Acta Crystallographica Section F
Structural Biology and Crystallization Communications
ISSN 1744-3091

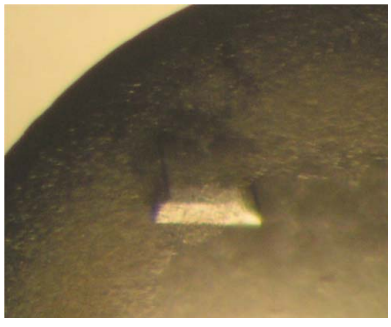
Expression, purification, crystallization and preliminary X-ray characterization of the GRP carbohydrate-recognition domain from *Homo sapiens*

Dongwen Zhou,[a,b] Jianping Sun,[a,b] Wei Zhao,[a,b] Xiao Zhang,[a,b] Yunyu Shi,[a,b] Maikun Teng,[a,b]* Liwen Niu,[a,b]* Yuhui Dong[c] and Peng Liu[c]

[a]Hefei National Laboratory for Physical Sciences at Microscale and School of Life Sciences, University of Science and Technology of China, 96 Jinzhai Road, Hefei, Anhui 230027, People's Republic of China, [b]Key Laboratory of Structural Biology, Chinese Academy of Sciences, 96 Jinzhai Road, Hefei, Anhui 230027, People's Republic of China, and [c]Beijing Synchrotron Radiation Facility, Institute of High Energy Physics, Chinese Academy of Sciences, 19B Yuquan Road, Beijing 100039, People's Republic of China

Correspondence e-mail: mkteng@ustc.edu.cn, lwniu@ustc.edu.cn

Received 7 February 2006
Accepted 9 April 2006



Galectins are a family of animal lectins which share similar carbohydrate-recognition domains (CRDs) and an affinity for β-galactosides. A novel human galectin-related protein named GRP (galectin-related protein; previously known as HSPC159) comprises only one conserved CRD with 38 additional N-terminal residues. The C-terminal fragment of human GRP (GRP-C; residues 38–172) containing the CRD has been expressed and purified. The protein was crystallized using the hanging-drop vapour-diffusion method from a solution containing 2% PEG 400 and 2*M* ammonium sulfate in 100 m*M* Tris–HCl buffer pH 7.5. Diffraction data were collected to a resolution limit of 2.0 Å at beamline 3W1A of Beijing Synchrotron Radiation Facility at 100 K. The crystals belong to the monoclinic space group *C*2, with unit-cell parameters $a = 123.07$, $b = 96.67$, $c = 61.56$ Å, $\beta = 118.72^\circ$. The estimated Matthews coefficient was 2.6 Å^3 Da^{-1}, corresponding to 51.8% solvent content.

1. Introduction

Galectins are a large family of animal lectins defined by their consensus sequences and their affinity for β-galactoside sugars (Drickamer & Taylor, 1993; Liu, 2000). In evolutionary terms galectins are ancient and they are widely distributed from the lower invertebrates to higher vertebrates (Hirabayashi *et al.*, 1992). About 15 mammalian galectins have been identified to date (Cooper, 2002; Gray *et al.*, 2004). The galectins can be classified into three groups: the prototype, which contains one carbohydrate-recognition domain (CRD), the chimera type, which has a proline/glycine-rich repetitive sequence connected to a CRD, and the tandem-repeat type, which contains two homologous CRDs in tandem separated by a short linker (Barondes *et al.*, 1994; Rabinovich, 1999). Prototype galectins are non-covalent homodimers composed of two identical CRDs, except for galectin-5, which exists as a monomer (Gitt *et al.*, 1995). The only member of the chimera type is galectin-3, which is predominantly found in mammals (Rabinovich, 1999). The tandem-repeat type includes galectin-4, galectin-6, galectin-8, galectin-9 and galectin-12 (Hotta *et al.*, 2001).

