Acta Crystallographica Section D
Biological Crystallography
ISSN 0907-4449

Crystallization and preliminary X-ray analysis of the tumor metastasis factor p37

Robbie Reutzel, Susan K. Boehlein, Lakshmanan Govindasamy, Ryan B. Brenes, Mavis Agbandje-McKenna, Sheldon M. Schuster and Robert McKenna*

Department of Biochemistry and Molecular Biology, University of Florida, Gainesville, FL 32610, USA

Correspondence e-mail: rmckenna@ufl.edu

Received 5 June 2002
Accepted 27 August 2002

P37, an outer-membrane bacterial protein from *Mycoplasma hyorhinis*, is a molecule whose presence on the surface of many tumor cells correlates highly with increased neoplastic invasivity and metastasis. P37 was overexpressed in *Escherichia coli*, purified by affinity chromatography and crystallized. Useful single crystals for X-ray diffraction structural studies have been grown by oil-immersion methods from a solution of 40% PEG 4000, 0.1 *M* ammonium bromide in a 0.1 *M* citrate buffer at pH 4.0. X-ray diffraction data were collected at the F2 beamline at CHESS with a crystal-to-CCD detector distance of 150 mm, collecting 1° oscillation slices with an exposure time of 30 s per frame. A 212° sweep of data (99.8% completeness) were collected from a single crystal under cryoconditions, with a maximal useful diffraction pattern to 1.8 Å resolution. The crystals are shown to be monoclinic and have been assigned to space group $P2_1$, with unit-cell parameters a = 50.02, b = 67.26, c = 59.89 Å, β = 108.29° and a scaling $R_{\rm sym}$ of 0.076 for 34 882 unique reflections. Packing considerations indicate that there is one molecule per asymmetric unit. It is expected that in the near future the structure of p37 will be obtained using phases from traditional heavy-atom isomorphous replacement and/or halide-soak methods. Elucidation of the structure of p37 may be paramount to producing new antibody-based anticancer therapeutic agents.

1. Introduction

The primary treatment strategies for cancer are mainly concerned with the removal and/or debulking of the primary tumor mass at its site of origin and then killing as many of the remaining cancer cells as possible. Unfortunately, tumors shed a small percentage of their origin cells, some of which find their way through the circulatory or lymphatic systems to distant sites. These cells then go through a complex and poorly understood process to form new tumor loci called metastases, that are most often the lethal component of cancer.

p37 is an outer-membrane bacterial protein from *Mycoplasma hyorhinis* whose presence on the surface of neoplastic cells has been correlated highly with the metastatic nature of these tumors. p37 was initially discovered in an effort to heighten the immune response of cancer patients by immunizing them with their own surgically debulked tumors (Fareed *et al.*, 1988). Immunized patients were found, by immunoblot analysis, to exhibit measurable antibody titers against several antigens, including a 38 kDa protein that had more commonly been designated p37 (Dudler *et al.*, 1988). The actual purified p37 protein has a molecular weight of 43.5 kDa but migrates to the 37 kDa mark on a silver-stained gel. Addition of purified recombinant p37 to the culture media of both melanoma and prostate carcinoma cell lines was shown to greatly increase the invasivity of these cells (Boehlein *et al.*, 2002). Furthermore, localization of p37 on the tumor cell membrane as well as neutralization of the metastatic potential of these cells by monoclonal antibody addition suggests that this protein does not require additional mycoplasmal proteins to exert its invasive effects. Similar findings have been reported using FS9 mouse sarcoma cells in culture (Steinemann *et al.*, 1984).

The *M. hyorhinis* operon that encodes this 403 amino-acid protein has also been shown to contain the genes for two additional proteins whose function is not yet known (Ilantzis *et al.*, 1993). Upon sequence analysis, it was revealed that this operon shares 41% sequence similarity to a periplasmic binding-protein-dependent transport system as found in certain Gram-negative bacteria (Dudler *et al.*, 1988; Gilson *et al.*, 1988). This may indicate that p37 works as part of a mycoplasmal transport system.

