Glycobiology, 2017, vol. 27, no. 1, 50–56
doi: 10.1093/glycob/cww083
Advance Access Publication Date: 9 September 2016
Communication

Communication

High-resolution crystal structures of *Colocasia esculenta* tarin lectin

Patricia R Pereira[2,3], Jennifer L Meagher[4], Harry C Winter[3], Irwin J Goldstein[3], Vânia M F Paschoalin[2], Joab T Silva[2], and Jeanne A Stuckey[1,3,4]

[2]Centro de Tecnologia, Universidade Federal do Rio de Janeiro (UFRJ), Avenida Athos da Silveira Ramos, 149., Rio de Janeiro 21941-909, Brazil, [3]Department of Biological Chemistry, University of Michigan, 1150 W. Medical Center Drive, Ann Arbor, MI 48109, USA, and [4]Center for Structural Biology, Life Sciences Institute, University of Michigan, 210 Washtenaw Ave, Ann Arbor, MI 48109, USA

[1]To whom correspondence should be addressed: Tel: +734-647-7534; Fax: +734-763-6492; e-mail: jass@umich.edu

Received 10 August 2015; Revised 11 August 2016; Accepted 12 August 2016

Abstract

Tarin, the *Colocasia esculenta* lectin from the superfamily of α-D-mannose-specific plant bulb lectins, is a tetramer of 47 kDa composed of two heterodimers. Each heterodimer possesses homologous monomers of ~11.9 (A chain) and ~12.7 (B chain) kDa. The structures of apo and carbohydrate-bound tarin were solved to 1.7 Å and 1.91 Å, respectively. Each tarin monomer forms a canonical β-prism II fold, common to all members of *Galanthus nivalis* agglutinin (GNA) family, which is partially stabilized by a disulfide bond and a conserved hydrophobic core. The heterodimer is formed through *domain swapping* involving the C-terminal β-strand and the β-sheet on face I of the prism. The tetramer is assembled through the dimerization of the B chains from heterodimers involving face II of each prism. The 1.91 Å crystal structure of tarin bound to Manα(1,3)Manα(1,6)Man reveals an expanded carbohydrate-binding sequence (QxDxNxVxY$x_{4/6}$WX) on face III of the β-prism. Both monomers possess a similar fold, except for the length of the loop, which begins after the conserved tyrosine and creates the binding pocket for the α(1,6)-terminal mannose. This loop differs in size and amino-acid composition from 10 other β-prism II domain proteins, and may confer carbohydrate-binding specificity among members of the GNA-related lectin family.

Key words: *Colocasia esculenta*, lectinn, Manα(1,3)Manα(1,6)Man, β-prism, tarin

Introduction

Lectins are carbohydrate-binding proteins of non-immune origin found in almost all living things (Goldstein et al. 1980) performing a multitude of functions, including cell proliferation (Pereira PR, Silva JT, et al. 2015), cell adhesion, cell signaling, glycoprotein clearance and pathogen recognition (Sharon 2008). This heterogeneous group of proteins varies widely in size, structure, molecular organization and constitution of their binding sites (Sharon 2008).

Plant lectins can be classified into 12 families of evolutionary and structurally related proteins (Van Damme et al. 2008). The *Galanthus nivalis* (snowdrop) agglutinin-related family resemble each other with respect to their primary sequences, carbohydrate-binding specificities and three-dimensional structures (Van Damme et al. 2007). The GNA was the first lectin characterized (PDB IDs 1MSA and 1JPC) and the X-ray structure showed that it adopts a β-prism II fold (Hester et al. 1995), conserved among the angiosperms, that is composed of three antiparallel four-stranded β-sheets arranged as a beta barrel (Van Damme et al. 2007). Each β-prism II domain contains one, two or three carbohydrate-binding sites (CBSs) (Sharma et al. 2007) displaying a conserved motif QXD XNXVXY on the third strand of each β-sheet (Ramachandraiah and Chandra 2000). Most of the plant GNA-related lectins have

subunits derived from primary translation products, comprising a single GNA domain ~110 residues in length (Van Damme et al. 1998); whereas other GNA-related lectins comprise two homologous GNA domains arranged in tandem. The supramolecular organization of the two domains of the GNA-related lectins is determined by the degree of oligomerization and post-translational processing of the precursors (Van Damme et al. 1998). Those with a single domain, exhibit exclusive specificity toward mannose and oligomannosides (Mo et al. 1999), whereas the two-domain lectins have two different CBSs that recognize oligomannoses and complex N-linked carbohydrates, respectively (Van Damme et al. 2007).

Tarin, the GNA-related lectin from *Colocasia esculenta*, accumulates in large amounts in vacuoles of the storage parenchyma cells of edible taro corms (de Castro et al. 1992). The *tar1* gene encodes a single polypeptide chain of 28 kDa. After translation, the signal peptide is removed and the protein is proteolyzed into two different, but homologous subunits (Carneiro et al. 1990). Our group purified and characterized tarin (Pereira et al. 2014) finding that it is a 47 kDa tetramer comprised of two identical heterodimers containing monomeric subunits of 11,900 Da and 12,700 Da (Pereira PR, Winter HC, et al. 2015). The GNA-related lectin from *Remusatia vivipara* (Bhat et al. 2010), a member of Araceae family is closely related to *C. esculenta* (Li et al. 2012) was also shown to be a 49.5 kDa tetramer composed of distinct 12 kDa and 12.7 kDa subunits (Bhat et al. 2010), and its crystal structure has been solved (PDB ID 3R0E) (Shetty et al. 2012).

Herein, we describe the crystal structures of apo and Manα(1,3)Manα(1,6)Man bound tarin.

