

# Chapter 7

## Lectins as HIV Microbicides

Leonardus M.I. Koharudin and Angela M. Gronenborn

**Abstract** Carbohydrate-binding proteins constitute a growing class of antiviral agents. They prevent infection by blocking virus entry into the host target cells and also block virus transmission from virus-infected cells to non-infected cells. In order to illustrate the molecular basis of their antiviral activity towards human immunodeficiency virus (HIV), a comprehensive review of the three-dimensional structures of various antiviral lectins, as well as modes and atomic determinants of their high-affinity oligosaccharide recognition, is presented here. The collective information derived from these studies aids in the understanding of carbohydrate recognition of the gp120 envelope protein by these antiviral lectins and may lead to novel directions in the development of alternative drug-leads for the prevention of HIV transmission.

**Keywords** Lectins • Antivirals • Cyanovirin-N (CV-N) • *Oscillatoria agardhii* agglutinin (OAA) • Griffithsin (GRFT) • Scytovirin (SVN) • *Microcystis viridis* lectin (MVL) • Actinohivin (AH) • Banana lectin (BanLec) • High mannose glycan • X-ray crystallography • NMR spectroscopy

### 7.1 Introduction

Human immunodeficiency virus (HIV) is a lentivirus that causes acquired immunodeficiency syndrome (AIDS), a progressive breakdown of the human immune system that eventually leads to life-threatening infection or death. Over the last 30 years, extensive and comprehensive research efforts have been directed at understanding the HIV life cycle in the fight against AIDS. As a result, several effective

---

L.M.I. Koharudin • A.M. Gronenborn (✉)  
Department of Structural Biology, University of Pittsburgh School of Medicine,  
3501 Fifth Avenue, Biomedical Science Tower 3/Room 1050, Pittsburgh, PA 15260, USA  
e-mail: amg100@pitt.edu

therapies and drugs that target different stages of the viral life cycle have become available. However, there is still no cure for HIV/AIDS to date, although valuable treatment options exist. The latter generally involves a combination, or “cocktails,” of several classes of drugs, directed predominantly against the key HIV enzymes. Highly Active Antiretroviral Therapy (HAART) has proved highly successful and has drastically reduced both new HIV infections and mortality (Pirrone et al. 2011; NIAID National Institute of Allergy and Infectious Diseases—Antiretroviral Therapy to Reduce the Transmission of HIV). Unfortunately, despite these major advances, the AIDS pandemic continues to pose a significant public health concern, as several million people worldwide are still infected with the virus (WHO World Health Organization—Data and Statistics).

New initiatives for preventing sexual transmission of HIV have been promoted and the development of microbicides for topical or ex-vivo use is one possible avenue (Chirenje et al. 2010; D’Cruz and Uckun 2004; Hladik and Doncel 2010; Minces and McGowan 2010; Turpin 2002; NIAID National Institute of Allergy and Infectious Diseases—Topical Microbicides; WHO World Health Organization—Microbicides). This approach will be particularly useful for curbing the escalating rate of HIV infection in women, notably in those regions of the world where social and psychological barriers are substantial and difficult to overcome (Turpin 2002; Team 2010; Minces and McGowan 2010; D’Cruz and Uckun 2004; Chirenje et al. 2010). For example, owing to economic and societal pressures, diagnosis and treatment of HIV infections may not be readily available or are stigmatized. Therefore, the potential use of microbicides when applied topically to genital mucosal surfaces is potentially a powerful strategy to significantly reduce transmission of sexually transmitted viral pathogens, given that application in cream form is discreet and can be completely controlled by women. Several candidates for use as such barrier applications include substances that directly interact with HIV virions, thus preventing viral entry into and fusion with the target cells, such as Carraguard<sup>®</sup>, cellulose sulfate, or PRO 2000<sup>®</sup>, tenofovir gel, or substances that enhance the natural vaginal defense mechanisms by maintaining an acidic pH, thereby protecting the vagina, such as Acidform<sup>®</sup>, BufferGel<sup>®</sup>, or *Lactobacillus crispatus* (Team 2010; Turpin 2002; D’Cruz and Uckun 2004; Hladik and Doncel 2010; Minces and McGowan 2010; Morrow and Hendrix 2010). Due to disappointing clinical trial outcomes for some of these candidates, new candidates, such as vaginal rings loaded with davifirine, maraviroc, or a combination of both, have been developed and are now undergoing active testing in ongoing clinical trials (IPM, International Partnership of Microbicides).

Carbohydrate-binding proteins (lectins) represent a novel therapeutic class of substances for the development of microbicides (Ziolkowska and Wlodawer 2006; Botos and Wlodawer 2005; Barrientos and Gronenborn 2005; Balzarini 2006). They block infection by binding to the sugars that decorate the surface of the HIV envelope (Env) glycoprotein gp120, cross-linking the trimeric Env in the closed non-fusogenic structure and thus rendering the virus unable to enter the host target cell, as well as blocking direct cell-to-cell transmission between virus-infected and non-infected cells. Lectins can also efficiently abrogate DC-SIGN-directed HIV-1 capture and subsequent transmission to T lymphocytes (Balzarini et al. 2007).

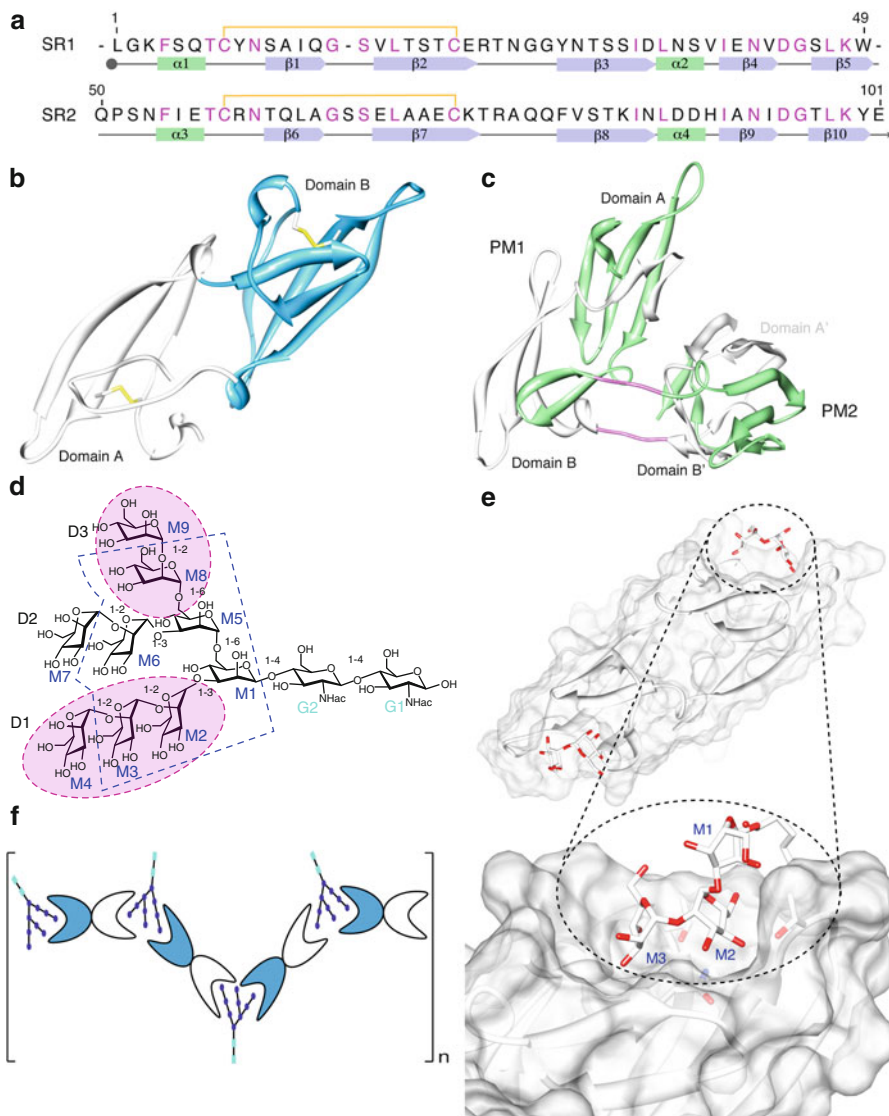
In order to provide a comprehensive view of the molecular basis of the anti-HIV properties of lectins, we describe here their atomic structures, distinct modes of glycan recognition, and binding epitopes on the oligosaccharide. The combined body of knowledge on these interactions may be leveraged for creating candidates as protein-based microbicides and may lead to novel directions in the development of alternative drug-leads for the prevention of HIV transmission.

## 7.2 Cyanovirin-N

Cyanovirin-N (CV-N) is a small (11 kDa) virucidal lectin originally identified in aqueous extracts from the cyanobacterium *Nostoc ellipsosporum* in a screen designed to discover anti-HIV compounds (Boyd et al. 1997). The amino acid sequence of CV-N comprises two tandem repeats, residues 1–50 (sequence repeat 1; SR1) and residues 51–101 (sequence repeat 2; SR2) (Gustafson et al. 1997). A pair of disulphide-bonded cysteines is present in each repeat, linking C8 to C22 in SR1 and C58 to C73 in SR2 (Fig. 7.1a) (Yang et al. 1999; Bewley et al. 1998; Gustafson et al. 1997). A monomeric structure of the native protein is only observed in solution, exhibiting a globular fold with an ellipsoidal shape, comprises two pseudo-symmetrical halves, termed domains A and B, respectively (Fig. 7.1b) (Bewley et al. 1998). Note that each tandem sequence repeats does not constitute an individual domain; instead, each domain is formed by strand exchange between the two repeats. The secondary structure elements in domain A are formed by a triple-stranded  $\beta$ -sheet ( $\beta$ 1,  $\beta$ 2, and  $\beta$ 3), a  $\beta$ -hairpin ( $\beta$ 9 and  $\beta$ 10), and two  $3_{10}$ -helical turns ( $\alpha$ 1 and  $\alpha$ 2), encompassing residues 1–38 and residues 90–101 (Fig. 7.1b). Likewise, domain B is composed of a triple-stranded  $\beta$ -sheet ( $\beta$ 6,  $\beta$ 7, and  $\beta$ 8), a  $\beta$ -hairpin ( $\beta$ 4 and  $\beta$ 5), and two  $3_{10}$ -helical turns ( $\alpha$ 3 and  $\alpha$ 4), comprising residues 39–89 (Fig. 7.1b).

CV-N was also found as a 3D domain-swapped dimer both in solution and crystal states (Fig. 7.1c) (Botos et al. 2002; Barrientos et al. 2004; Yang et al. 1999). High protein concentration is a major contributory factor for domain swapping of CV-N and the kinetic barrier between the domain-swapped dimer and the monomer can be overcome by increasing the temperature; thus for wild-type CV-N the domain swapped dimer is a trapped folding intermediate (Barrientos et al. 2004). Since half of one polypeptide chain is swapped between the two dimer halves, a “pseudomonomer” can be defined and this is formed by residues 1–50 of one chain and 51–101 of the other chain. Therefore, each pseudomonomer exhibits the same fold as the native monomer. Indeed all  $\phi$ / $\psi$  angles outside the hinge/loop region are within experimental error for the monomer and pseudomonomer and identical hydrophobic and charge interactions are present. It was also shown that the amino acid composition of the hinge-loop region that comprises residues Q50-N53 (Fig. 7.1a) plays a key role in domain swapping (Barrientos et al. 2002; Yang et al. 1999).

The solution NMR, crystal structures of CV-N, and a number of mutant variants have shed light on the molecular basis of CV-N antiviral activity. The protein binds



**Fig. 7.1** Structure and carbohydrate specificity of CV-N. **(a)** Sequence alignment of the first and second repeats of CV-N. **(b)** The overall structure of monomeric CV-N drawn in ribbon representation. **(c)** The structure of domain-swapped CV-N dimer in ribbon representation. **(d)** Chemical structure of Man-9. **(e)** Surface representation illustrating the interactions between CV-N and Man $\alpha$ 1–2Man $\alpha$  in both domains A and B, and between CV-N and synthetic hexamannose in domain A (zoom-in panel). **(f)** Schematic depiction of Man-9 binding by monomeric CV-N, illustrating the multisite and multivalent binding

readily to high mannose glycans via the reducing  $\text{Man}\alpha(1\text{--}2)\text{Man}$  ends of the D1 or D3 arms of Man-8 and Man-9 on the viral envelope glycoprotein gp120 (Fig. 7.1d) (Botos et al. 2002; Shenoy et al. 2001, 2002; Bewley 2001). It is this interaction with the glycans on HIV's gp120 and other mannosylated viral surface proteins that lies at the core of CV-N's inactivation of a wide range of enveloped viruses such as SIV, Ebola, influenza and hepatitis C, in addition to HIV (Ziolkowska and Wlodawer 2006; Barrientos et al. 2003, 2004; Barrientos and Gronenborn 2005; O'Keefe et al. 2003; Helle et al. 2006).

Two carbohydrate-binding sites, separated by  $\sim 35$  Å, are present on CV-N (Bewley 2001; Botos et al. 2002). Mapping these binding sites by NMR located the first binding site on domain A, comprising residues 1–7, 22–26, and 92–95. The second binding site is found on domain B and involves residues 41–44, 50–56, and 74–78 (Bewley 2001; Botos et al. 2002). The interaction mode of a disaccharide with the protein was revealed by the structure of CV-N in complex with  $\text{Man}\alpha(1\text{--}2)\text{Man}$ , determined in solution by NMR spectroscopy (Fig. 7.1e) (Bewley 2001). The disaccharide binds in a deep pocket on each domain in a stacked conformation. Hydrogen bonds with residues Lys3, Gln6, Thr7, Glu23, Thr25, Asn93, and Ile94 on domain A and with residues Glu41, Asp44, Ser52, Glu56, Thr57, Thr75, Arg76, and Gln78 on domain B appear to stabilize the interaction between the protein and the disaccharide (Bewley 2001).

The interactions of CV-N with the relevant high mannose glycans of gp120 were probed by solving X-ray structures of CV-N in complex with Man-9 (Fig. 7.1d) and a synthetic hexamannose (enclosed by blue dashed lines in Fig. 7.1d) (Botos et al. 2002). Since CV-N in the crystal is a domain-swapped dimer, four sugar-binding sites are present (Barrientos and Gronenborn 2002). In the two complex structures, only one sugar molecule was found interacting with domain A of the first pseudomonomer in the CV-N-Man-9 complex, while two sugars were bound to both A domains in the two pseudomonomers in the CV-N-hexamannose complex (Botos et al. 2002). No sugar was seen to interact with domain B in any pseudomonomer in either of the complexes. Instead, a well-defined, tightly bound buffer molecule (CHES) from the crystallization solution occupied the sugar-binding site on domain B in one of the two pseudomonomers (Botos et al. 2002).

The absence of any glycans in domain B in the domain-swapped crystal structures was explained by the different geometry of the sugar-binding site after domain swapping (Botos et al. 2002). Unlike in the monomer, where the binding pocket is intact on domain B and can accommodate a disaccharide in a stacked conformation (Bewley 2001), the position of the hinge and the relative orientation of the domains are altered upon domain swapping, and the close proximity of the hinge region to the binding site results in a slightly altered shape of the sugar-binding pocket (Botos et al. 2002). As a result, some of the essential protein–oligosaccharide hydrogen bonds may not be formed (Botos et al. 2002). For example, different positioning of the Ser52 O $\gamma$  atom interferes with its potential hydrogen-bonding to a mannose ring and perturbation of the side chain geometry of Asn53 in the hinge induces some steric hindrance, rendering sugar binding in the site on domain B unfavorable (Botos et al. 2002).

Unlike domain B, the binding site on domain A is located far away from the hinge region, and therefore is not affected by the geometry of the hinge-loop. It exhibits the same conformation in both the CV-N monomer and the 3D domain-swapped dimer (Botos et al. 2002). The very similar conformation of the sugar conformation when bound in domain A for Man $\alpha$ (1–2)Man in the solution structure and the mannose rings M2 and M3 of the D1 arm of hexamannose (see zoom-in panel of Fig. 7.1e) in the crystal structures is therefore not surprising.

