PROTEINS
STRUCTURE ■ FUNCTION ■ BIOINFORMATICS

STRUCTURE NOTE

Crystal structure of *Arabidopsis* translation initiation factor eIF-5A2

Yan-Bin Teng, Xiao-Xiao Ma, Yong-Xing He, Yong-Liang Jiang, Jin Du, Chengbin Xiang, Yuxing Chen, and Cong-Zhao Zhou*

Hefei National Laboratory for Physical Sciences at Microscale and School of Life Sciences, University of Science and Technology of China, Hefei, Anhui 230027, People's Republic of China

Key words: eukaryotic translation initiation factor; *Arabidopsis thaliana*; crystal structure; dimerization.

INTRODUCTION

The protein eukaryotic translation initiation factor 5A (eIF-5A) is a highly conserved eukaryotic translation initiation factor (eIF) found in eukaryotes and archaea.[1–3] Biochemical and molecular studies revealed that eIF-5A is the sole protein that contains a modified amino acid residue hypusine (Nε-(4-amino-2-hydroxybutyl)lysine).[4] The hypusination modification is made by two sequential reactions catalyzed by deoxyhypusine synthase (*EC 1.1.1.249*) and deoxyhypusine hydroxylase (*EC 1.14.99.29*).[5–7]

eIF-5A was originally purified and identified from immature red blood cells.[8] However, unlike the traditional translation initiation factors, eIF-5A is not essential for global protein synthesis[8,9] but might be involved in mRNA translocation across the nuclear envelope.[10,11] The hypusinated yeast eIF-5A was recently found to promote translation elongation.[12] Moreover, hypusine of the yeast eIF-5A has been found to be required for the sequence-specific interaction with RNA.[13] To help clarify these diverse and even somewhat controversial functions, seven structures of eIF-5A from various organisms have been solved (*Methanococcus jannaschii*, PDB codes: 1eif and 2eif[14]; *Pyrobaculum aerophilum*, PDB code: 1bkb[15]; *Pyrococcus horikoshii*, PDB code: 1iz6[16]; *Leishmania braziliensis*, PDB code: 1 × 6o; *Leishmania mexicana*, PDB code:1 × td; *Homo sapiens*, PDB code: 3cpf; *Saccharomyces cerevisiae*, PDB code: 3er0). They all share an overall structure of two domains, both of which resemble the nucleic acid binding fold.[14]

The plant *Arabidopsis thaliana* encodes three isoforms of eIF-5A: *At*eIF-5A1, 2, and 3 (GenBank Accession Numbers AF296082, BE039424, and AV526594). As the best investigated one, eIF-5A2 has been found to play a crucial role in plant growth and development by controlling cell proliferation and senescence.[17] Here, we report the crystal structure of eIF-5A2 at 2.3 Å resolution, which represents a novel dimerization pattern specifically conserved in all plants.

METHODS

Construction, expression, and purification of eIF-5A

The coding sequence of *eIF-5A* gene was amplified by PCR using the cDNA library of *A. thaliana* as the template. An additional sequence coding for a six-histidine tag was introduced at the 5′ end of the gene during PCR amplification. Then, the PCR product was cloned into a pET28a-derived vector and expressed at 37°C using the transformed *Escherichia coli* BL-21 (DE3) strain and 2 ×

Grant sponsor: Ministry of Agriculture of China; Grant number: 2009ZX08009-37B; Grant sponsor: Ministry of Science and Technology of China; Grant numbers: 2006CB910202, 2006CB806501.
*Correspondence to: Cong-Zhao Zhou, Hefei National Laboratory for Physical Sciences at Microscale and School of Life Sciences, University of Science and Technology of China, Hefei, Anhui 230027, People's Republic of China. E-mail: zcz@ustc.edu.cn
Received 31 May 2009; Revised 24 June 2009; Accepted 26 June 2009
Published online 7 July 2009 in Wiley InterScience (www.interscience.wiley. com).
DOI: 10.1002/prot.22530

