stability was calculated by averaging the Residue Stability Constants of each residue in ED3. [c] FG loop stability was calculated by averaging the Residue Stability Constants of residues 378-388. [d] The Maximum Residue Stability found in the ED3 for wild type and single mutants.

4.6 CONCLUSIONS

Monoclonal antibodies (MAbs) against the envelope (E) protein are capable of neutralizing flaviviruses *in vitro*, including DENV2 and WNV[14-16,100,101]. In particular, MAbs recognizing the domain 3 (ED3) of the envelope have been proven to be of high neutralizing potency[14-16,104]. Thus, the definition of dominant neutralizing epitopes against ED3 is of great interest for the development of an effective vaccine response or for the development of MAbs with therapeutic potential[3]. In the effort of dissecting dominant neutralizing epitopes in ED3 in flaviviruses, the dogmatic approach has been the study of the effect of single site mutations in MAb binding. Mutations that significantly decrease MAb binding affinity, relative to the wild type ED3, are defined as MAb-neutralization resistance mutations.

In the previous chapter we quantitatively described the mechanism of MAb-neutralization resistance exerted by different single mutations WNV in a single function of the mutational response of the energetic coupling that followed a Boltzmann distribution. The Boltzmann distribution suggested that there was an equilibrium process between different conformational states in the ensemble that included both the MAb-neutralization and non-neutralization resistant conformations. Within this ensemble of conformations, the MAb-neutralization resistant mutations were best able to shift this equilibrium towards conformations characterized with either high negative or high positive energetic coupling values in residues of the BC loop, the dominant neutralizing epitope in WNV ED3.

In order to test the generality of the results and conclusions derived from WNV ED3, we analyzed the mutational response on the energetic coupling of ED3 from DENV2. We found parallel results between DENV2 and WNV. Namely, a single Boltzmann distribution function was able to describe the mutational response of both the MAb-neutralization and non-neutralization resistant mutants. However, two main differences between DENV2 and WNV were observed:

1. The magnitude of perturbations of the energetic coupling by the single mutations was higher in DENV2 ED3 compared to WNV ED3.

2. The mutational response in DENV2 ED3 was observed in the FG loop and not in the BC loop as in the case of WNV. Remarkably, the FG loop in DENV2 has been suggested as the dominant neutralizing epitope in ED3.

The parallel results between DEVN2 ED3 and WNV ED3 support the generality of the continuum-like mechanism in MAb-neutralization resistance. And while there are differences in the relative magnitudes of effects of mutations and the locations of these perturbations between DENV2 and WNV; these differences correlate with the virus-specific antigenic determinants for each virus. Therefore, we can hypothesize that the uniqueness in the antigenic determinants in viral proteins are encoded in the thermodynamic signatures of each protein (e.g. depend on the localization of the response of MAb-neutralization resistant mutations on the energetic coupling of the viral protein) despite of the high structural and high sequence similarities that they may share, as in the case of the ED3 from DENV2 and WNV, and in general in Flaviviruses.

CHAPTER 5

DISCUSSION:

AN EXPANDED *MUTATIONAL SPACE* OF FLAVIVIRUSES TO EVADE ANTIBODY-MEDIATED NEUTRALIZATION

A viral strategy to allow survival and establish infection, in response to the selection pressure of the cellular and humoral immune system of the host, is the induction of mutations in viral epitopes of antigenic proteins (facilitated, in particular, by the low-fidelity RNA-dependent RNA polymerases of flaviviruses)[3,12]. This phenomenon, known as antigenic variation, has been reported in the envelope protein domain 3 (ED3) in flaviviruses as relevant to monoclonal antibody (MAb) evasion[12]. ED3 in flaviviruses is believed to be the putative-receptor binding domain[17] and is the major target of neutralizing MAbs. In West Nile virus (WNV), mutations that enable escape from recognition from neutralizing MAbs map to residues in ED3[14-16]. Similarly, other flaviviruses (dengue, Japanese encephalitis, and tick-borne encephalitis) efficiently escape antibody neutralization with mutations that map to ED3 as well[20,23,29,30,104,108,109].

During the antigenic variation process *in vivo*, the generation of random mutations throughout the flaviviral genome will result into MAb-neutralization resistant variants[3,12], which will become dominant if the mutations do not compromise some stage of the virus lifecycle. For example, the presence of mutations in ED3 that enables MAb evasion without the destruction of structural motifs involved in receptor-binding function or viral particle assembly (Zhang et al., *submitted*)[96]. These observations are in agreement with our solution biophysical studies that showed that single site mutations in ED3 from WNV

preserved the integrity of the fold and global solution structure as in the wild type protein (Figure 2.17, Chapter 2).

It is commonly understood that mutations that impair MAb-neutralization or decrease MAb binding affinity are located at sites that are part of the viral epitope. However, this interpretation is mostly based on MAb-antigen reactions; that is, there is no direct supporting structural evidence such as structural information of the antigen-antibody complex[110]. In WNV, as well as in other flaviviruses[16,28-30,38-40,111] and non-related viruses[112,113], the presence of naturally occurring mutations in regions that are not accessible to the solvent or to other macromolecules (e.g. MAbs) suggest that structural elements other than the epitope can play a role in MAb-neutralization resistance via long-range communication. We tested this hypothesis by investigating the effect of single site mutations in ED3 from WNV on the physicochemical properties of distantly positioned structural elements (e.g. the single tryptophan residue is > 20 away Å from the mutation sites). We observed that while the global solution structure was preserved between all single mutations, the dynamical properties and charge distribution of the single tryptophan residue in ED3 were being perturbed (Figures 2.12 and 2.14, Chapter 2). These results provided first evidence of long-range communication in ED3 from WNV.

We further investigated other long-range communication pathways between distant regions by determining the correlated fluctuations of distant residues in ED3 from WNV. To do so, we employed a previously validated ensemble-based description of the equilibrium, the COREX algorithm[91], which models protein fluctuations as local simultaneous unfolding reactions. Therefore, this approach is pathway-independent; namely, there is not a direct mechanical pathway that allows propagation of perturbations[95,96]. To investigate the long-range effects of naturally occurring mutations in ED3 from WNV, we determined networks of energetically coupled residues, and defined the perturbation of mutations over this network as the mutational response of the energetic coupling or $MR_{j,k}^{mut-wt}$ (equation 3.14, Chapter 3). The above mentioned experimental results validated the calculated long-range effects over the single tryptophan

residue due to the presence of single site mutations, positioned > 20 Å away or more (Figure 3.6, Chapter 3).

Upon investigation of the mutational response of the energetic coupling in WNV ED3, we found that the solvent-exposed BC loop was the region in ED3 most susceptible to perturbations caused by mutations of sites located throughout the protein (Figure 3.1, Chapter 3). However, MAb-neutralization resistant mutants were the most efficient ones in perturbing this region, by generating either high negative or high positive negative coupling (Figure 3.9, Chapter 3). Moreover, the rank-order plot of the mutational response followed a Boltzmann distribution function, which suggests that the ability to elicit MAb-neutralization resistance is a continuum. It is simply a manifestation of an equilibrium process between different populations of conformational states in ED3 that include both the *MAb-neutralization resistant* and *non-neutralization conformations*.

To provide an understanding of the physical nature of the mutations that generated either high negative or positive energetic coupling, we determined the effect of mutations on the residue stability constants of ED3 (equation 3.5, Chapter 3), which provides a metric of stability at the residue level. We observed that the MAb-neutralization resistant mutations affected mainly the residue stability of the BC loop relative to wild type ED3, and established a quantitative correlation of *low and high stability to low and high energetic coupling*, respectively, in that region. This correlation for WNV ED3 is shown in Figure 5.1, in which the rank-order plot of the mutational response of the energetic coupling in the BC loop is plotted together with the corresponding values of the residue stability in the same region for each mutant. It is obvious from Figure 5.1A that mutants with low stability (blue spheres) correlate with negative energetic coupling (blue squares), mutants with intermediate stability (green spheres) correlated with low positive coupling (green squares), and the mutants with the highest stabilities (red spheres) correlated with values of high positive energetic coupling (red squares). The differences in the local stability of the BC loop in ED3 indicates that this region in MAb-neutralization resistant mutants is found principally in two populations: 1) A less structured (or 'unfolded-like') conformation for mutants with *low*

stability ↔ *low energetic coupling* without jeopardizing the integrity of the fold of ED3 required for the viability of the virus; and 2), A 'rigid-like' conformation for mutants with *high stability* ↔ *high energetic coupling* because their local stability in the BC loop is as high as the stability of residues found in the core of the protein.

