9494–9504 *Nucleic Acids Research, 2016, Vol. 44, No. 19* *Published online 4 September 2016*
doi: 10.1093/nar/gkw776

Structural basis for single-stranded RNA recognition and cleavage by C3PO

Jing Zhang[1], Hehua Liu[1,2], Qingqing Yao[2], Xiang Yu[1,2], Yiqing Chen[1], Ruixue Cui[2], Baixing Wu[2], Lina Zheng[2], Junjun Zuo[1], Zhen Huang[3,4,*], Jinbiao Ma[2,*] and Jianhua Gan[1,*]

[1]Department of Physiology and Biophysics, School of Life Sciences, Fudan University, Shanghai 200438, China, [2]State Key Laboratory of Genetic Engineering, Collaborative Innovation Center of Genetics and Development, Department of Biochemistry, Institute of Plant Biology, School of Life Sciences, Fudan University, Shanghai 200438, China, [3]Department of Chemistry, Georgia State University, Atlanta, GA 30303, USA and [4]College of Life Sciences, Sichuan University, Chengdu 610041, China

Received June 1, 2016; Revised August 18, 2016; Accepted August 23, 2016

ABSTRACT

Translin and translin-associated factor-x are highly conserved in eukaroytes; they can form heteromeric complexes (known as C3POs) and participate in various nucleic acid metabolism pathways. In humans and *Drosophila*, C3POs cleave the fragmented siRNA passenger strands and facilitate the activation of RNA-induced silencing complex, the effector complex of RNA interference (RNAi). Here, we report three crystal structures of *Nanoarchaeum equitans* (*Ne*) C3PO. The apo-*Ne*C3PO structure adopts an open form and unravels a potential substrates entryway for the first time. The *Ne*C3PO:ssRNA and *Ne*C3PO:ssDNA complexes fold like closed football with the substrates captured at the inner cavities. The *Ne*C3PO:ssRNA structure represents the only catalytic form C3PO complex available to date; with mutagenesis and *in vitro* cleavage assays, the structure provides critical insights into the substrate binding and the two-cation-assisted catalytic mechanisms that are shared by eukaryotic C3POs. The work presented here further advances our understanding on the RNAi pathway.

INTRODUCTION

Translin (Tsn) and translin-associated factor-x (Trax) are members of the translin superfamily proteins (1,2). Tsn and Trax are highly conserved and coexist in many eukaryotes (3–6), including fission yeast, *Drosophila melanogaster* (*Drosophila*) and animals. They are involved in many important biological processes, such as normal cell growth, genome stability maintenance, neuronal development, spermatogenesis, RNA trafficking and RNA metabolism (7–11). Recently, it was demonstrated in *Homo sapiens* (*Hs*) and *Drosophila* that Tsn and Trax form heteromeric complex, also referred to as C3PO complex (12–14), which functions as the component 3 promoter of RISC (RNA-induced silencing complex) (15–19). RISC is the effector complex of RNA interference (RNAi), a conserved eukaryotic gene silencing mechanism triggered by small RNAs, such as small interfering RNAs (siRNAs) and microRNAs (miRNAs) (20–25). The RISC complex is assembled and functions in a stepwise manner (26–32). First, the siRNA duplex is loaded into the complex. Then, the passenger strand is cleaved by the argonaute protein (Ago, the catalytic component of RISC), resulting in nicked siRNA duplex. Finally, following the removal of the nicked passenger strand, the intact guide strand forms an active RISC complex with Ago, which degrades the target mRNA and causes the silencing.

In the RNAi pathways of higher eukaryotes, C3PO has been proposed to remove the nicked passenger strands and facilitates the formation of the active RISC complex (14). The crystal structures of *Drosophila* C3PO, *Hs*C3PO and C3PO-like from archaebacterium *Archaeoglobus fulgidus* (*Af*) have been reported (12,13,33). Unlike the truncated *Drosophila* C3PO, which adopts an open bowl-like overall fold, the full-length *Hs*C3PO assembles into a closed football-like shape. The molar ratio between *Hs*Tsn and *Hs*Trax subunits is 6:2 in the *Hs*C3PO structure. The *Af*C3PO complex, a homo-octamer of *Af*Trax, assembles similar to *Hs*C3PO. Six out of its eight subunits mimic eukaryotic Tsn (Tsn-like, referred to as TsnL), whereas the other two mimic eukaryotic Trax (Trax-like, referred to as TraxL). The two Trax (or TraxL) subunits in the C3PO complexes reside next to each other.

Though both Tsn and Trax subunits are important for the overall folding and substrate binding of C3PO, previous

*To whom correspondence should be addressed. Tel: +86 21 5163 0543; Fax: +86 21 5163 0541; Email: ganjhh@fudan.edu.cn
Correspondence may also be addressed to Jinbiao Ma. Tel: +86 21 5163 0542; Fax: +86 21 5163 0542; Email: majb@fudan.edu.cn
Correspondence may also be addressed to Zhen Huang. Tel: +1 404 413 5535; Fax: +1 404 413 5535; Email: huang@gsu.edu

mutagenesis studies showed that only the Trax subunit is responsible for the nuclease activity, which is divalent cation-dependent. It has been proposed that the catalytic center is composed of four conserved negative charged residues, corresponding to Glu123, Glu126, Asp193 and Glu197 in *Hs*Trax (13). These conclusions were further supported by the *Af*C3PO structure in complex with double-stranded (ds) RNA. Though human and *Drosophila* C3POs can also cleave some short dsRNA duplexes *in vitro*, they selectively cleave the two single-stranded (ss) RNA fragments resulting from the passenger strand degradation in the RNAi pathway. In *Neurospora carssa*, the C3PO has no significant role in the RNAi pathway, but it can degrade the 5′ pre-tRNA fragments, which are also single-stranded, during the tRNA-processing (9). Due to the lack of C3PO:ssRNA complex structure, the binding of C3PO and successive cleavage of ssRNA substrates is not clear. Similar to *Af*C3PO:dsRNA complex structure, the apo-structures of *Hs*C3PO and *Af*C3PO also adopt the closed football-like shape. Two distinct mechanisms were previously proposed for the formation of the catalytic C3PO:substrate complex: one is the 'dissociation and reassembly' model (12) and the other is the 'conformational switching' model (34). However, the conformational changes associated with the substrate entering (or product releasing from) the inner cavity (where the catalytic sites locate) of C3PO remain elusive.

To better understand the structural basis underlying the functions of C3PO, particularly the ssRNA substrates recognition and cleavage activity, we carried out biochemical and crystallographic studies of the Trax protein from *Nanoarchaeum equitans* (*Ne*) (35,36). *Ne*Trax is a predicted RNA-binding protein and is the only translin superfamily protein encoded by the genome of *N. equitans* (strain Kin4-M), which belongs to the *Nanoarchaeota* phylum of Archaea. *Ne*Trax shared some sequence similarity with both eukaryotic and bacterial translin superfamily proteins (Supplementary Figure S1). Compared with the Tsn subunits, *Ne*Trax was more similar to the Trax subunits of the eukaryotic C3POs. *In vitro* cleavage assays showed that *Ne*Trax has robust nuclease activity toward various types of substrates, including dsRNA, dsDNA, ssRNA and ssDNA. Three *Ne*Trax crystal structures were determined herein, including one apo-*Ne*Trax, one *Ne*Trax:ssRNA complex and one *Ne*Trax:ssDNA complex structure; overall shapes of *Ne*Trax complexes are similar to those of *Hs*C3PO and *Af*C3PO structures. Therefore, the *Ne*Trax protein is also referred to as *Ne*C3PO hereafter. The *Ne*C3PO complex contains six TsnL and two TraxL subunits; the two TraxL subunits are adjacently placed. Comparison of complex structures with the apo-*Ne*C3PO structure indicated that *Ne*C3PO can undergo large conformational change and create a wide gap between the two TraxL subunits, providing a plausible pathway for substrates entering (or products exiting from) the inner structural cavity. The substrate-binding mode and the catalytic mechanism were further verified by mutagenesis and *in vitro* catalytic assays, which also revealed the structural basis for the purine base preference of the translin superfamily proteins.

MATERIALS AND METHODS

DNA construction

The gene containing the codon-optimized cDNA (Supplementary Table S1) of the full-length wild-type (WT) *Ne*C3PO was purchased from GENERAY Biotechnology Co. Ltd. in Shanghai, China (http://www.geneary.com.cn). The target cDNA was amplified by polymerase chain reaction (PCR) and inserted into the BamHI/XhoI site of the plasmid pET28a to express His-*Ne*C3PO recombinant fusion protein. The plasmid DNA of WT *Ne*C3PO was used as template for PCR or overlap PCR reactions during the preparation of all *Ne*C3PO mutant constructs, including K19A, Q20A, S26A, K34A, E82Q, E85Q, D117N, E121Q, R124A, F160A, F160W, R163A, R164A, D167A and Y168A. The detailed sequences of the primers are also listed in Supplementary Table S1. Other procedures are similar to these utilized during the WT *Ne*C3PO DNA construction. The sequences of the plasmids were confirmed by DNA sequencing. All the recombinant strains are protected by 20% glycerol and stored at −80°C freezer till use.

