Characterisation of the Influenza A Virus Matrix Protein M1 and Its Role in Oligomerisation



Muhd Faiz-Hafiz Mohd Kipli

Hertford College

University of Oxford

A thesis submitted for the degree of

Doctor of Philosophy

Michaelmas 2020

Acknowledgements

Personal

First and foremost, I would like to say thank you and express my utmost gratitude to my supervisor, Jason Schnell. Without his guidance, I would neither be able to complete my DPhil project nor be able to publish a paper on my graduate work. Thank you for the unrelenting support that you have given me throughout my DPhil years. Your knowledge and encouragement were pivotal in completing the work described in this thesis. I am grateful to be taught by an expert who is very passionate about NMR and proteins.

Many thanks to Jolyon Claridge for supervising me in my first year of DPhil. Thank you for your insights and support. Your supervision has helped shape my experimental and analysis skills. Your optimism towards M1 has motivated me to continue working with M1.

Special thanks to my mentee, Alex Jiang. Even though your stay was for a short term, your work has helped me immensely in the experiments presented here and my development as a mentor. Your presence in the lab fills up the room with joy and laughter and certainly makes working with M1 exciting. Thanks for the good company on random nights out and especially for being a good friend.

Thank you, as well, to the Schnell lab, particularly to Anu Chandran and Felipe Ossa for fostering such a positive lab environment. I am glad to be working alongside you both.

I am also indebted to Amy Chu, Jodie Ford, Hannah Almutaqqi, Edwin Lock and Daniel Wells for proofreading this manuscript.

To Amy Chu, Celine Zhang, Nick Antypas, Sam Gérard - thanks for the excellent company and the nights of fun board games. To James Coates, thanks for making my Oxford life that much more colourful.

To my friends at Oxford: my experience being at Oxford has been amazing. I am grateful for the comradeship we have shared. To all members of OUGYM, Oxford Sirens and OU Pole: thanks for making my Oxford experience much fuller and transformative. I would never have thought to be good at sports, let alone win championships.

Special mention to Mollie and Emma, who both always remind me not to take life too seriously. I will forever cherish our friendship.

To my parents, I will be forever indebted to you both. Without your support and love, I would not be where I am today. I dedicate this DPhil to you. To my beloved sister, Rena and my brother, David, thank you for loving me unconditionally and for accepting me for who I am. You both have played a significant role in finishing this DPhil. To my sisters, thank you for the support you have given me throughout these years.

And finally, to Leon, you are the true wonder of my time in Oxford, thank you for your love and support. Before I met you, my future was uncertain. But you helped me through the hard times and supported me when I felt like giving up. Thank you for believing in me and for always pushing me in the right direction towards finishing this thesis and also in life.

Abstract

Influenza matrix protein 1 (M1) is the most abundant protein in the virus and forms the matrix layer beneath the lipid bilayer of the viral capsid, stabilising and providing shape to the three-dimensional structure of the virus particles. As a matrix protein, it is crucial for M1 to be able to associate and dissociate at the right stages of the influenza A virus life cycle. M1 polymerises upon binding to membranes. However, the physical triggers that covert M1 from a water-soluble component in the nucleus to an immobilised lattice of helical polymer on the membrane surface are still poorly understood. M1 oligomerisation correlates with pH, assembling at a high pH and disassembling at a low pH, which is driven primarily by the N-terminal domain (M1NT). In this thesis, the structural plasticity and oligomerisation of M1NT in solution was studied using NMR spectroscopy. We showed that M1NT was pertubed by sterol-containing compounds which induce a conformational change facilitating self-association in a pH dependent manner. The perturbed residues were consistent with the stacked-dimer interface, which are exposed at the surface and most sensitive to sterols. The effects of sterols on this interface is indirect and they most likely exert their effects through a reduction in water activity. The local changes observed are consistent with a priming of the N-terminal domain for polymerization. The exposed surface of the M1NT interface is hypothesised to be sensitive to changes in its environment and may serve to regulate polymerisation of M1 at the membrane surface. Chemical shift perturbations data and crystal structure models indicate the effects of conformational changes can be propagated from one subdomain interface to another.

Contents

List of Figures viii

List of Tables xi

List of Abbreviations xii

1 Introduction 1

1.1 Burden, Taxonomy and Nomenclature of Influenza A 1

1.2 Influenza Virus Life Cycle 3

1.2.1 Fusion and Nuclear Import 6

1.2.2 Viral Gene Expression 6

1.2.3 Nuclear export of vRNP complexes 7

1.2.4 Genome packaging 7

1.2.5 Viral Assembly and Budding 8

1.3 M1, the Matrix Protein of Influenza A Virus 9

1.3.1 M1 Structure 9

1.3.2 M1 Oligomerisation 12

1.3.3 M1 Scaffold Dissociation with pH 15

1.3.4 M1 Interaction with Lipid 16

1.3.5 M1 of Influenza A Virus and Influenza C Virus 19

1.3.6 M1 of influenza A Virus and Matrix Proteins of Other Enveloped RNA Viruses 20

1.4 Thesis Objectives 21

2 Materials and Methods 22

3 M1 N-Terminal Domain Structure & Dynamics 28

3.1 Expression and Purification of M1 28

3.2 NMR Characterisation of M1NT 29

3.2.1 Backbone Resonance Assignment of M1NT 29

3.2.2 Secondary Structure of M1NT 33

3.3 Dynamics of M1NT 34

3.4 Effects of Changes in pH and Ionic Strength 37

3.4.1 Effects of Changes in pH 37
3.4.2 Effects of Changes in KCl 44

4 M1 C-Terminal Domain 50
4.1 M1 C-terminal Domain Characterisation 50
4.1.1 M1CT Expression and Purification 51
4.1.2 NMR Characterisation of M1CT 53
4.1.3 M1NT with C-terminal Extensions 55

5 M1 Oligomerisation and the Hydrophobic Pocket 63
5.1 Sensitivity of The Hydrophobic Pocket of M1 to Solution Conditions 64
5.1.1 Surfactants 64
5.1.2 pH 70
5.1.3 Isopropyl Alcohol (IPA) 75
5.2 Oligomerisation 77
5.2.1 Cholesteryl Hemisuccinate (CHS) 77
5.2.1.1 CHAPS-solubilised CHS 77
5.2.1.2 LMNG-solubilised CHS 78
5.2.2 25-hydroxycholesterol (25-HC) 78
5.2.3 Mutants Based on 1EA3 Crystal Structure 81
5.2.4 Mutants based on 1AA7 Crystal Structure 88
5.2.5 CHAPS Binding of M1 and Its Mutants 91
5.2.5.1 Chemical Shift Perturbation (CSP) of T67 Resonance of Wild-type and Mutants M1 91
5.2.5.2 Backbone Confirmation of T67 Resonance of M1 Crystal Structures 96
5.2.5.3 Determination of T67 ψ Torsion Angle with CHAPS 97
5.2.6 Kosmotropes and Chaotropes 102
5.2.7 PHE 109

6 M1 interaction with proteins 112
6.1 Nuclear Export Protein (NEP) 112
6.1.1 Introduction 112
6.1.2 Expression and Purification of NEP 115
6.1.3 M1 Interaction with NEP 119
6.2 Tumour Susceptibility Gene 101 (TSG101) 121
6.2.1 Introduction 121
6.2.2 Expression and Purification of TSG101-UEV Domain 126
6.2.3 M1 Interaction with TSG101-UEV Domain 126

7 Discussion **129**
7.1 Solution Study of M1 Structure 131
7.1.1 NMR studies of M1NT and M1CT 131
7.1.2 pH Changes of M1NT 133
7.1.3 K^+ Changes of M1NT 138
7.2 Solution Study of M1 Oligomerisation 140
7.2.1 Oligomerisation of M1 141
7.2.2 Dual Sensing Model of Oligomerisation 145
7.3 Interaction of M1 with Binding Partners - NEP and TSG101 146
7.3.1 NEP 146
7.3.2 TSG 147
7.3.3 NEP and TSG101 148

8 Conclusion **149**
8.1 Conclusion 149

References **152**

List of Figures

1.1 Influenza A Virus : a) vRNA, b) Structure & 3) N-linked Glycan . 3
1.2 Influenza A Virus Life Cycle 5
1.3 Illustration of The Globular Domain of M1NT 10
1.4 M1NT Dimer Arrangements : Side-by-Side & Stacked 14

3.1 Purification of M1 N-terminal Domain - SEC and SDS-PAGE . . . 30
3.2 M1NT - SEC-MALS M1NT 31
3.3 HSQC M1NT at pH 7.0 & pH 5.5 32
3.4 Assigned HSQC M1NT at pH 7.0 33
3.5 Secondary Structure Prediction Derived From TALOS-n 35
3.6 M1NT - XRay vs TALOS-n 35
3.7 M1NT - R_1, R_2 and ^1H-^{15}N Het-NOE 36
3.8 M1NT - NMR Dynamics VS Xray B-factor 37
3.9 M1NT - XRay Crystal Contacts Bfactors 38
3.10 M1NT - Xray Crystal Contacts 39
3.11 M1NT - Changes in pH. 40
3.12 Perturbed Residues Upon Change in pH Mapped to M1NT Structure & The Correlation Between pH and Rotation Across Subdomain. . 41
3.13 M1 - pH Dependence of Intramolecular Subdomain Shifts and Tilting of Stacked Dimer Arrangements. 42
3.14 Structural Alignment of M1NT Stacked Dimer Shift Upon Change in pH 43
3.15 HSQC M1(224) at pH 7.0 - KCl Titration 46
3.16 HSQC M1(224) at pH 5.5 - KCl Titration - CSPs 47
3.17 HSQC M1(224) at pH 5.5 - KCl Titration 48
3.18 HSQC M1(224) at pH 5.5 - KCl Titration - CSPs 49

4.1 Nickel Purification of M1CT - 1 51
4.2 Nickel Purification of M1CT - 2 52
4.3 Ion Exchange Purification of M1CT - 3 52
4.4 HSQC M1CT at pH 5.5 54
4.5 M1 Construct Extension - 1 56

4.6 M1 Construct Extension - 2 57
4.7 HSQC M1 (1-224) at pH 7.0 58
4.8 HSQC M1 (1-232) at pH 5.5 59
4.9 Overlay HSQC of M1 (1-224) & M1 (1-232) at pH 5.5 60
4.10 Overlay HSQC M1(1-224) & M1CT at pH 5.5 61
4.11 Overlay HSQC M1(1-224) at pH 7 and M1CT at pH 5.5 62

5.1 Initial Detergents Screening 65
5.2 M1NT - DDM & AZ314 Titration 66
5.3 M1NT - C4C7CO_2 & LMNG Titration 67
5.4 Sterol-containing Surfactants 68
5.5 M1NT - CHAPS Titration 68
5.6 M1NT - CHAPSO Titration 69
5.7 M1NT - Cholate Titration 70
5.8 M1NT - Deoxycholate Titration 71
5.9 Illustration of the hydrophobic patch on M1NT 71
5.10 M1(224) - CHAPS Titration at pH 7.0 72
5.11 Quantification of the CSPs Induced by Sterol-containing Surfactants 73
5.12 Overlay of pH 7.0 & pH 5.5 - M1NT and M1(1-224) 74
5.13 M1NT - IPA Titration 76
5.14 CHAPS-solubilised CHS Titration - M1NT and M1(1-224) at pH 7 79
5.15 CHAPS solubilised CHS Titration - M1NT and M1(1-224) at pH 5.5 80
5.16 LMNG-solubilised CHS Titration - M1NT pH 7 81
5.17 25-HC Titration - M1NT pH 7 82
5.18 M1NT Dimer (PDB ID: 1EA3) 83
5.19 Hydrodrophobicity of M1NT Dimer (PDB ID: 1EA3) 84
5.20 CHAPS Titration of M1NT-L4A and M1NT-L4K 85
5.21 CHAPS Titration of M1NT-I51D and M1NT-E141A 86
5.22 CHAPS Titration of M1(224)-T67S 87
5.23 CHAPS-solubilised CHS Titration - M1NT-L4A & L4K 89
5.24 CHAPS-solubilised - CHS Titration M1NT-I51D & E141A 90
5.25 CHAPS-solubilised CHS Titration - M1(224)-T67S 91
5.26 M1NT Dimer (PDB ID: 1AA7) 92
5.27 M1NT Dimer (PDB ID: 1AA7) - Close View 93
5.28 CHAPS Titration of M1NT-M93A/D94A 94
5.29 CHAPS-solubilised CHS Titration - M1NT-M93A/D94A 95
5.30 Effects of CHAPS on Different M1NT Constructs 96
5.31 T67 Backbone Confirmation of M1NT Crystal Structures (PDB ID: 1EA3, 1AA7 & 4PUS) 98
5.32 CSP - Wild Type M1NT vs M1NT-L4A 100

5.33 CSP + CSI Difference - M1NT-L4A Upon CHAPS addition 101
5.34 M1NT Structural Alignment of Crystal Structures, 1EA3 & 1AA7 . 102
5.35 Hofmeister Series 103
5.36 Ranking of Kosmotropes & Chaotropes 103
5.37 M1NT - Chaotropes Titration : NaSCN & GdnHCl Titration . . . 104
5.38 M1NT - Kosmotropes Titration : Ammonium Sulfate 105
5.39 M1NT - Glycerol Titration 107
5.40 M1NT - CHAPS Titration per residue 108
5.41 PHE Titration - M1NT & M1(232) 110

6.1 Daisy Chain Model - The Nuclear Export Complex 115
6.2 NEP - Nickel Purification 117
6.3 NEP - Size Exclusion Chromatography (SEC) 117
6.4 NEP - HSQC 118
6.5 NEP - M1 Titration at pH 7.0 122
6.6 NEP - M1/CHAPS Titration at pH 7.0 123
6.7 NEP - M1 Oligomers Titration at pH 7.0 124
6.8 NEP - IE Purification 127
6.9 M1(224) - TSG101 Titratio at pH 7.0 128

7.1 The Difference in the Conformation of R134 and M135 in Crystal Structures - 1AA7 & 1EA3 134
7.2 M1 Oligomerisation Model - Figure 135
7.2 M1 Oligomerisation Model - Caption 136

List of Tables

5.1 Sterol-containing compounds with its associated critical micelle concentration (CMC) with the saturation concentration of M1NT CSPs 102

List of Abbreviations

1-D, 2-D . . . One- or two-dimensional, referring in this thesis to spatial dimensions in an NMR experiment

25-HC 25-hydroxycholesterol

AZ314 . . . . . Anzergent Detergent

C4C7CO$_2$ . . . Calixarene-based Detergent

CD . . . . . Circular Dichroism

CHAPS . . . . 3-((3-cholamidopropyl) dimethylammonio)- 1 - propanesulfonate

CHAPSO . . . 3-([3- Cholamidopropyl] dimethylammonio)-2-hydroxy-1-propanesulfonate

CSP . . . . . Chemical Shift Perturbation

CHS . . . . . Cholesteryl Hemisuccinate

CRM1 . . . . . Chromosomal Maintenance 1

cRNA . . . . . Complementary RNA

DDM . . . . . D-maltopyranoside

ER . . . . . Endoplasmic Reticulum

FT . . . . . Flow Through

HA . . . . . Haemagglutinin

HSQC . . . . . 2D ^{1}H-^{15}N Hetero Single Quantum Coherence

IE . . . . . Ion Exchange Chromatography

pI . . . . . Isoelectric Point

IPA . . . . . Isopropyl Alcohol

kb . . . . . Kilobases

LMNG . . . . Lauryl Maltose-Neopentyl Glycol

M1 . . . . . Matrix Protein 1

M2 . . . . . Matrix Protein 2 (Proton Channel)

M1CT . . . . . M1 C-terminal domain

M1NT . . . . . M1 N-terminal domain

MTOC . . . . Microtubule Organising Centre

NA . . . . . Neuraminidase

NEP . . . . . Nuclear Export Protein

NMR . . . . . Nuclear Magnetic Resonance

NLS . . . . . Nuclear Localisation Signal

NP . . . . . Nucleoprotein

PA . . . . . Polymerase Acidic Protein

PB1 . . . . . Polymerase Basic Protein 1

PB2 . . . . . Polymerase Basic Protein 2

SAN . . . . . Small-angle Neutron Scattering

SEC . . . . . Size-exclusion Chromatography

SPR . . . . . Surface Plasmon Resonance

vRNA . . . . . Viral RNA

vRNP . . . . . Viral Ribonucleoprotein

WHO . . . . . World Health Organisation

The reward of the young scientist is the emotional thrill of being the first person in the history of the world to see something or to understand something. Nothing can compare with that experience.

— Cecilia Payne-Gaposchkin

1 Introduction

1.1 Burden, Taxonomy and Nomenclature of Influenza A

Influenza A virus infections can cause severe burden, both in terms of the clinical and economical impact globally (Reed et al., 2015). The new estimate of the number of global deaths associated with influenza illnesses by the World Health Organisation (WHO) has increased from 250,000-500,000 to 290,000-650,000 (Iuliano et al., 2018). The new estimates only account for influenza-related respiratory death as opposed to the previous calculation, which includes respiratory and other deaths such as cardiovascular. The actual estimate of the burden caused by influenza is highly likely to be higher than the new estimate. A recent study suggested that non-respiratory deaths account for >50% of all influenza-related deaths (Muscatello et al., 2014; Redlberger-Fritz et al., 2012). In the United States alone, influenza-associated illnesses count from 9.2 million to 35.6 million illnesses in a single season (Rolfes et al., 2018).

The influenza virus belongs to the lipid-enveloped Orthomyxoviridae family. This is a family in which the virus is composed of negative-sense single-stranded RNA, and the genomes are segmented. Within this family of Orthomyxoviridae viruses, there are three main genera - Influenza A, B, C & D viruses. The reassortment of

genetic materials between the genome of influenza viruses of the same genera is common. This is because the genomes of these viruses are segmented, rendering it easier to exchange genetic information. Between these three genera of influenza, they infect a range of hosts. Influenza B viruses infect human, seals, and ferrets with humans being the main host. Meanwhile, influenza C virus has been isolated from humans, dogs and swine. Lastly, Influenza D infects swine and cattle. Among the three genera, influenza A virus can infect range of hosts including human, seals, wild birds, whales, swine, horses and chickens (Neumann and Kawaoka, 2015). Influenza C & D is relatively rare in comparison to A/B.

The genus of influenza A virus can be grouped into classes of serological subtypes. These subtypes are dependent on the viral haemagglutinin (HA), and neuraminidase (NA) glycoproteins. To date, there are known to be 18 HA and 11 NA subtypes. They are named according to and in order of: genera, host species, the geographical location of isolation, strain identification number, year of isolation and the HA and NA subtypes. The morphology that influenza A virus forms are pleomorphic and can be arranged into spherical, filamentous and bacilliform morphologies. They range from around 80 to 260nm in diameter.

Influenza virus genomes are small, approximately 13.6 kilobases (kb) and the genome is segmented into sizes of about 0.9 to 2.3 kb. The genome of the influenza A virus is made of 8 viral ribonucleoprotein (vRNP) complexes (Figure 1.1a & b). Each vRNP is made of a segment of the negative sense viral RNA (vRNA) packaged around the viral nucleoprotein (NP) and viral RNA trimeric polymeric complex – polymerase acidic protein (PA), polymerase basic protein 1 (PB1) and polymerase basic protein 2 (PB2) (Figure 1.1b). Influenza viruses also encode auxiliary proteins that have diverse functions during infection. This includes Nuclear Export Protein (NEP)/NS2, NS-1, and others. vRNPs are associated with the NEP and are enclosed by the matrix protein M1.

M1 is the most abundant protein in the virus and forms the matrix layer beneath the lipid bilayer of the viral capsid stabilising the three-dimensional structure of the virus (Fujiyoshi et al., 1994; Ruigrok et al., 1989). Additionally, it is

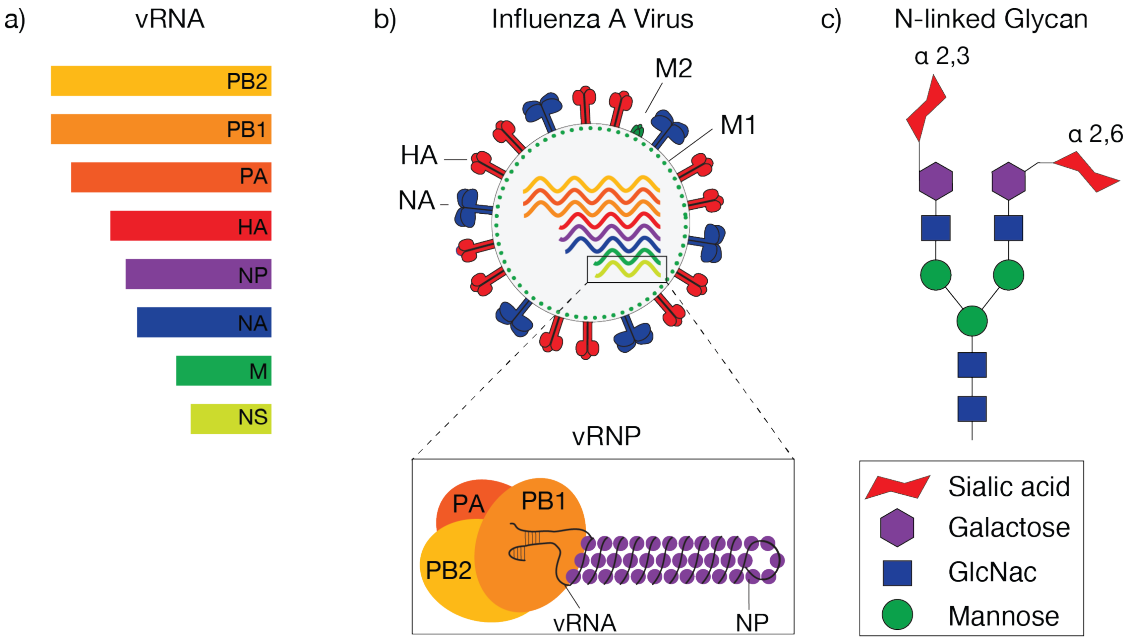


Figure 1.1: The influenza A virus structure and N-linked glycan. a) Schematic of the eight viral gene segments of the influenza A genome. The influenza A virus genome is segmented and made of 8 viral ribonucleoproteins (vRNP) complexes. Each segment represents a viral gene that codes for - PB2 (yellow), PB1 (orange), PA (dark orange), HA (red), NP (purple), NA (blue), M (green) and NS (light green). b) Influenza A virus. The matrix protein lies underneath the viral membrane as a support for the structure of the virus. The membrane proteins HA, NA and M2, are labelled with the vRNPs in the core of the virus. A single vRNP segment is also shown in detail along with its components - PB1, PB2, PA, vRNA and NP. c) α–2,3-linked sialic acid & α–2,6-linked sialic acid are shown in the figure.

also thought to form contacts between all other viral constituents. There are three transmembrane proteins embedded in the viral envelope: the predominant glycoproteins 1) HA and 2) NA that project out of the viral envelope as “spikes”, and 3) the proton channel, M2 protein.

1.2 Influenza Virus Life Cycle

To infect a host cell, viruses must first gain entry into the cell cytoplasm. This process is usually initiated by attachment to a receptor on the surface of the host cell. Influenza A virus mainly infects epithelial cells of the respiratory and gastrointestinal tract. The HA protein binds to sialic acid containing glycopoteins presented on the surface of the host cell membrane. There are several types of sialic acid, which can

be identified by the sialic acid linkage to the penultimate galactose of an N-linked glycan chain (Figure 1.1c) (Connor et al. 1994; Couceiro et al. 1993).

Generally, influenza A viruses can bind to a single type of sialic acid that is specific to each type of host cell. Sialic acid identity is an important factor for influenza virus-host tropism (Ito et al. 1998). In birds, avian influenza A viruses preferentially bind to sialic acid with a linkage between the C2 and C3 galactose, α–2,3-linked sialic acid via its HA protein (Figure 1.1c). This is usually found in the gastrointestinal tract of birds and the lower respiratory tract of mammals. This is different from mammalian influenza A viruses. The HA protein prefers to bind to sialic acid with a linkage between the C2 and C6 galactose, α–2,6-linked sialic acid. This is mainly found in the upper respiratory tract of mammals. Epithelial cells with a mixed population of sialic acids are possible, like in swine. Due to this fact, swine have been associated as a mixing vessel for the genetic reassortment between influenza viruses of avians and mammals (Ito et al., 1998).

Via receptor-mediated endocytosis, viruses are engulfed into the host cell by binding to the sialic acid receptor (Figure 1.2). Several endocytosis pathways can be used including caveolae-dependent, macropinocytosis-dependent, caveolae-independent, and clathrin-independent pathways but the most well-understood and widely accepted method is the clathrin-mediated endocytosis pathway (Edinger et al. 2014; Mercer et al. 2010). After binding, adaptor proteins are recruited and clathrin triskelions to the outside membrane where budding will form. Budding then occurs, forming a clathrin-coated pit. A mature pit will be cleaved from the plasma membrane through the use of membrane-binding and fission proteins such as dynamin forming a vesicle that then uncoats and forms an endosome. Via these pathways, influenza A viruses are able to enter the host cell within an early endosome where the genome of the virus and its transcriptional machinery needs to break through the viral capsid and membrane vesicle to allow for the hijacking of host cell machinery for reproduction.

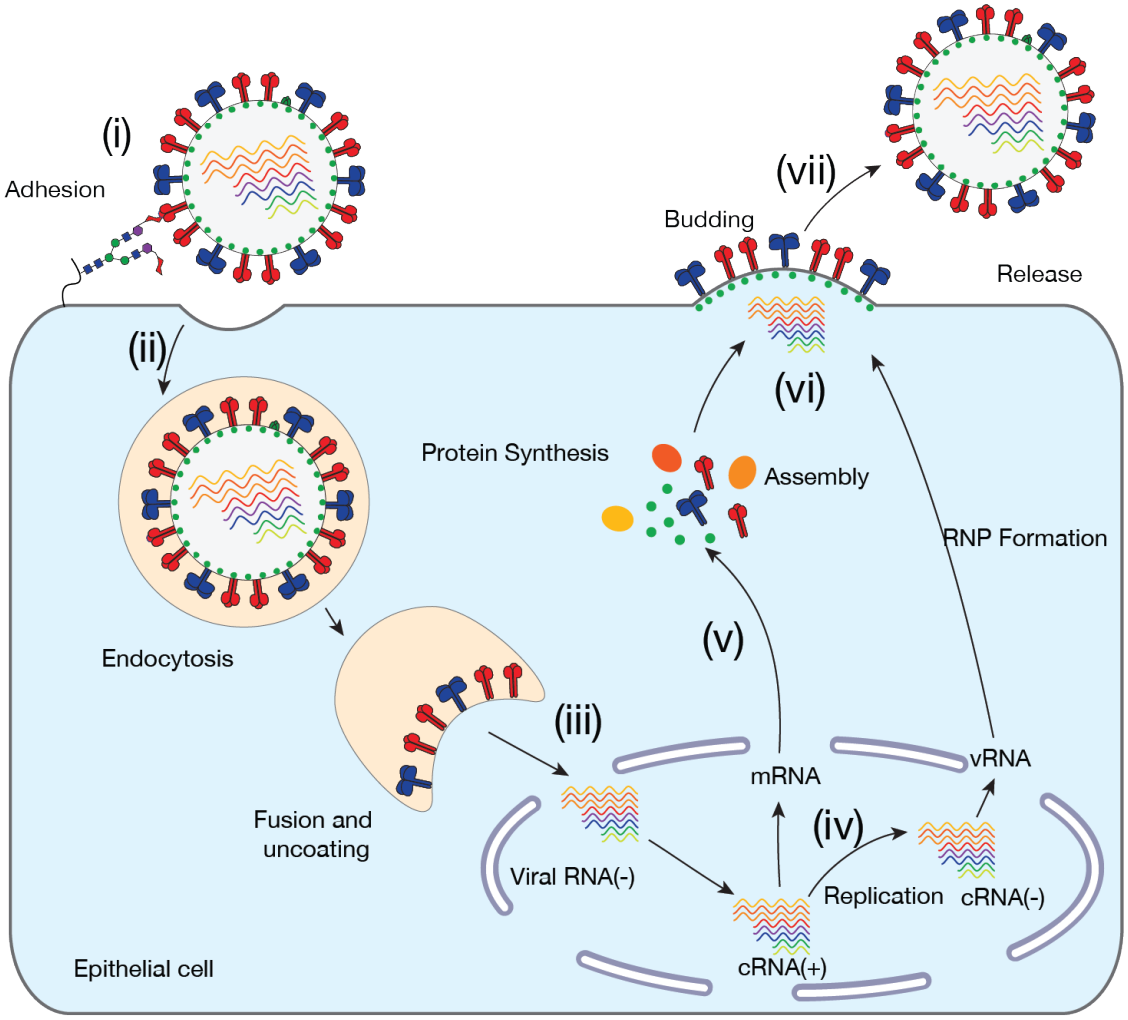


Figure 1.2: Influenza A Virus Life Cycle. (i) The influenza A virus first binds to the host cell surface glycans with sialic acid via its haemagglutinin (HA) that is presented at the outer surface of the virus particle. Receptor-mediated endocytosis is then initiated through the binding and fusion of the virus particle with the membrane of the host cell. (ii) This is then endocytosed into the host cell and is then trafficked to the endosome and fuses. The now, acidic endosome activates the M2 ion channel to pump protons into the core of the virus particle. Endosomal pH triggers membrane fusion, as well as the disintegration of the M1 scaffold and its interactions with the vRNPs. (iii) The vRNPs are then released into the cytoplasm of the host cells and transported into the nucleus where viral RNA replication begins. (iv) Newly synthesised vRNP are generated in the nucleus while (v) other components of the viral particles such as HA and NA are synthesised in the ER by the host cell ribosomes and newly made M1 has to then traffic back into the nucleus to facilitate vRNP export. (vi) Progeny vRNPs are then transported to the plasma membrane where other viral proteins such as HA and NA are concentrated at the lipid rafts region of the membrane. (vii) Here, new virus particles are assembled and buds off the host cell to infect other cells.

1.2.1 Fusion and Nuclear Import

Once the virus is enclosed in the early endosome, the virus then needs to fuse with the endosomal membrane. This allows the vRNP to be released in the cytoplasm (Figure 1.2). For this to happen, the HA protein undergoes a structural rearrangement, triggered by the reduced endosomal pH via endosomal trafficking and acidification (Hamilton et al. 2012). However, before the fusion, the HA protein must be cleaved by host proteases from the previous round of viral replication which converts the precursor HA0 into HA1 and HA2 subunits (Boulay et al. 1987; Chen et al. 1998), or within the endosome (Zhirnov et al. 2002). The structural rearrangement of the cleaved HA protein, HA2 that is triggered by the decrease in pH that will expose the hydrophobic fusion peptide of HA2 causing it to fuse with the viral membrane (Bullough et al. 1994).

The decrease in pH is caused by the pumping of protons by the viral ion channel M2, from the acidifying endosome into the core of the virus. When this occurs, weak interactions between the matrix protein, M1 and the vRNP are disrupted (Pinto et al. 1992). M1 and vRNP dissociate and the vRNP is released into the host cell cytoplasm. This exposes the nuclear localisation signals (NLS) on the vRNP where it translocates to nucleus and hijack the replication machinery of the host cell (Pinto et al. 1992).

1.2.2 Viral Gene Expression

Once in the nucleus, the vRNPs are transcribed to produce mRNA for the expression of viral proteins. Replication also occurs simultaneously via a complementary RNA (cRNA) intermediate (Figure 1.2). This is to provide negative sense RNAs to produce more vRNPs for transcription and the packaging into progeny virions (Fodor, 2013). The viral polymerase is responsible for both transcription and replication, and this occurs within the context of a vRNP. Viral gene expression is a tightly regulated process and was shown to express different genes at different times of the infection (Enami et al. 1985; Hatada et al. 1989; Shapiro et al., 1987). The expression of HA, NA and M1 are in the later stages of infection and consistent to their roles in

the export of vRNPs from the nucleus and structural roles to preserve the integrity of the progeny virions (Shapiro et al., 1987; Smith and Hay, 1982; Yángüez and Nieto, 2011). The regulation of viral gene expression has yet to be elucidated, but it has been proposed that these is regulated by a subtle difference in the promoter sequence of the viral genomic segments (Lee & Seong 1998; Lee et al. 2003)

1.2.3 Nuclear export of vRNP complexes

After the synthesis of vRNP complexes is completed, the complexes are transported from the nucleus to the cell membrane before it is being packaged into virus particles (Figure 1.2). The transport of vRNP complexes requires two vital proteins - M1 and NEP, which interact with the host cell export machinery (Chromosomal Maintenance 1, CRM1) that mediates the nuclear export through the nuclear membrane. Before nuclear export, the vRNPs are said to accumulate at the perinuclear microtubule organising centre (MTOC). A study by Lakdawala and colleagues suggested that vRNPs from subcomplexes within the cytoplasm will further coalesce forming a full genomic complement ready for packaging into viral particles (Lakdawala et al., 2014).

1.2.4 Genome packaging

A successful influenza virus infection would be that if each virion progeny contains the full complement of viral genomic segments upon budding from the host cell. This is likely to be achieved through selective packaging of the 8 different genomic segments instead of a random packaging strategy. It was shown through microscopic and biochemical studies that primarily progeny virion contain a single copy of each segment of the genome (Bergmann and Muster, 1995; Gao et al., 2012; Smith and Hay, 1982). It has been suggested that the viral genomic segments assemble through the interaction between RNA and RNA segments (Hutchinson et al., 2010). This has been demonstrated through tomographic reconstruction of budding virions where the relative positions of the 8 genomic vRNPs were conserved.

1.2.5 Viral Assembly and Budding

The virions then must leave the host cell, ready for the next round of infection of host cells (Figure 1.2). This is done by coalescing lipid raft domains on the apical surface of the cell membrane where they bud off from (Brown and Rose, 1992; Rodriguez Boulan and Sabatini, 1978; Scheiffele et al., 1999; Simons and Ikonen, 1997). Before budding, the newly synthesised viral transmembrane proteins - HA, NA and M2 at the endoplasmic reticulum (ER) are transported to the apical membrane via the Golgi apparatus. These proteins are then targeted to the lipid raft domains, which then cause the domain to broaden and coalesce (Chen et al., 2007; Leser and Lamb, 2005; Takeda et al., 2003). Through the expression of various combinations of viral proteins the gathering of the HA and NA at the lipid raft domains is proposed to cause membrane deformation and initiate the process of budding (Chen et al., 2007). Once the HA and NA are assembled at the lipid raft domains, M1 is said to bind to their cytoplasmic tails and polymerise, forming the matrix scaffold for the budding virion (Ali et al., 2000; Noton et al., 2007; Zhang et al., 2000). However, the exact mechanism of how M1 polymerises is still not understood. The scaffold M1 acts as a bridge for the incorporation of viral genomic segments complexes, which are packaged as a bundle of 8 segments (Noton et al., 2007). Through alanine substitution mutantation, M1 is proposed to recruit M2 (Chen et al., 2008; McCown and Pekosz, 2006), which leads to the curvature of the membrane forming the neck of the budding progeny virion. Subsequently, this results in membrane scission and the budding of the virions into the extracellular environment for the next round of the virus pathogenesis (Rossman et al., 2010; Rossman and Lamb, 2011). To prevent reattachment of the newly budded progeny after membrane scission, the sialic acid needs to be hydrolysed by the viral NA from the cell surface via its sialidase activity, thereby preventing reinfection (Air and Laver, 1989; Huang et al., 2008).

1.3 M1, the Matrix Protein of Influenza A Virus

The matrix protein of influenza A virus M1 is primarily responsible for the integrity of the virus underneath the viral envelope. M1 is associated with the viral membrane and connects the viral components between the outer membrane proteins hemagglutinin (HA), neuraminidase (NA) and Ion channel M2, and the core of the virus – the vRNPs. M1 versatility allows it to be involved in many important processes in the life cycle of influenza A virus such as maintainng the integreity of the virion shape, and the last stage of assembly and budding of new virions. Budding and assembly occur at the viral budozone, a domain of cholesterol or sphingolipid-rich membrane rafts (Veit & Thaa, 2011). The virus assembly is organised by M1, which acts as a hub for all other viral proteins which includes the vRNPs, the outer membrane proteins, including the ion channel and the viral membrane. The formation of the M1 scaffold protein to support the integrity of the new virion progeny is an important step before the budding process (Nayak et al., 2009). It was also shown that M1 is an essential determinant for virus particle morphologies (Chlanda et al., 2016). Residues 93 and 94 are located in M1 helix 6, which was previously shown to influence virion morphology (Burleigh et a., 2005). The exact details of how M1 self-associates and how that relates to its ability to bind to the membrane and disassociate during entry remains unclear and is one of the most critical questions in the field of influenza A virus research.

1.3.1 M1 Structure

Influenza Virus M1 protein is a 27.9 kDa protein comprised of 252 amino acid residues. M1 has two domains – the N-terminal domain (1-164 residues) and a shorter C-terminal domain (165-252 residues). M1 has the most conserved amino acid sequence as compared to other viral proteins (McCauley and Mahy, 1983). The N-terminal domain of M1 has been solved by X-ray crystallography over a range of pH values from pH 4.0 to pH 7.0 and higher (Arzt et al., 2001; Harris et al., 2001; Safo et al., 2014; Sha and Luo, 1997). The N-terminal domain consists of two 4-helical bundle subdomains (helices 1-4 and 6-9) which are connected by a

cross-over helix (helix 5) (Figure 1.3). Several studies have attempted to crystallise the full-length protein as well as the C-terminal domain but this has thus far proven unsuccessful (Arzt et al., 2001; Sha & Luo, 1997).

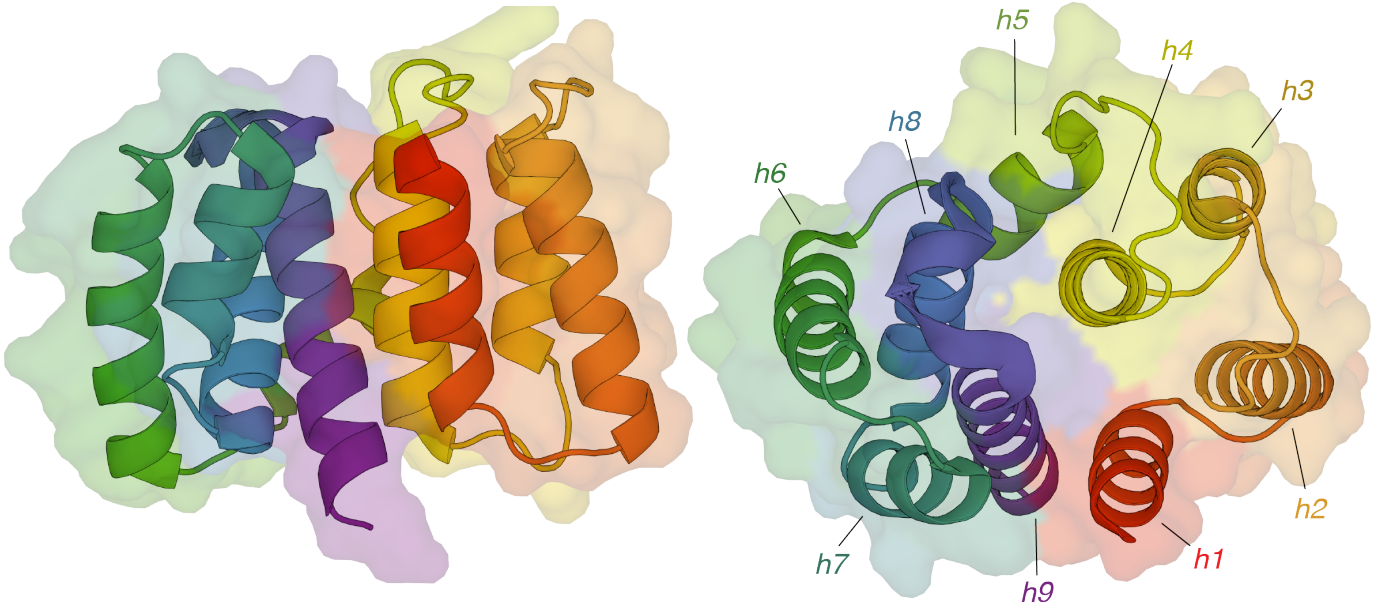


Figure 1.3: The globular domain of M1NT (PDB:1AA7). The N-terminal domain consists of two 4 helical bundle subdomains (helices 1-4, *h1* - *h4* and 6-9, *h6* - *h9*) which are connected by a cross-over helix (helix 5, *h5*).

The C-terminal domain appears to be largely unstructured in both acidic and neutral pH, making it difficult to crystallise. The C-terminal sequence is highly enriched in disorder-promoting residues (Dunker et al., 2001) composed - Met, Ala and Gln which may explain the unstructure nature of the domain. Other biophysical techniques have been used to characterise M1CT including circular dichroism (CD), dynamic light scattering, and analytical ultracentrifugation. These techniques have shown M1CT is a flexible and weakly structured region (Kordyukova et al., 2008; Ksenofontov et al., 2010; Shishkov et al., 2009, 2011; Zhang et al., 2012). CD is the only technique that analyses the structure of the C-terminal domain while other techniques examined the size of domain, which may appear larger due to the unusual size. The N-terminal domain and the C-terminal domain are connected by a long loop that is susceptible to proteolysis (Kordyukova et al., 2008).

Full-length M1 has been studied using Small-angle neutron scattering (SANS), with a reported radius of gyration of 2.9 nm which is much larger than expected for a 28 kDa globular protein (Arzt et al., 2001). The protein length was found to be 8.0 nm by Ruigrok and colleagues, which is also much larger than the 6.0 nm size that was observed by EM (Rui- grok et al., 2000). These studies also suggest that the C-terminus is unstructured. However, it is possible that M1CT forms a more stable structure when polymerised or in complex with binding partners. But the unstructured M1 C-terminus may also have functions in its disordered form. The conventional idea that proteins must fold in a certain way to perform biological functions has been revised as proteins with disordered structure that can perform its biological function has been found (Ge et al., 2010). The link between the three-dimensional structure of proteins and their functions forms the fundamental concepts in biochemistry. Intrinsically disordered proteins (IDPs) exist in a myriad of dynamic ensembles with no one specific stable structure. In this way, they are able to perform their variable biological functions. The C-terminal domain might behave as an IDP and fold upon binding with interacting partners. It has been suggested that viral proteins may contribute the ability of the virus to adapt to rapid changes to environmental conditions through its disordered region (Goh et al., 2008). The disordered region of the protein maybe able to act as a mechanism for the prevention of the binding of antibodies, diminishing immune response. This can act as viruses' mechanisms of adaptation and protection for survival (Goh et al., 2008). For example, HIV-1 is able to evade the immune response effectively in this way (Goh et al., 2008).

In full-length structures of M1 from Influenza A Virus (IAV), the M1CT is folded into a helical bundle, and both M1NT-M1NT and M1NT-M1CT intermolecular contacts are formed. It was shown that the C-terminal domain folded into three-helix globular domain in a linear polymer and the C-terminal domain was interacting with the membrane surface of the N-terminal domain from an adjacent M1 monomer (Peukes et. al, 2020).

1.3.2 M1 Oligomerisation

M1 has several functions in the influenza life cycle including the export of vRNP from the nucleus, recruitment of viral proteins during virus assembly and egression, regulation of viral transcription. Its amino acid sequence polymorphisms determine virus morphology (Bialas et al., 2012; Kordyukova et al., 2019). M1 expressed alone in mammalian cells is able to form virus-like particles (VLP), and this is attributed to its ability to oligomerise into long, helical polymers forming a lattice underneath the membrane coating the inside of virions (Calder et al., 2010; Gómez-Puertas et al., 2000; Noton et al., 2007). VLPs are molecules resembling viruses but are not infectious due to absence of viral genome. This ability to self-associate is critical for the formation of the matrix layer especially during the late stage of viral pathogenesis, the assembly, and budding stage (Ali et al., 2000; Avalos et al., 1997; Barman et al., 2001; Gómez-Puertas et al., 2000; Latham and Galarza, 2001; Wang et al., 2010).

M1 oligomerisation correlates with pH, where it assembles at neutral pH and disassembles at low pH (Shtykova et al., 2017; Zhang et al., 2012). The N-terminal domain alone shows a pH-dependent oligomerisation (Noton et al., 2007) and the C-terminal domain may also play a role in the oligomerisation pattern (Zhang et al., 2012). *In vitro*, M1 binds to lipid membranes that contain phosphotidylserine (PS) (Ruigrok et al., 2000) and *in vivo*, membrane binding was seen to promote M1 oligomerisation. Membrane binding primarily involved the N-terminal domain (Baudin et al., 2001; Höfer et al., 2019). M1 oligomerisation is also said to help with membrane bending, which is required for budding initiation. A study has shown that M1 was able to cause membrane curvature in the absence of other viral proteins. This was seen in lipid bilayer with negatively charged lipids. It was shown that M1 oligomerisation are needed to alter the lipid bilayer structure (Dahmani et al., 2019). It was also shown that oligomerisation is needed in order to allow virion release for the next cycle infection (Nayak et al., 2004; Nayak et al., 2009; Rossman and Lamb, 2011; Ruigrok et al., 2000).

M1, as a matrix protein, is crucial for the protein to be able to associate and dissociate at the right stages of the influenza A virus life cycle. As with

other orthomyxoviruses, the dissociation of scaffold matrix demands a low pH in a specific compartment (Batishchev et al., 2016; Lee, 2010; White and Whittaker, 2016; Zhirnov, 1992). This is a multi-step process that requires the disruption of important contacts that includes interactions between matrix proteins, the inner genome and the outer membrane proteins. Understanding the structure and how these interactions occur in terms of specific amino acids that are involved might help better understand how M1 oligomerises.

It is still poorly understood on what the physical triggers are that converts M1 from a water-soluble component in the nucleus and cytosol to an immobilised lattice of helical polymer on the membrane surface. Isolated full-length M1 is difficult to study in solution due to its high propensity to aggregate and is only weakly soluble. Moreover, membrane-attached oligomer is challenging to study at a high resolution due to its heterogeneity.

M1 N-terminal domain (M1NT) in crystal structures is frequently found to be part of a homodimer (Harris et al., 2001; Safo et al., 2014; Sha and Luo, 1997). When comparing between the crystal structures solved at acidic and neutral pHs, there is very little difference within a single monomeric subunit. However, the interactions between the monomers in the dimer are entirely different (Safo et al., 2014). Perhaps the most striking difference is the two primary arrangements of dimers between the crystals solved at pH 4.0 and pH 7.0 – PDB ID: 1AA7 (Safo et al., 2014) and 1EA3 (Arzt et al., 2001), respectively. 1AA7 crystal shows a side-by-side dimer arrangement in the asymmetric unit while the 1EA3 crystal shows a stacked dimer arrangement (Figure 1.4).This difference in arrangement may have some implications on the functionality of the protein

However, solution studies have indicated that M1 is primarily a monomer (Arzt et al., 2001; Shtykova et al., 2013) but these can co-exist with helical clusters present in small concentrations (Shtykova et al., 2013). The amount of clusters seems to increase with increasing pH. Above pH 6.0, the tendency for M1 to oligomerise increases as it gets to pH 7 (Shtykova et al., 2017). Zhang and colleagues have shown that M1 self-association possibly involves both the N-terminal and

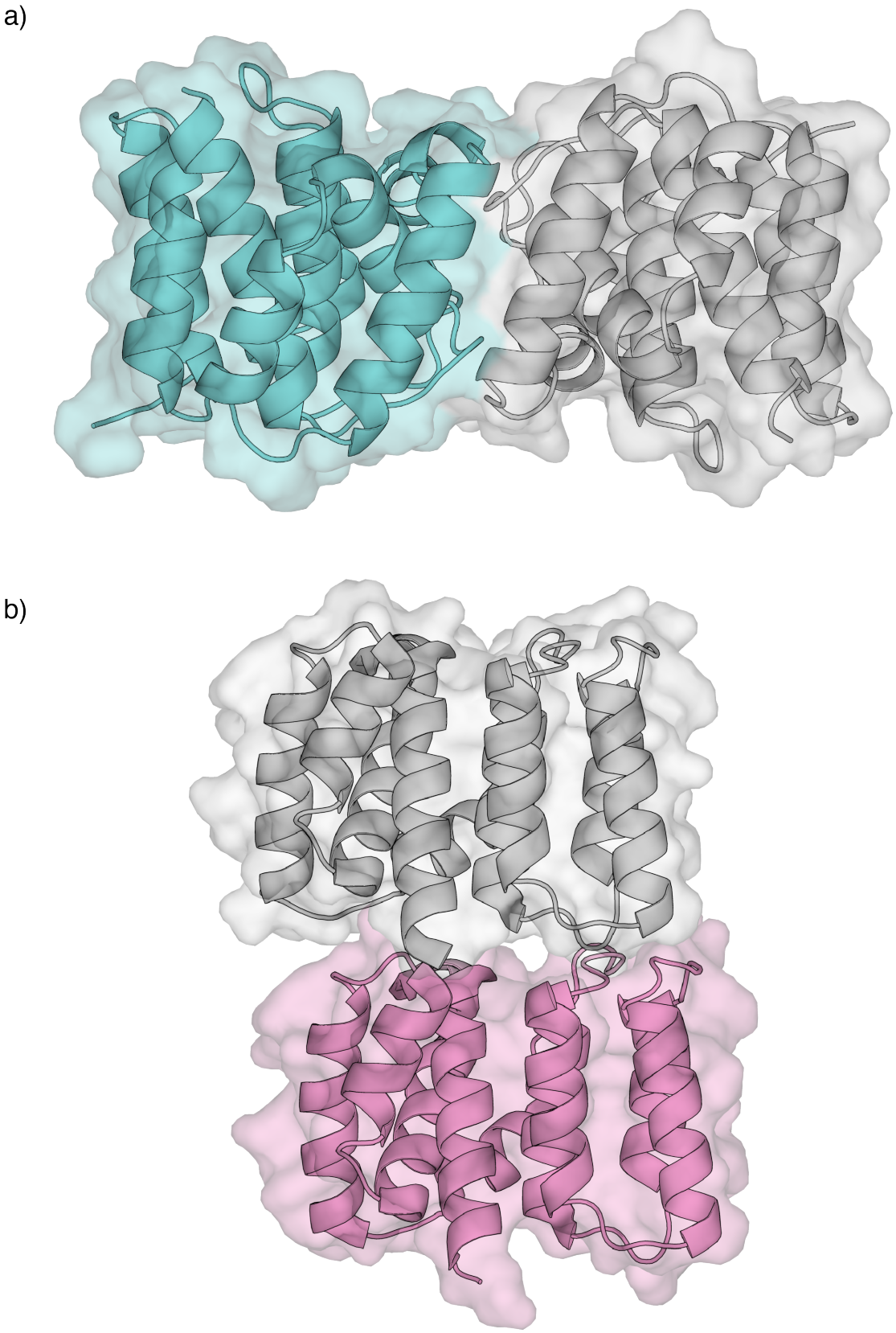


Figure 1.4: Illustration of the dimer in the asymetric unit a) side-by-side arrangement of M1NT (PDB ID: 1AA7) & b) stacked dimer arrangement of M1NT (PDB ID: 1EA3).

C-terminal domains at neutral pH (Zhang et al., 2012). The oligomers formed between full-length M1 and M1NT was shown to be different, with full-length protein forming oligomers with four-fold symmetry while M1NT forms oligomers with fivefold symmetry (Zhang et al., 2015). Furthermore, C-terminal residues have been implicated to be an essential determinant in the oligomerisation pattern. Zhang and colleagues have reported that residues S183 and T185 of the C-terminal domain are important in this regard (Zhang et al., 2015). Single point mutations at these residues have been shown to affect virus morphologies while double mutations disrupted virus replication (Zhang et al., 2015), which shows the importance of these residues for M1 functionality. Furthermore, another study has shown that the single histidine of the C-terminal domain play an important role in the self-association of M1 (Shtykova et al., 2017).

1.3.3 M1 Scaffold Dissociation with pH

The disintegration of M1 scaffold matrix is crucial for the critical step in releasing the viral genome after host infection as the viral particle fuses with the late endosome. Through limited proteolysis (Lip) experiment in vitro, studies have found that a decrease in pH to 6.0 of the internal core of the virions caused an irreversible change in M1 C-terminus allowing its dissociation from the vRNP (Li et al., 2014; Stauffer et al., 2014).

However, during this process, the interactions between M1 and the membrane remain (Batishchev et al., 2016). Studies have also verified that the M1 scaffold underneath the viral envelope undergoes a significant change as a whole scaffold when the pH drops from a neutral pH (Calder et al., 2010; Fontana et al., 2012; Fontana and Steven, 2013). Furthermore, a study done by Shtykova and colleagues on M1 self-association at different pH has shown that M1 changes at two distinct pH ranges: pH 6.0 to 7.0 and pH 5.0 to 5.5 and these changes may be due to the change in hydrogen bonds and electrostatic interactions between M1 individual monomers (Shtykova et al., 2017). At pH 6.5-7.0, M1 dissociates itself from the M1-vRNP interaction, releasing the vRNP. It should be noted that the vRNP is

released from M1 but still caught in the M1 scaffold. Once at the disintegration step, at a pH below 5.5, M1 starts to dissociate from the scaffold to allow the release of the vRNP into the cytoplasm.

1.3.4 M1 Interaction with Lipid

The lipid-binding regions of M1 were discovered in the 1980s and have been mapped to the hydrophobic regions within α-helix 4 and H7 & H8 of the N-terminal domain (Gregoriades and Frangione, 1981; Ye et al., 1987). However, there is still no agreement to the arrangement of M1 protein with respect to the lipid membrane and the region where the interaction of M1-lipid membrane occurs. Earlier studies proposed that there was a partial insertion of M1 into the hydrophobic section of the lipid membrane (Gregoriades and Frangione, 1981; Sha and Luo, 1997). Later studies suggested instead that membrane association is driven primarily by electrostatic interactions of M1 with the polar head of lipid molecules (Arzt et al., 2001; Baudin et al., 2001). According to the first hypothesis, M1 is partially inserted into the non-polar region of the inner lipid membranes through the hydrophobic regions of M1. The second hypothesis argues that M1 interacts with the polar head of lipid molecules through electrostatic interactions of the positively charged amino acids on the N-terminal domain. Through EM, Ruigrok and colleagues were not able to detect the possibility of insertion of M1 into the lipid membrane (Ruigrok et al., 1989). Furthermore, cryo-electron tomography has shown that the M1 layer underneath the lipid membrane and the inner lipid bilayer form distinct, separate layers (Yamaguchi et al., 2008).

There are two regions in M1 that can be putative regions of lipid binding: 1) Triplet Arginine residues – Arg76, Arg77 & Arg78, located at helix 5 (*h5*) (Das et al., 2012); 2) Polybasic region of Nuclear Localisation Signal (NLS) residues – Lys95, Lys98, Arg101, Lys102, Lys104, Arg105 at helix 6 (*h6*) (Liu and Ye, 2004; Thaa et al., 2009; Ye et al., 1995). These basic regions are both solvent-exposed and can interact with negatively charged polar phosphate heads of the lipid membrane (Sha and Luo, 1997).

Through mutagenesis studies, Das and colleagues mutated these triple basic residues to examine the effect on virus viability. They concluded that mutating either Arg77 or Arg78 to an alanine residue resulted in defective virus budding caused by the accumulation of vesicles of M1 underneath the membrane while mutating them into Asp residue were lethal to the virus viability. The same lethal effect was observed when both Arg77 and Arg88 were mutated into alanine residues (Das, et al., 2012). Furthermore, these mutants were shown to have defective intracellular trafficking and cannot be incorporated to form virus-like particles (VLPs) (Kerviel et al., 2016). Hence, this shows that the Triplet Arg residues are essential for the ability of M1 to form VLPs. It should make clear that these mutants fail to bud due to defective intracellular trafficking. Liu and Ye also have conducted mutagenesis studies on the NLS of M1 (residues 101-105) and found that mutations at these residues resulted in a defective viral replication and hence, slowed viral growth (Liu and Ye, 2002). Furthermore, M1 was not transported into the nucleus and remained in the cytoplasm of infected cells.

It was also shown that the polybasic NLS residues of M1 are essential for the translocation of M1 in the nucleus during the late steps of viral replication (Liu and Ye, 2002). Mutagenesis studies by Burleigh and colleagues have shown that the polybasic region of M1 is involved in budding due to its role in translocation (Burleigh et al., 2005). Mutations in these region caused alterations to the morphology of budded viruses. Also, the polybasic NLS of M1 has also been implicated to participate in membrane interaction through liposome interaction experiments *in vitro* (Baudin et al., 2001; Ruigrok et al., 2000). However, a more recent *in vivo* study demonstrated that polybasic NLS is not required to bind to membranes (Thaa et al., 2009). Their results do not exclude the possibility that the polybasic NLS and its surroundings are involved in membrane binding but instead, confirms that the NLS alone is not the only factor for membrane binding. Together with previous findings, this leads to the assumption that membrane binding involves different regions of M1 that work synergistically or on multiple redundant domains

via both its charged and hydrophobic regions (Kretzschmar et al.,1996; Shishkov et al., 1999; Kordyukova et al., 2008; Thaa et al., 2009).

A study by Baudin and colleagues showed that full-length M1 as well as M1NT bind effectively to negatively charged lipids, such as PS, at a concentration of 50% (Baudin et al., 2001) but not with lipids with a neutral charge, like phospatidylcholine (PC). However, an SPR study has found that M1 can bind to both neutral and negatively charged lipids, although the infinity for adsorption on charged lipids is about twice as much as compared to neutral lipids (Shishkov et al., 2009). They concluded that the binding of M1 to lipids could involve both electrostatic and hydrophobic interactions and this was also supported by other groups (Brevnov et al., 2016; Knyazev et al., 2009). Recent research using fluorescent probes has shown that M1 adsorption to lipid bilayer needs to comprise of PS and is not driven by the presence of either cholesterol or phosphatidylethanolamine (Shilova et al., 2017). Several studies performed *in vitro* and *in vivo* have shown the electrostatic interactions of M1 with lipid membranes (Batishchev et al., 2016; Baudin et al., 2001; Hilsch et al., 2014; Shilova et al., 2017). A study done by Hilsch and colleagues using both a model system and living cell to study how M1-lipid interaction affects M1 self-association, concluded that M1 binding to lipids and M1 self-association are not mutually exclusive processes and probably consequential processes (Hilsch et al., 2014). They showed that the interaction of M1 with lipids affect M1 self-association, where they detected oligomerisation of M1 on lipid membranes. In contrast, they do not detect any oligomerisation in the absence of lipid membrane in solution.

Despite the numerous studies to understand M1 self- association, it still remains unclear how this process occurs. Several questions regarding which subunit interactions are important for M1 self-association, how membrane interactions facilitate M1 self-association and how M1 conformationally changes in response to polymerisation & pH remain to be addressed.

1.3.5 M1 of Influenza A Virus and Influenza C Virus

The crystal structure of the homologous matrix protein full-length M1 of the Salmon Anemia influenza C virus (ISAV), M1-ISAV has been solved in recent years (Zhang et al., 2017). The M1-ISAV is a 22 kDa protein with 196 amino acids, and the crystal structure was obtained at pH 7.5. Although the calculated identity of amino acids between M1-ISAV and M1-influenza A Virus (M1-IAV) is 11%, this is the first full-length homologous matrix protein of an orthomyxovirus that has been crystallised and has given some insights to the organisation of the M1-ISAV domains.

The structure of M1-ISAV can be described as two helical bundles connected by a central α-helix : one helical bundle in the shape of a barrel made of 5 α-helices which is connected to a long flexible linker of the central α-helix and further packed with a bundle of two α-helices. There were some similarities and differences when M1-ISAV was compared to M1-IAV. The M1-ISAV was observed to assume an elbow shape with two globular domains - N-terminal domain and C-terminal domain. Like the M1-IAV, the N-terminal domain closely resembles the M1NT of M1-IAV. The C-terminal assumed a four helical bundle which is connected to the N-terminal domain via a flexible linker. This conformation has been predicted for the M1CT of the influenza A virus M1, M1-IAV (Shishkov et al., 2009, 2011).

Comparing the C-terminal domain of M1-ISAV to M1-IAV, it was shown that the C-terminal domain from ISAV is made of 3 α long helices and a short helix at the beginning of the extended linker region. This is in contrast to what was predicted for the M1 C-terminal domain of M1-IAV, which was predicted to form a tight four-helix bundle upon binding (Shishkov et al., 2009, 2011). In contrast to M1-IAV, the linker region of M1-ISAV is not susceptible to proteolytic attack possibly due to the absence of basic residues or hydrophobic residues. Also, the presence of cysteines residues, C13 of helix 2 and C115 from the linker that forms a disulphide bond tethering the linker to the N-terminal domain, could be providing further stabilisation.

Through liposome floatation assays, it was shown that M1-ISAV also binds to the lipid membrane through electrostatic interactions through the positively

charged loop of the N-terminal domain. This was similarly observed in M1-IAV (Baudin et al., 2001).

1.3.6 M1 of influenza A Virus and Matrix Proteins of Other Enveloped RNA Viruses

The matrix protein from other viruses that come from different families might be structurally, functionally and evolutionarily related. For example, the M1 of influenza A virus and the p17 protein of HIV-1 have been shown to be structurally related in the domain that binds the membrane and RNA (Harris et al., 1999). In comparison to the matrix protein of the orthomyxoviruses, the matrix protein of other enveloped RNA viruses contains both α- and β-structural elements (Sha and Luo, 1997).

The matrix protein from Ebola virus, VP40 also plays a vital role during the assembly and budding of the virus by maintaining the structural integrity of the virus (Hoenen et al., 2010; Stahelin, 2014; Wijesinghe et al., 2017). Interestingly, the VP40 has shown to adopt several different conformations that are distinct from each other to serve the function needed to adapt to the ebola virus life cycle (Bornholdt et al., 2013). There are three different structures of the VP40 that have been crystallised that shows the adaption of VP40 to different functions. The first structure shows VP40 as a monomer with two domains- N-terminal and C-terminals domains, both of which are β-sandwich domains (Dessen et al., 2000). A second structure assumes a butterfly-shaped dimer and this conformation was associated with its trafficking to the cell membrane. The last structure assumed a different conformation that uniquely binds to RNA regulating viral transcription in the nucleus of infected host cells (Bornholdt et al., 2013).

The VP40 of Ebola virus also binds to negatively charged lipids such as PS, similarly to the M1 protein of influenza A virus (Adu-Gyamfi et al., 2015; Gerstman et al., 2016). Like M1, the interaction between VP40 and the negatively charged lipid is said to be electrostatic. In VP40, this triggers a rearrangement into a linear hexamer that is crucial for budding in which the C-terminal domain binds to the

membrane (Scianimanico et al., 2000). Interestingly, in this hexamer conformation, the C-terminal domain of one monomer interacts with another from a VP40 dimer. As a whole, the hexamer constructs a filamentous morphology that was also seen in tomograms of virus particles. The negatively charged lipids that are present in the inner leaflet bilayer may be used by matrix proteins of other enveloped viruses as targets for binding.

These other enveloped viruses might be different in amino acid sequence, but it might be useful to understand how these different matrix proteins are structurally related to help guide our understanding of the M1 matrix protein and whether M1 also undergoes conformational change.

1.4 Thesis Objectives

There are gaps in our understanding of the structural and functional implications of M1. In this thesis, we will try to address critically important questions by studying the matrix protein, M1 of the influenza A virus using solution-state NMR as the primary tool to study its dynamics and structure.

- What are the structural differences between M1 at acidic and neutral pH?
- Does the C-terminal domain of M1 form a stable structure in solution?
- What are the dynamics and structural plasticity of different regions that may be involved in protein interactions, membrane binding and oligomerisation?
- How does the C-terminal domain affect the structure of the M1 N-terminal domain?
- How does membrane binding affect M1 structure and its ability to oligomerise? What is the conformational trigger that is associated with M1 oligomerisation?

With one little action
The chain reaction will never stop
Make it strong
Shine a light, and send it on

— Disney

2 Materials and Methods

Expression and Purification of WT M1NT, M1(1-224) and mutant Constructs

Constructs

All M1 constructs were based on the A/England/250/2009(H1N1) sequence (J9UXB0) and codon-optimized for E. coli expression. All M1 constructs contain a C-terminal 3C protease cut site followed by a $(His)_9$. These constructs were cloned into pCOLD II plasmid (Takara Bio) containing an ampicillin resistance gene.

M1NT constructs Expression and Purification

Plasmids were transformed (via heat shock) into *E. coli* strain BL21 (DE3) (NEB) and expressed as PolyHis-M1NT as His_9-tagged proteins. Overnight culture of transformed *E. coli* (200mL) was used to inoculate 2 L of high salt Luria Broth (LB) media. Cells were grown at 37°C containing ampicillin, 100 μg/mL to an OD_{600} of 0.5-0.8. Cells were cold shocked for 30 minutes on ice and left at 15°C for 10-15 minutes. The cultures were spun down at 3000xG for 15 minutes.

