Biochemical and Biophysical Research Communications 453 (2014) 826–832

Contents lists available at ScienceDirect

Biochemical and Biophysical Research Communications

journal homepage: www.elsevier.com/locate/ybbrc

High-resolution crystal structure reveals a HEPN domain at the C-terminal region of *S. cerevisiae* RNA endonuclease Swt1

Shuxia Peng *, Ke Zhou, Wenjia Wang, Zengqiang Gao, Yuhui Dong, Quansheng Liu

Multidisciplinary Research Center, Institute for High Energy Physics, Chinese Academy of Sciences, Beijing 100049, China

ARTICLE INFO

Article history:
Received 29 September 2014
Available online 16 October 2014

Keywords:
HEPN domain
Crystal structure
RNA endonuclease
SAXS
Swt1

ABSTRACT

Swt1 is an RNA endonuclease that plays an important role in quality control of nuclear messenger ribonucleoprotein particles (mRNPs) in eukaryotes; however, its structural details remain to be elucidated. Here, we report the crystal structure of the C-terminal (CT) domain of Swt1 from *Saccharomyces cerevisiae*, which shares common characteristics of higher eukaryotes and prokaryotes nucleotide binding (HEPN) domain superfamily. To study in detail the full-length protein structure, we analyzed the low-resolution architecture of Swt1 in solution using small angle X-ray scattering (SAXS) method. Both the CT domain and middle domain exhibited a good fit upon superimposing onto the molecular envelope of Swt1. Our study provides the necessary structural information for detailed analysis of the functional role of Swt1, and its importance in the process of nuclear mRNP surveillance.

1. Introduction

Nuclear mRNP quality control is an important step prior to mRNA export into cytoplasm, a fundamental process for successful eukaryotic gene expression. Swt1 (synthetically lethal with TREX) is an RNA endonuclease that associates with both the TREX complex (mRNA transcription and export complex) and the NPC (nuclear pore complex) *in vivo* [1]. The main function of Swt1 is to facilitate the degradation of aberrant pre-mRNAs at the nuclear periphery, preventing the leakage of defective mRNPs into the cytoplasm [2]. Analysis of the amino acid sequence as well as the prediction of the secondary structure of Swt1 from *Saccharomyces cerevisiae* (*S. cerevisiae*) showed that this endonuclease consists of three domains, namely N-terminal (NT) domain, middle (M) domain, and C-terminal (CT) domain. The middle domain contains a PIN (PilT N terminus) domain [3], a domain deemed responsible for the RNA nuclease activity observed in proteins containing this functional region. The crystal structures of several PIN domains of different proteins have been solved [4–10]. The PIN domain of Swt1 was shown to contain endonuclease activity *in vitro*, and appears to be essential for proper functioning of Swt1 in living cells [2].

In order to understand the precise mechanisms of Swt1 function in mRNA surveillance, it is important to analyze its molecular structure at high resolution. However, so far no three-dimensional structures of Swt1 have been analyzed. Whilst the structure of the Swt1 PIN domain has been predicted on the basis of its sequence similarity to other PIN domains [11], the structures of the NT and CT domains of Swt1 remain to be solved. Here, we report the crystal structure of the CT domain of *S. cerevisiae* Swt1, which was determined by using a single-wavelength anomalous dispersion (SAD) method. Furthermore, the low-resolution architecture of Swt1 full-length protein in solution was analyzed using small angle X-ray scattering (SAXS) method, thus providing novel insights into the domain organization of Swt1.

2. Materials and methods

2.1. Gene cloning, protein expression and purification

The genes encoding Swt1 full-length and truncated proteins were obtained by polymerase chain reaction (PCR) from *S. cerevisiae* genomic DNA, including Swt1 full-length (1-458AA), CT domain (312-458AA), PIN domain (120-315AA), N + M domain (1-315AA) and M + C domain (120-458AA). The amplified genes were cloned into vector PET28a and expressed in *Escherichia coli* strain BL21(DE3) with an N-terminal 6-His-tag. Cells were harvested by

Abbreviations: CT, C-terminal; HEPN domain, higher eukaryotes and prokaryotes nucleotide binding domain; mRNP, messenger ribonucleoprotein particle; NPC, nuclear pore complex; NT, N-terminal; PIN domain, PilT N terminus domain; SAD, single wavelength anomalous dispersion; SAXS, small angle X-ray scattering; SeMet, selenomethionine-substituted derivative; TREX, mRNA transcription export complex.

* Corresponding author at: 19B Yuquan Road, Shijingshan District, Beijing 100049, China.

E-mail address: pengsx@ihep.ac.cn (S. Peng).

http://dx.doi.org/10.1016/j.bbrc.2014.10.040
0006-291X/

centrifugation, resuspended in lysis buffer (50 mM Tris–HCl pH 8.0, 500 mM NaCl, 12.5% sucrose, 2 mM β-mercaptoethanol) and sonicated for about 30 min (6 s sonication/25 s pause for each cycle). Proteins were purified through a Ni^{2+} affinity fast flow column, subsequently followed by an anion exchange Q column and a size exclusion column Superdex200 (GE Healthcare), then concentrated to 20 mg/ml in buffer (20 mM Tris pH 8.0, 50 mM NaCl, 2 mM β-mercaptoethanol) for crystallization.

A Leu379Met mutation at the CT domain was introduced for preparing a selenomethionine-substituted (SeMet) derivative. SeMet-L379M-C domain was expressed in BL21(DE3) with selenomethionine supplemented in M9 medium. The expression and purification procedure for SeMet-L379M-C was the same as that for full-length protein.

