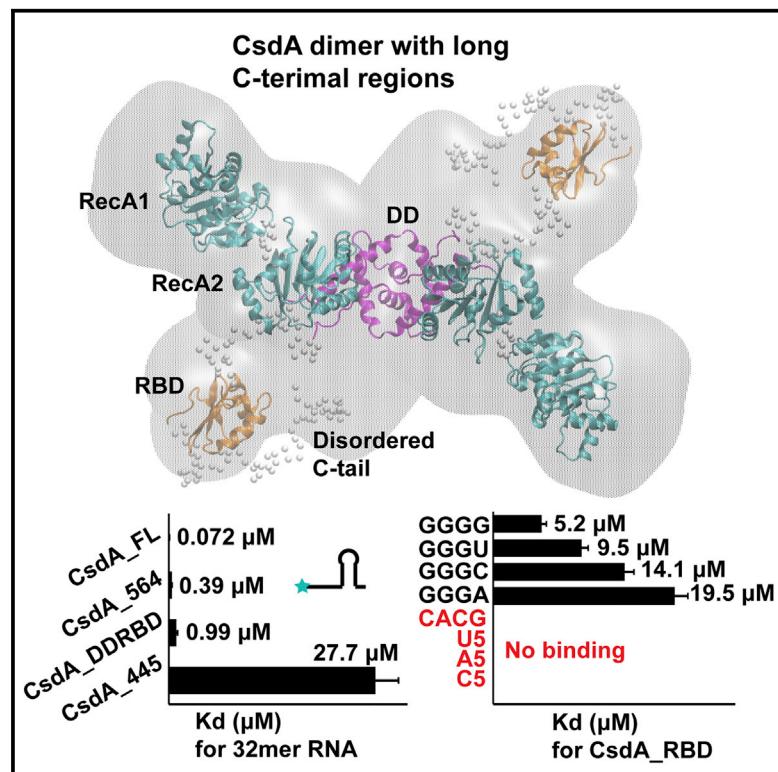


Structure

Insights into the Structure of Dimeric RNA Helicase CsdA and Indispensable Role of Its C-Terminal Regions

Graphical Abstract



Authors

Ling Xu, Lijun Wang, Junhui Peng, ..., Jihui Wu, Yajun Tang, Yunyu Shi

Correspondence

tangyj@ustc.edu.cn (Y.T.),
yyshi@ustc.edu.cn (Y.S.)

In Brief

Xu et al. demonstrate for the first time that DEAD-box RNA helicase CsdA is a stable dimer at low temperature. The long, flexible C-terminal regions of CsdA are essential for high enzymatic activity and strong RNA-binding affinity, and the RNA-binding domain prefers binding single-stranded G-rich RNA.

Highlights

- Ribosome biogenesis-related RNA helicase CsdA forms a dimer at low temperature
- SAXS data reveal that CsdA is flexible with long intrinsic disorder regions
- The C-terminal regions are critical for RNA binding and high enzymatic activities
- The RBD prefers binding to single-stranded G-rich RNA joined by H92 of 23S rRNA

Insights into the Structure of Dimeric RNA Helicase CsdA and Indispensable Role of Its C-Terminal Regions

Ling Xu,¹ Lijun Wang,¹ Junhui Peng,¹ Fudong Li,¹ Lijie Wu,² Beibei Zhang,¹ Mengqi Lv,¹ Jiahai Zhang,¹ Qingguo Gong,¹ Rongguang Zhang,^{2,3} Xiaobing Zuo,⁴ Zhiyong Zhang,¹ Jihui Wu,¹ Yajun Tang,^{1,*} and Yunyu Shi^{1,5,*}

¹Hefei National Laboratory for Physical Science at Microscale and School of Life Sciences, University of Science and Technology of China, Hefei, Anhui 230026, People's Republic of China

²National Center for Protein Science Shanghai, Shanghai Institutes of Biological Sciences, Chinese Academy of Sciences, Shanghai 200031, China

³National Laboratory of Biomacromolecules, Institute of Biophysics, Chinese Academy of Sciences, Beijing 100101, China

⁴X-ray Science Division, Advanced Photon Source, Argonne National Laboratory, Argonne, IL 60349, USA

⁵Lead Contact

*Correspondence: tangyj@ustc.edu.cn (Y.T.), yyshi@ustc.edu.cn (Y.S.)

<https://doi.org/10.1016/j.str.2017.09.013>

SUMMARY

CsdA has been proposed to be essential for the biogenesis of ribosome and gene regulation after cold shock. However, the structure of CsdA and the function of its long C-terminal regions are still unclear. Here, we solved all of the domain structures of CsdA and found two previously uncharacterized auxiliary domains: a dimerization domain (DD) and an RNA-binding domain (RBD). Small-angle X-ray scattering experiments helped to track the conformational flexibilities of the helicase core domains and C-terminal regions. Biochemical assays revealed that DD is indispensable for stabilizing the CsdA dimeric structure. We also demonstrate for the first time that CsdA functions as a stable dimer at low temperature. The C-terminal regions are critical for RNA binding and efficient enzymatic activities. CsdA_RBD could specifically bind to the regions with a preference for single-stranded G-rich RNA, which may help to bring the helicase core to unwind the adjacent duplex.

INTRODUCTION

DEAD-box proteins belong to a ubiquitous family of RNA helicases, which are widely found from prokaryotes to eukaryotes and participate in multiple cellular processes, such as pre-mRNA splicing, translation initiation, modulating RNA-protein complexes, RNA decay, and ribosome biogenesis (Linder, 2006; Linder and Jankowsky, 2011; Charollais et al., 2004). DEAD-box proteins of helicase superfamily II consist of a helicase core with two canonical tandem RecA-like domains that contain several conserved motifs and one of them is the D-E-A-D motif (Linder et al., 1989). The I, Ia, Ib, II, and III motifs are located in the first RecA-like domain while motifs IV, V, and VI

are in the second RecA-like domain, which provide RNA and ATP binding sites to remodel high-order structures of RNAs and RNA/protein complexes (Jankowsky and Fairman, 2007; Linder, 2006). In addition to their helicase cores, DEAD-box proteins contain auxiliary extensions in their N terminus and/or C terminus, which are thought to be essential to their specific recognition and interaction with different RNAs or proteins (Silverman et al., 2003; Yan et al., 2003). DEAD-box RNA helicases display common biochemical characteristics in that they have both ATPase and unwinding activities and ATP and RNA bind cooperatively to stabilize the flexible helicase core (Theissen et al., 2008; Lorsch and Herschlag, 1998; Polach and Uhlenbeck, 2002). According to these biochemical characteristics, it is believed that DEAD-box RNA helicases take advantage of the energy of ATP hydrolysis to unwind the secondary structures of RNAs (Sengoku et al., 2006; Del Campo and Lambowitz, 2009).

There are five DEAD-box RNA helicases, CsdA (cold-shock DEAD-box protein A, also named DeaD), DbpA, SrmB, RhIE, and RhIB, in *Escherichia coli* (*E. coli*), among which the first four have been reported to be involved in ribosome biogenesis. The *csdA* gene was originally recognized as a multi-copy suppressor of a gene mutation that encoded r-protein S2 (Toone et al., 1991). CsdA is associated with the pre50S particles and is involved in the biogenesis of the 50S subunit (Charollais et al., 2004), and plays an essential role in the survival of *E. coli* at low temperature, where, in strains with the deletion of the gene of CsdA protein, this results in a deficit in the biogenesis of the 50S subunit and an accumulation of 40S particles corresponding to an incompletely assembled 50S subunit (Charollais et al., 2004). Besides, there is evidence that the RNA degradosome is modified during cold shock, CsdA can replace DEAD-box protein Rh1B when expressed at a physiological level (Beran and Simons, 2001); and can be accompanied with RNase E and PNPase to form RNA degradosome (Py et al., 1996; Miczak et al., 1996). In addition, the overexpression of CsdA can stabilize inefficiently translated mRNA in *E. coli* (Iost and Dreyfus, 1994). Recently, it was reported that the translation of *rpoS* mRNA in *E. coli* at low temperature, which encodes stationary-phase

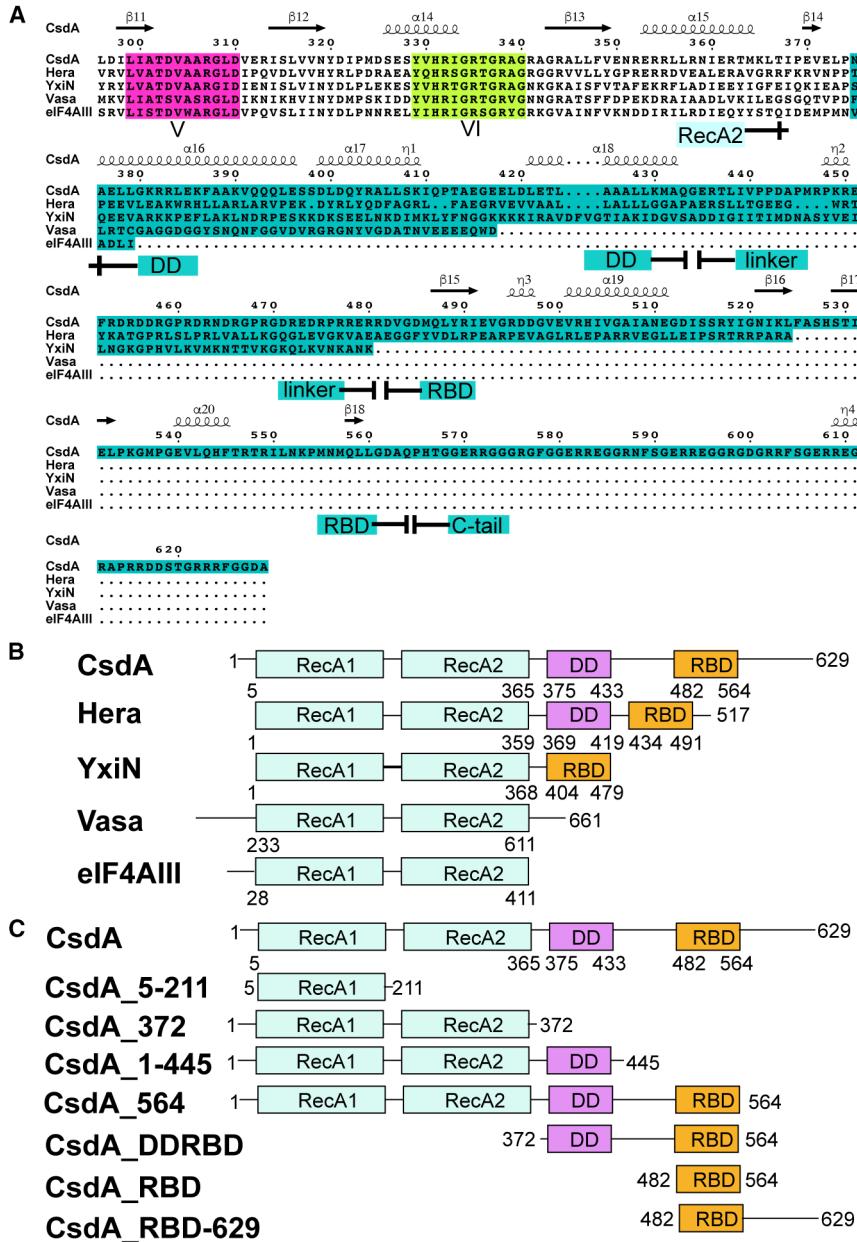


Figure 1. Sequence Alignment and Architecture of Different DEAD-Box Proteins

(A) Sequence alignment of the C terminus of DEAD-box proteins from *E. coli* CsdA, *T. thermophilus* Hera, *B. subtilis* YxiN, *D. melanogaster* Vasa, and human eIF4AIII. The alignment was created with Multalin. The C-terminal regions are highlighted in light blue. The conserved motifs in RecA2 are highlighted in magenta and lemon.

(B) Schematic presentation of domains in the different DEAD-box proteins. All of the DEAD-box proteins comprise of two tandem RecA-like domains (RecA1 and RecA2, colored in pale cyan). The different colors show the different regions, DD is colored in purple and RBD is colored in orange. (C) Domain schematic presentation of domains of different CsdA constructs. See also Figure S1.

conquer approach to study the structure of CsdA. We showed that CsdA not only consists of two tandem RecA-like domains but also contains previously uncharacterized dimerization domain (DD) and RNA-binding domain (RBD). We determined the structures of all domains of CsdA by X-ray crystallography or nuclear magnetic resonance (NMR). In addition, small-angle X-ray scattering (SAXS) assay helped us build ensemble structural models of CsdA_1–445 (residues 1–445) and CsdA_FL (full-length CsdA). *In vivo* and *in vitro* experiments suggested that DD is indispensable for the stability of CsdA. Our studies demonstrated for the first time that CsdA is a stable dimer at cold-shock situations. The ATPase and unwinding activity assays as well as the fluorescence polarization (FP) experiments indicated important roles of the C-terminal regions of CsdA for binding RNAs. Moreover, we found that CsdA_RBD prefers binding to single-stranded G-rich RNA, which may bring the helicase core to unwind the adjacent RNA duplex. Our findings further expand

the versatility of RBDs as an RNA-binding domain when compared with other RBDs.

RESULTS

The Complex Structure of RecA1 with AMP

The helicase core of CsdA is comprised of two RecA-like domains (RecA1 and RecA2) joined by a flexible linker and contains all conserved motifs (Figure S1). Although these DEAD-box proteins contain conserved helicase cores, they show significant differences in lengths and arrangements of their N/C-terminal extensions (Figures 1A and 1B), which would be related to their functions (Mohr et al., 2008; Rudolph and Klostermeier, 2009; Klostermeier and Rudolph, 2009). We tested crystallization of

and general stress response factor RpoS (σ^S), was not only regulated by the small RNA DsrA and the RNA chaperon Hfq, but was also activated by CsdA (Resch et al., 2010). The *csdA* gene encodes a multi-domain protein of 629 amino acids. Sequence alignments show that CsdA has a conserved core of ~365 amino acids and a long uncharacterized C-terminal region (Figures S1 and 1). The helicase core is related to ATP binding and RNA unwinding, while the C-terminal region may be involved in RNA recognition (Turner et al., 2007; Bizebard et al., 2004). However, the constitution of the C-terminal regions and the structure of CsdA are currently unclear.

