Cloning, expression in *Escherichia coli* and characterization of the recombinant Neu5Ac α 2,6Gal β 1,4GlcNAc-specific high-affinity lectin and its mutants from the mushroom *Polyporus squamosus*

Hiroaki TATENO¹, Harry C. WINTER and Irwin J. GOLDSTEIN²

Department of Biological Chemistry, University of Michigan Medical School, Ann Arbor, MI 48109-0606, U.S.A.

Lectin from the mushroom *Polyporus squamosus* (PSL) has a unique carbohydrate-binding specificity for sialylated glycoconjugates containing Neu5Ac α 2,6Gal β 1,4Glc/GlcNAc trisaccharide sequences of asparagine-linked glycoproteins. In the present study, we elucidate the molecular basis for its binding specificity as well as establish a consistent source of this useful lectin using a bacterial expression system. cDNA cloning revealed that PSL contains a ricin B chain-like (QXW)₃ domain at its N-terminus that is composed of three homologous subdomains (α , β and γ). A recombinant lectin was expressed in *Escherichia coli* as a fully active, soluble form. It agglutinated rabbit erythrocytes and showed the highest affinity for Neu5Ac α 2,6Gal β 1,4GlcNAc, but not for the sialyl α 2,3-linked isomer. We also investigated the structure–function relationship of PSL. A monomeric C-terminal

deletion mutant lacking 40 % of the lectin's molecular mass retained sugar-binding activity, indicating that the carbohydrate-binding sites are situated in the N-terminal portion of the lectin, whereas the C-terminal portion probably functions in oligomerization and structural stabilization. Mutant constructs that have single amino acid substitutions in the putative sugar-binding sites, based on sequence alignment with the ricin B-chain, indicate that the β and γ subdomains are most probably sugar-binding sites. The recombinantly expressed lectin will be a valuable reagent for the detection of the Neu5Ac α 2,6Gal β 1,4GlcNAc sequence of asparagine-linked glycans.

Key words: cDNA cloning, lectin, mushroom, *Polyporus squamosus*, recombinant expression system, sialic acid.

INTRODUCTION

Sialic acids are widely distributed as terminal carbohydrates on glycoconjugates of the eukaryotic cell surface [1]. Sialylation is crucial for a variety of cellular functions, such as cell–cell interaction, metastasis and progression of a variety of tumours, virus infection and the biological stability of glycoproteins. Lectins that recognize sialic acid-containing glycans have proven to be powerful tools for studying the biology of these structures [1]. *Limax flavus* agglutinin has been used as a reagent to detect sialic acid of any linkage, whereas SNA (*Sambucus nigra* agglutinin) and *Maackia amurensis* lectin can distinguish between α 2,3- and α 2,6-linked sialic acids; SNA is specific for Neu5Ac α 2,6Gal/GlcNAc, whereas *M. amurensis* lectin is specific for Neu5Ac α 2,3Gal β 1,4Glc/GlcNAc. They can be employed for the detection and preliminary characterization of sialylated glycans on cell surfaces and on complex carbohydrates in solution. As affinity matrices, they can be used for the isolation and resolution of sialic acid-containing glycoconjugates.

Isolectins (PSLs) with similar specificities but different affinities were purified from the mushroom *Polyporus squamosus*. The carbohydrate-binding properties of the high-affinity form were elucidated using four different approaches: precipitation inhibition assay, fluorescence quenching studies, glycolipid binding by lectin staining on high-performance TLC and frontal affinity chromatography coupled with electrospray ionization mass spectrometry [2,3]. From the results obtained by these assays, it was concluded that PSL exhibits the highest specificity and

affinity ($K_d \approx 10 \mu\text{M}$) towards non-reducing terminal Neu5Ac α 2,6Gal β 1,4Glc/GlcNAc groups of N-glycans.

In a further study, PSL was compared with SNA in the histochemical detection of the Neu5Ac α 2,6Gal β 1,4GlcNAc sequence of N-linked oligosaccharides [4]. A more restricted staining pattern was observed by PSL compared with SNA in paraffin sections from different rat organs. Pretreatment of sections with *N*-glycanase F abolished PSL staining, indicating that it reacts solely with asparagine-linked oligosaccharides. In contrast, SNA staining was only partially sensitive to *N*-glycanase F pretreatment, demonstrating that it was partly due to α 2,6-linked Neu5Ac present in serine/threonine-linked oligosaccharides. Both PSL and SNA staining were sensitive to sialidase treatment. Furthermore, PSL did not stain the mucus of sheep submandibular gland, which is extremely rich in serine/threonine-linked Neu5Ac α 2,6GalNAc. Collectively, these results indicated that PSL is a useful tool for the specific histochemical detection of α 2,6-linked Neu5Ac in asparagine-linked oligosaccharides, even though PSL also shows very weak reactivity towards *N*-acetylactosamine end groups.

In view of the importance of PSL as a reagent for identifying the Neu5Ac α 2,6Gal β 1,4Glc/GlcNAc epitope, it is of great interest to determine the molecular basis for its binding specificity as well as to provide a consistent and reliable source of the pure lectin. In the present study, we report the cloning and expression in *Escherichia coli* and characterization of the recombinant high-affinity PSL isolectin. We also constructed mutants by site-directed mutagenesis to shed light on the structure–function relationship of this useful lectin. Recombinantly expressed and

Abbreviations used: MOA, *Marasmius oreades* lectin; poly(A)⁺, polyadenylated; PSL, *Polyporus squamosus* lectin; RACE, rapid amplification of cDNA ends; rPSL, recombinant PSL; SNA, *Sambucus nigra* agglutinin; ST6Gal.I, β -galactoside α 2,6 sialyltransferase.

¹ Present address: Department of Molecular Biology, The Scripps Research Institute, 10550 North Torrey Pines Road, La Jolla, CA 92037, U.S.A.

² To whom correspondence should be addressed (email igoldste@umich.edu).

