Domain-Swapped Structure of the Potent Antiviral Protein Griffithsin and Its Mode of Carbohydrate Binding

Natasza E. Ziółkowska,[1] Barry R. O'Keefe,[2]
Toshiyuki Mori,[2,5] Charles Zhu,[2,3] Barbara Giomarelli,[2]
Fakhrieh Vojdani,[4] Kenneth E. Palmer,[4,6]
James B. McMahon,[2] and Alexander Wlodawer[1,*]
[1] Protein Structure Section
Macromolecular Crystallography Laboratory
[2] Molecular Targets Development Program
Center for Cancer Research
[3] Werner H. Kirsten Student Internship Program
National Cancer Institute
NCI-Frederick
Frederick, Maryland 21702
[4] Large Scale Biology Corporation
3333 Vaca Valley Parkway
Vacaville, California 95688

Summary

The crystal structure of griffithsin, an antiviral lectin from the red alga *Griffithsia* sp., was solved and refined at 1.3 Å resolution for the free protein and 0.94 Å for a complex with mannose. Griffithsin molecules form a domain-swapped dimer, in which two β strands of one molecule complete a β prism consisting of three four-stranded sheets, with an approximate 3-fold axis, of another molecule. The structure of each monomer bears close resemblance to jacalin-related lectins, but its dimeric structure is unique. The structures of complexes of griffithsin with mannose and N-acetyl-glucosamine defined the locations of three almost identical carbohydrate binding sites on each monomer. We have also shown that griffithsin is a potent inhibitor of the coronavirus responsible for severe acute respiratory syndrome (SARS). Antiviral potency of griffithsin is likely due to the presence of multiple, similar sugar binding sites that provide redundant attachment points for complex carbohydrate molecules present on viral envelopes.

Introduction

A number of lectins, particularly those that bind high-mannose sugars, have been shown to exhibit significant activity against human immunodeficiency virus (HIV), as well as against a number of other viruses (Botos and Wlodawer, 2005). For these reasons, several such proteins, among them cyanovirin-N (Boyd et al., 1997) and scytovirin (Bokesch et al., 2003), have been under active development as topical antiviral therapeutics. Most recently, another potent antiviral lectin, griffithsin, was isolated from the red alga *Griffithsia* sp., collected from the waters off New Zealand (Mori et al., 2005). Griffithsin was

*Correspondence: wlodawer@ncifcrf.gov
[5] Present address: Takeda Pharmaceutical Company, 2-17-85 Juso-homachi, Yodogawa-ku, Osaka 532-8686, Japan.
[6] Present address: Intrucept Biomedicine LLC, 707 West Monte Vista Ave., Vacaville, CA 95688.

shown to inhibit the cytopathic effects of different isolates of HIV-1 at concentrations as low as 43 pM (Mori et al., 2005), and its binding to viral envelope glycoproteins was shown to be directly dependent on their glycosylation states. These properties make griffithsin a promising potential candidate for development as a future pharmaceutical agent.

Knowledge of the three-dimensional structure of putative therapeutic proteins is very helpful in their development, yet the amino acid sequence of griffithsin was initially reported to be unique, giving no clues to its structure (Mori et al., 2005; Giomarelli et al., 2006). More careful analysis of the sequence of griffithsin indicates that this protein belongs to a diverse family of β-prism-I (or jacalin-related) lectins (Raval et al., 2004) and shares structural characteristics with proteins such as jacalin (Aucouturier et al., 1987), artocarpin (Jeyaprakash et al., 2004), and heltuba (Bourne et al., 1999) (Figure 1A). However, sequence identity between griffithsin and any of these other proteins is less than 30% (Mori et al., 2005) (see also Figure 1A). Thus, structural details, particularly of the carbohydrate binding site(s), cannot be readily modeled, necessitating solution of an experimental structure. We report here the structures of griffithsin in three crystal forms, with and without bound carbohydrate molecules.

Since the antiviral activity of griffithsin against HIV-1 has already been proven (Mori et al., 2005), we have investigated other viral targets of this lectin. We have now shown that griffithsin is a potent inhibitor of the life cycle of the coronavirus that causes severe acute respiratory syndrome (SARS), a life-threatening disease that first manifested itself only a few years ago. Such broad antiviral activity makes this newly characterized lectin a particularly promising target for the development as a novel antiviral agent.

Results

Description of the Fold of Griffithsin

The structure of griffithsin was solved by single-wavelength anomalous diffraction (SAD) with selenomethionine (SeMet)-labeled protein produced in *E. coli* and was refined at high resolution with material grown in plants. Three different crystal forms of griffithsin were studied (Table 1). The fold of griffithsin corresponds to the well-known β-prism motif (Chothia and Murzin, 1993), found mainly in a variety of lectins but also in some other proteins (Shimizu and Morikawa, 1996). The motif consists of three repeats of an antiparallel four-stranded β sheet that form a triangular prism. However, unlike other lectins whose sequences are compared in Figure 1A, griffithsin forms an intimate dimer (Figure 2) in which the first two β strands of one chain (consisting of 18 residues located at its N terminus) are associated with ten strands of the other chain (and vice versa). That arrangement is unique among the family members that have been described so far and is seen in all crystal forms of griffithsin, despite the presence of a significant nonnative N-terminal extension in one of

