

Structural and functional characterization of flavivirus non-structural protein 3 (NS3) in solution and atomic and enzymatic insights of vancomycin resistant enterococcus faecalis (V583) alkyl hydroperoxide subunit C

Pan, Ankita

2018

Pan, A. (2018). Structural And functional characterization of flavivirus non-structural protein 3 (NS3) in solution and atomic and enzymatic insights of vancomycin resistant enterococcus faecalis (V583) alkyl hydroperoxide subunit C. Doctoral thesis, Nanyang Technological University, Singapore.

<https://hdl.handle.net/10356/87918>

<https://doi.org/10.32657/10220/45578>



**Structural and functional characterization of flavivirus
non-structural protein 3 (NS3) in solution and atomic and
enzymatic insights of vancomycin resistant *Enterococcus
faecalis* (V583) alkyl hydroperoxide subunit C**

ANKITA PAN
SCHOOL OF BIOLOGICAL SCIENCES
2018

**Structural and functional characterization of flavivirus
non-structural protein 3 (NS3) in solution and atomic and
enzymatic insights of vancomycin resistant *Enterococcus*
faecalis (V583) alkyl hydroperoxide subunit C**

ANKITA PAN

SCHOOL OF BIOLOGICAL SCIENCES

A thesis submitted to Nanyang Technological University in partial
fulfilment of the requirement for the degree of

Doctor of Philosophy

2018

Acknowledgements

The satisfaction that accompany the completion of this work would be incomplete without mentioning the people who made it possible.

First and foremost, I would like to extend my sincere gratitude to my advisor, Dr. Gerhard Grüber. I am grateful, not only for his valuable advice regarding my work, but also for teaching me discipline, managing time efficiently and most importantly, to enjoy the result of one's hard work.

I would like to thank my thesis advisory committee members: Prof. Salil Bose, Prof. Susana Schochat Geifman, Prof. Mark Featherstone and Prof. Yoon Ho Sup for their valuable suggestions regarding my work. I am grateful to Prof. Shashi Bhushan for the analysis of EM data.

I would also like to take this opportunity to thank Dr. Ardina Grüber. Gracious and loving, she was a huge support for me and made me believe in myself. While I was traversing a difficult path in my life, I found you by my side.

I would like to thank Dr. Wan Geok Saw for her continuous help and advice in the work of the dengue project. I am extremely thankful to Dr. Malathy Sony Manimekalai who patiently guided me through the SAXS data collection and subsequent data analysis.

I would also like to thank Dr. Yew Kwang Toh who initially guided me through the cloning steps for the *E. faecalis* AhpC constructs. I am grateful to Dr. Asha M. Balakrishna for her help in collecting X-ray data of *EfAhpC*₁₋₁₇₂ at the National Synchrotron Radiation Research Centre (NSRRC, Hsinchu, Taiwan). Her invaluable advice in crystallography helped me in the crystallization and the structural determination of *EfAhpC*₁₋₁₇₂. Many thanks to Dr. Neelagandan Kamariah for his guidance concerning the NADH-dependent peroxidative assay for the *E. faecalis* AhpC proteins and for his scientific insights into the AhpC project.

I would also like to thank all former members of my lab. Special thanks to Dr. Wilson Nartey for cloning the *EfAhpC* construct, Dr. Phat Vinh Dip for the EM data analysis, Dr. Hendrik Seilaff, Dr. Dhiren Singh, Dr. Arvind Kumar for their helpful discussions and Ms. Sek Mun Foong for our animated discussions not only pertaining to science.

Many thanks go to my colleagues in this lab: Mr. Nebojsa Bogdanovič my fellow Geminian, Ms. Bharti Singal for her support and all the joy her endless conversations bring, Ms. Chui Fann Wong for meticulously going through this thesis and Ms. Sherilyn Chong Shi Min whose constant encouragement helped me to pull through times when I felt demotivated.

Thank you: Dr. Priya Ragunathan, Dr. Shin Joon, Dr. Harikishore Amaravadhi and Mr. Jack Aik Meng.

I gratefully acknowledge the financial support for my scholarship rendered by the Ministry of Education Tier 3 (MOE2012-T3-1-008) grant.

Words fall short to express my gratitude for my friends: Dipanwita, Ank and Ramya who has stood by me through thick and thin. Thank you Jeeva, Vishnu, Aditya and Annie for your support and making my stay here in Singapore memorable.

Lastly, I would like to express my deepest gratitude to my mother, father and grandfather who are always proud of my academic achievements and thank them for their selfless support, encouragement, unconditional love.

dedicated to ma and baba

TABLE OF CONTENT

LIST OF FIGURES	XI
LIST OF TABLES	XIV
ABBREVIATIONS	XVIII
ABSTRACT	XXII
1. INTRODUCTION	
1.1 Overview of flavivirus	1
1.1.1 Epidemiology of DENV and ZIKV	1
1.1.2 Flavivirus replicative cycle	2
1.1.3 Structure of the DENV and ZIKV particle	4
1.1.4 DENV and ZIKV genome assembly and replication	6
1.1.5 Non-structural proteins	11
1.1.5.1 Non-structural protein 1	11
1.1.5.2 Non-structural protein 2A	12
1.1.5.3 Non-structural protein 2B	12
1.1.5.4 Non-structural protein 3	13
1.1.5.5 Non-structural protein 4A	13
1.1.5.6 Non-structural protein 4B	13
1.1.5.7 Non-structural protein 5	14
1.1.6 NS3 structure and function	16
1.1.6.1 NS3 protease	18
1.1.6.2 NS3 helicase	20
1.2 Vancomycin resistant <i>Enterococcus faecalis</i> (V583) anti-oxidant system	21
1.2.1 Reactive oxygen species (ROS)	21
1.2.2 Anti-oxidant systems	23
1.2.3 The AhpR system	25
1.2.3.1 Alkyl hydroperoxide reductase subunit F	25
1.2.3.2 Alkyl hydroperoxide reductase subunit C	27
1.2.3.2.1 The active site of AhpC and its conformations	28
1.2.3.2.2 The catalytic cycle during peroxide reduction by AhpC	29
1.2.4 The <i>E. faecalis</i> anti-oxidant system	32
1.3 Goals of this thesis	33
2. MATERIAL AND METHODS	
2.1 Materials	37
2.1.1 Chemicals	37
2.1.2 Molecular biology materials	37
2.1.2.1 Enzymes and molecular biology kits	37
2.1.2.2 Template for generating gene constructs	37

2.1.2.3 Primers and synthetic oligonucleotides	38
2.1.2.4 Plasmids	38
2.1.2.5 <i>Escherichia coli</i> strains	38
2.1.3 Chromatography materials	39
2.1.4 Instruments	39
2.1.5 Consumables	39
2.1.6 Computer software	39
2.2 Methodology	40
2.2.1 Cloning of DENV-2 and ZIKV constructs	40
2.2.1.1 Cloning of DENV-2 NS2B ₁₈ NS3	40
2.2.1.2 Cloning of DENV-2 protease with linker (NS2B ₁₈ NS3 ₁₋₁₇₉)	44
2.2.1.3 Cloning of DENV2 helicase (NS3 ₁₆₉₋₆₁₈) with linker	46
2.2.1.4 Cloning of DENV-2 NS2B ₁₈ NS3 ₁₇₄ PPAVP ₁₇₉ mutant	47
2.2.1.5 Cloning of ZIKV NS2B ₄₈ NS3	47
2.2.1.6 Cloning of ZIKV NS2B ₄₈ NS3 protease with linker	50
2.2.1.7 Cloning of ZIKV NS2B ₁₈ NS3	52
2.2.1.8 Cloning of ZIKV NS2B ₁₈ NS3 mutants	54
2.2.2 Recombinant gene expression for flaviviral constructs	55
2.2.3 Induction test of viral proteins from DENV	55
2.2.4 Protein production of viral proteins from DENV	56
2.2.5 Purification of DENV-2 and ZIKV constructs	56
2.2.6 Small-angle X-ray scattering	59
2.2.6.1 SAXS data collection	59
2.2.6.2 SAXS data analysis	59
2.2.6.3 Protein folding and flexibility assessment using SAXS	60
2.2.6.4 Ab initio modeling using SAXS	61
2.2.6.5 Rigid-body modeling using SAXS	62
2.2.6.6 Ensemble modeling using SAXS	63
2.2.7 Ligation based cloning	64
2.2.7.1 Polymerase chain reaction	64
2.2.7.2 Agarose gel electrophoresis	65
2.2.7.3 Restriction digestion of insert or vector	66
2.2.7.4 Restriction digestion of insert/vector	66
2.2.7.5 CIAP treatment of digested vector	66
2.2.7.6 Ligation	67
2.2.8 Ligation independent cloning	67
2.2.8.1 PCR amplification	68
2.2.8.2 <i>DpnI</i> digestion	68
2.2.9 DNA precipitation	69
2.2.10 Transformation into DH5α cells	69
2.2.11 Isolation of plasmid	69
2.2.12 Recombinant gene expression	70
2.2.13 Induction test	70
2.2.14 Recombinant protein production and purification	70
2.2.15 Dynamic light scattering	72
2.2.16 Negative staining in electron microscope	72
2.2.17 NADH-dependent peroxidative assay	72

2.2.18 Ferrous oxidation xylenol orange assay	73
2.2.19 Crystallisation of <i>EfAhpC</i> ₁₋₁₇₂	74
2.2.19.1 Data collection and structure determination	75
2.2.19.2 Crystal structure determination of <i>EfAhpC</i> ₁₋₁₇₂	75
2.2.20 Protein sequence analysis	76

3. RESULTS – STRUCTURAL AND FUNCTIONAL CHARACTERIZATION OF FLAVIVIRAL NS3

3.1 Solution studies of DENV-2 NS3	80
3.1.1 Cloning of DENV-2 NS2B ₁₈ NS3 and purification of recombinant protein	80
3.1.2 Analysis of primary SAXS data for DENV-2 NS2B ₁₈ NS3	81
3.1.3 Construction of <i>ab initio</i> envelopes for DENV-2 NS2B ₁₈ NS3	83
3.1.4 Flexibility characterization of DENV-2 NS2B ₁₈ NS3	84
3.1.5 Cloning, production and purification of DENV-2 NS2B ₁₈ NS3 protease and helicase domains	86
3.1.6 SAXS of studies of DENV-2 protease and helicase domain	87
3.1.7 <i>Ab initio</i> model constructions for DENV-2 NS2B ₁₈ NS3 protease and helicase	90
3.1.8 Assessment of DENV NS3 domains flexibility	90
3.1.9 Cloning of DENV-2 NS2B ₁₈ NS3 _{174PPAVP179} and purification of the recombinant protein	91
3.1.10 Structural insights from SAXS experiment of DENV-2 NS2B ₁₈ NS3 mutant	91
3.1.11 ATPase activity for DENV-2 NS2B ₁₈ NS3 and helicase	92
3.2 Solution studies of ZIKV NS3	93
3.2.1 Cloning, expression and purification of recombinant ZIKV NS2B ₄₈ NS3	94
3.2.2 Generation of the ZIKV NS2B ₄₈ NS3 _{S135A} mutant	95
3.2.3 Cloning, expression and purification of recombinant ZIKV NS2B ₁₈ NS3	96
3.2.4 Primary SAXS data analysis of ZIKV NS2B ₁₈ NS3	98
3.2.5 Rigid body modelling of ZIKV NS2B ₁₈ NS3	100
3.2.6 <i>Ab initio</i> model constructions for ZIKV NS2B ₁₈ NS3	100
3.2.7 Flexibility characterization of ZIKV NS2B ₁₈ NS3	101
3.2.8 Generation of ZIKV NS3 protease and helicase domains	103
3.2.9 Primary SAXS data analysis for ZIKV NS3 domains	105
3.2.10 Assessment of ZIKV NS3 domains flexibility	106
3.2.11 <i>Ab initio</i> model constructions for ZIKV NS2B ₁₈ NS3 protease and helicase	107
3.2.12 Cloning, expression and purification of ZIKV NS2B ₁₈ NS3 mutants	108
3.2.13 Overall structural parameter of ZIKV NS2B ₁₈ NS3 _{F179A} and NS2B ₁₈ NS3 _{F179I}	112
3.2.14 ATPase activity of ZIKV NS2B ₁₈ NS3 and helicase	114

4. DISCUSSION – STRUCTURAL AND FUNCTIONAL CHARACTERIZATION OF FLAVIVIRAL NS3

4.1 Traits of NS3 in solution	118
4.1.1 Structural features of DENV-2 NS3 in solution	118
4.1.2 The linker residues of flaviviruses impart flexibility	119
4.1.3 The overall structural parameter of the mutant DENV-2 NS3 is similar to the WT	120
4.1.4 Both NS3 and its helicase domain were enzymatically active	120
4.2 ZIKV NS3 traits in solution	121
4.2.1 The recombinant ZIKV NS3 expressed in <i>E. coli</i> undergoes spontaneous auto-lysis	121

4.2.2 ZIKV NS3 is extended in solution, while the domains are compact	123
4.2.3 4.2.3 ZIKV NS3 linker is conserved among strains	123
4.2.4 ZIKV NS3 linker is important for the protein stability	124
4.2.5 ATPase inhibitory effects of the ZIKV NS2B ₁₈ NS3 and its helicase	125
5. RESULTS- ATOMIC AND ENZYMATIC INSIGHTS INTO EFAHPC	
5.1 Characterization of <i>EfAhpC</i>	128
5.1.1 Purification of <i>EfAhpC</i>	128
5.1.2 Enzymatic characterization of <i>EfAhpC</i>	129
5.1.3 Analysis of the oligomeric state of <i>EfAhpC</i> by DLS	130
5.1.4 Analysis of high molecular weight oligomers by EM	131
5.1.5 Crystallization of <i>EfAhpC</i>	132
5.2 Characterization of <i>EfAhpC</i>₁₋₁₇₂ and structural insights into <i>EfAhpC</i>₁₋₁₇₂	133
5.2.1 Cloning, expression and purification of <i>EfAhpC</i> ₁₋₁₇₂	133
5.2.2 Enzymatic characterization of <i>EfAhpC</i> ₁₋₁₇₂	134
5.2.3 Analysis of the oligomeric state of the truncated <i>EfAhpC</i>	135
5.2.4 Crystallisation and data collection of <i>EfAhpC</i> ₁₋₁₇₂ crystals	136
5.2.5 Crystal structure of <i>EfAhpC</i> ₁₋₁₇₂	140
5.3 Characterization of <i>EfAhpC</i>_{V78T} to investigate role of V78 in stable oligomer formation	144
5.3.1 Cloning, expression and purification of <i>EfAhpC</i> _{V78T}	144
5.3.2 Enzymatic characterization of <i>EfAhpC</i> _{V78T}	146
5.3.3 Analysis of destabilization of oligomer by DLS	146
5.4 Characterization of cysteine mutants to elucidate the role of C13 and C66 in <i>EfAhpC</i>	147
5.4.1 Generation of cysteine mutants and recombinant protein purification	147
5.4.2 Enzymatic characterization of cysteine mutants	149
5.4.3 Analysis of oligomeric state of cysteine mutants by DLS	150
5.5 Characterization of N- and C-terminal polar and charged residues unique to <i>EfAhpC</i>	150
5.5.1 Generation of N- and C-terminal mutants and purification of the enzyme variants	151
5.5.2 Enzymatic characterization of N- and C-terminal mutants	152
6. DISCUSSION- ATOMIC AND ENZYMATIC INSIGHTS INTO EFAHPC	
6.1.1 The <i>EfAhpC</i> forms an oxidized decamer	156
6.1.2 The structural intricacies of <i>EfAhpC</i>	156
6.1.3 The C-terminus does not play a role in oligomerization of <i>EfAhpC</i> but is crucial for peroxidative activity	160
6.1.4 The presence of valine at the oligomeric interface implicates a strong decamer formation	161
6.1.5 Residues C13 and C66 of <i>EfAhpC</i> contribute to decamer stability	162
6.1.6 Decamer destabilizing mutations of <i>EfAhpC</i> leads to reduced catalytic efficiency of <i>EfAhpC</i>	162
6.1.7 The unique residues at the N- and C-terminus contribute to the catalytic efficiency of the robust <i>EfAhpC</i>	162
7. CONCLUSION AND FUTURE PERSPECTIVE	143

8. REFERENCES	167
9. OTHERS	182

LIST OF FIGURES
INTRODUCTION

Figure 1.1.1: The viral replicative cycle.	4
Figure 1.1.2: Flavivirus maturation pathway.	4
Figure 1.1.3: Cryo-EM envelope of ZIKV virus.	6
Figure 1.1.4: Viral genome assembly and polyprotein produced.	7
Figure 1.1.5: Dengue virus genome with terminal untranslated regions.	8
Figure 1.1.6: The viral replication complex (RC).	9
Figure 1.1.7: The synthesis of viral RNA.	11
Figure 1.1.8: The NS1 crystal structure.	12
Figure 1.1.9: Topology model of NS4A based on DENV NS4A.	13
Figure 1.1.10: Topology model of NS4A based on DENV NS4B.	14
Figure 1.1.11: The NS5 crystal structure.	16
Figure 1.1.12: Cartoon representation of crystal structure of DENV-4 NS3 in two different conformations.	18
Figure 1.1.13: Crystal structure of the DENV-3 NS3 protease.	20
Figure 1.1.14: High resolution structure of DENV NS3 helicase.	21
Figure 1.2.1: Steps of ROS production.	22
Figure 1.2.2: Sites of ROS production.	23
Figure 1.2.3: OxyR and SoxRS system.	24
Figure 1.2.4: Anti-oxidant systems in human and <i>E. coli</i> .	25
Figure 1.2.5: Closed and open conformation of AhpF.	27
Figure 1.2.6: Dimer-decamer equilibrium of AhpC.	28
Figure 1.2.7: Conserved residues of the AhpC at the active site.	29
Figure 1.2.8: The mechanism of peroxide reduction by AhpC-AhpF complex in <i>E. coli</i> .	30
Figure 1.2.9: The <i>Ec</i> AhpC-AhpF ensemble in solution and the over-oxidation of C _P switching peroxidative function to chaperone function.	31
Figure 1.2.10: Anti-oxidant systems in catalase negative <i>E. faecalis</i> .	32

MATERIALS AND METHODS

Figure 2.2.1: Representation of overlap extension PCR used to generate the DENV-2 NS2B ₁₈ NS3.	41
Figure 2.2.2: Insertion of a stop codon into a plasmid using specifically designed forward (Fwd) and reverse (Rev) primers.	44
Figure 2.2.3: Schematic representation of flaviviral NS3 purification	57
Figure 2.2.4: Schematic representation of the reactions occurring in the continuous ATP hydrolysis assay.	58
Figure 2.2.5: Schematic representation of ligation dependent cloning.	64
Figure 2.2.5: Schematic representation of ligation independent cloning.	68
Figure 2.2.7: Stopped flow instrument.	73
Figure 2.2.8: Flow chart for solving the crystal structure of <i>Ef</i> AhpC ₁₋₁₇₂ .	76

RESULTS-STRUCTURAL AND FUNCTIONAL CHARACTERIZATION OF FLAVIVIRAL NS3

Figure 3.1.1: Amplification of gene encoding DENV-2 NS3 and recombinant DENV-2 NS2B ₁₈ NS3 production.	80
Figure 3.1.2: Purification of DENV-2 NS2B ₁₈ NS3.	81
Figure 3.1.3: SAXS studies of NS2B ₁₈ NS3.	82
Figure 3.1.4: <i>Ab initio</i> shape reconstruction of DENV-2 NS2B ₁₈ NS3.	83
Figure 3.1.5: Flexibility characterization of DENV-2 NS2B ₁₈ NS3.	85
Figure 3.1.6: Agarose gel showing amplified PCR-products.	86
Figure 3.1.7: Purification of DENV-2 NS3 single domains.	87
Figure 3.1.8: SAXS data analysis for DENV-2 single domains.	88
Figure 3.1.9: Distance distribution function of DENV-2 NS2B ₁₈ protease and helicase.	89
Figure 3.1.10: Solution envelopes of DENV-2 NS3 single domains.	90
Figure 3.1.11: Comparison of normalized Kratky-plot DENV-2 constructs.	90
Figure 3.1.12: Amplification of gene encoding DENV-2 NS2B ₁₈ NS3 mutant and recombinant DENV-2 NS2B ₁₈ NS3 production.	91

Figure 3.1.13: Purification of DENV-2 NS3 mutant.	91
Figure 3.1.14: SAXS studies of NS2B ₁₈ NS3 mutants.	92
Figure 3.1.15: Bar diagram representation of ATP hydrolyzing activity of DENV-2 NS2B ₁₈ NS3 and its helicase with linker.	93
Figure 3.2.1: Amplification of gene encoding ZIKV NS2B ₄₈ NS3 and recombinant ZIKV NS2B ₄₈ NS3 production.	94
Figure 3.2.2: Purification of ZIKV NS2B ₄₈ NS3.	95
Figure 3.2.3: Amplification of gene encoding ZIKV NS2B ₄₈ NS3 mutant and recombinant protein production.	95
Figure 3.2.4: Purification of ZIKV NS2B ₄₈ NS3 _{S135A} .	96
Figure 3.2.5: Amplification of gene encoding ZIKV NS2B ₁₈ NS3 mutant and recombinant protein production.	97
Figure 3.2.6: Purification of ZIKV NS2B ₁₈ NS3.	98
Figure 3.2.7: SAXS studies of ZIKV NS2B ₁₈ NS3.	99
Figure 3.2.8: <i>Ab initio</i> envelope reconstruction of ZIKV NS2B ₁₈ NS3.	101
Figure 3.2.9: Flexibility characterization of ZIKV NS2B ₁₈ NS3.	102
Figure 3.2.10: Agarose gel showing amplified PCR-products.	103
Figure 3.2.11: Recombinant ZIKV protease and helicase.	103
Figure 3.2.12: Purification of ZIKV protease.	104
Figure 3.2.13: Purification of ZIKV helicase.	105
Figure 3.2.14: SAXS studies of ZIKV NS2B ₁₈ NS3 single domains.	107
Figure 3.2.15: <i>Ab initio</i> envelope of ZIKV NS3 single domains.	108
Figure 3.2.16: Site-directed mutagenesis of ZIKV NS2B ₁₈ NS3.	109
Figure 3.2.17: Recombinant ZIKV NS3 mutant production.	110
Figure 3.2.18: Purification of ZIKV linker mutants.	111
Figure 3.2.19: Purification of ZIKV NS3 F179 mutants.	112
Figure 3.2.20: SAXS studies of ZIKV NS2B ₁₈ NS3 mutant.	113
Figure 3.2.22: ATP hydrolysis activity data of ZIKV proteins.	115

DISCUSSION -STRUCTURAL AND FUNCTIONAL CHARACTERIZATION OF FLAVIVIRAL NS3

Figure 4.1.1: Dengue NS3 adopts an extended conformation while HCV NS3 is compact.	120
Figure 4.1.2: Linker residues (₁₆₉ ERIGEPDYEVD ₁₇₉) in the DENV4 NS3 crystal structure.	121
Figure 4.2.1: The NS2B cofactor sequence alignment and crystal structure of ZIKV protease.	122
Figure 4.2.2: Interaction of ZIKV linker residues.	124

RESULTS-ATOMIC AND ENZYMATIC INSIGHTS INTO *EFAHPC*

Figure 5.1.1: Purification of <i>EfAhpC</i> .	129
Figure 5.1.2: <i>EfAhpC</i> is catalytically active.	130
Figure 5.1.3: <i>EfAhpC</i> forms high molecular weight oligomers.	131
Figure 5.1.4: Negative stained images of <i>EfAhpC</i> indicate that the protein forms a redox-independent decamer.	132
Figure 5.1.5: <i>EfAhpC</i> crystals.	133
Figure 5.2.1: Amplification of gene encoding <i>EfAhpC</i> ₁₋₁₇₂ mutant and recombinant <i>EfAhpC</i> ₁₋₁₇₂ production.	133
Figure 5.2.2: Purification of <i>EfAhpC</i> ₁₋₁₇₂ .	134
Figure 5.2.3: The C-terminal tail is crucial for peroxidative activity of <i>EfAhpC</i> .	135
Figure 5.2.4: <i>EfAhpC</i> ₁₋₁₇₂ like the full-length protein forms a decamer.	135
Figure 5.2.5: Picture of initial crystals of <i>EfAhpC</i> ₁₋₁₇₂ and diffraction pattern.	136
Figure 5.2.6: Optimization of <i>EfAhpC</i> ₁₋₁₇₂ by varying protein concentration.	137
Figure 5.2.7: Optimization of <i>EfAhpC</i> ₁₋₁₇₂ by varying protein concentration.	137
Figure 5.2.8: Optimization of <i>EfAhpC</i> ₁₋₁₇₂ crystal by overlaying paraffin/silicone oil.	137
Figure 5.2.9: Diffraction pattern of optimized <i>EfAhpC</i> ₁₋₁₇₂ .	138
Figure 5.2.10: Final optimized of <i>EfAhpC</i> ₁₋₁₇₂ crystals	139
Figure 5.2.11: The structural features of <i>EfAhpC</i> ₁₋₁₇₂ .	141
Figure 5.2.12: Structural intricacies of <i>EfAhpC</i> .	142
Figure 5.3.1: Amplification of gene encoding <i>EfAhpC</i> _{V78T} mutant and recombinant <i>EfAhpC</i> _{V78T} production.	145
Figure 5.3.2: Purification of <i>EfAhpC</i> _{V78T} .	145
Figure 5.3.3: The valine mutant has lower peroxidative activity than <i>EfAhpC</i> .	146

Figure 5.3.4: DLS data indicate that <i>EfAhpC</i> _{V78T} is unable to form stable decamer.	147
Figure 5.4.1: Site-directed mutagenesis of <i>EfAhpC</i> and respective recombinant protein production.	148
Figure 5.4.2: Purification of <i>EfAhpC</i> cysteine mutants.	149
Figure 5.4.3: The cysteine mutants show lower peroxidative activity than <i>EfAhpC</i> .	149
Figure 5.4.4: Destabilization of decamer in cysteine mutants is reiterated by DLS data.	150
Figure 5.5.1: Multiple sequence alignment of AhpC across different bacteria	151
Figure 5.5.2: Amplification of gene encoding N- and C-terminus mutants and recombinant <i>EfAhpC</i> _{V78T} production.	152
Figure 5.5.3: Purification of N- and C-terminal mutants.	152
Figure 5.5.4: N- and C- terminus mutants show reduced peroxidative activity.	153

DISCUSSION-ATOMIC AND ENZYMATIC INSIGHTS INTO *EFAHPC*

Figure 6.1.1: The connection of C _P loop dynamics and the oligomeric state of the AhpC.	157
Figure 6.1.2: The superposition of the <i>EfAhpC</i> and <i>EcAhpC</i> structure.	158
Figure 6.1.3: Comparison of the active site of AhpCs across various bacteria.	159
Figure 6.1.4: Two different conformations observed for the C _P loop at different chains in the crystal structure of <i>EfAhpC</i> ₁₋₁₇₂ .	160
Figure 6.1.5: Interaction of unique residues at the N- and C-terminus of <i>EfAhpC</i> .	164

LIST OF TABLES
MATERIALS AND METHODS

Table 2.1.1: Codon recognition in BL21-CodonPlus-RIL strains.	39
Table 2.2.1: Oligonucleotides synthesized, which corresponds to the sequence coding for 49 to 66 of DENV2 NS2B and Gly4-Ser-Gly4 linker	40
Table 2.2.2: Primers used for cloning DENV-2 NS2B ₁₈ NS3	41
Table 2.2.3: Components of PCR reaction mix for cloning the NS3 region of DENV-2 NS3 (Fragment2)	41
Table 2.2.4: Set-up of PCR program for cloning the NS3 region of DENV-2 NS3 (Fragment2)	42
Table 2.2.5: Amount of oligo 1 and 2 and fragment 2 mixed to obtain template to be used for overlap extension PCR	42
Table 2.2.6: Components of PCR reaction mix for overlap extension PCR	42
Table 2.2.7: PCR thermocycler program for overlap extension PCR	43
Table 2.2.8: Components of the reaction mix for setting up ligation	43
Table 2.2.9: Primers used site-directed mutagenesis to generate DENV-2 protease	44
Table 2.2.10: Components of PCR reaction mix	45
Table 2.2.11: PCR thermocycler program	45
Table 2.2.12: Reaction mix for DpnI digestion	45
Table 2.2.13: Primers used to generate DENV-2 helicase	46
Table 2.2.14: Components of PCR reaction mix for generating DENV-2 helicase	46
Table 2.2.15: PCR program set-up	46
Table 2.2.16: Primers for generating DENV-2 NS3 linker mutant	47
Table 2.2.17: Components of PCR reaction mix	48
Table 2.2.18: PCR thermocycler program	48
Table 2.2.19: Primers for site-directed mutagenesis	49
Table 2.2.20: Components of PCR mix	49
Table 2.2.21: PCR program	49
Table 2.2.22: Primers used to generate ZIKV protease	50
Table 2.2.23: Components of PCR mix	50
Table 2.2.24: Primers used to generate ZIKV helicase	51
Table 2.2.25: Components of PCR mix	51
Table 2.2.26: PCR thermocycler program	52
Table 2.2.27: Oligonucleotides synthesized which corresponds to the sequence coding for 49 to 66 of ZIKV NS2B and Gly4-Ser-Gly4 linker	52
Table 2.2.28: Components of PCR reaction mix for cloning NS3 region of ZIKV NS3 (Fragment 2)	53
Table 2.2.29: Set-up of PCR program for cloning NS3 region of ZIKV NS3 (Fragment 2)	53
Table 2.2.30: Primers used for overlap extension PCR for cloning ZIKV NS2B ₁₈ NS3	55
Table 2.2.31: Components of PCR reaction mix	65
Table 2.2.32: PCR thermocycler program	65
Table 2.2.33: CIAP treatment protocol	66
Table 2.2.34: Components of the reaction mix for setting up ligation	67
Table 2.2.35: Primers for amplification of <i>EfAhpC</i> ₁₋₁₇₂	67
Table 2.2.36: Primers for amplification of <i>EfAhpC</i> _{N2S,D10P,D14Q}	67
Table 2.2.37: Reaction mix for <i>DpnI</i> digestion	68
Table 2.2.38: Primers for amplification of <i>EfAhpC</i> mutants	69

RESULTS – STRUCTURAL AND FUNCTIONAL CHARACTERIZATION OF FLAVIVIRAL NS3

Table 3.1.1: Overall structural parameters for DENV-2 NS2B ₁₈ NS3	82
--	----

Table 3.1.2: Overall structural parameters of DENV-2 NS3 protease and helicase domains	89
Table 3.2.1: Overall structural parameters for ZIKV NS2B ₁₈ NS3	99
Table 3.2.2: Overall structural parameters of ZIKV NS3 protease and helicase domains	106
Table 3.2.3: Overall structural parameters of ZIKV NS2B ₁₈ NS3 _{F179A} and NS2B ₁₈ NS3 _{F179I}	114

RESULTS – ATOMIC AND ENZYMATIC INSIGHTS INTO *EFAHPC*

Table 5.1: Summary of in-house data collected from an <i>EfAhpC</i> ₁₋₁₇₂ crystal	138
Table 5.2: Summary of data collected at beamline 13B1 from an <i>EfAhpC</i> ₁₋₁₇₂ crystal.	139
Table 5.3: Summary of refinement statistics of <i>EfAhpC</i> ₁₋₁₇₂ crystal structure	106

Abbreviations

Å	Angstrom
ADP	Adenosine diphosphate
AhpC	Alkyl Hydroperoxide Reductase subunit C
AhpF	Alkyl Hydroperoxide Reductase subunit F
AhpR	Alkyl Hydroperoxide Reductase
ATP	Adenosine triphosphate
CTD	C-terminal domain
DENV	Dengue virus
DLS	Dynamic light scattering
DTT	Dithiothreitol
<i>E. coli</i>	<i>Escherichia coli</i>
EDTA	Ethylenediaminetetraacetic acid
<i>E. faecalis</i>	<i>Enterococcus faecalis</i>
IPTG	Isopropyl-β-D-thiogalactoside
kDa	kilo-Dalton
LB	Luria Bertoni
NAD	Nicotinamide adenine dinucleotide
NADH	Nicotinamide adenine dinucleotide reduced
NS	Non-structural
NSD	Normalized spatial discrepancy
NTA	Nitrilotriacetic acid
NTD	N-terminal domain
OD	Optical Density
PAGE	Polyacrylamide gel electrophoresis
PDB	Protein data bank
PEG	Polyethylene glycol
rmsd	Root mean square deviation
ROS	Reactive oxygen species
SDS	Sodium dodecyl sulphate
SEC	Size exclusion chromatography
<i>S. typhi</i>	<i>Salmonella typhimurium</i>
SAXS	Small angle X ray scattering

Tris	Tris-(hydroxymethyl) aminomethane
WT	Wild type
ZIKV	Zika virus

ABSTRACT

This thesis focuses on the non-structural protein 3 (NS3) of two flaviviral pathogens, the Dengue (DENV; DENV-2) and Zika virus (ZIKV), that have emerged as a major health concern in Singapore and world-wide. The NS3 comprises of an N-terminal serine protease domain and a C-terminal helicase domain which has nucleotide triphosphatase activities that are essential for RNA replication. NS3 is in constant motion and possess inherent flexibility. The inter- and intra-ensemble formation and crosstalk between the catalytically active helicase and protease require structural alterations and dynamic properties of the enzyme, allowing the catalytic centres to come in proximity during the enzymatic events.

The first chapter of this thesis describes the flexibility of DENV-2 NS3 in solution based on small angle X-ray scattering (SAXS) experiments. The data reveal that the overall structure of DENV-2 NS3 is extended and flexible in solution. Using the ensemble optimization method (EOM) it is demonstrated that NS3 is capable of existing in different conformations in solution. In this context, the importance of the linker residues in flexibility and domain–domain arrangement is shown by the compactness of the individual protease and helicase domains. A ₁₇₄PPAVP₁₇₉ linker stretch of the related Hepatitis C virus (HCV) NS3 was swapped into DENV-2 NS3 by mutagenesis to investigate the effect of this ‘PPxxP’ linker motif in the compaction of the molecule.

In addition, the overall low-resolution structure of the entire French Polynesia ZIKV NS3 in solution is presented for the first time. SAXS data of the entire ZIKV NS3 enabled rigid body modelling of the protein, which in turn sheds light on the domain-domain arrangement of the protein in solution. Solution studies of the individual protease and helicase domains show the compactness of the two domains as well as the contribution of the 10-residues linker region to the flexibility of NS3. Genetically engineered linker mutants of French Polynesia ZIKV NS3 enabled to identify linker residues being critical for the stability of the enzyme.

Enzymatic characterization of both DENV- and ZIKV NS3 as well as its helicase domain provide insight into contribution of the individual NS3 domains. Finally, the inhibitory effect of ATPase inhibitors on the enzymatically active DENV-2 and ZIKV NS3 are provided.

The second chapter of the thesis covers the structural and mechanistic aspects of the alkyl hydroperoxide reductase C of *Enterococcus faecalis* (EfAhpC) which together with subunit AhpF, a two-domain enzyme connected by a 41-residues linker, forms the so-called alkyl hydroperoxide reductase complex (AhpR), which is of paramount importance to restore redox homeostasis within the parasite. AhpCs are proposed to form enzymatically active homodimers and higher-oligomers. Such formations require flexibility and rearrangements of its C-terminal segment to come in proximity to the N-terminal segment of the second AhpC molecule, and to finally form a catalytic centre. In order to understand the dynamics and cross-talk of these enzymatic key segments of EfAhpC as well as the driving force(s) which modulates the equilibrium of dimer to higher oligomer formation, the first crystallographic structure of the *E. faecalis* AhpC (EfAhpC) was determined at 2.8 Å resolution, revealing a decamer-ring formed by five EfAhpC-dimers. The reported crystal structure provides insight into a transition state between a fully folded and locally unfolded conformation at the active site of the enzyme due to redox modulation. Amino acid substitutions of residues in the N- and C-termini as well as the oligomeric interphase of EfAhpC provide information into their structural and enzymatic roles. Mutagenesis, enzymatic and biophysical studies demonstrate the effect of the unusual existence of four cysteines in EfAhpC, which might optimize the functional adaptation of the *E. faecalis* enzyme under various physiological conditions.

1. Introduction

1.1 Overview of flavivirus

The flaviviruses belong to the *Flaviviridae* family. *Flaviviridae* derives its name from the Latin word *flavus*, which means yellow, for the hallmark jaundice (yellow discolouration of the skin) caused by infection of the yellow fever virus (YFV) [1]. The *Flaviviridae* family includes over 70 viruses and comprises of the genera *Flavivirus*, *Hepacivirus*, *Pestivirus* and *Pegivirus* [2]. These viruses contain a positive-sense, single-stranded RNA genome. Their genome is prone to high rates of mutation (up to 10^{-4} substitution per nucleotide), as the RNA polymerase lacks proofreading activity. Despite this, the evolution of flaviviruses is controlled by their need to replicate in vector or host [1]. Phylogenetic analyses have showed that two key branches: non-vector and vector borne emerged from a putative ancestor, and tick-borne and mosquito-borne clusters emerged from the vector-borne cluster [3]. The mosquito-borne cluster can further be divided into viruses borne by *Aedes* species such as YFV, dengue virus (DENV) and Zika virus (ZIKV), which can cause haemorrhagic disease syndromes, and viruses borne by *Culex* species such as Japanese encephalitis virus (JEV), West Nile Virus (WNV), Murray Valley encephalitis virus (MVEV), which cause encephalitic disease syndromes [1]. This thesis focuses on DENV and ZIKV, which are relevant in Singapore.

1.1.1 Epidemiology of DENV and ZIKV

Dengue is a viral haemorrhagic disease that has emerged a major concern in the tropical and sub-tropical regions of the world [4, 5]. There are four distinct serotypes of DENV, DENV 1-4 [5, 6]. Humans infected with DENV are usually asymptomatic [7], or they suffer from mild febrile phases that lead to a gamut of diseases that include asymptomatic dengue fever (DF), dengue shock syndrome (DSF) and dengue hemorrhagic fever (DHF) [8]. The incubation phase of the disease ranges from anywhere between 3 to 14 days, the first onset of symptom being sudden fever with headache and severe joint pain. The febrile phase of the disease is associated with high fever that lasts 2 to 14 days. The fever is biphasic: it breaks only to return within a few days. Eventually, the disease proceeds to a critical phase where there is leakage of plasma from blood vessels, accumulation of fluid in the chest and/or abdomen, and decreased blood supply to vital organs. Ultimately severe organ dysfunction might occur in infected individuals. The disease has become endemic in more than 125 countries in the world [4], affecting over 390 million humans till date [9]. According to the World Health Organization (WHO), more than 3.2 million cases across the Americas, South-East Asia and Western Pacific region were reported in 2015 (WHO fact sheet; 2017). WHO regions of Southeast Asia (SEA) [10] and

Western Pacific [11] combined bear almost 75% of the world-wide dengue burden [4]. In 2016, Western Pacific Region alone reported more than 37,5000 cases (WHO fact sheet; 2017). Most of the burden in this region comes from countries like Cambodia, Philippines, Vietnam, Malaysia and Singapore [4]. The number of dengue cases reported in Singapore in 2014, 2015 and 2016 are 18,335, 11,286 and 13,115, respectively [12]. The cumulative number of dengue cases for Singapore in 2017 was recorded as 2774 (NAE cases, online, accessed January 2, 2018).

ZIKV, closely related to the DENV and other flaviviruses, was declared a global emergency in 2016 [13]. Zika infection has been linked to foetal microcephaly and neurological disorders in adults. The most common disorders are Guillain-Barré syndrome, meningoencephalitis and acute myelitis [14-16] [16]. The first isolates of ZIKV came from a rhesus macaque monkey in Uganda in 1947. This African prototype was named MR-766 and was limited to the equatorial belt across Africa and Asia [17, 18]. Later in 1966, P6-740, the first non-African strain was isolated in Malaysia [17]. Till the early 2000s, only a handful of benign human ZIKV-related illness cases were recorded. The first largest outbreak of ZIKV outside the equatorial region occurred in 2007 on Yap Island [19]. The outbreak was characterized by cases with mild symptoms like fever, rash and/or conjunctivitis. In early to mid-2010s, infrequent manifestations of ZIKV infection were documented in Thailand, Cambodia, Malaysia, Indonesia and Philippines [19]. Widespread ZIKV infection ensued in French Polynesia where a small number of cases documented severe neurological disorders like the Guillain-Barré syndrome. Following the occurrence in French Polynesia, a large epidemic occurred in 2015 in Brazil. Increased number of new-borns with microcephaly resulted from the ZIKV infections in Brazil. Thereafter, the ZIKV has spread at an alarming rate all over the world and was also reported to circulate in the Pacific Islands like Fiji, American Samoa, Papua New Guinea, Palau, Islands of Cape Verde and South Asian countries like Vietnam, Thailand and Singapore [17]. The cumulative number of Zika cases for Singapore in 2017 was recorded as 67 (NAE cases, online, accessed January 2, 2018).

1.1.2 Flavivirus replicative cycle

Both DENV and ZIKV can infect cells from different tissues in various species and numerous cell lines including mosquito, human, monkey and mice cell lineages. This wide range of infectivity suggests that for mediating infection these viruses must bind to ubiquitous cell-surface receptors [20, 21]. In mosquito, both viruses replicate primarily in the midgut and salivary gland [22-25]. The replicative cycle of the virus begins with infection of the host cells

which include monocyte, macrophages and dendritic cells [26]. The attachment of the virus is the first step during this infection (Fig. 1.1.1). This is mediated by the envelope (E) proteins on the virion [26]. Host cell molecules like the DC-SIGN (dendritic-cell-specific ICAM3-grabbing non-integrin), CD14 (cluster of differentiation 14), HSP70/HSP90 (heat shock protein 70/90 kDa), mannose receptor, heparin sulfate, macrophage Fc γ receptor are anticipated to interact with the DENV E protein [20]. The proposed molecules for ZIKV interaction include DC-SIGN [21], TAM (Tyro, Axl, Mer) receptors and T cell immunoglobulin mucin domain (TIM) receptors [27]. After the attachment, the internalization of virus particles occurs by endocytosis via clathrin-coated vesicles [21, 28]. Upon entry, the virus fuses with the endosome, an event triggered by the low pH of the endosome (Fig. 1.1.1). The viral RNA is subsequently released. Due to the ability of the viral RNA to act as messenger RNA (mRNA), it remains in close association to the host rough endoplasmic reticulum (ER) where it is translated into a polyprotein (Fig. 1.1.1). Simultaneously there is induction of membrane structures in the host cell, which provide a membrane-bound niche for RNA replication and assembly of the virus particle. Processing of the polyprotein precedes viral genome replication and is mediated by the viral protease (NS3) or host proteases (*e.g.* signalase and furin) to initiate the release of individual proteins. After cleavage, the structural protein, capsid (C), and non-structural (NS) proteins NS3 and NS5, are released into the ER lumen, while other structural proteins, membrane precursor (prM) and envelope (E) proteins, and NS1, are released into cytoplasm [8, 29]. Besides the genomic RNA acting as mRNA, it also acts as a template for viral genome replication. The genome replication is orchestrated in the cytosol of infected cells. Once genome replication is completed, and the new RNA capped, it associates with the structural proteins (C, prM and E) to form the immature virus. In an immature virus, there are 90 heterodimers of E and prM, which extend as 60 trimeric spikes, giving the virus a rough appearance. Upon maturation, the cleavage of the 'pr' peptide is triggered by changes in the E protein conformation. This change in turn is initiated by acidic environment in the Trans-Golgi-Network (TGN) (Fig. 1.1.2A). On cleavage of the peptide by host furin protease, trimeric spikes lie flat against viral surfaces, imparting a smooth appearance to the virus (Fig. 1.1.2B-C) [8, 29, 30]. The mature virus is then secreted via exocytosis.

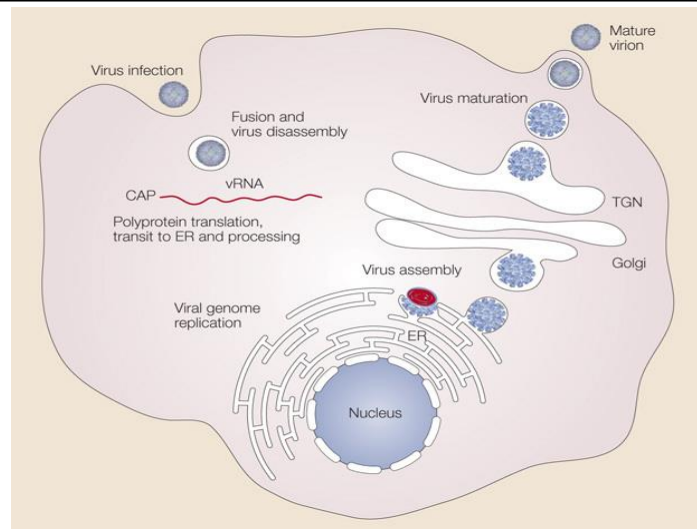


Figure 1.1.1: The viral replicative cycle. The viral infection commences with the attachment to the host, followed by fusion with endosome and release of the capped RNA. The polyprotein translated from the RNA is processed by host proteases and viral NS3. After replication of the RNA, it is packaged into viral particle, which undergoes a process of maturation as it passes through Golgi and TGN and is finally released into the environment (Figure was taken from [29]).

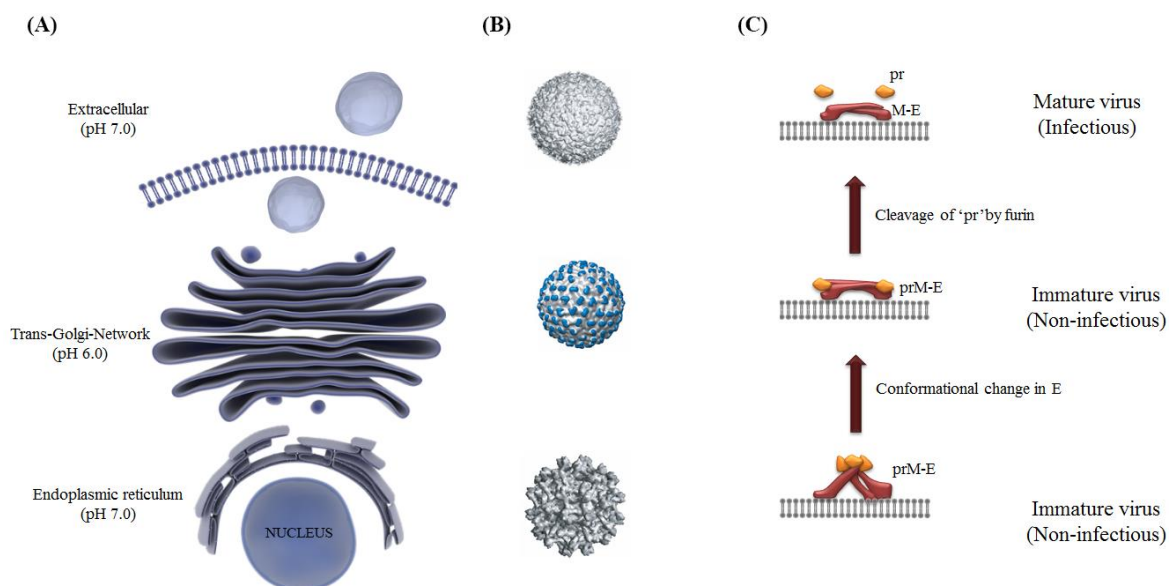


Figure 1.1.2: Flavivirus maturation pathway. (A) Immature virion bud into the endoplasmic reticulum (ER) and then passes on to the Trans-Golgi-Network (TGN). Here, the low pH induces conformational change in the protein 'E'. Furin cleavage of 'pr' from 'pr-M' occurs. Once the virus is released into the extracellular space, the 'pr' dissociates from the E protein. (B) Cryo-EM reconstruction of virus as it matures, transiting from a 'spiky' to a smooth appearance. (C) Pictorial depiction of the conformational changes occurring in 'E' and cleavage of 'pr' by furin as the immature virus particle passes from ER to TGN to the extracellular milieu where it is released as a mature virion (Figure modified from [8]).

1.1.3 Structure of the DENV and ZIKV particle

Cryo-electron microscopy (cryo-EM) studies of the mature DENV and ZIKV indicate that both viruses have a similar structure and morphology [30-32]. The mature viruses have an outer protein shell made up of 180 copies of a glycoprotein, the E protein, and a membrane protein, the M protein which are arranged in an icosahedral organization [20]. The E protein

(505 amino acids) is the major component of the virion envelope and is crucial for virus-host cell interactions like receptor binding during host cell invasion, membrane fusion and recognition of host immune system [30, 33]. Hidden beneath the E protein layer lies the M protein (75 amino acids). The icosahedral symmetry comprises of 60 repeating units of an E-M heterodimer, while the asymmetric unit consists of 3 individual E proteins [20, 30]. The E-M proteins are anchored to the underlying host-cell derived lipid bilayer (~4.5 nm thick) via their transmembrane domain. Multiple copies of the third structural protein of the virus, the capsid (C) protein are situated beneath the lipid bilayer. The nucleocapsid comprises of the C protein with the associated RNA genome. The C-terminal signal peptide of capsid protein is rooted into the host endoplasmic reticulum (ER). The C protein is a homodimer with an asymmetric charge distribution. The apolar residues interact with the lipid bilayer and the basic residues interact with genomic RNA [34].

Despite the overall morphology of the ZIKV being similar to the DENV (Fig. 1.1.3), the E protein of the ZIKV shows various dissimilarities to the DENV E protein. There are two sites of glycosylation for the E protein in DENV namely N67 and N153. Glycosylation of these residues ensure proper folding of the protein [20]. The glycosylation of this residue in DENV is also imperative for binding to DC-SIGN protein of the host cell for virus internalization. The N67 is mutated to aspartate in ZIKV and even though ZIKV interacts with DC-SIGN, they are postulated to have a different binding site [30, 33]. In fact, the N154 in ZIKV is thought to mediate the interaction with DC-SIGN [33]. The N154 glycosylation site has been implicated in neurovirulence, and this could explain the neural disorders associated with ZIKV. Apart from the differences in the sequence of E protein, the ZIKV was found to be thermally more stable than DENV. While the DENV E protein undergoes structural changes at 37° C, the virus retained its structural organization and infectivity even at 40° C [30]. Thus, the ZIKV is more adapted to survive harsh conditions and can be detected in [35], saliva [36] and semen [37] allowing sexual transmission of ZIKV.

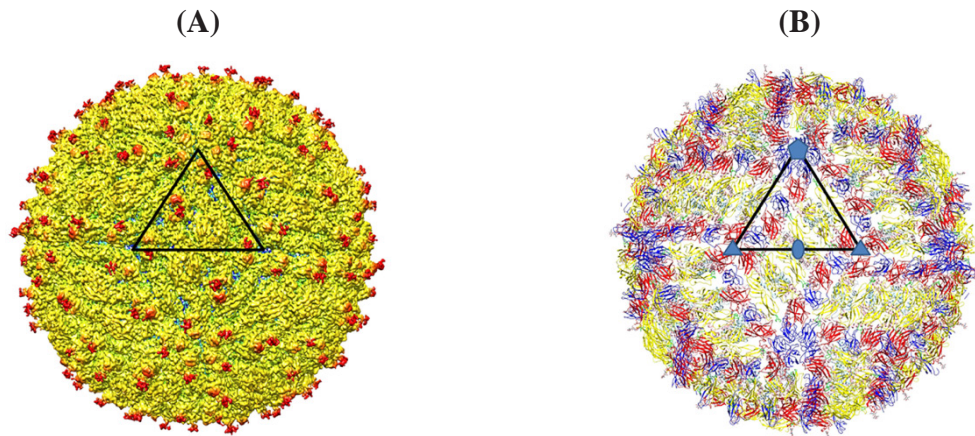


Figure 1.1.3: Cryo-EM envelope of ZIKV virus. (A) The ZIKV virus depicting the icosahedral symmetry. The black triangle represents the asymmetric unit. (B) The Ca backbone of the E and M proteins are shown. Orientation of the virus is same as in (A). The 3 domains of E protein (I, II and III) are depicted in red yellow and blue respectively. The figure was taken [31].

1.1.4 DENV and ZIKV genome assembly and replication

The RNA genome of flaviviruses is ~11 kb long. It comprises of a single open reading frame [38] which is flanked on either side by untranslated regions. The untranslated regions have certain secondary structures such as stem loops, and certain linear sequences that act as regulatory sequences during RNA replication [28]. Although the viral RNA is capped at the 5'-end (m7GppAmpN2), it lacks a poly (A) tail [28] (Fig. 1.1.4A). The single reading frame of the RNA genome is translated into a polyprotein precursor. This precursor encodes for three structural proteins C, prM and E, and seven non-structural proteins NS1, NS2A, NS2B, NS3, NS4A, NS4B, and NS5 [8]. This polyprotein precursor is processed during viral replication in the endoplasmic reticulum lumen via host proteases, including signalase and furin, and via viral proteases (NS3) at the cytoplasmic face [38] (Fig. 1.1.4B).

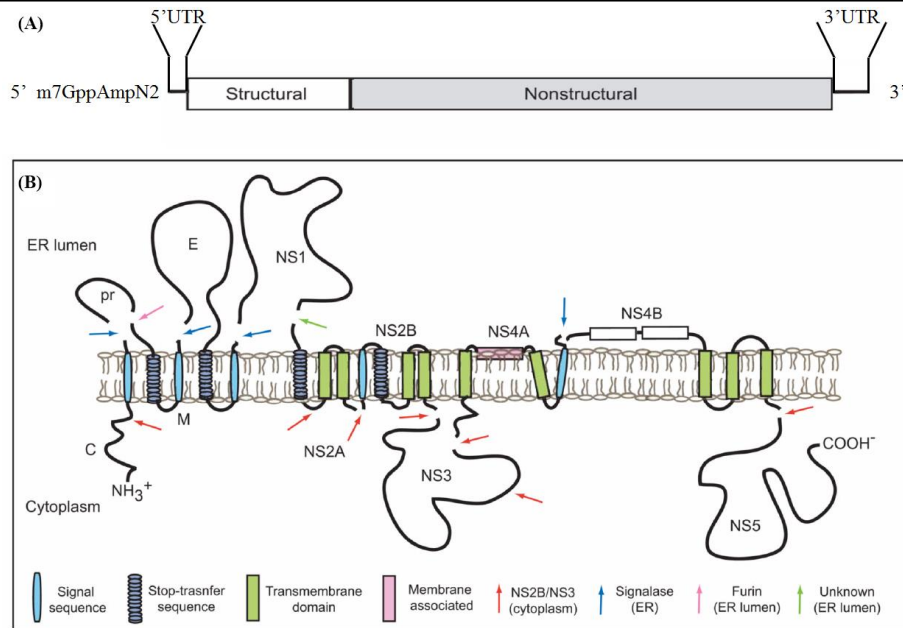


Figure 1.1.4: Viral genome assembly and polyprotein produced. (A) The open reading frame of the viral genome, the 5' cap (m7GppAmpN2) and the 5'-untranslated region (UTR) and 3'-UTR is shown in the figure. (B) The viral polyprotein that is translated from the open reading frame. The structural and nonstructural proteins of dengue virus are processed from the polyprotein by NS3 and host protease signalase (red and blue arrow respectively). The 'pr' protein is cleaved by furin (pink arrow) (The figure was taken from [8]).

Of the nonstructural proteins, NS3 and NS5 are well characterized. While NS3 is involved in the polyprotein processing, both NS5 and NS3 are involved in the viral replication process. They are multifunctional enzyme. NS3 has three activities: serine protease activity (together with NS2B), which is localized at the N-terminal domain, as well as helicase/NTPase- and RNA triphosphatase activity, which is localized at the C-terminal domain. The helicase/NTPase is required to unwind the double stranded RNA, formed during replication. The RNA triphosphatase activity is required for capping the RNA. Residing at the N-terminus, the S-adenosyl methyltransferase (MTase) domain of NS5 catalyzes the capping of RNA. At the C-terminus is the the RNA dependent RNA polymerase (RdRp) domain, which along with NS1 is essential for synthesis of RNA strand synthesis during replication [39].

While NS2B is crucial for the NS3 protease activity, the NS2A is needed for producing virus induced membrane during viral assembly. NS4A is a transmembrane protein like NS2A and NS2B and it induces the rearrangement of the membrane during viral replication. NS4B inhibits the immune response in host cells by inhibiting type 1 interferon response [26].

The genome replication is orchestrated in the cytosol. *Cis*-acting elements, *trans*-acting elements and viral induced membranes are necessary for the genome replication. The untranslated regions (UTR) on either side of the open reading frame constitute the *cis*-acting elements (Fig. 1.1.5). They act as promoters, allowing binding of the RNA polymerase. The 5'

UTR consists of stem loop A (SLA) and stem-loop B (SLB). Apart from the stem-loops, there are also linear sequences at the 5'-UTR that regulate the RNA replication. These sequences are: a 5'-upstream AUG-region (5'-UAR), a 5'-downstream AUG-region (5'-DAR), a 5'-cyclization sequence (5'-CS), and a C-coding region hairpin (cHP). At the 3'-UTR, after the stop codon, there is a variable region (VR) that controls viral growth and RNA synthesis. Further downstream are two secondary structures that modulate RNA synthesis, called A-2 and A3-domain. The latter is followed by a 3'-cyclization sequence (3'-CS), which along with the 5'-CS is involved in the circularization of the viral genome. Downstream of 3'-CS is a 3'SL structure. An 11-nucleotide sequence below the 3'SL is critical for RNA replication, and the stem-loop structure above the 3'SL is necessary for growth of the virus. Mutating or deleting either of the two regions severely compromises the genome replication and viral growth. CU_{OH}, the dinucleotide terminal of the dengue genome is involved in assembly of the replication complex (RC) (Fig. 1.1.6).

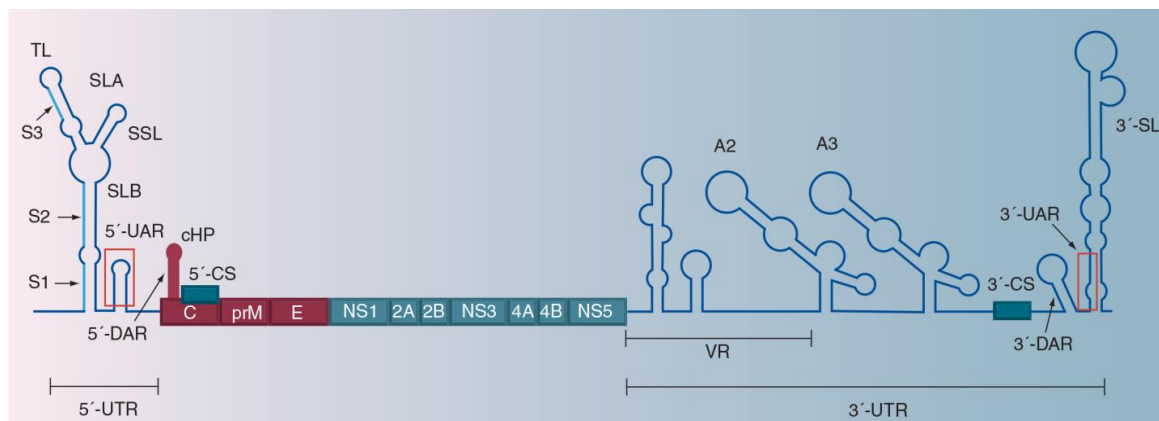


Figure 1.1.5: Dengue virus genome with the terminal untranslated regions. The untranslated regions (UTR) which act as *cis* elements in viral replication are shown. Stem loop structure SLA and SLB occur at the 5'-end. Linear regions at the 5'-end such as the 5'-upstream AUG region (5'-AUR), 5'-downstream AUG region (5'-DAR), 5'-cyclization sequences (5'-CS) and C-coding region hairpin (cHP) all play a role in genome replication. At the 3'-end are the variable region (VR) downstream of stop-codon, the A2- and A3-domain downstream of VR, 3'-stem loops (3'-SL) containing the 3'-UAR, and 3'- cyclization sequence 3'-CS (Figure was taken from [28]).

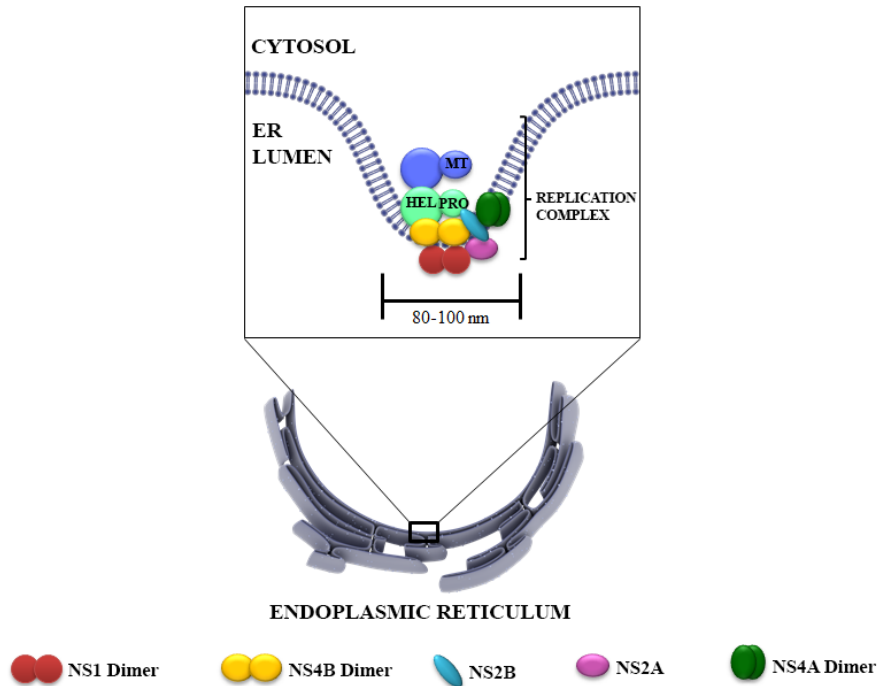


Figure 1.1.6: The viral replication complex (RC). The assembly of the non-structural proteins in the RC are shown. The figure shows dimerization of NS4B and oligomerization of NS4A. NS3 interacts with NS4B via the helicase domain (HEL) and with NS2B via the protease domain (PRO). NS5 consisting of the RNA dependent polymerase domain (POL) and MTase interact with the other nonstructural proteins in the RC via NS3 (modified from [40]).

The *trans*-acting factors include NS3 and NS5. They are accountable for all the enzymatic activity required for RNA replication. NS1, another *trans*-acting factor anchors the RC to the ER membrane. Several host proteins like the polypyrimidine tract binding protein (PTB), calreticulin, protein disulfide isomerase (PDI) also behave as *trans*-acting factor [28].

The viral induced membranes form a membrane enclosed environment at the cytosolic face of the ER where the viral nonstructural proteins, RNA and some host protein co-localize for genome replication.

The first step of RNA replication is the negative strand RNA synthesis [40]. The 5'UAR and 3'UAR have a complementary sequence and are involved in genome cyclization. The CS-sequence at the 5'- and 3'-end mediates the interaction between the 5'UAR and 3'UAR [28]. The circularized positive sense RNA is the template for negative strand by NS5 as it possesses the RNA-dependent polymerase domain [40]. The NS5 recognizes the SLA as the promoter; binds to it and initiates negative strand synthesis. The negative strand transiently remains associated with the positive sense strand as a double strand intermediate.

The NS3 helicase unwinds the double stranded intermediate in the next step where the positive strand RNA is synthesized. Following the unwinding by helicase, the RdRp domain will bind to the 3'-end of negative strand RNA to synthesize the positive sense RNA. The

double stranded intermediate is recycled for another round of positive sense RNA synthesis. The former positive sense strand is released as single strand of positive sense RNA [40].

The capping of positive sense RNA follows in the next step. The RNA triphosphatase activity of NS3 catalyzes the hydrolysis of 5'-terminal phosphate of the positive sense RNA. This 5'-diphosphate positive sense RNA then bind to the active site of the NS5 MTase for capping and methylation. The capping of RNA begins with MTase hydrolyzing GTP and forming a GMP-enzyme complex. GMP is then transferred to the 5'-diphosphate positive sense RNA. Once the cap is added, methylation occurs at at the N7-position of the guanine cap and ribose 2'-OH position of the first nucleotide. The capping process requires the presence of SLA at the 5'UTR, emphasizing that distinct viral length and sequences are important for this capping process. Since all separate reactions are carried out in one active site of the MTase, the 5'-end of RNA dissociates and re-associates with the NS5 at all steps [40] (Fig. 1.1.7). However, it is still unclear as to whether the 5'end of the RNA is capped while the nascent positive strand is being synthesized (co-transcriptional model) or the cap could be added as soon as the helicase unwinds the pre-existing full length positive strand RNA and subsequently hydrolyzes the 5'-terminal phosphate of the RNA (post-transcriptional model) [40].

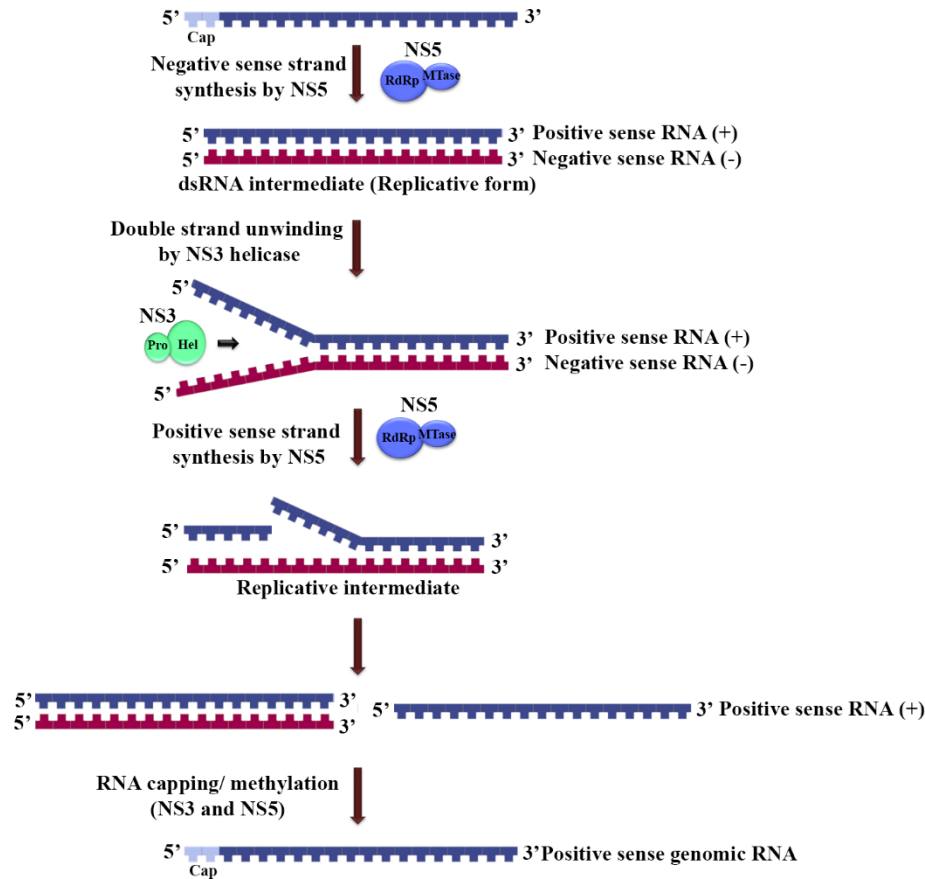


Figure 1.1.7: The synthesis of viral RNA. The negative strand is synthesized from the positive sense genomic RNA by NS5 resulting in a transient dsRNA which is the replicative form. The intermediate is then unwound by the helicase activity of NS3. Positive strand synthesis follows the unwinding, displacing the pre-existing positive strand as shown in the replicative intermediate. The displaced positive sense strand is then capped by the combined action of NS3 and NS5.

1.1.5 Non-structural proteins

1.1.5.1 Non-structural protein 1

NS1 is a 48 kDa glycoprotein. It is an antigenic marker and is used for screening DENV infection. It remains associated with lipids as a homodimer inside the cells and as a hexameric lipoprotein when it is secreted into the extracellular space. While the intracellular NS1 is crucial for the replication process, the secreted NS1 elicits immune response. The protein has three unique domains: a hydrophobic β -roll (1-29 amino acid residues), an α/β Wing domain (38-151 amino acid residues) and a central β -ladder (181-352 amino acid residues) [41] (Fig. 1.1.8). The protein contains two or three sites for glycosylation. N-linked glycosylation in DENV involves amino acids N130 and N207. Glycosylation of these residues are important for NS1 dimer stability and association with the ER membrane [42, 43]. NS1 can inactivate the complement system in host by activating Toll-like receptors (TLR) [12]. Currently NS1 has been identified as a marker in screening infections caused by flaviviruses and vaccines against NS1 are at various stages of development [41].

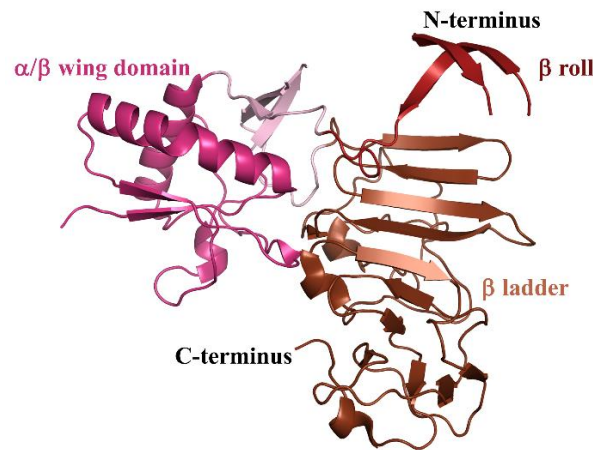


Figure 1.1.8: The NS1 crystal structure. The figure depicts a cartoon of crystal structure of DENV-2 NS1 (PDB ID: 4O6B; [44]). The figure depicts a monomer, and the ligand bound to the NS1 is not shown for clarity. The three domains of NS1 are distinguished by different colors. The N-terminal β -roll domain the α/β Wing domain and β -ladder domain are shown in red, magenta and brown, respectively.

1.1.5.2 Non-structural protein 2A

NS2A is a hydrophobic integral membrane protein of 22 kDa [45]. The N-terminal region of (1-68 amino acid residues) the protein is formed upon cleavage by unknown host proteases at the NS1/NS2A junction within the ER lumen, while the C-terminal region is generated in the cytosol by the NS3 protease [46]. At the N-terminus, residues 30-52 interact with the ER but does not span the lipid bilayer. The residues 69-209 form five trans-membrane segments that thread through the ER membrane. The terminal 210-218 residues form the C-terminus remain in the cytoplasm [47]. Co-localization of Kunjin virus (KUNV) NS2A with double stranded RNA (dsRNA) and its interaction with NS3, NS5 and 3' UTR indicate that NS2A is vital for synthesis of the genome [48]. This protein is also required for viral assembly [49] and DENV-2 NS2A alone has been implicated in abolishing interferon α/β response [50]. JEV NS2A inhibits dsRNA-activated protein kinase receptor (PKR) [51]. Recently it has been studied, that ZIKV NS2A inhibits glial cell proliferation and leads to premature differentiation of glial cells, thus causing neurological diseases in new born [52].