Due to the flexible linkers in these proteins, a limited number of whole-protein structures have been determined for RNA helicases with multiple domains. Here, we used a divide-and-

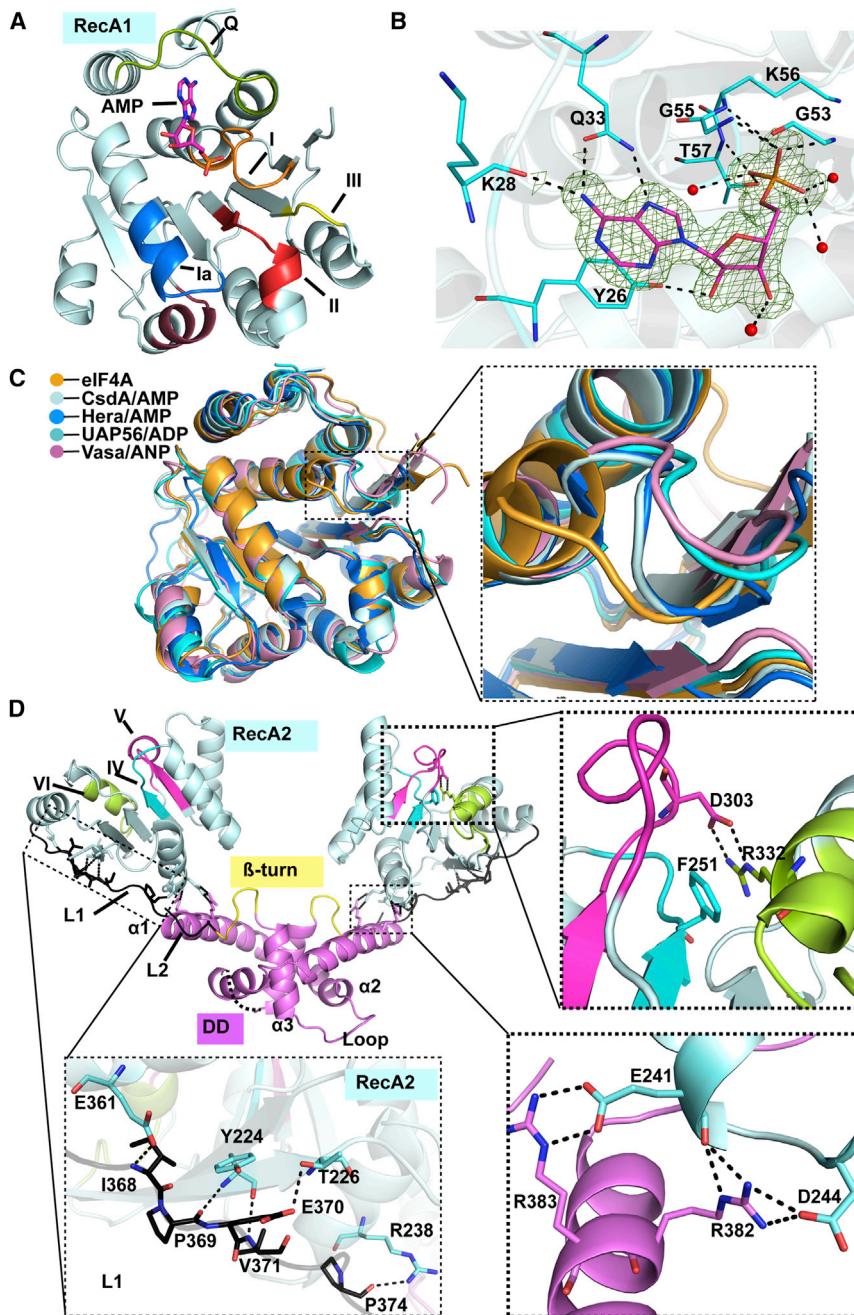


Figure 2. The Overall Domain Structures of CsdA

(A) Crystal structure of CsdA_RecA1/AMP complex. The consensus motifs are indicated by different colors. The AMP is shown in sticks.

(B) Close-up view of the interactions between AMP and CsdA_RecA1. Water molecules are shown as red spheres. Residues involved in polar interactions between proteins and AMP are indicated by dashed lines. The $2F_o - F_c$ electron density map of AMP was contoured at 1.0σ .

(C) Superposition of solved crystal structure of RecA1 of eIF4A (orange), CsdA/AMP (pale cyan), Hera/AMP (marine), UAP56/ADP (cyan), and Vasa/ANP (pink). The bound nucleotides are not shown. The sites that are nucleotides bound are highlighted in the box to the right.

(D) The crystal structure of CsdA_218–445. The consensus motifs V and VI are labeled in different colors. RecA2 is colored pale cyan and DD is violet. Residues F251 of motif IV stacks on R332 of motif VI, and D303 of motif V connects to R332 through two hydrogen bonds, as indicated by a dashed line and are highlighted in the dashed box in the upper panel. Residues involved in the interaction of RecA2 and L1 (black) are highlighted in the dash box of left bottom panel. Residues E241 and D244 in RecA2 are connected with residues R383 and R382 in DD through hydrogen bonds, as indicated by the dashed line and highlighted in the dashed box in the right bottom panel. Helices in DD are labeled as α_1 , α_2 , and α_3 . The missing loop of DD in one of the two chains is indicated using a broken line. The β turn between DD and the L2 linker (black) is colored in yellow. See also Figure S1.

the full-length and several different truncated constructs of CsdA (Figure 1C) and obtained single crystals for CsdA_5–211 (residues 5–211) (Figure 1C). However, due to the protein degradation in the process of crystallization, we only obtained single crystals for CsdA_218–445 (residues 218–445), although we used CsdA_1–445 (residues 1–445) constructs (Figure 1C) in crystallization.

We determined the structure of CsdA_5–211 bound to AMP (Figure 2A) by X-ray diffraction, hereafter referred to as CsdA_5–211/AMP complex. The structure of the complex was determined and refined to a final R factor of 0.1668 ($R_{\text{free}} = 0.2035$) at 1.6 \AA resolution by molecular replacement (Table 1),

using *Thermus thermophilus* RNA helicase Hera (PDB: 2GXS) as the model. The structure of CsdA_5–211 displays a RecA-like fold that is a typical module of RNA helicases and folds into a seven-stranded parallel β sheet surrounded by seven α helices. It contains six conserved motifs, Q-motif, I, Ia, Ib, II, and III, found throughout a large number of the first RecA-like domain of other DEAD-box RNA helicases (Cordin et al., 2006) (Figure S1). The AMP is seen in the electron density map (Figure 2B). Residues Y26, K28, and Q33 in the Q-motif interact with the adenine moiety of AMP, while the phosphate group is bound to residues G53, G55, K56, and T57 in motif I (also known as the P loop or the Walker A motif) (Figure 2B). Structural superposition shows that the whole structure of CsdA_5–211/AMP complex is similar to many other DEAD-box proteins bound with or without AMP, ADP, or AMP-PNP, except for some structural differences, especially in the P loop (Figure 2C). The P loop of the CsdA_5–211/AMP complex closely resembles that of Hera/AMP (PDB: 2GXQ) (Rudolph et al., 2006), and both loops move outside compared with the ligand-free eIF4A (PDB: 1FUU) (Caruthers et al., 2000) and move inside compared with UAP56/ADP (PDB: 1XTJ) (Shi et al., 2004) while the P loop of Vasa/ANP

Table 1. Crystallographic Data and Refinement Statistics

	CsdA_5-211/AMP	CsdA_218-445
Crystal Parameters		
Space group	P6 ₁	P2 ₁ 2 ₁ 2 ₁
a, b, c (Å)	57.99, 57.99, 104.36	55.62, 86.99, 117.85
α, β, γ (°)	90, 90, 120	90, 90, 90
Data Collection Statistics		
Wavelength (Å)	0.979	0.979
Resolution	28.99–1.60 (1.66–1.60) ^a	40–2.24 (2.33–2.24) ^a
R _{merge}	0.084 (0.591)	0.074 (0.607)
I/σI	12.50 (3.40)	27.89 (2.40)
Completeness	99.8 (100)	98.9 (95.9)
Redundancy	6.2	5.3
Refinement Statistics		
Protein atoms	1,554	3,531
B factors (Å ²)	20.78	58.70
AMP/water	1/205	0/183
B factors (Å ²)	24.25/32.34	–/54.05
R _{factor} /R _{free}	0.1668/0.2035	0.225/0.268
RMSD		
Bond lengths (Å)	0.010	0.009
Bond angles (°)	1.158	1.110
Ramachandran plot (% residues)		
Most favored	193 (99.0)	412 (95.0)
Additional allowed	2 (1.0)	21 (4.8)
Disallowed regions	0 (0.00)	2 (0.46)

^aValues in parentheses are for the highest-resolution shell. RMSD, root mean-square deviation.

(PDB: 2DB3) (Sengoku et al., 2006) moves a little outside when compared with that of UAP56/ADP. These step-by-step movements induced by different nucleotides may give appropriate conformations for binding and unwinding RNAs (Rudolph et al., 2006).

The CsdA_218–445 Dimer

The structure of CsdA_218–445 was determined by X-ray diffraction and refined to an R factor of 0.225, with an R_{free} of 0.268 at a resolution of 2.3 Å by molecular replacement (Table 1), using Hera (PDB: 3EAQ) as the model. The structure of CsdA_218–445 includes two RecA-like domains (RecA2) and two DDs, which form a V-shape dimer (Figure 2D). The V-shape conformation of the CsdA_218–445 dimer in solution was further confirmed by SAXS experiments (Figures 5 and S5). In the crystal structure, residues between DD and the partial L2 linker form a well-defined β turn structure (Figure 2D). Significantly, CsdA was never identified as a dimer in previous studies (Charollais et al., 2004; Iost and Dreyfus, 2006; Resch et al., 2010). RecA2 is located from residues 218 to 366, and the conserved motifs IV, V, and VI in the RecA2 of CsdA (Figure S1) lay the foundations for the helicase core. As in other DEAD-box RNA helicases, the RecA2 of CsdA consists of a central six-stranded parallel β sheet surrounded by six α helices (Figure 2D), and the conserved

motifs IV, V, and VI of RecA2 of CsdA are located close to each other in space. Some of the residues in these motifs interact with each other; for example, residue F251 in motif IV adopts π-cation interactions with R332 in motif VI, while the side chain of D303 in motif V interacts with R332 by hydrogen bonds (Figure 2D). The interaction network of RecA2 lays the foundation for the construction of the helicase core of CsdA. In addition, residues I368, P369, E370, V371, and P374 in the L2 linker interact with residues E361, Y224, T226, and R238 in RecA2 through hydrogen bonds, which stabilize the orientation of L2 to RecA2 (Figure 2D), while the hydrogen bonds formed by residues E241 and D244 in CsdA_RecA2, and R382 and R383 in CsdA_DD, seem to stabilize the orientation of CsdA_RecA2 to CsdA_DD (Figure 2D).

DD forms a dimer by approximately 2-fold symmetry and interlaces via α helices with extensive hydrogen bond interactions (Figure 3A) and tightly packed hydrophobic interactions (Figure 3B). The main hydrophobic interface of the dimer, comprised of the two α3 helices, especially residues L424, A425, L428, and A432 (Figure 3B), forms the hydrophobic core of DD. A Dali search for structural homologs (Holm and Rosenstrom, 2010), using CsdA_DD as a query, resulted in a distantly similar structure, *Helicobacter pylori* Hp0242 protein (Z score of 3.5, PDB: 2BO3) (Tsai et al., 2006) of unknown function. CsdA_DD and *T. thermophilus* helicase Hera_DD (PDB: 3EAQ) (Klostermeier and Rudolph, 2009) have a Z score of 3.9 (Figure 3C) by using Dali pairwise comparison. Recently, the crystal structure of the *Geobacillus stearothermophilus* DEAD-box RNA helicase CshA (PDB: 5IVL) has been reported, which also contains a DD (Huen et al., 2017). CsdA_DD and CshA_DD have a Z score of 4.0. Structural superposition of CsdA_DD to Hera_DD and CshA_DD show different packing of the three α helices (Figure 3C). The obvious differences in angle and length between CsdA_DD and Hera_DD helices result in a large root-mean-square deviation (RMSD) value of 3.0 Å between the Cα atoms of the two domains, while that of CsdA_DD and CshA_DD is 3.5 Å using a Dali pairwise comparison. The arrangement of the helices of CsdA_DD is different from those of Hera_DD or CshA_DD (Figure 3C). The orientation of the α1 helix in chain A of CsdA_DD rotates outward by approximately 15° and 30° when compared with that of Hera_DD and CshA_DD, respectively. The α2 and α3 helices in three DDs overlapped poorly. In addition, the loop between the α2 helix and α3 helices in chain B of CsdA_DD is longer than the corresponding loops in Hera_DD and CshA_DD (Figures 3C and S2B). Due to the different arrangements in the DDs, the angle of the V-shape dimer of CsdA is larger than that of Hera or CshA (Figure 3D).

Solution Structure of the RBD of CsdA

The sequence of the C-terminal regions of CsdA does not show any similarities to that of Hera, YxiN, Vasa, and eIF4AIll (Figure 1A). Sequence similarity searching indicates that a putative RBD (or CsdA_RBD) is located at the C-terminal region of CsdA, which is connected to the DD through a long linker with 39 amino acids (Figure 1C). We have determined the solution structure of CsdA_RBD by multi-dimensional heteronuclear NMR spectroscopy. All NMR spectral data were processed with NMRPipe and NMRDraw software (Delaglio et al., 1995). Two-dimensional ¹H/¹⁵N HSQC spectra reveal good dispersion

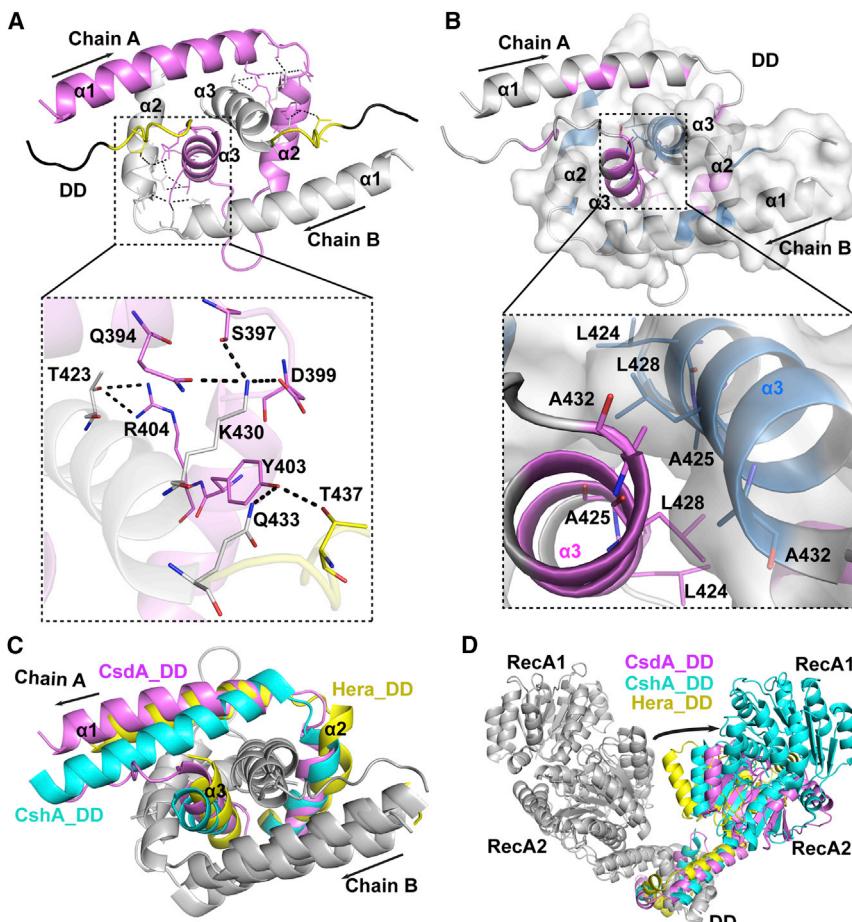


Figure 3. The Structural Detail of the DD of CsdA

(A) Residues involved in hydrogen bond interactions of the CsdA_DD. The two DD are shown as cartoons, and the two chains of two DD are indicated in violet and gray, respectively. Residues involved in hydrogen bond interactions are shown in sticks and highlighted in the dashed box below, with the hydrogen bonds in black dashed lines. An arrow indicates the direction of the chain.

