Published in final edited form as:

Crit Rev Biochem Mol Biol. 2016 September ; 51(5): 379–394. doi:10.1080/10409238.2016.1215406.

Oligomeric viral proteins: small in size, large in presence

Bhargavi Jayaraman^a, Amber M. Smith^a, Jason D. Fernandes^{b,c}, and Alan D. Frankel^a

^aDepartment of Biochemistry and Biophysics, University of California, San Francisco, CA, USA

^bUC Santa Cruz Genomics Institute, Santa Cruz, CA, USA

^cHoward Hughes Medical Institute, University of California, Santa Cruz, CA, USA

Abstract

Viruses are obligate parasites that rely heavily on host cellular processes for replication. The small number of proteins typically encoded by a virus is faced with selection pressures that lead to the evolution of distinctive structural properties, allowing each protein to maintain its function under constraints such as small genome size, high mutation rate, and rapidly changing fitness conditions. One common strategy for this evolution is to utilize small building blocks to generate protein oligomers that assemble in multiple ways, thereby diversifying protein function and regulation. In this review, we discuss specific cases that illustrate how oligomerization is used to generate a single defined functional state, to modulate activity via different oligomeric states, or to generate multiple functional forms via different oligomeric states.

Keywords

Viral proteins; oligomeric states; multifunctional; HIV-1 Capsid; Ebolavirus VP40; Flavivirus NS1; HIV-1 Rev

Introduction

It is well appreciated that protein function is intimately coupled to its sequence, tertiary and quaternary structure, and that structure also is influenced by extrinsic factors such as ligand binding, post-translational modifications and cellular environment. These structural properties specify protein function, which is acted upon through selective pressure and mutation to evolve protein fitness.

Understanding how proteins encoded by viruses achieve their fitness can be especially informative, as they generally evolve rapidly under selection pressures dictated by the host. Here, we examine some of the consequences of those pressures for evolving protein structure. In particular, viral proteins often rely on interactions with host factors to achieve their function and in many cases require the three dimensional structure of host complexes to adopt their folds (Tahirov et al., 2010; Xiao et al., 2010). Surveys of viral protein structure have shown that, compared to cellular proteins, viral proteins have less well-packed cores

Disclosure statement

The authors report no declaration of interests.

and tend toward intrinsic disorder (Tokuriki et al., 2009; Xue et al., 2012; Xue et al., 2014). Indeed, many viral proteins consist of short linear motifs (SLiMs) that are intrinsically disordered, potentially allowing interactions with multiple host protein partners by adapting to their structures (D'Orso and Frankel, 2010; Hagai et al., 2014).

From an evolutionary perspective, the use of SLiMs and intrinsically disordered regions can be advantageous given the selective pressures acting on viruses. For instance, with their high mutation rates (10^{-4} to 10^{-8} substitutions/nucleotide/generation for viruses versus 10^{-8} to 10^{-9} for eukaryotes) (McCulloch & Kunkel, 2008), viruses will favor robust protein architectures that can tolerate many mutations (Hagai et al., 2014). An intrinsically disordered scaffold by definition can accommodate many sequence changes without disrupting structure. Additionally, high mutation rates can impose genome size limitations on viruses (Bradwell et al., 2013), which in turn makes the use of SLiMs, as opposed to larger and more complex structured domains, advantageous.

While mutational robustness may be an obvious benefit, viral proteins are often under additional pressures to adapt to new environments. For example, viruses may move between different cell types or hosts, thereby drastically altering replication strategies that demand rapid adaptation (Racaniello, 2006; Woolhouse et al., 2001). The use of pliable disordered regions can allow refolding or reorganization of the viral protein upon entry into the new environment (Murzin, 2008; Roche et al., 2007). Because of the obligate nature of the virus–host interaction, viral proteins must rapidly evolve in order to avoid host countermeasures, sometimes leading to molecular “arms-races” in which host proteins or immune cells target existing domains within a viral protein (Daugherty & Malik, 2012). In such cases, a viral protein lacking its own independent structure that requires interaction with a functional partner to adopt structure, might more readily evade these host defenses (Xue et al., 2012).

The pressure for viral protein adaptability is also reflected in their propensity to be multifunctional (Bour & Strebel, 2000; Hale et al., 2008). Proteomics studies have demonstrated that viral proteins tend to have many diverse interactions with host proteins and consequently exhibit multifunctionality (Dyer et al., 2008; Jager et al., 2012). The ability of a viral protein to multitask drastically helps conserve genome space, demanding fewer viral proteins for replication.

Given the substantial selective pressures that shape viral protein structure and function, it is not surprising that viruses utilize a variety of innovative approaches during their evolution. One especially appealing approach has been to combine the use of intrinsic disorder and SLiMs with protein oligomerization to achieve both diversity of protein architecture and function. In essence, it is relatively easy for a virus to encode small “building blocks” that can oligomerize into larger protein assemblies. These assemblies in turn can interact with other host and viral components to integrate into more sophisticated functional machines. One advantage of oligomer use is to help limit genome size, without the need to encode heterologous subunits. The second is it can help achieve mutational robustness, utilizing flexible or dynamic protein–protein interfaces to generate multiple subunit arrangements. The third is flexibly constructed oligomers that assemble into different forms can then be used to generate different functional outputs. While the use of intrinsically disordered

proteins and oligomers to create multifunctional proteins is by no means unique to viruses, those characteristics are well suited to adapt to viral selection pressures and genomic constraints and are likely enriched compared to cellular proteins that have other ways, such as gene duplication, to diversify structure and function.

In this review, we focus on four examples of viral oligomers, each use core monomer building blocks in slightly different ways to build elaborate viral complexes. First, we examine the HIV-1 capsid, which assembles into oligomers of different sizes to generate a huge virion structure from thousands of monomers. Second, we look at the Ebolavirus matrix protein VP40, which also assembles oligomers of different sizes but where each oligomer serves a unique function. Third, we discuss the Flavivirus NS1 protein, which assembles oligomers of different size in combination with post-translational modifications and interactions with the host machinery to accomplish different functions. And fourth, we discuss HIV-1 Rev, which assembles a single oligomer for a single function but demonstrates extreme plasticity in that assembly. Together, these proteins illustrate how a single viral monomeric subunit can be assembled into a variety of oligomeric structures that can tune one function or serve multiple functions critical for viral replication.

Capsid proteins: assembling a complex structural cage using principles of symmetry and quasi-equivalence

Capsid proteins provide a striking example of how a small viral protein can act through an oligomeric assembly to form large and varied structures. Capsids are protein cages that encase the viral genome, ensuring their safe transport and delivery into a host cell. They are unique to viruses and their assembly involves fascinating geometry, symmetry and biology. They adopt different shapes (rods, helical, spherical, icosahedral etc.) and have different sizes, which mostly depend on the size of the viral genome encapsidated. Although many viral capsids are built from a single protein, some viruses such as picornavirus, herpesvirus or adenovirus, have capsids comprising multiple polypeptides. Other viruses, such as the P22 bacteriophage, also encode a scaffolding protein that facilitates capsid assembly but are not a part of the mature capsid (Prasad & Schmid, 2012).

Assembly of the HIV capsid

The Human Immunodeficiency Virus (HIV) capsid is one of the best-studied capsid proteins, and plays an integral role in the HIV life cycle. HIV is a lentivirus, a subfamily of retroviruses, and thus packages two copies of a single-stranded, positive-sense RNA genome within its capsid. The capsid protein itself is a part of larger 55 kDa Gag polyprotein. The immature virus assembles at the plasma membrane of the host cell, initially forming spherical particles composed of 2000–4000 copies of the Gag polyprotein, which extend radially into the core (Briggs et al., 2004). Transformation of the virus into the mature, infectious form occurs during budding/release from the host cell and is triggered by proteolytic processing of the Gag polyprotein. The viral protease cleaves the polyprotein into three major proteins, such as the membrane-associated matrix (MA) protein, the capsid (CA) protein and the RNA genome-associated nucleocapsid (NC) protein. This proteolytic processing causes a dramatic rearrangement in the CA, resulting in the immature spherical

virions maturing into infectious particles, with the CA adopting a fullerene cone shape (Campbell & Hope, 2015; Ganser-Pornillos et al., 2008; Sundquist & Krausslich, 2012).

