Since January 2020 Elsevier has created a COVID-19 resource centre with free information in English and Mandarin on the novel coronavirus COVID-19. The COVID-19 resource centre is hosted on Elsevier Connect, the company's public news and information website.

Review

CellPress

Bat-derived influenza-like viruses H17N10 and H18N11

Ying Wu[1], Yan Wu[1,2], Boris Tefsen[1], Yi Shi[1,3], and George F. Gao[1,2,3,4]

[1] CAS Key Laboratory of Pathogenic Microbiology and Immunology, Institute of Microbiology, Chineses Academy of Sciences, Beijing 100101, China
[2] University of Chinese Academy of Sciences, Chinese Academy of Sciences, Beijing 100049, China
[3] Research Network of Immunity and Health (RNIH), Beijing Institutes of Life Science, Chinese Academy of Sciences, Beijing 100101, China
[4] Office of Director-General, Chinese Center for Disease Control and Prevention (China CDC), Beijing 102206, China

Shorebirds and waterfowls are believed to be the reservoir hosts for influenza viruses, whereas swine putatively act as mixing vessels. The recent identification of two influenza-like virus genomes (designated H17N10 and H18N11) from bats has challenged this notion. A crucial question concerns the role bats might play in influenza virus ecology. Structural and functional studies of the two major surface envelope proteins, hemagglutinin (HA) and neuraminidase (NA), demonstrate that neither has canonical HA or NA functions found in influenza viruses. However, putative functional modules and domains in other encoded proteins are conserved, and the N-terminal domain of the H17N10 polymerase subunit PA has a classical structure and function. Therefore, potential genomic reassortments of such influenza-like viruses with canonical influenza viruses cannot be excluded at this point and should be assessed.

Influenza A and influenza A-like viruses

Influenza virus is a member of the Orthomyxoviridae family. There are three types of influenza virus based on its internal proteins of nucleoprotein and matrix protein, namely A, B, and C [1]. Among these, influenza A virus is the most prevalent pathogen for both humans and animals, causing the so-called seasonal flu. Influenza A virus was also the causative agent of four major pandemics, in other words the 1918 Spanish flu, 1957 Asian flu, 1968 Hong Kong flu, and the 2009 swine-origin pandemic flu (2009 pH1N1), as well as of the relatively milder pandemic, the 1977 Russian flu [2–4]. For the origin, genesis and ecology of influenza viruses, migratory birds (shorebirds and waterfowls) are regarded as reservoir hosts, providing a large pool of virus gene segments that can contribute to novel reassortant viruses [5–7]. Meanwhile, swine are believed to be 'mixing vessels' or at least intermediate hosts that influenza A viruses can utilize to 'jump' from poultry to humans [5].

Influenza A virus is an enveloped negative-stranded RNA virus, with a segmented genome of 8 pieces, which encode a total of 14 proteins [8]. There are three major envelope proteins embedded in the virus surface: hemagglutinin (HA), neuraminidase (NA), and matrix protein 2 (M2) [5]. HA is responsible for virus binding to susceptible cells, which harbor sialylated proteins (as virus receptors), eventually resulting in fusion to and entry into the cells. NA is a sialidase that enables mature viruses to be released from infected cells [5]. To date 16 HA subtypes and nine NA subtypes have been identified (Figure 1), and the different reassorted combinations are used to name the viruses, for example, the common seasonal flu viruses H1N1, H2N2, H3N2, and the sporadic human infections with H5N1 and the recent H7N9 [9–11].

In 2012, an astonishing story published in *Proc. Natl. Acad. Sci. USA* [12] caused some concerns because a new influenza virus genome (H17N10) was isolated from bats by next-generation sequencing (NGS). This raised a serious scientific and public health question as to the origin and evolution of influenza virus. If this genome can produce a real influenza virus that can cause human or animal infections, then the situation for influenza virus-caused diseases would be more complicated because bats are known to harbor many viruses and are regarded as a reservoir host for many human- and animal-infecting viruses, including the SARS (severe acute respiratory syndrome) coronavirus [13]. If this is the case, then our understanding of the influenza virus ecology will need to be rewritten. Therefore research on this NGS-identified novel genome is urgently needed for the sake of public health. Studies on the two surface envelope proteins, HA and NA (NA-like) [14–17], demonstrate that neither protein has the corresponding canonical influenza virus functions or structures [18]. Therefore the new genome does not represent a 'true' influenza virus, and it should be renamed influenza-like virus, at most. Whether or not the genomic segments will reassort with canonical influenza A virus genomes should be vigorously tested in the near future. Recently a similar virus genome, H18N11, was again identified by NGS, and neither HA nor NA have canonical structures or functions [19].

Corresponding authors: Wu, Y. (wuying@im.ac.cn); Gao, G.F. (gaof@im.ac.cn).
Keywords: bat-derived influenza-like virus; hemagglutinin (HA); neuraminidase (NA); PA; reassortment; H17N10; H18N11.