Protein expression and purification

The constructed expression vector was transformed into *Escherichia coli* BL21(DE3) competent cell and cultured at 37°C for 12 h in Lysogeny broth (LB) medium containing 50 μg/ml kanamycin. Every 20 ml revived bacterium suspension was inoculated into 1 L LB medium supplemented with kanamycin (50 μg/ml) and cultured at 37°C with continuous shaking (200 rpm). When the OD_{600} reaches 0.6–0.8, the protein expression was induced by addition of isopropyl β-D-1-thiogalacto-pyranoside (IPTG, the final concentration is 0.2 mM). The induced cultures were then grown at 18°C for an additional 18 h.

The expressed protein was purified in sequential steps. The cells were harvested by centrifugation at 4500 rpm (Beckman Coulter AllegraTM 25R centrifuge, TA-14-250 rotor) for 20 min, and the pellet was resuspended in the lysis buffer (Buffer A, 20 mM Tris–HCl pH 8.0, 500 mM NaCl, 25 mM imidazole pH 8.0) and lysed under high pressure via JN-02C cell crusher. The homogenate was clarified by centrifugation (17 000 rpm) at 4°C for 1 h and the supernatant was loaded onto HisTrap™ column (GE Healthcare) pre-equilibrated with Buffer A. The His-*Ne*C3PO was eluted from the column using elution buffer (Buffer B, 20 mM Tris–HCl pH 8.0, 500 mM NaCl, 500 mM imidazole pH 8.0) with a stage-wise gradient. The protein was further purified by gel filtration using a Superdex® 200 16/60 preparation grade column (GE Healthcare) in a gel-filtration buffer (Buffer C, 20 mM Tris–HCl pH 8.0, 100 mM NaCl). All the mutant proteins were produced and purified using the same procedures.

Selenomethionine-substituted *Ne*C3PO protein (Se-*Ne*C3PO) was expressed using M9 medium supplemented with 60 mg/l Se-Met (Sigma-Aldrich). After cell disruption, the expressed fusion protein was recovered by centrifugation in the insoluble fraction. For protein purification, inclusion bodies were suspended in 20 ml cold denaturing buffer (Buffer D, 50 mM Mes pH 6.0, 500 mM NaCl, 6 M guanidine hydrochloride) and the Se-*Ne*C3PO protein was

purified using HisTrap™ column (GE Healthcare). For refolding, the purified Se-*Ne*C3PO protein was successively dialyzed for 3 h against Buffer E (50 mM Mes pH 6.0, 500 mM NaCl) containing 6, 4, 2, 1, 0.5, 0.25 or 0.1 M guanidine hydrochloride, respectively. A final dialysis step was performed overnight against Buffer E without guanidine hydrochloride. The protein was further purified by gel filtration using a Superdex® 200 16/60 preparation grade column (GE Healthcare) in buffer F (50 mM Mes pH 6.0, 100 mM NaCl). All proteins were concentrated and stored in aliquots at –80°C.

Crystallization and x-ray diffraction data collection

All nucleic acids utilized in the structural studies were synthesized and purified in the laboratory, and dissolved in ddH_2O. The initial crystallization conditions for all crystals were identified using the Gryphon crystallization robot system (Arts Robbin Instrument) and the commercial crystallization kits (Hampton Research). The sitting-drop vapor diffusion method was utilized during the initial screening at 16°C; all the crystallization condition optimization procedures were performed using hanging-drop vapor diffusion method. The sample contains 27 mg/ml and 7.5 mg/ml proteins for the apo-*Ne*C3PO structure and the Se-*Ne*C3PO structure, respectively. The crystallization condition is composed of 5% v/v tacsimate (pH 7.0), 0.1 M HEPES (pH 7.0) and 10% w/v Polyethlene glycol (PEG) monomethyl ether 5000 for apo-*Ne*C3PO crystals, whereas it contains 0.1 M succinic acid pH 7.0, 12% w/v PEG 3350 for the Se-*Ne*C3PO crystals. Samples of the two complexes were prepared by mixing *Ne*C3PO (27 mg/ml) and substrates (composed of 10 consecutive rA or dA, respectively) with a molar ratio of 16:3. The *Ne*C3PO:ssRNA sample also contained 10 mM $CaCl_2$. Both samples were incubated on ice for 30 min prior to crystallization. The reservoir solution is composed of 10% PEG 300, 0.1 M citric acid pH 4.2, 0.2 M NaCl for the *Ne*C3PO:ssDNA crystals and it is composed of 10% (w/v) PEG 8000, 100 mM imidazole/HCl pH 8.0, 200 mM $Ca(OAc)_2$ for the *Ne*C3PO:ssRNA crystals.

The crystals were cryo-protected with their mother liquids supplemented with 25% (v/v) glycerol and flash-cooled in a 100-K nitrogen stream. The X-ray diffraction data were collected at the BL17U and BL19U beamlines of the Shanghai Synchrotron Radiation Facility. All diffraction data processing was carried out using the MOSFLM program (37) embedded in the CCP4 suite (38), HKL2000 program or HKL3000 program (39). The data collection and processing statistics are summarized in Table 1.

Structure determination and refinement

The structure was solved using selenomethionine single-wavelength anomalous diffraction (SAD) data of Se-*Ne*C3PO with the Autosol program (40) embedded in the Phenix suit (41); the figure of merit value is 0.28. The initial model was built using the Autobuilt program. The intimal model was then refined against the diffraction data of apo-*Ne*C3PO using Refmac5 program (42) of CCP4, which revealed the detailed orientations for more protein residues. The $2F_o–F_c$ and $F_o–F_c$ electron density maps were regularly calculated and used as guide for the building of the missing amino acids and solvent molecules using COOT (43). About 5% randomly selected data was set aside for free R-factor cross validation calculations during the refinement. The two complex structures were solved by molecular replacement with the Phaser program in the CCP4 suite; one tetramer (composed of one TsnL/TsnL homodimer and one TsnL/TraxL heterodimer) of the apo-*Ne*C3PO structure was used as the search mode. Nucleic acids, cations, water and other molecules were all build manually using COOT. The two structures were also refined using Refmac5 program of CCP4. The structural refinement statistics were summarized in Table 1.

In vitro cleavage assays

FAM-labeled $d(GT)_{12}$ and $d(AC)_{12}$ were purchased from GENERAY Biotechnology Co. Ltd. in Shanghai (http://www.geneary.com.cn). The reaction system (10 μl) contains 10 μM protein (for WT *Ne*C3PO or mutants), 0.8 μM $d(GT)_{12}$ or $d(AC)_{12}$, 10 mM $MgCl_2$, 20 mM Tris–HCl pH 8.0 and 100 mM NaCl. The mixtures were incubated at 37°C for 10 min or 40 min (reaction system) and the reactions were quenched by addition of 10 μl termination buffer (90% formamide, 20 mM ethylenediaminetetraacetic acid, 0.05% bromophenol blue, 0.05% xylene blue). Six microliter samples were loaded onto pre-warmed 18% urea sequencing gel and run at 50–55 W, 48–50°C for 90 min. The gel was imaged using Typhoon FLA 9000.

Substrate binding assays

Binding of *Ne*C3PO over the substrates were monitored using electrophoretic mobility shift assays (EMSA). Five microliter *Ne*C3PO, 1 μl ssDNA, 1 μl 100 mM $CaCl_2$, 2 μl 5× binding buffer (100 mM Tris–HCl pH 8.0, 500 mM NaCl) and 1 μl ddH_2O were mixed in a thin-wall Eppendorf tube to 10 μl reaction system. The final concentrations of *Ne*C3PO and DNAs were indicated on the figures. The reaction mixtures were incubated on ice for 1 h. Five microliter samples were mixed with 1 μl 6× loading buffer (15% Ficoll 400, 0.05% Bromophenol Blue, 0.05% Xylene Cyanol) and loaded onto pre-cooled 6% native polyacrylamide gel [0.5× Trsi-borate-EDTA (TBE) buffer]. Gels were run at 100 V for 30 min at 4°C in 0.5× TBE buffer. The gel was imaged using Typhoon FLA 9000.

