

Copyright
by
Jeff Borgeson
2010

**The Thesis Committee for Jeff A. Borgeson Certifies that this is the approved version of the
following thesis:**

**Structural Investigation of 1,8-Dinitro-4,5-Dihydroxyanthraquinone
and Implications for Docking as a Small Molecule Inhibitor
into the Protease of West Nile Virus**

Committee:

Stanley J. Watowich, PhD, Supervisor

Andres F. Oberhauser, PhD

Marc C. Morais, PhD

Dean, Graduate School

**Structural Investigation of 1,8-Dinitro-4,5-Dihydroxyanthraquinone
and Implications for Docking as a Small Molecule Inhibitor
into the Protease of West Nile Virus**

by

Jeffrey Allen Borgeson

Thesis

Presented to the Faculty of the Graduate School of

The University of Texas Medical Branch

in Fulfillment

of the Requirements

for the Degree of

MASTERS OF SCIENCE

The University of Texas Medical Branch

May, 2010

Acknowledgements

Contributions from many individuals have come together to foster the completion of this thesis. First and foremost, I would like to express my deepest gratitude towards my supervisor and mentor, Dr. Stanley Watowich, PhD, for providing me with an opportunity to train in his lab and for sharing wisdom throughout an important transition in my life. Dr. Watowich is an excellent research advisor, respectfully maintains high expectations, and provides intuitive scientific guidance. I am also very grateful to the talented and dedicated members of the Watowich lab, Kristen Lamb, Robert Malmstrom, and Suzanne Tomlinson, for their friendship and willingness to discuss their scientific knowledge. I would like to extend my appreciation to my committee members, Dr. Andres Oberhauser, Dr. Marc Morais, Dr. Kay Choi and Dr. Lee for their continual guidance, encouragement, support, and patience throughout the past year. The technical expertise of X-Ray Crystallography shared by Dr. Mark White was integral to the completion of this project, of which I am very thankful for. Thank you to Debora Botting for being supportive and committed to the Biochemistry and Molecular Biology department and its students. This thesis would not have been possible without the extra efforts of Dr. Cary Cooper, Dr. Dorian Coppenhaver, Dr. Norbert Herzog, Dr. Oberhauser, Dr. Sarita Sastry and Dr. Watowich. I am very fortunate for having the opportunity to get to know such remarkable professors during my time at the University of Texas Medical Branch. My final appreciation is extended to my friends and family for their unconditional love and support. I must greatly thank the central point in my life, my wife, Michelle Borgeson, for she is my constant guide and inspiration.

**Structural Investigation of 1,8-Dinitro-4,5-Dihydroxyanthraquinone
And Implications for Docking, as a Small Molecule Inhibitor, Into the
Protease of West Nile Virus**

Publication No. _____

Jeffrey Allen Borgeson, M.S.
The University of Texas Medical Branch, 2010

Supervisor: Stanley J. Watowich

Flaviviruses pose a global threat to human health and the development of a broad spectrum drug would decrease the disease burden. The small molecule, 1,8-dinitro-4,5-dihydroxyanthraquinone has shown to bind the proteases of the dengue and West Nile viruses while also reducing their titers in cell-based assays. Structure-based analog design will likely be performed in the near future to increase its activity. However, the binding mechanism and conformation of the lead is unknown. The virtual screen that discovered this inhibitor showed it as having an unorthodox bend in the middle of the anthraquinone structure upon binding the protease. Upon further investigation, tautomerism and a bent configuration may exist in the small molecule. The structure of the small molecule was investigated for structural significance upon binding the dengue and West Nile protease and to see how it affects virtual screening efforts when using AutoDock as the structure-based docking program. It can be concluded that a stable tautomer does not exist in our crystal and that the conformation portrays a slight binding nature which could allude to a bent structure in polar solvents. Co-crystallization of the protease and the small

molecule did not produce a crystal capable of solving the structure and virtual screening experiments would be virtually unaffected if the tautomer or bent structure was added to a small molecule database. The tautomer and bent structure may still provide slight differences in the binding affinity upon binding the West Nile NS2B-NS3 protease.

Table of Contents

Abstract.....	v
List of Tables.....	ix
List of Figures.....	x
 CHAPTER 1.....	 1
Introduction.....	1
1.1 EMERGING AND REMERGING FLAVIVIRUSES.....	1
1.2 NS2B-NS3 PROTEASE AS AN ANTIVIRAL TARGET IN DENGUE VIRUS AND WEST NILE VIRUS.....	5
1.3 INHIBITOR DISCOVERY FOR THE WEST NILE VIRUS AND DENGUE VIRUS.....	6
1.4 PROTON TAUTOMERISM	9
1.4.1 Small molecule inhibitors and structure based virtual screening.....	9
1.4.2 Anthraquinones.....	11
1.4.3 Virtual Screening.....	13
1.5 RATIONALE.....	14
1.5.1 Background problem.....	14
1.5.2 Central Hypothesis.....	16
1.5.3 The following aims were used to test this hypothesis.....	16
Aim 1: Structural Characterization of ARDP0006.....	16
Aim 2: Structural Characterization of ARDP0006 in complex with the NS2B-NS3pro of WNV.....	16
Aim 3: Determination of predicted docking poses of ARDP0006 in the in-silico, crystallized, tautomer form, to the WNV protease.....	17
 CHAPTER 2.....	 17
Methods.....	17
2.1 X-ray crystallography of ARDP0006.....	17
2.1.1 Crystal Method for ARDP0006.....	17

2.1.2 X-ray Detection.....	19
2.2 X-ray crystallography of the NS2B-NS3 protease in complex with ARDP0006.....	20
2.2.1 Expression and Purification of WNV S2B-NS3 protease.....	20
2.2.2 Protein and ARDP0006 preparation and crystal development.....	21
2.2.3 X-ray Detection.....	22
2.3 Structure Based Docking.....	22
2.3.1 Small molecule, ARDP0006 preparation.....	22
2.3.2 Docking with AutoDock 4.0.....	22
CHAPTER 3.....	25
Results.....	25
3.1 Structure Determination of ARDP0006 with X-ray crystallography.....	25
3.2 Structure Determination of WNV NS2B-NS3 protease in complex with ARDP0006.....	31
3.3 Structure Based Docking with AutoDock.....	33
CHAPTER 4.....	37
Discussion.....	37
4.1 Structure Determination of ARDP0006 with X-ray crystallography.....	37
4.1.1. X-ray crystal structures of anthraquinone derivatives.....	37
4.1.2. Hydrogen bonds induce resonance.....	38
4.1.3. Conformational implications deduced from the crystal structure.....	43
4.1.4. Tautomer implications deduced from the crystal structure.....	45
4.1.5. Implications for ARDP0006 in solution.....	47
4.2 Structure Determination of WNV NS2B-NS3 protease in complex with ARDP0006.....	48
4.3 Structure Based Docking with AutoDock.....	48
CHAPTER 5.....	41
Conclusions.....	52
Appendix.....	54
References.....	57
Vita.....	67

List of Tables

Table 3.1.1	Experimental Details: 1,8-dinitro-4,5-dihydroxyanthraquinone (ARDP0006).....	27
Table 3.1.2	Hydrogen Bonding of the ARDP0006 crystal.....	30
Table 3.3.1	Docking pose predictions among the three ARDP0006 structures.....	36
Table 4.1.1	X-ray crystal structures of anthraquinone derivatives chosen for comparisons...	38
Table 4.1.2	Anthraquinone comparisons.....	42
Table 4.2.1	Existing NS3 protease solved structures in the Protein Data Bank.....	48

List of Figures

Figure 1.1.1	Phylogenetic relationships of Flaviviruses.....	3
Figure 1.2.1	Diagram of the flavivirus polyprotein organization and processing.....	4
Figure 1.3.1	The bent structure of ARDP0006 disagrees with the x-ray structure.....	8
Figure 1.3.2	Possible binding sites of ARDP0006 that can interfere with the substrate-based inhibitor binding.....	7
Figure 1.4.1	Examples of small molecule tautomerism.....	11
Figure 1.4.2	Tautomers of anthraquinone variants.....	12
Figure 1.4.3	Proposed tautomerism of ARDP0006.....	13
Figure 2.1.1	Diagram of the experimental set-up for obtaining ARDP0006 crystals.....	18
Figure 3.1.1	ARDP0006 solved with x-ray crystallography.....	28
Figure 3.2.1	Progress towards x-ray determination of the protease-inhibitor complex.....	33
Figure 3.3.1	ARDP0006 molecules prepared for docking.....	35
Figure 3.3.2	Docking results.....	36
Figure 4.3.1	Intermolecular contacts made in docking results.....	50
Figure 4.3.2	Intermolecular contacts made in docking results for the tautomer structure.....	50

CHAPTER 1

Introduction

1.1 EMERGING AND REMERGING FLAVIVIRUSES

The dengue and West Nile viruses currently represent important emerging and reemerging infectious diseases and are currently NIH Category A and B Priority Pathogens, respectively. Dengue virus thrives in subtropical and tropical regions and exists in approximately 100 countries (NIH 2009), putting more than 2.5 billion people at risk (Gubler 2002a). Annually, dengue manifests nearly 500,000 cases of severe illness in the form of dengue hemorrhagic fever or dengue shock syndrome from the estimated 50-100 million that are infected (Gubler 2002 a, b, Halstead 1997). West Nile virus became the most widely distributed Flavivirus on the globe after its introduction in New York in 1999 and its expansion throughout most of the Americas (Hayes et al. 2005, Ebel et al. 2001, Laciotti et al. 1999). The enzootic cycle of birds and mosquitoes gave West Nile virus the ability to spread quickly from the US and Canada to the Central and South American countries in less than a decade (Bosch et al. 2007, Komar and Clark et al. 2006, Morales et al. 2006, Laciotti et al. 1999). Both viruses display a clear global health threat and are only two of many in the Flavivirus genus that are considered either currently or potentially pathogenic to humans.

The Flavivirus genus includes over seventy viruses, of which twenty-nine (Weissenböck et al. 2009, Dobler et al. 2009, Gould and Solomon et al. 2008) are known to inflict disease in humans. Some cause epidemics across the globe and can include high mortality rates that usually incorporate a form of encephalitis or hemorrhagic fever. Included are the virulent dengue, yellow fever, West Nile, Tick Borne encephalitis and Japanese encephalitis viruses. Most flaviviruses are considered arboviruses because they rely on mosquitoes or ticks to

complete their transmission cycle (Figure 1.1). Humans are only incidental hosts, as flaviviruses typically lack the ability to produce high enough titers to transmit from human to human. Dengue virus, however, is an exception since it lost the requirement for enzootic activity and can utilize humans as their reservoir and amplification hosts (Weaver and Reissen 2009). The potential for additional flaviviruses to evolve into infectious pathogens depends on many factors ranging from viral genetics to various epidemiological factors (Weaver and Reissen 2009). For instance, although West Nile and Japanese Encephalitis are both members of the Japanese encephalitis virus serocomplex (ICTVdB) (Figure 1.1.1), their geographical range and virulence have considerable differences. Although West Nile virus maintains a presence on every continent except Antarctica, it only develops into a severe neuroinvasive disease in less than one percent of all cases (Mostashari et al. 2001). Conversely, Japanese Encephalitis virus remains mostly anchored within the southeastern Asia region, while generating approximately a 33% mortality rate (Tsai 2000). Numerous flaviviruses have not yet been associated with human disease and remain segregated in their respective regions (Dobler et al. 2009, Weissenböck et al. 2009). However, their potential to emerge into threatening human pathogens is unknown and subject to change depending on how their genomes and epidemiological factors transform over time. Therefore, small molecules that inhibit both the dengue and West Nile viruses may improve future broad spectrum drug development efforts against flaviviruses.

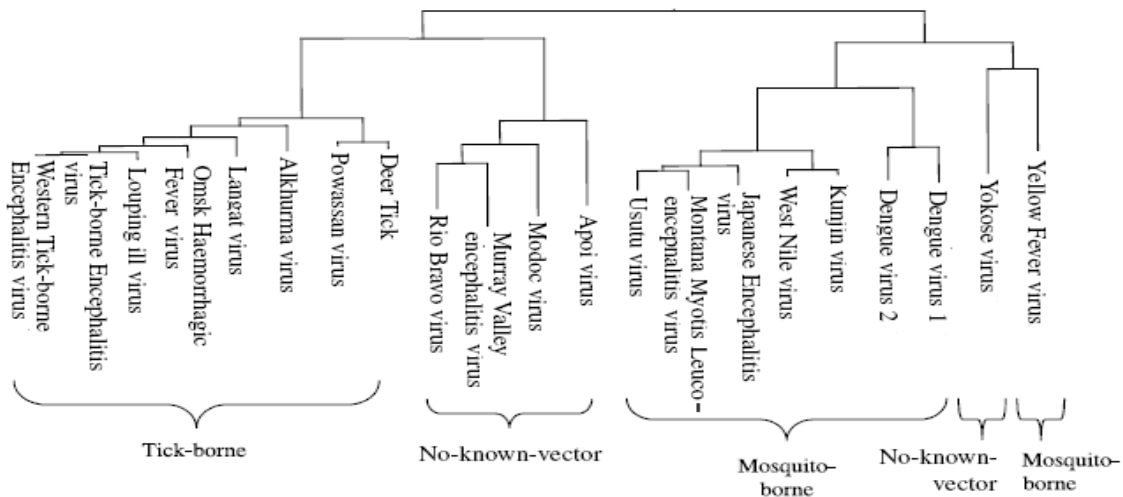


Figure 1.1.1. Phylogenetic relationships of the genus *Flavivirus* based on full genome sequences and vector classification. Adapted from (Cook and Holmes 2006).

Flaviviruses maintain a similar ~11 kB positive sense RNA genome that is packed into a nucleocapsid and ultimately surrounded by a virion consisting of an outer glycoprotein shell and an inner host derived lipid bilayer (Gollins and Porterfield 1986, Ishak et al. 1988). Viral entry is a three stage process: receptor binding, receptor-mediated endocytosis, and fusion of the viral and host cell membranes. Completion of the third step releases viral RNA into the cytoplasm where it mimics the host's mRNA and translates into a single polyprotein precursor by using the host's cellular machinery. The polyprotein weaves through the membrane of the ER Lumen leaving large amino acid sequences free on both sides. Host and viral proteases cleave the non-membrane bound polyprotein portions allowing for the formation of three structural (C-prM-E) and seven non-structural proteins (NS1-NS2A-NS2B-NS3-NS4A-NS4B-NS5) (Figure 1.2). The structural proteins eventually form part of the mature virion, whereas, the non-structural proteins conduct polyprotein processing, viral RNA synthesis, and viral morphogenesis (Murray et al. 2008, Lindenbach et al. 2007). Inhibiting these processes is possible with the design of small molecules that bind in strategic locations on the proteins essential for proliferation (Sampath and

Padmanabhan 2009). These proteins serve similar functions and maintain similar amino acid sequences among flaviviruses. Thus, broad spectrum antivirals may become achievable for flaviviruses (Ojwang et al. 2005, Shi et al. 2002).

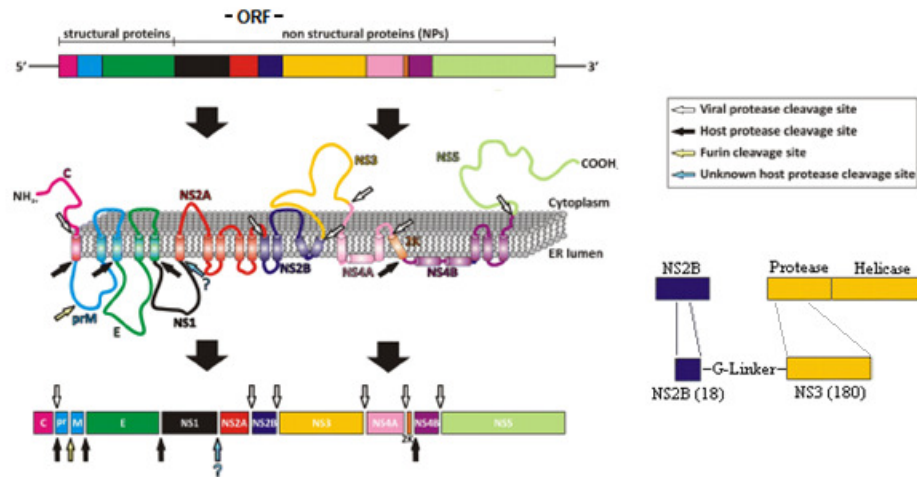


Figure 1.2.1. Diagram of the Flavivirus polyprotein organization and processing.

The top and bottom figure shows the linear organization of the structural and non-structural proteins within the open reading frame and as the cleaved products, respectively. The middle figure shows the polyprotein weaved through the ER lumen membrane. The arrows denote cleavage locations as predicted from numerous biochemical and cellular analyses. Lower right figure: The truncated, active form of NS2B-NS3pro is effectively studied by using only 18 residues of the NS2B and 180 residues of the NS3 protein attached by a glycine linker. Figure adapted from (Assenberg et al. 2009).

Despite the disease threat, effective treatments are lacking to combat flavivirus infections. Vaccines exist for YFV, JEV, and TBEV, except outbreaks continue to occur in developing countries (Stephenson 2006). Dengue virus and WNV vaccines are progressing but many years may pass before successful candidates emerge from clinical evaluation (Webster et al. 2009, Kramer et al. 2008, Martin et al. 2007, Monath et al. 2006). Dengue virus is especially challenging as it is currently believed that immunity of one dengue serotype may promote severe dengue disease when infected by a second serotype, a process known as the antibody-deficient enhancement of infection (Brown et al. 2009, Goncalvez et al. 2007, Vaughn et al. 2000). Antivirals are being pursued to inhibit important viral proteins and functions, but none have yet

been proven successful (Puig-Basagoiti et al. 2009, Li et al. 2009, Lescar et al. 2008, Dong et al. 2008, Malet et al. 2008).

1.2 NS2B-NS3 PROTEASE AS AN ANTIVIRAL TARGET IN DENGUE VIRUS AND WEST NILE VIRUS

Once cleaved from the polyprotein, the 618 amino acid NS3 protein folds into a protease domain on the N-terminal end accompanied by a helicase domain on the remaining C-terminal end (Assenberg et al. 2009). The 180 residue NS3 protease requires a central 40 residue portion of the NS2B polypeptide as a co-factor and remains active when separated from the helicase domain (Li et al. 2005, Leung et al. 2001, Yusof et al. 2000, Li et al. 1999, Arias et al. 1993, Falgout et al. 1993, 1991, Chambers et al. 1991). A genetically engineered Glycine⁴-Serine-Glycine⁴ linker between the C-terminal region of NS2B and the N-terminal of NS3 yields an active, recombinant NS2B-NS3 protease (NS2B-NS3pro) (Leung et al. 2001). The NS2B-NS3pro is identified as a trypsin-like serine protease and contains a catalytic triad composed of His⁵¹-Asp⁷⁵-Ser¹³⁵ (Bazan and Flatterick 1989, Gorbalenya et al. 1989). The heterodimer complex is necessary for viral proliferation as it cleaves at least five different polyprotein sites (Fig. 2): NS2A-NS2B-NS3-NS4A-NS4B-NS5 with a preference for Arg/Lys-Arg/Lys-Arg ↓ Gly/Ala at the P2, P1, and P1' sites ("↓" denotes the cleaved bond) (Shiryaev et al. 2007, Chambers et al. 2005, Chambers et al. 1993, Falgout et al. 1993, Zhang et al. 1992, Chambers et al. 1991, Falgout et al. 1991, Wengler et al. 1991, Chambers et al. 1990, Preugschat et al. 1990). Due to the important role played by the NS2B-NS3pro, it is as a key target for drug development (Leyssen et al. 2000). Phylogenetic clades were developed for the flavivirus NS3 protein based on cross neutralization and genetic analysis showing that the NS3pro shares structural similarities and is highly conserved among mosquito-borne flaviviruses with an amino acid

sequence identity of $\geq 70\%$ between DV, WNV, JEV, and YFV (Gaunt et al. 2001, Kuno et al. 1998, Ryan et al. 1998). Antiviral therapies against proteases in general proved successful for the HIV protease with the development of several approved inhibitors (Abbenante and Fairlie 2005, Wlodawer 2002, Leung et al. 2000). Additionally, inhibitors for the HCV serine protease have seen limited success (Chen et al. 2009, Abbenante and Fairlie 2005, Lamarre et al. 2003, Leung et al. 2000). For these reasons, NS3 protease has been acknowledged as one of the prime targets for antiviral development.