CV-N has been reported to bind Man-9 with nanomolar affinity (Bewley and Otero-Quintero 2001), while binding to a hexamannoside was found to be significantly weaker, with affinities in the low micromolar range (Botos et al. 2002). The latter is comparable to the affinity for Man $\alpha$ (1–2)Man (Matei et al. 2008; Bewley 2001; Bewley et al. 2002). The much tighter apparent binding to Man-9 (or Man-8) is explained with CV-N's multisite binding to both the D1 and D3 arms of Man-9 (or Man-8) that facilitates cross-linking (Fig. 7.1f). Indeed it was unambiguously demonstrated that CV-N interacts with the individual arms of Man-9 with very similar binding strength (Shenoy et al. 2002) and a CV-N protein that possesses only a single binding site binds Man-3 and Man-9 with identical affinity (Matei et al. 2008). Therefore, it is the formation of multivalent, multisite interactions between CV-N and oligosaccharides on gp120 that explains its unusually potent activity and makes CV-N a promising potential candidate for development as a future pharmaceutical agent.

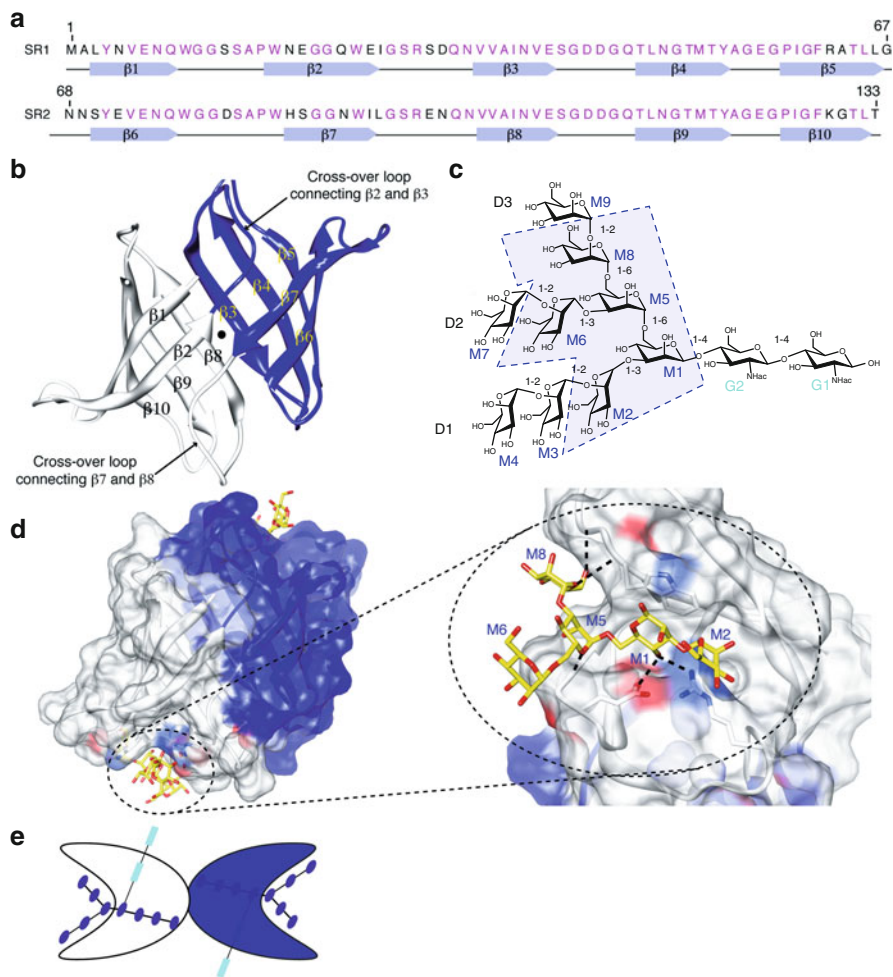
### 7.3 *Oscillatoria agardhii* Agglutinin

The *Oscillatoria agardhii* agglutinin (OAA) was isolated from the cyanobacterial *Oscillatoria agardhii* strain NIES-204 and is a protein of molecular mass of 13.9 kDa (Sato and Hori 2009; Sato et al. 2000). The amino acid sequence of OAA comprises 132 amino acids, arranged as two sequence repeats (Sato et al. 2007; Sato and Hori 2009). Residues 1–66 and residues 67–132 belong to sequence repeat 1 (SR1) and sequence repeat 2 (SR2), respectively. They are highly homologous with ~77 % sequence identity (51/66 residues) and ~86 % sequence similarity (57/66 residues) between SR1 and SR2 (Fig. 7.2a) (Koharudin et al. 2011; Sato et al. 2007).

The atomic structure of OAA was determined by X-ray crystallography (Koharudin and Gronenborn 2011; Koharudin et al. 2011). The overall architecture of OAA is a compact  $\beta$ -barrel that contains a continuous ten-stranded antiparallel  $\beta$ -sheet (Fig. 7.2a, b). Each of the amino acid sequence repeats folds into five  $\beta$ -strands, denoted as  $\beta$ 1– $\beta$ 5 and  $\beta$ 6– $\beta$ 10, for the first and second repeats, respectively (Fig. 7.2a). The two-sequence repeats are connected by a very short linker, comprising residues G67–N69.

The first two  $\beta$ -strands of each sequence repeat ( $\beta$ 1– $\beta$ 2 and  $\beta$ 6– $\beta$ 7) and the next three  $\beta$ -strands ( $\beta$ 3– $\beta$ 4– $\beta$ 5 and  $\beta$ 8– $\beta$ 9– $\beta$ 10) are positioned on opposite sides of the barrel (Fig. 7.2b), and the linkers connecting strands  $\beta$ 2 and  $\beta$ 3, and  $\beta$ 7 and  $\beta$ 8, respectively, cross at the top or the bottom of the barrel (Fig. 7.2b). As a result, the





**Fig. 7.2** Structure and carbohydrate specificity of OAA. **(a)** Sequence alignment of the first and second repeats of OAA. **(b)** The overall structure of OAA drawn in ribbon representation. **(c)** Chemical structure of Man-9. **(d)** Surface representation illustrating the interactions between OAA and  $\alpha 3, \alpha 6$ -mannopentaose in sites 1 and 2. **(e)** Schematic depiction of Man-9 binding by OAA in both sites 1 and 2

first two  $\beta$ -strands ( $\beta 1$ – $\beta 2$ ) in the first sequence repeat are located in the barrel between the  $\beta$ -strands from the second repeat, namely strands  $\beta 7$  and  $\beta 6$  on one side and strands  $\beta 10$ ,  $\beta 9$ , and  $\beta 8$  on the other side (Fig. 7.2b). Similarly, strands  $\beta 3$ ,  $\beta 4$ , and  $\beta 5$  of the first repeat are flanked either side by  $\beta 8$  and  $\beta 6$ , of the other sequence repeat (Fig. 7.2b). The swap of  $\beta$ -strands between the two sequence repeats creates an almost perfect C2 symmetric arrangement, rendering the conformation of the five  $\beta$  strands in each sequence repeat extremely similar (backbone atomic r.m.s.d. value of 0.66 Å).

The carbohydrate specificity of OAA for Man-9 was initially delineated in solution by NMR spectroscopy (Koharudin et al. 2011). Chemical shift mapping of  $^1\text{H}$  and  $^{15}\text{N}$  resonances for free and Man-9-bound OAA revealed that Man-9 binding to OAA is in slow exchange on the chemical shift scale (new bound resonances appear), suggesting relatively tight binding. The NMR titrations also revealed that OAA possesses two sugar-binding sites and that the Man-9 affinities for the two binding sites on OAA are distinct: the lower affinity site, site 1 (affected only at  $>1.1$  Man-9/protein molar ratio), comprises the loops connecting strands  $\beta 1$ – $\beta 2$ ,  $\beta 7$ – $\beta 8$ , and  $\beta 9$ – $\beta 10$ , and the higher affinity site (site 2) is located at the symmetrically related position and is made up by the loops connecting strands  $\beta 6$ – $\beta 7$ ,  $\beta 2$ – $\beta 3$ , and  $\beta 4$ – $\beta 5$ .

To further delineate the atomic details of the binding interface, OAA was co-crystallized with  $\alpha 3, \alpha 6$ -mannopentaose, the minimal unit of Man-9 recognized by OAA (enclosed by blue dashed lines and shaded in light blue in Fig. 7.2c) (Koharudin and Gronenborn 2011). Note that the binding between OAA and  $\alpha 3, \alpha 6$ -mannopentaose is in slow exchange on the NMR chemical shift scale, suggesting a relatively tight interaction, and that both binding sites demonstrated a very similar affinity for this sugar. The structure of  $\alpha 3, \alpha 6$ -mannopentaose-bound OAA was refined to 1.65 Å resolution, comparable to that of the apo OAA structure determined at 1.55 Å. A comparison between the apo- and  $\alpha 3, \alpha 6$ -mannopentaose-bound OAA structures reveals that they are very similar, with the compact ten-stranded antiparallel  $\beta$ -sheet barrel essentially identical. Only minor conformational differences, especially in the loops that are part of the two carbohydrate-binding sites, can be discerned. Note that clear additional electron density at opposite ends of the protein molecule permits the placement of a  $\alpha 3, \alpha 6$ -mannopentaose molecule into each site, and in the final density, an excellent fit of the atomic structure of two bound  $\alpha 3, \alpha 6$ -mannopentaose molecules, one per carbohydrate-binding site of OAA, into the density is noted (Fig. 7.2d).

The sugar-binding pockets of site 1 and site 2 on OAA are very similar (Fig. 7.2d) (Koharudin and Gronenborn 2011). OAA's carbohydrate recognition sites comprise short clefts, residing between the loops on the surface of the protein. Each binding site is formed primarily by residues in two loops that connect strands  $\beta 1$ – $\beta 2$  and  $\beta 9$ – $\beta 10$  and those connecting strands  $\beta 4$ – $\beta 5$  and  $\beta 6$ – $\beta 7$  for sites 1 and 2, respectively. These two loops contact the carbohydrate directly, in particular residues W10-G12 located in the loop that connects strands  $\beta 1$ – $\beta 2$  and E123-G124 in the connection between strands  $\beta 9$ – $\beta 10$  in binding site 1, and residues E56-G57 between strands  $\beta 4$  and  $\beta 5$  and W77-G79 in the loop between  $\beta 6$ – $\beta 7$  in binding site 2. The loops connecting strands  $\beta 7$ – $\beta 8$  in site 1 or strands  $\beta 2$ – $\beta 3$  in site 2 are slightly more remote, and contain amino acids with long side chains that can reach the carbohydrate, such as R95 in site 1 or R28 in site 2.

The zoom-in panel in Fig. 7.2d illustrates the interactions between OAA and  $\alpha 3, \alpha 6$ -mannopentaose in detail. The M5 $\alpha$ (1–6)M1 disaccharide unit of the  $\alpha 3, \alpha 6$ -mannopentaose is positioned most closely to the protein, and of the five-mannose carbohydrate moieties, the M2 $\alpha$ (1–3)M1 disaccharide is located deep inside the binding pocket while the M8 $\alpha$ (1–3)[M6 $\alpha$ (1–6)]M5 trisaccharide unit is pointing outwards. Overall, the branch point sits in the center of the binding site and all



mannose units are splayed out from the center. For the M2 $\alpha$ (1–3)M1 unit, the pyranose ring of M1 is stacked on top of the indole ring of the W10 or W77 side chains in site 1 or site 2, respectively. The pyranose ring of M2 is flanked by the long side chains of residues R95 or R28 in site 1 or site 2, respectively. On the other side of the cleft, where the M8 $\alpha$ (1–3)[M6 $\alpha$ (1–6)]M5 trisaccharide is located, the pyranose ring of M8 is flanked by residues in the loops connecting strands  $\beta$ 1– $\beta$ 2 and  $\beta$ 6– $\beta$ 7 in site 1 and 2, respectively, while the pyranose rings of M6 and M5 are flanked by residues in the  $\beta$ 9– $\beta$ 10 and  $\beta$ 4– $\beta$ 5 loops for site 1 and 2, respectively.

Of all the contacts in the binding sites it appears that the hydrophobic interaction between the aromatic side chain of W10 in site 1 and W77 in site 2 with the pyranose ring of M1 plays a critical role (Koharudin and Gronenborn 2011). In addition, several polar interactions are also observed. In particular, hydrogen bonds between the hydroxyl groups of the carbohydrate and main chain amide groups are present, augmented by several contacts with side chains (Koharudin and Gronenborn 2011). In binding site 1, hydrogen bonds are formed between the backbone amide of G11 and the C5 hydroxyl group of M8 (2.86 Å), the backbone amide of G12 and the C6 hydroxyl group of M8 (2.87 Å), the backbone amide of G124 and the C5 hydroxyl group of M5 (3.00 Å), and the backbone amide of G124 and the C6 hydroxyl group of M5 (3.15 Å). Side chain interactions in binding site 1 include hydrogen bonds between the C4 hydroxyl group of M1 and the side chain carboxyl group of E123 (2.81 Å) and between the C4 hydroxyl group of M1 and the terminal guanidinium group of R95 (2.89 Å).

Similarly, equivalent hydrogen bonds are found in binding site 2. Hydrogen bonds between the backbone amide of G78 and the C5 hydroxyl group of M8 (2.93 Å), the backbone amide of G79 and the C6 hydroxyl group of M8 (2.83 Å), the backbone amide of G57 and the C5 hydroxyl group of M5 (2.92 Å), and the backbone amide of G57 and the C6 hydroxyl group of M5 (3.20 Å) are present. Side chain hydrogen bonding is observed between the C4 hydroxyl group of M1 and the side chain carboxyl group of E56 (2.73 Å) and between the C4 hydroxyl group of M1 and the terminal guanidinium group of R28 (2.84 Å). Therefore, it can be concluded that the conformation of the two binding sites in the  $\alpha$ 3, $\alpha$ 6-mannopentaose-bound OAA structure is extremely similar.

In contrast to the low nanomolar activity of CV-N, OAA's antiviral potency is ~30-fold less (Koharudin et al. 2012). If a single binding contact between protein and sugar would be responsible for the activity, one would expect comparable inhibition with similar IC<sub>50</sub> values for CV-N and OAA, since they possess the same number of binding sites for carbohydrate. This clearly is not the case. This difference is caused by the fact that OAA can only recognize a single epitope of Man-8/9, namely the branched core mannose unit (Fig. 7.2e) (Koharudin and Gronenborn 2011), while CV-N binds to two epitopes, the Man $\alpha$ (1–2)Man units of the D1 or D3 arms (Botos et al. 2002; Bewley 2001) of the glycan (Fig. 7.1f). The multisite and multi-epitope interaction that CV-N can engage in is clearly not possible for OAA. On the other hand, the presence of distinctively different recognition epitopes on the glycan for OAA and CV-N may possibly be explored in a synergistic fashion when these core and terminal mannose-recognizing lectins are combined in targeting gp120 and blocking HIV infectivity.

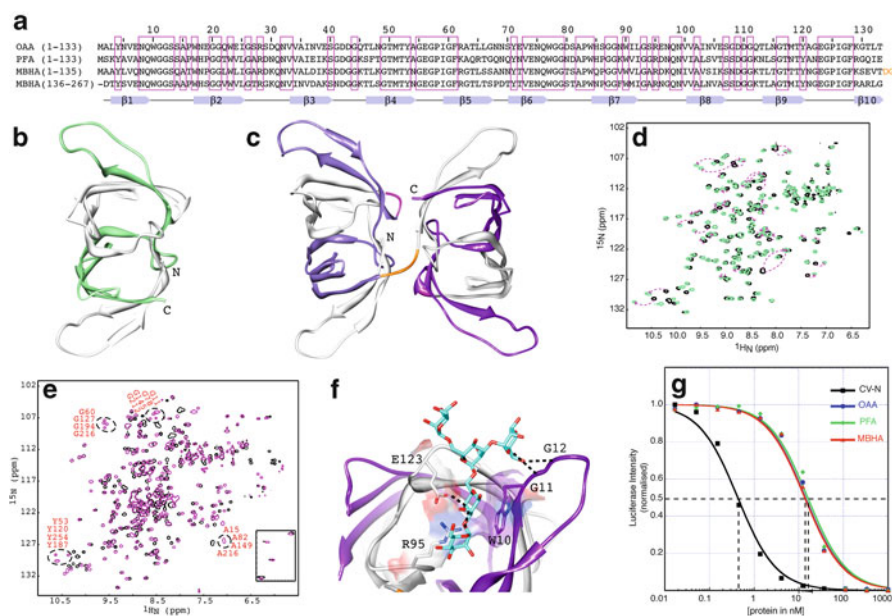
## 7.4 *Oscillatoria agardhii* Agglutinin Homolog Proteins

The compact,  $\beta$ -barrel-like architecture of the cyanobacterial OAA is very different from previously characterized lectin structures and unique among all available protein structures in the protein data bank (Koharudin et al. 2011). Most importantly, so far OAA's carbohydrate recognition of Man-9 is also unique, compared to all other antiviral lectins. While most of the known lectins that block HIV infectivity recognize the reducing or nonreducing end mannoses of Man-8/9, OAA recognizes the branched core unit of Man-8/9. These properties make OAA distinct and rare among all antiviral lectins (Koharudin and Gronenborn 2011; Koharudin et al. 2011).

