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

Cloning, expression, purification, crystallization and preliminary X-ray analysis of the 31 kDa *Vibrio cholerae* heat-shock protein VcHsp31

Samir Das,[a] Sanjay Dey,[b] Trina Roy[c] and Udayaditya Sen[d]*

[a]Structural Genomics Division, Saha Institute of Nuclear Physics, 1/AF Bidhan Nagar, Kolkata, West Bengal 700 064, India, [b]Department of Bioinformatics, SRM University, Ramapuram, Chennai 600 089, India, [c]PG Department of Biotechnology, St Xavier's College, 30 Mother Teresa Sarani (Park Street), Kolkata, West Bengal 700 016, India, and [d]Crystallography and Molecular Biology Division, Saha Institute of Nuclear Physics, 1/AF Bidhan Nagar, Kolkata, West Bengal 700 064, India

Correspondence e-mail: udayaditya.sen@saha.ac.in

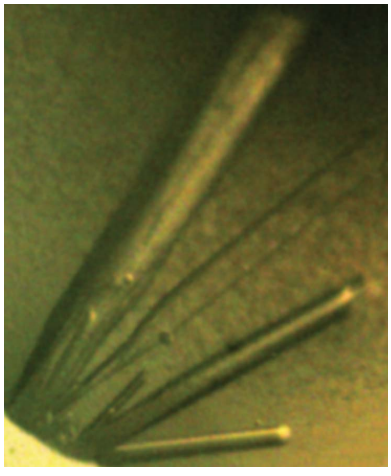
Received 8 July 2011
Accepted 14 August 2011

The Gram-negative bacterium *Vibrio cholerae*, which is responsible for the diarrhoeal disease cholera in humans, induces the expression of numerous heat-shock genes. VcHsp31 is a 31 kDa putative heat-shock protein that belongs to the DJ-1/PfpI superfamily, functioning as both a chaperone and a protease. VcHsp31 has been cloned, overexpressed and purified by Ni^{2+}–NTA affinity chromatography followed by gel filtration. Crystals of VcHsp31 were grown in the presence of PEG 6000 and MPD; they belonged to space group $P2_1$ and diffracted to 1.9 Å resolution. Assuming the presence of six molecules in the asymmetric unit, the Matthews coefficient was estimated to be 1.97 $Å^3$ Da^{-1}, corresponding to a solvent content of 37.4%.

1. Introduction

Environmental stresses on living cells, such as elevated temperature and various others, initiate the formation of misfolded proteins which may lead to malfunction of the cellular machinery. For the survival of cells, these misfolded proteins need to be either refolded into their active form or degraded to their constituent amino acids through a special set of damage-control proteins, which are communally known as heat-shock proteins (Hsps). Functionally, Hsps are molecular chaperones and proteases which play a pivotal role in cellular homeostasis either by assisting the correct folding of newly synthesized proteins and preventing their aggregation or by hydrolyzing those proteins which have become permanently misfolded, respectively (Quigley *et al.*, 2003). Expression of these heat-shock genes is controlled by specific proteins known as heat-shock transcription factors (HSFs), which undergo inactive-state to active-state conversion upon exposure to stressful conditions (Wu, 1995). Hsps generally work in four mechanistic ways: (i) by holding the nascent polypeptide up to a transition state which is capable of folding when released into the cytosol, (ii) by providing a confined region inside which folding can take place, (iii) by regulating protein degradation and (iv) by assisting protein translocation into different cellular compartments (Hartl & Hayer-Hartl, 2002; Arya *et al.*, 2007). Hydrophobic patches present in Hsps interact with solvent-exposed hydrophobic segments of misfolded proteins and control their accumulation (Lee *et al.*, 2003). The quality of proteins at the cellular level is maintained by Hsps in either the presence or the absence of ATP; for example, in *Escherichia coli* DnaK/Hsp70 acts as an ATP-dependent protease, whereas Hsp31 acts independently of any ATP/ADP cycle (Malki *et al.*, 2003). Surprisingly, the chaperone activity of Hsp31 is suppressed in the presence of ATP at elevated temperatures (Sastry *et al.*, 2002). Both induced and constitutively expressed Hsps promote cell survival in mammalian systems by reducing or inhibiting apoptosis.