0966-842X/$ – see front matter
 http://dx.doi.org/10.1016/j.tim.2014.01.010

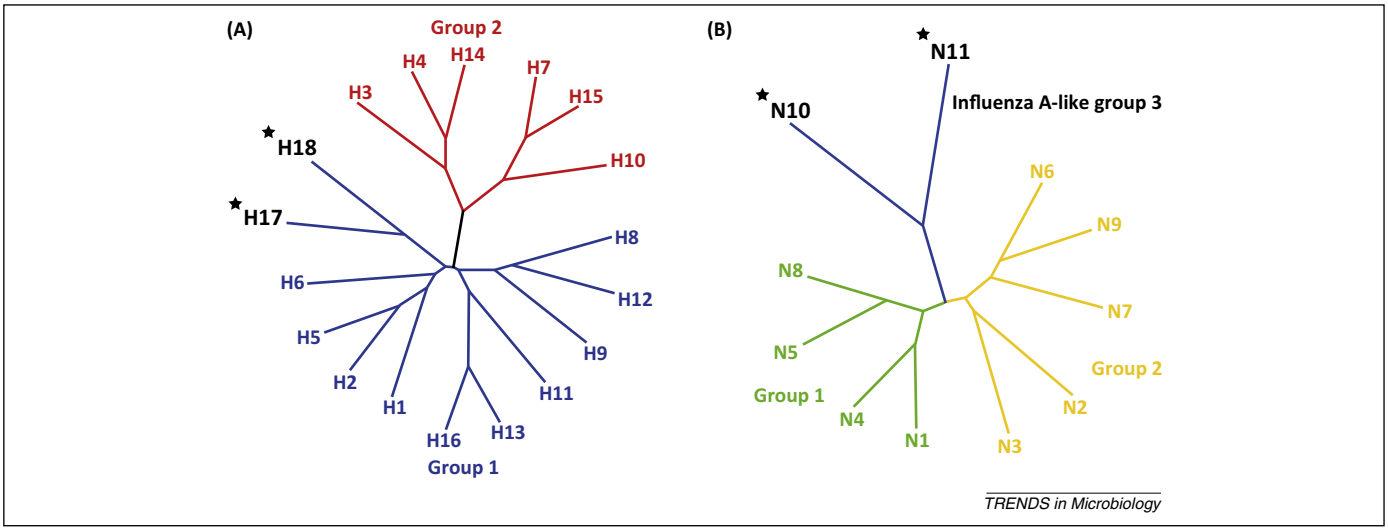

Figure 1. Phylogenetic trees of the hemagglutinin (HA) and neuraminidase (NA) genes of all known influenza A virus subtypes including H17/18 and NA-like N10/11. **(A)** The categorized HA molecules can be grouped into two groups, group 1 and group 2. The bat-derived H17 and H18 (each marked with a star) belong to group 1, displaying a typical sequence feature of group 1 HA molecules. **(B)** All the known NAs could also be grouped into two groups: group 1 and group 2. The bat-derived N10 and N11 (each marked with a star) do not belong to either group. They are NA homologs or could be called NA-like molecules. Here we propose that they be categorized as influenza A-like group 3.

In this review we summarize recent work on the functions and structures of both HA and NA derived from H17N10 and H18N11. In addition, all influenza proteins are examined using bioinformatics, and proteins from the six internal genes of the H17N10/H18N11 genomes are shown to conserve known structural and functional modules or domains. We also further discuss our recent work on the N-terminal domain of the polymerase subunit PA. We believe that the functions and structures of all the encoded proteins should be examined in detail to understand these influenza-like viruses better. Moreover, their reassortment potential should be assessed by reverse genetics experiments.

HAs H17 and H18 do not bind to the canonical sialic acid receptors

HA is the receptor binding protein of the influenza A virus, and is responsible for virus entry into host cells. Before the discovery of the H17 and H18 genes there were 16 subtypes of HA described, H1–H16. Based on their primary sequences, these HA molecules can be categorized into two groups (Figure 1): group 1 (H1, H2, H5, H6, H8, H9, H11, H12, H13, H16, H17, and H18) and group 2 (H3, H4, H7, H10, H14, and H15) [12,19–22]. Bat-derived H17 and H18 should be placed into group 1 based on their primary sequences (Figure 1) [12,16,17,19].

As expected from their primary sequences, the H17 and H18 structures resemble the structures of group 1 HAs rather than those of group 2. Previously solved HA structures demonstrate that there are group-specific features at sites where extensive conformational changes occur for HA activation, including the conformation of the interhelix loop and the rigid body orientation of the globular domain [22]. Taking H17 as an example, it displays a similar interhelix loop conformation to the HAs from group 1, and this is consistent with the phylogenetic analysis. Superimposition with other solved HA structures by means of the long central α-helices of HA2 revealed that the globular domains fall into three subgroups [16,23]: subgroup 1, including H1, H2, H5, and H9; subgroup 2, including H3, H7, and H14; and subgroup 3, consisting of H13, H16, H17 and H18. These differences may result from subtle variation in the interhelix loops among different HA subtypes and could signify different mechanisms during HA activation [16,23].

Influenza A virus enters susceptible cells through endocytosis after binding to cell surface receptors [24–26]. The receptors for influenza A viruses are sialic acids (SA) linked to cell surface glycolipids or glycoproteins [24–26]. After entering the cells, the virus fuses with the endosomal membrane under low pH and subsequently the genetic material is released into the cell [22,27–30]. Avian- and human-adapted influenza A viruses harbor distinct HA molecules that have different capacities to bind specifically linked SA-moieties. In Figure 2, the configurations of the α-2,3-linked SA and α-2,6-linked SA are shown. The molecular basis of the interaction of these specific receptors with the virus-derived HA has been extensively studied [22,27–30]. In the trimeric structure of the HA, resolved as early as in 1981 in an outstanding collaborative work by Wilson, Wiley, and Skehel [31], the HA head domain of each monomer was shown to be responsible for SA binding. We now know that there are three secondary elements and one base element involved directly in SA binding. The three secondary elements (the 130-loop, the 190-helix, and the 220-loop, in H3 numbering) form the edge portion, and four conserved residues (Y98, W153, H183, and Y195) form the base portion. These two portions usually form a shallow cavity to accommodate sialylated glycans. Typically, the SA moiety of the sialylated glycan forms several conserved hydrogen bonds with the 130-loop and the base residue Y98, and the remaining glycan moieties interact with the 220-loop or 190-helix.

Substitution of the residues in the three secondary elements is important for the HA protein to obtain avian or human SA receptor preference. Different HA subtypes

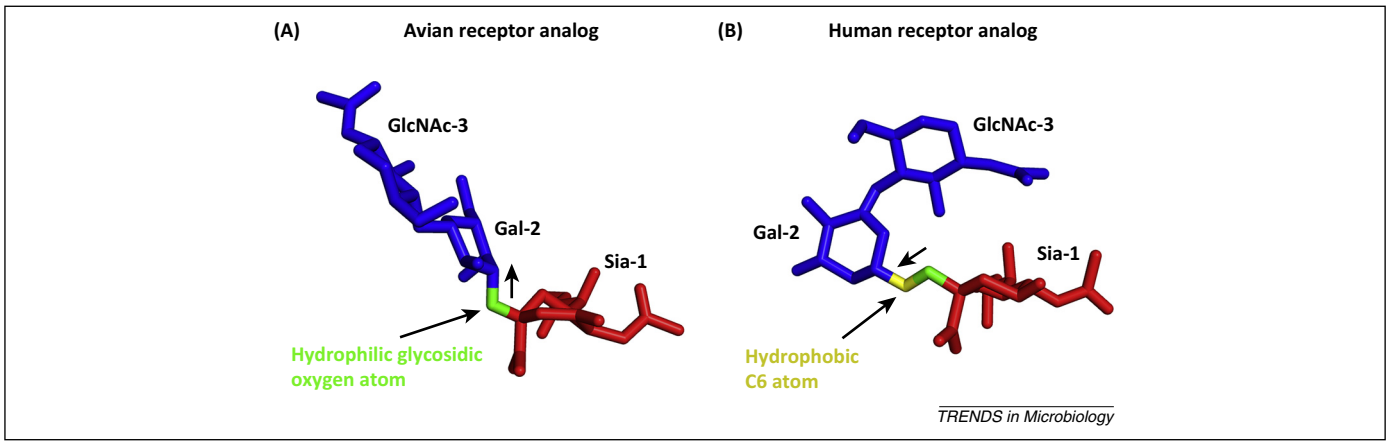

Figure 2. Configurations of avian and human receptor analogs. **(A)** The avian receptor analog. The Sia-1 (sialic acid, SA) is linked to Gal-2 (galactose) via an α2,3-glycosidic bond. The glycans extend forward and the hydrophilic glycosidic oxygen atom is exposed to the receptor binding site. **(B)** The human receptor analog. The Sia-1 is linked to Gal-2 via an α2,6-glycosidic bond. The glycans fold back and the hydrophobic C6 atom is exposed to the receptor binding site. In both (A) and (B) *N*-acetylglucosamine is abbreviated as GlcNAc-3; black arrows indicate the orientations of the glycosidic linkage.

