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

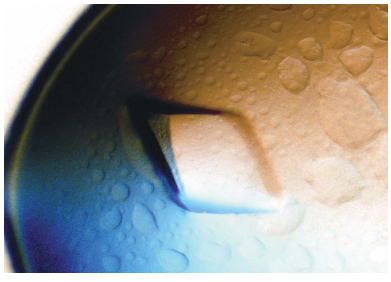
Crystallization and preliminary characterization of three different crystal forms of human saposin C heterologously expressed in *Pichia pastoris*

Robert Schultz-Heienbrok,[a] Natascha Remmel,[b] R. Klingenstein,[b] Maksim Rossocha,[a] Konrad Sandhoff,[b] Wolfram Saenger[a] and Timm Maier[a,c]*

[a]Institut für Chemie und Biochemie/Kristallographie, Freie Universität Berlin, Germany, [b]Kekule-Institut für Organische Chemie und Biochemie, Rheinische Friedrich-Wilhelms-Universität Bonn, Germany, and [c]Institute for Molecular Biology and Biophysics, Swiss Federal Institute of Technology, ETH Zürich, Switzerland

Correspondence e-mail: timm.maier@mol.biol.ethz.ch

Received 19 October 2005
Accepted 23 December 2005



The amphiphilic saposin proteins (A, B, C and D) act at the lipid–water interface in lysosomes, mediating the hydrolysis of membrane building blocks by water-soluble exohydrolases. Human saposin C activates glucocerebrosidase and β-galactosylceramidase. The protein has been expressed in *Pichia pastoris*, purified and crystallized in three different crystal forms, diffracting to a maximum resolution of 2.5 Å. Hexagonal crystals grew from 2-propanol-containing solution and contain a single molecule in the asymmetric unit according to the Matthews coefficient. Orthorhombic and tetragonal crystals were both obtained with pentaerythritol ethoxylate and are predicted to contain two molecules in the asymmetric unit. Attempts to determine the respective crystal structures by molecular replacement using either the known NMR structure of human saposin C or a related crystal structure as search models have so far failed. The failure of the molecular-replacement method is attributed to conformational changes of the protein, which are known to be required for its biological activity. Crystal structures of human saposin C therefore might be the key to mapping out the conformational trajectory of saposin-like proteins.

1. Introduction

Saposin-like proteins (SLPs) are a group of small (~80 amino acids in length) proteins present in all eukaryotes that exhibit diverse functions (Munford *et al.*, 1995). The key property of this protein superfamily is their interaction with biological membranes. SLPs display pairwise sequence identities of up to 60% and are consequently presumed to share a common fold. Experimentally derived structures of NKL (Liepinsh *et al.*, 1997), granulysin (Anderson *et al.*, 2003), saposin B (Ahn *et al.*, 2003) and saposin C (de Alba *et al.*, 2003; Hawkins *et al.*, 2005) are available. Members of the SLP superfamily are typically all-α proteins with four or five helices, of which helices I and IV/V as well as II and III are connected through conserved intramolecular disulfide bridges (Ahn *et al.*, 2003; Vaccaro *et al.*, 1995).

The SLP superfamily was named after the saposins (A, B, C and D), which are lysosomal activator proteins proteolytically derived from a single precursor protein, prosaposin. Saposins enable the breakdown of glycosphingolipids by activating lipid hydrolases (Furst & Sandhoff, 1992). Saposin C activates glucocerebrosidase and β-galactosylceramidase for decomposition of glucocerebrosides and galactosylceramides, respectively (O'Brien & Kishimoto, 1991). Mutations in either glucocerebrosidase or saposin C have been shown to cause Gaucher's disease, a metabolic storage disease resulting from sphingolipid accumulation (Horowitz & Zimran, 1994; Schnabel *et al.*, 1991).

Saposin C is an 81-amino-acid peptide with a single glycosylation site (Asn22) to which an N-linked oligosaccharide is covalently attached (Ito *et al.*, 1993). Recently, solution structures of human saposin C have been determined in different environments by NMR spectroscopy, showing that saposin C undergoes a conformational change in lipid micelles (de Alba *et al.*, 2003; Hawkins *et al.*, 2005). The molecule adopts a compact closed structure in the absence of detergent molecules (de Alba *et al.*, 2003) and opens up, exposing its hydrophobic core, when placed in a detergent environment (Hawkins *et al.*, 2005). In this transformation, the disulfide-bridged pairs of

helices move as rigid bodies around a hinge at the loops connecting helices I and II and helices IV and V.