The Gram-negative noninvasive enteric bacterium *Vibrio cholerae*, a pathogenic organism that is responsible for diarrhoea in various parts of the world, switches from a quiescent marine environment to a highly hostile milieu on entering the human intestine (JeevanJyot & Ghosh, 1995). It has been found that the low-pH, anaerobic conditions and the temperature upshift from 295 to 310 K on entering the human body cause amplified production of cholera toxin (Sahu *et al.*,



1994). This hostile environment triggers the induction of various heat-shock genes in *V. cholerae* which facilitate the breakdown or renaturation of misfolded proteins (Parsot & Mekalanos, 1990). The A5F0Q0 gene product of *V. cholerae* O395 is a 31 kDa hypothetical intracellular protease/amidase (hereafter referred to as VcHsp31) that shares 61% sequence identity with the *E. coli* chaperone Hsp31 (EcHsp31). A PFAM search showed that both EcHsp31 and VcHsp31 belong to the DJ-1/PfpI (PfpI, *Pyrococcus furiosus* protease I) superfamily. DJ-1, which is associated with autosomal recessive early onset of Parkinson's disease, is an intracellular Cys protease which is structurally related to the class I glutamine amidotransferase (GAT) superfamily. It catalyzes the removal of the amine group from glutamine and transfers it to a substrate to form a new carbon–nitrogen bond. All members of the DJ-1 superfamily are oligomers in nature, ranging from dimers to dodecamers (Fioravanti *et al.*, 2008). EcHsp31, which is organized as a homodimer, is a weak amidopeptidase towards a large number of protein substrates (Sastry *et al.*, 2009). EcHsp31 consists of 13 β-strands, 12 α-helices and various loops forming two α–β domains designated the A and P domains, while the putative catalytic triad consists of Cys-His-Asp residues. Structure-based sequence alignment has shown that Hsp31 homologues can be subdivided into three groups depending on (i) the presence of the P domain, (ii) the oligomerization state of the protein and (iii) the presence of Glu/Asp in the catalytic triad. EcHsp31 contains a complete P domain, an aspartate residue in the catalytic triad and an extended 22-residue linker region which regulates the accessibility of the active site (Quigley *et al.*, 2003). The dimerization interface of EcHsp31 forms a hydrophobic bowl that serves as a binding site for partially folded species. In contrast, the hydrophobic patches of human DJ-1 are located in a shallow groove formed mainly by two C-terminal helices (Lee *et al.*, 2003). The catalytic Cys residue of stress-response protein DR1199 from *Deinococcus radiodurans*, like other members of the DJ-1 superfamily, has been observed in an oxidized sulfenic acid form (Fioravanti *et al.*, 2008). The structures of EcHsp31 determined at two different temperatures (295 and 287 K) showed significant conformational variability in some crucial regions and these regions have been implicated in chaperone activity (Quigley *et al.*, 2004).

To understand the function of VcHsp31 at the molecular level, we have crystallized it using the hanging-drop vapour-diffusion method. Here, we report the cloning, expression, purification, crystallization and X-ray diffraction of VcHsp31.

