Acta Crystallographica Section D
Biological Crystallography
ISSN 0907-4449

Crystallization of a Golgi-associated PR-1-related protein (GAPR-1) that localizes to lipid-enriched microdomains

Matthew R. Groves,[a,b]*‡ Audrey Kuhn,[a] Astrid Hendricks,[b] Susanne Radke,[b] Ramon L. Serrano,[b]§ J. Bernd Helms[b]¶ and Irmgard Sinning[a,b]

[a]Structural Biology Program, EMBL, Meyerhofstrasse 1, 69117 Heidelberg, Germany, and [b]BZH, Im Neuenheimer Feld, 69120 Heidelberg, Germany

‡ Current address: EMBL Hamburg, Building 25A, DESY, Notkestrasse 85, 22603 Hamburg, Germany.
§ Current address: Stanford University Medical School, Department of Biochemistry B457, Stanford, California 94305-5307, USA.
¶ Current address: Department of Biochemistry and Cell Biology, Utrecht University, PO Box 80176, 3508 TD Utrecht, The Netherlands.

Correspondence e-mail: matthew.groves@embl-hamburg.de

Received 15 December 2003
Accepted 27 January 2004

The human Golgi-associated PR-1-related protein (GAPR-1) is closely related to plant pathogenesis-related (PR-1) proteins, which are upregulated in response to pathogen attack. Family members have been identified in a variety of organisms, together constituting the superfamily of PR-1 proteins. GAPR-1 is found within lipid-enriched microdomains on the cytosolic side of the endomembrane system. GAPR-1 is tightly anchored to membranes and absent from the cytosol, although it does not possess a membrane-spanning domain. Crystals of recombinantly expressed GAPR-1 have been grown that diffract to high (1.5 Å) resolution. Complete data sets have been collected on a trigonal crystal form ($P3_121/P3_221$), with unit-cell parameters $a = b = 73.5$, $c = 63.2$ Å. Molecular replacement using the NMR coordinates of tomato pathogenesis-related protein (28% identity) was unsuccessful and a search for heavy-metal derivatives or alternative phasing methods has been initiated.

1. Introduction

Human Golgi-associated PR-1 protein (GAPR-1) is a 17 kDa protein found within lipid-enriched microdomains of the Golgi complex (Eberle *et al.*, 2002). These microdomains have been shown to contain stochiometric amounts of only ten proteins, namely subunits of heterotrimeric G proteins, flotillin-1, caveolin-1, subunits of the vacuolar ATPase, GAPR-1 and an unknown protein (Gkantiragas *et al.*, 2001). In these microdomains, GAPR-1 is part of a core complex that is composed of the B subunit of vacuolar ATPase, flotillin-1, caveolin-1 and GAPR-1 (Gkantiragas *et al.*, 2001). The stability of this core complex has been shown to be mediated through protein–protein interactions rather than a lipid scaffold (Gkantiragas *et al.*, 2001). These protein–protein interactions may be involved in stable membrane-anchoring of GAPR-1.

Members of the PR-1 superfamily were originally identified through their induction in plants following infection with necrotizing viruses (Gianiniazzi *et al.*, 1970; van Loon & van Kammen, 1970). Since then, a number of secretory proteins with significant sequence homology have been identified, including proteins from the fruiting bodies of fungi (Schuren *et al.*, 1993), insect allergens (Lu *et al.*, 1993), mammalian CRISP proteins (Kjeldsen *et al.*, 1996; Kratzschmar *et al.*, 1996), human GliPR/RTVP-1 (Murphy *et al.*, 1995; Rich *et al.*, 1996) and proteins in snake or lizard venoms (Morrissette *et al.*, 1995; Brown *et al.*, 1999).

The physiological function(s) of members of the PR-1 superfamily remain to be established. The NMR solution structure of p14a (Fernandez *et al.*, 1997), a PR-1 family member from tomato sharing 28% identity with GAPR-1, allowed the determination of the arrangement of two conserved histidine/glutamate pairs, which have been suggested to be part of the active centre. Although the arrangement of these conserved residues was reminiscent of Zn^{2+} proteases, no metal ions have been found to be associated with PR-1 proteins. The structure of an additional member of the PR-1 superfamily has been elucidated by crystallography, *viz.* the wasp allergen Ves v 5. While Ves v 5 shares a similar fold to p14a, the conserved histidine/glutamate pairs have a significantly different arrangement (Henriksen *et al.*, 2001).