2.2. Crystallization and data collection

All of Swt1 full-length and truncated proteins have been screened for crystal growth. Crystals were obtained for the Swt1 CT domain by using the sitting drop vapor diffusion method. A volume of 1 μl protein solution was mixed with an equal volume of reservoir solution containing 0.1 M NaAc (pH 4.6), 2 M Na Formate, and equilibrated against 100 μl reservoir solution. Crystals appeared after about 2 weeks of growth at 18 °C. The crystallization conditions for the SeMet-L379M-C domain were identical to that used for the native CT domain. X-ray diffraction data were collected by adding saturated LiCl to the reservoir buffer as cryoprotectant. Crystal diffraction data of both native CT domain and SeMet-L379M-C domain were collected at the Beijing Synchrotron Radiation Facility (BSRF).

2.3. Structure determination and refinement

Diffraction data were processed and scaled using HKL-2000 [12]. The structure of SeMet-L379M-C domain was solved using the SAD method. Selenium sites were located using SHELX [13]. The initial phase was solved using the Autosol program embedded in the PHENIX software suite [14]. 80% of all residues were built with the program Autobuild [15]. The remaining residues were built manually using COOT [16], and the model was refined using the diffraction data in phenix.refine [17]. The SeMet-L379M-C domain structure was used as the search model to solve the native CT domain structure with PHASER using the molecular replacement method [18]. The final R_{work} and R_{free} for the refined model were 21.2% and 26.5%, respectively. The stereochemical quality of the CT domain structure was checked with the program PROCHECK [19]. A summary of data collection and structure refinement statistics is listed in Table 1. Figures were prepared using Pymol [20]. The DALI server was used for comparison of protein structures [21]. The structure was deposited into the protein data bank (PDB: 4PQZ).

2.4. Structure prediction

Swt1 full-length sequence was put into the website of Phyre2 (http://www.imperial.ac.uk/phyre/) to predict the three-dimensional structure of Swt1 [11]. The analysis revealed that the PIN domain of Swt1 (Lys129–Lys276) could be predicted using hSMG6 PIN domain structure (PDB: 2HWW) as a template. The confidence in the produced model was 100%.

2.5. SAXS measurement and data processing

Synchrotron radiation SAXS data were collected on the beamline 1W2A of the BSRF using a MARCCD detector. The scattering signal was recorded in the range of the momentum transfer $0.02 < s < 0.18$ Å^{-1}, where $s = (4\pi\sin\theta)/\lambda$, 2θ is the scattering angle, and $\lambda = 1.54$ Å is the X-ray wavelength. The measurements were carried out in a cuvette with an exposure time of 200 s. The scattering curves were primarily processed using PRIMUS [22]. To avoid protein concentration impacting on the interactions of the protein macromolecules in solution, a series of protein concentrations (2, 3 and 4 mg/ml of Swt1 full-length) in solution buffer (20 mM Tris–HCl, pH 8.0, 50 mM NaCl, and 2 mM DTT) were prepared and measured. After subtracting buffer scattering, the data were averaged and extrapolated to zero concentration. The resultant curve was used for all calculations as well as for the modeling. Distance distribution functions $p(r)$ and maximal diameter of the scattering object D_{max} were calculated from experimental data using GNOM [23]. The *ab initio* modeling of low-resolution shapes of the protein in solution was performed using GASBOR [24]. Ten models were calculated, and the most universal model was chosen as typical.

Table 1
Data collection and structure refinement statistics.

Data collection	Native-C	SeMet-L379M-C
Wavelength (Å)	1	0.9791
Space group	$P\,3_2\,2\,1$	$P\,3_2\,2\,1$
Cell dimensions	$a = b = 61.52$ Å, $c = 84.97$ Å; $\alpha = \beta = 90°$, $\gamma = 120°$	$a = b = 61.15$ Å, $c = 83.42$ Å; $\alpha = \beta = 90°$, $\gamma = 120°$
No. of molecules/asu	1	1
Resolution range[a] (Å)	50–2.30 (2.34–2.30)	50–2.70 (2.75–2.70)
Unique reflections[a]	8495 (420)	5261 (252)
Mean I/σ	61.89 (5.99)	88.17 (8.19)
Redundancy[a]	10.6 (10.4)	41.2 (33.6)
Completeness[a] (%)	97.9 (100)	100 (100)
R_{merge}[a] (%)	5.6 (39.5)	6.7 (46.8)
Refinement statistics		
Resolution range (Å)	33.22–2.30	
R_{work} (%)	21.2	
R_{free} (%)	26.5	
No. of protein atoms	1173	
No. of water molecules	25	
RMSD bond lengths (Å)	0.009	
RMSD bond angles (°)	1.141	
Ramachandran statistics		
Most favored regions (%)	96.27	
Allowed regions (%)	3.73	

[a] Values in parentheses are for the highest resolution shell.

2.6. Assays for nuclease activity in vitro

Analysis for testing nuclease activities of the purified proteins on fluorescence-labeled 49-nt poly(A)+ RNAs was performed at 30 °C in the buffer containing 10 mM Tris–HCl (pH 7.6), 50 mM NaCl, 2 mM DTT, 0.8 U/μl RNasin, 5% glycerol and 2.5 mM $MnCl_2$. Proteins were added to the reaction buffer containing 0.6 pmol of the fluorescence-labeled RNA substrate, and the total volume of 10 μl reaction mixtures were incubated at 30 °C for 30 min. Subsequently, the reactions were stopped by adding 10 μl 2× formamide-containing RNA loading buffer and heated for 10 min at 70 °C. Reaction products were analyzed on a denatured 12% polyacrylamide/8 M urea gel and visualized using a Typhoon FLA 7000 (GE Healthcare).

