STRUCTURAL BIOLOGY COMMUNICATIONS

ISSN 2053-230X

Cloning, expression, purification, crystallization and X-ray crystallographic analysis of recombinant human C1ORF123 protein

Siti Nurulnabila A. Rahaman,[a] Jastina Mat Yusop,[a] Zeti-Azura Mohamed-Hussein,[a,b] Kok Lian Ho,[c] Aik-Hong Teh,[d] Jitka Waterman[e] and Chyan Leong Ng[a]*

[a]Institute of Systems Biology, Universiti Kebangsaan Malaysia, 43600 UKM Bangi, Selangor, Malaysia, [b]School of Biosciences and Biotechnology, Faculty of Science and Technology, Universiti Kebangsaan Malaysia, 43600 UKM Bangi, Selangor, Malaysia, [c]Department of Pathology, Faculty of Medicine and Health Sciences, Universiti Putra Malaysia, 43400 UPM Serdang, Selangor, Malaysia, [d]Centre for Chemical Biology, Universiti Sains Malaysia, 11900 Bayan Lepas, Penang, Malaysia, and [e]Diamond Light Source, Harwell Science and Innovation Campus, Didcot OX11 0DE, England. *Correspondence e-mail: clng@ukm.edu.my

Received 2 December 2015
Accepted 2 February 2016

Edited by R. Sankaranarayanan, Centre for Cellular and Molecular Biology, Hyderabad, India

Keywords: C1ORF123; hypothetical protein; DUF866; polycystic ovary syndrome; bioinformatic analysis.

C1ORF123 is a human hypothetical protein found in open reading frame 123 of chromosome 1. The protein belongs to the DUF866 protein family comprising eukaryote-conserved proteins with unknown function. Recent proteomic and bioinformatic analyses identified the presence of C1ORF123 in brain, frontal cortex and synapses, as well as its involvement in endocrine function and polycystic ovary syndrome (PCOS), indicating the importance of its biological role. In order to provide a better understanding of the biological function of the human C1ORF123 protein, the characterization and analysis of recombinant C1ORF123 (rC1ORF123), including overexpression and purification, verification by mass spectrometry and a Western blot using anti-C1ORF123 antibodies, crystallization and X-ray diffraction analysis of the protein crystals, are reported here. The rC1ORF123 protein was crystallized by the hanging-drop vapor-diffusion method with a reservoir solution comprised of 20% PEG 3350, 0.2 *M* magnesium chloride hexahydrate, 0.1 *M* sodium citrate pH 6.5. The crystals diffracted to 1.9 Å resolution and belonged to an orthorhombic space group with unit-cell parameters a = 59.32, b = 65.35, c = 95.05 Å. The calculated Matthews coefficient (V_M) value of 2.27 Å^3 Da^{-1} suggests that there are two molecules per asymmetric unit, with an estimated solvent content of 45.7%.

1. Introduction

Open reading frame 123, which is located in the short arm of human chromosome 1, encodes a hypothetical protein known as C1ORF123 (Selvarajan & Shanmughavel, 2014). C1ORF123 consists of 160 amino acids with a calculated molecular weight of approximately 18 kDa. The C1ORF123 protein is exclusively found in eukaryotic cells and belongs to the DUF866 family of proteins of unknown function. To date, the only known protein structure from the DUF866 family is that of the *Plasmodium falciparum* homologue MAL13P1.257, which shares 26% sequence identity with C1ORF123 (Holmes *et al.*, 2006). No functional studies have yet been reported for MAL13P1.257.

In humans, C1ORF123 is expressed in various anatomical regions, including the brain, skeletal muscles and ovary. Bioinformatics analysis has annotated the C1ORF123 protein as a cellular protein (Selvarajan & Shanmughavel, 2014) and suggests that it is involved in the pathway related to an

Table 1
Identified interacting partners of C1ORF123 and its homologues from the STRING database (version 9.1).

