DNA-guided DNA cleavage at moderate temperatures by Clostridium butyricum Argonaute

Jorrit W. Hegge1,‡, Daan C. Swarts2,3,†, Stanley D. Chandradoss4, Tao Ju Cui4, Jeroen Kneppers1, Martin Jinek3, Chirlim Joo4 and John van der Oost1,∗

1Laboratory of Microbiology, Department of Agrotechnology and Food Sciences, Wageningen University, 6708 WE Wageningen, The Netherlands, 2Laboratory of Biochemistry, Department of Agrotechnology and Food Sciences, Wageningen University, 6708 WE Wageningen, The Netherlands, 3Department of Biochemistry, University of Zurich, CH-8057 Zurich, Switzerland and 4Kavli Institute of NanoScience, Department of BioNanoScience, Delft University of Technology, 2629 HZ Delft, The Netherlands

Received February 25, 2019; Revised April 05, 2019; Editorial Decision April 11, 2019; Accepted May 03, 2019

ABSTRACT

Prokaryotic Argonaute proteins (pAgos) constitute a diverse group of endonucleases of which some mediate host defense by utilizing small interfering DNA guides (siDNA) to cleave complementary invading DNA. This activity can be repurposed for programmable DNA cleavage. However, currently characterized DNA-cleaving pAgos require elevated temperatures (≥65°C) for their activity, making them less suitable for applications that require moderate temperatures, such as genome editing. Here, we report the functional and structural characterization of the siDNA-guided DNA-targeting pAgo from the mesophilic bacterium Clostridium butyricum (CbAgo). CbAgo displays a preference for siDNAs that have a deoxyadenosine at the 5′-end and thymidines at nucleotides 2–4. Furthermore, CbAgo mediates DNA-guided DNA cleavage of AT-rich double stranded DNA at moderate temperatures (37°C). This study demonstrates that certain pAgos are capable of programmable DNA cleavage at moderate temperatures and thereby expands the scope of the potential pAgo–based applications.

INTRODUCTION

Eukaryotic Argonaute proteins (eAgos) play a key role in RNA interference (RNAi) processes (1–3). As the core of the multiprotein RNA-induced silencing complex (RISC), eAgos bind small non-coding RNA molecules as guides to direct the RISC complex towards complementary RNA targets (3–5). Reflecting their physiological function, variation among eAgos is observed with respect to the presence or absence of a catalytic site, and to their potential to interact with other proteins (6). Depending on the eAgo and on the sequence complementarity between guide and target RNA, eAgo-guide complexes either catalyze endonucleolytic cleavage of the target RNA (7) or indirectly silence the target RNA by repressing its translation and promoting its degradation through recruitment of additional silencing factors (8). Independent of the mechanism, eAgo-mediated RNA binding generally results in sequence-specific silencing of gene expression. As such, eAgos can coordinate various cellular processes by regulating intracellular RNA levels.

Prokaryotes also encode Argonaute proteins (pAgos) (9,10). Various pAgos share a high degree of structural homology with eAgos as both pAgos and eAgos adopt the same four domain (N-PAZ-MID-PIWI) architecture (9–12). Despite their structural homology, several recently characterized pAgos have distinct functional roles and different guide and/or target preferences compared to eAgos. For example, several pAgos have been implicated in host defense by directly targeting DNA instead of RNA (13–16). One of the best characterized mechanisms that pAgos utilize is DNA-guided DNA interference, which is demonstrated for pAgos from Thermus thermophilus (TtAgo), Pyrococcus furiosus (PfAgo) and Methanocaldococcus jannaschii (MjAgo) (13–15,17–20). These pAgos use 5′-end phosphorylated small interfering DNAs (siDNAs) for recognition and successive cleavage of complementary DNA targets. This mechanism enables both TtAgo and PmAgo to mediate host defense against invading nucleic acids. Prokaryotes lack homologs of eukaryotic enzymes that are involved in guide biogenesis (21). Instead, both TtAgo and MjAgo—besides the canonical siDNA-

‡ To whom correspondence should be addressed. Tel: +31317483108; Email: john.vanderoost@wur.nl
† The authors wish it to be known that, in their opinion, the first two authors should be regarded as Joint First Authors.
∗ Present address: Stanley D. Chandradoss, Oxford Nanoimaging Ltd, Oxford, UK.
© The Author(s) 2019. Published by Oxford University Press on behalf of Nucleic Acids Research.
This is an Open Access article distributed under the terms of the Creative Commons Attribution Non-Commercial License (http://creativecommons.org/licenses/by-nc/4.0/), which permits non-commercial re-use, distribution, and reproduction in any medium, provided the original work is properly cited. For commercial re-use, please contact journals.permissions@oup.com
dependent target cleavage termed ‘slicing’—exhibit an alternative nuclease activity termed ‘chopping’ (14,17). Chopping facilitates autonomous generation of small DNA fragments from dsDNA substrates. Subsequently, these DNA fragments generated during chopping can serve as siDNAs for canonical slicing (14,17).

TiAgo and PfAgo can be programmed with short synthetic siDNA which allows them to target and cleave dsDNA sequences of choice in vitro (13,15). This activity has enabled the repurposing of PfAgo as an universal restriction endonuclease for in vitro molecular cloning (22). In addition, a diagnostic TiAgo-based application termed NAVIGATER (Nuclease Enrichment Via DNA Guided Argonaute from *T. thermophilus*) was developed which enables enhanced detection of rare nucleic acids with single nucleotide precision (23). In analogy with the now commonly used CRISPR-Cas9 and CRISPR-Cas12a enzymes (24–26), it has also been suggested that pAgos could be repurposed as next-generation genome editing tools (27).

However, due to the thermophilic nature (optimum activity temperature $\geq 65^\circ\text{C}$) and low levels of endonuclease activity at the relevant temperatures (20–37$^\circ\text{C}$), it is unlikely that the well-studied TiAgo, PfAgo and MjAgo are suitable for genome editing. The quest for a pAgo that can cleave dsDNA at moderate temperatures has resulted in the characterization of the Argonaute protein from *Natrornobacterium gregory* (NgAgo), which was claimed to be the first pAgo suitable for genome editing purposes (28). However, the study reporting this application has been retracted after a series of reproducibility issues (28–30). Instead, it has been suggested that NgAgo targets RNA rather than DNA (31).

Although considerable efforts have been made to elucidate the mechanisms and biological roles of pAgos, efforts have mainly focused on pAgo variants from (hyper)thermophiles. This has left a large group of mesophilic pAgos unexplored. We here report the characterization of the Argonaute protein from the mesophilic bacterium *Clostridium butyricum* (CbAgo). We demonstrate that CbAgo can utilize siDNA guides to cleave both ssDNA and dsDNA targets at moderate temperatures ($37^\circ\text{C}$). In addition, we have elucidated the macromolecular structure of CbAgo in complex with a siDNA guide and complementary ssDNA target in a catalytically competent state. CbAgo displays an unusual preference for siDNAs with a deoxyadenosine at the 5’-end and thymidines at nucleotides 2–4. The programmable DNA endonuclease activity of CbAgo provides a foundation for the development of pAgo-based applications at moderate temperatures.