An analogous correlation of *low/ high stability* ↔ *low/ high energetic coupling* in the analysis of single mutants in dengue virus type 2 (DENV2) ED3 was obtained (Figure 5.2). Despite of that the ED3 in DENV2 and WNV are homologous proteins, as judged by the high structural similarity and high sequence similarity, the correlation *low/ high stability* ↔ *low/ high energetic coupling* in DENV2 ED3 was localized to the FG loop, another solvent-exposed loop in the viral particle (Figure 4.1, Chapter 4).

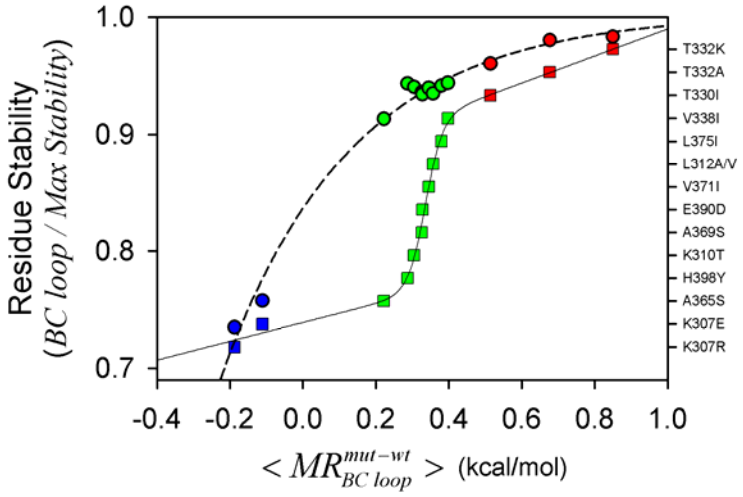

Figure 5.1. The Boltzmann distribution function of the rank-order plot of the mutational response in the BC loop (squares, solid d line) correlates with the rank order of the residue stability in the same region (spheres, dashed line) in ED3 from WNV. The residue stability of the BC loop was normalized to the maximum residue stability found in the protein. Symbols in blue, green and red indicate low, intermediate and high residue stability (or energetic coupling), respectively.

Thus, the equilibrium process between populations of different conformational states in ED3, represented by the Boltzmann distribution function of the energetic

coupling, was able to capture (implicitly) the physical nature (changes in stability) of the difference between *MAb-neutralization resistant* and *MAb non-resistant conformations* in two parallel, but different, viral protein antigens.

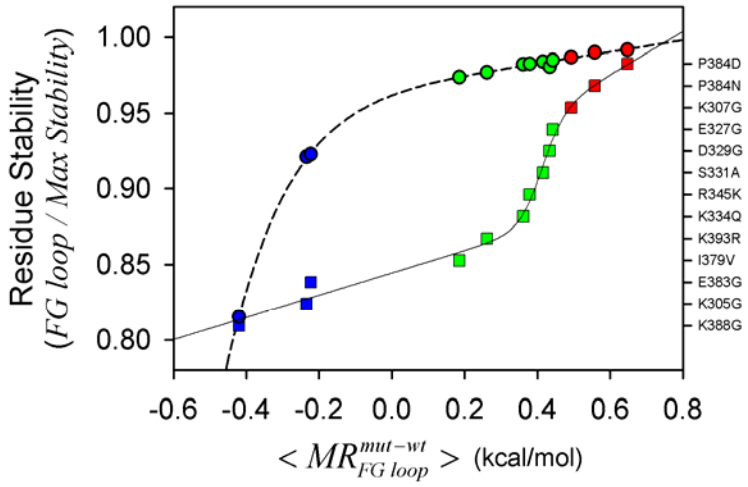

Figure 5.2. The Boltzmann distribution function of the rank-order plot of the mutational response in the FG loop (squares, solid d line) correlates with the rank order of the residue stability in the same region (spheres, dashed line) in ED3 from DENV2. The residue stability of the FG loop was normalized to the maximum residue stability found in the protein. Symbols in blue, green and red indicate low, intermediate and high residue stability (or energetic coupling), respectively.

The sharp equilibrium in the transition of the Boltzmann distribution between mutations with high negative and high positive energetic coupling values (Figures 5.1 and 5.2) indicates that ED3 is thermodynamically robust, but capable to respond to perturbations in the environment. For example, in the context of the virus lifecycle, a perturbation in the environment is equivalent to the selective pressure of neutralizing MAbs, and the ability of response of the virus by mutations allows the evasion of MAbs. For the virus to be able to respond against neutralizing MAbs, the equilibrium of ED3 must be located in the transition region of the Boltzmann distribution (with intermediate

energetic coupling). If the equilibrium of ED3 is located too far below or above the transition region (e.g. high positive or high negative energetic coupling), the response of the system upon a perturbation (by a mutation) will result in no significant changes of the populations towards conformations that are MAb-resistant like. Similarly, if the equilibrium transition were too broad, then the redistribution of the population of conformations towards a *MAb-neutralization resistant conformation* by mutations would be insignificant, and as a result the virus would be efficiently neutralized by MAbs. This behavior is illustrated in Figure 5.3A for WNV ED3 in which the equilibrium description of the rank-order plot of the energetic coupling represents as a phenomenological *observable* (<*obs*>), and the response of the equilibrium between high negative and high positive coupling, defined as $\frac{\partial < obs >}{\partial < MR_{BC\ loop}^{mut-wt} >}$, conforms to a sharp peak in the transition region (dashed line).

The sharpness of the peak indicates that in the transition region, the ED3 is able to respond very efficiently to changes in the environment. In terms of the biology of the virus, by having the wild type ED3 (or non-resistant mutants) positioned in the transition region, the virus will respond efficiently to neutralizing MAbs by decreasing MAb binding affinity by mutations energetically coupled to BC loop. Indeed, we observed a qualitative correlation between the response of the equilibrium of the energetic coupling and MAb binding affinity (or $\frac{\partial < obs >}{\partial < MR_{BC\ loop}^{mut-wt} >} \propto Log[Kd_{app}]$) of resistant and non-resistant mutants (Figure 5.3B).

In the case of DENV ED3, we observed that the response of the equilibrium in the FG loop ($\frac{\partial < obs >}{\partial < MR_{FG\ loop}^{mut-wt} >}$) between high negative and high positive coupling also conformed to a sharp peak in the transition region (Figure 5.4, dashed line), but in comparison to WNV ED3, the curve was shifted towards higher values of positive energetic coupling (Figure 5.3).

