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

Expression, purification, crystallization and preliminary X-ray analysis of a truncated soluble domain of human glioma pathogenesis-related protein 1

Nathalie Bonafé,[a] Bin Zhan,[b] Maria Elena Bottazzi,[b] Oriana A. Perez,[a] Raymond A. Koski[a] and Oluwatoyin A. Asojo[c]*

[a]L2 Diagnostics LLC, 300 George Street, New Haven, CT 06511, USA, [b]Department of Microbiology, Immunology, and Tropical Medicine, The George Washington University Medical Center, Washington DC 20037, USA, and [c]Department of Pathology and Microbiology, College of Medicine, Nebraska Medical Center, Omaha, NE 68198, USA

Correspondence e-mail: oasojo@unmc.edu

Received 12 August 2010
Accepted 5 September 2010

Glioma pathogenesis-related protein 1 (GLIPR1) is a member of the CAP superfamily that includes proteins from a wide range of eukaryotic organisms. The biological functions of most CAP proteins, including GLIPR1, are unclear. GLIPR1 is up-regulated in aggressive glioblastomas and contributes to the invasiveness of cultured glioblastoma cells. In contrast, decreased GLIPR1 expression is associated with advanced prostate cancer. Forced GLIPR1 overexpression is pro-apoptotic in prostate cancer cells and is being tested in clinical trials as an experimental prostate-cancer therapy. Human GLIPR1 was expressed as a truncated soluble protein (sGLIPR1), purified and crystallized. Useful X-ray data have been collected to beyond 1.9 Å resolution from a crystal that belonged to the orthorhombic space group $P2_12_12$ with average unit-cell parameters a = 85.1, b = 79.5, c = 38.9 Å and either a monomer or dimer in the asymmetric unit.

1. Introduction

Glioma pathogenesis-related protein 1 (GLIPR1), also known as related to testis-specific, vespid and pathogenesis-related 1 protein (RTVP-1), belongs to the cysteine-rich secretory protein, antigen 5 and pathogenesis-related 1 (CAP) superfamily (Gibbs *et al.*, 2008). CAP proteins are characterized by the CAP domain, historically referred to as the SCP (sperm coating glycoprotein) domain (NCBI domain cd00168 or Pfam domain PF00188), which has been reported in a diversity of proteins that are unrelated by phylogeny and have been isolated from bacteria, plants, animals and viruses (Geer *et al.*, 2002; Gibbs *et al.*, 2008). CAP proteins include the *Ancylostoma* secreted protein secreted by hookworms upon host entry (Hawdon *et al.*, 1996, 1999; Hawdon & Hotez, 1996), PR-1 pathogenesis-related proteins from plants (van Loon *et al.*, 2006) and CRISP proteins expressed in mammalian reproductive tracts, as well as venom allergens from insects and reptiles (Gibbs *et al.*, 2008). The CAP domain is a highly conserved cysteine-rich domain of at least 15 kDa with unknown function. CAP proteins have been implicated in various conditions ranging from plant responses to pathogens to human brain tumor growth (Cantacessi *et al.*, 2009; Gibbs *et al.*, 2008).

While several CAP protein structures have been solved (Gibbs *et al.*, 2008; Asojo *et al.*, 2005; Suzuki *et al.*, 2008), only one, the Golgi-associated PR-1 protein (GAPR-1; PDB code 1smb; Serrano *et al.*, 2004), is a human protein. This protein lacks the characteristic cysteines of the CAP domain (Gibbs *et al.*, 2008). 31 human CAP proteins have been identified, which can be subdivided into nine subfamilies (Gibbs *et al.*, 2008). GLIPR1, a member of the GLIPR1 subfamily, has a signal peptide followed by a CAP domain and a transmembrane domain. The predicted GLIPR1 C-terminal transmembrane domain is unique among mammalian CAP proteins (Gibbs *et al.*, 2008). GLIPR1 is normally expressed in fetal kidney and multiple adult tissues; however, its structure and biochemical functions have yet to be characterized. GLIPR1 mRNA was initially identified as a major up-regulated transcript in glioblastoma multiforme, the most aggressive form of human brain cancer, as well as within glioma cell lines (Gibbs *et al.*, 2008). GLIPR1 also has aberrant expression in prostate cancer, where it is down-regulated compared with normal human prostate (Ren *et al.*, 2004), suggesting that the

