IN SILICO STUDIES ON ZIKA NS3 HELICASE: BEDROCK FOR ANTIVIRAL DRUG DESIGN

SOFIAT YETUNDE OGUNTADE

216075817

A thesis submitted to the College of Health Sciences, University of KwaZulu-Natal, Westville, in fulfillment of the requirements of the degree of Master of Pharmacy

Supervisor

Prof Mahmoud Soliman

KwaZulu-Natal

2017

IN SILICO STUDIES ON ZIKA NS3 HELICASE: BEDROCK FOR ANTIVIRAL DRUG DESIGN

2017

SOFIAT YETUNDE OGUNTADE

216075817

A thesis submitted to the College of Health Sciences, University of KwaZulu-Natal, Westville Campus, in fulfilment of the requirements of the degree of

Master of Pharmacy

in

Pharmacy (Pharmaceutical Chemistry).

This is a thesis in which the chapters are written as a set of discrete research publications, with an overall introduction and conclusion.

This is to certify that the contents of this thesis are the original research work of Mrs Sofiat Yetunde Oguntade.

As the candidate's supervisor, I have approved this thesis for submission.

Supervisor:

Signed: ---------- Name: **Prof. Mahmoud E. Soliman** Date: ----------

ABSTRACT

Zika virus is a re-emerging infectious disease, which was declared to be a public health emergency of international concern due to its various reported complications ranging from microcephaly in newborn to Guillain-Barré Syndrome (GBS). Because less attention has been paid to this virus over time, literature has been lacking regarding the structural and conformational features of its proteins particularly the NS3 helicase protein. This dissertation has addressed two major aspects of Zika NS3 helicase protein: (i) the binding interactions and (ii) structural dynamics and conformational changes of the protein.

Investigations were carried out on the various Zika NS3 helicase ligand binding landscapes using 10 ligands via molecular docking, of which the best 3 were subjected to molecular dynamic simulations and several post dynamics analyses. Ivermectin, HMC-HO1α and lapachol emerged as the best 3 ligands. The result of the analysis showed that the binding of Ivermectin to ssRNA site and Lapachol and HMC-HO1α to the ATPase site induces a more compact protein structure, thus stabilizing residue fluctuations. The pharmacophoric characteristics found in Lapachol, HMC-HO1α and Ivermectin may be utilized in the design of a potent hybrid drug that can show efficient inhibition of a multitude of diseases including the detrimental co-infection of ZIKV, Dengue and Chikungunya.

Also in this study, a detailed structural dynamic analysis was carried out on the structural flexibility of the NS3 helicase protein after NITD008 binding via molecular dynamics simulation and other posts dynamic analysis including the Principal Component Analysis (PCA) and the Dynamic Cross Correlation (DCC) analysis. Result revealed a prominent shift in the P-Loop found at the ATP site of the helicase. This loop and helical flexible regions give new insight into the dynamic structural features of ZIKV NS3 Helicase. The PCA and DCC analysis result revealed a significant structural flexibility of the NITD008-NS3 Helicase system compared to the rigid unbound form of the protein. Furthermore, the NITD008-NS3 Helicase complex stability was also ensured via a 130ns molecular dynamic simulation, this has proven NITD008 as an effective potential inhibitor of the NS3 helicase protein. This research is of immense importance to the discovery of a potent Zika inhibitor and in medicine as it has proven potential inhibitors with a good binding affinity towards Zika NS3 helicase. Also, this study hopes to fill in the gap of information that has been missing regarding molecular studies on Zika virus to some extent.

DECLARATION 1 – PLAGIARISM

I, Sofiat Yetunde Oguntade, declare that

1. The research reported in this thesis, except where otherwise indicated, is my original work.

2. This thesis has not been submitted for any degree or examination at any other university.

3. This thesis does not contain other persons' data, pictures, graphs or other information, unless specifically acknowledged as being sourced from other persons.

4. This thesis does not contain other persons' writing, unless specifically acknowledged as being sourced from other researchers. Where other written sources have been quoted, then:

a. Their words have been re-written, but the general information attributed to them has been referenced.

b. Where their exact words have been used, then their writing has been placed in italics and inside quotation marks, and referenced.

5. This thesis does not contain text, graphics or tables copied from the Internet, unless specifically acknowledged, and the source being detailed in the thesis and in the references section.

A detail contribution to publications that form part and/or include research presented in this thesis is stated (include publications submitted, accepted, in *press* and published).

Signed ----------

DECLARATION 2 - LIST OF PUBLICATIONS

1. Oguntade, S.Y., Ramharack, P., Soliman M.E. (2017) Characterizing the Ligand Binding Landscape of Zika NS3 Helicase- Promising Lead Compounds as Potential Inhibitors. *Future virology*. (**Accepted**).

   Contribution:

   Sofiat Yetunde Oguntade: Contributed to the project by performing 70% literature review, experimental work, and data analysis, interpretation of results, manuscript preparation and writing.

   Pritika Ramharack: Contributed to the project by performing 30% experimental work, data analysis, manuscript preparation and writing.

   Mahmoud E. Soliman: Supervisor

2. Delving into Zika Virus Structural Dynamics- A Closer look at NS3 Helicase Loop flexibility and its Role in Drug Discovery. *RCS advances* (**Accepted**).

   Contribution:

   Sofiat Yetunde Oguntade: Contributed to the project by performing 70% literature review, experimental work, and data analysis, interpretation of results, manuscript preparation and writing.

   Pritika Ramharack: Contributed to the project by performing 30% experimental work, data analysis, manuscript preparation and writing.

   Mahmoud E. Soliman: Supervisor

RESEARCH OUTPUT

PUBLICATIONS

1. Oguntade, S.Y., Ramharack, P., Soliman M.E. (2017) Characterizing the Ligand Binding Landscape of Zika NS3 Helicase- Promising Lead Compounds as Potential Inhibitors. *Future virology*. (**Accepted**).

2. Ramharack, P., Oguntade, S.Y., Soliman M.E. (2017) Delving into Zika Virus Structural Dynamics- A Closer look at NS3 Helicase Loop flexibility and its Role in Drug Discovery. *RCS Advances* (**Accepted**).

DEDICATION

This work is dedicated to Almighty Allah, the Merciful One, the Giver of life, the One who has spared my life till this moment, the One who has made this possible and gave me the strength and opportunity to accomplish this.

This is also dedicated to my families: the Oguntade and Ogunlola family for their love and support throughout my master's program and every other day of my life as well as for teaching me many life skills that has contributed to my success. I appreciate them for being such a supportive family.

And lastly, this work is dedicated to the amazing man behind my success, without him, there will be no thesis. His encouragement, perseverance and endless support has kept me going, Dr Muhammed Olatunbosun Ogunlola and to my baby, Zayd Ayomide Ogunlola for supporting me throughout my period of study, you two are the best.

ACKNOWLDEGEMENTS

I would also like to express my profound gratitude to the following people:

- My Supervisor, Prof Mahmoud Soliman for his constant support during my master's degree.
- Dr Ndagi Umar, for his constant support and assistance.
- Dr Ojugbele, for his support during my master's degree.
- Pritika Ramharack for her assistance.

- To the UKZN molecular modeling and drug design research group (2016 group). It's been wonderful working with you all.
- To friends and families who has contributed positively towards my progress.
- To CHPC for their resources and technical support.
- UKZN College of Health Sciences for financial support.

I say a big thank you to everyone for being there for me in one way or the other, you are all highly appreciated.

TABLE OF FIGURES

Figure 2.1: Structure of ZIKV genome **10**
Figure 2.2: Replication cycle of Zika Virus..... **11**
Figure 2.3: Structure of NS3 ZIKV Helicase (PBD 5JMT) **15**
Figure 3.1: Graphical illustration of multi-dimensional potential energy surface..... **29**
Figure 4.1: Structure of NS3 ZIKV Helicase (PBD 5JMT). **39**
Figure 4.2: 2D structure, docked complexes and validation of the docked complexes. **46**
Figure 4.3: C-alpha RMSD backbone Plot for NS3 Helicase free and ligand bound conformations..... **47**
Figure 4.4: RMSF Plot for Lapachol, HMC-HO1α and Ivermectin systems..... **49**
Figure 4.5: Radius of gyration plot for Lapachol, HMC-HO1α and Ivermectin systems when compared to the free protein..... **50**
Figure 4.6: Free energy decomposition and ligand-residue interaction network at the ATPase site of the Lapachol- NS3 Helicase system. **53**
Figure 4.7: HMC-HO1α docked into the ATPase site of Zika NS3 helicase, illustrating ligand-residue interactions and active-site residue energy contributions. **54**
Figure 4.8: Free energy decomposition and ligand-residue interaction network at the ssRNA site of the Ivermectin- NS3 Helicase system..... **55**
Figure 5.1: Cartoon and surface representation of the three domains of the ZIKV helicase **68**
Figure 5.2: Energy contributions of the highest interacting residues at the ATPase active site.. ..... **78**
Figure 5.3: Superimposed conformation of structurally similar NITD008 and ATP docked at ATPase site of ZIKV NS3 Helicase. **79**
Figure 5.4: C-α backbone RMSD for NS3 Helicase APO enzyme and NITD-complex conformation. **81**
Figure 5.5: The RMSF of APO enzyme and NITD008-complex.. ..... **82**
Figure 5.6: The radius of gyration (Rg) plot illustrating the difference in enzyme compactness of the NITD008-complex compared to the APO enzyme. **83**
Figure 5.7: Structural Flexibility of the P-Loop (196-203), RNA-binding loop (244-255), and the 310 Helix (339-348) along the trajectory. **86**
Figure 5.8: Residue fluctuations at the P-Loop region. **87**
Figure 5.9: Residue fluctuations at the "325-348" region. **90**

Figure 5.10: Projection of Eigen values of the Cα backbone, during 130ns simulation, for Apo and NITD008-bound conformations of NS3 Helicase along the first two principal components. **91**

Figure S.1: 2D structures of potential ligand **105**

Figure S1: Complex of NITD008-NS3 Helicase with a Docking score of -7.7 kcal/mol. MM/GBSA calculations yielded a result of -30.00 kcal/mol. **109**

Figure S2: Complex of NITD008-NS3 Helicase with a Docking score of -7.6 kcal/mol. MM/GBSA calculations yielded a result of -13.67 kcal/mol. **110**

Figure S3: Complex of NITD008-NS3 Helicase with a Docking score of -7.1 kcal/mol. MM/GBSA calculations yielded a result of -11.86 kcal/mol. **111**

Figure S4: Complex of NITD008- NS3 Helicase with a Docking score of -7.1 kcal/mol. MM/GBSA calculations yielded a result of -23.99 kcal/mol. **112**

Figure S5: MM/GBSA calculations yielded a result of -5.90 kcal/mol, which was lower than that of the docking score of 6.9 kcal/mol.. ..... **113**

TABLE OF TABLES

Table 4.1: Calculated parameters for running accelerated molecular dynamics. **43**

Table 5.1: Summary of free binding Energy contributions to the NITD008-NS3 Helicase system..... **77**

Table S.1: Grid box parameter for molecular docking. **106**

Table S.2: Binding energy result from docking selected potential Zika NS3 helicase inhibitors. **106**

LIST OF ABBREVIATIONS

aMD	Accelerated Molecular Dynamics
ATP	Adenosine triphosphate
C	Capsid
CHIKV	Chikungunya Virus
C- Terminal	Carboxyl terminal
β	Beta
DCCM	Dynamic Cross Correlation Matrix
DENV	Dengue Virus
DNA	Deoxyribonucleic Acid
dsRNA	Double Stranded Ribonucleic Acid
FDA	Food and Drug Agency (USA)
E	Envelope
ER	Endoplasmic Reticulum
GAFF	General AMBER Force Field
HCV	Hepatitis C Virus
JEV	Japanese Encephalitis Virus
MM	Molecular Mechanics
MD	Molecular Dynamics
MTase	Methyl Transferase
MM/GB-SA	Molecular Mechanics General Born-Surface Area
MM/PB-SA	Molecular Mechanics Poisson Boltzman-Surface Area
MMV	Molegro Molecular Viewer
NS	Non-Structural

N-Terminal	Amino (NH_4) Terminal
NTPase	Nucleoside triphosphatase
PCA	Principal Component Analysis
PES	Potential Energy Surface
prM	Precursor Membrane
QSAR	Quantitative Structure-Activity Relationships
RdRp	RNA-dependent RNA polymerase
Rg	Radius of Gyration
RIN	Residue Interaction Network
RMSD	Root Mean Square Deviation
RMSF	Root Mean Square Fluctuation (RMSF)
RNA	Ribonucleic Acid
RTPase	RNA Triphosphatase
SASA	Solvent Accessible Surface Area
SEA	Substrate Envelope Analysis
SBDD	Structural Based Drug Design
ssRNA	Single Stranded Ribonucleic Acid
WHO	World Health Organization
WNV	West Nile Virus
YFV	Yellow Fever Virus
ZIKV	Zika Virus

TABLE OF CONTENTS

IN SILICO STUDIES ON ZIKA NS3 HELICASE: BEDROCK FOR ANTIVIRAL DRUG DESIGN ..... **ii**

ABSTRACT ..... **iii**

DECLARATION 1 – PLAGIARISM ..... **iv**

DECLARATION 2 - LIST OF PUBLICATIONS ..... **v**

RESEARCH OUTPUT ..... **vi**

DEDICATION ..... **vii**

ACKNOWLDEGEMENTS ..... **viii**

TABLE OF FIGURES ..... **ix**

TABLE OF TABLES ..... **xi**

LIST OF ABBREVIATIONS ..... **xii**

TABLE OF CONTENTS ..... **xiv**

CHAPTER 1 ..... **1**

1.1 Background and rationale of the study 1

1.2 Aim and objectives 2

1.3 Novelty and significance of this study 3

1.4 Thesis structure: 3

References 5

CHAPTER 2 ..... **7**

Background on Zika virus and Zika NS3 helicase **7**

2.1. Introduction 7

2.2. Brief history and epidemiology of Zika Virus 7

2.3. Transmission 8

2.4. Clinical Manifestation 8

2.5. Treatment 9

2.6. Structure of Zika Virus and Genome 9

2.7 Life Cycle of Zika Virus in a Host Cell 10

2.8. Potential Drug Target on Zika Virus 11

2.8.1. Structural Proteins 12

2.8.2 Non-Structural Proteins 12

2.8.3 Host Target 14

2.9 Zika NS3 Helicase 14
2.10 Potential Inhibitors of Zika NS3 Helicase 15
2.10.1 Ivermectine 16
2.10.2 Purine nucleoside analogues 16
2.10.3 Naphthaquinones 17
2.10.4 Adenosine analogue NITD008 17
2.11. Conclusion 18
References 18
CHAPTER 3 27
Introduction to Computational Chemistry 27
3.1 Introduction 27
3.2 Quantum mechanics 27
3.3 Schrodinger's equation 27
3.4 Born-Oppenheimer approximation 28
3.5 Potential Energy surface (PES) 29
3.6 Molecular mechanics 29
3.6.1 Force field 30
3.7 Molecular dynamic simulation 30
3.8 Molecular docking 31
3.9 Binding free energy. 32
3.10 Principal Component Analysis (PCA) 32
References 33
CHAPTER 4 36
Characterizing the Ligand Binding Landscape of Zika NS3 Helicase-Promising Lead Compounds as Potential Inhibitors. 36
Abstract 36
4.1. Introduction 37
4.2 Computational methodology 41
4.2.1. Protein structure preparation 41
4.2.2. Molecular docking 41
4.2.3 Molecular dynamic simulations 42
4.2.4 Thermodynamic calculations 43
4.2.5. Per-residue energy decomposition analysis 44

4.3. Results and Discussion 45
4.3.1 Docking result and validation 45
4.3.2 Molecular dynamics simulation and post molecular dynamics analysis 46
4.3.2.1 Systems stability 46
4.3.2.2 Root Mean Square Fluctuations (RMSF) 48
4.3.2.3 Radius of Gyration 49
4.3.2.4 Free Energy Calculations and Residue-Ligand Interaction Network 51
4.4. Conclusions 56
4.5. Future Perspective 57
4.6 Summary point 57
References 59
CHAPTER 5 65
Delving into Zika Virus Structural Dynamics: A Closer look at NS3 Helicase Loop flexibility and its Role in Drug Discovery 65
Abstract 65
5.1 Introduction 66
5.2 Computational Methods 70
5.2.1 System Preparation 70
5.2.2 Molecular Docking 70
5.2.3 Molecular Dynamic (MD) Simulations 71
5.2.4 Post-Dynamic Analysis 72
5.2.4.1 Binding Free Energy Calculations 72
5.2.4.2 Dynamic Cross-correlation Analysis (DCC) 74
5.2.4.3 Principal Component Analysis (PCA) 74
5.3 Results and Discussion 75
5.3.1 NITD008-NS3 Helicase Complex 75
5.3.1.1 Binding of NIT21D008 with ZIKV Helicase 75
5.3.1.2 Free Energy calculations 76
5.3.2 Systems Stability 79
5.3.2.1 Stability of NS3 Helicase APO and Bound System 80

5.3.2.2 Conformational Fluctuations of the NS3 Helicase 81
5.3.2.3 Distribution of Atoms around the NS3 Helicase Backbone 82
5.3.3 Investigation of the Dynamic Structural features ATP-Active Binding Region 84
5.3.3.1 Loop Flexibility and Distance metrics 84
5.3.3.2 Principal Component Analysis 90
5.4. Conclusion 92
References 94
CHAPTER 6 100
Conclusion and future recommendations 100
6.1. General conclusions 100
6.2. Recommendation and Future Studies 101
References 103
APPENDICES 104
Appendix 1. Supplementary material for Chapter 4 104
Appendix 2. Supplementary material for Chapter 5 107

CHAPTER 1

1.1 Background and rationale of the study

This chapter outlines the background and novelty of this research project. It then presents the Aims and Objectives of the study and concluded by outlining the structure of the following five chapters, specifically indicating in which ones the Objectives were addressed.

The Zika Virus (ZIKV) gained international notoriety shortly before the 2016 Olympics, when Brazil was infested with the mosquitoe-borne Zika virus. Concerns were raised about the grave consequences of those visiting the epidemic area, specifically pregnant women, due to the resulting microcephaly of newly born children and the Gullian barre Syndrome. Following this situation, the World Health Organization declared Zika virus outbreak to be a public emergency of international concern on the 1st of February 2016[1]. Approximately1.5million cases of the virus were reported in Brazil during the 2015 outbreak[2]. This was preceded by major outbreaks in Micronesia in 2007 and French Polynesia in 2013[3]. The virus has spread rapidly to new geographic areas including South and Central America and the potential of this virus to spread across the world and affect the lives of millions of people highlight the need to find a cure to Zika virus.

ZIKV belongs to the genus *flavivirus* and family *flaviviridae*[4,5], with other viruses of public health concern being Dengue Virus (DENV), West Nile (WNV), Yellow Fever Virus (YFV), Japenese Encephalitis Virus (JEV) and Chikungunya Virus (CHIKV)[4,6]. Antiviral agents against other flaviviral disease, such as Yellow Fever Virus (YFV), have been discovered, indicating the potential to develop an inhibitor to the Zika virus [7]. ZIKV genome cleaves onto the three structural and seven non-structural proteins, of which NS3 helicase is one of the non-structural forms[8]. These proteins play a vital role in the lifecycle and replication of the virus, thus forming an important target in the design of an inhibitor. The virus shares similar replication schemes and genome organizations with other flaviviruses[9]. Its NS3 helicase protein is the second largest viral protein and plays an essential role in the viral life cycle. In addition, it has a N-terminal protease and a C-terminal domain, both of which have RNA helicase, nucleoside and RNA triphosphatase activities that are involved in viral replication and RNA synthesis[10]. These multiple activities and the roles it plays in viral replication makes it an attractive target for designing a ZIKV inhibitor. Although numerous studies have reported the need for ZIKV drug discovery, neither FDA approved drugs nor vaccines are currently available. There is therefore the need to design new drug inhibitors that can destroy the virus, with a deeper understanding of

the structural, binding and conformational features of the proteins being necessary for this to occur. As there is a lack of literature in this regard, designing effective small drug molecule inhibitors may be challenging.

Ligand-protein binding is an important aspect in the field of molecular modeling and drug design that gives an insight into the suitability of the binding of a potential inhibitor (ligand) to a receptor. Molecular docking, a lock and key technique that fit in a 3dimensional structure of a ligand into the active site of the receptor can achieve this. The binding is associated with some energy functions (free binding energy) that estimates the degree of the binding and ensures the stability of the complex[11]. In addition, the binding is also associated with some interactions between the ligand and protein residues via some intermolecular forces. Upon ligand-protein binding, structural changes are usually observed in the protein which changes the conformation of the complex[12]. The protein dynamics is an important aspect that need to be monitored because the higher the degree of structural change the lesser the docking accuracy[13]. This is a challenge that can be resolved via Molecular dynamic (MD) simulation.

MD simulation is a valuable tool for understanding complex, dynamic phenomena that occur in biological systems. This includes conformational changes, protein stability, molecular recognition of proteins and complexes, and protein folding, which are very useful in drug design and structure determination. The simulation is a computational method that calculates the time dependent behavior of a molecular system and provides detailed information on its conformational changes and fluctuations. It also examines the dynamics, structure, and thermodynamics of the biological systems and their complexes[14]. MD simulation is also important in drug design due to its predicting of protein-ligand docking, in which a drug is modeled to target a specific site of the protein. The docked protein-ligand is usually rigid with poor flexibility, while the MD simulation allows for a proper degree of flexibility of the protein-ligand complex, thereby giving a better and more stable conformation [15].

1.2 Aim and objectives

This study has 2 major Objectives:

1. To provide a molecular understanding on the various drug binding landscape of Zika NS3 helicase protein. To accomplish this, the following objectives were outlined:

1.1. To carry out molecular docking on Zika NS3 helicase in complex with ten ligands and subject the best three complexes to subsequent molecular dynamic simulation and some post MD analysis.

1.2. To analyze the binding interactions between the ligands and protein residues of each complex by employing the Molecular Mechanics/Generalized-Born Surface Area method (MM/GBSA).

2. To investigate the structural dynamics and conformational changes of the protein both in its apo and bounded forms. To achieve this, the following objectives were outlined:

   2.1. To apply various metrics to describe the motion of the ATP binding region of the protein.

   2.2. To analyze various stages of the simulation in order to accurately probe the dynamics of the loop region.

1.3 Novelty and significance of this study

This study is the first and only computational study that has been carried out on Zika NS3 helicase protein to investigate its various ligand binding landscape, structural dynamics and conformational changes upon ligand binding. However, no experimental study has reported this till date. Several studies had been carried out on other NS3 helicase flaviviruses, such as Dengue Virus, Yellow Fever Virus and their inhibitors have been discovered but none on Zika Virus. Several studies have reviewed the prevalence of the disease in different countries including its manifestation and complications, but none have reported on the molecular dynamic simulation of the non-structural protein 3 of the virus. This work is the first investigation into the detailed dynamics and structure of the enzyme in response to different ligands.

The crystal structure of Zika NS3 helicase was reported last year[8,16] but since, then no serious work has been done on the protein. This gap of information forms the basis of this study, wherein a comprehensive study of the protein dynamics, free binding energy and landscape interaction of the protein is presented. The work presented in this thesis is considered a key cornerstone towards further molecular understanding of Zika virus and can serve as a road map for drug design and development

1.4 Thesis structure:

This work is presented in the following five chapters:

Chapter 2: Literature Review. This chapter provides a general overview on Zika virus, starting with a brief historical background on the Zika virus epidemic and epidemiology. It highlights aspects such as the transmission, clinical manifestation, as well as it genome, lifecycle, and potential drug targets. The NS3 helicase protein, as a drug target and the focus of this work, is addressed in detail, including its structure, and potential inhibitors.

Chapter 3: Methodology. This chapter details a general introduction to computational chemistry and various molecular modeling and simulation techniques, as well as their common applications. Theoretical descriptions of the computational methods are included, followed by relevant computational tools used in ZIKV research, with a focus on molecular dynamics simulations, quantum mechanics, molecular mechanics and docking, binding free energy calculations and principal component analysis.

Chapter 4: Accepted Manuscript. This is a research paper from this study. The paper is entitled "Characterizing the Ligand Binding Landscape of Zika NS3 Helicase- Promising Lead Compounds as Potential Inhibitors" and accepted in the journal of Future. It addressed Objective 1.1 and 1.2.

Chapter 5: Accepted Manuscript. This is a research paper from this study. The paper is entitled "Delving into Zika Virus Structural Dynamics- A Closer look at NS3 Helicase Loop flexibility and its Role in Drug Discovery" and submitted to the journal of RCS Advances. It addressed Objective 2.1 and 2.2.

Chapter 6: Conclusion. This expounds the concluding remarks of the entire thesis and future work plans.

References

(1) Petersen, E., Wilson, M. E., Touch, S., McCloskey, B., Mwaba, P., Bates, M., Dar, O., Mattes, F., Kidd, M., Ippolito, G., Azhar, E. I., and Zumla, A. (2016) Rapid Spread of Zika Virus in The Americas - Implications for Public Health Preparedness for Mass Gatherings at the 2016 Brazil Olympic Games. *Int. J. Infect. Dis.44*, 11–15.

(2) Bogoch, I. I., Brady, O. J., Kraemer, M. U. G., German, M., Creatore, M. I., Kulkarni, M. A., Brownstein, J. S., Mekaru, S. R., Hay, S. I., Groot, E., Watts, A., and Khan, K. (2016) Anticipating the international spread of Zika virus from Brazil. *Lancet387*, 335–336.

(3) Maestre, A. M., Caplivski, D., and Fernandez-Sesma, A. (2016) Zika Virus: More Questions Than Answers. *EBioMedicine5*, 2–3.

(4) Haddow, A. D., Schuh, A. J., Yasuda, C. Y., Kasper, M. R., Heang, V., Huy, R., Guzman, H., Tesh, R. B., and Weaver, S. C. (2012) Genetic characterization of zika virus strains: Geographic expansion of the asian lineage. *PLoS Negl. Trop. Dis.6*, 1–7.

(5) Martínez de Salazar, P., Suy, A., Sánchez-Montalvá, A., Rodó, C., Salvador, F., and Molina, I. (2016) Zika fever. *Enferm. Infecc. Microbiol. Clin.34*, 247–252.

(6) Hamel, R., Liegeois, F., Wichit, S., Pompon, J., Diop, F., Talignani, L., Thomas, F., Despres, P., Yssel, H., and Misse, D. (2016) Zika virus: epidemiology, clinical features and host-virus interactions. *Microbes Infect.18*, 441–449.

(7) Mastrangelo, E., Pezzullo, M., De burghgraeve, T., Kaptein, S., Pastorino, B., Dallmeier, K., De lamballerie, X., Neyts, J., Hanson, A. M., Frick, D. N., Bolognesi, M., and Milani, M. (2012) Ivermectin is a potent inhibitor of flavivirus replication specifically targeting NS3 helicase activity: New prospects for an old drug. *J. Antimicrob. Chemother.67*, 1884–1894.

(8) Tian, H., Ji, X., Yang, X., Xie, W., Yang, K., Chen, C., Wu, C., Chi, H., Mu, Z., Wang, Z., and Yang, H. (2016) The crystal structure of Zika virus helicase: basis for antiviral drug design. *Protein Cell.7*, 450–454.

(9) Yong-Qiang, D., Na-Na, Z., Chun-Feng, L., Min, T., Jia-Nan, H., Xu-Ping, X., Pei-Yong, S., and Cheng-Feng, Q. (2016) Adenosine analog NITD008 is a potent inhibitor of Zika virus. *Open Forum Infectious Disease.* 1–4.

(10) Luo, D., Vasudevan, S. G., and Lescar, J. (2015) The flavivirus NS2B-NS3 protease-helicase as a target for antiviral drug development. *Antiviral Res.118*, 148–158.

(11) Oledzki, P. R., Laurie, A. T. R., and Jackson, R. M. (2006) Specialist Review Protein–ligand docking and structure-based drug design. *Spec. Rev.* 1–17.

(12) Najmanovich, R., Kuttner, J., Sobolev, V., and Edelman, M. (2000) Side-chain flexibility in proteins upon ligand binding. *Proteins Struct. Funct. Genet.39*, 261–268.

(13) Erickson, J. A., Jalaie, M., Robertson, D. H., Lewis, R. A., and Vieth, M. (2004) Lessons in Molecular Recognition: The Effects of Ligand and Protein Flexibility on Molecular Docking Accuracy. *J. Med. Chem.47,* 45–55.

(14) Martin, K., and McCammon, J. A. (2002) Molecular dynamic simulations of biomolecules. *Nature structural Biology*, 9(9) 646-652.

(15) Pierre, T., and Philippe, D. (2012) Flexibility and binding affinity in protein– ligand, protein–protein and multi-component protein interactions: limitations of current computational approaches. *J. R. Soc. Interface,* 9, 20–33.

(16) Jain, R., Coloma, J., Garcia-Sastre, A., and Aggarwal, A. K. (2016) Structure of the NS3 helicase from Zika virus. *Nat. Struct. Mol. Biol.23*, 752–754.

CHAPTER 2

Background on Zika virus and Zika NS3 helicase

2.1. Introduction

This chapter briefly describes the historical background, transmission, manifestation, treatment, life cycle and the structure of Zika virus genome. This chapter also highlights the potential drug targets and then explains in detail the NS3 helicase protein as the focus of this work including its potential inhibitors.

2.2. Brief history and epidemiology of Zika Virus

The Zika virus (ZIKV) is a positive-sense, single stranded RNA arbovirus belonging to the spondweni serocomplex[1,2]. It is an emerging virus that was first discovered in the Zika forest in Uganda near lake Victoria in 1947 and was therefore named after the forest[1,2].The virus was isolated in the blood of a sentinel Rhesus monkey during research on the yellow fever virus, and a second isolation from the mosquito *Aedes africanus* was done at the same forest in January 1948[3]. Since then, sporadic isolations have been reported from a variety of mosquito specie in Africa (Nigeria, Gabon, Sierra Leone, and Cote d' Ivoire)[2,4], Asia (Cambodia, India, Indonesia, Malaysia, Pakistan, Phillipine, Singapore, Thailand and Vietnam), and central America [5,6]. In 1954, another isolation was made from the serum of a 10year old Nigerian girl who presented with fever and headache[6].In 2007, there was an outbreak in Micronesia, which was initially suspected to be Dengue Virus (DENV). Several serologic tests were carried out that later confirmed it to be ZIKV. Over 70% of the Yap resident were affected by the outbreak, although no death or hospitalization was recorded[7].

The most recent and devastating outbreak of ZIKV occurred in Brazil at the end of 2015, with an estimated 500,000-1,500,000 cases being reported[8]. The virus may have been introduced into Brazil by Asian travelers during the 2014 Football World Cup or Participants of the Oceanic countries of Va'a World Sprints canoe Championship the same year[9,10]. Autochthonous cases of ZIKV infection has been reported in 26 countries in South America as at January 2016, including Mexico, Paraguay, Venezuela, Barbados and Ecuador. Furthermore, ZIKV epidemic was reported for the first time in French Polynesia in 2013, with 8,510 suspected cases being reported, since which time autochthonous cases have been reported there, including Japan, Norway and France[5].

2.3. Transmission

Zika virus is mainly transmitted by the female aedes mosquitoe, the first isolation of the virus being made from arboreal *Aedes africanus.* Isolations have been reported from *A. apicoargenteus, A. luteocephalus, A. aegypti, A.vitattus,* and *A.furcifer* mosquitoes, which all belongs to the sub-genus Stegomyia[1]. *A.aegypti* (confined to tropical and subtropical regions) and *A.albopictus* (found in temperate, tropical and subtropical areas) have been implicated in Zika outbreaks, with the latter having a wider range than the former[11]. *A.albopictus* was reported to play an epidemic role in Gabon in 2007[12]. ZIKV is said to be epizootic and enzootic in non-human primates, which serve as their natural reservoir host (sylvatic cycle)[13]. The mass outbreak of the virus in Brazil has suggested that ZIKV can be transmitted through an urban transmission cycle in which the virus cycles between its vector and non-wild or domestic animal before affecting humans [14,15]. Based on virology and serology tests been carried out in Africa and Asia, non-human primates, including *C. albigenajohnstoni, Chlorocebussabaeus, Colobusabyssinicus, Erythrocebuspatas,* and *Pongopygmaeus* and mammals such as rodents and zebras, are suggested as possible vertebrate hosts for ZIKV transmission[16–18].

Non-vector routes have also been reported and include perinatal transmission following the viral crossing of the placenta during delivery[19]. The virus has also been detected in amniotic fluid, placenta and fetal tissues[20]. Sexual transmission following the detection of the virus in the semen of an infected patient has also been reported [21,22], as well as through blood transfusion [19,23]. The viral genome has been detected in urine[24], saliva[25] and breast milk[19], although transmission through breastfeeding has not been reported.

2.4. Clinical Manifestation

The Zika virus infection usually present as asymptomatic, with only one in five cases presenting with symptoms[26]. After the incubation period of 3-12days, patients present with acute onset of fever, followed by maculo-papular rash on the palm and soles, mild headache, malaise, myalgia, arthralgia, conjunctivitis and sometimes, abdominal pain[5,27]. Some of these symptoms are shared with other related flavivirus viruses, including dengue virus, yellow fever virus, West Nile, St. Louis encephalitis virus and Japanese encephalitis virus[28,29].Several cases of complications of ZIKV has been reported that include 42 cases of Guillian-Barrè Syndrome (GBS), an autoimmune disorder in which the immune system attacks healthy nerve cells of the peripheral nervous system[10], and other neurological conditions that were reported during the epidemic in French

Polynesia in 2013[4,30], which was also noted in Latin America in 2015. Other reported cases include, meningoencephalitis in the pacific islands[31] and myelitis in Guadeloupe[32].

The virus is the leading cause of microcephaly by prenatal transmission in South America[33]. This is a neurological birth defect in which a baby's head is significantly smaller than the heads of other children of the same sex and age[19]. The 2015 outbreak resulting in a 20 fold increase in number of infants born with this condition, with over 1,200 cases being reported[34]. According to research conducted by the Ministry of Health in Brazil, there is an increased risk of microcephaly and malformation during the first trimester of pregnancy. The World Health Organization (WHO) therefore declared a public health emergency of international concern on the issue of the suspected link between Zika and microcephaly on February 1, 2016[35–37] which ended on November 18, 2017. This was due to the high influx of research that has demonstrated the link between the virus and its complications[38].