Results and Discussion

Crystallographic studies

The crystal structures of tarin in its apo and carbohydrate-bound forms were refined to resolutions of 1.7 Å and 1.91 Å, respectively (Table I). Each subunit exhibits the canonical β-prism II fold common to all GNA-lectin members, comprising three subdomains of four-stranded antiparallel β sheets arranged perpendicular to the prism axis. The fold is partially stabilized by a disulfide bond between the cysteines on faces II and III of the prism, and by a conserved hydrophobic core (Figure 1A and D) (Hester et al. 1995; Shetty et al. 2012). Tarin heterodimeric unit (Figure 1B) contains two chains, A and B, that respectively correspond to the N- and C-terminal fragments generated by proteolysis from the 28 kDa tarin precursor protein (Bezerra et al. 1995). The B chain has 47.6% sequence identity to chain A and structurally aligns to A with an rmsd of 0.698 Å in the apo structure. Residue 93 in the B chain showed heterogeneity and was modeled and refined as both glutamine and arginine with 50% occupancy for all 12 instances of the B chain (4 in the native and 8 in the Manα(1,3)Manα(1,6)Man bound structures), which is consistent with the multiple pH dependent isoforms determined by Pereira and colleagues (Pereira PR, Winter HC, et al. 2015). The heterodimer is stabilized by a domain swap where the C-terminus of the A chain forms the fourth β-strand on face I of the B chain, and the C-terminus of the B chain reproduces this same interaction in the A chain (Figure 1B). This "C-terminal exchange" dimerization has been shown in the structure of the GNA-lectin homodimer (Hester et al. 1995) and the RVL heterodimer (Shetty et al. 2012). In addition to the swapped strand, the tarin dimer interface is quite extensive, having a buried interface of ~1727.9 Å^2 that is stabilized by 24 hydrogen bonds (H-bonds), two salt bridges and numerous van der Waals contacts.

Table I. Crystallography data collection and refinement statistics

Data collection	Native tarin	Trimannoside-bound tarin
SpaceGroup	$P2_1$	P1
Unit Cell a, b, c (Å)	47.700, 82.710, 122.700	78.280, 78.828, 92.935
α, β, γ (°)	90.00, 91.76, 90.00	76.19, 70.44, 59.80
Wavelength (Å)	0.97856	0.97872
Resolution (Å)[a]	1.72 (1.75–1.72)	1.91 (1.94–1.91)
Rsym (%)[b]	6.7 (38.1)	8.5 (35.2)
<I/σI>[c]	20 (5)	10 (5)
Completeness (%)[d]	99.3 (98.1)	97.3 (83.6)
Redundancy	4.0 (3.9)	3.9 (3.5)
Refinement		
Resolution (Å)	1.7	1.91
R-Factor[e]	0.188	0.209
Rfree[f]	0.208	0.244
Protein atoms	6770	13570
Water molecules	710	835
Unique reflections	104443	135823
R.m.s.d.[g]		
Bonds (Å)	0.007	0.010
Angles (°)	1.02	1.08
MolProbity score[h]	1.27	1.36
Clash score[h]	4.15	4.84

[a]Statistics for the highest resolution bin of reflections in parentheses.

[b]$R_{sym} = \Sigma_h \Sigma_j | I_{hj} - \langle I_h \rangle | / \Sigma_h \Sigma_j I_{hj}$, where I_{hj} is the intensity of observation j of reflection h and $\langle I_h \rangle$ is the mean intensity for multiply recorded reflections.

[c]Intensity of signal-to-noise ratio.

[d]Completeness of the unique diffraction data.

[e]R-factor = $\Sigma_h | |F_o| - |F_c| | / \Sigma_h |F_o|$, where F_o and F_c are the observed and calculated structure factor amplitudes for reflection h.

[f]R_{free} is calculated against a 10% random sampling of the reflections that were removed before structure refinement.

[g]Root mean square deviation of bond lengths and angles.

[h]Chen et al. (2010).

SEC–MALS analysis has shown that the biological unit of tarin is a tetramer with a mean molecular mass of 47 kDa (Pereira PR, Winter HC, et al. 2015). The tetramer is formed through the dimerization of the B chains from the heterodimers (Figure 1C). The interface involves face II of each prism and contains several prominent π–π stacking interactions: Phe57 and Phe73 from each chain interact in the middle of the interface, while His55 from one B chain and His62 from the other form π–π stacking at each end. The buried interface for the tetramer is 988 Å^2, comprising 24 H-bonds and 14 salt bridges. This tetrameric form is similar to the tetramer of RVL (Shetty et al. 2012), indeed the tetramers align with an rmsd of 0.789 Å. Although tarin was shown to contain 2–3% carbohydrate (Pereira PR, Winter HC, et al. 2015), no electron density corresponding to glycosylation was observed in either crystal structure.

Carbohydrate-binding sites: implications for specificity

Although Manα(1,3)Manα(1,6)Man did not inhibit hemagglutinating activity (Pereira PR, Winter HC, et al. 2015), we were able to obtain the crystal structure of the trimannoside bound to tarin by incubating the protein in a solution containing 145 molar excess of Manα(1,3)Manα(1,6)Man prior to crystallization. In the crystal

