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.

3.28 Microbial Lectins

N. Sharon, The Weizmann Institute of Science, Rehovot, Israel
I. Ofek, Sackler School of Medicine, Tel Aviv, Israel

3.28.1 Introduction

Numerous microorganisms produce lectins of different kinds that are specific for animal cell surface carbohydrates.[1,2] In viruses, the lectins occur as membrane constituents, in bacteria they are generally present on the cell surface, often as submicroscopic multisubunit surface appendages known as fimbriae or pili, and occasionally in intracellular soluble form, while in protozoa they are constituents of the cell envelope. Most of these lectins serve to tether the organisms to a substratum (e.g., soil bacteria to clays or marine bacteria to corals) in order to better access nutrients, escape deleterious agents, and withstand cleansing. In animals, they serve for the attachment (or adhesion) of the microorganisms to host tissues or cells, a prerequisite for infection, and are therefore among the virulence factors of the microorganisms. Detailed knowledge of the specificity and combining sites of these lectins is of importance not only for the understanding of a basic mechanism conserved throughout evolution, but also for the development of powerful inhibitors of the lectins for blocking microbial adhesion to tissues. This is the aim of antiadhesion therapy of microbial diseases.[3,4]

Most of the microbial lectins have only been poorly characterized. Their existence and specificity have been inferred primarily from inhibition experiments, in which the effect of different carbohydrates is examined on adhesion of the organisms to animal cells or on the agglutination of erythrocytes or other kinds of cell, such as yeasts.[1,2] Specificity for the carbohydrate of glycolipids or glycoproteins is usually assessed by binding of the bacterial lectins to these glycoconjugates separated on thin layer chromatograms or on blots of gel electrophoresis, respectively. Recently, the application of glycan microarrays for specificity studies of viral and bacterial lectins has been described.[5,6] In a few cases, the specificity and association constants of carbohydrates to the isolated lectins have been measured by physicochemical techniques, for example, by surface plasmon resonance.[64]

Binding of carbohydrates to lectins, including those of microbial sources, is generally characterized by a low affinity for monovalent ligands, a drawback balanced by multivalency, which provides high avidity for complex glycans or cell surfaces. In general, millimolar affinity is observed for lectin binding to monosaccharides, although cases of lectins

with micromolar affinities are known.[7] When oligosaccharides act as ligands, corresponding to an extended binding site on the lectins, increased affinity up to the micromolar range can be observed. These interactions are typified by a favorable enthalpy term, due to the high number of hydrogen bonds that is offset by an unfavorable entropy contribution, attributed either to solvent rearrangement or loss of ligand conformational flexibility.

An increase of several orders of magnitude in the affinity of carbohydrates to lectins can be achieved by suitable chemical derivatization, for example, by attachment to hydrophobic aglycones. Similar increase is obtained by their attachment to polymeric carriers, to form multivalent ligands such as dendrimers (see Chapter **3.36**).

For a small number of the microbial lectins, detailed information is available on their structure as well as that of their combining sites and their mode of interaction with ligands. It is based primarily on high-resolution X-ray crystallography of the lectins and their complexes with ligands (see Chapters **2.06** and **2.07**). Other inputs include binding experiments with different sugars and their derivatives using state-of-the-art techniques, site-directed mutagenesis of the proteins, and, to a limited extent, also NMR experiments and molecular modeling.

Such studies have shown that like the lectins themselves, the sites are diverse, even when their specificity is the same (although within a given protein family the sites may be similar).[8] In general, the sites appear to be preformed, since conformational changes occur rarely upon ligand binding. The amino acids that form the combining sites are not necessarily contiguous, and are brought together in space by the folding of the polypeptide chain. Changes in the spatial position of the amino acid residues may occur upon binding of the carbohydrate, bringing them to an orientation which improves stereocomplementarity with the ligand. The fit between the binding site on the protein and the carbohydrate is also affected by the shape of the ligand. Oligosaccharides are flexible molecules with considerable freedom of rotation around the glycosidic bonds connecting the individual monosaccharide constituents. As a result, they may assume different shapes, only one of which may fit the combining sites.

The types of bonds involved in the formation of lectin–carbohydrate complexes are in principle not different from those involved in the formation of complexes of proteins with other ligands, such as peptides, oligonucleotides, or various small molecules (see Chapter **3.21**). Binding between proteins and carbohydrates is stabilized primarily by a network of hydrogen bonds and hydrophobic interactions; in rare cases, electrostatic interactions (or ion pairing) and coordination with metal ions also play a role.[9–11] Bonding is sometimes mediated by one or more water molecules. Although in a single protein a limited set of amino acid residues contribute to the interactions with the ligand, in general most of the side chains of the 20 amino acids can participate in ligand binding.

3.28.2 Viral Lectins

Viruses contain sugar-specific surface proteins or glycoproteins that act as hemagglutinins and are therefore classified as lectins (**Table 1**). These lectins are structurally diverse and no common features can be discerned. Detailed information based largely on X-ray crystallography is available on the interaction with carbohydrates of influenza virus hemagglutinins. Complexes with ligands of the hemagglutinins of rotavirus, polyoma virus, and foot-and-mouth virus have also been well studied by the same technique. No other viral lectins have however been similarly characterized.

3.28.2.1 Influenza Viruses

Influenza A and B virions are studded with two different types of spikes, the hemagglutinin (HA), specific for *N*-acetylneuraminic acid that mediates receptor binding and pH-dependent membrane fusion and the neuraminidase (correct name, 'sialidase'), originally known as the viral receptor-destroying enzyme (RDE). In contrast, influenza C virus possesses only one type of spike, specific for 9-*O*-acetyl-*N*-acetylneuraminic acid, which combines all three functions, and is commonly referred to as the hemagglutinin-esterase-fusion protein (HEF). Influenza virus A is attracting much interest because an understanding of the structural determinants and molecular mechanisms involved in its binding to human cell receptors is central to the identification of viruses that pose a pandemic threat (see Chapter **4.25**).

The receptors recognized by influenza virus are sialic acids of cell surface glycoproteins and glycolipids. Sialic acids are usually found in either α2-3 or α2-6 linkages to galactose, the predominant penultimate sugar of N-linked carbohydrate side chains. The binding preference of a given HA for one or the other of these linkages correlates with the species specificity of the virus.[22,23] Thus, the HAs of all 15 antigenic subtypes found in avian influenza A and B viruses bind preferentially to sialic acid in α2-3 linkage, and it is this form of the sialosaccharide that predominates in avian enteric tracts where these viruses replicate. Swine influenza viruses bind sialic acid in α2-6 and sometimes also

Table 1 Viral lectins

Virus	*Specificity*	*References*
Flaviviruses		
Dengue virus	Heparan sulfate	12
Herpes viruses		
Herpes simplex	Heparan sulfate	13
Myxoviruses		
Orthomyxo		
Influenza A & B, human strains	Neu5Acα6Galβ4GlcNAc	See text
Porcine strains	Neu5Acα3/6Galβ4GlcNAc	”
Avian strains	Neu5Acα3Galβ4GlcNAc	”
Influenza C	$Neu5,9Ac_2α3Galβ4GlcNAc$	”
Paramyxo		14,15
Newcastle disease	Neu5Acα3Galβ4GlcNAc	
Sendai	Neu5Acα8Neu5Ac	
Rotavirus	Neu5Acα	See text
Nidoviruses		16,17
Coronavirus, bovine	$Neu5,9Ac_2$	
Totovirus, bovine	$Neu5,7,9Ac_3$	
Papoviruses		
Polyoma	Neu5Acα3Galβ4GlcNAc $Neu5Acα3Galβ3(Neu5Acα6)_{0,1}GalNAc$	See text
Picornaviruses		
Foot-and–mouth disease	Heparan sulfate	See text
Retroviruses		
HIV	ManOS, heparan sulfate	18–21

α2-3 linkages, and sialic acid in both linkages is detected in porcine tracheae. Human viruses of the H1, H2, and H3 subtypes that are known to have caused pandemics in 1918, 1957, and 1968, respectively, recognize 2-6-linked sialic acid, the major form found on cells of the human respiratory tract, and are now the focus of intense attention.

3.28.2.1.1 Influenza A virus hemagglutinin

This is the most thoroughly investigated viral lectin.[23,24] Its subunit is composed of two polypeptides, HA1 and HA2, m.w. 36 and 26 kDa, respectively, covalently linked by a single disulfide bond[22] (**Figure 1**). The hemagglutinin is a glycoprotein, with six N-linked oligosaccharides attached to HA1 and one to HA2. One of these glycans is an oligomannoside, while the others are complex bi- or triantennary structures, several of which contain sulfated galactose. X-ray crystallography showed that the hemagglutinin subunit consists of a hydrophilic, C-terminal domain, a hydrophobic membrane spanning region of 24–28 residues, an elongated triple-helical coiled stem, and a globular domain projecting 135 Å from the membrane. The globular domain consists of HA1 only, and contains the carbohydrate-binding site of the lectin. The subunits associate noncovalently to form trimers. The affinity of the hemagglutinin of the influenza virus for its receptor is downregulated by the two N-linked oligosaccharides attached to Asn123 and Asn129 in the vicinity of the combining site.[25] Thus, a mutant hemagglutinin devoid of these oligosaccharides bound to human erythrocytes much more strongly than the native lectin.

To date, crystal structures of several human, avian, and swine influenza hemagglutinins, mostly in complexes with the trisaccharide Neu5Acα6Galβ4Glc (Sia2-6Lac) or Neu5Acα3Galβ4Glc (Sia2-3Lac) and the pentasaccharide Neu5Acα6Galβ4GlcNAcβ3Galβ4Glc (LSTc) or Neu5Acα3Galβ4GlcNAcβ3Galβ4Glc (LSTa), analogs of the Neu5Acα2-6- and Neu5Acα2-3-human and avian receptors, respectively, have been determined. A complete description of all the structures obtained is beyond the scope of this chapter. Here, only selective examples will be given of the structures of HAs and how they interact with the corresponding ligands.

Early sequence studies have shown that for the HAs of the H2 and H3 human viruses, a minimum of two changes in receptor-binding site amino acids, Gln226 to Leu226 and Gly228 to Ser228, correlate with the shift from avian to human receptor binding.[26] By contrast, HAs of human H1 viruses acquire the ability to bind to human receptors while retaining Gln226 and Gly228. An understanding of how these changes occur, and a better insight into the combining sites of the HAs, have been provided by the recent study of John Skehel and co-workers of the structures of HAs from the 1918 pandemic virus (1918-human) with the use of HA expressed from the DNA recovered from tissues infected with virus in 1918 and from the prototype human (1934-human) and swine (1930-swine) H1 influenza viruses.[27]

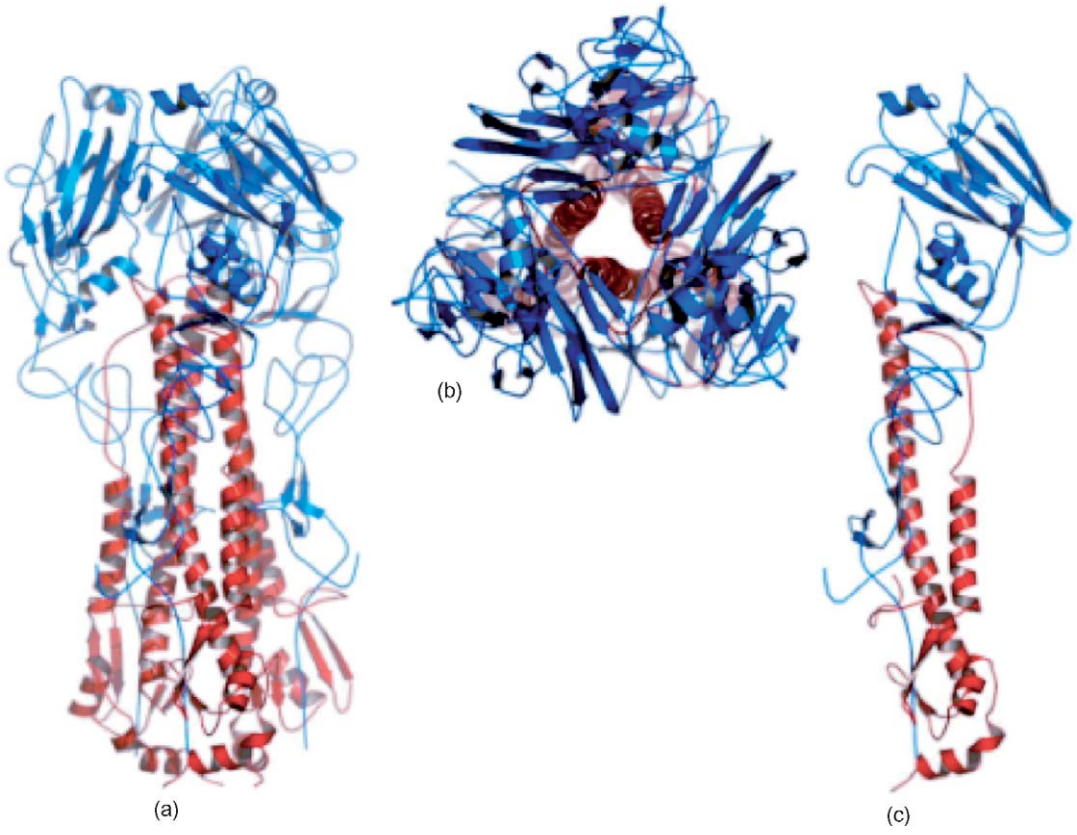

Figure 1 Influenza virus hemagglutinin trimer (a, side view, and b, top view) and monomer (c). In the monomer, the broad ribbon at the top represents the HA1 subunit and the narrow one the HA2 subunit. Courtesy J. J. Skehel, National Institute for Medical Research, Mill Hill, London.

- *Overall structure and receptor binding subdomain*. The overall trimeric structures of the HAs of the three H1 viruses (1918-human, 1934-human, and 1930-swine) HAs are similar (**Figure 2**), but they show notable differences to HAs of other subtypes with respect to the arrangements of the receptor-binding subdomain among others, both within the HA trimer and also within individual monomers. The receptor-binding sites are located as in other influenza A virus HAs at the membrane-distal tip of each subunit of the HA trimer. Three secondary structure elements, the 190 helix (residues 190–198), the 130 loop (residues 135–138), and the 220 loop (residues 221–228), form the sides of each site, with the base made up of the conserved residues Tyr98, Trp153, His183, and Tyr195. The conformations adopted by the 130 and 220 loops of the three H1 HAs are similar, but they are significantly different from those of the equivalent loops in the HAs of other influenza subtypes (**Figures 2b, 3,** and **4**). As observed with other HAs, the terminal sialic acids of the human and avian receptors interact with binding site residues through a series of conserved hydrogen bonds to the carboxyl and amide of the sialic acid (**Figure 3**).
- *The 1934-human HA/human receptor complex*. The electron density maps reveal well-ordered features for the Sia-**1**, Gal-**2**, and GlcNAc-**3** of the sialopentasaccharide (the numbers in bold refer to the positions of the monosaccharides in the ligand) in this complex (**Figure 3a**). Gal-**2** forms five hydrogen bonds that have not been previously observed in other HA/receptor complexes. Four bonds are possible between the 2- and 3-hydroxyls of Gal-**2** and the side chains of Lys222 and Asp225, and a fifth between the 4-hydroxyl of Gal-**2** and the main-chain amide of 227 that is mediated by a water molecule.
- *The 1934-human HA/avian receptor complex*. Again, only the Sia-**1**, Gal-**2**, and GlcNAc-**3** moieties of the sialopentasaccharide are ordered in this complex (**Figure 3b**). This observation is consistent with the results of hemagglutination assays showing dual binding specificity for this HA. The side-chain carbonyl of Gln226 forms a hydrogen bond with the 4-hydroxyl of Gal-**2**, as observed in other HA/avian receptor complexes. In addition, there is a previously unknown water-mediated interaction between the 4-hydroxyl of Gal-**2**, the main-chain carbonyl of residue 225, and the side chain of Lys222 (**Figure 3b**).

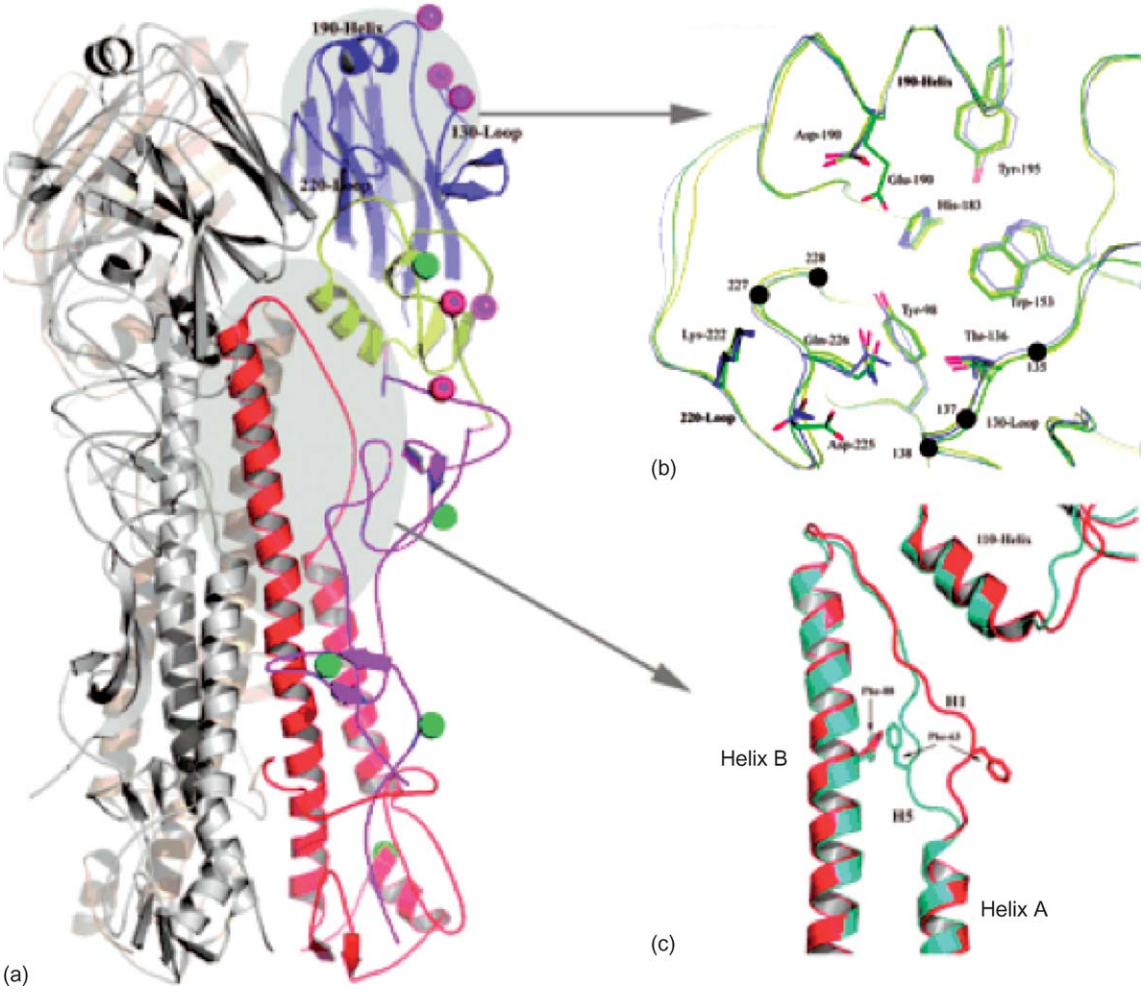

Figure 2 Structures of 1918-human, 1934-human, and 1930-swine HAs. a, Ribbon diagram of the trimer of 1918-human HA. Monomers 2 and 3 are in silver and gold, respectively, and the featured monomer is colored according to its individual subdomains: receptor binding (RB) in blue, vestigial esterase (E) in yellow, and fusion subdomains (F′ and F) in magenta and red, respectively. b, An expanded view of the superposed polypeptide backbone of the receptor-binding site of all three H1 HAs. The position of the three secondary structure units making up the site, the 190 helix, and the 130 and 220 loops, are indicated. Also shown are the side chains of some residues important for receptor binding. Certain Cα positions are indicated by black spheres for residues discussed in the text. Overall, the three H1 structures are very similar. c, An expanded view of a region of the F subdomain indicating differences between H1 and H5 subtype HAs in the position of the loop connecting helix A to helix B. Interactions between the C-terminal region of the loop and the RB and E subdomains (110 helix) influence the dispositions of the subdomains relative to the central coiled coil formed in the trimer by the B helices. Reproduced from Gamblin, S. J.; Haire, L. F.; Russell, R. J.; Stevens, D. J.; Xiao, B.; Ha, Y.; Vasisht, N.; Steinhauer, D. A.; Daniels, R. S.; Elliot, A.; Wiley, D. S.; Skehel, J. J. *Science* **2004**, *303*, 1838–1842.