2.5. Treatment

There is no specific treatment for the Zika virus, and generally entails providing supportive care and treating the symptoms. This might include bed rest, oxygen, adequate rehydration through intravenous fluid due to loss of fluid through sweating, and monitoring of vital signs. Antipyretic and analgesic, such as acetaminophen or dipyronecan, can be used to reduce the fever and arthralgia[3], while antihistamines can be used to treat the pruritus. The use of aspirin and non-steroidal anti-inflammatory drugs should be avoided until it has been confirmed that the infection is not Dengue or Chikungunya virus to prevent haemorrhagic complications[6,39].

2.6. Structure of Zika Virus and Genome

ZIKV is an envelope, icosahedra virus with non-segmented, positive sense, single stranded RNA genome of 10.7kb flanked by two untranslated 5' and 3' regions that contain a large polyprotein[40,41] (Figure 2.1A). This polyprotein contains more than 3000 amino acids, which cleaves into three structural proteins (envelope, E; membrane precursor, PrM; and capsid C) and seven Non-Structural proteins (NS1, NS2A, NS2B, NS3, NS4A, NS4B, and NS5), of which NS3 helicase protein is an important component for viral replication and RNA synthesis[13] (Figure 2.1B. The three structural proteins are responsible for the efficient regulation of virion assembly[42]. The interaction between the precursor membrane (prM) and the envelope protein (E) is important to prevent premature virus fusion, while the C protein associates with the RNA to conform the core of the virions[41].

The non-structural proteins assemble on the endoplasmic reticulum (ER) membranes to form a replication complex, which are membrane bound and morphologically distinct in their function

and composition[43]. According to the studies carried out by Saiz earlier this year, no studies has been carried out regarding the functions of the non-structural protein of ZIKV, although their functions can be inferred from other flaviviruses[41]. NS1, NS2A NS4A and NS4B are involved in RNA replication, NS2B complexes with NS3 to function as serine protease, NS2A is involved in the viral assembly, and NS3 possess RNA helicase and triphosphate activities. NS5 has methyltransferase and RNA guanyl-transferase activity, as well as capping and RNA-dependent RNA polymerase synthesis activity[13].

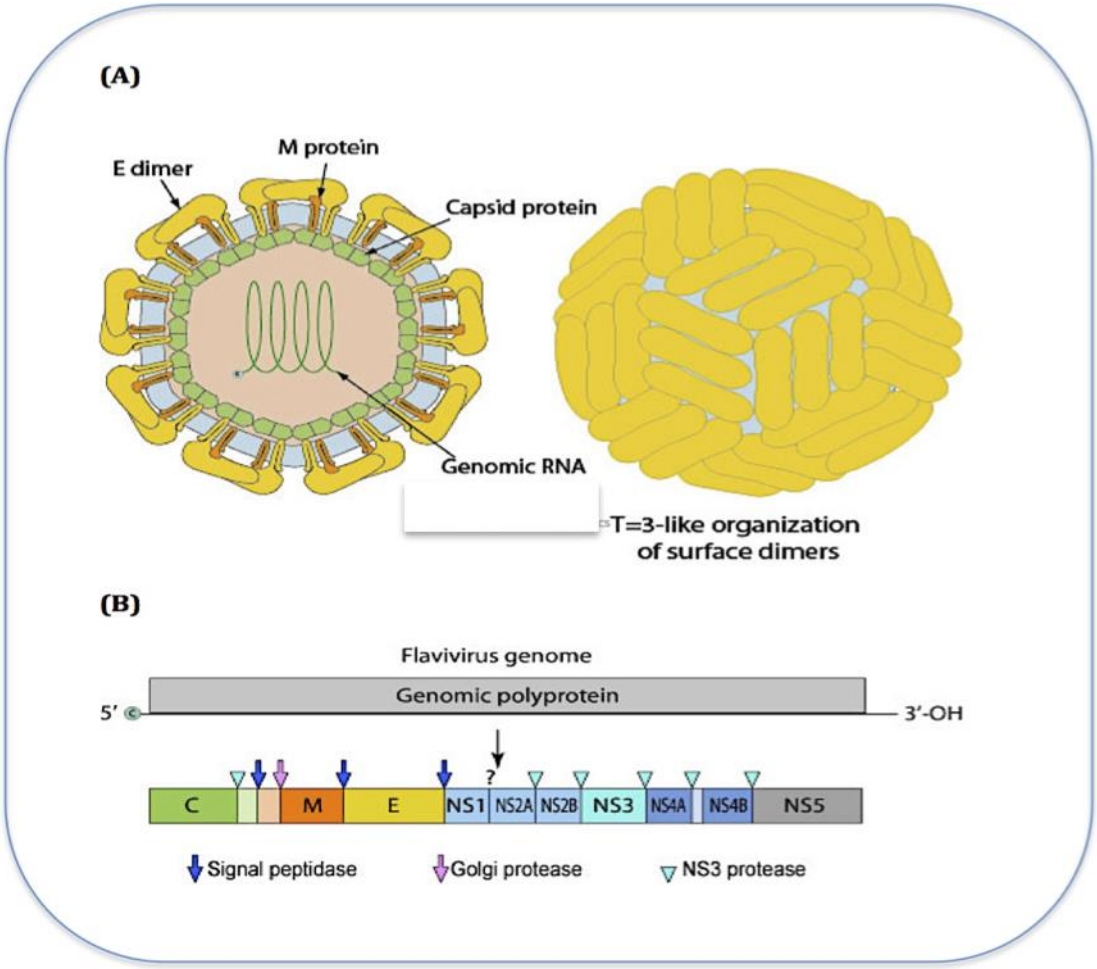


Figure 2.1 (A) Structure of ZIKV genome (Strain Mr766) (B) Schematic view of the organization of ZIKV genome[44].

2.7 Life Cycle of Zika Virus in a Host Cell

The life cycle of the Zika virus is similar to other known flaviviruses, being transmitted by a vector, the aedes mosquito. When a mosquito bites and sucks the blood of an infected human, it becomes infected, after which the virus makes its way into the mosquito's saliva [45]. Once the

infected mosquito bites another human, the virus is deposited from the saliva to the blood of the new host, who becomes infected. The replication cycle is shown in figure 2.2

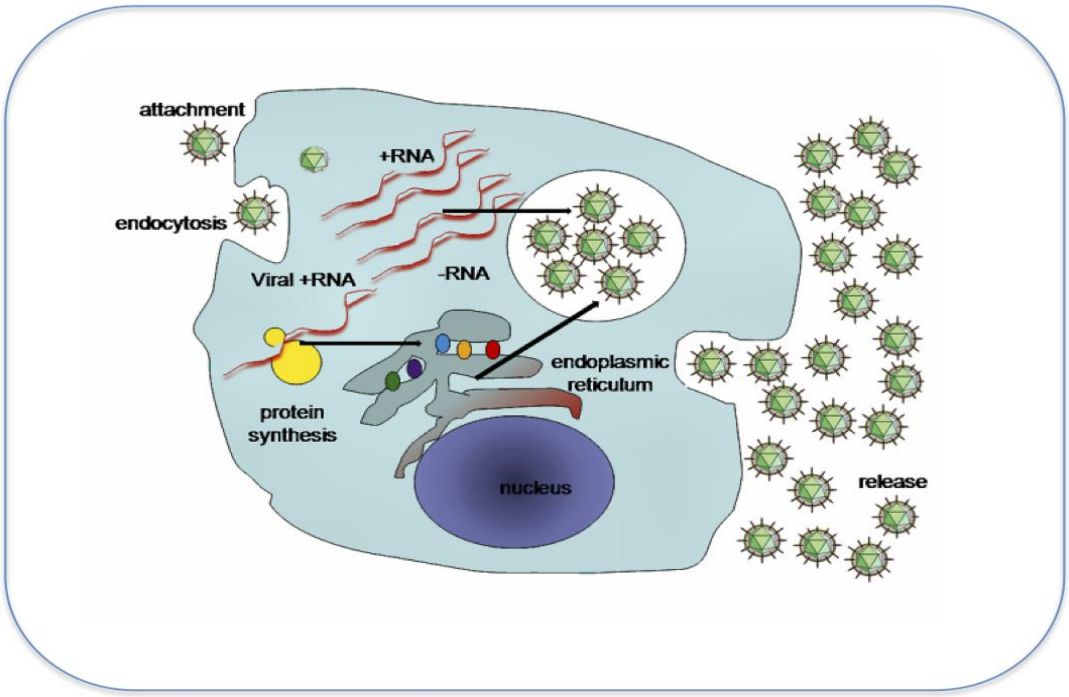


Figure 2.2 Replication cycle of Zika Virus[46]

When the virus is deposited into the human host, it finds its way into a cell to replicate. The virus becomes attached to the receptor of the host membrane using its envelope (E) protein. The binding of the E protein to the receptor induces the virion endocytosis, after which the viral membrane fuses with the host's endosomal membrane. Thereafter, the viral ssRNA is released into the host cytoplasm and is then translated into the polyprotein, which cleaves into the structural and non-structural proteins.[47]. Viral replication takes place at an intracellular viral factory in the endoplasmic reticulum and results in the formation of dsRNA, which is transcribed to additional ssRNA genome. Following the virion assembly within the endoplasmic reticulum, new virions are transported into the golgi apparatus and excreted to the intracellular space via exocytosis, where the new virion can infect other cells, thereby continuing the infection cycle[48](Figure 2.2).

2.8. Potential Drug Target on Zika Virus

Studies has shown that the structural and non-structural proteins of flaviviruses contain possible drug targets for antiviral drugs[49–51]. The structural protein contains the envelope, precursor membrane and the capsid protein that plays a vital role in the maturation of the virus and the non-

structural proteins.

2.8.1. Structural Proteins

The three structural proteins of flaviviruses are the Capsid (C), Precursor membrane (prM) and Envelope (E) membranes, each of which will be described further.

a. Capsid protein

The Capsid (C) protein is a small highly systolic basic protein approximately 11kDa that interact with the genomic RNA to form the nucleocapsid[52]. It localizes itself to three cellular compartments: the endoplasmic reticulum, cytoplasm and the nucleus. It plays a vital role in the assembly of the nucleo-capsid and its incorporation into nascent virions[53,54]. The region of the C protein that is required for dimerization has been characterized and is induced by RNA. However, the development of an *in vivo* assembly system could lead to identifying compounds that blocks the capsid dimerization or capsid RNA-interaction and subsequently inhibit either of this stage[52].

b. Precursor membrane protein

The precursor membrane (prM) is approximately19-21kDa, and is a glycoprotein that is translocated into the Endoplasmic Reticulum during polyprotein translation[42]. The prM contains 2 C terminals, the first anchoring the protein to the ER and the second acts as a signal for luminal translocation of envelope protein[55].It forms a vital part of the flaviviral envelope and assist in folding the E-protein. It also contribute to concealing the fusion loop of E to prevent premature fusion during virion release[54,56].

c. Envelope protein

The Envelope (E) protein is a large glycoprotein membrane (approximately 53kDa) that completely covers the surface of the mature virion. It is responsible for cell surface attachment/receptor binding, membrane fusion, and viral entry[57]. The E protein consist of three domains; central (called the β-barrel), second (involved in viral fusion during entry) and third (immunoglobulin-like domain that is the receptor binding region). Targeting this protein can disrupt the viral-host interaction and the entry of the virus into the host, thereby making it an ideal target for antiviral[58].

2.8.2 Non-Structural Proteins

A few selected non-structural proteins that are relevant to this study and will be discussed in this

section include NS1, NS2B, NS3 and NS5. Each of them will be briefly discussed.

a. NS1 protein

The NS1 is a glycoprotein (approximately 46kDa) found within and at the surface of the cell, which is then translocated into the ER and released from the envelope protein[57]. It plays an important role in viral replication[59], it has two N-glycosylation sites, the mutation of one or both of which can lead to a dramatic defect in RNA replication and subsequent viral production[60], makes it a good target for designing antiviral compound.

b. NS2B Protein

NS2B is a small membrane protein (approximately 14kDa), forms a complex with the NS3 protein and serves as a cofactor for NS3 serine protease activity. It has three conserved hydrophobic region flanking a hydrophilic domain of approximately 40 residues, this hydrophilic region is required to activate the NS3 protease domain[61]. Mutation of the NS2B-NS3 interaction can disrupt the serine protease activity, making it an ideal opportunity around which to design an inhibitor[62].

c. NS3 Protein

NS3 is a multifunctional protein (approximately70kDa) and contains several enzymatic activities that are involved in viral replication. It contains protease and helicase domains, the former complexes with the NS2B domain to function fully as a serine protease[63]. This serine protease domain is located at the N-terminal region of the protein, while the helicase domain is located at the C terminal, in addition to the RNA triphosphate (RTPase) activity[57]. The helicase domain participates in the RNA replication by displacing viral proteins bound to the dsRNA and disrupting the secondary structures that are formed by the ssRNA template[63]. The NS3 helicase has three subdomains, with the third subdomain forming the ssRNA binding tunnel. Studies has shown that this subdomain has been implicated in mediating the interaction between NS3 and NS5, disrupting this interaction being a possible strategy for designing antiviral drugs[64,65].

d. NS5 Protein

NS5 is the largest (approximately 103kDa) and the most conserved of the flaviviral protein[66]. Its N-terminal region comprises of the MTase S-adenosyl-methionine dependent methyltransferase (MTase) domain, while its C-terminal region comprises of the RNA-dependent RNA polymerase (RdRp) domain[43,67]. The RdRp plays an important role the in the lifecycle of the virus by conducting the de novo pathway and generating a viral template during RNA synthesis[68].The

MTase domain has two methylation caps, N7 and 2-O' methylation, which share a common core structured known as "SAM-dependent MTase fold that consist of α/β/α sheet. This protein is involved in virus-host interaction and interacts with its host environment[66].

2.8.3 Host Target

In attempting to discover inhibitors for any virus, it is important to consider the cell component of the host, which is essential for the lifecycle of the virus and itspathogenesis[52,69],these components being considered as potential target for developing antivirals. An important aspect in the lifecycle of the Zika virus is the entry of the virus into the host, the interruption of which could be a strategy to the discovery for an effective inhibitor against this virus[52,58]. Endoplasmic reticulum glycosidase are important host enzymes that catalyzes the removal of the three terminal glucose from the glycan at the glycosylation site, this being essential for the proper functioning of certain cellular and viral proteins[70]. Studies has shown that inhibitors of this enzymes inhibited the infection of many envelope viruses[71,72].Furthermore, some cellular proteins have been reported to bind to viral RNA, disruptions such as mutating the binding interactions could be key to developing an effective inhibitor[73].

2.9 Zika NS3 Helicase

Researchers are currently focusing on the structural and non-structural viral proteins to develop drugs[74], as clinical trials have shown that they are promising targets and play an important role in viral replication[49].In an effort to design an inhibitor against the Zika Virus, it is important to target the enzymatic activities that will affect the lifecycle of the virus[75]. One such enzyme is the NS3 helicase (Figure 2.3), which plays a significant role in viral replication and RNA synthesis.

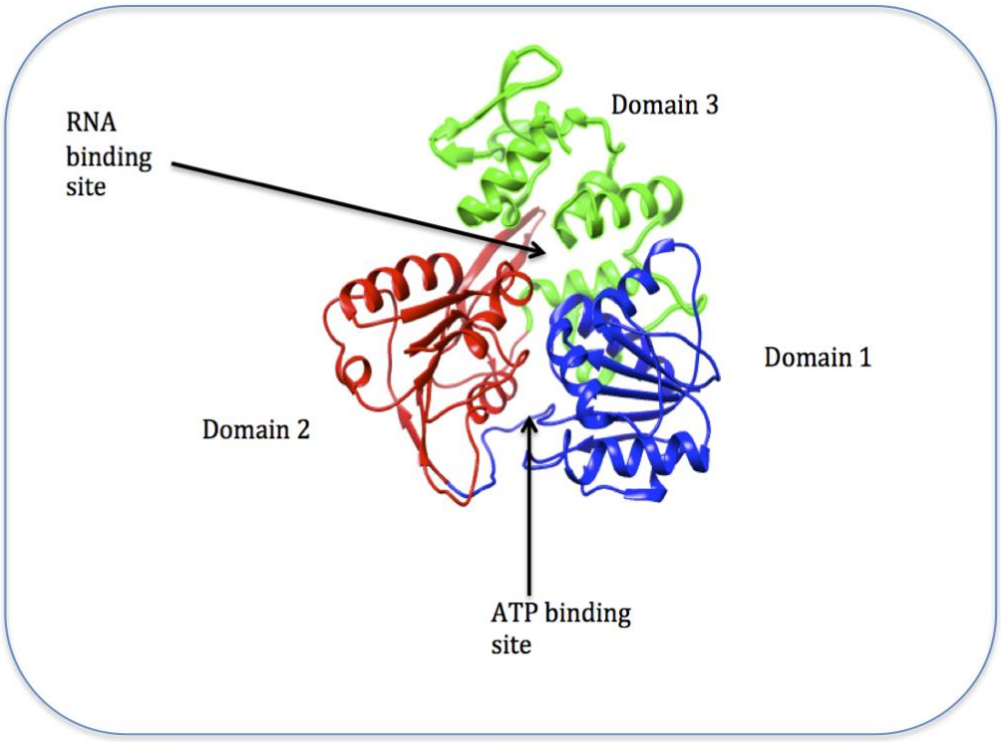


Figure 2.3 Structure of NS3 ZIKV Helicase (PBD 5JMT) [40]. Domain 1 (blue: residue 175-332) and domain 2 (red: residue 333-481) are seen facing each other, with domain 3 (green: residue 482-617) lying above the other 2 domains. The ATP binding site is located in the cleft between domains 1 and 2, and the ssRNA binding site is located at the tunnel that separates domain 3 from the other 2[75].

The structure of the Zika NS3 helicase is similar to other flavivirus, such as DENV, Yellow Fever Virus, Japanese Encephalitis Virus and Hepatitis C Virus (HCV), but the P-loop in each of these flaviviruses differs in their conformation[40]. The structure is characterized by three domain and two binding sites, with a protease domain at the N-terminus and a helicase domain at the C- terminus[43]. The two binding sites are the adenosine triphosphate (ATP) and single stranded ribonucleic acid (ssRNA) site. The first two domains are structurally similar, and comprises of a six-stranded β-sheet surrounded by a number of loop and helices, while domain 3 consist of four α helices bundle and two antiparallel β strand[43]. Domain 3 has been reported to interact with RNA dependent RNA polymerase NS5 in other flaviviruses, this helicase structure having two unique features: The P-loop, which is essential for NTpase binding, and a positively charged tunnel that interacts with domain 1 and 2.

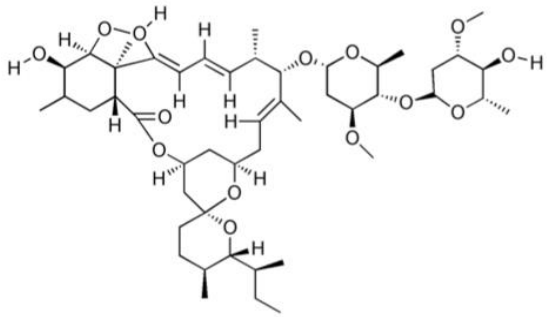
2.10 Potential Inhibitors of Zika NS3 Helicase

Although studies have been released regarding the discovery of Zika viruses[6,26,76], no FDA approved drugs are presently available. To explore Zika NS3 helicase protein, this research

utilized some clinically approved flavivirus NS3 small molecule inhibitors that have been proven to be effective for other flaviviral diseases.

2.10.1 Ivermectin

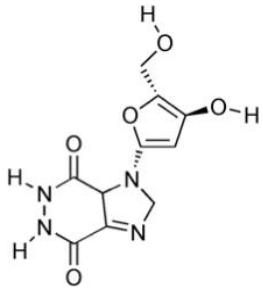
Ivermectin is a broadly used antihelmetic that has been proved to be a highly potent inhibitor for several flaviviruses, such as Yellow fever and dengue. Studies have shown that this drug has a good binding affinity towards the NS3 ssRNA binding pocket[77], and ivermectine was therefore used to analyse the RNA binding site of Zika NS3 helicase for this study, and to test it as a potential inhibitor of the virus.



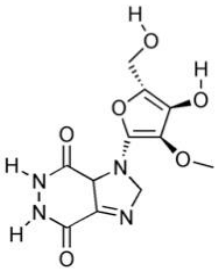
IVERMECTINE

2.10.2 Purine nucleoside analogues

In an experiment carried out by Borowski et al (2002), they reported how some purine nucleoside analogues inhibited the unwinding reaction of DNA/RNA of the West Nile Virus. They include; (1-(2'- deoxy- α -D-ribofuranosyl) imidazo [4,5-*d*] pyridazine-4,7(5*H*,6*H*) dione) (HMC-HO1α), 1-(2'-*O*-methyl- α -D-ribofuranosyl) imidazo [4,5-*d*] pyridazine- 4,7(5*H*,6*H*)-dione (HMC-HO4) and 1-(β-D-ribofuranosyl) imidazo [4,5-*d*] pyridazine-4,7(5*H*,6*H*)-dione (HMC-HO5)). They demonstrated that these nucleoside analogues possess the ability to interfere with the ATPase and helicase activity of West Nile Virus and Hepatitis C Virus (HCV) [78], which results in these compounds being explored on the ATP site of Zika NS3 helicase protein.



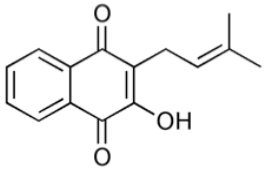
HMC-HO1α



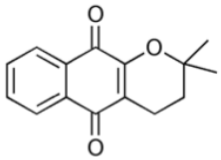
HMC-HO4

2.10.3 Naphthaquinones

Compounds with quinone as their core system are said to have positive biological activities including antiviral[79]. Naphthaquinones are a promising group of compounds with diverse biological activities some of which are used to treat various diseases. Lapachol is a member of this group, it exhibits a number of notable biological activities, which include antimalarial, antibacterial, trypanocidal, even antiviral. It has been used as a template to develop synthetic naphthaquinones, such as α lapachone and β lapachone which have shown considerable biological activities including antibacterial, fungal, anticancer, and antiviral activities[80,81]. It was suggested that these analogues acts at the level of the topoisomerase I and II enzyme which are vital for chromosome structure, DNA transcription and replication[82,83]. In this study, the various naphthaquinone compounds were docked into Zika NS3 helicase protein and the best was subjected to subsequent analysis.



LAPACHOL



ALPHA LAPACHONE

2.10.4 Adenosine analogue NITD008

Yin et al (2009) reported an adenosine analogue NITD008 that potentially inhibits dengue virus both *invivo* and *invitro*, and inhibits other flaviviruses, such as West Nile Virus and Yellow Fever Virus[84]. This experimental study proves that a nucleoside inhibitor could be developed to treat flaviviruses infections. In addition, Lo et al (2016) conducted research on the same adenosine analogue and reported how NITD008 was able to inhibit tick-borne flavivirus in an invitro experiment[85]. Finally, Yong-Qiang Deng et al (2015) demonstrated how NITD008 is a potent and

effective antiviral compound that can be used to design a ZIKV inhibitor[86]. This study therefore investigates the structure and conformation of the Zika NS3 helicase protein in an apo and bounded form using NITD008.

2.11. Conclusion

As an enveloped flavivirus, with a structural and non-structural component, interrupting the Zika virus transmission in humans will require drugs that can inhibit viral replication and RNA synthesis. This may be possible by exploring the NS3 protein, which is essential in the replication and RNA synthesis of the virus. This however, will be critical in understanding how the ZIKV NS3 helicase functions structurally, thus aiding in the design of effective inhibitors.

References

(1) Haddow, A. D., Schuh, A. J., Yasuda, C. Y., Kasper, M. R., Heang, V., Huy, R., Guzman, H., Tesh, R. B., and Weaver, S. C. (2012) Genetic characterization of zika virus strains: Geographic expansion of the asian lineage. *PLoS Negl. Trop. Dis.6*, 1–7.

(2) Martínez de Salazar, P., Suy, A., Sánchez-Montalvá, A., Rodó, C., Salvador, F., and Molina, I. (2016) Zika fever. *Enferm. Infecc. Microbiol. Clin.34*, 247–252.

(3) Zanluca, C., and Nunes, C. (2016) Zika virus - an overview. *Microbes Infect.18*, 295–301.

(4) Wikan, N., and Smith, D. R. (2016) Zika virus: History of a newly emerging arbovirus. *Lancet Infect. Dis.16*, e119–e126.

(5) Ioos, S., Mallet, H. P., Leparc Goffart, I., Gauthier, V., Cardoso, T., and Herida, M. (2014) Current Zika virus epidemiology and recent epidemics. *Med. Mal. Infect.44*, 302–307.

(6) Chan, J. F. W., Choi, G. K. Y., Yip, C. C. Y., Cheng, V. C. C., and Yuen, K. Y. (2016) Zika fever and congenital Zika syndrome: An unexpected emerging arboviral disease. *J. Infect.72*, 507–524.

(7) Lanciotti, R. S., Kosoy, O. L., Laven, J. J., Velez, J. O., Lambert, A. J., Johnson, A. J., Stanfield, S. M., and Duffy, M. R. (2008) Genetic and serologic properties of Zika virus associated with an epidemic, Yap State, Micronesia, 2007. *Emerg. Infect. Dis.14*, 1232–1239.

(8) Bogoch, I. I., Brady, O. J., Kraemer, M. U. G., German, M., Creatore, M. I., Kulkarni, M. A., Brownstein, J. S., Mekaru, S. R., Hay, S. I., Groot, E., Watts, A., and Khan, K. (2016) Anticipating the international spread of Zika virus from Brazil. *Lancet387*, 335–336.

(9) Pan American Health Organization. (2016) Epidemiological update. Neurological syndrome, congenital anomalies, and Zika virus infection. 17 January 2016. *World Heal. Organ.* 1–8.

(10) Musso, D., Nilles, E. J., and Cao-Lormeau, V. M. (2014) Rapid spread of emerging Zika virus in the Pacific area. *Clin. Microbiol. Infect.20*, O595–O596.

(11) Sampathkumar, P., and Sanchez, J. L. (2016) Zika Virus in the Americas: A Review for Clinicians. *Mayo Clin. Proc.91*, 514–521.

(12) Grard, G., Caron, M., Mombo, I. M., Nkoghe, D., Mboui Ondo, S., Jiolle, D., Fontenille, D., Paupy, C., and Leroy, E. M. (2014) Zika Virus in Gabon (Central Africa) – 2007: A New Threat from Aedes albopictus? *PLoS Negl. Trop. Dis.8*, e2681.

(13) Wong, S. S. Y., Poon, R. W. S., and Wong, S. C. Y. (2016) Zika virus infection-the next wave after dengue? *J. Formos. Med. Assoc.115*, 226–242.

(14) Diallo, D., Sall, A. A., Diagne, C. T., Faye, O., Faye, O., Ba, Y., Hanley, K. A., Buenemann, M., Weaver, S. C., and Diallo, M. (2014) Zika Virus Emergence in Mosquitoes in Southeastern Senegal, 2011. *PLoS One9*, e109442.

(15) Chan, J. F. W., To, K. K. W., Chen, H., and Yuen, K. Y. (2015) Cross-species transmission and emergence of novel viruses from birds. *Curr. Opin. Virol.10*, 63–69.

(16) Kilbourn, A. M., Karesh, W. B., Wolfe, N. D., Bosi, E. J., Cook, R. A., and Andau, M. (2003) Health Evaluation of Free-Ranging and Semi-Captive Orangutans (Pongo Pygmaeus Pygmaeus) in Sabah, Malaysia. *J. Wildl. Dis.39*, 73–83.

(17) Dick, G. W. A. (1952) Zika virus (II). Pathogenicity and physical properties. *Trans. R. Soc. Trop. Med. Hyg. 46* , 521–534.

(18) McCrae, A. W. R., and Kirya, B. G. (1982) Yellow fever and Zika virus epizootics and enzootics in Uganda. *Trans. R. Soc. Trop. Med. Hyg. 76* , 552–562.

(19) Besnard, M., Lastère, S., Teissier, A., Cao-Lormeau, V. M., and Musso, D. (2014) Evidence of perinatal transmission of zika virus, French Polynesia, December 2013 and February 2014. *Eurosurveillance19*, 8–11.

(20) Michael A. Johansson, Ph.D., Luis Mier‐y‐Teran‐Romero, Ph.D., Jennita Reefhuis, Ph.D., Suzanne M. Gilboa, P. D., and Susan L. Hills, M.B., B. S., Mlakar, J., Korva, M., Tul, N., Popović, M., Poljšak-Prijatelj, M., Mraz, J., Kolenc, M., Resman Rus, K., Vesnaver Vipotnik, T., Fabjan Vodušek, V., Vizjak, A., Pižem, J., Petrovec, M., Avšič Županc, T., Rasmussen, S. A., Jamieson, D. J., Honein, M. A., and Petersen, L. R. (2016) Zika Virus Associated with Microcephaly. *N. Engl. J. Med.374*, 1–3.

(21) Foy, B. D., Kobylinski, K. C., Foy, J. L. C., Blitvich, B. J., da Rosa, A. T., Haddow, A. D.,

Lanciotti, R. S., and Tesh, R. B. (2011) Probable Non-Vector-borne Transmission of Zika Virus, Colorado, USA. *Emerg. Infect. Dis.17*, 880–882.

(22) Musso, D., Roche, C., Robin, E., Nhan, T., Teissier, A., and Cao-Lormeau, V. M. (2015) Potential sexual transmission of zika virus. *Emerg. Infect. Dis.21*, 359–361.

(23) Chibueze, E. C., Tirado, V., and Olukunmi, O. (2016) Zika virus infection in pregnancy : a systematic review of disease course and complications. *Bull World Heal. Organ* Epub ahead of publication.

(24) Gourinat, A. C., O'Connor, O., Calvez, E., Goarant, C., and Dupont-Rouzeyrol, M. (2015) Detection of zika virus in urine. *Emerg. Infect. Dis.21*, 84–86.

(25) Musso, D., Roche, C., Nhan, T.-X., Robin, E., Teissier, A., and Cao-Lormeau, V.-M. (2015) Detection of Zika virus in saliva. *J. Clin. Virol.68*, 53–5.

(26) Petersen, E., Wilson, M. E., Touch, S., McCloskey, B., Mwaba, P., Bates, M., Dar, O., Mattes, F., Kidd, M., Ippolito, G., Azhar, E. I., and Zumla, A. (2016) Rapid Spread of Zika Virus in The Americas - Implications for Public Health Preparedness for Mass Gatherings at the 2016 Brazil Olympic Games. *Int. J. Infect. Dis.44*, 11–15.

(27) Hayes, E. B. (2009) Zika virus outside Africa. *Emerg. Infect. Dis.15*, 1347–1350.

(28) Mahfuz, M., Khan, A., Mahmud, H. Al, Hasan, M., Parvin, A., Rahman, N., and Rahman, S. M. B. (2014) Indian Journal of Pharmaceutical and Biological Research (IJPBR) In Silico Modeling and Immunoinformatics Probing Disclose the Epitope Based PeptideVaccine Against Zika Virus Envelope Glycoprotein. *Indian J. Pharm. Biol. Res.2*, 44–57.

(29) Plourde, A. R., and Bloch, E. M. (2016) A Literature Review of Zika Virus. *Emerg. Infect. Dis.22*, 1–15.

(30) Cao-Lormeau, V.-M., Blake, A., Mons, S., Lastère, S., Roche, C., Vanhomwegen, J., Dub, T., Baudouin, L., Teissier, A., Larre, P., Vial, A.-L., Decam, C., Choumet, V., Halstead, S. K., Willison, H. J., Musset, L., Manuguerra, J.-C., Despres, P., Fournier, E., Mallet, H.-P., Musso, D., Fontanet, A., Neil, J., and Ghawché, F. (2016) Guillain-Barré Syndrome outbreak associated with Zika virus infection in French Polynesia: a case-control study. *Lancet387*, 1531–1539.

(31) Sharman, J. P., Mato, A. R., and Keating, M. (2016) Zika Virus Associated with Meningoencephalitis. *N. Engl. J. Med.374*, 1592–1595.

(32) Mécharles, S., Herrmann, C., Poullain, P., Tran, T. H., Deschamps, N., Mathon, G., Landais, A., Breurec, S., and Lannuzel, A. (2016) Acute myelitis due to Zika virus infection. *Lancet387*, 1481.

(33) Charrel RN, Leparc-Goffart I, Pas S, de Lamballerie X, K. M. & R. C. (2016) State of knowledge on Zika virus for an adequate laboratory response . *Publ. Bull. World Heal. Organ. Type Res. emergencies Artic.171207* 1–29.

(34) Ventura, C. V., Maia, M., Bravo-Filho, V., Góis, A. L., and Belfort, R. (2016) Zika virus in Brazil and macular atrophy in a child with microcephaly. *Lancet387*, 228.

(35) Cauchemez, S., Besnard, M., Bompard, P., Dub, T., Guillemette-Artur, P., Eyrolle-Guignot, D., Salje, H., Van Kerkhove, M. D., Abadie, V., Garel, C., Fontanet, A., and Mallet, H. P. (2016) Association between Zika virus and microcephaly in French Polynesia, 2013-15: A retrospective study. *Lancet387*, 2125–2132.

(36) Zika, U. N., and Fund, M. T. (2016) Zika Virus and complications Update and funding request 2015–2016.

(37) Heymann, D. L., Hodgson, A., Sall, A. A., Freedman, D. O., Staples, J. E., Althabe, F., Baruah, K., Mahmud, G., Kandun, N., Vasconcelos, P. F. C., Bino, S., and Menon, K. U. (2016) Zika virus and microcephaly: Why is this situation a PHEIC? *Lancet387*, 719–721.