Cell pellets were then resuspended in 1 L of 2x M9 minimal media containing either 1 g/L ^{15}N-labelled ammonium chloride, or 1 g/L ^{15}N-labelled ammonium

sulfate and 3 g/L ^{13}C-labelled glucose, and expression was induced at 15°C with 1.5 mM IPTG (isopropyl-1-thio-β-Δ-galactopyanoside).

M9 minimal growth media was made with 48 mM Na_2HPO_4, 22mM KH_2PO_4, 8.5 mM NaCl, 1mM $MgSO_4$, 40 mg/mL Thiamine-HCl, 19 mM NH_4Cl (or $^{15}NH_4Cl$), 17 mM Glucose (or ^{13}C-labelled.

The culture was harvested by centrifugation at 4000xG for 15 minutes at 15°C after cultures grow overnight for 16-18 hours. Then the cells were resuspended in lysis buffer(20 mM Tris, 300 mM NaCl, 2 mM β-mercaptoethanol, 1 Roche EDTA-free inhibitor tablet per 50 mL lysis buffer, 1 mM EDTA, pH 7.5) and the cells were lysed using a high-pressure cell disruptor at 33 Kpsi. Insoluble materials were removed by centrifugation by at 27,216xG at 4°C for 15 min. Affinity chromatography was used as the first step of the purification. The lysate was run through a cOmplete™ His-Tag Purification Column (Roche), and the column was then washed with a lysis buffer before washing with low immidazole wash (30mM imidazole) to remove any non-specific proteins bound to the column. M1NT bound to the column was then eluted with elution buffer (300 mM imidazole). The eluted sample pH was adjusted to pH 5.0 and the salt concentration increased to 500 mM before it was concentrated for gel filtration using 10 kDa MWCO centrifugal concentrators up to a maximum of 150-200 µM. The sample was subjected to a size exclusion column (Superdex increase 200 pg 10/300GL) using the Akta Pure (GE Healthcare) that had been equilibrated in running buffer (20 mM Tris buffer, 500 mM NaCl, 1 mM EDTA, pH 5.5). The fractions with protein were collected and concentrated to 500 µL with 10 kDa MWCO centrifugal concentrators, buffer exchanged with NMR sample conditions (20mM Tris, 50mM NaCl, 2mM β-mercaptoethanol, 1 mM EDTA, and pH 5.5). For samples at neutral pH the concentrated sample was adjusted to pH 7.0 after concentration. Final yields for M1NT (residues 1-164), M1(1-224) and the M1NT(L4A) variant were 0.5 mg, 1 mg and 2 mg, respectively, per liter of minimal media.

M1CT - Expression and Purification

pFloat-SUMO-M1CT (Given by Dipali Patel) was transformed by heat shock into *E. coli* strain BL21 (DE3) (NEB) and expressed as SUMO fusion proteins. Overnight culture of transformed *E. coli* (200mL) was used to inoculate 2 litres of Luria Broth (LB) media. Cells were grown at 37°C containing kanamycin, 100 μg/mL to an OD_{600} of 0.5-0.6. The cultures were spun down at 3000xG for 15 minutes. Cell pellets were then resuspended in 1 L of 2x M9 minimal media containing either 1 g/L 15N-labelled ammonium chloride, or 1 g/L 15N-labelled ammonium sulfate and 3 g/L 13C-labelled glucose, and expression was induced with 1 mM IPTG ((isopropyl-1-thio-β-Δ-galactopyanoside). M9 minimal growth media was made with 48 mM Na_2HPO_4, 22 mM KH_2PO_4, 8.5 mM NaCl, 1 mM $MgSO_4$, 40 mg/mL Thiamine-HCl, 19 mM NH_4Cl (or $^{15}NH_4Cl$), 17 mM Glucose (or ^{13}C-labelled. Cells were harvested by centrifugation at 4000xG after induction with 1 mM IPTG (isopropyl-1-thio-β-Δ-galactopyanoside) at 37°C for 11-13 hours.

Harvested cells were resuspended in lysis buffer (20 mM Tris, 500 mM NaCl, 0.1 mM PMSF, pH 7) and lysed using a high-pressure cell disruptor at 33 Kpsi. Insoluble materials were removed by centrifugation by at 40,000xG at 4°C for 15 min. Lysate was loaded through a 5 mL HisTrap FF columns (GE Healthcare), and the column was then washed with lysis buffer to remove any residual proteins. M1CT bound to the column was then eluted with elution buffer (20 mM Tris, 500 mM NaCl, 500 mM imidazole, 0.1 mM PMSF, pH 7). Fusion protein was then cleaved using PreScission Protease, 5 units (GE healthcare) dialysed in Cleavage Buffer (50 mM Tris, 150 mM NaCl, 1 mM EDTA, 1 mM DTT, pH 8.0) overnight at 4°C. Cleaved proteins were run back through the 5 mL HisTrap FF columns (GE Healthcare). The flowthrough containing the proteins were then dialysed with a dialysis membrane in 20 mM Tris, pH 8.0 overnight at 4°C before being subjected to Ion Exchange (IE) chromatography using IE column to remove the fusion proteins. Identity of elution peak fractions and protease cleavage were confirmed by SDS-PAGE with a band at 19 MW KDa.

NEP Expression and Purification

His-tagged NEP construct was kindly provided by Duncan Paterson from Fodor's lab of Dunn School of Pathology, University of Oxford. Plasmids pET-M14-NEP were transformed into *E. coli* strain BL21 (DE3) and expressed his-NEP as his_4-tagged proteins. Overnight culture of transformed *E. coli* (50mL) was used to inoculate 2 litres of Luria Broth (LB) media. Cells were then grown at 37ř C containing kanamycin, 100 μg/mL to an OD_{600} of 0.8. The culture was then condensed to 500mL of M9 minimal media and was harvested by centrifugation at 4000xg after induction with 1 mM IPTG (isopropyl-1-thio-β-Δ-galactopyanoside) at 37°C for 4 hours.

NEP is mostly expressed in inclusion bodies (IB) and hence the IB needs to be solubilised. This is done by resuspending the harvested cells in lysis buffer and lysed using a high-pressure cell disruptor at 33 Kpsi. The IB were isolated by centrifugation at 40,000xG at 4°C for 15 min and washed with 1x PBS before resuspending in solubilization buffer (50 mM Tris, 200 mM NaCl, 6 M guanidine-HCl, 2 mM β-83 mercaptoethanol, 2 mM imidazole, 10% glycerol, pH 8.0) at 37°C for 11-13 hours. Any insoluble materials are removed by centrifugation at 40,000xG for 15 mins at 4°C.

Affinity chromatography was used as the first step of the purification and for refolding of NEP. Solubilised IB was loaded onto 5mL cOmplete™ His-Tag Purification Column Prepacked (Roche) at a flow rate of 5ml/min using an AKTA Prime (GE Healthcare) and the column was washed with solubilization buffer. Column-bound NEP was refolded in a gradient of 6 M to 0 M urea (50 mM Tris, 200 mM NaCl, 2 mM β mercaptoethanol, 2 mM imidazole, 10% glycerol, pH 8.0) at a length of 50mL at 1 ml/min. Refolded NEP was then eluted using elution buffer (50 mM Tris, 200 mM NaCl, 2 mM β-mercaptoethanol, 300 mM imidazole, 10%glycerol, pH 8.0). The His-tagged of NEP was then cleaved using PreScission Protease, 5 units (GE healthcare) dialysed with a dialysis membrane in Cleavage Buffer (50 mM Tris, 150 mM NaCl, 1 mM EDTA, 1 mM DTT, pH 8.0) overnight at room temperature.

The cleaved sample was then concentrated to 500 μL using a spin concentrator and the N-terminal His-tag was removed by size exclusion gel chromatography

(HiLoad Superdex 75 pg) using Akta Pure (GE Healthcare) that had been equilibrated in NMR buffer (20 mM Tris, 50 mM NaCl, pH 7). Identity of elution peak fractions and protease cleavage were confirmed by SDS-PAGE with a band at 14 MW KDa.

M1CT - Expression and Purification

pFloat-SUMO-M1CT was transformed into *E. coli* strain BL21 (DE3) and expressed as SUMO fusion proteins. Overnight culture of transformed *E. coli* (200mL) was used to inoculate 2 litres of Luria Broth (LB) media. Cells were grown at 37°C containing kanamycin, 100 μg/mL to an OD_{600} of 0.5-0.6. The cultures were spun down at 3000xG for 15 minutes. Cell pellets were then resuspended in 1 L of 2x M9 minimal media containing either 1 g/L ^{15}N-labelled ammonium chloride, or 1 g/L ^{15}N-labelled ammonium sulfate and 3 g/L ^{13}C-labelled glucose, and expression was induced with 1 mM IPTG ((isopropyl-1-thio-β-Δ-galactopyanoside). M9 minimal growth media was made with 48 mM Na_2HPO_4, 22 mM KH_2PO_4, 8.5 mM NaCl, 1 mM $MgSO_4$, 40mg/mL Thiamine-HCl, 19mM NH_4Cl (or $^{15}NH_4Cl$), 17mM Glucose (or ^{13}C-labelled. Cells were harvested by centrifugation at 4000xG after induction with 1 mM IPTG (isopropyl-1-thio-β-Δ-galactopyanoside) at 37°C for 11-13 hours.

Harvested cells were resuspended in lysis buffer (20 mM Tris, 500 mM NaCl, 0.1 mM PMSF, pH 7) and lysed using a high-pressure cell disruptor at 33 Kpsi. Insoluble materials were removed by centrifugation by at 40,000xG at 4°C for 15 min. Lysate was loaded through a 5 mL HisTrap FF columns (GE Healthcare), and the column was then washed with lysis buffer to remove any residual proteins. M1CT bound to the column was then eluted with elution buffer (20 mM Tris, 500 mM NaCl, 500mM imidazole, 0.1 mM PMSF, pH 7). Fusion protein was then cleaved using PreScission Protease (GE healthcare), 5 units dialysed in Cleavage Buffer (50 mM Tris, 150 mM NaCl, 1 mM EDTA, 1 mM DTT, pH 8.0) overnight at 4°C. Cleaved proteins were run back through the 5mL HisTrap FF columns (GE Healthcare). The flowthrough containing the proteins were then dialysed with a dialysis membrane in in 20mM Tris, pH 8.0 overnight at 4°C before being subjected to ion exchange

chromatography to remove the fusion proteins. Identity of elution peak fractions and protease cleavage were confirmed by SDS-PAGE with a band at 10 MW KDa.

SEC-MALS

Multiangle Ligjht Scaterring (MALS) experiments were performed during size exclusion chromatography on an analytical Superose 6 10/300 GL column (GE Healthcare) equilibrated with 50 mm HEPES 300 mM NaCl 10% glycerol, 5 mM $MgSO_4$, pH 8.0. Elution was monitored via online static light-scattering (DAWN HELEOS 8+, Wyatt Technology), differential refractive index (Optilab T- rEX, Wyatt Technology) and UV (SPD-20A, Shimadzu) detectors. M1NT had previously been purified by size-exclusion chromatography as described above. Data were analysed using the ASTRA software package (Wyatt Technology).

Nuclear Magnetic Resonance spectroscopy

Resonances of M1NT were assigned using BEST versions of 3D HNCA, HNCOCA, HNCACB, HNCOCACB, HNCO, and HNCACO spectra (Schanda et al., 2006) at 750 MHz (^{1}H) collected on ^{13}C, ^{15}N labelled protein (150 µM) at pH 5.5. Backbone amide assignments at pH 7.0 were obtained by titrating the pH from 5.5 to 7.0. T_1, T_2, heteronuclear NOE experiments were collected at 600 MHz (^{1}H) using ^{15}N-labelled M1NT(L4A) at pH 5.5. NMR data were processed using NMRPipe (Delaglio et al., 1995) and analysed using CCPNmr Analysis (Vranken et al., 2005) or CARA (Keller, 2005). Proton chemical shifts were referenced to DSS and ^{15}N and ^{13}C chemical shifts were referenced indirectly to proton.

In vitro Oligomerization of M1

Cholesteryl hemisuccinate (CHS) stock solutions (2%) were made up in 5% solutions of CHAPS. Solutions were sonicated until the solution became translucent. Protein concentrations were 30-40 µM M1NT, M1(1-224) and other mutants. Oligomerization was induced by the addition of 10-40 µM CHS at 30°C.

Ain't about how fast I get there.
Ain't about what's waiting on the other side.
It's the climb.

— Miley Cyrus

3

M1 N-Terminal Domain Structure & Dynamics

Contents

3.1 Expression and Purification of M1 28
3.2 NMR Characterisation of M1NT 29
3.2.1 Backbone Resonance Assignment of M1NT 29
3.2.2 Secondary Structure of M1NT 33
3.3 Dynamics of M1NT 34
3.4 Effects of Changes in pH and Ionic Strength 37
3.4.1 Effects of Changes in pH 37
3.4.2 Effects of Changes in KCl 44

3.1 Expression and Purification of M1

Several attempts at the expression and purification of the full-length M1 protein were made. Although M1 is abundant in the cell lysate, M1 purification posed a difficult challenge. Various methods of purification were attempted from fusion tags to refolding purification. However, these efforts were unsuccessful. A divide and conquer approach was then taken to express and purify the N-terminal and C-terminal domain.

The M1 protein is a 28 kDa protein, 252 residues long with a well-folded

N- terminal domain (M1NT) and an unstructured C-terminal domain (M1CT) (Shtykova et al., 2017). The M1NT domain spans from residue 1 to 164, with a molecular weight of 18 kDa while the M1CT domain encompasses residues 164 to 252 residues with a molecular weight of 10 kDa. The M1NT protein is consist of two 4-helix bundle subdomains (helices 1-4 and 6-9) connected by cross-over helix (helix 5) and an alpha-helical C-terminal domain.

Expression of the M1NT and M1CT was done using a bacterial *E.coli* expression system. All constructs produced had polyhistidine tags (his-tag) at the N-terminus unless indicated otherwise. The his-tagged M1 N-terminal domain (M1NT) was successfully expressed and isolated from *E.coli* BL21(DE3) cells. The expression of M1NT was confirmed by SDS-PAGE of cell extracts. Cells were harvested after 11-13 hours and lysed by a cell homogenizer. Purification of M1NT was carried out using affinity chromatography. After Ni-affinity chromatography, M1NT was passed through a gel filtration size-exclusion column (SEC) (Figure 3.1). The results in figure 3.1 shows a single band between 14-28 kDa corresponding to the molecular weight of M1NT following SEC.

Using the program ProtParam, and the known amino acid sequence of the M1NT, including the His tags, the theoretical molecular weight was calculated to be 19.4 kDa. This calculation is consistent with the SDS-PAGE result shown in figure 3.1. A sample of M1NT was adjusted to pH 7.0 and determined to be monomeric by size-exclusion chromatography with multi-angle light scattering (SEC-MALS). The average size of M1NT monomer 19.0 +/- 0.3 kDa at a concentration of 12.5 to 50 μM (Figure 3.2). This is consistent with previous reports for both recombinant M1 (Arzt et al., 2001) and M1 extracted from viruses (Shtykova et al., 2017).

3.2 NMR Characterisation of M1NT

3.2.1 Backbone Resonance Assignment of M1NT

Recombinant M1NT was produced using *E.Coli* expression system and labelled with ^{15}N and 13^{13}C isotopoes in minimal media. A 2D ^{1}H-^{15}N hetero single quantum coherence (HSQC) experiment of ^{15}N uniformly labelled his-tagged M1NT was

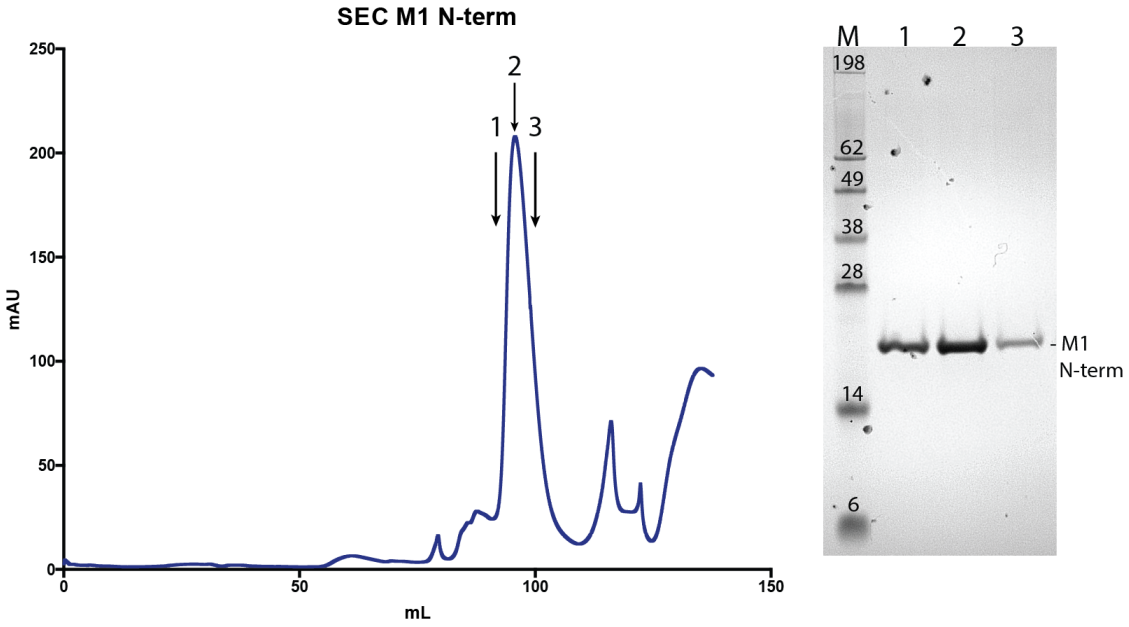


Figure 3.1: Purification of M1 N-term. After the BL21 (DE) *E. coli* cells were allowed to express the protein in pCold vector for 11-13 hours, they were lysed and spun down. The supernatant was run through the affinity chromatography (HIS) column, and the proteins were purified. After nickel purification, M1NT was further purified by size exclusion chromatography (SEC). The chromatogram is shown above, including the analysis by SDS-PAGE gel of the fractions marked 1,2 and 3. The elution contains a single band which corresponds to the expected molecular weight of M1NT - 19kDa.

then collected (Figure 3.3). Each cross peak represents a peak for each unique amide proton attached to ^{15}N, a bonded N-H pair in a peptide bond, essentially representing an amino acid in the sequence except for proline residues. The number of peaks shown in the spectrum are roughly consistent with the expected number of peaks. Furthermore, the peaks exhibited uniform linewidths and good dispersion for an α-helical protein.

In addition, the NH_2 groups of asparagine and glutamine amino acids can also be observed. The NMR spectra of M1NT at both pH 7.0 and 5.5 were well-dispersed for an alpha-helical protein and indicated a well-folded domain as expected from crystallographic studies.

The first and often the most time-consuming step towards characterising a protein's structure by NMR is the assignment of the backbone residues. The low maximum concentration of M1NT concentration determined to be the maximum before seeing aggregaration i.e. 200 μM) also presents a challenge. Analysis of

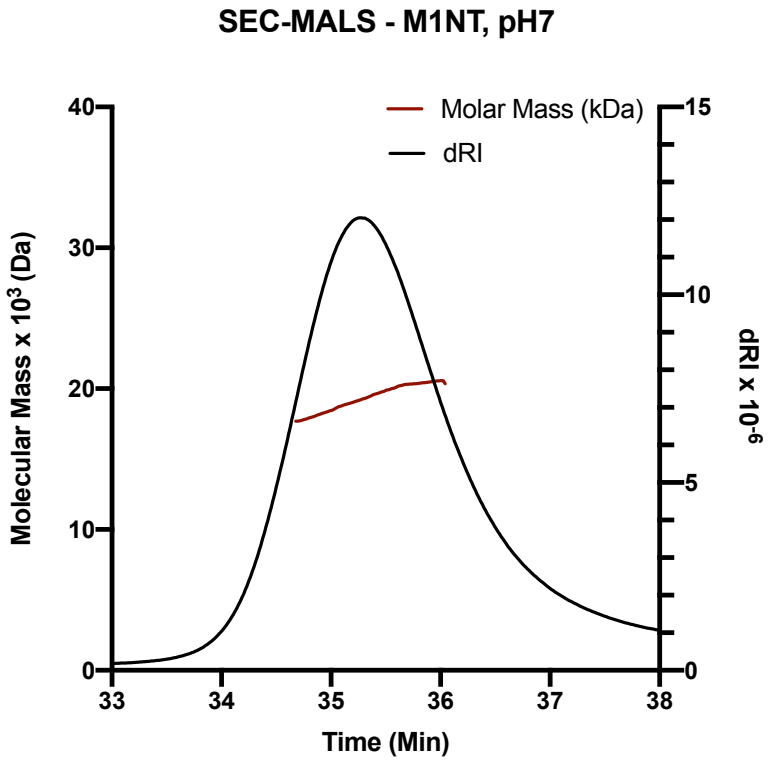


Figure 3.2: MALS experiments were performed during size exclusion chromatography on an analytical Superose 6 10/300 GL column (GE Healthcare) equilibrated with 50mm HEPES 300mM NaCl 10% glycerol, 5mM MgSO4 pH8.0. Elution was monitored via online static light-scattering (DAWN HELEOS 8+, Wyatt Technology), differential refractive index (dRI) (Optilab T- rEX, Wyatt Technology) and UV (SPD-20A, Shimadzu) detectors. Proteins had previously been purified by size-exclusion chromatography. Data were analysed using the ASTRA software package (Wyatt Technology).The elution contains a single band which corresponds to the molecular weight of M1NT - 19KDa.

proteins by NMR has traditionally required millimolar concentrations of highly pure protein. Using the expression and purification strategy that has been developed as described in Chapter 2, a high enough concentration of ^{15}N, ^{13}C-labelled protein samples were produced at a sufficiently high concentration (200 μM).

Extensive backbone assignment (85%) of M1NT was achieved using the standard suite of triple resonance experiments on a combination of ^{15}N, ^{15}N and ^{13}C labelled samples collected at 600 and 750 MHz. Sequential amino acid residues were initially assigned by establishing the connectivity of α-carbon (CA) resonances for i and *i–1* residues using HNCA and HN(CO)CA spectra. These connections were then validated using resonance information from the carbonyl (CO) and beta-carbon (CB) carbon experiments. In addition to triple resonance experiments, the connectivity of

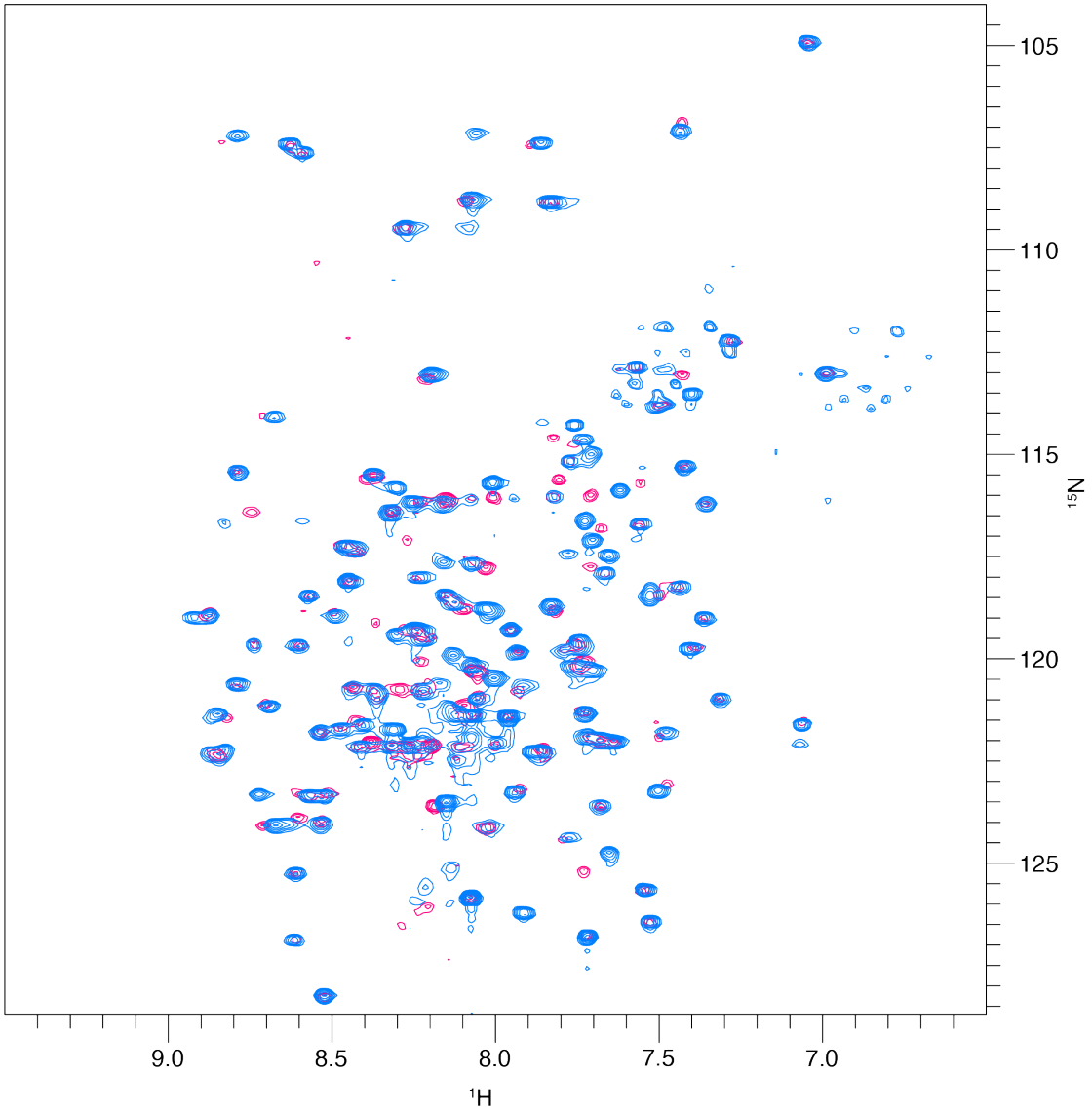


Figure 3.3: Overlay of ^{1}H-^{15}N TROSY spectrum of M1 N-term in 20 mM phosphate, 50 mM NaCl, 0.5 mM TCEP at pH 5.5 (red) and pH 7 (blue) collected on a 750 MHz (^{1}H) spectrometer.

M1NT backbone resonances was confirmed using sequential NOEs in ^{15}N NOESY-HSQC spectra, and the identity of amino acid residues was confirmed by the assignment of side-chain hydrogen atoms using ^{15}N TOCSY-HSQC.

Signals were missing or too weak to assign for residues in and around the loops between helices 6 and 7 and between helices 8 and 9 (Figure 3.4). This is due to the highly flexible nature of these residues.The near-complete backbone assignment of M1NT represents a crucial milestone in the study of this protein as it is an

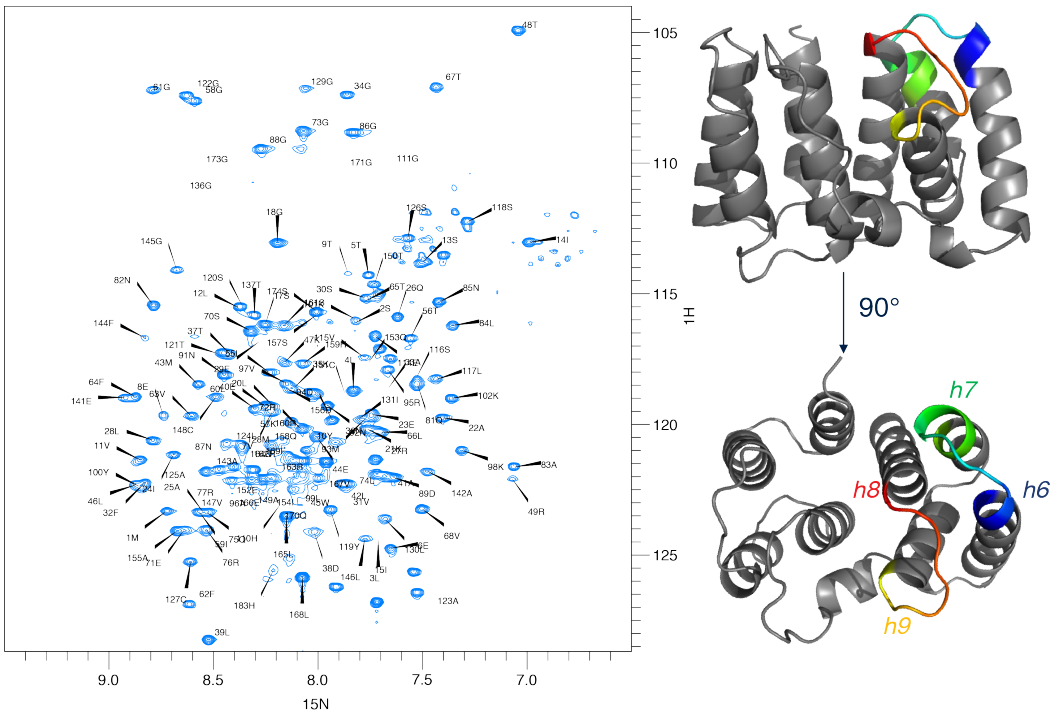


Figure 3.4: Assigned ^{1}H-^{15}N TROSY spectrum of M1 N-term in 20 mM phosphate, 50 mM NaCl, 0.5 mM TCEP at pH 5.5 collected on a 750 MHz (^{1}H) spectrometer (left). Only 85 % of the peaks have been assigned due to weak or missing signals for residues in and around the loops of helices 6 and 7, and helices 8 and 9 as shown on the M1NT structure (right).

opportunity for studying oligomerisation and other interactions in solution.

3.2.2 Secondary Structure of M1NT

TALOS is a database system for the empirical prediction of protein ϕ and ψ backbone torsion angles by using many different chemical shift assignments for a protein. TALOS's approach is derived from the observation that secondary chemical shifts are highly correlated to the type of protein secondary structure. TALOS makes quantitative predictions using secondary chemical shifts and the protein sequence to predict the protein backbone torsion angles ϕ and ψ with measured uncertainty. Through the ϕ and ψ angles, the secondary structure of the protein can be predicted.

Chemical shift-based analyses (using version TALOS-n) of the observable resonances indicated that the secondary structure correlated well with what is

seen in crystal structures (PDB: 1EA3 & 1AA7) (Figure 3.6). The backbone ϕ and ψ angles determined from chemical shifts (Shen and Bax, 2013) show that M1NT contains 8 helical regions (Figure 3.5). The first subdomain of the helical domain comprised the first four helices: *h1* (residues 2-14), *h2* (residues 18-36), *h3* (residues 38-48), *h4* (residues 54-67). The second subdomain is then followed by the cross-over helix, *h5* (residues 69-106) and the second sub-domain: *h6* (residues 90-105), *h7* (residues 114-119) , *h8* (residues 121-131) and *h9* (residues 141-174).

The secondary structure prediction from chemical shifts by TALOS correlates well with the known X-ray crystal structure, PDB ID: 1AA7 (Sha et al., 1997) as shown in figure 3.6. In TALOS prediction, helix 2, 3 and 9 extend from residues 18-36, 38-48 and 141-174, respectively. As compared to the X-ray structure which only extends only to 18-32, 38-47 and 141-157 residues for helix 2, 3 and 9, respectively. TALOS has predicted that the cross-over helix, *h5* spans from residues 70-76 and 79-88 in contrast with the X-ray structure which shows that residues 78-83 and 86-88 covers *h5*. Additionally, TALOS prediction and X-ray structure show that *h7* covers residues 114-119 and 109-118, respectively.

3.3 Dynamics of M1NT

A major caveat of X-ray crystallography is that it provides only a snapshot of a protein. In contrast NMR is able to capture the dynamics of a protein in solution at the picosecond to second time scale. Dynamics at this time scale can be characterised through the spin relaxation of ^{1}H, ^{13}C or ^{15}N. Through analysis, parameters such as R_1, R_2 and ^{1}H-^{15}N heteronuclear NOE (hetNOE) can be extracted to characterise the flexibility of the protein backbone. R_1 is described as the longitudinal relaxation rate and R_2 as the transverse relaxation rate (Kay et al., 1989). The hetNOE is the steady state Nuclear Overhauser Enhancement which results from the dipolar cross relaxation of the ^{15}N nucleus that is attach directly to the H proton of an amide (Kay et al., 1989).

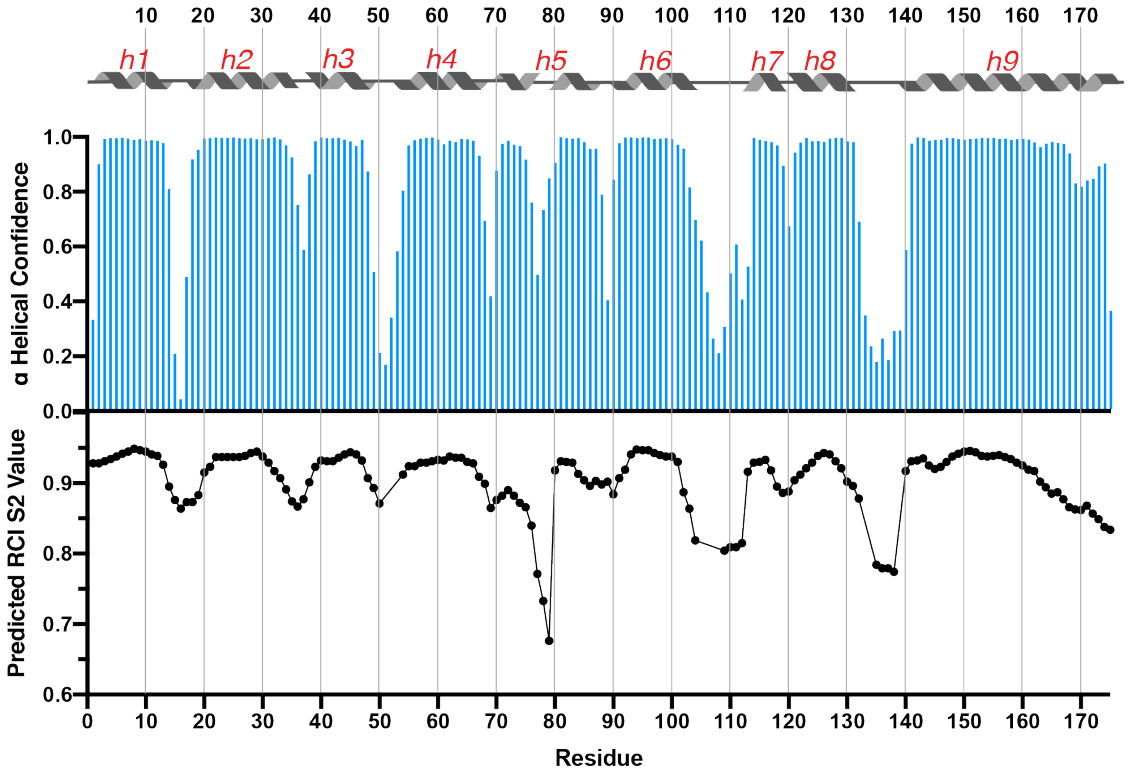


Figure 3.5: Secondary structure derived from TALOS-n. The backbone chemical shifts were analyzed using TALOS-n, and the derived secondary structure information was shown. The alpha-helical and coil regions are labelled as helices and solid lines, respectively.

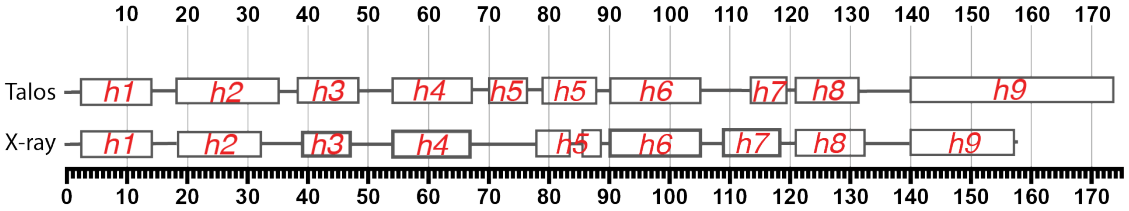


Figure 3.6: Comparison of the X-ray crystallographic M1NT secondary structure, 1AA7 (Sha et al., 1997) with experimentally observed and predicted structures using TALOS-n. The secondary structure elements are indicated with a box for alpha helices and a solid line for coils.

To evaluate the internal dynamics of M1NT, the backbone amide R_1 and R_2 relaxation rates were measured at 600 MHz using the M1NT construct with a stabilising L4A substitution, which prevents the protein from aggregating contributing to protein stability. It was not possible to produce a wildtype M1NT of sufficiently high concentration for NMR dynamic experiments (around 200μM). The ^{1}H-^{15}N heteronuclear NOE was also measured at 600 MHz.

As seen in figure 3.7, decreased ^{15}N R_2 and ^{1}H-^{15}N-heteronuclear NOEs indicated

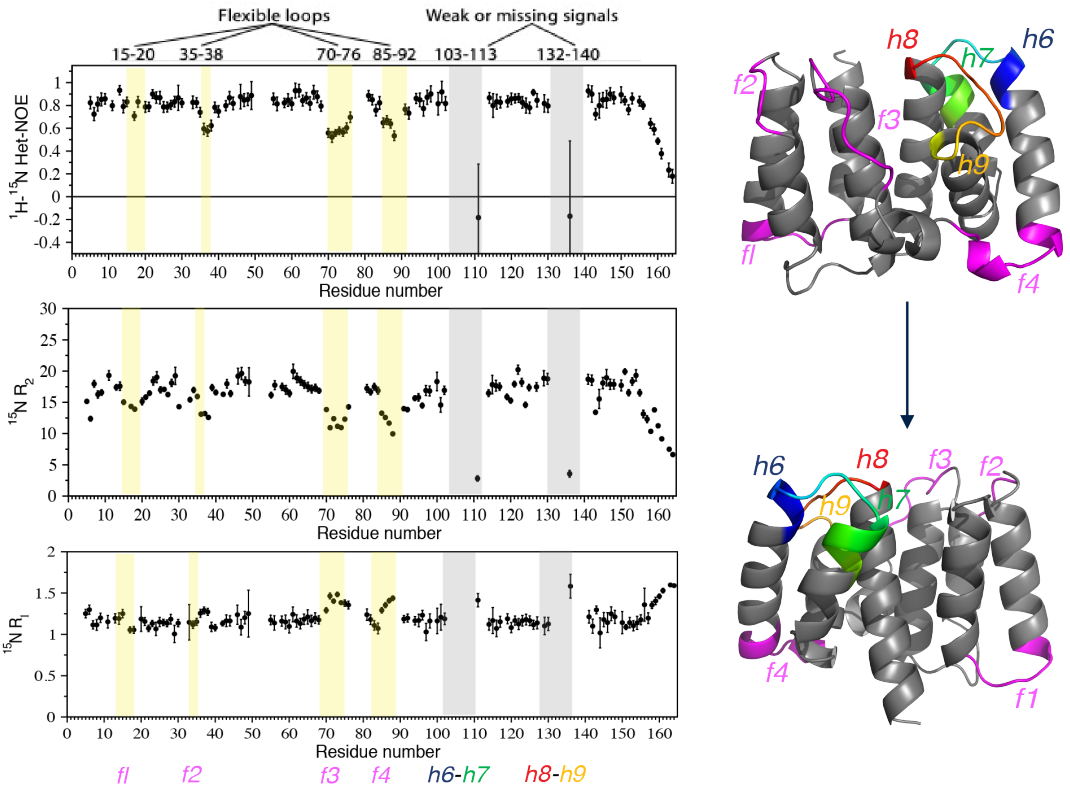


Figure 3.7: Assigned ^{1}H-^{15}N TROSY spectrum of M1 N-term L4A in 20 mM phosphate, 50 mM NaCl, 0.5 mM 2-mercaptoethanol at pH 5.5 collected on a 601 MHz (^{1}H). Only 85 % of the peaks are assigned due to weak or missing signals for residues in and around the loops of helices 6 and 7, and helices 8 and 9. The flexible loops are indicated in magenta from *f1* to *f4*.

flexibility in the loops between helices 1 to 2 (residues 15-20) and helices 2 to 3 (residues 35-38) as well as the regions connecting the cross-over helix 5 (residues 70-77 and 85-92). The relaxation properties of the glycines that could be assigned in the loops between helices 6 to 7 (residues 103-113) and between helices 8 to 9 (residues 132-140) confirmed that these regions are highly flexible. These regions are defined as highly flexible by TALOS-n automatically by a threshold figure.

Residues 69 to 77 are known to be disordered in M1 crystal structures in both neutral and acidic pH indicated by high B-factors (Harris et al., 2001; Safo et al., 2014) as shown in figure 3.8, labelled *f3*. Residues in the loops between *h6* and *h7* are found to be dynamic through NMR dynamic experiments. However, the low B-factor in the X-ray crystal structure of M1NT (PDB ID:1AA7), shows that this region is not flexible or dynamic. This can be explained by looking at the crystal contacts,

indicating that crystal contacts with an adjacent monomer stabilises this loop region as shown in figure 3.9. Residues 103-113 of the loop between helix 6 (*h6*) and helix 7 (*h7*) are stabilised by the residues in helix 1 (*h1*) and helix 9 (*h9*) as shown in figure 3.10. There is an electrostatic interaction between H110 with D156 of *h9* from an adjacent monomer and a hydrophobic interaction between I107 with Y10 of *h1*.

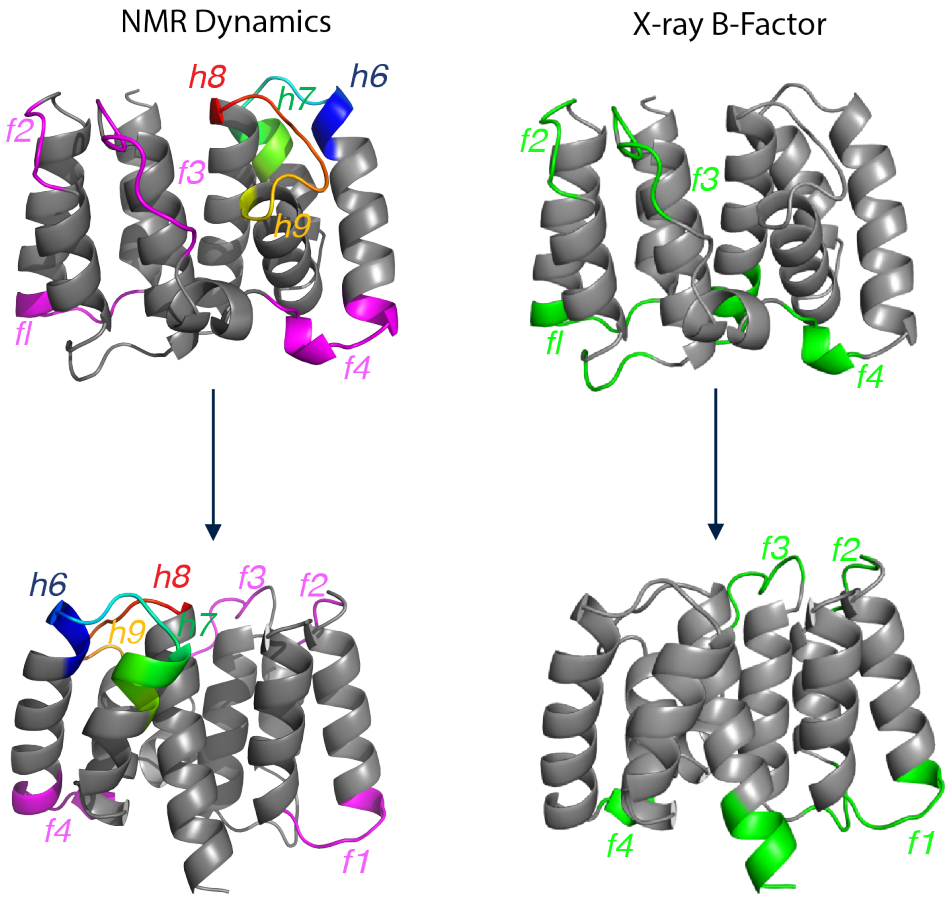


Figure 3.8: The M1NT domain structure (on the left) showing regions of flexibility by relaxation experiments (T1, T2 and NOE). The structure (1AA7) on the right shows an elevated X-ray B-factors of flexible regions. As depicted above (left), relaxation experiments indicate that the H6-H7 and H8-H9 loops are flexible while in contrast, the X-ray structure (right) show that these regions are not flexible.

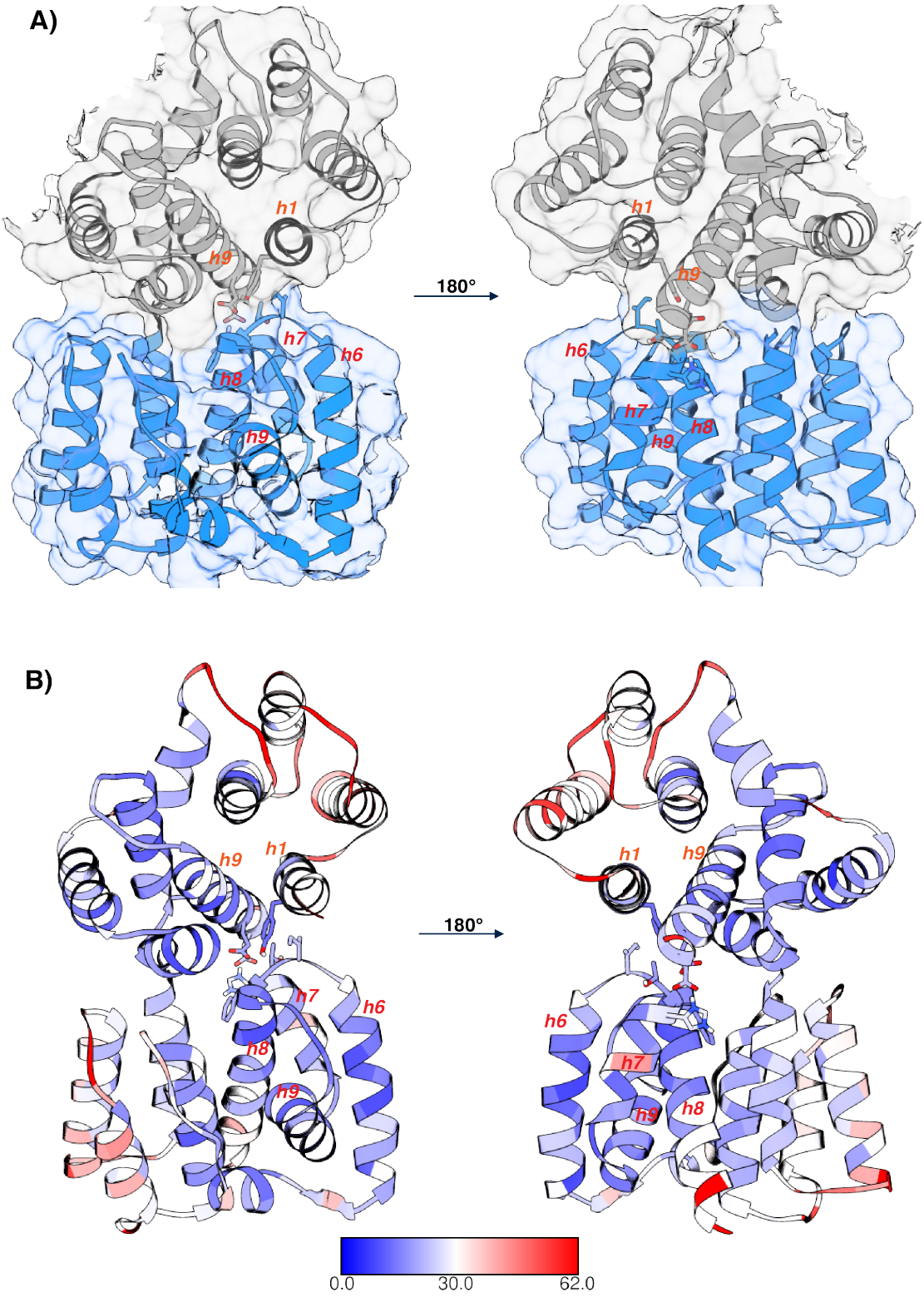


Figure 3.9: Crystal structure of M1NT, 1AA7 of A) monomer of a dimer (blue) with crystal contacts with an adjacent monomer (grey). B) A B-factor representation of the crystal contacts. The dynamic region of the loop region of helix 6 (*h6*) to helix 7 (*h7*) has a lower B-factors due to crystal contacts with the monomer of an adjacent molecule stabilised by the residue on helix 1 (*h1*) and helix 9 (*h9*).

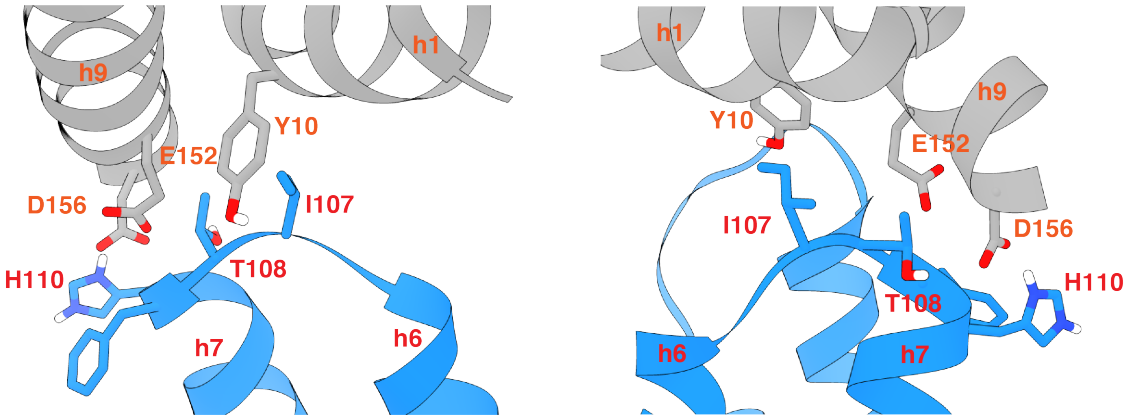


Figure 3.10: Close up of the crystal contacts that stabilised the dynamic region of the loop of helix 6 (*h6*) to helix 7 (*h7*) by the residues of the adjacent monomer of both helix 1 (*h1*) and helix 9 (*h9*). PDB: 1AA7

3.4 Effects of Changes in pH and Ionic Strength

3.4.1 Effects of Changes in pH

Several studies have reported that M1 has a strong tendency to oligomerise (Bui et al., 2000; Ruigrok et al. 2000; Calder et al., 2010) and form an organised structure underneath the virus envelope which is needed for virus assembly and budding (Ali et al., 2000; Avalos et al., 1997; Barman et al., 2001; Gómez-Puertas et al., 2000; Latham & Galarza, 2001; Wang et al., 2010). The interaction between M1 monomers facilitates membrane bending, which causes the asymmetry in the lipid bilayer required for bud initiation. Furthermore, the oligomerisation of M1 is needed for the final stage of budding closure in order for the virions to be released from the membrane (Nayak et al., 2004; Nayak et al., 2009; Rossman & Lamb, 2011; Ruigrok et al., 2000a).

The M1 constructs (M1NT & M1 (1-224)were tested for structural changes via chemical shift perturbations (CSPs) in response to changes in pH from pH 5.5 to pH 7. The CSPs between pH 5.5 to 6.0 and pH 6.5 to 7.0 were modest for both M1 (1-164) and M1 (1-224) (Figure 3.11). However, interestingly, significant CSPs were observed between pH 6.0 to pH 6.5. These changes were mapped to three histidines – 110H, 159H, and 162H, that all reside in helix 9 (*h9*). Residues from helix 1 (*h1*) –

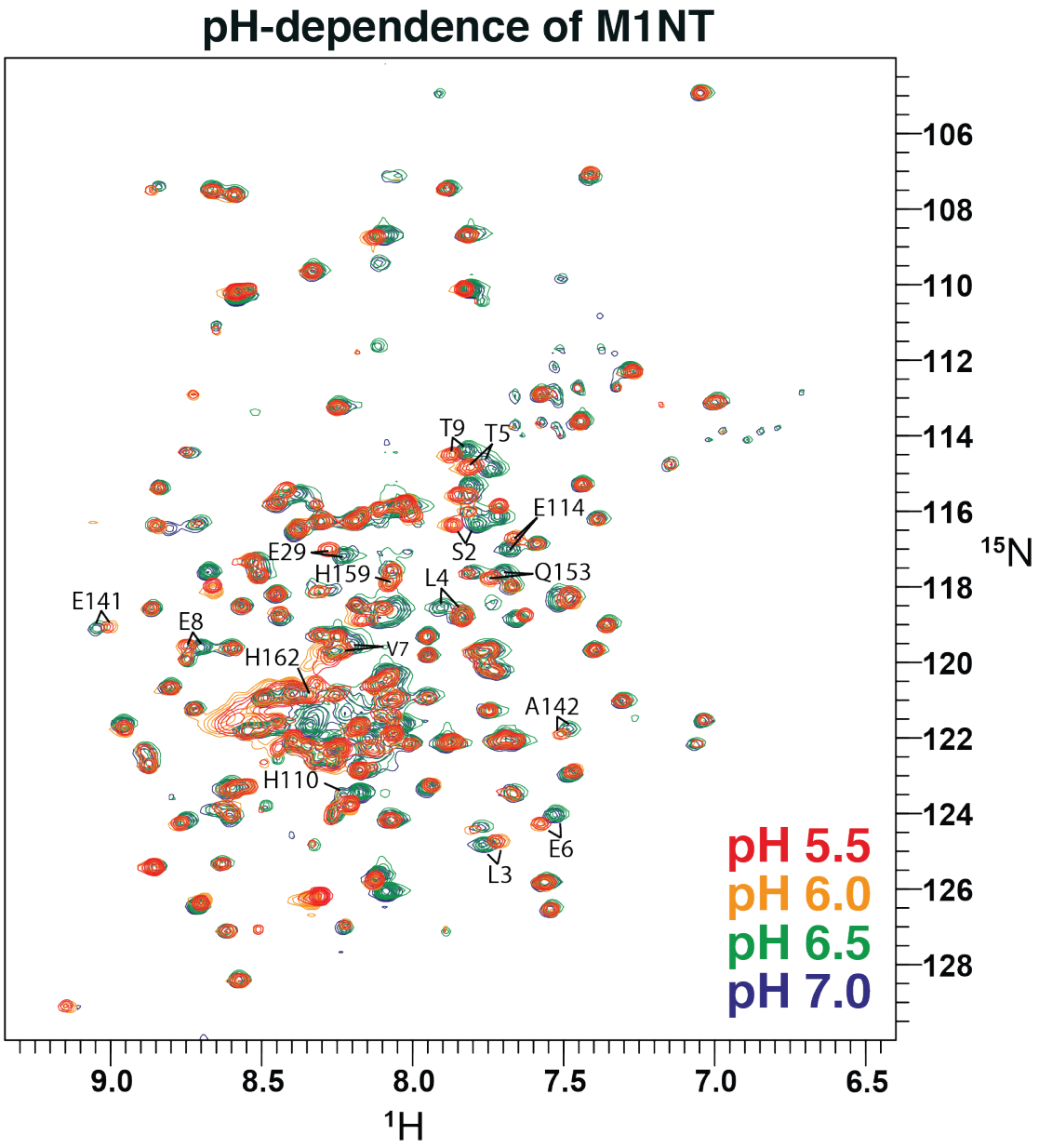


Figure 3.11: Spectral overlays of M1NT construct at a concentration of 50μM in 20mM Tris, 50mM NaCl at pH 5.5 (red), pH 6.0 (orange), pH 6.5 (green) or pH 7.0 (blue) at 15°C. M1NT was prepared at a low pH (pH 5.5) as described previously in Chapter 2. The pH was then gradually increased at 0.5 increments to pH 7.0

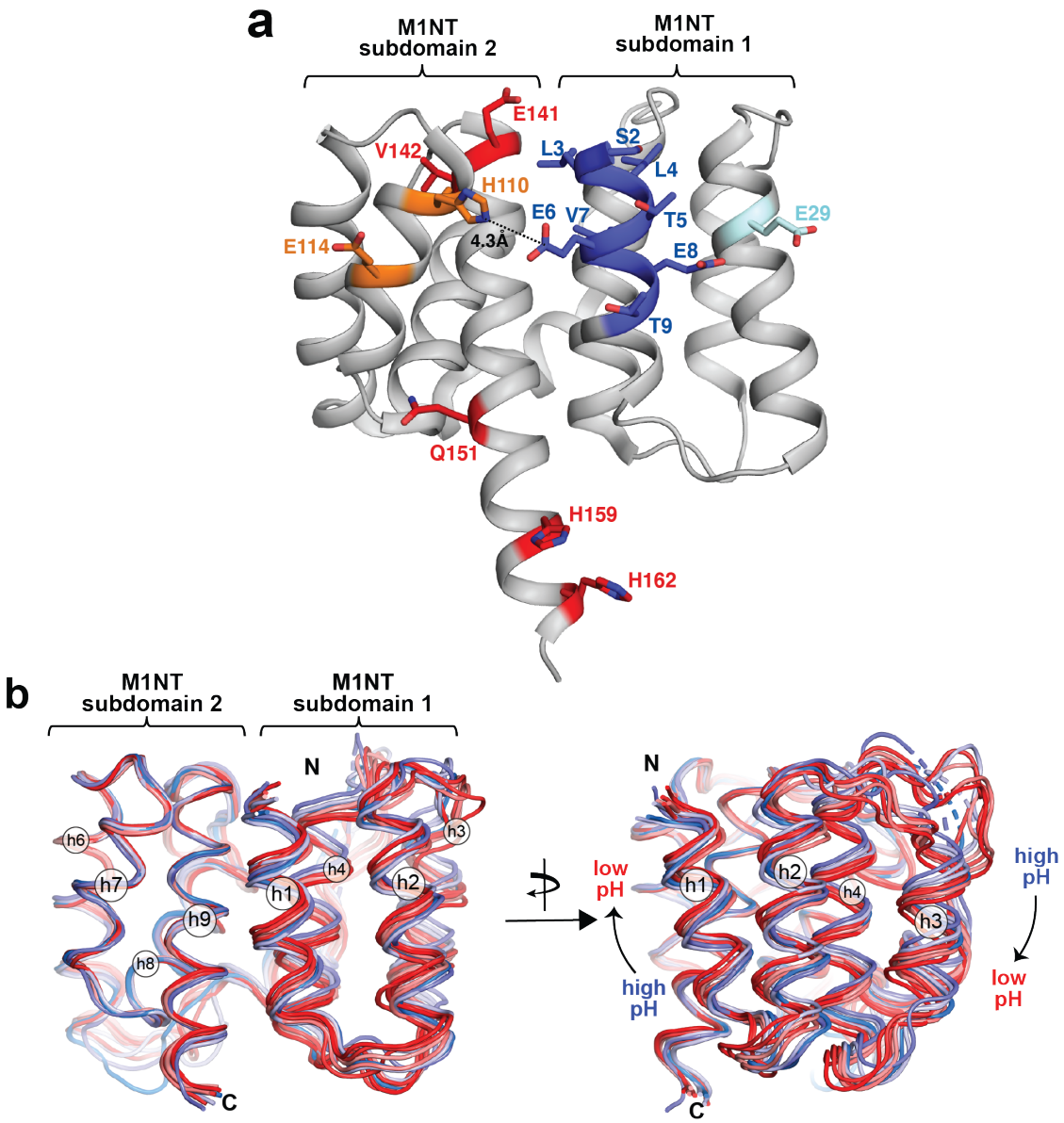


Figure 3.12: Perturbed Residues Upon Change in pH Mapped to M1NT Structure & The Correlation Between pH and Rotation Across Subdomain. a) M1NT structure, 5V7S (Chiang et al., 2017) illustrating the perturbed residues as shown in figure 3.7 mapped to helix 1 (dark blue), helix 2 (light blue), helix 6 (orange), and helix 9 (red). The distance between the $N_{\epsilon 2}$ of H110 in Helix 7 and the Cδ of E6 in Helix 1 is 4.3 Å. b) Overlays of the 9 pre-existing crystals structures derived from either A/Puerto Rico/8/1934 H1N1 or A/Wilson-Smith/1933 H1N1. Structures solved at pH 7 or above are coloured in shades of blue (5V7B, 1EA3, 5V8A) and structures below pH 7 are coloured in shades of red (1AA7, 4PUS, 5CQE, 5V6G, 5V7S). Structures are aligned on helices 6 to 9 of 1AA7 to highlight the pH-dependent relative subdomain rotation shown in the figure.

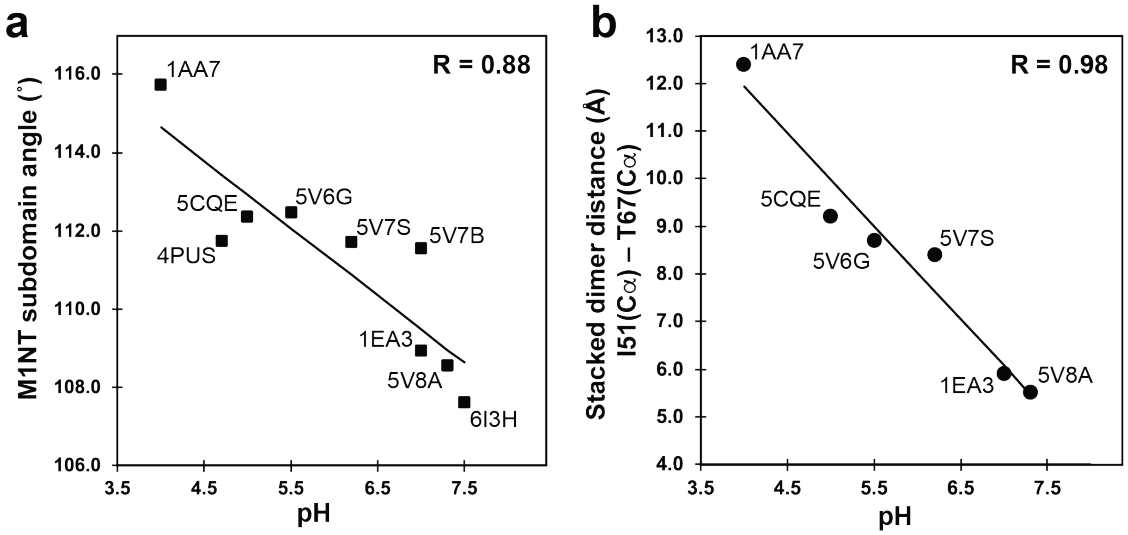


Figure 3.13: M1 - pH Dependence of Intramolecular Subdomain Shifts and Tilting of Stacked Dimer Arrangements. a) The angles between the M1NT subdomains have been plotted against pH for all 9 structures (PDB: 1AA7; 4PUS; 5CQE; 5V6G; 5V7S; 5V7B; 1EA3; 5V8A; 6I3H) from A/Puerto Rico/8/1934 H1N1 or A/Wilson-Smith/1933 H1N1. b) Plot of the intermolecular Cα-Cα distance between I151 and T67 as a function of pH for the six structures derived from A/Puerto Rico/8/1934 H1N1 or A/Wilson-Smith/1933 H1N1 that exhibit the stacked dimer interface. The linear regression and the R-value are shown in both figures 3.13a and 3.13b.

S2 to T9, were also perturbed (Figure 3.12a). As can be seen in Figure 3.11), M1(1-164) and M1(1-224) remained monomeric from low (5.5) pH to neutral (7.0) pH. In addition, the C-terminal domain in the M1(1-224) construct remains flexible and unfolded. Extending the C-terminal domain in the M1(224) construct does not affect the oligomerisation state - it remains a monomer in solution at low and neutral pH. However, the solubility of M1(224) decreases at pH 7.0 when compared to M1NT.