1.1.5.3 Non-structural protein 2B

NS2B is a 14 kDa protein, needed as a cofactor of NS3 protease, and is a membrane associated protein. The functional significance of this protein will be discussed with the NS3 part in section 1.1.7.

1.1.5.4 Non-structural protein 3

Since this protein is the focus of this thesis, it will be described in detail in section 1.1.7.

1.1.5.5 Non-structural protein 4A

NS4A is a hydrophobic protein of 16 kDa (127 residues) which is linked to NS4B by a signal peptide. This peptide contains 23 amino acids and is known as 2K, as it has a molecular weight of 2 kDa. It regulates the membrane modulating function of NS4A. The viral NS3 cleaves the signal peptide from the NS4A C-terminus [53]. This protein consists of two transmembrane domains [54] (Fig. 1.1.9). Amino acid residues 1-48 form an amphipathic helix and facilitate oligomerization of the NS4A [55]. In HCV, NS4A instead of NS2B is associated with NS3. Although NS4A and NS4B do not interact in WNV, it has been shown to interact via residues 40-76 with amino acids 84-146 of NS4B in DENV-2. This interaction correlates with the viral replication, indicating that interaction of both proteins is imperative for the DENV replicative cycle [56, 57]. The Akt-mTOR pathway crucial for the development of brain and regulation of autophagy is cooperatively suppressed by the ZIKV NS4A and NS4B [58].

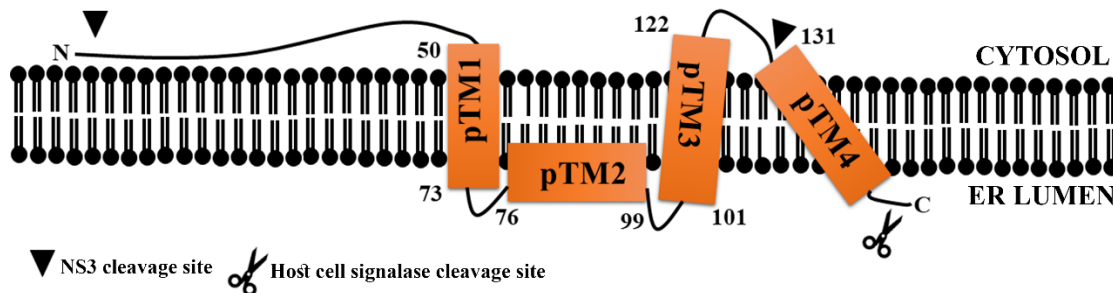


Figure 1.1.9: Topology model of NS4A based on DENV NS4A. The predicted arrangement of NS4A is shown. pTM1 and pTM2 are the two-transmembrane domains of NS4A. pTM1 spans the membrane from cytosol to lumen, whereas the pTM3 spans the membrane from the luminal side to cytosol. pTM2 is associated with the ER lumen. The pTM4 depicted is the 2K signal peptide. The sites of cleavage for the protein are shown. The figure has been modified from [54].

1.1.5.6 Non-structural protein 4B

The cleavage of 2K signal peptide from the N-terminus of NS4B is done by host signalases [50]. It is an integral transmembrane protein of 248 amino acids and contains three transmembrane domains (Fig. 1.1.10). Evidence suggests that the NS4B interacts with DENV helicase (domains II and III) and helps to unload the helicase from the single-stranded RNA (ssRNA)[59]. It interacts with NS4A and NS1 to regulate viral replication [57] .

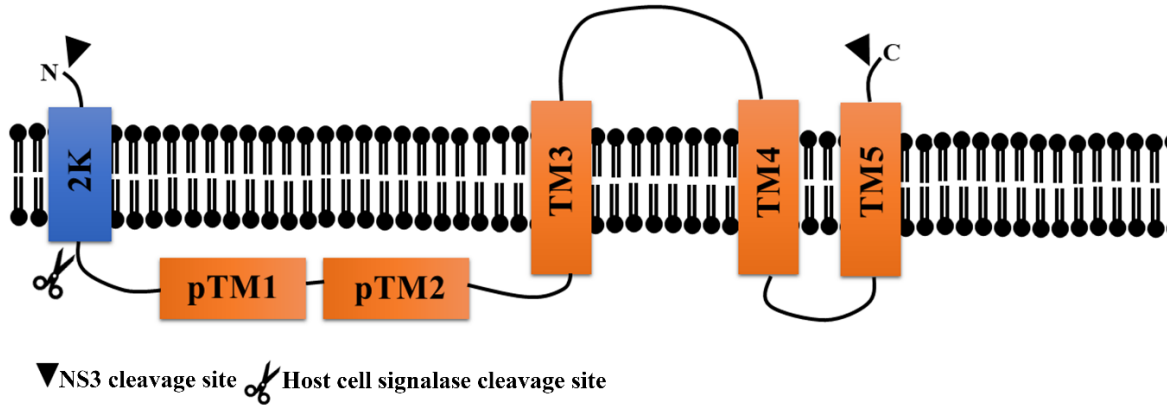


Figure 1.1.10: Topology model of NS4A based on DENV NS4B. The arrangement of the NS4B is revealed. TM3, TM4 and TM5 are the transmembrane domains of NS4B. pTM1 and pTM2 are located on the ER luminal side and do not span the ER membrane. The sites of cleavage for the protein are shown. The figure has been modified from [60].

1.1.5.7 Non-structural protein 5

NS5 is the most conserved protein amongst all other proteins of the virus [61]. NS5 is a two-domain enzyme playing a pivotal role in the RNA replication. At the N-terminus is the MTase. This domain has dual N7 and 2'-O methyltransferase activity. The MTase also provides the GTPase activity required for RNA capping. The structure of the MTase of DENV can be subdivided into three subdomains: an N-terminal GTP-binding subdomain (residues 7-54), a catalytic core subdomain (residues 55-222), and a C-terminal subdomain (residues 223-267) (Fig. 1.1.11). The N-terminal GTP-binding subdomain comprises the GTP-binding site, which provides GMP for cap formation. It is linked to the core subdomain via a positively charged protein stretch [62]. The core subdomain has the Ado-Met (SAM) binding site and the MTase catalytic site. The C-terminal subdomain's function is unclear.

The C-terminus of NS5 is an RNA-dependent RNA polymerase (RdRp) (Fig. 1.1.11), which first serves as a template for synthesis of negative sense RNA, which then serves as the template for positive strand synthesis [61]. As seen from the DENV-3 RdRp crystallographic structure (residues 265-900; PDB ID: 4CL1; [63, 64], the enzyme is structurally similar to other a U-shaped right-hand structure with three subdomains: fingers, thumb and palm [61, 62]. RdRp has a unique structure compared to other types of polymerases, where it has "fingertips" extending from the fingers subdomain that interacts with the thumb subdomain and encloses the active site. This interaction restricts the movement of these two subdomains and hence, no significant conformational change is observed upon binding of the substrates [65]. Six conserved motifs are identified within the DENV RdRp, and they are important for deoxy-nucleoside triphosphate (dNTP), RNA and metal ions binding, and catalysis [62]. The fingers and thumb subdomains are in charge of priming RNA and template binding, while the palm

subdomain play a central role in positioning the primer, divalent cations, templates and incoming nucleotides [66]. Together with the thumb subdomain, the fingers subdomain shapes a tunnel that directs the incoming RNA to the active site. The MTase and RdRp domains are physically linked by a linker suggesting that capping of RNA and viral replication are coupled. It has indeed been shown that the presence of the MTase has a stimulatory effect on the RdRp activity.

The determined crystallographic DENV-3 structure (PDB ID: 4V0R; [67]) revealed a compact structure with an overall dimension of 8.7 nm x 7.2 nm x 5.5 nm. In this structure the MTase domain was situated above the finger domains of RdRp and the GTP binding pocket, the catalytic triad and SAH binding pocket face away from the interdomain surface. The crystallographic structure has a fully resolved linker region ($_{263}\text{HVNAEPETPN}_{272}$), where the residues $_{263}\text{HVNA}_{266}$ form a short 3_{10} -helix. However, the low-resolution structure of DENV-3 NS5 obtained from small angle X-ray scattering (SAXS) experiments was extended with a maximum particle dimension of 11.5 nm, with the MTase domain situated above the RdRp domain [68]. SAXS data of all four DENV serotypes revealed that the overall dimension of NS5 from DENV 1-4 are similar, with DENV-4 resorting to a compact shape compared to the other serotypes [68]. SAXS studies of various mutants of DENV-4 NS5 suggested that the compactness of the protein was attributed to critical residues in the linker region namely K271, S266 and T267 [69].

As the viral replication occurs in host cytoplasm, NS5 is mainly located in the cytoplasm where it interacts with NS3. However, a high percentage of NS5 is localized in the nucleus of cells infected via DENV [70]. The nuclear NS5 is hyper-phosphorylated and does not interact with NS3. Two nuclear localization signals (NLS) were identified on DENV-2 NS5: β NLS (residues 320-368) and $\alpha\beta$ NLS (residues 369-405) [71]. The $\alpha\beta$ NLS is the determinant of nuclear localization and plays a role in viral RNA replication [71-73]. However, recently it has been revealed that the 18 residues at the C-terminus of NS5 is solely responsible for directing the NS5 either to the cytoplasm or the nucleus [74].

ZIKV NS5 shows 66% sequence homology with the DENV counterpart. The crystal structure for the individual domains of ZIKV NS5 are similar to other flavivirus single domains. The MTase adopts the classic $\alpha/\beta/\alpha$ fold and the crystal structure of ZIKV MTase can be overlapped with the MTases from other flaviviruses with a root mean square deviation (RMSD) of less than 0.73 Å. Like the DENV NS5, the presence of the MTase domain in ZIKV NS5 also stimulates the elongative RNA synthesis of RdRp domain [75]. The RdRp domain of NS5 adopts the characteristic polymerase motif, consisting of the thumb, palm and finger

domain. The domain arrangement in the full-length crystal structure of ZIKV NS5 from ZIKV strain MR766 (PDB ID: 5TMH and 5TFR; [76, 77], respectively) differs from DENV NS5. Superimposing the crystal structures of DENV and ZIKV NS5, an approximately 100° rotation of the MTase relative to the RdRp was observed in ZIKV NS5. The ZIKV MTase was situated at the back of RdRp while DENV MTase sits in the front of the RdRp domain (Wang et al. 2017). The domain arrangement in NS5 is similar to that described for JEV NS5. Unlike the DENV NS5, which exists as monomer, the ZIKV NS5 is dimeric in solution. This is shown by SAXS- and dynamic light scattering studies [78], and by the dimeric arrangement of the East African ZIKV (strain MR 766) NS5 inside the asymmetric unit (PDB ID: 5TFR; [76]).

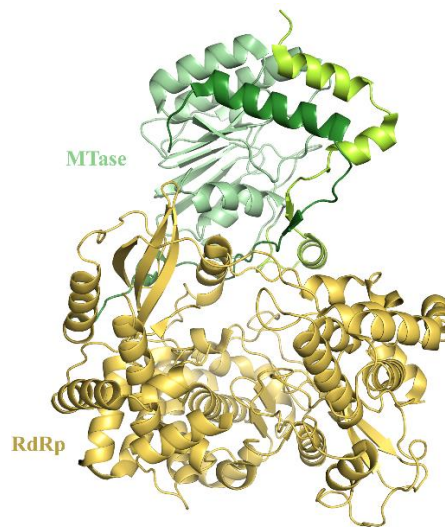


Figure 1.1.11: The NS5 crystal structure. The figure depicts a cartoon representation of structure of DENV-3 NS5 (PDB ID: 4V0R;[67]). The GTP-binding subdomain (light green), catalytic core subdomain (teal) and C-terminal subdomain (dark green) of the MTase is presented. The RdRp domain is colored in gold.

1.1.6 NS3 structure and function

The NS3 protein is an enzyme that is involved in a wide array of functions, encompassing polyprotein processing unwinding of double stranded RNA intermediate during genome replication and capping of viral genomic RNA. NS3 (618 amino acids) is the second largest protein in DENV with an approximate weight of 72 kDa. The enzyme is also the second highly conserved protein among flaviviruses [79]. The first 169 residues contain the chymotrypsin like serine protease that cleaves the polyprotein in *cis* and in *trans* orientation. For the protease to be functional, it requires another dengue viral protein, NS2B, as a cofactor. NS2B has two trans-membrane regions that are situated at the N- and C-terminus. In addition, it has a region of 47 amino acids that acts as a cofactor for the protease aiding proper folding of the domain. Without this hydrophilic segment of the NS2B, the NS3 protein is neither soluble nor catalytically active. The NS3 protease is linked to the helicase (residues 180-618) via a linker

which is 'acid rich' [80]. The linker region is not as conserved among flaviviruses as is the active sites for protease and helicase [81]. The flaviviral helicase located at the C-terminus, is an ATP-driven molecular motor [82]. It possesses RNA helicase, nucleoside and 5' RNA triphosphatase (RTPase) activities and can be divided into three subdomains. Although it has not been proven with certainty that the flaviviral helicases are involved in strand separation of the dsRNA intermediates [82], viruses with a faulty or impaired NS3 helicase do not replicate properly [83]. This indicates that the helicase is crucial for viral replication. The helicase is thought to be involved in resolution of secondary structures of the genomic RNA, and displacing protein cofactors during the replication process [83]. The RTPase activity initiates the viral RNA capping process by hydrolyzing the phosphoric anhydride bond of 5' RNA triphosphate [84]. The RTPase and ATP binding and hydrolysis site share the same active site [83, 84]. While both domains of NS3 can function independently, the presence of the protease in DENV promotes helicase activity in the full-length protein [85, 86]. On the contrary, the DENV-2 helicase alone has higher ATPase activity in comparison to the full-length protein [85].

The full-length NS3 protein from DENV-4 has been crystallized in two conformations [38, 81]. The wild-type (WT) DENV-4 NS3 crystallizes in the so-called first conformation, conformation I (PDB ID: 2VBC;[81]) (Fig. 1.1.12) which is an extended conformation. The protease domain is positioned at the entry of the ATPase active site amid the subdomains 1 and 2 of helicase. A linker mutant of the DENV-4 NS3 (insertion of glycine between E173 and P174) also adopts the conformation I (PDB ID: 2WZQ;[38]). An ADP-Mn²⁺ bound DENV-4 NS3 crystallizes in the second conformation, conformation II. This second conformation is an extended one, with the protease next to the ATPase catalytic site. However, the protease is rotated by 161° with respect to the helicase, which makes the ATP-binding site more accessible to solvents (PDB ID: 2WHX;[38]). This makes the second conformation of NS3 more active for both ATP hydrolysis and RNA-binding (Fig. 1.1.12). The domain-domain arrangement in ZIKV is still not clear as no high-resolution structure for the full-length ZIKV NS3 has been solved till date.

Besides its role in replication, NS3 can bind to the host cell nuclear receptor binding protein (NRBP) and undermine its role in ER-Golgi trafficking and is implicated in induction of membranous structure during viral replication [79]. NS3 is also involved in autophagy by interacting with the host cell proteins. Both DENV-2 and WNV NS3 have been shown to induce apoptosis in the host cells [87, 88]. In addition, studies on WNV NS3 revealed that NS3 is solely adequate to activate the caspase-8 activity that initiates apoptotic signaling [87, 89].

Furthermore, NS3 may work together with the host cell fatty acid biosynthetic pathway to form the replication complexes. NS3 was shown to interact with the fatty acid synthase (FASN) and localize FASN to the replication site [90]. This process requires a small GTPase, Rab18, which normally acts as a molecular switch in the regulation of membrane trafficking [91].

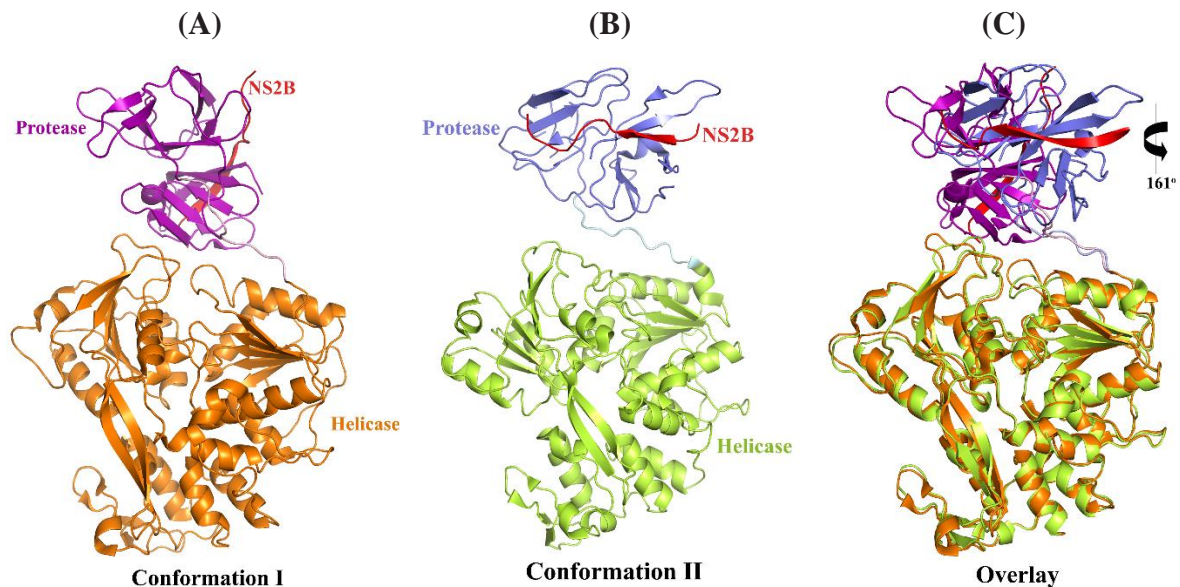


Figure 1.1.12: Cartoon representation of crystal structure of DENV-4 NS3 in two different conformations. (A) In conformation I of NS3 (PDB ID 2VBC;[81]), the protease is situated at the entrance of the ATP-binding site between subdomains 1 and 2 of the helicase. (B) In conformation II (PDB ID 2WHX; [92]), the protease is rotated by 161° with respect to the helicase. Conformation II was described to bind ADP-Mn²⁺ [92]. (C) Overlay of both structures depicting a rotated protease with respect to the helicase.

1.1.6.1 NS3 protease

The NS3 protease is crucial for viral infectivity. It has been shown that YFV was unable to infect target cells when genetic modifications were done to inactivate only the NS3 protease. Infectivity also reduced by 80% when a peptide inhibitor was used against DENV NS3 protease. The domain consists of a chymotrypsin like fold formed by two β -barrels (Fig. 1.1.13A), which are formed by six β -strands each. The cleft between the two β -barrels houses the conserved active site residues (H51, D75 and S135) [93] (Fig. 1.1.13A). The flaviviral NS3 protease necessitates the NS2B its folding. A difference of flaviviruses with hepaciviruses is, that NS4A is associated with the NS3 instead of NS2B [83]. NS2B has three membrane associated regions. The protein acts as a scaffold for the overall protease structure. NS2B can wrap around the protease domain, aiding in the formation of the protease active site [81]. There is a hydrophobic loop (29GLFG₃₂) in the NS3 protease domain that is oriented towards the hydrophilic domain (residues 49-96) of the NS2B that is anchored to the endoplasmic reticulum, facilitating NS3 protease and NS2B interaction. The presence of the N-terminal

hydrophilic region (residues 49-67) is adequate to solubilize the NS3 protease. However, the recombinant protein construct in DENV was enzymatically inactive. The hydrophilic residues 68-96 at the C-terminus of the NS2B form a β -hairpin (Fig. 1.1.13A). This contributes to building the substrate binding pocket of NS3 protease [83]. The sites of cleavage by viral protease include the NS2A-NS2B, NS2B-NS3, NS3-NS4A and NS4B-NS5 junction. The NS3 protease also cleaves upstream of the signal sequence at C-prM and NS4A-NS4B. The NS3 protease cleaves within NS2A and within NS3 itself [94] by recognizing the positively charged residues R/K at P1 and P2 position and short chain amino acid (G/S/A) at the P1' site of NS2A and at P2P1↓P1' within itself [95].

The expression of soluble and active recombinant NS3 protease was first achieved by linking 49-95 amino acid residues of WNV and DENV-2 NS2B to 1-169 residues of NS3 (PDB ID: 2FP7; [93] for WNV NS3 and PDB ID: 2FOM; [93] for DENV NS3). These residues were linked via a nonapeptide with the sequence G4-S-G4. The first NS3 protease crystal structure was obtained for DENV-2 in an open conformation, where the C-terminus of the NS2B cofactor was shifted away from the catalytic site of the protease in absence of an inhibitor [96]. NMR studies of an unlinked 50 amino acid residues stretch of NS2B and NS3 protease revealed that in absence of the physical linker the NS3 protease preferred the closed conformation [97]. Despite being unlinked, the recombinant NS3 protease was active [97] and the C-terminus of NS3 lined the protease active site. This closed conformation observed in DENV is similar to the closed conformation in WNV NS3 protease [98]. It is believed that the closed conformation of the protease is the best template for drug discovery [98]. Recently, the NS2B cofactor in ZIKV was linked to the NS3 protease by the terminal residues, ₁₂₇TGKR₁₃₀, from NS2B, and its respective atomic structure revealed also a closed conformation of the protease domain [13].

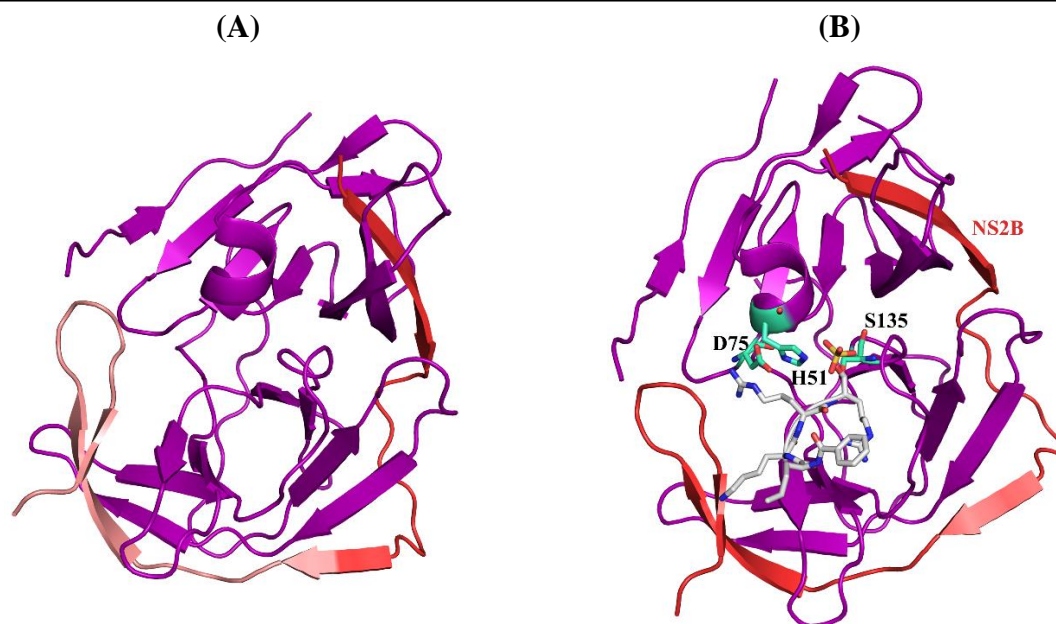


Figure 1.1.13: Crystal structure of the DENV-3 NS3 protease. (A) The cartoon representation of the crystal structure of DENV-4 NS2B/NS3 protease (PDB ID: 3U1I;[95]). The chymotrypsin like fold of NS3 protease is formed by two β -barrels that in turn is formed by six β -strands. The NS2B cofactor is shown in red. The hydrophilic region of NS2B that form a β -hairpin is indicated in a lighter shade of red. (B) The active site residues H51-D75-S135 are represented in sticks (teal). The peptide-based inhibitor, benzoyl-norleucine-K-R-R-aldehyde (sticks; grey) is shown to interact with the catalytic triad at the active site.

1.1.6.2 NS3 helicase

The NS3 helicase is a DEAH-RNA helicase, belonging to the superfamily 2 of helicases [85]. The helicase has three structural subdomains: domain I, II and III, each comprising of 150 amino acid residues. The domains are separated by series of clefts. Although domains I (181-326) and -II (327-481) show low sequence identity, both domains adopt a similar structural fold known as the RecA-like fold [85]. This fold is characterized by four α helices surrounding a central six-stranded parallel β -sheet [99] (Fig. 1.1.14). Domain III (482-618) is structurally different from the other two domains and is primarily composed of a bundle of four α helices surrounded by three shorter α helices and two anti-parallel β -strands (Fig. 1.1.14). This domain forms a single-stranded RNA binding channel. Seven conserved motifs have been identified in the NS3 helicase domain I and -II: motifs I, Ia, II-VI. Motif I, also known as the Walker A or P-loop (P for phosphate-loop), binds nucleotide phosphates, while motif II is involved in coordination with divalent metals such as Mg^{2+} or Mn^{2+} . Motif I and II are present in domain I. These two motifs along with motif VI (domain II) are involved in NTPase activity (specifically binding of ATP and ATP hydrolysis) of NS3 helicase ([100-102]. The NTPase active site is formed from K199 and T200 of the P-loop, residues D284 and E285 of motif II, and amino acids Q456 and R460 of motif VI. The K199 associates with the β -phosphate of NTP, while

D284 and E285 coordinates with divalent ions. The RecA-like domains are in parallel to the NTPase active site [85]. Mutagenesis studies reveal that flaviviral NS3 ATP-binding and hydrolysis active site is shared by the RNA 5'-triphosphatase active site. The RTPase activity is dependent on divalent metals [103]. The motifs Ia, IV and V are involved in binding of viral RNA. Motif III couples the ATP hydrolysis and RNA binding activity [81]. The coupling is brought about by concerted allosteric conformational changes [85].

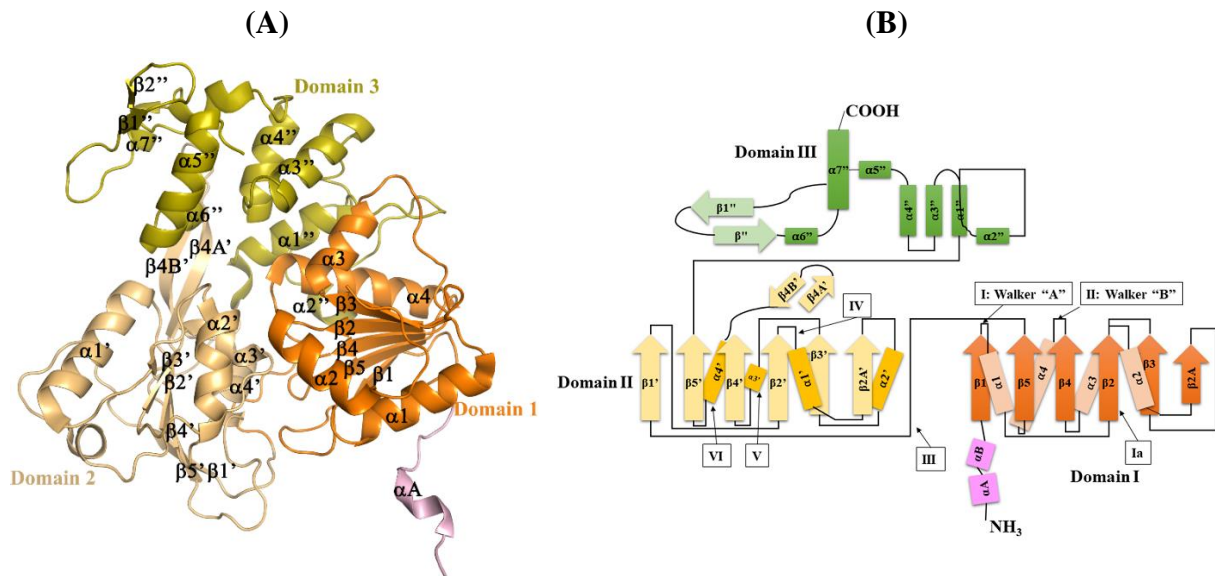


Figure 1.1.14: High resolution structure of DENV NS3 helicase. (A) Cartoon representation of the crystal structure of DENV-2 helicase (PDB ID: 2BMF; [85]). The helicase consists of three subdomains. Domains 1 (orange) and 2 (wheat) are crucial for RNA-binding and ATP hydrolysis. Domain 3 (olive) is essential for binding the single stranded RNA. The α helices and β sheets of each domain is numbered. (B) The topology of DENV NS3 helicases showing the overfolds of helicase. The secondary structural elements are numbered according to the ones shown in (A). The domains are colored according to the scheme in figure (A). The secondary elements are numbered according to the ones shown in (A), the β -sheets are colored in a lighter shade while the α -helices in a darker shade. The figure has been modified from [85].

1.2 Vancomycin resistant *Enterococcus faecalis* (V583) anti-oxidant system

1.2.1 Reactive oxygen species (ROS)

Life first evolved in the Archean Era around 3.8 billion years ago during which time the environment was anaerobic with high amounts of carbon dioxide in the atmosphere. With the emergence of the first oxygenic photosynthetic bacteria, the atmosphere gradually became saturated with oxygen. The oxygen being a reactive chemical created a crisis for the organisms and they had to devise methods to alleviate this stress. The anaerobes switched from anaerobic to aerobic metabolism. However, even the most aerobically active cell could not respire fast enough to keep the intracellular levels of oxygen below harmful levels. As a consequence of this switch to aerobic metabolism, reactive oxygen species (ROS) like superoxide (O_2^-), hydrogen peroxide (H_2O_2), hypochlorous acid ($HOCl$), singlet oxygen (1O_2), hydroxyl radicals

(HO \cdot), and lipid peroxides (ROOH) were generated in living organisms [104]. Of these, O $_2^{\cdot-}$, H $_2$ O $_2$, and HO \cdot are biologically relevant.

ROS are generated as a consequence of univalent addition of electrons to oxygen (Fig. 1.2.1). The electrons in the last shell of oxygen reside as unpaired, spin-aligned electron, allowing addition of only one electron at a time (Ames 1983). The sources of electrons for this step-wise reduction of oxygen are univalent electron donors such as flavins, transition metal ions and respiratory quinones.

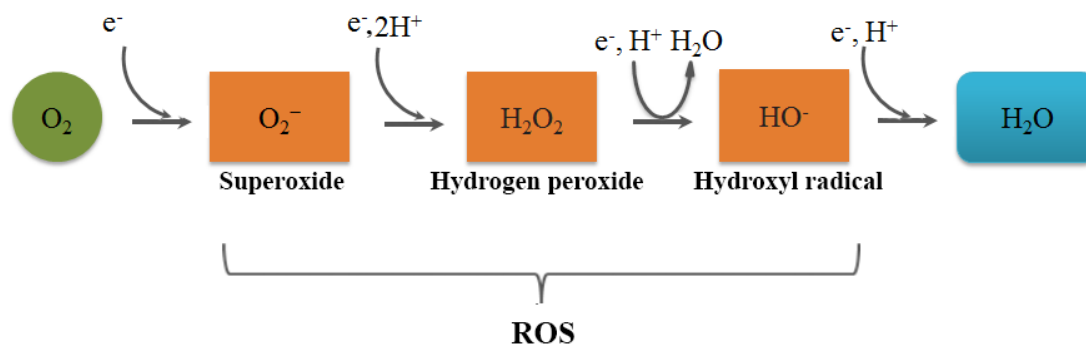


Figure 1.2.1: Steps of ROS production. The figure depicts a step-wise addition of electrons to molecular oxygen for generating ROS. The step-wise addition of electrons is made possible because of the fact that molecular oxygen has parallel unpaired electrons in their discrete orbital ensuring that electrons can be accepted one at a time (modified from [104]).

The endogenous sites in pro- and eukaryotes include the respiratory chain with its enzyme complexes I-V, where the NADH dehydrogenase (complex I) is the site for reduction of O $_2$ to O $_2^{\cdot-}$. When NAD $^+$ -dependent substrates are limited for complex I, electron flow occurs in reverse direction from complex II to complex I. This leads to over-production of ROS. Consequently, a H $^+$ -gradient is established across the inner mitochondrial membrane. This gradient dissipates through the F-ATP synthase (complex V). When ADP is unavailable, the movement of H $^+$ through the F-ATP synthase ceases and the flow of electrons is hindered, causing the respiratory chain to become more reduced. Under such condition, the generation of O $_2^{\cdot-}$ increases [105]. The Ubiquinone-cytochrome region of complex III generates O $_2^{\cdot-}$ from O $_2$ (Fig. 1.2.2). Several enzymes of the mitochondrial matrix, like aconitase, 1-galactono- γ lactone dehydrogenase (GAL), MnSOD (mitochondrial form of the Superoxide Dismutase), can also produce ROS [106].

Besides the NADH-, the NADPH-dependent electron transport involving CytP $_{450}$, produces O $_2^{\cdot-}$ in the endoplasmic reticulum of eukaryotes. Organic substrate, RH (where R is any alkyl group), reacts first with CytP $_{450}$ and is subsequently reduced by a flavoprotein to form a radical intermediate (CytP $_{450}$ -R \cdot). This oxygenated complex of this intermediate (CytP $_{450}$ -ROO \cdot) then becomes reduced by cytochrome *b*. Sometimes the

complexes may decompose releasing O_2^- . Apart from the afore-mentioned compartments in the eukaryotic cell, ROS can also be generated in the peroxisomes and at the plasma membrane [107]. Furthermore, in prokaryotes environmental factors such as ionizing radiation, redox-cycling compounds like menadione and paraquat can cause oxidative stress.

In plants, the electron transport chains in photosystem I (PSI) and PSII of chloroplast are the primary sources of ROS. Environmental stresses like drought, temperature and salt tend to increase the production of ROS. Also, conditions limiting CO_2 in plants increase the production of ROS in chloroplasts. Stress conditions in plants lead to decreased NADPH supply to chloroplasts and the respiratory chain is overloaded (Fig. 1.2.2) [106].

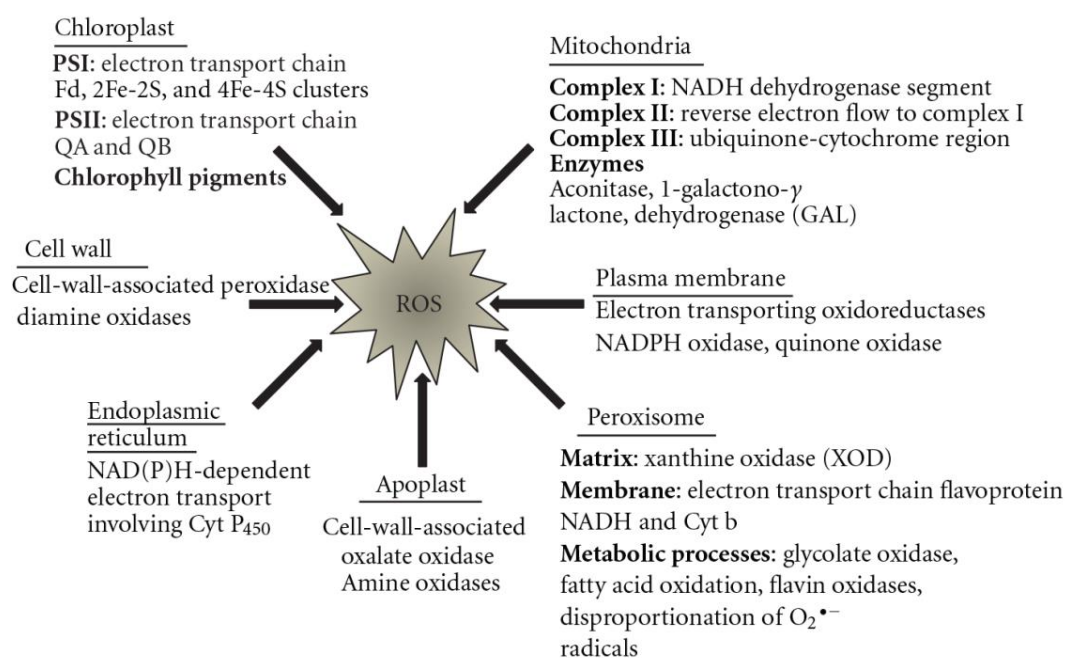


Figure 1.2.2: Sites of ROS production. Major endogenous sites for ROS production in eukaryotes include mitochondria, endoplasmic reticulum, and chloroplasts as well as the cell wall in plants. ROS is also generated at the plasma membrane and in peroxisomes (Figure modified from [106]).

1.2.2 Anti-oxidant systems

There are numerous enzyme systems that are responsible for combatting ROS. The soxRS (sox stand for Sry-related HMGbox) system is responsible for scavenging O_2^- . High concentrations of redox cycling compounds in the cell trigger the soxRS system. SoxR contains one sensory [2Fe-2S] cluster and is a homodimeric regulatory protein. On sensing redox active compounds, the [2Fe-2S] is oxidized, and soxR binds upstream of the *soxS* gene and induces the expression of *soxS*. The *soxS* protein then induces the expression of *sod* genes that encode the SODs (Fig.1.2.3). The latter neutralize superoxides by converting two molecules of superoxides into hydrogen peroxides and water [108]. The inability of the superoxides to

traverse cell membranes has led to the evolution of different types of SOD each residing in a specific cellular compartment [109]. All the various SODs contain metal centers which facilitate their reaction with O_2^- . Some examples are: iron- and manganese-co-factored enzymes (Fe-SOD and Mn-SOD) and copper–zinc-co-factored enzyme (Cu–Zn SOD) [104].

The OxyR-system is responsible for the regulation of H_2O_2 , which is slightly more complex than regulating O_2^- . The active site cysteine residues in OxyR react with H_2O_2 , resulting in a disulfide bonded form, that upregulate the expression of genes involved in combatting H_2O_2 (Fig. 1.2.3). The major enzymes scavenging H_2O_2 are peroxidases and catalases. Two catalases (KatE and KatG) and the alkyl hydroperoxide reductase system (AhpR) have been well studied in the model organism *Escherichia coli* [104]. The KatG belongs to the catalase-peroxidase family, whereas AhpR belongs to a class of thiol-peroxidases [104]. Although mammals have numerous peroxidases, they have only one catalase [110].

Besides these anti-oxidant enzymes, there are various non-enzymatic anti-oxidants, the most common being glutathione (GSH) and thioredoxin (Trx) [110]. Glutathione peroxidases (Gpx), a group of selenoproteins, are capable of reducing peroxides. Peroxidases that are dependent on the Trx system are known as peroxiredoxins (Prx) (Fig 1.2.4).

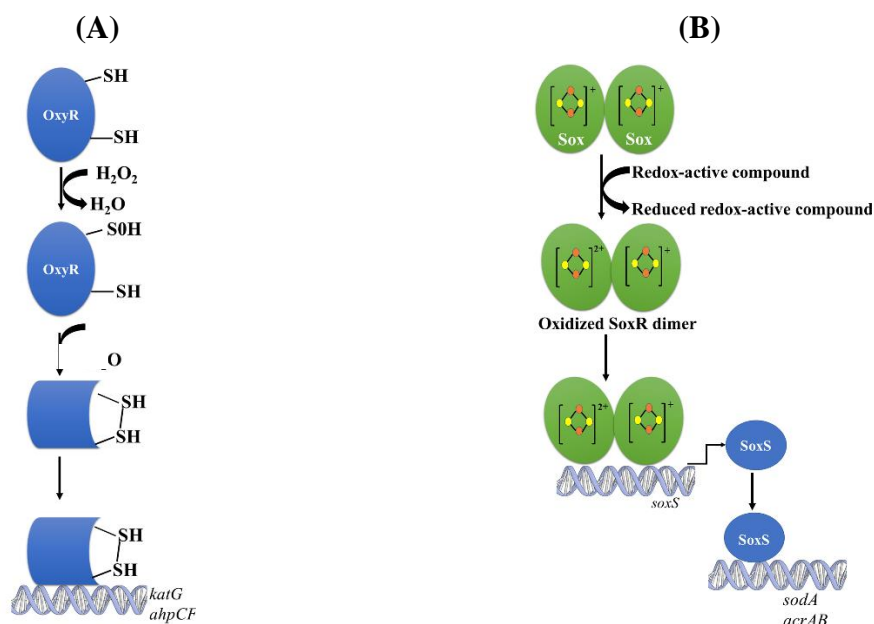


Figure 1.2.3: OxyR and SoxRS system. (A) When H_2O_2 reacts rapidly with a sensory residue in OxyR, there is activation of the OxyR-system. The resultant disulphide bond locks OxyR into a conformation that allows it to act as a positive transcriptional factor for *katG* and *ahpCF*. (B) SoxR is a [2Fe-2S] cluster containing homodimeric transcription factor. The protein is activated by direct oxidation of the cluster by redox active compounds. SoxR stimulates transcription of *soxS*, which in turn acts as a transcription factor for *sodA* and *acrAB* (Figure modified from [104]).

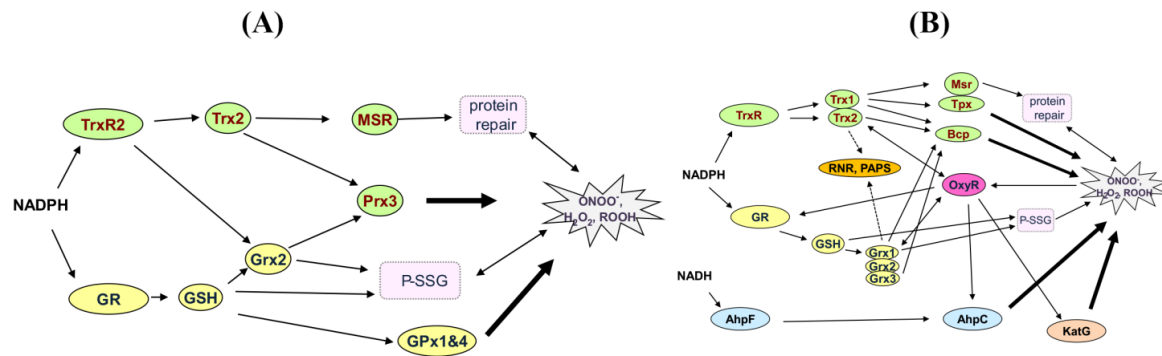


Figure 1.2.4: Anti-oxidant systems in human and *E. coli*. (A) In mitochondria of mammalian cells, thioredoxin (Trx) system is the main disulfide reductase system. Thioredoxin reductase (TrxR) reduces the Trx. TrxR2 is the mitochondrial isoform. Trx reduces methioninesulfoxide reductase (Msr) and is involved in repair of oxidized protein. The glutathione system serves as a back up to reduce thioredoxin in case electron transfer from TrxR is blocked. (B) In *E. coli*, the thioredoxin antioxidant system consists of the Trx1 and Trx2, which is reduced by the dedicated reductase TrxR. Once reduced, Trx can donate electrons to the three thiol peroxidases; thiol peroxidase (Tpx), bacterioferritin comigratory protein (BCP) and AhpC together with protein repair molecules like Msr and other molecules like Ribonuclease R (RNR), which is essential for DNA synthesis. Glutaredoxin (Grx) essentially possess similar functionalities as Trx. The dedicated peroxide quencher, AhpC is reduced by AhpF, which receive electrons from NADH (Figure taken from [111]).

1.2.3 The AhpR system

The AhpR complex is an essential peroxide inducible reaction to ROS in many bacteria [112] and is composed of two proteins: AhpF and AhpC. The AhpR ensemble is under the control of the OxyR H_2O_2 regulon. The gene encoding *ahpC* and *ahpF* are located at the same operon and transcription of both genes are regulated by OxyR, in response to stress. The *ahpF* *ahpC* can be co-transcribed as the *ahpF* lies upstream of *ahpC* [113]. The translation products of the genes are a 21 kDa protein (AhpC) and the 51 kDa flavoprotein (AhpF), respectively. The amino acid sequence of AhpF divulges that the protein is closely linked to the flavoprotein pyridine nucleotide and disulphide oxidoreductase family [114]. The AhpF bears 35% sequence identity with the *E. coli* TrxR. However, the N-terminal region of about 200 amino acid residues are absent in TrxR [115]. The AhpC has no amino acid sequence similarity to the Trx, instead they belong to the Prxs and are highly conserved among members of the animal kingdom [115]. The redox active cysteines of AhpF and AhpC are involved in reduction of H_2O_2 [113].

1.2.3.1 Alkyl hydroperoxide reductase subunit F

AhpF is a homo dimeric flavoprotein that consists an N-terminal domain (NTD) which contains a redox-active disulphide centre, an FAD-binding domain and an NADH/ redox-active disulphide domain [116, 117] (Fig. 1.2.5). The *Salmonella typhimurium* AhpF (*StAhpF*) crystal structure indicate that the NTD is a homologue of thioredoxin and protein disulfide

oxidoreductase (PDO), a thioredoxin-like (Trx-like) protein from *Pyrococcus furiosus* (*Pf*PDO). The NTD of *St*AhpF and *Pf*PDO both has two contiguous thioredoxin fold. While the *Pf*PDO has two redox active centers, one in each of the thioredoxin fold, the *St*AhpF has only one redox-active disulfide center. This redox active center (C129/C132) for *St*AhpF is present on the second thioredoxin fold and has a CXXC-motif [114, 118]. Limited proteolysis of *St*AhpF yielded two fragments, out of which the larger fragment of 37 kDa contained the C-terminal region consisting of the FAD and NADH/SS domains. This fragment can poorly reduce AhpC [114]. The high-resolution structure of the C-terminal region (212-521) of *E. coli* AhpF (*Ec*AhpF₂₁₂₋₅₂₁) correspond to the large fragment of *St*AhpF [119]. The redox-active center of this region (C345/C348) also has a CXXC-motif and remains closely associated with the isoalloxazine ring of the FAD [119]. Mutational studies of both redox active centers in *St*AhpF highlight that both centers are indispensable for AhpF activity [120]. The NTD of AhpF directly interacts with AhpC, reducing its redox-active disulfides [119, 121] and mediates the electron transfer from NADH via FAD to the C345/C348 reaction center and subsequently to AhpC. Additionally, it has been observed that generation of a chimeric protein of 1-207 residues of *St*AhpF linked to the full length TrxR from *E. coli*, conserved activity in AhpF/AhpC reductase assays [122]. Mutagenesis of C132S and C345S of the chimeric protein reiterated the role of both redox active centers in AhpC reductase activity [122]. This indicates that NTD of AhpF is important and plays a direct role in AhpC reduction [122]. Apart from the closed conformation of AhpF seen in *St*AhpF, an open conformation of Ahp was reported for the *E. coli* AhpF (*Ec*AhpF), where the NTD was rotated by 178° when compared to the *St*AhpF [116] (Fig. 1.2.5).

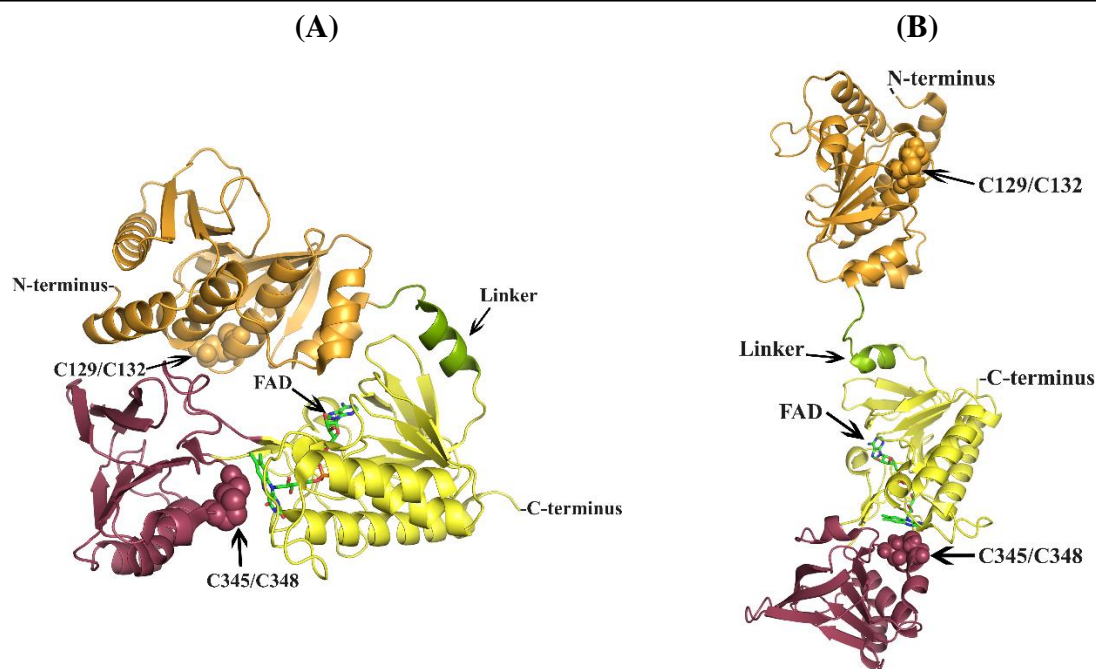


Figure 1.2.5: Closed and open conformation of AhpF. (A) Cartoon representation of crystal structure of *StAhpF* (PDB ID: 1HYU; [118]). AhpF binds FAD which is shown in sticks (green). This conformation is the closed conformation of the protein. The three domains: NTD (orange), FAD-binding domain (yellow) and NADH/ redox-active disulfide domain (brown) are demarcated in the cartoon. The linker is colored in green. (B) Cartoon representation of the high-resolution structure of *EcAhpF* (PDB ID: 4O5Q; [116]). AhpF binds FAD which is shown in sticks (green). This conformation represents the closed conformation of the protein. The different colors denote the three domains of AhpF as described for the *StAhpF*.

1.2.3.2 Alkyl hydroperoxide reductase subunit C

AhpC is the second subunit of AhpR and belongs to the Prx family. The latter are a class of thiol-based proteins that tightly regulate the levels of the intracellular peroxides [104]. They are ubiquitously present in all organisms from bacteria to humans. Six evolutionary subfamilies of Prxs have been identified: Prx1, Prx5, Prx6, Tpx, PrxQ and AhpE [123]. Besides their antioxidant role, these proteins act as modulators of H_2O_2 levels in the cells for signalling [124], transcription factors for controlling gene expression of other antioxidant enzymes [125], peroxinitrite reductases [126], molecular chaperones [127], foldases [128]. AhpC is a prototype of typical 2-Cys Prxs, which are the most abundant Prxs belonging to the subfamily Prx1. These proteins are extraordinary sensors of H_2O_2 and react approximately ten million times faster than free cysteine with peroxides to neutralize them [129]. They can reduce peroxide substrates with an efficiency in the order of $10^6 - 10^7 \text{ M}^{-1} \cdot \text{s}^{-1}$, and rely on the peroxidative (C_P) and the resolving (C_R) cysteines for catalysis [130]. Typical 2-Cys Prxs are obligate homodimers which form the active unit of the protein. In the dimeric arrangement, the N-terminus of one unit is juxtaposed to the C-terminus of the neighbouring dimer. This forms the dimeric interface and is held together by hydrogen bonds. These proteins can alter their quaternary structure from a dimer to a donut-shaped decamer [131] or even dodecamer [132] depending on the redox state,

and this lends mechanistic complexity to the AhpC [130]. Certain residues at the oligomeric interface (dimer-dimer interface) promote stable decamer formation as seen in case of *StAhpC*, where the mutation of T77 to valine stabilized the ring shaped decamer [133]. The modulation of the oligomeric state is also governed by several other factors which include protein phosphorylation [134], concentration of the protein [131], ionic strength [135] and pH [136] of the environment (Fig. 1.2.6).

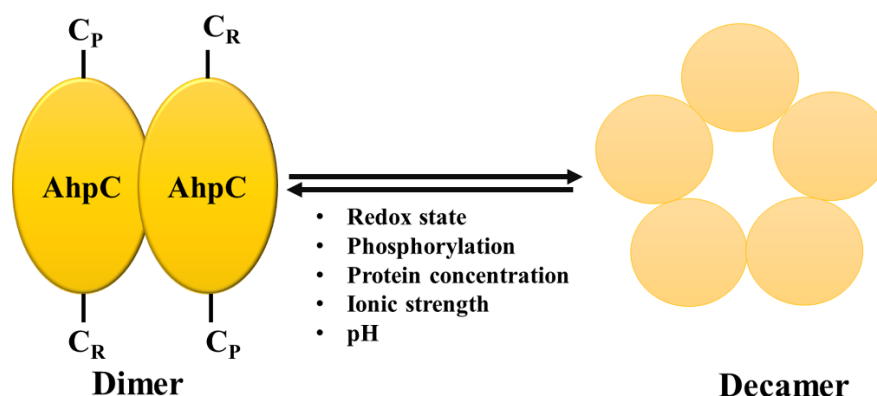


Figure 1.2.6: Dimer-decamer equilibrium of AhpC. The basic functional unit of AhpC is a dimer, where the N-terminus of one unit is in close proximity to the C-terminus of the other. The two conserved cysteines are designated as C_P and C_R and are present at the N- and C-terminus, respectively. The protein is capable of modulating its quaternary structure to form a ring shaped decamer. Each of the circle (yellow) of the decamer on the left represent a dimer. Thus, the ring is formed by a pentamer of dimers. The factors affecting the dimer-decamer equilibrium are listed in the diagram.

1.2.3.2.1 The active site of AhpC and its conformations

The AhpC contains two conserved cysteines: the peroxidative cysteine at the N-terminus and a resolving cysteine at the C-terminus. The C_P is present on the $\alpha 2$ -helix within a valosin-containing protein (VCP) motif, which is conserved within AhpCs (Fig. 1.2.7). The catalytic site is comprised of the P40, T44, and R120 (*E. coli* numbering). A negatively charged thiolate group is formed after deprotonation of the thiol group of the C_P which attacks the peroxide substrate. Residues T44 and R120 help to stabilize the thiolate anion for attacking the peroxide substrate, and converting it to generate water, with the formation of Cys-sulfenic acid at the active site (Fig. 1.2.7). Amino acid P40 protects the active site from water (Chuang et al., 2006). Besides the modulation of the quaternary structure, AhpC undergoes certain “conformational gymnastics” at the active site [130]. The two distinct conformations of the active site of AhpC are the fully folded state (FF) and the locally unfolded state (LU). The FF state is adopted in the reduced state, and this conformation facilitates substrate-binding. In this conformation, the C_P is placed at the bottom of the $\alpha 2$ -helix, and the C-terminal tail of AhpC is folds across the active site, separating the C_P and C_R by almost 14 Å. In the oxidized state, the AhpC active site adopts the LU state, where the $\alpha 2$ -helix is partially unwound, bringing the

C_P and C_R closer to mediate a disulfide bond formation. Mutation of C47 to serine in *StAhpC* resulted in a non-functional enzyme. Meanwhile, mutation of C166 to serine resulted in retainment 60% the peroxidative activity. This underlines the vital role of C47 in the peroxidatic activity [114]. The LU- and FF state also influence the oligomeric state of AhpC, with the FF state favoring the ring formation. Evidence has been shown in *StAhpC* that the decamer is crucial for effective catalytic activity of the protein [133]. Recent studies with *EcAhpC* indicated that in the oxidized state, the C-terminal tail is important for interacting with the AhpF, while in the reduced state, it stabilizes the decamer. Truncation of the C-terminal tail of *E. coli* leads to a loss of decamer formation and catalytic activity [112]. Thus, the C-terminus is vital for the domain-domain interaction in the alkyl hydroperoxide system.

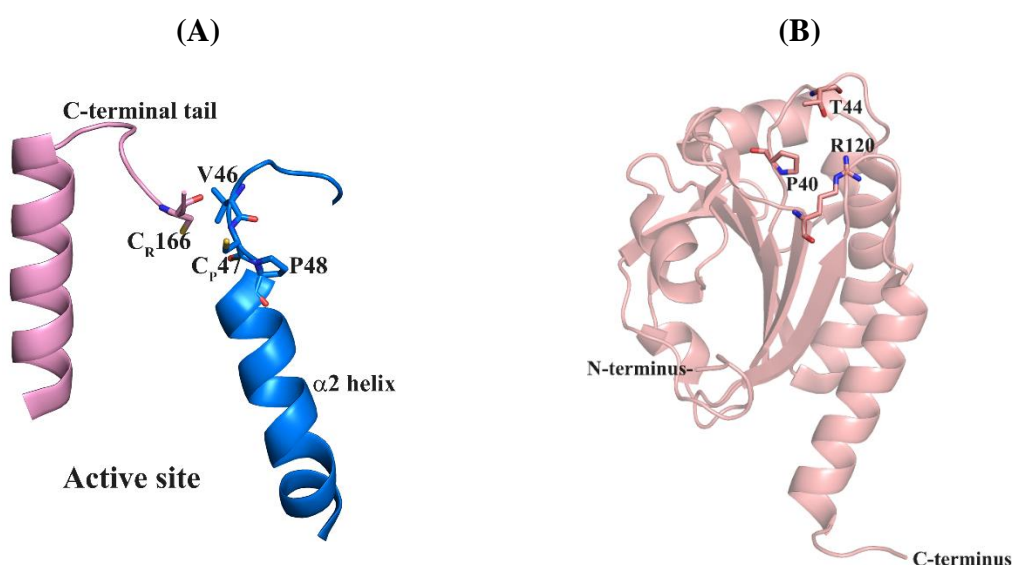


Figure 1.2.7: Conserved residues of the AhpC at the active site. (A) The FF state of *EcAhpC* showing the C_P and C_R at the active site of the protein. The C_P is present in the valosin-containing protein (VCP) motif, and the V46 and P48 of the motif is shown in sticks. For reasons of clarity only $\alpha 2$ -helix with its C_P , and C-terminus containing the C_R are shown. (B) Three other conserved residues at the active site are P40, T44 and R120. These residues are shown in sticks in the cartoon representation of crystal structure of *EcAhpC* (PDB ID:4O5R; [116]). The threonine and arginine stabilize the thiolate ion for attacking the peroxidative substrate by reducing the pKa at the active site. The proline shields the active site from water.

1.2.3.2.2 The catalytic cycle during peroxide reduction by AhpC

Currently, the best studied classical prokaryotic AhpR system is the *E. coli* AhpR ([112, 123, 137-140]. *EcAhpC* exists as an obligate homodimer and forming a ring of five dimers under reduced conditions. The C_P -loop of reduced *EcAhpC* acts as a redox sensitive molecular switch for oligomerization. During catalysis, the C_P at the N-terminus of one monomer is converted to sulfenic acid (C_P -SOH), the oxyacid of sulphur. With the formation of the C_P -SOH, the $\alpha 2$ -helix starts to unwind and shifts to the LU conformation. The LU conformation renders the C_P -SOH accessible for attack by the C_R , which is present at the C-terminus of the other monomer. This results in a disulfide bond formation between C_P and C_R .

[illegible]

The electron transfer pathway in AhpF is dictated by two factors. The first factor includes conformational changes that occur in the CTD and the second factor represents the movement of the NTD with respect to the CTD [137]. The movement of the domain is possible due to the linker that has inherent flexibility [140]. The conserved hydrophobic core of the AhpF NTD play a crucial role in the interaction of AhpF-AhpC [137]. In the oxidized state, the C-terminus of AhpC acts as an anchor for AhpC-AhpF assembly, wrapping around the NTD of AhpF, and slowing down the dissociation for effective transfer [140]. The C-terminus is extremely flexible under oxidized conditions and it binds the NTD of the AhpF, anchoring itself to the AhpF [116].

The electron transfer pathway in AhpF is dictated by two factors. The first factor includes conformational changes that occur in the CTD and the second factor represents the movement of the NTD with respect to the CTD [137]. The movement of the domain is possible due to the linker that has inherent flexibility [140]. The conserved hydrophobic core of the AhpF NTD play a crucial role in the interaction of AhpF-AhpC [137]. In the oxidized state, the C-terminus of AhpC acts as an anchor for AhpC-AhpF assembly, wrapping around the NTD of AhpF, and slowing down the dissociation for effective transfer [140]. The C-terminus is extremely flexible under oxidized conditions and it binds the NTD of the AhpF, anchoring itself to the AhpF [116].

Solution X-ray studies of the *E. coli* AhpC-AhpF reveal an ensemble of AhpC-AhpF in a stoichiometry of 10:2 (AhpC:AhpF) [138]. The dimeric AhpF molecule binds to the outer surface of the AhpC decameric ring (Fig. 1.2.9A). Such an arrangement brings the AhpF molecule in close proximity to the catalytic center at adjacent dimer or within a dimeric unit of AhpC [138].

In contrast to prokaryotic AhpC, the FF conformation persists on C_p-SOH formation in eukaryotic Prxs, which endorses the over-oxidation of C_p-SOH to C_p-SO₂H/ C_p-SO₃H. The over-oxidation inactivates the enzyme. An enzyme, sulfiredoxin, can reduce the C_p-SO₂H/ C_p-SO₃H, restoring peroxidative activity [141]. The GGLG-motif and the C-terminal helix containing the YF-motif in eukaryotic Prxs are responsible for this over-oxidation. A chimeric *EcAhpC* containing the YF-motif renders the bacterial AhpC sensitive to over-oxidation [123]. Under these circumstances there is a switching of peroxidase function of Prxs to molecular chaperone function [142]. The sulfinic acid intermediate forms high molecular weight species (Fig. 1.2.9B) with expression of molecular chaperone function and concurrent loss of peroxidase activity. Such switching of peroxidase activity to chaperone activity has also been observed in the case of *Pseudomonas putida* Prx, where the presence of an additional cysteine between the C_P and C_R promote the chaperone activity and compared to its counterpart from *Pseudomonas aeruginosa* which lacks the extra cysteine but share almost 93% homology [143].

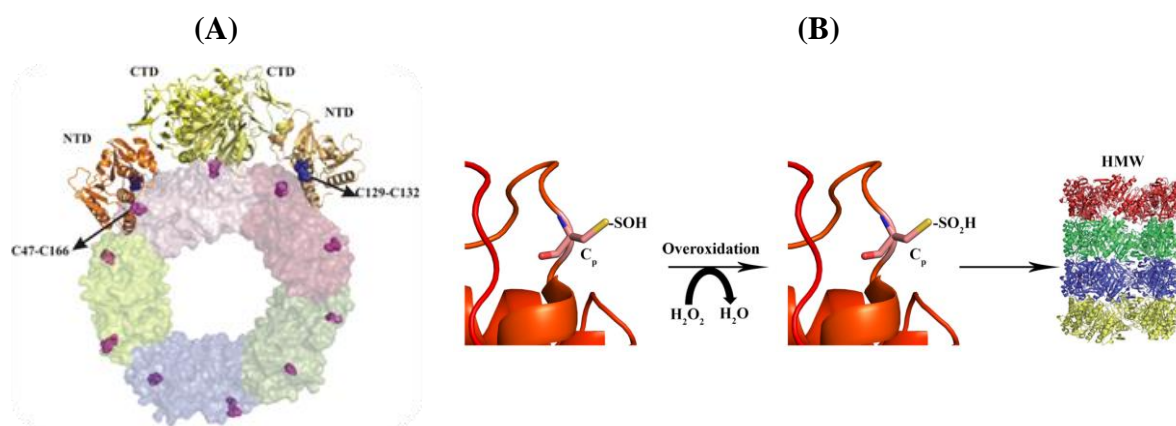


Figure 1.2.9: The *EcAhpC*-AhpF ensemble in solution and the over-oxidation of C_p switching peroxidative function to chaperone function. (A) The figure depicts how the dimer AhpF and the decameric AhpC interact in solution. It is postulated that AhpF and AhpC interact in a ratio of 2:10 as shown in the figure. Figure is taken from Kamriah et al., 2016. (B) The sulfinic acid upon over-oxidation forms sulfinic acid. The sulfinic acid state aids formation of high molecular weight (HMW) oligomer where the rings stack on top of each other as depicted (rings stacked shown in yellow, blue, green, red). Sulfinic acid formation causes a switch from peroxidative to molecular chaperone function.

1.2.4 The *E. faecalis* anti-oxidant system

While most prokaryotes use glutathione and/or catalases in their antioxidant systems (see figure 1.2.4 [111]), pathogenic bacteria like mycobacteria or *Enterococcus faecalis* use a mycothiol-dependent protein ensemble or a catalase negative system, as represented by *E. faecalis*, and therefore belonging to the class of catalase negative bacteria. Although the organism has the gene encoding catalase, the enzyme can only be active when heme is supplemented in the growing culture. Under such conditions, the peroxidases like AhpC and NADH-peroxidase become the focus of ROS scavenging (Fig 1.2.10).

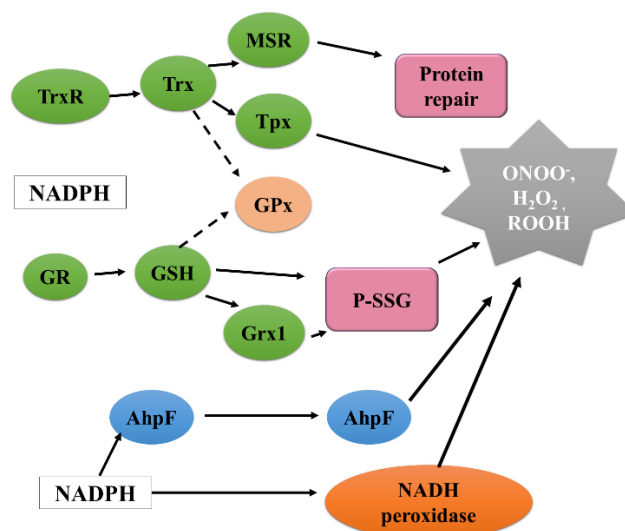


Figure 1.2.10: Anti-oxidant systems in catalase negative *E. faecalis*. The catalase enzyme is not functional in absence of heme. The major anti-oxidant systems are the three peroxidases AhpC, thiol peroxidase (Tpx) and NADH peroxidase. Msr are the virulence factors for the bacteria. The glutathione system (glutathione reductase (GR), glutathione (GSH) and glutaredoxin (Grx)) aid the thiol-dependent peroxidases in combatting ROS (Figure modified from [111]).

E. faecalis is an opportunistic gram-positive bacterium that, under the right conditions, becomes pathogenic and cause meningitis, endocarditis, urinary tract infections and bacteremia [144]. The bacterium is responsible for 90% of nosocomial infections [145]. Beside the fact that they are mainly found in the gastrointestinal track (GIT) of mammals and birds, they remain abundant in water and soil through fecal contamination [145]. It has been reported that *E. faecalis* is the leading cause of infection at surgical sites and bloodstream infections [146]. Compounding the pathogenicity of *E. faecalis* is the rapid development resistance towards antimicrobial agents, the most common being the Vancomycin-Resistant Enterococcus (VRE) [147, 148]. Because of robust genomic evolution, a quarter of the genome of *E. faecalis* is believed to be foreign and that most of the resistance genes are acquired via horizontal gene transfer [148]. Beside the acquisition of resistance genes, *E. faecalis* has evolved numerous virulence factors that ensure its survivability under harsh conditions of temperature, salinity, pH, etc. and during infection. Formation of biofilm, prolonged endurance during nutrient

deprivation, ability to alter host responses are few examples of virulence factors that promote pathogenicity [146]. Importantly, one of the major challenge of *E. faecalis*, and many other bacteria, during infection is oxidative stress they encounter in host systems [149, 150].

The ability of *E. faecalis* to survive the intolerant ROS condition of the macrophage implies the presence of robust antioxidant systems [151] that may serve as therapeutic targets for the prevention of at least the early stages of infections. Three reported peroxide scavengers in *E. faecalis* are the NADH peroxidase, AhpR and thiol peroxidase (Fig. 1.2.10) [152]. The NADH peroxidase and AhpR were found to be important for scavenging high amount of H₂O₂ which is endogenously generated from glycerol metabolism. The fast and complete degradation of the endogenously produced H₂O₂ necessitate not only the need for both enzymes [152], but most likely a need for a more robust catalytic activity. It is therefore important to gain insight into structures and unique features of these enzymes that confers their efficiency.

The *E. faecalis* *ahpC* and *ahpF* genes are regulated by HypR, a homologue of OxyR. *E. faecalis* (V583) AhpF (*EfAhpF*) is domain swapped when compared to other AhpFs [153]. This means that the typical N-terminal two-fold thioredoxin-like domain (NTD_N/C) is located at the C-terminus, whereas the pyridine nucleotide-disulfide oxidoreductase domain is encoded in the N-terminal part of its sequence. In *EfAhpF*, these two domains are connected via an unusually long linker region providing optimal communication between both domains. *EfAhpF* forms a dimer in solution similar to *EcAhpF* [153]. It was also recently discovered that the unique loop-helix- α 3 stretch (₄₀₉ILKDTPEPAKELLYGIEKM₄₂₆) is not present in homologue domains of other prokaryotic AhpFs. Deletion of the unique loop-helix part affects protein stability. Helix- α 3 is proposed to interact with the C-terminus of AhpC by providing efficient electron transfer for AhpC for the final reduction of H₂O₂. Because of the importance of the AhpR in antioxidant defence of the antibiotic-resistant strain *E. faecalis* (V583), the unique traits of its AhpF subunit [153], insight into varieties of *EfAhpC* in structure, mechanism and regulation are essential.

1.3 Goals of this thesis

NS3 is a two-domain enzyme connected by a 10/11- residues linker. To carry out the protease and helicase activity efficiently, the protein needs to undergo conformational changes. Recently, SAXS studies of NS5 indicated that the protein adopts multiple conformations in solution and that DENV-4 NS5 is compacted than other serotypes [68]. In this thesis, a systematic SAXS approach is utilized to investigate the DENV-2 and ZIKV NS3 and their

individual domains, to gain insights into the solution traits, compactness and flexibility of the protein. Till date no high- or low-resolution structure of the entire ZIKV NS3 exists. A soluble full-length NS3 was obtained, which enabled the investigation of domain-domain arrangement of NS3 and to construct a homology model of the full-length ZIKV NS3 based on the SAXS data. Mutants of both NS3s were genetically engineered, to elucidate the role of critical linker residues on the overall structural parameters of the protein. Enzymatic studies of both DENV-2 and ZIKV NS3 provide insight into the action of the potent ATPase inhibitors like resveratrol and quercetin.

E. faecalis is a gram-positive bacterium that despite lacking a functional catalase enzyme possesses a robust anti-oxidant system that allow the organism to survive intolerant ROS conditions of macrophages during infection [151]. The AhpR system, comprising of AhpC and AhpF, is a major player in combating ROS making it a valuable system to investigate for developing therapeutic targets for the prevention of infections. AhpC uses redox active cysteines to reduce H_2O_2 in an NADH-dependent pathway. The inter-subunit disulphide bond formed during catalysis is resolved by its interacting partner, AhpF. Since AhpC is involved in the direct catalysis of H_2O_2 reduction, gaining structural and mechanistic insights into this protein would help to understand its contribution towards the robust anti-oxidant system of the bacterium. To gain structural information about the *EfAhpC*, the recombinant full-length protein was purified, and attempts were made to crystallize the protein. A C-terminally truncated construct was generated to aid proper crystallization of the protein by removing the highly flexible tail. To determine the oligomeric state of the recombinant proteins, dynamic light scattering (DLS) and electron microscope was utilized. Mutagenesis of amino acid of polar and charged residues at N- and C-terminus, valine at the oligomeric interface and additional cysteine residues of *EfAhpC* was performed to elucidate the role of the unique residues in the protein. Also, NADH-dependent peroxidative activity was performed for the WT enzyme and its mutants, to gain knowledge regarding the effect of mutation of the said residues on the catalytic activity of the protein.