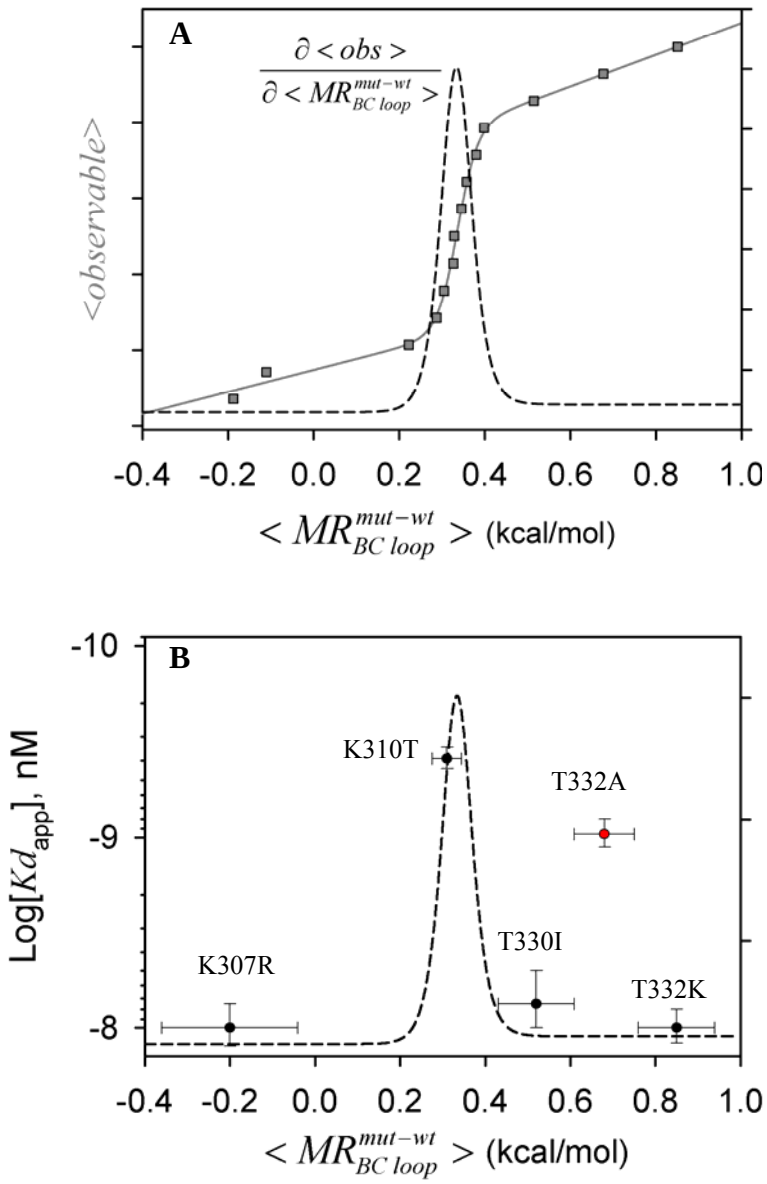

Figure 5.3. **A**, Response of the equilibrium (dashed line) on changes in energetic coupling of the BC loop due to mutations in ED3 from WNV. **B**, Correlation between the response of the equilibrium (dashed line) and decrease in MAb binding affinity. Error bars corresponds to deviation in the estimation Kd_{app} (vertical bars) and deviations in the mean of $< MR_{BC\ loop}^{mut-wt} >$ for that particular mutant (horizontal bars). The dissociation constants were obtained from Figure 21, Chapter 2.

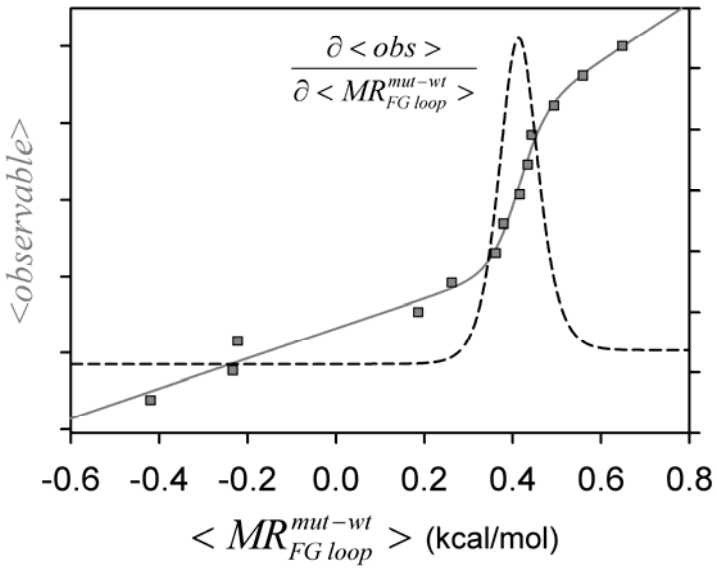

Figure 5.4. Response of the equilibrium (dashed line) on changes in energetic coupling of the FG loop due to mutations in ED3 from DENV2.

When the response of the equilibrium was correlated with the Kd_{app} of the single mutants in DENV2 ED3 (Figure 5.5), we observed, as in the case of WNV ED3, that the values of Kd_{app} delineated the bell-shape curve of the equilibrium response. However, outliers from the equilibrium response plot were observed for the MAb non-resistant mutants I379V and K393R (green arrows), and the MAb-resistant mutant K388G (blue sphere). By displacing the midpoint of equilibrium response to lower energetic coupling values (Figure 5.4, grey dashed curve), the outlier mutations I379V and K393R fall close to the peak of equilibrium response plot. This could suggest that the definition of the midpoint of the equilibrium transition in the rank-order plot of the mutational response of the energetic coupling was not accurate and probably the analyses of more mutants are required. In the case of WNV ED3, the mutant T332A also was an outlier in the equilibrium response plot (Figure 5.3B, red sphere).

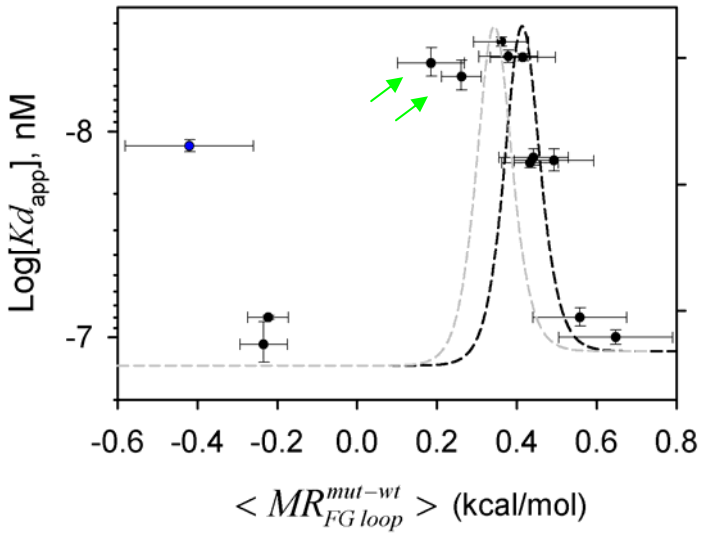

Figure 5.5. Correlation between the response of the equilibrium (black dashed line) and decrease in MAb binding affinity in DENV2 ED3. Vertical and horizontal error bars corresponds to deviations in Kd_{app} and deviations in the mean of $< MR^{mut-wt}_{FG\ loop} >$ for that particular mutant, respectively. For comparison, the response of the equilibrium in WNV ED3 was also included (grey dashed line). Dissociation constant were provided by Greg Gromowski[104].

Despite the small difference in the magnitude of the peak position (= midpoint of the equilibrium transition) and the localization of the energetic coupling perturbation between the ED3 from DENV2 and WNV, the parallel qualitative behavior of the mutational response suggests a common basic mechanism in MAb-neutralization resistance for both systems. Namely, the positioning of wild type ED3 or Mab non-resistant mutants near the transition region of the Boltzmann distribution of the energetic coupling allows the virus to counteract efficiently the neutralizing activity of MAbs by decreasing or increasing the local stability of the BC and FG loops (for WNV and DENV2, respectively). This decrease/ increase in stability in the loops BC or FG results in the same outcome: reduction of MAb binding affinity to the viral epitope. Thus, a quantitative correlation between the mutational response of the energetic coupling ($MR^{mut-wt}_{j,k}$) and the MAb-binding affinity (Kd_{app}) was established for the ED3 of WNV

and DENV2 (Figure 5.6). Specifically, mutations that decrease MAb binding affinity (with Kd_{app} ~4-100-fold > than for wild type ED3), either in DENV2 or WNV, fell into a negative or positive range of energetic coupling (Figure 5.6, blue and red shades, respectively). In the case of mutations with Kd_{app} similar to those for wild type ED3 fell into intermediate values of energetic coupling (Figure 5.6, green shade).