A

```
GRIFFITHSIN   1  --------SLT-HRKFGGS---GGSPFSGL------SSIAVRSG--S-YLDAIIID--   35
ARTOCARPIN    1  -----ASQTIT-VGSWGGP---GGNGWDEGSYTG-IRQIELSYK--E-AIGSFSVIYD   44
HELTUBA       1  ---MAASDIAVQAGPWGGN---GGKRWLQTAHGGKITSIIKGG--T-CIFSIQFVYK   49
PARKIA        1  -----SLKGMISVGPWGGS---GGNYWSFKANHA-ITEIVIHVK--D-NIKSISFK-D   44
CALSEPA       2  -----VPMDTI-SGPWGNN---GGNFWSFRPVNK-INQIVISYGGGGNNPIALTFSST  50
BANANA        1  ------MNGAIKVGAWGGN---GGSAFDMGPAVR-IISVKIFSG--D-VVDGVDVTFT  44
JACALIN      B1  NEQSGISQTVI-VGPWGAKSS--GKAFDDGAFTG-IREINLSYNKET-AIGDFQVVYD  A33
AHL          B1  DENSGKSQTVI-VGPWGAKVS--GKAFDDGAFTG-IREINLSYNKET-AIGDFQVVYD  A33
MPA          B1  -GRNGKSQSII-VGPWGDRVTN-GVTFDDGAYTG-IREINFEYNSET-AIGGLRVTYD  A33

GRIFFITHSIN  36  -------G-VHHGG------SGGNLSPTFTF-GSGEYISNMTIRSG---------DY   68
ARTOCARPIN   45  LNGDPF-SGPKHTS-----KLPYKNV-KIELKFPDEFLESVSGYTGPFSALATPTPV  85
HELTUBA      50  DKDNIEYHSGKFGV------L-GDKAETITF-AEDEDITAISGTFGAYYHM----TV  93
PARKIA       45  ASGDIS---GTFGGKDPRENEKGDEK-KIKIHWPTEYLKSISGSYGDYNGV----LV  93
CALSEPA      51  KADGSKDT-ITVGGGGPD-SITGTEM--VNI-GTDEYLTGISGTFGIYLDN----NV  97
BANANA       45  YYGKTE-T-RHYGG-------SGGTPH-EIVL-QEGEYLVGMAGEVANYHGA----VV  97
JACALIN     A34  LNGSPY-VGQNHKSFI----TGFTPV-KISLDFPSEYIMEVSGYTGNVSGY----VV  A80
AHL         A34  LNGSPY-VGQNHSSFI----SGFTPV-KISLDFPSEYITEVSGYTGNVSGY----VV  A80
MPA         A34  LNGMPF-VAEDHKSFI----TGFKPV-KISLEFPSEYIVEVSGYVGKVEGY----TV  A80

GRIFFITHSIN  69  IDNISFETNMGRRFGPYGGS-GGSANTLSNV---KVIQINGSAGDYLDSLDIYYEQY  121
ARTOCARPIN   86  VRSLTFKTNKGRTFGPYGDE-EGTYFNLP-IENGLIVGFKGRTGDLLDAIGIHMSL-  149
HELTUBA      94  VTSLTFQTN-KKVYGPFGTV-ASSSFSLP-LTKGKFAGFFGNSGDVLDSIGGVVVP-  147
PARKIA       94  IRSLSFITN-LTTYGPFGSTSGGGESFSIP-IADSVVVGFHGRAGYYLDALGIFVQPV  149
CALSEPA      98  LRSITFTTN-LKAHGPYGQK-VGTPFSSANVVGNEIVGFLGRSGYYVDAIGTYNRHK  153
BANANA       98  LGKLGFSTN-KKAYGPFGNT-GGTPFSLPIAAG-KISGFFGRGGKFLDAIGVYLEP-  141
JACALIN     A80  VRSLTFKTN-KKTYGPYGVT-SGTPFNLP-IENGLIVGFKGSIGYWLDYFSMYLSL-  A133
AHL         A80  VRSLTFKTN-KKTYGPYGVT-SGTPFNLP-IENGLIVGFKGSIGYWMDYFSMYLSL-  A133
MPA         A80  IRSLTFKTN-KQTYGPYGVT-NGTPFSLP-IENGLIVGFKGSIGYWLDYFSIYLSL-  A133
```

B

```
Sheet 2    19   SSIAVRSGSYLDAIIIDGV           HHGGSGGNLSPTFTFGSGE   56
Sheet 3    57   YISNMTIRSGDYIDNISFETNM        GRRFGPYGGSGGSANTLSNVKV  100
Sheet 1   101   IQINGSAGDYLDSLDIYYEQY 121  1  SLTHRKFGGSGGSPFSGL    18
```

Figure 1. Comparison of the Sequences of Griffithsin and Related Lectins

(A) Structure-based sequence comparison of griffithsin and eight other lectins with a similar fold. Residues that are strictly conserved among all nine proteins are shown on red background, and those of a similar character are highlighted in yellow. Residues that are strictly conserved among the other lectins but not in griffithsin are highlighted in green in the former and magenta in the latter.

(B) Structure-based sequence alignment of the three β sheets (blades) in griffithsin. Conserved residues that interact directly with the bound carbohydrates are highlighted in red, whereas conserved residues that do not make such contacts are shown on magenta background.

them. Thus, griffithsin can be described as a domain-swapped dimer, although it is not clear whether a corresponding monomeric forms does (or even could) exist. A considerable distance between the residues A18 and B19 (10.5 Å between their Cα atoms) is much larger than a corresponding distance in the domain-swapped proteins that have been shown to exist in both monomeric and dimeric forms, such as, for example, another antiviral lectin, cyanovirn-N (Yang et al., 1999). In our previous study that utilized size-exclusion chromatography, griffithsin was shown to exist as a homodimer in solution (Giomarelli et al., 2006).

Griffithsin contains three strictly conserved repeats of a sequence GGSGG (Figure 1B). These repeats were

Table 1. Data Collection and Refinement Crystallographic Data

	Crystal			Complex with Mannose	
	Unliganded Trigonal SeMet	Unliganded Orthorhombic	Complex with N-Glucosamine	1	2
Data Collection					
Cell parameters	a = b = 66.87 Å, c = 68.20 Å; α = β = 90°, γ = 120°	a = 33.92 Å, b = 64.86 Å, c = 96.08 Å; α = β = γ = 90°	a = 33.83 Å, b = 64.45 Å, c = 96.18 Å; α = β = γ = 90°	a = 53.77 Å, b = 39.46 Å, c = 57.24 Å; α = γ = 90° β = 91.0°	a = 53.74 Å, b = 39.44 Å, c = 57.16 Å; α = γ = 90° β = 91.0°
Space group	P3$_2$21	P2$_1$2$_1$2$_1$	P2$_1$2$_1$2$_1$	P2$_1$	P2$_1$
Molecules/a.u.	1	2	2	2	2
Resolution (Å)	20–2.0	15–1.3	30–2.0	30–1.78	30–0.94
Total reflections	260,626	827,314	182,881	286,011	1,416,707
Unique reflections	22,075 (11,942)[a]	53,034	13,606	21,739	148,669
Completeness (%)[b]	95.7 (66.6)	99.8 (99.1)	92.8 (56.0)	93.3 (73.7)	95.5 (75.0)
Avg. I/σ	33.6 (1.7)	14.0 (2.5)	18.9 (2.2)	37.2 (6.2)	14.16 (1.5)
R$_{merge}$ (%)[c]	3.2 (34.9)	6.2 (38.2)	8.5 (36.1)	4.1 (17.6)	6.4 (55.0)
Refinement Statistics					
R (%)[d]	18.89	15.91	19.93	15.07	13.55
R$_{free}$ (%)[e]	25.41	18.94	27.40	18.87	15.61
Rms deviations bond lengths (Å)	0.020	0.021	0.020	0.019	0.024
Rms deviations angles (degrees)	1.87	1.46	1.97	1.76	2.15
PDB code	2gux	2gty	2gue	2guc	2gud

[a] The number of unique reflections is given first with Friedel pairs unmerged and after merging in parentheses.
[b] The values in parentheses relate to the highest resolution shell.
[c] R$_{merge}$ = $\sum |I - <I>|/\sum I$, where I is the observed intensity, and $<I>$ is the average intensity obtained from multiple observations of symmetry-related reflections after rejections.
[d] R = $\sum ||F_o| - |F_c||/\sum |F_o|$, where F$_o$ and F$_c$ are the observed and calculated structure factors, respectively.
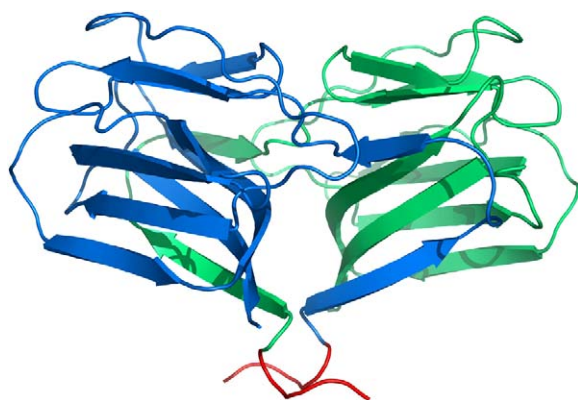[e] R$_{free}$ = defined by Brünger (Brünger, 1992).