- *The 1930-swine HA/human receptor complex*. The sialic acid of the receptor is located similarly in the 1930-swine and 1934-human HA receptor binding sites, but in the 1930 swine complex all five saccharides of the receptor analog are detected (Figure 3c). Lys222 again forms hydrogen bonds with the 2- and 3-hydroxyls of Gal-**2**, although in the present case this residue sits higher in the binding site. Asp190 hydrogen bonds to the amino nitrogen of GlcNAc-**3**, Ser193 hydrogen bonds to the 2-hydroxyl of Gal-**4**, and there is a water-mediated interaction between Thr189 and GlcNAc-**5**. The last three interactions have not been observed before in HA receptor complexes. In addition, the sialopentasaccharide exits the binding site in an orientation not previously seen, crossing the 190 helix near its N-terminus, about parallel to the threefold symmetry axis of the HA trimer.
- *The 1930-swine HA/avian receptor complex*. The electron density for the avian receptor analog bound to the 1930-swine HA is weak and mainly represents the sialic acid moiety (Figure 3d). A similar situation was observed for an

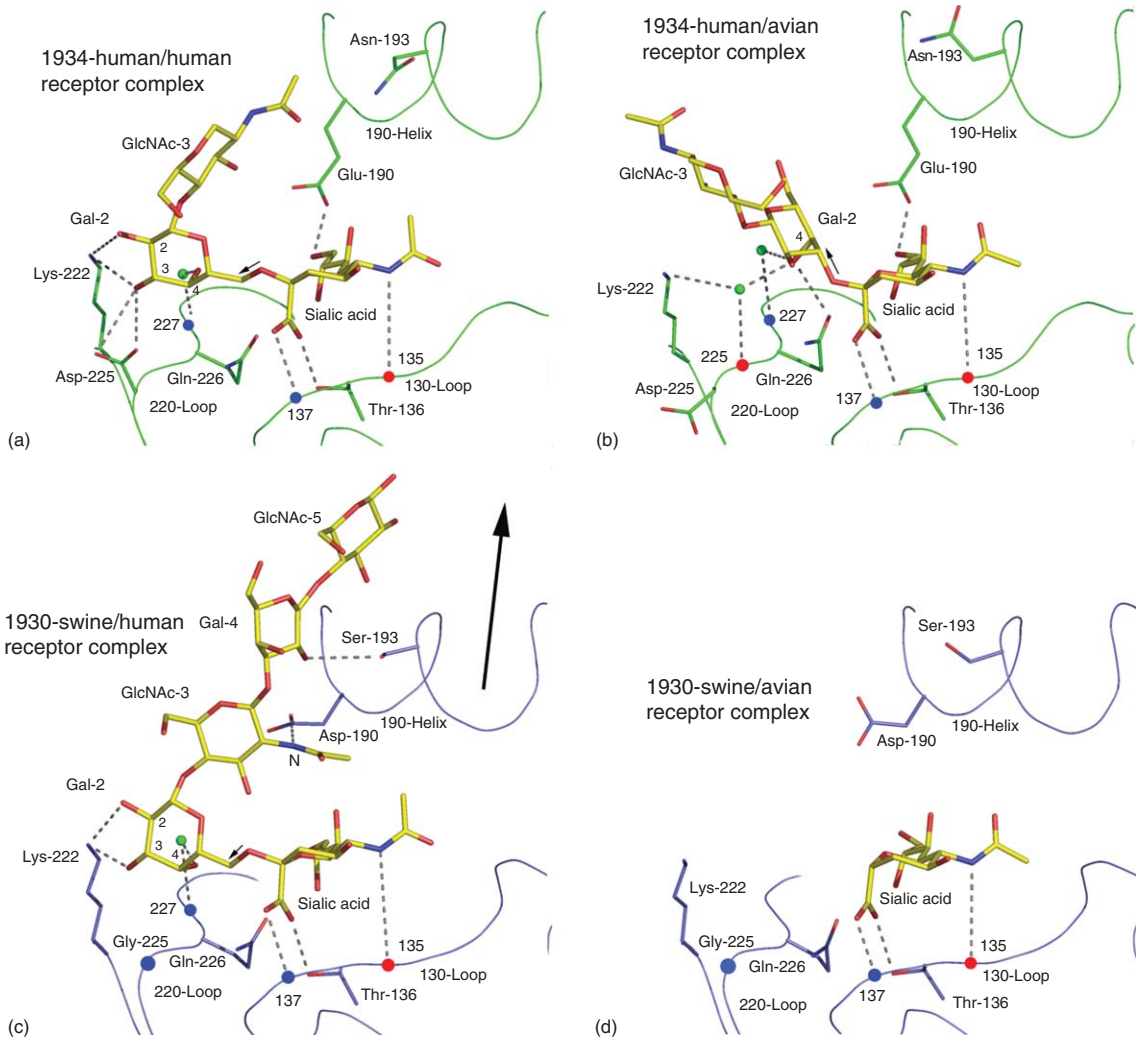

Figure 3 Interactions of 1934-human HA (top) and 1930-swine HA (bottom) with human receptor and with avian receptor analogs. The view of the receptor binding site is about the same as in **Figure 2b**. The three secondary structure components of the binding site are labeled in this backbone representation together with some of the most relevant side chains. The broken lines indicate potential hydrogen bond interactions between the protein and the receptors; residues making interactions via main-chain carbonyl groups are shown as red spheres, whereas those interacting via main-chain nitrogens are shown as blue spheres. In all four panels, the sialosaccharides are colored yellow for carbon atoms, blue for nitrogen, and red for oxygen. Water molecules are indicated by green spheres. a, 1934-human HA in complex with human receptor; b, 1934-human HA in complex with avian receptor; in both cases, the HA is colored in green for backbone and carbon atoms. c, 1930-Swine HA in complex with human receptor; d, 1930-swine HA in complex with avian receptor; in these cases, the backbone and carbon atoms of the HA are colored in blue. The small black arrows in (a), (b), and (c) indicate that for the two human receptor complexes the Sia-**1**–Gal-**2** linkage adopts a *cis*-conformation about the glycosidic bond, whereas for the avian complex it adopts a *trans*-conformation. The large black arrow in (c) indicates the direction of an axis parallel to the trimer threefold axis. Reproduced from Gamblin, S. J.; Haire, L. F.; Russell, R. J.; Stevens, D. J.; Xiao, B.; Ha, Y.; Vasisht, N.; Steinhauer, D. A.; Daniels, R. S.; Elliot, A.; Wiley, D. S.; Skehel, J. J. *Science* **2004**, *303*, 1838–1842.

H5 avian HA in complex with a human receptor analog, where only a subset of the atoms for the sialic acid could be located. These observations probably reflect the low affinity of the HAs for their respective ligands, consistent with the preference of the 1930-swine virus for human receptor in hemagglutination assays.

- *Human receptor complexes*. Complexes of the human receptor analog bound to 1934-human HA (green) and 1930-swine HA (blue) and to human H3 HA (red) are superimposed in **Figure 4a**. Perhaps the most important feature of this comparison is the difference in structure adopted by the 130 and 220 loops of the receptor-binding site between

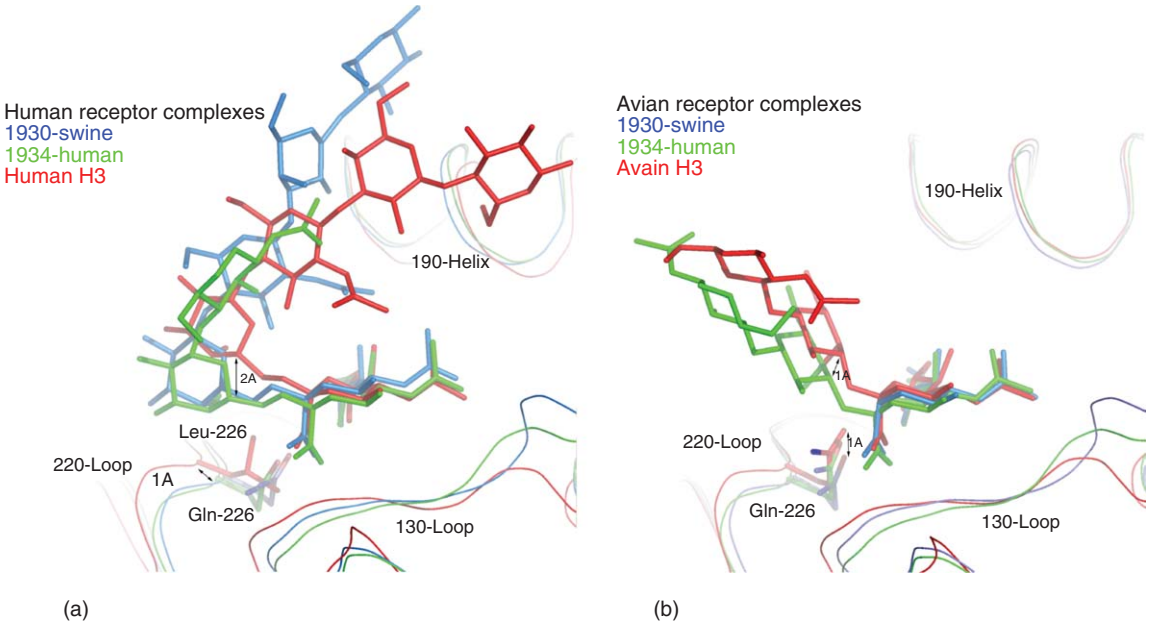

Figure 4 Differences in the orientations of bound receptors in the receptor-binding sites of three different HAs. The receptor-binding sites of 1934-human HA (green), 1930-swine HA (blue), and human H3 HA (red) are overlaid. The view matrix is about the same as in **Figure 3**. The sialopentasaccharides are colored according to the HAs to which they are bound. The side chains of Gln226 (H1 HAs) and Leu226 (H3 HA) are shown. a, Human receptor complexes. b, Avian receptor complexes. Reproduced Gamblin, S. J.; Haire, L. F.; Russell, R. J.; Stevens, D. J.; Xiao, B.; Ha, Y.; Vasisht, N.; Steinhauer, D. A.; Daniels, R. S.; Elliot, A.; Wiley, D. S.; Skehel, J. J. *Science* **2004**, *303*, 1838–1842.

the H1 and H3 HAs. One consequence of the change in the 130-loop structure is that the sialic acid of the receptor is tilted about 10° into the receptor-binding sites of the H1 HAs. This effect, together with different orientations about the glycosidic bond, contributes to Gal-**2** being located almost 2Å lower in the H1 HAs than in the human H3 HAs. Gal-**2** is able to adopt this position because structural differences in the 220 loop locate Gln226 lower in the binding site than the equivalent Leu226 of human H3 HA. Consequently, in the H1 HAs, Gal-**2** is located closer to the 220 loop and is able to form hydrogen bonds with Lys222. In the case of the 1934-human HA, Gal-**2** also interacts with Asp225. Thus, a combination of factors relating to the structure of the 130 and 220 loops enable the H1 HAs to make favorable hydrogen bond interactions with Gal-**2** of the human receptor. Gln226 plays an essentially passive role in this process, in marked contrast to the role played by Leu226 in the binding of human H3 HA to human receptor. In that case, Gal-**2** makes hydrophobic contacts with Leu226, and the higher position and the nature of this side chain are important for human receptor binding.

- *Avian receptor complexes*. Complexes of avian receptor analogs with 1934-human HA (green), 1930-swine HA (blue), and an avian H3 HA (red) are overlaid in **Figure 4b**. Again, the differences in the structure of the 130 loop between the H1 and H3 HAs result in the sialic acid of the avian receptor being located lower in the receptor binding site of the H1 HAs. Comparison of the 1934-human and avian H3 complexes also reveals that Gal-**2** of the avian receptor is located about 1Å lower in the binding site of the H1 complex, as is Gln226. In both complexes, the 4-hydroxyl of Gal-**2** hydrogen bonds with the side-chain carbonyl of Gln226 (**Figure 4b**), and the coordinated differences in position of the bound receptor and Gln226 enable this interaction to be conserved. It seems therefore that the 1934-human HA is able to bind the avian receptor in a manner reminiscent of avian HAs, with Gln226 playing a key role.

The above data for the different complexes show the overall structures of 1930-swine and 1934-human HA receptor-binding sites are very similar (**Figure 2b**), and that both contain a glutamine residue at position 226. It has been suggested that the 1930-swine HA binds less effectively to avian receptors than the 1934-human HA because the position adopted by Gln226 in 1934-human HA is about 1Å higher in its complex with the avian receptor than it is either in the human receptor complex or in the uncomplexed form (**Figure 5**). By contrast, the position of Gln226 in 1930-swine HA is about the same as uncomplexed and in the human and avian receptor complexes. The apparent

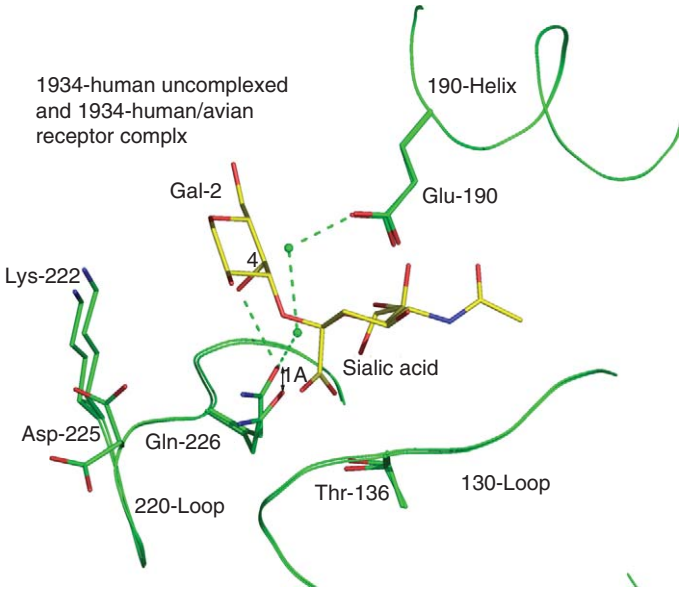

Figure 5 Superimposition of the binding site of 1934-human HA in its uncomplexed state and complexed with avian receptor analog. The HA is shown in green in both cases, and the avian receptor is colored as in **Figure 3**. Two water molecules, shown as green spheres, link Glu190 to Gln226 in the avian receptor complex. This hydrogen-bonded network is not formed in the uncomplexed structure or in the human receptor complex not shown. Reproduced from Gamblin, S. J.; Haire, L. F.; Russell, R. J.; Stevens, D. J.; Xiao, B.; Ha, Y.; Vasisht, N.; Steinhauer, D. A.; Daniels, R. S.; Elliot, A.; Wiley, D. S.; Skehel, J. J. *Science* **2004**, *303*, 1838–1842.

inability of Gln226 to adopt a higher position in the receptor-binding site seems to explain the failure of 1930-swine HA to interact as effectively with the avian receptor. This explanation is supported by the structural observation that in the 1934-human HA complex with avian receptor Glu190 interacts through two water molecules with Gln226 (**Figure 5**). This network of hydrogen bonds may be necessary to position Gln226 in the binding site for its interaction with Gal-**2**. Glu190 is conserved in avian H1 HAs, all of which specifically bind α2-3-linked receptors. By contrast, residue 190 of 1930-swine HA is an aspartic acid, which does not interact with either the 9-hydroxyl of Sia-**1** or Gln226 and is thus unable to facilitate binding to avian receptor.

Irrespective of the single amino acid difference of Asp or Gly at residue 225 between the sequences of 1918-human HAs, by recognizing human receptors, all would fulfill the first requirement of an epidemic virus: the ability to spread in the human population. The importance of this requirement was emphasized in the 1997 outbreak of H5 'chicken' influenza in Hong Kong, when the virus was extremely virulent but did not acquire the ability to bind 2-6-linked sialosides and was therefore unable to spread. With the ability to ensure the efficiency of the initial stages of virus infection, coupled with novel antigenicity, the human-1918 HA may have been the prime determinant of extensive mortality in the 1918 pandemic.

Previous crystallographic analysis of the trisaccharide human receptor analog, Sia2-6Lac, bound to human H3 HA, revealed an extended conformation[24] rather than the folded conformation of the pentasaccharide human receptor analog, LSTa. This conformation was presumed not to be representative of natural receptors because of the presence of glucose rather than *N*-acetylglucosamine at position **3** and the lack of glycosylation beyond this position. The mode of Sia2-6Lac binding, however, further demonstrates the plasticity of the HA receptor binding site in accommodating oligosaccharides in different conformations. By contrast, Sia2-6Lac binds in a manner very similar to that of the pentasaccharide, LSTa, again suggesting that there is no role in avian receptor binding for the saccharides beyond GlcNAc-**3**. It is therefore possible that some avian HAs are capable of binding sialic acid in α2-6 linkage to galactose, but are not capable of infecting humans because the saccharides beyond GlcNAc-**3** sterically clash with the HA. For an avian HA to change its binding specificity from avian to human receptors may therefore involve two sorts of changes, the first, to allow binding of the α2-6-linkage, either via mutation of Gln226 to Leu226 as in H2 and H3 HAs or via a specific positioning of Gln226 as in H1, and the second, to accommodate the additional saccharides linked to Gal-**2** in the natural receptor side chain.

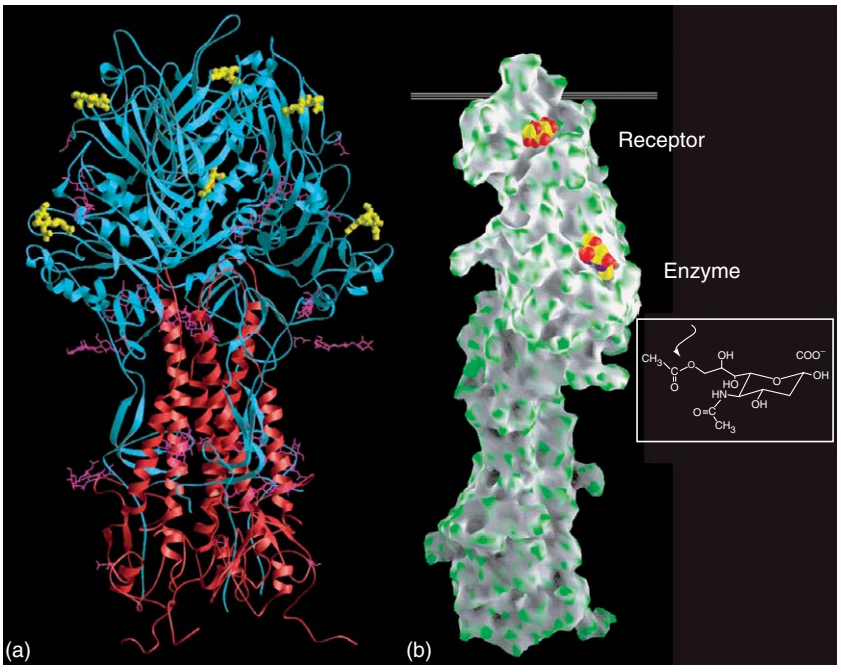

Figure 6 HEF glycoprotein structure. a, The structure of the HEF trimer. HEF1 (blue), HEF2 (red), receptor analog and enzyme inhibitor ligands, (yellow); N-linked carbohydrate ball and stick (purple). HEF1 is linked to HEF2 by a disulfide bond from Cys6 of HEF1 to Cys137 of HEF2. b, Monomer surface of HEF showing 9-O-acetylsialoside receptor-binding site (top) and 9-*O*-acetylesterase site (bottom). Inset, the esterase removes the acetyl group of 9-*O*-acetylsialic acid (see arrow). Reprinted by permission from Macmillan Publishers Ltd: *Nature* (Rosenthal, P. B.; Zhang, X.; Formanowski, F.; Fitz, W.; Wong, C. H.; Meier-Ewert, H.; Skehel, J. J.; Wiley, D. C. *Nature* **1998**, *396*, 92–96.), Copyright (1998).

3.28.2.1.2 Influenza C HEF

Like its homologous influenza A and B virus hemagglutinins, to which it bears limited primary sequence identity, HEF is a homotrimer of an N-glycosylated type I membrane glycoprotein. It consists of a 65 kDa N-terminal subunit HEF1 and a 30 kDa C-terminal subunit HEF2 that is membrane anchored. Despite an overall primary sequence identity of only 11%, the similarity between the three-dimensional (3-D) structure of HEF and the hemagglutinin is astounding (**Figure 6**).[16,28,29] The main differences are within the N-terminal subunits, which form the globular part of the spikes and which, in the case of HEF, contain the receptor-binding R and RDE esterase domains. The C-terminal HEF2 subunit, together with segments of HEF1, comprises the fusion domain F. In the trimer, HEF2 subunits constitute the elongated membrane-anchored stalk, at the core of which long central helices form a triple-stranded interphase.

Each HEF1 subunit contains two sites that can accommodate Neu5,9Ac_2. The actual receptor-binding sites are cavities at the tip of the globular head domain with Tyr127, Thr170, and Gly172 involved in ligand binding. Interaction with the 9-*O*-acetyl group, the most critical determinant for receptor recognition, is mediated among others by the hydroxyl group of Tyr224 and the guanido group of Arg236, which contact the acetyl carbonyl oxygen. Moreover, Phe225, Phe293, and Pro271 form a nonpolar pocket into which the acetyl methyl group can be fitted (**Figure 7**). The importance of the latter site for receptor recognition is underlined by the observations that mutations of residues adjacent to Pro271 (Asp269Asn, Thr270Leu or Ile, and Thr272Ile) are associated with increased binding efficiency to Neu5,9Ac_2 receptors.

3.28.2.2 Rhesus Rotavirus

The rotavirus hemagglutinin binds the alpha anomer of *N*-acetylneuraminic acid with an association constant (K_a) of $8 \times 10^5 M^{-1}$. It requires no additional carbohydrate moiety for binding to the sialic acid, does not distinguish between 3′ and 6′sialyllactose, and has approximately 10-fold lower affinity for *N*-glycolylneuraminic acid than for *N*-acetylneuraminic acid.[30] The rotavirus outer capsid is made of the coat glycoprotein VP7 and the spike protein VP4. Trypsin cleaves VP4 into an N-terminal fragment, VP8*, consisting of amino acids 46–231, that contains the viral hemagglutinin

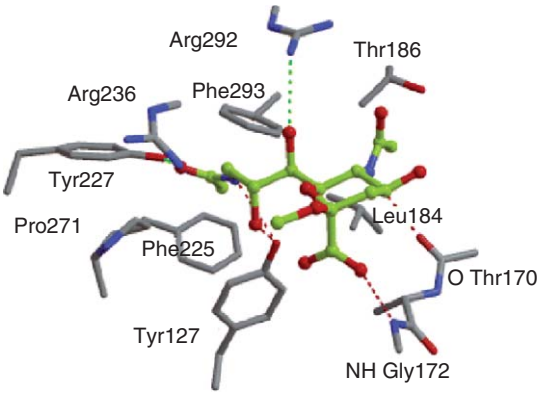

Figure 7 HEF receptor binding. Ligand bound to the receptor-binding site. Potential hydrogen bonds are indicated in green or red for those conserved in HA ligand binding. Four polar contacts are formed with the ligand identically in HEF and HA: two from the hydroxyl group of HEF1 Tyr127 (Y98 in HA1) to the 8-hydroxyl and 9-amide of the ligands and two from main-chain atoms: the carbonyl oxygen of HEF1 residue 170 (135 in HA1) to the 5-amide of the ligand and the amide of HEF1 172 (137 in HA1) to the carboxylate of the ligand. The acetyl methyl group binds in a nonpolar pocket unique to HEF, formed by Phe225 and 293, and Pro271; the acetyl carbonyl oxygen contacts the hydroxyl group of Tyr224 and the guanidino group of Arg236. Reprinted by permission from Macmillan Publishers Ltd: *Nature* (Rosenthal, P. B.; Zhang, X.; Formanowski, F.; Fitz, W.; Wong, C. H.; Meier-Ewert, H.; Skehel, J. J.; Wiley, D. C. *Nature* **1998**, *396*, 92–96.), Copyright (1998).