Interactions in the immature capsid

The ~230 residue CA protein contains two domains: a ~145 residue N-terminal domain (NTD) connected by a short flexible linker to a ~80 residue C-terminal domain (CTD) (Figure 1(A)). During the course of virion assembly and maturation, the CA protein arranges into distinct oligomers (hexamers and pentamers) that serve as repeating units to form the overall CA structure (Campbell & Hope, 2015; Ganser-Pornillos et al., 2008; Sundquist & Krausslich, 2012) (Figure 1). Several high resolution structures of the NTD, the CTD and the full length CA protein, together with cryo-electron microscopy, modeling and molecular dynamics simulations have aided in developing models of both the immature and mature virus CA structure, although heterogeneities in size and morphology have precluded determination of high resolution structures of immature assemblies (Byeon et al., 2009; Ganser-Pornillos et al., 2007; Gitti et al., 1996; Gres et al., 2015; Pornillos et al., 2009, Pornillos et al., 2011; Schur et al., 2015; Worthylake et al., 1999; Zhao et al., 2013). The NTD and CTD engage in different sets of interactions in order to build the CA assembly.

In the immature virion, the CA protein, as a part of the Gag polyprotein, forms continuous but incomplete hexameric lattices (Figure 1(B)). These lattices achieve a spherical shape through the introduction of curvature through irregularities and defects in lattice packing (regular hexameric lattices are normally flat) (Briggs et al., 2009). Each NTD interacts with five neighboring NTDs through homo-dimeric and homo-trimeric interfaces (Figure 1(C)). Two of the five NTDs are also part of the same hexameric unit and have additional contacts stabilizing the interaction. Mutations in any one interface do not appear to affect assembly provided that the other interfaces continue to stabilize the lattice (Schur et al., 2015).

In contrast to the NTDs, the CTDs play a vital role in assembling and stabilizing the hexameric lattices (Briggs et al., 2009; Schur et al., 2015). CA CTDs mediate homo-dimerization between hexamers (Figure 1(D)), and also stabilize each individual hexameric unit through interactions between adjacent CTDs (Figure 1(D)). Interestingly, there is little interaction between the NTD and CTDs in this arrangement (Schur et al., 2015).

Interactions in the mature capsid

Proteolytic processing during budding triggers a dramatic and spontaneous rearrangement of the CA protein to form the characteristic fullerene cone. The mature cone is built from ~250 hexameric units and 12 pentamers that are distributed asymmetrically in the broad and narrow ends of the cone to form a closed structure (Briggs et al., 2004; Ganser et al., 1999; Li et al., 2000).

The mature hexamer has a strikingly different arrangement of the CA subunits compared to the immature hexamer and is formed by an inner core comprising 6 NTDs (Figure 1(E)), surrounded by an outer ring formed by the corresponding CTDs (Ganser-Pornillos et al., 2007; Gres et al., 2015; Pornillos et al., 2009). The relative arrangement of the NTD and the CTD is also very different in the mature lattice (Figure 1(B,E)). While there is little contact

between the NTD and its covalently linked CTD, each NTD forms an extensive interface with its adjacent CTD, firmly securing each hexameric unit (Figure 1(G)). Interhexamer interactions are mediated by CTDs across twofold and threefold interfaces and involve substantial water-mediated interactions (Byeon et al., 2009; Pornillos et al., 2009; Gres et al., 2015; Zhao et al., 2013,) (Figure 1(F)).

An important aspect of CA assembly is the need for curvature in order to form a closed structure (rather than flat sheets). Thus, the local environment of CA molecules in an assembled CA shows minute changes such that the CA molecules are no longer exactly equivalent. Rather, small variations in bondings and flexibility at interfaces results in structures that are quasi-equivalent (Caspar & Klug, 1962). In the HIV CA, curvature is introduced through various means such as irregular lattice defects (predominant in immature particles) (Briggs et al., 2009), water-mediated contacts (Pornillos et al., 2011; Gres et al., 2015), conformational variability in CTD-mediated interactions (Pornillos et al., 2009; Zhao et al., 2013) and flexibility of the NTD–CTD linker resulting in changes in relative orientations of the NTD and CTD (Sundquist & Krausslich, 2012). A high resolution crystal structure of a CA pentamer illustrates the phenomenon of quasi-equivalence between hexamers and pentamers (Pornillos et al., 2011). The pentamer has a structural organization very similar to the hexamer with an NTD core surrounded by an outer CTD belt. The interfaces for the NTD core and the NTD–CTD interaction are very similar between the pentamer and hexamer with subtle differences in intersubunit bonding interactions, facilitated by the flexible linker (Figure 1(H)). Additional demonstration of quasi-equivalence comes from controlled dehydration studies on hexamers which showed that the same interfaces mediate packing in the absence of water through tighter packing and greater buried surface area (Gres et al., 2015).

Interestingly, the tertiary structure of the CA protein is conserved across many retroviruses although they bear little sequence homology (Pornillos et al., 2009; Schur et al., 2015) and assemble into CAs of different shapes. The NTDs of the N-tropic murine leukemia virus (N-MLV) CA protein form hexamers (Mortuza et al., 2004) with spherical CAs (Yeager et al., 1998) and the CA protein from the Rous sarcoma virus (RSV) assemble into hexamers and quasi-equivalent pentamers but the RSV CAs themselves are irregular polyhedra (Cardone et al., 2009).

Additional capsid functions

Although the major function of the CA is to encase the genome, many CA proteins are multifunctional and often play critical roles in viral replication and infectivity. The HIV CA protein interacts with several host proteins following entry into a host cell and which control early steps of the viral life cycle such as uncoating, reverse transcription of the RNA genome and nuclear entry of reverse-transcribed DNA (Ambrose & Aiken, 2014; Campbell & Hope, 2015). Two such proteins, CPSF6 and Nup153 are believed to be involved at the nuclear entry step of the viral life cycle. They specifically recognize the assembled hexamer in the mature CA, binding at the NTD–CTD interface (Bhattacharya et al., 2014; Price et al., 2014), illustrating the functional role of a specific oligomeric structure.

Thus, the HIV CA protein (and many other viral CA proteins) is a simple protein building block which forms a variety of different oligomers that are incorporated together to form two distinct forms of a larger and more complex complete CA lattice. The construction of this larger unit is achieved through symmetry and quasi-equivalence, and overall CA organizations serve multiple functions that include encasing the genome in addition to recognition by host factors.

Ebolavirus matrix protein VP40: a chameleon protein adopting cellular context-dependent structures

Ebola virus is a Filovirus with a ~19 kb long single stranded (–) sense RNA genome, organized into seven genes that encode eight or more polypeptides (Feldmann et al., 2003). The matrix protein, VP40 is the most abundant protein in viral particles and supports structural integrity of the viral particles (Scianimanico et al., 2000). VP40 serves multiple functions, including roles in assembly and budding, and RNA interactions. Each of these functions is achieved through the protein adopting altered structural and oligomeric states.

Domain organization of VP40

VP40 is ~325 residues, consisting of a ~40 residue disordered N-terminus followed by a ~150 residue NTD and a ~130 residue CTD (Figure 2(A)). The protein behaves as a dimer in solution with dimerization mediated by the NTD, burying >1500 Å² of surface area (Figure 2(B)). An intact dimerization surface is required for trafficking, assembly and budding of VLPs from the cell surface. Additionally, the CTDs form an interface with each other in this conformation, enabling VP40 dimers to assemble into filaments (Figure 2(C)), which appear to play a role in assembly and budding (Bornholdt et al., 2013).

VP40 transforms from a dimer to a hexamer at the membrane

At the membrane, VP40 interacts with the negatively charged membrane phospholipids through a charged basic patch on its CTD (disordered in the crystal structures). Structural studies show that, in the presence of membrane-mimicking additives, VP40 undergoes a drastic conformational change, triggering formation of hexamers, which likely facilitate viral matrix assembly (Bornholdt et al., 2013; Scianimanico et al., 2000). VP40 hexamers assemble through both an NTD–NTD interface observed in the in-solution dimer as well as a second NTD–NTD interface known as the “oligomerization interface”. The NTD surface that forms the oligomerization interface is buried by the CTD in VP40 dimers and is thus inaccessible (Figure 2(D)). It becomes accessible following a conformational displacement of the CTDs, which as a result become disordered and hence are not observed in the crystal structure (Figure 2(E)). However, the ends of the hexamer contain VP40 subunits with the same NTD and CTD arrangements as in the dimer. In the crystal, these hexamers further pack into filaments through the same CTD–CTD interface described, and likely reflect the matrix assembly process in the cell (Figure 2(F)) (Bornholdt et al., 2013).

A VP40 octameric ring regulates transcription

In addition to the dimeric and hexameric states, VP40 also assembles into an octameric RNA-binding ring (Bornholdt et al., 2013; Gomis-Ruth et al., 2003), formed by the NTDs interacting via the oligomerization interface observed in the membrane-associated hexamers, as well as yet another interface that also forms the RNA-binding pocket (the RNA-binding interface) (Figure 2(G)). This assembly requires two main conformational events: (1) The CTDs need to be displaced in order to make the NTD oligomerization interface accessible (as described above for hexamers), and (2) The N-terminal residues that form a part of the dimerization interface (residues 45–70) also need to be displaced to allow formation of the RNA-binding interface (Figure 2(G)). Mutational studies suggest that RNA could displace the CTD to facilitate ring formation.