RESULTS

*Ne*C3PO possesses various nuclease activities

Various substrates, including dsRNA, dsDNA, ssRNA and ssDNA, were utilized in the *in vitro* cleavage assays. Both dsRNA (Supplementary Figure S2A) and dsDNA (Supplementary Figure S2B) can be cleaved by *Ne*C3PO in a size-dependent manner. After reaction time of 1.5 h, the 12-bp dsRNA and dsDNA were completely degraded by *Ne*C3PO; in contrast, almost no cleavage activities were observed for 18-bp dsRNA and dsDNA. The *in vitro* catalytic assays (Supplementary Figure S2C) further revealed that the cleavage activity of *Ne*C3PO is higher in dsDNA than

Table 1. Data collection and structural refinement statistics

	Se-*Ne*C3PO	apo-*Ne*C3PO	*Ne*C3PO:ssDNA	*Ne*C3PO:ssRNA
Data collection				
Wavelength (Å)	0.97928	1.0000	1.00000	1.0000
Space group	P1	P1	$P2_1$	C2
Unit cell				
a,b,c (Å)	86.5, 92.8, 92.9	86.4, 92.7, 92.7	98.8, 128.2, 102.2	165.4, 77.6, 87.8
α, β, γ (°)	114.1, 113.6, 95.4	114.3, 113.4, 95.8	90.0, 95.2, 90.0	90.0, 118.8, 90.0
Resolution rang (Å)[α]	30.0-2.5(2.59–2.5)	30.0-2.4(2.49–2.4)	30.0–2.1(2.21–2.1)	30.0–1.9(1.97–1.9)
Completeness (%)[α]	96.4(96.8)	95.4(92.0)	96.4(92.5)	96.8(94.7)
R_{sym} (%)[α]	9.1(44.9)	6.5(33.3)	10.8(42.8)	10.5(45.7)
$I/\sigma I$[α]	17.7(2.3)	16.5(2.3)	9.6(3.1)	16.3(2.9)
Redundancy	2.5(2.5)	2.5(2.2)	6.5(5.9)	4.0(3.0)
Refinement				
Resolution		30.0–2.4	30.0–2.1	29.07–1.9
R_{work} (%)		22.9	21.6	20.6
R_{free} (%)		25.2	25.9	23.0
No. of reflections		80 578	137 238	71 365
r.m.s.d. bonds (Å)		0.008	0.008	0.008
r.m.s.d. angles (°)		1.358	1.127	1.161
Ramachandran plot (%)				
Most favored		99.4	99.4	99.3
Additional allowed		0.6	0.6	0.7

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

dsRNA. These observations were different from *Af*C3PO, which preferentially binds and cleaves dsRNAs (33).

*Ne*C3PO can also cleave ssRNAs and ssDNAs. However, unlike the duplexes, the size of the ssRNA and ssDNA substrates of *Ne*C3PO can be very long; it degraded both 49-nt ssRNA and ssDNA, but with different efficiency. The *in vitro* cleavage reactions were analyzed on denaturing gel (Supplementary Figure S2D); as estimated from the gel, the amount of remaining ssRNA in the presence of 1 μM *Ne*C3PO was less than the ssDNA in the presence of 10 μM *Ne*C3PO, suggesting that the cleavage activity of *Ne*C3PO is at least 10-fold higher in ssRNA than ssDNA.

*Ne*C3PO and ssRNA recognition

To gain more insights into the ssRNA-cleavage mechanism of *Ne*C3PO complexes, we determined the crystal structure of the *Ne*C3PO:ssRNA complex (Figure 1A and B). The octameric *Ne*C3PO structure can be divided into two tetramers, each containing one TsnL/TsnL homodimer and one TsnL/TraxL heterodimer. These tetramers were related by a 2-fold crystallographic axis in-between the two TsnL/TraxL heterodimers. RNA A10 (composed of ten consecutive adenine residues) was utilized in the crystallization procedure; in the structure, two RNA strands were captured in the inner cavity and located next to the interface between the two TsnL/TraxL heterodimers (Figure 1C). The structure was refined up to 1.9 Å resolution, resulting in well-defined electron density for six adenine residues (A1-A6) and the phosphate group of A7 (Figure 1D); the rest three residues are disordered.

The two RNA molecules spanned about 54 Å in distance (measured from the N6 atom of A6 to the N6 atom of A6′) and were related by the two-fold axis (Figure 2A); therefore, the interactions between *Ne*C3PO and the two RNA strands were identical (Figure 2B). A1 and A2 (the first and second residues of RNA A10) did not interact with the protein residue; instead, they were sandwiched between A2′ and A1′, and A1′ and A4, respectively. Besides the stacking interactions, A2 also formed two water-mediated hydrogen-bonds (H-bond) with A4. The distance between the O2′ atom of A2 and the water molecule was 2.5 Å, and the distances between the water molecule and the N7 and OP2 atoms of A4 were 2.6 Å and 3.0 Å, respectively. Unlike other residues, A3 was significantly twisted; instead of stacking with other nucleotides, its sugar puckers (mainly the O4 atom) and nucleobase formed several stacking interactions with the side chains of Ile27, which was located in the middle of helix α1 of the TraxL subunits. The nucleobase of A3 also formed two H-bonds with amino acid residues, one (2.6 Å) between the N6 atom of A3 and the backbone O atom of Asn75 and the other (2.7 Å) between the N7 atom of A3 and the side chain OG atom of Ser26 of the TraxL subunit.

Unlike the first three residues, the phosphate backbones of A4 to A6 were also involved in the interactions with the protein. The OP1 atom of A4 and the OP2 atom of A5 interacted with the NZ atom of Lys34 of the TraxL subunit, and the distances between them were 2.7 Å and 2.8 Å, respectively. The OP2 atom of A6 interacted with the NH2 atom of Arg124 of the TraxL subunit, and the distance between them was 3.1 Å. The phosphate group was further stabilized by one strong H-bond (2.4 Å) between the OP1 atom of A6 and the OH group of Tyr168 of the TraxL subunit. In addition, the sugar puckers (mainly the O4 atom) and the nucleobase of A6 were stabilized by stacking interactions with the guanidyl group of Arg164 of the TsnL subunit. Hydrophobic stacking interactions were also observed between the nucleobase of A5 and the side chain of Phe160 of the TsnL subunit.

Two-cation-assisted ssRNA cleavage mechanism

Previous studies showed that the activities of eukaryotic C3POs (including *Drosophila* C3PO and *Hs*C3PO) and *Af*C3PO are divalent cation-dependent (14,33). In consistence with these observations, we showed that *Ne*C3PO

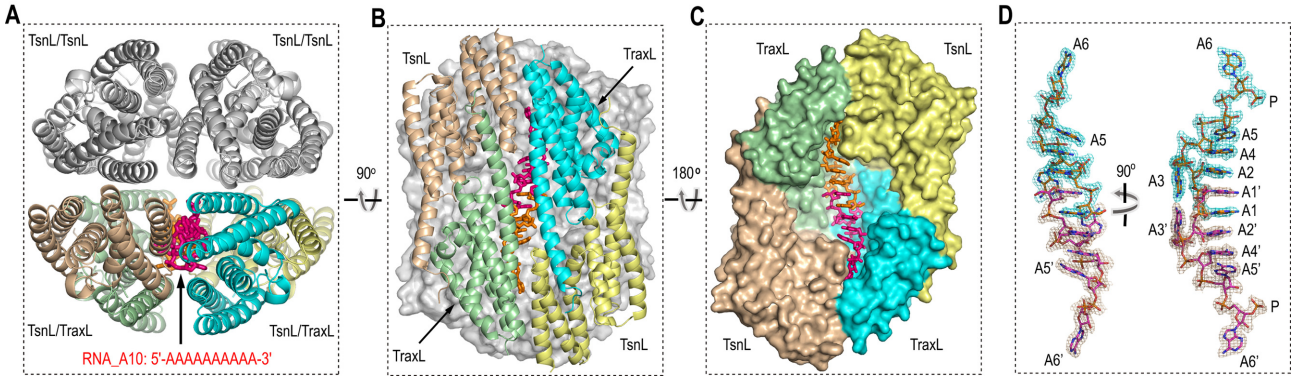

Figure 1. Overall structure of the *Ne*C3PO:ssRNA complex. (**A**) Cartoon representation showing the overall fold of the *Ne*C3PO:ssRNA complex. The two TsnL/TsnL homodimers are colored gray. TsnL and TraxL subunits of the two TsnL/TraxL heterodimers are colored brown and yellow, and green and cyan, respectively. The RNAs are shown as sticks in magenta and orange. (**B**) A 90° rotation of the image shown in A. The two TsnL/TsnL homodimers are shown as surface. (**C**) Surface and stick representation showing relative orientations between TsnL/TraxL heterodimers and RNAs. (**D**) Conformations of RNAs observed in the *Ne*C3PO:ssRNA complex structure. 2*Fo*–*Fc* electron density are contoured at 1.2σ and colored in cyan and brown for the two strands, respectively.