use different substitutions to achieve this goal. For H2 and H3 HAs, Q226L and G228S substitutions in the 220-loop are responsible for the switch between avian and human receptor binding specificities [30,32] whereas, for H1 HA, different combinations of substitutions at residues 190 and 225 are important for the SA binding preference [33,34]. For H5 HA, a single Q226L substitution is enough to change the receptor binding preference [35,36]. By contrast, for H7 HAs, Q226L substitution is not solely responsible for the acquisition of human receptor binding, and other amino acid substitutions also contribute to the receptor binding switch, especially G186 V substitution [37,38].

Multiple lines of evidence, including surface plasmon resonance (SPR) experiments, MDCK cell binding assays, and glycan microarray analysis, revealed that the bat-derived H17 and H18 do not bind to canonical human or avian receptors [16,17,19]. This lack of canonical receptor binding is likely due to specific structural features in the putative receptor binding sites of H17 and H18 HA. In the H17 and H18 structures (Figure 3) there is no obvious cavity to accommodate the sialylated glycans, due to strong interactions among three secondary elements (130-loop, 190-helix, and 220-loop) through a hydrogen bond and salt bridge network formed by residues D136, Q190, H226, and D228 [16,19]. Furthermore, the negatively charged D136 in the 130-loop (all canonical influenza HAs have an uncharged threonine or serine at this position) could result in a charge conflict with the negatively charged carboxyl group of SA, which is unfavorable for SA receptor binding [16,19]. Moreover, residue 98 (usually a conserved tyrosine) in the base of the receptor binding site is a phenylalanine in H17 and H18, and this could also affect SA receptor binding capacity. Thus, these five key residues likely contribute to the lack of SA receptor binding by H17 and H18 and make the putative binding cavity a much smaller, pseudo-binding site (a 'closed' site) (Figure 3).

The possibility remains that we may have not detected the binding of H17 and H18 to canonical human or avian receptors using the soluble protein *in vitro* because it is plausible that a stronger interaction may occur through receptor clustering *in vivo*. It is also possible that H17 and H18 may bind to canonical influenza receptors very weakly, below the level of our detection (which is a limitation of the available methods). However, together with the extensive amino acid changes in the receptor binding site of H17 and H18 protein, it is likely that the putative bat influenza virus has acquired a different receptor, possibly protein-based. There are many examples of closely related viruses that switch between protein and SA receptors, for example, paramyxoviruses. The most common type of the paramyxovirus attachment protein recognizes SA receptors. It is called hemagglutinin–neuraminidase (HN) and is found on viruses such as Newcastle disease virus (NDV) and human parainfluenza virus 3 (hPIV3), and recognizes SA receptors. The structures of the globular heads of HN proteins display a conserved β-sheet propeller motif, which was identified originally in influenza virus NA, and an SA binding site located in the central cavity of the proteins [39,40]. Unlike HN, the hemagglutinin (H) of measles virus (MV), which also belongs to the Paramyxoviridae family, possesses an inactivated SA receptor binding site and recognizes specific proteins, such as signal lymphocyte-activating molecule (SLAM), CD46, and nectin-4 [41–43]. The structure of the globular heads of H protein still reveals a conserved β-sheet propeller motif, and the specific protein receptors bind to the side part of H protein with different orientations [44]. In particular, the immunoglobulin (Ig)-like SLAM molecule binds to the H protein mainly through interactions between two β-sheets. Interestingly, the bat-derived influenza-like virus H17/H18 and N10/N11 proteins have similar Ig-like fold elements [14–17,19], which possibly provide the β-sheets that could interact with putative protein-based receptors. If the H17 and H18 protein should lose its trimerization state and expose its Ig-like fold element, it might be able to bind a specific protein receptor. In this case, the bat-derived influenza-like virus would abrogate the need for an active N10 and N11. This hypothesis or notion needs further experimental work to be confirmed or rejected. Thus, the bat-derived H17 and H18 are unique among the characterized HAs and might use a different entry mechanism.