1.3 INHIBITOR DISCOVERY FOR THE WEST NILE VIRUS AND DENGUE VIRUS PROTEASE

Discovering protease inhibitors against flaviviruses is challenging and research efforts have included high-throughput screening, inhibitors designed to mimic peptide substrates, and ligand or structure based virtual screening (Tomlinson et al. 2008a, Yin et al. 2006 a,b, Knox et al. 2006, Chanprapaph et al. 2005, Ganesh 2005, Leung et al. 2001). A previous study utilized structure-based virtual screening as a method of drug discovery and generated many hits against the solved structure of the inactive Dengue virus, serotype 2 (DV2), NS3 protease. Two of those hits exhibited antiviral activity against DV2 (Tomlinson et al. 2008b) and WNV (unpublished data) in cell based assays. Furthermore, kinetic assays conducted with both small molecules confirmed μM inhibition levels against the WNV and DV2 proteases (unpublished data). These results are promising as these two compounds display broad spectrum inhibition among two flaviviruses. However, it is still unknown as to which intermolecular contacts are being made between the small molecules and the protease or what the binding mechanism is. A structural investigation of the small molecules bound to the protease can aid in elucidating such details and highlight intermolecular contacts important to the binding affinity. Additionally, when the co-

complex is compared to the structure of ARDP0006 in solution, conclusions can be made as to how the small molecule is binding based upon the small molecule's structural changes. These details can assist in future analog design and optimization by pointing out important intermolecular contacts being made to the protease. Structure determination has enhanced analog design and lead optimization for numerous targets including the viral proteases of HIV and another Flaviviridae virus, HCV (Njoroge et al. 2008, Mitsuya et al. 2007, Thaisrivongs and Strohbach 1999, Kim et al. 1995). Drug optimization is a slow process and includes thorough knowledge and characterization of the target's active site and optimal binding interactions, including: steric hinderances, electrostatic forces, hydrophobicity, and the free energy of binding. For example, a drug candidate, VX-478, of the HIV-1 protease was developed after the following interactions were optimized: binding to aspartate residues, binding to a critical water molecule, minimizing of steric hindrances, and optimizing hydrogen bonding (Kim et al. 1995). Optimization of protease inhibitors for DV and WNV will take much effort. However, we must correctly assess the structure of our lead before fully understanding: how it interacts with the NS2B-NS3 protease, what interactions are preferred by the protease, and how to optimize interactions.

Reinvestigation of the lead 1,8-Dinitro-4,5-dihydroxyanthraquinone (designated as ARDP0006) from the virtual screening raised some questions about its structure in the docked pose (figure 1.3.1). Virtual screening methods which allow for unacceptable small molecule conformations can lead to inadequate results. Additionally, tautomerism can theoretically exist in ARDP0006, which could alter the structure or change its binding properties. Fortunately, ARDP0006 was found to reduce WNV and DV titers in cell-based assays and bind the protease

in biochemical experiments. However, a further structural description of ARDP0006 is needed to compare with the virtual screening result.

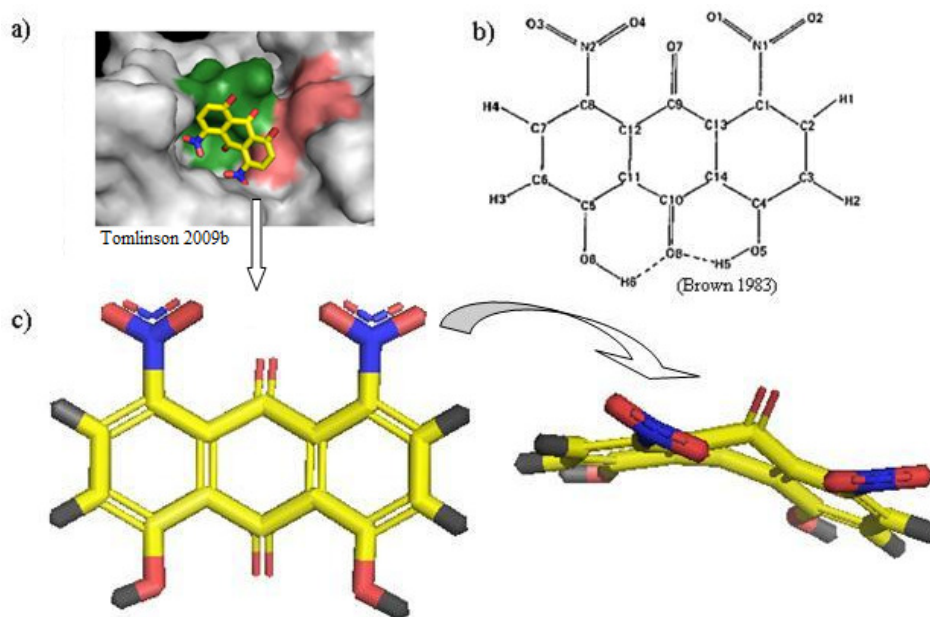


Figure 1.3.1. The bent structure of ARDP0006 disagrees with the x-ray determined structure.

a) The predicted docking pose of ARDP0006 in the active site of the DV2 NS3pro (Tomlinson et al. 2009b). b) The structure of ARDP0006 was previously solved via x-ray crystallography. The mean atomic deviation from the plane of the anthraquinone nucleus (C1-C14) is 0.028 Å. (Brown and Colclough 1983). c) The docked ARDP0006 shows a bent structure. Figures in c) by (PyMOL 2005). The docked result of ARDP0006 in c) was provided by Suzanne Tomlinson in PDB format.

Despite ARDP0006's structural uncertainties, a structural model, generated by a preliminary docking study, was created to envision possible binding sites for ARDP0006 (figure 1.3.2). There are three important subsite pockets (S1, S2, S3) in the active site and, when those subsites are mapped onto the 3D structure of a WNV NS2B-NS3 protease, we can visualize possible binding poses in which ARDP0006 would interfere with the binding of substrate-based inhibitors. The substrate-based inhibitor shown in complex with the protease (figure 1.3.2) is the result of a co-crystal structure previously solved (Erbel et al. in 2006, PDB: 2FP7). The substrate-based inhibitor contains the preferred three residues, Lys-Arg-Arg, that typically

proceeds the cleaving bond and, therefore, mimics naturally binding peptides as the P3, P2, and P1 residues, respectively (Shiryaev et al. 2007, Erbel et al. 2006, Knox et al. 2006).

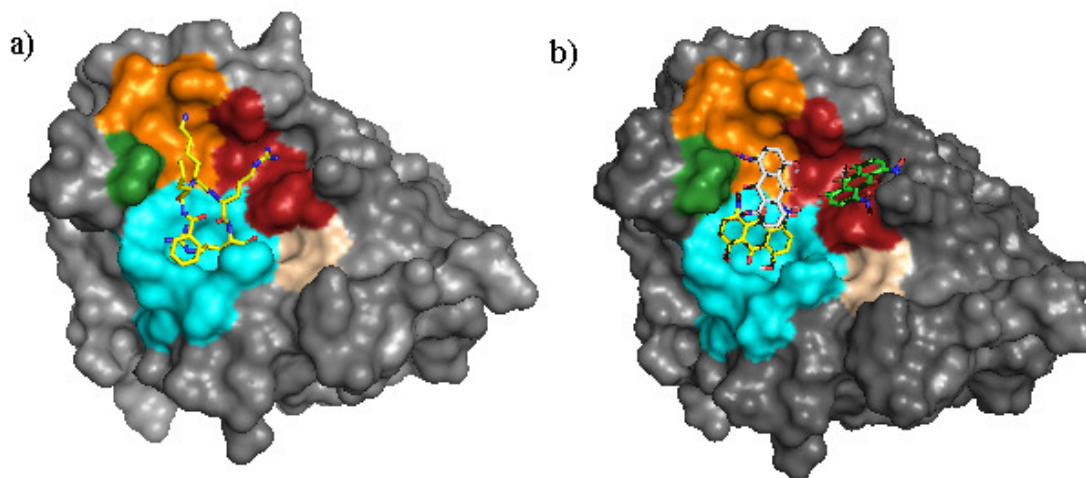


Figure 1.3.2. Possible binding poses of ARDP0006 that can interfere with the binding of the substrate-based inhibitor. a) The crystal structure of WNV NS2B-NS3pro was co-crystallized with a substrate based inhibitor that includes the P3, P2, and P1 residues typically recognized by the active site (Erbel et al. 2006). The substrate-based inhibitor represents the possible binding pose of naturally occurring, cleavable peptides. c) Three ARDP0006 small molecules were modeled into the active site at three possible binding subsites (S1-Cyan, S2-Red, S3-Orange) to show how it can bind in several locations and still interfere with the substrate-based inhibitor, therefore creating small molecule inhibition. The subsites, S4-Green and S1'-Peach are additional subsites in the protease. The surface residues associated with indicated subsites can be referenced to (Knox et al. 2006, Chappelle et al. 2005 Nall et al. 2004). Figures were created in (PyMOL 2005).

1.4 PROTON TAUTOMERISM

1.4.1. Small molecule inhibitors and structure based virtual screening

Small molecules with proton tautomerism maintain an equilibrium between isomers that are intrerconvertable in which hydrogens have the ability to bond with other atoms within the molecular structure (IUPAC 1997). In general, tautomers are capable of altering many properties of the molecule including hydrophobicity, pKa (Jang et al. 2002), overall structure, absorbance spectra (Ezaby et al. 1970), molecular fingerprinting (Oellien et al. 2006), and

electrostatics (Martin 2009). Mobile hydrogens create a shift of proton donors and acceptors within a molecule and can change the shape of the molecule by altering various properties of the involved bonds. Tautomer ratios in solution may flux significantly as changes in the solvent environment are made such as pH, concentration, and temperature. More importantly, proteins can prefer to bind a particular tautomer with differing strengths, thus creating concerns for small molecule inhibitor development and virtual screening. For example, the matrix metalloproteinase 8 binds to the minor isomer of the tautomer equilibrium (figure 1.4.1a) (Brandstetter et al. 2001); and a macrophage inhibitory factor acquires a separate affinity for each molecular tautomer (figure 1.4.1b) (Taylor et al. 1999).

Recent evaluations of potential tautomers in molecules within various databases have concluded that roughly 25-30% of small molecule tautomers are not included (Martin 2009, Millettii et al. 2009, Pilot 2008, Proudfoot 2005). Compound libraries that ignore tautomeric variants may cause potential leads to be overlooked in virtual screening studies. One study included 2,949 marked drugs estimated an increase of the dataset by 1.64-fold when considering each tautomer as a separate compound (Proudfoot 2005). The benefit of adding each tautomer to a molecular database in order to enhance results may not outweigh the cost of additional time required to conduct such a virtual screen unless the tautomers are proven to significantly improve screening results. In a small study, AutoDock predicted nearly identical binding energies, but separate poses for two proton tautomers of zanamivir, a neuraminidase inhibitor (Kalliokoski et al. 2009). With further evaluation, Kalliokoski et al. concluded that the time cost of considering each tautomer in structure based virtual-screening with AutoDock outweighs the benefits and suggested the best predicted form of a molecule provides adequate results.

Numerous cases like those displayed in figure 1.4.1 have given reason to further evaluate tautomerism and its effects on virtual screening (Kalliokoski et al. 2009, Oellien et al. 2006, Knox et al. 2005, Pospisil et al. 2003) and has prompted the development of new software and methodologies for tautomer enumeration and prediction (LigPrep, QUAKPAC, TAUTOMER, Milletti et al. 2009, Kalliokoski et al. 2009). Since it is unknown if ARDP0006 binds the NS2B-NS3 protease as a tautomer, we conducted a docking experiment to observe if the tautomer would produce different docking results.

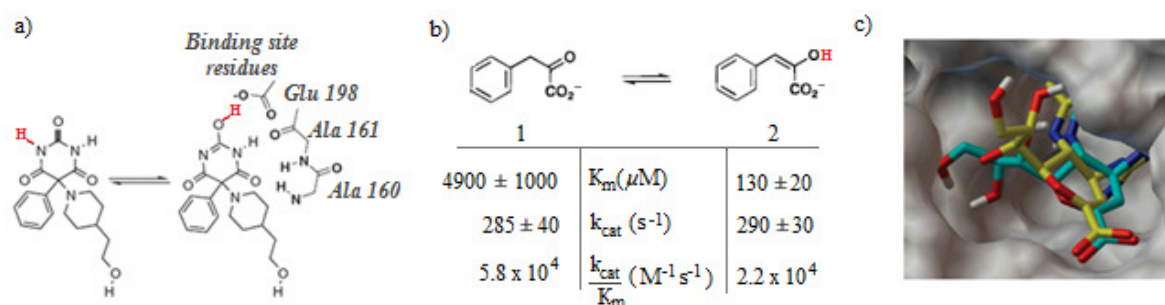


Figure 1.4.1. Examples of small molecule tautomers binding and docking differently: a) Matrix metalloproteinase 8 recognizes the minor conformation of the tautomer equilibrium (Brandstetter et al. 2001); b) Two tautomers recognized by macrophage inhibitory factor bind with differing affinities (Taylor et al. 1999); c) Two proton tautomers of zanamivir, a neuraminidase inhibitor, cause AutoDock to generate different poses while the predicted binding energies are nearly identical (-8.96 and -8.97 kcal/mol) (Kallioski et al. 2009).

1.4.2. Anthraquinones

Anthraquinones are a class of natural molecules with numerous derivatives, and cause several biological effects including antiviral (Hsiang et al. 2008, Shuangsoo et al. 2006, Semple et al. 2001), anticancer (Bhasin et al. 2008, Xio et al. 2008, Wu 2003), anti-inflammatory (Park et al. 2009), immunosuppressant (Huang et al. 1992), and vasorelaxant (Huang et al. 1991) activity. Due to the many applications of anthraquinone molecules, several structural derivatives have been determined and a few tautomer studies were conducted by utilizing x-ray crystallography, NMR, and spectroscopy methods (Ismail et al. 2009, Ng et al. 2005, Zain et al. 2005, Wu et al.

2003, Marzocchi et al. 2003, Yatsenko et al. 1997 & 1996, Fariña et al. 1992, Ulicky et al. 1987, Smulevich et al. 1987, Ezaby et al. 1970). Anthraquinone tautomerism influences absorbance spectra and spectroscopy methods have detected a dependency on the molecule's substituents, solvent polarity, and solvent pH (Yatsenko et al. 1997, Ezaby et al. 1970). However, precise determination of atomic positions and tautomer equilibrium is difficult with spectroscopy. With the use of x-ray crystallography, the 1,10 [Ia] and 1,4 [Ib] tautomers of Mordant Black 13 dye-free base and its butyl derivative were found in near equal amounts (Tafeenko et al. 1991 & 1987), and the 1,10 [IIa] and 9-imino, 10 tautomers [IIb] of an anthraquinone derivative were structurally determined (figure 1.4.2) (Yatsenko et al. 1997). Structural determination of these particular molecules reveals that tautomers can exist in crystal form in hydroxyanthraquinones similar to ARDP0006 in a stable conformation.

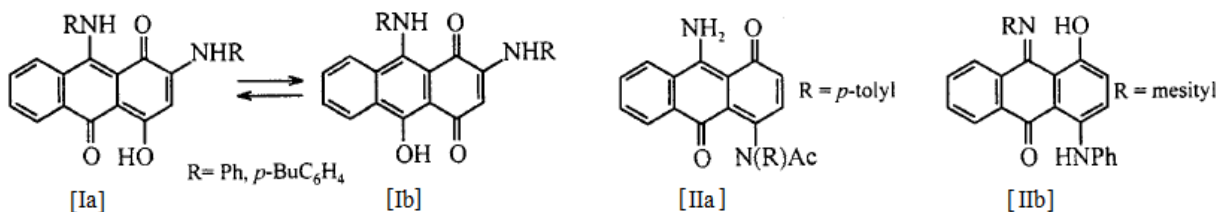


Figure 1.4.2. Tautomers of anthraquinone variants were determined with x-ray crystallography: 4-hydroxy-2,9-diphenylamino-1,10-anthraquinone [Ia], 4-hydroxy-2,9-di(4-n-butylphenyl) amino-1,10-anthraquinone [Ib] (Tafeenko et al. 1991, 1987), 4-(N-acetyl-*p*-tolylamino)-9-amino-1, 10-anthracenedione [IIa], 4-hydroxy-1-phenylamino-10-mesitylimino-9(10H)-anthracenone [IIb] (Yatsenko et al. 1997).

Although tautomerism is unknown to exist in ARDP0006, a mobile hydrogen has potential to have an effect on its overall structure. Although improbable due to pi bonding characteristics of anthraquinone's apparent three ring conjugated system, the bent structure has been considered for two reasons. Firstly, the two carbonyl carbons are not designated to include double bonds with their connecting carbons. This allows for the possibility of the anthraquinone

molecule to bend due to the lack of a resonance indication in the center ring. In literature, if the aromaticity is addressed, only the two outside rings are designated as aromatic structures (Ismail et al. 2009 a, b, Zain et al. 2005, Ulicky et al. 1987). Secondly, an unusual and unexplained case exists where an anthraquinone base is known to have a dihedral angle between the two outer benzene rings of 22.02° (Cao et al. 2007). The tautomer is important because, as indicated in the modeled tautomer (figure 1.4.3b), it would cause double bonds to shift and would create a more convincing conjugated bond system for the bottom portion of the central ring. Minor physical changes, like the one displayed in figure 1.4.3, which involves tautomerism and a bent anthraquinone conformation, can affect the molecule's binding affinity and mechanism with the protease and needs to be investigated.

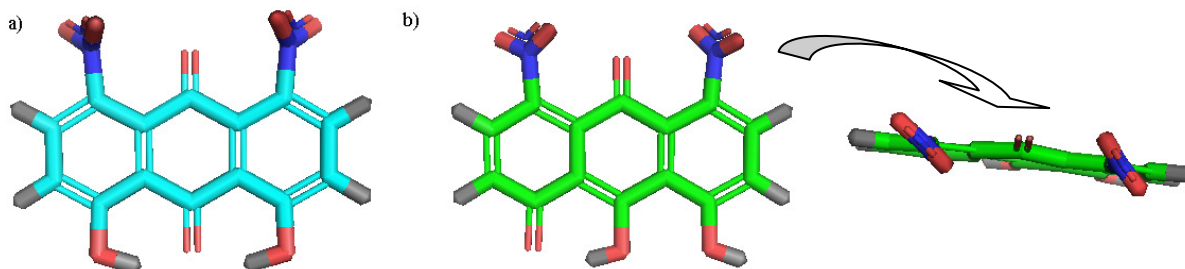


Figure 1.4.3. Possible planarity influences from tautomerization. a) displays the conventional conformation for ARDP0006. b) displays the tautomer conformation with the double bonds oriented in such a way that might enforce planarity in the middle ring. Figures modeled in (PvMOL 2005).

1.4.3 Virtual Screening

Structure based-docking is a tool that has been utilized extensively in coordination with virtual screening for drug discovery for the last two decades (Kitchen et al. 2004, Bajorath 2002, Gohlke and Klebe 2002, Langer and Hoffmann 2002, Walters et al. 1998). Evaluations of several docking programs have concluded success rates ranging 35%-70% for correctly docked poses (depending on the evaluation method) and a lack of accurate scoring functions (Cheng et al. 2009, Warren et al. 2006). However, this study concentrates on the docking pose

comparisons of the various ARDP0006 structure variants, not necessarily the accuracy of a correctly predicted pose. Docking as a methodology for the binding predictions of small molecules to protein sites has come a long way since it was pioneered in the early 1980's (Kuntz et al. 1982), and the importance of tautomerism may prove important to the future of virtual screening. This small docking study serves as our method for determining the importance of tautomerism when docking ARDP0006, and if the tautomeric form of ARDP0006 should be included in small molecule databases. Comparisons are made between the docking poses generated from three different conformations of ARDP0006: the crystal structure, tautomer modeled based on the crystal structure, computationally designed ARDP0006.

We found it useful to dock with the ARDP0006 compound found in a database of compounds in order to show the effect different ARDP0006 structures had when virtual screening uses a database, but ARDP0006 was not found in the database that our lab has access to. Therefore, ARDP0006 was designed in PyMOL (Delano 2002) and minimized in CHARMM (Brooks et al. 2009). This still represents a compound purely designed computationally.

AutoDock 4.0 (Huey et al. 2007, 2004) was used for this docking study. It is based on a semi-empirical free energy force field, and uses a Lamarckian Genetic Algorithm (LGA). It is widely cited in over 250 citations (as of November 15, 2009) and played an integral role in the discovery of an approved drug to treat HIV integrase (Rao et al. 2002).

1.5 RATIONALE

1.5.1 Background problem

The small molecule 1,8-dinitro-4,5-anthraquinone, ARDP0006, was generated from a previous virtual screening project and has *in-vitro* and *ex-vivo* activity against the dengue and West Nile viruses (Tomlinson et al. 2008b & unpublished data). Future experiments may entail

optimization and analog design which can be enhanced significantly by comparing the structural details of ARDP0006 in solution and in complex with the viral proteases. However, some structural uncertainties of unbound ARDP0006 are present and needed to be investigated in further detail. For example, there is the possibility of the protease preferring to bind a particular tautomer. If the structure of ARDP0006 is different than the assumed structure (figure 1.4.3a) while in solution, but binds the protease in the assumed form, we would be overlooking possible binding mechanisms that will be useful in optimizing the lead. Once the structure of ARDP0006 is known when in solution, and when bound in complex with the protease, we can begin to establish a binding mechanism.

The computationally designed and docked model of ARDP0006 has two notable differences compared to the x-ray crystallized solved structure (Brown and Colclough 1983). The previously docked computationally designed structure is unconventionally bent at its center and both hydroxyl hydrogens are rotated outward. According to the crystal structure, the mean atomic deviation from the plane of the anthraquinone nucleus (C1-C14) is a mere 0.028 Å, and the hydrogens are hydrogen bonded to the central carbonyl oxygen (Chapkanov 2007, Brown and Colclough 1983). The structural differences impede our ability in understanding how it binds to the protease (figure 1.3.1) and raises the question as to how accurate the docking is for this molecule. Angstrom level resolution of a co-crystal complex will most likely not elucidate the atomic details well enough to determine which tautomer the protease binds to if a tautomer is possible, so the structural details of ARDP0006 need to be elucidated to better understand the anthraquinone tautomerism possibilities between the hydroxyl and carbonyl at the angstrom level. Furthermore, since there is a possibility that the protease prefers to bind a tautomer structure of

ARDP0006 (figure 1.4.3b), an evaluation should be made as to if the structure based docking results would be significantly enhanced or changed when including the tautomer.