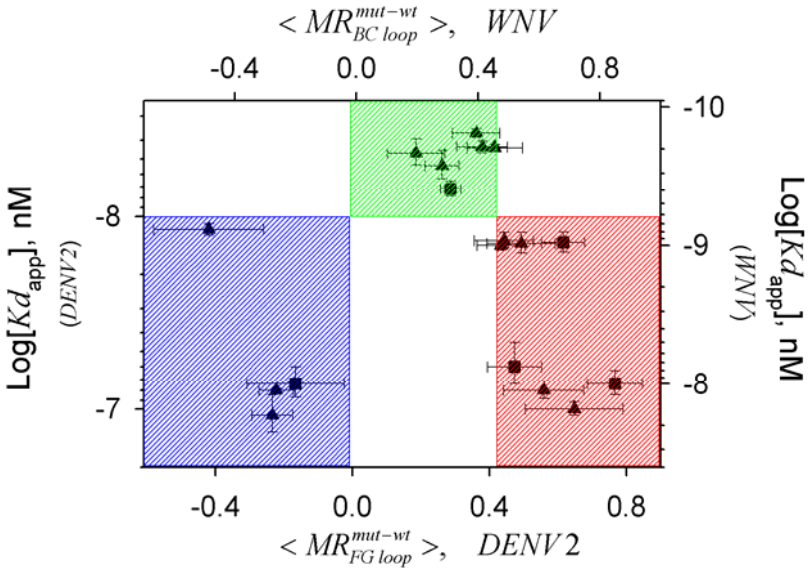

Figure 5.6. Correlation between energetic coupling perturbation and MAb binding affinity in ED3 from WNV (squares) and DENV2 (triangles). Blue, red and green shades indicate negative, positive and intermediate coupling.

The distribution of correlations in Figure 5.6 illustrates an 'energy-landscape map' allows us to establish a simple fundamental strategy for the mechanism of MAb-neutralization resistance based on where the calculated energetic coupling of the mutation falls: mutations that generate high negative or high positive energetic coupling will result, with higher probability, in reduction of MAb binding affinity; and mutations that do not shift the energetic coupling from an equilibrium point (= transition midpoint) will not result in decrease in MAb binding affinity. Thus, the regions in Figure 5.6 with white

background can be considered as not energetically allowed. Noteworthy, the energy-landscape map in Figure 5.6 is obtained by using energetic coupling values of specific regions in ED3: the BC loop in WNV and the FG loop in DENV2. These regions in ED3 were the most susceptible to perturbation due to mutations in ED3 and are virus-specific. That is, the distribution of mutations in Figure 5.6 is lost when the values of Kd_{app} are correlated with the energetic coupling in the BC loop for DENV2 and the FG loop for WNV as shown in Figure 5.7, where mutations that decrease MAb binding affinity fell into intermediate values of energetic coupling as in the case of Mab non-resistant mutations.

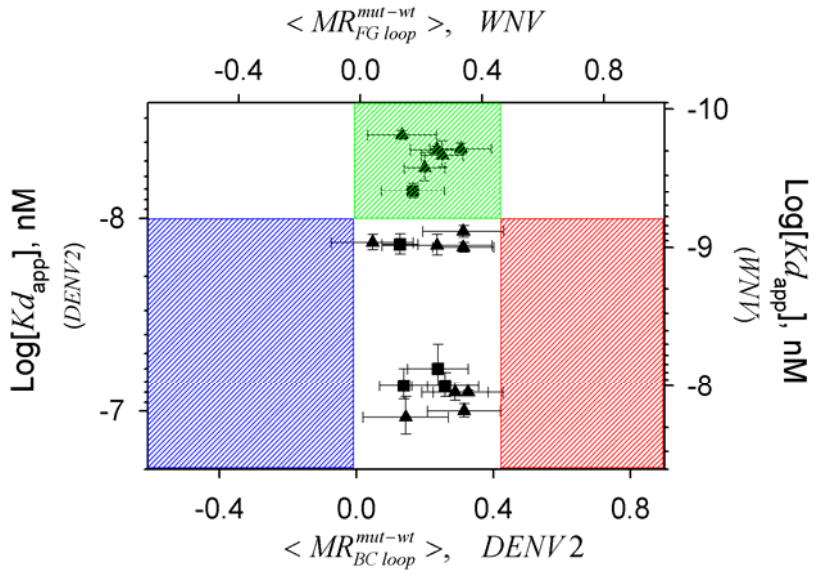

Figure 5.7. Lost correlation between energetic coupling perturbation and MAb binding affinity in ED3 from WNV (squares) and DENV2 (triangles) when the energetic coupling values corresponded to the FG loop for WNV and BC loop for DENV2. Blue, red and green shades indicate negative, positive and intermediate coupling.

Because the energy-landscape map in Figure 5.6 is pathway independent, then mutations anywhere in the ED3 can generate MAb-neutralization resistance, provided

that the substitution is energetically couple to the dominant epitopes in ED3 (BC loop or FG loop in ED3 from WNV or DENV2, respectively). This observation underscores the role of long-range interactions in MAb-neutralization resistance and provides, for the first time, a fundamental rationale for the presence of mutations outside the binding site (viral epitope) as it has been described in flaviviruses[16,28-30,38-40,111].

5.1 BIOLOGICAL SIGNIFICANCE OF LONG-RANGE COMMUNICATION IN VIRAL PROTEINS

ED3 in flaviviruses encodes critical epitopes recognized by potent virus type-specific neutralizing MAbs[12], and accordingly, it harbors most of the critical mutations that enable the virus to escape from neutralizing MAbs. However, ED3 also plays crucial roles during the multiplication cycle of flaviviruses: it is the putative cell-receptor-binding domain that triggers receptor-mediated endocytosis[17]; and during the fusion process between the viral and endosome membranes, that allows release of the genetic material into the cytoplasm, the envelope protein requires a dramatic domain reorganization in which ED3 plays a major role[114,115]. Thus the mutations selected in ED3 to generate MAb-escape mutants have to pass a 'quality control' in which the ED3 is able to preserve its functions. Results of this study provide insights on the strategy employed by the virus to achieve the goal of evading MAb binding without affecting its ability to carry out the essential viral functions. ***The strategy seems to reside at least in part in the property of energetic coupling among distant sites in the protein.***

In addition to performing different functions, ED3 is also the major target of neutralizing MAbs. ED3, being a small protein (~100 residues) with a small solvent-expose surface of residues (Figure 4.1, Chapter 4) is, therefore, under high selection pressure because there are limited number of residues available for mutations to generate a viable MAb-escape variant yet maintaining its normal functions. However, in this study it has clearly been demonstrated that decrease in MAb binding affinity correlates with the degree of energetic coupling on virus-specific regions of ED3, namely, high negative/

positive energetic coupling is correlated with low MAb binding affinity. Since the energetic coupling between residues is pathway-independent and long-range, then the mutation site can potentially be located anywhere in the protein (or even in a different domain of the envelope protein). Thus, ***the mutational space to generate MAb-neutralization resistance in ED3 is significantly expanded without compromising residues directly involved in other functions that are important for virus viability***.

Finally, we have observed a correlation between *low MAb binding affinity and high negative/ positive energetic coupling.* This correlation was established only in virus-specific structural element in the protein that corresponded to the dominant viral epitopes for DENV2 and WNV; namely, the FG and BC loops in ED3, respectively. Thus, **what would be the advantage of having one (or a very limited number) immunodominant epitope and not several?** Let us consider two scenarios:

1. *The viral protein antigen has numerous dominant epitopes*. This situation will generate a large diversity in the mutational space of the virus and the host will have to elicit a greater number of MAbs against the numerous epitopes. However, by having a large panel of MAbs with strong neutralizing capability, the chances of having MAbs binding to regions that are functionally important for the virus (e.g. residues involved in cell-receptor binding) increases significantly. In the case of ED3, which has a small surface of residues exposed to the solvent, this is a critical disadvantage because it is most likely these epitopes would most likely overlap with viral surfaces involved in cell-receptor binding.
2. *The viral protein antigen has one dominant epitope*. The diversity in the mutational space will be limited and the panel of MAbs produced by the host will react against this dominant epitope. However, the virus gains an advantage in that other regions functionally important for the virus most likely will not be compromised either by the MAbs (because they will react against the dominant epitope) nor by mutations because those residues will not be under high selection pressure.

The virus adopted the strategy of having ***one dominant epitope energetically coupled to other regions of the protein,*** the mutational space is expanded significantly (potentially to other domains as well) generating a large conformational diversity in the viral epitope but the mutations do not compromise residues that are solvent-exposed important for other functions such as cell-receptor binding and endosomal membrane fusion process. Therefore, the virus-specific antigenic determinants of neutralization emerge from the localization of the mutational response of the energetic coupling in ED3.

REFERENCES

1. Mackenzie, J. S., Gubler, D. J. and Petersen, L. R. (2004). Emerging flaviviruses: the spread and resurgence of Japanese encephalitis, West Nile and dengue viruses. *Nat. Med.* **10**, S98-S109.