(B) The extensively packed hydrophobic interactions in CsdA_DD. The two chains of DD are shown in the cartoon, while one of the chains is shown in the transparency surface. Residues of the two chains of DD involved in hydrophobic interactions are indicated in violet and blue. Residues L424, A425, L428, and A432 of the two α_3 helices forming the main hydrophobic interface of the dimer are highlighted in the dashed box below.

(C) Superposition of the DD of CsdA, Hera, and CshA. The CsdA_DD is colored in violet cartoon, the Hera_DD is colored in yellow cartoon, and the CshA_DD is colored in cyan cartoon. The corresponding helices in the DD are labeled as α_1 , α_2 and α_3 .

(D) Structural superposition of CsdA_RecA2-DD, Hera_RecA2-DD, and CshA_RecA2-DD. The differences in the angle of the three V-shape dimers are shown as arrows. See also Figure S2.

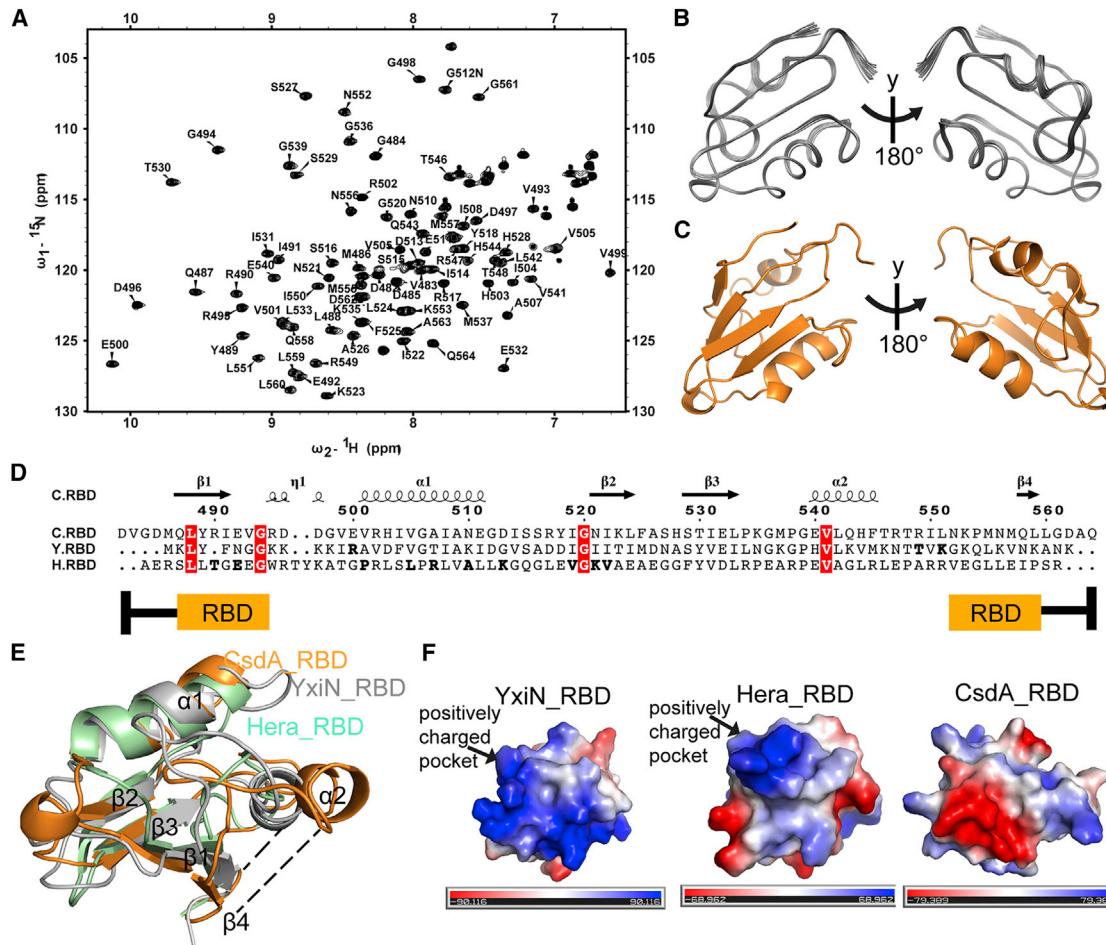
of the amide groups (Figure 4A). The assembly of the 20 lowest-energy structures is highly converged (Figures 4B and 4C). The RMSD of the 20 structures is 0.49 Å for the backbone and 0.92 Å for the heavy atoms in the well-defined region. The complete statistics for the 20 structures are listed in Table 2.

The structure of CsdA_RBD reveals a typical RRM-like fold, which is well characterized by four-stranded β sheets packed against two α helices (Clery et al., 2008). Commonly, RRMs have a conserved RNP1 sequence, [RK]-G-[FY]-[GA]-[FY]-[ILV]-X-[FY], and RNP2 sequence, [ILV]-[FY]-[ILV]-X-N-L. However, the corresponding RNP1 and RNP2 in CsdA_RBD are H-S-T-I-E-L-P-K-G and L-Y-R-I-E, respectively (Figure 4D), which are quite different from other common RRMs. A DALI search for structural homologs of CsdA_RBD reveals that CsdA_RBD shares low homogeneity with other RRMs, and the closest structural homolog with CsdA_RBD is the RBD of *Bacillus subtilis* RNA helicase YxiN (YxiN_RBD) (Z score 6.3, RMSD 2.3 Å, sequence similarity 32.9%). The structural differences are more significant between CsdA_RBD and the RBD of *T. thermophilus* Hera (Hera_RBD) (Z score 3.5, RMSD = 2.9 Å, sequence similarity 34.3%). Structural superpositions of CsdA_RBD, Hera_RBD (PDB: 4I67) (Steimer et al., 2013) and YxiN_RBD (PDB: 3MOJ) (Hardin et al., 2010) show differences between three RBDs are mainly in the arrangements of α_2 and β_4 ; the disposition of the β_4 to α_2 of CsdA_RBD is more open than that of YxiN_RBD and Hera_RBD (Figure 4E). Therefore,

the whole structure of CsdA_RBD is not as compact as those of YxiN_RBD and Hera_RBD (Figure 4E). Besides, the surface electrostatic potential of YxiN_RBD and Hera_RBD show positively charged pockets where RNA may bind (Hardin et al., 2010; Steimer et al., 2013), whereas no such pocket is seen in the corresponding region in CsdA_RBD (Figure 4F). Residues R417, T466, and K468 of YxiN_RBD have been shown to participate in specific protein-RNA interactions with the H92 stem-loop of 23s rRNA (Hardin et al., 2010), while the corresponding residues in CsdA_RBD are E500, R549, and L551 (Figure 4D). Similarly, residues T437, E440, P450, L454, R456, A459, K473, and V474 of Hera_RBD are involved in RNA binding, while the corresponding residues of CsdA_RBD are R490, V501, V505, A507, N510, D513, I519, N521, and I522 (Figure 4D). The differences suggest that the RNA-binding mode of CsdA_RBD may be different from those of YxiN_RBD and Hera_RBD since RRMs usually utilize various modes to bind RNA (Clery et al., 2008; Maris et al., 2005; Steimer et al., 2013).

Conformational Flexibilities of CsdA_1–445 and CsdA_FL Revealed by SAXS

We have obtained either crystal structures or solution structures of all the individual domains of CsdA. The narrow proton chemical shift range of NHs in HSQC spectra demonstrate that the C-tail 65 residues (following CsdA_RBD) are intrinsically disordered (Figure S4). To better understand the overall conformations of the helicase core domains of CsdA (CsdA_1–445) and the full-length CsdA (CsdA_FL) in solution, we performed SAXS experiments. The SAXS profiles with the corresponding

**Figure 4. Solution Structure of CsdA_RBD**

- (A) Two-dimensional $^1\text{H}/^{15}\text{N}$ HSQC spectrum of CsdA_RBD. The backbone assignments are labeled in the spectrum. The unlabeled resonances are the side-chain resonances of Asp and Glu.
- (B) The ensemble of the 20 lowest-energy structures of CsdA_RBD is shown in the ribbon.
- (C) Cartoon representation of the 20 lowest-energy structures of CsdA_RBD (orange).
- (D) Structural-based sequence alignment of CsdA_RBD (C.RBD), YxiN_RBD (Y.RBD), and Hera_RBD (H.RBD). The conserved residues are shown in red boxes. The residues involved in RNA binding in YxiN_RBD and Hera_RBD are highlighted in bold.
- (E) Structure comparison of CsdA_RBD (orange) with YxiN_RBD (gray) and Hera_RBD (light green). The distance differences between $\beta 4$ and $\alpha 2$ are shown in broken lines.
- (F) The electrostatic surface of YxiN_RBD, Hera_RBD, and CsdA_RBD. The negative surfaces are colored in red, while the positive surfaces are in blue. The positive RNA-binding pockets in YxiN_RBD and Hera_RBD are highlighted.

pair-distance distribution functions are shown in Figures 5A and 5B. Real space R_g values of CsdA_218–445, CsdA_1–445, and CsdA_FL in solution evaluated by GNOM (Semenyuk and Svergun, 1991) are 31.2 ± 0.1 Å, 46.6 ± 0.9 Å, and 53.6 ± 1.2 Å, while the D_{\max} values are 96.0, 158.0, and 182.0 Å, respectively. Molecular weights of CsdA_218–445, CsdA_1–445, and CsdA_FL assessed by SAXSMoW (Fischer et al., 2010) were 58.6, 97.6, and 135.9 kDa, respectively. Shape models of CsdA_218–445, CsdA_1–445, and CsdA_FL were generated by DAMMIN (Figure 5D), with χ values against experimental SAXS data of 0.88, 0.52, and 0.25, respectively. The crystal structure of CsdA_218–445 has a χ value of 1.1 against the experimental SAXS curve (Figure S5), and can be superimposed into its shape model quite well (Figures 5D and S5), with a normalized spatial discrepancy of 1.16. The relative locations

of RecA1 domains and C-terminal regions to the RecA2-DD domains are indicated in the corresponding low-resolution shape models (Figure 5D). The results suggest a V-shaped and extended conformation of CsdA in solution.

Due to the long linkers (nine amino acid residues) between RecA1 and RecA2 in the helicase core regions and C-terminal regions, both CsdA_1–445 and CsdA_FL show flexibilities in solution as indicated by a dimensionless Kratky plot (Figure 5C). To better characterize their flexibilities, Ensemble Optimization Method 2.0 (Tria et al., 2015) analyses were carried out for CsdA_1–445 and CsdA_FL, each with 200 cycles of genetic algorithms, respectively. One final five-conformer ensemble for CsdA_1–445 and one final six-conformer ensemble for CsdA_FL were obtained with final χ values of 0.27 and 0.31, respectively (Figures S6 and S7). Flexibilities of the linkers in the helicase

Table 2. Summary of Top20 CsdA_RBD Structures from CYANA and Xplor-NIH Calculations

	CYANA	Xplor-NIH in Vacuum	Xplor-NIH in Water
Experimental NMR restraints			
Distance constraints			
Total	1,290	1,290	1,290
Short range, $ i - j \leq 1$	693	693	693
Medium range, $1 < i - j < 5$	198	198	198
Long range, $ i - j \geq 5$	399	399	399
Dihedral-angle constraints	96	96	96
Structure statistics			
Average CYANA target function (\AA^2)	1.46		
Average total energy (kJ/mol)		-279.10	-1,787.08
Violations (mean and SD)			
Distance constraints (>0.5 \AA)	0	0	0
Dihedral-angle constraints (>5°)	0	0	0
van der Waals (>0.5 \AA)	0	0	0
Deviations from idealized geometry ^a			
Bond lengths (\AA)	0.001	0.004	0.008
Bond angles (°)	0.2	0.464	0.700
Ramachandran plot (%) ^b			
Residues in most favored regions	74.6	75.6	80.3
Residues in allowed regions	24.1	22.5	18.3
Residues in generous allowed regions	1.3	1.8	1.4
Residues in disallowed regions	0.0	0.1	0.0
Average pairwise RMSD (\AA) ^c			
Backbone	0.22	0.27	0.49
Heavy atom	0.69	0.77	0.92

^aDeviations from idealized geometry were generated by PSVS (Bhattacharya et al., 2007).

^bRamachandran plot were generated by PROCHECK (Laskowski et al., 1996).

^cRMSD values are calculated for ordered regions (residues 482–562).

core regions and tail regions can be indicated from the ensemble models, as the RecA1 domains and tail regions of CsdA can have various orientations (Figures 5E and S6). Details of the ensemble models of CsdA_1–445 and CsdA_FL can be found in *Supplemental Information* (Figures S6 and S7), and ensemble structural models of CsdA_FL are shown in Figure 5E. These results suggest that both the helicase core domains and the C-terminal regions of CsdA are flexible in solution.

DD Is Indispensable for the Stability of Dimeric CsdA

In Vitro and In Vivo

Both CsdA_564 (residues 1–564; Figure 1C) and CsdA_DDRBD (residues 372–564; Figure 1C), encompassing DD (residues 372–433), are apparent dimers in solution, which is judged by gel-permeation chromatography. To investigate the dimeric nature of CsdA and to find out whether the DD could help CsdA to

be a stable dimer in solution, we performed gel-permeation chromatography by mixing CsdA_564 and CsdA_DDRBD (Figure 1C). The heterodimer formed by CsdA_564 and CsdA_DDRBD with intermediate molecular weight would appear if monomer-dimer exchange occurs. We found that only two main peaks indicating each individual homodimer were observed in the chromatogram when CsdA_564 and CsdA_DDRBD were mixed and incubated at 24°C for 1 hr. If the incubation time was prolonged to 16 hr, very few heterodimers formed, which suggests no or very slow monomer-dimer exchange (Figure 6A). However, a peak with intermediate molecular weight appeared when incubating the mixture at 37°C for 1 hr, indicating fast monomer-dimer exchange. The controls, with incubation of each individual protein at 37°C for 1 hr and three peaks on the SDS-PAGE gel, confirms that the appearing third peak is the heterodimer formed by the DD exchange of CsdA_564 (RecA1-RecA2-DD-RBD) and CsdA_DDRBD (Figures 6A and S8A). All these experiments suggest that CsdA is a stable dimer at low temperature (24°C), whereas it can exchange rapidly at high temperature (37°C).