1.5.2. Central hypothesis

Based on the evidence that tautomers can change the binding properties of small molecules and that tautomers exist for other anthraquinone derivatives, we hypothesized that the crystallized and tautomer form ARDP0006 would produce different predicted intermolecular contacts than the computationally designed molecule when docked to the NS2B-NS3 protease of the West Nile virus.

1.5.3. The following aims were used to test this hypothesis.

Aim 1: Structural Characterization of ARDP0006

X-ray crystallography was utilized to solve the structure of the lead, ARDP0006. Although the structure was previously solved, a more thorough description was needed. The atomic details of our structure are supplemented with an assessment of other structures, thus producing conclusions on the form of ARDP0006 in solution. X-ray crystallography was used for because it was found useful in finding a stable tautomer in two similar anthraquinone structures (Yatsenko et al. 1997, Trafeenko et al. 1987)

Aim 2: Structural Characterization of ARDP0006 targeting NS2B-NS3pro of WNV

X-ray crystallography was attempted to solve the structure of the WNV NS2B-NS3 protease bound to the lead, ARDP0006. This aim was constructed to reveal where the leads bind on the protease and highlight the important interactions between the compound and the protease.

Aim 3: Determination of predicted docking pose differences of ARDP0006 in the in-silico, crystallized, and tautomer form, to the WNV protease

AutoDock was utilized to conduct docking experiments in order to predict binding poses of three ARDP0006 structures in the crystal structure of the NS2B-NS3 protease of WNV. This aim addresses the issue: Do the slight 3D structural differences of the crystal form, tautomer form, and *in-silico* form of ARDP0006, produce different docking predictions? *In-silico* ARDP0006 refers to the computationally designed 3D structure.

CHAPTER 2

Methods

2.1 X-RAY CRYSTALLOGRAPHY OF ARDP0006

2.1.1 Crystal Method for ARDP0006.

The compound, 1,8-Dinitro-4,5-dihydroxyanthraquinone, was commercially purchased in powder form from Sigma Aldrich® (www.sigmaaldrich.com; CAS: 81-55-0). The compound, like many anthraquinone derivatives, is only soluble at low concentrations in water. Therefore, 9.90 mg were mixed in 100% DMSO to create a 1 mL solution at a concentration of 30 mM. A second solution was created by mixing 100 μ L of the previous solution with 100 μ L of DMSO in order to create a solution with 100% DMSO and 15 mM ARDP0006. A third solution was created by mixing 100 μ L of the first solution with 100 μ L of ddH₂O in order to create a solution with 50% DMSO and 15 mM ARDP0006. Twenty-five μ L drops of each of the three ARDP0006 solutions were centrally positioned on glass cover slips and placed in a non-sealed, but covered, container along with 20 mL of 100% EtOH in a small, uncovered beaker. The drops were allowed to dehydrate at room temperature, overnight for about 16 hours while the EtOH aided in the dehydration process.

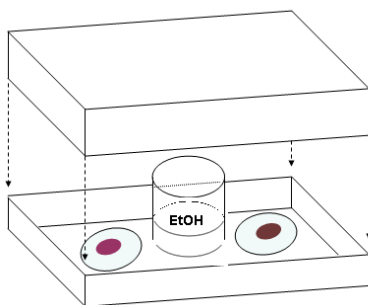


Figure 2.1.1. Diagram of the experimental set-up for obtaining ARDP0006 crystals.

2.1.2 X-ray Detection.

Promising crystals were submerged in paraffin oil with the use of a cryoloop and mounted in line with a MacScience MO6XHF (1.2 kW Cu 100 micron ultra fine focus) x-ray generator equipped with a Bunker SMART 2K CCD detector and a Cryo Industries of America CRYOCOOLER which is located at UTMB's Sealy Center for Structural Biology and Molecular Biophysics X-ray Crystallography Core facility. X-ray diffraction was performed at 105 K with the use of liquid nitrogen. Eight data collection runs were performed on the crystal of choice. The four high-resolution runs were performed at three different phi angles, 2θ of 100° , ω at 100° , an ω step of 0.5° , exposure times of 90 seconds, and a total of 360 frames per run. The four low-resolution runs were performed at four different phi angles, 2θ of 100° , ω at 50° , an ω step of 0.5° , exposure time of 60 seconds, and a total of 250 frames per run. Data collection was processed with PROTEUM (Brunker 2002), cell refinement on SAINT-Plus (Brunker 2001), and data reduction by SAINT-Plus. The program used to solve the structure via direct methods and to refine the structure was SHELX97 (Sheldrick 1997). Three refinements were performed including isotropic and anisotropic determination and an absorbance correction estimated by WinGX. All six hydrogens were located after the first refinement and were refined during the second. ORTEP-3 (Farriugia 1997) was used for molecular graphics and thermal plotting. WinGX was used for file and program organization and coordination during and after the structure solving process (Farrugia 1999).

2.2 X-RAY CRYSTALLOGRAPHY OF THE NS2B-NS3 PROTEASE IN COMPLEX WITH ARDP0006

2.2.1 Expression and Purification of WNV S2B-NS3 protease.

Frozen glycerol stocks of XL1-blue competent cells containing the correct plasmids were obtained from Suzanne Tomlinson and stored in -80°C. The plasmid construct for the West Nile NS2B-NS3 protease was originally obtained from Dr. Padmanabhan's laboratory. Expression and purification of both proteases were based on a previously described protocol (Yusof et al. 2000). Two test tubes containing 10 mL of 2xYT media and 50 µg/mL carbenicillin were inoculated with the glycerol stock, grown overnight at 37°C, and used to inoculate 2 L of 2xYT media with 50 µg/mL. The cultures were grown for nearly 4 hours at 37°C and 250 rpm until OD₆₀₀ is 0.7-0.8 AU, induced with 0.5 mM IPTG, and incubated for an additional 4 hours at 37°C and 250 rpm. The bacterial cultures were then centrifuged for 30 minutes at 3000 rpm, and the pellet is stored in -80°C.

Purification begins by resuspending the pellet in 10 mL of buffer A (500 mM NaCl and 50 mM HEPES-HCl [pH 7.0]), adding the lysis components (300 µg/mL lysozyme, 30 µg/mL DNase, and 10mM MgCl₂), and stirring the mixture at 4°C for 30 minutes. Triton-X 100 was added to a final concentration of 0.5% (v/v) and the mixture was stirred again at 4°C for 30 minutes to complete the lysis step. The solution was centrifuged for 20 minutes and 15,000 rpm at 4°C. The soluble fraction was applied to a nickel affinity column (1 mL of Nickel Sephadex beads were purchased from Amersham Biosciences - <http://www6.gelifesciences.com/>) pre-equilibrated with buffer A. The contaminating proteins were washed with 20 mL buffer A. Four additional buffers were made to complete the purification. Those four buffers were made with the contents of buffer A and increasing concentrations of imidazole (10 mM, 20 mM, 30 mM,

and 150 mM). In order from lowest concentration of imidazole to highest, 20 mL were poured in a stepwise fashion over the column to further elute contaminating proteins. The buffer A-150 mM Imidazole solution was then used to elute NS2B-NS3pro in 10 individual, 1 mL aliquots. The least contaminated fractions were pooled (analyzed by 2D SDS-PAGE) and concentrated to ~1-2 mL while also performing buffer exchange into buffer B (crystallization buffer: 50 mM NaCl, 10 mM HEPES-HCL [pH 7.5], and 10% glycerol (v/v)). The protein solution was then applied to a Biosep-SEC 3000 gel-filtration column (Phenomenex® - <http://www.phenomenex.com/>) with a 0.2-0.5 mL injection per run in buffer B. The peak fractions were then pooled and concentrated to the desired concentration for crystallization. The protein solutions were concentrated with Millipore's Amicon Ultra-15 centrifugal filter devices and the concentration was determined by UV spectrometry.

2.2.2 Protein and ARDP0006 preparation and crystal development.

A stock solution of ARDP0006 9.90 mg were mixed in 100% DMSO to create a 1mL stock solution at a concentration of 30 mM. After protein purification, the WNV NS2B-NS3 protease was concentrated to 30 mg/mL. A mixture of 400 μ L protein, 60 μ L inhibitor solution, and 140 μ L of crystalline buffer yielded a 20 mg/mL protein/3mM inhibitor solution. The solutions were kept in 4°C until crystal plates were created. First, large crystallography screens were performed at room temperature on the Phoenix Rigaku Edition crystallography robot owned by UTMB's Sealy Center for Structural Biology and Molecular Biophysics. We screened using 96-well plates, hanging drop method, and 0.4 μ L drops (0.2 μ L of protein solution plus 0.2 μ L of well solution). Crystals developed after 2-4 weeks. All promising conditions were optimized and scaled-up to 4 μ L drops in 24-well plates utilizing the hanging drop method. The plates were immediately stored in 4°C.

2.2.3 X-ray Detection.

The one crystal suitable for collecting x-ray data formed in one of the large screens with the 0.4 μ L drops (see results for details). The small crystal was mounted in paraffin oil and flash-frozen under a cryostream of nitrogen gas in line with the x-ray generator. X-ray diffraction was performed at 100 K on a MacScience MO6XHF (1.2 kW Cu 100 micron ultra fine focus) X-ray generator and a MacScience DIP2030H-VLM dual 30 cm diameter imaging-plate detector located at UTMB's Sealy Center for Structural Biology and Molecular Biophysics X-ray Crystallography Core facility. This system is suitable for protein crystals of cell dimensions 30-400 Å, can diffract up to 1.3 Å, and is equipped with a Cryo Industries of America CRYOCOOLER. A single frame, 22 hours exposure was taken in order to determine how well the crystal diffracted. Although some diffraction was observed, the data was too weak to continue.

2.3 STRUCTURE BASED DOCKING

2.3.1 Small molecule, ARDP0006 preparation.

An output PDB file was created from the ARDP0006 structure solved and refined by x-ray crystallography. From that structure, the tautomer structure was created simply from moving the hydrogen from O6 to the double bonded oxygen, O8 by using PyMOL (Delano 2002). Bond lengths change when double and single bond characteristics are altered, so those bond lengths were corrected according to (Allen et al. 1987). The *in-silico* conformation was designed in PyMOL (Delano 2002) and minimized in CHARMM (Brooks et al. 2009).

2.3.2 Docking with AutoDock 4.0.

The protein and small molecules were prepared for docking with AutoDock 4.0 while utilizing AutoDockTools 1.5.4. (Sanner 1999). Three separate docking experiments were

prepared and performed while each had its own respective directory: the protease with ARDP0006-crystal solved, the protease with ARDP0006-tautomer, and the protease with ARDP0006-computationally designed and minimized. The PDB used for docking was 2FP7. It represents the 3-D crystal structure of an active WNV NS2B-NS3 protease bound to a peptide inhibitor and was determined by x-ray crystallography (Erbel et al. 2006). PyMOL was used to remove the peptide substrate from the protease prior to docking experiments. PyMOL was also used to ambiguously place each of the ARDP0006 molecules (crystal structure, tautomer, and *in-silico* designed molecule) near the S1 pocket of the protease. Each ARDP0006 molecule was saved in its own separate PDB file and without the protease. The protease was then opened in AutoDock to prepare the molecule for docking: all hydrogens were added to the molecule and each amino acid was thoroughly inspected to eliminate errors (i.e. missing atoms, only one conformation exists, etc). The PDB file was then saved. The same steps were performed for each small molecule PDB file. The PDBQT files were created for the protease and each small molecule. The grid parameter files were created without allowing for any flexible residues. A large grid box was created which included all four subsite pocket as indicated in figure 1.3.2. AutoGrid was run separately for each ARDP0006 molecule with the protease. Upon successful AutoGrid completion, the docking files were prepared and repeated for each molecule. Default values were used for docking parameters except the Genetic Algorithm was increased to 50 and energy evaluations per the Genetic Algorithm run were increased to 1,000,000; which is much more intensive than the default values of 10 GA runs and 250,000 energy evaluations. Upon completion, 50 conformations were provided and grouped in clusters divided by predicted by energies and predicted pose similarities. The largest clusters coincidentally were the clusters that contained poses with the lowest predicted energies. This made choosing conformations simple.

Chosen conformations had the lowest predicted energy and were within the largest cluster size. The chosen conformations were written in PDB files and input into the LigPlot and LigEd programs (Wallace et al. 1995) along with the protease PDB file in order to help identify intermolecular contacts. The PDB files were also used in PyMOL to prepare figures that visualize the intermolecular contacts for the small molecule conformations. This ultimately displays bonding differences among the chosen conformations with the protease.

CHAPTER 3

Results

3.1 STRUCTURE DETERMINATION OF ARDP0006 WITH X-RAY CRYSTALLOGRAPHY

Many purple crystals grew from each of the three solutions (30 mM ARDP0006 in 100% DMSO, 15 mM ARDP0006 in 100% DMSO, and 15 mM ARDP0005 in 50:50 DMSO:ddH₂O), but most were thin, splintering rods. The 50% DMSO solution yielded a rectangular prism, rod-like purple crystal (450 x 70 x 50 μm), which was mounted in paraffin oil and flash frozen for data collection. Over 48 hours of data collection yielded 9976 reflections, 100% completeness and 0.83 Å resolution. Further details are shown in table 3.1.1.

The crystal packing was complex and assembled with various interactions (see figure 3.1.1). Four parallel molecules were stacked in three sheets with two molecules in the center layer sitting planar with each other. The top and bottom molecules were approximately 3.25(2) Å from the center sheet. The stacking alignment of the aromatic rings was staggered such that the rings asymmetrically overlapped, creating weak π - π interactions. The O8 and O6 atoms of one ARDP0006 molecule were centrally located beneath two adjoining rings of another anthraquinone molecule with an average distance to the six ring carbons of 3.575(2) Å. This can be referred to as anionic- π interactions in a parallel-displayed structure (Mooibroek et al. 2008). This is possible in π -electron depleted ring structures, which is characteristic of ARDP0006, and will be validated in the discussion section. The block of four molecules was rotated and repeated to create the crystal (figure 3.1.1a). Non-classical hydrogen bonds were formed between the carbon hydrogens, H4, H5, H6, H3 and the nitro oxygens, O4, O3, O1, and the

carbonyl oxygen, O7, respectfully (see details in figure 3.1.1 and table 3.1.2). The resonance nature of the nitro substituents cause both oxygens to remain partially negative and permitted them to bind with the carbon hydrogens. These interactions provided some influence on the final positions of the nitro groups which resulted in the nitro torsion angles to the anthraquinone base to be $83.3(3)^{\circ}$ (C7-C8-N2-O4) and $123.8(3)^{\circ}$ (C13-C1-N1-O2). Furthermore, the O6H1 and O5H2 hydroxyl groups intermolecularly interacted with O6 and O2 in nearby molecules, respectfully, but interactions were considered very weak because H1 and H2 were strongly hydrogen bonded intramolecularly with O8. This was evidenced by their positions and atomic displacement values shown in figure 3.3.1c (parameter details are located in tables 1 and 2 of the Appendix). The shortness of the hydrogen bonds (1.822\AA and 1.807\AA) is indicative of their strength and consequently prevented O8 from developing significant intermolecular contacts within the crystal.

Investigation of the ARDP0006 atomic structure revealed many bond distances matched their respective common values, but some were not within their range of expected lengths. Bond distances were compared with an extensive collection of x-ray and neutron diffraction results that had been extracted from the Cambridge Structural Database (CSD) (Allen et al. 1987). It has been widely referenced with over 500 citations. In ARDP0006, the carbonyl C9=O7 bond distance of $1.212(3)\text{\AA}$ complies with typical $Csp^2=O$ lengths and common benzequinoids, but the C10=O8 bond was abnormally long with a length of $1.241(2)\text{\AA}$. The C-C bonds that created the two outer benzene-like rings of the anthraquinone nucleus ranged from $1.364(3)\text{\AA}$ to $1.406(3)\text{\AA}$ and were within expected bond ranges for anthracenes (1.360\AA - 1.440\AA) (Allen et al. 1987). The two C-C bonds connecting the carbonyl carbon to the outer rings on the nitro side of ARDP0006 (C13-C9-C12) were notably much longer for benzenoid dimensions at $1.478(3)\text{\AA}$

Table 3.1.1. Experimental Details: 1,8-dinitro-4,5-dihydroxyanthraquinone (ARDP0006)

	Our Solved Structure	(Brown and Colclough 1983)
<i>Crystal Data</i>		
Chemical Formula	C ₁₄ H ₆ N ₂ O ₈	
M _r	330.21	-
Cell Setting, Space Group	Monoclinic, P21/c	Monoclinic, P21/a
a,b,c (Å)	6.630, 11.970, 15.450	15.664, 12.056, 6.704
β (°)	35.30	34.20
V (Å ³)	1220.83	1262.62
Z	4	4
Radiation Type	Cu 100 Kα	Cu 100 Kα
μ (mm ⁻¹)	1.324	1.29
Temp (K)	105	-
Crystal Form, color	Prism, Purple	Prism, Purple
Crystal Size (mm)	0.450 x 0.070 x 0.050	2.0 x 0.2 x 0.2
Solution grown in	50% DMSO	phenol
<i>Data Collection</i>		
Diffractometer	Brunker SMART 2K CCD	Weissenberg Camera
Reflections:		
Measured, Independent,	9976, 2136	2316, 171
Reflections with:	Fo > 4σ(Fo): 2110	I > 3σ(I): 1862
Rmerge	0.0349	0.044
Resolution	0.83 Å	-
<i>Refinement</i>		
Method	Direct Methods. Least Square Calculations	Patterson Synthesis. Least Square Calculations
Robserved	0.0543	0.058
wR(F ²)	0.1518	0.5σ
S	1.0680	-
Reflections used	2110	1862

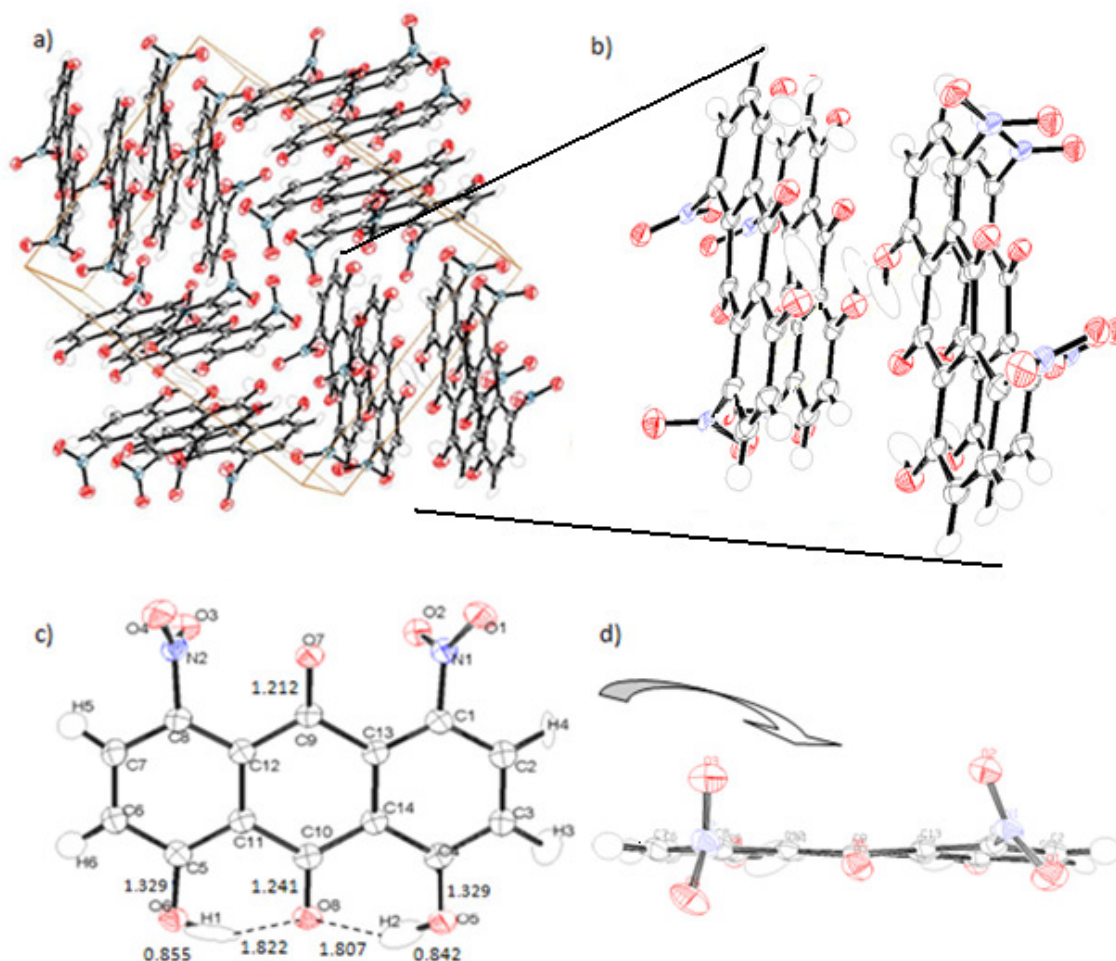


Figure 3.1.1. ARDP0006 was solved with x-ray crystallography. a) and b) show the crystal packing of the crystal where b) shows four parallel molecules stacked in a staggered fashion in three layers. The middle planar sheet is created by the two molecules with overlapping hydrogens in the center of the figure. c) The positions and anisotropic values of the atoms are represented as irregularly shaped atomic spheres. Important bond lengths are included in Å: C=O, C-O, CO-H, OH...O. d) The crystal maintains a near planar anthraquinone nucleus. The ellipsoids are drawn at the 50% probability level. The figures were created in ORTEP-3 (Farriugia 1997).

and 1.485(3) Å, but were within the normal quinoid range of values (Allen et al. 1987). However, the C10-C11 and C10-C14 bonds of 1.454(3) Å and 1.458(3) Å were shorter than benzequinoid bonds but still longer than anthracene or benzene bonds. They also did not correlate well with any of the $C_{aromatic}$ or quinone bond systems in the referenced collection. All

bonds within the nitro substituents (C1-N1, C8-N2, N1-O1, N1-O2, N2-O3, N2-O4) maintained normal $C_{\text{aromatic}}\text{-NO}_2$ values, but both C-OH bond lengths in ARDP0006 of 1.329(2) Å were shorter than the average $C_{\text{aromatic}}\text{-OH}$ distance of 1.362 Å. The H1 and H2 atoms hypothetically had the ability to rotate freely in solution making contacts with other molecules as the crystal formed, but they developed an intense attraction to the centrally located carbonyl oxygen, O8 with hydrogen bond lengths of 1.807(2) Å and 1.822 (4) Å. The lengths to their respective hydroxyl oxygens were 0.855(4) and 0.842(3) which is shorter than the 0.967 Å average length of O-H bonds. However, the thermal parameters show variability in their positions between the hydroxyls and O8 (figure 3.1.1c). H1 seems to experience more movement than H2 in the crystal, but there is no evidence that supports permanent transfer to O8. Seven bonds within ARDP0006 were inconsistent with their respective values according to the CSD (C10-C11, C10-C14, C10=O8, C4-O5H2, C5-O6H1, O6-H1, and O5-H2) and they are regionalized around the carbonyl, C10=O8. These peculiarities provided further insights into the structural details once our results were further explored and related structures were examined. These details will be elaborated in the discussion.

