15

Viral surface glycoproteins in carbohydrate recognition: structure and modelling

Jeffrey C. Dyason and Mark von Itzstein

SUMMARY

The importance of carbohydrate-recognizing viral surface glycoproteins in a range of clinically important viral infections has been identified and a number of these proteins have been investigated as possible drug discovery targets. As a part of these investigations several of these proteins have had their three-dimensional structures determined by either nuclear magnetic resonance spectroscopic or X-ray crystallographic methods. This structural information has provided an excellent basis for structure-assisted inhibitor design using computational chemistry methods. This chapter describes some of the most significant developments in the field of structure-based investigations of viral surface-resident carbohydrate-recognizing proteins. Specifically, an overview of these carbohydrate-recognizing proteins from four important human viruses, including influenza, dengue, rotavirus and parainfluenza, and associated structural investigations will be presented.

Keywords: Virus; Carbohydrate; Molecular modelling; Influenza; Dengue virus; Rotavirus; Parainfluenza

1. INTRODUCTION

For viruses to be able to replicate, they need to be able to adhere, then bind and invade and, finally, to escape from a susceptible host cell. Animal host cell surfaces are highly decorated by a range of glycan structures that are made up of both charged and neutral carbohydrates (Gagneux and Varki, 1999). These carbohydrates are involved in a range of different functions, including cell protection, a host of recognition events and cell–cell communication (Gagneux and Varki, 1999). As a consequence of this elaborate host cell surface carbohydrate display, viruses have had to develop a variety of strategies that can deal with this protective barrier to enable successful viral propagation. In fact, many viruses actively use these carbohydrates in the recognition and invasion process of their life cycle. Understanding the structure of these viral surface proteins is vital not only to help understand their function but also to enable the design of new drugs capable of mitigating the effects caused during infection.

2. INFLUENZA VIRUS

Influenza virus has been an affliction of mankind for centuries, with several major pandemics being reported, including the Spanish Flu which, at the end of World War I (1918–1919), killed an estimated 40 million people (Taubenberger *et al.*, 2001). More recently, the emergence of the highly pathogenic avian influenza virus strain, H5N1, although not easily transmissible between humans, has a reported mortality rate in man of over 50% (WHO, 2008).

Influenza virus contains a single-stranded segmented RNA genome and is a member of the orthomyxoviridae family that is further subdivided into three distinct types, influenza virus A, B and C (van Regenmortel *et al.*, 2000). Of most concern to the human population are types A and B, with type A the most likely to cause pandemics (Luscher-Mattli, 2000). Both types A and B have two surface-based carbohydrate recognizing proteins involved in the infectious cycle of the virus, namely haemagglutinin (HA), which has a lectin and fusion function and sialidase (also known as neuraminidase, NA) which has a glycohydrolase function. Type C has only one major carbohydrate recognizing protein and that is haemagglutinin-esterase-fusion protein (HEF), which combines the roles of the HA and NA of influenza A and B (Herrler *et al.*, 1988).

Currently, the major way of combating influenza virus A and B infection is via an annual vaccine injection, however, this is not foolproof and generally provides no protection against new strains of influenza virus, like that which may arise from the avian H5N1 strains. Also available are several drugs, the most successful of which are the NA inhibitors (von Itzstein, 2007), Relenza® and Tamiflu® (see below). However, mutant influenza strains resistant to Tamiflu® have been recently detected that potentially limit the available armament and clearly necessitates the ongoing search for new treatments

(von Itzstein, 2007). While influenza virus C is not common, it can cause respiratory tract infections generally in young children, however, in some cases, this may lead to more severe infections such as bronchiolitis and pneumonia (Wagaman *et al.*, 1989; Crescenzo-Chaigne and van der Werf, 2007). There is currently no vaccine or specific drug treatment for influenza C.

2.1. Haemagglutinin as a drug discovery target

Influenza virus HA is a homo-trimeric glycosylated protein located on the surface of the virion (Figure 15.1) and is responsible for the initial
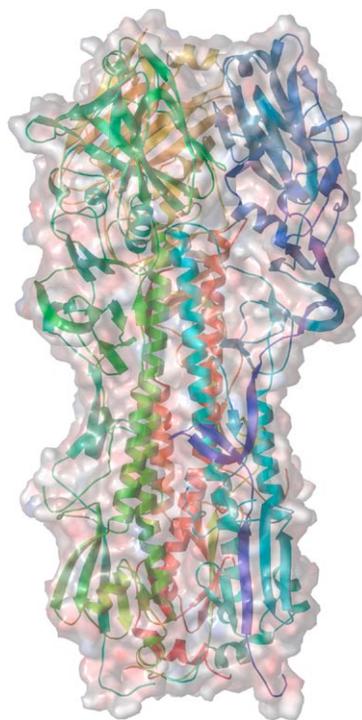

FIGURE 15.1   The secondary structure of an influenza virus haemagglutinin trimer is shown with a translucent surface. The surface is coloured by electrostatic potential with red denoting areas of negative charge while blue denotes areas of positive charge. (See colour plates section.)

attachment of the virion to the host cell surface, by recognition of terminal sialic acid moieties. A change in conformation of the protein is then responsible for the initiation of fusion with the cell membrane. The first structure of influenza virus A HA was published in 1981 (Wilson et al., 1981). Since that time, over 70 structures of various influenza virus HAs have been deposited in the Research Collabatory for Structural Bioinformatics (RCSB) Protein Data Bank (PDB) (Berman et al., 2000). During the 1980s and early 1990s, there were a number of attempts at using the available crystal structures of influenza virus HA to design small molecule HA binders that would hopefully inhibit the replication of influenza virus. A number of excellent reviews have appeared that detail these attempts (Luscher-Mattli, 2000; Matrosovich and Klenk, 2003).

However, it was realized that the low affinity of the sialic acid binding site would not produce the tight binding inhibitor required for a drug, leading to the design and synthesis of multivalent and polymeric inhibitors. These inhibitors have, however, not produced a suitable drug candidate, although work is still being undertaken (Ogata et al., 2007; Marra et al., 2008). A recent study by Nandi (2008) described an in silico search of a small subset in the ZINC database (Irwin and Shoichet, 2005) and led to the suggestion of some possible small molecule inhibitors of HA binding, although biological evaluation of these compounds is not yet reported.

Avian influenza virus HA preferentially recognizes terminal α-(2→3)-linked sialic acid containing glycoconjugates, whereas human-adapted influenza virus HA preferentially recognizes terminal α-(2→6)-linked sialic acid glycoconjugates. More recently, the crystal structures of HAs from the virus responsible for the 1918–1919 pandemic (Gamblin et al., 2004) and from the highly pathogenic avian influenza H5N1 strain (Stevens et al., 2006) have been elucidated. Knowledge of these structures has led to a greater understanding of the sialic acid

specificity of these strains and it appears as if only a few amino acid mutations are needed to enable the avian influenza virus HA to switch its specificity and recognize human receptors (Stevens et al., 2006; Yamada et al., 2006). Recent reports that influenza virus may recognize receptors other than the accepted α-(2→6)-linked sialic acids observed in the human lower respiratory tract (Rapoport et al., 2006; Nicholls et al., 2007), has inspired studies that may more completely characterize novel asialoglycan recognition by HA. For example, the von Itzstein group described a nuclear magnetic resonance (NMR) spectroscopy-based study that enables the interrogation of H5 (from H5N1), as a surface protein of virus-like particles, with novel inhibitors and potential novel glycan structures (Haselhorst et al., 2008).

