Anti-HIV I/II Activity and Molecular Cloning of a Novel Mannose/Sialic Acid-binding Lectin from Rhizome of Polygonatum cyrtonema Hua

Jie AN1#, Jin-Zhi LIU1#, Chuan-Fang WU1, Jian LI1, Lei DAI1, Els Van DAMME2, Jan BALZARINI3, Erik De CLERCQ3, Fang CHEN1*, and Jin-Ku BAO1*

1 College of Life Science, Sichuan University, Chengdu 610064, China; 2 Department of Molecular Biotechnology, Ghent University, 9000 Gent, Belgium; 3 Rega Institute for Medical Research, Katholieke University Leuven, B-3000 Leuven, Belgium

Abstract The anti-human immunodeficiency virus (HIV) I/II activity of a mannose and sialic acid binding lectin isolated from rhizomes of Polygonatum cyrtonema Hua was elucidated by comparing its HIV infection inhibitory activity in MT-4 and CEM cells with that of other mannose-binding lectins (MBLs). The anti-HIV activity of Polygonatum cyrtonema Hua lectin (PCL) was 10- to 100-fold more potent than other tested MBLs, but without significant cytotoxicity towards MT-4 or CEM cells. To amplify cDNA of PCL by 3′/5′-rapid amplification of cDNA ends (RACE), the 30 amino acids of N-terminal were determined by sequencing and the degenerate oligonucleotide primers were designed. The full-length cDNA of PCL contained 693 bp with an open reading frame encoding a precursor protein of 160 amino acid residues, consisting of a 28-residue signal peptide, a 22-residue C-terminal cleavage peptide and a 110-residue mature polypeptide which contained three tandemly arranged subdomains with an obvious sequence homology to the monocot MBL. However, only one active mannose-binding site (QDNVY) was found in subdomain I of PCL, that of subdomain II and III changed to HNNVY and PDNVY, respectively. There was no intron in PCL, which was in good agreement with other monocot MBLs. Molecular modeling of PCL indicated that its three-dimensional structure resembles that of the snowdrop agglutinin. By docking, an active sialic acid-binding site was found in PCL. The instabilization of translation initiation region (TIR) in mRNA of PCL benefits its high expression in rhizomes.

Key words Polygonatum cyrtonema Hua; mannose-binding lectin; anti-human immunodeficiency virus (HIV) I/II; molecular cloning; 3′/5′-rapid amplification of cDNA ends (RACE); sequence alignment; molecular modeling; docking

Plant lectins are reversible carbohydrate-binding proteins (or glycoproteins) of non-immuno origin that agglutinate cells and/or precipitate glycoconjugates. They possess at least one non-catalytic domain and are usually considered a heterogeneous group of proteins because of the apparent differences in molecular structure, sugar specificity and biological activities between individual lectins [1,2]. Recent advances in the biochemistry, molecular cloning and structural analysis of the lectins revealed the occurrence of seven families of structurally and evolutionarily related proteins which include the legume lectins, the monocot mannose-binding lectins (MBLs), the chitin-binding lectins composed of hevein domains, the type 2 ribosome inactivating proteins (RIPs II) and relevant lectins, jacalin-related lectins (amaranthin lectin family) and the cucurbitaceae phloem lectins [3]. Among them, the monocot MBLs (a superfamily of strictly mannose-specific lectins) have been found exclusively in a subgroup of the monocotyledonous plants. All monocot MBLs consist of subunits with a similar sequence and overall three-
diminutional structure. The first monocot MBL was reported in 1987, when a lectin with an exclusive specificity towards mannan was isolated from snowdrop (Galanthus nivalis) bulbs [4]. Since then, related lectins have been found in various tissues of monocot families: Alliaceae, Amaryllidaceae, Araceae, Bromeliaceae, Iridaceae, Liliaceae and Orchidaceae. Biochemical analysis and molecular cloning clearly indicated that all these lectins belong to a single superfamily of monocot mannos-binding proteins, which were named according to their origin and specificity [5,6]. At present, the monocot MBLs are still being studied intensively because of their interesting biological properties, such as potent inhibition towards retroviruses, which could possibly be applied in crop protection against insects and nematodes [3]. Two representative monocot MBLs, Galanthus nivalis agglutinin (GNA) and Narcissus pseudonarcissus agglutinin (NPA), have been previously reported to be effective inhibitors of the infection of HIV and feline immunodeficiency virus (FIV), respectively [7,8].