Recently, genes coding for OAA homologs were discovered in a number of other prokaryotic microorganisms, including cyanobacteria, proteobacteria, and chlorobacteria, as well as in a eukaryotic marine red alga (Koharudin et al. 2012; Sato and Hori 2009). Similar to OAA, these proteins, henceforth termed *Oscillatoria agardhii* agglutinin homologs (OAAHs), contain a sequence repeat of ~66 amino acids, with the number of repeats varying for different family members. For example, the 133-residue *Pseudomonas fluorescens* homolog, *Pseudomonas fluorescens* agglutinin (PFA), contains two sequence repeats, like OAA, while the *Myxococcus xanthus* homolog, *Myxococcus xanthus* hemagglutinin (MBHA), contains four sequence repeats over a length of 268 residues (Sato and Hori 2009; Koharudin et al. 2012). Apart from data for the founding member OAA, neither three-dimensional structures nor information about carbohydrate-binding specificities and antiviral activity is available up to now for any other member of the OAAH family. In order to further characterize this important lectin family, structural and carbohydrate specificity analyses for these two additional members of the OAAH family, PFA and MBHA, will be discussed in this chapter.

The amino acid sequences of PFA and MBHA share extensive sequence similarity to OAA, with ~62 % identity for pairwise alignment (Fig. 7.3a) (Koharudin et al. 2012; Sato and Hori 2009). Interestingly, the majority of the amino acid conservation is seen in the carbohydrate-binding regions, delineated previously in OAA (Fig. 7.2). This region encompasses the loops between  $\beta 1$ – $\beta 2$ ,  $\beta 7$ – $\beta 8$ , and  $\beta 9$ – $\beta 10$  in the first binding site and between  $\beta 6$ – $\beta 7$ ,  $\beta 2$ – $\beta 3$ , and  $\beta 4$ – $\beta 5$  in the second binding site. In addition, notable sequence conservation is seen throughout the secondary structure elements.

PFA assembles into a single, compact,  $\beta$ -barrel-like domain (Fig. 7.3b) as previously observed for OAA (Fig. 7.2b) (Koharudin et al. 2012). Each sequence repeat folds up into five  $\beta$ -strands, denoted as  $\beta 1$ – $\beta 5$  (colored in white) and  $\beta 6$ – $\beta 10$  (colored in green) for the first and second repeats, respectively. Different from OAA or PFA, MBHA contains four sequence repeats and in its crystal structure each two-sequence repeat folds into a  $\beta$ -barrel, resulting in a tandem arrangement of two barrels (Fig. 7.3c). The first barrel is composed of the first ten  $\beta$ -strands, colored in grey and purple, and the second barrel is made up by the second ten  $\beta$ -strands, also colored in grey and purple, respectively. A short linker comprising residues T133–G135 connects the first and second barrel (colored in orange). The structures of the first



**Fig. 7.3** Structure and carbohydrate specificity of PFA and MBHA. **(a)** Sequence alignment of OAA, PFA, and MBHA. **(b)** The overall structure of PFA drawn in ribbon representation. **(c)** The overall structure of MBHA drawn in ribbon representation. **(d)** Superposition of the two-dimensional  $^1\text{H}$ - $^{15}\text{N}$  HSQC spectra of free (*black*) and  $\alpha$ 3, $\alpha$ 6-mannopentaose-bound (*light green*) PFA at 1:3 molar ratio of PFA to sugar. **(e)** Superposition of the two-dimensional  $^1\text{H}$ - $^{15}\text{N}$  HSQC spectra of free (*black*) and  $\alpha$ 3, $\alpha$ 6-mannopentaose-bound (*magenta*) MBHA at 1:6 molar ratio of MBHA to sugar. **(f)** Surface representation illustrating the interactions between one of the binding sites of MBHA and  $\alpha$ 3, $\alpha$ 6-mannopentaose. **(g)** Anti-HIV activity assays for CV-N, OAA, PFA, and MBHA

(residues A2 to V132) and second (residues D136 to L266) barrel of MBHA are very similar, with a backbone atom r.m.s.d value of 0.63 Å.

The overall structures of these new family members closely resemble that of the founding member, OAA (Koharudin et al. 2012). The overall r.m.s.d values for backbone atoms between OAA (residues A2 to L132) and PFA (residues S2 to I132) is 0.50 Å, between OAA (residues A2 to L132) and the first domain of MBHA (residues A2 to V132) is 0.79 Å, and between OAA (residues A2 to L132) and the second domain of MBHA (residues D136 to L266) is 0.70 Å. Similarly, PFA is close in structure to each of the MBHA domains, with backbone atom r.m.s.d. values of 0.74 Å between PFA and the first domain (residues S2 to I132 of PFA and residues A2 to V132 of MBHA) and 0.60 Å between PFA and the second domain of MBHA (residues S2 to I132 of PFA and residues D136 to L266 of MBHA). This extensive structural similarity parallels the significant degree of amino acid identity throughout the protein sequences (Fig. 7.3a).

As discussed above, OAA's anti-HIV activity is associated with its binding to N-linked high mannose glycans on the viral envelope glycoprotein gp120 (Sato

et al. 2007; Koharudin et al. 2011). The epitope that is recognized by OAA on Man-8/9 is  $\alpha$ 3, $\alpha$ 6-mannopentaose (Man $\alpha$ (1–3)[Man $\alpha$ (1–3)[Man $\alpha$ (1–6)]Man $\alpha$ (1–6)]Man), the branched core unit of the triantennary high mannose structures (Koharudin and Gronenborn 2011). As expected from the sequence and structural conservation seen for PFA, MBHA and OAA, the two new OAAH members also share the carbohydrate specificity of OAA (Koharudin et al. 2011, 2012; Koharudin and Gronenborn 2011). Titrations of the proteins with  $\alpha$ 3, $\alpha$ 6-mannopentaose and monitoring the chemical shift changes by  $^1\text{H}$ - $^{15}\text{N}$  HSQC spectroscopy confirmed that both PFA and MBHA specifically and tightly interact with  $\alpha$ 3, $\alpha$ 6-mannopentaose (Fig. 7.3d, e, respectively) (Koharudin et al. 2012). Here again, similar to OAA,  $\alpha$ 3, $\alpha$ 6-mannopentaose binding to PFA and MBHA is in slow exchange on the NMR chemical shift scale, suggesting a relatively tight interaction.

Chemical shift differences between free PFA (black) and sugar-bound resonances at saturation (sugar:protein molar ratio of 3:1; green) (Fig. 7.3d) show that the most strongly perturbed resonances belong to residues in the loops, connecting strands  $\beta$ 1– $\beta$ 2 and  $\beta$ 9– $\beta$ 10 and between strands  $\beta$ 4– $\beta$ 5 and  $\beta$ 6– $\beta$ 7. Smaller changes are observed for residues in the loops connecting  $\beta$ 7– $\beta$ 8 and  $\beta$ 2– $\beta$ 3. Those resonances of PFA that were perturbed include residues N8-S14, W17-H18, I101, N118-Y120, E123-G124, I126-G127, and G130 in the first binding site and residues M51, Y53, E56-G57, I59-G60, Q76-A82, and W84-H85 in the second binding site, essentially equivalent to those in OAA (Koharudin and Gronenborn 2011; Koharudin et al. 2012). Therefore, the binding sites in both proteins are located in corresponding regions of the structures, consistent with the extensive sequence conservation in these sites.

NMR titration experiments were also used to determine which residues of MBHA interact with  $\alpha$ 3, $\alpha$ 6-mannopentaose. Based on the superposition of 2D  $^1\text{H}$ - $^{15}\text{N}$  HSQC spectra of apo (black) and  $\alpha$ 3, $\alpha$ 6-mannopentaose-bound (magenta) MBHA (Fig. 7.3e), it is clear that the affected MBHA residues are, not surprisingly, equivalent to those in OAA or PFA (Koharudin et al. 2012; Koharudin and Gronenborn 2011). However, for MBHA, complete saturation for all resonances was only achieved at a sugar:protein molar ratio of 6:1 (Fig. 7.3e), consistent with four sugar binding sites per polypeptide chain.

From the structure of the MBHA- $\alpha$ 3, $\alpha$ 6-mannopentaose complex specific contacts between MBHA side chains and the carbohydrate can be discerned (Koharudin et al. 2012). As shown in Fig. 7.3f, the aromatic side chain of W144 in site 3 (or W211 in site 4) plays a critical role, providing hydrophobic contacts for the pyranose ring of M1. In addition, several hydrogen bond contacts are also observed (Fig. 7.3f). Given the high sequence conservation in the carbohydrate-binding regions between MBHA and OAA (Fig. 7.3a), all specific contacts that are observed in the MBHA- $\alpha$ 3, $\alpha$ 6-mannopentaose complex are unsurprisingly identical to those observed previously in the OAA- $\alpha$ 3, $\alpha$ 6-mannopentaose complex (Koharudin et al. 2012). Note that in the structure of  $\alpha$ 3, $\alpha$ 6-mannopentaose-bound MBHA, only the two higher affinity sites, i.e., binding sites 3 and 4 that are both located in the second barrel of MBHA, are occupied by the glycan (Koharudin et al. 2012). No equivalent density was observed in the other two binding sites in the first barrel. However, it is well

established from NMR titration data that sites 1 and 2 can bind  $\alpha 3, \alpha 6$ -mannopentaose similar to binding sites 3 and 4, albeit with slightly reduced affinity.

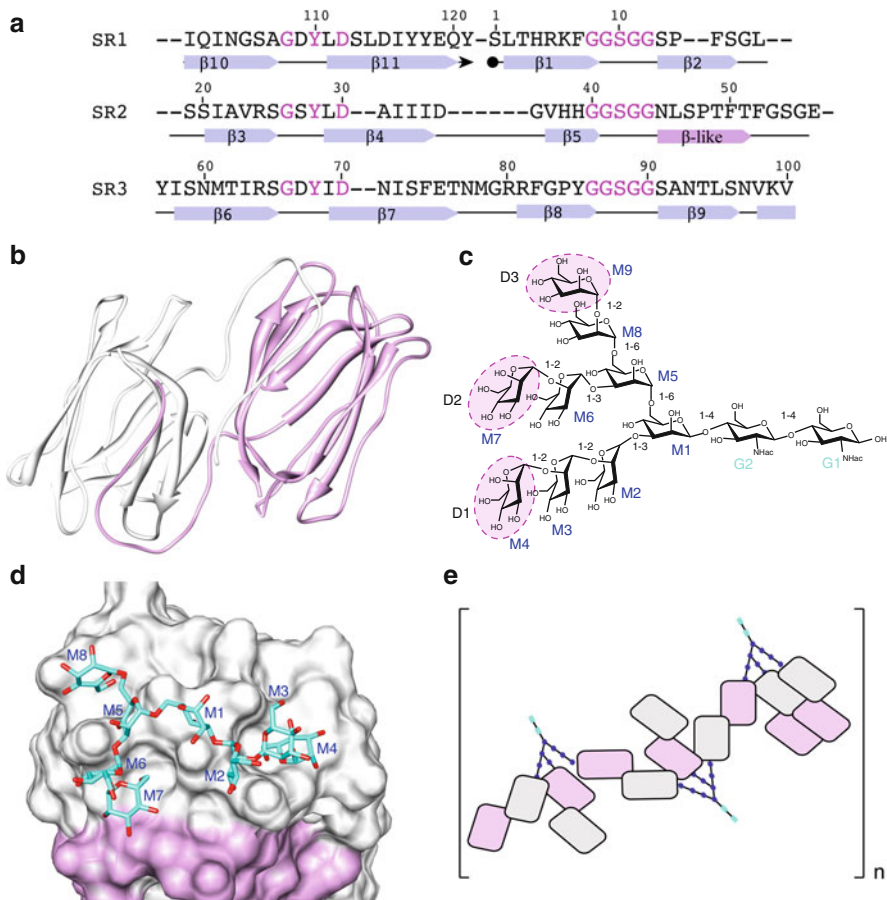
OAA's anti-HIV activity is mediated by the specific recognition of  $\alpha 3, \alpha 6$ -mannopentaose, the branched core unit of Man-8 and Man-9 (Koharudin and Gronenborn 2011), sugars on the HIV-1 envelope glycoprotein gp120. Similarly, it was shown that both PFA and MBHA also interact with this glycan, suggesting possible similar activities for these lectins (Koharudin et al. 2012). Indeed, the HIV assay data (Fig. 7.3g) clearly reveals that PFA and MBHA possess anti-HIV activity (Koharudin et al. 2012). All three OAAHs display very similar  $IC_{50}$  values, ranging from  $12 \pm 1$  nM for OAA and MBHA to  $15 \pm 1$  nM for PFA (Koharudin et al. 2012). As discussed above, these values are approximately 30-fold higher than that obtained for CV-N in the same experiment ( $0.4 \pm 0.1$  nM) (Koharudin et al. 2012). Again, it is interesting to note that, if solely avidity considerations on the protein were significant, one would expect MBHA to exhibit higher anti-HIV activity than OAA and PFA, given that it possesses four sugar-binding sites, compared to only two sites in latter lectins. However, the available data suggests that only two binding sites on the protein are required for any OAAH lectin family member to display anti-HIV activity and that the additional binding sites in some members merely increase the probability to engage the sugars of gp120.

## 7.5 Griffithsin

Griffithsin (GRFT), a lectin isolated from the red alga *Griffithsia* sp. that was collected from the waters off New Zealand, also belongs to those lectins that possess potent anti-HIV activity (Mori et al. 2005). The gene-encoding GRFT has not been isolated, but the amino acid sequence was obtained directly from protein purified from the cyanobacterium. The protein contains a single 121-amino acid chain (Fig. 7.4a), of which 120 residues are common amino acids while, surprisingly, the 31st residue (151 Da) does not appear to correspond to any standard amino acid and its identity is still unknown (Mori et al. 2005). Analysis of GRFT's sequence indicates the presence of three sequence repeats with residues 1–18 and residues 101–121 assigned to sequence repeat 1 (SR1), residues 19–56 to sequence repeat 2 (SR2), and residues 57–120 to sequence repeat 2 (SR3) (Fig. 7.4a). Two conserved regions are present; the GxYxD and the GGSGG motifs that are located in two distinct loop regions (Fig. 7.4a). Note that, like the OAAH lectin family members, there are no cysteine residues among its 121 amino acids.

The structure of GRFT was determined using recombinant proteins expressed in *E. coli* (Giomarelli et al. 2006) and *Nicotiana benthamiana* (Ziolkowska et al. 2006; O'Keefe et al. 2009). In both constructs residue 31 of GRFT was replaced by an alanine and this substitution did not seem to affect the carbohydrate-binding properties of the lectin. For the *E. coli* construct, the protein contains an N-terminal 6-His affinity tag followed by a thrombin cleavage site, extending the protein sequence by 17 amino acids (Giomarelli et al. 2006). Since the additional sequence





**Fig. 7.4** Structure and carbohydrate specificity of GRFT. **(a)** Sequence alignment of the three repeats of GRFT. **(b)** The overall structure of domain-swapped GRFT dimer, drawn in ribbon representation. **(c)** Chemical structure of Man-9. **(d)** Surface representation illustrating the detailed interactions between GRFT and synthetic Man-9. **(e)** Schematic depiction of Man-9 binding by GRFT dimer, illustrating the multisite and multivalent binding

could not be removed, these cloning artifact residues are present in the crystallized protein. On the other hand, the plant-expressed protein is a “native” construct, resembling more closely the authentic protein, although with an acetylated N terminus and mutated residue 31 (Ziolkowska et al. 2006; O’Keefe et al. 2009). Both constructs were crystallized and their structures determined. Interestingly, however, only a single molecule is present in the asymmetric unit of crystals of the His-tagged GRFT, whereas all crystal forms grown from the plant-expressed protein contain two molecules and diffract significantly better than the His-tagged crystals, most likely due to the absence of the N-terminal extension (Ziolkowska et al. 2006).