YT medium (OXOID LTD.) supplemented with 30 μg/mL kanamycin. When the cell culture reached an $OD_{600\ nm}$ of 0.6, protein expression was induced with 0.2 m*M* IPTG (BBI), and the cells were grown for a further 20 h at 16°C. Cells were collected by centrifugation, resuspended in 30 mL buffer containing 200 m*M* NaCl and 20 m*M* Tris-HCl, pH 7.5. Cells were lysed by three cycles of freezing/thawing followed by sonication. His-tagged proteins were purified using a Ni^{2+} affinity column. Eluted protein was further purified by gel filtration using a Superdex™ 75 column (GE Healthcare Bioscience) equilibrated in 200 m*M* NaCl, 20 m*M* β-mercaptoethanol, and 20 m*M* Tris-HCl, pH 7.5. The purity of the pooled fractions was checked by SDS-PAGE.

Crystallization, data collection, and processing

Crystals of eIF-5A2 were obtained by the hanging drop vapor diffusion method at 16°C. In each drop of crystallization, 1 μL protein sample at 10 mg/mL in the buffer of 20 m*M* NaCl, 20 m*M* Tris-HCl, pH 7.5, and 20 m*M* β-mercaptoethanol was mixed with 1 μL reservoir solution (25% ethylene glycol) and equilibrated against 0.5 mL reservoir solution. Typically, crystals appeared in 1–2 days and reached the maximum size of 300 × 400 × 0.05 μm in 1 week. The crystal was flash frozen at 100 K in a stream of nitrogen gas. Images of diffraction were collected using a MAR345dtb detector (MarResearch, Germany), with wavelength of 1.5418 Å and oscillation of 1°. X-ray crystallographic data were processed with *MOSFLM*[18] and scaled with *SCALA*.[19]

Structure solution and refinement

The structure was determined by molecular replacement method with the program *PHASER*,[20] using the structure of eIF-5A from *L. braziliensis* (PDB code 1 × 6o) as the search model. The initial model was refined by using the maximum likelihood method implemented in *REFMAC5*[21] as part of *CCP4i*[22] program suite and rebuilt interactively by using the σ_A-weighted electron density maps with coefficients 2m*Fo-Fc* and m*Fo-Fc* in the program *COOT*.[23] The final model was validated with the programs *PROCHECK*[24] and *MOLPROBITY*.[25] Structure factors and coordinates have been deposited in the Protein Data Bank (PDB http://www.rcsb.org/pdb) under the accession code of 3HKS. The final statistics and refinement parameters were listed in Table I. All the structure figures were prepared using the program *PyMol* (http://pymol.sourceforge.net/).[26]

Table I
Crystal Parameters, Data Collection and Structure Refinement Statistics

Data processing	
Space group	$P2_12_12_1$
Unit cell (Å), (°)	$a = 56.56$, $b = 79.83$, $c = 94.95$, $\alpha = \beta = \gamma = 90.00$
Resolution range (Å)	94.92–2.23 (2.28–2.23)[a]
Unique reflections	20,393 (1,293)
Completeness (%)	99.5 (96.4)
$\langle I/\sigma(I)\rangle$	22.1 (8.1)
R_{merge}[b] (%)	5.6 (23.9)
Average redundancy	5.7 (5.7)
Refinement statistics	
Resolution range (Å)	94.92–2.23
R-factor[c]/*R*-free[d] (%)	20.82/24.26
Number of protein atoms	2,358
Number of water atoms	169
RMSD[e] bond lengths (Å)	0.011
RMSD bond angles (°)	1.248
Mean B factors (Å^2)	22.15
Ramachandran plot[f] (residues, %)	
Most favored (%)	97.41
Additional allowed (%)	2.59
Outliers (%)	0
PDB entry	3HKS

[a]The values in parentheses refer to statistics in the highest bin.
[b]$R_{merge} = \sum_{hkl}\sum_i|I_i(hkl) - \langle I(hkl)\rangle|/\sum_{hkl}\sum_i I_i(hkl)$, where $I_i(hkl)$ is the intensity of an observation and $\langle I(hkl)\rangle$ is the mean value for its unique reflection; Summations are over all reflections.
[c]*R*-factor $= \sum_h|F_o(h) - F_c(h)|/\sum_h F_o(h)$, where F_o and F_c are the observed and calculated structure–factor amplitudes, respectively.
[d]*R*-free was calculated with 5% of the data excluded from the refinement.
[e]Root-mean square-deviation from ideal values.
[f]Categories were defined by Molprobity.

