Book Section:
Muniyandi, S, Pangratiou, G, Edwards, TA orcid.org/0000-0002-1467-3674 et al. (1 more author) (2018) Structure and Function of the Human Respiratory Syncytial Virus M2–1 Protein. In: Harris, JR and Bhella, D, (eds.) Virus Protein and Nucleoprotein Complexes. Subcellular Biochemistry, 88 . Springer Singapore , Singapore , pp. 245-260.

https://doi.org/10.1007/978-981-10-8456-0_11

Title: Structure and function of the human respiratory syncytial virus M2-1 protein

Keywords: HRSV; M2-1; transcription factor; anti-termination; structure; function, antivirals.

Abstract.

Human respiratory syncytial virus (HRSV) is a non-segmented negative stranded RNA virus and is recognized as the most important viral agent of lower respiratory tract infection worldwide, responsible for up to 199,000 deaths each year. The only FDA-approved regime to prevent HRSV-mediated disease is pre-exposure administration of a humanized HRSV-specific monoclonal antibody, which although being effective, is not in widespread usage due to its cost. No HRSV vaccine exists and so there remains a strong need for alternative and complementary anti-HRSV therapies. The HRSV M2-1 protein is a transcription factor and represents an attractive target for the development of antiviral compounds, based on its essential role in the viral replication cycle. To this end, a detailed analysis of M2-1 structure and functions will aid in identifying rational targets for structure-based antiviral drug design that can be developed in future translational research. Here we present an overview of the current understanding of the structure and function of HRSV M2-1, drawing on additional information derived from its structural homologues from other related viruses.

1.1. Introduction

RNA viruses can be divided into positive and negative sense groups based on the ability of their RNA genomes to act as messenger RNA (mRNA) for the production of proteins. In positive sense RNA viruses, the genome can be translated directly, while in negative sense RNA viruses, the input genome must undergo a copying event to produce coding sense mRNAs. The negative-sense RNA viruses (NSV) can be further divided into segmented (SNSV) and non-segmented negative strand viruses (NSNSV), depending on whether the genome is a single chain of ribonucleotides, or whether it is segmented into two or more separate RNA molecules. The group of NSNSVs encompasses pathogens of humans, animals and plants, and notable examples include rabies virus, Nipah virus, Ebola virus (EBOV) and human respiratory syncytial virus (HRSV), many of which are without effective preventative or therapeutic options to prevent disease (Palse et al. 1996; Tao and Ye 2010).

HRSV is the leading cause of lower respiratory tract illness in infants, elderly and immunocompromised individuals, causing bronchiolitis and pneumonia. Early and recurrent infections have been linked to the development of asthma in later life, which has a significant economic burden (Bohmwald et al. 2016). It has been estimated that all infants across the globe are infected with HRSV at least once during their first 2 years of life. Mild infections have a recovery time of 1-2 weeks, however HRSV infection can be serious and often fatal, and estimates for the number of fatalities per year are put between 50,000 - 199,000 (Scheltema et al. 2017; Shi et al, 2017). Presently, no vaccine exists for HRSV and the only options to reduce HRSV-mediated disease is preventative administration of a humanized mouse monoclonal antibody directed to a neutralizing epitope on the fusion protein (F) preventing host cell entry. This molecule is known as Palivizumab (or Synagis), and is effective, providing short-term protection to infants at high risk (Brady, 2015), although the cost of this treatment is a current barrier to its widespread usage (Hu and Robinson, 2010). Post exposure treatment for HRSV is limited to the FDA-approved administration of nebulized ribavirin (a guanine nucleotide analogue). However, this drug exhibits significant toxicity and its effectiveness has been questioned (Marcelin et al., 2014). Thus, there is an urgent need to develop new antivirals against HRSV.

This chapter describes the structure and function of the HRSV M2-1 protein, which is an essential transcription factor required for the synthesis of HRSV mRNAs. Determination of the high-resolution structure of M2-1, in combination with defining the structural basis of its function, could aid in the design of new antiviral compounds through structure-based drug design.

1.2. The Virion

HRSV is classified in the Pneumoviridae family of the order Mononegavirales. HRSV virions are pleiomorphic, with a cell derived lipid membrane surrounding an internal ribonucleocapsid (RNP) complex comprising the HRSV RNA genome in association with multiple viral proteins. The genome is approximately 15,200 nt in length and includes 10 genes coding for 11 proteins (Figure 1.1). Of these 11 proteins, 3 are associated with the RNP; the nucleoprotein (N), the RNA-dependent RNA polymerase (L), and the phoshoprotein (P). Three HRSV proteins are associated with the viral envelope, namely the fusion protein (F), the so-called attachment glycoprotein (G) and the small hydrophobic protein (SH) (Table 1). The matrix protein (M), is

situated below the viral envelope, and plays a pivotal role in virion assembly. NS1 and NS2 proteins are non-structural, being absent from the virion, and they play critical roles in overcoming the innate immune defenses of the infected host cell. The remaining two proteins are encoded by the M2 gene, and are translated from overlapping reading frames on a single M2 mRNA transcript. The M2-1 protein was initially characterized as a matrix protein that associated with the viral envelope (Collins and Wertz, 1985), but is currently described as an essential transcription factor (Collins et al, 1995; Collins et al, 1996), and may also provide a role in virion assembly (Li et al, 2008; Kiss et al, 2014). The M2-2 protein has been shown to influence the activity of the viral polymerase (Table 1.1) (Bermingham and Collins, 1999), and has been shown to be dispensible for virus multiplication, although viruses in which the M2-2 reading frame is not accessible are growth attenuated.

1.3. Transcription in HRSV

Common with all NSNSVs, the gene expression program of HRSV begins with the input negative sense RNA genome being copied to produce mRNA transcripts in a process known as primary transcription. This activity is mediated by the resident polymerase, brought into the infected cell with the input negative sense genome. These transcripts are subsequently translated into either structural proteins that are needed for subsequent rounds of RNA synthesis and for assembly into viral progeny, or they are non-structural proteins that modify the host cell environment facilitating virus multiplication. At a later time point, the input genome is replicated to form positive sense copies of the input genome, known as anti-genomes, and the events that mediate this switch in polymerase activity from a transcriptase to a replicase are not yet resolved. The promoters for both these activities are located within the initial 44 nucleotides of the genome 3' end, but the start sites for these district activities are different; replication begins at position 1, whereas transcription is initiated alongside position 3 (Tremaglio et al, 2013). Subsequently, the transcribing polymerase moves along the template in the 3' to 5' direction and recognizes and responds to gene start and gene end signals that flank each of the HRSV genes. These conserved sequences signal the initiation and termination of a single 5' capped and a 3' poly (A) tailed mRNA from each of the 10 HRSV genes (Kuo et al, 1996). In contrast, the replicating polymerase ignores these gene start and gene end signals, to synthesise a

complementary copy of the genome, that is enwrapped in nucleoprotein (N protein) and subsequently recruits a polymerase complex to form a helical RNP assembly.