2. Kuno, G. and Chang G. J. (2005). Biological transmission of arboviruses: reexamination of and new insights into components, mechanisms, and unique traits as wells as their evolutionary trends. *Clin. Micr. Rev.* **18**, 608-637.

3. Clyde, K., Kyle, J. L. and Harris, E. (2006). Recent advances in deciphering viral and host determinants of dengue virus replication and pathogenesis. *J. Virol.* **80**, 11418-11431.

4. Mukhopadhyay, S., Kuhn, R. J. and Rossmann, M. G. (2005). A structural perspective of the flavivirus life cycle. *Nature Rev. Microbiol.* **3**, 13-22.

5. Lindenbach, B. D. and Rice, C. M. in Fields Virology (eds Knipe, D. M. and Howley, P. M.) 991−1041 (Lippincott Williams & Wilkins, Philadelphia, 2001).

6. Briese, T., Jia, X. Y., Huang, C., Grady, L. J. and Lipkin, W. I. (1999). Identification of a Kunjin/West Nile-like flavivirus in brains of patients with New York encephalitis. *Lancet* **354**, 1261-1262.

7. Lanciotti, R. S., Roehrig, J. T., Deubel, V., Smith, J., Parker, M., Steele, K., Crise, B., Volpe, K. E., Crabtree, M. B., Scherret, J. H., Hall, R. A., MacKenzie, J. S., Cropp, C. B., Panigrahy, B., Ostlund, E., Schmitt, B., Malkinson, M., Banet, C., Weissman, J., Komar, N., Savage, H. M., Stone, W., McNamara, T. and Gubler D. J. (1999). Origin of the West Nile virus

responsible for an outbreak of encephalitis in the northeastern United States. *Science* **286**, 2333-2337.

8. Beasley, D. W., Davis, C. T., Estrada-Franco, J., Navarro-Lopez, R., Campomanes-Cortes, A., Tesh, R. B., Weaver, S. C. and Barrett, A. D. (2004). Genome sequence and attenuating mutations in West Nile virus isolate from Mexico. *Emerg. Infect. Dis.* **10**, 2221-2224.

9. Deardorff, E., Estrada-Franco, J., Brault, A. C., Navarro-Lopez, R., Campomanes-Cortes, A., Paz-Ramirez, P., Solis-Hernandez, M., Ramey, W. N., Davis, C. T., Beasley, D. W., Tesh, R. B., Barrett, A. D. and Weaver, S. C. (2006). Introductions of West Nile virus strains to Mexico. *Emerg. Infect. Dis.* **12**, 314-318.

10. Berrocal, L., Pena, J., Gonzalez, M. and Mattar, S. (2006). West Nile virus; ecology and epidemiology of an emerging pathogen in Colombia. *Rev. Salud Publica (Bogota)* **8**, 218-28.

11. Morales, M. A., Barrandeguy, M., Fabbri, C., Garcia, J. B., Vissani, A., Trono, K., Gutierrez, G., Pigretti, S., Menchaca, H., Garrido, N., Taylor, N., Fernandez, F., Levis, S. and Enria, D. (2006) West Nile virus isolation from equines in Argentina. *Emerg. Infect. Dis.* **12**, 1559-1561.

12. Diamond M. S. (2004). Evasion of innate and adaptive immunity by flaviviruses. *Immunol. Cell Biol.* **81**, 196-206

13. Engle, M. J and Diamond, M. S. (2003). Antibody prophylaxis and therapy against West Nile virus infection in wild-type and immunodeficient mice. *J. Virol.* **77**, 12941-12949.

14. Oliphant, T., Engle, M., Nybakken, G. E., Doane, C., Johnson, S., Huang, L., Gorlatov, S., Mehlhop, E., Marri, A., Chung, K. M., Ebel, G. D., Kramer, L. D., Fremont, D. H. and Diamond, M. S. (2005). Development of a humanized monoclonal antibody with therapeutic potential against West Nile virus. *Nat. Med.* **11**, 522-530.

15. Beasley, D. W. and Barrett, A. D. (2002). Identification of neutralizing epitopes within structural domain 3 of the West Nile virus envelope protein. *J. Virol.* **76**, 13097-13100.

16. Li, L., Barrett, A. D. and Beasley, D. W. (2005). Differential expression of domain 3 neutralizing epitopes on the envelope proteins of West Nile virus strains. *Virology* **335**, 99-105.

17. Rey, F. A., Heinz, F. X., Mandl, C., Kunz, C. and Harrison, S. C. (1995). The envelope glycoprotein from tick-borne encephalitis virus at 2 Å resolution. *Nature* **375**, 291–298.

18. Zhang, Y., Zhang, W., Ogata, S., Clements, D., Strauss, J. H., Baker, T. S., Kuhn, R. J. and Rossmann, M. G. (2004). Conformational changes of the flavivirus E glycoprotein. *Structure* **12**, 1607-1618.

19. Volk, D. E., Beasley, D. W., Kallick, D. A., Holbrook, M. R., Barrett, A. D. and Gorenstein, D. G. (2004). Solution structure and antibody binding studies of the envelope protein domain 3 from the New York strain of West Nile virus. *J. Biol. Chem.* **279**, 38755-38761.

20. Mandl, C.W., Allison, S.L., Holzmann, H., Meixner, T. and Heinz, F.X. (2000). Attenuation of tick-borne encephalitis virus by structure-based site-specific mutagenesis of a putative flavivirus receptor binding site. *J. Virol.* **74**, 9601-9609.

21. Bhardwaj, S., Holbrook, M., Shope, R.E., Barrett, A.D. and Watowich, S.J. (2001). Biophysical characterization and vector-specific antagonist activity of domain 3 of the tick-borne flavivirus envelope protein. *J. Virol.* **75**, 4002-4007.

22. Crill, W.D. and Roehrig, J.T. (2001). Monoclonal antibodies that bind to domain 3 of dengue virus E glycoprotein are the most efficient blockers of virus adsorption to Vero cells. *J. Virol.* **75**, 7769-7773.

23. Lin, C.W. and Wu, S.C. (2003). A functional epitope determinant on domain 3 of the Japanese encephalitis virus envelope protein interacted with neutralizing-antibody combining sites. *J. Virol.* **77**, 2600-2606.

24. Nybakken, G. E., Oliphant, T., Johnson, S., Burke, S., Diamond, M. S. and Fremont, D. H. (2005). Structural basis of West Nile virus neutralization by a therapeutic antibody. *Nature* **437**, 764-768.

25. Mukhopadhyay, S., Kim, B. S., Chipman, P. R., Rossmann, M. G. and Kuhn, R. J. (2003). Structure of West Nile virus. *Science* **302**, 248.

26. 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. and Strauss, J. H. (2002). Structure of dengue virus: implications for flavivirus organization, maturation, and fusion. *Cell* **108**, 717-725.

27. Zhang, W., Chipman, P. R., Corver, J., Johnson, P. R., Zhang, Y., Mukhopadhyay, S., Baker, T. S., Strauss, J. H., Rossmann, M. G. and Kuhn, R. J. (2003). Visualization of membrane protein domains by cryo-electron microscopy of dengue virus. *Nat. Struct. Biol.* **10**, 907-912.

28. Buckley, A. and Gould, E. A. (1985). Neutralization of yellow fever virus studied using monoclonal and polyclonal antibodies. *J. Gen. Virol.* **66**, 2523-2531.

29. Cecilia, D. and Gould, E. A. (1991). Nucleotide changes responsible for loss of neuroinvasiveness in Japanese encephalitis virus neutralization-resistant mutants. *Virology* **181**, 70-77.

30. Holzmann, H., Stiasny, K., Ecker, M., Kunz, C. and Heinz, F. X. (1997). Characterization of monoclonal antibody-escape mutants of tick-borne encephalitis virus with reduced neuroinvasiveness in mice. *J. Gen. Virol.* **78**, 31-37.

31. Kalia, V., Sarkar, S., Gupta, P. and Montelaro, R. C. (2005). Antibody neutralization escape mediated by point mutations in the intracytoplasmic tail of human immunodeficiency virus type 1 gp41. *J. Virol.* **79**, 2097-2107.