Galectins function in diverse important biological processes and the function of a given galectin can even vary from site to site depending on the nature of the available ligands (Drickamer & Taylor, 1993; Perillo *et al.*, 1998). Both galectin-1 and galectin-3 have been demonstrated to activate various cell types and to modulate cell adhesion and to be active *in vitro* in inducing pre-mRNA splicing (Dagher *et al.*, 1995; Vyakarnam *et al.*, 1997). In particular, galectin-1 can either promote or inhibit cell adhesion (Mahanthappa *et al.*, 1994;). Furthermore, both galectin-1 and galectin-9 can induce apoptosis of T-cells and thymocytes (Perillo *et al.*, 1995, 1997; Wada *et al.*, 1997). Galectin-3 also plays a role in multiple biological processes through interaction with specific ligands (Barondes *et al.*, 1994). For example, galectin-3 has been found to inhibit apoptosis through interaction with Bcl-2 family members (Hsu *et al.*, 2000) and to trigger activation of mast cells and basophils and play a role in inflammation through interaction with both IgE and the IgE receptor (Frigeri *et al.*, 1993; Liu, 1993).

Here, we describe the expression and preliminary X-ray characterization of a novel protein closely related to galectins, designated GRP (galectin-related protein; previously known as HSPC159). The

Table 1
Data-collection and reduction statistics.

Values in parentheses are for the last shell.

Space group	*C*2
Unit-cell parameters (Å, °)	a = 123.07, b = 96.67, c = 61.56, β = 118.72
Molecules per ASU	4
V_M (Å^3 Da^{-1})	2.6
Resolution limits (Å)	20–2.00 (2.07–2.00)
$I/\sigma(I)$	4.7 (1.8)
Observations	153674
Independent reflections	41481
R_{merge}† (%)	6.25 (22.13)
Completeness‡ (%)	97.6 (92.9)

† $R_{merge} = \sum_h \sum_j |I(h)_j - \langle I(h)\rangle| / \sum_h \sum_j I(h)_j$, where $I(h)_j$ is the observed intensity of reflection and $\langle I(h)\rangle$ is the mean intensity of reflection h. ‡ Completeness is the ratio of the number of reflections to that of possible reflections.

GRP sequence is highly conserved among different species. It shares consensus amino acids at 51 of the 64 most highly conserved residues in other galectins. On the other hand, GRP deviates significantly at five of the seven most critical residues for carbohydrate binding (Cooper, 2002). The molecular characteristics and biological functions of GRP have not been reported to date.

2. Materials and methods

2.1. Subcloning and expression

The GRP-C gene was amplified by PCR using the human brain cDNA library as a template. After amplification of the target gene by sense 5′-CATATGGTTCCATTTTGTGGGCACAT-3′ and antisense 5′-GCGGCCGCTCAGCCAAGCTTGGTG-3′ primers, the PCR fragment was cloned into pGEM-T vector (Promega) and the product was transfected into *Escherichia coli* DH5α cells and screened for positive clones by α-complementation (Ullmann *et al.*, 1967). The positive products were digested by *Nde*I and *Not*I and inserted into the plasmid vector p28 encoding a polypeptide with an N-terminal hexahistidine tag to facilitate purification. We altered the plasmid vector from pET28a+ (Novagen) to p28. Recombinants were then transfected into *E. coli* BL21(DE3) cells and screened for production of a ∼15 kDa protein to indicate a positive clone. The positive clone was again confirmed by DNA sequencing. Expression was performed in Luria–Bertani medium, which was incubated at 310 K and shaken at 240 rev min^{-1}. When the OD_{600} reached 0.6–0.8, the cell culture was supplied with isopropyl β-D-thiogalactopyranoside (IPTG) to a final concentration of 1 m*M* to induce protein production. The culture was incubated for an additional 12 h at 298 K. Cells from culture were first pelleted and then resuspended in lysis buffer containing 20 m*M* Tris–HCl pH 8.0, 200 m*M* NaCl. They were then lysed by sonication on ice. The supernatant containing soluble protein was finally collected by centrifugation.

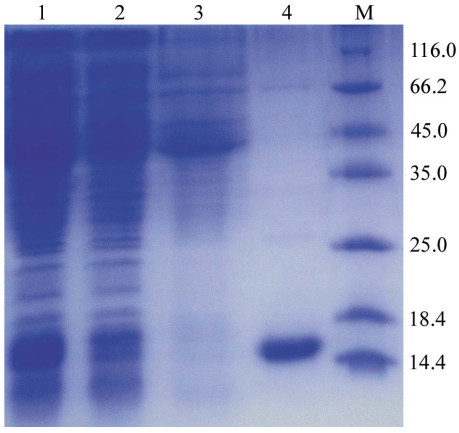


Figure 1
SDS–PAGE of GRP-C. Lane 1, cell lysate supernatant; lane 2, Ni–NTA column flowthrough; lane 3, wash fraction; lane 4, eluate; lane M, molecular-weight markers (kDa). Proteins were resolved by SDS–PAGE on a 12% gel, followed by staining with Coomassie Blue.