No.	Protein	Description	Experimental evidence	Reference
1	RHBDL1	Rhomboid, veinlet-like 1	Two-hybrid pooling assay	Giot *et al.* (2003)
2	TMBIM6	Transmembrane BAX inhibitor motif-containing protein 6; suppressor of apoptosis	Two-hybrid pooling assay	Giot *et al.* (2003)
3	CDKN1A	Cyclin-dependent kinase inhibitor 1A; cell-cycle regulator	Two-hybrid pooling assay	Stelzl *et al.* (2005)
4	SRP14	Signal recognition particle 14 kDa	Tandem affinity assay	Krogan *et al.* (2006)
5	SRP19	Signal recognition particle 19 kDa	Tandem affinity assay	Krogan *et al.* (2006)
6	NPUK68	UKp68-like protein; binds polyadenosine RNA oligonucleotides	Affinity capture-RNA assay	Batisse *et al.* (2009)
7	ZC3H14	Zinc-finger CCCH domain-containing protein 14; binds polyadenosine RNA oligonucleotides	Affinity capture-RNA assay	Batisse *et al.* (2009)
8	ANXA1	Annexin 1; calcium/phospholipid-binding protein which promotes membrane fusion and is involved in exocytosis; also regulates phospholipase A_2 activity	Two-hybrid assay	Vinayagam *et al.* (2011)
9	UBB	Ubiquitin B; protein degradation, maintenance of chromatin structure, regulation of gene expression, stress response, ribosome biogenesis and DNA repair	Affinity capture-MS assay	Danielsen *et al.* (2011)
1.	UBC	Ubiquitin C; protein degradation, maintenance of chromatin structure, regulation of gene expression, stress response, ribosome biogenesis and DNA repair	Affinity capture-MS assay	Danielsen *et al.* (2011)
11	LARP1 and LARP1B	La ribonucleoprotein domain family	Affinity capture-RNA assay	Schenk *et al.* (2012)
12	SSB	Sjögren syndrome antigen B (autoantigen La); binds to the 3′ poly(U) termini of nascent RNA polymerase III transcripts	Affinity capture-RNA assay	Schenk *et al.* (2012)

Table 2
Macromolecule-production information.

Source organism	*H. sapiens*
DNA source	Chemically synthesized (GenScript, USA)
Cloning vector	pUC57
Expression vector	pET-28b
Expression host	*E. coli* strain BL21 Rosetta-gami (DE3)
Complete amino-acid sequence of the construct†	MGSSHHHHHHSSGLVPRGSHMGKIALQLKATLEN-ITNLRPVGEDFRWYLKMKCGNCGEISDKWQYI-RLMDSVALKGGRGSASMVQKCLCARENSIEIL-SSIKPYNAEDNENFKTIVEFECRGLEPVDFQP-QAGFAVESGTAFSDINLQEKDWTDYDEKAQES-VGIYEVTHQFVKC

abnormality known as polycystic ovary syndrome (PCOS; Mohamed-Hussein & Harun, 2009). PCOS is a heterogeneous endocrine disorder that causes ∼10% of infertility in women (Diamanti-Kandarakis, 2008). Interestingly, a proteomic analysis of goat adipose tissue also identified the C1ORF123 homologue as one of the adipokines that may be involved in endocrine function (Restelli *et al.*, 2014). Other proteomics studies have found that the C1ORF123 protein is largely expressed in the hippocampus of people suffering from schizophrenia, bipolar disorder and methamphetamine-induced sensitization of the prefrontal cortex, as well as being a unique protein in the frontal cortex of aged rats associated with slow-wave sleep (SWS) (Schubert *et al.*, 2015; Wearne *et al.*, 2015; Vazquez *et al.*, 2009). This indicates the involvement of C1ORF123 in psychotic diseases or in age-related changes in brain function. A homologue of C1ORF123 has also been identified in the electric organ of the pacific electric ray *Torpedo californica* along with many neuromuscular junctions and presynaptic proteins, suggesting its role in synapse structure and maintenance (Mate *et al.*, 2011). C1ORF123 has also been identified as an O-GlcNAc transferase (OGT) interactor, indicating its possible role in the post-translational O-GlcNAcylation of proteins, which is important in many biological processes (Deng *et al.*, 2014). The network-based approach of the STRING database (Franceschini *et al.*, 2013) further deciphers the potential function of C1ORF123 and its homologues by the identification of interacting partners (Table 1). To better understand its biological function, we are working towards structural analysis of the human C1ORF123 protein. Here, we report the cloning, overexpression, purification, protein characterization and crystallization together with the initial X-ray crystallographic analysis of recombinant C1ORF123 (rC1ORF123).

2. Materials and methods

2.1. Protein production

The 492 bp coding sequence for human C1ORF123 (Gene ID 54987) was synthesized and cloned between the NdeI and HindIII restriction-endonuclease sites of the pUC57 cloning vector (GenScript, USA). The C1ORF123 gene was subcloned into the pET-28b vector using the same restriction enzymes to produce the pET-28b-C1ORF123 construct, which includes a 6×His fusion tag at the N-terminus of the recombinant protein. The calculated molecular weight of rC1ORF123, which contains 20 amino acids as a fusion tag at the N-terminus (Table 2), is approximately 20 kDa using *ProtParam* (Gasteiger *et al.*, 2005). Subsequently, the pET-28b-C1ORF123 construct was transformed into *Escherichia coli* strain BL21 Rosetta-gami (DE3) cells. A single colony of transformant was inoculated into 6 ml Luria–Bertani (LB) broth containing 50 mg ml^{-1} kanamycin and agitated

overnight in an incubator shaker at 250 rev min^{-1} and 310 K. The bacterial culture was then inoculated into 1 l LB broth supplemented with 50 mg ml^{-1} kanamycin and grown at 250 rev min^{-1} at 310 K. After the OD_{600} had reached 0.5–0.6, expression of the recombinant protein was induced by adding 1 m*M* isopropyl *β*-D-1-thiogalactopyranoside (IPTG). The culture was grown for a further 3–4 h at 310 K before the cells were harvested by centrifugation at 17 968*g*.

