Biochimie 104 (2014) 90–99

Contents lists available at ScienceDirect

Biochimie

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

Research paper

Human chimera-type galectin-3: Defining the critical tail length for high-affinity glycoprotein/cell surface binding and functional competition with galectin-1 in neuroblastoma cell growth regulation

Jürgen Kopitz [a, *], Sabine Vértesy [b], Sabine André [b], Sabine Fiedler [c], Martina Schnölzer [c], Hans-Joachim Gabius [b]

[a] Abteilung für Angewandte Tumorbiologie, Zentrum Pathologie, Klinikum der Ruprecht-Karls-Universität, Im Neuenheimer Feld 224, 69120 Heidelberg, Germany
[b] Institut für Physiologische Chemie, Tierärztliche Fakultät, Ludwig-Maximilians-Universität, Veterinärstraße 13, 80539 München, Germany
[c] Funktionelle Proteomanalyse, Deutsches Krebsforschungszentrum, Im Neuenheimer Feld 580, 69120 Heidelberg, Germany

ARTICLE INFO

Article history:
Received 15 April 2014
Accepted 26 May 2014
Available online 6 June 2014

Keywords:
Adhesion
Glycoprotein
Lectin
Matrix metalloproteinase
Neuroblastoma
Proliferation

ABSTRACT

Many human proteins have a modular design with receptor and structural domains. Using adhesion/growth-regulatory galectin-3 as model, we describe an interdisciplinary strategy to define the functional significance of its tail established by nine non-triple helical collagen-like repeats (I–IX) and the N-terminal peptide. Genetic engineering with sophisticated mass spectrometric product analysis provided the tools for biotesting, i.e. eight protein variants with different degrees of tail truncation. Evidently,various aspects of galectin-3 activity (*cis* binding and cell bridging) are affected by tail shortening in a different manner. Thus, this combined approach reveals an unsuspected complexity of structure–function relationship, encouraging further application beyond this chimera-type galectin.

1. Introduction

Human lectins are the natural tools to read glycan-encoded information and translate it into cellular effects [1]. This process critically depends on structural features of the protein. Di- or oligomeric displays of carbohydrate recognition domains (CRDs) enable ligand cross-linking or high-affinity binding of clustered ligands [2]. As a means toward this end, a CRD can be associated with modules that build the suited quaternary structure [3,4]. If phylogenetically conserved, presence of such sequence motifs in lectins is assumed to be of broad functional relevance. This is the case for the collagen-like tail of collectins and ficolins, which is extended by a relatively short N-terminal stretch (7–25 amino acids) equipped with cysteines: following proper folding of the collagen triple helix, disulphide bridging will then ensure stable lectin aggregation to tri- and multimers to facilitate homing in on bacterial and yeast surfaces (with their typical high-density presentation of binding sites) in host defence [5,6]. An association of a CRD with non-triple helical collagen (gelatin)-like repeats and a short starting peptide is invariably found in a different family of lectins, i.e. the galectins, in all vertebrates studied. Fittingly, this protein is referred to as its chimera-type member, termed galectin-3 (Gal-3) [7–9]. However, the physiological significance of this trimodular structural organization is in this case much less clear than for collectins.

In detail, the N-terminal peptide, which always contains two Ser acceptors of phosphorylation [10–13], is followed by a number of collagen-like repeats (nine for human Gal-3) completing the N-terminal tail of the CRD (Fig. 1). A survey of the structural organization of Gal-3 in vertebrates reveals a minimum number of eight with a range from eight to 13 repeat units (Supplementary material, Fig. S1). Underscoring similarity of this section to gelatin, gelatinase-A (matrix metalloproteinase-2, MMP-2), gelatinase-B (matrix metalloproteinase-9, MMP-9) and bacterial collagenases can shorten the non-CRD part, up to total removal to produce fully truncated Gal-3 (trGal-3) by bacterial enzymes [14–16]. In addition, MMPs-7 (matrilysin-1) and -13 (collagenase-3) as well as prostate-specific antigen (PSA), a Zn^{2+}-sensitive serine protease, can use Gal-3 as substrate [17–19]. Physiologically, the MMP-9-dependent

