The International Journal of Biochemistry & Cell Biology 72 (2016) 27–39

Contents lists available at ScienceDirect

The International Journal of Biochemistry & Cell Biology

journal homepage: www.elsevier.com/locate/biocel

Structural characterization of a *Vatairea macrocarpa* lectin in complex with a tumor-associated antigen: A new tool for cancer research

Bruno L. Sousa[a], José C. Silva-Filho[a], Prashant Kumar[b], Melissa A. Graewert[c], Ronniery I. Pereira[a], Rodrigo M.S. Cunha[d], Kyria S. Nascimento[a], Gustavo A. Bezerra[e], Plínio Delatorre[f], Kristina Djinovic-Carugo[e,g], Celso S. Nagano[h], Karl Gruber[b], Benildo S. Cavada[a,*]

[a] *Departamento de Bioquímica e Biologia Molecular, Universidade Federal do Ceará, Av. Mister Hull s/n, Bloco 907, Box 6043, 60440-970, Fortaleza, Ceará, Brazil*
[b] *Institute of Molecular Biosciences, University of Graz, Humboldtstrasse 50/3, A-8010 Graz, Austria*
[c] *European Molecular Biology Laboratory, c/o DESY, Notkestrasse 85, 22607 Hamburg, Germany*
[d] *Centro de Ciências Agrárias e Biológicas, Universidade do Vale do Acaraú, Rua Gerardo Rangel s/n, 62041-040, Sobral, Brazil*
[e] *Department of Structural and Computational Biology, Max F. Perutz Laboratories, University of Vienna, Vienna Biocenter (VBC), Vienna Biocenter Campus 5, 1030 Vienna, Austria*
[f] *Departamento de Biologia Molecular, Universidade Federal da Paraíba, Cidade Universitária, 58059-900, João Pessoa, Brazil*
[g] *Department of Biochemistry, Faculty of Chemistry and Chemical Technology, University of Ljubljana, Aškerčeva 5, SI-1000 Ljubljana, Slovenia*
[h] *Departamento de Engenharia de Pesca, Universidade Federal do Ceará, Av. Mister Hull s/n, Bloco 827, Fortaleza, Brazil*

ARTICLE INFO

Article history:
Received 18 September 2015
Received in revised form 5 December 2015
Accepted 31 December 2015
Available online 2 January 2016

Keywords:
Tn antigen
Cancer research
Crystal structure
Histochemistry
Legume lectin
Vatairea macrocarpa

ABSTRACT

Legume lectins are the most thoroughly studied group of lectins and have been widely linked to many pathological processes. Their use as immunohistochemistry markers for cell profiling and cancer diagnosis have made these molecules important tools for immunological studies and have stimulated the prospection and characterization of new lectins. The crystal structures of a recombinant seed lectin from *Vatairea macrocarpa* (rVML) and its complexes with GalNAcα1-O-Ser, GalNAc and α-lactose, have been determined at 1.90, 1.97, 2.70 and 1.83 Å resolution, respectively. Small angle X-ray scattering and calorimetry assays have confirmed the same pH stable oligomerization pattern and binding profiles proposed for its wild-type counterpart. *In silico* analyzes have explored the potential of this recombinant lectin as new tool for cancer research through a comparative profile with other legume lectins widely used for cancer diagnosis and prognosis. The results suggest the recognition of specific epitopes exhibited on different cancer cells as a process that relies on the disposition of hydrophobic clusters and charged regions around the lectin carbohydrate-binding site, favouring the anchorage of different groups in the antigen boundaries, highlighting the different potential of each analyzed lectin. In conclusion, the experimental results and comparative analysis show that rVML is as a promising tool for cancer research, able to bind with high affinity specific tumor-associated antigens, highly stable and easily produced.

1. Introduction

Plant lectins are a highly diverse class of carbohydrate-binding proteins of non-immune origin that have at least one non-catalytic domain. These proteins are able to bind specific sugar structures reversibly and with high specificity, but without enzymatically modifying them, triggering several important cellular processes (Lannoo and Van Damme, 2010; Liu et al., 2013; Macedo et al., 2015; Rüdiger and Gabius, 2002).

The remarkable anti-tumor properties of some plant lectins, resulting from their ability to induce programmed cell death and/or autophagocytosis in cancer cells (Fu et al., 2011; Liu et al., 2010; Yu et al., 2011), have attracted much attention for biomedical applications. In view of their binding specificities, plant lectins have become crucial tools for cancer diagnosis and prognosis, rather than simple recognition tools for identifying malignant tumors (Fritz et al., 2004; Heinrich et al., 2005; Liu et al., 2013; Majee and Biswas, 2013; Roth, 2011). They have also been widely used as candidate anti-tumor drugs in human cancers *in vitro*, and some of them

* Corresponding author. Tel.: +55 85 3366 9818; fax: +55 85 3366 9818.

http://dx.doi.org/10.1016/j.biocel.2015.12.016
1357-2725/

are further applied to pre-clinical and clinical therapies (Gemeiner et al., 2009; Li et al., 2011; Liu et al., 2010).

Alterations of glycosylation patterns at the membrane surface are consistent features of cancer cells. In tumors, this aberrant glycosylation is a consequence of the incomplete synthesis of carbohydrate chains, allowing higher expression of precursor carbohydrate moieties, such as T and Tn antigens (Galβ1-3GalNAcβα1-O-Ser/Thr)/(GalNAcα1-O-Ser/Thr) (Cao et al., 2008; Itzkowitz et al., 1989). These glycopeptide epitopes are associated with different types of tumors and represent attractive candidates, among the tumor-associated carbohydrate antigens, for the development of anticancer immune stimulation and drug targeting strategies (Brooks et al., 2010; Fuster and Esko, 2005; Lo-Man et al., 2004; Singh and Bandyopadhyay, 2007; Smart, 2004). The aberrant glycosylation pattern observed in cancer cells is well illustrated by alterations in the composition and structure of mucins, a heterogeneous group of *O*-glycosylated proteins greatly linked to tumorigenesis (Chung et al., 2015; Glavey et al., 2015). These proteins can be either secreted as the major component of the epithelial mucus (e.g. MUC2, MUC5AC, MUC5B, MUC6, MUC7, MUC9 and MUC19), acting as a protective barrier against numerous aggressions, or can be membrane-bound (e.g. MUC1, MUC3, MUC4, MUC12, MUC13, MUC15, MUC16, MUC17, MUC20, MUC21 and MUC22), participating in different cellular interactions and cell signaling. Among mucins, MUC2 and MUC1 have been extensively studied in view of their role in tumors resistance to chemotherapy (Jonckheere et al., 2014), highlighting their potential as targets for therapeutic purposes.

Legume lectins are the most thoroughly studied lectin family and have been widely reported to exhibit a number of links to many pathological processes, including malignancy (Liu et al., 2010; Van Damme et al., 1998). A few galactose/N-acetylgalactosamine (Gal/GalNAc)-binding lectins isolated from legume plants have proven to be useful markers for cancer histochemistry, such as VVLB4 (*Vicia villosa* isolectin 4), PHA (*Phaseolus vulgaris* agglutinin), DBL (*Dolichos biflorus* lectin), SBA (*Gycine max*) and PNA (*Arachis hypogaea*), among others (Bakalova and Ohba, 2003; Brooks and Carter, 2001; Konska et al., 2006; Shio et al., 2007). Nevertheless, despite the extensive efforts invested in characterizing new lectins, little is known about how these proteins interact and recognize complex antigens and glycoproteins, leading to important biological activities (Liu et al., 2009, 2008; Reynoso-Camacho et al., 2003). Recently, the recognition of a class of tumor-associated glycopeptides (MUC-1) by SBA and VVLB4 was investigated, emphasizing the process modulation by the peptide sequence close to the glycosylation point, demonstrating the influence of molecular boundaries in lectin affinity, a property directly related to the potential of these molecules as biological tools (Madariaga et al., 2014).

Legume lectins are commonly isolated from seeds and may exist as a complex mixture of isoforms, as in the case of lectins from *Robinia pseudoacacia* (Van Damme et al., 1995) and *Sophora japonica* (Van Damme et al., 1997). The presence of isoforms naturally contributes to the biological role of these proteins, offering an alternative strategy or an evolutionary adaptation (Zhu et al., 2006); at the same time, however, this is one of the main disadvantages of obtaining lectins from natural sources. Consequently, recombinant production, mainly in microbial hosts, has become a valuable strategy to circumvent this problem and may improve availability and homogeneity, as well as ensure a continuous supply of purified lectins with interesting activities, particularly for biomedical applications (Oliveira et al., 2014).

The seed lectin of *Vatairea macrocarpa* (VML) is a well-characterized homotetrameric and pH-stable Gal/GalNAc-specific lectin with remarkable biological effects, such as the release of chemotactic mediators by macrophages, leukocyte infiltration, induction of depressive-like behavior and the expression of neuroinflammatory markers in mice (Alencar et al., 2007, 2004; Calvete et al., 1998; Cavada et al., 1998; Gonçalves et al., 2013). Additionally, VML is able to recognize the Tn antigen (GalNAcα1-O-Ser/Thr), presenting interesting binding properties that suggest a potential application in cancer research (Dam et al., 2007). Recently, we have determined at high-resolution the crystal structure of wild-type VML (wtVML) in complex with Tn antigen (GalNAcα1-O-Ser) and performed computational calculations to analyze the structural bases for the recognition of naturally occurring MUC-2 (intestinal human mucin) fragments, establishing a comparative profile among different Tn-binding legume lectins (Sousa et al., 2015). The obtained results have strongly suggested VML as a promising biological tool for cancer research, showing a similar binding profile to other lectins commonly used as histological markers, such as VVLB4 and SBA.

Mature wtVML (alpha chain) consists in a double chain protein (gamma and beta) obtained as a mix of forms generated by its natural processing, which includes glycosylated and non-glycosylated forms. The oligosaccharide attached to wtVML (beta chain) was previously identified by mass spectrometry as Manα1-6[Manα1-3(Xylβ1-2)]Manβ1-4-GlcNAcβ1-4(Fucα1-3)GlcNAc, a typical plant N-glycan (Calvete et al., 1998). This major carbohydrate structure is identical to those from *R. pseudoacacia* and *Erythrina variegata* lectins, and the xylose units from these glycans seems to play a critical role in allergenicity as well as in regulating the targeting of these proteins to various organelles such as storage bodies (Altman, 2007; Wantyghem et al., 1992). This heterogeneity and allergenic potential are undesired features for research or biomedical applications, reinforcing the recombinant expression of VML in prokaryotic systems as the most appropriate strategy for producing a suitable biological tool.

Despite the broad number of studies concerning the potential applications of legume lectins in cancer research, only four structures of complex with Tn antigen are available, including VML along with VVLB4 (Babino et al., 2003), WBA I (*Psophocarpus tetragonolobus* lectin) (Kulkarni et al., 2005), and SBA (Madariaga et al., 2014). Since a necessity of further studies devoted to understanding the structural basis of the biological effects triggered by legume lectins is highlighted, we herein present the crystal structures of a recombinant *V. macrocarpa* seed lectin (rVML) and its complexes with Tn antigen (GalNAcα1-O-Ser), GalNAc and α-lactose. rVML was obtained as a non-glycosylated single-chain protein, retaining all the characteristic features observed for its wild-type counterpart, as indicated by small angle X-ray scattering (SAXS) and isothermal titration calorimetry (ITC). Molecular docking calculations were applied to investigate the potential of this recombinant lectin in recognizing typical antigens from different types of carcinoma, attesting its suitability through a comparative profile with other well-described Tn-binding legume lectins, including wtVML, SBA and VVLB4. The presented results introduce a stable and easily produced new protein tool for cancer research, which meets the requirements of availability, homogeneity, and purity for biological research assays.