The bacterial pellet was resuspended in lysis buffer (10 ml per gram of cell pellet) consisting of 25 m*M* Tris–HCl pH 7.5, 100 m*M* NaCl, 20 m*M* *β*-mercaptoethanol, 20 m*M* imidazole before being lysed by sonication (Qsonica, 30 cycles of 38% amplitude for 30 s each). The cell lysate was centrifuged at 17 968*g* at 277 K for 30 min to separate the soluble proteins from the cell debris. The supernatant was filter-sterilized with a 0.22 µm PVDF membrane filter before loading it onto an Ni–NTA-coupled HisTrap HP 5 ml column (GE Healthcare) which had been pre-equilibrated with binding buffer consisting of 25 m*M* Tris–HCl pH 7.5, 100 m*M* NaCl, 20 m*M* *β*-mercaptoethanol, 50 m*M* imidazole. The rC1ORF123 protein was eluted using a linear gradient of washing buffer consisting of 25 m*M* Tris–HCl pH 7.5, 100 m*M* NaCl, 20 m*M* *β*-mercaptoethanol, 500 m*M* imidazole. The protein eluted at an imidazole concentration of 79 m*M*. Fractions containing the rC1ORF123 protein were pooled and concentrated using Vivaspin concentrators fitted with a 3 kDa molecular-weight cutoff filter (Sartorius, Germany). The concentrated rC1ORF123 protein was further purified by size-exclusion chromatography (SEC) using a HiLoad 16/600 Superdex 75 pg gel-filtration column (GE Healthcare, USA) pre-equilibrated with size-exclusion buffer consisting of 25 m*M* Tris–HCl pH 7.5, 100 m*M* NaCl, 20 m*M* *β*-mercaptoethanol. The purity of the rC1ORF123 protein was verified using 12% SDS–PAGE (Fig. 1). The protein concentration of rC1ORF123 was assessed using the Bradford assay (Bio-Rad, USA). Fractions containing the rC1ORF123 protein purified by size-exclusion chromatography were also pooled and concentrated to 8.0 mg ml^{-1}.