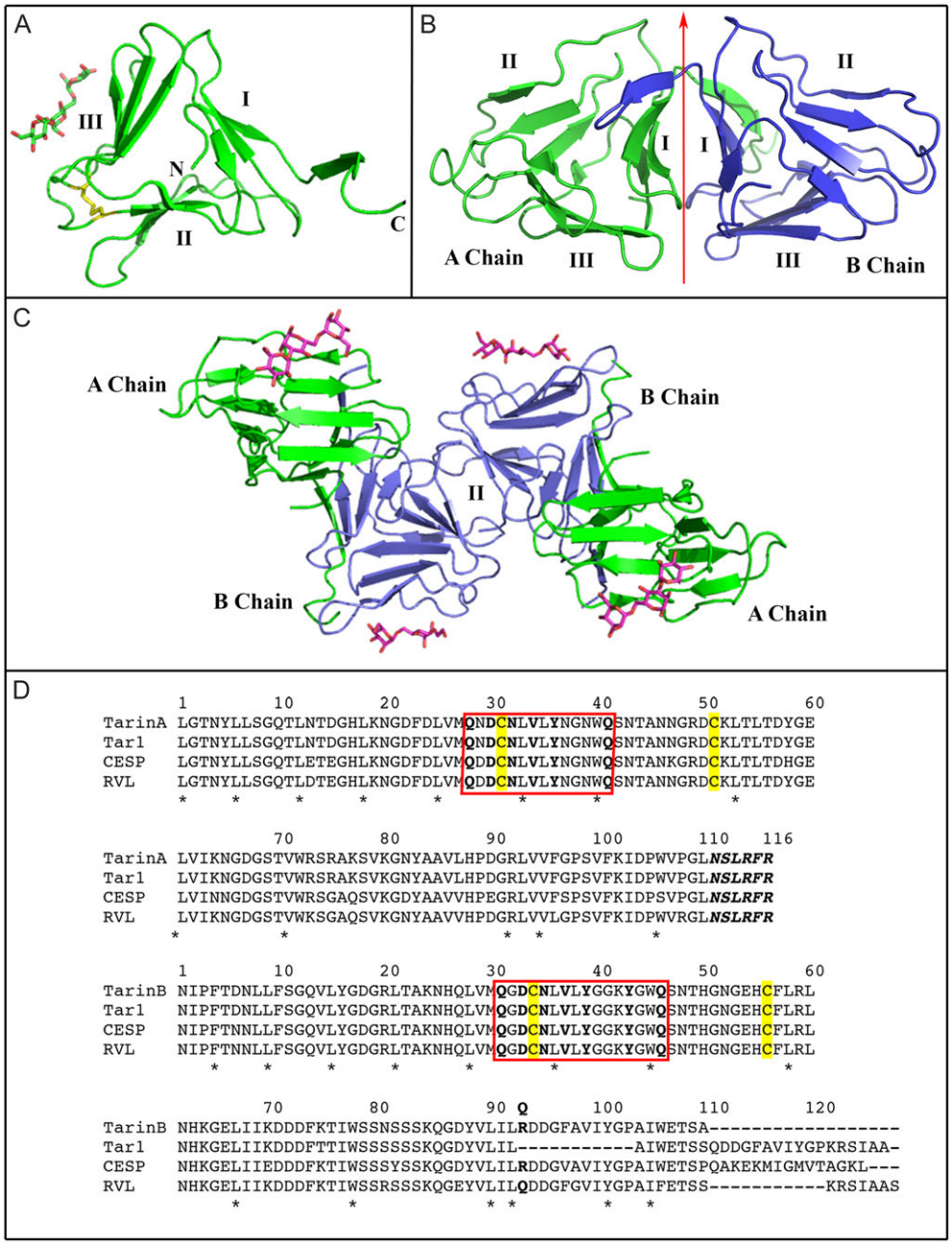
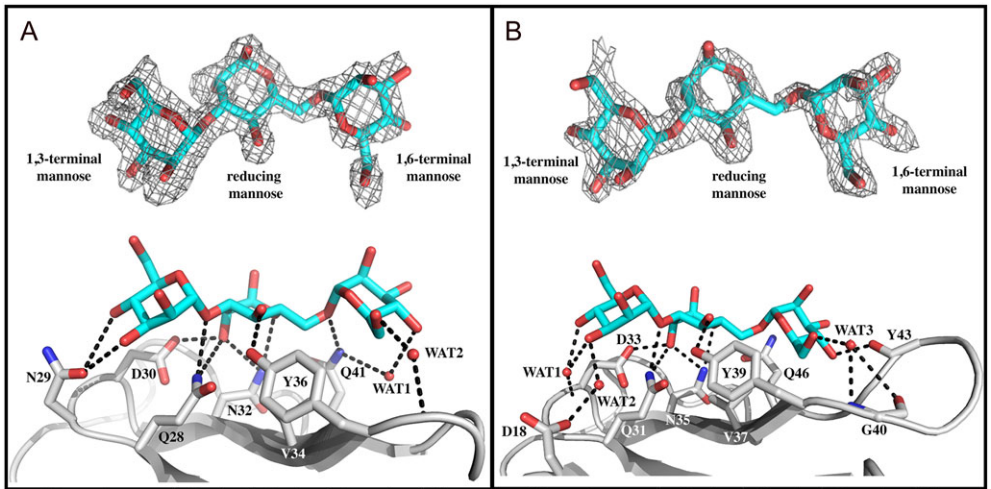


Fig. 1. *Crystal structure and higher order oligomers of tarin.* (**A**) The β-prism of tarin chain A bound to Manα(1,3)Manα(1,6)Man is shown as a green ribbon structure with disulfide bond in yellow and Manα(1,3)Manα(1,6)Man as sticks with carbons (green) and oxygens (red). The faces of the prism are labeled (I, II and III). (**B**) Tarin heterodimer, chain A (green) and chain B (blue), with the two-fold axis represented by a red arrow. (**C**) Tetrameric form of tarin showing B chains interacting via their II faces. Bound Manα(1,3)Manα(1,6)Man are shown as magenta sticks. (**D**) Sequence alignment of tarin (chains A and B), Tar1 gene (Bezerra et al. 1995), *C. esculenta* 12 kDa storage protein (CESP; GenBank: BAA03722.1) and *Remusatia vivipara* lectin (RVL: GenBank ID ACH41914.1). The expanded CBS for each chain is boxed in red with carbohydrate interacting residues in bold. The conserved proteolyzed linker sequence, not present in the structure, is in bold italics. The conserved cysteines are yellow and asterisks mark the conserved hydrophobic core.

structure, Manα(1,3)Manα(1,6)Man binds to the CBS on face III of each β-prism II domain (Figure 1C) offering insights into the binding specificity for each site.