2. Materials and methods

2.1. Wild-type VML purification

V. macrocarpa seeds were collected from trees growing on the Campus of the Federal University of Ceará (Fortaleza) in the northeastern part of Brazil. wtVML was purified from the seed crude extract by affinity chromatography on a Guar gum gel, following previously described procedures (Calvete et al., 1998; Cavada et al., 1998). The purity of the sample was confirmed by 12.5% gel electrophoresis.

2.2. Cloning and expression

The sequence encoding rVML was obtained as a synthetic gene (Epoch Biolabs, Missouri City, TX, USA), designed on the basis of the full-length wild-type protein sequence (UNIPROT P81371), and codon-optimized for expression in *Escherichia coli*. The gene was originally inserted into vector pET-32a (Novagen, Darmstadt, Germany), and the expressed protein was reported to be soluble and active, despite low yields (Alves-Filho et al., 2013). Based on that, the encoding sequence was subcloned into the vector pET-28a (Novagen, Darmstadt, Germany), as a strategy to increase the amount of soluble protein. The fragment was initially amplified by PCR using specific primers for insertion of restriction sites for NcoI and XhoI enzymes at the 5′ and 3′ ends, respectively, and then digested and introduced into a previously linearized pET-28a, which provides a C-terminal his-tag to the expressed protein. The construct correctness was assessed by DNA sequencing (LGC Genomics, Berlin, Germany). *E. coli* BL21 DE3 Codon Plus (Stratagene, La Jolla, CA, USA) was used as the expression host. All cells were grown in LB (lysogeny broth, Lennox) medium (Carl Roth GmbH, Karlsruhe, Germany) supplemented with kanamycin sulfate (50 mg L^{-1} final concentration). The recombinant protein was routinely expressed from cultures grown in baffled flasks at 37 °C and 120 rpm until an attenuance at 600 nm of ~0.6–0.8 was reached. Afterwards, the cultures were cooled to 30 °C, and expression was induced for 20 h by addition of 0.5 mM isopropyl thio-β-D-galactoside (Carl Roth GmbH). Protein expression and localization in cell extract fractions were monitored by SDS-PAGE.

Cells were harvested by centrifugation for 10 min at 4000 g and resuspended in buffer A (20 mM sodium phosphate pH 8.0, 250 mM NaCl and 20 mM imidazol). Sonication was applied for cell disruption, and the crude lysates were cleared by centrifugation for 30 min at 20,000 g and 4 °C.

2.3. Recombinant protein purification

The clarified lysate (soluble fraction) containing rVML was loaded onto a Ni-NTA agarose resin column (GE Healthcare) pre-equilibrated with buffer A. The retained fraction was subsequently eluted with buffer B (20 mM sodium phosphate pH 8.0, 250 mM NaCl and 500 mM imidazol) and concentrated by centrifugation (Millipore Amicon Ultra, 10,000 molecular weight cutoff), followed by loading onto a Superdex 200 16/60 size-exclusion column (GE Healthcare) pre-equilibrated with buffer C (10 mM L-malic acid, 10 mM MES, 10 mM Tris, 250 mM NaCl and 5% glycerol) (Newman, 2004). The samples were analyzed by SDS-PAGE, and positive fractions were pooled and concentrated by centrifugation, flash frozen with liquid nitrogen and stored at −80 °C. The protein concentration was determined using a Nanodrop spectrophotometer (Model 2000c, Peqlab, Erlangen, Germany), applying the extinction coefficient calculated based on the amino acid sequence using Protparam (Gasteiger et al., 2005).

2.4. Crystallization assays

Initially, rVML was concentrated by centrifugation (Millipore Amicon Ultra, 10,000 molecular weight cutoff) to 17.5 mg/mL. Crystallization attempts were performed for the uncomplexed protein, as well as for complexes with D-galactose, α-lactose, GalNac and Tn antigen (GalNAcα1-O-Ser). The concentration of protein solutions for all assays was kept at 17.5 mg/mL, based on pre-crystallization assays. For co-crystallization, ligands were added to the protein solutions to yield a ligand:protein ratio of 30 and then incubated at room temperature for 30 minutes prior to crystallization setups. Crystallization drops of 1 μL, which consisted of 0.5 μL of protein solution and 0.5 μL of reservoir solution, were set up by using the automatic protein crystallization system Oryx8 (Douglas Instruments). Different crystallization screens were used for initial screening purposes, such as JCSG-plus, Morpheus, MIDAS (Molecular Dimensions) and Index screen (Hampton Research). Among the screens tested, Index screen gave the best diffracting crystals for all crystallization attempts. Optimization assays were designed based on the initial results by ranging the composition of the selected crystallization conditions. All assays were performed using the sitting-drop vapor diffusion technique. 30% glycerol was added for cryoprotection prior to crystal mounting in a cryoloop and flash-cooling in liquid nitrogen.

2.5. Data collection and structure determination

The data sets were collected at the European Synchrotron Radiation Facility (Grenoble, France) at beam lines ID-29 and ID23-2 (see Table 1). Diffraction data were indexed, integrated and scaled using XDS (Kabsch, 1988) and programs from the CCP4 suite (Collaborative Computacional Project Number 4, 1994; Winn et al., 2011).

The structures were solved by molecular replacement with MOLREP (Vagin and Teplyakov, 2010, 1997) and PHASER (McCoy et al., 2007), using the monomer structure of wtVML (PDB code 4U2A) (Sousa et al., 2015) as the search template. Cycles of maximum likelihood refinement were carried out with the Refmac 5 program (Skubák et al., 2004; Vagin et al., 2004).

Initially, rigid body and restrained refinement cycles were performed with further alternation of manual model corrections using Coot (Emsley and Cowtan, 2004; Emsley et al., 2010) and restrained refinement cycles. After the addition of water and ligand molecules, an additional restrained refinement was carried out, leading to the R and R_{free} values listed in Table 1. The refinements of rVML and rVML-GalNAc structures were completed using TLS refinement, minimization, and individual B factor refinement with Refmac. The stereochemical quality of the final structures was assessed using MOLPROBITY and PROCHECK (Davis et al., 2007; Laskowski et al., 1993), considering B-factors as additional quality parameters. Figures were designed using the Pymol Molecular Graphics System (Schrödinger, LLC.). Atomic coordinates for the crystal structures of rVML (metalized without ligands) and rVML in complex with Tn antigen (rVML-Tn), GalNAc (rVML-GalNAc) and α-lactose (rVML-Lac) have been deposited in the Protein Data Bank under the codes 4XXA, 4XTP, 4XTM and 4WV8, respectively. Omit maps of the ligands were built using the Omit program (Vellieux and Dijkstra, 1997).

2.6. Isothermal titration microcalorimetry

The recombinant protein (rVML) was initially dialyzed against buffer D (50 mM Tris-HCl pH 7.6, 150 NaCl, 5 mM $CaCl_2$ and 5 mM $MnCl_2$) for 12 h. This first step aimed to ensure complete metal incorporation into the protein, since metal-binding is required to stabilize the carbohydrate-binding site (CBS) structure. A second dialysis step against 20 mM sodium phosphate pH 7.6 and NaCl 150 mM (PBS) was performed to prepare the samples for analysis, since PBS possesses a lower enthalpy of ionization compared to Tris buffer, thereby generating fewer artifactual signals (Goldberg et al., 2002). At the same time, lyophilized wtVML was dissolved in buffer D and treated using the same preparation steps as rVML. The protein concentration was determined using a Nanodrop spectrophotometer (Model 2000c, Peqlab, Erlangen, Germany), as described above. The ligands, GalNAc and Tn antigen (GalNAcα1-O-Ser), were dissolved in PBS and loaded in the injection syringe. The ITC measurements were performed at 25 °C with an ITC200 microcalorimeter (MicroCal, Inc.). The lectin solution (100 μM) was placed in the 200 μL sample cell. Titrations were

Table 1
Statistics of data collection, refinement and structure quality.

Parameter	rVML	rVML – Tn	rVML – GalNAc	rVML – Lac
Data collection				
Beamline	ESRF ID29	ESRF ID29	ESRF ID29	ESRF ID23-2
Wavelength (Å)	0.915	0.915	0.915	0.872
Space group	$P3_121$	I222	I222	C2
Unit cell parameters (Å)				
a	96.08	56.09	56.35	152.64
b	96.08	86.02	86.11	97.43
c	85.10	128.88	129.02	78.67
β				94.6
Total number of reflections	359699	136723	22707	314995
Number of unique reflections	36249	22033	7221	100813
Molecules per asymmetric unit	2	1	1	4
Resolution range (Å)	83.21–1.75 (1.84–1.75)	46.99–1.97 (2.07–1.97)	47.15–2.70 (2.83–2.70)	82.05–1.83 (1.93–1.83)
R_{merge} (%)	3.9 (22.7)	11 (52.0)	11 (39.0)	8.1 (26.2)
Completeness (%)	99.9 (99.6)	97.7 (85.3)	81.7 (84.8)	99.4 (98.6)
Multiplicity	9.9 (9.6)	6.2 (5.8)	3.2 (3.1)	3.1 (3.1)
I/σ(I)	10.1 (4.94)	10.0 (2.8)	6.4 (2.3)	4.10 (2.74)
Molecular replacement				
Phaser				
Z-score (rotation function)	4.5			9.4
Z-score (translation function)	9.9			9.4
Molrep				
wRfac/Score (%)		44.3/73.8	38.3/72.4	
Refinement				
Resolution range (Å)				
R_{factor} (%)	19.2	19.8	20.5	15.7
R_{free} (%)	20.0	20.5	24.4	16.6
Residues in the asymmetric unit	469	234	233	950
Water molecules	153	135	31	680
Temperature factor				
Average B values (Å^2)	34	29	39	16
Root-mean-square (r.m.s.) deviations				
Bond lengths (Å)	0.019	0.014	0.012	0.019
Bond angles (°)	1.942	1.463	1.396	1.880
Ramachandran plot				
Residues in most favored regions (%)	97	97	95	97
Residues in additional allowed regions (%)	3	3	5	3
Residues in generously allowed regions (%)	0	0	0	0

performed with 25 injections of 1.5 μL of the respective ligand solution (1.5 mM) every 180 s. Blank measurements were performed by injecting PBS without ligands. Two independent titrations were performed for each ligand.

The high-precision automated peak-shape analysis software NITPIC (Keller et al., 2012) was used with default parameters for the integration and baseline adjustments of all thermograms used in the present work. In addition to integrating the thermograms, NITPIC also assigns error bars to the integrated isotherms. The resulting isotherms were fitted with the SEDPHAT software (Houtman et al., 2007; Zhao et al., 2014), using the one site hetero-association (A+B ↔ AB) model with the following variables: incfA (fraction of binding-incompetent protein), log10(Ka) (logarithm of the macroscopic binding constant) and HAB (enthalpy change of the binding in kcal/mol). Non-linear regression fitting of the data was carried out using three fitting algorithms (Marquardt-Levenberg, Nelder-Mead simplex and simulated annealing) until all three converged at the same solution. This yielded values for Ka and the enthalpy change. Other thermodynamic parameters were calculated using the equation $\Delta G = -RT\ln(K_a) = \Delta H - T\Delta S$, where ΔG, ΔH and ΔS are the changes in free energy, enthalpy and entropy of binding, respectively. T is the absolute temperature, and R the ideal gas constant (1.98 cal mol^{-1} K^{-1}). Errors associated with the fitting parameters were estimated based on the covariance matrix.