Residues highly perturbed due to pH changes from pH 6.0 to 6.5 reside at *h1* and *h9* which may suggest that changes in pH cause shifts in the intermolecular domain between the two subdomains – subdomain 1 (*h1* to *h4*) and subdomain 2 (*h6* to *h9*) (Figure 3.12b). X-ray structures of M1 (Harris et al., 2001; Safo et al., 2014; Sha & Luo, 1997) shows that there are small differences between neutral and acidic pH structures within a single protein monomer. These differences are mostly attributed in the loop regions. However, the interactions between monomers are completely different. Further analyses of M1NT structures solved over a range of pH were done to determine if there was a correlation between pH changes and

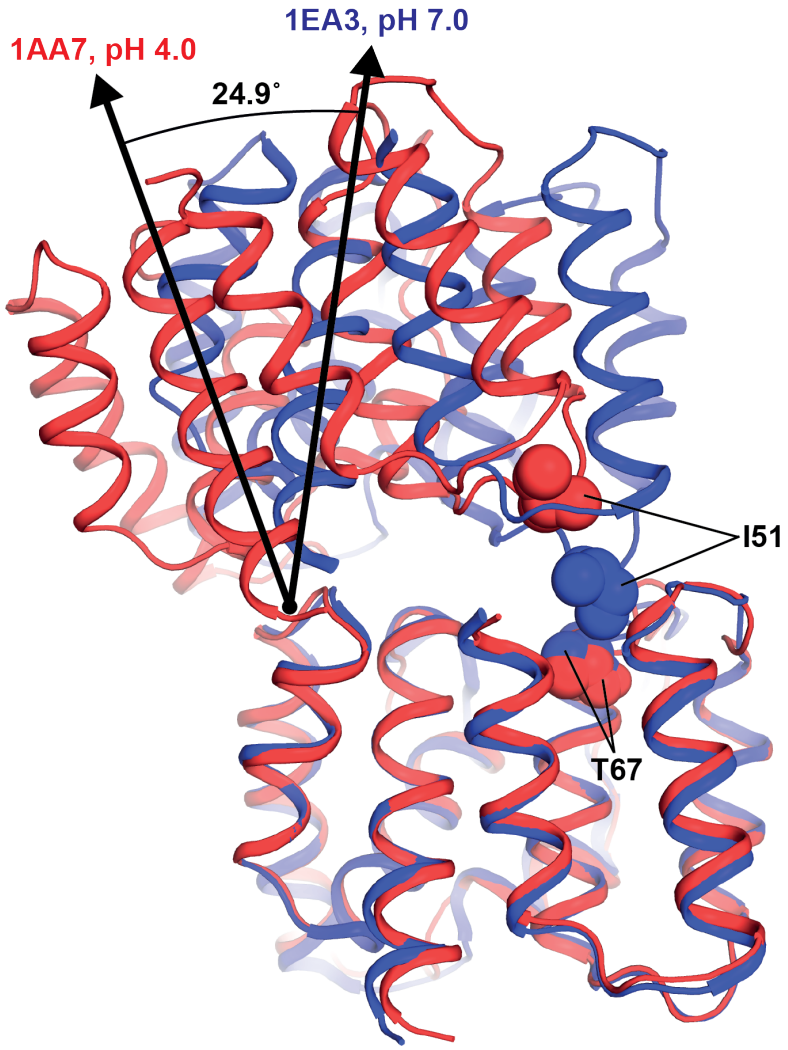


Figure 3.14: Structural alignment of the M1NT stacked dimers observed in 1AA7 (red) and 1EA3 (blue) which illustrates in the rotation of helix 9 by $\sim 25°$C. The subunit tilting results in the largest changes at the subdomain 1 interface, highlighted by the different distances between I51 and T67, represented as spheres.

subdomain movements. M1NT crystal structures deposited in the Protein Data Bank, of which 9 structures out of 11 are derived from one of the two viral strains – A/Puerto Rico/1934 H1N1 (PR8) or A/Wilson-Smith/1933 H1N1 (WSN). These two strains differed by two conservative substitutions - I14V & S115A. The 9 structures were structurally aligned on subdomain 2 and intriguingly there was a correlation between the pH and rotation of subdomain 1. The rotation of the subdomains were quantified and plotted as a function of pH (Figure 3.13b).

Earlier studies carried out by Harris and colleagues have observed the tilting

between stacked dimers that was observed on M1NT structures that were determined at neutral (pH 7.0) and acidic pH (pH 5.5) (Harris et al., 2001) (Figure 3.14). Out of the 9 structures that were analysed previously, 6 of them have a stacked dimer arrangement. The tilt between stacked dimers was quantified by measuring the subdomain 1 distance between I51 of one subunit of one monomer and T67 of another subunit (Figure 3.14). These residues were chosen due to the stacked dimer is formed from two interfaces, and in subdomain 1 the interface is centered on I51, which is in a hydrophobic loop and protrudes into a shallow concave surface near T67 on the adjacent subunit. Hence, this can be sorted according to whether the subdomain rotation correlates with packing at the I51 interface or the T67 interface. A strong correlation can be observed when this distance is plotted against pH (Figure 3.13b) which is consistent with the study done by Harris and colleagues (Harris et al., 2001).

This tilting of the stacked dimer arrangement seen when pH is decreased may play a role in the physiological context during influenza virus infection. This may relate to the disintegration of oligomeric M1 by possibly destabilising the interaction between M1 monomers which forms the critical step in releasing the viral genome after host infection as the viral particle fuses with the late endosome. The CSPs seen at the subdomain interface may suggest a pH-dependent mechanism at which the oligomeric M1 disintegrates into its monomeric form (Calder et al., 2010; Sha and Luo, 1997; Shtykova et al., 2017; Zhang et al., 2015). The histidines at helix 9, *h9* – H159 and H162, might also play a role in regulating the intermolecular domain tilt as these histidines approach H110 of an adjacent subunit.

3.4.2 Effects of Changes in KCl

A study by Stauffer et al. showed that the Influenza A Virus needs a stepwise process after entry into the host cell for efficient uncoating and infection. The first step involves the M2 ion channel being activated through the exposure of the virus in the early endosomes to a pH between 6.5 to 6.0 with low K^+/high Na^+ concentrations. The second step is triggered through the exposure of the core to a pH of <6.0 to 5.4 and by influx of K^+ ions. This exposure to both high concentration of H^+

and K^+ promote more specificity and thereby increasing efficiency of uncoating and infectivity. Through limited proteolysis (LiP) experiment, the M1 C-terminus was shown to undergo irreversible changes when the virus was exposed to a low pH and K^+ concentration (Stauffer et al. 2014).

M1 constructs were tested for structural changes in response to changes in K^+ concentrations at neutral and acidic pH. The CSPs between the neutral (7.0) and acidic (5.5) pH at different K+ concentrations were modest for the M1(1-224) construct (Figure 3.15 and 3.17). The M1(224) construct composed of the N-terminal domain with an extension of the C-terminal domain lacking only 28 amino acids. The modest change in CSPs as the concentration of KCl increases to 80mM suggests there are no significant or dramatic structural changes in either the N-terminus and C-terminus domains. This is at least true for M1 or M1(224) monomers. There is a possibility that K^+ would affect the M1 structure in the polymerised form or when bound to vRNPs.

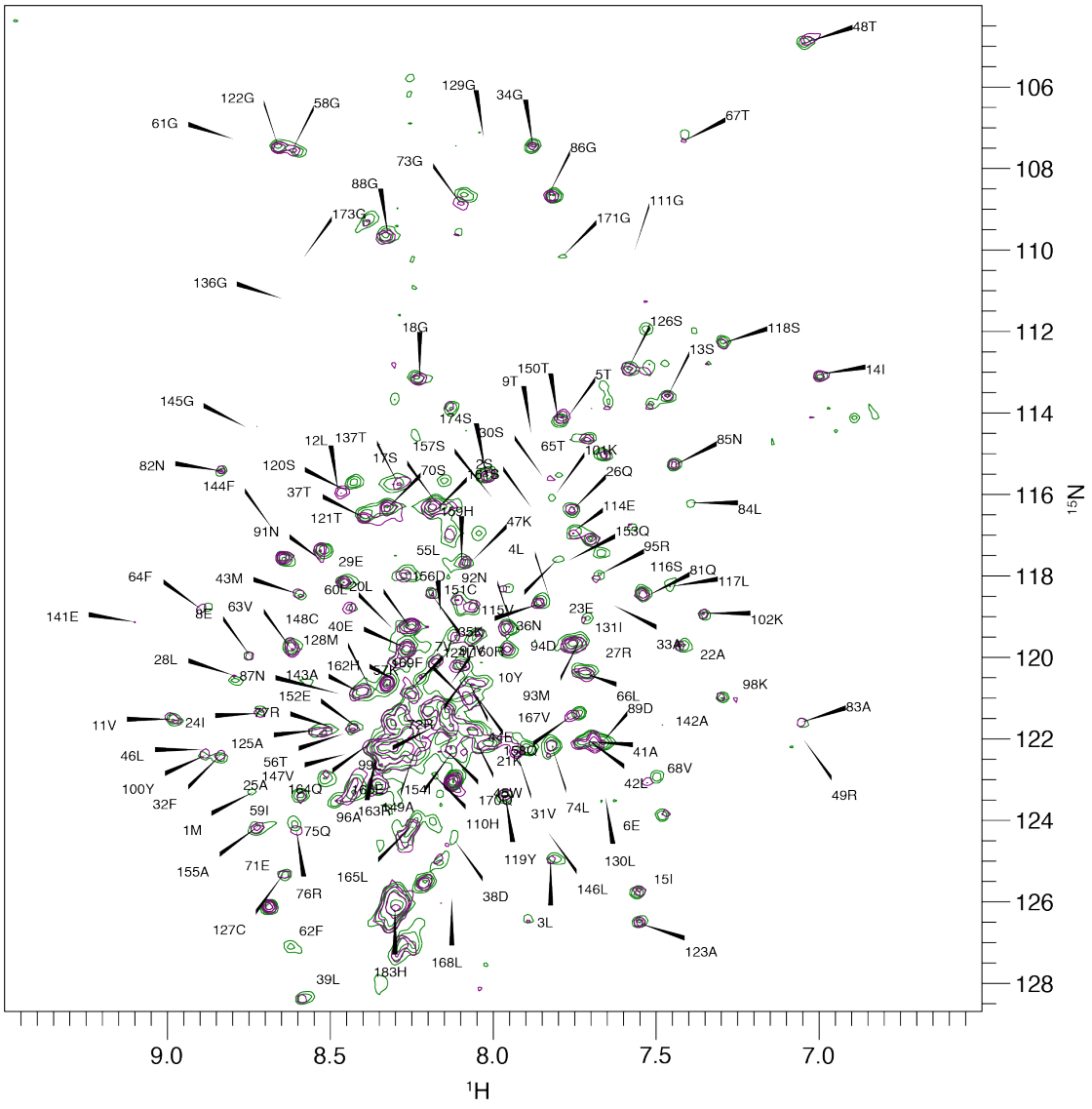


Figure 3.15: Spectral overlays of M1(224), pH 7.0., 50mM NaCl (green) showing chemical shift perturbations upon addition of 80mM KCl (purple). Only modest changes in chemical shift perturbations can be seen. M1 (1-224) was prepared at a low pH (pH 5.5) as described previously in Chapter 2. The pH was then gradually increased to pH 7.

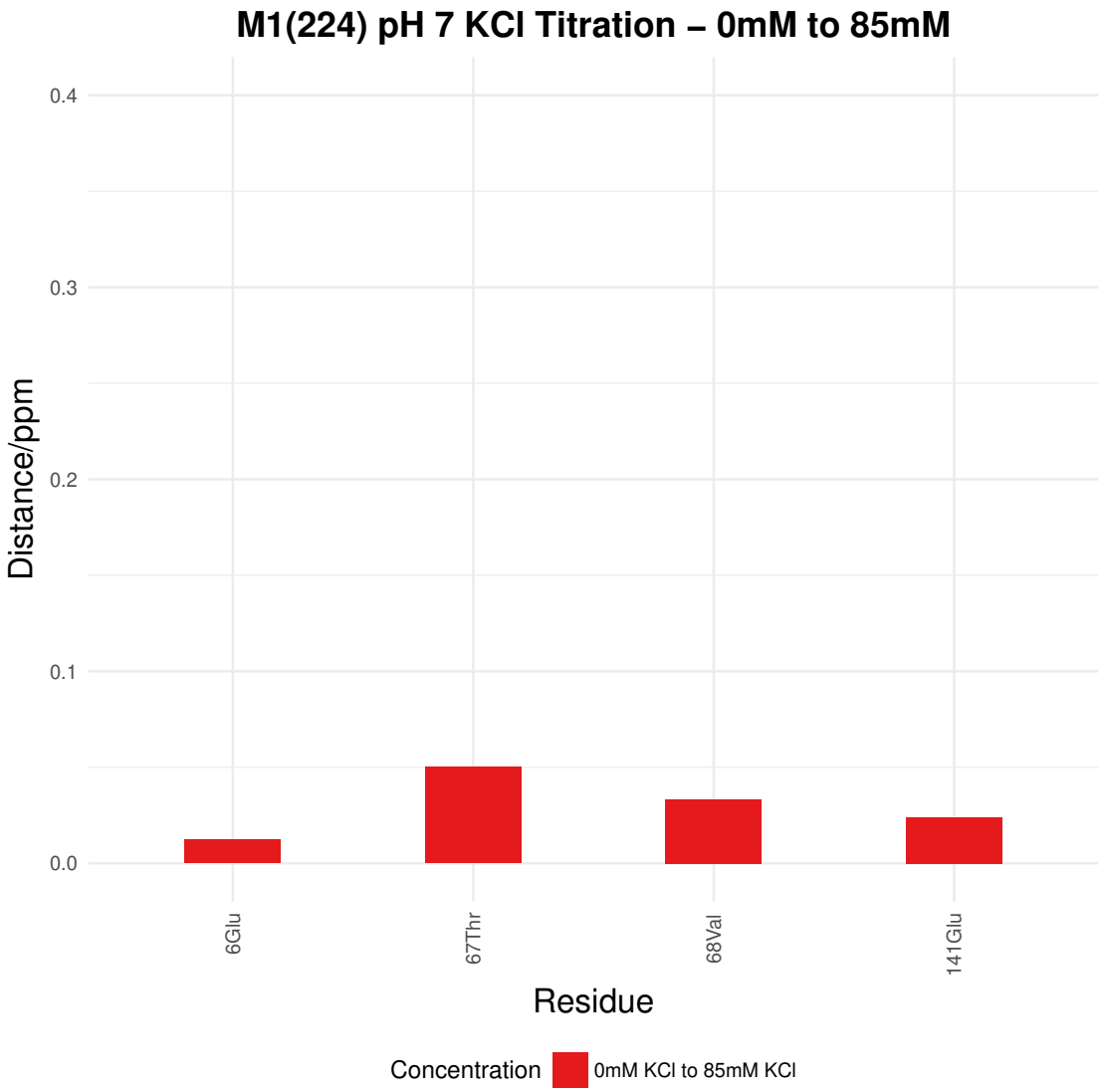


Figure 3.16: Chemical shift perturbations of selected residues of M1(224) at pH 7 after KCl concentration was increased to 85mM from 0mM. These above residues were selected because they were strongly perturbed and well resolved resonances. Only a few small shifts of peaks can be seen to suggest any structural changes.

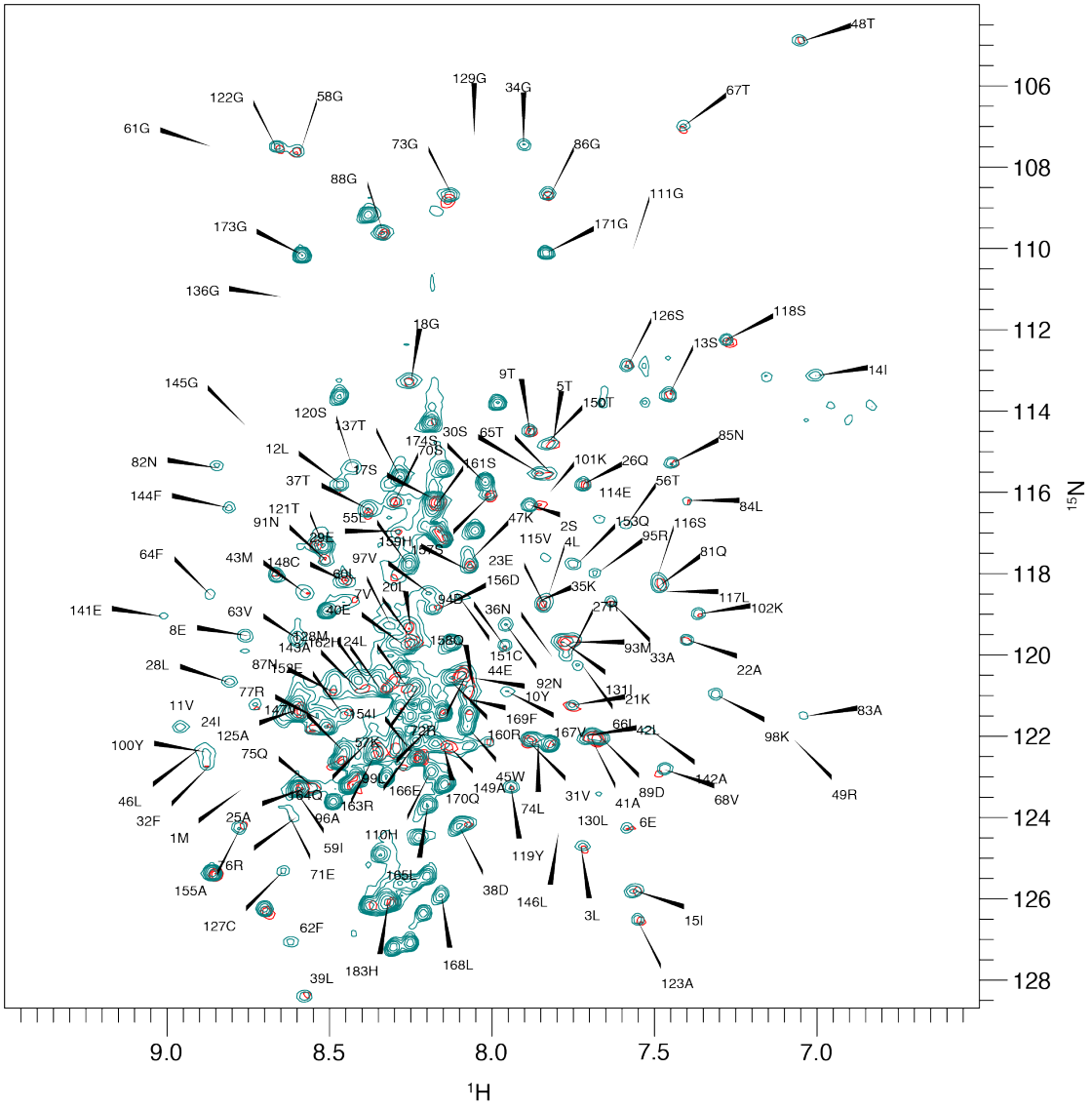


Figure 3.17: Spectral overlays of M1(224), pH 5.5, 50mM NaCl (Teal) showing chemical shift pertubations upon addition of 80mM KCl (Red). As with at pH 7, only modest changes in chemical shift pertubations can be seen.M1 (1-224) was prepared at a low pH (pH 5.5) as described previously in Chapter 2.

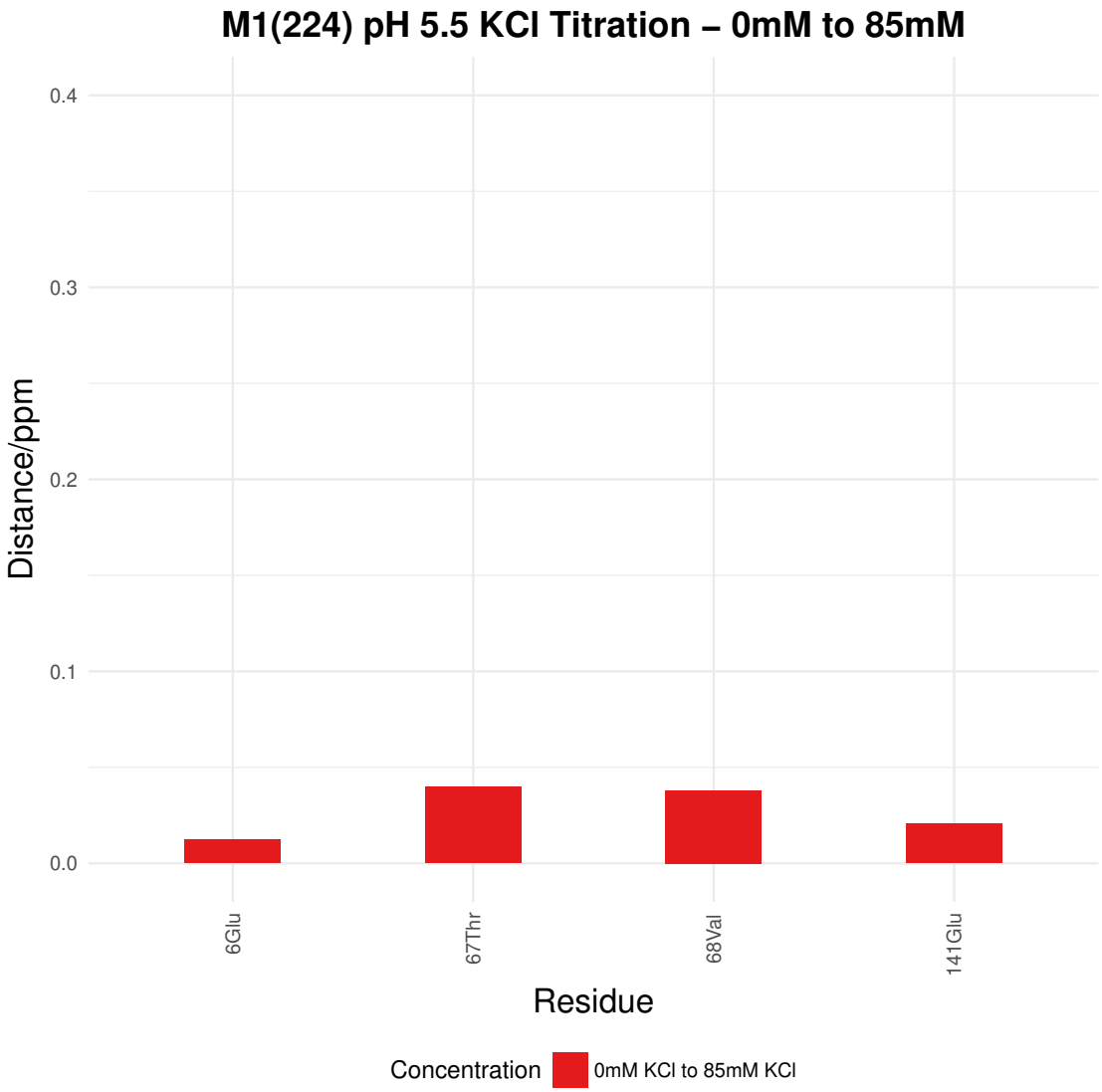


Figure 3.18: Chemical shift perturbations of selected residues of M1(224) at pH 5.5 after KCl concentration was increased to 85mM from 0mM. These above residues were selected because they were strongly perturbed and well resolved resonances. Only a few small shifts of peaks can be seen to suggest any structural changes.

I'll spread my wings, and I'll learn how to fly
I'll do what it takes till I touch the sky
And I'll make a wish
Take a chance
Make a change
And breakaway

— Kelly Clarkson

4 M1 C-Terminal Domain

Contents

4.1 M1 C-terminal Domain Characterisation 50
4.1.1 M1CT Expression and Purification 51
4.1.2 NMR Characterisation of M1CT 53
4.1.3 M1NT with C-terminal Extensions 55

4.1 M1 C-terminal Domain Characterisation

The M1 C-terminal domain (M1CT) is weakly structured, as indicated by biophysical techniques which have been used to characterise M1CT which found that M1CT, finding it to be flexible and intrinsically disordered (Kordyukova et al., 2008; Ksenofontov et al., 2010; Shishkov et al., 2009, 2011; Zhang et al., 2012). Studies on full-length M1 have also shown consistent results (Arzt et al., 2001; Ruigrok et al., 2000). Being highly unstructured may also suggest that M1CT could act as an IDP and fold upon binding with interacting partners. It is highly likely that M1CT has a more folded structure in the oligomerised form of the full-length M1 (Zhang et al., 2017; Peukes et al., 2020). It's important to note that the quantities of full-length M1 used in the above studies are very small as compared to what is needed to study full-length M1 in NMR. Furthermore, the full-length

M1 is purified from viruses in which methods attempted to produce this protein in this thesis uses bacterial expression. The following sections will discuss my attempts to characterise M1CT using NMR.

4.1.1 M1CT Expression and Purification

M1 C-terminal (M1CT) domain is 88 amino acids long and spans the residues 165 to 252 with a molecular weight of 10 kDa. An M1CT construct fused to a His-SUMO tag (his-SUMO M1CT) was expressed using the pFloat vector expression system in BL21 (DE3) *E.coli*. Cells were harvested after 11-13 hours, suspended in lysis buffer and lysed using a homogenizer. The first step of purification of M1CT was carried out using Ni-affinity chromatography (Figure 4.1A). The theoretical molecular weight of his-SUMO M1CT is calculated to be 22.93 kDa. The bands from the fractions of the eluted peak in the nickel purification chromatogram ran at their expected molecular weights, as shown in the SDS-PAGE (Figure 2.1B).

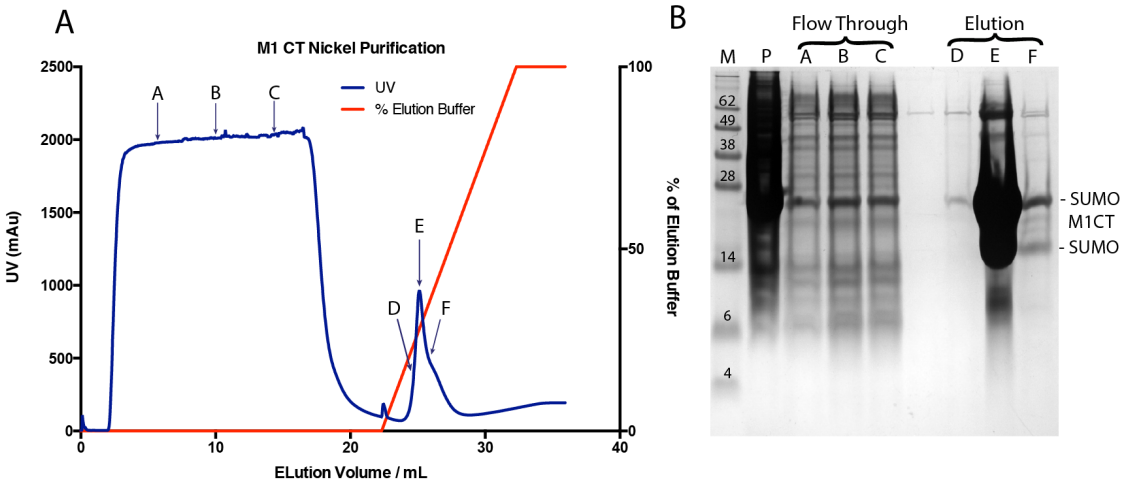


Figure 4.1: Nickel purification of M1CT. A) Nickel purification of M1CT chromatogram. The blue line is the UV trace at 280nm. The red line is the percentage of the elution buffer running through the column. The cell pellet is marked as P shown in B. The flow-through is marked as A, B and C while the eluted protein is marked as D, E and F. B) SDS-PAGE analysis of the fractions collected (A,B and C - Flow-through; D,E and F - Elution fractions).

The fractions from the eluted peak were then pooled, and the SUMO fusion protein is then cleaved by HRV3C protease. These were then run through the NI column to remove the fusion protein (Figure 4.2). The uncleaved and cleaved M1CT

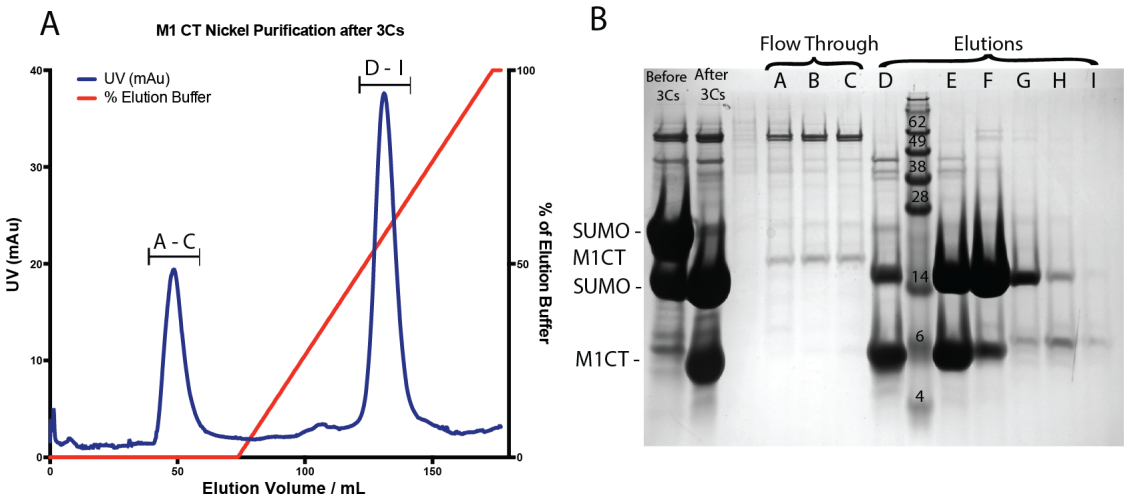


Figure 4.2: Nickel purification of M1CT after HRV3C cleavage. A) The flow-through is marked as A, B and C while the elution fractions are marked as D, E and F. B) SDS-PAGE analysis of the uncleaved and cleaved M1CT as well as fractions collected (A-C - Flow-Through; D-I - Elution fractions) from the Ni purification in Fig. 4.2. It can be observed that SUMO fusion protein eluted together with the cleaved M1CT.

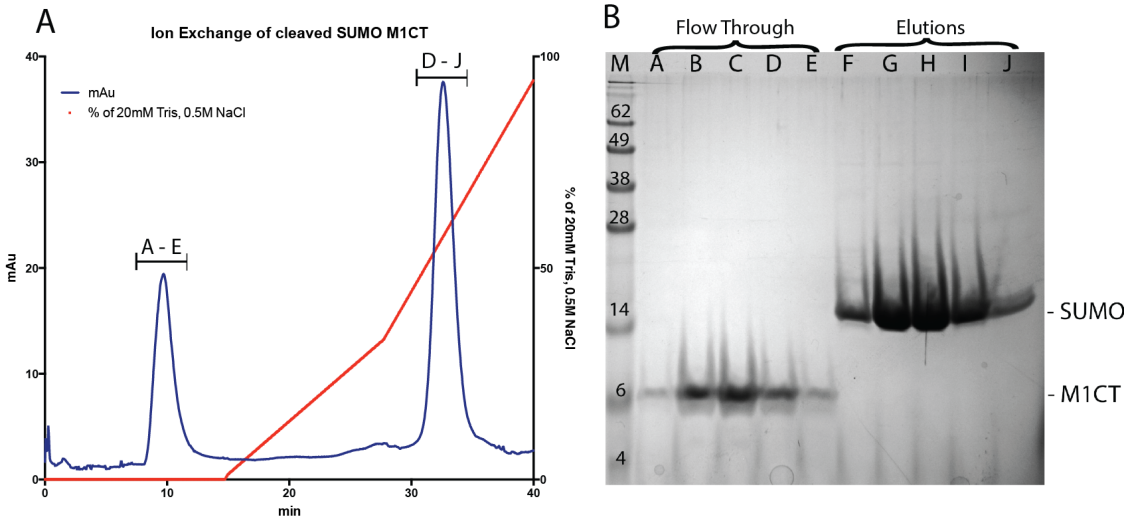


Figure 4.3: Ion exchange (IE) purification of SUMO M1 C-term after HRV3C cleavage. A) IE chromatogram of cleaved M1CT - The blue line represents the UV trace at 280nm, and the red line represents the percentage of the buffer at increasing NaCl concentration. The flow-through is marked as A-E while the elution fractions are marked as F-J. B) SDS-PAGE analysis of the fractions from the flow-through and eluted peaks. The flow-through consisted of pure M1CT as indicated by the single band.

are shown in figure 4.2B, and it can be observed that the band for SUMO is at the expected molecular weight of 12kDa. However, M1CT is not at its expected molecular weight. The theoretical molecular weight M1CT is calculated to be approximately 9.7kDa but appeared to be just above the 6 kDa protein marker. This is probably due to disordered nature of M1CT and do not run at expected positions in SDS-PAGE. The plasmid used for the expression of M1CT was sent for sequencing, and it showed the correct sequence. It should be noted that disordered proteins often do not run at the expected positions in gels or columns due to its unglobular nature of the protein. Another observation seen from the nickel purification of M1CT after HRV3C cleavage is the absence of cleaved M1CT in the flow-through (FT) where cleaved M1CT was expected to be present (figure 4.2B). The cleaved M1CT and the SUMO tag seemed to be eluted together as shown in the elution lanes in figure 4.2B. It is unclear whether cleaved M1CT interacts with SUMO tag or M1CT sticks to the HIS column.

A second purification step was needed to remove the SUMO tag, as well as other impurities therefore, the eluted fractions from the second Ni purification were pooled and subjected to ion exchange (IE) chromatography through a Q column (Figure 4.3). The pI of M1CT is approximately 9.4 while the pI of SUMO is approximately 5.3. The IE was performed at pH 8 where the M1CT passed through the column when it had an overall positive charge as shown by the first peak in figure 4.3A. The SUMO protein bound to the Q-resin as expected due to an overall negative charge and was eluted at increasing NaCl concentrations, as shown by the second peak in figure 4.3A. The confirmation of M1CT and SUMO tag being present in fractions of the first peak and second peak, respectively, from the IE chromatogram, was confirmed by SDS-PAGE analysis shown in figure 4.3B.

4.1.2 NMR Characterisation of M1CT

A HSQC experiment of ^{15}N uniformly labelled cleaved M1CT was then collected (Figure 4.4). The NMR spectrum of M1CT at pH 5.5 indicated that it is an intrinsically disordered domain. It exhibited poorly dispersed NMR spectra,

suggesting that it did not adopt a stable tertiary structure. An NMR spectrum of M1CT at pH 7.0 was intended to be collected but due to time constraint, this was not done.

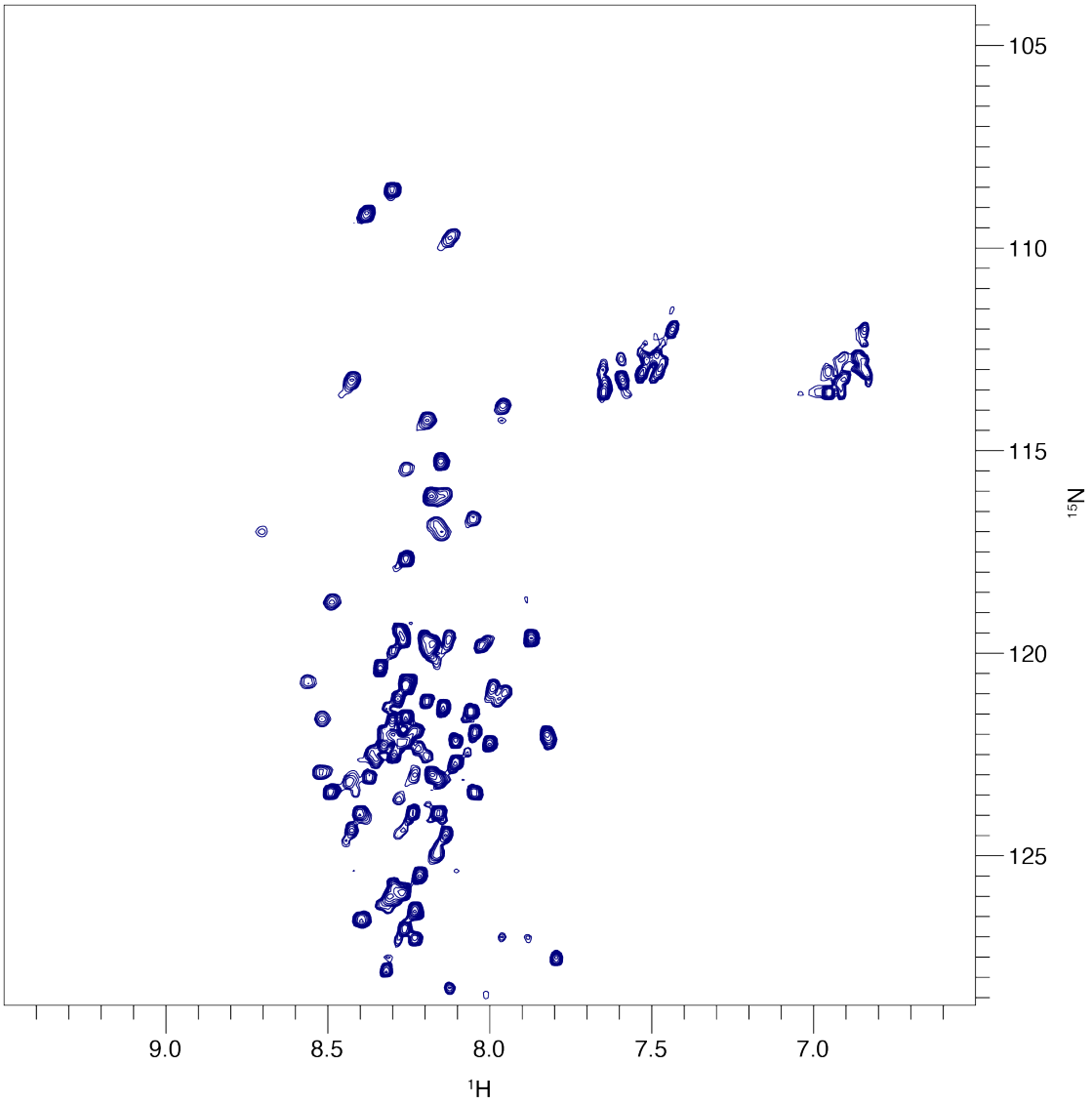


Figure 4.4: Assigned ^{1}H-^{15}N TROSY spectrum of M1 C-term in 20 mM phosphate, 50 mM NaCl, 0.5 mM TCEP at pH 5.5 collected on a 750 MHz (^{1}H) spectrometer.

M1CT adopting a non-stable tertiary structure is consistent with previous reports showing that the C-terminal domain (165 to 252) is flexible and intrinsically disordered (Kordyukova et al., 2008; Ksenofontov et al., 2010; Shishkov et al., 2009, 2011; Zhang et al., 2012). The M1NT domain of the M1 comprised more than 65% of the full-length protein and unlike the M1NT domain which is fully resolved, the

C- terminal domain of M1 has not been fully crystallised despite several attempts (Arzt et al., 2001; Sha Luo, 1997). The C-terminal domain was missing in the crystal structures, although crystallisation drops were set up with M1 full-length protein (Sha Luo, 1997). However, Baudin and colleagues noted that when M1 full-length protein was concentrated higher than 2-3 mg/ml, the protein degrades such that the C-terminus is removed. This might explain the difficulties in the crystallisation of the M1 full-length protein.

Other biophysical techniques such as circular dichroism (CD) spectroscopy, limited proteolysis and MALDI-TOF MS analysis of the proteolysis products, dynamic light scattering and analytical ultracentrifugation, have been used to investigate the structural features of the C-terminal domain (Kordyukova et al., 2008; Ksenofontov et al., 2010; Shishkov et al., 2009, 2011; Zhang et al., 2012). All the results indicate that the C-terminal domain is flexible and weakly structured. However, there is also a possibility that M1CT forms a more stable folded structure in the context of full length M1 and this will be tested in the next section.

4.1.3 M1NT with C-terminal Extensions

It has been reported that M1 has a strong tendency to oligomerise (Bui et al., 2000; Ruigrok et al. 2000; Calder et al., 2010) forming an organised structure underneath the virus envelope which is needed for virus assembly and budding (Ali et al., 2000; Avalos et al., 1997; Barman et al., 2001; Gomez-Puertas et al., 2000; Latham and Galarza, 2001; Wang et al., 2010). Although it has been reported that the N-terminal domain mediates the oligomerisation of M1, there is little information on the contribution of the C-terminus of M1 towards M1-M1 self-association. The C-terminal domain accounts for one-third of the full length protein and is connected to the M1NT domain via a long loop sensitive to proteolytic cleavage (Kordyukova et al., 2008).

To gain insights into whether the addition of native C-terminal domain sequence affected the M1NT or the M1CT conformation, we trialled purification, solubility, and characterisation of several M1 N-terminal constructs with increasingly extended

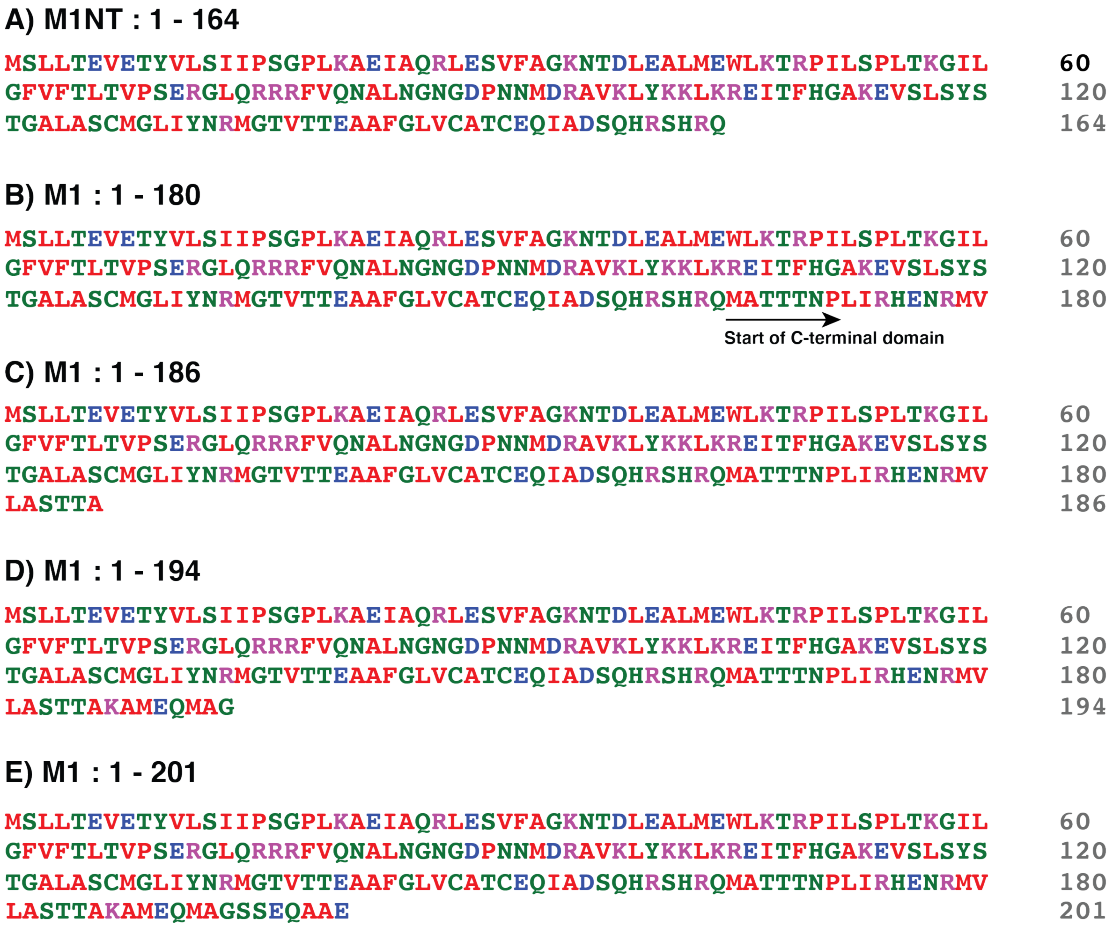


Figure 4.5: M1 construct extension : A) 1 to 164 residues, B) 1 to 180 residues, C) 1 to 186 residues, D) 1 to 194 residues and E) 1 to 201 residues.

C-terminal domains. Ten constructs with increasing amounts of C-terminal sequence were tested: residues 1-180, 1-186, 1-194, 1-201, 1-209, 1-224, 1-232, 1-236, 1-241 and 1-252 (Figure 4.5 and 4.6). Test expression and purification of each of the constructs was performed. The solubility of the constructs were assessed by UV absorbance after concentrating using centrifugal filters and qualitatively monitoring for visible aggregates qualitatively during sample concentration. M1NT (1-164) (Figure 4.5A) could be concentrated to $\sim$ 450 μM at pH 7.0 and was stable for $\sim$ 4 days at 30°C. M1(1-164) was $\sim$ 4-fold more soluble at pH 5.5.

It was found that the solubility limits at pH 7.0 decreases as the length of the C-terminus increases. An M1(1-224) construct (Figure 4.6G) had a concentration limit of 20µM at pH 5.5 and exhibited an intermediate propensity to self-assemble, which is in good agreement with previous findings (Zhang, 2015). Monomeric



Figure 4.6: M1 construct extension : F) 1 to 209 residues, G) 1 to 224 residues, H) 1 to 232 residues, I) 1 to 236 residues and J) 1 to 241 residues.

samples of constructs longer than M1(1-224) could not be prepared consistently at sufficient concentration for NMR studies. The longest construct purified successfully was M1(1-232). The spectra of M1(1-224) and M1(1-232) are shown in figures 4.7 and 4.8, respectively. Overlay spectra of M1(1-224) and M1(1-232) at pH 5.5 are shown in figures 4.9 and 4.11.

Overlay of M1(1-224) with the M1CT construct indicated that the C-terminal region remained unfolded and exhibited minor chemical shift changes (Figure 4.10 and 4.11). A comparison of the chemical shifts of M1(1-164) and M1(1-224) indicated only small chemical shift changes in the M1NT. Addition of the unlabelled M1CT construct in trans to an ^{15}N labelled M1NT sample at pH 7.0 failed to detect interactions of M1CT with soluble (monomeric) M1NT, which was consistent with the absence of strong interactions between the M1NT and the M1CT in the

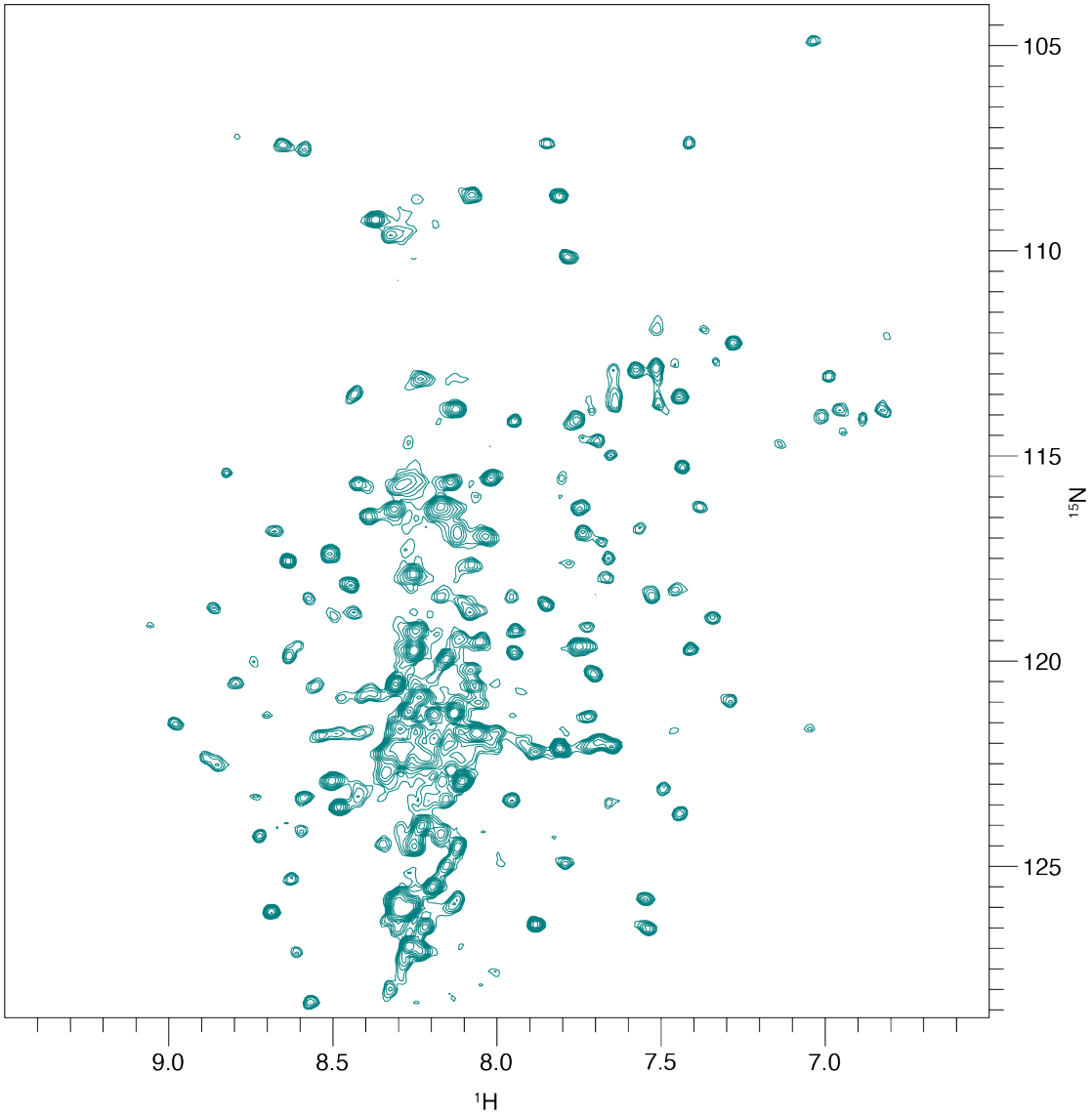


Figure 4.7: ^{1}H-^{15}N TROSY spectrum of M1 (1-224) in 20 mM phosphate, 50 mM NaCl, 0.5 mM 2-mercaptoethanol at pH 7 collected on a 750 MHz (^{1}H) spectrometer.

monomeric state (Data not shown).

NMR data on M1CT suggest it to be monomeric in solution at pH 5.5 as the rotational correlation coefficient is consistent of a monmeric protein and not of a dimer (Figure 4.4). The high degree of flexibility of M1CT indicates that it is unlikely to have a stable dimerization domain. Overlaying the spectra of M1(1-224) and M1CT (165-252) at both neutral and acidic pHs show that the C-terminal domain remained unfolded and exhibited only minor chemical shift changes ()Figure 4.10 and 4.11). The longest construct purified, M1(1-232) length has also shown

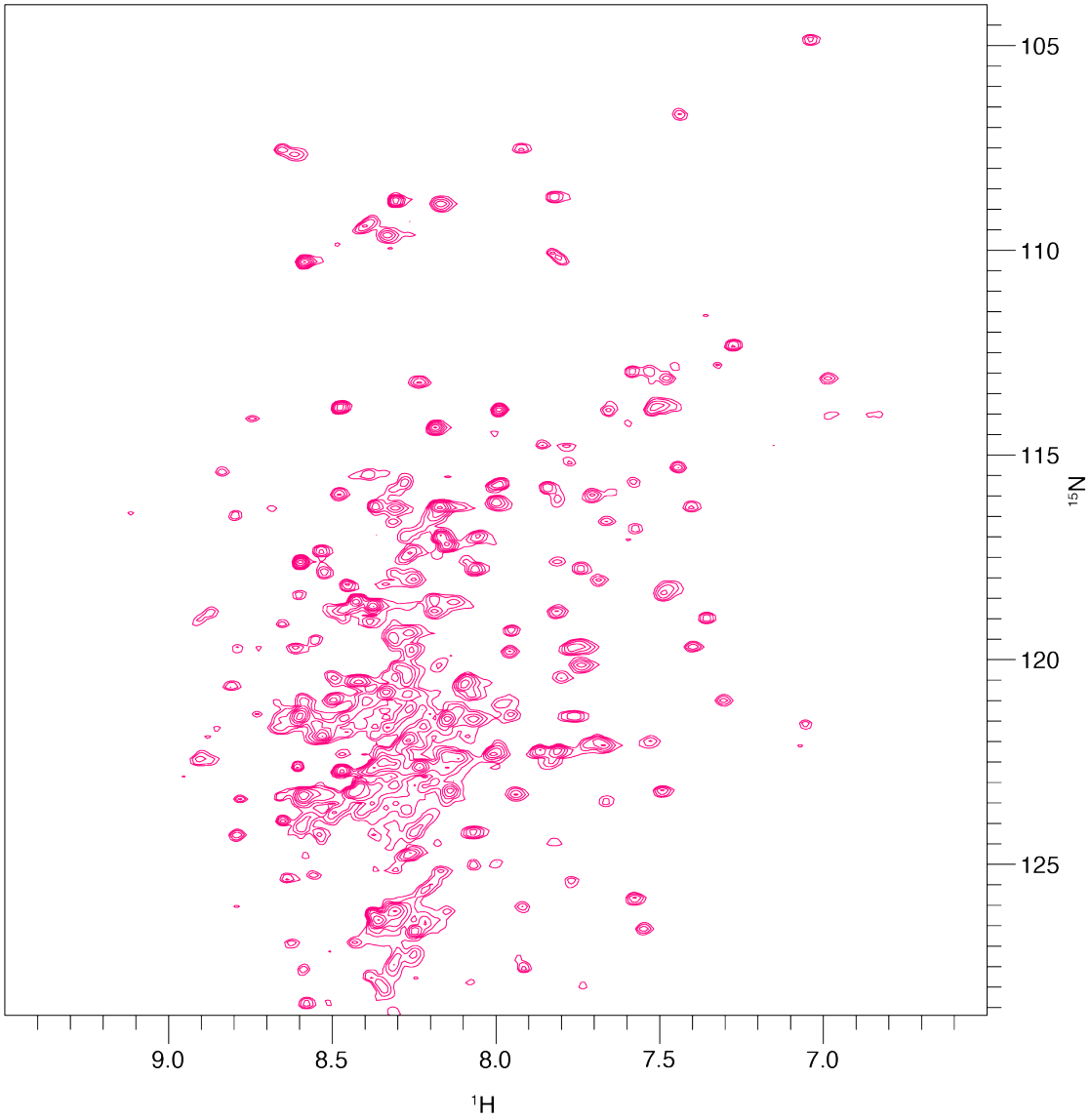


Figure 4.8: ^{1}H-^{15}N TROSY spectrum of M1 (1-232) in 20 mM phosphate, 50 mM NaCl, 0.5 mM 2-mercaptoethanol at pH 5.5 collected on a 750 MHz (^{1}H) spectrometer.

the C-terminal domain in an unfolded state. The C-terminal domain is likely to form a more stable folded domain in the oligomerised form of M1 (Peukes et al., 2020 ;Selzer et al. 2020; Zhang et al., 2017).

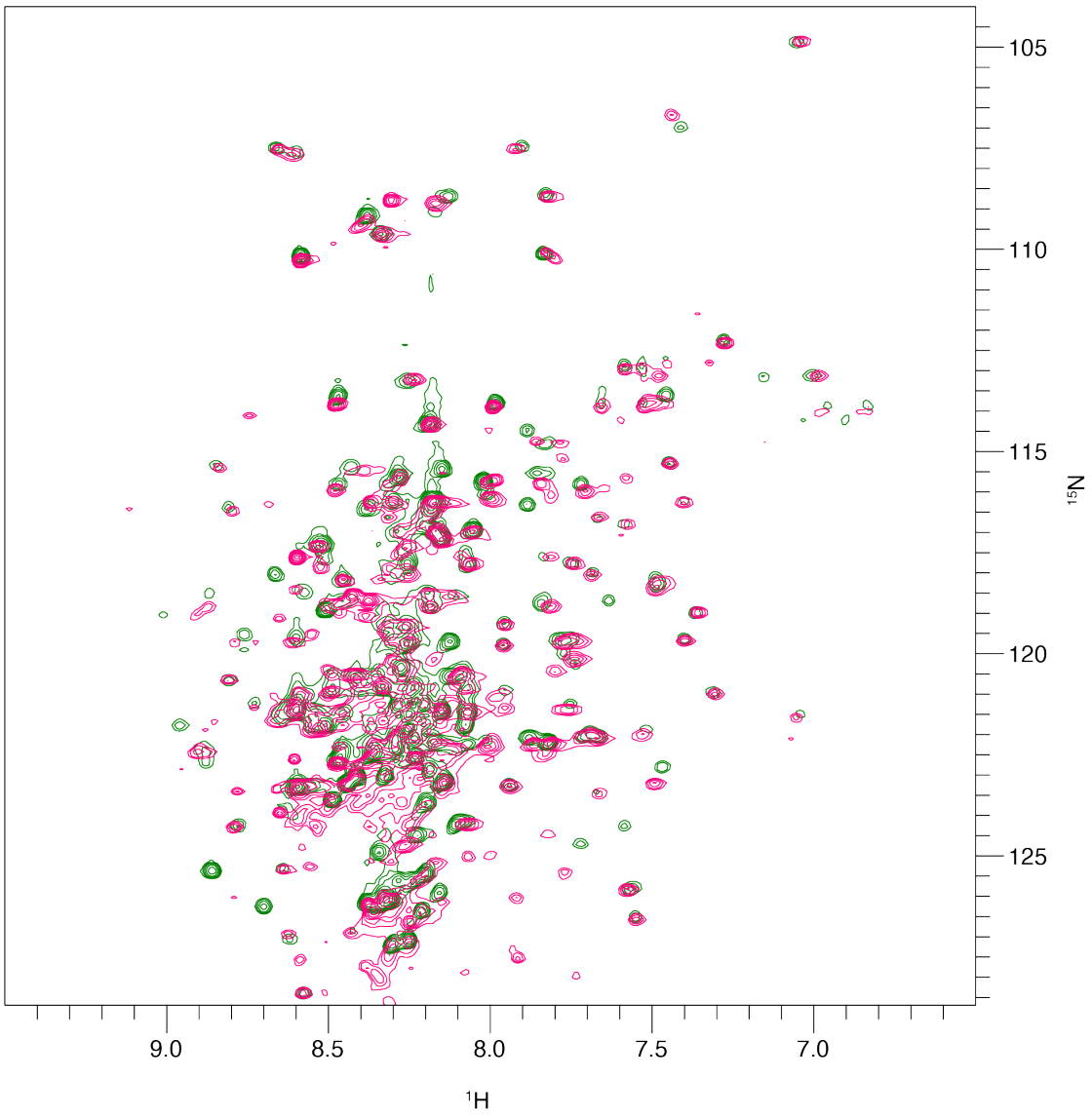


Figure 4.9: Overlay of ^{1}H-^{15}N TROSY spectrum of M1 (1-224) (Green) at pH 5.5 and M1 (1-232) (Pink) in 20 mM phosphate, 50 mM NaCl, 0.5 mM 2-mercaptoethanol at pH 5.5 collected on a 750 MHz (^{1}H) spectrometer.

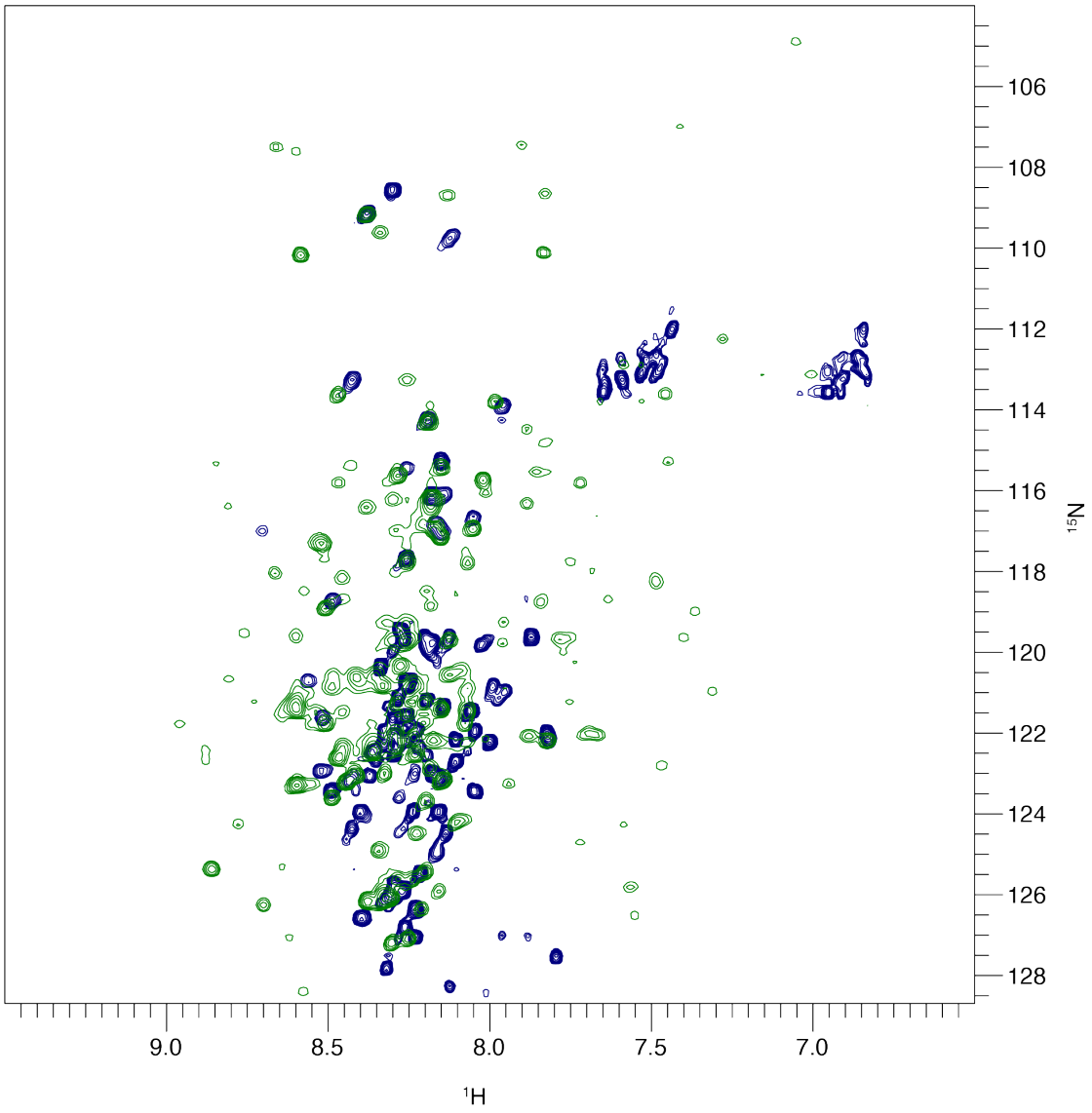


Figure 4.10: Overlay of ^{1}H-^{15}N TROSY spectrum of M1 (1-224) (green) and M1CT (165-252) (blue) in 20 mM phosphate, 50 mM NaCl, 0.5 mM 2-mercaptoethanol at pH 5.5 collected on a 750 MHz (^{1}H) spectrometer.

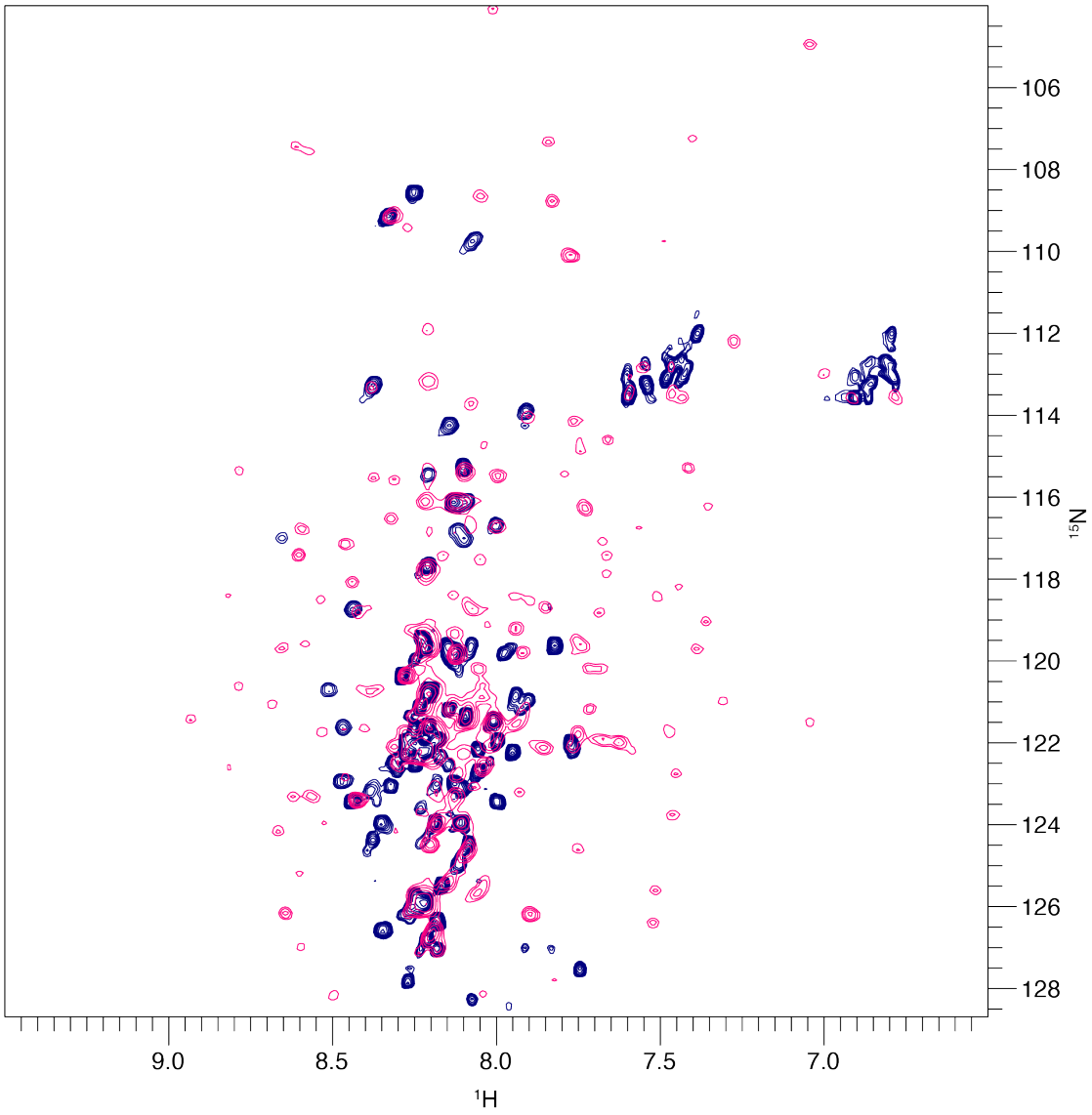


Figure 4.11: Overlay of ^{1}H-^{15}N TROSY spectrum of M1 (1-224) (pink) at pH 7 and M1CT (165-252) (navy blue) in 20 mM phosphate, 50 mM NaCl, 0.5 mM 2-mercaptoethanol at pH 5.5 collected on a 750 MHz (^{1}H) spectrometer.