The VP40 octameric ring has been shown to be a negative regulator of viral transcription and is required for the Ebola virus life cycle (Bornholdt et al., 2013; Hoenen et al., 2010; Hoenen et al., 2005). Interestingly, although the dimerization interface is not a part of the octameric ring, VP40 dimers appear to serve as precursors for both the membrane-associated hexamers and the octameric RNA-binding ring (Bornholdt et al., 2013).

Thus, VP40 undergoes both tertiary and quaternary structural rearrangements – via a variety of different interaction surfaces – in response to cues from cellular localization and context (such as membrane association or RNA binding). Each of these rearranged states then fulfills important and independent functions in the viral life cycle. Recent structural studies of VP40 from Marburg virus, a related virus in the filovirus family have shown that Marburg-VP40 has an NTD structure similar to that of Ebola-VP40, forms dimers like Ebola-VP40 and assembles into similar oligomeric rings, suggesting that such structural rearrangements may be conserved across filoviruses (Oda et al., 2016; Timmins et al., 2003).

Flavivirus NS1 protein: oligomeric states drive functionality in virus replication and immune modulation

Flaviviruses are small spherical, positive-sense, single-stranded, enveloped RNA viruses (Lindenbach & Rice, 2003), consisting of 70 members including the West Nile, dengue, yellow fever and Zika viruses. Flavivirus particles are transmitted to humans through arthropods (typically mosquito or tick), with Zika virus as the only documented example of a human to human transmission (Foy et al., 2011). The genome of Flavivirus is ~11 kb in length and encode the mRNA template for the production of all 10 viral proteins (Muller & Young, 2013). Translation of the long open reading frame produces a large polyprotein that is cleaved by a combination of viral and host proteases to generate three structural proteins and seven nonstructural (NS) proteins (Chambers et al., 1990). The seven NS proteins make up the RNA replication complex that is responsible for generating both the minus- and positive-strands of RNA necessary to carry out the viral life cycle (Saeedi & Geiss, 2013; Westaway et al., 2003).

NS1 is a multifunctional player in the Flavivirus life cycle and required for early viral RNA replication, countering immune response (Chambers et al., 1990; Muller & Young, 2013),

and production of infectious viral particles (Scaturro et al., 2015). NS1 is a unique nonstructural protein in that it exists in multiple oligomeric forms, undergoes glycosylation and has multiple localizations (intracellular, cell membrane-associated and secreted), correlated with distinct functions (Chambers et al., 1990; Muller & Young, 2013). It is highly conserved throughout the Flavivirus genus (Muller & Young, 2013).

NS1 has a molecular weight of 46–55 kDa, is 352 amino acids in length, contains 12 invariant cysteine residues and two conserved glycosylation sites (Hahn et al., 1988; Muller & Young, 2013; Pryor & Wright, 1993; Wallis et al., 2004). The range in molecular weight corresponds to the various observed glycosylation patterns dependent on the infecting Flavivirus (Muller & Young, 2013). During translation as a part of the viral polyprotein, NS1 is translocated to the ER lumen via an N-terminal signal sequence (Falgout et al., 1989; Zhang & Padmanabhan, 1993). Following cleavage at its N- and C-termini in the ER, NS1 becomes glycosylated and rapidly assembles into a dimer (Winkler et al., 1988; Winkler et al., 1989). It is then trafficked to different cellular locations, namely intracellular sites of viral RNA replication where it associates with other viral nonstructural proteins that form the replication complex, the cell surface, associated with lipid rafts, or further processed and secreted as soluble hexamers via the Golgi. (Muller & Young, 2013; Schlesinger et al., 1990; Somnuk et al., 2011).

Domain architecture of NS1

NS1 consists of three domains, namely the β -roll dimerization domain (residues 1–29), the wing domain (residues 30–180) and the β -ladder domain (residues 181–352) (Figure 3(A)) (Akey et al., 2014). The wing domain can be divided into two sub domains, an α/β subdomain and a discontinuous connector subdomain. The 12 invariant cysteines form six disulfide bonds within the NS1 monomer and are located throughout the three domains (Akey et al., 2014).

The dimer interface of NS1 is generated by a domain swap of the β -roll dimerization domain and the formation of the β -ladder where each monomer contributes 9 rungs (Figure 3(B)). The β -roll dimerization domain and the connector subdomain of the wing domain create a hydrophobic protrusion on one side of the NS1 dimer plane formed by the β -ladder. Newly synthesized NS1 is a monomer that quickly dimerizes after glycosylation, generating a highly hydrophobic surface that associates with the ER membrane (Crooks et al., 1994; Winkler et al., 1988; Winkler et al., 1989). Detergent molecules co-crystallized in the structure suggest that this side of the dimer interacts with the ER membrane (Akey et al., 2014). The other side of the β -ladder comprises the glycosylation sites, a ~50 residue ordered loop (the spaghetti loop) and a disordered loop from the wing domain (Figure 3(C,D)).

Glycosylation is important in NS1 function

Two conserved glycosylation sites were identified biochemically at Asn 207 and Asn130 (Hahn et al., 1988; Pryor & Wright, 1994; Somnuk et al., 2011), with a third glycosylation site at Asn175 in some NS1 species such as the West Nile virus (Akey et al., 2014). The

glycan at Asn207 is a high-mannose glycan while the glycan at Asn130 is processed into more complex sugars in the golgi secretory pathway (Pryor & Wright, 1994) (Flamand et al., 1999). Mutagenesis studies suggest different roles for the two glycans in regulating NS1 stability, cell surface expression, secretion and binding to components of the human complement system (Somnuk et al., 2011).

An NS1 dimer is essential for RNA replication

Early investigation into the function of NS1 revealed its necessity for RNA replication (Mackenzie et al., 1996). Deletion of NS1 prevented the accumulation of minus-stranded RNA (Lindenbach & Rice, 1997), while perturbations to the NS1 dimer interface reduced RNA replication, indicating that an NS1 dimer is necessary for viral RNA replication (Hall et al., 1999; Pryor & Wright, 1993). NS1 has been shown to directly interact with NS4B, a transmembrane protein in the viral replication complex, through two residues (Arg10-Gln11) located on the hydrophobic β -roll dimerization domain (Youn et al., 2012). Mutational studies have further identified residues in core of the wing and β -ladder domains clustering on hydrophobic protrusions that point toward the β -roll domain which likely faces the ER membrane (Scaturro et al., 2015). However, the precise role of NS1 in viral replication is still poorly understood.

Secreted NS1 forms hexamers that can modulate the immune system

NS1 is secreted only in human cells and not mosquito cells, via the golgi secretory pathway following additional processing of its glycans at Asn130 and circulates in the bloodstream of infected patients (Muller & Young, 2013). Secreted NS1 is hexameric and associated with lipids, existing as a high-density lipoprotein (Figure 3(E)) (Crooks et al., 1994; Mackenzie et al., 1996). Reduced NS1 secretion in the presence of chemical lipid inhibitors and the ability of NS1 to coat and reorganize heterogeneous liposomes into discrete lipoprotein particles highlights the importance of NS1 association with lipids (Akey et al., 2014; Gutsche et al., 2011).

The NS1 hexamer has a barrel-like architecture, built as a trimer of dimers (Akey et al., 2014; Gutsche et al., 2011; Muller et al., 2012) (Figure 3(E)). The interior of the hexamer consists of the three β -rolls that make up the dimer interface while the exterior contains the spaghetti loops, glycosylation sites and the disordered wing-domain loop (Akey et al., 2014).

Secreted NS1 is an important marker of Flavivirus infection (Alcon et al., 2002; Young et al., 2000), interacts with multiple components of the immune system and plays multiple roles in the pathogenesis of flaviviruses (Muller & Young, 2013). It can form immune complexes and activate the complement pathway (Avirutnan et al., 2006). Conversely, it can also engage in immune evasion by binding to players in the complement system and inhibiting the complement pathways (Avirutnan et al., 2010; Avirutnan et al., 2011; Chung et al., 2006). Secreted NS1 from Dengue virus can also bind to uninfected cells of specific types through interactions with heparin sulfate and chondroitin sulfate E, likely leading to vascular leakage (Avirutnan et al., 2007).