The GRFT fold is a  $\beta$ -prism-I motif (Ziolkowska et al. 2006) that consists of three repeats of an antiparallel four-stranded  $\beta$ -sheet (Fig. 7.4a) that form a triangular prism (Fig. 7.4b). Surprisingly, GRFT forms a domain-swapped dimer in which the first two  $\beta$ -strands of one chain are associated with ten strands of the other chain and vice versa (Fig. 7.4b), completely distinct from other  $\beta$ -prism-I lectins (Chandra et al. 2006; Raval et al. 2004).

The carbohydrate specificity of GRFT was initially probed using soluble gp120 in an inhibition-binding assay with monosaccharide glucose, mannose, and *N*-acetylglucosamine (Mori et al. 2005). The detailed contacts between the protein and mannose were later elucidated in the crystal structure of GRFT-mannose complex determined at a resolution of 1.8 Å (Ziolkowska et al. 2006). A comparison between apo- and mannose-bound GRFT structures yields very similar protein conformations with an r.m.s.d value of 0.46 Å, indicating that sugar binding does not induce any large conformational changes in the protein. Six mannose molecules were found bound in two groups of three mannoses. Each mannose group engages in direct intermolecular contacts with each monomer of the GRFT domain-swapped dimer. Therefore, it appears that a GRFT-swapped dimer contains a total of six nearly identical carbohydrate-binding sites, i.e., three sites per GRFT pseudomonomer. Further interaction studies with various mannose disaccharides confirmed that, unlike CV-N or the OAAH lectin family members, each GRFT-binding site indeed can bind an individual mannose monosaccharide located at each reducing end of Man-8/9 oligosaccharides (enclosed by magenta dashed lines and shaded in light magenta in Fig. 7.4c) (Ziolkowska et al. 2006, 2007).

The detailed atomic interactions between GRFT and Man-9 are seen in the crystal structure of an engineered monomeric GRFT (mGRFT) in complex with synthetic nanomannoside oligosaccharides (Moulaei et al. 2010a). This mGRFT was generated by inserting two or four amino acids at the dimerization interface and, different from the GRFT-swapped dimer, the mGRFT contains only three sugar-binding sites (Moulaei et al. 2010a). Similar to the mannose bound GRFT complex structure, all three binding sites are occupied. Binding sites 1 and 3 are occupied by mannoses M4 and M7 from the D1 and D2 arms of a single nanomannoside in the asymmetric unit, respectively (Fig. 7.4d), whereas site 2 is occupied by mannose M6 from the D2 arm of another nanomannoside in the symmetrically related molecule (Fig. 7.4c). Note that all three binding sites are created by the equivalent loops, connecting the first and fourth strand of each  $\beta$ -sheet, containing strictly conserved GGSGG sequences (Fig. 7.4a). This positions the three sugar-binding sites in an almost perfect equilateral triangle at the edges of the protein, with the carbohydrate molecules separated by about 15 Å from each other (Fig. 7.4d) (Moulaei et al. 2010a).

The antiviral activity of GRFT has been tested on T-lymphoblastic cells and antiviral activity was found at picomolar concentrations, rendering GRFT the most potent antiviral lectin to date (Ziolkowska et al. 2006; Moulaei et al. 2010a; Mori et al. 2005). While both dimer and monomer forms of GRFT are active against HIV, the dimeric protein possesses higher potency than mGRFT. The binding of GRFT and mGRFT to the viral envelope protein gp120 is also different, with mGRFT displaying

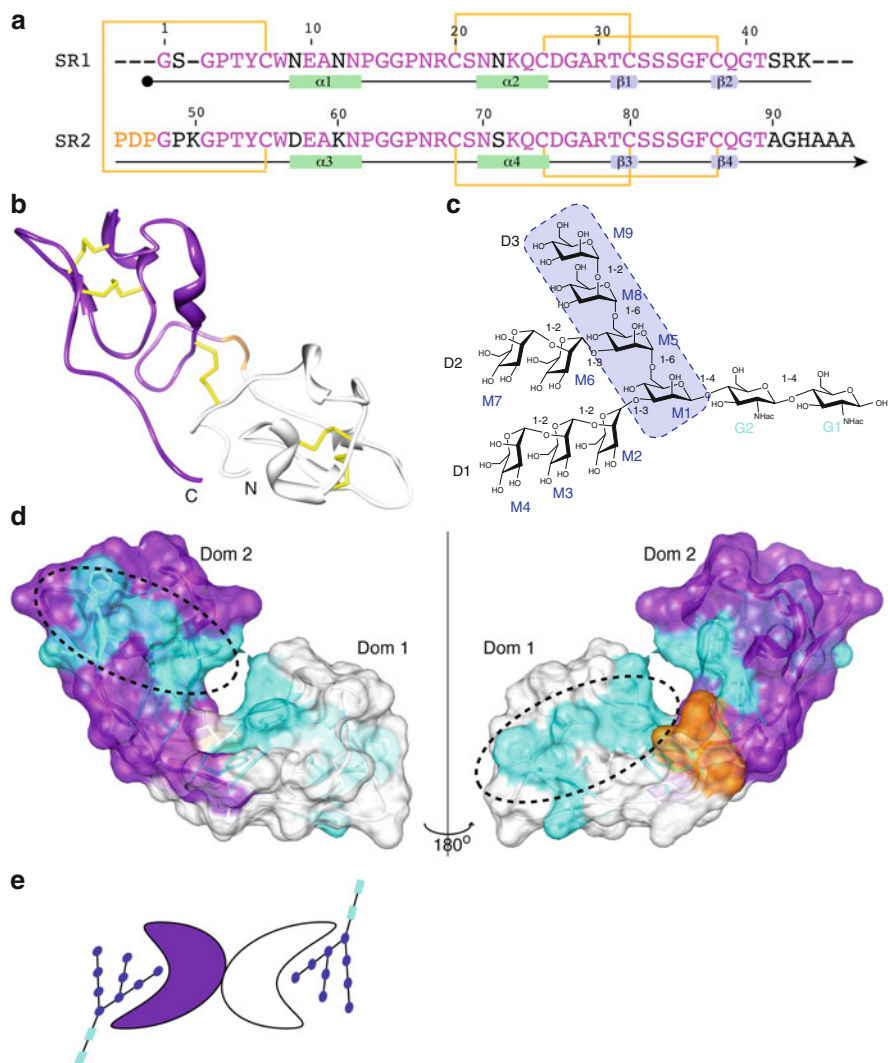
approximately 50-fold lower avidity (Moulaei et al. 2010a), despite similar enthalpies and dissociation constants of both GRFT and mGRFT for binding nonamannoside. Therefore, it seems that interactions between individual high mannose oligosaccharides and individual monomeric units of GRFT do not suffice to create potent antiviral activity. Thus, only when cross-linking of multiple high mannose oligosaccharides on gp120 is induced can significant antiviral potency of GRFT be shown (Fig. 7.4e), similarly to the original observation for CV-N (Fig. 7.1f). Interestingly, in a recent study in which a single or all three binding sites of the GRFT were destroyed was interpreted such that GRFT's activity is not caused by simply binding to gp120, but that the structure of gp120 or its oligomeric state is affected (Xue et al. 2012).

## 7.6 Scytovirin

The antiviral lectin scytovirin (SVN) was isolated from aqueous extracts of the cyanobacterium *Scytonema varium* as part of a program investigating anti-HIV activity in natural product extracts (Bokesch et al. 2003). A single chain of SVN contains 95 amino acids, with a molecular mass of 9713 Da (Bokesch et al. 2003; Xiong et al. 2006a). Similar to the CV-N or OAA discussed above, SVN also displays an internal sequence duplication (Fig. 7.5a) (Bokesch et al. 2003), with high (~75 %) sequence identity between the N-terminal sequence (residues 1–48) and its C-terminal counterpart (residues 49–95). The sequence contains a large number of cysteine residues, ten in total (Bokesch et al. 2003) and, as seen in the crystal structures of natural and recombinant SVN, five disulfide bond are formed between C7–C55, C20–C32, C26–C38, C68–C80, and C74–C86 (Fig. 7.5a) (Moulaei et al. 2007, 2010b).

The crystal structures of this antiviral lectin were solved at 1.3 and 1.0 Å resolution for the natural and recombinant proteins, respectively (Moulaei et al. 2007). Unlike CV-N or GRFT, SVN is strictly monomeric (Fig. 7.5b), with no indication of oligomerization under any conditions (Moulaei et al. 2007). However, similar to OAA or PFA or the monomer of CV-N, the SVN structure revealed two highly symmetric domains, domain 1 (D1) and domain 2 (D2) (Fig. 7.5b). The structures of the two domains are virtually identical with an r.m.s.d value of 0.25 Å (Moulaei et al. 2007) and the fold of SVN is novel, compared to other proteins in the database up to now.

SVN has been shown to bind specific oligosaccharides on gp41 and gp120 (Bokesch et al. 2003). In particular, it was shown to bind to a tetrasaccharide substructure of the high mannose oligosaccharides that decorate these HIV-1 envelope glycoproteins (McFeeters et al. 2007; Adams et al. 2004). In order to demonstrate carbohydrate binding of the protein, NMR titration experiments were used to monitor the chemical shift perturbation of SVN using the Man $\alpha$ (1 $\rightarrow$ 2)Man $\alpha$ (1 $\rightarrow$ 6)Man $\alpha$ (1 $\rightarrow$ 6)Man tetrasaccharide (enclosed by magenta dashed lines and shaded in light magenta in Fig. 7.5c) (McFeeters et al. 2007). The results indicated that one binding site was observed in each domain of SVN and that the two sites interact differently with the same tetrasaccharide (McFeeters et al. 2007). At low NMR magnetic field (500 MHz), residues in D1 exhibit intermediate exchange whereas



**Fig. 7.5** Structure and carbohydrate specificity of SVN. **(a)** Sequence alignment of the first and second repeats of SVN. **(b)** The overall structure of SVN, drawn in ribbon representation. **(c)** Chemical structure of Man-9. **(d)** Surface representation illustrating the interactions between SVN and the Man $\alpha$ (1 $\rightarrow$ 2)Man $\alpha$ (1 $\rightarrow$ 6)Man $\alpha$ (1 $\rightarrow$ 6)Man tetrasaccharide in sites 1 and 2 based on NMR titration data. **(e)** Schematic depiction of Man-9 binding by SVN in both sites 1 and 2

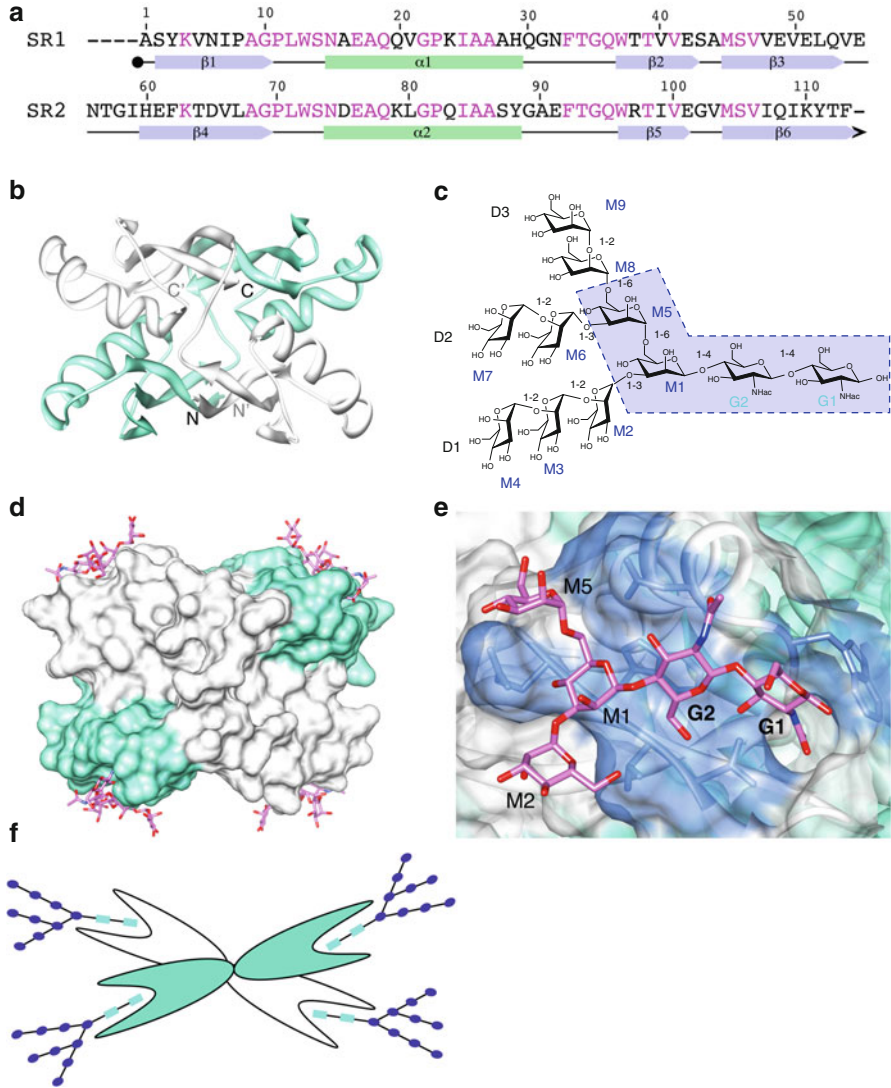
residues in D2 exhibit fast exchange. However, when the titrations were conducted at high NMR magnetic field (800 MHz), the intermediate exchange becomes slow and the fast exchange becomes intermediate, suggesting that the carbohydrate affinity of D1 is slightly tighter than that of D2. Using a two-site model fitting, dissociation constant values were calculated as  $\sim 30 \mu\text{M}$  for D1 and  $\sim 160 \mu\text{M}$  for D2 (McFeeters et al. 2007).

Residues whose resonances exhibit backbone amide chemical shift changes upon addition of tetrasaccharide are almost identical for the two domains (Fig. 7.5d) (McFeeters et al. 2007) and mapping of the affected residues onto the protein structure reveals very similar binding grooves on the surface of each domain of SVN. Three aromatic residues (residues Y6, W8, F37 in D1 and the corresponding residues Y54, W56, F85 in D2) in the two domains are all clustered near the binding sites and experience chemical shift perturbations, suggesting tetrasaccharide binding (McFeeters et al. 2007). As to the differences, it was speculated that they may partially be explained by the differences in residue composition in the two domains of SVN that are involved in glycan binding (McFeeters et al. 2007). For example, residue N9 that is located in D1 shows large backbone amide chemical shift changes in the presence of tetrasaccharide while the corresponding residue D57 in D2 only exhibited minor chemical shift perturbations upon binding. Similarly, no chemical shift perturbations are seen for G76 in D2, different from the significant change observed for the equivalent G28 in D1. Since G1 and S2 are close to residues with the largest chemical shift perturbations in D1, this suggests that an ordered N-terminus is necessary for tight binding. Since SVN does not bind to the  $\text{Man}\alpha(1\rightarrow6)\text{Man}\alpha(1\rightarrow6)\text{Man}$  trisaccharide (Adams et al. 2004; McFeeters et al. 2007), it seems that the reducing end  $\text{Man}\alpha(1\rightarrow2)\text{Man}$  in the tetrasaccharide is required in the interaction.