(residues 93–208) and a C-terminal fragment, VP5*, the function of which is to permeabilize the membrane to which the virus binds. No crystals of free VP8* could be obtained, but it crystallized in complex with methyl α-linked *N*-acetylneuraminic acid. NMR analyses of VP8* in solution, and X-ray studies of its complex with the above ligand, revealed the same basic protein structure.[31] It is a single, compactly folded globular domain with two cysteines (C203 and C216) and two prolines (P68 and P182) in the *cis*-configuration. The crystals contain a mixture of molecules with the peptide bond G156–P157 in either the *cis*- or *trans*-configuration.

The tight fold of the β-sandwich, the cross-bracing of the β-sheets by the β-ribbon and the C-terminal α-helix, the short loops between the strands and the dense hydrophobic cores between the major structural elements all suggest a compact, rigid structure, that accounts for the protease resistance and stability of the VP8* core (**Figure 8**). The β-sandwich of the rotavirus sialic acid binding domain has the same fold as the S-carbohydrate recognition domain (S-CRD) of the galectins, despite the absence of a significant sequence similarity between the two (9% identity with human galectin-3 in structurally equivalent residues).

The sialic acid-binding site of the rhesus rotavirus hemagglutinin lies above the cleft between the two β-sheets and appears to be an open-ended, shallow groove. Of the four amino acids (Arg101, Tyr155, Tyr188, and Ser190) seen to be involved in sialic acid binding, the last three were previously identified as likely ligand-binding residues by mutagenesis studies. In addition to the seven hydrogen bonds (**Figure 9**), the sialic acid makes several van der Waals contacts with the side chains of six amino acids, three of which are tyrosines.

3.28.2.3 Polyoma Virus

This virus is a nonenveloped, icosahedrically symmetrical particle, with a circular, double-stranded DNA genome.[32] Its carbohydrate binding site is located in viral protein 1 (VP1, m.w. *c.* 42kDa), the major constituent of the outer shell (capsid) of the virion. Each virion contains 360 copies of VP1, arranged in pentamers. VP1 has two antiparallel β-sheets with a topology that resembles the jellyroll fold; some loops that connect the β-strands are extensive and contain additional secondary structural elements (**Figure 10**). The most striking feature of the capsid is the way the individual pentamers are tied together by the C-terminal arms of the monomers; the last 63 residues emerge from each monomer and protrude into a subunit of another pentamer, where they form a β-strand that augments a sheet in the target subunit.

Two types of strain of murine polyoma virus are known that differ in their tumorogenicity in mice and in their specificity for sialic acid oligosaccharides. The critical difference in the structure of the viral protein (VP1) that contains the carbohydrate-binding site of these strains is in residue 91, which is glycine in the poorly tumorogenic strains and glutamic acid in the highly tumorogenic ones.[33] Crystallographic studies at low resolution of the intact poorly tumorogenic virus particle in complex with Sia2-3Lac[34] and with a branched hexasaccharide containing both α2-3- and α2-6-linked *N*-acetylneuraminic acid[35,36] have located the combining site in a shallow groove and have

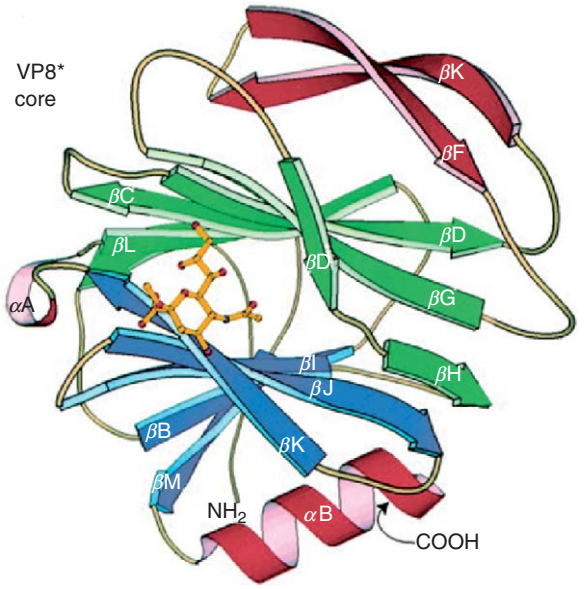

Figure 8 Ribbon diagram of the sialic acid-binding domain of rotavirus. The central structural feature of VP8* is an 11-stranded antiparallel β-sandwich, formed from a five-stranded and a six-stranded β-sheet, with an interrupted top strand. The two β-sheets are joined by five short intersheet loops, as well as by a brief stretch of a parallel β-structure between strand βH′ of the six-stranded sheet and strand βJ of the five-stranded one. The cleft between the sheets is filled by a dense core of hydrophobic side chains, contributed by all strands of the sheets, except for βH. The domain contains three other structural elements, namely a short α-helix (αA) as part of the intersheet loop, a longer α-helix (αB) at the C-terminus, and an extended β-ribbon made up of strands βE and βF. Sialic acid is shown as a stick and ball model. Reproduced with permission from Dormitzer, P. R.; Sun, Z. Y.; Wagner, G.; Harrison, S. C. *Embo J.* **2002**, *21*, 885–897.

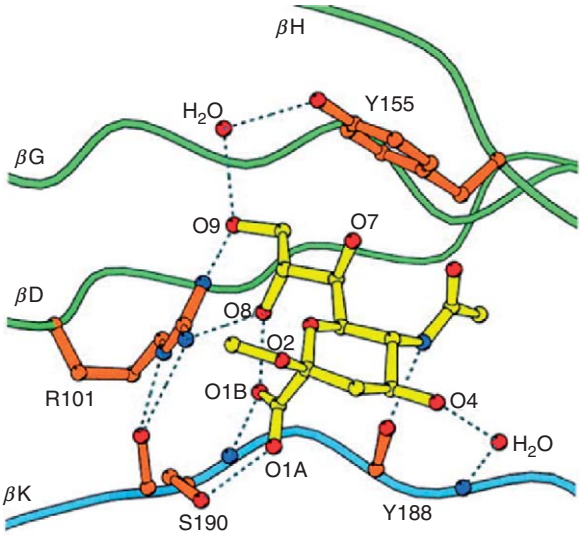

Figure 9 Binding of MeαNeu5Ac by rhesus rotavirus. The hydrogen bonds anchoring the ligand to the protein are: two bonds from the side-chain guanidinium of Arg101 to the ligand glycerol chain (atoms O8 and O9); one from the side-chain hydroxyl of Tyr155 to the glycerol side chain (O9) via a water bridge; one from the side-chain hydroxyl of Ser190 to the carboxylate (atom O1A); one from the main-chain amide of Ser190 to the carboxylate (atom O1B); one from the main-chain carbonyl of Tyr188 to the acetamide nitrogen of the ligand; and one from the main-chain amide to 4-OH via a water bridge. The ligand is shown as a ball-and-stick model; hydrogen bonds are indicated by dotted lines. From Dormitzer, P. R.; Sun, Z. Y.; Wagner, G.; Harrison, S. C. *Embo J.* **2002**, *21*, 885–897.

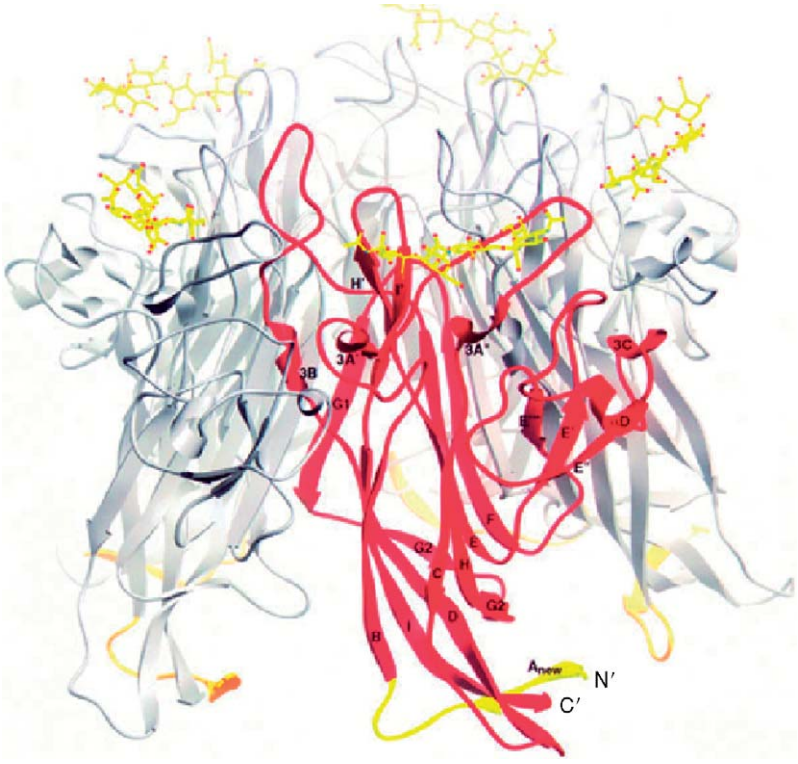

Figure 10 Structure of the recombinant VP1 pentamer of polyoma virus. Ribbon drawing of the VP1 pentamer complexed with the disialylated oligosaccharide Neu5Acα3Galβ3(Neu5Acα6)GlcNAc. One monomer is shown in red, the others in gray. The receptor fragments are shown as ball-and-stick models. The rearranged N-terminal segments are shown in yellow (red monomer) and orange. Reproduced from Stehle, T.; Harrison, S. C. *Embo J.* **1997**, *16*, 5139–5148.

shown that the sialic acid as well as the galactose form contacts with the protein (Figure 11). The inability of the highly tumorogenic strains to bind the branched ligand is due to electrostatic repulsion between the side chain of Glu91 and the carboxylate of the α2-6-linked *N*-acetylneuraminic acid. These results have been confirmed and extended in a high-resolution crystallographic study of a pentamer of VP1 from the poorly tumorogenic virus in complex with the above pentasaccharide.[32]

3.28.2.4 Foot-and-Mouth Disease Virus

This virus has the basic picornavirus structure (Figure 12). Its icosahedral capsid comprises 60 copies each of four virus-encoded proteins, VP1–VP4. X-ray crystallography of the complex of the virus with its ligand, heparin, revealed that the combining site is located in a shallow depression on its surface, consists of contributions from the three major capsid proteins, VP1, VP2 and VP3 and is located at the junction of these proteins. VP1, VP2, and VP3, virtually identical for the type A and O viruses.[37,38] There were virtually no changes in the protein to accommodate the sugar. In the complex with a sulfated heparin trisaccharide $HSO_3$6GlcNH(HSO_3)α[(HSO_3)Idoα][(HSO_3)2GlcN(HSO_3)], the key binding residue was Arg56 of VP3, which interacts ionically with two of the sulfates of the trisaccharide. The importance of this residue is supported by the finding that field isolates of the virus, where the arginine is replaced by histidine, do not bind heparin. Additionally, nonionic interactions observed in the virus–ligand complex include bonds between His195 of VP1 and the iduronic acid, and a stacking interaction between the imidazole ring of this amino acid and the hydrophobic face of the glucosamine that follows the iduronic acid.[38] In strain O_1BFS, the most tightly bound heparin residues were GlcN2, Ido3, and GlcN4, and, for $A10_{61}$, these were the only ones for which there was clear electron density.

The ligands were the basic and polar residues Arg56 and Asn88 of VP3, Thr134 and Arg135 of VP2, and Lys193 of VP1, and bridging water molecules were important (Figure 13b). As for virus strain O_1BFS, there was no direct involvement of the VP1 C-terminal residues 200RHKQI205 in binding; however, the neighboring residues His195 (O_1BFS) (Figure 4b) and Lys-193 ($A10_{61}$) of VP1 (Figure 13a) did contact the heparan substrate (HS), and the C-terminus may stabilize these residues in a suitable position for HS binding.

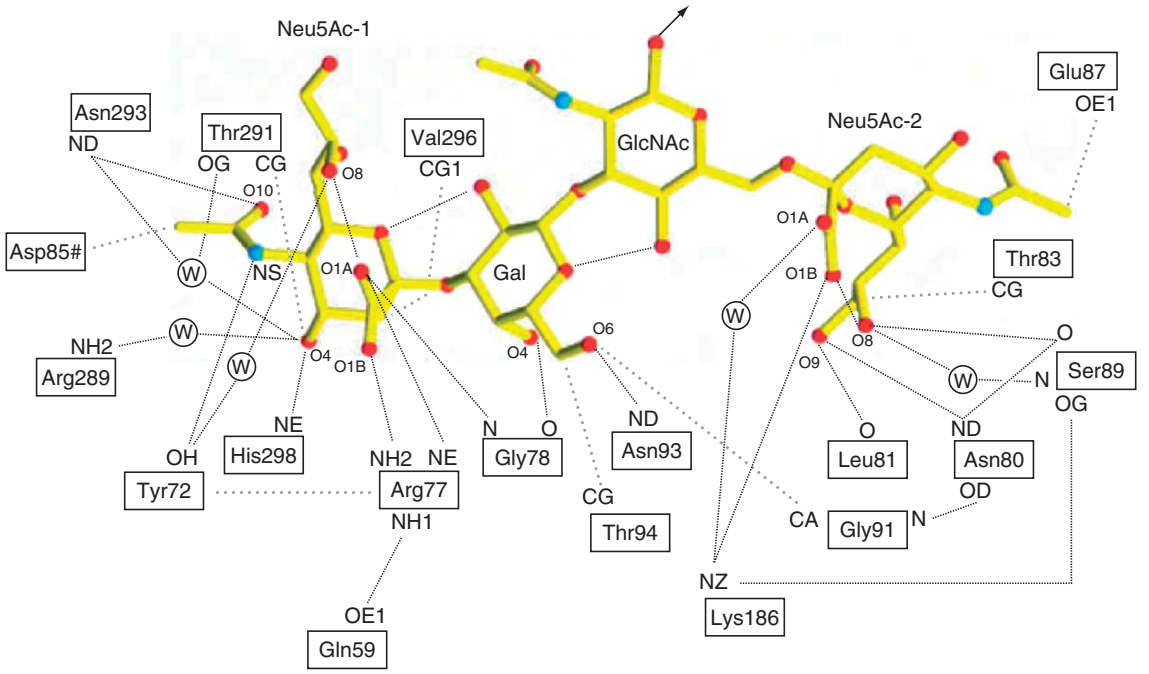

Figure 11 Schematic view of the combining site of polyoma virus with the bound tetrasaccharide Neu5Acα3Galβ3(Neu5Acα6)GlcNAc. Hydrogen bonds are represented by thin broken lines, and hydrophobic contacts are shown as thick gray broken lines. Asp85, located at the tip of the BC2-loop of the clockwise VP1 neighbor, approaches the *N*-acetyl group of Neu5Ac-**1**. The small circles labeled W′ represent water molecules Reproduced from Stehle, T.; Harrison, S. C. *Embo J.* **1997**, *16*, 5139–5148.

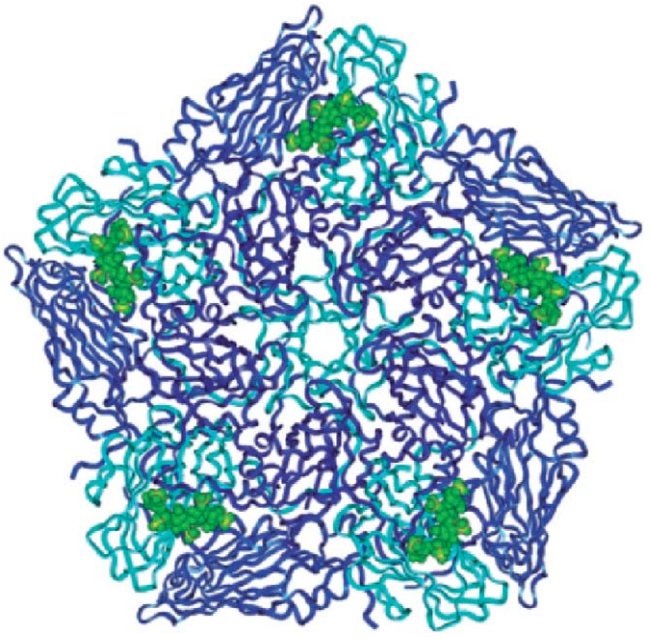

Figure 12 Model of foot-and-mouth disease virus in complex with the heparin trisaccharide $HSO_3$6GlcNH(HSO_3)α(HSO_3)Idoα][(HSO_3)2GlcN(HSO_3)]. The structure shown is a pentamer; 12 such structures make up the icosahedral viral capsid. The trisacharide is shown as a CPK model. Reproduced from Mulloy, B.; Linhardt, R. J. *Curr. Opin. Struct. Biol.* **2001**, *11*, 623–628.

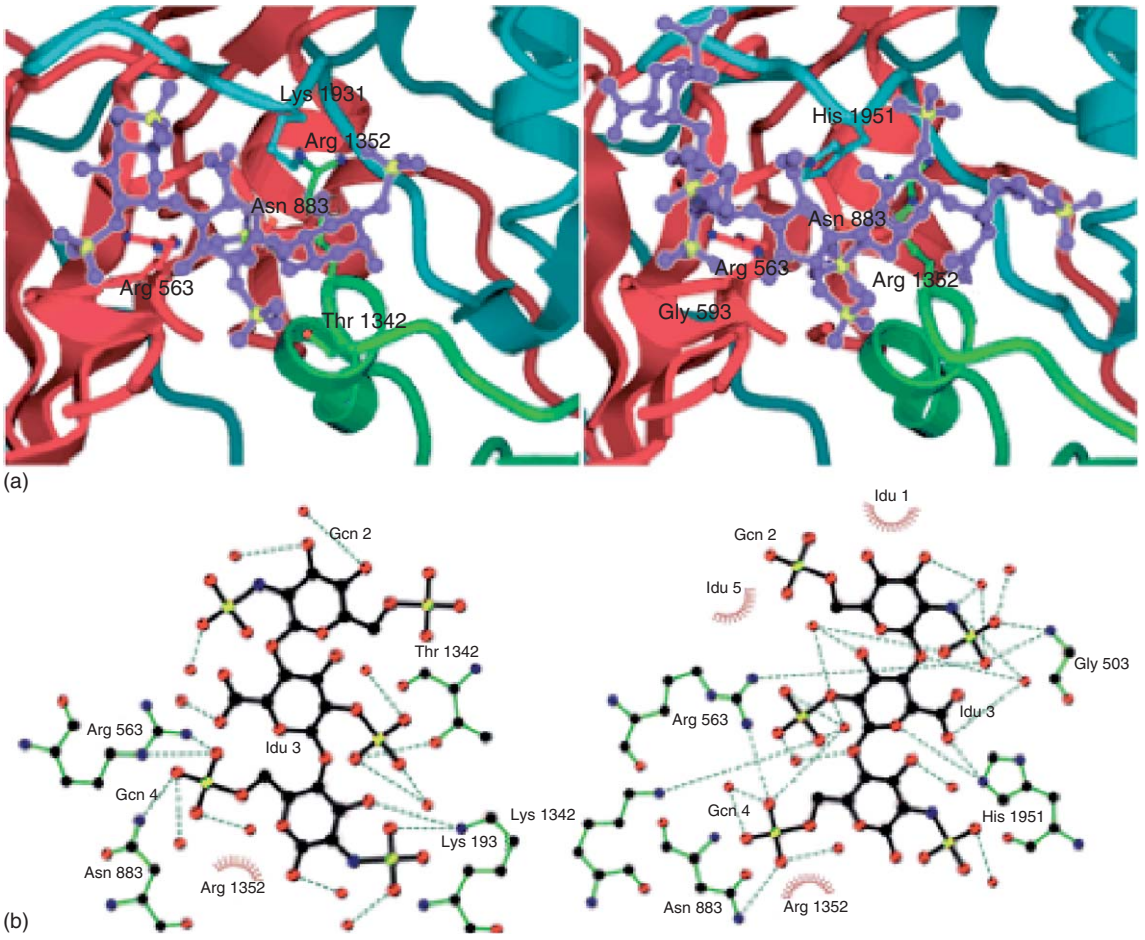

Figure 13 Heparin-binding site of different strains ($A10_{61}$ and O_1BFS) of foot-and-mouth virus. a, Ligand-binding site in $A10_{61}$ (left) and O_1BFS (right). Viral proteins are depicted as ribbons with standard color coding. The side chains of ligand residues (labeled) are drawn as balls and sticks in the same color as the corresponding protein. They are labeled with the residue number and the chain ID as the least significant digit,for example, 563 is residue 56 of VP3. The sugar residues are depicted as purple balls and sticks, with sulfur atoms in yellow. b, Sugar–protein interactions. $A10_{61}$ (left) and O_1BFS (right) are shown each with the pentasulfated trisaccharide GlcN-IdoA-GlcN (its constituents marked as Gcn 2, Idu 3, and Gcn 4, respectively). Only protein side chains that interact directly are shown and these are rearranged to clarify the hydrogen-bonding pattern. Ligand bonds are drawn in purple and nonligand bonds in brown. Hydrogen bonds are depicted by olive-green, dashed lines. Nonligand residues involved in hydrophobic contacts are shown as red, fringed semicircles. Water molecules are colored red. Reproduced from Fry, E. E.; Newman, J. W.; Curry, S.; Najjam, S.; Jackson, T.; Blakemore, W.; Lea, S. M.; Miller, L.; Burman, A.; King, A. M.; *et al*. *J. Gen. Virol.* **2005**, *86*, 1909–1920.

Two ligating residues were conserved between the type A and O complexes: Arg-56 of VP3 and Arg-135 of VP2 (**Figures 13a** and **13b**). The former switches from a histidine to an arginine in type O viruses on adaptation to tissue culture and is a key ligand in both complexes. In the O_1BFS complex, it acts as a bidentate ligand, stabilizing sulfate groups from rings 2 and 4, whereas in the $A10_{61}$ complex, it only interacts with ring 4. Arg-135 makes a hydrophobic interaction in both complexes and appears to polarize Asn-88 of VP3, increasing its affinity for HSO_3-6-NH-SO_3.

