Influenza virus non-structural protein NS1: interferon antagonism and beyond

Daniel Marc

1Pathologie et Immunologie Aviaire, INRA, UMR1282 Infectiologie et Santé Publique, 37380 Nouzilly, France
2Université François Rabelais, UMR1282 Infectiologie et Santé Publique, 37000 Tours, France

Most viruses express one or several proteins that counter the antiviral defences of the host cell. This is the task of non-structural protein NS1 in influenza viruses. Absent in the viral particle, but highly expressed in the infected cell, NS1 dramatically inhibits cellular gene expression and prevents the activation of key players in the IFN system. In addition, NS1 selectively enhances the translation of viral mRNAs and may regulate the synthesis of viral RNAs. Our knowledge of the virus and of NS1 has increased dramatically during the last 15 years. The atomic structure of NS1 has been determined, many cellular partners have been identified and its multiple activities have been studied in depth. This review presents our current knowledge, and attempts to establish relationships between the RNA sequence, the structure of the protein, its ligands, its activities and the pathogenicity of the virus. A better understanding of NS1 could help in elaborating novel antiviral strategies, based on either live vaccines with altered NS1 or on small-compound inhibitors of NS1.

Introduction

Influenza viruses are enveloped viruses whose genome, ~13 kb in size, is made up of eight ssRNAs of negative polarity. Transcription and replication of viral RNAs take place in the nucleus of the infected cell, generating at least 18 different species of positive-strand viral RNAs. These include (i) the eight complementary RNAs that act as templates for the synthesis of new genomic RNAs, (ii) eight unspliced mRNAs and (iii) at least two spliced mRNAs. The viral mRNAs collectively encode up to 15 proteins, which schematically fall in three classes: (i) the four membrane-associated proteins comprise the two major surface glycoproteins haemagglutinin (HA) and neuraminidase (NA), the ion channel M2, and the matrix protein M1, which coats the inner face of the viral envelope; (ii) the three subunits of the viral polymerase (PA, PB1 and PB2) are tightly associated with the viral ribonucleoproteins (RNPs), in which the viral RNAs are bound to the nucleoprotein (NP); and, finally, (iii) nuclear export protein (NEP, formerly known as NS2) is also present in the virion, whilst NS1, PB1-F2, PA-X, PA-N155, PA-N182 and N40 are expressed in the infected cells, but absent from the viral particle. The unspliced viral mRNAs encode eight full-length products, and three PB1- and PA-related N-terminally truncated products that are not uniformly expressed among different strains (PB1-N40, PA-N155 and PA-N182) (Muramoto et al., 2013; Wise et al., 2009). In addition, PB1-F2 is encoded in the (+1) reading frame of PB1 and PA-X is produced from the PA mRNA through a ribosomal frameshift (Jagger et al., 2012). Finally, ~10–15 % of the M- and NS-specific mRNAs are spliced (Robb et al., 2010), thus encoding M2 and NEP, respectively.

Amongst the viral proteins, non-structural protein NS1 is probably that which is most involved in the interactions between the virus and the host cell, notably in antagonizing the antiviral response. It is therefore a key player in the viral cycle (Hale et al., 2008b). The purpose of this review is to establish relationships between the RNA and the protein, structure and functions of NS1 in the viral cycle, and to make parallels with other viral proteins.

Three-dimensional structure of NS1

Three genera, five proteins

Influenza A NS1 is generally a 230 aa protein, of which we now know at least three distinct variants: alleles (or clades) A and B, and the NS1s of the H17 and H18 bat viruses (Tong et al., 2013). Allele B NS1 is found exclusively in some avian viruses (but not all of them), with rare exceptions in mammals (Guo et al., 1992), whilst viruses harbouring allele A NS1 are by far the most frequent. Each of the two other genera of influenza B and influenza C viruses has its own NS1 (281 and 246 aa, respectively) (Alamgir et al., 2000; Yin et al., 2007). Whether NS1 orthologues exist in distantly related orthomyxoviruses, such as infectious salmon anemia virus, is less clear (Mérou et al., 2011); however, it is known that the orthomyxovirus Thogoto virus has no NS1 gene and that its IFN-antagonist activities rely on ML, a splice variant of the matrix protein (Buettner et al., 2010).
Influenza A virus NS1

Three-dimensional structure

NS1 is made up of two structural domains that are joined by a flexible linker region (Qian et al., 1994). The RNA-binding domain (RBD, Fig. 1a) is a homodimer of aa 1–73, whilst aa 80–230 form the effector domain (ED, Fig. 1b). Structures of the ED and RBD generally have been solved independently, using bacterially expressed recombinant proteins representing either domain. The structure of full-length NS1 was solved for three variants, although it was necessary to introduce the double substitution R38A–K41A to prevent the aggregation of the recombinant protein (Bornholdt & Prasad, 2008; Carrillo et al., 2014).

RBD. The structure of the RBD (Fig. 1a), which was solved by both crystallography and NMR (Chien et al., 1997; Liu et al., 1997), is composed of six α-helices (three from each monomeric unit). Antiparallel helices α2 and α2’ bind to the major groove of dsRNA, thanks to several basic amino acids that include Arg38 and Lys41 (Cheng et al., 2009). The RBD of influenza B viruses shows a very similar homodimeric structure (Fig. 2b). In spite of the distant relationship between the two polypeptide sequences, comparison of their respective α2 helices reveals a similar array of several conserved basic amino acids (Yin et al., 2007). Moreover, we can remark that RNA binding by two positively charged, antiparallel α-helices is a feature shared by a number of RNA- or dsRNA-binding proteins that also act as homodimers: the B2 proteins of Flock House virus and Nodamura virus, two positive-strand RNA viruses (Chao et al., 2005; Körber et al., 2009), the Rop protein of plasmid CoEI (Struble et al., 2008), and the core proteins of Dengue virus and West Nile virus (Dokland et al., 2004; Ma et al., 2004). Fig. 2 emphasizes the structural similarity between these proteins (left panels), but also reveals differences in the distribution of charged amino acids in the RNA-binding site, suggesting distinct modes of RNA binding.

ED. The monomeric unit of the ED is composed of three α-helices and seven β-strands (Fig. 1b). The latter form a large antiparallel β-sheet that surrounds one side of the long helix α5. This 'α-helix β-crescent fold' has no structural equivalent, except that it is remotely reminiscent of the 'hotdog fold' (Bornholdt & Prasad, 2006; Dillon & Bateman, 2004). However, jFATCAT-precalculated structure homologies at the Protein Data Bank (PDB) website (Prlc et al., 2010) reveal that a number of proteins share a conserved core structure with the ED of NS1 consisting of a long α-helix lying upon a four-strand antiparallel β-sheet. These include ribosomal proteins S11 and L18 (PDB ID: 3V2C and 1ILY, respectively), O6-methylguanine DNA methyltransferase of archaea Pyrococcus kodakaraensis (PDB ID: 1MG7), and a model PIWI protein from Archaeoglobus fulgidus (PDB ID: 2W42), which was shown to regulate RNA silencing (Parker et al., 2004).