Figure 2. A Ribbon Cα Trace of a Griffithsin Dimer, Based on the Coordinates of the His-Tagged Version of the Protein

Molecule A is colored blue, and molecule B green, with the N-terminal extension that resulted from the cloning procedure colored red.

predicted to represent boundaries between subdomains (Mori et al., 2005), since similar sequences are often reported for typical flexible linkers. However, in griffithsin, these sequences are found in loops connecting the first and fourth strand of each sheet rather than in interdomain linkers and are very well ordered. As discussed below, the main chain amide of the last residue of each of these sequences participates in creation of a ligand binding site, and the strict conservation of this sequence may be related to the presence of three monosaccharide binding sites on each molecule of griffithsin.

Comparisons with Related Lectins

The structure of griffithsin was compared to the structures of a number of proteins that display a similar overall fold, including mannose binding lectins. For the purpose of such comparisons, we utilized a compact domain of griffithsin derived from the coordinates of the high-resolution, orthorhombic form of the carbohydrate-free protein. The domain used for the comparison consisted of residues A1–A18 and B19–B121. For other proteins used in these comparisons for which multiple structures exist, we preferentially chose the structures of complexes with carbohydrates and, in case of multiple molecules in the asymmetric unit, arbitrarily chose

the first molecule present in a file if bound to a carbohydrate.

We found artocarpin (PDB code 1vbo) (Jeyaprakash et al., 2004) to be the closest homolog of griffithsin, with an rms deviation of only 1.44 Å between 110 pairs of Cα atoms. The deviations for the closely related lectins including banana lectin (2bmz) (Meagher et al., 2005), heltuba (1c3m) (Bourne et al., 1999), *Artocarpus hirsuta* lectin (1toq) (Rao et al., 2004), *Maclura pomifera* agglutinin (1jot) (Lee et al., 1998), and jacalin (1ugw) (Jeyaprakash et al., 2003) were only slightly larger, at 1.60 Å (113 pairs), 1.64 Å (112 pairs), 1.72 Å (112 pairs), 1.78 Å (111 pairs), and 1.80 Å (113 pairs), respectively. The deviation from molecule D of calsepa (PDB code 1ouw; that molecule chosen due to the presence of bound mannose) is 2.05 Å for 115 pairs, whereas the deviation for the first of the three covalently linked domains of *Parkia* lectin was 1.97 Å for 116 pairs. By comparison, the deviation from the coordinates of vitelline membrane outer layer protein I (PDB code 1vmo) (Shimizu and Morikawa, 1996), a protein that may exhibit enzymatic rather than strictly carbohydrate binding properties, was 3.0 Å for 116 pairs of Cα atoms.

The first of the repeated broad GGSGG loops in griffithsin (residues 8–12) is located between strands 1 and 2 that are swapped between the two molecules forming the dimer. This region is close to the principal carbohydrate binding sites in this family of lectins, yet it shows only partial structural conservation, limited to glycines 8 and 12. Jacalin, *Maclura pomifera* agglutinin, and *Artocarpus hirsuta* lectin contain two separate chains, with a break between chains B and A falling in exactly this area. The trace of that chain in other lectins, on the other hand, is very similar to that in griffithsin (Figure 3), although the amino acid sequence is not very highly conserved between them (Figure 1A). A loop connecting the two inner strands of the same sheet (Gly108-Tyr110 in griffithsin), on the other hand, follows an almost identical path in all the compared structures. This structural element interacts closely with the carbohydrate molecules (or with a sulfate in the structure of griffithsin solved in the absence of a sugar ligand), and it is not surprising that its sequence is also very similar.

Griffthsin differs very considerably from all other compared lectins in the conformation of the adjacent
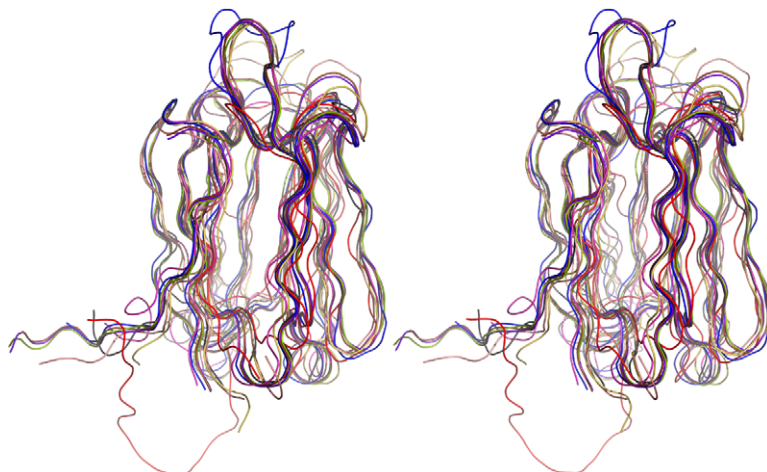

Figure 3. Superposition of the Backbone Tracing of Griffithsin and Other Related Lectins

A compact domain of griffithsin was defined as residues A1–A18 and B19–B121. The following colors were used: griffithsin, red; artocarpin (1vbo), blue; heltuba (1c3m), magenta; *parkia* lectin (1zgs), yellow; calsepa (1ouw), pink; banana lectin (2bmz), black; jacalin (1ugw), violet; AHL (1toq), brown; MPA (1jot), green.

β-hairpin, consisting of Gly66 and Asp67. This loop is much shorter than in any other lectins, leaving this part of the structure much more open. With the exception of artocarpin, in which this loop is even longer, all other lectins are similar in this area.

Another conserved broad loop with the GGSGG sequence, 86–90, is located near the secondary carbohydrate binding site in calsepa and follows a course quite different than in banana and *parkia* lectins, yet similar to the path in the other proteins. However, since that site is created in calsepa also by the previously described loop that is significantly shorter in griffithsin, it is unlikely that an analogous site could exist in the latter.

Another short inner loop formed by residues 26–28 in griffithsin bears considerable similarity to the equivalent loops in banana and *parkia* lectins as well as artocarpin but is quite different in the remaining proteins. That loop, together with the last broad GGSGG loop (40–44) creates a second carbohydrate binding site in banana lectin. The latter loop differs very significantly between most of these proteins, with only its path in banana lectin having any resemblance to its trace in griffithsin.

Only 12 residues are strictly conserved in all the compared structures (Figure 1A). It appears that almost all, if not all of them, are crucial in establishing the β-prism-I fold. As mentioned above, Gly8 and Gly12 are part of a broad loop in the first sheet and are found in the vicinity of the principal carbohydrate binding site. Interestingly, even though these residues are conserved also in the two-chain lectins such as jacalin, they are located on two separate chains that are not connected directly. Glu56, Phe74, Thr76, and Asn77 are all part of a partially buried cluster that fills the space between sheets 2 and 3 and maintains their spacing. The side chains of Glu56 and Thr76 make a strong (2.6 Å) hydrogen bond, and this feature is also strictly conserved. Gly83, Pro84, and Gly86 are part of an outer strand in sheet 3 and are located in a region that shows surprising conservation, including the *cis* conformation of the Gly-Pro peptide bond. Gly105 and Gly108 are found in a very highly conserved loop connecting the inner strands of sheet 1, whereas the nearby Asp112 may be the only side chain of a residue directly involved in carbohydrate binding to be found in all the structures.