**MATERIALS AND METHODS**

**Plasmid construction**

The CbAgo gene was codon harmonized for *Escherichia coli* BI21 (DE3) and inserted into a pET-His6 MBP TEV cloning vector (obtained from the UC Berkeley MacroLab, Addgene #29565) using ligation independent cloning (LIC) using oligonucleotides oDS067 and oDS068 (Supplementary Table S1) to generate a protein expression construct that encodes the CbAgo polypeptide sequence fused to an N-terminal tag comprising a hexahistidine sequence, a maltose binding protein (MBP) and a Tobacco Etch Virus (TEV) protease cleavage site.

**Generation of the double mutant**

CbAgo double mutant (D541A, D611A) was generated using an adapted Quick Directed Mutagenesis Kit instruction manual (Stratagene). The primers were designed using the web-based program primerX (http://bioinformatics.org/primerx).

**CbAgo expression and purification**

The CbAgo WT and DM proteins were expressed in *E. coli* BI21(DE3) Rosetta™ 2 (Novagen). Cultures were grown at $37^\circ\text{C}$ in LB medium containing 50 $\mu$g ml$^{-1}$ kanamycin and 34 $\mu$g ml$^{-1}$ chloramphenicol until an OD$_{600\text{nm}}$ of 0.7 was reached. CbAgo expression was induced by addition of isopropyl β-D-1-thiogalactopyranoside (IPTG) to a final concentration of 0.1 mM. During the expression cells were incubated at $18^\circ\text{C}$ for 16 h with continuous shaking. Cells were harvested by centrifugation and lysed by sonication (Bandalone, Sonopuls. 30% power, 1 son for 2 son for 5 min) in lysis buffer containing 20 mM Tris–HCl pH 7.5, 250 mM NaCl, 5 mM imidazole, supplemented with an EDTA free protease inhibitor cocktail tablet (Roche). The soluble fraction of the lysate was loaded on a nickel column (HisTrap Hp, GE Healthcare). The column was extensively washed with wash buffer containing 20 mM Tris–HCl pH 7.5, 250 mM NaCl and 30 mM imidazole. Bound protein was eluted by increasing the concentration of imidazole in the wash buffer to 250 mM. The eluted protein was dialysed at $4^\circ\text{C}$ overnight against 20 mM HEPES pH 7.5, 250 mM KCl and 1 mM dithiothreitol (DTT) in the presence of 1 mg TEV protease (32) to cleave of the His6-MBP tag. Next, the cleaved protein was diluted in 20 mM HEPES pH 7.5 to lower the final salt concentration to 125 mM KCl. The diluted protein was applied to a heparin column (HiTrap Heparin HP, GE Healthcare), washed with 20 mM HEPES pH 7.5, 125 mM KCl and 1 mM dithiothreitol (DTT) in the presence of 1 mg TEV protease (32) to cleave of the His6-MBP tag. Next, the cleaved protein was diluted in 20 mM HEPES pH 7.5 to lower the final salt concentration to 125 mM KCl. The diluted protein was applied to a heparin column (HiTrap Heparin HP, GE Healthcare), washed with 20 mM HEPES pH 7.5, 125 mM KCl and eluted with a linear gradient of 0.125–2 M KCl. Next, the eluted protein was loaded onto a size exclusion column (Superdex 200 16/600 column, GE Healthcare) and eluted with 20 mM HEPES pH 7.5, 500 mM KCl and 1 mM DTT. Purified CbAgo protein was diluted in size exclusion buffer to a final concentration of 5 $\mu$M. Aliquots were flash frozen in liquid nitrogen and stored at $-80^\circ\text{C}$.

**Co-purification nucleic acids**

To 500 pmol of purified CbAgo in SEC buffer CaCl$_2$ and protease K (Ambion) were added to final concentrations of 5 mM CaCl$_2$ and 250 $\mu$g/ml protease K. The sample was incubated for 4 h at 65°C. The nucleic acids were separated from the organic fraction by adding Roti phenol/chloroform/isoamyl alcohol pH 7.5–8.0 in a 1:1 ratio. The top layer was isolated and nucleic acids were precipitated using ethanol precipitation by adding 99% ethanol in a 1:2 ratio supplied with 0.5% linear polymerized acrylamide as a carrier. This mixture was incubated overnight at $-20^\circ\text{C}$ and centrifuged in a table centrifuge at 16 000 g for 30 min. Next, the nucleic acids pellet was washed with...
70% ethanol and solved in 50 µl MilliQ water. The purified nucleic acids were treated with either 100 µg/ml RNase A (Thermo), 2 units DNase I (NEB) or both for 1 h at 37°C and resolved on a denaturing urea polyacrylamide gel (15%) and stained with SYBR gold.

**Single stranded activity assays**

Unless stated otherwise 5 pmol of each ChAgo, siDNA and target were mixed in a ratio of 1:1:1, in 2× reaction buffer containing 20 mM Tris–HCl pH 7.5 supplemented with 500 µM MnCl₂. The target was added after the ChAgo and siDNA had been incubated for 15 min at 37°C. Then the complete reaction mixture was incubated for 1 h at 37°C. The reaction was terminated by adding 2× RNA loading dye (95% formamide, 0.025% bromophenol blue, 5 mM EDTA) and heating it for 5 min at 95°C. After this, the samples were resolved on a 20% denaturing (7 M urea) polyacrylamide gel. The gel was stained with SYBR gold nucleic acid stain (Invitrogen) and imaged using a G:BOX Chemi imager (Syngene).

**Double stranded activity assay**

In two half reactions 12.5 pmol of ChAgo was loaded with either 12.5 pmol of forward or reverse siDNA in reaction buffer containing 10 mM Tris–HCl pH 7.5, 10 µg/ml BSA and 250 µM MnCl₂. The half reactions were incubated for 15 min at 37°C. Next, both half reactions were mixed together and 120 ng target plasmid was added after which the mixture was incubated for 1 h of 37°C. After the incubation the target plasmid was purified from the mixture using a DNA clean and concentrate kit (DNA Clean & Concentrator™-5, Zymogen) via the supplied protocol. The purified plasmid was subsequently cut using either EcoRI-HF (NEB) or SapI-HF (NEB) in Cutsmart buffer (NEB) for 30 min at 37°C. A 6× DNA loading dye (NEB) was added to the plasmid sample prior to resolving it on a 0.7% agarose gel stained with SYBR gold (Invitrogen). In the double stranded assays in which linear plasmid was targeted, 300 ng pUC IDT, linearized by EcoRI, was added to both half reactions and incubated for 1 h of 37°C. The reaction products were resolved on a 0.7% agarose gel.