3.28.3 Bacterial Lectins

3.28.3.1 Surface-Bound Lectins

A variety of bacterial species and genera express surface lectins, frequently of more than one type and with distinct specificities[2,3,39] (**Table 2**). A few bacteria are known to produce intracellular and soluble lectins. The diverse

Table 2 Bacterial surface lectins

Organism	*Carbohydrate*
C. jejuni	Fucα2Galβ4GlcNAc
E. coli Type 1	Manα3Manα6Man
P	Galα4Gal
S	Neu5Acα3Galβ3GalNAc
CFA/1	Neu5Acα8
F1C	GalNAcβ4Galβ
F17	GlcNAc
K1	GlcNAcβ4GlcNAc
K99	Neu5Acα3Galβ4Glc
H. influenzae	$Neu5Ac\alpha3_{0,1}Gal\beta4GlcNAc\beta3Gal\beta4GlcNAc$
H. pylori	Neu5Acα3Galβ4GlcNAc Fucα2Galβ3(Fucα4)Gal
K. pneumoniae	Man
N. gonorrhoea	Galβ4GlcNAc
N. meningitidis	$Neu5Ac\alpha3_{0,1}Gal\beta4GlcNAc\beta3Gal\beta4GlcNAc$
P. aeruginosa	Galβ3Glc(NAc)β3Galβ4Glc
S. typhimurium	Man
S. pneumoniae	$Neu5Ac\alpha3_{0,1}Gal\beta4GlcNAc\beta3Gal\beta4GlcNAc$
S. suis	Galα4Galβ4Glc

Modified from Sharon, N. Biochim. Biophys. Acta 2006, 1760, 527–537.

specificities of most of the surface lectins are among the factors determining the organ and animal tropism of the bacteria. It is not known whether individual cells co-express multiple lectins or if each lectin is confined to a distinct cell subpopulation of the species. We know, however, that the expression of the surface lectins is commonly regulated by a mechanism known as phase variation, that controls the back-and-forth conversion of the lectin-expressing cells in a bacterial population to nonexpressing ones.[40] In *Escherichia coli, Klebsiella pneumoniae*, and *Salmonella* spp., the lectins often are in the form of submicroscopic hair-like appendages, named fimbriae or pili, that protrude from the surface of the cells. During the fimbriated phase, a typical Gram-negative bacterium carries 200–500 peritrichously arranged fimbriae.

The most prevalent and best characterized bacterial surface lectins with respect to structure, biosynthesis, and function are those expressed by *E. coli*, namely the mannose-specific type 1 fimbriae, the galabiose-specific P fimbriae, and the *N*-acetylglucosamine-specific F17 fimbriae. They all consist of an assembly of different classes of subunit and belong to the family of flexible 'Cup' fimbriae, referred to as such because they are assembled by the chaperone/usher pathway.[41] The carbohydrate-binding subunit is typically located at the end of the assembled structure.

3.28.3.1.1 Type 1 fimbriae

Type 1 fimbriae (**Figures 14** and **15**) are expressed by a large number of *E. coli* strains, and are found in more than 95% of *E. coli* isolates from intestinal and extraintestinal infections such as urinary ones.[1,3,39,42] They are also produced by other enterobacterial species, among them of *K. pneumoniae, Salmonella typhimurium*, and *Salmonella enteritidis*. The affinity of different phenotypes of *E. coli* type 1 or of the isolated fimbriae to mannose or MeαMan may differ within a factor of 15 and they can be functionally subdivided into either low-mannose-binding (M1L) or high-mannose-binding (M1H) phenotypes.[43] *E. coli* exhibiting these two basic phenotypes have been found to predominate in different niches. Most isolates from the large intestine of healthy humans (*c.* 80%) express a distinct M1L phenotype, whereas most isolates from urinary tract infections (>70%) express M1H variants. These naturally occurring variations dramatically change the tissue tropism of *E. coli* and can be a major factor in shifting the bacterial adaptation from commensal to pathologic habitats, a phenomenon known as pathoadaptation.[44] Typically, type 1 fimbriae of *E. coli* have a considerably higher affinity (up to 40 times as compared to mannose) for oligosaccharides such as Manα3-Manβ4GlcNAc or Manα6(Manα3)Manα6(Manα3)Man that are constituents of cell surface glycoproteins.[45]

In the urinary tract, the fimbriae mediate binding of the bacteria to the oligomannosides $Man_9GlcNAc_2$ to $Man_6GlcNAc_2$ of uroplakin Ia, a major glycoprotein of urothelial apical plaques.[46,47] Anchorage of *E. coli* to the urothelial surface via type 1 fimbriae–uroplakin I interactions may play a role in their colonization of the bladder and eventual ascent through the ureters, against urine flow, to invade the kidneys.

Structurally, the fimbriae are 1–2 mm long and 7 nm thick fibers, that are rod like, made up largely of repeating immunoglobulin-like FimA subunits (m.w. 17 kDa) arranged helically in a structure referred to as 'shaft' (**Figure 16**).[48,49]

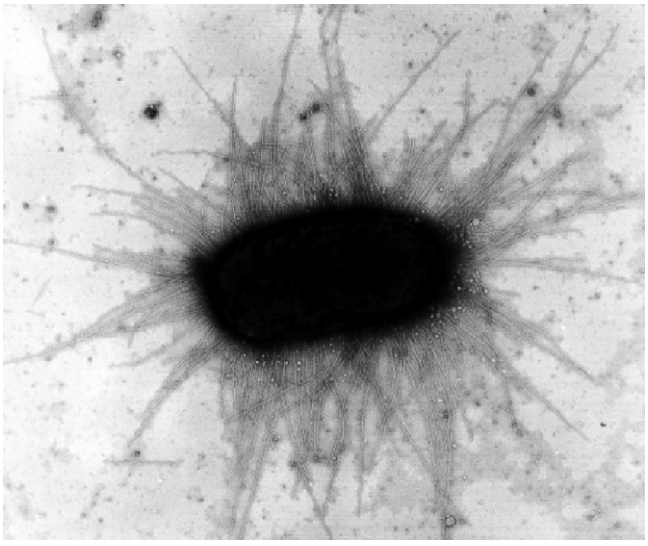

Figure 14 Type 1 fimbriated *E. coli.* Magnification ×20000. Courtesy Dr. David Hasty.

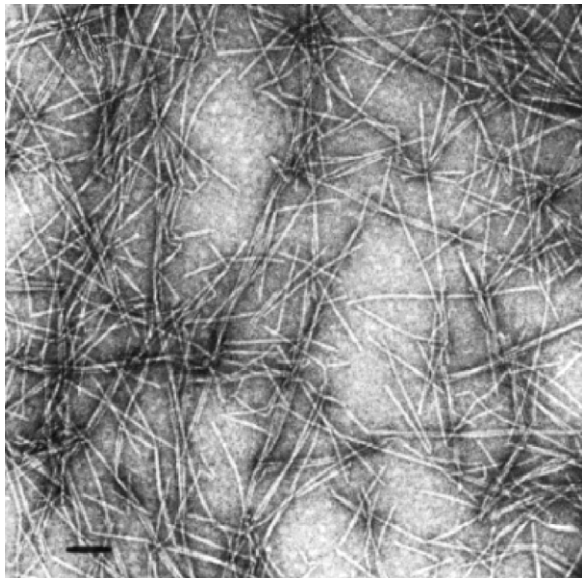

Figure 15 Electron micrograph of isolated type 1 fimbriae. Reproduced from Eshdat, Y.; Silverblatt, F. J.; Sharon, N. *J. Bacteriol.* **1981**, *148*, 308, with permission from American Society for Microbiology.

The shaft is joined to a short 3 nm thick distal tip fibrillum that consists of two adapter proteins, FimF and FimG, and a third of a different kind, FimH (m.w. 29–31 kDa). The latter is the only subunit that possesses a carbohydrate-binding site and is thus responsible for the sugar-binding activity of the fimbriae.[50] FimH is also present in small numbers at intervals along the fimbrial filament, but only the subunit at the tip appears to be able to mediate mannose-specific adhesive interactions, whereas the subunits at the other positions are inaccessible to the carbohydrate ligand.[51] Minor sequence variations in FimH alleles from different clinical isolates have been shown to correlate with altered carbohydrate-binding profiles of the fimbriae.[44]

The FimH subunits of *E. coli* and *K. pneumoniae* are 88% homologous. Still, the two organisms differ in their fine specificity, for example, in their relative affinity for Manα3Manβ4GlcNAc and *p*-nitrophenyl α-mannoside.[52–54] Other aromatic α-mannosides are also high-affinity ligands (up to 1000 times stronger than MeαMan) for *E. coli* type 1, suggesting the presence of a hydrophobic binding region next to the monosaccharide-combining site of FimH. With

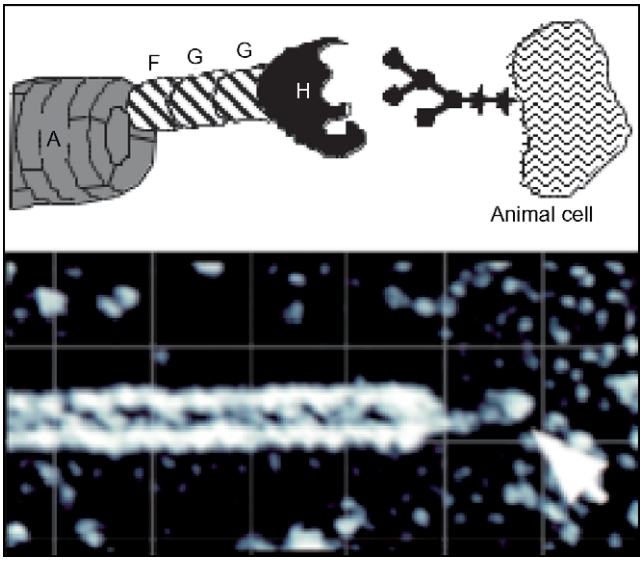

Figure 16 Structure of type 1 fimbriae. Top, schematic representation, depicting the distal end of the shaft as well as the tip of a fimbrium (or pilus) and a typical attachment site on host cells; A, F, G, and H denote the fimbrial subunits. Bottom, electron micrograph, in which arrowhead indicates the fimbrial tip containing the FimH subunit. Reproduced from Schilling, J. D.; Mulvey, M. A.; Hultgren, S. J. *J. Infect. Dis.* **2001**, *183*(*1*), S36–S40.

several *Salmonella* species examined, aromatic α-mannosides, as well as the trisaccharide Manα2Manβ4GlcNAc, were weaker inhibitors than MeαMan. The combining site of *Salmonella* species appears thus to be smaller than that of *E. coli* and *K. pneumoniae*, and to be devoid of an adjoining hydrophobic region. Different combining sites were found in other mannose-specific bacterial lectins. Therefore, although classified together on the basis of their monosaccharide (primary) specificity, these lectins differ in their fine specificity. Swapping experiments with genetically engineered hybrid fimbriae, in which the FimH of one species (e.g., *E. coli*) was presented on the shaft of the other species (*K. pneumoniae*), have shown that the shaft plays a role in modulating the specificity of fimbriae, probably by imposing conformational constraints on the carbohydrate-binding subunit.[55]

The first 3-D structure of a fimbrial carbohydrate-binding subunit to be solved was that of FimH in complex with the FimC chaperone (see below) to which the mannose analog cyclohexylbutanoyl-*N*-hydroxyethyl-D-glucamide was bound.[56] FimH is seen folded into two all-β class domains connected by a short extended linker (**Figure 17**). One of these, located in the N-terminal half of the subunit (residues 1–156), is the lectin domain, with the mannose-binding site at its tip. The C-terminal half (residues 160–279), known as the pilin domain, serves to anchor the subunit to the fimbriae. It binds in the cleft of FimC, although there is only limited contact between FimH and the C-terminal domain of the chaperone. The carbohydrate recognition domain of FimH is an 11-stranded elongated β-barrel with a jellyroll-like topology, while the pilin domain has an immunoglobulin fold that lacks the seventh (C-terminal) β-strand present in the canonical immunoglobulin fold. In the FimC–FimH complex, the missing β-strand is donated by the seventh strand of the N-terminal domain of the chaperone to complete the immunoglobulin-like fold of FimH. This kind of 'donor strand complementation' is thought to initiate folding of FimH directly on the chaperone, thus accounting for the function of the latter in biogenesis of the fimbriae.[57]

The 3-D structure of the FimC–FimH in complex with bound mannose has also been elucidated by X-ray crystallography.[58] Although mannose exists in solution as a mixture of α- and β-anomers, only the former was found in the crystal (**Figure 18**). It is buried at a deep and negatively charged site at the edge of FimH, opposite to the region through which the latter combines with the chaperone. All the mannose hydroxyls, except the anomeric one, interacted extensively with combining site residues, almost all of which are situated at the ends of β-strands or in the loops extending from them. Part of the hydrogen-bonding network is identical to that found in mannose complexes of other lectins. Thus, Asp54 makes cooperative hydrogen bonds with the 4- and 6-OH and Asn140 with the 3- and 4-OH of the ligand, similar to those made by the equivalent residues with the same sugar in the combining sites of legume lectins with the same primary specificity.[59] In addition, the N-terminal amine of the FimH polypeptide is H-bonded to the 2-OH, 6-OH, and the ring oxygen of the mannose.[58] This is one of the rare cases in which an N-terminal amine of a protein participates in ligand binding. Phe142 of FimH interacts hydrophobically

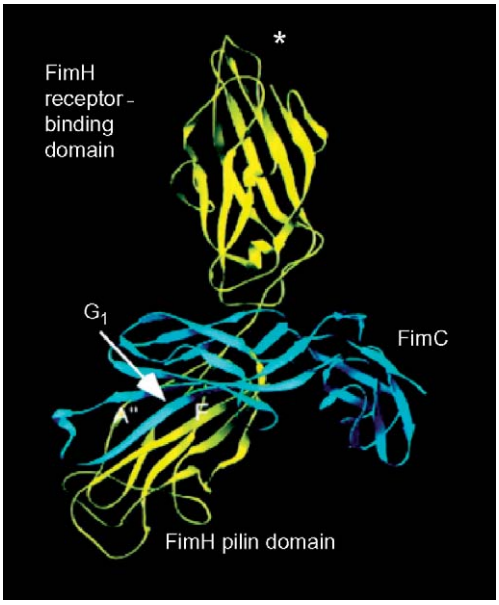

Figure 17 Crystal structure of FimH, the carbohydrate-binding subunit of type 1 fimbriae, in complex with the FimC chaperone. The asterisk indicates the mannose-binding pocket of FimH. Also shown is the insertion of the G1 strand of the FimC chaperone into the hydrophobic groove formed between the A and F strands of the FimH pilin domain, an interaction known as donor-strand complementation. Reproduced from Schilling, J. D.; Mulvey, M. A.; Hultgren, S. J. *J. Infect. Dis.* **2001**, *183*(*1*), S36–S40.

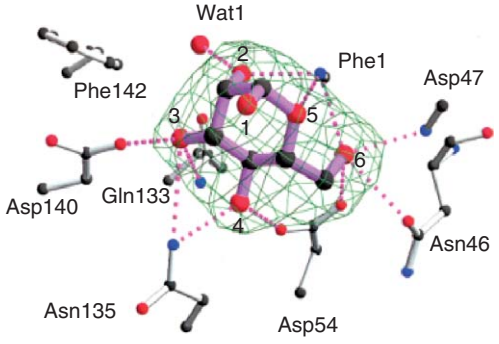

Figure 18 The α-anomer of mannose in the combining site of FimH. The eight marked residues combine with the mannose by hydrogen bonds and hydrophobic interactions. Contact residues are shown as ball-and-stick model. Wat stands for water. From Hung, C. S.; Bouckaert, J.; Hung, D.; Pinkner, J.; Widberg, C.; DeFusco, A.; Auguste, C. G.; Strouse, R.; Langermann, S.; Waksman, G.; *et al*. *Mol. Microbiol.* **2002**, *44*, 903–915.[59]

with the C2–C3 bond of the mannose. The same residue, together with Ile13, Tyr 48, and Ile52, form part of a hydrophobic ridge that surrounds the site and which may help to direct the ligand into it. Site-directed mutagenesis showed that combining site residues Asp54, Gln133, Asn135, and Asp140 are essential for carbohydrate binding by FimH, since their replacement by alanine, asparagine, or aspartic acid resulted in complete loss of this activity.

Examination of the binding site region of FimH has provided confirmation of the suggestion, made some two decades ago, that the combining site of this lectin is extended, and fits best mannose-containing trisaccharides such as Manα3Man*β*4GlcNAc or Manα6(Manα3)Man.[45] X-ray crystallography and modeling studies of the complexes of FimH with hydrophobic mannosides[60] have furnished a molecular explanation for the high affinity of type 1 fimbriae to such compounds (**Figure 19**). It is likely that the hydrophobic character of the ridge of the mannose-binding site of FimH distinguishes *E. coli* type 1 fimbriae from those of *Salmonella* species that neither exhibit an increased affinity for mannose with hydrophobic substituents nor for the oligomannosides mentioned earlier.

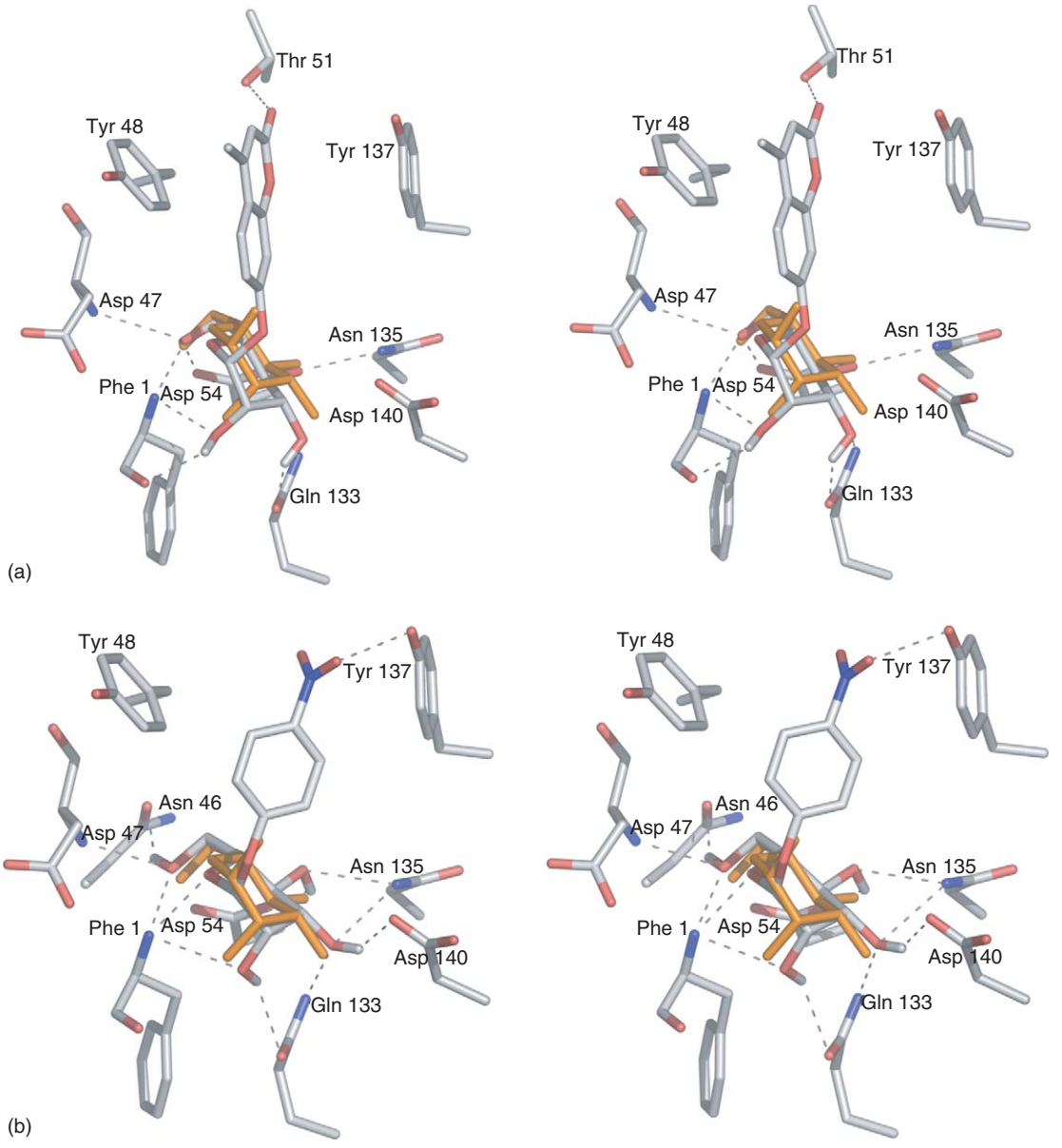

Figure 19 Stereodiagrams of MeUmbαMan (a) and *p*NPheαMan (b) docked in the FimH-binding site. The crystallographically determined position of the α-anomer of mannose is shown in orange for comparison. From Bouckaert, J.; Berglund, J.; Schembri, M.; De Genst, E.; Cools, L.; Wuhrer, M.; Hung, C. S.; Pinkner, J.; Slattegard, R.; Zavialov, A.; *et al*. *Mol. Microbiol.* **2005**, *55*, 441–455.

Nearly all mutations in the combining site of FimH abolished or decreased its binding not only to mannose, but also to urinary epithelial cells, indicating that the site may be highly conserved.[44,61] Support for this conclusion comes from the finding that there are very few variations in the sequences of the mannose-binding site of over 200 uropathogenic strains of *E. coli* examined, in contrast to enterohemorrhagic strains of *E. coli*, in which there are sequence variations at this site.[61] On the other hand, replacement of residues 185–279 within the FimH pilin domain with a corresponding segment of the type 1C fimbrial subunit FocH has led to a loss of the multivalent mannotriose-specific binding property accompanied by the acquisition of a distinct mannose-specific (i.e., monovalent, M1H) binding capability. Bacteria expressing the monovalent hybrid FimH were capable of binding strongly to uroepithelial tissue culture cells and guinea pig erythrocytes. They could not, however, agglutinate yeasts or bind human buccal cells, functions readily accomplished by the *E. coli* expressing mannotriose-specific FimH variants. Based on the relative potency of

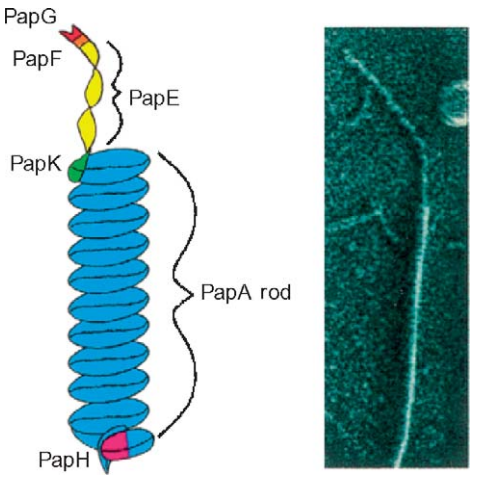

Figure 20 Structure of P fimbriae (pili). On the left, a schematic diagram of a single P fimbrium, showing the location of each subunit within the structure; on the right, electronmicrograph showing the two subassemblies of a P fimbrium. Reproduced from Dodson, K. W.; Pinkner, J. S.; Rose, T.; Magnusson, G.; Hultgren, S. J.; Waksman, G. *Cell* **2001**, *105*, 733–743.

inhibitory compounds of different structures, it was concluded that the receptor-binding site within the monovalent FimH–FocH subunit has an extended structure with an overall configuration similar to that within the multivalent FimH of natural origin.