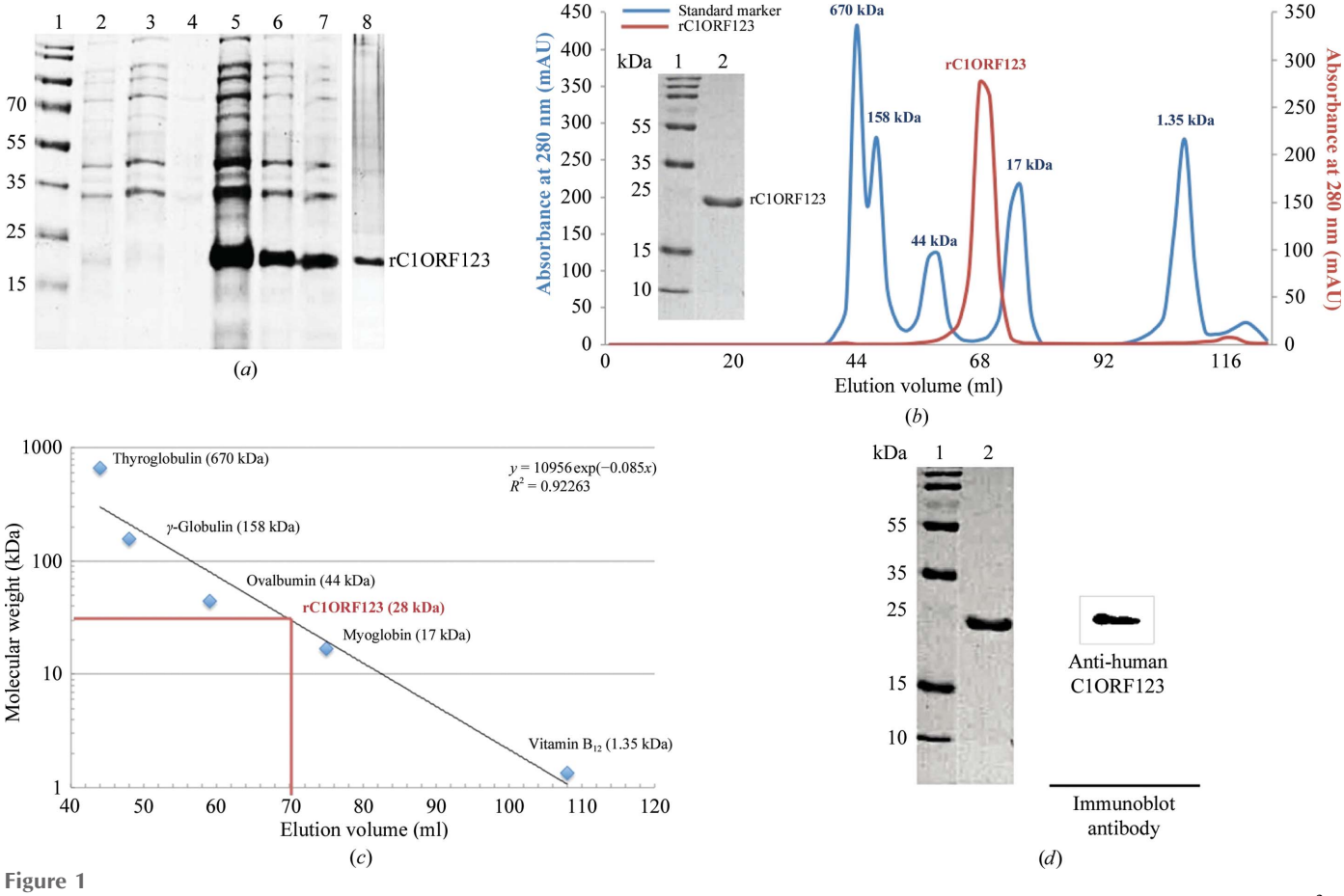

Figure 1

Overexpression, purification and identification using a Western blot of recombinant C1ORF123 protein. (*a*) Analysis of the overexpressed and Ni^{2+}–NTA affinity-purified rC1ORF123 protein using SDS–PAGE (12%). Lane 1, PageRuler prestained protein ladder (Thermo Scientific, USA; labelled in kDa); lanes 2, 3 and 4, total protein, pellet and supernatant of crude extract before IPTG induction, respectively, as a negative control; lanes 5, 6 and 7, total protein, pellet and supernatant of crude extract after IPTG induction, respectively; lane 8, rC1ORF123 purified using Ni^{2+}–NTA affinity chromatography. (*b*) Size-exclusion chromatography (SEC) results using a HiLoad 16/600 Superdex 75 pg gel-filtration column (GE Healthcare, UK). The rC1ORF123 protein eluted as a single peak between those for myoglobin (17 kDa) and ovalbumin (44 kDa), suggesting that the rC1ORF123 protein may be a monomer in solution. rC1ORF123 purified by SEC is shown on 12% SDS–PAGE, with the protein band corresponding to ∼20 kDa. (*c*) Calibration curve for size-exclusion chromatography on a Superdex 16/600 75 pg gel-filtration column indicating the molecular weight of rC1ORF123 to be ∼28 kDa. (*d*) The recombinant rC1ORF123 protein was positively detected by a Western blot using anti-human C1ORF123 antibodies (right panel); the band at ∼20 kDa corresponding to purified rC1ORF123 is shown (left panel).

Table 3
Crystallization information.

Method	Sitting-drop vapor diffusion (initial crystallization) and hanging-drop vapor diffusion (crystal optimization)
Plate type	96-well MRC plates (initial crystal screening) and 24-well plates (crystal optimization)
Temperature (K)	293
Protein concentration (mg ml^{-1})	8.0 (initial screening) and 7.8 (optimization)
Buffer composition of protein solution	25 m*M* Tris–HCl, 0.1 *M* sodium chloride tribasic pH 7.5, 20 m*M* β-mercapthoethanol
Composition of reservoir solution	0.2 *M* magnesium chloride hexahydrate, 0.1 *M* sodium citrate tribasic buffer pH 6.5, 20% polyethylene glycol 3350
Volume and ratio of drop	1 µl, 1:1 ratio of protein:reservoir solution (initial crystal screening); 2 µl, 1:1 ratio of protein: reservoir solution (crystal optimization)
Volume of reservoir	80 µl (initial crystallization) and 1.0 ml (crystal optimization)

2.2. Verification of recombinant C1ORF123 protein using a Western blot

A Western blot was performed according to the protocol described by Mahmood & Yang (2012). rC1ORF123 proteins were run on a 12% SDS–PAGE gel and transferred to a nitrocellulose membrane. After blocking with bovine serum albumin (BSA) for 1 h at room temperature, the membrane was probed with rabbit anti-human C1ORF123 antibodies (Sigma–Aldrich) for 2 h at room temperature followed by two 10 min PBST washes with agitation. The membrane was then incubated at room temperature for 1 h with a goat anti-rabbit IgG secondary antibody conjugated to horseradish peroxidase (HRP) (Sigma–Aldrich). The membrane was washed twice with PBST for 10 min and then soaked in SuperSignal West Pico chemiluminescent substrate (Thermo Scientific) for ∼5 min. The results were obtained using a chemiluminescent imaging system.