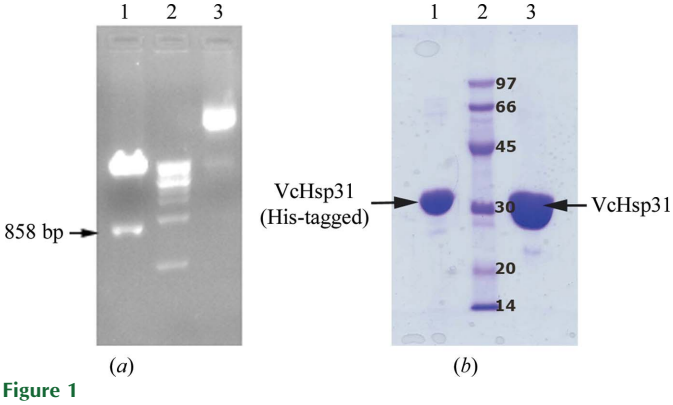


Figure 1
(*a*) Confirmation of the VcHsp31 clone by restriction digestion with *Nde*I and *Bam*HI, showing a band at 858 bp. Lane 1, VcHsp31 clone; lane 2, 1 kb DNA ladder; lane 3, pET28a$^+$-VcHsp31 construct. (*b*) The homogeneity of the purified VcHsp31 was checked by 12% SDS–PAGE. 6×His-tagged VcHsp31 is shown in lane 1; lane 2 contains molecular-weight markers (labelled in kDa) and lane 3 contains VcHsp31 without a histidine tag.

2. Materials and methods

2.1. Cloning, expression and purification of VcHsp31

The *VcHsp31* gene was cloned into kanamycin-resistant pET28a$^+$ (Novagen) vector using specific primers (forward, 5′-**TGGACT-G**CATATGATGACGACAGTAATGAATGATAAAC-3′; reverse. 5′-**GTA**GGATCCTTATTTCTGCTCAACTTCAGCAAGTAAT-GCC-3′). The primers were synthesized (NeuProCell) with adaptors (bold) for the restriction enzymes *Nde*I and *Bam*HI (underlined). Chromosomal DNA of *V. cholerae* strain O395 was used as a template to amplify the region encoding VcHsp31. Both the 858 bp *VcHsp31* PCR amplicon and the pET28a$^+$ vector were double-digested with the restriction enzymes *Nde*I and *Bam*HI and were purified from 1.0% agarose gel using a gel-extraction kit (Qiagen). The two double-digested DNA fragments were then ligated. The clones were selected appropriately using *E. coli* XL1-Blue cells with kanamycin resistance (Fig. 1*a*). VcHsp31 protein was finally over-expressed in *E. coli* BL21 (DE3) cells in the presence of kanamycin as a fusion protein with a 6×His tag at the N-terminus followed by a thrombin cleavage site.

For overexpression, a single colony was picked, transferred into 100 ml LB broth and grown overnight. 1 l LB broth was inoculated with 10 ml overnight culture and the culture was grown at 310 K until the OD_{600} reached 0.6; the cells were then transferred to 293 K. The cells were induced with 0.1 m*M* IPTG at 293 K overnight. After induction, the cells were harvested at 4500*g* for 20 min and the pellet