Studies using truncated HRSV genomes in which RNP complexes are reconstituted in cells by supplying the protein and RNA components in trans have shown that replication requires P and L proteins, whereas transcription requires P, L and also M2-1 (Yu et al, 1995; Collins et al, 1995; Collins et al, 1996). In the absence of M2-1, the abundance of full length transcription products is reduced, suggesting that the L protein may be poorly processive, especially during the transcription of long genes (Fearns and Collins, 1999). Inclusion of M2-1 in the transcription assay increases the abundance of full length transcripts, and taken together, these observations have led to the proposal that M2-1 is a processivity factor, allowing the RdRp to reach the end of each transcriptional unit (Collins et al, 1995). In addition, the presence of M2-1 in reconstitution assays appears to influence the ability of the polymerase to terminate transcription at the HRSV conserved gene end signals, leading to increased synthesis of multi-cistronic mRNAs spanning two or more transcriptional units (Hardy and Wertz, 1998). Whether the polymerase processivity and anti-termination activities of M2-1 are related has not been resolved; it is possible that increased polymerase processivity is observed when M2-1 prevents the polymerase from terminating at sequences resembling gene ends, spuriously located within genes (Sutherland et al, 2001).

More recently, a post-transcriptional role of M2-1 has recently been proposed (Rincheval et al, 2017), as well as recent reports that M2-1 is found in association with M and RNPs within the virion, thus implicating a structural role of M2-1 in virion assembly (Kiss et al, 2014).

Attempts to rescue infectious HRSV from cDNAs from which the M2-1 gene has been omitted have been unsuccessful (Collins et al, 1995), leading to the proposal that M2-1 is essential for virus multiplication, a proposition that is consistent with its apparent and critical role in ensuring full length transcription of all 10 HRSV mRNAs. However, the possibility that the requirement for M2-1 coding capacity is for one of the alternative non-transcription functions cannot be eliminated. Regardless, the requirement of M2-1 for the HRSV lifecycle means it represents a promising target for the design of antiviral molecules that disrupt one or more of its proposed functions, and thus block HRSV multiplication and thus disease.

1.4. Macromolecular interactions involved in M2-1 mediated antitermination

M2-1 is soluble as a tetramer and directly interacts with RNA and with P in a competitive manner (Blondot et al, 2012). It is not clear whether this dual interaction with RNA and P is strictly mutually exclusive, or can occur simultaneously on different monomers within the tetramer. P is also a tetramer, and residues within its N-terminus (residues 1-28) interact with N, supplying RNA-free N monomers to encapsidate the newly synthesized replication products in the form of RNPs (Renner et al, 2016). This chaperone activity of P maintains N in a soluble form that otherwise has a tendency to aggregate as helical assemblies, possibly through interacting with host cell RNAs. The C-terminal domain of P, particularly the last 9 residues, interacts with N proteins that are components of the viral RNPs, rather than free N monomers (Tran et al, 2007). Just N-terminal, proximal to this N-interaction region of P, lies the L-protein interacting site (residues 216-239, with crucial hydrophobic residues at 216, 223 and 227) (Sourimant et al, 2015).

The aforementioned proteins, N, P, L and M2-1 form the minimal requirement for the in vitro reconstitution of the HRSV transcription machinery. The variable, overlapping and multivalent molecular interactions involved in formation of the transcriptase and replicase complexes, as well as the large size, represents a significant bottleneck in the in vitro stoichiometric reconstitution and structural analysis of this assembly. All of these proteins, partly based on structural and sequence analysis, appear to have rigid domains connected by hinges facilitating domain movements (Tanner et al, 2014; Tawar et al, 2009; Liang et al, 2015). This can be expected to yield a megadalton complex that is inherently flexible thereby resulting in conformational heterogeneity in the final assembly, a problem traditionally difficult to address via structural biology. Detailed structural understanding of the molecular interactions involved in HRSV anti-termination would reveal details of the inter-molecular interfaces that could be used for structure-based inhibitor design of antiviral therapeutics, in addition to expanding our understanding of the process.

Functional and structural homologues of the above-mentioned replicase and transcriptase complex associated proteins exist in all other NSNSVs, which also share a broadly similar transcriptional programme. However, it is interesting to note that few NSNSVs possess a separate protein product that acts as a transcription factor, with only the various M2-1 proteins expressed by the pneumoviruses and VP30 expressed by the filoviruses, including Ebola virus (EBOV). Like M2-1, VP30 appears to be dispensible

for replication but essential for RdRp transcriptional activities, which is dependent on VP30 phosphorylation at specific sites (Biedenkopf et al, 2016; Lier et al, 2017). Presumably, the many other viruses that do not express such a protein either have incorporated protein modules that possess similar functions elsewhere, or utilize a transcription mechanism that has no such requirement. In any case, similarities between M2-1 and VP30 at both the structural and functional levels mean that the knowledge gained from understanding M2-1 structure and function will have important benefits for the other viruses within Pneumoviridae and Filoviridae families.

1.5. Crystal structure of HRSV M2-1 protein

1.5.1. Overexpression and purification of M2-1 for structural studies

Full length HRSV (strain A2) M2-1 (194 residues) has a requirement for zinc ions during high level expression of folded protein. M2-1 crystallized in two different space groups under conditions containing polyethylene glycol and cadmium ions. X-ray diffraction data was collected and the crystal structure of M2-1 was solved to 2.5 Å using the anomalous signal from the cadmium ions and from the bound zinc atoms (Tanner et al, 2014).