Although p37 has not been implicated directly, mycoplasmal infection has been linked to many cancers. One group of researchers has reported that 48% of patients suffering from

gastric cancer were also positive for *M. hyorhinis* infection (Sasaki *et al.*, 1995). Another group was able to locate p37 in paraffin-embedded carcinoma cells by immunohistochemistry and then independently verify the presence of mycoplasma DNA by PCR, indicating that greater than 50% of colon carcinoma, gastric carcinoma, esophageal cancer and lung cancer samples were positive for *M. hyorhinis* infection. Other gastrointestinal afflictions such as chronic superficial gastritis, intestinal metaplasia and gastric ulcer have shown a 25% mycoplasma infection rate (Huang *et al.*, 2001). These results indicate that p37 may be linked to tumor invasivity and increased metastasis as well as several gastrointestinal disorders.

An interesting aspect of the possible structure of p37 is its sequence similarity with both hemagglutinin A (Sauter *et al.*, 1992) of avian influenza virus and cathepsin B (Jia *et al.*, 1995) from *Schistosoma mansoni* found using the web-based *BLAST* server (Altschul *et al.*, 1997). Residues 186–343 of p37 have been shown to share a partial pairwise sequence homology (33/159 residues) with the distal head domain (residues 118–271) of hemagglutinin A of avian influenza virus and residues 153–254 of p37 share a partial pairwise sequence homology (38/100 residues) with several β-strands (residues 228–329) of cathepsin B from *S. mansoni*. A model based on the similarity of p37 to hemagglutinin A has been built and may be useful as a search model for molecular-replacement phasing methods (Boehlein *et al.*, 2002).

In this paper, we report the purification, crystallization and preliminary X-ray diffraction data analyses of p37. Structural determination of the p37 protein will help in understanding its function and could be used to develop new antibody-based cancer therapies.

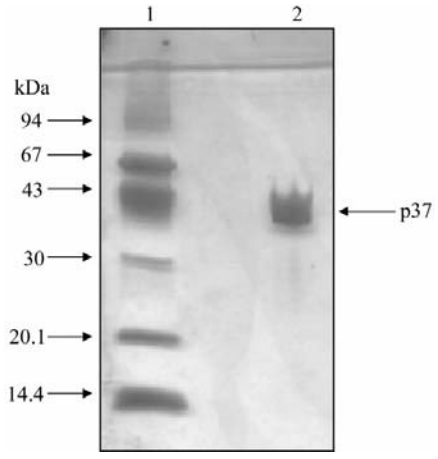

Figure 1
SDS silver-stained gel of purified p37 sample used for crystallization. Lane 1, low molecular-weight standards; lane 2, p37 sample (10 mg ml^{-1} in 50 m*M* Tris–HCl pH 7.5).

2. Methods and results

2.1. Expression and purification

A 380-amino-acid truncated form of p37 was expressed and purified by the method described by Boehlein *et al.* (2002). This truncated form did not express the N-terminal 23-residue long signal peptide and membrane-anchoring diglyceride. This enables the expression of a more soluble form of the protein. To express p37, the coding region was amplified by PCR and the resulting cDNA was sequenced. This product was then ligated into a pET31f1m1 expression vector and used to transform BL21DE3pLysS bacteria. Cells were selected for using ampicillin and chloramphenicol resistance on agar plates and freshly transformed colonies were used to inoculate media.

The cells were lysed by vortexing the pellet in 1/10 the original volume in a 20 m*M* phosphate buffer pH 7.8, followed by sonication for three 15 s cycles. The resulting crude cell lysate was then centrifuged at 40 000*g* for 20 min at 277 K to remove cell debris. The clear supernatant was then subjected to ion-exchange chromatography using the Econo-System (BioRad). A 5 ml cation-exchange S column was attached to the bottom of a 50 ml anion-exchange Q column and equilibrated with 20 m*M* phosphate buffer pH 7.95 at a flow rate of 2.5 ml min^{-1}. The soluble cellular extract was loaded on the column. The flow-through containing p37 was pooled. This pooled sample was adjusted to pH 6.1 with 2 *M* acetic acid and loaded onto a 5 ml cation exchanger, Econo-Pac S cartridge, equilibrated with 20 m*M* NaOAc pH 6.1. The column was washed with 5% 20 m*M* NaOAc pH 6.1 with 1 *M* NaCl (buffer *B*) and the p37 protein was eluted off the column with 15% buffer *B*. The eluted sample was then concentrated using a Centriprep 10 spin column (Millipore, Bedford, MA, USA).

The purity and concentration of the p37 sample were assessed by silver-stained SDS gel electrophoresis (Fig. 1) and UV spectroscopy optical density readings, respectively.