Six residues that are conserved in all other structures are not preserved in griffithsin (Figure 1). Four of them (Arg5, Phe7, Ile63, and Tyr68) are part of a cluster between sheets 1 and 3. The strand containing the latter two residues is shifted considerably in griffithsin compared to all other structures and thus may need differently sized residues to maintain its location. Ile103 (Phe in all other structures) is deeply buried and its nature depends on the other neighbors found in its vicinity. The side chain of Tyr68 points at the unique carbohydrate binding site of griffithsin (see below), and this residue may be partially responsible for the properties of this lectin. In addition to these residues, a deletion in the aligned sequences, located between residues 18 and 19 of griffithsin, corresponds to an isoleucine in all other structures. Deletion of this residue may be related to the domain-swapped nature of griffithsin that differentiates this protein from all the other related lectins. It should be noted that the site of domain swapping that follows the second β strand is topologically different than the location of the separation between chains A and B of jacalin, the latter following the first β strand.

The structure presented here does not shed much new light on the nature and/or importance of residue 31, which was reported not to be among the 20 standard amino acids (Mori et al., 2005). An alanine inserted at that position in the artificial genes used for the work described here is buried, but the side chain of His38 at which it is pointing (at a distance of 3.3 Å to CE1) could easily move if a larger residue were to substitute for Ala31. Although this residue follows Asp30, the side chain of which interacts directly with a bound sugar (see below), it does not appear itself to be implicated in carbohydrate binding since its side chain points into a direction opposite to that of Asp30. The residues that follow the other two aspartic acids involved in carbohydrate binding are asparagine and serine; thus the identity of these residues does not appear to be an important contributor to the carbohydrate binding properties of griffithsin.

The Mode of Binding of Carbohydrates

Cocrystallization of griffithsin with mannose yielded a new, monoclinic crystal form with excellent diffraction properties, containing a griffithsin dimer in the asymmetric unit. A comparison of the coordinates resulting from the two high-resolution structures of the unliganded and mannose bound griffithsin yielded rms difference of 0.46 Å. The largest differences were observed for SerA88, GlyB11, and GlyB53, as well as in the orientation of the side chains of Tyr28, Tyr68, Asp70, and Tyr110 of both molecules. Native griffithsin incorporated SO_4^{-2} ions in places corresponding to the localization of Man1, Man4, Man5, and Man6 in the structure of the complex with mannose and a molecule of ethylene glycol in place of Man2.

Seven mannose molecules were found to be bound to each dimer of griffithsin in the 1.8 Å structure but only six in the atomic-resolution structure. The latter carbohydrates are very well ordered in both structures, whereas the seventh molecule in the lower-resolution structure, bound in a site created through intermolecular interactions in the crystal, shows some indications of disorder. It is likely that low-occupancy mannose might also be present in this site in the atomic-resolution structure, but such variability suggests that ligand binding in this site is not very specific, and thus site 7 will not be considered further. However, the six principal sites are arranged in two groups of three, one on molecule A and the other on molecule B, and are very similar in both structures. The only significant difference is that whereas the electron density in the lower-resolution structure (Figure 4) could be interpreted as resulting from the presence of exclusively α mannose, clear indications of the presence of an admixture of β mannose are seen in the atomic-resolution structure. Similar presence of both anomers of mannose was also indicated in the *Ralstonia solanacearum* lectin, solved at atomic resolution (PDB code 1vyy; no other reference available). This minor component will not be considered in the further discussion.

The interactions between Man1, Man2, and Man3 with molecule A are virtually identical to the interactions of Man4, Man5, and Man6 with molecule B. Man1
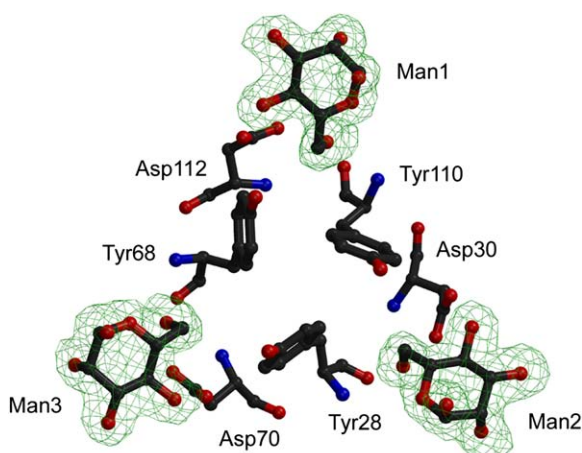
Figure 4. Mannose Binding Sites 1–3, Created Principally by Molecule A of Griffithsin

The omitmap $F_o - F_c$ electron density map was calculated at 1.8 Å resolution, based on the final coordinates of the structure of crystal 1 refined after the removal of the mannoses, and it was contoured at the 2.7 σ level.

(and Man4) are found in a site that corresponds to the major carbohydrate binding site described for a variety of lectins from this family. Due to domain swapping that exists in griffithsin, however, this site is formed by the residues from both molecules A and B. The protein side chain most responsible for the specific interactions with the mannose is AspA112, which interacts through its carboxylate with the hydroxyl oxygens O6 and O4 of the mannose. Oxygen O6 also accepts a hydrogen bond from the main chain amide of TyrA110, whereas O5 is a recipient of a bond from the N of AspA109. Hydroxyl O3 accepts a proton from the main chain amide of GlyB12, while O2 and O1 interact only with solvent (water or glycol). The interactions for Man4 are the same with the opposite chains. This pattern of interactions is very similar for the other mannose residues. Man2 (and Man5) bind in an area anchored by AspA30 and B30, respectively, with the hydrogen bonds to the main chain amides of residues 27, 28, and 44. Similarly, Man3 (and Man6) interact primarily with Asp70, with the other hydrogen bonds supplied by the main chain amides of residues 67, 68, and 90 (Figure 5).

The presence of three carbohydrate binding sites that form an almost perfect equilateral triangle on the edge of the lectin (Figure 4) is unprecedented for this family, although seen in lectins with β-prism-II fold (Hester et al., 1995; Wright et al., 2000). However, the distances

between the sites in griffithsin are ∼14 Å, whereas corresponding distances in *Galanthus nivalis* agglutinin are ∼25 Å. While the presence of site 1 was reported for all β-prism-I lectins, site 2 has so far only been seen in banana lectin (Meagher et al., 2005). As seen in Figure 1, Asp112 is strictly conserved in all lectins that were compared, explaining the conservation of this site. Not surprisingly, the only other lectin that has an equivalent of Asp30 is banana lectin since that residue is crucial for creation of site 2. Site 3, however, appears to be unique to griffithsin, and Asp70 is not conserved at all in any other β-prism-I lectin.

The three carbohydrate binding sites of griffithsin are formed from the parts of the structure with extensive sequence conservation. A tyrosine precedes the aspartic acid by two positions, and a glycine by four. Either a leucine or isoleucine is immediately preceding the aspartate, and position −3 is occupied by an aspartic acid or a serine. With the sequence and the structure of this region highly conserved, it is not surprising that the main chain amides of these residues can make very similar interactions with the sugars in all three sites. The sequence of site 1 is conserved among other lectins. Gly108 is strictly conserved, and the positions corresponding to Tyr110 and Leu111 in griffithsin are occupied by similar residues in all β-prism-I lectins compared in Figure 1A. Sequences corresponding to site 2 are not well conserved, except for banana lectin where a glycine precedes by four positions the aspartic acid involved in carbohydrate binding. Sequences corresponding to site 3 show also limited conservation. For example, Gly66 is conserved in all other lectins except for banana lectin. A position corresponding to Tyr68 in griffithsin is occupied by valine, and the position of Ile69 is occupied by similar residues in the other lectins.