2.3. Mass-spectrometry

To confirm the sequence and the molecular weight of the expressed rC10RF123 protein, a single protein band with a molecular weight of approximately 20 kDa was excised from the 12% SDS–PAGE gel (Fig. 1*a*) and used for protein identification by mass spectrometry. Peptides obtained after trypsin digestion of rC1ORF123 were extracted and analyzed by matrix-assisted laser desorption/ionization time-of-flight/time-of-flight mass spectrometry (MALDI-TOF/TOF MS) using a 5800 Proteomics Analyzer mass spectrometer (Applied Biosystems/SCIEX; Bringans *et al.*, 2008) by Proteomics International Pty Ltd (Australia). Protein identification was carried out using the *Mascot* sequence-matching software (Matrix Science) based on the Ludwig NR database.

2.4. Crystallization

The rC1ORF123 protein was purified to homogeneity in a buffer consisting of 25 m*M* Tris–HCl pH 7.5, 100 m*M* NaCl, 20 m*M* β-mercaptoethanol. It was then concentrated to 8 mg ml^{-1} and used for initial crystallization screening with commercially available crystallization screening kits such as Index, Crystal Screen and Crystal Screen 2 (Hampton Research). Screening was performed using the sitting-drop vapor-diffusion method in standard 96-well MRC crystallization plates (Molecular Dimensions). Drops consisting of 0.5 µl rC1ORF123 protein and 0.5 µl reservoir solution were equilibrated against 80 µl reservoir solution at 293 K. Initial crystal hits were obtained from several crystallization conditions. The crystallization conditions were further optimized using the hanging-drop vapor-diffusion method in 24-well plates with crystallization drops consisting of 1 µl protein solution (concentrated to 7.8 mg ml^{-1}) and 1 µl reservoir solution. Single crystals were obtained after 5 d from drops comprised of reservoir solution consisting of 0.2 *M* magnesium chloride hexahydrate, 0.1 *M* sodium citrate tribasic pH 6.5, 20% PEG 3350. The best crystals, with typical dimensions of ∼400 × 100 × 50 µm (Fig. 2*a*), were selected for X-ray diffraction analysis. The crystallization of rC1ORF123 is summarized in Table 3.

Table 4
Data collection and processing.

Values in parentheses are for the outer shell.

Diffraction source	I02, Diamond Light Source
Wavelength (Å)	0.9797
Temperature (K)	100
Detector	Pilatus 6M
Crystal-to-detector distance (mm)	340.75
Rotation range per image (°)	0.2
Total rotation range (°)	180
Space group	$P2_12_12$ or $P2_12_12_1$
a, *b*, *c* (Å)	59.32, 65.35, 95.05
α, β, γ (°)	90, 90, 90
Mosaicity (°)	0.86
Resolution range (Å)	30.90–1.90 (1.94–1.90)
Total No. of reflections	159233
No. of unique reflections	29180
Completeness (%)	98.1 (97.2)
Multiplicity	5.5 (5.3)
$\langle I/\sigma(I)\rangle$	13.8 (3.0)
$R_{\rm meas}$†	0.065 (0.526)
Overall *B* factor from Wilson plot (Å^2)	35.3

† $R_{\rm meas} = \sum_{hkl}\{N(hkl)/[N(hkl)-1]\}^{1/2}\sum_i |I_i(hkl) - \langle I(hkl)\rangle|/\sum_{hkl}\sum_i I_i(hkl)$, where $N(hkl)$ is the multiplicity of reflection *hkl*.

2.5. Data collection and processing

Prior to flash-cooling in liquid nitrogen, the rC1ORF123 crystals were immersed for 5 min at 293 K in cryoprotectant solution consisting of 0.1 *M* sodium citrate buffer pH 6.5, 0.2 *M* magnesium chloride tribasic, 22% PEG 3350, 20% glycerol. X-ray diffraction data were collected on the I02 beamline at Diamond Light Source, UK at 100 K in a nitrogen-gas stream and at a wavelength of 0.9797 Å. A total of 900 images were collected with 0.2° rotation range per image using a Pilatus 6M detector. The data were indexed and integrated using *MOSFLM* (Leslie & Powell, 2007) *via* the *iMosflm* interface (v.7.1.1) (Battye *et al.*, 2011). The crystal lattice is orthorhombic, and *POINTLESS* (Evans, 2006) suggests that the crystal is most likely to belong to either space group $P2_12_12$ or $P2_12_12_1$. For further analysis, the data were scaled and merged with *AIMLESS* (Evans & Murshudov,

2013) in space group *P*222. The data-collection and processing statistics are summarized in Table 4.