To further investigate how the DD affects the activity of CsdA, we used a classical NADH-coupled method to perform *in vitro* ATPase (Bessman, 1963) and unwinding activity assays (Bizebard et al., 2004). CsdA was shown to unwind an RNA duplex with 3' or 5' single-stranded extensions (Bizebard et al., 2004). We observed that CsdA_1–445 possesses the conventional ATPase (Figure 6B) and unwinding activities in RNA and ATP-dependent manners (Figure 6C, lane 2) when compared with the controls. Conversely, both the ATPase and unwinding activities of CsdA_372 (residues 1–372; Figure 1C, with the truncation of the DD) disappeared (Figures 6B and 6C, lane 3). It seems that DD is indispensable for the ATPase and unwinding activities of CsdA. We also conducted *in vivo* assays to further elucidate whether the deletion of the *csdADD* gene affects the activity of CsdA in *E. coli*. The isogenic *csdA*- and *csdADD*-deficient Top10 strains (Δ *csdA*, Δ *csdADD*) were obtained as described previously (Datsenko and Wanner, 2000; Resch et al., 2010) with minor changes. In previous studies (Resch et al., 2010; Sledjeski et al., 1996) and our research, the synthesis of RpoS was shown not to have significant differences in *E. coli* wild-type (WT) strain and Δ *csdA* strain at 37°C (Figure 6D, lanes 1 and 2), whereas RpoS synthesis was scarcely detected in Δ *csdA* strain at 24°C (Figure 6D, lane 3) compared with the WT strain (Figure 6D, lane 5). RpoS synthesis was rescued in the Δ *csdA* strain that was transfected with pBADCsdA plasmids (Figure 6D, lane 4). However, it is still not clear whether the DD is dispensable for the translation activation of *rpoS* mRNA at high temperature or cold-shock situation. To answer this question, we produced an isogenic *csdADD* deletion strain and incubated it at 37°C and 24°C until the A_{600} reached 0.5. By western blotting, the levels of the synthesis of RpoS were similar in the Δ *csdADD* strain and WT at 37°C (Figure 6D, lane 8). However, RpoS synthesis was rarely detected in the Δ *csdADD* strain at 24°C (Figure 6D, lane 6), while the synthesis of RpoS at 24°C was rescued in the Δ *csdADD* strain transfected with pBADCsdA plasmids (Figure 6D, lane 7). These findings are very different from previous studies that showed that many other RNA helicase monomers can express well *in vitro* and possess RNA unwinding activities as monomers (Samatanga and Klostermeier, 2014; Sengoku et al., 2006; Yang and Jankowsky, 2006). We further

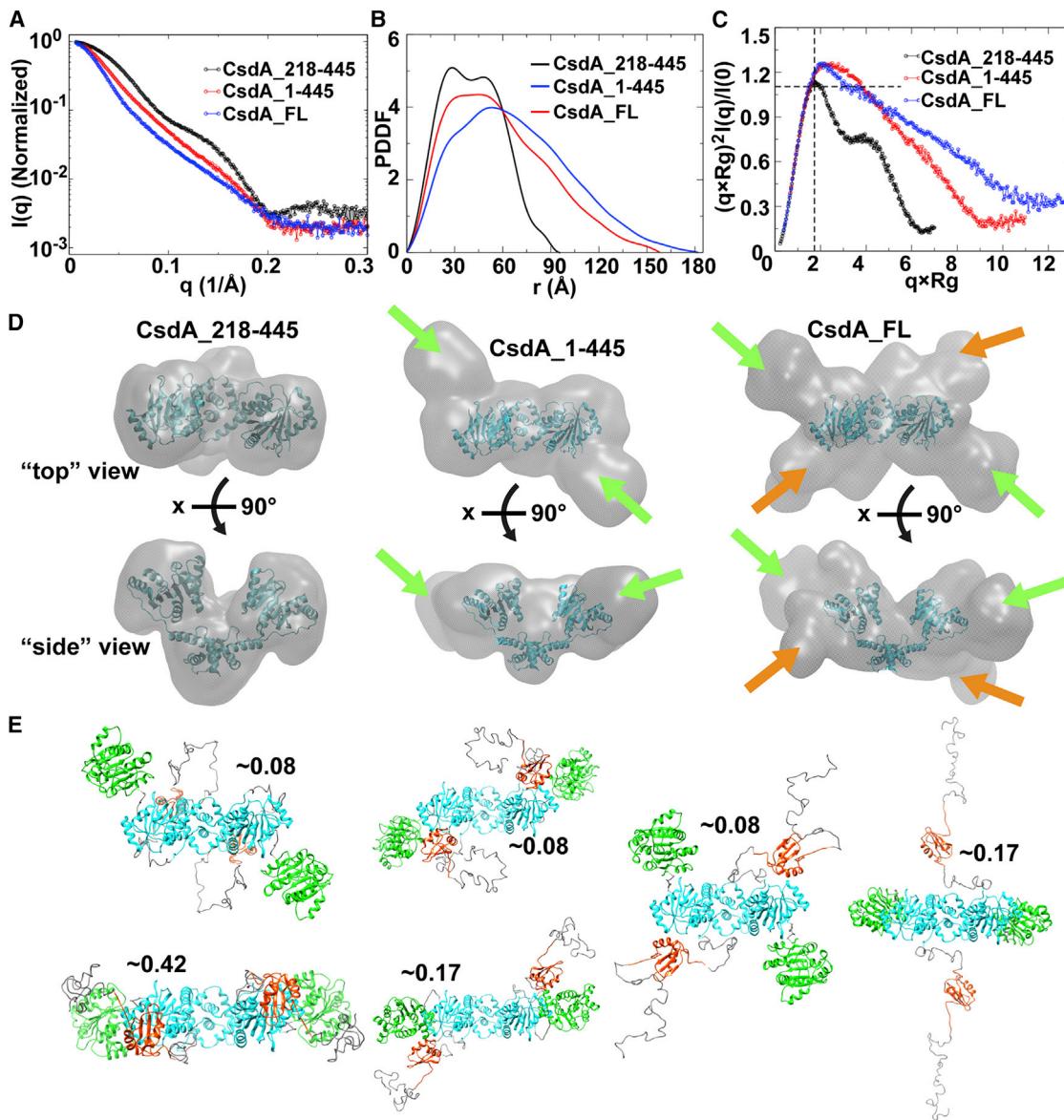


Figure 5. SAXS Profiles and Low-Resolution Shape Models of CsdA_218-445, CsdA_1-445, and CsdA_FL

(A) Normalized SAXS profiles of CsdA_218-445 (black), CsdA_1-445 (red), and CsdA_FL (blue). For all the three proteins, only data points up to $q = 0.3 \text{ \AA}^{-1}$ are plotted.

(B) Pair-distance distribution function (PDDF) curves of CsdA_218-445 (black), CsdA_1-445 (red), and CsdA_FL (blue). All the three curves have relatively flattened peaks, suggesting that the corresponding proteins are of multi-domains.

(C) R_g -based dimensionless Kratky plots of CsdA_218-445 (black), CsdA_1-445 (red), and CsdA_FL (blue). The peak of rigid, globular protein is indicated by dotted lines. Both CsdA_1-445 and CsdA_FL show flexibilities according to the dimensionless Kratky plots.

(D) Shape models of CsdA_218-445, CsdA_1-445, and CsdA_FL. The crystal structure of CsdA_218-445 can be superimposed to its shape model quite well, with a normalized spatial discrepancy of 1.16. Putative RecA1 domains and the C-terminal regions are indicated with green arrows and orange arrows, respectively.

(E) Ensemble structural models of CsdA_FL with individual fraction of occupancy. The RecA1 domains are shown in green, RecA2-DD in cyan, and RBD in orange. Disordered regions are in gray. See also Figures S4-S7.

asked how DD deletion affects the activity of CsdA *in vitro* and *in vivo*. We have noticed that CsdA_372 expressed poorly and precipitated easily when DD was deleted. In addition, dynamic light scattering showed that DD deletion significantly decreased the stability of CsdA_372, and the protein easily aggregated under any conditions (data not shown). DD has no enzymatic activity or RNA-binding affinity (data not shown). We also found

that the mutations on the DD of CsdA_1-445, containing A425R, A426R, A427R, L428R, and L429R, had an effect on the dimerization of CsdA_1-445; the CsdA_1-445 R5 mutant behaves as a monomer (Figures S8B and S8C). The monomer is easy to precipitate and has undetectable ATPase or unwinding activity (Figure S8D), suggesting that the DD plays a critical role in overall structural stability.

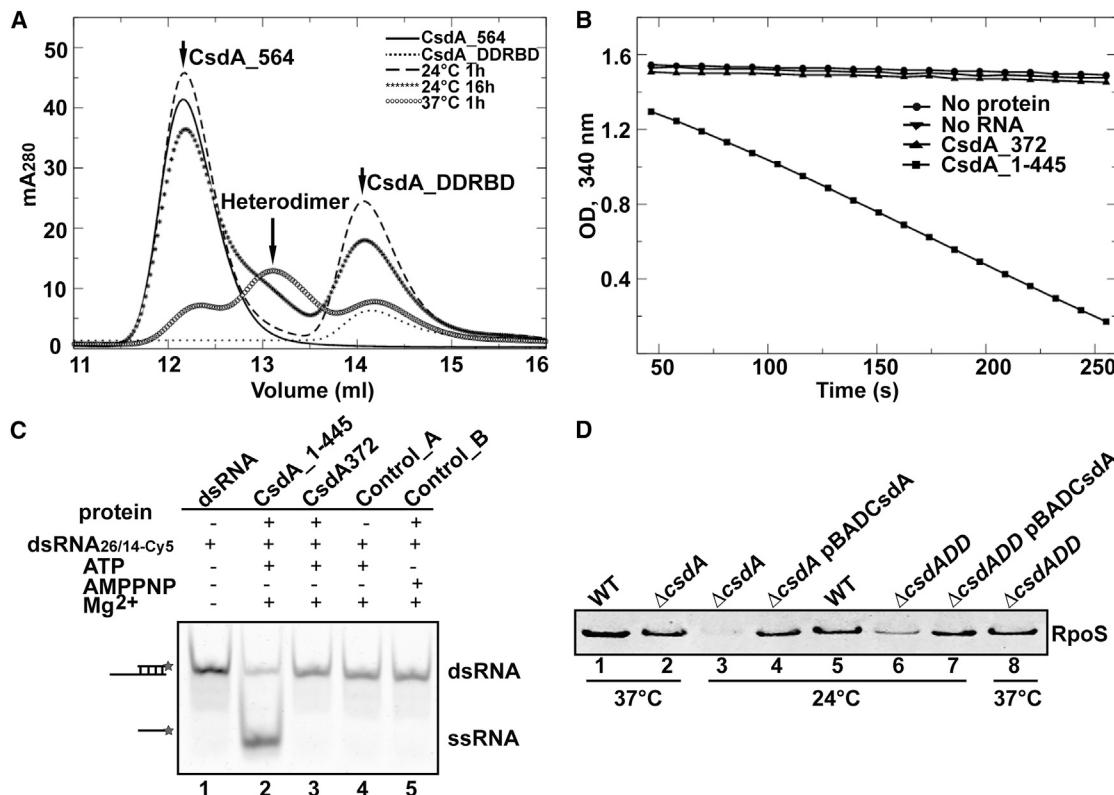


Figure 6. ATPase and Unwinding Activities Assays Show that DD Is Essential to the Stability of CsdA

(A) Size-exclusion chromatography of CsdA_564 (solid), CsdA_DDRBD (dot), and a mixture of them, after incubation at 24°C for 1 hr (dashed) and 16 hr (star) as well as 37°C (circle) for 1 hr. Heterodimer occurred in the middle of the CsdA_564 and CsdA_DDRBD (arrow), indicating the dimeric nature of CsdA in solution at low and high temperature.

(B) The ATPase activity of CsdA_1-445 and CsdA_372. The RNA used in the ATPase activity assay is polyU₁₀ RNA. No protein and no RNA stand for protein and RNA was added into the reaction system, respectively.

(C) The unwinding activity of CsdA_1-445 and CsdA_372. Control_A shows that protein was not added into the reaction system. Control_B shows that AMP-PNP instead of ATP was added into the reaction system. The 5'-Cy5-labeling dsRNA_{26/14-Cy5} is shown in the black star. The dsRNA_{26/14-Cy5} bands are indicated as dsRNA and the unwound single-stranded RNAs are indicated as ssRNA.

(D) The expression of RpoS in the WT, *ΔcscdA* and *ΔcscdADD* strains at 37°C and 24°C. The expression levels of RpoS were determined by quantitative western blotting. See also Figure S8.

The ATPase and Unwinding Activities of CsdA_564 and CsdA_1–445

The helicase core of DEAD-box RNA helicases catalyzes RNA duplex separation, while the functions of the auxiliary regions differ (Linden et al., 2008; Yan et al., 2003; Pugh et al., 1999). As shown in our experiments, CsdA_1-445 displayed normal RNA-dependent ATPase activity (Figure 6B) and ATP-dependent unwinding activity (Figure 6C). To evaluate the functions of the C-terminal regions of CsdA, we quantitatively monitored the unwinding activities of CsdA_1-445, CsdA_564, and CsdA_DDRBD, and the ATPase activities of CsdA_1-445 and CsdA_564. In unwinding assays, we found that CsdA_1-445 displayed unwinding activity with successive increases in the intensity of the unwound single-stranded RNAs in 50 min (Figures 7A and 7C). Interestingly, upon the addition of CsdA_564, ~80% RNA duplex was unwound into single-stranded RNA in only 5 min (Figures 7B and 7C). However, when CsdA_DDRBD was added into the reaction, no significant differences in the intensities of the unwound single-stranded RNAs were detected after 50 min (Figure 7C). These results demonstrate that the RBD and

the long linker between DD and RBD could enhance the unwinding activity of CsdA_564. In ATPase activity assays, the ATPase activity of CsdA_564 in the presence of 73mer RNA (a 73 nt RNA fragment of *E. coli* 23s rRNA, nucleotides 2,508–2,580, containing the stem of hairpin 92; [Figure S3A](#)) ([Kossen et al., 2002](#)) is stronger than that of CsdA_1–445 ([Figure 7D](#)). These results provide important insights into the contributions of the C-terminal regions in the RNA unwinding activity and ATPase activity of CsdA.