Fig. 1. NS1 structure. (a) Schematic representation of the RBD (dimer). Antiparallel helices α2 and α2’ (green) of the dimeric structure bind to the major groove of dsRNA. The axis of the RNA duplex helix is roughly parallel to those of helices α2 and α2’. The cartoon was drawn after the RBD structure of NS1 from virus A/ Udorn/72 (H3N2) [Protein Data Bank (PDB) ID: 1AIL]. (b) Schematized structure of the ED. Seven antiparallel β-strands form a large β-sheet that surrounds one side of the long helix α5. The RBD is connected to strand β1 through the linker region formed by aa 73–88 (see Fig. 6 below for numbering of β-strands and α-helices). In this oversimplified drawing, the large β-sheet was unfolded and flattened, in order to highlight the connectivity and topology of the building blocks (after PDB ID: 2GX9) (Bornholdt & Prasad, 2006). Shaded areas correspond to the sites of interactions with CPSF30 (30 kDa subunit of cleavage and polyadenylation specificity factor; aa 103, 106 and 183–187), p85-β of phosphatidylinositol 3-kinase (PI3K) (aa 89, 95–100, 142, 145 and 164) and PKR (protein kinase RNA-activated; aa 123–127). (c) Three distinct quaternary structures observed for full-length NS1: ‘semi-open’ conformation of NS1 from an H6N6 virus (left), closed conformation of the same full-length NS1 harbouring a deletion of aa 80–84 (centre) and ‘open’ conformation (right) of the full-length NS1 of H5N1 virus VN04, which also harbours the aa 80–84 deletion. The RNA-binding interface is perpendicular to the plane of the figure (adapted from Carrillo et al., 2014). (d) NS1 fibrils. Alternate dimerization of the RBD and ED results in long fibrils, which in turn can associate around a threefold axis to form long tubular structures that can accommodate a dsRNA molecule in their ~20 Å diameter inner tunnel (only one of the three tubule-forming fibrils is represented) (drawing adapted after Bornholdt & Prasad, 2008). The hatched blue rectangles in the RBD depict the approximate position of basic amino acids R38, R38’, K41 and K41’, which point towards the interior of the tunnel.
et al., 2009). The ED structure has been solved for several variants of NS1: allele B NS1 of an avian influenza A virus (Hale et al., 2008a; Kerry et al., 2011), human H3N2 viruses (Aramini et al., 2011), A/PR8/34 (H1N1) (Bornholdt & Prasad, 2006; Kerry et al., 2011), pandemic pdmH1N1 (2009) (PDB ID: 3M5R) and A/Vietnam/1203/2004 (H5N1) (Bornholdt & Prasad, 2008). Whatever the variant analysed, the overall fold is conserved and close inspection of the structures reveals a highly conserved array of charged, surface-exposed amino acids, in spite of the variability of the peptide sequences (Fig. 3).

Dimerization of the structural domains of NS1

Whilst both full-length NS1 and the isolated RBD make obligate dimers (Nemeroff et al., 1995; Qian et al., 1995), the ED forms weak dimers (Aramini et al., 2011; Bornholdt & Prasad, 2006, 2008; Hale et al., 2008a; Kerry et al., 2011). Two distinct interfaces of dimerization have been identified: whilst a strand–strand interface was identified in the crystal structures of the ED of A/PR8/34 (H1N1) and of the full-length NS1 of A/Vietnam/1203/2004 (H5N1) (Bornholdt & Prasad, 2006, 2008), other authors found that only the helix–helix interface was strictly conserved across several variants of ED. Several criteria characterize the helix–helix dimer as a transient dimer, whilst the strand–strand interaction is even weaker (Aramini et al., 2011; Kerry et al., 2011).

The structures observed for both the ED and the full-length proteins raise two questions. (i) What is the biological relevance of the multiple dimerization modes of the ED? (ii) How do the RBD and ED interact with each other? These issues, which are discussed briefly below, were examined in more detail in the recent review by Hale (2014). First, weak dimerization of the ED may provide a way to regulate distinct biological activities of NS1, as it allows reversible interactions with itself or with other partners and also confers flexibility to the relative orientation of the two ED monomers in the helix–helix dimer (Hale, 2014; Kerry et al., 2011). With regard to the interdomain contacts and the overall quaternary structure (schematized in Fig. 1c), the three full-length structures now available reveal that the linker region allows large conformational changes, such as the ‘helix-open’ and ‘helix-closed’ states that have been hypothesized to be associated with distinct properties of NS1 (Carrillo et al., 2014; Hale, 2014; Kerry et al., 2011). The two domains of the full-length protein VN04 (from an H5N1 virus) were observed to dimerize independently, forming long linear chains of alternating RBD and ED dimers (Fig. 4).
Dimer chains, in turn, associate around a threefold axis, forming long tubular structures containing an inner tunnel whose diameter could accommodate a dsRNA (Bornholdt & Prasad, 2008). Whether these crystal structures are biologically relevant is unknown, although they may be related to the long fibrillar structures observed in vitro when recombinant NS1 is mixed with dsRNAs (Bornholdt & Prasad, 2008) and to the NS1-containing filamentous or crystalline structures that are sometimes observed in the cytoplasm of infected cells (Newby et al., 2007; Terrier et al., 2012; Yoshida et al., 1981). However, the structures that were solved recently for two variants of the full-length NS1 from an avian H6N6 virus reveal that both the length and the sequence of the linker region can dramatically alter the quaternary state of NS1 (Fig. 1c) and its mode of oligomerization. These distinct quaternary states probably confer particular properties to the variant NS1s and to the virus, as mutations that modify the length or the sequence of the linker region can dramatically alter the quaternary state of NS1 (Fig. 1c) and its mode of oligomerization. These distinct quaternary states probably confer particular properties to the variant NS1s and to the virus, as mutations that modify the length or the sequence of the linker region can dramatically alter the quaternary state of NS1 (Fig. 1c) and its mode of oligomerization. These distinct quaternary states probably confer particular properties to the variant NS1s and to the virus, as mutations that modify the length or the sequence of the linker region can dramatically alter the quaternary state of NS1 (Fig. 1c) and its mode of oligomerization. These distinct quaternary states probably confer particular properties to the variant NS1s and to the virus, as mutations that modify the length or the sequence of the linker region can dramatically alter the quaternary state of NS1 (Fig. 1c) and its mode of oligomerization.

Several messages in NS1 mRNA

The mRNA that is transcribed from the segment 8 viral RNA either encodes NS1, if unspliced, or NEP after being spliced (Fig. 4a). As a consequence, and as NEP’s coding sequence extends for ~140 nt beyond the NS1 stop codon, the mRNA of NS1, like that of M1, has a 3’ UTR that is much longer than that of the six other viral mRNAs. NS1 mRNA harbours at least two sequence motifs to which the RBD of NS1 was shown to bind in vitro (Marc et al., 2013): AGC(A/G)AAAG, which is part of the universal sequence that is invariably present at the 5’ end of all positive-strand RNAs of influenza A viruses, and UG(A/C)UUGAAG, of which there are two copies in the 3’ UTR of NS1 mRNA, at positions 735 (15 nt downstream of the NS1 stop codon) and 817 (48 nt upstream of the NEP stop codon). In addition to these sequence motifs in the UTRs, two elements of secondary structure were predicted in the coding sequence, in the positive strand of segment 8 (a). Schematized structure of the ED, also showing the interaction sites of PKR and CPSF30 (shaded areas). (b–e) Space-filling representation of the ED structure from four distantly related variants: human H3N2 Udorn strain, avian strain Albany (allele B), H5N1 strain VN04 and pdmH1N1 (2009). Charged amino acids (red: acidic amino acids Asp and Glu; blue: basic amino acids Arg, Lys and His) are highlighted, in order to emphasize the highly conserved array that they form at the surface of the ED. W187 (top) is part of the helix–helix interface and is also close to the CPSF30-binding site (see Fig. 1b). As can be seen in Fig. 6 below, amino acids D92, E96, E97, E152, E153, K126, K131, D125, W187 and R200 are strictly or highly conserved. A number of these conserved amino acids are either part of binding sites (E186: CPSF30; 123–127: PKR) or are associated with phenotype-altering substitutions (D92 and R200) (see Table 2). Structures were rendered using software Cn3D (PDB IDs are indicated). Numbering of amino acids corresponds with that of the multiple alignment in Fig. 6.