The final interactions in the three sites are provided by the main chain amides of Gly12, Gly44, and Gly90, respectively. All these glycines are the C-terminal residues of the strictly conserved sequence GGSGG, and this observation may indeed explain the unprecedented conservation of this feature of the primary structure of griffithsin. Gly12, together with Gly8, is strictly conserved in all compared lectins. The sequence GGSGG is also present in banana lectin in site 2. Gly86 is strictly conserved, and Gly90 is present in all lectins except for heltuba.

While the carbohydrate molecules observed in the cocrystal structure are remarkably clear, we have also observed limited binding of monosaccharides in the orthorhombic crystals soaked in either mannose or N-acetylglucosamine, although in none of these structures
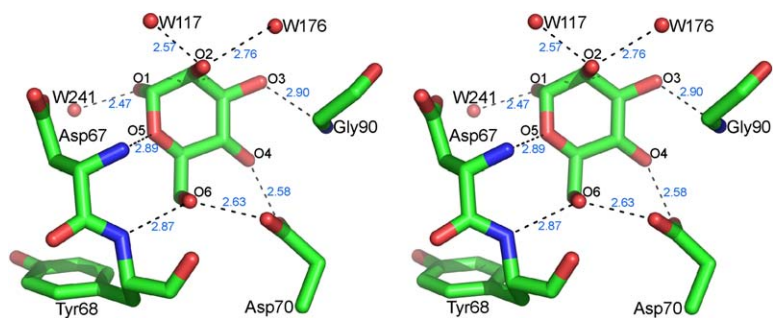

Figure 5. Details of the Interactions between Man3 and Griffithsin

The locations and lengths of hydrogen bonds are representative for all six principal carbohydrate binding sites.

Table 2. The Anti-SARS Virus Activity of Grifithsin

Assay Format	EC$_{50}$ (μM)	EC$_{90}$ (μM)
Visual cytopathic effect	0.28	n.r.
Neutral red	0.96	n.r.
Virus yield	n.r.	0.34

n.r., not recorded.

Table 3. Anti-HIV Activity of Several Lectins

Lectin	EC$_{50}$ (nM)
Jacalin[a] (O'Keefe et al., 1997)	>227
Concanavalin A[a] (Charan et al., 2000)	98
Urtica diocia agglutinin[a] (Charan et al., 2000)	105
P. tetragonolobus lectin[a] (Charan et al., 2000)	52
N. pseudonarcissus lectin[a] (Charan et al., 2000)	96
Myrianthus holstii lectin[a] (Charan et al., 2000)	150
Scytovirin (Bokesch et al., 2003)	0.3
Cyanovirin-N (Boyd et al., 1997)	0.1
Griffithsin[a] (Mori et al., 2005)	0.04

[a] Monosaccharide-specific lectins.

were all six principal sites occupied. Site 6 is clearly occupied by N-acetylglucosamine in a conformation similar to that of Man6 in the cocrystal structure, although other sites are occupied by other ligands present in the mother liquor. Ethylene glycol is found in site 5, sulfate molecules are found in sites 1 and 4, site 2 is occupied by only water molecules, whereas site 3 is involved in intermolecular contacts in this crystal form and is thus not accessible to ligands. With the carbohydrate binding sites at a variable distance from the symmetry-related protein molecules, it is not certain whether the preferential binding of a carbohydrate molecule in site 6 indicates higher affinity or simply easier accessibility. The latter explanation seems to be more probable in view of the similarity of the sites as shown in the cocrystal structure.

To further test the antiviral activity of griffithsin, the protein was submitted for testing against the coronavirus that was responsible for SARS outbreak in 2004. It had been reported previously that the coronavirus bored a surface "spike" protein that was heavily glycosylated and was able to bind to the monosaccharide-specific human lectin DC SIGN (Yang et al., 2004). Based on the demonstrated promiscuity of griffithsin binding to oligosaccharides, it was hypothesized that griffithsin should be capable of binding to the spike protein of the SARS virus and inhibiting subsequent infection. The results of three different antiviral assays with the SARS virus showed that griffithsin could indeed inhibit both viral replication and the cytopathicity induced by the virus (Table 2). Though griffithsin was active against the virus at nanomolar concentrations, it was not as active against the SARS virus as it is against HIV (Table 3).

Discussion

Although both the His-tagged and the plant-expressed griffithsin crystallized easily, crystals grown from the plant-produced material (resembling more closely the authentic protein) diffracted significantly better. Crystals of the His-tagged griffithsin contained only a single molecule in the asymmetric unit, whereas both crystal forms grown from the plant-expressed material contained two. However, only the latter crystals diffracted to atomic (or at least near-atomic) resolution, both in the presence and absence of bound carbohydrate. In all crystals that did not contain bound carbohydrate, the binding sites exhibited considerable disorder, with even the side chains of the strictly conserved tyrosines 28, 68, and 110 showing significant flexibility. The presence of carbohydrates, whether soaked in or cocrystallized, stabilized the binding sites and significantly lowered the temperature factors of the residues involved in their binding.

Griffithsin is an unusual protein in many ways, and consequently, the determination of its three-dimensional structure is illuminating on many levels. These include

its unusual domain-swapped dimeric structure, its three repeated domains, and its six independent binding sites for monosaccharides. The antiviral lectin cyanovirin-N has previously been reported to form domain-swapped dimers but they differed from the griffithsin dimers in that half of the each molecule was involved in swapping, rather than two β strands out of 12. Both monomers and dimers of cyanovirin-N could be isolated, whereas we have seen no indications of the existence of griffithsin in a monomeric form. In addition, only four carbohydrate binding sites are present in each dimer of cyanovirin-N, rather than six in griffithsin.

The presence of three almost identical domains in each molecule of griffithsin is a unique adaptation of the β-prism-I family and an advance in our knowledge of the structural complexity of this group of lectins. The other structurally closely related lectins (Figures 1A and 3) contain either a single carbohydrate binding site, or at most two such sites, in each molecule. Although griffithsin shares structural homology with many mannose- or N-acetylglucosamine-specific lectins, its reported biological activity against HIV was >1000-fold better than those previously reported for several monosaccharide-specific lectins (Charan et al., 2000). Of those lectins most closely related to griffithsin, only jacalin has been previously reported to exhibit anti-HIV activity (Corbeau et al., 1995). Jacalin was, however, found to be relatively weak in its anti-HIV activity, not reaching an EC$_{50}$ level at a high dose of 10 μg/ml (227 nM) (O'Keefe et al., 1997) in the same whole-cell assay system in which griffithsin displayed an EC$_{50}$ = 43 pM (Mori et al., 2005). The source of this enhanced potency and, presumably, higher affinity for HIV envelope glycoproteins, is of considerable interest. An explanation is suggested from the structure of griffithsin, which offers no less than six separate binding sites for mannose in a compact domain-swapped dimer. The enhanced binding potential for the high-mannose oligosaccharides present on the HIV envelope glycoprotein gp120 to be gained from this multivalency is significant. The sites on each monomer are much closer than in the lectins that belong to the β-prism-II family, presumably allowing tighter binding of complex carbohydrates. Similar enhancement was seen in previous studies on the anti-HIV protein cyanovirin-N, which detailed the increased relative binding affinity, due to multivalent interactions, of this domain-swapped dimer for high mannose oligosaccharides (Shenoy et al., 2002). Initial attempts at the crystallization of griffithsin with larger oligosaccharides have shown signs that similar multivalent binding to these carbohydrate structures occurs

with griffithsin, leading to rapid precipitation of the complexes (data not shown).