1.5.2. Molecular architecture of M2-1

Monomeric M2-1 consists of three distinct regions linked by unstructured flexible sequences (Figure 1.2a). These are the zinc binding domain (ZBD: residues 7-25), a tetramerization helix (residues 32-49) and the core domain (residues 69-172). Connecting the tetramerization helix and the core domain is a flexible linker (residues 52-67) that is poorly resolved in the crystal structure, and which includes two sites that can be phosphorylated (S58 and S61) with important consequences relating to the function of M2-1 as a transcription factor. The last 20 residues of the C-terminus are not resolved in the electron density as they are unstructured, and many of these residues are dispensable, as determined by the rescue of infectious HRSV with these residues deleted from the M2-1 ORF. Overall, the molecule has 9 alpha helices and no beta strands (Figure 1.2a). Each monomer of M2-1 interacts with other protomers forming a highly stable tetramer.

M2-1 tetramerisation is mediated by the oligomerization helix (residues 32-49, α1) that buries a series of hydrophobic residues (L36, L43, I46 and M50) on one helix

face within a four-helix bundle. There is also extensive interaction between the ZBD of one monomer within the NTD with the core of an adjacent monomer, significantly increasing the buried surface area in the tetramer and consequently increasing its predicted stability. In the context of the tetramer, the ZBD lies on the N-terminal face of the molecule, in close proximity to the RNA binding surface (Figure 1.2b), that also includes residues R3 and R4 from the extreme N-terminus of M2-1.

The crystal structure of M2-1 provides a framework for the rational analysis of residues involved in direct M2-1 interactions with RNA, P and M. In addition, the structure will provide a framework on which to explore the mechanisms by which M2-1 performs its various assigned or proposed functions, as well as understand how these multiple functions may be regulated.

1.5.3. RNA binding by M2-1

Electrostatic potential calculations on M2-1 revealed that its surface is saturated by positively charged residues (Figure 1.3) that form four prominent tracts on the tetramer. In order to charaterise potential RNA ligands, and thus provide further information regarding the mechanism of M2-1 function, quantitative RNA binding studies were performed using fluorescence anisotropy (FA) with different RNA sequences. These included polyribonucleotides of A, C, G and U, as well as RNAs representing various positive and negative sense HRSV gene-end sequences. These sequences were chosen as all functional studies performed to date reveal M2-1 is active using mini-genomes in which all HRSV-specific sequences have been eliminated, excepting conserved gene start and end sequences. The highest binding affinity of all was determined for the polyribonucleotide poly(A) (K_d = 19.1 nM), and comparison of the binding specificities of HRSV gene end sequences indicated that positive sense sequences exhibited consistently higher affinities than negative sense complements, but RNA affinities broadly corresponded with their A content. Variation of up to 5-fold were found in M2-1 binding affinity to the different gene-end sequences tested here, with SH gene-end binding the tightest (K_d = 46 nM). These data provide evidence that the function of M2-1 involves binding RNA sequences that are located at the 3' end of HRSV mRNAs, and this location is consistent with its proposed role in influencing the process of transcription termination.

1.5.4. Mapping M2-1 residues that interact with RNA.

With M2-1:RNA interactions established through FA and identification of high-affinity RNA ligands, the contribution of specific M2-1 residues towards RNA binding was examined through a variety of M2-1 mutants. These were selected based on the tetramer electrostatistics, and the previous NMR based RNA-binding analysis of the M2-1 globular core (58-177) (Blondot et al, 2012) (Figure 1.4). Multiple residues that lie within the core resulted in the reduced RNA binding affinity, including S58D/S61D, K92A, K92D, K150A, R151A, R151D, K150A/R151A and K159A. In addition, residues R3 and R4 that lie within the N-terminal tip of M2-1 were also found to contribute to RNA binding, showing that residues involved in this activity are not restricted only to the core, but extended outside.

The S58D/S61D double mutant represents a mimic of phosphorylated M2-1, and these two serine residues are located above the positively charged surface in the 3D structure of M2-1, in a position that is consistent with the observed influence over RNA binding. The reduced affinity of this phosphomimetic mutant for RNA suggests that M2-1 phosphorylation during the infectious cycle may modulate binding affinity to target RNAs, thereby potentially affecting M2-1 function. Comparison between the crystal structures of native M2-1 and the phosphomimetic M2-1 mutant S58D/S61D suggests that phosphorylation does not cause extensive structural rearrangements within the M2-1 tetramer, thus suggesting these functional changes may be due to alterations in surface electrostatic potential, and possibly its interaction with an RNA binding partner.

1.5.5. Identification of M2-1 residues involved in transcription factor activity

The aforementioned residues that influence RNA binding were assessed for their contribution towards M2-1 antitermination activity using a bicistronic HRSV minigenome (Tanner et al, 2014), in which two transcription units were separated by functional gene start and gene end sequences, forming a gene junction. This minigenome transcription system was set up so that the expression of luciferase from the downstream gene was dependent on the presence of M2-1, with the implication being that M2-1 was required in order for the polymerase to complete transcription of the long upstream gene, and thus gain access to the transcription start site of the downstream luciferase gene. Minigenome analysis found that all previously identified mutants that exhibited a reduction in RNA binding also showed reduced luciferase

expression activity in cellulo, thus suggesting that the RNA-binding activity of M2-1 was tied to its ability to act as a transcription factor.

1.6 HRSV M2-1 structure: Similarities and differences to other known related structures

The four-helix bundle formed by the oligomerization helix strongly favors the tetrameric organization of M2-1, with the multi-protomer contacts of the ZBD increasing oligomer stability. Comparison of the M2-1 X-ray structure with the NMR structure of the M2-1 core domain (Blondot et al, 2012) (residues 58-177) reveals a similar core fold, suggesting that the core domain can fold independently in the absence of all other M2-1 residues, although residues outside the core contribute to RNA binding as described above (Blondot et al, 2012). Superposition of the full-length M2-1 structure and the EBOV VP30 C-terminal domain shows structural similarities, with helices 3, 6 and 7 aligning well to each other, while helices 4 and 5 share the same orientation (Figure 1.5). This close structural alignment implies strong evolutionary ties, and also raises the possibility that these polymerase accessory proteins may act through common mechanisms. As described above, the EBOV VP30 is one of only a small number of polymerase accessory proteins within the broad group of mononegaviruses, and like M2-1 it possess a ZBD and its function is modulated by dynamic phosphorylation (Lier et al, 2017). For both M2-1 and VP30, dynamic reversible phosphorylation is required for their respective transcription factor functionality, with VP30 phospho-ablatant mutants being inactive (Biedenkopf et al, 2016).