Nothing in life is to be feared, it is only to be understood. Now is the time to understand more, so that we may fear less.

— Marie Curie

5

M1 Oligomerisation and the Hydrophobic Pocket

Contents

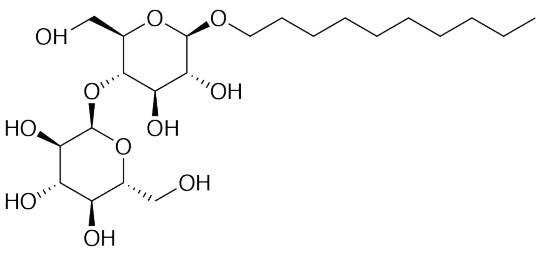
5.1 Sensitivity of The Hydrophobic Pocket of M1 to Solution Conditions 64
5.1.1 Surfactants 64
5.1.2 pH 70
5.1.3 Isopropyl Alcohol (IPA) 75
5.2 Oligomerisation 77
5.2.1 Cholesteryl Hemisuccinate (CHS) 77
5.2.2 25-hydroxycholesterol (25-HC) 78
5.2.3 Mutants Based on 1EA3 Crystal Structure 81
5.2.4 Mutants based on 1AA7 Crystal Structure 88
5.2.5 CHAPS Binding of M1 and Its Mutants 91
5.2.6 Kosmotropes and Chaotropes 102
5.2.7 PHE 109

5.1 Sensitivity of The Hydrophobic Pocket of M1 to Solution Conditions

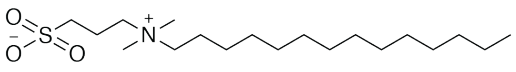
M1 plays many important roles in the influenza life cycle, and its ability to oligomerise is critical for various functions which include the formation of the matrix layer during its viral pathogenesis (Ali et al., 2000; Avalos et al., 1997; Barman et al., 2001; Gómez-Puertas et al., 2000; Latham and Galarza, 2001; Wang et al., 2010). M1 alone is able to form virus-like particles (VLPs) as a result of M1 self-associating into a long, helical oligomers beneath the membrane (Calder et al., 2010; Gómez-Puertas et al., 2000; Noton et al., 2007). *In vitro*, M1 forms oligomers efficiently at a neutral pH and disintegrates at low pH (Shtykova et al., 2017; Zhang et al., 2012), and the M1NT domain alone shows a pH-dependent oligomerisation (Noton et al., 2007). A study done by Ruigrok et al., has demonstrated that M1 binds to lipid membranes with PS and is predominanly driven by the M1NT domain (Baudin et al., 2001; Ruigrok et al., 2000; Höfer et al., 2019). *In vivo*, membrane binding was seen to promote M1 oligomerisation (Hilsch et al., 2014) and is primarily driven by the N-terminal domain (Baudin et al., 2001; Höfer et al., 2019). The physical trigger that converts M1 from the monomeric to the oligomeric form is still not understood. Identifying this trigger will help us better understand how M1 multimerises. The objectives of the following sections will examine whether M1's association with membrane mimetics can be observed through CSPs and to test whether membrane binding induces conformational changes that might be associated with oligomerisation.

5.1.1 Surfactants

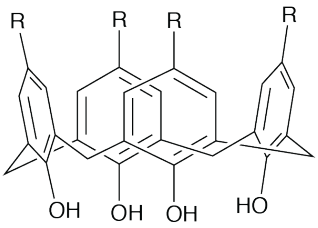
Since membrane-binding induces M1 oligomerisation, we tested the effects of several membranes-mimicking surfactants on M1NT structure. We first tested n-dodecyl-β-D-maltopyranoside (DDM), Anzergent (AZ314) and, Calixarene-based detergent ($C4C7CO_2$) as shown in Figure 5.1. These surfactants interacted strongly with M1 leading to a severe loss of signals, a narrowing of signal dispersion, and visible aggregates in the samples (Figure 5.2 and 5.3a). Lauryl maltose-neopentyl glycol



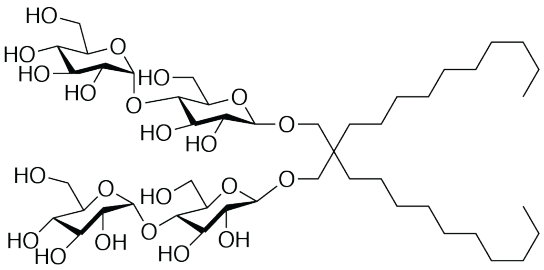
n-Dodecyl-β-D-maltopyranoside
(DDM)



AZ314



Calixarene-based detergent
C4C7CO₂



Lauryl Maltose Neopentyl Glycol
(LMNG)

Figure 5.1: Initial Screening of Detergents tested: n-dodecyl-β-D-maltopyranoside (DDM), AZ413, Calixarene-based detergent ($C4C7CO_2$) and Lauryl Maltose Neopentyl Glycol (LMNG).

(LMNG) was then tested, and in contrast, the result has shown that it did not interact with M1NT (Figure 5.3b).

Sterol-containing surfactant, 3-((3-cholamidopropyl) dimethylammonio)- 1 -propanesulfonate (CHAPS) was also tested (Figure 5.4). Interestingly, CHAPS induced highly specific CSPs and did not aggregate the M1NT sample, remaining stable in a soluble form, as shown in figure 5.5. From this result, other sterol-containing surfactants were tested, including 3-([3-Cholamidopropyl] dimethylammonio)-2-hydroxy-1-propanesulfonate (CHAPSO), cholate, and deoxycholate as shown in figure 5.4. These sterol-containing surfactants differing by their attached moieties were tested to observe if the same effects of CSPs were produced (Figure 5.6, 5.7 & 5.8).

Intriguingly, the same specific CSPs were observed for all the sterol-containing surfactants tested. These sterol-containing surfactants strongly perturbed the

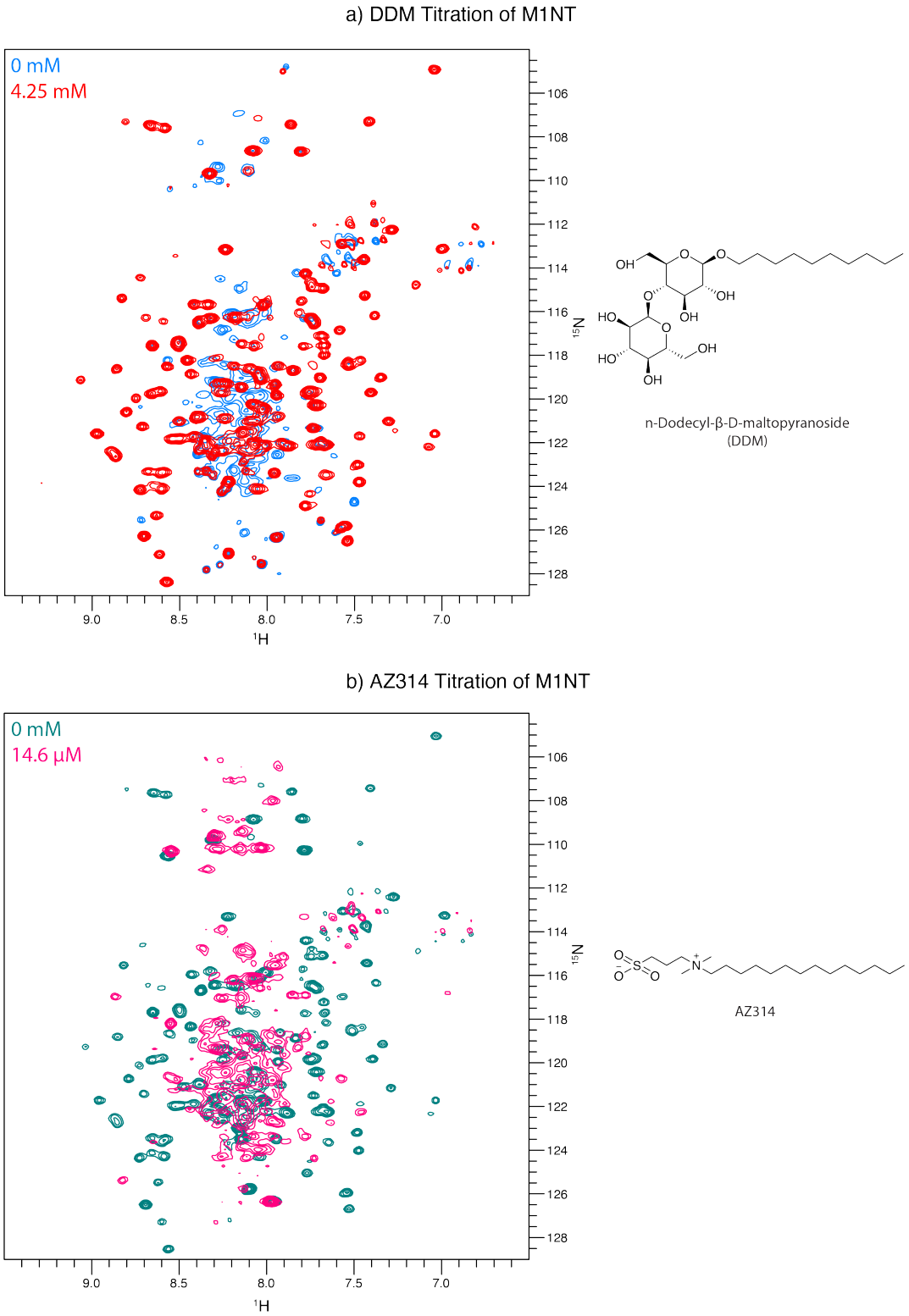


Figure 5.2: DDM and AZ314 titration on M1NT. a) Overlay of M1NT at 40 μM before (blue) and after DDM titration of 4.25 mM (red). b) Overlay of M1NT at 40 μM before (teal) and after AZ314 titration of 14.6 μM (pink). DDM and AZ314 titration of M1NT result in severe loss of signals and visible aggregates. Spectra collected at 15°C, pH 7 on 750 MHz spectrometers.

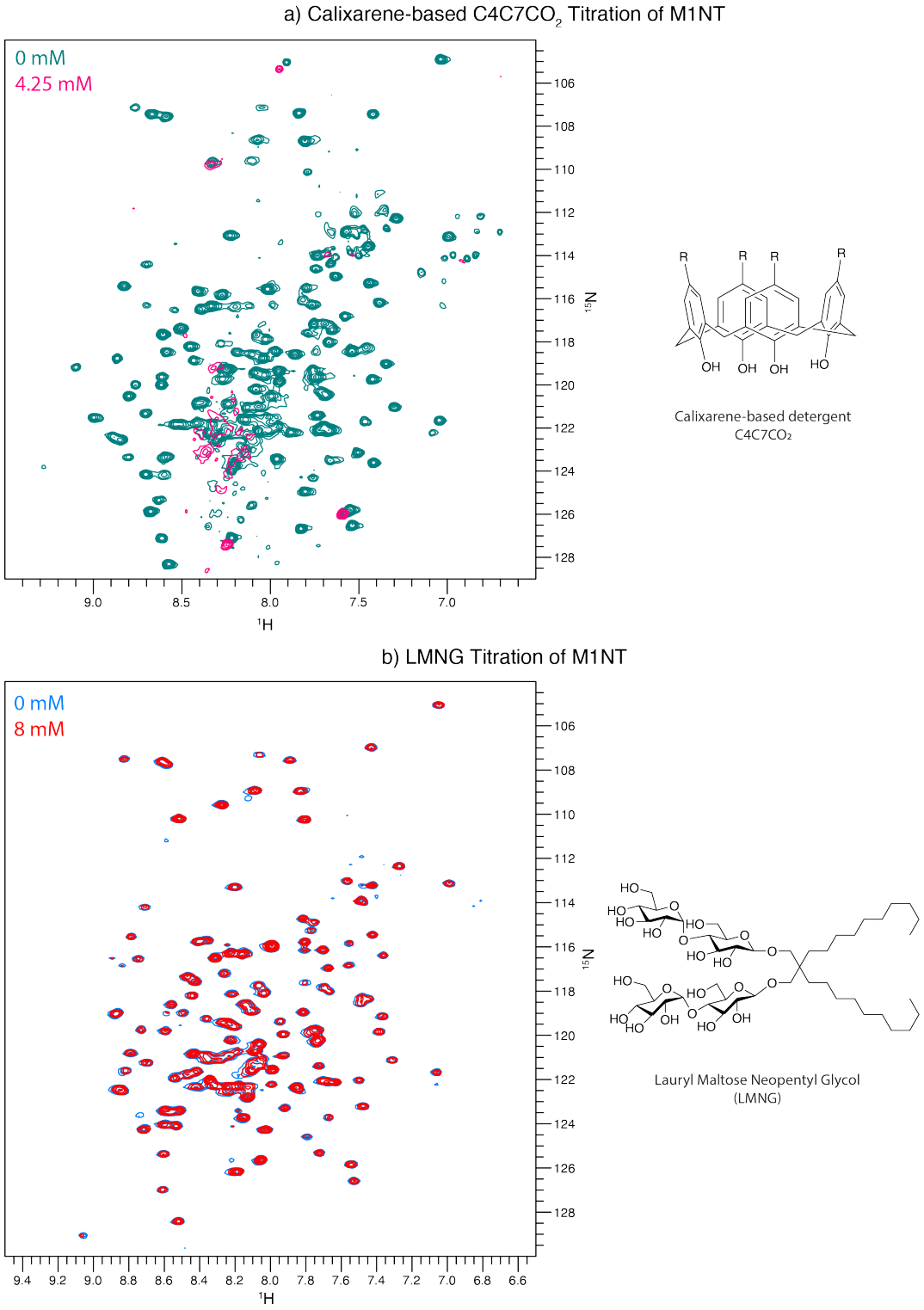


Figure 5.3: C4C7CO_2 and LMNG titration on M1NT. a) Overlay of M1NT at 40 μM before (teal) and after C4C7CO_2 titration of 4.25 mM (pink). b) Overlay of M1NT at 40 μM before (blue) and after LMNG titration of 8 mM (pink). C4C7CO_2 titration of M1NT results in severe loss of signals and visible aggregates while LMNG did not interact with M1NT. Spectra collected at 15°C for C4C7CO_2 and 30° for LMNG, both at pH 7 on 750 MHz spectrometers.

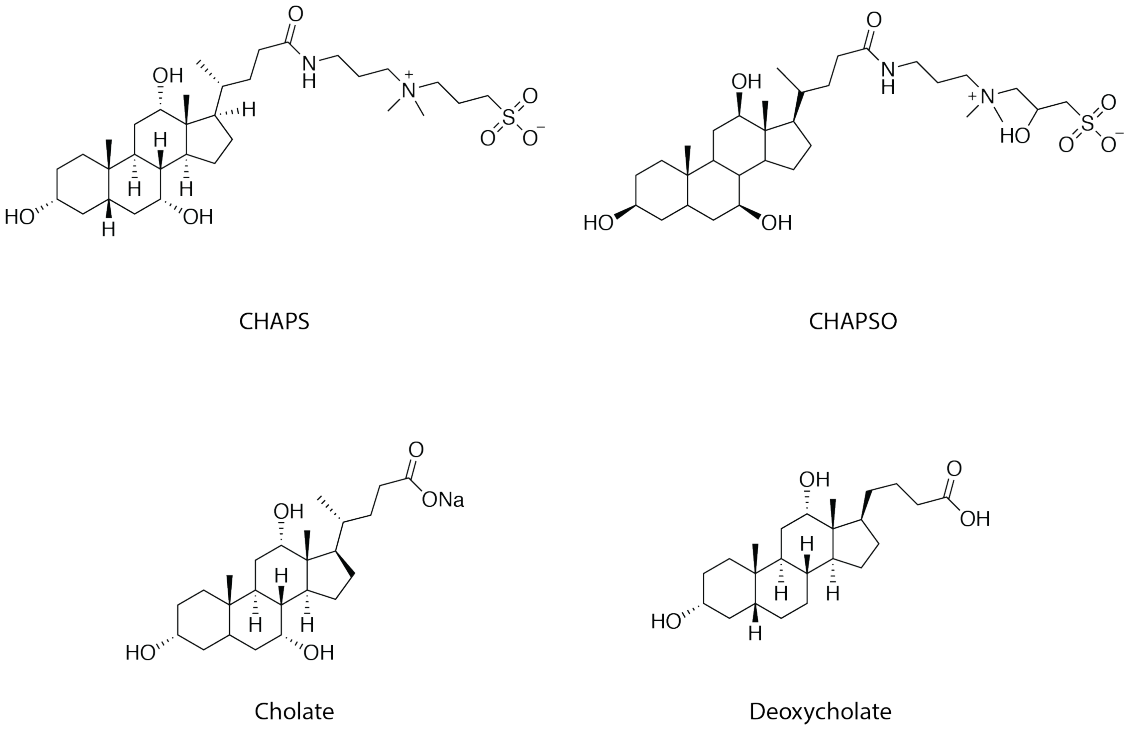


Figure 5.4: Sterol-containing surfactants: CHAPS, CHAPSO, Cholate and Deoxycholate.

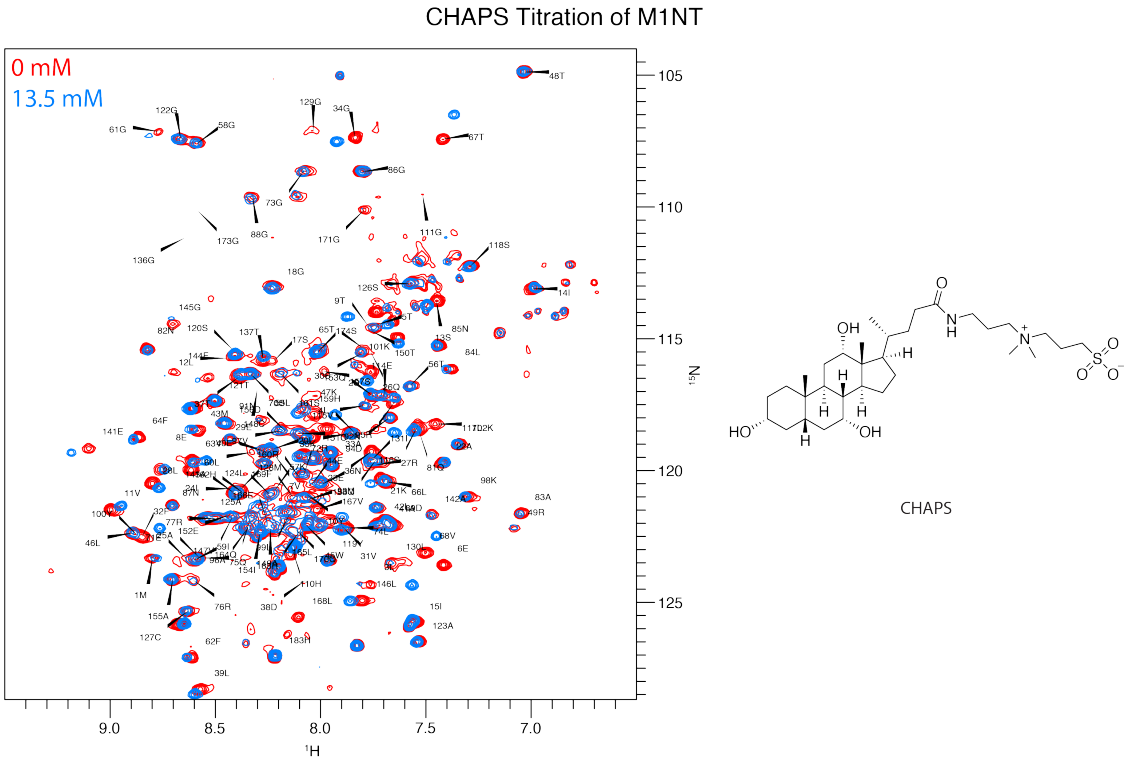


Figure 5.5: Overlay of M1NT at 40 μM before (blue) and after CHAPS titration to 13.5 mM (red). Spectra collected at 15°C, pH 7 on 750 MHz spectrometers.

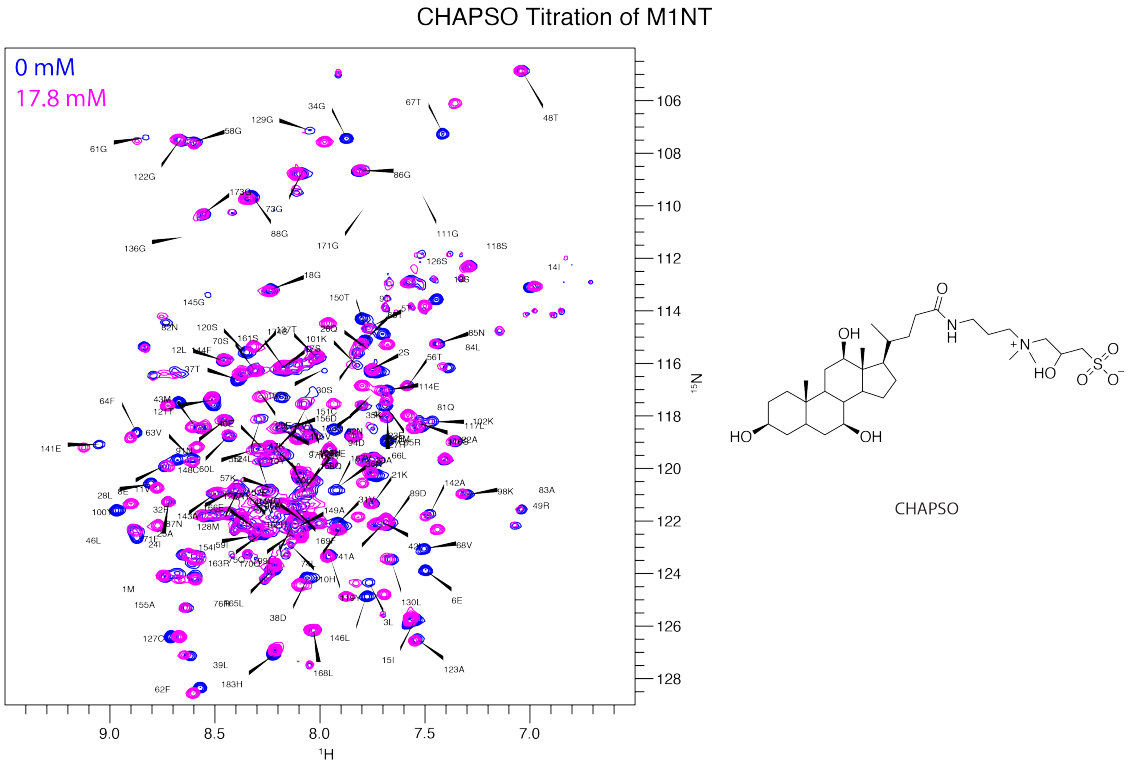


Figure 5.6: Overlay of M1NT at 40 μM before (dark blue) and after CHAPSO titration to 17.8 mM (magenta). Spectra collected at 15°C, pH 7 on 750 MHz spectrometers.

backbone amides of residues L4, G34, T67, and E141. Interestingly, these residues are mapped to the hydrophobic patch shown in the crystal structure in figure 5.9. The longer construct, M1 (1-224) which includes most of the C-terminus and has a higher propensity to self-associate, was also tested with the sterol-containing surfactants. Results also show that the surfactants induced the same specific CSPs (Figure 5.10). In this case, the CSPs exhibited were larger than those with M1NT (1-164). These titration of sterol-containing surfactants indicate a fast exhange regime.

To quantify the CSPs upon titration, T67 resonance was used because it is well separated in the 2D ^{1}H-^{15}N HSQC experiment of M1 constructs (Figure 5.5, 5.6, 5.7 & 5.8). From figure 5.11a, we can see the CSP for T67 resonance for each of the sterol compounds (CHAPS, CHAPSO, Cholate and Deoxycholate) tested on M1NT. The CSPs varied between compounds with Cholate and CHAPSO both exhibiting similarly large changes. This is in contrast to deoxycholate with the least change seen, which maybe due to the solubility limit of deoxycholate (0.24 g/L).

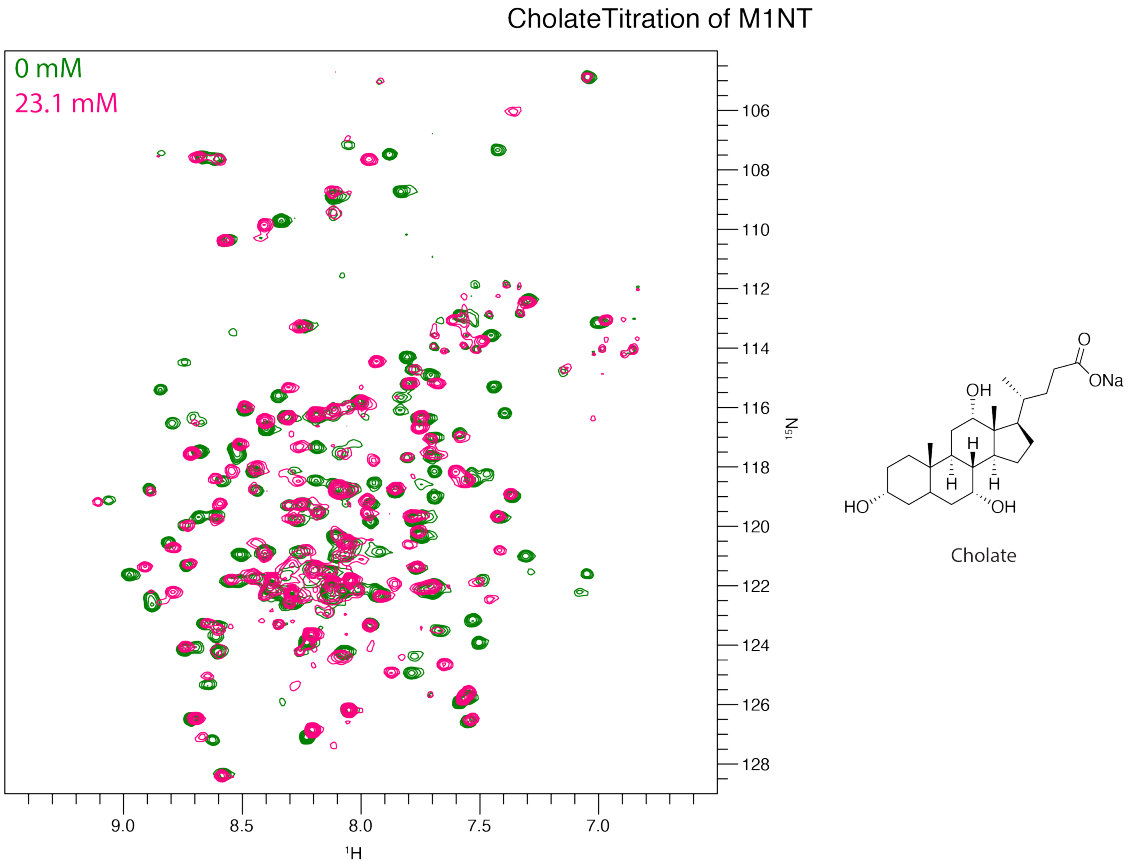


Figure 5.7: Overlay of M1NT at 40 μM before (green) and after Cholate titration to 23.1 mM (pink). Spectra collected at 15°C, pH 7 on 750 MHz spectrometers.

Interestingly, when the CSPs for T67 resonance were compared between M1NT and the longer construct, M1(1-224) which shows a higher propensity to self-associate, M1NT exhibited a greater CSP than M1(1-224) (Figure 5.11b).

5.1.2 pH

In the initial stages of virus infection, the virion becomes acidified in the process of endocytosis (Rossman and Lamb, 2011). The transmembrane M2 proton pump gets activated in the acidified endosome (pH 5) as virions enter the host cell. The protons are pumped into the virion interior, which results in dissociation of M1 from the membrane and vRNP uncoating (Harris et al., 2001). This was demonstrated by Fontana et al. through the use of cryo-electron tomography which showed the thinning of M1 layer from the viral envelope (Fontana et al., 2012). Apart from M1 interactions with viral proteins, M1 also takes a role in interacting with the

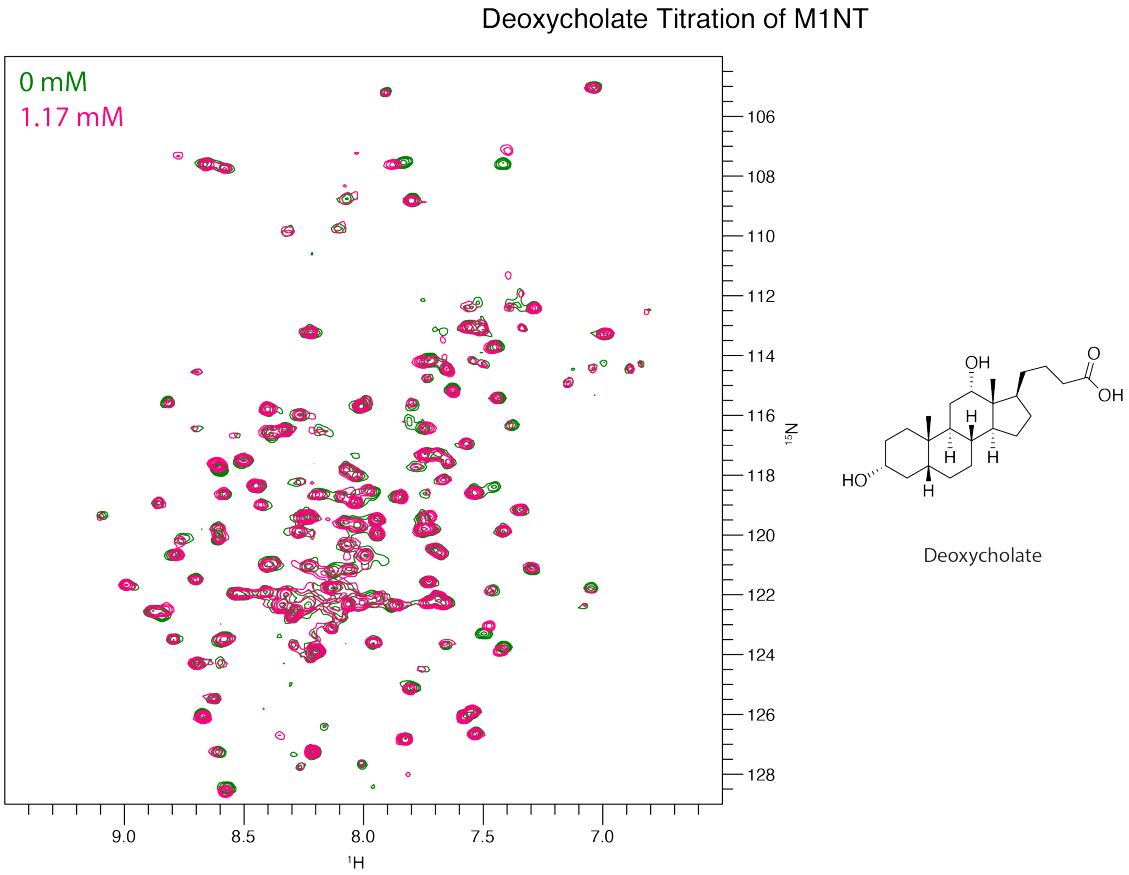


Figure 5.8: Overlay of M1NT at 40 μM before (green) and after Deoxycholate titration of 1.17 mM (pink). Spectra collected at 15°C, pH 7 on 750 MHz spectrometers.

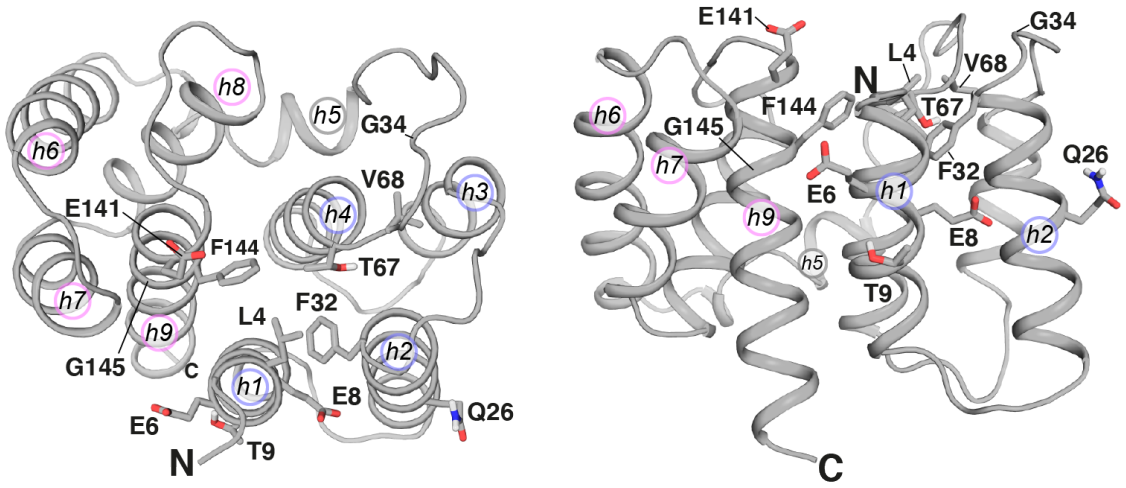


Figure 5.9: The specific CSPs observed from the effects of sterol-containing surfactants were mapped to 1AA7 M1NT crystal structure. These residues that were perturbed - L4, E6, E8, T9, Q26, F32, G34, T67, V68, E141, F144 & G145 are illustrated in the figure which forms the only hydrophobic patch on M1NT. Perturbed residues are clustered in helix 1 (N-terminus), helix 2 (C-terminus), helix 4 (C-terminus), and helix 9 (N-terminus).

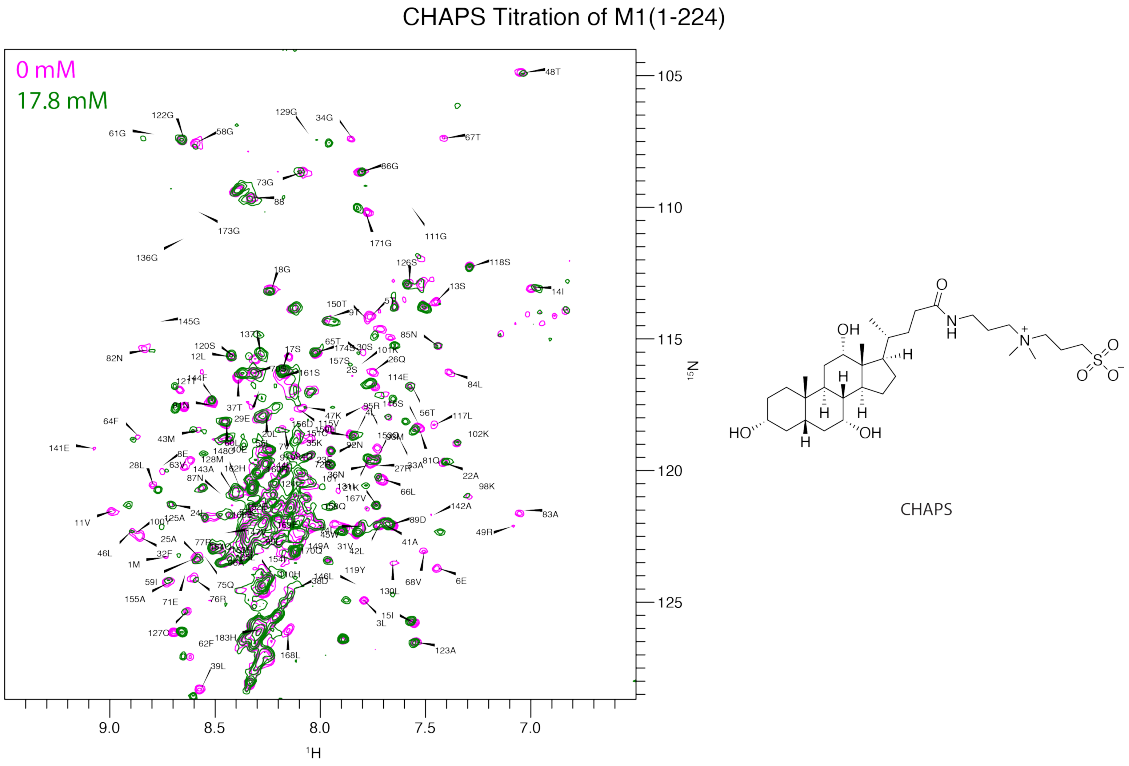


Figure 5.10: Overlay of M1(1-224) at 40 μM before (magenta) and after CHAPS titration to 17.8 mM (green). Spectra collected at 30°C, pH 7 on 750 MHz spectrometers.

host membrane in the late stages of virus assembly and budding. M1 versatility in function is believed to be associated with its ability to have multiple conformations corresponding to different pHs in various stages in virus replication (Safo et al., 2014).

The crystal structures available for M1NT at acidic or neutral pH may not be enough to address all the multiple potential conformations during viral replication, and perhaps through solutions studies may shed more light on these different conformational changes. Since sterol-containing surfactants such as the ones shown in figure 5.4 induce specific CSPs at the hydrophobic patch, the chemical shift perturbations (CSPs) of the residues at the hydrophobic patch at pH 5.5 and pH 7 were examined for both M1NT and M1(1-224) constructs (Figure 5.12). Significant CSPs were observed between pH 6.0 to pH 6.5 , and modest CSPs were observed between pH 5.5 to 6.0 and pH 6.5 to 7.0 for both M1NT & M1(1-224) constructs (Figure 3.11)). When the CSPs for T67 and G34 resonances were compared at pH 5.5 and 7, the CSPs interestingly moved in the same direction as the CSPs induced

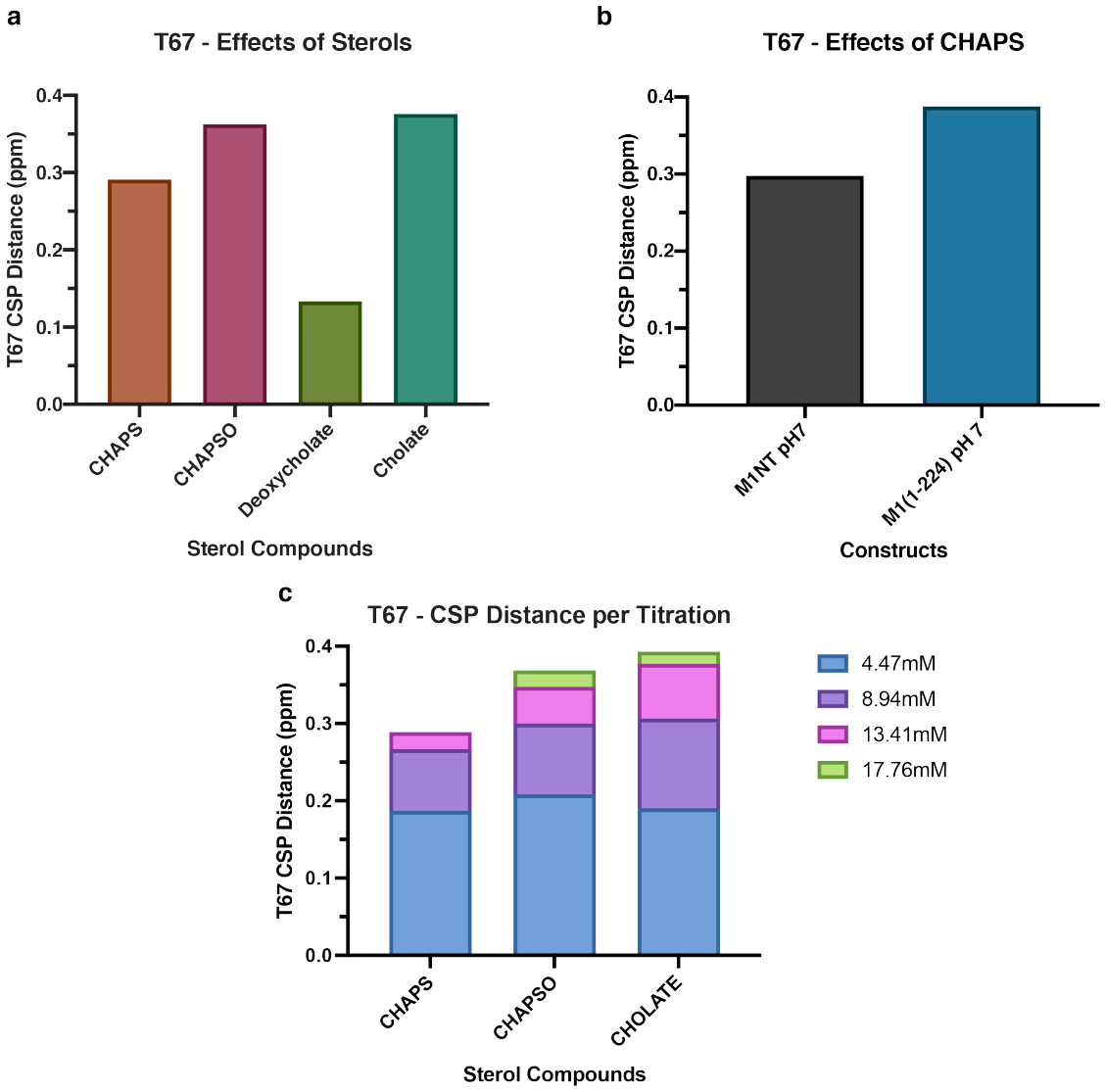


Figure 5.11: Quantifying the chemical shift perturbations of the T67 resonance on the hydrophobic patch on a) the different sterol compounds tested, b) different M1 constructs: M1NT(1-164) & M1(1-224), and c) CSPs per titration point for each sterols. These CSPs were measured at the saturation level of titration for each sterol compound. The concentration of M1NT and M1(1-224) used was 40-50μM and M1(1-224) at 30-40uM. The formula used to calculate the CSPs distance moved is $d = \sqrt{\frac{1}{2}\left[\delta_H^2 + \left(\frac{1}{5} \cdot \delta_N^2\right)\right]}$, where d is the distance moved in ppm, δ_{H} and $delta_{\mathrm{N}}$ are the chemical shifts changes in H^1 and N^{15} shifts in ppm.

pH 5.5 vs pH 7:

a) M1NT

b) M1(1-224)

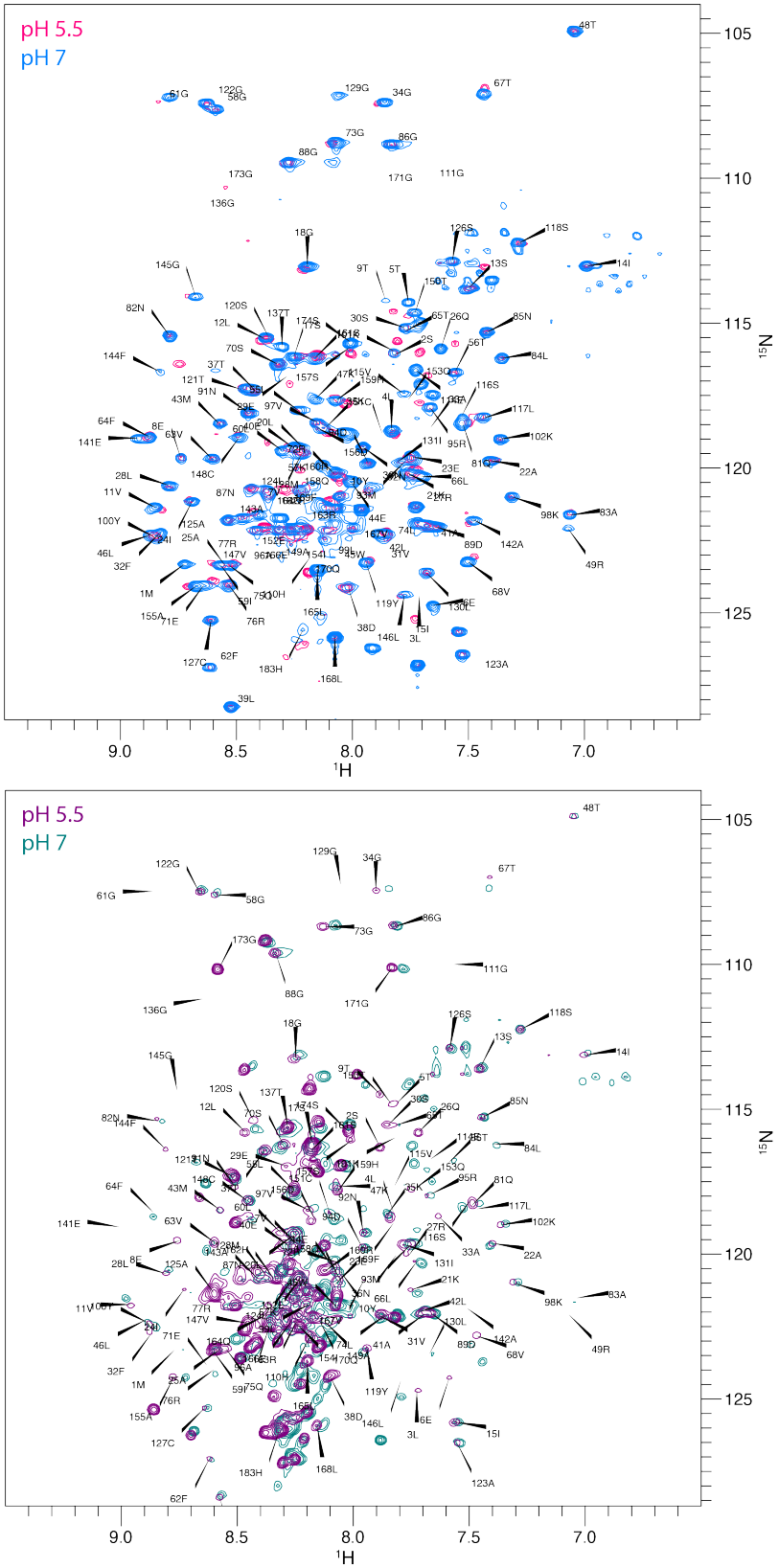


Figure 5.12: Overlay of a) M1NT b) M1(1-224) at pH 5.5 & pH 7. Spectra collected at 30°C on 750 MHz spectrometers. M1NT(1-164) & M1(1-224) was prepared at a low pH (pH 5.5) as described previously in Chapter 2. For samples at pH 7, The pH was then gradually increased to pH 7. The concentration of both constructs are 50-60μM in 20mM Tris, 50mM NaCl.

by sterol-containing compounds for both constructs as shown in figure 5.5 and 5.10.

The movement of the same residues at the hydrophobic pocket in response to pH changes from pH 5.5 to 7 and CHAPS titration may provide insights into the understanding of M1 oligomerisation. This interface is sensitive to changes in solution conditions. In addition, tilting of the stacked dimer arrangement seen in M1 structures when pH is decreased may also play a role in oligomerisation.

5.1.3 Isopropyl Alcohol (IPA)

The M1NT structure has been crystallised by several groups at both neutral and acidic pHs. There are two primary arrangements of the dimers that can be observed from all the crystal structures of M1NT - 1) stacked and 2) side-by-side arrangements. The 1EA3 crystal structure is the only crystal structure that shows a stacked dimer arrangement, as shown in figure 5.18. The 1EA3 is solved at neutral pH, pH 7.5 (Arzt et al., 2001). On the other hand, the 1AA7 and 4PUS, which show side-by- side arrangements, were solved at acidic pHs, pH 4.7 and pH 4.3, respectively (Safo et al., 2014; Sha and Luo, 1997). When comparing the crystallisation conditions between these crystals, aside from the pH and other standard crystallisation additives, the 1EA3 crystallisation conditions include the addition of 5% Isopropyl alcohol (IPA). Thus we decided to test the effect of titrating IPA into M1NT at pH 5.5 and 7.0, and the result is shown in figure 5.13. Interestingly, the chemical shift perturbations (CSPs) of L3, E6, G34, T67, V68 and E141 resonances moved in the same direction as the CSPs shown by sterol-containing detergents (Figures 5.5, 5.6, 5.7 & 5.8).

To summarise, several membrane mimicking detergents were tested on M1NT and it was found that DDM and AZ314 interacted strongly, aggregating M1NT within minutes of addition (Figure 5.2). This was in contrast to when LMNG was added, where M1NT was completely unperturbed (Figure 5.3b). When CHAPS was tested on M1NT, it was observed that CHAPS induced highly specific CSPs and remained in the solubilised form (Figure 5.5). Interestingly, when other sterol-containing surfactants containing a similar headgroup such as deoxycholate were tested, these surfactants also induced the same specific CSPs, and the protein

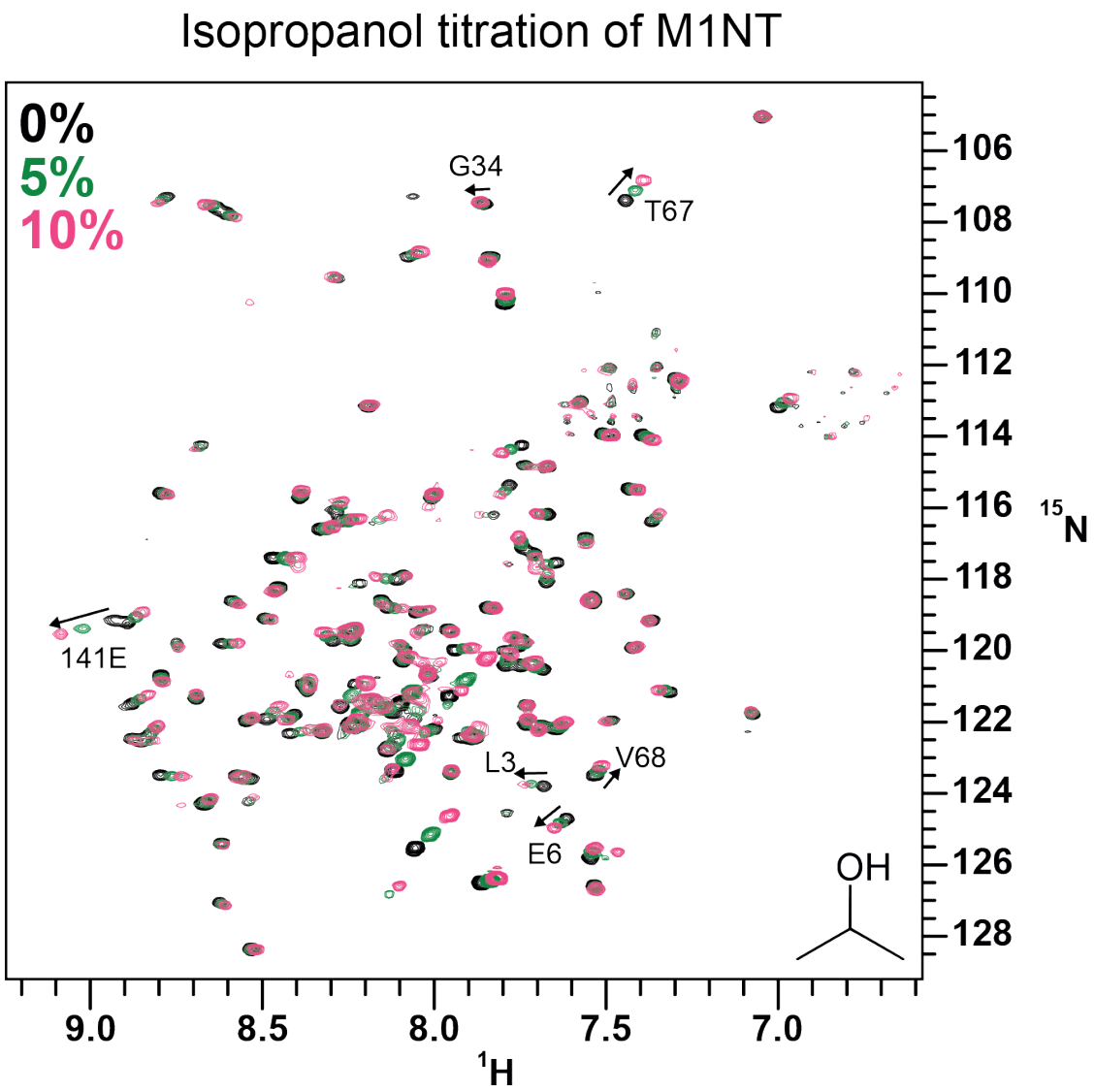


Figure 5.13: Overlay of M1NT at 40 μM at 0%, 5% & 10% isopropyl alcohol (IPA) at pH 7. The movements of L3, E6, G34, T67, V68 and E141 are indicated in the figure. Spectra collected at 30°C on 750 MHz spectrometers.

remained stable in solution (Figure 5.6, 5.7 & 5.8). In addition, IPA was also shown to have a similar effect on M1NT. All the tested sterol-containing surfactants (Figure 5.4) have a common head group, and IPA suggests that the hydroxyl groups of the head group probably induce these highly specific CSPs. These sterol-containing surfactants strongly perturbed L4, G34, T67 and E141 backbone residues which were mapped to the solvent-exposed N-terminal surface of M1NT (Figure 5.9). Similarly, M1(1-224) construct was also shown to have these residues highly perturbed by the sterol-containing surfactants (Figure 5.10). Intriguingly, the longer construct,

M1(1-224), which has a higher propensity to self-associate, exhibited a greater CSP when CHAPS was added as compared to the M1NT construct (Figure 5.11) which, could suggest that the C-terminal domain could exert more effect in the chemical shift changes on the solvent-exposed N-terminal surface of M1NT.

5.2 Oligomerisation

5.2.1 Cholesteryl Hemisuccinate (CHS)

5.2.1.1 CHAPS-solubilised CHS

It was clear from the similar results achieved for the different sterol compounds tested that some aspect of the sterol group induced specific CSPs on M1. The similar effects from IPA suggest that the hydroxyl groups were responsible. We had then sought to examine further the role of the sterol ring by testing sterol-like compounds that lacked the sterol ring hydroxyls, for example, cholesterol. Since cholesterol is largely insoluble in water, we used cholesteryl hemisuccinate (CHS), which can be solubilised in detergents like CHAPS. CHS was added to M1NT(1-164) or M1(1-224) at approximately stoichiometric concentrations and this led to a loss of signals for M1NT at pH 7.0 (Figure 5.14A). Intriguingly, the sample became cloudy as soon as CHS was added but did not precipitate out of solution and was stable for 1-2 days.

Increasing the temperature up to 50°C was not sufficient to recover the M1NT signals. CHAPS-solubilised CHS was also tested with M1(1-224) construct, and a loss of signals was also observed, as shown in figure 5.14B. However, signals from the C-terminal region were still detectable. This indicates that the C-terminus remained disordered.

These experiments were tested again for both M1NT and M1(1-224) at pH 5.5. Interestingly, for both M1NT and M1(1-224), there were no loss of signals or any CSPs, as shown in figure 5.15a and figure 5.15b, respectively. Altogether, these experiments suggest that addition of CHAPS-solubilised CHS to M1NT or M1 (1-224) at pH 7.0, resulted in the irreversible formation of large complexes, but the C-terminal region for M1(1-224) remained disordered.

Due to the limitations of NMR with size, we are unable to study these large complexes through this technique. As the large complexes formed, the tumbling rate is slower with faster relaxation rates and subsequently causes the peaks to become broader and weaker. Increasing the temperature up to 50°C to increase the overall molecular tumbling rate did not recover any of the signals from the peaks. Further efforts have been made to characterise these large complexes. However, the heterogeneity and large size of the oligomers prevented characterisation by dynamic light scattering. Electron microscopy (EM) has also proven unfrutiful as we were unable to determine conclusively the particles in EM images that were M1 - these experiment trials were subsequently dropped due to time constraint.

5.2.1.2 LMNG-solubilised CHS

Among all the detergents that were tested previously, sterols-containing surfactants (Figure 5.4) induced highly specific CSPs and detergents shown in figure 5.1 interacted with M1NT strongly resulting in severe loss of signals, a narrowing of signal dispersion, and visible aggregates in the sample. LMNG was the only detergent tested that did not interact with M1NT constructs on its own. This was an opportunity to test whether CHS solubilised in LMNG could also induce M1 oligomerisation. The result shown in figure 5.16 shows that LMNG-solubilised CHS did not perturb M1NT signals (no CSPs observed) and hence did not trigger self-association in contrast to what was seen in CHAPS-solubilised CHS in figure 5.14.

5.2.2 25-hydroxycholesterol (25-HC)

We then tried using another cholesterol compound, 25-hydroxycholesterol (25-HC) alone and, in combination with CHAPS. It was observed that both 25-HC and CHAPS-solubilised 25-HC did not induce oligomerisation (Figure 5.17). This implies that the specific combination of bile acid-derived compound with CHS was needed to induce M1 oligomerisation.

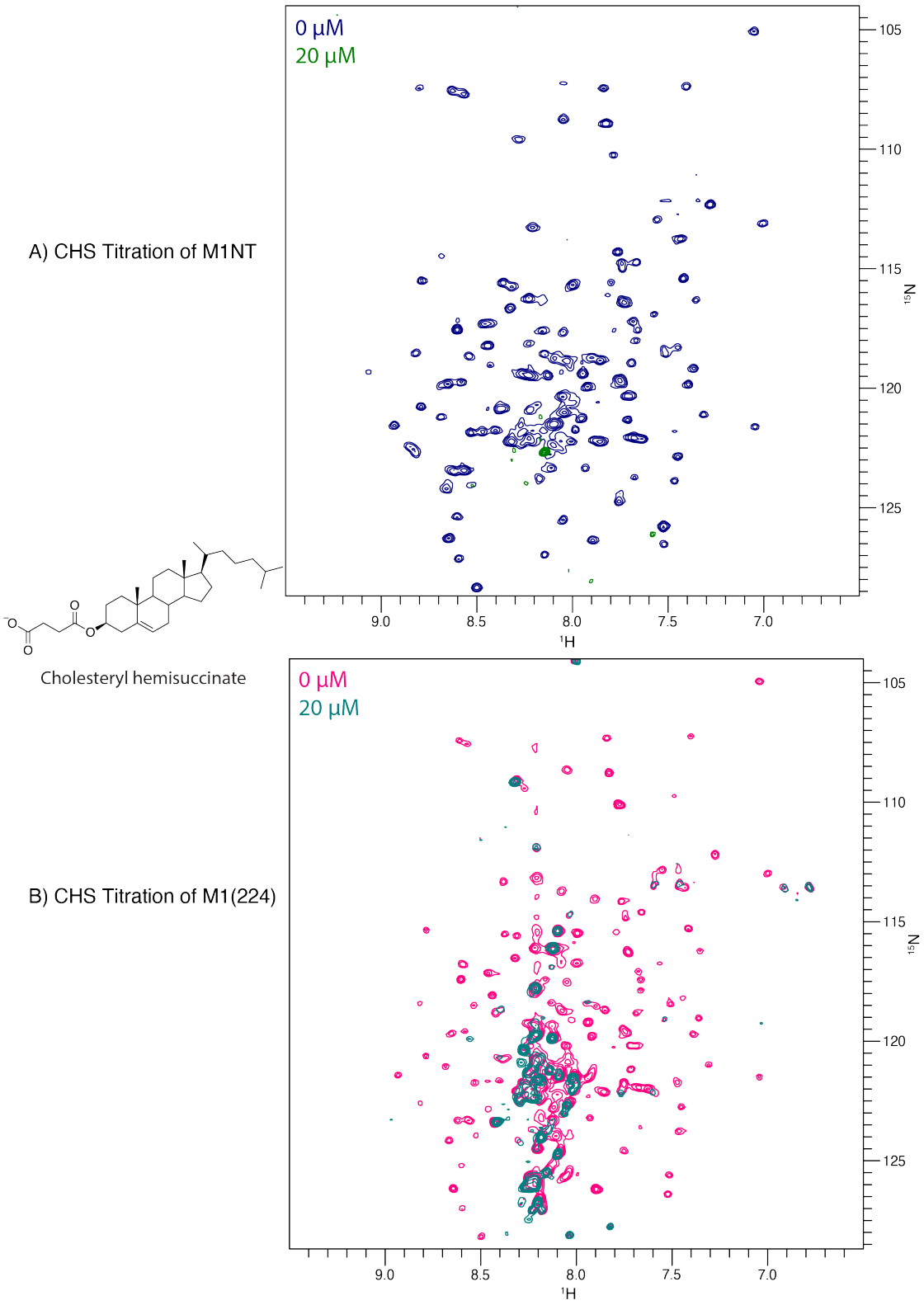


Figure 5.14: CHS titration of a) M1NT and b) M1(1-224) at pH 7. a) Overlay of M1NT at 30 μM before (blue) and after CHS titration of 20 μM (green). b) Overlay of M1(1-224) at 30 μM before (pink) and after CHS titration of 20 μM (teal). Spectra collected at 30°C, pH 7 on 750 MHz spectrometers.

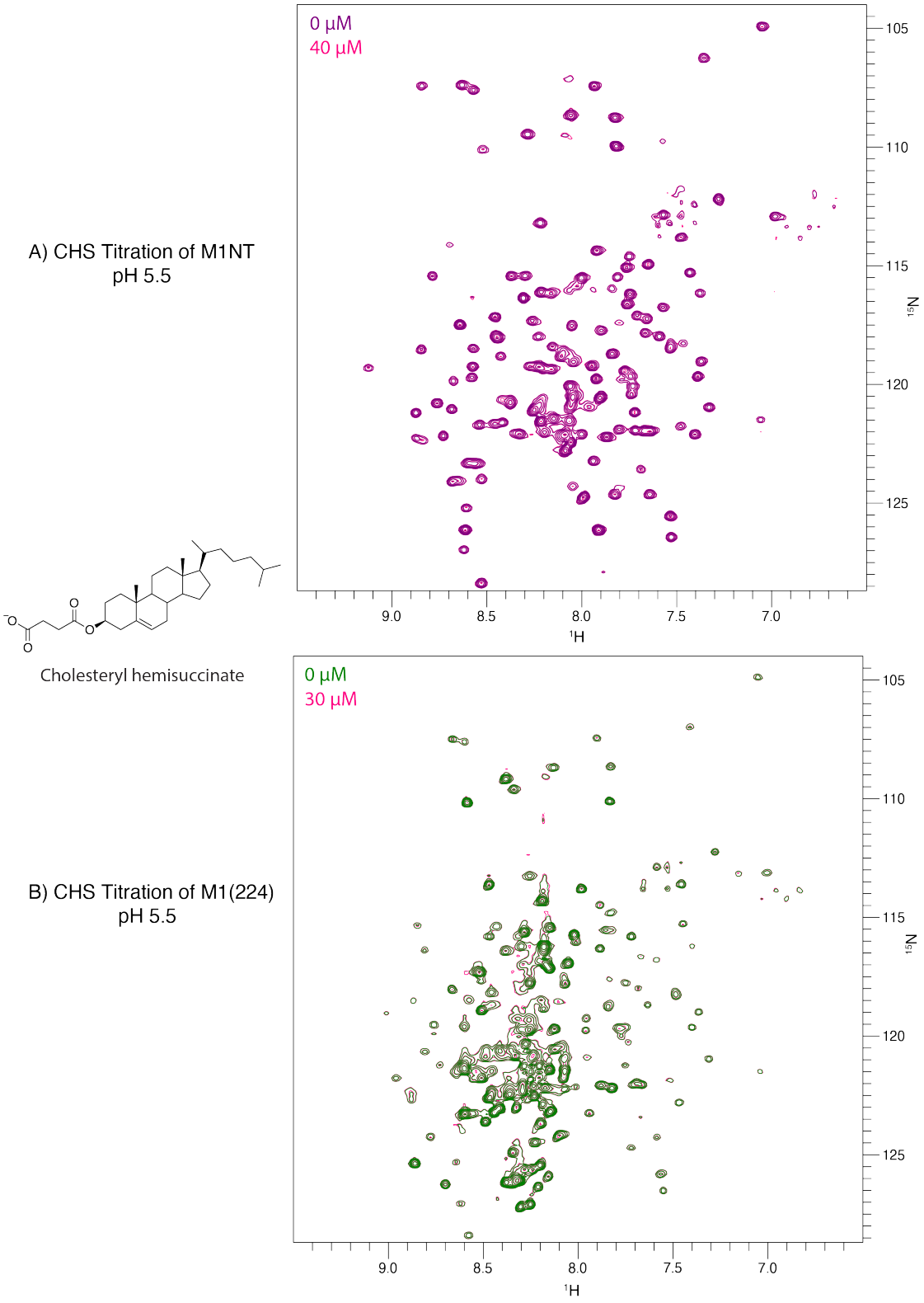


Figure 5.15: CHS titration of a) M1NT and b) M1(1-224) at pH 5.5. a) Overlay of M1NT at 40μM before (purple) and after CHS titration of 40 μM (magenta). b) Overlay of M1(1-224) at 30 μM before (green) and after CHS titration of 20 μM (pink). Spectra collected at 30°C, pH 5.5 on 750 MHz spectrometers.

