Available online at www.sciencedirect.com
ScienceDirect
Journal of Structural Biology 158 (2007) 255–266
Journal of Structural Biology
www.elsevier.com/locate/yjsbi

Carbohydrate binding properties and carbohydrate induced conformational switch in sheep secretory glycoprotein (SPS-40): Crystal structures of four complexes of SPS-40 with chitin-like oligosaccharides

Devendra B. Srivastava, Abdul S. Ethayathulla, Janesh Kumar, Rishi K. Somvanshi, Sujata Sharma, Sharmistha Dey, Tej P. Singh *

Department of Biophysics, All India Institute of Medical Sciences, New Delhi 110029, India

Received 23 September 2006; received in revised form 1 November 2006; accepted 7 November 2006
Available online 17 November 2006

Abstract

Crystal structures of four complexes of sheep secretory glycoprotein (SPS-40) with *N*-acetylglucosamine oligosaccharides ($GlcNAc_n$, $_{n=3-6}$) have been determined at moderate resolutions. The binding studies of SPS-40 have been carried out using fluorescence spectroscopy and Surface Plasmon Resonance (SPR). Structure determinations of four complexes have shown a novel binding pattern of $GlcNAc_n$ molecules to SPS-40. The results indicate that the most preferred recognition region in the carbohydrate binding groove in SPS-40 is at subsites −4 to −2 among which subsite −2 provides the maximum interactions with carbohydrate residues. These structures have also shown that the interactions of $GlcNAc_3$ and $GlcNAc_4$ do not perturb the protein structure and those of $GlcNAc_5$ induce partial conformational changes while in the case of $GlcNAc_6$ the partially closed binding groove opened up completely. As in other SPX-40 structures, SPS-40 structure contains three overlapping flexible surface segments, His188–His197, Phe202–Arg212 and Phe244–Pro260 with several charged residues protruding outwardly. It creates a cluster of positive charges with a flexible base thus indicating a good scope of promoting the intermolecular interactions. This protein is glycosylated at Asn39 and may recognize other receptors having sugar binding sites. It appears that SPS-40 may involve both carbohydrate and protein bindings. The systematic carbohydrate-binding studies and the detailed structural results of four protein–carbohydrate complexes provide an excellent insight into the mechanism of carbohydrate binding. These are the first studies of this kind on secretory glycoproteins and their interactions with carbohydrates.

Keywords: Secretory glycoprotein; Sugar-binding; Protein binding; Dry secretion; Chito-oligosaccharides

1. Introduction

Tissue homeostasis is a result of a coordinated regulation of proliferation and elimination of cells. Developmentally regulated removal of cells is mostly controlled by programmed cell death (PCD). Pregnancy induces a massive development of mammary alveolar structures which at birth embody the differentiated secretory epithelium required for milk production. After cessation of lactation a collapse of the lobulo-alveolar structures occurs which is paralleled by a reductive remodeling of the gland. This process is termed as involution and it is characterized by a proteolytic degradation of the extracellular matrix and a loss of secretory epithelial cells mainly by programmed cell death (Strange et al., 1992). A 40 kDa glycoprotein is expressed in high concentrations during this period of involution which is similar to a prominent protein reported earlier in the whey secretions of non-lactating cows which was thought to be an important marker protein for mammary function during involution (Rejman and Hurley, 1988; Kumar et al., 2006).

* Corresponding author. Fax: +91 11 2658 8663.
E-mail address: tps@aiims.aiims.ac.in (T.P. Singh).

1047-8477/$ - see front matter © 2006 Elsevier Inc. All rights reserved.
doi:10.1016/j.jsb.2006.11.002