**Crystallization**

To reconstitute the ChAgo DM–siDNA–target DNA complex, siDNA and target DNA were pre-mixed at a 1:1 ratio, heated to 95°C, and slowly cooled to room temperature. The formed dsDNA duplex (0.5 M) was mixed with ChAgo DM in SEC buffer at a 1:1:4 ratio (ChAgo DM:duplex DNA), and MgCl₂ was added to a final concentration of 5 mM. The sample was incubated for 15 min at 20°C to allow complex formation. The complex was crystallized at 20°C using the hanging drop vapour diffusion method by mixing equal volumes of complex and reservoir solution. Initial crystals were obtained at a ChAgo DM concentration of 5 mg/ml with a reservoir solution consisting of 4 M sodium formate. Data was collected from crystals grown obtained using a complex concentration of 4.3 mg/ml and reservoir solution containing 3.8 M Sodium Formate and 5 mM NiCl₂ at 20°C. For cryoprotection, crystals were transferred to a drop of reservoir solution and flash-cooled in liquid nitrogen.

X-ray diffraction data were measured at beamline X06DA (PXIII) of the Swiss Light Source (Paul Scherrer Institute, Villigen, Switzerland). Data were indexed, integrated and scaled using AutoPROC (35). Crystals of the ChAgo–siDNA–target DNA complex diffracted to a resolution of 3.55 Å and belonged to space group P6₁ 2 2, with one copy of the complex in the asymmetric unit. The structure was solved by molecular replacement using PhaserMR (34). As search model, the structure of TrAgo in complex with guide and target DNA strands (PDB: 5GQ9) was used after removing loops and truncating amino acid side chains. Phases obtained using the initial molecular replacement solution were improved by density modification using phenix.resolve (35) and phenix.morph_model (36). The atomic model was built manually in Coot (37) and refined using phenix.refine (37). The final binary complex model contains ChAgo residues 1–463 and 466–748, guide DNA residues 1–16, and target DNA residues (–18)(–1).

**Structure analysis**

Core root means square deviations (rmsd) of structure alignments were calculated using Coot SSM superpose (38). Hydrogen bonding interactions were analysed using PDBePISA (39), while other interactions with the Mg²⁺ ion were determined manually in Coot (37). Figures were generated using PyMOL (Schrödinger).

**Single-molecule experimental set-up**

Single-molecule fluorescence FRET measurements were performed with a prism-type total internal reflection fluorescence microscope. Cy3 and Cy5 molecules were excited with 532 and 637 nm wavelength, respectively. Resulting Cy3 and Cy5 fluorescence signal was collected through a 60× water immersion objective (UPlanSApo, Olympus) with an inverted microscope (IX73, Olympus) and split by a dichroic mirror (635dcxr, Chroma). Scattered laser light was blocked out by a triple notch filter (NF01-488/532/635, Semrock). The Cy3 and Cy5 signals were recorded using a EM-CCD camera (iXon Ultra, DU-897U-C50-#BV, Andor Technology) with exposure time 0.1 s. All single-molecule experiments were done at room temperature (22 ± 2°C).

**Fluorescent DNA and RNA preparation**

The RNAs with amine-modification (amino-modifier C6-U phosphoramidite, 10-3039, Glen Research) were purchased from STPharm (South Korea) and DNAs with amine-modification (internal amino modifier iAmMC6T) Ella biotech (Germany). The guide and target strands were labeled with donor (Cy3) and acceptor (Cy5), respectively, using the NHS-ester form of Cy dyes (GE Healthcare). 1 µl of 1 mM of DNA/RNA dissolved in MilliQ H₂O was added to 5 µl labeling buffer of (freshly prepared) sodium tetraborate (380 mg/10ml, pH 8.5). 1 µl of 20 mM dye (1 mg in 56 µl DMSO) was added and incubated overnight at room temperature in the dark, followed by washing and ethanol precipitation. The labeling efficiency was ~100%.
Single-molecule sample preparation

A microfluidic chamber was incubated with 20 μl Streptavidin (0.1 mg/ml, Sigma) for 30 s. Unbound Streptavidin was washed with 100 μl of buffer T50 (10 mM Tris–HCl pH 8.0, 50 mM NaCl buffer). The 50 μl of 50 pM acceptor-labeled target construct were introduced into the chamber and incubated for 1 min. Unbound labeled constructs were washed with 100 μl of buffer T50. The ChAgo binary complex was formed by incubating 10 nM purified ChAgo with 1 nM of donor-labeled guide in a buffer containing 50 mM Tris–HCl pH 8.0 (Ambion), 1 mM MnCl2, and 100 mM NaCl (Ambion) at 37°C for 20 min. For binding rate (k_on) measurements, the binary complex was introduced into the fluidics chamber using syringe during the measurement. The experiments were performed at the room temperature (23 ± 1°C).

For fluorescence Guide Loading Experiments before immobilizing ChAgo on the single-molecule surface, 1 μl of 5 μM His-tagged apo-ChAgo was incubated with 1 μl of 1 μg/ml biotinylated anti-6x His antibody (Abcam) for 10 min. Afterward, the mixture was diluted 500× in T50 and 50 μl were loaded in the microfluidic channel for 30 s incubation, followed by washing with 100 μl of T50 buffer. Cy3-labeled ssDNA (0.1 nM) was applied to the microfluidic chamber in imaging buffer (50 mM Tris–HCl pH 8.0, 100 mM NaCl, 1 mM MnCl2, 1 mM Trolox ((±)-6-hydroxy-2,5,7,8-tetramethylchromane-2-carboxylic acid), supplemented with an oxygen-scavenging system (0.5 mg/ml glucose oxidase (Sigma), 85 mg/ml catalase (Merck) and 0.8% (v/v) glucose (Sigma)).

Single-molecule data acquisition and analysis

CCD images of time resolution 0.1 or 0.3 s were recorded, and time traces were extracted from the CCD image series using IDL (ITT Visual Information Solution). Co-localization between Cy3 and Cy5 signals was carried out with a custom-made mapping algorithm written in IDL. The extracted time traces were processed using Matlab (MathWorks) and Origin (Origin Lab).

The binding rate (k_on) was determined by first measuring the time between when ChAgo binary complex was introduced to a microfluidic chamber and when the first ChAgo-guide docked to a target; and then fitting the time distribution with a single-exponential growth curve, \( A(1 - e^{-kt}) \). The dissociation rate was estimated by measuring the dwell time of a binding event. A dwell time distribution was fitted by single-exponential decay curve (\( Ae^{-t/\Delta t} \)).

Fluorescence competition experiments

MBP-tagged ChAgo was immobilized on the quartz surface using an anti-MBP antibody. An equimolar mixture of let7 DNA guide (Cy3 labeled) and let7 RNA guide (Cy5 labeled) in imaging buffer was introduced to the microfluidic chamber. After 5 min, 10 snapshots of independent fields of view with simultaneous illumination were collected to estimate the amount of guide molecules bound to protein. Movies were taken for 200 s (2000 frames) at continuous illumination of Cy3 and Cy5 molecules to determine the dwell times of the binding events. Dwell times were binned in a histogram and fitted with a single exponential decay curve.