The reducing mannose is the main driving force in binding to both sites as each has 1–2 hydrophobic interactions and 5 H-bonds to the protein with no water mediated H-bonds (Figure 2). The 1,3-terminal mannose binds via only 1–3 H-bonds directly to the protein. The well-ordered density of the 1,3-terminal mannose in the A chain is due to crystal packing. Lys24 and Asn25 from an adjacent tetramer form H-bonds with hydroxyls O3 and O2 on the solvent exposed side of the 1,3-terminal mannose, effectively sandwiching the carbohydrate in the binding site.

The binding site of the 1,6-terminal mannose heralds the greatest difference between the A and B monomers of tarin in the trimannoside-bound structure. The 1,6-terminal mannose binds via 1 H-bond directly to the A chain, but has eight interactions with the



Mannose Atom	Protein Atom	Protein Residue	Interaction Type	Distance (Å)
1,3-terminal mannose				
O1	NE2	Q28	H-bond	3.03
O3	OD1	N29	H-bond	2.72
O4	OD1	N29	H-bond	2.76
Reducing mannose				
O2	NE2	Q28	H-bond	3.35
O2	OD2	D30	H-bond	2.60
O2	ND2	N32	H-bond	2.98
O4	OH	Y36	H-bond	2.68
O5	ND2	N32	H-bond	3.34
C6	CG1/CG2	V34	Hydrophobic	
1,6-terminal mannose				
O1	NE2	Q41	H-bond	2.80
O4	O	WAT1	H-bond	2.98
WAT1	NE2	Q41	H-bond	3.10
O6	O	WAT2	H-bond	3.26
WAT2	N	N37	H-bond	3.26

Mannose Atom	Protein Atom	Protein Residue	Interaction Type	Distance (Å)
1,3-terminal mannose				
O1	NE2	Q31	H-bond	3.00
O3	O	WAT1	H-bond	3.15
O3	O	WAT2	H-bond	2.71
O4	O	WAT1	H-bond	2.89
WAT1	N	G32	H-bond	3.25
WAT1	N	D33	H-bond	2.99
WAT2	OD1	D18	H-bond	2.67
WAT2	OE1	Q31	H-bond	3.57
O4	CB/CG	D33	Hydrophobic	
Reducing mannose				
O2	NE2	Q31	H-bond	3.22
O2	OD2	D33	H-bond	2.52
O2	ND2	N35	H-bond	2.88
O4	OH	Y39	H-bond	2.69
O5	ND2	N35	H-bond	3.03
C6	CG1/CG2	V37	Hydrophobic	
C6	CD	Q46	Hydrophobic	
1,6-terminal mannose				
O4	O	Y43	H-bond	2.69
O6	O	WAT3	H-bond	2.97
WAT3	O	G40	H-bond	3.58
WAT3	N	G40	H-bond	3.17
O1	CB/CG/CD	Q46	Hydrophobic	
O4	CA/C	W45	Hydrophobic	
C5	CG1	V37	Hydrophobic	
C6	CG1	V37	Hydrophobic	
C6	O	L38	Hydrophobic	
C6	CA	G44	Hydrophobic	
C6	N	W45	Hydrophobic	

Fig. 2. *Carbohydrate-binding sites of tarin.* Manα(1,3)Manα(1,6)Man binding to the CBS on face III of (**A**) A chain and (**B**) B chain. Top views show trimannoside as sticks, oxygens (red) and carbons (cyan), with their F_o–F_c electron density maps contoured at 3σ shown as a black grid. The bottom view of each site is rotated 90° about the *X*-axis with regards to the top view. Side chains that interact with Manα(1,3)Manα(1,6)Man are shown as sticks with carbons (gray), nitrogens (blue) and oxygens (red). H-bonds are depicted as dashed lines. Conserved interacting water molecules are shown as red spheres. Noncovalent interactions between the CBSs and Manα(1,3)Manα(1,6)Man, including water mediated interactions, are listed below each picture. The H-bond distance is the average distance between the ligand and protein atom in all A or B chains of the asymmetric unit.

B chain. An overlay of the two monomers show a similar fold with an rmsd of 0.958 Å. The difference lies in the length of the loop connecting strands 3 and 4 of β-sheet III. This loop is three residues long (NGN) in A chain, and five (GGKYG) in B chain. This produces a 5 Å linear expansion of the B chain CBS allowing the 1,6-terminal mannose to lie ~2 Å closer to the surface of the protein interacting directly with the backbone atoms (Figure 2B). Based on the number of interactions of the individual monosaccharides and the overall buried surface area of the trimannoses (318 $Å^2$ and 352.6 $Å^2$ for sites A and B, respectively), it appears that the B chain has a higher affinity for Manα(1,3)Manα(1,6)Man than the A chain.