The nucleotide sequence data reported have been submitted to DDBJ, GenBank® and EBI Nucleotide Sequence Databases under the accession numbers AB120706 (PSL 1a) and AB120707 (PSL 1b).

purified PSL retains the specificity and affinity observed for the native protein. The expression system allows us to prepare milligram quantities of this unique lectin for biomedical studies. The structure–function relationship of PSL is also discussed.

EXPERIMENTAL

Materials

Saccharides, their derivatives, α_2 -macroglobulin, goat anti-rabbit IgG alkaline phosphatase conjugate and Sigma-Fast BCIP (5-bromo-4-chloroindol-3-yl phosphate)/NBT (Nitro Blue Tetrazolium) alkaline phosphatase substrate were purchased from Sigma. Gal β -O-(CH₂)₈CONH-Synsorb (β -D-galactosyl-Synsorb) was obtained from Chembiomed (Edmonton, AB, Canada) and DEAE-Sephacel was from Amersham Biosciences. Rabbit anti-PSL IgG was a gift from Dr F. Yagi (Department of Biochemical Science and Technology, Kagoshima University, Kagoshima, Japan).

Purification of the lectin

Lectin was purified from wild-collected *P. squamosus* as described previously [2]. Briefly, a crude extract of the carpophores was applied on to a β -D-galactosyl-Synsorb column, which had been equilibrated with PBSA (10 mM PBS containing 1 M ammonium sulphate), washed with PBSA and eluted with 0.2 M lactose in PBSA. The affinity-purified lectin was further purified by anion-exchange chromatography on a DEAE-Sephacel column, which had been equilibrated with 50 mM phosphate buffer (pH 7.8). The adsorbed protein was eluted with 50 mM phosphate buffer (pH 7.8) containing 1 M sodium chloride. The DEAE-unbound (pass-through, designated PSL1) and DEAE-bound (designated PSL2) peak fractions were collected, dialysed against PBS and concentrated by ultrafiltration.

Protein estimation

Protein concentration was determined by the method of Lowry et al. [5] using BSA as a standard.

PAGE, SDS/PAGE and Western blotting

Native gel electrophoresis using 10 % (w/v) acrylamide slab gel was performed in an alkaline buffer system (Tris/glycine, pH 8.3) [6]. SDS/PAGE analysis in the presence and absence of 2-mercaptoethanol using a 15 % slab gel was performed in the Tris/glycine buffer system [7]. BenchMark prestained protein molecular-mass standards used in SDS/PAGE were obtained from Invitrogen. For immunostaining, gels were blotted on to PVDF membranes using 10 mM sodium carbonate buffer (pH 10.0) containing 10 % (v/v) methanol. Blots were stained with a polyclonal rabbit anti-PSL antibody and visualized with anti-rabbit IgG alkaline phosphatase and Sigma Fast BCIP/NBT alkaline phosphatase substrate.

Subunit structure

The subunit structure was determined by gel filtration through a G2000-SWXL Progel-TSK column (30 $\times$ 0.78 cm; Supelco, Bellefonte, PA, U.S.A.) using a Beckman System Gold HPLC system in PBS with or without 20 mM lactose at a flow rate of 1 ml/min. Lectins (20 μ l, 0.06–1 mg/ml) were injected, and the effluent was monitored at 280 nm.

Haemagglutination assay and hapten inhibition assay

The haemagglutinating activity of the lectin was determined by a 2-fold serial dilution procedure using formaldehyde-treated [8] 2 % rabbit erythrocytes as described previously [9]. The titre was defined as the reciprocal of the highest dilution exhibiting observable haemagglutination.

For hapten inhibition of haemagglutination, 10 μ l of the serially diluted saccharide solutions was incubated with 10 μ l containing eight agglutinin units of the lectin in microtitre plates, followed by the addition, after 30 min, of a 2 % rabbit erythrocyte suspension (20 μ l). The lowest concentration of saccharide that visibly decreased the extent of agglutination was defined as the minimum inhibitory concentration.

Quantitative precipitation assay

Assays were performed by the microprecipitation procedure described in [10]. Different amounts of glycoproteins or polysaccharides ranging from 0 to 100 μ g were added to 20 μ g of the purified lectin in a total volume of 120 μ l of PBS. After incubation at 37 °C for 1 h, the reaction mixtures were stored at 4 °C for 48 h. The precipitates formed were washed twice with 500 μ l of ice-cold PBS, dissolved in 0.05 M NaOH and assayed for protein by the method of Lowry et al. [5] using BSA as a standard.

Hapten inhibition was performed in the same system with increasing amounts of potentially inhibitory saccharides added to the reaction mixture containing 20 μ g of the purified lectin and 40 μ g of α_2 -macroglobulin in a final volume of 120 μ l of PBS (pH 7.2). The concentration of saccharide required for 50 % inhibition was interpolated from the corresponding inhibition curves.

Peptide sequencing and analysis

Peptide sequences were determined by the Protein Structure Facility at the University of Michigan Core Facilities.

RNA isolation and cDNA cloning

Fresh tissues of the mushroom were ground to a powder with a pestle under liquid nitrogen. Total cellular RNA was isolated with Concert Plant RNA reagent (Invitrogen, Carlsbad, CA, U.S.A.) and, subsequently, poly(A)⁺ (polyadenylated) RNA was isolated with Micro-FastTrack 2.0 kit (Invitrogen). Using this method, approx. 6 μ g of poly(A)⁺ RNA/g of mushroom was isolated.