FRET targeting experiments of ATTT and AAAA guide target combinations

100 pM of target construct annealed with biotin handle were flushed in the microfluidic chamber. After incubation of 1 min, the microfluidic chamber was rinsed with 100 μl T50 buffer. 10 nM of apo-ChAgo was loaded with 1 nM of ATTT seed DNA guide or with AAAA seed DNA guide at 37°C for 30 min in imaging buffer after which the mixture is introduced inside the microfluidic chamber. Movies of 200 s were taken at continuous illumination of the Cy3 signal. Site specific protein target interactions were identified as FRET signals and were further analyzed.

RESULTS

ChAgo mediates siDNA-guided ssDNA cleavage

ChAgo was successfully expressed in E. coli from a codon-optimized gene using a T7-based pET expression system and purified (Supplementary Figure S1A). To determine the guide and target binding characteristics of ChAgo, we performed single-molecule experiments using Förster resonance energy transfer (FRET). We immobilized either Cy5-labeled single stranded RNA or DNA targets (FRET acceptor) on a polymer-coated quartz surface (Figure 1A). Next, we introduced ChAgo in complex with either a Cy3-labeled siRNA or siDNA guide (FRET donor) and recorded the interactions. Strikingly, ChAgo could utilize both siRNAs and siDNAs to bind DNA or RNA targets (Figure 1B). To test which guide is preferentially bound by ChAgo we performed a competition assay in which ChAgo was immobilized into the microfluidic chamber, and an equimolar mixture of siDNA and siRNAs was introduced. While only short-lived interactions (average dwell time: 0.48 s) were observed for siRNA, siDNA was strongly bound (average dwell time: 44 s) by ChAgo (Figure 1C). This results suggests that ChAgo utilizes siDNA rather than siRNA as a guide.

ChAgo is phylogenetically closest related to the clade of halobacterial pAgos, among which also pAgo from Natriunobacterium gregoryi (NgAgo) can be found (Figure 1D and Supplementary Figure S2). A multiple sequence alignment of ChAgo with other pAgos (Supplementary Figure S1B) suggests that ChAgo contains the conserved DEDX catalytic residues (where X can be a D, H or N) which are essential for nuclease activity in ‘slicening’ Agos (40). In the case of ChAgo, this concerns residues D541, E577, D611 and D727.

To confirm whether ChAgo indeed is an active nuclease, we performed in vitro activity assays in which ChAgo was loaded with either synthetic siDNAs or siRNAs (21 nucleotides in length). Next the complexes were incubated at 37°C with 45-nucleotide complementary single stranded RNA or DNA target oligonucleotides. While no activity was found in any of the combinations in which siRNAs or target RNAs were used, ChAgo was able to cleave target DNAs in a siDNA-dependent manner (Figure 1E). In
Figure 1. ChAgo exhibits DNA-guided DNA endonuclease activity at 37°C. (A) Left: Overview of the single molecule assay to determine the binding characteristics of ChAgo. Right: FRET diagram of a ChAgo–siDNA complex that has three complementary base pairs (2–4nt) to the DNA target. Indicated is the dwell time (Δτ). (B) Comparison of the binding rates (k_on) of ChAgo in complex with siDNA or siRNA to bind DNA or RNA targets. The rates are similar for each nucleic acid type guide and target. N is the number of base paired nucleotides. (C) Dwell time histograms showing ChAgo preferentially binds siDNAs in siDNA-siRNA competition experiments. (D) Schematic phylogenetic tree of characterized pAgos. (E) ChAgo exhibits DNA-guided DNA endonuclease activity. Upper panel: Sequence of the synthetic let7 miRNA-based siDNA guide and target DNA sequences that were used for the in vitro cleavage assays. Lower panel: ChAgo, guides and targets were mixed in a 1:1:1 molar ratio and incubated for 1 h at 37°C. Catalytic mutant ChAgoDM was used as a control. Cleavage products were analysed by denaturing polyacrylamide electrophoresis. (F) ChAgo displays highest activity at 37°C. ChAgo and siDNA were mixed and pre-incubated at various temperatures for 10 min. Next, target DNA was added and the sample was incubated for 1 h at the same temperature. ChAgoDM was used as a control. Cleavage products were analysed by denaturing polyacrylamide electrophoresis. (G) Quantified data of a ChAgo-mediated siDNA-guided ssDNA cleavage turnover experiment using 5 pmol target DNA and increasing concentrations of ChAgo-siDNA (1.25–20 pmol). (H) FRET diagram showing that a cleavage compatible ChAgo-siDNA remains bound to a fully complementary target DNA (N = 21) during the entire the measurement (340 seconds).
agreement with the predicted DEDD catalytic site (Supplementary Figure S1B), alanine substitutions of two of aspartic acids (D541A, D611A) in the expected catalytic tetrad abolished the nuclease activity, demonstrating that the observed siDNA-guided ssDNA endonucleolytic activity was indeed catalyzed by the DEDD catalytic site. To further investigate the full temperature range at which ChAgo is active, we performed additional cleavage assays at temperatures ranging from 10 to 95°C. While ChAgo displayed the highest activity at its physiologically relevant temperature (37°C), ChAgo also catalyzed siDNA-guided target DNA cleavage at temperatures as low as 10°C and as high as 50°C (Figure 1F).

When ChAgo–siDNA complexes and target ssDNA substrates (45 nt) were mixed in equimolar amounts, cleavage of the target DNA was not complete after one hour incubation (Figure 1E). Therefore, we investigated the substrate turnover kinetics of ChAgo by monitoring the cleavage assays in a time course using variable ChAgo:siDNA:target DNA ratios (Figure 1G). A rapid burst of activity was observed during the first minute, likely indicating the first target binding and cleavage event. This stage was followed by a slow steady state, suggesting that under these conditions the ChAgo–siDNA complex slowly dissociates from the cleaved target DNA product before being able to bind and cleave a new target DNA strand. The cleavage kinetics were confirmed using single-molecule assays which demonstrated that the ChAgo-siDNA complex remains bound to the DNA target (N = 21) for several minutes (Figure 1H), which prevents ChAgo–siDNA complexes from binding and cleaving new DNA targets. Thus, while ChAgo functions as a multi-turnover nuclease enzyme, its steady-state rate is limited by product release.