32. Edwards, T. G., Wyss, S., Reeves, J. D., Zolla-Pazner, S., Hoxie, J. A., Doms, R. W. and Baribaud, F. (2002). Truncation of the cytoplasmic domain induces exposure of conserved regions in the ectodomain of human immunodeficiency virus type 1 envelope protein. *J. Virol.* **76**, 2683-2691.

33. Chambers, T.J., Hahn, C.S., Galler, R. and Rice, C.M. (1990). Flavivirus genome organization, expression, and replication. *Annu. Rev. Microbiol.* **44**, 649-688.

34. Granwehr, B., Lillibridge, K., Higgs, S., Mason, P., Aronson, J., Campbell, J. and Barrett, A.D. (2004). West Nile virus: where are we now? *Lancet Infect. Dis.***4**, 547-556.

35. McMinn, P.C. (1997). The molecular basis of virulence of the encephalitogenic flaviviruses. *J. Gen. Virol.* **78**, 2711-2722.

36. Modis, Y., Ogata, S., Clements, D. and Harrison, S.C. (2003). A ligand-binding pocket in the dengue virus envelope glycoprotein. *Proc. Natl. Acad. Sci.*, **100**, 6986-6991.

37. Kaufmann, B., Nybakken, G. E., Chipman, P. R., Zhang, W., Diamond, M. S., Fremont, D. H., Kuhn, R. J. and Rossmann, M. G. (2006). West Nile virus in complex with the Fab fragment of a neutralizing monoclonal antibody. *Proc. Natl. Acad. Sci.* **103,** 12400-12404.

38. Heinz, F. X., Mandl, C., Berger, R., Tuma, W. and Kunz, C. (1984). Antibody-induced conformational changes result in enhanced avidity of antibodies to different antigenic sites on the tick-borne encephalitis virus glycoprotein. *Virology* **133**, 25-34.

39. Cecilia, D., Gadkari, D. A., Kedarnath, N. and Ghosh, S. N. (1988). Epitope mapping of Japanese encephalitis virus envelope protein using monoclonal antibodies against an Indian strain. *J. Gen. Virol.* **69**, 2741-2747.

40. Guirakhoo, F., Heinz, F. X. and Kunz, C. (1989). Epitope model of tick-borne encephalitis virus envelope glycoprotein E: analysis of structural properties, role of carbohydrate side chain, and conformational changes occurring at acidic pH. *Virology* **169**, 90-99.

41. Kawahara, K. and Tanford, C. (1966). Viscosity and density of aqueous solutions of urea and guanidine hydrochloride. *J. Biol. Chem.* **241**, 3228-3232.

42. Yu, S., Wuu, A., Basu, R., Holbrook, M. R., Barrett, A. D. and Lee J. C. (2004). Solution structure and structural dynamics of envelope protein domain 3 of mosquito- and tick-borne flaviviruses. *Biochemistry* **43**, 9168-9176.

43. Pace, C. N., Vajdos, F., Fee, L., Grimsley, G. and Gray, T. (1995). How to measure and predict the molar absorption coefficient of a protein. *Protein Sci.* **4**, 2411-2423.

44. Bolen, D. W. and Santoro, M. M. (1988). Unfolding free energy changes determined by the linear extrapolation method. 1. Unfolding of phenylmethanesulfonyl alpha-chymotrypsin using different denaturants. *Biochemistry* **27**, 8069-8074.

45. Lacowicz, J. R. (1999). Principles of Fluorescence Spectroscopy, Chapter 1, Plenum Press, New York.

46. Longworth, J. W. and Rahn, R. O. (1967). Energy transfer in poly-L-tyrosine as a function of the degree of ionization of the phenolic hydroxyls. I. Room-temperature fluorescence of model compounds and poly-L-tyrosine. *Biochim. Biophys. Acta* **147**, 526-535.

47. Stern, O. and Volmer, M. (1919). Über die abklingungszeit der fluoreszenz. *Phys. Z.* **20**, 183-188.

48. Eftink, M. R. and Ghiron, C. A. (1981). Fluorescence Quenching Studies with Proteins. *Anal. Biochem.* **114**, 199-227.

49. Ragone, R., Colonna, G., Balestrieri, C., Servillo, L. and Irace, G. (1984). Determination of tyrosine exposure in proteins by second-derivative spectroscopy. *Biochemistry* **23**, 1871-1875.

50. Savitzky, A., and Golay, M. J. E. (1964) Smoothing and differentiation of data by simplified least squares procedures. *Anal. Chem.* **36**, 1627-1639.

51. Parker, C. A. and Rees W. T. (1960). Correction of Fluorescence Spectra and Measurement of Fluorescence Quantum Efficiency. *Analyst* **85**, 587-592

52. Gill, J. E. (1969). The Fluorescence Excitation Spectrum of Quinine Bisulfate. *Photochem. Photobiol.* **9**, 313-322.

53. Schuck, P. (2000). Size-distribution analysis of macromolecules by sedimentation velocity ultracentrifugation and lamm equation modeling. *Biophys. J.* **78**, 1606-1619.

54. Cohn, E. J. and Edsall, J. T. (1943). Proteins, Amino Acids and Peptides as Ions and Dipolar Ions. Chapter 4, Rheinhold, New York.

55. Makhatadze, G. I., Loladze, V. V.., Ermolenko, D. N., Chen, X. and Thomas S. T. (2003). Contribution of surface salt bridges to protein stability: guidelines for protein engineering. *J. Mol. Biol.* **327**, 1135-1148.

56. Loladze, V. V. and Makhatadze, G. I. (2002). Removal of surface charge-charge interactions from ubiquitin leaves the protein folded and very stable. *Protein Sci.* **11**, 174-177.

57. Loladze, V. V., Ibarra-Molero, B., Sanchez-Ruiz, J. M. and Makhatadze, G. I. (1999). Engineering a thermostable protein via optimization of charge-charge interactions on the protein surface. *Biochemistry* **38**, 16419-16423.

58. Lacowicz, J. R. (1999). Principles of Fluorescence Spectroscopy, Chapter 8, Plenum Press, New York.

59. Bujalowski, W. and Klonowska, M. M. (1994). Close proximity of tryptophan residues and ATP-binding site in Escherichia coli primary replicative helicase DnaB protein. Molecular topography of the enzyme. *J. Biol. Chem.* **269**, 31359-31371.

60. Eftink, M. R. and Ghiron, C. A. (1976). Fluorescence quenching of indole and model micelle systems. *J. Phys. Chem.* **80**, 486-493.

61. Fraczkiewicz, R. and Braun, W. (1998). Exact and efficient analytical calculation of the accessible surface areas and their gradients for macromolecules. *J. Comp. Chem.* **19**, 319-333.

62. Lacowicz, J. R. (1999). Principles of Fluorescence Spectroscopy, Chapter 16, Plenum Press, New York.

63. Berman, H. A. and Taylor P. (1978). Fluorescent phosphonate label for serine hydrolases, pyrenebutyl methylphosphonofluoridate: reaction with acetylcholinesterase. *Biochemistry* **17**, 1704-1713.

64. Eftink, M. R., Zajicek, J. L. and Ghiron, C. A. (1977). A hydrophobic quencher of protein fluorescence: 2,2,2-trichloroethanol. *Biochim. Biophys. Acta* **491**, 473-481.

65. Ando, T. and Asai, H. (1980). Charge effects on the dynamic quenching of fluorescence of 1,N6-ethenoadenosine oligophosphates by iodide, thallium (I) and acrylamide. J. Biochem. (Tokyo) **88**, 255-264.

66. Wetlaufer, D. B., Edsall, J. T. and Hollingworth, B. R. (1958). Ultraviolet difference spectra of tyrosine groups in proteins and amino acids. *J. Biol. Chem.* **233**, 1421-1428.

67. Donovan, J. W. (1969). Changes in ultraviolet absorption produced by alteration of protein conformation. *J. Biol. Chem.* **244**, 1961-1967.

68. Brahms, S., Brahms, J., Spach, G. and Brack, A. (1977). Identification of β,β-turns and unordered conformations in polypeptide chains by vacuum ultraviolet circular dichroism. *Proc. Natl. Acad. Sci.,* **74**, 3208-3212.