2.2. Sialidase as a drug discovery target

The NA of influenza virus A and B has been the major focus in the development of drugs to combat influenza infection with two drugs already available on the market, namely the structure-based designed zanamivir (Relenza®, GlaxoSmithKline) and oseltamivir (Tamiflu®, Roche). The development of these drugs has been adequately reviewed elsewhere (Laver, 2006; von Itzstein, 2007, 2008). Influenza virus NA is a receptor-destroying enzyme which cleaves the HA receptor, sialic acid, from a range of sialyloligosaccharides allowing newly formed viral progeny to escape the surface of infected cells and go on to infect more cells.

Influenza virus NA is a glycosylated homo-tetramer which is tethered to the virus surface by a long protein stalk at its C terminal (Figure 15.2). The first structure of influenza virus NA was published in 1983 (Colman et al., 1983; Varghese et al., 1983) and there exists nearly 100 structures in the current PDB, including a number of mutant proteins as well as a number of inhibitor-enzyme complexes.
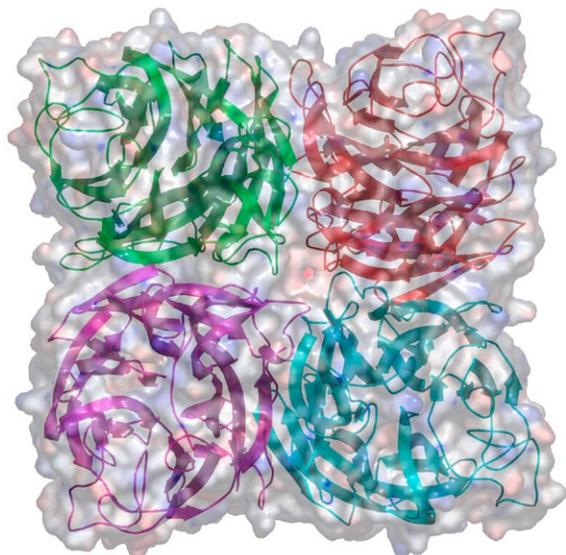
FIGURE 15.2 The influenza virus sialidase tetramer is shown with a translucent surface, coloured by electrostatic potential overlaying the monomer secondary structure. Each monomer has six four-stranded anti-parallel β-sheets arranged as if on the blades of a propeller. (See colour plates section.)

FIGURE 15.3 Superimposition of two complexed N9 sialidase structures of influenza virus showing the induced fit of Glu276 when oseltamivir carboxylate (orange carbon atoms) is bound, compared with Neu5Ac2en (green carbon atoms). (See colour plates section.)

The inhibitor-enzyme complexes in particular allow further study to understand the interactions involved and the design of new inhibitors to combat the existence of mutated influenza viruses that are less susceptible to the existing drugs, particularly Tamiflu®. This wealth of structural information has facilitated the use of structure-based drug design techniques, such as docking, molecular mechanics/Poisson-Boltzmann surface area (MM/PBSA) simulations, comparative molecular field analysis (CoMFA), etc., in providing not only an explanation of how inhibitors bind to the influenza virus NAs but also in the design of new inhibitors. A selection of publications gives an overview of the use of these techniques (Stoll *et al.*, 2003; Masukawa *et al.*, 2003; Armstrong *et al.*, 2006; Platis *et al.*, 2006; Zhang *et al.*, 2006; Abu Hammad *et al.*, 2007).
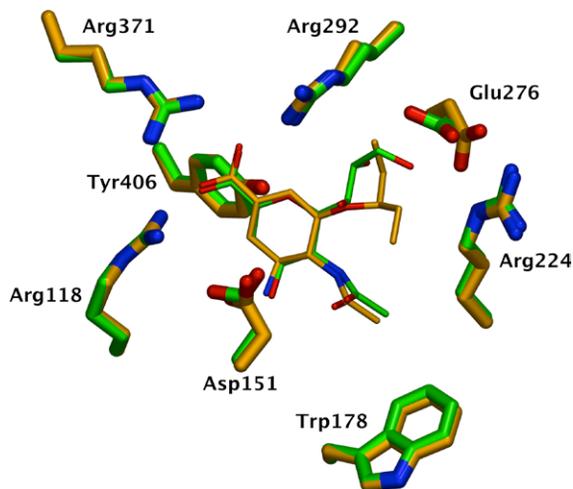
One aspect that is not handled consistently in many of these studies is the ability of the protein to undergo induced-fit when the glycerol side chain of the sialic acid-based inhibitor zanamivir is replaced with a hydrophobic group, e.g. the active form of oseltamivir, oseltamivir carboxylate (Figure 15.3). In this example, to accommodate the hydrophobic group of oseltamivir carboxylate, glutamic acid residue 276 (Glu276) flips to form an internal interaction with arginine 224 (Arg224) that results in a hydrophobic pocket where, previously, there was a negatively-charged hydrogen bonding group. Such induced-fit phenomena are not easily predicted or indeed adequately handled by theoretical models. Relying on this induced fit, as oseltamivir carboxylate does, also makes it a "hot-spot" for the virus to generate mutants that render the drug less able to inhibit the receptor-destroying enzyme function of the NA. This outcome has already been observed in a histidine-tyrosine (His274Tyr) N1 mutant where oseltamivir carboxylate is 265-fold less potent
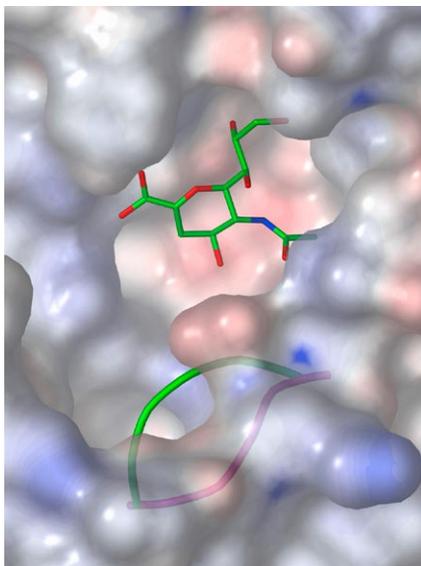
FIGURE 15.4 Crystal structures of the influenza virus N8 sialidase. The surface and the magenta coloured 150 loop are from a structure of the uncomplexed protein while the Neu5Ac2en and green-coloured 150 loop are from a structure where Neu5Ac2en was soaked into the protein crystals. The surface is coloured by electrostatic potential. (See colour plates section.)

compared to zanamivir which is only a factor of 2 less potent (Collins *et al.*, 2008).

