Hui Zhao,^{a,b} Hong Wang,^{a,b}
Huihui Liu,^{a,b} Maikun Teng^{a,b,*}
and Xu Li^{a,b,*}

^aSchool of Life Sciences, University of Science and Technology of China, 96 Jinzhai Road, Hefei, Anhui 230026, People's Republic of China, and ^bKey Laboratory of Structural Biology, Chinese Academy of Sciences, 96 Jinzhai Road, Hefei, Anhui 230026, People's Republic of China

Correspondence e-mail: mkteng@ustc.edu.cn, sachem@ustc.edu.cn

Received 11 July 2012

Accepted 21 August 2012

Crystallization and preliminary crystallographic studies of the W2 domain of *Drosophila melanogaster* eukaryotic translation initiation factor 5C domain-containing protein

The *Drosophila melanogaster* eukaryotic translation initiation factor 5C domain-containing protein (ECP) is composed of two independently folded domains which belong to the basic leucine-zipper and W2 domain-containing protein (BZW) family. Based on the sequence similarity between the C-terminal W2 domain of ECP and some eukaryotic translation initiation factors (such as eIF2B ϵ , eIF4 γ , eIF5 *etc.*), ECP has been speculated to participate in the translation initiation process. Structural information on the C-terminal W2 domain of ECP would be helpful in understanding the specific cellular function of this protein. Here, the W2 domain of ECP was expressed and crystallized. Crystals grown by the hanging-drop vapour-diffusion method diffracted to 2.70 Å resolution and belonged to space group *I*4, with unit-cell parameters $a = b = 81.05$, $c = 57.44$ Å. The Matthews coefficient suggested that there was one molecule per asymmetric unit in the crystal.

1. Introduction

Translation initiation is the rate-limiting step during polypeptide synthesis and hence is an important point of regulation (Bieniossek *et al.*, 2006). Previously, two major mechanisms of initiation were considered to be essential: prokaryotic, based on Shine–Dalgarno (SD) interaction, and eukaryotic, which requires a cap structure and ribosomal scanning. A number of eukaryotic translation initiation factors (eIFs) are involved in the cap-dependent initiation process [core factors eIF1, eIF1A, eIF2, eIF2B *etc.* as well as auxiliary factors eIF6, poly(A)-binding protein (PABP) and others; Jackson *et al.*, 2010]. A crucial step in cap-dependent initiation is the recognition of the start codon during the ribosomal scanning process of the 43S pre-initiation complex. This complex consists of the small ribosomal subunit bound to the eIF2–GTP–Met–tRNA_i ternary complex (TC) and translation-initiation factors eIF1, eIF1A, eIF3 and eIF5 bound to the TC (Asano *et al.*, 2000). Base pairing of Met–tRNA_i with the proper AUG codon triggers GTP hydrolysis by the γ -subunit of eIF2. GTP hydrolysis evokes the release of eIF2–GDP and presumably of other bound translation-initiation factors and therefore prepares for the joining of the large subunit (Bieniossek *et al.*, 2006).

eIF5 is a protein that is known to be phylogenetically conserved from *Cryptococcus neoformans* to *Homo sapiens*. It has been proposed to act as a classical GTPase-activating protein (GAP) by providing an arginine finger (Paulin *et al.*, 2001). eIF5 possesses two independently folded domains: an N-terminal leucine-zipper motif and a C-terminal W2 domain (Wang *et al.*, 2006). The N-terminal domain of eIF5 acts as a GAP, inserting an arginine finger (Arg15) into the active site of eIF2 γ (Paulin *et al.*, 2001). The C-terminal domain of eIF5 (eIF5-CTD) mediates binding to eIF1, eIF2 β , eIF3C and eIF4G and can form a multifactor complex with eIF3/eIF1/eIF2/GTP/Met–tRNA_i^{Met} that can exist free of the ribosome (Asano *et al.*, 2000). eIF5 also acts as a GDP-dissociation inhibitor (GDI; Jennings & Pavitt, 2010). The minimal domains for eIF5-CTD binding have been mapped within eIF2 β , eIF3C and eIF4G (Singh *et al.*, 2004). In eIF5 the binding motif is located at the very C-terminus and is formed by a bipartite region dominated by aromatic and acidic amino acids called AA-boxes or the W2 domain (two invariant tryptophans; Asano *et al.*, 1999). A similar W2 domain can also be found in the C-terminus of all eukaryotic eIF2B ϵ proteins (the catalytic subunit of

eIF2B) and in mammalian eIF4G (Bieniossek *et al.*, 2006). The W2 domain is also named the eukaryotic translation initiation factor 5C terminal domain (eIF5C).

