Acta Crystallographica Section D
Biological Crystallography
ISSN 0907-4449

Expression, purification, crystallization and preliminary X-ray analysis of the cathelicidin motif of the protegrin-3 precursor

Jean Frédéric Sanchez,[a] François Hoh,[a] Marie-Paule Strub,[a] Jean Marc Strub,[b] Alain Van Dorsselaer,[b] Robert Lehrer,[c] Tomas Ganz,[c] Alain Chavanieu,[a] Bernard Calas,[a] Christian Dumas[a] and André Aumelas[a]*

[a]Centre de Biochimie Structurale, UMR 5048 CNRS-UM1/UMR 554 INSERM-UM1, Université Montpellier 1, Faculté de Pharmacie, 15 Avenue Charles Flahault, 34060 Montpellier CEDEX 2, France, [b]Laboratoire de Spectrométrie de Masse Bio-Organique, 1 Rue Blaise Pascal, Faculté de Chimie, 67000 Strasbourg, France, and [c]Department of Medicine, Center for the Health Sciences, Los Angeles, CA 90024, USA

Correspondence e-mail: aumelas@cbs.univ-montp1.fr

Received 10 May 2001
Accepted 24 July 2001

Numerous precursors of antibacterial peptides with unrelated sequences share a similar prosequence which belongs to the cathelicidin family of proteins. The three-dimensional structure of this cathelicidin motif, which contains two disulfide bonds, has not yet been reported. The cathelicidin motif (ProS) of the protegrin-3 precursor was overexpressed in *Escherichia coli* as a His-tagged protein. The His_6 tag was removed by thrombin cleavage. ProS was purified to homogeneity and single crystals were obtained by the hanging-drop vapour-diffusion method at pH 3–4. Preliminary X-ray diffraction analysis indicated that these crystals belong to the hexagonal space group $P6_122$ or $P6_522$, with unit-cell parameters $a = b = 51.42$, $c = 134.25$ Å. These crystals diffracted beyond 2.75 Å (1.9 Å at ESRF) and contain one molecule per asymmetric unit.

1. Introduction

Numerous precursors of antibacterial peptides with unrelated sequences, including the protegrin precursors, share a similar prosequence of 96–101 residues and belong to cathelicidin family (Gennaro *et al.*, 1998; Gennaro & Zanetti, 2000; Ritonja *et al.*, 1989; Scocchi *et al.*, 1997; Storici *et al.*, 1996; Wang *et al.*, 1998; Zanetti *et al.*, 1995, 2000; Zhao *et al.*, 1994). Protegrins (PG-1 to PG-5, 16–18 residues) are a family of five antibacterial peptides isolated from porcine leucocytes (Kokryakov *et al.*, 1993). They are initially synthesized as a 149-residue precursor (147 residues for PG-2) devoid of antibacterial activity, in which three regions have been identified: the signal peptide (sequence 1–29), the prosequence (sequence 30–130) and the protegrin sequence (sequence 131–148 for PG-1 and PG-3). The Gly149 is removed in a well known amidation step (Merkler, 1994; Suzuki *et al.*, 1990; Zhao *et al.*, 1997). The protegrin prosequence, referred to as ProS, contains four cysteines engaged in two disulfide bonds. Their 1–2, 3–4 arrangement was clearly established for probactenecin-7, a bovine cathelicidin (Storici *et al.*, 1996). Taking into account the high degree of sequence identity between the probactenecin and ProS sequences (74–78%) and the alignment of the four cysteines, the disulfide-bond arrangement of ProS was assumed to be identical to that of probactenecin. Therefore, the two disulfide bonds of ProS are Cys85–Cys96 and Cys107–Cys124. The detailed three-dimensional structure of this widespread cathelicidin motif is not yet known. To determine the structure of the cathelicidin motif of the protegrin-3 precursor, ProS was overexpressed in *E. coli*. Here, we report the overexpression, purification, crystallization and preliminary X-ray diffraction analysis of the recombinant ProS.

