International Journal of Biological Macromolecules 94 (2017) 114–121

Contents lists available at ScienceDirect

International Journal of Biological Macromolecules

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

Comprehensive analysis of α 2–3-linked sialic acid specific *Maackia amurensis* leukagglutinin reveals differentially occupied *N*-glycans and C-terminal processing

Gnanesh Kumar B.S, Avadhesha Surolia *

Molecular Biophysics Unit, Indian Institute of Science, Bangalore, Karnataka 560012, India

ARTICLE INFO

Article history:
Received 11 July 2016
Received in revised form 4 October 2016
Accepted 4 October 2016
Available online 5 October 2016

Keywords:
C-terminal processing
Legume lectin
Macroheterogeneity

ABSTRACT

Seeds of *Maackia amurensis* constitutes two sialic acid specific agglutinins known as leukagglutinin and hemagglutinin. *Maackia amurensis* leukagglutinin (MAL) recognizes α2-3-linked sialic acid present mainly in *N*-glycans and composed of two disulfide linked monomers. It exhibits potential *N*-glycosylation sites (four PNGs) which have been assumed to undergo differential occupancy. In this study we have characterized the site specific macro- and microheterogeneity of monomers in detail by analysing *N*-glycopeptides and peptides through liquid chromatography coupled to ion trap mass spectrometer in MS^3 mode ($LC–MS^n$). We observed the presence of mainly paucimannose *N*-glycans at Asn_{61}, Asn_{113} and Asn_{191} whereas a high mannose type with varying $Man_{5–9}$ occurs at Asn_{179}. Interestingly Asn_{179} and Asn_{191} exhibited differential occupancy which was evident by the presence of non-glycosylated peptides. This has contributed to the difference in molecular mass of monomers upon SDS-PAGE. Further the presence of disulfide linked peptides confirmed the covalent linkage of monomers which also undergoes uniform C-terminal processing.

1. Introduction

N-Glycosylation of proteins is the most conserved and complex posttranslational modification in eukaryotes that begins with the transfer of fourteen sugar precursor ($Glc_3Man_9GlcNAc_2$) to a consensus Asn-X-Thr/Ser sequence. Sometime this motif may undergo differential occupancy by oligosaccharide precursor resulting in macroheterogeneity whereas the extent of processing and modification of the precursor leads to microheterogeneity. Thus *N*-glycosylation leads to structurally more diverse forms of glycoproteins known as glycoforms [1,2]. This process also results in the *N*-glycans diversity across organisms at species level and at cellular level [2]. In both plants and animals, high mannose *N*-glycans remain structurally similar and due to dissimilarity in processing of glycans, plants complex and paucimannose type are characterized by the presence of β (1–2) xylose and α (1–3) fucose in the pentasaccharide core [3]. In contrast, animals complex type glycans are bi or tri antennary decorated with terminal sialic acid, galactose, *N*-acetylgalactosamine and fucose [4].

Agglutinins (leukoagglutinin and hemagglutinin) present in the seeds of *Maackia amurensis* are distinctive among legume lectins due to their specificity towards α 2–3-linked sialic acid present in *N*- or *O*-glycans of several animal glycoproteins [5,6]. Only few lectins are reported to recognize exogenous sialic acid across the plant kingdom such as *Sambucus* sps. agglutinin (α2–6-linked sialic acid) [7,8]. This unusual specificity enabled them as important biochemical tool in sialic acid research [9]. Furthermore, the dimeric *M. amurensis* leukoagglutinin comprises disulfide linked monomers with a molecular mass of 70000 Da [10], a rare structural property among seed lectins. Additionally, the seeds also contain *M. amurensis* hemagglutinin (MAH), a tetramer of non covalently linked 33,000 Da monomers [11].