To investigate if the length and composition of this loop may confer carbohydrate-binding specificity among the GNA-related lectin family members, a structural alignment of the A and B CBSs of tarin with ten β-prism II proteins was performed (Figure 3A). Although the carbohydrate-binding consensus sequence is highly conserved among all β-prism II proteins, the loop defining the binding site of 1,6-terminal mannose varies in both size and content. The loop is bracketed by the conserved tyrosine that was part of the original consensus sequence and interacts with the reducing sugar, along with the newly noted tryptophan residue. The spatial position of the tyrosine and tryptophan is the same for all β-prism II proteins, leaving the intervening loop region to differ in sequence and conformation. RVL adopts the same binding site conformations as tarin. As the closest relative to tarin, it folds into a similar oligomeric structure (Shetty et al. 2012). The apo A chains share 92% sequence identity and structurally align with an rmsd of 0.476 Å, while the apo B chains share 95% homology and align with an rsmd of 0.376 Å. The remaining members of the β-prism II super-family adopt the loop conformation of GNA (Wright and Hester 1996).

Overlays of tarin's A and B CBSs with the trimannoside-bound form of GNA show that the GNA structure has a proline residue in

A

Protein	PDB ID	Consensus Sequence	RMSD with Tarin A (Å)	Sequence Identity with Tarin A (%)	RMSD with Tarin B (Å)	Sequence Identity with Tarin B (%)
TarinA	5T20	QNDCNLVLY**NGN---**WQ	0.000	100.00	0.958	47.57
TarinB	5T20	QGDCNLVLY**GGKY-G**WQ	0.958	47.57	0.000	100.00
RVL	3R0EA	QDDCNLVLY**NGN---**WQ	0.476	92.45	0.941	46.53
RVL	3R0EB	QGDCNLVLY**GGKY-G**WQ	0.936	49.04	0.376	94.50
PCL	3A0E	QGDCNFVLY**DSGK*P*V**WA	0.866	37.62	1.013	43.00
Curculin	2DPF	QNKCNLVKY**QNGR*Q*I**WA	0.851	46.00	0.881	42.57
Neoculin	2D04A	QNNCNLVKY**QHGR*Q*I**WA	0.752	41.41	0.713	44.44
Neoculin	2D04B	QNKCNLVKY**QNGR*Q*I**WA	0.885	46.00	0.897	42.57
GNA	1JPC	QEDCNLVLY**DVDK*P*I**WA	1.119	37.26	0.800	41.58
NPL	1NPL	QEDCNLVLY**DVDK*P*I**WA	1.107	36.27	0.797	39.60
ScaFet	1DLP	QSDCNLVLF**DSDV*R*V**WA	0.947	33.00	0.877	34.00
Gastrodianin	1XD6	QNDCNLVLY**DNNR*A*V**WA	0.959	35.48	0.888	38.30
ASA	1KJ1A	QEDCNLVLY**DHST*S*V**WA	1.030	35.35	1.173	38.38
ASA	1KJ1D	QEDCNLVLY**DHST*S*V**WT	0.982	35.05	0.912	39.00
ScaMan	1B2P	QTDCNLVLY**DNNN*P*I**WA	0.802	35.00	0.701	40.59

