Preliminary X-ray crystallographic analysis of an engineered variant of human chimera-type galectin-3 with a shortened N-terminal domain

Andrea Flores-Ibarra,^a Federico M. Ruiz,^a Sabine Vértesy,^b Sabine André,^b Hans-Joachim Gabius^b and Antonio Romero^{a*}

Received 3 November 2014

Accepted 1 January 2015

Keywords: collagen; galectin-3; glycome; lipopolysaccharide.

^aDepartment of Chemical and Physical Biology, Centro de Investigaciones Biológicas, CSIC, Ramiro de Maeztu 9, 28040 Madrid, Spain, and ^bInstitut für Physiologische Chemie, Tierärztliche Fakultät, Ludwig-Maximilians-Universität, Veterinärstrasse 13, 80539 München, Germany. *Correspondence e-mail: romero@cib.csic.es

How lectins translate sugar-encoded information into cellular effects not only depends on glycan recognition. Other domains of the protein can contribute to the functional profile of a lectin. Human galectin-3 (Gal-3), an adhesion/growth-regulatory galectin, is composed of three different domains and is thus called a chimera-type protein. In addition to the carbohydrate-recognition domain, this lectin encompasses an N-terminal domain consisting of a peptide harbouring two phosphorylation sites and nine non-triple-helical collagen-like repeats. This region plays an as yet structurally undefined role in Gal-3 aggregation and ligand recognition. To date, crystallization of full-length Gal-3 has not been achieved. With the aim of providing structural insights into this modular organization, a Gal-3 variant was crystallized maintaining the terminal peptide and three of the nine collagen-like repeats. The crystals belonged to the orthorhombic space group $P2_12_12_1$, with unit-cell parameters $a = 94.04$, $b = 97.96$, $c = 236.20$ Å, and diffracted to a resolution of 3.3 Å.

1. Introduction

With the growing realisation that the unsurpassed structural versatility of glycans forms the basis of the sugar code (Gabius *et al.*, 2011), increasing attention is being paid to the structural analysis of sugar receptors (lectins). Among them, the adhesion/growth-regulatory galectins exemplify that distinct aspects of the glycome are translated into cellular effects (Kasai & Hirabayashi, 1996; Cooper, 2002; Kaltner & Gabius, 2012). Indeed, the productive coupling of suitable pairs of lectins and ligands is important for, for example, driving tumour cells into anoikis by a tumour suppressor (André *et al.*, 2007; Amano *et al.*, 2012). As analysis of the carbohydrate-recognition domains (CRDs) of galectins has advanced, two issues in structure–activity correlations have become prominent challenges, *i.e.* the analysis of natural variants (Ruiz *et al.*, 2014) and the role of non-CRD sections in the case of galectin-3 (Gal-3).

Gal-3 has a unique structural design. The presence of the N-terminal peptide with two sites for serine phosphorylation and non-triple-helical collagen-like (Pro/Gly-rich) repeats, as shown in Fig. 1, explains why Gal-3 is referred to as a chimera-type family member (Kasai & Hirabayashi, 1996). In vertebrates, the number of repeats ranges from eight to 13 (Cooper, 2002; Kopitz *et al.*, 2014). By working with full-length and proteolytically truncated proteins of human Gal-3, this N-terminal domain has been implicated in high-affinity cell-surface binding, potential for aggregation (dimerization or pentamerization, favoured by the presence of oligovalent

ligands) and binding processes, for example to *Neisseria* lipopolysaccharides (Ochieng *et al.*, 1993; Kopitz *et al.*, 2001; Ahmad *et al.*, 2004; Quattroni *et al.*, 2012; Dawson *et al.*, 2013). Interestingly, this protein is monomeric in solution, with indications of the presence of two conformers based on a bimodal charge distribution obtained by nano-ESI mass spectrometry (Kopitz *et al.*, 2003; Göhler *et al.*, 2010). Whereas crystallization of the CRD proved to be feasible (Seetharaman *et al.*, 1998), no corresponding study has been completed to date for the full-length protein. Because recent NMR spectroscopy experiments also revealed the possibility of intra-protein interactions between the CRD and the N-terminal domain (Berbís *et al.*, 2014), the need for crystallographic analysis of full-length Gal-3 is obvious.

In order to address this issue in a strategic stepwise manner, we took advantage of recently performed protein engineering to shorten the N-terminal region (Kopitz *et al.*, 2014). With the aim of obtaining initial structural information on both non-CRD modules, we focused on a variant with the N-terminal peptide and three of the nine collagen-like repeats, termed Gal-3(N VII–IX) (Fig. 1). Here, we report the first information on the crystallization of a Gal-3 protein with a shortened N-terminal sequence.