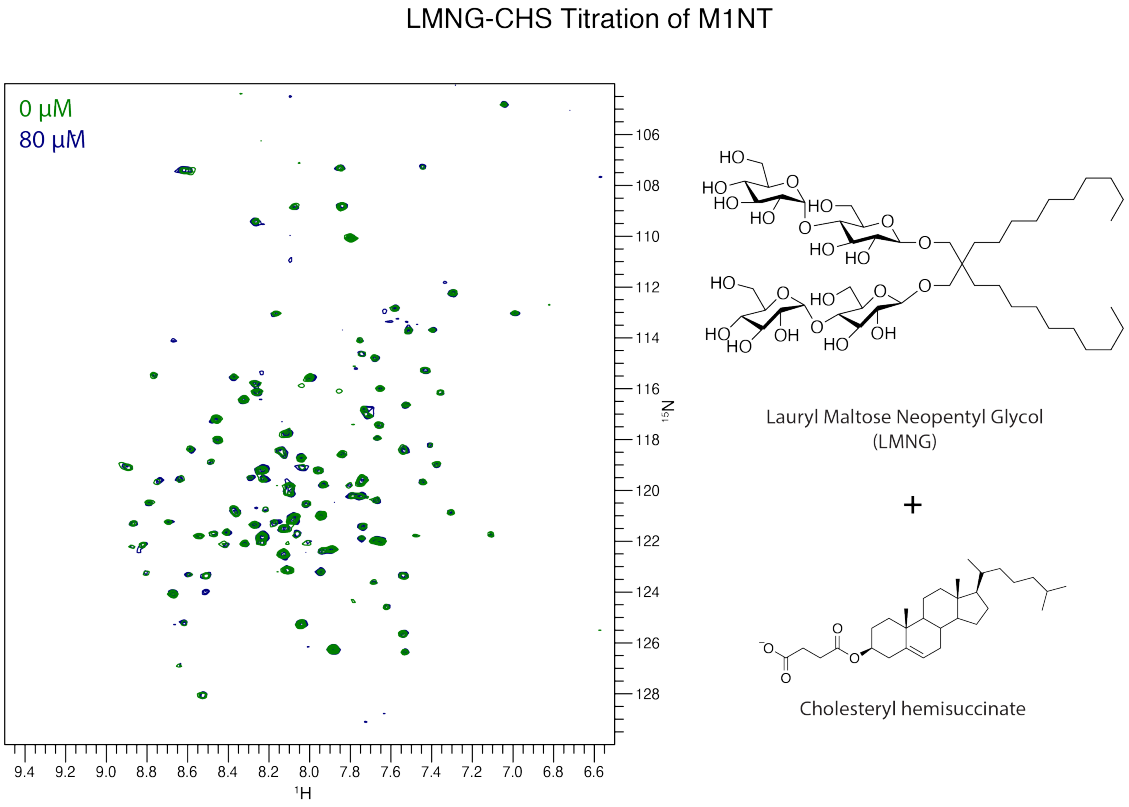


Figure 5.16: LMNG-solubilised CHS titration of M1NT. Overlay of M1NT in the presence of LMNG at 80 μM before (green) and after CHS titration of 80 μM (blue). Spectra collected at 30°C, pH 7 on 750 MHz spectrometers.

5.2.3 Mutants Based on 1EA3 Crystal Structure

From available crystal structures of M1NT, there are two primary arrangements: stacked and side-by-side arrangement. The crystal structure 1EA3 shows a stacked dimer arrangement while the 1AA7 structure shows a side-by-side arrangement. Mapping the strongly perturbed residues L4, G34, T67 and E141 by the presence of sterol-containing surfactants onto the available crystal structures of M1NT has shown that these residues reside at the ends of helices 1, 2, 4 and 9, respectively, as shown in figure 5.18a. These residues are at the centre of a slightly concave, hydrophobic surface on the N-terminal side of the M1NT domain, as shown in figure 5.18 and figure 5.19. Interestingly, figure 5.19b and 5.19c show that the hydrophobic patch forms one of the interfaces between the stacked dimers in the crystal structure of 1EA3 (figure 5.19a).

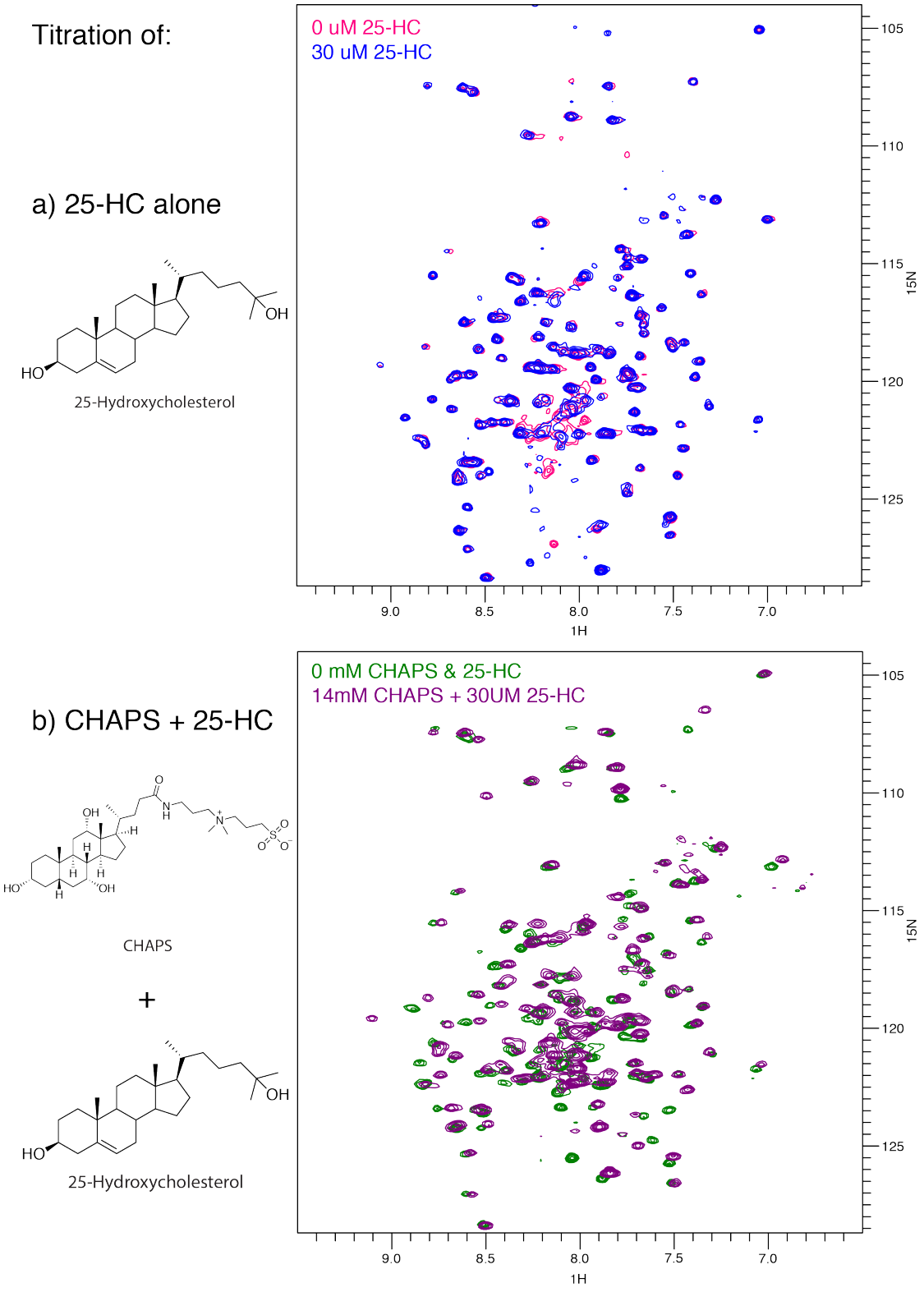


Figure 5.17: 25-HC titration of M1NT. a) Overlay of M1NT at 30 μM before (pink) and after 25-HC titration of 30 μM (blue). b) Overlay of M1NT at 30 μM before (green) and after addition of 14 mM CHAPS and 25-HC titration of 30 μM (purple). Spectra collected at 30°C, pH 7 on 750 MHz spectrometers.

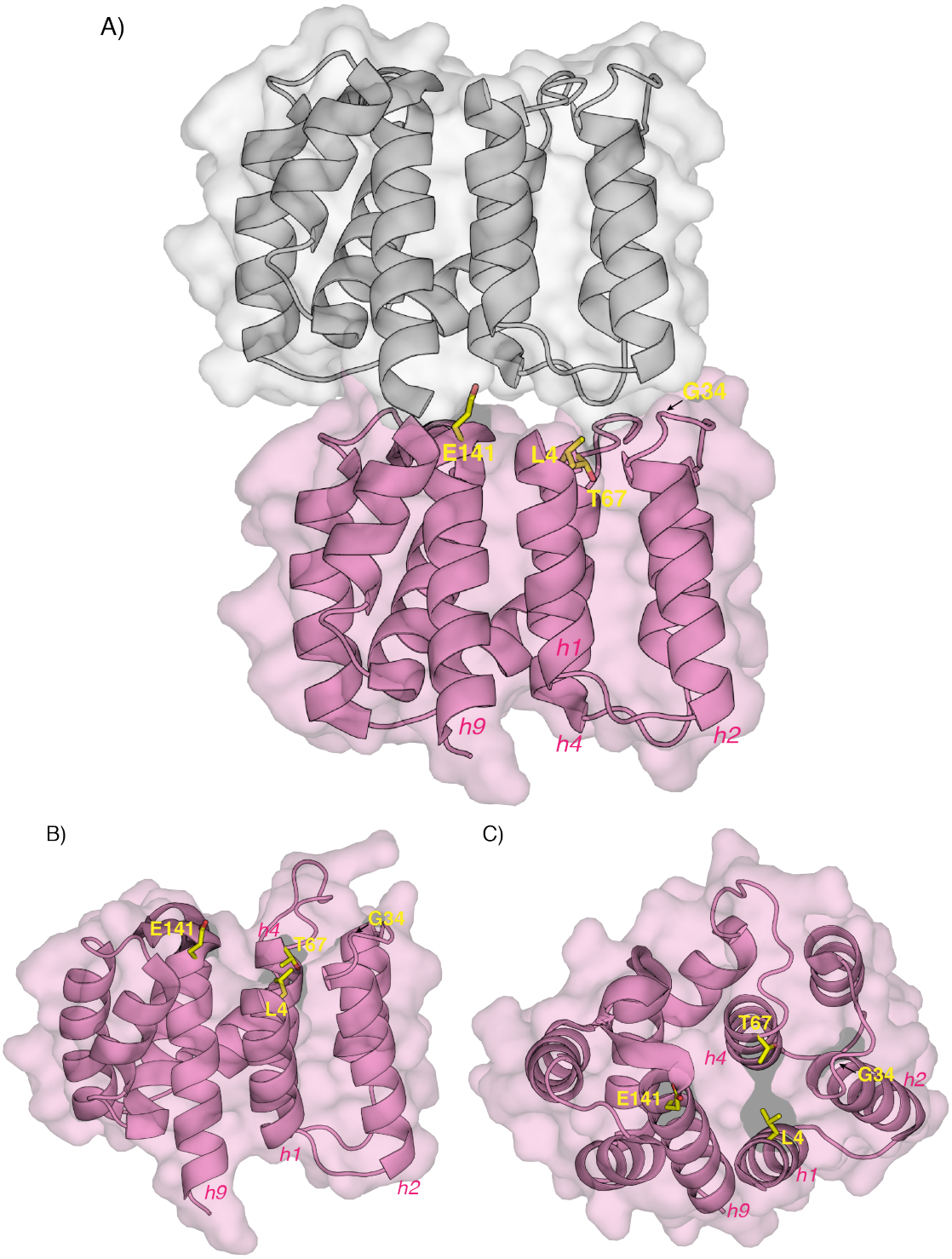


Figure 5.18: Illustration of the a) stacked dimer arrangement of M1NT, 1EA3. b) and c) - monomer of M1NT, 1EA3. It can be seen in b) that the hydrophobic patch has a slightly concave shape. The residues on the hydrophobic patch perturbed by sterols are labelled in yellow and the helix labelled in red.

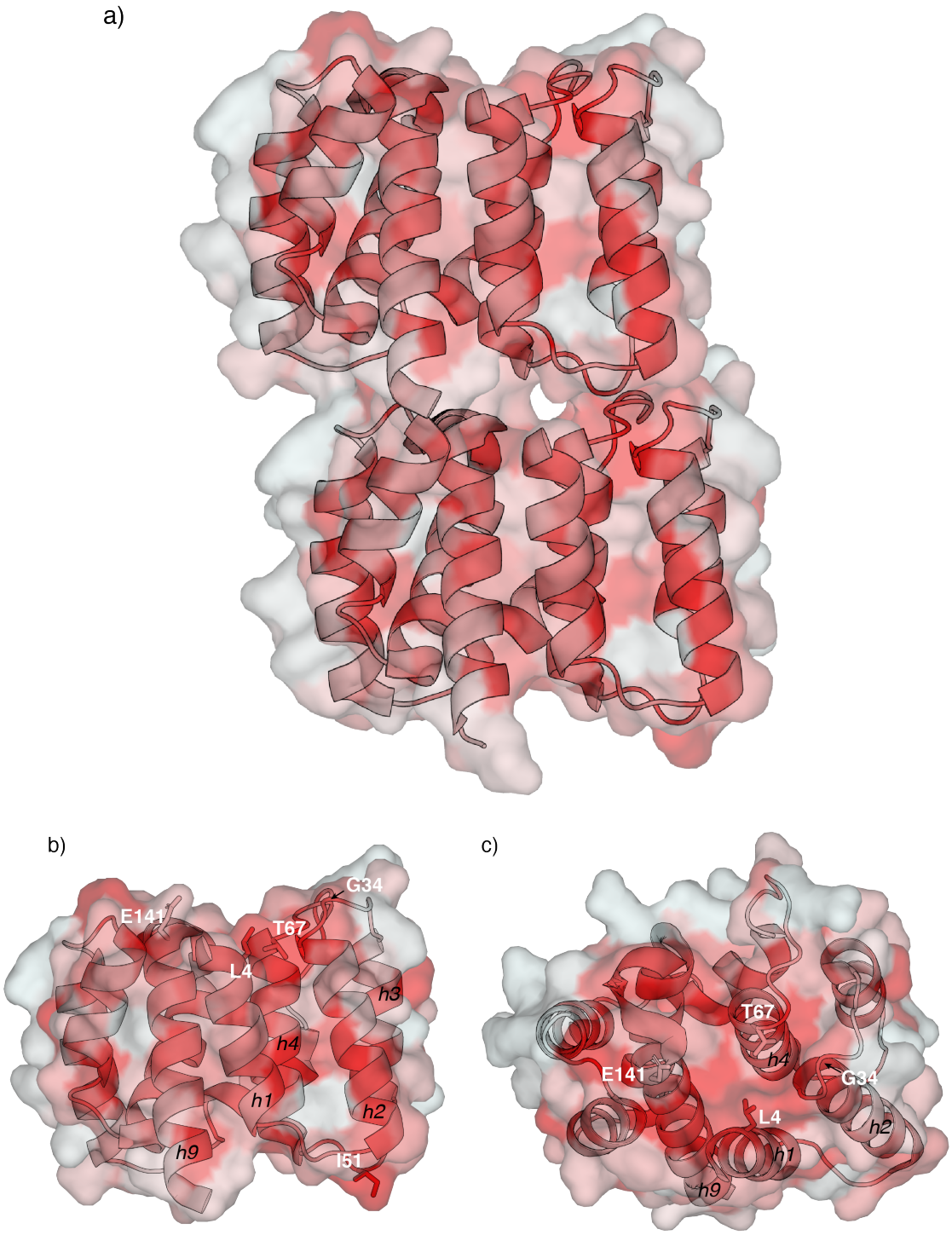


Figure 5.19: Hydrophobicity illustration of the crystal structure of M1NT, 1EA3 in which red shows hydrophobic residues and white shows non-hydrophobic residues. a) Dimer of M1NT, 1EA3. b) and c) - monomer of M1NT, 1EA3 showing the hydrophobicity of the hydrophobic patch at the interface of the dimer at which the subunits interact. The residues of the hydrophobic patch are shown in white, and the helices are labelled in black.

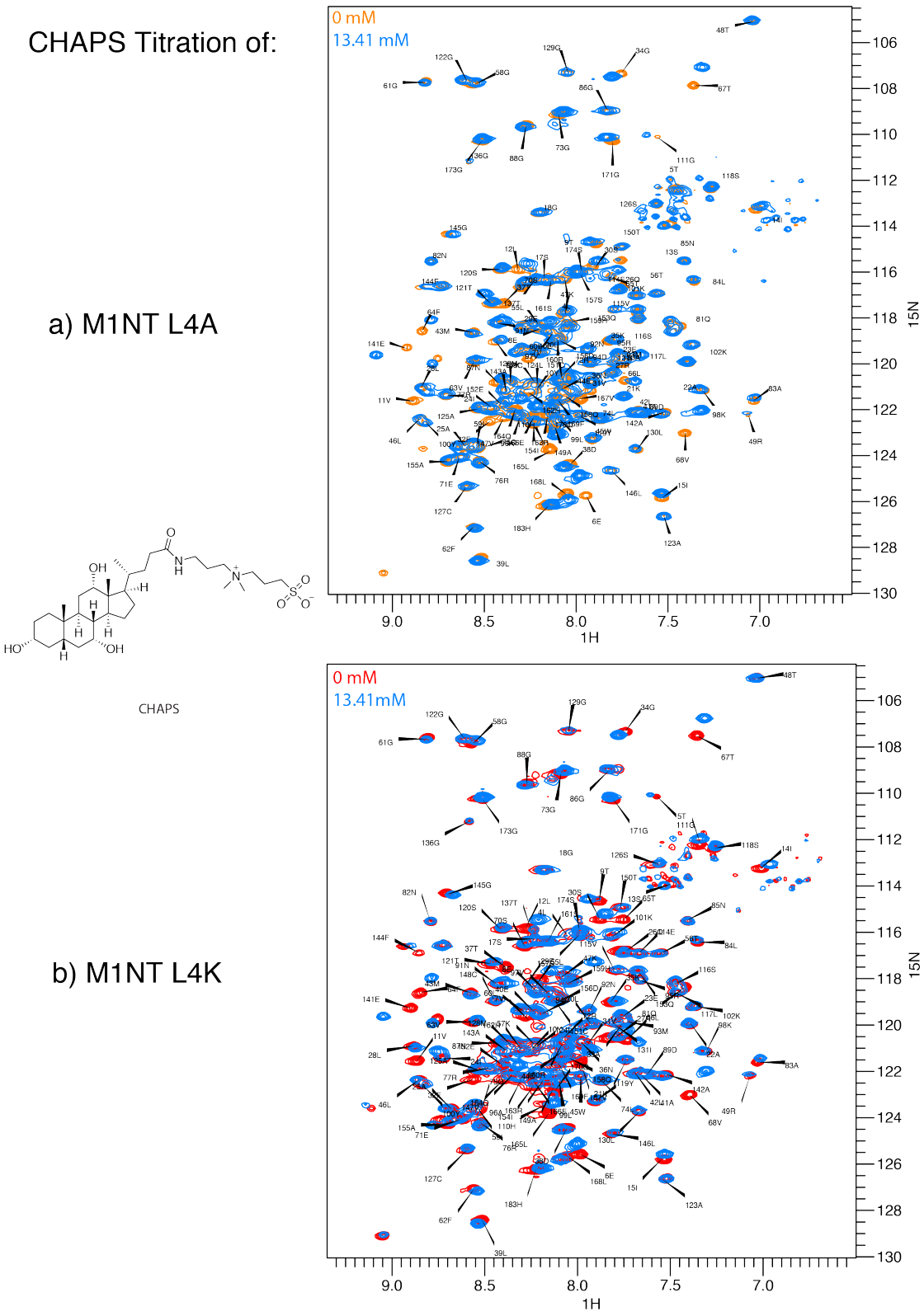


Figure 5.20: CHAPS titration of M1NT Mutants: a) M1NT-L4A and b) M1NT-L4K. a) Overlay of M1NT-L4A at 50 μM before (orange) and after CHAPS titration to 13.41 mM (blue). b) Overlay of M1NT-L4K at 50 μM before (red) and after CHAPS titration to 13.41 mM (blue). Spectra collected at 30°C, pH 7 on 750 MHz spectrometers.

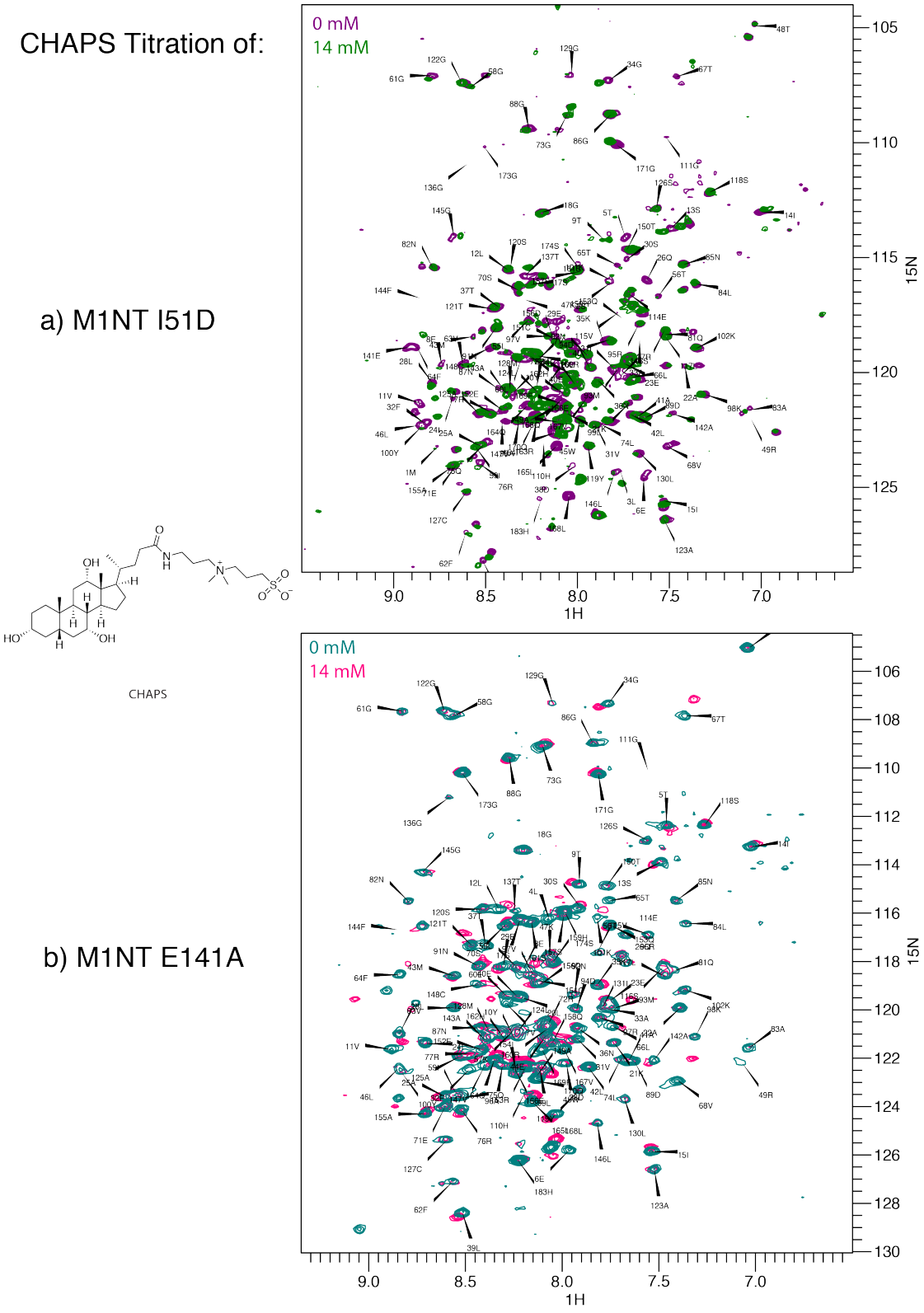


Figure 5.21: CHAPS titration of M1NT Mutants : a) M1NT-I51D and b) M1NT-E141A. a) Overlay of M1NT-I51D at 50 μM before (purple) and after CHAPS titration to 14 mM (green). b) Overlay of M1NT-E141A at 50 μM before (green) and after CHAPS titration to 14 mM (Magenta). Spectra collected at 30°C, pH 7 on 750 MHz spectrometers.

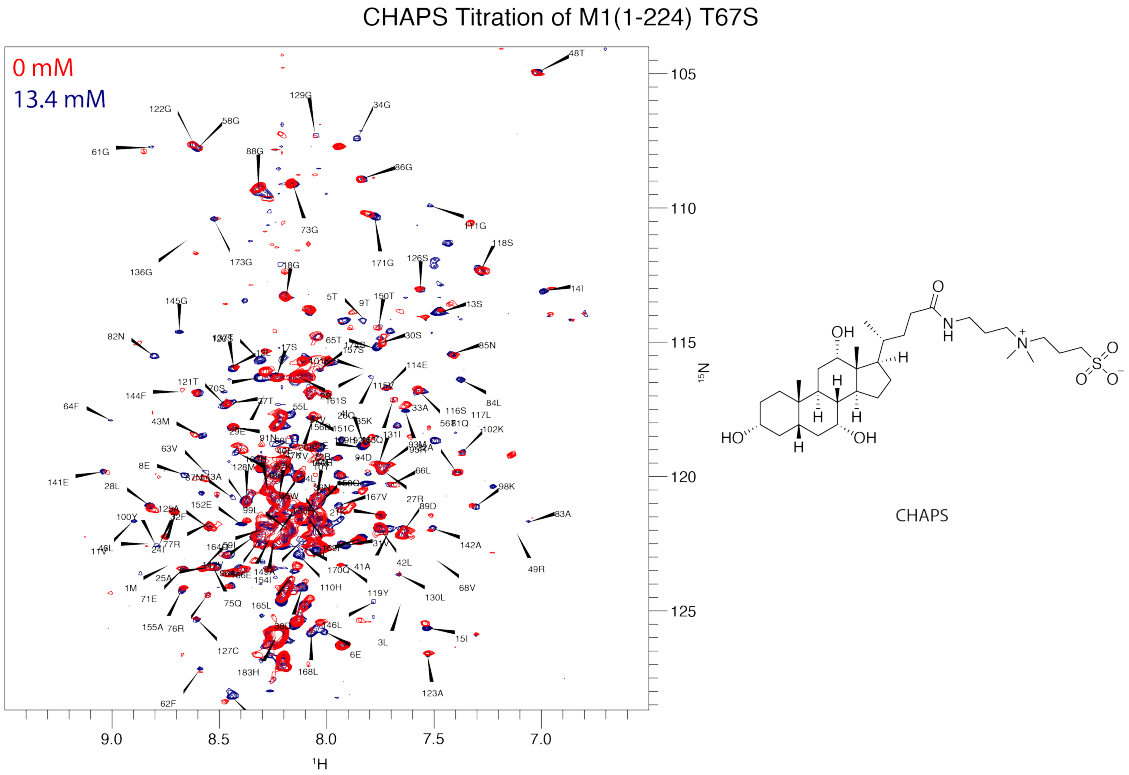


Figure 5.22: CHAPS titration of M1(1-224)-T67S mutant. Overlay of M1(1-224)-T67S at 50 μM before (red) and after CHAPS titration to 13.4 mM (blue). Spectra collected at 30°C, pH 7 on 750 MHz spectrometers.

Next, we tried to test whether the interface of the subunit packing seen in 1EA3 is involved in CHS-induced association. This was done by introducing the single-site mutations L4A, L4K, T67, I51D and E141A. Each single-site mutation is an individual construct. Expression and purification were done similarly as wild ty pe M1 as described in Chapter 2. The introduction of isoleucine in place of the aspartic acid at residue 51, I51D (figure 5.19b) diminished the hydrophobicity of the loop between helix 3 and helix 4 that protrudes itself into the hydrophobic surface cavity of an adjacent subunit. Furthermore, the E141A mutation was used to disrupt an inter-subunit electrostatic interaction with the serine, 54S at the amino-terminal end of helix 4.

These mutants show a well-folded M1NT from their spectra, and they remained soluble at higher concentrations (450μM) than wildtype. Furthermore, these mutant samples remained stable for longer (about 7-8 days vs. 2-3 days) than the wildtype M1NT sample. First, we tested if these mutants still bind to CHAPS.

All the mutants were able to respond to the addition of CHAPS, as shown in figures 5.20, 5.21 and 5.22.

Oligomerisation was then tested on these mutants by adding CHS and CHAPS at both acidic and neutral pH. Interestingly, results show that none of these mutants self-associate at either pH 5.5 and pH 7, as shown in figures 5.23, 5.24 & 5.25.

5.2.4 Mutants based on 1AA7 Crystal Structure

As mentioned above, the two main arrangements frequently seen in M1NT crystal structures are stacked (1EA3) and side-by-side arrangement (1AA7). We then tested mutagenesis residues found at the interface of the side-by-side dimer shown in M1NT 1AA7 crystal structure. A double mutant of M93A/D94A was produced based on the interface formed by the M1NT dimer 1AA7 crystal, as shown in figure 5.26. Expression and purification were done similarly as wild ty pe M1 as described in Chapter 2. These mutations are within helix 6 on the large buried surface area. M93 are buried at the dimer interface and forms hydrophobic interactions with V97 and Y100 (helix 6) as well as L130 (helix 8) of another subunit (Figure 5.26b and 5.27b). Furthermore, D94 forms salt bridges with K98 and R101 of helix 6 of another subunit as shown in figure 5.26a and 5.27a. The M93A/D94A double mutant should diminish these interactions, as mentioned previously.

The M93A/D94A double mutant shows a well-folded M1NT from their spectra (As shown by the blue spectrum in Figure 5.28), and it remained soluble at the same concentration as the wildtype. CHAPS binding was first tested, and it was shown that the double mutant (M93A/D94A) was still able to bind to CHAPS (Figure 5.28). CHS-induced self-association was then tested on the double mutant by adding CHS and CHAPS at both pH 5.5 and pH 7. The M93A/D94A double mutant behaved similarly to wildtype, undergoing CHS-induced self-association at pH 7.0 but not pH 5.5, as shown in figure 5.29.

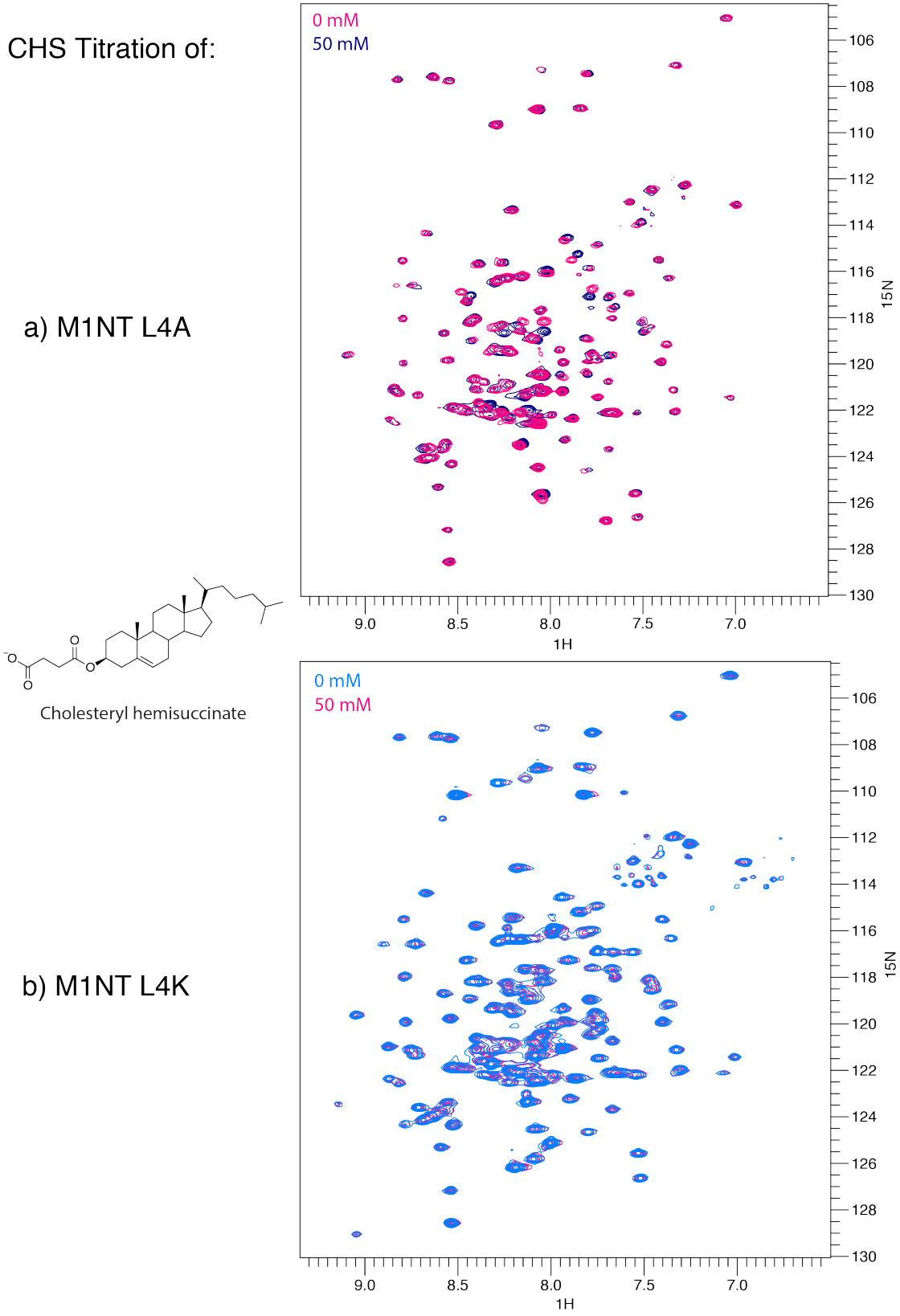


Figure 5.23: CHAPS-solubilised CHS Titration of M1NT-L4A, pH 7 and b) M1NT-L4K, pH 7. a) Overlay of M1NT-L4A with 14mM CHAPS at 50 μM before (red) and after CHS titration of 50 μM (blue). b) Overlay of M1NT-L4K at 50 μM before (light blue) and after CHS titration of 50 μM (pink). Spectra collected at 30 °C, pH 7 on 750 MHz spectrometers.

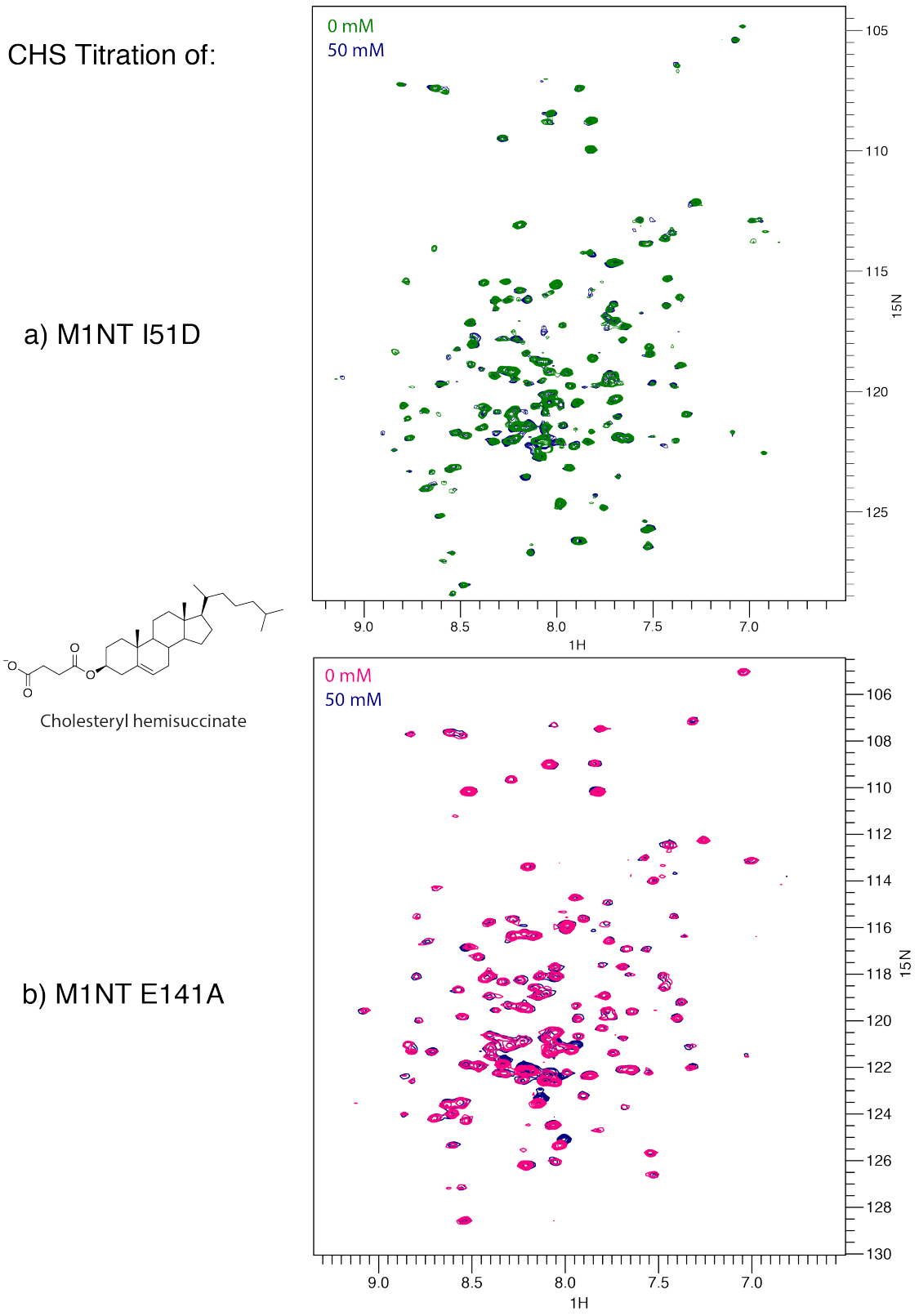


Figure 5.24: CHAPS-solubilised CHS Titration of a) M1NT-I51D, pH 7 and b) M1NT-E141A, pH 7. a) Overlay of M1NT-I51D at 50 μM before (green) and after CHS titration of 50 μM (blue). b) Overlay of M1NT-E141A at 50 μM before (pink) and after CHS titration of 50 μM (blue). Spectra collected at 30 °C, pH 7 on 750 MHz spectrometers.

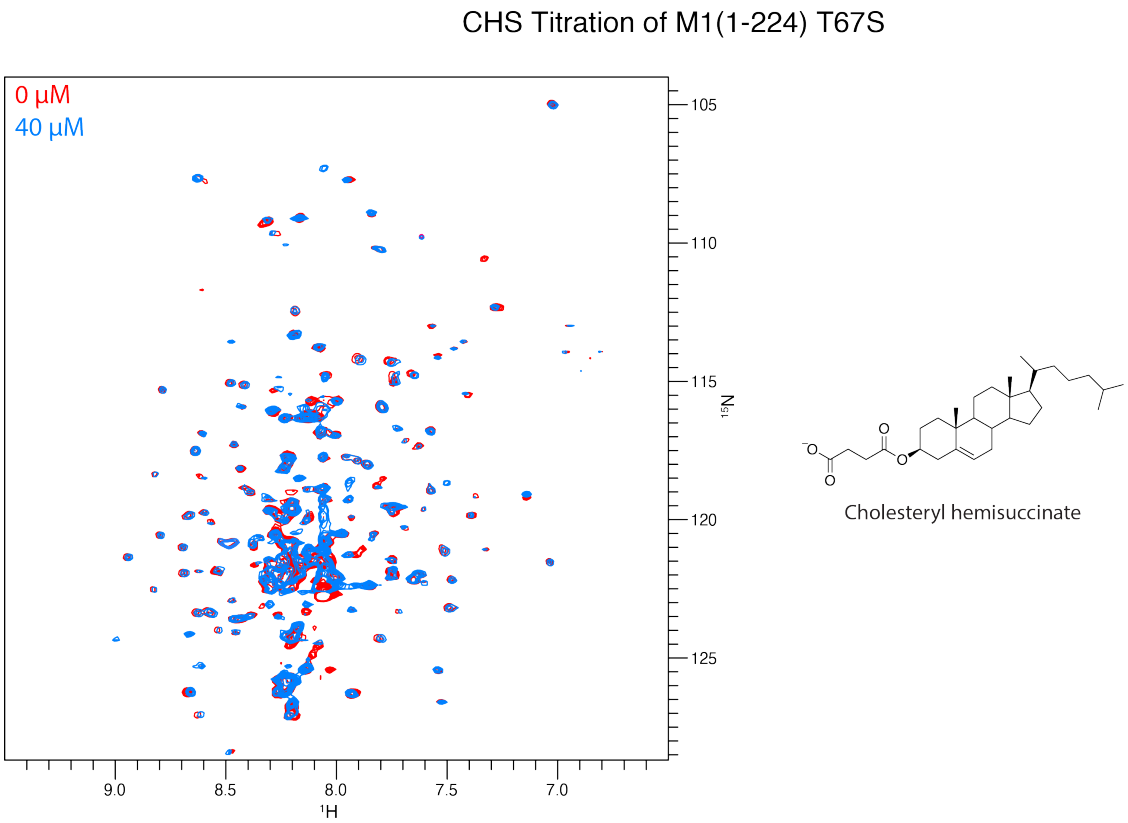


Figure 5.25: CHAPS-solubilised CHS Titration of a) M1(224)-T67S, pH 7. Overlay of M1(224)-T67S at 40 μM before (red) and after CHS titration of 40 μM (light blue). Spectra collected at 30 °C, pH 7 on 750 MHz spectrometer.

5.2.5 CHAPS Binding of M1 and Its Mutants

5.2.5.1 Chemical Shift Perturbation (CSP) of T67 Resonance of Wild-type and Mutants M1

As mentioned previously, the specific combination of a bile acid-derived compound and CHS at pH 7.0 was required for M1 self-association (Figure 5.14. Mutants based on the M1NT 1EA3 crystal structure, M1NT-L4A, L4K, I51D, T67S or E141A were not able to induce M1-self association at both pH 7.0 and pH 5.5 (Figure 5.23, 5.24 & 5.25). However, a double mutant based on the M1NT 1AA7 structure, M1NT-M93A/D94A was able to induce M1-self association. The T67 resonance CSP was measured for all mutants and compared with the wild type of M1 (Figure 5.30). The result shows that the CSPs for the wildtype at pH 7 and 5.5 show a difference in T67 resonance movement upon CHAPS binding.

Interestingly, the spectra of the wildtype at pH 5.5 and mutants that were based

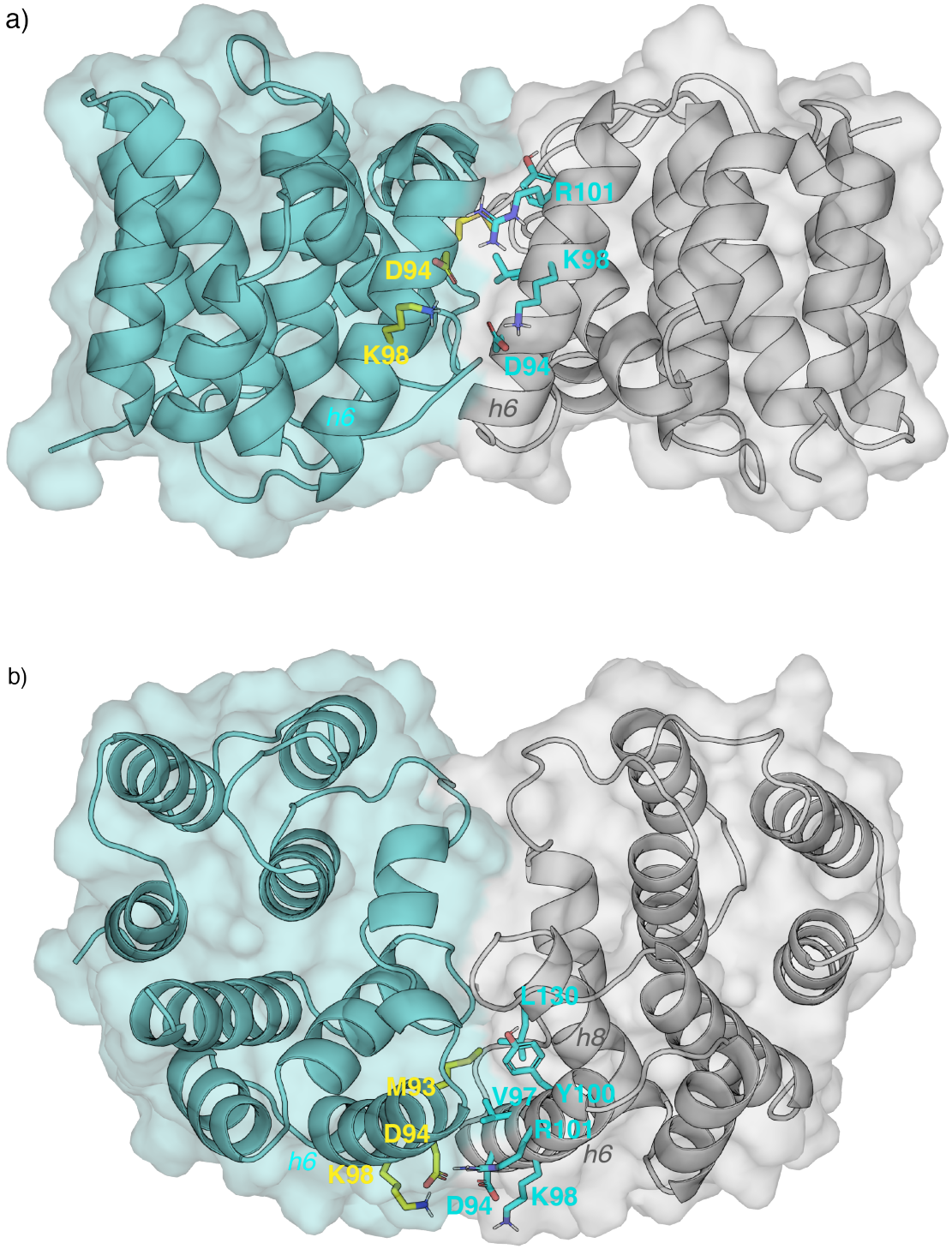


Figure 5.26: Illustration of the side-by-side dimer arrangement of M1NT, 1AA7. a) - side view and b) - top view show the dimer arrangement of two subunits of M1NT. A double mutant of M1NT-M93A/D94A was made to test the importance of the side-by-side dimer interface. a) shows the important contacts of D94 that forms salt bridges to K98 with other subunits and b) shows the critical contacts of M93 interaction with V97, Y100 and L130.

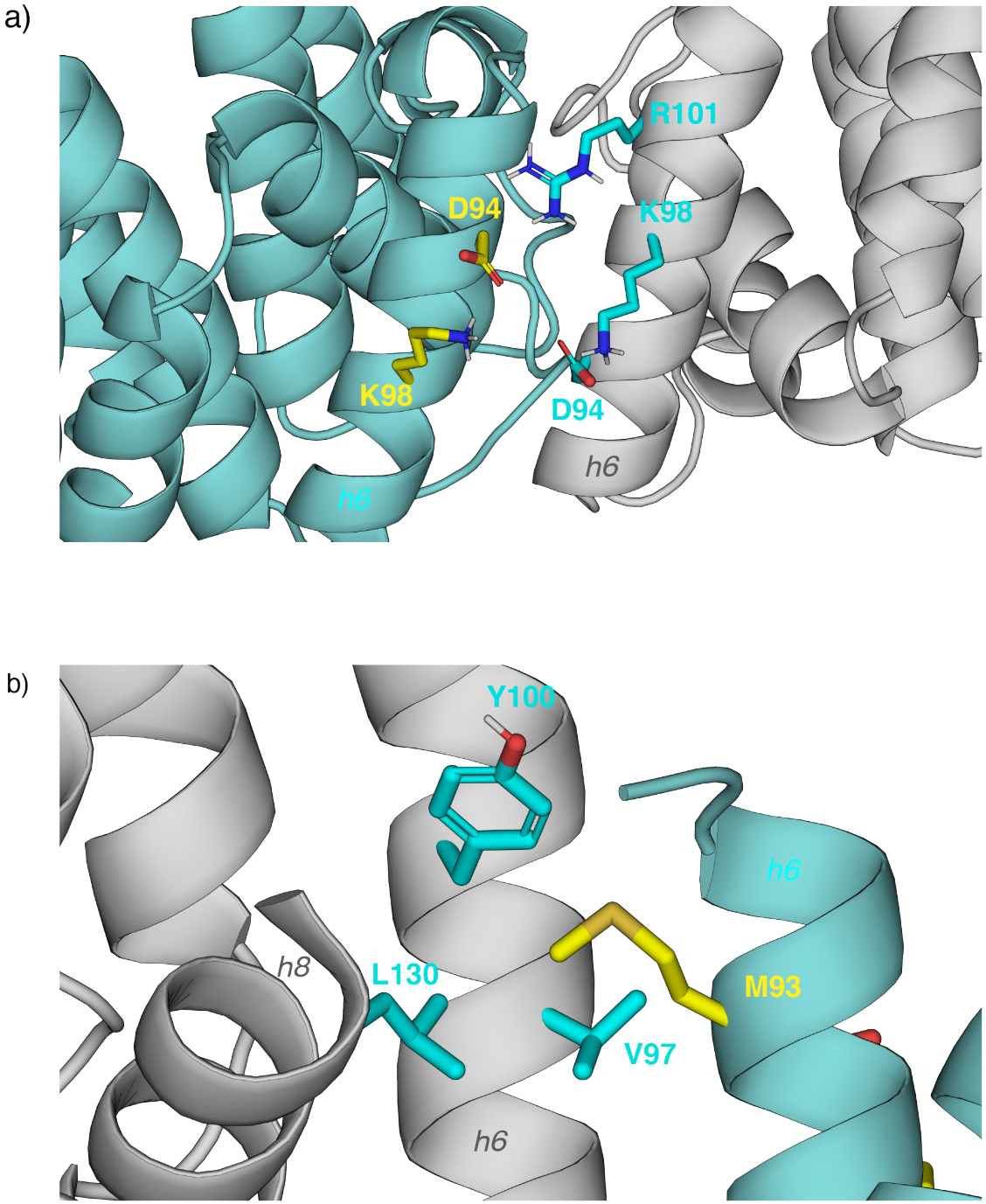


Figure 5.27: Illustration of the important contacts of M93 and D94 in M1NT, 1AA7 crystal. a) shows the important contacts of D94 that forms salt bridges to K98 and R101 with other subunits of helix 6 and b) shows the important contacts of M93 interaction with V97, Y100 and L130 from helix 6 and 8.

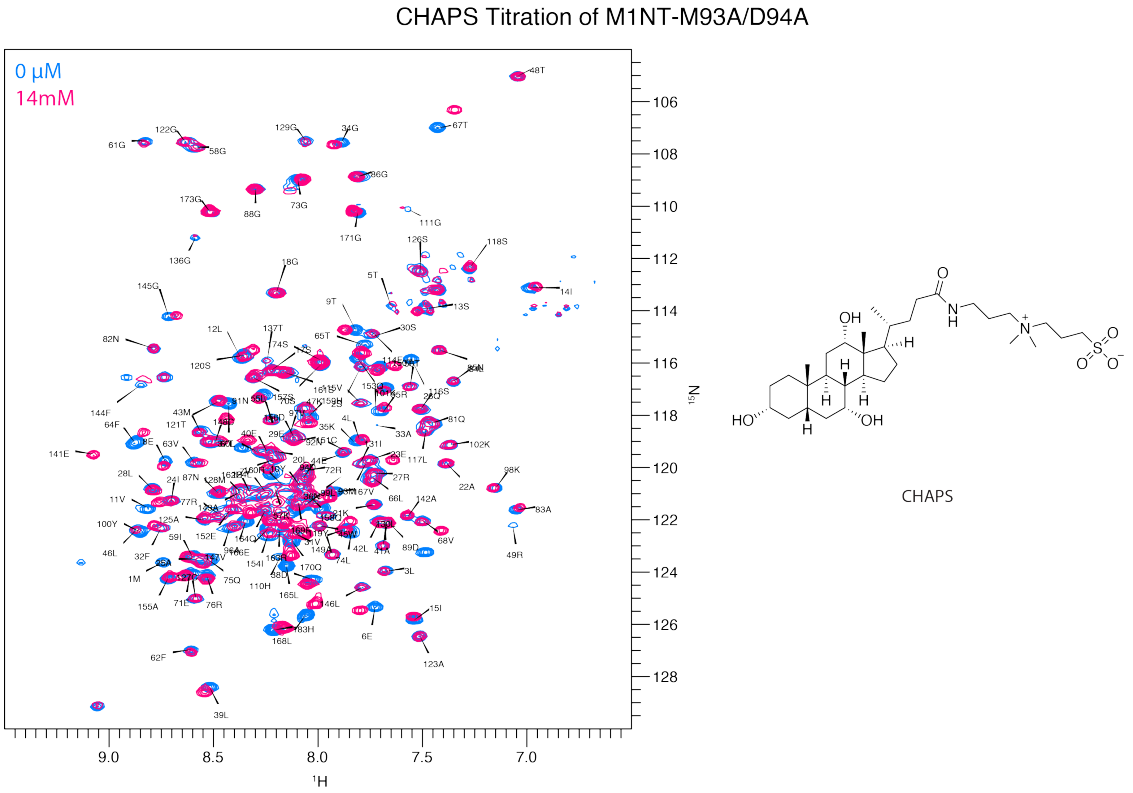


Figure 5.28: CHAPS titration of M1NT-M93A/D94A Mutant. a) Overlay of M1NT-L4A at 50 μM before (blue) and after CHAPS titration to 14 mM (magenta). Spectra collected at 30°C, pH 7 on 750 MHz spectrometers.

on the 1EA3 structure (L4A, L4K, I51D and E141A) indicate a comparable CSP. As mentioned previously, the wildtype M1NT and M1(1-224) were not able to induce M1 self-association upon the addition of CHAPS-solubilised CHS at pH 5.5. This could suggest that the T67 movement correlates with the ability to oligomerise. Furthermore, when the CSP for the double mutant M1NT-M93A/D94A based on the 1AA7 structure was compared with the wildtype M1NT at pH 7, it can be seen that both CSPs are of a comparable magnitude. M1NT-M93A/D94A at pH 7 was able to self-associate under CHAPS-solubilised CHS oligomerisation.

Intriguingly, the longer construct, M1(1-224), which have a higher propensity to self-associate, exhibited a greater CSP when CHAPS was added (FIgure 5.10) as compared to the M1NT construct (Figure 5.30). When pH was lowered to pH 5.5, at which M1 does not efficiently oligomerise, there was a decrease in the CSP observed for T67 resonance. Thus, the sterol-induced perturbations of the T67

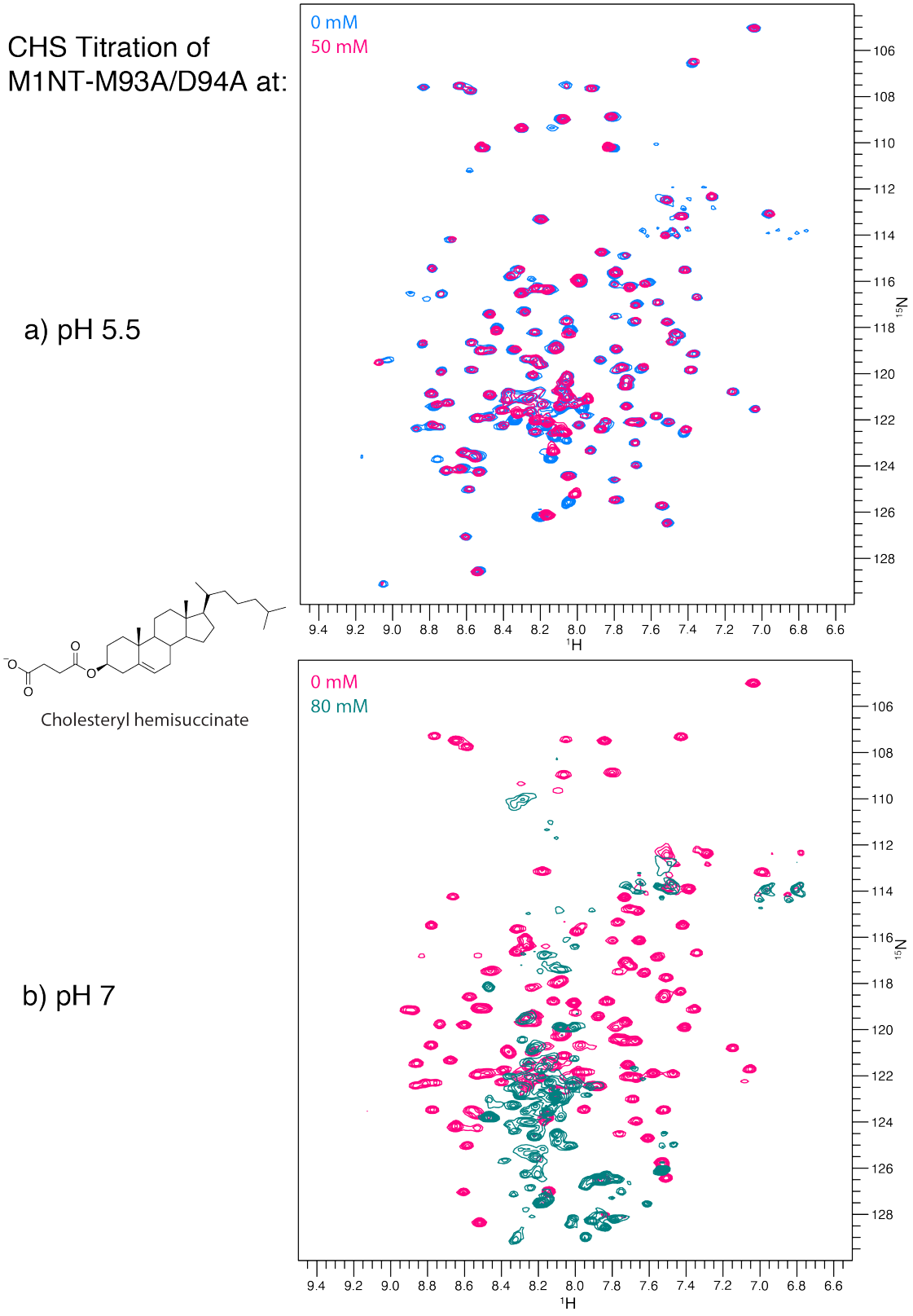


Figure 5.29: CHAPS-solubilised CHS Titration of M1NT-M93A/D94A at a) pH 5.5 and b) pH 7. a) Overlay of M1NT-M93A/D94A at 50μM before (blue) and after CHS titration of 50 μM (pink). b) Overlay of M1NT-M93A/D94A at 80μM before (pink) and after CHS titration of 80 μM (teal). Spectra collected at 30°C, pH 7 on 750 MHz spectrometers

surface patch correlated with the propensity of M1NT to self-associate.

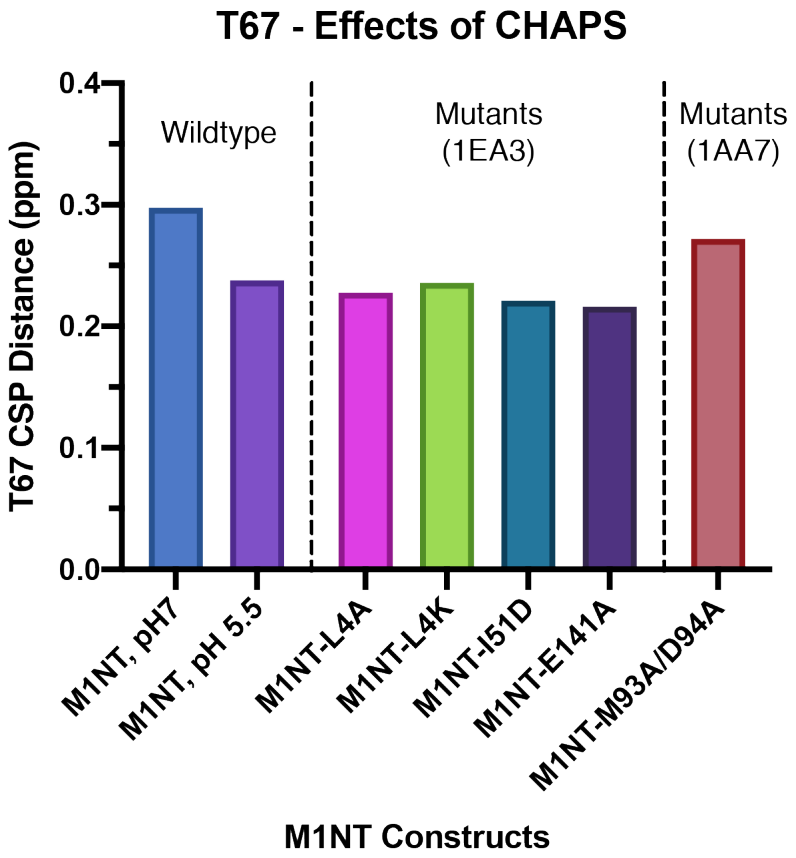


Figure 5.30: Effects of CHAPS at saturation concentration on the T67 resonance on the different constructs of M1NT - Wildtype (pH 7 and 5.5) and Mutants - L4A, L4K, I51D, E141A and M93A/D94A. All M1 constructs are at 50μM at pH 7.0.

5.2.5.2 Backbone Confirmation of T67 Resonance of M1 Crystal Structures

Among all the available M1NT crystal structures, the 1EA3 solved at pH 7 is the most unique dimer, organising with a stacked dimer arrangement (Figure 5.18a). Other crystal structures such as 1AA7 and 4PUS solved at acidic pH showed a side-by-side dimer arrangement (Figure 5.26). Furthermore, M1NT mutants based on 1EA3 structure were not able to induce self-association upon the addition of CHAPS-solubilised CHS (Figure 5.23, 5.24 & 5.25). This is in contrast with the M1NT double mutant based on the 1AA7 structure, M1NT-M93A/D94A, which was able to induce M1-M1 self-association when CHAPS-solubilised CHS was added (Figure 5.29).

A more in-depth analysis was then done to compare the 1EA3 and 1AA7 structures. Interestingly, it was found that the T67 in the 1EA3 structure had a distinct backbone conformation that is extended. This distinct confirmation was compared with other crystal structures of M1NT such as 1AA7 and 4PUS. The T67 is uniquely extended, Φ = -80° and Ψ = -165°, in comparison with other structures with confirmations of Φ = -93 ± 17° and Ψ = -22 ± 21°.