The Cys3-His ZBD of M2-1 is noncanonical with no structural homologs found by the DALI server (Figure 1.1a), and similarity to the VP30 ZBD is unknown due to the current lack of a VP30 ZBD structural model. Sequence searches suggest that the M2-1 ZBD is related to the ZBD of Nup475, which was also shown to bind cadmium (Amann et al, 2003).

From the crystal structure of M2-1 it is possible to visualize a track that bound RNA might take across the surface of each monomer, linking residues that are known to influence RNA binding, and also encompassing a contiguous region of positive charge. RNA and P have overlapping binding sites on the M2-1 surface and within one protomer, thus binding is mutually exclusive. M2-1 binds P in a 1:1 stoichiometry at high affinity suggesting that each M2-1 tetramer potentially binds to a P-tetramer.

However, the possibility that each monomer within a tetramer might be occupied by either RNA or P ligands cannot be ruled out.

M2-1 from the closely related human metapneumovirus in the Pneumoviridae family shares 38% sequence homology with HRSV M2-1 (Leyrat et al, 2016). HMPV M2-1 is also tetrameric with an overall domain organization that is the same as that of HRSV M2-1. The final ten residues in the HMPV M2-1 C-terminus were not resolved in the structure, suggesting it is disordered, as reported for HRSV M2-1. Interestingly, in the crystal structure of HMPV M2-1, one of the monomers has its helical core domain flipped out from the rest of the molecule by 60 Å when compared to symmetrically arranged protomers in the HRSV M2-1 subunit. This structural plasticity in HMPV M2-1 was further explored by molecular dynamics simulations and small angle X-ray scattering measurements, which were consistent with the possibility that HMPV M2-1 can adopt both open and closed conformations in solution, with dynamic exchange between the two states, with the oligomerization domain as the most rigid part of the structure. This flexibility and domain swinging in HMPV M2-1 was postulated to increase the surface area of the molecule and allow sampling of a larger volume in solution that helps to trap RNA fragments, first proposed through a 'fly-casting' mechanism (Shoemaker et al, 2000).

HMPV M2-1 was also crystallized in the presence of adenosine monophosphate (AMP) and a five nucleotide long DNA fragment. Although these two molecules are non-native binding partners of M2-1, their location on the M2-1 surface corresponded to the RNA binding site determined for HRSV M2-1 from NMR and mutagenesis studies. A model describing the HPMV M2-1 RNA binding site proposes that the RNA may bind along an extended and contiguous tract of positive surface charge, extending between the N-terminal and C-terminal faces of each monomer. The length of this surface is consistent with binding a oligoribonucleotide of 13 nucleotides in length, which coincides with the length of the HRSV gene end consensus sequence.

1.7. A role for M2-1 in virion assembly?

The RNP within the HRSV virion comprises the minimal essential components required for transcription namely RNA, N, P, L and M2-1. This RNP is enclosed within the lipid bilayer with a matrix protein (M) in between them, and this assembly appears to be mediated by M2-1:M interactions. In a study involving super-resolution microscopy of filamentous HRSV particles isolated from infected cells, the M2-1 and

M proteins were localized within the enveloped virion with M2-1 being closer to the genomic RNA than M (Kiss et al, 2014). Thus, M protein, by simultaneously interacting with two chemically different entities (M2-1 and the phospholipid bilayer), appears to act as a molecular scaffold. This overall plan of virion architecture is maintained among other members of the Order Mononegavirales, such as those classified in the Rhabdoviridae and Filoviridae families, where the corresponding matrix enwraps the RNP. However, the resulting RNP and virion morphologies are less consistent; in the case of the rhabdoviruses, virions are bullet shaped and the RNP core is found as a highly regular and rigid helical assembly, surrounded by a matrix protein sheath. In the case of filoviruses (e.g. EBOV) the RNPs are helical, but exhibit a more extended morphology that is consistent with the filamentous shape of EBOV virions, which suggest RNP flexibility.

HRSV M appears able to adopt multiple different oligomeric states in solution, although it was crystallized as a monomer (Money et al, 2009). M has N- and C-terminal domains connected by a 13-residue region, and is globally similar to EBOV VP40 in fold, but different in topology. M has a distinct negatively-charged lobe at the N-terminus and a positively charged area of 600 $Å^2$ extending from the NTD to CTD. The N-terminus of the M protein directly interacts with M2-1, and these two regions are proposed to be responsible for simultaneously interacting with M2-1 and the lipid bilayer, consistent with its role as a virion scaffold. This arrangement would presumably require M2-1 to form a two dimensional layer below M, and also in contact with the RNP. From the crystal structure of HRSV M2-1, it is currently unclear how the tetramer can arrange into a planar lattice along with M, due to its tendency to multimerize as a highly stable tetramer. Whether the alternative conformation of the HMPV M2-1 protein could adopt such as layer, with one core domain per tetramer being extended, is an interesting possibility.

1.8. Other M2-1 functions.

Following entry into cells, many NSNSVs are reported to induce the formation of dense cytoplasmic structures that contain viral components, and in HRSV infected cells these sites are known as Inclusion Bodies (IB) (Norrby et al, 1970). Functionally analogous structures have been identified in many other NSNSVs including rhabdoviruses and filoviruses (Lahaye et al, 2009; Dolnik et al, 2015; Baskerville et al, 1985) and they are often used as histological markers for confirming viral infections. Using a combination

of state-of-the-art fluorescence microscopy, super-resolution microscopy and pulse-chase techniques, HRSV IBs have been shown to represent the sites of active viral RNA synthesis (Rinchavel et al, 2017), where both transcription and replication take place. Rather than being amorphous aggregates, IBs have a complex organization, and contain a spherical sub-structure named as the 'Inclusion Body Associated Granule' (IBAG). These are found to be enriched with nascent mRNA and also M2-1, but devoid of other proteins of the replication complex (N, P and L).

IBAG formation is strictly dependent on viral RNA synthesis and these structures are dynamic, undergoing continuous assembly and disassembly cycles, as inferred from time-lapse fluorescence studies. The IBAGs were not found to be translationally active, suggesting that they may represent a transient compartment where newly synthesized viral mRNAs that are associated with M2-1 are stored prior to translation in the cytosol. While the interaction between M2-1 and mRNAs is consistent with the previous RNA binding studies described above, that identified highest affinities for poly(A) sequences, the established strong association of M2-1 with nascent viral RNA in IBAGs may reflects a post-transcriptional function of M2-1 that is novel and needs to be explored.