One of the most interesting advances to come from recent NA structure determination and analysis is the existence of two distinct groups of influenza virus NAs, group 1 and 2 (Russell *et al.*, 2006). The main difference between the groups appears to be a very flexible active site loop that includes one of the main catalytic residues, aspartic acid 151 (Asp151). The group 1 NAs consisting of N1, N4, N5 and N8 subtypes have the very flexible 150-loop which opens up a large new pocket in the active site in their uncomplexed form, whereas the group 2 NAs consisting of N2, N3, N6, N7 and N9 show no movement of this 150-loop (Figure 15.4). While the current commercially available NA-targeting drugs have all been designed against the group 2 NAs, they generally appear to be equally effective against the group 1 NAs (Govorkova

et al., 2001; Mishin *et al.*, 2005), although some strain variability has been observed. A complication with the current avian H5N1 strain (an example of a group 1 NA) induced disease is that in humans it manifests differently to the normally observed human influenza disease resulting in a higher mortality. In fact, the H5N1 strain appears to have more in common with the Spanish flu strain of 1918–1919 (H1N1) where healthier individuals appear to be more susceptible than the usual influenza target population of the old, young and immunocompromised. It has been proposed that this is a direct consequence of these pandemic or pandemic-like viruses' ability to cause an immune event referred to as a cytokine storm (Peiris *et al.*, 2007).

2.3. Haemagglutinin esterase fusion protein

The HEF of influenza virus C is structurally very similar to HA found on the surface of influenza virus A and B, in that it is a homotrimeric glycosylated surface protein. The major differences are in:

(i) the receptor that is recognized, sialic acid-containing glycoconjugates are recognized by HA while 9-*O*-acetylated sialic acid-glycoconjugates are recognized by HEF; and
(ii) the fact that HEF contains both a lectin function as well as a receptor-destroying enzyme, in contrast to influenza virus A and B that have two distinct proteins for these functions.

The receptor-destroying enzyme of HEF is an esterase and cleaves an *O*-acetate group from the 9'-position of sialic acids. The crystal structure of HEF was published in 1998 (Rosenthal *et al.*, 1998) with the receptor-destroying enzyme active site being very similar in nature to a serine esterase and being completely separate to the receptor-binding site. Modelling experiments using a forced glycosidic angle torsion search have helped explain NMR results obtained

recently within the von Itzstein group (Mayr et al., 2008) for both HEF and bovine coronavirus esterase. This led to the understanding of the essential pharmacophoric groups required for binding to the receptor-destroying enzyme active site and the knowledge that the aglycon group does not appear to be involved in substrate binding to this site. Similar haemagglutinin esterases (≈30% homology) have also been identified in corona and toroviruses (de Groot, 2006) and, although these viruses are not closely related to influenza virus C, an understanding of the structure and function of the haemagglutinin esterase will provide more information on the mechanism of cell recognition and entry.

2.4. Future directions

Influenza virus is very adaptable, with resistant mutants already appearing for at least one of the commercially available NA inhibitors. The ongoing structural and computational studies on the surface-resident carbohydrate recognizing proteins will help not only to identify new and potent drugs but also to understand more fully the mechanism of action of these proteins. This is especially of interest to the group 1 NAs including the current avian influenza (H5N1), where a flexible loop in the active site provides anti-influenza drug designers with another target to improve the potency of existing drugs and the chance to develop new group 1-specific drugs.

3. PARAINFLUENZA

The paramyxoviridae family includes as members the human parainfluenza viruses, Newcastle disease virus and the measles and mumps viruses (Lamb et al., 1996). Of particular interest are human parainfluenza viruses and Newcastle disease virus, both of which have a surface-resident carbohydrate recognizing protein known as haemagglutinin-neuramindase (HN) which is involved in both cellular recognition and as a receptor-destroying enzyme (Lamb et al., 1996). There are four human parainfluenza viruses (hPIV) strains identified, hPIV-1, hPIV-2, hPIV3 and hPIV-4, all of which cause respiratory tract symptoms, e.g. coughing and wheezing, predominantly in infants. The human parainfluenza viruses are considered the second most important cause of lower respiratory tract disease in young children, behind respiratory syncytial virus (Henrickson, 2003). Strains hPIV-1 and hPIV-2 are commonly associated with the respiratory tract disease croup, while hPIV-3 can cause bronchiolitis and pneumonia and hPIV-4 is rarely detected (Henrickson, 2003). Newcastle disease virus is also known as avian paramyxovirus 1 and has a devastating effect on the poultry industry worldwide (Seal et al., 2000).

3.1. Haemagglutinin-neuraminidase as a drug discovery target

Similar to the HEF of influenza virus C, HN includes both a recognition element for terminal sialic acids as well as a receptor-destroying (sialidase/neuraminidase) function to facilitate virion budding. The first crystal structures of a HN from Newcastle disease virus were published in 2000 (Crennell et al., 2000), showing it to be very similar to the influenza virus NA in its overall shape and active site configuration, however, the active site appeared to be more flexible, probably allowing it to function as both a sialic acid-recognizing lectin and a NA. A more recent Newcastle disease virus HN structure (Zaitsev et al., 2004) has revealed the presence of a second sialic acid binding site which is proposed to be involved in the fusion process but not in the initial recognition of sialic acids. At that time, these Newcastle disease virus HN structures were the only available structures for structure-based drug design of inhibitors against human parainfluenza viruses and led to two reports of designed inhibitors for this enzyme system

(Alymova et al., 2004; Tindal et al., 2007). These studies made use of the 2-deoxy-2,3-didehydro-N-acetylneuraminic acid (Neu5Ac2en) template which was also the precursor of the influenza virus NA inhibitor Relenza® and both also targeted a large cavity around the O-4 binding region of the Neu5Ac2en template. The best of these inhibitors led to micromolar inhibition of the NA function of the most serious of the human parainfluenza viruses (i.e. hPIV-3). A crystal structure has been reported for one of these designed inhibitors (Ryan et al., 2006) in complex with Newcastle disease virus HN and demonstrates that the ligand is bound into the active site in the predicted manner (Figure 15.5).

More recently, several crystal structures of the HN from hPIV-3 have been published (Lawrence et al., 2004), although, to the best of our knowledge, no structure-based drug design studies have been reported. Furthermore, crystal structures of the HN from PIV5, formerly known as simian virus 5, have been published (Yuan et al., 2005) and reveal an overall architecture that appears to be similar to those of Newcasle disease virus and hPIV-3.

3.2. Future directions

In addition to providing drug design opportunities, these HN structures, along with structures of the fusion (F) protein, provide more information on the mechanism of cell entry by this family of viruses. This will lead to a greater understanding of how these viruses propagate, making it easier to combat their spread. The original structure-based drug design studies for an anti-human parainfluenza virus drug were based on the available Newcastle disease virus HN structures, more detailed design work can now be undertaken using the hPIV-3 structures. Also, there are a number of non-sialic acid based templates that have been used in the development of influenza virus NA (e.g. cyclohexene, cyclopentane and pyrrolidine), these alternative templates could be used in the structure-based drug design process to extend the range of compounds as potential HN inhibitors.

