FEBS Letters 587 (2013) 3620–3625

journal homepage: www.FEBSLetters.org

Conformational change of a unique sequence in a fungal galectin from *Agrocybe cylindracea* controls glycan ligand-binding specificity

Naoyuki Kuwabara [a], Dan Hu [b], Hiroaki Tateno [b], Hisayoshi Makyio [a], Jun Hirabayashi [b], Ryuichi Kato [a,*]

[a] *Structural Biology Research Center, Photon Factory, Institute of Materials Structure Science, High Energy Accelerator Research Organization (KEK), 1-1 Oho, Tsukuba, Ibaraki 305-0801, Japan*

[b] *Research Center for Stem Cell Engineering, National Institute of Advanced Industrial Science and Technology (AIST), Central 4, 1-1-1 Umezono, Tsukuba, Ibaraki 305-8568, Japan*

ARTICLE INFO

Article history:
Received 14 August 2013
Revised 28 August 2013
Accepted 28 August 2013
Available online 10 September 2013

Edited by Miguel De la Rosa

Keywords:
Galectin
Glycan
Saccharide
ACG
X-ray crystallography
Protein engineering

ABSTRACT

A fungal galectin from *Agrocybe cylindracea* (ACG) exhibits broad binding specificity for β-galactose–containing glycans. We determined the crystal structures of wild-type ACG and the N46A mutant, with and without glycan ligands. From these structures and a saccharide-binding analysis of the N46A mutant, we revealed that a conformational change of a unique insertion sequence containing Asn46 provides two binding modes for ACG, and thereby confers broad binding specificity. We propose that the unique sequence provides these two distinct glycan-binding modes by an induced-fit mechanism.

Structured summary of protein interactions:
ACG and **ACG** bind by x-ray crystallography (View interaction)

1. Introduction

Galectins constitute an evolutionarily conserved family of β-galactoside-binding proteins [1]. Members of the galectin family are found in mammals, amphibians, fish, nematodes, sponges, and some fungi [2]. Galectins carry out intra- and extra-cellular functions through glycoconjugate-mediated recognition mediated by a conserved carbohydrate recognition domain (CRD). The CRD consists of ~130 amino acids, which fold into a β-sandwich structure comprising two anti-parallel β-sheets (the F-sheet and S-sheet), with the concave surface of the S-sheet forming a groove to accommodate the glycan ligands. In addition to the basic β-galactose-binding site, which is made up of about six residues shared by most members of the family, galectins also contain adjacent binding sites that determine the fine specificity of each galectin for longer β-galactoside-containing glycans such as sialylated glycans, blood-group antigens, poly-*N*-acetyllactosamine, and sulfated glycans [3].

ACG, a fungal galectin from *Agrocybe cylindracea*, recognizes a wide range of glycan epitopes; these include not only typical structures such as Galβ1–4GlcNAc (LacNAc) and Galβ1–3GalNAc (T antigen), but also many derivatives in which the C3 position of the terminal β-galactose is substituted with Siaα2-3, Sulfo-3, or GalNAcα1-3 [4,5]. Recently, we assessed the importance of the β-galactoside-binding amino-acid residues of ACG by substituting five key residues with alanine [6]. Surprisingly, single amino-acid substitution at Pro45 (P45A) or Asn46 (N46A) in ACG drastically changed the specificity profile: the mutant proteins acquired a distinguishably increased binding affinity for GalNAcα1–3Galβ-containing glycans, but completely lost the wild-type protein's affinity for other β-galactosides.

To elucidate the structural mechanisms underlying broad (or specific) substrate recognition, we solved the crystal structures of wild-type ACG and the N46A mutant in complex with lactose or blood group A tetraose (A-tetra, GalNAcα1–3(Fucα1-2)Galβ1–4GlcNAc), and compared these structures with those of the ligand-free forms. The *cis* conformation of Pro45 was observed in the lactose-bound form, whereas the *trans* conformation was observed in the ligand-free and A-tetra-bound forms, suggesting that *cis–trans* conversion of Pro45 accompanying the conformational change of the unique insertion sequence governs the induced-fit mechanism that is responsible for the broad specificity of ACG. In this study, we revealed the structural basis by which substitutions of Pro45, Asn46,

Abbreviations: ACG, *Agrocybe cylindracea* galectin; A-tetra, GalNAcα1–3(Fucα1-2)Galβ1–4GlcNAc; CRD, carbohydrate recognition domain; RMSD, root-mean-square deviation; PEG, polyethylene glycol

* Corresponding author. Fax: +81 29 879 6179.
E-mail address: ryuichi.kato@kek.jp (R. Kato).

0014-5793/$36.00
http://dx.doi.org/10.1016/j.febslet.2013.08.046

and Asn47 residues abolish the basic β-galactose-binding mode while preserving affinity for GalNAcα1–3Galβ-containing glycans.