Polygonatum cyrtonema Hua is a typical representative of the monocot Liliaceae family and is also an important traditional Chinese herbal medicine. The rhizomes of this plant have been used as a tonic, which benefits the spleen, lung and kidney, and as a traditional medicine for hypertension and diabetes without side effects in Chinese medicine science for about 2000 years. The polysaccharide from Polygonatum cyrtonema has an active effect on the immune system of mice [9] and the mannos/sialic acid-binding protein from the rhizome has multi-biological activities, such as hemagglutination, mitogen and calcium channel block [10,11]. In addition, anti-tumor studies on Polygonatum cyrtonema lectin (PCL) indicated that the lectin exhibited a potent suppressive activity on gastric tumor cell lines SGC and HSC [12].

In this paper, we describe the anti-HIV-I/II activity of lectin and the molecular cloning of the lectin gene from rhizome of Polygonatum cyrtonema Hua. More bioinformatic analysis based on the sequence of PCL is also presented which may enable us to further understand its mechanism at the molecular level.

Materials and Methods

**Polygonatum cyrtonema Hua plant**

The young rhizomes of Polygonatum cyrtonema Hua were collected from Changning County (Yibin, China), then instantly frozen in liquid nitrogen and stored at −70°C before use.

**Purification of PCL and mass spectrometry analysis**

PCL was purified according to a previous study [13]. Matrix-assisted laser desorption/ionization-time of flight (MALDI-TOF) mass spectrometry was performed using a Voyager-RP mass spectrometer (PerSeptive, Foster City, USA), according to the method of Woo et al. [14].

**Antiviral assay for HIV**

PCL was evaluated for its inhibitory effect against HIV-I/II-induced cytopathicity in CEM and MT-4 cells, and other four monocot MBLs, GNA [14], Lycoris radiata agglutinin (LRA), Listera ovata agglutinin (LOA) and Scilla campanulata agglutinin (SCA) [15–17], were tested as controls. The method of the anti-HIV assay has been described previously [18]. Briefly, CEM and MT-4 cells (4.5×10^5 cells/ml) were suspended in fresh culture medium and infected with HIV-I and HIV-II at 100 CCID<sub>50</sub> per milliliter of cell suspension (1 CCID<sub>50</sub> being the dose infective for 50% of cell cultures). Then, 100 µl of the infected cell suspension was transferred to microplate wells, mixed with 100 µl of the appropriate dilutions of the test lectins (i.e. final concentrations of 200, 40, 8, 1.6, 0.32, and 0.062 µg/ml respectively) and further incubated at 37 ºC. After 4–5 d, syncytium formation was examined microscopically in the infected CEM and MT-4 cell cultures. Antiviral activity was expressed as EC<sub>50</sub>, the 50% effective concentration corresponded to the lectin concentration required to inhibit HIV-induced syncytium formation by 50% in the virus-infected CEM or MT-4 cell cultures. The 50% cytotoxicity (CC<sub>50</sub>) corresponded to the lectin concentration required to reduce the viability of CEM or MT-4 cells by 50%.

**Analysis of N-terminal amino acid sequence**

The N-terminal amino acid sequence of PCL was determined by automated Edman degradation. The purified PCL was subjected to 12.5% sodium dodecyl sulfate-polyacrylamide gel electrophoresis (SDS-PAGE) and electrotransferred to a PVDF membrane (Millipore, Bedford, USA) that was stained with Coomassie brilliant blue R-250, and the bands corresponding to the lectin subunits were excised from the membrane. Microsequencing was performed using a PROCISE amino acid sequencer (Applied Biosystems, Foster City, USA) equipped with an HPLC system. The phenylthiohydantoin amino acids were identified in a 140 ºC microgradient using the phenylthiohydantoin amino acid analyzer based on their retention times.

http://www.abbs.info; www.blackwellpublishing.com/abbs
RNA and DNA isolation

The total RNA and DNA were extracted using the RNA and DNA extraction kit (TaKaRa Bio, Shiga, Japan) according to the manufacturer’s instructions.

3’/5’-RACE of Polygonatum cyrtonema lectin gene

cDNA synthesis was performed with the 3’-RACE kit (TaKaRa Bio). First, RNA was reversely transcribed with a cDNA synthesis primer and primer 1 (5’-TCGGATCC- CCAAATCTCTTSTTCACTGESCA-3’) was designed according to the N-terminal amino acid sequence of PCL. The 3’-RACE was performed essentially according to the manufacturer’s instructions. Polymerase chain reaction (PCR) was performed under the following conditions: cDNA was denatured at 94 °C for 5 min, followed by 30 cycles of amplification (94 °C for 30 s, 55 °C for 30 s and 72 °C for 2 min) and 7 min at 72 °C. The PCR product was purified and cloned into pUC18-T vector (TaKaRa Bio) for sequencing.