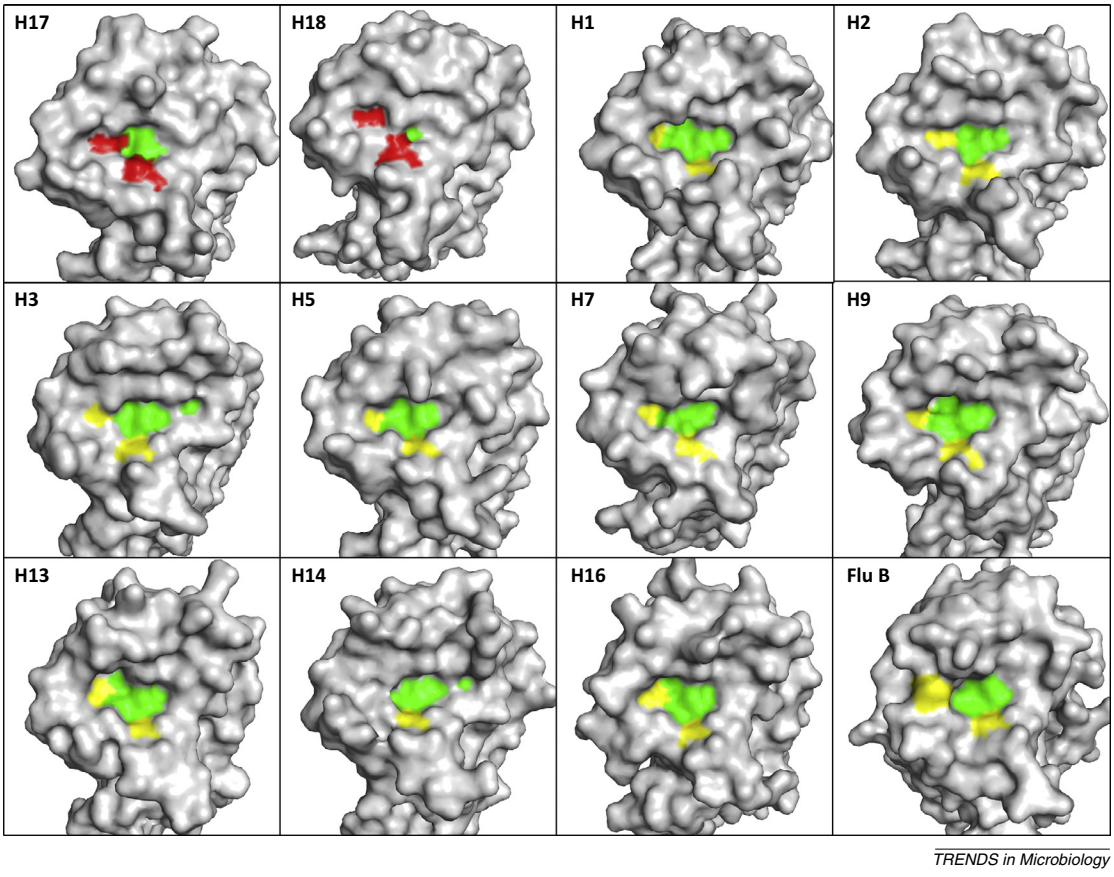

Figure 3. Receptor-binding 'head' domains of all the hemagglutinins (HAs) with known structures (including influenza A and B) showing the SA binding sites. This figure clearly shows that H17 and H18 have a smaller SA-binding cavity – with two acidic amino acids (labeled in red) in comparison to the non-charged amino acids in all the other HA molecules (labeled in yellow). Note, the base of the receptor binding groove is labeled in green; both H17 and H18 show a 'closed' conformation, and that of H18 is even smaller.

NA-like molecules N10 and N11 are not sialidases

NA is the second virus-surface envelope protein and is a sialidase, responsible for cleavage of SA from glycans on the host cell surface to release the emerging progeny virus, prevent virus aggregation, and help virus migration [45,46]. Therefore it has an opposite function from HA, SA-binding versus SA-releasing. If bat-derived HA does not bind SA, consequently it would have no need for a sialidase to help release the virus particles. There are nine subtypes (N1–N9) of NA identified so far, before the discovery of the NA-like N10 or N11 genes from bat. Similarly to HA, NA has also been classified into two groups, group 1 and group 2, based on primary sequences [47] (Figure 1). N1, N4, N5, and N8 are group 1 members and N2, N3, N6, N7, and N6 belong to group 2. It is clearly shown from the phylogenetic tree (Figure 1) that N10 and N11 are genetically far away from NA molecules found on canonical influenza A viruses (N1–N9), thereby forming a unique cluster, which we propose to name group 3.

The structure of the NA head domain is a homotetramer, each monomer of which is composed of six topologically identical β-sheets arranged in a propeller formation [48]. With the exception of N7, the structures of all NAs (including the newly discovered bat-derived NA-like molecules, N10 and N11) have been solved [49–53] (Figure 2). Analysis of the active site of N1-N9 shows that eight residues (R118, D151, R152, R224, E276, R292, R371, and Y406, N2 numbering) which directly interact with the substrate, and 11 framework residues (E119, R156, W178, S179, D198, I222, E227, H274, E277, N294, and E425) which support the catalytic residues, are all highly conserved [14].

With the exception of the 2009 pandemic H1N1 virus N1, all members of group 1 NA have an additional cavity (150-cavity) adjacent to the active site [54]. Crystal structures reveal that the 150-cavity results from the open form of the 150-loop (formed by amino acids 147–152, N2 numbering). By contrast, the 150-loop of group 2 NAs display closed conformations which result in 150-cavity deficiency. The open form of the 150-loop has been considered as a new opportunity to develop antiviral compounds [49]. Moreover, D151 on the 150-loop was confirmed to be related to enzymatic activity. The D151G mutant appears to mediate the weak hydrolytic and enhanced SA binding activities [55].

Thanks to fast technology development, the crystal structures of NA-like molecules N10 and N11 have been immediately solved after the genome sequencing, and they maintain an overall tetrameric NA structure but with distinct features, especially in the deficiency of the enzymatic active site or 150-cavity [14,15,19] (Figure 4). In both structures the active sites are a more opened form with no clue as to how the SA is accommodated. In fact, when scrutinized the primary sequences of both N10 and N11