An adapter-ligated cDNA library was constructed with the Marathon cDNA amplification kit (Clontech, Palo Alto, CA, U.S.A.). Two degenerate forward primers (PSLF1, TTYGAR-GGNCAAYGGNATHATAYCAYATHCC; PSLF2, GCNACNGAY-GGNACNGCNATHATHGGNTA) were designed from the N-terminal amino acid sequence (FEGHGIYHIP) and a cyanogen bromide cleavage fragment (ATDGTAIIGY) respectively of the native PSL1 for RACE (rapid amplification of cDNA ends). 3'-RACE was performed with a combination of primers, i.e. AP1 (adapter primer 1; Invitrogen), PSLF1/PSLF2 and Platinum *Taq* DNA polymerase (Invitrogen), as follows. DNA was denatured at 94 °C for 3 min, followed by three-step cycles (40 cycles) at 92 °C for 0.5 min, 55 °C for 0.5 min and 68 °C for 1 min; denaturation was further extended at 68 °C for 5 min. Each of the amplified fragments was cloned using Zero Blunt TOPO PCR Cloning kit (Invitrogen). Inserted DNA was sequenced with T7 and SP6 primers by the DNA Sequencing Core Facility of University of Michigan, and two different genes (termed *PSL1a* and *PSL1b*), including poly(A)⁺, were obtained. Gene-specific reverse

primers (PSLaR1, CACAGAGTAGCGAGCAACGGGATGGG-TG; PSLbR1, CACCAAACATCAAGCAGCGACACGAGGCC) were designed for *PSL1a* and *PSL1b* respectively, and 5'-RACE was performed with AP1 and PSL1aR1/PSL1bR1 and with *Pfx* DNA polymerase (Invitrogen).

Construction, expression and purification of rPSL (recombinant PSL)

The protein-coding regions of *PSL1a* and *PSL1b* were amplified by PCR using synthetic oligonucleotides incorporating *Nde*I and *Bam*HI restriction enzyme sites for cloning purposes. The approx. 1 kb PCR products were cloned into the expression vector pET-43a (Novagen), yielding pET-PSL1a and pET-PSL1b.

A deletion mutant (PSL1a Δ C) was generated by PCR using GeneTailor Site-Directed Mutagenesis System (Invitrogen), pET-PSL1a as a template and the mutagenesis primers PSL1aD1F (CCCCGCGCTTCGGCAGGATTGAAAGAGTTATC) and PSL1aD1R (AATCCTGCCGAAGCGCGGGGCAGGCAAGCA) according to the manufacturer's instructions. The nucleotides of PSL1aD1F shown underlined (see above) have been altered to introduce a stop codon to generate a deletion mutant of PSL1a (Met¹-Asp¹⁶⁹) and the nucleotide sequence of mutant clones was verified by DNA sequencing. A deletion mutant (PSL1a Δ C-C163S, where C163S stands for Cys¹⁶³ $\rightarrow$ Ser) was generated using pET-PSL1a Δ C as a template and the mutagenesis primers PSL1a Δ C-C163SF (TGCATGCCACTCTGCTTGCCAGCCCCGCGCT) and PSL1a Δ C-C163SR (GGCAAGCAGAGTGGCATGCACGTTTCGCGCCG), wherein the underlined nucleotide of PSL1aD2F (see above) is altered to change Cys¹⁶³ to Ser¹⁶³. All other mutants with single amino acid substitutions were also generated using the GeneTailor Site-Directed Mutagenesis System. All mutant constructs were verified by DNA sequencing.

Nova Blue (DE3) strain of *E. coli* harbouring the expression vector was precultured in 5 ml of Luria-Bertani medium containing 50 μ g/ml ampicillin at 37 °C for 3 h and was added to 1 litre of the same medium. After the absorbance *A* at 600 nm reached 0.4–0.6, 1 ml of 1 M isopropyl β -D-thiogalactoside was added to the medium, and the cells were further cultured at 25 °C overnight. The induced cells were collected by centrifugation, resuspended in a lysis buffer (PBSA supplemented with 1 % Nonidet P40 and 1 mM PMSF) and sonicated at 4 °C. The insoluble fraction was removed by centrifugation at 10 000 *g* for 30 min at 4 °C. Recombinant lectin was purified from the soluble fraction by absorption on to a β -D-galactosyl-Synsorb column in PBSA and elution by PBSA supplemented with 0.2 M lactose.

Thermostability

To examine their thermostability, lectin solutions (1 mg/ml in PBS) were incubated for 30 min at 40, 50, 55, 60 or 65 °C. After the lectin solution was cooled on ice, its haemagglutination activity was assayed as described above.

CD spectroscopy

CD spectra were measured on a Jasco J-715 spectropolarimeter (Japan Spectroscopic Co.) using a protein concentration of 0.1 mg/ml in 10 mM phosphate buffer (pH 7.8) and an optical path length of 1 mm. Buffer baselines were measured under the same conditions and subtracted from the corresponding spectrum. The results were transformed to mean molar ellipticity per residue and smoothed using the standard analysis program. The secondary structure was calculated by the method of Yang et al. [11].

Figure 1 Native PAGE (A) and SDS/PAGE of native and recombinant lectins and deletion mutants in the absence (B) or presence (C) of 2-mercaptoethanol

Lane a, affinity-purified PSL; lane b, PSL1; lane c, PSL2; lane d, affinity-purified rPSL1a; lane e, PSL1a Δ C; lane f, PSL1a Δ C-C163S; lane st, BenchMark prestained protein molecular-mass standards.

Isothermal titration calorimetry

Isothermal titration calorimetry was performed in a VP-ITC (MicroCal, Northampton, MA, U.S.A.) at 25 °C. Lectins in PBS were titrated (4 μ l additions) with ligand in the same buffer. The sialyllactose used was from human milk, 80 % of which is α 2,6-linked. Concentrations were calculated on the basis of the α 2,6-isomer only. Integration of the raw titration data and curve-fitting were performed using Origin ver. 7 software supplied with the instrument.

Processing of sequence data

Multiple sequence alignment was performed by the Clustal W program [12]. A search for homologous sequences was performed using the FASTA program.