Another class of high-affinity ligand for type 1 fimbriated *E. coli* are mannose-derived neoglycoproteins and dendrimers (see Chapter **3.23**).[61a] The former are proteins to which varying numbers of α-mannose residues are covalently attached, while the latter are multifunctional spherical branched polymers of well-defined molecular size that carry large numbers of such residues on their surface.

3.28.3.1.2 P fimbriae

In contrast to type 1 *E. coli*, that recognize structures present only in glycoproteins, P-fimbriated *E. coli* are specific for galabiose (Galα4Gal), a structure found in glycoproteins (see Chapter **3.14**) as well as in membrane glycosphingolipids of the globo-series.[62–64] These bacteria bind the disaccharide when it is present either at the nonreducing position or at an internal one of such glycolipids. They adhere mainly to the upper part of the kidney, where galabiose is more abundant. P fimbriae are similar to type 1 fimbriae in that they too are composite structures consisting of a long, rigid rod and a short, flexible, open helical part (**Figure 20**). The rigid section is about 7 nm in diameter and is composed mainly of about 1200–2400 copies of the PapA subunit (16–22 kDa) arranged in a tightly packed right-handed helix. The short flexible tip, consisting of PapE monomers, is 2 nm in diameter and is joined to the rigid rod by PapK adapter subunit.

The carbohydrate-binding subunit, PapG, is located exclusively at the N-terminal domain of the flexible tip and appears to be the sole determinant of binding specificity. It has mostly a β-sheet structure that can be subdivided into two regions (**Figure 21**). One is in the form of a β-barrel similar to the corresponding region of FimH; the other, with a structure that has not been encountered elsewhere, contains the carbohydrate-binding site. Like in type 1 fimbriae, donor strand complementation between PapG and PapK has been observed in the crystal structure of the chaperone-subunit complex of the P fimbriae. As with type 1 fimbria, allelic variations in PapG are found with distinct receptor specificity.

The combining site of PapG is extended and accommodates galabiose, the primary specificity determinant, binding the disaccharide by an array of hydrogen bonds, some mediated by water molecules, and hydrophobic interactions, with space for additional monosaccharide residues at its nonreducing end (**Figure 22**).

3.28.3.1.3 F17 fimbriae

These fimbriae, produced by enterotoxigenic *E. coli*, are 3 nm wide, flexible and wire-like organelles, built up of the major pilin subunit F17-A and exposing the F17-G subunit at their tip.[65] They mediate binding of the bacteria to *N*-acetylglucosamine-presenting receptors on the microvilli of the intestinal epithelium of ruminants, leading to diarrhea or septicemia. Binding of F17 *E. coli* to the microvilli is inhibited by *N*-acetylglucosamine as well as its

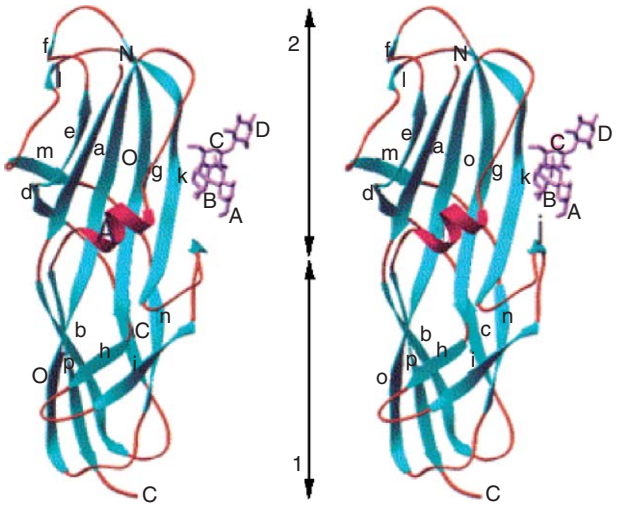

Figure 21 Ribbon presentation of the N-terminal region of PapG. The upper part, made up of seven β-strands, forms a β-barrel. The lower part is composed of a six-stranded central antiparallel β-sheet, flanked on one side by two double-stranded β-sheets and, on the other side, by an α-helix and a large loop connecting this helix to the central β-sheet. The upper part contains the carbohydrate-binding site, shown as a stick model of the tetrasaccharide ligand GbO4 (GalNAcβ3Galα4Galβ4Glc), starting with A at the nonreducing end. N- and C-mark the respective ends of the polypeptide. Regions 1 and 2 of the structure are indicated by the vertical black arrows and labeled 1 and 2, respectively. Reproduced from Dodson, K. W.; Pinkner, J. S.; Rose, T.; Magnusson, G.; Hultgren, S. J.; Waksman, G. *Cell* **2001**, *105*, 733–743.

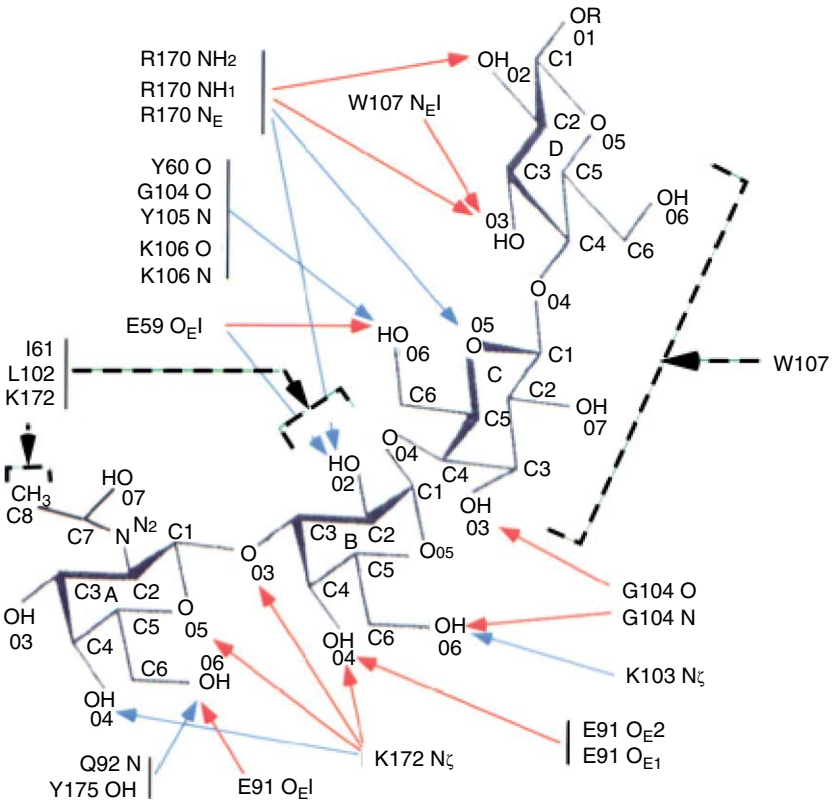

Figure 22 Schematic representation of interactions between PapG and its ligand. Polar interactions, either direct or water mediated, are shown by arrows. Brackets and arrows shown as dashed lines indicate contacts with aromatic/hydrophobic platforms. Reproduced from Dodson, K. W.; Pinkner, J. S.; Rose, T.; Magnusson, G.; Hultgren, S. J.; Waksman, G. *Cell* **2001**, *105*, 733–743.

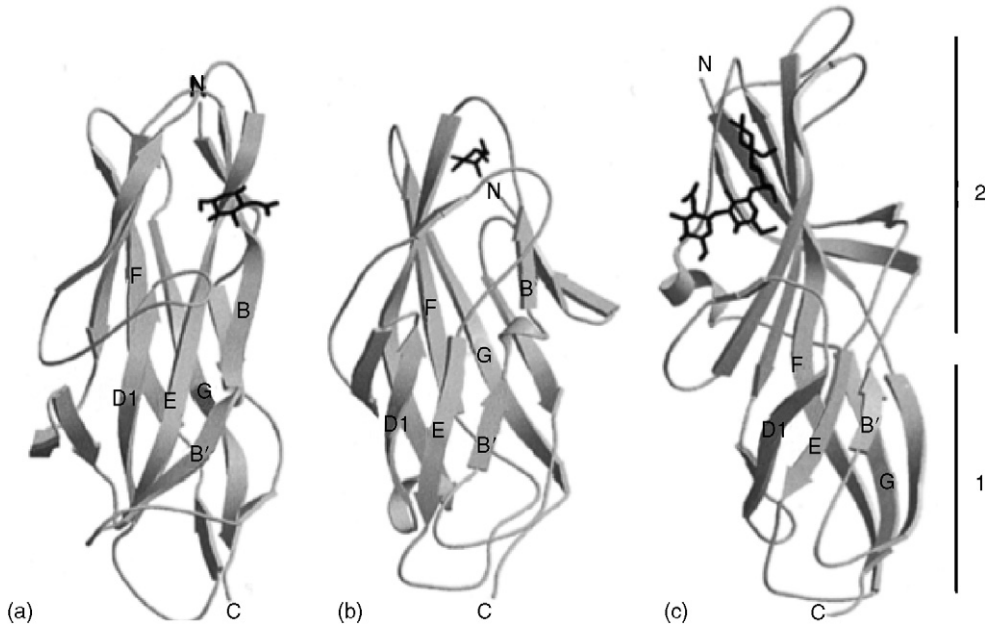

Figure 23 Localization of the sugar-binding sites of F17a-G (a), FimH (b), and PapGII (c). The proteins (represented in gray) were superimposed, based on the structural core of the immunoglobulin fold, which was identified in the three carbohydrate-binding subunits. Representative carbohydrate ligands are shown in black. The C-termini of the lectin domains, which precede the linker to the pilin domain, coincide approximately. Structurally equivalent strands are labeled with their names as defined for F17-G. The two parts of the PapGII domain are indicated: part 1 has the immunoglobulin-like core, whereas part 2 holds the sugar-binding site. From Buts, L.; Bouckaert, J.; De Genst, E.; Loris, R.; Oscarson, S.; Lahmann, M.; Messens, J.; Brosens, E.; Wyns, L.; De Greve, H. *Mol. Microbiol.* **2003**, *49*, 705–715.

β1-4-linked oligomers. The appearance and molecular organization of F17 fimbriae are similar to those of type 1 and P fimbriae, being composed of a flexible tip fibrillum with an open helical structure connected to the end of a tightly wound helical rod made up of their major fimbrial subunit.

The high-resolution crystal structure of the lectin domain of F17-G, in complex with *N*-acetylglucosamine, revealed that the monosaccharide is bound on the side of the ellipsoid-shaped protein in a conserved site around which all natural variations of F17-G are clustered [65] (for a study of the lectin domain of a closely related strain, see Ref.: 66). Despite lack of any sequence identity, it was unexpectedly found that the lectin domain of F17-G is similar to that FimH and PapG, and that all three share the immunoglobulin-like fold of the structural components (pilins) of their fimbriae (**Figure 23**). Analogous to the other fimbrial subunits, F17-G is a two-domain protein in which a C-terminal pilin domain is linked with an N-terminal carbohydrate-specific lectin domain.

The binding site of the lectin domain of F17a-G, one of the several cloned variants of F17-G examined in complex with the *N*-acetylglucosamine, is formed by the carbonyl group of Ala43, the side chains of Asp88, Thr89, Trp109, Ser117, Thr118, Gln119, and the nitrogen of Gly120 (**Figure 24**). Interactions between the carbohydrate and the protein include 11 possible hydrogen bonds, of which four are mediated by water molecules, as well as the hydrophobic stacking of the Trp109 side chain against the C5 and C6 atoms of the sugar. The *N*-acetyl group of *N*-acetylglucosamine contributes significantly to the affinity of the ligand to F17a G, due to a good complementarity of van der Waals surfaces between this group and the side chains of Thr118 and Asn44, as well as the carbonyl group of Ala43.

3.28.3.2 Soluble Bacterial Lectins

Only small number of bacterial lectins that are soluble are known (**Table 3**). The 3-D structures of all these in complexes with ligands have been investigated by X-ray crystallography and in some cases also by NMR.

3.28.3.2.1 *Pseudomonas aeruginosa* lectins

Two soluble lectins, originally isolated from *Pseudomonas aeruginosa* by Nechama Gilboa-Garber in the 1970s,[67,68] have been well characterized.[7] The galactose-specific PA-IL is a tetramer of four identical subunits, each consisting of

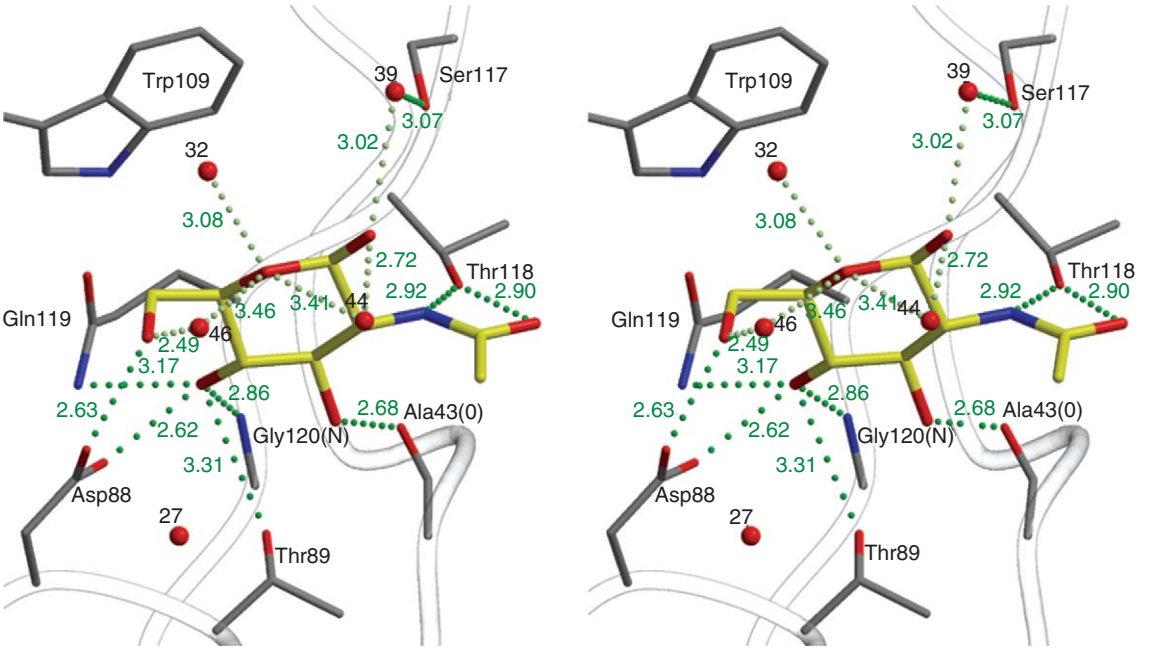

Figure 24 Stereo view of *N*-acetylglucosamine in the binding site of F17a-G. Hydrogen bonds are indicated by green dotted lines with heavy atom distances in Å. Water molecules are represented as small spheres. From Buts, L.; Bouckaert, J.; De Genst, E.; Loris, R.; Oscarson, S.; Lahmann, M.; Messens, J.; Brosens, E.; Wyns, L.; De Greve, H. *Mol. Microbiol.* **2003**, *49*, 705–715.

Table 3 Soluble bacterial lectins[a]

Source	*Name*	*Specificity*
Pseudomonas aeruginosa	PA-IL	Galactose
	PA-IIL	L-Fucose > mannose
Ralstonia solanacearum	RSL	L-Fucose
	RS-IIL	Mannose > L-fucose
Nostoc ellipsosporum	Cyanovirin N	M(9), M(8)
Microcystis viridis	MVL	

[a]For references, see text.

121 amino acids (m.w. 12.7 kDa). The subunit of the fucose- and mannose-specific PA-IIL, also a tetramer, is similar in size to that of PA-IL, but its primary structure is different.[69] The activities of both proteins are dependent upon the presence of divalent cations. The *Pseudomonas*-soluble lectins are found mainly in the bacterial cytoplasm. They are released by a subpopulation of the bacterial cells and serve to tether the bacteria to a substratum by binding to ill-defined structures on the bacterial surface and to a receptor on animal cells that contain the corresponding sugar.

PA-IL displays medium-range affinity for galactose, with a K_a of $3.4 \times 10^4 M^{-1}$.[70,71] Among monosaccharides, the specificity is strictly for galactose, with the exception of *N*-acetylgalactosamine, albeit with a lower affinity than that for galactose. The presence of a hydrophobic group on the anomeric position of the monosaccharide, either in α- or β-configuration, enhances the affinity, as observed with other lectins, such as *E. coli* type 1 fimbriae, with the tightest binding obtained for phenyl β-thiogalactoside. As for disaccharides, only those containing a terminal α-galactose residue are recognized, such as Galα3Gal, Galα4Gal, and Galα6Glc, to which the lectin displays the highest affinity; PA-IL binds also to a large number of glycoconjugates.[72,73] The latter include compounds containing terminal unsubstituted Galα4Gal disaccharide, that is, the human P1 and Pk blood group antigens, present on either red blood cell glycosphingolipids or on pigeon egg white glycoproteins, as well as Galα3Gal, the human B blood group antigen, and the xenoantigen present on pig tissues.

PA-IIL is characterized by an unusually high affinity for L-fucose, with a K_a of $1.6 \times 10^6 M^{-1}$ and binds the Le^a trisaccharide [Galβ3(Fucα4)GlcNAc] more strongly, with a K_a of $4.7 \times 10^6 M^{-1}$.[7,74] Thermodynamic studies indicate that these interactions are dominated by enthalpy. The entropy contribution is slightly favorable when binding to fucose and to the highest-affinity ligand Le^a. The high-resolution X-ray structures of two complexes of PA-IIL with Le^a and LNnFP-V (lacto-*N*-neofucopentaose-5, Galβ4GlcNAcβ4(Fucα3)Galβ4Glc), allowed the precise determination of the conformation of a trisaccharide and a pentasaccharide (see later, **Figure 28**). The different types of interaction between the oligosaccharides and the protein involve not only hydrogen bonding, but also calcium- and water-bridged contacts, allowing a rationalization of the thermodynamic data.

Three crystal structures of PA-IL have been reported thus far: the native lectin that contains one calcium ion per monomer, the calcium-free lectin, and the lectin with bound monosaccharide that contains both calcium and galactose (**Figure 25**).[75] In all cases, the quaternary structure is a tetramer, each of which monomers adopts a small jellyroll β-sandwich fold, consisting of two curved sheets, each one constructed from four antiparallel β strands.

The most striking feature of the PA-IIL structure is the involvement of the C-terminal carboxyl group (Gly114) of each monomer in the carbohydrate-binding site of the other monomer. Dimerization of the PA-II subunit is the result of a head-to-tail association of two monomers that make contact through the curved five-stranded β sheet. It largely involves hydrophobic contacts in the strands outside the Greek key motif of the protein and several hydrogen bonds at the extremity of these strands.[7,76] The buried surface is $\sim 1500 Å^2$ per monomer. The dimers form a tetramer mainly by the antiparallel association of β-strands comprised of amino acids 79–85 (strand 6) from each dimer with their counterparts in the other dimer. The surface buried by the association of the dimers into tetramers is $\sim 500 Å^2$ per monomer.

Both lectins are unusual in binding monosaccharides via direct interaction with metal ions.[7] The sugar-binding sites of PA-IL and PA-IIL contain one and two calcium ions, respectively. In the PA-IL-galactose complex, one sugar residue is located in each monomer, close to the calcium atom. Oxygen atoms O3 and O4 of galactose are directly involved in the coordination sphere of the calcium ion (**Figure 26**). The lectin–sugar interaction also involves six direct hydrogen bonds between the sugar hydroxyl groups and the protein and an additional one bridged by a water molecule. Some limited hydrophobic contacts are established between the apolar face of the galactose residue and certain hydrophobic amino acids.

In the PA-IIL-fucose complex, the fucose residue is locked onto a pair of calcium ions. Such a protein–carbohydrate binding mode involving two calcium ions has never been observed before (**Figure 27**). Three hydroxyl groups of fucose participate in the coordination of those calcium ions: O2 to the first calcium, O4 to the second calcium, and O3 to both of them. In addition to these electrostatic interactions, the three hydroxyl groups of the monosaccharide participate in hydrogen bonds with acidic groups of the calcium-binding site. The ring oxygen of fucose is involved in one hydrogen bond with the main chain, resulting altogether in a network of seven direct hydrogen bonds between the

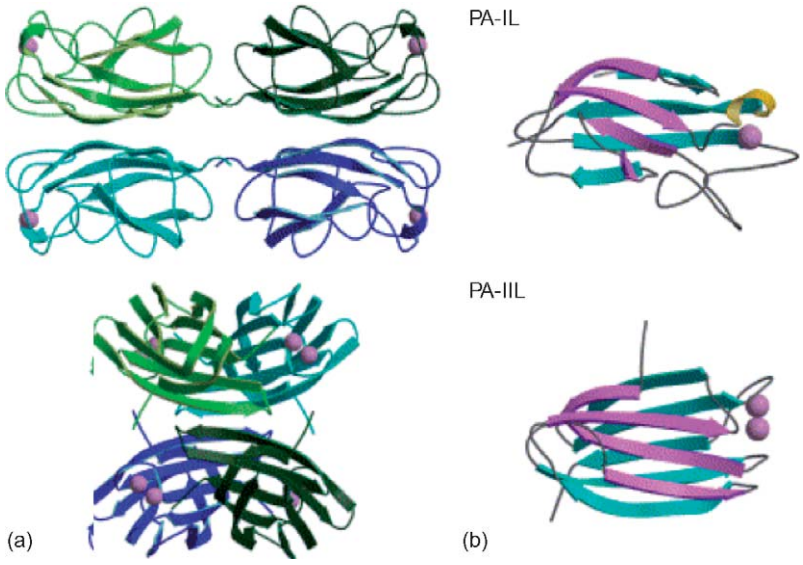

Figure 25 Crystal structures PA-IL and PA-IIL. a, Tetramers with stick representation of monosaccharides and space-filling representation of calcium ions. b, Monomers with the two β-sheets represented by different colors. From Imberty, A.; Wimmerova, M.; Mitchell, E. P.; Gilboa-Garber, N. *Microbes Infect.* **2004**, *6*, 221–228.