69. Chang, C. T., Wu, C. S. and Yang, J. T. (1978). Circular dichroic analysis of protein conformation: inclusion of the beta-turns. *Anal. Biochem.* **91**, 13-31.

70. Johnson, W. C. (1999). Analyzing protein circular dichroism spectra for accurate secondary structures. *Proteins* **35**, 307-312.

71. Provencher, S. W. and Glockner, J. (1981). Estimation of globular protein secondary structure from circular dichroism. *Biochemistry* **20**, 33-37.

72. Eisinger, J. (1969). Intramolecular energy transfer in adrenocorticotropin. *Biochemistry* **8**, 3902-3908.

73. Lehrer, S. S. (1971). Solute perturbation of protein fluorescence. The quenching of the tryptophyl fluorescence of model compounds and of lysozyme by iodide ion. *Biochemistry* **10**, 3254-3263.

74. Makhatadze, G. I. and Privalov, P. L. (1993). Contribution of hydration to protein folding thermodynamics. I. The enthalpy of hydration. *J. Mol. Biol.* **232**, 639-659.

75. Loladze, V. V.., Ermolenko, D. N. and Makhatadze, G. I. (2002). Thermodynamic consequences of burial of polar and non-polar amino acid residues in the protein interior. *J. Mol. Biol.* **320**, 343-357.

76. Schellman, J. A. (1978). Solvent denaturation. *Biopolymers* **17**, 1305-1322.

77. Smith, C. K., Bu, Z., Anderson, K. S., Sturtevant, J. M., Engelman, D. M. and Regan, L. (1996). Surface point mutations that significantly alter the structure and stability of a protein's denatured state. *Protein Sci.* **5**, 2009-2019.

78. Myers, J. K., Pace, C. N. and Scholtz, J. M. (1995). Denaturant m values and heat capacity changes: relation to changes in accessible surface areas of protein unfolding. *Protein Sci.* **4**, 2138-2148.

79. Pace, C. N., Laurents, D. V. and Erickson, R. E. (1992). Urea denaturation of barnase: pH dependence and characterization of the unfolded state. *Biochemistry* **31**, 2728-2734.

80. Sundberg, E. J. and Mariuzza, R. A. (2002). Molecular recognition in antibody-antigen complexes. *Adv. Protein Chem.* **61**, 119-160.

81. Davies, D. R. and Cohen, G. H. (1996). Interactions of protein antigens with antibodies. *Proc. Natl. Acad. Sci.* **93**, 7-12.

82. Davies, D. R., Sheriff, S. and Padlan, E. A. (1988). Antibody-antigen complexes. *J. Biol. Chem.* **263**, 10541-10544.

83. Braden, B. C. and Poljak, R. J. (1995). Structural features of the reactions between antibodies and protein antigens. *FASEB J.* **9**, 9-16.

84. Maillard, R. A., Jordan, M., Beasley, D. W., Barrett A. D. and Lee, J. C. Resistance of West Nile virus to Antibody-mediated Neutralization: Long-range Communications in the Envelope Protein Domain III is the Key (*submitted*).

85. Lin, C. W. and 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**, 2600-2606.

86. Tainer, J. A., Getzoff, E. D., Alexander, H., Houghten, R. A., Olson, A. J., Lerner, R. A. and Hendrickson, W. A. (1984). The reactivity of anti-peptide antibodies is a function of the atomic mobility of sites in a protein. *Nature* **312**, 127-134.

87. Westhof, E., Altschuh, D., Moras, D., Bloomer, A. C., Mondragon, A., Klug, A. and Van Regenmortel, M. H. (1984). Correlation between segmental mobility and the location of antigenic determinants in proteins. *Nature* **311**, 123-126.

88. Zimmermann, J., Oakman, E. L., Thorpe, I. F., Shi, X., Abbyad, P., Brooks, C. L. III, Boxer, S. G. and Romesberg, F. E. (2006). Antibody evolution constrains conformational heterogeneity by tailoring protein dynamics. *Proc. Natl. Acad. Sci.* **103**, 13722-13727.

89. Jimenez, R., Salazar, G., Yin, J., Joo, T. and Romesberg, F. E. (2004). Protein dynamics and the immunological evolution of molecular recognition. *Proc. Natl. Acad. Sci.* **101**, 3803-3808.

90. Fieser, T. M., Tainer, J. A., Geysen, H. M., Houghten, R. A. and Lerner, R. A. (1987). Influence of protein flexibility and peptide conformation on reactivity of monoclonal anti-peptide antibodies with a protein alpha-helix. *Proc. Natl. Acad. Sci.* **84**, 8568-85672.

91. Hilser, V. J. and Freire E. (1996). Structure-based calculation of the equilibrium folding pathway of proteins. Correlation with hydrogen exchange protection factors. *J. Mol. Biol.* **262**, 756-772.

92. Hilser, V. J., Dowdy, D., Oas, T. G. and Freire E. (1998). The structural distribution of cooperative interactions in proteins: analysis of the native state ensemble. *Proc. Natl. Acad. Sci.* **95**, 9903-9908.

93. Whitten, S. T., Garcia-Moreno, E. B. and Hilser V. J. (2005). Local conformational fluctuations can modulate the coupling between proton binding and global structural transitions in proteins. *Proc. Natl. Acad. Sci.* **102**, 4282-4287.

94. Pan, H., Lee, J. C. and Hilser V. J. (2000). Binding sites in Escherichia coli dihydrofolate reductase communicate by modulating the conformational ensemble. *Proc. Natl. Acad. Sci.* **97**, 12020-12025.

95. Liu, T., Whitten, S. T. and Hilser V. J. (2006). Ensemble-based signatures of energy propagation in proteins: a new view of an old phenomenon. *Proteins* **62**, 728-738.

96. Planells-Cases, R., Ferrer-Montiel, A. V., Patten, C. D. and Montal M. (1995). Mutation of conserved negatively charged residues in the S2 and S3 transmembrane segments of a mammalian K^+ channel selectively modulates channel gating. *Proc. Natl. Acad. Sci.* **92**, 9422-9426.

97. Tytgat, J. and Hess P. (1992). Evidence for cooperative interactions in potassium channel gating. *Nature* **359**, 420-423.

98. Gubler, D. J. (1998). Dengue and dengue hemorrhagic fever. *Clin. Microbiol. Rev.* **11**, 480-496.

99. Kaufman, B. M., Summers, P. L., Dubois, D. R., Cohen, W. H., Gentry, M. K., Timchak, R. L., Burke, D. S. and Eckels, K. H. (1989). Monoclonal antibodies for dengue virus prM glycoprotein protect mice against lethal dengue infection. *Am. J. Trop. Med. Hyg.* **41**, 576-580.

100. Kaufman, B. M., Summers, P. L., Dubois, D. R. and Eckels, K. H. (1987). Monoclonal antibodies against dengue 2 virus E-glycoprotein protect mice against lethal dengue infection. *Am. J. Trop. Med. Hyg.* **36**, 427-434.

101. Hiramatsu, K., Tadano, M., Men, R. and 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**, 437-445.

102. Trirawatanapong, T., Chandran, B., Putnak, R. and Padmanabhan, R. (1992). Mapping of a region of dengue virus type-2 glycoprotein required for binding by a neutralizing monoclonal antibody. Gene **116**, 139-150.

103. Gromowski, D. G. and Barrett A. D. Characterization of an Antigenic Site on the Envelope Protein Domain III (ED3) of Dengue 2 Virus that Represents a Dominant, Type-Specific, Antibody-Mediated Neutralization Determinant *(in preparation)*.

104. Wang, E., Ni, H., Xu, R., Barrett, A. D., Watowich, S. J., Gubler, D. J. and Weaver, S. C. (2000). Evolutionary relationships of endemic/epidemic and sylvatic dengue viruses. *J. Virol.* **74**, 3227-3234.

105. Twiddy, S. S., Holmes, E. C. and Rambaut, A. (2003). Inferring the rate and time-scale of dengue virus evolution. *Mol. Biol. Evol.* **20**, 122-129.

106. Gaunt, M. W., Sall, A. A., de Lamballerie, X., Falconar, A. K., Dzhivanian, T. I. and Gould, E. A. (2001). Phylogenetic relationships of flaviviruses correlate with their epidemiology, disease association and biogeography. *J. Gen. Virol.* **82**, 1867-1876.