RESULTS AND DISCUSSION

Purification of the lectins

According to a previous report [2], two lectin fractions, namely DEAE-unbound (high-affinity) and DEAE-bound (low-affinity) lectins, were separated by affinity chromatography on β -galactosyl-Synsorb, followed by anion-exchange chromatography on DEAE-Sephacel. Only the DEAE-unbound high-affinity lectin was characterized. It was reported that each of the two lectins was electrophoretically pure [2]. However, native PAGE of these fractions on 10 % acrylamide gels showed that the DEAE-unbound fraction is a mixture of proteins with different ionic charges, whereas the DEAE-bound fraction appears to be homogeneous (Figure 1A). In the present study, we designate the DEAE-unbound high-affinity fraction as PSL1 and the DEAE-bound fraction as PSL2.

Amino acid sequencing

The N-terminal amino acid sequence (PFEGHGIYHPSVNTAN-VRI) of PSL1 was determined in a previous report [2]. After cyanogen bromide cleavage of PSL1, two large peptide sequences were detected: DLAGHAATDGTAIIGYRPTG and FGSKSSGA-HAYNWFVERGNF.

Molecular cloning of PSL1

3'-RACE with the adapter primers and the degenerate primers that were designed from the N-terminal amino acid sequence and one of the cyanogen bromide cleavage fragments yielded 1.3 and 1 kb products respectively. Cloning and sequencing of the

Figure 2 Multiple amino-acid sequence alignment of PSL1a and PSL1b with MOA

Breaks for maximum alignment are indicated by dashes. Invariant amino acid residues are boxed and shaded. Conserved amino acid residues are shaded. Solid and dashed underlines denote the N-terminal amino acid sequence and the sequences determined by amino-acid sequence analysis of the isolated peptides generated by cyanogen bromide cleavage of PSL1 respectively. Lower-case letters denote the various amino acids from the N-terminal amino acid sequence of PSL1. The circled asparagine residues denote the potential N-glycosylation sites. The closed arrow indicates the C-terminal amino acid (Asp) of the deletion mutants. The open arrow indicates Cys¹⁶³, which is replaced by Ser¹⁶³ of the PSL1aΔC-C163S.

5'-RACE products generated two different full-length nucleotide sequences including polyadenylation (termed PSL1a and PSL1b). PSL1a and PSL1b contain 71 and 60 bp 5'-untranslated regions, followed by the 858 and 876 bp open reading frames encoding the 286 and 292 amino acid residues, and the 231 and 219 bp 3'-untranslated regions respectively. Neither of them contains an adenylation signal sequence. The calculated molecular masses of PSL1a (31.04 kDa) and PSL1b (31.998 kDa) without an N-terminal methionine residue are in good agreement with the molecular mass of the native lectin (30 kDa) estimated by SDS/PAGE.

The deduced amino acid sequences revealed that PSL1a contains the sequence of the cyanogen bromide cleavage peptides of PSL1, although it appeared to contain a slightly different N-terminal amino acid sequence (Figure 2). PSL1b contains different but similar sequences. No signal sequence could be discerned in the deduced amino acid sequences of the PSL1a and PSL1b, indicating their synthesis on free polysomes. PSL1a and PSL1b have six and three potential N-linked glycosylation sites (Asn-Xaa-Ser/Thr) respectively in their sequences (Figure 2). However, as indicated previously, no carbohydrate was detected in the native PSL1 [2].

As shown in Figure 2, a FASTA search of PSL1a and PSL1b revealed that they show highest sequence identity to the Gal α 1, 3Gal β 1,4GlcNAc-specific lectin from the mushroom *Marasmius oreades* [MOA (*Marasmius oreades* lectin)] [13], despite their unique carbohydrate-binding properties. MOA showed much higher sequence identity to PSL1b compared with PSL1a. These two PSL isolectins also contain three tandemly repeated homologous subdomains, all of which have the conserved (Q/N)/XW motif, the common feature of the ricin B-chain lectin domains. This is documented in Figure 3. A schematic represen-

tation of PSL1, incorporating these homologous subdomains, is shown in Figure 4 along with constructs described below.

Construction and isolation of recombinant lectins

To express recombinant lectins, the entire open reading frames of PSL1a and PSL1b were cloned into the expression vector pET-43a and introduced into *E. coli* strain Nova Blue (DE3) cells. Active rPSL1a was isolated from *E. coli* extracts by a single chromatographic step on a β -D-galactosyl-Synsorb column. The yield of rPSL1a purified by affinity chromatography ranged from 4 to 7 mg from a 1 litre culture. However, rPSL1b accumulated as aggregates in inclusion bodies and a soluble form was not obtained.

Molecular mass and subunit structure

Native PAGE showed that the purified recombinant lectin is homogeneous (Figure 1A). After SDS/PAGE and boiling for 5 min in the presence or absence of 2-mercaptoethanol in SDS sample buffer, rPSL1a gave a single band at 30 kDa by SDS/PAGE (Figures 1B and 1C). rPSL1a was subjected to size-exclusion chromatography in the presence or absence of a haptenic sugar (20 mM lactose), wherein it migrated as a single peak at an apparent molecular mass of 50 kDa, having the same mobility as the native PSL1 (Figure 5). The lectin might interact weakly with the column matrix despite the fact that it is a silica-based non-carbohydrate matrix, since the peak shows tailing. Together with SDS/PAGE analysis, it is clear that the rPSL1a exists as a dimer of approx. 30 kDa subunits, similar to the native lectin. The N-terminal amino acid sequence of rPSL1a was H₂N-SFQGHGI, indicating that its N-terminal methionine residue was removed.

Figure 3 Multiple sequence alignment of tandemly repeated subdomains (α , β and γ) of PSL1 α with the 1 α and 2 γ subdomains in the ricin B-chain

Amino acid residues involved in galactose binding in ricin B as determined by X-ray crystallography (*) and amino acid residues involved in the formation of the hydrophobic core (+) are indicated. The conserved (Q/N)XW motif is boxed. The mutated amino acid residues are shown underlined and in boldface.