Here, we report three different crystal forms of human saposin C heterologously expressed in *Pichia pastoris*. Structure determination by molecular replacement using the known SLP structures has so far been unsuccessful, suggesting that the crystal forms represent one or more conformational states of saposin C that have not been previously observed.

2. Materials and methods

2.1. Cloning and overexpression

The DNA sequence for human saposin C was amplified by PCR from the vector pBHE0 (Henseler *et al.*, 1996) using primers designed to introduce a C-terminal His_6 tag and appropriate restriction sites, resulting in an additional Arg residue before the His_6 tag. The PCR product was digested with *Sna*BI and *Mfe*I and subcloned into pPIC9K digested with *Eco*RI and *Sna*BI, generating the plasmid pPIC9K-C. *P. pastoris* strain GS115 (Invitrogen) was transformed by electroporation with linearized (*Sac*I restriction) pPIC9K-C and cells were grown on RD plates. Transformants were screened for multi-copy insertion by growth on YPD-Agar containing antibiotic G418 (Invitrogen) at concentrations up to 4 mg ml^{-1}. The phenotype of a clone growing at the highest G418 concentration was determined to be Mut+ by growth on methanol as the sole carbon source. For high cell-density fermentation, 3.5 l fermentation basal salts medium supplemented with 2 ml l^{-1} PTM1 trace salts solution was inoculated with 150 ml of a 24 h culture in MGY medium in a Bioflo3000 bioreactor (New Brunswick scientific). The temperature was adjusted to 303 K and the pH to 5.0 by addition of 20% NH_4OH. Oxygen was added to the air flow to maintain the level of dissolved oxygen. When glycerol from the MGY medium was metabolized, a 4 h feed-batch phase with a feed of 15 ml l^{-1} h^{-1} of 50% aqueous glycerol was started. The pH was then adjusted to and kept at 3.0 and the glycerol feed was stopped. Cells were adapted to methanol by a slowly increasing feed of methanol supplemented with 2 ml l^{-1} trace salts solution until a feed rate of 12 ml l^{-1} h^{-1} was reached. Induction with methanol was then continued for 96 h, after which the medium was harvested by centrifugation for 15 min at 1000*g* followed by a second centrifugation for 30 min at 15 000*g*.

2.2. Protein purification

Saposin C was captured from the medium by cation-exchange chromatography on Poros HS20 (Applied Biosystems) using 50 m*M* sodium phosphate pH 4 for washing and a linear gradient of the same buffer supplemented with 1 *M* NaCl for elution. The protein was loaded onto Ni^{2+}–NTA Superflow (Qiagen) after adjusting the pH to 7.5 with 0.1 *M* NaOH. The column was washed with 50 m*M* sodium phosphate buffer pH 6.0 and the protein was eluted with 300 m*M* NaCl in 50 m*M* sodium phosphate buffer pH 4.0. The buffer was exchanged to 25 m*M* sodium citrate pH 4.0, 100 m*M* NaCl and the protein was concentrated to either 8 or 25 mg ml^{-1} as determined by BSA assay (BioRad). The purity of the sample was judged on a silver-stained SDS precast polyacrylamide gel to be >99%. The identity of the protein was confirmed by MALDI–TOF mass spectrometry (Peter Franke, Freie Universität Berlin, Germany).

2.3. Crystallization

Crystallization of saposin C was achieved by vapour diffusion using either the hanging-drop or sitting-drop method. Initial crystallization conditions were established from various sparse-matrix and grid screens. The crystals were further optimized by varying the precipitant concentration, pH and additives. One hexagonal, one tetragonal and one orthorhombic crystal form were obtained.

The crystallization condition for the hexagonal crystal form was identified in 2-propanol *versus* pH grid screens: the optimized reservoir solution contained 50 m*M* sodium citrate pH 4.0, 32.5%(*v*/*v*) 2-propanol. 2 µl reservoir solution was mixed with 2 µl protein solution (8 mg ml^{-1}) in a hanging-drop setup. The orthorhombic crystal form grew from reservoir solution containing 50 m*M* Bis-Tris pH 6.5, 30%(*v*/*v*) pentaerythritol ethoxylate 15/4 (15/4 EO/OH) and 50 m*M* ammonium sulfate (Hampton Research Index Screen condition No. 57), mixing 1 µl reservoir solution with 1 µl protein solution (25 mg ml^{-1}) in a sitting-drop setup in Greiner 96-well plates with round-shaped wells and concave bottom and equilibrating against 100 µl reservoir solution. Systematic screening around the former condition in a hanging-drop setup mixing 2 µl reservoir solution with 2 µl protein solution (25 mg ml^{-1}) resulted in the growth of a tetragonal crystal form with the following optimized conditions: 42%(*v*/*v*) pentaerythritol ethoxylate 15/4 (15/4 EO/OH), 50 m*M* sodium acetate, 300 m*M* $MgSO_4$. All hanging-drop crystallization was carried out with siliconized cover slips using 24-well plates from Nelipak, Venray, The Netherlands with a reservoir volume of 750 µl.