4. DENGUE VIRUS

The flaviviruses are members of the Flaviviridae family and are transmitted to humans by either mosquito or tick bites (Solomon and Mallewa, 2001). Flavivrus infection can lead to a number of severe and debilitating symptoms, e.g. fever-arthralgia-rash, viral haemorrhagic fever and neurological disease (Solomon and Mallewa, 2001). The mosquito-borne viruses include Dengue virus (DENV), of which there are four strains (DENV1, DENV2, DENV3 and DENV4), yellow fever virus, West Nile virus and Japanese encephalitis virus, while the tick-borne viruses include tick-borne encephalitis virus, Langat virus and Omsk haemorrhagic fever virus.
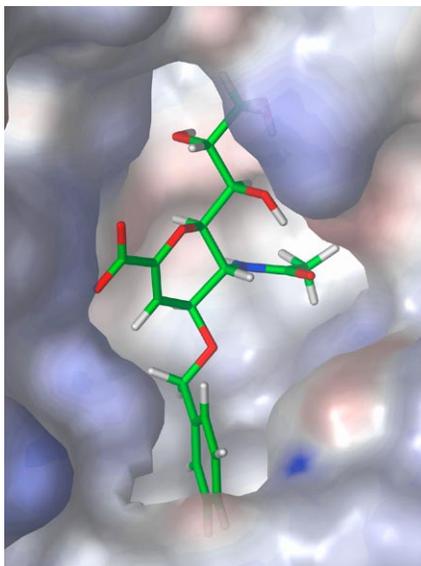

FIGURE 15.5 Crystal structure of 4-O-benzyl-Neu5Ac2en soaked into the haemagglutinin-neuraminidase of Newcastle disease virus, showing the large pocket targeted by the benzyl group of the designed inhibitor. (See colour plates section.)

4.1. E-Glycoprotein and E-glycoprotein-domain 3

The surface of flavivirus is predominantly covered by E-glycoprotein (EGP) which forms dimers that are tethered to the surface by a stalk region. The EGP-associated glycan of the virus has been shown to be important for DENV infection through its interaction with dendritic cell-specific ICAM-3-grabbing non-integrin, DC-SIGN (Navarro-Sanchez et al., 2003; Tassaneetrithep et al., 2003) (see Chapter 34). This interaction and its involvement in the life cycle have been recently reviewed (Perera et al., 2008). Of note, EGP is the major site of host cell receptor binding and of host-mediated antibody neutralization. Both the mosquito and tick-borne virus EGPs share significant sequence and structural homology. The crystal structures of EGP from several sources show that it is comprised of three distinct domains (Figure 15.6); namely DENV2 (Modis et al., 2003; Zhang et al., 2004), DENV3 (Modis et al., 2005) and tick-borne encephalitis virus (Rey et al., 1995). After binding and during the fusion process, the EGP dimers dissociate and reform as trimers; the structures of this trimeric fusion state are also available, i.e. DENV2 (Modis et al., 2004),

tick-borne encephalitis virus (Bressanelli et al., 2004) and West Nile virus (Nybakken et al., 2006).

Of these three domains, it has been found that Domain 3 (D3) is mostly responsible for the virus binding to target host cells (Crill and Roehring, 2001). Also, D3 by itself has been shown to block virus binding. In addition to the crystal structures mentioned above, there have been structures of the D3 sub-unit from DENV2 (Huang et al., 2008), DENV4 (Volk et al., 2007), West Nile virus (Volk et al., 2004), Japanese encephalitis virus (Wu et al., 2003), Langat virus (Wu et al., 2003) and haemorrhagic fever virus (Volk et al., 2006) determined by NMR spectroscopic techniques.

While the receptor for EGP has not been definitively established, previous studies have identified the glycosaminoglycan, heparan sulfate and a range of protein species as proposed mammalian cell surface receptors (Clyde et al., 2006). A recent work (Aoki et al., 2006) has shown that there is an association between DENV and the mammalian cell surface glycolipid, paragloboside. This suggests that EGP and, in particular D3, are carbohydrate-recognizing proteins with small variations in sequence between the members of this family leading to differing cell receptors.
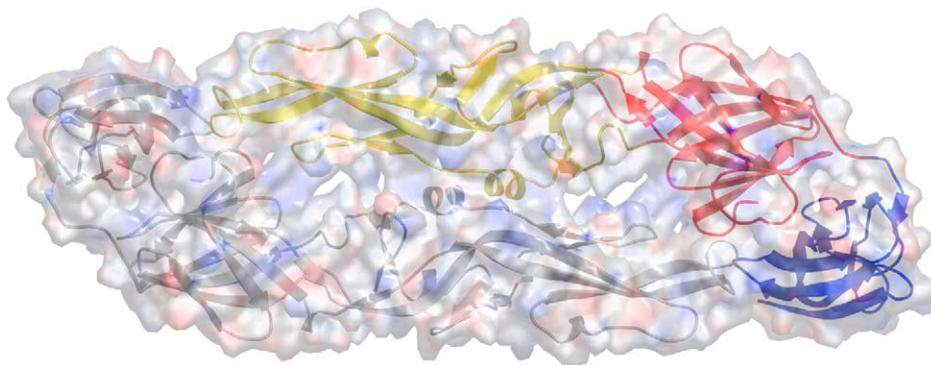

FIGURE 15.6 Crystal structure of the E-glycoprotein dimer of dengue virus. The red-coloured ribbon is Domain 1, the yellow-coloured ribbon is Domain 2, while the blue-coloured ribbon is Domain 3. The surface is coloured by electrostatic potential. (See colour plates section.)

The available D3 structures from different flaviviruses also allows us to come to a greater understanding of how various neutralizing antibodies bind to D3 and EGP (Wu *et al.*, 2003; Nybakken *et al.*, 2005; Kaufmann *et al.*, 2006; Gromowski and Barrett, 2007) and facilitate the design of possible therapeutics capable of halting the spread of these arthropod-borne viruses (Figure 15.7).

4.2. Future directions

Access to the available structures and molecular modelling techniques, for example blind autodocking (Hetenyi and van der Spoel, 2002) and saturation transfer difference (STD-) NMR spectroscopic techniques (Mayer and Meyer,

FIGURE 15.7 Superimposition of the available flavivirus Domain 3 structures. Green, DENV2; red, DENV3; blue, Japenese encephalitis virus; cyan, Langat virus; magenta, Omsk haemorrhagic fever virus; yellow, tickborne encephalitis virus; and brown, West Nile virus. The surface represents the remaining portion of the EGP dimer and is coloured by electrostatic potential. A major difference can be seen in loop 1 for tick-borne encephalitis virus, Omsk haemorrhagic fever virus and Langat virus which bulges out because there is a 4 residue deletion in the loop 2 portions of these proteins. (See colour plates section.)