2.2. Crystallization and X-ray analysis

Initial crystal screening used the hanging-drop vapor-diffusion crystal-growth method (McPherson, 1982) at room temperature with conditions specified in the Hampton Research Cryoscreen kit. Preliminary crystals were grown from a drop consisting of 2 µl of purified p37 (10 mg ml^{-1} in 50 m*M* Tris–HCl pH 7.5) and 2 µl of a precipitant containing 0.085 *M* Tris–HCl pH 8.5, 0.17 *M* sodium acetate trihydrate, 25.5%(*w*/*v*) polyethylene glycol 4000 and 15%(*v*/*v*) glycerol (condition No. 22, Hampton Research Crystal Screen cryo-kit). Long thin needle-shaped crystals (of dimensions 0.01 × 0.01 × 0.5 mm) were grown overnight. In an effort to slow the crystal growth and to attempt to grow larger crystals, crystal trays were set up and placed at 277 K. Unfortunately, this did not yield better quality crystals. Hence, these initial crystals were microseeded in 2 µl of purified p37 and 5 µl of precipitant. This produced crystals of larger size (0.03 × 0.02 × 0.06 mm), but they were still very thin and difficult to mechanically manipulate (Fig. 2*a*). A single crystal was placed on a thin nylon crystal-mounting loop (Hampton Research) and flash-frozen (Oxford Cryosystems) for data collection at 105 K. X-ray diffraction data were collected using an R-AXIS IV++ image-plate system equipped with Osmic mirrors and a Rigaku HU-H3R CU rotating-anode generator operating at 50 kV and 100 mA. All data were collected using an

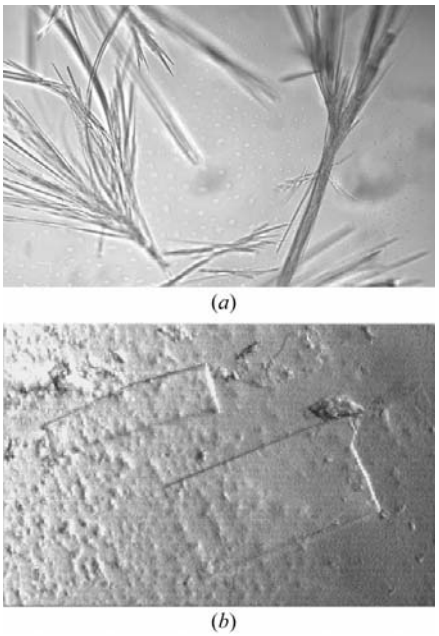

Figure 2
Crystals of p37. (*a*) Thin needle-like crystals obtained from initial crystallization screening. (*b*) Plate-like crystals obtained from refined crystallization conditions. Image (*a*) was taken with a Zeiss Axioplan 2 microscope under polarized light. Image (*b*) was taken with a Bio-Rad 1024 ES confocal microscope with an Olympus IX 70 transmission.

exposure time of 5 min per image and an oscillation range of 1.0°. The 2θ angle was fixed at 0° and the crystal-to-detector distance was fixed at 100 mm. Only 60 images were collected owing to X-ray decay that caused sequential worsening of the diffraction data. Of these images, only the first 40 were indexed and scaled, using *DENZO* and *SCALEPACK* (Otwinowski, 1992), as primitive orthorhombic (Laue group *P*222), with unit-cell parameters a = 52.65, b = 66.73, c = 119.72 Å. The data were of poor quality, diffracting to 3.5 Å resolution, and scaled with an R_{sym} of 0.213 (0.552 in the outer resolution shell) for 1704 unique reflections (27.4% completeness). Because of this lack of data, an unambiguous space-group assignment could not be made. Hence, because of the experimental difficulties associated with this crystal form of p37 further studies were halted.

In an attempt to obtain more useful crystals, a 600 µl sample of purified p37 (10 mg ml^{-1} in 50 m*M* Tris–HCl pH 7.5) was sent to the Hauptman–Woodward Institute for automated crystallization screening using the oil-immersion microbatch method (Chayen *et al.*, 1990). This method uses fractions of microlitre quantities and surveys 1536 conditions over several weeks. These results gave a lead condition which was scaled up to be a practical procedure in the laboratory. This was first achieved by placing 2.0 µl of the purified p37 sample in 2.0 µl of a solution containing 40% polyethylene glycol 4000 (PEG 3350), 0.1 *M* ammonium bromide, 0.1 *M* citrate pH 4.0 as the precipitant under 1 ml paraffin oil at 310 K in a 24-well Linbro plate (Hampton Research). After a short time interval (less than 48 h), clear plate-like crystals had grown. The condition volumes were increased and variation of the protein-to-precipitant ratio finally yielded optimal crystallization conditions at 7 µl protein and 3 µl precipitant. This produced crystals of dimensions 0.5 × 0.3 × 0.2 mm (Fig. 2*b*).