3. Results and discussion

3.1. Overall structure of the S. cerevisiae Swt1 CT domain

The crystal of the Swt1 CT domain belongs to space group $P3_221$, and contains only one molecule per asymmetric unit. The crystal structure of Swt1 CT domain was solved at 2.3-Å resolution.

Residues Pro323 to Thr458 could be traced. The structure of the CT domain is composed exclusively of α helices, including three long antiparallel helices (α1, α5 and α6), two shorter helices (α2 and α4) that stack at a slight angle to the long helices, and another shorter α helix (α3) forming a helix-loop-helix motif with helix α2 at an angle of about 90 degrees. Embedded in between the α-helices are two $3_{(10)}$-helices (η1, η2), one being located between helices of α1 and α2, and the other between helices α3 and α4 (Fig. 1A). The topology of the CT domain structure is shown in Fig. 1B. A disulfide bond is formed by two residues of Cys391 in the two adjacent molecules, which are symmetric with each other via a crystallographic twofold rotation axis (Fig. 1C). The electrostatic potential surface of the CT domain of Swt1 is shown in Fig. 1D.

3.2. Sequence similarity of Swt1 protein family

Multiple sequence alignment showed that Swt1 proteins from different species are highly conserved. Apart from the conserved residues in the NT and the middle domains, the conserved residues in the CT domain of Swt1 include Gly354, Tyr362, Thr368, Trp381, Val384, Phe385, Leu416, Phe419, Leu427 and Leu442 (Supplementary Fig. 1A). Residues of Tyr362 and Thr368 are located on the loop between η1 and α2, and residues of Trp381, Val384 and Phe385 are located on the loop between helix α2 and α3 to form the right-angled kink of the two helices (Fig. 1E), and most of the remaining conserved residues are located on α helices.

3.3. Structural comparison of the Swt1 CT domain to the superfamily of HEPN domains

After analysis of the crystal structure of the CT domain, we found that this conserved domain shares common characteristics with other proteins of the HEPN domain superfamily. For example, the residue's number of the CT domain is 136, which is close to the size of the originally defined HEPN domain (approximately 100–120 amino acids long). It is composed exclusively of α helices, four of which (α1, α2, α5, α6) form an up-and-down helical bundle. The core of this fold is comprised of α-hairpins, with the N- and C-termini spatially juxtaposed (Fig. 2A). Several helical elements (α3, η2, α4) are found inserted in between the long α helices, with the sequence of this insert being poorly conserved (Supplementary Fig. 1A). The partly conserved Ex3[KR] motif (x is any amino acid) containing a conserved acidic residue in HEPN domain is present as Dx3R in the CT domain of yeast Swt1 (Fig. 1E). This motif is found conserved as a Ex3K motif in its human Swt1 homologue (Supplementary Fig. 1A). Taken together, these characteristics suggest that the CT domain of Swt1 is a member of the HEPN domain superfamily, which is in agreement with an earlier prediction using computational methods [25].

Interestingly, whilst *S. cerevisiae* is a lower eukaryote, the HEPN domain superfamily is present primarily in higher eukaryotes and prokaryotes. In order to investigate similarities between these domains, we compared the structure of the CT domain with those of HEPN domains from bacteria as well as higher eukaryotes [26,27]. For instance, a comparison of the crystal structures of the Swt1 CT domain (Fig. 2Aa) and the HEPN domain protein from the bacterium *Thermotoga maritime* (PDB: 1O3U, Fig. 2Ab) resulted in a *Z* score of 4.3, with an RMSD of 2.6 Å (Fig. 2B). In addition, when comparing the crystal structure of Swt1 CT domain with that of the human sacsin HEPN domain (PDB: 3O10, Fig. 2Ac), the *Z* score was 3.9, with an RMSD of 2.8 Å in Dali server (Fig. 2C). Together, these results suggest that the 3D-structure of the Swt1 CT domain closely resembles those of HEPN domains. Interestingly, the sequence identities of HEPN domains can be less than 10% [21].