The specificity of griffithsin for monosaccharides was at first thought to be a disadvantage for this protein's potential as a prophylactic or therapeutic agent. However, in some ways, the specificity displayed by griffithsin may be an advantage. It was previously speculated that griffithsin, due to its more promiscuous carbohydrate binding, might be active against viruses such as the coronavirus responsible for the recent SARS outbreak (Mori et al., 2005). This prediction was based on the ability of griffithsin to bind a broader spectrum of oligosaccharide structures than other antiviral lectins such as cyanovirin-N (Bolmstedt et al., 2001) and scytovirin (Adams et al., 2004). This prediction has proved to be true as recent data from anti-SARS assays has shown that griffithsin is indeed active against this virus at nanomolar concentrations, while both cyanovirin-N and scytovirin were inactive (data not shown). Whether or not this protein will prove to have any clinical utility as a prophylactic or therapeutic agent for the treatment or prevention of viral infections remains to be answered, but its unique structure and interactions with carbohydrates, in addition to its unusual potency against HIV, provide ample justification for continued study.

Experimental Procedures

Cloning and Bacterial Expression of Griffithsin
His-tagged griffithsin was expressed and purified as previously reported (Giomarelli et al., 2006). Protein labeled with SeMet (Calbiochem, San Diego, CA) was expressed in a similar manner. Briefly, His-griffithsin expressing strain BL21(DE3) cells were grown overnight at 37°C on an LB agar plate containing kanamycin at a concentration of 30 μg/ml. A single colony was picked from the plate and inoculated in a 500 ml culture shake flask containing 125 ml Luria-Bertani (LB) medium (Difco, Gaithersburg, MD). The overnight culture was grown at 37°C with 250 rpm shaking for 9 hr. The total of 2 liters of Overnight Express Autoinduction Systems 2 medium (Novagen, Madison, WI) was equally divided into five 2 liters baffled culture shake flasks. Kanamycin was added to this medium to a final concentration of 30 μg/ml. Each flask then received 8 ml of the original seed culture and was grown at 37°C with 300 rpm shaking. After 16 hr of growth, cells were collected for extraction and purification of SeMet-labeled His-griffithsin as previously reported (Giomarelli et al., 2006). His-tagged protein produced in *E. coli* was used for all of the in vitro studies on the antiviral activity of griffithsin against the SARS virus.

Purification of Griffithsin from *Nicotiana* Tissue
Nicothiana benthamiana plants (24 days postsowing) were grown in a controlled environment (Shivprasad et al., 1999) and inoculated with infectious recombinant tobamovirus vectors 1 that carried a synthetic cDNA of griffithsin. Twelve days post inoculation (dpi), infected leaf and stem material was harvested at one leaf above the inoculation leaves, and tissue was homogenized in 20 mM acetate buffer (pH 4.0) containing 250 mM NaCl, 15 mM ascorbic acid, and 5.3 mM sodium meta-bisulfite at tissue-to-buffer ratio of 1:1.25 w/v. Green juice was recovered by passing the homogenate through four layers of cheesecloth. Recovered green juice was adjusted to pH 4.0 and then clarified by centrifugation at 7500 × g for 30 min (S1 fraction). S1 fraction was dialyzed overnight at 4°C against phosphate-buffered saline (PBS) (pH 7.4) with 7000 MWCO tubing and centrifuged at 75,000 × g for 30 min to remove insoluble proteins and particles. Prior to column chromatography, the pH of the S1 fraction was adjusted to 4.0 by using HCl and filtered through a 0.45 μm filter. Clarified S1 was loaded on a SP Sepharose Fast Flow (GE Healthcare, Piscataway, NJ) column equilibrated with 20 mM acetate buffer (pH 4.0) (solution A). After washing the unbound proteins from the column, griffithsin was eluted with a linear NaCl

gradient from 0 to 250 mM in solution A, with subsequent pH adjustment of eluted protein fractions to pH 7.0 by using 1 M phosphate buffer (pH 8.0) to a final concentration of 20 mM. Fractions were analyzed by SDS-PAGE with 16% Tris glycine SDS-Page gels and by matrix-assisted laser desorption-ionization time of flight (MALDI-TOF) mass spectrometry with a PerSeptive Biosystem Voyager-DE STR (Applied Biosystems, Foster City, CA) to verify fractions for pooling. Protein was then loaded on a Source 15RPC column (GE Healthcare, Piscataway, NJ) equilibrated with 20 mM phosphate (pH 7.4) (solution B) and eluted with a 2.5% n-propanol gradient, followed by a 6% linear n-propanol gradient in solution B. Fractions containing >99.5% pure GRFT were pooled and dialyzed against PBS (pH 7.4) by using 7000 MWCO dialysis tubing.

The plant-expressed griffithsin did not bear a His-tag, but was acetylated on its N terminus. This material, like the *E. coli* produced griffithsin, contained an alanine residue at position 31 to replace the unidentified amino acid reported in the original publication (Mori et al., 2005).

Severe Acute Respiratory Syndrome Virus Assays
Three types of assays were used to assess the activity of griffithsin against the severe acute respiratory syndrome (SARS) coronavirus.
Inhibition of Viral Cytopathic Effect
Griffithsin was tested for its ability to inhibit the cytopathic effect (CPE) of the SARS virus (strain 200300592). For the CPE assay, four log10 dilutions of griffithsin (e.g., 1000, 100, 10, 1 μg/ml) were added to triplicate wells of a 96-well flat-bottomed microplate containing the Vero 76 cell monolayer; within 5 min, the SARS virus was added, the plate was sealed and incubated at 37°C, and CPE was read microscopically when untreated infected controls developed a 3 to 4+ CPE (approximately 72 to 120 hr). A known positive control drug (interferon α-n3) was evaluated in parallel with griffithsin. The data were expressed as 50% effective concentrations (EC_{50}).
Increase in Neutral Red Dye Uptake
This test was run to validate the CPE inhibition seen in the initial test and utilizes the same 96-well microplates after the CPE has been read. Neutral red (NR) was added to the medium, and the resulting dye uptake by the Vero 76 cells was read on a computerized microplate reader as previously published (McManus, 1976). An EC_{50} was also determined from this dye uptake.
Decrease in Virus Yield Assay
Following positive CPE inhibition and NR dye uptake results, griffithsin was retested both by CPE inhibition and, using the same plate, for effect on reduction of virus yield assay. Frozen and thawed eluates from each cup were assayed for virus titer by serial dilution onto fresh monolayers of Vero 76 cells. Development of CPE in these cells is an indication of presence of infectious virus. As in the initial tests, interferon α-n3 was run in parallel as a positive control. The 90% effective concentration (EC_{90}), which is the test drug concentration that inhibits virus yield by 1 log10, was determined from these data.