2. Materials and methods

2.1. Overexpression and purification of the His-tagged ProS

A PCR product containing the coding sequence of ProS was generated from the plasmid pBluescript including the pro-protegrin-3 cDNA (Panyutich *et al.*, 1997). The cDNA encoding for the ProS sequence (residues 30–130) was subcloned into the *Nde*I/*Bam*HI sites of the pET-15b plasmid vector (Novagen) leading to the N-terminus in frame fusion with the His tag. *E. coli* strain BL21(DE3), transformed with the recombinant plasmid, was grown in 400 ml of medium (20 g l^{-1} tryptone, 10 g l^{-1} yeast extract, 5 g l^{-1} sodium citrate, 5 g l^{-1} KH_2PO_4 adjusted to pH 7.0) containing 100 mg l^{-1} ampicillin. After overnight growth at 310 K, the culture was used to inoculate 4 l of the same medium supplemented with ampicillin (100 mg l^{-1}), $MgSO_4$ (10 m*M*), glucose (5 g l^{-1}), biotin (1 mg l^{-1}), thiamine (10 mg l^{-1}) and nicotinamide (10 mg l^{-1}) and grown at 310 K for 3 h. Expression was induced for 4 h by addition of 1 m*M* IPTG. The cells were harvested by centrifugation for 15 min at 8000 rev min^{-1}. Cells, 20 g wet weight, were resuspended in 50 ml of ice-cold buffer *A* (100 m*M* Tris–HCl pH 7.5, 150 m*M* NaCl) containing 5 m*M* benzamidine. The mixture was homogenized and freezed/thawed three times. The lysate was

Table 1
Data-collection statistics.

Values in parentheses refer to data in the last resolution shell (2.80–2.75 Å).

Space group	$P6_122$ or $P6_522$
Unit-cell parameters (Å)	
$a = b$	51.42
c	134.25
Resolution range (Å)	23–2.75
Asymmetric unit content	Monomer
Total No. of reflections	21019
No. of unique reflections	3110
Redundancy	6.6
Average $I/\sigma(I)$	16.1 (3.6)
Completeness (%)	91.3 (58)
$R_{\rm merge}$† (%)	6.2 (28.6)

† $R_{\rm merge} = \sum |I_h - \langle I_h \rangle| / \sum I_h$.

kept on ice and probe-sonicated for 1 min with 0.1 s bursts at 340 W. Particulate material was then removed by centrifugation at 20 000*g* for 30 min at 277 K. The supernatant was further clarified by centrifugation at 90 000*g* for 30 min and applied to a Q-Sepharose (Amersham Pharmacia Biotech) column (20 × 2.6 cm) equilibrated with buffer *A*.

The flowthrough containing the protein was applied to an Ni-NTA Superflow (Qiagen) column (2 × 2.5 cm) equilibrated with buffer *A*. The column was washed with 30 ml of buffer *A*, 50 ml of buffer *A* containing 1 *M* NaCl and then re-equilibrated with buffer *A*. The protein was eluted with 0.5 *M* imidazole in the above buffer and 10 ml fractions of eluent were analyzed by Coomassie-staining SDS–PAGE gels. The protein-containing fractions 1–4 were pooled and concentrated 20 times using an Ultrafree (Millipore) filter with a MW cutoff of 5000 Da. Imidazole was removed by several concentration/dilution cycles with buffer *A*. The concentration of the sample was estimated by UV–visible spectroscopy using a calculated extinction coefficient of 2800 $\rm mol^{-1}$ $\rm cm^{-1}$ at 280 nm.