2. Material and Methods

2.1 Materials

2.1.1 Chemicals

All the chemicals used for these studies were of at least analytical grade. Chemicals used for media preparation and buffer preparation for protein purification in this study were purchased from Sigma Aldrich (St. Louis, MO, USA), Merck (Darmstadt, Germany), Gibco (Singapore) and VWR (Darmstadt, Germany). Electrophoresis chemicals (agarose, SDS, glycine, APS etc.) were purchased from Bio-Rad (Hercules, CA, USA) and VWR (Darmstadt, Germany). The following items were purchased from:

Antibiotics	Calbiochem, Sigma and Gibco (Invitrogen)
BSA	GERBU (Heidelberg, Germany)
IPTG	Fermentas (Vilnius, Lithuania)
Imidazole	Sigma Aldrich (St. Louis, MO, USA)
Pefabloc ^{SC}	BIOMOL (Hamburg, Germany)
Thrombin	Sigma Aldrich (St. Louis, MO, USA)

2.1.2 Molecular biology materials

2.1.2.1 Enzymes and molecular biology kits

<i>DpnI</i>	New England Biolabs (Frankfurt, Germany)
KAPA Hi-Fi DNA Polymerase	Roche (Mannheim)
Nucleobond AX mediprep Kit	MN & Co (Düren, Germany)
Pfu DNA Polymerase	Fermentas (Glen Burnie, MD, USA)
QIAquick PCR purification kit	Qiagen (Hilden, Germany)
QIAquick gel extraction kit	Qiagen (Hilden, Germany)
QIAprep Spin miniprep plasmid kit	Qiagen (Hilden, Germany)
Restriction enzymes	Fermentas and New England Biolabs
T4 DNA Ligase	New England Biolabs

2.1.2.2 Template for generating gene constructs

The plasmid DNA for amplifying the gene encoding *E. faecalis* AhpC was kindly provided by Prof. Greg Cook (University of Otago, New Zealand). The cDNA for amplification of genes encoding DENV-2 NS3 and its single domains were provided by Prof. Liu Ding Xiang (School of Biological Sciences, Nanyang Technological University). The plasmid for cloning the ZIKV NS3 and its gene constructs was synthesized from Integrated DNA technologies (IDT).

2.1.2.3 Primers and synthetic oligonucleotides

Primers for generating the gene constructs for both *E. faecalis* AhpC and flaviviral NS3 were synthesized by IDT. The oligos 1 and 2 for the overlap extension polymerase chain reaction (PCR) were also synthesized by IDT.

2.1.2.4 Plasmids

All the *E. faecalis* AhpC constructs were cloned into a modified pET-9d vector [154]. The plasmid allows the expression of His-tagged protein at the N-terminus and also confers resistance towards kanamycin to facilitate the selection of positive clones.

All the flaviviral NS3 constructs were cloned into a modified pET-32b plasmid [81] for gene expression. The plasmid was used for the expression of thioredoxin-His₆ fused protein, with a thrombin cleavage site enabling thioredoxin-His₆ tag cleavage. This vector confers ampicillin resistance.

2.1.2.5 *Escherichia coli* strains

Electrocompetent DH5 α cell strain were used for all the above cloning processes. The genotype of the strain is F⁻, ϕ 80*lacZ* Δ M15, Δ (*lacZYA-argF*), U169, *recA1*, *endA1*, *hsdR17*(rK⁻, mK⁺), *gal*⁻, *phoA*, *supE44*, λ -*thi* -1 *gyrA96*, *relA1*. The *recA1* mutation inactivates the recombinases and prevents homologous recombination, while the *endA1* mutation inactivates the endonucleases, thus preventing degradation of the transformed plasmid.

Electrocompetent BL21 DE3 cell strain were used for the expression of *E. faecalis* AhpC and its mutants. The genotype of this strain was: F⁻, *ompT*, *hsdS_B* (rB⁻, mB⁻), *gal*, *dcm*, *lon* (DE3). These strains utilize the T7 RNA polymerase promoter to facilitate the overexpression of the desired gene product. These strains lack the Lon and OmpT proteases which would ensure that the recombinantly produced proteins are not degraded.

Electrocompetent BL21-CodonPlus-RIL strains were used for the expression of DENV and ZIKV constructs. Genotype: F⁻, *ompT*, *hsdS*(rB⁻, mB⁻), *dcm*⁺, Tet^r, *gal*, *endA*, Hte [*argU*, *ileY*, *leuW*, Cam^r]. These cells contain extra copies of the *argU*, *ileY*, and *leuW* tRNA genes. These genes encode tRNAs that recognize specific codons (Table 2.1.1). This helps to confine translation of heterologous proteins from organisms that have AT-rich genomes. In addition, these strains additionally confer chloramphenicol (Cam^r) resistance. The utilization of this strain provides an advantage in which there is tighter regulation of gene expression.

Strain	tRNA genes (codon recognition of gene product)	Antibiotic resistance
BL21-CodonPlus-RIL strain	<i>argU</i> (AGA, AGG), <i>ileY</i> (AUA), <i>leuW</i> (CUA)	Cam

Table 2.1.1: Codon recognition in BL21-CodonPlus-RIL strains.

2.1.3 Chromatography materials

Immobilized affinity chromatography

Ni²⁺-NTA agarose QIAGEN (Hilden, Germany)

Gel filtration

Superdex 75 HR (10/30) GE Healthcare (Uppsala, Sweden)

Superdex 200 HR (10/30) GE Healthcare (Uppsala, Sweden)

Superdex 200 HR Increase (10/30) GE Healthcare (Uppsala, Sweden)

2.1.4 Instruments

ÄKTA FPLC GE Healthcare (Uppsala, Sweden)

BioSpec-Nano Shimadzu, Tokyo, Japan

MetalJet-NANOSTAR Bruker AXS (Karlsruhe, Germany)

Micropulser Electroporator Bio-Rad (Hercules, CA, USA)

PCR Thermocycler:

 Biometra T personal Biometra (Göttingen, Germany)

 Biometra T gradient Biometra (Göttingen, Germany)

Sonoplus Sonicator Bendelin (Berlin, Germany)

Ultraspec 2100 Pro-Spectrophotometer Amersham Biosciences (Bucks, UK)

2.1.5 Consumables

Amicon ultra centricons (10-50 kDa) Millipore (Co-cork, Ireland)

Microcentrifuge and centrifuge tubes Axygen, Greiner

Syringe Filters (0.2/0.45 µM) Sartorius (Bradford, USA)

Syringe, needles, serological pipettes BD Biosciences

2.1.6 Computer software

ATSAS 2.8.3 EMBL (Hamburg, Germany)

Adobe Photoshop cs6 Adobe Systems, San Jose, CA, USA

Coot	Emsley and Cowtan, 2004 [155]
HKL2000 package	Zbyszek and Wladek, 1997 [156]
MolProbity	Davis <i>et al.</i> , ASER McCoy <i>et al.</i> , 2007 [157]
PROCHECK	Laskowski <i>et al.</i> , 1993 [158]
PyMOL 4.3.0	DeLano Scientific LLC, USA [159]
Ramachandran plot	Ramachandran <i>et al.</i> , 1963 [160]
Refmac5 (CCP4 package)	Murshudov <i>et al.</i> , 1997 [161]
UCSF Chimera	Pettersen <i>et al.</i> , 2004 [162]
Vector NTI 10.3.0	Invitrogen

2.2 Methodology

2.2.1 Cloning of DENV-2 and ZIKV constructs

2.2.1.1 Cloning of DENV-2 NS2B₁₈NS3

The full length DENV-2 NS2B₁₈NS3 (GenBank: GQ398264) was generated using overlap extension PCR. Two oligonucleotides were synthesized, which represents Fragment 1 for the overlap PCR (Fig. 2.2.1). Oligo 1 contained the *Nco*I restriction site (underlined) and the sequence corresponds to nucleotides encoding residues 49 to 66 of DENV2 NS2B and Gly4-Ser-Gly4 linker. Oligo 2 was the reverse complement of oligo 1. The sequences of oligo 1 and 2 are given in the table below:

Fragment 1	Oligonucleotide sequence 5'-3'	Tm
Oligo 1	ACG <u>TCC</u> ATG <u>GCC</u> GAC TTG GAA CTG GAG AGA GCC GCT	75.3 °C
	GAC GTA AGA TGG GAA GAA CAG GCA GAG GGT GGC GGA	
	GGT AGT GGT GGA GGC GGT GCC GGA GT	
Oligo 2	ACT CCG GCA CCG CCT CCA CCA CTA CCT CCG CCA CCC	75.3 °C
	TCT GCC TGT TCT TCC CAT CTT ACG TCA GCG GCT CTC	
	TCC AGT TCC AAG TCG GCC ATG GAC GT	

Table 2.2.1: Oligonucleotides synthesized, which corresponds to the sequence coding for 49 to 66 of DENV2 NS2B and Gly4-Ser-Gly4 linker

Fragment 2 contained the nucleotides encoding residues 1 to 618 of DENV2 NS3. This fragment was amplified using an NS3-forward primer 5'- GAG GCG GTG CCG GAG TAT TGT GGG ATG TC-3' and an NS3-reverse primer 5'- TCA GAA GCT TAC TTT CTT CCA GCT GCG AAT TCC-3' with *Hind*III restriction enzyme site (underlined). The two fragments were subsequently mixed and amplified using a new forward primer with the *Nco*I restriction site (underlined) and NS3 reverse primer that was used to clone NS3 fragment.

Primers	Oligonucleotide sequence 5'-3'	Tm
Forward primer	ACG TCC ATG GCC GAC TTG GAA CTG GAG AGA	67 °C
Reverse primer	TCA GAA GCT TAC TTT CTT CCA GCT GCG AAT TCC	62.8 °C

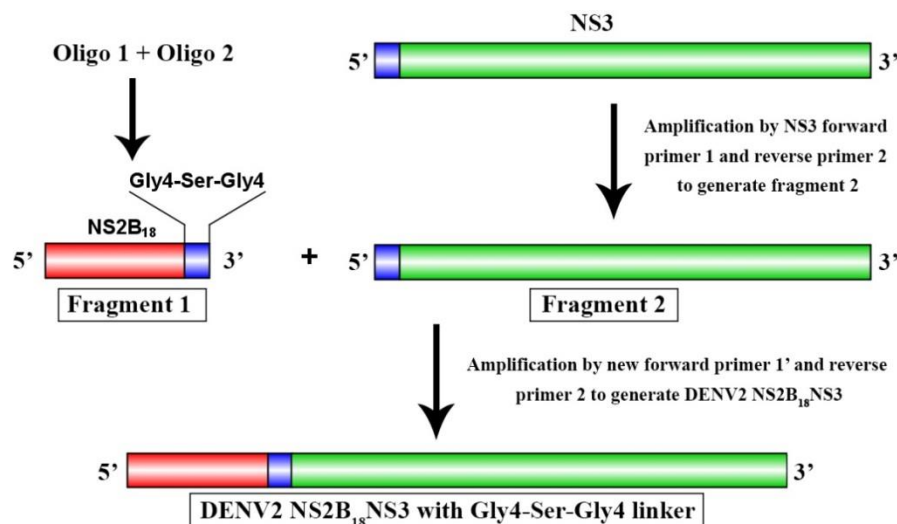
Table 2.2.2: Primers used for cloning DENV-2 NS2B₁₈NS3

Figure 2.2.1: Representation of overlap extension PCR used to generate the DENV-2 NS2B₁₈NS3. Fragment 2 was generated by amplifying NS3 region from DENV2 with forward primer 1 and reverse primer 2. The fragment 2 generated was mixed with fragment 1 and amplified using a new forward primer 1' and the same reverse primer (primer 2) used to generate the fragment 2.

DENV-2 cDNA was used as a template for the amplification of fragment 2 (NS3-region). PCR was set up in a total volume of 50 µl on ice as mentioned below with the appropriate concentration of each constituent as:

Reagents	Amount
5x KAPA HiFi-buffer (with 10 mM Mg ₂ SO ₄)	10 µl
DNTP's (10 mM)	1.5 µl
Forward primer (100 µM)	0.3 µl
Reverse primer (100 µM)	0.3 µl
Template (plasmid DNA) (10 ng/µl)	0.5 µl
MilliQ water	36.4 µl
KAPA HiFi-DNA polymerase	1.0 µl

Table 2.2.3: Components of PCR reaction mix for cloning the NS3 region of DENV-2 NS3 (Fragment2)

All reagents were pipetted into a PCR tube that was kept on ice all the time. The reagents were gently mixed and short centrifuged. The PCR-thermocycler (Biometra T personal) was preheated to 99 °C before reaction tubes were placed inside the machine. The following PCR thermocycler program was used for amplification:

Cycle steps	Temperature	Time
Initial denaturation	96 °C	3 min
Cycle denaturation	96 °C	20 sec
Annealing	58 °C	15 sec
Extension	72 °C	270 sec
Final extension	72 °C	5 min
End	4 °C	

Table 2.2.4: Set-up of PCR program for cloning the NS3 region of DENV-2 NS3 (Fragment2)

After amplification, 5 µl of the PCR product was applied onto an analytical agarose gel (1.5%) to check its quality. The gel was kept in ethidium bromide solution for 20 min and observed under UV light. The size of the PCR-product was confirmed with the appropriate DNA marker. Subsequently, the fragment 2 was mixed with fragment 1 (oligo 1 + oligo 2) in equimolar ratio for being used for the overlap PCR. To obtain an equimolar ratio of fragment 1 and fragment 2, the following constituents were added into a PCR tube:

Oligonucleotides	Amount
<div> <div> Oligo 1 Oligo 2 </div> <div> } </div> <div> Fragment 1 </div> </div>	0.5 µl
Fragment 2	0.6 µl

Table 2.2.5: Amount of oligo 1 and 2 and fragment 2 mixed to obtain template to be used for overlap extension PCR

Overlap extension PCR was then set up in a total volume of 50 µl on ice as mentioned below with the appropriate concentration of each constituent as:

Reagents	Amount
5x KAPA HiFi-buffer (with 10 mM Mg ₂ SO ₄)	10 µl
DNTP's (10 mM)	1.5 µl
Forward primers (100 µM)	0.3 µl
Reverse primers (100 µM)	0.3 µl
Equimolar ratio of fragment 1+2	1.6 µl
MilliQ water	35.3 µl
KAPA HiFi DNA polymerase	1.0 µl

Table 2.2.6: Components of PCR reaction mix for overlap extension PCR

PCR thermocycler (Biometra T personal) was preheated to 99 °C before reaction tubes were placed inside the machine. The following PCR thermocycler program was used for amplification:

Cycle steps	Temperature	Time
Lid	99 °C	
Initial denaturation	96 °C	3 min
Cycle denaturation	96 °C	20 secs
Annealing	65 °C	15 secs
Extension	72 °C	270 secs
Final extension	72 °C	5 min
End	4 °C	

} 30 cycles

Table 2.2.7: PCR thermocycler program for overlap extension PCR

After amplification, the PCR-product was tested on a 1.5% agarose gel. The amplified product, corresponding to the correct size of the construct, was excised from the gel and purified by QIAquick gel extraction kit (Qiagen) as per the manufacturer's protocol. The product was finally eluted in 30 µl of elution buffer (10 mM Tris-HCl, pH 8.5; Elution Buffer (EB), QIAGEN). The purified DNA was double digested overnight at 37 °C by using an appropriate amount of *NcoI* and *HindIII* enzymes. After overnight digestion, the reaction mixture was purified by QIAquick PCR purification kit (Qiagen). Subsequently, ligation reaction between *NcoI* and *HindIII* digested modified pET32b and digested PCR-product was setup at room temperature. The reagents used were as follows:

Reagents	Amount
Ligase buffer (10x)	1.5 µl
Vector (V)	60 ng
Insert (I)	variable (3 or 5 times of vector)
T4 Ligase	1 µl
MilliQ water	variable

Table 2.2.8: Components of the reaction mix for setting up ligation

The ligation reaction was performed for 3 hours at room temperature. DNA precipitation was done following ligation. 90 µl of MilliQ water was added to the ligation reaction mix to make a final volume of 100 µl. 1 ml of butanol was added into the mixture and mixed thoroughly. DNA was precipitated by centrifuging at 13,000 x *g* (using Eppendorf Mini-centrifuge) for 10 minutes and the supernatant was discarded gently. The pellet was resuspended in 200 µl of 70% ethanol and centrifuged again for 5 minutes. Thereafter, the supernatant was again removed, and the pellet was re-dissolved in 10 µl of sterile water.

5 µl of dissolved DNA was used for the transformation into *E. coli* DH5α cells. Single colony was picked from the transformation plate on the next day and was incubated in LB media containing 30 µg/ml ampicillin at 37 °C. The plasmid DNA was isolated using standard protocol QIAprep Spin Miniprep Kit (Qiagen). It was subsequently double digested and applied

onto a 1.5% agarose gel, to confirm the presence of the ligated insert. The size of insert and vector were compared with marker. The sequences of selected plasmid were confirmed by DNA sequencing (1st Base, Singapore).

2.2.1.2 Cloning of DENV-2 protease with linker (NS2B₁₈NS3₁₋₁₇₉)

The protease with linker (NS2B₁₈NS3₁₋₁₇₉) construct was created by introducing a stop codon (TAA) after residue 179 of NS3 by site-directed mutagenesis. The stop-codon (underlined) was incorporated into the plasmid using specifically designed forward and reverse primers. The insertion sequence was incorporated towards the 5'-end of the forward primer, while the reverse primer anneals back-to-back with the 5'-end of the complementary region of the forward primer (Fig 2.2.2).

Primers	Oligonucleotide sequence 5'-3'	Tm
Forward primer	GAG ATT GAA <u>TAA</u> GAC ATT TTT CGA AAG AAA AGA TTG ACC	59.2 °C
Reverse primer	AAA AAT GTC <u>TTA</u> TTC AAT CTC TGG ATT GTC TTC AAT GCT	60 °C

Table 2.2.9: Primers used site-directed mutagenesis to generate DENV-2 protease

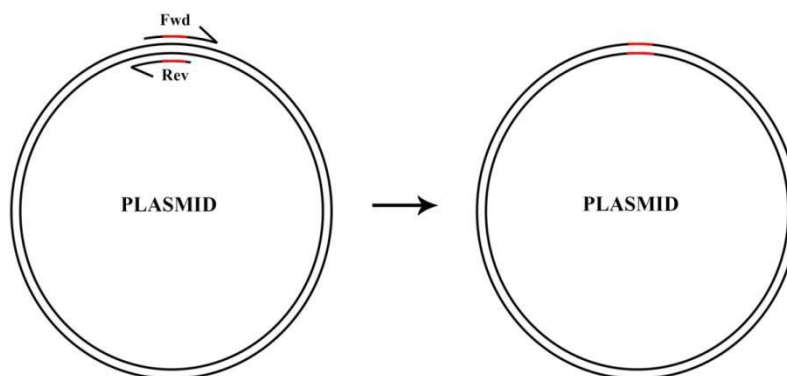


Figure 2.2.2: Insertion of a stop codon into a plasmid using specifically designed forward (Fwd) and reverse (Rev) primers. The stop codon is inserted into the plasmid of interest by site-directed mutagenesis.

The whole pET32b plasmid carrying DENV-2 NS2B₁₈-G₄SG₄-NS3 fusion construct was used for generating the DENV-2 protease with linker construct via site-directed mutagenesis. The reaction was set up in a total volume of 50 µl on ice as mentioned below with the appropriate concentration of each constituent as:

Reagents	Amount
5x KAPA HiFi-buffer (with 10mM Mg ₂ SO ₄)	10 µl
DNTP's (10 mM)	1.5 µl
Forward primers (100 µM)	0.3 µl
Reverse primers (100 µM)	0.3 µl
Template (plasmid DNA) (10 ng/µl)	1.0 µl
MilliQ water	35.9 µl
KAPA HiFi DNA polymerase	1.0 µl

Table 2.2.10: Components of PCR reaction mix

The following PCR thermocycler program was used for amplification where the lid of the thermocycler was preheated to 99 °C:

Cycle steps	Temperature	Time
Initial denaturation	96 °C	3 min
Cycle denaturation	96 °C	20 secs
Annealing	55 °C	15 secs
Extension	72 °C	150 secs
Final extension	72 °C	5 min
End	4 °C	

Table 2.2.11: PCR thermocycler program

Thereafter, a 1.5 % agarose gel was used to prove the quality and quantity of the PCR-product. The gel was stained with ethidium bromide solution for 20 min and observed under UV light. To remove the methylated DNA template 10 µl of the reaction mixture was treated with *DpnI* (NEB).

Reagents	Amount
PCR product	10.0 µl
Cut Smart buffer (10x)	1.5 µl
<i>DpnI</i>	0.3 µl
MilliQ water	34.0 µl

Table 2.2.12: Reaction mix for *DpnI* digestion

After 1 h incubation at 37 °C *DpnI* enzyme was heat inactivated at 80 °C for 20 minutes. Following heat inactivation of the enzyme, DNA was precipitated as described earlier. The DNA pellet was dissolved in 10 µl of water and DNA was used for the transformation into *E. coli* DH5α cells. Single colony was picked from the transformation plate on the next day and was incubated in LB media supplemented with 30 µg/ml ampicillin at 37 °C. The plasmid DNA was isolated, and the sequences of selected plasmid were confirmed by DNA sequencing (1st Base, Singapore).

2.2.1.3 Cloning of DENV2 helicase (NS3₁₆₉₋₆₁₈) with linker

The DENV2 helicase (NS3₁₆₉₋₆₁₈) with linker (residues 169-179) was amplified using the following primers incorporating the *NcoI* and *HindIII* sites (in bold):

Primers	Oligonucleotide sequence 5'-3'	Tm
Forward primer	GGT TCT GCC ATG GCA GAA AAA AGC ATT GAA GAC AAT CCA GAG	59.2 °C
Reverse primer	TCA GAA GCT TAC TTT CTT CCA GCT GCG AAT TCC	60 °C

Table 2.2.13: Primers used to generate DENV-2 helicase

The whole pET32b plasmid, carrying DENV-2 NS2B₁₈-G₄SG₄-NS3 fusion construct, was used for generating the DENV-2 helicase with linker construct. The PCR-reaction was set up in a total volume of 50 µl on ice as mentioned below with the appropriate concentration of each constituent as:

Reagents	Amount
5x KAPA HiFi-buffer (with 10mM Mg ₂ SO ₄)	10 µl
DNTP's (10 mM)	1.5 µl
Forward primers (100 µM)	0.3 µl
Reverse primers (100 µM)	0.3 µl
Template (plasmid DNA) (10 ng/µl)	1.0 µl
MilliQ water	35.9 µl
KAPA HiFi-DNA polymerase	1.0 µl

Table 2.2.14: Components of PCR reaction mix for generating DENV-2 helicase

Amplification of the NS3-helicase with linker was completed using the following PCR thermocycler program where the lid of the thermocycler was preheated to 99 °C:

Cycle steps	Temperature	Time
Initial denaturation	96 °C	3 min
Cycle denaturation	96 °C	20 sec
Annealing	57 °C	15 sec
Extension	72 °C	200 sec
Final extension	72 °C	5 min
End	4 °C	

Table 2.2.15: PCR program set-up

After amplification, the PCR-product was digested with *NcoI* and *HindIII*, and ligated into the modified pET32b vector. The plasmid was then transformed in *E. coli* DH5α cells. Single colonies were picked from the transformation plate on the next day and were incubated in LB media containing 30 µg/ml ampicillin at 37 °C. The plasmid DNA was isolated, subsequently double digested and applied onto a 1.5% agarose gel, to confirm the presence of

the ligated insert. The sequences of selected plasmid were confirmed by DNA sequencing (1st Base, Singapore).

2.2.1.4 Cloning of DENV-2 NS2B₁₈NS3₁₇₄PPAVP₁₇₉ mutant

To investigate the effect of the substitution of PPxxP motif from the hepaciviruses on DENV NS3, a linker mutant was generated. The amino acid residues 174-179 of the DENV-2 NS3 linker was substituted by the ₁₇₄PPAVP₁₇₉ residues from HCV NS3. The mutant was created via site-directed mutagenesis, using DENV-2 NS2B₁₈NS3 fused to modified pET32b as the template DNA. The primers encoding the mutations (underlined) are listed as following:

Primers	Oligonucleotide sequence 5'-3'	T _m
Forward primer	GAA GAC <u>CCA CCA GCG GTG CCC</u> GAT GAC ATT TTT CGA AAG AAA AGA	68.6 °C
Reverse primer	GTC ATC <u>GGG CAC CGC TGG TGG GTC</u> TTC AAT GCT TTT TTC AGT CTG	69.7 °C

Table 2.2.16: Primers for generating DENV-2 NS3 linker mutant

The PCR reaction mix and thermocycler set-up were the same as described in section 2.2.1.2. The only difference was that the annealing temperature was set at 62°C. The same treatment with *DpnI* enzyme (see section 2.2.1.2) was applied to 10 µl of PCR mixture to remove the methylated DNA template. 5 µl of the sample were then precipitated and transformed into DH5α *E. coli* cells. Plasmids were isolated from the transformants using QIAprep Spin Miniprep Kit, and the sequences of mutated sites were verified by DNA sequencing (1st Base, Singapore).

2.2.1.5 Cloning of ZIKV NS2B₄₈NS3

The gene encoding French Polynesia ZIKV NS3 (strain H/PF/2013) linked to 48 amino acids (49-97) of NS2B (GenBank: KJ776791) was synthesized (Integrated DNA Technologies, Singapore) and used as the template for subsequent molecular cloning. The construct was cloned by Dr. Ardina Grüber from our lab. To transfer the full-length NS3 plus NS2B (49-97) gene into modified pET32b plasmid, the gene encoding NS2B₄₈NS3 was amplified using the forward primer 5'-AGA TCC ATG GTG GAC ATG TAT ATT-3' and reverse primer 5'-TGA TAA GCT TAT CTT TTC CCA GC-3', with the incorporation of restriction sites *NcoI* and *HindIII* (underlined). The reaction mix for the PCR was as follows:

Reagents	Amount
5x Pfu-buffer (with 10mM Mg ₂ SO ₄)	5 µl
DNTP's (10 mM)	1 µl
Forward Primers (100 µM)	1 µl
Reverse Primers (100 µM)	1 µl
Template (plasmid DNA) (40 ng/µl)	1.0 µl
MilliQ water	40 µl
Pfu-DNA polymerase	1.0 µl

Table 2.2.17: Components of PCR reaction mix

The reaction mix was kept on ice and the DNA polymerase was added as the last component. The mixture was then placed in a thermocycler and using the PCR program:

Cycle steps	Temperature	Time	
Initial denaturation	95 °C	4 min	
Cycle denaturation	95 °C	30 sec	} 16 cycles
Annealing	48 °C	45 sec	
Extension	72 °C	4 min	
Final extension	72 °C	10 min	
End	4 °C		

Table 2.2.18: PCR thermocycler program

After amplification, the presence of PCR-product was checked on a 1.5% agarose gel. The DNA corresponding to the correct size of construct was cut out immediately and purified by QIAquick gel extraction kit (Qiagen) as per the manufacturer's protocol. The product was finally eluted in 30 µl of EB buffer (Qiagen). The purified DNA was double digested at 37 °C for 1 h by taking appropriate amount of *Nco*I and *Hind*III Fast digestion enzymes. After enzyme digestion and purification of the PCR-product using QIAquick PCR purification kit (Qiagen), ligation reaction with the modified pET32b and PCR-product was setup at room temperature as described earlier.

The ligation reaction was performed at room temperature for 3 hours. The reaction was terminated by precipitation of DNA from the rest of the reaction mixture by adding 90 µl of MilliQ water to make a final volume of 100 µl. 1 ml of butanol was added into the mixture and mixed thoroughly. Then the mixture was centrifuged at 13,000 x g (using Eppendorf Mini-centrifuge) for 10 min. and the supernatant was discarded gently. The pellet was resuspended in 200 µl of 70% ethanol and centrifuged again for 5 minutes. Supernatant was again removed, and the pellet was re-dissolved in 6 µl of elution buffer (10 mM Tris-HCl, pH 8.5; Elution Buffer (EB), QIAGEN).

6 µl of ligation mixture was used for the transformation into *E. coli* DH5α cells. Single colony was picked from the transformation plate on the next day and was incubated in LB

media containing 30 µg/ml ampicillin at 37 °C. The plasmid DNA was isolated using standard protocol (QIAGEN mini-prep kit). It was subsequently double digested and analysed with a 1.5% agarose gel, to confirm the presence of the ligated insert. The size of insert and vector were compared with appropriate controls and markers. The sequences of the selected plasmid were obtained from DNA sequencing results (1st Base, Singapore).

A mutant of ZIKV NS2B₄₈NS3, NS2B₄₈NS3_{S135A}, was generated by site-directed mutagenesis replacing S135 by A, to prevent auto-proteolytic cleavage. The mutation (underlined) was incorporated into the plasmid using specifically designed forward and reverse primers stated in the following table:

Primers	Oligonucleotide sequence 5'-3'	T _m
Forward primer	GCA GGA ACT <u>GCC</u> GGA TCT CCA ATC CTA GAC AAG	65.3 °C
Reverse primer	GAT TGG AGA TCC <u>GGC</u> AGT TCC TGC TGG GTA ATC	65.6 °C

Table 2.2.19: Primers for site-directed mutagenesis

The reaction mix for the PCR was as follows:

Reagents	Amount
5x KAPA HiFi-buffer (with 10mM Mg ₂ SO ₄)	10 µl
DNTP's (10 mM)	1.5 µl
Forward Primers (100 µM)	0.3 µl
Reverse Primers (100 µM)	0.3 µl
Template (plasmid DNA) (10 ng/µl)	1.0 µl
MilliQ water	35.9 µl
KAPA HiFi-DNA polymerase	1.0 µl

Table 2.2.20: Components of PCR mix

The reaction mix was kept on ice and the polymerase was added as the last component. The mixture was then placed in a thermocycler and using the PCR program:

Cycle steps	Temperature	Time	
Initial denaturation	96 °C	3 min	
Cycle denaturation	96 °C	20 sec	} 20 cycles
Annealing	62 °C	15 sec	
Extension	72 °C	270 sec	
Final extension	72 °C	5 min	
End	4 °C		

Table 2.2.21: PCR program

Analytical agarose gel (1.5%) was used to confirm the successful amplification and the presence of desired PCR-product. The PCR-product was stained using ethidium bromide technique for 20 min and visualized under UV light. 20 µl of the desired product was afterwards

treated with *DpnI* (NEB) in order to remove the methylated DNA template. The reaction was set up according to the set-up described in table 2.2.12. and incubation lasted for 1 hour at 37 °C.

After 20 minutes incubation at 80 °C to inactivate the *DpnI* enzyme, precipitation of DNA was done as earlier described. The DNA pellet obtained by precipitation was dissolved in 10 µl of sterile water and transformed into *E. coli* DH5α cells. Single colony was picked from the transformation plate on the next day and was incubated in LB-media containing 30 µg/ml ampicillin at 37 °C. The plasmid DNA was isolated, and the sequences of the selected plasmid were confirmed by DNA sequencing (1st Base, Singapore).

2.2.1.6 Cloning of ZIKV NS2B₄₈NS3 protease with linker

The protease with linker (NS2B₄₈NS3₁₋₁₇₉) construct was created by introducing a stop codon (TAG) after residue 179 of NS3 by site-directed mutagenesis. The stop-codon (underlined) was incorporated into the plasmid using specifically designed forward and reverse primers. The insertion sequence was incorporated towards the 5'-end of the forward primer, while the reverse primer anneals back-to-back with the 5'-end of the complementary region of the forward primer.

Primers	Oligonucleotide sequence 5'-3'	Tm
Forward primer	GAG CCT <u>TAG</u> ATC GAT GCT GAA GAA GAA GCA GC	63.1 °C
Reverse primer	CAT CGA <u>TCT</u> <u>AAG</u> GCT CGA AGC ACT CAA CAG	61.7 °C

Table 2.2.22: Primers used to generate ZIKV protease

The lid of the PCR thermocycler was preheated to 99 °C before reaction tubes were placed inside the machine. The reaction mix for the PCR was as follows:

Reagents	Amount
5x KAPA HiFi-buffer (with 10mM Mg ₂ SO ₄)	10 µl
DNTP's (10 mM)	1.5 µl
Forward Primers (100 µM)	0.4 µl
Reverse Primers (100 µM)	0.4 µl
Template (plasmid DNA) (10 ng/µl)	1.0 µl
MilliQ water	35.7 µl
KAPA HiFi-DNA polymerase	1.0 µl

Table 2.2.23: Components of PCR mix

The NS2B₄₈NS3_{S135A} in the modified pET32b plasmid was used as template for cloning the ZIKV protease. The reaction mix was kept on ice and the DNA polymerase was added as the last component. The PCR program was set up according to the protocol mentioned in table 2.2.11.

Once the presence of amplified product was confirmed, *DpnI* digestion was performed as previously described followed by DNA precipitation to desalt the product. 7 µl of sterile water added to dissolve the precipitated DNA which was subsequently utilized for transformation into DH5α *E. coli* cells by electroporation. The following day plasmids were isolated from the transformants using QIAprep Spin Miniprep Kit, and the sequences of mutated sites were confirmed by DNA sequencing (1st Base, Singapore).

2.2.1.6 Cloning of ZIKV NS2B₄₈NS3 helicase with linker

The ZIKV helicase (NS3₁₆₉₋₆₁₈) with linker (residues 169-179) was generated using the primers described in the table 2.2.24. The *Bam*HI and *Hind*III restriction sites in the forward and reverse primers are underlined.

Primers	Oligonucleotide sequence 5'-3'	Tm
Forward primer	ATT AGG ATC CAG GAG GGA GGA AGA GAC TCC	63 °C
Reverse primer	TAC GAA GCT TAT CTT TTC CCA GCG GCA AAC TC	63.1 °C

Table 2.2.24: Primers used to generate ZIKV helicase

Amplification of ZIKV protease was done using the NS2B₄₈NS3_{S135A} in the modified (pET 32b) plasmid as the template with the following PCR program:

Reagents	Amount
10x <i>Pfu</i> buffer with 20 mM Mg ₂ SO ₄	5.0 µl
DNTP's (10 mM)	1.0 µl
Forward Primers (100 µM)	0.4 µl
Reverse Primers (100 µM)	0.4 µl
Template (plasmid DNA) (50 ng/µl)	1.0 µl
MilliQ water	41.2 µl
<i>Pfu</i> DNA polymerase	1.0 µl

Table 2.2.25: Components of PCR mix

The mixture was then placed in a thermocycler and using the PCR program:		
Cycle steps	Temperature	Time
Initial denaturation	96 °C	3 min
Cycle denaturation	96 °C	20 secs
Annealing	58 °C	15 secs
Extension	72 °C	200 secs
Final extension	72 °C	5 min
End	4 °C	

Table 2.2.26: PCR thermocycler program

After amplification, the PCR-product was digested with *Bam*HI and *Hind*III and ligated into the modified pET32b vector. The plasmid was then transformed in *E. coli* DH5 α cells. Single colonies were picked from the transformation plate on the next day and were incubated in LB media containing 30 μ g/ml ampicillin at 37 °C. The plasmid DNA was isolated and subsequently double digested and applied onto a 1.5% agarose gel, to confirm the presence of the ligated insert. The sequences of the selected plasmid were confirmed by DNA sequencing (1st Base, Singapore).

2.2.1.7 Cloning of ZIKV NS2B₁₈NS3

The ZIKV NS2B₁₈NS3 was cloned using the same approach as that for cloning DENV-2 NS2B₁₈NS3. Two oligonucleotides representing fragment 1 were synthesized which contained the nucleotides encoding residue 49 to 67 of NS2B and the Gly4-Ser-Gly4 linker. The sequences of oligo 1 and 2 are given in the table below:

Fragment 1	Oligonucleotide sequence 5'-3'	T _m
Oligo 1	GTG GAC ATG TAT ATT GAA AGA GCA GGT GAC ATC ACA TGG GAA AAA GAT GCG GAA GGT GGC GGA GGT AGT GGT GGA GGC GGT AGT GGT GCT	72.9 °C
Oligo 2	AGC ACC ACT ACC GCC TCC ACC ACT ACC TCC GCC ACC TTC CGC ATC TTT TTC CCA TGT GAT GTC ACC TGC TCT TTC AAT ATA CAT GTC CAC	72.9 °C

Table 2.2.27: Oligonucleotides synthesized which corresponds to the sequence coding for 49 to 66 of ZIKV NS2B and Gly4-Ser-Gly4 linker

Fragment 2 contained the nucleotides encoding residues 1 to 618 of ZIKV NS3. This fragment was amplified using an NS3-forward primer 5'- GGT GGA GGC GGT AGT GGT GCT CTA TGG GAT GTG-3' and previously described reverse primer for helicase (Table 2.2. 24).

ZIKV NS2B₄₈NS3 in the modified pET32b plasmid was used as a template for the amplification of fragment 2 (NS3-region). PCR was set up in a total volume of 50 µl on ice as mentioned below with the appropriate concentration of each constituent as:

Reagents	Amount
5x KAPA HiFi-buffer (with 10mM Mg ₂ SO ₄)	10 µl
DNTP's (10 mM)	1.5 µl
Forward primer (100 µM)	0.4 µl
Reverse primer (100 µM)	0.4 µl
Template (plasmid DNA) (10 ng/µl)	0.5 µl
MilliQ water	35.7 µl
KAPA HiFi-DNA polymerase	1.0 µl

Table 2.2.28: Components of PCR reaction mix for cloning NS3 region of ZIKV NS3 (Fragment 2)

A gradient PCR was set-up at the thermocycler with two different annealing temperatures as described in table 2.2.29.

Cycle steps	Temperature	Time	
Initial denaturation	96 °C	3 min	
Cycle denaturation	96 °C	20 sec	} 20 cycles
Annealing	57.4 / 62 °C	15 sec	
Extension	72 °C	270 sec	
Final extension	72 °C	5 min	
End	4 °C		

Table 2.2.29: Set-up of PCR program for cloning NS3 region of ZIKV NS3 (Fragment 2)

The fragments 1 and 2 were subsequently mixed in equimolar ratio a new forward primer with the *Nco*I restriction site (underlined) and helicase reverse primer.

Primers	Oligonucleotide sequence 5'-3'	T _m
Forward primer	TCT <u>GCC</u> ATG <u>GCA</u> GTG GAC ATG TAT ATT GAA AGA	62.3 °C
Reverse primer	TAC <u>GAA</u> <u>GCT</u> <u>TAT</u> CTT TTC CCA GCG GCA AAC TC	63.1 °C

Table 2.30: Primers used for overlap extension PCR for cloning ZIKV NS2B₁₈NS3

Overlap extension PCR was then set up in a total volume of 50 µl on ice as described in table 2.2.6. The lid of the PCR thermocycler was preheated to 99 °C before reaction tubes were placed inside the machine. The following PCR thermocycler program was used for amplification:

Cycle steps	Temperature	Time	
Lid	99 °C		
Initial denaturation	96 °C	3 min	
Cycle denaturation	96 °C	20 sec	} 20 cycles
Annealing	62 °C	15 sec	
Extension	72 °C	270 sec	
Final extension	72 °C	5 min	
End	4 °C		

Table 2.31: PCR thermocycler program for overlap extension PCR

After amplification, the PCR-product was checked on a 1.5% agarose gel. The amplified product corresponding to the correct size of the construct was excised from the gel and purified by QIAquick gel extraction kit (Qiagen) as per the manufacturer's protocol. After digestion of the amplified product with *NcoI* and *HindIII*, ligation with the digested vector was set as previously described. The ligated product was precipitated and transformed into *E. coli* DH5 α cells. Single colony was picked from the transformation plate on the next day and was incubated in LB media containing 30 μ g/ml ampicillin at 37 °C. Purified plasmid was double digested using *NcoI/HindIII* enzymes and the presence of ligated insert was confirmed by 1.5% agarose gel. The sequences of selected plasmid were validated by DNA sequencing (1st Base, Singapore).

2.2.1.8 Cloning of ZIKV NS2B₁₈NS3 mutants

To investigate the importance of the highly conserved ZIKV NS3 linker residues, various linker mutants were generated by substituting single amino acid residues in the ZIKV linker. The mutants were created using site-directed mutagenesis, using ZIKV NS2B₁₈NS3 fused to modified pET32b as the template DNA. The primers encoding the mutations (underlined) are listed as following:

Mutant Name	Primers (5'-3')	T _m (°C)
NS2B ₁₈ NS3 _{T174A}	F 5'- GAG GAA GAG <u>GCT</u> CCT GTT GAG TGC TTC GAG CCT -3'	67.6
	R 5'- CTC AAC AGG <u>AGC</u> CTC TTC CTC CCT CCT CCC TTG-3'	67.5
NS2B ₁₈ NS3 _{V176A}	F 5'- GAG ACT CCT <u>GCT</u> GAG TGC TTC GAG CCT TCG ATG-3'	66.5
	R 5'- GAA GCA CTC <u>AGC</u> AGG AGT CTC TTC CTC CCT CCT-3'	67.1
NS2B ₁₈ NS3 _{E177A}	F 5'- ACT CCT GTT <u>GCG</u> TGC TTC GAG CCT TCG ATG CTG-3'	68.5
	R 5'-CTC GAA GCA <u>CGC</u> AAC AGG AGT CTC TTC CTC CCT-3'	67.5
NS2B ₁₈ NS3 _{C178A}	F 5'-CCT GTT GAG <u>GCT</u> TTC GAG CCT TCG ATG CTG AAG-3'	65.9
	R 5'-AGG CTC GAA <u>AGC</u> CTC AAC AGG AGT CTC TTC CTC-3'	66.0
NS2B ₁₈ NS3 _{F179A}	F 5'-GAG ACT CCT GTT GAG TGC <u>GCC</u> GAG CCT TCG ATG-3'	67.2
	R 5'-AGG CTC <u>GGC</u> GCA CTC AAC AGG AGT CTC TTC CTC-3'	69.0
NS2B ₁₈ NS3 _{F179I}	F 5'-GAG ACT CCT GTT GAG TGC <u>ATC</u> GAG CCT TCG ATG-3	67.5
	R 5'-AGG CTC <u>GAT</u> GCA CTC AAC AGG AGT CTC TTC CTC-3'	68.0

Table 2.2.30: Primers used for overlap extension PCR for cloning ZIKV NS2B₁₈NS3 (F=Forward; R=Reverse)

The same treatment with *DpnI* enzyme (see section 2.2.1.2) was applied to 10 µl of PCR mixture to remove the methylated DNA template. 5 µl of the sample were then precipitated and transformed into DH5α *E. coli* cells. Plasmids were isolated from the transformants using QIAprep Spin Miniprep Kit, and the sequences of mutated sites were verified by DNA sequencing (1st Base, Singapore).

2.2.2 Recombinant gene expression for flaviviral constructs

The plasmids with the confirmed sequences were transformed into the expression *E. coli* strain, BL21-CodonPlus(DE3)-RIL competent cells by electroporation at a constant voltage of 2.5 kV (Micropulser electroporator, Bio-Rad).

2.2.3 Induction test of viral proteins from DENV

Cells with genes encoding full-length NS3, the single domains as well the mutants were grown in 2 ml LB media supplemented with 100 µg/ml ampicillin and 34 µg/ml chloramphenicol, at 37 °C with the shaking speed of 180 rpm. When cell cultures reached an attenuance (OD₆₀₀) of 0.6, the cells were split into two aliquots, and the final concentration of 1 mM IPTG was added into one of them for protein production. The cells were then incubated at 37 °C for two hours, and subsequently centrifuged and re-suspended in 60 µl of 1x lysis buffer (0.25 M Tris-HCl, pH 6.8, 9.2% SDS, 40% glycerol, 0.2% bromophenol blue) in the

presence of 12.5 mM DTT. The cell suspension was heated at 95 °C for 5 min, and 12 µl of sample was loaded onto a 17% SDS-polyacrylamide gel to check the presence of protein production. Afterwards, the SDS gel was stained with Coomassie staining solution (0.2% w/v Coomassie brilliant blue, 40% v/v methanol, 1% v/v acetic acid) for 5-10 min, and de-stained with de-staining solution 1 (25% v/v isopropanol, 10% v/v acetic acid), followed by de-staining solution 2 (7.5% v/v methanol, 10% acetic acid).

2.2.4 Protein production of viral proteins from DENV

Cells were grown in LB media under the condition stated in section 2.3. Protein production was induced at 18 °C for 16 hours. After induction, the cells were harvested at 7,000 x g for 12 min at 4 °C.

2.2.5 Purification of DENV-2 and ZIKV constructs

Cells producing the DENV-2 NS3-proteins were lysed on ice by sonication with an ultrasonic homogenizer (Bandelin, KE76 tip) for 3 x 1 min in buffer A (20 mM sodium phosphate, pH 7.5, 500 mM NaCl, 2 mM Pefabloc^{SC} and 0.8 mM DTT). After sonication, the cell lysate was centrifuged at 10,000 x g for 35 min at 4 °C. The resulting supernatant was passed through a filter (0.45 µm; Millipore) and supplemented with Ni²⁺-NTA resin preequilibrated in buffer A. The His-tagged protein could bind to the matrix for 1 hour at 4 °C by mixing on a sample rotator (Neolab) after which the Ni²⁺-NTA was thoroughly washed with 10 column volumes of buffer A without DTT. The protein was subsequently eluted with an imidazole gradient (20–500 mM). The fractions containing the protein of interest were identified by SDS–PAGE, pooled together and dialyzed overnight with thrombin in buffer B (20 mM sodium phosphate, pH 7.5, 200 mM NaCl and 1 mM DTT). The sample was incubated with Ni-NTA resin the following day for 1 hour at 4 °C to remove cleaved Trx-His₆ tag (Fig. 2.2.3).

While the NS2B₁₈NS3 was applied to SuperdexTM 200 HR 10/300 column (GE Healthcare), the protease and helicase with linker was applied to a SuperdexTM 75 HR 10/300 column (GE Healthcare) in buffer C (20 mM Tris/HCl, pH 7.5, 200 mM NaCl and 1 mM DTT). Fractions, containing the protein of interest, were pooled and concentrated. The purity and homogeneity of the recombinant proteins were analysed by a 17% SDS gel, which became stained with Coomassie Brilliant Blue R250. Protein concentrations were determined by the bicinchoninic acid assay (BCA; Pierce, Rockford, IL., USA) as well as by its absorbance at

280 nm in a spectrophotometer (BioSpec-Nano, Shimadzu) using an extinction coefficient $10,005 \text{ M}^{-1} \text{ cm}^{-1}$.

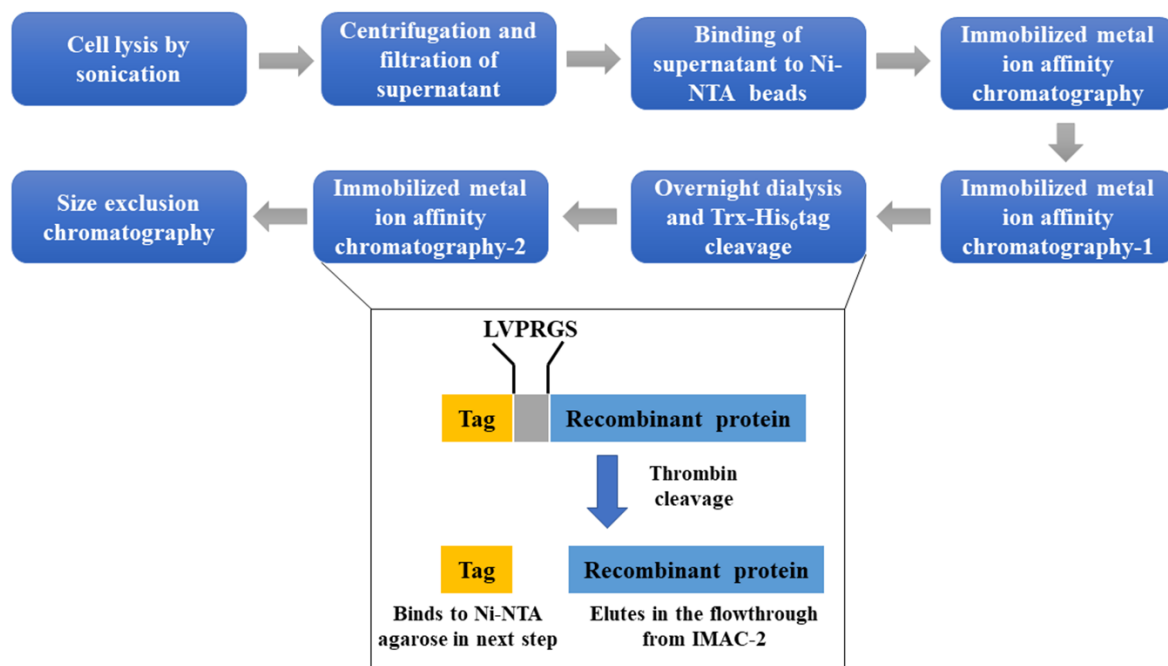


Figure 2.2.3: Schematic representation of flaviviral NS3 purification. The top panel depicts the steps in a flow diagram. The bottom panel shows the process occurring during overnight dialysis, namely N-terminal tag cleavage. The thrombin recognizes the sequence LVPRGS and the cleavage tag. This tag binds to the Ni-NTA beads while the recombinant protein is obtained in the flow-through of the second affinity chromatography and subsequently applied to size exclusion column.

2.2.6 ATP hydrolysis assay for DENV-2 and ZIKV NS2B₁₈NS3 and helicase

The specific ATP hydrolysis activity of NS2B₁₈NS3 and helicase was measured by a continuous ATP hydrolysis assay. The basis of this assay is a reaction where ATP is constantly regenerated and coupled to the oxidation of NADH (Fig. 2.2.4). On hydrolysis of ATP by the investigated protein, ADP is generated. The enzyme pyruvate kinase (PK) present in the assay catalyses the conversion of phosphoenol-pyruvate (PEP) to pyruvate and facilitates transfer of the phosphate moiety, generated from ATP hydrolysis, to ADP. Thus, ATP is regenerated. The pyruvate generated is converted to lactate by another enzyme present in the assay, lactate dehydrogenase (LDH). During this conversion, one molecule of NADH is oxidized to NAD⁺. NADH shows strong absorbance at 340 nm while NAD⁺ does not (Fig. 2.2.4). The highest occupied molecular orbital in the nicotinamide ring of NAD⁺ is a pi-bonding orbital whereas in the dihydronicotinamide ring of NADH it is the non-bonding orbital - the lone pair of electrons on the nitrogen. The energy required to move an electron from a non-bonding orbital to a pi* orbital is much smaller than that required to move an electron from pi-bonding orbital

to a π^* orbital. This assay spectroscopically measures the consumption of NADH, which is proportional to the rate of steady-state ATP hydrolysis.

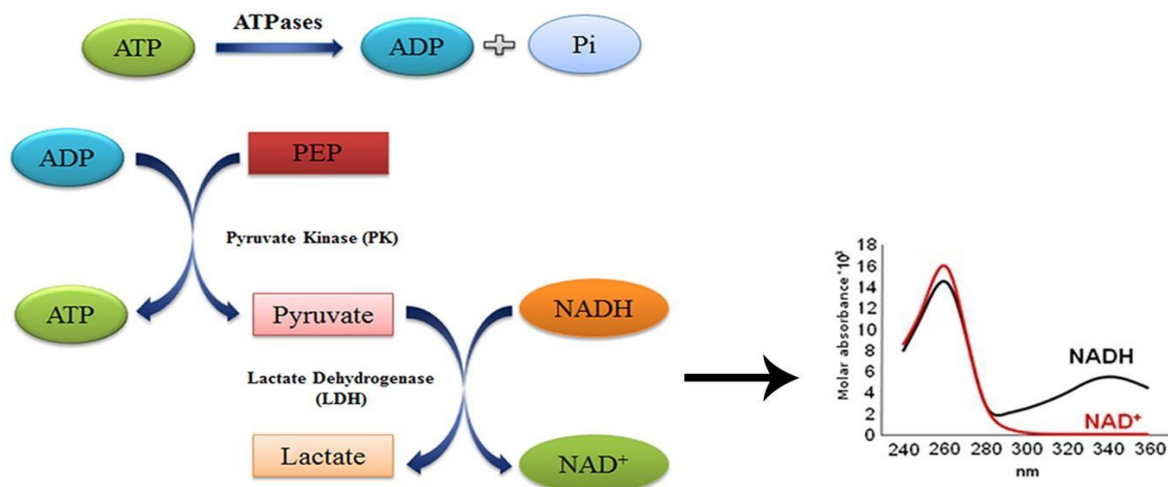


Figure 2.2.4: Schematic representation of the reactions occurring in the continuous ATP hydrolysis assay. ATPases are the proteins we are investigating that are known to hydrolyse ATP to ADP and P_i . ATP is continuously regenerated while NADH is oxidized to NAD^+ .

The change in absorbance was measured for 250 s in 2 s intervals at 37 °C after adding 5 μ g of NS2B₁₈NS3 and helicase to 1 ml reaction solution (25 mM HEPES pH 7.5, 25 mM KCl, 5 mM MgCl₂, 5 mM KCN, 2 mM phosphoenolpyruvate, 2 mM ATP, 0.35 mM NADH, 30 units L-lactic acid dehydrogenase, 30 units pyruvate kinase). All the measurements were performed in triplicate and its activity was derived by fitting the linear section of the slope from the averaged data.

The inhibition measurements were performed on NS2B₁₈NS3 and helicase. The reaction mixture was prepared by mixing 150 μ M of inhibitor (Resveratrol, quercetin, NBD-Cl) into the reaction solution stated above. After adding 5 μ g of NS2B₁₈NS3 and helicase into the reaction mixture, the change of absorbance was recorded for 250 s in 2 s intervals at 37 °C. For IC₅₀ determination of resveratrol to NS2B₁₈NS3 and helicase, the same procedure was used as described above, in a concentration range of 25 - 400 μ M. For each resveratrol concentration, the data were recorded in triplicate, and the averaged data were used to derive the ATPase activity. The ATPase activity was then normalized to 0 and 1, and plotted against the concentration of resveratrol. The experimental data were then fitted with dose response function using the program Origin Pro 9.0 (OriginLab Corporation), and the IC₅₀ was estimated from the program.

2.2.6 Small-angle X-ray scattering

2.2.6.1 SAXS data collection

SAXS-data of all NS3 DENV-2 proteins were collected with the BRUKER NANOSTAR SAXS instrument, equipped with a Metal-jet X-ray source and VANTEC 2000-detector system. The X-ray radiation was generated from a liquid gallium alloy with microfocus X-ray source ($KD = 1.3414 \text{ \AA}$ with a potential of 70 kV and a current of 2.857 mA). The X-rays were filtered through Montel mirrors and collimated by a two-pinhole system. The sample to detector distance was set at 0.67 m and the sample chamber and X-ray paths were evacuated. This setup covers a range of momentum transfer of $0.016 < s < 0.4 \text{ \AA}^{-1}$ ($s = 4\pi \sin(\theta)/\lambda$, where 2θ is the scattering angle). SAXS experiments were carried out at 15 °C for a concentration series ranging from 1.2 to 5.4 mg/ml and in buffer C (20 mM Tris/HCl, pH 7.5, 200 mM NaCl and 1 mM DTT) for a sample volume of 40 μl in a vacuum tight quartz capillary. Corresponding to each protein sample, data were collected for a buffer under identical experimental conditions, providing a background scattering curve. The data were collected for 30 minutes and for each measurement a total of six frames at 5 min intervals were recorded. The data were flood-field, spatial corrected and processed using the build-in SAXS software. The data were tested for possible radiation damage by comparing the six data frames, which indicated no change. The scattering of the corresponding buffer was subtracted, and the difference curves normalized for the concentration as well as initial intensity.

2.2.6.2 SAXS data analysis

The primary SAXS data processing steps were performed using the program package PRIMUS [163]. To identify the presence of inter-particle interaction between protein molecules with increasing protein concentrations (concentration dependence), the scattering profiles from different protein concentrations were overlapped and compared at low angle region. If all scattering profiles overlapped nicely, it was evaluated as no concentration dependence in the scattering profiles. If the low angle of scattering profiles shifted upward or downward, it was evaluated that there is concentration dependence presented in the scattering profiles, and the data will either be extrapolated to infinite dilution or merged together.

Two important parameters, including forward intensity ($I(0)$) and radius of gyration (R_g), were determined from Guinier approximation and pair-distance distribution ($P(r)$) function. The forward intensity ($I(0)$) is the radiation intensity scattered through zero angle ($q = 0$). This parameter cannot be measured, and it can only be determined through extrapolation,

because the detected intensity cannot be differentiated from the un-scattered intensity that passes through the sample [164, 165]. $I(0)$ is not dependent on the particle shape and this value can be used to determine the molecular mass of a particle [164]. The radius of gyration (R_g) is the root-mean-squared distance of each scatterer from the center of particle. This parameter is shape-dependent, whereby the objects with same volume but with different shapes have different R_g values [164, 165].

The Guinier approximation [166, 167] states that at very low angle ($q < 1/R_g$), the intensity can be represented as:

$$I(q) = I(0)e^{\frac{-q^2 R_g^2}{3}}$$

This means that when the scattering intensity are plotted as $\ln I(q)$ vs q^2 , the intensity exists as a linear function, and the R_g can be calculated from the slope of the curve and $I(0)$ can be derived from the y-intercept of the curve. However, this relationship is only valid for $qR_g < 1.3$ for globular protein. Protein aggregation and inter-particle interaction for all the protein samples were analyzed using the scattering data from the Guinier region [166], where the non-linear behavior at this region indicates the formation of aggregation. In time-resolved SAXS data analysis, this R_g value is computed using a command-line program AUTORG [168] to speed up the R_g estimation.

To validate these parameters, the R_g and $I(0)$ were also computed based on the pair-distance distribution ($P(r)$) function. This method considers the whole scattering data while computing these values and thus, the computation of both parameters can be more precise [165]. The $P(r)$ function describes all the distances between two electrons in a protein particle as a paired-set, and can be computed using the software GNOM [169]. This function is derived from the scattering intensity by an inverse Fourier transformation [169, 170] and their relationship is represented as:

$$p(r) = \frac{r^2}{2\pi^2} \int q^2 I(q) \frac{\sin qr}{qr} dq$$

Additionally, the maximum dimension (D_{max}) of particle can also be determined from the $P(r)$ function, where D_{max} is the maximum pair-distance that can be found within the particle.

2.2.6.3 Protein folding and flexibility assessment using SAXS

To qualitatively access the globular nature of the particle, a Kratky plot can be generated by transforming the scattering pattern of the particle as $q^2 I(q)$ vs q . For a globular protein with

a well-defined and solvent-accessible surface, its Kratky plot will exhibit a typical bell-curve profile with a well-defined maximum. Conversely, for a random chain, the Kratky plot exhibits a plateau or slightly increase at high q region. The Kratky plot cannot distinguish clearly between partially unfolded and fully folded particles since both exhibit a similar profile. To overcome this, a normalized Kratky plot was generated ($(qR_g)^2(I(q)/I(0))$ vs qR_g), by normalizing the intensity $I(q)$ with the forward intensity $I(0)$, and q with the radius of gyration (R_g) of the particle [164, 171, 172]. Therefore, the normalized Kratky plot is independent to the particle size and molecular mass, where $I(0)$ is related to the particle molecular mass. This allows to compare the well-folded globular particle, extended folded particle and fully disordered particle. For well-folded globular particles, despite their particle sizes, their normalized Kratky plots exhibit a well-defined bell-curve shape, with a maximum $(qR_g)^2(I(q)/I(0))$ value of 1.104 at qR_g value of $\sqrt{3}$. For the fully disordered particles, their normalized Kratky plots lose the bell-curve shape, and the $(qR_g)^2(I(q)/I(0))$ values increase with the increased qR_g values. For the particles exhibiting the shape in-between, their normalized Kratky plots exhibit a profile that lies in between these two extremes.

2.2.6.4 Ab initio modeling using SAXS

The *ab initio* low-resolution models of the proteins were reconstructed using two programs DAMMIN [173] and DAMMIF [174]. Both DAMMIN and DAMMIF generates the *ab initio* models based on the bead modeling method, where the particle is represented as a collection of dummy atoms within a search volume, and each dummy atom can come from the particle or the solvent. A search volume is a sphere that has the diameter equal to the D_{\max} determined from $P(r)$ function. This program considers the low angle data in a range of $q < 0.2 \text{ nm}^{-1}$. DAMMIF utilizes the simulated annealing to construct the bead model, and disconnected models are immediately discarded from the simulation, which speed up the simulation time. The calculation of theoretical scattering amplitudes of the models is only performed to the connected models, and a final model that has a lowest discrepancy (χ^2) between the theoretical scattering profile and the experimental data is selected.

To validate the reproducibility of the shape reconstructions, at least 10 independent *ab initio* reconstructions were performed for each scattering data, and all the reconstructed models were used as input for the program suite DAMAVER [175]. DAMAVER has a set of programs that align these reconstructed models, select the typical model, and construct an averaged model. The normalized spatial discrepancy (NSD) values are computed between two models,

where the two models are similar if the NSD is close to one. Subsequently, a final mean NSD over all model-pairs and their dispersion of NSD are calculated. The reconstructed models which have a high NSD dispersion are then discarded, while the rest of the models are averaged. The averaged *ab initio* model was selected as the representative model. Beside this, the averaged excluded volume (V_{Ex}) computed from DAMAVER using the DAMMIN/DAMMIF models as input, was used to estimate the molecular mass of the protein (161).

2.2.6.5 Rigid-body modeling using SAXS

Rigid-body modeling of the DENV-2 protease was performed using the program CORAL [176] to build the missing regions of NS2B and NS3 linker in the crystal structure. CORAL allows the translation and rotation of the domains from multiple components of a complex. However, unlike SASREF [177], which explores the conformational space by random rotation of domains and does not account for the missing residues, the N- and the C-terminus of the domains are constrained with a pre-generated library that consists of the self-avoiding DR random loops (length from 5 to 100 residues) in CORAL. If no linker with appropriate length to connect two domains is found, this conformation is rejected. For successful queries, the random loops are inserted into the domains, and its scattering intensity will be computed to compare with the experimental data [176].

Rigid-body modeling for ZIKV NS2B₁₈NS3 was performed using the software BUNCH [177], by docking the high resolution crystallographic structures of single domains (PDB IDs: 5LCO (protease) and 2BMJ (helicase)) against the experimental data. BUNCH performs the modeling using a combination of rigid-body and *ab initio* modeling approach. The high-resolution structures act as the rigid bodies, and BUNCH employs a simulated annealing procedure to search for the optimal orientations of these rigid bodies and connects them with the dummy residue (DR) chains that correspond to the flexible linkers, to simultaneously fit the experimental scattering data. Then the theoretical scattering intensity from the high-resolution domain structures, together with its DR-chain, are computed using CRY SOL [178] and compared with the experimental scattering intensity, generating the discrepancy χ^2 value. The χ^2 value is used to analyze the goodness of fit between the experimental and calculated data, and is defined as:

$$\chi^2 = \frac{1}{N_k - 1} \sum_{j=1}^{N_k} \left[\frac{I(q_j) - cI_{calc}(q_j)}{\sigma(q_j)} \right]^2$$

N_k is the number of experimental point, c is the scaling factor, $I(q_j)$ is the intensity of experimental data at point j , and $I_{calc}(q_j)$ is the intensity of calculated data at point j . A better fit between the experimental and calculated data gives rise to a lower χ value. Therefore, in most of the data fitting and model building programs, χ value is minimized to generate a best solution that explains the experimental data. The rigid-body models are superimposed into the reconstructed low-resolution envelopes using the software SUPCOMB [179], where it aligns two models by minimizing the normalized spatial discrepancy (NSD), which is a qualitative measurement of the similarity between the sets of three-dimensional points from two models.

2.2.6.6 Ensemble modeling using SAXS

For a system containing the homogenous random particles in solution, the observed scattering pattern is the sum of the scattering patterns produced by each particle averaged over all the orientations. Meanwhile, for a system containing flexible particles that adopts different conformations in solution, each conformation will contribute significantly to the final scattering pattern. Therefore, the SAXS data of NS3 were further analyzed by considering an ensemble of conformations. The Ensemble Optimization Method (EOM) [180] quantitatively assesses the flexibility that is present in the particle, by fitting the averaged theoretical scattering pattern from an ensemble of conformations to the experimental scattering pattern. Using the updated EOM version, EOM 2.0 [181], a random pool of 10,000 independent models was generated to approximate the conformational space. In this step, the crystallographic structures of the protease and the helicase were used as the inputs, and the linker region and missing portion of NS2B linked via physical G4SG4 linker in the high-resolution structures were created based on the given protein sequences. For each model presented in the random pool, the theoretical scattering pattern is automatically computed using CRY SOL. After that, genetic algorithm (GA) is performed to select the subsets of conformations from the random pool and compared with the experimental scattering data, to formulate the discrepancy (χ^2) between two scattering patterns. The final ensemble with the minimum discrepancy was then selected. For each protein, the GA was repeated 100-times, and the averaged R_g and D_{max} from each final ensemble were computed and represented as R_g and D_{max} distributions. The final ensemble with the least discrepancy was selected and the conformations were reported as representative models. The distribution of R_g values describes the system flexibility. The broad R_g distribution of selected pool indicates the randomness of the particle, while the narrow R_g distribution suggests that the particle is rigid. Besides this, the distributions with R_g average

values above the R_g average values from the random pool are considered as extended, whereas models with values below the average values as compact. When this is applied to the multi-domain proteins, the former indicates a long inter-domain distance, where the domains are away from each other, and the latter indicates a short inter-domain distance, where the domains are closer to each other. This qualitative characterization of the particles can be quantified as R_{flex} , where the selected ensemble R_{flex} values close to the one extracted from the random pool R_{flex} are considered as an indication of random motion, while the R_{flex} values smaller or larger than the one extracted from the random pool R_{flex} indicate a rigid or flexible system. The complementary metrics R_σ represents the ratio of the standard deviations for the selected pool and random pool. The R_σ value is expected to be less than one when the selected ensemble R_{flex} is smaller than the random pool R_{flex} .

2.2.7 Ligation based cloning

Shown below is the schematic representation of ligation dependent cloning utilized in this study:

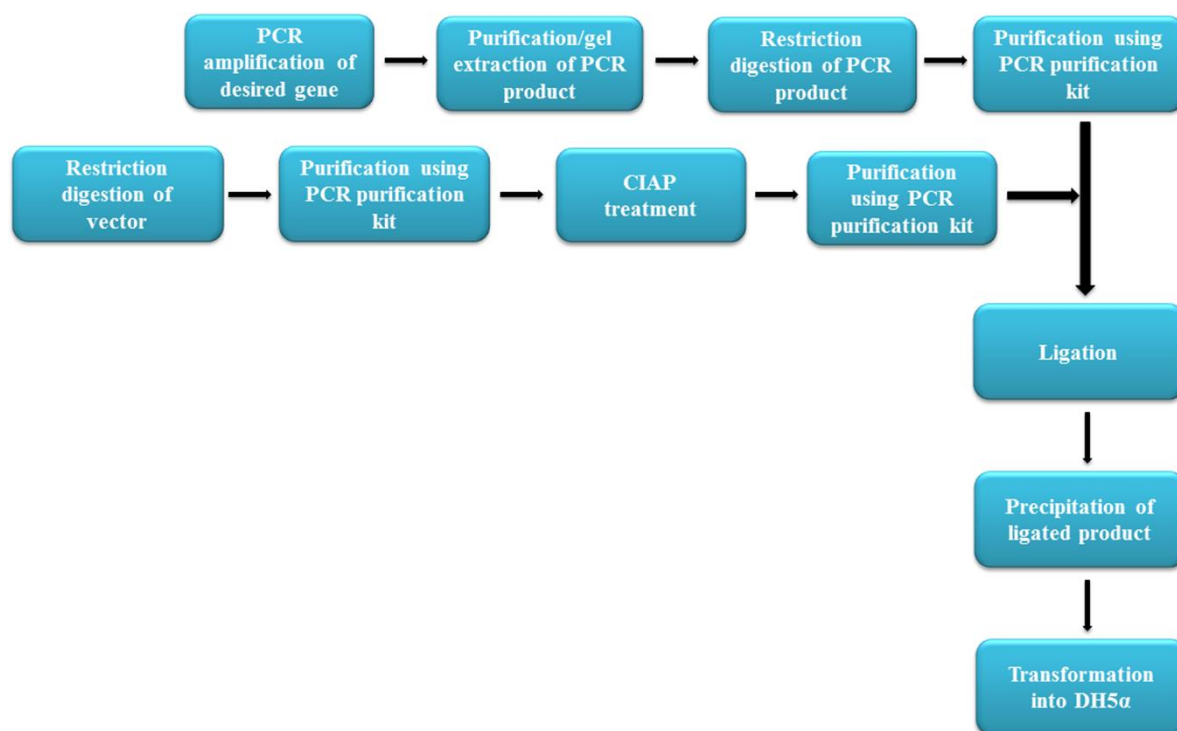


Figure 2.2.5: Schematic representation of ligation dependent cloning.

2.2.7.1 Polymerase chain reaction

The PCR mix was set up in a total volume of 50 μ l with the components listed in table 2.2.31. All the components were kept on ice all the time.

Reagents	Amount
Buffer ‡	10 µl/5 µl*
dNTP's (10 mM)	1.0 µl/1.5 µl*
Forward primers (100 µM)	0.3 µl
Reverse primers (100 µM)	0.3 µl
Template (plasmid DNA) (10 ng/µl)	1.0 µl
MilliQ water	Make up to 50 µl
DNA polymerase ‡	1.0 µl

Table 2.2.31: Components of PCR reaction mix. (‡ - 10x Pfu buffer (with 20 mM Mg₂SO₄) was used with Pfu DNA polymerase and 5x KAPA HiFi-buffer (with 10 mM MgCl₂) was used KAPA HiFi-DNA polymerase. *- 10 µl of buffer and 1.0 µl of dNTP mix was used in case of Pfu DNA polymerase while 5 µl of buffer and 1.5 µl of dNTP mix was used in case of HiFi DNA polymerase).

This set-up was followed for all constructs generated by this approach of cloning. The type of polymerase, DNA template, as well as primers used for each construct are defined when the constructs are described later. Once the reagents were added to the PCR tubes, they were mixed by pipetting and thereafter placed into a thermocycler.

The thermocycler lid temperature was set to 99 °C and pre-heated before reaction tubes were placed into the machine. The thermocycler program for all constructs is described as:

Cycle steps	Temperature	Time
Initial denaturation	95 °C	3 min
Cycle denaturation	95 °C/98 °C ‡	20 secs
Annealing	Dependent on primer T _m	15 secs
Extension	72 °C	2 min/kb or 30 sec/kb *
Final extension	72 °C	5 min
End	4 °C	

Table 2.2.32: PCR thermocycler program (‡ - 95 °C in case of Pfu DNA polymerase and 98 °C in case of KAPA HiFi-DNA polymerase. *- 2 min/kb used in case of Pfu DNA polymerase and 30 sec/kb used in case of KAPA HiFi-DNA polymerase).

The annealing temperature was set at approximately 5 °C lower than the melting temperature (T_m) of the primers. The T_m are mentioned when the primers are described. The 25 cycles were performed for amplification, unless otherwise specified.