role of GLIPR1 in cancer may be dependent on cellular context. When an adenoviral vector overexpressing human *Glipr1* was transfected into prostate, lung and colon cancer cell lines, increased rates of apoptosis were observed (Ren *et al.*, 2002, 2004). Viral delivery of *Glipr1* into a murine model of prostate cancer decreased tumor growth and metastasis (Satoh *et al.*, 2003) and a human gene-transfer approach is now being evaluated in clinical trials for prostate cancer gene therapy (Thompson, 2010; Kadmon *et al.*, 2010). In contrast, an adenoviral vector overexpressing *Glipr1* in glioma cells induced anchorage-independent growth, invasiveness and resistance to apoptosis, while siRNA silencing of GLIPR1 decreased cell proliferation and induced cell apoptosis (Rosenzweig *et al.*, 2006). As part of efforts to facilitate the design of novel cancer therapeutics, as well as to reveal possible GLIPR1 functions, we have expressed, purified and crystallized a recombinant truncated soluble form of GLIPR1 (sGLIPR1).

2. Materials and methods

2.1. Recombinant glycoprotein production

A yeast codon-optimized cDNA sequence encoding the human GLIPR1 protein soluble domain (sGLIPR1; amino-acid residues 22–220; nucleotides 212–808; GenBank accession No. NM_006851) was cloned into a *Pichia pastoris* expression vector from Invitrogen. The expression vector GLIPR1(22–220)/pPICZαA was designed to direct the synthesis and secretion of sGLIPR1 into the culture supernatant. The vector encoded a *Pichia* secretory signal sequence fused to the 5′ end of the *Glipr1* cDNA. After signal sequence processing, EAEAEF was predicted to constitute the N-terminal end of the secreted protein. The recombinant plasmid was linearized and transformed into *P. pastoris* strain X33. Small-scale expression of the transformed colonies was induced with methanol for up to 3 d. Cell supernatants were collected after centrifugation and tested for sGLIPR1 by SDS–PAGE and by Western blots using anti-GLIPR1$_{75–95}$ peptide antibody. Expression was induced in one of the positive clones in a 4 l culture for 72 h with 0.5% methanol. The predicted molecular weight of sGLIPR1 is 23 295 Da after removal of the signal peptide. The sGLIPR1 was purified by SP-Sepharose cation-exchange chromatography (HiTrap SP XL column, GE Healthcare) in 50 m*M* MES buffer pH 6.1 followed by S-200 Superdex (GE Healthcare) size-exclusion chromatography in the same buffer. Analysis of purified sGLIPR1 by electrospray mass spectrometry as well as N-terminal amino-acid sequencing and mass-spectrometric analysis of sGLIPR1 tryptic fragments were performed by the W. M. Keck Foundation Biotechnology Resource Laboratory at Yale University.

2.2. Crystallization

Initial crystallization screens were carried out on samples of sGLIPR1 using sparse-matrix screens from Qiagen including Classic Screens I and II and Cryo Screen. The experiments were carried out at 293 K. Crystals were grown by vapor diffusion in hanging drops that were equilibrated against 1 ml crystallization solution in a NeXtal EasyXtal tool crystallization plate (Qiagen, USA). Drops were prepared by mixing 2 µl protein solution with an equal volume of crystallization solution. All protein solutions used in crystallization experiments consisted of 12 mg ml^{-1} protein in 50 m*M* Tris–HCl pH 7.5. After optimization of the various hit conditions, the best crystals were obtained in 48 h by vapor diffusion from hanging drops consisting of a mixture of 2 µl recombinant sGLIPR1 solution and 1.5 µl precipitant solution [0.17 *M* ammonium sulfate, 0.085 *M* sodium cacodylate pH 6.5, 25.5%(*w*/*v*) PEG 8000 and 15%(*v*/*v*) glycerol]. The largest of these crystals was 0.5 × 0.3 × 0.1 mm in size and typically grew in clusters as illustrated in Fig. 1.

2.3. Diffraction experiments

Since sGLIPR1 crystals grew in solutions that contained adequate cryoprotectant, they were directly flash-cooled in a stream of N_2 gas at 113 K prior to collecting diffraction data. Data sets were collected using an Xcalibur PX Ultra four-circle kappa platform with a 165 mm diagonal Onyx CCD detector and a high-brilliance sealed-tube Enhance Ultra (Cu) X-ray source (Oxford Diffraction, Oxford, England) operating at 50 kV and 40 mA. Data collection was performed using the *CrysAlis Pro* software to select the best orientations of the crystals, keeping the longest axis fixed to maximize spot separation. A diffraction image showing the crystal quality is illustrated in Fig. 2. Complete data sets for each crystal were collected

Figure 1
The largest sGLIPR1 crystal is 0.5 mm on the longest side and 0.1 mm on the smallest side and was separated with cryotools prior to cryocooling in a stream of liquid N_2 for data collection.