Figure 4 Schematic representation of PSL1 and the deletion mutants PSL1 α Δ C and PSL1 α Δ C-C163S

Haemagglutination and haemagglutination inhibition

The minimum concentration for haemagglutination activity of rPSL1 α against formaldehyde-treated rabbit erythrocytes was estimated to be 0.3 μ g/ml, which is indistinguishable from the native PSL1 (0.5 μ g/ml).

The sugar-binding specificity of rPSL1 α was compared with that of native PSL1 and PSL2 by the haemagglutination inhibition assay (Table 1). It is evident that Neu5Ac α 2,6Gal β 1,4GlcNAc is the best inhibitor for rPSL1 α (IC_{50} = 0.0014 mM), being 4500-fold more inhibitory than D-galactose (IC_{50} = 6.3 mM), whereas its α 2,3 isomer was not inhibitory up to 0.36 mM. PSL2 also appeared to be a Neu5Ac α 2,6Gal β 1,4GlcNAc-specific lectin as indicated by the haemagglutination inhibition assay (Table 1).

Quantitative precipitation and precipitation inhibition

rPSL1 α gave a precipitin reaction with human α_2 -macroglobulin similar to the native PSL1, whereas PSL2 did not precipitate with α_2 -macroglobulin (Figure 6). The precipitation inhibition assay showed that rPSL1 α is highly specific for the Neu5Ac α 2,6Gal β 1,4GlcNAc trisaccharide sequence, but not its α 2,3-linked isomer (Figure 7).

Construction, isolation and characterization of C-terminal deletion mutants

To address the question as to whether the ricin B chain-like domain at the N-terminus of PSL1 α functions as a carbohydrate-recognition domain, we designed a C-terminal deletion mutant. Termed PSL1 α Δ C, the C-terminal mutant of rPSL1 α (Met¹-Asp¹⁶⁹, calculated molecular mass 18.09877 kDa) was constructed by site-directed mutagenesis (Figure 4) and expressed in *E. coli*. PSL1 α Δ C bound to a β -galactosyl-Synsorb column and was eluted by lactose. The yield of PSL1 α Δ C was approx. 10 mg from 1 litre of culture. Native PAGE showed that the purified mutant is electrophoretically pure (Figure 1A). After SDS/PAGE in the absence of 2-mercaptoethanol, PSL1 α Δ C gave a single band at approx. 38 kDa (Figure 1B), whereas in its presence, a single band was observed at 18 kDa (Figure 1C). Size-exclusion chromatography of intact PSL1 α Δ C showed a single peak at approx. 38 kDa, indicating it to be a disulphide-linked homodimer of 18 kDa monomers (Figure 5), whereas rPSL1 α is a non-covalently linked dimer of 30 kDa monomers (Figures 1B and 1C). PSL1 α Δ C contains a single cysteine residue at position 163, near its C-terminus. As such, it forms a disulphide-linked homodimer (Figure 4). PSL1 α Δ C agglutinated rabbit erythrocytes, but did not precipitate α_2 -macroglobulin (Figure 6). The

Figure 5 Size-exclusion chromatography on G2000-SWXL Progel-TSK in PBS with 20 mM lactose at a flow rate of 1 ml/min

Lectins (20 μ l, 0.06–1 mg/ml) were injected, and the effluent was monitored at absorbance A_{280} unit full scale of 0.32. Elution positions of standards: BSA (66 kDa), hen ovalbumin (44 kDa), soya-bean trypsin inhibitor (21.5 kDa) and cytochrome *c* (13 kDa) are indicated by the arrows.

minimum concentration required for haemagglutination activity of PSL1a Δ C against formaldehyde-treated rabbit erythrocytes was estimated to be 4 μ g/ml, approx. 8-fold less active than the intact recombinant form (0.5 μ g/ml). The haemagglutination inhibition assay showed that the sugar-binding specificity of PSL1a Δ C was essentially identical with that of the intact rPSL1a (Table 1). The N-terminal amino acid sequence of PSL1a Δ C was H₂N-SFQGHGI, confirming that the deletion that took place was of the C-terminus.

Cys¹⁶³ of PSL1a Δ C, which forms a disulphide-linked dimer, was replaced by a serine residue. The resulting single-chain peptide, PSL1a Δ C-C163S, also bound to β -galactosyl-Synsorb and

Figure 6 Precipitation of native and recombinant lectins and deletion mutants with α_2 -macroglobulin

●, PSL1; ▲, PSL2; ○, rPSL1a; △, PSL1a Δ C; □, PSL1a Δ C-C163S.

Figure 7 Sugar hapten inhibition of rPSL1a- α_2 -macroglobulin precipitation

Inhibitory sugars were incubated with 20 μ g of lectin in 120 μ l total volume of PBS at 4 °C. ●, Neu5Ac α 2,6Gal β 1,4GlcNAc; ▲, Neu5Ac α 2,3Gal β 1,4GlcNAc; ○, *N*-acetylglucosamine; △, lactose; □, galactose.

was eluted by lactose. After SDS/PAGE, either in the absence or presence of 2-mercaptoethanol, PSL1a Δ C-C163S showed a single band at 18 kDa (Figures 1B and 1C). After native PAGE, PSL1a Δ C-C163S showed greater mobility when compared with PSL1a Δ C, suggesting that it differs in molecular structure from PSL1a Δ C, even though they have identical amino acid sequences except at position 163. Gel-exclusion chromatography of PSL1a Δ C-C163S showed a single peak at 18 kDa when the concentration of protein was varied in separate runs from 0.06 to 1 mg/ml. Similar shape and position of the peaks in each case indicates that no concentration-dependent association of monomers occurs (Figure 5). Collectively, these results ascertain that PSL1a Δ C-C163S is a monomer of approx. 18 kDa. Nevertheless,

Table 1 Inhibition of haemagglutination activity of native and recombinant lectins by saccharides

Pr, potency relative to galactose with PSL1; LacNAc, *N*-acetylglucosamine; Me- β -LacNAc, methyl β -*N*-acetylglucosamine.