5.2.5.3 Determination of T67 ψ Torsion Angle with CHAPS

The CSPs induced by sterol-containing surfactants are highly specific and involved the T67 resonance. However, no NOEs can be detected directly with M1. Hence, we then tested whether these specific shifts are associated with a similar change in the Ψ backbone angle by measuring the 13C' and 13Cα chemical shifts in the absence and presence of CHAPS. The rationale behind this is that we saw a change in the backbone angle between the crystal structures - PDB:1EA3 & 1AA7 structures and we wanted to know if CHAPS had an effect in the backbone angle (detailed explanation can be find later in the chapter). The wild-type M1NT is less stable at high concentrations for weeks at 30°C in the NMR spectrometer, and hence, the mutant M1NT-L4A was instead used for these experiments. M1NT-L4A responds to the addition of CHAPS similarly to the wild-type M1NT as can be seen in figure 5.20a. M1NT-L4A is also highly stable at high concentrations at a more extended period of time at 30°C, which makes it a suitable sample for these experiments. It should be noted that M1NT-L4A is used as a mutant for investigation as it does not form oligomers.

The wild-type M1NT and mutant M1NT-L4A were compared at pH 7.0, and they showed similar spectra with some perturbations as can be seen in figure 5.32a. These chemical shift changes seen between the wild-type and mutant M1NT, M1NT-L4A are quantified and plotted in figure 5.32b. Apart from helix 1 where the mutation L4A is, the largest perturbations are in the G34 loop. In addition, there are also changes around T67 and E141. When these hotspots were mapped, it can be seen in figure 5.32c that L4A packs against F32, which is close to the G34

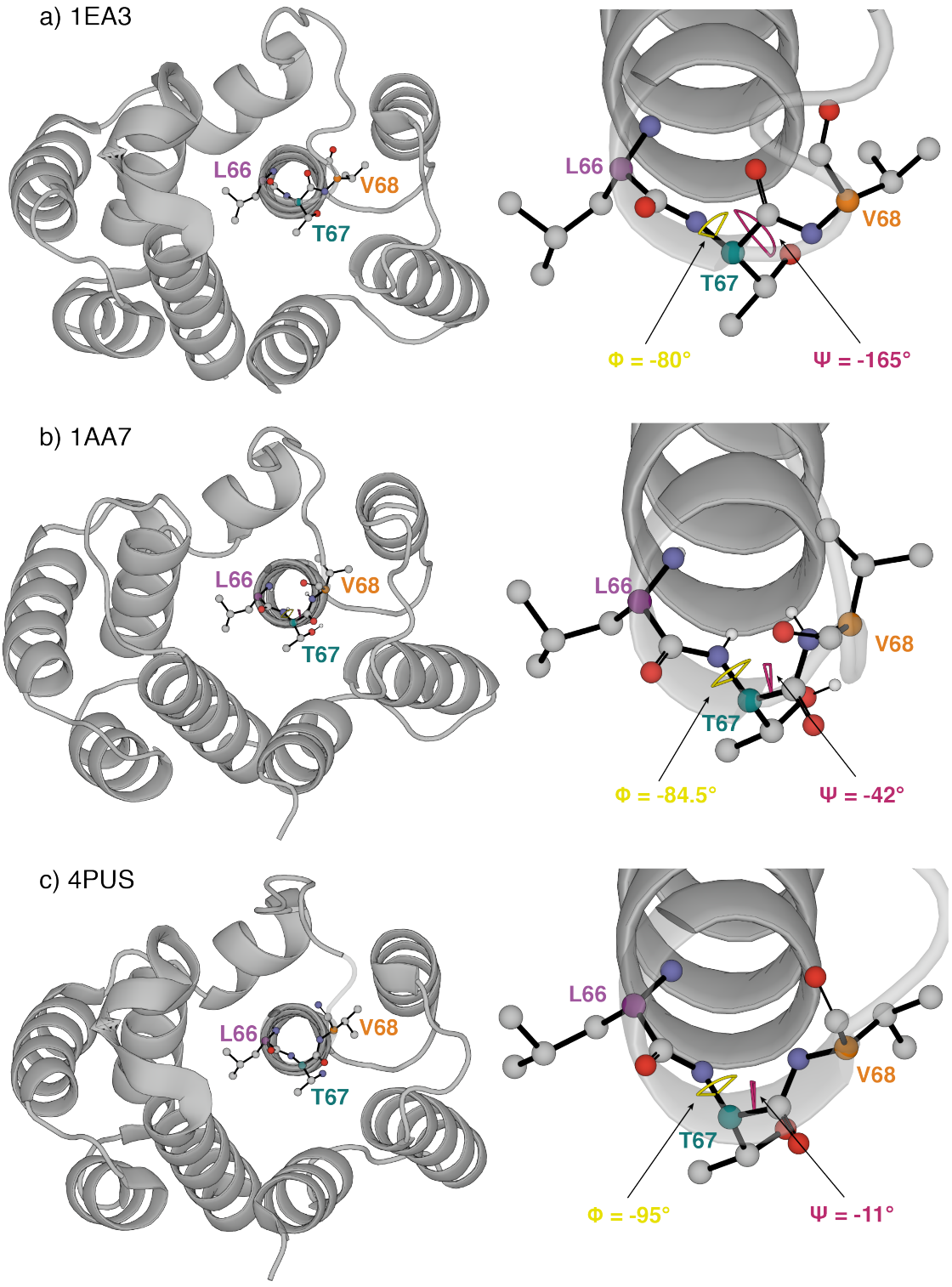


Figure 5.31: T67 backbone confirmation of various M1NT crystals - a) 1EA3, b) 1AA7 & c) 4PUS. The T67 backbone confirmation for a) 1EA3 crystal is extended (Φ = -80° and Ψ = -165°) as compared to other crystal structures for example - b) 1AA7 (Φ = -84.5°and Ψ = -42°) c) 4PUS (Φ = -95° and Ψ = -11°).

loop of helix 2-helix 3 (h2-h3), and also against T67 and F144 of helix 4 & 9. Is it interesting that a change in the L4A residue in helix 1 at the hydrophobic patch propagates changes to the end of helix 9 via allosteric effect. These perturbations may probe some insights into the inability of the mutant M1NT-L4A and other mutants to self-associate upon addition of CHAPS-stabilised CHS.

The chemical shift changes were measured (Figure 5.33a) and seen to be shifted 0.56 ppm upfield for the T67 13C' upon the addition CHAPS (Figure (5.33b). Hence, this is an indication of a more extended conformation. Furthermore, the 3-bond Cα-Cα coupling (^{3}J(Cα-Cα)) was also measured, and this measurement reflects the backbone conformation (Hennig, 2000). This was measured for the T67-V68 inter-residue, and the ^{3}J(cα-cα was 1.51 Hz in the absence, and 0.50 Hz in the presence, of CHAPS. The difference of 1.1 Hz confirms a more extended backbone conformation when CHAPS is present.

Interestingly, perturbations were also seen for the backbone ^{13}C chemical shifts for the residues F32, L66, and F144 (Figure (5.33). These residues, as can be seen previously in figure 5.32c, are in the same plane and tightly packed against L4 and T67. Furthermore, it can be seen that structural changes in the hydrophobic patch upon addition of CHAPS can be transmitted between residue T67, and the residues of helix 9 (E141 and F144) via allosteric conformational change. When structural alignment is done on the outer helices 2, 3, and 6 of 1EA3 and 1AA7, the monomers indicate that helices 1, 4, 7, and 9 are shifted toward the N-terminal face of the M1NT domain, where the hydrophobic patch is (Figure 5.33 & 5.34). Furthermore, the ^{13}C' and ^{13}Cα of residues at the end of helix 9 showed a slight downfield shift which is an indication of increased helix formation (Figure 5.33b). This suggests that the changes that occur at the N-terminal face of the M1NT can be propagated through allosteric conformational change to the M1 C-terminal tail, and vice versa.

CHAPS effects on the CSPs of M1NT is likely to be an indirect effect, as NOEs between M1NT and CHAPS could not be detected either from the backbone amides or the side-chain methyls. High concentrations of sterol-containing surfactants are also capable to cause the CSPs observed in M1NT. In light of this, the critical micelle

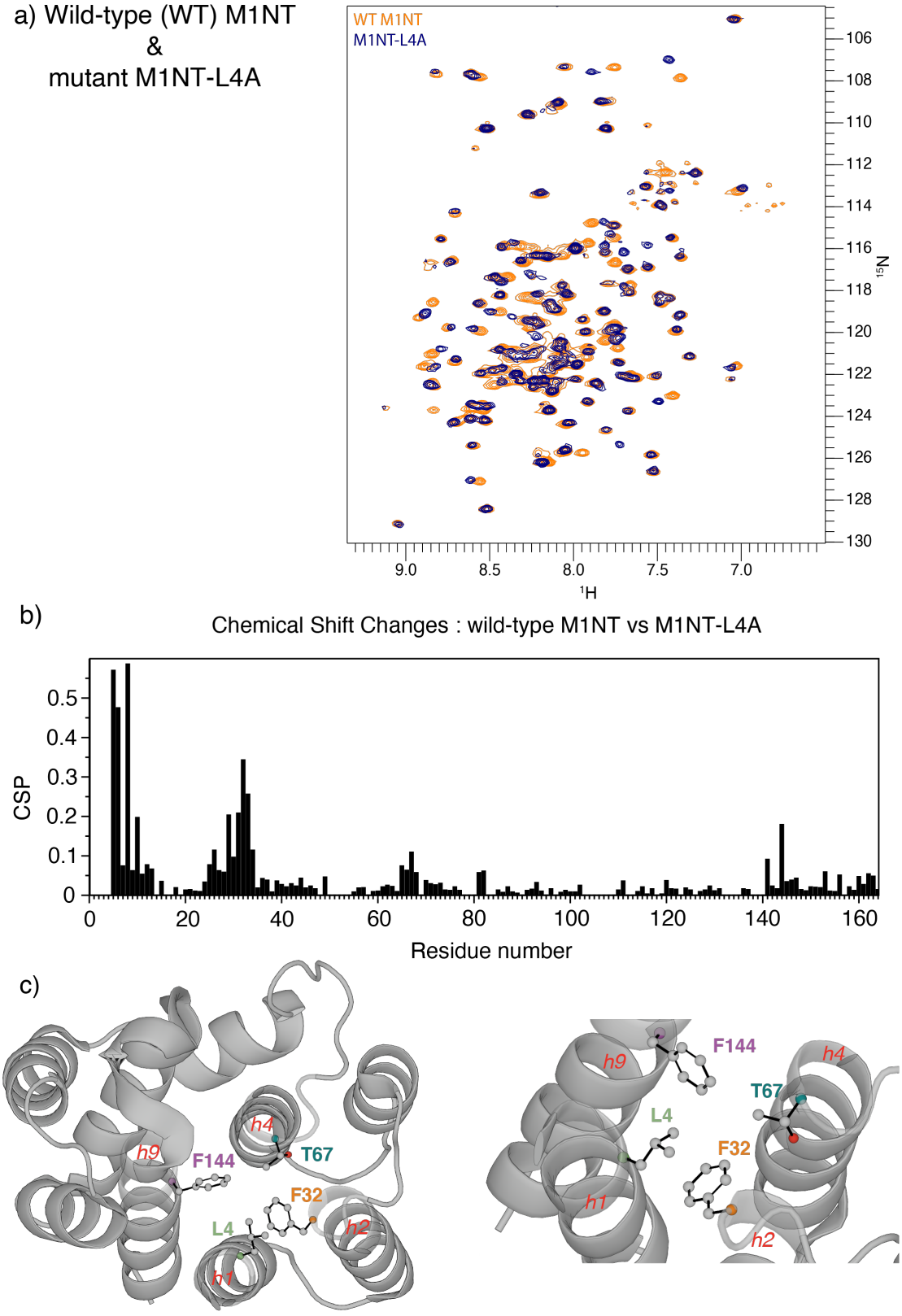


Figure 5.32: Wild-type (WT) & mutant M1NT-L4A. a) Overlay of WT M1NT at 50 μM (Orange) and mutant M1NT-L4A at 60 μM (blue). Spectra collected at 30°C, pH 5.5 on 750 MHz spectrometers. b) Quantification of the chemical shift perturbations (CSPs) between WT M1NT and mutant M1NT-L4A. c) L4 residue in helix 1 (*h1*) can be seen to be tightly packed against F32, T67 and F144 of helix 2 (*h2*), helix 4 (*h4*) and helix 9 (*h9*), respectively. The crystal structure used is PDB:1AA7

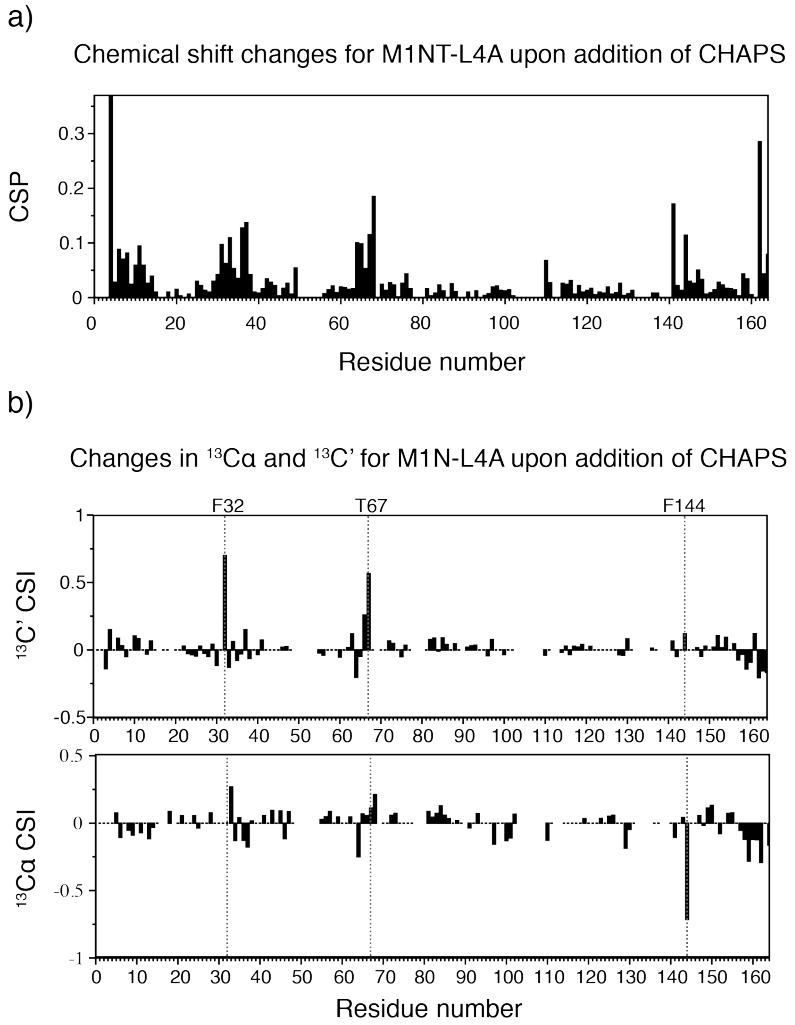


Figure 5.33: Chemical shift changes (CSPs) and Chemical Shift Index (CSIs) difference of M1NT-L4A at 200μM upon addition of CHAPS at pH 7.0, 30°C. a) The CSP per residue when CHAPS upon addition of CHAPS were recorded and plotted. As expected, the CSPs at the hydrophobic patch residues changed the most. b) The CSI difference for ^{13}C' and ^{13}Cα per residue was also measured after the addition of CHAPS. F32 and T67 ^{13}C' chemical shifts have been shifted 0.63 and 0.51 ppm upfield, respectively. Furthermore, F144 ^{13}Cα has been shifted downfield by 0.65 ppm.

concentrations (CMC) was checked for each sterol containing surfactants that were used from CHAPS to deoxycholate. Interestingly, the saturation concentrations of maximum exhibited CSPs correlated well with the CMCs of each sterol surfactants, as seen in table 5.1. This possibly suggests that CHAPS may exert its effect through its sterol monomers and possibly through water activity or water ordering/structure as no NOEs can be detected directly with M1. In addition, IPA, which has a different chemical structure to sterol-containing compounds, has a similar effect

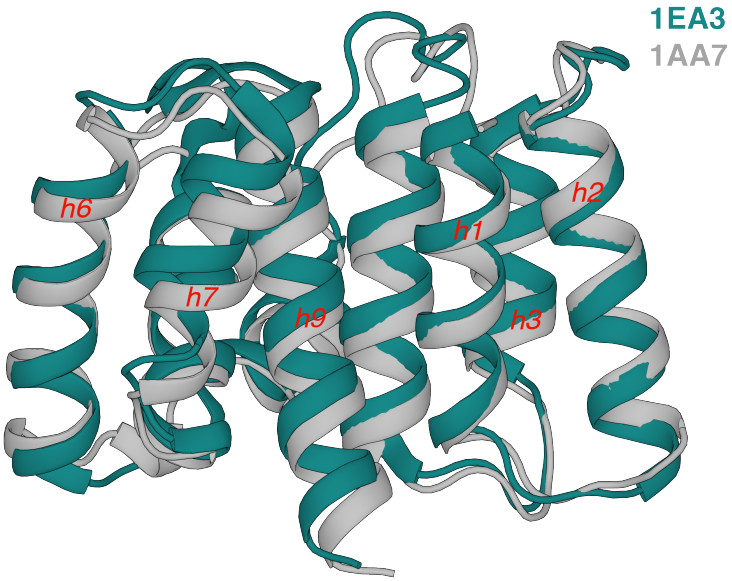


Figure 5.34: Structural alignment between M1NT crystals - 1EA3 (Teal) and 1AA7 (gray). Alignment of the outer helices 2, 3, and 6 of these two crystals shows that the helices 1, 4, 7 and 9 are shifted toward the N-terminal face of the M1NT domain, where the hydrophobic path is.

to CHAPS. This further suggest that water activity or water ordering/structure may play a role in exerting the observed specific CSPs.

Table 5.1: Sterol-containing compounds with its associated critical micelle concentration (CMC) with the saturation concentration of M1NT CSPs

Sterols	Critical Micelle Concentration (CMC)	Saturation Concentration of Sterol
CHAPS	10-14mM	13.4 mM
CHAPSO	12-17mM	17.9 mM
Cholate	13-18mM	17.9 mM

5.2.6 Kosmotropes and Chaotropes

As described above, the effect of sterol-containing compounds such as CHAPS, and IPA on the hydrophobic patch of M1NT might be attributed to water activity or water ordering/structure. Crystal structures of M1NT have shown a large number of bound water molecules that have been modelled near the hydrophobic patch where residue T67 resides. To investigate if water structure plays a role in the effects exerted by sterol-containing compounds and IPA, kosmotropes and chaotropes were tested.

Water is made of the two most reactive elements in the periodic table and is one of the simplest and most ubiquitous compounds known to exist. Water is regarded as a highly structured liquid due to an extensive and dynamic network of hydrogen bonds with localised and structured clustering (Perera, 2011). The ordering or structuring of liquid water depends on two finely balanced processes: 1) enthalpic process & 2) entropic process, to achieve the minimum energy. In an enthalpic process, a drive towards the most optimum four hydrogen bonds for each molecule of water is preferred. This is to maximise the strength (enthalpy) of hydrogen bonds while in contrast, the entropic process drives towards reducing the number and strength of the hydrogen bonds to favour the greater freedom of movement (entropy). These two interactions can be easily shifted from one to another with the changes in physical conditions, solutes, and surfaces.

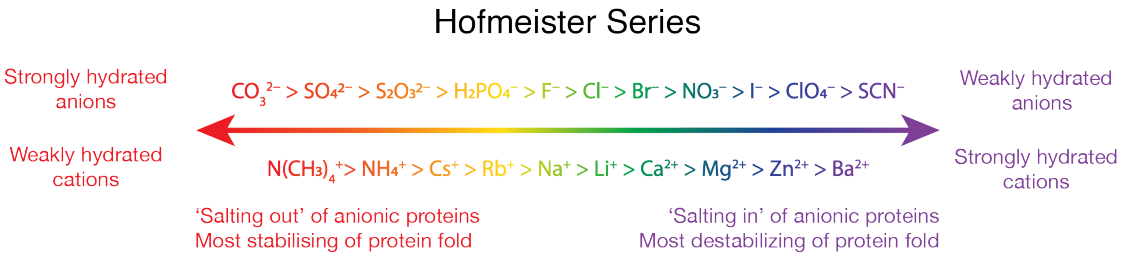


Figure 5.35: The Hofmeister Series.

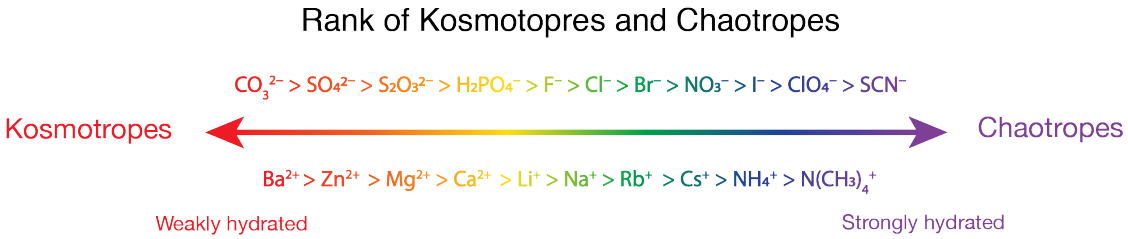


Figure 5.36: Approximate order of kosmotropes and chaotropes in strength. Cl- for the anions and Na+ for the cations are considered as neutral. Any ions on the left side and right side of Cl- and Na+ are considered kosomotropes and chaotropes, respectively. The chaotropes are strongly hydrated and kosmotropes are weakly hydrated.

Electrolytes dissociate in water into ions and are surrounded by water molecules, hence become hydrated. The former structure of bulk water changes when ions dissolve in water. The dipolar water molecules are reordering themselves around the

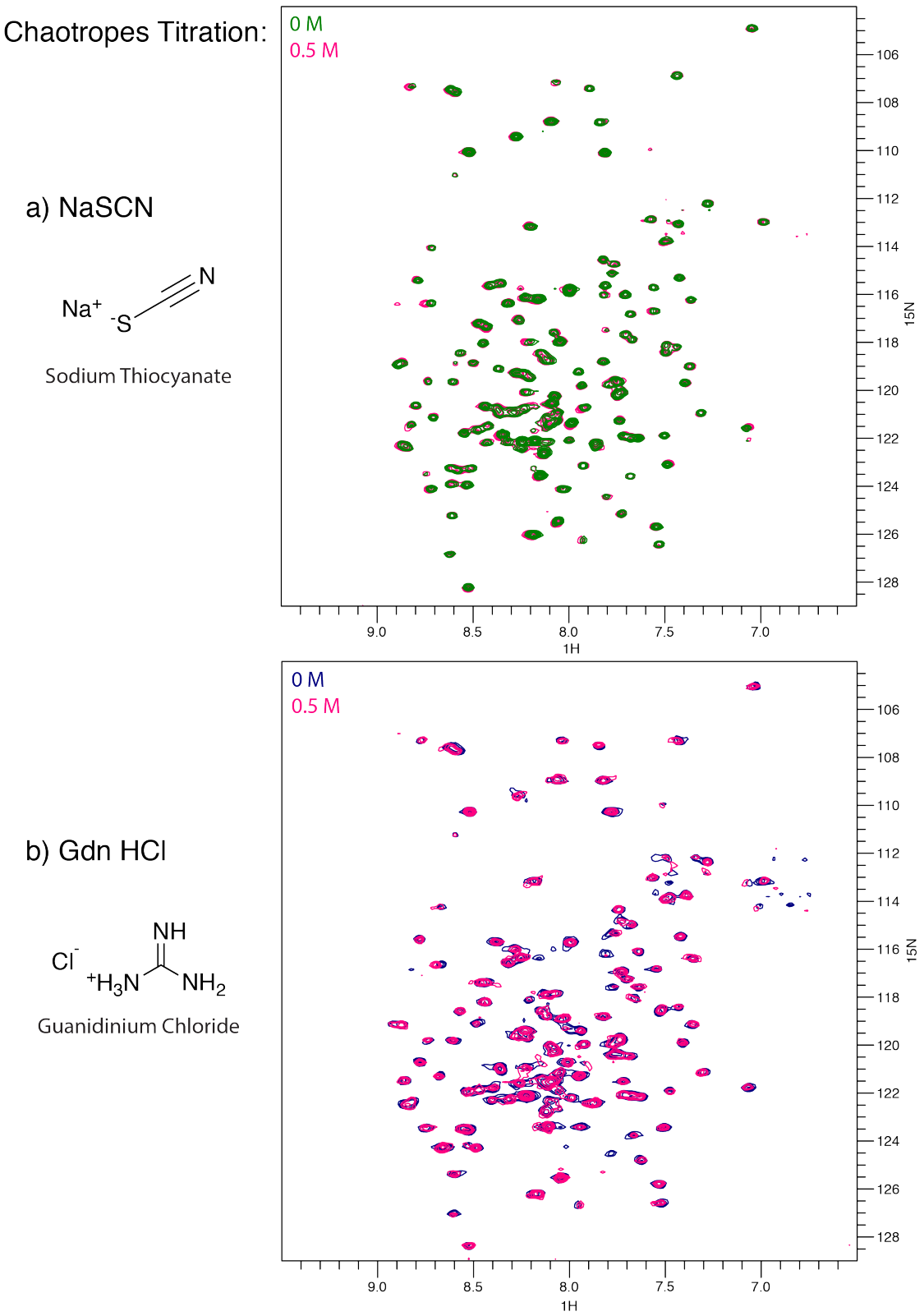


Figure 5.37: Sodium Thiocyanate (NaSCN) and Guanidinium Chloride (GdnHCl) titration of M1NT. a) Overlay of M1NT at 40 μM before (green) and after NaSCN titration of 0.5M (pink). b) Overlay of M1NT at 40 μM before (blue) and after GdnHCl titration of 0.5M (pink). Spectra collected at 25°C, pH 7 on 750 MHz spectrometers.

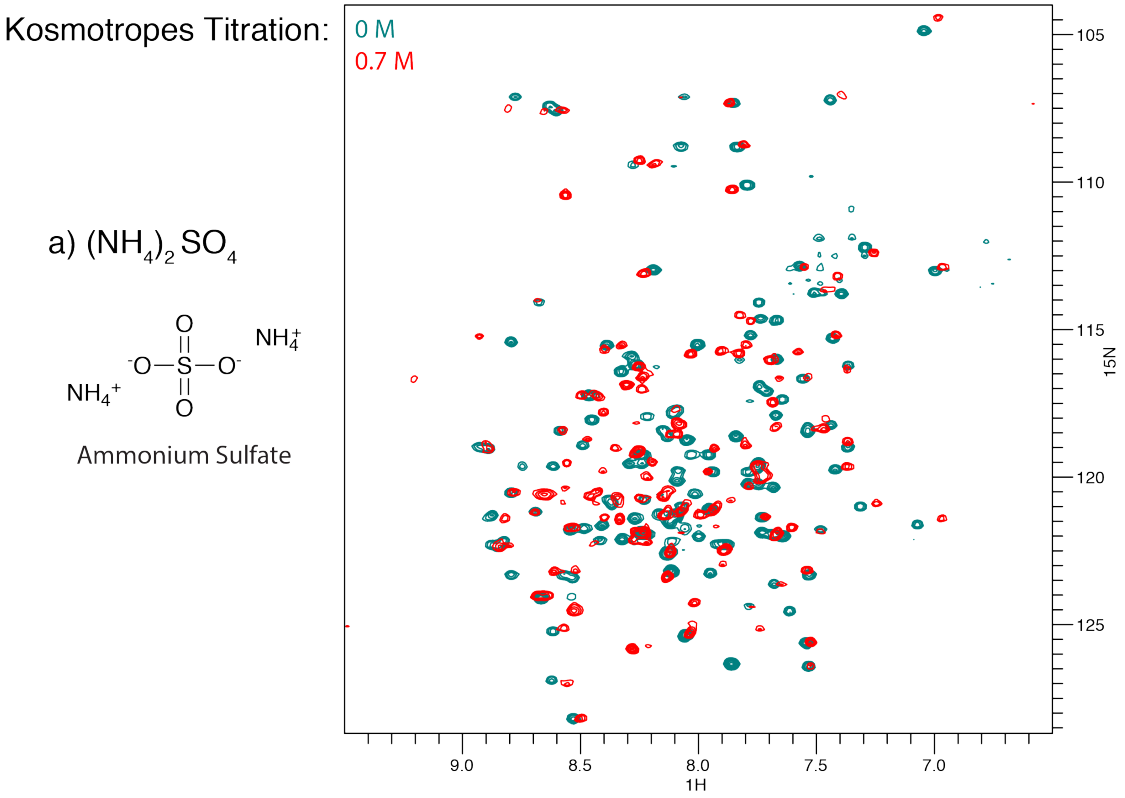


Figure 5.38: Kosmotropes - Ammonium Sulfate ($(NH_4)_2SO_4$) titration of M1NT. a) Overlay of M1NT at 40 μM before (green) and after NaSCN titration of 0.7 M (pink). Spectra collected at 25°C, pH 7 on 750 MHz spectrometers.

ions forming hydration shell(s) to accommodate the newly introduced large electric field around the ions. When discussing the effects of ions on the structure of water, often the effects beyond these hydration shells are generally focused on (Marcus, 2009). The terms “structure making” and “structure breaking", or the more or less equivalent terms - kosmotropes and chaotropes, respectively, have been used extensively to describe the effect of different ions (Bernal and Fowler, 1933; Hribar et al., 2002) on the surrounding water network in the biophysical context (Collins and Washabaugh, 1985; MacDougall, 1937). However, there is no clear definition of what these terms - kosmotropes and chaotropes, actually mean. Recently, these concepts have been challenged by several (Lyubartsev et al., 2001; Omta et al., 2003a, 2003b) when it comes to concentrated solutions, although the validity for diluted solutions still persists (Mancinelli et al., 2007; Schwenk et al., 2004).

Structural change of water ordering has immense significance as it has been

used to interpret a wide range of phenomena involving ions (Plumridge and Waigh, 2002; Zangi and Berne, 2006; Zou et al., 2002). In general, ionic behaviours behave similarly to the Hofmeister series (Kang et al., 2020). The Hofmeister series was discovered in 1888 and is originally based on various ions' abilities to precipitate hen egg-white proteins. It was shown that the series has the same effect on biological macromolecules and is probably the most well-known phenomena ascribed to structure breaking and structure making (Collins, 1997; Collins and Washabaugh, 1985; Plumridge and Waigh, 2002).

The Hofmeister ion series can be ranked, and the series has a negative correlation for anions and cations with the hydration levels (Figure 5.35). Anions are more strongly hydrated than that of cations of the same ionic radius - the Hw (water hydrogen) can approach 0.8Å closer than Ow (water oxygen) can approach cations (Plumridge and Waigh, 2002). Hence, small ions are much more hydrated than larger ions. Large monovalent ions with low charge density have a weaker interaction with water than water-to-water interactions and thus, have minimal effect on the hydrogen-bonding beyond its hydration shell. This is in contrast with small or multivalent ions with high charge density which interferes with the hydrogen bonding between water molecules, creating local order (Kang, 2020).

On the left-hand side of the Cl- in the anions series are anions well-hydrated and deemed to be kosmotropes, and on the right-hand side of Cl- in the anions series are less hydrated and classified to be chaotropes (Figure 5.35). For the cations, the left side of the series tends to stabilise the protein and the right side where the divalent cations reside increases the solubility of protein, promoting unfolding (Kang, 2020). Kosmotropes are effective at stabilising native protein folds while chaotropes are more effective in precipitating the protein (Kang, 2020; Marcus, 2009). The approximate ranking of kosmotropes and chaotropes for both anions and cations are shown in figure 5.36.

To understand whether conformational changes that have been observed by sterol-containing surfactants are due to the changes in water structure, a few chaotropes and kosmotropes have been tested. Like detergents, Sodium Isothiocynate (NaSCN)

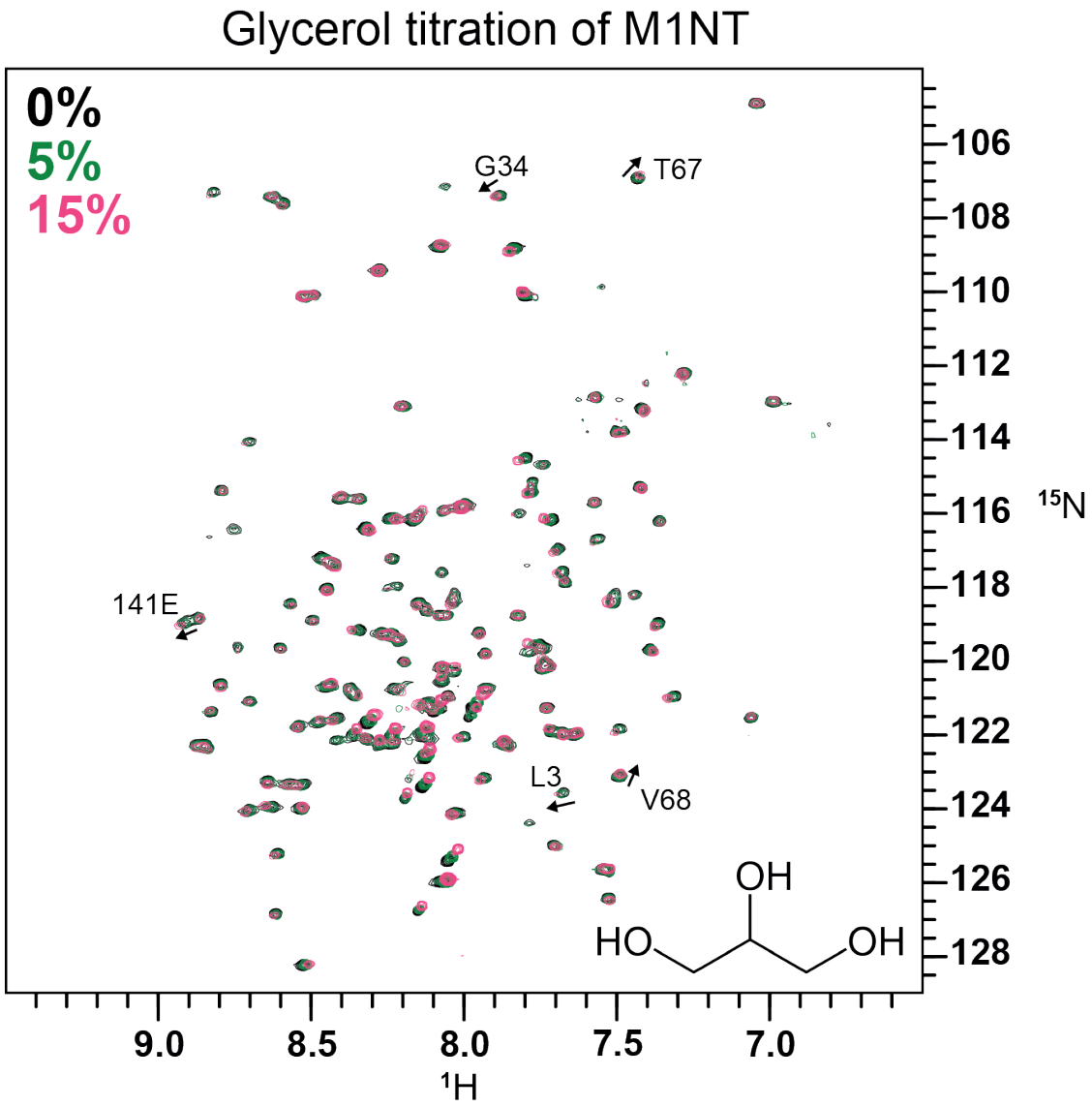


Figure 5.39: Glycerol titration of M1NT - Spectral overlay of M1NT alone (black) and in the presence 5% (green) or 15% (magenta) glycerol. All spectra were recorded at 25°C and pH 5.5. Deuterium lock artefacts with glycerol were corrected for by aligning spectra on the crosspeak for residue T48.

and guanidine chloride (Gdn HCl) are chaotropes (Cray et al., 2013), and when tested on M1NT, pH7.0 at 0.5M, NaSCN & Gdn HCl both showed little to no effects on the NMR spectrum (figure 5.37). This is in contrast with the kosmotrope that was tested - Ammonium Sulfate ($(NH_4)_2SO_4$), that had large effects even from 0.1M. Even though CSPs can be seen in the spectra, these CSPs are for a large number of the residues in the spectra, which indicates a low level of specificity (Figure 5.38). These CSPs are not strongly correlated with those seen with CHAPS

(Figure 5.40) and isopropanol (Figure 5.13).

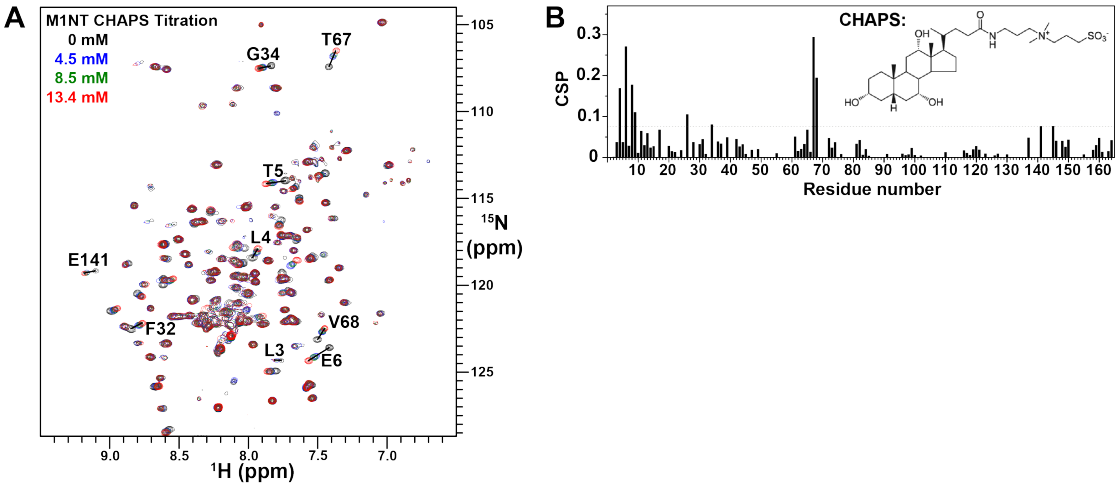


Figure 5.40: A) spectral overlay of 15N-labeled M1NT at pH 7.0 with increasing amounts of CHAPS indicating a saturable effect close to the critical micelle concentration of the detergent. B) per-residue chemical shift perturbations of M1NT at 13.4 mM CHAPS.

Interestingly, two structures with stong intermolecular stacking were crystallised in the presence of compounds that reduce water activity - 5% isopropanol (IPA) (1EA3) and 10% glycerol (5WCO). IPA and glycerol can strongly affect water activity (Zhu and Grant, 1996). The addition of IPA to M1NT has a similar effect to that of CHAPS and the CSPs occurred in the same direction (Figure 5.13). Glycerol also reduces water activity and was used in the crystallisation conditions for the ISA virus M1 protein (Zhang et al., 2017). The addition of 5% or 10% glycerol also showed similar shifts to CHAPS and IPA but at a lesser magnitude (Figure 5.39) consistent with its smaller effects on water activity (Marcolli and Peter, 2005).

Taken together, these results suggest that the effect of CHAPS on the hydrophobic patch might be attributed to water activity. The effects appear not enough to be caused indirectly by changes in water structure alone, since neither chaotropic nor kosmotropic additives specifically perturbed residues in the hydrophobic cavity. Thus it is likely that the changes observed are caused by a reduction in water activity at the hydrophobic patch upon CHAPS addition.

5.2.7 PHE

M1 protein is involved in multiple roles in the replication of the virus, and this includes the uncoating of virions, exporting vRNP complex in nuclear export, and the assembly and budding of new virions (Rossman and Lamb, 2011). M1 protein self-associates underneath the lipid bilayer of the virion, forming a multimer that supports the virion structure and is critical to virion integrity (Xie et al., 2013). The oligomerisation and the breakdown of M1 complexes are crucial for effective viral replication (Fontana and Steven, 2013). M1 acts as a tether, connecting the membrane proteins such as HA and NA with the core of the virion, the vRNP. The functional versatility of M1 is owed to its ability to form multiple conformational states in response to different localised conditions or environments such as pH and proximity to lipid surroundings (Arzt et al., 2001; Batishchev et al., 2016; Harris et al., 2001; Hilsch et al., 2014; Safo et al., 2014; Shtykova et al., 2013; Zhang et al., 2012).

Mosier et al. hypothesised that the M1-M1 self-association can be destabilised by disrupting this formation of the large M1-M1 interface by using small molecules as inhibitors. This was deemed as the next generation of antiviral agents (Mosier et al., 2016). The destabilising effect of these small molecules could interfere with the formation of the M1 layer underneath the viral membrane rendering unstable virions with impaired ability to replicate. Furthermore, M1 exhibits a critical feature, which is that it has the most highly conserved sequence of any influenza protein. This makes M1 an intriguing candidate for drug target capabilities as there are minimal selection pressures (Furuse et al., 2009; Heiny et al., 2007; Ito et al., 1991; Mosier et al., 2016).

Mosier et al., conducted a virtual drug library screen (>70,000 commercially available small molecules) against a structure of the isolated M1 N-terminal domain (PDF:4PUS & 1EA3) and the top 10 'hit' compounds were identified. Through *in vitro* replication assay, each hit compound was tested to identify compounds with dose-dependent inhibition and no to little cytotoxicity. The molecule PHE (MW 400) was identified to fit these characteristics and using surface plasmon resonance

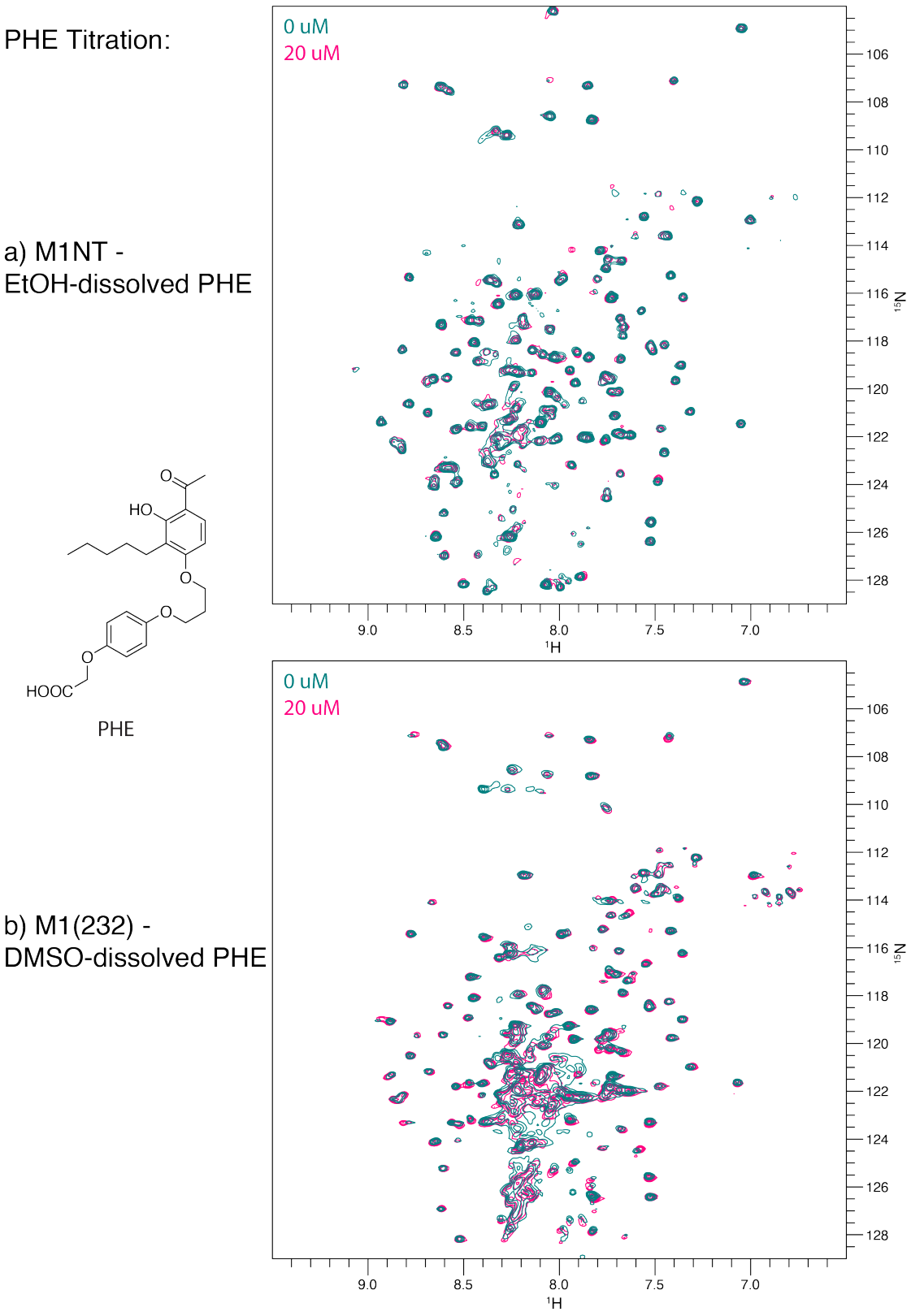


Figure 5.41: PHE titration. a) Overlay of M1NT at 40 μM before (teal) and after ethanol-dissolved PHE titration (pink). b) Overlay of M1(232) at 40 μM before (teal) and after DMSO-dissolved PHE titration (pink). Spectra collected at 25°C, pH 7 on 750 MHz spectrometers.

(SPR), Mosier et al. found that PHE binds to M1NT with an affinity of 870 ± 150 nM at pH 7.4. Furthermore, they also found that PHE was able to disrupt M1 self-association with high potency (Mosier et al. 2016).

We then tested PHE on M1NT with the aim to examine whether PHE binds to M1 and subsequently can disrupt oligomerisation with CHS at pH 7.0 - this will shed light the residues involved in Oligomerisation. PHE was added to samples of purified and isotopically labelled samples of M1NT to determine if PHE binds to M1NT and whether a binding site could be identified. There was no significant effect on the M1NT chemical shifts when ethanol-dissolved PHE was added to M1NT (Figure 5.41a). A different organic solvent was also used, DMSO, to dissolve PHE in order to match the solvent used by Mosier et al., to test whether PHE binds to a longer construct, M1(232) and it was found that the small CSPs were consistent with nonspecific effects i.e. random peaks with unsignificant amount of shift that can be considered to be true (due to PHE) and not due to change in solvent (Figure 4.40b). From these results, we could not confirm that PHE binds to M1NT. Despite the virtual screen being carried out against the structure of the isolated N-terminal domain, it remains possible that PHE binding only occurs in the context of the full-length protein. In addition, the M1NT construct used for NMR studies differs in sequence at 5 different amino acid residues from that used by Mosier et al. and 0.05% of Tween 80 (other buffer conditions are identical) was also present in the biotinylated M1 sample used in the SPR study. These differences might explain the lack of binding of PHE to M1NT.

You might work but I work harder
You might fight but I fight smarter
Might not be on the top of the world but hey
I'm on my way

— Charlie Brown

6
M1 interaction with proteins

Contents

6.1 Nuclear Export Protein (NEP) 112
6.1.1 Introduction 112
6.1.2 Expression and Purification of NEP 115
6.1.3 M1 Interaction with NEP 119
6.2 Tumour Susceptibility Gene 101 (TSG101) 121
6.2.1 Introduction 121
6.2.2 Expression and Purification of TSG101-UEV Domain . 126
6.2.3 M1 Interaction with TSG101-UEV Domain 126

6.1 Nuclear Export Protein (NEP)

6.1.1 Introduction

The genome of the influenza A virus is made up of 8 viral ribonucleoproteins (vRNP) complexes. Each vRNP is made of a segment of the viral RNA packaged around the nucleoprotein and viral RNA polymerase. vRNPs are associated with the nuclear export proteins (NEP) and are enclosed by the matrix protein M1. M1 is the most abundant protein in the virus and forms the matrix layer beneath the lipid bilayer of the viral capsid stabilising the three-dimensional structure (Fujiyoshi et al., 1994; Ruigrok et al., 1989). Additionally, it is also thought to form contacts

between several other viral constituents such as CRM1.

Apart from forming a significant role in structurally stabilising the architecture of the virus particles, M1 has other essential roles that include regulating the transport of vRNPs and recruiting other viral proteins in the budding process during virus assembly. In influenza A virus-infected cells, replication and transcription of the viral genome occur in the nucleus. To be packaged into viral particles at the plasma membrane, encapsulated viral genomes must be exported from the nucleus. The transport of the packaged vRNP is mediated by a cellular host protein called Crm1/exportin1 (Elton et al., 2001; Ma et al., 2001; Watanabe et al., 2001), which is part of the nuclear transport receptors family, importin β (Weis, 2003). With the help of GTPase Ran, Crm1 is able to bind to its package (vRNP), which contains a leucine-rich nuclear export signal (NES) (Fornerod et al., 1997; Fukuda et al., 1997; Stade et al., 1997). Crm1 is known to mediate the export of RNA of other enveloped virions such as HIV-1 and influenza via their adaptor proteins, Rev and NEP (nuclear export protein), respectively (Meyer and Malim, 1994; Fischer et al., 1995; Askjaer et al., 1998).

NEP is a 14.5 kDa protein in which only a crystal structure of the C-terminus is available (Akarsu et al., 2003). The NEP C-terminal domain is a dimer forming an antiparallel 4-helix bundle. NEP contains a leucine-rich NES which is analogous to the adaptor protein of HIV-1 Rev by acting as a bridge connecting the Crm1 to allow export out of the nucleus (Martin and Helenius, 1991; Whittaker et al., 1996; Bui et al., 2000). vRNPs are bound by M1 with high affinity (Baudin et al., 2001), and studies have shown that NEP associates with M1 mediating the export of vRNPs from the nucleus (Yasuda et al., 1993; Ward et al., 1995, Akarsu et al., 2003). Recombinant virus with no NEP has been shown to have defective vRNP export (Neumann et al., 2000) and antibodies against NEP prevent the export of vRNP (O'Neill et al., 1998). Furthermore, there is also evidence through biochemical assays that NEP interacts with Crm1 (Elton et al., 2001; Paragas et al., 2001). Pull-down experiments have shown that the C-terminal domain of NEP interacts with the N-terminal domain of M1 (Akarsu et al., 2003).

It has also been reported that M1 binds to vRNP at the C-terminal domain (Baudin et al., 2001), and binds to NEP at the N-terminal domain (Akarsu et al., 2003). Studies have shown that NEP binds to M1 lacking the residues 1 to 88 (Yasuda et al., 1993), and Akarsu and colleagues have shown through pull-down experiments that NEP binds to M1 NLS basic motifs of residues 101 to 105, RKLKR. M1 NLS was also reported to bind to negatively charged liposomes by Baudin and colleagues (Baudin et al., 2001). Replacing the basic residues through alanine mutagenesis studies have shown that NEP binding to M1 was severely reduced (Akarsu et al., 2003). Furthermore, another study showed that a mutation at the NLS basic motif of M1 at either lysines residues caused M1 not to be transported into the nucleus of infected host cells, and the virus became non-viable after cell transfection (Liu and Ye, 2002). However, mutations at either arginine residues have no significant effect on the export of vRNP after cell transfection (Liu and Ye, 2002). The NLS in M1 contributes to binding to NEP and masking this NLS will ensure that the newly synthesised vRNPs are transported out of the nucleus to the cytoplasm for budding (Arkasu et al., 2003). This is to prevent the vRNPs from being transported back to the nucleus (Arkasu et al., 2003).

Pull-down and mutagenesis experiments have shown that the Trp78 residue of NEP which resides at the helical hairpin was responsible for the binding to M1 NLS. Trp78 is exposed to the solvent and is surrounded by the side chains of glutamate residues (Arkasu et al., 2003). Furthermore, it was found that these glutamate residues were not conserved in influenza B virus and are replaced by basic residues and may suggest that the importance of the positive charges. Replacing the glutamate residues surrounding Trp78 with alanine residues retain significant binding for M1, findings from pull-down experiments suggest. However, replacing Trp78 itself led to severe disruption in binding (Arkasu et al., 2003). These interactions between Trp78 and the surrounding glutamate residues might be involved in the electrostatic interactions with the basic NLS M1 motif to allow the export of M1 out of the nuclues. Trp residue represents a critical residue in recognition of NLS motif such as in the interaction of nuclear importin α with

NLS motif where the Trp residue interacts with Lys residues from the NLS (Conti et al., 1998; Conti and Kuriyan, 2000).

A daisy-chain model was proposed by Arkasu and colleagues to describe the interaction between NEP and M1 as shown in figure 6.1. This model involves a complex of Crm1, NEP, M1 and vRNP interacting with one another without any interference between individual components. vRNP is proposed to bind to the M1 C-terminus domain, the M1 N-terminal domain via its NLS motif to the C-terminal domain of NEP, and the N-terminal of NEP, in turn, binds to the Crm1. This mimics a similar interaction with the use of two adaptor proteins in the U snRNA export system which is also mediated by Crm1, where Crm1 recognises the adaptor protein PHAX, and PHAX subsequently binds to another adaptor protein, CBC, which then binds to U snRNA (Ohno et al., 2000).

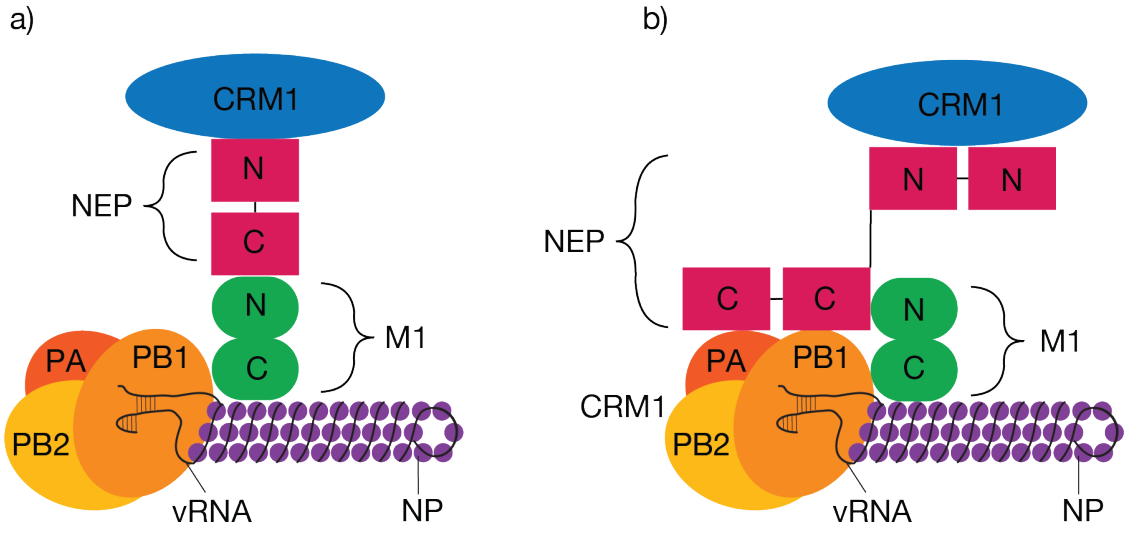


Figure 6.1: The daisy chain model of the nuclear export complex. a) Initial daisy chain model involving Crm1, NEP, M1 and vRNP. In this model, the Crm1 is said to mediate the export of the vRNP via NEP. M1 interacts with the vRNP, and NEP acts as a bridge by binding the Crm1 at the N-terminus and binding to N-terminal domain of M1 at the C-terminus via its nuclear localisation signal (NLS). M1 then binds to the vRNP at its C-terminus. b) Refined model of the daisy chain model. In this model, NEP also interacts with the viral polymerase as an additional binding site and hence, stabilises the M1 and vRNP interaction. Also, the NEP N-terminus is said to be flexible and binds to the Crm1.

6.1.2 Expression and Purification of NEP

We then wanted to test whether NEP binds to M1NT to define its interaction partners. His-tagged NEP construct was kindly provided by Ducan Paterson from Fodor's lab of Dunn School of Pathology, University of Oxford. His-tagged NEP was over-expressed using the pET expression system in BL21 (DE3) *E. coli* cells. High-level expression of recombinant proteins in *E. coli* often induce the aggregation of the expressed protein into inclusion bodies (IB) (Chrunyk et al., 1993; Freedman and Wetzel, 1992; Singh et al., 2015; Williams et al., 1982). This was the case for the expression of NEP, which was found predominantly in the IB instead of the soluble fraction. The use of high temperature, high inducer concentration, as well as the use of strong promoter systems, can give a high expression of the desired protein. However, this can overwhelm the protein quality control system resulting in the partially folded and misfolded protein to aggregate into IB (Carrió et al. 2000; Singh et al. 2015).

It has been proven that IB can have protein molecules in native-like conformation and methods have been developed to solubilise these IBs while retaining the secondary structures of proteins, resulting in the high recovery of bioactive protein. The IBs were isolated by lysing the *E. coli* cell pellet with a french press. The IB was then solubilised by resuspending in 6 M guanidine hydrochloride buffer. This was then passed through the Ni affinity column where it bound and refolded slowly by decreasing the concentration of 6 M guanidine hydrochloride to 0 M on the column (Figure 6.2). The calculated theoretical molecular weight of polyhis-tagged NEP is calculated to be 16.1 kDa and the eluted fractions are consistent with this, as shown with SDS-PAGE in Figure 6.2.

The eluted fractions from Ni purification and refolding was then pooled and run through a size exclusion column. This results in a single narrow peak and the eluted fractions from this peak were analysed by SDS-PAGE (numbered 1,2 and 3 in the SEC chromatogram (figure 6.3). The bands shown in the SDS-PAGE are at the expected molecular weight of polyhis-tagged NEP. Soluble NEP is known to exist as a monomer, consistent with the single peak in the SEC chromatogram.

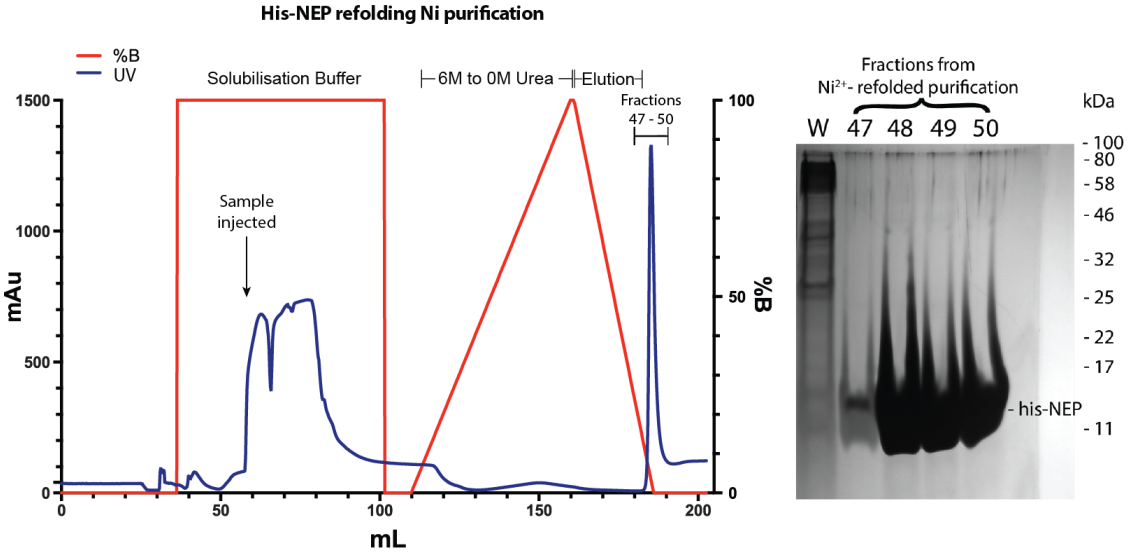


Figure 6.2: Nickel column purification and refolding of NEP and the analysis of the fractions on a SDS-PAGE gel (NEB). Lanes 47, 48, 49 and 50 represent the fractions from the peak at the elution step of the Ni purification chromatogram. NEP was expressed using BL21 (DE) *E.coli* cells and induced at 37°C for 4 hours. Detailed method is described in chapter 2.

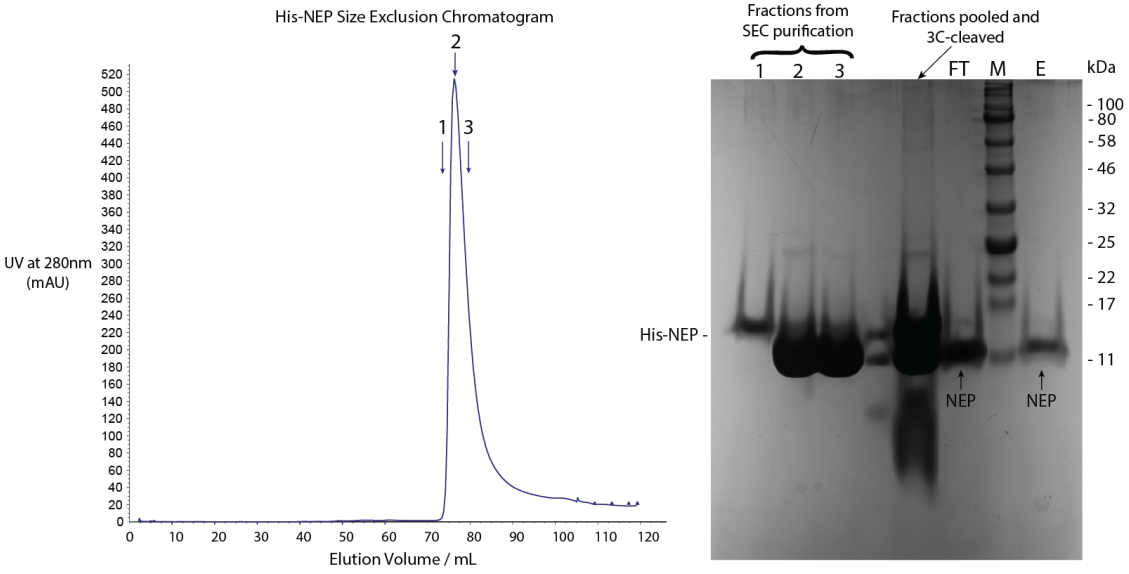


Figure 6.3: Size exclusion of His-tagged NEP (figure on the left). SDS-PAGE gel of fractions from size exclusion chromatography. In addition, samples from Ni-column purification after HRV3C protease cleavage of His-tagged NEP were also analysed (figure on the right). Lanes numbered as 1,2 and 3 represent the fractions labelled on the SEC chromatogram. FT - flow through, M - marker, E - elution.

The fractions from the SEC were then pooled and the polyhis-tag was cleaved by HRV3C protease. This was then run back onto the Ni-column to remove the polyhis-tag and analysed again on a SDS-PAGE (figure 6.3). The protein is shown to be entirely cleaved as the bands are at the expected molecular weight of cleaved NEP (14.5 kDa). The presence of a band in the elution lane (E in figure 6.3) is likely to be caused by insufficient washing of the column to elute the cleaved protein before adding elution buffer.