2.2.7.2 Agarose gel electrophoresis

To visualize the amplified PCR product, as well as for gel extraction, DNA samples were mixed with 1/6 volume of 6x loading dye (Thermo Scientific) and loaded onto a 1.5 % agarose gel. The gel was prepared with 1x TAE buffer (40 mM Tris, 1% (v/v) acetic acid, 1 mM EDTA, pH 8.0). The gel was run at 100 V. Bands were visualized under UV-light after staining with ethidium bromide for 15 minutes.

2.2.7.3 Restriction digestion of insert or vector

For extraction of DNA from the agarose gel, the desired band on the agarose gel was excised with a scalpel under UV-light. DNA was then extracted using the QIAquick gel extraction kit following manufacturer's protocol. DNA was eluted in 30 μ l of EB buffer (10 mM Tris-HCl, pH 8.5; Qiagen).

2.2.7.4 Restriction digestion of insert/vector

The required amount (100 ng - 1 μ g) of DNA was digested with restriction enzymes. The reaction was set up in 50 μ l, with the appropriate amount of restriction enzymes (2-5 U of restriction enzyme/ μ g of DNA to be digested). Double digestion was set with buffers compatible with both restriction enzymes. The reaction mixtures were then incubated at 37 °C for 1 hour when FastDigest enzymes (Thermo Scientific) were used and overnight when normal restriction enzymes (NEB) were used.

DNA was purified using QIAquick PCR purification kit following the manufacturer's protocol. DNA was eluted in 30 μ l of EB buffer.

2.2.7.5 CIAP treatment of digested vector

To prevent recircularization of linearized vector, calf intestinal alkaline phosphatase (CIAP) treatment was performed. The reaction was set up in 50 μ l as follows:

Reagents	Amount
Vector (V) (100 ng)	1 μ l
Alkaline phosphatase 10X reaction buffer	5 μ l
CIAP enzyme (1 U/ μ l)	1 μ l
MilliQ water	up to 50 μ l

Table 2.2.33: CIAP treatment protocol

The reaction mix was incubated at 37 °C for 30 minutes followed by purification of the dephosphorylated vector using QIAquick PCR purification kit.

2.2.7.6 Ligation

The ligation of the desired insert and vector backbone was set up in 10 μ l as follows:

Reagents	Amount
Ligase buffer (10x)	1 μ l
Dephosphorylated vector	60 ng
Insert	variable (3 or 5 molar excess of vector)
T4 Ligase (40,000 U/ml)	0.5 μ l
MilliQ water	up to 10 μ l

Table 2.2.34: Components of the reaction mix for setting up ligation

The ligation mix was incubated at room temperature for 2 hours before proceeding to DNA precipitation of ligated product. After ligation DNA precipitation was performed as explained later.

The full-length wild-type (WT) *E. faecalis* AhpC was cloned by Dr. Wilson Nartey from our lab. This *E. faecalis* AhpC gene in pET9d plasmid was then used as a template to clone the C-terminally truncated *E. faecalis* AhpC (*EfAhpC*₁₋₁₇₂). *EfAhpC*₁₋₁₇₂ was cloned using the method described above using *Pfu* DNA polymerase with the primers described in table 2.2.35. The restriction sites *Nco*I and *Sac*I are underlined. The annealing temperature was set at 48.8 °C.

Primers	Oligonucleotide sequence 5'-3'	T _m (°C)
Forward primer	CAT <u>GCC</u> ATG <u>GCA</u> ATG AAT TTA ATT AAT CAA	49.7
Reverse primer	TTC <u>GAG CTC</u> TTA TGG TTG CCA ATT GGC TGG	57.9

Table 2.2.35: Primers for amplification of *EfAhpC*₁₋₁₇₂

The N-terminal triple mutant was cloned using the primers listed in table 2.2.36. This *E. faecalis* AhpC gene in pET 9d plasmid was then used as a template. *Pfu* DNA polymerase was used for the amplification. The restriction sites *Nco*I and *Sac*I are underlined.

Primers	Oligonucleotide sequence 5'-3'	T _m (°C)
Forward primer	CTG <u>CCC</u> ATG <u>GCA</u> ATG TCT TTA ATT AAT CAA AAA CTT TTT CCA TTT GAA TGT CAG GCC TAT CAC	66.8
Reverse primer	CCG <u>GAG CTC</u> TTA AAT TTT ACC AAC TAA ATC TAA GCT TGG CGC AAT TGT TTC TTC TCC TGG TTG	67.9

Table 2.2.36: Primers for amplification of *EfAhpC*_{N2S,D10P,D14Q}

2.2.8 Ligation independent cloning

Shown below is the schematic representation of ligation independent cloning utilized in this study:

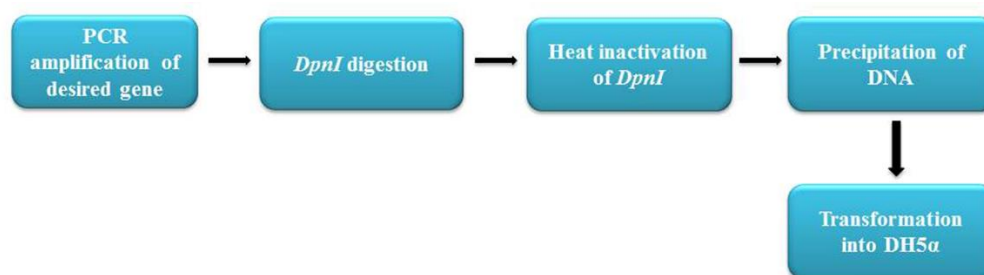


Figure 2.2.6: Schematic representation of ligation independent cloning.

2.2.8.1 PCR amplification

The PCR reaction mix was set up in 50 μ l as described in section 2.2.7.1; table 2.2.31. KAPA HiFi DNA polymerase was used in this protocol and the PCR thermocycler program was modified according to the program suited to this polymerase (refer Table 2.2.32). KAPA HiFi polymerase was used in the reaction mix and the annealing temperature was set at 62 °C for generating all constructs cloned by this method.

2.2.8.2 *DpnI* digestion

To remove the methylated DNA template, *DpnI* digestion was set up in a total reaction volume of 50 μ l.

Reagents	Amount
PCR product	variable
Cut Smart buffer (10x)	1.0 μ l
<i>DpnI</i> (20,000 U/ml)	0.3 μ l
MilliQ water	up to 50 μ l

Table 2.2.37: Reaction mix for *DpnI* digestion

The reaction mix was incubated at 37 °C for 1 hour and subsequently inactivated at 80 °C for 20 minutes, followed by DNA precipitation.

Site-directed mutagenesis of *E. faecalis* AhpC was done using this ligation independent method. The primer for cloning each of the mutants are described in table 2.2.38. The mutant gene sequences are underlined. The double cysteine mutant *EfAhpC*_{C13N,C66V} was created in two steps using primers used to generate the *EfAhpC*_{C13N} and *EfAhpC*_{C66V} mutants.

Mutant	Oligonucleotide sequence	T _m (°C)
<i>EfAhpC</i> _{C13N}	F 5'- GAC TTT GAA <u>AAT</u> GAT GCC TAT CAC GAT GGT -3'	59
	R 5'- ATA GGC ATC <u>ATT</u> TTC AAA GTC AAA AAG TTT -3	55
<i>EfAhpC</i> _{C66V}	F 5'- GAA TTA AAT <u>GTT</u> GAG GTT TAC TCT GTT TCA GAA GAT -3'	57.5
	R 5'- GTA AAC CTC <u>AAC</u> ATT TAA TTC CTG TAA ATG CGC ATA -3	59.2
<i>EfAhpC</i> _{V78T}	F 5'- AGT CAT TAT <u>ACG</u> CAT AAA GCA TGG GCT GAC GCT ACA -3'	64.4
	R 5'- TGC TTT ATG <u>CGT</u> ATA ATG ACT ATC TTC TGA AAC AGA -3	59.1
<i>EfAhpC</i> _{N169K}	F 5'- GTA GCA GAA CAT GGT GAT AAA GTT TGT CCA GCC <u>AAA</u> TGG CAA CCA GG -3'	65
	R 5'- GGT TGC CAT <u>TTG</u> GCT GGA CAA ACT TTA TCA CCA TGT TCT GCT AC -3	65

Table 2.2.38: Primers for amplification of *EfAhpC* mutants

2.2.9 DNA precipitation

90 µl of MilliQ water was added to the ligation reaction mix/ *DpnI* to make a final volume of 100 µl. 1 ml of butanol was added into the mixture and mixed thoroughly. DNA was precipitated by centrifuging at 13,000 x g (using Eppendorf Mini-centrifuge) for 10 minutes and the supernatant was discarded gently. The pellet was resuspended in 200 µl of 70% ethanol and centrifuged again for 5 minutes. The ethanol was then discarded, and the DNA pellet was dissolved in 10 µl of sterile water.

2.2.10 Transformation into DH5α cells

5-7 µl of 10-80 ng of the precipitated plasmid was added to *E. coli* DH5α cells which were thawed on ice. Cells were incubated on ice for 5 minutes and subsequently transferred to pre-chilled electroporation cuvettes (0.2 cm, BioRad) without introducing bubbles. The cuvettes. Electroporation was performed with the MicroPulser electroporator (BioRad) at 2.5 kV. Luria Bertoni (LB) media was immediately added to the electroporation cuvette and the cells transferred to a 2 ml microcentrifuge tubes and cells recovered by shaking at 180 rpm (Infors HT Minitron shaker) at 37 °C for an hour. The cells were then plated on to LB agar plates containing kanamycin (for constructs cloned in pET9d) and incubated overnight.

2.2.11 Isolation of plasmid

Plasmid DNA was isolated from the positive clones identified on the LB agar plates and grown overnight in LB media. The plasmid was extracted using QIAprep Spin miniprep plasmid kit following the manufacturer's protocol. The plasmids were eluted in 50 µl of EB

buffer. To ensure the correct sequence of the desired genes was present, the plasmids were sequenced.

2.2.12 Recombinant gene expression

Sequence verified plasmids of the different *EfAhpC* gene constructs were transformed into *E. coli* BL21 (DE3) cells for protein production. Electroporation was performed according to the protocol described in above in section 2.2.4. Once the cells were revived after transformation, they were plated on to LB agar plates with kanamycin.

2.2.13 Induction test

Reagents for 4x lysis buffer:	250 mM Tris-HCl, pH 6.8
	9.2% SDS
	40% Glycerol
	0.2% Bromophenol blue
	1 mM DTT

3-4 single colonies were picked from the LB agar plate. Cells were grown to OD₆₀₀ of 0.6 at 37 °C with shaking at 180 rpm. The colonies were then induced with IPTG with a final concentration of 1 mM. The induced culture was incubated at 37 °C with shaking at 180 rpm for 4 hours. Respective controls were left uninduced under identical conditions. Cells were then pelleted down and resuspended with 60 µl of 2x lysis buffer in the presence of 1 mM DTT. Resuspended cells were heated at 95 °C for 5 min and 15 µl of each sample was loaded onto a 17% SDS polyacrylamide gel.

2.2.14 Recombinant protein production and purification

For large scale production of recombinant proteins, an overnight culture of *E. coli* BL21 (DE3) cells was prepared in LB media supplemented with kanamycin (30 µg/ml). This overnight culture was subsequently scaled up to a starting OD₆₀₀ of 0.1. Once the cultures reached an optical density of OD₆₀₀ of 0.6–0.7, IPTG (1 mM) was added to induce protein production. Following incubation of the culture at 37 °C, with shaking at 180 rpm for 4 hours, the cells were harvested by centrifugation at 7,500 x g for 15 minutes at 4 °C on Beckman Coulter Avanti J26X centrifuge. The cell pellet was frozen in liquid nitrogen and stored in -80 °C until further use.

For purification, cells producing the *EfAhpC* recombinant proteins were lysed on ice by sonication with an ultrasonic homogenizer (Bandelin, KE76 tip) for 3 x 1 min in buffer A.

Buffer A (Lysis buffer)	50 mM Tris-HCl, pH 7.5
	200 mM NaCl
	2 mM Pefabloc ^{SC}
	0.8 mM DTT

After sonication, the cell lysate was centrifuged at 10,000 x g for 35 min at 4 °C. The resulting supernatant was passed through a filter (0.45 µm; Millipore). Subsequently, the purification of His₆-tagged recombinant proteins was carried out in two steps. The first step of purification was immobilized metal affinity chromatography (IMAC). Buffers used for this step are enlisted below:

Buffer B (wash buffer)	Buffer C (elution buffer)
50 mM Tris-HCl, pH 7.5	50 mM Tris-HCl, pH 7.5
200 mM NaCl	200 mM NaCl
	Imidazole gradient (0-500 mM)

The filtered supernatant obtained after centrifugation was incubated with Ni²⁺-NTA resin that was pre-equilibrated in buffer B, to allow His₆-tagged protein to bind to the matrix. The resin was incubated for 1 hour at 4 °C by mixing on a sample rotator (Neolab). The Ni²⁺-NTA was thoroughly washed with 10 column volumes of buffer B. The bound tagged protein was subsequently via batch elution in buffer C. The fractions containing *EfAhpC* recombinant proteins were identified by SDS-PAGE. The fractions were then pooled together and concentrated.

The concentrated protein was then subjected to next step of purification. Buffers for the size exclusion chromatography are enlisted below:

Buffer D	Buffer E
50 mM Tris-HCl, pH 7.5	50 mM Tris-HCl, pH 7.5
200 mM NaCl	500 mM NaCl

In this step, a Superdex 200 HR 10/30 column (GE Healthcare) was used and the protein eluted with buffer D. To obtain the reduced form of protein, 1 mM DTT was added to buffer D during gel filtration. The purity and homogeneity of *EfAhpC* sample was analyzed by loading onto a 17% SDS-PAGE gel and stained with Coomassie Brilliant Blue R250.

When *EfAhpC*₁₋₁₇₂ was purified for crystallization purpose, the protein was eluted from Superdex 200 HR 10/30 column by buffer E. For obtaining the reduced protein, DTT was added to the buffer for gel filtration column to a final concentration of 1 mM

Protein concentrations were determined at 280 nm in a spectrophotometer (BioSpec-Nano, Shimadzu).

2.2.15 Dynamic light scattering

Dynamic light scattering of the reduced and oxidized form of *EfAhpC* and its mutants were carried out using the Malvern Zetasizer Nano ZS spectrophotometer. The samples were measured in a low-volume quartz batch cuvette (ZEN2112, Malvern Instruments) using 12 μ l of 1 mg/ml of the respective protein solution equilibrated with buffer A. After 60 s of equilibration time, the backscattering at 173° was detected for all proteins. Scattering intensities were analyzed using the instrument software to calculate the hydrodynamic diameter (D_H), size, and volume distribution.

2.2.16 Negative staining in electron microscope

The data was collected and analyzed by Dr. Phat Van Dip for our group. For negative staining image collection, 2.5 μ l (25 μ g/ml) of oxidized *EfAhpC* was applied onto copper grids coated with a thin carbon film. The grids were stained with 2% uranyl acetate (v/w). Electron micrographs were recorded on a FEI Tecnai spirit (T12) electron microscope at a voltage of 120 kV and 66,350x magnification. Images were recorded on a 4K eagle camera (FEI). Individual *EfAhpC* particles were selected semi-automatically, aided by the e2boxer tool, in Swarm mode, from EMAN2 [182] software package. In total, 30 and 25 images were collected for oxidized and reduced WT *EfAhpC*. A total of 2090 and 2623 particles were selected for oxidized and reduced *EfAhpC*, respectively. The 2D class averaging was performed using EMAN2 [182] and RELION software [183].

2.2.17 NADH-dependent peroxidative assay

The NADH-dependent peroxidase assay was monitored at 340 nm following the decrease in NADH absorbance at a stopped flow spectrophotometer Applied Photophysics SX20 (Applied Photophysics, UK) (Fig. 2.2.7A). In the stopped flow instrument, solutions of two from two drive syringes are forced to mix in a mixing chamber. After mixing, the solution then passes through a cell and the flow stopped where the change in absorbance is probed (Fig. 2.2.7B). The change in absorbance is recorded as a function of time.

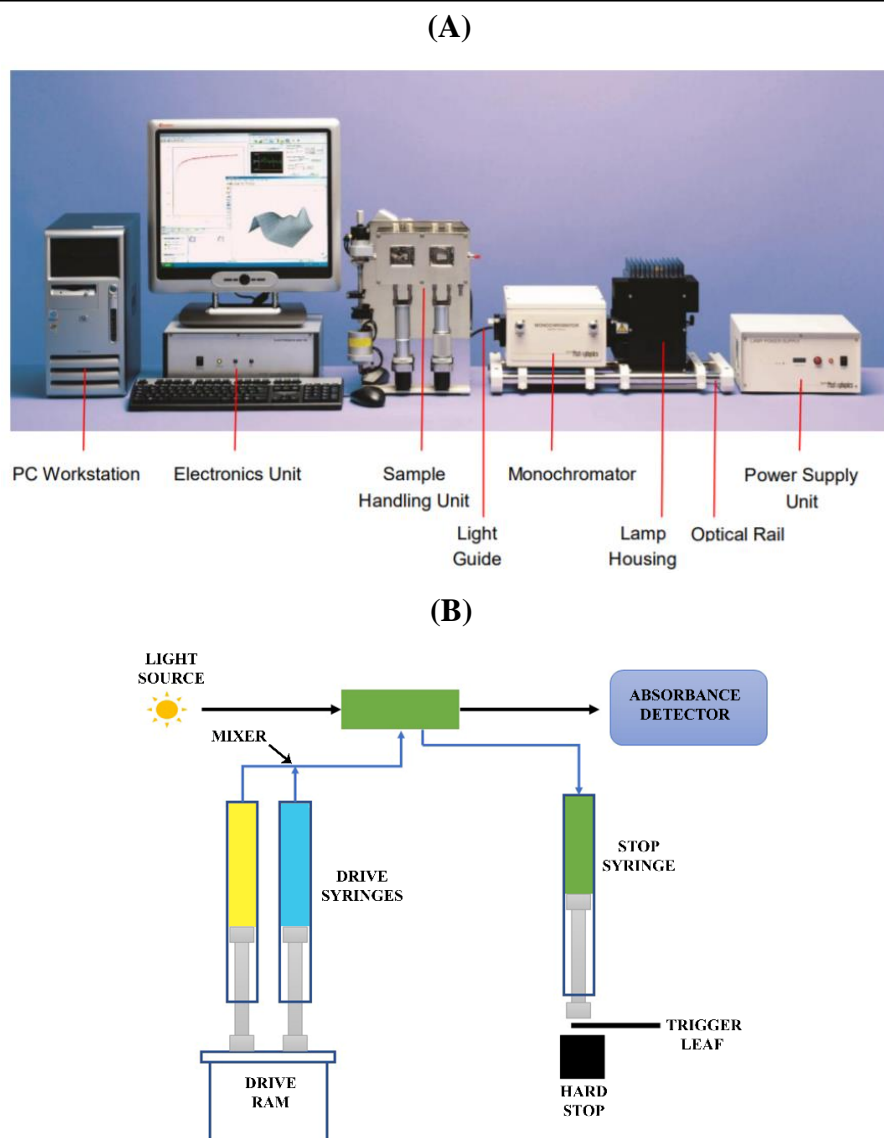


Figure 2.2.7: Stopped flow instrument. (A) The instrument set-up of stopped flow instrument, Applied Photophysics SX20. The instrument has a lamp housing from where the monochromator selects a single wavelength to pass into the sample handling unit. The change in absorbance is then detected at the electronic unit. (B) The cartoon representation of the sample handling unit. The two drive syringes contain the two reactants whose activity is to be measured; the reactants are mixed in the mixer and flow stopped by stop syringe and the reaction is then measured as an absorbance change as a function of time.

In the stopped flow set-up, one of the drive syringes contain 100 mM HEPES; pH 7.5, 200 mM ammonium sulphate, 1 mM EDTA, 30 μ M *EfAhpC*, 0.2 μ M of WT *EfAhpF* and 2 mM H_2O_2 . The other syringe has 600 μ M NADH. When mixed in the mixer chamber, the concentrations of all the components were halved. The data were then collected in triplicate and the averaged data are presented in this study.

2.2.18 Ferrous oxidation xylenol orange assay

Peroxide decomposition catalyzed by *EfAhpC* in the presence of WT *EfAhpF* was measured using the ferrous oxidation xylenol (FOX) assay. The assay is based on the principle

that peroxides like H_2O_2 oxidize Fe^{2+} to Fe^{3+} at an acidic pH. The Fe^{3+} then forms a colored adduct with xylenol orange which can be measured spectrophotometrically at 560 nm. AhpC reduces H_2O_2 to H_2O and the reduction in H_2O_2 concentration can be monitored over time by monitoring the change in absorbance of the colored adduct. A standard curve of absorbance at known concentrations of H_2O_2 can be plotted. The unknown concentration of H_2O_2 in a reaction can then be determined from this curve.

The reaction mixture (100 μl total volume) contained 50 mM HEPES buffer pH 7.5, 15 μM *EfAhpC* or mutant protein, 100 μM H_2O_2 , 250 μM NADH and 0.4 μM WT *EfAhpF*. At several time intervals, 10 μl aliquots were removed from the reaction mixture, added to 190 μl of FOX-reagent and incubated at room temperature for 30 min. The sample absorbance was measured at $\lambda = 560$ nm to quantitate the residual peroxide. The assay was done in triplicates, and the data were averaged and plotted.

2.2.19 Crystallisation of *EfAhpC*₁₋₁₇₂

Crystals drops of recombinant *EfAhpC*₁₋₁₇₂ in both oxidized and reduced condition were set up using vapour diffusion in Hampton Research crystal screen HR-110 and HR-112, Emerald Biosystems Wizard screen I and Wizard screen II, and Molecular Dimensions Morpheus Screens I and II. Hanging drops were set up on siliconized cover slip with 2 μl droplet size per well in 1:1 ratio of AhpC in buffer (50 mM Tris/HCl, pH 7.5, 500 mM NaCl and 1 mM DTT) and precipitant solution at 18 °C. Crystals were obtained only for the reduced condition, where the buffer for protein was supplemented with 1 mM DTT. Initial microcrystals were observed in Molecular Dimensions Morpheus screen I and Emerald Biosystems Wizard Screen I. Crystals were first optimised for the Morpheus screen. Large crystals were obtained with optimization only in Morpheus screen I #32. Once protein crystals were confirmed, the crystallisation condition was further optimised by systematically performing a grid screen by varying the concentrations of precipitant, salt, buffer pH and protein concentration. Cryo-protectant solution was also optimised by varying the percentages of glycerol. The final optimised crystals were quickly dipped in cryoprotectant solution and were flash-frozen in liquid nitrogen at 100 K. Optimised *EfAhpC*₁₋₁₇₂ crystals were tested in-house at 100 K on a Rigaku RAXIS IV image plate detector with a Rigaku RA-Micro 7 HFM rotating copper anode generator (Rigaku/MSD). The optimized crystals diffracted poorly. Hence, we proceeded with optimizing the crystals in Wizard Screen I.

Well-formed crystals were obtained in Wizard screen I #25 (0.2 M sodium chloride, 0.1 M Tris, pH 8.5 and 30% (v/v) PEG 400). This condition was further optimised via Hampton Research additive screen after confirming that the crystals were protein crystals. The final crystals obtained showed reflections in the diffraction image using the in-house machine (Rigaku) using 5 mg/ml protein concentration. Cryo-protectant solution was also optimised by testing different percentages of glycerol. The crystals were quickly dipped in cryoprotectant solution and were flash-frozen in liquid nitrogen at 100 K. All the diffraction data were indexed, integrated and scaled using the HKL2000 suite program and were kindly guided by Dr. Asha M. Balakrishna from our group.

2.2.19.1 Data collection and structure determination

Together with Dr. Asha M. Balakrishna, single-wavelength datasets for *EfAhpC*₁₋₁₇₂ crystals were collected at 140 K on beamline 13B1 at the National Synchrotron Radiation Research Centre (NSRRC, Hsinchu, Taiwan) using the ADSC Quantum 315 CCD detector. Staff at beamline 13B1 (NSRRC) kindly provided technical assistance. All the diffraction data were indexed, integrated and scaled using the HKL2000 suite program [156]. All the crystals belong to the tetragonal space group C2 and have similar unit cell parameters.

2.2.19.2 Crystal structure determination of *EfAhpC*₁₋₁₇₂

Initial phases for AhpC structure determination from *E. faecalis* were obtained by molecular replacement method using the *E. coli* AhpC structure (PDB ID: 4O5R; [116]) as model using the program PHASER [157]. Subsequent rigid body and restrain refinement was done with Refmac5; the starting model was improved manually using Coot [184]. Refinement was done until convergence and the geometry of the final model was validated with MolProbity [185]. Structural comparison analysis was carried out by the SUPERPOSE program [186] as included in the CCP4 suite (Fig. 2.2.8).

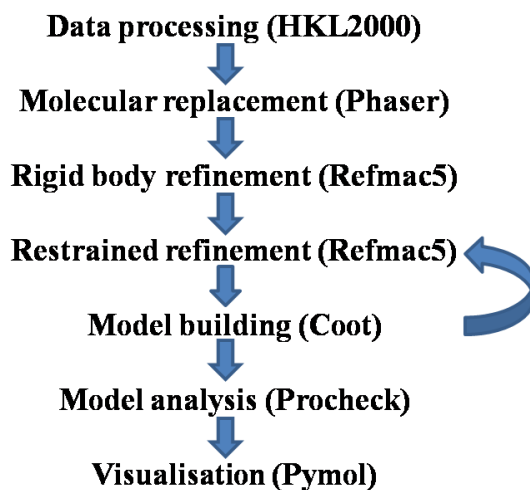


Figure 2.2.8: Flow chart for solving the crystal structure of *EfAhpC1-172*.

2.2.20 Protein sequence analysis

Multiple sequence alignments were performed using an interactive sequence alignment viewing, editing and analysis tool, Jalview 2.9 [187]. The protein sequences were retrieved from NCBI (Ref) and UniProt database [188]. The sequences were subsequently submitted to the web-server, Clusatal Omega [189] via Jalview 2.9, to perform multiple sequence alignment. The results were viewed in Jalview 2.9, and the sequence similarity was coloured according to the percentage of the residues in each column that agree with the consensus sequence.

3. Results – Structural and Functional Characterization of Flaviviral NS3

3.1 Solution studies of DENV-2 NS3

The atomic resolution structure of DENV-4 NS2B₁₈NS3 (PDB ID: 2VBC;[81]) revealed an extended conformation. A second conformation of DENV-4 NS2B₁₈NS3 (PDB ID: 2WHX, [38]) have been trapped using crystallography technique, showing these two protein domains adopts certain flexibility through its linker region (Fig 1.1.12; section 1.1.6). However, how the communication between the two domains of NS3 occur and how they are arranged relative to each other in solution is still elusive, as well as the degree of flexibility and dynamics of these domains inside NS3. To gain more insight into how these two protein domains undergo conformational changes in solution, SAXS experiments were carried out. Since DENV-2 is one of the more predominant serotypes in Southeast Asia, it has been selected for this thesis.

3.1.1 Cloning of DENV-2 NS2B₁₈NS3 and purification of recombinant protein

To study the DENV-2 NS2B₁₈NS3, the gene encoding the entire protein was constructed by overlap PCR of two gene fragment. One fragment encodes 18 amino acid residues for NS2B and the nonapeptide linker, and the other fragment coding for the 618 amino acid residues for DENV-2 NS3. The amplified product represented in figure 3.1.1A (~1900 bp), was double digested with *NcoI* and *HindIII* to ensure insertion of the gene in the correct orientation into the modified pET32b plasmid [81]. Once DNA sequencing confirmed the correct plasmid sequence, it was transformed in to *E. coli* BL21-CodonPlus (DE3)-RIL cells for protein production. Overproduction of recombinant DENV-2 NS2B₁₈NS3 was identified in the induced lane on the SDS-PAGE gel (Fig. 3.1.1B).

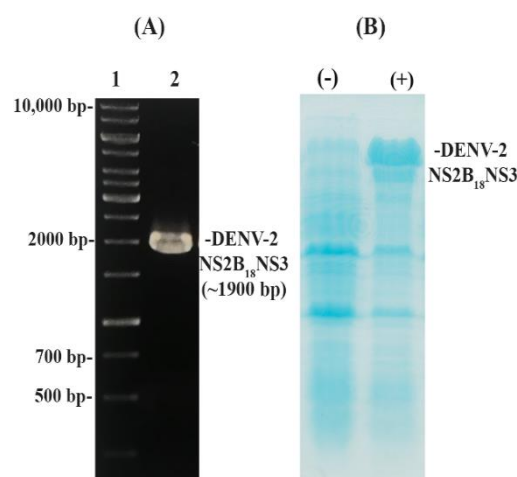


Figure 3.1.1: Amplification of gene encoding DENV-2 NS3 and recombinant DENV-2 NS2B₁₈NS3 production. (A) Overlap PCR amplified product (lane 2) and DNA ladder in lane 1. DNA sample of 5 μ l was loaded on an agarose gel (1.5%). (B) SDS-PAGE (17%) showing the expression of recombinant protein. Protein production was induced with 1 mM IPTG. (+) and (-) indicates protein obtained from cells induced with IPTG and uninduced cells, respectively

This protein when purified by affinity chromatography eluted from the Ni-NTA column in an imidazole gradient from 35- 70 mM (Fig. 3.1.2). These fractions were then pooled and dialysed overnight against buffer B (materials and methods, section 2.2.18) and the thioredoxin-tag was cleaved by thrombin. When the tag-free protein was subjected to a second round of affinity chromatography, the protein was identified in the flow-through, while the cleaved tag eluted from the column in an imidazole concentration of 500 mM. The recombinant

protein and the tag were identified by a 17% SDS-PAGE gel (Fig. 3.1.2). The flow-through was then concentrated and applied to a size-exclusion column (SEC). A single prominent peak was observed on the elution profile (Fig. 3.1.2). When a sample of this peak was loaded on a 17% SDS-PAGE gel, a single band of DENV-2 NS2B₁₈NS3 was observed at approximately ~70 kDa (Fig. 3.1.2). Fractions from the shaded region of the peak, which corresponds to the top 15% of maximum UV absorbance, were pooled and concentrated for SAXS experiments.

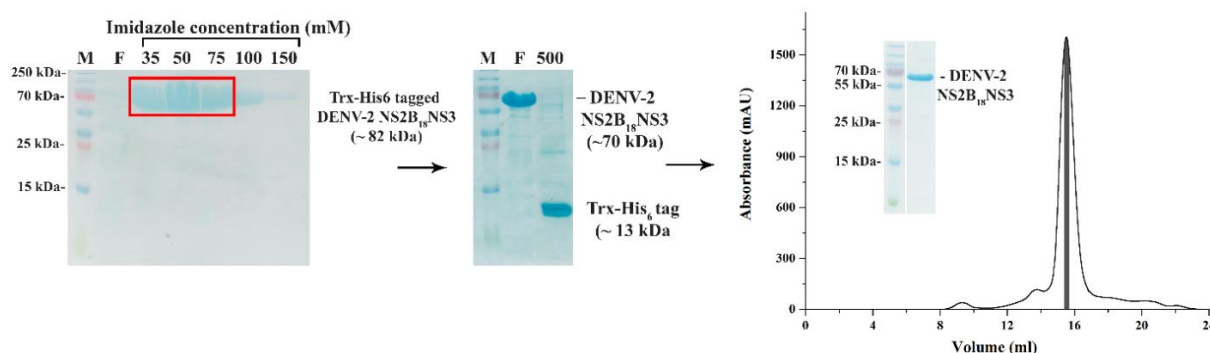


Figure 3.1.2: Purification of DENV-2 NS2B₁₈NS3. The protein was subjected to two rounds of affinity chromatography and the protein was eluted by batch method using an imidazole gradient. The first SDS-PAGE gel shows the imidazole concentrations at which the tagged protein eluted. (M= Protein marker; F= flow-through). The red box indicates the fractions pooled for overnight cleavage of tag. After a second affinity chromatography, the protein eluted in the flow-through (F). On application to SEC, a single peak was observed in the chromatogram. Shaded grey region corresponds to 15% of maximum UV absorbance, from where fractions were pooled for further studies. *Inset:* Pure and homogeneous DENV-2 NS2B₁₈NS3.

3.1.2 Analysis of primary SAXS data for DENV-2 NS2B₁₈NS3

To avoid freezing and thawing of the protein, SAXS measurements of recombinant DENV-2 NS2B₁₈NS3 were performed with freshly purified protein. The protein was concentrated to 1.2 mg/ml, 3.3 mg/ml, and 5.4 mg/ml, respectively. The Guinier plot at very low q was linear for all the protein concentrations, indicating absence of protein aggregates (Fig. 3.1.3A). To improve the data quality, the three datasets were merged. The merged scattering curve was used for primary data analysis. The R_g value, which is related to the slope of the Guinier plot, was calculated to be 3.18 ± 0.05 nm. This value equates with the R_g value of 3.25 ± 0.02 nm derived from the $P(r)$ function (table 3.1.1). The asymmetric nature of the distance-distribution curve with a tail is indicative of an elongated molecule in solution. The estimated maximum particle dimension (D_{max}) was 10.3 ± 1 nm (Fig. 3.1.3B) (table 3.1.1).

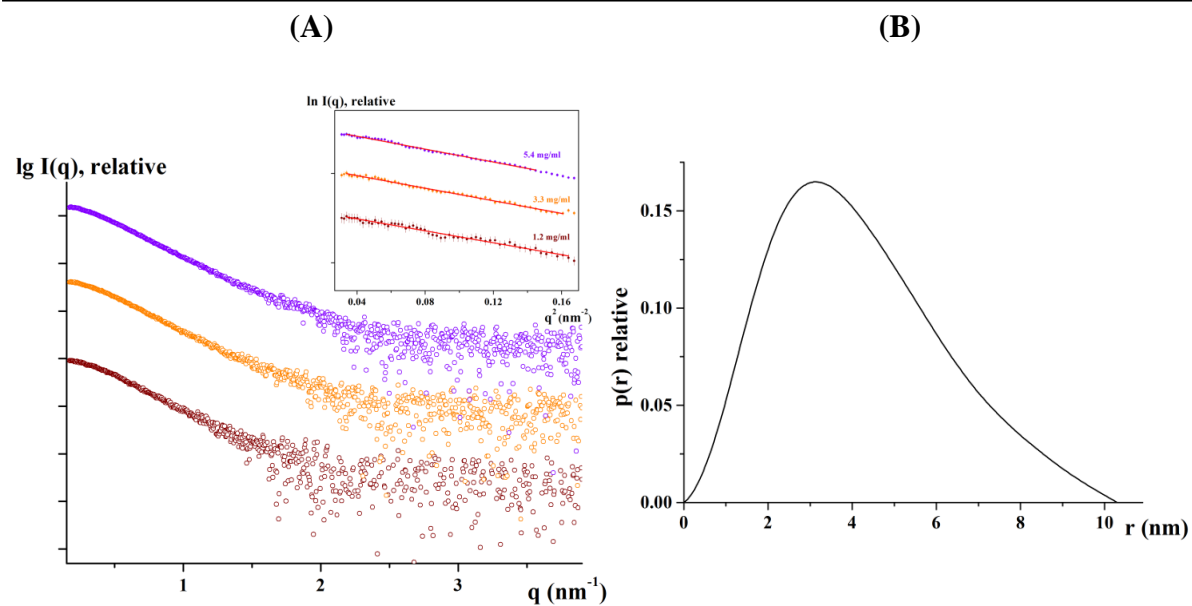


Figure 3.1.3: SAXS studies of NS2B₁₈NS3. (A) Scattering pattern (○) of DENV-2 NS2B₁₈NS3 at 1.2 mg/ml (brown), 3.3 mg/ml (orange) and 5.4 mg/ml (purple). *Inset:* A linear Guinier plot at low angles indicated protein sample were not aggregated. (B) Pair-wise distance distribution $P(r)$ curve was generated after merging all three datasets. The curve was bell-shaped but slightly skewed to the right with a tail indicating an elongated molecule in solution.

Protein sample	DENV-2 NS2B ₁₈ NS3
Concentration range (mg ml ⁻¹)	1.20 – 5.40
Structural parameters†	
$I(0)$ (cm ⁻¹) [from $P(r)$]	103.1 ± 0.76
R_g (nm) [from $P(r)$]	3.25 ± 0.02
$I(0)$ (cm ⁻¹) (from Guinier)	102.2 ± 1.20
R_g (nm) (from Guinier)	3.18 ± 0.05
D_{max} (nm)	10.3 ± 1
Porod volume estimate (nm ³)	~98
DAMMIF excluded volume (nm ³)	~134
Dry volume from sequence (nm ³) ‡	~87.2
Molecular mass determination	
Calculated monomeric MM (kDa)	~72.0
Molecular mass MM (kDa) [from <i>Porod invariant</i>]	66.5 ± 6.7
Molecular mass MM (kDa) [from <i>excluded volume</i>]	70.3 ± 7.0

Table 3.1.1: Overall structural parameters for DENV-2 NS2B₁₈NS3

3.1.3 Construction of *ab initio* envelopes for DENV-2 NS2B₁₈NS3

The low-resolution shapes of NS2B₁₈NS3 were reconstructed *ab initio* using DAMMIF. The program was run twenty times independently to ascertain reproducibility. After comparing the individual models using the program DAMAVER, the final averaged model with a NSD (normalized spatial discrepancy) of 0.709 ± 0.05 has been reported.

SUPCOMB was used to align the the *ab initio* envelope to the first crystallographic structure of DENV-4 NS2B₁₈NS3 (PDB ID: 2VBC;[81]). The two 3D models aligned, indicating that the solution shape agrees with the crystal structure (Fig. 3.1.4A). Apart from the above-mentioned crystal structure of DENV-4 NS2B₁₈NS3, two other crystallographic structures have been reported. These two structures have two different conformations. The 2WZQ (E₁₇₃GP₁₇₄ mutant protein;[38]) adopts conformation I while 2WHX (WT NS2BNS3-ADP-Mn²⁺ complex; [38]) adopted conformation II. In the first conformation, the protease is positioned near the of the ATP-binding site the helicase. In comparison, the protease is rotated by 161° with respect to the helicase in conformation II. To elucidate which conformation is dominant in solution, CRY SOL was used to evaluate the fit of the three crystal structures with the solution experimental data. The theoretical scattering patterns extracted from 2VBC, 2WZQ and 2WHX fit nicely to the experimental data of DENV2 NS3, with the lowest χ^2 value observed for 2WHX ($\chi^2 = 0.596$) (Fig. 3.1.4B).

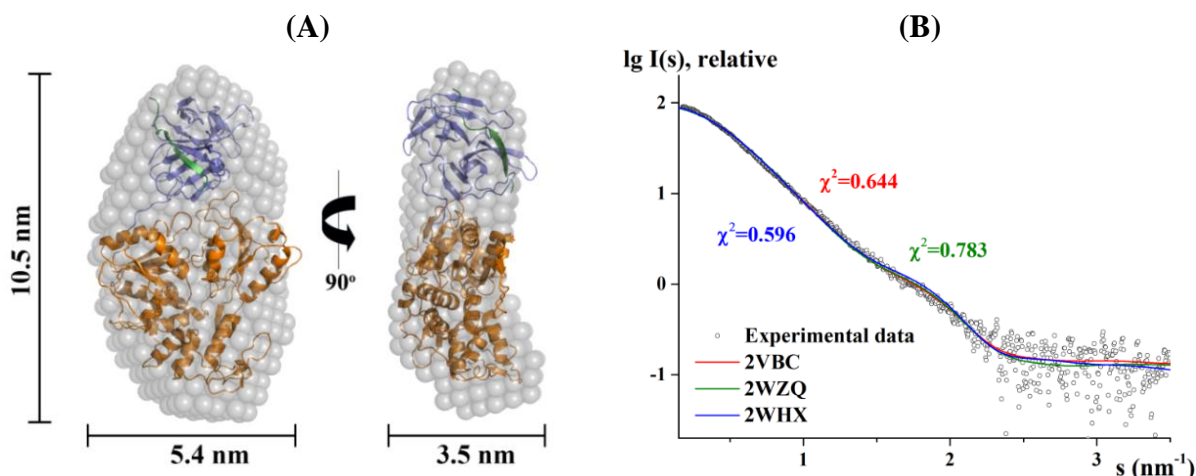


Figure 3.1.4: *Ab initio* shape reconstruction of DENV-2 NS2B₁₈NS3. (A) *Ab initio* DAMMIF envelope of DENV-2 NS2B₁₈NS3 (grey) superimposed onto the cartoon representation of the crystallographic structure of DENV-4 NS2B₁₈NS3 (PDB ID: 2VBC), with NS2B in green, protease in violet and the helicase in orange. (B) Merged scattering pattern (○; black) and calculated scattering profiles (—; red: crystal structure PDB ID 2VBC; green: 2WZQ; blue: 2WHX) of DENV-2 NS2B₁₈NS3.

3.1.4 Flexibility characterization of DENV-2 NS2B₁₈NS3

Considering that two conformations of DENV-4 NS2B₁₈NS3 existed in crystallographic structures, further analysis of SAXS data were performed to evaluate the probable mobility in the protein. The normalized Kratky-plot provides the information about the conformational behavior of the particle being examined. For a well-folded globular protein such as lysozyme, it will have a distinct bell curve profile. The bell curve profile of normalized Kratky plot of DENV2 NS2B₁₈NS3 (Fig. 3.1.5A) was shifted from a well-defined bell curve profile as seen in a globular lysozyme protein, showing that DENV-2 NS2B₁₈NS3 is extended and less rigid in solution. The flexibility quantification was done with the program EOM 2.0 to assess the potential flexibility of the protease with respect to the helicase (Fig. 3.1.5B). As a result, a discrepancy $\chi^2 = 0.225$ was obtained with EOM 2.0 by selecting an ensemble of five models with diverse orientations of the protease with respect to the helicase. The selected ensembles had a unimodal R_g distribution as shown in figure 3.1.5B. On flexibility quantification, (*ensemble* $R_{flex} = \sim 81\%$ vs *pool* $R_{flex} = \sim 87\%$), the motion of the protease with respect to the helicase was determined to be random. The quality of the data was further confirmed by the control value $R_\sigma = 0.89$ (expected to be lower than 1.0 when *ensemble* $R_{flex} < \text{pool } R_{flex}$).

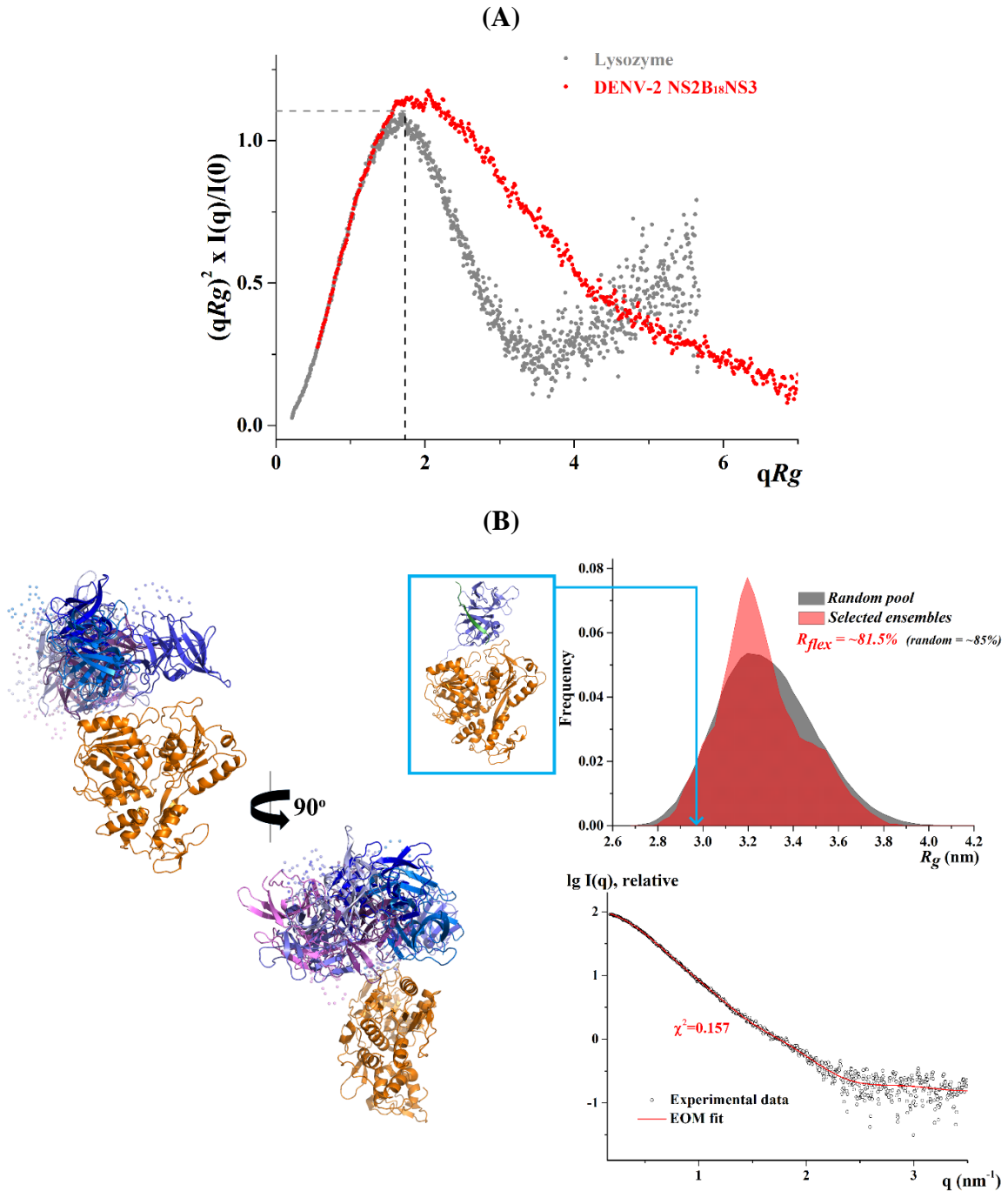


Figure 3.1.5: Flexibility characterization of DENV-2 NS2B₁₈NS3. (A) Normalized Kratky plot for DENV-2 NS2B₁₈NS3 (•; red) compared to that of a globular lysozyme protein (•; grey). The bell curve profile of DENV-2 NS2B₁₈NS3 was shifted from a well-defined bell curve profile as seen for lysozyme. This indicates that DENV-2 NS2B₁₈NS3 is extended and flexible in solution. (B) Flexibility characterization of DENV-2 NS2B₁₈NS3 using EOM 2.0. (Right top) Comparison between the R_g distributions (random pool, grey; selected ensemble, red) shows a random motion as confirmed by R_{flex} ($\sim 81.5\%$ versus $\sim 85\%$, with 85% as the threshold value for randomness) and $R\sigma = 0.89$. The R_g which corresponds to the DENV-4 NS2B₁₈NS3 crystal structure (PDB ID: 2VBC) is labelled with a light blue arrow. (Right bottom) SAXS pattern (black circles) and calculated ensemble scattering profiles (red line) with a χ^2 of 0.157. (Left) Cartoon representation of ensemble of seven possible orientations for the protease domain (in shades of purple) relative to the helicase domain (orange).

3.1.5 Cloning, production and purification of DENV-2 NS2B₁₈NS3 protease and helicase domains

The gene encoding DENV-2 protease (NS2B₁₈NS3₁₋₁₇₉) was obtained by site-directed mutagenesis and appeared as a one clean band at about 6000 bp (Fig. 3.1.6A). The helicase (NS3₁₆₉₋₆₁₈) encoding gene fragment (1350 bp) was amplified and ran as a single band closer to the 1500 bp of the marker (Fig. 3.1.6B).

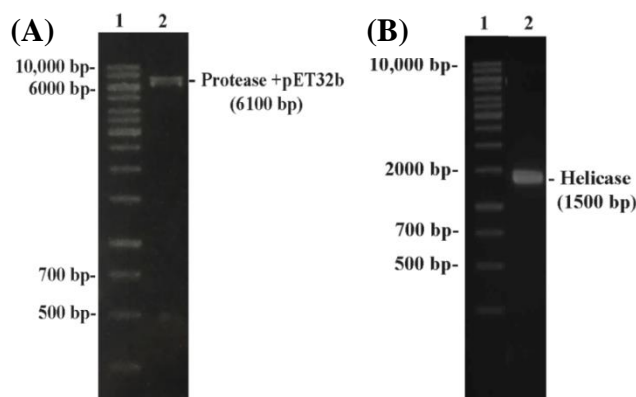


Figure 3.1.6: Agarose gel showing amplified PCR-products. (A) Gene product of DENV-2 protease with linker after site-directed mutagenesis (lane 2). (B) PCR amplified gene product of DENV-2 helicase with linker (lane 2). Lane 1 is a DNA ladder in both cases.

The plasmid for both constructs with the confirmed gene sequence was then transformed into BL21-CodonPlus (DE3)-RIL. Over-production of recombinant protein was detected for both constructs in the induced lane of a SDS-PAGE gel (Fig. 3.1.7A-B, first gel). Both the proteins eluted in the 35 - 100 mM imidazole gradient of the first affinity chromatography (Fig. 3.1.7A-B, second gel). These fractions were pooled and dialysed against buffer B overnight for Trx-His₆ cleavage with thrombin. Following a second affinity chromatography, the tag-free protein eluted in the flow-through. This was confirmed by an SDS-PAGE gel (Fig. 3.1.7A-B, third gel). The cleaved tag could bind to the matrix and was eluted from the column by 500 mM of imidazole (Fig. 3.1.7A-B, third gel). The flow-through from step 2 was then concentrated and applied to SEC, where both proteins eluted from the column as homogenous population as shown in (Fig. 3.1.7C-D).

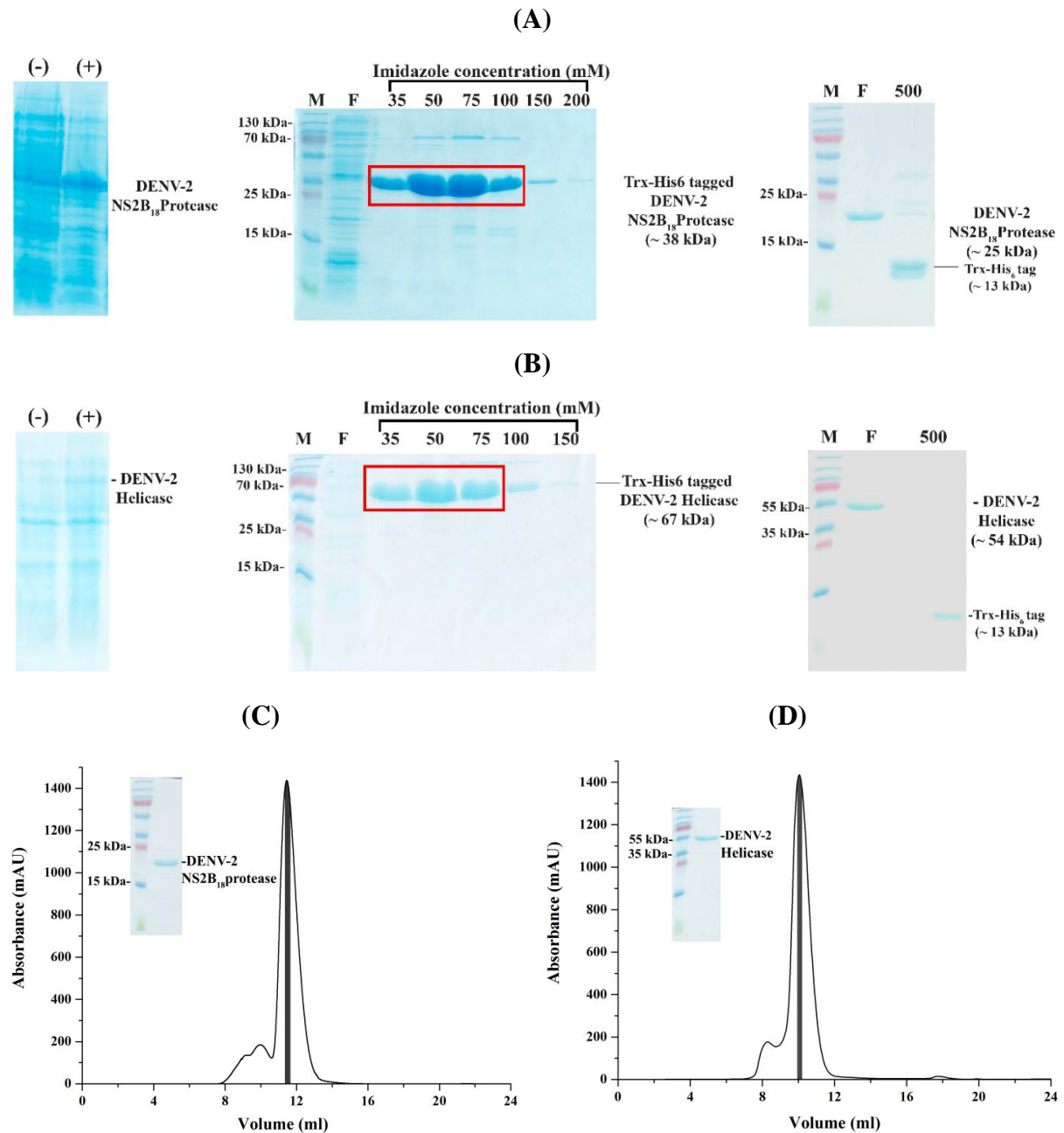


Figure 3.1.7: Purification of DENV-2 NS3 protease and helicase domains. The proteins were eluted by batch method from Ni-NTA column by an imidazole gradient. After induction of protease (A) and helicase (B) was seen in the cell induced with IPTG (+), the recombinant protein was purified by affinity chromatography. The second SDS-PAGE gel in (A) and (B) depict the imidazole concentrations at which the tagged proteins eluted. (M- Protein marker; F- flow-through). The red box indicates the fractions pooled for overnight cleavage of tag. After second affinity chromatography, both protease and helicase eluted in the flow-through (F). This was confirmed by the third gel in (A; protease) and (B; helicase). On application to SEC, a single peak was observed in the chromatograph for both (C) protease and (D) helicase.

3.1.6 SAXS of studies of DENV-2 protease and helicase domain

To elucidate the contribution of the protease and/or the helicase towards the flexible behavior of DENV-2 NS3, SAXS data of DENV-2 protease and helicase was collected with freshly purified protein. SAXS-patterns of DENV-2 protease (1.6, 3.0 and 4.7 mg/ml) and

helicase (1.3, 2.5, 3.0 and 5.0 mg/ml) were recorded (Fig. 3.1.8A-B). Since no concentration dependent increase in R_g was observed, the 3.0 mg/ml dataset of the protease, and the 2.5 mg/ml dataset of the helicase was used for further analysis (Fig. 3.1.8A-B). The Guinier plots at low q from all three protein concentrations showed linearity at low angles. There was no indication of aggregation of protein samples. The R_g values derived from Guinier approximation and the $P(r)$ function for protease are 1.92 ± 0.02 and 1.95 ± 0.03 nm, respectively, and for helicase are 2.59 ± 0.03 and 2.55 ± 0.02 , respectively (Fig. 3.1.9) (table 3.1.2). The D_{max} was estimated to be 6.3 ± 1 nm and 8.4 ± 1 nm, respectively, for protease and helicase. The $P(r)$ function curve for both curves were skewed to the right with a small tail. The appearance of this tail is attributed to the linker for both the constructs.

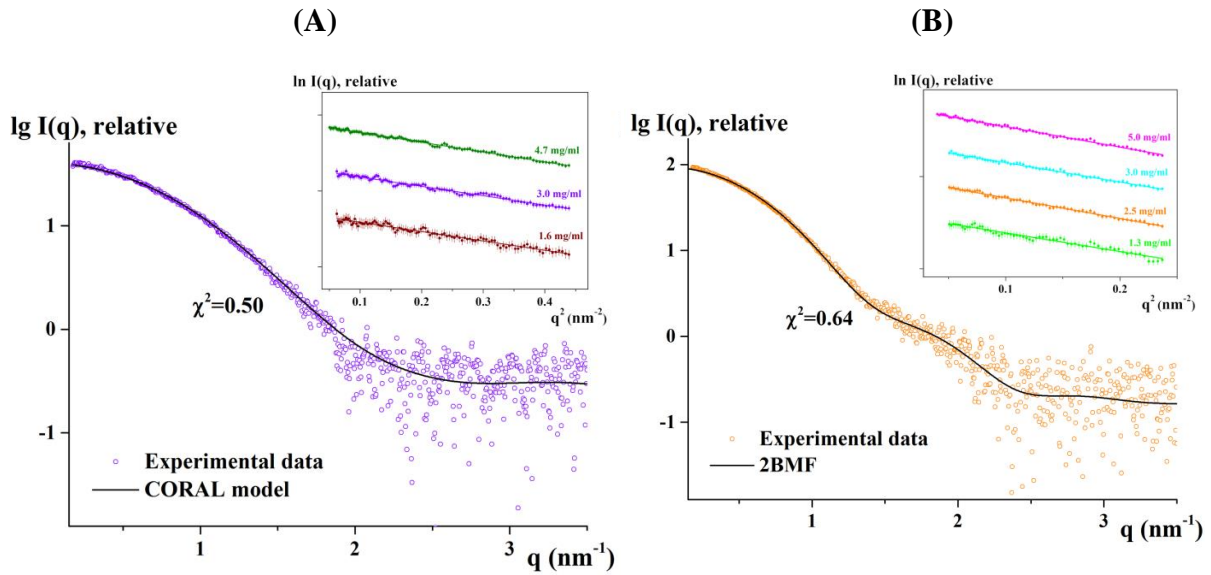


Figure 3.1.8: SAXS Data analysis for DENV-2 protease and helicase domains. (A) SAXS-scattering pattern of DENV-2 NS2B₁₈protease (○; purple). The 3.0 mg/ml dataset was used for data analysis. Fit of the theoretical scattering curve of the CORAL model (—; black) of protease to the experimental data. *Insert:* The Guinier region of the data of protease collected at three different concentrations (1.6, 3.0 and 4.7 mg/ml) (B) SAXS-scattering pattern of helicase (○; orange). The 3.0 mg/ml dataset was used for data analysis. *Insert:* The Guinier region the data of DENV-2 helicase collected at four different concentrations (1.3, 2.5 3.0 and 5.0 mg/ml). Fit of the theoretical scattering curve of the crystal structure of helicase (PDB ID: 2BMF) (—; black) to the experimental data.

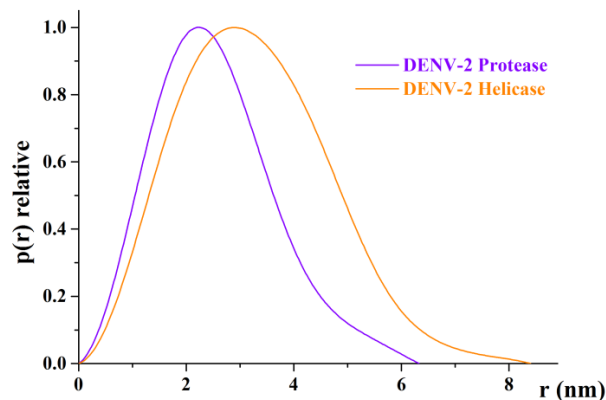


Figure 3.1.9: Distance distribution function of DENV-2 NS2B₁₈NS3 protease and helicase. Distance distribution function $P(r)$ of protease (magenta) and helicase (orange). Both the curves are skewed to the right.

Protein sample	DENV-2 Protease	DENV-2 Helicase
Concentration range (mg ml ⁻¹)	1.60 – 4.70	1.30 – 5.00
Temperature (K)	288.15	288.15
Structural parameters[†]		
$I(0)$ (cm ⁻¹) [from $P(r)$]	41.49 ± 0.3	97.08 ± 0.56
R_g (nm) [from $P(r)$]	1.95 ± 0.02	2.55 ± 0.02
$I(0)$ (cm ⁻¹) (from Guinier)	41.28 ± 0.3	98.13 ± 0.78
R_g (nm) (from Guinier)	1.92 ± 0.02	2.59 ± 0.03
D_{max} (nm)	6.3 ± 1	8.4 ± 1
Porod volume estimate (nm ³)	~ 31.5	~ 73
DAMMIF excluded volume (nm ³)	~ 45.4	~ 94
Dry volume from sequence (nm ³) [‡]	~ 26.4	~ 62.3
Molecular mass determination		
Calculated monomeric MM (kDa)	~ 21.8	~ 51.4
Molecular mass MM (kDa) [from <i>Porod invariant</i>]	18.9 ± 2	43.8 ± 4.4
Molecular mass MM (kDa) [from <i>excluded volume</i>]	22.6 ± 2	46.7 ± 4.7
Molecular mass MM (kDa) [from V_c]	24.3 ± 2	50.1 ± 5.0

Table 3.1.2: Overall structural parameters of DENV-2 NS3 protease and helicase domains

3.1.7 *Ab initio* model constructions for DENV-2 NS2B₁₈NS3 protease and helicase

The low-resolution shapes of protease and helicase were reconstructed using DAMMIF, and the final averaged models have been reported (Fig. 3.1.10). The crystal structure of the protease (PDB ID: 2FOM; [93]) does not include the N-terminal 18 amino acid residues from NS2B or the linker residues as at the C-terminus. These residues for the protease were thus built using CORAL and this model was subsequently used for the superimposition with *ab initio* envelope (Fig. 3.1.10A). The helicase *ab initio* envelope fitted nicely with the DENV-4 helicase crystal structure (PDB ID: 2BMF; [85]) (Fig. 3.1.10B).

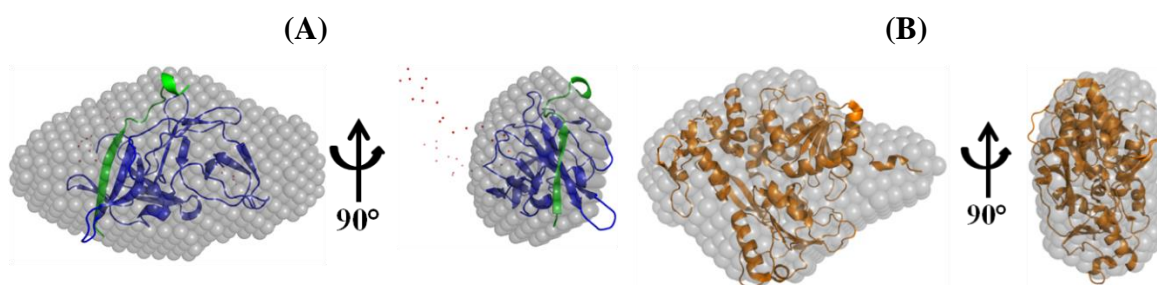


Figure 3.1.10: Solution envelopes of DENV-2 NS3 protease and helicase domains. (A) Overlay of DAMMIF envelope onto the cartoon representation of the CORAL model of protease. The protease is colored in blue with the NS2B in green. The red dots represent the missing residues that were modeled. (B) Overlay of DAMMIF envelope onto the cartoon representation of the crystallographic structure of DENV-4 helicase (PDB ID: 2BMF), with helicase colored in orange.

3.1.8 Assessment of DENV NS3 domains flexibility

Superimposition of the normalized Kratky-plots of DENV-2 NS2B₁₈NS3 and the two domains of DENV-2 NS3 (Fig. 3.1.11), revealed only a slight change of the bell curve profile for both helicase and protease domain when compared with a well-defined bell curve profile of lysozyme. Thus, the individual domains do not contribute much towards the flexibility of DENV-2 NS3.

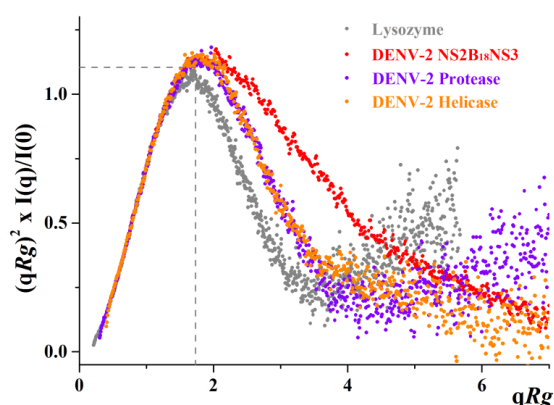


Figure 3.1.11: Comparison of normalized Kratky-plot DENV-2 constructs. Normalized Kratky-plot of DENV-2 NS2B₁₈NS3 (●; red) compared to the individual domains, protease (●; magenta) and helicase (●; orange), and the compact globular lysozyme (●; grey) with a peak (---; grey), represents the theoretical peak. Compared to the extended and flexible full-length protein, both domains were compacted and the flexibility of the NS3 protein is attributed to the linker.

3.1.9 Cloning of DENV-2 NS2B₁₈NS3_{174PPAVP179} and purification of the recombinant protein

The PCR amplified product for DENV-2 NS2BNS3_{174PPAVP179} mutant is shown in figure 3.1.12A. The top more prominent band (~8000 bp) was excised from the agarose gel and DNA extracted. This fragment was used in the consequent cloning steps. Once the plasmid with the correct gene sequence was transformed for over-expression in BL21-CodonPlus (DE3)-RIL, protein production for the recombinant protein was checked (Fig. 3.1.12B) The mutant eluted in the imidazole gradient of 35-100 mM from the first Ni-NTA column (Fig. 3.1.13; first gel) and in the flow-through in the second Ni-NTA step (Fig. 3.1.13; second gel). The protein which eluted from the SEC was homogeneous and pure (Fig. 3.1.13; inset).

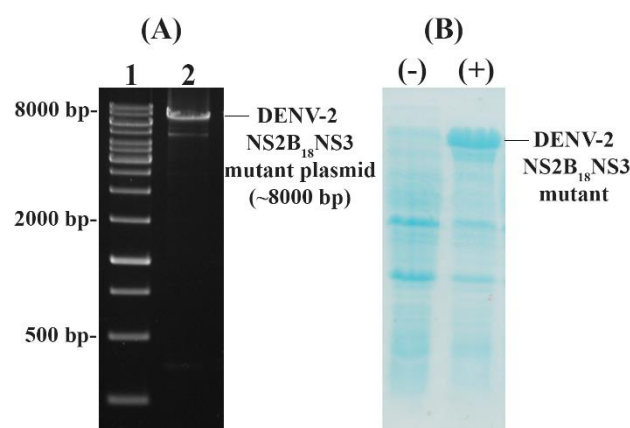


Figure 3.1.12: Amplification of gene encoding DENV-2 NS3 mutant and recombinant DENV-2 NS2B18NS3 production. (A) Amplified product (lane 2) and DNA ladder in lane 1. DNA sample of 5 μ l was loaded on an agarose gel (1.5%). (B) SDS-PAGE (17%) showing the expression of recombinant protein. Protein production was induced with 1 mM IPTG. (+) and (-) indicates protein obtained from cells induced with IPTG and uninduced cells, respectively.

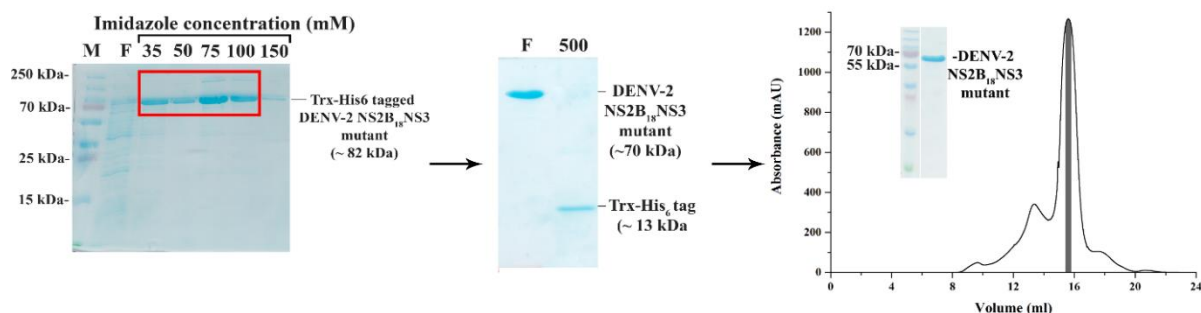


Figure 3.1.13: Purification of DENV-2 NS3 mutant. The proteins were eluted by batch method from Ni-NTA column by an imidazole gradient. Fractions indicated by the red box are pooled for the subsequent tag cleavage. On cleavage of the thioredoxin tag the protein elutes in the flow-through (F) while the tag elutes with 500 mM imidazole (500) from the second Ni-NTA column. A single peak is observed in SEC chromatogram. When fractions from the grey region (top 15% of peak) were concentrated and run on a SDS-PAGE gel (inset), it ran as a single band around 70 kDa.

3.1.10 Structural insights from SAXS experiment of DENV-2 NS2B₁₈NS3 mutant

SAXS data were collected at concentrations of 1.1 mg/ml, 3.0 mg/ml and 5.0 mg/ml (Fig. 3.1.14A). There was no indication of aggregation of protein sample (linearity of the

Guinier plot; Fig. 3.1.14A). The three datasets were merged for further data analysis. The calculated R_g value from Guinier approximation and that derived from $P(r)$ function were 3.28 ± 0.03 nm and 3.25 ± 0.08 nm, respectively. The D_{max} was estimated to be 10.8 ± 1 nm (Fig. 3.1.14B). The overall structural parameters of the mutant were similar to the WT, indicating that replacing the DENV-2 NS3 linker residues with residues of the HCV NS3 linker did not compact the extended DENV-2 NS3. The dimensionless Kratky plot further indicates that the flexibility of the mutant recombinant protein is also not altered (Fig. 3.1.14C).