3. Results and discussion

The rC1ORF123 protein with an N-terminal 6×His tag (rC1ORF123) was successfully overexpressed and purified to homogeneity using affinity chromatography (Ni–NTA) and size-exclusion chromatography (Superdex 75). rC1ORF123 migrated as a single protein band on SDS–PAGE with a molecular weight of approximately 20 kDa (Fig. 1*a*) according to standard molecular-weight markers. Size-exclusion chromatography (SEC) showed the elution of a single peak containing rC1ORF123 with a retention time indicating that the molecular weight of the rC1ORF123 protein lies between those of myoglobin (17 kDa) and ovalalbumin (44 kDa) (Fig. 1*b*). The SEC calibration curve further indicates the molecular weight of rC1ORF123 to be ∼28 kDa (Fig. 1*c*), which implies that the protein is most likely to exist as a monomer in solution. For protein validation, a Western blot was performed using anti-human C1ORF123 antibody against the rC1ORF123 protein. rC1ORF123 was positively detected by the antibody, which confirmed that the rC1ORF123 protein is similar to human C1ORF123 (Fig. 1*d*). To further verify the identity of rC1ORF123, the protein was validated by MALDI-TOF/TOF MS (Applied Biosystems/SCIEX). There were 62 peptides that matched 38% of the protein sequence of the human C1ORF123 protein (Fig. 3). Both the Western blot and the MALDI/TOF results confirmed that rC1ORF123 is identical to the human C1ORF123 protein.

The purified rC1ORF123 was concentrated to 8.0 mg ml^{-1} and used for crystallization screening. Initial crystal hits were obtained from several crystallization conditions that contained 0.2 *M* magnesium or calcium ions, medium-size polyethylene glycol (PEG 3350 or PEG 8000) and buffers (sodium cacodylate, bis-tris or HEPES) with a pH in the range between 5.5 and 7.5. Crystals suitable for X-ray diffraction analysis were obtained after optimization from conditions that consisted of 0.2 *M* magnesium chloride, 0.1 *M* sodium citrate pH 6.5, 20%(*w*/*v*) PEG 3350 (Fig. 2*a*). The crystals were flash-cooled in liquid nitrogen after the addition of an additional 20%

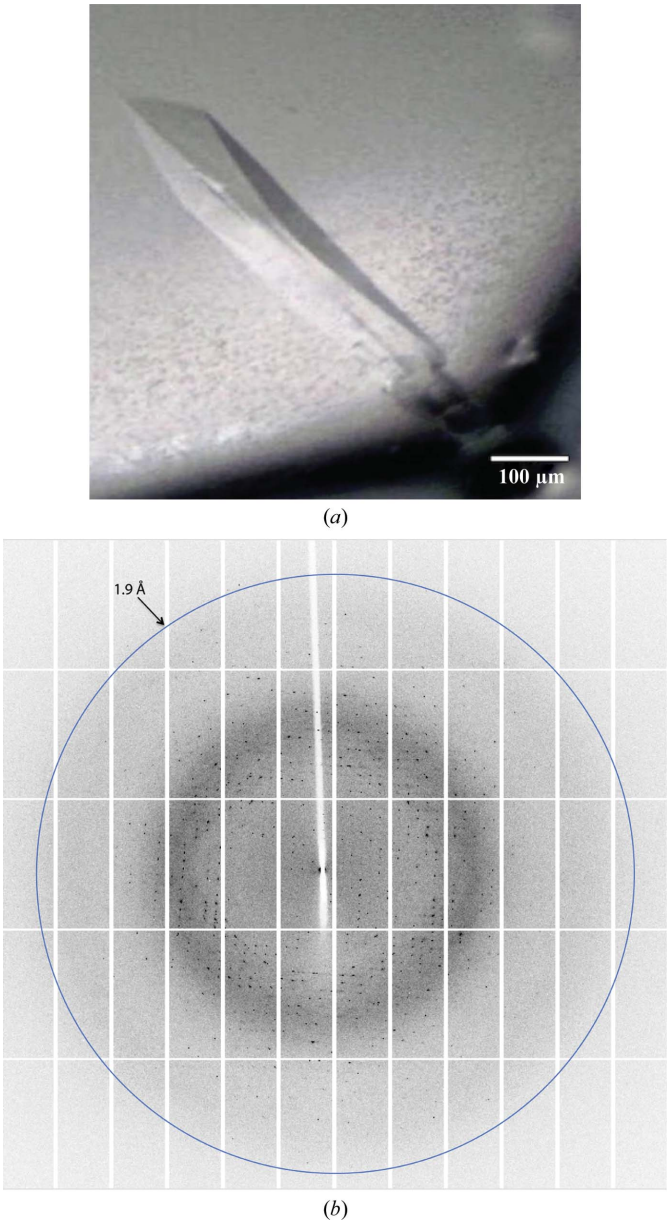

Figure 2
Protein crystal and diffraction images of recombinant human C1ORF123. (*a*) Crystal of rC1ORF123 grown using reservoir solution consisting of 0.1 *M* sodium citrate tribasic pH 6.5, 0.2 *M* magnesium chloride, 20% PEG 3350. (*b*) A diffraction image of an rC1ORF123 protein crystal which diffracted to 1.9 Å resolution on the I02 beamline at Diamond Light Source, UK. (*c*) Stereographic projection of the self-rotation function from rC1ORF123 data as calculated by *MOLREP* (Leslie & Powell, 2007). The three perpendicular twofold crystallographic axes of point group 222 are present in the $\kappa = 180^\circ$ section. The map also shows two pairs of noncrystallographic symmetry (NCS) twofold axes which are almost parallel to the crystallographic *x* and *y* axes.