Nearly all legume lectins found in seeds undergo co- or post translational modifications during synthesis and transport including signal peptide cleavage, *N*-glycosylation, assembly of subunits and finally C-terminal processing [12]. MAL and MAH are also subjected to signal peptide cleavage followed by *N*-glycosylation and has been found to comprise four and three

Abbreviations: aa, amino acid; CID, collision induced dissociation; ESI, electrospray ionization; LC, liquid chromatography; MAL, *Maackia amurensis* leukagglutinin; MAH, *M. amurensis* hemagglutinin; PNGs, potential *N*-glycosylation sites.

* Corresponding author.
E-mail address: surolia@mbu.iisc.ernet.in (A. Surolia).

http://dx.doi.org/10.1016/j.ijbiomac.2016.10.007
0141-8130/

potential *N*-glycosylation sites in the subunits respectively [13,14]. The structural characterization of MAL by Imberty et al. has shown the occupancy of atleast three sites in each monomer [15] and more recently Kim et al. [16] determined the presence of high mannose and paucimannose *N*-glycans after treating with PNGase F and PNGase A respectively. Further it has been shown that the high mannose type is largely associated with the lectin activity and an attempt has been made to determine the site occupancy of high mannose type [16]. However, the studies on site specificity of all *N*-glycans in both the monomers of MAL has not been addressed in detail.

Here we performed the mass spectrometric analysis of MAL for determining site specific macro- and microheterogeneity of *N*-glycans by employing LC–MS^n approach [17,18]. The (glyco)peptides derived from in-solution and in-gel proteolytic digestion were separated on liquid chromatography coupled to electrospray ionization ion trap MS and analysed using MS^3 mode which allowed for unambiguous assignment of glycan sequence to the corresponding site [18]. The analysis confirmed the presence of paucimannose *N*-glycans at Asn_{61} and Asn_{113}, whereas high mannose and another paucimannose type exhibited differential occupancy at Asn_{179} and Asn_{191} respectively. Additionally, the analysis also revealed the C-terminal processing of MAL monomers a feature noted, so far, mostly in hetero tetrameric legume lectins.

2. Materials and methods

2.1. Materials

Maackia amurensis lectin, ammonium bicarbonate, 1,4-dithiothreitol (DTT), iodoacetamide (IAA), sequence grade chymotrypsin and thermolysin, solvents (HPLC grade purity) were purchased from Sigma-Aldrich (MO, USA), Precision Plus Protein™ Dual Color Standards was from Bio-Rad (CA, USA). All other reagents used were of high quality and procured from local firm.

2.2. In-solution digestion

400–500 pmol of MAL (reconstituted in 100 mM ABC) either reduced and alkylated with 10 mM DTT for 1 h at 56 °C followed by brief incubation at 95 °C and alkylated using 55 mM IAA in dark at room temperature or without reduction and alkylation was subjected to proteolysis using 2–3 μl of thermolysin (0.5 μg/μl) at 65 °C and chymotrypsin (0.1 μg/μl) at 37 °C overnight.

2.3. SDS PAGE and in-gel digestion

The nature of subunits of MAL was analysed using 12% SDS-PAGE under reducing and non reduced conditions [19]. The protein bands after coomassie staining and followed by destaining were subjected to in-gel digestion with chymotrypsin according to Shevchenko et al. [20].

2.4. LC–MS^n of (glyco)peptides and analysis

The mass spectrometry analysis of (glyco) peptides resulting from proteolysis was performed as previously described [18]. In brief, the liquid chromatography (LC) was carried out on Agilent HP1100 Series, using a reverse phase column (Agilent Poroshell 120, SB C18, 4.6 mm × 150 mm, 2.7 μm) in a linear gradient with ACN/H_2O/0.1% formic acid for 60 min for the separation of (glyco) peptides. The eluting analytes were analysed on the HCT Ultra PTM Discovery (Bruker Daltonics) ion trap mass spectrometer in positive ion mode. The tandem mass spectrometry (MS^n) was performed in Auto MS^3 mode where precursor ions were fragmented by CID using helium gas as collision gas and for subsequent MS^3 the single most abundant fragment ions from MS^2 were automatically subjected to fragmentation.