RNA-Binding Properties of the C-Terminal Regions of CsdA and CsdA RBD

Compared with the catalytic role of the helicase core of RNA helicases, additional regions may mediate binding of RNAs, such as Hera_RBD (Klostermeier and Rudolph, 2009) and YxIN_RBD (Kossen et al., 2002; Hardin et al., 2010). Our study shows that the ATPase and unwinding activities of CsdA_564 are higher than those of CsdA_1–445. To better understand the function of the flanking regions following DD, FP experiments were performed. CsdA FL (residues 1–629; Figure 1C), CsdA 564,

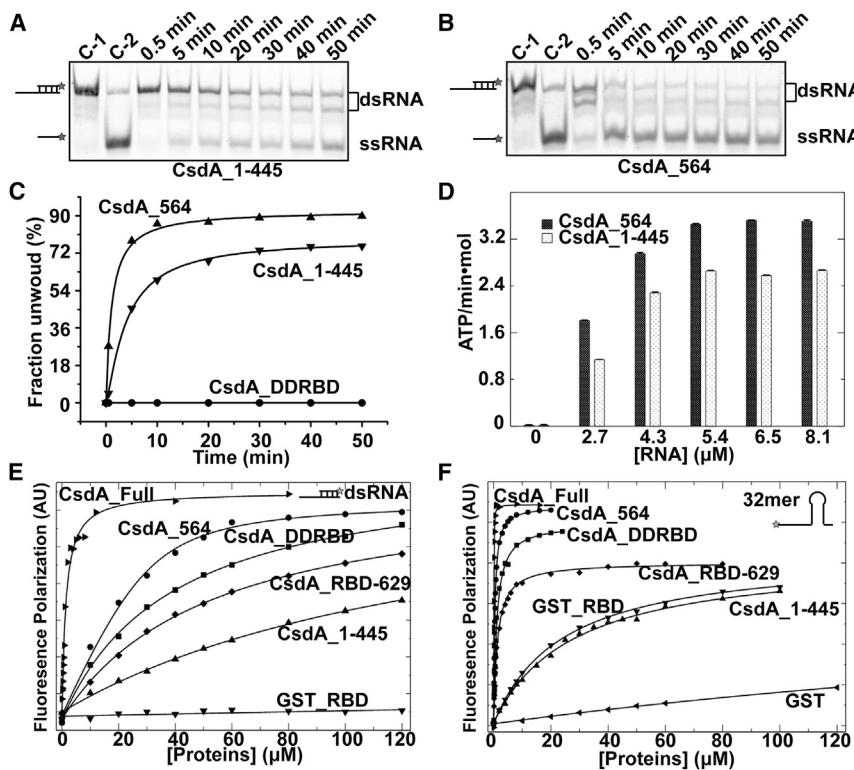


Figure 7. Unwinding and ATPase Activities and Binding Ability of CsdA Constructs

(A) Unwinding activity of CsdA_1-445 in the presence of ATP over time. Controls are indicated as C-1 and C-2. C-1 shows that only dsRNA_{26/14-Cy5} was added in the reaction system. C-2 shows that dsRNA_{26/14-Cy5} were denatured at 90°C for 5 min. The 5'-Cy5-labeling dsRNA_{26/14-Cy5} is labeled in the black star. The dsRNA_{26/14-Cy5} bands are indicated as dsRNA, unwound single-stranded RNA is indicated as ssRNA.

(B) Unwinding activity of CsdA_564 in the presence of ATP over time.

(C) Quantitation of unwinding activity of CsdA_564 and CsdA_1-445 in (A and B), and CsdA_DDRBD (data not shown). A fraction of 100 stands for 100% unwound ssRNA.

(D) ATPase activity of CsdA_1-445 and CsdA_564 in the existence of 73mer RNA.

(E) The RNA-binding affinities of different CsdA constructs to dsRNA_{26/14-FAM} determined by FP experiments. The 5'-FAM-labeling dsRNA_{26/14-FAM} is shown in the black star. The K_D values are shown in Table S1.

(F) The RNA-binding affinities of different CsdA constructs to 32mer RNA determined by FP experiments. The 5'-FAM-labeling 32mer RNA is shown in the black star. The K_D values are shown in Table S1. K_D values and the corresponding standard errors were determined as described in the STAR Methods. See also Figure S3 and Table S1.

CsdA_1-445, CsdA_DDRBD, GST_RBD, CsdA_RBD-629 (Figure 1C), and 5'-FAM-labeled dsRNA_{26/14-FAM}, 32mer RNA (a 32 nt RNA containing the regions comprising of the stem of hairpin 92 of 23S rRNA; Figure S3B) were selected for RNA-binding affinity assays. The equilibrium dissociation constants, K_D, were analyzed by fitting the FP data of 5'-FAM-labeled RNAs binding to different CsdA segments. The FP data were fitted similar to Wang et al. (2011), as described in the STAR Methods. CsdA_FL displays the highest RNA-binding affinity in the presence of both dsRNA_{26/14-FAM} and 32mer RNA (Figures 7E and 7F; Table S1). In dsRNA_{26/14-FAM}-binding assays, RNA-binding affinities of the different CsdA constructs containing the C-terminal regions are all higher than those of CsdA_1-445. The RNA-binding affinities of CsdA_FL (1.30 μM), CsdA_564 (21.60 μM), CsdA_DDRBD (39.0 μM), and CsdA_RBD-629 (52.80 μM) are ~16-, ~7-, ~4-, and ~3-fold higher than that of CsdA_1-445 (~147 μM) (Figure 7E; Table S1), suggesting that the C-terminal regions of CsdA might enhance RNA binding.

In the 32mer RNA-binding assays, the differences in 32mer RNA-binding affinities among different CsdA constructs are similar to those in dsRNA_{26/14-FAM}-binding affinity assays (Figure 7F; Table S1). The RNA-binding affinities of CsdA_FL (0.07 μM), CsdA_564 (0.39 μM), CsdA_DDRBD (0.99 μM), and CsdA_RBD-629 (1.37 μM) are ~385-, ~71-, ~28-, and ~20-fold than that of CsdA_1-445 (27.70 μM) (Figure 7F; Table 2). It is likely that the C-terminal regions are essential for CsdA to bind RNA strongly.

From the above results, we also found that RNA-binding affinities of 32mer RNA to different CsdA constructs are higher than those of dsRNA_{26/14-FAM} (Figures 7E and 7F; Table S1). Especially, CsdA_RBD exhibits a K_D of 24 μM for 32mer RNA, while

the RNA-binding affinity of CsdA_RBD to dsRNA_{26/14-FAM} is hardly detected, indicating that CsdA_RBD prefers binding the 32mer RNA. To further investigate how CsdA_RBD binds RNA, we performed further NMR analysis of CsdA_RBD in the presence of 32mer RNA. Chemical shifts changing upon titrating RNAs with different concentrations were recorded using ¹H,¹⁵N-HSQC NMR spectra. Several obvious changes in chemical shifts and peak intensity were observed for CsdA_RBD backbone resonances assigned to residues H503, I504, N510, I514, S516, N521, I522, R549, and L551 (Figure S9A). These titration assays allowed us to identify the residues of CsdA_RBD involved in RNA binding (Figures S9B and S9C). Moreover, these residues are not conserved in Hera_RBD or YxiN_RBD (Figure 4D). The WT CsdA_RBD could bind 32mer RNA as revealed by electrophoretic mobility shift assays (Figure S9D). However, the mutation of the CsdA_RBD with both R549A and L551A had a significant effect on the RNA-binding affinity; the CsdA_RBD mutant no longer binds to 32mer RNA (Figure S9E), suggesting that these residues are crucial for CsdA_RBD to bind RNA.

To further investigate whether the binding of the 32mer RNA to CsdA_RBD involves the single-stranded region (referred to as 15ssRNA, followed by the stem region, referred to as 17stem in the 32mer RNA), we performed FP assays on 15ssRNA and 17stem RNA using different C-terminal constructs containing RBD (CsdA_DDRBD, CsdA_RBD, and CsdA_RBD-629). All these constructs show similar RNA-binding affinities with those in 32mer RNA-binding assays (Table S1). Notably, the RNA-binding affinities of CsdA_DDRBD (1.65 μM) and CsdA_RBD-629 (8.30 μM) are ~10- and ~8-fold higher than those in the 17stem RNA-binding assays, respectively (Table S1). In addition,

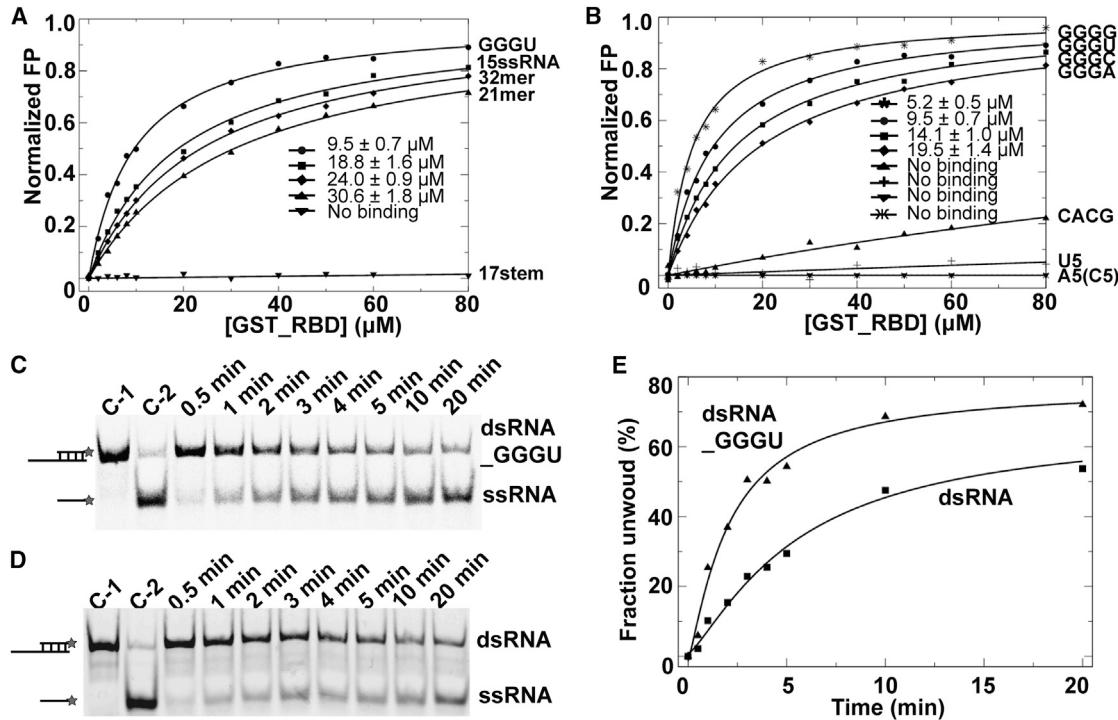


Figure 8. The RNA-Binding Affinities of CsdA_RBD to Different RNAs

(A and B) The RNA-binding affinities of CsdA_RBD to different RNA fragments determined by FP experiments. K_D values and the corresponding standard errors were determined as described in the STAR Methods.

(C and D) Unwinding activity of CsdA_564. C-1 shows that only dsRNA was added in the reaction system. C-2 shows that dsRNA was denatured at 90°C for 5 min. The 5'-Cy5-labeling dsRNA is labeled in the black star. The dsRNA_{26/14-Cy5}-GGGU bands are indicated as dsRNA_GGGU, unwound single-stranded RNA is indicated as ssRNA. The dsRNA_{26/14-Cy5} bands are indicated as dsRNA.

(E) Quantitation of unwinding activity of CsdA_564 and CsdA_1–445 in (C and D). A fraction of 100 stands for 100% unwound ssRNA. See also Figure S3B and Table S2.

CsdA_RBD has a K_D of 18.80 μM to 15ssRNA (Table S1), while no RNA-binding affinities to 17stem RNA can be detected, suggesting that CsdA_RBD prefers binding the single-stranded region, 15ssRNA.

Interestingly, we found that RNA-binding affinities of CsdA_RBD (30.60 μM) to the 21mer RNA (lacking the 11mer RNA in the 5' end of 32mer RNA) is similar to 32mer RNA (24.00 μM) and 15ssRNA (18.80 μM) (Figure 8A; Table S1). It is likely that CsdA_RBD might recognize the 4mer GGGU sequence joined by the 17stem since CsdA_RBD has no RNA-binding affinity to the 17stem RNA. To further investigate whether CsdA_RBD prefers binding to a single-stranded GGGU sequence, we conducted FP assays utilizing different RNA mutants (Figures 8A and 8B). We found that the RNA-binding affinities of CsdA_RBD to single-stranded CACG RNA, A₅ RNA, C₅ RNA, and U₅ RNA are hardly detected when compared with GGGG (5.20 μM) and GGGU (9.50 μM) RNAs. We also determined the RNA-binding affinities of CsdA_RBD to GGGC RNA (14.10 μM) and GGGA (19.50 μM) RNA. Further RNA unwinding assays show that the unwinding ability of CsdA_564 for dsRNA_{26/14-Cy5}-GGGU (with a GGGU segment in the 5' overhang) is higher than for dsRNA_{26/14-Cy5} (Figures 8C, 8D, and 8E). Together with the above results, CsdA_RBD might prefer binding to single-stranded G-rich RNAs.

We also determined the ATPase activities of CsdA_1–445 and CsdA_564 in the existence of RNAs used in the RNA-binding affinity assays. For short RNAs, the ATPase activities of CsdA_564 are similar to CsdA_1–445, while for 32mer RNA, 15ssRNA, or dsRNA_{26/14}, the ATPase activities of CsdA_564 are stronger than that of CsdA_1–445 (Table S2). Moreover, the ATPase activity of CsdA_564 activated by 32mer RNA is apparently higher than those activated by other RNAs (Table S2), suggesting that the higher RNA-binding affinity for 32mer RNA could improve the enzymatic activity of CsdA_564.

DISCUSSION

CsdA is a cold-shock DEAD-box protein that plays an important role in *E. coli* in a cold-shock situation (Resch et al., 2010; Charnois et al., 2004). We not only determined the structures of RecA1 and RecA2, but also found two previously uncharacterized domains, a DD and an RBD in the C-terminal regions of CsdA. These findings further demonstrate the diversity of C-terminal regions in DEAD-box RNA helicases. A majority of RNA helicases function as monomers; e.g., the *B. subtilis* YxiN, the yeast Mss116p, *Drosophila* MLE, and hepatitis C virus NS3 (Wang et al., 2008; Del Campo and Lambowitz, 2009; Prabu et al., 2015; Dumont et al., 2006). Our study demonstrated that CsdA is a stable dimer in solution at low temperature (Figure 6A).

and that DD is responsible for the normal function (Figures 6B–6D, S8B, and S8C) and structural stability of the CsdA dimer, suggesting that a stable dimeric conformation may be critical to its function at low temperature.

