Biochimica et Biophysica Acta 1810 (2011) 150–161

Contents lists available at ScienceDirect

Biochimica et Biophysica Acta

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

Human galectin-3 (Mac-2 antigen): Defining molecular switches of affinity to natural glycoproteins, structural and dynamic aspects of glycan binding by flexible ligand docking and putative regulatory sequences in the proximal promoter region

Mickaël Krzeminski [a,1], Tanuja Singh [b,c,1], Sabine André [b], Martin Lensch [b], Albert M. Wu [c,*], Alexandre M.J.J. Bonvin [a,*], Hans-Joachim Gabius [b,*]

[a] Bijvoet Center for Biomolecular Research, Faculty of Science, Utrecht University, Padualaan 8, 3584 CH Utrecht, The Netherlands
[b] Institute of Physiological Chemistry, Faculty of Veterinary Medicine, Ludwig-Maximilians-University, Veterinärstr. 13, 80539 Munich, Germany
[c] Glyco-Immunochemistry Research Laboratory, Institute of Molecular and Cellular Biology, College of Medicine, Chang-Gung University, Kwei-san, Tao-yuan 333, Taiwan

A R T I C L E I N F O

Article history:
Received 27 August 2010
Received in revised form 29 October 2010
Accepted 2 November 2010
Available online 8 November 2010

Keywords:
Glycosylation
Lectin
Macrophage
Modeling
Mucin
Sialylation

A B S T R A C T

Background: Human galectin-3 (Mac-2 antigen) is a cell-type-specific multifunctional effector owing to selective binding of distinct cell-surface glycoconjugates harboring β-galactosides. The structural basis underlying the apparent preferences for distinct glycoproteins and for expression is so far unknown.
Methods: We strategically combined solid-phase assays on 43 natural glycoproteins with a new statistical approach to fully flexible computational docking and also processed the proximal promoter region *in silico*.
Results: The degree of branching in *N*-glycans and clustering of core 1 *O*-glycans are positive modulators for avidity. Sialylation of *N*-glycans in α2–6 linkage and of core 1 *O*-glycans in α2–3 linkage along with core 2 branching was an unfavorable factor, despite the presence of suited glycans in the vicinity. The lectin–ligand contact profile was scrutinized for six natural di- and tetrasaccharides enabling a statistical grading by analyzing flexible docking trajectories. The computational analysis of the proximal promoter region delineated putative sites for Lmo2/c-Ets-1 binding and new sites with potential for RUNX binding.
General significance: These results identify new features of glycan selectivity and ligand contact by combining solid-phase assays with *in silico* work as well as of reactivity potential of the promoter.

1. Introduction

Protein–glycan recognition is being delineated as a key molecular interaction route based on the exceptional capacity of oligosaccharides for high-density information coding and the emergence of various families of lectins [1]. Despite the glycomic diversity on cell surfaces this process has an exquisite target selectivity which underlies the precision of the post-binding biosignaling [2]. To define the nature of the mechanisms and the structural basis for the inherent selectivity are thus questions with conspicuous biorelevance. Here, we present a combined approach toward this end, studying binding preferences of a multifunctional human lectin. We tested various ligand structures and topology of presentation in combination with flexible ligand docking and statistical analysis of contact.

The Mac-2 antigen was originally defined as rat macrophage marker present abundantly on thioglycollate-induced peritoneal exudate cells [3]. It was later confirmed to be expressed in the monocyte–macrophage lineage and detected also in epithelial and stromal cells as well as in diverse tumor cell types [4–6]. The human Mac-2 protein, known as major non-integrin binding partner for laminin and IgE receptor, was then identified to be a β-galactoside-specific lectin [7]. The characteristics of the C-terminal sequence section, the lectin part, led to its assignment to the galectin family, herein as galectin-3 [8,9]. Its lectin property plays an important role in defense mechanisms. For instance, it enables macrophages to target di-*N*-acetylated lactose (LacdiNAc) of helminth parasites and the major xenoantigen (α-Gal) [10,11]. Evidently, distinct glycan epitopes serve as sugar-encoded signals for the lectin in immune defense. This