(38) WHO. (2016) WHO | Fifth meeting of the Emergency Committee under the International Health Regulations (2005) regarding microcephaly, other neurological disorders and Zika virus. *World Heal. Organ.*

(39) Shuaib, W., Stanazai, H., Abazid, A. G., and Mattar, A. A. (2016) Re-Emergence of Zika Virus: A Review on Pathogenesis, Clinical Manifestations, Diagnosis, Treatment, and Prevention. *Am. J. Med.129*, 879.e7–879.e12.

(40) Tian, H., Ji, X., Yang, X., Xie, W., Yang, K., Chen, C., Wu, C., Chi, H., Mu, Z., Wang, Z., and Yang, H. (2016) The crystal structure of Zika virus helicase: basis for antiviral drug design. *Protein Cell7*, 450–454.

(41) Saiz, J. C., Vázquez-Calvo, Á., Blázquez, A. B., Merino-Ramos, T., Escribano-Romero, E., and Martín-Acebes, M. A. (2016) Zika virus: The latest newcomer. *Front. Microbiol.7*, 1–19.

(42) Roby, J. A., Setoh, Y. X., Hall, R. A., and Khromykh, A. A. (2015) Post-translational regulation and modifications of flavivirus structural proteins. *J. Gen. Virol.96*, 1551–1569.

(43) Bollati, M., Alvarez, K., Assenberg, R., Baronti, C., Canard, B., Cook, S., Coutard, B., Decroly, E., de Lamballerie, X., Gould, E. A., Grard, G., Grimes, J. M., Hilgenfeld, R., Jansson, A. M., Malet, H., Mancini, E. J., Mastrangelo, E., Mattevi, A., Milani, M., Moureau, G., Neyts, J., Owens, R. J., Ren, J., Selisko, B., Speroni, S., Steuber, H., Stuart, D. I., Unge, T., and Bolognesi, M. (2010)

Structure and functionality in flavivirus NS-proteins: Perspectives for drug design. *Antiviral Res.87*, 125–148.

(44) Aryal Sagar. (2015) Zika Virus- Structure, Genome, Symptoms, Transmission, Pathogenesis, Diagnosis - Online Microbiology Notes. *Microbiol. info*.

(45) Disease, Z. V. Zika Virus Disease and Transmission. Available from: http://necsi.edu/research/social/pandemics/transmission.pdf. (Accessed 10 Dec 2016.

(46) Ask Scientific. (2016) Zika Virus Life Cycle and pathogenicity in humans. Avaiable from https://www.askscientific.com/zika-virus-life-cycle-pathogenicity-humans/ (Accessed 23 Feb 2017).

(47) ViralZone. (2015) Zika virus (strain Mr 766). *Expasy Bioinformation Resour. Portal*. Available from http://viralzone.expasy.org/all_by_species/6756.html (Accessed10 Jan 2017).

(48) Sreedharan.J. (2016) The Zika Virus: A new Threat from mosquito. *Scientia11*, 9–17.

(49) Malet, H., Massé, N., Selisko, B., Romette, J. L., Alvarez, K., Guillemot, J. C., Tolou, H., Yap, T. L., Vasudevan, S. G., Lescar, J., and Canard, B. (2008) The flavivirus polymerase as a target for drug discovery. *Antiviral Res.80*, 23–35.

(50) Luo, D., Vasudevan, S. G., and Lescar, J. (2015) The flavivirus NS2B-NS3 protease-helicase as a target for antiviral drug development. *Antiviral Res.118*, 148–158.

(51) Sironi, M., Forni, D., Clerici, M., and Cagliani, R. (2016) Nonstructural Proteins Are Preferential Positive Selection Targets in Zika Virus and Related Flaviviruses. *PLoS Negl. Trop. Dis.10*, e0004978.

(52) Sampath, A., and Padmanabhan, R. (2009) Molecular targets for flavivirus drug discovery. *Antiviral Res.81*, 6–15.

(53) Martins, I. C., Gomes‐Neto, F., Faustino, A. F., Carvalho, F. A., Carneiro, F. A., Bozza, P. T., Mohana‐Borges, R., Castanho, M. A. R. B., Almeida, F. C. L., Santos, N. C., and Da Poian, A. T. (2012) The disordered N-terminal region of dengue virus capsid protein contains a lipid-droplet-binding motif. *Biochem. J.444*, 405–415.

(54) Katoh, H., Okamoto, T., Fukuhara, T., Kambara, H., Morita, E., Mori, Y., Kamitani, W., and Matsuura, Y. (2013) Japanese encephalitis virus core protein inhibits stress granule formation through an interaction with Caprin-1 and facilitates viral propagation. *J. Virol.87*, 489–502.

(55) Hsieh, S. C., Zou, G., Tsai, W. Y., Qing, M., Chang, G. J., Shi, P. Y., and Wang, W. K. (2011) The C-terminal helical domain of dengue virus precursor membrane protein is involved in virus

assembly and entry. *Virology410*, 170–180.

(56) Yu, I.-M., Holdaway, H. a, Chipman, P. R., Kuhn, R. J., Rossmann, M. G., and Chen, J. (2009) Association of the pr peptides with dengue virus at acidic pH blocks membrane fusion. *J. Virol.83*, 12101–7.

(57) Lindenbach, B. D., and Rice, C. M. (2003) Molecular Biology of Flaviviruses. *Adv. Virus Res.59*, 23–61.

(58) Rushika Perera. (2009) Closing the door on flaviviruses: Entry as a target for antiviral drug design. *Natl. Inst. Heal.80*, 11–22.

(59) Mackenzie, J. M., Jones, M. K., and Young, P. R. (1996) Immunolocalization of the dengue virus nonstructural glycoprotein NS1 suggests a role in viral RNA replication. *Virology220*, 232–240.

(60) Flamand, M., Megret, F., Mathieu, M., Lepault, J., Rey, F. a, and Deubel, V. (1999) Dengue virus type 1 nonstructural glycoprotein NS1 is secreted from mammalian cells as a soluble hexamer in a glycosylation-dependent fashion. *J. Virol.73*, 6104–10.

(61) Yusof, R., Clum, S., Wetzel, M., Murthy, H. M. K., Padmanabhan, R., Yusof, R., Clum, S., Wetzel, M., Murthy, H. M. K., and 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 Purified NS2B / NS3 Serine Protease of Dengue Virus Type 2 Exhibits *275*, 9963–9969.

(62) Chambers, T. J., Nestorowicz, a, and Rice, C. M. (1995) Mutagenesis of the yellow fever virus NS2B/3 cleavage site: determinants of cleavage site specificity and effects on polyprotein processing and viral replication. *J. Virol.69*, 1600–1605.

(63) Natarajan, S. (2010) NS3 protease from flavivirus as a target for designing antiviral inhibitors against dengue virus. *Genet. Mol. Biol.33*, 214–219.

(64) Brooks, A. J., Johansson, M., John, A. V., Xu, Y., Jans, D. A., and Vasudevan, S. G. (2002) The interdomain region of dengue NS5 protein that binds to the viral helicase NS3 contains independently functional importin β1 and importin α/β-recognized nuclear localization signals. *J. Biol. Chem.277*, 36399–36407.

(65) Tay, M. Y. F., Saw, W. G., Zhao, Y., Chan, K. W. K., Singh, D., Chong, Y., Forwood, J. K., Ooi, E. E., Grüber, G., Lescar, J., Luo, D., and Vasudevan, S. G. (2015) The C-terminal 50 amino acid residues of dengue NS3 protein are important for NS3-NS5 interaction and viral replication. *J. Biol. Chem.290*, 2379–2394.

(66) Faye, O., Freire, C. C. M., Iamarino, A., Faye, O., de Oliveira, J. V. C., Diallo, M., Zanotto, P. M. A., and Sall, A. A. (2014) Molecular Evolution of Zika Virus during Its Emergence in the 20th Century. *PLoS Negl. Trop. Dis.8*, 36.

(67) Lim, S. P., Noble, C. G., and Shi, P.-Y. (2015) The dengue virus NS5 protein as a target for drug discovery. *Antiviral Res.119*, 57–67.

(68) Selisko, B., Wang, C., Harris, E., and Canard, B. (2014) Regulation of Flavivirus RNA synthesis and replication. *Curr. Opin. Virol.9*, 74–83.

(69) Ramharack, P. and, and Soliman, M. E. S. (2016) Zika virus drug targets: a missing link in drug design and discovery – a route map to fill the gap. *RSC Adv.6*, 68719–68731.

(70) Qu, X., Pan, X., Weidner, J., Yu, W., Alonzi, D., Xu, X., Butters, T., Block, T., Guo, J. T., and Chang, J. (2011) Inhibitors of endoplasmic reticulum α-glucosidases potently suppress hepatitis C virus virion assembly and release. *Antimicrob. Agents Chemother.55*, 1036–1044.

(71) Chang, J., Wang, L., Ma, D., Qu, X., Guo, H., Xu, X., Mason, P. M., Bourne, N., Moriarty, R., Gu, B., Guo, J. T., and Block, T. M. (2009) Novel imino sugar derivatives demonstrate potent antiviral activity against flaviviruses. *Antimicrob. Agents Chemother.53*, 1501–1508.

(72) Block, T. M., Lu, X., Mehta, A. S., Blumberg, B. S., Tennant, B., Ebling, M., Korba, B., Lansky, D. M., Jacob, G. S., and Dwek, R. A. (1998) Treatment of chronic hepadnavirus infection in a woodchuck animal model with an inhibitor of protein folding and trafficking. *Nat. Med.4*, 610–4.

(73) Davis, W. G., Blackwell, J. L., Shi, P.-Y., and Brinton, M. a. (2007) Interaction between the cellular protein eEF1A and the 3'-terminal stem-loop of West Nile virus genomic RNA facilitates viral minus-strand RNA synthesis. *J. Virol.81*, 10172–87.

(74) Baharuddin, A., Hassan, A. A., Sheng, G. C., Nasir, S. B., Othman, S., Yusof, R., Othman, R., and Rahman, N. A. (2014) Current approaches in antiviral drug discovery against the Flaviviridae family. *Curr. Pharm. Des.20*, 3428–3444.

(75) Jain, R., Coloma, J., Garcia-Sastre, A., and Aggarwal, A. K. (2016) Structure of the NS3 helicase from Zika virus. *Nat. Struct. Mol. Biol.23*, 752–754.

(76) Maestre, A. M., Caplivski, D., and Fernandez-Sesma, A. (2016) Zika Virus: More Questions Than Answers. *EBioMedicine5*, 2–3.

(77) Mastrangelo, E., Pezzullo, M., De burghgraeve, T., Kaptein, S., Pastorino, B., Dallmeier, K., De lamballerie, X., Neyts, J., Hanson, A. M., Frick, D. N., Bolognesi, M., and Milani, M. (2012) Ivermectin is a potent inhibitor of flavivirus replication specifically targeting NS3 helicase activity:

New prospects for an old drug. *J. Antimicrob. Chemother.67*, 1884–1894.

(78) Borowski, P., Lang, M., Haag, A., Schmitz, H., Choe, J., Chen, H., and Hosmane, R. S. (2002) Characterization of Imidazo [4 , 5- d] Pyridazine Nucleosides as Modulators of Unwinding Reaction Mediated by West Nile Virus Nucleoside Triphosphatase / Helicase : Evidence for Activity on the Level of Substrate and / or Enzyme *46*, 1231–1239.

(79) Hussain, H., Krohn, K., Ahmad, V. U., Miana, G. A., and Green, I. R. (2007) Lapachol: an overview. *Arkivoc2007*, 145–171.

(80) Ferreira, V. F., Ferreira, S. B., and da Silva, F. D. C. (2010) Strategies for the synthesis of bioactive pyran naphthoquinones. *Org. Biomol. Chem.8*, 4793–802.

(81) Lopes, J. N., Cruz, F. S., Docampo, R., Vasconcellos, M. E., Sampaio, M. C. R., Pinto, A. V, and Gilbert, B. (1978) In vitro and in vivo evaluation of the toxicity of 1,4-naphthoquinone and 1,2-naphthoquinone derivatives against Trypanosoma cruzi. *Ann. Trop. Med. Parasitol.72*, 523–531.

(82) Boothman, D. A., Trask, D. K., and Pardee, A. B. (1989) Inhibition of potentially lethal DNA damage repair in human tumor cells by beta-lapachone, an activator of topoisomerase I. *Cancer Res.49*, 605–12.

(83) Anderson, R. D., and Berger, N. A. (1994) International Commission for Protection Against Environmental Mutagens and Carcinogens. Mutagenicity and carcinogenicity of topoisomerase-interactive agents. *Mutat. Res.309*, 109–42.

(84) Yin, Z., Chen, Y.-L., Schul, W., Wang, Q.-Y., Gu, F., Duraiswamy, J., Kondreddi, R. R., Niyomrattanakit, P., Lakshminarayana, S. B., Goh, A., Xu, H. Y., Liu, W., Liu, B., Lim, J. Y. H., Ng, C. Y., Qing, M., Lim, C. C., Yip, A., Wang, G., Chan, W. L., Tan, H. P., Lin, K., Zhang, B., Zou, G., Bernard, K. A., Garrett, C., Beltz, K., Dong, M., Weaver, M., He, H., Pichota, A., Dartois, V., Keller, T. H., and Shi, P.-Y. (2009) An adenosine nucleoside inhibitor of dengue virus. *Proc. Natl. Acad. Sci. U. S. A.106*, 20435–9.

(85) Saito, T., and Sadoshima, J. (2016) In vitro antiviral activity of adenosine analog NITD008 against tick-borne flaviviruses *116*, 1477–1490.

(86) Yong-Qiang, D., Na-Na, Z., Chun-Feng, L., Min, T., Jia-Nan, H., Xu-Ping, X., Pei-Yong, S., and Cheng-Feng, Q. (2016) Adenosine analog NITD008 is a potent inhibitor of Zika virus. *Open Forum Infectious Disease.* 1–4.

CHAPTER 3

Introduction to Computational Chemistry

3.1 Introduction

Computational chemistry, also known as molecular modeling, is a subfield of theoretical chemistry, which focuses on solving chemically related problems by calculation[1]. It deals with the computer simulation and modeling systems, such as biomolecules, drugs, and organic or inorganic molecule. Computational chemistry consists of molecular mechanics and quantum mechanics, where the former is concerned with the physical laws of molecular systems, and the later deals with the behavior of matter on an atomic level. This chapter starts with an overview of quantum mechanics, Schrodinger's equation, a brief introduction into the various theoretical and computational tools that best describes computational chemistry, which includes Born-oppenheimer's approximation, quantum mechanics, molecular mechanics, and molecular dynamic simulations amongst others.

3.2 Quantum mechanics

Quantum mechanics is a branch of physics that is concerned with the principles that describes the interaction of matter and energy on the atomic and subatomic levels. The term quantum mechanics is attributed to Max Planck in 1900 and was derived from the Latin word *quanta* meaning "how much". J.J Thompson discovered that the laws of classical Newtonian physics cannot be applied at the atomic level, resulting in new equations and principles being developed to mathematically describe the behavior of matter and energy on a quantum scale. One of the important principles of quantum mechanics is the wave particle duality, which states that matter and light can act as either waves or particles depending on the circumstance of observation, this being better explained by solving the Schrodinger equation.[2]

3.3 Schrodinger's equation

Schrodinger equation was developed by Erwin Schrodinger, which predicts the future behavior of a dynamic system, and is a fundamental physics equation for describing quantum mechanical behavior. The Schrodinger equation has two types, the time dependent and the time independent equations, the former describing the quantum system evolution[3].

Schrodinger's equation is given as

$$\mathbf{H}\psi = \mathbf{E}\psi \qquad \text{Eq. 3.1}$$

Where H is the Hamiltonian operator, ψ is a wave function and E is the energy of the state.

The Hamiltonian equation is also given as the sum of the kinetic (T) and energy potential (V)

$$\mathbf{H} = \mathbf{T} + \mathbf{V} \qquad \text{Eq. 3.2}$$

Where H is the Hamiltonian operator, T is the kinetic energy and V is the potential energy

In a more detailed form, Hamiltonian equation can be given as

$$\boldsymbol{H} = -\frac{h^2}{2m_e}\sum_i \nabla_i^2 - \frac{h^2}{2}\sum_A \frac{1}{M_A}\nabla_A^2 - \sum_A \sum_i \frac{z_A e^2}{r_{Ai}} + \sum_i \sum_{j>i} \frac{e^2}{r_{ij}} + \sum_A \sum_{B>A} \frac{z_A z_B e^2}{R_{AB}} \qquad \text{Eq. 3.3}$$

Where *A* and *B* are the nuclei, *i* and *j* are the electrons, *MA* mass of nucleus *A*, *me* mass of an electron, R_{AB} the distance between nuclei *A* and *B*, r_{ij} the distance between electrons *i* and *j*, ^{Z}A the charge of nucleus *A*, r_{Ai} the distance between nucleus *A* and electron *i*.

3.4 Born-Oppenheimer approximation

In 1927, Max Born and J. Robert Oppenheimer developed the Born-Oppenheimer approximation, leads to a simplification of the Schrodinger's equation by assuming that the nuclei and the electron in a molecule can be separated. It considers the momenta of the electron and nuclei to be of similar magnitude and solves the time dependent Schrodinger equation, resulting from Hamiltonian equation, by assuming the nuclei of a molecule to be stationary. It then solves the electronic ground-state first, calculated the energy of the system and solve for nuclear motion[4]. The energy of a molecule is a function of the electron coordinates, and depends on the nuclear coordinate parameter that usually defines the geometry, which makes the concept of Potential Energy Surface (PES) possible.

$$\boldsymbol{T}^{elec} = \left[-\frac{h^2}{8\pi^2 m}\sum_i^{electrons}\left(\frac{\partial^2}{\partial x^2} + \frac{\partial^2}{\partial y^2} + \frac{\partial^2}{\partial z^2}\right)\right] \qquad \text{Eq. 3.4}$$

Schrodinger's equation for fixed nuclei electrons is given as

$$\mathbf{H}^{\mathbf{elec}}\boldsymbol{\varphi}^{\mathbf{elec}}(\mathbf{r},\mathbf{R}) = \mathbf{E}^{\mathbf{eff}}(\mathbf{R})\boldsymbol{\varphi}^{\mathbf{elec}}(\mathbf{r},\mathbf{R}) \qquad \text{Eq. 3.5}$$

Solving this equation for other fixed positions of concern will produce a Potential Energy Surface (PES).

3.5 Potential Energy surface (PES)

The Potential Energy surface (PES) is a central concept in computational chemistry that describes the energy of a molecule in terms of its structure. It is a graphical or mathematical relationship between the energy of a molecule and its geometry[5]. PES is multi-dimensional, and for a diatomic molecule is a two-dimensional plot, with the potential energy at the bond distance on the Y axis and the internuclear separation at the X-axis. For larger systems, there are many dimensions of the surfaces and there are degrees of freedom within the molecule. PES is often represented by the illustration in Figure 3.1.

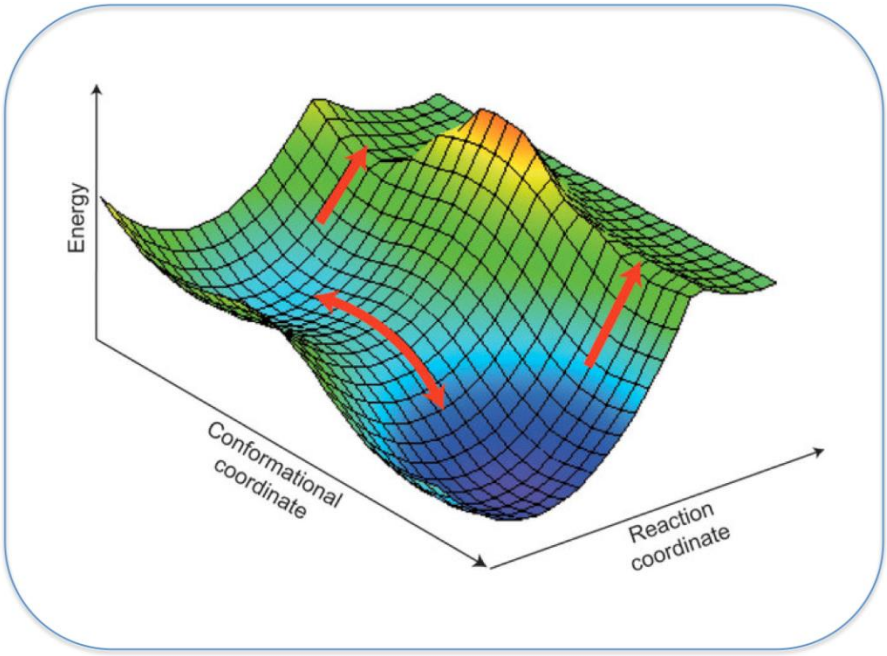


Figure 3.1. Graphical illustration of multi-dimensional potential energy surface[6].

3.6 Molecular mechanics

Molecular Mechanics (MM) is a computational method that computes the potential energy surface for the arrangement of atoms using potential functions that are derived using classical physics. MM aims to predict the structure and physical properties of molecules by calculating the steric energy (lowest energy due to conformation) of a molecule, which is assumed to rise from specific interactions with another molecule. These interactions have the ability of bonds to stretch and bend beyond their equilibrium lengths and angles. In MM, atoms are considered to be spheres, while bonds are represented as springs[7]. The deformation of the spring can be used to describe the stretching, bending and twisting ability of a bond. Non-bonded atoms also interact via electrostatic attraction/repulsion, van der Waals attraction and stearic repulsion. To measure the contribution of

each of the interactions in a molecule, the stearic energy of the molecule is given as the sum of the energies of the interactions.

$$E_{tot} = E_{str} + E_{bend} + E_{tor} + E_{vdw} + E_{elec} \quad \text{Eq.3.6}$$

Where E_{tot} is the total energy, $E_{str}$is the bond-stretching energy, E_{bend} is the angle-bending energy, E_{tors} is the torsional energy, E_{vdw} is the van der Waals energy and E_{elec} is the electrostatic energy. The bond stretching, bending and torsion interactions are called bonded interactions, while the Van der Waals and electrostatic interactions are between non-bonded atoms.

3.6.1 Force field

A force field is a mathematical expression describing the dependence of the energy of a system on the coordinates of its particles[8]. It serves as the basic model for molecular mechanics and molecular dynamics calculations. Force fields predict the entire energy of a system based upon the location and orientation of the atoms and molecules relative to each other[9]. The most commonly used force field include AMBER[10], GROMOS[11], CHARMM[12] and NAMD[13]. These force fields consists of an analytical form of the interatomic potential and a set of parameters entering into this form[8]. The parameters are typically obtained either from *ab initio* or semi-empirical quantum mechanical calculations, or by fitting to experimental data, such as neutron, X-ray and electron diffraction, NMR, infrared, Raman and neutron spectroscopy[8]. All force fields have different parameters; thus, they must be adjusted to give the results of the forces acting within a molecule. The AMBER force field was used for this study, with the General AMBER Force Field (GAFF) parameters being applied for ligand parameterization and the standard AMBER force field for the protein.

3.7 Molecular dynamic simulation

Molecular dynamics (MD) is a technique for computationally simulating complex systems modeled at the atomic level. MD is the solution of the classical equations of motion for atoms and molecules to obtain the time evolution of a system, and is applied to many-particle systems, as a general analytical solution is not possible, making it necessary to use numerical methods and computers[14]. It allows the generation of atomic trajectories of many particle systems by integrating Newton's equations of motion for atoms on an energy surface;

$$\mathbf{F}i = mid^2r_i\,(t) \quad \text{E.q 3.7}$$

$$dt^2$$

Where $r_i\,(t) = (x_i\,(t), y_i\,(t), z_i\,(t))$ is the (i) particle position vector, F_i the force acting on

(i) particle at time t and m_i the particle mass.

The aim of computer simulations is to facilitate the understanding of the properties of assemblies of molecules in terms of their structure and the microscopic interactions between them, which serves as a complement to conventional experiments[15].Computer simulations act as a bridge between microscopic length and time scales and the macroscopic world of the laboratory with exact and accurate predictions. It acts as a bridge between theory and experiment, in the sense that a theory can be tested by conducting a simulation and the result of the simulation can be compared with the experimental result[16]. In MD calculations, the system is either treated with MM or QM methods, and the results are usually represented as a trajectory that specifies how the particle position and velocity varies with time.

3.8 Molecular docking

Molecular docking is a valuable tool in Structural Based Drug Design (SBDD), and aim to predict the ligand-receptor complex structure using computation methods. To achieve this, a three-dimensional representation of the protein target must be available, can be obtained either from experimental information such as X-ray crystallography or NMR or from a homology model[17]. Thereafter, the ligand binding site is located and docking can commence using docking programs such as AutoDock[18], DOCK[19], PhDOCK[20], Surflex[20], FlexX [21]and GLIDE [22]. Docking can be related to the "lock and key" mechanism that was first proposed by Fischer over 100 hundred years ago. Later, the theory of molecular recognition was developed by Koshland to encompass conformational changes during the binding event[23]. Docking entails two things; sampling and scoring. Sampling refers to generating putative ligand binding conformations near a binding site of a protein, while scoring predicts the binding tightness for individual ligand conformations with a physical or empirical energy function[24]. The ligand-receptor binding energy is calculated as follows:

$$\boldsymbol{E_{binding}=E_{target}+E_{ligand}+E_{target-ligand}} \qquad \text{E.q 3.8}$$

Docking consists of two types: rigid and flexible docking. In rigid docking, the protein and ligand are kept rigid, while in flexible docking, flexibility is allowed either for the protein or ligand, or both. The docking method used in this research study was the advanced version of AutoDock,

AutoDock Vina[18], as it enables predicting the flexible ligands' binding affinity to the NS3 helicase protein.

3.9 Binding free energy.

Free energy calculations have become an important approach in computational chemistry, as it enables a detailed investigation of the energetic factors that are responsible for the molecular stability or binding affinity of molecular systems[25]. They are typically based on molecular dynamics simulations of the receptor--ligand complex. The Molecular Mechanics Poisson–Boltzmann Surface Area (MM-PBSA)[26] is a popular approach to estimate the binding free energies molecular systems. Other methods that are available to calculate free energies, includes Free Energy Perturbation[27], Replica Exchange Free Energy Perturbation[28], and Thermodynamic Integration[29]. These methods are highly rigorous and demanding, and become more expensive as the size of the system increases[26]. In the MM-PBSA, the binding free energy is the sum of the energy and configurational entropy that is associated with complex formation in the gas-phase, and the difference in solvation free energies between the complex and the unbound molecules. A brief explanation about free binding energy calculation is provided in Chapter 4.

3.10 Principal Component Analysis (PCA)

Principal Components Analysis (PCA) also known as essential dynamics is one of the sophisticated technique used in the analysis of trajectory for understanding the dynamics of biological system at molecular level[30]. It describes harmonic atomic displacement and identifies major conformational changes in protein structures during MD simulation[31]. PCA describes the direction of motion of the protein (eigenvectors) and amplitude their amplitude (eigenvalues)[32]. This technique uses sophisticated underlying mathematical principles to transforms a number of possibly correlated variables into a smaller number of variables, called principal components[33]. It reduces the original data by containing large variables into principal component, the easy spotting of the resulting trends, patterns and outliers would have been impossible without PCA[33]. Each principal component is a linear combination of the original variables, while the amount of information expressed by each principal component is called its variance. Principal components are useful, because utilizing only the first few principal components that explain a substantial proportion of the total variation can reduce the complexity in interpretation that can be caused by having a large number of interrelated variables.

References

(1) Jensen, F. (2007) Introduction to computational chemistry. John Wiley & Sons.

(2) Tannor, J. D. (2010) introduction to quantum mechanics: a time dependent perspective. John Wiley and Sons, Houston

(3) 3. Lewars, E. G. (2011) Computational Chemistry, *Comp. Theor. Chem.*, 9-43.

(4) Bechstedt, F. (2015) Born-Oppenheimer Approximation, in *Many-Body Approach to Electronic Excitations*, pp 3–12.

(5) Einstein, A. The concept of Potential Energy Surface. Computational Chemistry, Introduction to the Theory and Applications of Molecular and Quantum Mechanics; Springer, 2011; pp 9-11 .

(6) Glowacki, D. R., Harvey, J. N., and Mulholland, A. J. (2012) Taking Ockham's razor to enzyme dynamics and catalysis. *Nat Chem4*, 169–176.

(7) Bornemann, F. A., Nettesheim, P., Schu, C., and Introduction, I. (1996) Quantum-classical molecular dynamics as an approximation to full quantum dynamics ¨. *J. Chem. Phys. 105*, 1074–1083.

(8) González, M. A. (2011) Force fields and molecular dynamics simulations. *Collect. SFN12*, 169–200.

(9) Handley, C. M., and Popelier, P. L. A. (2010) Potential Energy Surfaces Fitted by Artificial Neural Networks. *J. Phys. Chem. A114*, 3371–3383.

(10) Pearlman, D. A., Case, D. A., Caldwell, J. W., Ross, W. S., Cheatham, T. E., DeBolt, S., Ferguson, D., Seibel, G., and Kollman, P. (1995) AMBER, a package of computer programs for applying molecular mechanics, normal mode analysis, molecular dynamics and free energy calculations to simulate the structural and energetic properties of molecules. *Comput. Phys. Commun.91*, 1–41.

(11) Christen, M., Hunenberger, P. H., Bakowies, D., Baron, R., Burgi, R., Geerke, D. P., Heinz, T. N., Kastenholz, M. A., Krautler, V., Oostenbrink, C., Peter, C., Trzesniak, D., and Van Gunsteren, W. F. (2005) The GROMOS software for biomolecular simulation: GROMOS05. *J. Comput. Chem.26*, 1719–1751.

(12) Gumbart, J., Aksimentiev, A., Tajkhorshid, E., Wang, Y., and Schulten, K. (2007) Molecular dynamics simulations of proteins in lipid bilayers. *J. Comput. Chem.15*, 80–90.

(13) Phillips, J. C., Braun, R., Wang, W., Gumbart, J., Tajkhorshid, E., Villa, E., Chipot, C., Skeel, R. D., Kale, L., and Schulten, K. (2005) Scalable molecular dynamics with NAMD. *J. Comput. Chem.26*, 1781–1802.

(14) Petrenko, R., and Meller, J. (2010) Molecular Dynamics, in *Encyclopedia of Life Sciences*. John Wiley & Sons, Ltd, Chichester, UK.

(15) Allen, M. (2004) Introduction to molecular dynamics simulation. *Comput. Soft Matter From Synth. Polym. to proteins23*, 1–28.

(16) Allen, M. (2004) Introduction to molecular dynamics simulation. *Comput. Soft Matter From Synth. Polym. to proteins23*, 1–28.

(17) Oledzki, P. R., Laurie, A. T. R., and Jackson, R. M. (2006) Specialist Review Protein–ligand docking and structure-based drug design. *Spec. Rev.* 1–17.

(18) Morris, G. M., Goodsell, D. S., Halliday, R. S., Huey, R., Hart, W. E., Belew, R. K., and Olson, A. J. (1998) Automated docking using a lamarckian genetic algorithm and an empirical binding free energy function. *J. Comput. Chem.19*, 1639–1662.

(19) Ewing, T. J. A., Makino, S., Skillman, A. G., and Kuntz, I. D. (2001) DOCK 4 . 0 : Search strategies for automated molecular docking of flexible molecule databases. *Comput. Mol. Des.15*, 411–428.

(20) Joseph-mccarthy, D., Joseph-mccarthy, D., and Alvarez, J. C. (2014) Automated generation of MCSS-derived pharmacophoric DOCK site points for searching multiconformation Databases. *Proteins Struct. Funct. Bioinforma.51*, 189–202.

(21) Warren, G. L., Andrews, C. V, Capelli, A., Clarke, B., LaLonde, J., Lambert, M. H., Lindvall, M., Nevins, N., Semus, S. F., Senger, S., Tedesco, G., Wall, I. D., Woolven, J. M., Peishoff, C. E., and Head, M. S. (2006) A critical assessment of docking programs and scoring functions.Capelli, A.LaLonde, J.Semus, S. F.Head, M. S. *J. Med. Chem.49*, 5912–5931.

(22) Friesner, R. A., Banks, J. L., Murphy, R. B., Halgren, T. A., Klicic, J. J., and Mainz, D. T. (2004) Glide: a new approach for rapid, accurate docking and scoring. 1. Method and assessment of docking accuracy. *J. Med. Chem.47*, 1750–1759.

(23) Koshland, D. E. (2004) Crazy, but correct. *Nature432*, 447.

(24) Huang, S. Y., and Zou, X. (2010) Advances and challenges in Protein-ligand docking. *Int. J. Mol. Sci.11*, 3016–3034.

(25) Homeyer, N., and Gohlke, H. (2012) Free Energy Calculations by the Molecular Mechanics Poisson À Boltzmann Surface Area Method. *Mol. Inform.31*, 114–122.

(26) Iii, B. R. M., Mcgee, T. D., Swails, J. M., Homeyer, N., Gohlke, H., and Roitberg, A. E. (2012) MMPBSA . py : An Efficient Program for End-State Free Energy Calculations. *J. Chem. theory Comput.* 1–27.

(27) Zwanzig, R. W. (1954) High-Temperature Equation of State by a Perturbation Method. I. Nonpolar Gases. *J. Chem. Phys.22*, 1420–1426.

(28) Bergonzo, C., Henriksen, N. M., Roe, D. R., Swails, J. M., Roitberg, A. E., and Cheatham, T. E. (2014) Multidimensional replica exchange molecular dynamics yields a converged ensemble of an RNA tetranucleotide. *J. Chem. Theory Comput.10*, 492–499.

(29) Olivares, W., and McQuarrie, D. A. (1975) On the theory of ionic solutions. *Biophys. J.15*, 143–162.

(30) Appiah-Kubi, P., and Soliman, M. E. S. (2016) Dual anti-inflammatory and selective inhibition mechanism of leukotriene A4 hydrolase/aminopeptidase: insights from comparative molecular dynamics and binding free energy analyses. *J. Biomol. Struct. Dyn.1102*, 1–16.

(31) Jorgensen, W. L., Chandrasekhar, J., Madura, J. D., Impey, R. W., and Klein, M. L. (1983) Comparison of simple potential functions for simulating liquid water. *J. Chem. Phys.79*, 926.