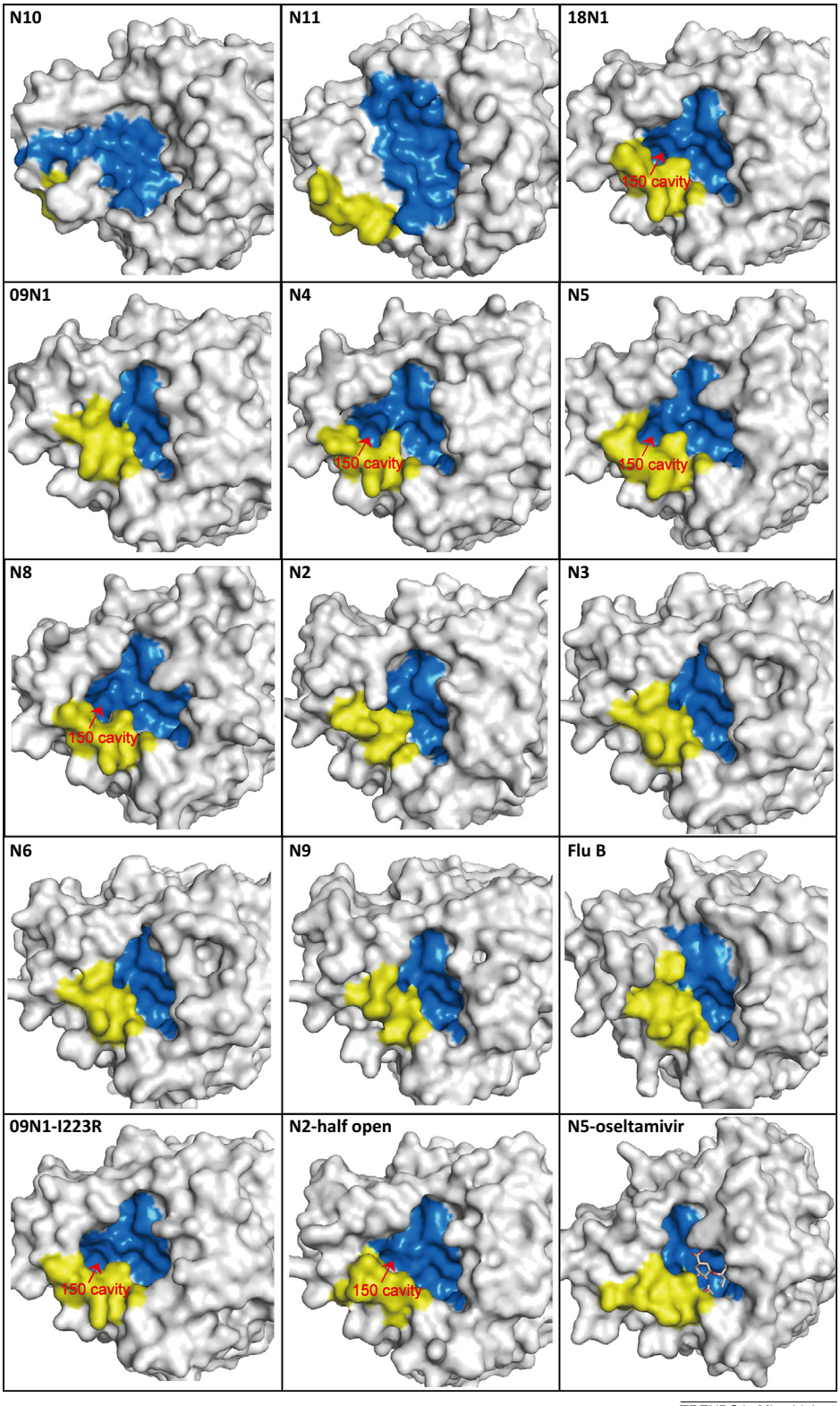

Figure 4. Enzymatic active sites and the adjacent 150-loop structures of all the neuraminidases (NAs) with known structures (including influenza B virus) and NA-like N10 and N11. All the NA structures have been crystallographically resolved except for N7. Amino acids of the 150-loop are labeled in yellow and the substrate-binding groove is labeled in blue. Members from group 1 have a 150-cavity (except for N1 of 2009 pH1N1), which group 2 members otherwise lack. N10 and N11 also lack a 150-cavity, and have a more open structure which does not support SA binding. Under some circumstances, 09N1 can open (I223R mutation) and N2 can half-open (oseltamivir carboxylate-induced). After binding to the inhibitors, group 1 members usually close their 150-cavity (N5 is shown with oseltamivir bound).

show diversified amino acid usages in both the active and the framework sites. Only three of eight active site (R118, R224, and E276) and three (four for N11) of 11 framework (S179, D198, E425, and W178-N11) amino acids are conserved. Therefore the structures do not support a neuraminidase function. Indeed, no sialidase enzymatic activities were detected in either N10 or N11 proteins when tested *in vitro* [14,15,19]. Therefore they are not canonical NA proteins, and instead are NA-like molecules. N10 and N11 might function as a receptor binding protein via their Ig-like folds, as discussed in the previous section. Thus, alterations of the conserved NA active site architecture observed in both bat-derived N10 and N11, combined with evidence for a lack of sialidase activity, indicates that they may have an unknown function distinct from canonical influenza NAs.

Conserved functional modules in other putative encoded proteins

Until now there has been no evidence that these virus genomes can assemble into a real virus, or at least it might not be a viable virus in non-bat cells because attempts to propagate the virus *in vitro* have so far failed [12]. Owing to the distinct functional and structural properties of both H17/18 and N10/11, a comprehensive analysis of the functional modules was performed in the other encoded proteins in addition to the HA and NA. Figure 5 shows that the functional modules in other