2001; Haselhorst *et al.*, 2007a,b) in combination with glycan array studies (Day *et al.*, 2007), will provide exciting opportunities for the identification of not only the host cell based glycan receptor but also characterization of the binding site on D3. This information combined with structure-based drug design techniques can then be applied to search for a possible inhibitor of the initial recognition and binding events involved in cellular invasion by these arthropod-borne viruses. As presented in the influenza virus HA example above, small molecule-based inhibitors may not bind with sufficient affinity to produce an entity capable of preventing virus replication, however, multivalent inhibitors may provide valuable alternatives.
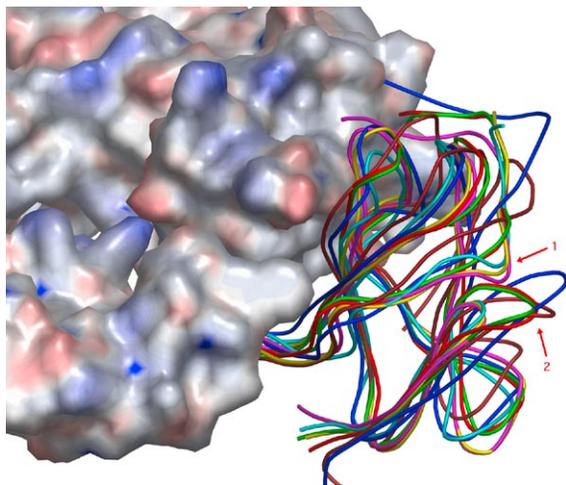
5. ROTAVIRUS

Rotavirus is a member of the *Reoviridae* family and is the most frequent causative agent of severe diarrhoea in young children accounting for 25–55% of all hospital admissions for diarrhoea and 600 000 deaths every year (Glass *et al.*, 2006). While advances in the development of vaccines have been made (Heyse *et al.*, 2008) for the human population, questions about safety and efficacy still remain (Franco and Greenberg, 2001). Furthermore, rotavirus also impacts on productivity in the livestock industry (Yu *et al.*, 2008) and therefore further work is needed to better understand the rotavirus lifecycle and carbohydrate involvement for drug and new vaccine development opportunities.

5.1. VP8*

The rotavirus virion surface spike protein VP4 is responsible for receptor binding and cell penetration (Ludert *et al.*, 1996; Zarate *et al.*, 2000; Arias *et al.*, 2002). The protein VP4 is cleaved by trypsin into the N-terminal VP8*

fragment which has been shown to be responsible for haemagglutination and the VP5* fragment which is believed to be involved in cellular fusion (Fiore *et al.*, 1991; Ciarlet *et al.*, 2002; Graham *et al.*, 2003). There has been some controversy over the absolute requirement of sialic acid in the initial binding of VP8* of the virus to the cell surface (Delorme *et al.*, 2001; Ciarlet *et al.*, 2002; Isa *et al.*, 2006). Currently, rotavirus strains are classified as either sialidase-sensitive, i.e. they no longer infect sialidase-treated cells inferring a sialic acid dependence of infection, or sialidase-insensitive as they infect cells regardless of sialidase treatment, thereby suggesting sialic acid independence of infection.

To examine carbohydrate recognition further, the first structures (NMR and X-ray) of a VP8* from a Rhesus rotavirus strain (sialidase-sensitive) was published in 2002 (Dormitzer *et al.*, 2002). Subsequently, the crystal structures of a number of native and site-directed mutant Rhesus rotavirus VP8* proteins have been reported (Kraschnefski *et al.*, 2009). The only other VP8* structure from a sialidase-sensitive strain that has been published is from the porcine CRW-8 strain (Blanchard *et al.*, 2007). All of the crystal structures of the above VP8* proteins have included the ligand Neu5Acα2Me (Figure 15.8). The binding site shows that the sugar ring is positioned between two tyrosine (Tyr) residues, while the carboxylic acid moiety of the sialic acid forms hydrogen bonds with a serine (Ser) residue and finally an arginine (Arg) residue forms hydrogen bonds to the glycerol side chain of the sialic acid. This sialic acid binding motif is different to that seen in influenza virus haemagglutinin and sialidase, but it has been reported in several sialyl-Lewis X recognizing bacterial toxins (Baker *et al.*, 2007).

Further to these structural investigations, and more relevant to human rotavirus infection, two structures of VP8* proteins from the sialidase-insensitive human strains Wa (Kraschnefski
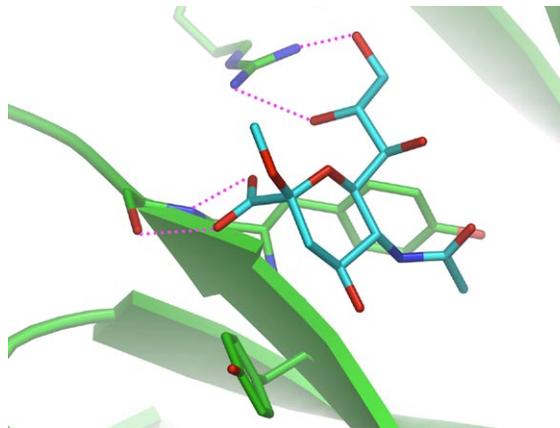

FIGURE 15.8   Crystal structure of Neu5Acα2Me (cyan-coloured carbons) bound to the VP8* protein of rotavirus strain CRW-8. Hydrogen bonds to the serine (Ser) and arginine (Arg) are shown with magenta-coloured dotted lines while the two tyrosine (Tyr) residues which help position the sugar ring are also shown. (See colour plates section.)

et al., 2005; Blanchard *et al.*, 2007) and DS-1 (Monnier *et al.*, 2006) have been determined. Finally, a recent publication (Yu *et al.*, 2008) has described the crystallization and preliminary X-ray diffraction analysis of VP8* protein from the sialidase-sensitive bovine rotavirus strain NCDV.

Molecular modelling and STD NMR experiments within our group (Haselhorst *et al.*, 2007a) have shown a very good correlation between the binding of a number of sialic acid derivatives to Rhesus rotavirus VP8* and the degree to which these compounds can inhibit rotavirus infection of cells. Also, information obtained on the binding epitope of sialic acid disaccharides showed that the penultimate sugar residue does not interact with the protein, but remains predominantly solvated. This is in very good agreement with molecular dynamics calculations that showed that this sugar (galactose) was unable to adopt a conformation where it is in close proximity to the surface (Haselhorst *et al.*, 2007a).

**RESEARCH FOCUS
BOX**

- Identification of next generation influenza virus sialidase inhibitors is required to replace existing drugs as resistance develops.
- Understanding the carbohydrate recognition processes involved in dengue virus and rotavirus infection is necessary to enable the

development of useful chemotherapeutic agents.
- A systematic investigation of host–virus carbohydrate recognition phenomena would greatly enhance the field.

5.2. Future directions

The knowledge of the binding site and ligand for the VP8* protein from human rotavirus strains provides exciting opportunities for the design of inhibitors to prevent virus entry and infection. For example, the use of computational chemistry methods in exploring the available structural data in combination with other techniques, such as NMR spectroscopy, will be invaluable in advancing the role of carbohydrates in the rotavirus lifecycle.

6. CONCLUSIONS

A large number of clinically important viruses utilize carbohydrate viral surface protein interactions to propagate infection and disease. The identification and structural characterization of the proteins and glycans essential in these recognition phenomena opens up new avenues for both drug discovery and vaccine development. While only a limited number of viral surface protein structures have been determined, already structure-based drug design techniques have afforded new classes of drugs. Although significant advances have been made in targeting virus surface glycoproteins for the

discovery of drugs, there remains an urgent need for further investigations (Research Focus Box). The science of glycovirology is still in its infancy but shows great promise for the delivery of new weapons against pathogens that have very few drug treatments available.