The structure's overall planarity was investigated. As figure 3.1.1d suggests, the anthraquinone nucleus was virtually flat with small deviations in bonded C1, C2, and C13 atoms. The ring involving C1 is aromatic, but the torsion angle indicates the attached O1-N1-O2 nitro group was twisted more than the other nitro by 40.5° and influenced the planarity of the ring. Crystal contacts and steric hindrances influence the positioning of the nitro. Hydrogen bonds are probable between the O1 atom and H6 and H3 and between the O2 atom and H2, H3, and H4 in nearby anthraquinones. More importantly, both nitros situated themselves such that they avoided interference with the π -stacking of the four molecule layers. Furthermore, torsion angles were

measured to understand the planarity in detail. The anthraquinone nucleus was not entirely planar and not uniformly bent. The torsion angles indicated that the molecule was bent at the C9 atom with an angle of 6.0° and, in contrast, was bent at the C10 atom with an angle of 0.1° (torsion angle details are found in table 4.1.2 and table 3 of the appendix). Additionally, the O7 carbonyl atom was elevated out of the anthraquinone plane with a torsion angle of 10.9° while the O8 atom was only elevated by 0.8°. Crystal packing was not influencing the position of O7 so steric effects from the nitro groups were the only remaining explanation. However, the subtle bending and twisting of the central ring and the C9=O7 carbonyl in the anthraquinone base of ARDP0006 was not caused by the N1 nitro, and was somewhat unexpected since the anthraquinone base was seemingly a three ring conjugated system. The atoms involved with the intramolecular hydrogen bonds maintained a planar configuration while the central ring and carbonyl group on the nitro side of the molecule were allowed to bend at the pivoting atom, C9. This indicated that a form of aromatic stabilization involving the C10 atom was caused by the hydrogen bonds.

Table 3.1.2 Hydrogen Bonding of the ARDP0006 crystal

D—H...A	D—H	H...A	D...A	D—H...A
Intramolecular Contacts (Å)				
O6—H1...O8	0.855(4)	1.822(4)	2.556(4)	149.22(4)
O5—H2...O8	0.842(3)	1.807(3)	2.569(3)	148.65(3)
Primary intermolecular Contacts (Å)				
O6—H1...O6i	0.855(4)	2.479(4)	2.582(4)	88.34(4)
O5—H2...O2ii	0.842(3)	2.442(3)	2.946(3)	120.39(3)
O5—H2...O7iv	0.842(3)	3.295(3)	3.190(3)	75.78(3)
C3—H3...O7i	0.974(3)	2.693(3)	3.275(3)	118.35(3)
C6i—H6i...O1ii	1.037(4)	2.681(4)	3.408(4)	130.99(4)
C7iii—H5iii...O3	0.967(3)	2.534(3)	3.304(3)	134.93(3)
C2—H4...O4ii	1.003(3)	2.585(3)	3.076(3)	110.76(3)

The structural data of ARDP0006 was not identical to the same compound (1,8-dinitro-4,5-dihydroxyanthraquinone) previously solved (Brown and Colclough 1983), but the structures correlated with similar dimensions. The advantages for our structure included: more reflections, lower residual factors for the merged data and refinement process, much improved isotropic temperature parameters, the addition of refined anisotropic thermal parameters, and a better understanding of the crystal packing and planarity of the overall molecule (additional details are compared in table 3.1.1). Both crystals were solved with virtually identical space groups ($P2_1/c$ vs $P2_1/a$). There were two notable differences between the two structures. The two C-OH bond distances were first noted as being different and were shorter by an average of 0.015 Å in ARDP0006. The C-OH length disparities were a direct result of a delicate intramolecular hydrogen bonding network. The second difference was noticed in the nitro torsion angles to the anthraquinone plane which happen to be 83.3° and 123.8° in our structure and 83.8° and 62.9° in the previously determined structure. Both structures similarly acquired one N-O bond slightly longer than the other N-O bond in both nitro groups. The oxygens of each nitro group made asymmetric intermolecular contacts in the crystal and slight differences were appropriate without further explanation. Nevertheless, both atomic structures sustained the same seven bond discrepancies on the hydroxyl side of the molecule of which do not belong to appropriate bond systems in the CSD data collection.

3.2 STRUCTURE DETERMINATION OF WNV NS2B-NS3 PROTEASE IN COMPLEX WITH ARDP0006.

We aimed to determine the exact binding location and correlating intermolecular contacts experimentally by utilizing x-ray crystallography. The results would have elucidated the structure of ARDP0006 when bound and would have significantly aided structure-based analog

design since the intermolecular contacts would be displayed. By understanding what conformation the molecule is in while in solution (aim1) and when bound to the protease (aim 2), we would have obtained clues on the binding mechanism and a better understanding of which binding interactions were taking place.

For our co-crystal trials, it was necessary to obtain large quantities of protein at high purity. The NS2B-NS3 protease construct was engineered with a Histidine tag so we could utilize Ni affinity chromatography. This was the primary method of purification, but high purity was not obtained until a High Pressure Liquid Chromatography (HPLC) step was performed. This yielded a protein solution with greater than 95% purity (figure 2.2a).

Obtaining protein co-crystals requires a delicate balance of protein, buffer, salt, precipitant, pH, and temperature. Screening for the correct ingredients requires large amounts of protein, so large screens were conducted with drop sizes of 0.4 μ L per well on 96 well plates. Upon the discovery of small microcrystals, smaller screens with larger, 4 μ L drops were created. Fortunately, crystals were found in two conditions. The microcrystals in figure 2.3.1b-iii were produced in the following conditions: 20 mg/mL NS2B-NS3 protein, 50mM NaCl, 10mM HEPES, 0.1M sodium acetate, 1.3M tri-sodium citrate, 0.1M bis-tris. The other two figures in 2.3.1b were produced in the following conditions: 20 mg/mL NS2B-NS3 protein, 50mM NaCl, 10mM HEPES, 5% glycerol, 0.2M ammonium sulfate, 20.9 w/v PEG 4000, 0.1M Tris, and pH 7.44. The thin, rod-like crystal in the left figure was mounted in paraffin oil under a cryostream of nitrogen gas and the space group was estimated to be P2, with unit cell parameters were estimated as: $a = 73 \text{ \AA}$, $b = 77 \text{ \AA}$, $c = 153 \text{ \AA}$, and $\beta = 102^\circ$. The diffraction pattern, seen in figure 2.3.1c was too weak to obtain sufficient data which is understandable for such a small crystal with dimensions of $5\mu\text{m} \times 5\mu\text{m} \times 120\mu\text{m}$. The promising larger crystal in figure 2.3.1b-ii has a

width of 80 μm and length of 250 μm , but diffraction was unattainable. The microcrystals in figure 2.3.1b iii were too small to mount and obtain x-ray data.

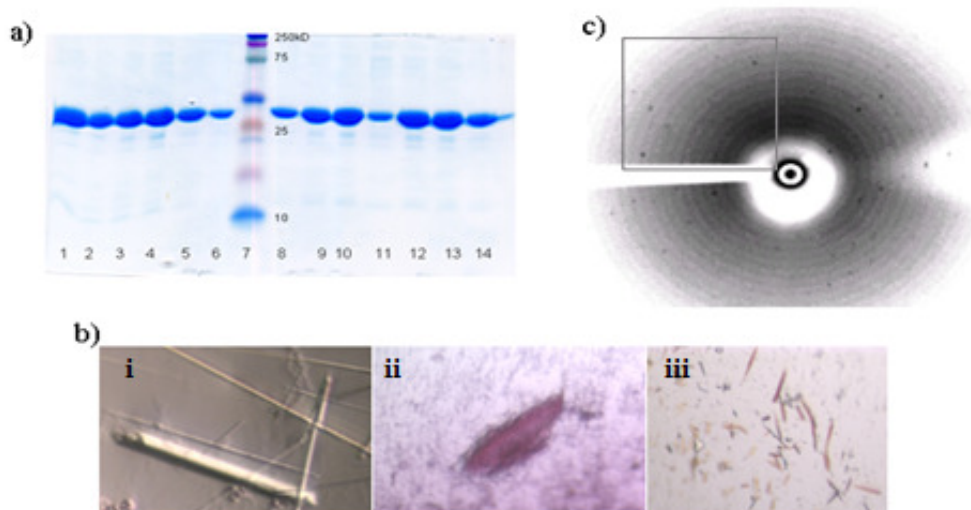


Figure 3.2.1. Progress towards X-ray determination of the protease and inhibitor complex. a) This shows an SDS-PAGE of purified WNV NS2B-NS3pro in lanes 1-6 and 8-14. Size of the protease is 27kD. B) Crystals were obtained. c) shows a weak diffraction pattern giving promise towards finding optimized and diffractable crystals.

3.3 STRUCTURE BASED DOCKING WITH AUTODOCK

Before docking, three ARDP0006 structures were created and prepared for docking: (1) a molecule was designed with PyMOL and minimized with CHARMM (figure 3.3.1a); (2) a PDB file was generated from the x-ray crystal structure solved in aim one (figure 3.1.1b); and (3) the crystal structure, (2), was altered to include the tautomer variant (figures 3.3.1c). The tautomeric hydrogen was assumed and designed to make an intramolecular hydrogen bond with the newly formed carbonyl because there is no reason why it should not reflect the same hydrogen bonding activity as the original crystal structure when disregarding possible solvent contacts being made. Structure (1) was created computationally and energetically minimized which resulted in a stable and planar molecule with both intramolecular hydrogen bonds. This result did not reproduce the

computationally generated bent structure as seen from the previously docked result in figure 1.3.1a and 1.3.1c which leads to the noticeable conclusion that either the two programs used vastly different calculations in finding a stable conformation or the bent structure was a product of the docking enumeration of ARDP0006 into the provided protease target. Structures (2) and (3) were essentially identical with the exception of the transferred hydrogen from O6 to O8 and the newly formed C5=O6 bond which was altered to reflect a common carbonyl bond length of 1.217 Å. The two primary differences between the three structures included the rotated nitros with respect to the anthraquinone and the tautomer variant in (3). AutoDock was allowed to rotate the C-N bonds and hydroxyl bonds during docking, so neither of the two differences was significant. Notably, both (1) and (2) include a planar anthraquinone nucleus and both hydroxyls are turned inward to create hydrogen bonds with the carbonyl oxygen. These two similarities were retained throughout all three docking experiments despite freedom allowed for the hydroxyl bond rotation.

Three docking experiments were conducted with molecules (1) – (3). Each was docked separately into the active WNV NS2B-NS3 protease crystal structure that was previously solved by x-ray crystallography when in complex with a peptide inhibitor (Erbel et al. 2006, PDB: 2FP7). An additional three docking experiments were performed in the same manner but with the previously solved apo crystal structure of the DV NS3 protease lacking the NS2B cofactor (Murthy et al. 1999; PDB: 1BEF). These dockings would have been effectively compared against the WNV NS2B-NS3 protease dockings and the original screening result, but the x-ray structure representative of the Murthy DV NS3 protein model was recently found to be fake (Grant 2009) and the results are not shown.

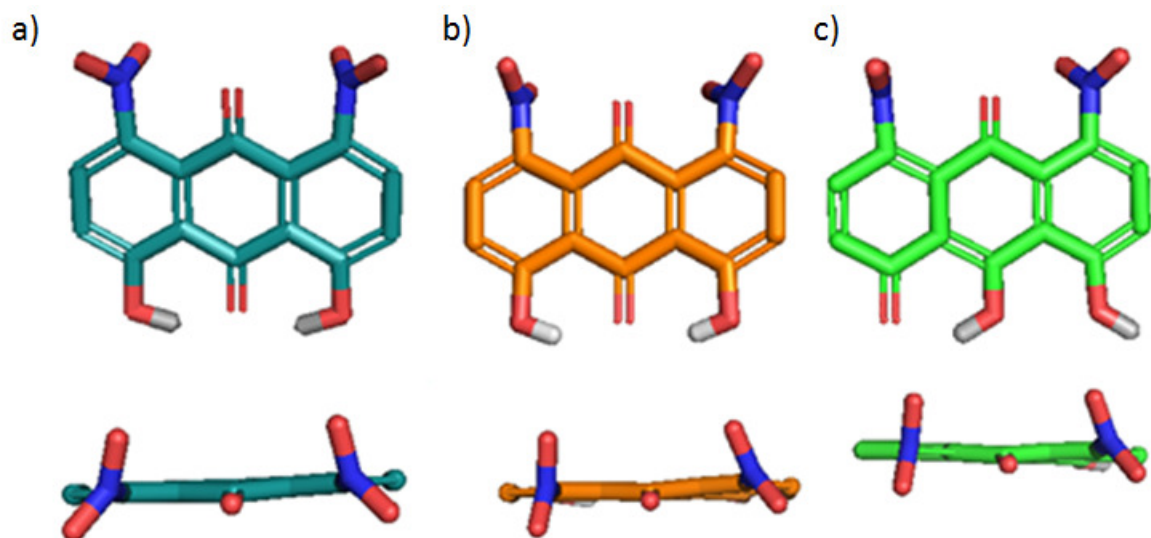


Figure 3.3.1. ARDP0006 molecules prepared for docking: the computationally designed model with PyMOL and minimized with CHARMM is cyan. The structure solved by x-ray crystallography is orange. The tautomer variant is green. (models were prepared in PyMOL).

The docking experiment with the WNV protease returned fifty conformations with 1,000,000 energy evaluations per conformation for each of the three docking experiments. The predicted poses were grouped into about half a dozen clusters with similar poses and predicted binding energies. A fairly common practice for selecting a pose is to choose the largest cluster and then the lowest energy conformation in the cluster. The predicted pose with the lowest energy is not necessarily the most sought after pose because energy rankings and docking scores are not regarded to be as reliable as the predicted poses (Cheng et al. 2009, Warren et al. 2006, Kontoyianni et al. 2004). The predicted poses and energies were very similar among the three ARDP0006 docking experiments. Of the total 150 predicted poses, 133 were found to bind in the S2 pocket of the protease, which gives a strong basis for choosing conformations bound there (table 3.3.1). Fortunately, the largest clusters for each docking experiment included the lowest predicted ΔG calculations and the poses in those clusters all possessed common poses. The three chosen conformations for evaluation, shown in figure 3.3.2a, were in the largest cluster and

obtained predicted binding energies within 0.30 kcal/mol of each other, which is practically negligible for computational predictions. As the figure 3.3.2a shows, the poses are nearly identical with a very low average RMS between the three molecules of 72 nm (nitro atoms and hydrogens were excluded). This variation of poses is much less than the pose variation of the same ARDP0006 molecule generated from the docking experiment. All 44 docked poses of the crystal structure in the S2 pocket are shown in figure 3.3.2b and clearly display more variety than the poses in 3.3.1a. Thus, the computationally designed and minimized, crystal structure, and tautomer structural variations of ARDP0006 do not provide any valuable differences or advantages in binding predictions.

Table 3.3.1. Docking pose predictions of the 3 ARDP0006 structures.

ARDP0006 Conformation	<i>in-silico</i>	Crystal structure	Tautomer	Total
# of poses docked in S2 pocket of 2FP7	43/50	44/50	46/50	133/150

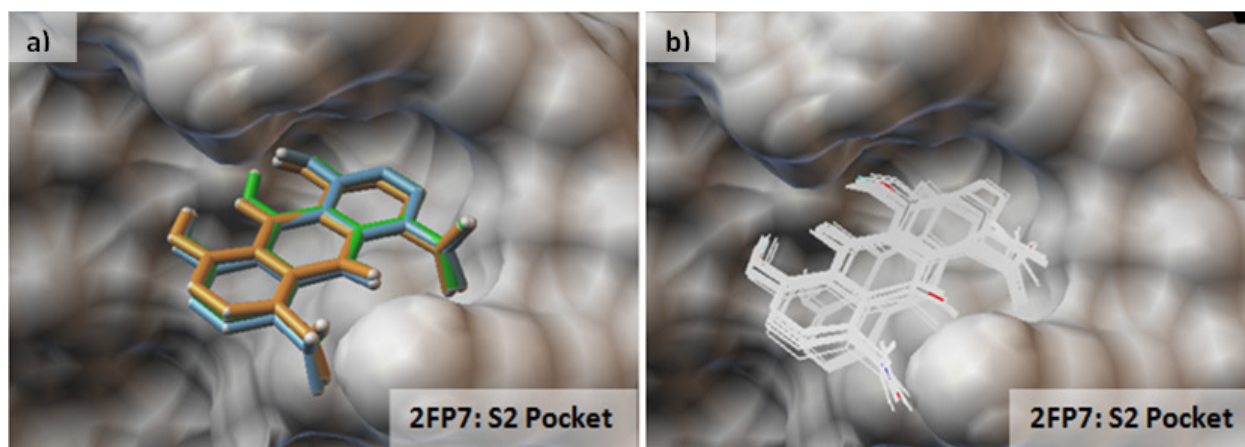


Figure 3.3.2. Docking results. a) Each of the three conformations docked in the previously determined crystal structure of the WNV NS2B-NS3 protease (2FP7). Colors correlate to figure 4.3.1: Cyan=*in-silico*, orange= x-ray crystal structure, Green=tautomer. b) Multiple docking poses of ARDP0006's crystal structure docked into the WNV NS2B-NS3 protease. Figures designed in AutoDock 4.0.

CHAPTER 4

Discussion

4.1. STRUCTURE BASED DETERMINATION OF ARDP0006 WITH X-RAY CRYSTALLOGRAPHY

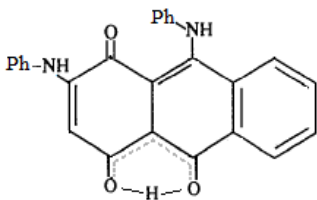
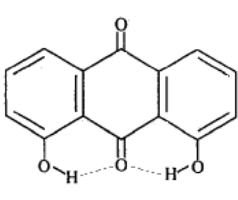
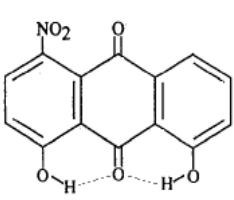
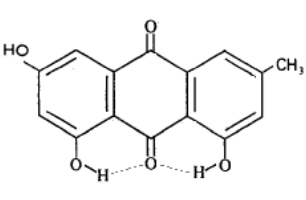
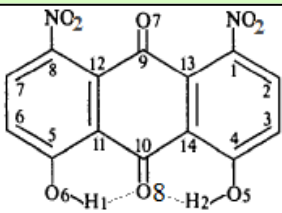
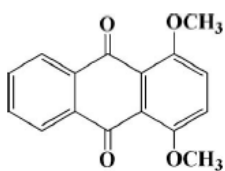
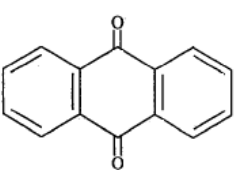
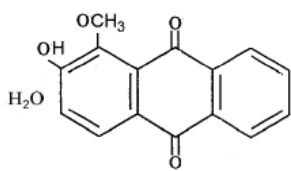
4.1.1. X-ray crystal structures of anthraquinone derivatives

The following discussion of ARDP0006 involves comparisons with many anthraquinone derivatives. Although dozens of anthraquinone structures are known, only seven, besides ARDP0006, are included for detailed evaluation. Of the molecules chosen, five are 4,5-dihydroxyanthraquinone analogs (compounds [I]-[V]) and three lack hydroxyls at the 4 and 5 positions of the anthraquinone base (compounds [VI] – [VIII]). The compounds are listed in table 4.1.1 and their bonds and correlating properties are listed in table 4.1.2. Compounds [II] – [VIII] contain simple substituents in order to minimize compounding chemical and electronic effects on the nature of the anthraquinone. Included is alizarin black-B [I] (Trafeenko et al. 1987), 4,5-dihydroxyanthraquinone [II] (Zain et al. 2005), 1-nitro-4,5-dihydroxyanthraquinone [III] (Yatsenko et al. 1996), 2-methyl-4,5,7-dihydroxyanthraquinone [IV] (Zhu et al. 2006), ARDP0006 [V], 1,4-dimethoxyanthraquinone [VI] (Cao et al. 2007), anthraquinone [VII] (Prakash 1967), and 2-hydroxy-1-methoxyanthraquinone [VIII] (Liu and Jiao 2009).