Saccharide	PSL1		PSL2 Pr	rPSL1a Pr	PSL1a Δ C Pr	PSL1a Δ C-C163S Pr
	IC ₅₀ (mM)*	Pr				
Galactose	6.3	[1]†	0.5	1	0.5	0.5
Me- β -Gal	3.1	2	1	2	0.5	0.5
Gal β 1,4Glc (lactose)	1.6	4	4	4	4	4
Gal β 1,4GlcNAc (LacNAc)	1.6	4	4	4	4	4
Gal β 1,4GlcNAcOMe (Me- β -LacNAc)	0.5	12.6	12.6	12.6	12.6	12.6
Neu5Ac α 2,3Gal β 1,4GlcNAc	> 0.36	< 17.5	< 17.5	< 17.5	< 17.5	< 17.5
Neu5Ac α 2,6Gal β 1,4GlcNAc	0.0014	4500	2250	4500	2250	2250

* Minimum concentration required for 50% inhibition of haemagglutination activity of lectins against formaldehyde-treated rabbit erythrocytes.

† The value of 1 is the referenced relative potency.

Figure 8 Representative calorimetric titration of rPSL1a and its mutants, showing the titration of RPSL1aΔC-C163S, 0.055 mM in PBS with 5.2 mM sialyllactose as described

In all titrations used to obtain the data in Table 2, saturation of the binding sites was obtained, and the initial injection, usually smaller than expected, was deleted for curve-fitting. Multiple iterations were performed for constant χ^2 values, which were achieved only with the one-site model. Error values are for the fitting of each curve and were similar to those shown in all titrations.

PSL1aΔC-C163S agglutinated formaldehyde-treated rabbit erythrocytes at a minimum concentration of approx. 8 $\mu\text{g}/\text{ml}$. The haemagglutination induced by PSL1aΔC-C163S was specifically inhibited by the Neu5Ac α 2,6Gal β 1,4GlcNAc trisaccharide sequence, but not by the α 2,3-linked isomer, the same as the intact recombinant form (Table 1). Similar to the disulphide-linked dimer (PSL1aΔC), PSL1aΔC-C163S also failed to precipitate α_2 -macroglobulin (Figure 6). The N-terminal amino acid sequence of PSL1aΔC-C163S was H₂N-SFQGHGI, indicating that it was properly expressed and isolated.

The properties of the C-terminal deletion mutant indicated unequivocally that the carbohydrate-recognition domain of PSL is confined to the N-terminal portion of the molecule.

Titration calorimetric determination of carbohydrate binding

Thermodynamic parameters of the recombinant and deletion mutants for the binding of sialyllactose were determined. A representative titration is shown in Figure 8. As shown in Table 2, both rPSL1a and monomeric PSL1aΔC-C163S exhibited high affinity towards sialyllactose, having a K_d value of approx. 1.3 μM and a stoichiometry of 1 mol/mol of monomer. rPSL1a was also titrated with methyl β -N-acetyllactosaminide, with which it showed a K_d value of approx. 50 μM and a stoichiometry of 1 mol/mol of monomer. These values are very similar to K_i values of native PSL from the fluorescence quenching inhibition assay described previously [2]. The stoichiometry of rPSL1a and PSL1aΔC-C163S showed that each has a single high-affinity site for α 2,6-linked sialic acid per monomer. The disagreement between the results obtained by haemagglutination and calorimetric assays might be explained by the presence of a second binding site having a much lower affinity ($K_d \geq 1 \text{ mM}$) and/or other specificities, which would escape detection by calorimetric determination, but can interact with glycoconjugates on the cell surface.

CD spectra

The secondary structures of the native and recombinant proteins and deletion mutants were compared by CD spectroscopy (results not shown). The spectra were very similar to each other. The contents of α -helix, β -structure, β -turn and random coil structure of PSL1 were calculated to be 18.4, 56.7, 0 and 24.8% respectively based on the method of Yang et al. [11]. The CD spectrum of rPSL1a was identical with that of PSL1. Deletion of the C-terminal domain of rPSL1a resulted in a decrease in the content of α -helix (9.6–9.8%) with a concomitant increase in the content of random coil (31.4%). The spectra of the two deletion mutants were similar to each other, indicating that they were properly folded during their expression in *E. coli*. The CD spectra also indicate that the N-terminal domain consists mainly of a β -structure, whereas the C-terminal domain has a somewhat higher proportion of α -helix.

Identification of sugar-binding sites

Figure 3 shows a multiple sequence alignment of tandemly repeated subdomains (α , β and γ) of PSL1a with the 1 α and 2 γ subdomains in the ricin B-chain. Although ricin B-chain is composed of two tandemly repeated (QXW)₃ subdomains (1 α , 1 β , 1 γ , 2 α , 2 β and 2 γ), two of the six QXW subdomains have been shown to be involved in lactose binding (1 α and 2 γ) by X-ray crystallography [14,15], although there is other evidence for a third binding site in the 1 β subdomain [16,17]. In the crystallographically identified 1 α and 2 γ domains, the haptenic galactose is located in a cleft in which conserved residues contribute

Table 2 Thermodynamic parameters of sugar binding to lectins

Ligand	Lectin	[Subunit] (mM)	Stoichiometry	K_d (μM)	K_a (10^5 M^{-1})	$-\Delta H$ (kcal/mol)*	$-\Delta G$ (kcal/mol)	$-T\Delta S$ (kcal/mol)
Sial-Lac†	rPSL1a	0.049	0.977	1.28	7.81	8.62	8.03	0.50
Sial-Lac†	PSL1aΔC-C163S	0.055	0.992	1.37	7.29	8.83	8.00	0.83
Sial-Lac†	PSL1a-N47A	0.042	1.28	1.67	6.00	8.18	7.88	0.3
Me- β -LacNAc	rPSL1a	0.076	1.00	49.3	0.203	5.17	5.87	-0.70

* 1 cal = 4.184 J.