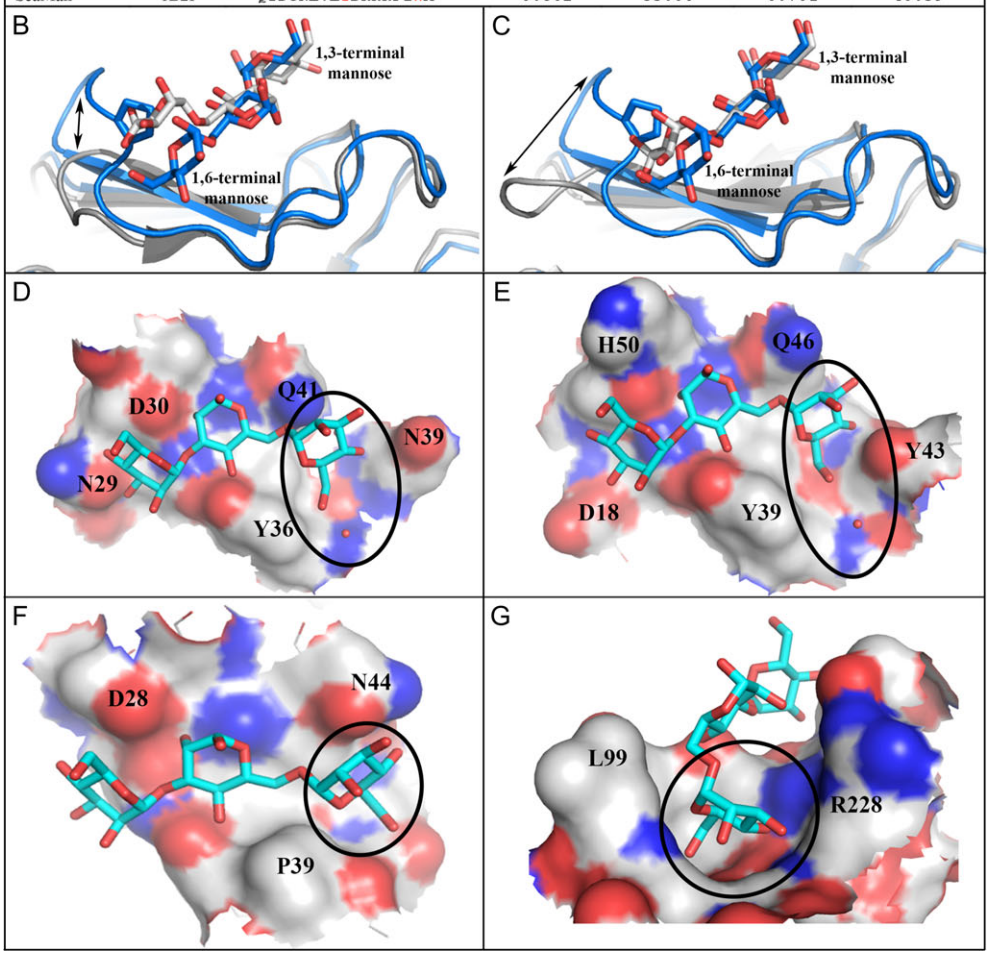


Fig. 3. *Conformational differences in carbohydrate-binding sites between β-prism II proteins.* (**A**) Sequence alignment of the expanded conserved CBSs of the ten β-prism II family members. The sequence identity and rsmd from structural alignments with tarin carbohydrate-bound forms are listed. The conserved Tyr and Trp residues are in red. The residues comprising the 1,6-terminal mannose-binding site are in bold italics. RVL, *Remusatia vivipara* lectin; PCL, *Polygonatum cyrtonema* Hua lectin; Curculin, sweet protein from *Curculigo latifolia*; Neoculin, taste-modifying protein from Hypoxidaceae family; GNA, *Galanthus nivalis* agglutinin; NPL, *Narcissus pseudonarcissus* lectin; ScaFet, fetuin-binding lectin from *Scilla campanulata*; *Gastrodianin,* mannose-binding proteins from Orchidaceae family; ASA, *Allium sativum* agglutinin; ScaMan, mannose-binding lectin from *Scilla campanulata*. Tarin CBS A (**B**) and B (**C**) were structurally aligned to GNA (PDB ID: 1JPC). In both panels, the tarin backbone is depicted in gray cartoon and bound to Manα(1,3)Manα(1,6)Man in gray sticks. GNA structure is represented as a blue cartoon with Pro39 and the bound trimannoside in blue sticks. GNA represents the β-prism II proteins listed in A, except for RVL, which adopts the tarin loop conformation. Black arrows indicate differences in the conformational loop of tarin and GNA. (**D**)–(**G**) Surface diagrams of the carbohydrate-binding residues from the CBS of tarin A chain (D), tarin B chain (E), GNA (F) and Concanavalin A (G) with carbons (gray), nitrogens (blue) and oxygens (red). Each bound trimannose is shown as cyan sticks. The black circles represent the binding site of the 1,6-terminal mannose.

the loop region, occupying the same position as 1,6-terminal mannose in the tarin structure (Figures 3B and C and Supplementary Figure S1). Therefore, the position of 1,6-terminal mannose in the GNA structure is displaced up to 4.5 Å with regards to its position in tarin. Although the proline forces a kink in the loop, this conformation is not dependent on the presence of a proline residue

within the sequence, but on the actual number of residues in the loop. β-prism II proteins that display 6 residues between the conserved tyrosine and tryptophan have the side chain of their fifth amino acid structurally aligning with Pro39 of GNA. The fifth position of this loop is very diverse in sequence and may determine the binding specificity.

In addition to the sequence and conformation of the loop, the overall surface conformation of the binding site may infer whether the protein will preferentially bind terminal mannoses of high-mannose glycans or non-terminal mannoses in complex N-glycans. Pereira PR, Winter HC, et al. 2015 determined that tarin preferred binding non-terminal mannoses, as opposed to other trimannose binding proteins that preferentially bind terminal mannoses, such as snowdrop GNA (Fouquaert et al., 2009) and Concanavalin A (Naismith and Field 1996). The surface of the CBS of Concanavalin A forms a deep pocket to accommodate the binding of the 1,6-terminal mannose (Figure 3G). The CBS surface of GNA snowdrop is only the size of the trimannose with residues Asn44 and Pro39 effectively sandwiching the 1,6-terminal residue (Figure 3F). The CBS surfaces of tarin are larger and flatter than GNA (Figure 3D and E). This increase in surface area, especially on the B chain, could allow for the recognition of an additional saccharide unit or simply reduce steric impediments for linked saccharide units from complex N-glycans. Further studies of tarin bound to complex N-glycans are needed to fully map the entire CBS and understand its specificity.

The structures presented here have provided evidence that differences in size, shape and amino-acid content of the loop, starting after the tyrosine in the conserved carbohydrate-binding sequence of GNA-related lectins, can contribute to the carbohydrate-binding specificity of these proteins, thus expanding the carbohydrate-binding sequence to $QxDxNxVxYx_{4/6}WX$.

Materials and methods

Tarin purification

A crude extract from taro tubers [*C. esculenta* (L.) Schott] (Rio de Janeiro, Brazil) was prepared as previously described (Roy et al. 2002). Tarin was purified to homogeneity. Extract was loaded onto a Cibacron Blue Sepharose 3GA column pre-equilibrated with 10 mM Tris–HCl pH 7.5 and protein eluted by 0.4 M NaCl at a rate of 1.7 mL/min. Two mL fractions were collected during elution step until the absorbance at 280 nm dropped below 0.01 (Pereira et al. 2014). The purified tarin was dialyzed against distilled water, overnight at 4°C, freeze dried and resolubilized in a minimum volume of elution buffer for crystallization studies.

Crystallization and data collection

Native tarin crystals were grown by sitting drop vapor diffusion, mixing equal volumes of tarin at 6.5 mg/mL in 10 mM Tris, pH 7.5 and 0.4 M NaCl with a well solution containing 25–35% PEG 3350, 0.2 M lithium sulfate and 0.1 M Hepes, pH 7.0 or 0.1 M Bis-Tris, pH 6.5 at 20°C. For the trimannoside-bound structure, tarin at 6.5 mg/mL was pre-incubated with 20 mM trimannoside for 4 h at room temperature. Crystals of trimannoside-bound tarin grew under similar conditions as the native crystals. All crystals grew within one week. Diffraction data were collected at LS-CAT 21-ID-G (apo) and 21-ID-F (trimannoside) at the Advance Photon Source at Argonne National Laboratory. The data were processed with HKL2000 (Otwinowski and Minor 1997). Native tarin crystallized in space group P21 with 4 heterodimers in the asymmetric unit. Trimannose-bound tarin crystallized in space group P1 with 8 heterodimers in the asymmetric unit.