Alizarin black-B, [I], is more complex structure while including an amine at C9, a carbonyl at C8, two phenols, and an intramolecular hydrogen bond involved between the amine at C9 and a nearby carbonyl. Compound [I] was included in this comparison for three reasons: it still obtains a carbonyl and hydroxyl group on C10 and C5, it was crystallized with a tautomer transition state between the carbonyl and hydroxyl oxygens, and obtains valuable bond

information for comparison. More specifically, [II] was chosen due to the simplicity of obtaining only the 4,5-dihydroxy- substituents; [III] was included because it is only one nitro short of being identical to ARDP0006; [VI] was chosen because it is the only known anthraquinone crystal structure to possess a significant bend between the outer benzene rings (22.02°); and [VII] was selected because it is the original anthraquinone base with no substituents that would interfere with the electronic structure and, ultimately, the bond lengths.

Table 4.1.1: X-ray crystal structures of anthraquinone derivatives chosen for comparisons

			
[I]	[II]	[III]	[IV]
			
[V]	[VI]	[VII] (R=12.5%)	[VIII]

[I]: alizarin black-B (Trafeenko et al. 1987); [II]: 4,5-dihydroxyanthraquinone (Zain et al. 2005); [III]: 1-nitro-4,5-dihydroxyanthraquinone (Yatsenko et al. 1996); [IV]: 2-methyl-4,5,7-trihydroxyanthraquinone (emodin) (Zhu et al. 2006); [V]: ARDP0006; [VI]: 1,4-dimethoxyanthracene-9,10-dione (2007 et al. Cao); [VII]: anthraquinone (Prakash et al. 1967); and [VIII]: 2-hydroxy-1-methoxyanthraquinone (Liu and Jiao 2009).

4.1.2. Hydrogen bonds induce resonance

The x-ray crystallography results of ARDP0006 required further explanation about the nature of its structure. The carbonyl bond elongation (noted as C10=O8 in ARDP0006) has been observed in several anthraquinones, along with the other six bond peculiarities surrounding the carbonyl (table 4.1.2), and has been directly credited to intramolecular hydrogen bonding

(H1...O8...H2) (Ismail et al. 2009, Zhu et al. 2007, Boonnack et al. 2005, Zain et al. 2005, Ng et al. 2005, Yatsenko et al. 1996, Ulicky et al. 1987). A resonance model proposes that the hydrogens induce electron delocalization and that 4,5-dihydroxyanthraquinones consist of two electronic configurations (Ulicky et al. 1987). The first and primary resonance form consists of the typical anthraquinone representation with aromatic structures as the outer two rings and equivalent carbonyl bonds attached to the center ring. The second resonance form includes five important features that verify each bond irregularity in the 4,5-dihydroxyanthraquinones (comparisons can be seen in table 4.1.2). 1) The π -bond of the C10=O8 bond becomes polarized. The oxygen atom converts from sp^2 hybridization to sp^3 so it can accommodate the transferred electron. This is supported because the C10=O8 bond lengths in 4,5-dihydroxyanthraquinones are typically in the range of 1.240-1.255 Å, which closely resembles anionic-delocalized double bond character (1.235-1.255 Å) (Allen et al. 1987). 2) Consequently, a positive charge delocalizes within the anthraquinone nucleus which subtly increases the local dipole moments of substituents, including the carbon hydrogens. 3) The benzene π -electron cloud extends from the two outer rings to C10 which includes the C10 p-orbital electron in the conjugated system. This resonance characteristic is supported in many structures with the observed shortening of the C11-C10-C14 bonds to the approximate length of 1.455 Å (table 4.1.2 compounds [I] – [V]). 4) The anthraquinone nucleus is electron deficient due to the collective efforts of the delocalized positive charge and the electron withdrawing nature of the nitros. 5) The hydroxyls are electrostatically attracted to the positively charged ring structures which tightened the C-O-H bonds. The partial zwitterionic character also influenced crystal packing. The nitro groups were attracted to the increased dipole of the H1, H2, H3, and H4 atoms while O8 and O6 made weak anionic- π interactions with anthraquinone ring structures. This resonance

model helps in understanding the unconventional bond dimensions in ARDP0006 and agrees with several analogous 4,5 dihydroxy- structures.

In contrast, several related anthraquinone structures lack the hydroxyl-carbonyl interactions and indirectly provide insight into the tautomerism and planarity of ARDP0006. A correlation between two of the central ring C-C bonds and the C10=O8 carbonyl length was previously made in order to estimate bond order and imply a percentage of the second resonance structure as a part of the overall structure (Ulicky et al. 1987). They denoted Δ as the difference between the bond length C10-C11 and C11-C12. Smaller difference calculations successfully correlated with longer C10=O8 bond lengths which suggested a larger percentage of the resonance structure. Although this estimate is simplistic and imprecise for estimating bond order, it was effectively carried out to further prove the relevance of the resonance model and to attempt in making connections with the tautomerism and planarity of the anthraquinone nucleus. The discussion is supplemented with and supported by a thorough analysis of comparable anthraquinone structures solved by x-ray crystallography (comparisons can be viewed in table 4.1.2).

The hydrogen bond induced resonance structure clearly cannot exist if an anthraquinone lacks the hydroxyls near either quinone carbonyl. Compounds [V] – [VIII] further support the resonance model because they do not comprise of the seven peculiar bond distances in compounds [I] – [V] and conversely portray typical quinoid $C_{sp^2}=O$, benzenoid, and quinoid central ring C-C bond lengths with averages of 1.222Å, 1.378Å, and 1.478Å, respectfully (Allen et al. 1987). Table 4.1.2 was created to organize relevant structural information and ranked each examined anthraquinone by their Δ calculation from smallest to largest. The compound with the smallest Δ calculation was alizarin black-B, [I], with a value of -0.046 and the largest was

compound [VIII] with a value of 0.093. Each of the 4,5-dihydroxy- compounds had values of 0.054 and smaller while the non-4,5-dihydroxy- molecules obtained values of 0.072 and larger. Carbonyl bonds in the resonance containing structures ranged from an abnormally long length of 1.301(3) Å to 1.241(2) Å corresponding well with Δ values of -0.046 and 0.054, respectfully. More importantly, the anthraquinones lacking the resonance form showed no direct relationship between Δ and their carbonyl lengths. Their carbonyl lengths were 1.215(2) Å, 1.213 Å, and 1.219(2) Å which corresponded to a Δ value of 0.072, 0.082, and 0.093, respectfully. Thus, simple approximate cutoff values can be established such that, if the Δ value is near 0.070 or larger, then the resonance structure contribution can be removed from consideration. Moreover, if the Δ value is near 0.005 or becomes negative, the percentage of the resonance structure is suggested to be near 50% or more (Ulicky 1987) and, due to the crystallized intermediate hydrogen in [I], tautomerism should be considered. From this simple approximation, ARDP0006 contains approximately 25% of the resonance structure and no consideration for tautomerism.

Table 4.1.2.
Anthraquinone
comparisons

	[I]	[II]	[III]	[IV]	[V]	[VI]	[VII]	[VIII]
	4,5-Dihydroxyanthraquinones					Non-4,5 Dihydroxyanthraquinones		
H-Bonded quinone oxygen?	Yes	Yes	Yes	Yes	Yes	No	No	No

Ranked from smallest to largest: The difference (Δ) of two significant bond distances (d) in the central ring (Å).

$$\Delta = d(\text{C10}=\text{C11}) - d(\text{C11}=\text{C12})$$

Δ	-0.046	0.041	0.046	0.046	0.054	0.072	0.082*	0.093
----------	--------	-------	-------	-------	-------	-------	--------	-------

Significant bond length comparisons between crystal structures of 4,5-dihydroxy- and non-4,5-dihydroxy- anthraquinone (Å)

C9=O7	1.262(3)	1.219(5)	1.210(3)	1.222(4)	1.212(2)	1.215(2)	1.213	1.219(2)
C10=O8	1.301(3)	1.248(5)	1.242(2)	1.266(4)	1.241(2)	1.215(2)	1.213	1.218(2)
C9-C12	1.423(3)	1.478(4)	1.492(3)	1.490(4)	1.478(3)	1.490 (2)	1.495	1.485(3)
C9-C13	1.467(3)	1.478(4)	1.490(3)	1.480(4)	1.485(3)	1.487(2)	1.472	1.479(3)
C10-C11	1.395(3)	1.447(3)	1.465(3)	1.450(4)	1.458(3)	1.490 (2)	1.472	1.487(3)
C10-C14	1.460(3)	1.447(3)	1.462(3)	1.455(4)	1.454(3)	1.487(2)	1.495	1.487(3)
C11-C12	1.439(3)	1.406(3)	1.419(3)	1.404(4)	1.404(3)	1.417(3)	1.401	1.392(3)
C13-C14	1.410(3)	1.406(3)	1.407(3)	1.413(4)	1.406(3)	1.392(3)	1.401	1.408(3)
C4-O5H2	1.306(3)	1.343(4)	1.331(3)	1.358(4)	1.329(3)	–	–	–
C5-O6H1	–	1.343(4)	1.337(3)	1.355(4)	1.329(3)	–	–	–
π -stacking in crystal?	N/A	Yes	Yes	Yes	Yes	No	Yes	Yes

Anthraquinone planarity – Torsion angles (Φ)

$\phi_1(\text{C8-C12-C9-C13})^\circ$	N/A	0.4(3)°	3.7(2)°	1.1°	6.0(2)°	21.7(2)°	N/A	3.0(2)°
$\phi_2(\text{C5-C11-C10-C14})^\circ$	N/A	1.0(3)°	0.6(2)°	0.8°	0.1(2)°	24.1(2)°	N/A	2.1(2)°
$\Delta(\phi_1)^\circ = \phi_1 - \phi_2$	N/A	-0.6°	3.1°	0.3°	5.9°	-2.4°	N/A	0.9°

Carbonyl planarity with the anthraquinone plane – Torsion angles (Φ)

$\phi_3(\text{C8-C12-C9-O7})^\circ$	N/A	1.0(3)°	6.5(2)°	1.9°	10.9(2)°	25.5(2)°	N/A	3.3(2)°
$\phi_4(\text{C5-C11-C10-O8})^\circ$	N/A	0.4(3)°	0.3(2)°	0.9°	0.8(2)°	25.5(2)°	N/A	1.6(2)°
$\Delta(\phi_3)^\circ = \phi_3 - \phi_4$	N/A	0.6°	6.2°	1.0°	10.1°	0.0°	N/A	1.7°

*This Δ calculation was averaged between C9-C12 and C9-C13 whereas $\Delta = \{[d(\text{C9-C12}) + d(\text{C9-C13})]/2\} - d(\text{C11-C12})$.

†Compound [I] is unique and contains a carbonyl on C8 instead of C9 and it is H-bonded to the amine on C9. Thus, C9=O8 is really C8=O8 for [I] and its distance is longer than typical Csp²=O bonds. However, C10=O8 is the point of interest.

4.1.3. Conformational implications deduced from the crystal structure

Without the stabilizing nature of resonance, the anthraquinone nucleus may prefer to bend in solution at the pivotal atomic positions, C9 and C10. After searching for anthraquinone structures in literature, 1,4-dimethoxyanthraquinone (Cao et al. 2007), [VI], was the only structure found not to form its crystal with π - π stacking interactions. Coincidentally, [VI] was reported to have a 22.02° angle between the two outer benzene rings. Factors contributing to the bent structure were not reported or apparent in the crystal packing interactions so evidence points to the lack of the stabilizing π - π stacking interactions in the crystal. This suggests either a form of non-planar resonance or an absence of localized π -bonding for the center ring. In addition, some structure reports included a measurement of the dihedral angle between the two outer benzene rings which ranged from 0° to about 5° . Furthermore, the carbonyls in some structures were non-planar. Thus, bent anthraquinones may be a favorable conformation when not in crystal form.

Since the anthraquinone structures seems to have a bending nature and the hydrogen induced resonance should establish some degree of planar stability, then one side of the molecule may be more planar than the other, as long as some twisting is allowed. Measuring torsion angles on both sides of the molecule is a good method of determining if the planarity is symmetric. The measured angles were expected to be small due to the fact that the molecules were crystal packed in parallel to each other with π - π stacking interactions. However, a small trend in the measurements has given reason to believe that the nonequivalent planarity could be magnified when the molecules are in solution. This would further indicate that the 4,5-dihydroxy- resonance structure in ARDP0006 could stabilize its planarity in solution. The

possibility that ARDP0006 may bend to nearly 20° is important to the binding affinity including the overall fit, steric hindrances, and distances between atomic electrostatic attractions.

Bond lengths and angles for some structures were not specified in literature, so to further investigate the structural details, the Crystallographic Information Framework (CIF) files of the selected structures were downloaded from the International Union of Crystallography website (<http://journals.iucr.org/>) and all measurements were made using ORTEP-3. The torsion angles involved at the pivotal C9 and C10 atoms were measured in order to assess the bending nature of the anthraquinone base and carbonyls on both sides of the molecule. Torsion angles and parameters were designated as follows: $\phi_1 = \phi(\text{C8-C12-C9-C13})$ represents torsion angle of the non-resonance side of the anthraquinone nucleus in all the compounds; $\phi_2 = \phi(\text{C5-C11-C10-14})$ represents torsion angle of the opposite side which happens to be the resonance side of compounds [I] – [V]; $\phi_3 = \phi(\text{C8-C12-C9-O7})$ represents the carbonyl oxygen angle of elevation from the plane of the anthraquinone on the non-resonance side of each molecule, $\phi_4 = \phi(\text{C5-C11-C10-O8})$ represents the carbonyl oxygen angle of elevation from the plane of the anthraquinone on the opposite side which happens to be the resonance side of compounds [I] – [V]; $\Delta(\phi_{12}) = (\phi_1 - \phi_2)$ and $\Delta(\phi_{34}) = (\phi_3 - \phi_4)$ is the difference in angles between the resonance side and non-resonance side of the molecule with respect to the anthraquinone bend and the carbonyl bend. All measurements are included in table 4.1.3.

A subtle trend was observed such that the resonance side is slightly more planar than the non-resonance side of molecules [II] – [V]. The largest $\Delta(\phi_{12})$ and $\Delta(\phi_{34})$ values were found in [III] and [V] as 3.1° and 6.3°, and 5.9° and 10.1°, respectfully. Molecules [II] and [IV] have no substituents to disturb the π - π stacking interactions which allowed their molecules to become more tightly packed and maintain a more planar conformation. Thus, their ϕ and $\Delta(\phi)$ values

were low. Compounds [VI] and [VIII] showed an average $\Delta(\phi)$ of 1.67° which is expected since they have no clear reason to produce different $\Delta(\phi)$ values except for their respective crystal packing interactions. Compounds [I] and [VII] could not be assessed due to the lack of information and inability to obtain their CIF files. Compounds [I], [III] and [V] are the only structures in the selected set of anthraquinones whose substituents are largely out-of-plane and interfere with the π - π stacking. Accordingly, [III] and [V] have the larger ϕ and $\Delta(\phi)$ values, but the torsion angles are still much less than in [VIII]. This evidence marginally supports the concept that anthraquinones in solution may be capable of bending up to 20° and the induced resonance would stabilize their planarity in 4,5-dihydroxyanthraquinones. At the very least, this indicates reasonable doubt that the anthraquinones favor planar conformation in solution. All measurements and calculations are in table 4.1.2.

4.1.4. Tautomer implications deduced from the crystal structure

Clearly both hydroxyl hydrogens strongly interact with the carbonyl and induce a resonance structure, but tautomerism is observed when one of the hydrogens transfers from the hydroxyl to the carbonyl. The atomic displacement ellipsoid of H1 showed more positional deviation than H2 between its hydroxyl and O8. Two nearby oxygens, O2ⁱⁱ and O7^{iv}, were weakly attracted intermolecularly with H2 while H1 was only weakly involved with one nearby oxygen, O6ⁱ. This possibly restrained H2 from interacting with O8 as strongly as H1 did (figures 3.1.1c and table 3.1.1 confirm these hydrogen bonding details). When plotted with ORTEP-3, the 50% atomic displacement ellipsoid of H1 extends halfway to O8. This clearly displays the hydrogen's attraction to O8, but there is no evidence of any hydrogen permanently anchoring to O8 in the crystal, which would be indicative of an energetically favorable and stable tautomer.

Excited state charge transfer or proton exchange may be occurring in a small portion of ARDP0006 molecules in the crystal. Previous spectroscopy experiments including Raman spectroscopy, infrared, and UV-VIS have also observed strong hydrogen bonding in hydroxy anthraquinones and have shown moderately dissimilar results whereas excited state tautomerism and proton tunneling (two minima exist for the proton between the two oxygens in the ground state) exist in non-polar solvents (Marzocchi 2003, Smulevich 1987, Ezaby1970). If the proton was transferring equally or was equally delocalized between the two oxygens we would expect to observe equal bond lengths for C10-O8 and C1-O6, and the hydrogen would be more centrally positioned between both oxygens like was observed in the unique crystal structure of alizarin blue-black B (Trafeenko et al. 1987). In alizarin blue-black B, the C-O lengths for the hydroxyl and carbonyl were 1.306(3) Å and 1.301(3) Å, respectfully, while the hydrogen's anisotropic ellipsoid spanned the entire distance between the hydroxyl to the carbonyl. If half the ARDP0006 molecules in the crystal structure permanently obtained the tautomer variant O8-H, then we would have observed either nearly equal values for the C-O bond lengths or both tautomers could have appeared simultaneously in one crystal like was observed for 4-nitro-5-methoxyimidazole and its tautomer (Kubicki et al. 2003). Over 80 space groups ranging from each of the seven lattice systems were chosen to solve the ARDP0006 atomic structure in an exhaustive attempt to find a unit cell with some evidence of two individual tautomers, but none provided a dependable model of ARDP0006 that was different than the one created by solving the structure with the $P2_1/c$ space group. (Figure 3.1.1 shows the refined crystal structure of ARDP0006 solved with the $P2_1/c$ space group and is graphically represented by ORTEP-3). Furthermore, if the hydrogens were simply attracted to O8 and restrained from proton or electron exchange, then the structural model would not show such atomic deviation in the hydrogen's

position. In conclusion, although H1 shows more interaction with O8 than H2 in the crystal, there was no evidence suggesting the existence of an energetically stable tautomer where a hydrogen atom is permanently bound to O8 in any portion of the molecules in the crystal. With that being said, evidence suggests a small amount of either charge transfer or proton exchange may be occurring between the hydroxyls and carbonyl which supports previous studies that investigated tautomerism in non-polar solution. This conclusion agrees well with the Δ calculation that estimated the amount of resonance structure to be nearly 25% because of the active interaction between H1 and H2 with O8.

4.1.5. Implications for ARDP0006 in solution

Although the crystal structure and reviewed spectroscopy studies have provided an opportunity to examine the atomic structure, electronic configuration, overall conformation, and possible instances of tautomerism in crystal form and in non-polar solvents, the structure's hydrogen bonding features in a solution with more physiological-like conditions should be investigated. In protic solvents that exhibit moderate polarity and hydrogen binding capabilities, the hydroxyl hydrogens may interact with the solvent instead of O8. In such conditions, ARDP0006 may not display any fraction of the resonance induced structure and may exist in a bent configuration. By understanding the molecule's conformation before binding the protease and after binding the protease will aid in elucidating the binding mechanism.

Additionally, ARDP0006 should be considered for drug absorption potential. The molecule passes the "Lipinski's Rule of Five" criterion which is a set of empirically derived rules created to describe the chemical properties commonly found in a large dataset (2,245 compounds) of marketed oral drugs (Lipinski et al. 2001). However, if the ARDP0006 exhibits zwitterionic properties and the resonance structure is present in polar solvents, then the

molecules would be regarded as poorly absorbed. This gives additional motivation for investigating the electronic configuration and structure while in polar solution.

4.2 STRUCTURE DETERMINATION OF WNV PROTEASE IN COMPLEX WITH ARDP0006

Fully diffractable crystals were not achieved, so no structural information was obtained. Although an inhibitor bound protease structure was not solved, a weak diffraction pattern was collected from one small crystal and consistent with protein diffraction. X-ray crystallography has proven to be a successful method for solving the structure of other DV and WNV NS3 protease structures (Table 2.4.1). These conclusions provide evidence that future experiments may lead to a solved structure if larger crystals can be produced.

Table 4.2.1. Existing NS3 protease solved structures in the Protein Data Bank.
<http://www.rcsb.org>

PDB	Molecules	Resolution	R-Value	R-Free	Reference
<i>West Nile virus</i>					
3E90	NS2B-NS3 + Naph-KKR-H	2.45	0.197	0.244	Robin 2009
2GGV	NS2B-NS3; H51A mutant	1.80	0.193	0.228	Aleshin 2007
2IJO	NS2B-NS3 + BPTI	2.30	0.212	0.264	Aleshin 2007
2FP7	NS2B-NS3 + Bz-Nle-KRR-H	1.68	0.182	0.219	Erbel 2006
<i>Dengue virus</i>					
2FOM	NS2B-NS3	1.50	0.177	0.209	Erbel 2006

4.3 STRUCTURE BASED DOCKING WITH AUTODOCK

As indicated in the results section, there is no apparent difference or advantage in docking the crystal structure or tautomer structure of ARDP0006 over the computationally

designed structure. To further evaluate the differences in the docking poses of both possible ARDP0006 tautomers, intermolecular contacts were measured between the protease and the small molecules. For this analysis, the two poses chosen obtained the lowest binding energy predictions but were still within the same cluster. The two small molecules were still within 0.30 kcal/mol for their predicted binding energies and obtained low RMS values, but these conformations chosen would be the ones commonly selected from a typical virtual screening experiment. Figure 3.3.3 displays a stereo view of the crystal structure docked and figure 3.3.4 displays a stereo view of the tautomer variant's docked pose. As the figures show, the same three hydrogen bonds were made with a slight distance difference of 0.1 Å due to the proposed tautomeric hydrogen. The Histidine, His-51, is close enough to attempt at making a weak π - π interaction depending on how the protease fluctuates upon binding either of the small molecules. According to LigPlot (Wallace et al. 1995), Trp-50 and Asp-75 make additional hydrophobic contacts with the anthraquinone nucleus of both ARDP0006 molecules. Thus the most apparent contacts have been elucidated and do not display differences significant enough to affect a virtual screening experiment.