Influenza A virus NS1

http://vir.sgmjournals.org 2597
The first motif (nt 107–174, motif 1 in Fig. 4a) can fold into two alternative secondary structures and corresponds at the protein level to the peptide 28GDAPFLDRLRRDQKSLRGRGST49, which makes up the helix α2 (i.e. the RNA-binding site) of the RBD. Interestingly, the mfold (http://mfold.rna.albany.edu/?q=mfold)-predicted structure of this motif is very similar to that of NS1-specific RNA aptamers (Fig. 4b, d). Of note, the double substitution R38A–K41A, which is frequently introduced to invalidate the RBD, dramatically alters the structure of this motif (Fig. 4c). This perhaps may also contribute to the severe attenuation of these viruses, and could partly explain the large phenotype differences observed between viruses harbouring the double substitution (R38A–K41A) and the single-site mutants (R38A or K41A) (Donelan et al., 2003; Jiao et al., 2008; Lalime & Pekosz, 2014; Newby et al., 2007).

The second RNA structural motif (nt 524–576, motif 2 in Fig. 4a) has the potential to switch between a hairpin and a pseudoknot. It encompasses the 3′-splice acceptor site of NS1 mRNA and corresponds at the protein level to the peptide sequence167PGHTNEDVKNAIGVLI182 in the ED (Gultyaev et al., 2010; Moss et al., 2011). Remarkably, the latter structural motif is supported by both nucleotide covariations and structure probing of the RNA, and it is conserved not only in both alleles of NS1 from influenza A viruses, but also in the NS1 mRNA of influenza B viruses (Gultyaev et al., 2010). Of note, a putative pseudoknot was also predicted at the 3′-splice acceptor site of M1 mRNA (Moss et al., 2011).

**Post-translational modifications, localization and binding partners**

**Synthesis and post-translational modifications**

NS1 is produced abundantly during the whole course of the viral cycle, to the point of forming crystal-like structures in the infected cells (Terrier et al., 2012; Yoshida et al., 1981). In the infected cell, a significant fraction of NS1 harbours phosphorylated amino acids. These include Thr215, Ser42 and Ser48, although these are not strictly conserved (Hale et al., 2009; Hsiang et al., 2012). Although replacement of Thr215 was observed to reduce viral replication, phosphorylation of NS1 does not appear to play a critical role in the viral cycle, as indicated by two independent studies, both carried out with H3N2 virus A/Udorn/72 (Hale et al., 2009; Hsiang et al., 2012). In addition to phosphorylation, two types of ubiquitination have been observed: SUMO1-linkage to Lys70 and Lys219 (Santos et al., 2007).
NS1 is found in both the nuclear and cytoplasmic compartments. It is imported into the nucleus through interactions of importin-α with its nuclear localization signal (NLS), which is made up by basic amino acids (Arg35, Arg38 and Lys41) in helix α2 of its RBD (Melén et al., 2007). Conversely, a nuclear export signal in its ED allows its cytoplasmic redistribution (Li et al., 1998). In the nucleus, it sometimes concentrates in the nucleolus (Melén et al., 2007). The subcellular distribution and the nuclear-to-cytoplasmic fraction ratio of NS1 show a wide range of variations, depending on the virus strain, cell type, early or late steps of the viral cycle, or whether NS1 expression is observed in transfected or infected cells (Forbes et al., 2013; Li et al., 1998; Newby et al., 2007; Volmer et al., 2010).

RNA ligands

One of the first properties observed for NS1 was the binding of RNAs through its N-terminal domain (Qian et al., 1995). RNA partners of NS1 were identified through various approaches, and include small nuclear RNAs U6 and U6atac involved in splicing (Lu et al., 1994; Qiu et al., 1995; Wang & Krug, 1998), polyadenylated mRNAs (Qiu & Krug, 1994), viral genomic RNAs (Hatada et al., 1997), dsRNAs (Hatada & Fukuda, 1992), and viral mRNAs (Marc et al., 2013; Maríon et al., 1997a). The diversity of RNA ligands identified likely reflects the high affinity of the RBD towards RNA and also suggests a low specificity of interaction. However, the identification of these ligands was probably related to the methods used in the 1990s. Deep-sequencing methods that have been developed since, such as HITS-CLIP (high-throughput sequencing of RNA isolated by cross-linking immunoprecipitation; Darnell, 2010), would probably help to better identify the RNAs that are bound to NS1 in the infected cell. However, a SELEX (systematic evolution of ligands by exponential enrichment) approach allowed us to identify at least two virus-specific motifs that strongly bind recombinant NS1 in vitro (Marc et al., 2013). AGCAAAAG is strictly conserved at the 5’ end of influenza A virus positive-strand RNAs, whilst UGAUUGAAG is highly conserved in the 3’ UTR of NS1 mRNA (see above and Fig. 4a). In vitro, the RBD binds cooperatively to aptamers containing these sequences with a subnanomolar affinity. Interestingly, both motifs are specific for the viral positive-strand RNAs, which suggests that NS1, in addition to its IFN-antagonist activities, may exert biological activities towards viral positive-strand RNAs. Interestingly, the similarity of the mfold-predicted structures between NS1 aptamers and the RNA motif corresponding to helix α2 of the RBD (motif 1 in Fig. 4a) suggests the hypothesis that NS1 may also bind this structural motif in its own mRNA. In addition, we showed that NS1 also recognized a short duplex motif that is reminiscent of the canonical splice donor site, which may partly explain its splicing inhibitor activity by competition with spliceosomal protein U1C (Marc et al., 2013). The ED is apparently not involved in these interactions, at least in vitro, although we cannot rule out the possibility that it may somehow regulate the RNA-binding properties of the RBD in vivo.