Structure determination and refinement

Tarin structure was solved by molecular replacement, inputting the sequence of tarin from *C. esculenta* (GenBank: CAA53717.1) in the Balbes server (Long et al. 2008), which in turn, used the fetuin-binding lectin (PDB ID 1DLP) as a search model. The trimannoside-binding tarin structure was solved by molecular replacement using Phaser (McCoy et al. 2007), with the apo heterodimer as a search model. The tarin sequence was determined by the high-resolution electron density maps, sequence alignments of various isoforms of tarin (Bezerra et al. 1995; Van Damme et al. 1995) and mass spectrometry data (data not shown).

Difference electron density maps of the trimannoside-bound structure showed well-defined electron density for one trimannoside bound per monomer in 15 of the 16 monomers. The binding site of the P monomer is partially occupied, and due to crystal packing, the F monomer shows electron density for a disaccharide unit. Both structures went through iterative rounds of electron density fitting and refinement in Coot (Emsley and Cowtan 2004) and Buster (Bricogne et al. 2011), respectively. Calculations of protein interfaces including noncovalent interactions and buried surface areas were completed by the PISA server (Krissinel and Henrick 2007). Protein-ligand interactions were calculated using Molecular Operating Environment software (Inc CCG 2011). Structural representations were created using PyMOL (Schrödinger 2015). For structural analysis and comparisons, rmsd calculations were performed in Coot using the secondary-structure matching algorithm (Krissinel and Henrick 2004) on all Cα within the aligned monomers.

List of abbreviations

ASA, Allium sativum agglutinin; CBS, carbohydrate-binding site; CESP, *Colocasia esculenta* storage protein; GNA, *Galanthus nivalis* agglutinin; SEC, size exclusion chromatography; MALS, multiangle light scattering; NPL, *Narcissus pseudonarcissus* lectin; PCL, *Polygonatum cyrtonema* Hua lectin; PISA, protein interfaces, surfaces and assemblies; RVL, *Remusatia vivipara* lectin; ScaFet, fetuin-binding lectin from *Scilla campanulata*; ScaMan, mannose-binding lectin from *Scilla campanulata*.

Accession numbers

Coordinates and structure factors for apo tarin and trimannoside-bound tarin have been deposited in the Protein Data Bank, with PDB IDs 5T1X and 5T20, respectively.

Supplementary data

Supplementary data for this article is available online at http://glycob.oxfordjournals.org/.

Funding

P.R. Pereira was supported by a fellowship from CNPq (Conselho Nacional de Desenvolvimento Científico e Tecnológico) [Grant 142475/2008-4], CAPES (Coordenação de Aperfeiçoamento de Pessoal de Nível Superior) and the Fulbright Program [Grant BEX 3421/10-4]. The use of the Advanced Photon Source was supported by the U.S. Department of Energy, Office of Science, Office of Basic Energy Sciences, under

Contract No. DE-AC02-06CH11357. Use of the LS-CAT Sector 21 was supported by the Michigan Economic Development Corporation and the Michigan Technology Tri-Corridor for the support of this research program [Grant 085P1000817].

Acknowledgements

We thank Dr. D. Smith from LS-CAT for his help with X-ray data collection. We thank Dr. Marcus A. Nadruz Coelho from the Jardim Botânico of Rio de Janeiro for confirming taro identification. We thank Jorginaldo W. Oliveira from RFA herbarium from Universidade Federal do Rio de Janeiro, Departamento de Botânica, for helping with the exsiccate. A voucher specimen of *C. esculenta* was deposited at the RFA Herbarium under the designation RFA-39.962.

Conflict of interest

None declared.

References

Bezerra IC, Castro LA, Neshich G, de Almeida ER, Grossi de Sá MF, Mello LV, Monte-Neshich DC. 1995. A corm-specific gene encodes tarin, a major globulin of taro (*Colocasia esculenta* L. Schott). *Plant Mol Bio*. 28: 137–144.

Bhat GG, Shetty KN, Nagre NN, Neekhra VV, Lingaraju S, Bhat RS, Inamdar SR, Suguna K, Swamy BM. 2010. Purification, characterization and molecular cloning of a monocot mannose-binding lectin from Remusatia vivipara with nematicidal activity. *Glycoconj J*. 27:309–320.

Bricogne G, Blanc E, Brandl M, Flensburg C, Keller P, Paciorek W, Roversi P, Sharff A, Smart OS, Vonrhein C, et al. 2011. *BUSTER version 2.11.1* Cambridge, United Kingdom, Global Phasing Ltd.

Carneiro M, Rodrigues CA, De Castro LAB, Da Silva MC, Coutinho MV. 1990. Isolation and characterization of the major albumin from *Colocasia esculenta* corms. *Plant Sci*. 67:39–46.

Chen VB, Arendall WB, Headd JJ, Keedy DA, Immormino RM, Kapral GJ, Murray LW, Richardson JS, Richardson DC. 2010. MolProbity: All-atom structure validation for macromolecular crystallography. *Acta Crystallogr D*. 66:12–21.

de Castro LA, Carneiro M, Neshich Dde C, De Paiva GR. 1992. Spacial and temporal gene expression patterns occur during corm development. *Plant Cell*. 4:1549–1559.