Abbreviations: BSM, bovine submaxillary mucin; Gal, D-galactose; GalNAc, *N*-acetyl-D-galactosamine; Glc, D-glucose; GlcNAc, *N*-acetyl-D-glucosamine; gp, glycoprotein; HADDOCK, High Ambiguity Driven DOCKing; hGal-3, human galectin-3; LacNAc, *N*-acetyllactosamine; LacdiNAc, di-*N*-acetylated lactose; Man, D-mannose; OSM, ovine submaxillary mucin; PSM, porcine submaxillary mucin; THGP, Tamm–Horsfall urinary glycoprotein; TBS-T, Tris/HCl-buffered saline containing 0.05% Tween 20; TF, Thomsen–Friedenreich antigen (Galβ1–3GalNAcα)

* Corresponding authors. A.M. Wu is to be contacted at tel.: +886 3 211 8966; fax: +886 3 211 8456, +886 3 211 8700. A.M.J.J. Bonvin, tel.: +31 30 2533850; fax: +31 30 2537623. H.-J. Gabius, tel.: +49 89 2180 2290; fax: +49 89 2180 2508.
E-mail addresses: amwu@mail.cgu.edu.tw (A.M. Wu), a.m.j.j.bonvin@uu.nl (A.M.J.J. Bonvin), gabius@lectins.de (H.-J. Gabius).
[1] These authors contributed equally to this manuscript.

0304-4165/$ – see front matter
doi:10.1016/j.bbagen.2010.11.001

also holds true for adhesion/growth-regulatory events triggered by galectin-3, observed for example with thymocytes or carcinoma/neuroblastoma cells [2,12–15].

To exert its specific functions, human galectin-3 (hGal-3) requires a discriminatory reactivity to the different types of oligosaccharides based on a substituted β-galactoside core in natural glycoconjugates. Binding and inhibition assays, frontal affinity chromatography, glycan arrays and titration calorimetry revealed a strong reactivity with β1–4-linked galactosides, especially oligomeric *N*-acetyllactosamine (LacNAc), and histo-blood group ABH epitopes with indications for a minor influence of α2–3-sialylation and *N*-glycan branching in free glycans, and assay-dependent variation of avidity for *N*- *vs O*-glycans [16–25]. Due to the rather common presence of theoretically suited epitopes at the termini of glycan branches, binding of hGal-3 to many cell surface glycoconjugates can be expected. But, however, this lectin is very selective when engaging in contacts with cell surface glycoconjugates.

Despite the abundance of glycoconjugates with terminal β-galactosides only a limited set of glycoproteins/glycolipids qualifies as physiological galectin-3 ligand, and this with cell-type selectivity. Testing for example colon cancer cells, carcinoembryonic antigen, laminin and lysosome-associated membrane proteins 1/2 as well as mucins (especially the mucin MUC1 *via O*-glycans and the *N*-glycan of MUC1-C at Asn36) were bioactive [15,26–28]. The glycoproteins CD29, CD43, CD45, CD71, and CD98 (H-chain) and the Mac-2-binding protein were counterreceptors in the case of thymocytes, for hippocampal neurons the neural cell adhesion molecule L1 and for neuroblastoma cells ganglioside GM1 [13,14,29]. Thus, the presence of a cognate carbohydrate structure combined with other, not yet precisely defined topological factors appears to underlie the distinct binding properties. It is noteworthy in this context that hGal-3's tridomain structure, especially its five collagenase-sensitive tandem-repeat modules, endows the only chimera-type member of this lectin family with unique properties for cooperativity of binding and oligomerization in the presence of multivalent ligands [30–33]. This characteristic gives special reason to test natural glycoproteins with different degrees of valency and branch-end structure. Compared to free *N*- and *O*-glycans the spatial distribution of the glycan chains attached to a protein carrier is assumed to partake in regulation of lectin affinity [34–36]. To what extent factors imposed by glycan presentation on a natural protein affect the binding of this potent effector in the immune system and beyond has not yet been addressed. In order to take the fine-specificity analysis of hGal-3 to the topological level, probing well-characterized glycoproteins comparatively under identical conditions is a promising approach.