This glycoprotein has been isolated from the dry secretions of Sheep and is named as SPS-40. It seems to act as a protective signalling factor during extensive tissue remodeling in the early phase of involution. SPS-40 is homologous to the human macrophage chitinase (Renkema et al., 1995; Fusetti et al., 2002) (HCHT) (52% sequence identity) as well as to another mammalian chitinase-like protein (YM1) (Sun et al., 2001; Tsai et al., 2004) (47% sequence identity). The amino acid sequence of SPS-40 indicates that it does not have chitin hydrolyzing activity of chitinases due to a mutation of an active site residue from Glu to Leu. Several other animal glycoproteins including a 39 kDa human cartilage glycoprotein (Johansen et al., 1993; Hakala et al., 1993) (HCGP-39), bovine chondrocyte chitinase-like protein (CLP-1) (GenBank Accession Number AF011373), porcine heparin-binding glycoprotein (Shackelton et al., 1995) (GP38k), rat cartilage glycoprotein (RGP39) (GenBank Accession Number AF062038) and a breast regression protein (Morrison and Leder, 1994) (BRP39) are highly homologous to SPS-40 (identities ranging from 95 to 68%) and constitute a single family named hereafter as SPX-40 proteins. So far, crystal structures of only five native proteins MGP-40 (Mohanty et al., 2003), HCGP-39 (Houston et al., 2003; Fusetti et al., 2003), SPC-40 (Kumar et al., 2006), SPS-40 (Srivastava et al., 2006) and of complexes of HCGP-39 with various chitin like oligosaccharides have been determined (Houston et al., 2003; Fusetti et al., 2003). The structural results shown that the five native proteins have similar overall foldings although the carbohydrate-binding grooves of SPX-40 (Kumar et al., 2006; Srivastava et al., 2006) differ significantly from those of chitinases and HCGP-39. The complexes of HCGP-39 with carbohydrates have also been reported showing different binding behaviour for short and long oligosaccharides (Houston et al., 2003; Fusetti et al., 2003). Therefore, the carbohydrate-binding data of chitinases and HCGP-39 could not be correlated. In order to understand the mechanism of binding of oligosaccharides with SPX-40 proteins particularly where they lack a chitin hydrolyzing capability, we have determined the crystal structures of SPS-40 proteins with various oligosaccharides of different lengths. There are several questions regarding the function of SPS-40 and its interactions with oligosaccharides: Do the interactions of carbohydrates with SPS-40 suggest a specific recognition role for these proteins? What is the optimum size of carbohydrate chain? Which class of carbohydrates do these proteins prefer? How many subsites are accessible in the carbohydrate-binding groove? Which of the subsites are more preferred for binding? In order to answer these questions, we report the structures of four complexes of SPS-40 with four *N*-acetyl glucosamine oligosaccharides ($GlcNAc_{n,n=3-6}$). The results of these studies clearly indicate the binding preferences of various oligosaccharides and accompanied conformational changes.

2. Experimental procedure

2.1. Purification

Fresh mammary secretions were collected from sheeps maintained at the Indian Veterinary Research Institute, Izatnagar, India. Secretions were collected on days 4, 8, 12, 16 and 20 after the onset of involution period (day 0 being the last milking day). In order to protect the protein from degradation by enzymes whose concentrations are also increased during involution (Aslam and Hurley, 1997), protease inhibitors (Fang and Sandholm, 1995) were added. The samples were pooled and skimmed. These were diluted twice with 50 mM Tris–HCl (pH 7.8). Cation exchanger CM-Sephadex C-50 ($7\,gl^{-1}$) equilibrated in 50 mM Tris–HCl, pH 7.8 was added and stirred slowly for about 1 h with a mechanical stirrer. The gel was allowed to settle and the solution was decanted. In order to remove the unbound proteins, the protein bound gel was washed with an excess of 50 mM Tris–HCl (pH 7.8). The washed protein bound gel was packed in a column (25×2.5 cm) and washed with the same buffer containing 0.2 M NaCl, which removed other impurities. The remaining proteins were eluted with the same buffer containing 0.5 M NaCl. The eluted protein solution was desalted and was again passed through CM-Sephadex C-50 column (10×2.5 cm) which was pre-equilibrated with 50 mM Tris–HCl (pH 7.8) and eluted with a linear gradient of 0.05–0.5 M NaCl in the same buffer. The elution profile showed three peaks. The peak corresponding to 0.3 M NaCl was pooled and was concentrated using an Amicon ultrafiltration cell (Bedford, USA). The concentrated samples were passed through a Sephadex G-100 column (100×2 cm) using 50 mM Tris–HCl, pH 7.8 containing 0.5 M NaCl. The second peak in this final chromatographic step corresponded to a molecular mass of 40 kDa as indicated by Matrix-assisted laser desorption/ionization-time of flight mass spectrometery (MALDI-TOF) (Kratos-Shimadzu, Kyoto, Japan) and sodium dodecyl sulfate polyacrylamide gel electrophoresis (SDS–PAGE). The protein samples were blotted on a Polyvinylidene fluoride (PVDF) membrane. The sequence of first 20 amino acid residues from the N-terminus was determined using a protein sequencer PPSQ-20 (Shimadzu, Kyoto, Japan) that showed an identity of more that 90% with the already known N-terminal sequence of goat signalling protein (Mohanty et al., 2003).

2.2. Fluorescence studies of protein–carbohydrate binding

In order to evaluate the binding characteristics of carbohydrates to SPS-40, various sugar compounds were used that included, monosaccharides such as glucose (Glc), *N*-acetylglucosamine (GlcNAc), glucosamine (GlcN), galactose (Gal), *N*-acetylgalactosamine (GalNAc) and mannose (MAN); disaccharides such as $GlcNAc_2$, lactose and trehalose and oligosaccharides such as $GlcNAc_3$, $GlcNAc_4$, $GlcNAc_5$ and $GlcNAc_6$ (Sigma Chemical Co., St. Louis, USA). For the