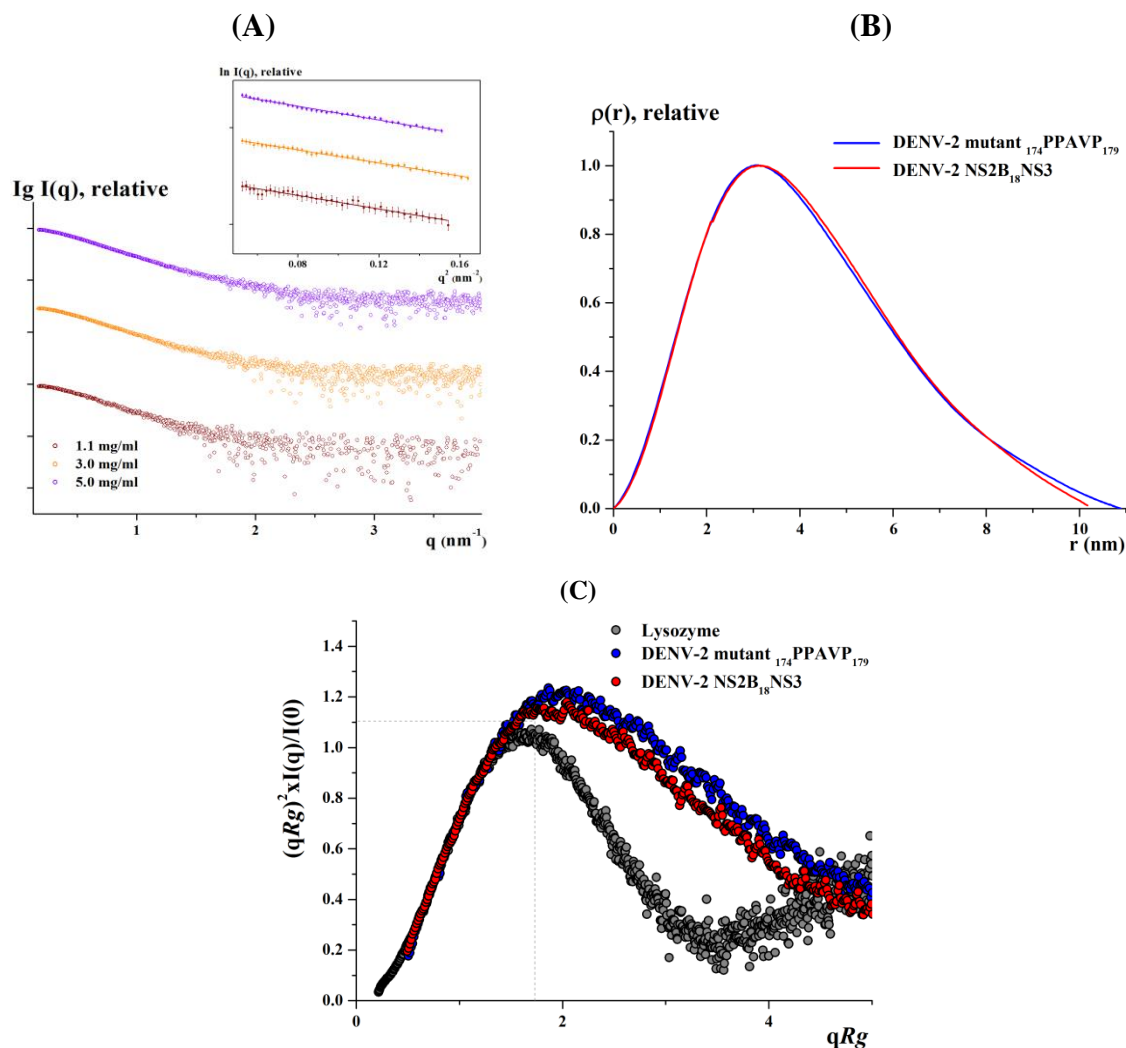


Figure 3.1.14: SAXS studies of NS2B $_{18}$ NS3 mutant. (A) Scattering pattern (\circ) of DENV-2 NS2B $_{18}$ NS3 mutant at 1.1 mg/ml (brown), 3.0 mg/ml (orange) and 5.0 mg/ml (magenta). *Inset:* Linearity of the Guinier plot at low angles indicates no aggregation of protein sample. (B) Pair-wise distance distribution $P(r)$ curve of WT (—; red) and mutant (—; blue). The merged curve was used to generate the curve for both proteins. The curve was bell-shaped but slightly skewed to the right with a tail indicating an elongated molecule in solution. (C) Comparison of the dimensionless Kratky plot of WT (\bullet ; red) and mutant protein (\bullet ; blue) indicate that the flexibility of both proteins is similar.

3.1.11 ATPase activity for DENV-2 NS2B $_{18}$ NS3 and helicase

Since the DENV2 NS2B $_{18}$ NS3 has a helicase at its C-terminus, continuous ATP hydrolysis was studied to characterize the DENV-2 helicase ATPase activity. In these studies,

the entire DENV-2 NS2B₁₈NS3 as well as the helicase domain was investigated, to get insight into the ATP hydrolysis of the enzyme. The ATPase activity of DENV2 NS2B₁₈NS3 was determined to be 474.5 U/mM (6.59 U/mg). In comparison, the DENV-2 helicase with linker (NS3₁₆₈₋₆₁₈) showed a higher ATPase activity (564.9 U/mM, 10.99 U/mg) compared to the entire NS3 complex (Fig. 3.1.15A). The assay confirmed that both, the recombinant full length NS3 as well as the helicase with linker were enzymatically active.

Three ATPase inhibitors, resveratrol, quercetin and NBD-Cl were used to evaluate the inhibitory effect to the ATP hydrolysis activity of DENV-2 NS2B₁₈NS3 and its helicase with linker. A fixed inhibitor concentration (150 μ M), was used for both proteins. Resveratrol dropped ATP hydrolytic activity of NS2B₁₈NS3 and helicase by 40% and 68%, respectively. In comparison, quercetin reduced the ATPase activity of NS2B₁₈NS3 to 47% and the helicase one to only 81%. This interesting phenomenon confirms a recent computational study, demonstrating that quercetin binds to NS3 protease. In contrast to resveratrol and quercetin, NBD-Cl boosted the ATP hydrolysis activity in both NS2B₁₈NS3 and helicase (Fig. 3.1.15B).

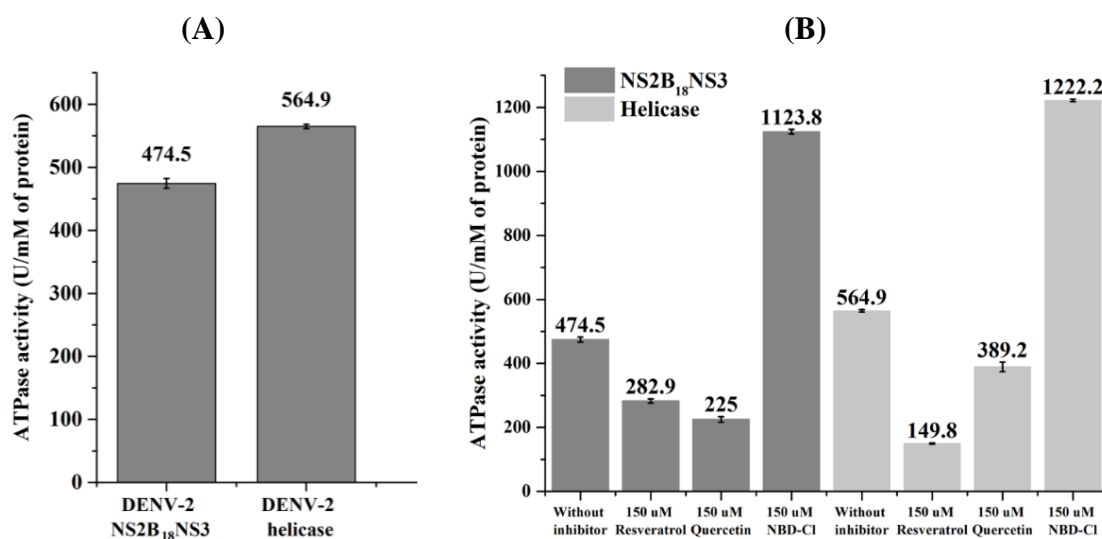


Figure 3.1.15: Bar diagram representation of ATP hydrolysis activity for DENV-2 NS2B₁₈NS3 and helicase. (A) DENV-2 NS2B₁₈NS3 and the helicase using an NADH-dependent continuous ATP hydrolysis assay. (B) Use of inhibitors to study the effect on hydrolytic activity of DENV-2 NS2B₁₈NS3 and the NS3-helicase.

3.2 Solution studies of ZIKV NS3

After the SAXS studies for ZIKV NS3 were conducted, the crystal structure of the individual domain of ZIKV NS3 was solved (protease PDB ID: 5LC0; [190] and helicase PDB ID: 5GJB; [191]). Despite the protease and helicase structure being available, the full-length NS3 structure has not been solved. To gain more insight into how the ZIKV NS3 domains are

arranged in solution and how the protein domains undergo structural alterations, SAXS experiments were performed.

3.2.1 Cloning, expression and purification of recombinant ZIKV NS2B₄₈NS3

PCR amplification of the gene encoding ZIKV NS2B₄₈NS3 showed two bands on an 1.5% agarose gel (Fig. 3.2.1A). The upper band corresponds to the correct size of the gene (~2022 bp), this band was then excised from the gel, purified, digested with restriction enzymes and ligated with modified pET32b vector. The plasmid containing the correct insert sequence was subsequently transformed into *E. coli* BL21-CodonPlus (DE3)-RIL cells for production of protein. Protein production on induction of gene expression by IPTG was confirmed by a 17% SDS-PAGE gel (Fig. 3.2.1B).

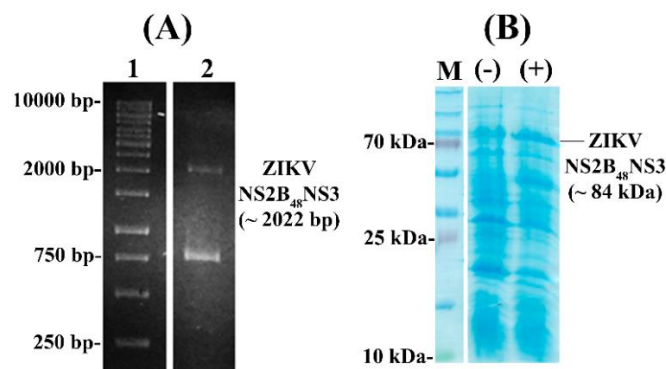


Figure 3.2.1: Amplification of gene encoding ZIKV NS2B₄₈NS3 and recombinant ZIKV NS2B₄₈NS3 production. (A) Amplified product (lane 2) and DNA ladder in lane 1. DNA sample of 5 μ l was loaded on an agarose gel (1.5%). (B) SDS-PAGE (17%) showing the expression of recombinant protein. Protein production was induced with 1 mM IPTG. (+) and (-) indicates protein samples from cells induced with IPTG or uninduced cells, respectively.

Like its DENV-2 counterpart, the first step to obtain a pure recombinant ZIKV NS2B₄₈NS3 was affinity chromatography. The Trx-His₆ tagged protein eluted from the Ni-NTA column at an imidazole concentration of 35 - 100 mM. Several bands (apart from the correct size of the protein ~84 kDa) were observed when the samples were run on a 17% SDS-PAGE gel, indicating degradation (Fig. 3.2.2). Following the Trx-His₆ tag by thrombin, the protein was subjected to another round of affinity chromatography, where the protein eluted in flow-through and the cleaved tag in the fraction eluted with 500 mM of imidazole. However, the protein was co-purified with the degraded product(s) (Fig. 3.2.2). The protein was then concentrated and applied to a SEC column where three prominent peaks were observed (Fig. 3.2.2). Samples from peak II and III were applied to a SDS-PAGE gel (Fig. 3.2.2, inset). Peak III, which had the maximum absorbance, was identified as the degraded product. ZIKV NS2B₄₈NS3 was identified at peak II which still showed degradation on the SDS-PAGE gel

(Fig. 3.2.2, inset). Since, a homogenous population of purified protein was not observed on the SDS-PAGE gel, the recombinant protein was rendered unsuitable for SAXS studies.

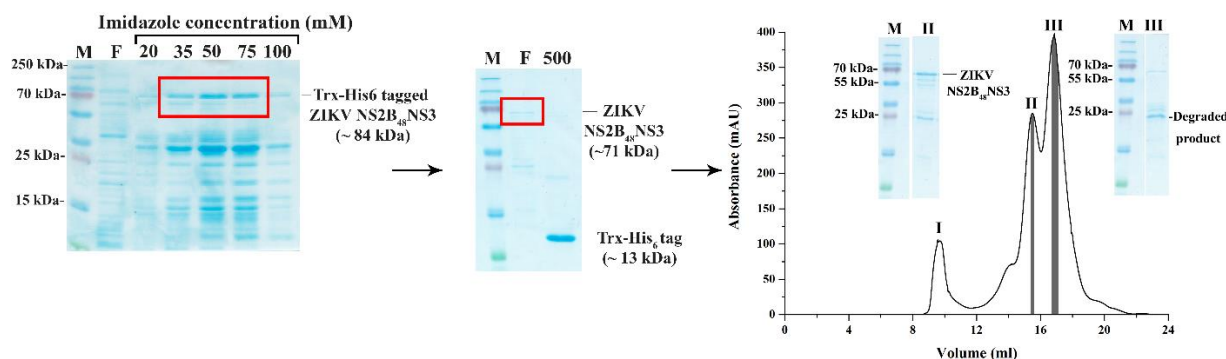


Figure 3.2.2: Purification of ZIKV NS2B₄₈NS3. The first SDS-PAGE gel shows the imidazole concentrations at which the tagged protein eluted from Ni-NTA column (M= Protein marker; F= flow-through). Several bands were observed indicating probable degradation. The red box indicates the fractions pooled for overnight cleavage of tag. After the second affinity chromatography, the protein eluted in the flow-through (F) and appeared degraded as well. The SEC chromatogram reveals three peaks. Shaded grey region corresponds to 15% of maximum UV absorbance, from where fractions were pooled and concentrated. When the concentrated fractions from peak II and III were run on a gel (inset of the chromatograph), II was identified as the ZIKV NS2B₄₈NS3 and III as a degraded product.

3.2.2 Generation of the ZIKV NS2B₄₈NS3_{S135A} mutant

Since the ZIKV NS2B₄₈NS3 showed auto-proteolysis, the next approach to obtain a pure and homogenous ZIKV NS3 was to mutate the catalytic serine (S135) in the protease to prevent the probable self-cleavage by the protease. The plasmid containing the ZIKV NS2B₄₈NS3 sequence was used as a DNA template for PCR. The amplified product appeared as a single band at about 8000 bp (vector (5899 bp) + insert (2022 bp)) in an agarose gel (Fig. 3.2.3A). After confirming the mutated sequence, the plasmid was transformed into BL21-CodonPlus (DE3)-RIL cells for production of recombinant protein, which was confirmed by an SDS-PAGE gel (Fig. 3.2.3B).

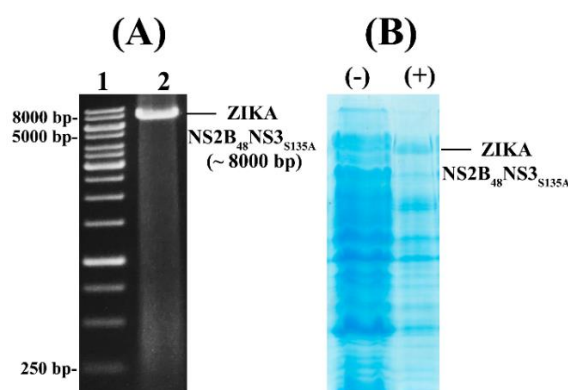


Figure 3.2.3: Amplification of gene encoding ZIKV NS2B₄₈NS3 mutant and recombinant protein production. (A) Amplified product (lane 2) and DNA ladder in lane 1. DNA sample of 5 μ l was loaded on an agarose gel (1.5%). (B) SDS-PAGE (17%) showing the expression of recombinant protein. Protein production was induced with 1 mM IPTG. (+) and (-) indicates cells induced with IPTG or uninduced cells, respectively.

The protein was purified in three steps as described for the WT enzyme. The SDS-PAGE gels after first and second affinity chromatography, showed a similar profile of ‘laddering’ (Fig. 3.2.4). When the protein was finally applied the SEC, an elution profile like the WT enzyme was observed (Fig. 3.2.4). However, peak III showed had slightly lesser UV absorbance than peak II, indicating that the mutant protein revealed lower degradation than ZIKV NS2B₄₈NS3. However, when peak II was run on SDS-PAGE gel, a lower band (25 kDa) other than the band for ZIKV NS2B₄₈NS3_{S135A} (71 kDa) was still observed (Fig. 3.2.4, inset). Although the mutation of the catalytic serine reduced the auto-proteolysis, it did not completely prevent it. Hence a new construct, ZIKV NS2B₁₈NS3, was generated.

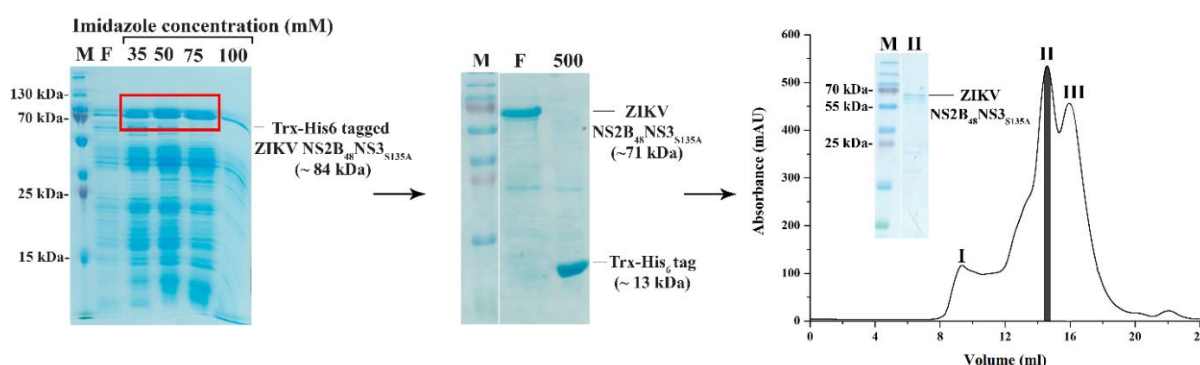


Figure 3.2.4: Purification of ZIKV NS2B₄₈NS3_{S135A}. The first SDS-PAGE gel shows the imidazole concentrations at which the tagged protein eluted from Ni-NTA column (M= Protein marker; F= flow-through). The red box indicates the fractions pooled for overnight cleavage of tag. After the second affinity chromatography, the protein eluted in the flow-through (F). The SEC chromatogram reveals three peaks. When fractions from peak II were run on a gel (inset of chromatograph), II was identified as ZIKV NS2B₄₈NS3. Peak III represents degradation products (not shown).

3.2.3 Cloning, expression and purification of recombinant ZIKV NS2B₁₈NS3

A new ZIKV NS3 construct with a shorter NS2B (18 amino acids) sequence was created by overlap extension PCR as described for DENV-2 NS3. The agarose gel confirmed the presence of the amplified gene (Fig. 3.2.5A), which was digested and ligated into the modified pET 32b plasmid. The verified sequenced plasmid was transformed into BL21-

CodonPlus(DE3)-RIL cells. A 17% denaturing gel confirmed the production of ZIKV NS2B₁₈NS3 by the expression strain (Fig. 3.2.5B).

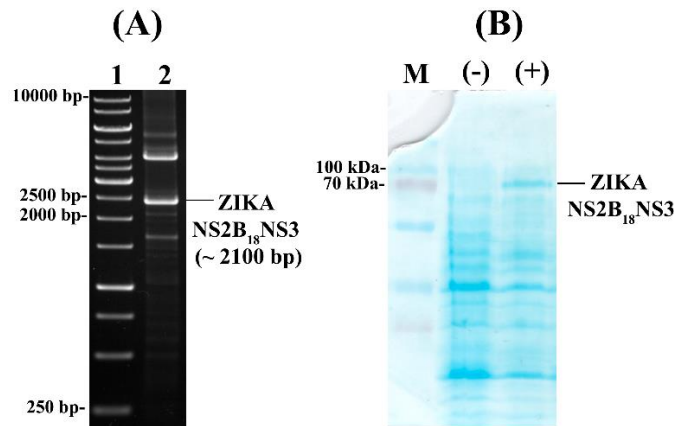


Figure 3.2.5: Amplification of gene encoding ZIKV NS2B₁₈NS3 mutant and recombinant protein production. (A) Amplified product (lane 2) and DNA ladder in lane 1. Unspecific binding of primers leading to several bands was observed in the gel. The band corresponding to the correct gene size was excised and used for subsequent steps. DNA sample of 5 μ l was loaded on an agarose gel (1.5%). (B) SDS-PAGE (17%) showing the expression of recombinant protein. Protein production was induced with 1 mM IPTG. (+) and (-) indicates protein obtained from cells induced with IPTG and uninduced cells, respectively

After first purification step, the protein eluted from the Ni-NTA column at an imidazole concentration of 35 - 100 mM as shown in figure 3.2.6. More than one band were still observed on the SDS-PAGE gel, however, the band corresponding to ZIKV NS2B₁₈NS3 was more prominent than the degraded products. After tag cleavage and a second affinity chromatography step, the recombinant protein was applied to SEC. The elution chromatogram recorded a much higher UV absorbance for peak II (Fig. 3.2.6). The UV absorbance of peak III was considerably lower, indicating noticeably lesser degradation of ZIKV NS2B₁₈NS3. This was confirmed by an SDS-PAGE gel (Fig. 3.2.6, inset). Since a more homogenous protein was obtained for this construct, the ZIKV NS2B₁₈NS3 was utilized in SAXS studies. The elution fractions corresponding to the top 15% of peak II absorbance (grey shaded area in (Fig. 3.2.6)) were pooled utilized for SAXS measurements.

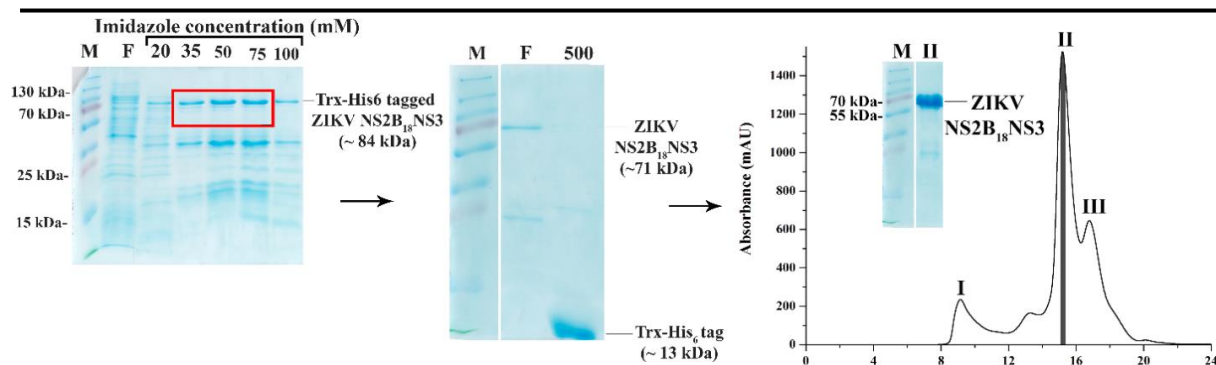


Figure 3.2.6: Purification of ZIKV NS2B₁₈NS3. The first SDS-PAGE gel shows the imidazole concentrations at which the tagged protein eluted from Ni-NTA column (M= Protein marker; F= flow-through). Probable degradation was observed in the gel. The red box indicates the fractions pooled for overnight cleavage of tag. After second affinity chromatography, the protein eluted in the flow-through (F). Three peaks were observed in the SEC chromatogram. The fractions from the region marked in grey was pooled, concentrated and run on gel. A homogenous population of protein was observed.

3.2.4 Primary SAXS data analysis of ZIKV NS2B₁₈NS3

The SAXS data for ZIKV NS2B₁₈NS3 was recorded with freshly purified protein at four protein concentrations: 1.2, 2.4, 3.0 and 4.5 mg/ml (Fig. 3.2.7A). Linearity of Guinier region at all four protein concentrations confirmed that the protein was not aggregated (Fig. 3.2.7A, inset). The R_g values derived from the Guinier approximation were 3.25 ± 0.14 (1.2 mg/ml), 3.25 ± 0.04 (2.4 mg/ml), 3.14 ± 0.04 (3.0 mg/ml) and 3.31 ± 0.03 nm (4.5 mg/ml) (table 3.2.1). No concentration dependent increase in R_g values was observed and data collected for any one concentration could be used for further analysis. In this case, the SAXS data collected at 2.4 mg/ml concentration was used. The R_g value obtained from the $P(r)$ function was 3.30 ± 0.02 nm, which confirmed the value determined from Guinier approximation. The maximum particle dimension (D_{max}) was estimated to be 10.4 ± 1 nm (Fig. 3.2.7B) (table 3.2.1). The estimated molecular weight from the Porod volume was 63.1 ± 6.3 kDa, revealing that the protein is monomer in solution.

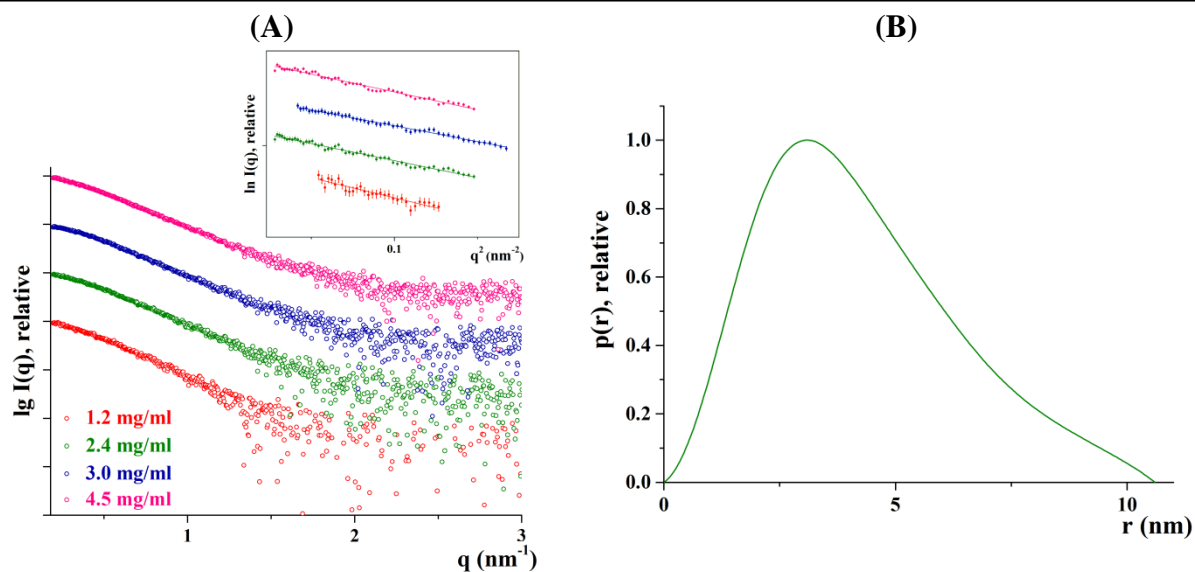


Figure 3.2.7: SAXS studies of ZIKV NS2B₁₈NS3. (A) Scattering pattern (\circ) of ZIKV NS2B₁₈NS3 at 1.2 mg/ml (red), 2.4 mg/ml (green), 3.0 mg/ml (blue) and 4.5 mg/ml (magenta). *Inset:* Linearity of the Guinier plot at low q indicates no aggregation of protein sample. (B) Pair-wise distance distribution $P(r)$ curve was of ZIKV NS2B₁₈NS3. The bell-shaped curve had a tail indicating an elongated molecule in solution.

Protein sample	ZIKV NS2B ₁₈ NS3
Concentration range (mg ml ⁻¹)	1.20 – 4.50
Structural parameters[†]	
$I(0)$ (cm ⁻¹) [from $P(r)$]	105.5 ± 1.04
R_g (nm) [from $P(r)$]	3.26 ± 0.04
$I(0)$ (cm ⁻¹) (from Guinier)	105.2 ± 0.68
R_g (nm) (from Guinier)	3.30 ± 0.02
D_{max} (nm)	10.4 ± 1
Porod volume estimate (nm ³)	~ 101
DAMMIF excluded volume (nm ³)	~ 129
Dry volume from sequence (nm ³) [‡]	~ 86
Molecular mass determination	
Calculated monomeric MM (kDa)	~ 71.2
Molecular mass MM (kDa) [from <i>Porod invariant</i>]	63.1 ± 6.3
Molecular mass MM (kDa) [from <i>excluded volume</i>]	64.5 ± 6.5

Table 3.2.1: Overall structural parameters for ZIKV NS2B₁₈NS3

3.2.5 Rigid body modelling of ZIKV NS2B₁₈NS3

Till date no high-resolution structure for ZIKV NS3 is available. Thus, BUNCH (Petoukhov et al., 2005) which models multidomain proteins based on SAXS data, was used to perform rigid body modelling. The high-resolution structures of ZIKV protease (PDB ID: 5LC0) and ZIKV helicase (PDB ID 5GJB) were used as rigid bodies for BUNCH to determine the optimal positions and orientation of the subunits and linker residues and the flexible linker connecting the two subunits was constructed based on plausible conformations of the linkers that could be attached to the terminal residues of the domains. For reproducibility of the model, the BUNCH algorithm was run ten independent times, and the model with the least χ^2 of 0.68 was reported as a representative model (Fig. 3.2.8).

3.2.6 *Ab initio* model constructions for ZIKV NS2B₁₈NS3

The solution shape of ZIKV NS2B₁₈NS3 was constructed using the program DAMMIF [174], which reconstructs low-resolution shape of proteins *ab initio*. To ascertain the reproducibility of the low-resolution shape, ten independent DAMMIF calculations were performed. The program DAMVER [175] was then used to compare the individual models and obtain an averaged model once the most ‘typical’ model of the ten models was selected. The averaged model of ZIKV NS2B₁₈NS3 with NSD of 0.685 ± 0.04 is reported in (Fig. 3.2.8A). Since the NSD of all ten independent models were less than 1, it indicated that the shape reconstruction was reproducible.

The BUNCH model of ZIKV NS2B₁₈NS3 was then superimposed onto the *ab initio* solution shapes using SUPCOMB (Fig. 3.2.8A). The rigid body model and the solution envelope overlapped nicely with an NSD of 1.35. CRY SOL was used to evaluate the fit of the theoretical scattering curve to the experimental data (Fig. 3.2.8B).

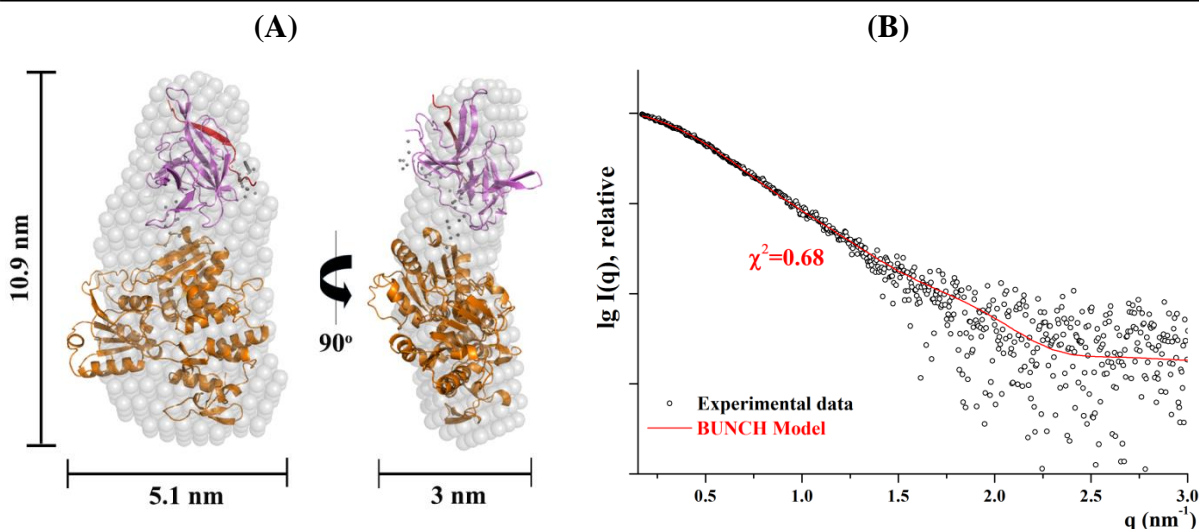
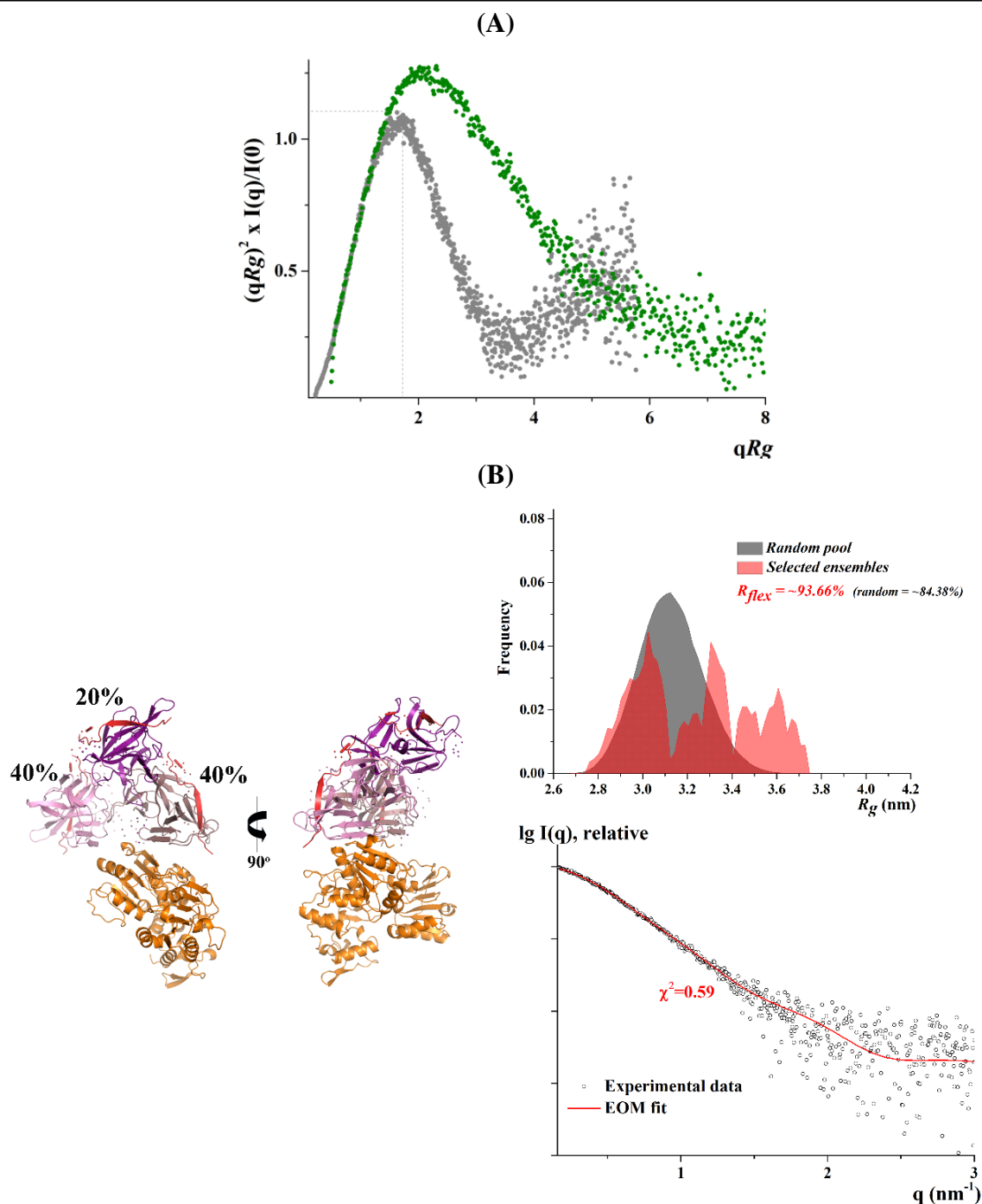


Figure 3.2.8: *Ab initio* envelope reconstruction of ZIKV NS2B₁₈NS3. (A) DAMMIF envelope superposed onto the cartoon of the BUNCH model of ZIKV NS3 with the NS2B cofactor (red), the protease domain (purple) and the helicase domain (orange). Grey dots represent the unresolved region in the high-resolution structure. Front (left) and side (right) views are displayed (B) Scattering curve of ZIKV NS2B₁₈NS3 superimposed with the theoretical scattering pattern derived from BUNCH model.

3.2.7 Flexibility characterization of ZIKV NS2B₁₈NS3

On analyzing the normalized Kratky plot, the ZIKV NS2B₁₈NS3 was found to be flexible and it adopts an extended conformation. When compared to the reference protein lysozyme, the ZIKV NS2B₁₈NS3 showed a broader bell-curved profile, with a shift in the peak maxima (Fig. 3.2.9A).

To further characterize this dynamic nature of the protein, EOM 2.0 [181] was performed. The ensemble R_g distribution range from 2.8 – 3.7 nm, indicating a wide range of motion (Fig. 3.8.9B). An ensemble R_{flex} value of 93.66 % as compared to pool R_{flex} value of 84.38 % indicate a random motion of the protease. Three representative ensemble models were selected by EOM 2.0 had a value of $\chi^2 = 0.59$ (Fig. 3.2.9B).



3.2.8 Generation of ZIKV NS3 protease and helicase domains

The gene encoding the ZIKV protease was amplified from the plasmid containing ZIKV NS2B₄₈NS3 gene sequence. The PCR amplified product after site-directed mutagenesis appeared as a clean band at ~ 6000 bp (Fig. 3.2.10A). The helicase encoding gene fragment (1347 bp) was also cloned from the same template DNA and ran as a (Fig. 3.2.10B). The plasmid for both constructs with the confirmed gene sequence was then transformed into BL21-CodonPlus (DE3)-RIL cells.

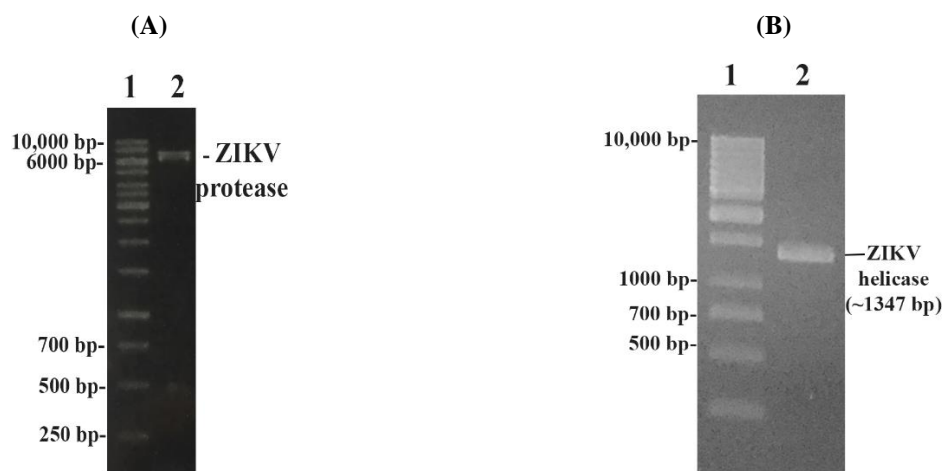


Figure 3.2.10: Agarose gel showing amplified PCR-products. (A) Gene product of ZIKV protease with linker after site-directed mutagenesis (lane 2). (B) PCR amplified gene product of ZIKV helicase with linker (lane 2). Lane 1 is a DNA ladder in both cases.

A 17% SDS-PAGE gel confirmed the protein production for both constructs as shown in figure 3.2.11. The band for ZIKV protease with the Trx-His₆ tag was observed at 38 kDa (Fig. 3.2.11A), while the band for Trx-His₆ tagged ZIKV helicase was observed at about 63 kDa (Fig. 3.2.11B).

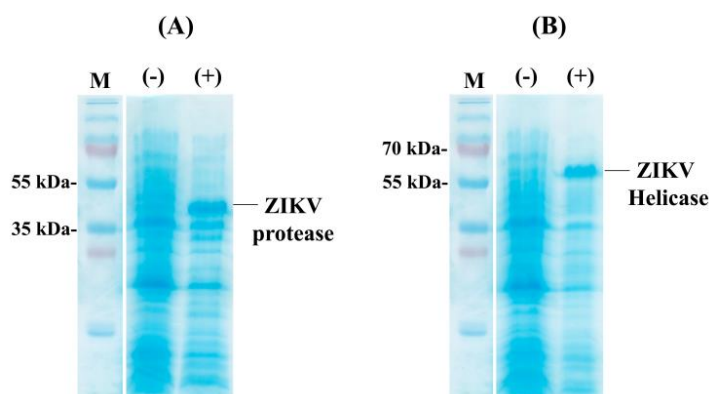


Figure 3.2.11: Recombinant ZIKV protease and helicase. SDS-PAGE gels showing the expression of (A) ZIKV protease and (B) ZIKV helicase. Protein production was induced with 1 mM IPTG. (+) and (-) indicates protein obtained from cells induced with IPTG and uninduced cells, respectively (M= protein ladder).

Multiple bands were observed on the SDS-PAGE gel after the first affinity chromatography column, indicating probable degradation of the ZIKV protease (Fig. 3.2.12). However, after Trx-His₆ tag cleavage and a second affinity chromatography step, no degraded product was observed in the denaturing gel (Fig. 3.2.12; second gel). One band at around 25 kDa could be seen on the denaturing gel with the cleaved Trx-His₆ tag in the fraction eluted with 500 mM imidazole. Three distinct peaks were observed in the elution diagram (Fig. 3.2.12). Peak III with the maximum UV absorbance in the chromatogram was identified as the fraction containing the recombinant protein. Peak I and II were identified as oligomeric populations of ZIKV protease in the SDS-PAGE gel (Fig. 3.2.12). Fraction of the grey shaded area in figure 3.2.12, which correspond to the top 15% of peak II were pooled and samples used for acquisition of SAXS data.

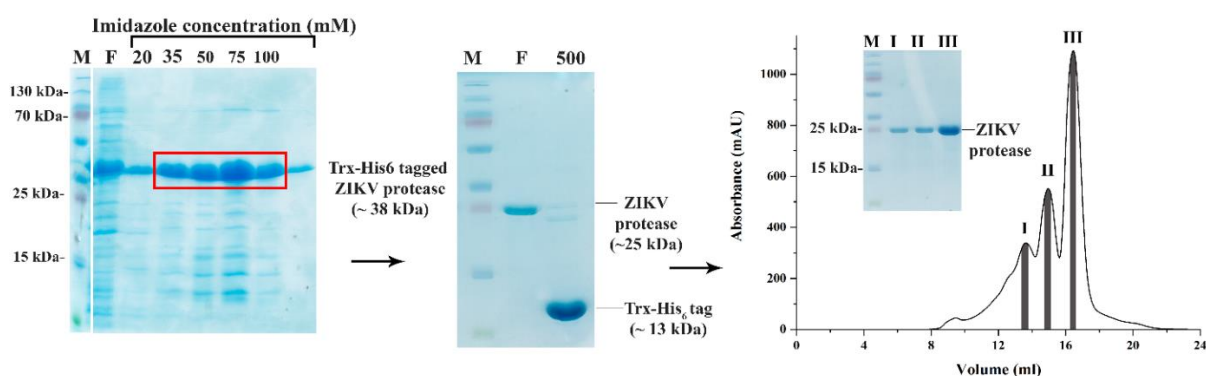


Figure 3.2.12: Purification of ZIKV protease. The first SDS-PAGE gel shows the imidazole concentrations at which the tagged protein eluted from Ni-NTA column (M= Protein marker; F= flow-through). The red box indicates the fractions pooled for overnight cleavage of tag. After the second affinity chromatography, the protein eluted in the flow-through (F) Three peaks were noticed in the SEC chromatogram. When fractions from the areas shaded in grey in the chromatogram were run on a gel (inset), ZIKV protease was identified in peak III. Peak I and II were oligomeric fractions of the protease.

In comparison to the ZIKV protease, the SDS-PAGE gel of the ZIKV helicase after the first affinity column contained fewer bands apart from the Trx-His₆ tagged ZIKV helicase at ~63 kDa (Fig. 3.2.13). After tag-cleavage and a second affinity chromatography, one single band was observed at 50 kDa (Fig. 3.2.13). The protein was then applied to SEC, where a single peak for helicase was observed (Fig. 3.2.13). When fractions were pooled from this peak (shaded region; Fig. 3.2.13) and run on gel, a pure and homogenous ZIKV helicase was observed. The protein was then concentrated to three concentrations for SAXS experiments.

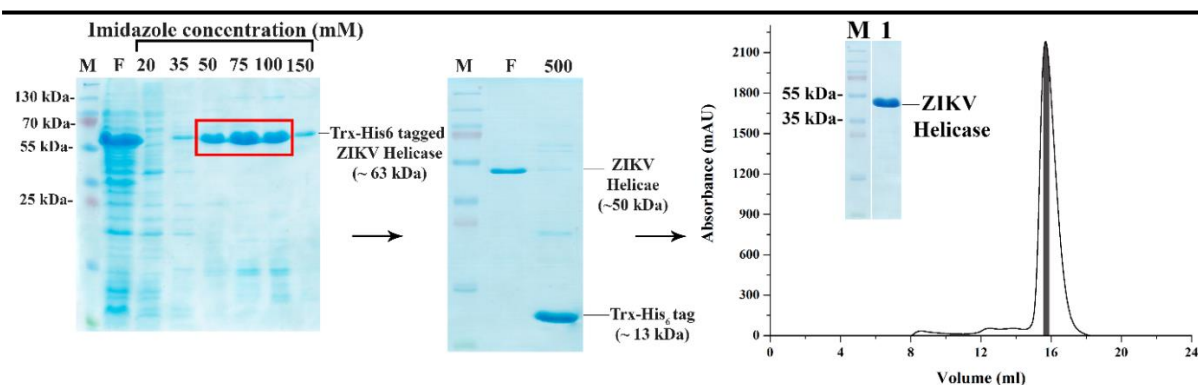


Figure 3.2.13: Purification of ZIKV helicase. The first SDS-PAGE gel shows the imidazole concentrations at which the tagged protein eluted from the Ni-NTA column (M=Protein marker; F=flow-through). The red box indicates the fractions pooled for overnight cleavage of the tag. After second affinity chromatography, the protein eluted in the flow-through (F). On application to SEC, a single peak was observed. When fractions from 15% of peak maximum was pooled, concentrated and run on gel (inset), a homogeneous and pure helicase was observed.

3.2.9 Primary SAXS data analysis for ZIKV NS3 domains

SAXS data for ZIKV protease was collected at 1.6, 2.1 and 3.6 mg/ml of the recombinant protein (Fig. 3.2.14A). Guinier region analysis indicated no aggregation at all three protein concentrations (Fig. 3.2.14A, inset) and the 2.1 mg/ml dataset was used for further analysis. The R_g value extracted from Guinier approximation was 2.05 ± 0.04 nm and that derived from $P(r)$ function was 2.01 ± 0.01 nm. The recombinant protein had an estimated D_{max} of 6.1 ± 1 nm in solution (Fig. 3.2.14B) (table 3.2.2).

The Guinier plots for all three protein concentrations (1.3, 2.4 and 3.2 mg/ml) of ZIKV helicase were linear indicating no inter-particle interaction (Fig. 3.2.14C, inset). The 2.4 mg/ml dataset was used to derive the R_g value from Guinier approximation and the $P(r)$ function, which were 2.31 ± 0.03 nm and 2.35 ± 0.01 nm, respectively. The D_{max} value obtained from $P(r)$ function was 6.5 ± 1 nm (Fig. 3.2.14D) (table 3.2.2).

Protein sample	ZIKV Protease	ZIKV Helicase
Concentration range (mg ml ⁻¹)	1.60 – 3.60	1.30 – 5.00
Temperature (K)	288.15	288.15
Structural parameters†		
$I(0)$ (cm ⁻¹) [from P(r)]	40.94 ± 0.52	63.82 ± 0.62
R_g (nm) [from P(r)]	2.05 ± 0.02	2.31 ± 0.03
$I(0)$ (cm ⁻¹) (from Guinier)	41.25 ± 0.33	64.65 ± 0.36
R_g (nm) (from Guinier)	2.08 ± 0.01	2.35 ± 0.01
D_{max} (nm)	6.1 ± 0.6	6.5 ± 0.7
Porod volume estimate (nm ³)	~37	~67
DAMMIF excluded volume (nm ³)	~55	~93
Dry volume from sequence (nm ³) ‡	~30	~61
Molecular mass determination		
Calculated monomeric MM (kDa)	~25.4	~50.6
Molecular mass MM (kDa) [from <i>Porod invariant</i>]	23.1 ± 2.3	41.8 ± 4.2
Molecular mass MM (kDa) [from <i>excluded volume</i>]	27.5 ± 2.8	46.5 ± 4.7

Table 3.2.2: Overall structural parameters of ZIKV NS3 protease and helicase domains

3.2.10 Assessment of ZIKV NS3 domains flexibility

Both the ZIKV helicase and the protease domain are compact when compared to the full-length protein. Both domains adopt a bell-curved profile similar to the globular lysozyme in the normalized Kratky plot (Fig. 3.2.14E). The helicase has a peak maximum very close to that of well folded lysozyme ($\sqrt{3}$, 1.104) and is slightly more compact than the protease domain.

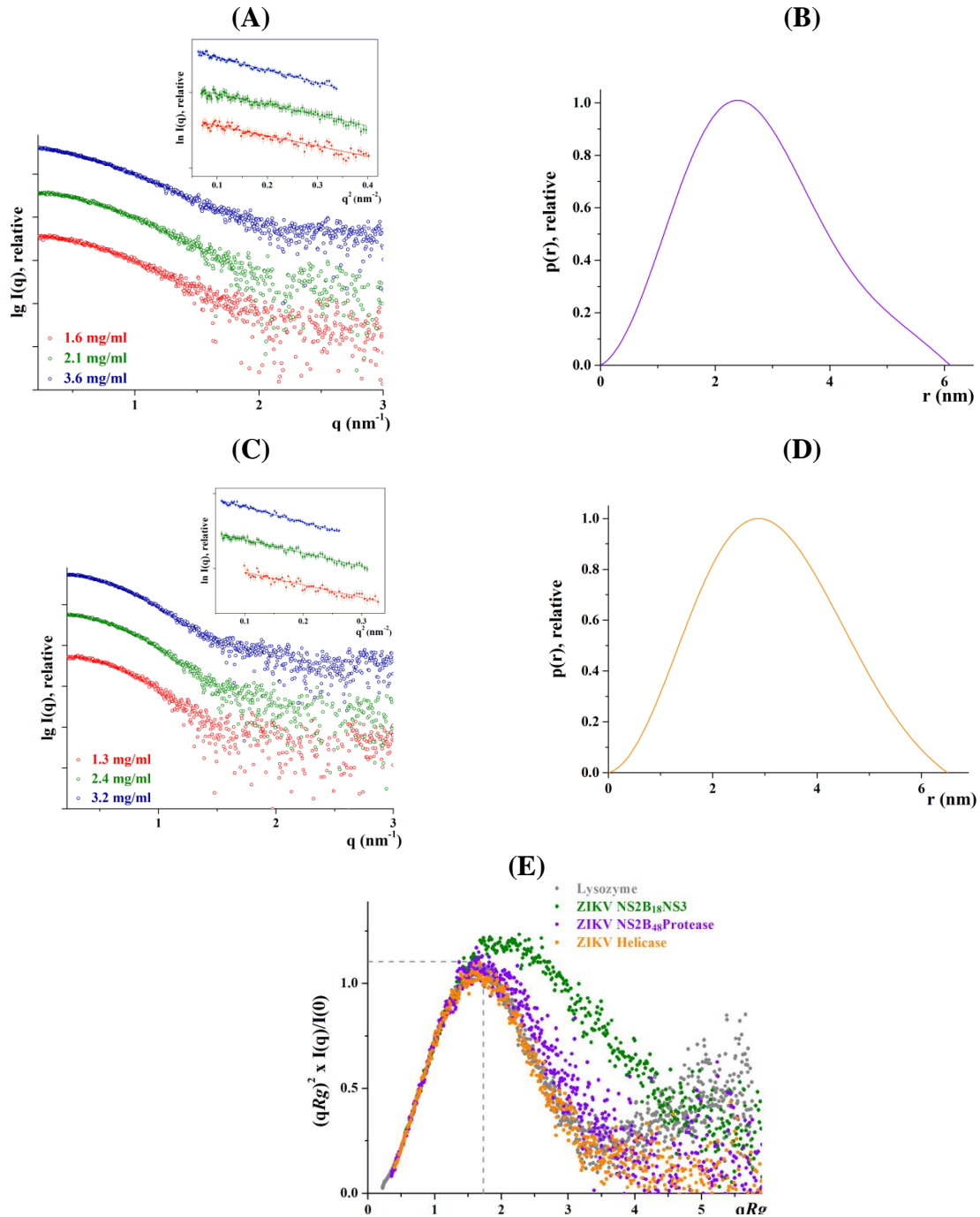


Figure 3.2.14: SAXS studies of ZIKV NS2B₁₈NS3 protease and helicase domains. Scattering pattern (\circ) of (A) ZIKV protease at 1.6 mg/ml (red), 2.1 mg/ml (green) and 3.6 mg/ml (blue) and (C) ZIKV helicase recorded at 1.3 mg/ml (red), 2.4 mg/ml (green) and 3.2 mg/ml (blue). *Inset:* Linearity of the Guinier plot at low angles indicates no sample aggregation. $P(r)$ curve of (B) protease and (D) helicase. Both curves indicate a compact globular protein. (E) Comparison of the dimensionless Kratky plot of ZIKV NS2B₁₈NS3 (\bullet ; green), protease (\bullet ; magenta) and helicase (\bullet ; orange).

3.2.11 *Ab initio* model constructions for ZIKV NS2B₁₈NS3 protease and helicase

Low resolution shapes for the ZIKV NS3 domains were reconstructed using DAMMIF. The program was run ten times independently for each protein. The final averaged model for the two domains are depicted in figure 3.2.15 that are convergent with NSD values of $0.53 \pm$

0.01 and 0.54 ± 0.01 for the protease and the helicase, respectively. The *ab initio* envelopes superimposed well with the crystal structure of protease (PDB ID: 5LC0) and helicase (PDB ID: 5GJB) (Fig. 3.2.15).

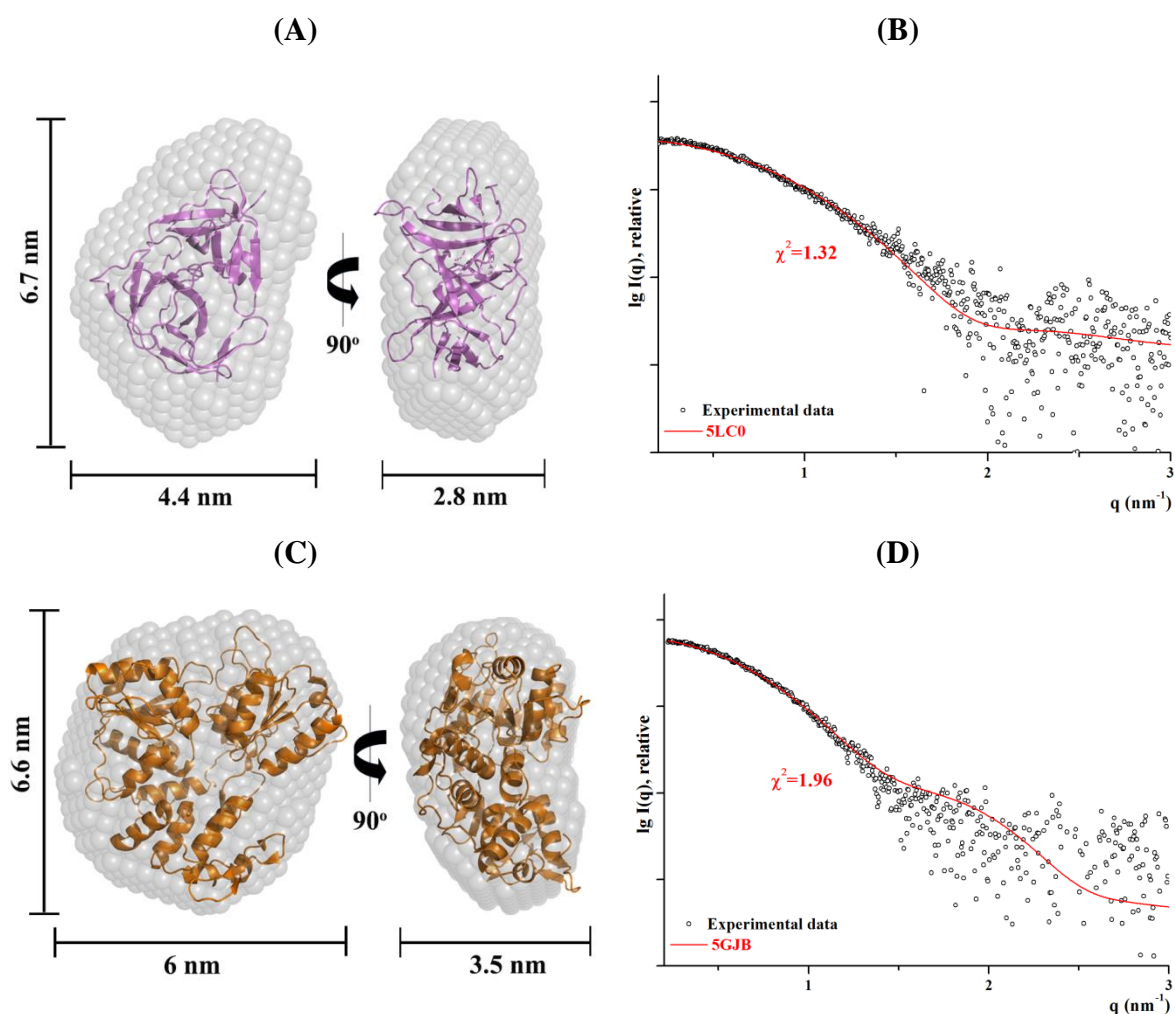


Figure 3.2.15: Solution envelopes of ZIKV protease and helicase. Overlay of DAMMIF envelope onto the cartoon representation of the crystal structure of (A) protease and (C) helicase. The protease is colored in purple and helicase in orange. The fit of the experimental data of (B) protease and (D) helicase to the theoretical scattering curve from their respective crystal structures.

3.2.12 Cloning, expression and purification of ZIKV NS2B₁₈NS3 mutants

Single amino acid residues such as T174, V176, E177, C178, and F179 were mutated to A to investigate the importance of linker residues on the overall structural parameters of the protein. Besides mutating the residue F179 to alanine, this residue was also mutated to isoleucine to investigate the consequence of substituting the DENV-2 linker residue in its place. The resultant PCR products were run on a gel. All the mutants revealed a single band corresponding to ~8000 bp of the DNA ladder (Fig. 3.2.16).

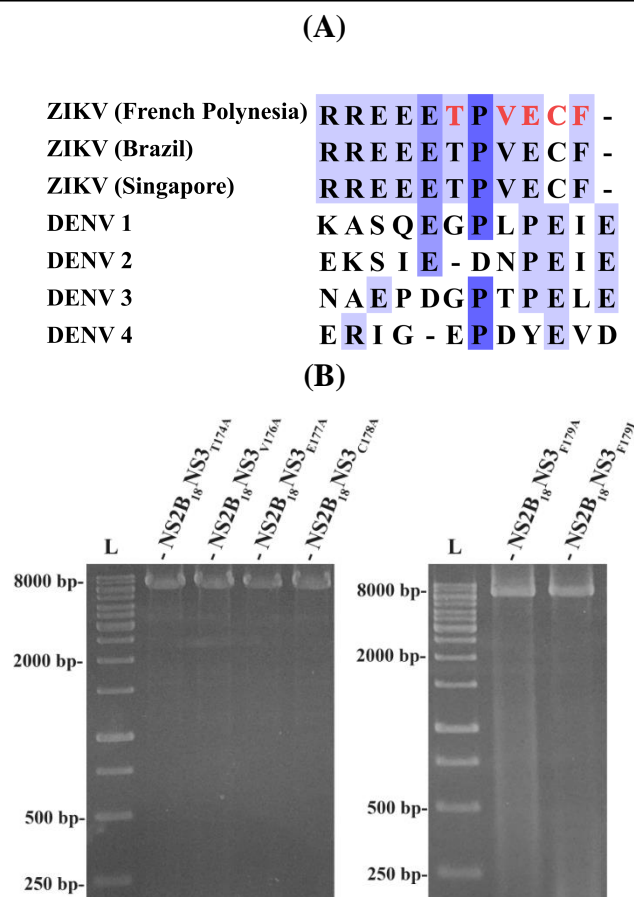


Figure 3.2.16: Site-directed mutagenesis of ZIKV NS2B₁₈NS3. (A) The sequence alignment of the NS3 linker region. The residues highlighted in red in the sequence were mutated to alanine. (B) The figure depicts the 1.5% agarose gel on which the mutated PCR products were run. Single bands are observed for all the mutants. DNA ladder is loaded in lane marked L.

The sequence verified plasmids were transformed into BL21-CodonPlus (DE3)-RIL cells for overproduction of recombinant proteins. Protein production for all constructs was confirmed by the presence of a ~ 84 kDa protein band in an SDS-PAGE gel (Fig. 3.2.17). Both F179 mutants had higher production of the recombinant proteins when compared to rest of the mutants.

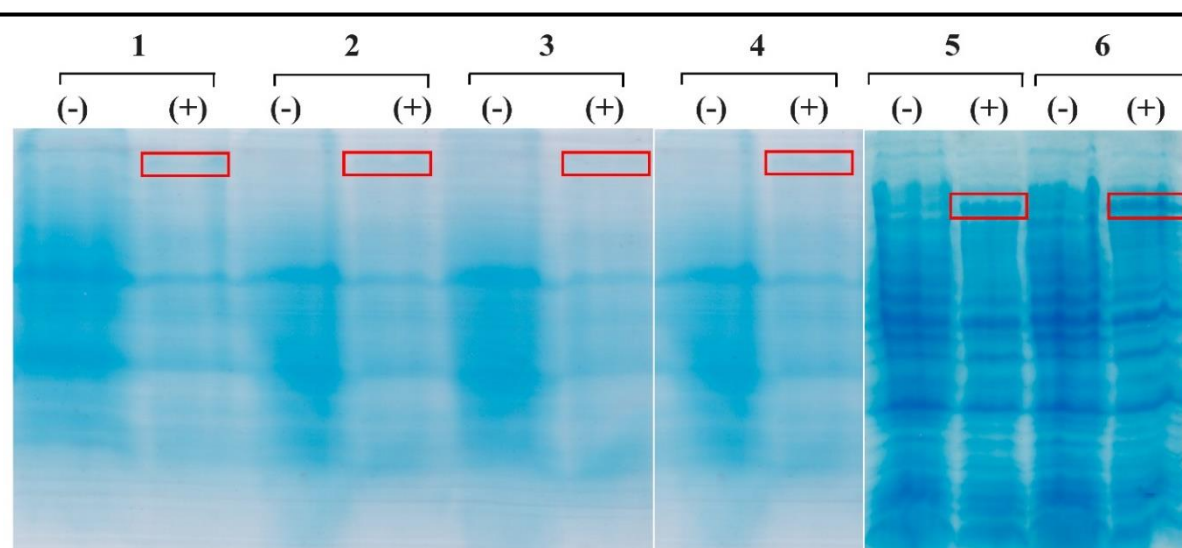


Figure 3.2.17: Recombinant ZIKV NS3 mutant production. SDS-PAGE gels showing production of (1) NS2B₁₈NS3_{T174A}, (2) NS2B₁₈NS3_{V176A}, (3) NS2B₁₈NS3_{E177A}, (4) NS2B₁₈NS3_{C178A}, (5) NS2B₁₈NS3_{F179A} and (6) NS2B₁₈NS3_{F179L}. The recombinant protein bands are highlighted in black boxes. (-) non-induced and (+) induced samples. Samples in (5) and (6) were run on a separate gel than the other samples and the bands for the recombinant protein appear lower as the dye front ran out for this gel. Also, the gel with samples 5 and 6 were stained for a longer time, hence the protein bands appear darker.

The recombinant protein variants were purified by affinity chromatography as the first step to obtain a pure protein for SAXS. The SDS-PAGE gels for all the mutants showed a severe laddering in all the lanes (Fig. 3.2.18). For all the variants, the respective protein eluted from the Ni-NTA column in an imidazole gradient from 35 – 100 mM. The fractions indicated in figure 3.2.18 were pooled and the tag was cleaved by thrombin, before proceeding to the next affinity chromatography. The flow-through for NS2B₁₈NS3_{T174A}, NS2B₁₈NS3_{V176A} was concentrated and applied to SEC. The protein that eluted from SEC appeared degraded and was unsuitable for further studies (Fig. 3.2.18E).

The NS2B₁₈NS3_{E177A}, NS2B₁₈NS3_{C178A} eluted in the flow-through from the Ni-NTA column in a second step (Fig. 3.2.18C-D). Fractions marked in the SEC chromatogram (Fig. 3.2.18E) were applied on an SDS-PAGE gel and it was observed that mostly degraded protein eluted from the column. These samples were not suitable for SAXS studies (Figure 3.2.18F).

Interestingly for the NS2B₁₈NS3_{F179A} and NS2B₁₈NS3_{F179L}, the protein co-eluted with the Trx-His₆ tag in 500 mM imidazole from the Ni-NTA column (Fig. 3.2.19A-B). To remove the high concentration of imidazole before applying the protein samples to the SEC column, buffer exchange was performed while concentrating the protein. Two distinct peaks for both proteins were observed in the chromatograms and fractions from the shaded region of each peak were run in an SDS-PAGE gel. Peak II was identified as the cleaved tag. Fractions from the first peak containing the protein were pooled, concentrated for SAXS measurement (Fig. 3.2.19C).

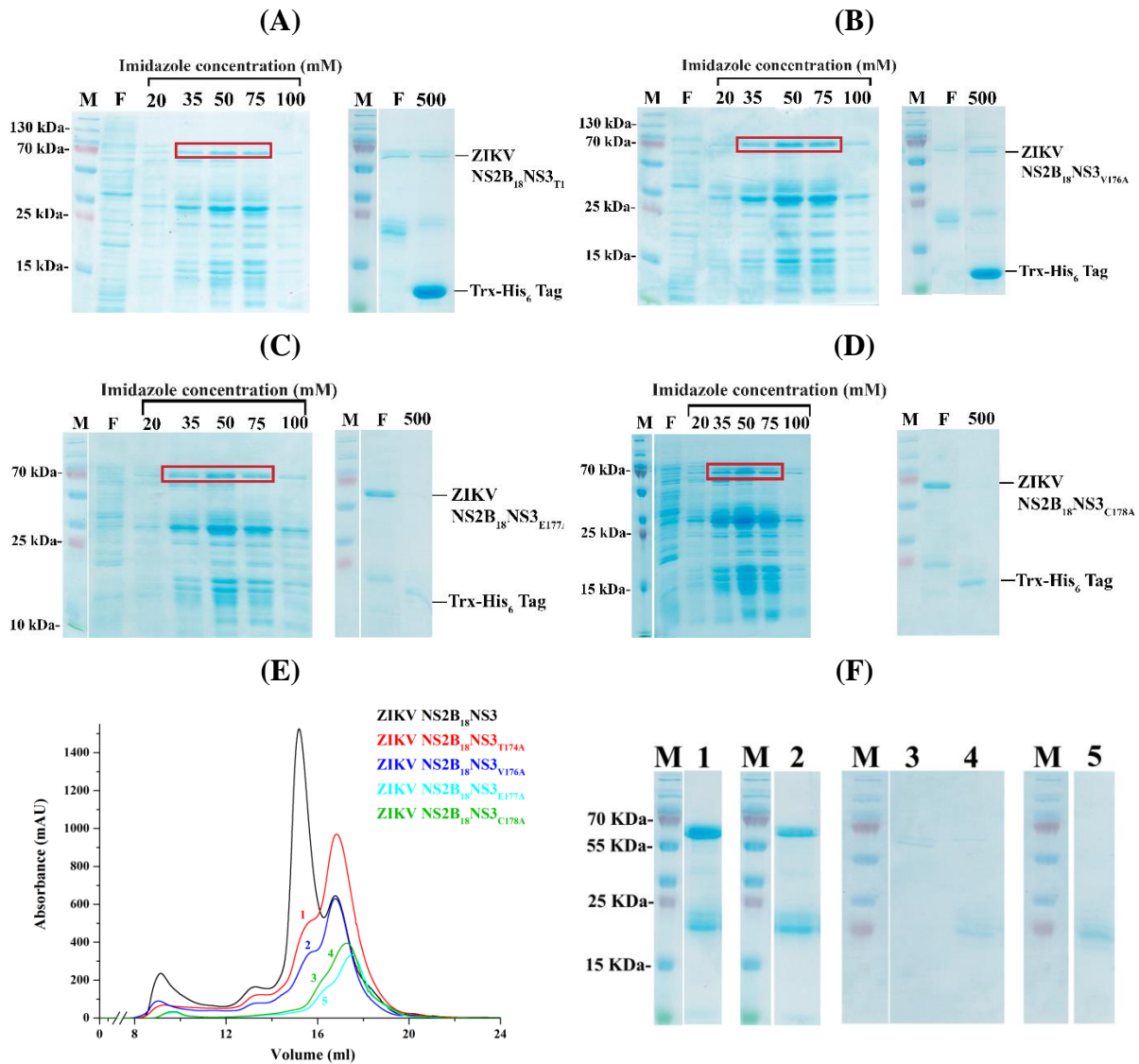


Figure 3.2.18: Purification of ZIKV linker mutants. The SDS-PAGE gel from first and second affinity chromatography for (A) ZIKV NS2B₁₈NS3_{T174A}, (B) ZIKV NS2B₁₈NS3_{V176A}, (C) ZIKV NS2B₁₈NS3_{E177A} and (D) ZIKV NS2B₁₈NS3_{C178A}. The mutant proteins appear degraded in both SDS-PAGE gels. (E) Comparison of SuperdexTM 200 Increase 10/300 column elution diagrams of purified ZIKV NS2B₁₈NS3 (black) and NS2B₁₈NS3_{T174A} (red), NS2B₁₈NS3_{V176A} (blue), NS2B₁₈NS3_{E177A} (cyan) and NS2B₁₈NS3_{C178A} (green). (F) SDS-PAGE gels of purified ZIKV linker mutants. Protein marker has been loaded in the lanes marked M. The peak fraction of NS2B₁₈NS3_{T174A} (1), NS2B₁₈NS3_{V176A} (2), NS2B₁₈NS3_{E177A} (5) and NS2B₁₈NS3_{C178A} (3, 4) were applied on the gel. Degradation of the purified linker mutants was observed on the gel, rendering not suitable for SAXS studies.