RESULTS AND DISCUSSION

Overall structure

The crystal structure of eIF-5A2 was refined to the resolution of 2.3 Å, with two subunits (A and B) in one asymmetric unit. The majority of residues are well fitted in the electron density map, except for the N-terminal His-tag and residues Met1-Ala15 in both subunits, and the last two residues Gly158 and Lys159 in subunit B. The side chains of Arg87, Asp106, and Asp135 in subunit A, and those of His52, Glu132 in subunit B are fitted in the electron density map at a lower occupancy, using the program *PHOENIX*.[27] At the surface, six molecules of ethylene glycol are well fitted in the flat but slightly longer pieces of electron density [Fig. 1(A)]. The final refinement and validation statistics are listed in Table I.

The structure of eIF-5A2 comprises two distinct domains of antiparallel β-sheet that resembles the classic architecture of eIF-5A [Fig. 1(A)]. The highly conserved N-terminal domain has an SH3-like barrel motif,[16] composed of strands β1–β6 and a 3_{10} helix. Strands β3–β6 forms a distorted semiopen β-barrel. The hypusine modification site Lys51 is located at the protruding loop between strands β3 and β4. The C-terminal domain resembles the oligonucleotide-binding fold (OB fold)[30], and it contains strands β7–β12 and helix α1.

Comparative structural analysis with the DALI server (http://ekhidna.biocenter.helsinki.fi/dali_server)[31] indicates that both the N- and C-terminal domains have the nucleic acid binding fold as previous reports.[14,16] The