1.9. Conclusions and future outlook

Solving the X-ray crystal structure of M2-1 represents an important step towards understanding the mechanism behind the complex functions of M2-1. However, complete clarity of M2-1 functions has yet to be achieved, and M2-1 does not operate in isolation from other viral components. Multiple protein-protein and protein-nucleic acid interactions execute the anti-termination process, and high resolution structures of these other components will greatly aid in revealing the molecular mechanisms behind M2-1 activities. Though homology modeling and other theoretical methods can help to some extent, accurate understanding requires experimentally determined 3D structures. More important than isolated structures are the structures of binary and ternary complexes of the HRSV transcriptase and replicase complexes.

Structure determination of M2-1 in complex with specific RNA sequences and with full-length and/or crucial fragments of P is the next step for structural biologists. High-resolution structures that represent snap shots of the anti-termination process remain a difficult task, however the recent technical advancement in cryo-electron microscopy (cryo-EM) including the development of direct detectors, stable specimen

stages, high energy electron guns, phase plates, movie mode imaging and advanced image processing strategies, the field of cryo-EM is witnessing a ‘resolution revolution’. A range of different biological specimens that have otherwise previously resisted structure determination like large complexes, membrane proteins, fibrous assemblies, low-abundant protein complexes, difficult to crystallize huge assemblies like viruses, are now feasible targets for structural studies. Unfortunately, there still remains difficulty in structure determination of conformationally flexible, heterogeneous protein assemblies of megadalton size. Further insight into the structure and function of M2-1 can be obtained through the application of recent advances in electron microscopy paired with crystal structures of sub-complexes, in combination of a thorough understanding of its potentially diverse functions from observations made in cellulo.

Acknowledgements

Work on HRSV M2-1 in the Barr and Edwards laboratories is supported in part by funding from the Medical Research Council (Grant: MR/L007290/1)

References

1. Amann BT, Worthington MT, Berg JM (2003) A Cys3His zinc-binding domain from Nup475/tristetraprolin: a novel fold with a disklike structure. Biochemistry 42:217-221.
2 Baskerville SP, Fisher-Hoch GH, Neild, AB, Dowsett (1985) Ultrastructural pathology of experimental Ebola haemorrhagic fever virus infection. J. Pathology 147, pp. 199-209
3. Bermingham A, Collins PL (1999) The M2-2 protein of human respiratory syncytial virus is a regulatory factor involved in the balance between RNA replication and transcription. Proc Natl Acad Sci U S A. 1999 Sep 28;96(20):11259-64.
4. Biedenkopf N, Schlereth J, Grunweller A et al. (2016) RNA Binding of Ebola Virus VP30 Is Essential for Activating Viral Transcription. J Virol 90:7481-7496
5. Blondot ML, Dubosclard V, Fix J et al. (2012) Structure and functional analysis of the RNA- and viral phosphoprotein-binding domain of respiratory syncytial virus M2-1 protein. PLoS Pathog 8:e1002734
6. Bohmwald K, Espinoza JA, Rey-Jurado E et al. (2016) Human Respiratory Syncytial Virus: Infection and Pathology. Semin Respir Crit Care Med 37:522-537
7. Brady MT; American Academy of Pediatrics Committee on Infectious Diseases; American Academy of Pediatrics Bronchiolitis Guidelines Committee. Updated guidance for palivizumab prophylaxis among infants and young children at increased risk of hospitalization for respiratory syncytial virus infection. Pediatrics. 2014;134(2):415-420.
8. Collins PL, Hill MG, Camargo E et al. (1995) Production of infectious human respiratory syncytial virus from cloned cDNA confirms an essential role for the transcription elongation factor from the 5' proximal open reading frame of the M2 mRNA in gene expression and provides a capability for vaccine development. Proc Natl Acad Sci U S A 92:11563-11567
9. Collins PL, Hill MG, Cristina J, Grosfeld H. (1996) Transcription elongation factor of respiratory syncytial virus, a nonsegmented negative-strand RNA virus. Proc Natl Acad Sci U S A. 1996 Jan 9;93(1):81-5
10. Cowton VM, Mcgivern DR, Fearns R (2006) Unravelling the complexities of respiratory syncytial virus RNA synthesis. J Gen Virol 87:1805-1821
11. Dolnik O, Stevermann L, Kolesnikova L et al. (2015) Marburg virus inclusions: A virus-induced microcompartment and interface to multivesicular bodies and the late endosomal compartment. Eur J Cell Biol 94:323-331
12. Fearns R. and Collins PL (1999) Role of the M2-1 transcription antitermination protein of respiratory syncytial virus in sequential transcription. J Virol. 73:5852–5864
13. Hardy RW, Wertz GW (1998) The product of the respiratory syncytial virus M2 gene ORF1 enhances readthrough of intergenic junctions during viral transcription. J Virol 72:520-526
14. Hu J, Robinson JL (2010) Treatment of respiratory syncytial virus with palivizumab: a systematic review. World J Pediatr 6:296-300
15. Kiss G, Holl JM, Williams GM et al. (2014) Structural analysis of respiratory