2. Materials and methods

2.1. Purification and crystallization

Construction and purification of wild-type and N46A-mutant ACG, each fused with a FLAG tag at its N-terminus, were described previously [6]. Blood group A antigen tetraose type 2 (A-tetra: GalNAcα1–3(Fucα1–2)Galβ1–4GlcNAc) was purchased from Oligo-Tech (ELICITYL, France). For crystallization, protein-containing buffer was exchanged with 10 mM Tris–HCl, 150 mM NaCl (pH 7.5) by dialysis, followed by concentration up to 20 mg/ml protein. Initial crystallization trials were performed using an automated protein crystallization and monitoring system [7]. After screening, ACG was crystallized by the hanging-drop vapor diffusion method at 20 °C in buffer containing 30% PEG 4000, 0.1 M Na citrate (pH 5.6), and 0.2 M ammonium acetate, with or without 5 mM lactose. We could not obtain crystals of the ligand-free form of wild-type ACG, even though the protein sample was subjected to several rounds of dialysis before crystallization attempts. Lactose-bound N46A-mutant crystal was also not obtained because of its low affinity for lactose. The N46A mutant in complex with A-tetra was crystallized in 30% PEG 1500 and 5 mM A-tetra by the streak seeding method using wild-type crystals as seeds.

2.2. Structure determination and refinement

Crystals were cryo-protected by soaking them in reservoir solution supplemented with 2.5–5.0% ethylene glycol. After flash cooling, diffraction data were collected at 100 K using a CCD detector (ADSC Quantum 270) on beamline AR-NE3A of the Photon Factory at KEK (Tsukuba, Japan). The data were processed using the HKL2000 software suite [8]. Phases were determined by the molecular replacement method using chain A of 1WW7 (wild-type ACG ligand-free form) as a search model [9]. Subsequent crystallographic refinements were conducted using the programs COOT [10] and REFMAC5 [11]. The statistics of data collection and refinements are summarized in Table 1. RMSD and ligand-recognition plots were prepared using lsqkab [12] and LIGPLOT+ [13], respectively.

3. Results and discussion

3.1. Overall structure

Four crystal structures (wild-type ACG with lactose, wild-type ACG with A-tetra, N46A mutant with A-tetra, and ligand-free N46A mutant) were determined by X-ray crystallography at resolutions between 1.3 and 2.2 Å (Table 1). In these structures, each asymmetric unit contains a ACG dimer, consistent with the fact that ACG forms a dimer in solution [4]. Because the two protomers of the dimer are almost identical to each other, we refer to chain A of each crystal structure in the following discussion. The space groups of these structures are hexagonal primitive, in contrast to those of previously reported structures, which belonged to space groups $P2_12_12_1$ or $P1$ [5]. In this study, we crystallized ACG fused to a FLAG-tag at the N-terminus and to a KLAAALE sequence from the vector at the C-terminus. These additional sequences might affect the crystal packing and change the space group. In spite of the different space groups, overall our structures were similar to previously reported ones (Fig. 1A). The overall RMSD values between the wild-type lactose complex and the previously reported structures (PDB ID: 1WW6, lactose complex; PDB ID: 1WW7, ligand-free form) are 0.3–0.5 Å. Thus, any differences in crystal packing between the structures do not appear to seriously distort the overall structure of ACG.

Fig. 1B and C show Cα comparisons between our four structures and the previously reported ones (PDB IDs: 1WW6 and 1WW7). The RMSD between the wild-type A-tetra complex and other ACG structures is 0.9 Å. Almost all of the structures are similar with each other; however, a striking difference is observed in the unique sequence (Ala40–Asn47) (Figs. 1B and C and 2). The RMSD values of the unique sequence are higher than 2.0 Å. In addition, another region (Ser141–Ser149), structurally adjacent to the unique sequence, also has a relatively high RMSD value. This comparison revealed that the conformation of the Ala40–Asn47 region is changed

Table 1
Data collection and refinement statistics.

Ligand	Wild-type lactose	Wild-type A-tetra	N46A A-tetra	N46A ligand-free
Data collection statistics				
Wavelength (Å)	1.0000	1.0000	1.0000	1.0000
Resolution range (Å)[a]	35.0–1.9 (1.97–1.90)	35.0–1.35 (1.40–1.35)	35.0–1.6 (1.66–1.60)	35.0–2.2 (2.28–2.20)
Space group	$P6_1$	$P6_5$	$P6_5$	$P6_1$
Unit cell (Å)	$a = b = 125.4$, $c = 56.6$	$a = b = 102.9$, $c = 74.7$	$a = b = 102.8$, $c = 74.6$	$a = b = 124.9$, $c = 57.0$
Total reflections	452885	1068800	665598	291328
Unique reflections	40147	98506	59060	25919
Multiplicity[a]	11.3 (11.3)	10.8 (8.1)	11.3 (11.2)	11.2 (11.2)
Completeness (%)[a]	100.0 (99.8)	99.9 (99.9)	99.9 (100.0)	99.8 (99.9)
$\langle I/\sigma I \rangle$[a]	9.0 (4.0)	33.2 (4.92)	28.92 (7.0)	6.5 (3.5)
Refinement statistics				
R_{work} (%)	17.3	14.0	14.8	21.1
R_{free} (%)	19.6	15.7	17.4	24.4
Number of atoms	2626	3011	2883	2418
Macromolecules	2405	2504	2449	2302
Ligands	46	111	107	10
Water	175	396	327	106
Protein residues	320	318	313	304
RMS bond length (Å)	0.016	0.009	0.011	0.008
RMS bond angles (°)	1.89	1.20	1.28	1.12
Ramachandran favored (%)	94	98	98	96
Ramachandran outliers (%)	1.9	0.0	0.0	0.0

[a] Statistics for the highest-resolution shell are shown in parentheses.