SVN is capable to interact with a single epitope of Man-8/9 (Fig. 7.5e), similar to that of OAA (Fig. 7.2e). Interestingly, however, SVN is active at low nanomolar concentrations against T-tropic strains and primary isolates of HIV-1, but is 300-fold less effective against M-tropic strains (Bokesch et al. 2003). Also, unlike CV-N that inactivates the virus on contact (Boyd et al. 1997), the viral inhibition by SVN is reversible (Bokesch et al. 2003). Pretreatment and removal of SVN caused normal susceptibility to HIV infection in uninfected CEM-SS cells and a normal infectivity in cell-free virus (Bokesch et al. 2003). It is also worth pointing out (1) that an individual domain of SVN has been reported to possess anti-HIV activity, with Domain 1 exhibiting similar activity as the full-length protein, while Domain 2 was much less active (Xiong et al. 2006b), and (2) that the truncated individual domain of SVN loses activity when its N-terminus is changed (Xiong et al. 2006b). Although there is no doubt that SVN binds to oligosaccharides and thereby inhibits HIV infection, no crystal structure is available at present of a complex between the protein and its carbohydrate ligand.

## 7.7 *Microcystis viridis* Lectin

*Microcystis viridis* lectin (MVL) is a 113 amino acid protein (~13 kDa) that was isolated from the freshwater bloom-forming cyanobacterium *Microcystis viridis* NIES-102 (Yamaguchi et al. 1999). The amino acid sequence of MVL comprises two highly homologous sequence repeats, SR1 and SR2, each containing 54 amino acids with about 50 % identity (Fig. 7.6a) (Yamaguchi et al. 1999). The two sequence



**Fig. 7.6** Structure and carbohydrate specificity of MLV. **(a)** Sequence alignment of the first and second repeats of MLV. **(b)** The overall structure of SVN homodimer, drawn in ribbon representation. **(c)** Chemical structure of Man-9. **(d)** Surface representation illustrating the interactions between MLV homodimer and the Man $\alpha$ (1-6)Man $\beta$ (1-4)GlcNAc $\beta$ (1-4)GlcNAc tetrasaccharide in sites 1 and 2. **(e)** Detailed interactions between MLV and the Man $\alpha$ (1-6)Man $\beta$ (1-4)GlcNAc $\beta$ (1-4)GlcNAc tetrasaccharide in one of the four binding sites. **(f)** Schematic depiction of Man-9 binding by MLV

repeats are joined by a five residue-linker (Fig. 7.6a). Up to now, the MVL protein sequence remains unique, as there is no significant similarity or homology to any reported protein sequences in the database.

The crystal structure of MVL shows that the protein forms a stable, symmetric homodimer (Fig. 7.6b) (Williams et al. 2005). As expected on the basis of the observed sequence duplication (Fig. 7.6a), each monomer is formed by two similar domains, domains A and B. Each domain contains a three-stranded antiparallel  $\beta$ -sheet (formed by strands  $\beta$ 1,  $\beta$ 2, and  $\beta$ 3 in domain A and strands  $\beta$ 4,  $\beta$ 5, and  $\beta$ 6 in domain B) and a single  $\alpha$ -helix that is packed against one face of the sheet ( $\alpha$ 1 and  $\alpha$ 2 located between strands  $\beta$ 1 and  $\beta$ 2 and between strands  $\beta$ 4 and  $\beta$ 5 in domains A and B, respectively) (Fig. 7.6a, b) (Williams et al. 2005). Three bulges in the  $\beta$ -strands, between residues 34–35, 40–41, and 44–45 in domain A and residues 93–94, 99–100, and 103–104 in domain B distort the normal twist of the  $\beta$ -sheet, such that it wraps around the  $\alpha$ -helix. The interactions between domains A and B of a single monomer are very limited as the domain is boomerang-shaped, connected by residues 55–59 (Williams et al. 2005). Indeed, the long axes of the two domains are oriented approximately orthogonal to each other; and the C-domain can be superimposed on the N-domain by an  $\sim 180^\circ$  rotation about an axis that bisects the angle between the two domains. In the dimer, two boomerang-shaped monomers interlock and each domain from one monomer contacts both domains from the second monomer (Fig. 7.6b).

Two sugar-binding sites per monomer were identified on MVL using NMR titrations (Bewley et al. 2004). A number of other mannose-containing carbohydrates were tested for binding and it was shown that MVL does bind  $\text{Man}_6\text{GlcNAc}_2$  with low micromolar affinity, but not  $\alpha$ - and  $\beta$ -linked dimannosides, disaccharides  $\text{Man}\beta(1\text{--}4)\text{GlcNAc}$  and  $\text{GlcNAc}\beta(1\text{--}4)\text{GlcNAc}$ , or mannotriose (Bewley et al. 2004). It was therefore suggested that high-affinity carbohydrate binding requires the presence of mannose and glucosamine residues and at least a tetrasaccharide core structures, such as  $\text{Man}\alpha(1\text{--}6)\text{Man}\beta(1\text{--}4)\text{GlcNAc}\beta(1\text{--}4)\text{GlcNAc}$  (enclosed by blue dashed lines and shaded in light blue in Fig. 7.6c).

The carbohydrate specificity of MVL was later confirmed by X-ray analysis (Williams et al. 2005). A crystal structure of the complex of MVL with  $\text{Man}_3\text{GlcNAc}_2$  shows four independent carbohydrate-binding sites per homodimer, two each within a single polypeptide chain (Fig. 7.6d) (Williams et al. 2005). No significant conformational changes in MVL are induced by sugar binding. The  $\text{Man}\beta(1\text{--}4)\text{GlcNAc}\beta(1\text{--}4)\text{GlcNAc}$  (M1-G2-G1) trisaccharide core sits tightly in each binding site and the reducing GlcNAc residue (G1) sugar unit is clearly essential for defining the specificity of carbohydrate binding as it is buried deep inside the binding pocket (Fig. 7.6e).

The details of the interaction between MVL and the carbohydrate are provided by the X-ray structure: In the G1 sugar unit, the acetyl methyl group fits into a deep hole and is in van der Waals contact with the side chains of Pro-11, Trp-13, the methyl groups of Leu-12 and Thr-39 in domain A (Fig. 7.6e) or of Pro-70, Trp-72, the methyl groups of Leu-71 and Thr-98 in domain B (not shown), whereas the acetyl oxygen atom is hydrogen-bonded to the backbone amide of Ser-43 in domain



A (Fig. 7.6e) and that of Gly-102 in domain B (not shown). The acetyl NH group of the G1 sugar unit is hydrogen-bonded to the backbone oxygen atom of Leu-12 in domain A (Fig. 7.6e) and of Leu-71 in domain B (not shown), and the O3 atom of the G1 pyranose ring is hydrogen-bonded to the side-chain hydroxyl group of Thr-39 in domain A (Fig. 7.6e) and of Thr-98 (not shown) in domain B. The reducing hydroxyl group (O1) from the G1 unit protrudes away from the binding site and remains solvent-accessible such that an N-linked Asn would not disrupt binding.

For the G2 sugar unit, the acetyl oxygen and O6 atoms of the subsequent GlcNAc2 unit are hydrogen-bonded to the backbone amide groups of Asn-15 and Thr-39 in domain A (Fig. 7.6e) and of Asn-74 and Thr-98 in domain B (not shown), respectively. In addition, the pyranose ring of the G2 unit is stacked on top of the six-membered ring of Trp-37 in domain A (Fig. 7.6e) and of Trp-96 in domain B (not shown). For the M1 unit, the pyranose ring of M1 is stacked on top of the five membered ring of Trp-37 in domain A (Fig. 7.6e) and of Trp-96 in domain B (not shown) while the O4 atom of M1 forms a hydrogen bond with the side-chain amide group of Gln-36 in domain A (Fig. 7.6e) or Gln-95 in domain B (not shown), stabilizing the protein interaction with the M1-G2-G1 reducing end trisaccharide core.

The branched mannose residues, Man2 and Man5, extend up and away from the binding cleft. They, however, still form hydrogen bonds with MVL. In the M2 sugar unit, a hydrogen bond is formed between the side chain of Thr-38 and the O6 hydroxyl group of M2 in domain A. Since the corresponding residue in domain B is Arg-97, no equivalent hydrogen bond is possible. Finally, for the M5 sugar unit, a hydrogen bond is present between the O4 atom and the side chain of Gln-36 in domain A (Fig. 7.6e) and of Gln-95 in domain B, as well as a water-bridged hydrogen bond between the O5 atom and the side-chain amide of Asn-15 in domain A (Fig. 7.6e) and of Asn-74.

In summary, MVL uniquely recognizes the  $\text{Man}\alpha(1 \rightarrow 6)\text{Man}\beta(1 \rightarrow 4)\text{-GlcNAc}\beta(1 \rightarrow 4)\text{GlcNAc}$  tetrasaccharide core structure of high mannose oligosaccharides (Bewley et al. 2004; Williams et al. 2005) and inhibits HIV-1 cell fusion with an  $\text{IC}_{50}$  value of 30–40 nM, depending on the HIV-1 strain (Bewley et al. 2004). Similar to OAA or SVN, MVL does not precipitate when it interacts with Man9, indicating that no cross-linking occurs (Fig. 7.6f) as was observed with CV-N or GRFT (Figs. 7.1f and 7.4e). However, the question still remains as to whether all four binding sites are required for antiviral potency since no mutagenesis studies deleting individual binding sites on MVL have been reported.

## 7.8 Actinohivin (AH)

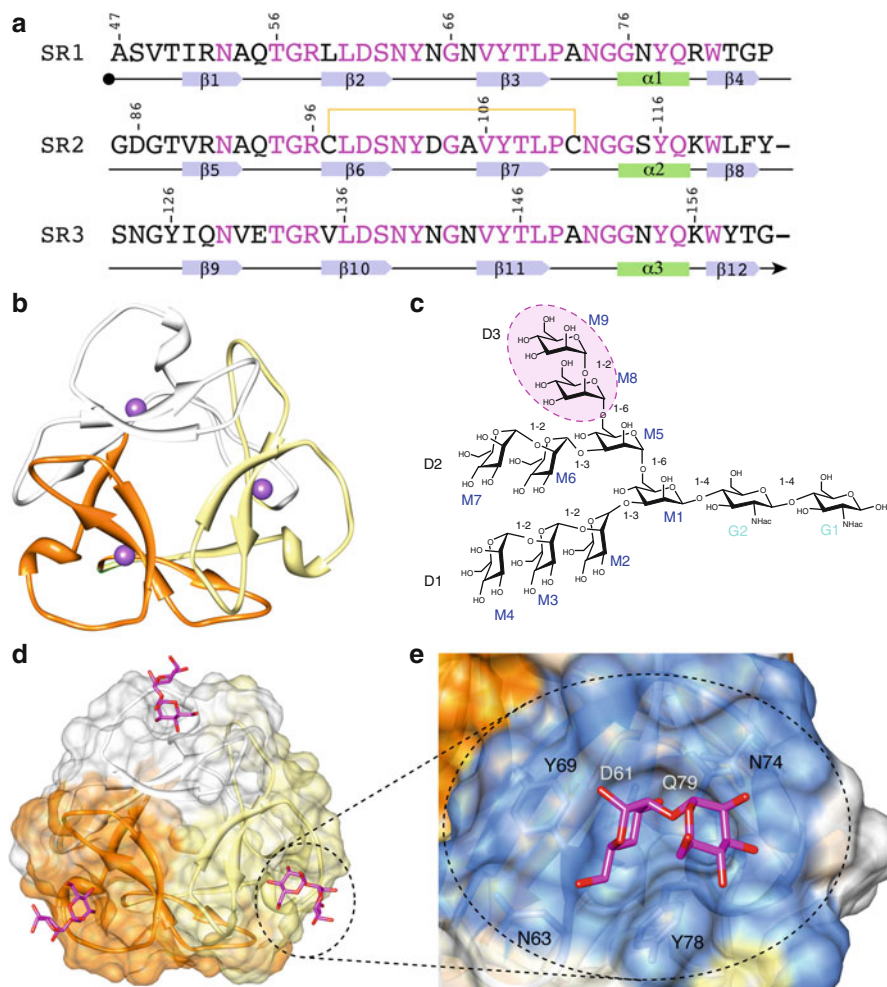
AH is an anti-HIV lectin from the actinomycete *Longisporum Albid* (actinomycete strain K97-0003) (Inokoshi et al. 2001; Chiba et al. 2001). It contains a 114-residue chain of molecular mass ~12.5 kDa and its amino acid sequence was determined by Edman degradation (Chiba et al. 2001). Like GRFT, three highly conserved internal sequence repeats are present, comprising residues 1–38, 39–77, and 78–114 designated as SR1, SR2, and SR3, respectively (Fig. 7.7a) (Chiba et al. 2001).

The crystal structure of AH was determined by X-ray crystallography at 1.2 Å resolution (Tanaka et al. 2009). Two protein molecules are present per asymmetric unit with both chains adopting very similar conformations with a pairwise C $\alpha$  atom r.m.s.d value of 0.17 Å (Tanaka et al. 2009). The overall structure of AH is composed of three domains that are very similar to each other (Fig. 7.7b). Each domain contains a four-stranded antiparallel  $\beta$ -sheet and a short  $3_{10}$  helix (Tanaka et al. 2009). Three  $\beta$ -strands ( $\beta_1$ ,  $\beta_2$ , and  $\beta_3$ ), the first long loop, the first short  $3_{10}$ -helix, and the last  $\beta$ -strand ( $\beta_{12}$ ) form domain 1. The second and third domains are then formed by a stretch of continuous amino acids that make up the next four  $\beta$ -strands ( $\beta_4$ ,  $\beta_5$ ,  $\beta_6$ , and  $\beta_7$ ), the second long loop and the second short  $3_{10}$ -helix for domain 2 and domain 3 contains the next four  $\beta$ -strands ( $\beta_8$ ,  $\beta_9$ ,  $\beta_{10}$ , and  $\beta_{11}$ ), the third long loop, and the third short  $3_{10}$ -helix. AH can be regarded as a cyclic assembly given that the last  $\beta$ -strand is a component of the first domain. All three domains exhibit pseudo threefold symmetry, in which the three  $\beta$ -sheets form a triangular barrel. Note that, like CV-N, where each domain is formed by strand exchange between the two sequence repeats, each domain of AH also exhibits strand exchange between two sequence repeats within AH (domains 1, 2, and 3 being composed of SR3 and SR1, SR1 and SR2, and SR2 and SR3, respectively) (Fig. 7.7a).

The carbohydrate specificity of AH was established using frontal affinity chromatography (FAC) (Tanaka et al. 2009). Various types of glycans were tested with an AH-immobilized column. Relatively tight affinities were obtained for Man<sub>9</sub> and Man<sub>8</sub> with  $K_d$  values of  $2.9 \times 10^{-4}$  and  $2.1 \times 10^{-4}$  M, respectively, and it was shown that AH recognizes the Man $\alpha$ (1–2)Man disaccharide epitope of these high mannose oligosaccharides (Tanaka et al. 2009). The Man $\alpha$ (1–2)Man epitope located in the D1 arm is most effectively bound, followed by the Man $\alpha$ (1–2)Man epitope in the D3 and D2 arms. It was also shown that the presence of both Man $\alpha$ (1–2)Man epitopes on the D1 and D3 arms in Man-8/9 are essential for the high affinity binding to AH (Fig. 7.7c) (Tanaka et al. 2009), similar to the findings with CV-N (Fig. 7.1d) (Bewley and Otero-Quintero 2001; Botos et al. 2002). A binding study between AH and glycosylated gp120 by surface plasmon resonance resulted in similar conclusions (Hoorelbeke et al. 2010).

Structures of AH alone and in complex with Man $\alpha$ (1–2)Man disaccharides were determined (Hoque et al. 2012). A superposition between the apo and Man $\alpha$ (1–2)Man-bound AH yielded backbone and all heavy atom r.m.s.d. values of  $\sim 0.3$  and  $\sim 0.8$  Å, respectively, suggesting that no large conformation changes occurred upon Man $\alpha$ (1–2)Man binding (Hoque et al. 2012). Three Man $\alpha$ (1–2)Man molecules found bound to an AH molecule (Fig. 7.7d), each in a stacked conformation and buried in a shallow pocket made up by the three  $\beta$ -strands and a long loop (Fig. 7.7d) (Hoque et al. 2012), similar to what was previously observed in the complex between CV-N and Man $\alpha$ (1–2)Man (Bewley 2001).