Crystallization of Unliganded Griffithsin
Plant-expressed griffithsin and an *E. coli*-expressed SeMet derivative were purified as reported earlier (Mori et al., 2005; Giomarelli et al., 2006). The protein samples were concentrated to about 20 mg/ml. Crystallization was carried out by the hanging drop vapor diffusion method (Wlodawer and Hodgson, 1975). The plant-expressed protein and SeMet derivative crystallized under identical conditions (0.2 M ammonium sulfate and 30% PEG 4000 at low pH), with the largest crystals reaching the size of 0.2 × 0.2 × 0.05 mm. Before flash freezing, the crystals were transferred into a cryoprotectant solution containing 5% ethylene glycol. X-ray data were collected at the SER-CAT beamline 22-ID, the Advanced Photon Source (APS), Argonne, Illinois, on a MAR300 CCD detector. All data were processed with the HKL2000 package (Otwinowski and Minor, 1997) (Table 1).

We have obtained two different crystal forms of unliganded griffithsin. The protein expressed in *E. coli* crystallized in the trigonal space group P3$_2$21, with a single molecule in the asymmetric unit. This protein construct was prepared with an N-terminal 6-His affinity tag followed by a putative thrombin cleavage site, extending the sequence by 17 amino acids (Mori et al., 2005; Giomarelli et al., 2006). However, enzymatic removal of the tag proved to be impractical and

thus the crystallized protein consisted of 138 amino acids, rather than the 121 that are present in plant-expressed griffithsin. Isomorphous crystals were grown for both the plant-expressed protein as defined above and for a similarly produced SeMet derivative.

Determination of the Structure of Unliganded Griffithsin

The structure of griffithsin in the trigonal crystal form was solved by single-wavelength anomalous diffraction (SAD), initially with HKL3000 (Minor et al., 2006), followed by density modification with RESOLVE (Terwilliger, 2003) and by automated tracing with ARP/wARP (Perrakis et al., 1999). The latter program built 123 out of 138 residues expected in this construct, with only a single one-residue break at Asn67. That residue as well as one residue each on the N and C termini were added with the program O (Jones and Kjeldgaard, 1997) during subsequent refinement and rebuilding. The structure was refined with REFMAC5 (Murshudov et al., 1997) utilizing data extending to 2.0 Å resolution (Table 1). The resulting model contains all 121 residues expected in the structure of authentic griffithsin, as well as additional five residues derived from the N-terminal extension. The rest of the N-terminal extension, including the His-tag, was found to be disordered.

The second crystal form, grown from material expressed in plants and containing only 121 residues in a protein chain, was orthorhombic (P2$_1$2$_1$2$_1$) and contained two griffithsin molecules in the asymmetric unit. These crystals diffracted better than the trigonal ones, and data were collected to near-atomic resolution. The structure was solved by molecular replacement with PHASER (Storoni et al., 2004) and was refined with REFMAC5 (Murshudov et al., 1997) at 1.3 Å resolution. Both chains could be traced end-to-end, and even their terminal residues, including acetylated N-terminal serines, were clearly observed in the electron density.

Crystallization and Structure Determination of Carbohydrate Complexes

A complex of griffithsin with N-acetylglucosamine was prepared by soaking the orthorhombic crystals described above in a reservoir solution supplemented with 50 mM N-acetylglucosamine. X-ray data were collected on a MAR345 image plate detector mounted on a Rigaku rotating anode X-ray source. CuKα radiation was focused by an MSC/Osmic mirror system. Data were processed with the HKL2000 package (Otwinowski and Minor, 1997) (Table 1). The structure of the complex was refined with REFMAC5 (Murshudov et al., 1997) at 2.0 Å resolution, with the structure of unliganded griffithsin in the orthorhombic crystal form as the initial model. A complex of griffithsin with mannose was prepared by soaking in a similar manner and data were collected, but that structure was not refined further when cocrystals became available.

A distinct crystal form of a complex of griffithsin with mannose was obtained by cocrystallization by the hanging drop vapor diffusion method. The crystals were grown in 1.8 M MgSO$_4$, 0.1 M MES (pH 6.5) with 1:10 molar ratio of griffithsin monomers to mannose. Crystals belonged to space group P2$_1$ (Table 1) and diffracted to 1.8 Å resolution on the home source described above and to 0.9 Å on a synchrotron. Full data sets were collected on two crystals of the complex. A complete data set extending to 1.8 Å resolution was collected on the home source from crystal 1. High-resolution data were collected at the synchrotron to a nominal resolution of 0.9 Å from crystal 2, but since that crystal suffered from the presence of considerable ice rings, these data were merged with another low-resolution data set collected on crystal 1, this time on the synchrotron. Data from these two crystals were scaled together with HKL2000 package (Otwinowski and Minor, 1997), retaining reflections extending only to 0.94 Å. The structure that utilized home data was solved with PHASER (Storoni et al., 2004) by using the orthorhombic structure of unliganded griffithsin and was refined with REFMAC5 (Murshudov et al., 1997). The synchrotron structure was refined starting directly from that model (Table 1).

References

Adams, E.W., Ratner, D.M., Bokesch, H.R., McMahon, J.B., O'Keefe, B.R., and Seeberger, P.H. (2004). Oligosaccharide and glycoprotein microarrays as tools in HIV glycobiology; glycan-dependent gp120/protein interactions. Chem. Biol. 11, 875–881.

Aucouturier, P., Mihaesco, E., Mihaesco, C., and Preud'homme, J.L. (1987). Characterization of jacalin, the human IgA and IgD binding lectin from jackfruit. Mol. Immunol. 24, 503–511.

Bokesch, H.R., O'Keefe, B.R., McKee, T.C., Pannell, L.K., Patterson, G.M., Gardella, R.S., Sowder, R.C., Turpin, J., Watson, K., Buckheit, R.W., Jr., and Boyd, M.R. (2003). A potent novel anti-HIV protein from the cultured cyanobacterium Scytonema varium. Biochemistry 42, 2578–2584.

Bolmstedt, A.J., O'Keefe, B.R., Shenoy, S.R., McMahon, J.B., and Boyd, M.R. (2001). Cyanovirin-N defines a new class of antiviral agent targeting N-linked, high-mannose glycans in an oligosaccharide-specific manner. Mol. Pharmacol. 59, 949–954.

Botos, I., and Wlodawer, A. (2005). Proteins that bind high-mannose sugars of the HIV envelope. Prog. Biophys. Mol. Biol. 88, 233–282.

Bourne, Y., Zamboni, V., Barre, A., Peumans, W.J., Van Damme, E.J., and Rouge, P. (1999). Helianthus tuberosus lectin reveals a widespread scaffold for mannose-binding lectins. Struct. Fold. Des. 7, 1473–1482.

Boyd, M.R., Gustafson, K.R., McMahon, J.B., Shoemaker, R.H., O'Keefe, B.R., Mori, T., Gulakowski, R.J., Wu, L., Rivera, M.I., Laurencot, C.M., et al. (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, 1521–1530.

Brünger, A.T. (1992). The free R value: a novel statistical quantity for assessing the accuracy of crystal structures. Nature 355, 472–474.

Charan, R.D., Munro, M.H., O'Keefe, B.R., Sowder, R.C.I.I., McKee, T.C., Currens, M.J., Pannell, L.K., and Boyd, M.R. (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.