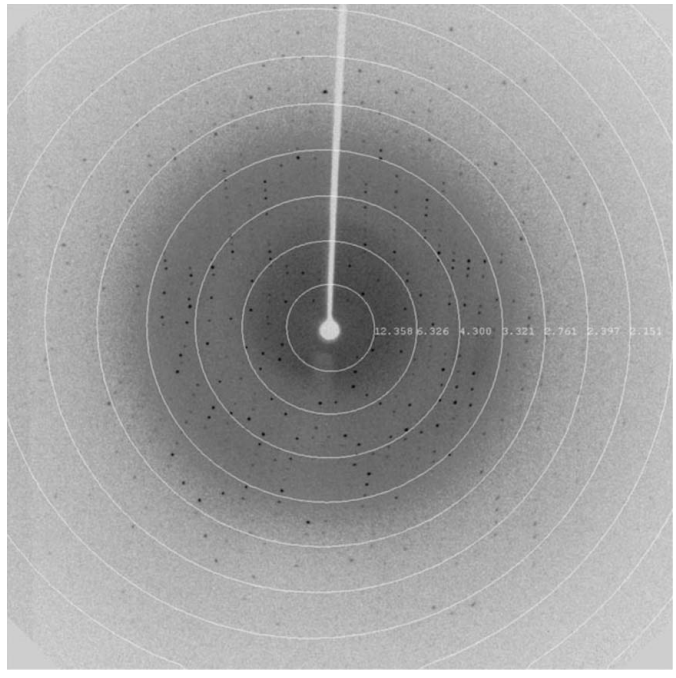

Figure 2
The crystal has visible diffraction spots beyond 2.0 Å resolution and the diffraction pattern indicates that the crystal is of a protein and not a salt.

Table 1
Data-collection and reduction statistics.

	Overall	Inner shell	Outer shell
Low-resolution limit (Å)	29.04	29.04	1.95
High-resolution limit (Å)	1.85	5.85	1.85
$R_{\rm merge}$†	0.093	0.033	0.496
Total No. of observations	279138	10465	16237
Total No. of unique reflections	23229	830	3305
$\langle I/\sigma(I)\rangle$	20.1	44.1	3.2
Completeness (%)	99.9	99.2	99.4
Multiplicity	12.0	12.6	4.9

† $R_{\rm merge} = \sum_{hkl}\sum_i |I_i(hkl) - \langle I(hkl)\rangle| / \sum_{hkl}\sum_i I_i(hkl)$, where $I_i(hkl)$ and $\langle I(hkl)\rangle$ are the intensity of measurement of I and the mean intensity of the reflection with indices hkl, respectively.

using a crystal-to-detector distance of 65 mm and an exposure time of 120 s per 0.8° oscillation. X-ray data sets were processed using the program *CrysAlis Pro* (Oxford Diffraction). Crystallographic data statistics are given in Table 1. We expect either one or two monomers in the asymmetric unit (Matthews, 1968; Kantardjieff & Rupp, 2003). The presence of a monomer in the asymmetric unit is based on the volume of the unit cell being 26 992 Å^3 and would correspond to a Matthews coefficient of 3.11 Å^3 Da^{-1} (60% solvent), while the presence of a dimer in the asymmetric unit would correspond to a Matthews coefficient of 1.56 Å^3 Da^{-1} (21% solvent).

3. Results and discussion

CAP-superfamily proteins contain predicted signal peptides, which is consistent with their extracellular localization or their localization to specific intracellular compartments. We chose a *P. pastoris* eukaryotic expression vector to direct the synthesis and secretion of the soluble domain of human GLIPR1 into the culture supernatant. Passage through the secretory pathway of *P. pastoris* cells promoted post-translational folding and modification of sGLIPR1. Purification of sGLIPR1 from culture medium yielded more than 100 mg per litre at over 95% purity. LC MS/MS analysis of sGLIPR1 positively identified eight peptides, which were sufficient to establish the identity of the recombinant protein. The N-terminal amino-acid sequence was EAEAEF, confirming that the protein was properly processed. Electrospray mass-spectrometric analysis identified a predominant species with a mass of 24 449 kDa, which is 1154 Da greater than the predicted molecular weight of sGLIPR1 without post-translational modification. Deconvulated data suggested glycosyl modifications by showing a series of peaks of molecular weight ranging between 24 200 and 26 800 and separated by 162 Da. Glycosylation of sGLIPR1 was confirmed by enzymatic digestion with PNGase F and Endo H. These results are in agreement with the single predicted N-linked glycosylation site at residue Asn71 in the mature protein and did not hinder crystallization. If glycosylation and heterogeneity prove to be a problem during structure determination, we will deglycosylate the sample and repeat the crystallization studies.