Our studies suggest that the C-terminal regions strongly enhance the enzymatic activities and RNA-binding affinities of CsdA, which is similar to other DEAD-box RNA helicases, including Hera and CshA, which have similar dimeric conformations (Huen et al., 2017; Klostermeier and Rudolph, 2009). In Hera or CshA, the RNA-binding affinities can increase about 50–70 times in the presence of their C-terminal regions. However, in CsdA, the linker between DD and RBD and the C-tail are longer than those in Hera and CshA. Due to these long, basic amino acid-rich linkers or glycine-rich tails in CsdA, the RNA-binding affinities can be about ~100–400 times higher for the full-length CsdA than for the helicase core domains. It is consistent with the previous study that some RNA helicases containing a C-terminal glycine-rich region are capable of binding single-stranded nucleic acids (Nakajima et al., 1997). Unlike DbpA and its *B. subtilis* homolog YxiN, which can specifically bind to hairpin 92 of 23S rRNA via their RBDs (Diges and Uhlenbeck, 2001; Wang et al., 2006; Hardin et al., 2010), our results show that CsdA can bind to different RNAs with μ M or sub- μ M affinities. These findings suggest that the C-terminal regions play critical roles in RNA binding of CsdA. The C-terminal regions of CsdA are more complex than other DEAD-box RNA helicases in *E. coli*. Our SAXS models of CsdA give us a low-resolution picture of this dimeric multi-domain protein, suggesting that the C-terminal regions are dynamic and flexible in solution, which could be related to the diverse functions of this protein at low temperature.

Our studies also show that CsdA_RBD prefers binding single-stranded G-rich RNA, which is not reported before. Although CsdA_RBD has a typical RRM-like motif and the RNA recognition sequences of CsdA_RBD and Hera_RBD are similar (GGGPy for Hera_RBD) (Steimer et al., 2013), structural studies have shown that CsdA_RBD and Hera_RBD or YxiN_RBD differ in their surface electrostatic potentials (Figure 4F). In addition, the RNA recognition modes of the CsdA_RBD and the YxiN_RBD are different: CsdA_RBD prefers binding single-stranded G-rich RNA, while YxiN_RBD specifically recognizes the loop region of hairpin 92 (Kossen et al., 2002). Since some mutations on CsdA_RBD destroy the stability of the protein, we failed to verify all the shifted amino acids in chemical shift perturbation assays in detail, but our results have shown that the likely RNA-binding site in CsdA_RBD is different from Hera_RBD or YxiN_RBD. Thus, although the complex structure of CsdA_RBD and RNA are unknown, our findings suggest that CsdA_RBD might further expand the versatility of RBDs as RNA-binding domains (Steimer et al., 2013).

As mentioned above, the sequence preference of CsdA_RBD is similar to Hera_RBD, which prefers binding a single-stranded GGGPy sequence and, upon Hera_RBD binding to RNA, the helicase core of Hera would then unwind the adjacent duplex (Steimer et al., 2013). In CsdA, due to the RNA-binding preference of CsdA_RBD, CsdA can bind to 32mer RNA with an affinity of ~20–30 times higher than other non-preferred RNAs and can unwind dsRNA with a 5' overhang GGGU sequence more efficiently. Based on these results, it is likely that CsdA might

generally bind to the exposed single-stranded RNA with a G-rich sequence through its RBD, while the long, basic amino acid-rich linkers or tails in its C-terminal regions strongly strengthen the binding, and then the helicase core may unwind the neighboring or adjacent RNA duplex. It is already known that CsdA is involved in the assembly of the 50S large ribosomal subunit (Charollais et al., 2004). The deletion of *csdA* leads to a severe deficit of free 50S subunits and accumulation of 40S particles at low temperatures (Charollais et al., 2004). 23S rRNA is an important component of both pre50S and 50S subunit of ribosome (Shajani et al., 2011). CsdA_RBD could specifically bind to single-stranded G-rich regions, which might have special biological implication for unwinding the duplex nearby. Future studies should focus on the mechanism of how the dimeric CsdA unwinds RNAs with complicated tertiary and secondary structures.

STAR★METHODS

Detailed methods are provided in the online version of this paper and include the following:

- KEY RESOURCES TABLE
- CONTACT FOR REAGENT AND RESOURCE SHARING
- EXPERIMENTAL MODEL AND SUBJECT DETAILS
- METHOD DETAILS
 - Protein Purification
 - Crystallization and Data Collection
 - NMR Spectroscopy and Data Processing
 - Solution Structure Calculation
 - SAXS Data Collection and Analysis
 - RNA Transcription
 - Mutagenesis
 - Electrophoresis Mobility Shift Assay
 - ATPase Activity Assay
 - Helicase Activity Assay
 - Fluorescence Polarization Assay
 - Fluorescence Polarization Analysis
 - Gene Deletion and Western Blotting
- DATA AND SOFTWARE AVAILABILITY

SUPPLEMENTAL INFORMATION

Supplemental Information includes nine figures and two tables and can be found with this article online at <https://doi.org/10.1016/j.str.2017.09.013>.

AUTHOR CONTRIBUTIONS

Conceptualization, L.X., Y.T., and Y.S.; Investigation, L.X., L.W., F.L., L.W., B.Z., M.L., J.Z., Q.G., R.Z., and J.W.; Methodology, J.P. and Z.Z.; Writing – Original Draft, L.X.; Writing – Review & Editing, L.X., Y.T., and Y.S.; Supervision and Coordination, Y.S.

ACKNOWLEDGMENTS

The authors thank Prof. Ke Ruan for helpful discussions. The authors also thank the staff at BL17U and Na Li, Guangfeng Liu, and Hongjin Wu of BL19U2 of the Shanghai Synchrotron Radiation Facility (SSRF), for their assistance with the X-ray and SAXS data collection. This work was supported by Ministry of Science and Technology of China (2016YFA0500700); the Strategic Priority Research Program of the Chinese Academy of Science (XDB08010101 and XDB08030302); Chinese National Natural Science Foundation

(31400629); Chinese National Natural Science Foundation (31330018); and Chinese National Natural Science Foundation (31270760).

Received: June 20, 2017

Revised: August 21, 2017

Accepted: September 20, 2017

Published: October 26, 2017

REFERENCES

- Beran, R.K., and Simons, R.W. (2001). Cold-temperature induction of *Escherichia coli* polynucleotide phosphorylase occurs by reversal of its autoregulation. *Mol. Microbiol.* 39, 112–125.
- Bessman, M.J. (1963). Deoxynucleoside monophosphate kinases – deoxynucleoside monophosphate + ATP reversible deoxynucleoside diphosphate + ADP. *Method. Enzymol.* 6, 166–176.
- Bhattacharya, A., Tejero, R., and Montelione, G.T. (2007). Evaluating protein structures determined by structural genomics consortia. *Proteins* 66, 778–795.
- Bizebard, T., Ferlenghi, I., lost, I., and Dreyfus, M. (2004). Studies on three *E. coli* DEAD-box helicases point to an unwinding mechanism different from that of model DNA helicases. *Biochemistry* 43, 7857–7866.
- Caruthers, J.M., Johnson, E.R., and McKay, D.B. (2000). Crystal structure of yeast initiation factor 4A, a DEAD-box RNA helicase. *Proc. Natl. Acad. Sci. USA* 97, 13080–13085.
- Charollais, J., Dreyfus, M., and lost, I. (2004). CsdA, a cold-shock RNA helicase from *Escherichia coli*, is involved in the biogenesis of 50S ribosomal subunit. *Nucleic Acids Res.* 32, 2751–2759.
- Clery, A., Blatter, M., and Allain, F.H.T. (2008). RNA recognition motifs: boring? Not quite. *Curr. Opin. Struct. Biol.* 18, 290–298.
- Clore, G.M., and Gronenborn, A.M. (1994). Multidimensional heteronuclear nuclear-magnetic-resonance of proteins. *Methods Enzymol.* 239, 349–363.
- Cordin, O., Banroques, J., Tanner, N.K., and Linder, P. (2006). The DEAD-box protein family of RNA helicases. *Gene* 367, 17–37.
- Datsenko, K.A., and Wanner, B.L. (2000). One-step inactivation of chromosomal genes in *Escherichia coli* K-12 using PCR products. *Proc. Natl. Acad. Sci. USA* 97, 6640–6645.
- Del Campo, M., and Lambowitz, A.M. (2009). Structure of the yeast dead box protein Mss116p reveals two wedges that crimp RNA. *Mol. Cell* 35, 598–609.
- Delaglio, F., Grzesiek, S., Vuister, G.W., Zhu, G., Pfeifer, J., and Bax, A. (1995). Nmrpipe – a multidimensional spectral processing system based on unix pipes. *J. Biomol. NMR* 6, 277–293.
- Diges, C.M., and Uhlenbeck, O.C. (2001). *Escherichia coli* DbpA is an RNA helicase that requires hairpin 92 of 23S rRNA. *EMBO J.* 20, 5503–5512.
- Dumont, S., Cheng, W., Serebrov, V., Beran, R.K., Tinoco, I., Jr., Pyle, A.M., and Bustamante, C. (2006). RNA translocation and unwinding mechanism of HCV NS3 helicase and its coordination by ATP. *Nature* 439, 105–108.
- Emsley, P., and Cowtan, K. (2004). Coot: model-building tools for molecular graphics. *Acta Crystallogr. D Biol. Crystallogr.* 60, 2126–2132.
- Fischer, H., Neto, M.D., Napolitano, H.B., Polikarpov, I., and Craievich, A.F. (2010). Determination of the molecular weight of proteins in solution from a single small-angle X-ray scattering measurement on a relative scale. *J. Appl. Crystallogr.* 43, 101–109.
- Guntert, P., Mumenthaler, C., and Wuthrich, K. (1997). Torsion angle dynamics for NMR structure calculation with the new program DYANA. *J. Mol. Biol.* 273, 283–298.
- Guzman, L.M., Belin, D., Carson, M.J., and Beckwith, J. (1995). Tight regulation, modulation, and high-level expression by vectors containing the arabinose Pbad promoter. *J. Bacteriol.* 177, 4121–4130.
- Hardin, J.W., Hu, Y.X., and McKay, D.B. (2010). Structure of the RNA binding domain of a DEAD-box helicase bound to its ribosomal RNA target reveals a novel mode of recognition by an RNA recognition motif. *J. Mol. Biol.* 402, 412–427.
- Holm, L., and Rosenstrom, P. (2010). Dali server: conservation mapping in 3D. *Nucleic Acids Res.* 38, W545–W549.
- Huen, J., Lin, C.L., Golzarroshan, B., Yi, W.L., Yang, W.Z., and Yuan, H.S. (2017). Structural insights into a unique dimeric DEAD-box helicase CshA that promotes RNA decay. *Structure* 25, 469–481.
- Humphrey, W., Dalke, A., and Schulten, K. (1996). VMD: visual molecular dynamics. *J. Mol. Graph. Model.* 14, 33–38.
- lost, I., and Dreyfus, M. (1994). Messenger-RNAs can be stabilized by DEAD-box proteins. *Nature* 372, 193–196.
- lost, I., and Dreyfus, M. (2006). DEAD-box RNA helicases in *Escherichia coli*. *Nucleic Acids Res.* 34, 4189–4197.
- Jankowsky, E., and Fairman, M.E. (2007). RNA helicases – one fold for many functions. *Curr. Opin. Struct. Biol.* 17, 316–324.
- Klostermeier, D., and Rudolph, M.G. (2009). A novel dimerization motif in the C-terminal domain of the *Thermus thermophilus* DEAD box helicase Hera confers substantial flexibility. *Nucleic Acids Res.* 37, 421–430.
- Konarev, P.V., Petoukhov, M.V., Volkov, V.V., and Svergun, D.I. (2006). ATSAS 2.1, a program package for small-angle scattering data analysis. *J. Appl. Crystallogr.* 39, 277–286.
- Konarev, P.V., Volkov, V.V., Sokolova, A.V., Koch, M.H.J., and Svergun, D.I. (2003). PRIMUS: a Windows PC-based system for small-angle scattering data analysis. *J. Appl. Crystallogr.* 36, 1277–1282.
- Kossen, K., Karginov, F.V., and Uhlenbeck, O.C. (2002). The carboxy-terminal domain of the DExD/H protein YxiN is sufficient to confer specificity for 23 S rRNA. *J. Mol. Biol.* 324, 625–636.
- Kozin, M.B., and Svergun, D.I. (2001). Automated matching of high- and low-resolution structural models. *J. Appl. Crystallogr.* 34, 33–41.
- Laskowski, R.A., MacArthur, M.W., Moss, D.S., and Thornton, J.M. (1993). Procheck - a program to check the stereochemical quality of protein structures. *J. Appl. Crystallogr.* 26, 283–291.
- Laskowski, R.A., Rullmann, J.A.C., MacArthur, M.W., Kaptein, R., and Thornton, J.M. (1996). AQUA and PROCHECK-NMR: programs for checking the quality of protein structures solved by NMR. *J. Biomol. NMR* 8, 477–486.
- Linden, M.H., Hartmann, R.K., and Klostermeier, D. (2008). The putative RNase P motif in the DEAD box helicase Hera is dispensable for efficient interaction with RNA and helicase activity. *Nucleic Acids Res.* 36, 5800–5811.
- Linder, P. (2006). Dead-box proteins: a family affair – active and passive players in RNP-remodeling. *Nucleic Acids Res.* 34, 4168–4180.
- Linder, P., and Jankowsky, E. (2011). From unwinding to clamping - the DEAD box RNA helicase family. *Nat. Rev. Mol. Cell Biol.* 12, 505–516.
- Linder, P., Lasko, P.F., Ashburner, M., Leroy, P., Nielsen, P.J., Nishi, K., Schnier, J., and Slonimski, P.P. (1989). Birth of the D-E-A-D box. *Nature* 337, 121–122.
- Lorsch, J.R., and Herschlag, D. (1998). The DEAD box protein eIF4A. 1. A minimal kinetic and thermodynamic framework reveals coupled binding of RNA and nucleotide. *Biochemistry* 37, 2180–2193.
- Maris, C., Dominguez, C., and Allain, F.H.T. (2005). The RNA recognition motif, a plastic RNA-binding platform to regulate post-transcriptional gene expression. *FEBS J.* 272, 2118–2131.
- Mccoy, A.J., Grosse-Kunstleve, R.W., Adams, P.D., Winn, M.D., Storoni, L.C., and Read, R.J. (2007). Phaser crystallographic software. *J. Appl. Crystallogr.* 40, 658–674.
- Miczak, A., Kaberdin, V.R., Wei, C.L., and Linchao, S. (1996). Proteins associated with RNase E in a multicomponent ribonucleolytic complex. *Proc. Natl. Acad. Sci. USA* 93, 3865–3869.
- Mohr, G., Del Campo, M., Mohr, S., Yang, Q.S., Jia, H.J., Jankowsky, E., and Lambowitz, A.M. (2008). Function of the C-terminal domain of the DEAD-box protein Mss116p analyzed in vivo and in vitro. *J. Mol. Biol.* 375, 1344–1364.
- Murshudov, G.N., Vagin, A.A., and Dodson, E.J. (1997). Refinement of macromolecular structures by the maximum-likelihood method. *Acta Crystallogr. D Biol. Crystallogr.* 53, 240–255.
- Nakajima, T., Uchida, C., Anderson, S.F., Lee, C.G., Hurwitz, J., Parvin, J.D., and Montminy, M. (1997). RNA helicase A mediates association of CBP with RNA polymerase II. *Cell* 90, 1107–1112.