Based on the sequence of the 3’-RACE product, the specific primers 2 (5’-GACCACA TTGCGGTCTTGCT -3’) and 3 (5’-GCCTGAGTCGTACAACAAA-3’) were designed to amplify the 5’ end of PCL. According to the manufacturer’s instructions in the SMART-RACE cDNA amplification kit (Clontech, San Jose, USA), RNA was reversely transcribed followed by tailing cDNA with SMART II oligonucleotide. The first round of PCR was performed with primer 2 and SMART III. The PCR product was diluted 50-fold for the second round of amplification with primer 3 and SMART III. The PCR product was purified and cloned into pUC18-T vector for sequencing.

Generation of Polygonatum cyrtonema lectin full-length DNA sequence

Based on the nucleotide sequence of the 3’/5’-RACE products, gene-specific primers 4 (5’-GAAGCATGC-GCAGCTAGTAGTAGTAGTAGTC-3’) and 5 (5’-AGGGTG- GACGATAGTAGTAGTAGTGTTG-3’) were designed for the amplification DNA of PCL from genomic DNA. The thermal cycling program was the same as that utilized for 3’/5’-RACE.

Bioinformatic analysis of PCL sequence and structure

DNA sequence and the associated molecular information were analyzed by DNA Tools 6.0 (Carlsberg Research Center, Copenhagen, Denmark), Vector NTI 9.0 (Infor Max, Bethesda, USA), HCAdraw 2 (Doriane, Paris, France), Swiss-Model (Biozentrum, Basel, Switzerland) and AutoDock 3.0 (The Scripps Research Institute, La Jolla, USA).

Results

Antiviral assay for HIV

All four mannose-specific plant lectins (i.e. GNA, LRA, LOA and SCA) were proved highly effective at inhibiting HIV-I- and HIV-II-induced cytopathicity in MT-4 and CEM cells. Their EC50 values ranged from 0.43 µg/ml to 9.00 µg/ml (Table 1) without significant cytotoxicity towards MT-4 and CEM cells, and the CC50 values ranged in 6.50–83.00 µg/ml. PCL exhibited inhibition on HIV-I at EC50 of 0.08 µg/ml in MT-4 cells and 0.05 µg/ml in CEM cells, and on HIV-II at EC50 of 0.08 µg/ml in MT-4 cells and 0.10 µg/ml in CEM cells. These results mean that the anti-HIV potency (EC50) of PCL was 10- to 100-fold higher (more potent) than those of the other four lectins. The cytotoxicity of PCL against MT-4 and CEM cells was much lower than its EC50 towards HIV-I/II, and the CC50 value was 73.00 µg/ml and 74.00 µg/ml, respectively. This was 10-fold less toxic than two mannose-specific plant lectins in the total four investigated ones, as their CC50 values ranged in 6.50–8.60 µg/ml. The CC50 values of PCL were almost 1000-fold higher than its EC50 values. Therefore, PCL was at least 10- to 100-fold more inhibitory against HIV-I and HIV-II in MT-4 and CEM cells.

Molecular mass and N-terminal amino acid sequence of PCL

Upon SDS-PAGE, purified PCL gave a single band.
corresponding to a molecular mass of approximately 12 kDa (data not shown). Similarly, MALDI-TOF mass spectrum showed that the molecular mass of PCL was 11.92 kDa (Fig. 1).

Using automated Edman degradation, 30 amino acids at N-terminal were sequenced as follows: VNSLSSPNSLFTGHSLEVGSYRLMPGC. Therefore, the degenerate primer was synthesized according to the underlined amino acid sequence.