The detailed atomic interactions in each binding pocket in AH are depicted in Fig. 7.7e. Given the pseudo-symmetry of protein, identical conformations of Man $\alpha$ (1–2)Man are seen in all three domains of AH. Therefore, only the interactions between Man $\alpha$ (1–2)Man in the first domain of AH are described here. The first mannose residue (M1) of Man $\alpha$ (1–2)Man sits at the edge of the rim and does not form



**Fig. 7.7** Structure and carbohydrate specificity of AH. **(a)** Sequence alignment of the three repeats of AH. **(b)** The overall structure of AH, drawn in ribbon representation. **(c)** Chemical structure of Man-9. **(d)** Surface representation illustrating the interactions between AH and the Man $\alpha$ 1–2Man $\alpha$  disaccharides in the three available binding sites. **(e)** Detailed interactions between AH and Man $\alpha$ 1–2Man $\alpha$  disaccharide in one of the three available binding sites

hydrogen bonds to the protein (Fig. 7.7d, e) (Hoque et al. 2012), while the second mannose residue (M2) is buried and interacts closely with the protein via four hydrogen bonds (Hoque et al. 2012). These hydrogen bonds are between one of the two carboxyl O atoms of the Asp15 side chain and the hydroxyl group attached to the C3 atom of M2 (2.5 Å) and between the other Asp15 carboxyl O and the hydroxyl group attached to the C4 atom of M2 (2.7 Å). Hydrogen bonds are also present between the amino nitrogen of the Asn28 side chain and the hydroxyl group attached to the C3

atom of M2 (2.9 Å) and between the hydroxyl group of the Tyr23 side chain and the hydroxyl group attached to the C4 atom of M2 (2.8 Å). In addition, the Tyr32 side chain is wedged between Man $\alpha$ (1–2)Man, engaged in hydrophobic contacts with the C5 and C6 atoms.

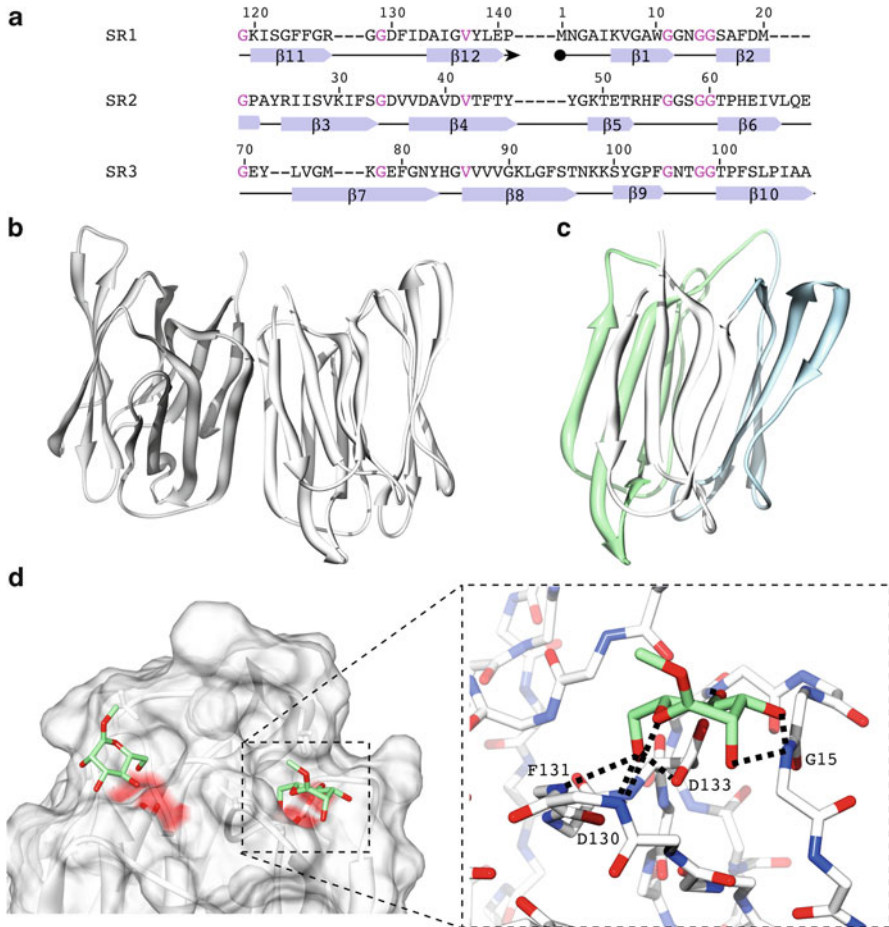
Like for the other antiviral lectins described above, the antiviral activity of AH is associated with binding to gp120-attached high mannose oligosaccharides (Chiba et al. 2001; Takahashi et al. 2005, 2011). Like seen with CV-N, it is very likely that AH also induces cross-linking, given that AH has three binding sites and recognizes two epitopes on the D1 or D3 arms of Man-8/9. With an IC<sub>50</sub> of 60–700 nM and relatively low toxicity to MT-4 cells, this lectin warrants further investigation and possible development as a microbicide.

## 7.9 Banana Lectin (BanLec)

Banlec is a lectin isolated from the ripened fruit of the banana, *Musa acuminata* or *Musa paradisiac* (Koshte et al. 1990; Peumans et al. 2000). It is a ~30 kDa homodimeric lectin with each monomer containing 141 residues. Like in GRFT, three internal amino acid sequence repeats are present, comprising residues 1–20 and 119–141 as SR1, 21–69 as SR2, and 70–118 as SR3 (Fig. 7.8a) (Singh et al. 2005). A conserved GxxGG motif is also present in a distinct loop region (Fig. 7.8a).

The crystal structure of BanLec was solved at ~2.5 Å resolution for the carbohydrate-free protein (Singh et al. 2004). It crystallized with two molecules in the asymmetric unit (Fig. 7.8b), consistent with the solution dimer state (Peumans et al. 2000). The monomer contains three four-stranded antiparallel  $\beta$ -sheets shaped like a prism with pseudo threefold symmetry, similar to other jacalin-like lectins (Fig. 7.8c). The first SR, consisting of strands 1, 12, 11, and 2, makes up one face of the prism in a pseudo Greek key motif. The second and third faces of the prism are also Greek key motifs containing strands 5, 4, 3, and 6 and strands 9, 8, 7, and 10, respectively (Fig. 7.8a, c).

Unlike GRFT or AH that possess 3 SRs and 3 carbohydrate-binding sites in each monomer, BanLec has only two primary binding sites based on the crystal structure of BanLec in complex with methyl- $\alpha$ -D-mannoside (Singh et al. 2005). In the first site, the interactions involve hydrogen bonds between the O3 hydroxyl group and the amide nitrogen of G15, between the O4 hydroxyl group and the amide nitrogen of G15, between the O4 hydroxyl group and the carboxyl group of D133, between the O5 hydroxyl group and the amide nitrogen of D130, between the O6 hydroxyl group and the amide nitrogen of D130, between the O6 hydroxyl group and the amide nitrogen of F131, between the O6 hydroxyl group and the carbonyl oxygen of F131, and between the O6 hydroxyl group and the carboxyl group of D133. In the second site, hydrogen bonds are formed between the O3 hydroxyl group and the amide nitrogen of G60, between the O4 hydroxyl group and the carboxyl group of D38, between the O5 hydroxyl group and the amide nitrogen of D35, between the O6 hydroxyl group and the amide nitrogen of G34, between the O6 hydroxyl group



**Fig. 7.8** Structure and carbohydrate specificity of BanLec. (a) Sequence alignment of the three repeats of BanLec. (b) The overall structure of BanLec homodimer, drawn in ribbon representation. (c) The structure of a BanLec monomer showing the presence of three  $\beta$ -strand sheets, drawn in ribbon representation and colored in *light grey*, *light blue*, and *light green*, corresponding to SR1, SR2, and SR3, respectively. (d) Surface representation and detailed interactions between BanLec monomer and methyl- $\alpha$ -D-mannoside monosaccharides in the two available binding sites

and the amide nitrogen of V36, between the O6 hydroxyl group and the carbonyl oxygen of V36, and between the O6 hydroxyl group and the carboxyl group of D38. In addition to the interacting residues specified above, V86 makes a hydrophobic contact with the sugar in the first site. The presence of two binding sites in BanLec was confirmed by the structure of BanLec from a different species bound to two different ligands (Meagher et al. 2005). In this report, the authors also suggest that there is a preference in the primary binding site for the reducing ends of the disaccharides (Meagher et al. 2005).



The anti-HIV activity of BanLec arises, as with the other lectins mentioned above, through interactions with high mannose sugars present on HIV gp120. Using temperature-sensitive viral entry studies (Swanson et al. 2010), it was shown that BanLec blocks HIV-1 cellular entry with comparable potency as other anti-HIV lectins, such as Griffithsin. Based on these results, BanLec also has potential for use as an antiviral microbicide.

## 7.10 Other Antiviral Lectins

In general, lectins are capable of reversibly binding to carbohydrate moieties of complex sugars without altering their covalent structure. They are found in a wide variety of species, including prokaryotes, sea corals, algae, fungi, higher plants, invertebrates, and vertebrates. Their functions involve in many biological processes, including host–pathogen interactions, cell–cell communication, induction of apoptosis, cancer metastasis and differentiation, targeting of cells, in addition to binding carbohydrates (Sharon 2008; Sharon and Lis 1989, 2004). While this chapter mainly focuses on the structures and carbohydrate specificities of nine lectins that are well characterized, any lectins that can interact with oligosaccharides on gp120 may possess antiviral activity and can potentially be developed as an HIV microbicide.

Examples of such lectins are proteins from the sea coral *Gerardia savaglia* (named *Gerardia savaglia* agglutinin; GSA) (Kljajic et al. 1987), mannose-specific plant lectins from the Amaryllidaceae family, such as *Narcissus pseudonarcissus* agglutinin (NPA) (Balzarini et al. 1991), *Galanthus nivalis* agglutinin (GNA) (Hester and Wright 1996; Wright and Hester 1996) and the *Hippeastrum* hybrid agglutinin (HHA) (Van Damme et al. 1988), *Listera ovata* agglutinin (LOA) from the Orchidaceae family (Van Damme et al. 1987, 1994), *Epipactis helleborine* agglutinin (EHA) (Van Damme et al. 1987, 1994), *Cymbidium* hybrid agglutinin (CHA) (Van Damme et al. 1987, 1994), *Allium porrum* agglutinin (APA) from the Alliaceae family (Balzarini et al. 1992), plant lectins derived from the Moraceae dicot family, such as jacalin (Bourne et al. 2002) or from the Cecropiaceae dicot family, such as *Myrianthus holstii* lectin (Charan et al. 2000), concanavalin A (ConA) derived from the jack bean (*Canavalia ensiformis*) (Edelman et al. 1972), *Urtica dioica* agglutinin (UDA) derived from rhizomes (Balzarini et al. 1992), or mammalian lectins such as DC-SIGN (Geijtenbeek et al. 2000; Feinberg et al. 2001).

Among these additional lectins, some require a metal ion for sugar binding, such as the C-type lectins that use a  $\text{Ca}^{2+}$  atom. Additionally, some lectins recognize a particular glycosidic linkage between the different carbohydrate units. For example, UDA and monocot lectins interact with terminal carbohydrates (Mannose or GlcNAc), while others preferentially bind to oligosaccharide cores, such as DC-SIGN or ConA. Similarly, some lectins interact with a single terminal carbohydrate unit, while others engage with several units of the oligosaccharide. Since the carbohydrate specificity for some of these lectins has not yet been elucidated, future



structural studies will undoubtedly yield important new insights that may further our understanding of carbohydrate–protein interactions.

As a general rule, the affinities of these lectins for monosaccharides are relatively weak, exhibiting  $K_d$  values in the millimolar range. In addition, for some lectins, binding is promiscuous, and multiple specificities for Man, GlcNAc, and Fuc on the same site have been observed. This can be explained by the limited ways in which the few hydroxyl groups, the only common feature of these carbohydrates, can be positioned in space (Drickamer 1995). Even slight changes in the sugar-binding sites can dramatically change their binding selectivity. Thus, one may wonder where the selectivity in these proteins originates from. To some degree, the abundance of extended, sugar-binding sites that can accommodate more than one type of carbohydrate and multiple binding sites may play a role. In an extended binding site, several sugar residues of an oligosaccharide can be positioned within the protein, thereby significantly increasing the affinity, when compared to the binding of an individual monosaccharide. Similarly, multiple binding sites per molecule can compensate for the weak affinity of individual sites and can be involved in binding terminal sugars of multi-antennary oligosaccharides, thereby creating multivalent contacts.

At this juncture it seems prudent to ask the question why these lectins are not considered as candidates in the development of protein-based microbicides. The reasons are not altogether clear, but some possibilities can be suggested: First, these “orphan” lectins are generally fairly large, especially if they dimerize or multimerize; they also possess relatively weak affinity towards high mannose oligosaccharides, compared to the most potent antiviral lectins that were discussed above. Third, and most importantly, most of these lectins interfere with normal cellular activities, and, therefore exhibit mild to severe toxicity, making them undesirable.

## 7.11 Conclusion and Future Strategy

This chapter aimed to provide a comprehensive review of the structures, the distinct modes of glycan recognition, and the atomic details of the protein–glycan interactions for nine different antiviral lectins. These lectins are candidates for development as protein-based microbicides. The sole common feature of these eight proteins is the presence of internal sequence repeats. For example, CV-N, OAA, PFA, SVN, and MVL possess two sequence repeats, while GRFT, AH, and BanLec contain three sequence repeats and MBHA encompasses four sequence repeats. Interestingly, the number of sequence repeat often corresponds to the number of binding sites and domains that are observed in each lectin. However, the sequence repeats are not equivalent to the domains, since the individual domains in these lectins involve strand exchange between sequence repeats. The extensive sequence similarity of the repeats, ranging from ~32 % in CV-N to ~86 % in OAA, goes hand-in-hand with structural similarity between the individual domains, with the exception of GRFT, where the three repeats result in three binding sites, but not three individual domains.

Everything else about these lectins is unique. Some lectins exhibit domain swapping and multimerization, while others exist solely as monomers. Although CV-N and GRFT can domain swap, they are either found naturally as a monomer in solution or can be engineered into a monomer, as in the case of GRFT. The nature of domain swapping in CV-N and GRFT, however, is distinctly different; in CV-N, half of the protein chain swaps while in GRFT only the first two  $\beta$ -strands, out of twelve in total, are involved in the swap. MVL and BanLec, on the other hand, do not domain swap, but instead are homodimerizing. The remaining five lectins (OAA, PFA, MBHA, SVN, and AH) are all monomers. For OAA and PFA no indication of dimerization was noted, even at very high protein concentration ( $\sim 200$  mg/mL).

As mentioned above, the number of sequence repeats directly translates into the number of sugar-binding sites. Lectins with two sequence repeats contain two binding sites, such as monomeric CV-N, OAA, PFA, and SVN. mGRFT and AH possess three sequence repeats and therefore three sugar-binding sites. MBHA and the CV-N domain-swapped dimer contain four sequence repeats and, therefore, four sugar-binding sites. Finally, the lectin with the highest number of binding sites is GRFT, which has a triple sequence repeat and in the domain-swapped dimer consequently six binding sites. An exception to this rule is BanLec, which possesses three sequence repeats but only two sugar-binding sites.

The most interesting difference between all these lectins is their carbohydrate specificity: each lectin recognizes a distinct epitope of Man-8/9. For example, CV-N binds to the terminal  $\text{Man}\alpha(1\rightarrow 2)\text{Man}$  di- and tri-mannosides in the D1 or D3 arms (Fig. 7.1d). Interestingly, while AH also recognizes this epitope, the conformation of the bound  $\text{Man}\alpha(1\rightarrow 2)\text{Man}$  disaccharide is different in the two lectins (Figs. 7.1e and 7.7d); Both mannoses are tightly bound, forming a total of nine hydrogen bonds with CV-N, while only one of the two mannoses is buried in AH, with a total of four hydrogen bonds. This may explain the higher potency of CV-N, compared to AH. The specificity of all the other lectins is significantly different. For example, all OAAH members recognize the branched mannose core of Man-8/9 (Fig. 7.2c), GRFT recognizes any of the terminal mannoses on the D1, D2, or D3 arms of Man-8/9 (Fig. 7.4c), SVN binds only to the  $\text{Man}\alpha(1\rightarrow 2)\text{Man}\alpha(1\rightarrow 6)\text{Man}\alpha(1\rightarrow 6)\text{Man}$  tetrasaccharide (Fig. 7.5c), and MVL to the  $\text{Man}\alpha(1\rightarrow 6)\text{Man}\beta(1\rightarrow 4)\text{GlcNAc}\beta(1\rightarrow 4)\text{GlcNAc}$  tetrasaccharide (Fig. 7.6c). While the X-ray structure of BanLec complexed with a methyl- $\alpha$ -D-mannoside has been solved, it is not clear whether its mode of Man-9 binding is similar to that of GRFT.