A single crystal taken from its well using a nylon crystal-mounting loop (Hampton Research) was flash-frozen prior to data collection at 105 K. X-ray diffraction data were collected at CHESS F-2 beamline (λ = 0.9169 Å) using the ADSC Quantum 4 CCD detector system. The crystal-to-detector distance used was 150 mm and the 2θ angle was fixed at 0°, with a 1° oscillation angle and 30 s exposure time for each frame. 212 images were collected to a maximum useful resolution of 1.8 Å (Fig. 3). The data were indexed and scaled using *DENZO* and *SCALEPACK* (Otwinowski, 1992), with the assigned Laue group *P*2 and unit-cell parameters a = 50.02, b = 67.26, c = 59.89 Å, β = 108.29°. The data set contained 34 882 unique reflections (99.8% complete, 99.8% in the outer resolution shell), resulting in an R_{sym} of 0.076 (0.256 in the outer resolution shell) (Table 1). By inspection of the intensities of the 0k0 reflections, we can infer the assignment of a twofold screw axis along the b axis that implies the space group to be $P2_1$. From the unit-cell volume and the molecular weight of p37, a V_M value (Matthews, 1968) of 2.608 Å^3 Da^{-1} was calculated for one molecule per asymmetric unit (with a solvent fraction of 0.529 assuming a partial specific volume of 0.74 cm^3 g^{-1}) using *CNS* version 1.1 (Brünger *et al.*, 1998). Table 2 gives a full summary of the data-collection statistics.

Although an initial attempt at molecular-replacement phasing has been made using the partial homology model of p37 (Boehlein *et al.*, 2002) as a search model, no clear solutions were found from rotation- and translation-function searches using *CNS* version 1.1 (Brünger *et al.*, 1998). This failure is most likely because the model is only 'at best' representing the N-terminal domain of p37. Although more efforts will be made to use the molecular-replacement method, we have also started to use traditional heavy-atom (Patterson, 1963) and halide-soak techniques (Dauter *et al.*, 2000) to phase the data set.

Table 1
Statistics of data collection.

Resolution shells (Å)	No. unique reflections	Completeness (%)	R_{sym}†
50.00–3.88	3556	99.5	0.057
3.88–3.08	3521	99.9	0.055
3.08–2.69	3496	99.9	0.061
2.69–2.44	3513	99.9	0.076
2.44–2.27	3461	99.9	0.088
2.27–2.13	3484	99.8	0.097
2.13–2.03	3456	99.9	0.117
2.03–1.94	3478	99.8	0.181
1.94–1.86	3444	99.7	0.194
1.86–1.80	3473	99.8	0.256
Total	34882	99.8	0.076

† $R_{sym} = \sum[\mathrm{ABS}(I - \langle I \rangle)]/\sum(I)$, where I is the intensity of an individual reflection and $\langle I \rangle$ is the average intensity for this reflection; the summation is over all intensities.

Table 2
Summary of data statistics.

Values in parentheses are for the outer resolution shell.

Temperature (K)	105
Resolution range (Å)	20–1.8
Space group	$P2_1$
Unit-cell parameters (Å)	a = 50.02, b = 67.26, c = 59.89, β = 108.29°
Wavelength (Å)	0.9169
Solvent content (%)	52.9
Crystal dimensions (mm)	0.5 × 0.3 × 0.2
V_M (Å^3 Da^{-1})	2.61
Total No. of reflections	247158
No. unique reflections	34882
R_{sym}	0.076 (0.256)
Completeness (%)	99.8 (99.8)
Average $I/\sigma(I)$	5.8