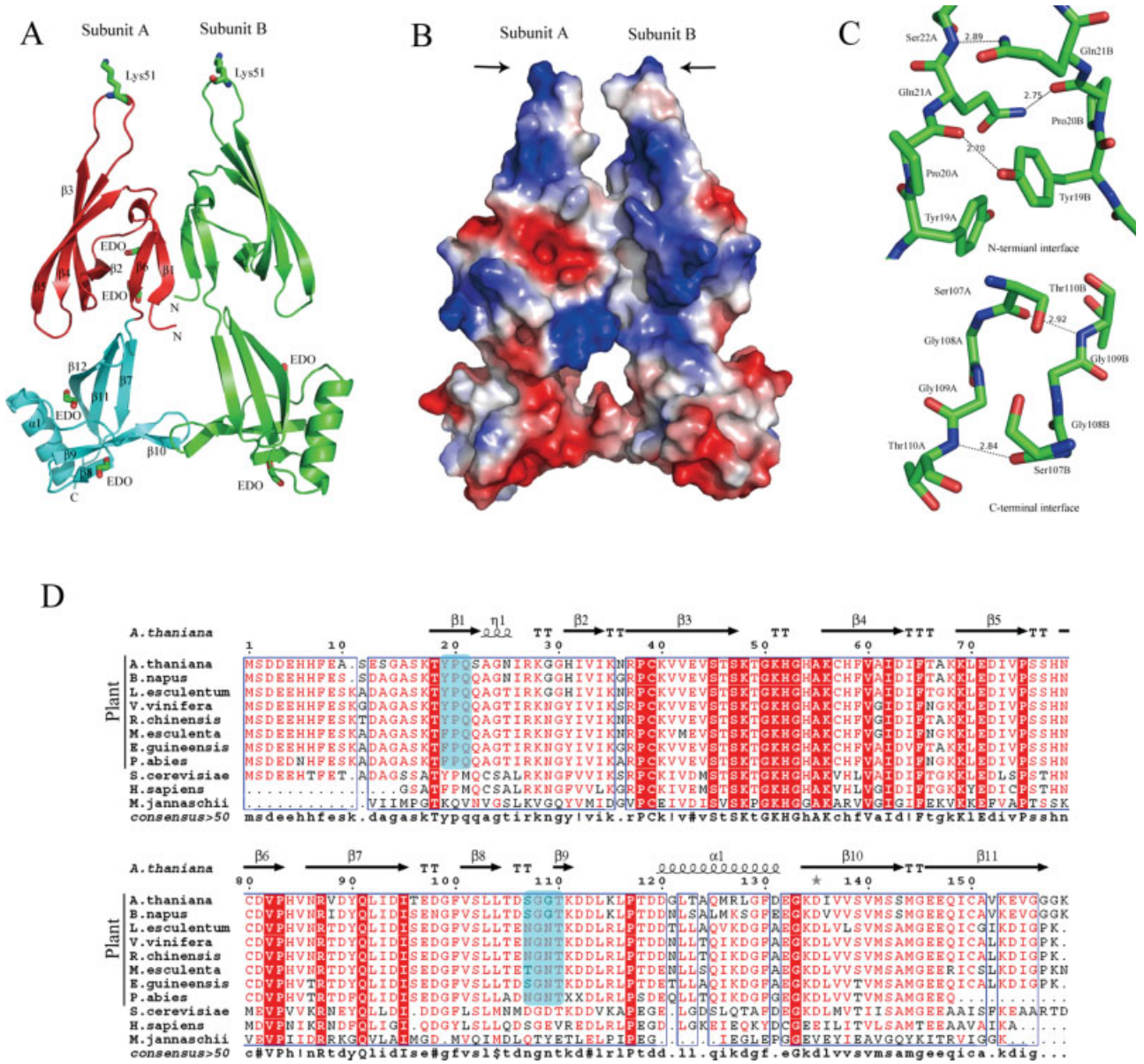

Figure 1

The overall structure and dimerization pattern of *A. thaliana* eIF-5A2. **(A)** The overall structure. The conserved residue Lys51 and ethylene glycol molecules are shown by sticks. The N- and C-terminal domains in subunit A are colored in red and cyan, and subunit B in green, respectively. **(B)** Surface representation of electrostatic potential. The surface potential is displayed as a color gradient from red (negative) to blue (positive). The black arrows point to the possible RNA binding sites. **(C)** The dimeric interfaces. The hydrogen bonds are shown by dashed lines. **(D)** Alignment of eIF-5A sequences. The upper eight sequences are from the plants. Residues Tyr19, Pro20, and Gln21 at the N-terminal interface and Ser107-Thr110 at the C-terminal interface are marked in cyan. All sequences were obtained from NCBI databases. The multialignment was performed using the programs MultAlin[28] and ESPript.[29]

recombinant eIF-5A2 overexpressed in *Pichia pastoris* with the post-translational modification has a higher affinity toward nucleic acids compared with the unmodified one overexpressed in *E. coli*, as indicated by the absorbance at 260 nm (data not shown). It is in agreement with the previous reports that the exposed positively charged residues, especially the hypusine, can bind the phosphate backbone of RNA by electrostatic interactions.[32] This positively charged region at the protruding loop of the hypusination site is also present in eIF-5A2 [Fig. 1(B)].

The novel crystal dimer structure in plants

The homodimerization of eIF-5A has been found in several organisms. In *M. jannaschii* IF-5A dimer (PDB

code:2eif), the two β3 strands interact with each other via six hydrogen bonds to form a continuous six-stranded antiparallel β-sheet, leading to the two active site loops pointing to the opposite directions.[14] In human eIF-5A dimer (PDB code: 3cpf), the two chains are linked via a disulfide bond. The dimeric structure of eIF-5A2 formed between subunits A and B was significantly different than those observed in other structures. The two subunits form a dimer in parallel via an interface of 400 $Å^2$ between the N-terminal domains, and one of 280 $Å^2$ between the C-terminal domains, respectively [Fig. 1(A)]. In fact, a small fraction of eIF-5A2 also exists as homodimer in solution, as detected by gel filtration (data not shown).