syncytial virus reveals the position of M2-1 between the matrix protein and the ribonucleoprotein complex. J Virol 88:7602-7617
16. Kuo L, Grosfeld H, Cristina J et al. (1996) Effects of mutations in the gene-start and gene-end sequence motifs on transcription of monocistronic and dicistronic minigenomes of respiratory syncytial virus. J Virol 70:6892-6901
17. Lahaye X, Vidy A, Pomier C et al. (2009) Functional characterization of Negri bodies (NBs) in rabies virus-infected cells: Evidence that NBs are sites of viral transcription and replication. J Virol 83:7948-7958
18. Leyrat C, Renner M, Harlos K et al. (2014) Drastic changes in conformational dynamics of the antiterminator M2-1 regulate transcription efficiency in Pneumovirinae. Elife 3:e02674
19. Li D, Jans DA, Bardin PG, Meanger J, Mills J, Ghildyal R. Association of respiratory syncytial virus M protein with viral nucleocapsids is mediated by the M2-1 protein. J Virol. 2008 Sep;82(17):8863-70. doi: 1128/JVI.00343-08. Epub 2008 Jun 25.
20. Liang B, Li Z, Jenni S et al. (2015) Structure of the L Protein of Vesicular Stomatitis Virus from Electron Cryomicroscopy. Cell 162:314-327
21. Lier C, Becker S, Biedenkopf N (2017) Dynamic phosphorylation of Ebola virus VP30 in NP-induced inclusion bodies. Virology 512:39-47
22. Marcelin JR, Wilson JW, Razonable RR et al. (2014) Oral ribavirin therapy for respiratory syncytial virus infections in moderately to severely immunocompromised patients. Transpl Infect Dis 16:242-250
23. Money VA, Mcphee HK, Mosely JA et al. (2009) Surface features of a Mononegavirales matrix protein indicate sites of membrane interaction. Proc Natl Acad Sci U S A 106:4441-4446
24. Norrby, E., Marusyk, H. & Orvell, C. (1970) Morphogenesis of respiratory syncytial virus in a green monkey kidney cell line (Vero). J Virol. 6, 237–242
25. Palese P, Zheng H, Engelhardt OG et al. (1996) Negative-strand RNA viruses: genetic engineering and applications. Proc Natl Acad Sci U S A 93:11354-11358
26. Renner M, Bertinelli M, Leyrat C et al. (2016) Nucleocapsid assembly in pneumoviruses is regulated by conformational switching of the N protein. Elife 5:e12627
27. Rincheval V, Lelek M, Gault E et al. (2017) Functional organization of cytoplasmic inclusion bodies in cells infected by respiratory syncytial virus. Nat Commun 8:563
28. Scheltema NM, Gentile A, Lucion F et al. (2017) Global respiratory syncytial virus-associated mortality in young children (RSV GOLD): a retrospective case series. Lancet Glob Health 5:e984-e991
29. Shi, T., McAllister DA., O'Brian KL et al (2017) Global, regional and national disease burden estimates of acute lower respiratory infections due to respiratory syncytial virus in young children in 2015: a systematic review and modelling study. Lancet. 2017 Sep 2;390(10098):946-958. doi: 10.1016/S0140-6736(17)30938-8
30. Shoemaker BA, Portman JJ, Wolynes PG (2000) Speeding molecular recognition by using the folding funnel: the fly-casting mechanism. Proc Natl Acad Sci U S A 97:8868-8873
31. Sourimant J, Rameix-Welti MA, Gaillard AL et al. (2015) Fine mapping and characterization of the L-polymerase-binding domain of the respiratory syncytial virus phosphoprotein. J Virol 89:4421-4433.
32. Sutherland KA, Collins PL, Peeples ME. (2001) Synergistic effects of gene-end signal mutations and the M2-1 protein on transcription termination by respiratory

syncytial virus. Virology Sep 30;288(2):295-307.
33. Tanner SJ, Ariza A, Richard CA et al. (2014) Crystal structure of the essential transcription antiterminator M2-1 protein of human respiratory syncytial virus and implications of its phosphorylation. Proc Natl Acad Sci U S A 111:1580-1585
34. Tao YJ, Ye Q (2010) RNA virus replication complexes. PLoS Pathog 6:e1000943
35. Tawar RG, Duquerroy S, Vonrhein C et al. (2009) Crystal structure of a nucleocapsid-like nucleoprotein-RNA complex of respiratory syncytial virus. Science 326:1279-1283
36. Tran TL, Castagne N, Bhella D et al. (2007) The nine C-terminal amino acids of the respiratory syncytial virus protein P are necessary and sufficient for binding to ribonucleoprotein complexes in which six ribonucleotides are contacted per N protein protomer. J Gen Virol 88:196-206
37. Tremaglio CZ, Noton SL, Deflube LR et al. (2013) Respiratory syncytial virus polymerase can initiate transcription from position 3 of the leader promoter. J Virol 87:3196-3207
38. Yu Q, Hardy RW, Wertz GW (1995) Functional cDNA clones of the human respiratory syncytial (RS) virus N, P, and L proteins support replication of RS virus genomic RNA analogs and define minimal trans-acting requirements for RNA replication. J Virol 69:2412-2419