Structure of ChAgo in the cleavage-competent conformation

To investigate the molecular architecture of ChAgo in light of its biochemical activity, we crystallized ChAgoDM in complex with both a 21 nt siDNA and a 19 nt DNA target, and solved the structure of the complex at 3.54 Å resolution (Figure 2 and Supplementary Table S1). Like other Agos, ChAgo adopts a bilobed conformation in which one of the lobes comprises the N-terminal, linker L1, and PAZ domains, which are linked by linker L2 to the other lobe comprising the MID and PIWI domains. Nucleotides 2–16 of the siDNA constitute a 15 base-pairs A-form-like duplex with the target DNA (Figure 2A). The 5′-terminal nucleotide of the siDNA is anchored in the MID domain pocket, where the 5′-phosphate group of the siDNA makes numerous interactions with MID domain residues and the C-terminal carboxyl group of ChAgo (Supplementary Figure S3), as observed for in structures of TtAgo and MjAgo bound to siDNAs (12,19,41). To test whether the interactions with the 5′-phosphate group of the siDNA are important for ChAgo activity, we performed target DNA cleavage assays in which we used siDNAs with a 5′ phosphate or a 5′ hydroxyl group (Supplementary Figure S4). As observed for other pAgos (42,43), ChAgo is able to utilize both siDNAs for target DNA cleavage, but it cleaves target DNA much more efficiently when the siDNA contains a 5′-phosphate group. This is in agreement with the siDNA–protein interactions predicted form the crystal structure. Furthermore, the backbone phosphates of the siDNA seed segment appear to form hydrogen-bonding and ionic interactions with specific residues in the MID, PIWI and L1 domains (Supplementary Figure S3). At the distal end of the siDNA-target DNA duplex, the N-domain residue His35 caps the duplex by stacking onto the last base pair. After this point, the remaining 3′-terminal nucleotides of the siDNA are disordered, while the target DNA bends away from the duplex and enters the cleft between the N-terminal and PAZ domains. In agreement with other ternary pAgo complexes (18,44,45), the PAZ domain pocket, which normally binds the 3′ end of the guide in a binary Ago-guide complex, is empty.

ChAgo is phylogenetically closely related to TtAgo (Figure 1D). However, ChAgo is 63 amino acids (9.2%) longer than TtAgo (748 amino acids versus 685 amino acids) and ChAgo and TtAgo share only 23% sequence identity. Superposition of the ChAgo complex structure with the structure of TtAgo bound to a siDNA and DNA target (PDB: 4NCB) (Figure 2C) reveals that the macromolecular architecture and conformation of these TtAgo and ChAgo structures are highly similar (Core root mean square deviation of 3.0 Å over 563 residues). In the TtAgo structure, which is thought to represent a catalytically competent state, a ‘glutamate finger’ side chain (Glu512TrtAgo) is inserted into the catalytic site completing the catalytic DDED tetrad (44). Similarly, the corresponding residue in ChAgo (Glu577) is located within a flexible loop and is positioned near the other catalytic residues (Figure 2D; Asp541, Asp611 and Asp727). All pAgos and eAgos characterized to date cleave the target strand in between nucleotide 10 and 11 of the target strand. In line with the consensus, the catalytic residues of ChAgo perfectly align with the scissile phosphate linking these nucleotides in our structure (Figure 2D). This observation implies that this structure represents the cleavage competent conformation of ChAgo.

Only 15 siDNA–target DNA base pairs are formed in the complex, which suggests that additional siDNA-target DNA binding is not essential for target DNA cleavage. To determine the minimum siDNA length that ChAgo requires for target binding, we performed single-molecule fluorescence assays. First, ChAgo was immobilized on a surface and next it was incubated with S′-phosphorylated Cy3-labeled siDNAs (Figure 2E). These assays demonstrate that ChAgo can bind siDNAs with a minimal length of 12 nucleotides. Next, we determined the minimum siDNA length for ChAgo-siDNA mediated target DNA cleavage (Figure 2F). In line with the observation that the ChAgo adopts a cleavage-competent confirmation when only 14 base pairs are formed, ChAgo can cleave target DNAs when programmed with siDNAs as short as 14 nt (forming 13 siDNA-target DNA base pairs) under the tested conditions. This resembles the activity of P/βAgo, MjAgo and MpAgo, which require siDNAs with a minimal length of 15 nt to catalyze target DNA cleavage (14,15,43). Only TtAgo has been reported to mediate target DNA cleavage with siDNAs as short as 9 nt (12).
Figure 2. Structure of ChAgo in complex with a siDNA and a DNA target. (A) Upper panel: Schematic diagram of the domain organization of ChAgo. L1 and L2 are linker domains. Lower panel: Sequences of the siDNA (red) and target DNA (blue). Nucleotides that are unordered in the structure are coloured grey. See also Supplementary Table S1. (B) Overall structure of the ChAgo-siDNA-target DNA complex. Domains are coloured according to the colour scheme in panel A. (C) Structural alignment of ChAgo (green) and TtAgo (light purple; PDB: 4NCB). Core Root Mean Square Deviation of 3.0 Å over 563 residues. (D) Close-up views of the aligned DDED catalytic sites of ChAgo (green) and TtAgo (light purple; PDB: 4NCB). Modelled side chains of D541 and D611 in ChAgo are indicated with green asterisks. The glutamate finger of both pAgos (E512 in TtAgo or E577 in ChAgo) are inserted into the catalytic site. The scissile phosphate between nucleotide −10 and −11 of the target DNA strand (blue) is indicated with a black asterisk in the left panel. (E) Total internal reflection microscopy (TIRM) was used to determine the minimal length for siDNA to be bound by ChAgo. Left panel: Graphical overview of the TIRM method. Right panel: Histogram with TIRM results demonstrated that synthetic siDNAs of at least 12 nt in length are efficiently bound by ChAgo. The red line indicates the total number of countable molecules within the microscope image. The raw microscope images are given in Supplementary Figure S5. (F) ChAgo mediates target DNA cleavage with siDNAs as short as 14 nucleotides. ChAgo was incubated with siDNA and target DNA in a 1:1:1 ratio. Cleavage products were analyzed by denaturing polyacrylamide electrophoresis.

**ChAgo associates with plasmid-derived siDNAs in vivo**

It has previously been demonstrated that certain pAgos co-purify with their guides and/or targets during heterologous expression in *Escherichia coli* (13,16). To determine whether ChAgo also acquires siDNAs during expression, we isolated and analyzed the nucleic acid fraction that co-purified with ChAgo. Denaturing polyacrylamide gel electrophoresis revealed that ChAgo co-purified with small nucleotides with a length of ~12–19 nucleotides (Figure 3A). These nucleic acids were susceptible to DNase I but not to RNase A treatment, indicating that ChAgo acquires 12–19 nucleotide long siDNAs in vivo, which fits with its observed binding and cleavage activities in vitro (Figures 1 and 2).

We cloned and sequenced the siDNAs that co-purified with ChAgo to determine their exact length and sequence. The majority of the siDNAs had a length of 16 nucleotides and are complementary to the plasmid used for expres-
The sequence of the siDNA affects ChAgO activity

To investigate if the 5′-terminal nucleotide of the siDNA affects the activity of ChAgO, we performed cleavage assays. ChAgO was loaded with siRNA guides with varied nucleotides at position 1 (g1N) and incubated with complementary target DNAs (Figure 4A). Surprisingly, the highest cleavage rates were observed when ChAgO was loaded with siDNAs containing a 5′-T, followed by siDNAs containing 5′-A. ChAgO bound 5′-G or 5′-C siDNAs displayed slightly lower initial cleavage rates. Also for other pAgos the g1N preference observed in vivo is not reflected in the in vitro activities; TtAgO (which preferentially co-purifies with g1C siDNAs) as well as PfAgO and MpAgO (of which the in vivo g1N preferences are unknown) demonstrate no clear preference for a specific g1N during in vitro cleavage reactions (13,17,43). Instead, the preference of TtAgO for 5′-C siDNAs is determined by specific recognition of a guanosine nucleotide in the corresponding position t(-1) in the target DNA (17). Indeed, TtAgO structures and models have revealed base-specific interactions with target strand guanine, while base-specific interactions with the 5′-terminal cytidine in the siDNA are less obvious (17). Similarly, we predict no obvious base-specific interactions with the 5′-terminal cytidine in the structure of the ChAgO complex (Supplementary Figure S7). When we investigated potential base-specific interactions with the base at the opposing target strand t(-1) position, we observed that the t(-1) thymine base is not placed in the t(-1) binding pocket as has been observed in TtAgO, RsAgO and hAGO2 (17,46,47). Instead, the thymine bases are flipped and stacks on Phe557 that also caps the siDNA-target DNA duplex (Supplementary Figure S7). At present, we are unable to rationalize the preferential co-purification of 5′-adenosine siDNAs with ChAgO.