To have a better view, the two interfaces between the pairs of N- and C-terminal domains are shown in Figure 1(C). The N-terminal interface is stabilized by the surrounding residues via three hydrogen bonds: Ser22A-N ~ Gln21B-Nε; Gln21A-Nε ~ Pro20B-O; and Pro20A-O ~ Tyr19B-Oε [Fig. 1(C)]. The smaller C-terminal interface possesses two main chain hydrogen bonds of Ser107A ~ Thr110B and Thr110 ~ ASer107B [Fig. 1(C)].

The structure of eIF-5A2 represents the first structure of plant eIF-5A, indicating a rather distinct dimerization pattern from the previous eIF-5A of known structure. After performing a multiple sequence-alignment by the programs MultAlin[28] and ESPript,[29] we noticed that the residues Tyr19 (substituted by Phe in several cases), Pro20, and Gln21 are highly conserved in all species of plant [Fig. 1(D)]. All these three residues are crucial to stabilize the N-terminal interface [Fig. 1(C)]. The residues contributing to the C-terminal interface are also conserved [Fig. 1(D)]. Thus, we proposed that this dimerization pattern is unique in plants. Further biochemical and biophysical analyses are needed to prove whether this form of dimer could facilitate the RNA binding.

REFERENCES

1. Chen KY, Liu AY. Biochemistry and function of hypusine formation on eukaryotic initiation factor 5A. Biol Signals 1997;6: 105–109.
2. Park MH, Lee YB, Joe YA. Hypusine is essential for eukaryotic cell proliferation. Biol Signals 1997;6:115–123.
3. Schnier J, Schwelberger HG, Smit-Mcbride Z, Kang HA, Hershey JW. Translation initiation factor 5A and its hypusine modification are essential for cell viability in the yeast *Saccharomyces cerevisiae.* Mol Cell Biol 1991;11:3105–3114.
4. Cooper HL, Park MH, Folk JE, Safer B, Braverman R. Identification of the hypusine-containing protein hy+ as translation initiation factor eIF-4D. Proc Natl Acad Sci USA 1983;80:1854–1857.
5. Park MH, Wolff EC. Cell-free synthesis of deoxyhypusine. Separation of protein substrate and enzyme and identification of 1,3-diaminopropane as a product of spermidine cleavage. J Biol Chem 1988;263:15264–15269.
6. Park MH, Wolff EC, Folk JE. Hypusine: its post-translational formation in eukaryotic initiation factor 5A and its potential role in cellular regulation. BioFactors (Oxford, England) 1993;4:95–104.
7. Wolff EC, Park MH, Folk JE. Cleavage of spermidine as the first step in deoxyhypusine synthesis. The role of NAD. J Biol Chem 1990;265:4793–4799.
8. Kemper WM, Berry KW, Merrick WC. Purification and properties of rabbit reticulocyte protein synthesis initiation factors M2Balpha and M2Bbeta. J Biol Chem 1976;251:5551–5557.
9. Benne R, Hershey JW. The mechanism of action of protein synthesis initiation factors from rabbit reticulocytes. J Biol Chem 1978; 253:3078–3087.
10. Rosorius O, Reichart B, Kratzer F, Heger P, Dabauvalle MC, Hauber J. Nuclear pore localization and nucleocytoplasmic transport of eIF-5A: evidence for direct interaction with the export receptor CRM1. J Cell Sci 1999;112:2369–2380.
11. Zanelli CF, Valentini SR. Is there a role for eIF5A in translation? Amino Acids 2007;33:351–358.
12. Saini P, Eyler DE, Green R, Dever TE. Hypusine-containing protein eIF5A promotes translation elongation. Nature 2009;459: 118–121.