Figure 1.1

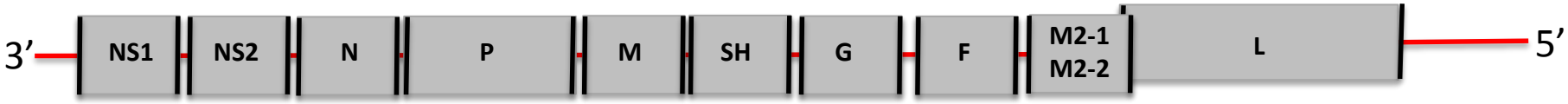


Figure 1.2

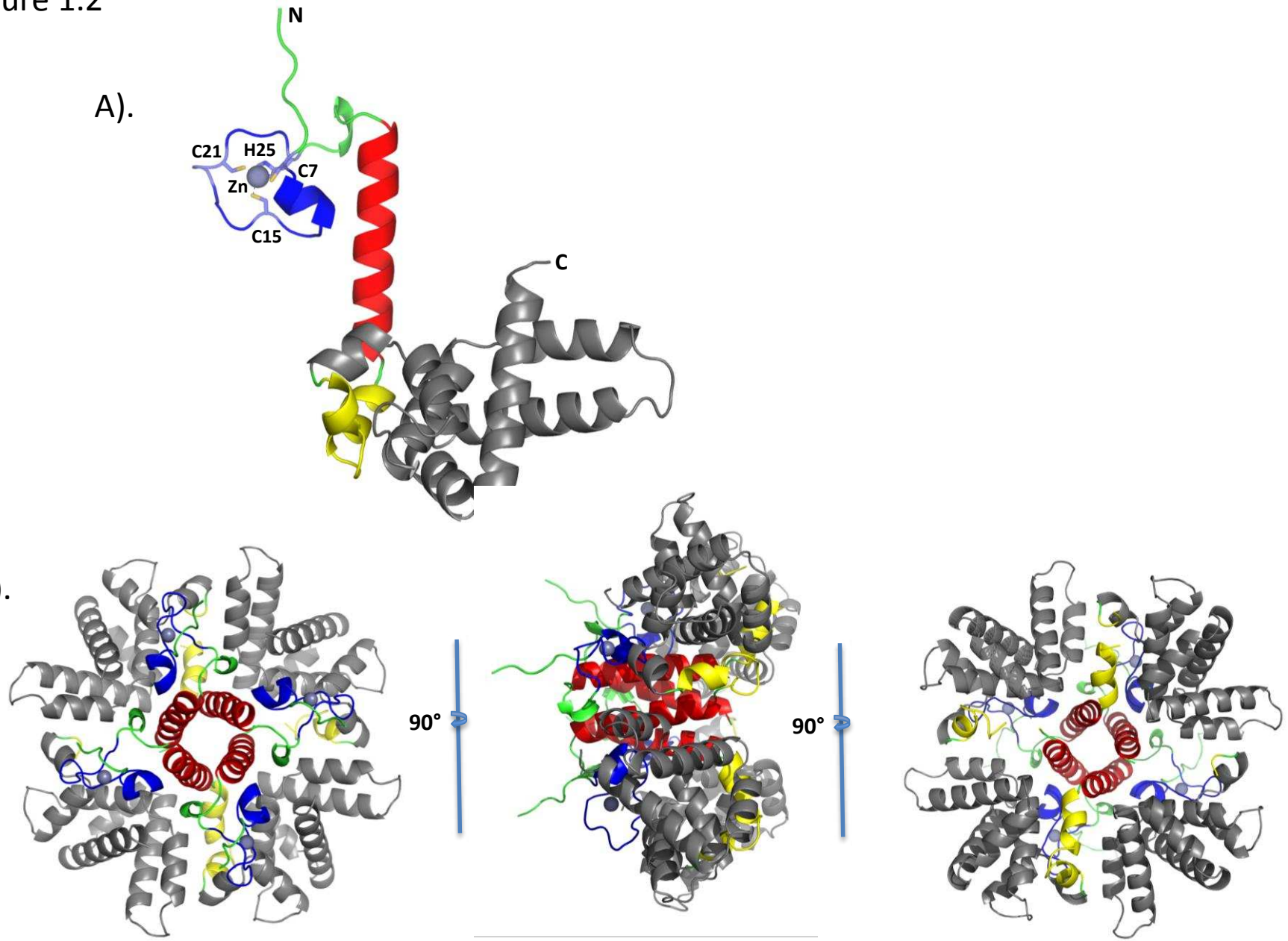


Figure 1.3

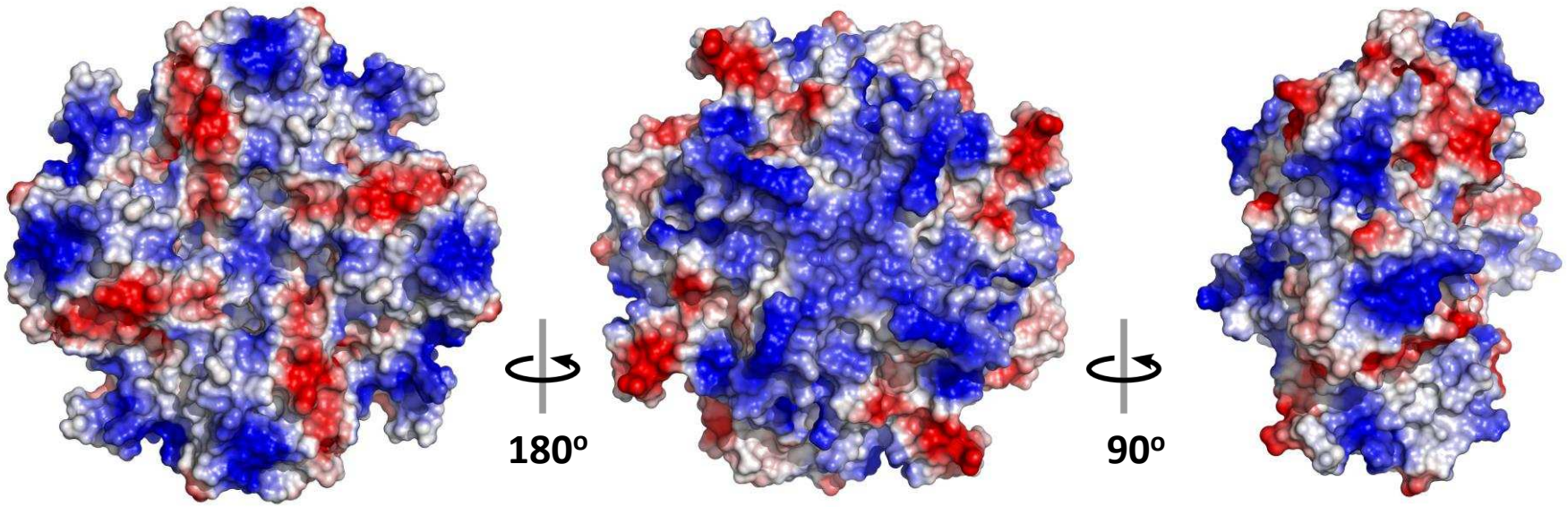


Figure 1.4

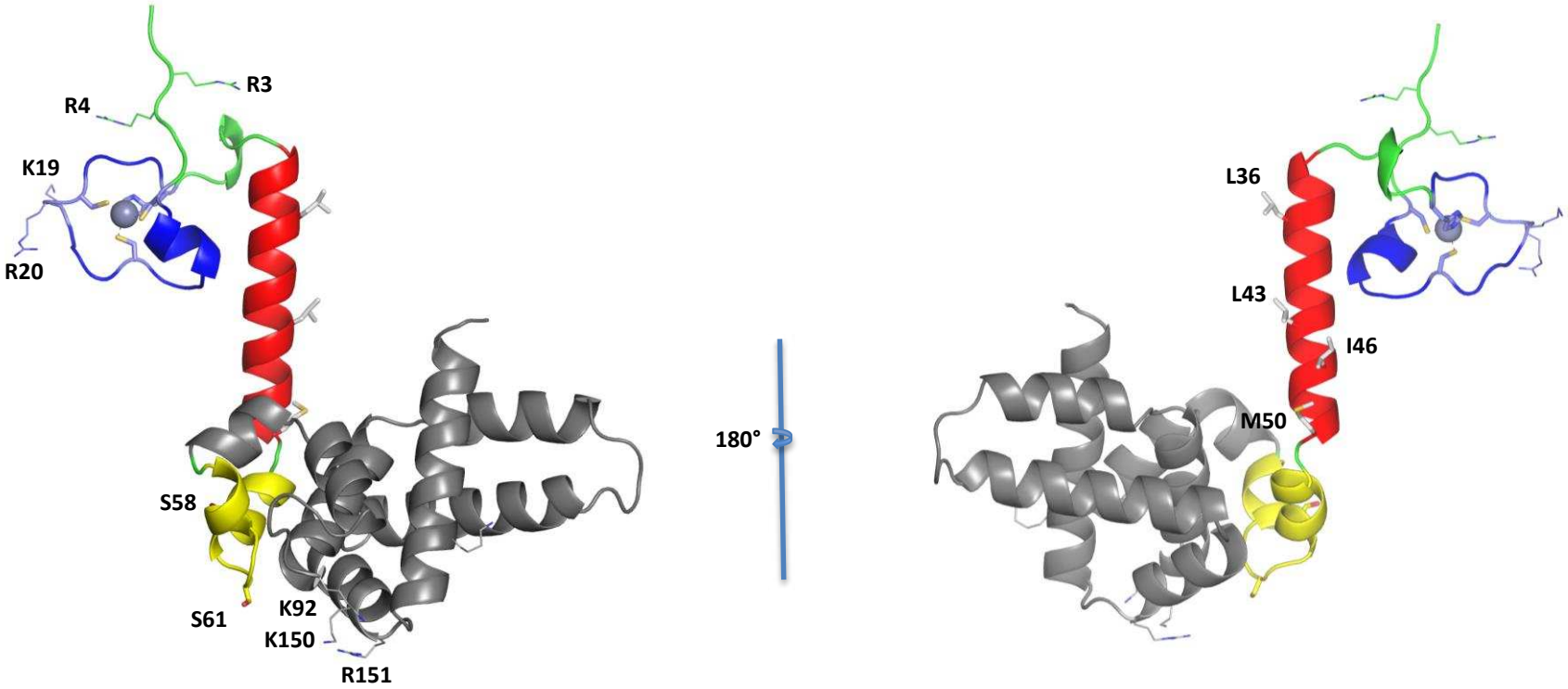
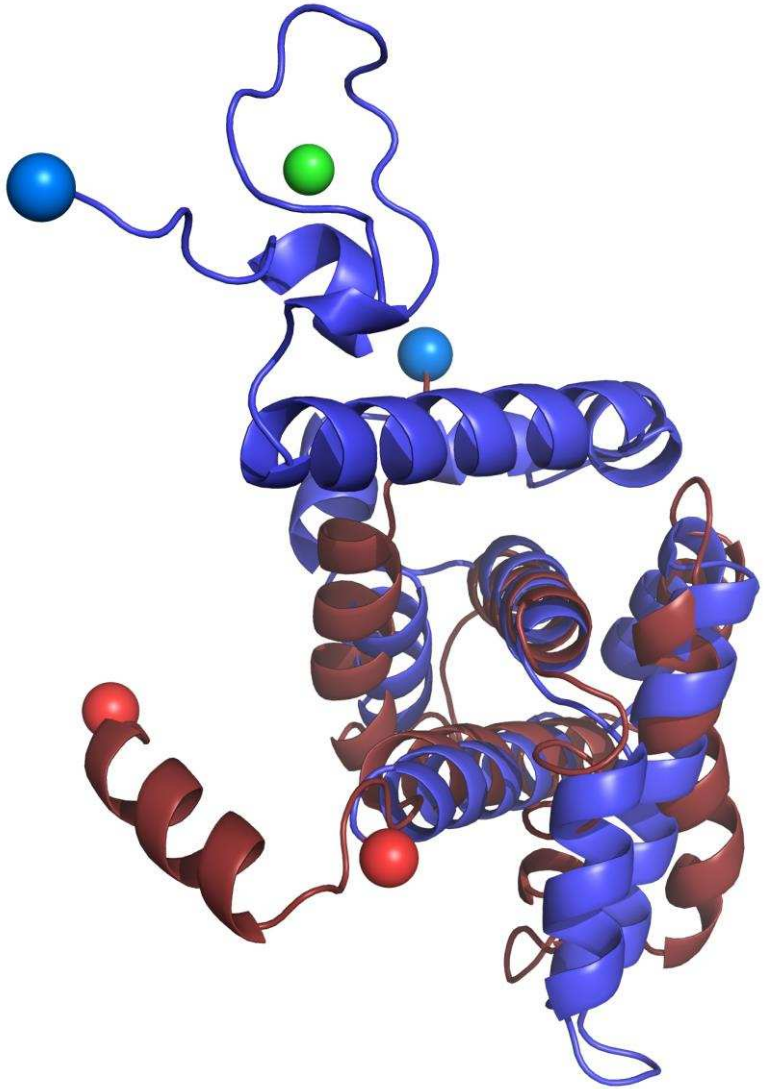


Figure 1.5



Gene position	Protein	Function
1.	Non-structural protein (NS1)	Type I and III IFN antagonist. IFN α/β antagonist mediating antiviral state, suppressing maturation of dendritic cells and T-lymphocyte response. Inhibits phosphorylation of IFN response element 3 disrupting binding to IFN promoter and decreases STAT2 production through degradation.
2.	Non-structural protein (NS2)	Type I interferon antagonist. Causes degradation of STAT2 and interacts with RIG-I to suppress IFN synthesis.
3.	Nucleoprotein (N)	Sequesters viral RNA forming a nucleocapsid (NC) (protein-RNA complex) which is helical. Associates with RNA forming the ribonucleocapsid (RNP) complex
4.	Phosphoprotein (P)	L-protein cofactor that interacts with the NC to place L onto the RNA. Also interacts with M2-1
5.	Matrix protein (M)	Drive HRSV assembly and budding; vital for virus particle formation, having positive and hydrophobic domain important for cytoplasmic membrane binding
6.	Short hydrophobic protein (SH)	Forms a pentameric ion channel, and is able to inhibit tumour necrosis factor alpha (TNFα) signaling, perhaps helping HRSV evade the immune system
7.	Glycoprotein (G)	Involved in viral attachment to the host cell
8.	Fusion protein (F)	Required for fusion of host cell via membranes and promotes syncytia
=9.	M2-1	Transcription factor, with other potential roles as discussed in the text
=9.	M2-2	Inhibits viral transcription up-regulating RNA therefore mediates the regulatory switch from transcription to RNA replication
10.	L-protein	RNA dependent RNA polymerase (RdRp) transcribing/replicating the viral genome

Table 1.1. List of proteins coded by the HRSV genome and their function