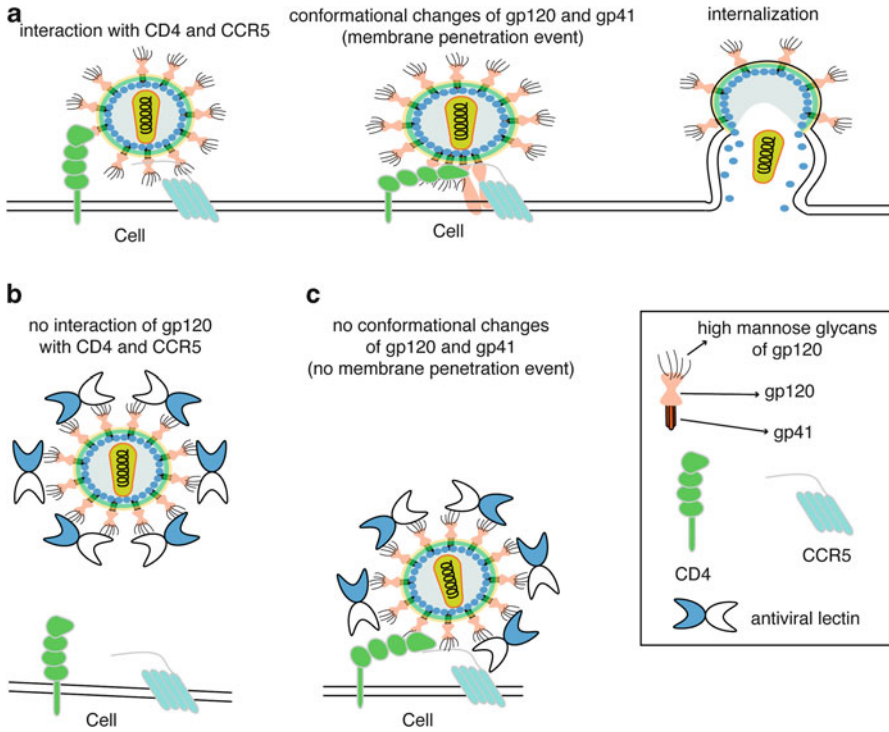
The difference in their carbohydrate specificity can be related to the number of epitopes available on Man-8/9 for each lectin. Thus, a single epitope is available for OAA, PFA, MBHA, SVN, and MVL, two epitopes for CV-N and AH, and three epitopes for GRFT.

There also appears to be a correlation between the number of binding sites on the protein, the number of epitopes recognized on Man-8/9 and antiviral potency. GRFT and BanLec are active against HIV-1 at picomolar concentrations, thus more potent than CV-N. CV-N, which is active at sub-nanomolar concentrations, is more potent than OAA, PFA, MBHA, SVN, MVL, and AH. The latter all display  $\text{IC}_{50}$  values in the low to medium nanomolar ranges. It appears that the number of bound epitopes

on the high mannose oligosaccharides is more critical for potency than the number of binding sites on the protein. This is derived from the comparison between monomeric and dimeric CVN, or between OAA (or PFA) and MBHA, with two and four binding sites, respectively. In both cases, no or very little difference in antiviral activity was observed. On the other hand, comparing CV-N and OAA, both of which possess two binding sites on the protein and interact with two sugar epitopes in the case of CV-N and only one in the case of OAA reveals that CV-N exhibits higher potency than OAA. Likewise, CV-N is also more potent than the other lectins, such as PFA and SVN, each with two binding sites but interacting with only one sugar epitope, or MBHA and MVL, each with four binding sites but again, contacting only one epitope. Comparing CV-N and AH, both of which can recognize the same number epitopes on Man-8/9, with AH possessing three binding sites and CV-N only two, CV-N is the more potent of the two.

In summary, the intrinsic antiviral activities of CV-N, OAA, PFA, MBHA, GRFT, SVN, MVL, AH, and BanLec render all of these lectins promising candidates for microbicide development, especially as components in antiviral preparations that can be applied topically. They specifically target the high mannose sugars that decorate the major envelope glycoprotein gp120 of HIV (Fig. 7.9). Indeed, preliminary studies on CV-N have been very promising, with results showing that this lectin was highly efficient in preventing infection by chimeric SIV/HIV-1 viruses in macaques, when delivered by either vaginal or rectal routes (Tsai et al. 2003, 2004). Safety concerns, however, remain as lectins can induce mitogenic activity in PBMC cultures, notably when exposed over a period of several days (Huskens et al. 2008; Buffa et al. 2009). Interestingly, attachment of CV-N to polyethylene glycol (PEG) polymer chains appears to be effective in reducing mitogenic activity in vitro (Zappe et al. 2008). GRFT, on the other hand, seems to be devoid of mitogenic activity (O'Keefe et al. 2009). The other lectins are presently being investigated for such effects. Given the differences in binding, toxicity, and anti-HIV activity, it is prudent to consider more than one lectin in the development of a lectin-based anti-HIV microbicide. This surely should increase the chance for success in this endeavor.

What then would be the best way for delivering these lectins? A number of modes are being considered. The most straightforward is to include them as components in microbicidal preparations for vaginal or rectal application, or possibly as components in multifunctional contraceptive gels. Another method of delivery of lectins, potentially even more promising, is their in situ expression by modified bacteria, similar to those naturally found in the vagina. This concept has already been explored for CV-N, using the human commensal bacterium *Streptococcus gordonii* (Giomarelli et al. 2002) or the engineered strain of the natural vaginal *Lactobacillus jensenii* (Liu et al. 2006). Therefore, we are confident that the other lectins discussed above will also be tested in similar ways. They should be complementary to CV-N, given their differences in sugar epitope specificities. Further developments of these antiviral lectins as pharmacological agents will contribute to find novel avenues to interfere with HIV transmission, still an urgent need to prevent suffering and deaths of vulnerable populations worldwide.



**Fig. 7.9** Mode of HIV inhibition by antiviral lectins. **(a)** In the absence of antiviral lectins, the interaction between gp120 and CD4 introduces a conformational change that allows the fusion peptide of gp41 to penetrate the cell membrane, leading to viral-cell membrane fusion and HIV capsid deposition into the cell. **(b)** and **(c)** In the presence of antiviral lectins, they bind to the high mannose glycans on gp120/41, preventing the required conformational change, thereby blocking infection

## References

- Adams EW, Ratner DM, Bokesch HR, McMahon JB, O'Keefe BR, Seeberger PH (2004) Oligosaccharide and glycoprotein microarrays as tools in HIV glycobiology; glycan-dependent gp120/protein interactions. *Chem Biol* 11(6):875–881
- Balzarini J (2006) Inhibition of HIV entry by carbohydrate-binding proteins. *Antiviral Res* 71(2–3):237–247
- Balzarini J, Schols D, Neyts J, Van Damme E, Peumans W, De Clercq E (1991) Alpha-(1–3)- and alpha-(1–6)-D-mannose-specific plant lectins are markedly inhibitory to human immunodeficiency virus and cytomegalovirus infections in vitro. *Antimicrob Agents Chemother* 35(3): 410–416
- Balzarini J, Neyts J, Schols D, Hosoya M, Van Damme E, Peumans W, De Clercq E (1992) The mannose-specific plant lectins from *Cymbidium* hybrid and *Epipactis helleborine* and the (N-acetylglucosamine)*n*-specific plant lectin from *Urtica dioica* are potent and selective inhibitors of human immunodeficiency virus and cytomegalovirus replication in vitro. *Antiviral Res* 18:191–207
- Balzarini J, Van Herrewege Y, Vermeire K, Vanham G, Schols D (2007) Carbohydrate-binding agents efficiently prevent dendritic cell-specific intercellular adhesion molecule-3-grabbing

- nonintegrin (DC-SIGN)-directed HIV-1 transmission to T lymphocytes. *Mol Pharmacol* 71(1):3–11
- Barrientos LG, Gronenborn AM (2002) The domain-swapped dimer of cyanovirin-N contains two sets of oligosaccharide binding sites in solution. *Biochem Biophys Res Commun* 298(4):598–602
- Barrientos LG, Gronenborn AM (2005) The highly specific carbohydrate-binding protein cyanovirin-N: structure, anti-HIV/Ebola activity and possibilities for therapy. *Mini Rev Med Chem* 5(1):21–31
- Barrientos LG, Louis JM, Botos I, Mori T, Han Z, O'Keefe BR, Boyd MR, Wlodawer A, Gronenborn AM (2002) The domain-swapped dimer of cyanovirin-N is in a metastable folded state: reconciliation of X-ray and NMR structures. *Structure* 10(5):673–686
- Barrientos LG, O'Keefe BR, Bray M, Sanchez A, Gronenborn AM, Boyd MR (2003) Cyanovirin-N binds to the viral surface glycoprotein, GP1,2 and inhibits infectivity of Ebola virus. *Antiviral Res* 58(1):47–56
- Barrientos LG, Lasala F, Delgado R, Sanchez A, Gronenborn AM (2004) Flipping the switch from monomeric to dimeric CV-N has little effect on antiviral activity. *Structure* 12(10):1799–1807
- Bewley CA (2001) Solution structure of a cyanovirin-N:Man alpha 1-2Man alpha complex: structural basis for high-affinity carbohydrate-mediated binding to gp120. *Structure* 9(10):931–940
- Bewley CA, Otero-Quintero S (2001) The potent anti-HIV protein cyanovirin-N contains two novel carbohydrate binding sites that selectively bind to Man(8) D1D3 and Man(9) with nanomolar affinity: implications for binding to the HIV envelope protein gp120. *J Am Chem Soc* 123(17):3892–3902
- Bewley CA, Gustafson KR, Boyd MR, Covell DG, Bax A, Clore GM, Gronenborn AM (1998) Solution structure of cyanovirin-N, a potent HIV-inactivating protein. *Nat Struct Biol* 5(7):571–578
- Bewley CA, Kiyonaka S, Hamachi I (2002) Site-specific discrimination by cyanovirin-N for alpha-linked trisaccharides comprising the three arms of Man(8) and Man(9). *J Mol Biol* 322(4):881–889
- Bewley CA, Cai M, Ray S, Ghirlando R, Yamaguchi M, Muramoto K (2004) New carbohydrate specificity and HIV-1 fusion blocking activity of the cyanobacterial protein MVL: NMR, ITC and sedimentation equilibrium studies. *J Mol Biol* 339(4):901–914
- Bokesch HR, O'Keefe BR, McKee TC, Pannell LK, Patterson GM, Gardella RS, Sowder RC II, Turpin J, Watson K, Buckheit RW Jr, Boyd MR (2003) A potent novel anti-HIV protein from the cultured cyanobacterium *Scytonema varium*. *Biochemistry* 42(9):2578–2584
- Botos I, Wlodawer A (2005) Proteins that bind high-mannose sugars of the HIV envelope. *Prog Biophys Mol Biol* 88(2):233–282
- Botos I, O'Keefe BR, Shenoy SR, Cartner LK, Ratner DM, Seeberger PH, Boyd MR, Wlodawer A (2002) Structures of the complexes of a potent anti-HIV protein cyanovirin-N and high mannose oligosaccharides. *J Biol Chem* 277(37):34336–34342
- Bourne Y, Astoul CH, Zamboni V, Peumans WJ, Menu-Bouaouiche L, Van Damme EJ, Barre A, Rouge P (2002) Structural basis for the unusual carbohydrate-binding specificity of jacalin towards galactose and mannose. *Biochem J* 364:173–180
- Boyd MR, Gustafson KR, McMahan JB, Shoemaker RH, O'Keefe BR, Mori T, Gulakowski RJ, Wu L, Rivera MI, Laurencot CM, Currens MJ, Cardellina JH II, Buckheit RW Jr, Nara PL, Pannell LK, Sowder RC II, Henderson LE (1997) Discovery of cyanovirin-N, a novel human immunodeficiency virus-inactivating protein that binds viral surface envelope glycoprotein gp120: potential applications to microbicide development. *Antimicrob Agents Chemother* 41(7):1521–1530
- Buffa V, Stieh D, Mamhood N, Hu Q, Fletcher P, Shattock RJ (2009) Cyanovirin-N potently inhibits human immunodeficiency virus type 1 infection in cellular and cervical explant models. *J Gen Virol* 90(pt 1):234–243
- Chandra NR, Kumar N, Jeyakani J, Singh DD, Gowda SB, Prathima MN (2006) Lectindb: a plant lectin database. *Glycobiology* 16(10):938–946

- Charan RD, Munro MH, O'Keefe BR, Sowder RCH, McKee TC, Currens MJ, Pannell LK, Boyd MR (2000) Isolation and characterization of *Myrianthus holstii* lectin, a potent HIV-1 inhibitory protein from the plant *Myrianthus holstii*. *J Nat Prod* 63:1170–1174
- Chiba H, Inokoshi J, Okamoto M, Asanuma S, Matsuzaki K, Iwama M, Mizumoto K, Tanaka H, Oheda M, Fujita K, Nakashima H, Shinose M, Takahashi Y, Omura S (2001) Actinohivin, a novel anti-HIV protein from an actinomycete that inhibits syncytium formation: isolation, characterization, and biological activities. *Biochem Biophys Res Commun* 282(2):595–601
- Chirenje ZM, Marrazzo J, Parikh UM (2010) Antiretroviral-based HIV prevention strategies for women. *Expert Rev Anti Infect Ther* 8(10):1177–1186
- D'Cruz OJ, Uckun FM (2004) Clinical development of microbicides for the prevention of HIV infection. *Curr Pharm Des* 10(3):315–336
- Drickamer K (1995) Multiplicity of lectin-carbohydrate interactions. *Nat Struct Biol* 2(6):437–439
- Edelman GM, Cunningham BA, Reeke GN Jr, Becker JW, Waxdal MJ, Wang JL (1972) The covalent and three-dimensional structure of concanavalin A. *Proc Natl Acad Sci U S A* 69(9):2580–2584
- Feinberg H, Mitchell DA, Drickamer K, Weis WI (2001) Structural basis for selective recognition of oligosaccharides by DC-SIGN and DC-SIGNR. *Science* 294(5549):2163–2166
- Geijtenbeek TB, Kwon DS, Torensma R, van Vliet SJ, van Duinhoven GC, Middel J, Cornelissen IL, Nottet HS, KewalRamani VN, Littman DR, Figdor CG, van Kooyk Y (2000) DC-SIGN, a dendritic cell-specific HIV-1-binding protein that enhances trans-infection of T cells. *Cell* 100(5):587–597
- Giomarelli B, Provvedi R, Meacci F, Maggi T, Medagliani D, Pozzi G, Mori T, McMahon JB, Gardella R, Boyd MR (2002) The microbicide cyanovirin-N expressed on the surface of commensal bacterium *Streptococcus gordonii* captures HIV-1. *AIDS* 16(10):1351–1356
- Giomarelli B, Schumacher KM, Taylor TE, Sowder RC II, Hartley JL, McMahon JB, Mori T (2006) Recombinant production of anti-HIV protein, griffithsin, by auto-induction in a fermentor culture. *Protein Expr Purif* 47(1):194–202
- Gustafson KR, Sowder RC II, Henderson LE, Cardellina JH II, McMahon JB, Rajamani U, Pannell LK, Boyd MR (1997) Isolation, primary sequence determination, and disulfide bond structure of cyanovirin-N, an anti-HIV (human immunodeficiency virus) protein from the cyanobacterium *Nostoc ellipsosporum*. *Biochem Biophys Res Commun* 238(1):223–228
- Helle F, Wychowski C, Vu-Dac N, Gustafson KR, Voisset C, Dubuisson J (2006) Cyanovirin-N inhibits hepatitis C virus entry by binding to envelope protein glycans. *J Biol Chem* 281(35):25177–25183
- Hester G, Wright CS (1996) The mannose-specific bulb lectin from *Galanthus nivalis* (snowdrop) binds mono- and dimannosides at distinct sites. Structure analysis of refined complexes at 2.3 Å and 3.0 Å resolution. *J Mol Biol* 262(4):516–531
- Hladik F, Doncel GF (2010) Preventing mucosal HIV transmission with topical microbicides: challenges and opportunities. *Antiviral Res* 88(suppl 1):S3–S9
- Hoorelbeke B, Huskens D, Ferir G, Francois KO, Takahashi A, Van Laethem K, Schols D, Tanaka H, Balzarini J (2010) Actinohivin, a broadly neutralizing prokaryotic lectin, inhibits HIV-1 infection by specifically targeting high-mannose-type glycans on the gp120 envelope. *Antimicrob Agents Chemother* 54(8):3287–3301
- Hoque MM, Suzuki K, Tsunoda M, Jiang J, Zhang F, Takahashi A, Ohbayashi N, Zhang X, Tanaka H, Omura S, Takenaka A (2012) Structural insights into the specific anti-HIV property of actinohivin: structure of its complex with the alpha(1–2)mannobiose moiety of gp120. *Acta Crystallogr D Biol Crystallogr* 68(pt 12):1671–1679
- Huskens D, Vermeire K, Vandemeulebroucke E, Balzarini J, Schols D (2008) Safety concerns for the potential use of cyanovirin-N as a microbicidal anti-HIV agent. *Int J Biochem Cell Biol* 40(12):2802–2814
- Inokoshi J, Chiba H, Asanuma S, Takahashi A, Omura S, Tanaka H (2001) Molecular cloning of actinohivin, a novel anti-HIV protein from an actinomycete, and its expression in *Escherichia coli*. *Biochem Biophys Res Commun* 281(5):1261–1265