More recently, another PR-1 family member, Tex31, was identified by Milne *et al.* (2003), who proposed that Tex31 might have serine protease-like activity against a conotoxin precursor. Antifungal activity of PR-1 proteins in plants has been reported (Niderman *et al.*, 1995) and PR-1 proteins have been shown to play a central role in plant immune response during systemic acquired resistance, which is similar to the innate immune system of animals (Kitajima & Sato, 1999; Zhang *et al.*, 1999; Klessig *et al.*, 2000; Maleck *et al.*, 2000).

A role in the immune system of the PR-1 superfamily is also suggested by the identification of a human glioma pathenogenesis-related protein, GliPR (20% identity). GliPR

has been found in all glioma cell lines and tumours studied and high levels of GliPR expression can be induced with phorbol ester in macrophages (Murphy *et al.*, 1995), which are active at the front line of the human immune system. GAPR-1 was also found to be expressed in a tissue-specific manner, being highly expressed in immunocompetent cells and organs, consistent with a possible role in the innate immune system. Thus, the PR superfamily of proteins may represent a common ground in immune response to pathogens in both plants and animals.

A detailed biochemical analysis of GAPR-1 indicated that it is unique amongst PR-1 proteins in two respects. Firstly, it is N-myristolated *in vivo* and found only in lipid-enriched microdomains of the Golgi complex (Eberle *et al.*, 2002). Secondly, GAPR-1 does not possess a signal sequence and is not sorted into secretory granules or vacuoles, in contrast to all other members of the PR-1 family (Eberle *et al.*, 2002).

We report here the crystallization of human GAPR-1. Complete data sets have been collected both in-house to 2.1 Å resolution and at ESRF beamline ID14-EH2 to 1.5 Å resolution.

2. Materials and methods

2.1. Expression, purification and crystallization

The ORF of GAPR-1 was obtained from EST44799 (Genbank accession No. aa339686), cloned into pBluescript SK$^-$ (obtained from ATCC, Rockville, USA; Eberle *et al.*, 2002) and subsequently the full-length mature sequence was cloned into pQE60 vector using *Nco*I and *Bgl*II restriction sites. GAPR-1 was overexpressed in *Escherichia coli* (M15Rep4) by induction with 1 m*M* IPTG for 4 h at 310 K. The soluble fraction of the cell lysate was fractionated on DEAE Sepharose. The flow-through was loaded onto a Macroprep High S column and the bound GAPR-1 was eluted from the column by an NaCl gradient. Subsequent gel filtration on a Sephadex G-200 column resulted in GAPR-1 that was purified to apparent homogeneity. The expression products were checked by dynamic light scattering, mass spectrometry and equilibrium centrifugation for homogeneity and oligomerization state prior to crystallization. The protein was exchanged into a buffer containing 10 m*M* Tris–HCl pH 7.5, 50 m*M* NaCl and 2 m*M* DTT and then concentrated using Amicon Centricon 10 kDa filters to a concentration of 12.6 mg ml^{-1} as estimated by Bradford reagent (Biorad). Initial screening was performed at room temperature (295.5 K) by hanging-drop vapour diffusion (McPherson, 1976) using sparse-matrix kits from Hampton Research (Jancarik & Kim, 1991; Hampton Research, Laguna Niguel, CA, USA). Drops were prepared on siliconized cover slips and equilibrated against 1 ml of reservoir solution. Screens were prepared by combining equal volumes of reservoir solution and protein solution (2 µl). Conditions producing crystals from the initial screens were refined to produce crystals suitable for X-ray diffraction analysis.

2.2. Data collection

Crystals were transferred to a cryoprotectant solution [reservoir solution with an additional 15%(*v*/*v*) glycerol] and then picked up using a fibre loop and flash-cooled in a stream of nitrogen gas at 100 K. In-house diffraction data were collected from a single crystal on a MAR345 (MAR Research) image-plate detector using Cu $K\alpha$ radiation from a rotating-anode X-ray generator operating at 50 kV and 100 mA. A higher resolution data set was collected at the ID14 EH-2 ESRF beamline on a MAR CCD detector. The programs *MOSFLM* (Leslie, 1992), *SCALA* (Evans, 1997), *AMoRe* (Navaza, 1994) and *CNS* (Brünger *et al.*, 1998) were used for data processing and analysis.