(32) Cocco, S., Monasson, R., and Weigt, M. (2013) From Principal Component to Direct Coupling Analysis of Coevolution in Proteins: Low-Eigenvalue Modes are Needed for Structure Prediction. *PLoS Comput. Biol.9*.

(33) Jolliffe, I. T. (2005) Principal component analysis. *Appl. Opt.44*, 6486.

CHAPTER 4

Characterizing the Ligand Binding Landscape of Zika NS3 Helicase-Promising Lead Compounds as Potential Inhibitors.

Sofiat Oguntade[a], Pritika Ramharack [a] and Mahmoud E. S. Soliman[a*]

[a]Discipline of Pharmaceutical Sciences, School of Health Sciences, University of KwaZulu-Natal, Westville Campus, Durban 4001, South Africa

*Corresponding Author: Mahmoud E.S. Soliman email: soliman@ukzn.ac.za

1. Dean and Head of School of Health Sciences, Full Professor: Pharmaceutical Sciences, University of KwaZulu-Natal, Westville Campus, Durban 4001, South Africa.

2. Department of Pharmaceutical Organic Chemistry, Faculty of Pharmacy, Zagazig University, Zagazig, Egypt.

3. College of Pharmacy and Pharmaceutical Sciences, Florida Agricultural and Mechanical University, FAMU, Tallahassee, Florida 32307, USA.

Email: soliman@ukzn.ac.za

Telephone: +27 (0) 31 260 8048, Fax: +27 (0) 31 260 7872

Abstract

Aim: This study aims to provide insight into the binding features of the ATPase and ssRNA sites of the NS3 helicase. **Methods:** Clinically approved *flavivirus* inhibitors were docked to the corresponding active sites of the protein and the three best compounds were validated with molecular dynamic simulations. **Result:** Binding of Ivermectin to ssRNA site and Lapachol and HMC-HO1α to the ATPase site allowed for conformational rigidity of the Zika NS3 helicase, thus stabilizing residue fluctuations and allowing for protein stability. Favorable free binding energies were also noted between compounds and the helicase, thus supporting the intermolecular forces at the helicase active site. **Conclusion:** The

pharmacophoric characteristics found in Lapachol, HMC-HO1α and Ivermectin may be utilized in the design of a potent hybrid drug that can show efficient inhibition of a multitude of diseases including the detrimental co-infection of ZIKV, Dengue and Chikungunya.

Keywords:

Zika virus, NS3 helicase, molecular dynamic simulations, free binding energy

4.1. Introduction

Zika virus (ZIKV) is a positive-sense, single stranded RNA arbovirus belonging to the genus *flavivirus* and family *flaviviridae* [1]. The virus was first discovered in a forest in Uganda called the Zika forest near lake victoria in 1947, thus coining the virus's name [2,3]. The virus was then isolated in the blood of a sentinel Rhesus monkey during research on the Yellow fever virus [4], while a second isolation was done in 1948 at the same site [5]. ZIKV virus has a wide geographical distribution including Africa (Uganda, Egypt, Gabon), Asia (India, Malaysia, Vietnam, Thailand, Indonesia), and Micronesia [6]. This has been demonstrated through viral isolations and serologic studies [7,8]. Although isolations of the virus were analyzed, researchers only detected the virus in humans in 1952 when neutralizing antibodies were picked up in infected sera. Scientists Boorman and Porterfield subsequently studied the transmission of viruses from mosquito to primates and based on further isolations from both mosquito and monkey concluded that mosquitoes acted as vectors for ZIKV [1].

The rapid spread of the virus across continents is primarily due to vector transmission via the *Aedes aegypti, Aedes albopictus and Aedes africanus* mosquito [2]. These Vectors are endemic to tropical and sub-tropical areas. However, due to evolving climates, the mosquitoes have expanded their habitat, thus increasing the number of mosquitoes as vectors of *flaviviruses* [9–11]. However, other routes of transmission have been reported, including, sexual transmission [10,12], perinatal transmission, and blood transfusion [13]. The

symptoms following ZIKV viral infection are mild headache, maculopapular rash, fever, malaise, conjunctivitis and arthralgia. These symptoms are shared with other related *flavivirus*es, including Dengue virus, Yellow fever virus, West Nile, St. Louis encephalitis virus and Japanese Encephalitis virus [15,16]. The most recent and devastating outbreak of ZIKV occurred in Brazil, at the end of 2015. The virus has, to date, rampaged South America by being evidenced as a leading cause of microcephaly by prenatal transmission [17]. Increasing scientific evidence now shows that the virus can pass through the blood-brain-barrier and infect neural cells, thus playing a role in diseases such as microcephaly and Gullian-Barré Syndrome [18].

The ZIKV genome contains 10.7kb single stranded RNA, which contains a large polyprotein, which cleaves into 3 structural proteins (envelope, E; membrane precursor, PrM; and capsid, C) and seven non-structural proteins (NS1, NS2A, NS2B, NS3, NS4A, NS4B, and NS5), of which, the NS3 helicase plays a pivotal role in viral replication and RNA synthesis. Presently, researchers are focusing on the structural and non-structural viral proteins for the development of drugs [19], due to their crucial characteristics in viral replication [20]. The NS3 helicase (Figure 4.1) has three domains and two binding sites, being the adenosine triphosphate (ATP) and single stranded ribonucleic acid (ssRNA) site [21,22]. Inhibiting both the ATP and ssRNA sites will be crucial in the inhibition of the NS3 helicase as studies have shown that each domain may act independently from the other [23]. However, the close proximity of the two binding sites can bring about the possibility of designing a single inhibitor that can span both sites [3].

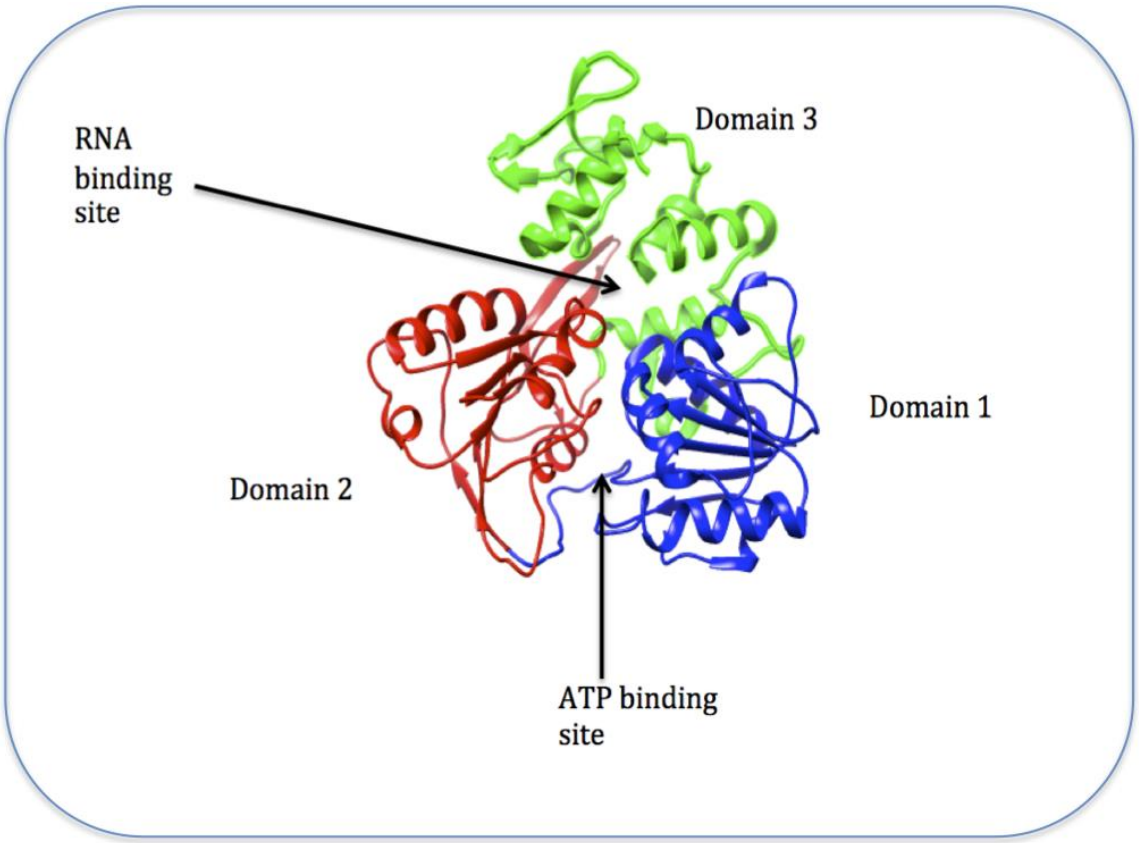


Figure 4.1: Structure of NS3 ZIKV Helicase (PBD 5JMT) [21]. Domain 1 (blue: residue 175-332) and domain 2 (red: residue 333-481) are seen facing each other and domain 3 (green: residue 482-617) lying above the other 2 domains. The ATP binding site is located in the cleft between domain 1 and domain 2 and the ssRNA binding site is located at the tunnel that separates domain 3 from the other 2 domains [22].

Due to the rapid spread of the disease on a global scale and the detrimental long-term complications, researchers such as Barrows *et al* (2016) have turned to 'repurposing' *flavivirus* FDA approved drugs rather than the lengthy process of designing and synthesizing new drugs [4]. One of the most widely used anthelmintic drugs, Ivermectin, has been evidenced to have potent inhibitory effects on *flaviviruses* by acting as a competitive inhibitor of viral ssRNA at the RNA binding site of the NS3 helicase [6,7]. Barrows *et al* (2016)

validated Ivermectin as a potent ZIKV inhibitor in an *in vitro* screening study, alongside 17 other FDA-approved *flavivirus* drugs as well as daptomycin, which had no previous anti-viral activity [4]. Another potential drug candidate against ZIKV is the adenosine nucleoside analog, NITD008, which has been reported to have competitive inhibitory properties against adenosine substrates *in vitro* and *in vivo* [8]. However, other reports have also shown elevated toxicity levels in preclinical animal testing [9].

One of the major challenges of ZIKV is its ability to co-infect the host. Multiple cases reporting Chikungunya, Dengue and Zika co-infection have been identified, leading to potentially exacerbated neurological effects on the host and fetus [10]. By this end, identifying potential inhibitors against ZIKV that have already been approved as a Dengue or Chikungunya treatment would be beneficial as it would be less toxic than administering multiple drugs to a patient [11].

Although studies have reported ZIKV drug discovery, no FDA approved drugs are presently available. There is also a lack in literature regarding the structural and conformational features of the protein, thus designing effective novel small drug molecule inhibitors may be challenging.

In this study, we have utilized clinically approved *flavivirus* NS3 small molecule inhibitors to analyze the binding affinity and stability of the ZIKV NS3 domains via Molecular Dynamics (MD) simulations, thus mapping out binding hotspots and landscaping interactions of the complexes. In addition to this, we will employ Accelerated Molecular Dynamics (aMD) to run the simulation for a longer time frame to ensure sufficient conformational sampling and accurate physical force field. Accelerated MD is an enhanced sampling technique that operates by modifying potential energy, reducing the height of local barriers and accelerating transition between different low energy states [24-26]. This will enable the sampling of

distinct bimolecular conformations and rare barrier-crossing events that cannot be easily accessed in a conventional MD simulation, thereby improving the efficiency of convectional MD [27].

4.2 Computational methodology

4.2.1. Protein structure preparation

The crystal structure of the *Escherichia coli strain of* Zika virus NS3 helicase was retrieved from protein data bank (PDB: 5JMT). It was then prepared for molecular docking by stripping it off water molecules using UCSF CHIMERA [28] and adding the necessary hydrogen atoms using Molegro Molecular Viewer (MMV) [29].

4.2.2. Molecular docking

Molecular docking was performed on 10 ligands: 6 Naphthoquinones (Lapachol, Atovaquone, Parvaquone, Buparvaquone, α-Lapachone, β-Lapachone)[30], 3 purine nucleoside analogues (1-(2'-deoxy-α-D-ribofuranosyl)imidazo[4,5-*d*]pyridazine-4,7(5*H*,6*H*)dione)(HMC-HO1α), 1-(2'-*O*-methyl- α-D-ribofuranosyl)imidazo[4,5-*d*]pyridazine- 4,7(5*H*,6*H*)-dione(HMC-HO4) and 1-(β-D-ribofuranosyl)imidazo [4,5-*d*]pyridazine-4,7(5*H*,6*H*)-dione(HMC-HO5)) [31,32] and Ivermectin [33]. Each of the compounds were then downloaded from PubChem [34], converted to mol2 format and assessed using MMV to ensure that they display the correct bond angle and hybridization state. The 2D structures of the ligands are given in the supplementary material (Figure S1).

Docking was carried out with the Autodock Vina software [35]. Ivermectin was docked at the ssRNA binding site, while the rest were docked at the ATPase binding site. The grid box

parameters for the 2 sites are given in Figure S2. Of the 10 ligands docked into the active site of NS3 protein structure, the best 3 complexes were chosen and subsequently subjected to 10ns of accelerated-MD.

4.2.3 Molecular dynamic simulations

Molecular Dynamic (MD) simulations were performed on the 3 complexes using the graphics processor unit (GPU) version of the PMEMD engine provided with the AMBER 14 package [36,37]. The Antechamber module was used to generate atomic partial charges for the ligands using general amber force field GAFF force field [38]. The protein was described using the FF14SB of the Amber force field [39]. The LEAP module in AMBER 14 was used to generate topologies for the system by adding protons and counter ions to neutralize the system [36]. Subsequently, the complexes were then solvated in a TIP3P [40] octahedron water box with 8Å away from the water box edge. The Periodic boundary conditions were employed and the particle-mesh Ewald method (PME) in AMBER 14 was used to treat the long-range electrostatic interactions with a non-bonding cut-off distance of 10 Å.

Minimization of the systems were performed with a restraint potential of 500 kcal/mol $Å^2$ to treat the solute for 1000 steepest descent steps using the SANDER module of the AMBER 14 program, followed by 1000 steps of conjugate gradient minimization. The systems were then minimized over 1000 steps with unrestrained conjugate gradient. Gradually, the systems were heated from 0 to 300 K for 50ps, such that the system maintained a fixed number of atoms and a fixed volume, that is, a canonical (NVT) ensemble.

The entire system was then equilibrated at 300 K with a 2fs time step in the NPT ensemble for 500ps, and Berendsen temperature coupling [39] was used to maintain a constant pressure at 1 bar. The SHAKE algorithm [40] was employed on all atoms to constrain the bonds of all

hydrogen atoms. With no restraints imposed, an initial production run was performed for 10ns in an isothermal isobaric (NPT) ensemble using a Berendsen barostat with a target pressure of 1 bar and a pressure-coupling constant of 2ps. The systems were subsequently subjected to 10ns of accelerated MD using a set of parameters calculated from the potential energy of the converged system (Table 4.1). Coordinates were saved every 1ps and the trajectories were analyzed every 1ps using the PTRAJ module of AMBER 14. Each system was consequently subjected to post molecular dynamic analysis including root mean square fluctuation (RMSF), root mean square deviation (RMSD) and radius of gyration (Rg). Included in analysis was the ligand-residue profile [25]. Visualization of trajectories was conducted in Chimera [28], while the results were analyzed, and plots were generated with aid of Origin software [41].

Table 4.1: Calculated parameters for running accelerated molecular dynamics**.**

	ethreshP (kcal/mol)	ethreshD (Kcal/mol)	alphaP (kcal/mol)	alphaD (kcal/mol)
Lapachol System	-144728	9424.98	1404.4	355.2
Ivermectin System	-144600	9465.1	1425.2	355.2
HMC-HO1αSystem	-144698	9455.25	1404.4	355.2

4.2.4 Thermodynamic calculations

Over the years, molecular mechanics/generalized-born surface area (MM/GBSA) method of binding free energy calculations have proved to be a practicable means of understanding the ligand-residue landscape binding in various biological macromolecules [42–45]. Therefore,

MM/GBSA approach was employed to calculate the binding free energies of Ivermectin, Lapachol andHMC-HO1α bound to NS3 helicase protein. To achieve this, 1,000 snapshots were extracted from each of the 20ns trajectories. The following equation describes the calculations of binding free energy.

$$\Delta G_{bind}=G_{complex}-G_{receptor}-G_{ligand} \quad (1)$$

$$\Delta G_{bind}=E_{gas}+G_{sol}-TS \quad (2)$$

$$E_{gas}=E_{int}+E_{vdw}+E_{ele} \quad (3)$$

$$G_{sol}=G_{GB}+G_{SA} \quad (4)$$

$$G_{SA}=\gamma SASA \quad (5)$$

The term E_{gas} denotes the gas-phase energy that consists of the internal energy E_{int}, Coulomb energy E_{ele}, and the van der Waals energies E_{vdw}. E_{gas} was directly estimated from the FF14SB force field terms. The solvation free energy, G_{sol}, is estimated from the energy contribution from the polar states, G_{GB} and non-polar states, G_{SA}. The non-polar solvation energy, G_{SA}, is determined from the solvent accessible surface area (SASA) using a water probe radius of 1.4 Å, whereas the polar solvation, G_{GB}, contribution is estimated by solving the generalized born (GB) equation. S and T denote the total entropy of the solute and temperature, respectively.

4.2.5. Per-residue energy decomposition analysis

Per-residue free energy decomposition was carried out to obtain the contribution of each residue to the total binding free energy profile between the inhibitors Ivermectin, Lapachol and HMC-HO1α with the NS3 helicase protein. This was achieved using the MM/GBSA

method in AMBER 14[25].

4.3. Results and Discussion

4.3.1 Docking result and validation

Molecular docking is one of the routinely used methods in molecular modeling and drug design. It is used to predict the conformation of small molecule (ligand) within the appropriate binding site, making it a valuable tool in drug discovery [46,47]. Furthermore, molecular docking ranks docked compounds based on the binding affinity of the ligand to the receptor (Figure S3).

In this study, 10 compounds were chosen to dock into the NS3 helicase based on their inhibitory characteristics at *flavivirus* ATPase/ ssRNA sites [30,31,33]. Of the 10 compounds, the best 3 were chosen for subsequent conformational and binding mode analysis. Lapachol and HMC-HO1α were chosen from the naphthoquinones and purine nucleoside analogues respectively because they portrayed the most optimal docked conformation from the molecular docking studies that were carried out. Ivermectin was docked into the ssRNA site due to its high potency as a *flavivirus* inhibitor [33,48].

Validation of molecular docking was done by superimposing each of the docked complexes with the PDB structures of their natural substrates for ssRNA and ATPase site (PDB code: 5GJB and 5GJC). This was also achieved by comparing the RMSD of each ligand with the natural substrate. The best conformation with the lowest R.MSD (< 1.5 Å) value for each ligand were chosen. The results of the superimposition are shown in Figure 4.2. (5JMT-green, 5GJB and 5GJC - magenta, ATP and RNA- red, Lapachol, HMC-HO1α and Ivermectin –blue).

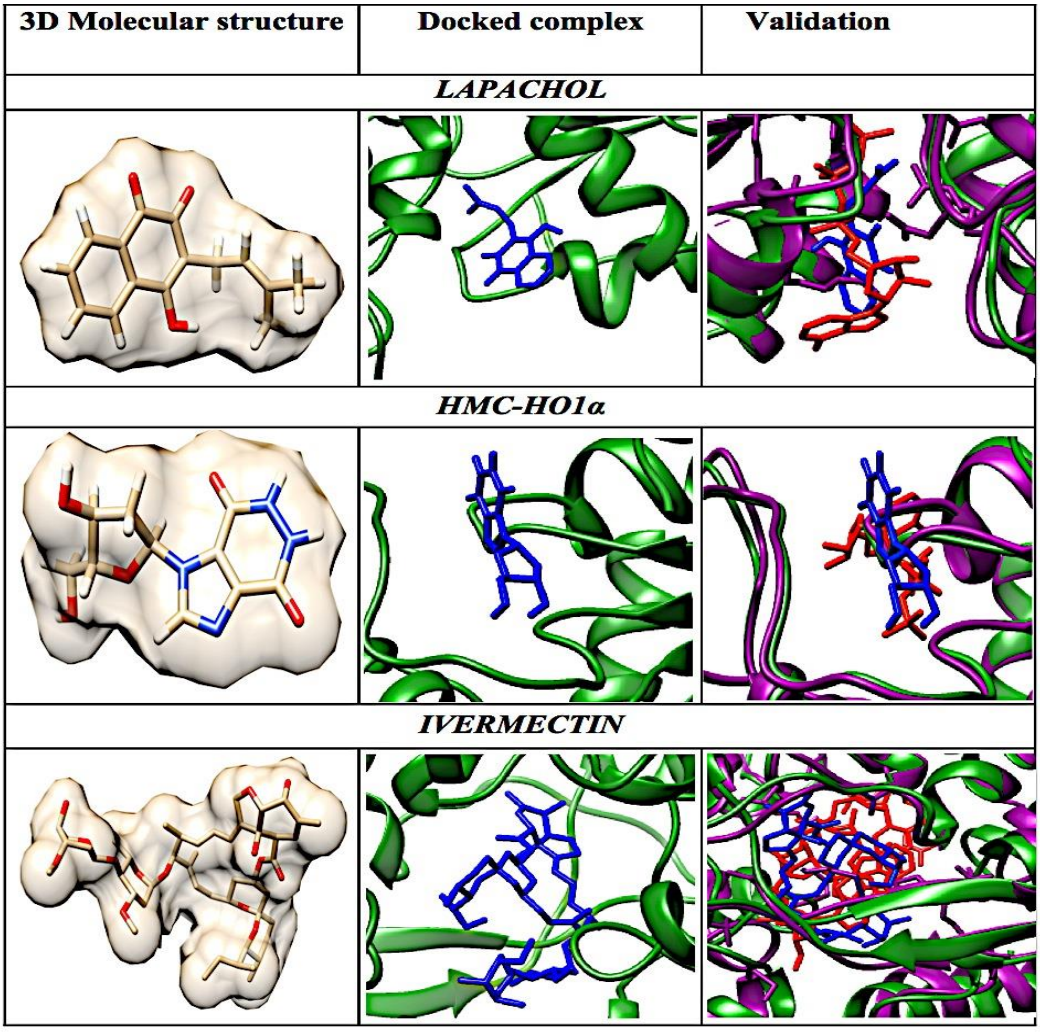


Figure 4.2: 2D structure, docked complexes and validation of the docked complexes.

4.3.2 Molecular dynamics simulation and post molecular dynamics analysis

A frequently overlooked side of molecular docking is the flexibility of the binding target. The ligand and receptor usually undergo conformational changes before binding and sometimes the ligand fits in with little mobility. To ensure the stability of the complex, the 3 complexes were subjected to aMD [47,49,50].

4.3.2.1 Systems stability

The stability of the systems was investigated by assessing the Root Mean Square Deviation

(RMSD) regarding the Cα-backbone atoms of the 3D structure during the simulation (Figure 4.3). Equilibrium was attained after 2,000ps and the overall average RMSD value for Lapachol, HMC-HO1α and Ivermectin measured 0.98Å, 0.97Å and 0.95Å respectively. The results at the ATPase site exhibited similar stability between the HMC-HO1α-ATPase system and the Lapachol-ATPase system, whereas, the Ivermectin-ssRNA complex demonstrated the lowest average RMSD from all three systems. This is indicative of a more stable complex, justifying Ivermectin as a potent *flavivirus* inhibitor. The RMSD plot further postulates that the binding of the three ligands at two different active sites of the protein still allowed for conformational rigidity compared to the unstable free protein, which yielded an elevated average of 1.76Å. It can therefore be deduced that all three ligands allowed for structural stability of the NS3 Helicase protein.

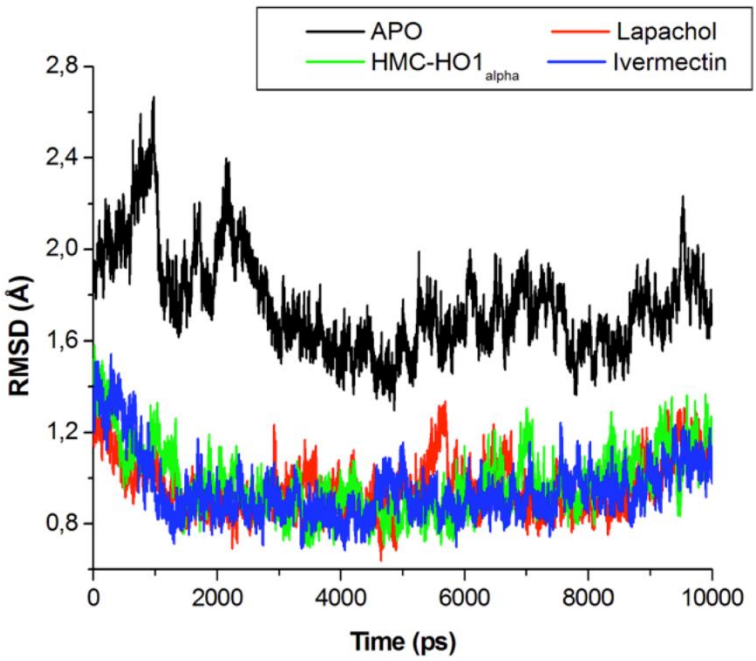


Figure 4.3: C-alpha RMSD backbone Plot for NS3 Helicase free and ligand bound conformations. The ligands Lapachol, HMC-HO1α and Ivermectin are seen to stabilize the protein as compared to the fluctuating free protein.

4.3.2.2 Root Mean Square Fluctuations (RMSF)

Root Mean Square of Fluctuations (RMSF) were analyzed to show the mobility of each of the residues found in the protein, thereby giving an insight into the flexibility of the protein [51]. Figure 4.4 depicts the RMSF of the residues for each system for the duration of the simulation. High fluctuations were observed at certain residues for each of the systems, with the free protein showing the greatest fluctuations during the simulation (1.61Å). All three ligand-bound systems showed C-α residue fluctuations at residues 72-79 and 409-411. The HMC-HO1α-ATPase system specifically showed flexibility at the "172-176" region, whereas the Ivermectin-ssRNA illustrated the lowest fluctuations of all four systems. This correlates with the RMSD stability of the systems, demonstrating the free protein to have highly unstable residues with large fluctuations compared to the ligand-bound systems. The RMSF of Ivermectin also correlated with the RMSD plot, illustrating a relatively stable system after ligand binding.

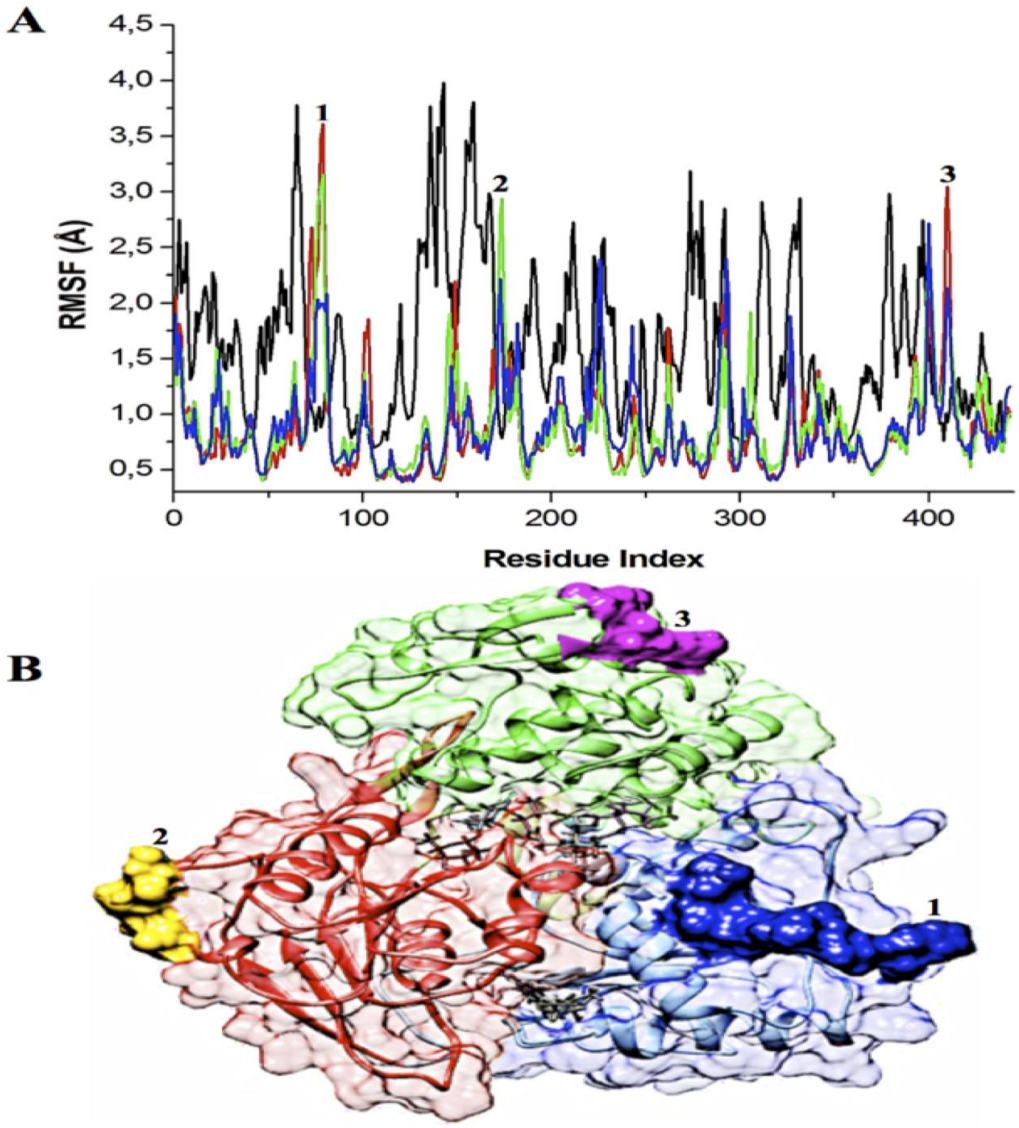


Figure 4.4: RMSF Plot for Lapachol, HMC-HO1α and Ivermectin systems. (A) Lapachol (0.88Å) showed a higher stability at the ATPase site compared to HMC-HO1α (0.90Å) while Ivermectin (0.85Å) showed the most favorable stability of all the systems, (B) NS3 Helicase residue flunctuations at regions "72-79" loop (Navy), 2- the "172-176" helix (Gold) and 3- the "409-411" loop (Magenta).

4.3.2.3 Radius of Gyration

To further validate the stability of the systems, the overall protein shape and folding was

measured by analyzing the radius of gyration (Rg) of the protein. This gave an insight into the distribution of C-α atoms within the protein [51,52]. The plots for all the systems are shown in Figure 4.5. From the graph, a difference can be seen in the compactness of the three systems from the beginning of the simulation. Ivermectin shows a lower average Rg (22.23Å) when compared to HMC-HO1α (22.30Å) and Lapachol (22.33Å), indicating that Ivermectin exhibits a very good structural stability at the ssRNA site when it binds to the ZIKV NS3 helicase. Also at the ATPase site, the result indicates that HMC-HO1α is more compact and therefore exhibit more stability than Lapachol. The Rg of the Apo protein correlates with the RMSD and RMSF results, showing a wide distribution of C-α atoms for the duration of the simulation, thus indicating unstable fluctuations of the protein's residues in the absence of a ligand.

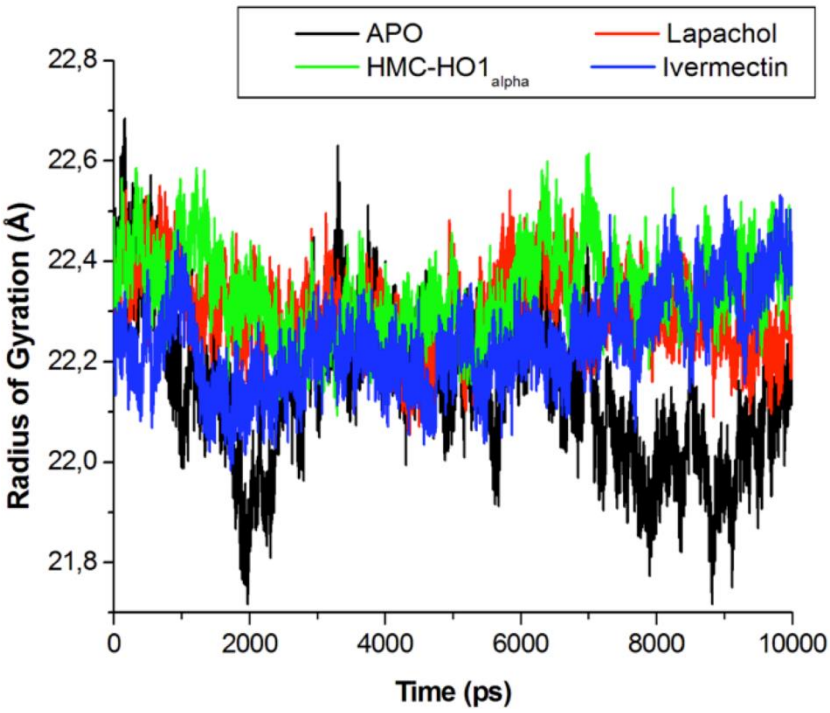


Figure 4.5: Radius of gyration plot for Lapachol, HMC-HO1α and Ivermectin systems when compared to the free protein.

4.3.2.4 Free Energy Calculations and Residue-Ligand Interaction Network

Studies have shown that free binding energies calculations are important parameters for the validation of ligand-protein binding [43]. Based on the Systems' stability, we can deduce that during the simulation, binding of the three best-docked molecules, being, Lapachol, HMC-HO1α and Ivermectin, stabilized the fluctuating free protein. This may be due to non-covalent interactions taking place between the ligands and the active site's residues. To estimate the binding affinities of each of the ligands to the protein, the binding free energies were calculated using the Molecular Mechanics/Generalized-Born Surface Area method (MM/GBSA) [54]. Table 4.2 summarizes the binding free energy of HMC-HO1α-ATPase and Lapachol-ATPase systems to be -42.81kcal/mol and -39.32kcal/mol respectively. The non-polar solvation (-103.51kcal/mol) contributed greatly towards the total binding free energy of HMC-HO1α-NS3 helicase system while other favorable binding contributions also came from intermolecular electrostatic interactions (-62.53kcal/mol) and van der Waals interactions (-40.98kcal/mol). Lapachol-NS3 helicase system had its greatest binding contribution from non-polar solvation energy (-65.96kcal/mol), followed by van der Waals interactions (-38.23kcal/mol) and then intermolecular electrostatic interactions (-27.73kcal/mol). A polar solvation of 60.69kcal/mol and 26.64kcal/mol for HMC-HO1α-NS3 helicase system and Lapachol-NS3 helicase system respectively were also observed. This indicates that HMC-HO1α has a preferable binding energy than Lapachol at the ATPase site. Ivermectin had a relatively higher binding energy (-84.56kcal/mol) at the ssRNA site with the greater energy contribution from the non-polar solvation (-136.32kcal/mol) and van der Waals interactions (-104.36kcal/mol).