2.7. Small angle X-ray scattering

SAXS data of rVML at different pH values were collected at the PETRA III, P12 beamline (DESY synchrotron, Hamburg) (Blanchet et al., 2015). Using a PILATUS 2 M pixel detector at a sample-detector distance of 3.1 m and at an energy of 10 keV (λ = 0.124 nm), the range of momentum transfer $0.1 < s < 0.45$ nm^{-1} was covered ($s = 4\pi \sin\theta/\lambda$, where 2θ is the scattering angle). Measurements were performed with protein solutions prepared in four different buffers, covering a large pH range (citrate, pH 3.0; phosphate, pH 5.0; HEPES, pH 7.0; CHES, pH 9.0), and at three different protein concentrations (1.0, 5.0 and 10.0 mg/mL). Buffers were prepared at 100 mM concentration and supplemented with NaCl 150 mM and 5% glycerol. The automated sample changer was employed to load the samples and constantly remove the irradiated sample. Twenty successive exposures of 50 milliseconds exposure time were collected and compared to detect and discard the radiation damage effects. The data was normalized to the intensity of the transmitted beam and radially averaged; the scattering of the buffer was

subtracted and the difference curves were scaled for protein concentration. These primary data processing steps were performed using the automated data pipeline SASFLOW (Franke and Svergun, 2009).

SAXS analysis was performed using various programs of the ATSAS 2.6 package (Petoukhov et al., 2012). The forward scattering I(0) and the radius of gyration R_g were extracted from the Guinier approximation calculated with the AutoRG function within PRIMUS (Konarev et al., 2003; Petoukhov et al., 2007). These parameters were also computed from the entire scattering patterns using the indirect transform package GNOM (Svergun, 1992), also providing the pair distribution function, p(r), of the particle and the maximum size D_{max}. The molecular weight (MW) of the solute was evaluated by comparison of the forward scattering with that from a reference solution of bovine serum albumin ($MW_{Monomer-Dimer}$ = 72 kDa; $I(0)_{BSA}$ = 1917). Further MW estimations were retrieved from the excluded volume calculated with Porod's law and from *ab initio* reconstructions as previously described (Petoukhov et al., 2012). The latter were generated with the DAMMIF program (Franke and Svergun, 2009). For the *ab initio* reconstruction of the complex, were performed 10 DAMMIF runs. To check the stability of the solution the individual models were superimposed onto each other (SUPCOMB (Kozin and Svergun, 2001)). Using the DAMAVER program (Volkov and Svergun, 2003) the common structural feature was determined and used as a starting model for a final round of DAMMIN (Svergun, 1999).

The theoretical scattering from the high resolution model of the rVML (PDB code: 4XTP) was calculated with the CRYSOL program (Svergun et al., 1995) and compared with the respective scattering profile. Using the Coral program, additional residues were added to the model to improve the general fit (Petoukhov et al., 2012). The Oligomer program (Konarev et al., 2003) was employed for a better description of the experimental scattering profile of rVML. Accordingly, the volume fractions of the protein in different oligomeric states were analyzed. Model representations displayed in the figures were generated with Pymol Graphics System (Schrödinger, LLC.).

2.8. Molecular docking calculations

Molecular docking calculations were performed to assess the ability of rVML to recognize glycoprotein fragments decorated with Tn antigen, similar to the previously conducted analysis for wtVML (Sousa et al., 2015). For this purpose, eight different fragments of human mucins were used, seven of MUC2, structurally characterized by nuclear magnetic resonance (NMR) (Borgert et al., 2012), and one of MUC1, obtained from its crystal complex with SBA (PDB ID 4D69) (Madariaga et al., 2014). MUC2 is a typical secreted, gel-forming mucin from the human gastrointestinal tract that has been linked to inflammation and cancer, while MUC1 consists of a transmembrane mucin from normal secretory epithelium that is found at high levels over the entire surface of diverse types of carcinoma cells (Kufe, 2009). The MUC2 fragments used were decorated with either three (Ac-Pro-Thr*-Thr-*Thr*-Pro-Leu-Lys-NH_2 – PDB ID 2LHW, MUC2-1), two (Ac-Pro-Thr-*Thr-*Thr-Pro-Leu-Lys-NH_2 – PDB ID 2LHZ, MUC2-2; Ac-Pro-Thr*-Thr-Thr*-Pro-Leu-Lys-NH_2 – PDB ID 2LHY, MUC2-3; Ac-Pro-Thr-Thr*-Thr*-Pro-Leu-Lys-NH_2 – PDB ID 2LHX, MUC2-4) or one (Ac-Pro-Thr*-Thr-Thr-Pro-Leu-Lys-NH_2 – PDB ID 2LI2, MUC2-5; Ac-Pro-Thr-Thr*-Thr-Pro-Leu-Lys-NH_2 – PDB ID 2LI1, MUC2-6; Ac-Pro-Thr-Thr-Thr*-Pro-Leu-Lys-NH_2 – PDB ID 2LI0, MUC2-7) α-*O*-GalNAc units covalently linked to threonine residues (Tn antigen), while the MUC1 fragment (MUC1-2) presents a single glycosylation (Pro-Asp-Thr*-Arg), as indicated by *. For the purpose of clarity, Tn residues located at Thr2, Thr3 and Thr4 on MUC2 fragments will hereinafter be designated as Tn1, Tn2 and Tn3, respectively.

The same set of calculations involving these mucin fragments was also performed with five different Tn-binding legume lectin structures. The analyzed lectins were VVLB4 (*V. villosa*) (Babino et al., 2003), SBA (*G. max*) (Dessen, 1995), DBL (*D. biflorus*) (Hamelryck et al., 1999), WBA I (*P. tetragonolobus*) (Kulkarni et al., 2005) and GSA-I-B4 (isolectin IV from *Griffonia simplicifolia*) (Tempel et al., 2002). The PDB IDs of the structures investigated in this study are as follows: 1N47 (VVLB4), 1SBF (SBA), 1LU2 (DBL), 2D3S (WBA I) and 1HQL (GSA). The files were modified to contain coordinates related to a single polypeptide chain. Similarly, the files of MUC2 fragments were also changed to present only one NMR model.

Molecular docking calculations were performed in AutoDock Vina, version 1.1.2, which applies an iterated local search global optimizer for the optimization procedure, where the succession of each step consists of a mutation and local optimization, with the acceptance decisions made according to the Metropolis criterion. It uses the efficient quasi-Newton method Broyden–Fletcher–Goldfarb–Shanno (BFGS) for local optimization (Trott and Olson, 2009). The Autodock graphical interface AutoDockTools, version 1.5.6, was used to keep polar hydrogens and add partial charges to the proteins and ligands using the Kollman United charges (Morris et al., 1998). The proteins and carbohydrate ligands were treated as rigid and flexible molecules, respectively. The search space for the docking calculations, selected based on the mucin fragments' length, was defined by a 50 Å × 50 Å × 50 Å cube centered on the lectin binding site. Exhaustiveness was set to 12, and all other parameters were used as default. For each ligand, the nine top-ranked generations based on the predicted binding affinity (in kilocalories per mole) were analyzed. The solutions were first chosen based on the coordination of Tn residues at the CBS, considering as essential interactions for carbohydrate binding: (i) hydrogen bonds between hydroxyls 3-OH and 4-OH to a specific aspartic acid residue and (ii) the stabilization of the N-acetyl group (Rougé et al., 2011). Results matching these criteria were further ranked based on the theoretical binding energy (given as a negative score in kcal/mol).

Complex ligands such as mucin fragments present a considerable amount of rotatable bonds that may hamper the detection of appropriate solutions for docking calculations (Trott and Olson, 2009). In view of that, the crystal structure of the complex between SBA and MUC1-2 fragment was used for a redocking calculation and the RMSD (root mean square deviation) of the atomic positions between the best docking result, following the above criteria, and the crystal model was analyzed to validate the system.

3. Results

3.1. Structure analysis

High expression levels were observed for rVML, with an average yield of 90 mg of pure protein per liter of induced cell culture. Nevertheless, despite extensive crystallization attempts, diffracting crystals were only obtained for the metal-containing, uncomplexed form of rVML and for the co-crystallized complexes with GalNAc (rVML-GalNAc), α-lactose (rVML-Lac) and the Tn antigen (rVML-Tn). Diffraction data were collected from these crystals to resolutions of 1.90, 2.70, 1.83 and 1.97 Å, respectively. The recombinant lectin was obtained as a single-chain protein with the same amino acid sequence as wtVML, but with two (Met and Ala) and eight (Leu, Glu and 6xHis) extra residues added to its N- and C-terminus, respectively, comprising 250 amino acids. The presence of extra amino acids in recombinant protein sequences, often as a

result of cloning strategies, should mostly be avoided for specific biological or structural purposes, but they may also confer extra stability to intrinsically flexible termini. As previously reported, the crystal structure of wtVML presented no electron density for the last six C-terminal residues, which was probably caused by the flexibility of this portion. Based on that, no residues were removed from rVML as an attempt to obtain further structural information about these missing residues. Fortunately, two monomers from rVML-Lac presented density for all six C-terminal residues, whereas the other structures analyzed in this study show electron density for only one or two additional residues. The structure of rVML-Lac thus provides full structural information for the whole protein sequence. No density related to the fusion tag (Leu-Glu-6xHis) was observed.

Crystals were obtained in three different space groups ($P3_121$, C2 and I222) (see Table 1), but no significant changes in the protein backbone structure were observed, indicating a negligible influence of crystal packing or ligand binding on the overall lectin structure. Nevertheless, clear evidence of structure stabilization is given by the faster growth of crystals for the complexes with GalNAc and Tn antigen (overnight) when compared to other crystallization assays (2 weeks on average). This is most likely a consequence of the higher number of interactions established because of the N-acetyl substituent at C-2. The asymmetric unit contains a single monomer in the case of rVML-GalNAc and rVML-Tn, a DB58 dimer (dimerization pattern first observed for the vegetative lectin of *D. biflorus*) in the case of uncomplexed rVML, and a complete tetramer in the case of rVML-Lac (built by two opposite DB58 dimers) (Buts et al., 2001). All structures contain two divalent metal ions (Ca^{2+} and Mn^{2+}) per subunit and a ligand molecule bound to the lectin CBS in the case of co-crystallized structures. For uncomplexed rVML, glycerol molecules are bound at both available CBS (DB58 dimer), illustrating how different molecules can also be hosted by lectins in the absence of specific ligands (Bezerra et al., 2011; Delatorre et al., 2013).

Additionally, since rVML was produced as a single-chain lectin, clear electron density is observed for the loop built up by the Asn113-Gln120 segment, which is poorly defined in the wild-type lectin structure, most likely caused by its flexibility. This loop corresponds to a glycosylation/cleavage site for wtVML related to its natural processing (Sousa et al., 2015).