2.4. Data collection and processing

X-ray diffraction data from the hexagonal and the tetragonal crystals were both measured at beamline ID29 at ESRF, Grenoble, France on a Quantum 4 CCD area detector (ADSC, Area Detector System Corporation) at 100 K. The hexagonal crystal was taken from the mother liquor and flash-frozen in liquid nitrogen, whereas the tetragonal crystal was incubated in reservoir solution supplemented with 22% glycerol as a cryoprotectant prior to freezing in liquid nitrogen. X-ray data collection of the orthorhombic crystal form was performed at room temperature in a glass capillary using an in-house Enraf–Nonius FR571 rotating-anode generator equipped with an Osmic MaxFlux mirror system and a MAR 345 imaging-plate detector. Programs from the *HKL* software suite were used for data reduction (Otwinowski & Minor, 1997).

3. Results and discussion

To overcome the limitations of other expression systems such as missing glycosylation (*Escherichia coli*), poor yield (natural sources) or high expense (chemical synthesis), human saposin C was over-expressed in the methylotrophic yeast *P. pastoris*. The expression construct included a C-terminal His_6 tag which was not cleaved off prior to crystallization. It has previously been shown that human saposin C expressed with a C-terminal His tag is fully functional, as assayed by binding to phosphatidylserine and activation of glucocerebrosidase (Qi & Grabowski, 2001; Qi *et al.*, 1994). Compared with *E. coli*, overexpression in *P. pastoris* offers the additional benefit of extremely facilitated downstream purification since the cells secrete mainly the overexpressed protein. Thus, saposin C could be directly captured from the medium. The protein expression and purification yield was about 40 mg of protein per litre of medium, which is comparable to the yield typically obtained with an *E. coli* expression system.

The expectation that overexpression in yeast cells also resulted in glycosylation of Asn22 was not fulfilled. Analysing the sample by MALDI–TOF revealed a peak at 10 252.2 Da that agreed well with

Table 1
Data-collection statistics of the three different crystal forms.

Values in parentheses refer to the highest resolution shell.

	Hexagonal			
	High resolution	SAD	Tetragonal	Orthorhombic
X-ray source	ESRF, ID29	ESRF, ID29	ESRF, ID29	Rotating-anode generator
Wavelength (Å)	0.871	1.879	0.933	1.542
Space group	$P6_{1/3/5}22$	$P6_{1/3/5}22$	$P4_{1/3}2_12$	$C222_1$
Unit-cell parameters (Å)	a = 53.7, c = 117.0	a = 53.6, c = 117.1	a = 48.9, c = 154.3	a = 57.0, b = 93.5, c = 88.3
Resolution (Å)	30.0–2.5 (2.54–2.45)	30.0–3.0 (3.11–3.00)	30–2.6 (2.69–2.6)	40–2.4 (2.49–2.4)
Mosaicity (°)	0.6	1.4	0.8	0.5
Total observations	26693	83253	61224	34121
Unique reflections	4057	2284	6286	9532
Completeness	99.5 (99.7)	100 (100)	99.9 (100)	99.2 (99.3)
$I/\sigma(I)$	15.2 (4.2)	47.4 (7.4)	22.8 (5.3)	19.0 (1.83)
R_{sym}	10.7 (43.7)	8.6 (74.5)	8.6 (52.8)	6.4 (72.3)
Molecules per ASU	1	1	2	2
Matthews coefficient (Å^3 Da^{-1})	2.4	2.4	2.3	2.9

the length of the cloned sequence including the His_6 tag with an additional arginyl residue but without glycosylation. However, glycosylation is neither essential for binding to phospholipids nor for activation of lysosomal hydrolases (Hiraiwa *et al.*, 1993; Vielhaber *et al.*, 1996). Four other peaks were found at weights incrementally decreasing by 137 Da, indicating that the histidyl residues were either partially cleaved off or only partially expressed.