The active site residues of proteins are important for the protein's functionality; therefore, it is important to understand the interactions of these potential inhibitors with the amino acids

residues in the protein [52]. To gain more insight into the contribution of each residue towards the binding of the ligand, per residue interaction energy decomposition analysis was carried out on the three systems.

Table 4. 2: Binding free energy analysis (kcal/mol) for inhibitor- NS3 helicase complexes.

Energy Components (kcal/mol)					
Compound	**ΔE_{vdW}**	**ΔE_{elec}**	**ΔG_{gas}**	**ΔG_{solv}**	**ΔG_{bind}**
Ivermectin	-104.36±3.95	-32.26±7.87	-136.32±11.07	52.07± 5.34	-84.56±7.77
HMC-HO1α	-40.98±3.40	-62.53±10.17	-103.51±8.78	60.69± 7.39	-42.81±4.16
Lapachol	-38.23±2.99	-27.73±6.87	-65.96±5.71	26.64± 4.12	-39.32±3.52

Lapachol-ATPase System

At the ATPase binding site, Lapachol illustrated a favorable energy contribution with residues Glu112 (-3.05kcal/mol), sharing the highest total energy, while other contributions came from residues Leu20 (-0.24kcal/mol), Gly23 (-0.29kcal/mol), Ala24 (-0.25kcal/mol), Glu57 (-0.30kcal/mol), Ala43 (-0.32kcal/mol), Asn243 (-0.89kcal/mol) and Arg285 (-0.8kcal/mol). However, Lys26 (0.45kcal/mol) and Arg288 (0.8kcal/mol) showed unfavorable energy contributions (Figure 4.6).

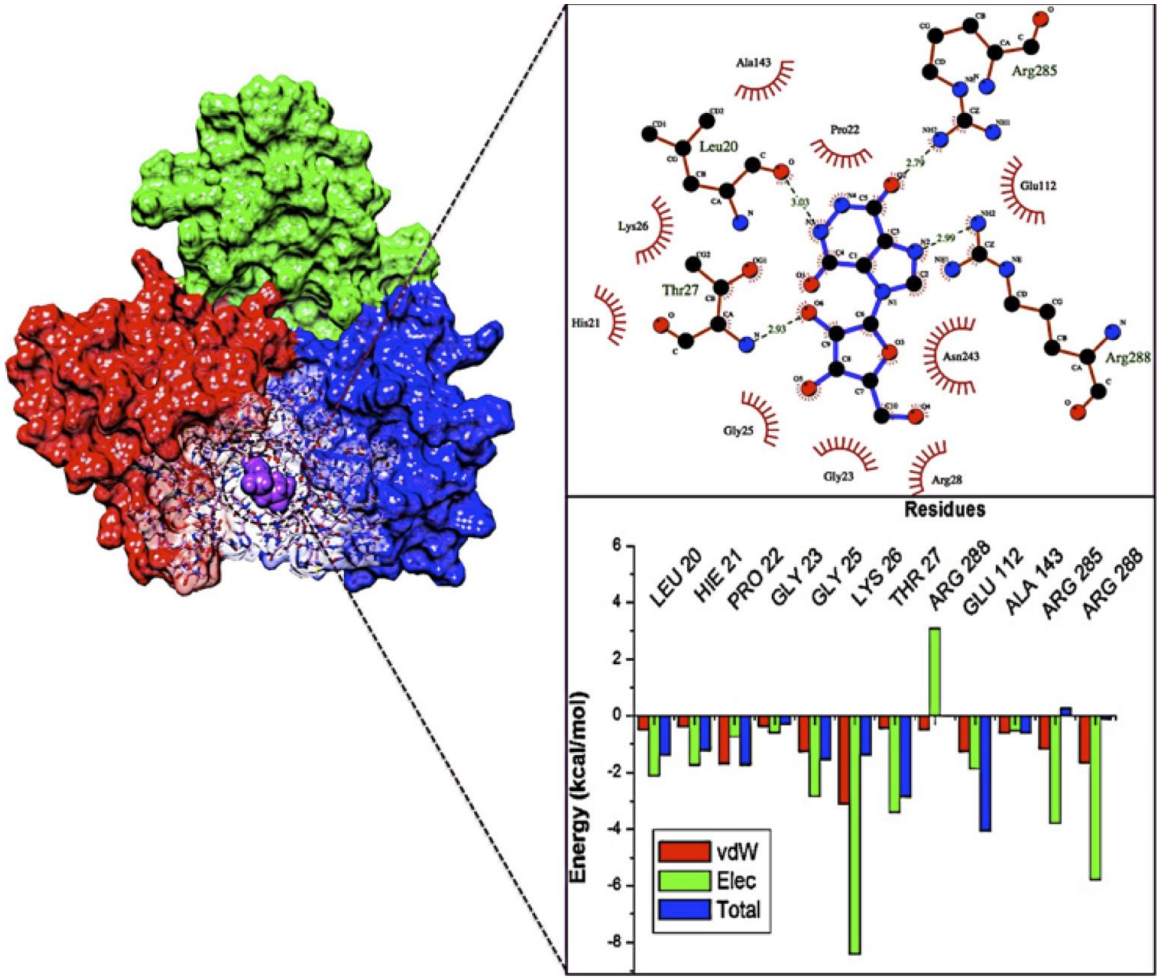


Figure 4.6: Free energy decomposition and ligand-residue interaction network at the ATPase site of the Lapachol- NS3 Helicase system.

HMC-HO1α-ATPase System

As evident from Figure 4.6 and 4.7, HMC-HO1α and Lapachol interact with the ATPase active sites residues by forming a hydrogen bond with residue Arg285 and hydrophobic interactions with residues His21, Gly23, Glu112 and Ala143. In addition, HMC-HO1α exhibited hydrophobic interactions with residues Pro22, Lys26, Gly25, Arg28, and Asn243. Subsequent to HMC-HO1α binding at the ATPase site, significant energy contributions came from residues Leu20 (-1.412kcal/mol), His21 (-1.24kcal/mol), Pro22 (-1.75kcal/mol), Gly25 (-1.55kcal/mol), Lys26 (-1.40kcal/mol), and Thr27 (-2.87kcal/mol), with the highest

contribution coming from Glu112 (-4.08kcal/mol), while the unfavorable energy contributions came from residue Arg285 (0.26kcal/mol).

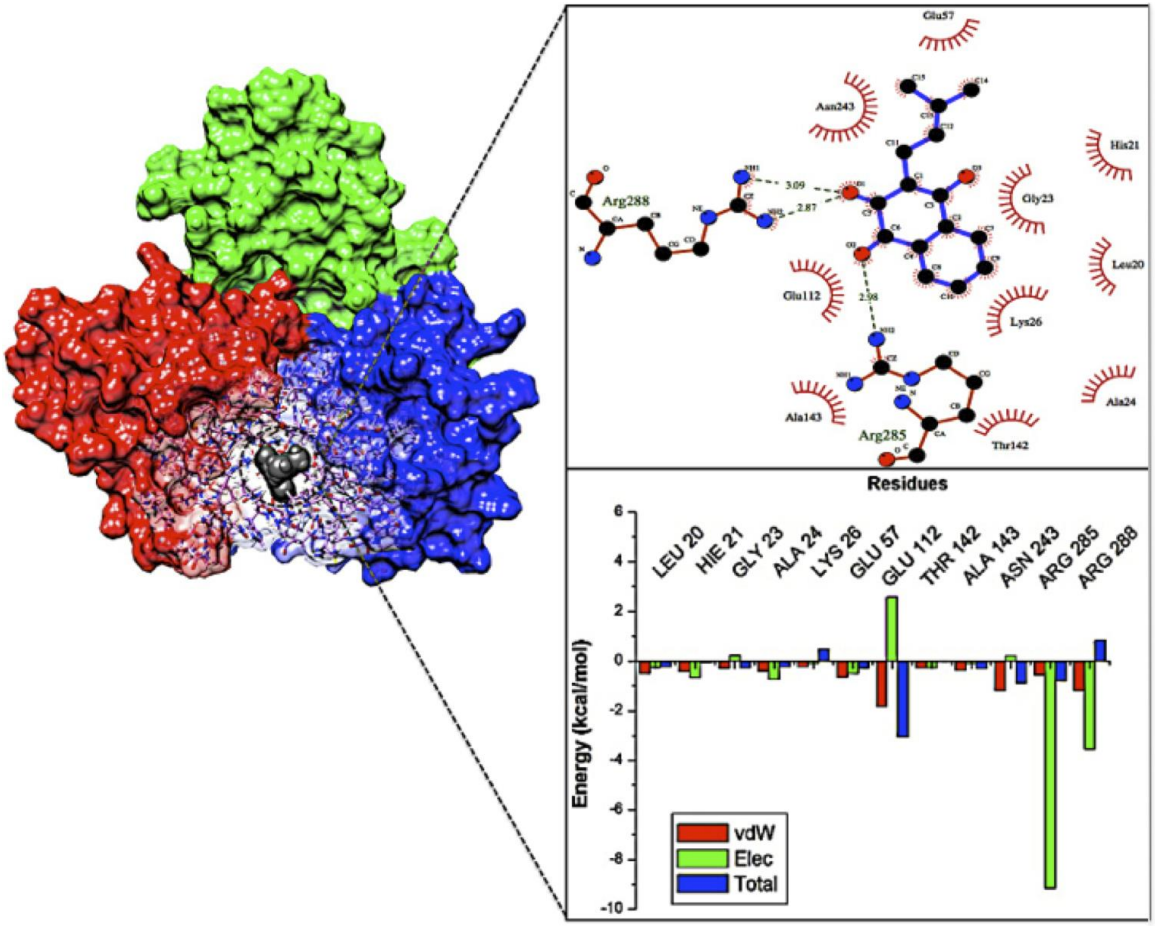


Figure 4.7: HMC-HO1α docked into the ATPase site of Zika NS3 helicase, illustrating ligand-residue interactions and active-site residue energy contributions.

The ligand-residue interaction network elucidates on the binding interactions between Ivermectin and the ssRNA active site residues (Figure 4.8), forming a hydrogen bond with Arg214 and hydrophobic interactions with residues Ser94, Leu319, Asp117, Val369, Pro368, Thr235, Met240, Glu218, Lys215, Ala90 and Met362. The plot also reveals that residue Arg214 (-5.84kcal/mol) had the highest total energy contribution to the binding of Ivermectin to the NS3 helicase protein at the ssRNA site. Other favorable energy contributions came

from residues Ala90 (-1.48kcal/mol), Ser119 (-2.464kcal/mol), Thr235 (-1.52kcal/mol), Asp236 (-1.65kcal/mol), Leu319 (-1.11kcal/mol), Met 362 (-1.87kcal/mol), Pro368 (-1.14kcal/mol) and Val369 (-1.32kcal/mol) while the unfavorable energy contribution came from Asp117 (3.0kcal/mol).

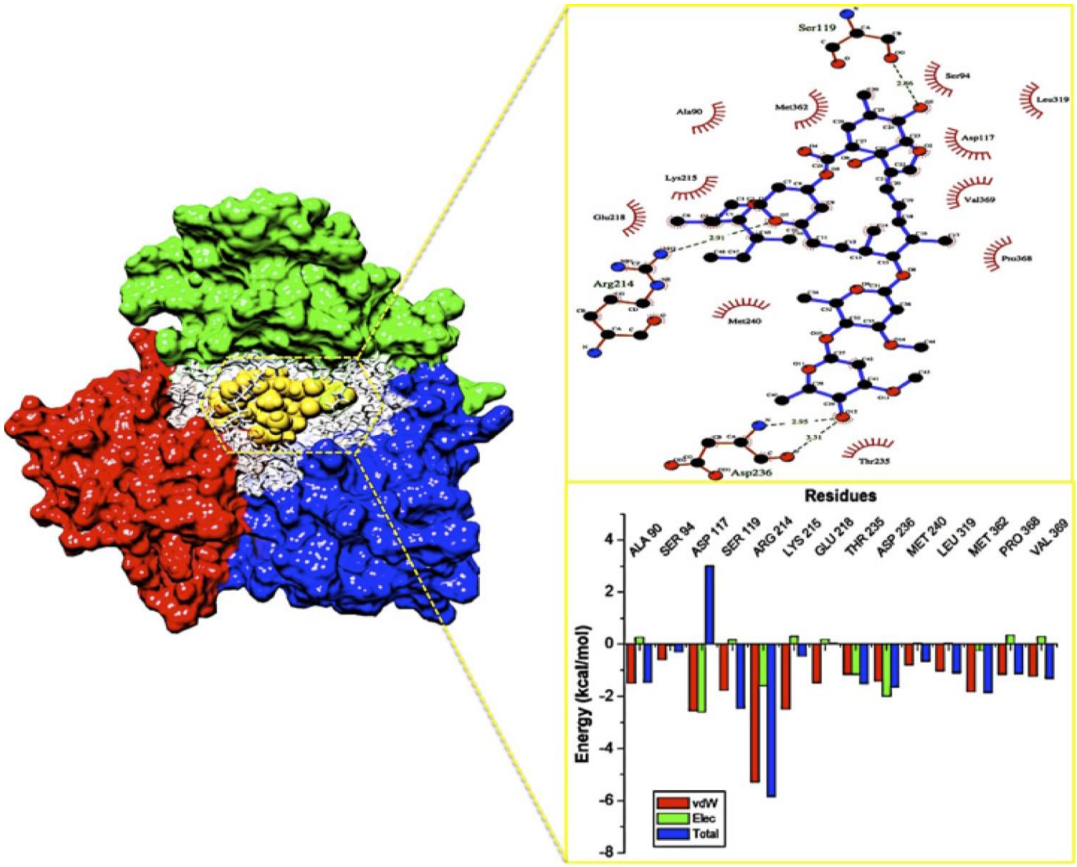


Figure 4.8: Free energy decomposition and ligand-residue interaction network at the ssRNA site of the Ivermectin- NS3 Helicase system.

4.4. Conclusions

In this study, we report the binding analysis of three potential inhibitors of Zika NS3 helicase at the ATPase site (Lapachol and HMC-HO1α) and ssRNA site (Ivermectin). Results showed that the binding of Ivermectin to ssRNA site and Lapachol and HMC-HO1α to the ATPase site allows for conformational rigidity of the Zika NS3 helicase, thus stabilizing residue fluctuations. The interactions between the active site residues and ligands allowed for key structural flexibility at two loop regions of the NS3 helicase, thus allowing for protein stability and a possible structural mechanism of action for competitive inhibition of natural substrates.

This study aims to contribute toward the repurposing of potent *flavivirus* inhibitors against the devastating ZIKV epidemic. This strategy overcomes the concept of "shooting the dark" with experimental screening as the compounds utilized in the study have already been synthesized, thus reducing the drug discovery time-line. These potential inhibitors have been pre-clinically tested against other arboviruses and have proven to be effective [6,12–14]. Drugs such as Ivermectin have multiple functions, including anti-parasitic and more recently anti-viral properties [7,11,15]. Lapachol and HMC-HO1α have been shown to have potent effects as *flavivirus* inhibitors including Dengue and Yellow fever virus[13].

The findings of this study provide fundamental insights toward the structural dynamics of the two active site regions on the NS3 Helicase and the ligand-receptor interaction network. The pharmacophoric characteristics found in Lapachol, HMC-HO1α and Ivermectin may be utilized in the design of a potent hybrid drug that is able to show efficient inhibition of a multitude of diseases including the detrimental co-infection of ZIKV, Dengue and Chikungunya.

4.5. Future Perspective

To our knowledge, this is the first account of detailed computational investigations aimed to provide an insight into the binding features of Lapachol, HMC-HO1α and Ivermectin to ZIKV NS3 helicase. Based on the structural dynamics of the two active site regions on the NS3 Helicase and the ligand-receptor interaction network, it may be noted that the chemical characteristics found in these flavivirus inhibitors play a fundamental role in releasing a potent multi-purpose inhibitor against arboviruses. This will allow for pregnant females in endemic areas to take the drug as a precautionary measure against arboviruses such as Dengue and ZIKV. By having a lower toxicity and higher efficiency, a multi-purpose drug will be safe to consume by pregnant females and may diminish the risk of drug resistance due to the multiple diseases it is effective against. Distribution on a global scale and at lower cost compared to a vaccine that may need optimal storage conditions.

4.6 Summary point

- The Zika virus (ZIKV) has emerged as a pathogen of major health concern. The rapid spread of the virus has led to an uproar in the medical domain as scientists frantically race to develop effective vaccines and small molecules to inhibit the virus.
- Targeted therapy against specific viral maturation proteins is necessary in halting the replication of the virus in the human host, thus decreasing host–host transmission.
- ZIKV co-infection with dengue and chikungunya has caused major concern in the scientific community due to the detrimental fetal neurological effects it leads to.
- Scientists have now turned to ‘repurposing’ *Flavivirus* US FDA approved drugs rather than the lengthy process of designing and synthesizing new drugs.
- Multipurpose drugs such as Ivermectin, Lapachol and HMC-HO1α have proven to have favorable inhibitory effects on the ZIKV NS3 helicase enzyme.

- Free energy calculations and residue–ligand interaction networks have shown Lapachol, HMC-HO1α and Ivermectin to trigger key structural flexibility at two-loop regions of the NS3 helicase.
- These flexible loops present a possible structural mechanism for competitive inhibition of natural substrates.
- The pharmacophoric characteristics found in Lapachol, HMC-HO1α and Ivermectin may be utilized in the design of a potent hybrid drug that is able to show efficient inhibition of a multitude of diseases including the detrimental co-infection of ZIKV, dengue and chikungunya.

Acknowledgements

The authors acknowledge the College of Health Sciences, UKZN, and the National Research Foundation for their financial support and the Center for High Performance Computing (http://www.chpc.ac.za) for their computational resources.

References

1. Sreedharan.J. The Zika Virus: A new Threat from mosquito. *Scientia*. 11, 9–17 (2016).

2. Haddow AD, Schuh AJ, Yasuda CY, *et al.* Genetic characterization of zika virus strains: Geographic expansion of the asian lineage. *PLoS Negl. Trop. Dis.* 6(2), 1-7 (2012).

3. Martínez de Salazar P, Suy A, Sánchez-Montalvá A, Rodó C, Salvador F, Molina I. Zika fever. *Enferm. Infecc. Microbiol. Clin.* 34(4), 247–252 (2016).

4. Barrows NJ, Campos RK, Powell ST *et al.* A Screen of FDA-Approved Drugs for inhibition of Zia Virus infection. Cel Host & microbe 20(2), 259-270 (2016).

5. Wikan N, Smith DR. Zika virus: History of a newly emerging arbovirus. *Lancet Infect. Dis.* 16(7), 119–126 (2016).

6. Chan JFW, Choi GKY, Yip CCY, Cheng VCC, Yuen KY. Zika fever and congenital Zika syndrome: An unexpected emerging arboviral disease. *J. Infect.* 72(5), 507–524 (2016).

7. Heang V, Yasuda CY, Sovann L, *et al.* Zika virus infection, Cambodia, 2010. *Emerg. Infect. Dis.* 18(2), 349–351 (2012).

8. Kutsuna S, Kato Y, Takasaki T, *et al.* Two cases of zika fever imported from french polynesia to Japan, December to January 2013. *Eurosurveillance*. 19(4), 1–4 (2014).

9. Hamel R, Dejarnac O, Wichit S, *et al.* Biology of Zika Virus Infection in Human Skin Cells. *J. Virol.* 89(17), 8880–96 (2015).

10. Foy BD, Kobylinski KC, Foy JLC, *et al.* Probable Non-Vector-borne Transmission of Zika Virus, Colorado, USA. *Emerg. Infect. Dis.* 17(5), 880–882 (2011).

11. Hayes EB. Zika virus outside Africa. *Emerg. Infect. Dis.* 15(9), 1347–1350 (2009).

12. Musso D, Roche C, Robin E, Nhan T, Teissier A, Cao-Lormeau VM. Potential sexual transmission of zika virus. *Emerg. Infect. Dis.* 21(2), 359–361 (2015).

13. Besnard M, Lastère S, Teissier A, Cao-Lormeau VM, Musso D. Evidence of perinatal transmission of zika virus, French Polynesia, December 2013 and February 2014. *Eurosurveillance*. 19(13), 8–11 (2014).

14. Mahfuz M, Khan A, Mahmud H Al, *et al.* In Silico Modeling and Immunoinformatics Probing Disclose the Epitope Based PeptideVaccine Against Zika Virus Envelope Glycoprotein. *Indian J. Pharm. Biol. Res.* 2(4), 44–57 (2014).

15. Plourde AR, Bloch EM. A Literature Review of Zika Virus. *Emerg. Infect. Dis.* 22(7), 1–15 (2016).

16. Charrel RN, Leparc-Goffart I, Pas S, de Lamballerie X KM& RC. State of knowledge on Zika virus for an adequate laboratory response. *Publ. Bull. World Heal. Organ. Type Res. Emergencies Artic.*, 1–29 (2016).

17. Cao-Lormeau V-M, Blake A, Mons S, *et al.* Guillain-Barré Syndrome outbreak associated with Zika virus infection in French Polynesia: a case-control study. *Lancet*. 387(10027), 1531–1539 (2016).

18. Baharuddin A, Hassan AA, Sheng GC, *et al.* Current approaches in antiviral drug discovery against the Flaviviridae family. *Curr. Pharm. Des.* 20(21), 3428–3444 (2014).

19. Malet H, Massé N, Selisko B, *et al.* The flavivirus polymerase as a target for drug discovery. *Antiviral Res.* 80(1), 23–35 (2008). ``

20. Tian H, Ji X, Yang X, *et al.* The crystal structure of Zika virus helicase: basis for antiviral drug design. *Protein Cell*. 7(6), 450–454 (2016).

21. Jain R, Coloma J, Garcia-Sastre A, Aggarwal AK. Structure of the NS3 helicase from Zika virus. *Nat. Struct. Mol. Biol.* 2(July), 1–4 (2016).

22. Incicco JJ, Gebhard LG, González-Lebrero RM, Gamarnik A V., Kaufman SB. Steady-State NTPase Activity of Dengue Virus NS3: Number of Catalytic Sites, Nucleotide Specificity and Activation by ssRNA. *PLoS One*. 8(3), 1-8 (2013).

23. Amber C. Amber 2015 Reference Manual.

24. Miao Y, Feixas F, Eun C, McCammon JA. Accelerated molecular dynamics simulations of protein folding. *J. Comput. Chem.* 36(20), 1536–49 (2015).

25. Lindert S, Bucher D, Eastman P, Pande V, Mccammon JA. Accelerated Molecular Dynamics Simulations with the AMOEBA Polarizable Force Field on Graphics Processing Units. *J. Chem. Theory Comput.* 9, 1–2 (2013).

26. Pettersen EF, Goddard TD, Huang CC, *et al.* UCSF Chimera - A visualization system for exploratory research and analysis. *J. Comput. Chem.* 25(13), 1605–1612 (2004).

27. Windows MM V, X MOS. Molegro Molecular Viewer User Manual. (2011).

28. Da Costa ECB, Amorim R, Da Silva FC, *et al.* Synthetic 1,4-pyran naphthoquinones are potent inhibitors of Dengue virus replication. *PLoS One*. 8(12), 1–11 (2013).

29. Borowski P, Lang M, Haag A, *et al.* Characterization of Imidazo [4 , 5- d] Pyridazine Nucleosides as Modulators of Unwinding Reaction Mediated by West Nile Virus Nucleoside Triphosphatase / Helicase : Evidence for Activity on the Level of Substrate and / or Enzyme. 46(5), 1231–1239 (2002).

30. Briguglio I, Piras S, Corona P, Carta A. Inhibition of RNA Helicases of ssRNA+ Virus Belonging to Flaviviridae, Coronaviridae and Picornaviridae Families. *Int. J. Med. Chem.*

2011, 1–22 (2011).

31. Mastrangelo E, Pezzullo M, De burghgraeve T, *et al.* Ivermectin is a potent inhibitor of flavivirus replication specifically targeting NS3 helicase activity: New prospects for an old drug. *J. Antimicrob. Chemother.* 67(8), 1884–1894 (2012).

32. Kim S, Thiessen PA, Bolton EE, *et al.* PubChem substance and compound databases. *Nucleic Acids Res.* 44(1), 1202–1213 (2016).

33. Steffen C, Thomas K, Huniar U, Hellweg A, Rubner O, Schroer A. TmoleX--A Graphical User Interface for TURBOMOLE. *J. Comput. Chem.* 31, 2967–2970 (2010).

34. Case DA, Cheatham TE, Darden T, *et al.* The Amber bimolecular simulation programs. J. Comput. Chem. 26(16), 1668–1688 (2005).

35. Cornell WD, Cieplak P, Bayly CI, *et al.* A second-generation force field for the simulation of proteins, nucleic acids, and organic molecules. *J. Am. Chem. Soc.* 117(19), 5179–5197 (1995).

36. Wang J, Wolf RM, Caldwell JW, Kollman PA, Case DA. Development and testing of a general Amber force field. *J. Comput. Chem.* 25 `(9), 1157–1174 (2004).

37. Goetz AW, Williamson MJ, Xu D, Poole D, Grand SL, Walker RC. Routine microsecond molecular dynamics simulations with amber - part i: Generalized born. *J. Chem. Theory Comput.* 8, 1542–1555. (2012).

38. Jorgensen WL, Chandrasekhar J, Madura JD, Impey RW, Klein ML. Comparison of simple potential functions for simulating liquid water. *J. Chem. Phys.* 79(2), 926 (1983).

39. Seifert E. Origin Pro 9.1: Scientific data analysis and graphing software - Software review. *J. Chem. Inf. Model.* 54(5), 1552 (2014).

40. Greenidge PA, Kramer C, Mozziconacci JC, Wolf RM. MM/GBSA binding energy prediction on the PDB bind data set: Successes, failures, and directions for further improvement. *J. Chem. Inf. Model.* 53(1), 201–209 (2013).

41. Godschalk F, Genheden S, Söderhjelm P, Ryde U. Comparison of MM/GBSA calculations based on explicit and implicit solvent simulations. *Phys. Chem. Chem. Phys.* 15(20), 7731–9 (2013).

42. Tsui V, Case DA. Theory and applications of the Generalized Born solvation model in macromolecular simulations. *Biopolymers*. 56(4), 275–291 (2000).

43. Kollman PA, Massova I, Reyes C, *et al.* Calculating Structures and Free Energies of Complex Molecules: Combining Molecular Mechanics and Continuum Models. *Acc. Chem. Res.* 33(12), 889–897 (2000).

44. Meng X-Y, Zhang H-X, Mezei M, Cui M. Molecular docking: a powerful approach for structure-based drug discovery. *Curr. Comput. Aided. Drug Des.* 7(2), 146–57 (2011).

45. Ferreira LG, Dos Santos RN, Oliva G, Andricopulo AD. Molecular docking and structure-based drug design strategies. *Molecules*, 20(7), 13384-13421 (2015).

46. De Oliveira AS, Da Silva ML, Flávia A, *et al.* NS3 and NS5 Proteins: Important Targets for Anti-Dengue Drug Design. *J. Braz. Chem. Soc*. 25(10), 1759–1769 (2014).

47. Jr. FRS. Molecular Dynamics Simulations of Protein Dynamics and their relevance to drug discovery. *Pharmacol.* 10(6), 738–744 (2010).

48. Arodola OA, Soliman MES. Could the FDA-approved anti-HIV PR inhibitors be promising anticancer agents? An answer from enhanced docking approach and molecular dynamics analyses. *Drug Des. Devel. Ther.* 9, 6055–6065 (2015).

49. Sing A SM. Understanding the cross-resistance of oseltramivir to HINI and H5NI influenza A neuraminidase mutation using multidimensional computational analyses. , 1–18 (2015).

50. Appiah-Kubi P, Soliman MES. Dual anti-inflammatory and selective inhibition mechanism of leukotriene A4 hydrolase/aminopeptidase: insights from comparative molecular dynamics and binding free energy analyses. *J. Biomol. Struct. Dyn.* 1102, 1–16 (2016).

51. Karubiu W, Bhakat S, Soliman MES. Compensatory role of double mutation N348I/M184V on nevirapine binding landscape: insight from molecular dynamics simulation. *Protein J.* 33(5), 432–446 (2014).

52. Kumalo HM, Soliman ME. Per-Residue Energy Footprints-Based Pharmacophore Modeling as an Enhanced In Silico Approach in Drug Discovery: A Case Study on the Identification of Novel Secretase1 (BACE1) Inhibitors as Anti-Alzheimer Agents. *Cell. Mol. Bioeng.* 9(1), 175–189 (2016).

CHAPTER 5
Delving into Zika Virus Structural Dynamics: A Closer look at NS3 Helicase Loop flexibility and its Role in Drug Discovery

Pritika Ramharack[A], Sofiat Oguntade[A] Mahmoud E. S. Soliman[A*]

[A]Molecular Modeling and Drug Design Research Group, School of Health Sciences, University of KwaZulu-Natal, Westville Campus, Durban 4001, South Africa

*Corresponding Author: Mahmoud E.S. Soliman email: soliman@ukzn.ac.za

1. Dean and Head of School of Health Sciences, Full Professor: Pharmaceutical Sciences, University of KwaZulu-Natal, Westville Campus, Durban 4001, South Africa.
2. Department of Pharmaceutical Organic Chemistry, Faculty of Pharmacy, Zagazig University, Zagazig, Egypt.

3. College of Pharmacy and Pharmaceutical Sciences, Florida Agricultural and Mechanical University, FAMU, Tallahassee, Florida 32307, USA.

Email: soliman@ukzn.ac.za

Telephone: +27 (0) 31 260 8048, Fax: +27 (0) 31 260 7872

Abstract

The Zika virus has emerged as a pathogen of major health concern. The rapid spread of the virus has led to an uproar in the medical domain as scientists frantically race to develop effective vaccines and small molecules to inhibit the virus. In the past year, there has been a flood of Zika knowledge published including its characteristics, transmission routes and its role in disease conditions such as Microcephaly and Gullian-Barŕe syndrome. Targeted therapy against specific viral maturation proteins is necessary in halting the replication of the virus in the human host, thus decreasing host-host transmission. This prompted us to investigate the structural properties of the Zika NS3 Helicase when bound to ATP-competitive inhibitor, NITD008. In this study, comparative molecular dynamic simulations were employed for Apo and bound protein to

demonstrate the molecular mechanism of the Helicase. Results clearly revealed that NITD008-binding caused significant residue fluctuations at the P-loop compared to the rigid nature of the Apo conformation. The NITD008-helicase complex also revealed residues 339-348 to transition from a $_3$10-Helix to a stable α-helix. These protein fluctuations were verified by investigation of dynamic cross correlation and principal component analysis. The fundamental dynamic analyses presented in this report is crucial in understanding Zika NS3 Helicase function, thereby giving insights toward an inhibition mechanism. The information reported on the binding mode at the ATPase active site may also assist in designing of effective inhibitors against this detrimental viral target.

Keywords:

Zika NS3 Helicase, ATPase active site, Molecular dynamic simulations, P-Loop Flexibility.

5.1 Introduction

The re-emerging Zika virus (ZIKV) has evolved into a catastrophic epidemic over the past year, with scientific community announcing that the long-term effects associated with the virus will have to be dealt with in the decades to follow [1]. The virus was declared an international public health emergency by the World Health Organization [2], based on growing evidence of the virus being linked with congenital neurological diseases such as Guillain-Barŕe, cranial nerve dysfunction and Microcephaly [3,4]. The ZIKV made its devastating re-appearance in Brazil and has now spread on a global scale, with an estimated 75 countries with reported mosquito-borne ZIKV transmission as of December 2016 [5].

Zika virus is an arthropod-borne *flavivirus* initially discovered in the Zika forest area of Uganda in 1947 [6]. Of the *flavivirus* genera, ZIKV is most closely related to the Spodweni virus from the Spodweni group; however, ZIKV shares structural similarities with other *flaviviruses*, including Dengue virus and West Nile virus [7]. The ZIKV genome is made up of structural proteins, being the capsid, precursor membrane and envelope form the viral particle and seven non-structural proteins, being NS1, NS2A, NS2B, NS3, NS4A, NS4B and NS5, which participate in the replication of the RNA genome, virion assembly and invasion of the innate immune system [8–10]. In our previous review, we explicated on the key viral target proteins, including the multifunctional viral replication NS3 helicase protein[11]. The ZIKV helicase comes from the superfamily helicases, SF2 [12], with the inhibition of either one of the binding sites, the RNA-binding groove or the ATP-binding site (Figure 1), leading to the virus becoming incapable of sufficient maturation and replication. The structural characteristics of the ZIKV NS3 protein includes three domains: domain I (residues 182-327), domain II (residues 328-480) and domain III (residues 481-617), as well as a P-Loop (residues 196-203) which is located at the ATP-binding site of domain I [12,13].

The co-crystallization of $MnATP^{-2}$ and RNA with ZIKV helicase, reported by Tian *et al* (2016) and Cao *et al* (2016), have paved the way to understanding the mechanism by which these substrates bind to the enzyme, initiating viral RNA replication [14,15]. Despite the flood of integrated knowledge on ZIKV over the past year, the molecular and structural mechanism for helicase inhibition is yet to be established [12].