In order to characterize the seed segment of ChAgO, and to test whether the seed length changes depending on the nature of the guide and the target (i.e. DNA versus RNA), we performed additional single-molecule binding assays. The length of seed was determined based on the minimal number of complementary nucleotide pairs between guide and target that were required to achieve a stable binding event. When only nt 2–4 of the siDNA are complementary to the DNA and RNA targets, ChAgO–siDNA complexes bound to the DNA target with an average dwell time 58-fold longer compared to RNA target-binding (Figure 4B). When nt 2–7 of the siDNA were complementary to the target, the ChAgO–siDNA complex stably bound to both to target DNA and RNA beyond our observation time of 300 s. This suggests ChAgO prefers DNA targets above RNA targets and that the seed segment of the siDNAs bound by ChAgO comprises nt 2–7.

Next, we set out to investigate whether ChAgO displays a preference for siDNAs with thymidines at nt 2–4 in vitro, similar to the observed sequence preference for siDNAs that co-purified with ChAgO in vivo. ChAgO was incubated with siDNAs in which nt 2–4 were varied and complementary target DNAs were added. In contrast to the 5′-base preference, the preference for thymidines in the nt 2–4 segment that we observed in vivo is also reflected in vitro: ChAgO displays the highest target cleavage rates when programmed with siDNAs containing thymidines at nt 2–4 (Figure 4C). To confirm these findings, we performed single-molecule assays in which we compared the target binding properties of ChAgO–siDNA complexes containing siDNAs with either thymidines or a deoxyadenosines at position 2–4. These assays demonstrate that the dwell time of ChAgO loaded with siDNA containing thymidines at nt 2–4 on a target was 18-fold longer compared to ChAgO loaded with siDNA containing deoxyadenosines at nt 2–4 (Figure 4D). Combined, these data indicate that ChAgO displays a preference for siDNAs containing thymidines at position 2–4.

A pair of ChAgO–siDNA complexes can cleave double stranded DNA

Thermophilic pAgos have successfully been used to generate double stranded DNA breaks in plasmid DNA (13,15).
Figure 4. The siDNA sequence affects CbAgo activity. (A) CbAgo has no strong 5′-end nucleotide preference. CbAgo was incubated with siDNA with varied 5′-end and incubated with complementary DNA targets. Cleavage products were analysed by denaturing polyacrylamide electrophoresis and quantified. Graphs display the amount of target DNA cleaved. Error bars indicate the standard variation of three independent experiments. (B) Histograms displaying dwell time of CbAgoDM–siDNA complexes binding either DNA or RNA targets with a varied sequence complementarity (N = number of complementary nucleotides between the siDNA and the target, starting at nt 2. Thus N 3 = nt 2 – 4). The photobleaching limit is reached where the signal is deactivated (300 s). (C) CbAgo preferentially utilizes siDNAs with thymidines at position 2–4. CbAgo–siDNA complexes with siDNA in which nt 2–4 varied were incubated with complementary DNA targets. Cleavage products were analysed by denaturing polyacrylamide electrophoresis and quantified. Graphs display the amount of target DNA cleaved. Error bars indicate the standard variation of three independent experiments. (D) Histograms displaying dwell time of CbAgoDM in complex with a 5′-ATTT siDNA or 5′-AAAA siDNA binding to a target DNA. Interactions of CbAgo are on average ~18-fold longer with the siDNA containing a 5′-ATTT motif compared to interactions with siDNAs containing a 5′-AAAA motif.

As each pAgo–siDNA complex targets and cleaves a single strand of DNA only, two individual pAgo–siDNA complexes are required for dsDNA cleavage, each targeting another strand of the target dsDNA. Although all pAgos characterized so far appear to lack the ability to actively unwind or displace a dsDNA duplex substrate, it has been proposed that, at least in vitro, thermophilic pAgos rely on elevated temperatures (>65°C) to facilitate local melting of the dsDNA targets to target each strand of the DNA individually. However, CbAgo is derived from a mesophilic organism and we therefore hypothesize that it is able to mediate protection against invading DNA at moderate temperatures (37°C). To test if CbAgo can indeed cleave dsDNA targets at 37°C, we incubated apo-CbAgo and pre-assembled CbAgo–siDNA complexes with a target plasmid. Previous studies showed that the ‘chopping’ activity of siDNA-free apo-TtAgo and apo-MjAgo can result in plasmid linearization or degradation, respectively (14,17). We observed that apo-CbAgo converted the plasmid substrate from a supercoiled to open-circular state, possibly by nicking one of the strands, but did not observe significant linearization or degradation of the plasmid DNA (Figure 5A). When the plasmid was targeted by CbAgo loaded with a single siDNA, we also observed loss of supercoiling (Figure 5A). As this activity was not observed with nuclease-deficient CbAgoDM, we conclude that apo–CbAgo and
**Figure 5.** Double stranded plasmid DNA cleavage by CbAgo. (A) Two CbAgo–siDNA complexes can generate double stranded DNA breaks in plasmid DNA. CbAgo–siDNA complexes were pre-assembled and incubated with target plasmid DNA. Cleavage products were analysed by agarose gel electrophoresis (Supplementary Figure S8B) and quantified. The spacing between both CbAgo–siDNA target sites affects the linearization efficiency (nucleotide spacing between the predicted cleavage sites: +15 nt, +10 nt, +5 nt, 0 nt, −5 nt, −10 nt, −15 nt, a single siDNA, no siDNA). With 0 nt spacing, both CbAgo–siDNA complexes are exactly on top of each other. (B) Schematic overview of the pUC IDT target plasmid. Blue arrows indicate target sites while percentages indicate the GC-content of the 100 bp segments in which these target sites are located. (C) Pre-assembled CbAgo-siDNA complexes targeting various pUC IDT segments were incubated with pUC IDT. Cleavage products were incubated with EcoRI or SapI and were further analysed by agarose gel electrophoresis. The GC-content of the segments in which the target sites were located are indicated by the percentage (in blue).