3.2. Overall structure

The tertiary structure of the rVML monomer displays the characteristic legume lectin fold that consists of a β-sandwich presenting the jelly roll motif made up of a flat six-stranded and a curved seven-stranded β-sheet interconnected by loops of variable length (Fig. 1a) (Loris et al., 1998). The models have good stereochemistry, with most non-Pro and non-Gly residues displaying main-chain dihedral angles that fall within the most favoured or additionally allowed regions of the Ramachandran plot (Laskowski et al., 1993).

The carbohydrate binding sites are well structured in all crystal models and present the same residue configuration as the wild-type lectin. The CBS consists of a shallow cavity in the protein surface (Fig. 1b) formed by four separate loops: Glu80-Asp89, Ala96-Thr121, Asp127-His139 and Ser213-Thr222. The four key residues (Asp89, Gly107, Phe129 and Asn131) to which the carbohydrate binding capability is attributed occupy similar positions in comparison to other legume lectins (Loris et al., 1998). The remaining residues which compose the CBS and to which the carbohydrate specificity is attributed are Ser216 and Leu215. As seen for wtVML, His219 also contributes to the formation of CBS boundaries in rVML, providing the co-crystallized ligands with better anchorage to the 6-OH hydroxyl group through van der Waals interactions.

Similarly to other legume lectins, the CBS on rVML is located near to a conserved metal binding site, presenting the divalent ions Ca^{2+} and Mn^{2+} (Fig. 1c). The presence of these ions allows the formation of a non-proline *cis* peptide bond between Ala87 and Asp88, which is essential for CBS proper fold. It is well known among legume lectins that the absence of these divalent metals induces a *cis* to *trans* isomerization of the above mentioned peptide bond, which is incompatible with carbohydrate binding (Garcia-Pino et al., 2006). Studies involving the demetallization of Concanavalin A have reported that other divalent ions (e.g. Zn^{2+} and Cd^{2+}) are also able to bind the metal binding site, although affecting the carbohydrate binding (Bouckaert et al., 2000). At rVML's metal binding site, the Ca^{2+} ion is coordinated by the main chain carbonyl of Phe129 and the side chains of Asp127, Asn131 and Asp134, whereas the Mn^{2+} ion is coordinated by the carboxylate groups of Glu125, Asp127 and Asp134, as well as the imidazole group of His139. The coordination spheres of the ions are completed by two structural waters for each metal ion, clearly visible in the electron density (Fig. 1c).

Table 2
Comparison of hydrogen bonding distance in the coordination of Tn antigen to wtVML and rVML. Numbering for rVML residues is indicated within brackets.

Tn antigen group	VML group	Distances (Å)	
		wtVML[a]	rVML
3-OH	ND2 / Asn129 (131)	2.92	2.99
3-OH	N / Gly105 (107)	2.95	2.73
3-OH	OD1 / Asp87 (89)	2.51	2.60
4-OH	OD2 / Asp87 (89)	2.72	2.53
4-OH	N / Leu213 (215)	3.1	3.28
6-OH	OG / Ser214 (216)	2.62	2.77
7-OH	N / Gly105 (107)	2.94	2.85

[a] Sousa et al., 2015.

Despite presenting different asymmetric unit contents, all crystal lattices are built up by tetramers (Fig. 1a), which can be observed directly in the case of rVML-Lac and are generated by crystal symmetry operations for all other structures. As previously reported, tetramers represent the natural biological assembly of wtVML and are an essential feature for the biological role of lectins (Calvete et al., 1998). The tetramerization pattern was further confirmed by a PISA server analysis, with oligomers composed of two DB58 dimers (X1 interface), conserving the well-structured signature motif Ser-Tyr-Ile-Val-Ser around residue 188 (186 for wtVML), previously reported as typical for this type of dimer (Brinda et al., 2004). These observations allow us to conclude that the structures presented here are in close agreement with real biological assemblies.

As mentioned before, the C-terminal residues not detected in wtVML are clearly seen for two monomers in rVML-Lac, providing a better understanding of the lectin's overall structure (Supplementary Fig. S1). The central channel formed upon tetramerization is partially blocked by the C-terminus, establishing contacts with both opposite monomers involved in the tetramer interface (type II or canonical). Based on that, we can assume that the lectin tetramer in solution presents partially free C-termini which contribute to oligomer stabilization, acting like dynamic cage bars that would prevent the access of large molecules to the channel. This assumption is corroborated by the exclusive presence of water molecules in the central channel in all wild-type and recombinant VML structures analyzed so far.

3.3. rVML – GalNac / Tn antigen / α-lactose interactions

The set of interactions established by rVML with GalNAc and Tn antigen are similar to that observed for wtVML (Fig. 2a), with minor differences with respect to hydrogen bond lengths (Table 2). The sugar pyranose ring is stacked between the aromatic side chain of Phe129 and the aliphatic side chain of Leu215. Hydrogen bonds are established by hydroxyls 3-OH, 4-OH and 6-OH, as well as by

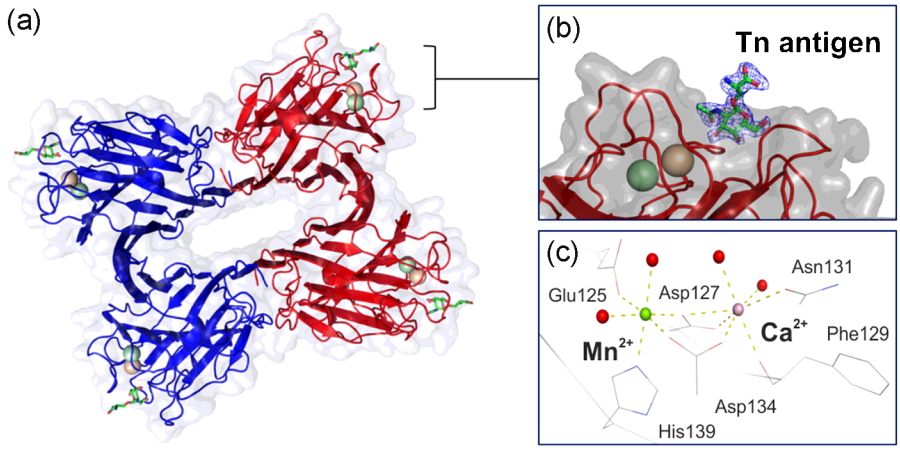

Fig. 1. Overall structure of rVML tetramer complexed with Tn antigen. (a) – Peptide chains are represented as ribbons colored in blue (dimer A) and red (dimer B) within the visible surface (light blue). (b) – Carbohydrate binding site of rVML showing an omit map (contoured at 1.5σ) of the Tn antigen (green sticks). Calcium and manganese ions are represented as spheres colored in wheat and green, respectively. (c) – Metal binding site presenting all residues (grey lines) involved in metal ion coordination (yellow dashes). Calcium and manganese ions, as well as water molecules, are represented as spheres coloured in pink, green and red, respectively.

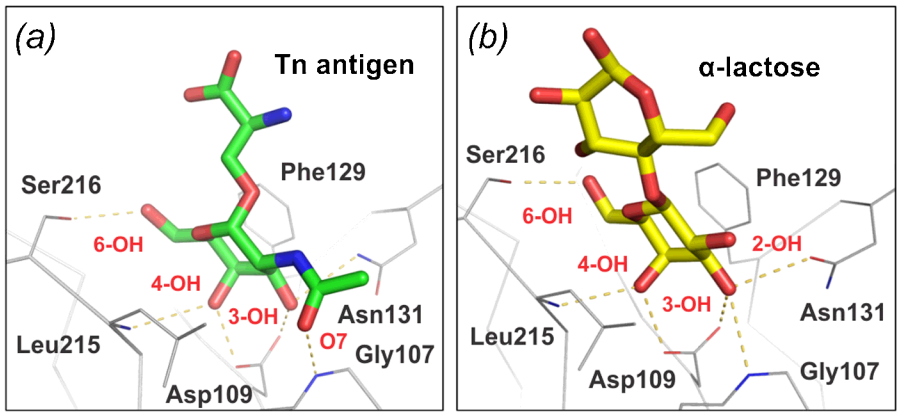

Fig. 2. rVML carbohydrate binding site ribbon representation showing the hydrogen bonding interactions with Tn antigen (a) and α-lactose (b). The polypeptide chains and the amino acid residues involved in ligand coordination are represented as ribbons and lines colored in grey; Tn antigen and α-lactose are represented as sticks colored in green and yellow; and hydrogen bonds are represented as yellow dashes, respectively.

the carbonyl group (O7), following the same pattern as previously described (Sousa et al., 2015). The hydrophobic pocket formed by the side chains of Phe108 and Trp133 accomodates the N-acetyl group at the C-2 of both ligands, equivalent to the situation in wtVML and other GalNAc-binding lectins.

Crystals for rVML-Tn and rVML-GalNAc were obtained in the same crystallization conditions as the wild-type lectin crystals (citrate buffer pH 3.5 and ammonium sulphate). Therefore, similar to wtVML, sulphate ions are observed bound to both recombinant structures (seven for rVML-Tn and one for rVML-GalNAc), most likely replacing structural waters. In the case of the complex with Tn antigen, hydrogen bonds to the seryl moiety (serine residue) are established solely through one of these sulphate molecules that links the Tn ammonium group to the NH of Thr132 and ND2 of Asn131. In both structures, citrate ions are also observed to be bound at a secondary site close to the CBS, presenting the same coordination as seen for wtVML.

The coordination of α-lactose (β-D-Galactose-1→4-α-D-Glucose) by rVML occurs mainly through the galactose moiety of this disaccharide (Fig. 2b). In a manner similar to that of other monosaccharides bound to CBS, the pyranose ring of galactose is stacked by Phe129 and Leu215. The hydrogen bonding pattern established for the galactose moiety differs from that observed for GalNAc, since no substituent is placed at C-2. In this case, the direct hydrogen bond formed between the carbonyl group (O7) in the N-acetyl substituent to Gly107 NH group is replaced by a water bridge to hydroxyl 2-OH. Additionally, the glucose moiety of α-lactose is stabilized by van der Waals interactions between hydroxyls 2-OH and 3-OH, as well as the glycosidic bond O, and the main chain carbonyl of Leu215. These weak interactions allow the glucose moiety to adopt distinct conformations, as can be seen in the different subunits of rVML-Lac tetramer. Water bridges may also play a role in the coordination of glucose, as observed between Leu215 main chain and hydroxyl 2-OH in one monomer.

3.4. Calorimetric assays

The binding affinities of both wild-type and recombinant VML were compared through ITC assays to assess suitability of the recombinant lectin as a tool for biological purposes (Fig. 3). The analysis indicates that both forms of the lectin behave similarly, as seen by the K_d values (Table 3), indicating that the wild-type lectin (K_d of 29.26 ± 5.07) is able to bind the Tn antigen (GalNAcα1-O-Ser) only marginally better when compared to its recombinant counterpart (K_d of 55.40 ± 2.84), with even lower differences for binding with GalNAc (K_d of 31.40 ± 0.75 and 26.97 ± 0.89, respectively).