Drosophila melanogaster eIF5C domain-containing protein (ECP) is a 422-amino-acid protein which belongs to the basic leucine-zipper and W2 domain-containing protein (BZW) family. It possesses an *S*-adenosylmethionine-dependent methyltransferase domain (Tyr124–Met217) and a C-terminal W2 domain (Lys244–Glu415). Three phosphorylation sites, Ser407, Ser412 and Ser414, have been verified to be located at the C-terminus of the W2 domain using a mass-spectrometry-based proteomics method (Zhai *et al.*, 2008) and may be modified by casein kinase II (CKII) in the elongation-initiation process. There is only 27% identity over 96 amino acids to the C-terminal domain of eIF2B ϵ (PDB entry 3jui; J. Wei, H. Xu, C. Zhang, M. Wang, F. Gao & W. Gong, unpublished work) and 22% identity over 119 amino acids to the C-terminal domain of human eIF5 (PDB entry 2iu1; Bieniossek *et al.*, 2006). However, no structural and functional studies of this specific domain of ECP have been reported to date. Therefore, the functions of the ECP W2 domain still need to be verified by further study of their cell biology and structural biology. Here, a preliminary crystallographic study of the C-terminal W2 domain of the eIF5C domain-containing protein (ECP) from *D. melanogaster* is reported.

2. Materials and methods

2.1. Cloning and expression

Primers of sense strand 5'-GGAATTCCATATGAAGCTACAT-AAGGCTCAGGCTAGCC-3' and antisense strand 5'-GACACTCGAGTTACTCAGATTCCGATTCTTCTTCGGC-3' (Invitrogen) were used to amplify the ECP W2 domain (Lys244–Glu415, ECP^{244–415}) gene by polymerase chain reaction from the *D. melanogaster* ECP cDNA-containing plasmid (a gift from Professor Wei Xie of Southeast University, People's Republic of China). The PCR fragment was digested by the *Nde*I and *Xho*I restriction endonucleases and then inserted into expression vector p28 (Novagen; a

modification based on pET28-a) with a hexahistidine tag (MGHH-HHHH) at the N-terminus of the recombinant ECP^{244–415}. After sequencing, the plasmid was transformed into *Escherichia coli* BL21 (DE3) cells (Novagen). The transformant was grown in 1.6 l Luria–Bertani (LB) medium containing 50 μ g ml⁻¹ kanamycin at 310 K. When an OD₆₀₀ of 0.6–0.8 was reached, 0.5 mM isopropyl β -D-1-thiogalactopyranoside (IPTG) was added for induction. After 16 h induction at 289 K, the cells were harvested by centrifugation at 6000g for 10 min.

2.2. Purification

The harvested cells were suspended in buffer A (50 mM Tris–HCl pH 8.0, 500 mM NaCl) and lysed by sonication on ice. The soluble portion was obtained after centrifugation at 14 000g for 30 min and was applied onto a Ni–NTA column (Qiagen) pre-equilibrated with buffer A. The bound protein was eluted with buffer A containing 300 mM imidazole. After ultrafiltration to 2 ml using a Millipore 10 kDa centrifugal device, the target protein was purified using a Superdex 200 (GE Healthcare) gel-filtration chromatography column previously equilibrated with buffer A. The yield of the protein after purification was approximately 20 mg l⁻¹.