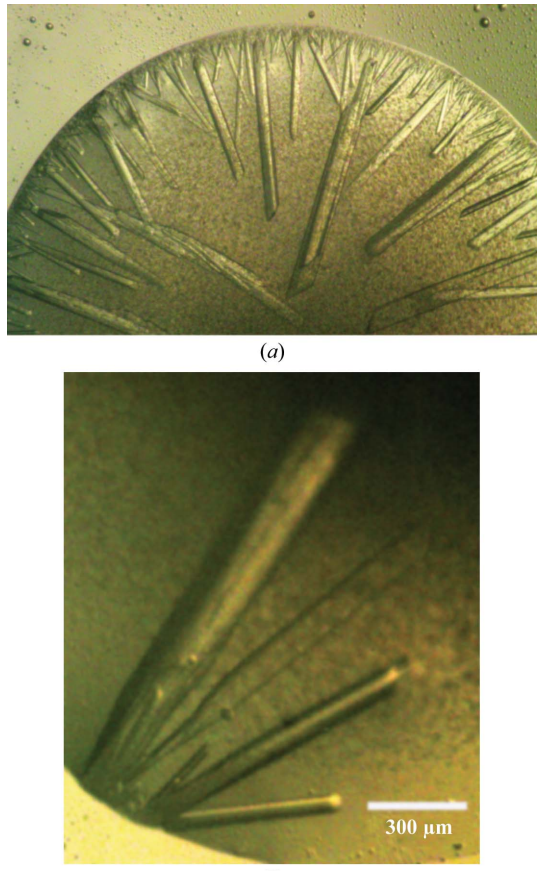


Figure 2
Crystals of VcHsp31 grown in the presence of 5% PEG 6000 at pH 5.0. (*a*) Nucleation starts from the periphery of the drops; the crystals diffracted poorly. (*b*) Large crystals (1.2 × 0.2 × 0.2 mm) appeared when 5% PEG 6000 + 8% MPD at pH 5.0 was used as the precipitant.

Table 1
Data-collection and data-processing parameters for the VcHsp31 crystal.

Values in parentheses are for the outermost resolution shell.

Space group	$P2_1$
Unit-cell parameters (Å, °)	a = 73.514, b = 79.094, c = 133.132, β = 95.216
Oscillation range (°)	1
Maximum resolution (Å)	1.90 (1.93–1.90)
No. of molecules in asymmetric unit	6
Matthews coefficient (V_M) (Å^3 Da^{-1})	1.97
Solvent content (%)	37.43
No. of unique reflections	113580
Multiplicity	1.79
Mosaicity (°)	0.36
Completeness (%)	94.8 (91.0)
R_{merge}† (%)	5.19 (28.0)
Average $I/\sigma(I)$	6.9 (2.0)

† $R_{merge} = \sum_{hkl}\sum_i |I_i(hkl) - \langle I(hkl)\rangle| / \sum_{hkl}\sum_i I_i(hkl)$, where $I_i(hkl)$ is the observed intensity of the ith measurement of reflection hkl and $\langle I(hkl)\rangle$ is the mean intensity of reflection hkl calculated after scaling.

was resuspended in 25 ml ice-cold lysis buffer consisting of 50 m*M* Tris–HCl pH 8.0 and 300 m*M* NaCl. PMSF (final concentration of 1 m*M*) and 10 mg lysozyme were added to the resuspended solution and it was lysed by sonication on ice. The cell lysate was then centrifuged (12 000*g* for 50 min) at 277 K and 6×His-tagged VcHsp31 was isolated from the supernatant using Ni^{2+}–NTA-based immobilized metal-ion affinity chromatography (Qiagen). The 6×His-tagged VcHsp31 protein was eluted with lysis buffer containing 150 m*M* imidazole (Sigma). The purity of the eluted fraction was checked by 12% SDS–PAGE (Fig. 1*b*) and was found to be nearly homogeneous. The eluted protein was dialyzed overnight against thrombin cleavage buffer (20 m*M* Tris–HCl pH 8.0, 150 m*M* NaCl, 2.5 m*M* $CaCl_2$), concentrated using an Amicon Ultra centrifugation unit (molecular-weight cutoff 10 000 Da) and the 6×His tag was cleaved with 1 U thrombin by overnight incubation at 277 K. The protein was further purified by gel filtration using a Sephacryl S-100 (GE Healthcare) column (78 × 1.4 cm) pre-equilibrated with thrombin cleavage buffer containing 0.02% sodium azide at 277 K. The homogeneity of the purified protein was checked on 12% SDS–PAGE (Fig. 1*b*).