107. Wu, K. P., Wu, C. W., Tsao, Y. P., Kuo, T. W., Lou, Y. C., Lin, C. W., Wu, S. C. and 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**, 46007-46013.

108. Holzmann, H., Heinz, F. X., Mandl, C. W., Guirakhoo, F. and 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.

109. Zhang, S., Bovshik, E. I., Maillard, R. A., Volk, D. E., Schein, C., Gorenstein, D. G., Lee, J. C., Barrett, A. D. and Beasley, D. W. C. Critical role of a Tyrosine corner motif in the envelope protein receptor-binding domain of mosquito-borne flaviviruses (*submitted*)

110. Yewdell, J. W. and Gerhard, W. (1981). Antigenic characterization of viruses by monoclonal antibodies. *Annu. Rev. Microbiol.* **35**, 185-206.

111. Goncalvez, A. P., Purcell R. H. and Lai, C.-J. (2004). Epitope determinants of a chimpanzee Fab Antibody That Efficiently Cross-Neutralizes Dengue Type 1 and Type 2 Viruses Map to Inside and in Close Proximity to Fusion Loop of the Dengue Type 2 Virus Envelope Glycoprotein. *J. Virol.* **78**, 12919-12928.

112. Kalia, V., Sarkar, S., Gupta, P. and Montelaro, R. C. (2005). Antibody neutralization escape mediated by point mutations in the intracytoplasmic tail of human immunodeficiency virus type 1 gp41. *J. Virol.* **79**, 2097-2107.

113. Edwards, T. G., Wyss, S., Reeves, J. D., Zolla-Pazner, S., Hoxie, J. A., Doms, R. W. and Baribaud, F. (2002). Truncation of the cytoplasmic domain induces exposure of conserved regions in the ectodomain of human immunodeficiency virus type 1 envelope protein. *J. Virol.* **76**, 2683-2691.

114. Bressanelli, S., Stiasny, K., Allison, S. L., Stura, E. A., Duquerroy, S., Lescar, J., Heinz, F. X. and Rey, F. A. (2004). Structure of a flavivirus envelope glycoprotein in its low-pH-induced membrane fusion conformation. *EMBO J.* **23**, 728-738.

115. Modis, Y., Ogata, S., Clements, D. and Harrison, S. C. (2004). Structure of the dengue virus envelope protein after membrane fusion. *Nature* **427**, 313-319.

116. Hilser, V. J. (2001). Modeling the native state ensemble. *Methods Mol. Biol.* **168**, 93-116.

VITA

Rodrigo Alejandro Maillard Reyes was born on March 15, 1979 in Lima-Peru to parents Carlos Alejandro Maillard and Irene Andrea Reyes. After high school, Rodrigo pursued a major in Biology at Universidad Peruana Cayetano Heredia, where he met his wife, Maria Fe Lanfranco. Soon after he started his undergraduate studies in 1997, Rodrigo participated in diverse research projects. In fall 1998 he did a research internship at Arizona State University in the department of Molecular and Cellular Biology under the guidance of Drs. Douglas E. Chandler and Robert McGaughey. In 1999 Rodrigo worked in the isolation and purification of plant-derived secondary metabolites with Dr. Abraham Vaisberg at Universidad Peruana Cayetano Heredia. And before graduating he met Dr. Jorge Arevalo from the Tropical Medicine Institute "Alexander Von Humboldt" in Lima-Peru and conducted DNA sequence analysis of Leishmania and related parasites. Rodrigo was able to present his latest undergraduate research work in the Molecular Evolution meeting in Italy in June 2002, just before starting his graduate studies. In August 2002, he joined the new graduate track in Biophysical, Structural and Computational Biology at The University of Texas Medical Branch in Galveston, Texas. After completion on his first year of laboratory rotations and coursework, Rodrigo joined the laboratory of Dr. James C. Lee and conducted his dissertation research focusing on the molecular mechanism of antibody-mediated neutralization resistance in flaviviruses. Rodrigo has presented his research orally at the Pediatric Dengue Vaccine Development meeting in 2005 and 2006, at the national Protein Society meeting (2006) and at the Gibbs Conferences in Biothermodynamics (2006). Following the completion of his dissertation, Rodrigo will continue his scientific training starting on March 2007 in the field of single molecule biophysics under the supervision of Dr. Carlos Bustamante at the University of California in Berkeley.

Education

B.Sc., May 2002, Universidad Peruana Cayetano Heredia

Publications

Maillard, R. A., Jordan, M., Beasley, D. W., Barrett, A. and Lee, J. C. Resistance of West Nile Virus to Antibody-mediated Neutralization: Long-range Communications in the Envelope Protein Domain III is the Key (submitted to Journal of Biological Chemistry).

Maillard, R. A., Liu, T., Barrett, A. D., Hilser, V. J. and Lee J. C. Resistance to Antibody-mediated Neutralization in West Nile virus is Modulated by the Redistribution of Conformational States of the Envelope Protein Domain 3 (in preparation)

Maillard, R. A., Gromowski, D. G., Liu, T., Barrett, A. D., Hilser V. J. and Lee, J. C. Identification of Antigenic Determinants of Neutralization in Flaviviruses by Mutational Response of the Ensemble of the Envelope Protein Domain 3 (in preparation)

Zhang, S., Bovshik, E. I., Maillard, R. A., Volk, D. E., Schein, C., Gorenstein, D. G., Lee, J. C., Barrett, A. D. and Beasley, D. W. C. Critical role of a Tyrosine corner motif in the envelope protein receptor-binding domain of mosquito-borne flaviviruses (submitted to PLOS Pathogens)

Yu, S., Maillard, R. A. and Lee, J. C. Perturbation of Structural Dynamics and Thermodynamics of cAMP binding in E. coli cAMP Receptor Protein Induced by Replacements of Side-chain at Residue 138 (submitted to Journal of Biological Chemistry).

Galletto, R., Jezewska, M. J., Maillard, R. and Bujalowski, W. (2005). The Nucleotide-Binding Site of the Escherichia coli DnaC Protein. Molecular Topography of DnaC Protein – Nucleotide Cofactors Complexes. *Cell Biochemistry and Biophysics* **43**, 331-353.

Galletto, R., Maillard, R., Jezewska, M. J. and Bujalowski, W. (2004). Global Conformation of the Escherichia coli Replication Factor DnaC Protein in Absence and Presence of Nucleotide Cofactors. *Biochemistry* **43**, 10988-11001.

Abstracts

Maillard, R. A., Liu, T., Barrett, A. D., Hilser, V. J. and Lee, J. C. Neutralization Resistance of West Nile Virus is Modulated by the Energy of Coupling Among a Network of Residues in the Domain III of the Viral Envelope Protein. Sealy Center for Structural Biology and Molecular Biophysics 11th Symposium, Galveston, Texas, 2006 AND The Protein Society Meeting, 2006

Maillard, R. A., Jordan, M., Beasley, D. W., Barrett, A. D. and Lee, J. C. Long-Range Communications in Envelope Protein Domain III of Neutralization-Resistant Mutants of West Nile Virus. Infectious Disease Workshop, London, UK, 2005

Maillard, R. A., Jordan, M. and Lee, J. C. Neutralizing-Resistant Mutations of West Nile Virus Envelope Protein Domain III Exert Structural Perturbations at Distant Sites. 5th McLaughlin Symposia in Infection & Immunity, Galveston, Texas, 2005

Bujalowski, W., Galletto, R., Maillard, R. A. and Jezewska, M. J. Global Conformations of the Escherichia coli Replication Factor DnaC Protein in Absence and Presence of Nucleotide Cofactors. 49th Biophysical Society Meeting, 2005

Maillard, R. A., Jordan, M. and Lee, J. C. Solution Properties and Energetics of Envelope Protein Domain III from West Nile Virus: Comparing Wild Type and Neutralization Resistance Mutants. 49th Biophysical Society Meeting, 2005

Maillard R, Arévalo J, Guerra D and Mirko J. Zimic. Nuclear DNA Coding Sequence Analysis supports that Nucleotide Composition correlates with the Phylogenetic Organization of Microevolutive Lineages. Molecular Evolution Conference, Sorrento, Italy, 2002