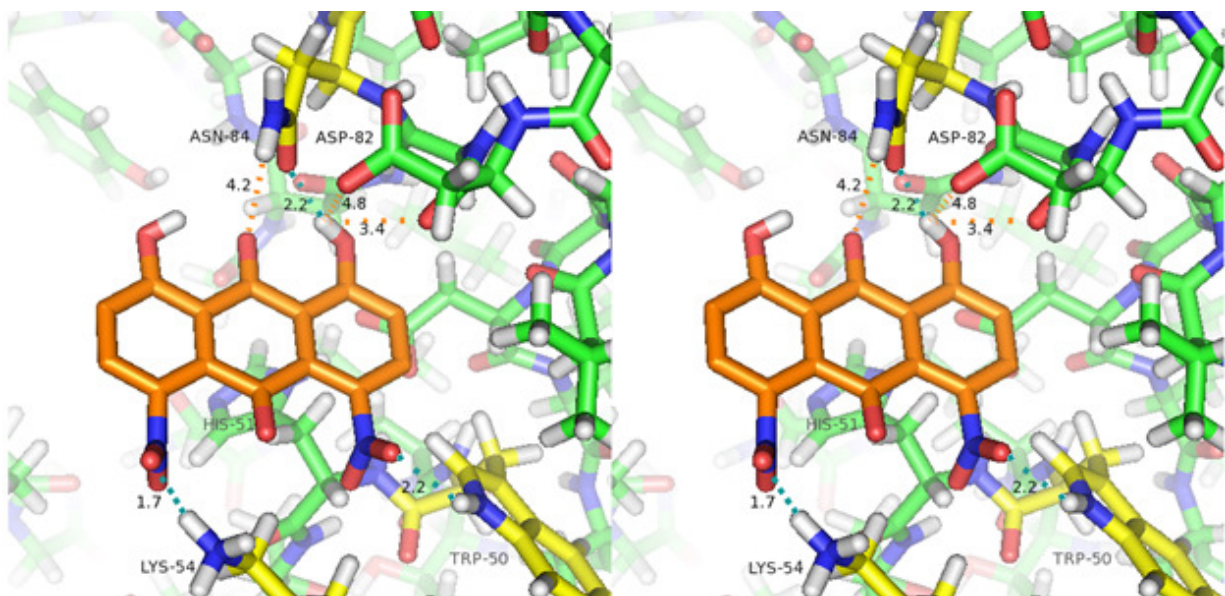


Figure 4.3.1. Intermolecular contacts made in docking results for the crystal structure. This docking result is visualizing the intermolecular contacts being made between the crystal structure of ARDP0006 and the WNV NS2B-NS3 protease in the S2 pocket in stereo view. Modeled in PyMOL.

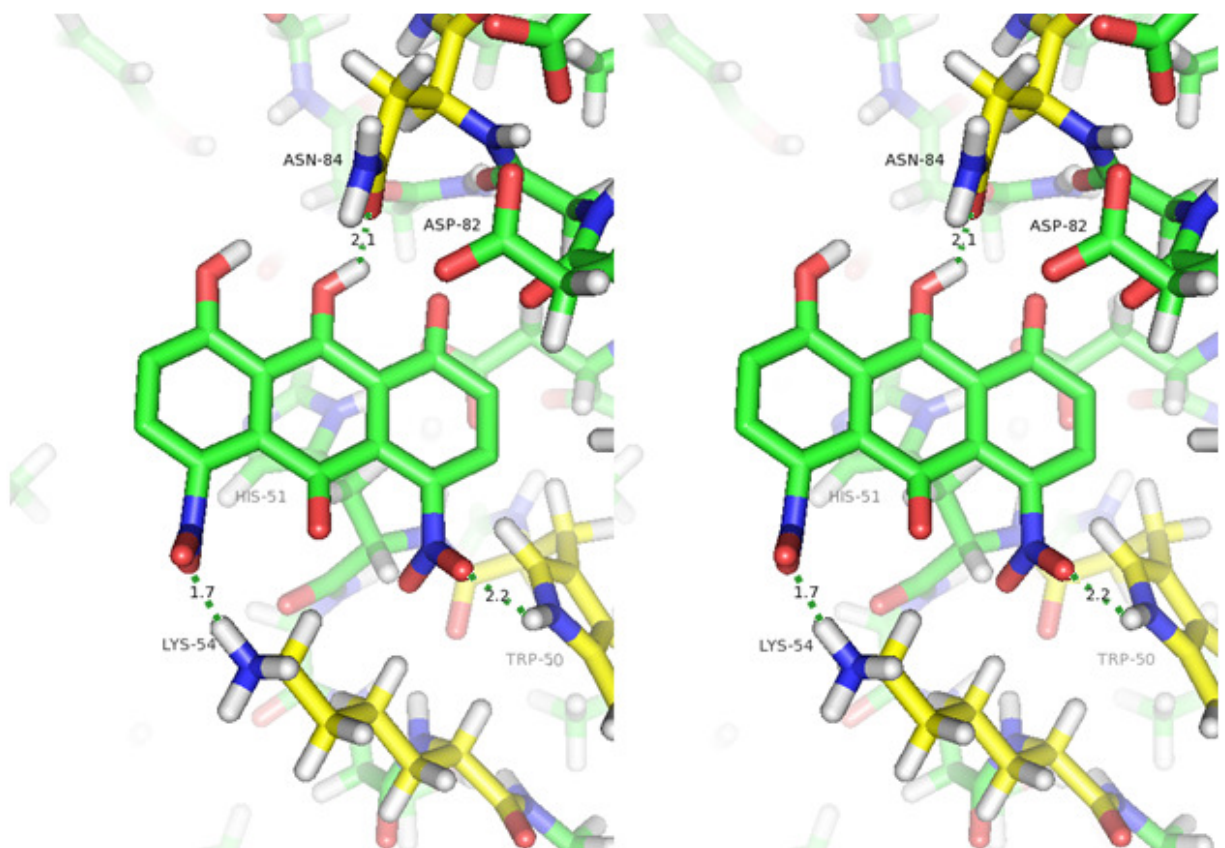


Figure 4.3.2. Intermolecular contacts made in docking results for the tautomer structure. This docking result is visualizing the intermolecular contacts being made between the tautomer variant of ARDP0006 and the WNV NS2B-NS3 protease in the S2 pocket in stereo. Modeled in PyMOL.

The comparison above is, however, dependent on the notion that the intramolecular bonds remain bonded in solution and upon binding the protease. If the hydrogens become solvated and break away from the carbonyl, then three important changes would occur. 1) the hydrogen in the original crystal structure has an opportunity to bind with the backbone carbonyl of Gly-83 and the an acidic oxygen of aspartic acid, Asp-82 (as elucidated by the orange dashed lines in figures 3.3.3 and 3.3.4). The distance between these potential hydrogen bonding opportunities and the tautomeric hydrogen is too long and does not exist. 2) The nearby residue, Asn-84, of the NS2B cofactor has a better opportunity to hydrogen bond with the O8 carbonyl with the hydrogen of the amine side chain of the original crystal structure. The tautomeric hydrogen ultimately shields any involvement between the O8 oxygen and Asn-84. 3) The anthraquinone nucleus may be capable of bending with the lack of the induced resonance structure, which would moderately affect the steric hindrances and hydrogen bond distances. Thus, the set of docking experiments conducted in this study, where the hydrogens remain intramolecularly bonded, is somewhat incomplete. Still the results indicate virtually no differences or advantages between the computationally designed, crystal structure, tautomer. I would expect to observe minor differences if the hydrogens were allowed to be unbound to O8. However, I would suspect very little influence on the overall docking outcome because the binding energy would only be affected slightly and the molecule would most certainly be predicted to bind in the S2 pocket.

CHAPTER 5

Conclusions

The aim of this project was to determine the exact binding location and correlating intermolecular contacts between the West Nile protease and the small molecule ARDP0006. The docking result in a previous study opened new questions about the overall structure of the ARDP0006. Although co-crystallization was pursued, the minor structural uncertainties of ARDP0006 were most likely not obtainable in the co-crystal complex. Furthermore, the co-crystal trials were not obtained to give desired structural information. Thus, the small molecule structure was solved with x-ray crystallography in order to gain a better understanding of the structure. The structure of ARDP0006 was more complex than originally suspected due to the hydrogen bond induced resonance structure. The final conclusion for aim two is that although the structure displays resonance in crystal form and some form of proton exchange in non-polar solvent, determination of the hydrogens intramolecularly bind or interact with solvent in polar solution is needed. This will help in understanding if the molecule is allowed to bend, and if the hydrogens are allowed to freely interact with the protease upon binding or if they remain bound intramolecularly to the carbonyl. In further detail, neither a bent structure nor a stable tautomer was observed in the crystal. The bent structure may be possible if the intramolecular hydrogens do not interact with the carbonyl due to the lack of the stabilizing nature of a resonance structure. Thus, although the docking result in figure 1.3.1c seems like a docking artifact, it may be a possible conformation if the hydrogens are not bound to the carbonyl. This can only be confirmed once the structure of ARDP0006 is determined in a polar solvent.

Although the bent conformation of ARDP0006 may affect its binding affinity with the protease, the docking results indicate that the minor structural changes, including the tautomer and planarity, will not affect the docking results. Therefore, as far as virtual screening is concerned, there is no reason to further investigate the small molecule's detailed structure in polar solvent. With that being said, the binding mechanism and/or binding affinity still has a possibility of being affected with either the bent conformation or the tautomer.

Appendix

Table 1. Atomic coordinates (x, y, z), isotropic parameters – U (\AA^2), and atomic displacement parameters (\AA^2)

	x/a	y/b	z/c	U	U11	U22	U33	U12	U13	U23
H(1)	-0.329(4)	1.013(2)	0.465(3)	0.06771	0.04(2)	0.04(2)	0.13(4)	-0.01(2)	0.03(2)	0.04(2)
H(2)	-0.139(4)	0.855(2)	0.291(2)	0.04499	0.03(2)	0.03(1)	0.07(2)	-0.01(1)	-0.02(1)	0.00(1)
H(3)	0.225(3)	0.829(2)	0.162(1)	0.03237	0.04(1)	0.04(1)	0.01(1)	0.01(1)	-0.01(1)	0.00(1)
H(4)	0.471(4)	0.971(2)	0.144(1)	0.03628	0.06(2)	0.04(1)	0.01(1)	0.02(1)	0.02(1)	-0.01(1)
H(5)	-0.016(4)	1.373(2)	0.573(2)	0.04255	0.05(2)	0.04(1)	0.03(1)	0.00(1)	0.01(1)	-0.04(1)
H(6)	-0.286(4)	1.234(2)	0.584(1)	0.03974	0.02(1)	0.10(2)	0.00(1)	0.00(1)	-0.01(1)	-0.01(1)
O(1)	0.69610(18)	1.13558(10)	0.22408(8)	0.03118	0.0171(6)	0.0408(7)	0.0347(8)	0.0007(5)	-0.0026(5)	0.0021(5)
O(2)	0.44811(19)	1.24966(9)	0.19218(8)	0.02803	0.0291(7)	0.0307(6)	0.0237(7)	0.0018(5)	-0.0007(5)	0.0023(4)
O(3)	0.22166(19)	1.44947(10)	0.42009(8)	0.03339	0.0313(7)	0.0326(7)	0.0353(8)	-0.0009(5)	-0.0021(6)	0.0056(5)
O(4)	0.4118(2)	1.3760(1)	0.5265(1)	0.03575	0.0312(8)	0.0483(8)	0.0251(7)	-0.0089(6)	-0.0110(6)	-0.0027(5)
O(5)	-0.0569(2)	0.8323(1)	0.2577(1)	0.02844	0.0252(8)	0.0315(7)	0.0280(8)	-0.0053(5)	-0.0010(6)	-0.0045(5)
O(6)	-0.3444(2)	1.0641(1)	0.4965(1)	0.02996	0.0237(7)	0.0351(8)	0.0312(8)	-0.0031(6)	0.0028(5)	-0.0018(5)
O(7)	0.42043(17)	1.23641(10)	0.37070(7)	0.02554	0.0202(7)	0.0336(6)	0.0217(7)	-0.0036(4)	-0.0041(5)	-0.0024(4)
O(8)	-0.1999(17)	0.95331(10)	0.37498(8)	0.02659	0.0224(6)	0.0312(6)	0.0254(7)	-0.0046(5)	-0.0019(5)	-0.0004(5)
N(1)	0.5159(2)	1.1603(1)	0.2182(1)	0.02393	0.0212(7)	0.0326(7)	0.0170(7)	0.0005(5)	-0.0039(5)	-0.0013(5)
N(2)	0.2624(2)	1.3771(1)	0.4746(1)	0.02596	0.0245(8)	0.0310(8)	0.0215(8)	-0.0006(6)	-0.0023(6)	-0.0046(5)
C(1)	0.3718(2)	1.0738(1)	0.2383(1)	0.02322	0.0174(8)	0.0307(8)	0.0203(8)	0.0017(6)	-0.0051(6)	0.0016(6)
C(2)	0.3716(3)	0.9779(1)	0.1878(1)	0.02544	0.0220(8)	0.0328(9)	0.0201(9)	0.0032(6)	-0.0057(7)	-0.0005(6)
C(3)	0.2275(3)	0.8979(2)	0.1968(1)	0.0258	0.0230(9)	0.0315(9)	0.0214(9)	0.0025(7)	-0.0062(7)	-0.0023(7)
C(4)	0.0813(3)	0.9123(1)	0.2549(1)	0.02387	0.0214(8)	0.0292(8)	0.0189(8)	-0.0002(6)	-0.0096(6)	0.0011(6)
C(5)	-0.1938(3)	1.1353(1)	0.4876(1)	0.02428	0.0180(8)	0.0323(8)	0.0213(8)	0.0017(6)	-0.0051(6)	0.0028(6)
C(6)	-0.1817(3)	1.2273(2)	0.5428(1)	0.02761	0.0243(9)	0.0349(9)	0.0231(9)	0.0030(7)	-0.0012(7)	-0.0001(7)
C(7)	-0.0292(3)	1.3036(2)	0.5387(1)	0.02659	0.0254(9)	0.0335(9)	0.0199(9)	0.0024(7)	-0.0029(7)	-0.0014(7)
C(8)	0.1116(3)	1.2885(1)	0.4787(1)	0.02449	0.0215(8)	0.0295(8)	0.0211(8)	0.0006(6)	-0.0057(6)	0.0001(6)
C(9)	0.2663(2)	1.1815(1)	0.3640(1)	0.0218	0.0194(8)	0.0272(8)	0.0173(8)	0.0008(6)	-0.0068(6)	0.0020(6)
C(10)	-0.0619(2)	1.0224(1)	0.3705(1)	0.02287	0.0185(8)	0.0285(8)	0.0196(9)	0.0022(6)	-0.0092(6)	0.0036(6)
C(11)	-0.0503(2)	1.1201(1)	0.4272(1)	0.02274	0.0187(8)	0.0297(8)	0.0183(8)	0.0013(6)	-0.0066(6)	0.0008(6)
C(12)	0.1066(2)	1.1983(1)	0.4234(1)	0.02257	0.0189(8)	0.0286(8)	0.0188(8)	0.0018(6)	-0.0057(6)	0.0018(6)
C(13)	0.2386(2)	1.0885(1)	0.3004(1)	0.02204	0.0178(8)	0.0285(8)	0.0182(8)	0.0017(6)	-0.0069(6)	0.0012(6)
C(14)	0.0878(2)	1.0076(1)	0.3083(1)	0.02216	0.0186(8)	0.0301(8)	0.0159(8)	0.0009(6)	-0.0085(6)	0.0002(6)

Table 2. Geometric Parameters of 1,8 dinitro 4,5 dihydroxyanthraquinone (ARDP0006)

Bond Angles (°)				Bonds (Å)			
H(2)-O(5)-C(4)	106.2(18)	H(1)-O(6)-C(5)	106.8(22)	H(3) -C(3)	0.98(3)	H(5) -C(7)	0.98(3)
O(4)-N(2)-O(3)	125.3(2)	O(4)-N(2)-C(8)	118.8(2)	H(4) -C(2)	0.99(3)	H(6) -C(6)	0.98(3)
O(3)-N(2)-C(8)	115.8(2)	C(12)-C(11)-C(5)	119.4(2)	H(2) -O(5)	0.82(3)	H(1) -O(6)	0.79(4)
C(12)-C(11)-C(10)	120.5(2)	C(5)-C(11)-C(10)	120.1(2)	O(7) =C(9)	1.212(2)	O(8) =C(10)	1.241(2)
C(11)-C(12)-C(9)	120.7(2)	C(11)-C(12)-C(8)	118.3(2)	O(5) -C(4)	1.329(3)	O(6) -C(5)	1.329(3)
C(9)-C(12)-C(8)	121.0(2)	C(10)-C(14)-C(4)	120.1(2)	N(1) -O(2)	1.214(2)	N(2) -O(4)	1.215(2)
C(10)-C(14)-C(13)	120.5(2)	C(4)-C(14)-C(13)	119.4(2)	N(1) -O(1)	1.226(2)	N(2) -O(3)	1.221(2)
O(6)-C(5)-C(11)	123.1(2)	O(6)-C(5)-C(6)	116.5(2)	N(1) -C(1)	1.462(3)	N(2) -C(8)	1.463(3)
C(11)-C(5)-C(6)	120.4(2)	O(8)-C(10)-C(11)	119.9(2)	C(10) -C(14)	1.454(3)	C(9) -C(13)	1.485(3)
O(8)-C(10)-C(14)	120.5(2)	C(11)-C(10)-C(14)	119.6(2)	C(10) -C(11)	1.458(3)	C(9) -C(12)	1.478(3)
O(5)-C(4)-C(14)	123.4(2)	O(5)-C(4)-C(3)	117.0(2)	C(14) -C(13)	1.406(3)	C(11) -C(5)	1.405(3)
C(14)-C(4)-C(3)	119.6(2)	N(1)-C(1)-C(13)	122.1(2)	C(14) -C(4)	1.405(3)	C(11) -C(12)	1.404(3)
N(1)-C(1)-C(2)	115.7(2)	C(13)-C(1)-C(2)	122.1(2)	C(5) -C(6)	1.390(3)	C(12) -C(8)	1.375(3)
O(7)-C(9)-C(12)	121.4(2)	O(7)-C(9)-C(13)	121.0(2)	C(4) -C(3)	1.392(3)	C(6) -C(7)	1.368(3)
C(12)-C(9)-C(13)	117.5(2)	O(2)-N(1)-O(1)	124.5(2)	C(1) -C(13)	1.374(3)	C(3) -C(2)	1.369(3)
O(2)-N(1)-C(1)	117.7(2)	O(1)-N(1)-C(1)	117.7(2)	C(1) -C(2)	1.387(3)	C(7) -C(8)	1.387(3)
H(6)-C(6)-C(5)	117.2(16)	H(6)-C(6)-C(7)	122.7(16)				
C(5)-C(6)-C(7)	120.1(2)	H(5)-C(7)-C(6)	124.4(15)				
H(5)-C(7)-C(8)	116.1(15)	C(6)-C(7)-C(8)	119.3(2)				
N(2)-C(8)-C(12)	121.4(2)	N(2)-C(8)-C(7)	116.0(2)				
C(12)-C(8)-C(7)	122.6(2)	H(3)-C(3)-C(4)	118.6(14)				
H(3)-C(3)-C(2)	120.6(14)	C(4)-C(3)-C(2)	120.9(2)				
C(14)-C(13)-C(1)	118.7(2)	C(14)-C(13)-C(9)	120.3(2)				
C(1)-C(13)-C(9)	120.7(2)	H(4)-C(2)-C(1)	118.8(13)				
H(4)-C(2)-C(3)	122.0(14)	C(1)-C(2)-C(3)	119.1(2)				

Table 3. Torsion angles.