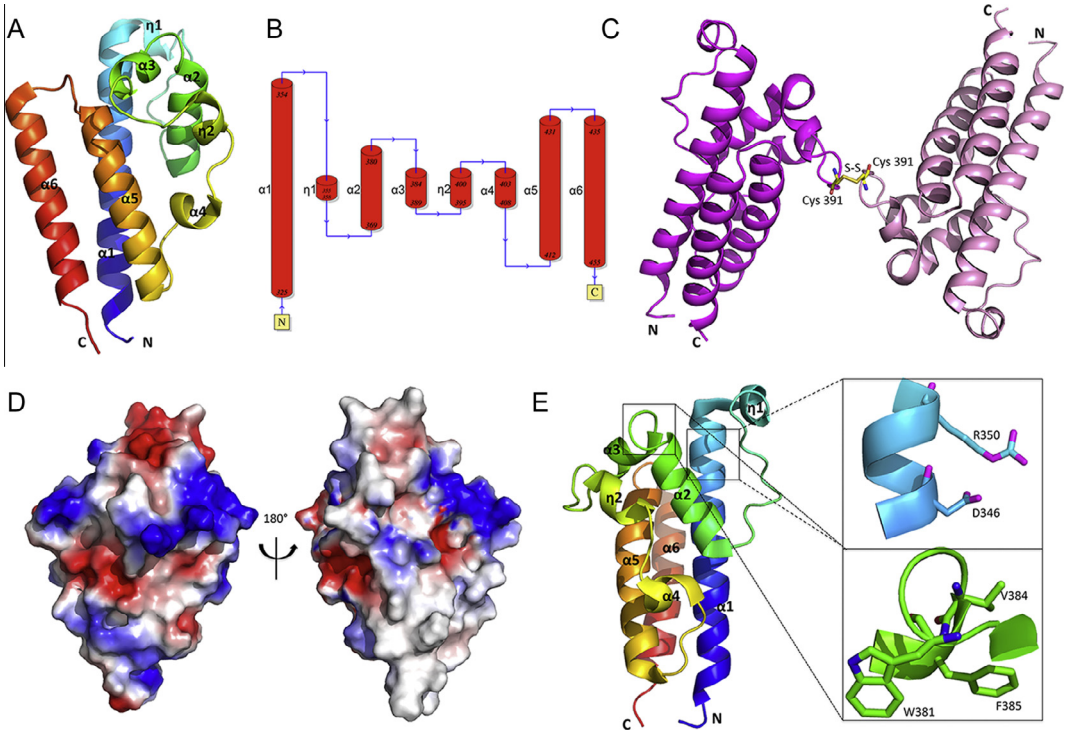

Fig. 1. Structure of the CT domain of *S. cerevisiae* Swt1. (A) Crystal structure of Swt1 CT domain. (B) Topology diagram of the CT domain. (C) The disulfide bond between two molecules; (D) electrostatic potential surface of the CT domain. The surface is colored from blue to red for positively to negatively charged regions. (E) The partly conserved D3xR motif in the α1 helix and the conserved amino acid residues in the loop between the α2 and α3 helices. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

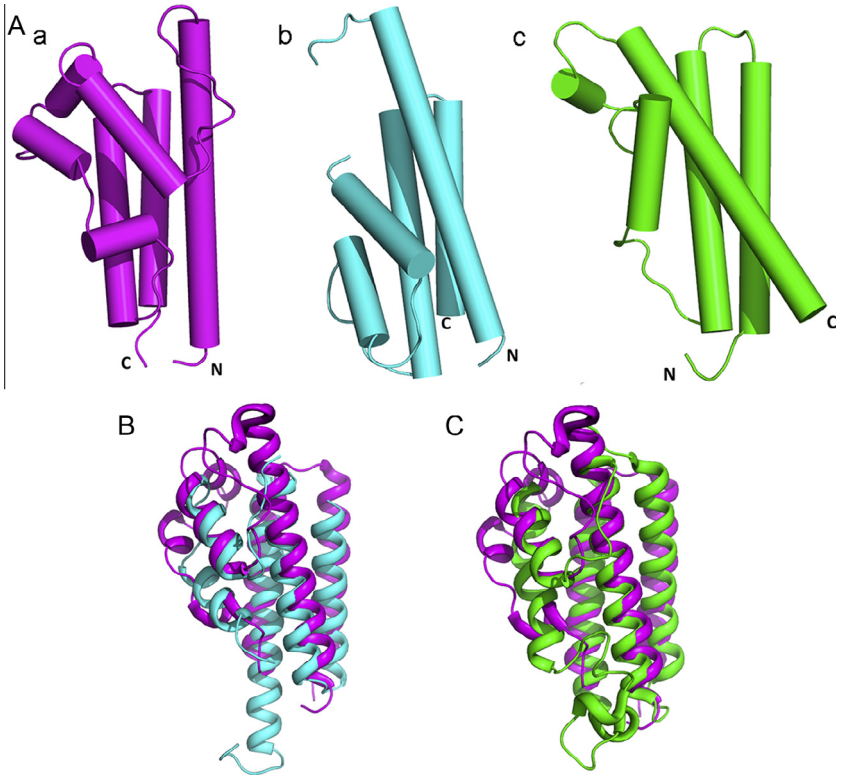

Fig. 2. Structural comparison of the CT domain of Swt1 with other HEPN domains. (A): a, Structure of Swt1 CT domain from *S. cerevisiae* (PDB: 4PQZ, magenta); b, Structure of HEPN domain protein from *Thermotoga maritime* (PDB: 1O3U, cyan); c, Structure of human sacsin HEPN domain (PDB: 3O10, green). (B) Structural alignment of 4PQZ to 1O3U. (C) Structural alignment of 4PQZ to 3O10. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

3.4. Structure prediction of the PIN domain

S. cerevisiae Swt1 can be subdivided into three domains, according to its primary sequence and the prediction of its secondary structure (Fig. 3A). In order to understand the structural details of Swt1, we used the program Phyre2 to predict the structure of its PIN domain. The predicted structure of the PIN domain included residues Lys129 to Lys276, and used the structure of the hSMG6 PIN domain as a template (Supplementary Fig. 1B). As shown in Fig. 3B, analysis of the superimposition revealed four highly conserved acidic residues (aspartate or glutamate), as well as a single polar residue (serine) in the PIN domain of Swt1, all of which were proposed to represent the putative active sites for nuclease activity.

3.5. SAXS analysis of the low-resolution structure of Swt1

To identify the entire contour of full-length Swt1, and to understand the domain organization of this protein, we analyzed the low-resolution envelope of Swt1 in solution using SAXS method, which predicts the intact structural conformation of proteins. Experimental scattering patterns of Swt1 in solution were averaged and extrapolated to zero concentration (Fig. 3C, curve 1). Ten independent GASBOR models were generated, and the one with the smallest normalized spatial discrepancy (NSD) value was chosen as a typical model (NSD = 1.311). This model provided a good approximation to the experimental data, with a discrepancy value of $\chi = 0.86$ (Fig. 3C, curve 2). The protein structural model was represented by an ensemble of dummy residues equal to the residual number of Swt1, and the envelope of Swt1 was found to resemble a "W" shape (Fig. 3D).