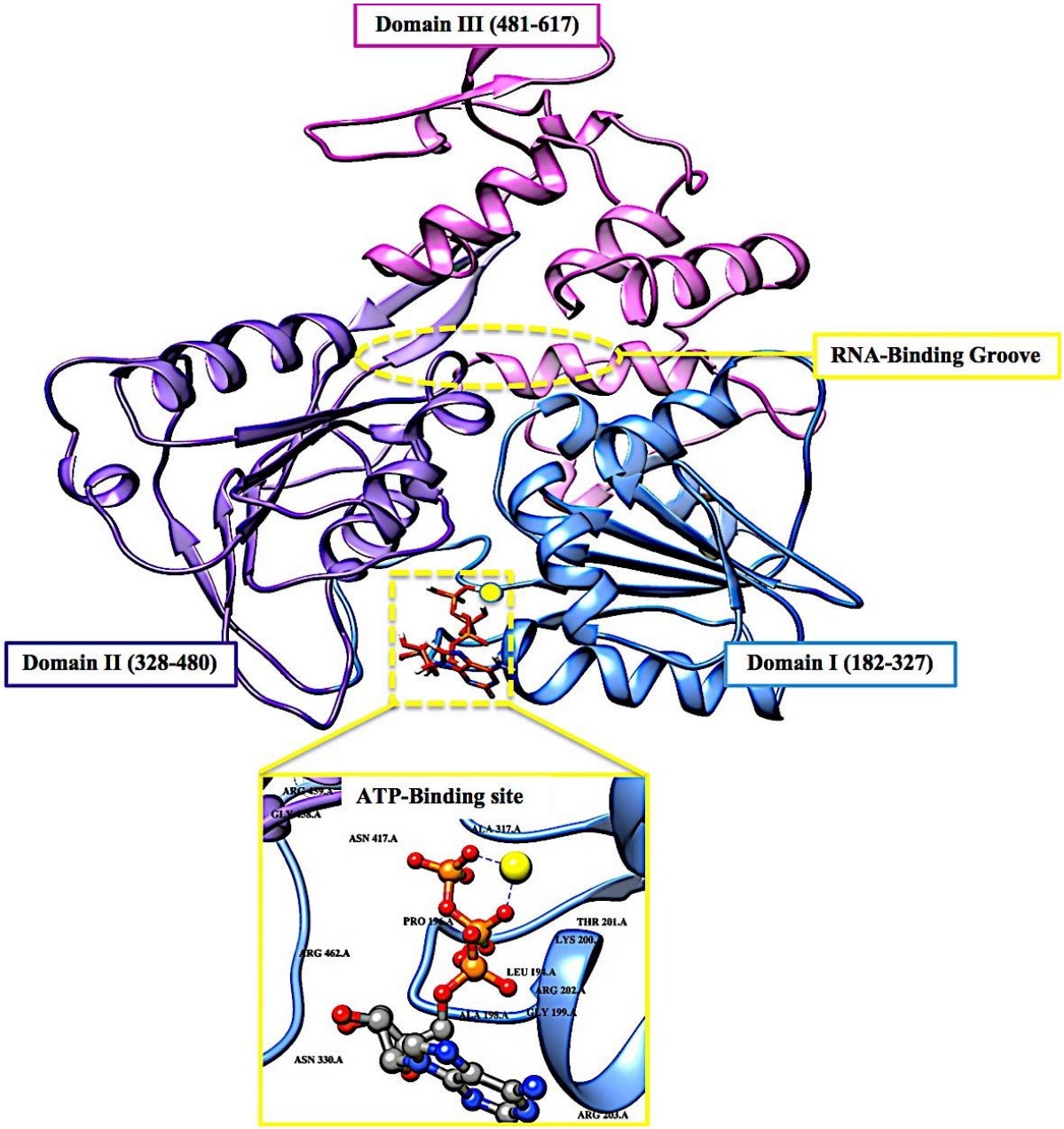


Figure 5.1: Cartoon and surface representation of the three domains of the ZIKV helicase and the two active-binding regions (yellow) that form profound hydrophobic cavities in the electrostatic surface area, allowing ATP and ssRNA to bind.

Another battle being fought by researchers is the discovery of new modes of transmission of the virus, from initially being transmitted from vector to host, to now being inclusive of blood

transfers from host to host as well as secondary sexual transmission [16–18]. This has allowed for rapid diffusion of the virus between continents. In the plethora of strategic characteristics of the virus, its ability to target neuronal cells has been one of the most problematic tasks that pharmaceutical chemists have had to overcome [19–24]. The design of ZIKV inhibitors will not only need to be target-specific, effective and have minimal toxicity, but it will also have to pass through the blood-brain-barrier [25].

Although there are currently vaccine clinical trials under way, there are still no FDA approved small molecule inhibitors against the virus [26–30]. This may be due to many reasons including time-consuming experimental testing of large libraries of compounds or minimal literature available on the functionality of the virus in host cells. These possible barriers have prompted us to utilize computational drug design tools, such as molecular dynamic (MD) simulations to explore the conformational landscape of this biological system's ATP-binding region. The crystallographic structures have revealed evidence of residue mobility, including the rotation of motor domains, however, the precise structural characteristics of the helicase upon small molecule binding, is yet to be determined [12,31–36].

In this study we investigate the conformational changes at the ATP-binding region after a 130ns MD simulation of the free enzyme state as well as a NITD008-bound complex [37]. This study will be critical in understanding how the ZIKV NS3 helicase functions structurally, thus aiding in the design of effective, target-specific inhibitors.

5.2 Computational Methods

5.2.1 System Preparation

The ZIKV NS3 helicase in complex with ATP and a magnesium ion (PDB code: 5GJC)[14] was obtained from RSCB Protein Data Bank [38]. The 3-D structure of the experimental ZIKV inhibitor, NITD008, was obtained from PubChem[39] and prepared on Molegro Molecular Viewer (MMV) [40]. In the ZIKV crystal structure of the ATP-bound helicase, residues A247-S253 were absent, thus the free enzyme (PDB code: 5JMT)[13] was utilized in the docking of NITD008. Deng *et al* (2016) reported conclusive *in vivo* evidence of the inhibition of ZIKV by NITD008. The compound is classified as an adenosine nucleoside analog that competitively inhibits ATP, thus sharing an active site [37].

5.2.2 Molecular Docking

Molecular docking is a conventional method in computational chemistry which is utilized in the prediction optimized geometric conformations of a ligand within an appropriate binding site [41]. The Molecular docking software utilized included Raccoon [42], Autodock Graphical user interface supplied by MGL tools [43] and AutoDockVina [44] with default docking parameters. Prior to docking, Gasteiger charges were added to NITD008 and the non-polar hydrogen atoms were merged to carbon atoms. Water molecules were removed, and polar hydrogen was added to the crystal structure of the NS3 helicase. NITD008 was then docked into the ATPase binding pocket of the NS3 helicase (by defining the grid box with spacing of 1 Å and size of 32 × 26 × 30 pointing in x, y and z directions respectively). Due to the lack of experimental data describing ZIKV approved inhibitors, validation of molecular docking based on the lowest energy pose

becomes unreliable [45]. To overcome any experimental bias, the five best conformational poses, based on binding affinities (kcal/mol), were subjected to MD simulations.

5.2.3 Molecular Dynamic (MD) Simulations

Molecular dynamic (MD) simulations provide a robust tool to explore the physical movements of atoms and molecules, thus providing insights on the dynamical evolution of biological systems. The MD simulation was performed using the GPU version of the PMEMD engine provided with the AMBER package, FF14SB variant of the AMBER force field [46] was used to describe the protein.

ANTECHAMBER was used to generate atomic partial charges for the ligand by utilizing the Restrained Electrostatic Potential (RESP) and the General Amber Force Field (GAFF) procedures. The Leap module of AMBER 14 allowed for addition of hydrogen atoms, as well as Na^+ and Cl^- counter ions for neutralization to both the Apo- and Bound system.

Both systems were then suspended implicitly within an orthorhombic box of TIP3P water molecules such that all atoms were within 10Å of any box edge.

An initial minimization of 2000 steps were carried out with an applied restraint potential of 500 kcal/mol $Å^2$ for both solutes, were performed for 1000 steps using a steepest descent method followed by a 1000 step of conjugate gradients. An additional full minimization of 1000 steps were further carried out by conjugate gradient algorithm without restrain.

A gradual heating MD simulation from 0K to 300K was executed for 50ps, such that the system maintained a fixed number of atoms and fixed volume, i.e., a canonical ensemble (NVT). The solutes within the system are imposed with a potential harmonic restraint of 10kcal/mol $Å^2$ and

collision frequency of $1.0ps^{-1}$. Following heating, an equilibration estimating 500ps of each system was conducted; the operating temperature was kept constant at 300K. Additional features such as a few atoms and pressure where also kept constant mimicking an isobaric-isothermal ensemble (NPT). The systems pressure was maintained at 1 bar using the Berendsen barostat.

The total time for the MD simulation conducted was 130ns. In each simulation, the SHAKE algorithm was employed to constrict the bonds of hydrogen atoms. The step size of each simulation was 2fs and an SPFP precision model was used. The simulations coincided with isobaric-isothermal ensemble (NPT), with randomized seeding, constant pressure of 1 bar maintained by the Berendsen barostat, a pressure-coupling constant of 2ps, a temperature of 300K and Langevin thermostat with collision frequency of $1.0ps^{-2}$.

5.2.4 Post-Dynamic Analysis

The coordinates of the free enzyme and NITD008 complex were each saved every 1ps and the trajectories were analyzed every 1ps using PTRAJ, followed by analysis of RMSD, RMSF and Radius of Gyration using the CPPTRAJ module employed in AMBER 14 suit.

5.2.4.1 Binding Free Energy Calculations

Binding free energy calculations is an important end point method that may elucidate on the mechanism of binding between a ligand and enzyme, including both enthalpic and enthropic contributions [47]. To estimate the binding affinity of the docked systems, the free binding energy was calculated using the Molecular Mechanics/GB Surface Area method (MM/GBSA) [48]. Binding free energy was averaged over 15000 snapshots extracted from the 130ns trajectory. The

free binding energy (ΔG) computed by this method for each molecular species (complex, ligand and receptor) can be represented as:

$$\Delta G_{bind} = G_{complex} - G_{receptor} - G_{ligand} \quad (1)$$

$$\Delta G_{bind} = E_{gas} + G_{sol} - TS \quad (2)$$

$$E_{gas} = E_{int} + E_{vdw} + E_{ele} \quad (3)$$

$$G_{sol} = G_{GB} + G_{SA} \quad (4)$$

$$G_{SA} = \gamma SASA \quad (5)$$

The term E_{gas} denotes the gas-phase energy, which consist of the internal energy E_{int}; Coulomb energy E_{ele} and the van der Waals energies E_{vdw}. The E_{gas} was directly estimated from the FF14SB force field terms. Solvation free energy, G_{sol}, was estimated from the energy contribution from the polar states, G_{GB} and non-polar states, G. The non-polar solvation energy, SA. G_{SA}, was determined from the solvent accessible surface area (SASA), using a water probe radius of 1.4 Å, whereas the polar solvation, G_{GB}, contribution was estimated by solving the GB equation. S and T denote the total entropy of the solute and temperature respectively.

To obtain the contribution of each residue to the total binding free energy profile at the ATPase site, per-residue free energy decomposition was carried out at the atomic level for imperative residues using the MM/GBSA method in AMBER 14 suit.

The system displaying the most favorable binding interaction and energy contributions were subjected to further analysis.

5.2.4.2 Dynamic Cross-correlation Analysis (DCC)

Dynamic cross correlation is a widespread method in MD simulations in which the correlation coefficients of motions between atoms of a protein may be quantified [49]. The dynamic cross correlation between the residue-based fluctuations during simulation was calculated using the CPPTRAJ module incorporated in AMBER 14. The formula used to describe dynamic cross correlation is given below:

$$Cy = \frac{< \Delta \mathrm{ri}.\, \Delta \mathrm{rj} >}{\left(< \Delta r_i^2 >< \Delta r_j^2 >\right)^{\frac{1}{2}}}$$

The cross-correlation coefficient (Cij) varies within a range of −1 to +1 of which the upper and lower limits correspond to a fully correlated and anti-correlated motion during the simulation process. Where, i and j stands for i^{th} and j^{th} residue respectively and Δr_i or Δr_j represents displacement vectors correspond to i^{th} and j^{th} residue respectively. The generated dynamic cross correlation matrix was constructed in Origin software.

5.2.4.3 *Principal Component Analysis (PCA)*

Principal component analysis (PCA) is a covariance-matrix-based mathematical technique that is able to demonstrate atomic displacement and the loop dynamics of a protein [50]. Prior to processing the MD trajectories for PCA, the trajectories of the free enzyme (APO) and the NITD008-bound complex (Complex) were stripped of solvent and ions using the PTRAJ module in AMBER 14. The stripped trajectories were then aligned against their corresponding fully minimized structures. PCA was performed for C-α atoms on 900 snapshots each. Using in-house scripts, the first two principal components were calculated and the covariance matrices were generated. The first two principal components (PC1 and PC2) generated from each trajectory

were averaged for both the free-enzyme and NITD008-complex. The first two principal components (PC1 and PC2) were computed and a 2 X 2 covariance matrix were generated using Cartesian coordinates of Cα atoms. PC1 and PC2 correspond to first two eigenvectors of covariant matrices. Origin software [51] was used to construct PCA plots.

5.3 Results and Discussion

5.3.1 NITD008-NS3 Helicase Complex

5.3.1.1 Binding of NIT21D008 with ZIKV Helicase

Research into ZIKV inhibitors has been minimal before 2016. However, NITD008, a *flavivirus* adenosine analogue was evidenced, both *in vitro* and *in vivo*, to inhibit ZIKV replication. The adenosine nucleoside analogue competes with natural ATP substrates, which are incorporated into the growing RNA chain. By this substitution, NITD008 is incorporated into the RNA chain, thus terminating the RNA elongation and inhibiting ZIKV maturation [37].

Molecular docking has become a major computational tool that is used to predict the orientation of a ligand at a binding site on the receptor. Results from docking often display multiple predicted orientations of the ligand within the active pocket [52].

In this study, NITD008 docked at the ATP-binding site in 6 favorable conformations (Figures S2-S6), with the highest binding-affinity being -8.2 kcal/mol. Scoring functions often attempt to reproduce experimental binding affinities, but most software do not always yield the best prediction. Validation of the docked structure with experimentally known drugs was also not possible due to the lack of FDA inhibitors against ZIKV [45,53,54].

In an attempt to improve the binding affinity prediction of NITD008, all 6 predicted complexes were subjected to 130ns molecular dynamic simulations, allowing for more realistic receptor flexibility in an implicit solvent. Each complex was then analyzed using the accurate, MM/GBSA, free binding energy calculation to determine the most favorable pose of NITD008 at the NS3 ATPase active site [47,55–57].

5.3.1.2 Free Energy calculations

The total binding free energy for each of the 6 poses of the NITD008- NS3 helicase complex were calculated using the MM/GBSA approach to better understand the various energy contributions within the binding pocket and assess which binding pose would show the most favorable intermolecular interactions at the helicase active site. Per residue decomposition analysis was also assessed and the residue-ligand interaction network of each pose were depicted as "ligplot" maps (Figures S2-S6). Of the six systems, the pose with the highest docking score, -8.2 kcal/mol, showed the most favorable free binding energy (-55.90 kcal/mol) supported the molecular docking score, indicating a favorable structural pose of NITD008 at the binding site.

The thermodynamic energy contribution of NITD008 to the total binding free energy of the complex surmounts to the stability of NITD008 in the ATP binding pocket and thus the stability of the complex during the simulation. Table 4.1 summarizes the free binding energy of the system considering the energies of the NS3 helicase and NITD008.

Table 5.1: Summary of free binding Energy contributions to the NITD008-NS3 Helicase system.

	Energy Components (kcal/mol)				
	Δ E_{vdW}	**ΔE_{elec}**	**ΔG_{gas}**	**ΔG_{solv}**	**ΔG_{bind}**
ZIKV HELICASE	-3429.35 ± 30.09	-28758.51 ± 159.37	-32187.86 ± 155.05	-5121.93 ± 115.09	-37309.79 ± 71.27
NITD008	-4.69 ± 0.85	18.12 ± 5.27	13.43 ± 5.28	-221.12 ± 3.35	-207.68 ± 3.72
COMPLEX	-37.71 ± 4.12	-382.94 ± 28.72	-420.64 ± 28.59	364.75 ± 22.80	-55.90 ± 7.71

Figure 5.2 represents the residue interaction plot of NITD008 within the active site. The active site residues Gly199, Lys200 and Glu286 formed stable hydrogen bonds with highly electronegative oxygen atoms of NITD008. The residues pocketing NITD008 within the active site included Gly197, Ala198, Gly199, Lys200, Thr201, Arg202, Glu288, Gly415, Asn417 and Arg456.

It was also interesting to note that the most favorable NITD008-pose shared five active residues with the ATP-bound helicase reported by Tian *et al* (2016). The crystal structure of the ATP-bound helicase showed Lys200 to stabilize the triphosphate of the ATP [14]. The Lys200 of the NITD008-bound helicase showed a similar stabilizing hydrogen bond with the terminal hydroxyl group located on the ribose of NITD008.

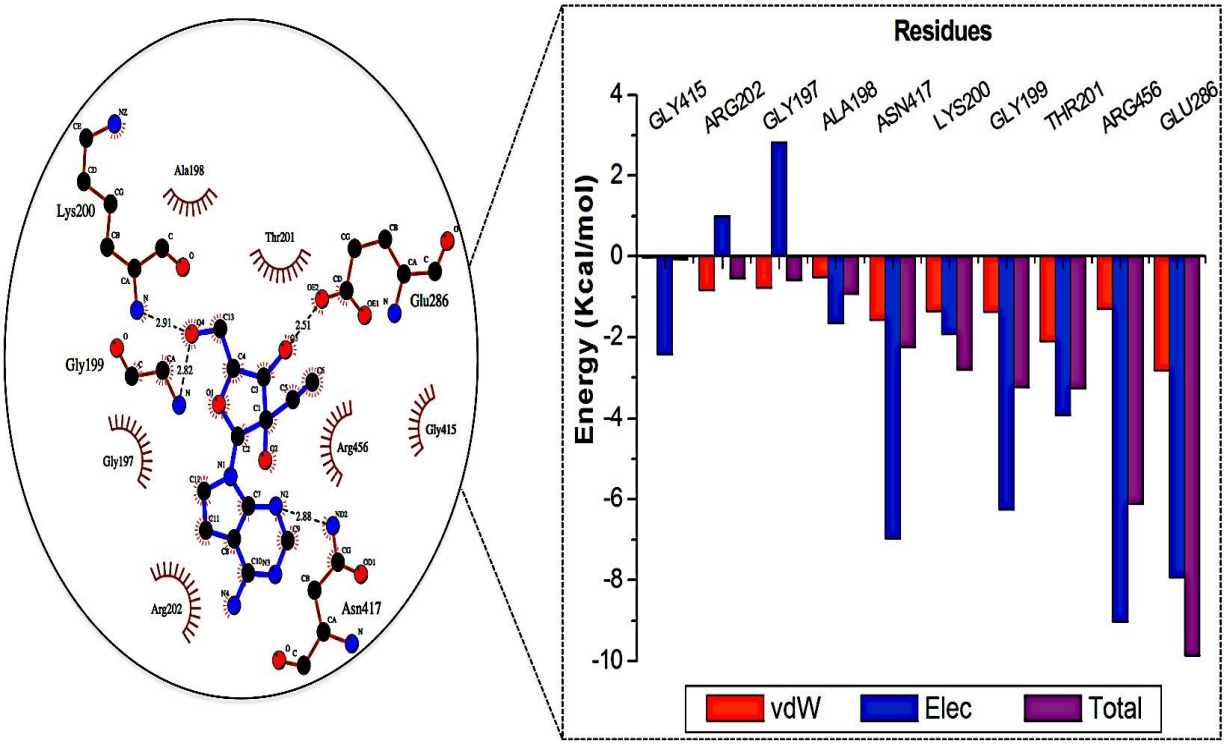


Figure 5.2: Energy contributions of the highest interacting residues at the ATPase active site. The residue ligand interaction network illustrates stabilizing hydrophobic interactions pocketing NITD008 at the active site. The highest energy contribution was a hydrogen bond interaction shared between Glu286 and the 3rd oxygen of the ribose component of NITD008.

Superimposition of NITD008-docked NS3 helicase with the ATP-NS3 helicase complex demonstrated both compounds to bind in a hydrophilic conformation despite the carbon and acetylene substitutions at N-7 of the purine and the 2' position of the ribose, respectively (Figure 5.3). The structural similarities between NITD008 and ATP, as well as the active site residue interactions and accurate free-binding energy prompted the further analysis of NITD008-complex.

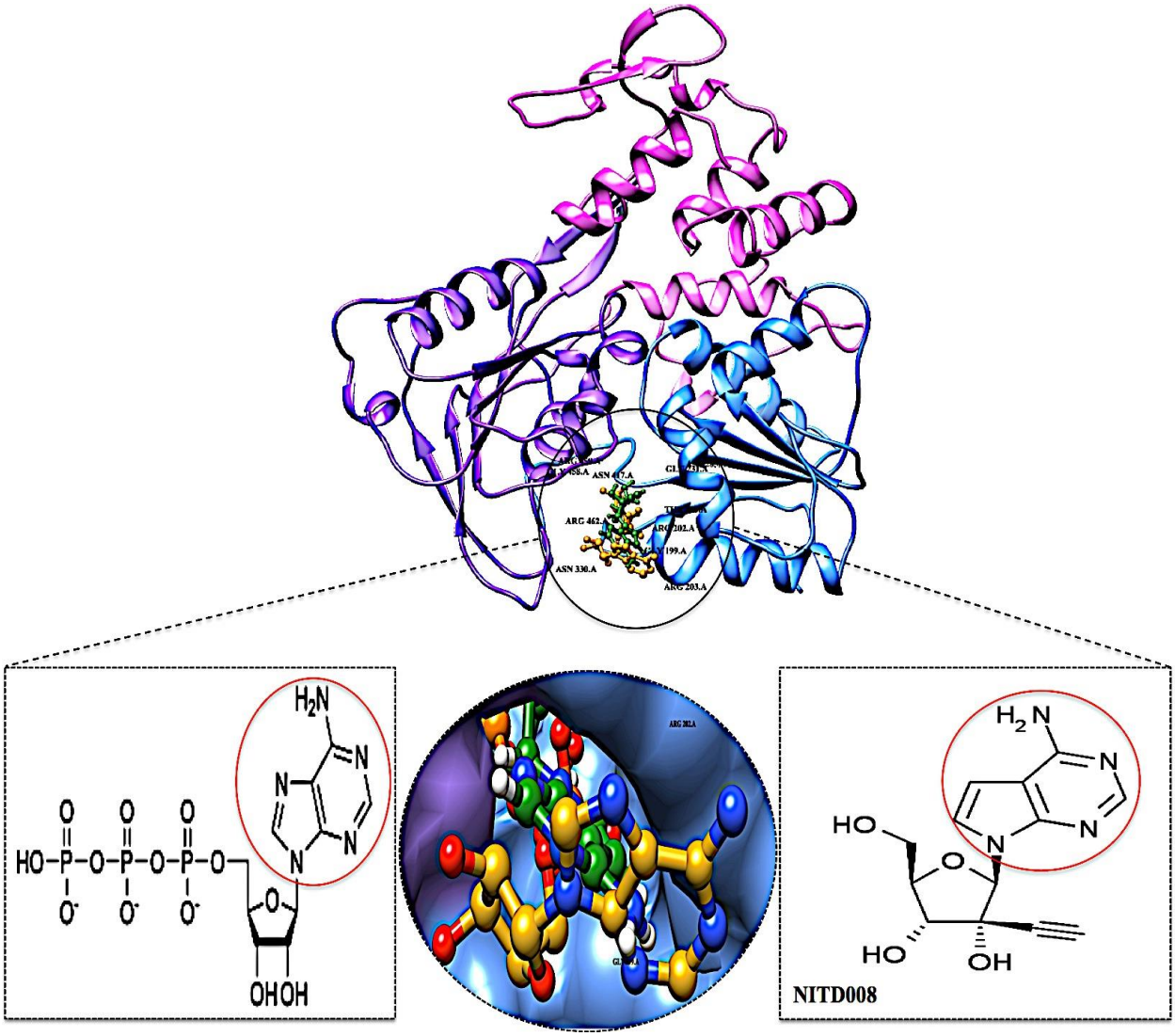


Figure 5.3: Superimposed conformation of structurally similar NITD008 and ATP docked at ATPase site of ZIKV NS3 Helicase.

5.3.2 Systems Stability

The length of a MD simulation is paramount when establishing insights into the structural dynamics of a biological system. With an extended simulation time, a system is able to reach convergence, thus becoming stable. To assure the equilibration of the simulation, the potential

energy and temperature where monitored (Figure S1). The average potential energy (-145774 kcal/mol) was measured at 300K, suggesting a stable conformation at this temperature.

5.3.2.1 Stability of NS3 Helicase APO and Bound System

The C-α backbone root mean square deviations (RMSD) were monitored throughout the 130ns MD simulation for both the free (APO) enzyme and the complex. Both systems reached convergence after 60ns (RMSD deviation < 2 Å). It can be noted that the C-α backbone atoms in both systems stabilized after a 40ns time, although, fluctuations in rigidity did increase during the 47-52ns period in the NITD008 complex (Figure 5.4). This could possibly be due to the occurrence of conformational changes because of the bond interactions taking place between NITD008 and the active site residues as seen in the Per-residue energy decomposition.

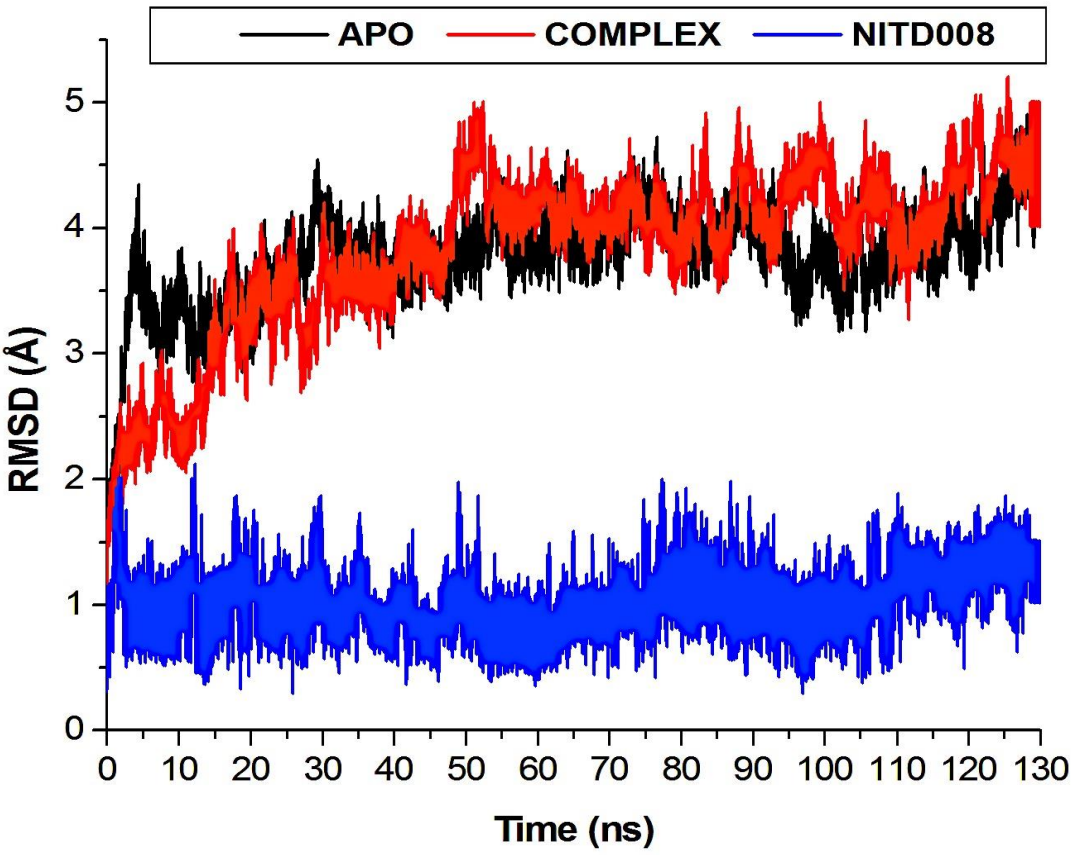


Figure 5.4: C-α backbone RMSD for NS3 Helicase APO enzyme and NITD-complex conformation. The average C-α RMSD was calculated to be 3.62 Å and 3.77 Å, respectively. Increased fluctuations occurred at 47-52ns in the NITD008-complex.

5.3.2.2 Conformational Fluctuations of the NS3 Helicase

To better understand the structural changes that may be occurring upon ligand binding, the root mean square fluctuation (RMSF) of the C-α atoms of each residue in the APO system and NITD008-complex were calculated. Figure 5.5 clearly demonstrates greater flexibility of residues of the NITD008-complex when compared to the APO enzyme. Fluctuations take place between residues 198-204, which form distinct hydrophobic and hydrogen bond interactions with NIT008D at the active site. This region, the P-Loop, is found in all *flavivirus* helicases and has been shown to have flexibility during binding of ATP [14]. The P-loop adopts structural modifications to accommodate the binding of ATP and Mn^{2+}. This flexibility extends greatly in comparison to the APO enzyme, thus verifying ZIKV P-loop flexibility upon ligand-binding. Other fluctuations occurred in domain II, and I around the ATP-active site, at residues 244-248 and 325-348.

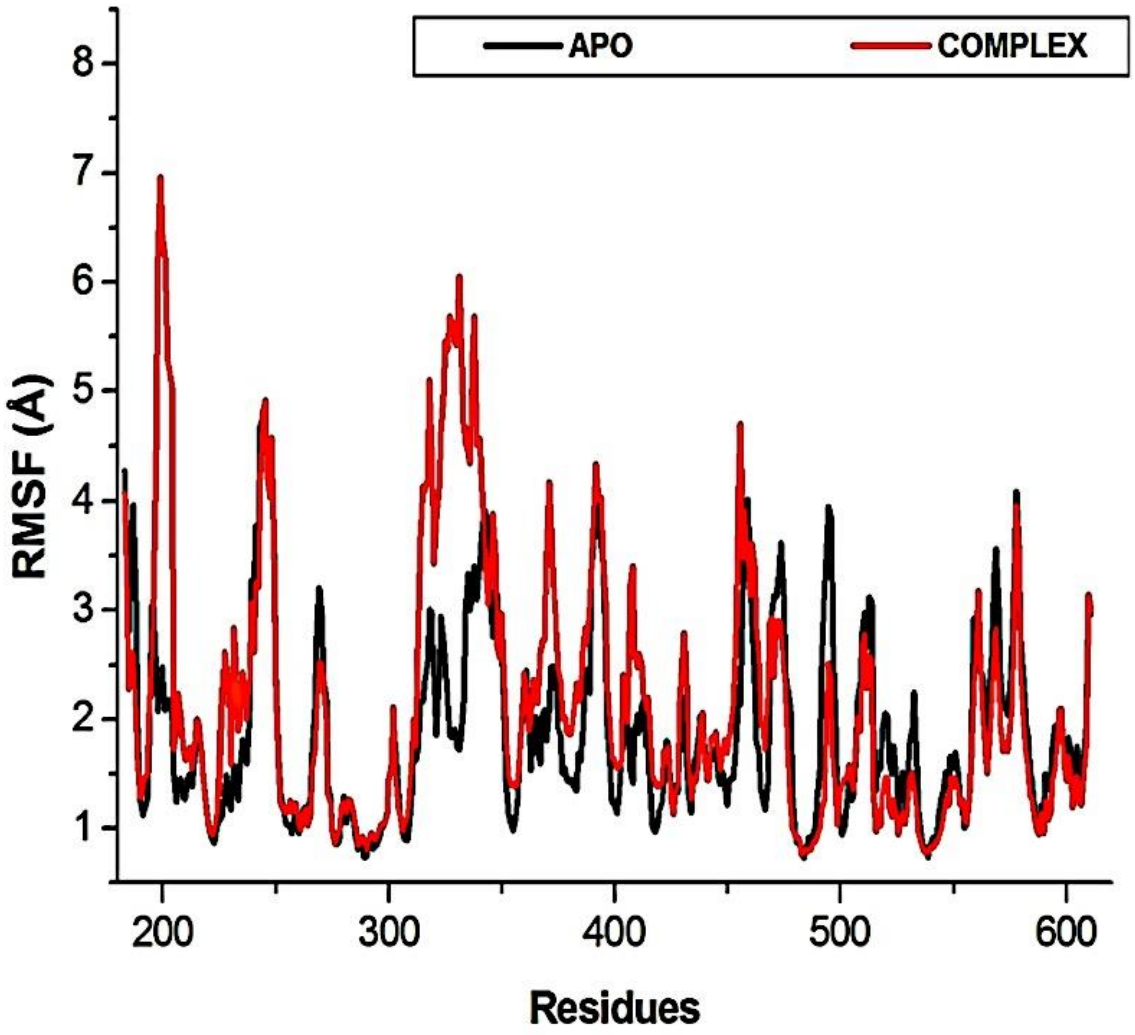


Figure 5.5: The RMSF of APO enzyme and NITD008-complex. The structural flexibility in domain I and II is highly attributed to the binding of NITD008 to the ATP-active site. This is substantiated by the average RMSF of the NITD008-complex (2.17 Å), which is significantly higher than that of the APO enzyme (1.90 Å).

5.3.2.3 Distribution of Atoms around the NS3 Helicase Backbone

The radius of gyration around the C-α atoms can measure the shape and folding of NS3 helicase before and after NITD008 binding. The radius of gyration measures the distribution of atoms

from the center of mass (COM), thus indicating how compact a system is. Both the APO (22.05 Å) and NITD008 (22.17 Å) showed very similar structural compactness, however, there was an atomic distribution in the NITD008-complex from 40-58ns (Figure 5.6). This correlates with the escalated instability of the complex at 47-52ns demonstrated in the RMSD plot.