References

Abu Hammad, A.M., Afifi, F.U., Taha, M.O., 2007. Combining docking, scoring and molecular field analyses to probe influenza neuraminidase-ligand interactions. J. Mol. Graph. Model. 26, 443–456.

Alymova, I.V., Taylor, G., Takimoto, T., et al., 2004. Efficacy of novel hemagglutinin-neuraminidase inhibitors BCX 2798 and BCX 2855 against human parainfluenza viruses in vitro and in vivo. Antimicrob. Agents Chemother. 48, 1495–1502.

Aoki, C., Hidari, K.I., Itonori, S., et al., 2006. Identification and characterization of carbohydrate molecules in mammalian cells recognized by dengue virus type 2. J. Biochem. 139, 607–614.

Arias, C.F., Isa, P., Guerrero, C.A., et al., 2002. Molecular biology of rotavirus cell entry. Arch. Med. Res. 33, 356–361.

Armstrong, K.A., Tidor, B., Cheng, A.C., 2006. Optimal charges in lead progression: a structure-based neuraminidase case study. J. Med. Chem. 49, 2470–2477.

Baker, H.M., Basu, I., Chung, M.C., Caradoc-Davies, T., Fraser, J.D., Baker, E.N., 2007. Crystal structures of the staphylococcal toxin SSL5 in complex with sialyl Lewis X reveal a conserved binding site that shares common features with viral and bacterial sialic acid binding proteins. J. Mol. Biol. 374, 1298–1308.

Berman, H.M., Westbrook, J., Feng, Z., et al., 2000. The Protein Data Bank. Nucleic Acids Res. 28, 235–242.

Blanchard, H., Yu, X., Coulson, B.S., von Itzstein, M., 2007. Insight into host cell carbohydrate-recognition by human and porcine rotavirus from crystal structures of the virion spike associated carbohydrate-binding domain (VP8*). J. Mol. Biol. 367, 1215–1226.

Bressanelli, S., Stiasny, K., Allison, S.L., et al., 2004. Structure of a flavivirus envelope glycoprotein in its low-pH-induced membrane fusion conformation. EMBO J. 23, 728–738.

Ciarlet, M., Ludert, J.E., Iturriza-Gomara, M., et al., 2002. Initial interaction of rotavirus strains with N-acetylneuraminic (sialic) acid residues on the cell surface correlates with VP4 genotype, not species of origin. J. Virol. 76, 4087–4095.

Clyde, K., Kyle, J.L., Harris, E., 2006. Recent advances in deciphering viral and host determinants of dengue virus replication and pathogenesis. J. Virol. 80, 11418–11431.

Collins, P.J., Haire, L.F., Lin, Y.P., et al., 2008. Crystal structures of oseltamivir-resistant influenza virus neuraminidase mutants. Nature 453, 1258–1261.

Colman, P.M., Varghese, J.N., Laver, W.G., 1983. Structure of the catalytic and antigenic sites in influenza virus neuraminidase. Nature 303, 41–44.

Crennell, S., Takimoto, T., Portner, A., Taylor, G., 2000. Crystal structure of the multifunctional paramyxovirus hemagglutinin-neuraminidase. Nat. Struct. Biol. 7, 1068–1074.

Crescenzo-Chaigne, B., van der Werf, S., 2007. Rescue of influenza C virus from recombinant DNA. J. Virol. 81, 11282–11289.

Crill, W.D., Roehrig, J.T., 2001. Monoclonal antibodies that bind to domain III of dengue virus E glycoprotein are the most efficient blockers of virus adsorption to Vero cells. J. Virol. 75, 7769–7773.

Day, C., Tiralongo, J., Hartnell, R.D., von Itzstein, M., Wilson, J.C., Korolik, V., 2007. Optimisation of glycan and small molecule arrays for analysis of Campylobacter chemotaxis and adherence. Zoonoses Pub. Health 54 (Suppl. 1), 99.

de Groot, R.J., 2006. Structure, function and evolution of the hemagglutinin-esterase proteins of corona- and toroviruses. Glycoconj. J. 23, 59–72.

Delorme, C., Brussow, H., Sidoti, J., et al., 2001. Glycosphingolipid binding specificities of rotavirus: identification of a sialic acid-binding epitope. J. Virol. 75, 2276–2287.

Dormitzer, P.R., Sun, Z.Y., Wagner, G., Harrison, S.C., 2002. The rhesus rotavirus VP4 sialic acid binding domain has a galectin fold with a novel carbohydrate binding site. EMBO J. 21, 885–897.

Fiore, L., Greenberg, H.B., Mackow, E.R., 1991. The VP8 fragment of VP4 is the rhesus rotavirus hemagglutinin. Virology 181, 553–563.

Franco, M.A., Greenberg, H.B., 2001. Challenges for rotavirus vaccines. Virology 281, 153–155.

Gagneux, P., Varki, A., 1999. Evolutionary considerations in relating oligosaccharide diversity to biological function. Glycobiology 9, 747–755.

Gamblin, S.J., Haire, L.F., Russell, R.J., et al., 2004. The structure and receptor binding properties of the 1918 influenza hemagglutinin. Science 303, 1838–1842.

Glass, R.I., Parashar, U.D., Bresee, J.S., et al., 2006. Rotavirus vaccines: current prospects and future challenges. Lancet 368, 323–332.

Govorkova, E.A., Leneva, I.A., Goloubeva, O.G., Bush, K., Webster, R.G., 2001. Comparison of efficacies of RWJ-270201, zanamivir, and oseltamivir against H5N1, H9N2, and other avian influenza viruses. Antimicrob. Agents Chemother. 45, 2723–2732.

Graham, K.L., Halasz, P., Tan, Y., et al., 2003. Integrin-using rotaviruses bind α2β1 integrin α2 I domain via VP4 DGE sequence and recognize αXβ2 and αVβ3 by using VP7 during cell entry. J. Virol. 77, 9969–9978.

Gromowski, G.D., Barrett, A.D., 2007. Characterization of an antigenic site that contains a dominant, type-specific neutralization determinant on the envelope protein domain III (ED3) of dengue 2 virus. Virology 366, 349–360.

Haselhorst, T., Blanchard, H., Frank, M., et al., 2007a. STD NMR spectroscopy and molecular modeling investigation of the binding of N-acetylneuraminic acid derivatives to rhesus rotavirus VP8* core. Glycobiology 17, 68–81.

Haselhorst, T., Munster-Kuhnel, A.K., Oschlies, M., Tiralongo, J., Gerardy-Schahn, R., von Itzstein, M., 2007b. Direct detection of ligand binding to Sepharose-immobilised protein using saturation transfer double difference (STDD) NMR spectroscopy. Biochem. Biophys. Res. Commun. 359, 866–870.

Haselhorst, T., Garcia, J.M., Islam, T., et al., 2008. Avian influenza H5-containing virus-like particles (VLPs): host-cell receptor specificity by STD NMR spectroscopy. Angew. Chem. Int. Ed. Engl. 47, 1910–1912.