- Otwinowski, Z., and Minor, W. (1997). Processing of X-ray diffraction data collected in oscillation mode. *Methods Enzymol.* 276 (Pt A), 307–326.
- Polach, K.J., and Uhlenbeck, O.C. (2002). Cooperative binding of ATP and RNA substrates to the DEAD/H protein DbpA. *Biochemistry* 41, 3693–3702.
- Prabu, J.R., Muller, M., Thomae, A.W., Schussler, S., Bonneau, F., Becker, P.B., and Conti, E. (2015). Structure of the RNA helicase MLE reveals the molecular mechanisms for uridine specificity and RNA-ATP coupling. *Mol. Cell* 60, 487–499.
- Pugh, G.E., Nicol, S.M., and Fuller-Pace, F.V. (1999). Interaction of the *Escherichia coli* DEAD box protein DbpA with 23 S ribosomal RNA. *J. Mol. Biol.* 292, 771–778.
- Py, B., Higgins, C.F., Krisch, H.M., and Carpousis, A.J. (1996). A DEAD-box RNA helicase in the *Escherichia coli* RNA degradosome. *Nature* 381, 169–172.
- Resch, A., Vecerek, B., Palavra, K., and Blasi, U. (2010). Requirement of the CsdA DEAD-box helicase for low temperature riboregulation of rpoS mRNA. *RNA Biol.* 7, 796–802.
- Rudolph, M.G., Heissmann, R., Wittmann, J.G., and Klostermeier, D. (2006). Crystal structure and nucleotide binding of the *Thermus thermophilus* RNA helicase Hera N-terminal domain. *J. Mol. Biol.* 361, 731–743.
- Rudolph, M.G., and Klostermeier, D. (2009). The *Thermus thermophilus* DEAD box helicase Hera contains a modified RNA recognition motif domain loosely connected to the helicase core. *RNA* 15, 1993–2001.
- Samatanga, B., and Klostermeier, D. (2014). DEAD-box RNA helicase domains exhibit a continuum between complete functional independence and high thermodynamic coupling in nucleotide and RNA duplex recognition. *Nucleic Acids Res.* 42, 10644–10654.
- Schwieters, C.D., Kuszewski, J.J., Tjandra, N., and Clore, G.M. (2003). The Xplor-NIH NMR molecular structure determination package. *J. Magn. Reson.* 160, 65–73.
- Semenyuk, A.V., and Svergun, D.I. (1991). GNOM - a program package for small-angle scattering data-processing. *J. Appl. Crystallogr.* 24, 537–540.
- Sengoku, T., Nureki, O., Nakamura, A., Satoru, K.I., and Yokoyama, S. (2006). Structural basis for RNA unwinding by the DEAD-box protein *Drosophila* vasa. *Cell* 125, 287–300.
- Shajani, Z., Sykes, M.T., and Williamson, J.R. (2011). Assembly of bacterial ribosomes. *Annu. Rev. Biochem.* 80, 501–526.
- Shen, Y., Delaglio, F., Cornilescu, G., and Bax, A. (2009). TALOS plus: a hybrid method for predicting protein backbone torsion angles from NMR chemical shifts. *J. Biomol. NMR* 44, 213–223.
- Shi, H., Cordin, O., Minder, C.M., Linder, P., and Xu, R.M. (2004). Crystal structure of the human ATP-dependent splicing and export factor UAP56. *Proc. Natl. Acad. Sci. USA* 101, 17628–17633.
- Silverman, E., Edwalds-Gilbert, G., and Lin, R.J. (2003). DExD/H-box proteins and their partners: helping RNA helicases unwind. *Gene* 312, 1–16.
- Sledjeski, D.D., Gupta, A., and Gottesman, S. (1996). The small RNA, DsrA, is essential for the low temperature expression of RpoS during exponential growth in *Escherichia coli*. *EMBO J.* 15, 3993–4000.
- Steimer, L., Wurm, J.P., Linden, M.H., Rudolph, M.G., Wohner, J., and Klostermeier, D. (2013). Recognition of two distinct elements in the RNA substrate by the RNA-binding domain of the *T. thermophilus* DEAD box helicase Hera. *Nucleic Acids Res.* 41, 6259–6272.
- Svergun, D.I. (1999). Restoring low resolution structure of biological macromolecules from solution scattering using simulated annealing. *Biophys. J.* 76, 2879–2886.
- Theissen, B., Karow, A.R., Kohler, J., Gubaev, A., and Klostermeier, D. (2008). Cooperative binding of ATP and RNA induces a closed conformation in a DEAD box RNA helicase. *Proc. Natl. Acad. Sci. USA* 105, 548–553.
- Toone, W.M., Rudd, K.E., and Friesen, J.D. (1991). Dead, a new *Escherichia coli* gene encoding a presumed Atp-dependent RNA helicase, can suppress a mutation in rpsB, the gene encoding ribosomal protein S2. *J. Bacteriol.* 173, 3291–3302.
- Tria, G., Mertens, H.D.T., Kachala, M., and Svergun, D.I. (2015). Advanced ensemble modelling of flexible macromolecules using X-ray solution scattering. *IUCrJ* 2, 207–217.
- Tsai, J.Y., Chen, B.T., Cheng, H.C., Chen, H.Y., Hsiao, N.W., Lyu, P.C., and Sun, Y.J. (2006). Crystal structure of HP0242, a hypothetical protein from *Helicobacter pylori* with a novel fold. *Proteins* 62, 1138–1143.
- Turner, A.M.W., Love, C.F., Alexander, R.W., and Jones, P.G. (2007). Mutational analysis of the *Escherichia coli* DEAD box protein CsdA. *J. Bacteriol.* 189, 2769–2776.
- Wang, S., Overgaard, M.T., Hu, Y., and Mckay, D.B. (2008). The *Bacillus subtilis* RNA helicase YxiN is distended in solution. *Biophys. J.* 94, L1–L3.
- Wang, S.Y., Hu, Y.X., Overgaard, M.T., Karginov, F.V., Uhlenbeck, O.C., and Mckay, D.B. (2006). The domain of the *Bacillus subtilis* DEAD-box helicase YxiN that is responsible for specific binding of 23S rRNA has an RNA recognition motif fold. *RNA* 12, 959–967.
- Wang, W.W., Wang, L.J., Zou, Y., Zhang, J.H., Gong, Q.G., Wu, J.H., and Shi, Y.Y. (2011). Cooperation of *Escherichia coli* Hfq hexamers in DsrA binding. *Genes Dev.* 25, 2106–2117.
- Yan, X.M., Mouillet, J.F., Ou, Q.L., and Sadovsky, Y. (2003). A novel domain within the DEAD-box protein DP103 is essential for transcriptional repression and helicase activity. *Mol. Cell Biol.* 23, 414–423.
- Yang, Q.S., and Jankowsky, E. (2006). The DEAD-box protein Ded1 unwinds RNA duplexes by a mode distinct from translocating helicases. *Nat. Struct. Mol. Biol.* 13, 981–986.

STAR★METHODS**KEY RESOURCES TABLE**

REAGENT or RESOURCE	SOURCE	IDENTIFIER
Antibodies		
Anti-Sigma S	Abcam	Cat#ab81737; RRID: AB_2564522
Goat anti-Mouse IgG (H+L) Secondary	ThermoFisher	Cat#35518; RRID: AB_614942
Antibody, DyLight 680		
Bacterial and Virus Strains		
BL21(DE3)	Novagen	Cat#69450
Top10	ThermoFisher	Cat#C404010
Chemicals, Peptides, and Recombinant Proteins		
AMP-PMP	Sigma-Aldrich	Cat#25612-73-1
ATP	Sangon Biotech	Cat#A600311
Deposited Data		
Solution structure, RRM-like domain of DEAD-box protein, CsdA	This paper	5B88
Crystal structure, DEAD-box RNA helicase	This paper	5GI4
Crystal structure, DEAD-box RNA helicase	This paper	5GJU
Oligonucleotides		
UUUUUUUUUUUUUUUU	Takara	N/A
73mer RNA of E. coli 23s rRNA (nucleotides 2508-2580)	RNA transcription	N/A
dsRNA _{26/14-Cy5}	Takara	N/A
dsRNA _{26/14-Cy5-GGGU}	Takara	N/A
GCUUUACGGUGCUA	Takara	N/A
RNAs used in FP assays	Takara	N/A
Recombinant DNA		
Plasmid: pET28	(Wang et al., 2011)	N/A
Plasmid: pBAD18	(Guzman et al., 1995)	N/A
Software and Algorithms		
HKL2000	(Otwinowski and Minor, 1997)	http://www.hkl-xray.com/hkl-2000
Phaser	(Mccoy et al., 2007)	www CCP4.ac.uk/html/phaser.html
COOT	(Emsley and Cowtan, 2004)	https://www2.mrc-lmb.cam.ac.uk/personal/pemsley/coot/
REFMAC5	(Murshudov et al., 1997)	http://www CCP4.ac.uk/
PROCHECK	(Laskowski et al., 1993)	www CCP4.ac.uk/html/procheck_man/index.html
PyMOL	Schrödinger, LLC	https://www.pymol.org/
VMD	(Humphrey et al., 1996)	www.ks.uiuc.edu/Research/vmd/
NMRDraw	(Delaglio et al., 1995)	https://spin.niddk.nih.gov/bax/software/NMRPipe/
NMRpipe	(Delaglio et al., 1995)	https://spin.niddk.nih.gov/bax/software/NMRPipe/
Sparky	T.D. Goddard and D.G. Kneller, SPARKY 3	https://www.cgl.ucsf.edu/home/sparky/
Talos+	(Shen et al., 2009)	https://spin.niddk.nih.gov/bax/software/TALOS/
Cyana	(Guntert et al., 1997)	www.cyana.org/
Xplor-NIH	(Schwieters et al., 2003)	https://nmr.cit.nih.gov/xplor-nih/
PSVS	(Bhattacharya et al., 2007)	psvs-1_5-dev.nesg.org/
ATSAS package	(Konarev et al., 2006)	https://www.embl-hamburg.de/biosaxs/software.html

(Continued on next page)

Continued

REAGENT or RESOURCE	SOURCE	IDENTIFIER
PRIMUS	(Konarev et al., 2003)	https://www.embl-hamburg.de/biosaxs/primus.html
GNOM	(Semenyuk and Svergun, 1991)	https://www.embl-hamburg.de/biosaxs/gnom.html
DAMMIN	(Svergun, 1999)	https://www.embl-hamburg.de/biosaxs/dammin.html
SUPCOMB	(Kozin and Svergun, 2001)	https://www.embl-hamburg.de/biosaxs/supcomb.html
EOM 2.0	(Tria et al., 2015)	https://www.embl-hamburg.de/biosaxs/eom.html

CONTACT FOR REAGENT AND RESOURCE SHARING

Further information and requests for resources and reagents should be directed to and will be fulfilled by the Lead Contact, Yunyu Shi (yyshi@ust.edu.cn).

EXPERIMENTAL MODEL AND SUBJECT DETAILS

The *E.coli* CsdA expression recombinants CsdA_FL (residues 1-629, the full length of CsdA), CsdA_5-211 (residues 5-211), CsdA_372 (residues 1-372), CsdA_1-445 (residues 1-445), CsdA_564 (residues 1-564), CsdA_218-445 (residues 218-445), CsdA_RBD (residues 482-564) and CsdA_DDRBD (residues 372-564), CsdA_DD (residues 372-445), CsdA_RBD-629 (residues 482-629) were cloned into the modified pET28 TEV expression vector with an N-terminal hexa-histidine (6 × His) TEV tag (Wang et al., 2011) using common molecular cloning techniques and verified by DNA sequencing. CsdA_5-211 and CsdA_1-445 recombinants were expressed in *E.coli* strain BL21 (DE3) using LB medium at 37°C for 6 hours by adding 0.2 mM and 1 mM isopropyl-β-d-thiogalactopyranoside (IPTG) at $A_{600} \sim 0.8\text{-}1.0$, while CsdA_FL, CsdA_564, CsdA_218-445, CsdA_RBD, CsdA_DDRBD, CsdA_DD and CsdA_RBD-629 were induced at 16°C for 16 hours by adding 0.5 mM IPTG. The $^{15}\text{N}/^{13}\text{C}$ -labeled-CsdA_RBD protein used for NMR experiments was expressed by culturing the bacteria in LeMaster and Richards (LR) medium supplemented with 0.5 g/L ^{15}N labeled NH₄Cl and 2.5 g/L ^{13}C labeled glucose as unique nitrogen and carbon sources. Cells were harvested by centrifugation at 5000 rpm and stored at -80°C.

METHOD DETAILS**Protein Purification**

E.coli cells were re-suspended in binding buffer (20 mM Tris-HCl, 1 M NaCl, pH 8.0), sonicated and centrifuged at 13000 rpm to separate the precipitates and supernatants, which were then further loaded onto nickel-chelating columns (GE Healthcare) for purification. All recombinants were washed and eluted using 40 mM and 500 mM imidazole in binding buffer. Pooled fractions were digested at 20°C for 3 hours using tobacco etch virus (TEV) protease for His-TEV tag cleavage. Uncleaved recombinants and His-tagged TEV protease were removed by reloading the fractions onto nickel-chelating columns while the first flow-through fractions were concentrated. CsdA_5-211, CsdA_DDRBD, CsdA_DD, CsdA_RBD and CsdA_RBD-629 were subjected to a 16/60 Superdex 75 column (GE Healthcare) for further purification while CsdA_372, CsdA_1-445, CsdA_564 and CsdA_FL were purified using 16/60 Superdex 200 column (GE Healthcare). The purified $^{15}\text{N}/^{13}\text{C}$ -labeled-CsdA_RBD protein was concentrated to 1.3 mM after dialysis into NMR buffer (50 mM Na₂HPO₄, 150 mM NaCl, pH 6.5).

Crystallization and Data Collection

The purified CsdA_5-211 and CsdA_1-445 were dialyzed overnight against the storage buffer (50 mM Na₂HPO₄, 200 mM NaCl) and concentrated to 10 mg/ml mixed with 2 mM AMP-PNP for crystallization. Crystals of CsdA_5-211 were obtained from hanging drop vapor diffusion approaches at 20°C by mixing 1.0 μl of reservoir solution, 2.1 M DL-malic acid, pH 7.0, with 1.0 μl of protein that was incubated with 2 mM AMP-PNP and equilibrated against 200 μL reservoir solution. The CsdA_1-445 proteins were incubated with 2 mM AMP-PNP and 2 mM MgCl₂ using a reservoir solution of 0.2 M trimethylamine N-oxide dihydrate, 0.1 M Tris pH 8.5, and 20% w/v polyethylene glycol monomethyl ether 2,000. For data collection, all crystals that were selected from the hanging drops were soaked in a cryo-protectant solution consisting of 30% (v/v) glycerol in reservoir solution and flash cooled in liquid N₂ immediately. For all crystals, X-ray diffraction data were recorded at beamline BL17U at the Shanghai Synchrotron Radiation Facility (SSRF). The native data were processed using HKL2000 (Otwinowski and Minor, 1997) and the structures were determined by molecular replacement by Phaser (McCoy et al., 2007). The structure models were further built and refined using COOT (Emsley and Cowtan, 2004) and REFMAC5 (Murshudov et al., 1997) while the high quality of the structures were validated using PROCHECK (Laskowski et al., 1993). All structural images were created using PyMOL (Schrodinger, LLC. The PyMOL Molecular Graphics System, Version 1.8, 2015) or VMD (Humphrey et al., 1996).