2.2. Protein purification

The supernatant containing soluble protein was loaded onto an Ni–NTA column (Novagen) that had been pre-equilibrated with binding buffer (identical to the lysis buffer). Once all unbound proteins had been washed from the column using washing buffer (20 m*M* Tris–HCl pH 8.0, 200 m*M* NaCl, 30 m*M* imidazole), GRP-C protein was eluted from the column using elution buffer (20 m*M* Tris–HCl pH 8.0, 200 m*M* NaCl, 200 m*M* imidazole). The protein was ultrafiltered to remove imidazole and salt and concentrated to a final concentration of 10 mg ml^{-1}. The concentration of the protein was estimated using BCA Protein Assay Kit (Pierce). Both phenylmethylsulfonyl fluoride (PMSF) and EDTA were used to inhibit the degradation of GRP-C.

2.3. Crystallization

Initial crystallization of GRP-C was carried out by the hanging-drop vapour-diffusion method using Crystal Screens I and II (Hampton Research). The protein eventually crystallized with reservoir solution containing 2% PEG 400 and 2 *M* ammonium sulfate in 100 m*M* Tris–HCl buffer pH 7.5; drops contained 1 µl precipitant solution mixed with 1 µl protein solution. The droplet was equilibrated against 400 µl reservoir solution and crystals grew at 277 K.

2.4. Collection and reduction of X-ray diffraction data

X-ray diffraction data were collected at 100 K on beamline 3W1A of the Beijing Synchrotron Radiation Facility at the Institute of High Energy Physics, Chinese Academy of Sciences. A single crystal was transferred to a cryoprotectant composed of reservoir solution with 25% glycerol and immediately flash-frozen in liquid nitrogen before being mounted on the goniometer. The diffraction data processing was performed using *AUTOMAR* v.1.4 (MAR Research GmbH). The data-collection and processing statistics are listed in Table 1.

3. Results and discussion

By sequence *BLAST*, GRP shows a comparatively low overall identity with other galectins; for example, 28% with human galectin-3 and 29% with human galectin-7. On the other hand, it is notable that GRP may differ from other galectins in carbohydrate-binding activity (Cooper, 2002). GRP probably has some novel biological functions. Further studies need to be carried out in the future.

Intact GRP with a C-terminal hexahistidine tag was initially expressed as inclusion bodies. The GRP gene was then cloned into a plasmid vector with an N-terminal hexahistidine tag and was expressed as a soluble protein in very small quantities at low temperature. By exploration, we found that GRP-C with an N-terminal hexahistidine tag was highly expressed as a soluble protein at room temperature. Thus, we conjecture that the coil formed by the N-terminal 38 residues may impede the correct folding of GRP in *E. coli* BL21(DE3) cells. The pure protein pooled from the Ni–NTA

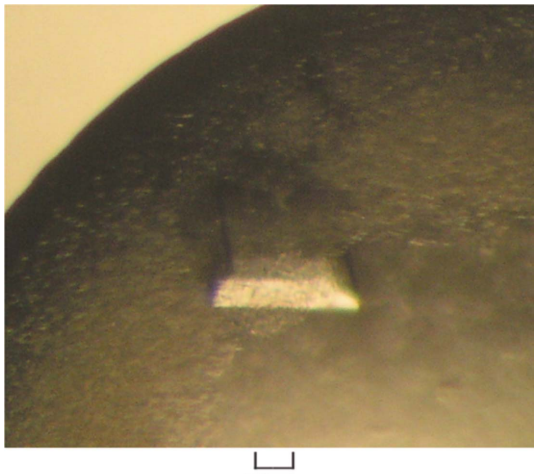


Figure 2
Photomicrography of a crystal of GRP-C.

column degrades very rapidly and 2 m*M* PMSF and 5 m*M* EDTA can inhibit the degradation satisfactorily. SDS–PAGE and dynamic light-scattering assays showed that GRP-C exists as a monomer in solution. The protein molecular weight and purity were estimated at about 15 kDa and >90%, respectively, using SDS–PAGE (Fig. 1). The crystals of GRP-C grew to typical dimensions of about 0.8 × 0.6 × 0.2 mm after a month at 277 K (Fig. 2). The crystals diffracted X-rays to a resolution limit of 2.0 Å and belonged to the monoclinic space group *C*2. The solvent content of the crystals has been estimated at 51.8% and the V_M ratio as 2.6 Å^3 Da^{-1} (Matthews, 1968). Structural determination of GRP-C is in progress.