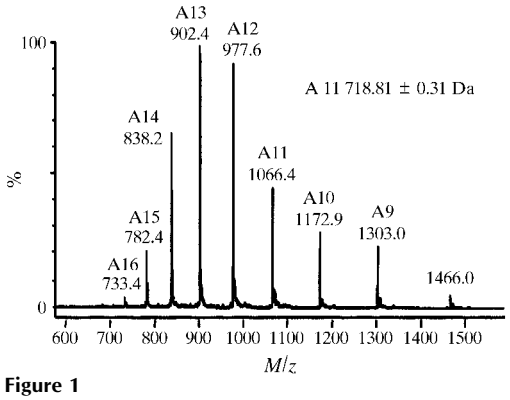

Figure 1
Molecular mass spectrum of the recombinant ProS. Mass analysis of the recombinant ProS was performed using the electrospray mass spectrometry technique (ES-MS) on a VG Bio-Q quadrupole with a mass range of 4000 Da (Bio-Tech, Manchester, England) in the positive mode. The protein was desalted on Zip-Tip (Millipore) and 10 pmol were used for mass analysis. The calibration was performed using the multiply charged ions produced by a separate introduction of horse heart myoglobin (16 951.4 Da).

2.2. His-tag cleavage and purification of ProS

The protein was digested at 298 K in the presence of 2 m*M* $CaCl_2$ with 10 units of thrombin per milligram of protein. The cleavage was complete after 1 h. The protein was purified by gel filtration on a Sephadex HR-100 (Pharmacia Biotech) column (100 × 1.6 cm) equilibrated with buffer *B* (50 m*M* sodium phosphate, 50 m*M* NaCl pH 7.0). 1 ml fractions were collected at a flow rate of 0.6 ml $\rm min^{-1}$. The protein was eluted as a single peak and the corresponding fractions pooled and concentrated. About 10–12 mg of purified ProS were obtained from 20 g wet weight of cells. Purity was controlled by Coomassie-staining SDS–PAGE gels, ^{1}H NMR and mass spectrometry (Fig. 1).

2.3. Crystallization

The hanging-drop vapour-diffusion method was used to screen and to optimize the crystallization conditions. The drops contained equal volumes (1 µl) of ProS protein solution (35 mg $\rm ml^{-1}$ in 10 m*M* Tris–HCl buffer pH 7.0) and reservoir solution. The optimized reservoir solution consisted of 1.2–1.4 *M* ammonium sulfate and 0.1 *M* sodium citrate buffer at pH 3.8. Two types of hexagonal diffraction-quality crystals were reproducibly obtained within a week at 291 K. From drops containing 1.2 *M* ammonium sulfate reservoir solution, crystals reached dimensions of about 0.2 × 0.1 × 0.05 mm, whereas larger crystals were obtained from drops containing 1.4 *M* ammonium sulfate reservoir solution (Fig. 2). Prior to data collection, crystals were transferred to cryoprotectant mother liquor supplemented with 30% glycerol and flash-frozen.

3. Results and discussion

Well diffracting crystals of the recombinant ProS protein, corresponding to the cathelicidin protegrin-3 precursor, were obtained. X-ray data were collected at 100 K on a MAR Research image-plate detector (Hamburg, Germany) mounted on a Rigaku RU-200 rotating-anode generator (Tokyo, Japan) operating at 40 kV and 90 mA. The crystal-to-detector distance was 140 mm and the oscillation range was 0.5° per film (25 min exposure time). The images were processed and scaled with the *HKL* suite of programs (Otwinowski & Minor, 1997). The autoindexing procedure indicated that the two types of crystals belong to a primitive hexagonal space group, with unit-cell parameters $a = b = 51.42$, $c = 134.25$ Å. The lattice was assigned to one of the enantiomorphic space groups $P6_122$ or $P6_522$. Considering a molecular weight of 11 718 Da for the protein, the Matthews coefficient was determined to be 2.19 $\rm Å^3\ Da^{-1}$, assuming one molecule per asymmetric unit, with a corresponding solvent content of 44%. Table 1 summarizes the data-collection statistics. While good-quality data could be collected to 2.75 Å resolution on a conventional X-ray source, the same crystal diffracted beyond 1.9 Å resolution using synchrotron radiation (ESRF, Grenoble, FIP beamline).