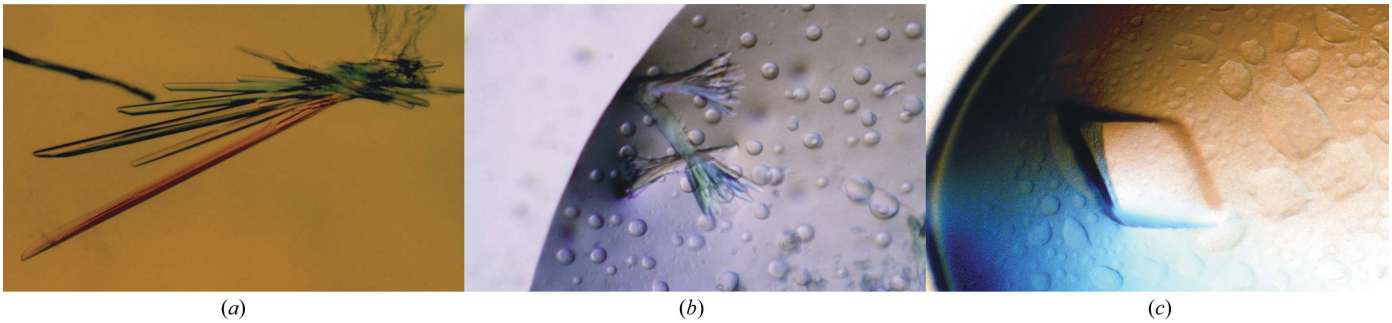
Three different crystal forms of human saposin C were obtained using two different precipitating agents. Hexagonal crystals were obtained using 2-propanol as a precipitant and grew as needles up to a length of 750 µm (Fig. 1*a*). The Bravais lattice of the crystals is primitive hexagonal with point group *P*622. However, the analysis of systematic absences along 00*l* was ambiguous and leaves the space group open to be either $P6_322$ or one of the enantiomeric space groups $P6_122$ and $P6_522$. Determination of the true space group must await structure determination.

The tetragonal crystal forms grew in the presence of pentaerythritol ethoxylate 15/4. Typical crystals obtained from these conditions are shown in Fig. 1(*b*). The needle-like crystals reached a maximum length of 300 µm, contained two molecules in the asymmetric unit and have unit-cell parameters a = 48.9, c = 154.3 Å. From the systematic absences the space group could be determined to be either $P4_12_12$ or $P4_32_12$.

Whereas the tetragonal crystals were obtained in sitting drops using pentaerythritol ethoxylate 15/4 as the precipitant, an orthorhombic crystal form was obtained in a hanging-drop vapour-diffusion setup using increased pentaerythritol ethoxylate 15/4 concentrations and magnesium chloride as an additive. Typical crystals are shown in Fig. 1(*c*). The crystal dimensions reached approximately 500 × 300 × 200 µm. No flash-freezing protocol could be established for these crystals and thus their diffraction had to be measured at room temperature using a rotating-anode source. The unit-cell parameters obtained were a = 57.0, b = 93.5, c = 88.3 Å. The space group could be unambiguously determined as $C222_1$ from systematic absences. Data-collection statistics for each of the three different crystal forms are compiled in Table 1.

Interestingly, phase determination by molecular replacement using various approaches as implemented in *AMoRe* (Navaza, 1994), *MOLREP* (Vagin & Teplyakov, 1997) and *Phaser* (Read, 2001) including ensemble searches failed for all saposin C crystal forms using either the SapC NMR structures (PDB codes 1m12 and 1sn6) or structures of other SLPs, such as the crystal structure of granulysin (PDB code 1l9l) or the NMR structure of NKL (PDB code 1nkl), as search models. Since data-collection statistics gave no indication of twinning and since the identity of the saposin protein has been unambiguously established by overexpression in *P. pastoris*, MALDI–TOF mass spectrometry and SDS–PAGE analysis, we conclude that the failure of structure determination by molecular replacement arises from considerable conformational differences between the search models and the crystallized saposin C. Comparison of the two known NMR structures of saposin C has already demonstrated flexibility in this protein (de Alba *et al.*, 2003; Hawkins *et al.*, 2005). Whether each of the three different crystal forms described here represents a distinct conformational state remains to be seen. Based on the two NMR structures, we are currently calculating theoretical intermediate structures using Cartesian interpolation (Vonrhein *et al.*, 1995) and normal-mode analysis (Suhre & Sanejouand, 2004) in order to generate improved search models.