Thus, NS1 is a viral protein that utilizes different oligomeric states such as the Ebolavirus VP40 but in addition, also undergoes glycosylation by host enzymes to traffic to different cellular destinations and perform discrete functions. NS1 is less hydrophobic than a typical integral membrane protein, supporting its ability to be associated with the membrane and to be secreted (Winkler et al., 1989). Intracellular NS1 is essential for viral RNA replication, while the role of cell surface-associated NS1 is still unclear. Secreted NS1 on the other hand, by being circulated in the blood stream of infected individuals, adopts myriad roles in immune activation, evasion and pathogenesis. NS1 from different Flaviviruses are likely to have species-specific variations such as glycosylation sites, host interaction partners and mechanism of action.

HIV-1 rev protein: plastic assembly of the Rev-RNA oligomer facilitates a critical viral function

The evolution of viral regulatory proteins provides further evidence for the benefits of using flexible subunit arrangements and small modules to construct larger functional complexes. HIV is a complex retrovirus that encodes regulatory and accessory proteins in its ~9 kb genome, in addition to the *gag*, *pol* and *env* gene products which are common to all retroviruses (Tang et al., 1999). Full-length viral transcripts are produced in the nucleus during transcription of the integrated proviral DNA. HIV relies on overlapping reading frames and alternative splicing to generate all the proteins from these transcripts. In order to circumvent normal host processes that retain unspliced RNAs in the nucleus, HIV depends on the function of the regulatory protein, Rev (produced from a fully spliced message), to export the singly spliced and unspliced viral RNA transcripts to the cytoplasm (Feinberg et al., 1986; Sodroski et al., 1986). Those unspliced transcripts encode the virion structural proteins and provide the genomic RNA for packaging.

Rev domain architecture and function

Rev is a 116 amino acid protein (~13 kDa) and consists of three regions: a hydrophobic oligomerization domain (OD) that flanks an arginine-rich motif (ARM), which serves as both the nuclear localization signal (NLS) and as the RNA-binding domain, and a leucine-rich domain that contains the nuclear export signal (NES) which interacts with the Crm1 nuclear export receptor (Figure 4(A)) (Pollard & Malim, 1998). After Rev is synthesized in the cytoplasm, it is imported into the nucleus by its NLS and the utilization of multiple host importins (Arnold et al., 2006). Inside the nucleus, Rev binds to the Rev response element (RRE) (Figure 4(B)), a highly structured portion of the viral mRNA located in the *env* gene and present only in singly spliced and unspliced transcripts (Hadzopoulou-Cladaras et al., 1989, Malim et al., 1989, Mann et al., 1994). A key feature of Rev is its ability to utilize different surfaces of its alpha-helical ARM to recognize a wide variety of RNA binding sites, and the plasticity of the hydrophobic oligomerization interfaces to generate a large ribonucleoprotein (RNP) complex that recruits the host export machinery.

The RRE provides the scaffold and dictates the conformation of the RNP

A variety of studies highlight the RNA-binding and subunit plasticity of Rev and its use of disordered regions and conformational rearrangements in mediating interactions. At the level of RNA recognition, biochemical studies have shown that Rev assembles on the RRE in several discrete subunit steps (Pond et al., 2009; Robertson-Anderson et al., 2011). Rev binds cooperatively to the RRE, with nucleation beginning at stem IIB, with a second RRE binding site identified at the junction between stems IIA, IIB and IIC and a third at stem I (Figure 4(B)) (Bai et al., 2014; Daly et al., 1993; Daugherty et al., 2008; Holland et al., 1990). Specific Rev binding was also observed with an isolated stem IA hairpin (Daugherty et al., 2008), which contains a similar asymmetric purine-rich internal loop as in stem IIB, suggesting that at least a portion of binding specificity is dictated by the RNA structure (Daugherty et al., 2008). Stem IA likely forms an intermediate binding site in the overall assembly of the Rev/RRE RNP (Bai et al., 2014; Jayaraman et al., 2015). Interestingly, although each Rev subunit uses a single alpha-helical ARM to bind to the RRE, the mode of RNA recognition is unique to each site (Daugherty et al., 2008; Jayaraman et al., 2015). Mutagenesis revealed that Rev utilizes different amino acids to recognize stem IIB and stem IA, in combination with structural studies that corroborated different binding strategies for each site (Daugherty et al., 2008). At the IIB site, Rev makes base-specific contacts with the unpaired bases, while several other ARM residues contact the phosphate backbone (Figure 4(C)) (Battiste et al., 1996; Jayaraman et al., 2014). At the IIA/IIB/IIC junction the interactions are largely electrostatic and the helical register of the ARM is different from that observed with stem IIB (Figure 4(C)) (Jayaraman et al., 2014). The residues identified for stem IA binding reside on the opposite side of the ARM utilized in binding stem IIB and the junction site (Figure 4(C)) (Daugherty et al., 2008; Jayaraman et al., 2014). Taken together, the results support both sequence-specific and structural-context-dependent recognition for each RRE-binding site, and illustrate how the diversity of Rev-RNA recognition allows the evolution of a larger RNP complex using small binding motifs.

The structure of the RNA has co-evolved with Rev to generate binding sites for the multiple Rev subunits and to help organize their arrangement using the structure of the RNA as a scaffold. The RRE is ~350 nucleotides in length although many experimenters use a truncated version (~240 nt) that is sufficient for viral replication (Huang et al., 1991; Mann et al., 1994). The number of Rev monomers that bind to the RRE to form the HIV Rev/RRE RNP is dependent on the length of the RRE (Cook et al., 1991; Daly et al., 1989; Daly et al., 1993; Daugherty et al., 2010a). The RRE secondary structure was predicted to contain four distinct stem loops and one branched stem loop (Malim et al., 1989). Selective 20-hydroxyl acylation analyzed by primer extension (SHAPE) analysis for the RRE has reported alternative structures, comprising either four- or five-stem loops (Bai et al., 2014; Legiewicz et al., 2008) (Figure 4(B)). A small angle X-ray scattering (SAXS) structure of the 4-stem loop RRE revealed an “A” like structure, where each leg contains an identified Rev-binding site (Fang et al., 2013). This architecture of the RRE positions each binding site in an optimal distance and orientation for Rev binding (Fang et al., 2013). Biochemical experiments revealed that the two RRE conformations exist in solution and that the five-stem loop conformation is more active in promoting viral replication compared with its four-stem

loop counterpart (Sherpa et al., 2015). The existence of multiple RNA conformers may modulate the assembly of the Rev/RRE RNP, or its interactions with host factors, and may take further advantage of Rev's pliability.

The hydrophobic OD provides ample opportunities for rev dimer configurations

The plasticity of protein–protein interactions between Rev subunits underlies the inherent assembly of the RNP complex. Three crystal structures of Rev have been determined (Figure 4(D)), and each forms a dimer regardless whether Rev is free (without RNA), bound to a Fab or to an RRE fragment containing IIB and an adjacent bulge created to mimic the stem II junction (Daugherty et al., 2010b, DiMattia et al., 2010, Jayaraman et al., 2014). The N-terminal eight residues of Rev appear disordered, and the C-terminal region including the NES that interacts with the Crm1 export receptor is intrinsically disordered (Daugherty et al., 2010a). Comparison of the RNA-free structure and the Fab-bound structure reveals two different protein–protein interfaces, one between monomers and another generating a higher order arrangement of dimers, with different crossing angles between the subunits (Daugherty et al., 2010b, DiMattia et al., 2010). A third dimer crossing angle was observed in the structure of Rev bound to stem IIB and its adjacent site (Jayaraman et al., 2014). Taken together, these structures reveal a flexible dimer interface that can repack its hydrophobic residues to accommodate a variety of dimer crossing angles while retaining the higher oligomer protein–protein interface. The structure of Rev bound to RNA also exposed the ability of the RNA to compensate energetically unfavorable Rev conformations (Jayaraman et al., 2014). The RRE scaffold in combination with the ability of Rev to adopt multiple dimer interfaces suggests that intermediate binding states (dimer, tetramer, etc.) generate the specificity for the subsequent binding site (Bai et al., 2014). Such intermediate assembly states might simply reflect the pathways of RNP assembly or, in principle, might provide other states used to generate functional complexes with other host factors.