* Corresponding author. Tel.: +49 6221 564227; fax: +49 6221 565981.
E-mail address: juergen.kopitz@med.uni-heidelberg.de (J. Kopitz).

http://dx.doi.org/10.1016/j.biochi.2014.05.010
0300-9084/

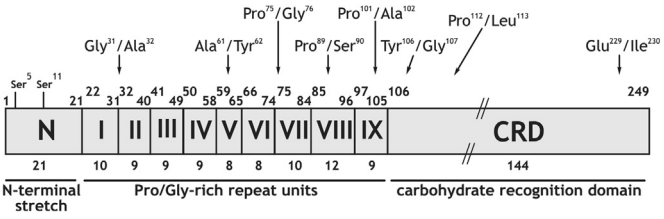

Fig. 1. Structural organization of the trimodular design of human Gal-3. Human Gal-3 is composed of an N-terminal stretch (N), nine Pro/Gly-rich repeat units (I–IX) and the CRD. The positions of the two serine residues as acceptors for phosphorylation are illustrated and seven cleavage sites for MMPs and PSA at positions Gly^{31}/Ala^{32} (MMP-2), Ala^{61}/Tyr^{62} (MMP-2, MMP-7, MMP-9, MMP-13), Pro^{75}/Gly^{76} (MMP-7), Pro^{89}/Ser^{90} (MMP-13), Pro^{101}/Ala^{102} (MMP-13), Tyr^{106}/Gly^{107} (PSA) and Pro^{112}/Leu^{113} (MMP-7) are indicated by arrows. The collagenase from *Clostridium histolyticum* commonly used for complete tail removal hits the PSA-specific site and also the linkage between Glu^{229}/Ile^{230} close to the end of the CRD.

processing appears to be a regulatory step in endochondral bone formation on both sides of the chondro-osseous junction in mice [20]. The main cleavage site for MMPs in human Gal-3 is at the Ala61/Tyr62 bond in the fifth repeat [16], as highlighted along with all other known sites in Fig. 1. It is predisposed in the repeats due to the presence of a His (not Pro) residue distally. Of note, a single nucleotide polymorphism (C to A; rs 4644), entering a Pro also at this site [21] and hereby shifting MMP-2-dependent cleavage to the Gly31/Ala32 bond, correlates with breast cancer incidence [22]. This association intimates clinical relevance of certain types of proteolytic processing. Apparently, the how of tail shortening can have so far not fully disclosed *in vivo* significance, assumedly acting as molecular functionality switch when surpassing a critical level.

Reflecting the literature status, the collagen-like region is considered as a means towards Gal-3 oligomerization via tail-to-tail interaction (as in other proteins with such a tail such as synexin-VII), with low propensity when free in solution but favoured by presence of suitable polyvalent ligands and on cell surfaces [23–31]. Interestingly, the bimodal charge distribution of Gal-3 in nano-electrospray ionization mass spectrometry (nESI MS) with occurrence of a highly charged form indicates presence of a rather linear form in solution [32], as do the Y-shaped proteins (probably dimers) seen in electron micrographs [25]. In addition, the CRD, too, can be involved in Gal-3 aggregation when loaded with ligand [33–35]. Using calorimetry to study this process, the dissociation constants for Gal-3 binding of free oligosaccharides and N-glycans of the glycoprotein asialofetuin are independent of the quaternary structure [36,37]. Intracellularly, the deletion of the tail completely blocked the routing of Gal-3 to late endosomes/lysosomes in human endothelial cells [38], either repeat-based sequences or the starting peptide were found involved in secretion and nuclear delivery in different cell types [39–41]. In all these cases, truncations by enzymatic cleavage or genetic engineering were instrumental to probe into structure-activity correlations. Since these experiments were mostly focused on defining effects of complete tail removal, it is still an intriguing and open issue to relate the individual sections of the tail to functions.