13. Xu A, Chen KY. Hypusine is required for a sequence-specific interaction of eukaryotic initiation factor 5A with postsystematic evolution of ligands by exponential enrichment RNA. J Biol Chem 2001;276:2555–2561.
14. Kim KK, Hung LW, Yokota H, Kim R, Kim SH. Crystal structures of eukaryotic translation initiation factor 5A from Methanococcus jannaschii at 1.8 A resolution. Proc Natl Acad Sci USA 1998; 95:10419–10424.
15. Peat TS, Newman J, Waldo GS, Berendzen J, Terwilliger TC. Structure of translation initiation factor 5A from *Pyrobaculum aerophilum* at 1.75 A resolution. Structure 1998;6:1207–1214.
16. Yao M, Ohsawa A, Kikukawa S, Tanaka I, Kimura M. Crystal structure of hyperthermophilic archaeal initiation factor 5A: a homologue of eukaryotic initiation factor 5A (eIF-5A). J Biochem 2003;133:75–81.
17. Feng H, Chen Q, Feng J, Zhang J, Yang X, Zuo J. Functional characterization of the *Arabidopsis* eukaryotic translation initiation factor 5A-2 that plays a crucial role in plant growth and development by regulating cell division, cell growth, and cell death. Plant Physiol 2007;144:1531–1545.
18. Leslie AGW. Recent changes to the MOSFLM package for processing film and image plate data. Joint CCP4 + ESF-EAMCB Newsletter on Protein. Crystallography 1992;26.
19. Evans PR. Data reduction. In: Proceedings of CCP4 Study Weekend on Data Collection and Processing, Warrington, 1993;114–122.
20. Mccoy AJ, Grosse-Kunstleve RW, Adams PD, Winn MD, Storoni LC, Read RJ. Phaser crystallographic software. J Appl Crystallogr 2007;40:658–674.
21. Murshudov GN, Vagin AA, Dodson EJ. Refinement of macromolecular structures by the maximum-likelihood method. Acta Crystallogr D Biol Crystallogr 1997;53:240–255.
22. Bailey S. The Ccp4 Suite – programs for protein crystallography. Acta Crystallogr D 1994;50:760–763.
23. Emsley P, Cowtan K. Coot: model-building tools for molecular graphics. Acta Crystallogr 2004;60:2126–2132.
24. Laskowski RA, Macarthur MW, Moss DS, Thornton JM. Procheck - a program to check the stereochemical quality of protein structures. J Appl Crystallogr 1993;26:283–291.
25. Davis IW, Leaver-Fay A, Chen VB, Block JN, Kapral GJ, Wang X, Murray LW, Arendall WB, Snoeyink J, Richardson JS, Richardson DC. MolProbity: all-atom contacts and structure validation for proteins and nucleic acids. Nucl Acids Res 2007;35:W375–W383.

26. Delano WL. The PyMOL molecular graphics system. San Carlos, CA, USA: DeLano Scientific LLC. Available at: http://www.pymol.org
27. Adams PD, Grosse-Kunstleve RW, Hung LW, Ioerger TR, Mccoy AJ, Moriarty NW, Read RJ, Sacchettini JC, Sauter NK, Terwilliger TC. PHENIX: building new software for automated crystallographic structure determination. Acta Crystallogr 2002;58:1948–1954.
28. Corpet F. Multiple sequence alignment with hierarchical clustering. Nucl Acids Res 1988;16:10881–10890.
29. Gouet P, Robert X, Courcelle E. ESPript/ENDscript: extracting and rendering sequence and 3D information from atomic structures of proteins. Nucl Acids Res 2003;31:3320–3323.
30. Murzin AG. OB(oligonucleotide/oligosaccharide binding)-fold: common structural and functional solution for non-homologous sequences. EMBO J 1993;12:861–867.
31. Holm L, Sander C. Protein structure comparison by alignment of distance matrices. J Mol Biol 1993;233:123–138.
32. Burd CG, Dreyfuss G. Conserved structures and diversity of functions of RNA-binding proteins. Science 1994;265:615–621.