The major host factor used by Rev for RNA export is the Crm1 nuclear export receptor (Fornerod et al., 1997, Neville et al., 1997) and its interaction with the Rev/RRE RNP again highlights how viruses take advantage of oligomerization to achieve complex functions. Crm1 is an abundant exporter of many protein cargos, which all share the presence of a leucine-rich NES (Hutten & Kehlenbach, 2007). The crystal structures of Crm1-host protein complexes show cargo interactions with Crm1 monomers (Dong et al., 2009, Monecke et al., 2009), however the Rev/RRE RNP was found to bind across the interface of a novel Crm1 dimer (Booth et al., 2014). The precise way in which the RNP is organized on the Crm1 dimer remains to be determined, but it appears that the RNA scaffold plays an important role in presenting the Rev subunits (and NESs) to Crm1. One interesting piece of genetic evidence comes from the selection of viral resistance mutants to a dominant-negative version of Rev containing mutations in its NES (M10), where resistance mutations were found to arise in the RRE (RRE61); this RNA was structurally different from the wild-type RRE (Legiewicz et al., 2008). As the Rev NES loosely interacts with a conserved highly rigid NES cleft on Crm1 (Guttler et al., 2010), it seems likely that the manner of presentation of

Rev subunits to the Crm1 dimer, guided by the RRE scaffold, underlies formation of the functional export complex (Booth et al., 2014).

The RRE61 mutation provides evidence that Rev can adapt to different RRE structures during viral evolution. The plasticity of Rev, particularly with its ability to accommodate different Rev dimerization or oligomerization interfaces, is likely an important property that leads to mutational robustness. The hydrophobic and modular nature of the bipartite OD facilitates a wide range of possible packing arrangements, leading to diverse dimer crossing angles (Figure 4(E)). Coupled with the ability of the Rev ARM to recognize different RNA sites via different interfaces, Rev is able to adapt to an evolving RNA scaffold to maintain functional export complexes with small, flexible protein motifs.

Conclusions

The four viral proteins highlighted here, HIV-1 Capsid, Ebola virus VP40, Flavivirus NS1 and HIV-1 Rev, illustrate a variety of ways in which a small viral protein can oligomerize to form large and diverse architectures. While protein oligomeric structures as such are widespread in biology, these examples illustrate how a single-protein sequence can assemble into more than one functionally important structural state, and can that depend on factors such as cellular localization, post-translational modifications and interaction partners.

The HIV CA protein utilizes two domains, connected by a flexible linker, to arrange into a variety of oligomeric forms – including pentamers and hexamers – and then further arranges these substructures into two distinct superstructures – the mature and immature viral CA. The wide diversity of CA substructures is achieved through the use of a flexible linker to alter the relative positions of the domains, while the domains themselves engage in a wide range of packing interactions. However, the large-scale changes to packing during maturation are triggered by proteolytic cleavage of the Gag polyprotein, by yet unknown mechanisms. Interestingly, the overall tertiary structure of CA domains from different retroviruses is highly conserved despite poor sequence identity between CA primary sequences.

The Ebolavirus matrix protein VP40 is similar to HIV CA in that it also has two domains connected by a linker. These domains also engage in discrete packing interactions to build different oligomeric states with defined functions in the viral life cycle. However, unlike the CA protein, VP40 utilizes changes to tertiary structure such as unfolding of parts of domains to reveal previously buried surfaces. Interactions with the cellular membrane and other viral or host factors trigger these dramatic transformations in VP40 assembly.

The Flavivirus protein NS1 is made up of multiple domains, utilizes domain-swapping for dimerization and localizes to the ER membrane during the early stages of viral infection where it plays a role in viral RNA replication. It is expressed on cell surfaces and also secreted in later stages of the viral life cycle as a hexamer, associated with lipids. It also undergoes glycosylation at multiple sites which help in oligomer stability and secretion. Also, by being secreted into the bloodstream, NS1 makes itself available to interact with

multiple host proteins and members of the immune system, thereby directly affecting pathogenesis.

Lastly, the HIV-1 Rev oligomer is different from the other examples in that it is much smaller and has only one known function. However, it demonstrates extreme plasticity by employing a ~60 amino-acid structured region to form adaptable hydrophobic and RNA-binding interfaces, which allows pliable recognition of a viral RNA structure.

This review has attempted to emphasize one evolutionary strategy in which oligomerization of viral proteins can be used to adapt to different protein machines. All four cases make use of pliability in protein structure to align and tether critical interaction interfaces in a variety of manners. The resulting diversity of complex architectures allows the construction of large viral machinery that would otherwise be evolutionary expensive to encode solely within a viral genome. Additionally, the diversity of complexes that can be built from a single protein sequence allows adaptation to diverse sets of functions through its ability to generate distinct oligomeric states, as well as tuning and regulation of existing functions. It is evident that the rapid evolution of viruses has allowed them to develop a large bag of tricks to adapt and evolve proteins in the face of rapidly changing selection pressures.

Acknowledgments

Funding

This work was supported by NIH P50GM082250 grant to A.D.F.

References

- Akey DL, Brown WC, Dutta S, et al. Flavivirus NS1 structures reveal surfaces for associations with membranes and the immune system. *Science*. 2014; 343:881–885. [PubMed: 24505133]
- Alcon S, Talarmin A, Debruyne M, et al. Enzyme-linked immunosorbent assay specific to Dengue virus type 1 non-structural protein NS1 reveals circulation of the antigen in the blood during the acute phase of disease in patients experiencing primary or secondary infections. *J Clin Microbiol*. 2002; 40:376–381. [PubMed: 11825945]
- Ambrose Z, Aiken C. HIV-1 uncoating: connection to nuclear entry and regulation by host proteins. *Virology*. 2014; 454–455:371–379.
- Arnold M, Nath A, Hauber J, Kehlenbach RH. Multiple importins function as nuclear transport receptors for the Rev protein of human immunodeficiency virus type 1. *J Biol Chem*. 2006; 281:20883–20890. [PubMed: 16704975]
- Avirutnan P, Fuchs A, Hauhart RE, et al. Antagonism of the complement component C4 by flavivirus nonstructural protein NS1. *J Exp Med*. 2010; 207:793–806. [PubMed: 20308361]
- Avirutnan P, Hauhart RE, Somnuk P, et al. Binding of flavivirus nonstructural protein NS1 to C4b binding protein modulates complement activation. *J Immunol*. 2011; 187:424–433. [PubMed: 21642539]
- Avirutnan P, Punyadee N, Noisakran S, et al. Vascular leakage in severe dengue virus infections: a potential role for the nonstructural viral protein NS1 and complement. *J Infect Dis*. 2006; 193:1078–1088. [PubMed: 16544248]
- Avirutnan P, Zhang L, Punyadee N, et al. Secreted NS1 of dengue virus attaches to the surface of cells via interactions with heparan sulfate and chondroitin sulfate E. *PLoS Pathog*. 2007; 3:183.
- Bai Y, Tambe A, Zhou K, Doudna JA. RNA-guided assembly of Rev-RRE nuclear export complexes. *Elife*. 2014; 3:e03656. [PubMed: 25163983]