† The sialyllactose used was from human milk, 80% of which is α 2,6-linked. Concentrations were calculated on the basis of the α 2,6-isomer only.

Figure 9 Identification of sugar-binding sites

E. coli harbouring the expression vector was cultured in 1 litre of Luria–Bertani medium at 37 °C and protein expression was induced by adding isopropyl β -D-thiogalactoside. The cells were collected, resuspended in 20 ml of lysis buffer and sonicated at 4 °C. The insoluble fraction was removed by centrifugation at 10 000 *g* for 30 min at 4 °C. (A) The activity of soluble extracts containing mutant lectins was analysed by haemagglutination assay; HA, haemagglutination activity. (B) The soluble extracts (0.5 ml) were incubated with β -galactosyl-Synsorb resin (0.5 ml) at room temperature (20–23 °C). After 1 h, the resin was extensively washed with PBSA and the bound lectin was eluted by 100 μ l of PBSA supplemented with 0.2 M lactose. The eluted fractions (5 μ l) were resolved by SDS/PAGE, transferred on to PVDF membranes and subjected to Western blotting with anti-PSL antiserum. (C) The soluble *E. coli* crude extracts were resolved by SDS/PAGE, transferred on to PVDF membranes and subjected to Western blotting with anti-PSL antiserum.

to form a compact hydrophobic core. In the ricin 1 α domain (Figure 3), Asp²², Gln³⁵, Trp³⁷, Asn⁴⁶ and Gln⁴⁷ (Asp²³⁴, Ile²⁴⁶, Tyr²⁴⁸, Asn²⁵⁵ and Gln²⁵⁶ in the 2 γ domain) are involved in the binding of galactose. Trp⁴⁹ (Trp²⁵⁸ in the 2 γ domain) is an essential aromatic residue in the hydrophobic core of the globular domain. Comparison with the structure of ricin suggested that the β subdomain in PSL1a is the most probable carbohydrate-recognition domain, since this unit contains all of the amino acid residues required for both sugar binding and formation of hydrophobic core (Asp⁷², Tyr⁸⁷, Gln⁹⁵ and Trp⁹⁷). On the other hand, the α subdomain may not harbour a carbohydrate-binding site, since position 22 is occupied by an alanine residue in place of an asparagine, which is important for galactose binding in the ricin B-chain.

We investigated the role of the α , β and γ subdomains of rPSL1a using site-directed mutagenesis of amino acids corresponding to residues identified as being involved in ligand binding in the ricin B-chain (Figure 3). We mutated the (Q/N)XW sequence to AXW in each subdomain, termed N47A, Q95A and Q142A. We also mutated Ile³⁷, Tyr⁸⁷ and Trp¹³³, which are putatively essential for sugar binding, to alanine residues; these mutants are termed I37A, Y87A and W133A. All mutants were expressed in *E. coli* cells and the soluble crude extracts containing the mutants were resolved by SDS/PAGE and transferred on to PVDF membranes. Staining of the blots with anti-PSL antiserum revealed the presence of a major band at 30 kDa (Figure 9C) and lesser amounts of several bands between 16 and 21 kDa, which are probably truncated forms of the intact 30 kDa subunit, presumably generated by bacterial proteases.

Mutants N47A, Q95A, Y87A and W133A were expressed in the soluble form to the same extent as the wild-type PSL1a, suggesting that these mutated sites had little impact on the stability of

Figure 10 Thermostability of native and recombinant lectins and deletion mutants

Results are representative of three independent experiments. ●, PSL1; ▲, PSL2; ○, rPSL1a; △, PSL1a Δ C; □, PSL1a Δ C-C163S.

these molecules (Figure 9). Q142A and I37A were also detected in soluble *E. coli* extracts, but the amounts of these mutants expressed were decreased. The soluble extracts were incubated with β -galactosyl-Synsorb and eluted with lactose. As shown in Figure 9(B), the activities of Q95A, Q142A and Y87A were abrogated, whereas N47A, I37A and W133A retained activity. Mutations of Ile³⁷ and Asn⁴⁷ in the α subdomain showed little effect on activity. The purified N47A showed high affinity for α 2,6-linked sialyllactose ($K_d = 1.67 \mu$ M), similar to that of rPSL1a analysed by isothermal titration calorimetry (Table 2).

These results indicate that the β and γ subdomains are most probably the sugar-binding sites in the ricin B-like domain, with the α subdomain being inactive. One of the β and γ subdomains (most probably the β subdomain, since mutation of Tyr⁸⁷, but not Trp¹³³, abolishes activity) is the high-affinity binding site for α 2,6-linked sialic acid structures, whereas both have low affinity for β -galactosyl structures. Because of this low affinity, both subdomains may be necessary to effect a sufficiently strong binding to β -galactosyl-Synsorb resin. This explanation also accounts for the residual haemagglutination activity of the monomeric truncation mutants, which can involve both sialylated and non-sialylated glycans occurring on the erythrocyte membrane.

Thermostability

Thermostabilities of the native and recombinant proteins and deletion mutants were compared (Figure 10). Although the native and recombinant forms retained their activity up to 55 °C, the deletion mutants, PSL1a Δ C and PSL1a Δ C-C163S, lost most of their activity within 30 min at 40 °C, indicating that the C-terminal domain probably functions to stabilize the three-dimensional structure of the lectin.