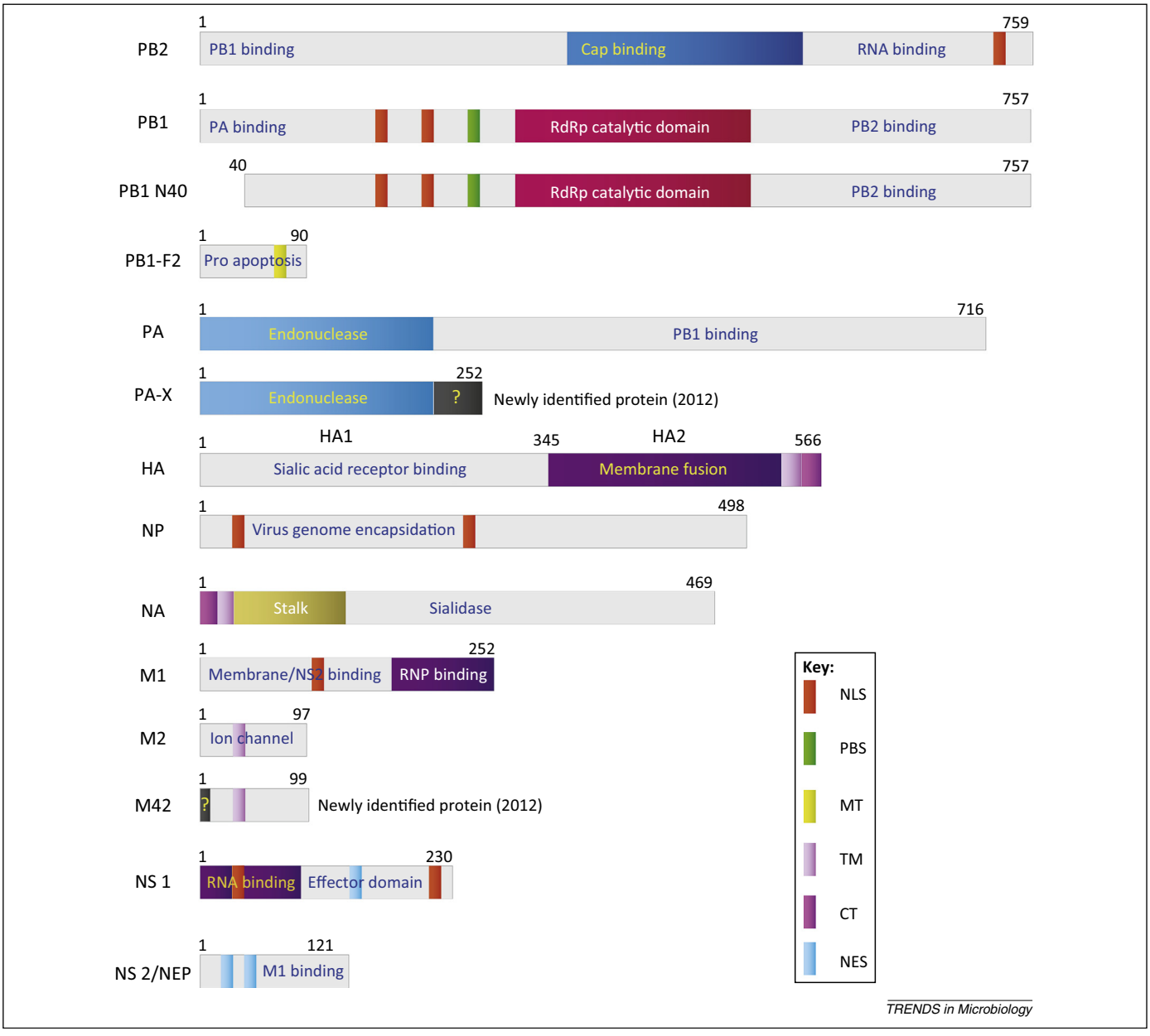

Figure 5. Schematic of the 14 influenza A virus-encoded proteins showing the functional and structural modules or domains. All the known modules are indicated and are highly conserved in the genomes of both H17N10 and H18N11. Abbreviations: '?', unknown functions; CT, cytoplasmic tail; MT, mitochondrial targeting sequence; NES, nuclear export signal; NLS, nuclear localization signal; PBS, promoter binding site; RdRp, RNA-dependent RNA polymerase; RNP, ribonucleoprotein; TM, transmembrane anchor. Amino acid numbering is based on all the protein sequences of the 1918 pandemic H1N1 virus [Genbank: ABA55038.1(PB2), ABA55039.1(PB1/PB1N40), ABW36320.1(PB1-F2), ABA55040.1(PA), AGG82783.1(PA-X), AAD17229.1(HA), AAV48837.1(NP), AAF77036.1(NA), AAN06597.1(M1), AAN06598.1(M2/M42), AAK14368.1(NS1), AAK14369.1(NS2)].

proteins are conserved (explicitly described in detail in the legend of Figure 5) but their real roles need to be experimentally tested in the future. Recent work on the N-terminal domain of the polymerase subunit PA paved the road for such a research [56].

In influenza virus there are six further genes, in addition to HA and NA, which encode 12 proteins, with some newly identified protein products [8,56–59]. These proteins include PA, PB1, PB2, PB1-N40, PB1-F2, PA-X, NP, M1, M2, M42, NS1, and NS2 (Figure 5). Two additional proteins of N-truncated PAs, PA-N155, and PA-N182, have also been reported [60]. M1 is the third virus-surface envelope protein and NP is responsible for viral RNA binding to form ribonucleoprotein (vRNP), which then combines with PB1, PB2, and PA to form the vRNP complex, a core apparatus responsible for proper virus replication [5]. Previous studies on influenza virus have shown that NP binds to viral RNA destined for transcription by PB1, that harbors a canonical RNA-dependent RNA polymerase domain, in conjunction with PB2, with cap-binding capacity, that together form the polymerase complex [61]. PB1 also binds directly to the C-terminal domain of PA [62,63]. The N-terminal domain of PA (PAn) functions as an endonuclease to generate small RNA primers that are essential for initiation of viral gene transcription [64,65]. This cleavage of host-derived RNA occurs in a sequence-specific manner [66,67]. Tong and coworkers demonstrated that the combination of NP, PB1, PB2, and PA derived from bat H17N10 is able to transcribe RNA in an influenza A mini-genome reporter assay, and this activity was abrogated when PB1 was removed from the complex [12].

In studying the N-terminal domain of PA from H17N10 [56], the relative distance was examined between H17N10 PAn and other PAn molecules of which the crystal structures have been solved {A/goose/Guangdong/1/96 (avian H5N1) [65], A/Victoria/3/1975 (human H3N2) [64], A/Vietnam/1203/2004 (avian H5N1) [68], and A/California/04/2009 (human pH1N1) [14]}. The sequence identity between PAn from the four influenza A viruses analyzed is 93–98%, whereas the PAn sequence identity of H17N10 with the four influenza A strains is only 71–72%. The sequence identity between PAn from influenza B and PAn from the four influenza A viruses analyzed is much lower (31–34%), and is in the same range as the identity with H17N10 PAn (35%). PAn from influenza C is at a greater distance, sharing 22–28% identity with the PAn from the four influenza A viruses and 21% with H17N10 PAn. Clearly, based on this comparison, H17N10 PAn is evolutionarily distant from the PAn domains of other viruses used in this comparison.

The crystal structure of H17N10 PAn was solved and was shown to be highly similar to PAn domains of other influenza viruses, folding into five α-helices and seven β-strands [56]. The catalytically conserved residues, including H41, E80, D108, E119, and K134, are ordered in close proximity such that they can accommodate a manganese ion in the center. As expected from the structure, H17N10 PAn possesses manganese-dependent endonuclease activity on both single-stranded DNA and RNA. When the endonuclease activity of H17N10 PAn was compared to that of avian (H5N1) and human (pH1N1) origins, no obvious differences were observed. Similarly to H5N1 PAn [65], the activity of H17N10 PAn is lost by substituting the key active histidine residue at position 41 for an alanine. These observations demonstrate the functional competence of H17N10 PAn as an influenza viral endonuclease, indicating that H17N10 possesses at least one protein domain that suggests an influenza-like life cycle. More experimental work on the other protein molecules or their functional modules should be done to reveal the complete life cycle.

Box 1. Outstanding questions

- Are H17N10 and H18N11 real viruses or merely DNA relics harbored in bats? If they are real viruses, can we grow the virus and under what conditions? Can we isolate a real virus from bats rather than only a genome?
- Can the genomic segments of the H17N10 and H18N11 reassort with canonical influenza viruses?
- Are the cellular receptors for the H17N10 and H18N11 truly different from that of the canonical influenza A viruses? If so, can we identify their receptors?
- Is bat-derived influenza A-like virus widespread in the world, in addition to South America, the initial place of isolation?
- Can we turn the H17 or H18 molecules into SA or SA-linked glycopeptides that bind in the same way as the canonical influenza A virus? The same question needs to be answered for N10 and N11 – as well as to how we can turn them into a real neuraminidase (sialidase); this will require mutagenesis work.