- Battiste JL, Mao H, Rao NS, et al. Alpha helix-RNA major groove recognition in an HIV-1 rev peptide-RRE RNA complex. *Science*. 1996; 273:1547–1551. [PubMed: 8703216]
- Bhattacharya A, Alam SL, Fricke T, et al. Structural basis of HIV-1 capsid recognition by PF74 and CPSF6. *Proc Natl Acad Sci USA*. 2014; 111:18625–18630. [PubMed: 25518861]
- Booth DS, Cheng Y, Frankel AD. The export receptor Crm1 forms a dimer to promote nuclear export of HIV RNA. *Elife*. 2014; 3:e04121. [PubMed: 25486595]
- Bornholdt ZA, Noda T, Abelson DM, et al. Structural rearrangement of ebola virus VP40 begets multiple functions in the virus life cycle. *Cell*. 2013; 154:763–774. [PubMed: 23953110]
- Bour S, Strebel K. HIV accessory proteins: multifunctional components of a complex system. *Adv Pharmacol*. 2000; 48:75–120. [PubMed: 10987089]
- Bradwell K, Combe M, Domingo-Calap P, Sanjuan R. Correlation between mutation rate and genome size in riboviruses: mutation rate of bacteriophage Q β . *Genetics*. 2013; 195:243–251. [PubMed: 23852383]
- Briggs JA, Riches JD, Glass B, et al. Structure and assembly of immature HIV. *Proc Natl Acad Sci USA*. 2009; 106:11090–11095. [PubMed: 19549863]
- Briggs JA, Simon MN, Gross I, et al. The stoichiometry of Gag protein in HIV-1. *Nat Struct Mol Biol*. 2004; 11:672–675. [PubMed: 15208690]
- Byeon IJ, Meng X, Jung J, et al. Structural convergence between Cryo-EM and NMR reveals intersubunit interactions critical for HIV-1 capsid function. *Cell*. 2009; 139:780–790. [PubMed: 19914170]
- Campbell EM, Hope TJ. HIV-1 capsid: the multifaceted key player in HIV-1 infection. *Nat Rev Microbiol*. 2015; 13:471–483. [PubMed: 26179359]
- Cardone G, Purdy JG, Cheng N, et al. Visualization of a missing link in retrovirus capsid assembly. *Nature*. 2009; 457:694–698. [PubMed: 19194444]
- Caspar DL, Klug A. Physical principles in the construction of regular viruses. *Cold Spring Harb Symp Quant Biol*. 1962; 27:1–24. [PubMed: 14019094]
- Chambers TJ, Hahn CS, Galler R, Rice CM. Flavivirus genome organization, expression, and replication. *Annu Rev Microbiol*. 1990; 44:649–688. [PubMed: 2174669]
- Chung KM, Liszewski MK, Nybakken G, et al. West Nile virus nonstructural protein NS1 inhibits complement activation by binding the regulatory protein factor H. *Proc Natl Acad Sci USA*. 2006; 103:19111–19116. [PubMed: 17132743]
- Cook KS, Fisk GJ, Hauber J, et al. Characterization of HIV-1 REV protein: binding stoichiometry and minimal RNA substrate. *Nucleic Acids Res*. 1991; 19:1577–1583. [PubMed: 2027765]
- Crooks AJ, Lee JM, Easterbrook LM, et al. The NS1 protein of tick-borne encephalitis virus forms multimeric species upon secretion from the host cell. *J Gen Virol*. 1994; 75:3453–3460. [PubMed: 7527836]
- D'orso I, Frankel AD. HIV-1 Tat: Its Dependence on Host Factors is Crystal Clear. *Viruses*. 2010; 2:2226–2234. [PubMed: 21709748]
- Daly TJ, Cook KS, Gray GS, et al. Specific binding of HIV-1 recombinant Rev protein to the Rev-responsive element in vitro. *Nature*. 1989; 342:816–819. [PubMed: 2481237]
- Daly TJ, Doten RC, Rennert P, et al. Biochemical characterization of binding of multiple HIV-1 Rev monomeric proteins to the Rev responsive element. *Biochemistry*. 1993; 32:10497–10505. [PubMed: 8399195]
- Daugherty MD, Booth DS, Jayaraman B, et al. HIV Rev response element (RRE) directs assembly of the Rev homooligomer into discrete asymmetric complexes. *Proc Natl Acad Sci USA*. 2010a; 107:12481–12486. [PubMed: 20616058]
- Daugherty MD, D'orso I, Frankel AD. A solution to limited genomic capacity: using adaptable binding surfaces to assemble the functional HIV Rev oligomer on RNA. *Mol Cell*. 2008; 31:824–834. [PubMed: 18922466]
- Daugherty MD, Liu B, Frankel AD. Structural basis for cooperative RNA binding and export complex assembly by HIV Rev. *Nat Struct Mol Biol*. 2010b; 17:1337–1342. [PubMed: 20953181]
- Daugherty MD, Malik HS. Rules of engagement: molecular insights from host-virus arms races. *Annu Rev Genet*. 2012; 46:677–700. [PubMed: 23145935]

- Dimattia MA, Watts NR, Stahl SJ, et al. Implications of the HIV-1 Rev dimer structure at 3.2 Å resolution for multi-meric binding to the Rev response element. *Proc Natl Acad Sci USA*. 2010; 107:5810–5814. [PubMed: 20231488]
- Dong X, Biswas A, Suel KE, et al. Structural basis for leucine-rich nuclear export signal recognition by CRM1. *Nature*. 2009; 458:1136–1141. [PubMed: 19339969]
- Dyer MD, Murali TM, Sobral BW. The landscape of human proteins interacting with viruses and other pathogens. *PLoS Pathog*. 2008; 4:e32. [PubMed: 18282095]
- Falgout B, Chanock R, Lai CJ. Proper processing of dengue virus nonstructural glycoprotein NS1 requires the N-terminal hydrophobic signal sequence and the downstream nonstructural protein NS2a. *J Virol*. 1989; 63:1852–1860. [PubMed: 2522997]
- Fang X, Wang J, O'carroll IP, et al. An unusual topological structure of the HIV-1 Rev response element. *Cell*. 2013; 155:594–605. [PubMed: 24243017]
- Feinberg MB, Jarrett RF, Aldovini A, et al. HTLV-III expression and production involve complex regulation at the levels of splicing and translation of viral RNA. *Cell*. 1986; 46:807–817. [PubMed: 3638988]
- Feldmann H, Jones S, Klenk HD, Schnittler HJ. Ebola virus: from discovery to vaccine. *Nat Rev Immunol*. 2003; 3:677–685. [PubMed: 12974482]
- Flamand M, Megret F, Mathieu M, et al. Dengue virus type 1 nonstructural glycoprotein NS1 is secreted from mammalian cells as a soluble hexamer in a glycosylation-dependent fashion. *J Virol*. 1999; 73:6104–6110. [PubMed: 10364366]
- Fornerod M, Ohno M, Yoshida M, Mattaj JW. CRM1 is an export receptor for leucine-rich nuclear export signals. *Cell*. 1997; 90:1051–1060. [PubMed: 9323133]
- Foy BD, Kobylinski KC, Chilson Foy JL, et al. Probable non-vector-borne transmission of Zika virus, Colorado, USA. *Emerging Infect Dis*. 2011; 17:880–882. [PubMed: 21529401]
- Ganser BK, Li S, Klishko VY, et al. Assembly and analysis of conical models for the HIV-1 core. *Science*. 1999; 283:80–83. [PubMed: 9872746]
- Ganser-Pornillos BK, Cheng A, Yeager M. Structure of full-length HIV-1 CA: a model for the mature capsid lattice. *Cell*. 2007; 131:70–79. [PubMed: 17923088]
- Ganser-Pornillos BK, Yeager M, Sundquist WI. The structural biology of HIV assembly. *Curr Opin Struct Biol*. 2008; 18:203–217. [PubMed: 18406133]
- Gitti RK, Lee BM, Walker J, et al. Structure of the amino-terminal core domain of the HIV-1 capsid protein. *Science*. 1996; 273:231–235. [PubMed: 8662505]
- Gomis-Ruth FX, Dessen A, Timmins J, et al. The matrix protein VP40 from Ebola virus octamerizes into pore-like structures with specific RNA binding properties. *Structure*. 2003; 11:423–433. [PubMed: 12679020]
- Gres AT, Kirby KA, Kewalramani VN, et al. STRUCTURAL VIROLOGY. X-ray crystal structures of native HIV-1 capsid protein reveal conformational variability. *Science*. 2015; 349:99–103. [PubMed: 26044298]
- Gutsche I, Coulibaly F, Voss JE, et al. Secreted dengue virus nonstructural protein NS1 is an atypical barrel-shaped high-density lipoprotein. *Proc Natl Acad Sci USA*. 2011; 108:8003–8008. [PubMed: 21518917]
- Guttler T, Madl T, Neumann P, et al. NES consensus redefined by structures of PKI-type and Rev-type nuclear export signals bound to CRM1. *Nat Struct Mol Biol*. 2010; 17:1367–1376. [PubMed: 20972448]
- Hadzopoulou-Cladaras M, Felber BK, Cladaras C, et al. The rev (trs/art) protein of human immunodeficiency virus type 1 affects viral mRNA and protein expression via a cisacting sequence in the env region. *J Virol*. 1989; 63:1265–1274. [PubMed: 2783738]
- Hagai T, Azia A, Babu MM, Andino R. Use of host-like peptide motifs in viral proteins is a prevalent strategy in host-virus interactions. *Cell Rep*. 2014; 7:1729–1739. [PubMed: 24882001]
- Hahn YS, Galler R, Hunkapiller T, et al. Nucleotide sequence of dengue 2 RNA and comparison of the encoded proteins with those of other flaviviruses. *Virology*. 1988; 162:167–180. [PubMed: 2827375]
- Hale BG, Randall RE, Ortin J, Jackson D. The multifunctional NS1 protein of influenza A viruses. *J Gen Virol*. 2008; 89:2359–2376. [PubMed: 18796704]