3. Results and discussion

Two conditions from the initial screen produced microcrystals from which further screens were designed. Lens-shaped crystals large enough for X-ray analysis were obtained in 20%(*v*/*v*) PEG 8K, 100 m*M* bis-Tris pH 7.0 and 200 m*M* magnesium acetate. These crystals appeared after 3–4 d and were suitable for analysis after 6–7 d (0.4 × 0.2 × 0.2 mm). Diffraction from these crystals was measured in-house to 3.0 Å. Hexagonal crystals were obtained in 30% PEG 4K, 100 m*M* Tris pH 8.0 and 140 m*M* magnesium chloride. These crystals appeared after 7–8 d and were suitable for analysis after 10–12 d (0.4 × 0.4 × 0.4 mm). The hexagonal crystals had unit-cell parameters similar to those of the lens-shaped crystals and shared the same point group, but diffracted in-house to 2.1 Å. A complete data set to 1.5 Å was collected on a MAR CCD detector at beamline ID14 EH-2 at the ESRF (Table 1). A low-resolution pass was made using a crystal-to-detector distance of 200 mm and an oscillation range of 100° in 1° steps. A high-resolution pass was made at a crystal-to-detector distance of 110 mm and an oscillation range of 130° in 0.5° steps.

Table 1
X-ray data-collection statistics of the GAPR-1 crystals on ESRF beamline ID14-EH2.

Values in parentheses correspond to the highest resolution shell (1.56–1.49 Å).

Space group	Trigonal ($P3_121/P3_221$)
Unit-cell parameters	
$a = b$ (Å)	73.5
c (Å)	63.2
$\alpha = \beta$ (°)	90
γ (°)	120
Wavelength (Å)	0.93
Temperature (K)	100
Exposure time (s)	2 (10)
Oscillation range per frame (°)	1 (0.5)
Resolution (Å)	44.1–1.50 (1.56–1.49)
Observed reflections	235802 (23839)
Unique reflections	30262 (3937)
$R_{\rm merge}$ (%)	9.9 (25.1)
Completeness (%)	97.2 (90.4)
$I/\sigma(I)$	3.6 (2.8)
Unique reflections $I > 3\sigma(I)$ (%)	68.3
Multiplicity	7.8 (6.1)

Autoindexing yielded unit-cell parameters $a = b = 73.5$, $c = 63.2$ Å, $\alpha = \beta = 90$, $\gamma = 120°$ in the trigonal point group $P321$. The $00l$ axial reflections are markedly stronger for $l = 3n$ (n integer), indicating that the space group is either $P3_121$ or $P3_221$. Data completeness and $I/\sigma(I)$ were 97.2% and 3.6, respectively, overall and 90.4% and 2.8, respectively, for the 1.56–1.49 Å highest resolution shell. The overall $R_{\rm merge}$ values for the data set were 9.3 and 22.4% for the 1.56–1.49 Å highest-resolution shell. Using a molecular weight of 17 kDa for monomeric GAPR-1, the predicted Matthews coefficient ($V_{\rm M}$) for a single molecule in the asymmetric unit is 2.81 Å^3 Da^{-1} (Matthews, 1968), which is within the usual range for protein crystals, whereas the Matthews coefficient for two molecules in the asymmetric unit is 1.3 Å^3 Da^{-1}. Similarly, the predicted solvent content for a single molecule in the asymmetric unit is 56% and for two molecules the predicted solvent content is 11%, as calculated using the *CCP*4 utility *MATTHEWS_COEFF* (Collaborative Computational Project, Number 4, 1994). Both the dynamic light-scattering experiment and analytical ultracentrifugation indicated that the sample was monodisperse and found in a monomeric solution state. However, gel-filtration data indicate that a small fraction (∼5%) of the protein runs as an apparent dimer.

Poly Ser/Ala/Gly models and the generation of an average structure from the NMR ensemble, as well as individual members of the NMR p14a ensemble, were used as

molecular-replacement search models. However, both the p14a (PDB code 1cfe) and the Ves v 5 (PDB code 1qnx) search models are of low homology (<35% sequence identity) and analysis of the data by a number of molecular-replacement algorithms has failed to yield a clear result to date. Efforts are now under way to identify heavy-metal derivatives of this crystal form.

References

Brown, R. L., Haley, T. L., West, K. A. & Crabb, J. W. (1999). *Proc. Natl Acad. Sci. USA*, **96**, 754–759.

Brünger, A. T., Adams, P. D., Clore, G. M., DeLano, W. L., Gros, P., Grosse-Kunstleve, R. W., Jiang, J.-S., Kuszewski, J., Nilges, M., Pannu, N. S., Read, R. J., Rice, L. M., Simonson, T. & Warren, G. L. (1998). *Acta Cryst.* D**54**, 905–921.