Emsley P., Cowtan K. 2004. *Coot*: Model-building tools for molecular graphics. *Acta Crystallogr D*. 20:2126–2132.

Fouquaert E, Smith DF, PeumanS WJ, Proost P, Balzarini J, Savvides SN, Van Damme EJM. 2009. Related lectins from snowdrop and maize differ in their carbohydrate-binding specificity. *Biochem Biophys Res Commun*. 380:260–265.

Goldstein IJ, Hughes RC, Monsigny M, Osawa T, Sharon N. 1980. What should be called a lectin? *Nature*. 285:66–66.

Hester G, Kaku H, Goldstein IJ, Wright CS. 1995. Structure of mannose-specific snowdrop (*Galanthus nivalis*) lectin is representative of new plant lectin family. *Nature Struct Biol*. 2:472–479.

Inc CCG. 2011. *Molecular Operating Environment (MOE)* 1010 Sherbooke St. West, Suite #910, Montreal, QC, Canada, H3A 2R7, Chemical Computing Group Inc.

Krissinel E., Henrick K. 2004. Secondary-structure matching (SSM), a new tool for fast protein structure alignment in three dimensions. *Acta Crystallogr D*. 60:2256–2268.

Krissinel E., Henrick K. 2007. Inference of macromolecular assemblies from crystalline state. *J Mol Biol*. 372:774–797.

Li R, Yi T, Li H. 2012. Is Remusatia (araceae) monophyletic? Evidence from three plastid regions. *Int J Mol Sci*. 13:71–83.

Long F, Vagin A, Young P, Murshudov GN. 2008. BALBES: A molecular replacement pipeline. *Acta Crystallogr D*. 64:125–132.

McCoy AJ, Grosse-Kunstleve RW, Adams PD, Winn MD, Storoni LC, Read RJ. 2007. Phaser crystallographic software. *J Appl Cryst*. 40:658–674.

Mo H, Rice KG, Evers DL, Winter HC, Peumans WJ, Van Damme EJ, Goldstein IJ. 1999. Xanthosoma sagittifolium tubers contain a lectin with two different types of carbohydrate-binding sites. *J Biol Chem*. 274: 33300–33305.

Naismith JH, Field RA. 1996. Structural basis of trimannoside recognition by Concanavalin A. *J Biol Chem*. 271:972–976.

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

Pereira PR, Del Aguila EM, Vericimo MA, Zingali RB, Paschoalin VM, Silva JT. 2014. Purification and characterization of the lectin from taro (*Colocasia esculenta*) and its effect on mouse splenocyte proliferation in vitro and in vivo. *Protein J*. 33:92–99.

Pereira PR, Silva JT, Verícimo MA, Paschoalin VMF, Teixeira GAPB. 2015. Crude extract from taro (*Colocasia esculenta*) as a natural source of bioactive proteins able to stimulate haematopoietic cells in two murine models. *J Funct Foods* in press doi:10.1016/j.jff.2015.07.014.

Pereira PR, Winter HC, Verícimo MA, Meagher JL, Stuckey JA, Goldstein IJ, Paschoalin VMF, Silva JT. 2015. Structural analysis and binding properties of isoforms of tarin, the GNA-related lectin from *Colocasia esculenta*. *Biochim Biophys Acta*. 1854:20–30.

Ramachandraiah G, Chandra NR. 2000. Sequence and structural determinants of mannose recognition. *Proteins*. 39:358–364.

Roy A, Banerjee S, Majumder P, Das S. 2002. Efficiency of mannose-binding plant lectins in controlling a homopteran insect, the red cotton bug. *J Agric Food Chem*. 50(23): 6775–6779.

Schrödinger LLC. 2015. *The PyMOL Molecular Graphics System*, Version 1.8.

Sharma A, Chandran D, Singh DD, Vijayan M. 2007. Multiplicity of carbohydrate-binding sites in beta-prism fold lectins: Occurrence and possible evolutionary implications. *J Biosci*. 32:1089–1110.

Sharon N. 2008. Lectins: Past, present and future. *Biochem Soc T*. 36(6): 1457–1469.

Shetty KN, Bhat GG, Inamdar SR, Swamy BM, Suguna K. 2012. Crystal structure of a beta-prism II lectin from Remusatia vivipara. *Glycobiology*. 22:56–69.

Van Damme EJ, Goossens K, Smeets K, Leuven FV, Verhaert P, Peumans WJ. 1995. The major tuber storage protein of araceae species is a lectin. Characterization and molecular cloning of the lectin from *Arum maculatum* L. *Plant Physiol*. 107:1147–1158.

Van Damme EJ, Peumans WJ, Barre A, Rougé P. 1998. Plant lectins: A composite of several distinct families of structurally and evolutionary related proteins with diverse biological roles. *Crit Rev Plant Sci*. 17:575–692.

Van Damme EJ, Nakamura-Tsuruta S, Smith DF, Ongenaert M, Winter HC, Rouge P, Goldstein IJ, Mo H, Kominami J, Culerrier R, et al. 2007. Phylogenetic and specificity studies of two-domain GNA-related lectins: Generation of multispecificity through domain duplication and divergent evolution. *Biochem J*. 404:51–61.

Van Damme EJ, Lannoo N, Peumans WJ. 2008. Chapter 3 plant lectins. In: Jean-Claude K, Michel D, editors. *Advances in botanical research*. Vol. 3, Academic Press: Elsevier, p. 107–209.

Wright CS, Hester G. 1996. The 2.0 A structure of a cross-linked complex between snowdrop lectin and a branched mannopentaose: Evidence for two unique binding modes. *Structure*. 4:1339–1352.