Protein partners

More than 50 proteins have been shown, through various approaches, to interact with NS1. The list of partners is not exhaustive as (i) novel approaches frequently yield new candidates (de Chassey et al., 2013; Shapira et al., 2009) and (ii) the set of binding partners may vary depending on polymorphisms in NS1 amino acid sequence, which has been clearly demonstrated for some well-known partners (de Chassey et al., 2013; Shapira et al., 2009; Twu et al., 2007). Schematically, NS1 partners fall into four groups. The first group consists of nuclear proteins acting in the maturation and nucleo-cytoplasmic export of cellular mRNAs: CPSF30 (30 kDa subunit of cleavage and polyadenylation specificity factor), poly(A)-binding protein II (PABPII), NFX1/TAP and several proteins of the nuclear export machinery, nucleolin. The second group contains cytoplasmic proteins that play a role in mRNA transport and translation: Staufen, translation initiation factor eIF4G1 and poly(A)-binding protein I (PABPI). Several proteins involved in signalling pathways and in various steps of the antiviral defence can be grouped in a third class: RIG-I, TRIM-25, PKR (protein kinase RNA-activated), PACT (PKR activator), as well as the regulatory subunit p85-β of phosphatidylinositols 3-kinase (PI3K) and PDZ-containing proteins. PDZ domains are adaptor modules that allow a protein to specifically interact with partners through recognition of their C-terminal peptide. PDZ-domain-containing proteins take part in several signalling pathways; a few hundred of them have been described in humans and several virus families express PDZ-interacting proteins (for a review, see Javier & Rice, 2011). NS1’s PDZ-binding motif is made up of its 4 aa C-terminal peptide and, more specifically, consists of the motif 227ESEV230 in NS1 of avian influenza viruses (Fan et al., 2013; Obenauer et al., 2006). Lastly, NS1 also interacts with the viral polymerase complex and, more specifically, with NP, within multiprotein complexes containing NS1, CPSF30 and the polymerase subunits (Kuo & Krug, 2009). See Table 1. As shown in Table 1, most of these proteins were identified through the yeast two-hybrid method, and the interaction was generally confirmed in the infected cell through glutathione S-transferase pull-down and co-immunoprecipitation. Most NS1 partners interact with the ED. Some of them, like importin-α isoforms, interact directly with the...
RBD (Melén et al., 2007), whilst others interact indirectly via the bound RNA – in which case the interaction is abolished by RNase treatment. Some interactions are conditional, which probably reflects the existence of multimeric complexes, e.g. the PA subunit of the viral polymerase was shown to interact with NS1 within complexes containing both NS1 and CPSF30 (Kuo & Krug, 2009). Conversely, some interactions may be mutually exclusive: it has been proposed that interaction of NS1 with dsRNAs may prevent its binding to CPSF30 (Aramini et al., 2011). It is important again to emphasize that polymorphisms in the NS1 peptide sequence may result in variations in the set of NS1 partners (de Chassey et al., 2013; Rajsbaum et al., 2012; Shapira et al., 2009). For instance, depending on their C-terminal sequence that makes up the PDZ-binding motif, distinct NS1 variants will not interact with the same PDZ modules. Other interactions involving amino acids from the ED or RBD have been shown to be abolished by substitution of these critical amino acids (Table 2).

A significant number of NS1-interacting proteins are dsRNA-binding proteins (de Chassey et al., 2013) and one may therefore ask whether these are genuine protein–protein interactions or RNA-induced protein–protein interactions. To answer this question, it is important to determine whether RNA or protein plays a direct role in the interaction. Hence, for direct RNA-protein interactions, removal of the RNA by RNase treatment abolishes the interaction. In indirect RNA-protein interactions, the bound RNA is not affected by RNase treatment and RNase treatment only abolishes the interaction. Interestingly, some interactions are conditional, e.g. the NS1 protein from human stems of HAV was found to interact with the PA subunit of the viral polymerase, whilst that from swine stems of HAV interacts only in the absence of the PA subunit (Collins et al., 2002). It appears that the composition of the viral polymerase affects the interaction. Some indirect RNA-protein interactions are abolished by RNase treatment, in which case the interaction occurs via the bound RNA. The interaction may be abolished by RNase treatment due to conformational changes in the protein, which may then unmask the PDZ-binding domain or may render it inaccessible. Some interactions may not be abolished by RNase treatment, in which case the interaction is via the bound RNA. These interactions are abolished by RNase treatment due to conformational changes in the protein, which may then unmask the PDZ-binding domain or may render it inaccessible. Some interactions may not be abolished by RNase treatment, in which case the interaction is via the bound RNA.

**Table 1. NS1-interacting proteins**

<table>
<thead>
<tr>
<th>Group</th>
<th>Protein</th>
<th>Site of interaction in NS1</th>
<th>Method of identification*</th>
<th>RNA dependency</th>
<th>Reference</th>
</tr>
</thead>
<tbody>
<tr>
<td>Nuclear proteins/RNA metabolism</td>
<td>CPSF30</td>
<td>ED</td>
<td>Y2H (CoIP)–CoCry</td>
<td>ND</td>
<td>Das et al. (2008), Nemeroff et al. (1998)</td>
</tr>
<tr>
<td></td>
<td>PABPII</td>
<td>ED</td>
<td>Y2H (GST-PD)</td>
<td>No</td>
<td>Chen et al. (1999)</td>
</tr>
<tr>
<td></td>
<td>NFX1, Rac1 (+ Nup98)</td>
<td>ED + RBD</td>
<td>GST-PD</td>
<td>No</td>
<td>Satterley et al. (2007)</td>
</tr>
<tr>
<td></td>
<td>p15</td>
<td>ED</td>
<td>GST-PD</td>
<td>No</td>
<td>Satterley et al. (2007)</td>
</tr>
<tr>
<td></td>
<td>E1B-AP5</td>
<td>ED + RBD</td>
<td>GST-PD</td>
<td>Yes</td>
<td>Satterley et al. (2007)</td>
</tr>
<tr>
<td></td>
<td>NS1-BP (nuclear protein)</td>
<td>NS1</td>
<td>Y2H (GST-PD)</td>
<td>ND</td>
<td>Wolff et al. (1998)</td>
</tr>
<tr>
<td></td>
<td>Nucleolin</td>
<td>RBD</td>
<td>Phage display</td>
<td>ND</td>
<td>Murayama et al. (2007)</td>
</tr>
<tr>
<td></td>
<td>RNA helicase A (DHx9)</td>
<td>NS1</td>
<td>StrepTag-NS1</td>
<td>Yes</td>
<td>Lin et al. (2012)</td>
</tr>
<tr>
<td></td>
<td>ADAR1</td>
<td>RBD</td>
<td>Y2H</td>
<td>No</td>
<td>de Chassey et al. (2013), Tawaratsumida et al. (2014)</td>
</tr>
<tr>
<td></td>
<td>RNA helicase DDX21</td>
<td>RBD</td>
<td>GST-PD</td>
<td>No</td>
<td>Chen et al. (2014)</td>
</tr>
<tr>
<td></td>
<td>hnRNP-U</td>
<td>NS1</td>
<td>Tandem affinity purification</td>
<td>ND</td>
<td>Pichlmair et al. (2012)</td>
</tr>
<tr>
<td>Cytosolic/ribosomal proteins</td>
<td>elf4GI</td>
<td>Region 1–113 of NS1</td>
<td>GST-PD (CoIP)</td>
<td>No</td>
<td>Aragón et al. (2000)</td>
</tr>
<tr>
<td></td>
<td>hStaufen</td>
<td>NS1</td>
<td>Y2H (GST-PD)</td>
<td>No</td>
<td>Falcón et al. (1999)</td>
</tr>
<tr>
<td></td>
<td>PABPII</td>
<td>RBD</td>
<td>CoIP</td>
<td>No</td>
<td>Burgui et al. (2003)</td>
</tr>
<tr>
<td></td>
<td>NS1-1-</td>
<td>NS1</td>
<td>Y2H (GST-PD)</td>
<td>ND</td>
<td>Wolff et al. (1996)</td>
</tr>
<tr>
<td></td>
<td>Importin-α</td>
<td>RBD</td>
<td>GST-PD</td>
<td>ND</td>
<td>Melén et al. (2007)</td>
</tr>
<tr>
<td>IFN system and signalling pathways</td>
<td>p85-β subunit of PI3K</td>
<td>ED</td>
<td>GST-PD, CoIP, CoCry</td>
<td>No</td>
<td>Hale et al. (2010)</td>
</tr>
<tr>
<td></td>
<td>RIG-I</td>
<td>NS1</td>
<td>CoIP, colocalization</td>
<td>Yes</td>
<td>Pichlmair et al. (2006)</td>
</tr>
<tr>
<td></td>
<td>TRIM-25</td>
<td>NS1</td>
<td>CoIP</td>
<td></td>
<td>Gack et al. (2009)</td>
</tr>
<tr>
<td></td>
<td>PKR</td>
<td>ED (123–127)</td>
<td>GST-PD</td>
<td></td>
<td>Li et al. (2006a), Min et al. (2007)</td>
</tr>
<tr>
<td></td>
<td>PACT</td>
<td>NS1</td>
<td>GST-PD</td>
<td>Yes</td>
<td>Li et al. (2006a), Tawaratsumida et al. (2014)</td>
</tr>
<tr>
<td></td>
<td>Dah, PSD-95</td>
<td>C terminus (PDZ-binding motif)</td>
<td>PDZ domain array analysis</td>
<td>ND</td>
<td>Obenauer et al. (2006)</td>
</tr>
<tr>
<td></td>
<td>Scribble, Dlg1, MAGI-1,-2 and -3</td>
<td>C terminus (PDZ-binding motif)</td>
<td>GST-PD</td>
<td>ND</td>
<td>Golebiowski et al. (2011)</td>
</tr>
<tr>
<td></td>
<td>Ubc9</td>
<td>NS1</td>
<td>Y2H, GST-PD</td>
<td>ND</td>
<td>Xu et al. (2011)</td>
</tr>
<tr>
<td></td>
<td>Herc5</td>
<td>RBD</td>
<td>GST-PD, CoIP</td>
<td>No</td>
<td>Zhao et al. (2010)</td>
</tr>
<tr>
<td>Viral proteins</td>
<td>NP</td>
<td>RBD</td>
<td>CoIP</td>
<td>No</td>
<td>Marión et al. (1997b), Robb et al. (2011)</td>
</tr>
<tr>
<td></td>
<td>NP and PA</td>
<td>NS1</td>
<td>CoIP</td>
<td>ND</td>
<td>Kuo &amp; Krug (2009)</td>
</tr>
</tbody>
</table>