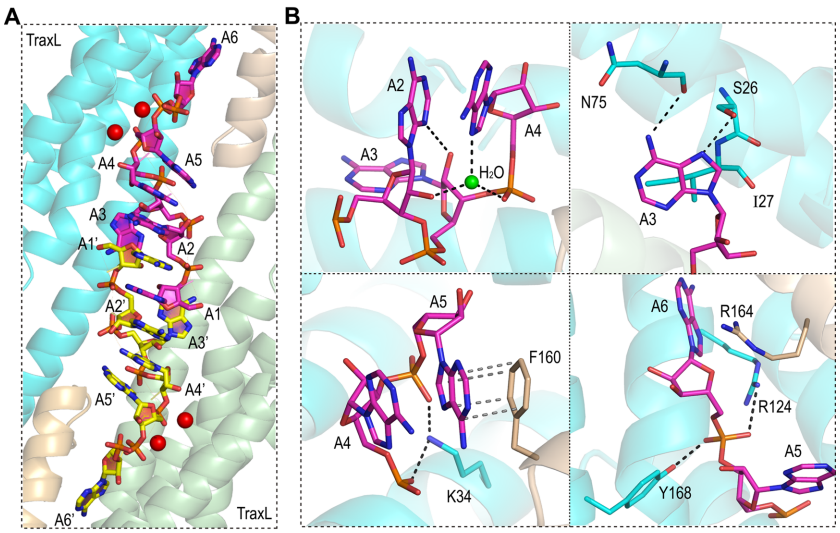

Figure 2. *Ne*C3PO and ssRNA recognition. (**A**) Close-up view showing the orientations of RNAs bound at the interface between the two TsnL/TraxL heterodimers. The RNAs are shown as sticks with the C atoms colored in magenta for one strand and yellow for another strand. (**B**) Detailed interactions between RNA and protein residues. Backbones of TraxL and TsnL subunits are shown as cartoons in cyan and light brown, respectively.

was active in the presence of Mg^{2+} or Mn^{2+}, but it was inactive in the presence of Ca^{2+} (Figure 3A). The crystallization solution of the *Ne*C3PO:ssRNA complex contained 200 mM $Ca(OAc)_2$ with two Ca^{2+} ions (Ca1 and Ca2) captured in each of the two catalytic sites (Figure 3B and Supplementary Figure S3A). The four conserved catalytic residues (Glu82, Glu85, Asp117 and Glu121) were all involved in the coordination of the two Ca^{2+} ions. These four residues formed two metal-binding sites, the prominent metal-binding site (M1) and the second metal-binding site (M2). Ca1 was located at the M1 site and was six-coordinated in an octahedral geometry. Beside its interaction with the OE2 atom of Glu85 and the OE1 atom of Glu121, Ca1 also coordinated with the OP1 atom (at S_p position) of A5 and three water molecules. Ca2 was located at the M2 site and was five-coordinated. Similar to Ca1, Ca2 also coordinated with A5 and Glu85 but with the OP1 atom and the OE1 atom, respectively; the other three coordinating interactions were between Ca2 and the OE1 atom of Glu82, the OD2 atom of Asp117, and one water molecule, respectively.

Like C3POs, the activities of many other RNA-cleaving nucleases, including RNase H and RNase III, are also divalent cation (Mg^{2+} and Mn^{2+})-dependent (44–47). Though the overall structures are very different, *Ne*C3PO and RNase H shared remarkable similarity at their active site composition and cation-coordination (Supplementary Figure S3B and C), suggesting that *Ne*C3PO also follows the two-cation-assisted mechanism in RNA-cleavage reaction. Via deprotonation, the M1 site cation activated the catalytic water molecule, which was located 3.4 Å away from the 5′ phosphorus of the scissile bond between A4 and A5. Also similar to RNase III and RNase H, one H-bond was formed between the OP2 atom (at R_p position) of A6 and the catalytic water molecule in the *Ne*C3PO structure, which

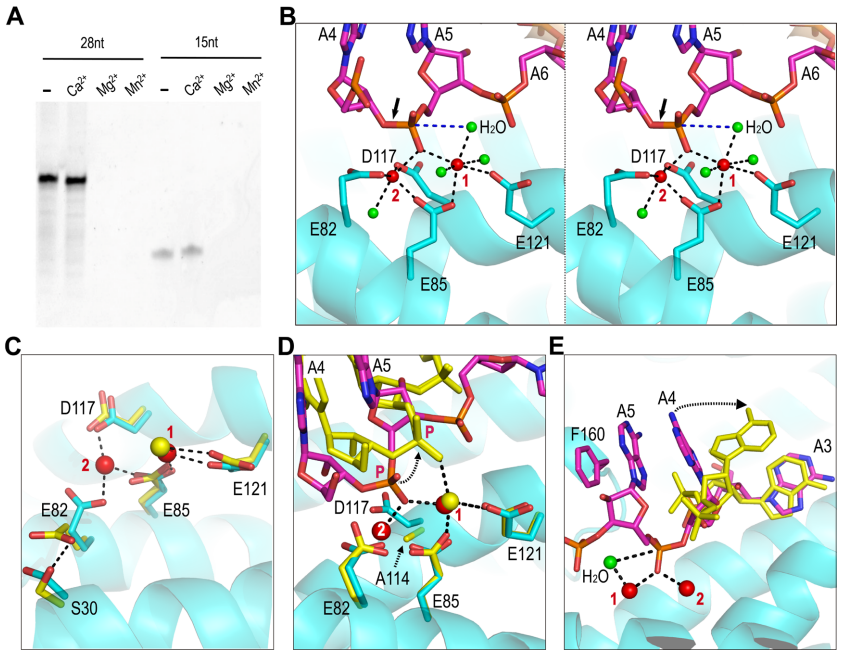

Figure 3. Cation-assisted cleavage activity of *Ne*C3PO. (**A**) *In vitro* cleavage assays performed in the presence of various cations. (**B**) Stereoview showing coordinations of the Ca^{2+} located in the TraxL subunits of the *Ne*C3PO:ssRNA structure. The TraxL subunit is shown as cartoon in cyan, and catalytic residues and RNAs are shown as sticks with the C atoms colored in cyan and magenta, respectively. Ca^{2+} ions and water molecules are shown as red and green spheres, respectively. (**C**) Comparison of the TraxL and TsnL subunits of the *Ne*C3PO:ssRNA structure showing the conformational changes of the catalytic residues. Comparisons of the *Ne*C3PO:ssRNA structure with the *Af*C3PO:dsRNA structure (**D**) and the *Ne*C3PO:ssDNA structure (**E**) showing the impact of cations on catalytic complex assembly. In C and D, catalytic residues of *Ne*C3PO TsnL subunit and *Af*C3PO TraxL subunit are shown as sticks with C atoms colored in yellow, and the associated cations are shown as yellow spheres. In D and E, the dsRNA and ssDNA bound at the TraxL subunits of *Af*C3PO:ssRNA and *Ne*C3PO:ssDNA structures are shown as sticks in yellow. Backbones of the TsnL subunit of *Ne*C3PO:ssRNA, TraxL subunit of *Af*C3PO:dsRNA and TraxL subunit of *Ne*C3PO:ssDNA structures are omitted for clarity in C–E.

helped orient the water molecule for the in-line attack on the 5′ phosphorus. M2 site cation coordinated with the OP1 atom of A5 and it was only 3.0 Å away from the 3′ leaving oxygen; it can facilitate the bond breakage by neutralizing the negative charge developing on the oxygen anion.

In the *Ne*C3PO:ssRNA structure, Ca^{2+} ions were also observed at the metal-binding sites of the six TsnL subunits (Figure 3C); however, unlike the TraxL subunits, only one Ca^{2+} ion was located at the M1 site and coordinated with the side chains of Glu85 and Glu121, which adopted similar conformations in all the eight subunits. Both Asp117 and Glu82 did not coordinated with any cation in the TsnL subunits. The conformations of Asp117 were conserved in the structure; however, when compared with the TraxL subunits, the side chains of Glu82 of the TsnL subunits rotated about 180° around the CB-CG bond and formed one H-bond (2.5 Å) with the OG atom of Ser30. M1 site cations were also observed in the *Af*C3PO:dsRNA structure (Figure 3D); however, due to the mutation of D114 (corresponding to Asp117 of *Ne*C3PO), the M2 site cations were not observed in the structure, the two residues flanking the scissile bond all shifted away from the catalytic sites (33). In the *Ne*C3PO:ssDNA complex structure (Figure 3E) grown in the absence of $CaCl_2$, the residue corresponding to A5 of the *Ne*C3PO:ssRNA structure was disordered. The one corresponding to A4 was stable but the nucleobase was tilted about 90°, when compared with the A4 of the *Ne*C3PO:ssRNA structure. These observations indicated that, besides catalysis, the cations and their precise interactions with the conserved protein residues were also critical for the catalytic complex assembly of C3POs.