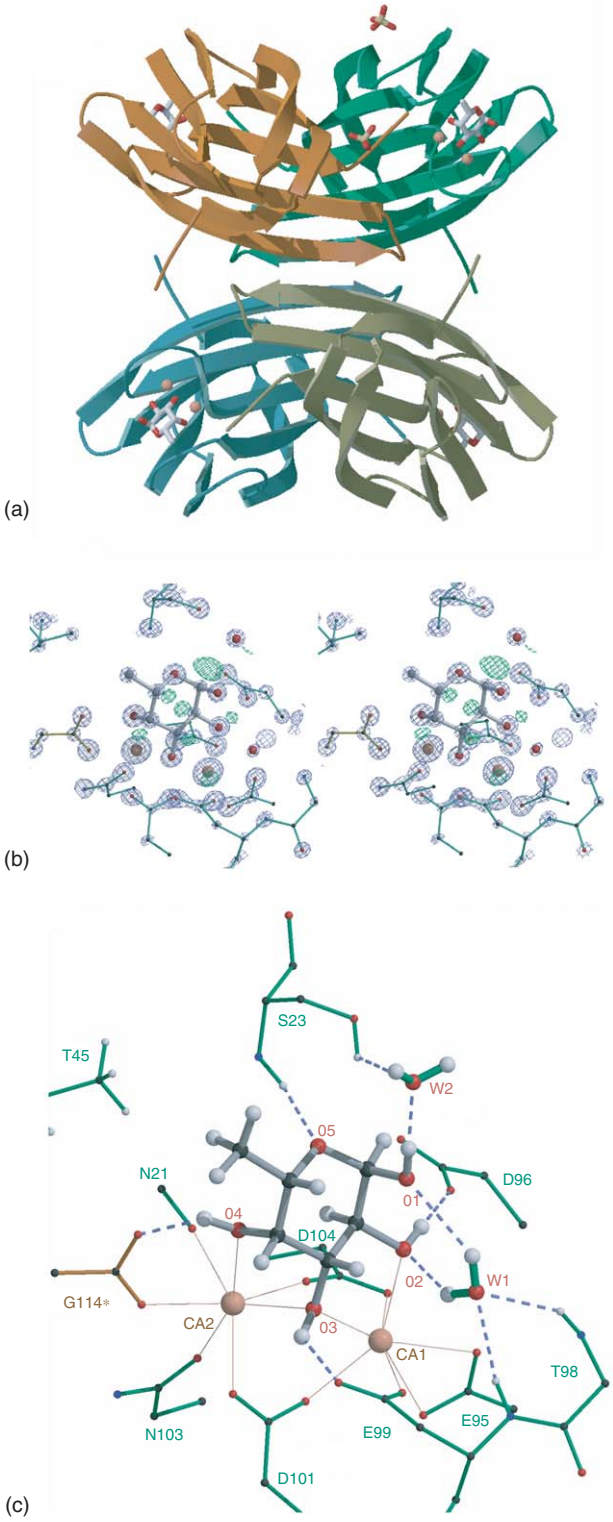

Figure 26 Structure of the PA-IIL-fucose complex. a, Ribbon representation of the tetramer consisting of the asymmetric unit with stick representation of sugar and sulfate and CPK representation of the calcium ions. b, Electron density map around the monosaccharide residue, water molecules, and calcium ions, omitting the hydrogen atoms riding the fucose carbons. c, Interactions of PA-IIL with calcium ions and fucose with modeled hydrogen atoms. Coordination contacts are indicated by orange solid lines and hydrogen bonds by blue dashed lines. The asterisk symbol indicates the terminal glycine from the other monomer. Reproduced from *Proteins* Mitchell, E. P.; Sabin, C.; Snajdrova, L.; Pokorna, M.; Perret, S.; Gautier, C.; Hofr, C.; Gilboa-Garber, N.; Koca, J.; Wimmerova, M.; Imberty, A.; Copyright © (2005, Wiley); Reprinted with permission of Wiley-Liss, Inc., a subsidary of John Wiley & Sons, Inc.

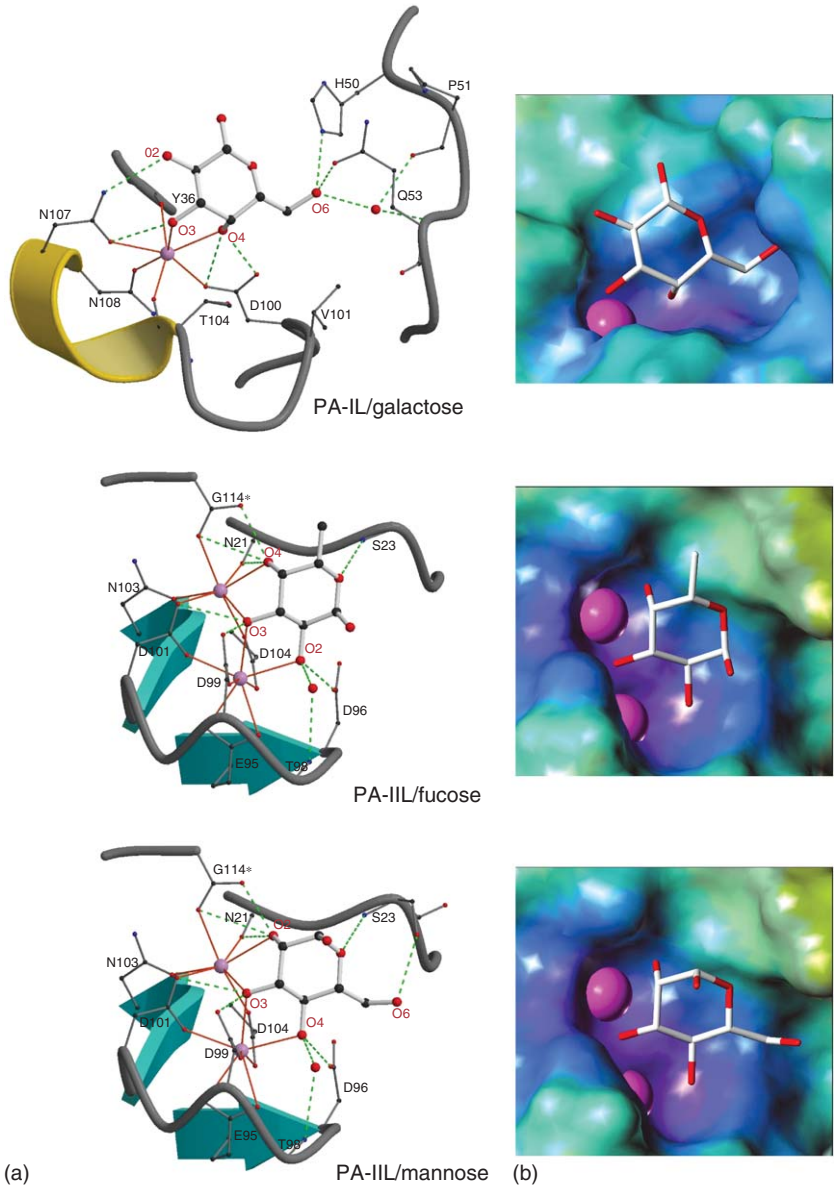

Figure 27 Binding sites in the crystal structures of the PA-IL-galactose, PA-IIL-fucose, and PA-IIL-mannose complexes. a, Stick representation of the amino acids involved in binding. Ca^{2+} coordination bonds are shown as solid orange lines; hydrogen bonds as dashed green lines. Color coding: red, oxygen; blue, nitrogen; black, carbon; pink, Ca^{2+}. b, Electrostatic surface representation (color coding from violet for negative to orange for positive) of the protein-binding site, with Ca^{2+} as large pink spheres and monosaccharides as stick models From Imberty, A.; Wimmerova, M.; Mitchell, E. P.; Gilboa-Garber, N. *Microbes Infect.* **2004**, *6*, 221–228.

sugar and the protein. A water molecule mediates an additional hydrogen bond. The methyl group at position 6 of fucose is located in a shallow pocket with hydrophobic character.

Analysis of the mode of binding of Le^a by PA-IIL revealed a large number of contacts between the sugar and the protein (Figure 28). All monosaccharide units of the trisaccharide interact with the protein surface either directly (hydrogen bonding and coordination of calcium) or indirectly through bridging water molecules. In addition to the nine direct hydrogen bonds between the fucose of Le^a and the protein described, the GlcNAc residue establishes a hydrogen bond between its O6 atom and the main-chain carbonyl group of Asp96. Two water molecules play an important bridging role in the interaction: one bridges O1 and O2 of fucose to the amide nitrogen of Thr98, whereas the other bridges the Gal residue (O2 and O1) to the side chain of Ser23. Hydrophobic contacts are very limited and

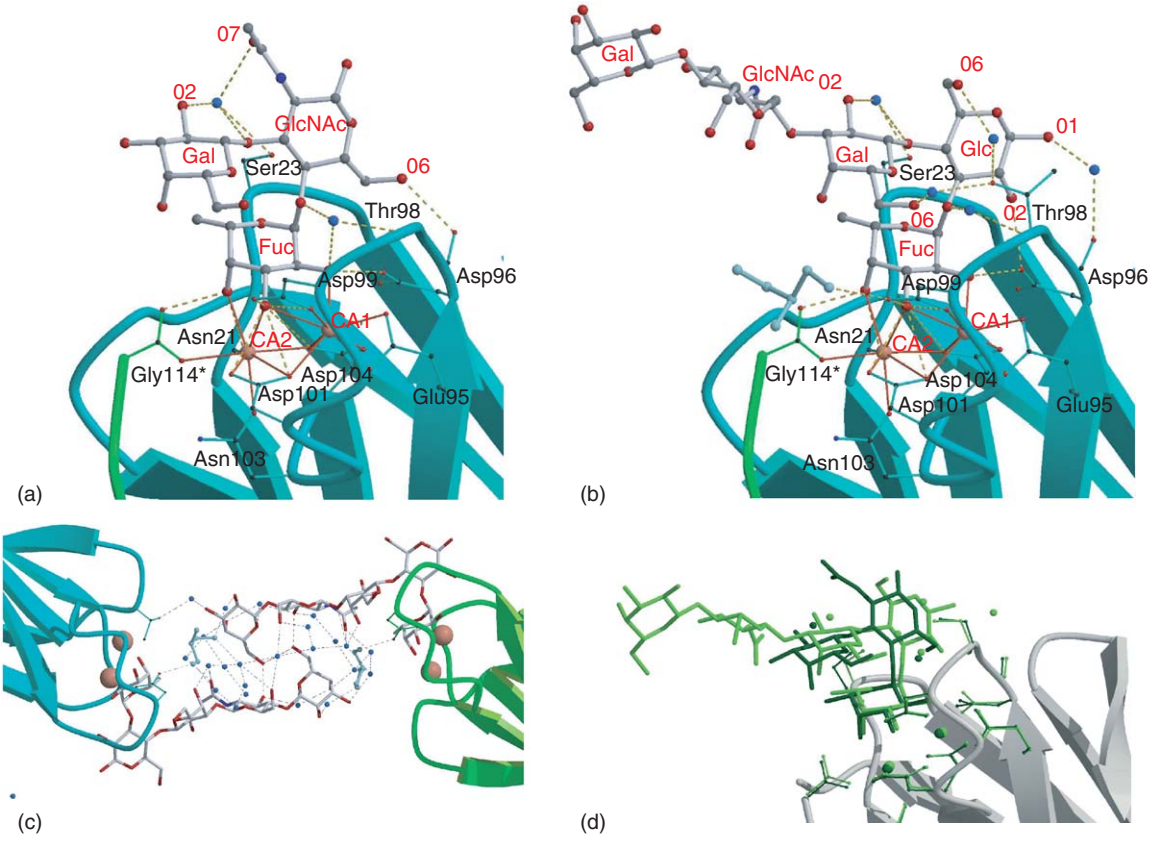

Figure 28 Complexes of PA-IIL with oligosaccharides. a, Interactions of PA-IIL with calcium ions and Le^a. Coordination contacts are indicated by orange solid lines and hydrogen bonds by green dashed lines. b, Same representation for the interaction of PA-IIL with calcium ions and LNnFP-V. c, Interaction between two pentasaccharides in the PA-IIL–LNnFP-V complex. Hydrogen bonds are represented by dotted lines. d, Superimposition of the binding sites of the PA-IIL–Le^a and PA-IIL–LNnFP-V complexes. From Perret, S.; Sabin, C.; Dumon, C.; Pokorna, M.; Gautier, C.; Galanina, O.; Ilia, S.; Bovin, N.; Nicaise, M.; Desmadril, M.; *et al*. *Biochem. J.* **2005**, *389*, 325–332.

only involve the methyl group of fucose with that of Thr145. The conformation adopted by Le^a in the binding site is similar to that predicted by those earlier studies, confirming the rigidity of the trisaccharide.

Among the four independent molecules of the PA-IIL–LNnFP-V complex, clear electron density can be observed for two complete pentasaccharides (Figure 28), whereas only a tetrasaccharide and a trisaccharide can be located for the two remaining sites. The fact that two of the oligosaccharides can be completely visualized is due the rigidity of their linear part resulting from fortuitous crystal packing that generates extensive contacts between two adjacent sugars centered on a pseudo twofold axis of symmetry (Figure 28c). It is this interaction that allows the very high-resolution structure of milk pentasaccharide to be described here for the first time. The protein-binding site, together with the calcium ions and fucose moiety, are identical with those described previously. Additional interactions with the protein surface are established by the glucose and galactose residues of the Le^x glucose analog moiety (Figure 28b). The glucose establishes a direct hydrogen bond between its hydroxyl group at position 2 and the side chain of Asp-96. In addition, five interactions are mediated by a water molecule, involving O1 of fucose, O1 and O6 of glucose, and O1, O2, and O6 of galactose. Hydrophobic interactions are observed for the methyl group of fucose, but also for C6 of galactose, which interacts with the carbon of a glycerol molecule located at the protein surface. The other part of the pentasaccharide, that is, the Glcβ4GlcNAc linked to position 3 of galactose, does not interact with the protein surface, but instead is involved in extensive interactions with the neighboring carbohydrate in the crystal lattice (Figure 28c).

PA-IL and PA-IIL are the first calcium-dependent lectins of bacterial origin to be structurally characterized. They do not seem to be related evolutionarily, even though they both adopt a jellyroll fold and they both use calcium for carbohydrate binding. In PA-IL, calcium, together with a unique network of hydrogen bonds, generates a binding site

endowed with a selective specificity for galactose. In contrast, in PA-IIL, the two calcium ions are associated with a very high affinity (micromolar range) and a wide specificity for binding different monosaccharides. The occurrence of metal ions in lectins is well known. Carbohydrate binding by legume lectins depends on the presence of manganese and calcium ions, whose main role is to maintain the architecture of the binding site and not to make contacts with the sugar ligand. Direct calcium–sugar interactions have been observed in only two families of animal lectins: pentraxins and C-type lectins. The first family includes serum amyloid P component (SAP) and C-reactive protein (CRP), which contain two close calcium ions involved in binding of diverse ligands. The lectin activity of SAP differs from that of PA-IIL, since it is only directed to sulfated or phosphorylated carbohydrates. Despite the lack of sequence or fold similarity, the galactose-binding mode of PA-IL shows a striking resemblance to those of two C-type animal lectins complexed with galactose, that is, the QPDWG mutant of rat mannose-binding protein A and the tunicate galactose-specific C-type lectin, TC14, from *Polyandrocarpa misakiens*.[7,78] Nevertheless, in PA-IL, the occurrence of an additional strong interaction with the galactose O6 group can explain the 10-fold higher galactose affinity compared to that of the C-type animal lectins.

3.28.3.2.2 *Ralstonia solanacearum* lectins

Two soluble bacterial lectins have been recently isolated from the plant pathogen *Ralstonia solanacearum*, each with different affinity to fucose. One of these is RSL (subunit M(r) 9.9 kDa), related to fungal lectins,[79] and the other is RS-IIL (subunit M(r) 11.6kDa), a tetrameric lectin, with high sequence similarity to PA-IIL.[80]

RSL consists of 90 amino acids with a tandem repeat in its amino acid sequence. It binds fucose with K_a of $5 \times 10^5 M^{-1}$. Among the monosaccharides tested, only L-galactose and D-arabinose, differing from fucose just by the substituent at C6, displayed some binding, albeit with affinity 4–5 times lower than that of fucose. RSL shows a preference for binding to Fucα2Gal and Fucα6Gal epitopes. It possesses two binding sites per monomer and an unusually high affinity for Fucα2Gal-containing oligosaccharides (K_a of $4 \times 10^6 M^{-1}$ for 2-fucosyllactose). The lectin has been crystallized with MeαFuc and with the highest-affinity ligand 2-fucosyllactose. The X-ray crystal structure of the RSL-MeαFuc complex revealed that each monomer consists of two small four-stranded antiparallel sheets (**Figure 29**). Trimerization through a threefold or pseudo-threefold axis generates a six-bladed β-propeller architecture, very similar to that previously described for the fungal lectin of *Aleuria aurantia*. Each monomer presents two fucose-binding sites, resulting in six symmetrically arranged sugar-binding sites for the β-propeller.

The hydrogen-bonding network between fucose and protein consists of six bonds. The fucose hydroxyl O2 is bridged to the main chain by the amide group of Ala40a in site 1 (Ala85 in site 2). Hydroxyl O3 receives hydrogen from NE1 of Trp81 (Trp36 of neighboring monomer in site 2) and donates to the carboxyl group of Glu28 (Glu73 in site 2). The other acidic oxygen of this carboxyl receives a hydrogen bond from O4 of fucose that is also hydrogen-bonded to NE of Arg17 (Arg62 in site 2). Finally, the ring oxygen O5 receives a hydrogen bond from the terminal NH_2 of the same arginine. In the high-resolution structure of RSL complexed with MeαFuc, water molecules are observed bound to the most accessible oxygens of fucose, O1, and O2, but they are not present in all sites and do not bridge to the protein. Hydrophobic contacts play an important role in the binding of fucose; Trp76 (Trp31 of the neighboring monomer in site 2) stacks to the fucose hydrophobic face (C3, C5, and C6) with distances smaller than 4 Å. On the other face of fucose, the sulfur atom of Cys30 (Cys75 in site 2) establishes a hydrophobic contact with the methine carbon C2. The methyl group at C6 is inserted in a hydrophobic pocket made by isoleucine residues 59 and 61 (monomer Pro14 and Ile16 in site 2) and tryptophan residues 10 and 81 (Trp53 of the same monomer and Trp16 of the neighboring one in site 2).

The same fucose-binding mode is observed in all of the complexes studied, with a very limited variation in the fucose orientation. The only difference between site 1 and site 2 is the Ile59/Pro14 substitution in the hydrophobic pocket. In all binding sites, the hydrogen bond network is fully conserved.

In the RSL-FucLac complex (intramonomer (site 1) and intermonomer (site 2)), the trisaccharide adopts very similar conformations and establishes the same contacts with the protein. The interactions involving the fucose residue have been described above. The galactose does not interact with the protein surface but makes an interresidue Gal.O2-Fuc.O3 hydrogen bond with the fucose residue. The glucose residue is folded back on the protein surface with its hydrophobic face, made by the CH of C1, C3, and C5, establishing van der Waals interactions with the plane surface created by a side chain of Asp32 that creates a salt bridge with Arg17 (Asp77 and Arg62 in site 2). In site 2, there is an additional stabilization by O6 interacting with the side chain of Trp53 through a bridging water molecule.