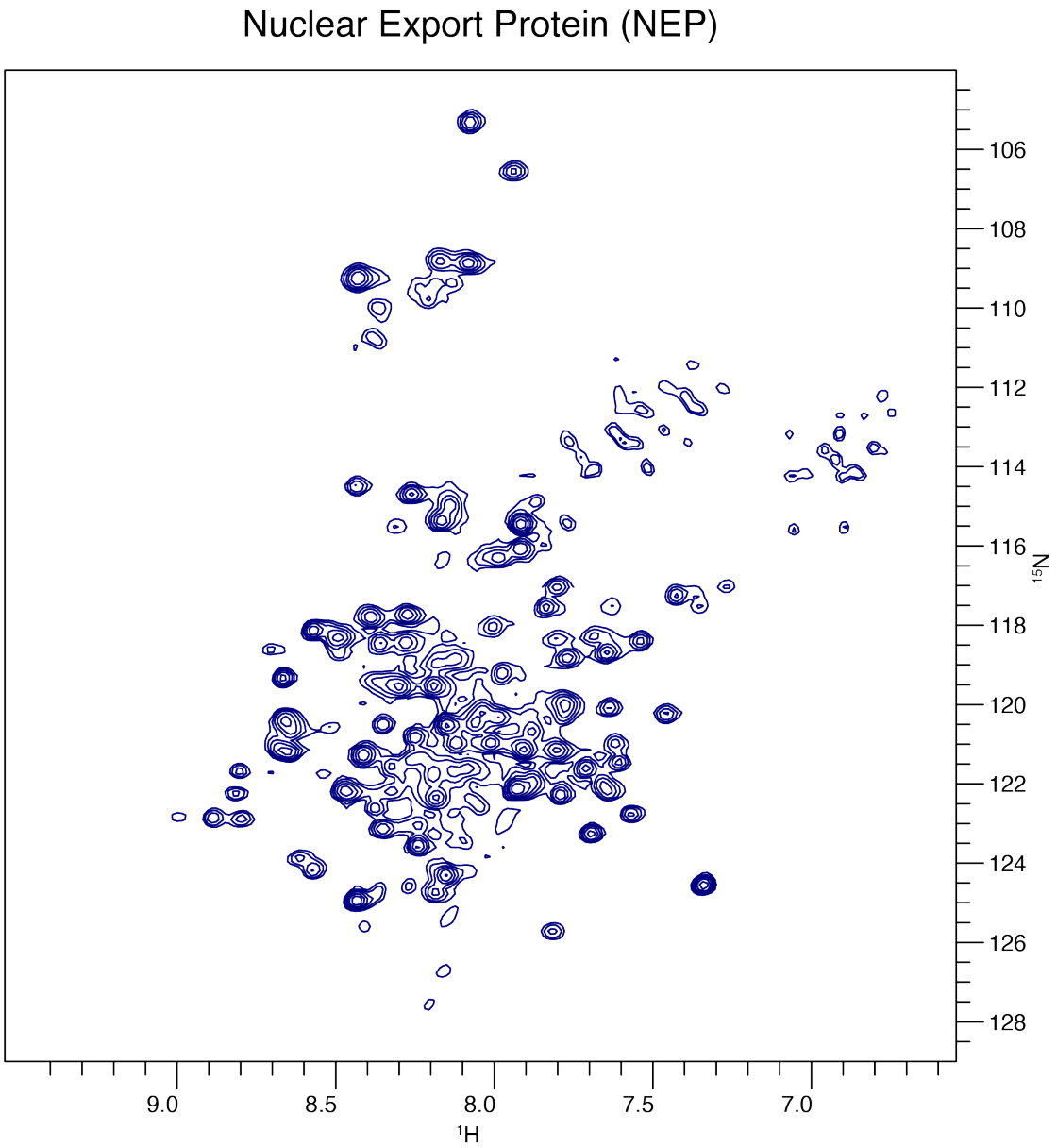


Figure 6.4: HSQC spectrum of 50 μM NEP in 20 mM Tris, 100 mM NaCl, pH 7. Spectrum was collected with 650MHz spectrometer at 30°C.

Plasmids were transformed (via heat shock) into *E. coli* strain BL21 (DE3) (NEB) and expressed as PolyHis-M1NT as His_9-tagged proteins. Overnight culture of transformed *E. coli* (200mL) was used to inoculate 2 L of high salt Luria Broth (LB) media. Cells were grown at 37°C containing ampicillin, 100 μg/mL to an OD_{600} of 0.5-0.8. Cells were cold shocked for 30 minutes on ice and left at 15°C for 10-15 minutes. The cultures were spun down at 3000xG for 15 minutes.

Cell pellets were then resuspended in 1 L of 2x M9 minimal media containing either 1 g/L ^{15}N-labelled ammonium chloride, or 1 g/L ^{15}N-labelled ammonium sulfate and 3 g/L ^{13}C-labelled glucose, and expression was induced at 15°C with 1.5 mM IPTG (isopropyl-1-thio-β-Δ-galactopyanoside).

To collect the NMR spectrum of NEP, it was labelled with isotopes. Plasmids pET-M14-NEP were transformed into *E. coli* strain BL21 (DE3) and expressed his-NEP as his_4-tagged proteins. Overnight culture of transformed *E. coli* (50mL) was used to inoculate 2 litres of Luria Broth (LB) media. Cells were then grown at 37řC containing kanamycin, 100 μg/mL to an OD_{600} of 0.8. The culture was then condensed to 500mL of M9 minimal media containing 1 g/L ^{15}N-labelled ammonium chloride , and expression was induced at 15°C with 1.5 mM IPTG (isopropyl-1-thio-β-Δ-galactopyanoside). Cells were then harvested by centrifugation at 4000xg after induction with 1 mM IPTG (isopropyl-1-thio-β-Δ-galactopyanoside) at 37°C for 4 hours. The samples are then purified as described in chapter 2.

A HSQC experiment of ^{15}N uniformly labelled NEP was then performed (Figure 6.4). The spectrum of NEP showed characteristics of a well-folded protein including homogenous peak linewidths and dispersion typical of an alpha-helical protein.

6.1.3 M1 Interaction with NEP

M1NT-NEP titration experiments were performed to examine if there is any interaction between the two proteins. Firstly, M1NT was tested on isotopically-labelled ^{15}N NEP as shown in figure 6.5a. There were no significant changes when M1NT was added to NEP. NEP was then tested on M1NT as shown in figure 6.5b. There were also no significant perturbations that can be seen with the M1NT peaks.

These experiments were then repeated at different concentrations with M1 at 50μM and NEP at 50μM, 100μM and 200μM. However, none had shown any significant changes which confirms no binding of M1NT to NEP or NEP to M1NT.

Further experiments were done by priming M1NT with CHAPS and to examine whether CHAPS induces an M1NT conformation necessary for binding to NEP (see Chapter 5). The addition of M1NT with CHAPS was first tested on NEP (Figure 6.6a). As can be seen in figure 6.6a, there were no significant changes that can be seen in the overlay spectra suggesting no interaction. A longer construct of M1, M1(224) that includes the C-terminus was also tested to examine if a C-terminal extension facilitates binding of M1 to NEP and shows no significant CSPs. There were also no significant changes in the spectra of M1NT primed with CHAPS after the addition of NEP as shown in figure 6.6*b*. These experiments confirm that there was no binding between NEP and M1 primed with CHAPS.

Lastly, there is a possibility that NEP could be binding to M1 oligomers, and hence this was also tested. NEP was used to examine changes to CSPs upon addition of M1 oligomers. CHAPS were first added to NEP, as shown in figure 6.7*a*. There was a slight change in the spectra of NEP with CHAPS. CHS was then added, followed by M1NT, as shown in figure 6.7*b*. When CHS was added, there were no changes in the spectra of NEP. However, there are slight changes in the spectra when M1NT was added afterwards as shown in figure 6.7. But this result is not convincing enough to show binding between NEP and M1.

As was discussed above, the daisy chain model described the interaction of M1, NEP, vRNP and Crm1 in which each interaction in the chain are independent of one another. However, there is also a possibility that the interaction of M1 and NEP requires all individual components to be present. A study proposed by Brunotte et al., suggest a refined daisy chain model in which NEP interacts with the vRNP via its polymerase as well as M1 through its C-terminus and the N-terminus binds to Crm1 (Figure 6.1). They showed that M1 alone is not able to interact with vRNP through an *in vitro* reconstitution assay of the vRNP complex. The interaction between M1 and vRNP requires NEP C-terminal domain (Brunotte et al.). The

C-terminal domain of NEP contains both M1 and the viral polymerase binding sites. This refined model suggests that NEP acts more than just an adaptor protein that bridges M1 and CRM1, and instead, also acts by supporting the M1/NEP interaction by binding to the viral polymerases, PB1 and PB2. To this date, there are no other experiments with recombinantly produced M1 and NEP that have shown these two proteins interact. Is it also a possiblility that the interaction of M1 and NEP may require vRNP or at least the viral polymerases to be present.

6.2 Tumour Susceptibility Gene 101 (TSG101)

6.2.1 Introduction

For complete pathogenesis of the influenza virus, virions with encapsulated viral genomes must be bud off from the plasma membrane of infected host cells. For budding to occur, this complex process requires orchestrated help from both the virus and host cell proteins. Particularly, the matrix protein of RNA viruses or M1 protein of influenza A virus is involved in regulating the assembly of virions and budding from host cells. Multiple studies have shown that the matrix protein of different viruses is capable of forming virus-like particles (VLPs): human-immunodeficiency virus (HIV) through its p6 Gag (Accola et al., 2000; Martin-Serrano et al., 2001), and Ebola virus through its viral protein (VP), VP40 (Martin-Serrano et al., 2001).

It is known that tumour susceptibility gene (TSG101) protein in the host cell interacts with HIV and Ebola viruses proteins to assist the late step of virion pathogenesis, budding (Garrus et al., 2001; Martin-Serrano et al., 2001; Pornillos et al., 2002; VerPlank et al., 2001). TSG101 is involved in the ESCRT-1 transport and is a part of a complex subunit of the endosomal sorting complex. In addition, TSG101 is highly conserved amongst mammalian species (Li and Cohen, 1996). The TSG101 is usually exposed on the cell surface but not typically thought as a cell surface protein. It plays a part in the regulation of the transport of the endocytosed proteins from the multivesicular body (MVB) to the proteasome (Tu et al., 2010). However, during viral infection, TSG101 is hijacked and is used to facilitate budding of virion particles and has been shown to be ESCRT independent.

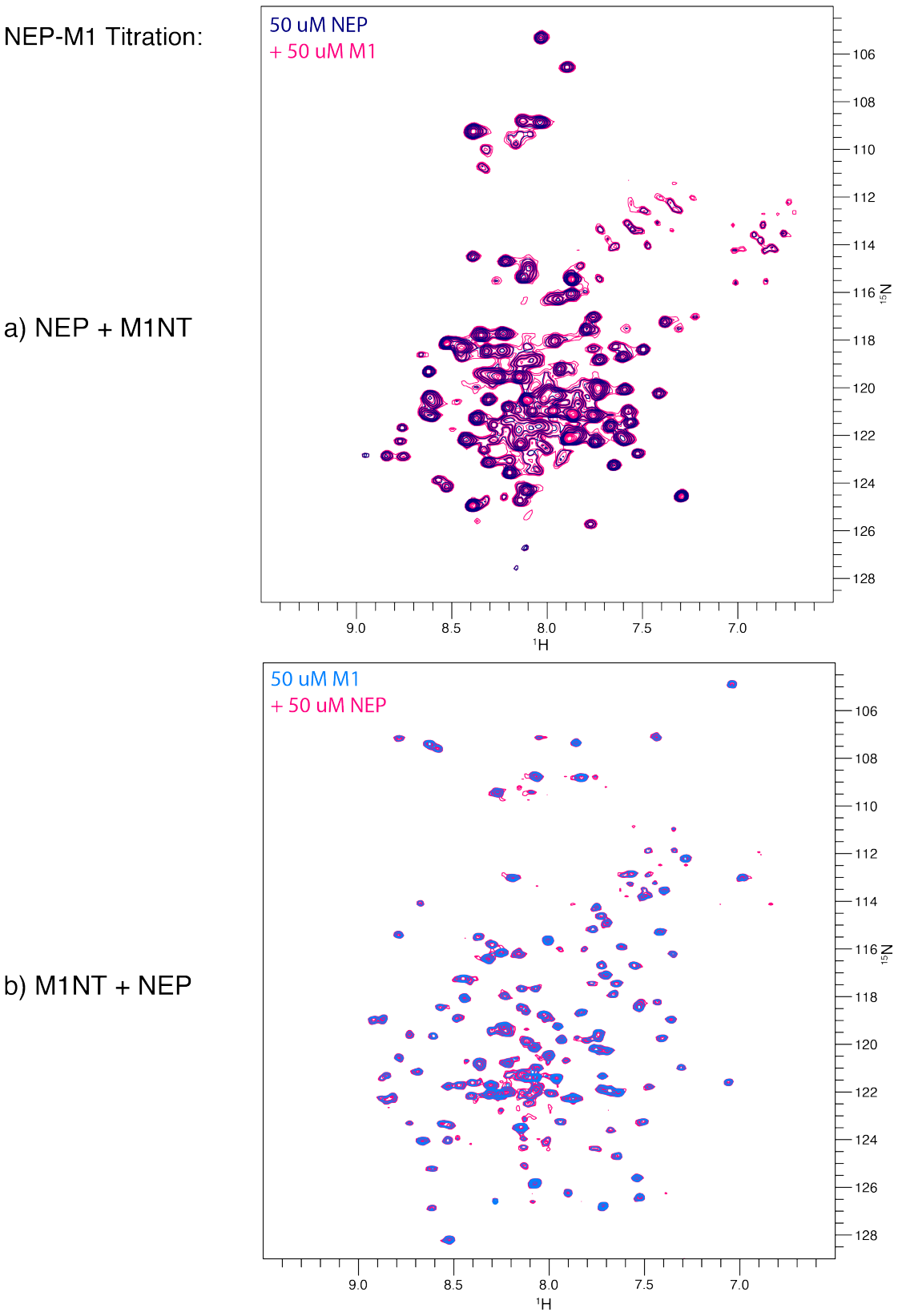


Figure 6.5: NEP - M1 titration. a) Overlay of NEP at 50 μM before (blue) and after addition of M1NT at 50 μM (pink). b) Overlay of M1NT at 50 μM before (light blue) and after addition of NEP at 50 μM (pink). Spectra collected at 25°C, pH 7 on 750 MHz spectrometers.

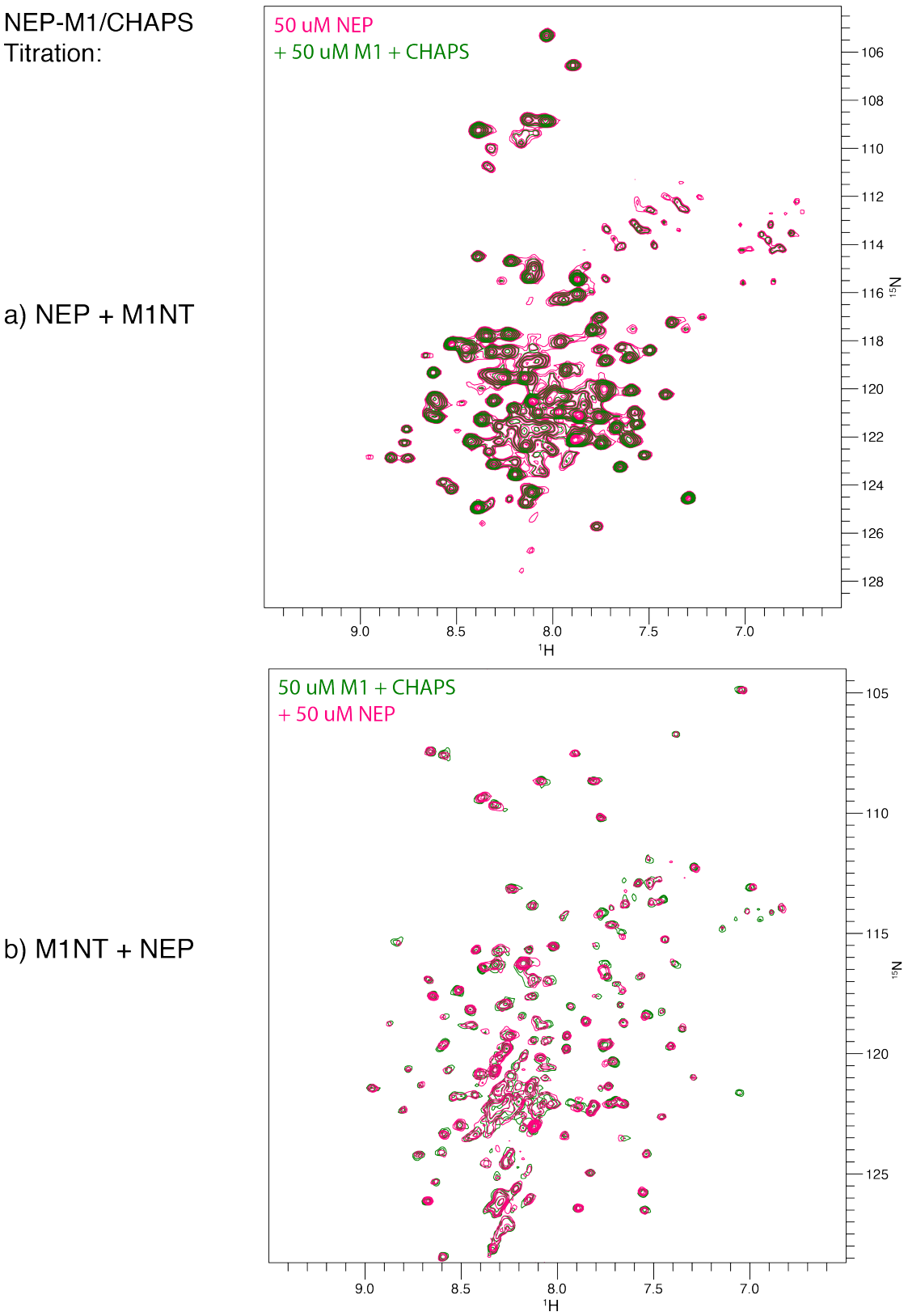


Figure 6.6: NEP - M1/CHAPS Titration. a) Overlay of NEP at 50 μM before (blue) and after addition of M1NT/CHAPS at 50 μM (pink). b) Overlay of M1NT with CHAPS at 50 μM before (light blue) and after addition of NEP at 50 μM (pink). Spectra collected at 25°C, pH 7 on 750 MHz spectrometers.

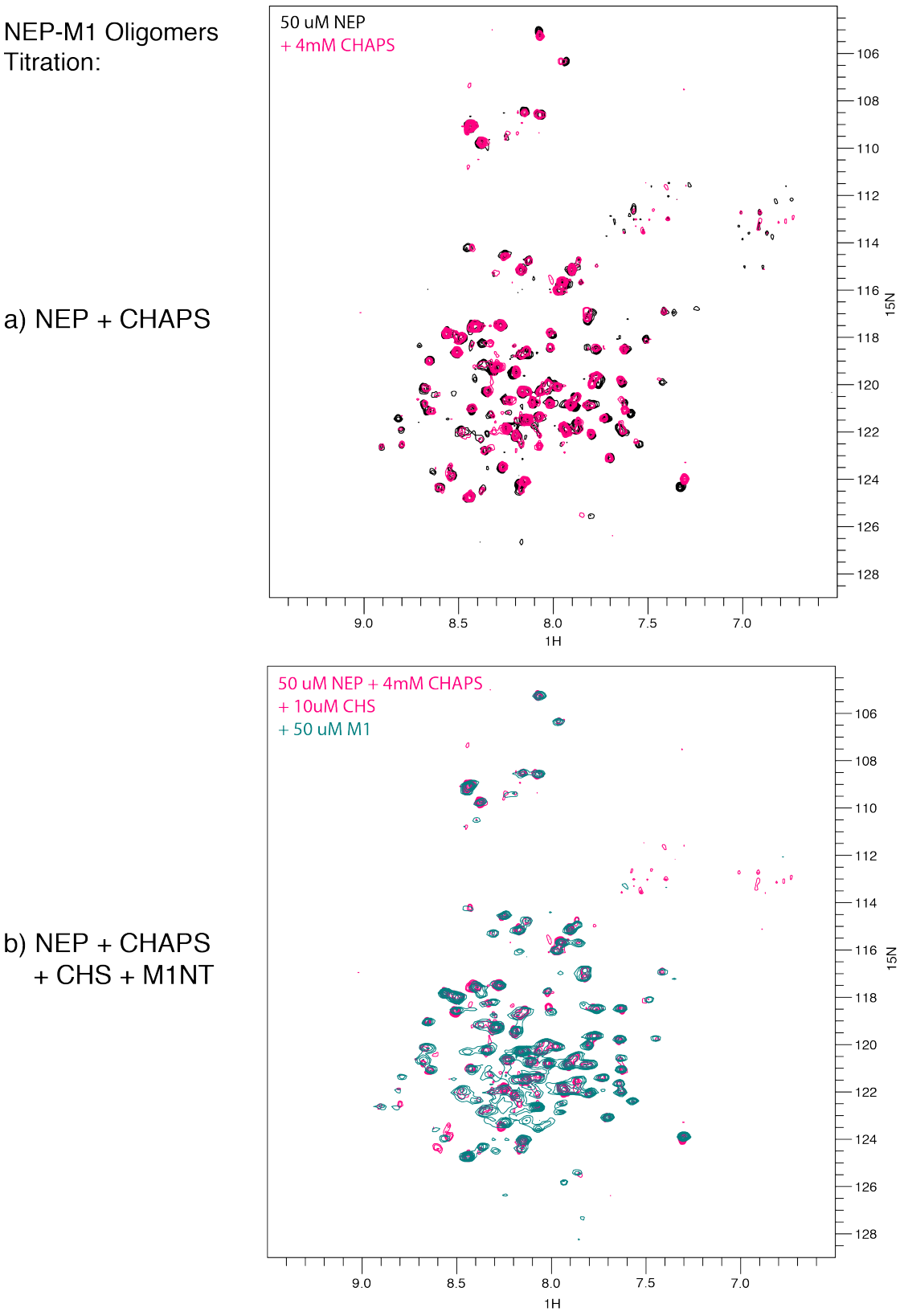


Figure 6.7: NEP - M1 Oligomers Titration. a) Overlay of NEP at 50 μM before (black) and after addition of CHAPS at 4 mM (pink). b) Overlay of NEP with CHAPS & CHAPS-solubilised CHS at 50 μM before (pink) and after addition of M1NT at 50 μM (Teal). Spectra collected at 25°C, pH 7 on 750 MHz spectrometers.

This was also seen in other enveloped viruses such as cells infected by HIV (Garrus et al., 2001; Martin-Serrano et al., 2001; Mazzé and Degrève, 2006; Perez and Nolan, 2001; VerPlank et al., 2001).

In HIV-infected cells, the p6 Gag protein of the HIV which contains the P(T/S)AP motif that interacts with the domain of TSG101, the ubiquitin E2 variant (UEV) domain (Garrus et al., 2001; Myers and Allen, 2002; VerPlank et al., 2001). The p6 Gag protein modulates the assembly of new HIV virions and budding and like M1, is the structural protein. The P(T/S)AP or YXXL are motif within the late domain, and this reflects their functioning role at the end of the pathogenesis viral cycle that involves budding of fresh progeny. They are known to be highly conserved in sequence to regulate interactions between viral and host proteins (Freed, 2002). In this case, the late domain resides at the C-terminal domain of the p6 Gag protein of HIV. Through this interaction, TSG101 is hijacked to facilitate budding of virions by recruiting other cellular components (Pornillous et al.,2003). A study shown by Bonavia and colleagues has demonstrated that similarly to HIV-infected cells, TSG101 may play a role in the budding of influenza virions, and interestingly, was exposed to the surface of the influenza-infected host cell (Bonavia et al., 2010). Furthermore, M1 was shown to interact with TSG101 via its UEV domain. It should be noted that TSG101 is not required for IAV budding and TSG101 involvement is outside of canonical ESCRT-mediated scission.

The UEV domain of TSG101 (TSG101-UEV) interacts with the late-domains of the HIV, p6 Gag and Ebola virus, VP40 (Freed, 2003; Martin-Serrano et al., 2001). The M1 protein also contains the canonical late-domain motif, YKRL (residues 100-103), and this could be the binding site for the TSG101-UEV domain (Bonavia et al., 2010). However, intriguingly, M1 also has another canonical late-domain motif, PTGP between helix 1 and helix 2 (residues 16-19), which could also be the binding site for the TSG101-UEV domain. TSG101 may play a role in for viral replication, and hence understanding the interaction between M1 and the TSG101-UEV domain may provide insights into understanding the late budding stage of the influenza pathogenesis.

TSG101 could be involved in the interaction of M1 with the membrane or other viral or host proteins in the budding stage of the virions. Thus, it is important to explore the possibility of the interaction of M1 with TSG101 through NMR experiments and if so, how does this affect M1 structure and interactions.

6.2.2 Expression and Purification of TSG101-UEV Domain

Attempts were made to purify the TSG101-UEV domain (residues 1-145) of TSG101 using a previously published protocol (Pornillos et al., 2002). Similar to that for M1 constructs, the plasmid, coding for TSG101-UEV domain, was transformed into BL21 *E. coli*. TSG101-UEV was expressed and purified by the salting-out method via ammonium sulfate precipitation. At different ammonium sulfate saturation points, the pelleted precipitate was resuspended in TSG-UEV resuspension buffer (50mM Tris, 50mM NaCl, 2mM -mercaptoethanol, pH 6.5, 25mM MOPS).

The TSG101-UEV domain precipitated from ammonium sulfate salting-out process starting from 28% to 75% saturation. The resuspended pellets of the precipitate, when subjected to ion-exchange chromatography, eluted the TSG101-UEV domain at 300-400mM of NaCl (Figure 6.8). However, as seen on the SDS-PAGE gel, the TSG101-UEV domain ran lower than 14 kDa, which contrasted its calculated molecular weight of 16.6kDa. However, the expected mass was confirmed by mass spectrometry.

6.2.3 M1 Interaction with TSG101-UEV Domain

Other enveloped viruses such as the p6 Gag of HIV and VP40 of the Ebola virus interacts with TSG101 UEV domain. Here, we used NMR spectroscopy to examine if the influenza M1 virus also interacts with TSG101. The longer M1 construct (M1(1-224)), which includes a portion of the C-terminus, was used to check its interaction with TSG101-UEV domain. TSG101-UEV domain was added into M1(224) sample, and results are shown in figure 5.9. There were no significant changes when TSG was added to M1. Experiements were done at different concentrations with M1(224)

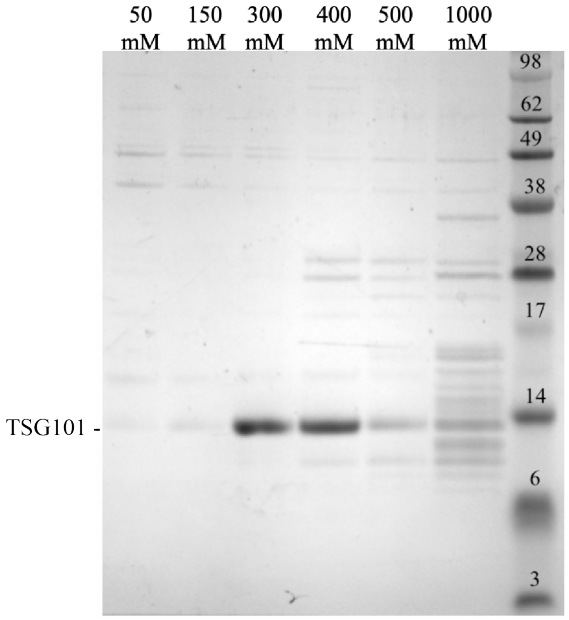


Figure 6.8: TSG101-UEV Salting out, ammonium sulfate purification - an SDS-PAGE gel of ion-exchange chromatography of the TSG101-UEV domain. NaCl concentrations are noted across the top of the lanes, and protein standards (in kDa) are shown.

at 60μM and TSG101 at 30μM, 60μM and 120μM (Data not shown). However, no significant changes were detected.

The C-terminal of M1(224) remained unchanged when TSG was added, and this is consistent with studies done by Bonavia and colleagues (Bonavia et al., 2010). Though, our finding did not see any interaction between TSG-UEV domain and M1-N terminal domain. There is a possibility that TSG101 full sequence is needed to bind to M1. However, it has been shown that the UEV domain is the important motif for binding to the late domain motif of other enveloped viruses e.g. HIV p6 Gag or Ebola virus VP40. In addition, the UEV domain alone is sufficient to bind to the late domain of the HIV p6 Gag late domain motif (Bonavia et al., 2010). The C-terminal of TSG101 was not needed for this interaction, and the UEV-domain is mediating this interaction.

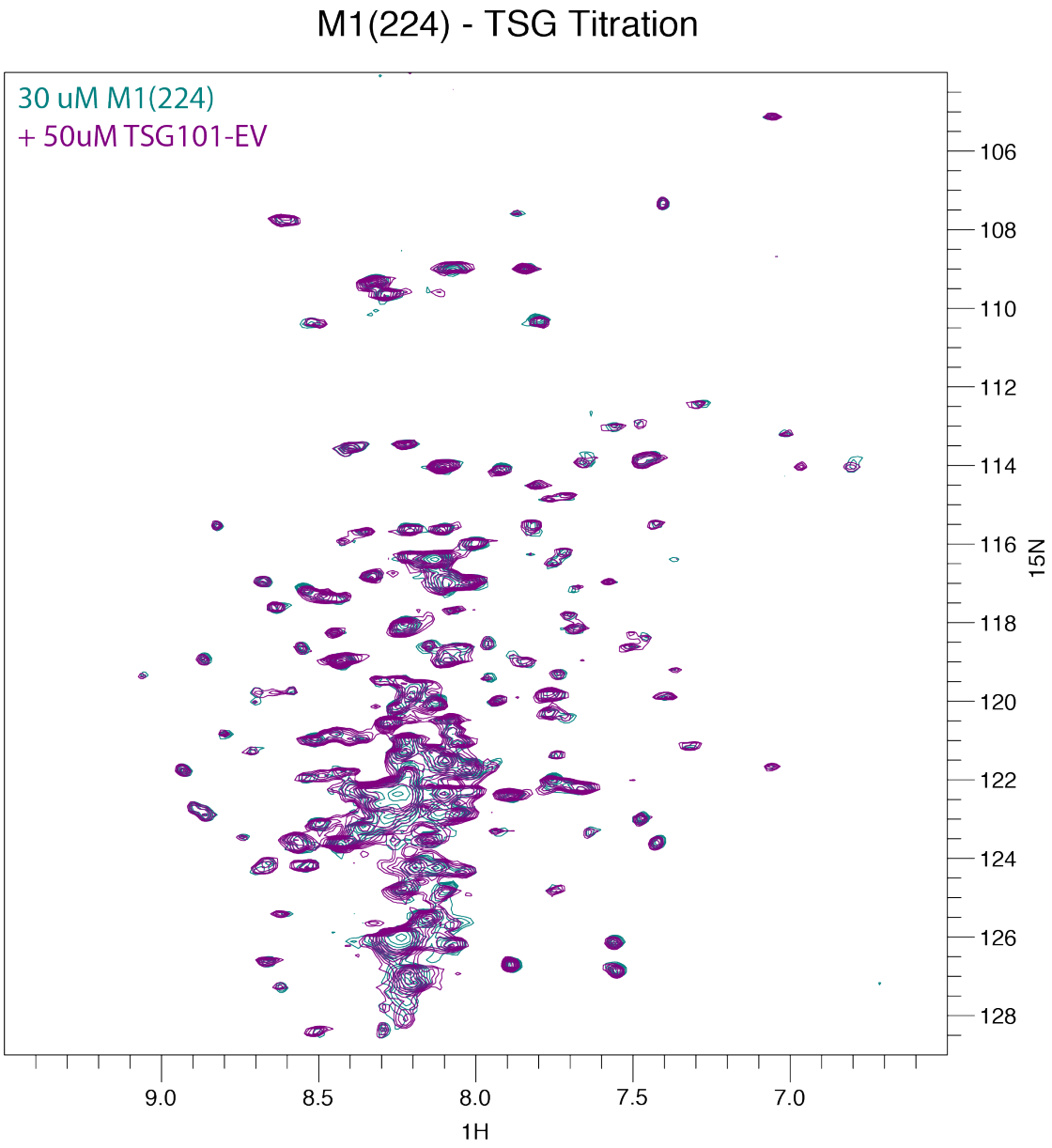


Figure 6.9: M1(224) - TSG101-UEV titration. Overlay of M1(224) at 30 μM before (teal) and after 50 μM TSG titration (purple). Spectra collected at 25°C, pH 7 on 750 MHz spectrometers in 20mM Tris, 50mM NaCl buffer condition. M1(1-224) was prepared at a low pH (pH 5.5) as described previously in Chapter 2. The pH was then gradually increased to pH 7.

Bersusah susah dahulu, bersenang senang kemudian

Initial hardship to enjoy the fruits of labour in the future

— Dad

7 Discussion

Contents

7.1 Solution Study of M1 Structure 131
7.1.1 NMR studies of M1NT and M1CT 131
7.1.2 pH Changes of M1NT 133
7.1.3 K^+ Changes of M1NT 138
7.2 Solution Study of M1 Oligomerisation 140
7.2.1 Oligomerisation of M1 141
7.2.2 Dual Sensing Model of Oligomerisation 145
7.3 Interaction of M1 with Binding Partners - NEP and TSG101 146
7.3.1 NEP 146
7.3.2 TSG 147
7.3.3 NEP and TSG101 148

Matrix proteins exist in many enveloped viruses including HIV-1, Ebola and influenza viruses. They are typically the most abundant proteins in enveloped viruses and act as bridges connecting the viral core and the envelope where viral membrane proteins and glycoproteins reside. Matrix proteins often play multiple essential roles in the lifecycle of viruses from host infection to the later stages including assembling of viral components and budding of fresh virions progeny (Calder et al., 2010; Ge et al., 2010; Liljeroos et al., 2011).

The functional versatility of matrix proteins enables them to protect the virion

genome and act as structural support for the viral envelope while also inducing disassembly of virion components when intracellular conditions are appropriate. These processes typically require interaction with the membrane, and the details of the molecular interactions of how the matrix protein interacts with membranes and with each other are not fully understood (Baudin et al., 2001; Gregoriades, 1980; Gregoriades and Frangione, 1981; Leser and Lamb, 2017; Ruigrok et al., 2001, 2000; Wijesinghe et al., 2017; Ye et al., 1987). Understanding how these matrix proteins interact with each other and other components of the virus and host cell, and how they respond to conditions appropriate to the different steps in the viral life cycle, can reveal mechanisms that can be exploited for therapeutic intervention.

M1, the matrix protein of the influenza A virus, is the structural protein of the virion, lying underneath the viral envelope. M1 is responsible for the varying shapes of viruses that range from spherical to filamentous (Chlanda Zimmerberg, 2016). M1 associates with both the viral membrane and several other viral components thereby acting as a bridge between the membrane and the central components of the virus including the vRNPs, the M2 ion channel, hemagglutinin (HA), and neuraminidase (NA). M1 also has a key role in the last stage of the pathogenesis as a hub for interactions at the viral budozone, which is a region of high concentration of membrane rafts and cholesterol (Simons and Vas, 2004; Lingwood and Simons, 2010; Veit Thaa, 2011). The M1 is thought to gather all the viral components, orchestrating the formation of the virus progeny before budding off from the host cells. However, before budding can occur, a crucial step is for M1 to oligomerise to form the scaffold M1 matrix that forms the enveloped layer supporting the virus shape before budding off as mature virions (Nayak et al., 2009).

A crucial question regarding M1 function is: How exactly does M1 appropriately time polymerisation at the membrane and dissociation during infection? To gain critical insights into the versatile functioning of M1, it is essential to study this protein in solution at high resolution. In this PhD thesis, I have reported the first detailed solution NMR study of M1 in my efforts to understand M1 oligomerisation and interactions.

7.1 Solution Study of M1 Structure

7.1.1 NMR studies of M1NT and M1CT

Here, I have reported the near-complete backbone NMR assignment of M1NT at acidic pH (pH 5.5; Figure 3.4) showing a well-folded alpha-helical domain, which is consistent with crystallographic studies. NMR data and SEC-MALS were consistent with a monomer form of M1NT in solution. Chemical shift-based analyses using TALOS-n are consistent with crystallographic data indicating that M1NT contains two four-helix bundles (*h1* to *h4* and *h6* to *h9*) connected by a cross-over helix, *h5* (residues 70 - 88) (Figure 3.5 & 3.6)

Amide ^{1}H and ^{15}N relaxation measurements provided insights into the internal dynamics of M1NT and are complementary to the X-ray crystallographic studies. As expected, the loop region of M1NT shows flexibility in the loops *f1*, *f2*, *f3* and *f4* (Figure 3.7). Intriguingly, two further regions were highly flexible - the loops between helices 6 to 7 (residues 103-113) and 8 to 9 (residues 132-140). When compared to the available X-ray crystal structure (1AA7) these regions are not flexible or dynamic (Figure 3.9) due to crystal contacts that stabilised these regions. These highly flexible regions may play an important role in the functioning of M1 protein such as its oligomerisation.

The 10 kDa M1 C-terminal (M1CT) region consisting of residues 165 to 252 as well as the full-length M1 protein have not yet been crystallised yet, although it has been this has been attempted several times (Arzt et al., 2001; Sha Luo, 1997). The NMR spectrum of M1CT indicates that it is disordered on its own, with most backbone amide cross-peaks between 8.0 and 8.5 ppm, suggesting that it did not adopt a stable tertiary structure (Figure 4.4). This is consistent with findings of previous reports showing that M1CT is weakly structured at pH 7 (Kordyukova et al., 2008; Ksenofontov et al., 2010; Shishkov et al., 2009, 2011; Zhang et al., 2012). Other biophysical techniques have been used to investigate the structural features of the C-terminal domain, and this includes circular dichroism (CD), dynamic light scattering and analytical ultracentrifugation. All the results

indicate that the C-terminal domain is flexible and weakly structured (Kordyukova et al., 2008; Ksenofontov et al., 2010; Shishkov et al., 2009, 2011; Zhang et al., 2012). In addition, NMR data on M1CT indicates it to be monomeric in solution at pH 5.5. The high degree of flexibility of M1CT indicates that it is unlikely to have a stable dimerization domain. This is in contrast to reports by Zhang and colleagues, which indicate that M1CT is a dimer in solution independent of the pH and concentration based on their gel filtration data K. Zhang et al. 2015. However, the high flexibility of M1CT might cause misinterpretation of the size as unfolded proteins elute similarly to a larger protein due to a larger radius of gyration. It is also possible that M1CT forms a more stable folded structure in the full length of M1.

The possibility that M1CT could form a more stable structure in the full length M1 was tested. Although it was not possible to purify the M1 full-length, a near full-length construct M1(224) that is 29 residues shorter was purified instead. Overlaying the spectra of M1(1-224) and M1CT at both neutral and acidic pHs show that the C-terminal domain remained unfolded and exhibited only minor chemical shift changes (Figure 4.10 & 4.11). The longest construct purified M1(1-232) which is just 21 residues shorter from the full-length also showed the C-terminal domain in an unfolded state. It can be safe to assume that the full-length M1 would likely show an unfolded C-terminal domain. The C-terminal domain sequence is mainly consist of disorder-promoting amino acids. However, it is likely that the C-terminal domain forms a more stable folded structure in the polymerised form of full-length M1 (Zhang et al., 2017). In full-length structures of M1 from Influenza A Virus (IAV), the M1CT is folded into a helical bundle, and both M1NT-M1NT and M1NT-M1CT intermolecular contacts are formed. It was shown that the C-terminal domain folded into three-helix globular domain in a linear polymer and the C-terminal domain was interacting with the membrane surface of the N-terminal domain from an adjacent M1 monomer. Furthermore, the C-terminal domain of M1 might also act as an IDP and fold upon binding with interacting partners such as M1 monomers. This may facilitate interactions with other viral proteins and host cell proteins as similarly seen in HIV-1 viral proteins (Goh et al., 2008).

7.1.2 pH Changes of M1NT

In both neutral pH and acidic pH, M1NT and M1(224) remain monomeric. In addition, the C-terminal domain in the M1(1-224) construct remains flexible and unfolded. Extending the C-terminal domain in the M1(224) construct does not affect the oligomerisation state—it remains as a monomer in solution at both pHs. This result is also consistent with the gel filtration data through the purification of M1CT and also other studies (Arzt et al., 2001; Shtykova et al., 2013). M1 extracted from M1-vRNP complexes was also found to be monomeric at acidic pH and when neutralised at a higher pH, stays monomeric (Zhirnov, 1992). This is in contrast with the findings from Sha et al. that show that both the full-length M1 and truncated M1NT migrate as dimers in gel filtration experiments at low and neutral pHs (Sha et al., 1997). However, these findings highlight that this could be an anomalous migration due to the unstructred C-terminal domain (Arzt et al., 2001). Furthermore, reports by Zhang et al., through their sedimentation velocity analytical ultracentrifugation and gel filtration experiments, have identified that intact M1 and M1NT in the dimeric and higher oligomer forms at neutral pH (Zhang et al, 2015). Full-length M1 was found to be in the dimeric form while M1NT was only found in its monomeric states in acidic pH.

Crystal structures have been reported at both high and low pH. However, structural studies by X-ray crystallography have suggested that the N-terminal domain of the Influenza A Virus M1 protein exists as a dimer (Harris et al., 2001; Safo et al., 2014; Sha and Luo, 1997) contrary to what has been observed in solution by NMR. This is probably because in crystals, protein may form crystal contacts that are possibly not occuring in physiological conditions The crystal structure at neutral pH (PDB ID: 1EA3) has been reported by Arzt et al., and at low pH (PDB IDs: 1AA7 and 4PUS) by Sha and Luo, and Safo et al. respectively (Arzt et al., 2001; Safo et al., 2014; Sha and Luo, 1997). These structures have been used as representatives for high and low pH protein structures. Comparison of the monomers of the M1NT domain between these crystals resulted in least squared superpositions with an RMSD value of 0.8.

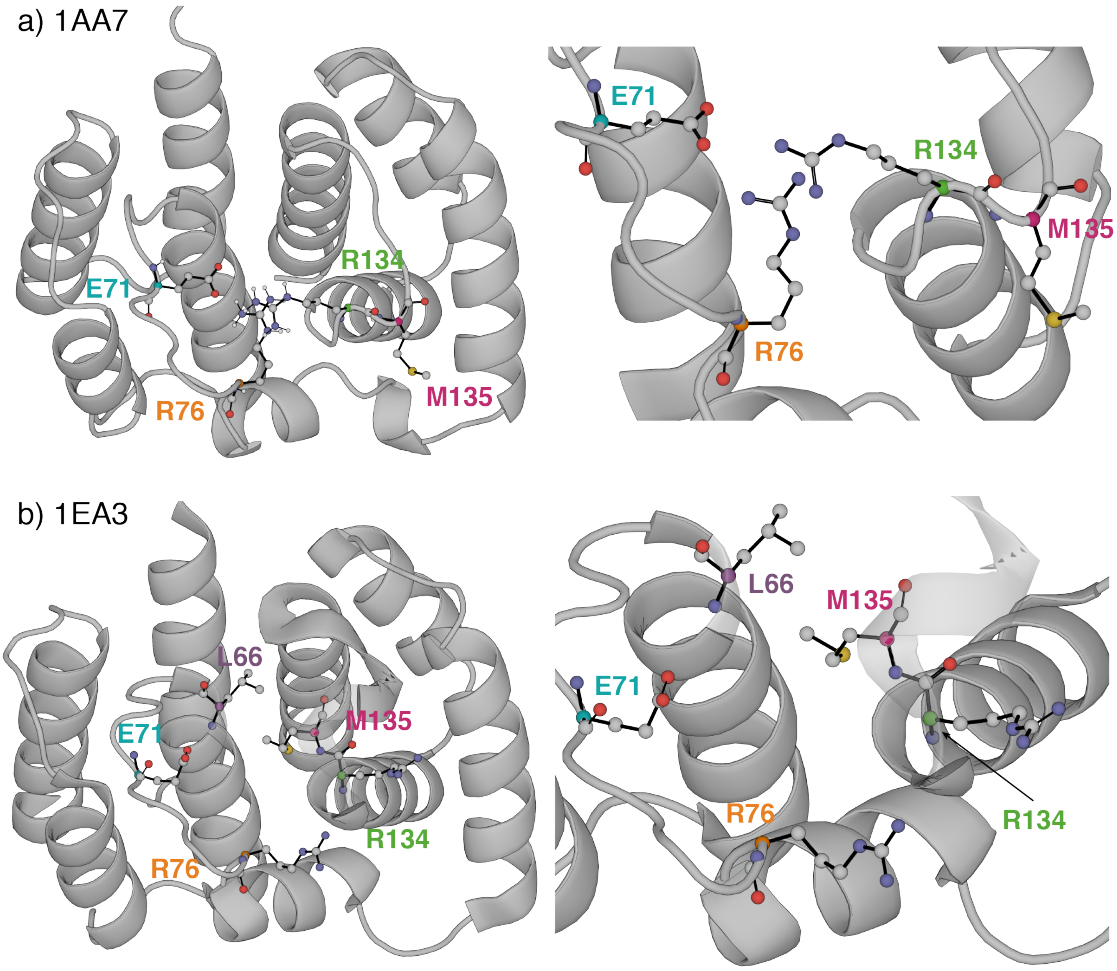


Figure 7.1: The repacking of R134 and M135 in 1AA7 and 1EA3 crystal structures. a) In 1AA7, M135 side chain is flipped “out" and away from the R134 sidechain, close to the protein surface. b) In 1EA3, M135 side chain is flipped “in" being tightly buried into the core of M1NT with L66 within the hydrophobic surface patch. R134 and R76 are flipped outward and contribute to the positively charged surface involved in membrane binding.

Significant CSPs were observed on M1 constructs in response to pH changes between pH 6.0 and pH 6.5 (Figure 3.11), and these highly perturbed residues reside on helix 1 (h1) and helix 9 (h9). Analysis of 9 M1NT crystal structures suggests that there is a correlation between a change in pH and the shifting between intramolecular subdomains - subdomain 1 (h1 to h4) and subdomain 2 (h6 to h9) (Figure 3.12 & 3.13*a*). Early studies of M1NT have found large changes in how the M1NT stacked dimers is arranged as a result of pH-dependent rotation across domains resulting in tilting (Harris et al., 2001). However, none of the earlier work noticed the subtle intramolecular subdomain movements. This is consistent with the analysis done on

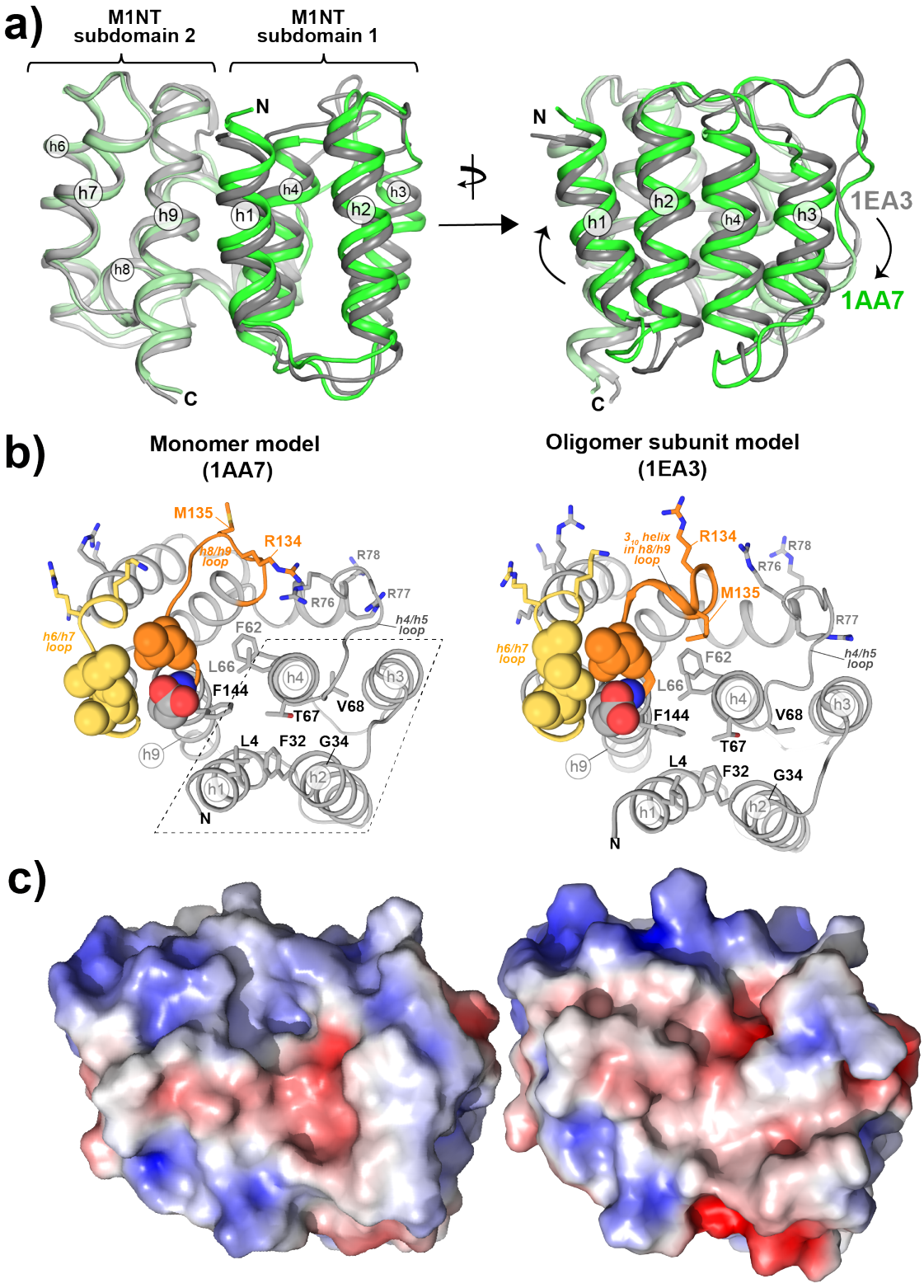


Figure 7.2: (Caption next page.)

Figure 7.2: (Figures previous page.) a) Comparing the two crystal structures - 1EA3 (Neutral pH, stacked dimer subunit) and 1AA7 (low pH). Both are derived from A/Puerto Rico/8/1934 H1N1. Here the two structures are aligned on the Cα atoms of the C-terminal domain (helices 6 to 9) to demonstrate the relative shift of the N-terminal subdomain (helices 1 to 4). b) 1AA7 and 1EA3 can be used as models for the monomer and oligomer subunit structures. The hydrophobic patch (shown in orange and yellow) which is sensitive to water can be seen centred around T67 of helix 4 (h4). The view is oriented such that the membrane-binding surface is facing up. Helices 1 to 4 form subdomain 1 as indicated by the dashed lines. The formation of the 3_{10} helix (residues 134 to 137 of helix 8/helix 9 (h8/h9 loop) in orange) of 1EA3 causes R134 to flip out facing the membrane-interacting surface while the M135 is packaged into the subunit interface. The close packing of residues T139/E141 of h8/h9 loop and residues I107/T108 of h6/h7 loop may help reduce the loop mobility seen for the monomer in solution. c) The electrostatic properties of 1AA7 (left) and 1EA3 (right) are shown with surface representations with the same orientation as in b). The conformational changes between 1AA7 and 1EA3 that are seen in T67/V68 and induced by CHAPS result in a rearrangement of the helix 4/helix 5 loop (h4/h5 loop) that remodels the oligomer interface. Large surface changes are also seen near the membrane binding surface with the formation of the 3_{10} helix of the h8/h9 loop in 1EA3.

the 6 M1NT crystal structures with stacked dimer arrangements, showing a strong correlation of the tilting of stacked dimers at lower pH (Figure 3.13*b*) and may be related to the disrupted form of M1 oligomers when exposed to low pH due to conformational change (Calder et al., 2010a; Sha and Luo, 1997; Shtykova et al., 2017; Zhang et al., 2015). Hence, subdomain rotation between subdomain 1 and subdomain 2 of M1NT might correlate with the geometry of arrangement of the stacked dimers and might play a role in regulating oligomerisation.

Solved M1NT crystals have similar confirmations between monomers but have different associations between crystal dimers. For both the low pH structures, 4PUS and 1AA7, and other M1NT crystal structures, a side-by-side packing arrangement in the asymmetric unit can be seen (Safo et al., 2014, and Sha and Luo, 1997). However, the unique arrangement of 1EA3 in the asymmetric unit at neutral pH shows a stacked dimer arrangement. Through M1NT characterisation in solution as well as analysis of M1NT crystal structures, specific features of 1EA3 are proposed to be relevant to the membrane-bound oligomerised form of M1. The uniquely extended conformation of T67 was observed in 1EA3 upon addition of CHAPS which is likely to exert its effect through a reduction in water activity; this may be

needed for a close packing of subunits. In addition, an interesting feature that was seen in 1EA3 is the single turn of 3_{10} helix found in the loop between helix 8/helix 9 (*h8*/*h9*). This 3_{10} helix is stabilised by an H-bond formed between R134 and A137. Resonances from *h8*/*h9* loop were not detected and hence not assigned in the HSQC except for G136. Dynamic experiments have shown that this region is disordered in the water-solubilised, monomeric form of M1 studied here. Interestingly, the 3_{10} helix is found in all 6 M1NT crystal structures that were solved at pH 7 or above and were not present in all (4 out of 5 structures) but one crystal structure, 4PUS that was determined at pH below 7. 4PUS was solved at pH 4.7 but does not contain a stacked dimer arrangement.

Examinations of the associations between crystal dimers between crystals with and without the presence of the 3_{10} helix, taking 1EA3 and 1AA7 as models, have shown that the largest changes occur within the residues from 132 to 140 as well as the membrane-facing side of the M1NT. As shown in figure 7.1, M135 and R134 show the most dramatic repacking. In the 1AA7 without the 3_{10} helix, M135 is seen to be in the "out" conformation exposing itself rather than being packed into the hydrophobic surface patch formed by F62 and L66 of helix 4 (7.2*b*). The sidechain of R134 is packed against M1NT, forming a salt-bridge and π-stacking interactions with E71 and R76, respectively. Formation of the 3_{10} helix seen in 1EA3 results in the interactions formed by R134 to be broken and M135 to be in the "in" conformation tucked tightly into the core of M1NT. R134 and R76 can be seen to flip out into the exposed solvent onto a positively charged surface that is expected to be the membrane-binding face of the M1NT (Arzt et al., 2001; Das et al., 2012; Höfer et al., 2019; Kerviel et al., 2016; Peukes et al., 2020; Shishkov et al., 1999). Thus, this results in the large changes seen in the electrostatic profile at the membrane binding surface. The 3_{10} helix might be stabilised by interactions with negatively charged membranes or membrane mimetics (Figure 7.1c).

Furthermore, M135 is packed into the core of M1NT against residues L66 and F62 of helix 4 as well as residues of the hydrophobic surface patch. The relevance of this repacking of M135 in the “in” position into the hydrophobic surface

patch for the neutral pH stacked oligomer is unknown, but M135 is found to be highly conserved in influenza A (Darapaneni, 2015) and hydrophobic sidechains are preferred at this position (Hom et al., 2019).

M135 might be sensitive to solution conditions as some CSPs were seen for L66. Further stabilisation may come from closer contacts between residues 139/141 of the helix 8/helix 9 loop and 107/108 of the helix 6/helix 7 loop (Figure 7.1b). The flexibility of these loops observed in the solubilised monomer form might help regulate the oligomerisation of M1. The loops form a large part of the subunit interface, and the differences between the two crystal structures might give insights into its role in the oligomerisation of M1.

7.1.3 K^+ Changes of M1NT

The core of the virus is composed of M1 and vRNPs. The core is stabilised by the interaction between the vRNPs and M1 matrix formed underneath the membrane. Through binding to both the membrane and to the vRNPs, the M1 matrix envelope provides a stable structure and stability for the virus particles. The entry of the Influenza A virus into the host cells involves several steps which include endocytosis and fusion at late endosomes. During this process, the viral core disassembles into its components, uncoating at the membrane.

A study by Stauffer et al. has shown that the Influenza A Virus needs stepwise priming after entry into the host cell for efficient uncoating and infection (Stauffer et al., 2014). The uncoating process begins before the core is delivered into the cytosol. It was found that the priming of the core requires a two-step process mediated by the M2 ion channel. The first step involves the activation of the M2 ion channel through the exposure of the virus in the early endosomes at a pH between 6.0 to 6.5 with low K^+/high Na^+ concentrations. Upon activation, protons are pumped into the virus through the M2 channel acidifying the core of the virus. This triggers a conformational change of M1 that induces the weakening of M1-M1 and M1-vRNP interactions (Zhirnov, 1990).

The second step is triggered through the exposure of the core to a pH of below 6.0 to 5.4 and by an influx of K^+ ions. Stauffer and colleagues suggested that this possibly weakens the remaining M1-vRNP and vRNP-vRNP associations. The pre-acidified core is exposed to an increase in the concentration of protons and a high concentration of K^+ replacing the high concentration of Na+. This exposure of both high levels of H^+ and K^+ promote priming efficiency of uncoating and infectivity.

Stauffer et al. have shown that exposure of the virus to a pH 5.8 and low K^+ concentration causes irreversible changes of the M1 C-terminus through trypsin-based limited proteolysis (LiP) experiment. An *in vitro* study by Harris et al., also shows that a peptide sequence between residues 139 to 166 of M1 (the linker region) undergoes a conformational change under acidic conditions (Harris et al., 2001). It was suggested that this conformational change occurs in the linker region from an alpha-helical conformation to a partially unfolded form. Changes are similarly seen in the linker region and flanking sequences by Stauffer through LiP analysis. This is possibly due to the changes in the conserved region of the highly conserved Zinc finger motif of the linker that causes changes in the interaction between the N and C-terminal domains of M1, weakening the M1-vRNP interactions.

The M1 C-terminal domain is said to bind with the vRNPs (Baudin et al., 2001). The acidic condition possibly changes the conformation of the M1 linker region. Furthermore, the K^+ ions could also weaken the contacts between the vRNPs and the C-terminal of M1 as it is known that K^+ ions affect the secondary structure of RNAs. Taken altogether, this causes the changes in the M1 matrix and the vRNP bundle stability. The combinatorial effect of both low pH and high K^+ concentration may explain the effective uncoating and infectivity.

The M1 construct M1(224) was tested for any structural changes in response to both pH and K^+ concentrations. Despite, the implicated evidence caused by H^+ and K^+ ions mentioned above, there were no significant chemical shift perturbations between neutral (pH 7.0) and acidic (pH 5.5) pHs (Figure 3.15 and 3.17) when different concentrations of K^+ at various pHs were tested. There were no structural changes that were seen in either the N-terminus and C-terminus domains. This

experiment was tested using M1 monomers which raised the possibility that M1 may need to be in its polymerised form or to be bound to the viral component (vRNPs) in order to be able to see any significant changes caused by differing K^+ concentrations. This needs to be further tested to confirm this hypothesis.

7.2 Solution Study of M1 Oligomerisation

Studies have shown that M1 has a strong propensity to self-associate (Shtykova et al., 2017; Zhang et al., 2012). This ability to self-associate is critical for the formation of the matrix layer especially during the late stage of the viral pathogenesis, the assembly, and budding stage (Ali et al., 2000; Avalos et al., 1997; Barman et al., 2001; Gómez-Puertas et al., 2000; Latham and Galarza, 2001; Wang et al., 2010). It was also shown that this oligomerisation is required for the next cycle infection (Nayak et al., 2009, 2004; Rossman and Lamb, 2011; Ruigrok et al., 2000). M1's tendency to oligomerise increases with increasing pH. At low pH, M1 exists mostly as monomers (Arzt et al., 2001; Shtykova et al., 2013) and it was shown that a small fraction of M1 also exists as helical clusters (Shtykova et al., 2013). The proportion of clusters increase beyond pH 6.0 and hence, increasing the tendency of M1 to oligomerise at pH 7 (Shtykova et al., 2017). It should be noted that the full length of M1 is used for these studies where the C-terminal domain might facilitate the formation of this clusters.

A study done by Baudin and colleagues has shown that full-length M1 as well as M1NT bind effectively to negatively charged lipids, phosphatidylserine (PS) (Baudin et al., 2001). Furthermore, a study done by Hilsch and colleagues using both model systems and living cells concluded that the binding of M1 to lipids and M1 self-association are not mutually exclusive processes and are probably concomitant processes (Hilsch et al., 2014). They showed that the interaction of M1 with lipids increases M1 self-association, where they detected oligomerisation of M1 on lipid membranes. As here, they do not detect any oligomers in the absence of lipid membrane in solution and M1 does not form VLPs without membrane targeting.

The mechanistic details of how M1 self-associate and how membrane binding affects the oligomerisation of M1 remains a mystery and represents one of the most important questions in understanding Influenza A Virus pathogenesis. Despite the numerous studies that have been done to understand M1 self-association, how exactly M1 self-associates is poorly understood. Particularly, is is unknown which subunit interactions are important for M1 self-association and how M1 conformationally changes in response to polymerisation are the questions for which answers are not yet known. Below, I discuss my efforts in answering these questions using NMR spectroscopy.

7.2.1 Oligomerisation of M1

In Chapter 3, several membrane mimicking surfactants were tested on M1NT structure and we found that sterol-containing membrane mimetics induced specific conformational changes in M1NT. These surfactants, including CHAPS and CHAPSO, were found to perturb residues at the hydrophobic patch of M1NT, specifically around T67 residue—the backbone amides of residues L4, G34, T67, and E141. High concentrations of sterol-containing compounds were required to induce the conformational changes observed in the M1NT and saturation concentrations correlated well with critical micelle concentrations indicating that the effects were mediated by sterol monomers as beyond the CMC bicelles has no effect on the hydrophobic patch of M1.

The perturbed residues were found to be from the subdomain of highly conserved N-terminal subdomain 1 (helix 1, *h1* to helix 4, *h4*) (Hom et al., 2019). Similar results from the different sterol compounds tested suggest that some aspects of the sterol group might induce specific CSPs on M1. However, isopropanol and glycerol, which are chemically very different compounds to the sterol compounds were also tested and were found to perturb residues in a similar way to the addition of sterol compounds. This suggests that these compounds act indirectly on M1, and probably through a reduction in water activity (Marcolli and Peter, 2005; Zhu and Grant, 1996). Furthermore, the effects appear not to be mediated by an increase in

water structure since chaotropic additives did not specifically perturb residues in the hydrophobic cavity. Kosmotropes, on the other hand, may affect water structures; however, they did not perturb residues similarly to sterols compounds.

One hypotheses of the effects of reduced water activity is an extension of the backbone of T67. Analysis of the M1NT crystals, the closely packed stacked dimer arrangement seen in 1EA3, and the side-by-side or lateral dimer arrangement seen in 1AA7, has shown that there was a difference in the T67 conformation between the two crystals. T67 appears to be uniquely in an extended backbone conformation in 1EA3 as compared with its conformation in all other structures of M1NT. Hence, the effect of reduced water activity was consistent with the extended backbone conformation seen in 1EA3 crystal structure which forms the tightly packed stacked dimer structure.

The perturbed residues that is possibly caused by water activity are not just at the subdomain 1, but the residues (E141 and F144) in helix h9 were also strongly perturbed. Helix 9 forms more than half of the subdomain interface as it packs against helices 1 and 4 and the strong perturbations observed are consistent with the movement of the subdomain (Arzt et al., 2001). Furthermore, T67 and V68, which are very sensitive to reductions in water activity, reside at the start of the loop between helix 4 and the subdomain linker helix 5. As previously discussed in Chapter 3, the analysis of crystal structures confirms that the subdomains can rotate relative to each other (Figure 3.12 & 3.13). Subdomain rotation may also provide a mechanism which M1 uses to coordinate between the two oligomerisation surfaces. The similarities of perturbed residues caused by reduced water activity and change in pH, especially in helices 1 and 9, suggest that subdomain rotation might be involved in regulating the oligomerisation of M1 as conformational change at the hydrophobic patch is propagated across the C-terminus of M1NT which subsequently might affect C-terminal domain of M1 through conformational change.

The C-terminus of M1 might also plays a role in regulating oligomerisation. In full-length M1, the C-terminus was seen to adopt a stable tertiary structure and was packed against the N-terminal domain of an adjacent subunit to stabilise

oligomerisation (Peukes et al., 2020; Zhang et al., 2017). Although the constructs used here contain partial C-terminus sequence, a longer construct of M1 with C-terminal extension, M1 (1-224), was found to be more sensitive to changes in water activity. This correlates with its higher propensity to self-associate as compared to M1NT (Zhang et al., 2015). Upon the reduction of water activity, the downfield shifts of 13C' and 13Cα chemical shifts were observed at the end of helix 9, which indicates an increase in helix formation (Figure 5.33). This may suggest a mechanism of allostery between the N-terminal interface that is sensitive to water activity and the C-terminal domain.

Cholesterol in the form of cholesteryl hemisuccinate (CHS) solubilised in CHAPS was found to mediate M1 self-association on both M1NT and M1(224) at pH 7.0, which led to the loss of most signals from the M1NT domain (Figure 5.14). However, at pH 5.5, CHS-mediated association was not seen (Figure 5.15). When these residues specifically perturbed by sterol surfactants were examined, it was observed that the residues form one of the interfaces involved in the stack arrangement of the 1EA3 crystal structure (Figure 5.18a).

Mutants based on the 1EA3 crystal structure residues forming the interface of the dimer were produced and tested for their ability to form CHS-mediated associations. The M1NT single-site mutants L4A, L4K, I51D, T67S and E141A, were unable to self-associate in the presence of CHS at both neutral and acidic pHs. Moreover, the M93A/D94A double mutant of M1NT, which we predicted to disrupt the side-by-side dimer interface seen in the 1AA7 crystal structure, was able to self-associate at neutral pH. Similar to the wild type M1NT, it was unable to self-associate at acidic pH. This suggests that the stacked dimer interface and not the lateral interface were involved in oligomerisation.