- IPM (International Partnership of Microbicides). <http://www.ipmglobal.org/our-work/ipm-product-pipeline/maraviroc>. Accessed 10 May 2013.
- Kljajic Z, Schroder HC, Rottmann M, Cuperlovic M, Movsesian M, Uhlenbruck G, Gasic M, Zahn RK, Muller WE (1987) A D-mannose-specific lectin from *Gerardia savaglia* that inhibits nucleocytoplasmic transport of mRNA. *Eur J Biochem* 169(1):97–104
- Koharudin LM, Gronenborn AM (2011) Structural basis of the anti-HIV activity of the cyanobacterial *Oscillatoria agardhii* agglutinin. *Structure* 19(8):1170–1181
- Koharudin LM, Furey W, Gronenborn AM (2011) Novel fold and carbohydrate specificity of the potent anti-HIV cyanobacterial lectin from *Oscillatoria agardhii*. *J Biol Chem* 286(2):1588–1597
- Koharudin LM, Kollipara S, Aiken C, Gronenborn AM (2012) Structural insights into the anti-HIV activity of the *Oscillatoria agardhii* agglutinin homolog lectin family. *J Biol Chem* 287(40):33796–33811
- Koshte VL, van Dijk W, van der Stelt ME, Aalberse RC (1990) Isolation and characterization of BanLec-I, a mannoside-binding lectin from *Musa paradisiac* (banana). *Biochem J* 272(3):721–726
- Liu X, Lagenaur LA, Simpson DA, Essenmacher KP, Frazier-Parker CL, Liu Y, Tsai D, Rao SS, Hamer DH, Parks TP, Lee PP, Xu Q (2006) Engineered vaginal *Lactobacillus* strain for mucosal delivery of the human immunodeficiency virus inhibitor cyanovirin-N. *Antimicrob Agents Chemother* 50(10):3250–3259
- Matei E, Furey W, Gronenborn AM (2008) Solution and crystal structures of a sugar binding site mutant of cyanovirin-N: no evidence of domain swapping. *Structure* 16(8):1183–1194
- McFeeters RL, Xiong C, O’Keefe BR, Bokesch HR, McMahan JB, Ratner DM, Castelli R, Seeberger PH, Byrd RA (2007) The novel fold of scytovirin reveals a new twist for antiviral entry inhibitors. *J Mol Biol* 369(2):451–461
- Meagher JL, Winter HC, Ezell P, Goldstein JJ, Stuckey JA (2005) Crystal structure of banana lectin reveals a novel second sugar binding site. *Glycobiology* 15(10):1033–1042
- Minces LR, McGowan I (2010) Advances in the development of microbicides for the prevention of HIV infection. *Curr Infect Dis Rep* 12(1):56–62
- Mori T, O’Keefe BR, Sowder RC II, Bringans S, Gardella R, Berg S, Cochran P, Turpin JA, Buckheit RW Jr, McMahan JB, Boyd MR (2005) Isolation and characterization of griffithsin, a novel HIV-inactivating protein, from the red alga *Griffithsia* sp. *J Biol Chem* 280(10):9345–9353
- Morrow KM, Hendrix C (2010) Clinical evaluation of microbicide formulations. *Antiviral Res* 88(suppl 1):S40–S46
- Moulaei T, Botos I, Ziolkowska NE, Bokesch HR, Krumpke LR, McKee TC, O’Keefe BR, Dauter Z, Wlodawer A (2007) Atomic-resolution crystal structure of the antiviral lectin scytovirin. *Protein Sci* 16(12):2756–2760
- Moulaei T, Shenoy SR, Giomarelli B, Thomas C, McMahan JB, Dauter Z, O’Keefe BR, Wlodawer A (2010a) Monomerization of viral entry inhibitor griffithsin elucidates the relationship between multivalent binding to carbohydrates and anti-HIV activity. *Structure* 18(9):1104–1115
- Moulaei T, Stuchlik O, Reed M, Yuan W, Pohl J, Lu W, Haugh-Krumpke L, O’Keefe BR, Wlodawer A (2010b) Topology of the disulfide bonds in the antiviral lectin scytovirin. *Protein Sci* 19(9):1649–1661
- NIAID (National Institute of Allergy and Infectious Diseases (NIAID))—Topical Microbicides. <http://www.niaid.nih.gov/topics/hivaids/research/prevention/pages/topicalmicrobicides.aspx>. Accessed 6 March 2013.
- NIAID (National Institute of Allergy and Infectious Diseases)—Antiretroviral Therapy to Reduce the Transmission of HIV). <http://www.niaid.nih.gov/topics/HIVAIDS/Research/prevention/Pages/art.aspx>. Accessed 6 March 2013.
- O’Keefe BR, Smee DF, Turpin JA, Saucedo CJ, Gustafson KR, Mori T, Blakeslee D, Buckheit R, Boyd MR (2003) Potent anti-influenza activity of cyanovirin-N and interactions with viral hemagglutinin. *Antimicrob Agents Chemother* 47(8):2518–2525

- O'Keefe BR, Vojdani F, Buffa V, Shattock RJ, Montefiori DC, Bakke J, Mirsalis J, d'Andrea AL, Hume SD, Bratcher B, Saucedo CJ, McMahon JB, Pogue GP, Palmer KE (2009) Scaleable manufacture of HIV-1 entry inhibitor griffithsin and validation of its safety and efficacy as a topical microbicide component. *Proc Natl Acad Sci U S A* 106(15):6099–6104
- Peumans WJ, Zhang W, Barre A, Houles Astoul C, Balint-Kurti PJ, Rovira P, Rouge P, May GD, Van Leuven F, Truffa-Bachi P, Van Damme EJ (2000) Fruit-specific lectins from banana and plantain. *Planta* 211(4):546–554
- Pirrone V, Thakkar-Rivera N, Jacobson JM, Wigdahl B, Krebs FC (2011) Combinatorial approaches to the prevention and treatment of HIV-1 infection. *Antimicrob Agents Chemother* 55(5):1831–1842
- Raval S, Gowda SB, Singh DD, Chandra NR (2004) A database analysis of jacalin-like lectins: sequence-structure-function relationships. *Glycobiology* 14(12):1247–1263
- Sato T, Hori K (2009) Cloning, expression, and characterization of a novel anti-HIV lectin from the cultured cyanobacterium, *Oscillatoria agardhii*. *Fish Sci* 75:743–753
- Sato Y, Murakami M, Miyazawa K, Hori K (2000) Purification and characterization of a novel lectin from a freshwater cyanobacterium, *Oscillatoria agardhii*. *Comp Biochem Physiol B Biochem Mol Biol* 125(2):169–177
- Sato Y, Okuyama S, Hori K (2007) Primary structure and carbohydrate binding specificity of a potent anti-HIV lectin isolated from the filamentous cyanobacterium *Oscillatoria agardhii*. *J Biol Chem* 282(15):11021–11029
- Sharon N (2008) Lectins: past, present and future. *Biochem Soc Trans* 36(pt 6):1457–1460
- Sharon N, Lis H (1989) Lectins as cell recognition molecules. *Science* 246(4927):227–234
- Sharon N, Lis H (2004) History of lectins: from hemagglutinins to biological recognition molecules. *Glycobiology* 14(11):53R–62R
- Shenoy SR, O'Keefe BR, Bolmstedt AJ, Cartner LK, Boyd MR (2001) Selective interactions of the human immunodeficiency virus-inactivating protein cyanovirin-N with high-mannose oligosaccharides on gp120 and other glycoproteins. *J Pharmacol Exp Ther* 297(2):704–710
- Shenoy SR, Barrientos LG, Ratner DM, O'Keefe BR, Seeberger PH, Gronenborn AM, Boyd MR (2002) Multisite and multivalent binding between cyanovirin-N and branched oligomannosides: calorimetric and NMR characterization. *Chem Biol* 9(10):1109–1118
- Singh DD, Saikrishnan K, Kumar P, Dauter Z, Sekar K, Surolia A, Vijayan M (2004) Purification, crystallization and preliminary X-ray structure analysis of the banana lectin from *Musa paradisiaca*. *Acta Crystallogr D Biol Crystallogr* 60(pt 11):2104–2106
- Singh DD, Saikrishnan K, Kumar P, Surolia A, Sekar K, Vijayan M (2005) Unusual sugar specificity of banana lectin from *Musa paradisiaca* and its probable evolutionary origin. *Crystallographic and modelling studies*. *Glycobiology* 15(10):1025–1032
- Swanson MD, Winter HC, Goldstein IJ, Markovitz DM (2010) A lectin isolated from bananas is a potent inhibitor of HIV replication. *J Biol Chem* 285(12):8646–8655
- Takahashi A, Inokoshi J, Chiba H, Omura S, Tanaka H (2005) Essential regions for antiviral activities of actinohivin, a sugar-binding anti-human immunodeficiency virus protein from an actinomycete. *Arch Biochem Biophys* 437(2):233–240
- Takahashi A, Inokoshi J, Hachiya A, Oka S, Omura S, Tanaka H (2011) The high mannose-type glycan binding lectin actinohivin: dimerization greatly improves anti-HIV activity. *J Antibiot (Tokyo)* 64(8):551–557
- Tanaka H, Chiba H, Inokoshi J, Kuno A, Sugai T, Takahashi A, Ito Y, Tsunoda M, Suzuki K, Takenaka A, Sekiguchi T, Umeyama H, Hirabayashi J, Omura S (2009) Mechanism by which the lectin actinohivin blocks HIV infection of target cells. *Proc Natl Acad Sci U S A* 106(37):15633–15638
- Team CPISAS (2010) Expanded safety and acceptability of the candidate vaginal microbicide Carraguard(R) in South Africa. *Contraception* 82(6):563–571
- Tsai CC, Emau P, Jiang Y, Tian B, Morton WR, Gustafson KR, Boyd MR (2003) Cyanovirin-N gel as a topical microbicide prevents rectal transmission of SHIV89.6P in macaques. *AIDS Res Hum Retroviruses* 19(7):535–541

- Tsai CC, Emau P, Jiang Y, Agy MB, Shattock RJ, Schmidt A, Morton WR, Gustafson KR, Boyd MR (2004) Cyanovirin-N inhibits AIDS virus infections in vaginal transmission models. *AIDS Res Hum Retroviruses* 20(1):11–18
- Turpin JA (2002) Considerations and development of topical microbicides to inhibit the sexual transmission of HIV. *Expert Opin Investig Drugs* 11(8):1077–1097
- Van Damme EJM, Allen AK, Peumans WJ (1987) Leaves of the orchid twayblade (*Listera ovata*) contain a mannose-specific lectin. *Plant Physiol* 85:566–569
- Van Damme EJM, Allen AK, Peumans WJ (1988) Related mannospecific lectins from different species of the family Amaryllidaceae. *Physiol Plant* 73:52–57
- Van Damme EJM, Smeets K, Torrekens S, Van Leuven F, Peumans WJ (1994) Characterization and molecular cloning of the mannosebinding lectins from the Orchidaceae species *Listera ovata*, *Epipactis helleborine* and *Cymbidium hybrid*. *Eur J Biochem* 221:769–777
- WHO (World Health Organization (WHO)—Microbicides). <http://www.who.int/hiv/topics/microbicides/microbicides/en>. Accessed 6 March 2013.
- WHO (World Health Organization—Data and Statistics). <http://www.who.int/hiv/data/en/index.html>. Accessed 6 March 2013.
- Williams DC Jr, Lee JY, Cai M, Bewley CA, Clore GM (2005) Crystal structures of the HIV-1 inhibitory cyanobacterial protein MVL free and bound to Man3GlcNAc2: structural basis for specificity and high-affinity binding to the core pentasaccharide from n-linked oligomannoside. *J Biol Chem* 280(32):29269–29276
- Wright CS, Hester G (1996) The 2.0 Å structure of a cross-linked complex between snowdrop lectin and a branched mannopentaose: evidence for two unique binding modes. *Structure* 4(11):1339–1352
- Xiong C, O’Keefe BR, Botos I, Wlodawer A, McMahon JB (2006a) Overexpression and purification of scytovirin, a potent, novel anti-HIV protein from the cultured cyanobacterium *Scytonema varium*. *Protein Expr Purif* 46(2):233–239
- Xiong C, O’Keefe BR, Byrd RA, McMahon JB (2006b) Potent anti-HIV activity of scytovirin domain 1 peptide. *Peptides* 27(7):1668–1675
- Xue J, Gao Y, Hoorelbeke B, Kagiampakis I, Zhao B, Demeler B, Balzarini J, Liwang PJ (2012) The role of individual carbohydrate-binding sites in the function of the potent anti-HIV lectin griffithsin. *Mol Pharm* 9(9):2613–2625
- Yamaguchi M, Ogawa T, Muramoto K, Kamio Y, Jimbo M, Kamiya H (1999) Isolation and characterization of a mannan-binding lectin from the freshwater cyanobacterium (blue-green algae) *Microcystis viridis*. *Biochem Biophys Res Commun* 265(3):703–708
- Yang F, Bewley CA, Louis JM, Gustafson KR, Boyd MR, Gronenborn AM, Clore GM, Wlodawer A (1999) Crystal structure of cyanovirin-N, a potent HIV-inactivating protein, shows unexpected domain swapping. *J Mol Biol* 288(3):403–412
- Zappe H, Snell ME, Bossard MJ (2008) PEGylation of cyanovirin-N, an entry inhibitor of HIV. *Adv Drug Deliv Rev* 60(1):79–87
- Ziolkowska NE, Wlodawer A (2006) Structural studies of algal lectins with anti-HIV activity. *Acta Biochim Pol* 53(4):617–626
- Ziolkowska NE, O’Keefe BR, Mori T, Zhu C, Giomarelli B, Vojdani F, Palmer KE, McMahon JB, Wlodawer A (2006) Domain-swapped structure of the potent antiviral protein griffithsin and its mode of carbohydrate binding. *Structure* 14(7):1127–1135
- Ziolkowska NE, Shenoy SR, O’Keefe BR, McMahon JB, Palmer KE, Dwek RA, Wormald MR, Wlodawer A (2007) Crystallographic, thermodynamic, and molecular modeling studies of the mode of binding of oligosaccharides to the potent antiviral protein griffithsin. *Proteins* 67(3):661–670