- Hall RA, Khromykh AA, Mackenzie JM, et al. Loss of dimerisation of the nonstructural protein NS1 of Kunjin virus delays viral replication and reduces virulence in mice, but still allows secretion of NS1. *Virology*. 1999; 264:66–75. [PubMed: 10544130]
- Hoenen T, Jung S, Herwig A, et al. Both matrix proteins of Ebola virus contribute to the regulation of viral genome replication and transcription. *Virology*. 2010; 403:56–66. [PubMed: 20444481]
- Hoenen T, Volchkov V, Kolesnikova L, et al. VP40 octamers are essential for Ebola virus replication. *J Virol*. 2005; 79:1898–1905. [PubMed: 15650213]
- Holland SM, Ahmad N, Maitra RK, et al. American Society for Microbiology. Human-Immunodeficiency-Virus Rev Protein Recognizes a Target Sequence in Rev-Responsive Element Rna Within the Context of Rna Secondary Structure. *Journal of Virology*. 1990; 64:5966–5975. [PubMed: 2243382]
- Huang XJ, Hope TJ, Bond BL, et al. Minimal Rev-response element for type 1 human immunodeficiency virus. *J Virol*. 1991; 65:2131–2134. [PubMed: 2002556]
- Utten S, Kehlenbach RH. CRM1-mediated nuclear export: to the pore and beyond. *Trends Cell Biol*. 2007; 17:193–201. [PubMed: 17317185]
- Jager S, Cimerancic P, Gulbahce N, et al. Global landscape of HIV-human protein complexes. *Nature*. 2012; 481:365–370.
- Jayaraman B, Crosby DC, Homer C, et al. RNA-directed remodeling of the HIV-1 protein Rev orchestrates assembly of the Rev-Rev response element complex. *Elife*. 2014; 3:e04120. [PubMed: 25486594]
- Jayaraman B, Mavor D, Gross JD, Frankel AD. Thermodynamics of Rev-RNA interactions in HIV-1 Rev-RRE assembly. *Biochemistry*. 2015; 54:6545–6554. [PubMed: 26422686]
- Legiewicz M, Badorrek CS, Turner KB, et al. Resistance to RevM10 inhibition reflects a conformational switch in the HIV-1 Rev response element. *Proc Natl Acad Sci USA*. 2008; 105:14365–14370. [PubMed: 18776047]
- Li S, Hill CP, Sundquist WI, Finch JT. Image reconstructions of helical assemblies of the HIV-1 CA protein. *Nature*. 2000; 407:409–413. [PubMed: 11014200]
- Lindenbach BD, Rice CM. trans-Complementation of yellow fever virus NS1 reveals a role in early RNA replication. *J Virol*. 1997; 71:9608–9617. [PubMed: 9371625]
- Lindenbach BD, Rice CM. Molecular biology of flaviviruses. *Adv Virus Res*. 2003; 59:23–61. [PubMed: 14696326]
- Mackenzie JM, Jones MK, Young PR. Immunolocalization of the dengue virus nonstructural glycoprotein NS1 suggests a role in viral RNA replication. *Virology*. 1996; 220:232–240. [PubMed: 8659120]
- Malim MH, Hauber J, Le SY, et al. The HIV-1 rev transactivator acts through a structured target sequence to activate nuclear export of unspliced viral mRNA. *Nature*. 1989; 338:254–257. [PubMed: 2784194]
- Mann DA, Mikaelian I, Zimmel RW, et al. A molecular rheostat. Co-operative rev binding to stem I of the rev-response element modulates human immunodeficiency virus type-1 late gene expression. *J Mol Biol*. 1994; 241:193–207. [PubMed: 8057359]
- Mcculloch SD, Kunkel TA. The fidelity of DNA synthesis by eukaryotic replicative and translesion synthesis polymerases. *Cell Res*. 2008; 18:148–161. [PubMed: 18166979]
- Monecke T, Guttler T, Neumann P, et al. Crystal structure of the nuclear export receptor CRM1 in complex with Snurportin1 and RanGTP. *Science*. 2009; 324:1087–1091. [PubMed: 19389996]
- Mortuza GB, Haire LF, Stevens A, et al. High-resolution structure of a retroviral capsid hexameric amino-terminal domain. *Nature*. 2004; 431:481–485. [PubMed: 15386017]
- Muller DA, Landsberg MJ, Bletchly C, et al. Structure of the dengue virus glycoprotein non-structural protein 1 by electron microscopy and single-particle analysis. *J Gen Virol*. 2012; 93:771–779. [PubMed: 22238236]
- Muller DA, Young PR. The flavivirus NS1 protein: molecular and structural biology, immunology, role in pathogenesis and application as a diagnostic biomarker. *Antiviral Res*. 2013; 98:192–208. [PubMed: 23523765]
- Murzin AG. Biochemistry. Metamorphic proteins. *Science*. 2008; 320:1725–1726. [PubMed: 18583598]

- Neville M, Stutz F, Lee L, et al. The importin-beta family member Crm1p bridges the interaction between Rev and the nuclear pore complex during nuclear export. *Curr Biol*. 1997; 7:767–775. [PubMed: 9368759]
- Oda S, Noda T, Wijesinghe KJ, et al. Crystal structure of Marburg Virus VP40 reveals a broad, basic patch for matrix assembly and a requirement of the N-terminal domain for immunosuppression. *J Virol*. 2016; 90:1839–1848.
- Pollard VW, Malim MH. The HIV-1 Rev protein. *Annu Rev Microbiol*. 1998; 52:491–532. [PubMed: 9891806]
- Pond SJ, Ridgeway WK, Robertson R, et al. HIV-1 Rev protein assembles on viral RNA one molecule at a time. *Proc Natl Acad Sci USA*. 2009; 106:1404–1408. [PubMed: 19164515]
- Pornillos O, Ganser-Pornillos BK, Kelly BN, et al. X-ray structures of the hexameric building block of the HIV capsid. *Cell*. 2009; 137:1282–1292. [PubMed: 19523676]
- Pornillos O, Ganser-Pornillos BK, Yeager M. Atomic-level modelling of the HIV capsid. *Nature*. 2011; 469:424–427. [PubMed: 21248851]
- Prasad BV, Schmid MF. Principles of virus structural organization. *Adv Exp Med Biol*. 2012; 726:17–47. [PubMed: 22297509]
- Price AJ, Jacques DA, McEwan WA, et al. Host cofactors and pharmacologic ligands share an essential interface in HIV-1 capsid that is lost upon disassembly. *PLoS Pathog*. 2014; 10:e1004459. [PubMed: 25356722]
- Pryor MJ, Wright PJ. The effects of site-directed mutagenesis on the dimerization and secretion of the NS1 protein specified by dengue virus. *Virology*. 1993; 194:769–780. [PubMed: 8389081]
- Pryor MJ, Wright PJ. Glycosylation mutants of dengue virus NS1 protein. *J Gen Virol*. 1994; 75:1183–1187. [PubMed: 8176380]
- Racaniello VR. One hundred years of poliovirus pathogenesis. *Virology*. 2006; 344:9–16. [PubMed: 16364730]
- Robertson-Anderson RM, Wang J, Edgcomb SP, et al. Single-molecule studies reveal that DEAD box protein DDX1 promotes oligomerization of HIV-1 Rev on the Rev response element. *J Mol Biol*. 2011; 410:959–971. [PubMed: 21763499]
- Roche S, Rey FA, Gaudin Y, Bressanelli S. Structure of the prefusion form of the vesicular stomatitis virus glyco-protein G. *Science*. 2007; 315:843–848. [PubMed: 17289996]
- Saeedi BJ, Geiss BJ. Regulation of flavivirus RNA synthesis and capping. *Wiley Interdiscip Rev RNA*. 2013; 4:723–735. [PubMed: 23929625]
- Scaturro P, Cortese M, Chatel-Chaix L, et al. Dengue Virus Non-structural Protein 1 Modulates Infectious Particle Production via Interaction with the Structural Proteins. *PLoS Pathog*. 2015; 11:e1005277. [PubMed: 26562291]
- Schlesinger JJ, Brandriss MW, Putnak JR, Walsh EE. Cell surface expression of yellow fever virus non-structural glycoprotein NS1: consequences of interaction with antibody. *J Gen Virol*. 1990; 71:593–599. (Pt. [PubMed: 2138210])
- Schur FK, Hagen WJ, Rumlova M, et al. Structure of the immature HIV-1 capsid in intact virus particles at 8.8 Å resolution. *Nature*. 2015; 517:505–508. [PubMed: 25363765]
- Scianimanico S, Schoehn G, Timmins J, et al. Membrane association induces a conformational change in the Ebola virus matrix protein. *Embo J*. 2000; 19:6732–6741. [PubMed: 11118208]
- Sherpa C, Rausch JW, Le Grice SF, et al. The HIV-1 Rev response element (RRE) adopts alternative conformations that promote different rates of virus replication. *Nucleic Acids Res*. 2015; 43:4676–4686. [PubMed: 25855816]
- Sodroski J, Goh WC, Rosen C, et al. A second post-transcriptional trans-activator gene required for HTLV-III replication. *Nature*. 1986; 321:412–417. [PubMed: 3012355]
- Somnuk P, Hauhart RE, Atkinson JP, et al. N-linked glycosylation of dengue virus NS1 protein modulates secretion, cell-surface expression, hexamer stability, and interactions with human complement. *Virology*. 2011; 413:253–264. [PubMed: 21429549]
- Sundquist WI, Krausslich HG. HIV-1 assembly, budding, and maturation. *Cold Spring Harb Perspect Med*. 2012; 2:a006924. [PubMed: 22762019]