Cleavage assay and purine preference

To provide more insights into the catalytic mechanism of *Ne*C3PO and further verify the substrate binding modes observed in the two complex structures, we constructed several *Ne*C3PO mutants and carried out *in vitro* cleavage assays using two FAM-labeled ssDNAs, $(AC)_{12}$ (Figure 4A) and $(GT)_{12}$ (Figure 4B). Among the four cation-coordinating residues, Glu82, Asp117, and Glu121 were involved in the coordination of one cation, whereas Glu85 coordinated with two cations (Figure 3B). Interestingly, when compared with the WT protein, the cleavage activities of E82Q, D117N and E121Q mutants were much weaker; the activity of E85Q was also weaker than that of the WT but much stronger than those of E82Q, D117N and E121Q. These observations suggested that Glu85 was mainly involved in the coordination of the two cations.

Among the nucleic acid interacting residues, mutations of Lys19 had no obvious impact on the cleavage activity of the protein (K19A). However, alanine substitutions of Ser26 and Arg164 all led to weakened cleavage activities of the mutant proteins (S26A and R164A), compared with the WT. Mutations of other residues, including Lys34, Arg124, Arg163 and Tyr168, caused more significant reduction on the cleavage activities of the mutants (K34A,

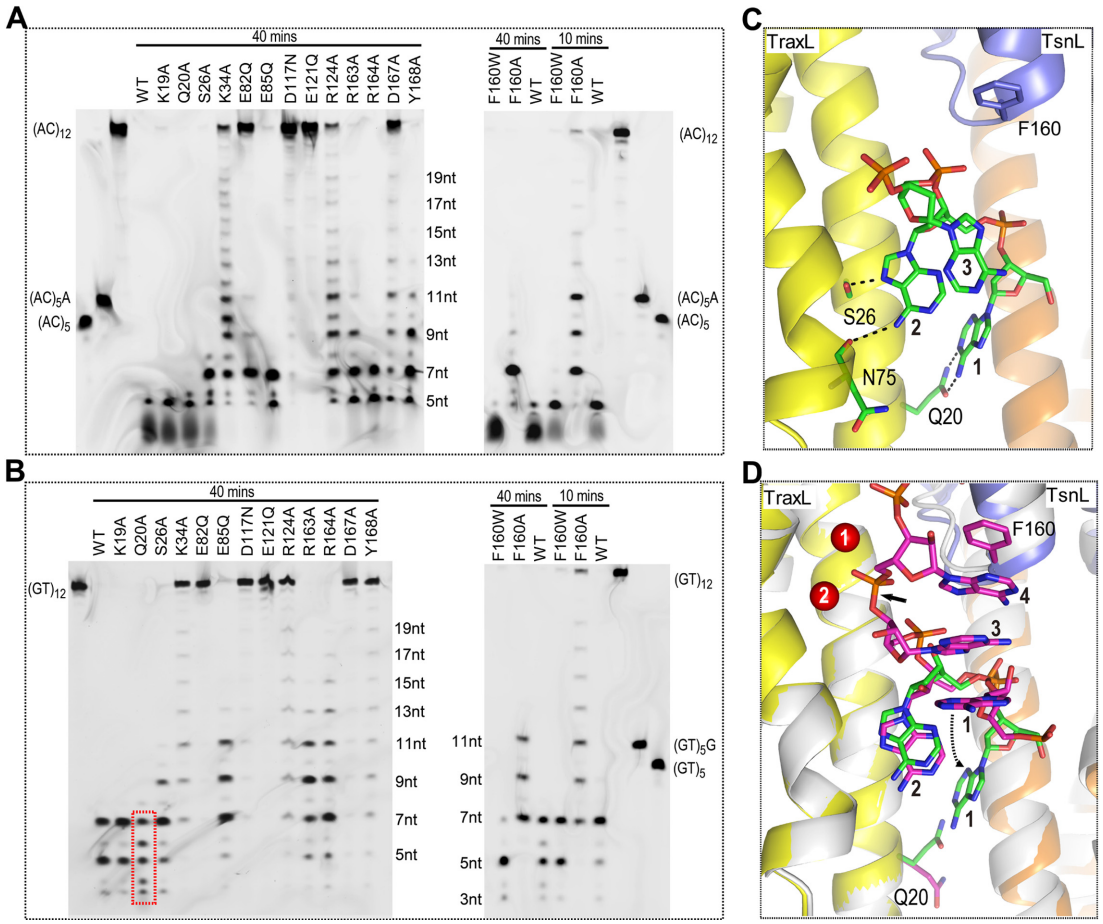

Figure 4. Substrate cleavage and purine preference of *Ne*C3PO. *In vitro* $(AC)_{12}$ (**A**) and $(GT)_{12}$ (**B**) cleavage reactions catalyzed by WT *Ne*C3PO and mutants. (**C**) Interactions between ssDNA and *Ne*C3PO observed in the *Ne*C3PO:ssDNA structure. Protein backbones are shown as cartoons in yellow, brown and blue. ssDNA and the interacting residues are shown as sticks with C atoms colored in green. (**D**) Structural superposition of *Ne*C3PO:ssDNA and *Ne*C3PO:ssRNA. The *Ne*C3PO:ssRNA structure is colored white for the protein backbone and magenta for C atoms of the protein side chains and ssRNA. The *Ne*C3PO:ssDNA structure is colored as seen in C.

R124A, R163A and Y168A). These results were all consistent with the structural observations.

Interestingly, the *in vitro* cleavage assays with both substrates showed an obvious pattern of products. Comparison with the FAM-labeled markers, including $(AC)_5$ and $(AC)_5A$ in Figure 4A, and $(GT)_5$ and $(GT)_5G$ in Figure 4B, indicated that *Ne*C3PO preferentially cut after the purine residues. In the *Ne*C3PO:ssRNA structure, the side chain of Phe160 stacked with the nucleobase of A6 (or A6′). Replacing Phe160 with residues that has larger (Trp160 for F160W) or smaller (Ala160 for F160A) side chains showed certain enhancing or weakening effect on the overall cleavage activity of the mutant proteins, compared with the WT. However, the mutant proteins still selectively cut after the purine residues. These observations indicated that Phe160 plays important role in the substrate binding, but is not the main cause of *Ne*C3PO preference to purine.

In the *Ne*C3PO:ssDNA structure, the side chain of Asn20 of the TraxL subunit interacted with A1 of the substrate forming two H-bonds, one (∼2.9 Å) between the OE1 atom of Asn20 and the N6 atom of A1 and the another (∼2.9 Å) between the NE2 atom of Asn20 and the N1 atom of A1 (Figure 4C). Replacing Asn20 with Ala20 had no obvious impact on the overall cleavage rate; however, compared with the WT, the mutation caused obvious change on the product pattern (Figure 4B). Besides the 5-nt $(GT)_2G$ and 7-nt $(GT)_3G$ products, there were also significant amount of 4-nt $(GT)_2$ and 6-nt $(GT)_3$ products generated in the presence of Q20A mutant with the intensities of $(GT)_2G$, $(GT)_3$, and $(GT)_3G$ comparable with each other. These results showed that Gln20 plays critical role in the cleavage site selection of *Ne*C3PO.

As analyzed on the denaturing gel (Figure 4B), the shortest product produced by *in vitro* *Ne*C3PO cleavage reaction was three nucleotides in length. Superimposition of *Ne*C3PO:ssRNA and *Ne*C3PO:ssDNA structures (Figure 4D) can generate a plausible model that can explain the formation of the 3-nt product. Again, Gln20 was involved in the product formation by interacting with the first nucleotide and orienting it in a conformation similar to A1 in the *Ne*C3PO:ssDNA structure. The 2nd nucleotide can be modeled based on either structure. In-between the third and fourth nucleotides, whose conformations were similar

to those of the *Ne*C3PO:ssRNA structure, was the cleavage site.

Substrate entryway and associated conformational changes

Similar to the *Hs*C3PO and *Af*C3PO structures, the *Ne*C3PO:ssRNA and *Ne*C3PO:ssDNA structures also folded into closed football-like shape with the catalytic residues located on the surface of the inner cavity (Figure 5A). On the surface of the complexes, there were four gaps between the four dimers; however, all the gaps were too narrow (<6 Å in width). Substrates, such as RNA A10 and DNA A10 captured in the complex structures, and the 14-bp dsDNA and dsRNA utilized in the *in vitro* cleavage assays, were not allowed to access the interior cavity through these narrow gaps.