Chothia, C., and Murzin, A.G. (1993). New folds for all-β proteins. Structure 1, 217–222.

Corbeau, P., Pasquali, J.L., and Devaux, C. (1995). Jacalin, a lectin interacting with O-linked sugars and mediating protection of CD4+ cells against HIV-1, binds to the external envelope glycoprotein gp120. Immunol. Lett. 47, 141–143.

Giomarelli, B., Schumacher, K.M., Taylor, T.E., Sowder, R.C., Hartley, J.L., McMahon, J.B., and Mori, T. (2006). Recombinant production of anti-HIV protein, griffithsin, by auto-induction in a fermentor culture. Protein Expr. Purif. 47, 194–202.

Hester, G., Kaku, H., Goldstein, I.J., and Wright, C.S. (1995). Structure of mannose-specific snowdrop (Galanthus nivalis) lectin is representative of a new plant lectin family. Nat. Struct. Biol. 2, 472–479.

Jeyaprakash, A.A., Katiyar, S., Swaminathan, C.P., Sekar, K., Surolia, A., and Vijayan, M. (2003). Structural basis of the carbohydrate specificities of jacalin: an X-ray and modeling study. J. Mol. Biol. *332*, 217–228.

Jeyaprakash, A.A., Srivastav, A., Surolia, A., and Vijayan, M. (2004). Structural basis for the carbohydrate specificities of artocarpin: variation in the length of a loop as a strategy for generating ligand specificity. J. Mol. Biol. *338*, 757–770.

Jones, T.A., and Kjeldgaard, M. (1997). Electron-density map interpretation. Methods Enzymol. *277*, 173–208.

Lee, X., Thompson, A., Zhang, Z., Ton-that, H., Biesterfeldt, J., Ogata, C., Xu, L., Johnston, R.A., and Young, N.M. (1998). Structure of the complex of *Maclura pomifera* agglutinin and the T-antigen disaccharide, Galβ1,3GalNAc. J. Biol. Chem. *273*, 6312–6318.

McManus, N.H. (1976). Microtiter assay for interferon: microspectrophotometric quantitation of cytopathic effect. Appl. Environ. Microbiol. *31*, 35–38.

Meagher, J.L., Winter, H.C., Ezell, P., Goldstein, I.J., and Stuckey, J.A. (2005). Crystal structure of banana lectin reveals a novel second sugar binding site. Glycobiology *15*, 1033–1042.

Minor, W., Cymborowski, M., Otwinowski, Z., and Chruszcz, M. (2006). HKL-3000: the integration of data reduction and structure solution - from diffraction images to an initial model in minutes. Acta Crystallogr. D Biol. Crystallogr., in press.

Mori, T., O'Keefe, B.R., Sowder, R.C., Bringans, S., Gardella, R., Berg, S., Cochran, P., Turpin, J.A., Buckheit, R.W., Jr., McMahon, J.B., and Boyd, M.R. (2005). Isolation and characterization of griffithsin, a novel HIV-inactivating protein, from the red alga *Griffithsia* sp. J. Biol. Chem. *280*, 9345–9353.

Murshudov, G.N., Vagin, A.A., and Dodson, E.J. (1997). Refinement of macromolecular structures by the maximum-likelihood method. Acta Crystallogr. D Biol. Crystallogr. *53*, 240–255.

O'Keefe, B.R., Beutler, J.A., Cardellina, J.H., Gulakowski, R.J., Krepps, B.L., McMahon, J.B., Sowder, R.C., Henderson, L.E., Pannell, L.K., Pomponi, S.A., and Boyd, M.R. (1997). Isolation and characterization of niphatevirin, a human-immunodeficiency-virus-inhibitory glycoprotein from the marine sponge *Niphates erecta*. Eur. J. Biochem. *245*, 47–53.

Otwinowski, Z., and Minor, W. (1997). Processing of X-ray diffraction data collected in oscillation mode. Methods Enzymol. *276*, 307–326.

Perrakis, A., Morris, R., and Lamzin, V.S. (1999). Automated protein model building combined with iterative structure refinement. Nat. Struct. Biol. *6*, 458–463.

Rao, K.N., Suresh, C.G., Katre, U.V., Gaikwad, S.M., and Khan, M.I. (2004). Two orthorhombic crystal structures of a galactose-specific lectin from *Artocarpus hirsuta* in complex with methyl-α-D-galactose. Acta Crystallogr. Biol Crystallogr. *60*, 1404–1412.

Raval, S., Gowda, S.B., Singh, D.D., and Chandra, N.R. (2004). A database analysis of jacalin-like lectins: sequence-structure-function relationships. Glycobiology *14*, 1247–1263.

Shenoy, S.R., Barrientos, L.G., Ratner, D.M., O'Keefe, B.R., Seeberger, P.H., Gronenborn, A.M., and Boyd, M.R. (2002). Multisite and multivalent binding between cyanovirin-N and branched oligomannosides: calorimetric and NMR characterization. Chem. Biol. *9*, 1109–1118.

Shimizu, T., and Morikawa, K. (1996). The β-prism: a new folding motif. Trends Biochem. Sci. *21*, 3–6.

Shivprasad, S., Pogue, G.P., Lewandowski, D.J., Hidalgo, J., Donson, J., Grill, L.K., and Dawson, W.O. (1999). Heterologous sequences greatly affect foreign gene expression in tobacco mosaic virus-based vectors. Virology *255*, 312–323.

Storoni, L.C., McCoy, A.J., and Read, R.J. (2004). Likelihood-enhanced fast rotation functions. Acta Crystallogr. Biol Crystallogr. *60*, 432–438.

Terwilliger, T.C. (2003). SOLVE and RESOLVE: automated structure solution and density modification. Methods Enzymol. *374*, 22–37.

Wlodawer, A., and Hodgson, K.O. (1975). Crystallization and crystal data of monellin. Proc. Natl. Acad. Sci. USA *72*, 398–399.

Wright, L.M., Reynolds, C.D., Rizkallah, P.J., Allen, A.K., Van Damme, E.J., Donovan, M.J., and Peumans, W.J. (2000). Structural characterisation of the native fetuin-binding protein *Scilla campanulata* agglutinin: a novel two-domain lectin. FEBS Lett. *468*, 19–22.

Yang, F., Bewley, C.A., Louis, J.M., Gustafson, K.R., Boyd, M.R., Gronenborn, A.M., Clore, G.M., and Wlodawer, A. (1999). Crystal structure of cyanovirin-N, a potent HIV-inactivating protein, shows unexpected domain swapping. J. Mol. Biol. *288*, 403–412.

Yang, Z.Y., Huang, Y., Ganesh, L., Leung, K., Kong, W.P., Schwartz, O., Subbarao, K., and Nabel, G.J. (2004). pH-dependent entry of severe acute respiratory syndrome coronavirus is mediated by the spike glycoprotein and enhanced by dendritic cell transfer through DC-SIGN. J. Virol. *78*, 5642–5650.

Accession Numbers

Coordinates and structure factors have been deposited in the PDB with accession codes 2gux (SeMet structure of the *E. coli*-expressed griffithsin), 2gty (plant-expressed, unliganded griffithsin), 2guc (a complex with mannose at 1.8 Å resolution), 2gud (mannose complex at 0.94 Å resolution), and 2gue (N-acetylglucosamine complex).