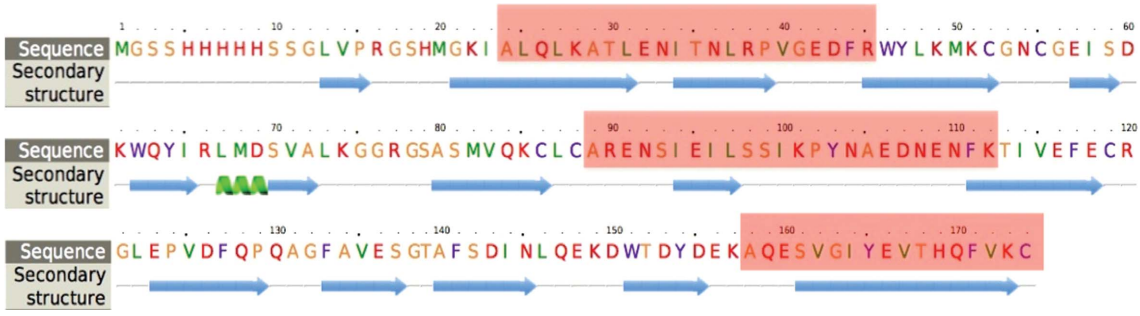

Figure 3
The secondary-structure prediction for rC1ORF123 determined by *Phyre*2 (Kelley *et al.*, 2015) suggests that C1ORF123 contains 15 β-strands, one α-helix and 14 loop regions. Peptide sequences obtained using mass spectrometry (MALDI-TOF/TOF MS; Applied Biosystems/SCIEX) matching those of C1ORF123 are highlighted in red.

glycerol to the solution as a cryoprotectant. The best rC1ORF123 crystal, with dimensions of ~400 × 100 × 50 µm, diffracted to 1.9 Å resolution (Fig. 2*b*). Diffraction data were collected with 98.1% completeness on the I02 beamline at Diamond Light Source. The data were indexed and integrated using *MOSFLM* (Leslie & Powell, 2007) and were scaled and merged with *AIMLESS* (Evans & Murshudov, 2013). Indexing indicates that the crystal lattice is orthorhombic, with unit-cell parameters a = 59.32, b = 65.35, c = 95.05 Å. However, *POINTLESS* (Evans, 2006) suggests that the actual space group is either $P2_12_12_1$ or $P2_12_12$, and it has yet to be determined by further structural analysis. The crystallographic parameters and data-collection statistics are shown in Table 4. The calculated Matthews coefficient (V_M; Matthews, 1968) value of 2.28 Å^3 Da^{-1} implies that the crystal consists of two rC1ORF123 molecules per asymmetric unit with an estimated solvent content of 46.1%. The self-rotation function (Crowther, 1972) was calculated from the rC1ORF123 data using *MOLREP* (Vagin & Teplyakov, 2010). The self-rotation function map shows three peaks in the κ = 180° section that correspond to the three perpendicular crystallographic twofold axes of the point group (222; Fig. 2*c*). The map also shows two pairs of noncrystallographic symmetry (NCS) twofold axes which are almost parallel to the crystallographic x and y axes. The corresponding self-rotation function peaks are at approximately $\varphi = \pm 10°$ and $\varphi = \pm 80°$, suggesting that rC1ORF123 forms a dimer in the crystal. This is in good agreement with Holmes *et al.* (2006), who suggest that the *P. falciparum* homologue MAL13P1.257 might also form a weak dimer based on its crystal structure. Secondary-structure prediction using *Phyre*2 (Kelley *et al.*, 2015) suggests that C1ORF123 forms 15 β-strands (61% sequence coverage) and one α-helix (2% sequence coverage) and contains 14 loop regions (Fig. 3). This is similar to the *P. falciparum* homologue MAL13P1.257 (Holmes *et al.*, 2006), which shares only 26% sequence identity with C1ORF123. Currently, we are working towards the structure determination of rC1ORf123 by molecular replacement. Comparison of human rC1ORF123 with its homologue from a protozoan parasite along with analysis of their conserved regions and structural differences may help us to understand the biological function of the DUF866 protein family.

Acknowledgements

The authors would like to acknowledge the Ministry of Science, Technology and Innovation (MOSTI), Malaysia for financial support through the ScienceFund grant (02-01-02-SF0993).