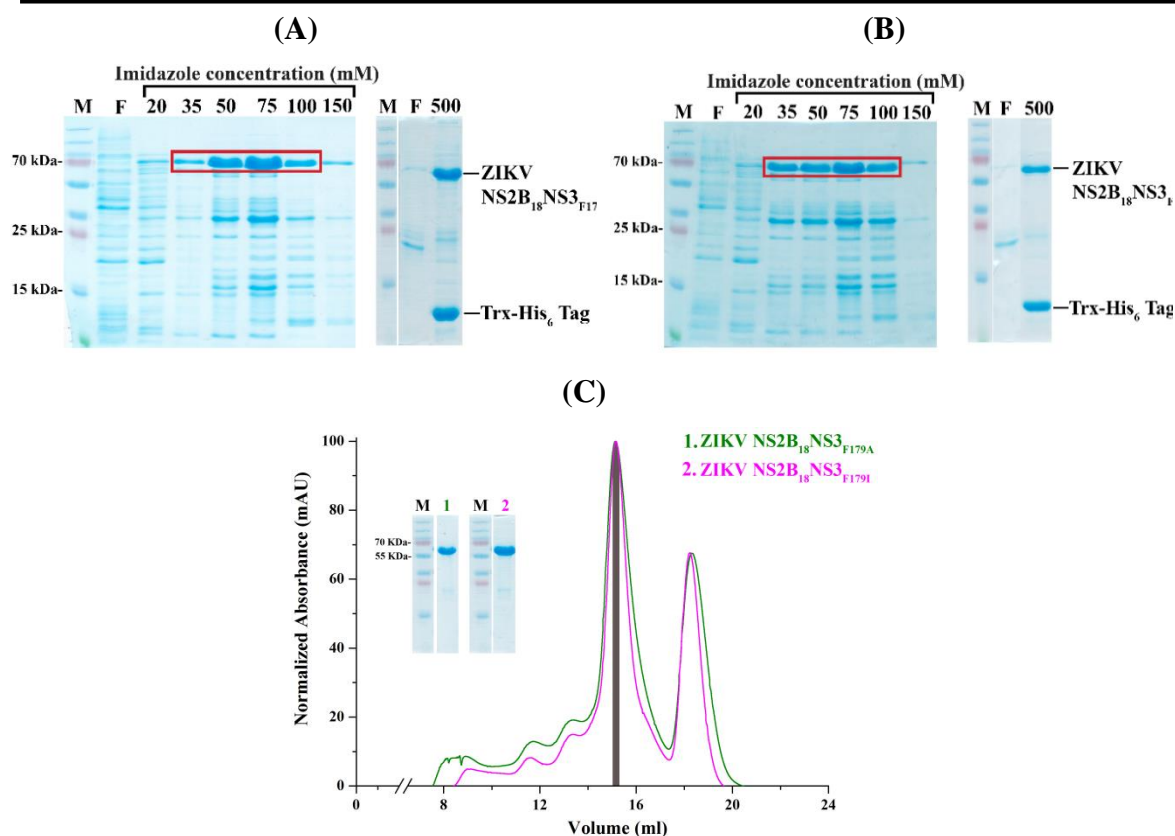


Figure 3.2.19: Purification of ZIKV NS3 F179 mutants. The SDS-PAGE gel from the first and second affinity chromatography for (A) ZIKV NS2B₁₈NS3_{F179A}, (B) ZIKV NS2B₁₈NS3_{F179I}. (C) Elution diagrams of ZIKV NS2B₁₈NS3_{F179A} (green) and NS2B₁₈NS3_{F179I} (magenta) purified using Superdex™ 200 Increase 10/300 column. The fractions from the shaded region were utilized for SAXS studies. Inset: the SDS-PAGE gels show homogeneous populations of NS2B₁₈NS3_{F179A} (green, 1) and NS2B₁₈NS3_{F179I} (magenta, 2), which are suitable for SAXS studies. Protein marker has been loaded in the lanes marked M.

3.2.13 Overall structural parameter of ZIKV NS2B₁₈NS3_{F179A} and NS2B₁₈NS3_{F179I}

SAXS-patterns of ZIKV NS2B₁₈NS3_{F179A} (2.5 mg/ml) and NS2B₁₈NS3_{F179I} (6.0 mg/ml) were recorded as shown in figure 3.2.20A. The Guinier plots appeared linear, indicating there was no aggregation of protein (Fig. 3.2.20A, inset). From the Guinier approximation, the R_g for NS2B₁₈NS3_{F179A} was 3.38 ± 0.06 and 3.31 ± 0.05 nm for NS2B₁₈NS3_{F179I}. The $P(r)$ function is shown in figure 3.2.20B. A D_{max} of 11.0 ± 1 nm for both mutants, signifying that like the WT ZIKV NS2B₁₈NS3, the mutant proteins are also elongated in solution. The R_g values obtained from the $P(r)$ function were 3.48 ± 0.02 nm and 3.43 ± 0.02 nm for ZIKV NS2B₁₈NS3_{F179A} and NS2B₁₈NS3_{F179I}, respectively.

The normalized Kratky plot of the mutant proteins provided the information about the conformational behavior of these protein. In comparison to the globular lysozyme, the normalized Kratky plot of both ZIKV NS2B₁₈NS3_{F179A} and NS2B₁₈NS3_{F179I} exhibited a wide bell-shaped profile skewed to the right (Fig. 3.2.20C), indicating that both are extended and

display a array of motion. The profile is comparable to that of the WT ZIKV NS2B₁₈NS3 (Fig. 3.2.20C) [78].

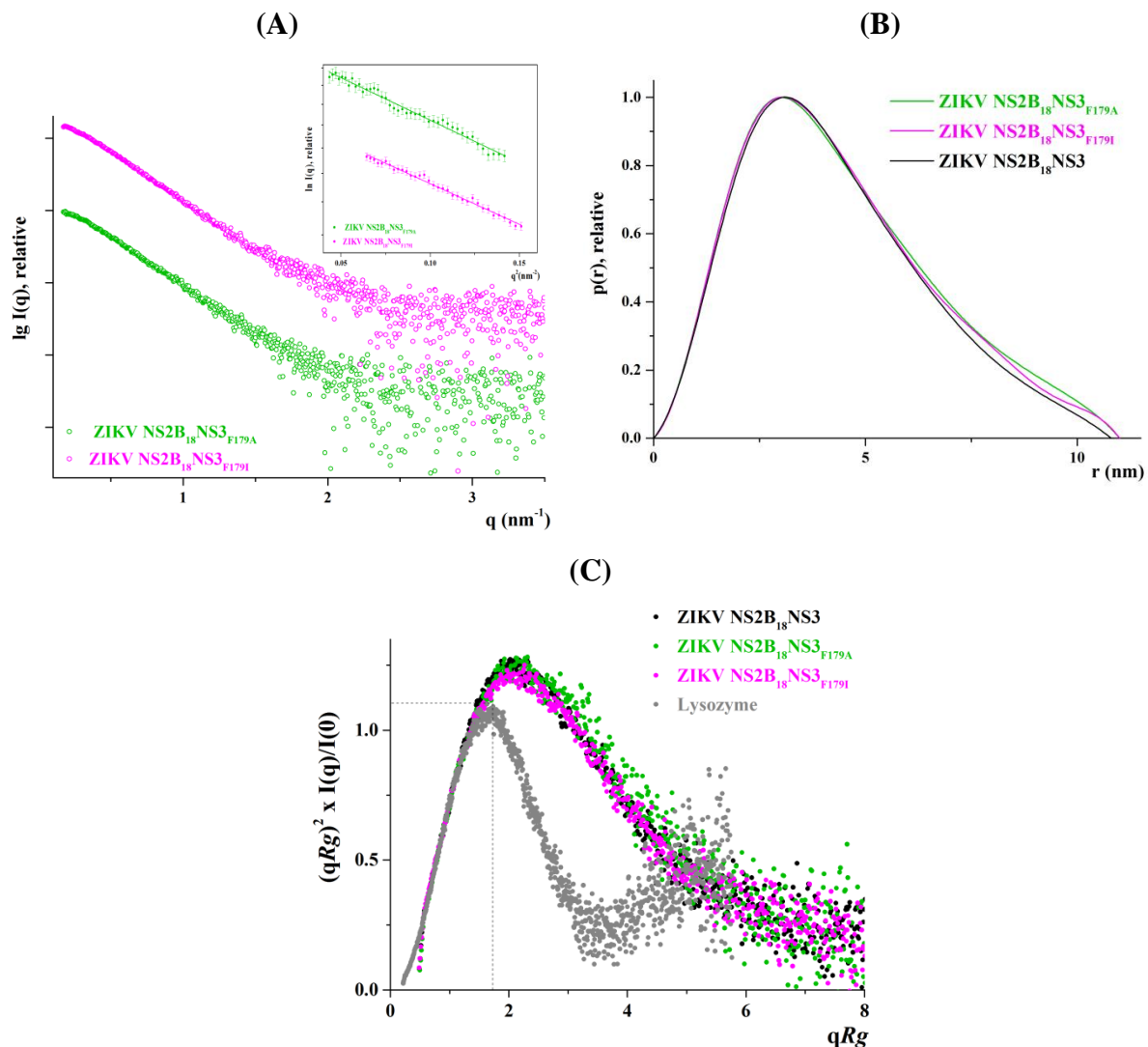


Figure 3.2.20: SAXS studies of ZIKV NS2B₁₈NS3 mutant. (A) Scattering pattern (\circ) of ZIKV NS2B₁₈NS3_{F179A} mutant at 2.5 mg/ml (green), and ZIKV NS2B₁₈NS3_{F179I} 6.0 mg/ml (magenta) *Inset:* No aggregation of protein sample was observed as seen from the linearity of Guinier plot. (B) Pair-wise distance distribution $P(r)$ curve of WT (—; black), ZIKV NS2B₁₈NS3_{F179A} (—; green), and ZIKV NS2B₁₈NS3_{F179I} (—; magenta). (C) Comparison of the dimensionless Kratky plot of WT (\bullet ; black), ZIKV NS2B₁₈NS3_{F179A} (\bullet ; green), and ZIKV NS2B₁₈NS3_{F179I} (\bullet ; magenta). The plot indicates that the mutants are flexible like the WT protein.

Protein sample	ZIKV NS2B ₁₈ NS3 _{F179A}	ZIKV NS2B ₁₈ NS3 _{F179I}
Concentrations (mg ml ⁻¹)	2.5	6.0
Structural parameters		
$I(0)$ (arbitrary unit) (from Guinier)	106.61 ± 1.36	78.86 ± 0.97
R_g (nm) (from Guinier)	3.38 ± 0.06	3.31 ± 0.05
$I(0)$ (arbitrary unit) [from $P(r)$]	106.80 ± 0.81	79.46 ± 0.56
R_g (nm) [from $P(r)$]	3.48 ± 0.02	3.43 ± 0.02
D_{max} (nm)	11.0 ± 1	11.0 ± 1
Porod volume estimate (V_p) (nm ³)	~93	~96
DAMMIF excluded volume (V_{ex}) (nm ³)	~118	~118
Dry volume from sequence (nm ³) ‡	~86	~86
Molecular mass determination		
Calculated monomeric MM (kDa) [from sequence*]	~71.27	~71.31
MM from Porod invariant ($V_p/1.6$) (kDa)	58.55 ± 5.8	60.26 ± 6.0
MM from excluded volume ($V_{ex}/2$) (kDa)	59 ± 5.9	59 ± 5.9

Table 3.2.3: Overall structural parameters of ZIKV NS2B₁₈NS3_{F179A} and NS2B₁₈NS3_{F179I}

3.2.14 ATPase activity of ZIKV NS2B₁₈NS3 and helicase

Here, ATP hydrolysis of the recombinant ZIKV NS3 and helicase from French Polynesia ZIKV was measured. An ATP hydrolytic activity of 596.7 U/mM of NS2B₁₈NS3 was determined, which was slightly lower compared to the single helicase domain (646.9 U/mM of protein; Fig. 3.2.22A). The ATPase activity of the ZIKV NS2B₁₈NS3 was significantly higher compared to values of the DENV-2 (474.5 U/mM of protein). The same is the case for the ZIKV helicase, whose activity was 15% (564.9 U/mM of protein) higher compared to the DENV-2 helicase.

Two ATPase inhibitors resveratrol and quercetin were used to further analyze the hydrolytic activity of the protein. As shown in figure 3.2.22, resveratrol (150 µM) and quercetin (150 µM) reduced the ATPase activity by ~74% and ~28%, respectively. In comparison, the inhibition by resveratrol is lower in the ZIKV helicase (51%) compared to the entire NS2B₁₈NS3, whereas quercetin reduced ATPase activity stronger in ZIKV NS2B₁₈NS3 than in its single helicase (15% inhibition).

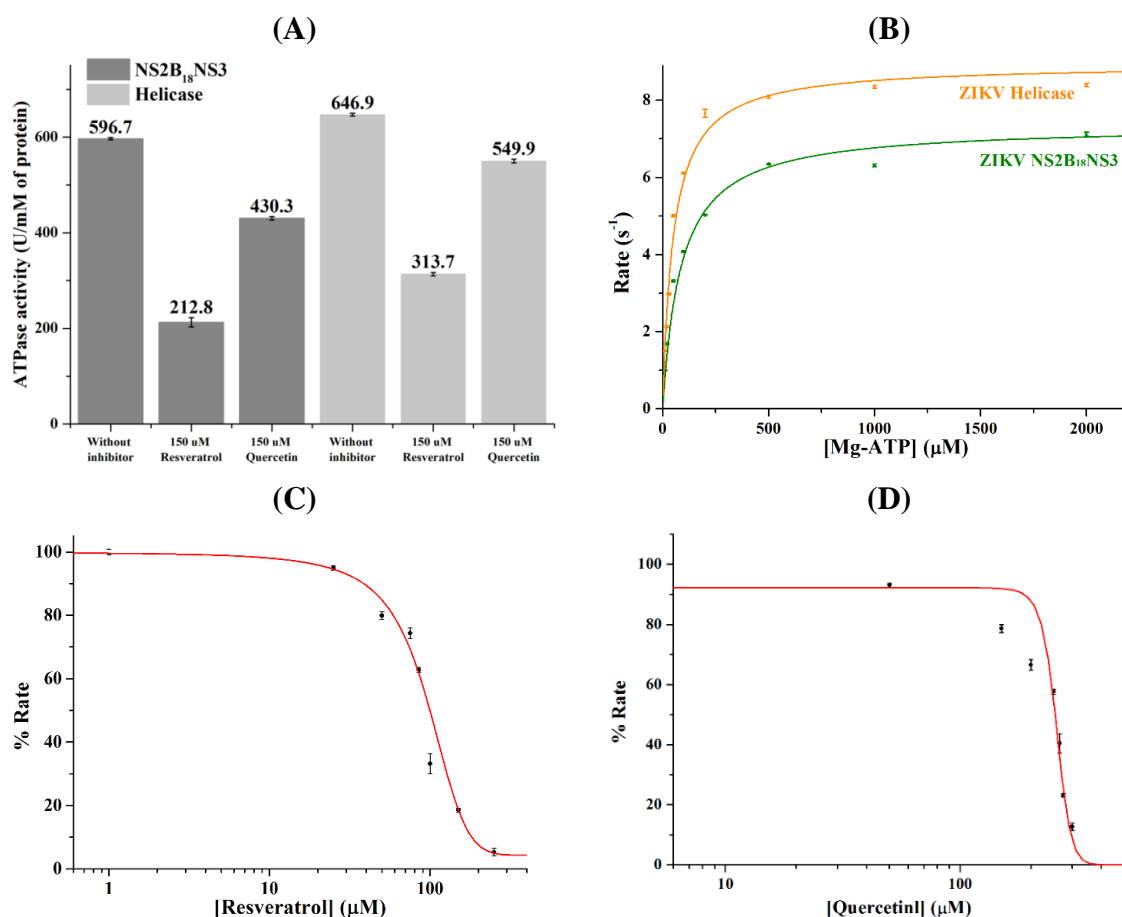


Figure 3.2.22: Hydrolysis activity assay of ZIKV proteins. (A) Bar diagram representation of ATP hydrolysis activity of ZIKV NS2B₁₈NS3 and helicase without and with inhibitors. (B) The Michaelis-Menten plots of ZIKV NS2B₁₈NS3 (green) and helicase (orange) with their fit lines for the determination of the enzymatic kinetic parameters of each protein. ATP hydrolysis of ZIKV NS2B₁₈NS3 was inhibited by (C) resveratrol with the IC₅₀ of $94.25 \pm 5.02 \mu\text{M}$, and (D) quercetin with the IC₅₀ of $256.17 \pm 13.3 \mu\text{M}$.

4. Discussion – Structural and Functional Characterization of Flaviviral NS3

4.1 Traits of NS3 in solution

Flaviviruses have emerged as major concern for human health with over 50-100 million cases [85] occurring each year and yet no effective antiviral therapy exists that effectively treats infections caused by the viruses [26, 83]. On studying the flaviviral replication cycle, it was deemed that NS3 and NS5 are indispensable for viral replication. Hence, these proteins are attractive drug targets Luo et al., 2015). In this thesis, the NS3 from dengue (serotype 2) and zika were studied by SAXS to gain insight into the solution structure and conformation of the protein. Understanding the characteristic behavior of the protein in solution along with its biochemical characterization, could go a long way to help in drug discovery against this protein.

The NS3 is a two domain multi-enzyme coded in a single polypeptide. The N terminal 1-169 residues are a chymotrypsin like serine protease that cleaves the viral polyprotein and is a crucial step in viral protein maturation [192]. The C-terminal helicase (179-618) possesses RNA helicase, RNA-dependent NTPase and 5' RNA triphosphatase activities [26]. The helicase has three subdomains, whereby domains I and II adopt a RecA-like fold, while domain III forms a single-stranded RNA binding tunnel. Based on biochemical, genetic and biophysical studies [71, 193, 194], the interacting sites between NS3 and NS5 were mapped to the subdomain III of NS3 helicase (residues 566-585) and bNLS-region of the NS5 RdRp (residues 320-368).

4.1.1 Structural features of DENV-2 NS3 in solution

Till date, two conformations were recognized from the crystal structures of DENV-4 NS2B₁₈NS3, called conformation I (PDB ID: 2VBC [80]) and conformation II (PDB ID: 2WHX [38]. By comparing the two conformations it is observed, that the protease in conformation II is rotated 161° compared to the one in conformation I [38]. In addition, in conformation I, the protease blocks the ATP-binding site of the helicase whereas in conformation II the ATP-binding site is accessible for ADP-binding due to the rotation of the protease. To ascertain the domain arrangements in DENV-2 NS3, the solution structure of DENV-2 NS3FL was derived from the SAXS experiment.

The Guinier analysis of the protein indicates a monodisperse solution. The estimated particle dimension of 10.3 ± 1 nm, as well as the right skewed $P(r)$ curve indicates that the DENV-2 NS3 is an extended molecule. The estimated molecular weight from the Porod volume (~ 97 nm³) was 65.5 ± 6.6 kDa, showing that DENV-2 exists as a monomer in solution.

The normalized Kratky plot reaffirmed that the molecule is extended in solution as well as being highly flexible. When this flexibility was quantitated using EOM 2.0, seven representative models of probable conformations of NS3 in solution were identified. The ensemble of conformations has a unimodal R_g distribution with a main peak around 3.1 to 3.2 nm. Thus, the inter-domain distance of the two domains of DENV-2 NS3 is predominantly extended in solution. Also, the R_g of the crystal structure (PDB ID: 2VBC) (2.98 nm) was within the range of R_g distribution for the selected ensemble pool. The CRY SOL fitting shows (Fig. 3.1.4B) that conformation II could be the more preferred conformation in solution. Knowing that the protease domain of this protein cleaves in *cis* at the NS2A-NS2B junction and in *trans* at the NS3-NS4A junction [192, 195], it is not surprising that the protease domain can adopt a flexible motion, so that it can re-orientate and extend itself, to carry out its functions.

4.1.2 The linker residues of flaviviruses impart flexibility

The dimensionless Kratky plot (Fig. 3.1.5A; section 3.1.4) indicates that both the protease and helicase domain of DENV-2 NS3 is compacted and rigid in solution. The flexibility of the full-length protein is thus attributed to the linker. The slight tail observed in the $P(r)$ curve could be due to the presence of the linker residues (11/12 amino acids) in both domains. In flaviviruses the linker residues show sequence variation and is an acid rich linker [38] (Fig 4.1.1A). The less conserved amino acid sequence of flavivirus NS3 linker permits the protein to adopt an elongated conformation [83]. Evolutionarily, the length of the flaviviral NS3 inter-domain linker has been optimized so that it can impart flexibility to the protein when undergoing conformational changes during viral polyprotein processing and RNA replication.

In contrast, the linker connecting the HCV NS3 domains has high sequence conservation among various strains with an invariant PPxxP motif [196]. This proline rich conserved linker levies a compact conformation for the HCV NS3 (Fig 4.1.1B). Thus, while DENV NS2B₁₈NS3 resorts to an extended conformation, the crystallographic structure of the related NS3₃₋₆₃₁NS4A₂₁₋₃₂ (PDB ID: 1CU1; [197]) from HCV adopts a more globular conformation, with the NS3 protease active site positioned towards the molecular interior [198] (Fig 4.1.1B).

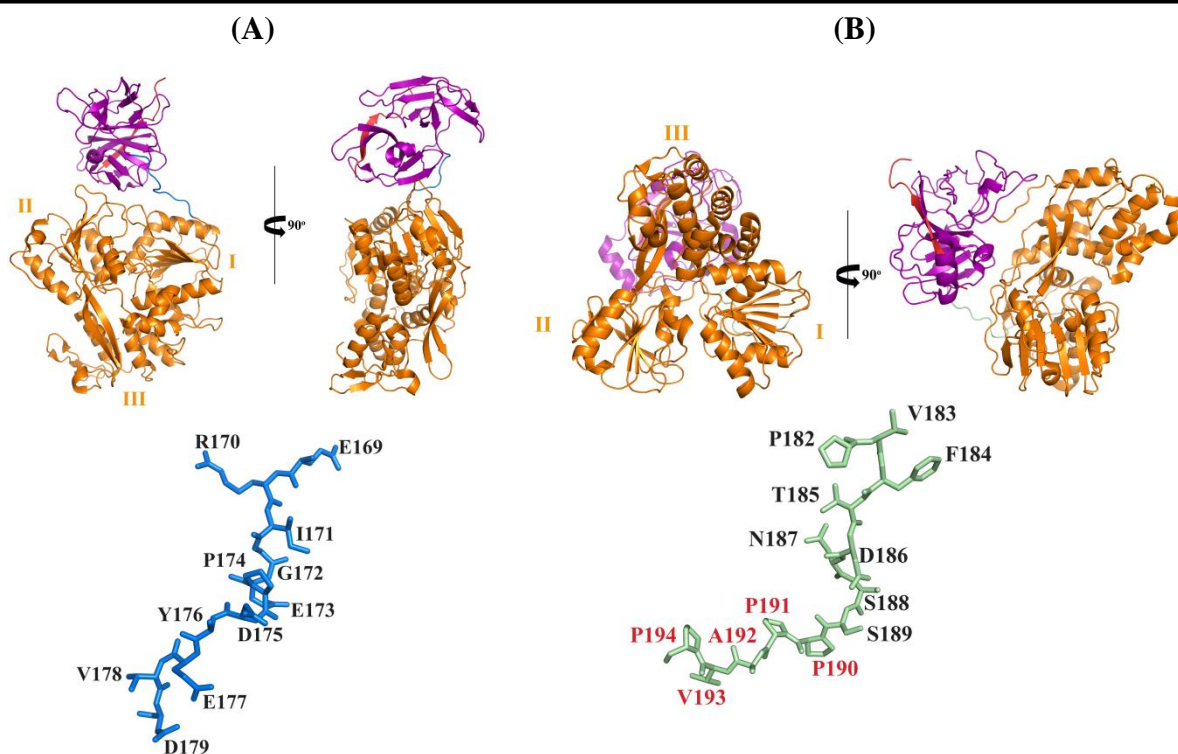


Figure 4.1.1: Dengue NS3 adopts an extended conformation while HCV NS3 is compact. (A) The top part shows the cartoon representation of the crystal structure (PDB ID: 2VBC) from DENV-4 NS3, with the NS2B cofactor in red, protease in purple, linker in blue and helicase in orange. The three domains of helicase are designated as I, II, III. The front and the side views are shown. The protease of the dengue NS3 is situated above the helicase domain. The bottom part shows the acid rich region of the linker in sticks. (B) The top part shows the cartoon representation of the crystal structure (PDB ID: 1CU1) of HCV NS3 with the NS4A cofactor in red, protease in purple, linker in blue and helicase in orange. Here also the three domains of helicase are marked as I, II, III. The front and side view of the structures are shown. The proline rich linker of HCV NS3 is shown at the bottom in sticks. The invariant PPxxP motif is labeled in red. Comparing both the linkers, the HCV NS3 itself is compact, allowing the protease to be positioned at the back of helicase.

4.1.3 The overall structural parameter of the mutant DENV-2 NS3 is similar to the WT

When the invariant PPAVP motif of HCV NS3 linker was substituted into the DENV-2 NS3, the extended feature of the protein did not alter. Instead, the flexibility of the mutant protein was slightly higher than WT DENV-2 NS3 (Fig. 3.1.14C). Although a solution study of the entire HCV NS3₃₃₋₆₃₁NS4A₂₁₋₃₂ is missing, it is speculated that the compact form observed in the crystal structure of HCV NS3 could be caused by crystal packing, where C-terminus of NS3 came in proximity to the protease active site [197].

4.1.4 Both NS3 and its helicase domain were enzymatically active

The DENV-2 helicase in our study, in comparison to the NS3, possesses higher ATP hydrolysis activity. This in contrast to the previous studies where the DENV-2 NS2B₄₇NS3 and its helicase domain have similar ATPase activity [84]. This difference could be a consequence of different constructs being used for the assay as well as the set-up of the assay

[83]. The use of ATPase inhibitors such as resveratrol and quercetin reveal that both inhibitors inhibit the full-length as well as helicase. While resveratrol strongly inhibits DENV-2 helicase, the quercetin exerts a stronger inhibitory effect on the full-length protein. Computational studies have previously revealed that quercetin binds allosterically to DENV-3 NS3 [199]. The interaction is predicted to be mediated by R73, N74, K 88, I123, G124, N152, Q167 of the protease domain. Of these K88 also interacts with the linker, which hints that the quercetin binding to the protease can affect linker residues. The linker residues of the DENV NS3 also interacts with the P-loop residues and can thus influence the ATPase activity of the protein. So, binding of quercetin to the protease may strongly inhibit ATPase activity of NS3, D175 of the linker interacts with R202 of the P-loop via salt-bridge interactions and residues G173 and D179 hydrogen bonds with the I182 and R202. Any alteration of these interaction could also lead to altered NS3 activity and this region may be targeted for inhibiting the protein.

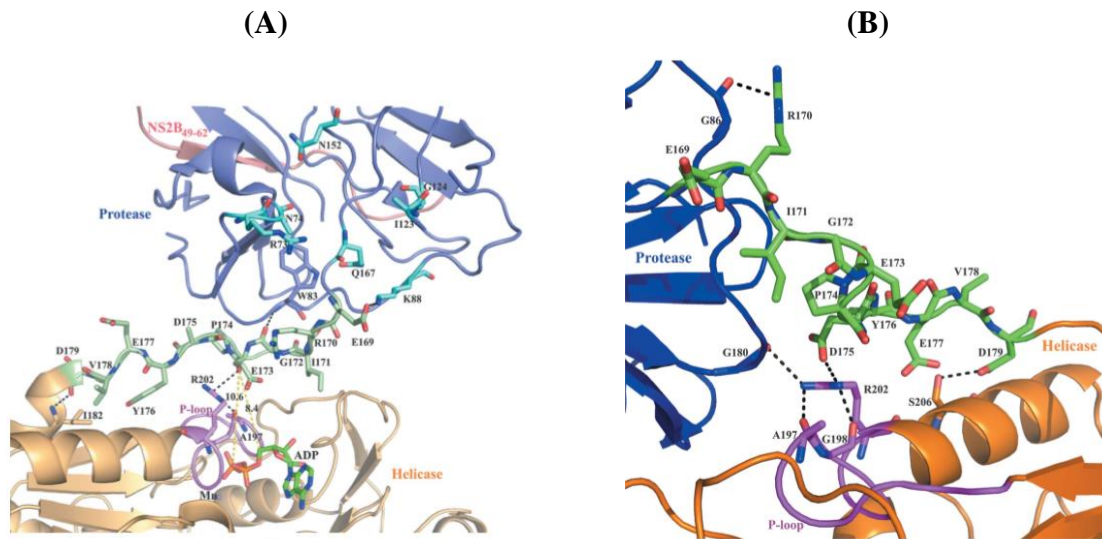


Figure 4.1.2: Linker residues (₁₆₉ERIGEPDYEVD₁₇₉) in DENV4 NS3 crystal structure. (A) Representation of the DENV-4 NS3 (PDB ID 2WHX) with the zoomed in view of the linker residues. Conformation II of the NS3 is depicted in this figure. The figure shows ADP (green sticks) and manganese (violet spheres) are bound to the NS3 active site. The protease and helicase domains are shown in blue and orange, respectively. The linker region is shown in green and the P-loop of the helicase domain is in magenta. The R73, N74, K 88, I123, G124, N152, Q167 of the protease domain are predicted to interact with quercetin (residues colored in cyan) [199]. E169 of the linker interacts with K88 of the protease domain via salt bridge and G172 forms backbone hydrogen bonds with W83 (residues shown in sticks). Apart from interacting with protease, the E173 and D179 of the linker also interacts with R202 and I182 of the helicase via hydrogen bonds, respectively. (B) This figure shows conformation I, where the linker residue R170 hydrogen bonds with G86 of the protease (PDB ID 2VBC). D175 of the linker interacts with R202 of the P-loop via salt-bridge interactions. In addition, R202 also hydrogen-bonds with G80 of the protease and A197 and Gly198 of the P-loop in the helicase. These interactions reveal that the linker residue via its interaction with the two domains, influence their activity.

4.2 ZIKV NS3 traits in solution

4.2.1 The recombinant ZIKV NS3 expressed in *E. coli* undergoes spontaneous auto-lysis

When the ZIKV NS3 was co-expressed with 48 amino acid residues from the NSB region, auto-lysis was observed in the resulting fusion protein (Fig. 3.2.6). The central region of NS2B

(49-97) is crucial for protease to fold correctly (Fig 4.2.1A). In helping the protease to fold correctly, the NS2B helps form the active site of the protease to generate a catalytically active protease [198]. The auto-lysis observed in the recombinant protein was hypothesized to be due the active protease. To prevent this auto-cleavage, the proposed critical residue S135 in the catalytic triad of the NS3 protease was mutated (Fig. 4.2.1B). However, this recombinant mutant was not stable. Therefore, a shorter NS2B (18 amino acid residues) was fused to the NS3 to generate the ZIKV NS2B₁₈NS3. While the 48 amino acids are required to form a catalytically active protease, only 18 amino acids (49-67) are required to properly fold the NS3 protease. ZIKV NS2B₁₈NS3 was used to study the solution structure of the ZIKV NS3.

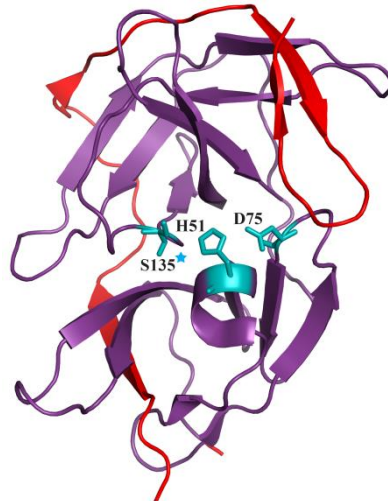
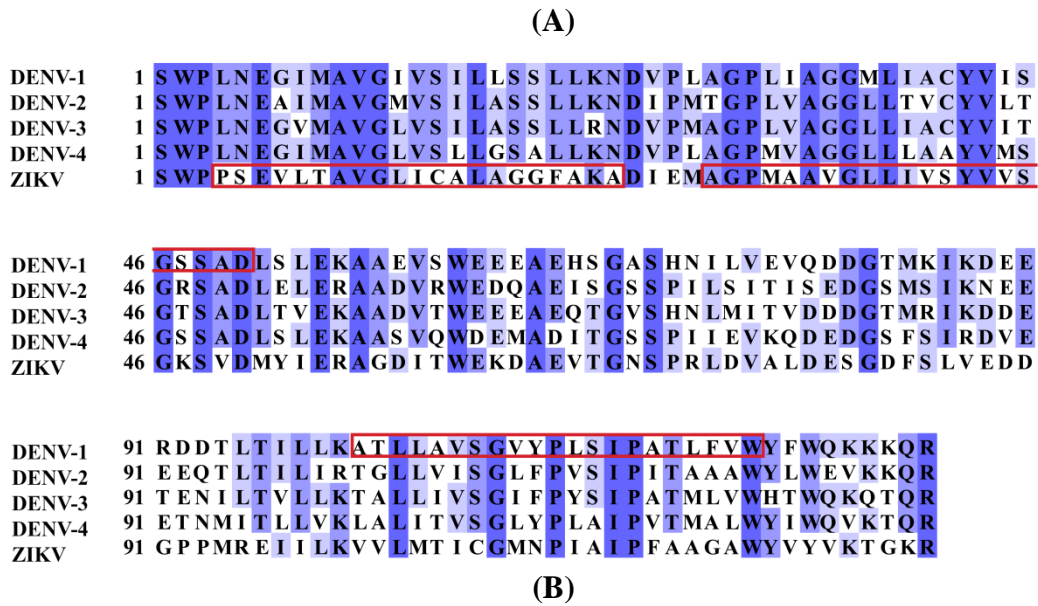


Figure 4.2.1: The NS2B cofactor sequence alignment and crystal structure of ZIKV protease. (A) The sequence alignment of NS2B from flaviviruses, the four serotypes of dengue and zika. Sequence alignment was done by Clustal omega. The predicted transmembrane region (4-24; 29-49; and 101-121) of NS2B is highlighted in red boxes. TOPCONS was used as the transmembrane prediction server. (B) The NS2B cofactor of ZIKV protease is highlighted in red (49-97 residues). It wraps around the protease facilitating proper folding of the protease and subsequently the protease active site. The catalytic triad is (S¹³⁵-H⁵¹-D⁷⁵) at the active site is show. The S135 that was mutated in ZIKV NS2B₄₈NS3_{S135} is designated by a blue star.

4.2.2 ZIKV NS3 is extended in solution, while the domains are compact

The data presented demonstrate that ZIKV NS3 is monomeric in solution. The estimated molecular weight derived from the Porod volume ($\sim 101 \text{ nm}^3$) was $63.1 \pm 6.3 \text{ kDa}$, which matches the monomer molecular weight. The protein was extended and flexible in solution. The flexibility of the protein was further quantitated using EOM 2.0. Three representative ensemble models for the probable conformation of the helicase were obtained. The large scale of R_g distribution indicated a wide variety of motion in the protein. Like in DENV NS2B₁₈NS3, the two domains of ZIKV NS3 are compact and rigid and the flexibility arises mainly due to the linker

4.2.3 4.2.3 ZIKV NS3 linker is conserved among strains

A comparison of the linker residue sequences of ZIKV NS3 from the French Polynesia, Brazil and Singapore lineages represents the identity of the RREEETPVECF-linker sequence (Fig. 4.2.2). In contrast, the NS3s of the four DENV serotypes show significant differences in the linker region (Fig. 4.2.2). A closer look to this region inside the structurally well understood DENV-4 NS3 (PDB ID: 2WHX) [200] highlights that the linker residue E169 and G172 have interaction with K88 and W83 of the protease, respectively, whereby linker residues E173 and D179 have hydrogen bonding with the helicase residues R202 and I182, respectively. R202 resides in the P-loop region and stabilizes the transition state during ATP hydrolysis. Therefore, any changes in the linker residues will directly affect ATP hydrolysis as a consequence of the altered interaction with the two enzyme domains. When the full length French Polynesia ZIKV NS3 was modelled based on the DENV-4 NS3 structure (PDB ID: 2WHX) [200], by individually docking the protease (PDB ID: 5LC0) and helicase (PDB ID: 5GJB) domains of ZIKV NS3 and manually building the linker region, the ZIKV NS3 linker residues T174 and P175 formed hydrogen bonds with W83 and Y68 of the protease, respectively, and the V176 interacted with the P-loop residue R203 of the helicase (Fig. 4.2.2). While the interaction with protease residue W83 and helicase residue R203 are preserved in both, the ZIKV NS3 model showed two interactions of the linker with the protease and one with the helicase. Based on the interaction of T174 and V176, these residues were mutated to investigate how the overall structural parameters of the protein changes. Comparing the sequence of the ZIKV and DENV linker region, it was seen that at 177 the uncharged residues of DENV linker was replaced by a charged E in ZIKV. Also, the highly conserved E178 was replaced by C in ZIKV and I/L/V by F in ZIKV.

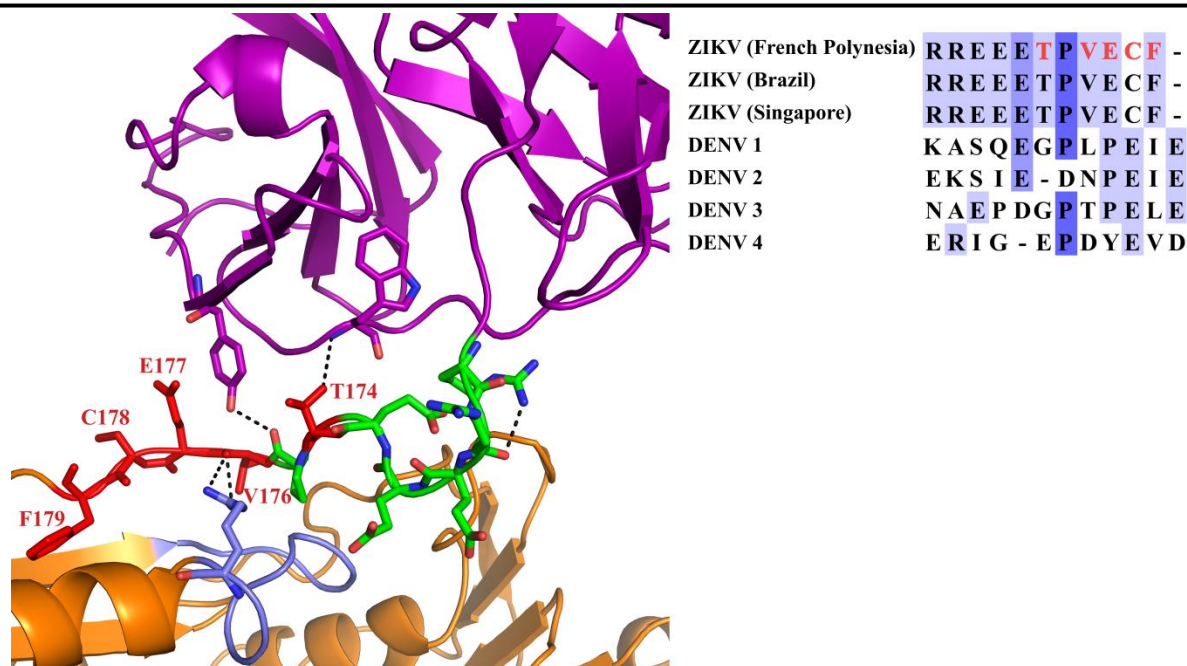


Figure 4.2.2: Interaction of ZIKV linker residues. Linker residues ($_{169}$ RREEETPVECF $_{179}$) in the ZIKV NS3 model generated based on the structure of DENV-4 NS3 (PDB ID: 2WHX) [200]. The protease (PDB ID: 5LC0) and helicase (PDB ID: 5GJB) domains of ZIKV where individually docked on to the full length DENV-4 NS3 structure and the linker residues were manually built, and energy minimized. The protease and helicase domains are highlighted in purple and orange, respectively. The linker residues are depicted in green and the P-loop is colored in blue. The ZIKV NS3 linker residues T174 and P175 interact with W83 and Y68 of the protease, respectively, and the linker residue V176 interacts with the P-loop residue R203 of the helicase. The residues mutated, namely T174, V176, E177, C178, F179 are shown in sticks and highlighted in red. Multiple sequence alignments of ZIKV- and DENV NS3 (Right inset) Multiple sequence alignment of linker of NS3 proteins of ZIKV isolated from French Polynesia (GenBank: [KJ776791](#)), Brazil (GenBank: [KU729217](#)), Singapore (GenBank: [KX827309](#)) and NS3 from DENV-1 (GenBank: [EU081230](#)), DENV-2 (GenBank: [GQ398264](#)), DENV-3 (GenBank: [AY662691](#)) and DENV-4 (GenBank: [AY776330](#)). The mutated residues are highlighted in red in the sequence.

4.2.4 ZIKV NS3 linker is important for the protein stability

The recombinant protein production for ZIKV NS2B $_{18}$ NS3 $_{T174A}$, ZIKV NS2B $_{18}$ NS3 $_{V176A}$, ZIKV NS2B $_{18}$ NS3 $_{E177A}$ and ZIKV NS2B $_{18}$ NS3 $_{C178A}$ was low compared to the WT ZIKV NS2B $_{18}$ NS3 (Fig. 3.2.5B and Fig. 3.2.17). The mutation of the T174 and V176 linker residues led to a less stable ZIKV NS3 (Fig. 3.2.18F). Though substantial amount of the mutant protein was obtained, the sample eluting from the SEC was not clean. Since two populations were observed on the SDS-PAGE gel, the samples were considered polydisperse and unsuitable for SAXS data collection. Mutating the T174 could lead to altered interaction of the mutated residue with protease and could lead to the destabilization of the protein. Similarly, the altered interaction of V176A led to higher protein degradation in the mutated recombinant protein. The purification of ZIKV NS2B $_{18}$ NS3 $_{E177A}$ and ZIKV NS2B $_{18}$ NS3 $_{C178A}$ was unsuccessful as mostly degraded protein eluted from the SEC (evident from SDS-PAGE gels in Fig. 3.2.18F). The residues T174, V176, E177 and C178 are important for the stability of the ZIKV NS3. Since the degradation was more prominent in ZIKV NS2B $_{18}$ NS3 $_{E177A}$ (Fig. 3.2.18E-F), this residue was thought to be a crucial residue in the linker for protein stability. On the other hand,

the production of the mutants ZIKV NS2B₁₈NS3_{F179A} and ZIKV NS2B₁₈NS3_{F179IA} was adequate when compared to the WT enzyme and the other linker mutants (Fig. 3.2.5 and Fig 3.2.17). A clean homogeneous protein was obtained from SEC on purification of the mutants, thus indicating unlike the mutation of aforementioned linker residues, mutation of F179 did not influence ZIKV NS3 stability.

4.2.5 ATPase inhibitory effects of the ZIKV NS2B₁₈NS3 and its helicase

The ATP hydrolysis data of the French Polynesia ZIKV NS2B₁₈NS3 and helicase revealed the effects of resveratrol and quercetin, with resveratrol significantly reducing ATPase activity of the entire recombinant NS2B₁₈NS3 and its helicase. Of interest is, that resveratrol is a potent inhibitor of both the NS2B₁₈NS3 as well as the helicase, and that this molecule has a stronger inhibitory effect to both proteins compared to quercetin. Very recently, two synthetic resveratrol analogs were identified to have anti-DENV activity, targeting viral RNA translation and viral replication, and indicating the therapeutic value of resveratrol derivatives [201]. Moreover, resveratrol has previously been studied for antiviral effects associated with inhibitions of viral replication, protein synthesis, gene expression, and nucleic acid synthesis for a number of viruses including influenza virus, hepatitis C virus, herpes simplex virus, human immunodeficiency virus and many more [202, 203]. The inhibitory role of quercetin in the RNA replication machinery was shown previously, where a reduction of the related DENV-2 RNA levels by 67% was observed [204]. The potency of resveratrol inhibiting ATPase activity of French Polynesia ZIKV NS3 described may shed light for lead optimization experiments in the future.

5. Results- Atomic and Enzymatic Insights into *EfAhpC*

The ability of *E. faecalis* to survive the intolerant ROS conditions of the host macrophage suggests the presence of robust antioxidant systems [151]. Thus, these systems may serve as therapeutic targets for the prevention of at least the early stages of infections. *E. faecalis* requires the presence of heme for its catalase to be functional. In the absence of a functional catalase, the AhpR system is crucial for maintaining low levels of intracellular H₂O₂.

Like other 2-Cys peroxiredoxins, *EfAhpC* is proposed to undergo redox dependent oligomerization. The oligomeric state and its redox modulation are studied by dynamic light scattering and electron microscopy of the protein. Since no high-resolution structure exists for this protein, attempts were made to crystallize the protein to gain atomic insight into the structure. The full-length protein did not crystallize, and a truncated protein was generated for crystallization purpose. Insight was obtained through amino acid substitutions of residues in the N- and C-terminus as well as the unique cysteine residues C13 and C66 of *EfAhpC*, which provide insights into their structural and enzymatic roles.

5.1 Characterization of *EfAhpC*

5.1.1 Purification of *EfAhpC*

For purifying the oxidized form of *EfAhpC*, care was taken to avoid contamination with any sort of reducing agent. The first step of purification was immobilized metal ion affinity chromatography using Ni²⁺-NTA agarose beads (Qiagen). The recombinant protein produced with the N-terminal His₆-tag has affinity for Ni²⁺ and binds to the Ni-NTA agarose beads. The protein was eluted from a column containing the beads by batch elution with varying imidazole concentration (0 - 500 mM). The fractions eluted were loaded on a 17% SDS-PAGE gel. Clear bands were identified on the gel in the fractions eluted with 100 – 500 mM of imidazole (Fig. 5.1.1A). Fractions eluted with 125-250 mM of imidazole were pooled together, concentrated and applied to SEC to obtain highly pure protein in oxidised form for further studies (Fig. 5.1.1B). The eluted protein was run on a 17% SDS-PAGE gel to check the purity and homogeneity of the protein. A single sharp band on the SDS-PAGE gel indicated that the protein was pure and homogenous (Fig. 5.1.1B; inset).

For the reduced *EfAhpC*, the same fractions from the Ni²⁺-NTA step were pooled and concentrated and applied onto the SEC. The protein was eluted with the same buffer described earlier but with 1 mM of DTT. As shown in figure 5.1.1, the reduced protein eluted at the same volume of (around 12 ml) as the oxidised one.

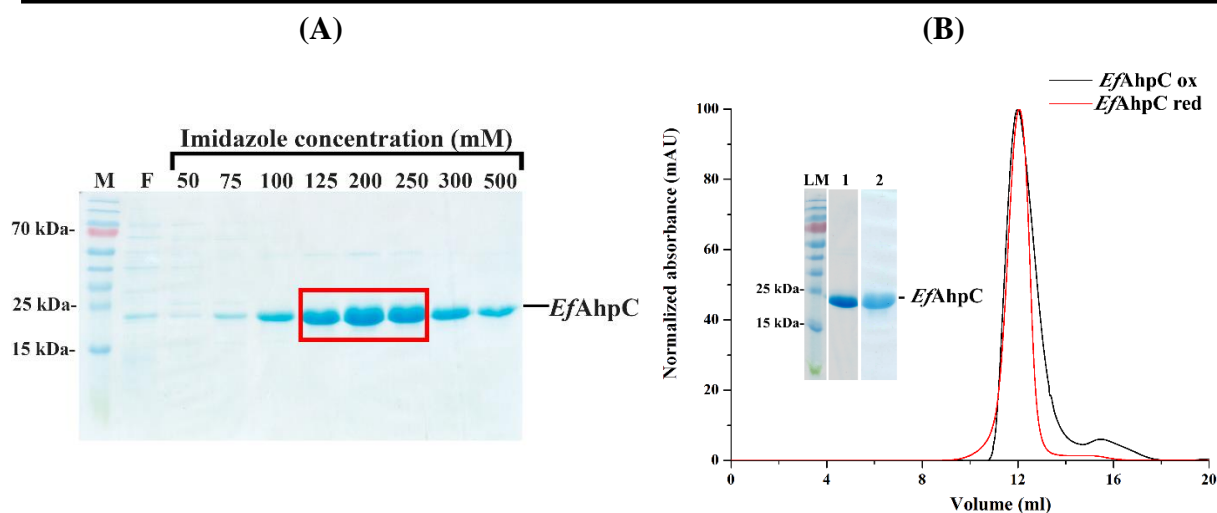


Figure 5.1.1: Purification of *EfAhpC*. (A) The 17% SDS-PAGE gel shows the imidazole concentrations at which the tagged protein eluted from the affinity chromatography column (M= Protein marker; F= flow-through). The red box indicates the fractions pooled for the next step of purification. (B) On application to SEC, a single peak was observed in the chromatogram for both oxidized (—; black) and reduced (—; red) protein. *Inset*: 17% SDS-PAGE gel of the eluted SEC fractions. LM= protein marker; 1= *EfAhpC* ox and 2= *EfAhpC* red.

5.1.2 Enzymatic characterization of *EfAhpC*

The two subunits of the *E. faecalis* AhpR system, the NADH-dependent AhpF and AhpC, orchestrate the detoxification of H_2O_2 by catalysing the reduction of H_2O_2 to H_2O . This detoxification step is brought about by transferring reducing equivalents from NADH to H_2O_2 . An NADH-dependent assay was performed to measure the catalytic activity, where the drop in NADH absorbance is monitored spectrophotometrically at 340 nm. Due to the ability of NADH to differentially absorb at 340 nm, the transfer of electrons between NADH via FAD and the C345/C348 catalytic centres inside the C-terminal domain of AhpF results in the drop of NADH-absorbance at 340 nm.

A sharp drop in absorbance of NADH at 340 nm was observed at the stopped flow instrument (Fig. 5.1.2A) which suggested that the purified recombinant *EfAhpC* is catalytically active and is capable of H_2O_2 reduction in the presence of *EfAhpF*. This is further reiterated by the FOX assay, which directly measures the decrease in H_2O_2 concentration as it is decomposed by AhpC as shown in figure 5.1.2B.

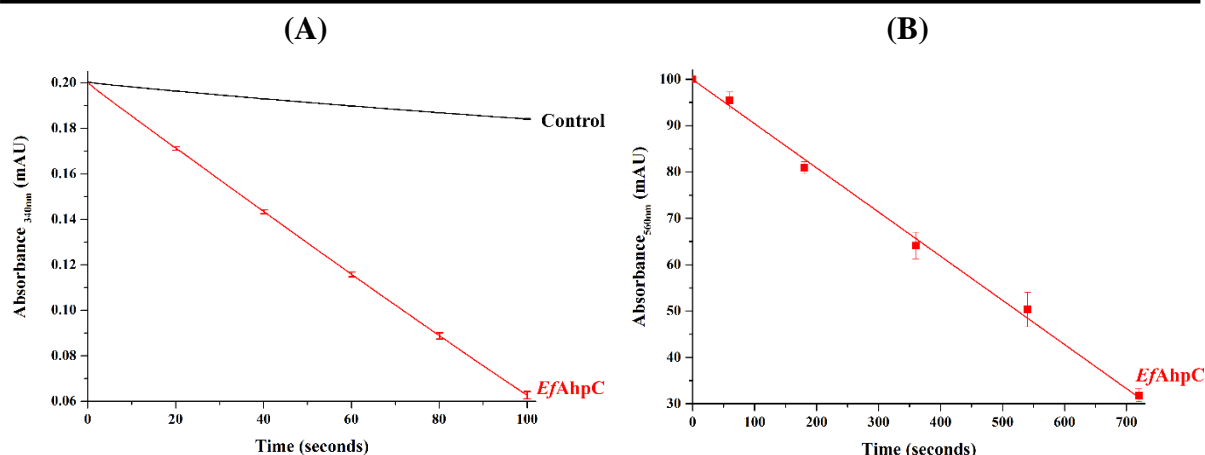


Figure 5.1.2: *EfAhpC* is catalytically active. (A) NADH-dependent peroxidative activity of *EfAhpC* (red). The control (—; black) represents the reaction which does not have AhpC, just WT *EfAhpF*. The sharp decrease in NADH absorbance indicates that the *EfAhpC*-AhpF complex catalyses the detoxification of H₂O₂. (B) The FOX assay shows rapid decomposition of H₂O₂ by *EfAhpC*.

5.1.3 Analysis of the oligomeric state of *EfAhpC* by DLS

To analyze the redox modulated oligomeric state of the *EfAhpC*, DLS was performed with both oxidized and reduced *EfAhpC*. A hydrodynamic diameter of 12.11 ± 2.1 d.nm (Fig. 5.1.3A) and estimated molecular weight of 228 ± 40.8 kDa was recorded for the oxidized *EfAhpC*. For the reduced protein, the hydrodynamic diameter and estimated molecular weight recorded was 12.4 ± 1.45 d.nm (Fig. 5.1.3A) and 242 ± 28.3 kDa, respectively. The correlation function (Fig. 5.1.3B) of both the oxidized and reduced protein indicate, that they are of the same size, suggesting that redox modulation does not affect the protein size. Since the monomeric molecular weight of *EfAhpC* from the protein sequence (using ExPasy ProtParam tool) was determined to be 20 kDa, the molecular weight estimated from DLS data indicate that both the oxidized and reduced protein form higher weight oligomers.

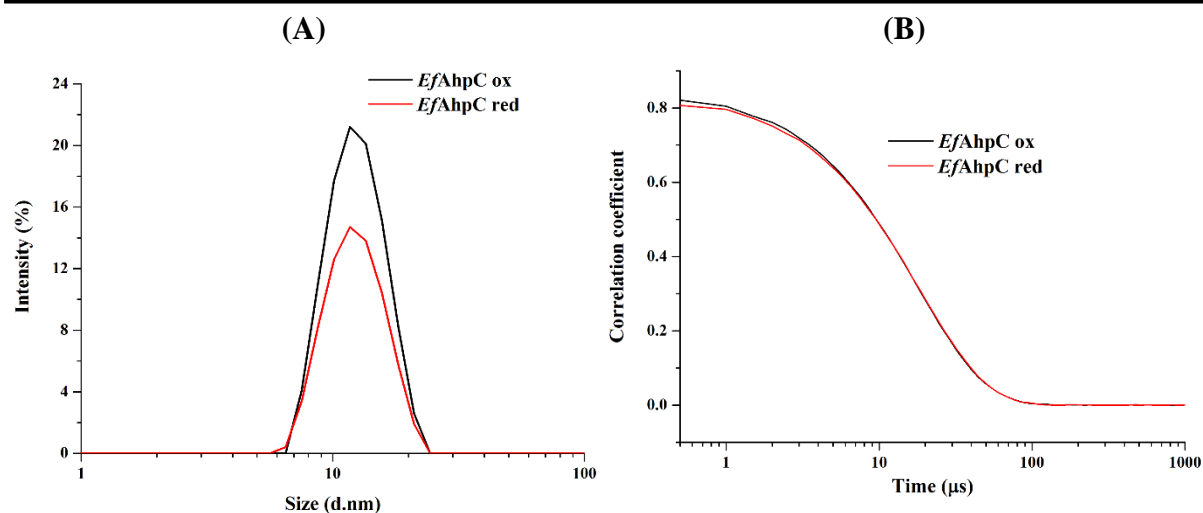


Figure 5.1.3: *EfAhpC* forms high molecular weight oligomers. (A) DLS experiments were used to study the hydrodynamic diameter of oxidized (—; black) and reduced (—; red) *EfAhpC*. The oxidized form revealed a diameter of 12.11 nm and the reduced protein a diameter of 12.4 nm, indicating the protein forms higher oligomers independent to the redox state. (B) The correlation coefficient for oxidized (—; black) and reduced (—; red) *EfAhpC*.

5.1.4 Analysis of high molecular weight oligomers by EM

As the DLS data indicated that the *EfAhpC* in both oxidized and reduced state form high molecular weight oligomers, negative stain was utilized to study the oligomer under electron microscopy. The EM data analysis of both oxidized and reduced protein was done by Dr. Phat Van Dip and Dr. Shashi Bhushan. The micrograph of both oxidized and reduced protein depicted ring-like particles (Fig. 5.1.4A-B). The class averages using EMAN2 and RELION software for the oxidized and reduced *EfAhpC* form revealed 5-mer ring-shaped particles that have an outer diameter of approximately 130 Å and an inner diameter of 40 Å (Fig. 5.1.4C-D). The EM data confirm that *EfAhpC* forms a stable decamer regardless of the redox state of the protein.

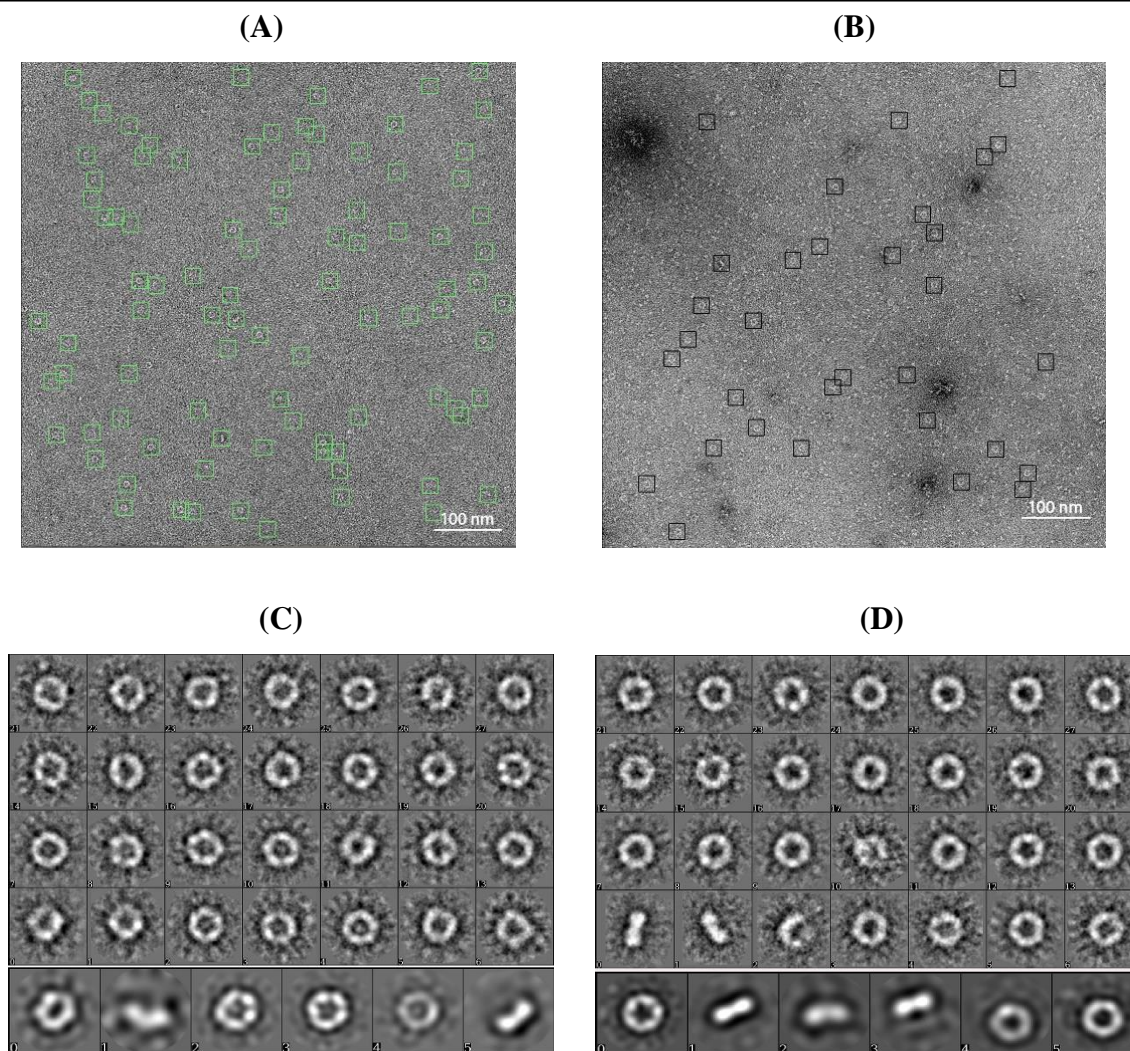


Figure 5.1.4: Negative stained images of *EfAhpC* indicate that the protein forms a redox-independent deacmer. (A) EM images of oxidized *EfAhpC*. Particles for class averaging are highlighted in boxes (*green*). Ring and rod-shaped particles were found and selected using EMAN2 software. (B) EM image of reduced *EfAhpC*. Particles for class averaging are highlighted in boxes (*black*). (C) 2D class average images of isolated *EfAhpC* particle in oxidized form obtained using RELION software. (D) 2D class average image of isolated *EfAhpC* particle in oxidized form obtained using RELION software.

5.1.5 Crystallization of *EfAhpC*

To gain structural insights into the redox independent deacmer, attempts were made to crystallize the full-length protein using commercially available crystallization screens. However, only 2D plate like crystals (Fig. 3.3.5) were observed in conditions in which the protein did crystallize, and these crystals did not produce considerable diffraction pattern when checked in-house at the RIGAKU instrument. It was suspected that the flexible C-terminus of *EfAhpC* could contribute to the poor tendency of the protein to crystallize and diffract. Based on the success of structure determination of the truncated *E. coli* AhpC (*EcAhpC*₁₋₁₇₂) (Dip et al, 2014), a truncated *EfAhpC* was generated to gain atomic insights into the stable deacmer.

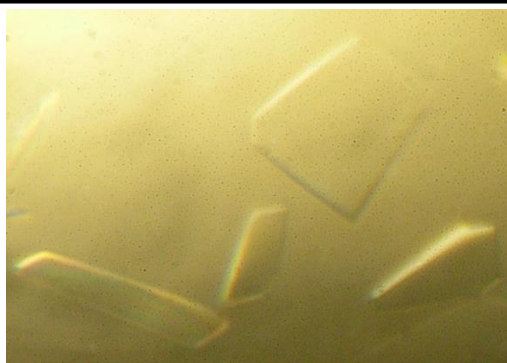


Figure 5.1.5: *EfAhpC* crystals. 2D plate like crystals obtained for *EfAhpC*, which did not diffract.

5.2 Characterization of *EfAhpC*₁₋₁₇₂ and structural insights into *EfAhpC*₁₋₁₇₂

5.2.1 Cloning, expression and purification of *EfAhpC*₁₋₁₇₂

A single band (~516 bp) was observed in an 1.5% agarose gel (Fig. 5.2.1A) when the amplified product of the gene encoding *EfAhpC*₁₋₁₇₂ was run on the gel. The digested insert was ligated into the modified pET 9d vector. Once the sequence of the plasmid was confirmed by sequencing, the plasmid was transformed into *E. coli* BL21 DE3 cells for expression. Over-production of protein was induced by IPTG. Recombinant *EfAhpC*₁₋₁₇₂ production from the expression strain was identified by a 17% SDS-PAGE gel (Fig. 5.2.1B).

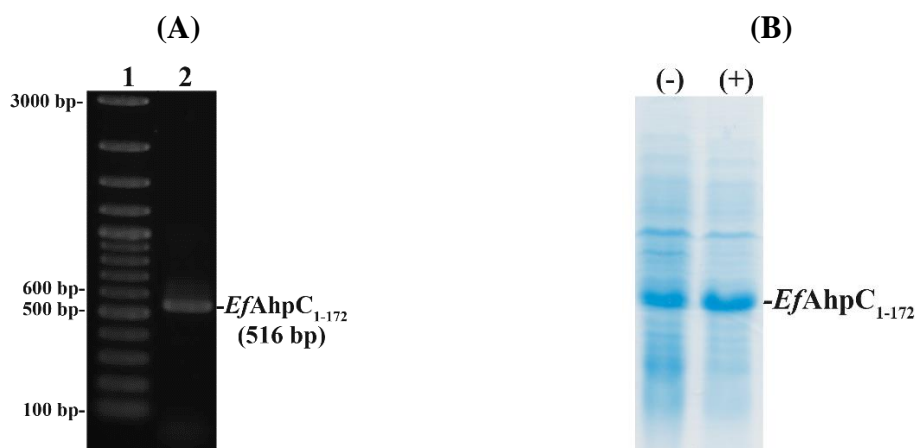


Figure 5.2.1: Amplification of gene encoding *EfAhpC*₁₋₁₇₂ and recombinant *EfAhpC*₁₋₁₇₂ production. (A) Amplified product (lane 2) and DNA ladder in lane 1. DNA sample of 5 μ l was loaded on an agarose gel (1.5%). (B) SDS-PAGE (17%) showing the expression of recombinant protein. Protein production was induced with IPTG. (+) and (-) indicates induced and non-induced cells, respectively. A band of the recombinant protein was observed in both lanes due to the leaky expression of pET9d.

The oxidized *EfAhpC*₁₋₁₇₂ was purified in the first step by affinity chromatography. Fractions eluted from the column by an imidazole gradient were loaded on a 17% SDS-PAGE gel. The recombinant protein was identified in an imidazole gradient of 75-500 mM (Fig. 5.2.2A). The 125-250 mM imidazole fractions were pooled together, concentrated and applied

to a size exclusion column. A single peak (at approximately 12 ml) was observed in the chromatogram. When fractions from 15 % of the peak absorbance maxima were applied on a gel, a clean band was observed between 25 -15 kDa of the protein marker (Fig. 5.2.2B, inset).

For obtaining the reduced *EfAhpC*₁₋₁₇₂, the protein was purified in the same way except for the SEC buffer, which contained DTT (final concentration of 1 mM). The protein eluted from the affinity column in the same imidazole gradient as the oxidized protein, and from the SEC at the same elution volume as the oxidized protein (Fig. 5.2.2B).

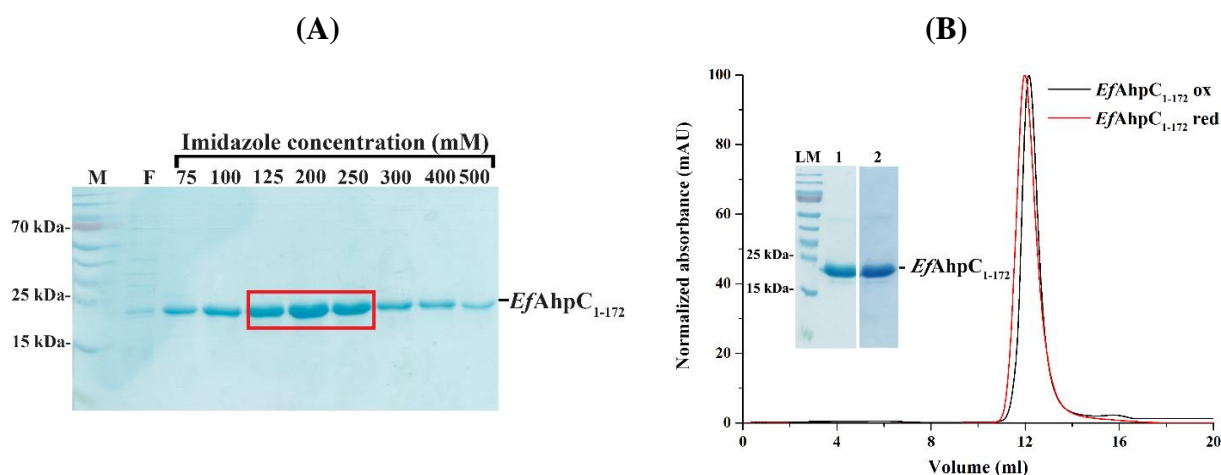


Figure 5.2.2: Purification of *EfAhpC*₁₋₁₇₂. (A) The SDS-PAGE gel shows the imidazole concentrations at which the tagged protein eluted from the affinity chromatography column (M= Protein marker; F= flow-through). The red box indicates the fractions pooled for the next step of purification. (B) On application to SEC, a single peak was observed in the chromatogram for both oxidized (—; black) and reduced (—; red) protein. *Inset*: SDS-PAGE gel of the eluted SEC fractions. LM= protein marker; 1= *EfAhpC*₁₋₁₇₂ oxidized and 2= *EfAhpC*₁₋₁₇₂ reduced.

5.2.2 Enzymatic characterization of *EfAhpC*₁₋₁₇₂

The rate of decline in absorbance of NADH at 340 nm was slower for *EfAhpC*₁₋₁₇₂ suggesting that the truncated protein was catalytically less active in comparison to the full-length protein (Fig. 5.2.3A). The FOX assay also showed that the peroxide decomposition is almost nil when compared to the full-length protein as shown in figure 5.2.3B. This indicates that the C-terminus is essential for the peroxidative activity of the protein.

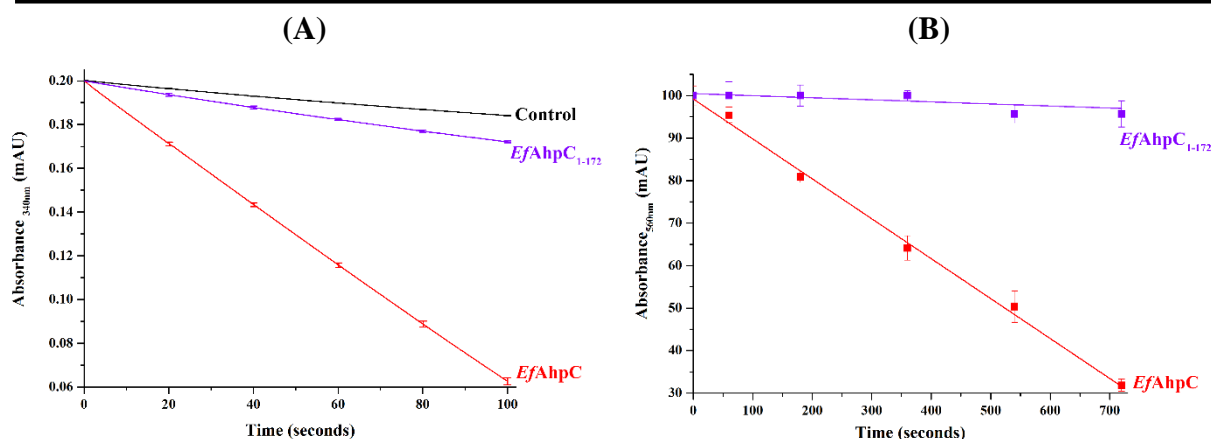


Figure 5.2.3: The C-terminal tail is crucial for peroxidative activity of *EfAhpC*. (A) NADH-dependent peroxidative activity of *EfAhpC*₁₋₁₇₂ (—; violet) compared to that of *EfAhpC* (—; red). The control (—; black) represents the reaction which does not have AhpC, just WT *EfAhpF*. (B) The FOX assay for *EfAhpC*₁₋₁₇₂ (—; violet) show almost no decomposition of H₂O₂ in comparison to *EfAhpC* (—; red).

5.2.3 Analysis of the oligomeric state of the truncated *EfAhpC*

DLS data of the oxidized and reduced protein at 3 mg/ml revealed a hydrodynamic diameter of 11.9 ± 0.66 d.nm and 12.1 ± 1.40 d.nm (Fig. 5.2.4A). The correlation function for the oxidized and reduced data are depicted in figure 5.2.4B. These values are similar to the values described for oxidized and reduced full-length *EfAhpC*. Even though the absence of the C-terminus truncation reduces the catalytic efficiency of the recombinant protein, it does not affect the stable decamer formation. This gives the impression that other factors, apart from the C-terminus, contribute to the stability of the *EfAhpC* decamer.

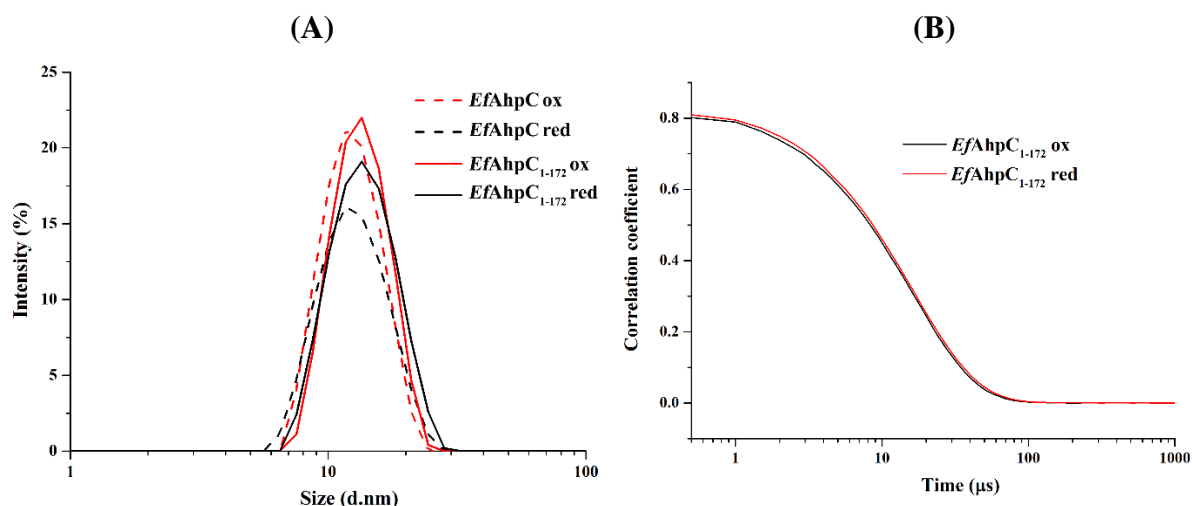


Figure 5.2.4: *EfAhpC*₁₋₁₇₂ like the full-length protein forms a decamer. (A) DLS experiments were used to study the hydrodynamic diameter of oxidized (—; black) and reduced (—; red) *EfAhpC*₁₋₁₇₂. Comparison of DLS profile of *EfAhpC*₁₋₁₇₂ with *EfAhpC* (---; oxidized in red and reduced in black) protein reveals that despite a truncated C-terminus, the protein is capable of forming a decamer. (B) The correlation coefficient for oxidized (—; black) and reduced (—; red) *EfAhpC*₁₋₁₇₂.