Collaborative Computational Project, Number 4 (1994). *Acta Cryst.* D**50**, 760–763.

Eberle, H. B., Serrano, R. L., Füllekrug, J., Schlosser, A., Lehmann, W. D., Lottspeich, F., Kaloyanova, D., Wieland, F. T. & Helms, J. B. (2002). *J. Cell Sci.* **115**, 827–838.

Evans, P. (1997). *Jnt CCP4/ESF–EACBM Newsl. Protein Crystallogr.* **33**, 22–24.

Fernandez, C., Szyperski, T., Bruyere, T., Ramage, P., Mosinger, E. & Wüthrich, K. (1997). *J. Mol. Biol.* **266**, 576–593.

Gianiniazzi, S., Martin, C. & Vallee, J. C. (1970). *C. R. Hebd. Sceances Acad. Sci., Ser. D*, **270**, 2383–2386.

Gkantiragas, I., Brügger, B., Stüven, E., Kaloyanova, D., Li, X.-Y., Löhr, K., Lottspeich, F., Wieland, F. T. & Helms, B. H. (2001). *Mol. Biol. Cell*, **12**, 1819–1833.

Henriksen, A., King, T. P., Mirza, O., Monsalve, R. I., Meno, K., Ipsen, H., Larsen, J. N., Gajhede, M. & Spangfort, M. D. (2001). *Proteins*, **45**, 438–448.

Jancarik, J. & Kim, S.-H. (1991). *J. Appl. Cryst.* **24**, 409–411.

Kitajima, S. & Sato, F. (1999). *J. Biochem. (Tokyo)*, **125**, 1–8.

Kjeldsen, L., Cowland, J. B., Johnsen, A. H. & Borregaard, N. (1996). *FEBS Lett.* **380**, 246–250.

Klessig, D. F., Durner, J., Noad, R., Navarre, D. A., Wendehenne, D., Kumar, D., Zhou, J. M., Shah, J., Zhang, S., Kachroo, P., Trifa, Y., Pontier, D., Lam, E. & Silva, H. (2000). *Proc. Natl Acad. Sci. USA*, **97**, 8849–8855.

Kratzschmar, J., Haendler, B., Eberspaecher, U., Roostermann, D., Donner, P. & Schleuning, W. D. (1996). *Eur. J. Biochem.* **236**, 827–836.

Leslie, A. G. W. (1992). *Jnt CCP4/ESF–EACBM Newsl. Protein Crystallogr.* **26**, 27–33.

Loon, L. C. van & van Kammen, A. (1970). *Virology*, **40**, 190–211.

Lu, G., Villalba, M., Coscia, M. R., Hoffman, D. R. & King, T. P. (1993). *J. Immunol.* **150**, 2823–2830.

McPherson, A. (1976). *Methods Biochem. Anal.* **25**, 249–345.

Maleck, K., Levine, A., Eulgem, T., Morgan, A., Schmid, J., Lawton, K. A., Dangl, J. L. & Dietrich, R. A. (2000). *Nature Genet.* **26**, 403–410.

Matthews, B. (1968). *J. Mol. Biol.* **33**, 491–497.

Milne, T. J., Abbenante, G., Tyndall, J. D., Halliday, J. & Lewis, R. J. (2003). *J. Biol. Chem.* **278**, 31105–31110.

Morrissette, J., Kratzschmar, J., Haendler, B., el-Hayek, R., Mochca-Morales, J., Martin, B. M., Patel, J. R., Moss, R. L., Scheuning, W. D. & Coronado, R. (1995). *Biophys. J.* **68**, 2280–2288.

Murphy, E. V., Zhang, Y., Zhu, W. & Biggs, J. (1995). *Gene*, **159**, 131–135.

Navaza, J. (1994). *Acta Cryst.* A**50**, 157–163.

Niderman, T., Genetet, I., Bruyere, T., Gees, R., Stintzi, A., Legrand, M., Fritig, B. & Mosinger, E. (1995). *Plant Physiol.* **108**, 17–27.

Rich, T., Chen, P., Furman, F., Huynh, N. & Isreal, M. A. (1996). *Gene*, **180**, 125–130.

Schuren, F. H., Asgeirsdotti, S. A., Kothe, E. M., Scheer, J. M. & Wessels, J. G. (1993). *J. Gen. Microbiol.* **139**, 2083–2090.

Zhang, Y., Fan, W., Kinkema, M., Li, X. & Dong, X. (1999). *Proc. Natl Acad. Sci. USA*, **96**, 6523–6528.