The T67 resonance CSP upon addition of CHAPS was also measured for all mutants and compared with the wild type of M1. Interestingly, it shows that the T67 movement correlates well with the ability to oligomerise. Wildtype M1NT at pH 5.5 and mutants based on 1EA3 crystal structure shows a comparable CSP and are unable to self-associate. In contrast, wildtype M1NT at pH 7.0 and the

double mutant M1NT-M93A/D94A based on crystal structure 1AA7 shows that they were able to self-associate. It may be that a critical level of shift is required to induce self-association as this is consistent with the larger shift seen in M1(1-224) upon interaction with CHAPS and also, the higher propensity of M1(1-224) to self-associate compared to M1NT. pH also plays an important role as results have shown that M1NT, M1(1-224) and the double mutant, M1NT-M93A/D94A was able to self-associate at pH 7.0 and not at pH 5.5. This is consistent with previous reports (Arzt et al., 2001; Calder et al., 2010b; Fontana et al., 2012; Fontana and Steven, 2013; Shtykova et al., 2017, 2013).

The combination of CHAPS and CHS induces the assembly of M1NT domain into large complexes (oligomers) mediated by the same residues that are both sensitive to water activity. Furthermore, oligomerisation required the cholesterol derivative CHS. Oligomerisation was not observed with 25-hydroxycholesterol, which lacks the negatively charged succinate. Oligomerisation was also not seen in the presence of CHS with LMNG in which LMNG on its own does not interact with M1 or induce a similar shift seen with sterol compounds, glycerol or isopropanol. This indicates that the negatively charged succinate moiety of CHS is necessary to induce M1 self-association. This may relate to M1's ability to interact with negatively charged membranes via its strong basic surface (Ruigrok, 2000; Baudin, 2001; Thaa et al., 2009) which may facilitate direct interactions with the CHS micelles.

Results from M1NT mutants are consistent with CHS-induced oligomerisation occurring at the stacked dimer interface. However, it is unknown what the relationship is between binding to CHS and CHAPS micelles in the context of physiological membranes. Cholate and deoxycholate induced similar perturbations like CHAPS and CHAPSO but also appear to not bind directly to M1NT. In addition, they did not induce oligomerisation despite being negatively charged. When cholate and deoxycholate are compared to CHS, the position of negative charges is very different and hence presumably these negative charges are different also in their positioning within a micelle. This might suggest that there is some dependency on the geometry of the charges for a direct interaction. Further studies

to understand the mechanism of CHS-induced oligomerisation are needed. In addition to the geometry of the charges, a combination of preferred pH and a critical level of chemical shift perturbation resulting in conformational change of M1, is needed to be induced for self-association.

7.2.2 Dual Sensing Model of Oligomerisation

Looking at the two main crystals arrangements, 1AA7 (side-by-side dimer arrangement) and 1EA3 (stacked dimer arrangement), may serve as a useful model for the conversion of the monomeric M1NT to the membrane-bound, polymerised form, respectively (Figure 7.2*b*). Comparing the two crystal structures, the most substantial conformational change occurs at the membrane facing side of M1NT and within residues 132-140. As discussed in Chapter 5, within these residues in 1AA7, the most dramatic change is the repacking of R134 and M135 where the M135 is seen to be flipped out exposing itself while the R134 sidechain is packed against M1NT which forms a salt-bridge and pi-stacking interactions with E71 and R76, respectively. In 1EA3, the interactions formed by R134 are broken, and M135 is flipped in and tucked tightly into the core of M1NT. Additionally, R76 was also flipped into the exposed solvent. R134 and R76 have been implicated to be at the membrane-binding face of M1NT (Arzt et al., 2001; Das et al., 2012; Höfer et al., 2019; Kerviel et al., 2016; Shishkov et al., 1999).

Below, a model of allosteric communication is proposed in which membrane binding induces the formation of stacked-dimer interfaces facilitating polymer assembly. Membrane binding is known to mediate M1 oligomerisation, although the exact mechanism of the process is unknown (Hilsch et al., 2014). M1 uses the proximity to the membrane as a sensor via the reduced water activity near the membrane. Studies performed on 1,2-dipalmitoyl-sn-glycero-3-phosphatidylcholine (DPPC) membrane show that the rate of water diffusion and water activity decreases when in proximity to the membrane (Fisette et al., 2016). At the interfacial region, the boundary formed between the membrane and aqueous solvent, water molecules are in a frozen-like conformation, unlike in bulk water (Tielrooij et al., 2009).

This is due to the interactions of the phosphate head groups hydrogen bonding to two water molecules strongly. This results in reduced water activity and can be used as a sensor when in proximity in the membrane, which may induce a conformational change in M1. Sensitivity to water can provide a mechanism for restricting protein-protein association to the membrane surface.

Under the dual-sensing model, the M1 N-terminal domain is sensitive to both water activity, suggesting that the protein senses a combination of both changes in water activity , and the negative charges that are unique to the membrane surface. The electrostatic interactions drive attachment to the host cell membrane, increasing the local M1 concentration. The sensing of a combination of the right pH (neutral pH), changes in water activity due to its proximity to the membrane and interactions of R134 and R76 to the negatively charged phospholipids induces the repacking of the hydrophobic surface patch. This conformational change in turn propagates its change via helix 9 (F144) to the C-terminal side of the M1NT and hence induces polymerisation. This model restricts M1 only to the unique environment of the membrane surface through its dual sensing of negative charge and reduced water activity. At the membrane-water interface, water activity is greatly reduced, and water ordering increases (Damodaran, 2015; Nagle and Tristram-Nagle, 2000). This can affect solute hydration and hence the energetics of protein-protein interactions. These conformational changes are allosteric in nature and allow for conformational changes at one subunit surface to be propagated to the C-terminal side. Thus, this may pave the way to regulating the direction of oligomerisation.

7.3 Interaction of M1 with Binding Partners - NEP and TSG101

7.3.1 NEP

Several studies have implicated the interaction of M1 and NEP. Baudin and colleagues have reported that the C-terminal domain of M1 binds to NEP while another report suggests that the N-terminal domain of M1 is primarily responsible for the binding of NEP (Baudin et al., 2001; Noton et al., 2007). Furthermore,

another study suggests that the C-terminal domain of M1 is responsible for the binding of NEP (Stauffer et al., 2014). NEP is a critical component in the export of newly synthesised vRNPs for the budding of progeny virions, which requires the formation of a nuclear export complex in a daisy chain manner (Figure 6.1). This nuclear export complex consists of vRNP, M1, and NEP (Akarsu et al., 2003; Huang et al., 2013; Neumann et al., 2000; O'Neill et al., 1998). The binding sites of M1 on NEP have been mapped to the residues 50-121 (Reuther et al., 2014; Shimizu et al., 2011). M1 is unable to interact directly with CRM1 (Arkasu et al., 2003), and instead, NEP is proposed to form a bridge for M1/CRM1 interaction acting as a mediator by binding to the M1 C-terminal domain (Arkasu et al., 2003). This enables the export of vRNP from the nucleus to the membrane where the budding of virions via the CRM1-dependent export pathway occurs (Fakuda et al., 1997; Neumann et al., 2000).

Results show that there are no significant changes in CSPs seen in these spectra, indicating there is no binding between either NEP with M1NT or M1NT with NEP. These experiments were repeated at different molar equivalents and in addition, a longer construct of M1 with C-terminal extension, M1(1-224) was also tested, but there were still no interactions indicating neither concentration nor the presence of C-terminal region of M1 was the limiting factor. Furthermore, the possibility that NEP would bind to the polymerised form of M1 was also tested, and the results were also negative.

7.3.2 TSG

TSG101 was also tested with M1 and failed to show any interaction. The construct used with TSG101 contains a C-terminus extension, M1(1-224), although the isolated N-terminal domain was the domain that was seen to be interacting with TSG101 in other studies and this confirms our results that the C-terminal domain of M1(224) remains unchanged (Bonavia et al., 2010). This is consistent with studies by Bonavia and colleagues, where they showed that the C-terminal domain does not contribute binding to M1 (Bonavia et al., 2010). It is also possible that

interactions between TSG101 and M1 are happening on a timescale that doesn't allow detection by NMR. Isothermal calorimetry was used to detect this interaction. However, the M1 sample was not homogenous due to M1-M1 interactions and hence, the data was not interpretable.

A study by Bonavia et al. has shown that TSG101 was involved in facilitating the budding of virions of influenza-infected cells and M1 was shown to interact with TSG101 via its UEV domain (Bonavia et al., 2010). Like the matrix proteins of other enveloped viruses of HIV, p6 Gag and Ebola virus, VP40 which contains late domains that interact with the UEV domain of TSG, M1 also contains the canonical late-domain motif, YKRL (Martin-Serrano et al., 2001, Freed, 2003; Bonavia et al., 2010). Furthermore, M1 contains another putative late-domain motif, PTGP between helix 1 and helix 2 that could also serve as a binding site for TSG101. Although for M1, IAV budding does not require TSG101, the ESCRT pathway or the late domain motif. Through NMR, no binding can be seen between M1 and TSG101-UEV domain.

7.3.3 NEP and TSG101

Despite all the previous studies that have been performed implicating the interaction of M1 with NEP, and TSG101, our results from NEP- or TSG101-M1 binding experiments have failed to show any interactions. There is a possibility that the interaction of +M1 and NEP requires vRNP to be present at the same time or some cellular context is missing such as CRM1. This may explain the lack of binding in the results seen between M1 and NEP through NMR. There is also a possibility that M1 C-terminal region is involved in the interaction with NEP where it needs to undergo some conformational change. As with TSG101, perhaps the full length of TSG101 is needed. As with NEP, another reason could be that other interacting binding partners need to be present at the same time.

You might work but I work harder
You might fight but I fight smarter
Might not be on the top of the world but hey
I'm on my way

— Charlie Brown

8 Conclusion

Contents

8.1 Conclusion 149

8.1 Conclusion

M1 is involved in many different stages of the Influenza A virus pathogenesis cycle of infection and hence, is involved in a large number of interactions with cellular host and viral factors including TSG101 (Bonavia et al., 2012), and the influenza M2 protein (McCown Pekosz, 2005, 2006; Iwatsuki-Horimoto et al., 2006; Chen, Leser, Jackson, Lamb, 2008; Wang et al., 2010), NEP (Baudin et al., 2001; Noton et al., 2007; Bialas et al., 2014; Stauffer et al., 2014), vRNP (Noda et al., 2012) and glycoproteins: HA and NA (Simons Garoff, 1980; Ali et al., 2000; Barman et al., 2001). M1 interactions with these binding partners may also regulate M1 oligomerisation. However, there have been no structural studies or biophysical measurements of any of these interactions. The work conducted here did not detect any interactions with two of M1's binding partners (NEP and TSG101).

These interactions with M1 may need multiple binding partners at once, or require some cellular context that is missing. Perhaps, M1 needs to undergo large

conformational changes such as what has been seen in other enveloped matrix proteins like VP40 of Ebola virus (Bornholdt et al., 2013) to interact with its binding partners. The lack of structure in the C-terminal region and the high propensity for self-association has hindered studies of longer versions of M1, which may reveal important conformational changes. We have shown here that the addition of C-terminal residues affects the structural plasticity of M1NT. There is a possibility of undiscovered conformational changes in the C-terminal region that may be involved in regulating M1 oligomerisation. The C-terminal domain of full-length M1 was seen to adopt a stable tertiary structure packed against an adjacent subunit of the N-terminal domain stabilising the polymer structure (Peukes et al., 2020; Zhang et al., 2017).

It was found that the combination of CHAPS and CHS were able to form large complexes which is possibly the oligomerisation of M1NT that perhaps mediated by the same residues that are both sensitive to water activity and found at the stacked dimer interface. The geometry of the negatively charged succinate of CHS possibly plays a part in inducing oligomerisation. Furthermore, cholesterol may also play a role in triggering oligomerisation. Shilova and colleagues show that M1 adsorption to the lipid membrane is not driven by cholesterol but only requires PS. However, this does not exclude the possibility that cholesterol is needed for polymerisation (Shilova et al., 2017). Moreover, influenza A virus is known to assemble and bud from the viral budozone, which includes a region of high concentration of membrane rafts and cholesterol (Veit Thaa, 2011).

M1NT of the influenza A virus was found to be sensitive to water activity. In this thesis we proposed a dual-sensing mechanism in which M1 senses both changes in water activity and negative charges found at the membrane surface. This will result in the M1NT domains to be repackaged to allow for oligomerisation facilitating polymer assembly. The allosteric conformational change at the surface of the N-terminal domain is manifested to the C-terminal side and possibly serves to regulate the direction of polymerisation. Matrix proteins of other enveloped viruses are structurally diverse; however, these matrix proteins must be able to sense

and respond through conformational changes to their unique local environment at the membrane surface. So far, the mechanism in which the polymerisation was manifested itself was unknown. This work, however, opens up the possibility that other unrelated virus matrix proteins may also be able to sense the unique aqueous environment of the membrane surface to regulate assembly.

Some of the regions identified here as being involved in priming M1 for oligomerisation are phosphorylated, including helix 1 (S2/T5 and T9/Y10) and the loop between helices 2 and 3 (T37) (Hutchinson et al., 2012). Thus, we speculate that phosphorylation also plays a role in regulating M1 polymerisation. Notably, the S2/T5 site is very near to L4, which was exploited here in the L4A mutant used to disrupt oligomerisation and stabilise the monomer. Additionally, S140, which is near to the conformational switch, is strongly predicted to be phosphorylated by PKC, a kinase known to act on M1 (Reinhardt and Wolff, 2000). This opens up future experiments that can help understand whether phosphorylation at these sites plays a role in regulating M1 oligomerisation. Another interesting area to investigate is the potential for targeting the hydrophobic pocket in which T67 resides for novel anti-viral drugs. M1 is the most abundant and most highly conserved protein of the influenza A virus (Ito et al. 1991) and our work has shown that this interface plays an important role in M1 oligomerisation.

References

Accola, M. A., B. Strack, and H. G. Göttlinger (June 2000). "Efficient Particle Production by Minimal Gag Constructs Which Retain the Carboxy-Terminal Domain of Human Immunodeficiency Virus Type 1 Capsid-P2 and a Late Assembly Domain". eng. In: *Journal of Virology* 74.12, pp. 5395–5402.

Adu-Gyamfi, Emmanuel et al. (Feb. 2013). "The Ebola Virus Matrix Protein Penetrates into the Plasma Membrane: A Key Step in Viral Protein 40 (VP40) Oligomerization and Viral Egress". eng. In: *The Journal of Biological Chemistry* 288.8, pp. 5779–5789.

Air, G. M. and W. G. Laver (1989). "The Neuraminidase of Influenza Virus". eng. In: *Proteins* 6.4, pp. 341–356.

Akarsu, Hatice et al. (Sept. 2003). "Crystal Structure of the M1 Protein-Binding Domain of the Influenza A Virus Nuclear Export Protein (NEP/NS2)". In: *The EMBO Journal* 22.18, pp. 4646–4655.

Ali, Ayub et al. (Sept. 2000). "Influenza Virus Assembly: Effect of Influenza Virus Glycoproteins on the Membrane Association of M1 Protein". In: *Journal of Virology* 74.18, pp. 8709–8719.

Amorim, Maria Joao et al. (May 2011). "A Rab11- and Microtubule-Dependent Mechanism for Cytoplasmic Transport of Influenza A Virus Viral RNA". eng. In: *Journal of Virology* 85.9, pp. 4143–4156.

Arzt, S. et al. (Jan. 2001). "Combined Results from Solution Studies on Intact Influenza Virus M1 Protein and from a New Crystal Form of Its N-Terminal Domain Show That M1 Is an Elongated Monomer". eng. In: *Virology* 279.2, pp. 439–446.

Arzt, Steffi et al. (Feb. 2004). "Structure of a Knockout Mutant of Influenza Virus M1 Protein That Has Altered Activities in Membrane Binding, Oligomerisation and Binding to NEP (NS2)". en. In: *Virus Research* 99.2, pp. 115–119.

Askjaer, P. et al. (Dec. 1998). "The Specificity of the CRM1-Rev Nuclear Export Signal Interaction Is Mediated by RanGTP". eng. In: *The Journal of Biological Chemistry* 273.50, pp. 33414–33422.

Avalos, R. T., Z. Yu, and D. P. Nayak (Apr. 1997). "Association of Influenza Virus NP and M1 Proteins with Cellular Cytoskeletal Elements in Influenza Virus-Infected Cells." en. In: *Journal of Virology* 71.4, pp. 2947–2958.

Bagchi, Biman (Mar. 2012). "From Anomalies in Neat Liquid to Structure, Dynamics and Function in the Biological World". en. In: *Chemical Physics Letters* 529, pp. 1–9.

Barman, Subrata et al. (Aug. 2001). "Transport of Viral Proteins to the Apical Membranes and Interaction of Matrix Protein with Glycoproteins in the Assembly of Influenza Viruses". en. In: *Virus Research* 77.1, pp. 61–69.

Batishchev, O. V. et al. (Jan. 2016). "pH-Dependent Formation and Disintegration of the Influenza A Virus Protein Scaffold To Provide Tension for Membrane Fusion". en. In: *Journal of Virology* 90.1, pp. 575–585.

Baudin, Florence et al. (Mar. 2001). "In Vitro Dissection of the Membrane and RNP Binding Activities of Influenza Virus M1 Protein". en. In: *Virology* 281.1, pp. 102–108.

Bergmann, M. and T. Muster (Dec. 1995). "The Relative Amount of an Influenza A Virus Segment Present in the Viral Particle Is Not Affected by a Reduction in Replication of That Segment". eng. In: *The Journal of General Virology* 76 (Pt 12), pp. 3211–3215.

Bernal, J. D. and R. H. Fowler (Aug. 1933). "A Theory of Water and Ionic Solution, with Particular Reference to Hydrogen and Hydroxyl Ions". In: *The Journal of Chemical Physics* 1.8, pp. 515–548.

Bialas, Kristy M., Emily A. Desmet, and Toru Takimoto (Nov. 2012). "Specific Residues in the 2009 H1N1 Swine-Origin Influenza Matrix Protein Influence Virion Morphology and Efficiency of Viral Spread In Vitro". In: *PLoS ONE* 7.11.

Bialas, Kristy M. et al. (Feb. 2014). "Specific Nucleoprotein Residues Affect Influenza Virus Morphology". en. In: *Journal of Virology* 88.4, pp. 2227–2234.

Bonavia, Aurelio et al. (Nov. 2010). *Antibody Targeting of TSG101 on Influenza-Infected Cells*. English. https://www.dovepress.com/antibody-targeting-of-tsg101-on-influenza-infected-cells-peer-reviewed-article-VAAT.

Bornholdt, Zachary A. et al. (Aug. 2013). "Structural Rearrangement of Ebola Virus VP40 Begets Multiple Functions in the Virus Life Cycle". English. In: *Cell* 154.4, pp. 763–774.

Boulay, F. et al. (Sept. 1987). "The Influenza Hemagglutinin Precursor as an Acid-Sensitive Probe of the Biosynthetic Pathway". eng. In: *The EMBO journal* 6.9, pp. 2643–2650.

Brevnov, V. V., N. V. Fedorova, and A. V. Indenbom (Nov. 2016). "Formation of the Layer of Influenza A Virus M1 Matrix Protein on Lipid Membranes at pH 7.0". en. In: *Russian Chemical Bulletin* 65.11, pp. 2737–2744.

Brown, Deborah A. and John K. Rose (Feb. 1992). "Sorting of GPI-Anchored Proteins to Glycolipid-Enriched Membrane Subdomains during Transport to the Apical Cell Surface". en. In: *Cell* 68.3, pp. 533–544.

Bruce, Emily A., Paul Digard, and Amanda D. Stuart (June 2010). "The Rab11 Pathway Is Required for Influenza A Virus Budding and Filament Formation". en. In: *Journal of Virology* 84.12, pp. 5848–5859.

Brunotte, Linda et al. (July 2014). "The Nuclear Export Protein of H5N1 Influenza A Viruses Recruits Matrix 1 (M1) Protein to the Viral Ribonucleoprotein to Mediate Nuclear Export". en. In: *Journal of Biological Chemistry* 289.29, pp. 20067–20077.

Bui, M. et al. (Feb. 2000). "Role of the Influenza Virus M1 Protein in Nuclear Export of Viral Ribonucleoproteins". eng. In: *Journal of Virology* 74.4, pp. 1781–1786.

Bullough, P. A. et al. (Sept. 1994). "Structure of Influenza Haemagglutinin at the pH of Membrane Fusion". eng. In: *Nature* 371.6492, pp. 37–43.

Burleigh, L.M. et al. (2005). "Influenza A Viruses with Mutations in the M1 Helix Six Domain Display a Wide Variety of Morphological Phenotypes". In: *Journal of Virology* 79.2, pp. 1262–1270.

Calder, Lesley J. et al. (June 2010). "Structural Organization of a Filamentous Influenza A Virus". en. In: *Proceedings of the National Academy of Sciences* 107.23, pp. 10685–10690.

Calero, Carles, H. Eugene Stanley, and Giancarlo Franzese (Apr. 2016). "Structural Interpretation of the Large Slowdown of Water Dynamics at Stacked Phospholipid Membranes for Decreasing Hydration Level: All-Atom Molecular Dynamics". In: *Materials* 9.5.

Carrió, M. Mar, Rafael Cubarsi, and Antonio Villaverde (2000). "Fine Architecture of Bacterial Inclusion Bodies". en. In: *FEBS Letters* 471.1, pp. 7–11.

Chase, Geoffrey P. et al. (Sept. 2011). "Influenza Virus Ribonucleoprotein Complexes Gain Preferential Access to Cellular Export Machinery through Chromatin Targeting". eng. In: *PLoS pathogens* 7.9, e1002187.

Chen, Benjamin J. et al. (July 2007). "Influenza Virus Hemagglutinin and Neuraminidase, but Not the Matrix Protein, Are Required for Assembly and Budding of Plasmid-Derived Virus-like Particles". eng. In: *Journal of Virology* 81.13, pp. 7111–7123.

Chen, Benjamin J. et al. (Oct. 2008). "The Influenza Virus M2 Protein Cytoplasmic Tail Interacts with the M1 Protein and Influences Virus Assembly at the Site of Virus Budding". eng. In: *Journal of Virology* 82.20, pp. 10059–10070.

Chen, J., J. J. Skehel, and D. C. Wiley (Sept. 1998). "A Polar Octapeptide Fused to the N-Terminal Fusion Peptide Solubilizes the Influenza Virus HA2 Subunit Ectodomain". eng. In: *Biochemistry* 37.39, pp. 13643–13649.

Chiang, Meng-Jung et al. (Jan. 2017). "Maintaining pH-Dependent Conformational Flexibility of M1 Is Critical for Efficient Influenza A Virus Replication". In: *Emerging Microbes & Infections* 6.1, pp. 1–11.

Chlanda, Petr et al. (Apr. 2016). "The Hemifusion Structure Induced by Influenza Virus Haemagglutinin Is Determined by Physical Properties of the Target Membranes". en. In: *Nature Microbiology* 1.6, pp. 1–8.

Chrunyk, B. A. et al. (Aug. 1993). "Inclusion Body Formation and Protein Stability in Sequence Variants of Interleukin-1 Beta". eng. In: *The Journal of Biological Chemistry* 268.24, pp. 18053–18061.

Collins, K. D. (Jan. 1997). "Charge Density-Dependent Strength of Hydration and Biological Structure". en. In: *Biophysical Journal* 72.1, pp. 65–76.

Collins, Kim D. and Michael W. Washabaugh (Nov. 1985). "The Hofmeister Effect and the Behaviour of Water at Interfaces". en. In: *Quarterly Reviews of Biophysics* 18.4, pp. 323–422.

Connor, Robert J. et al. (Nov. 1994). "Receptor Specificity in Human, Avian, and Equine H2 and H3 Influenza Virus Isolates". en. In: *Virology* 205.1, pp. 17–23.

Conti, E. et al. (July 1998). "Crystallographic Analysis of the Recognition of a Nuclear Localization Signal by the Nuclear Import Factor Karyopherin Alpha". eng. In: *Cell* 94.2, pp. 193–204.

Conti, Elena and John Kuriyan (Mar. 2000). "Crystallographic Analysis of the Specific yet Versatile Recognition of Distinct Nuclear Localization Signals by Karyopherin α". en. In: *Structure* 8.3, pp. 329–338.

Couceiro, J. Nelson S. S., James C. Paulson, and Linda G. Baum (Aug. 1993). "Influenza Virus Strains Selectively Recognize Sialyloligosaccharides on Human Respiratory Epithelium; the Role of the Host Cell in Selection of Hemagglutinin Receptor Specificity". en. In: *Virus Research* 29.2, pp. 155–165.

Cray, Jonathan A. et al. (2013). "A Universal Measure of Chaotropicity and Kosmotropicity". en. In: *Environmental Microbiology* 15.1, pp. 287–296.

Dahmani, Ismail, Kai Ludwig, and Salvatore Chiantia (Aug. 2019). "Influenza A Matrix Protein M1 Induces Lipid Membrane Deformation via Protein Multimerization". en. In: *Bioscience Reports* 39.8.

Damodaran, Srinivasan (2015). "Water at Biological Phase Boundaries: Its Role in Interfacial Activation of Enzymes and Metabolic Pathways". eng. In: *Sub-Cellular Biochemistry* 71, pp. 233–261.

Darapaneni, Vivek (Jan. 2015). "Large-Scale Analysis of Influenza A Virus Sequences Reveals Universally Conserved Residues of Matrix Proteins". en. In: *American Journal of Current Microbiology* 3.1, pp. 1–13.

Das, Subash C. et al. (Feb. 2012). "The Highly Conserved Arginine Residues at Positions 76 through 78 of Influenza A Virus Matrix Protein M1 Play an Important Role in Viral Replication by Affecting the Intracellular Localization of M1". In: *Journal of Virology* 86.3, pp. 1522–1530.

Delaglio, F. et al. (Nov. 1995). "NMRPipe: A Multidimensional Spectral Processing System Based on UNIX Pipes". eng. In: *Journal of biomolecular NMR* 6.3, pp. 277–293.

Disalvo, E. (Jan. 2015). *Membrane Hydration: The Role of Water in the Structure and Function of Biological Membranes*. Vol. 71.

Disalvo, E. Anibal (n.d.). *Membrane Hydration - The Role of Water in the Structure and Function of Biological Membranes | E. Anibal Disalvo | Springer*. https://www.springer.com/gp/book/9783319190594.

Edinger, Thomas O., Marie O. Pohl, and Silke Stertz (2014). "Entry of Influenza A Virus: Host Factors and Antiviral Targets". In: *Journal of General Virology,* 95.2, pp. 263–277.

Elton, D. et al. (Jan. 2001). "Interaction of the Influenza Virus Nucleoprotein with the Cellular CRM1-Mediated Nuclear Export Pathway". eng. In: *Journal of Virology* 75.1, pp. 408–419.

Enami, M., R. Fukuda, and A. Ishihama (Apr. 1985). "Transcription and Replication of Eight RNA Segments of Influenza Virus". eng. In: *Virology* 142.1, pp. 68–77.

Engelhardt, Othmar G. and Ervin Fodor (2006 Sep-Oct). "Functional Association between Viral and Cellular Transcription during Influenza Virus Infection". eng. In: *Reviews in Medical Virology* 16.5, pp. 329–345.

Fischer, Utz et al. (Aug. 1995). "The HIV-1 Rev Activation Domain Is a Nuclear Export Signal That Accesses an Export Pathway Used by Specific Cellular RNAs". en. In: *Cell* 82.3, pp. 475–483.

Fisette, Olivier et al. (Sept. 2016). "Hydration Dynamics of a Peripheral Membrane Protein". In: *Journal of the American Chemical Society* 138.36, pp. 11526–11535.

Fodor, E. (2013). "The RNA Polymerase of Influenza a Virus: Mechanisms of Viral Transcription and Replication". eng. In: *Acta Virologica* 57.2, pp. 113–122.

Fontana, Juan and Alasdair C. Steven (May 2013). "At Low pH, Influenza Virus Matrix Protein M1 Undergoes a Conformational Change Prior to Dissociating from the Membrane". en. In: *Journal of Virology* 87.10, pp. 5621–5628.

Fontana, Juan et al. (Mar. 2012). "Structural Changes in Influenza Virus at Low pH Characterized by Cryo-Electron Tomography". en. In: *Journal of Virology* 86.6, pp. 2919–2929.

Fornerod, Maarten et al. (Sept. 1997). "CRM1 Is an Export Receptor for Leucine-Rich Nuclear Export Signals". en. In: *Cell* 90.6, pp. 1051–1060.

Fournier, Emilie, Vincent Moules, Boris Essere, Jean-Christophe Paillart, Jean-Daniel Sirbat, Annie Cavalier, et al. (Dec. 2012). "Interaction Network Linking the Human H3N2 Influenza A Virus Genomic RNA Segments". eng. In: *Vaccine* 30.51, pp. 7359–7367.

Fournier, Emilie, Vincent Moules, Boris Essere, Jean-Christophe Paillart, Jean-Daniel Sirbat, Catherine Isel, et al. (Mar. 2012). "A Supramolecular Assembly Formed by Influenza A Virus Genomic RNA Segments". In: *Nucleic Acids Research* 40.5, pp. 2197–2209.

Freed, Eric O. (May 2002). "Viral Late Domains". In: *Journal of Virology* 76.10, pp. 4679–4687.

— (Feb. 2003). "The HIV–TSG101 Interface: Recent Advances in a Budding Field". English. In: *Trends in Microbiology* 11.2, pp. 56–59.

Freedman, R. B. and R. Wetzel (Aug. 1992). "Protein Engineering". eng. In: *Current Opinion in Biotechnology* 3.4, pp. 323–325.

Fujiyoshi, Y. et al. (Jan. 1994). "Fine Structure of Influenza A Virus Observed by Electron Cryo-Microscopy." In: *The EMBO Journal* 13.2, pp. 318–326.

Fukuda, Makoto et al. (Nov. 1997). "CRM1 Is Responsible for Intracellular Transport Mediated by the Nuclear Export Signal". en. In: *Nature* 390.6657, pp. 308–311.

Furuse, Yuki et al. (May 2009). "Evolution of the M Gene of the Influenza A Virus in Different Host Species: Large-Scale Sequence Analysis". In: *Virology Journal* 6.1, p. 67.

Gao, Qinshan et al. (July 2012). "The Influenza A Virus PB2, PA, NP, and M Segments Play a Pivotal Role during Genome Packaging". eng. In: *Journal of Virology* 86.13, pp. 7043–7051.

Garrus, J. E. et al. (Oct. 2001). "Tsg101 and the Vacuolar Protein Sorting Pathway Are Essential for HIV-1 Budding". eng. In: *Cell* 107.1, pp. 55–65.

Gavazzi, Cyrille et al. (Jan. 2013). "An in Vitro Network of Intermolecular Interactions between Viral RNA Segments of an Avian H5N2 Influenza A Virus: Comparison with a Human H3N2 Virus". eng. In: *Nucleic Acids Research* 41.2, pp. 1241–1254.

Gc, Jeevan B. et al. (Oct. 2016). "The Ebola Virus Protein VP40 Hexamer Enhances the Clustering of PI(4,5)P2 Lipids in the Plasma Membrane". eng. In: *Physical chemistry chemical physics: PCCP* 18.41, pp. 28409–28417.

Ge, Peng et al. (Feb. 2010). "Cryo-EM Model of the Bullet-Shaped Vesicular Stomatitis Virus". eng. In: *Science (New York, N.Y.)* 327.5966, pp. 689–693.

Goh, Gerard Kian-Meng, A. Keith Dunker, and Vladimir N. Uversky (Oct. 2008). "A Comparative Analysis of Viral Matrix Proteins Using Disorder Predictors". en. In: *Virology Journal* 5.1, p. 126.

Gómez-Puertas, P. et al. (Dec. 2000). "Influenza Virus Matrix Protein Is the Major Driving Force in Virus Budding". eng. In: *Journal of Virology* 74.24, pp. 11538–11547.

Gregoriades, A (Nov. 1980). "Interaction of Influenza M Protein with Viral Lipid and Phosphatidylcholine Vesicles." In: *Journal of Virology* 36.2, pp. 470–479.

Gregoriades, A. and B. Frangione (Oct. 1981). "Insertion of Influenza M Protein into the Viral Lipid Bilayer and Localization of Site of Insertion." en. In: *Journal of Virology* 40.1, pp. 323–328.

Halder, Umesh Chandra et al. (May 2013). "Phosphorylation Drives an Apoptotic Protein to Activate Antiapoptotic Genes: Paradigm of Influenza A Matrix 1 Protein Function". eng. In: *The Journal of Biological Chemistry* 288.20, pp. 14554–14568.

Hamilton, Brian S., Gary R. Whittaker, and Susan Daniel (July 2012). "Influenza Virus-Mediated Membrane Fusion: Determinants of Hemagglutinin Fusogenic Activity and Experimental Approaches for Assessing Virus Fusion". In: *Viruses* 4.7, pp. 1144–1168.

Harris, A., B. Sha, and M. Luo (Apr. 1999). "Structural Similarities between Influenza Virus Matrix Protein M1 and Human Immunodeficiency Virus Matrix and Capsid Proteins: An Evolutionary Link between Negative-Stranded RNA Viruses and Retroviruses". eng. In: *The Journal of General Virology* 80 (Pt 4), pp. 863–869.

Harris, A. et al. (Oct. 2001). "The Crystal Structure of the Influenza Matrix Protein M1 at Neutral pH: M1-M1 Protein Interfaces Can Rotate in the Oligomeric Structures of M1". eng. In: *Virology* 289.1, pp. 34–44.

Harris, Audray et al. (Dec. 2006). "Influenza Virus Pleiomorphy Characterized by Cryoelectron Tomography". eng. In: *Proceedings of the National Academy of Sciences of the United States of America* 103.50, pp. 19123–19127.

Hatada, E. et al. (Apr. 1989). "Control of Influenza Virus Gene Expression: Quantitative Analysis of Each Viral RNA Species in Infected Cells". eng. In: *Journal of Biochemistry* 105.4, pp. 537–546.

Heiny, A. T. et al. (Nov. 2007). "Evolutionarily Conserved Protein Sequences of Influenza A Viruses, Avian and Human, as Vaccine Targets". en. In: *PLOS ONE* 2.11, e1190.

Hilsch, Malte et al. (Aug. 2014). "Influenza A Matrix Protein M1 Multimerizes upon Binding to Lipid Membranes". In: *Biophysical Journal* 107.4, pp. 912–923.

Hoenen, T. et al. (July 2010). "Oligomerization of Ebola Virus VP40 Is Essential for Particle Morphogenesis and Regulation of Viral Transcription". en. In: *Journal of Virology* 84.14, pp. 7053–7063.

Höfer, C. T. et al. (Jan. 2019). "Structural Determinants of the Interaction between Influenza A Virus Matrix Protein M1 and Lipid Membranes". eng. In: *Biochimica Et Biophysica Acta. Biomembranes* 1861.6, pp. 1123–1134.

Hom, Nancy et al. (July 2019). "Deep Mutational Scan of the Highly Conserved Influenza A Virus M1 Matrix Protein Reveals Substantial Intrinsic Mutational Tolerance". en. In: *Journal of Virology* 93.13.

Hribar, Barbara et al. (Oct. 2002). "How Ions Affect the Structure of Water". In: *Journal of the American Chemical Society* 124.41, pp. 12302–12311.

Huang, I.-Chueh et al. (May 2008). "Influenza A Virus Neuraminidase Limits Viral Superinfection". en. In: *Journal of Virology* 82.10, pp. 4834–4843.

Huang, Shengping et al. (Jan. 2013). "A Second CRM1-Dependent Nuclear Export Signal in the Influenza A Virus NS2 Protein Contributes to the Nuclear Export of Viral Ribonucleoproteins". en. In: *Journal of Virology* 87.2, pp. 767–778.

Hutchinson, Edward C. et al. (Feb. 2010). "Genome Packaging in Influenza A Virus". eng. In: *The Journal of General Virology* 91.Pt 2, pp. 313–328.

Hutchinson, Edward C. et al. (Nov. 2012). "Mapping the Phosphoproteome of Influenza A and B Viruses by Mass Spectrometry". en. In: *PLOS Pathogens* 8.11, e1002993.

Ito, Toshihiro et al. (Sept. 1998). "Molecular Basis for the Generation in Pigs of Influenza A Viruses with Pandemic Potential". In: *Journal of Virology* 72.9, pp. 7367–7373.

Ito, T et al. (Oct. 1991). "Evolutionary Analysis of the Influenza A Virus M Gene with Comparison of the M1 and M2 Proteins." In: *Journal of Virology* 65.10, pp. 5491–5498.

Iuliano, A. Danielle et al. (Mar. 2018). "Estimates of Global Seasonal Influenza-Associated Respiratory Mortality: A Modelling Study". eng. In: *Lancet (London, England)* 391.10127, pp. 1285–1300.

Kang, Beibei et al. (Mar. 2020). "Hofmeister Series: Insights of Ion Specificity from Amphiphilic Assembly and Interface Property". en. In: *ACS Omega* 5.12, pp. 6229–6239.

Keller, Rochus Leonhard Josef (2005). "Optimizing the Process of Nuclear Magnetic Resonance Spectrum Analysis and Computer Aided Resonance Assignment". en. Doctoral Thesis. ETH Zurich.

Kerviel, Adeline et al. (Nov. 2016). "Involvement of an Arginine Triplet in M1 Matrix Protein Interaction with Membranes and in M1 Recruitment into Virus-Like Particles of the Influenza A(H1N1)Pdm09 Virus". en. In: *PLOS ONE* 11.11, e0165421.

Knyazev, D. G., V. A. Radyukhin, and V. S. Sokolov (Mar. 2009). "Intermolecular Interactions of Influenza M1 Proteins on the Model Lipid Membrane Surface: A Study Using the Inner Field Compensation Method". en. In: *Biochemistry (Moscow) Supplement Series A: Membrane and Cell Biology* 3.1, pp. 81–89.

Kordyukova, L. V. et al. (2008). "Influenza A Virus M1 Protein Structure Probed by in Situ Limited Proteolysis with Bromelain". eng. In: *Protein and Peptide Letters* 15.9, pp. 922–930.

Kordyukova, Larisa V. et al. (Feb. 2019). "Matrix Proteins of Enveloped Viruses: A Case Study of Influenza A Virus M1 Protein". eng. In: *Journal of Biomolecular Structure & Dynamics* 37.3, pp. 671–690.

Kretzschmar, E., M. Bui, and J. K. Rose (June 1996). "Membrane Association of Influenza Virus Matrix Protein Does Not Require Specific Hydrophobic Domains or the Viral Glycoproteins". eng. In: *Virology* 220.1, pp. 37–45.

Ksenofontov, A. L. et al. (Aug. 2011). "Intrinsically Unstructured Regions in the C Domain of the Influenza Virus M1 Protein". en. In: *Molecular Biology* 45.4, p. 634.

Lakdawala, Seema S. et al. (Dec. 2016). "Correction: Influenza A Virus Assembly Intermediates Fuse in the Cytoplasm". en. In: *PLOS Pathogens* 12.12, e1006121.

Latham, T. and J. M. Galarza (July 2001). "Formation of Wild-Type and Chimeric Influenza Virus-like Particles Following Simultaneous Expression of Only Four Structural Proteins". eng. In: *Journal of Virology* 75.13, pp. 6154–6165.

Lee, K. H. and B. L. Seong (Aug. 1998). "The Position 4 Nucleotide at the 3' End of the Influenza Virus Neuraminidase vRNA Is Involved in Temporal Regulation of Transcription and Replication of Neuraminidase RNAs and Affects the Repertoire of Influenza Virus Surface Antigens". eng. In: *The Journal of General Virology* 79 (Pt 8), pp. 1923–1934.

Lee, Kelly K (Apr. 2010). "Architecture of a Nascent Viral Fusion Pore". In: *The EMBO Journal* 29.7, pp. 1299–1311.

Lee, M.-T. Michael et al. (Mar. 2003). "Activation of Influenza Virus RNA Polymerase by the 5' and 3' Terminal Duplex of Genomic RNA". eng. In: *Nucleic Acids Research* 31.6, pp. 1624–1632.

Leser, George P. and Robert A. Lamb (Nov. 2005). "Influenza Virus Assembly and Budding in Raft-Derived Microdomains: A Quantitative Analysis of the Surface Distribution of HA, NA and M2 Proteins". eng. In: *Virology* 342.2, pp. 215–227.

— (May 2017). "Lateral Organization of Influenza Virus Proteins in the Budozone Region of the Plasma Membrane". en. In: *Journal of Virology* 91.9.

Li, L. and S. N. Cohen (May 1996). "Tsg101: A Novel Tumor Susceptibility Gene Isolated by Controlled Homozygous Functional Knockout of Allelic Loci in Mammalian Cells". eng. In: *Cell* 85.3, pp. 319–329.

Li, Sai et al. (Apr. 2014). "pH-Controlled Two-Step Uncoating of Influenza Virus". In: *Biophysical Journal* 106.7, pp. 1447–1456.

Liljeroos, Lassi et al. (Nov. 2011). "Electron Cryotomography of Measles Virus Reveals How Matrix Protein Coats the Ribonucleocapsid within Intact Virions". eng. In:

Proceedings of the National Academy of Sciences of the United States of America 108.44, pp. 18085–18090.

Liu, Teresa and Zhiping Ye (Dec. 2002). "Restriction of Viral Replication by Mutation of the Influenza Virus Matrix Protein". In: *Journal of Virology* 76.24, pp. 13055–13061.

— (Sept. 2004). "Introduction of a Temperature-Sensitive Phenotype into Influenza A/WSN/33 Virus by Altering the Basic Amino Acid Domain of Influenza Virus Matrix Protein". en. In: *Journal of Virology* 78.18, pp. 9585–9591.

Lyubartsev, A. P., K. Laasonen, and A. Laaksonen (Feb. 2001). "Hydration of Li+ Ion. An Ab Initio Molecular Dynamics Simulation". In: *The Journal of Chemical Physics* 114.7, pp. 3120–3126.

Ma, K., A. M. Roy, and G. R. Whittaker (Apr. 2001). "Nuclear Export of Influenza Virus Ribonucleoproteins: Identification of an Export Intermediate at the Nuclear Periphery". eng. In: *Virology* 282.2, pp. 215–220.

MacDougall, F. H. (May 1937). "Ions in Solution. By R. W. Gurney". In: *The Journal of Physical Chemistry* 41.5, pp. 766–766.

Mancinelli, R. et al. (June 2007). "Perturbation of Water Structure Due to Monovalent Ions in Solution". en. In: *Physical Chemistry Chemical Physics* 9.23, pp. 2959–2967.

Marcolli, C. and Th Peter (June 2005). "Water Activity in Polyol/Water Systems: New UNIFAC Parameterization". English. In: *Atmospheric Chemistry and Physics* 5.6, pp. 1545–1555.

Marcus, Yizhak (Mar. 2009). "Effect of Ions on the Structure of Water: Structure Making and Breaking". In: *Chemical Reviews* 109.3, pp. 1346–1370.

Martin-Serrano, J., T. Zang, and P. D. Bieniasz (Dec. 2001). "HIV-1 and Ebola Virus Encode Small Peptide Motifs That Recruit Tsg101 to Sites of Particle Assembly to Facilitate Egress". eng. In: *Nature Medicine* 7.12, pp. 1313–1319.

Martin, Kelsey and Ari Heleniust (Oct. 1991). "Nuclear Transport of Influenza Virus Ribonucleoproteins: The Viral Matrix Protein (M1) Promotes Export and Inhibits Import". en. In: *Cell* 67.1, pp. 117–130.

Mazzé, F. M. and L. Degrève (2006). "The Role of Viral and Cellular Proteins in the Budding of Human Immunodeficiency Virus". eng. In: *Acta Virologica* 50.2, pp. 75–85.

McCauley, J W and B W Mahy (May 1983). "Structure and Function of the Influenza Virus Genome." In: *Biochemical Journal* 211.2, pp. 281–294.

McCown, Matthew F. and Andrew Pekosz (Aug. 2006). "Distinct Domains of the Influenza A Virus M2 Protein Cytoplasmic Tail Mediate Binding to the M1 Protein and Facilitate Infectious Virus Production". en. In: *Journal of Virology* 80.16, pp. 8178–8189.

Mercer, Jason, Mario Schelhaas, and Ari Helenius (2010). "Virus Entry by Endocytosis". In: *Annual Review of Biochemistry* 79.1, pp. 803–833.

Meyer, B. E. and M. H. Malim (July 1994). "The HIV-1 Rev Trans-Activator Shuttles between the Nucleus and the Cytoplasm". eng. In: *Genes & Development* 8.13, pp. 1538–1547.

Miyake, Yasuyuki et al. (Apr. 2019). "Influenza Virus Uses Transportin 1 for vRNP Debundling during Cell Entry". en. In: *Nature Microbiology* 4.4, pp. 578–586.

Momose, Fumitaka et al. (2011). "Apical Transport of Influenza A Virus Ribonucleoprotein Requires Rab11-Positive Recycling Endosome". eng. In: *PloS One* 6.6, e21123.

Mosier, Philip D. et al. (Aug. 2016). "Broad Spectrum Anti-Influenza Agents by Inhibiting Self-Association of Matrix Protein 1". en. In: *Scientific Reports* 6.1, p. 32340.

Muscatello, David J. et al. (2014). "Mortality Attributable to Seasonal and Pandemic Influenza, Australia, 2003 to 2009, Using a Novel Time Series Smoothing Approach". eng. In: *PloS One* 8.6, e64734.

Myers, Erin L. and Jane F. Allen (Nov. 2002). "Tsg101, an Inactive Homologue of Ubiquitin Ligase E2, Interacts Specifically with Human Immunodeficiency Virus Type 2 Gag Polyprotein and Results in Increased Levels of Ubiquitinated Gag". eng. In: *Journal of Virology* 76.22, pp. 11226–11235.

Nagle, J. F. and S. Tristram-Nagle (Nov. 2000). "Structure of Lipid Bilayers". eng. In: *Biochimica Et Biophysica Acta* 1469.3, pp. 159–195.

Nayak, Debi P., Eric Ka-Wai Hui, and Subrata Barman (Dec. 2004). "Assembly and Budding of Influenza Virus". en. In: *Virus Research*. Mechanisms of Enveloped Virus Release 106.2, pp. 147–165.

Nayak, Debi P. et al. (Aug. 2009). "Influenza Virus Morphogenesis and Budding". en. In: *Virus Research*. Virus Research - 25th Anniversary Issue 143.2, pp. 147–161.

Neumann, G., M. T. Hughes, and Y. Kawaoka (Dec. 2000). "Influenza A Virus NS2 Protein Mediates vRNP Nuclear Export through NES-Independent Interaction with hCRM1". eng. In: *The EMBO journal* 19.24, pp. 6751–6758.

Neumann, Gabriele and Yoshihiro Kawaoka (May 2015). "Transmission of Influenza A Viruses". In: *Virology* 0, pp. 234–246.

Noda, Takeshi et al. (Jan. 2012). "Three-Dimensional Analysis of Ribonucleoprotein Complexes in Influenza A Virus". en. In: *Nature Communications* 3.1, p. 639.

Noton, Sarah L. et al. (2007). "Identification of the Domains of the Influenza A Virus M1 Matrix Protein Required for NP Binding, Oligomerization and Incorporation into Virions". In: *Journal of General Virology,* 88.8, pp. 2280–2290.

O'Neill, R E, J Talon, and P Palese (Jan. 1998). "The Influenza Virus NEP (NS2 Protein) Mediates the Nuclear Export of Viral Ribonucleoproteins." In: *The EMBO Journal* 17.1, pp. 288–296.

Omta, Anne Willem et al. (Dec. 2003a). "Influence of Ions on the Hydrogen-Bond Structure in Liquid Water". In: *The Journal of Chemical Physics* 119.23, pp. 12457–12461.

— (July 2003b). "Negligible Effect of Ions on the Hydrogen-Bond Structure in Liquid Water". en. In: *Science* 301.5631, pp. 347–349.

Paragas, Jason et al. (Aug. 2001). "Influenza B and C Virus NEP (NS2) Proteins Possess Nuclear Export Activities". In: *Journal of Virology* 75.16, pp. 7375–7383.

Paterson, Duncan and Ervin Fodor (Dec. 2012). "Emerging Roles for the Influenza A Virus Nuclear Export Protein (NEP)". en. In: *PLOS Pathogens* 8.12, e1003019.

Perera, Aurélien (Oct. 2011). "On the Microscopic Structure of Liquid Water". In: *Molecular Physics* 109.20, pp. 2433–2441.

Perez, Omar D and Garry P Nolan (Nov. 2001). "Resistance Is Futile: Assimilation of Cellular Machinery by HIV-1". en. In: *Immunity* 15.5, pp. 687–690.

Peukes, Julia et al. (Sept. 2020). "The Native Structure of the Assembled Matrix Protein 1 of Influenza A Virus". en. In: *Nature*, pp. 1–4.

Plumridge, T. H. and R. D. Waigh (2002). "Water Structure Theory and Some Implications for Drug Design". en. In: *Journal of Pharmacy and Pharmacology* 54.9, pp. 1155–1179.

Pornillos, Owen et al. (Nov. 2002). "Structure of the Tsg101 UEV Domain in Complex with the PTAP Motif of the HIV-1 P6 Protein". en. In: *Nature Structural Biology* 9.11, pp. 812–817.

Redlberger-Fritz, Monika et al. (July 2012). "Attributable Deaths Due to Influenza: A Comparative Study of Seasonal and Pandemic Influenza". eng. In: *European Journal of Epidemiology* 27.7, pp. 567–575.

Reed, Carrie et al. (Mar. 2015). "Estimating Influenza Disease Burden from Population-Based Surveillance Data in the United States". en. In: *PLOS ONE* 10.3, e0118369.

Reinhardt, J. and T. Wolff (May 2000). "The Influenza A Virus M1 Protein Interacts with the Cellular Receptor of Activated C Kinase (RACK) 1 and Can Be Phosphorylated by Protein Kinase C". eng. In: *Veterinary Microbiology* 74.1-2, pp. 87–100.

Reuther, Peter et al. (Jan. 2014). "Adaptive Mutations in the Nuclear Export Protein of Human-Derived H5N1 Strains Facilitate a Polymerase Activity-Enhancing Conformation". en. In: *Journal of Virology* 88.1, pp. 263–271.

Rodriguez Boulan, E. and D. D. Sabatini (Oct. 1978). "Asymmetric Budding of Viruses in Epithelial Monlayers: A Model System for Study of Epithelial Polarity". eng. In: *Proceedings of the National Academy of Sciences of the United States of America* 75.10, pp. 5071–5075.

Rolfes, Melissa A. et al. (Jan. 2018). "Annual Estimates of the Burden of Seasonal Influenza in the United States: A Tool for Strengthening Influenza Surveillance and Preparedness". eng. In: *Influenza and Other Respiratory Viruses* 12.1, pp. 132–137.

Rossman, Jeremy S. and Robert A. Lamb (Mar. 2011). "Influenza Virus Assembly and Budding". In: *Virology* 411.2, pp. 229–236.

Rossman, Jeremy S. et al. (Sept. 2010). "Influenza Virus M2 Protein Mediates ESCRT-Independent Membrane Scission". eng. In: *Cell* 142.6, pp. 902–913.

Ruigrok, R. W. et al. (Feb. 2000). "Membrane Interaction of Influenza Virus M1 Protein". eng. In: *Virology* 267.2, pp. 289–298.

Ruigrok, Rob W. H., Lesley J. Calder, and Stephen A. Wharton (Nov. 1989). "Electron Microscopy of the Influenza Virus Submembranal Structure". en. In: *Virology* 173.1, pp. 311–316.

Ruigrok, Rob W. H et al. (June 2000). "Structural Characterization and Membrane Binding Properties of the Matrix Protein VP40 of Ebola virus11Edited by J. Karn". en. In: *Journal of Molecular Biology* 300.1, pp. 103–112.

Ruigrok, Rob et al. (Oct. 2001). "Role of Influenza Virus M1 Protein in the Viral Budding Process". en. In: *International Congress Series* 1219, pp. 397–404.

Safo, Martin K. et al. (2014). "Crystal Structures of Influenza A Virus Matrix Protein M1: Variations on a Theme". eng. In: *PloS One* 9.10, e109510.

Schanda, Paul, Hélène Van Melckebeke, and Bernhard Brutscher (July 2006). "Speeding Up Three-Dimensional Protein NMR Experiments to a Few Minutes". In: *Journal of the American Chemical Society* 128.28, pp. 9042–9043.

Scheiffele, P. et al. (Jan. 1999). "Influenza Viruses Select Ordered Lipid Domains during Budding from the Plasma Membrane". eng. In: *The Journal of Biological Chemistry* 274.4, pp. 2038–2044.

Schwenk, Christian F., Thomas S. Hofer, and Bernd M. Rode (Mar. 2004). ""Structure Breaking" Effect of Hydrated Cs+". In: *The Journal of Physical Chemistry A* 108.9, pp. 1509–1514.

Scianimanico, S. et al. (Dec. 2000). "Membrane Association Induces a Conformational Change in the Ebola Virus Matrix Protein". eng. In: *The EMBO journal* 19.24, pp. 6732–6741.

Selzer, Lisa et al. (Sept. 2020). "Full-Length Three-Dimensional Structure of the Influenza A Virus M1 Protein and Its Organization into a Matrix Layer". en. In: *PLOS Biology* 18.9, e3000827.

Sha, Bingdong and Ming Luo (Mar. 1997). "Structure of a Bifunctional Membrane-RNA Binding Protein, Influenza Virus Matrix Protein M1". en. In: *Nature Structural Biology* 4.3, pp. 239–244.

Shapiro, G. I., T. Gurney, and R. M. Krug (Mar. 1987). "Influenza Virus Gene Expression: Control Mechanisms at Early and Late Times of Infection and Nuclear-Cytoplasmic Transport of Virus-Specific RNAs". eng. In: *Journal of Virology* 61.3, pp. 764–773.

Shilova, L. A. et al. (July 2017). "Study of Adsorption of Influenza Virus Matrix Protein M1 on Lipid Membranes by the Technique of Fluorescent Probes". en. In: *Biochemistry (Moscow), Supplement Series A: Membrane and Cell Biology* 11.3, pp. 225–230.

Shimizu, Teppei et al. (Jan. 2011). "Crucial Role of the Influenza Virus NS2 (NEP) C-Terminal Domain in M1 Binding and Nuclear Export of vRNP". eng. In: *FEBS letters* 585.1, pp. 41–46.

Shishkov, Alexander V. et al. (July 1999). "The in Situ Spatial Arrangement of the Influenza A Virus Matrix Protein M1 Assessed by Tritium Bombardment". en. In: *Proceedings of the National Academy of Sciences* 96.14, pp. 7827–7830.

Shishkov, Alexander V. et al. (Oct. 2009). *The In Situ Structural Characterization of the Influenza A Virus Matrix M1 Protein within a Virion*. en. https://www.eurekaselect.com/85105/article.

Shishkov, Alexander et al. (Dec. 2011). "Spatial Structure Peculiarities of Influenza A Virus Matrix M1 Protein in an Acidic Solution That Simulates the Internal Lysosomal Medium". eng. In: *The FEBS journal* 278.24, pp. 4905–4916.

Shtykova, Eleonora V. et al. (Dec. 2013). "Structural Analysis of Influenza A Virus Matrix Protein M1 and Its Self-Assemblies at Low pH". en. In: *PLOS ONE* 8.12, e82431.

Shtykova, Eleonora V. et al. (Dec. 2017). "Influenza Virus Matrix Protein M1 Preserves Its Conformation with pH, Changing Multimerization State at the Priming Stage Due to Electrostatics". en. In: *Scientific Reports* 7.1, p. 16793.

Simons, K. and E. Ikonen (June 1997). "Functional Rafts in Cell Membranes". eng. In: *Nature* 387.6633, pp. 569–572.

Singh, Anupam et al. (Mar. 2015). "Protein Recovery from Inclusion Bodies of Escherichia Coli Using Mild Solubilization Process". In: *Microbial Cell Factories* 14.1, p. 41.

Smith, G. L. and A. J. Hay (Apr. 1982). "Replication of the Influenza Virus Genome". en. In: *Virology* 118.1, pp. 96–108.

Stahelin, Robert V. (June 2014). "Membrane Binding and Bending in Ebola VP40 Assembly and Egress". In: *Frontiers in Microbiology* 5.

Stauffer, Sarah et al. (Nov. 2014). "Stepwise Priming by Acidic pH and a High K+ Concentration Is Required for Efficient Uncoating of Influenza A Virus Cores after Penetration". eng. In: *Journal of Virology* 88.22, pp. 13029–13046.

Takeda, Makoto et al. (Dec. 2003). "Influenza Virus Hemagglutinin Concentrates in Lipid Raft Microdomains for Efficient Viral Fusion". eng. In: *Proceedings of the National Academy of Sciences of the United States of America* 100.25, pp. 14610–14617.

Taubenberger, Jeffery K. and John C. Kash (June 2010). "Influenza Virus Evolution, Host Adaptation and Pandemic Formation". In: *Cell host & microbe* 7.6, pp. 440–451.

Thaa, Bastian, Andreas Herrmann, and Michael Veit (Jan. 2009). "The Polybasic Region Is Not Essential for Membrane Binding of the Matrix Protein M1 of Influenza Virus". en. In: *Virology* 383.1, pp. 150–155.

Tielrooij, K. J. et al. (Nov. 2009). "Dielectric Relaxation Dynamics of Water in Model Membranes Probed by Terahertz Spectroscopy". en. In: *Biophysical Journal* 97.9, pp. 2484–2492.

Tu, Chun et al. (Sept. 2010). "Endosomal-Sorting Complexes Required for Transport (ESCRT) Pathway-Dependent Endosomal Traffic Regulates the Localization of Active Src at Focal Adhesions". en. In: *Proceedings of the National Academy of Sciences* 107.37, pp. 16107–16112.

Veit, Michael and Bastian Thaa (2011). "Association of Influenza Virus Proteins with Membrane Rafts". eng. In: *Advances in Virology* 2011, p. 370606.

VerPlank, L. et al. (July 2001). "Tsg101, a Homologue of Ubiquitin-Conjugating (E2) Enzymes, Binds the L Domain in HIV Type 1 Pr55(Gag)". eng. In: *Proceedings of the National Academy of Sciences of the United States of America* 98.14, pp. 7724–7729.

Vranken, Wim F. et al. (June 2005). "The CCPN Data Model for NMR Spectroscopy: Development of a Software Pipeline". eng. In: *Proteins* 59.4, pp. 687–696.

Wang, Dan et al. (May 2010). "The Lack of an Inherent Membrane Targeting Signal Is Responsible for the Failure of the Matrix (M1) Protein of Influenza A Virus To Bud into Virus-Like Particles". en. In: *Journal of Virology* 84.9, pp. 4673–4681.

Ward, A. C. et al. (1995). "Expression and Analysis of the NS2 Protein of Influenza A Virus". eng. In: *Archives of Virology* 140.11, pp. 2067–2073.

Watanabe, Ken et al. (Aug. 2001). "Inhibition of Nuclear Export of Ribonucleoprotein Complexes of Influenza Virus by Leptomycin B". en. In: *Virus Research* 77.1, pp. 31–42.

Weis, Karsten (Feb. 2003). "Regulating Access to the Genome: Nucleocytoplasmic Transport throughout the Cell Cycle". en. In: *Cell* 112.4, pp. 441–451.

White, Judith M. and Gary R. Whittaker (June 2016). "Fusion of Enveloped Viruses in Endosomes". eng. In: *Traffic (Copenhagen, Denmark)* 17.6, pp. 593–614.

Whittaker, G., M. Bui, and A. Helenius (May 1996). "Nuclear Trafficking of Influenza Virus Ribonuleoproteins in Heterokaryons". eng. In: *Journal of Virology* 70.5, pp. 2743–2756.

Whittaker, G, I Kemler, and A Helenius (Jan. 1995). "Hyperphosphorylation of Mutant Influenza Virus Matrix Protein, M1, Causes Its Retention in the Nucleus." In: *Journal of Virology* 69.1, pp. 439–445.

Wijesinghe, Kaveesha J. et al. (Apr. 2017). "Detection of Lipid-Induced Structural Changes of the Marburg Virus Matrix Protein VP40 Using Hydrogen/Deuterium Exchange-Mass Spectrometry". In: *The Journal of Biological Chemistry* 292.15, pp. 6108–6122.

Williams, D. C. et al. (Feb. 1982). "Cytoplasmic Inclusion Bodies in Escherichia Coli Producing Biosynthetic Human Insulin Proteins". en. In: *Science* 215.4533, pp. 687–689.

Xie, Hang et al. (Jan. 2013). "The Compensatory G88R Change Is Essential in Restoring the Normal Functions of Influenza A/WSN/33 Virus Matrix Protein 1 with a Disrupted Nuclear Localization Signal". en. In: *Journal of Virology* 87.1, pp. 345–353.

Yamaguchi, Masashi et al. (May 2008). "Zernike Phase Contrast Electron Microscopy of Ice-Embedded Influenza A Virus". eng. In: *Journal of Structural Biology* 162.2, pp. 271–276.

Yángüez, Emilio and Amelia Nieto (Mar. 2011). "So Similar, yet so Different: Selective Translation of Capped and Polyadenylated Viral mRNAs in the Influenza Virus Infected Cell". eng. In: *Virus Research* 156.1-2, pp. 1–12.

Yasuda, Jiro et al. (Sept. 1993). "Molecular Assembly of Influenza Virus: Association of the NS2 Protein with Virion Matrix". en. In: *Virology* 196.1, pp. 249–255.

Ye, Z. P. et al. (Feb. 1987). "Functional and Antigenic Domains of the Matrix (M1) Protein of Influenza A Virus." en. In: *Journal of Virology* 61.2, pp. 239–246.

Ye, Z., D. Robinson, and R. R. Wagner (Mar. 1995). "Nucleus-Targeting Domain of the Matrix Protein (M1) of Influenza Virus". eng. In: *Journal of Virology* 69.3, pp. 1964–1970.

Zangi, Ronen and B. J. Berne (Nov. 2006). "Aggregation and Dispersion of Small Hydrophobic Particles in Aqueous Electrolyte Solutions". In: *The Journal of Physical Chemistry B* 110.45, pp. 22736–22741.

Zhang, J., A. Pekosz, and R. A. Lamb (May 2000). "Influenza Virus Assembly and Lipid Raft Microdomains: A Role for the Cytoplasmic Tails of the Spike Glycoproteins". eng. In: *Journal of Virology* 74.10, pp. 4634–4644.

Zhang, Ke et al. (May 2012). "Dissection of Influenza A Virus M1 Protein: pH-Dependent Oligomerization of N-Terminal Domain and Dimerization of C-Terminal Domain". en. In: *PLOS ONE* 7.5, e37786.

Zhang, Ke et al. (2015). "Two Polar Residues within C-Terminal Domain of M1 Are Critical for the Formation of Influenza A Virions". en. In: *Cellular Microbiology* 17.11, pp. 1583–1593.

Zhang, Wenting, Chengzhi Cai, et al. (Feb. 2017). "Subcellular Localization and Interactions of Infectious Salmon Anemia Virus (ISAV) M1 and NEP as Well as Host Hsc70". In: *Virology Journal* 14.1, p. 30.

Zhang, Wenting, Wenjie Zheng, et al. (Aug. 2017). "Crystal Structure of an Orthomyxovirus Matrix Protein Reveals Mechanisms for Self-Polymerization and Membrane Association". en. In: *Proceedings of the National Academy of Sciences* 114.32, pp. 8550–8555.

Zhirnov, O. P. (May 1990). "Solubilization of Matrix Protein M1/M from Virions Occurs at Different pH for Orthomyxo- and Paramyxoviruses". eng. In: *Virology* 176.1, pp. 274–279.

— (Jan. 1992). "Isolation of Matrix Protein M1 from Influenza Viruses by Acid-Dependent Extraction with Nonionic Detergent". en. In: *Virology* 186.1, pp. 324–330.

Zhirnov, Oleg P., Mine R. Ikizler, and Peter F. Wright (Sept. 2002). "Cleavage of Influenza a Virus Hemagglutinin in Human Respiratory Epithelium Is Cell Associated and Sensitive to Exogenous Antiproteases". eng. In: *Journal of Virology* 76.17, pp. 8682–8689.

Zhu, Haijian and David J. W. Grant (Aug. 1996). "Influence of Water Activity in Organic Solvent + Water Mixtures on the Nature of the Crystallizing Drug Phase. 2. Ampicillin". en. In: *International Journal of Pharmaceutics* 139.1, pp. 33–43.

Zou, Qin et al. (Feb. 2002). "The Molecular Mechanism of Stabilization of Proteins by TMAO and Its Ability to Counteract the Effects of Urea". In: *Journal of the American Chemical Society* 124.7, pp. 1192–1202.