Cloning of PCL full-length cDNA

Based on the primer 1 designed for the amplification of the 3' end of PCL cDNA, a 550-bp fragment was obtained. Two specific primers (primer 2 and 3), which were designed according to the 5'-RACE fragment, were used for the amplification of 5'-PCL cDNA and a 290-bp fragment was obtained by nested PCR. Finally, full-length cDNA and its deduced amino acid sequence were obtained by analysis of cDNA 3' and 5' end sequence. The full-length cDNA of PCL (Fig. 2) was 693 bp and contained a 480-bp open reading frame (ORF) encoding a 160-amino acid protein. Using the gene specific primers 4 and 5, the full-length DNA and deduced amino acid sequences were obtained by PCR on genomic DNA, which was extracted from young rhizomes of Polygonatum cyrtonema Hua. By comparison, it was completely the same among the nucleotide sequence and amino acid sequence of PCL from cDNA and genomic DNA. Comparison of the nucleotide acid sequence and the amino acid sequence of PCL from cDNA and genomic DNA revealed completely same among them. This result indicated that there was no intron in PCL, which was in good agreement with those of other Amaryllidaceae species, such as Galanthus nivalis. According to the rules of predicting lectin signal peptide [19] and the cleavage site of C-terminal [20], a 28-amino acid signal peptide (cleavage site between A28 and V29) and a 22-amino acid C-terminal cleavage peptide (cleavage site between A138 and V139) were identified from the PCL full-length cDNA sequence. So the precursor protein of PCL consists of a 28-residue signal peptide, a 22-residue C-terminal cleavage peptide and a 110-residue mature polypeptide with a molecular weight of 11.92 kDa, and a predicted isoelectric point of 7.0. These results are consistent with the accurate molecular weight of 11.920451 kDa by MALDI-TOF mass spectrometry (Fig. 1).

Sequence comparison between various lectins

The alignment of the amino acid sequences encoding the PCL, Orchidaceae lectin, Liliaceae lectin, Alliaceae lectin and Amaryllidaceae lectin clearly indicated that
PCL belonged to the monocot MBL superfamily. Homolog analysis showed the identity between PCL and GNA, Cymbidium hybrid agglutinin (CA), Allium sativum agglutinin (ASA) and SCA was 61%, 59%, 55.1% and 55%, respectively. The three mannose-binding boxes of all the lectins were strictly conserved and therefore were functional, except the second and third boxes in PCL (Fig. 3). The QDNSVY in PCL was replaced by HNNVY and PDNVY in the subdomain II and III of PCL.

Molecular modeling and docking of PCL

The amino acid sequence of PCL exhibits higher identity and homology with the snowdrop lectin, GNA. Whereas the three-dimensional structure of GNA has been resolved by X-ray diffraction analysis, its coordinates can be used to model the structure of lectins with homologous amino acid sequences. To determine if the overall folding of PCL and the structure of the carbohydrate-binding sites also resembled that of GNA [21], hydrophobic cluster analysis (HCA) [22,23] and molecular modeling were carried out using GNA as a model protein. Molecular modeling of PCL was carried out using the Swiss-Model program (Biozentrum) [24].

In addition to these sequence similarities, the minor insertions or deletions mainly occurred in loops. Structure similarities suggested that both proteins had very similar three-dimensional structures, so the localization of the 12 strands of β-sheet, occurring along the HCA plot of GNA, were readily recognized on the HCA plots (Figs. 4 and 5). The three tandem subdomains were connected by loops to form a 12-strand β-barrel containing three putative mannose-binding sites, which were located in the clefts...
formed by the three bundles of β-sheet. However, the three mannose-binding sites in PCL had undergone some changes from those found in GNA. Gln57 and Asp59 of the binding site of subdomain II of GNA were replaced by His58 and Asn60 in PCL, and Gln26 was replaced by Pro26 in subdomain III, which suggested that these two binding sites were non-reactive because, as was shown by docking experiments, no hydrogen bond could be formed between mutant amino acid and C3-OH of mannose (Fig. 6). Consequently, the PCL monomer probably possessed only one active mannose-binding site instead of three as in GNA monomer, which results from the replacement of several amino acid residues involved in the saccharide-binding sites by His and Asn in subdomain II and Pro in subdomain III of PCL (Figs. 7 and 8). Owing to the presence of these residues, the hydrogen bonds, which were required to anchor sugars into the binding site, could no longer be formed. As a result, the affinity of PCL toward mannose should be slightly reduced compared with that of GNA, but it still kept most of its ability to bind mannose as the first mannose-binding site kept the highest affinity to mannose in most of the monocot MBLs [25]. Because PCL is a mannose- and sialic acid-binding lectin, to investigate the binding ability of sialic acid to PCL, we searched through the three-dimensional structure of PCL by docking and found three putative binding regions in PCL [26]. Given the interaction of PCL, one of the three binding regions at the N-terminal of PCL may not work [21]. Due to the steric hindrance, another binding region near the active mannose-binding domain could not work. Hence, only one region may be active in binding to the sialic acids of gp120 (Fig. 9). The lowest energy for binding and docking is −5.16 Kcal/mol and 7.26 Kcal/mol, respectively.