Financial support for this project to LN and MT was provided by research grants from Chinese National Natural Science Foundation (grant Nos. 30121001, 30025012, 30130080 and 30571066), the '973' and '863' Plans of the Chinese Ministry of Science and Technology (grant Nos. G1999075603, 2004CB520801 and 2002BA711A13) and the Chinese Academy of Sciences (grant No. KSCX1-SW-17).

References

Barondes, S. H., Cooper, D. N. W., Gitt, M. A. & Leffler, H. (1994). *J. Biol. Chem.* **269**, 20807–20810.
Cooper, D. (2002). *Biochim. Biophys. Acta*, **1572**, 209–231.
Dagher, S. F., Wang, J. L. & Patterson, R. J. (1995). *Proc. Natl Acad. Sci. USA*, **92**, 1213–1217.
Drickamer, K. & Taylor, M. E. (1993). *Annu. Rev. Cell Biol.* **9**, 237–264.
Frigeri, L. G., Zuberi, R. I. & Liu, F. T. (1993). *Biochemistry*, **32**, 7644–7649.
Gitt, M. A., Wiser, M. F., Leffler, H., Herrmann, J., Xia, Y. R., Massa, S. M., Cooper, D., Lusis, A. & Barondes, S. H. (1995). *J. Biol. Chem.* **270**, 5032–5038.
Gray, C. A., Adelson, D. L., Bazer, F. W., Burghardt, R. C., Meeusen, E. N. & Spencer, T. E. (2004). *Proc. Natl Acad. Sci. USA*, **101**, 7982–7987.
Hirabayashi, J., Satoh, M. & Kasai, K. (1992). *J. Biol. Chem.* **267**, 15485–15490.
Hotta, K., Funahashi, T., Matsukawa, Y., Takahashi, M., Nishizawa, H., Kishida, K., Matsuda, M., Kuriyama, H., Kihara, S., Nakamura, T., Tochino, Y., Bodkin, N. L., Hansen, B. C. & Matsuzawa, Y. (2001). *J. Biol. Chem.* **276**, 34089–34097.
Hsu, D. K., Yang, R.-Y., Yu, L., Pan, Z., Salomon, D. R., Fung-Leung, W.-P. & Liu, F.-T. (2000). *Am. J. Pathol.* **156**, 1073–1083.
Liu, F.-T. (1993). *Immunol. Today*, **14**, 486–490.
Liu, F.-T. (2000). *Clin. Immunol.* **97**, 79–88.
Mahanthappa, N. K., Cooper, D. N., Barondes, S. H. & Schwarting, G. A. (1994). *Development*, **120**, 1373–1384.
Matthews, B. W. (1968). *J. Mol. Biol.* **33**, 491–497.
Perillo, N. L., Marcus, M. E. & Baum, L. G. (1998). *J. Mol. Med.* **76**, 402–412.
Perillo, N. L., Pace, K. E., Seilhamer, J. J. & Baum, L. G. (1995). *Nature (London)*, **378**, 736–739.
Perillo, N. L., Uittenbogaart, C. H., Nguyen, J. T. & Baum, L. G. (1997). *J. Exp. Med.* **185**, 1851–1858.
Rabinovich, G. A. (1999). *Cell Death Differ.* **6**, 711–721.
Ullmann, A., Jacob, F. & Monod, J. (1967). *J. Mol. Biol.* **24**, 339–343.
Vyakarnam, A., Dagher, S. F., Wang, J. L. & Patterson, R. J. (1997). *Mol. Cell Biol.* **17**, 4730–4737.
Wada, J., Ota, K., Kumar, A., Wallner, E. I. & Kanwar, Y. S. (1997). *J. Clin. Invest.* **99**, 2452–2461.