Concluding remarks

Bat-derived influenza A viruses, both H17N10 and H18N11, may not follow the canonical rule for their life cycle as other influenza A viruses, at least based on the recent studies of HA and NA proteins. So far we have not isolated the virus from bats, and we suspect that this genome or virus should be, at most, called influenza-like virus or even a completely new name for the sake of public health concerns in view of the wide public knowledge about potential influenza pandemics. However, it is possible that the determined genomic sequences do not correspond to a functional virus because they may contain mutations originating from the non-functional viruses present in the bats. More work therefore needs to be done, such as the isolation of the virus or more sequences to be determined from the bats. Otherwise, the reassortment ability of the six internal gene segments of either the H17N10 or the H18N11 genome should be examined soon. If reassortment ever occurred, though doubtful, our understanding of influenza virus ecology would be changed and the range of reservoir and 'mixing vessel' hosts for influenza virus would need to be expanded. Furthermore, we are also expecting more HxNy viruses to be discovered from bats in the future (Box 1).

Acknowledgments

Work in the laboratory of G.F.G. is supported by the 973 Project of Ministry of Science and Technology of China (MOST; grant 2011CB504703) and the Key Project of the National Natural Science Foundation of China (NSFC; grant 81330082). G.F.G. is a leading principal investigator of the NSFC Innovative Research Group (Grant No. 81321063).

References

1 Palese, P. and Shaw, M.L. (2007) Orthomyxoviridae: the viruses and their replication. In *Fields Virology* (4th edn) (Knipe, D.M. and Howley, P.M., eds), pp. 1647–1690, Lippincott Williams & Wilkins

2 Cox, N.J. and Subbarao, K. (2000) Global epidemiology of influenza: past and present. *Annu. Rev. Med.* 51, 407–421

3 Kobasa, D. *et al.* (2004) Enhanced virulence of influenza A viruses with the haemagglutinin of the 1918 pandemic virus. *Nature* 431, 703–707

4 Novel Swine-Origin Influenza A (H1N1) Virus Investigation Team (2009) Emergence of a novel swine-origin influenza A (H1N1) virus in humans. *N. Engl. J. Med.* 360, 2605–2615

5 Webster, R.G. *et al.* (1992) Evolution and ecology of influenza A viruses. *Microbiol. Rev.* 56, 152–179

6 Hale, B.G. *et al.* (2010) Structural insights into phosphoinositide 3-kinase activation by the influenza A virus NS1 protein. *Proc. Natl. Acad. Sci. U.S.A.* 107, 1954–1959

7 Liu, D. *et al.* (2013) Origin and diversity of novel avian influenza A H7N9 viruses causing human infection: phylogenetic, structural, and coalescent analyses. *Lancet* 381, 1926–1932

8 Liu, X. *et al.* (2013) Insights into the roles of cyclophilin A during influenza virus infection. *Viruses* 5, 182–191

9 Palese, P. (2004) Influenza: old and new threats. *Nat. Med.* 10, S82–S87

10 Beigel, J.H. *et al.* (2005) Avian influenza A (H5N1) infection in humans. *N. Engl. J. Med.* 353, 1374–1385

11 Gao, R. *et al.* (2013) Human infection with a novel avian-origin influenza A (H7N9) virus. *N. Engl. J. Med.* 368, 1888–1897

12 Tong, S. *et al.* (2012) A distinct lineage of influenza A virus from bats. *Proc. Natl. Acad. Sci. U.S.A.* 109, 4269–4274

13 Shi, Z. (2013) Emerging infectious diseases associated with bat viruses. *Sci. China Life Sci.* 56, 678–682

14 Kowalinski, E. *et al.* (2012) Structural analysis of specific metal chelating inhibitor binding to the endonuclease domain of influenza pH1N1 (2009) polymerase. *PLoS Pathog.* 8, e1002831

15 Zhu, X. *et al.* (2012) Crystal structures of two subtype N10 neuraminidase-like proteins from bat influenza A viruses reveal a diverged putative active site. *Proc. Natl. Acad. Sci. U.S.A.* 109, 18903–18908

16 Sun, X. *et al.* (2013) Bat-derived influenza hemagglutinin H17 does not bind canonical avian or human receptors and most likely uses a unique entry mechanism. *Cell Rep.* 3, 769–778

17 Zhu, X. *et al.* (2013) Hemagglutinin homologue from H17N10 bat influenza virus exhibits divergent receptor-binding and pH-dependent fusion activities. *Proc. Natl. Acad. Sci. U.S.A.* 110, 1458–1463

18 Garcia-Sastre, A. (2012) The neuraminidase of bat influenza viruses is not a neuraminidase. *Proc. Natl. Acad. Sci. U.S.A.* 109, 18635–18636

19 Tong, S. *et al.* (2013) New world bats harbor diverse influenza A viruses. *PLoS Pathog.* 9, e1003657

20 Air, G.M. (1981) Sequence relationships among the hemagglutinin genes of 12 subtypes of influenza A virus. *Proc. Natl. Acad. Sci. U.S.A.* 78, 7639–7643

21 Nobusawa, E. *et al.* (1991) Comparison of complete amino acid sequences and receptor-binding properties among 13 serotypes of hemagglutinins of influenza A viruses. *Virology* 182, 475–485

22 Gamblin, S.J. and Skehel, J.J. (2010) Influenza hemagglutinin and neuraminidase membrane glycoproteins. *J. Biol. Chem.* 285, 28403–28409

23 Lu, X. *et al.* (2013) Structure and receptor binding specificity of hemagglutinin H13 from avian influenza A virus H13N6. *J. Virol.* 87, 9077–9085

24 Sauter, N.K. *et al.* (1989) Hemagglutinins from two influenza virus variants bind to sialic acid derivatives with millimolar dissociation constants: a 500 MHz proton nuclear magnetic resonance study. *Biochemistry* 28, 8388–8396

25 Takemoto, D.K. *et al.* (1996) A surface plasmon resonance assay for the binding of influenza virus hemagglutinin to its sialic acid receptor. *Virology* 217, 452–458

26 Gambaryan, A.S. *et al.* (1997) Specification of receptor-binding phenotypes of influenza virus isolates from different hosts using synthetic sialylglycopolymers: non-egg-adapted human H1 and H3 influenza A and influenza B viruses share a common high binding affinity for 6′-sialyl(N-acetyllactosamine). *Virology* 232, 345–350

27 Bullough, P.A. *et al.* (1994) Structure of influenza haemagglutinin at the pH of membrane fusion. *Nature* 371, 37–43

28 Harrison, S.C. (2008) Viral membrane fusion. *Nat. Struct. Mol. Biol.* 15, 690–698

29 Skehel, J.J. and Wiley, D.C. (2000) Receptor binding and membrane fusion in virus entry: the influenza hemagglutinin. *Annu. Rev. Biochem.* 69, 531–569

30 Wiley, D.C. and Skehel, J.J. (1987) The structure and function of the hemagglutinin membrane glycoprotein of influenza virus. *Annu. Rev. Biochem.* 56, 365–394