Concluding remarks

Inasmuch as β -galactoside α 2,6 sialyltransferase (ST6Gal.I) is one of the most frequently up-regulated glycosyltransferase in human cancers, its products, Neu5Ac α 2,6Gal β 1,4GlcNAc-containing glycoproteins, are also overexpressed. In some malignancies such as colon cancer, nearly 100% of the patients show an increased expression of this ST6Gal.I [18]. The elderberry bark lectin (SNA) has long been used as a probe to detect the expression of ST6Gal.I products. However, SNA recognizes Neu5Ac α 2,6Gal β 1,4GlcNAc (the ST6Gal.I product) as well as Neu5Ac α 2,6Gal/GalNAc (sialyl-Tn), which is the product of a different sialyltransferase. Since PSL recognizes only the Neu5Ac α 2,6Gal β 1,4GlcNAc of asparagine-linked glycans,

the recombinant lectin should be a useful probe for the ST6Gal.I product.

We thank Dr D. Aminoff (Department of Biological Chemistry, The University of Michigan, Ann Arbor, MI, U.S.A.) for a critical reading of this paper, Dr R. Triebel for use of the calorimeter, Dr A. Gafni for use of the Jasco J-715 spectropolarimeter, Dr J. Schauerte for helpful advice on CD analysis and Dr F. Yagi for anti-PSL antiserum. This work was supported by a grant (GM29470) from the National Institutes of Health.

REFERENCES

- Varki, A. (1999) Sialic acid. In *Essentials of Glycobiology* (Varki, A., Cummings, R., Esko, J., Freeze, H., Hart, G. and Marth, J., eds.), pp. 195–209, Cold Spring Harbor Laboratory Press, Plainview, NY.
- Mo, H., Winter, H. C. and Goldstein, I. J. (2000) Purification and characterization of a Neu5Ac α 2,6Gal β 1,4Glc/GlcNAc-specific lectin from the fruiting body of the polypore mushroom *Polyporus squamosus*. *J. Biol. Chem.* **275**, 10623–10629.
- Zhang, B., Palcic, M. M., Mo, H., Goldstein, I. J. and Hinds Gaul, O. (2001) Rapid determination of the binding affinity and specificity of the mushroom *Polyporus squamosus* lectin using frontal affinity chromatography coupled mass spectrometry. *Glycobiology* **11**, 141–147.
- Toma, V., Zuber, C., Winter, H. C., Goldstein, I. J. and Roth, J. (2001) Application of a lectin from the mushroom *Polyporus squamosus* for the histochemical detection of the Neu5Ac α 2,6Gal β 1,4Glc/GlcNAc sequence of N-linked oligosaccharides: a comparisons with the *Sambucus nigra* lectin. *Histochem. Cell Biol.* **116**, 183–193.
- Lowry, O. H., Rosebrough, N. J., Farr, A. L. and Randall, R. J. (1951) Protein measurements with the Folin phenol reagent. *J. Biol. Chem.* **193**, 265–275.
- Davis, B. J. (1964) Disc electrophoresis. II. Method and application to human serum proteins. *Ann. N.Y. Acad. Sci.* **121**, 404–427.
- Laemmli, U. K. (1970) Cleavage of structural proteins during the assembly of the head of bacteriophage T4. *Nature (London)* **227**, 680–685.
- Novak, T. P. and Barondes, S. H. (1975) Agglutinin from *Limulus polyphemus*. Purification with formalinized horse erythrocytes as the affinity adsorbent. *Biochim. Biophys. Acta* **393**, 115–123.
- Winter, H. C., Mostafapour, K. and Goldstein, I. J. (2002) The mushroom *Marasmius oreades* lectin is a blood group type B agglutinin that recognizes the Gal α 1,3Gal and Gal α 1,3Gal β 1,4GlcNAc porcine xenoplasentation epitopes with high affinity. *J. Biol. Chem.* **277**, 14996–15001.
- Mo, H., Winter, H. C., Van Damme, E. J. M., Peumans, W. J., Misaki, A. and Goldstein, I. J. (2001) Carbohydrate binding properties of banana (*Musa acuminata*) lectin I. Novel recognition of internal α 1,3-linked glucosyl residues. *Eur. J. Biochem.* **268**, 2609–2615.
- Yang, J. T., Wu, C.-S. and Martinez, H. M. (1986) Calculation of protein conformation from circular dichroism. *Methods Enzymol.* **130**, 208–269.
- Thompson, J. D., Higgins, D. G. and Gibson, T. J. (1994) CLUSTAL W: improving the sensitivity of progressive multiple sequence alignment through sequence weighting, position-specific gap penalties and weight matrix choice. *Nucleic Acids Res.* **22**, 4673–4680.
- Kruger, R. P., Winter, H. C., Simonson-Leff, N., Stuckey, J. A., Goldstein, I. J. and Dixon, J. E. (2002) Cloning, expression, and characterization of the Gal α 1,3Gal high affinity lectin from the mushroom *Marasmius oreades*. *J. Biol. Chem.* **277**, 15002–15005.
- Rutenber, E., Ready, M. and Robertus, J. D. (1987) Structure and evolution of ricin B chain. *Nature (London)* **326**, 624–626.
- Rutenber, E. and Robertus, J. D. (1991) Crystallographic refinement of ricin to 2.5 Å. *Proteins* **10**, 260–269.
- Frankel, A. E., Burbage, C., Fu, T., Tagge, E., Chandler, J. and Willingham, M. C. (1996) Ricin toxin contains at least three galactose-binding sites located in B chain sub-domains 1 α , 1 β , and 2 γ . *Biochemistry* **35**, 14749–14756.
- Steeves, R. M., Denton, M. E., Barnard, F. C., Henry, A. and Lambert, A. H. (1999) Identification of three oligosaccharide binding sites in ricin. *Biochemistry* **38**, 11677–11685.
- Dall'Olio, F., Chiricolo, M., Mariani, E. and Facchini, A. (2000) β -Galactoside α 2,6-sialyltransferase in human colon cancer: contribution of multiple transcripts to regulation of enzyme activity and reactivity with *Sambucus nigra* agglutinin. *Int. J. Cancer* **88**, 58–65.

Received 10 March 2004/5 May 2004; accepted 3 June 2004

Published as BJ Immediate Publication 3 June 2004, DOI 10.1042/BJ20040391