The radius of gyration (R_g) of Swt1 calculated by Guinier approximation was identified as 35 Å. Computational analysis of the molecular mass (MM) by SAXS resulted in a mass of approx. 50 kDa, which is in agreement with the MM of the Swt1 monomer in solution. The maximum diameter D_{max} of the particle was 105 Å; in addition, the protein exhibited a flexible and extended conformation in solution. Structural analysis of the SAXS-generated Swt1 model revealed three protein domains, together with flexible linkers (resulting from the secondary structural prediction) connecting the adjacent domains [28]. The surface of the predicted PIN structure is of cubic shape, with a length of the diagonal line of approx. 40-Å (Fig. 3Ea). In comparison, the surface of the crystal structure of the CT domain exhibits the shape of an extended, triangular pyramid, with a width of 30-Å and a height of 50-Å (Fig. 3Eb). In order to provide a better understanding of the overall domain organization of Swt1, the rigid structures of the PIN domain and the CT domain were placed into the low-resolution envelope of Swt1 manually. Our analysis revealed that both domains fit well into the GASBOR model of full-length Swt1 (Fig. 3Ec).

3.6. Analysis of nuclease activity of Swt1 in vitro

In vitro, Swt1 full-length protein degrades RNA given the presence of manganese ions (Mn^{2+}). In order to analyze nuclease activity *in vitro*, several truncations of Swt1 were expressed and purified, including N + M domain, M + C domain, M (PIN) domain and CT domain. Our *in vitro* results showed that the PIN domain possesses RNA endonuclease activity of Swt1, whereas the CT domain displayed no such activity (Fig. 4A and B). In order to investigate the influence of the CT domain on endonuclease activity, we tested RNA degradation activity by deleting either NT or CT domains. Our *in vitro* results showed that deleting the CT domain slightly decreased the RNA degradation efficiency (Fig. 4A and C). In addition, we tested Nuclease activity as a function of time. As

