Available online at www.sciencedirect.com
ScienceDirect
Carbohydrate Research 343 (2008) 48–55
Carbohydrate RESEARCH

N-Glycosylation in *Chrysosporium lucknowense* enzymes

Alexander V. Gusakov,* Alexey I. Antonov and Boris B. Ustinov

Division of Chemical Enzymology, Department of Chemistry, M. V. Lomonosov Moscow State University, Moscow 119899, Russia

Received 24 September 2007; received in revised form 16 October 2007; accepted 18 October 2007
Available online 26 October 2007

Abstract—Twenty-eight enzymes, encoded by different genes and secreted by different mutant strains of *Chrysosporium lucknowense*, were subjected to MALDI-TOF MS peptide fingerprinting followed by analysis of the MS data using the GlycoMod tool from the ExPASy proteomic site. Various N-linked glycan structures were discriminated in the *C. lucknowense* proteins as a result of the analysis. N-Glycosylated peptides with modifications matching the oligosaccharide compositions contained in the GlycoSuiteDB were found in 12 proteins. The most frequently encountered N-linked glycan, found in 9 peptides from 7 proteins, was $(Man)_3(GlcNAc)_2$, that is, the core pentasaccharide structure forming mammalian-type high-mannose and hybrid/complex glycans in glycoproteins from different organisms. Nine out of 12 enzymes represented variably N-glycosylated proteins carrying common $(Hex)_{0–4}(HexNAc)_{0–6} + (Man)_3(GlcNAc)_2$ structures, most of them being hybrid/complex glycans. Various glycan structures were likely formed as a result of the enzymatic trimming of a 'parent' oligosaccharide with different glycosidases. The N-glycosylation patterns found in *C. lucknowense* proteins differ from those reported for the extensively studied enzymes from *Aspergilli* and *Trichoderma* species, where high-mannose glycans of variable structure have been detected.

Keywords: *Aspergillus* sp.; *Chrysosporium lucknowense*; Glycoproteins; MALDI-TOF mass spectrometry; N-Glycosylation; *Trichoderma reesei*

1. Introduction

Fungal glycoside hydrolases[1,2] and other enzymes belonging to different protein families are often glycosylated, carrying both O-linked and N-linked glycans. The glycosylation is the most frequently encountered post-translational modification in those proteins. The molecules of many glycoside hydrolases have a modular structure, that is, they consist of a catalytic module, flexible peptide linker, and carbohydrate-binding module.[3,4] Linker peptides, which are rich in Ser and Thr residues, are typically O-glycosylated.[4–6] The N-glycosylation seems to be restricted to the catalytic modules, and it is usually absent in other parts of enzyme molecules. The fungal carbohydrate-binding modules are not glycosylated as a rule.

Although a large quantity of information on the glycosylation of yeast proteins has been published, the number of species belonging to filamentous fungi and studied from this point of view is not high. The N-glycosylation has been extensively studied in different enzymes from *Aspergilli*[7–17] as well as in *Trichoderma reesei* cellulases and acetylxylan esterase.[6,18–28] It has been shown that, in most cases, the N-linked oligosaccharides represent mammalian-type high-mannose glycans ($Man_xGlcNAc_2$), which may be phosphorylated (in *T. reesei*). After the secretion of mature proteins, the high-mannose-type glycans may be subjected to enzymatic trimming by an array of hydrolytic enzymes, including α-mannosidases, endoglycosidases F or H, phosphatase, etc. Thus, depending on the fungal strain and fermentation conditions, the observed glycosylation patterns display different variations.[6,22,24,25] Up to 11 mannose residues have been detected in the N-linked oligosaccharide from cellobiohydrolase I of *T. reesei*.[6] A linear $Hex_{7-26}GlcNAc_2$ series of glycans containing up to 24 mannose residues and up to three β-Gal*f* residues has been found in α-galactosidase A from *Aspergillus niger*.[14]

* Corresponding author. Tel.: +7 495 939 5966; fax: +7 495 939 0997; e-mail: avgusakov@enzyme.chem.msu.ru

0008-6215/$ - see front matter
doi:10.1016/j.carres.2007.10.014

Chrysosporium lucknowense is a filamentous fungus belonging to ascomycetes. It produces a wide range of enzymes that catalyze the biodegradation of cellulose, hemicelluloses, and other polysaccharides. During the last few years, we isolated and characterized about 30 enzymes secreted by different strains of *C. lucknowense*; data on purification and properties of some of the enzymes have been published previously.[29–32] The characterization also included MALDI-TOF mass spectrometry (MS) peptide fingerprinting for the isolated proteins. The full fungal genome of the *C. lucknowense* was sequenced in 2005, and then the genome annotation was carried out (see for instance http://www.dyadic.com/wt/dyad/pr_1143820822). About 200 carbohydrate-active enzymes have already been identified in the *C. lucknowense* genome at this time, so this fungus is very interesting as a source of glycosidases and other enzymes acting on carbohydrates. The amino acid sequences of *C. lucknowense* glycoside hydrolases demonstrate the highest similarity (up to 70–85%) to known enzymes from *Chaetomium* spp., *Humicola* spp., *Neurospora crassa*, while displaying less similarity (up to 60–70%) to the fungal enzymes of genera *Trichoderma (Hypocrea)*, *Aspergillus*, *Phanerochaete*.