Sample crystals permitted the collection of data to beyond 1.9 Å resolution in the orthorhombic space group $P2_12_12$, with average unit-cell parameters a = 85.1, b = 79.5, c = 38.9 Å. Efforts are under way to solve the structure by molecular replacement using reptilian CRISPs, which share about 35% sequence identity with sGLIPR. Solving the first structure of the first member of this CAP-protein subfamily will help to illuminate their unique structural features and provide new insights into their functions.

This work was supported in part by National Cancer Institute R43CA123991 SBIR grant to NB. OAA is supported by National Cancer Institute Mentored Career development grant for under-represented minorities K01CA113486.

References

Asojo, O. A., Goud, G., Dhar, K., Loukas, A., Zhan, B., Deumic, V., Liu, S., Borgstahl, G. E. & Hotez, P. J. (2005). *J. Mol. Biol.* **346**, 801–814.
Cantacessi, C., Campbell, B. E., Visser, A., Geldhof, P., Nolan, M. J., Nisbet, A. J., Matthews, J. B., Loukas, A., Hofmann, A., Otranto, D., Sternberg, P. W. & Gasser, R. B. (2009). *Biotechnol. Adv.* **27**, 376–388.
Geer, L. Y., Domrachev, M., Lipman, D. J. & Bryant, S. H. (2002). *Genome Res.* **12**, 1619–1623.
Gibbs, G. M., Roelants, K. & O'Bryan, M. K. (2008). *Endocr. Rev.* **29**, 865–897.
Hawdon, J. M. & Hotez, P. J. (1996). *Curr. Opin. Genet. Dev.* **6**, 618–623.
Hawdon, J. M., Jones, B. F., Hoffman, D. R. & Hotez, P. J. (1996). *J. Biol. Chem.* **271**, 6672–6678.
Hawdon, J. M., Narasimhan, S. & Hotez, P. J. (1999). *Mol. Biochem. Parasitol.* **99**, 149–165.
Kadmon, D., Sonpavde, G., Jain, R. K., Ayala, G. E., Ittmann, M. M., Kurosaka, S., Edamura, K., Tabata, K., Milesand, B. J. & Thompson, T. C. (2010). In the press.
Kantardjieff, K. A. & Rupp, B. (2003). *Protein Sci.* **12**, 1865–1871.
van Loon, L. C., Rep, M. & Pieterse, C. M. (2006). *Annu. Rev. Phytopathol.* **44**, 135–162.
Matthews, B. W. (1968). *J. Mol. Biol.* **33**, 491–497.
Ren, C., Li, L., Goltsov, A. A., Timme, T. L., Tahir, S. A., Wang, J., Garza, L., Chinault, A. C. & Thompson, T. C. (2002). *Mol. Cell. Biol.* **22**, 3345–3357.
Ren, C., Li, L., Yang, G., Timme, T. L., Goltsov, A., Ji, X., Addai, J., Luo, H., Ittmann, M. M. & Thompson, T. C. (2004). *Cancer Res.* **64**, 969–976.
Rosenzweig, T., Ziv-Av, A., Xiang, C., Lu, W., Cazacu, S., Taler, D., Miller, C. G., Reich, R., Shoshan, Y., Anikster, Y., Kazimirsky, G., Sarid, R. & Brodie, C. (2006). *Cancer Res.* **66**, 4139–4148.
Satoh, T., Timme, T. L., Saika, T., Ebara, S., Yang, G., Wang, J., Ren, C., Kusaka, N., Mouraviev, V. & Thompson, T. C. (2003). *Hum. Gene Ther.* **14**, 91–101.
Serrano, R. L., Kuhn, A., Hendricks, A., Helms, J. B., Sinning, I. & Groves, M. R. (2004). *J. Mol. Biol.* **339**, 173–183.
Suzuki, N., Yamazaki, Y., Brown, R. L., Fujimoto, Z., Morita, T. & Mizuno, H. (2008). *Acta Cryst.* D**64**, 1034–1042.
Thompson, T. C. (2010). *Yonsei Med. J.* **51**, 479–483.