Henrickson, K.J., 2003. Parainfluenza viruses. Clin. Microbiol. Rev. 16, 242–264.

Herrler, G., Durkop, I., Becht, H., Klenk, H.D., 1988. The glycoprotein of influenza C virus is the haemagglutinin, esterase and fusion factor. J. Gen. Virol. 69, 839–846.

Hetenyi, C., van der Spoel, D., 2002. Efficient docking of peptides to proteins without prior knowledge of the binding site. Protein Sci. 11, 1729–1737.

Heyse, J.F., Kuter, B.J., Dallas, M.J., Heaton, P., 2008. Evaluating the safety of a rotavirus vaccine: the REST of the story. Clin. Trials 5, 131–139.

Huang, K.C., Lee, M.C., Wu, C.W., Huang, K.J., Lei, H.Y., Cheng, J.W., 2008. Solution structure and neutralizing

antibody binding studies of domain III of the dengue-2 virus envelope protein. Proteins 70, 1116–1119.

Irwin, J.J., Shoichet, B.K., 2005. ZINC – a free database of commercially available compounds for virtual screening. J. Chem. Infect. Model. 45, 177–182.

Isa, P., Arias, C.F., Lopez, S., 2006. Role of sialic acids in rotavirus infection. Glycoconj. J. 23, 27–37.

Kaufmann, B., Nybakken, G.E., Chipman, P.R., et al., 2006. West Nile virus in complex with the Fab fragment of a neutralizing monoclonal antibody. Proc. Natl. Acad. Sci. USA 103, 12400–12404.

Kraschnefski, M.J., Scott, S.A., Holloway, G., Coulson, B.S., von Itzstein, M., Blanchard, H., 2005. Cloning, expression, purification, crystallization and preliminary X-ray diffraction analysis of the VP8* carbohydrate-binding protein of the human rotavirus strain Wa. Acta Crystallogr. Sect. F Struct. Biol. Cryst. Commun. 61, 989–993.

Kraschnefski, M.J., Bugarcic, A., Fleming, F., et al., 2009. Effects on sialic acid recognition of amino acid mutations in the carbohydrate-binding cleft of the rotavirus spike protein. Glycobiology 19, 194–200.

Lamb, R.A., Kolakofsky, D., 1996. Paramyxoviridae: the viruses and their replication. In: Fields, B.N., Knipe, D.M., Howley, P.M. (Eds.), Fundamental Virology. Lippincott-Raven, Philadelphia, pp. 577–604.

Laver, W., 2006. Antiviral drugs for influenza: Tamiflu® past, present and future. Future Virol. 1, 577–586.

Lawrence, M.C., Borg, N.A., Streltsov, V.A., et al., 2004. Structure of the haemagglutinin-neuraminidase from human parainfluenza virus type III. J. Mol. Biol 335, 1343–1357.

Ludert, J.E., Feng, N., Yu, J.H., Broome, R.L., Hoshino, Y., Greenberg, H.B., 1996. Genetic mapping indicates that VP4 is the rotavirus cell attachment protein in vitro and in vivo. J. Virol. 70, 487–493.

Luscher-Mattli, M., 2000. Influenza chemotherapy: a review of the present state of art and of new drugs in development. Arch. Virol. 145, 2233–2248.

Marra, A., Moni, L., Pazzi, D., Corallini, A., Bridi, D., Dondoni, A., 2008. Synthesis of sialoclusters appended to calix[4]arene platforms via multiple azide-alkyne cycloaddition. New inhibitors of hemagglutination and cytopathic effect mediated by BK and influenza A viruses. Org. Biomol. Chem. 6, 1396–1409.

Masukawa, K.M., Kollman, P.A., Kuntz, I.D., 2003. Investigation of neuraminidase-substrate recognition using molecular dynamics and free energy calculations. J. Med. Chem. 46, 5628–5637.

Matrosovich, M., Klenk, H.D., 2003. Natural and synthetic sialic acid-containing inhibitors of influenza virus receptor binding. Rev. Med. Virol. 13, 85–97.

Mayer, M., Meyer, B., 2001. Group epitope mapping by saturation transfer difference NMR to identify segments of a ligand in direct contact with a protein receptor. J. Am. Chem. Soc. 123, 6108–6117.

Mayr, J., Haselhorst, T., Langereis, M.A., et al., 2008. Influenza C virus and bovine coronavirus esterase reveal a similar catalytic mechanism: new insights for drug discovery. Glycoconj. J. 25, 393–399.

Mishin, V.P., Hayden, F.G., Gubareva, L.V., 2005. Susceptibilities of antiviral-resistant influenza viruses to novel neuraminidase inhibitors. Antimicrob. Agents Chemother. 49, 4515–4520.

Modis, Y., Ogata, S., Clements, D., Harrison, S.C., 2003. A ligand-binding pocket in the dengue virus envelope glycoprotein. Proc. Natl. Acad. Sci. USA 100, 6986–6991.

Modis, Y., Ogata, S., Clements, D., Harrison, S.C., 2004. Structure of the dengue virus envelope protein after membrane fusion. Nature 427, 313–319.

Modis, Y., Ogata, S., Clements, D., Harrison, S.C., 2005. Variable surface epitopes in the crystal structure of dengue virus type 3 envelope glycoprotein. J. Virol. 79, 1223–1231.

Monnier, N., Higo-Moriguchi, K., Sun, Z.Y., Prasad, B.V., Taniguchi, K., Dormitzer, P.R., 2006. High-resolution molecular and antigen structure of the VP8* core of a sialic acid-independent human rotavirus strain. J. Virol. 80, 1513–1523.

Nandi, T., 2008. Proposed lead molecules against hemagglutinin of avian influenza virus (H5N1). Bioinformation 2, 240–244.

Navarro-Sanchez, E., Altmeyer, R., Amara, A., et al., 2003. Dendritic-cell-specific ICAM3-grabbing non-integrin is essential for the productive infection of human dendritic cells by mosquito-cell-derived dengue viruses. EMBO Rep. 4, 723–728.

Nicholls, J.M., Chan, M.C., Chan, W.Y., et al., 2007. Tropism of avian influenza A (H5N1) in the upper and lower respiratory tract. Nat. Med. 13, 147–149.

Nybakken, G.E., Oliphant, T., Johnson, S., Burke, S., Diamond, M.S., Fremont, D.H., 2005. Structural basis of West Nile virus neutralization by a therapeutic antibody. Nature 437, 764–769.

Nybakken, G.E., Nelson, C.A., Chen, B.R., Diamond, M.S., Fremont, D.H., 2006. Crystal structure of the West Nile virus envelope glycoprotein. J. Virol. 80, 11467–11474.

Ogata, M., Murata, T., Murakami, K., et al., 2007. Chemoenzymatic synthesis of artificial glycopolypeptides containing multivalent sialyloligosaccharides with a gamma-polyglutamic acid backbone and their effect on inhibition of infection by influenza viruses. Bioorg. Med. Chem. 15, 1383–1393.