31 Wilson, I.A. *et al.* (1981) Structure of the haemagglutinin membrane glycoprotein of influenza virus at 3 Å resolution. *Nature* 289, 366–373

32 Liu, J. *et al.* (2009) Structures of receptor complexes formed by hemagglutinins from the Asian Influenza pandemic of 1957. *Proc. Natl. Acad. Sci. U.S.A.* 106, 17175–17180

33 Zhang, W. *et al.* (2013) Molecular basis of the receptor binding specificity switch of the hemagglutinins from both the 1918 and 2009 pandemic influenza A viruses by a D225G substitution. *J. Virol.* 87, 5949–5958

34 Gamblin, S.J. *et al.* (2004) The structure and receptor binding properties of the 1918 influenza hemagglutinin. *Science* 303, 1838–1842

35 Zhang, W. *et al.* (2013) An airborne transmissible avian influenza H5 hemagglutinin seen at the atomic level. *Science* 340, 1463–1467

36 Xiong, X. *et al.* (2013) Receptor binding by a ferret-transmissible H5 avian influenza virus. *Nature* 497, 392–396

37 Shi, Y. *et al.* (2013) Structures and receptor binding of hemagglutinins from human-infecting H7N9 influenza viruses. *Science* 342, 243–247

38 Xiong, X. *et al.* (2013) Receptor binding by an H7N9 influenza virus from humans. *Nature* 499, 496–499

39 Lawrence, M.C. *et al.* (2004) Structure of the haemagglutinin–neuraminidase from human parainfluenza virus type III. *J. Mol. Biol.* 335, 1343–1357

40 Yuan, P. *et al.* (2011) Structure of the Newcastle disease virus hemagglutinin–neuraminidase (HN) ectodomain reveals a four-helix bundle stalk. *Proc. Natl. Acad. Sci. U.S.A.* 108, 14920–14925

41 Naniche, D. *et al.* (1993) Human membrane cofactor protein (CD46) acts as a cellular receptor for measles virus. *J. Virol.* 67, 6025–6032

42 Tatsuo, H. *et al.* (2000) SLAM (CDw150) is a cellular receptor for measles virus. *Nature* 406, 893–897

43 Muhlebach, M.D. *et al.* (2011) Adherens junction protein nectin-4 is the epithelial receptor for measles virus. *Nature* 480, 530–533

44 Zhang, X. *et al.* (2013) Structure of measles virus hemagglutinin bound to its epithelial receptor nectin-4. *Nat. Struct. Mol. Biol.* 20, 67–72

45 Palese, P. *et al.* (1974) Characterization of temperature sensitive influenza virus mutants defective in neuraminidase. *Virology* 61, 397–410

46 Liu, C. *et al.* (1995) Influenza type A virus neuraminidase does not play a role in viral entry, replication, assembly, or budding. *J. Virol.* 69, 1099–1106

47 Assad, F.A. *et al.* (1980) A revision of the system of nomenclature for influenza viruses: a WHO memorandum. *Bull. World Health Organ.* 58, 585–591

48 Varghese, J.N. *et al.* (1983) Structure of the influenza virus glycoprotein antigen neuraminidase at 2.9 Å resolution. *Nature* 303, 35–40

49 Russell, R.J. *et al.* (2006) The structure of H5N1 avian influenza neuraminidase suggests new opportunities for drug design. *Nature* 443, 45–49

50 Singh, S. *et al.* (1995) Structure-based inhibitors of influenza virus sialidase. A benzoic acid lead with novel interaction. *J. Med. Chem.* 38, 3217–3225

51 Li, Q. *et al.* (2013) Functional and structural analysis of influenza virus neuraminidase N3 offers further insight into the mechanisms of oseltamivir resistance. *J. Virol.* 87, 10016–10024

52 Wang, M. *et al.* (2011) Influenza A virus N5 neuraminidase has an extended 150-cavity. *J. Virol.* 85, 8431–8435

53 Varghese, J.N. *et al.* (1995) Three-dimensional structure of the complex of 4-guanidino-Neu5Ac2en and influenza virus neuraminidase. *Protein Sci.* 4, 1081–1087

54 Li, Q. *et al.* (2010) The 2009 pandemic H1N1 neuraminidase N1 lacks the 150-cavity in its active site. *Nat. Struct. Mol. Biol.* 17, 1266–1268

55 Zhu, X. *et al.* (2012) Influenza virus neuraminidases with reduced enzymatic activity that avidly bind sialic acid receptors. *J. Virol.* 86, 13371–13383
56 Tefsen, B. *et al.* (2013) The N-terminal domain of PA from bat-derived influenza-like virus H17N10 has endonuclease activity. *J. Virol.* 88, 1935–1941
57 Chen, W. *et al.* (2001) A novel influenza A virus mitochondrial protein that induces cell death. *Nat. Med.* 7, 1306–1312
58 Wise, H.M. *et al.* (2009) A complicated message: identification of a novel PB1-related protein translated from influenza A virus segment 2 mRNA. *J. Virol.* 83, 8021–8031
59 Jagger, B.W. *et al.* (2012) An overlapping protein-coding region in influenza A virus segment 3 modulates the host response. *Science* 337, 199–204
60 Muramoto, Y. *et al.* (2013) Identification of novel influenza A virus proteins translated from PA mRNA. *J. Virol.* 87, 2455–2462
61 Sugiyama, K. *et al.* (2009) Structural insight into the essential PB1-PB2 subunit contact of the influenza virus RNA polymerase. *EMBO J.* 28, 1803–1811
62 He, X. *et al.* (2008) Crystal structure of the polymerase PA(C)–PB1(N) complex from an avian influenza H5N1 virus. *Nature* 454, 1123–1126
63 Obayashi, E. *et al.* (2008) The structural basis for an essential subunit interaction in influenza virus RNA polymerase. *Nature* 454, 1127–1131
64 Dias, A. *et al.* (2009) The cap-snatching endonuclease of influenza virus polymerase resides in the PA subunit. *Nature* 458, 914–918
65 Yuan, P. *et al.* (2009) Crystal structure of an avian influenza polymerase PA(N) reveals an endonuclease active site. *Nature* 458, 909–913
66 Noble, E. *et al.* (2012) Endonuclease substrate selectivity characterized with full-length PA of influenza A virus polymerase. *Virology* 433, 27–34
67 Datta, K. *et al.* (2013) Characterization of PA-N terminal domain of influenza A polymerase reveals sequence specific RNA cleavage. *Nucleic Acids Res.* 41, 8289–8299
68 Dubois, R.M. *et al.* (2012) Structural and biochemical basis for development of influenza virus inhibitors targeting the PA endonuclease. *PLoS Pathog.* 8, e1002830