These results are in good agreement with previously reported data (Dam et al., 2007) and are corroborated by structural alignments of the binding sites (and their boundaries) from both lectin forms (data not shown), where only minor differences are detected,

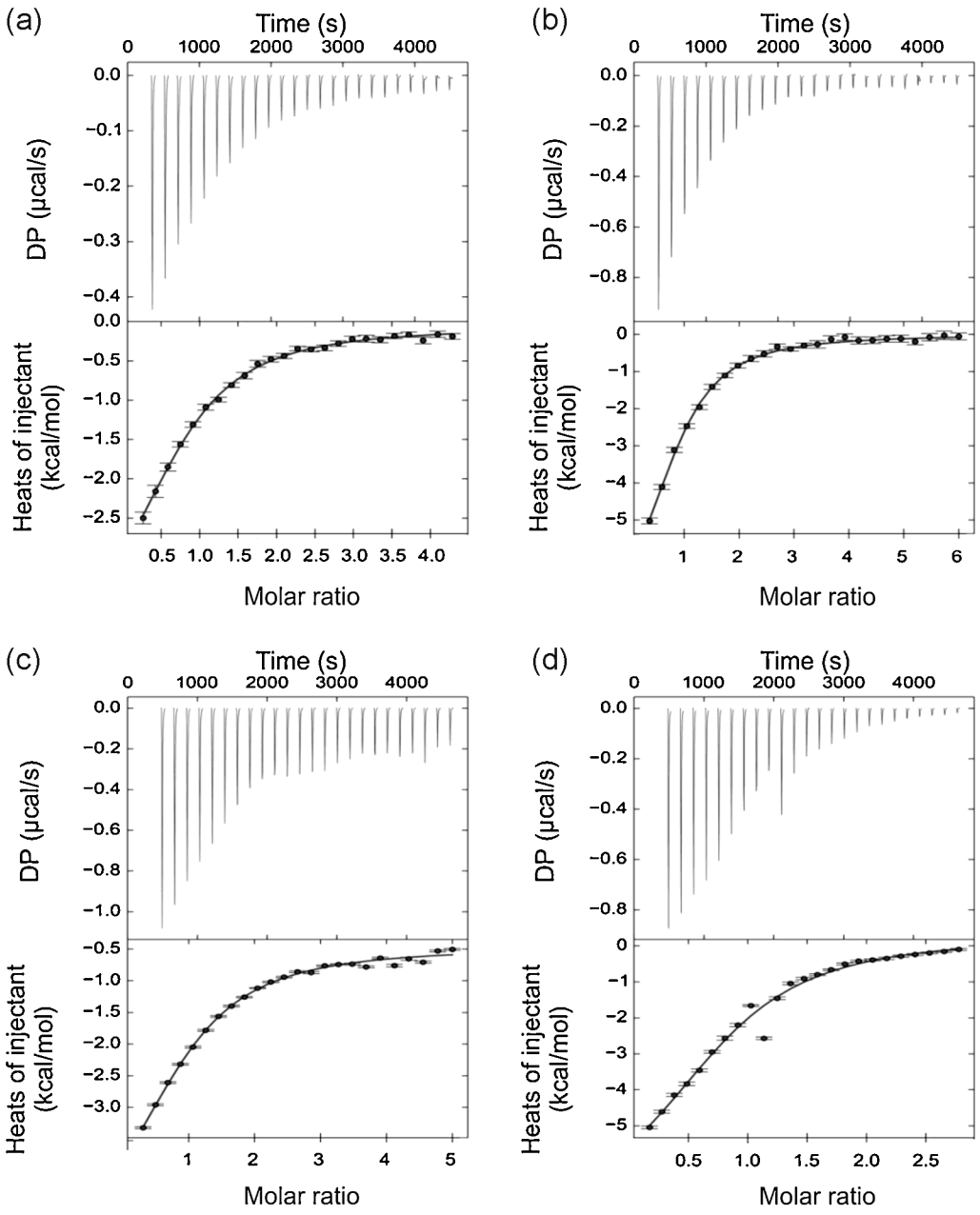

Fig. 3. Representative Isothermal Titration Calorimetry (ITC) profiles for wtVML with Tn antigen (a) and GalNAc (b), as well as for rVML with Tn antigen (c) and GalNAc (d). The top panel of each figure shows the thermogram, the middle panel shows the fitted isotherm obtained using the SEDPHAT program, and the bottom panel shows the residuals of each fit.

which are probably related to the subtle alterations among the hydrogen bond lengths as described earlier (Table 2). These results prove that the recombinant production of VML as a single-chain protein does not influence lectin binding capability.

3.5. *Small angle X-ray scattering*

To gain further insights into the oligomerization state of rVML, we performed small angle X-ray scattering (SAXS) experiments. Concentration series were measured at different pH values (pH 3, 5, 7, and 9). For all the samples measured, there was no obvious concentration dependency despite a slight tendency to form aggregates at higher concentrations. This observation was more severe at the higher pH values (pH 7 and 9). For the following analysis the SAXS profile collected at pH 5 was used by merging the data from the highest concentration (10 mg/mL) and the lowest (1 mg/mL). The primary parameters derived from the scattering profile are summarized in Supplementary Table S1. The molecular models and experimental SAXS data have been deposited on *SASBDB* (Small Angle Scattering Biological Data Bank; accession code: SASBDBS2).

The scattering profiles collected at different pH values were in good agreement to each other with the exception of the q-range 1–1.5 nm^{-1}. Here, the profile measured at pH 9 showed the biggest difference (Fig. 4a). Running the Oligomer program indicated that a good fit could be achieved by including around 15% of monomeric fraction for this pH value.

Ab initio models, as well as molecular weight estimations for rVML suggested that the sample is mostly tetrameric in solution (Fig. 4b). Indeed, the theoretical scattering curve of the tetrameric assembly of the crystal structure for rVML-Tn (PDB ID 4XTP), showed a reasonable good fit to experimental data (Chi^2 (Crysol): 4.5), whereby the curves coincide extremely well at low angles

Table 3
Thermodynamic binding parameters for wtVML and rVML.

Protein	Ligand	K_d	ΔG	ΔH	$T\Delta S$	T (K)
wtVML		(μM)	(kcal/mol)	(kcal/mol)	(kcal/mol)	
	Tn antigen	29.26 ± 5.07	−6.16 ± 0.10	-4.63 ± 0.47	1.53 ± 0.48	298
	GalNAc	31.40 ± 0.75	−6.12 ± 0.01	-6.66 ± 0.04	−0.54 ± 0.05	298
rVML						
	Tn antigen	55.40 ± 2.84	−5.79 ± 0.03	−5.11 ± 0.17	0.67 ± 0.18	298
	GalNAc	26.97 ± 0.89	−6.21 ± 0.02	−7.18 ± 0.13	−0.97 ± 0.14	298

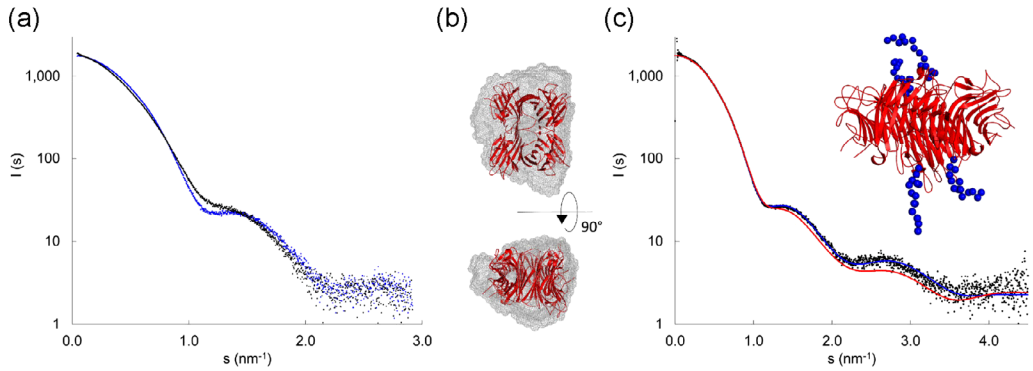

Fig. 4. SAXS analysis of rVML. (a) Comparison of data collected at pH 5 (blue) and pH 9 (black). (b) Overlay of the ab initio model (mesh) with the crystal structure (red ribbon). (c) The inlay shows the crystal structure as red ribbon representation and the added dummy model residues are shown as blue spheres. The theoretical scattering profiles of the crystal structure (red) and the extended model (blue) are compared with the experimental scattering profile (black).

($q < 1\ nm^{-1}$). This suggests, that the overall shape and size of the protein in solution do, indeed, agree very well with the crystal structure. However, some variations of the internal structure in solution are observed. Therefore, the fit could, for example, be improved by adding the eight missing residues for the recombinant sequence (Pro-Ser-Asp-Asp-Ser-Asn and the additional residues Leu-Glu) at the C-terminal (Chi^2 (Crysol): 2.8), and even more improvement could be achieved by first removing the last six detected C-terminal residues from the crystal structure, which seemed rather rigid as a result of crystal packaging, and then modelling 14 additional residues with the Coral program (Chi^2 (Crysol): 0.9; Fig. 4c).

3.6. *In silico analysis of legume lectins as biological tools for cancer research*

The possibility of using rVML as a tool for cancer research (diagnostic/prognostic) was explored through molecular docking calculations performed between this lectin and eight fragments of human mucin (MUC2 and MUC1) differently decorated with Tn antigen, which have been widely used as biomarkers of cancer progression (Kudelka et al., 2015).

The previous analysis involving wtVML and MUC2 fragments was based on a rigid docking calculation, where both lectin and mucin fragments were treated as rigid bodies, and was carried out considering the correlation type search as the sum between shape and electrostatic forces. The obtained results suggested that wtVML was able to bind only MUC2-1 (Sousa et al., 2015). Nevertheless, a more extensive recognition pattern was observed when mucin fragments were considered as flexible molecules. In fact, the current analysis has detected that rVML is able to bind fragments MUC2-1, MUC2-3, MUC2-4 and MUC2-7, coordinating the Tn residue located at Thr4 of each fragment (Tn3), similar to that observed for the crystal structures. However, the peptide backbone orientation is not conserved. MUC2-3 and MUC2-4 present orientations similar to that observed for the complex between wtVML and fragment 1 (mode A) (Sousa et al., 2015), with three hydrophobic clusters conserved in the former complex, but only two in the

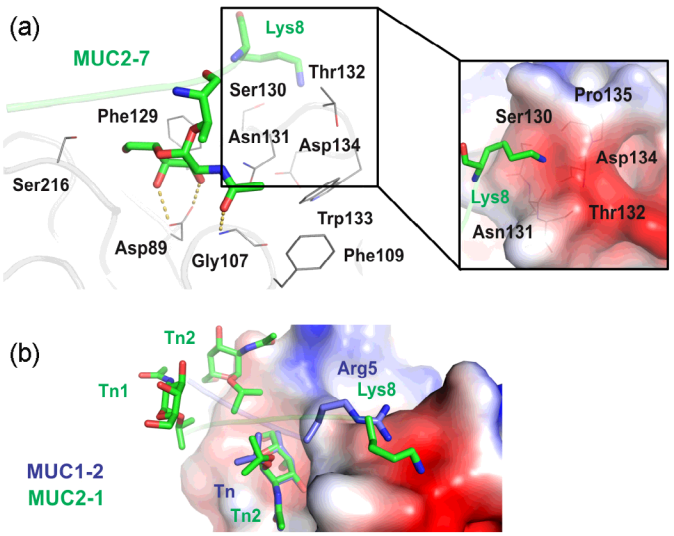

Fig. 5. Results of the docking simulations for rVML. (a) Orientation B of MUC2-7 with the Tn moiety accommodated into lectin CBS. Also shown is the orientation of Lys8 toward the negatively charged region (inset) formed by the side chain of Ser130 and the main chain carbonyl groups of Asn131, Thr132, Asp134 and Pro135. (b) Binding of MUC2-1 (green) and MUC1-1 (blue) to rVML CBS. Their respective C-terminal positive residues (Lys8 and Arg5, respectively) are oriented toward the negative region, which seems to be the structural factor involved in the more energetically favourable accommodation mode B (see text).

latter. On the other hand, MUC2-7 presents the Lys8 residue oriented to a negatively charged region close to the binding site and formed by the side chain hydroxyl (OG) from Ser128 and the main chain carbonyl groups from Asn131, Thr132, Asp134 and Pro135 (mode B) (Fig. 5). Interestingly, the theoretical binding energies for these three fragments are almost identical (Table 4), indicating that both backbone orientations are energetically favourable. Furthermore, despite the similarity within the results obtained for binding of MUC2-1, MUC2-3 and MUC2-4 to rVML, a stronger binding energy is attributed to fragment 1 (Table 4), which could be

Table 4
Theoretical binding energies of the molecular docking calculations between the tested lectins and mucin fragments. Values in parentheses indicate the best results.