A second strategy we are pursuing to determine the structure of saposin C is to obtain experimental phases by exploiting the anom-



(*a*) (*b*) (*c*)

Figure 1
(*a*) Hexagonal crystals of human saposin C with a maximum length of 750 µm. (*b*) Tetragonal crystals with a maximum length of 300 µm. (*c*) An orthorhombic crystal of 500 µm in the longest direction.

alous signal of the eight S atoms (three disulfide bridges, two methionine residues) in the protein. For this type of SAD experiment the hexagonal crystals were used to collect highly redundant data sets at an X-ray wavelength of 1.8786 Å in order to maximize the anomalous signal of sulfur (Table 1). We expect that this and future measurements with further improved data quality will allow phase determination by sulfur SAD.

From structure determination of the three different crystal forms, we expect to obtain a detailed picture on the dynamic behaviour of human saposin C and likely of other members of the saposin-like family. Determining structures of saposin C in different conformational states by X-ray crystallography may ultimately define the trajectory of its opening motion, which is a prerequisite and specificity-determining factor for substrate interaction in the whole family of SLP proteins.

We thank Claudia Alings for technical assistance and Peter Franke for providing mass spectra. We are also grateful for beamtime and support during data collection at the ESRF in Grenoble. This work was supported by Fonds der Chemischen Industrie and Deutsche Forschungsgemeinschaft.

References

Ahn, V. E., Faull, K. F., Whitelegge, J. P., Fluharty, A. L. & Prive, G. G. (2003). *Proc. Natl Acad. Sci. USA*, **100**, 38–43.
Alba, E. de, Weiler, S. & Tjandra, N. (2003). *Biochemistry*, **42**, 14729–14740.
Anderson, D. H., Sawaya, M. R., Cascio, D., Ernst, W., Modlin, R., Krensky, A. & Eisenberg, D. (2003). *J. Mol. Biol.* **325**, 355–365.
Furst, W. & Sandhoff, K. (1992). *Biochim. Biophys. Acta*, **1126**, 1–16.
Hawkins, C. A., Alba, E. & Tjandra, N. (2005). *J. Mol. Biol.* **346**, 1381–1392.
Henseler, M., Klein, A., Glombitza, G. J., Suziki, K. & Sandhoff, K. (1996). *J. Biol. Chem.* **271**, 8416–8423.
Hiraiwa, M., Soeda, S., Martin, B. M., Fluharty, A. L., Hirabayashi, Y., O'Brien, J. S. & Kishimoto, Y. (1993). *Arch. Biochem. Biophys.* **303**, 326–331.
Horowitz, M. & Zimran, A. (1994). *Hum. Mutat.* **3**, 1–11.
Ito, K., Takahashi, N., Takahashi, A., Shimada, I., Arata, Y., O'Brien, J. S. & Kishimoto, Y. (1993). *Eur. J. Biochem.* **215**, 171–179.
Liepinsh, E., Andersson, M., Ruysschaert, J. M. & Otting, G. (1997). *Nature Struct. Biol.* **4**, 793–795.
Munford, R. S., Sheppard, P. O. & O'Hara, P. J. (1995). *J. Lipid Res.* **36**, 1653–1663.
Navaza, J. (1994). *Acta Cryst.* A**50**, 157–163.
O'Brien, J. S. & Kishimoto, Y. (1991). *FASEB J.* **5**, 301–308.
Otwinowski, Z. & Minor, W. (1997). *Methods Enzymol.* **276**, 307–326.
Qi, X. & Grabowski, G. A. (2001). *J. Biol. Chem.* **276**, 27010–27017.
Qi, X., Leonova, T. & Grabowski, G. A. (1994). *J. Biol. Chem.* **269**, 16746–16753.
Read, R. J. (2001). *Acta Cryst.* D**57**, 1373–1382.
Schnabel, D., Schroder, M. & Sandhoff, K. (1991). *FEBS Lett.* **284**, 57–59.
Suhre, K. & Sanejouand, Y. H. (2004). *Nucleic Acids Res.* **32**, W610–W614.
Vaccaro, A. M., Salvioli, R., Barca, A., Tatti, M., Ciaffoni, F., Maras, B., Siciliano, R., Zappacosta, F., Amoresano, A. & Pucci, P. (1995). *J. Biol. Chem.* **270**, 9953–9960.
Vagin, A. & Teplyakov, A. (1997). *J. Appl. Cryst.* **30**, 1022–1025.
Vielhaber, G., Hurwitz, R. & Sandhoff, K. (1996). *J. Biol. Chem.* **271**, 32438–32446.
Vonrhein, C., Schlauderer, G. J. & Schulz, G. E. (1995). *Structure*, **3**, 483–490.