Peiris, J.S., de Jong, M.D., Guan, Y., 2007. Avian influenza virus (H5N1): a threat to human health. Clin. Microbiol. Rev. 20, 243–267.

Perera, R., Khaliq, M., Kuhn, R.J., 2008. Closing the door on flaviviruses: entry as a target for antiviral drug design. Antiviral Res. 80, 11–22.

Platis, D., Smith, B.J., Huyton, T., Labrou, N.E., 2006. Structure-guided design of a novel class of benzyl-sulfonate inhibitors for influenza virus neuraminidase. Biochem. J. 399, 215–223.

Rapoport, E.M., Mochalova, L.V., Gabius, H.J., Romanova, J., Bovin, N.V., 2006. Search for additional influenza virus to cell interactions. Glycoconj. J. 23, 115–125.

Rey, F.A., Heinz, F.X., Mandl, C., Kunz, C., Harrison, S.C., 1995. The envelope glycoprotein from tick-borne encephalitis virus at 2 A resolution. Nature 375, 291–298.

Rosenthal, P.B., Zhang, X., Formanowski, F., et al., 1998. Structure of the haemagglutinin-esterase-fusion glyco-protein of influenza C virus. Nature 396, 92–96.

Russell, R.J., Haire, L.F., Stevens, D.J., et al., 2006. The structure of H5N1 avian influenza neuraminidase suggests new opportunities for drug design. Nature 443, 45–49.

Ryan, C., Zaitsev, V., Tindal, D.J., et al., 2006. Structural analysis of a designed inhibitor complexed with the hemagglutinin-neuraminidase of Newcastle disease virus. Glycoconj. J. 23, 135–141.

Seal, B.S., King, D.J., Sellers, H.S., 2000. The avian response to Newcastle disease virus. Dev. Comp. Immunol. 24, 257–268.

Solomon, T., Mallewa, M., 2001. Dengue and other emerging flaviviruses. J. Infect 42, 104–115.

Stevens, J., Blixt, O., Tumpey, T.M., Taubenberger, J.K., Paulson, J.C., Wilson, I.A., 2006. Structure and receptor specificity of the hemagglutinin from an H5N1 influenza virus. Science 312, 404–410.

Stoll, V., Stewart, K.D., Maring, C.J., et al., 2003. Influenza neuraminidase inhibitors: structure-based design of a novel inhibitor series. Biochemistry 42, 718–727.

Tassaneetrithep, B., Burgess, T.H., Granelli-Piperno, A., et al., 2003. DC-SIGN (CD209) mediates dengue virus infection of human dendritic cells. J. Exp. Med. 197, 823–829.

Taubenberger, J.K., Reid, A.H., Janczewski, T.A., Fanning, T.G., 2001. Integrating historical, clinical and molecular genetic data in order to explain the origin and virulence of the 1918 Spanish influenza virus. Philos. Trans. Roy. Soc. Lond., Sect. B, Biol. Sci. 356, 1829–1839.

Tindal, D.J., Dyason, J.C., Thomson, R.J., et al., 2007. Synthesis and evaluation of 4-O-alkylated 2-deoxy-2,3-didehydro-N-acetylneuraminic acid derivatives as inhibitors of human parainfluenza virus type-3 sialidase activity. Bioorg. Med. Chem. Lett. 17, 1655–1658.

van Regenmortel, M.H.V., Fauquet, C.M., Bishop, D.H.L. et al., (Eds.), 2000. Virus Taxonomy: Seventh Report of the International Committee on Taxonomy of Viruses. Academic Press, San Diego.

Varghese, J.N., Laver, W.G., Colman, P.M., 1983. Structure of the influenza virus glycoprotein antigen neuraminidase at 2.9 Å resolution. Nature 303, 35–40.

Volk, D.E., Beasley, D.W., Kallick, D.A., Holbrook, M.R., Barrett, A.D., Gorenstein, D.G., 2004. Solution structure and antibody binding studies of the envelope protein domain III from the New York strain of West Nile virus. J. Biol. Chem. 279, 38755–38761.

Volk, D.E., Chavez, L., Beasley, D.W., Barrett, A.D., Holbrook, M.R., Gorenstein, D.G., 2006. Structure of the envelope protein domain III of Omsk hemorrhagic fever virus. Virology 351, 188–195.

Volk, D.E., Lee, Y.C., Li, X., et al., 2007. Solution structure of the envelope protein domain III of dengue-4 virus. Virology 364, 147–154.

von Itzstein, M., 2007. The war against influenza: discovery and development of sialidase inhibitors. Nat. Rev. Drug Discov. 6, 967–974.

von Itzstein, M., 2008. Avian influenza virus, a very sticky situation. Curr. Opin. Chem. Biol. 12, 102–108.

Wagaman, P.C., Spence, H.A., O'Callaghan, R.J., 1989. Detection of influenza C virus by using an in situ este-rase assay. J. Clin. Microbiol. 27, 832–836.

WHO. (2008). World Health Organization – Avian Influenza. http://www.who.int/csr/disease/avian_influenza/en/index.html.

Wilson, I.A., Skehel, J.J., Wiley, D.C., 1981. Structure of the haemagglutinin membrane glycoprotein of influenza virus at 3Å resolution. Nature 289, 366–373.

Wu, K.P., Wu, C.W., Tsao, Y.P., et al., 2003. Structural basis of a flavivirus recognized by its neutralizing antibody: solution structure of the domain III of the Japanese encephalitis virus envelope protein. J. Biol. Chem. 278, 46007–46013.

Yamada, S., Suzuki, Y., Suzuki, T., et al., 2006. Haemagglutinin mutations responsible for the binding of H5N1 influenza A viruses to human-type receptors. Nature 444, 378–382.

Yu, X., Guillon, A., Szyczew, A.J., et al., 2008. Crystallization and preliminary X-ray diffraction analysis of the carbohydrate-recognizing domain (VP8*) of bovine rotavirus straon NCDV. Acta Crystallogr. Sect. F Struct. Biol. Cryst. Commun. F64, 509–511.

Yuan, P., Thompson, T.B., Wurzburg, B.A., Paterson, R.G., Lamb, R.A., Jardetzky, T.S., 2005. Structural studies of the parainfluenza virus 5 hemagglutinin-neuraminidase tetramer in complex with its receptor, sialyllactose. Structure 13, 803–815.

Zaitsev, V., von Itzstein, M., Groves, D., et al., 2004. Second sialic acid binding site in Newcastle disease virus hemagglutinin-neuraminidase: implications for fusion. J. Virol. 78, 3733–3741.

Zarate, S., Espinosa, R., Romero, P., Mendez, E., Arias, C.F., Lopez, S., 2000. The VP5 domain of VP4 can mediate attachment of rotaviruses to cells. J. Virol. 74, 593–599.

Zhang, J., Yu, K., Zhu, W., Jiang, H., 2006. Neuraminidase pharmacophore model derived from diverse classes of inhibitors. Bioorg. Med. Chem. Lett. 16, 3009–3014.

Zhang, Y., Zhang, W., Ogata, S., et al., 2004. Conformational changes of the flavivirus E glycoprotein. Structure 12, 1607–1618.