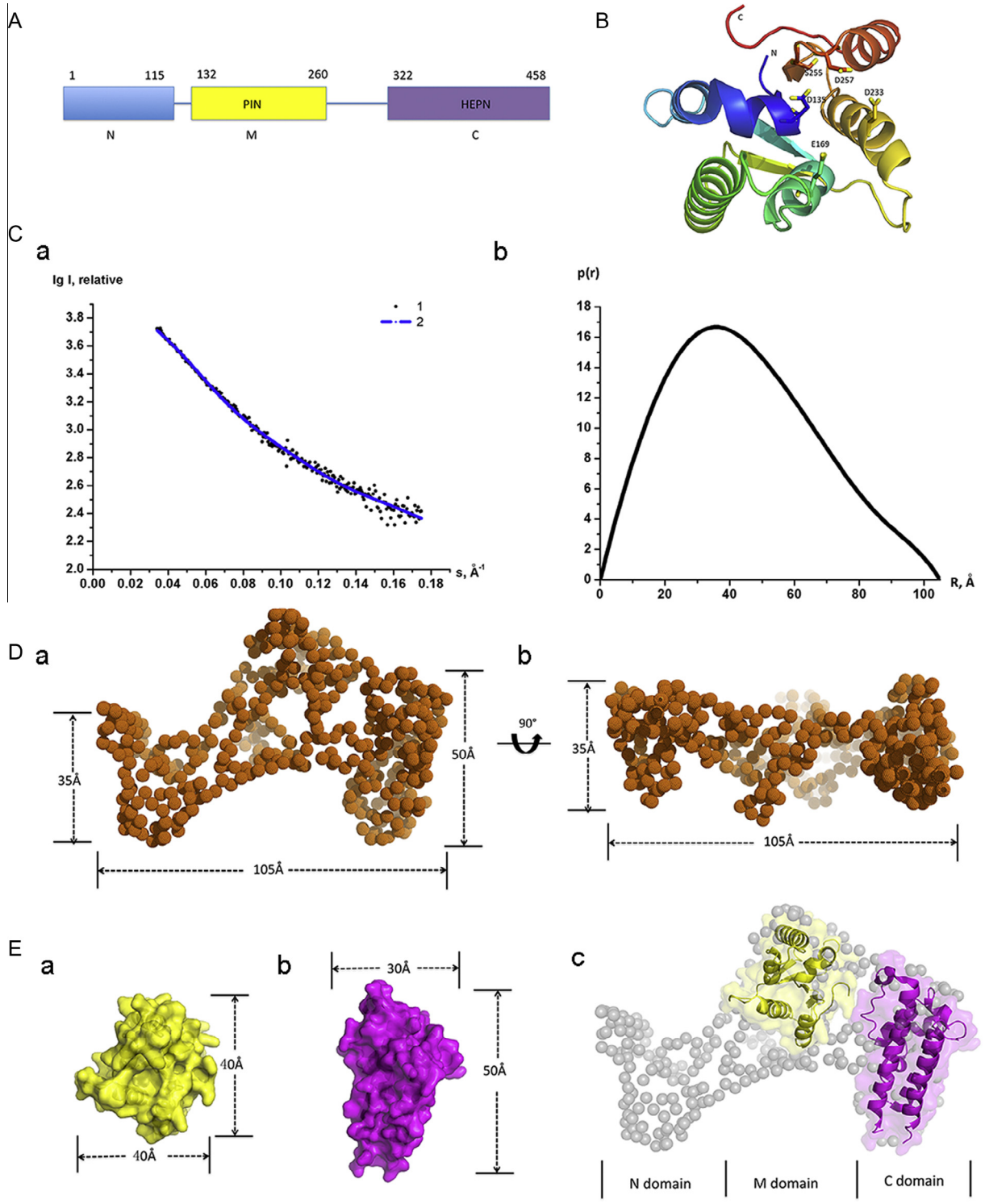

Fig. 3. The low-resolution structural model of Swt1 solved by SAXS method. (A) Three domains of yeast Swt1. Numbers on top indicate the amino acids of Swt1. (B) The predicted structure of the PIN domain (Lys129–Lys276), with the conserved residues as putative active sites indicated in sticks. (C): a, Experimental scattering curve of Swt1 full-length in solution and subsequent model reconstruction; 1, experimental scattering curve extrapolated to zero concentration; 2, scattering pattern calculated from the *ab initio* model; b, The distance distribution functions $p(r)$ calculated using GNOM. (D) GASBOR model of Swt1 in solution. (E): a, Surface of the predicted structure of the PIN domain; b, surface of the crystal structure of the CT domain; c, docking of PIN and CT domains into the GASBOR model.

shown in Fig. 4D, deletion of the Swt1 CT domain reduced the rate of RNA degradation relative to that of the full-length protein.

The CT domain of Swt1 shares most common features of HEPN domains, most of which in bacteria possess RNA nuclease activity [29,30]. However, our analysis of RNA degradation activity *in vitro* indicates that the HEPN-like domain of *S. cerevisiae* Swt1 does not possess endonuclease activity. This observation is probably due to the fact that the CT domain of *S. cerevisiae* Swt1 lacks a conserved Rx4-6H motif. In HEPN domains, this motif is primarily responsible for metal-independent nuclease activity [25]. Although the