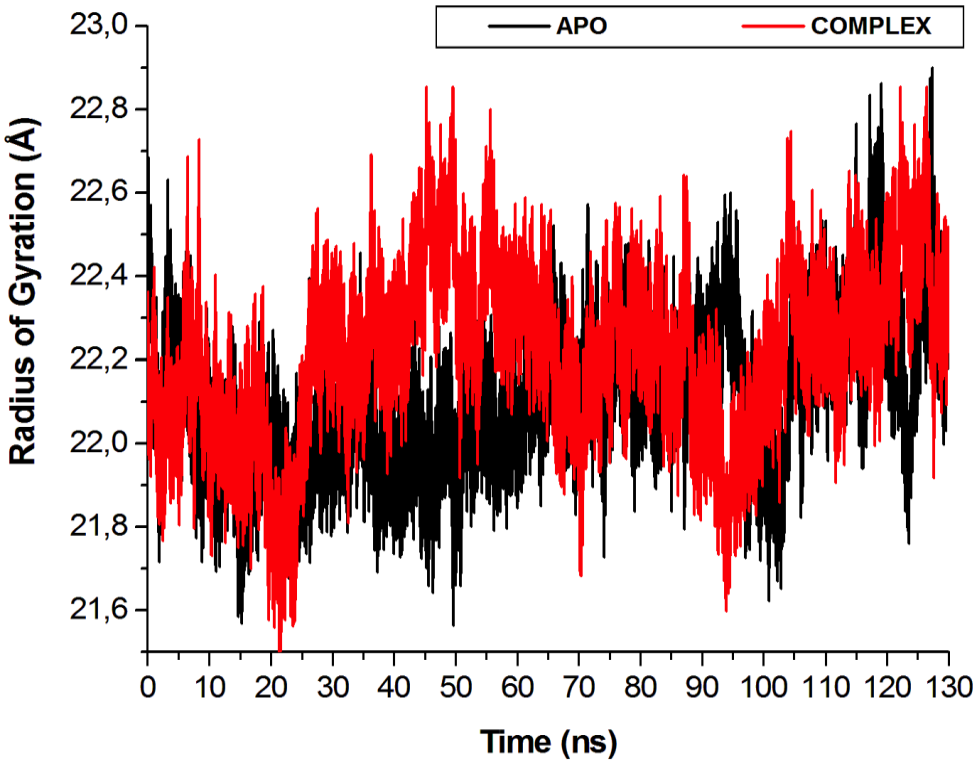


Figure 5.6: The radius of gyration (Rg) plot illustrating the difference in enzyme compactness of the NITD008-complex compared to the APO enzyme.

The flexibility calculated from the RMSD, RMSF and Rg encouraged us to explore the dynamic structural modifications of the NS3 Helicase after NITD008 binding.

5.3.3 Investigation of the Dynamic Structural features ATP-Active Binding Region

5.3.3.1 Loop Flexibility and Distance metrics

The ZIKV NS3 Helicase is made up of three known flexible loops that are common to all *flaviviruses*: The P-loop (residues 196-203), the RNA-binding loop (residues 244-255) and the β-hairpin loop (residues 431-444). These loops may vary in size depending on the type of virus; however, they all have the same fundamental structural flexibility. The RMSF plot demonstrated major fluctuations at the P-loop as well as the RNA-binding loop, the β-hairpin loop however, showed no significant conformational change compared to the APO enzyme. The plot also illustrated a flexible "325-338" region. Figure 5.7 depicts three snapshots of the APO enzyme and NITD008-complex, taking at different intervals along the trajectory. Clear conformational shifts are illustrated along the trajectory in both APO and bound systems.

To further investigate the conformational changes of the NS3 Helicase upon ligand binding, dynamic cross-correlation matrix (DCCM) analysis was performed at different conformational positions of the Cα backbone atoms of the free protein and ligand-bound complex. Highly correlated motions of residues are represented in the red to yellow regions, whereas, the negative/anti-correlated movements of residue Cα atoms are represented by blue-navy regions. It is evident from the correlation map that more globally correlated motion is observed in the case of the free protein, confirming conformational shifts after ligand binding. The latter residues of the NS3 Helicase, being residues 500-600, displayed anti-correlated movements in both the Apo and Bound complex, supporting the residue fluctuations in figure 5. Figure 7 also depicts anti-correlation motions at residues "340-390", which may be explained by the snapshots, in which, the flexible region in the NITD008-bound complex was converted from a $_{3}10$-helix to a α-helix.

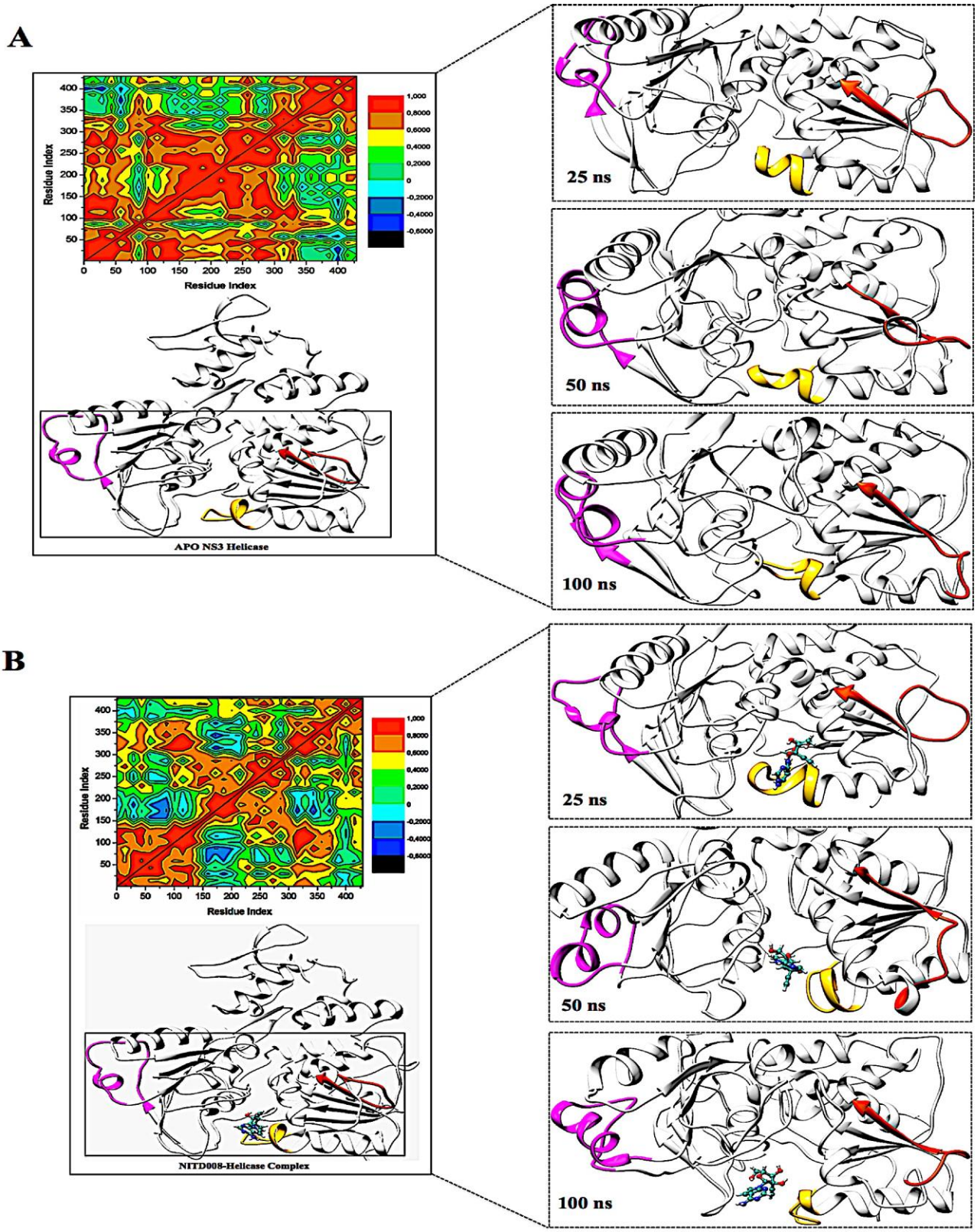

A
Residue Index
Residue Index
APO NS3 Helicase
25 ns
50 ns
100 ns
B
Residue Index
Residue Index
NITD008-Helicase Complex
25 ns
50 ns
100 ns

Figure 5.7: Structural Flexibility of the P-Loop (196-203), RNA-binding loop (244-255), and the 310 Helix (339-348) along the trajectory. The RNA-binding loop (orange) showed the loop shifting down in the Apo structure but an upward shift in the NITD008- Helicase complex. The P-Loop (Yellow) shifted away from the active site in the bound complex but closed in on the active site when no ligand was present. In the Apo structure, the helix-loop-helix stayed, with vibrational movement during the simulation, although, in the bound complex, the $_3$10 Helix (Pink) was modified into a α-helix due to ligand motional shifts further into the hydrophobic pocket.

The P-loop clearly illustrates that when NS3 Helicase is in its APO form and exposed to a 130ns simulation, the P-loop closes on the active site by uncoiling the α-helix at Arg203 to form part of the loop. The loop tip (Ala198) and the adjacent catalytic residue (Gly451) had an average distance of 9.71Å compared to the NITD008-complex distance of 12.75Å, whereby, as NITD008 becomes more stable at the active site and forms bond interactions, the P-loop is directed away from NITD008 and a larger catalytic space becomes available for the ligand as it forms stable hydrophobic interactions deeper within the hydrophobic pocket (Figure 5.8).

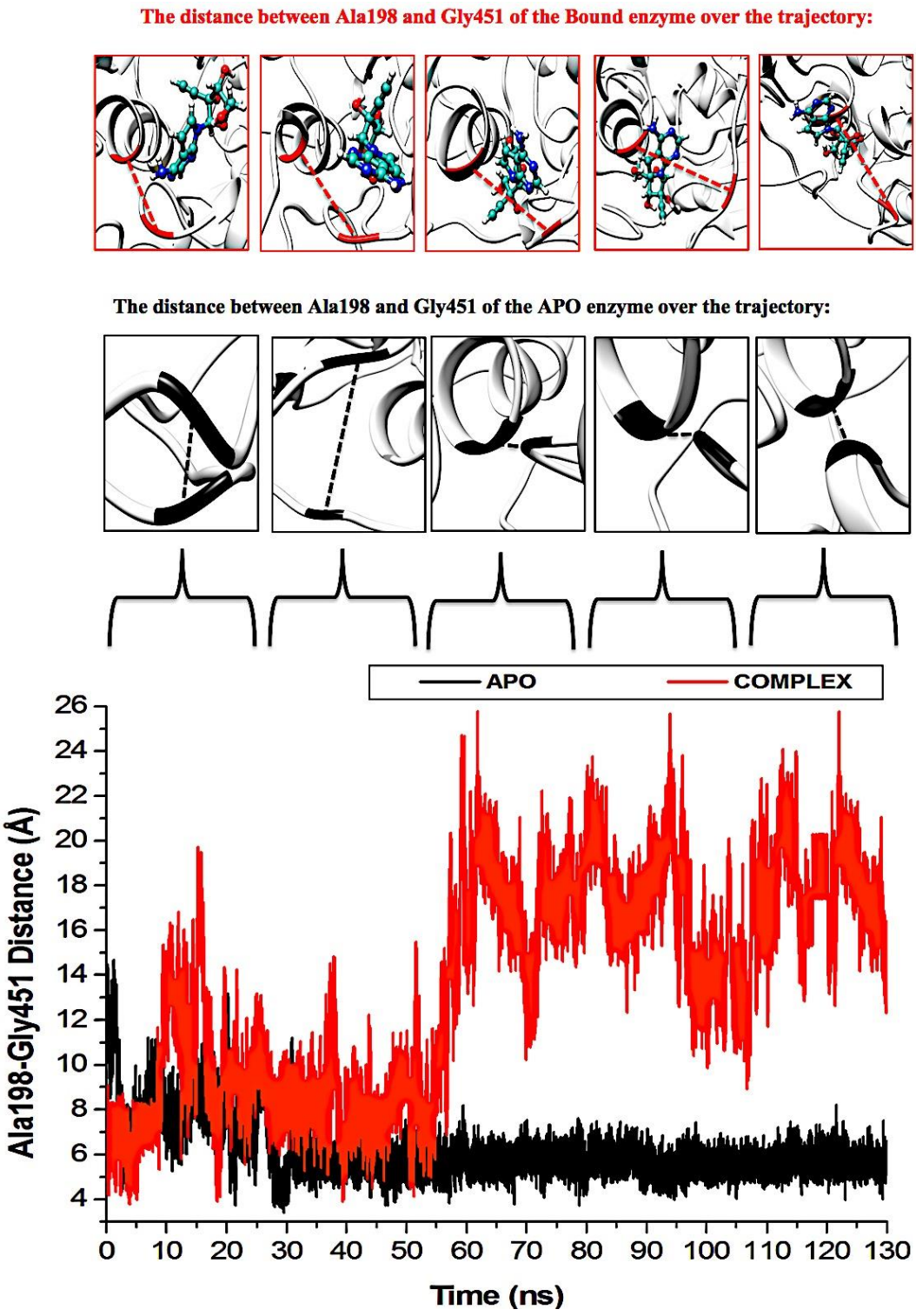
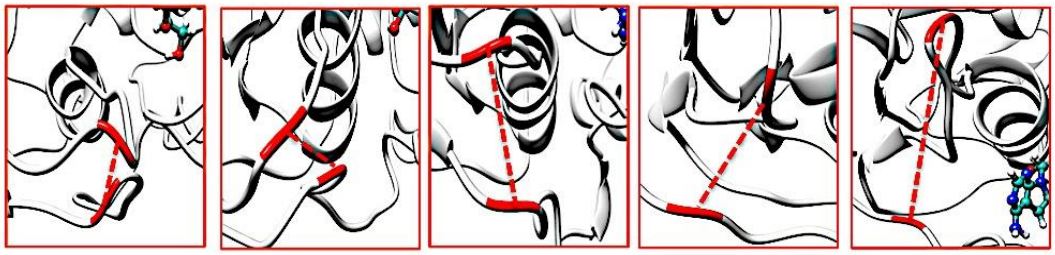


Figure 5.8: Residue fluctuations at the P-Loop region.The Apo enzyme illustrates closing of the loop at the active site due to a vacant hydrophobic pocket. Subsequent to ligand binding and the

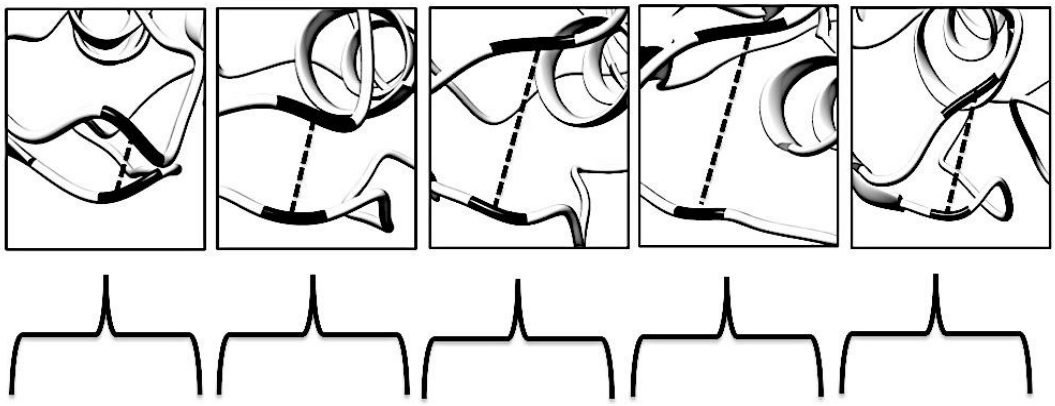
initiation of stabilizing hydrogen and hydrophobic bond interactions, the P-loop shifts down to accommodate the ligand, thus increasing the size of the hydrophobic pocket.

The “325-348” region demonstrates opposing conformational modifications between the APO and complex systems compared to that of the P-loop. The Distance between the two catalytic residues from the loop tips; residue Ser324 and residue Asn448, measured for the APO and NITD008-complex was 6.34Å and 8.34Å, respectively (Figure 5.9). The NITD008-complex had a greater distance between the residues due to the unraveling of 2 β-sheets found in domain II. This led to a “325-338” loop shift behind the active site and the “339-348” region being modified from a $_3$10 Helix to a α-Helix (Figure 7). The $_3$10 Helix conversion could be due to many reasons including changes in pH, interactions with other proteins and in this case, ligand binding. The ligand-protein interactions lead to distances between nitrogen and oxygen atoms from the protein backbone to fluctuate and as NITD008 moved further into the hydrophobic pocket, these fluctuations and hydrogen bond conversions caused the $_3$10 helix to convert to an α-helix. These changes are important in illustrating the conformational fluctuations upon ligand binding.

The distance between Ser324 and Asn458 of the Bound enzyme over the trajectory:



The distance between Ser324 and Asn458 of the APO enzyme over the trajectory:



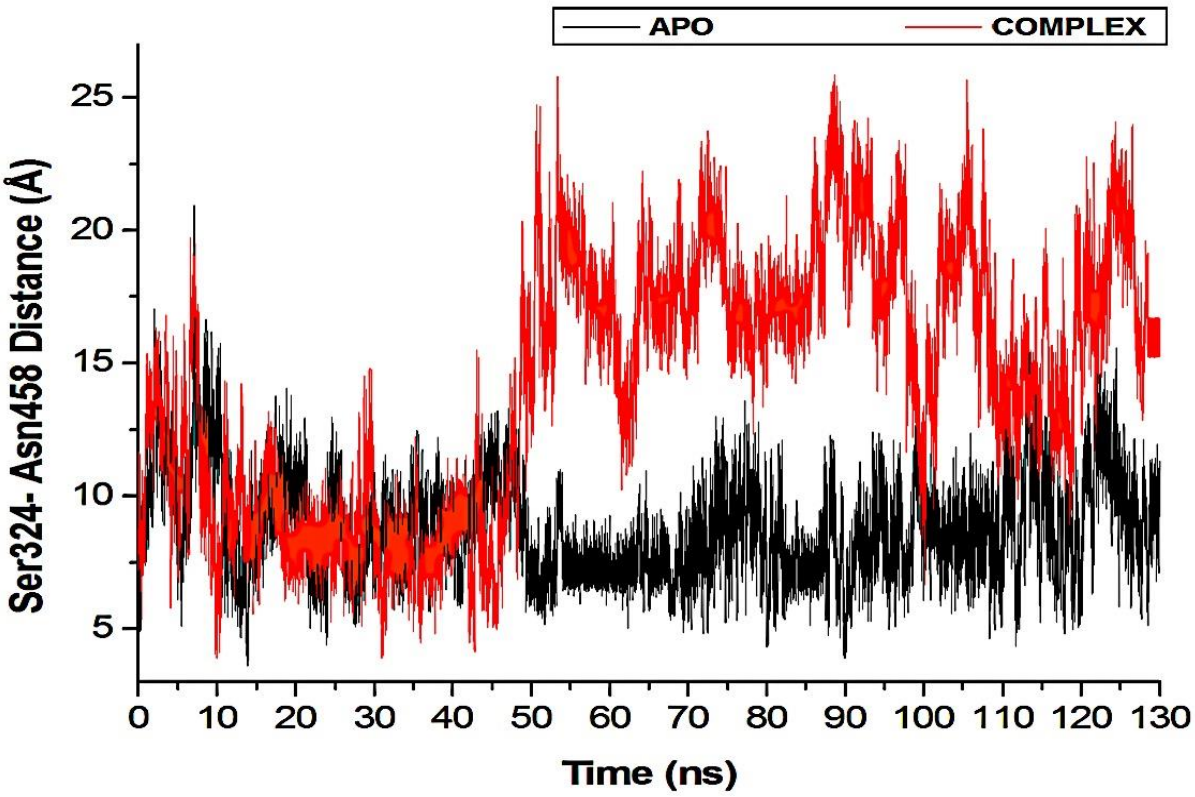


Figure 5.9: Residue fluctuations at the "325-348" region. The Apo enzyme illustrates widening of the loops of the Apo enzyme. The rear loop shifts down as the P-loop closes in on the active site. The largest fluctuation is seen after system stabilization at 40-60ns. The NITD008-Helicase enzyme shows instability in both loops throughout the simulations, although, there was no widening of the loops as the rear loop shifted back rather than downward movement seen in the Apo system.

5.3.3.2 Principal Component Analysis

Conformational transitions of the free protein and NITD008-bound complex were characterized using PCA, a technique that has been widely employed to present experimentally detected conformational variations. Figure 5.10 highlights the motional shifts across two principle components in the case of NITD008-bound and unbound NS3 Helicase. It is evident that eigenvectors computed from the respective simulations varied immensely between the two systems, further elaborating on the dynamic conformational fluctuations from free to ligand-bound protein. The unbound system shows restricted structural motions of residue Cα atoms, whereby the NITD008-bound system shows a larger spatial occupancy, thus substantiating the rigidity of the unbound system. This corresponds with the stability of the systems, illustrating greater distribution of the atoms around the center of mass and the system stability deviations for the NITD008-bound system. Correlation from analysis of both the free and bound protein demonstrates structural loop flexibility after binding of NITD008 to the ATPase active site.

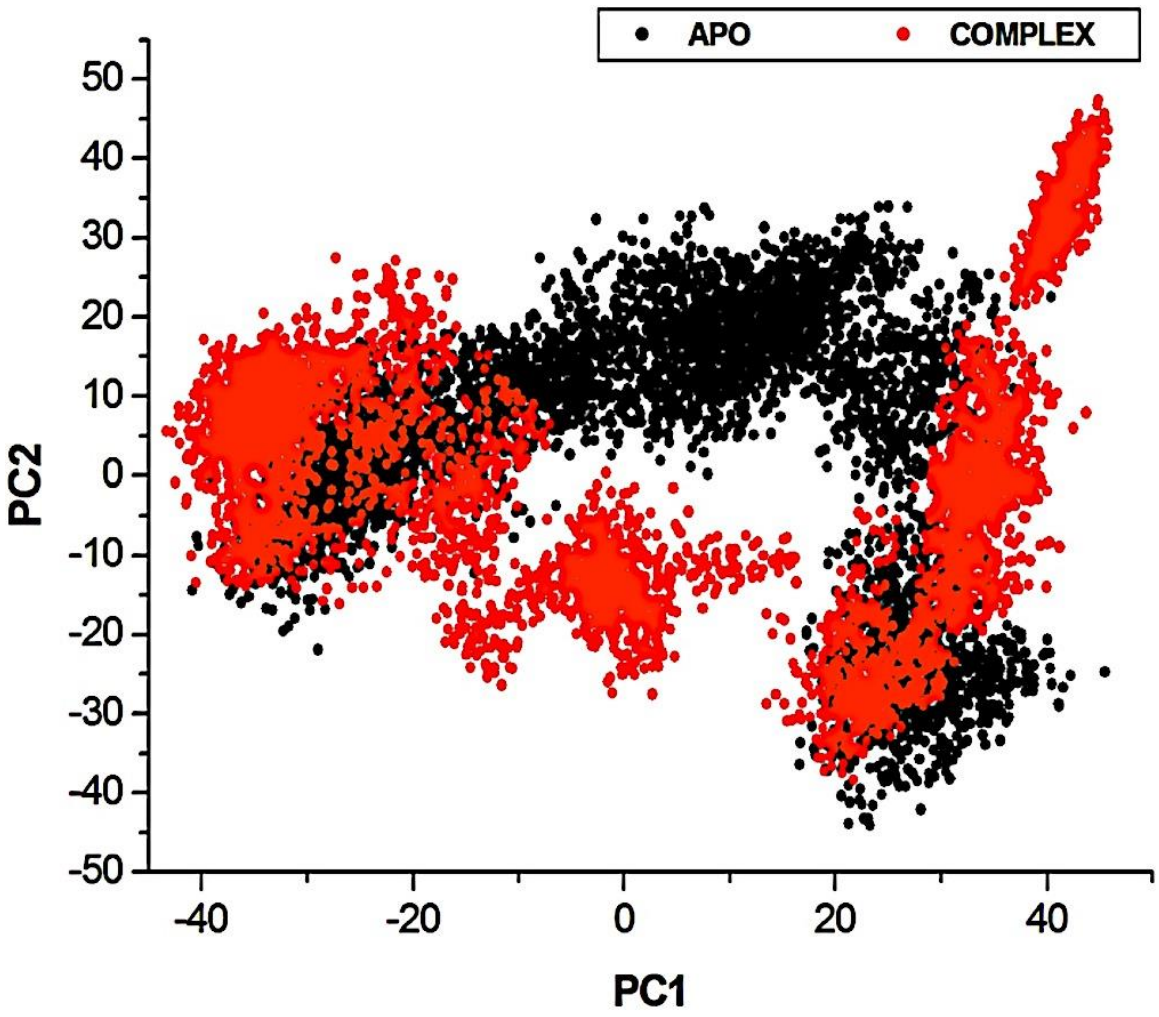


Figure 5.10: Projection of Eigen values of the Cα backbone, during 130ns simulation, for Apo and NITD008-bound conformations of NS3 Helicase along the first two principal components. The X- and Y- axis, PC1 and PC2, respectively, represent a covariance matrix after elimination of eigenvectors (rotational movements). Each point between the single-directional motions represents a unique conformation during the simulation, whereby, similar structural conformations overlap in the graph.

5.4. Conclusion

The detailed MD analyses provided in this report demonstrate the structural alterations in ZIKV NS3 Helicase loop flexibility subsequent to binding of potent inhibitor, NITD008 [37]. Molecular simulations revealed profound motional shifts of the ZIKV P-Loop at the ATPase active site. This flexibility was revealed in the RMSF analysis and verified by graphical investigation of the loop at different time intervals during the simulation. Investigation into the dynamic cross-correlation of the unbound and bound systems as well as a plot of conformational poses along the first two principal components resulted in strongly significant structural flexibility of the NITD008-NS3 Helicase system compared to the rigid unbound protein. The P-loop has demonstrated similar motional shifts in other *flaviviruses* as well as in ZIKV, when natural substrate, ATP binds at the active site. The competitive inhibitor, NITD008, has been proven to effectively constrain ZIKV replication both *in vitro* and *in vivo*. Complex stability measured through the 130ns simulation showed consistency of NITD008 at the ATPase active site and binding free energy calculations and residue-ligand networks revealed strong stabilizing hydrophobic and hydrogen bond interactions pocketing NITD008 in the active site. Further conformational changes were illustrated by the "325-338" loop shift behind the active site and the "339-348" region being modified from a fluctuating $_3$10 Helix to a more stable α-Helix.

Crystallographic studies have identified the P-loop, specifically Lys200, to be critical in stabilizing the triphosphate moiety of an NTP, thus allowing flexibility upon ligand binding and activation [12–14]. To augment these key findings, Lys200 showed strong hydrogen bonds with the NTP-analogue, NITD008. Other active-hotspot residues included P-loop residues: Gly197-Arg202, Ala198, Glu286, Gly415, Asn417 and Arg456. The insights demonstrating the above binding landscape of the ZIKV NS3 Helicase will aid researchers in the identification of

targeted-small molecule inhibitors through structure based drug design and to utilize pharmacophore models in screening for effective drugs with minimal toxicity.

Future experimental analysis is needed to fully understand these loop shifts toward inhibition of the enzyme as well as investigations into possible mutational resistance as seen in other *flavivirus* Helicase NTPase sites.

Acknowledgements

The authors acknowledge the National Research Foundation for their financial support (UID: 102103) and the Center for High Performance Computing (http://www.chpc.ac.za) for their computational resources.

References

1 L. Broxmeyer and R. Kanjhan, *Mod. Res. Inflamm.*, 2016, **5**, 20–30.

2 Centers for Disease Control, *Centers Dis. Control Prev. Zika Virus Home*, 2016, 1–12.

3 S. A. Rasmussen, D. J. Jamieson, M. A. Honein and L. R. Petersen, *N. Engl. J. Med.*, 2016, **374**, 1981–1987.

4 A. M. Palomo, *J. Public Health Policy*, 2016, **37**, 133–135.

5 WHO, 2016, 1–12.

6 O. Faye, C. C. M. Freire, A. Iamarino, O. Faye, J. V. C. de Oliveira, M. Diallo, P. M. Zanotto and A. A. Sall, *PLoS Negl. Trop. Dis.*, 2014, **8**, 1–10.

7 R. Tilak, S. Ray, V. W. Tilak and S. Mukherji, *Med. J. Armed Forces India*, 2016, **72**, 157–163.

8 C. G. Noble, Y. L. Chen, H. Dong, F. Gu, S. P. Lim, W. Schul, Q. Y. Wang and P. Y. Shi, *Antiviral Res.*, 2010, **85**, 450–462.

9 M. Mahfuz, A. Khan, H. Al Mahmud, M. Hasan, A. Parvin, N. Rahman and S. M. B. Rahman, *Indian J. Pharm. Biol. Res.*, 2014, **2**, 44–57.

10 C. Zanluca, C. N. Duarte and D. Santos, *Microbes Infect.*, 2016, **18**, 295–301.

11 P. Ramharack and M. E. S. Soliman, *RSC Adv.*, 2016, **6**, 68719–68731.

12 R. Jain, J. Coloma, A. Garcia-Sastre and A. K. Aggarwal, *Nat. Struct. Mol. Biol.*, 2016, **2**, 1–4.

13 H. Tian, X. Ji, X. Yang, W. Xie, K. Yang, C. Chen, C. Wu, H. Chi, Z. Mu, Z. Wang and H. Yang, *Protein Cell*, 2016, **7**, 450–454.

14 H. Tian, X. Ji, X. Yang, Z. Zhang, Z. Lu, K. Yang, C. Chen, Q. Zhao, H. Chi, Z. Mu, W. Xie, Z. Wang, H. Lou, H. Yang and Z. Rao, *Protein Cell*, 2016, **7**, 562–570.

15 X. Cao, Y. Li, X. Jin, Y. Li, F. Guo and T. Jin, *Nucleic Acids Res.*, 2016, **44**, 10505–10514.

16 E. D'Ortenzio, S. Matheron, X. de Lamballerie, B. Hubert, G. Piorkowski, M. Maquart, D. Descamps, F. Damond, Y. Yazdanpanah and I. Leparc-Goffart, *N. Engl. J. Med.*, 2016, **374**, 2195–2198.

17 A. C. Gourinat, O. O, Connor, E. Calvez, C. Goarant and D.-R. M, *Emerg. Infect. Dis.*, 2015, **21**, 84–86.

18 M. J. Turmel, M. J. P. Hubert, Y. M. V Maquart, M. Guillou-Guillemette and I. Leparc-Goff, *Lancet*, 2016, **6736**, 2501–2501.

19 A. R. Plourde and E. M. Bloch, *Emerg. Infect. Dis.*, 2016, **22**, 1–15.

20 J.-M. Anaya, C. Ramirez-Santana, I. Salgado-Castaneda, C. Chang, A. Ansari, M. E. Gershwin, R. Martines, J. Bhatnagar, M. Keating, L. Silva-Flannery, A. Muehlenbachs, J. Gary, C. Woods, A. Parker, B. Wakerley, A. Uncini, N. Yuki, J. Anaya, Y. Shoenfeld, A. Rojas-Villarraga, R. Levy, M. Dalakas, B. Wakerley, N. Yuki, S. Kivity, M. Arango, M. Ehrenfeld, O. Tehori, Y. Shoenfeld, J. Anaya, A. Denman, B. Rager-Zisman, T. Kolter, H. Willison, N. Yuki, R. Lardone, N. Yuki, F. Irazoqui, G. Nores, I. Kostovic, R. Ghiulai, M. Sarbu, Z. Vukelic, C. Ilie, A. Zamfir, T. Bell, E. Field, H. Narang, O. Faye, C. Freire, A.

Iamarino, O. Faye, J. Oliveira, M. Diallo, A. Nahmias, S. Nahmias and D. Danielsson, *BMC Med.*, 2016, **14**, 1–3.

21 N. L. Bayless, R. S. Greenberg, T. Swigut, J. Wysocka and C. A. Blish, *Cell Host Microbe*, 2016, **20**, 423–428.

22 D. Olagnier, M. Muscolini, C. B. Coyne, M. S. Diamond and J. Hiscott, *DNA Cell Biol.*, 2016, **35**, 367–372.

23 J. B. Brault, C. Khou, J. Basset, L. Coquand, V. Fraisier, M. P. Frenkiel, B. Goud, J. C. Manuguerra, N. Pardigon and A. D. Baffet, *EBioMedicine*, 2016, **10**, 71–76.

24 H. Li, L. Saucedo-Cuevas, J. A. Regla-Nava, G. Chai, N. Sheets, W. Tang, A. V. Terskikh, S. Shresta and J. G. Gleeson, *Cell Stem Cell*, 2016, **19**, 593–598.

25 T. J. Nowakowski, A. A. Pollen, E. Di Lullo, C. Sandoval-Espinosa, M. Bershteyn and A. R. Kriegstein, *Cell Stem Cell*, 2016, **18**, 591–596.

26 J. Cohen, *Science (80-.).*, 2016, **351**, 543–544.

27 E. Kim, G. Erdos, S. Huang, T. Kenniston, L. D. Falo and A. Gambotto, *EBioMedicine*, 2016, **13**, 315–320.

28 T. C. Pierson and B. S. Graham, *Cell*, 2016, 167, 625–631.

29 G. W. A. Dick, S. F. KIitchen and A. J. Haddow, *Trans. R. Soc. Trop. Med. Hyg.*, 1969, **63**, 708–737.

30 R. W. Malone, J. Homan, M. V Callahan, J. Glasspool-Malone, L. Damodaran, A. D. B. Schneider, R. Zimler, J. Talton, R. R. Cobb, I. Ruzic, J. Smith-Gagen, D. Janies, J.

Wilson, D. Hone, S. Hone, S. Bavari, V. Soloveva and S. Weaver, *PLoS Negl. Trop. Dis.*, 2016, **10**, 1–26.

31 A. N. Hazin, A. Poretti, D. Di Cavalcanti Souza Cruz, M. Tenorio, A. van der Linden, L. J. Pena, C. Brito, L. H. V. Gil, D. de Barros Miranda-Filho, E. T. de A. Marques, C. M. Turchi Martelli, J. G. B. Alves and T. A. Huisman, *N. Engl. J. Med.*, 2016, **374**, 2193–2195.