Unlike the two complex structures, the apo-*Ne*C3PO structure has an open barrier conformation (Figure 5B), indicated that the C3PO complex formation may follow the 'conformational switching' mechanism. Among the four gaps on the surface of *Ne*C3PO, three were between TsnL subunits while the fourth was between TraxL subunits. The sizes of the three TsnL–TsnL gaps were comparable to those in the complex structures, but the TraxL–TraxL gap (∼17 Å in width) is much wider than the other three. This gap represents a plausible substrate entryway for the single-stranded nucleic acids and we speculated that it can undergo conformational changes to create an even wider gap to accommodate other substrates in solution, including the A-form RNA duplex (∼18 Å in diameter). As revealed by structural comparisons, the overall fold of the four dimers (especially the two TsnL/TsnL homodimers) of the apo-*Ne*C3PO structure was very similar to that of the complex structures. The conformational differences mainly occurred on the interface between the two TsnL/TraxL heterodimers due to the tilt of the TsnL subunit and the rotation of the TraxL subunit (Figure 5C).

Besides the overall orientation, structural comparisons also revealed significant conformational differences in the residues located on the two heterodimer interfaces. In the apo-*Ne*C3PO structure, the residues were far away from each other. Due to the lack of interaction, they became very flexible, exhibiting weak electron density and high B values. However, in the *Ne*C3PO:ssRNA complex structure, these residues were all well-defined and form extensive interactions between each other, including two salt bridges and one H-bond (Figure 5D). One salt bridge is between Glu121 of the TraxL subunit and Arg163 of the TsnL subunit; the other is between Arg124 of the TraxL subunit and Asp167 of the TsnL subunit, which is involved in another salt-bridge interaction with Arg164 of the same TsnL subunit. The H-bond (2.7 Å) is formed between the OH group of Tyr38 of the TraxL subunit and the NE atom of Arg163 of the TsnL subunit. Compared to the *Ne*C3PO:ssRNA structure, the conformations of Tyr38, Glu121 and Arg124 of TraxL subunit were similar in the *Ne*C3PO:ssDNA complex structure, but they did not interact with the side chains of Arg163 and Arg164 of the TsnL subunit, which all point toward the N-terminus of helix α6 (the last helix of *Ne*C3PO) (Figure 5D). Besides the side chains, the backbone of α6 also shifted by 1.2 Å, which might due to the disordering of the nucleotide A5 that stacks with Phe160 from downside in the *Ne*C3PO:ssRNA structure. These observations suggested that the interactions between the local protein residues and the nucleic acid substrate are also critical for the catalytic C3PO complex assembly.

Additional substrate binding sites

Besides the common substrate binding sites located at the interface between the two TsnL/TraxL heterodimers, the *Ne*C3PO:ssDNA structure also revealed several additional nucleic acid binding sites (Supplementary Figure S4A), which can be divided into two groups (group A and group B). The group A was composed of four sites, which were all located at the interfaces between the two TsnL molecules within the TsnL/TsnL homodimers (Supplementary Figure S4B) and has similar interactions with the nucleobases (Supplementary Figure S4C). Formed by the side chains of Met71 and Phe72 from one TsnL molecule and the side chains of Tyr73, Phe77, Trp109 and Phe113 from the partner TsnL molecule, the nucleobase binding sites were very hydrophobic in nature, which formed extensive hydrophobic interactions with the nucleobases of the adenines. Moreover, the N6 atom of the nucleobase also formed one H-bond (∼2.8 Å) with the backbone O atom of Met71, which further stabilized the nucleobase.

Similar to group A, the group B also contained four nucleobase binding sites; however, their locations and conformations were different (Supplementary Figure S4D). Among the four sites, two were located at the interface between the two TsnL/TraxL heterodimers; the orientations and interactions of the bound adenine nucleobases were similar to those of A6 and A6′ observed in the *Ne*C3PO:ssRNA structure (Figure 2). The other two binding sites were between the TsnL and TraxL subunits within the TsnL/TraxL heterodimers and exhibited similar interactions with the nucleobases as the group A binding sites.

As revealed by the *Ne*C3PO:ssRNA structure, the major substrate binding site can accommodate eight nucleobases (including A1, A2, A4, A5, A1′, A2′, A4′ and A5′) stacking on each other. *Ne*C3PO can cleave the 49-nt ssRNA and ssDNA substrates very efficiently (Supplementary Figure S2D), whereas it only has very low cleavage activity toward the 77-nt ssDNA (Supplementary Figure S2E). Most likely, the 49-nt substrates are cleaved by *Ne*C3PO in the closed form. The additional nucleic acid binding sites may involve in the binding and packing of the substrates into the inner cavity. The 77-nt substrates may be too long to be efficiently packed into the inner cavity of *Ne*C3PO in the closed form, which led to lower efficiency in the catalytic process.

DISCUSSION

In this study, we have performed the structural and biochemical studies of *Ne*C3PO. Three structures were determined: one apo-*Ne*C3PO structure and two complex structures (*Ne*C3PO:ssRNA complex and *Ne*C3PO:ssDNA complex). Both complex structures folded like a closed football with the substrates bound inside the inner cavity of the structures. One plausible substrate entryway is revealed by the apo-*Ne*C3PO structures, which adopts the open form

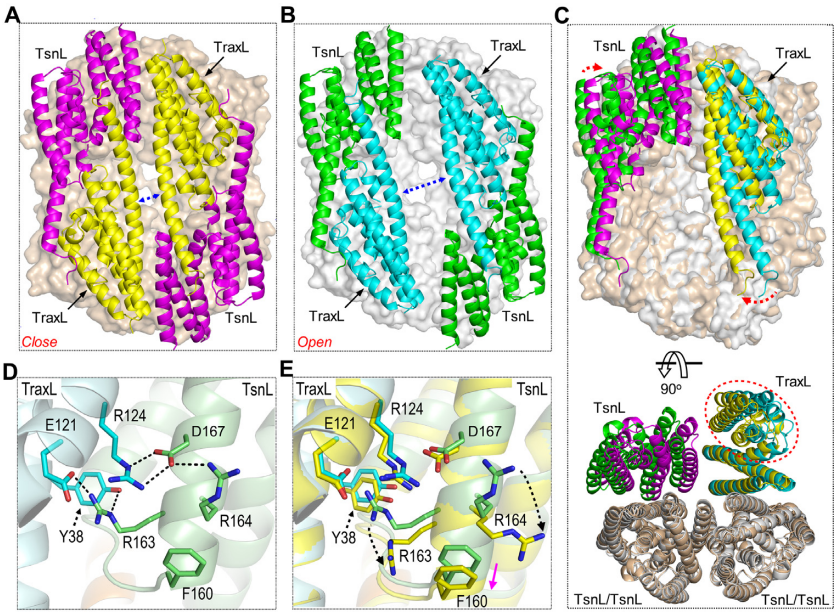

Figure 5. Conformational changes and substrate entryway. Overall folds of the closed (**A**) and open (**B**) form of *Ne*C3PO, based on the *Ne*C3PO:ssRNA and the apo-*Ne*C3PO structures, respectively. (**C**) Superposition showing conformational changes of TsnL and TraxL subunits. (**D**) Interactions between TsnL and TraxL subunits observed in the *Ne*C3PO:ssRNA structure. (**E**) Superposition showing conformational changes of conserved Arg residues. In A–C, TsnL/TsnL homodimers are colored light brown and white for the closed and open forms, respectively. TsnL and TraxL subunits are colored magenta and yellow, and green and cyan for the closed and open structures, respectively. For clarity, one TsnL and one TraxL subunit are omitted in C. In D and E, protein backbones and C atoms of the side chains are colored green and cyan, respectively, for TsnL and TraxL subunits of the closed form, whereas both subunits of the open form are colored yellow in E.

and has one wide gap between the two TraxL subunits (Figure 5B). In light of the recently reported open form *Hs*Tsn structure (34), our apo- and complex structures of *Ne*C3PO suggested that 'conformational switching' may be a common feature of the translin superfamily proteins. Though it has not been captured in any crystal structure of eukaryotic C3PO, we speculated that the eukaryotic complexes may also adopt the open football form during substrate loading.

The topology and subunit assembly of the *Ne*C3PO structures are very similar to those of the *Hs*C3PO structure (Supplementary Figures S5A and C). The sequence of *Ne*C3PO is also similar to the eukaryotic translin superfamily proteins, especially the Trax proteins. The identity between *Ne*C3PO and *Hs*Trax is about ∼22% (Supplementary Figure S1). Besides the four residues (Glu82, Glu85, Asp117 and Glu121) involved in cation coordination, other residues, such as Ser26, Lys34 and Arg124 involved in substrate binding in the *Ne*C3PO:ssRNA structure (Figure 2B), are also conserved in the Trax subunits of *Hs*C3PO complex (Supplementary Figure S1). These structure and sequence conservations indicated that the TraxL subunits of *Ne*C3PO and the Trax subunits of eukaryotic C3PO complexes may share the similar mechanisms in the binding and cleavage of the substrates, such as ssRNAs.