Initial attempts to solve the structure by molecular replacement using various structures supposed to share a similar global fold (PDB entries 1cew, 1stf, 1dvd, 1mnl, 1eqk, 1g96) as search models have so far proven unsuccessful. Screening for heavy-atom derivatives is now under way.

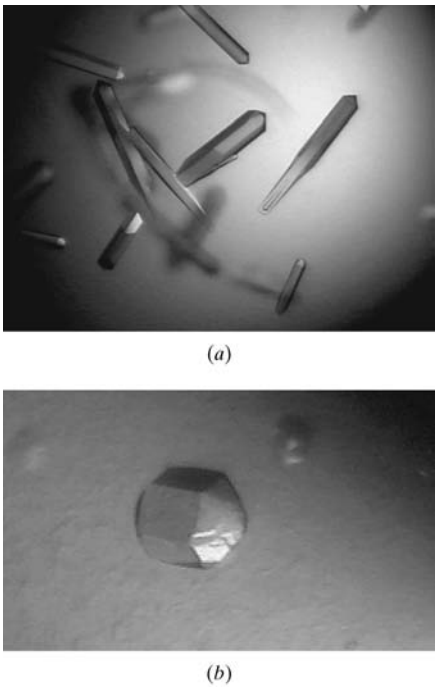

Figure 2
Photomicrographs of the two types of crystals obtained from drops containing (*a*) 1.2 *M* and (*b*) 1.4 *M* ammonium sulfate of the recombinant ProS overexpressed in *E. coli*. Typical dimensions are about 0.1 × 0.1 × 0.08 mm and 0.2 × 0.05 × 0.05 mm for *A* and *B*, respectively.

This research was supported by a grant from the Ministère de l'Education Nationale, de la Recherche et de la Technologie (JFS). We acknowledge the opportunity to test preliminary diffraction data at the ESRF synchrotron (Grenoble, France) and are indebted to Dr S. L. Salhi for editorial revision of the manuscript.

References

Gennaro, R., Scocchi, M., Merluzzi, L. & Zanetti, M. (1998). *Biochim. Biophys. Acta*, **1425**, 361–368.

Gennaro, R. & Zanetti, M. (2000). *Biopolymers*, **55**, 31–49.

Kokryakov, V. N., Harwig, S. S., Panyutich, E. A., Shevchenko, A. A., Aleshina, G. M., Shamova, O. V., Korneva, H. A. & Lehrer, R. I. (1993). *FEBS Lett.* **327**, 231–236.

Merkler, D. J. (1994). *Enzyme Microb. Technol.* **16**, 450–456.

Otwinowski, Z. & Minor, W. (1997). *Methods Enzymol.* **276**, 307–326.

Panyutich, A., Shi, J., Boutz, P. L., Zhao, C. & Ganz, T. (1997). *Infect. Immun.* **65**, 978–985.

Ritonja, A., Kopitar, M., Jerala, R. & Turk, V. (1989). *FEBS Lett.* **255**, 211–214.

Scocchi, M., Wang, S. & Zanetti, M. (1997). *FEBS Lett.* **417**, 311–315.

Storici, P., Tossi, A., Lenarcic, B. & Romeo, D. (1996). *Eur. J. Biochem.* **238**, 769–776.

Suzuki, K., Shimoi, H., Iwasaki, Y., Kawahara, T., Matsuura, Y. & Nishikawa, Y. (1990). *EMBO J.* **9**, 4259–4265.

Wang, Y., Agerberth, B. & Johansson, J. (1998). *J. Protein Chem.* **17**, 522–523.

Zanetti, M., Gennaro, R. & Romeo, D. (1995). *FEBS Lett.* **374**, 1–5.

Zanetti, M., Gennaro, R., Scocchi, M. & Skerlavaj, B. (2000). *Adv. Exp. Med. Biol.* **479**, 203–218.

Zhao, C., Liaw, L., Lee, I. H. & Lehrer, R. I. (1997). *FEBS Lett.* **410**, 490–492.

Zhao, C., Liu, L. & Lehrer, R. I. (1994). *FEBS Lett.* **346**, 285–288.