2.2. Crystallization

For crystallization, thrombin-cleaved VcHsp31 in buffer consisting of 20 m*M* Tris–HCl pH 8.0, 150 m*M* NaCl, 2.5 m*M* $CaCl_2$ was concentrated to 12 mg ml^{-1} using an Amicon Ultra centrifugation unit (molecular-weight cutoff 10 000 Da) in the presence of 4 m*M* DTT. Crystallization was performed by the hanging-drop vapour-diffusion method using 24-well crystallization trays (Hampton Research, Laguna Niguel, California, USA). Grid Screen Ammonium Sulfate, Grid Screen PEG 6000, Crystal Screen and Crystal Screen 2 from Hampton Research (Jancarik & Kim, 1991) were used to explore the initial crystallization conditions. Typically, 2 µl protein solution (12 mg ml^{-1}) was mixed with an equal volume of the screening solution and equilibrated over 700 µl of the latter as the reservoir solution. Initially, poor crystals were observed which grew from the periphery of drops in the presence of 5% PEG 6000 pH 5.0 incubated at both 277 and 293 K (Fig. 2*a*). This promising crystallization condition was further optimized in the presence of different concentrations of various additives such as glycerol, dioxane, 2-methyl-1,3-propanediol (MPD) and ethylene glycol. Among them, the best crystals (1.2 × 0.2 × 0.2 mm; Fig. 2*b*) were obtained when 4 µl protein solution was mixed with 2 µl precipitant solution consisting of 5% PEG 6000 in 0.1 *M* citrate pH 5.0 and 8% MPD and equilibrated against a reservoir solution consisting of 15% PEG 6000 in 0.1 *M* Tris–HCl buffer pH 8.5 and 5% MPD at 293 K for 4–5 d. The rod-shaped crystals thus obtained diffracted to 1.9 Å resolution.

2.3. Data collection and processing

Crystals of VcHsp31 were fished out from the crystallization drops using a 10 µm nylon loop and flash-cooled in a stream of nitrogen (Oxford Cryosystems) at 100 K. A diffraction data set was collected using an in-house MAR Research image-plate detector of diameter 345 mm and Cu $K\alpha$ radiation generated by a Bruker–Nonius FR591 rotating-anode generator equipped with Osmic MaxFlux confocal optics and operated at 50 kV and 65 mA. Data were processed and scaled using *AUTOMAR* (http://www.marresearch.com/automar/run.html). Data-collection and processing statistics are given in Table 1.

3. Results

The crystals grown in the absence of MPD at 293 K diffracted to only 3.4 Å resolution, whereas the crystals grown in the presence of 5% PEG 6000 and 8% MPD were of very good quality and diffracted to 1.9 Å resolution. However, both crystals belonged to the primitive monoclinic space group $P2_1$, with unit-cell parameters a = 73.514, b = 79.094, c = 133.132 Å, β = 95.216°. Packing considerations, based on the molecular weight of 32 kDa, indicated the presence of six molecules in the asymmetric unit, corresponding to a Matthews coefficient V_M (Matthews, 1968) of 1.97 Å^3 Da^{-1} and a solvent content of 37.4%. As the crystals grown in the absence of MPD diffracted poorly with spot streaking, data collection was not pursued from these crystals. A total of 87 frames were collected from a crystal grown in the presence of PEG 6000 and 8% MPD with a crystal-to-detector distance of 150 mm. The exposure time for each image was 4 min and the oscillation range per image was maintained at 1°. A

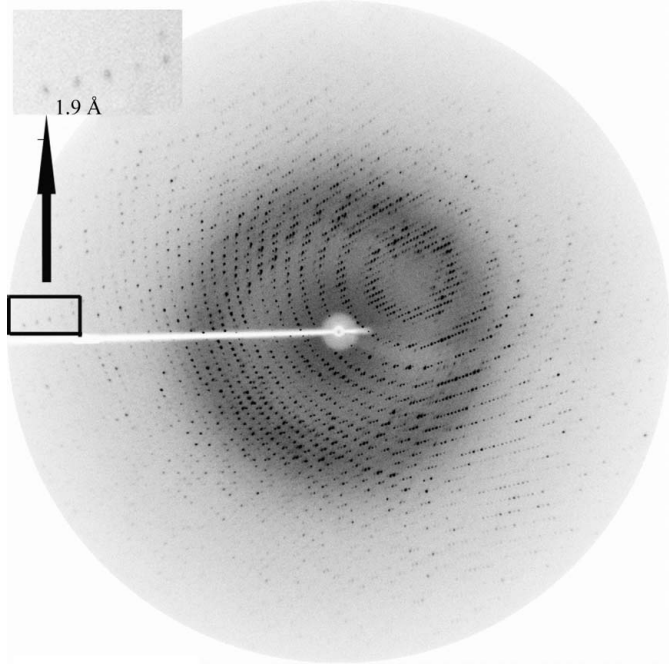


Figure 3
X-ray diffraction image of the VcHsp31 crystal; the edge of the detector, which corresponds to a resolution of 1.9 Å, is enlarged to show the quality of the reflections in the highest bin.