H(2) - O(5) - C(4) - C(14)	6.8	C(9) - C(12) - C(8) - C(7)	-176.3	N(1) - C(1) - C(2) - H(4)	3.4
H(2) - O(5) - C(4) - C(3)	-174.5	C(8) - C(12) - C(9) - C(13)	-174	C(2) - C(1) - N(1) - O(2)	116.8
H(1) - O(6) - C(5) - C(11)	4.2	C(4) - C(14) - C(10) - O(8)	-2.4	C(2) - C(1) - N(1) - O(1)	-59.1
H(1) - O(6) - C(5) - C(6)	-174.9	C(10) - C(14) - C(4) - O(5)	-0.1	N(1) - C(1) - C(2) - C(3)	-173.6
O(4) - N(2) - C(8) - C(12)	-99.3	C(4) - C(14) - C(10) - C(11)	178.4	C(13) - C(1) - C(2) - H(4)	-179.5
O(4) - N(2) - C(8) - C(7)	83.2	C(10) - C(14) - C(4) - C(3)	-178.8	C(2) - C(1) - C(13) - C(14)	-4.9
O(3) - N(2) - C(8) - C(12)	83.3	C(13) - C(14) - C(10) - O(8)	177	C(2) - C(1) - C(13) - C(9)	168.7
O(3) - N(2) - C(8) - C(7)	-94.1	C(13) - C(14) - C(10) - C(11)	-2.3	C(13) - C(1) - C(2) - C(3)	3.5
C(12) - C(11) - C(5) - O(6)	-178.6	C(10) - C(14) - C(13) - C(1)	-177.2	O(7) - C(9) - C(13) - C(14)	163
C(5) - C(11) - C(12) - C(9)	176.7	C(10) - C(14) - C(13) - C(9)	9.2	O(7) - C(9) - C(13) - C(1)	-10.5
C(12) - C(11) - C(5) - C(6)	0.5	C(13) - C(14) - C(4) - O(5)	-179.5	C(12) - C(9) - C(13) - C(14)	-12.1
C(5) - C(11) - C(12) - C(8)	-1	C(4) - C(14) - C(13) - C(1)	2.1	C(12) - C(9) - C(13) - C(1)	174.4
C(12) - C(11) - C(10) - O(8)	179.3	C(4) - C(14) - C(13) - C(9)	-171.5	H(6) - C(6) - C(7) - H(5)	-4
C(12) - C(11) - C(10) - C(14)	-1.4	C(13) - C(14) - C(4) - C(3)	1.9	H(6) - C(6) - C(7) - C(8)	-179.7
C(10) - C(11) - C(12) - C(9)	-1.8	O(6) - C(5) - C(6) - H(6)	-0.9	C(5) - C(6) - C(7) - H(5)	176.1
C(10) - C(11) - C(12) - C(8)	-179.5	O(6) - C(5) - C(6) - C(7)	178.9	C(5) - C(6) - C(7) - C(8)	0.5
C(5) - C(11) - C(10) - O(8)	0.8	C(11) - C(5) - C(6) - H(6)	179.9	H(5) - C(7) - C(8) - N(2)	0.3
C(10) - C(11) - C(5) - O(6)	-0.1	C(11) - C(5) - C(6) - C(7)	-0.2	H(5) - C(7) - C(8) - C(12)	-177.1
C(5) - C(11) - C(10) - C(14)	-179.9	O(5) - C(4) - C(3) - H(3)	-1.9	C(6) - C(7) - C(8) - N(2)	176.3
C(10) - C(11) - C(5) - C(6)	179	O(5) - C(4) - C(3) - C(2)	177.9	C(6) - C(7) - C(8) - C(12)	-1.1
C(11) - C(12) - C(9) - O(7)	-166.7	C(14) - C(4) - C(3) - H(3)	176.9	H(3) - C(3) - C(2) - H(4)	3.6
C(11) - C(12) - C(9) - C(13)	8.4	C(14) - C(4) - C(3) - C(2)	-3.3	H(3) - C(3) - C(2) - C(1)	-179.5
C(11) - C(12) - C(8) - N(2)	-175.9	C(13) - C(1) - N(1) - O(2)	-60.3	C(4) - C(3) - C(2) - H(4)	-176.2
C(11) - C(12) - C(8) - C(7)	1.4	N(1) - C(1) - C(13) - C(14)	172	C(4) - C(3) - C(2) - C(1)	0.7
C(8) - C(12) - C(9) - O(7)	10.9	C(13) - C(1) - N(1) - O(1)	123.8		
C(9) - C(12) - C(8) - N(2)	6.4	N(1) - C(1) - C(13) - C(9)	-14.4		

References

- Abbenante, G., Fairlie, D. (2005). Protease inhibitors in the clinic. *Medicinal Chemistry*, 1:71–104.
- Allen, F., Kennard, O., Watson, D. (1987). Tables of bond lengths determined by X-ray and neutron diffraction. Part 1. Bond lengths in organic compounds. *Journal of the Chemical Society Perkin Transactions 2*. 12: S1-S19.
- Arias, C., Preugschat, F., Strauss, J. (1993). Dengue 2 virus NS2B and NS3 form a stable complex that can cleave NS3 within the helicase domain. *Virology*. 193: 888–899.
- Artsob, H., Gubler, D., Enria, D., Morales, M., Pupo, M., Bunning, M., Dudley, J., (2008). West Nile Virus in the New World: Trends in the Spread and Proliferation of West Nile Virus in the Western Hemisphere. *Zoonoses and Public Health*. 56: 357-369.
- Assenberg, R., Mastrangelo, E., Walter, T., Verma, A., Milani, M., Owens, R., Stuart, D., Grimes, J., Mancini, E. (2009). Crystal Structure of a novel conformational state of the Flavivirus NS3 protein: Implications for Polyprotein and Viral Replication. *Journal of Virology*. 83(24):12895-12906.
- Bajorath, J. (2002). Integration of virtual and high-throughput screening. *Nature Reviews: Drug Discovery*. 1: 882–894.
- Bazan, J., Fletterick, J. (1989). Detection of a trypsin-like serine protease domain in flaviviruses and pestiviruses. *Virology*. 171: 637–639.
- Berman, K., Kwo, P. (2009). Boceprevir, an NS3 Protease inhibitor of HCV. *Clinics in Liver Disease*. 13(3):429-439.
- Bosch, I., Herrera, F., Navarro, J., Lentino, M., Dupuis, A., Maffei, J., Jones, M., Fernandez, E., Perez, N., Perez-Eman, J., Guimaraes, A., Barrera, R., Valero, N., Ruiz, J., Velasquez, G., Martinez, J., Comach, G., Komar, N., Spielman, A. and Kramer, L., (2007). West Nile virus, Venezuela. *Emerging Infectious Diseases* 13(4): 651-653.
- Boonnak, N., Chantrapromma, S., Fun, H., Anjum, S., Ali, S., Rahman, A., Karalai, C. (2005). 3-(3,7-Dimethylocta-2,6-dienyloxy)-1,8-dihydroxy-6-methyl-9,10-anthraquinone. *Acta Crystallographica*. E61, 410-412.
- Brandstetter, H., Grams, F., Glitz, D., Lang, A., Huber, R., Bode, W., Krell, H-W., Engh, R. (2001). The 1.8-Å crystal structure of a matrix metalloproteinase 8-barbiturate inhibitor complex reveals a previously unobserved mechanism for collagenase substrate recognition. *Journal of Biological Chemistry*. 276:17405–17412.
- Brooks, B., Brooks III, C., Mackerell, C., Nilsson, L., Petrella, R., Roux, B., Won, Y., Archontis, G., Bartels, C., Boresch, S., Caflisch, A., Caves, L., Cui, Q., Dinner, A., Feig, M., Fischer, S., Gao, J., Hodoscek, M., Im, W., Kuczera, K., Lazaridis, T., Ma, J., Ovchinnikov, V., Paci, E., Pastor, R., Post, C., Pu, J., Schaefer, M., Tidor, B., Venable, R., Woodcock, H., Wu, X., Yang, W., York, D., Karplus, M. (2009). CHARMM: The Biomolecular simulation Program. *Journal of Computational Chemistry*: 30: 1545-1615.
- Brown, M., Huang, Y., Marshall, J., King, C., Hoskin, D., Anderson, R. (2009). Dramatic caspase-dependent apoptosis in antibody-enhanced dengue virus infection of human mast cells. *Journal of Leukocyte Biology*. 85: 71–80.
- Brown, M., Colclough, M. (1983). 1,8-Dinitro-4,5-dihydroxyanthraquinone, *C₁₄H₆N₂O₈. *Acta Crystallographica*. C(39): 300-302.
- Brunker. (2001). SAINT-Plus. Version 1.6. Brunker AXS Inc., Madison, WI, USA.

- Brunker. (2002). PROTEUM. Version 1.38. Brunker AXS Inc., Madison, WI, USA.
- Cao, L., Wang, Y., Gao, M., Zhou, B. (2007). 1,4-Dimethoxyanthracene-9,10-dione. *Acta Crystallographica*. E63, 1876–1877.
- Chanprapaph, S., Saparpakorn, P., Sangma, C., Niyomrattanakit, P., Hannongbua, S., Angsuthanasombat, C., Katzenmeier, G. (2005). Competitive inhibition of the dengue virus NS3 serine protease by synthetic peptides representing polyprotein cleavage sites. *Biochemical and Biophysical Research Communications*. 330: 1237-1246.
- Chambers. T., Hahn, C., Galler, R., Rice, C. (1990). Flavivirus genome organization, expression, and replication. *Annual Review of Microbiology*. 44:649-688.
- Chambers, T., Grakoui, A., Rice, C. (1991). Processing of the yellow fever virus nonstructural polyproteins: A catalytically active NS3-proteinase domain and NS2b are required for cleavages at dibasic sites. *Journal of Virology*. 65: 6042-6050.
- Chambers, T., Nestorowicz, A., Amberg, S., Rice, C. (1993). Mutagenesis of the yellow fever virus NS2B protein: effects on proteolytic processing, NS2B-NS3 complex formation, and viral replication. *Journal of Virology*. 67(11):6797-6807.
- Chambers, T., Droll, D., Tang, Y., Liang, Y., Ganesh, V., Murthy, K., Nickells, M. (2005). Yellow fever virus NS2B-NS3 protease: characterization of charged-to-alanine mutant and revertant viruses and analysis of polyprotein-cleavage activities. *Journal of General Virology*. 86(5):1403-1413.
- Chapkanov, A. (2007). 1,8-Dinitro-4,5-dihydroxyanthraquinone-Solid-state linear-polarized IR-spectroscopy of oriented colloid particles as a suspension in nematic liquid crystal. *Spectrochimica Acta Part A*. 71: 1481–1484.
- Chappell, K., Nall, T., Stoermer, M., Fang, N., Tyndall, J., Fairlie, D., Young, P. (2005) Site-directed mutagenesis and kinetic studies of the West Nile Virus NS3 protease identify key enzyme–substrate interactions. *Journal of Biological Chemistry*. 280: 2896–2903.
- Cheng, T., Li, X., Li, Y., Liu, Z., Wang, R.. (2009). Comparative Assessment of Scoring Functions on a Diverse Test Set. *Journal of Chemical Information and Modeling*. 49: 1079–1093
- Cook, S., Holmes E. (2006). A multigene analysis of the phylogenetic relationships among the flaviviruses (Family: Flaviviridae) and the evolution of vector transmission. *Archives of Virology*. 151: 309–325.
- DeLano, W. (2002). The PyMOL Molecular Graphics System. DeLano Scientific, Palo Alto, CA, USA.
- Dobler, G. (2009). Zoonotic tick-borne flaviviruses. *Veterinary Microbiology*. 140(3-4): 221-228.
- Dong, H., Zhang, B., Shi, P. (2008). Flavivirus methyltransferase: a novel antiviral target. *Antiviral Research*. 80(1): 1-10.
- Ebel, G., Dupuis, A., Ngo, K., Nicholas, D., Kauffman, E., Jones, A., Young, D., Maffei, J., Shi, P., Bernard, K., Kramer, L. (2001). Partial genetic characterization of West Nile virus strains, New York State. *Emerging and Infectious Diseases*. 7:650–653.
- Erbel, P., Schiering, N., D`Arcy, A., Renatus, M., Kroemer, M., Lim, P., Yin, Z., Keller, H., Vasudevan, S., Hommel, U. (2006). Structural basis for the activation of flaviviral NS3 proteases from dengue and West Nile virus. *Nature Structural & Molecular Biology*. 13: 372-373.
- Ezaby, M., Salem, T., Zewail, A., Issa, R. (1970). Spectral studies of some hydroxy-derivatives of anthraquinones. *Journal of the Chemical Society B: Physical Organic*. 7: 1293-1296.

- Falgout, B., Pethel, M., Zhang, Y., Lai, C. (1991). Both nonstructural proteins NS2b and NS3 are required for the proteolytic processing of dengue virus nonstructural proteins. *Journal of Virology*. 65: 2467–2475.
- Falgout, B., Miller, R., Lai, C. (1993). Deletion analysis of dengue virus type 4 nonstructural protein NS2B: identification of a domain required for NS2B-NS3 protease activity. *Journal of Virology*. 67: 2034–2042.
- Fariña, F., Molina, M., Noheda, P., Paredes, C. (1992). Polycyclic Hydroxyquinones. XXVII. Tautomerism in 1,4-Dihydroxy-9,10-anthraquinone Monimines. Cycloaddition Reactions of Their 1,4-Anthraquinoid Tautomers. *Tetrahedron*. 48(39):8437-8450.
- Farrugia, L.(1997). *ORTEP*-3 for Windows - a version of *ORTEP*-III with a Graphical User Interface (GUI). *Journal of Applied Crystallography*. 30: 565.
- Farrugia, L. (1999). *WinGX* suite for small-molecule single-crystal crystallography. *Journal of Applied Crystallography*. 32: 837-838.
- Ganesh, V., Muller, N., Judge, K., Luan, C., Padmanabhan, R., Murthy, K. (2005). Identification and characterization of nonsubstrate based inhibitors of the essential dengue and West Nile virus proteases. *Bioorganic & Medicinal Chemistry Letters*. 13: 257-264.
- Gaunt, M., Sall, A., de Lamballerie, X., Falconar, A., Dzhibanian, T., Gould, E. (2001). Phylogenetic relationships of flaviviruses correlate with their epidemiology, disease association and biogeography. *Journal of General Virology*. 82(8): 1867-1876.
- Ghosh, A., Dawson, Z., Mitsuya, H. (2007). Darunavir, a conceptually new HIV-1 protease inhibitor for the treatment of drug-resistant HIV. *Bioorganic & Medicinal Chemistry Letters*. 15(24):7576-7580.
- Gohlke, H., Klebe, G. (2002). Approaches to the description and prediction of the binding affinity of small-molecule ligands to macromolecular receptors. *Angewandte Chemie (International ed. in English)*. 41: 2644–2676.
- Gollins, S., Porterfield, J. (1986). pH-dependent fusion between the flavivirus West Nile and liposomal model membranes. *Journal of General Virology*. 67:157–166.
- Goncalvez, A., Engle, R., St Claire, M., Purcell, R., Lai, C. (2007). Monoclonal antibody-mediated enhancement of dengue virus infection in vitro and in vivo and strategies for prevention. *Proceedings of the National Academy of Sciences of the United States of America*. 104: 9422–9427.
- Gorbalenya, A., Donchenko, A., Koonin, E., Blinov, V. (1989). N-terminal domains of putative helicases of flavi- and pestiviruses may be serine proteases. *Nucleic Acids Research*. 17: 3889–3897.
- Grant, B. (2009). Crystallographer faked Data. *The Scientist* [online]. 18 Dec. 2009. Feb 2010. <<http://www.the-scientist.com/blog/display/56226/>>.
- Gould, E., Solomon, T. (2008). Pathogenic flaviviruses. *Lancet*. 371:500–509.
- Gubler, D. (1994). Perspectives on the prevention and control of dengue hemorrhagic fever. *Gaoxiong Yi Xue Ke Xue Za Zhi*. Suppl. 10: S15–S18.
- Gubler, D. (1997). Dengue and dengue hemorrhagic fever; its history and resurgence as a global public health problem. In: Gubler, D., Kuno, G., Editors. *Dengue and Dengue Hemorrhagic Fever*. London, United Kingdom: CAB International Press: 1–22
- Gubler, D. (2002)a. Epidemic dengue/dengue hemorrhagic fever as a public health, social and economic problem in the 21st century. *Trends Microbiology*. 10(2): 100–103

- Gubler, D. (2002)b. The global emergence/resurgence of arboviral diseases as public health problems. *Archives of Medical Research*. 33: 330–342.
- Halstead, S., Epidemiology of dengue and dengue hemorrhagic fever. In: Gubler, D., Kuno, G, Editors. *Dengue and Dengue Hemorrhagic Fever*. London, United Kingdom: CAB International Press: 23–44.
- Hayes, E., Komar, N., Nasci, R., Montgomery, S., O’Leary, D., Campbell, G. (2005). Epidemiology and transmission dynamics of West Nile virus disease. *Emerging and Infectious Diseases*. 11: 1167–1173
- Huang, H., Chang, J., Tung, S., Wu, R., Foegh, M., Chu, S. (1992). Immunosuppressive effect of emodin, a free radical generator . *European Journal of Pharmacology*. 211: 359–364.
- Huang, H., Chu, S., Chao, P. (1991). Vasorelaxants from Chinese herbd, emodin and scoparone, possess immunosuppressive properties. *European Journal of Pharmacology*. 198: 211–213.
- Huang, H., Lee, C., Chao, P., Chen, C., Chu, S. (1991). Vasorelaxant effect of emodin, an anthraquinone from a Chinese herb. *European Journal of Pharmacology*. 205: 289–294.
- Huey, R., Goodsell, D., Morris, G., Olson, A. (2004). Grid-based hydrogen bond potentials with improved directionality. *Lett. Drug Des. Discovery*. 1: 178–183.
- Huey, R., Morris, G., Olson, A., Goodsell, D. (2007). A semiempirical free energy force field with charge-based desolvation. *Journal of Computational Chemistry*. 28: 1145–1152.
- Irwin, J., Shoichet, B. (2005). ZINC - A Free Database of Commercially Available Compounds for Virtual Screening. *Journal of Chemical Information and Modeling*. 45(1). 177-182.
- Ishak, R., Tovey, D., Howard C. (1988). Morphogenesis of yellow fever virus 17D in infected cell cultures. *Journal of General Virology*. 69:325-335.
- Ismail, N., Osman, C., Ahmed, R., Awang, K., Ng, S. (2009)a. 1,3-Dihydroxy-2-methoxymethyl-9,10-anthraquinone from *Rennellia elliptica* Korth. *Acta Crystallographica*. E(65): 1433-1434.
- Ismail, N., Osman, C., Ahmed, R., Awang, K., Ng, S. (2009)b. 1-Hydroxy-2-methoxy-6-methyl-9,10-anthraquinone from *Rennellia elliptica* Korth. *Acta Crystallographica*. E(65): 1435.
- IUPAC. Compendium of Chemical Terminology, 2nd ed. (the "Gold Book"). Compiled by A. D. McNaught and A. Wilkinson. Blackwell Scientific Publications, Oxford (1997). XML on-line corrected version: <http://goldbook.iupac.org> (2006) created by M. Nic, J. Jirat, B. Kosata; updates compiled by A. Jenkins.
- Jang, Y., Goddard, WA III, Noyes, K., Sowers, L., Hwang, S., Chung, D. (2002). First principles calculations of the tautomers and pK(a) values of 8-oxoguanine: implications for mutagenicity and repair. *Chemical Research in Toxicology*. 15:1023–1035.
- Kalliokoski, T., Heikki, S., Lahtela-Kakkonen, M., Poso, A. (2009). The Effect of Ligand-Based Tautomer and Protomer Prediction on Structure-Based Virtual Screening. *Journal of Chemical Information and Modeling*. 49(12):2742-2748.
- Kim, E., Baker, C., Dwyer, M., Murcko, B., Rao, B., Tung, R., Navia, M. (1995). Crystal Structure of HIV-1 Protease in Complex with VX-478, a Potent and Orally Bioavailable Inhibitor of the Enzyme. *Journal of American Chemistry*. 117: 1181-1182.
- Kitchen, D., Decornez, H., Furr, J., Bajorath, J. (2004). Docking and Scoring in Virtual Screening for Drug Discovery: Methods and Applications. *Nature Reviews: Drug Discovery*. 3(11): 935-949.