32 N. Gruba, J. I. Rodriguez Martinez, R. Grzywa, M. Wysocka, M. Skorenski, M. Burmistrz, M. Lecka, A. Lesner, M. Sienczyk and K. Pyrc, *FEBS Lett.*, 2016, **590**, 3459–3468.

33 B. D. Cox, R. A. Stanton and R. F. Schinazi, *Antivir. Chem. Chemother.*, 2016, **24**, 118–126.

34 J. Lei, G. Hansen, C. Nitsche, C. D. Klein, L. Zhang and R. Hilgenfeld, *Science (80-.).*, 2016, **353**, 503–5.

35 D. Luo, T. Xu, R. P. Watson, D. Scherer-Becker, A. Sampath, W. Jahnke, S. S. Yeong, C. H. Wang, S. P. Lim, A. Strongin, S. G. Vasudevan and J. Lescar, *EMBO J.*, 2008, **27**, 32090–3219.

36 A. V. Chernov, S. A. Shiryaev, A. E. Aleshin, B. I. Ratnikov, J. W. Smith, R. C. Liddington and A. Y. Strongin, *J. Biol. Chem.*, 2008, **283**, 17270–17278.

37 Y. Q. Deng, N. N. Zhang, C. F. Li, M. Tian, J. N. Hao, X. P. Xie, P. Y. Shi and C. F. Qin, *Open Forum Infect. Dis.*, 2016, **3**, 1–4.

38 H. M. Berman, T. Battistuz, T. N. Bhat, W. F. Bluhm, E. Philip, K. Burkhardt, Z. Feng, G.

L. Gilliland, L. Iype, S. Jain, P. Fagan, J. Marvin, D. Padilla, V. Ravichandran, N. Thanki, H. Weissig and J. D. Westbrook, *Biol. Crystallogr.*, 2002, **58**, 899–907.

39 S. Kim, P. A. Thiessen, E. E. Bolton, J. Chen, G. Fu, A. Gindulyte, L. Han, J. He, S. He, B. A. Shoemaker, J. Wang, B. Yu, J. Zhang and S. H. Bryant, *Nucleic Acids Res.*, 2016, **44**, 1202–1213.

40 S. Kusumaningrum, E. Budianto, S. Kosela, W. Sumaryono and F. Juniarti, *J. Appl. Pharm. Sci.*, 2014, **4**, 47–53.

41 Z. Yang, K. Lasker, D. Schneidman-Duhovny, B. Webb, C. C. Huang, E. F. Pettersen, T. D. Goddard, E. C. Meng, A. Sali and T. E. Ferrin, *J. Struct. Biol.*, 2012, **179**, 269–278.

42 S. Cosconati, S. Forli, A. L. Perryman, R. Harris, D. S. Goodsell and A. J. Olson, *Expert Opin. Drug Discov.*, 2010, **5**, 597–607.

43 M. F. Sanner, *Scripps Res. Inst.*, 2008, **26**, 1–12.

44 O. Trott and A. J. Olson, *J. Comput. Chem.*, 2010, **31**, 445–461.

45 N. Huang, B. K. Shoichet and J. J. Irwin, *J. Med. Chem.*, 2012, **49**, 6789–6801.

46 P. C. Nair and J. O. Miners, *Silico Pharmacol.*, 2014, **2**, 1–4.

47 M. Ylilauri and O. T. Pentikäinen, *J. Chem. Inf. Model.*, 2013, **53**, 2626–2633.

48 T. Hou, J. Wang, Y. Li and W. Wang, *J. Chem. Inf. Model.*, 2011, **51**, 69–82.

49 V. Gosu and S. Choi, *Sci. Rep.*, 2014, **4**, 1–13.

50 A. M. Martinez and A. C. Kak, *Trans. Pattern Anal. Mach. Intell.*, 2001, **23**, 228–233.

51 E. Seifert, *J. Chem. Inf. Model.*, 2014, **54**, 1552.

52 D. Ramirez and J. Caballero, *Int. J. Mol. Sci.*, 2016, **17**, 1–15.

53 X.-Y. Meng, H.-X. Zhang, M. Mezei and M. Cui, *Curr. Comput. Aided. Drug Des.*, 2011, **7**, 146–57.

54 L. G. Ferreira, R. N. Dos Santos, G. Oliva and A. D. Andricopulo, *Molecular docking and structure-based drug design strategies*, 2015, vol. 20.

55 P. A. Greenidge, C. Kramer, J. C. Mozziconacci and R. M. Wolf, *J. Chem. Inf. Model.*, 2013, **53**, 201–209.

56 J. M. Hayes and G. Archontis, *InTech*, 2011, 171–190.

57 F. Godschalk, S. Genheden, P. Söderhjelm and U. Ryde, *Phys. Chem. Chem. Phys.*, 2013, **15**, 7731–9.

CHAPTER 6

Conclusion and future recommendations

This chapter outlines the general conclusion of the study as well as recommendations for future research based on its findings.

6.1. General conclusions

The research study reported in this thesis is an in-silico studies on Zika NS3 Helicase: bedrock for antiviral drug design. It aims are (a) to provide a molecular understanding on the various drug binding landscape of Zika NS3 helicase protein (b) to investigate the structural dynamics and conformational changes of the protein. This work has accomplished the aims of this study; the results from this work confirmed the following conclusions:

The results of Objective 1 found that:

1. Ivermectin, HMC-HO1α and Lapachol emerged as the best three ligands out of the 10 clinically proven flaviviral NS3 small molecule inhibitor that were docked into the Zika NS3 helicase protein.

2.The binding of Ivermectin to ssRNA site and Lapachol and HMC-HO1α to the ATPase site allows for conformational rigidity of the Zika NS3 helicase, thus stabilizing residue fluctuations.

3.Ivermectine interacted with the protein at the ssRNA binding site through hydrophobic interactions, and hydrogen and covalent bond formation with the protein residues. While HMC-HO1α and Lapachol interacted with the ATPase active sites residues by forming hydrogen bonds and hydrophobic interactions. However, HMC-HO1α also formed a covalent interaction with residue Arg288 of the protein.

4. Free binding energy calculations showed Ivermectin to have a relatively higher binding energy (-84.56kcal/mol) at the ssRNA site than HMC-HO1α (-42.81kcal/mol), and Lapachol (-39.32kcal/mol) at the ATPase site. Furthermore, residue Glu112 had the highest energy contribution to the binding of HMC-HO1α (-4.08kcal/mol), and Lapachol (-3.05kcal/mol) to NS3 helicase protein binding at the ATPase site, while residue Arg214 had the highest total energy contribution to the binding of Ivermectin (-5.84kcal/mol) to the NS3 helicase protein at the ssRNA site.

The results of Objective 2 found that

5. Investigating the dynamic cross-correlation of the unbound and bound systems resulted in strongly significant structural flexibility of the NITD008-NS3 Helicase system compared to the rigid unbound protein. This include:

i) The RNA-binding loop shifting down in the Apo structure but shifting upward in the NITD008- Helicase complex.
ii) The P-Loop shifting away from the active site in the bound complex but being closed on the active site when no ligand was present.
iii) The "325-338" loop shifting behind the active site and the "339-348" region being modified from a fluctuating $_3$10 Helix to a more stable α-Helix.

6. The principal component analysis revealed a motional shift from a restricted structural motion of the unbounded form of the protein to a structural loop flexibility of the NITD008-NS3 Helicase system.

The results presented herein have clearly contributed towards the molecular understanding of drug-enzyme interactions and structural and conformational evolution upon ligand binding. It has also made it clear that inhibition of the non-structural protein of Zika virus at the replication stage of its life cycle will have great advantages to the clinical therapeutics of the virus. Furthermore, the computational techniques presented in this work has proven to be beneficial tools towards drug discovery and development. In addition, this research could contribute substantially to the design of novel inhibitor against the Zika virus, which will eventually bring about a cure to the notorious virus. Therefore, resulting to an improved quality of life of those affected by the virus, especially the pregnant women and children that were born with microcephally, Gullian-Barrè syndrome and other neurological diseases that are associated to the virus thereby, improving the health of the public at large.

6.2. Recommendation and Future Studies

The work presented in this thesis have identified lead compounds with good binding affinities. However, the lack of sufficient experimental data which could serve as a reference point for the presented models was a major challenge. Also, the lack of literatures with information on the structural conformations of the protein was also a challenge. These challenges serve as a source of motivation for this work.

As this is the first study reporting on drug-enzyme interactions and structural and conformational evolution upon ligand binding of Zika virus NS3 helicase, the following recommendations are noted to take this direction of research further:

- A better free binding energy result can be achieved by using experimental data as a reference point.
- Synthesis and prospective biological testing of these lead compounds is still required rationalizing their activity against Zika NS3 helicase.
- Experimental assessments are needed to fully understand the loop shifts towards inhibiting the enzyme, and longer dynamic simulations may need to be run to validate the consistency of the loop flexibility.
- Sophisticated computational approaches, such as, coarse grained molecular dynamics[1], Substrate Envelope Analysis (SEA)[2] and Quantitative Structure-Activity Relationships (QSAR)[3], would provide a better understanding of the enzyme dynamics during the catalytic course.

References

(1) Siuda, I., and Thøgersen, L. (2013) Conformational flexibility of the leucine binding protein examined by protein domain coarse-grained molecular dynamics. *J. Mol. Model.19*, 4931–4945.

(2) Shen, Y., Altman, M. D., Ali, A., Nalam, M. N. L., Cao, H., Rana, T. M., Schiffer, C. A., and Tidor, B. (2013) Testing the substrate-envelope hypothesis with designed pairs of compounds. *ACS Chem. Biol.8*, 2433–2441.

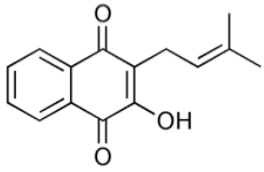
(3) Hirst, J. D., McNeany, T. J., Howe, T., and Whitehead, L. (2002) Application of non-parametric regression to quantitative structure-activity relationships. *Bioorg. Med. Chem.10*, 1037–41.

APPENDICES

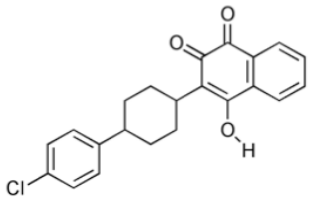
Appendix 1. Supplementary material for Chapter 4

Characterizing the Ligand Binding Landscape of Zika NS3 Helicase- Promising Lead Compounds as Potential Inhibitors.

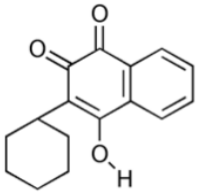
Sofiat Oguntade[1], Pritika Ramharack[1] and Mahmoud E. S. Soliman[1*]



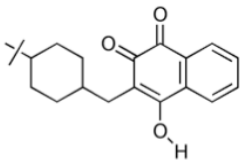
LAPACHOL



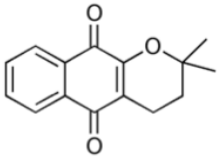
ATOVAQUONE



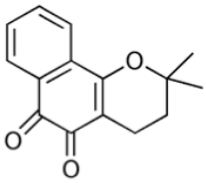
PARVAQUONE



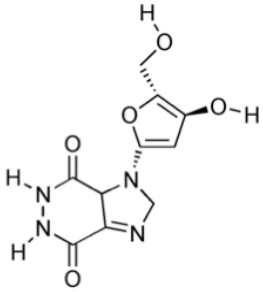
BUPARVAQUONE



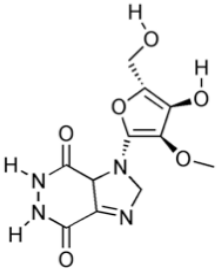
ALPHA LAPACHONE



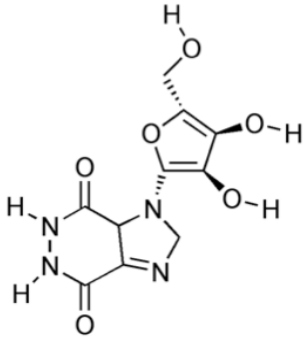
BETA LAPACHONE



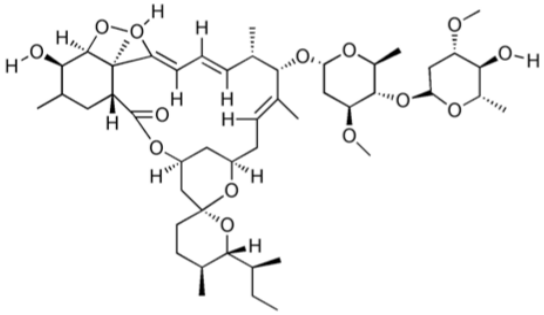
HMC-HO1α



HMC-HO4



HMC-HO5



IVERMECTINE

Figure S1: 2D structures of potential ligand

Table S.1: Grid box parameter for molecular docking.

Site	Centre	dimension
ssRNA	X=-7.8, Y=20.7, Z=-4.8	X=20, Y=56, Z=28
ATP	=-19.8, Y=20.1, Z=-23.2	X=26, Y=28, Z=40

Table S.2: Binding energy result from docking selected potential Zika NS3 helicase inhibitors.

Site	Potential ligands	Docking Score (kcal/mol)
ATP	**Naphthaquinones**	
	Atovaquone	-9.1
	Buparvaquone	-8.7
	Parvaquone	-8.2
	α-Lapachone	-7.8
	β-Lapachone	-7.5
	Lapachol	-6.3
	purine nucleoside analogues	
	HMC-HO5	-7.0
	HMC-HO4	-7.0
	HMC-HO1α	-6.9
ssRNA	Ivermectin	-6.4

Appendix 2. Supplementary material for Chapter 5

Delving into Zika Virus Structural Dynamics- A Closer look at NS3 Helicase Loop flexibility and its Role in Drug Discovery

Pritika Ramharack[A], Sofiat Oguntade[A] Mahmoud E. S. Soliman[A*]

[A]Molecular Modeling and Drug Design Research Group, School of Health Sciences, University of KwaZulu-Natal, Westville Campus, Durban 4001, South Africa

Supplementary Data

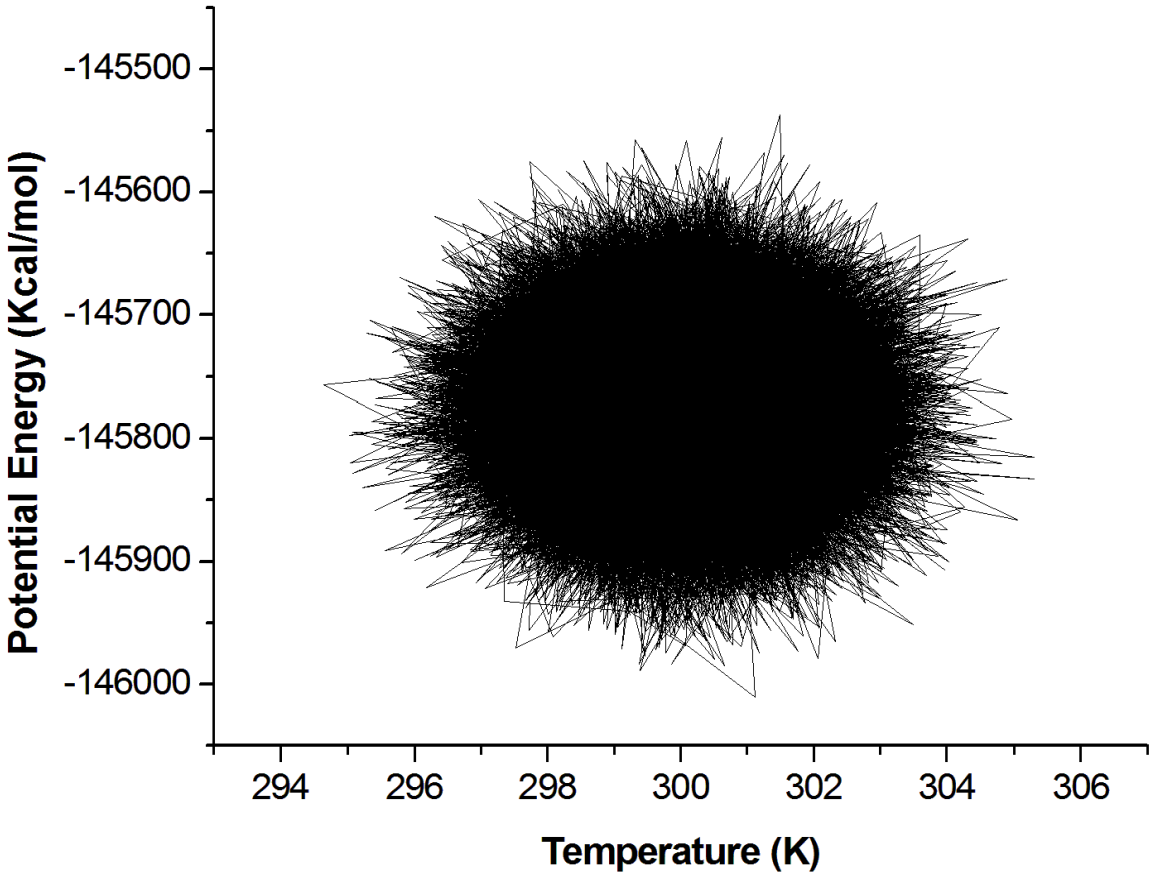


Figure S1: Potential Energy Fluctuations of the NITD008-NS3 Helicase System at varying temperatures during the 100ns simulation. The average temperature of the system was 300K and the average potential energy was -145774 kcal/mol.

Energy Components (kcal/mol)					
	Δ E_{vdW}	**ΔE_{elec}**	**ΔG_{gas}**	**ΔG_{solv}**	**ΔG_{bind}**
ZIKV NS3 Helicase	-3573.27 ± 30.81	-30020.42 ± 127.46	-33593.70 ± 140.40	-4481.84 ± 97.48	-38075.53 ± 73.37
NITD008 (Pose 1)	-4.82 ± 1.08	-12.26 ± 5.98	7.43 ± 6.09	-24.39 ± 1.89	-16.96 ± 5.29
Complex	-32.11 ± 4.09	-31.36 ± 8.72	-63.46 ± 9.18	33.47 ± 4.56	**-30.00 ± 5.58**

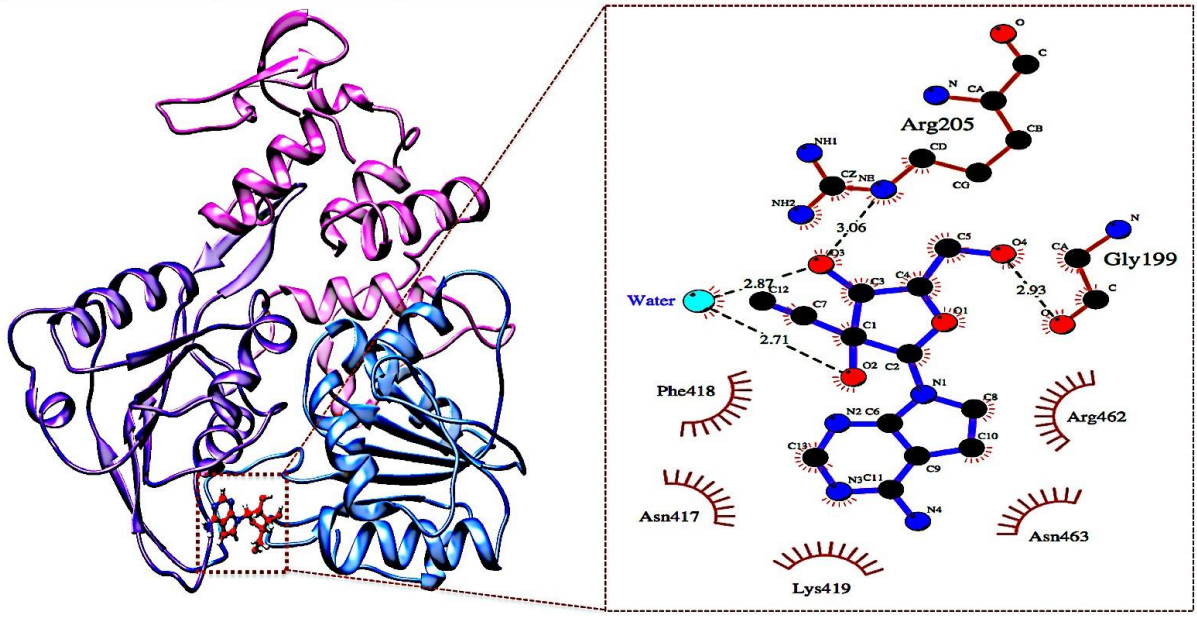


Figure S2: Complex of NITD008-NS3 Helicase with a Docking score of -7.7 kcal/mol. MM/GBSA calculations yielded a result of -30.00 kcal/mol. The ligand shifted further out of the hydrophobic pocket after 150ns of the simulation. This may possibly be due to the ligand not interacting with the stabilizing residues of the P-loop**.**

Energy Components (kcal/mol)					
	Δ E_{vdW}	**ΔE_{elec}**	**ΔG_{gas}**	**ΔG_{solv}**	**ΔG_{bind}**
ZIKV NS3 Helicase	-3567.30 ± 31.08	- 29710.32 ± 175.80	-33277.62 ± 178.68	-4790.36 ± 137.78	-38067.98 ± 76.12
NITD008 (Pose 2)	-5.47 ± 1.01	-6.34 ± 7.85	0.87 ± 7.35	-27.84 ± 3.89	-26.97 ± 5.17
Complex	-24.12 ± 2.75	-9.35 ± 7.29	-33.47 ± 6.78	19.80 ± 5.62	**-13.67 ± 3.01**

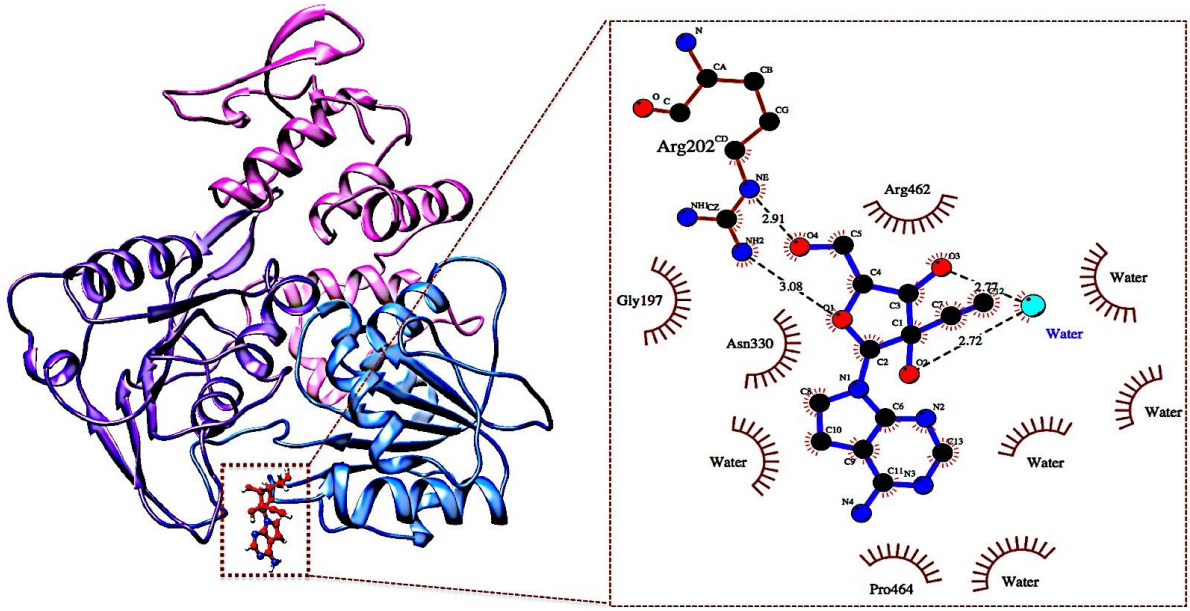


Figure S3: Complex of NITD008-NS3 Helicase with a Docking score of -7.6 kcal/mol. MM/GBSA calculations yielded a result of -13.67 kcal/mol.The ligand docked out of the hydrophobic pocket and during the simulation, due to the lack of stabilizing interactions, the ligand moved further out of the active site and into the solvent.

Energy Components (kcal/mol)					
	ΔE_{vdW}	ΔE_{elec}	ΔG_{gas}	ΔG_{solv}	ΔG_{bind}
ZIKV NS3 Helicase	-3564.79 ± 26.89	- 29878.89 ± 131.37	-33443.68 ± 133.78	-4633.21 ± 88.03	-38076.89 ± 68.96
NITD008 (Pose 3)	-5.56 ± 0.60	10.53 ± 6.58	4.96 ± 6.51	-28.45 ± 3.51	-23.49 ± 4.78
Complex	-17.75 ± 4.65	-14.45 ± 7.46	-32.21 ± 8.76	20.35 ± 5.95	**-11.86 ± 6.37**

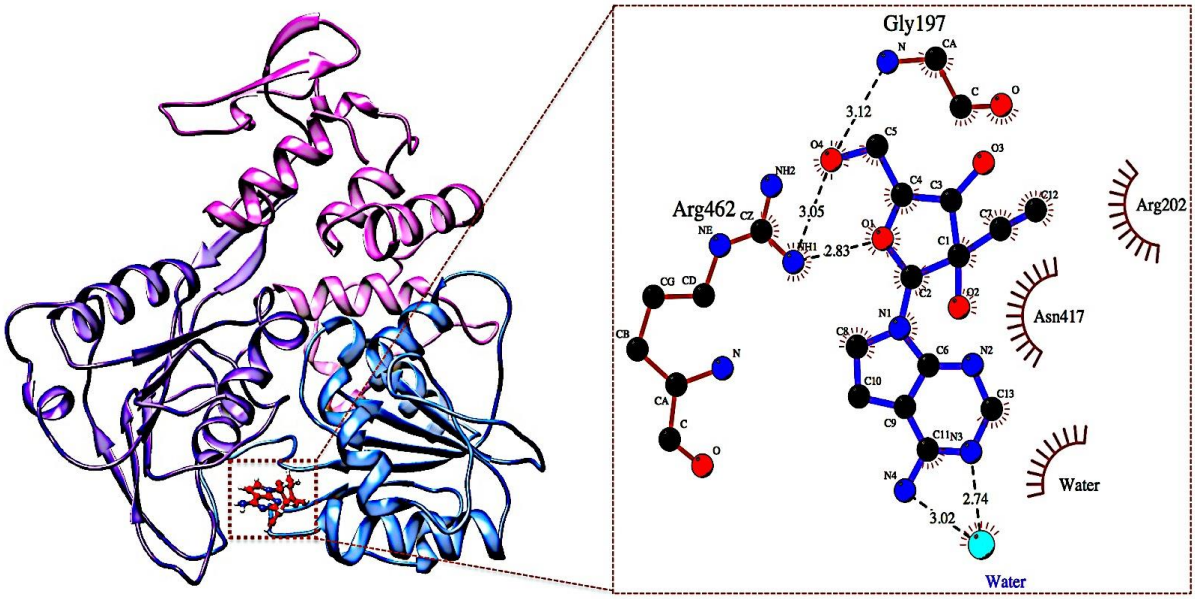


Figure S4: Complex of NITD008-NS3 Helicase with a Docking score of -7.1 kcal/mol. MM/GBSA calculations yielded a result of -11.86 kcal/mol.This ligand showed a similar pose to that of the -7.6 kcal/mol-docked pose, however, there was only one residue, Arg462, which showed stabilizing hydrogen bonds with the terminal oxygen located on the ribose group of NITD008.

Energy Components (kcal/mol)					
	ΔE_{vdW}	ΔE_{elec}	ΔG_{gas}	ΔG_{solv}	ΔG_{bind}
ZIKV NS3 Helicase	-3546.66 ± 28.91	- 30102.45 ± 107.82	-33649.11 ± 103.65	-4466.57 ± 89.33	-38115.68 ± 48.08
NITD008 (Pose 4)	-5.86 ± 0.44	12.17 ± 4.43	6.31 ± 4.48	-28.40 ± 2.23	-22.09 ± 3.39
Complex	-27.60 ± 3.22	-25.88 ± 6.98	-53.48 ± 6.59	29.49 ± 4.00	**-23.99 ± 4.06**

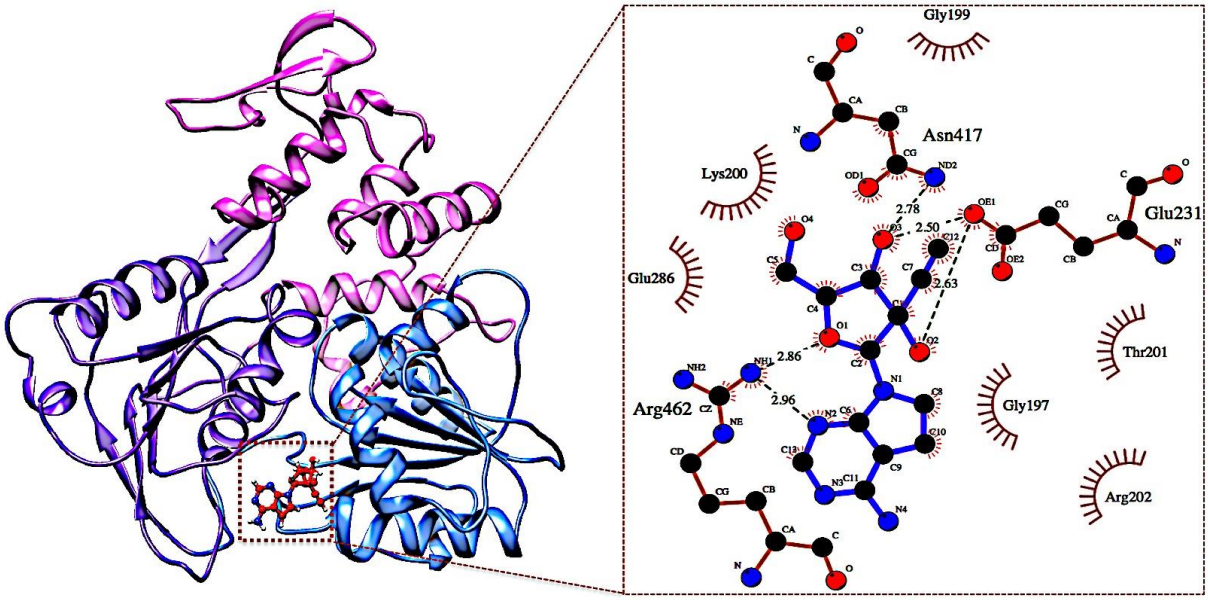


Figure S5: Complex of NITD008- NS3 Helicase with a Docking score of -7.1 kcal/mol. MM/GBSA calculations yielded a result of -23.99 kcal/mol.This pose showed the same docking score as the above ligand, however, three residues: Arg462, Asn417, and Glu231, were involved in stabilizing hydrogen bonds.

Energy Components (kcal/mol)					
	ΔE_{vdW}	ΔE_{elec}	ΔG_{gas}	ΔG_{solv}	ΔG_{bind}
ZIKV NS3 Helicase	-3541.38 ± 29.25	- 29907.78 ± 118.22	-33449.15 ± 118.80	-4626.31 ± 114.92	-38075.46 ± 56.93
NITD008 (Pose 5)	-5.54 ± 0.70	12.28 ± 5.43	6.74 ± 5.32	-26.45 ± 2.46	-19.71 ± 4.02
Complex	-13.69 ± 3.83	-7.75 ± 6.43	-21.44 ± 7.73	15.54 ± 5.95	**-5.90 ± 3.06**

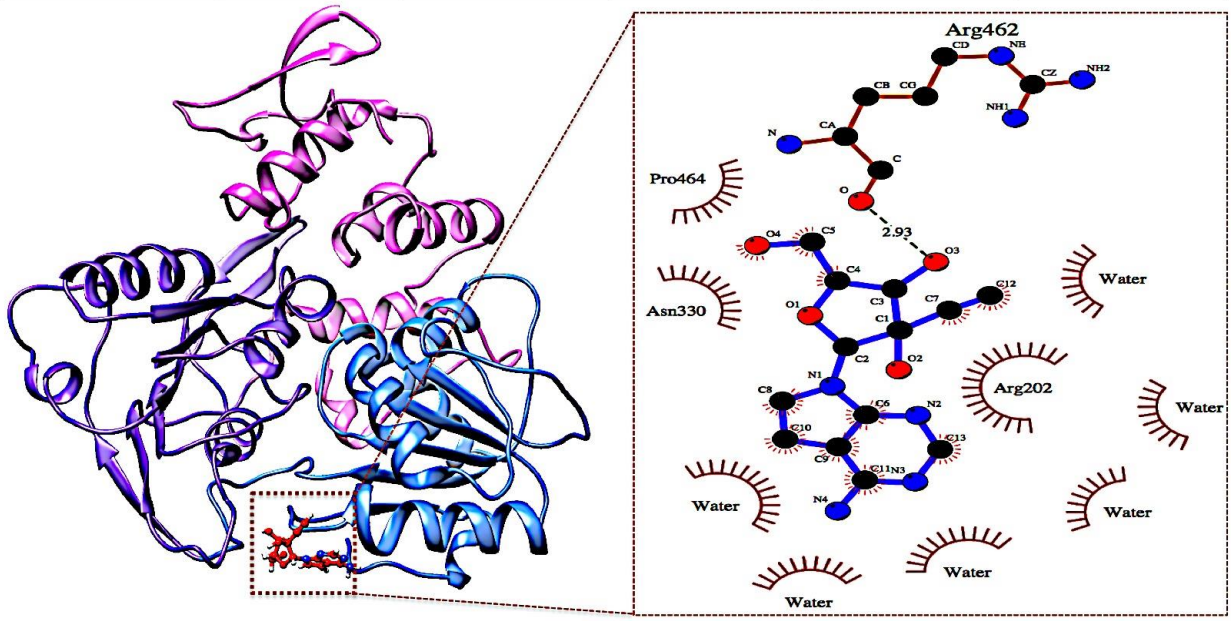


Figure S6: MM/GBSA calculations yielded a result of -5.90 kcal/mol, which was lower than that of the docking score of 6.9 kcal/mol. This was due to the ligand binding out of the active site of the enzyme, thus leading to minimal intermolecular forces at the hydrophobic pocket.