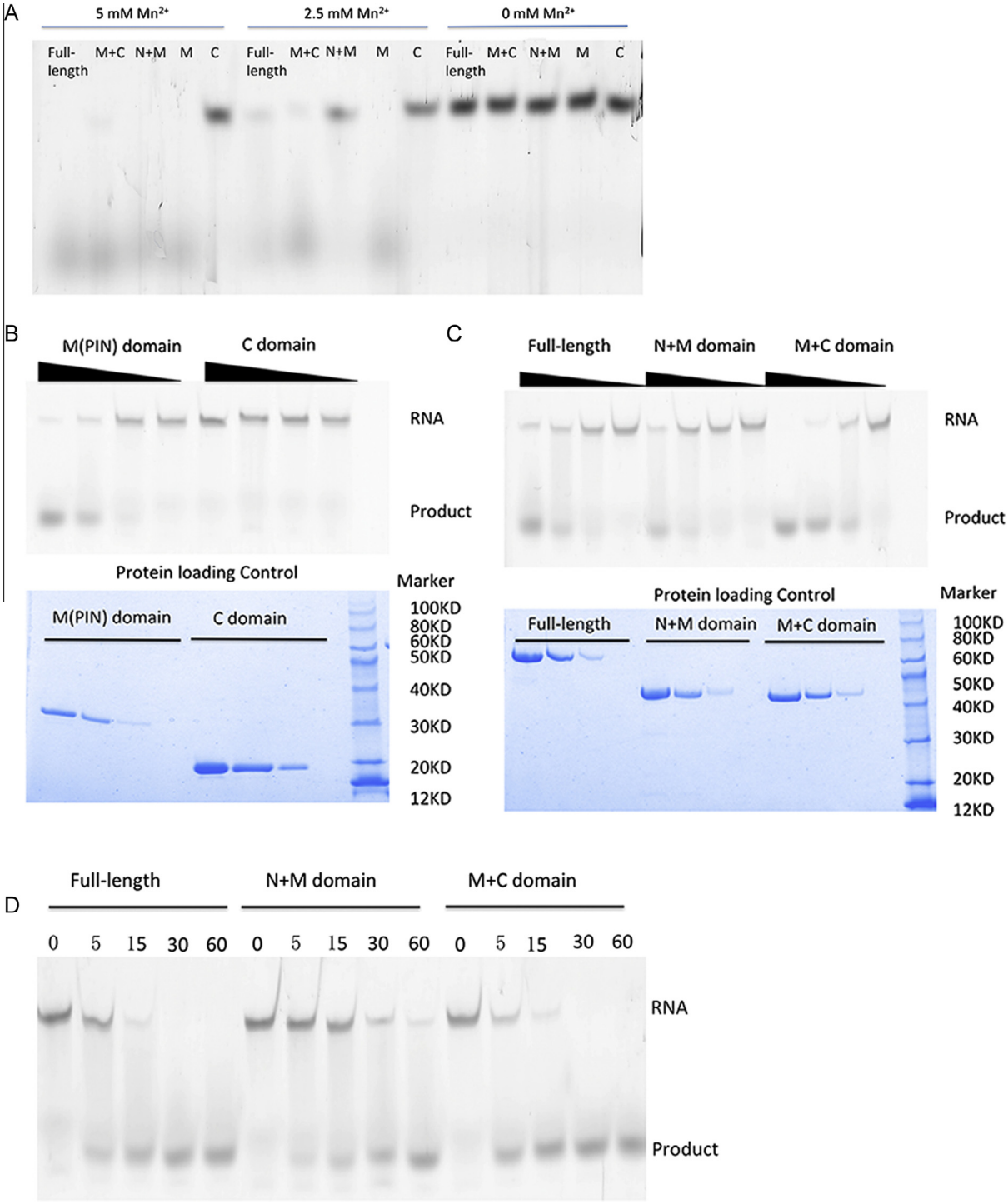

Fig. 4. Nuclease activity of Swt1 proteins on single-stranded 49-nt poly(A)+ RNA. (A) Nuclease activity assays of Swt1 full-length, M + C, N + M, M and CT domains in the presence of Mn^{2+} at 5, 2.5 and 0 mM concentrations. 20 pmol of each protein were added to the total 10 μl reaction buffer containing 0.6 pmol fluorescence-labeled RNA, and were incubated for 30 min at 30 °C. (B) and (C) Nuclease activities of the middle PIN domain, CT domain, Swt1 full-length, N + M and M + C domains of yeast Swt1. 50, 20, 5, 0 pmol of proteins were added to the total 10 μl reaction buffer, containing 0.6 pmol fluorescence-labeled RNA in the presence of 2.5 mM Mn^{2+}, incubated at 30 °C for 30 min. (D) Nuclease activities of Swt1 full-length, N + M and M + C domains at different time points. 100 pmol of proteins were added to the 50 μl reaction buffer containing totally 3 pmol fluorescence-labeled RNA in the presence of 2.5 mM Mn^{2+} at 30 °C, each of 10 μl reaction sample was taken out and the reaction was stopped at the indicated time points of 0, 5, 15, 30 and 60 min.

HEPN-like domain of Swt1 could slightly influence the nuclease activity *in vitro*, the specific function of this domain remains to be identified. Thus, we can only speculate that the HEPN domain might function as a non-catalytic domain responsible for sensing or recognizing unspliced mRNPs, which are specifically targeted by Swt1. Moreover, Swt1 from *Candida dubliniensis* contains the conserved Rx4H motif in the η2-α4 insert as shown by multiple sequence alignment (Supplementary Fig. 1A). We therefore cannot exclude the possibility that this HEPN-like domain of Swt1 may play a role as RNA endonuclease in other species.

We reveal here for the first time the crystal structure of the CT domain of *S. cerevisiae* Swt1, which is a conserved domain in Swt1 protein family and belongs to the HEPN domain superfamily. Since the HEPN domain has not been reported in lower eukaryotes before [31,32], our study provides evidence that the family of HEPN domains is widespread in prokaryotes, both lower and higher eukaryotes; however, its function needs to be further investigated. The low-resolution structural model of Swt1 obtained by SAXS was primarily reported, which revealed the most likely domain organization of Swt1. Our results presented here reveal for the first time that in *S. cerevisiae* Swt1, the HEPN domain is adjacent to the C-terminal of the PIN domain. Since both the PIN domain and the HEPN domain are common toxin domains in different bacterial toxin–antitoxin systems, we speculate that these two domains may have been acquired simultaneously and modified throughout eukaryotic evolution. Our experimental results presented here should provide important structural information required for further clarification on how Swt1 functions in mRNA quality control at the nuclear periphery in living cells.

Acknowledgments

We thank all researchers and staff of the Beijing Synchrotron Radiation Facility involved in this study for their assistance with data collection, and Weiqun Shen (Peking University) for her assistance in the N-terminal sequence analysis. We thank Juan Li (IHEP), Chaofei Chen (Anhui University), Peng Cao (IBP) and Yuanyuan Chen (IBP) for their technical assistance. We thank all members of the Liu lab for critical discussion of the manuscript. We thank Dr. T. Juelich (Peking University) for linguistic assistance during the preparation of this manuscript. This work was supported by the National Natural Science Foundation of China (Grant Nos. 31300623, 31170689, 11179022) and the National Basic Research Program 973 of China (Grant No. 2012CB917201).

Appendix A. Supplementary data

Supplementary data associated with this article can be found, in the online version, at http://dx.doi.org/10.1016/j.bbrc.2014.10.040.