Fragment	Energy (Kcal/mol)						
	rVML	wtVML	VVLB4	SBA	WBA I	DBL	GSA
MUC2-1	−7.1 (1)	−6.8 (1)	−5.3 (2)	−5.8 (5)	−5.9 (5)	–	–
MUC2-2	–	–	−7.0 (5)	–	−7.0 [a] (1)	–	–
MUC2-3	−5.8 (8)	−5.8 (6)	−6.1 (2)	−6.2 (1)	−6.5 (3)	−5.6[a] (6)	–
MUC2-4	−5.7 (3)	−7.2 (3)	−5.8 (8)	–	−5.8 (6)	–	–
MUC2-5	–	–	–	−6.2 (9)	−6.7 (1)	−5.9 (8)	–
MUC2-6	–	–	–	–	–	–	–
MUC2-7	−5.8 (9)	–	−6.4 (4)	–	−7.2 (1)	−6.7 (2)	–
MUC1-2	−7.3 (1)	−6.8 (3)	−6.6 (5)	−6.6 (2)	−7.2 (2)	−6.0 (5)	–

[a] Binding through Tn2.

explained by electrostatic interactions between residues Lys8 on MUC2-1 and Asp105 on the lectin, highlighting the high flexibility of lysine residues.

Surprisingly, wtVML was not able to recognize MUC2-7 fragment at any of the above described orientations, even when the exhaustiveness parameter was set to 15. This fragment exhibits a high flexibility, mainly caused by its single glycosylation (Borgert et al., 2012), which probably hinders a stable coordination by wtVML, as also seen for SBA. Nevertheless, docking results indicate that rVML is able to interact with this fragment (with low energy) (Table 4), despite presenting the same groups as wtVML. This fact is probably caused by minor alterations that affect the protein structure and the orientation of side chains, generating different binding site volumes and variable bond lengths for ligand interaction (see Table 2).

As expected, docking calculations following the current strategy demonstrated that MUC2-1 fragment interacts with wtVML in a manner similar to that observed for its recombinant counterpart (mode B), but different from previously reported rigid analysis (mode A) (Sousa et al., 2015). The obtained binding energies (Table 4) indicate that both orientations may occur simultaneously, pointing to a versatile binding mechanism that seems to be modulated by the peptide sequence close to the glycosylation point (Madariaga et al., 2014), as well as by the density of Tn residues for each fragment. The role of the above-described negatively charged region located close to the lectin CBS also becomes evident when docking MUC1-2 fragment to rVML, where interactions established to residue Arg5 on the peptide are clearly visualized (Fig. 5). This feature strongly resembles what was previously reported for the crystal structure of SBA-MUC1-2 complex, where the interaction between residues Arg5 on the peptide and Glu113 on SBA favours a stronger anchorage of this fragment (Madariaga et al., 2014). For comparison purposes, the crystal structure of SBA was also used for docking calculations following the same strategy. Interestingly, the interactions involving residue Arg5 observed for the crystal complex were not detected after the calculations, most likely reflecting the high flexibility of this residue. This observation also indicates a more favourable interaction of MUC1-2 to rVML (Table 4), mostly related to this negatively charged area. This region creates an electrostatically more attractive environment for positively charged residues on the fragment, allowing a conformational stabilization that would hardly be achieved through interactions with a single amino acid residue, as seen for SBA.

As previously reported (Sousa et al., 2015), we have also compared herein the binding modes of rVML to mucin fragments with those observed for other Tn-binding legume lectins, using the current set of fragments and docking strategy. The comparative analysis showed that the CBS structure per se and, most importantly, the residues composing its boundaries greatly influence the capacity of these lectins to interact with different mucin fragments (Madariaga et al., 2014). MUC2-1 illustrates this fact by binding solely to rVML, VVLB4, SBA, and, with a lower energy, to WBA I (Table 4). The orientation of this fragment follows mode A for VVLB4 and SBA and mode B for rVML, as discussed above, suggesting that the negatively charged region of rVML, which is weakly charged for the other two lectins (Supplementary Fig. S2), promotes the second orientation that seems to be energetically more favourable. On the other hand, the lower binding energy of WBA I for MUC1-2 appears to be linked to (i) an upward conformation of Lys216 induced by Pro213 and (ii) the absence of any negatively charged spot on CBS boundaries, as observed for rVML, preventing an energetically more favourable coordination of this fragment (Supplementary Fig. S3).

The importance of CBS boundaries for the binding of complex ligands is supported by the lack of proper interactions between MUC2-1 and either GSA or DBL. For GSA, the presence of a bulky glutamic acid at position 106, which is normally replaced by a glycine in other lectins, prevents the accommodation of the N-acetyl group from Tn residues, making its binding energetically unfavourable (Tempel et al., 2002), in fact, no fragments of MUC2 or MUC1-2 bound to this lectin (Supplementary Fig. S4). In the case of DBL, the replacement of an aromatic residue at CBS by a leucine and the insertion of a second amino acid between the asparagine and tryptophan residues, which also compose the binding site, seems to interfere with anchorage of complex ligands (Supplementary Fig. S5), such as glycoproteins, by disrupting hydrophobic cluster 1 (Sousa et al., 2015), despite not affecting the binding of GalNAc/Tn moieties (Hamelryck et al., 1999). Furthermore, this insertion partially disrupts the negatively charged region, which would prevent the coordination of positive residues at this spot (Supplementary Fig. S3), explaining the lower binding energy of DBL for this construct (Table 4).

The docking system (Autodock Vina) suitability for complex ligands (such as mucin fragments) was analyzed based on a redocking calculation involving the crystal structure of the complex between SBA and MUC1-2 (Madariaga et al., 2014). The calculated RMSD between the best docking result for MUC1-2 (found following the same used criteria) and the crystal structure was 2.5, which decreases to 2.0 after removing the flexible Arg5 residue. Thus, despite the considerable amount of rotatable bounds present in the mucin fragments, the adopted system was considered reliable for the current purpose.

4. Discussion

Herein we present the structural characterization of a recombinant Tn antigen-binding lectin from *V. macrocarpa* (rVML), which retains all the main features of its wild-type counterpart, including carbohydrate specificity and oligomer stability in a wide pH range. Four different crystal structures, which were obtained in different space groups, have been determined for this recombinant lectin, including complexes with GalNAc, α-lactose and Tn antigen. These models represent the first structures of a recombinant Tn-binding legume lectin, providing important structural information that has

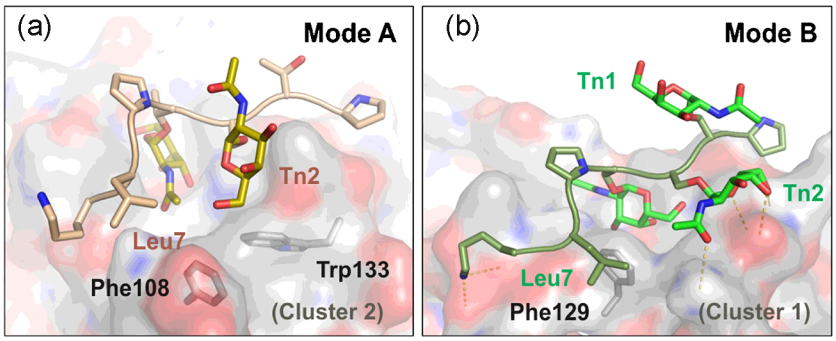

Fig. 6. Hydrophobic clusters involved in the coordination of *O*-glycosylated fragments. (a) Tn2 and Leu7 (MUC2-4) are stabilized by a hydrophobic stacking coordinated by Cluster 2. (b) Mutual stacking between Tn 1 and Tn2 expose polar groups within Tn moieties, while Leu7 is coordinated by Cluster 1 (MUC2-7). Leu7 also interacts with the non polar portion of Lys8 side chain, generating a more compact conformation than that observed for mode B. Hydrogen bonds are represented as yellow dashes.

not been detected for wtVML, as well as emphasizing the reliability of rVML for biological applications.

In silico assays have demonstrated the high potential of rVML as a new tool for cancer research through a comparative structural analysis involving different legume lectins widely used for cancer diagnosis and prognosis. The obtained results show that rVML presents a differential selectivity for glycoproteins with a high density of Tn residues, which agrees with previously reported results by Dam and co-workers for wtVML (Dam et al., 2007). Nevertheless, the high theoretical energy detected for the anchorage of single glycosylated MUC1-2 fragment to rVML indicates a great dependency of binding site boundaries in determining lectin affinity for specific antigens that interact with amino acid groups close to the glycosylation point on these epitopes. All lectins analyzed have demonstrated this same structural dependency, recognizing different groups of glycoprotein fragments with variable binding energies and stressing a differential potential for antigen detection for each one.

The coordination of all fragments herein analyzed seems to be closely related to the stabilization of their hydrophobic groups, especially the pyranose rings and N-acetyl groups from Tn moieties, as well as Leu7 for MUC2 fragments (Fig. 6). These hydrophobic groups are mainly stabilized by the previously described hydrophobic clusters 1 (Phe129) and 2 (Phe108 and Trp133) (Sousa et al., 2015), but also by interacting with each other or with other parts of the fragment, adopting more or less compact conformations, depending on the availability and disposition of anchor groups on the protein surface (Fig. 6a). Conformational modes A and B of MUC2 fragments are clearly stabilized through a hydrophobic pattern of interactions involving clusters 1 and 2, respectively.

The presented results strongly support that legume lectins possess different potentials for biological purposes, even among those commonly used as histological markers, such as SBA and VVLB4. Despite sharing the same monosaccharide affinity, these GalNAc binding lectins present variations within binding site boundaries, allowing them to recognize different specific epitopes spread over different types of cancer cells. These differences are mainly related to the disposition of hydrophobic clusters and charged regions around the lectin carbohydrate binding site, which favours the anchorage of different groups close to Tn moieties on glycoproteins. This feature is directly dependent on the epitope's amino acid sequence, as previously postulated (Madariaga et al., 2014), as well as on the density of Tn residues along the glycoprotein. The clustering of specific glycans induced by its higher density may increase the avidity of lectins for particular glycosylated receptors, raising their local concentrations and modulating their biological activities by favoring a multimeric organization on cell surfaces (Elola et al., 2015).