References

Batisse, J., Batisse, C., Budd, A., Böttcher, B. & Hurt, E. (2009). *J. Biol. Chem.* **284**, 34911–34917.
Battye, T. G. G., Kontogiannis, L., Johnson, O., Powell, H. R. & Leslie, A. G. W. (2011). *Acta Cryst.* D**67**, 271–281.
Bringans, S. D., Kendrick, T., Lui, J. & Lipscombe, R. J. (2008). *Rapid Commun. Mass Spectrom.* **22**, 3450–3454.
Crowther, R. A. (1972). *The Molecular Replacement Method*, edited by M. G. Rossmann, pp. 173–178. New York: Gordon & Breach.
Danielsen, J. M. R., Sylvestersen, K. B., Bekker-Jensen, S., Szklarczyk, D., Poulsen, J. W., Horn, H., Jensen, L. J., Mailand, N. & Nielsen, M. L. (2011). *Mol. Cell. Proteomics*, **10**, M110.003590.
Deng, R.-P., He, X., Guo, S.-J., Liu, W.-F., Tao, Y. & Tao, S.-C. (2014). *Proteomics*, **14**, 1020–1030.
Diamanti-Kandarakis, E. (2008). *Expert Rev. Mol. Med.* **10**, e3.
Evans, P. (2006). *Acta Cryst.* D**62**, 72–82.
Evans, P. R. & Murshudov, G. N. (2013). *Acta Cryst.* D**69**, 1204–1214.
Franceschini, A., Szklarczyk, D., Frankild, S., Kuhn, M., Simonovic, M., Roth, A., Lin, J., Minguez, P., Bork, P., von Mering, C. & Jensen, L. J. (2013). *Nucleic Acids Res.* **41**, D808–D815.
Gasteiger, E., Hoogland, C., Gattiker, A., Duvaud, S., Wilkins, M. R., Appel, R. D. & Bairoch, A. (2005). *The Proteomics Protocols Handbook*, edited by J. M. Walker, pp. 571–607. Totowa: Humana Press.
Giot, L. *et al.* (2003). *Science*, **302**, 1727–1736.
Holmes, M. A., Buckner, F. S., Van Voorhis, W. C., Mehlin, C., Boni, E., Earnest, T. N., DeTitta, G., Luft, J., Lauricella, A., Anderson, L., Kalyuzhniy, O., Zucker, F., Schoenfeld, L. W., Hol, W. G. J. & Merritt, E. A. (2006). *Acta Cryst.* F**62**, 180–185.
Kelley, L. A., Mezulis, S., Yates, C. M., Wass, M. N. & Sternberg, M. J. E. (2015). *Nature Protoc.* **10**, 845–858.
Krogan, N. *et al.* (2006). *Nature (London)*, **440**, 637–643.
Leslie, A. G. W. & Powell, H. R. (2007). *Evolving Methods for Macromolecular Crystallography*, edited by R. J. Read & J. L. Sussman, pp. 41–51. Dordrecht: Springer.
Mahmood, T. & Yang, P.-C. (2012). *N. Am. J. Med. Sci.* **4**, 429–434.
Mate, S. E., Brown, K. J. & Hoffman, E. P. (2011). *Skelet. Muscle*, **1**, 20.
Matthews, B. W. (1968). *J. Mol. Biol.* **33**, 491–497.
Mohamed-Hussein, Z. A. & Harun, S. (2009). *Theor. Biol. Med. Model.* **6**, **18**.

Restelli, L., Codrea, M. C., Savoini, G., Ceciliani, F. & Bendixen, E. (2014). *J. Proteomics*, **108**, 295–305.
Schenk, L., Meinel, D., Strässer, K. & Gerber, A. (2012). *RNA*, **18**, 449–461.
Schubert, K. O., Föcking, M. & Cotter, D. R. (2015). *Schizophr. Res.* **167**, 64–72.
Selvarajan, S. & Shanmughavel, P. (2014). *Eur. J. Appl. Sci. Technol.* **1**, 43–49.
Stelzl, U. *et al.* (2005). *Cell*, **122**, 957–968.
Vagin, A. & Teplyakov, A. (2010). *Acta Cryst.* D**66**, 22–25.
Vazquez, J., Hall, S. C. & Greco, M. A. (2009). *Brain Res.* **1298**, 37–45.
Vinayagam, A., Stelzl, U., Foulle, R., Plassmann, S., Zenkner, M., Timm, J., Assmus, H. E., Andrade-Navarro, M. A. & Wanker, E. E. (2011). *Sci. Signal.* **4**, rs8.
Wearne, T. A., Mirzaei, M., Franklin, J. L., Goodchild, A. K., Haynes, P. A. & Cornish, J. L. (2015). *J. Proteome Res.* **14**, 397–410.