The purity of the target protein was estimated by SDS–PAGE (Fig. 1). The identity of the protein was further confirmed by LC–MS/MS peptide-mapping experiments. Briefly, the target band of the SDS–PAGE was in-gel-digested with trypsin and the digested peptides were chromatographically separated (Jupiter 5 μ m C18 300A, Phenomenex) using a Surveyor system (Thermo Fisher Scientific, California, USA) connected to an LTQ mass spectrometer (Thermo Fisher Scientific, California, USA). The chromatographic method used a flow rate of 90 nl min⁻¹ with a step gradient from mobile phase A containing 0.1% formic acid in water to mobile phase B containing 0.1% formic acid in acetonitrile. MS/MS fragmentation was performed using collision-induced dissociation (CID) with an activation *Q* of 0.250, an activation time of 30.0 ms, 35% normalized collision energy and an isolation width of 1.0 Da. MS/MS data were compared using the *SEQUEST* software (Thermo Fisher Scientific, California, USA).

2.3. Crystallization

The recombinant ECP^{244–415} was concentrated to 5 mg ml⁻¹ in buffer A (calculated from the OD₂₈₀ using a molar absorption coefficient of 29 575 M⁻¹ cm⁻¹; Eppendorf BioPhotometer plus) by centrifugal ultrafiltration (Millipore) prior to crystallization trials. Initial crystallization trials were performed with a Mosquito liquid-handling robot (TTP LabTech) using the vapour-diffusion method in 96-well crystallization plates at 289 K. Drops were prepared by mixing 0.25 μ l protein solution containing 5 mg ml⁻¹ protein with 0.25 μ l reservoir solution and were equilibrated against 100 μ l reservoir solution under 576 different conditions based on the crystallization screens Crystal Screen, Crystal Screen 2, Index, SaltRx, Grid1, Complex and Complex pH kits from Hampton Research and Molecular Dimensions. 3 d later, the best crystals were observed in drops consisting of 18% (w/v) polyethylene glycol 2000 (PEG 2000), 0.1 M MES pH 6.0. Subsequent screening was performed by varying the pH and the PEG 2000 concentration. Additive Screen (Hampton Research) was also used to optimize the conditions. Drops were prepared by mixing 1 μ l protein solution containing 5 mg ml⁻¹ protein with 1 μ l reservoir solution and were equilibrated against 200 μ l reservoir solution in 24-well crystallization plates using the hanging-drop vapour-diffusion method at 289 K. The optimal crystals

Figure 1
SDS–PAGE analysis of the ECP W2 domain. The protein was analyzed on 15% SDS–PAGE and was stained with Coomassie Blue. The ~21 kDa band in lane 1 coincides with the theoretical molecular weight of recombinant ECP^{244–415} (21.24 kDa). Lane M contains molecular-mass markers (labelled in kDa).

Figure 2

Crystal of recombinant ECP^{244–415} grown under the condition 12% (w/v) PEG 2000, 0.1 M caesium chloride, 0.1 M MES pH 6.0.

Table 1

Data-collection statistics.

Values in parentheses are for the highest resolution shell.

Space group	I4
Wavelength (Å)	0.979
Unit-cell parameters (Å)	$a = b = 81.05$, $c = 57.44$
Resolution limits (Å)	50.0–2.70 (2.80–2.70)
No. of observed reflections	21458
No. of unique reflections	5304
Completeness (%)	92.8 (100.0)
Moscaicity (°)	0.72
Multiplicity	4.4 (4.9)
$R_{p.i.m.}^{\dagger}$ (%)	4.0 (21.4)
$R_{r.i.m.}^{\ddagger}$ (%)	8.4 (47.3)
$\langle I/\sigma(I) \rangle$	15.2 (5.4)
V_M (Å ³ Da ^{−1})	2.36
No. of subunits per asymmetric unit	1
Solvent content (%)	47.82

$\dagger R_{p.i.m.} = \sum_{hkl} \{1/[N(hkl) - 1]\}^{1/2} \sum_i |I_i(hkl) - \langle I(hkl) \rangle| / \sum_{hkl} \sum_i I_i(hkl)$. $\ddagger R_{r.i.m.} = \sum_{hkl} \{N(hkl)/[N(hkl) - 1]\}^{1/2} \sum_i |I_i(hkl) - \langle I(hkl) \rangle| / \sum_{hkl} \sum_i I_i(hkl)$, where $I_i(hkl)$ is the i th observation of reflection hkl , $\langle I(hkl) \rangle$ is the weighted average intensity of all observations i of reflection hkl and $N(hkl)$ is the multiplicity.

appeared in drops consisting of 12% (w/v) PEG 2000, 0.1 M caesium chloride, 0.1 M MES pH 6.0.