Based on these findings, rVML has been presented here as a new tool for cancer research, free of isoforms and easily purified at high yields, with binding profiles comparable to SBA and VVLB4. Biological assays are clearly necessary to confirm these findings and to deeply explore the potential of rVML and other legume lectins as biological tools.

Acknowledgments

Results shown in this report are derived from work performed at European Synchrotron Radiation Facility (Grenoble, France). We thank Dr. Dmitri Svergun (EMBL/DESY) for technical support in the SAXS analysis, the Vienna International PostDoctoral Program (VIPS), and David Martin for helping with the English language editing of the manuscript. This research was supported by the Fundação Cearense de Apoio ao Desenvolvimento Científico e Tecnológico - FUNCAP, Conselho Nacional de Desenvolvimento Científico e Tecnológico - CNPq, Coordenação de Aperfeiçoamento de Pessoal de Nível Superior – CAPES.

Appendix A. Supplementary data

Supplementary data associated with this article can be found, in the online version, at http://dx.doi.org/10.1016/j.biocel.2015.12.016.

References

Alencar, N.M.N., Assreuy, A.M.S., Criddle, D.N., Souza, E.P., Soares, P.M.G., Havt, A., et al., 2004. Vatairea macrocarpa lectin induces paw edema with leukocyte infiltration. Protein Pept. Lett. 11, 195–200.

Alencar, N.M.N., Assreuy, A.M.S., Havt, A., Benevides, R.G., de Moura, T.R., de Sousa, R.B., et al., 2007. Vatairea macrocarpa (Leguminosae) lectin activates cultured macrophages to release chemotactic mediators. Naunyn Schmiedebergs Arch. Pharmacol. 374, 275–282.

Altman, F., 2007. The role of protein glycosylation in allergy. Int. Arch. Allergy Immunol. 142, 99–115.

Alves-Filho, J., Garcia Nascimento, A.S., Gondim, A.C., Pereira, R.I., Maranguape, R., et al., 2013. Isoform Characterisation, heterologous expression and functional analysis of two lectins from vatairea macrocarpa. Protein Pept. Lett. 20, 1204–1210.

Babino, A., Tello, D., Rojas, A., Bay, S., Osinaga, E., Alzari, P.M., 2003. The crystal structure of a plant lectin in complex with the Tn antigen. FEBS Lett. 536, 106–110.

Bakalova, R., Ohba, H., 2003. Interaction of soybean agglutinin with leukemic T-cells and its use for their in vitro separation from normal lymphocytes by lectin-affinity chromatography. Biomed. Chromatogr. 17, 239–249.

Bezerra, E.H.S., Rocha, B.A.M., Nagano, C.S., Bezerra, G.D.A., Moura, T.R.De, Bezerra, M.J.B., et al., 2011. Structural analysis of ConBr reveals molecular correlation between the carbohydrate recognition domain and endothelial NO synthase activation. Biochem. Biophys. Res. Commun. 408, 566–570.

Blanchet, C.E., Spilotros, A., Schwemmer, F., Graewert, M.a., Kikhney, A., Jeffries, C.M., et al., 2015. Versatile sample environments and automation for biological

solution X-ray scattering experiments at the P12 beamline (PETRA III. DESY). J. Appl. Crystallogr. 48, 431–443.
Borgert, A., Heimburg-Molinaro, J., Song, X., Lasanajak, Y., Ju, T., Liu, M., et al., 2012. Deciphering structural elements of mucin glycoprotein recognition. ACS Chem. Biol. 7, 1031–1039.
Bouckaert, J., Loris, R., Wyns, L., 2000. Zinc/calcium and cadmium/cadmium substituted concanavalin A: interplay of metal binding, pH and molecular packing. Acta Crystallogr. D Biol. Crystallogr. 56, 1569–1576.
Brinda, K.V., Mitra, N., Surolia, A., 2004. Determinants of quaternary association in legume lectins. Prot. Sci. 13, 1735–1749.
Brooks, C.L., Schietinger, A., Borisova, S.N., Kufer, P., Okon, M., Hirama, T., et al., 2010. Antibody recognition of a unique tumor-specific glycopeptide antigen. Proc. Natl. Acad. Sci. USA 107, 10056–10061.
Brooks, S., Carter, T.M., 2001. N-acetylgalactosamine, N-acetylglucosamine and sialic acid expression in primary breast cancers. Acta Histochem. 103, 37–51.
Buts, L., Dao-Thi, M.H., Loris, R., Wyns, L., Etzler, M., Hamelryck, T., 2001. Weak protein-protein interactions in lectins: the crystal structure of a vegetative lectin from the legume Dolichos biflorus. J. Mol. Biol. 309, 193–201.
Calvete, J.J., Santos, C.F., Mann, K., Grangeiro, T.B., Nimtz, M., Urbanke, C., et al., 1998. Amino acid sequence, glycan structure, and proteolytic processing of the lectin of Vatairea macrocarpa seeds. FEBS Lett. 425, 286–292.
Cao, Y., Merling, A., Karsten, U., Goletz, S., Punzel, M., Kraft, R., et al., 2008. Expression of CD175 (Tn), CD175s (sialosyl-Tn) and CD176 (Thomsen-Friedenreich antigen) on malignant human hematopoietic cells. Int. J. Cancer 123, 89–99.
Cavada, B.S., Santos, C.F., Grangeiro, T.B., Nunes, E.P., Sales, P.V., Ramos, R.L., et al., 1998. Purification and characterization of a lectin from seeds of Vatairea macrocarpa Duke. Phytochemistry 49, 675–680.
Collaborative Computacional Project Number 4. The CCP4 Suite: Programs for Protein Crystallography. Acta Crystallogr 1994;50:760-3.
Chung, S., Gnanapragassam, V.S., Jain, M., Rachagani, S., Ponnusamy, M.P., Batra, S.K., 2015. Pathobiological implications of mucin glycans in cancer: sweet poison and novel targets. Biochim. Biophys. Acta 1856, 211–225.
Dam, T.K., Gerken, T., Cavada, B.S., Nascimento, K.S., Moura, T.R., Brewer, C.F., 2007. Binding studies of alpha-GalNAc-specific lectins to the alpha-GalNAc (Tn-antigen) form of porcine submaxillary mucin and its smaller fragments. J. Biol. Chem. 282, 28256–28263.
Davis, I.W., Leaver-Fay, A., Chen, V.B., Block, J.N., Kapral, G.J., Wang, X., et al., 2007. MolProbity: All-atom contacts and structure validation for proteins and nucleic acids. Nucleic Acids Res. 35, 375–383.
Delatorre, P., Silva-Filho, J.C., Rocha, B.A.M., Santi-Gadelha, T., Nóbrega, R.B., Gadelha, C.A.A., et al., 2013. Interactions between indole-3-acetic acid (IAA) with a lectin from Canavalia maritima seeds reveal a new function for lectins in plant physiology. Biochimie 95, 1697–1703.
Dessen, A., 1995. X-ray crystal structure of the soybean agglutinin cross-linked with a biantennary analog. Biochemistry 34, 4933–4942.
Elola, M.T., Blidner, A.G., Ferragut, F., Bracalente, C., Rabinovich, G.A., 2015. Assembly, organization and regulation of cell-surface receptors by lectin-glycan complexes. Biochem. J. 469, 1–16.
Emsley, P., Cowtan, K., 2004. Coot: Model-building tools for molecular graphics. Acta Crystallogr. Sect. D Biol. Crystallogr. 60, 2126–2132.
Emsley, P., Lohkamp, B., Scott, W.G., Cowtan, K., 2010. Features and development of Coot. Acta Crystallogr. Sect. D Biol. Crystallogr. 66, 486–501.
Franke, D., Svergun, D.I., 2009. DAMMIF, a program for rapid ab-initio shape determination in small-angle scattering. J. Appl. Crystallogr. 42, 342–346.
Fritz, P., Dippon, J., Kierschke, T., Siegle, I., Möhring, A., Moisa, A., et al., 2004. Impact of Mistletoe Lectin Binding in Breast Cancer. Anticancer Res. 24, 1187–1192.
Fu, L., Zhou, C., Yao, S., Yu, J., Liu, B., Bao, J., 2011. Plant lectins: targeting programmed cell death pathways as antitumor agents. Int. J. Biochem. Cell Biol. 43, 1442–1449.
Fuster, M.M., Esko, J.D., 2005. The sweet and sour of cancer: glycans as novel therapeutic targets. Nat. Rev. Cancer 5, 526–542.
Garcia-Pino, A., Buts, L., Wyns, L., Loris, R., 2006. Interplay between metal binding and cist/trans isomerization in legume lectins: structural and thermodynamic study of *P. angolensis* lectin. J. Mol. Biol. 361, 153–167.
Gasteiger, E., Hoogland, C., Gattiker, A., Duvaud, S., Wilkins, M.R., Appel, R.D., et al., 2005. Protein Identification and Analysis Tools on the ExPASy Server. Proteomics Protoc. Handb., 571–607.
Gemeiner, P., Mislovicová, D., Tkác, J., Svitel, J., Pätoprstý, V., Hrabárová, E., et al., 2009. Lectinomics II. A highway to biomedical/clinical diagnostics. Biotechnol. Adv. 27, 1–15.
Glavey, S.V., Huuynh, D., Reagan, M.R., Manier, S., Moschetta, M., Kawano, Y., et al., 2015. The cancer glycome: carbohydrates as mediators of metastasis. Blood Rev. 29, 269–279.
Goldberg, R.N., Kishore, N., Lennen, R.M., 2002. Thermodynamic quantities for the ionization reaction of buffers. J. Phys. Chem. Ref. Data 31, 231–370.
Gonçalves, F.M., Freitas, A.E., Peres, T.V., Rieger, D.K., Ben, J., Maestri, M., et al., 2013. Vatairea macrocarpa lectin (VML) induces depressive-like behavior and expression of neuroinflammatory markers in mice. Neurochem. Res. 38, 2375–2384.
Hamelryck, T.W., Loris, R., Bouckaert, J., Strecker, G., Imberty, A., Fernandez, E., et al., 1999. Carbohydrate Binding, Quaternary Structure and a Novel Hydrophobic Binding Site in Two Legume Lectin Oligomers from Dolichos biflorus. J. Mol. Biol. 286, 1161–1177.
Heinrich, E.L., Welty, L.A.Y., Banner, L.R., Oppenheimer, S.B., 2005. Direct targeting of cancer cells: A multiparameter approach. Acta Histochem. 107, 335–344.
Houtman, J.C.D., Brown, P.H., Bowden, B., Yamaguchi, H., Appella, E., Samelson, L.E., et al., 2007. Studying multisite binary and ternary protein interactions by global analysis of isothermal titration calorimetry data in SEDPHAT: application to adaptor protein complexes in cell signaling. Protein Sci. 16, 30–42.
Itzkowitz, S.H., Yuan, M., Montgomery, C.K., Kjeldsen, T., Takahashi, H.K., Bigbee, W.L., et al., 1989. Expression of Tn, Sialosyl-Tn, and T Antigens in Human Colon Cancer Expression of Tn Sialosyl-Tn, and T Antigens in Human Colon Cancer1. Cancer Res. 49, 197–204.
Jonckheere, N., Skrypek, N., Van, Seuningen, 2014. Mucins and tumor resistance to chemotheurapeutic drugs. Biochim. Biophys. Acta 1846, 142–151.
Kabsch, W., 1988. Automatic indexing of rotation diffraction patterns. J. Appl. Crystallogr. 21, 67–72.
Keller, S., Vargas, C., Zhao, H., Piszczek, G., Brautigam, C.a., Schuck, P., 2012. High-precision isothermal titration calorimetry with automated peak-shape analysis. Anal. Chem. 84, 5066–5073.
Konarev, P.V., Volkov, V.V., Sokolova, A.V., Koch, M.H.J., 2003. Svergun DI.: a Windows PC-based system for small-angle scattering data analysis. J. Appl. Crystallogr. 36, 1277–1282.
Konska, G., Guerry, M., Latour, M.D.E., Guillot, J., 2006. Study of the expression of Tn antigen in different types of human breast cancer cells using VVA-B 4 lectin. Oncol. Rep. 15, 305–310.
Kozin, M.B., Svergun, D.I., 2001. Automated matching of high- and low-resolution structural models. J. Appl. Crystallogr. 34, 33–41.
Kudelka, M.R., Ju, T., Heimburg-Molinaro, J., Cummings, R.D., 2015. Simple sugars to complex disease-mucin-type O-glycans in cancer. Adv. Cancer Res. 126, 53–135.
Kufe, D.W., 2009. Mucins in cancer: function, prognosis and therapy. Nat. Rev. Cancer 9, 874–885.
Kulkarni, K.A., Sinha, S., Katiyar, S., Surolia, A., Vijayan, M., Suguna, K., 2005. Structural basis for the specificity of basic winged bean lectin for the Tn-antigen: a crystallographic, thermodynamic and modelling study. FEBS Lett. 579, 6775–6780.
Lannoo, N., Van Damme, E.J.M., 2010. Nucleocytoplasmic plant lectins. Biochim. Biophys. Acta Gen. Subj. 1800, 190–201.
Laskowski, R.A., MacArthur, M.W., Moss, D.S., Thornton, J.M., 1993. PROCHECK: a program to check the stereochemical quality of protein structures. J. Appl. Crystallogr. 26, 283–291.
Li, W., Yu, J., Xu, H., Bao, J., 2011. Concanavalin A: a potential anti-neoplastic agent targeting apoptosis, autophagy and anti-angiogenesis for cancer therapeutics. Biochem. Biophys. Res. Commun. 414, 282–286.
Liu, B., Bian, H.J., Bao, J.K., 2010. Plant lectins: Potential antineoplastic drugs from bench to clinic. Cancer Lett. 287, 1–12.
Liu, B., Min, G.W., Bao, J.K., 2009. Induction of apoptosis by concanavalin A and its molecular mechanisms in cancer cells. Autophagy 5, 432–433.
Liu, Z., Liu, B., Zhang, Z.T., Zhou, T.T., Bian, H.J., Min, M.W., et al., 2008. A mannose-binding lectin from Sophora flavescens induces apoptosis in HeLa cells. Phytomedicine 15, 867–875.
Liu, Z., Luo, Y., Zhou, T.-T., Zhang, W.-Z., 2013. Could plant lectins become promising anti-tumour drugs for causing autophagic cell death? Cell Prolif. 46, 509–515.
Lo-Man, R., Vichier-Guerre, S., Perraut, R., Dériaud, E., Huteau, V., BenMohamed, L., et al., 2004. A fully synthetic therapeutic vaccine candidate targeting carcinoma-associated Tn carbohydrate antigen induces tumor-specific antibodies in nonhuman primates. Cancer Res. 64, 4987–4994.
Loris, R., Hamelryck, T., Bouckaert, J., Wyns, L., 1998. Legume lectin structure. Biochim. Biophys. Acta 1383, 9–36.
Macedo, M., Oliveira, C., Oliveira, C., 2015. Insecticidal activity of plant lectins and potential application in crop protection. Molecules 20, 2014–2033.
Madariaga, D., Mart, N., Somovilla, V.J., Coelho, H., Valero-gonza, J., Castro-lo, J., et al., 2014. Detection of tumor-associated glycopeptides by lectins: the peptide context modulates carbohydrate recognition. ACS Chem. Biol. 10, 747–756.
Majee, S.B., Biswas, G.R., 2013. Exploring plant lectins in diagnosis, prophylaxis and therapy. J. Med. Plant Res. 7, 3444–3451.
McCoy, A.J., Grosse-Kunstleve, R.W., Adams, P.D., Winn, M.D., Storoni, L.C., Read, R.J., 2007. Phaser crystallographic software. J. Appl. Crystallogr. 40, 658–674.
Morris, G.M., Goodsell, D.S., Halliday, R.S., Huey, R., Hart, W.E., Belew, R.K., et al., 1998. Automated docking using a lamarckian genetic algorithm and and empirical binding free energy function. J. Comp. Chem. 19, 1639–1662.
Newman, J., 2004. Novel buffer systems for macromolecular crystallization. Acta Crystallogr. Sect. D Biol. Crystallogr. 60, 610–612.
Oliveira, C., Teixeira, J.a, Domingues, L., 2014. Recombinant production of plant lectins in microbial systems for biomedical application - the frutalin case study. Plant Sci. 5, 390.
Petoukhov, M.V., Franke, D., Shkumatov, A.V., Tria, G., Kikhney, A.G., Gajda, M., et al., 2012. New developments in the ATSAS program package for small-angle scattering data analysis. J. Appl. Crystallogr. 45, 342–350.
Petoukhov, M.V., Konarev, P.V., Kikhney, A.G., Svergun, D.I., 2007. ATSAS 2.1 - Towards automated and web-supported small-angle scattering data analysis. J. Appl. Crystallogr. 40, 223–228.
Reynoso-Camacho, R., Meji, E.G., Loarca-Piña, G., 2003. Purification and acute toxicity of a lectin extracted from tepary bean (Phaseolus acutifolius). Nutrition Rev. 41, 21–27.
Roth, J., 2011. Lectins for histochemical demonstration of glycans. Histochem. Cell Biol. 136, 117–130.