To answer this question with respect to carbohydrate-dependent cell surface binding and mediation of cell-bridging contacts, we first genetically engineered cDNAs, optimized recombinant protein production, then purified a panel of human Gal-3 variants with increasing degree of tail shortening, verified the predicted sequence by mass spectrometry and proved their glycan-binding capacity (for overview of the molecular design of the eight variants, please see Fig. 2). Using a cell model with marked difference between full-length and truncated Gal-3 in i) affinity to surface-presented ligands (ganglioside GM1) in direct cell binding and ii) competition

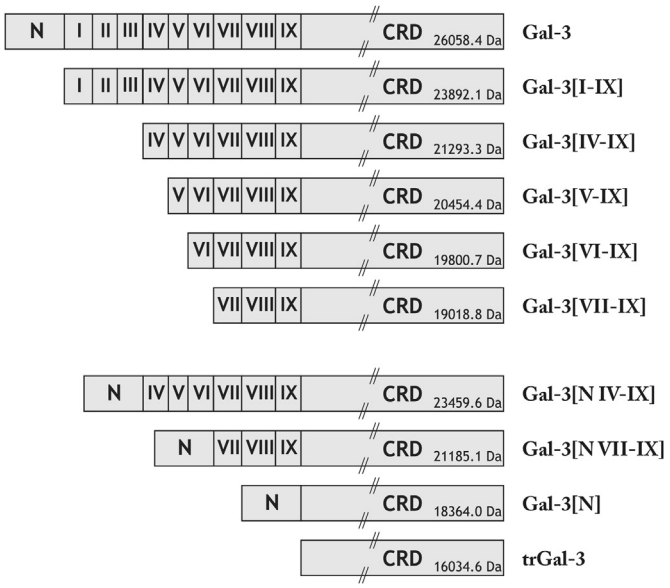

Fig. 2. Structural organization of the Gal-3 variants. Upper part: the five Gal-3 variants without the N-terminal stretch (N) and consecutively increased degree of removal of the collagen-like repeat units. Bottom part: the three variants with N-terminal stretch and blockwise reduction of number of collagen-like repeat units as well as the fully truncated Gal-3 (trGal-3). The predicted molecular mass is given in each case for a protein after removal of N-formyl Met and without N-terminal acetylation (for detailed mass and sequence information, please see Supplementary material, Fig. S2 (predicted amino acid sequence and N-terminal sequence determined by analysis of reISD spectra) and Table S2 (predicted masses and masses determined by analysis of MALDI-TOF spectra of intact protein)).

with growth-regulatory galectin-1 (Gal-1) [27,42–45], the variants' properties were systematically determined under strictly identical conditions. These experiments were flanked by measuring capacity to establish cell–cell adhesion (haemagglutination). Both assays revealed a critical length of the tail for Gal-3 activity.

2. Materials and methods

2.1. Cloning, protein expression and purification

Design of the cDNAs for the variants followed two different protocols depending on the intended type of tail shortening, i.e. either common omission of the N-terminal peptide stretch followed by stepwise reduction of the number of repeat units (Fig. 2, upper part with five variants) or maintained starting peptide followed by blockwise repeat removal up to complete deletion of all units (Fig. 2, bottom part with three variants). In the five cases of the first category, a flanking reverse oligonucleotide with an internal *Hind*III restriction site and a variant-type-specific forward oligonucleotide with an internal *Nde*I restriction site were synthesized and applied in polymerase chain reactions (PCR) with Phusion High Fidelity Polymerase™ (New England BioLabs, Frankfurt, Germany) according to the manufacturer's instructions (for primer sequences, please see Supplementary material, Table S1). The amplified fragments were purified by gel electrophoresis in a 1.5% agarose matrix and then further used for cloning (please see below). The cDNAs for variants harbouring the N-terminal stretch were constructed by a 2-step PCR procedure, as exemplarily described in detail here for Gal-3 [N IV–IX]. Two fragments with overlapping regions were generated using the forward flanking oligonucleotide (5′-gctcatatggcagacaatttttcgctc-3′) and reverse flanking oligonucleotide (5′-cgaaagcttttagatctggacataggacaaggtg-3′) with a *Nde*I or *Hind*III restriction site, respectively, and an oligonucleotide pair specific for