RS-IIL recognizes fucose but displays much higher affinity for mannose and fructose, which is opposite to the preference of PA-IIL. The binding loop of the cations is fully conserved in RS-IIL as in PA-IIL, whereas the preference for mannose versus fucose can be attributed to the change of a three-amino-acid sequence in the 'specificity loop'. The RS-IIL crystal contains one monomer of 113 amino acids per asymmetric unit, together with two Ca^{2+} ions

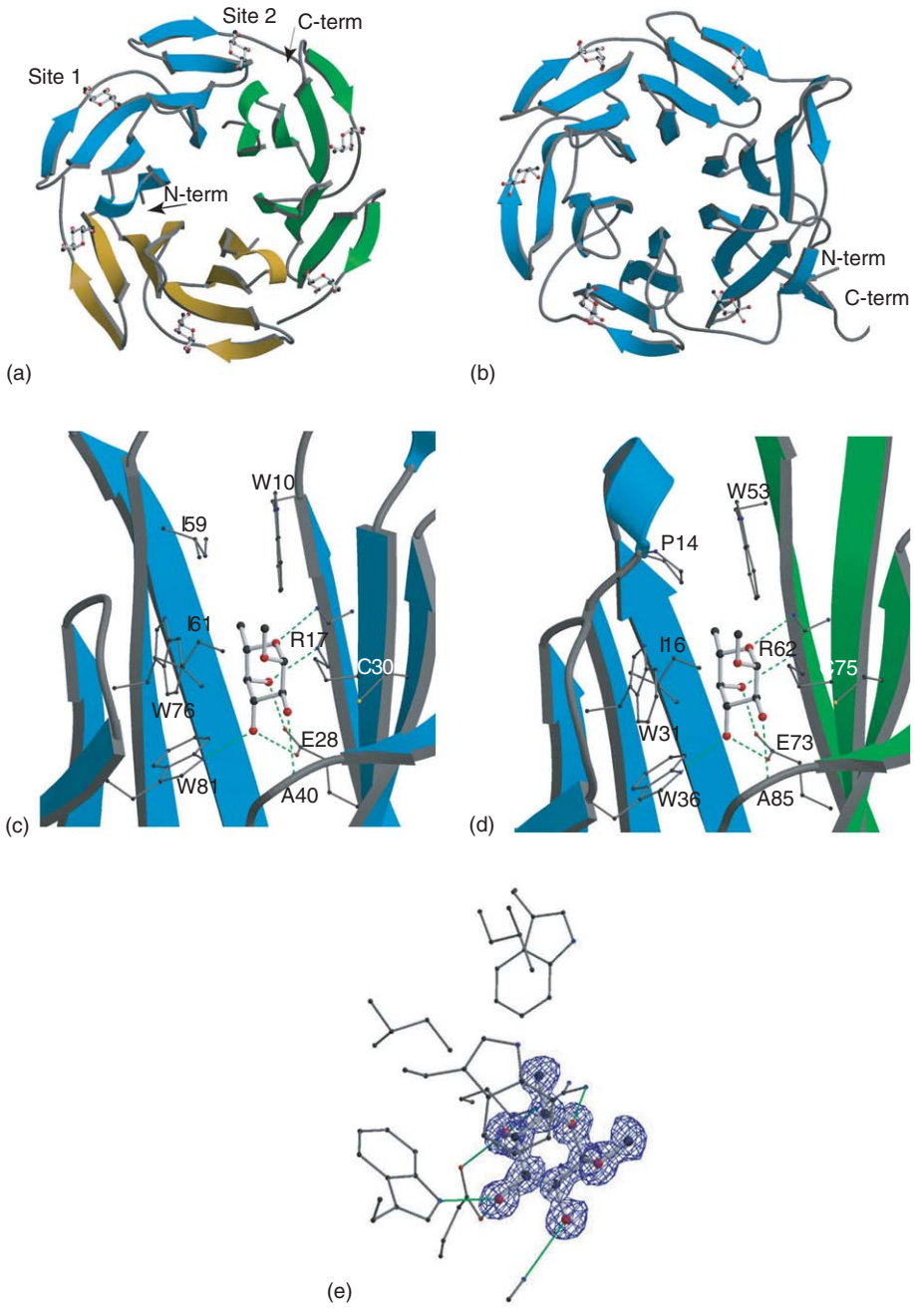

Figure 29 RSL trimer complexed with fucose. a, Ribbon diagram of the complex with fucose (shown as sticks). b, Same representation for *Aleuria aurantia* agglutinin (AAL) monomer complexed with fucose. c, Fucose in binding site 1 with hydrogen bonds represented by dashed lines. d, Same representation for binding site 2. e, Final electron density map around the fucose molecule. From Kostlanova, N.; Mitchell, E. P.; Lortat-Jacob, H.; Oscarson, S.; Lahmann, M.; Gilboa-Garber, N.; Chambat, G.; Wimmerova, M.; Imberty, A. *J. Biol. Chem.* **2005**, *280*, 27839–27849.

and one MeαMan molecule (Figure 30). The overall topology of RS-IIL is that of a nine-stranded antiparallel β-sandwich consisting of two sheets of four and five strands respectively. In the tetrameric quaternary structure, one interface between two monomers is created by strong complementary packing of the one hydrophobic curved β-sheet to its counterpart, and the other interface is created via sheet extension (Figure 30). Both monomeric and tetrameric structures are very similar to those described in detail for several crystal structures of PA-IIL.[77,78] The binding site is occupied by two Ca^{2+} ions and a mannose residue. The two ions are located close together, 3.75 Å, and

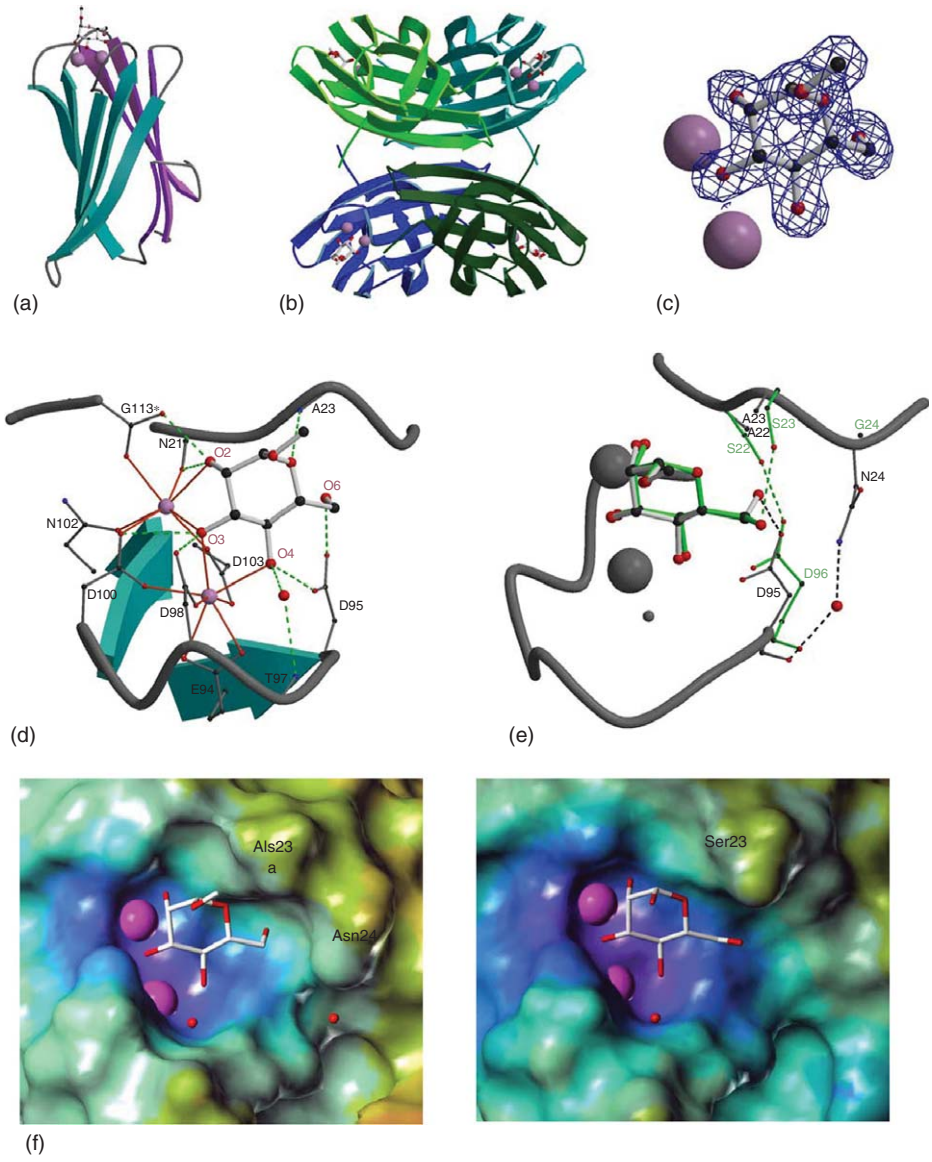

Figure 30 a, Crystal structure of the RSL-II-MeαMan complex with stick representation of the calcium ions. b, Representation of the molecular tetramer. c, Final electron density map around the monosaccharide residue. d, Calcium and MeαMan binding site with hydrogen bonds represented as dashed green lines and coordination contacts as orange solid lines. Calcium ions and water are represented by violet spheres respectively. e, Superimposition of MeαMan in RS-L binding site (grey and black lines and labels) and αMan in PA-IIL binding site (green lines and labels) from structure 1OVS. f, Electrostatic surface representation (color coding from violet for negative to orange for positive) of the protein binding site with Ca^{2+} as large pink spheres and monosaccharides as stick models with RS-IIL on the left and PA-II-MeαMan (1OVS) on the right. From Sudakevitz, D.; Kostlanova, N.; Blatman-Jan, G.; Mitchell, E. P.; Lerrer, B.; Wimmerova, M.; Katcoff, D. J.; Imberty, A.; Gilboa-Garber, N. *Mol. Microbiol.* **2004**, *52*, 691–700.

both display a classical seven-ligand coordination, mainly involving one loop, the conserved 'calcium-binding loop', via the side chains of Glu94, Asp98, Asp100, Asn102, and Asp103. A second loop is involved through the side chain of Asn21. In addition, the acidic group of the C-terminus Gly113 of the other monomer, which is involved in the head-to-tail interaction, also participates in the coordination of the Ca^{2+} ions. All amino acids of the calcium-binding site are strictly conserved between PA-IIL and RS-IIL, and the Ca^{2+} ions are bound in a similar manner.

In the RSLII-MeαMan complex (Figure 30), O2, O3, and O4 participate directly in the coordination of the Ca^{2+} ions, with a special position for O3, which participates in the coordination sphere of both. The O3, O4, and O6

hydroxyl groups of mannose are involved in hydrogen bonds with acidic amino acids of the calcium-binding loop extending from Glu94 to Asp103. The O2 hydroxyl group establishes hydrogen bonds with the backbone oxygen of Asn21 as well as with the C-terminus from the adjacent monomer. The ring oxygen O5 receives a hydrogen bond from the backbone nitrogen of Ala23. One water molecule makes an additional interaction by bridging between O4 and the backbone nitrogen of Thr97. Hydrophobic contacts appear to be very limited: only the methyl group of the MeαMan ligand is in relatively close position to Ala23. The comparison with the binding mode of mannose in PA-IIL shows similar contacts except in the region of O6.

3.28.3.2.3 Cyanovirin-N

Cyanovirin-N (CVN) is an 11 kDa lectin from the cyanobacterium (blue-green alga) *Nostoc ellipsosporum* with potent virucidal activity, that possesses the ability to inactivate all strains of human immunodeficiency virus (HIV) and simian immunodeficiency virus (SIV) as well as other viruses such as those of influenza and Ebola.[81,82] In addition, it binds very strongly the mammalian oligomannosides Man(9) and Man(8) (**Figure 31**), its minimal target being Manα2Manα, the termini of the branches of Man(8) and Man(9). The lectin consists of 101 amino acid residues organized in two domains, with a high degree of internal sequence and structural similarity. The solution structure solved by NMR revealed that CVN is an elongated, largely β-sheet protein that displays internal twofold pseudo-symmetry[83,84] (**Figure 32**). The two sequence repeats (residues 1–50 and 51–101) share 32% sequence identity and are superimposable;

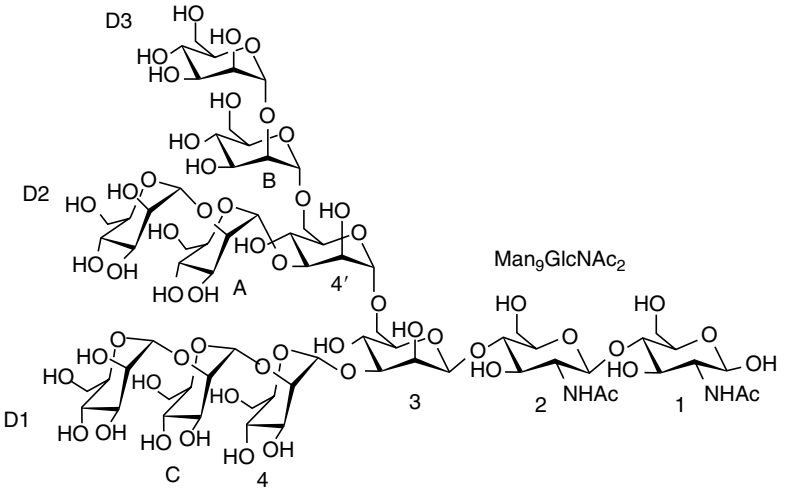

Figure 31 Structure of the N-linked undecasaccharide $Man_9GlcNAc_2$ (Man(9)) where the outer mannose residues are labeled as D1, D2, and D3. In the corresponding decasaccharide (Man(8)), the D2 arm is missing. From Bewley, C. A.; Kiyonaka, S.; Hamachi, I. *J. Mol. Biol.* **2002**, *322*, 881–889.

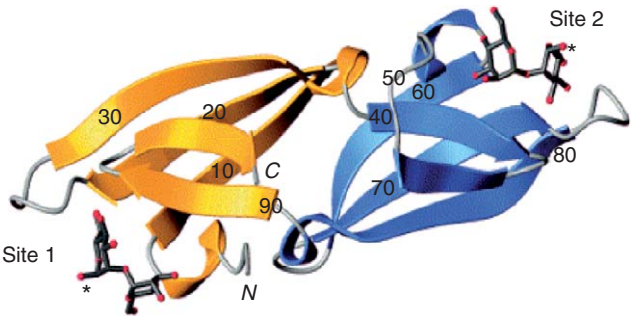

Figure 32 Ribbon drawing of the solution structure of a 1:2 CVN:Manα2Manα complex (PDB coordinates 1IIY) The opposing carbohydrate-binding sites are colored yellow and blue, and labeled as site 1 and site 2, respectively. The disaccharide Manα2Manα is shown in gray bonds with oxygen atoms colored red, and the reducing hydroxyl group is indicated by an asterisk. From Bewley, C. A.; Kiyonaka, S.; Hamachi, I. *J. Mol. Biol.* **2002**, *322*, 881–889.

they do not form, however, separate domains, since the overall fold is dependent on numerous contacts between them. Rather, two symmetrically related domains are formed by strand exchange between the two repeats.

Among the properties of CVN is its ability to block HIV envelope-mediated fusion. The affinity of oligosaccharide ligands to CVN was measured by their ability to inhibit the fusion-blocking activity of CVN in an HIV fusion assay. By this assay, it was found that CVN binds to gp120 with an equilibrium association constant K_a of $2.4\times10^7\,M^{-1}$ and an apparent stoichiometry of 2 equiv. of CVN per gp120; Man(8) acts as a divalent ligand (2 CVN:1 saccharide) with a K_a of $5.4\times10^7\,M^{-1}$, and Man(9) functions as a trivalent ligand (3 CVN:1 saccharide) with a K_a of $1.3\times10^8\,M^{-1}$. Isothermal titration calorimetry experiments of CVN binding to Man(9) confirmed the nanomolar affinity of K_a of $1.5\times10^8\,M^{-1}$, and the fitted data indicated a stoichiometry equal to approximately 1:1 CVN:Man(9).

The 1:1 stoichiometry at micromolar concentrations suggested that CVN has not only a high-affinity binding site relevant to the studies at nanomolar concentrations, but a lower-affinity site as well that facilitates cross-linking of CVN-oligomannose at micromolar concentrations or higher. The high-affinity site has a K_a of $7.2\times10^6\,M^{-1}$ and the low-affinity site a K_a of $6.8\times10^5\,M^{-1}$, as determined by isothermal titration calorimetry. The specificity studies indicated that the minimum structure required for high-affinity binding comprises the Manα2Manα that represents the terminal disaccharide of all three arms of Man(9). High-resolution NMR experiments demonstrate that CVN binds the above disaccharide via two distinct binding sites of differing affinities located on opposite ends of the protein.[83] Further NMR analysis of the solution structure of a complex of CVN-Manα2Manα at a molar ratio 1:2 revealed that the lectin recognizes the stacked conformation of the disaccharide through a deep hydrophilic-binding pocket on one side of the protein (site 2) and a semicircular cleft on the other (site 1)[85] (**Figure 33**). With the prominent exception of the C1 hydroxyl group of the reducing mannopyranose ring, the bound disaccharide is positioned so that each hydroxyl group is involved in a direct or water-mediated hydrogen bond to the polar or charged side chains comprising the binding pocket. The approximately 40 Å spacing of the two binding sites provides a simple model for CVN:gp120 binding. The CVN: Manα2Manα complex provides the first high-resolution structure of a mannose-specific protein–carbohydrate complex with nanomolar affinity and presents a new carbohydrate-binding motif, as well as a new class of carbohydrate-binding protein, that facilitates divalent binding via a monomeric protein.

Site 1 of CVN is able to discriminate between the three related trisaccharides Meα(Manα2Manα2Man), Meα-(Manα2Manα3Man), and Meα(Manα2Manα6Man) with remarkable selectivity, and binds these trisaccharides with K_a values ranging from $8.1\times10^3\,M^{-1}$ to $6.6\times10^6\,M^{-1}$. Site 2 is less selective in that it binds all three trisaccharides with similar K_a values ranging from 1.7 to $3.7\times10^5\,M^{-1}$, but overall binds these trimannosides with higher affinities than site 1.

3.28.3.2.4 *Microsystis viridis* lectin

Recently, another cyanobacterial protein known as MVL, originally isolated from a laboratory culture of *Microcystis viridis* NIES-102.[86] MVL has also been shown to inhibit HIV-1 fusion at nanomolar concentrations.[87] Like cyanovirin, MVL is unusual in that it binds oligosaccharides with very high affinity in the absence of multivalent interactions.

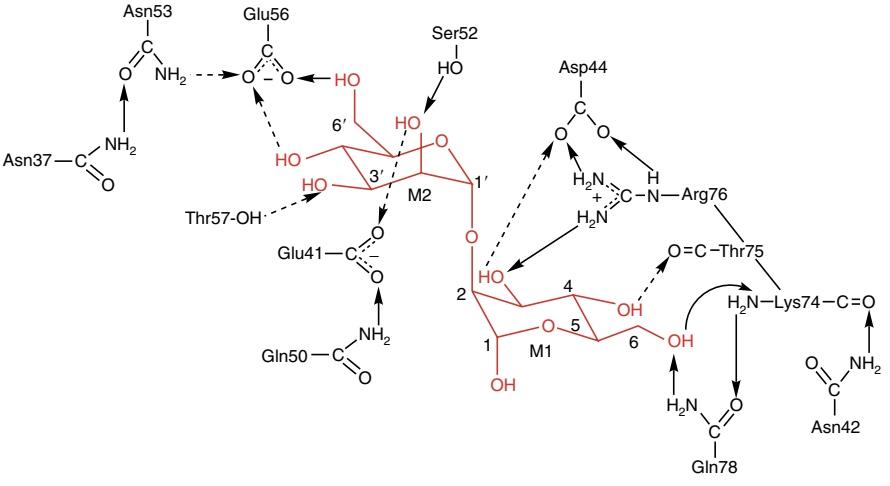

Figure 33 Schematic showing the potential hydrogen-bonding network between mannose hydroxyls, interfacial side chains, and proximal polar residues (for example, Asn37 and Gln50) of the high-affinity domain of CVN. Solid arrows indicate distances between donor and acceptor atoms that allow for direct hydrogen bonds, and dashed arrows denote distances that would accommodate water-mediated hydrogen bonds. From Bewley, C. A. *Structure* **2001**, *9*, 931–940.

However, it is is distinct among oligomannose-binding proteins in that it exhibits fine specificity for the Manα6Manβ4GlcNAcβ4GlcNAc tetrasaccharide core found in N-linked oligomannosides.

MVL in solution is a monodisperse homodimer with four independent carbohydrate-binding sites. Each subunit comprises two homologous domains, each consisting of 54 residues with 50% sequence identity between them, connected by a five-residue linker. MVL exhibits no significant sequence similarity to any other known protein family. X-ray crystallography of MVL free and bound to the pentasaccharide $Man_3GlcNAc_2$ confirmed that the protein is a homodimer stabilized by an extensive intermolecular interface between monomers[87,88] (Figure 34). Each

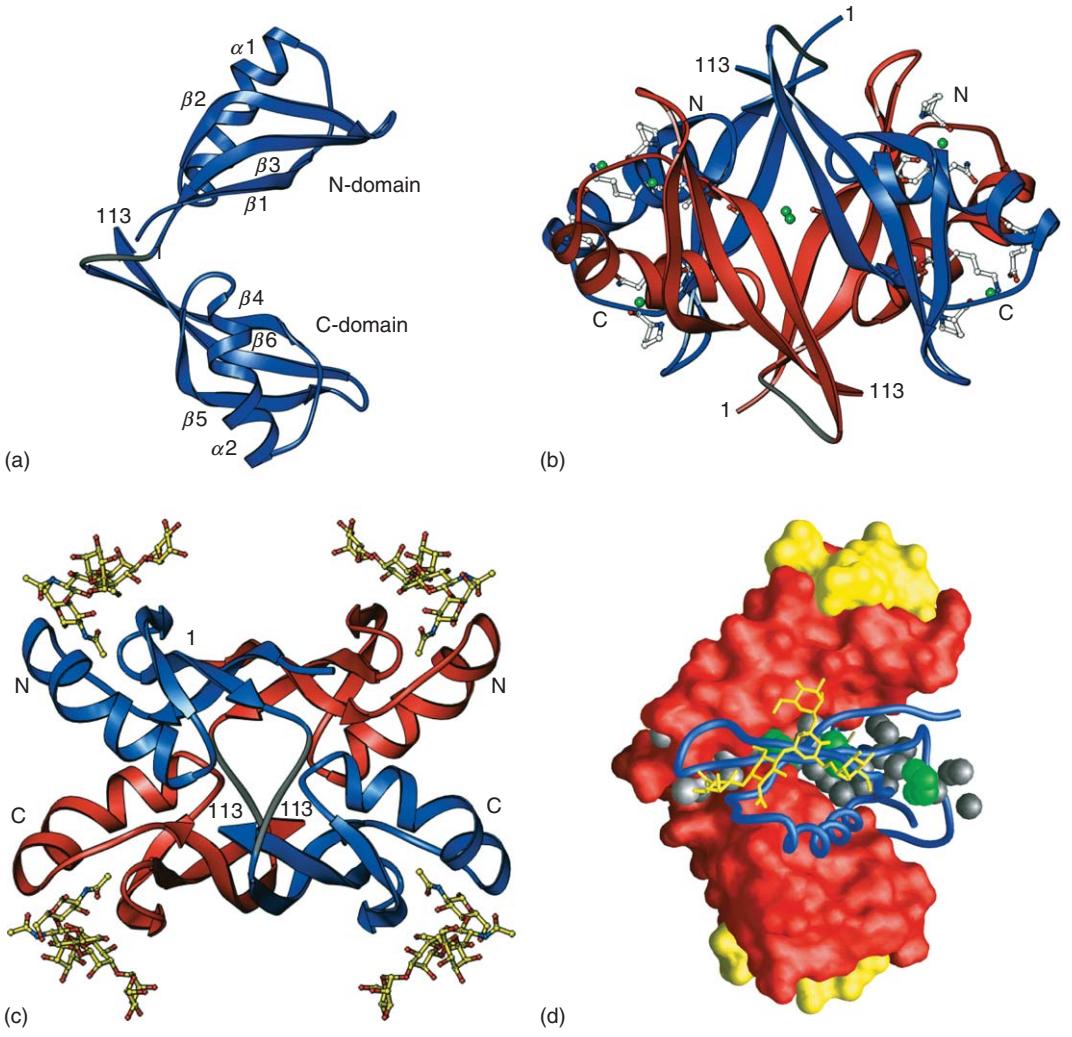

Figure 34 Overview of the crystal structures of MVL free and bound to $Man_3GlcNAc_2$. a, Ribbon diagram of a single monomer of MVL from the crystal structure of the free protein. In the N-domain, strands 1, 2, and 3 and helix 1 comprise residues 2–10, 32–41, 45–53, and 14–29, respectively; in the C-domain, strands 4, 5, and 6 and helix 2 comprise residues 61–69, 91–100, 104–112, and 73–88, respectively. b, Ribbon diagram of the free MVL dimer with one subunit in blue and the other in red, together with water molecules (green spheres) and associated side chains that participate in water-bridged hydrogen bonds between the two monomers. c, Ribbon diagram of the MVL-$Man_3GlcNAc_2$ complex with the bound carbohydrate depicted as cylinders and spheres (carbon, yellow; nitrogen, blue; and oxygen, red). d, Combined surface, tube, stick, and space-filling representation of the MVL dimer illustrating the water-filled channel formed between monomers in the dimer. A surface diagram of one monomer (red) and tube diagram of the second monomer (blue) are shown with space-filling rendering of water (gray) and ethylene glycol (green) molecules that fill the channel between the monomers. A surface diagram of carbohydrate bound to the surface model of one monomer of MVL and a stick diagram of carbohydrate bound to the tube model of the second monomer are shown in yellow. From Williams, D. C., Jr.; Lee, J. Y.; Cai, M.; Bewley, C. A.; Clore, G. M. *J. Biol. Chem.* **2005**, *280*, 29269–29276.