5.2.4 Crystallisation and data collection of *EfAhpC*₁₋₁₇₂ crystals

To gain insight into the *EfAhpC*₁₋₁₇₂ structure, hanging drops of oxidised and reduced *EfAhpC*₁₋₁₇₂ at 5 mg/ml and 8 mg/ml were set-up. Initial micro-crystals were observed in Molecular Dimensions Morpheus screen I (#32) and Emerald Biosystems Wizard Screen I (#25) for the reduced protein alone. For the oxidized form, no crystals were observed. Initially, crystals in the Morpheus screen (0.09 M NPS, 0.1 M Buffer 2 (Sodium HEPES; MOPS (acid)] pH 7.5, 37.5% MPD_P1K_P3350) were optimised. The initial crystals diffracted to 8 Å (Fig. 5.2.5) in-house.

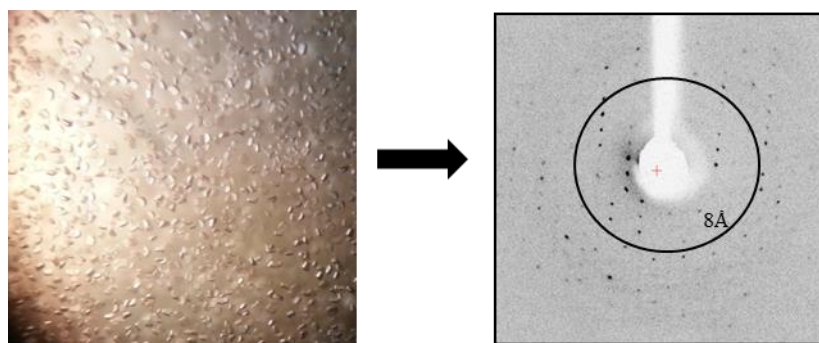


Figure 5.2.5: Picture of initial crystals of *EfAhpC*₁₋₁₇₂ and diffraction pattern. Initial *EfAhpC*₁₋₁₇₂ crystals, which were crystallized in 0.09 M NPS, 0.1 M Buffer 2 (Sodium HEPES; MOPS (acid)) pH 7.5, 37.5% MPD_P1K_P3350. (right) The best diffraction pattern of *EfAhpC*₁₋₁₇₂ crystals before optimization. The circle shows resolution shell for 8 Å.

To improve the quality of the crystals and diffraction resolution, a series of optimizations were done by varying crystallization parameters. A systematic grid screen was performed by varying the concentrations of precipitant, salt, and buffer. The best crystals were however obtained in the original condition.

Further optimizations for the crystals were done by varying protein concentration (3 mg/ml, 5 mg/ml, 8 mg/ml) (Fig. 5.2.6) and varying drop size (*EfAhpC*₁₋₁₇₂ in buffer (50 mM Tris/HCl, pH 7.5, 500 mM NaCl and 1 mM DTT) and precipitant solution in the ratio 1:1, 2:1, 2:2)) (Fig. 5.2.7). The crystals were tested in-house to check for the quality of crystals. The optimal protein concentration for obtaining well-formed crystals was found to be 5 mg/ml and the optimal drop size was the one in which *EfAhpC*₁₋₁₇₂ and precipitant solution were in a ratio of 2:1.

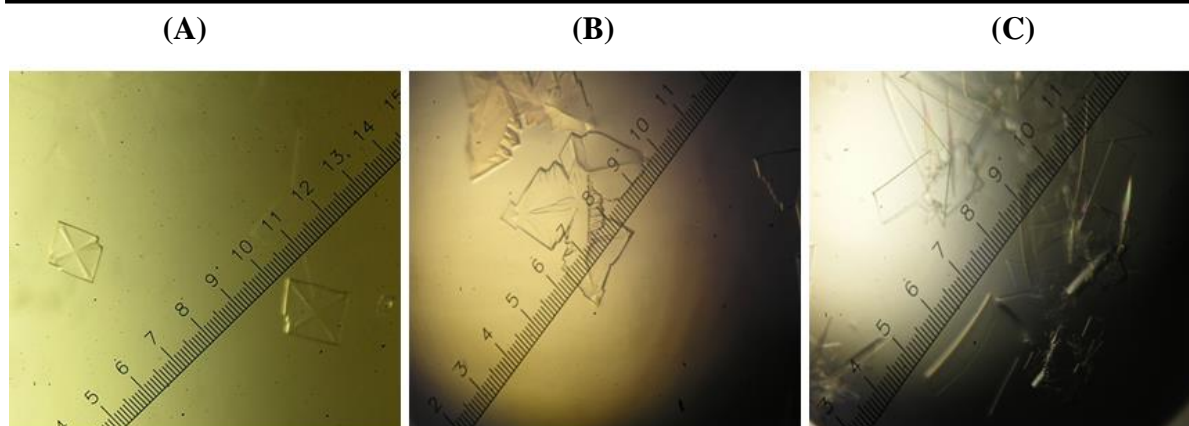


Figure 5.2.6: Optimization of *EfAhpC*₁₋₁₇₂ by varying protein concentration. Crystals of *EfAhpC*₁₋₁₇₂ with varying protein concentrations of 3 mg/ml (A), 5 mg/ml (B), and 8 mg/ml (C). Better diffracting crystals were obtained with 5 mg/ml.

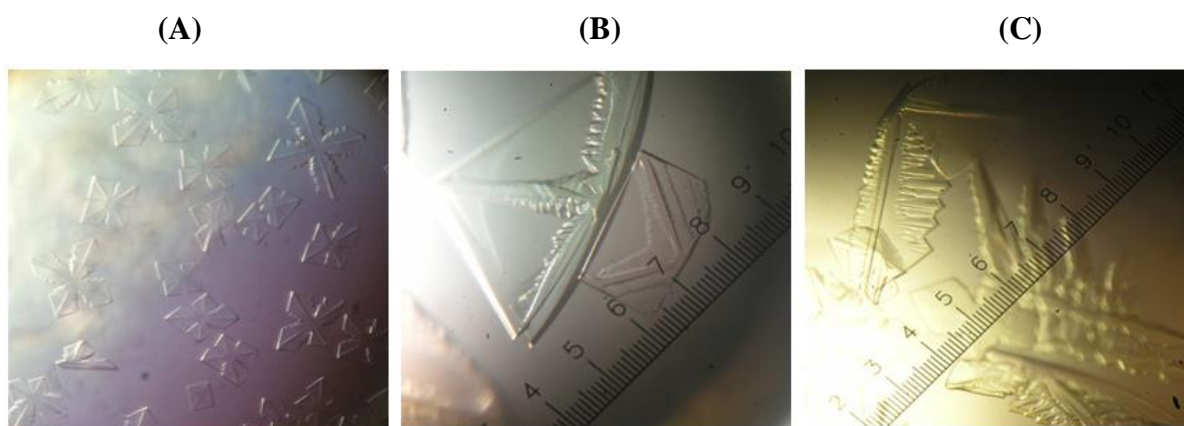


Figure 5.2.7: Optimization of *EfAhpC*₁₋₁₇₂ by varying protein concentration. Crystals of *EfAhpC*₁₋₁₇₂ with varying ratio of the purified recombinant protein in buffer and Morpheus 1 #32 solution. The ratios of protein: precipitant is 1:1 (A), 2:1 (B), and 2:2 (C). Better diffracting crystals were obtained in the protein:precipitant ratio of 2:1(v/v).

The nucleation process in Morpheus I (#32) was initiated within a day leading to formation of crystals. To slow down the nucleation process, the reservoir solution in the well was overlaid with paraffin and silicon oil (Fig. 5.2.8). However, the diffraction pattern did not improve when these crystals were tested in-house.

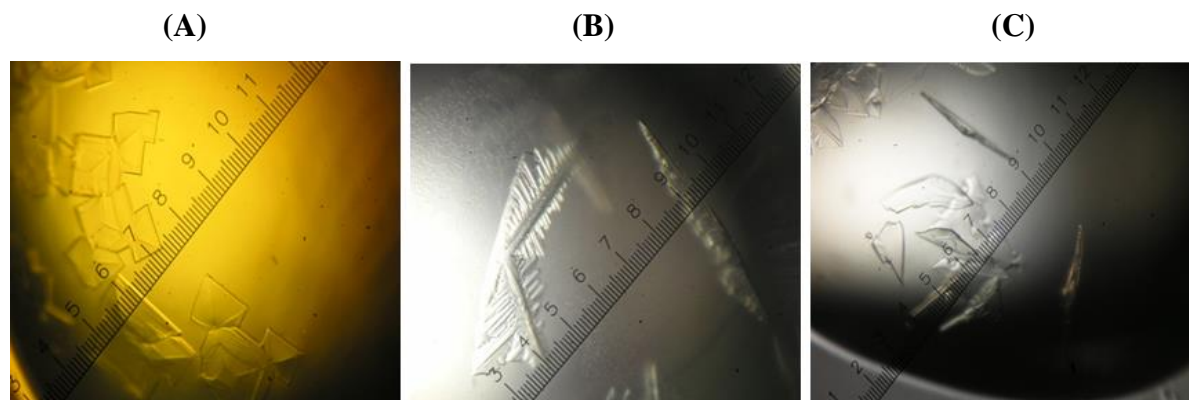


Figure 5.2.8: Optimization of *EfAhpC*₁₋₁₇₂ by adding paraffin and/ or silicone oil in reservoir. Crystals of *EfAhpC*₁₋₁₇₂ after covering the reservoir solution by paraffin (A), by silicon oil (B), and combination of paraffin and silicon oil in the ratio 4:1 (C). Crystal that were obtained without overlay of silicone or/and paraffin oil diffracted better than the ones obtained with overlay of oil.

After optimization, the final optimal condition for crystals (Fig. 5.2.9) was found to be in the original Morpheus I (#32) solution with the protein concentration of 5 mg/ml, with a protein: precipitant ratio of 2:1 (v/v) without any paraffin or silicone overlay in the reservoir. The optimized crystals were harvested and tested in in-house diffracted to 6 Å.

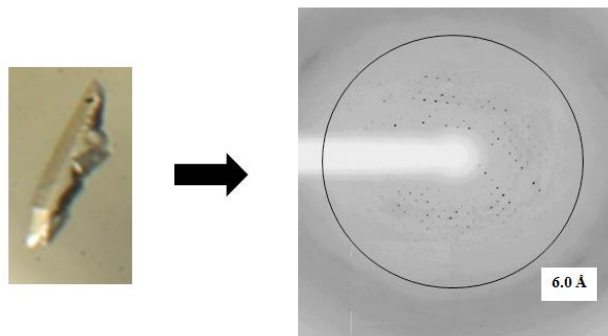


Figure 5.2.9: Diffraction pattern of optimized *EfAhpC*₁₋₁₇₂. Optimized crystals of *EfAhpC*₁₋₁₇₂ in Morpheus I (#32) set up with 5 mg/ml protein concentration and the diffraction pattern of the crystal. Circles show resolution shell for 6 Å.

The final cryo-protectant solution comprised of 0.09 M NPS, 0.1 M Buffer 2 (Sodium HEPES; MOPS (acid)) pH 7.5, 37.5% MPD_P1K_P3350, and 25% (v/v) glycerol. A single-wavelength data set with the optimised crystal was collected at the in-house machine, where temperature was maintained at -173.15 °C (100 K) (Table 5.1).

Data collection parameters	
Temperature (K)	100
Wavelength (Å)	1.00
Resolution range (Å)	30-5.9
Crystal System	Orthorhombic
Space group	C222 ₁
Unit cell parameters (Å)	$a = 94.80, b = 204.87, c = 103.47$
R _{sym}	10 (26)
Completeness (%)	98.6 (99.7)
Average I/σ(I)	15.3 (2.0)
No. of molecules in asymmetric unit	5 molecules
Solvent content (%)	53

Table 5.2.: Summary of in-house data collected from an *EfAhpC*₁₋₁₇₂ crystal

Since the resolution of the crystals optimized in the Morpheus I (#32) did not improve beyond 6 Å, crystals obtained in the Wizard I (#25) condition were optimized. A systematic grid screen was done as described previously and optimal crystals were obtained in the original Wizard screen I #25 (0.2 M sodium chloride, 0.1 M Tris, pH 8.5 and 30% (v/v) PEG 400). An additional screen using Hampton Research additive screen was done to improve the crystal quality. The crystals obtained in Wizard screen I #25 (Fig. 5.2.10) with the additive 40% (v/v) acetonitrile diffracted in-house to a resolution of 3.5 Å. The crystals were quickly soaked into

cryoprotectant solution containing 25% (w/v) glycerol in crystallization liquid and flash-cooled in liquid nitrogen at 100 K. With the assistance of Dr. Asha M. Balakrishna from our lab, single wavelength datasets for *EfAhpC*₁₋₁₇₂ were collected at 140 K on beamline 13B1 at the National Synchrotron Radiation Research Center (NSRRC, Hsinchu, Taiwan). Data were collected as a series of 0.5° oscillation images with 7 sec exposure time and a detector distance of 400 mm. All diffraction data were indexed, integrated and scaled using the HKL2000 suite of programs. The summary of data collection is shown table 5.2.

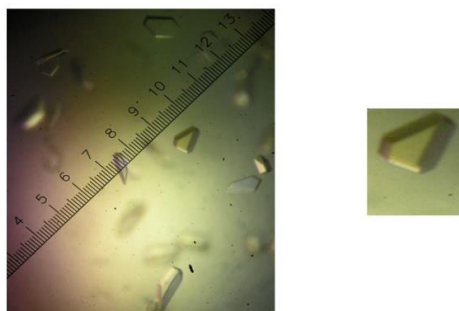


Figure 5.2.10: Final optimized *EfAhpC*₁₋₁₇₂ crystal. The figure shows optimized crystals of *EfAhpC*₁₋₁₇₂ in Wizard I (#25) with a protein concentration of 5 mg/ml that was tested at the synchrotron. This crystal diffracted to 2.9 Å at the beamline 13B1 at the National Synchrotron Radiation Research Center (NSRRC, Hsinchu, Taiwan).

Data collection parameters	
Wavelength (Å)	1.00
Space group	C2
Unit cell parameters (Å)	
<i>a</i> (Å)	103.20
<i>b</i> (Å)	124.67
<i>c</i> (Å)	97.80
$\alpha = \beta$ (°)	90
γ (°)	113.6
Resolution range (Å)	30-2.85
Solvent content (%)	47.80
Total number of reflections	82,039
Number of unique reflections	25,934
<i>I</i> / σ ^a	8.6 (1.6)
Completeness (%)	99.1(95.4)
R merge ^b (%)	9.8 (39.5)
Redundancy	3.2 (2.2)

Table 5.2: Summary of data collected at beamline 13B1 from an *EfAhpC*₁₋₁₇₂ crystal.

(^a = Values in parentheses refer to the corresponding values of the highest resolution shell (2.95 – 2.85 Å). ^b Rmerge = $\sum \sum_i |I_h - \bar{I}_h| / \sum \sum_i I_h$, where \bar{I}_h is the mean intensity for reflection *h*.)

Based on the CC1/2 statistics, the data were cut off at 2.85 Å [205] (correlation coefficient between two random halves of the data set where CC1/2 > 10%). PHENIX [206] was used to calculate CC1/2 (76.5% for the highest resolution shell and 99.2% for the entire data set), supporting our high-resolution cutoff determination. The presence of five molecules in the

asymmetric unit with a corresponding solvent of 58.7% was established by Matthews coefficient [207]. The scaling statistics shown in table 5.2 indicate that the diffraction data were of reasonable quality. No disulphide bond was observed among the five *EfAhpC*₁₋₁₇₂ molecules (A-E) which could be a result of the highly disordered C-terminus and possibly due to radiation damage and due the presence of DTT in the purification buffer.

5.2.5 Crystal structure of *EfAhpC*₁₋₁₇₂

The *EfAhpC*₁₋₁₇₂ structure was solved at a resolution of 2.8 Å. The crystal structure of *EfAhpC*₁₋₁₇₂ contains five molecules (chains A-E) in the asymmetric unit, forming a half ring shape conformation. The decameric ring is generated by its crystallographic two-fold symmetry operation (A'-E') (Fig. 5.2.11A). Each *EfAhpC*₁₋₁₇₂ monomer consists of seven stranded β -sheets located in the center of the molecule, flanked at one side by four and on the other side by two α -helices (Fig. 5.2.11B). Each monomer contains two conserved cysteine residues, the peroxidative cysteine (C_P47) and the resolving cysteine (C_R166). Helix α 2 contains C_P47 which is involved in the local unfolding during catalysis. C_P47 is surrounded by three additional conserved residues P40, S44 and R120, which activate and stabilize the C_P-thiolate for peroxidation (Fig. 5.2.11C). P40, S44 and C_P47 are found in a contiguous segment with a universally conserved PXXXT/SXVCP sequence motif, whereas R120 is distantly located on β 6, but in close proximity to C_P47 in the so-called FF confirmation. In *EfAhpC* a serine is present in the PXXXT/SXVCP sequence motif instead of a threonine. Apart from the conserved C_P47 and C_R166, *E. faecalis* has two extra cysteines, C13 on β 1 and C66 on β 4, the roles of which are not known yet (Fig. 5.2.11C). The *EfAhpC*₁₋₁₇₂ monomer contains 172 amino acids and all main chain residues fit well into the electron density, except for the C-terminal residues beyond C_R166, which are disordered. The side chain density for all the amino acids till 162 are very well resolved.

The *EfAhpC*₁₋₁₇₂ monomer has two interfaces, a dimer- and oligomer interface. The dimer interface is mainly stabilized by salt bridge- and hydrogen bond interactions, which are formed between β 7 and β 7' of each monomer, which altogether forms a combined 14 stranded β -sheet structure (Fig. 5.2.11C). The dimer interface is tightly connected by salt bridges between E147 and K151 and hydrogen bonds between the residues K132, Y134, E135, I136 and N137, of β 7 and β 7', residues D138, G142 and N144 present in the loop between β 7 and α 5 and residues E147, R150 and K151 present on α 5 of each monomer. The oligomeric assembly is mainly stabilized by hydrophobic interactions and hydrogen bond interactions (Fig. 5.2.11D).

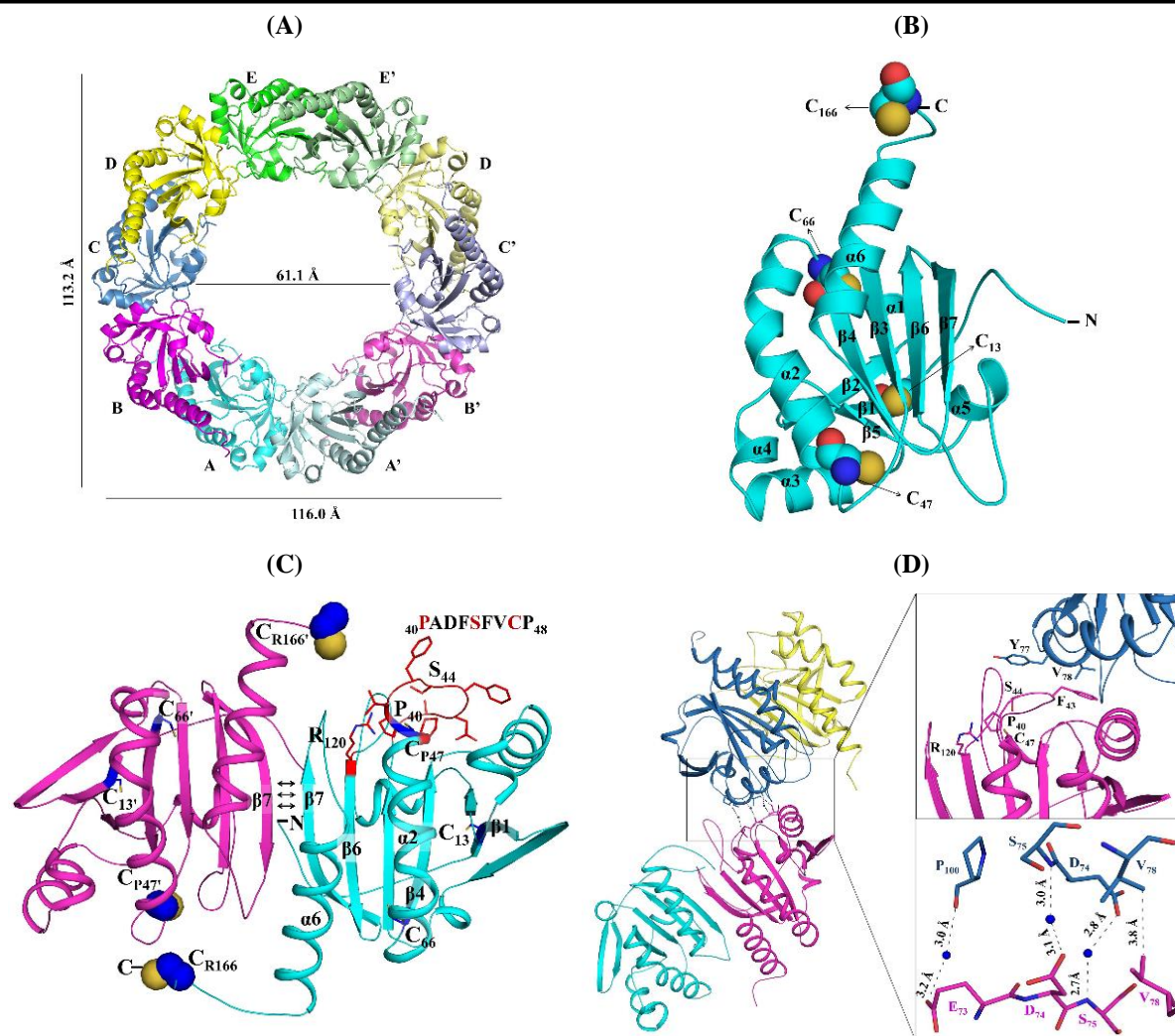


Figure 5.2.11. The structural features of *EfAhpC*₁₋₁₇₂. (A) The crystal structure of reduced *EfAhpC*₁₋₁₇₂ in decameric form with an outer diameter of 116.0 Å x 113.2 Å and inner diameter of 61.1 Å. The asymmetric unit is comprised of five chains (A-E), each colored differently and its symmetry mates are illustrated in lighter shade (A'-E'). (B) Cartoon representation of one monomer of *EfAhpC*₁₋₁₇₂ shown in cyan, the α -helices and β -strands are labelled. The peroxidative cysteine C₄₇ and the resolving cysteine C₁₆₆ as well as the two extra cysteines C₁₃ and C₆₆ are shown as sphere. (C) The dimer interface of *EfAhpC*₁₋₁₇₂ is shown between chain D (cyan) and E (magenta). The interactions (---; black) between the anti-parallel β -strand of the two molecules at the dimer interface generates a continuous 14-stranded β -sheet in the dimer. The dimer interface is tightly connected by salt bridges and hydrogen bonds. The conserved PXXXT/SXVCP motif is shown in red represented in sticks. (D) The oligomeric interface is shown with all the chains of *EfAhpC* coloured differently; chain A; cyan, -B; magenta, -C; deep blue, -D; yellow (For clarity chain E is not shown). The dimer-dimer interface is formed between chain B and C (magenta and deep blue). Top inset: the residues in the catalytic triad, C₄₇, S₄₄, R₁₂₀ and the conserved P₄₀ at the dimer-dimer interface with the conserved Y₇₇' and V₇₈' of the adjacent chain are shown. Bottom inset: the microenvironment in the oligomeric interface of V₇₈ and V₇₈' is presented.

The chains A and B of the determined crystallographic structure showed also that *EfAhpC*₁₋₁₇₂ in the reduced state mimics the FF conformation, whereby its catalytic residue C₄₇ is in a helical conformation, and in close proximity with residues of the catalytic triad. In addition, the S γ -atom of C₄₇ is hydrogen bonded to the guanidine group of R₁₂₀ (Fig. 5.2.12A). Interestingly, chains C and D of *EfAhpC*₁₋₁₇₂ reveal another conformation of C₄₇ in the active site, where the α 2-helix is partially unfolded and the C₄₇ is repositioned in the loop,

exposing it to the resolving residue C_{R166}', and paving its way into a possible disulphide bond formation of the oxidized structure (Fig. 5.2.12B). Although no intermolecular disulphide bond is observed between the cysteine residues in the presented *EfAhpC*₁₋₁₇₂ structure, this conformation observed in the C and D chains seems to be a transition state in-between the FF and LU conformation. In this transition state, the hydrogen bond between the S_γ-atom of C_{P47} and the guanidine group of R120 is disrupted (Fig. 5.2.12B). The r.m.s.d between the two conformations is 0.624 Å for 161 C_α atoms in each chain.

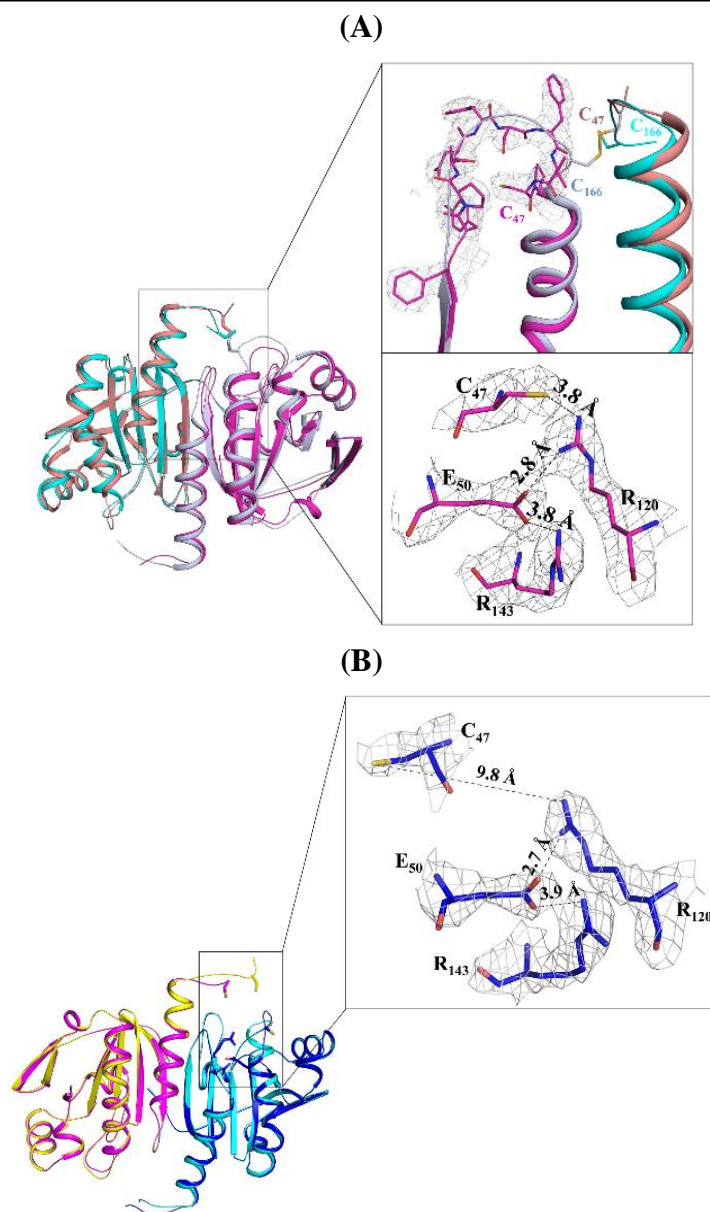


Figure 5.2.12. Structural intricacies of EfAhpC (A) The superposition of the EfAhpC dimer (chain A and B; cyan and magenta) and EcAhpC structure (PDB: 4O5Q; deep salmon & blue white). (*Top inset*) The final 2FoFc map at the active site $_{38}\text{FYPADFTFVCP}_{48}$, of EfAhpC is shown; the map is contoured at 1σ . In EcAhpC, there is a local unfolding of the helix, and Cp47 forms an intermolecular disulphide bond with Cr166. In comparison, in EfAhpC, there is no intermolecular disulphide bond. (*Bottom inset*) The hydrogen bond network, which stabilizes the Cp47, is shown and the interactions within 5 Å are shown in dotted lines. All the residues involved in the hydrogen bonding are presented as sticks and are labelled. The 2FoFc map of the active residues is contoured at 1σ . (B) The superposition of the EfAhpC dimer (chain A and B; cyan and magenta) as well as chain C and D; deep blue and yellow) is shown. In this conformation, the Sy atom of Cp47 faces away from the guanine group of R120. All the residues involved in the hydrogen bonding are presented as sticks and are labelled. In addition, the 2FoFc map of the active residues is contoured at 1σ .

Refinement parameters	
Refinement Statistics	
R factor ^a (%)	20.8
R free ^b (%)	25.2
Number of amino acid residues	830
Ramachandran statistics	
Most favoured (%)	92.5
Additionally allowed (%)	7.5
Generously (%)	0.0
Disallowed (%)	0.0
R.M.S. Deviations	
Bond lengths (Å)	0.004
Bond angles (°)	0.786
Mean atomic B values	
Overall	59.8
Wilson B Factor	56.23
CC1/2	99.2 (76.5)
CC*	99.8 (93.1)
Clash score	2.39
MolProbity score:	1.36

Table 5.3: Summary of refinement statistics of *EfAhpC*₁₋₁₇₂ crystal structure. (a = R-factor = $\Sigma ||FO| - |FC|| / \Sigma |FO|$, where FO and FC are measured and calculated structure factors, respectively. b = R-free = $\Sigma ||FO| - |FC| / \Sigma |FO|$, calculated from 5% of the reflections selected randomly and omitted during refinement.)

5.3 Characterization of *EfAhpC*_{V78T} to investigate role of V78 in stable oligomer formation

EfAhpC contains a valine at the oligomeric interface as opposed to threonine observed in *StAhpC* and *EcAhpC*. While the threonine forms a hydrogen bond at the oligomer interface, valine forms a hydrophobic interaction with the V78 from the neighboring chain. Studies have shown that substituting the T77 in *StAhpC* with valine resulted in a stable decamer formation [133]. To confirm the importance of V78 at the oligomeric interface of *EfAhpC*, the residue was mutated to threonine.

5.3.1 Cloning, expression and purification of *EfAhpC*_{V78T}

The valine mutant of *EfAhpC* was generated by site-directed mutagenesis. Two prominent bands were observed when the amplified gene was run on an 1.5% agarose gel. The band for the plasmid containing the mutated gene sequence was observed at around 5000 bp (Fig. 5.3.1A), this was excised from the gel and DNA extracted. Following *DpnI* digestion, the plasmid was transformed into expression strain and protein production induced with IPTG. The

protein production was assessed by a 17% SDS-PAGE gel, where sufficient protein production was observed in the induced samples (Fig. 5.3.1B).

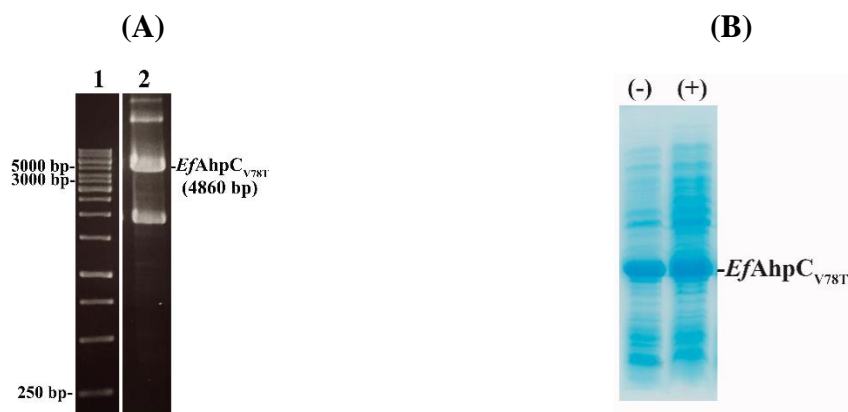


Figure 5.3.1: Amplification of gene encoding *EfAhpC_{V78T}* mutant and recombinant *EfAhpC_{V78T}* production. (A) Amplified product (lane 2) and DNA ladder in lane 1. DNA sample of 5 μ l was loaded on an agarose gel (1.5%). (B) SDS-PAGE (17%) showing the expression of recombinant protein. Protein production was induced with IPTG (1 mM). (+) and (–) indicates protein obtained from induced and uninduced cells, respectively. Band for recombinant protein was observed in both lane due to the leaky expression of pET 9d.

The protein was purified as described for the WT enzyme. Interestingly, *EfAhpC_{V78T}* eluted already at 75-150 mM of imidazole from the Ni^{2+} -NTA column (Fig. 5.3.2A). In comparison, WT *EfAhpC* eluted at 100-500 mM imidazole (Fig. 5.1.1A). The fractions were then pooled, concentrated and applied to SEC. In comparison to the WT, the mutant protein eluted at a later volume and a broader peak was observed in the chromatogram (Fig. 5.3.2B).

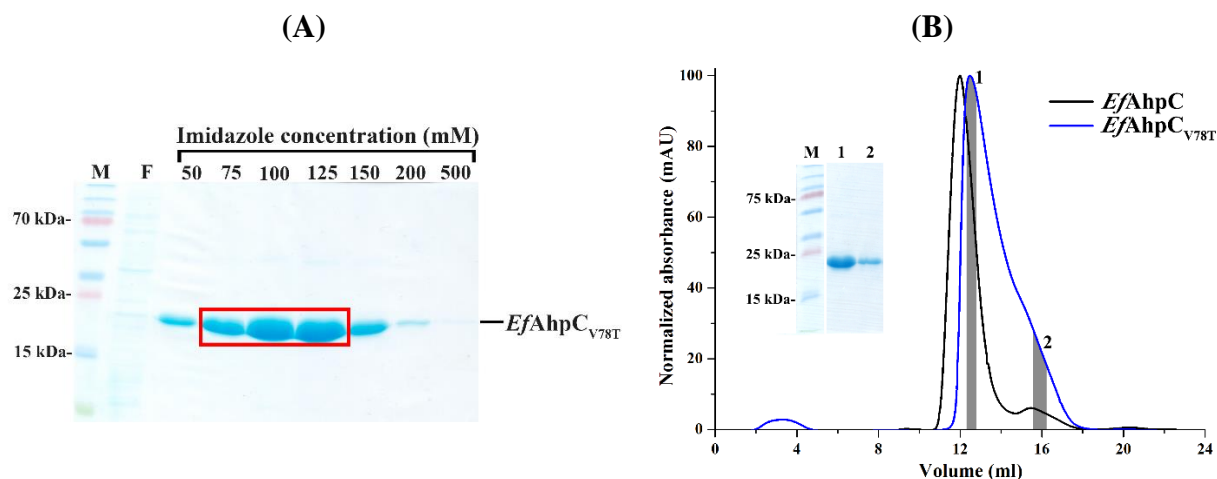


Figure 5.3.2: Purification of *EfAhpC_{V78T}*. (A) The SDS-PAGE gel shows that the mutant protein eluted at a lower imidazole concentration from the affinity chromatography column than *EfAhpC* (M= Protein marker; F= flow-through). The red box indicates the fractions pooled for SEC. (B) Comparison of SEC elution profiles of *EfAhpC* (—; black) and *EfAhpC_{V78T}* (—; blue) shows that the mutant elutes later than WT. Fractions from the main peak (1) and shoulder region (2) marked in grey, were run on SDS-PAGE gel (*inset*; lane 1 and 2 respectively) to assess quality of protein. (M= marker).

5.3.2 Enzymatic characterization of *EfAhpC*_{V78T}

The rate of decline in absorbance of NADH at 340 nm was slower for *EfAhpC*_{V78T} in comparison to *EfAhpC*, suggesting a significant drop in enzyme activity (Fig. 5.3.3A). The FOX assay also showed a lower rate of peroxide decomposition as shown in figure 5.3.3B. Since effective catalysis of peroxide reduction is linked to stable decamer, the lower catalytic efficiency for the valine mutant could be attributed to the disruption of the stable decamer as a result of mutating V78 to T.

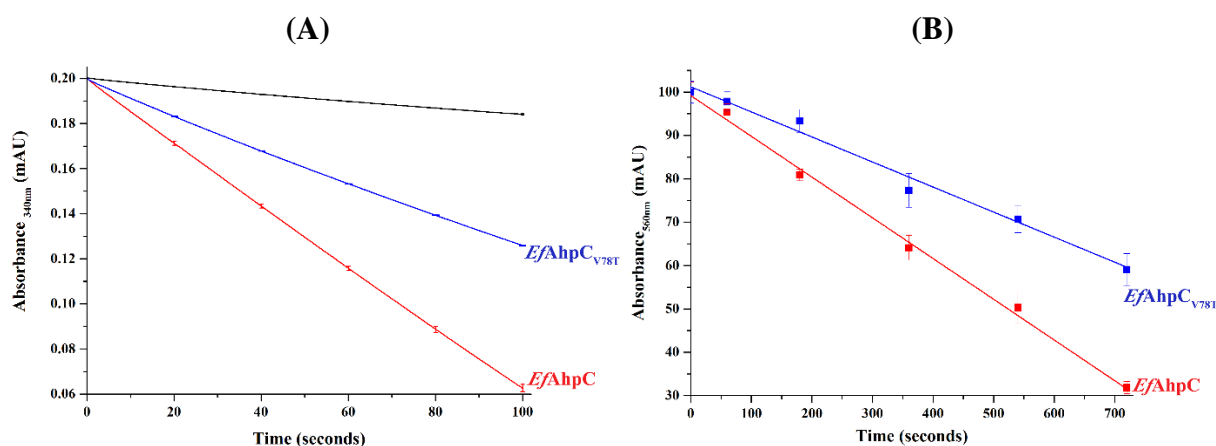


Figure 5.3.3: The valine mutant has lower peroxidative activity than *EfAhpC*. (A) NADH-dependent peroxidative activity of *EfAhpC*_{V78T} (—; blue) compared to that of *EfAhpC* (—; red). The control (—; black) represents the reaction which does not have AhpC, just WT *EfAhpF*. (B) The FOX assay for *EfAhpC*_{V78T} (—; violet) show lower decomposition of H₂O₂ in comparison to *EfAhpC* (—; red).

5.3.3 Analysis of destabilization of oligomer by DLS

Fraction from the main peak (#1 in Fig. 5.3.2B) as well as from the shoulder (#2 in Fig. 5.3.2B) were pooled, concentrated to 1 mg/ml and used for DLS studies (Fig. 5.3.4A). A hydrodynamic diameter of 10.96 ± 1.82 d.nm corresponding to an estimated molecular weight of 180 ± 29.9 was observed for fraction #1, and a diameter of 6.67 ± 2.4 d.nm (56.4 ± 20.8 kDa) was calculated for the shoulder fraction (#2), indicating destabilization of the decamer in the mutant protein. The correlation coefficient for both the data are shown in figure 5.3.4B.

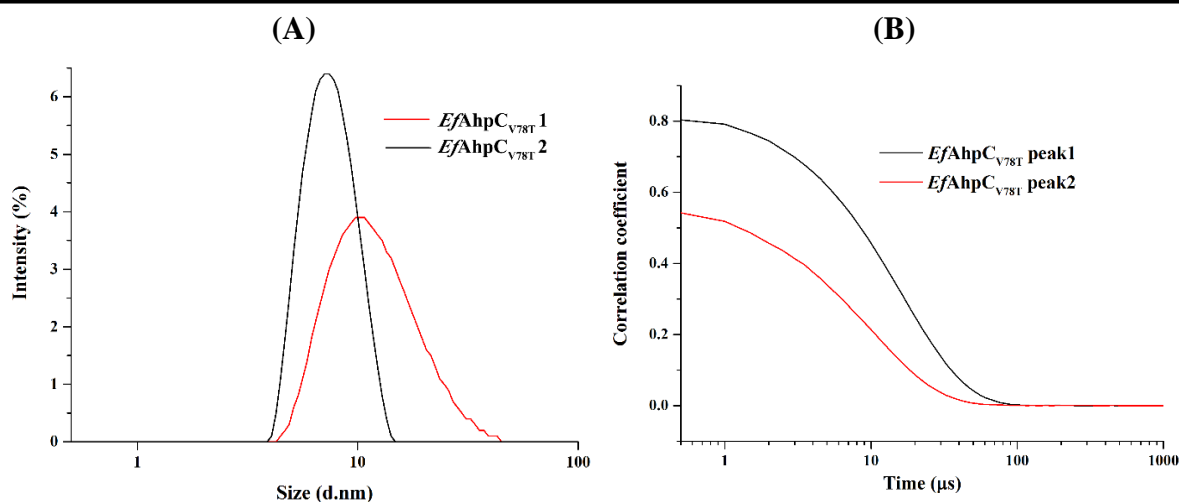


Figure 5.3.4: DLS data indicate that *EfAhpC_{V78T}* is unable to form stable decamer. (A) DLS experiments were used to study the hydrodynamic diameter of mutant from peak 1 in purification (—; black) and peak 2 (—; red). (B) The correlation coefficient shows a faster decay for sample measured from peak 2 indicating that this particle is smaller in size than the one from peak 1.

5.4 Characterization of cysteine mutants to elucidate the role of C13 and C66 in *EfAhpC*

Although the *EfAhpC* is classified as a 2-Cys peroxiredoxins, the protein is unique in the sense that it contains two extra cysteine residues apart from the conserved peroxidative and resolving cysteines. Whether these residues play a role in oligomerization or contribute to the catalysis was investigated by generating two single- (*EfAhpC_{C13N}*, *EfAhpC_{C66V}*) and a double cysteine mutant (*EfAhpC_{C13N,C66V}*).

5.4.1 Generation of cysteine mutants and recombinant protein purification

The single and double cysteine mutants were generated by site-directed mutagenesis. A single band was observed for all three gene constructs when loaded on an 1.5% agarose gel at around 5000 bp (Fig. 5.4.1A). After *DpnI* digestion, the mutated plasmids were transformed into BL21 DE3 cells for gene expression and protein production. As shown in figure 5.4.1B, thick band were observed for all three mutants, indicating high production of recombinant mutant proteins (Fig. 5.4.1B).

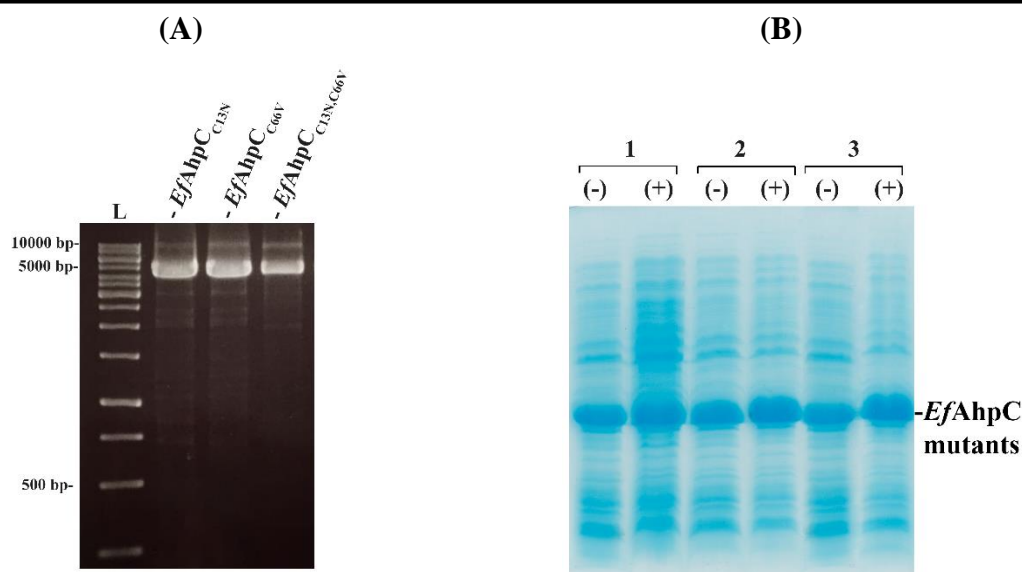


Figure 5.4.1: Site-directed mutagenesis of *EfAhpC* and respective recombinant protein production. (A) 1.5% agarose gel on which the mutated PCR products were run. Single bands were observed for all the mutants. DNA ladder was loaded in lane marked L. (B) Recombinant *EfAhpC* mutant production. SDS-PAGE gels showing production of 1) *EfAhpC*_{C13N} 2) *EfAhpC*_{C66V} 3) *EfAhpC*_{C13N,C66V}. (-) non-induced and (+) induced samples.

The cysteine mutants were purified using the protocol described for the WT protein. For all the three mutants, the protein was identified in the 100-250 mM imidazole fraction eluted from the affinity chromatography. The 125-250 mM fractions were pooled in the case of each mutant and subsequently applied to gel filtration column. The SDS-PAGE gel of the affinity chromatography fractions of the double mutant is shown in figure 5.4.2A. The single mutant *EfAhpC*_{C66V} eluted at an elution volume like the WT enzyme, whereby the single mutant *EfAhpC*_{C13N} eluted later (Fig. 5.4.2B). The shift to longer elution volumes increased with the double mutant *EfAhpC*_{C13N,C66V}, which in addition showed a shoulder between 14.5 to 18 ml, reflecting destabilization of the decamer formation (Fig. 5.4.2B).

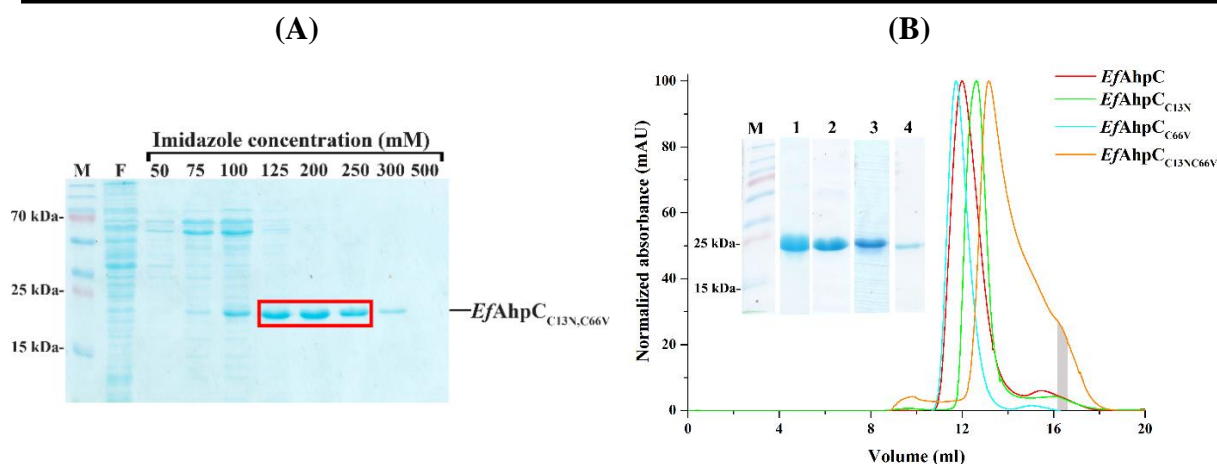


Figure 5.4.2: Purification of *EfAhpC* cysteine mutants. (A) The SDS-PAGE gel shows the double cysteine mutant affinity chromatography fractions (M= Protein marker; F= flow-through). The red box indicates the fractions pooled for SEC. (B) Comparison of S200 gel filtration elution profile of *EfAhpC* (red), *EfAhpC*_{C13N} (green), *EfAhpC*_{C66V} (cyan) and *EfAhpC*_{C13N,C66V} (orange). The elution profiles of *EfAhpC*_{C13N} and *EfAhpC*_{C13N,C66V} are shifted to the right, and for the double cysteine mutant a small shoulder was observed around 16 ml (shaded grey). (Inset) SDS-PAGE of the purified *EfAhpC*_{C13N} (lane 1), *EfAhpC*_{C66V} (lane 2) and *EfAhpC*_{C13N,C66V} *EfAhpC*_{V78T} (lane 3; shoulder-lane 4). Protein marker is used in the lane marked M.

5.4.2 Enzymatic characterization of cysteine mutants

To evaluate the effect of the cysteine mutants on enzyme activity, when NADH-dependent peroxidase activity was measured, the mutants showed reduced NADH-oxidation compared to WT *EfAhpC*. The *EfAhpC*_{C13N,C66V}, shows the lowest activity, followed by *EfAhpC*_{C13N} and *EfAhpC*_{C66V}, respectively (Fig. 5.4.3A). Also, the results of the FOX assay in figure 5.4.3B show that H₂O₂ reduction was lower in the case of *EfAhpC*_{C13N} and even more in *EfAhpC*_{C13N,C66V}. These assays demonstrate that the substitutions of the two additional cysteines lower enzymatic activity.

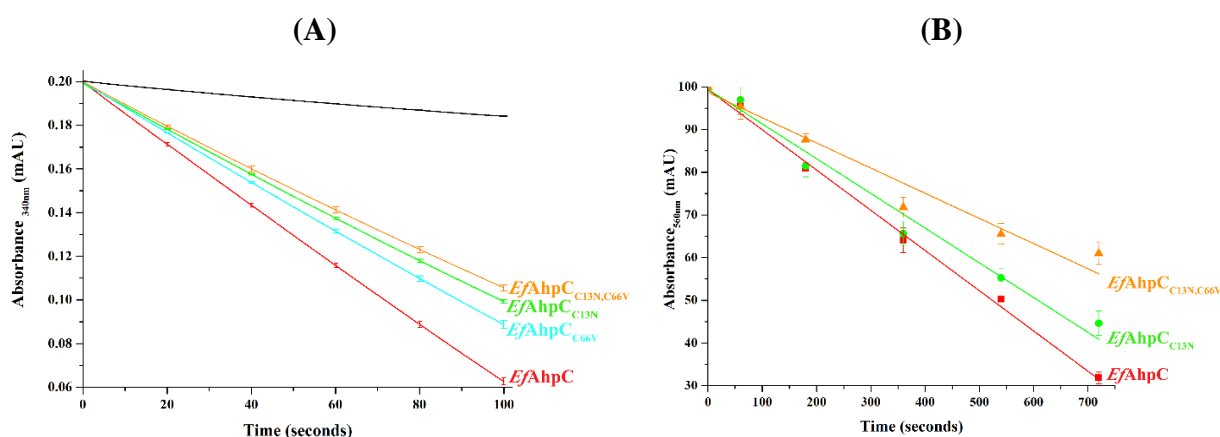


Figure 5.4.3: The cysteine mutants show lower peroxidative activity than *EfAhpC*. (A) Comparison of NADH-dependent peroxidase assay of *EfAhpC*_{C13N} (green), *EfAhpC*_{C66V} (cyan) and *EfAhpC*_{C13N,C66V} (orange) with *EfAhpC* (red), depicting that the double cysteine mutant among all the cysteine mutants, is the least efficient in catalysing H₂O₂ reduction. (B) Ferrous oxidation xylenol orange (FOX) assay corroborate the result for the cysteine mutants.

5.4.3 Analysis of oligomeric state of cysteine mutants by DLS

DLS experiments of the main peak fractions denoted in figure 3.3.25B of the two single mutants revealed a hydrodynamic diameter of 11.04 ± 1.53 d.nm (*EfAhpC*_{C13N}), 11.98 ± 4.98 d.nm (*EfAhpC*_{C66V}). For *EfAhpC*_{C13N,C66V}, the main peak fraction showed a hydrodynamic diameter of 11.34 ± 4.76 d.nm, whereas as the shoulder fraction a diameter of 7.53 ± 3.50 d.nm. Comparing the hydrodynamic diameter values of the cysteine mutants to WT *EfAhpC* (12.1 and 12.4 d.nm; section 5.1.3), destabilization of the decamer is confirmed and the destabilizing effect is more pronounced in the double cysteine mutant (Fig. 5.4.4).

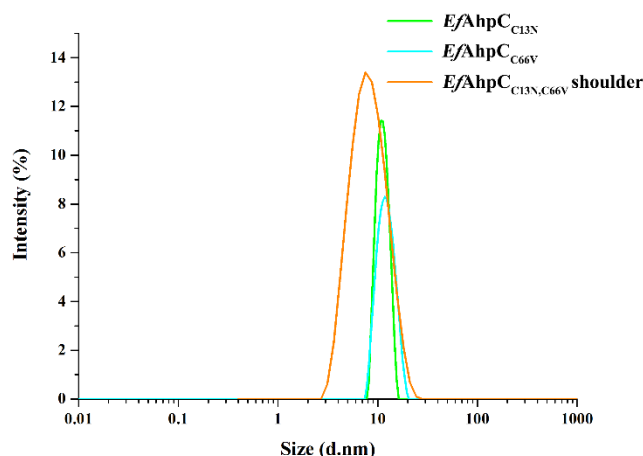


Figure 5.4.4: Destabilization of decamer in cysteine mutants is reiterated by DLS data. The dynamic light scattering data for *EfAhpC*_{C13N} (—; green), *EfAhpC*_{C66V} (—; cyan), and the shoulder fraction of *EfAhpC*_{C13N,C66V} (—; orange).

5.5 Characterization of N- and C-terminal polar and charged residues unique to *EfAhpC*

The sequence alignment revealed differences in the conserved residues at the N- and C-terminus (Fig. 5.5.1). The presented crystal structure reveals that amino acid N2 at the dimeric interface is hydrogen bonded to residue I4 of the same chain and G108 of the neighboring chain. In addition, residue D14 has a backbone interaction with A97 of the same chain. The C-terminus is important for the catalytic activity of the AhpC and N169 is different from the conserved K169 in their *E. coli* and *S. typhi* counterpart.

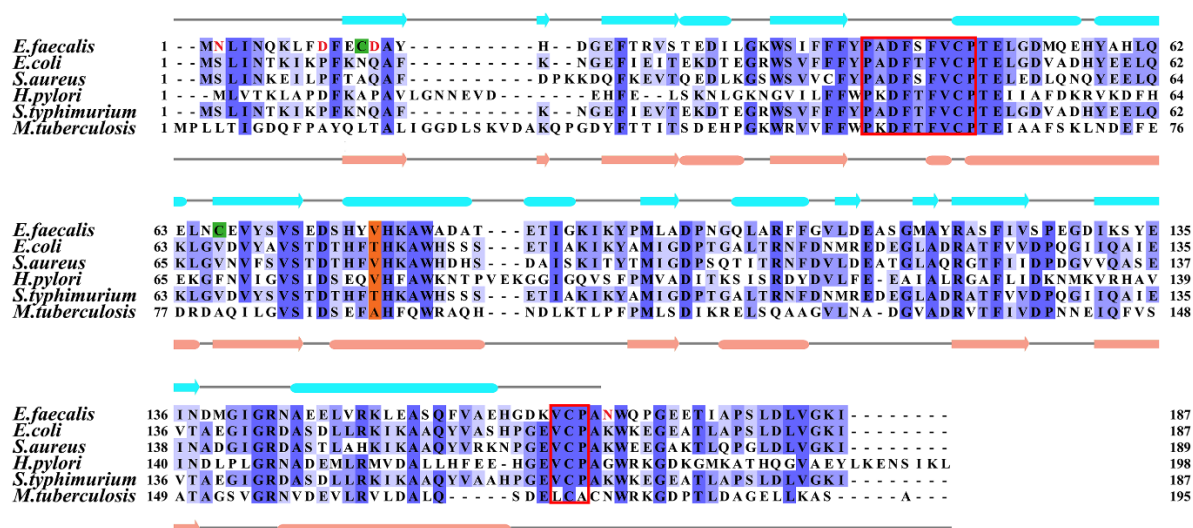


Figure 5.5.1: Multiple sequence alignment of AhpC across different bacteria. The sequence alignment of AhpCs from *E. faecalis* (GenBank: H7C7A0), *E. coli* (GenBank: P0AE08), *S. aureus* (GenBank: P0A0B7), *H. pylori* (GenBank: P21762), *S. typhimurium* (GenBank: P0A251) and *M. tuberculosis* (GenBank: P9WQB7) are shown. The secondary structure elements coloured in cyan is the *EfAhpC*₁₋₁₇₂ structure and that coloured in salmon is the structure of *S. typhimurium* AhpC (PDB ID: 1N8J). The conserved PXXXT/SXVCp-motif and VCpP motif are highlighted in a red box. The V77/T78 in the dimer-dimer interface is emphasized in orange. The C13 and C66 of *EfAhpC* are highlighted in green. The N- and C-terminal residues that were mutated are indicated in red.

5.5.1 Generation of N- and C-terminal mutants and purification of the enzyme variants

The N2, D10 and D14 at the N-terminus was mutated to S, P and Q respectively to generate a triple mutant *EfAhpC*_{N2S,D10P,D14Q}. The construct was generated by ligation dependent method of cloning with the help of mutagenic primer. The agarose gel of the amplified product revealed a single band between the 500 and 750 bp of the DNA ladder (correct size approximately 558 bp) (Fig. 5.5.2A). The digested insert was ligated into pET 9d. The N169 was mutated to K by site-directed mutagenesis. The mutated plasmid run as a single band corresponding to around 5000 bp of the marker (Fig. 5.5.2A). The plasmids for both constructs containing the correct gene sequence was transformed into BL21 for protein production (Fig. 5.5.2B).

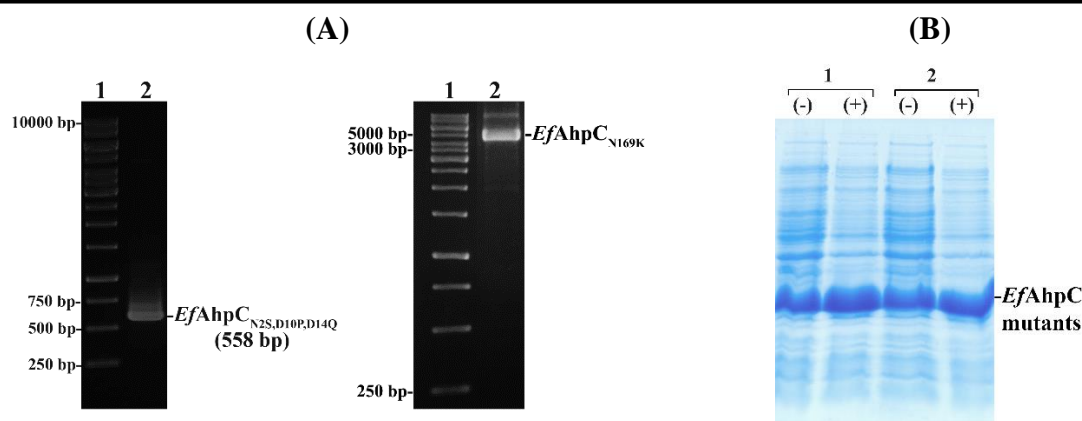


Figure 5.5.2: Amplification of gene encoding N- and C- terminus mutants and recombinant protein production. (A) Left: Amplified product (lane 2) of *EfAhpC*_{N2S,D10P,D14Q} and DNA ladder in lane 1. Right: Amplified product (lane 2) of *EfAhpC*_{N169K} and DNA ladder in lane 1. (B) Recombinant *EfAhpC* mutant production. SDS-PAGE gels showing production of 1) *EfAhpC*_{N2S,D10P,D14Q} and 2) *EfAhpC*_{N169K}. (+) and (-) indicates the presence and absence of IPTG, respectively.

The mutants were purified according to the protocol developed for *EfAhpC*, with both the recombinant proteins eluting from the affinity column at an imidazole gradient of 100-300 mM imidazole (Fig. 5.5.3A). On application to SEC, a single peak was observed for both proteins in the chromatogram as shown in figure 5.5.3B. Both protein eluted around 12 ml like the wild type. The SDS-PAGE gel for both proteins shown a single band after the gel filtration step indicating a homogeneous and pure protein for both constructs (Fig. 5.5.3B; inset).

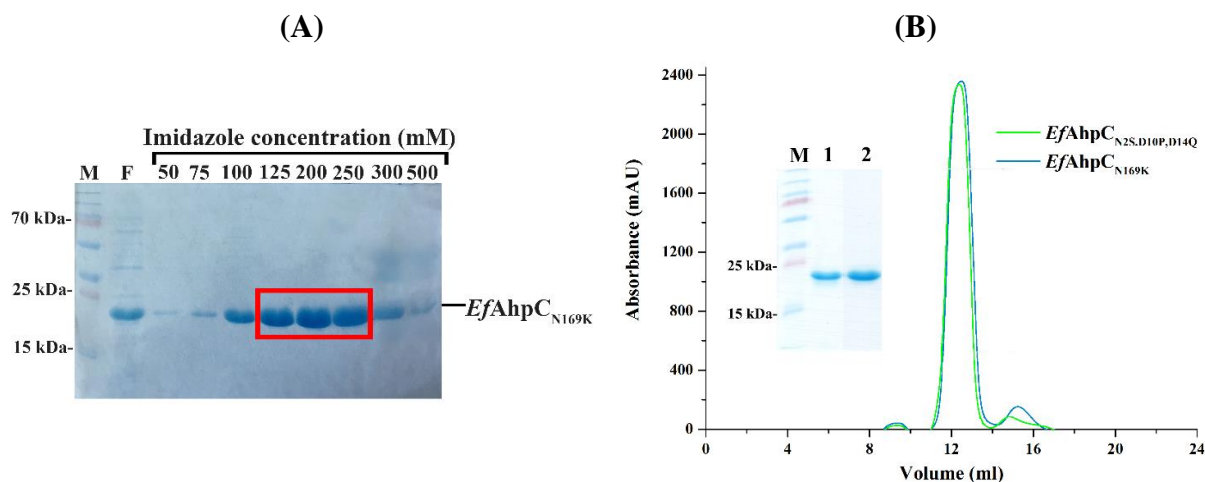


Figure 5.5.3: Purification of N- and C- terminus mutants. (A) The SDS-PAGE gel shows the imidazole concentrations at which the *EfAhpC*_{N169K} eluted from the affinity chromatography column (M= Protein marker; F= flow-through). The red box indicates the fractions pooled for the next step of purification. (B) On application to SEC, a single peak was observed in the chromatogram for both *EfAhpC*_{N2S,D10P,D14Q} (—; green) and *EfAhpC*_{N169K} (—; blue) protein. Inset: SDS-PAGE gel of the eluted SEC fractions. LM = protein marker; 1= *EfAhpC*_{N2S,D10P,D14Q} and 2= *EfAhpC*_{N169K}.

5.5.2 Enzymatic characterization of N- and C-terminal mutants

Both *EfAhpC*_{N2S,D10P,D14Q} and *EfAhpC*_{N169K} showed a reduced NADH dependent peroxidative activity with a slightly higher effect of enzyme inhibition of *EfAhpC*_{N2S,D10P,D14Q}

(about 25%; Fig. 5.5.4A). The ferrous oxidation xylenol (FOX) assay revealed that the amount of reduced H_2O_2 was lower in case of *EfAhpC*_{N2S,D10P,D14Q} and *EfAhpC*_{N169K}, confirming the effect of inhibition of both mutant enzymes (Fig. 5.5.4B). The mutation of residues at the N- and C- terminus affects peroxidative activity of the *EfAhpC*.

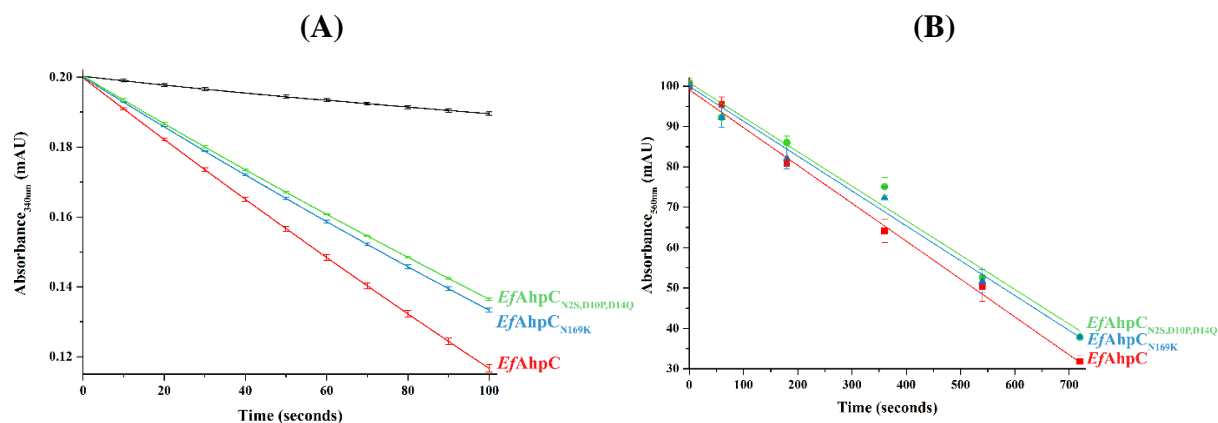


Figure 5.5.4: N- and C- terminus mutants show reduced peroxidative activity. (A) The NADH-dependent peroxidase activity of the N-terminus triple mutant *EfAhpC*_{N2S,D10P,D14Q} (green) is significantly lower than WT *EfAhpC* (red), whereas the C-terminus of the *EfAhpC*_{N169K} mutant (blue) is slightly more efficient than the triple mutant in catalyzing H_2O_2 reduction. (B) The FOX assay confirms that both the *EfAhpC*_{N2S,D10P,D14Q} (green) as well as the *EfAhpC*_{N169K} mutant (blue) affect H_2O_2 reduction.

6. Discussion- Atomic and Enzymatic Insights into *EfAhpC*

6.1.1 The *EfAhpC* forms an oxidized decamer

EfAhpC is a 2-Cys peroxiredoxin belonging to the Prx1 sub family with the basic functional unit being a homodimer. These peroxiredoxins can undergo conformational changes based on the redox state of the thiol group, with the oxidized state C_P favouring the dimer. In case of *E. coli*. [112], *S. typhii* [133], *Leishmania* [208], *Pseudomonas* [143] and *H. pylori* [209] AhpCs five conformational states have been identified depending on the redox state of the enzyme: the reduced dimer, the reduced decamer, the oxidized dimer, the hyperoxidized decamer and the hyperoxidized hyperaggregates [210, 211]. The SEC-, DLS- and EM studies of *EfAhpC* presented indicate a sixth state, which is the oxidized decamer. The electron micrographs of the oxidized and reduced protein reflect that the decamerization of the *EfAhpC* is insensitive to redox modulation of the thiol group during catalysis.

6.1.2 The structural intricacies of *EfAhpC*

Besides the mechanistic complexity of the 2-Cys peroxiredoxins to change their quaternary structure during catalysis based on the redox state of the thiol group, they undergo conformational changes at the active site. As seen in the case of most AhpCs, in the reduced state of the cysteines, the active site is substrate ready and is most competent for substrate binding. The α 2-helix on which the C_P resides, is in a FF conformation. Currently, the C_R on the C-terminus is moved across the active site, separating the C_P and C_R by 14 Å. In the oxidized state, α 2-helix is partially unfolded with the C_P adopting a LU conformation, which brings the C_P and C_R closer to facilitate the disulphide bond formation between the C_P and C_R. The C-terminal tail becomes also disordered during unwinding [123, 130]. The active site conformational changes influence the quaternary structure changes with the FF state favouring decamer formation and the LU state favouring dimerization [123, 130] (Fig. 6.1.1).

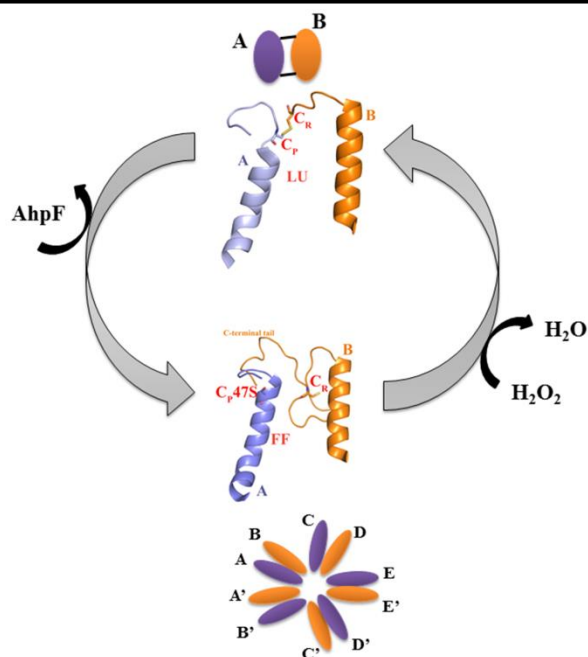


Figure 6.1.1: The connection of C_P loop dynamics and the oligomeric state of the AhpC. Besides redox modulation, the C_P loop dynamics also affects the dimer-decamer equilibrium. Only the $\alpha 2$ -helix containing C_P and $\alpha 6$ -helix containing C_R are shown for ease of depiction. Shown at the bottom of the picture depicts the fully folded (FF) conformation of C_P loop, where the catalytic cysteines are reduced and are almost 14 Å apart. In this conformation the active site pocket is substrate ready enabling attack of H_2O_2 as depicted. The decamer ring is favoured in this state. The dimers (A-blue and B-orange) assemble as pentamers to form the donut shaped ring. Since the *StAhpC*(C47S) mutant has the density for the C-terminus, the cartoon representation of the *StAhpC* crystal structure (PDB ID: 1N8J) is depicted. The top depicts the partially unwound $\alpha 2$ -helix adopting the locally unwound conformation (LU). This state favors the dimer as depicted by A and B linked by disulphide bridges (-, black). The cartoon representation of the oxidized *StAhpC* crystal structure (PDB ID: 1YF1) containing the disulphide bonded C_P and C_R are shown.

To elucidate such structural and mechanistic changes in *EfAhpC*, the crystal structure of the reduced *EfAhpC*₁₋₁₇₂ has been solved to 2.85 Å, revealing five molecules (chains A-E) in the asymmetric unit, and forming a half ring shape conformation. In the presented *EfAhpC*₁₋₁₇₂ structure both the cysteines are in the reduced state, and the C-terminal residues beyond C_R are disordered. The overall topology of reduced *EfAhpC*₁₋₁₇₂ is similar to the oxidized *EcAhpC* structure (PDB ID: 4O5R; [116]). The r.m.s.d. value between *EfAhpC*₁₋₁₇₂ and *EcAhpC* for 161 C α -atoms is 0.988 Å. This indicates that the C-terminus truncation did not modify the structural traits of the protein. However, while C_{P47} becomes exposed in the oxidized *EcAhpC* due to the local unfolding of the $\alpha 2$ -helix, resulting in an intermolecular disulfide bond with $C_{R166'}$ (C_{PS} - SC_R), the helix in the reduced state of the *EfAhpC*₁₋₁₇₂ structure adopts a partially folded conformation (Fig. 6.1.2). Thus, C_{P47} and $C_{R166'}$ move apart from a distance of 10.42 Å (*EcAhpC*) to 13.68 Å (*EfAhpC*₁₋₁₇₂) in the two homodimers with the respective sulphurs being oriented in opposite directions (Fig. 6.1.2).