2. Materials and methods

2.1. Protein expression and purification

The protein was produced in *Escherichia coli* BL21 (DE3)-pLysS cells (Promega, Mannheim, Germany) after induction with 400 μ M isopropyl β -D-thiogalactopyranoside for 16 h at 295 K and was purified from bacterial extracts by affinity chromatography on lactosylated Sepharose 4B as the crucial step, as described previously in detail (Kopitz *et al.*, 2014). The protein was precipitated by increasing the saturation of ammonium sulfate to 80% in phosphate-buffered saline (PBS) with 4 mM DTT and 20 mM lactose, and the tubes were kept at 277 K. The final concentration of the protein was 17.5 mg ml⁻¹. Suspensions of precipitated material were dialyzed in PBS containing 4 mM β -mercaptoethanol with a Slide-A-Lyzer Dialysis Cassette (3500 MWCO, Thermo Scientific, Rockford, USA). After affinity chromatography to remove any inactive material, solutions were first concentrated and then fractionated by gel filtration on a HiPrep 16/60

Figure 1

Schematic illustration of the design of full-length Gal-3 with the CRD, the N-terminal peptide (N) and the nine Pro/Gly-rich repeat units (top) and its proteolytically trimmed form (middle). Engineering to combine N and CRD with the repeat units VII–IX shortens the N-terminal region to generate a variant with reduced number of non-triple-helical collagen-like repeats (bottom; from Kopitz *et al.*, 2014 with permission).

Table 1

Data collection and processing.

Values in parentheses are for the outer shell.

Wavelength (Å)	0.9794
Temperature (K)	100
Detector	PILATUS 6M
Crystal-to-detector distance (mm)	502.34
Rotation range per image (°)	0.5
Total rotation range (°)	160
Exposure time per image (s)	0.5
Space group	<i>P</i> 2 ₁ 2 ₁ 2 ₁
<i>a</i> , <i>b</i> , <i>c</i> (Å)	94.04, 97.96, 236.20
Mosaicity (°)	0.18
Resolution range (Å)	46.11–3.30 (3.48–3.30)
Total No. of reflections	193899 (28560)
No. of unique reflections	33543 (4836)
Completeness (%)	99.8 (99.2)
Multiplicity	5.8 (5.9)
$\langle I/\sigma(I) \rangle$	10.1 (2.8)
$R_{\text{merge}}^{\dagger}$	0.169 (0.686)
$R_{\text{r.i.m.}}^{\ddagger}$	0.186 (0.753)
Overall <i>B</i> factor from Wilson plot (Å ²)	60.8

$\dagger R_{\text{merge}} = \frac{\sum_{hkl} \sum_i |I_i(hkl) - \langle I(hkl) \rangle|}{\sum_{hkl} \sum_i I_i(hkl)}$, $\ddagger R_{\text{r.i.m.}} = \frac{\sum_{hkl} \{ [N(hkl)/[N(hkl) - 1]]^{1/2} \sum_i |I_i(hkl) - \langle I(hkl) \rangle| \}}{\sum_{hkl} \sum_i I_i(hkl)}$, where $I_i(hkl)$ is the value of the *i*th intensity measurement and $\langle I(hkl) \rangle$ is the mean value of all measurements of *I*(*hkl*).

Sephacryl S-100 high-resolution column (GE Healthcare, Freiburg, Germany) equilibrated with 20 mM sodium/potassium phosphate pH 7.0 with 150 mM sodium chloride, 4 mM β -mercaptoethanol and 5 mM lactose in an ÄKTApriime system (GE Healthcare, Uppsala, Sweden). Protein-containing fractions were concentrated to 12 mg ml⁻¹. Protein concentration was determined by measuring the absorbance at 280 nm using the calculated extinction coefficient of 9970 M⁻¹ cm⁻¹ (<http://web.expasy.org/protparam>).

2.2. Crystallization

Screening was conducted in 96-well sitting-drop plates (Swissci MRC, Molecular Dimensions, Suffolk, England) at 295 K using a Cartesian Honeybee System (Genomic Solutions, Irvine, USA) and the JBScreen Classic (Jena Bioscience, Jena, Germany), Wizard Classics I–III (Emerald Bio, Bainbridge Island, USA) and Index (Hampton Research, Aliso Viejo, USA) commercial kits. The nanodrops were 0.4 μ l in volume, consisting of 0.2 μ l protein solution (12 mg ml⁻¹) and 0.2 μ l precipitant, and were equilibrated against 50 μ l reservoir solution. Further attempts to optimize the initial crystallization conditions were performed using optimization screens prepared with a Tecan Freedom EVO liquid-handling robot (Männedorf, Switzerland). The selected conditions were (i) 18% (w/v) PEG 8000, 100 mM Tris–HCl pH 8.5, 200 mM lithium sulfate and (ii) 22% PEG 8000, 100 mM MES pH 6.5, 200 mM ammonium sulfate. Different grids varying the PEG molecular weight (PEG 6000, PEG 8000 and PEG 10 000), the PEG concentration (15–30%) and the pH from 6.0 (0.1 M MES) to 9.0 (0.1 M Tris–HCl) were prepared. Additionally, micro-seeding and counter-diffusion techniques were performed with the same two described conditions. Seed stocks were prepared by mechanical homogenization of crystals with a glass fibre with a ball at the tip. For seeding, the