**ND,** Not determined.

*Y2H, Yeast two-hybrid; GST-PD, GST (glutathione S-transferase) pull-down; CoIP, co-immunoprecipitation; CoCry, crystal structure of the complex. When two methods are indicated, the first method was used for identification, whilst the second method was used for confirmation.
interactions or whether they rely primarily on dsRNA–protein interactions. Indeed, binding of NS1 to specific dsRNA motifs may well drive its association with other dsRNA-binding proteins that share the same targets. NS1 may thereby alter the activity of these proteins or compete with them for the target dsRNA. Although a number of studies have probed the RNA dependency of the interaction by RNase treatment of the complex, one may question whether RNase A can readily access and degrade a protein-bound dsRNA.

### Multiple activities in the infected cell

Through its multiple interactions in the infected cell, NS1 exerts several activities which all favour viral multiplication. Schematically, these activities are directed towards (i) the antiviral response, (ii) the metabolism of cellular mRNAs and (iii) viral RNAs (Fig. 5). NS1 expression by itself seems to be detrimental to the cell, as judged by the difficulties encountered in establishing cell lines that constitutively express NS1 (van Wielink et al., 2011).

### Table 2. Amino acids involved in NS1 interactions and phenotype-linked substitutions

<table>
<thead>
<tr>
<th>Amino acid/position</th>
<th>Localization</th>
<th>Substitution</th>
<th>Properties</th>
<th>Role</th>
<th>Reference</th>
</tr>
</thead>
<tbody>
<tr>
<td>Arg35</td>
<td>RBD helix x2</td>
<td>R35A</td>
<td>Not viable</td>
<td>RBD dimerization</td>
<td>Lalime &amp; Pekosz (2014)</td>
</tr>
<tr>
<td>Arg38</td>
<td>RBD helix x2</td>
<td>R38A</td>
<td>Attenuated</td>
<td>RNA binding, NLS</td>
<td>Lalime &amp; Pekosz (2014), Newby et al. (2007)</td>
</tr>
<tr>
<td>Glu40</td>
<td>RBD helix x2</td>
<td>Q40K</td>
<td>Increased IFN sensitivity</td>
<td>RNA binding, NLS, ISG15 linkage</td>
<td>Lalime &amp; Pekosz (2014), Yin et al. (2007), Jiao et al. (2008)</td>
</tr>
<tr>
<td>Lys41</td>
<td>RBD helix x2</td>
<td>K41A</td>
<td>Attenuated</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Ser42</td>
<td>RBD helix x2</td>
<td>S42P</td>
<td>Thermosensitive, attenuated</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Ala60</td>
<td>RBD helix x3</td>
<td>A60G</td>
<td>Increased IFN sensitivity</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Lys62</td>
<td>RBD helix x3</td>
<td>K62N</td>
<td>Thermosensitive</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Leu69</td>
<td>RBD helix x3</td>
<td>L69R</td>
<td>Increased IFN sensitivity</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Deletion 80–84</td>
<td>Interdomain linker</td>
<td></td>
<td>Enhanced pathogenicity</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Tyr89</td>
<td>ED, strand β1</td>
<td>D92E, D92Y</td>
<td>p85-β interaction</td>
<td>Modulates virulence</td>
<td>Hale et al. (2010), Seo et al. (2002), Wu et al. (2014)</td>
</tr>
<tr>
<td>Asp92</td>
<td>ED, end of strand β1</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>95–100</td>
<td>ED, helix x4</td>
<td>F103S</td>
<td>Reduced CPSF30 binding</td>
<td>CPSF30 interaction</td>
<td>Hale et al. (2010), Kochs et al. (2007), Twu et al. (2007)</td>
</tr>
<tr>
<td>Phe103</td>
<td>ED, between helix x4 and strand β2</td>
<td>M106I</td>
<td>Reduced CPSF30 binding</td>
<td>CPSF30 interaction</td>
<td>Kochs et al. (2007), Twu et al. (2007)</td>
</tr>
<tr>
<td>Met106</td>
<td>ED, between helix x4 and strand β2</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>123–127</td>
<td>ED, between strands β3 and β4</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Ala132</td>
<td>ED, strand β4</td>
<td>A132T</td>
<td>Thermosensitive, attenuated</td>
<td></td>
<td>Hatada et al. (1990)</td>
</tr>
<tr>
<td>Leu141–Leu146</td>
<td>ED, strand β5</td>
<td></td>
<td></td>
<td>Nuclear export signal</td>
<td>Li et al. (1998)</td>
</tr>
<tr>
<td>Glu142</td>
<td>ED, strand β5</td>
<td></td>
<td></td>
<td>p85-β interaction</td>
<td>Hale et al. (2010)</td>
</tr>
<tr>
<td>Ile145</td>
<td>ED, strand β5</td>
<td></td>
<td></td>
<td>p85-β interaction</td>
<td>Hale et al. (2010)</td>
</tr>
<tr>
<td>Ala149</td>
<td>ED, strand β5</td>
<td>A149V</td>
<td>Attenuated</td>
<td></td>
<td>Li et al. (2006b)</td>
</tr>
<tr>
<td>Pro164</td>
<td>ED, between strand β6 and helix x5</td>
<td>A149V</td>
<td></td>
<td>p85-β interaction</td>
<td>Hale et al. (2010)</td>
</tr>
<tr>
<td>180–187</td>
<td>ED, helix x5</td>
<td>W187C</td>
<td>Increased IFN sensitivity</td>
<td>CPSF30 interaction</td>
<td>Das et al. (2008)</td>
</tr>
<tr>
<td>Trp187</td>
<td></td>
<td>R200I</td>
<td>Increased IFN sensitivity</td>
<td>ED dimerization</td>
<td>Wu et al. (2014)</td>
</tr>
<tr>
<td>Arg200</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Thr215</td>
<td>ED</td>
<td>K217T</td>
<td>Increased IFN sensitivity</td>
<td>Phosphorylation site</td>
<td>Hale et al. (2009)</td>
</tr>
<tr>
<td>Lys217</td>
<td>ED</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Lys219 and Lys221</td>
<td>ED</td>
<td></td>
<td>SUMOylation site</td>
<td></td>
<td></td>
</tr>
<tr>
<td>227–230</td>
<td>ED</td>
<td></td>
<td>PDZ-binding motif</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
Conversely, NS1 is virtually indispensable for the virus, whose replication potential is reduced dramatically by mutations in the NS1 gene, especially if these invalidate the RBD (Donelan et al., 2003; Egorov et al., 1998; García-Sastre et al., 1998).