Forecast of PCL mRNA structure and translation initiation region analysis

The translation initiation region (TIR) has an important...
impact on gene expression, and the secondary structure of this region will influence gene expression directly. Too high a concentration of GC in TIR will highly stabilize the secondary structure, thus the transcription unit can not pass smoothly, thus inhibiting gene expression [27]. For the calculation of RNA $\Delta G$ and the forecast of secondary structure, refer to http://www.genebee.msu.su/services/rna2_reduced.html [28]. After analysis, the GC concentration in PCL TIR was 37%, while it was 50% in PCL cDNA. The $\Delta G$ in TIR was $-3.19$ KJ/mol and no typical stem-loop structure was formed (Fig. 10). The high instability of TIR in PCL mRNA can make the transcription unit pass smoothly and so benefited high expression of protein production, which accorded with the role of PCL as the predominant protein in rhizomes of Polygonatum cyronema Hua.
Discussion

In this study, we tested the HIV inhibitory activity of lectin from the rhizomes of *Polygonatum cyrtonema* Hua, in MT-4 and CEM cells and compared this with GNA, LRA, LOA and SCA, the classic anti-HIV MBLs. All the tested MBLs were found to be active without significant cytotoxicity. However, PCL exhibited a much lower EC50 value and the most effective anti-HIV activity, which was 10- to 100-fold more potent than other tested MBLs. Also, the cytotoxicity was lower than other MBLs from Amaryllidaceae, Liliaceae and Orchidaceae.

In the monocot MBL superfamily, GNA was researched early and its anti-HIV activity mechanism was clearly represented. GNA had high affinity for α-(1-3)-D-mannose oligomers and the crystal structure of GNA has been resolved [25]. It was potent inhibitor of the HIV-induced cytopathicity and directly interfered the virus-cell membrane fusion process by binding to the high-mannose glycans on gp120, the crucial envelope glycoprotein of HIV-I during the infection and blocking the binding of HIV to target cells [29]. A number of studies has clearly shown that the binding of MBL to HIV was dependent on the high-mannose glycans on gp120, and gp120 was extensively glycosylated with N-linked complex and high mannose carbohydrates accounting for about half of the molecular weight [30,31]. Research showed that gp120 of HIV-I contains approximately 50% of the high-mannose-type oligosaccharides. All 24 N-linked glycosylation sites were used in gp120, and 11 of them contained hybrid and/or high-mannose structures, while 13 of the sites contained complex-type oligosaccharides. Furthermore, many of the complex oligosaccharides were sialylated [32–34]. The anti-HIV potency (EC50) of PCL was approximately 10-fold higher, while the cytotoxicity (CC50) was about 10–fold less toxic than GNA in MT-4 and CEM cells. It might be that unlike the other MBLs binding mannose only, PCL binds not only to mannose oligomers but also to sialic acid on gp120 because it is a mannose- and sialic acid-binding lectin [10]. Therefore, PCL could effectively block the infection of HIV toward MT-4 and CEM cells, then inhibit the syncytia formation by binding more saccharides on gp120 and exhibit more effective anti-HIV activity.

Using degenerate primers, which were specifically designed according to the N-terminal sequences of PCL, and the RACE-PCR technique, the full-length cDNA of PCL was obtained. PCL closely resembles the classic monocot MBLs with respect to its molecular structure and amino acid sequence. However, in contrast with most other lectins in this family, it differs from the above-mentioned lectins chiefly in its higher affinity for both mannose and sialic acid, and also differs with respect to its biological activities [10]. Analysis of the primary structure of PCL revealed the presence of three mannose binding sites (QDNVY) as other MBLs, but the subdomain II and III had undergone some changes. The Gln58 and Asp60 were replaced by His58 and Asn60 in subdomain II, and Gln58 was replaced by Pro26 in subdomain III. The carbohydrate binding capability may have changed because these two binding sites were non-reactive. Three sialic acid-binding regions were verified in PCL by docking and only one of them is active. However, we know little about the sialic acid binding domain and the amino acids involved in the binding of sialic acid. These issues should be resolved by further studies, including site-directed mutagenesis and X-ray crystallographic analysis.

References

12. Bao JK, Zeng ZK. Study on molecular stability and biological activity of *Polygonatum cyrtonema* Hua lectin II. Sheng Wu Hua Xue Xue Zhi 1996, 12:
20 Van Damme EJ. Molecular cloning and characterization of multiple isoforms of the snowdrop (Galanthus nivalis L.) lectin. Planta 1991, 186: 35–43
30 Ji X, Gewurz H, Spear GT. Mannose binding lectin (MBL) and HIV. Mol Immunol 2005, 42: 145–152

Edited by
Zheng-Wu QI

©Institute of Biochemistry and Cell Biology, SIBS, CAS