Journal of Chromatography A, 1217 (2010) 225–234

Contents lists available at ScienceDirect

Journal of Chromatography A

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

Separation by hydrophobic interaction chromatography and structural determination by mass spectrometry of mannosylated glycoforms of a recombinant transferrin-exendin-4 fusion protein from yeast

Melissa D. Zolodz[a,1], John T. Herberg[b,1], Halyna E. Narepekha[a], Emily Raleigh[b], Matthew R. Farber[a], Robert L. Dufield[a], Denis M. Boyle[b,*]

[a] Analytical Research and Development, Pfizer Global Biologics, Chesterfield, MO 63017, USA
[b] Bioprocess Research and Development, Pfizer Global Biologics, Chesterfield, MO 63017, USA

ARTICLE INFO

Article history:
Available online 30 October 2009

Keywords:
Hydrophobic interaction chromatography (HIC)
Mannosylation
Transferrin
Yeast
Mass spectrometry
ETD
NPLC/MS

ABSTRACT

Obtaining sufficient amounts of pure glycoprotein variants to characterize their structures is an important goal in both functional biology and the biotechnology industry. We have developed preparative HIC conditions that resolve glycoform variants on the basis of overall carbohydrate content for a recombinant transferrin-exendin-4 fusion protein. The fusion protein was expressed from the yeast *Saccharomyces cerevisiae* from high density fermentation and is post-translationally modified with mannose sugars through O-glycosidic linkages. Overall hydrophobic behavior appeared to be dominated by the N-terminal 39 amino acids from the exendin-4 and linker peptide sequences as compared to the less hydrophobic behavior of human transferrin alone. In addition, using LC techniques that measure total glycans released from the pure protein combined with new high resolution technologies using mass spectrometry, we have determined the locations and chain lengths of mannose residues on specific peptides derived from tryptic maps of the transferrin-exendin-4 protein. Though the protein is large (80,488 kDa) and contains 78 possible serine and threonine residues as potential sites for sugar addition, mannosylation was observed on only two tryptic peptides located within the first 55 amino acids of the N-terminus. These glycopeptides were highly heterogeneous and contained between 1 and 10 mannose residues scattered among the various serine and threonine sites which were identified by electron transfer dissociation mass spectrometry. Glycan sequences from 1 to 6 linear mannose residues were detected, but mannose chain lengths of 3 or 4 were more common and formed 80% of the total oligosaccharides. This work introduces new technological capabilities for the purification and characterization of glycosylated variants of therapeutic recombinant proteins.

1. Introduction

Hydrophobic interaction chromatography (HIC) is widely used in the purification of proteins. HIC exploits the binding between hydrophobic patches on the surface of proteins and a nonpolar ligand immobilized to a solid support [1,2]. The main factors affecting protein retention to a hydrophobic support are the type and concentration of salt used [3,4], and the hydrophobicity and density of the immobilized ligand [5,6]. Temperature, pH and other conditions employed during binding such as chromatographic flow rate can also exert major effects. In classical HIC, protein binding is typically achieved by high salt concentrations that specifically promote these interactions, while desorption is controlled either by isocratic elution or by gradual removal of these salts from chromatography columns [7]. Other components, often nonpolar but water soluble additives can also be used to control elution selectivity [8,9]. The complex selectivity behavior of proteins on HIC media has been estimated from intrinsic protein properties relating to their hydrophobic surface properties [4,10], and has been proposed as resulting from specific structural changes, i.e., spreading or unfolding of the protein during adsorption [11], the distribution of salt ions and water molecules during binding [12], or by different preferred binding orientations of the protein or a combination of different mixed mode interactions [13,14].

The usefulness of HIC for the purification of proteins comes from an ability to separate closely related variants by varying chromatographic selectivity through mobile phase changes while still retaining high adsorptive capacities for preparative purposes. Fur-

* Corresponding author at: Pfizer Inc., BB3D, 700 Chesterfield Village Parkway West, Chesterfield, MO 63017, USA. Tel.: +1 636 247 6394; fax: +1 636 247 6511.
E-mail address: denis.m.boyle@pfizer.com (D.M. Boyle).
[1] These authors contributed equally to this work.

0021-9673/$ – see front matter © 2009 Elsevier B.V. All rights reserved.
doi:10.1016/j.chroma.2009.10.057

thermore, requirements in the biotechnology industry demand uniformly high purity preparations with a high degree of structural knowledge of the components, which in turn require highly resolving and scaleable technologies along with the analytical tools to produce them. The large scale synthesis of heterologous proteins from recombinant systems nearly always results in small amounts of structural heterogeneities of the desired protein as part of the complex mixture [15,16]. These arise during translation or result from chemical modifications incurred during purification (post-translational modifications, sequence variations generated from proteolysis or transcriptional/translational errors, and degradation products) [17], and are important because any of these can potentially affect the target activity profile of a proposed therapeutic protein. Variants at the primary sequence level such as truncations or mis-incorporations [18,19], oxidations/deamidation [20] and glycosylation [21,22] variants, or multimeric aggregated impurities [23] frequently have altered charge properties to allow separation of species by traditional ion exchange methods, but this is not always the case. The use of HIC in preparative mode to separate uncharged protein variants from the primary target species, such as glycoprotein isoforms in semi-purified preparations, is especially appropriate.