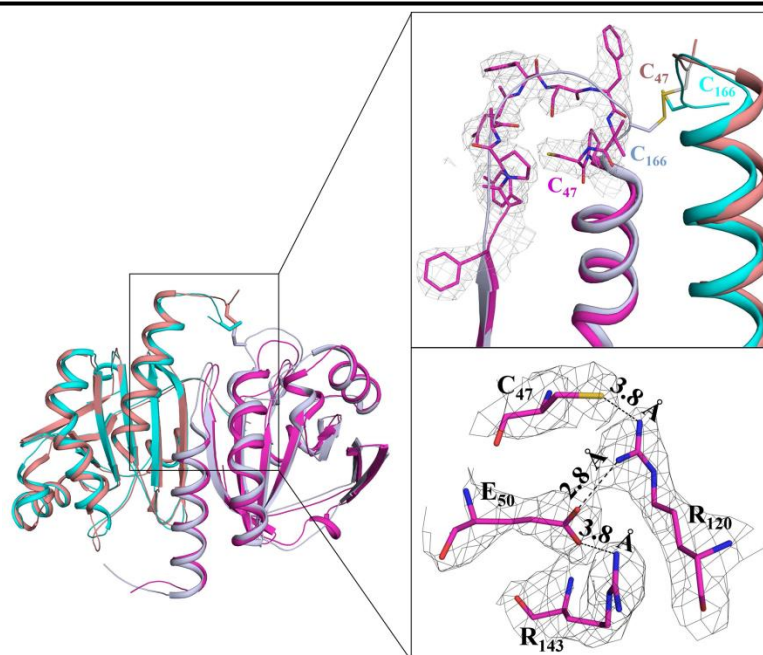


Figure 6.1.2: The superposition of the *EfAhpC* and *EcAhpC* structure. (Top inset) The final 2FoFc map at the active site $_{38}\text{FYPADFTFVCP}_{48}$, of *EfAhpC* is shown; the map is contoured at 1σ . In *EcAhpC*, there is a local unfolding of the helix, and Cp_{47} forms an intermolecular disulphide bond with C_{166} . In comparison, in *EfAhpC*, no intermolecular disulphide bond was observed. (Bottom inset) The hydrogen bond network, which stabilizes the Cp_{47} , is shown and the interactions within 5 Å are shown in dotted lines. All the residues involved in the hydrogen bonding are presented as sticks and are labelled. The 2FoFc map of the active residues is contoured at 1σ .

All AhpCs contain the local sequence motif PXXXT/SXXC_P at the active site, which is essentially invariant in all Prx families. So far, the structural details of the redox modulated dimer to decamer equilibrium described for the prokaryotic *EcAhpC* and *StAhpC* contain a threonine within the PXXXT/SXXC_P motif [112, 131, 208, 212, 213]. In the FF conformation, C_P is located in the first turn of helix α_2 , bounded by the threonine and an arginine, forming together the catalytic triad which is located close to the oligomeric interface (Fig. 6.1.3). The C γ of threonine is involved in a CH- π H bond with an aromatic residue (Y/F) of the adjoining dimer and this bond is conserved among all 2-Cys peroxiredoxins (Fig. 6.1.3; [129]). In case of *EfAhpC*, the atomic structure of *EfAhpC*₁₋₁₇₂ reveals that the respective serine residue (S44) inside the PXXXSXXC_P motif of *EfAhpC* does not exhibit such CH- π H bond with the adjacent Y77' (Fig. 6.1.3). Since the threonine of the PXXXT/SXXC_P motif provides the physical interaction between the active site and the oligomeric interface, the presence of S44 in the active-site triad of *EfAhpC* explains in part the differences seen in the absence of the equilibrium of dimer to higher quaternary structures seen for most 2-Cys Prxs.

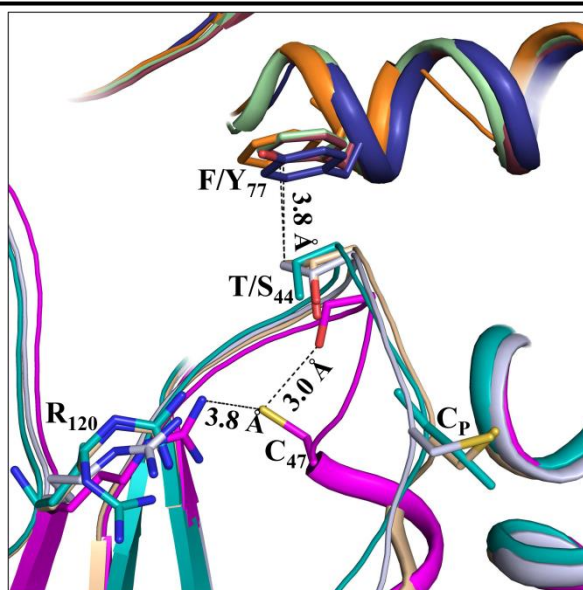


Figure 6.1.3: Comparison of the active site of AhpCs across various bacteria. Overlap of the catalytic triad residues (Cp47, T/S44, R120) and the CH- π hydrogen bond between T44 and Y/F77 of the 2-Cys Prx structures, *EfAhpC* (pink and deep blue), *EcAhpC* (PDB: 4O5Q; deep salmon and raspberry), *StAhpC* (PDB: 1YEP; wheat and pale green), *StAhpC* (PDB: 4XRA; teal and orange). Only the oligomer interface between chains B and C is shown for clarity. All the catalytic residues and F/Y77 of the adjacent chain are shown as sticks; the numbering is according to the *EfAhpC* sequence

Recent mutational studies of the eukaryotic yeast homologue proteins, containing T-S variants (Tsa1 and Tsa2, respectively), demonstrate that serine in the catalytic triad play a pivotal role in the decamer stabilization upon disulfide formation [129]. Since S44 is present in the PXXXT/SXXC_P motif for *EfAhpC*, it could explain the stable decamer in the oxidized state. Furthermore, serine does not possess an extra C γ which is seen in the case of threonine. This extra C γ in the threonine residue causes steric hindrance with nearby residues and disfavors the decameric state during the transition from the FF to LU states [129]. Amino acid S44 in *EfAhpC* could thus be a contributing factor to the stable decamer observed for this protein.

Although *EfAhpC*₁₋₁₇₂ was crystallized under reduced conditions, in the presented structure two chains (C and D) reveal the α 2-helix being partially unfolded. This does not mimic the LU conformation in the oxidized state but rather adopts an intermediate conformation in-between the catalytic FF- and LU conformation. During the LU state, the C_P and C_R' disulphide bonded active site becomes stabilized which weakens the oligomer building interface. The fact that helix α 2 is slightly unwound in two of the monomers in *EfAhpC*₁₋₁₇₂, irrespective of the absence of an intra-disulfide, signifies that the C_P loop possesses a high degree of conformational freedom in the reduced state and underpins the mechanism suggested by Wood and coworkers, in which the C_P loop is in a dynamic equilibrium between the fully folded and locally unfolded conformations in the reduced state but becomes trapped in the

locally unfolded conformation upon disulfide formation (Fig. 6.1.4).

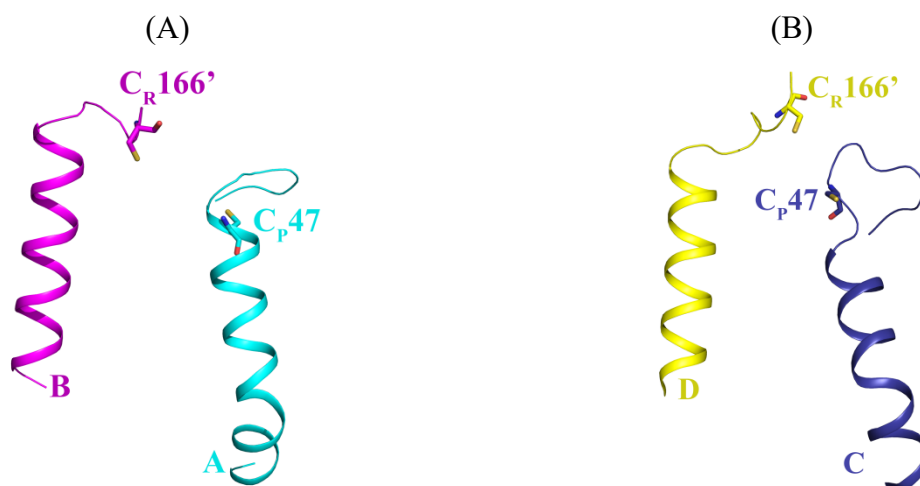


Figure 6.1.4: Two different conformations observed for the C_P loop at different chains in the crystal structure of *EfAhpC*₁₋₁₇₂. (A) Two chains (A and B) in the presented crystal structure adopt the folded conformation (FF). (B) In two chains (C and D) the α 2-helix was found to be partially unfolded. This partially unwound state is not the locally unfolded conformation (LU) where a disulphide bridge exists between C_P and C_R. The state observed was an intermediate state between LU and FF. The observed conformations signify that the C_P loop possesses a high degree of conformational freedom.

6.1.3 The C-terminus does not play a role in oligomerization of *EfAhpC* but is crucial for peroxidative activity

Although the conformational state of the C_P-loop is the switch in determining dimer-decamer equilibrium [131], the C-terminus also plays a role in stabilizing the dimer in the oxidized state and decamer in the reduced state [112]. Majority of the current high-resolution structures of reduced AhpCs has a disordered C-terminus, rendering it difficult to gain atomic insights into this flexible tail. The only exception is the C_P47S mutant structure of *S. typhi* which has a stabilizing effect on the FF conformation, since the C_P47S to R119 hydrogen bond is stronger than formed by the sulphur atom (PDB ID:1N8J; [213]).

In the reduced form, two regions in the *StAhpC*(C_P47S) (PDB ID:1N8J; [213]) mutant at the C-terminus were identified that were involved in hydrogen bonding. Region 1 included the interactions of C166, A168 and K169 forming four hydrogen bonds with the T49, G139, G142, and R143 of the adjoining dimer. The second region of hydrogen bonding involves the residues G185 and I187. These residues hydrogen bond with F43, F45, S86, and T88 of the neighboring dimer. Residue I187 also interacts with P48 and T49 of the neighboring dimer. Amino acids F43 and F45 are located at the decameric building interface. These interactions thus stabilize the decamer. However, when the α 2 helix is locally unwound to facilitate disulphide bridge formation between C_P and C_R, the interaction of F43 and F45 with the C-terminus is disrupted, destabilizing the decamer [112]. Truncation of the C-terminus in

EcAhpC revealed that the *EcAhpC* was unable to form a decamer [112]. Interestingly, the truncation of *EfAhpC* C-terminal tail did not disrupt decamer formation in either the oxidized or reduced states. This leads to the idea that there are other factors at play which stabilize the decamer building interface irrespective of the interaction of the oligomer building residues with the C-terminus.

However, C-terminal tail deletion does disrupt the catalytic activity of the protein. Due to the truncation, the C-terminal tail is unable to bind to the binding groove of the N-terminal domain of the AhpF [137]. Since the interaction of AhpF with AhpC is hindered, the disulphide bond between the catalytic cysteines are not resolved and the catalytic efficiency is lower. Such reduction in NADH-dependent peroxidative activity due to C-terminus truncation was also observed in the truncated constructs of *EcAhpC* [112].

6.1.4 The presence of valine at the oligomeric interface implicates a strong decamer formation

Mutational studies of the *StAhpC* have shown that mutating the T77 (according to *S. typhii* numbering) to I and D disrupt the decamer while substituting the same residue with V promotes decamer stability [133]. Since valine is naturally present at the oligomeric surface of *EfAhpC*, it was hypothesized that the redox-modulation independent decamer could be a result of V78 (according to *E. faecalis* numbering). This hypothesis was tested by mutating V78 to T. SEC- and DLS studies indicated that mutation of this residues destabilizes the decamer and that V78 is a key player in stable decamer formation.

As highlighted by the *EfAhpC*₁₋₁₇₂ structure, residue V78 adopts the same conformation as T77 in *StAhpC* and has hydrophobic interaction with V78' along with the hydrogen bond interactions at the oligomer interface in the *EfAhpC* structure, which contribute to the stabilization of the decamer (Fig. 5.2.11 in section 5.2.5). The sequence alignment in figure in introduction shows that a valine is also naturally present instead of threonine in the *H. pylori* AhpC (*HpAhpC*). The crystallographic structure of *HpAhpC* reveals that helix $\alpha 5$ (residues N158-F162) is shifted towards the same motif of the second subunit of the dimer, increasing the area of interaction of the dimer interface with the residues E151 and R154 being involved in hydrogen bonding [133]. In comparison, in *EfAhpC*, amino acids E147 and R150 make salt bridge interactions at the same positions. In case of *EcAhpC* and *StAhpC* the glutamic acid is replaced by an aspartic acid [112, 213].

6.1.5 Residues C13 and C66 of *EfAhpC* contribute to decamer stability

Substitution of one of the two additional cysteine residues of *EfAhpC*, C13, resulted in destabilization of the decameric ring as observed from the elution profile of *EfAhpC*_{C13N} (Fig. 5.3.2B in section 5.3.1), which was even stronger in the *EfAhpC*_{C13N,C66V} double mutant and also reflected by the significant change in the DLS data (Fig. 5.3.4 in section 5.3.3). Interestingly, C13 is located near the oligomeric interface and therefore, may contribute to stabilize this interface. Taken together, the novel stable decamer formation independent to redox modulation of *EfAhpC* is a concerted action of at least the critical residues S44 and V78 and the *E. faecalis* specific additional cysteine residues C13 and C66.

6.1.6 Decamer destabilizing mutations of *EfAhpC* leads to reduced catalytic efficiency of *EfAhpC*

The T77I and T77D mutation of *StAhpC*, which destabilizes the decamer, showed reduced peroxidative activity. It is believed that the decamer is essential for proper functioning of AhpCs. The probable reason for this is that the interaction of residues at the oligomeric interface helps to stabilize a loop involved in the formation of the active site. The V78 to T mutation in the *EfAhpC* resulted in destabilized decamers and the NADH-dependent activity of the mutant enzyme was consequently lower than the WT enzyme (Fig. 5.3.3 in 5.3.2). A similar drop in NADH-dependent peroxidative activity is also observed in the cysteine mutants of AhpC (Fig. 5.4.3 in section 5.4.2) and this is linked to the destabilization of decamer building ability of the *EfAhpC*_{C13N} and *EfAhpC*_{C13N,C66V}.

6.1.7 The unique residues at the N- and C-terminus contribute to the catalytic efficiency of the robust *EfAhpC*

A comparison of the protein sequence of *EfAhpC*, *EcAhpC* and *StAhpC* demonstrates that the conserved residues S2, P10, and Q14 at the N-terminus of *EcAhpC* and *StAhpC* are replaced in *EfAhpC* by N2, D10, and D14, respectively. According to the *EfAhpC*₁₋₁₇₂ structure, the amide group of amino acid N2 forms an intra-chain hydrogen bond with I4 and an inter-chain hydrogen bond with G109 of the neighboring chain, stabilizing the dimer interface (Fig. 6.1.5A-B). Residue D14 is present near the oligomeric interface (Fig. 6.1.5A-B), and small structural perturbations at this region could ultimately affect the peroxidative capability of the protein. Although the oligomeric state of the triple mutant *EfAhpC*_{N2S,D10P,D14Q} is not altered, an altered peroxidative activity for this mutant is presented, which could be due

to altered interaction with the *EfAhpF* due to the change in structural perturbation. Although the C-terminal residues beyond 162 are not well resolved in the *EfAhpC*₁₋₁₇₂ crystal structure, a comparison with the oxidized *EcAhpC*₁₋₁₇₂ [112] and reduced *StAhpC* structure [213] reveal that the C-terminal tail is present between the oligomeric interface and the active site loop in the oxidized state and in the reduced state is positioned closer to the dimeric interface (Fig. 6.1.5A-B). Particularly, N169 (K169 in *EcAhpC* and K168 in *StAhpC*) is present in-between the catalytic site and oligomeric interface in the oxidized state and near the dimer interface in reduced state (Fig. 6.1.5A-B). Therefore, substitution of *EfAhpC* N169 by the respective K169 of *EcAhpC* might interfere with catalytic efficiency of the protein as seen for the *EfAhpC*_{N169K} mutant, which has a reduced peroxidative activity when compared to WT *EfAhpC*.

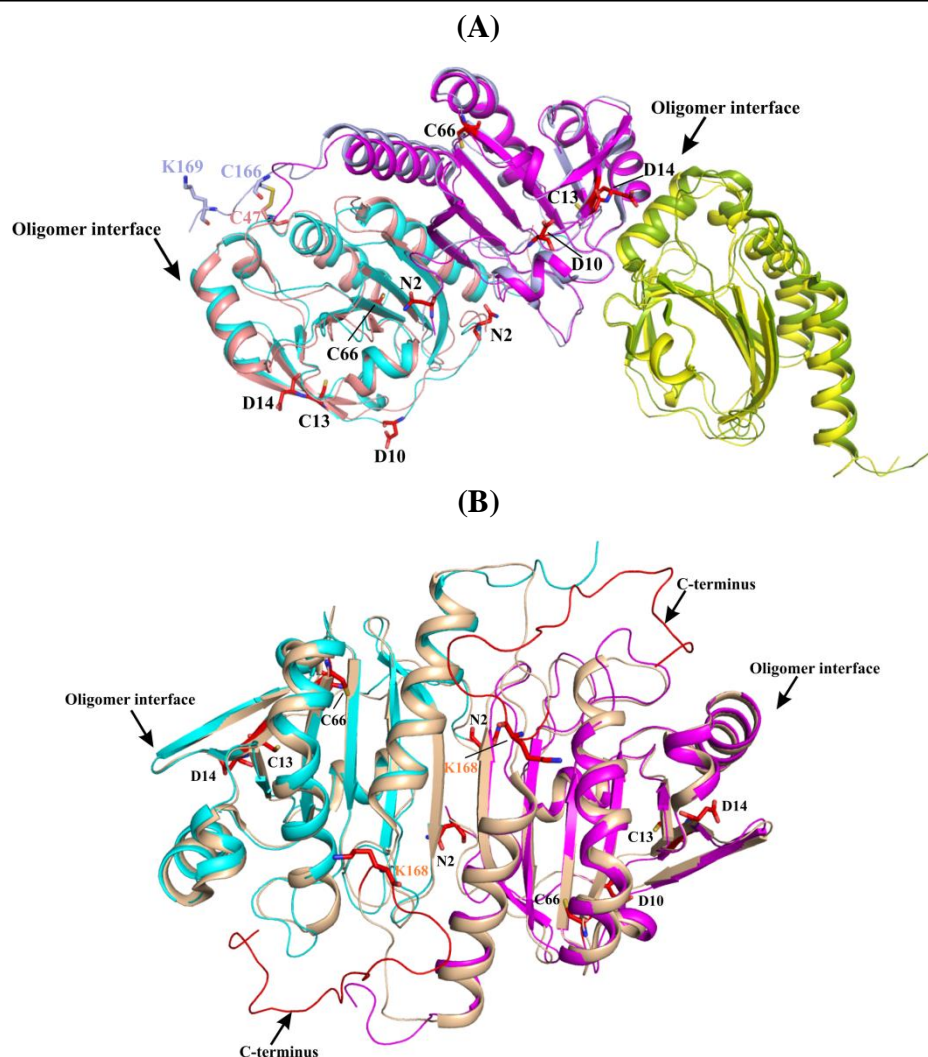


Figure 6.1.5: Interaction of unique residues at N- and C-terminus of *EfAhpC*. (A) Superimposition of the *EfAhpC*- (chain A and B; cyan and magenta, and neighboring oligomer chain C; yellow) and oxidized *EcAhpC* structure (PDB: 4O5Q; dimer chains; deep salmon and light blue and neighboring oligomer chain; green). The N-terminal residues N2, D10, D14 and the two novel cysteines C13 and C66, which were mutated, are represented in sticks (red). The C-terminal residue N169, which was mutated, is represented as sticks in light blue. The amino acids D10, C13 and D14 are located near the oligomeric interface, indicating their role in stabilizing the oligomeric interface, while N2, that is present at the dimeric interface, stabilizes the dimer interface via hydrogen bonding. Perturbations of these interactions by amino acid substitution could cause altered peroxidative activity of the mutant protein. The C-terminal tail in the oxidized structure is present between the oligomeric interface and the catalytic site. (B) Superimposition of the *EfAhpC* structure (chains A and B; cyan and magenta) and reduced *StrAhpC* (PDB: 1N8J; beige) show that the C-terminal tail is present closer to the dimeric interface.

7. Conclusion and Future Perspectives

In this thesis, SAXS has been utilized to characterize the flaviviral NS3 in solution. This technique has recently gained popularity for its capacity to characterize structural aspects of macromolecules in solutions. The application of SAXS is varied, and it ranges from determination of low-resolution structures of macromolecules, flexibility characterization, to probing conformational modifications in response to alterations in the environment of the sample being studied. In this study, SAXS, along with existing structural information obtained from NMR and X-ray crystallography, extends the structural knowledge of NS3 protein and provide dynamic information. DENV NS5 and NS3 are two important proteins that form the catalytic core for viral genome replication. Previously, SAXS studies with DENV NS5 revealed a different domain-domain arrangement of NS5 [68] as compared to the reported crystallographic structure (PDB ID: 4V0R; [67]), with the MTase positioned on top of RdRp domain. The DENV-2 NS3 was found be an extended molecule in solution, and CRYSOLE fitting of the experimental data to the theoretical scattering curve generated from the crystallographic structures revealed that, unlike DENV NS5, the solution structure of NS3 is similar to the reported extended crystallographic structures of DENV NS3 (PDB IDs: 2VBC; 2WHX; [80, 92]). DENV-2 NS3 is a flexible molecule, and EOM studies reveal that besides the two reported conformations of NS3, the molecule can exist in several conformers in solution, which corroborates the fact that NS3 must undergo a wide range of motion to perform its crucial roles in the virus replicative cycle. This mobility was found to be mediated by the linker of protein. Swapping of the PPAVP residues of the related HCV NS3 into linker of the DENV-2 NS3 did not alter the conformation of the engineered mutant.

Till date, the full-length high-resolution structure of the ZIKV NS3 is missing. With help of SAXS data acquired for ZIKV NS2B₁₈NS3, rigid body modelling was performed to gain insights into the domain-domain arrangement of the protein. Like the DENV counterpart, the ZIKV NS3 is also extended in solution, and can adopt multiple conformations to be able to interact with other components of the replication complex. Mutagenesis studies of the ZIKV NS3 linker residues reveal that T174, V176, E177 and C178 are crucial for the stability of the protein. The importance of these residues for the ZIKV NS3 function and how these residues affect the overall structure of the protein can be probed in future.

Both the flaviviral NS3 and helicase were enzymatically active. The activities could be inhibited by ATPase inhibitors like resveratrol and quercetin. However, the effect of quercetin on ATP hydrolysis of ZIKV NS3 is lower compared to the DENV counterpart which could mean that the inhibitor acts differently in both proteins.

Future studies with the DENV- and ZIKV NS3 will focus on studying the conformational changes that occur during substrate-binding, such as RNA, using time-resolved SAXS (TR-SAXS). TR-SAXS allows to collect data at different time points of a reaction, enabling to further our knowledge regarding the mechanism(s) of replication. Further, binding kinetics of NS3 and NS5 and the interaction of these protein with RNA using biophysical techniques like ITC.

The AhpR system of *E. faecalis* enables the organism to combat intolerant ROS conditions of macrophages of the host. Recently, the *EfAhpF* has been studied in much details, revealing a unique swapping of the catalytic domains of AhpF. In this thesis, systematic efforts were made to study the AhpC counterpart. DLS and EM studies showed that the *EfAhpC*, unlike the *EcAhpC*, does not undergo redox modulation of its quaternary structure and exists as decamer in even in the oxidized state. The C-terminally truncated protein was used for structural studies and the first crystal structure of *EfAhpC* was solved at 2.85 Å. This truncated protein in spite of being catalytically inactive, did not lose its ability to form a decameric ring. The critical role the V78 plays in the decamer formation was established in this study. The additional cysteines in this protein in a concerted effect contribute to the proper functionality of the protein as well as its oligomer building capability. The specific N- (N2, D10, D14) and C-terminal (N169) influence peroxidative activity of the protein. This study helps to understand the variation of nature in key enzymes like AhpC-AhpF complex.

Since the thioredoxin like domain is situated at the C-terminus of *EfAhpF* and the pyridine nucleotide-disulfide oxidoreductase domain at the N-terminal, it would be interesting to investigate the mechanism of the interaction of *EfAhpC* and *EfAhpF* in future. NMR titration studies can be utilized to investigate the region of interaction of *EfAhpC* with *EfAhpF*. Also, *in-vivo* mutagenesis can be performed for AhpC, to identify critical residues of the enzyme for survival of the pathogen.

8. References

1. Huang, Y.-J.S., et al., *Flavivirus-Mosquito Interactions*. Viruses, 2014. **6**(11): p. 4703-4730.
2. Simmonds, P., et al., *ICTV Virus Taxonomy Profile: Flaviviridae*. The Journal of General Virology, 2017. **98**(1): p. 2-3.
3. Kuno, G., et al., *Phylogeny of the Genus Flavivirus*. Journal of Virology, 1998. **72**(1): p. 73-83.
4. Murray, N.E.A., M.B. Quam, and A. Wilder-Smith, *Epidemiology of dengue: past, present and future prospects*. Clinical Epidemiology, 2013. **5**: p. 299-309.
5. Gubler, D.J., *Dengue, Urbanization and Globalization: The Unholy Trinity of the 21(st) Century*. Tropical Medicine and Health, 2011. **39**(4 Suppl): p. 3-11.
6. Halstead, S.B., *Etiologies of the Experimental Dengues of Siler and Simmons**. The American Journal of Tropical Medicine and Hygiene, 1974. **23**(5): p. 974-982.
7. World Health Organization. Regional Office for South-East Asia., *Comprehensive guidelines for prevention and control of dengue and dengue haemorrhagic fever*. Rev. and expanded. ed. SEARO Technical publication series. 2011, New Delhi, India: World Health Organization Regional Office for South-East Asia. xiv, 196 p.
8. Perera, R. and R.J. Kuhn, *Structural proteomics of dengue virus*. Current Opinion in Microbiology, 2008. **11**(4): p. 369-377.
9. Bhatt, S., et al., *The global distribution and burden of dengue*. Nature, 2013. **496**: p. 504.
10. Shepard, D.S., E.A. Undurraga, and Y.A. Halasa, *Economic and Disease Burden of Dengue in Southeast Asia*. PLOS Neglected Tropical Diseases, 2013. **7**(2): p. e2055.
11. Arima, Y. and T. Matsui, *Epidemiologic update of dengue in the Western Pacific Region, 2010*. Western Pacific Surveillance and Response Journal : WPSAR, 2011. **2**(2): p. 4-8.
12. Chan, H.B.Y., C.H. How, and C.W. Ng, *Definitive tests for dengue fever: when and which should I use?* Singapore Med J, 2017. **58**(11): p. 632-635.
13. Phoo, W.W., et al., *Structure of the NS2B-NS3 protease from Zika virus after self-cleavage*. Nature Communications, 2016. **7**: p. 13410.
14. Broutet, N., et al., *Zika Virus as a Cause of Neurologic Disorders*. New England Journal of Medicine, 2016. **374**(16): p. 1506-1509.
15. Carteaux, G., et al., *Zika Virus Associated with Meningoencephalitis*. New England Journal of Medicine, 2016. **374**(16): p. 1595-1596.
16. Mécharles, S., et al., *Acute myelitis due to Zika virus infection*. The Lancet. **387**(10026): p. 1481.
17. Song, B.-H., et al., *Zika virus: History, epidemiology, transmission, and clinical presentation*. Journal of Neuroimmunology, 2017. **308**: p. 50-64.
18. Dick, G.W., S.F. Kitchen, and A.J. Haddow, *Zika virus. I. Isolations and serological specificity*. Trans R Soc Trop Med Hyg, 1952. **46**(5): p. 509-20.
19. Duffy, M.R., et al., *Zika Virus Outbreak on Yap Island, Federated States of Micronesia*. New England Journal of Medicine, 2009. **360**(24): p. 2536-2543.
20. Rodenhuis-Zybert, I.A., J. Wilschut, and J.M. Smit, *Dengue virus life cycle: viral and host factors modulating infectivity*. Cell Mol Life Sci, 2010. **67**(16): p. 2773-86.
21. Saiz, J.C., et al., *Zika Virus: the Latest Newcomer*. Front Microbiol, 2016. **7**: p. 496.
22. Sriurairatna, S. and N. Bhamarapravati, *Replication of dengue-2 virus in Aedes albopictus mosquitoes. An electron microscopic study*. Am J Trop Med Hyg, 1977. **26**(6 Pt 1): p. 1199-205.
23. Molina-Cruz, A., et al., *Effect of mosquito midgut trypsin activity on dengue-2 virus infection and dissemination in Aedes aegypti*. Am J Trop Med Hyg, 2005. **72**(5): p. 631-7.
24. Salazar, M.I., et al., *Dengue virus type 2: replication and tropisms in orally infected Aedes aegypti mosquitoes*. BMC Microbiology, 2007. **7**(1): p. 9.
25. Wong, L.P. and S. AbuBakar, *Health Beliefs and Practices Related to Dengue Fever: A Focus Group Study*. PLOS Neglected Tropical Diseases, 2013. **7**(7): p. e2310.
26. Sampath, A. and R. Padmanabhan, *Molecular targets for flavivirus drug discovery*. Antiviral Res, 2009. **81**(1): p. 6-15.
27. Hamel, R., et al., *Biology of Zika Virus Infection in Human Skin Cells*. Journal of Virology, 2015. **89**(17): p. 8880-8896.

28. Alcaraz-Estrada, S.L., M. Yocupicio-Monroy, and R.M.d. Angel, *Insights into dengue virus genome replication*. Future Virology, 2010. **5**(5): p. 575-592.
29. Mukhopadhyay, S., R.J. Kuhn, and M.G. Rossmann, *A structural perspective of the flavivirus life cycle*. Nat Rev Micro, 2005. **3**(1): p. 13-22.
30. Kostyuchenko, V.A., et al., *Structure of the thermally stable Zika virus*. Nature, 2016. **533**: p. 425.
31. Prasad, V.M., et al., *Structure of the immature Zika virus at 9 Å resolution*. Nature Structural & Molecular Biology, 2017. **24**: p. 184.
32. Zhang, X., et al., *Dengue structure differs at the temperatures of its human and mosquito hosts*. Proceedings of the National Academy of Sciences of the United States of America, 2013. **110**(17): p. 6795-6799.
33. Shi, Y. and G.F. Gao, *Structural Biology of the Zika Virus*. Trends in Biochemical Sciences. **42**(6): p. 443-456.
34. Ma, L., et al., *Solution structure of dengue virus capsid protein reveals another fold*. Proceedings of the National Academy of Sciences of the United States of America, 2004. **101**(10): p. 3414-3419.
35. Ann-Claire, G., et al., *Detection of Zika Virus in Urine*. Emerging Infectious Disease journal, 2015. **21**(1): p. 84.
36. Barzon, L., et al., *Isolation of infectious Zika virus from saliva and prolonged viral RNA shedding in a traveller returning from the Dominican Republic to Italy, January 2016*. Eurosurveillance, 2016. **21**(10): p. 30159.
37. Mansuy, J.M., et al., *Zika virus in semen and spermatozoa*. The Lancet Infectious Diseases. **16**(10): p. 1106-1107.
38. Luo, D., et al., *Flexibility between the protease and helicase domains of the dengue virus NS3 protein conferred by the linker region and its functional implications*. J Biol Chem, 2010. **285**(24): p. 18817-27.
39. Nomaguchi, M., et al., *De Novo Synthesis of Negative-Strand RNA by Dengue Virus RNA-Dependent RNA Polymerase In Vitro: Nucleotide, Primer, and Template Parameters*. Journal of Virology, 2003. **77**(16): p. 8831-8842.
40. Klema, V.J., et al., *Dengue Virus Nonstructural Protein 5 (NS5) Assembles into a Dimer with a Unique Methyltransferase and Polymerase Interface*. PLOS Pathogens, 2016. **12**(2): p. e1005451.
41. Rastogi, M., N. Sharma, and S.K. Singh, *Flavivirus NS1: a multifaceted enigmatic viral protein*. Virology Journal, 2016. **13**(1): p. 131.
42. Somnuk, P., et al., *N-linked glycosylation of Dengue virus NS1 protein modulates secretion, cell-surface expression, hexamer stability, and interactions with human complement*. Virology, 2011. **413**(2): p. 253-264.
43. Pryor, M.J. and P.J. Wright, *Glycosylation Mutants of Dengue Virus NS1 Protein*. Journal of General Virology, 1994. **75**(5): p. 1183-1187.
44. Akey, D.L., et al., *Flavivirus NS1 structures reveal surfaces for associations with membranes and the immune system*. Science, 2014. **343**(6173): p. 881-5.
45. Chambers, T.J., D.W. McCourt, and C.M. Rice, *Yellow fever virus proteins NS2A, NS2B, and NS4B: identification and partial N-terminal amino acid sequence analysis*. Virology, 1989. **169**(1): p. 100-9.
46. Falgout, B. and L. Markoff, *Evidence that flavivirus NS1-NS2A cleavage is mediated by a membrane-bound host protease in the endoplasmic reticulum*. Journal of Virology, 1995. **69**(11): p. 7232-7243.
47. Xie, X., et al., *Membrane Topology and Function of Dengue Virus NS2A Protein*. Journal of Virology, 2013. **87**(8): p. 4609-4622.
48. Mackenzie, J.M., et al., *Subcellular localization and some biochemical properties of the flavivirus Kunjin nonstructural proteins NS2A and NS4A*. Virology, 1998. **245**(2): p. 203-15.
49. Kümmerer, B.M. and C.M. Rice, *Mutations in the Yellow Fever Virus Nonstructural Protein NS2A Selectively Block Production of Infectious Particles*. Journal of Virology, 2002. **76**(10): p. 4773-4784.

50. Muñoz-Jordán, J.L., et al., *Inhibition of interferon signaling by dengue virus*. Proceedings of the National Academy of Sciences of the United States of America, 2003. **100**(24): p. 14333-14338.
51. Tu, Y.-C., et al., *Blocking Double-Stranded RNA-Activated Protein Kinase PKR by Japanese Encephalitis Virus Nonstructural Protein 2A*. Journal of Virology, 2012. **86**(19): p. 10347-10358.
52. Yoon, K.-J., et al., *Zika-Virus-Encoded NS2A Disrupts Mammalian Cortical Neurogenesis by Degrading Adherens Junction Proteins*. Cell Stem Cell. **21**(3): p. 349-358.e6.
53. Lin, M.H., et al., *Membrane undulation induced by NS4A of Dengue virus: a molecular dynamics simulation study*. J Biomol Struct Dyn, 2014. **32**(10): p. 1552-62.
54. Miller, S., et al., *The Non-structural Protein 4A of Dengue Virus Is an Integral Membrane Protein Inducing Membrane Alterations in a 2K-regulated Manner*. Journal of Biological Chemistry, 2007. **282**(12): p. 8873-8882.
55. Stern, O., et al., *An N-Terminal Amphipathic Helix in Dengue Virus Nonstructural Protein 4A Mediates Oligomerization and Is Essential for Replication*. Journal of Virology, 2013. **87**(7): p. 4080-4085.
56. Zou, J., et al., *Mapping the Interactions between the NS4B and NS3 Proteins of Dengue Virus*. J Virol, 2015. **89**(7): p. 3471-83.
57. Zou, J., et al., *Characterization of Dengue Virus NS4A and NS4B Protein Interaction*. Journal of Virology, 2015. **89**(7): p. 3455-3470.
58. Liang, Q., et al., *Zika Virus NS4A and NS4B Proteins Deregate Akt-mTOR Signaling in Human Fetal Neural Stem Cells to Inhibit Neurogenesis and Induce Autophagy*. Cell Stem Cell. **19**(5): p. 663-671.
59. Zou, J., et al., *Mapping the Interactions between the NS4B and NS3 Proteins of Dengue Virus*. Journal of Virology, 2015. **89**(7): p. 3471-3483.
60. Miller, S., S. Sparacio, and R. Bartenschlager, *Subcellular Localization and Membrane Topology of the Dengue Virus Type 2 Non-structural Protein 4B*. Journal of Biological Chemistry, 2006. **281**(13): p. 8854-8863.
61. Lim, S.P., et al., *A Crystal Structure of the Dengue Virus Non-structural Protein 5 (NS5) Polymerase Delineates Interdomain Amino Acid Residues That Enhance Its Thermostability and de Novo Initiation Activities*. The Journal of Biological Chemistry, 2013. **288**(43): p. 31105-31114.
62. Davidson, A.D., *Chapter 2 New Insights into Flavivirus Nonstructural Protein 5*, in *Advances in Virus Research*. 2009, Academic Press. p. 41-101.
63. Zhao, Y., et al., *Molecular basis for specific viral RNA recognition and 2'-O-ribose methylation by the dengue virus nonstructural protein 5 (NS5)*. 2015.
64. Lambert, S.M., et al., *The crystal structure of NS5A domain 1 from genotype 1a reveals new clues to the mechanism of action for dimeric HCV inhibitors*. Protein Sci, 2014. **23**(6): p. 723-34.
65. Choi, K.H., *Viral Polymerases*, in *Viral Molecular Machines*, M.G. Rossmann and V.B. Rao, Editors. 2012, Springer US: Boston, MA. p. 267-304.
66. Boehr, D.D., et al., *Structure, Dynamics, and Fidelity of RNA-Dependent RNA Polymerases*, in *Nucleic Acid Polymerases*, K.S. Murakami and M.A. Trakselis, Editors. 2014, Springer Berlin Heidelberg: Berlin, Heidelberg. p. 309-333.
67. Zhao, Y., et al., *A Crystal Structure of the Dengue Virus NS5 Protein Reveals a Novel Interdomain Interface Essential for Protein Flexibility and Virus Replication*. PLOS Pathogens, 2015. **11**(3): p. e1004682.
68. Saw, W.G., et al., *Structural insight and flexible features of NS5 proteins from all four serotypes of Dengue virus in solution*. Acta Crystallographica Section D, 2015. **71**(11): p. 2309-2327.
69. Subramanian Manimekalai, M.S., et al., *Identification of the critical linker residues conferring differences in the compactness of NS5 from Dengue virus serotype 4 and NS5 from Dengue virus serotypes 1-3*. Acta Crystallographica Section D, 2016. **72**(6): p. 795-807.
70. Kapoor, M., et al., *Association between NS3 and NS5 proteins of dengue virus type 2 in the putative RNA replicase is linked to differential phosphorylation of NS5*. J Biol Chem, 1995. **270**(32): p. 19100-6.

71. Brooks, A.J., et al., *The interdomain region of dengue NS5 protein that binds to the viral helicase NS3 contains independently functional importin beta 1 and importin alpha/beta-recognized nuclear localization signals*. J Biol Chem, 2002. **277**(39): p. 36399-407.
72. Kumar, A., et al., *Nuclear Localization of Dengue Virus Nonstructural Protein 5 Does Not Strictly Correlate with Efficient Viral RNA Replication and Inhibition of Type I Interferon Signaling*. Journal of Virology, 2013. **87**(8): p. 4545-4557.
73. Pryor, M.J., et al., *Nuclear Localization of Dengue Virus Nonstructural Protein 5 Through Its Importin α/β -Recognized Nuclear Localization Sequences is Integral to Viral Infection*. Traffic, 2007. **8**(7): p. 795-807.
74. Tay, M.Y.F., et al., *The C-terminal 18 Amino Acid Region of Dengue Virus NS5 Regulates its Subcellular Localization and Contains a Conserved Arginine Residue Essential for Infectious Virus Production*. PLOS Pathogens, 2016. **12**(9): p. e1005886.
75. Zhao, B., et al., *Structure and function of the Zika virus full-length NS5 protein*. Nature Communications, 2017. **8**: p. 14762.
76. Upadhyay, A.K., et al., *Crystal structure of full-length Zika virus NS5 protein reveals a conformation similar to Japanese encephalitis virus NS5*. Acta Crystallographica. Section F, Structural Biology Communications, 2017. **73**(Pt 3): p. 116-122.
77. Wang, B., et al., *The structure of Zika virus NS5 reveals a conserved domain conformation*. Nature Communications, 2017. **8**: p. 14763.
78. Saw, W.G., et al., *Structural features of Zika virus non-structural proteins 3 and -5 and its individual domains in solution as well as insights into NS3 inhibition*. Antiviral Research, 2017. **141**: p. 73-90.
79. Chua, J.J., M.M. Ng, and V.T. Chow, *The non-structural 3 (NS3) protein of dengue virus type 2 interacts with human nuclear receptor binding protein and is associated with alterations in membrane structure*. Virus Res, 2004. **102**(2): p. 151-63.
80. Luo, D., et al., *Crystal Structure of the NS3 Protease-Helicase from Dengue Virus*. Journal of Virology, 2008. **82**(1): p. 173-183.
81. Luo, D., et al., *Crystal structure of the NS3 protease-helicase from dengue virus*. J Virol, 2008. **82**(1): p. 173-83.
82. Lescar, J., et al., *Towards the design of antiviral inhibitors against flaviviruses: the case for the multifunctional NS3 protein from Dengue virus as a target*. Antiviral Res, 2008. **80**(2): p. 94-101.
83. Luo, D., S.G. Vasudevan, and J. Lescar, *The flavivirus NS2B-NS3 protease-helicase as a target for antiviral drug development*. Antiviral Research, 2015. **118**: p. 148-158.
84. Gebhard, L.G., S.B. Kaufman, and A.V. Gamarnik, *Novel ATP-Independent RNA Annealing Activity of the Dengue Virus NS3 Helicase*. PLOS ONE, 2012. **7**(4): p. e36244.
85. Xu, T., et al., *Structure of the Dengue Virus Helicase/Nucleoside Triphosphatase Catalytic Domain at a Resolution of 2.4 Å*. Journal of Virology, 2005. **79**(16): p. 10278-10288.
86. Yon, C., et al., *Modulation of the Nucleoside Triphosphatase/RNA Helicase and 5'-RNA Triphosphatase Activities of Dengue Virus Type 2 Nonstructural Protein 3 (NS3) by Interaction with NS5, the RNA-dependent RNA Polymerase*. Journal of Biological Chemistry, 2005. **280**(29): p. 27412-27419.
87. Ramanathan, M.P., et al., *Host cell killing by the West Nile Virus NS2B-NS3 proteolytic complex: NS3 alone is sufficient to recruit caspase-8-based apoptotic pathway*. Virology, 2006. **345**(1): p. 56-72.
88. Shafee, N. and S. AbuBakar, *Dengue virus type 2 NS3 protease and NS2B-NS3 protease precursor induce apoptosis*. J Gen Virol, 2003. **84**(Pt 8): p. 2191-5.
89. Prikhod'ko, E.A., et al., *The NS3 protein of hepatitis C virus induces caspase-8-mediated apoptosis independent of its protease or helicase activities*. Virology, 2004. **329**(1): p. 53-67.
90. Heaton, N.S., et al., *Dengue virus nonstructural protein 3 redistributes fatty acid synthase to sites of viral replication and increases cellular fatty acid synthesis*. Proceedings of the National Academy of Sciences of the United States of America, 2010. **107**(40): p. 17345-17350.
91. Tang, W.-C., et al., *Rab18 Facilitates Dengue Virus Infection by Targeting Fatty Acid Synthase to Sites of Viral Replication*. Journal of Virology, 2014. **88**(12): p. 6793-6804.

92. Luo, D., et al., *Flexibility between the Protease and Helicase Domains of the Dengue Virus NS3 Protein Conferred by the Linker Region and Its Functional Implications*. The Journal of Biological Chemistry, 2010. **285**(24): p. 18817-18827.
93. Erbel, P., et al., *Structural basis for the activation of flaviviral NS3 proteases from dengue and West Nile virus*. Nat Struct Mol Biol, 2006. **13**(4): p. 372-3.
94. Oliveira, A.S.d., et al., *NS3 and NS5 proteins: important targets for anti-dengue drug design*. Journal of the Brazilian Chemical Society, 2014. **25**: p. 1759-1769.
95. Noble, C.G., et al., *Strategies for development of dengue virus inhibitors*. Antiviral Research, 2010. **85**(3): p. 450-462.
96. Erbel, P., et al., *Structural basis for the activation of flaviviral NS3 proteases from dengue and West Nile virus*. Nat Struct Mol Biol, 2006. **13**(4): p. 372-373.
97. Kim, Y.M., et al., *NMR Analysis of a Novel Enzymatically Active Unlinked Dengue NS2B-NS3 Protease Complex*. The Journal of Biological Chemistry, 2013. **288**(18): p. 12891-12900.
98. Chen, W.N., et al., *The dengue virus NS2B-NS3 protease retains the closed conformation in the complex with BPTI*. FEBS Lett, 2014. **588**(14): p. 2206-11.
99. Ye, J., et al., *RecA-like motor ATPases—lessons from structures*. Biochimica et Biophysica Acta (BBA) - Bioenergetics, 2004. **1659**(1): p. 1-18.
100. Enemark, E.J. and L. Joshua-Tor, *On Helicases and other motor proteins*. Current opinion in structural biology, 2008. **18**(2): p. 243-257.
101. Wu, J., et al., *Structure of the Flavivirus Helicase: Implications for Catalytic Activity, Protein Interactions, and Proteolytic Processing*. Journal of Virology, 2005. **79**(16): p. 10268-10277.
102. Xu, T., et al., *Towards the design of flavivirus helicase/NTPase inhibitors: crystallographic and mutagenesis studies of the dengue virus NS3 helicase catalytic domain*. Novartis Found Symp, 2006. **277**: p. 87-97; discussion 97-101, 251-3.
103. Wang, C.C., et al., *Analysis of the nucleoside triphosphatase, RNA triphosphatase, and unwinding activities of the helicase domain of dengue virus NS3 protein*. FEBS Lett, 2009. **583**(4): p. 691-6.
104. Imlay, J.A., *The molecular mechanisms and physiological consequences of oxidative stress: lessons from a model bacterium*. Nature reviews. Microbiology, 2013. **11**(7): p. 443-454.
105. Turrens, J.F., *Mitochondrial formation of reactive oxygen species*. The Journal of Physiology, 2003. **552**(Pt 2): p. 335-344.
106. Sharma, P., et al., *Reactive Oxygen Species, Oxidative Damage, and Antioxidative Defense Mechanism in Plants under Stressful Conditions*. Journal of Botany, 2012. **2012**: p. 26.
107. Cabiscol, E., et al., *Oxidative stress promotes specific protein damage in Saccharomyces cerevisiae*. Journal of Biological Chemistry, 2000.
108. Mishra, S. and J. Imlay, *Why do bacteria use so many enzymes to scavenge hydrogen peroxide?* Archives of biochemistry and biophysics, 2012. **525**(2): p. 145-160.
109. Sies, H., *Strategies of antioxidant defense*. European Journal of Biochemistry, 1993. **215**(2): p. 213-219.
110. Covarrubias, L., et al., *Function of reactive oxygen species during animal development: Passive or active?* Developmental Biology, 2008. **320**(1): p. 1-11.
111. Lu, J. and A. Holmgren, *The thioredoxin antioxidant system*. Free Radical Biology and Medicine, 2014. **66**: p. 75-87.
112. Dip, P.V., et al., *Key roles of the Escherichia coli AhpC C-terminus in assembly and catalysis of alkylhydroperoxide reductase, an enzyme essential for the alleviation of oxidative stress*. Biochimica et Biophysica Acta (BBA) - Bioenergetics, 2014. **1837**(12): p. 1932-1943.
113. Jacobson, F.S., et al., *An alkyl hydroperoxide reductase from Salmonella typhimurium involved in the defense of DNA against oxidative damage. Purification and properties*. Journal of Biological Chemistry, 1989. **264**(3): p. 1488-96.
114. Poole, L.B. and H.R. Ellis, *Flavin-dependent alkyl hydroperoxide reductase from Salmonella typhimurium. 1. Purification and enzymatic activities of overexpressed AhpF and AhpC proteins*. Biochemistry, 1996. **35**(1): p. 56-64.
115. Chae, H.Z., S.J. Chung, and S.G. Rhee, *Thioredoxin-dependent peroxide reductase from yeast*. Journal of Biological Chemistry, 1994. **269**(44): p. 27670-27678.

116. Dip, P.V., et al., *Structure, mechanism and ensemble formation of the alkylhydroperoxide reductase subunits AhpC and AhpF from Escherichia coli*. Acta Crystallographica Section D, 2014. **70**(11): p. 2848-2862.
117. Roberts, B.R., et al., *Oxidized and synchrotron cleaved structures of the disulfide redox center in the N-terminal domain of Salmonella typhimurium AhpF*. Protein Science : A Publication of the Protein Society, 2005. **14**(9): p. 2414-2420.
118. Wood, Z.A., et al., *Structure, mechanism and regulation of peroxiredoxins*. Trends in Biochemical Sciences. **28**(1): p. 32-40.
119. Bieger, B. and L.-O. Essen, *Crystal structure of the catalytic core component of the alkylhydroperoxide reductase AhpF from Escherichia coli* Edited by K. Nagai. Journal of Molecular Biology, 2001. **307**(1): p. 1-8.
120. Li Calzi, M. and L.B. Poole, *Requirement for the Two AhpF Cystine Disulfide Centers in Catalysis of Peroxide Reduction by Alkyl Hydroperoxide Reductase*. Biochemistry, 1997. **36**(43): p. 13357-13364.
121. Poole, L.B., et al., *AhpF can be dissected into two functional units: tandem repeats of two thioredoxin-like folds in the N-terminus mediate electron transfer from the thioredoxin reductase-like C-terminus to AhpC*. Biochemistry, 2000. **39**(22): p. 6602-15.
122. Reynolds, C.M. and L.B. Poole, *Attachment of the N-terminal domain of Salmonella typhimurium AhpF to Escherichia coli thioredoxin reductase confers AhpC reductase activity but does not affect thioredoxin reductase activity*. Biochemistry, 2000. **39**(30): p. 8859-69.
123. Kamariah, N., et al., *Transition steps in peroxide reduction and a molecular switch for peroxide robustness of prokaryotic peroxiredoxins*. 2016. **6**: p. 37610.
124. Rhee, A., R. Cheong, and A. Levchenko, *The application of information theory to biochemical signaling systems*. Physical biology, 2012. **9**(4): p. 10.1088/1478-3975/9/4/045011.
125. Delaunay, A., et al., *A Thiol Peroxidase Is an H₂O₂ Receptor and Redox-Transducer in Gene Activation*. Cell. **111**(4): p. 471-481.
126. Bryk, R., P. Griffin, and C. Nathan, *Peroxyntirite reductase activity of bacterial peroxiredoxins*. Nature, 2000. **407**: p. 211.
127. Jang, H.H., et al., *Two Enzymes in One*. Cell. **117**(5): p. 625-635.
128. Zito, E., et al., *Oxidative protein folding by an endoplasmic reticulum localized peroxiredoxin*. Molecular cell, 2010. **40**(5): p. 787-797.
129. Tairum, C.A., et al., *Catalytic Thr or Ser Residue Modulates Structural Switches in 2-Cys Peroxiredoxin by Distinct Mechanisms*. 2016. **6**: p. 33133.
130. Hall, A., P.A. Karplus, and L.B. Poole, *Typical 2-Cys Peroxiredoxins: Structures, mechanisms and functions*. The FEBS journal, 2009. **276**(9): p. 2469-2477.
131. Wood, Z.A., et al., *Dimers to doughnuts: redox-sensitive oligomerization of 2-cysteine peroxiredoxins*. Biochemistry, 2002. **41**(17): p. 5493-504.
132. Wong, C.F., et al., *AhpC of the mycobacterial antioxidant defense system and its interaction with its reducing partner Thioredoxin-C*. Scientific Reports, 2017. **7**: p. 5159.
133. Parsonage, D., et al., *Analysis of the link between enzymatic activity and oligomeric state in AhpC, a bacterial peroxiredoxin*. Biochemistry, 2005. **44**(31): p. 10583-92.
134. Aran, M., et al., *ATP-dependent modulation and autophosphorylation of rapeseed 2-Cys peroxiredoxin*. FEBS Journal, 2008. **275**(7): p. 1450-1463.
135. Kitano, K., et al., *Stimulation of peroxidase Activity by Decamerization Related to Ionic Strength: ApPC Protein from *Amphibacillus xylanus**. The Journal of Biochemistry, 1999. **126**(2): p. 313-319.
136. König, J., et al., *The plant-specific function of 2-Cys peroxiredoxin-mediated detoxification of peroxides in the redox-hierarchy of photosynthetic electron flux*. Proceedings of the National Academy of Sciences, 2002. **99**(8): p. 5738-5743.
137. Nartey, W., et al., *NMR studies reveal a novel grab and release mechanism for efficient catalysis of the bacterial 2-Cys peroxiredoxin machinery*. FEBS Journal, 2015. **282**(23): p. 4620-4638.
138. Kamariah, N., et al., *Low resolution solution structure of an enzymatic active AhpC10:AhpF2 ensemble of the Escherichia coli Alkyl hydroperoxide Reductase*. Journal of Structural Biology, 2016. **193**(1): p. 13-22.

139. Kamariah, N., et al., *Essential role of the flexible linker on the conformational equilibrium of bacterial peroxiredoxin reductase for effective regeneration of peroxiredoxin*. Journal of Biological Chemistry, 2017. **292**(16): p. 6667-6679.
140. Kamariah, N., et al., *Crystallographic and solution studies of NAD⁺- and NADH-bound alkylhydroperoxide reductase subunit F (AhpF) from Escherichia coli provide insight into sequential enzymatic steps*. Biochimica et Biophysica Acta (BBA) - Bioenergetics, 2015. **1847**(10): p. 1139-1152.
141. Jönsson, T.J., et al., *Reduction of Cysteine Sulfinic Acid in Peroxiredoxin by Sulfiredoxin Proceeds Directly through a Sulfinic Phosphoryl Ester Intermediate*. The Journal of Biological Chemistry, 2008. **283**(35): p. 23846-23851.
142. Conway Myra, E. and C. Lee, *The redox switch that regulates molecular chaperones*, in *Biomolecular Concepts*. 2015. p. 269.
143. An, B.C., et al., *An additional cysteine in a typical 2-Cys peroxiredoxin of Pseudomonas promotes functional switching between peroxidase and molecular chaperone*. FEBS Letters, 2015. **589**(19, Part B): p. 2831-2840.
144. Murray, B.E., *The life and times of the Enterococcus*. Clinical Microbiology Reviews, 1990. **3**(1): p. 46-65.
145. Gilmore, M.S., F. Lebreton, and W. van Schaik, *Genomic Transition of Enterococci from Gut Commensals to Leading Causes of Multidrug-resistant Hospital Infection in the Antibiotic Era*. Current opinion in microbiology, 2013. **16**(1): p. 10-16.
146. Stuart, C.H., et al., *Enterococcus faecalis: Its Role in Root Canal Treatment Failure and Current Concepts in Retreatment*. Journal of Endodontics. **32**(2): p. 93-98.
147. Gonzales, R.D., et al., *Infections due to vancomycin-resistant Enterococcus faecium resistant to linezolid*. The Lancet, 2001. **357**(9263): p. 1179.
148. Paulsen, I.T., et al., *Role of Mobile DNA in the Evolution of Vancomycin-Resistant Enterococcus faecalis*. Science, 2003. **299**(5615): p. 2071-2074.
149. Nathan, C.F., *Mechanisms of macrophage antimicrobial activity*. Transactions of The Royal Society of Tropical Medicine and Hygiene, 1983. **77**(5): p. 620-630.
150. Thomas, E.L., R.I. Lehrer, and R.F. Rest, *Human Neutrophil Antimicrobial Activity*. Reviews of Infectious Diseases, 1988. **10**: p. S450-S456.
151. Riboulet, E., et al., *Relationships between Oxidative Stress Response and Virulence in Enterococcus faecalis*. Journal of Molecular Microbiology and Biotechnology, 2007. **13**(1-3): p. 140-146.
152. La Carbona, S., et al., *Comparative study of the physiological roles of three peroxidases (NADH peroxidase, Alkyl hydroperoxide reductase and Thiol peroxidase) in oxidative stress response, survival inside macrophages and virulence of Enterococcus faecalis*. Molecular Microbiology, 2007. **66**(5): p. 1148-1163.
153. Toh, Y.K., et al., *Novel insights into the vancomycin-resistant Enterococcus faecalis (V583) alkylhydroperoxide reductase subunit F*. Biochimica et Biophysica Acta (BBA) - General Subjects.
154. Grüber, G., et al., *Expression, purification, and characterization of subunit E, an essential subunit of the vacuolar ATPase*. Biochemical and Biophysical Research Communications, 2002. **298**(3): p. 383-391.
155. Emsley, P. and K. Cowtan, *Coot: model-building tools for molecular graphics*. Acta Crystallogr D Biol Crystallogr, 2004. **60**(Pt 12 Pt 1): p. 2126-32.
156. Zbyszek, O. and M. Wladek, *Processing of X-ray diffraction data collected in oscillation mode*. Meth. Enzymol., 1997. **276**: p. 307-326.
157. McCoy, A.J., et al., *Phaser crystallographic software*. Journal of Applied Crystallography, 2007. **40**(Pt 4): p. 658-674.
158. Laskowski, R.A., et al., *PROCHECK: a program to check the stereochemical quality of protein structures*. Journal of Applied Crystallography, 1993. **26**(2): p. 283-291.
159. DeLano, W., *The PyMOL Molecular Graphics System*. DeLano Scientific, 2002.
160. Ramachandran, G.N., C. Ramakrishnan, and V. Sasisekharan, *Stereochemistry of polypeptide chain configurations*. J Mol Biol, 1963. **7**: p. 95-9.

-
161. Murshudov, G.N., A.A. Vagin, and E.J. Dodson, *Refinement of macromolecular structures by the maximum-likelihood method*. Acta Crystallogr D Biol Crystallogr, 1997. **53**(Pt 3): p. 240-55.
162. Pettersen, E.F., et al., *UCSF Chimera--a visualization system for exploratory research and analysis*. J Comput Chem, 2004. **25**(13): p. 1605-12.
163. Konarev, P.V., et al., *PRIMUS: a Windows PC-based system for small-angle scattering data analysis*. Journal of Applied Crystallography, 2003. **36**(5): p. 1277-1282.
164. Putnam, C.D., et al., *X-ray solution scattering (SAXS) combined with crystallography and computation: defining accurate macromolecular structures, conformations and assemblies in solution*. Quarterly Reviews of Biophysics, 2007. **40**(3): p. 191-285.
165. Jacques, D.A. and J. Trehella, *Small-angle scattering for structural biology—Expanding the frontier while avoiding the pitfalls*. Protein Science : A Publication of the Protein Society, 2010. **19**(4): p. 642-657.
166. Guinier., A., *Diffraction of X-rays of very small angles-application to the study of ultramicroscopic phenomenon*. Ann. Phys., 1939. **12**: p. 161-237.
167. Guinier, A.a.F., G., *Small angle scattering of X-rays*. Wiley, New York, 1995.
168. Petoukhov, M.V., et al., *ATSAS 2.1 - towards automated and web-supported small-angle scattering data analysis*. Journal of Applied Crystallography, 2007. **40**(s1): p. s223-s228.
169. Svergun, D., *Determination of the regularization parameter in indirect-transform methods using perceptual criteria*. Journal of Applied Crystallography, 1992. **25**(4): p. 495-503.
170. Glatter, O., *A new method for the evaluation of small-angle scattering data*. Journal of Applied Crystallography, 1977. **10**(5): p. 415-421.
171. Durand, D., et al., *NADPH oxidase activator p67phox behaves in solution as a multidomain protein with semi-flexible linkers*. Journal of Structural Biology, 2010. **169**(1): p. 45-53.
172. Reyes, F.E., Schwartz, C.R., Tainer, T.A., and Rambo, R.P., *Chapter 11- Methods for using new conceptual tools and parameters to assess RNA structure by small-angle X-ray scattering. in Methods Enzymol. (Donald, H.B. -A. ed.)*. Academic Press, 2014: p. 235-263.
173. Svergun, D.I., *Restoring low resolution structure of biological macromolecules from solution scattering using simulated annealing*. Biophys J., 1999: p. 2879-2886.
174. Franke, D. and D.I. Svergun, *DAMMIF, a program for rapid ab-initio shape determination in small-angle scattering*. Journal of Applied Crystallography, 2009. **42**(2): p. 342-346.
175. Volkov, V.V. and D.I. Svergun, *Uniqueness of ab initio shape determination in small-angle scattering*. Journal of Applied Crystallography, 2003. **36**(3 Part 1): p. 860-864.
176. Petoukhov, M.V., Franke, D., Shkumatov, A.V., Tria, G., Kikhney, A.G., Gajda, M., Gorba, C., Mertens, H.D.T., Konarev, P.V. and Svergun, D.I., *New developments in the ATSAS program package for small-angle scattering data analysis*. J. Appl. Cryst., 2012. **45**: p. 342-350.
177. Petoukhov, M.V. and D.I. Svergun, *Global Rigid Body Modeling of Macromolecular Complexes against Small-Angle Scattering Data*. Biophysical Journal, 2005. **89**(2): p. 1237-1250.
178. Svergun D.I., B.C.a.K.M.H.J., *CRY SOL - a Program to Evaluate X-ray Solution Scattering of Biological Macromolecules from Atomic Coordinates*. J. Appl. Cryst. , 1995. **28**(768-773).
179. Kozin, M.B. and D.I. Svergun, *Automated matching of high- and low-resolution structural models*. Journal of Applied Crystallography, 2001. **34**(1): p. 33-41.
180. Bernadó, P., et al., *Structural Characterization of Flexible Proteins Using Small-Angle X-ray Scattering*. Journal of the American Chemical Society, 2007. **129**(17): p. 5656-5664.
181. Tria, G., et al., *Advanced ensemble modelling of flexible macromolecules using X-ray solution scattering*. IUCrJ, 2015. **2**(Pt 2): p. 207-217.
182. Tang, G., et al., *EMAN2: An extensible image processing suite for electron microscopy*. Journal of Structural Biology, 2007. **157**(1): p. 38-46.
183. Scheres, S.H.W., *RELION: Implementation of a Bayesian approach to cryo-EM structure determination*. Journal of Structural Biology, 2012. **180**(3): p. 519-530.
184. Emsley, P. and K. Cowtan, *Coot: model-building tools for molecular graphics*. Acta Crystallographica Section D, 2004. **60**(12 Part 1): p. 2126-2132.
185. Davis, I.W., et al., *MolProbity: all-atom contacts and structure validation for proteins and nucleic acids*. Nucleic Acids Research, 2007. **35**(Web Server issue): p. W375-W383.

-
186. Krissinel, E. and K. Henrick, *Secondary-structure matching (SSM), a new tool for fast protein structure alignment in three dimensions*. Acta Crystallographica Section D, 2004. **60**(12 Part 1): p. 2256-2268.
187. Waterhouse, A.M., et al., *Jalview Version 2—a multiple sequence alignment editor and analysis workbench*. Bioinformatics, 2009. **25**(9): p. 1189-1191.
188. *UniProt: the universal protein knowledgebase*. Nucleic Acids Research, 2017. **45**(D1): p. D158-D169.
189. Sievers, F. and D.G. Higgins, *Clustal Omega, Accurate Alignment of Very Large Numbers of Sequences*, in *Multiple Sequence Alignment Methods*, D.J. Russell, Editor. 2014, Humana Press: Totowa, NJ. p. 105-116.
190. Lei, J., et al., *Crystal structure of Zika virus NS2B-NS3 protease in complex with a boronate inhibitor*. Science, 2016. **353**(6298): p. 503-5.
191. Tian, H., et al., *Structural basis of Zika virus helicase in recognizing its substrates*. Protein Cell, 2016. **7**(8): p. 562-70.
192. Falgout, B., et al., *Both nonstructural proteins NS2B and NS3 are required for the proteolytic processing of dengue virus nonstructural proteins*. J Virol, 1991. **65**(5): p. 2467-75.
193. Johansson, M., et al., *A small region of the dengue virus-encoded RNA-dependent RNA polymerase, NS5, confers interaction with both the nuclear transport receptor importin-beta and the viral helicase, NS3*. J Gen Virol, 2001. **82**(Pt 4): p. 735-45.
194. Tay, M.Y., et al., *The C-terminal 50 Amino Acid Residues of Dengue NS3 Protein Are Important for NS3-NS5 Interaction and Viral Replication*. J Biol Chem, 2015. **290**(4): p. 2379-94.
195. Preugschat, F., C.W. Yao, and J.H. Strauss, *In vitro processing of dengue virus type 2 nonstructural proteins NS2A, NS2B, and NS3*. J Virol, 1990. **64**(9): p. 4364-74.
196. Kohlway, A., et al., *The Linker Region of NS3 Plays a Critical Role in the Replication and Infectivity of Hepatitis C Virus*. Journal of Virology, 2014. **88**(18): p. 10970-10974.
197. Yao, N., et al., *Molecular views of viral polyprotein processing revealed by the crystal structure of the hepatitis C virus bifunctional protease-helicase*. Structure, 1999. **7**(11): p. 1353-63.
198. Li, L., et al., *Structure-guided Discovery of a Novel Non-peptide Inhibitor of Dengue Virus NS2B-NS3 Protease*. Chem Biol Drug Des, 2014.
199. de Sousa, L.R., et al., *Flavonoids as noncompetitive inhibitors of Dengue virus NS2B-NS3 protease: inhibition kinetics and docking studies*. Bioorg Med Chem, 2015. **23**(3): p. 466-70.
200. Luo, D., et al., *Flexibility between the protease and helicase domains of the Dengue virus NS3 protein conferred by the linker region and its functional implications*. Journal of Biological Chemistry, 2010. **285**(24): p. 18817-18827.
201. Han, Y.-S., et al., *Identification of resveratrol analogs as potent anti-dengue agents using a cell-based assay*. Journal of Medical Virology, 2017. **89**(3): p. 397-407.
202. Abba, Y., et al., *Antiviral Activity of Resveratrol against Human and Animal Viruses*. Advances in Virology, 2015. **2015**: p. 184241.
203. Campagna, M. and C. Rivas, *Antiviral activity of resveratrol*. Biochemical Society Transactions, 2010. **38**(1): p. 50-53.
204. Zandi, K., et al., *Antiviral activity of four types of bioflavonoid against dengue virus type-2*. Virology Journal, 2011. **8**(1): p. 1-11.
205. Diederichs, K. and P.A. Karplus, *Better models by discarding data?* Acta Crystallographica Section D: Biological Crystallography, 2013. **69**(Pt 7): p. 1215-1222.
206. Adams, P.D., et al., *PHENIX: a comprehensive Python-based system for macromolecular structure solution*. Acta Crystallographica Section D: Biological Crystallography, 2010. **66**(Pt 2): p. 213-221.
207. Matthews, B.W., *Solvent content of protein crystals*. Journal of Molecular Biology, 1968. **33**(2): p. 491-497.
208. Morais, M.A.B., et al., *How pH Modulates the Dimer-Decamer Interconversion of 2-Cys Peroxiredoxins from the Prx1 Subfamily*. The Journal of Biological Chemistry, 2015. **290**(13): p. 8582-8590.
209. Chuang, M.-H., et al., *Proteomic analysis of proteins expressed by Helicobacter pylori under oxidative stress*. PROTEOMICS, 2005. **5**(15): p. 3895-3901.

-
210. Liebthal, M., et al., *Redox-Dependent Conformational Dynamics of Decameric 2-Cysteine Peroxiredoxin and its Interaction with Cyclophilin 20-3*. Plant Cell Physiol, 2016. **57**(7): p. 1415-1425.
 211. Dietz, K.-J., *Peroxiredoxins in Plants and Cyanobacteria*. Antioxidants & Redox Signaling, 2011. **15**(4): p. 1129-1159.
 212. Perkins, A., et al., *Peroxiredoxins: Guardians Against Oxidative Stress and Modulators of Peroxide Signaling*. Trends in biochemical sciences, 2015. **40**(8): p. 435-445.
 213. Wood, Z.A., L.B. Poole, and P.A. Karplus, *Peroxiredoxin evolution and the regulation of hydrogen peroxide signaling*. Science, 2003. **300**(5619): p. 650-3.



Author's publication related to this project

1. Saw, W.G.¹, **Pan, A.¹**, Manimekalai, M.S.S., Grüber, A., and Grüber, G. (2018) Structural insight and flexibility features of NS3 and NS5 proteins of Dengue and Zika virus in solution. *Progress in Biophysics and Molecular Biology*. (Manuscript to be submitted)
2. Guanhua Zhu, Ankita Pan, Gerhard Grüber, Lanyuan Lu. (2018) Conformational states of Zika virus Non-Structural proteins 3 determined by molecular dynamics simulations. *Progress in Biophysics and Molecular Biology*. (Manuscript to be submitted)
3. **Pan, A.**, Balakrishna, A.M., Nartey, W., Kohlmeier A., Dip, P.V., Bhushan, S., Grüber, G. (2017) Atomic structure and enzymatic insights into the vancomycin-resistant *Enterococcus faecalis* (V583) alkylhydroperoxide reductase subunit C. *Free Rad. Biol. & Medi.* **115**, 252-265. (Impact factor: 5.6)
4. **Pan, A.¹**, Saw, W.G.¹, Manimekalai, M.S.S., Grüber, A., Shin, J., Matsui, T., Weiss, T. and Grüber, G. (2017) Structural features of NS3 of Dengue virus serotypes 2 and 4 in solution and insight into RNA binding and the inhibitory role of quercetin. *Acta Crystallogr.* **D73**, 402-419. (Impact factor: 2.1)
5. Saw, W.G.¹, **Pan, A.¹**, Manimekalai, M.S.S. and Grüber, G. (2017) Structural features of Zika virus non-structural proteins 3 and -5 and its individual domains in solution as well as insights into NS3 inhibition. *Antiviral Res.* **141**, 73-90. (Impact factor: 4.2)
6. Manimekalai, M.S.S., Saw, W.G., **Pan, A.**, Grüber, A. and Grüber, G. (2016) Identification of the critical linker residues conferring differences in compactness of DENV-4 NS5 from Dengue virus serotypes 1-3. *Acta Crystallogr.* **D72**, 795-807. (Impact factor: 2.1)

(¹- authors contributed equally)

Posters and abstracts

1. **Pan, A.**, Saw, W.G., Manimekalai, M.S.S., Grüber, A., and Grüber, G. (2017) “Small angle X-ray scattering studies of flaviviral NS3”. 24th Congress and General Assembly of the International Union of Crystallography
2. **Pan, A.**, Saw, W.G., Manimekalai, M.S.S., Grüber, A., and Grüber, G. (2017) “Solution studies of DENV-2 NS3”. EMBO courses and workshops: structural and biophysical methods for biological macromolecules in solution.

Award

IUCr Journal Poster Prize in Instrumentation and Methods for the poster “Small angle X-ray scattering studies of flaviviral NS3” at the 24th Congress and General Assembly of the International Union of Crystallography