**IFN antagonism**

NS1 inhibits the antiviral defences of the infected cell, either through preventing the activation of IFN-inducing proteins or through inhibiting effector proteins. NS1 prevents RIG-I activation, which normally triggers the antiviral response upon being activated by dsRNAs (Mibayashi et al., 2007; Rehwinkel et al., 2010; Schmidt et al., 2009). It was also shown to prevent the antiviral response mediated by MDA5 in duck cells (Wei et al., 2014). In addition to its direct inhibition of RIG-I, NS1 also blocks the RIG-I signal transduction pathway through interacting with ubiquitin E3 ligases TRIM-25 and Riplet, and preventing RIG-I ubiquitination (Gack et al., 2009; Rajsbaum et al., 2012). In addition to the RIG-I pathway, NS1 prevents the activation of protein kinase PKR, which is also a major component of the antiviral defence (Li et al., 2006a; Lu et al., 1995; Min & Krug, 2006; Min et al., 2007). Altogether, the blockade of these multiple pathways leads to several defects in the activation of the IFN system, preventing the nuclear translocation of transcription factors NFκB and IFN-regulatory factor 3 (Rand et al., 2014; Talon et al., 2000; Wang et al., 2000). NS1 also acts downstream, antagonizing the effects of an activated IFN system: through sequestering dsRNAs, it prevents the activation of 2’–5’ oligo-adenylate synthetases and of RNase L involved in the antiviral defence (Min & Krug, 2006). Thanks to this downstream activity of NS1, viral replication can take place even in cells that have been set in an active antiviral state by a prior IFN treatment (Min & Krug, 2006; Newby et al., 2007). NS1 therefore acts on a number of strategic, highly connected proteins (de Chassey et al., 2013), which are also targeted by other specialized proteins of DNA or RNA viruses (de Chassey et al., 2013; Pichlmair et al., 2012; Versteeg & García-Sastre, 2010). However, NS1 is not the sole actor involved in evading the IFN response and recent studies show that the entire genome of influenza viruses contributes to the IFN antagonism (Liedmann et al., 2014; Pérez-Cidoncha et al., 2014). In addition to its IFN-antagonist activities, NS1 activates PI3K, through binding its regulatory subunit p85-β and blocking its allosteric control of the catalytic subunit p110. This results in multiple effects that can vary depending on the step of the viral cycle, from inhibition of apoptosis to induction of type I IFNs (Hale et al., 2010; Hrincius et al., 2011).

**NS1 activity towards viral and cellular RNAs**

The RNA-binding specificity of NS1 towards positive-strand viral RNAs suggests that some NS1 activities might specifically target viral mRNAs and perhaps also cRNAs. With more than 26 species of virus-derived RNAs in the infected cell (Chiang et al., 2008), the virus may require specialized activities, first to distinguish virus derived from cellular RNAs, then to sort viral genomic RNAs from positive-strand RNAs, and perhaps to distinguish cRNAs, which must stay in the nucleus, from cytoplasm-bound mRNAs. NS1 may well play an important role here and at least some of its activities have been well characterized. The importance of the RNA-binding activity of NS1 is supported by the observation that mutations that invalidate the RBD are much more detrimental to the virus than truncations of the ED (Donelan et al., 2003; Egorov et al., 1998; Jiao et al., 2008).

**NS1 regulates the activity of the viral polymerase.**

Earlier studies showed that NS1 could enhance RNA synthesis in an in vitro assay using RNP cores purified from virions, apparently by preventing the dissociation of RNA polymerase from its template (Shimizu et al., 1994). More recent data confirm that NS1 regulates the synthesis of viral RNA and the transcription/replication balance (Falcón et al., 2004; Min et al., 2007; Wang et al., 2010): (i) mutant viruses with truncations in the ED showed a thermosensitive (ts) phenotype, with a reduced production of viral RNA at the non-permissive temperature (Falcón et al., 2004), and (ii) complementation by NS1 of a minireplicon assay system generally resulted in a reduced production of
the three species of viral RNAs (i.e. mRNA, cRNA and viral RNA) and modified their ratios in a manner depending on the NS1 variant analysed (Wang et al., 2010). Although the underlying mechanism is unknown, it is tempting to speculate that this is achieved through interactions of NS1 with: (i) the nucleoprotein and the viral RNPs (see above; Robb et al., 2011), and (ii) RNA motifs in the positive-strand viral RNAs (Marc et al., 2013). In addition, NS1 may also regulate the synthesis of viral RNAs through inhibiting cellular factors that themselves regulate the activity of the viral polymerase, such as RNA helicase DDX21 or the heterogeneous nuclear RNP hnRNP-U (Chen et al., 2014; Pichlmair et al., 2012).

**Splicing and mRNA export.** It had long been known that cellular mRNAs in influenza-infected cells could not reach the cytoplasm and were degraded in the nucleus (Fortes et al., 1994; Katze & Krug, 1984), and experiments pointed to the key role of NS1. NS1 was shown to inhibit the cleavage and polyadenylation of host pre-mRNAs notably through its interaction with the poly(A)-containing mRNAs (Chen et al., 1998; Qiu & Krug, 1994) and with CPSF30 (Nemeroff et al., 1998). NS1 of H3N2 virus A/Udorn/72 was also shown to inhibit pre-mRNA splicing both in vivo and in vitro through its association with the spliceosome and with U6 small nucleolar RNA (Lu et al., 1994; Qiu et al., 1995).