2.4. Data collection and processing

For data collection, crystals were flash-cooled in liquid nitrogen using a cryoprotectant solution consisting of 10% (v/v) glycerol in 12% (w/v) PEG 2000, 0.1 M caesium chloride, 0.1 M MES pH 6.0 and were then transferred into a liquid-nitrogen stream. X-ray diffraction data were collected on beamline 17U1 of the Shanghai Synchrotron Radiation Facility (SSRF) using a Jupiter CCD detector. All frames were collected at 100 K using a 1° oscillation angle. The crystal-to-detector distance was set to 400 mm. The complete diffraction data set was subsequently processed using *HKL-2000* (Otwinowski & Minor, 1997).

3. Results and discussion

In this work, recombinant ECP^{244–415} protein was successfully expressed, purified and crystallized. The ~21 kDa band on the SDS-PAGE gel was determined to be the recombinant ECP^{244–415} protein by an LC-MS/MS peptide-mapping experiment; the 153 amino acids of ECP^{244–415} covered 88.95% of the 172-amino-acid sequence. Tetragonal crystals grew in 3 d (Fig. 2). A total of 180 diffraction images were recorded from a single crystal. Diffraction data were collected from ECP^{244–415} crystals and were processed to 2.70 Å resolution. The quality indicators in Table 1 clearly indicate the presence of useful data to a much higher resolution and we will reprocess the data when phasing is successful. The crystal belonged to space group *I4*, with unit-cell parameters $a = b = 81.05$, $c = 57.44$ Å. The Matthews coefficient of 2.36 Å³ Da^{−1} indicated that there was only one monomer in the asymmetric unit, with a solvent content of 47.82%. Detailed data-processing statistics are given in Table 1. The molecular-replacement method was used to determine the structure using *MOLREP* (Vagin & Teplyakov, 2010) from the *CCP4* package (Winn *et al.*, 2011). However, a molecular-replacement solution could not be found. Experiments to obtain crystals of selenomethionine-substituted protein are currently under way.

We are grateful to the staff members at the SSRF for the collection of diffraction data. Financial support for this project was provided by the Chinese Ministry of Science and Technology (grant Nos. 2012CB917200 and 2009CB825500), the Chinese National Natural Science Foundation (grant Nos. 31130018, 30900224 and 10979039) and the Science and Technological Fund of Anhui Province for Outstanding Youth (grant No. 10040606Y11).

References

- Asano, K., Clayton, J., Shalev, A. & Hinnebusch, A. G. (2000). *Genes Dev.* **14**, 2534–2546.
- Asano, K., Krishnamoorthy, T., Phan, L., Pavitt, G. D. & Hinnebusch, A. G. (1999). *EMBO J.* **18**, 1673–1688.
- Bieniossek, C., Schütz, P., Bumann, M., Limacher, A., Uson, I. & Baumann, U. (2006). *J. Mol. Biol.* **360**, 457–465.
- Jackson, R. J., Hellen, C. U. & Pestova, T. V. (2010). *Nature Rev. Mol. Cell Biol.* **11**, 113–127.
- Jennings, M. D. & Pavitt, G. D. (2010). *Small GTPases*, **1**, 118–123.
- Otwinowski, Z. & Minor, W. (1997). *Methods Enzymol.* **276**, 307–326.
- Paulin, F. E., Campbell, L. E., O'Brien, K., Loughlin, J. & Proud, C. G. (2001). *Curr. Biol.* **11**, 55–59.
- Singh, C. R., Yamamoto, Y. & Asano, K. (2004). *J. Biol. Chem.* **279**, 49644–49655.
- Vagin, A. & Teplyakov, A. (2010). *Acta Cryst.* **D66**, 22–25.
- Wang, D., Xu, M., Liu, J., Wan, Y., Deng, H., Dou, F. & Xie, W. (2006). *Biochim. Biophys. Acta*, **1760**, 1428–1433.
- Winn, M. D. *et al.* (2011). *Acta Cryst.* **D67**, 235–242.
- Zhai, B., Villén, J., Beausoleil, S. A., Mintseris, J. & Gygi, S. P. (2008). *J. Proteome Res.* **7**, 1675–1682.