droplets (0.5 μ l protein solution and 0.5 μ l reservoir solution) were equilibrated for 1–2 d. Finally, 0.2 μ l of the seed solution was added to each sample using a 5 μ l Hamilton syringe (Reno, USA). The counter-diffusion method was performed using the GCB-Domino Crystallization Kit (Triana Science and Technology, Granada, Spain) using capillaries of 0.2 mm inner diameter.

2.3. Data collection and processing

For X-ray data collection, crystals with approximate dimensions of $0.05 \times 0.05 \times 0.03$ mm were transferred into a cryoprotectant solution composed of the mother liquor plus 20% (w/v) PEG 400 and cooled in liquid nitrogen prior to data collection. X-ray diffraction data were collected using a PILATUS 6M detector (Dectris) on the BL13 XALOC beamline at the ALBA synchrotron, Barcelona, Spain. Crystals were rotated with an oscillation angle of 0.5° and an

Figure 2
Crystals of Gal-3(N VII–IX). The average crystal size is $0.05 \times 0.05 \times 0.03$ mm.

exposure time of 0.5 s per image. The data were processed using *XDS* (Kabsch, 2010) and were integrated and scaled using *AIMLESS* (Evans, 2011; Evans & Murshudov, 2013). A summary of the data-collection and processing statistics is given in Table 1.

3. Results and discussion

Extensive broad-scale screening for conditions for obtaining crystals of Gal-3(N VII–IX) (Fig. 1) led to positive results with two different solutions, *i.e.* conditions B3 and C1 of JBScreen Classic 5. They both contained PEG 8000 as the main precipitant. In comparison, larger crystals were obtained with condition B3 [18% (w/v) PEG 8000, 100 mM Tris–HCl pH 8.5, 200 mM lithium sulfate]. These crystals are shown in Fig. 2. Although much effort has been put into improving the diffraction quality of the crystals by designing optimized conditions and applying seeding and counter-diffusion techniques, no better crystals have been obtained thus far. Data analysis revealed that the crystals belonged to the orthorhombic space group $P2_12_12_1$, as indicated by systematic absences. Data processing showed signs of radiation damage after exposure ($\varphi = 200^\circ$) and thus the resolution was limited to 3.3 Å. Appreciable signal [$I/\sigma(I) \simeq 10.1$ (2.8)] could be observed with 99.8% completeness and an R_{merge} of 0.169. A diffraction pattern is shown in Fig. 3.

A distinctive characteristic of the data is that the Matthews coefficient (Matthews, 1968) is indicative of the presence of between six ($V_M = 4.28 \text{ Å}^3 \text{ Da}^{-1}$, solvent content 71.04%) and 12 ($V_M = 2.14 \text{ Å}^3 \text{ Da}^{-1}$, solvent content 42.08%) monomers in the asymmetric unit. Self-rotation analysis showed peaks corresponding to one fourfold axis (parallel to the z axis) and twofold axes (in the xy plane) (Fig. 4), in line with the presence

Figure 3
X-ray diffraction image ($\varphi = 60^\circ$) of Gal-3(N VII–IX) with a resolution ring at 3.2 Å.

Figure 4

Self-rotation function of Gal-3(N VII–IX) crystals. (a) A self-rotation search with $\kappa = 90^\circ$ was used to identify the noncrystallographic fourfold axis. (b) $\kappa = 180^\circ$ section. Calculations were performed using *MOLREP* with an integration radius of 30 Å and data between 15 and 4 Å resolution (the crystallographic *a* axis is along $\varphi = 0^\circ$, $\psi = 90^\circ$ and the *c** axis is perpendicular to the plane of the figure).

of eight to 12 molecules in the asymmetric unit. Structure solution by molecular replacement (Vagin & Teplyakov, 2010; McCoy *et al.*, 2007), using PDB entry 1a3k (Seetharaman *et al.*, 1998) as a search model, is currently in progress. Whether the aggregation is owing to contacts *via* the CRD, as recently observed for the chicken prototype galectin CG-2 (Ruiz *et al.*, 2013), or the non-CRD part (or a combination of these) is an open question. Having succeeded in crystallizing a Gal-3 variant with a tail containing both types of module, although shortened, will provide the answer to this question. Equally important from this perspective, the identified conditions for crystal formation may now help to obtain, step by step, structural information on engineered variants with extended tail lengths, with the aim of completing the study with physiological (full-length) Gal-3.