Like the Trax subunits, the Tsn subunits of eukaryotic C3PO also adopt similar overall fold as the *Ne*C3PO protein (Supplementary Figure S5B). Due to the lack of the catalytic residues required for the phosphodiester hydrolysis, the Tsn subunits of the C3PO are not involved in the substrate cleavage reaction, but they are necessary for substrate binding. The *Ne*C3PO:ssRNA structure (Figure 2B) showed that two residues, Phe160 and Arg164, of the TsnL subunits of *Ne*C3PO form stacking interactions with the substrates. Although the Phe160 is not conserved, the *Hs*Tsn has a Lys residue (Lys193) at the position corresponding to Arg164 of *Ne*C3PO. *Hs*Tsn also has an Arg residue (Arg192) corresponding to Arg163 of *Ne*C3PO. In the *Ne*C3PO:ssRNA structure, Arg164 interacts with A6 of the substrate (Figure 2B). Though Arg163 does not form direct interaction with the substrate, it forms one H-bond with the side chain of Glu121 (Figure 5D). As shown by the *in vitro* cleavage assays (Figure 4A and B), replacing Arg163 or Arg164 by Ala caused obvious reduction on the catalytic activity of the mutant proteins (R163A and R164A). Similar results were also observed during the mutagenesis studies of Arg192 and Lys193 of the Tsn subunits of *Hs*C3PO, suggesting that the two conserved residues may play similar functional roles in eukaryotes and prokaryotes.

The eukaryotic Trax proteins are unstable on their own, but the eukaryotic Tsn proteins are stable and can form homo-octameric structures. *Hs*Tsn can bind both ssRNAs and ssDNAs and it has very high affinity (up to nM level) toward certain substrates (such as GU-repeating sequence). As revealed by *in vitro* cleavage assays, *Ne*C3PO can cleave single-stranded nucleic acids; the EMSA assays further showed that *Ne*C3PO has higher binding affinity toward the GT-repeating sequence (especially longer sequences), compared to the AC-repeating sequence (Supplementary Figure S6). The overall folds of the closed form *Ne*C3PO and the *Hs*Tsn homomeric structures are similar to each other, suggesting that they may utilize common strategies in substrate binding, such as positively charged residues. In the

*Ne*C3PO:ssDNA structure, Gln20 interacts with the nucleobase of A1 of the substrate (Figure 4C). Guanine has one N atom (N1) at the position 1 and one O atom (O6) at the position 6; compared with the N1 and N6 atoms of adenine, the N1 and O6 atoms of guanine are more suitable in forming H-bonds with the OE1 and NE2 atoms of Gln20, which may contribute to the stronger binding affinity of the GT-repeating sequences (Supplementary Figure S6). The eukaryotic Tsn proteins have one Glu residue at the position corresponding to Gln20 of *Ne*C3PO. Although it needs to be further verified, the highly conservation indicates that this Glu residue may play certain role in the GU-repeating sequence binding.

Besides *N. equitans*, the gene encoding a single translin superfamily protein has also been found in many other prokaryotes (48), including five archaea phyla (*Crenarchaeota*, *Euryarchaeota*, *Thaumarchaeota*, *Nanoarchaeota* and *Korarchaeota*) and the bacterial phylum Chloroflexi (which represents an early diverging clade of eubacteria and is near the root of the tree of life) (49). The sequence similarities of bacterial proteins are very high (up to 60%) and they also share very high similarity with archaeal proteins, including *Af*C3PO that adopts similar topology and overall fold as *Ne*C3PO. The most obvious conformational differences between *Ne*C3PO and *Af*C3PO are located at the α5–α6 linker and the N-terminus of α6 (Supplementary Figures S5B–D). Compared to *Ne*C3PO, the α5-α6 linker of *Af*C3PO is longer by four residues; among them, the linker residue Lys158 is involved in the dsRNA substrate interaction. *N. equitans* belongs to the *Nanoarchaeota* phyla of archaea, and it has the fastest evolutionary rate among the archaea kingdom (36). The identity between *Ne*C3PO and *Af*C3PO is <15% (Supplementary Figure S1). Compared with *Af*C3PO, *Ne*C3PO is more similar to eukaryotic homologs in sequence and size of the α5–α6 linker. All these observations suggested that, most likely, the strong ssRNA cleavage activities of *Ne*C3PO and eukaryotic C3PO are the consequence of evolution. Similar to *Ne*C3PO, some eukaryotic C3POs, such as *Drosophila* C3PO, also have weak ssDNA-cleaving capacity (12). However, no dsDNA cleavage activity has been reported for any C3POs previously; whether this activity is unique to *Ne*C3PO or common to the eukaryotic C3PO proteins and the functional role of this cleavage needs to be further investigated in the future.

ACCESSION NUMBERS

Structure factors and coordinates have been deposited in the Protein Data Bank under accession codes 5JR9, 5JRC and 5JRE for apo-*Ne*C3PO, *Ne*C3PO:ssRNA and *Ne*C3PO:ssDNA structures, respectively.

SUPPLEMENTARY DATA

Supplementary Data are available at NAR Online.

ACKNOWLEDGEMENTS

We thank the BL17U and BL19U beamline staff at the Shanghai Synchrotron Radiation Facility for assistance during data collection and Prof. Xinhua Ji and members of the Gan and Ma laboratories for insightful discussions.

FUNDING

National Natural Science Foundation of China [31370728, 31230041, 2157020958]; National Basic Research and Sichuan S&T Programs of China [2011CB966304, 2012CB910502, 16GH0134]; USA NIH [R01GM095881, R42ES026935]. Funding for open access charge: National Natural Science Foundation of China [31370728].
Conflict of interest statement. None declared.

REFERENCES

1. Aoki,K., Ishida,R. and Kasai,M. (1997) Isolation and characterization of a cDNA encoding a Translin-like protein, TRAX. *FEBS Lett.*, **401**, 109–112.
2. Aoki,K., Suzuki,K., Sugano,T., Tasaka,T., Nakahara,K., Kuge,O., Omori,A. and Kasai,M. (1995) A novel gene, translin, encodes a recombination hotspot binding-protein associated with chromosomal translocations. *Nat. Genet.*, **10**, 167–174.
3. Finkenstadt,P.M., Jeon,M. and Baraban,J.M. (2002) Trax is a component of the Translin-containing RNA binding complex. *J. Neurochem.*, **83**, 202–210.
4. Gupta,G.D., Makde,R.D., Kamdar,R.P., D'Souza,J.S., Kulkarni,M.G., Kumar,V. and Rao,B.J. (2005) Co-expressed recombinant human Translin-Trax complex binds DNA. *FEBS Lett.*, **579**, 3141–3146.
5. Laufman,O., Ben Yosef,R., Adir,N. and Manor,H. (2005) Cloning and characterization of the Schizosaccharomyces pombe homologs of the human protein Translin and the Translin-associated protein TRAX. *Nucleic Acids Res.*, **33**, 4128–4139.
6. Claussen,M., Koch,R., Jin,Z.Y. and Suter,B. (2006) Functional characterization of Drosophila Translin and Trax. *Genetics*, **174**, 1337–1347.
7. Li,Z., Wu,Y. and Baraban,J.M. (2008) The Translin/Trax RNA binding complex: clues to function in the nervous system. *Biochim. Biophys. Acta*, **1779**, 479–485.
8. Wu,X.Q. and Hecht,N.B. (2000) Mouse testis brain ribonucleic acid-binding protein/translin colocalizes with microtubules and is immunoprecipitated with messenger ribonucleic acids encoding myelin basic protein, alpha calmodulin kinase II, and protamines 1 and 2. *Biol. Reprod.*, **62**, 720–725.
9. Li,L.D., Gu,W.F., Liang,C.Y., Liu,Q.H., Mello,C.C. and Liu,Y. (2012) The translin-TRAX complex (C3PO) is a ribonuclease in tRNA processing. *Nat. Struct. Mol. Biol.*, **19**, 824–830.
10. Yang,S.C., Cho,Y.S., Chennathukuzhi,V.M., Underkoffler,L.A., Loomes,K. and Hecht,N.B. (2004) Translin-associated factor X is post-transcriptionally regulated by its partner protein TB-RBP, and both are essential for normal cell proliferation. *J. Biol. Chem.*, **279**, 12605–12614.
11. Yang,S.C. and Hecht,N.B. (2004) Translin associated protein X is essential for cellular proliferation. *FEBS Lett.*, **576**, 221–225.
12. Tian,Y., Simanshu,D.K., Ascano,M., Diaz-Avalos,R., Park,A.Y., Juranek,S.A., Rice,W.J., Yin,Q., Robinson,C.V., Tuschl,T. *et al.* (2011) Multimeric assembly and biochemical characterization of the Trax-translin endonuclease complex. *Nat. Struct. Mol. Biol.*, **18**, U658–U651.
13. Ye,X.C., Huang,N.A., Liu,Y., Paroo,Z., Huerta,C., Li,P., Chen,S., Liu,Q.H. and Zhang,H. (2011) Structure of C3PO and mechanism of human RISC activation. *Nat. Struct. Mol. Biol.*, **18**, U650–U643.
14. Liu,Y., Ye,X.C., Jiang,F., Liang,C.Y., Chen,D.M., Peng,J.M., Kinch,L.N., Grishin,N.V. and Liu,Q.H. (2009) C3PO, an endoribonuclease that promotes RNAi by facilitating RISC activation. *Science*, **325**, 750–753.
15. Kaya,E. and Doudna,J.A. (2012) Guided tour to the heart of RISC. *Science*, **336**, 985–986.
16. MacRae,I.J., Ma,E., Zhou,M., Robinson,C.V. and Doudna,J.A. (2008) In vitro reconstitution of the human RISC-loading complex. *Proc. Natl. Acad. Sci. U.S.A.*, **105**, 512–517.
17. Preall,J.B. and Sontheimer,E.J. (2005) RNAi: RISC gets loaded. *Cell*, **123**, 543–545.