Control of protein glycosylation patterns in particular is an important consideration for the choice of expression system for biotechnology products. The glycosylation pattern of circulating proteins or peptides significantly affects their fate *in vivo*, such as their resistance to proteolysis [24] (including transferrin [25]), effects upon plasma half-life [26], immunogenicity [27–29] or bioactivity [30–32]. Both the total amount and specific type of glycan structures attached to the parent molecule must be evaluated during manufacturing to maintain a consistent pharmacokinetic profile and avoid undesirable side effects. Human glycoproteins have both N- and O-linked glycan structures and most marketed recombinant proteins are currently produced from mammalian cell platforms [33]. Heterologous proteins produced from yeast also contain both N- and O-linked glycoproteins, and typically O-linked glycoproteins are produced as long linear chains of mannose [2,34]. For example, transferrin has been produced from both *Pichia pastoris* [35–37] and *Saccharomyces cerevisiae* [38]; this is consistent with the use of these systems to produce several protein biotherapeutic candidates [33]. Other key benefits include the development of yeast strains that can be efficiently and stably transformed with foreign genes, secrete large amounts of protein, and show vigorous growth at high density with straightforward scale-up on inexpensive substrates.

In this paper, we used HIC to separate glycosylated isoforms (mannosylated glycoproteins) of a purified recombinant transferrin-exendin-4 fusion protein (Tf-exendin) produced in the yeast *S. cerevisiae*. The extent of glycosylation is not predictable in yeast proteins and they are usually produced as complex mixtures [34]. Addition of the first mannose residue on serine and threonine residues occurs in the ER, while linear elongation of mannose residues through mannosyltransferase enzymes occurs in the Golgi [21,39]. Also presented are new mass spectrometric glycan analysis methods to characterize the mannosylated glycoprotein variants. Due to its complex degree of modification, several new technologies were optimized to characterize the extent of this heterogeneous post-translational modification. Recombinant proteins grown in yeast provide a unique set of analytical challenges but are viable candidates for characterization utilizing mass spectrometry such as linear ion trap mass spectrometry. We first coupled traditional glycan normal phase HPLC (NPLC) to ESI mass spectrometry to determine glycan length. Since O-linked glycosylations sites are not constricted to any consensus site sequence, determining the location of the oligosaccharides and elucidating site occupancy can be a cumbersome and difficult task [34]. A key technology in the present work is the dual use of Collision Induced Dissociation (CID) and Electron Transfer Dissociation (ETD) [40,41]. With CID, the weakest bonds break first (such as labile glycosylation bonds) which prevents determining the site of glycosylation [42]. However, with ETD, an electron from an electron donor anion (such as fluoranthene) produces charged peptide ions and fragmentation occurs along the peptide backbone almost randomly [43]. ETD is a particularly effective tool to locate the sites of glycosylation. The oligosaccharide remains attached to the serine/threonine residues providing ions with both the peptide sequence, the site of the glycan and the length of the glycan. We can use the property of neutral loss of the sugar via CID to locate the glycopeptides in the tryptic peptide map to trigger this ETD process. This method is termed "neutral loss-dependent ETD" (NL-dependent ETD) and is a form of tandem mass spectrometry used in LC/MS peptide mapping experiments to determine peptide sequence, glycosylation site and glycan size on each amino acid [44]. Tandem LC/MS via neutral loss of sugar moiety with CID followed by ETD on the precursor ion was optimized to locate glycosylation sites. The linear ion trap (LIT) mass spectrometer with its dual fragmentation methodologies is a particularly valuable tool in the characterization of recombinant yeast proteins. These data were used to provide structural assignments for both the length and location of mannose residues to relate these structural characteristics to their observed behavior in HIC.

2. Materials and methods

2.1. Preparation of purified transferrin fusion protein

The 80,488 kDa fusion protein consists of exendin-4, a 39 amino acid GLP-1 peptide agonist [45] at the N-terminus, then a 12 amino acid linker sequence (PEAPTDPEAPTD) from human transferrin followed by the complete sequence of 679 residues for transferrin. Two N-linked glycosylation sites (S466A, T664A) in the molecule have been eliminated, but there are 78 serine and threonine residues as possible O-linked glycosylation sites. Recombinant human transferrin (hTf) purified from *S. cerevisiae* without N-linked sites was used as a control in the studies. Briefly, protein was purified as a secreted protein from broth obtained from 11 L of high density, fed-batch fermentation broth of *S. cerevisiae*. Soluble Tf-exendin protein (1.8 g protein/L cell-free supernatant solution) was collected following removal of cells by a continuous operation using centrifugation, re-centrifugation of resuspended solids, microfiltration through a 1000 K molecular weight cut off (MWCO) membrane followed by retention and concentration by ultrafiltration using a 30 K MWCO membrane. Purification was accomplished as follows: SO_3 fractogel chromatography, α-mannosidase (Jack bean) digestion of the product pool for 24 h. at pH 5.0 and enzyme: protein ratio of 25.5:1 (w/w) in 1 mM $ZnSO_4$, Butyl Sepharose-HP chromatography, ultrafiltration/diafiltration (50 K MWCO), Poros 50HQ chromatography followed by a final ultrafiltration/diafiltration. Purified protein was stored frozen at 20 mg/mL in buffer consisting of 20 mM L-histidine, 20 mM sodium phosphate, 90 g/L trehalose, 0.1 g/L L-methionine, pH 7.3.

2.2. Hydrophobic interaction chromatography (HIC) conditions

All chemicals used for the preparation of buffers were USP grade and were obtained from VWR Scientific (Westchester, PA, USA). Buffers were prepared and pH adjusted at room temperature (18–22 °C). Butyl Sepharose-HP resin was obtained from GE Healthcare (Piscataway, NJ, USA) and Toyopearl Butyl 650S was obtained from Tosoh Bioscience (Montgomeryville, PA, USA). Chromatography experiments were performed at room temperature by using a GE Healthcare AKTA Explorer 100 automated chromatography