Rougé, P., Peumans, W.J., Van Damme, E.J.M., Barre, A., Singh, T., Wu, J.H., Wu, A.M., 2011. Structure-function relationships of plant lectins that specifically recognize T and Tn antigens. In: Wu, A.M. (Ed.), The Molecular Immunology of Complex Carbohydrates. , 3rd ed. Springer, pp. 157–170.
Rüdiger, H., Gabius, H.J., 2002. Plant lectins: Occurrence, biochemistry, functions and applications. Glycoconj. J. 18, 589–613.
Shio, Y., Suzuki, H., Kawaguchi, T., Ohsugi, J., Higuchi, M., Fujiu, K., et al., 2007. Carbohydrate status detecting by PNA is changeable through cancer prognosis from primary to metastatic nodal site: A possible prognostic factor in patient with node-positive lung adenocarcinoma. Lung Cancer 57, 187–192.
Singh, R., Bandyopadhyay, D., 2007. A target molecule for cancer therapy. Cancer Biol. Ther. 6, 481–486.
Skubák, P., Murshudov, G.N., Pannu, N.S., 2004. Direct incorporation of experimental phase information in model refinement. Acta Crystallogr. Sect. D Biol. Crystallogr. 60, 2196–2201.
Smart, J.D., 2004. Lectin-mediated drug delivery in the oral cavity. Adv. Drug Deliv. Rev. 56, 481–489.
Sousa, B.L., Silva-Filho, J.C., Kumar, P., Ilário, R., et al., 2015. High-resolution structure of a new Tn antigen-binding lectin from Vatairea macrocarpa and a comparative analysis of Tn-binding legume lectins. Int. J. Biochem. Cell Biol. 59, 103–110.
Svergun, D., Barberato, C., Koch, M.H., 1995. CRYSOL - A program to evaluate X-ray solution scattering of biological macromolecules from atomic coordinates. J. Appl. Crystallogr. 28, 768–773.
Svergun, D.I., 1992. Determination of the regularization parameter in indirect-transform methods using perceptual criteria. J. Appl. Crystallogr. 25, 495–503.
Svergun, D.I., 1999. Restoring low resolution structure of biological macromolecules from solution scattering using simulated annealing. Biophys. J. 76, 2879–2886.
Tempel, W., Tschampel, S., Woods, R.J., 2002. The xenograft antigen bound to Griffonia simplicifolia lectin 1-B4. X-ray crystal structure of the complex and molecular dynamics characterization of the binding site. J. Biol. Chem. 277, 6615–6621.
Trott, O., Olson, A.J., 2009. Software news and update autodock vina: improving the speed and accuracy of docking with a new scoring function, efficient optimization, and multithreading. J. Com. Chem. 31, 455–461.
Vagin, A., Teplyakov, A., 1997. MOLREP: an Automated Program for Molecular Replacement. J. Appl. Crystallogr. 30, 1022–1025.
Vagin, A., Steiner, R., Lebedev, A., Potterton, L., McNicholas, S., Long, F., et al., 2004. REFMAC5 dictionary: Organization of prior chemical knowledge and guidelines for its use. Acta Crystallogr. Sect. D Biol. Crystallogr. 60, 2184–2195.
Vagin, A., Teplyakov, A., 2010. Molecular replacement with MOLREP. Acta Crystallogr. Sect. D Biol. Crystallogr. 66, 22–25.
Van Damme, E.J., Barre, A., Smeets, K., Torrekens, S., Van Leuven, F., Rougé, P., et al., 1995. The bark of Robinia pseudoacacia contains a complex mixture of lectins. Characterization of the proteins and the cDNA clones. Plant Physiol. 107, 833–843.
Van Damme, E.J.M., Barre, A., Rougé, P., Peumans, W.J., 1997. Molecular cloning of the bark and seed lectins from the Japanese pagoda tree (Sophora japonica). Plant Mol. Biol. 33, 523–536.
Van Damme, E.J.M., Peumans, W.J., 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.
Vellieux, F.M.D., Dijkstra, B.W., 1997. Computation of Bhat's OMIT maps with different coefficients. J. Appl. Cryst. 30, 396–399.
Volkov, V.V., Svergun, D.I., 2003. Uniqueness of ab initio shape determination in small-angle scattering. J. Appl. Crystallogr. 36, 860–864.
Wantyghem, J., Platzer, N., Giner, M., Derappe, C., Goussant, Y., 1992. Structural analysis of the carbohydrate chain of glycopeptides isolated from Robinia pseudoacacia seed lectins. Carbohydr. Res. 236, 181–193.
Winn, M.D., Ballard, C.C., Cowtan, K.D., Dodson, E.J., Emsley, P., Evans, P.R., et al., 2011. Overview of the CCP4 suite and current developments. Acta Crystallogr. Sect. D Biol. Crystallogr. 67, 235–242.
Yu, Q.-J., Li, Z.-Y., Yao, S., Ming, M., Wang, S.-Y., Liu, B., et al., 2011. In silico analysis of molecular mechanisms of Galanthus nivalis agglutinin-related lectin-induced cancer cell death from carbohydrate-binding motif evolution hypothesis. Appl. Biochem. Biotechnol. 165, 1037–1046.
Zhao, H., Piszczek, G., Schuck, P., 2014. SEDPHAT – A platform for global ITC analysis and global multi-method analysis of molecular interactions. Methods 77, 137–148.
Zhu, Y., Ng, P.M.L., Wang, L., Ho, B., Ding, J.L., 2006. Diversity in lectins enables immune recognition and differentiation of wide spectrum of pathogens. Int. Immunol. 18, 1671–1680.