Through these multiple interactions with nuclear factors involved in mRNA maturation, NS1 blocks the splicing and nucleo-cytoplasmic export of cellular mRNAs (Lu et al., 1994; Robb et al., 2010; Satterly et al., 2007), thereby contributing to the host’s shut-off. In contrast to its splice inhibitor activity on polymerase II transcripts, NS1 has only a modest effect, if at all, on the splicing of the viral M mRNA (Robb & Fodor, 2012; Salvatore et al., 2002), whilst it does not appear to impact the splicing of its own mRNA (Robb et al., 2010): in two cellular models using two distinct viruses, it was observed that the spliced NS2 mRNA represented ~15% of the total NS mRNA and remained constant during the viral cycle. What is the basis of NS1’s differential splice inhibitor activity on polymerase II transcripts versus viral mRNAs? One possibility is that the splicing machinery acting on viral mRNAs is different from that acting on cellular mRNAs, rendering it immune to NS1 inhibition. This is supported by the observation that the viral polymerase interacts with a complex formed by the spliceosomal factors RED and SMU1 (Fournier et al., 2014). If, on the contrary, the splicing machinery is the same, another possibility is that viral mRNAs in some way prevent NS1 from inhibiting their own splicing. This could be achieved through the specific recognition by NS1 of sequence motifs in viral mRNAs (see above): they all harbour the RBD-binding motif AGC(A/G)AAAG in their 5’ UTR, whilst NS1 mRNA additionally contains two putative RBD-binding motifs UG(A/C)UUGAAG in its 3’ UTR (Fig. 4a). Yet another possibility is that splicing of M1 and NS1 mRNAs is primarily regulated by the RNA structure motif that encompasses the splice acceptor site (see above) in both mRNAs (Dela-Moss et al., 2014). NS1 was suggested to be involved in the nuclear export of viral mRNAs (Schneider & Wolff, 2009), in a manner similar to the Rev protein of human immunodeficiency virus, which allows the export of intron-containing viral RNAs (Daugherty et al., 2010). This hypothesis has not been explored thoroughly, although it is in agreement with the fact that NS1 interacts with both viral mRNAs and several factors of the mRNA export pathway (Satterly et al., 2007; Wang et al., 2008; York & Fodor, 2013). In addition, under this hypothesis, NS1 might perhaps participate in sorting the two types of viral positive-strand RNAs (e.g. NP-bound cRNAs, which must stay in the nucleus, versus mRNAs that are associated with nuclear export factors) (York & Fodor, 2013).

**Translation.** In spite of the fact that viral mRNAs resemble cellular mRNAs at their 5’ end, harbouring a 5’-cap structure stolen from cellular mRNAs, they are nevertheless translated preferentially. This results from the binding of NS1 to the AGCAAAAG motif in their 5’ UTR (de la Luna et al., 1995; Marc et al., 2013; Park & Katze, 1995), within a complex associating PABP and translation initiation factor eIF4G1 (Burgui et al., 2003). The critical role of NS1 in enhancing the translation of viral mRNAs is supported by the observation that NS1 ts mutants show a much reduced synthesis of viral proteins at the non-permissive temperature (Hatada et al., 1990). In addition to its effect in the translation of full-length products, whether NS1 may regulate the synthesis of alternative translation products of the PB1 and PA-mRNAs (PB1-N40, PB1-F2, PA-X, PA-N155 and PA-N182) is unknown, although it may provide yet another means to modulate the viral cycle, as was shown for the nsp1β protein of porcine reproductive and respiratory syndrome virus (Li et al., 2014). However, an issue that to our knowledge has not been explored is whether NS1 might somehow assist the viral polymerase in the sequence-specific protection of viral mRNAs from cleavage by the cap-snatching mechanism (Shih & Krug, 1996).

**Silencing and microRNAs.** As well as its activities on viral and cellular mRNAs, NS1 was shown to bind small interfering RNA (siRNAs) and to suppress RNA silencing in both Drosophila and plants, provided it had a functional RBD (Bucher et al., 2004; Li et al., 2004). Furthermore, as with Ebola virus VP35 and vaccinia virus E3L, NS1 was shown to suppress short hairpin RNA-induced RNAi in some mammalian cells (Haasnoot et al., 2007). Intriguingly, we noticed that the mfold-predicted structures of the RNA aptamers that we raised against recombinant protein NS1 (Marc et al., 2013) are reminiscent of the pre-mRNA structures (Starega-Roslan et al., 2011). Altogether, these data lead us to speculate that NS1, like the structurally related protein B2 of Flock House virus, may suppress IFN-independent antiviral activities through interacting with dsRNAs, siRNAs, and RNase III enzymes Dicer and Drosha (Qi et al., 2012; Shapiro et al., 2014).
NS1 polymorphisms, pathogenicity and host range

Along with HA and NA, NS1 contributes substantially to the variability of the virus genome (Obenauer et al., 2006), and indeed it has been shown to be particularly permissive to mutations (Heaton et al., 2013). Even if they result from a biased sampling, the >25 000 NS1 sequences in GenBank reveal a large number of polymorphisms in the peptide sequence, including length variations.

Whilst NS1 is most often a 230 aa protein, mutations that either suppress the stop codon at position 231 or create a premature stop codon result in length variations. For instance, from the late 1940s until the middle of the 1980s, NS1 of human influenza A viruses harboured the C-terminal, 7 aa extension 231RRNKMAD237. Conversely, NS1 of pdmH1N1 (2009), like that of most swine viruses, has only 219 aa (Fig. 6). In addition to truncations, short internal deletions are also observed: from 2000, NS1 of almost all highly pathogenic H5N1 viruses harbour a deletion of aa 80–84 in the linker region. More extensive deletions, which can be obtained through reverse genetics (Cauthen et al., 2007; Egorov et al., 1998; García-Sastre et al., 1998), are rarely observed in field isolates, confirming the requirement of NS1 for normal growth of the virus. Such viruses, such as those expressing an invalidated RBD NS1 (Donelan et al., 2003; Newby et al., 2007), replicate poorly, and show a dramatic decrease in their pathogenicity and in their ability to counter the IFN system.

As well as length variations, several polymorphisms are observed in the peptide sequence of NS1, in spite of the constraint resulting from the overlapping ORF of NEP after NS1 codon 168 (Fig. 4a). These can arise from at least three distinct origins: (i) adaptation to a new host (Bhatt et al., 2013; Noronha et al., 2012; Vijaykrishna et al., 2011), (ii) adaptation to a new viral gene constellation following reassortment (Neverov et al., 2014), and (iii) random variations that are favoured by the fact that the segment is isolated in a given species and no longer participates in the influenza virus gene pool of wild waterfowl. On the one hand, it seems that each host species somehow imprints NS1 with characteristic sequence signatures (Noronha et al., 2012; Obenauer et al., 2006). On the other hand, isolated evolution in a given species (such as NS1 of classical swine viruses) does not necessarily lead to a restricted host range. Different viruses expressing distinct variants of NS1 can successfully infect a given species: more than four different NS1 can be easily distinguished in influenza A viruses infecting humans (seasonal H1N1 prior to 2009, pdmH1N1, seasonal H3N2, and avian H5N1 and H7N9 viruses), which indicates that several distinct NS1 variants are fully functional in a given host (Knepper et al., 2013).

Several of these polymorphisms are depicted in Fig. 6. Allele B NS1 (avBH5N3) is easily recognized by the peptide motif 211LLSMRD217, along with >40 other substitutions. It is found exclusively in avian viruses (with rare exceptions), which probably reflects the fact that it harbours some avian-specific host range determinants and indeed it was shown to confer specific properties to the virus phenotype (Adams et al., 2013; Munir et al., 2011). Each of the two alleles of NS1 is subdivided into proteotypes (Obenauer et al., 2006), or variant types (Noronha et al., 2012), some of which seem to be associated exclusively with a species or a species group (Obenauer et al., 2006). The first eight sequences in Fig. 6 illustrate the diversity of proteotypes within allele A. NS1 of the 1918 virus shares several characteristics with that of avian viruses (avAH5N1 and avAH6N1), whilst that of pdmH1N1 is typical of North-American swine viruses (lines swU5 and pdmH1N1). NS1 of human H3N2 viruses evolved from that of seasonal human viruses of the twentieth century (H1N1 then H2N2) and contains the characteristic motif 186EWNDN190, which appeared in 1960 in H2N2 viruses. Some substitutions between variant proteotypes change amino acids involved in NS1 interactions with host partners. They either abolish or restore these interactions, probably with some host range specificity, thus modulating the spectrum of NS1 activities and the viral phenotype. For instance, substitutions of aa 103 and 106 modulate the stability of the NS1–CPSF30 interaction, resulting in an altered virus phenotype (Das et al., 2008; Kochs et al., 2007; Twu et al., 2007). Another example is the sequence of the PDZ-binding motif, which is relatively host specific (227ESEV230 in avian viruses, 227RSKV230 in human viruses): substitutions result in variations in the set of PDZ ligands and in the viral phenotype (Jackson et al., 2008; Obenauer et al., 2006; Soubies et al., 2010). Altogether, these phenotype-associated polymorphisms do not suffice to confer a high-pathogenicity phenotype, which generally results from the acquisition of a multibasic cleavage site in the HA of avian influenza A viruses: there is no known naturally occurring mutation in NS1 that increases dramatically, by itself, the viral pathogenicity. Rather, even if the spectrum of its activities can vary, the biological functionality of NS1 seems to be generally conserved, probably resulting from a conservation of its core functions, and only in rare instances have field viruses been isolated that harboured mutations conferring an attenuated phenotype (Cauthen et al., 2007; Jiao et al., 2008; Li et al., 2006b).