Acknowledgements

We cordially thank the staff of the XALOC beamline at the ALBA synchrotron for their generous support. This work was funded by grants BFU2011-24615 and CSD2009-00088 from the Spanish Ministry of Science and Innovation, the Regional Government of Madrid (S2010/BMD-2353) and the European Union's Seventh Framework Program FP7/2007-2013 under REA grant agreement No. 317297 ('GLYCOPHARM'). We also thank Drs B. Friday and H. Kaltner for inspiring discussions.

References

- Ahmad, N., Gabius, H.-J., André, S., Kaltner, H., Sabesan, S., Roy, R., Liu, B., Macaluso, F. & Brewer, C. F. (2004). *J. Biol. Chem.* **279**, 10841–10847.
- Amano, M., Eriksson, H., Manning, J. C., Detjen, K. M., André, S., Nishimura, S.-I., Lehtiö, J. & Gabius, H.-J. (2012). *FEBS J.* **279**, 4062–4080.
- André, S. *et al.* (2007). *FEBS J.* **274**, 3233–3256.
- Berbis, M. A., André, S., Cañada, F. J., Pipkorn, R., Ippel, H., Mayo, K. H., Kübler, D., Gabius, H.-J. & Jiménez-Barbero, J. (2014). *Biochem. Biophys. Res. Commun.* **443**, 126–131.
- Cooper, D. N. W. (2002). *Biochim. Biophys. Acta*, **1572**, 209–231.
- Dawson, H., André, S., Karamitopoulou, E., Zlobec, I. & Gabius, H.-J. (2013). *Anticancer Res.* **33**, 3053–3059.
- Evans, P. R. (2011). *Acta Cryst. D* **67**, 282–292.
- Evans, P. R. & Murshudov, G. N. (2013). *Acta Cryst. D* **69**, 1204–1214.
- Gabius, H.-J., André, S., Jiménez-Barbero, J., Romero, A. & Solís, D. (2011). *Trends Biochem. Sci.* **36**, 298–313.
- Göhler, A., André, S., Kaltner, H., Sauer, M., Gabius, H.-J. & Dose, S. (2010). *Biophys. J.* **98**, 3044–3053.
- Kabsch, W. (2010). *Acta Cryst. D* **66**, 125–132.
- Kaltner, H. & Gabius, H.-J. (2012). *Histol. Histopathol.* **27**, 397–416.
- Kasai, K.-i. & Hirabayashi, J. (1996). *J. Biochem.* **119**, 1–8.
- Kopitz, J., André, S., von Reitzenstein, C., Versluis, K., Kaltner, H., Pieters, R. J., Wasano, K., Kuwabara, I., Liu, F.-T., Cantz, M., Heck, A. J. R. & Gabius, H.-J. (2003). *Oncogene*, **22**, 6277–6288.
- Kopitz, J., Vértess, S., André, S., Fiedler, S., Schnölzer, M. & Gabius, H.-J. (2014). *Biochimie*, **104**, 90–99.
- Kopitz, J., von Reitzenstein, C., André, S., Kaltner, H., Uhl, J., Ehemann, V., Cantz, M. & Gabius, H.-J. (2001). *J. Biol. Chem.* **276**, 35917–35923.
- Matthews, B. W. (1968). *J. Mol. Biol.* **33**, 491–497.
- McCoy, A. J., Grosse-Kunstleve, R. W., Adams, P. D., Winn, M. D., Storoni, L. C. & Read, R. J. (2007). *J. Appl. Cryst.* **40**, 658–674.
- Ochieng, J., Platt, D., Tait, L., Hogan, V., Raz, T., Carmi, P. & Raz, A. (1993). *Biochemistry*, **27**, 4455–4460.
- Quattroni, P., Li, Y., Lucchesi, D., Lucas, S., Hood, D. W., Herrmann, M., Gabius, H.-J., Tang, C. M. & Exley, R. M. (2012). *Cell. Microbiol.* **14**, 1657–1675.

- Ruiz, F. M., Fernández, I. S., López-Merino, L., Lagartera, L., Kaltner, H., Menéndez, M., André, S., Solís, D., Gabius, H.-J. & Romero, A. (2013). *Acta Cryst.* **D69**, 1665–1676.
- Ruiz, F., Scholz, B. A., Buzamet, E., Kopitz, J., André, S., Menéndez, M., Romero, A., Solís, D. & Gabius, H.-J. (2014). *FEBS J.* **281**, 1446–1464.
- Seetharaman, J., Kanigsberg, A., Slaaby, R., Leffler, H., Barondes, S. H. & Rini, J. M. (1998). *J. Biol. Chem.* **273**, 13047–13052.
- Vagin, A. & Teplyakov, A. (2010). *Acta Cryst.* **D66**, 22–25.