We analyze here 43 glycoprotein preparations with differences in branch-end and core structures including sialylation and in degree of branching as well as a bacterial polysaccharide. These binding assays are supplemented with fine-specificity monitoring using 41 glycan compounds as inhibitors of hGal-3/glycoprotein binding. The experimental part is combined with *in silico* calculations to monitor the range of interactions in complexes with six cognate di- and tetrasaccharides. Using HADDOCK version 2.1 [37,38], a flexible docking approach that has been introduced previously to the study of protein–carbohydrate interactions [39], we visualize the dynamics within the interaction profile of a lectin for the first time at a statistical level. The presented results contribute to the efforts to explain the origin of affinity, specificity and eventually intrafamily differences in physiological ligand selection, with implications for drug design. These differences are also manifested on the level of the expression patterns of galectins, for example detected by immunohistochemical fingerprinting of human tumors with prognostic relevance [40–42]. To help delineate the underlying sequence signals, we present a comparative *in silico* dissection of the proximal promoter region of the hGal-3 gene with those of three proto-type galectins, a step to spot putative sites for transcription factors and discern non-uniform features.

2. Materials and methods

2.1. Galectin-3

Human galectin-3 was obtained by recombinant production and purified by affinity chromatography on lactosylated Sepharose 4B, prepared after resin activation with divinyl sulfone, as crucial step. It was subjected to rigorous quality controls by one- and two-dimensional gel electrophoresis, gel filtration and mass spectrometry, then biotinylated under activity-preserving conditions with the *N*-hydroxysuccinimide ester derivative of biotin [12,31]. The degree of biotinylation was determined by two-dimensional gel electrophoresis and mass spectrometric fingerprinting [43,44]. Lack of influence of the labeling on lectin activity was ascertained by solid-phase and cell-binding assays using asialofetuin and human SW480 colon carcinoma cells, respectively [45,46].

2.2. Glycoproteins and saccharides

Sources and further processing of the glycoproteins have been given previously in the first reports on galectin reactivity [47,48]. The predominant carbohydrate determinants at branch ends with affinity to galectins are listed in Table 1. Due to its importance for galectin binding, information on branch-end sialylation is as follows for the tested glycoproteins:

- *N*-glycans of bovine lactoferrin are exclusively α2–6-sialylated, and in addition terminal LacdiNAc and Galα1–3Galβ1–4GlcNAc (B-like) epitopes are present [49]. The same applies to human glycophorin, which can naturally present terminal galactose in one arm due to incomplete sialylation [50].
- Porcine thyroglobulin, human α_1-acid glycoprotein and bovine fetuin all contain α2–3/6-sialylated *N*-acetyllactosamine with branch specificity, e.g. α2–3 sialylation resides in the β1–4-branch of the α1–3-arm (for further information, please see [51]).

The *Pneumococcus* type 14 polysaccharide was a generous gift from the late Dr. E. A. Kabat (Department of Microbiology, Columbia University, NY, USA). Mono-, di- and oligosaccharides used were obtained from Dextra (Reading, Berkshire, UK) or Sigma (Munich, Germany).

2.3. Enzyme-linked lectinosorbent assay

The assay was performed under conditions deliberately kept constant to allow comparison of the different galectins tested [39,47,48,52]. In detail, the volume of each reagent solution applied to wells of the plate was 50 μl/well, and all incubations, except for coating, were performed at 20 °C. The reagents, if not indicated otherwise, were diluted with Tris-buffered saline (TBS; 0.05 M Tris–HCl, 0.15 M NaCl, pH 7.35) containing 0.05% Tween 20 (TBS-T). TBS-T was used for washing the plates between incubation steps.

The surface of the 96-well microtiter plate wells (Nunc-Immuno, Kamstrupvej, Denmark) was coated with glycoproteins dissolved in 0.05 M sodium carbonate buffer (0.05 M $NaHCO_3$/0.05 M Na_2CO_3, pH 9.6) overnight at 4 °C. After washing the plate, solution with biotinylated hGal-3 (5 μg/ml) was added and the plate was incubated for 30 min. The plates were then carefully washed to remove any free lectin. The ExtrAvidin/alkaline phosphatase solution (diluted 1:10,000; Sigma) was added thereafter to detect the specifically bound hGal-3 by its biotin moieties. Carbohydrate dependence of binding was routinely ascertained by haptenic inhibition. After 1 h the plates were washed at least four times to remove free conjugate and incubated with a solution of *p*-nitrophenyl phosphate (Sigma