However, numerous mutant viruses harbouring loss-of-function substitutions in NS1 have been obtained experimentally through either site-directed or high-throughput random approaches (Wu et al., 2014). Altogether, the combined sets of naturally occurring and experimentally obtained NS1 mutants are useful in deciphering the structure–activity relationships of NS1 (see Table 2).

Interestingly, a comparison including the most distantly related variants from bats reveals highly conserved regions (Fig. 6). The latter include not only the basic amino acids of helix α2, but also a number of highly conserved motifs that may underlie critical functions, such as 131KANFS135, 186EWNDN190. Of note, Trp187 in the latter motif is involved in ED dimerization.

Viral phenotype can therefore be modified by mutations in the NS segment and a fortiori by reassortment events that
change the NS1 allele or its proteotype. This is well illustrated by the complex genesis of Asian H5N1 viruses: whilst their precursor A/Goose/Guangdong/1/96 (H5N1) had an NS1 of allele B (Duan et al., 2008), several new clades emerged, first in 1997 in Hong Kong and again after 2000, which all possessed an NS1 of allele A (Duan et al., 2008; Vijaykrishna et al., 2008). This allele exchange probably contributed to broaden the virus host range, leading to the emergence of H5N1 viruses in mammals, including humans. In birds, reassortment events that change the allele of NS1 or its proteotype are not uncommon: homosubtypic isolates have been isolated that harbour quasi-identical genomic constellations that only differ by the NS1 allele (Briand et al., 2010). In mammals, several reassortments in swine flu viruses have contributed to the emergence and spread to humans of new H1N1 and H3N2

**Fig. 6.** Polymorphism of NS1 protein sequence. Structure elements are represented by grey rods (α-helices) or open rectangles (β-strands). Only differences relative to the consensus sequence (above and below) are shown (hyphens are deletions). Peptide sequences are representative of human seasonal viruses H3N2 (huH3N2) and H1N1 before 2009 (huH1N1), pdmH1N1, classical swine (swUS), avian allele A (avAH6N1 & avAH5N1) and avian allele B (avBH5N3), human 1918 virus (hu1918), human H7N9 (huH7N9), and of a recently characterized bat virus (batH18). GenBank accession numbers of the peptide sequences are (in their order of alignment): ACU80272, ADV69058, ACP41110.1, AF153262, AER57956, AF333238, AGX00869, AHD24960, ADA69390 and AGX84939. Alignment was performed with the ‘showalign’ tool of the software suite EMBoss (Rice et al., 2000).
viruses harbouring a typically swine-origin NS segment (Nelson et al., 2012; Smith et al., 2009).

**Novel antiviral strategies targeting NS1**

Truncations, deletions or substitutions that invalidate NS1 dramatically reduce the replication potential and pathogenicity of the virus, which suggests the possibility of targeting NS1 to combat viral infection. Current strategies rely on vaccines and antiviral compounds. Whilst most influenza vaccines for humans, pigs or horses generally consist of inactivated viruses or purified subunits, live-attenuated vaccines present the twofold advantage of inducing a better cellular response and of eliciting a local immunity following intranasal administration. NS1 stands as a good candidate for virus attenuation and several studies have demonstrated the vaccinal efficiency of genetically engineered live-attenuated viruses in which NS1 or its ED is partially or totally deleted (Ferko et al., 2004; Kappe et al., 2012; Mössler et al., 2013; Pica et al., 2012; Zhou et al., 2010), or with an engineered NS segment harbouring deoptimized codons (Nogales et al., 2014). However, these promising candidates will probably require further modifications, notably to prevent reassortment events with circulating viruses (Zhou et al., 2010).

Antiviral compounds targeting NS1 may enrich the arsenal of drugs against influenza, which is currently limited to neuraminidase inhibitors and M2-antagonists adamantanes (Palmer, 2011). Several approaches are being explored (reviewed by Engel, 2013): screening of compounds that block the biological activities of NS1 in cell-based assays (Basu et al., 2009; Mata et al., 2011), bioinformatic searches of conserved binding sites and computer-assisted drug design (Darapaneni et al., 2009), screening of compounds that measurably alter the physical properties of recombinant NS1 (Barman et al., 2014) or block the NS1–RNA interaction in vitro (Cho et al., 2012), antisense oligonucleotides that target the NS1 segment (Wu et al., 2008), or NS1-directed DNA aptamers (Woo et al., 2013). These attempts have already identified some compounds that are active in vitro at micromolar or submicromolar concentrations (Engel, 2013) and that need further optimization in order to be fully operational.

**Conclusion and perspectives**

Our broad knowledge about NS1 is to some extent measured by the number of NS1-related publications: we know its atomic structure, its viral and cellular partners, its biological activities, and thousands of sequences from viruses isolated in various hosts. We should, however, acknowledge that to a certain extent this knowledge is superficial, and that our understanding is patchy and far from complete. Do we really know its structure and the underlying molecular mechanics? Whilst the structure of RBD is well defined, several issues are left unanswered with regard to the ED and the full-length protein. How exactly does the ED dimerize and how is it regulated? How do the two domains interact with each other? Does RNA binding induce a conformational change that modifies the quaternary state of the protein, as was observed for other RNA-binding proteins (Kowalski et al., 2011)? Are there other, yet unknown, biological activities associated with NS1 and particularly with its ED? Is it possible to assess the adaptation level to a given host species and predict the risk associated? Answering these questions will need multidisciplinary approaches, from the biophysical characterization of the molecule to the evaluation of its biological activities in cells, from the individual host to the population level, using all the tools available in molecular biology, biophysics and bioinformatics.

**Acknowledgements**

I thank Nadia Naffakh for her critical review of the manuscript and my colleagues for their encouraging support. This work was funded in part by Institut Carnot Santé Animale.

**References**


Influenza A virus NS1


evade recognition by the host viral RNA sensor RIG-I. *Cell Host Microbe* 5, 439–449.


Li, S., Min, J. Y., Krug, R. M. & Sen, G. C. (2006a). Binding of the influenza A virus NS1 protein to PKR mediates the inhibition of its activation by either PACT or double-stranded RNA. *Virology* 349, 13–21.

Influenza A virus NS1


The influenza fingerprints: NS1 and M1 proteins contribute to specific host cell ultrastructure signatures upon infection by different influenza A viruses. *Virology* 432, 204–218.