CbAgo–siDNA complexes are generating nicks in dsDNA plasmid targets with their DEDD catalytic site. When using two CbAgo–siDNA complexes, each targeting one strand of the plasmid, we observed that a fraction of the target plasmid DNA becomes linearized (Figure 5A). This implies that CbAgo-siDNA complex-mediated nicking of each of the target plasmid DNA strands resulted in the generation of a double stranded DNA break. Next, we investigated if the spacing between the two siDNAs affects the ability of CbAgo to cleave the plasmid. The most efficient plasmid linearization was achieved when the siDNAs were orientated exactly or almost opposite to each other (Figure 5A).

Finally, we investigated whether the GC-content of the target DNA plays a role during DNA targeting by CbAgo. For TtAgo, it has been observed that AT-rich DNA is cleaved more efficiently than GC-rich DNA (17), possibly because AT-rich DNA is more prone to unwinding. To test if such preference also exists for CbAgo, we designed a target plasmid containing 16 gene fragments of 100 bp complementary to sequences from the human genome, with an increasing GC content (Figure 5B). The target plasmid was digested with a restriction enzyme (SapI or EcoRI) either before or after the incubation with two CbAgo–siDNA complexes (Figure 5B and Supplementary Figure S8A). Interestingly, CbAgo could cleave the negatively supercoiled plasmid, but not the linearized plasmid, suggesting that, like TtAgo, CbAgo relies on the negative supercoiled state of the target plasmid to facilitate local DNA melting before targeting and cleavage can take place (13). Furthermore, CbAgo–siDNA complexes were only able to generate dsDNA breaks in gene fragments with a GC-content of 31% or lower (Figure 5C). We hypothesize that, at moderate temperatures, the supercoiled state of the target plasmid allows for local DNA unwinding especially in AT-rich DNA regions. Taken together, this suggests that, like thermophilic pAgos, CbAgo relies on unwinding of the dsDNA before targeting and cleavage can take place.

**DISCUSSION**

Several prokaryotic Argonaute proteins have been demonstrated to protect their host against invading nucleic acids,
such as plasmid DNA (13,15,16). Similar to TrAgo and RsAgo, ChAgo co-purifies with guides which are preferentially acquired from the plasmid used for its heterologous expression in E. coli. In addition, ChAgo mediates programmable DNA-guided DNA cleavage in vitro. This suggests that, similar to the phylogenetically related TrAgo, also ChAgo can interfere with plasmid DNA via DNA-guided DNA interference.

Sequencing of the nucleic acids that co-purified with ChAgo revealed that ChAgo preferentially associates with siDNAs with a 5′-ATTT-3′ sequence at their 5′ end. It was previously shown that RNA guides utilized by eAgos can be divided into functional segments. These segments are (from 5′ to 3′) the anchor nucleotide (nt 1), the seed (nt 2–8) and sub-seed segments (nt 2–4/2–5), the central (nt 9–12), the 3′ supplementary (nt 13–16) and the tail (nt 17–21) segments (48,49). Extending this knowledge to the siDNAs that co-purified with ChAgo, ChAgo preferentially associates with siDNAs that have a 5′-terminal adenosine anchor (nt 1) and a thymidine-rich sub-seed. In RNAi pathways, the preference for a specific 5′-terminal nucleotide is important for guide RNA loading into a subset of eAgos (50–52). Similarly, several pAgos including RsAgo, TrAgo, and now ChAgo also preferentially associate with specific 5′-terminal nucleotides in vivo (13,16). However, for both ChAgo and TrAgo, there is no clear preference for siDNAs with that specific 5′-base during cleavage assays in vitro. Rather than having a functional importance, the preference of pAgo for a specific nucleotide at the siDNA 5′ end might be a consequence of siDNA generation and/or loading, as has been demonstrated for TrAgo (17). Several studies on human Ago2 have described the importance of the sub-seed segment (nt 2–4/2–5) in its RNA guides (48,53,54). For hAgo2, a complete match between the guide RNA sub-seed segment and the target RNA triggers a conformational change that first exposes the remainder of the seed (nt 5–8/6–8), and eventually the rest of the guide. This facilitates progressive base paring between the guide RNA and the target (55). However, a specific nucleotide preference in the sub-seed segment, as we have observed for ChAgo, has not been described for any other Argonaute protein. The preference for the T-rich sub-seed is not only observed in the in vivo acquired siDNAs, but also plays a clear role during target binding and cleavage assays in vitro. This may reflect a structural preference for these thymidines in the cleft of the PIWI domain. We have not been able to obtain diffracting crystals of ChAgo in complex with siDNAs that have a 5′-ATTT-3′ sequence at the 5′-end. Future research will thus be necessary to determine the structural basis the apparent preference for these nucleotides at these positions. We hypothesize that this bias might reflect the mesophilic nature of ChAgo, which might have better access to AT-rich dsDNA fragments, both for siDNA acquisition and for target cleavage.

Several DNA-targeting pAgos have been repurposed for a range of molecular applications among which a cloning, recombineering and nucleic acid-detection method (22,23,56,57). Additionally, the potential repurposing of pAgo for genome editing applications has previously been discussed (27). However, all characterized DNA-cleaving pAgos to date originate from thermophilic prokaryotes and are solely active at elevated temperatures, which limits the potential repurposing of pAgos for applications that require moderate temperatures, such as genome editing. The biochemical characterization of ChAgo reported herein is the first example of a pAgo that catalyzes siDNA-guided dsDNA cleavage at 37°C, indicating that the pool of mesophilic pAgos contains candidates that—in theory—can be utilized for potential applications that require moderate temperatures, such as genome editing. If ChAgo or other mesophilic pAgos could be harnessed for genome editing, they will have certain advantages over the currently well-established genome editing tools CRISPR-Cas9 and CRISPR-Cas12a. While CRISPR-based genome editing tools can be programmed with a guide RNA to target DNA sequences of choice, target DNA cleavage additionally requires the presence of a protospacer adjacent motif (PAM) next to the targeted sequence (5′-NGG-3′ for Cas9 and 5′-TTTT-3′ for Cas12a) (58). This limits the possible target sites of Cas9 and Cas12a. In contrast, pAgos do not require a PAM for DNA targeting, which would make them much more versatile tools compared to CRISPR-associated nucleases. However, PAM binding by Cas9 and Cas12a also promotes unwinding of dsDNA targets (59–61) which subsequently facilitates strand displacement by the RNA guide, and eventually R-Loop formation. The absence of such mechanism in pAgos might explain their limited nuclease activity on dsDNA targets.

Here, we have demonstrated that ChAgo does not strictly rely on other proteins when targeting AT-rich dsDNA sequences in vitro. As such, this study provides a foundation for future efforts to improve double stranded DNA target accessibility of pAgos and to facilitate the further development of pAgo-based applications at moderate temperatures.

DATA AVAILABILITY

Atomic coordinates and structure factors reported in this paper have been deposited in the Protein Data Bank (PDB) with the accession number PDB: 6QZK.

SUPPLEMENTARY DATA

Supplementary Data are available at NAR Online.

ACKNOWLEDGEMENTS

We are grateful to Meitian Wang, Vincent Olieric and Takashi Tomizaki at the Swiss Light Source (Paul Scherrer Institute, Villigen, Switzerland) for assistance with X-ray diffraction measurements.