References

[1] S. Rother, E. Clausing, A. Kieser, K. Strasser, Swt1, a novel yeast protein, functions in transcription, J. Biol. Chem. 281 (2006) 36518–36525.
[2] M. Skruzny, C. Schneider, E. Hurt, et al., An endoribonuclease functionally linked to perinuclear mRNP quality control associates with the nuclear pore complexes, PLoS Biol. 7 (2009) 164–174.
[3] P.M. Clissold, C.P. Ponting, PIN domains in nonsense-mediated mRNA decay and RNAi, Curr. Biol. 10 (2000) R888–890.
[4] I. Levin, R. Schwarzenbacher, I.A. Wilson, et al., Crystal structure of a PIN (PilT N-terminus) domain (AF0591) from *Archaeoglobus fulgidus* at 1.90 Å resolution, Proteins 56 (2004) 404–408.
[5] J. Jeyakanthan, E. Inagaki, C. Kuroishi, T.H. Tahirov, Structure of PIN-domain protein PH0500 from *Pyrococcus horikoshii*, Acta Crystallogr. Sect. F: Struct. Biol. Cryst. Commun. 61 (2005) 463–468.
[6] D. Takeshita, S. Zenno, M. Tanokura, et al., Crystal structure of the PIN domain of human telomerase-associated protein EST1A, Proteins 68 (2007) 980–989.
[7] R.D. Bunker, J.L. McKenzie, E.N. Baker, V.L. Arcus, Crystal structure of PAE0151 from *Pyrobaculum aerophilum*, a PIN-domain (VapC) protein from a toxin–antitoxin operon, Proteins 72 (2008) 510–518.
[8] J. Lu, M. Sun, K. Ye, Structural and functional analysis of Utp23, a yeast ribosome synthesis factor with degenerate PIN domain, RNA 19 (2013) 1815–1824.
[9] F. Glavan, I. Behm-Ansmant, E. Izaurralde, E. Conti, Structures of the PIN domains of SMG6 and SMG5 reveal a nuclease within the mRNA surveillance complex, EMBO J. 25 (2006) 5117–5125.
[10] V.L. Arcus, K. Backbro, E.N. Baker, et al., Distant structural homology leads to the functional characterization of an archaeal PIN domain as an exonuclease, J. Biol. Chem. 279 (2004) 16471–16478.
[11] L.A. Kelley, M.J.E. Sternberg, Protein structure prediction on the Web: a case study using the Phyre server, Nat. Protoc. 4 (2009) 363–371.
[12] Z. Otwinowski, W. Minor, Processing of X-ray diffraction data collected in oscillation mode, Macromol. Crystallogr. Part A 276 (1997) 307–326.
[13] G.M. Sheldrick, SHELX: applications to macromolecules, in: S. Fortier (Ed.), Direct Methods for Solving Macromolecular Structures, Kluwer Academic Publishers, Dordrecht, 1998, pp. 401–411.
[14] P.D. Adams, P.V. Afonine, P.H. Zwart, et al., PHENIX: a comprehensive Python-based system for macromolecular structure solution, Acta Crystallogr. Sect. D Biol. Crystallogr. 66 (2010) 213–221.
[15] T.C. Terwilliger, R.W. Grosse-Kunstleve, P.D. Adams, et al., Iterative model building, structure refinement and density modification with the PHENIX AutoBuild wizard, Acta Crystallogr. Sect. D Biol. Crystallogr. 64 (2008) 61–69.
[16] P. Emsley, K. Cowtan, Coot: model-building tools for molecular graphics, Acta Crystallogr. Sect. D Biol. Crystallogr. 60 (2004) 2126–2132.
[17] P.V. Afonine, R.W. GrosseKunstleve, P.D. Adams, et al., Towards automated crystallographic structure refinement with phenix.refine, Acta Crystallogr. Sect. D Biol. Crystallogr. 68 (2012) 352–367.
[18] A.J. McCoy, R.W. Grosse-Kunstleve, R.J. Read, et al., Phaser crystallographic software, J. Appl. Crystallogr. 40 (2007) 658–674.
[19] R.A. Laskowski, M.W. Macarthur, D.S. Moss, J.M. Thornton, Procheck – a program to check the stereochemical quality of protein structures, J. Appl. Crystallogr. 26 (1993) 283–291.
[20] The PyMOL Molecular Graphics System, Schrödinger, LLC.
[21] L. Holm, J. Park, DaliLite workbench for protein structure comparison, Bioinformatics 16 (2000) 566–567.
[22] P.V. Konarev, V.V. Volkov, D.I. Svergun, et al., PRIMUS: a Windows PC-based system for small-angle scattering data analysis, J. Appl. Crystallogr. 36 (2003) 1277–1282.
[23] D.I. Svergun, Determination of the regularization parameter in indirect-transform methods using perceptual criteria, J. Appl. Crystallogr. 25 (1992) 495–503.
[24] D.I. Svergun, M.V. Petoukhov, M.H.J. Koch, Determination of domain structure of proteins from X-ray solution scattering, Biophys. J . 80 (2001) 2946–2953.
[25] V. Anantharaman, K.S. Makarova, L. Aravind, et al., Comprehensive analysis of the HEPN superfamily: identification of novel roles in intra-genomic conflicts, defense, pathogenesis and RNA processing, Biol. Direct 8 (2013) 15.
[26] G. Kozlov, A.Y. Denisov, K. Gehring, et al., Structural basis of defects in the sacsin HEPN domain responsible for autosomal recessive spastic ataxia of Charlevoix-Saguenay (ARSACS), J. Biol. Chem. 286 (2011) 20407–20412.
[27] H. Erlandsen, J.M. Canaves, I.A. Wilson, et al., Crystal structure of an HEPN domain protein (TM0613) from *Thermotoga maritima* at 1.75 A resolution, Proteins 54 (2004) 806–809.
[28] L. Slabinski, L. Jaroszewski, A. Godzik, et al., XtalPred: a web server for prediction of protein crystallizability, Bioinformatics 23 (2007) 3403–3405.
[29] K.S. Makarova, Y.I. Wolf, E.V. Koonin, Comprehensive comparative-genomic analysis of type 2 toxin–antitoxin systems and related mobile stress response systems in prokaryotes, Biol. Direct 4 (2009) 19.
[30] A. Kimelman, A. Levy, R. Sorek, et al., A vast collection of microbial genes that are toxic to bacteria, Genome Res. 22 (2012) 802–809.
[31] M. Grynberg, H. Erlandsen, A. Godzik, HEPN: a common domain in bacterial drug resistance and human neurodegenerative proteins, Trends Biochem. Sci. 28 (2003) 224–226.
[32] A. Padavannil, C. Jobichen, J. Sivaraman, et al., Dimerization of VirD2 binding protein is essential for agrobacterium induced tumor formation in plants, PLoS Pathog. 10 (2014).
[33] R. Chenna, H. Sugawara, J.D. Thompson, et al., Multiple sequence alignment with the clustal series of programs, Nucleic Acids Res. 31 (2003) 3497–3500.
[34] P. Gouet, E. Courcelle, D.I. Stuart, F. Metoz, ESPript: analysis of multiple sequence alignments in PostScript, Bioinformatics 15 (1999) 305–308.