All the mass spectra were analysed using Data analysis software, version 4.1. *N*-glycopeptides are characterized by the presence of oxonium ions at *m/z* 366 $(HexNAcHex_1)^{+1}$ and 528 $(HexNAcHex_2)^{+1}$ that arises from the cleavage of the core HexNAc residues. We performed manual search for these ions and from evaluation of B- and Y- ions, the glycan sequence was deduced. The most intense fragment ions mainly $Y_1/Y_{1\alpha}$ $(Peptide + HexNAc)^{+n}$ corresponded to the intact peptide mass linked with proximal HexNAc of glycan. The evaluation of the b- and y- ions in corresponding MS^3 of $Y_1/Y_{1\alpha}$ ions further confirmed the peptide backbone which allowed for proper assignment of *N*-glycans to corresponding site [17,18].

3. Results

3.1. Analysis of N-glycopeptides

The thermolysin digestion of intact MAL covered the *N*-glycopeptides representing all the potential *N*-glycosylation sites (Fig. 1A and B). The MS analysis of *N*-glycopeptides that eluted from LC at 6.5 min and 22.2 min, indicated the presence of high mannose *N*-glycans (Fig. S1A and B). The MS^2 of a doubly charged precursor ions at *m/z* 1254.40 revealed a sequential loss of 162 u from the intact ions that corresponding to hexose (Fig. 2A). The evaluation of almost all Y- and B- ions revealed the presence of 9 hexoses (mannoses) and the mass of Y_1 ion (846.32^{+1}) corresponded with the mass of peptide $LAPNKT_{176–181}$ with proximal HexNAc (GlcNAc)(Fig. 2A). Further the fragmentation of the Y_1 ions confirmed the peptide sequence (Fig. 2B). The precursor ions bearing varying mannose residues from Man_{5-8} were also observed (Fig. S1A and B). Interestingly the presence of a peptide LAPNKT at *m/z* 643.32^{+1} that eluted after its *N*-glycopeptide precursor ions at *m/z* 1254.40^{+2} exhibited the non-glycosylation of Asn_{179} (Fig. 2C) (Table 1).

The MS^2 analysis of rest of the *N*-glycopeptides (Fig. S1C–E) showed the presence of typical paucimannose *N*-glycans at Asn_{61}, Asn_{113} and Asn_{191} where the spectra are characterized by fragment ions arising from the loss of pentose, deoxyhexose and hexose(s) from the intact *N*-glycopeptide precursor ions (Figs. S2–S3 and Fig. 3) [21]. Additionally the variants were also observed either with the lack of pentose or hexose and even deoxyhexose (Fig. S1C–E). The differential occupancy of *N*-glycans was also found at Asn_{191} which possess a paucimannose type. The evidence for its occurrence came from the presence of peptide precursor ions at *m/z* 1022.46^{+1} corresponding to LVYPSNQTT (Fig. 3C) which eluted after its *N*-glycopeptide precursor ions (Table 1). Furthermore, the in-solution digestion with chymotrypsin also confirmed these glycan sequence with respective site of attachment (Table S1). A representative MS/MS of a doubly charged *N*-glycopeptide precursor ions at *m/z* 1311.25^{+2} which was also observed in thermolytic digestion (Fig. S1B) clearly demonstrated a high mannose *N*-glycan at Asn_{179} (Fig. S4).

3.2. N-glycopeptides derived from in-gel digestion

The nonreduced MAL migrated as a single species on SDS-PAGE whereas reduced sample resolved in to two closely related subunits (Fig. S5) indicating the presence of a disulfide linkages between the two monomers in the dimer. The nature of *N*-glycans in these two gel resolved monomers were examined further by subjecting them to chymotryptic digestion. The LC–MS^n analysis of chymotryptic digest of upper band revealed the *N*-glycopeptides corresponding to Asn_{61}, Asn_{179} and Asn_{191} along with a non glycosylated pep-