18. Gregory,R.I., Chendrimada,T.P., Cooch,N. and Shiekhattar,R. (2005) Human RISC couples microRNA biogenesis and posttranscriptional gene silencing. *Cell*, **123**, 631–640.
19. Ameres,S.L., Martinez,J. and Schroeder,R. (2007) Molecular basis for target RNA recognition and cleavage by human RISC. *Cell*, **130**, 101–112.
20. Drinnenberg,I.A., Weinberg,D.E., Xie,K.T., Mower,J.P., Wolfe,K.H., Fink,G.R. and Bartel,D.P. (2009) RNAi in budding yeast. *Science*, **326**, 544–550.
21. van Rij,R.P. and Andino,R. (2004) RNAi—a guide to gene silencing. *Science*, **303**, 1978–1979.
22. Tabara,H., Grishok,A. and Mello,C.C. (1998) RNAi in C-elegans: soaking in the genome sequence. *Science*, **282**, 430–431.
23. Vastenhouw,N.L., Brunschwig,K., Okihara,K.L., Muller,F., Tijsterman,M. and Plasterk,R.H.A. (2006) Long-term gene silencing by RNAi. *Nature*, **442**, 882–882.
24. Lu,R., Maduro,M., Li,F., Li,H.W., Broitman-Maduro,G., Li,W.X. and Ding,S.W. (2005) Animal virus replication and RNAi-mediated antiviral silencing in Caenorhabditis elegans. *Nature*, **436**, 1040–1043.
25. Sijen,T. and Plasterk,R.H.A. (2003) Transposon silencing in the Caenorhabditis elegans germ line by natural RNAi. *Nature*, **426**, 310–314.
26. Sontheimer,E.J. and Carthew,R.W. (2004) Argonaute journeys into the heart of RISC. *Science*, **305**, 1409–1410.
27. Song,J.J., Smith,S.K., Hannon,G.J. and Joshua-Tor,L. (2004) Crystal structure of argonaute and its implications for RISC slicer activity. *Science*, **305**, 1434–1437.
28. Wang,Y.L., Sheng,G., Juranek,S., Tuschl,T. and Patel,D.J. (2008) Structure of the guide-strand-containing argonaute silencing complex. *Nature*, **456**, U209–U234.
29. Wang,Y.L., Juranek,S., Li,H.T., Sheng,G., Tuschl,T. and Patel,D.J. (2008) Structure of an argonaute silencing complex with a seed-containing guide DNA and target RNA duplex. *Nature*, **456**, U921–U972.
30. Nakanishi,K., Weinberg,D.E., Bartel,D.P. and Patel,D.J. (2012) Structure of yeast Argonaute with guide RNA. *Nature*, **486**, 368–374.
31. Elkayam,E., Kuhn,C.D., Tocilj,A., Haase,A.D., Greene,E.M., Hannon,G.J. and Joshua-Tor,L. (2012) The structure of human argonaute-2 in complex with miR-20a. *Cell*, **150**, 100–110.
32. Maiti,M., Lee,H.C. and Liu,Y. (2007) QIP, a putative exonuclease, interacts with the Neurospora Argonaute protein and facilitates conversion of duplex siRNA into single strands. *Genes Dev.*, **21**, 590–600.
33. Parizotto,E.A., Lowe,E.D. and Parker,J.S. (2013) Structural basis for duplex RNA recognition and cleavage by Archaeoglobus fulgidus C3PO. *Nat. Struct. Mol. Biol.*, **20**, 380–386.
34. Eliahoo,E., Marx,A., Manor,H. and Alien,A. (2015) A novel open-barrel structure of octameric translin reveals a potential RNA entryway. *J. Mol. Biol.*, **427**, 756–762.
35. Waters,E., Hohn,M.J., Ahel,I., Graham,D.E., Adams,M.D., Barnstead,M., Beeson,K.Y., Bibbs,L., Bolanos,R., Keller,M. *et al.* (2003) The genome of Nanoarchaeum equitans: Insights into early archaeal evolution and derived parasitism. *Proc. Natl. Acad. Sci. U.S.A.*, **100**, 12984–12988.
36. Brochier,C., Gribaldo,S., Zivanovic,Y., Confalonieri,F. and Forterre,P. (2005) Nanoarchaea: representatives of a novel archaeal phylum or a fast-evolving euryarchaeal lineage related to Thermococcales? *Genome Biol.*, **6**, R42.
37. Battye,T.G., Kontogiannis,L., Johnson,O., Powell,H.R. and Leslie,A.G. (2011) iMOSFLM: a new graphical interface for diffraction-image processing with MOSFLM. *Acta Crystallogr. D Biol. Crystallogr.*, **67**, 271–281.
38. Potterton,E., Briggs,P., Turkenburg,M. and Dodson,E. (2003) A graphical user interface to the CCP4 program suite. *Acta Crystallogr. D Biol. Crystallogr.*, **59**, 1131–1137.
39. Minor,W., Cymborowski,M., Otwinowski,Z. and Chruszcz,M. (2006) HKL-3000: the integration of data reduction and structure solution–from diffraction images to an initial model in minutes. *Acta Crystallogr. D Biol. Crystallogr.*, **62**, 859–866.
40. Terwilliger,T.C., Adams,P.D., Read,R.J., McCoy,A.J., Moriarty,N.W., Grosse-Kunstleve,R.W., Afonine,P.V., Zwart,P.H. and Hung,L.W. (2009) Decision-making in structure solution using Bayesian estimates of map quality: the PHENIX AutoSol wizard. *Acta Crystallogr. D Biol. Crystallogr.*, **65**, 582–601.
41. Adams,P.D., Grosse-Kunstleve,R.W., Hung,L.W., Ioerger,T.R., McCoy,A.J., Moriarty,N.W., Read,R.J., Sacchettini,J.C., Sauter,N.K. and Terwilliger,T.C. (2002) PHENIX: building new software for automated crystallographic structure determination. *Acta Crystallogr. D Biol. Crystallogr.*, **58**, 1948–1954.
42. Murshudov,G.N., Skubak,P., Lebedev,A.A., Pannu,N.S., Steiner,R.A., Nicholls,R.A., Winn,M.D., Long,F. and Vagin,A.A. (2011) REFMAC5 for the refinement of macromolecular crystal structures. *Acta Crystallogr. D Biol. Crystallogr.*, **67**, 355–367.
43. Emsley,P. and Cowtan,K. (2004) Coot: model-building tools for molecular graphics. *Acta Crystallogr. D Biol. Crystallogr.*, **60**, 2126–2132.
44. Gan,J., Shaw,G., Tropea,J.E., Waugh,D.S., Court,D.L. and Ji,X. (2008) A stepwise model for double-stranded RNA processing by ribonuclease III. *Mol. Microbiol.*, **67**, 143–154.
45. Gan,J.H., Tropea,J.E., Austin,B.P., Court,D.L., Waugh,D.S. and Ji,X.H. (2006) Structural insight into the mechanism of double-stranded RNA processing by ribonuclease III. *Cell*, **124**, 355–366.
46. Nowotny,M., Gaidamakov,S.A., Crouch,R.J. and Yang,W. (2005) Crystal structures of RNase H bound to an RNA/DNA hybrid: substrate specificity and metal-dependent catalysis. *Cell*, **121**, 1005–1016.
47. Mueser,T.C., Nossal,N.G. and Hyde,C.C. (1996) Structure of bacteriophage T4 RNase H, a 5′ to 3′ RNA-DNA and DNA-DNA exonuclease with sequence similarity to the RAD2 family of eukaryotic proteins. *Cell*, **85**, 1101–1112.
48. Gupta,G.D., Kale,A. and Kumar,V. (2012) Molecular evolution of translin superfamily proteins within the genomes of eubacteria, archaea and eukaryotes. *J. Mol. Evol.*, **75**, 155–167.
49. Cavalier-Smith,T. (2006) Rooting the tree of life by transition analyses. *Biol. Direct*, **1**, 19.