data set with an overall completeness of 94.8% was obtained to 1.9 Å resolution (Table 1, Fig. 3) with an R_{merge} of 5.19%.

A *BLAST* (Altschul *et al.* 1990) search for a homologous structure showed that Hsp31 from *E. coli* (EcHsp31) has the highest identity (61%) to VcHsp31. Therefore, the coordinates of EcHsp31 (PDB entry 1n57; Quigley *et al.*, 2003) were used as a search model in molecular-replacement (MR) calculations using *MOLREP* from the *CCP*4 package (Winn *et al.*, 2011). Before proceeding with MR calculations, waters were deleted from the coordinates of the search model. Additionally, mismatched residues and Met residues (present as selenomethionine) were truncated to alanine. Six molecules of this modified model in the asymmetric unit produced a correlation coefficient of 61% with an *R* factor of 38.1% using data between 30.0 and 4.0 Å resolution. Rigid-body refinement followed by a few cycles of NCS restrained positional refinement with *CNS* (Brünger *et al.*, 1998), using data between 30 and 2.8 Å resolution, led to an *R* factor of 40.8% (R_{free} = 41.4%). A sixfold NCS-averaged electron-density map calculated using this model showed an unambiguous polypeptide chain. Refinement and model building is in progress.

References

Altschul, S. F., Gish, W., Miller, W., Myers, E. W. & Lipman, D. J. (1990). *J. Mol. Biol.* **215**, 403–410.

Arya, R., Mallik, M. & Lakhotia, S. C. (2007). *J. Biosci.* **32**, 595–610.

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.

Fioravanti, E., Durá, M. A., Lascoux, D., Micossi, E., Franzetti, B. & McSweeney, S. (2008). *Biochemistry*, **47**, 11581–11589.

Hartl, F. U. & Hayer-Hartl, M. (2002). *Science*, **295**, 1852–1858.

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

JeevanJyot & Ghosh, A. (1995). *Microbiology*, **141**, 2101–2109.

Lee, S.-J., Kim, S. J., Kim, I.-K., Ko, J., Jeong, C.-S., Kim, G.-H., Park, C., Kang, S.-O., Suh, P.-G., Lee, H.-S. & Cha, S.-S. (2003). *J. Biol. Chem.* **278**, 44552–44559.

Malki, A., Kern, R., Abdallah, J. & Richarme, G. (2003). *Biochem. Biophys. Res. Commun.* **301**, 430–436.

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

Parsot, C. & Mekalanos, J. J. (1990). *Proc. Natl Acad. Sci. USA*, **87**, 9898–9902.

Quigley, P. M., Korotkov, K., Baneyx, F. & Hol, W. G. (2003). *Proc. Natl Acad. Sci. USA*, **100**, 3137–3142.

Quigley, P. M., Korotkov, K., Baneyx, F. & Hol, W. G. (2004). *Protein Sci.* **13**, 269–277.

Sahu, G. K., Chowdhury, R. & Das, J. (1994). *Infect. Immun.* **62**, 5624–5631.

Sastry, M. S., Korotkov, K., Brodsky, Y. & Baneyx, F. (2002). *J. Biol. Chem.* **277**, 46026–46034.

Sastry, M. S., Zhou, W. & Baneyx, F. (2009). *Protein Sci.*, **18**, 1439–1447.

Winn, M. D. *et al.* (2011). *Acta Cryst.* D**67**, 235–242.

Wu, C. (1995). *Annu. Rev. Cell Dev. Biol.* **11**, 441–469.