The availability of the protein amino acid sequences from the proprietary annotated *C. lucknowense* genome combined with the MS data for secreted proteins allowed one to discriminate possible modifications in peptides derived from *C. lucknowense* proteins. This paper describes N-glycosylation patterns observed in different *C. lucknowense* enzymes, which were found by the analysis of MS data using the GlycoMod tool from the ExPASy proteomic site.[33]

2. Results

Twenty-eight enzymes in total, encoded by different genes and secreted by different strains of *C. lucknowense*, were subjected to MALDI-TOF MS peptide fingerprinting, followed by analysis of the MS data with the MASCOT program for identification of enzymes in the proprietary *C. lucknowense* translated protein database as well as proteomic tools (MotifScan, PeptideMass, FindPept, FindMod, GlycoMod) from the ExPASy site created and maintained by the Swiss Institute of Bioinformatics (http://cn.expasy.org/tools/). The proteins analyzed belonged to different classes of enzymes (hydrolases and oxidoreductases) and included nine cellulases (*endo*-(1→4)-β-glucanases, EC 3.2.1.4, and *exo*-cellobiohydrolases, EC 3.2.1.91), six *endo*-(1→4)-β-xylanases (EC 3.2.1.8), two cellobiose dehydrogenases (EC 1.1.99.18), β-glucosidase (EC 3.2.1.21), β-xylosidase (xylan (1→4)-β-xylosidase, EC 3.2.1.37), *exo*-β-D-glucosaminidase (EC 3.2.1.-), glucoamylase (EC 3.2.1.3), chitinase (EC 3.2.1.14), arabinogalactan *endo*-(1→4)-β-galactosidase (EC 3.2.1.89), laminarinase (glucan (1→3)-β-glucosidase, EC 3.2.1.58), xyloglucanase (xyloglucan-specific *exo*-β-(1→4)-glucanase, EC 3.2.1.155), non-classified β-(1→3)-glucanase belonging to the GH17 family (EC 3.2.1.-), acetylxylan esterase (EC 3.1.1.72), and alkaline serine protease (EC 3.4.21.-).

Although the MotifScan tool found potential N-glycosylation sites (the consensus sequence Asn-Xaa-Ser/Thr, where Xaa cannot be Pro) in most of the proteins analyzed (in 23 enzymes out of 28 in total), the analysis of the MS data with the GlycoMod tool revealed the possible N-glycosylated peptides with modifications matching the oligosaccharide compositions contained in the GlycoSuiteDB[34,35] only in 12 proteins. They are listed in Table 1.

To demonstrate how the analysis was carried out, we will examine the *C. lucknowense* xylanase (Xyl IV) that represents a good example of a variably glycosylated protein. The amino acid sequence of the Xyl IV is shown in Figure 1. The enzyme consists of 375 amino acid residues (including a signal peptide). The MALDI-TOF mass spectrum of peptides obtained after a tryptic digestion of the protein band from the SDS-PAGE gel is shown in Figure 2. Using the PeptideMass and FindPept tools, 22 specific tryptic peptides matching the sequence of Xyl IV were found. They are shown in bold in Figure 1. The GlycoMod tool found additional nine peaks in the mass spectrum of the Xyl IV digest that corresponded to two peptides (shaded in Fig. 1) carrying variable modifications of Asn-98 and Asn-285 with oligosaccharides (Table 2). Any N-linked glycans bound to the third potential N-glycosylation site (Asn-356) were not discriminated.

The peak with m/z 2645.2 in Figure 2 corresponds to the first N-glycosylated peptide in the Xyl IV sequence (Fig. 1), where the $(Hex)_3(GlcNAc)_2$ oligosaccharide is bound to the Asn-98 residue. This oligosaccharide seems to represent the well-known conserved $(Man)_3(GlcNAc)_2$ core structure that forms mammalian-type high-mannose and hybrid/complex glycans in glycoproteins from different organisms.[37] Two other peaks (with m/z 2807.3 and 2969.3), differing cumulatively from the first mentioned peak by 162.1 and 162.0 Da, that is, by the mass of an anhydrohexose residue, correspond to the same peptide modified with glycans containing the same core pentasaccharide that is longer by one and two Hex residues, respectively (Table 2, Nos. 2 and 3). Two additional peaks (with m/z 3010.4 and 3172.4), differing cumulatively from the peak 2645.2 by 365.1 Da (HexNAc + Hex) and 162.0 Da (Hex), correspond to the same first peptide modified with $(Hex)_1(HexNAc)_1 + (Man)_3(GlcNAc)_2$ and $(Hex)_2(HexNAc)_1 + (Man)_3(GlcNAc)_2$ oligosaccharides, respectively. These oligosaccharides represent hybrid/complex glycan structures according to the nomenclature accepted in the GlycoSuiteDB.