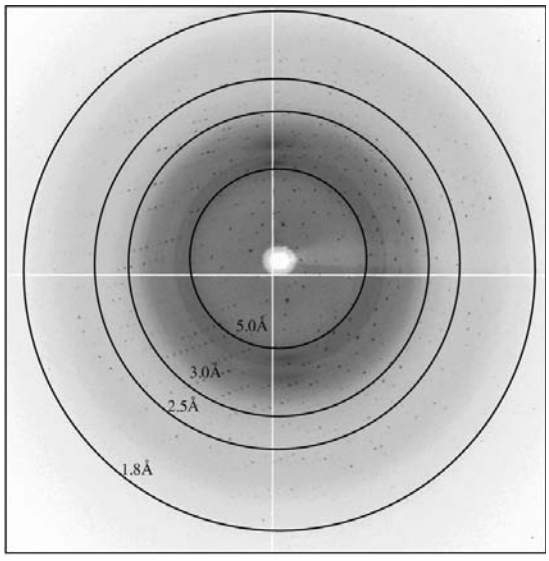

Figure 3
X-ray diffraction pattern of p37 collected on the F2 beamline (λ = 0.9169 Å) with the ADSC Quantum 4 CCD detector system at CHESS. These data were collected with a crystal-to-detector distance of 150 mm, an exposure time of 30 s per image and an oscillation range of 1.0° at 105 K. Concentric rings show the demarcation of 1.8, 2.5, 3.0 and 5.0 Å resolution shells.

The authors wish to thank Tim Vaught for taking optical photographs of our crystals. We also thank Jennifer Wolfley and George DeTitta at the Hauptman–Woodward Institute for crystal screening. We would also like to extend our gratitude to the staff at CHESS, especially Art Weaver and David Schuller, for their invaluable help in data collection. This work was supported by the University of Florida Research Opportunity Fund (SS) and a University of Florida Alumni Fellowship (RR).

References

Altschul, S. F., Madden, T. L., Schaffer, A. A., Zhang, J., Zhang, Z., Miller, W. & Lipman, D. J. (1997). *Nucleic Acids Res.* **25**, 3389–3402.
Boehlein, S. K., Brenes, R. B., Sheng, S., Reutzel, R., McKenna, R., Agbandje-McKenna, M. & Schuster, S. M. (2002). Submitted.

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

Chayen, N. E., Shaw Stewart, P. D., Maeder, D. L. & Blow, D. M. (1990). *J. Appl. Cryst.* **23**, 297–302.

Dauter, Z., Dauter, M. & Rajashankar, K. R. (2000). *Acta Cryst.* D**56**, 232–237.

Dudler, R., Schmidhauser, C., Parish, R. W., Wettenhall, R. E. H. & Schmidt, T. A. (1988). *EMBO J.* **7**, 3971–3974.

Fareed, G. C., Mendiaz, E., Sen, A., Juilare, G. J. F., Weisenburger, T. H. & Totanes, T. J. (1988). *Biol. Res. Mod.* **7**, 11–23.

Gilson, E., Alloing, G., Schmidt, T., Claverys, J.-P., Dudler, R. & Hofnung, M. (1988). *EMBO J.* **7**, 3971–3974,

Huang, S., Li, J. Y., Wu, J., Meng, L. & Shou, C. C. (2001).*World J. Gastroenterol.* **7**, 266–269.

Ilantzis, C., Thomson, D. M. P., Michaelidou, A., Benchimol, S. & Stanners, C. P. (1993). *Microbiol. Immunol.* **37**, 119–128.

Jia, Z., Hasnain, S., Hirama, T., Lee, X., Mort, J. S., To, R. & Huber, C. P. (1995). *J. Biol Chem.* **270**, 5527–5533.

McPherson, A. (1982). *Preparation and Analysis of Protein Crystals.* New York: Wiley & Sons.

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

Otwinowski, Z. (1992). *An Oscillation Data Processing Suite for Macromolecular Crystallography*, Yale University, New Haven, CT, USA.

Patterson, A. L. (1963). *Acta Cryst.* **16**, 1255–1256.

Sasaki, H., Igaki, H., Ishizuku, T., Kogoma, Y., Sugimura, T. & Terada, M. (1995). *Jpn J. Cancer Res.* **86**, 791–794.

Sauter, N. K., Hanson, J. E., Glick, G. D., Brown, J. H., Crowther, R. L., Park, S. J., Skehel, J. J. & Wiley, D. C. (1992). *Biochemistry*, **31**, 9609–9621.

Steinemann, C., Fenner, M., Binz, H. & Parish, R. W. (1984). *Proc. Natl Acad. Sci. USA*, **81**, 3747–3750.