Authors contributions: J.W.H. and J.v.d.O. conceived the project and designed the biochemical experiments, which were performed by J.H. and J.K. Single-molecule experiments were designed by S.C., T.J.C. and C.J. and performed by S.C. and T.J.C. X-ray crystallographic analysis was designed and performed by D.C.S. under the supervision of M.J., J.W.H., D.C.S., C.H., M.J., C.J. and J.v.d.O. wrote the manuscript. All authors read and approved the manuscript.
FUNDING

Meitian Wang, Vincent Olieric and Takashi Tomizaki at the Swiss Light Source (Paul Scherrer Institute, Villigen, Switzerland) for assistance with X-ray diffraction measurements; Netherlands Organization of Scientific Research (NWO) [ECO grant 71013002 and NWO-TOP grant 714.015.001 to J.v.d.O.]; Swiss National Science Foundation (SNSF) Project Grant [SNSF 31003A_149393 to M.J.]; European Molecular Biology Organization (EMBO) [ALTF 179-2015 and aALTF 509-2017 to D.C.S.]; M.J. is International Research Scholar of the Howard Hughes Medical Institute and Valleé Scholar of the Bert L & N Kuggie Valleé Foundation; Vidi [864.14.002 to C.J.] of the Netherlands Organization for Scientific research.

Conflict of interest statement. None declared.

REFERENCES

1. Ketting, R.F. (2010) MicroRNA biogenesis and function: an overview. Adv. Exp. Med. Biol., 700, 1–14.
2. Joshua-Tor,L. and Hannon,G.J. (2011) Ancestral roles of small RNAs: an ago-centric perspective. Cold Spring Harb. Perspect. Biol., 3, 1–11.
3. Meister,G. (2013) Argonaute proteins: functional insights and emerging roles. Nat. Rev. Genet., 14, 447–459.
4. Pratt,A.J. and MacRae,I.J. (2009) The RNA-induced silencing complex: a versatile gene-silencing machine. J. Biol. Chem., 284, 17997–17999.
5. Bartel,D.P. (2009) MicroRNAs: target recognition and regulatory functions. Cell, 136, 215–223.
6. Kuhn,C.D. and Joshua-Tor,L. (2013) Eukaryotic Argonautes come into focus. Trends Biochem. Sci., 38, 263–271.
7. Huttner,G. and Simard,M.J. (2008) Argonaute proteins: key players in RNA silencing. Nat. Rev. Mol. Cell Biol., 9, 32–32.
8. Ketting,R.F. (2011) The many faces of RNAi. Dev. Cell, 20, 148–161.
9. Makarova,K.S., Wolf,Y.L., van der Oost,J. and Koonin,E.V. (2009) Prokaryotic homologs of Argonaute proteins are predicted to function as key components of a novel system of defense against mobile genetic elements. Biol. Direct, 4, 29.
10. Swarts,D.C., Koehorst,J.J., Westra,E.R., Schaap,P.J. and Van Der Oost,J. (2015) Effects of argonaute on gene expression in Thermus thermophilus. Acta Crystallogr. Sect. D Biol. Crystallogr., 61, 49–52.
11. Zander,A., Holzmeister,P., Klose,D., Tinnefeld,P. and Grohmann,D. (2014) Single-molecule FRET supports the two-state model of Argonaute action. RNA Biol., 11, 45–56.
12. Willkomm,S., Oellig,C.A., Zander,A., Restle,T., Keegan,R., Grohmann,D. and Schneider,S. (2017) Structural and mechanistic insights into the DNA-guided DNA endonuclease activity of an archaeal Argonaute. Nat. Microbiol., 17035, 1–7.
13. Swarts,D.C., Makarova,K., Wang,Y., Nakanishi,K., Berenguer,J., Brouns,S.J.J.J. and Patel,D.J. (2017) Autonomous generation and loading of DNA guides by bacterial Argonaute. Mol. Cell, 65, 985–998.
14. Olovnikov,I., Kan,K., Sachidanda,R., Newman,D. and Aravin,A. (2013) Bacterial Argonaute samples the transcriptome to identify foreign DNA. Mol. Cell, 51, 594–605.
15. Swarts,D.C., Szczepaniak,M., Sheng,G., Chandradoss,S.D., Zhu,Y., Wang,Y., Swarts,D.C., Szczepaniak,M., Sheng,G., Chandradoss,S.D. et al. (2017) Autonomous generation and loading of DNA guides by bacterial Argonaute. Mol. Cell, 65, 985–998.
41. Wang, Y., Sheng, G., Juranek, S., Tuschl, T. and Patel, D.J. (2008) Structure of the guide-strand-containing argonaute silencing complex. Nature, 456, 209–213.

42. Willkomm, S., Makarova, K. and Grohmann, D. (2018) DNA-silencing by prokaryotic Argonaute proteins adds a new layer of defence against invading nucleic acids. FEMS Microbiol. Rev., 42, 376–387.

43. Kaya, E., Doxzen, K.W., Knoll, K.R., Wilson, R.C., Strutt, S.C., Kranzusch, P.J. and Doudna, J.A. (2016) A bacterial Argonaute with noncanonical guide RNA specificity. Proc. Natl. Acad. Sci. U.S.A., 113, 4057–4062.

44. Sheng, G., Zhao, H., Wang, J., Rao, Y., Tian, W., Swarts, D.C., van der Oost, J., Patel, D.J. and Wang, Y. (2014) Structure-based cleavage mechanism of Thermus thermophilus Argonaute DNA guide strand-mediated DNA target cleavage. Proc. Natl. Acad. Sci. U.S.A., 111, 652–657.

45. Wang, Y., Juranek, S., Li, H., Sheng, G., Wardle, G.S., Tuschl, T. and Patel, D.J. (2009) Nucleation, propagation and cleavage of target RNAs in Ago silencing complexes. Nature, 461, 754–761.

46. Liu, Y., Esyunina, D., Olovnikov, I., Teplova, M., Kulbachinskii, A., Aravin, A.A. and Patel, D.J. (2018) Accommodation of helical imperfections in Rhodobacter sphaeroides Argonaute ternary complexes with guide RNA and target DNA. Cell Rep., 24, 453–462.

47. Swarts, D.C. and Jinek, M. (2018) Cas9 versus Cas12a/Cpf1: Structure–function comparisons and implications for genome editing. Wiley Interdiscip. Rev. RNA, 9, 1–19.

48. Anders, C., Niewoehner, O., Duerst, A. and Jinek, M. (2014) Structural basis of PAM-dependent target DNA recognition by the Cas9 endonuclease. Nature, 513, 569–573.

49. Yamano, T., Zetsche, B., Ishitani, R., Zhang, F., Nishimasu, H. and Nureki, O. (2017) Structural basis for the canonical and non-canonical PAM recognition by CRISPR-Cpf1. Mol. Cell, 67, 633–645.

50. Frank, F., Hauver, J., Nagar, B. (2012) Arabidopsis Argonaute MID domains use their nucleotide specificity loop to sort small RNAs. EMBO J., 31, 3588–3595.