- Tahirov TH, Babayeva ND, Varzavand K, et al. Crystal structure of HIV-1 Tat complexed with human P-TEFb. *Nature*. 2010; 465:747–751. [PubMed: 20535204]
- Tang H, Kuhen KL, Wong-Staal F. Lentivirus Replication and Regulation. *Annu Rev Genet*. 1999; 33:133–170. [PubMed: 10690406]
- Timmins J, Schoehn G, Kohlhaas C, et al. Oligomerization and polymerization of the filovirus matrix protein VP40. *Virology*. 2003; 312:359–368. [PubMed: 12919741]
- Tokuriki N, Oldfield CJ, Uversky VN, et al. Do viral proteins possess unique biophysical features? *Trends Biochem Sci*. 2009; 34:53–59. [PubMed: 19062293]
- Wallis TP, Huang CY, Nimkar SB, et al. Determination of the disulfide bond arrangement of dengue virus NS1 protein. *J Biol Chem*. 2004; 279:20729–20741. [PubMed: 14981082]
- Westaway EG, Mackenzie JM, Khromykh AA. Kunjin RNA replication and applications of Kunjin replicons. *Adv Virus Res*. 2003; 59:99–140. [PubMed: 14696328]
- Winkler G, Maxwell SE, Ruemmler C, Stollar V. Newly synthesized dengue-2 virus nonstructural protein NS1 is a soluble protein but becomes partially hydrophobic and membrane-associated after dimerization. *Virology*. 1989; 171:302–305. [PubMed: 2525840]
- Winkler G, Randolph VB, Cleaves GR, et al. Evidence that the mature form of the flavivirus nonstructural protein NS1 is a dimer. *Virology*. 1988; 162:187–196. [PubMed: 2827377]
- Woolhouse ME, Taylor LH, Haydon DT. Population biology of multihost pathogens. *Science*. 2001; 292:1109–1112. [PubMed: 11352066]
- Worthylake DK, Wang H, Yoo S, et al. Structures of the HIV-1 capsid protein dimerization domain at 2.6 Å resolution. *Acta Crystallogr D Biol Crystallogr*. 1999; 55:85–92. [PubMed: 10089398]
- Xiao A, Wong J, Luo H. Viral interaction with molecular chaperones: role in regulating viral infection. *Arch Virol*. 2010; 155:1021–1031. [PubMed: 20461534]
- Xue B, Blocquel D, Habchi J, et al. Structural disorder in viral proteins. *Chem Rev*. 2014; 114:6880–6911. [PubMed: 24823319]
- Xue B, Mizianty MJ, Kurgan L, Uversky VN. Protein intrinsic disorder as a flexible armor and a weapon of HIV-1. *Cell Mol Life Sci*. 2012; 69:1211–159. [PubMed: 22033837]
- Yeager M, Wilson-Kubalek EM, Weiner SG, et al. Supramolecular organization of immature and mature murine leukemia virus revealed by electron cryo-microscopy: implications for retroviral assembly mechanisms. *Proc Natl Acad Sci USA*. 1998; 95:7299–7304. [PubMed: 9636143]
- Youn S, Li T, McCune BT, et al. Evidence for a genetic and physical interaction between nonstructural proteins NS1 and NS4B that modulates replication of West Nile virus. *J Virol*. 2012; 86:7360–7371. [PubMed: 22553322]
- Young PR, Hilditch PA, Bletchly C, Halloran W. An antigen capture enzyme-linked immunosorbent assay reveals high levels of the dengue virus protein NS1 in the sera of infected patients. *J Clin Microbiol*. 2000; 38:1053–1057. [PubMed: 10698995]
- Zhang L, Padmanabhan R. Role of protein conformation in the processing of dengue virus type 2 nonstructural polyprotein precursor. *Gene*. 1993; 129:197–205. [PubMed: 8325506]
- Zhao G, Perilla JR, Yufenyuy EL, et al. Mature HIV-1 capsid structure by cryo-electron microscopy and all-atom molecular dynamics. *Nature*. 2013; 497:643–646. [PubMed: 23719463]

Figure 1.

The HIV CA protein assembles into different oligomers during the viral life cycle to form the complex capsid structure. (A) Domain organization and structure of the HIV CA protein monomer showing NTD (blue) and CTD (green) (PDB accession: 4XFX). (B–D) CA arrangement in immature particles (PDB accession: 4USN). (B) CA hexamers showing the relative arrangement of NTD and CTD, with the NTD and CTD from one of the polypeptide chains shown in bright blue and green, respectively. (C) NTD and (D) CTD arrangement in a hexamer with interactions across twofold and threefold axes shown in red. (E–H) CA arrangement in a mature hexamer (PDB accession: 4XFX). (F) NTD arrangement in a mature hexamer. (G) CTD arrangement in mature hexamer with interactions across twofold

and threefold axes shown in red. (H) Intra-hexamer interactions between NTD and its adjacent CTD. (I) CA arrangement in a mature pentamer (PDB accession: 3P05). (see colour version of this figure at www.informahealthcare.com/).

Author Manuscript

Author Manuscript

Author Manuscript

Author Manuscript

Figure 2.

The Ebolavirus matrix protein VP40 is a transformer protein that adopts cellular context-dependent structures. (A) Domain organization and structure of the Ebolavirus matrix protein VP40 monomer showing NTD (blue) and CTD (green) (PDB accession: 4LDB). (B) VP40 dimer using the NTD dimerization surface (PDB accession: 4LDB). (C) VP40 CTD-CTD interaction surface (PDB accession: 4LDB). (D) VP40 NTD-NTD oligomerization surface (PDB accession: 4LDD). (E) VP40 hexamer structure formed using the NTD dimerization surface shown in (B) and NTD oligomerization surface shown in (D) (PDB

accession: 4LDD). (E) Formation of VP40 filaments from hexamers shown in (D) through CTD–CTD interaction surface shown in (C) (PDB accession: 4LDD). (G) VP40 octameric ring structure formed from the NTDs using the NTD oligomerization surface shown in (D) and also an RNA-binding interface (PDB accession: 4LDM and 1H2C). (see colour version of this figure at www.informahelthcare.com/).

Author Manuscript

Author Manuscript

Author Manuscript

Author Manuscript

Figure 3.

Flavivirus NS1 oligomeric states and localization determine the function of NS1. (A) Domain organization and structure of NS1 monomer showing the β -roll (red), wing (blue) and β -ladder (green) domains (PDB accession: 4O6B). (B) NS1 dimer with one subunit colored by domain while the other is colored in gray. The wing domain is broken into two subdomains; α/β subdomain (blue) and the connector subdomain (light blue). (C) The β -roll domain (red) and the greasy finger loop within the wing domain (light blue) are responsible for generating the hydrophobic protrusion that interacts with the ER membrane and

participates in viral RNA replication. (D) NS1 electrostatic surface map reveals the hydrophobic nature of the β -roll and greasy finger protrusion (electropositive colored in blue (+5 kT), electronegative colored in red (−5 kT), electrostatic surface potential determined at pH 7). (E) The secreted NS1 hexamer is a trimer of dimers, with the hydrophobic portion of NS1 located in the center. The disordered regions of NS1 are exposed to interact with the host immune response. (see colour version of this figure at www.informahealthcare.com/).

Author Manuscript

Author Manuscript

Author Manuscript

Author Manuscript

Figure 4.

HIV-1 Rev plasticity facilitates the assembly of ribonucleoprotein complexes. (A) Domain organization and structure of HIV Rev monomer (blue; bipartite oligomerization domain (OD), green; arginine rich motif (ARM), light blue; nuclear export sequence) (PDB accession: 3LPH and 3NBZ). (B) Secondary structure of the Rev response element (RRE) with both the 5 stem-loop and 4 stem-loop structures shown. (C) ARM interactions with the known Rev response element (RRE) binding sites. Residues colored in yellow make direct interactions with the RNA. Residues in gray have been shown through mutagenesis to

prevent RNA binding (PDB accession: 4PMI). (D) Observed Rev dimer crossing angles in the determined structures of Rev (PDB accession: 3LPH, 4PMI and 2×7L). (E) Effect of RNA on the packing of the hydrophobic residues of the bipartite OD. Each monomer is colored in either blue or green (PDB accession: 3LPH and 4PMI). (see colour version of this figure at www.informahealthcare.com/).

Author Manuscript

Author Manuscript

Author Manuscript

Author Manuscript