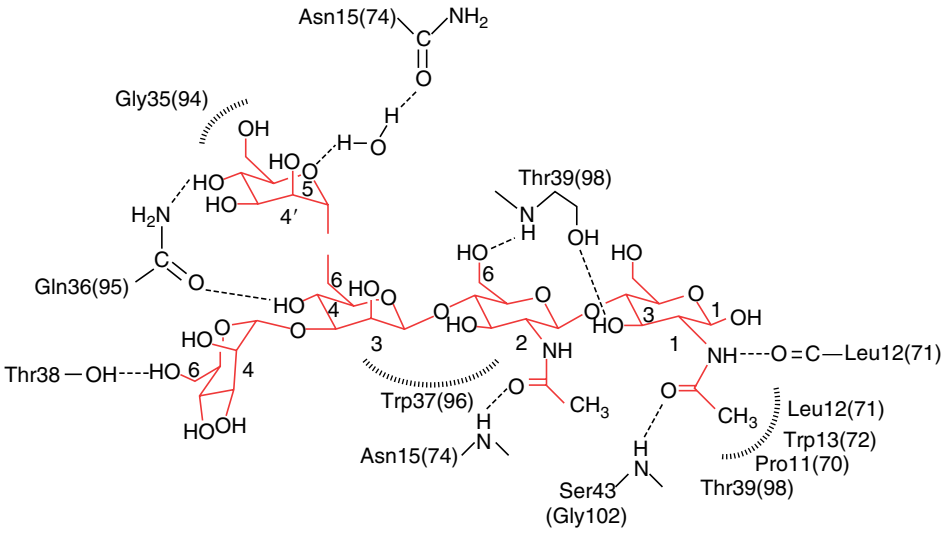

Figure 35 Schematic diagram of hydrogen-bonding interactions between MVL and $Man_3GlcNAc_2$. The carbohydrate structure is shown in red and the contacting protein residues in black. Dashed lines indicate hydrogen bonds, and the dashed arcs represent hydrophobic contacts. From Williams, D. C., Jr.; Lee, J. Y.; Cai, M.; Bewley, C. A.; Clore, G. M. *J. Biol. Chem.* **2005**, *280*, 29269–29276.

monomer contains two structurally homologous domains with high sequence similarity connected by a short five-amino-acid residue linker. Intriguingly, a water-filled channel is observed between the two monomers. $Man_3GlcNAc_2$ binds to a preformed cleft at the distal end of each domain such that a total of four independent carbohydrate molecules associate with each homodimer. The binding cleft provides shape complementarity, including the presence of a deep hydrophobic hole that accommodates the *N*-acetyl methyl at the reducing end of the carbohydrate, and specificity arises from seven to eight intermolecular hydrogen bonds, one water-mediated bond, and a number of hydrophobic interactions (**Figure 35**).

References

1. Ofek, I.; Hasty, D. L.; Doyle, R. J. *Bacterial Adhesion to Animal Cells and Tissues.* ASM Press: Washington, DC, 2003; 416pp.
2. Sharon, N.; Lis, H. *Lectins*, Kluwer/Springer: Dordrecht, The Netherlands, 2003; 454pp.
3. Ofek, I.; Hasty, D. L.; Sharon, N. *FEMS Immunol. Med. Microbiol.* **2003**, *38*, 181–191.
4. Sharon, N. *Biochim. Biophys. Acta.* **2006**, *1760*, 527–537.
5. Blixt, O.; Head, S.; Mondala, T.; Scanlan, C.; Huflejt, M. E.; Alvarez, R.; Bryan, M. C.; Fazio, F.; Calarese, D.; Stevens, J.; Razi, N.; Stevens, D. J.; Skehel, J. J.; van Die, I.; Burton, D. R.; Wilson, I. A.; Cummings, R.; Bovin, N.; Wong, C. H.; Paulson, J. C. *Proc. Natl. Acad. Sci. USA* **2004**, *101*, 17033–17038.
6. Ratner, D. M.; Adams, E. W.; Su, J.; O'Keefe, B. R.; Mrksich, M.; Seeberger, P. H. *Chembiochem* **2004**, *5*, 379–382.
7. Imberty, A.; Wimmerova, M.; Mitchell, E. P.; Gilboa-Garber, N. *Microbes. Infect.* **2004**, *6*, 221–228.
8. Lis, H.; Sharon, N. *Chem. Rev.* **1998**, *98*, 637–674.
9. Lee, Y. C.; Lee, R. T. *Acc. Chem. Res.* **1995**, *28*, 321–327.
10. Lemieux, R. U. *Acc. Chem. Res.* **1996**, *29*, 373–380.
11. Sharon, N. *Biochemist* **2006**, *28*, 13–17.
12. Chen, Y.; Maguire, T.; Hileman, R. E.; Fromm, J. R.; Esko, J. D.; Linhardt, R. J.; Marks, R. M. *Nat. Med.* **1997**, *3*, 866–871.
13. Spillmann, D. *Biochimie* **2001**, *83*, 811–817.
14. Lamb, R. A.; Kolakofsky, D. Paramyxoviridae: The viruses and their replication. In *Fields Virology*, 3rd ed.; Fields, B. N., Knippe, D.M, Howley, D. R., Eds.; Lippincott-Raven: Philadelphia, PA, 1996; pp 1177–1204.
15. Villar, E.; Barroso, I. M. *Glycoconj. J.* **2006**, *23*, 5–17.
16. de Groot, R. J. *Glycoconj. J.* **2006**, *23*, 59–72.
17. Schwegmann-Weisels, C; Herrler, G. *Glycoconj. J.* **2006**, *23*, 51–58.
18. Vives, R. R.; Imberty, A.; Sattentau, Q. J.; Lortat-Jacob, H. *J. Biol. Chem.* **2005**, *280*, 21353–21357.
19. Haidar, M.; Gluckman, J. C.; Gattegno, L. *Glycobiology* **1992**, *2*, 429–435.
20. Mbemba, E.; Gluckman, J. C.; Gattegno, L. *Glycobiology* **1994**, *4*, 13–21.
21. Rider, C. C. *Glycoconj. J.* **1997**, *14*, 639–642.
22. Wiley, D. C.; Skehel, J. J. *Annu. Rev. Biochem.* **1987**, *56*, 365–394.
23. Skehel, J. J.; Wiley, D. C. *Annu. Rev. Biochem.* **2000**, *69*, 531–569.
24. Russell, R. J.; Stevens, D. J.; Haire, L. F.; Gamblin, S. J.; Skehel, J. J. *Glycoconj. J.* **2006**, *23*, 85–92.

25. Ohuchi, M.; Ohuchi, R.; Feldmann, A.; Klenk, H. D. *J. Virol.* **1997**, *71*, 8377–8384.
26. Connor, R. J.; Kawaoka, Y.; Webster, R. G.; Paulson, J. C. *Virology* **1994**, *205*, 17–23.
27. Gamblin, S. J.; Haire, L. F.; Russell, R. J.; Stevens, D. J.; Xiao, B.; Ha, Y.; Vasisht, N.; Steinhauer, D. A.; Daniels, R. S.; Elliot, A.; Wiley, D. S.; Skehel, J. J. *Science* **2004**, *303*, 1838–1842.
28. Rosenthal, P. B.; Zhang, X.; Formanowski, F.; Fitz, W.; Wong, C. H.; Meier-Ewert, H.; Skehel, J. J.; Wiley, D. C. *Nature* **1998**, *396*, 92–96.
29. Dormitzer, P. R.; Sun, Z. Y.; Blixt, O.; Paulson, J. C.; Wagner, G.; Harrison, S. C. *J. Virol.* **2002**, *76*, 10512–10517.
30. Dormitzer, P. R.; Sun, Z. Y.; Wagner, G.; Harrison, S. C. *EMBO. J.* **2002**, *21*, 885–897.
31. Stehle, T.; Harrison, S. C. *EMBO. J.* **1997**, *16*, 5139–5148.
32. Freund, R.; Garcea, R. L.; Sahli, R.; Benjamin, T. L. *J. Virol.* **1991**, *65*, 350–355.
33. Stehle, T.; Yan, Y.; Benjamin, T. L.; Harrison, S. C. *Nature* **1994**, *369*, 160–163.
34. Stehle, T.; Harrison, S. C. *Structure* **1996**, *4*, 183–194.
35. Stehle, T.; Gamblin, S. J.; Yan, Y.; Harrison, S. C. *Structure* **1996**, *4*, 165–182.
36. Fry, E. E.; Lea, S. M.; Jackson, T.; Newman, J. W.; Ellard, F. M.; Blakemore, W. E.; Abu-Ghazaleh, R.; Samuel, A.; King, A. M.; Stuart, D. I. *EMBO. J.* **1999**, *18*, 543–554.
37. Mulloy, B.; Linhardt, R. J. *Curr. Opin. Struct. Biol.* **2001**, *11*, 623–628.
38. Fry, E. E.; Newman, J. W.; Curry, S.; Najjam, S.; Jackson, T.; Blakemore, W.; Lea, S. M.; Miller, L.; Burman, A.; King, A. M.; Stuart, D. I. *J. Gen. Virol.* **2005**, *86*, 1909–1920.
39. Nataro, J. P.; Cohen, P. S.; Mobley, H. L. T.; Weiser, J. N., Eds.; In *Colonization of Mucosal Surfaces*; ASM Press: Washington, DC, 2005; 456pp.
40. Gally, D. L.; Bogan, J. A.; Eisenstein, B. I.; Blomfield, I. C. *J. Bacteriol.* **1993**, *175*, 6186–6193.
41. Remaut, H.; Waksman, G. *Curr. Opin. Struct. Biol.* **2004**, *14*, 161–170.
42. Soto, G. E.; Hultgren, S. J. *J. Bacteriol.* **1999**, *181*, 1059–1071.
43. Sokurenko, E. V.; Chesnokova, V.; Dykhuizen, D. E.; Ofek, I.; Wu, X. R.; Krogfelt, K. A.; Struve, C.; Schembri, M. A.; Hasty, D. L. *Proc. Natl. Acad. Sci. USA* **1998**, *95*, 8922–8926.
44. Sokurenko, E. V.; Feldgarden, M.; Trintchina, E.; Weissman, S. J.; Avagyan, S.; Chattopadhyay, S.; Johnson, J. R.; Dykhuizen, D. E. *Mol. Biol. Evol.* **2004**, *21*, 1373–1383.
45. Sharon, N. *FEBS Lett.* **1987**, *217*, 145–157.
46. Zhou, G.; Mo, W. J.; Sebbel, P.; Min, G.; Neubert, T. A.; Glockshuber, R.; Wu, X. R.; Sun, T. T.; Kong, X. P. *J. Cell. Sci.* **2001**, *114*, 4095–4103.
47. Xie, B.; Zhou, G.; Chan, S. Y.; Shapiro, E.; Kong, X. P.; Wu, X. R.; Sun, T. T.; Costello, C. E. *J. Biol. Chem.* **2006**, *281*, 14644–14653.
48. Madison, B.; Ofek, I.; Clegg, S.; Abraham, S. N. *Infect. Immun.* **1994**, *62*, 843–848.
49. Ofek, I.; Hasty, D. L.; Abraham, S. N.; Sharon, N. *Adv. Exp. Med. Biol.* **2000**, *485*, 183–192.
50. Schembri, M. A.; Kjaergaard, K.; Sokurenko, E. V.; Klemm, P. *J. Infect. Dis.* **2001**, *183*(*Suppl 1*), S28–S31.
51. Abraham, S. N.; Hasty, D. L.; Simpson, W. A.; Beachey, E. H. *J. Exp. Med.* **1983**, *158*, 1114–1128.
52. Schilling, J. D.; Mulvey, M. A.; Hultgren, S. J. *J. Infect. Dis.* **2001**, *183*(*1*), S36–S40.
53. Firon, N.; Ofek, I.; Sharon, N. *Carbohydr. Res.* **1983**, *120*, 235–249.
54. Firon, N.; Ashbenazi, S.; Miralman, D.; Ofek, I.; Sharon, N. *Infect. Immun.* **1987**, *55*, 472–476.
55. Duncan, M. J.; Mann, E. L.; Cohen, M. S.; Ofek, I.; Sharon, N.; Abraham, S. N. *J. Biol. Chem.* **2005**, *280*, 37707–37716.
56. Choudhury, D.; Thompson, A.; Stojanoff, V.; Langermann, S.; Pinkner, J.; Hultgren, S. J.; Knight, S. D. *Science* **1999**, *285*, 1061–1066.
57. Normark, S. *Proc. Natl. Acad. Sci. USA* **2000**, *97*, 7670–7672.
58. Hung, C. S.; Bouckaert, J.; Hung, D.; Pinkner, J.; Widberg, C.; DeFusco, A.; Auguste, C. G.; Strouse, R.; Langermann, S.; Waksman, G.; Hultgren, S. J. *Mol. Microbiol.* **2002**, *44*, 903–915.
59. Sharon, N.; Lis, H. *Adv. Exp. Med. Biol.* **2001**, *491*, 1–16.
60. Bouckaert, J.; Berglund, J.; Schembri, M.; De Genst, E.; Cools, L.; Wuhrer, M.; Hung, C. S.; Pinkner, J.; Slattegard, R.; Zavialov, A.; Choudhury, D.; Langermann, S.; Hultgren, S. J.; Wyns, L.; Klemm, P.; Oscarson, S.; Knight, S. D.; De Greve, H. *Mol. Microbiol.* **2005**, *55*, 441–455.
61. Sokurenko, E. V.; Schembri, M. A.; Trintchina, E.; Kjaergaard, K.; Hasty, D. L.; Klemm, P. *Mol. Microbiol.* **2001**, *41*, 675–686.

61a. Pieters, R. J. *Trends Glycosci. Glycotech.* **2004**, *16*, 234–254.

62. Stromberg, N.; Nyholm, P. G.; Pascher, I.; Normark, S. *PNAS* **1991**, *88*, 9340–9344.
63. Larsson, A.; Ohlsson, J.; Dodson, K. W.; Hultgren, S. J.; Nilsson, U.; Kihlberg, J. *Bioorg. Med. Chem.* **2003**, *11*, 2255–2261.
64. Dodson, K. W.; Pinkner, J. S.; Rose, T.; Magnusson, G.; Hultgren, S. J.; Waksman, G. *Cell* **2001**, *105*, 733–743.
65. Buts, L.; Bouckaert, J.; De Genst, E.; Loris, R.; Oscarson, S.; Lahmann, M.; Messens, J.; Brosens, E.; Wyns, L.; De Greve, H. *Mol. Microbiol.* **2003**, *49*, 705–715.
66. Merckel, M. C.; Tanskanen, J.; Edelman, S.; Westerlund-Wikstrom, B.; Korhonen, T. K.; Goldman, A. *J. Mol. Biol.* **2003**, *331*, 897–905.
67. Gilboa-Garber, N.; Mizrahi, L.; Garber, N. *Can. J. Biochem.* **1977**, *55*, 975–981.
68. Gilboa-Garber, N. *Biochim. Biophys. Acta* **1972**, *273*, 165–173.
69. Gilboa-Garber, N.; Katcoff, D. J.; Garber, N. C. *FEMS Immunol. Med. Microbiol.* **2000**, *29*, 53–57.
70. Garber, N.; Guempel, U.; Belz, A.; Gilboa-Garber, N.; Doyle, R. J. *Biochim. Biophys. Acta* **1992**, *1116*, 331–333.
71. Stoitsova, S. R.; Boteva, R. N.; Doyle, R. J. *Biochim. Biophys. Acta* **2003**, *1619*, 213–219.
72. Chen, C. P.; Song, S. C.; Gilboa-Garber, N.; Chang, K. S.; Wu, A. M. *Glycobiology* **1998**, *8*, 7–16.
73. Lanne, B.; Ciopraga, J.; Bergstrom, J.; Motas, C.; Karlsson, K. A. *Glycoconj. J.* **1994**, *11*, 292–298.
74. Perret, S.; Sabin, C.; Dumon, C.; Pokorna, M.; Gautier, C.; Galanina, O.; Ilia, S.; Bovin, N.; Nicaise, M.; Desmadril, M.; Gilboa-Garber, N.; Wimmerova, M.; Mitchell, E. P.; Imberty, A. *Biochem. J.* **2005**, *389*, 325–332.
75. Cioci, G.; Mitchell, E. P.; Gautier, C.; Wimmerova, M.; Sudakevitz, D.; Perez, S.; Gilboa-Garber, N.; Imberty, A. *FEBS Lett.* **2003**, *555*, 297–301.
76. Loris, R.; Tielker, D.; Jaeger, K. E.; Wyns, L. *J. Mol. Biol.* **2003**, *331*, 861–870.
77. Mitchell, E.; Houles, C.; Sudakevitz, D.; Wimmerova, M.; Gautier, C.; Perez, S.; Wu, A. M.; Gilboa-Garber, N.; Imberty, A. *Nat. Struct. Biol.* **2002**, *9*, 918–921.
78. Poget, S. F.; Legge, G. B.; Proctor, M. R.; Butler, P. J.; Bycroft, M.; Williams, R. L. *J. Mol. Biol.* **1999**, *290*, 867–879.
79. Kostlanova, N.; Mitchell, E. P.; Lortat-Jacob, H.; Oscarson, S.; Lahmann, M.; Gilboa-Garber, N.; Chambat, G.; Wimmerova, M.; Imberty, A. *J. Biol. Chem.* **2005**, *280*, 27839–27849.
80. Sudakevitz, D.; Kostlanova, N.; Blatman-Jan, G.; Mitchell, E. P.; Lerrer, B.; Wimmerova, M.; Katcoff, D. J.; Imberty, A.; Gilboa-Garber, N. *Mol. Microbiol.* **2004**, *52*, 691–700.
81. Dey, B.; Lerner, D. L.; Lusso, P.; Boyd, M. R.; Elder, J. H.; Berger, E. A. *J. Virol.* **2000**, *74*, 4562–4569.

82. Barrientos, L. G.; O'Keefe, B. R.; Bray, M.; Sanchez, A.; Gronenborn, A. M.; Boyd, M. R. *Antiviral. Res.* **2003**, *58*, 47–56.
83. Bewley, C. A.; Gustafson, K. R.; Boyd, M. R.; Covell, D. G.; Bax, A.; Clore, G. M.; Gronenborn, A. M. *Nat. Struct. Biol.* **1998**, *5*, 571–578.
84. Bewley, C. A.; Kiyonaka, S.; Hamachi, I. *J. Mol. Biol.* **2002**, *322*, 881–889.
85. Bewley, C. A. *Structure* **2001**, *9*, 931–940.
86. Yamaguchi, M.; Ogawa, T.; Muramoto, K.; Kamio, Y.; Jimbo, M.; Kamiya, H. *Biochem. Biophys. Res. Commun.* **1999**, *265*, 703–708.
87. Bewley, C. A.; Cai, M.; Ray, S.; Ghirlando, R.; Yamaguchi, M.; Muramoto, K. *J. Mol. Biol.* **2004**, *339*, 901–914.
88. Williams, D. C., Jr.; Lee, J. Y.; Cai, M.; Bewley, C. A.; Clore, G. M. *J. Biol. Chem.* **2005**, *280*, 29269–29276.

Biographical Sketch

Nathan Sharon is Emeritus Professor at the Weizmann Institute of Science, which he joined in 1954, after receiving his Ph.D. Degree from the Hebrew University, Jerusalem. Between 1956 and 1959 he was postdoctoral fellow at the laboratories of Fritz Lipmann and Roger W. Jeanloz, both at the Massachusetts General Hospital, and of Dan Koshland at Brookhaven National Laboratory, and in subsequent years was, among others, visiting professor at University of California, Berkeley, and Harvard University, and Fogarty Scholar at the NIH. He has made seminal contributions to glycobiolology in his studies amino sugars and glycoproteins, of lectin-carbohydrate interactions and their role in cell recognition, and the application of the knowledge accrued to medical problems such as the treatment of microbial infections (together with I. Ofek) and bone marrow transplantation. In addition, he has been a tireless and highly effective advocate for glycobiology worldwide. Sharon's numerous publications include several widely cited reviews, as well as three books, 'Complex Carbohydrates' (1975), and 'Lectins', with H. Lis (1989 and 2006), which also appeared in Japanese. He is the recipient of many awards and honors, which include the Israel Prize in biochemical research, membership of the Israel Academy of Sciences and of the European Molecular Biology Organization, Honorary Doctorate from the University of Paris, and Honorary Membership of the American Society of Biochemistry and Molecular Biology and of the American Society of Microbiology.

Itzhak Ofek is Professor of Human Microbiology at Tel Aviv University, which he joined in 1979. After receiving his Ph.D. from the Hebrew University, Jerusalem, in 1972 he spent three years as postdoctoral fellow at the laboratories of Gene Stollerman and the late Edwin Beachey, both at the University of Tennessee, Department of Medicine Infectious Diseases Division, Memphis, TN. During 1975 and 1979 he was faculty member of The Departments of Microbiology at College of Medicine in Memphis, TN and Hadassah Medical School, in Jerusalem, Israel. He was visiting professor at the Department of Pathology, Washington University, St Louis with Erika Crouch and at the Research Center for Infectious Diseases, Wurzburg, Germany with Jorg Hacker. He has made extensive research of bacterial adhesion with special emphasis on the role of lectin–carbohydrate interactions (with N. Sharon) and on innate immunity against bacterial infections (with Yona Keissari, Hanny Sahly, and Erika Crouch). Ofek's publications include more than 150 original articles, several of which are highly cited, as well as many reviews and two books (coauthored with the late Ron Doyle and with David Hasty), one of which appeared in Italian as well. He also co-edited three books, two of which are on microbial adhesion and one on innate immunity. Ofek together with Sharon have been the main advocates for anti-adhesion therapy of infectious diseases with agents such as carbohydrates or natural food products.