NMR Spectroscopy and Data Processing

All NMR spectra of 1.3 mM $^{15}\text{N}/^{13}\text{C}$ -labeled-CsdA_RBD mixed with 10% or 100% D₂O in NMR buffer were recorded at 298K on a Bruker DMX600 spectrometer with a cryoprobe. The proton chemical shift was referenced relative to the frequency of the 1H resonance of water, which was further calibrated by DSS. The ^{13}C and ^{15}N chemical shifts were referenced indirectly based on the ratio of the proton frequency to the carbon/nitrogen frequency (Clore and Gronenborn, 1994). Backbone and side-chain resonance assignments were obtained by recording the following spectra: 2D $^1\text{H}/^{15}\text{N}$ HSQC and $^1\text{H}/^{13}\text{C}$ HSQC, 3D CBCA(CO)NH, CBCANH, (H)C(CO)NH-TOCSY, HBHA(CO)NH, H(C)(CO)NH-TOCSY, HNCO, HN(CA)CO, HCCH-COSY and HCCH-TOCSY. Distance restraints were obtained from 3D $^1\text{H}/^{15}\text{N}$ -NOESY-HSQC and $^1\text{H}/^{13}\text{C}$ -NOESY-HSQC. All NMR spectral data were processed with NMRPipe and NMRDraw software (Delaglio et al., 1995). All assignments were generated using Sparky software T. D. Goddard and D. G. Kneller, SPARKY 3, University of California, San Francisco.

Solution Structure Calculation

The solution structure of CsdA_RBD was calculated using the program CYANA (Guntert et al., 1997) with NOE distance restraints and 96 torsion angle restraints. Dihedral angle (Φ and Ψ) restraints were generated from the backbone and $^{13}\text{C}\beta$ chemical shifts using the program TALOS+ (Shen et al., 2009). NOESY peak assignments were performed automatically with a chemical shift tolerance of 0.02 ppm for 1H and 0.2 ppm for heavy atoms. Structure calculations started from 100 random conformers for which 10,000 torsion angle dynamics steps were performed during the standard CYANA simulated annealing protocol. We further refined the structures with Xplor-NIH program (Schwieters et al., 2003) in vacuum and in explicit water, respectively, with the final distance restraints (noe restraints) and dihedral restraints generated from CYANA calculations. 20 conformers without violations and with the lowest final energies were selected to form a representative ensemble. Structural qualities were assessed by PSVS (Bhattacharya et al., 2007) and PROCHECK (Laskowski et al., 1996). See Table S1 in Supplemental Information for detailed statistical analysis of the ensembles.

SAXS Data Collection and Analysis

The highly purified CsdA_218-445, CsdA_1-445 and CsdA_FL were dialyzed against SAXS buffer (20 mM Tris-HCl, 200 mM NaCl, pH 7.5) and concentrated to 10 mg/ml. The concentration for CsdA_218-445, CsdA_1-445 and CsdA_FL were set to 1 mg/ml, 3 mg/ml, and 5 mg/ml. SAXS data for CsdA_1-445 and CsdA_FL were collected at beamline 12ID-B of the Advanced Photon Sources (APS) at Argonne National Laboratory (Chicago, IL, USA) using a wavelength of 1.033 Å. The SAXS data for CsdA_218-445 were collected at SSRF BL19U2. The data were analyzed in the ATSAS package (Konarev et al., 2006) following the standard procedures. After subtracting buffer scattering, the data curves from different concentrations were scaled and merged using PRIMUS (Konarev et al., 2003). GNOM (Semenyuk and Svergun, 1991) was employed for estimating the particle maximum dimension (D_{max}) and real space R_g of the proteins and calculation of the pair distance distribution function (PDDF). Low-resolution dummy atom shape model of CsdA_218-445, CsdA_1-445 and CsdA_FL was generated by DAMMIN (Svergun, 1999). SUPCOMB (Kozin and Svergun, 2001) was used to superimpose the crystal structure of CsdA_218-445 to the corresponding shape model generated by DAMMIN. To better address the flexibilities of CsdA_1-445 and CsdA_FL, we carried out EOM 2.0 (advanced ensemble optimized method) (Tria et al., 2015) analysis. The ensemble model building procedure of CsdA_1-445 was as follows. First, a pool of 10,000 random structures of the CsdA_1-445 dimer was generated by treating the two RecA1 domains and the RecA2-DD dimer as rigid bodies, respectively. Then, 200 cycles of genetic algorithm were run using data points up to 0.2 Å⁻¹. The size and fractions of occupancy of the ensemble was optimized during the genetic algorithm. For CsdA_FL, the model building procedure was the same as that of CsdA_1-445, except that the rigid bodies were two RecA1 domains, RecA2-DD dimer and two RBD domains.

RNA Transcription

The RNA used in ATPase experiments is a 73mer RNA of *E. coli* 23s rRNA that contains the stem of hairpin 92 of 23S rRNA (nucleotides 2508-2580) and was transcribed and purified *in vitro*. The DNA template for transcribing 73mer RNA was synthesized in Takara Bio, Inc., and dissolved in DEPC water to a final concentration of 60 µM. The reaction mixture was formed by 10 mM Tris, 10mM DTT, 10 mM NTPs, 40 mM MgCl₂, 0.3 µM T7 template, 0.3 µM DNA templates, 3mg/ml T7 RNA polymerase. The reaction was performed at 37°C for 4h. After transcription, the transcription products were treated with 0.1 of total volume (0.1V) 0.5 M EDTA, 0.1V 5 M NaCl, 3V absolute alcohol and incubated at -40°C overnight. Then the transcription products were centrifuged, discarded the supernatant and dissolved the precipitate RNA in 1.5 ml DEPC-treated water. Add equal volume RNA loading buffer (Takara), 90°C for 5 min and ice-cool for 5min. The RNA samples were separated from the 10% denaturing polyacrylamide gel and purified by Elutrap (Whatman). The final 73mer RNA were dialyzed into DEPC-treated water and concentrated to 1mM.

Mutagenesis

The CsdA_1-445 R5 mutant (five residues in the DD, A425-L429, are all mutated to R) and CsdA_RBD-R549AL551A mutant were created by using the QuickChange Mutagenesis Kit (Takara). The mutated coding DNAs were transformed into BL21 (DE3) and confirmed by DNA sequencing.

Electrophoresis Mobility Shift Assay

The RNA used in binding assay is 5'-FAM-labeling 32mer RNA. The binding buffer was 20 mM Tris, 150 mM NaCl, pH 7.5. The concentration of 32mer RNA is 40nM and the concentration of the protein is from 100 μ M with 2-fold dilutions. The reactions were incubated for 1.5h at room temperature and separated on 6% native polyacrylamide gels. The RNA bands were analyzed using Typhoon 7000 (GE Healthcare).

ATPase Activity Assay

The ATPase activities of CsdA_1-445, CsdA_564, CsdA_372 and the control were measured in the existence of either poly U₁₄ (UUUUUUUUUUUUUU) or a 73mer RNA of *E. coli* 23s rRNA (nucleotides 2508-2580). The ATPase activities of CsdA_1-445 and CsdA_564 were also measured in the existence of other RNAs used in the RNA-binding affinity assay. The reactions were all performed at 25 °C using a classical method as previously described (Bessman, 1963). In the pyruvate kinase coupling lactase dehydrogenase assay, the hydrolysis of ATP is coupled to the oxidation of NADH continuously, which is detected by measuring the decrease of absorbance at 340 nm. The reaction system contains 10 μ M proteins, 400 μ M phosphoenolpyruvate, 10 U/ml pyruvate kinase, 10 U/ml lactate dehydrogenase, 2 mM ATP, 2 mM MgCl₂, 0.25 mM NADH and 2-8 μ M RNA (for poly U₁₄ is 3 μ M). The reaction buffer is 10 mM Hepes, 75 mM NaCl, pH 7.5. The reaction was monitored at CLARIOstar (BMG LABTECH). Data from three independent experiments were analyzed using Origin 8.0 software.

Helicase Activity Assay

Helicase activity assays were carried out for the CsdA_1-445, CsdA_564 and CsdA_372, CsdA_DDRBD, CsdA_1-445 R5 mutant (five residues in the DD, A425-L429, are all mutated to R) and the control. The RNA substrates designed in helicase assays were Cy5-labeled-duplex RNA purchased from Takara Bio, Inc. Helicase assays were performed using previous protocols (Bizebard et al., 2004) with minor changes. Cy5-labeled-duplex RNA (dsRNA_{26/14-cy5}: Cy5-labeled-GCUUUACGGUGCUA, AACAAAACA AAAUAGCACCGUAAAGC) or (dsRNA_{26/14-cy5}-GGGU: Cy5-labeled-GCUUUACGGUGCUA, AACAAAACGGGUAGCACCGU AAGC) were 0.1 μ M, unlabeled single strand “trap” RNA (GCUUUACGGUGCUA) which is counterpart of the labeled duplex was 2 μ M, the concentration of proteins was at 30 μ M. 2mM ATP/MgCl₂ (or AMP-PNP/MgCl₂), 2units/ μ L RNase inhibitor (Takara). The reaction buffer was 20mM Hepes, 100mM NaCl, pH 7.5, 10% glycerol. The reactions were incubated at 25 °C and initiated by addition of 0.1 μ M labeled RNA. 10 μ L aliquots were removed from reaction mixtures every 5 min from 0 min to 50 min (or 0 min to 30 min). Aliquots were stopped by mixing with 2.5 μ L of a stop buffer (20 mM EDTA, 50% glycerol, 3mg/ml proteinase K) for another 10 min at 25 °C and put on ice before electrophoresis. 10 μ L aliquots were subjected to 18% native PAGE (19 : 1 acrylamide/bis) in 0.5 \times TBE buffer on ice. The RNA bands were analyzed using Typhoon 7000 (GE Healthcare).

Fluorescence Polarization Assay

The FP assay was carried out as previously described (Wang et al., 2011) with minor changes. The lyophilized 5'-FAM (carboxyfluorescein)-labeled RNA duplex (dsRNA_{26/14-FAM}: 5'-FAM-labeled-GCUUUACGGUGCUA, AACAAAACAAAAUAGCACCGUAAAGC) and 5'-FAM-labeled 32mer RNA (5'-FAM- GCAGGUCCAAGGGU**GGGC**UGUU**GCC**CAUU) that is a 32mer RNA containing the regions (show in bold) comprising of the stem of hairpin 92 of 23S rRNA (Steimer et al., 2013), 21mer RNA (5'-FAM-GGGU**GGGC**CUG UUC**GCC**CAUU-3'), 15ssRNA (5'-FAM- GCAGGUCCAAGGGU), 17stem RNA (5'-FAM-**GGGC**CUGUUC**GCC**CAUU), GGGU RNA (5'-FAM-GGGU-3'), GGGG RNA (5'-FAM-GGGG-3'), GGGA RNA (5'-FAM-GGGA-3'), GGGC RNA (5'-FAM-GGGC-3'), CACG RNA (5'-FAM-CACG-3'), A₅ RNA (5'-FAM-AAAAA-3'), C₅ RNA (5'-FAM-CCCCC-3'), U₅ RNA (5'-FAM-UUUUU-3') were purchased from Takara Bio, Inc., and dissolved in DEPC water to a final concentration of 0.1 mM. The 0.1 mM 5'-FAM-labeled RNA was diluted to 1 μ M in FP buffer (20 mM Tris, pH 7.5, 200 mM NaCl). Different concentrated CsdA constructs, CsdA_FL, CsdA_564, CsdA_1-445, CsdA_DDRBD, CsdA_RBD629, GST_RBD, CsdA_DD were diluted in FP buffer to a series of concentrations. The FAM-labeled RNAs were diluted to 60 nM in FP buffer before the assay. The total volume of the FP assay system was adjusted to 200 μ L using the FP buffer. The FAM-labeled RNAs were mixed with CsdA constructs and incubated for 10 min. The equilibrium dissociation constants of different RNAs and CsdA constructs were determined by measuring the FP. Samples were excited at 485 nm and fluorescence was detected at 525 nm. The FP assay was collected in a SpectraMax M5 (Molecular Devices) plate reader at 20°C.

Fluorescence Polarization Analysis

The fluorescence polarization of RNA upon protein binding was fitted similar to Wang et al. (Wang et al., 2011) and was expressed as following:

$$FP = FP_{ini} + \frac{\max}{2nR} \left(K_d + P + nR - \sqrt{-4nPR + (K_d + P + nR)^2} \right)$$

where FP is the fluorescence intensity, FP_{ini} the initial FP of RNA without any protein, P the protein concentration, R the concentration of labeled RNA, n the binding stoichiometry (Protein : RNA ratio), K_d the equilibrium dissociation constant. Data were fitted with stoichiometry n fixed to 1, which stands for one monomer molecule binds to one RNA molecule, unless a good fitting is not obtained. Standard errors were obtained by fitting the data to the above equation.

Gene Deletion and Western Blotting

The isogenic *csdA* and *csdADD* deficient strain Top10 (Δ *csdA*, Δ *csdADD*) were obtained as previously described (Datsenko and Wanner, 2000; Resch et al., 2010) with minor changes. Wild-type (WT) full-length and mutant *csdA* genes were inserted into the pBAD18-kanamycin plasmid (Guzman et al., 1995) containing the arabinose pBAD promoter. The wild-type full-length plasmids (pBADCsdA) were transformed into Δ *csdA* and Δ *csdADD* deficient strains by electroporation. The bacterial strains were grown in Luria Broth (LB) medium supplemented with kanamycin (100 µg/ml). Cultures were grown in LB medium at 37°C or at 24°C to an A_{600} of 0.5 and harvested and boiled in binding buffer. Equal amounts of total protein were separated by SDS-polyacrylamide gel electrophoresis (SDS-PAGE) thereupon and transferred onto a nitrocellulose membrane using a wet blotting method. The membranes were washed three times and every one time for 8 minutes with TBST after incubation with primary anti-Sigma S antibodies (Abcam) at 4°C overnight, followed by incubation with 1: 15000 secondary antibodies conjugated with IRDye 800 CW fluorescent dyes (Takara) at room temperature for 1 h. The bands were detected at 700 nm and 800 nm using an Odyssey Image scanner (Gene Company).

DATA AND SOFTWARE AVAILABILITY

Atomic coordinates and structure factors for the reported crystal structures have been deposited in the PDB under ID codes 5B88 (RRM-like domain of DEAD-box protein, CsdA), 5GI4 (DEAD-box RNA helicase) and 5GJU (DEAD-box RNA helicase).