- Komar, N., Clark, G. (2006). West Nile virus activity in Latin America and the Caribbean. *Pan American Journal of Public Health*. 19:112–117.
- Kontoyianni, M., Sokol, G., McClellan, L. (2004). Evaluation of Library Ranking Efficacy in Virtual Screening. *Journal of Computational Chemistry*. 26(1):11-22.
- Knox, A., Meegan, M., Carta, G., Lloyd, D. (2005). Considerations in Compound Database Preparation - “Hidden” Impact on Virtual Screening Results. *Journal of Chemical Information and Modeling*. 45: 1908–1919.
- Knox, J., Ma, N., Yin, Z., Patel, S., Wang, W., Chan, W., Ranga Rao, K., Wang, G., Ngew, X., Patel, V., Beer, D., Lim, S., Vasudevan, S., Keller, T. (2006). Peptide Inhibitors of West Nile NS3 Protease: SAR Study of Tetrapeptide Aldehyde Inhibitors. *Journal of Medicinal Chemistry*. 49: 6585–6590.
- Kramer, L., Styer, L., Ebel, G. (2008). A Global Perspective on the Epidemiology of West Nile Virus. *Annual Review of Entomology*. 53: 61-81.
- Kuno, G., Chang, G., Tsuchiya, K., Karabatsos, N., Cropp, C. (1998). Phylogeny of the genus *Flavivirus*. *Journal of Virology*. 72 (1):73-83.
- Kubicki, M. (2004). Two tautomers in one crystal: 4(5)-nitro-5(4)-methoxyimidazole. *Acta Crystallographica*. B(60): 191-196.
- Kuntz, I., Blaney, J., Oatley, S., Langridge, R., Ferrin, T. (1982). A geometric approach to macromolecule–Ligand interactions. *Journal of Molecular Biology*. 161, 269–288.
- Langer, T., Hoffmann, R. (2002). Virtual screening: an effective tool for lead structure discovery. *Current Pharmaceutical Design*. 7: 509–527.
- Lamarre, D., Anderson, P., Bailey, M., Beaulieu, P., Bolger, G., Bonneau, P., Bös, M., Cameron, D., Cartier, M., Cordingley, M., Faucher, A., Goudreau, N., Kawai, S., Kukolj, G., Lagacé, L., LaPlante, S., Narjes, H., Poupart, M., Rancourt, J., Sentjens, R., St. George, R., Simoneau, B., Steinmann, G., Thibeault, D., Tsantrizos, Y., Weldon, S., Yong, C., Llinàs-Brunet, M. (2003). An NS3 protease inhibitor with antiviral effects in humans infected with hepatitis C virus. *Nature*. 426: 186–189.
- Lanciotti, R., Roehrig, J., Deubel, V., Smith, J., Parker, M., Steele, K., Crise, B., Volpe, E., Crabtree, B., Scherret, H., Hall, A., MacKenzie, S., Cropp, B., Panigrahy, B., Ostlund, E., Schmitt, B., Malkinson, M., Banet, C., Weissman, J., Komar, N., Savage, M., Stone, W., McNamara, T., Gubler, D. (1999). Origin of the West Nile virus responsible for an outbreak of encephalitis in the northeastern United States. *Science*. 286:2333–2337
- Lescar, J., Luo, D., Xu, T., Sampath, A., Lim, S., Canard, B., Vasudevan, S. (2008). Towards the design of antiviral inhibitors against flaviviruses: the case for the multifunctional NS3 protein from Dengue virus as a target. *Antiviral Research*. 80(2):94-101
- Leung, D., Schroder, K., White, H., Fang, X., Stoermer, J., Abbenante, G., Martin, J., Young, R., Fairlie, P. (2001). Activity of recombinant dengue 2 virus NS3 protease in the presence of a truncated NS2B co-factor, small peptide substrates, and inhibitors. *Journal of Biological Chemistry*. 276: 45762–45771.
- Leung, D.; Abbenante, G.; Fairlie, D. (2000). Protease inhibitors: current status and future prospects. *Journal of Medical Chemistry*. 43, 305–341.
- Leyssen, P., De Clercq, E., Neyts, J. (2000). Perspectives for the treatment of infections with *Flaviviridae*. *Clinical Microbiology Reviews*. 13: 67-82.
- Li, Z., Khaliq, M., Zhou, Z., Post, C., Kuhn, R., Cushman, M. (2009). Design, Synthesis, and Biological Evaluation of Antiviral Agents Targeting *Flavivirus* Envelope Proteins. *Journal of Medicinal Chemistry*. 51 (15): 4660-4671.

Li, H., Clum, S., You, S., Ebner, K., Padmanabhan, R., (1999). The serine protease and RNA-stimulated nucleoside triphosphatase and RNA helicase functional domains of dengue virus type 2 NS3 protein converge within a region of 20 amino-acids. *Journal of Virology*. 73: 3108–3116.

Li, J., Lim, S., Beer, D., Patel, V., Wen, D., Tumanut, C., Tully, D., Williams, J., Jiricek, J., Priestle, P., Harris, J., Vasudevan, S., (2005). Functional profiling of recombinant NS3 proteases from all four serotypes of dengue virus using tetrapeptide and octapeptide substrate libraries. *Journal of Biological Chemistry* 280: 28766–28774.

LigPrep; Schrödinger LLC: New York, NY. <<http://www.schrodinger.com/products/14/10/>>. (Accessed 02/2010).

Lipinski, C., Lombardo, F., Dominy, B., Feeney, P. (2001). Experimental and computational approaches to estimate solubility and permeability in drug discovery and development settings. *Advanced Drug Delivery Reviews*. 46(1-3): 3-26.

Lindenbach, B., Thiel, H., Rice, C. (2007). *Flaviviridae: the viruses and their replication*. In Knipe, D., Howley, P., (ed.), *Fields Virology*, 5th Edition, 5th ed. Lippincott-Raven Publishers, Philadelphia, PA.

Liu, Z., Jiao, Y. (2009). 2-Hydroxy-1-methoxyanthraquinone monohydrate. *Acta Crystallographica*. E(65): 1523.

Malet, H., Masse, N., Selisko, B., Romette, J., Alvarez, K., Guillemot, J., Tolou, H., Yap, T., Vasudevan, S., Lescar, J., Canard, B. (2008). The flavivirus polymerase as a target for drug discovery *Antiviral Research*. 80 (1):23-35.

Martin, Y. (2009). Let's not forget tautomers. *Journal of Computationally Aided Molecular Design*. 23(10): 693-704.

Martin, J., Pierson, T., Hubka, S., Rucker, S., Gordon, I., Enama, M., Andrews, C., Xu, Q., Davis, B., Nason, M., Fay, M., Koup, R., Roederer, M., Bailer, R., Gomez, P., Mascola, J., Chang, G., Nabel, G., Graham, B. (2007). A West Nile virus DNA vaccine induces neutralizing antibody in healthy adults during a phase 1 clinical trial. *Journal of Infectious Diseases*. 196:1732–1740.

Marzocchi, M., Mantini, A., Casu, M., Smulevich, G. (2003). Intramolecular hydrogen bonding and excited state proton transfer in hydroxyanthraquinones as studied by electronic spectra, resonance Raman scattering, and transform analysis. *Journal Chemical Physics*. 108 (2): 534-549.

Milletti F, Storchi L, Sforza G, Cross S, Cruciani G. (2009). Tautomer enumeration and stability prediction for virtual screening on large chemical databases. *Journal of Chemical Information and Modeling*. 49, 68–75.

Monath T, Liu J, Kanasa-Thanan N, Myers G, Nichols R, Deary A, McCarthy K, Johnson C, Ermak T, Shin S, Arroyo J, Guirakhoo F, Kennedy J, Ennis F, Green S, Bedford P. (2006). A live, attenuated recombinant West Nile virus vaccine. *Proceedings of the National Academy of Sciences of the United States of America*. 103:6694–6699.

Mooibroek, T., Gamez, P., Reedijk, J. (2008). Lone Pair- π interactions: a supramolecular bond? *CrystEngComm*. 10: 1501-1515.

Morales, M., Barrandeguy, M., Fabbri, C., Garcia, J., Vissani, A., Trono, K., Gutierrez, G., Pigretti, S., Menchaca, H., Garrido, N., Taylor, N., Fernandez, F., Levis, S., Enría, D. (2006). West Nile virus isolation from equines in Argentina. *Emerging Infectious Diseases*. 12:1559–1561

Mostashari F., Bunning M., Kitsutani P., Singer, D., Nash, D., Cooper, M., Katz, N., Liljebjelke, K., Biggerstaff, B., Fine, A., Layton, M., Mullin, S., Johnson, A., Martin, D., Hayes, E., Campbell, G.. (2001). Epidemic West Nile Encephalitis, New York, 1999: results of a household-based seroepidemiological survey. *Lancet*. 358(9278):261-264.

Murray, C., Jones, C., Rice, C. (2008). Architects of assembly: roles of Flaviviridae non-structural proteins in virion morphogenesis. *Nature Reviews in Microbiology*. 6:699-708.

Murthy, H., Clum, S., Padmanabhan, R., (1999). Dengue virus NS3 serine protease. Crystal structure and insights into interaction of the active site with substrates by molecular modeling and structural analysis of mutational effects. *Journal of Biological Chemistry*. 274: 5573–5580.

Nall, T., Chappell, K., Stoermer, M., Fang, N., Tyndall, J., Young, P., Fairlie, D. (2004). Enzymatic characterization and homology model of a catalytically active recombinant West Nile virus NS3 protease. *Journal of Biological Chemistry*. 279: 48535–48542.

Ng, S., Abdul, I., Fun, H., Boonsri, S., Chantrapromma, S., Prawat, U. (2005) 1,3-Dihydroxy-2-methyl-9,10-anthraquinone. *Acta Crystallographica*. E61: 3656–3658

Njoroge F., Chen K., Shih N., Piwinski J. (2008). Challenges in modern drug discovery: a case study of boceprevir, an HCV protease inhibitor for the treatment of hepatitis C virus infection. *Accounts of Chemical Research*. 41(1):50-59.

Oellien, F., Cramer, J., Beyer, C., Ihlenfeldt, W., Selzer, P. (2006). The impact of tautomer forms on pharmacophore-based virtual screening. *Journal of Chemical Information and Modeling*. 46:2342–2354.

Ojwang, J., Ali, S., Smee, D., Morrey, J., Shimasaki, C., Sidwell, R. (2005). Broad-spectrum inhibitor of viruses in the Flaviviridae family. *Antiviral Research* 68:49-55.

Park, M., Kwon, H., Sung, M. (2009). Evaluation of aloin and aloe-emodin as anti-inflammatory agents in aloe by using murine macrophages. *Bioscience, Biotechnology, and Biochemistry*. 73(4):828-832.

Permana, D., Lajis, H., Othman, G., Ali, M., Aimi, N., Kitajima, M., Takayama, H. (1999). Anthraquinones from *hedyotis herbacea*. 62(10):1430-1431.

Pilot, P. (2008). Scitegic chemistry components Accelrys. <http://accelrys.com/products/datasheets/chemistry-component-collection.pdf>.

Pospisil, P., Ballmer, P., Scapozza, L., Folkers, G. (2003). Tautomerism in Computer-Aided Drug Design. *Journal of Receptors and Signal Transduction Research*. 23: 361–371.

Prakash, A. (1967). Refinement of the crystal structure of anthraquinone. *Acta Crystallographica*. 22: 439.

Preugschat, F., Yao, C., Strauss, J. (1990). In vitro processing of dengue virus type 2 nonstructural proteins NS2A, NS2B, and NS3. *Journal of Virology*. 64: 4364–4374.

Proudfoot, J. (2005). The evolution of synthetic oral drug properties. *Bioorganic & Medicinal Chemistry Letters*. 15:1087–1090.

Puig-Basagoiti, F., Qing, M., Dong, H., Zhang, B., Zou, G., Yuan, Z., Shi, P. (2009). Identification and characterization of Inhibitors of West Nile virus. *Antiviral Research*. 83(1): 71-79.

QUAKPAC; OpenEye Scientific Software: Santa Fe, NM.
<<http://www.eyesopen.com/products/applications/quacpac.html>>. (Accessed 02/2010).

Rao, G., Bhatnagar, S., Ahuja, V. (2002). Structure-based design of a novel peptide inhibitor of HIV-1 integrase: a computer modeling approach. (1):31-38.

Sanner, M. (1999). Python: a programming language for software integration and development. *Journal of Molecular Graphics and Modeling*. 17(1): 57–61.

- Sampath, A., Padmanabhan, R. (2009). Molecular Target for Flavivirus drug discovery. *Antiviral Research*. 81: 6–15
- Sheldrick, G. (2008). A short history of *SHELX*. *Acta Crystallographica*. A(64): 112-122.
- Shiryaev, S., Kozlov, I., Ratnikov, B., Smith, J., Lebl, M., Strongin, A. (2007). Cleavage preferences distinguishes the two-component NS2B-NS3 serine proteases of Dengue and West Nile viruses. *Journal of Biochemistry*. 401:743-752.
- Shuangsoo, D., Zhengguo, Z., Yunru, C., Xin, Z., Baofeng, W., Lichao, Y., Yan'an, C. (2006). Inhibition of the replication of hepatitis B virus in vitro by emodin. *Medical Science Monitor*. 12(9): BR302-306.
- Shi, P. (2002). Strategies for the identification of inhibitors of West Nile virus and other flaviviruses. *Current Opinion in Investigational Drugs*. (11):1567-1573.
- Smulevich, G., Foggi, P., Feis, A., Marzocchi, P. (1987). Fluorescence excitation and emission of 1,8-dihydroxyanthraquinone -d₀ and -d₂ in n-octane at 10 K. *Journal of Chemical Physics*. 87: 5664-5669.
- Stephenson, J. (2006). Developing vaccines against flavivirus diseases: past success, present hopes and future challenges. *Novartis Found Symposium*. 277: 193–201.
- Ryan, M., Monaghan, S., Flint, M. (1998). Virus encoded proteinases of the Flavivirus. *Journal of General Virology*. 79(5): 947-959.
- Tafeenko, V., Popov, S., Medvedev, S. (1987). “Conjugated molecular systems with the hydroxy ketone unit, the structure of Alizarine Black-Blue-B” [translated title]. *Doklady Akademii Nauk*. 294(5): 1136-1139.
- Tafeenko, V., Popov, S., Medvedev, S. (1991). X-ray structural investigation of the base of the dye Alizarine Blue Black B at room temperature and at -150°C. *Journal of Structural Chemistry*. 32(3): 406-409. Translated from: *Zhurnal Strukturnoi Khimii*. 32(3): 106-109.
- Tafeenko, V., Bogdan, T., Medvedev, S., Kozyrev, A., Popov, S. (1991). Crystal and molecular structures of dibutyl alizarin blue-black B. *Journal of Structural chemistry*. 32(5): 774-776. Translated from: *Zhurnal Strukturnoi Khimii*. 32(5): 169-171.
- TAUTOMER; Molecular Networks GmbH: Erlangen, Germany. <http://www.molecular-networks.com/products/tautomer> (Accessed 02/2010)
- Taylor, A., Johnson, W., Czerwinski, R., Li, H., Hackert, M., Whitman, C. (1999). Crystal structure of macrophage migration inhibitory factor complexed with (E)-2-fluoro-p-hydroxycinnamate at 1.8 Å resolution: implications for enzymatic catalysis and inhibition. *Biochemistry* 38:7444–7452.
- Thiel, H., Collett, M., Gould, E., Heinz, F., Houghton, M., Meyers, G., Purcell, R. and Rice, C. (2006) . Index of Viruses - Flaviviridae. In: ICTVdB - The Universal Virus Database, version 4. Büchen-Osmond, C (Ed), Columbia University, New York, USA. http://www.ncbi.nlm.nih.gov/ICTVdb/Ictv/fs_index.htm
- Thaisrivongs, S, Strohbach, J. (1999). Structure-based discovery of Tipranavir disodium (PNU-140690E): a potent, orally bioavailable, nonpeptidic HIV protease inhibitor. *Biopolymers*. 51(1):51-58.
- Tomori, O. (2004). Yellow fever: the recurring plague. *Critical Reviews in Clinical Laboratory Sciences*. 41:391-427.
- Tomlinson S, Malmstrom R, Watowich S., (2009)a. New Approaches to Structure-based discovery of dengue protease inhibitors. *Infectious disorders drug targets*. 9(3): 327-343.

Tomlinson, S., Malmstrom, R., Russo, A., Mueller, N., Watowich, S. (2009)b. Structure-based discovery of dengue virus protease inhibitors. *Antiviral Research Journal*. 82(3):110-114.

Tsai, T. (2000). New initiatives for the control of Japanese encephalitis by vaccination: minutes of a WHO/CVI meeting, Bangkok, Thailand, 13-15 October 1998. *Vaccine*. 18: Suppl. 2:1-25.

Ulický, L., Kettmann, V., Soldánová, J., Betiba, V. (1987). Structure of 1-Acetyl-2,4,5,7-tetrahydroxy-9,10-anthraenedione. *Acta Crystallographica*. C43: 335-339.

United States. National Institutes of Health. (2009). Biodefense: NIAID Category A, B, and C Priority Pathogens. 10 Dec. 2009. 3 Mar 2010. <<http://www3.niaid.nih.gov/topics/BiodefenseRelated/Biodefense/research/CatA.htm>>.

Vaughn, D., Green, S., Kalayanarooj, S., Innis, B., Nimmannitya, S., Suntayakorn, S., Endy, T., Raengsakulrach, B., Rothman, A., Ennis, F., Nisalak, A. (2000). Dengue viremia titer, antibody response pattern, and virus serotype correlate with disease severity. *Journal of Infectious Diseases*. 181: 2–9.

van der Meulen, K. Penssaert, M. and Nauwynck, H. (2005). West Nile virus in the vertebrate world. *Archives of Virology*. 150 (4), 637–657.

Virus Taxonomy, Classification and Nomenclature of Viruses, 8th ICTV Report of the International Committee on Taxonomy of Viruses. Fauquet, C., Mayo, M., Maniloff, J., Desselberger, U., Ball, L. (EDS) (2005). Elsevier/Academic Press: 1259.

Wallace A, Laskowski R, Thornton J. (1995). LIGPLOT: a program to generate schematic diagrams of protein-ligand interactions. *Protein Engineering*. 8:127–134.

Warren, G., Andrews, C., W., Capelli, A., Clarke, B., Lalonde, J., Lambert, M., Lindvall, M., Nevins, N., Semus, S., Senger, S., Tedesco, G., Wall, I., Woolven, Peishoff, C., Head, M. (2006). A Critical Assessment of Docking Programs and Scoring functions. *Journal of Medicinal Chemistry*. 49: 5912-5931.

Walters, W., Stahl, M., Murcko, M. (1998). Virtual screening - an overview. *Drug Discovery Today* 3(4), 160–178.

Watkin, D., Pearce, L., Prout, C. (1993). CAMERON . A Molecular Graphics Package. Chemical Crystallography Laboratory, University of Oxford, England.

Weaver, S., Reisen, W., (2010). Present and Future Arboviral Threats, *Antiviral Research*. 85(2):328-345.

Webster, D., Farrar, J., Rowland-Jones, S., (2009). Progress towards a dengue vaccine. *Lancet Infectious Diseases* 9: 678-687.

Weissenböck, H., Hubálek, Z., Bakonyi, T., Nowotny, N. (2009). Zoonotic mosquito-borne flaviviruses: Worldwide presence of agents with proven pathogenicity and potential candidates of future emerging diseases. *Veterinary Microbiology*. 140(3-4): 271-280.

Weisberg, S., Jacobson, M. (2009). Telaprevir: hope on the horizon, getting closer. *Clinics in Liver Disease*. (3):441-452.

Wengler, G., Czaya, G., Farber, P., Hegemann, J. (1991). In vitro synthesis of West Nile virus proteins indicates that the amino-terminal segment of the NS3 protein contains the active centre of the protease which cleaves the viral polyprotein after multiple basic amino acids. *Journal of General Virology*. 72: 851–858

World Health Organization. (2000). Strengthening implementation of the global strategy for dengue fever/dengue haemorrhagic fever prevention and Control. Report of the Informal Consultation. WHO, Geneva, Switzerland. 18–20 Oct. 1999.

- Wlodawer, A. (2002). Rational approach to AIDS drug design through structural biology. *Annual Review Medicine*. 53:595-614.
- Wu, T., Lin, D., Shi, L., Damu, A., Kuo, P., Kuo, Y. (2003). Cytotoxic anthraquinones from the stems of *Rubia wallichiana* Decne. *Chemical & Pharmaceutical Bulletin*. 51(8):948-950.
- Yatsenko, A., Paseshnichenko, K., Popov, S. (1996). 4,5-Dihydroxy-1-nitroanthracene-9,10-dione. *Acta Crystallographica*. C(52): 3258-3260.
- Yatsenko, A., Tafeenko, V., Zhukov, S., Medvedev, S., Popov, S. (1997). Structures and Tautomeric Interconversions of Anthraquinone Imine Derivatives. *Structural Chemistry*. 8(3): 197-204.
- Yin, Z., Patel, S., Wang, W., Wang, G., Chan, W., Rao, K., Alam, J., Jeyaraj, D., Ngew, X., Patel, V., Beer, D., Lim, S., Vasudevan, S., Keller, T. (2006)a. Peptide inhibitors of dengue virus NS3 protease. Part 1: Warhead. *Bioorganic & Medicinal Chemistry Letters*. 16: 36-39.
- Yin, Z., Patel, S., Wang, W., Chan, W., Rao, K., Wang, G., Ngew, X., Patel, V., Beer, D., Knox, J., Ma, N., Ehrhardt, C., Lim, S., Vasudevan, S., Keller, T. (2006)b. Peptide inhibitors of dengue virus NS3 protease. Part 2: SAR study of tetrapeptide aldehyde inhibitors. *Bioorganic & Medicinal Chemistry Letters*. 16: 40-43.
- Yusof, R., Clum, S., Wetzel, M., Murthy, H., Padmanabhan, R., (2000). Purified NS2B/NS3 Serine Protease of Dengue Virus Type 2 Exhibits Cofactor NS2B Dependence for Cleavage of Substrates with Dibasic Amino Acids in vitro. *Journal of Biological Chemistry*. 275: 9963–9969.
- Zain, S., Ng, S. (2005). 1,8-Dihydroxyanthraquinone. *Acta Crystallographica*. E61, 2921–2923.
- Zhang, L., Mohan, P., Padmanabhan, R., (1992). Processing and localization of dengue virus type-2 polyprotein precursor NS3-NS4a-NS4b-NS5. *Journal of Virology*. 66: 7549–7554.
- Zhu, L., Zhao, Z., Yu, S. (2008). 2-Hydroxy-1,6,7,8-tetramethoxy-3-methylantraquinone. E64: 371.

Vita

Jeff A. Borgeson was born in Kansas City, Missouri on August 23, 1983 to Jeff L. Borgeson and Connie L. Borgeson. He graduated from The University of Missouri in 2006 where he earned a Bachelor's of Science Degree in Physics and minored in Mathematics. Shortly thereafter he joined the PhD program at the University of Texas Medical Branch in August of 2007. In 2008, he began his master's research in the lab of Dr. Stanley Watowich in the Biophysics, Structural, and Computational program within the Biochemistry and Molecular Biology Department.

Jeff served as the Graduate Student Organization President (2008-2009), represented the graduate school student body as a senator to the Student Government Association (2008-2009), and served on several committees.

Jeff and Michelle Panethiere (former surname) were married in 2008 and currently reside at 2008 Leisure Lane, League City, Texas.

This Dissertation was written and typed by Jeff Borgeson.

Education

B.S., December 2006, University of Missouri, Columbia, Missouri.